Neural-network quantum states for exotic hadrons and nuclei

Wei-Lin Wu (吴伟霖)

Peking University

Based on: PRL136 (2026) 071901; arXiv: 2606.09254

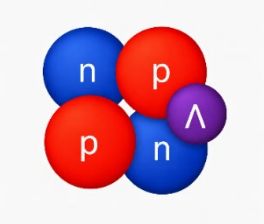
Coauthors: Lu Meng (孟璐), Liang-Zhen Wen (温梁臻), Shi-Lin Zhu (朱世琳)

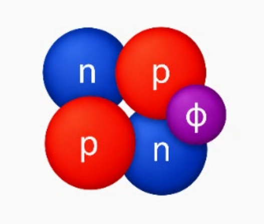
北京大學 PEKING UNIVERSITY

Contents

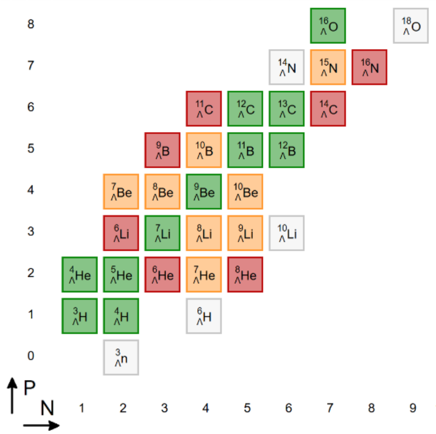
北京大學
PEKING UNIVERSITY

Exotic hadrons and nuclei

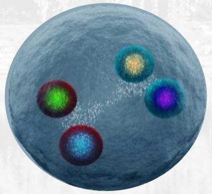




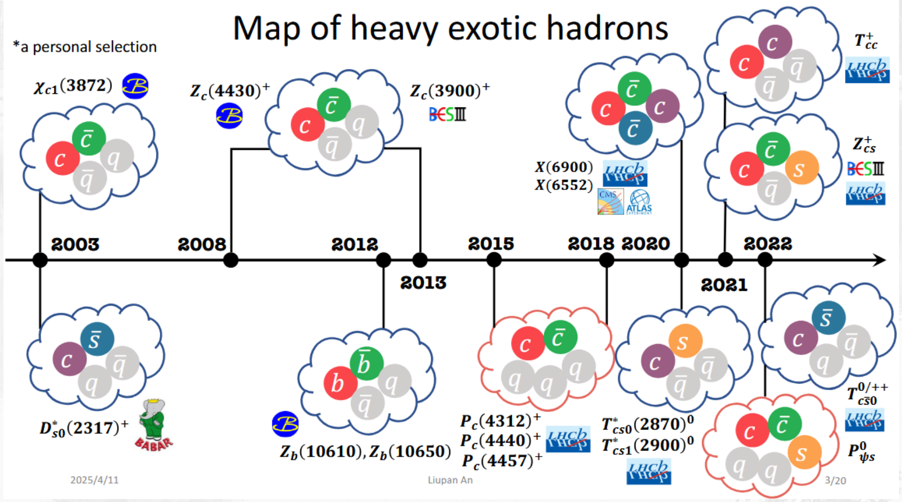
Hypernuclei



hypernuclei.kph.uni-mainz.de

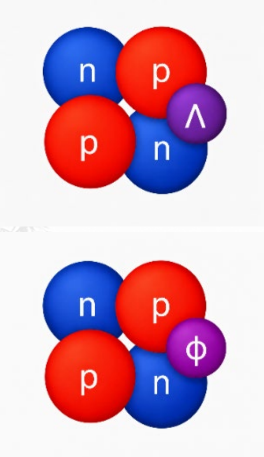


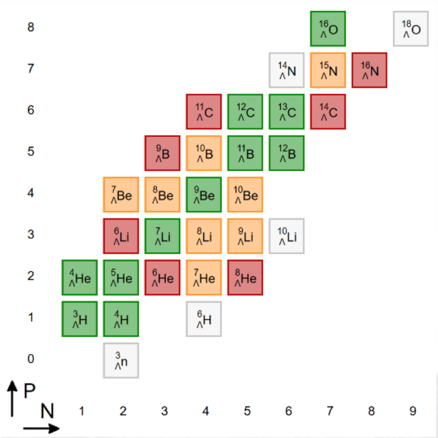
Multiquark



Credit: Liu-Pan An

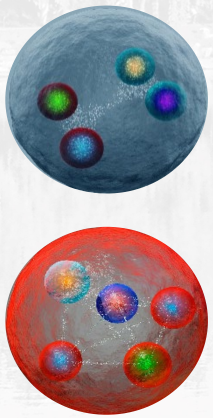
Exotic hadrons and nuclei

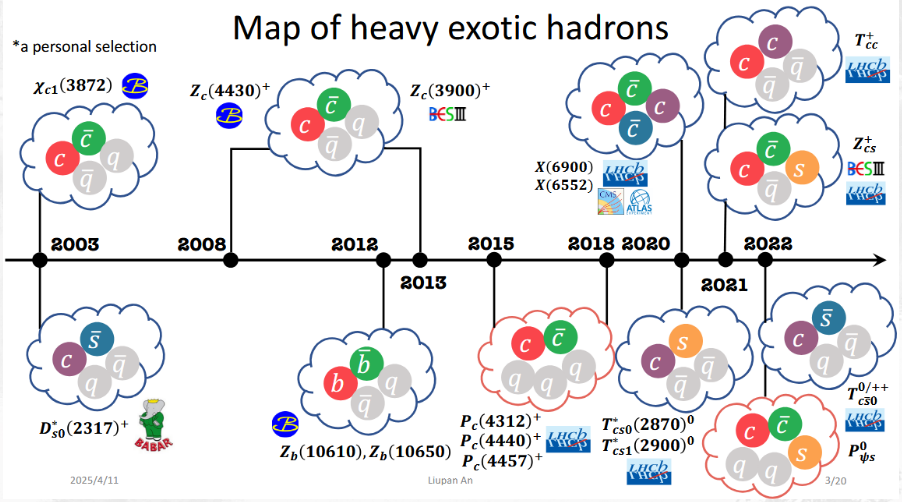




Hypernuclei

hypernuclei.kph.uni-mainz.de





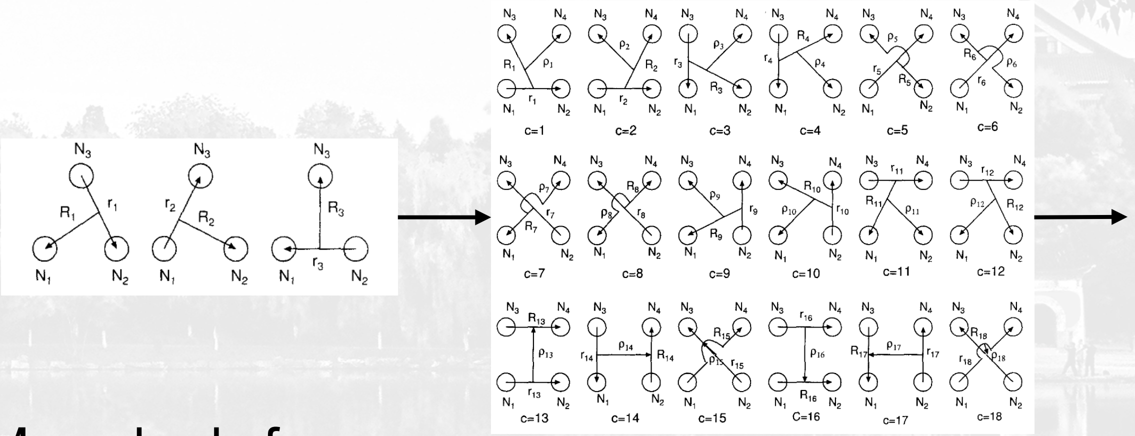
Multiquark

Credit: Liu-Pan An

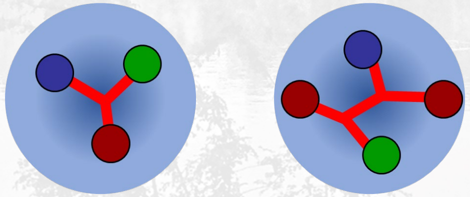
$$H(x_1,\cdots x_N)\psi(x_1,\cdots x_N) = E\psi(x_1,\cdots x_N)$$

Few-body system in strong interaction:

- Dimensionality curse



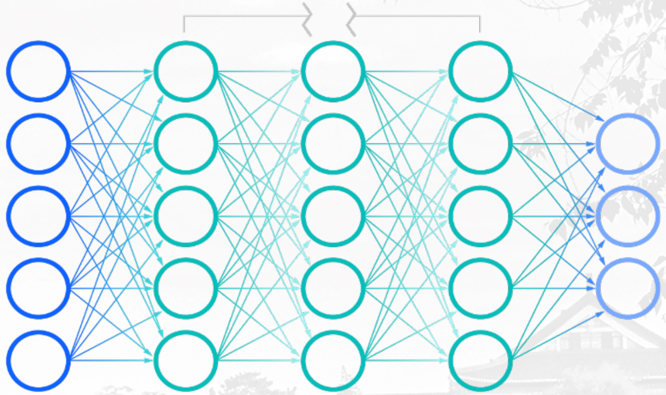
- Many-body force



PRL 86, 18 (2001) ;
PRD 72, 014505 (2005) ;
PRD 96, 074508 (2017)

Neural-network Quantum State (NQS)

DNN: flexible representation of high-dimensional function



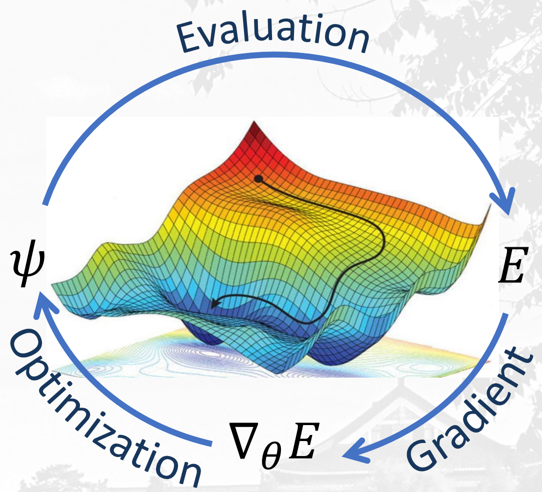
NQS: $\psi_\theta(r_i, s_i, \dots)$ represented by DNN Science 355, 602 (2017)

- Polynomial scaling with N
- Capable of describing different configurations
- Physical symmetries encoded in the architecture

Unsupervised learning: no additional data, trained by physical equation

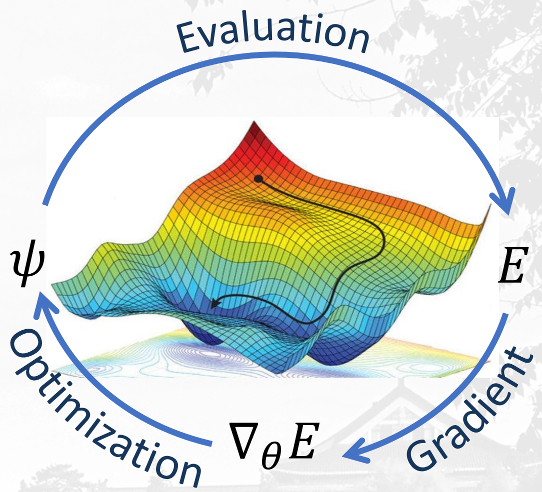
Variational Monte Carlo (VMC)

Variational Principle. $E[\theta] = \langle\psi_\theta|H|\psi_\theta\rangle / \langle\psi_\theta|\psi_\theta\rangle \geq E_0$



- Minimizing $E[\theta]$ approaches ground state
- Traditional ansatz: based on *a priori* knowledge, limited expressiveness
- **NQS: an expressive, flexible ansatz**

Variational Monte Carlo (VMC)



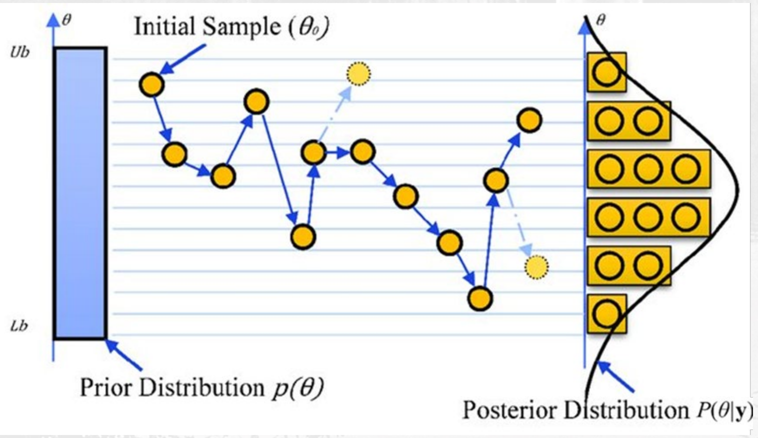
Variational Principle. $E[\theta] = \langle \psi_\theta | H | \psi_\theta \rangle / \langle \psi_\theta | \psi_\theta \rangle \geq E_0$

- Minimizing $E[\theta]$ approaches ground state
- Traditional ansatz: based on *a priori* knowledge, limited expressiveness
- **NQS: an expressive, flexible ansatz**

Monte Carlo Estimation. High-dim integrals via importance sampling:

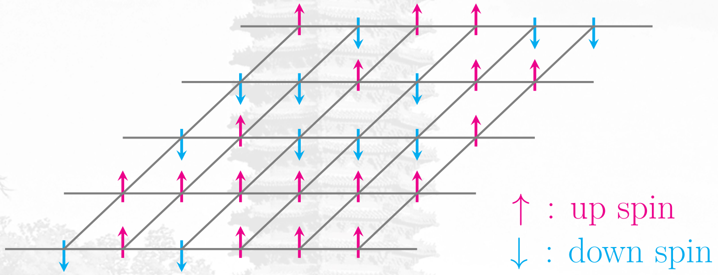
$$E[\theta] = \frac{\int |\psi_\theta(r)|^2 \frac{H\psi(r)}{\psi(r)} dr}{\int |\psi_\theta(r)|^2 dr} \approx \frac{1}{N} \sum_i E_L(r_i) \qquad E_L(r) = \frac{H\psi(r)}{\psi(r)} = \frac{T\psi(r)}{\psi(r)} + V(r)$$

- Samples $r_i \sim |\psi_\theta|^2$ generated by Metropolis-Hastings algorithm
- Efficient evaluation with statistical error $\sim 1/\sqrt{N}$
- easily handles complicated potentials, only needs $V(r_i)$



Applications in many-body physics

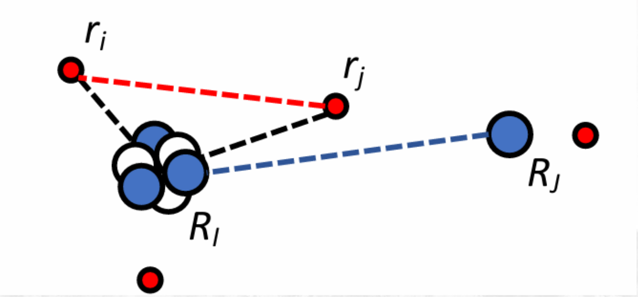
Spin Models



Science 355, 602 (2017), …

see also Timur's talk

Atomic & Molecular Systems

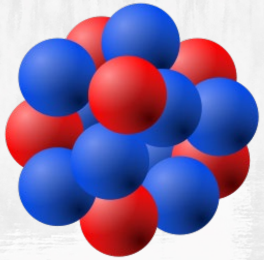


FermiNet: Phys. Rev. Res. 2, 033429 (2020).

PauliNet: Nat. Chem. 12, 891 (2020).

…

Nuclear Physics



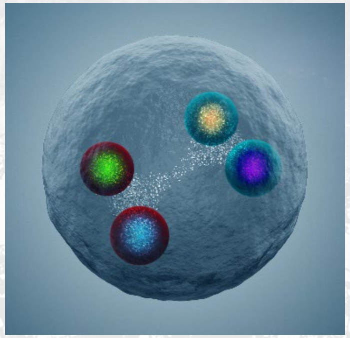
$A \leq 4$: PRL 127, 022502 (2021).

Neutron-α scattering: PRL 135, 172502 (2025).

Hypernuclei: PLB 874, 140285 (2026), …

Mesic nuclei: 2606.09254

Hadron Physics



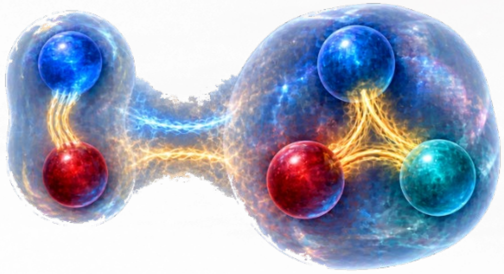
DeepQuark: PRL 136, 071901 (2026)

Contents

北京大學
PEKING UNIVERSITY

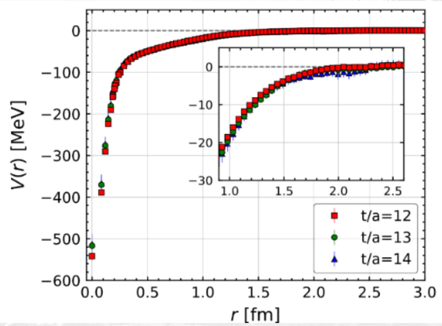
Meson-Nucleus System

- Meson $(s\bar{s}, c\bar{c})$ - Nucleon (qqq)
  - No common valence quarks
  - Probe of gluonic dynamics



- Lattice HAL QCD potential: attractive, but insufficient for MN bound state

PRD 106, 074507 (2022); PLB 860, 139178 (2025)

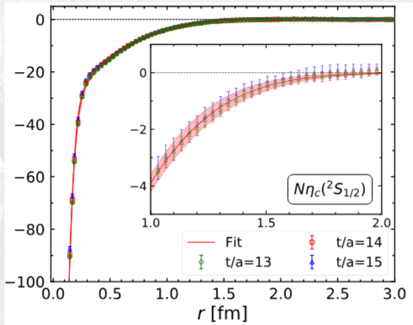


$N\phi(4S_{3/2})$

- M-A bound states
  - Prediction of M-A bound states using QCD van der Waals force

PRL 64, 1011 (1990); PRC 63, 022201 (2001)

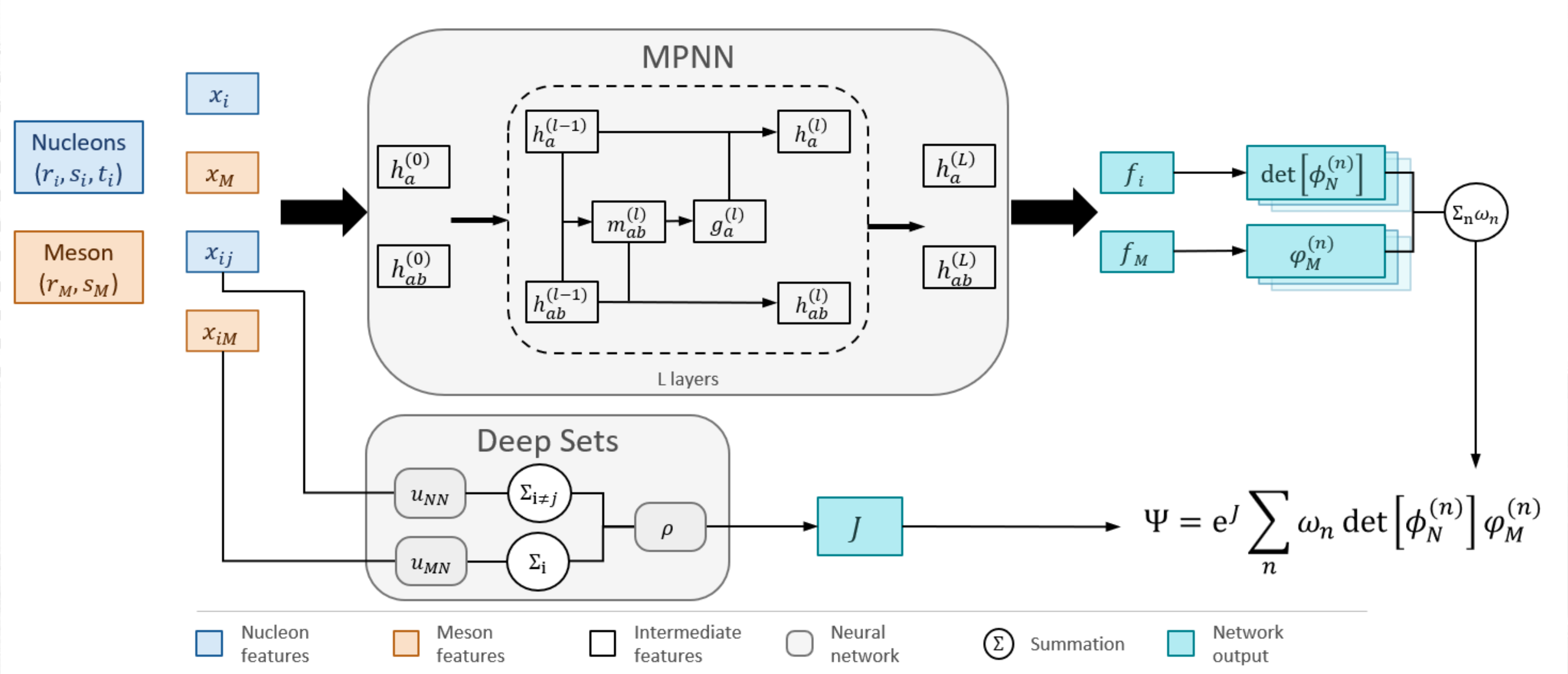
  - Now: state-of-the-art calculations + first-principles potential
  - Solving the full $(A+1)$-body Hamiltonian



$N\eta_c(^2S_{1/2})$

$$H = \sum_i \frac{p_i^2}{2m_i} + V_{NN} + V_{3N} + V_{MN}$$

Meson-Nucleus NQS



Slater determinant

- Enforces fermionic antisymmetry
- Message-passing neural network
  - ➢ $\phi_i(x_j) \rightarrow \phi_i(x_j; x_1, \ldots, x_N)$

Jastrow factor

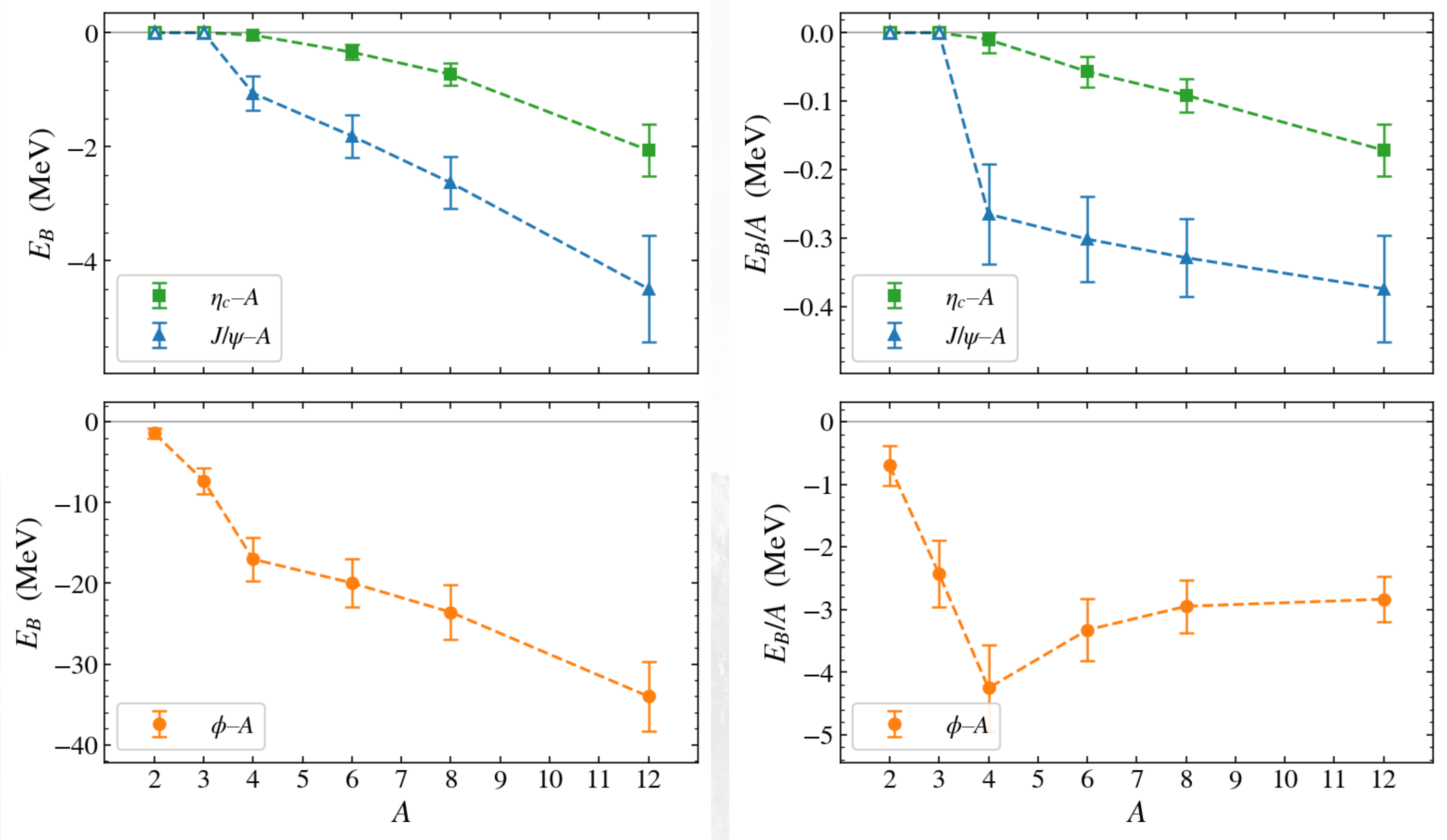
- Permutation invariant
- Deep Sets neural network
  - ➢ $J = \sum_{i<j} u(x_i, x_j) \rightarrow J(x_1, \ldots, x_N)$

Beyond single-particle orbitals and pairwise correlations

Binding energy

- Binding mass number:
  - $A \geq 2$ for ϕ
  - $A \geq 4$ for J/ψ
  - $A \geq 6$ for η_c
- E_B/A - A:
  - Nearly linear for charmonium
  - Peak at $A = 4$ for ϕ

→ Feature of short-range attractive interaction



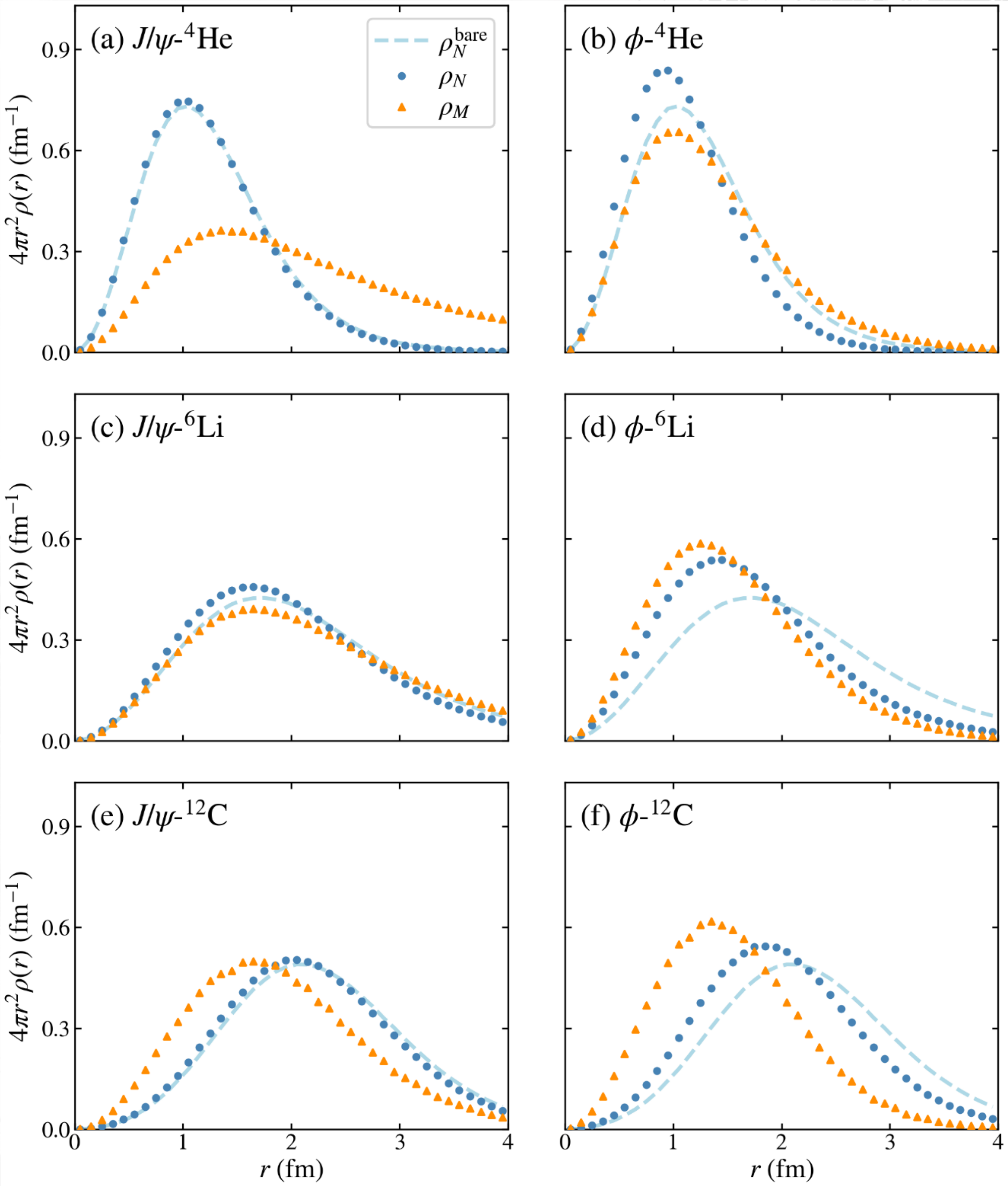
Error bar: uncertainty from HAL QCD potential

Spatial Structure

- Normalized radial density:

$$\rho_N(r) = \frac{1}{4\pi r^2}\frac{\langle\Psi|\sum_i \delta(|\boldsymbol{r}_i - \boldsymbol{R}_{\rm n}| - r)|\Psi\rangle}{A\,\langle\Psi|\Psi\rangle}, \qquad \rho_M(r) = \frac{1}{4\pi r^2}\frac{\langle\Psi|\delta(|\boldsymbol{r}_M - \boldsymbol{R}_{\rm n}| - r)|\Psi\rangle}{\langle\Psi|\Psi\rangle},$$

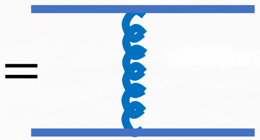
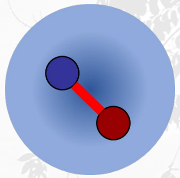
- Charmonium-Nucleus:
  - Meson halo → nuclear interior
- ϕ-Nucleus:
  - Meson embedded inside nucleus
  - "Glue" effect: nuclear distribution is compressed
  - Short-range interaction strength ~ density overlap: maximum at $^4\mathrm{He}$



Contents

北京大學
PEKING UNIVERSITY

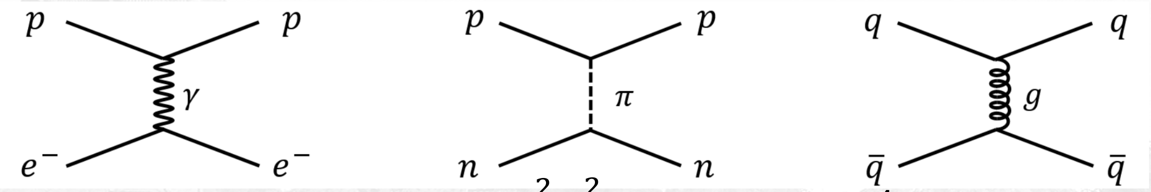
Challenges in multiquark systems

Solving multiquark spectrum in quark potential model: $V_{ij}(r) =$  $+$ 

one-gluon exchange

linear confinement

● Stronger interactions among quarks



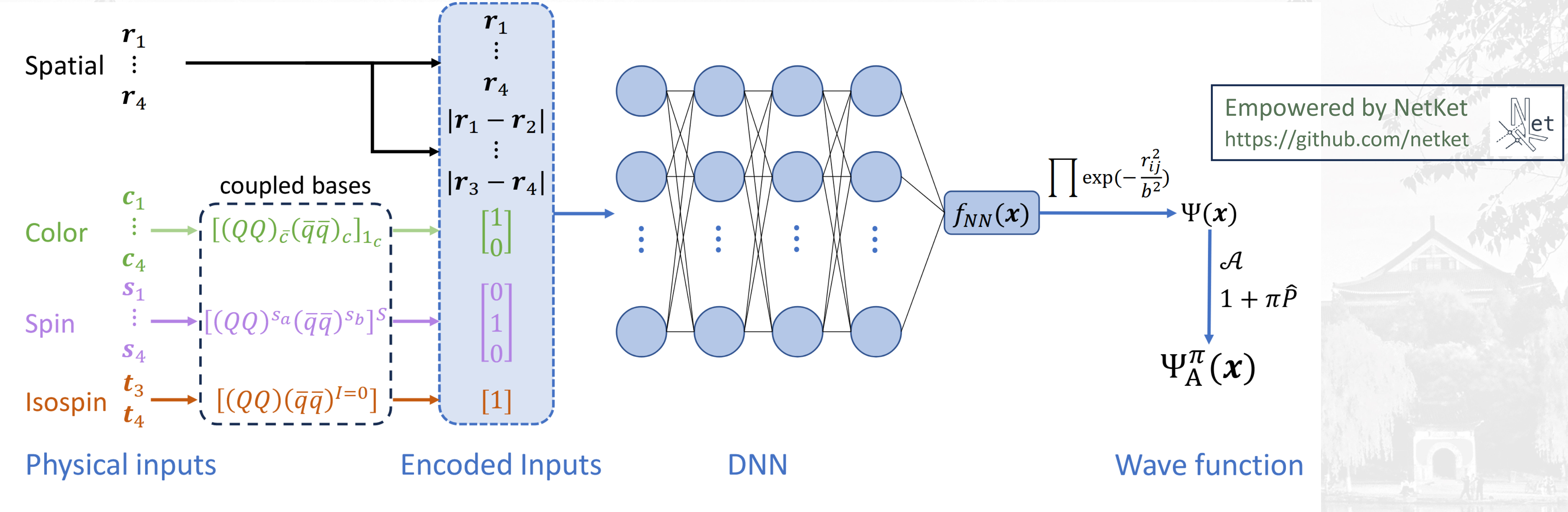
$e^2 = 0.1$ $\quad$ $\frac{m_\pi^2 g_A^2}{16\pi f_\pi^2} \sim 0.1$ $\quad$ $\frac{4}{3}\alpha_s(m_c) \sim 0.4$

- ➢ significant correlations, no evidence of shell structure
- ➢ Slater wave function is not a natural ansatz

● Symmetry constraints

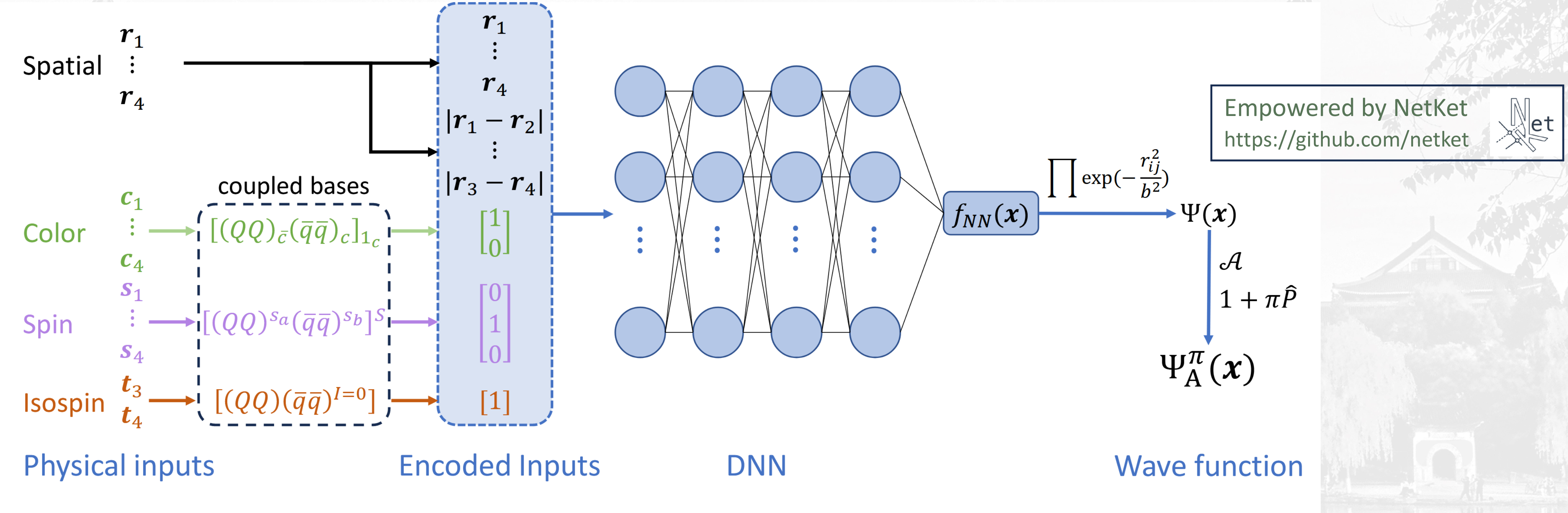
- ➢ color projection to ensure SU(3) singlet
- ➢ spin projection to target states of interest $\quad$ $[cc\bar{q}\bar{q}]_{S=1}$: T_{cc} state; $[cc\bar{q}\bar{q}]_{S=0}$: DD threshold

DeepQuark Architecture



- Coupled bases instead of single-particle (s_i, t_i, c_i)
  - satisfy discrete symmetry
  - describe multichannel coupling

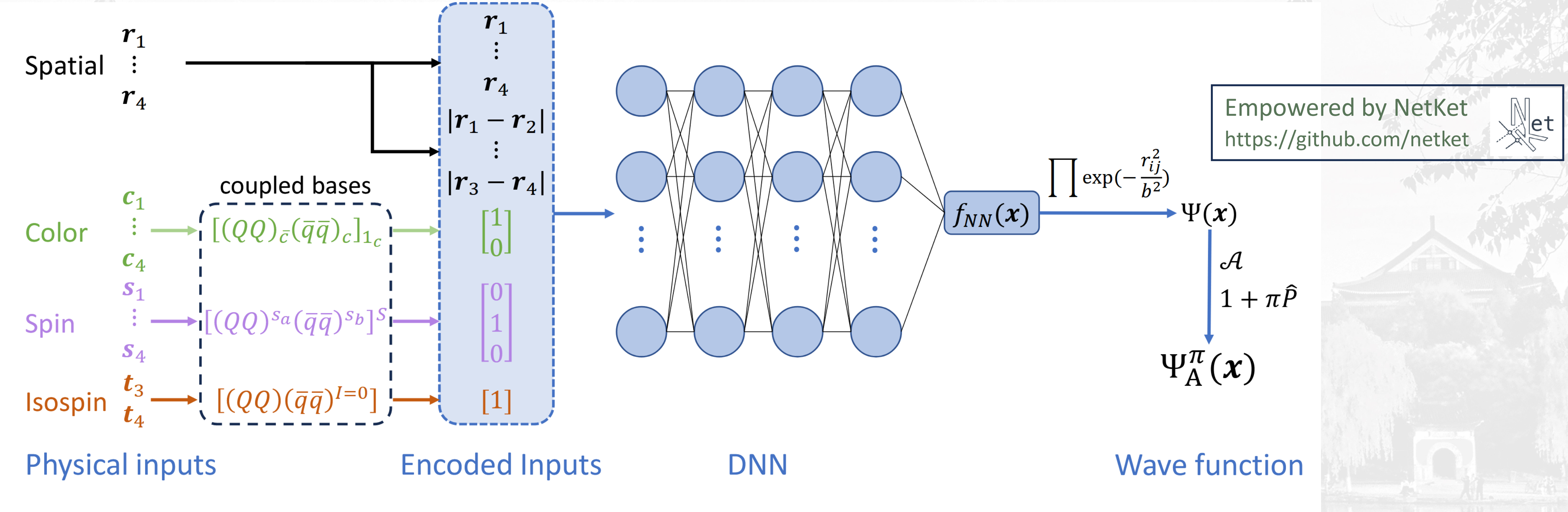
DeepQuark Architecture



- Coupled bases instead of single-particle (s_i, t_i, c_i)
  - ➢ satisfy discrete symmetry
  - ➢ describe multichannel coupling
- Fully connected network with only $O(10^3)$ parameters

vs. GEM for tetraquark:~10^4 bases; $\mathrm{A} \leq 4$ nuclei: ~10^4 paras

DeepQuark Architecture



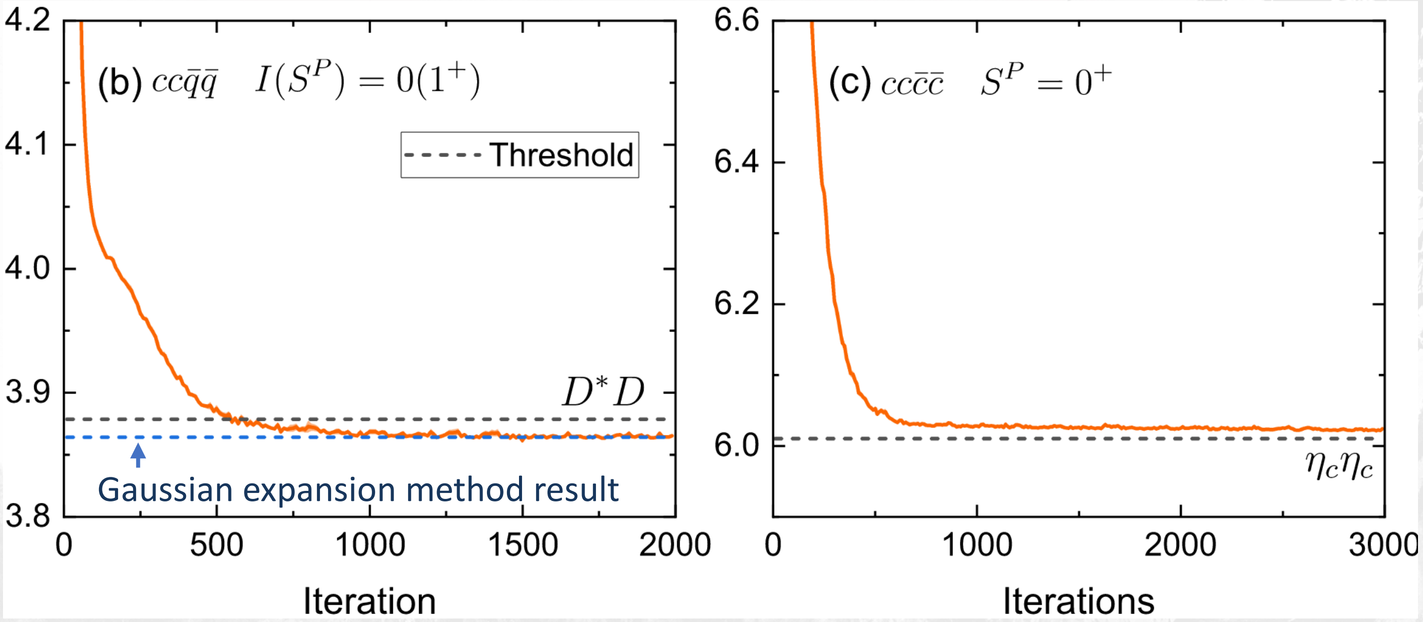
- Coupled bases instead of single-particle (s_i, t_i, c_i)
  - satisfy discrete symmetry
  - describe multichannel coupling
- Fully connected network with only $O(10^3)$ parameters

vs. GEM for tetraquark:~10^4 bases; $A \leq 4$ nuclei: ~10^4 paras

- Antisymmetrization $\mathcal{A} = \sum_{\mathcal{P}}(-1)^{\mathcal{P}}\mathcal{P}$
  - $O(n!)$ for identical particles
  - manageable up to hexaquark
  - the price for enforcing symmetry

Benchmark: Tetraquarks

- Convergence within 2000 iterations
- Agreement with known results



- $cc\bar{q}\bar{q}$: molecular state, agree with exp.
- $bb\bar{q}\bar{q}$: compact tetraquark
- $QQ\bar{Q}\bar{Q}$: no bound states

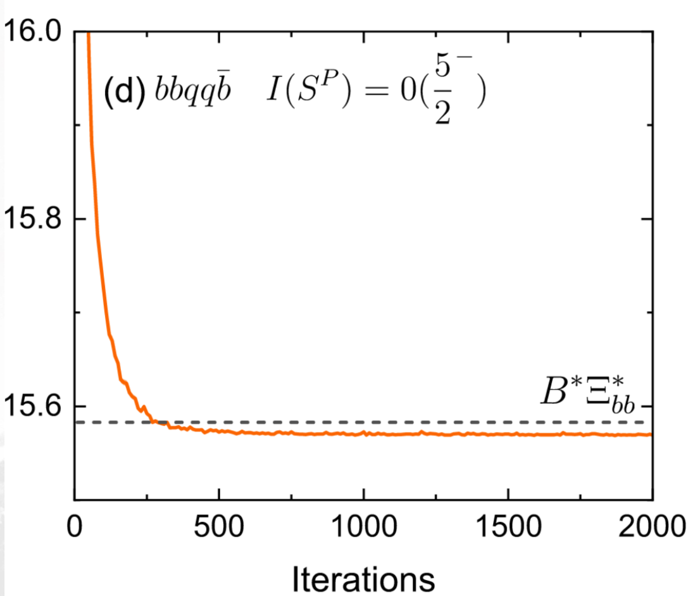
	$I(S^P)$	ΔE	$P_{\bar{3}_c\otimes 3_c}$	$P_{6\otimes 6_c}$	r_{QQ}	$r_{\bar{q}\bar{q}}$	$r_{Q\bar{q}}$
$cc\bar{q}\bar{q}$	$0(1^+)$	-15	55%	45%	1.24	1.41	1.06
$bb\bar{q}\bar{q}$	$0(1^+)$	-153	97%	3%	0.33	0.78	0.69

Extension to Pentaquarks

- Five-body dynamical calculations without approximations via DeepQuark; unattainable from conventional methods yet
- Moderate increase in computational cost compared to tetraquark

  (2000 iters: ~40/60 mins on an A100 GPU for tetra/penta-quark)
- Pentaquark bound states $QQ\bar{Q}qq$

  PRD 88, 054014 (2013)
  PRD 110, 076014 (2024)
  - T_{cc} analogue under heavy diquark-antiquark symmetry: $(QQ)_{\bar{3}_c} \to \bar{Q}$
  - Possible search in D-wave $J/\psi\Lambda_c$



	$I(S^P)$	Thresh.	ΔE
$cc\bar{c}qq$	$0(1/2^-, 3/2^-)$	$\eta_c\Lambda_c, J/\psi\Lambda_c$	Unbound
	$0(5/2^-)$	$\bar{D}^*\Xi^*_{cc}$	-3
$bb\bar{b}qq$	$0(1/2^-, 3/2^-)$	$\eta_b\Lambda_b, \Upsilon\Lambda_b$	Unbound
	$0(5/2^-)$	$B^*\Xi^*_{bb}$	-14

$\to J/\psi\Lambda_c$ (D-wave)

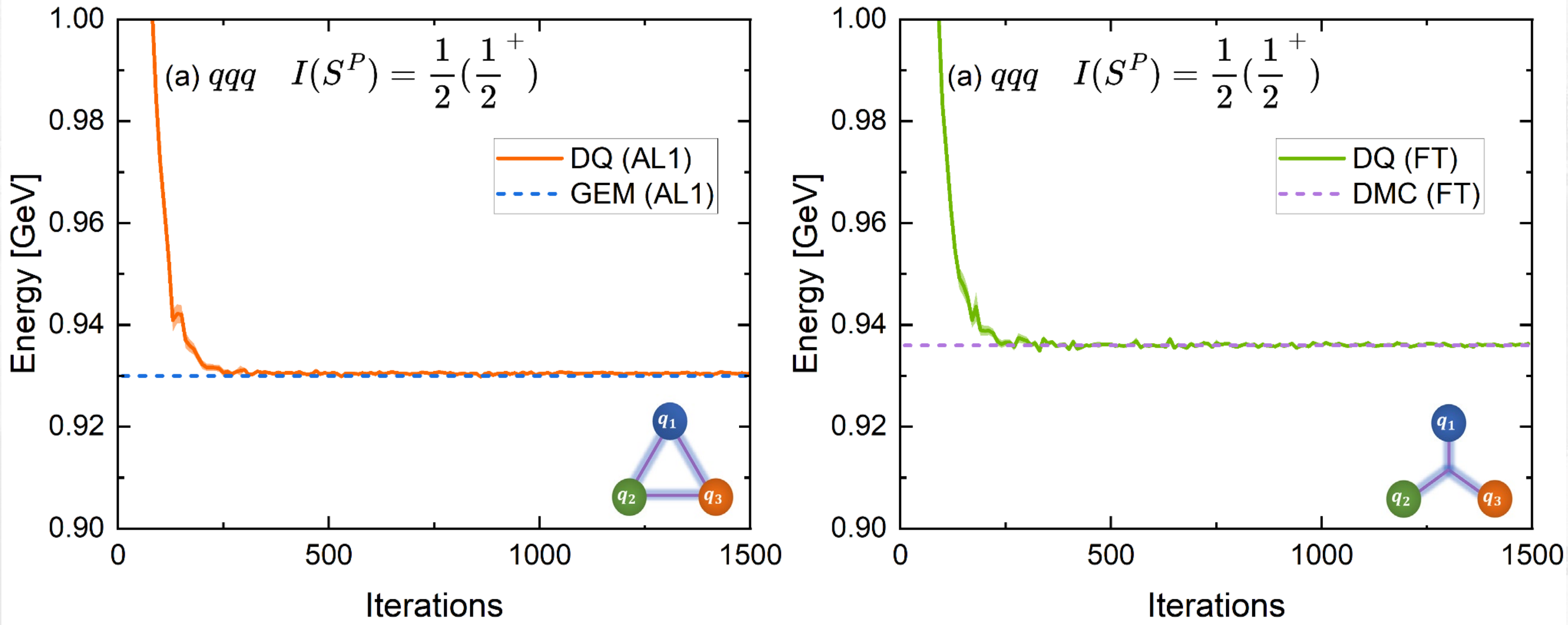
Beyond pairwise interactions

- Lattice QCD favors flux-tube confinement interactions for baryons PRL 86, 18-21 (2001)

$$V_{\text{conf}} = \sum_{i<j} \lambda r_{ij} \to \sigma L_{\min}$$

$$L_{\min} = \begin{cases} \left[\frac{1}{2}(a^2+b^2+c^2) + \frac{\sqrt{3}}{2}\sqrt{d(d-2a)(d-2b)(d-2c)}\right]^{1/2} & \max(\theta_a,\theta_b,\theta_c) < \frac{2\pi}{3}, \\ d - \max(a,b,c) & \max(\theta_a,\theta_b,\theta_c) > \frac{2\pi}{3}, \end{cases}$$

- DeepQuark efficiently incorporates flux-tube type interactions using VMC
- Ready for exploring different confinement mechanisms in tetraquarks



Summary

- DNN + VMC framework
  - DNN as many-body wave function: efficient, general, unbiased
  - VMC: solves ground state using variational principle; handles complicated interactions
- Mesic-Nuclei bound states:
  - Full $(A+1)$-body calculations up to $A = 12$
  - Bound states: $A \geq 2$ for ϕ, $A \geq 4$ for J/ψ, $A \geq 6$ for η_c
  - E_B/A peaks at ϕ-$^4\mathrm{He}$
- DeepQuark:
  - Architecture for strongly correlated multiquark systems with symmetry constraints
  - Outperforms traditional methods starting from pentaquarks
  - Capable of flux-tube interactions

Thanks for your attention!

Backup

Stochastic Reconfiguration

Parameter optimization: Stochastic Reconfiguration (SR)

PRB 71, 241103 (2005)

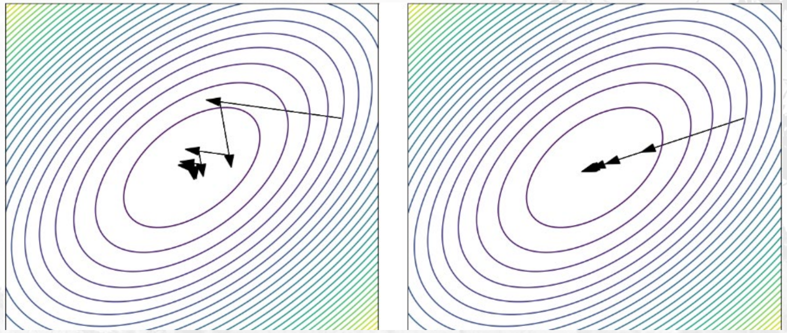
$$\theta^{i+1} = \theta^{i} - \eta S^{-1} \nabla_{\theta^i} E[\theta^i]$$

- S: Quantum Fisher information matrix

$$S_{ij} = \frac{\left\langle \partial_{\theta_i}\psi \middle| \partial_{\theta_j}\psi \right\rangle}{\langle\psi|\psi\rangle} - \frac{\langle \partial_{\theta_i}\psi|\psi\rangle \left\langle \psi \middle| \partial_{\theta_j}\psi \right\rangle}{\langle\psi|\psi\rangle}$$

  - ➢ Metric on wave-function manifold
  - ➢ Faster, more stable convergence

$\theta^{i+1} = \theta^{i} - \eta \nabla E$ $\qquad$ $\theta^{i+1} = \theta^{i} - \eta S^{-1} \nabla E$

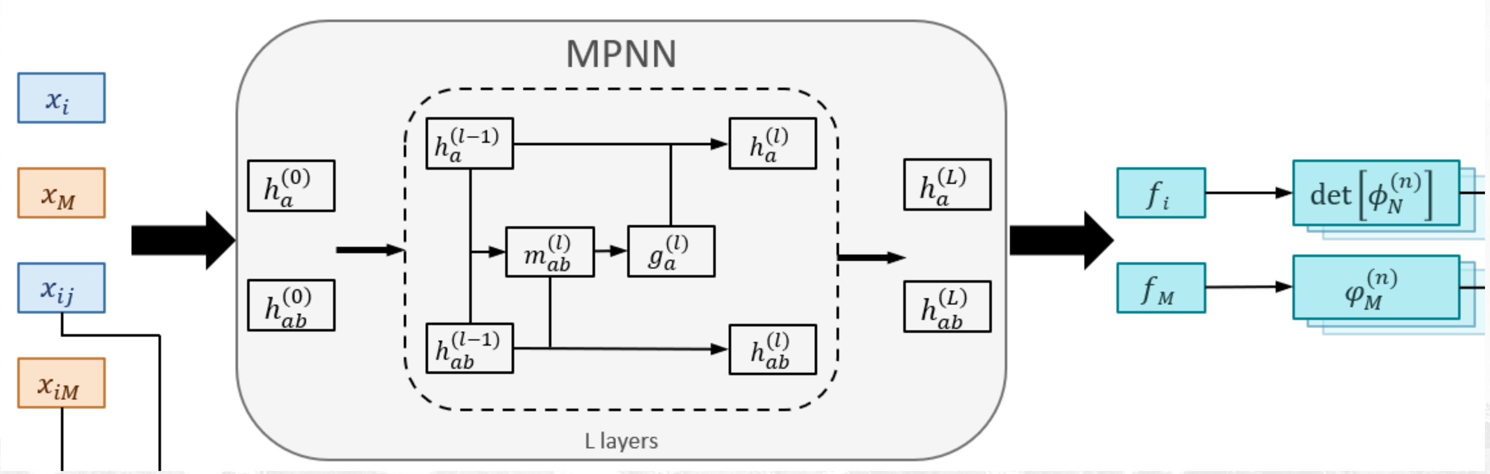


Nuclear ground states

TABLE I. Nuclear ground-state energies and S^2 expectation values. E uses the lattice nucleon mass $m_N = 954$ MeV, while E_{phys} uses the physical $m_N = 938.9$ MeV. Statistical errors are shown in parentheses. Experimental values from [61] are listed for comparison.

Nucleus	E (MeV)	E_{phys} (MeV)	$\langle S^2 \rangle$	Exp. (MeV)
^{2}H	−2.416(0)	−2.240(1)	2.00	−2.224
^{3}H	−8.921(1)	−8.470(3)	0.75	−8.482
^{3}He	−8.250(1)	−7.808(2)	0.75	−7.718
^{4}He	−29.072(1)	−28.161(3)	0.00	−28.30
^{6}Li	−32.478(4)	−31.134(12)	2.01	−31.99
^{8}Be	−57.873(7)	−55.689(26)	0.02	—
^{12}C	−91.621(16)	−87.724(28)	0.03	−92.16

Message-passing neural network



Backflow-transformed orbitals $\phi_i(x_j; x_1, \ldots, x_N)$:

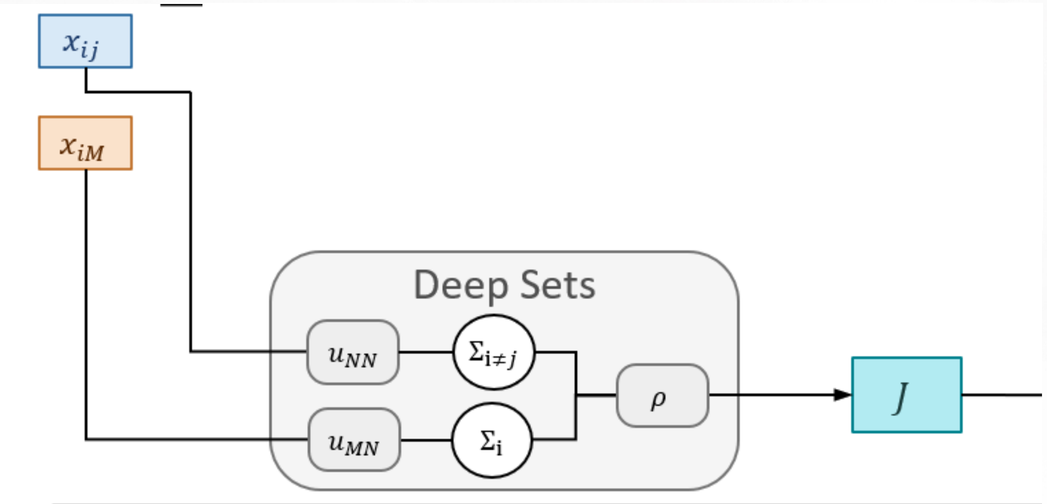
- Antisymmetry: nucleon exchange $i \leftrightarrow j$ swaps orbital rows $\phi_{i\mu} \leftrightarrow \phi_{j\mu}$
- Input $h_i^{(0)}: (r_i, s_i, t_i)$, $h_{ij}^{(0)}: (r_{ij}, s_i, s_j, t_i, t_j)$
- Message-passing iteration
- Output

$$\boldsymbol{m}_{ab}^{(l)} = M^{(l)}(\boldsymbol{h}_a^{(l-1)}, \boldsymbol{h}_b^{(l-1)}, \boldsymbol{h}_{ab}^{(l-1)}),$$
$$\boldsymbol{m}_a^{(l)} = \sum_{b \neq a} \boldsymbol{h}_b^{(l-1)} \odot \boldsymbol{m}_{ab}^{(l)},$$
$$\boldsymbol{h}_a^{(l)} = \boldsymbol{h}_a^{(l-1)} + U^{(l)}(\boldsymbol{h}_a^{(l-1)}, \boldsymbol{m}_a^{(l)}),$$
$$\boldsymbol{h}_{ab}^{(l)} = \boldsymbol{h}_{ab}^{(l-1)} + V^{(l)}(\boldsymbol{h}_{ab}^{(l-1)}, \boldsymbol{m}_{ab}^{(l)}).$$

$$\boldsymbol{f}_a = \boldsymbol{h}_a^{(L)} + \sum_{b=1}^{A+1} \lambda(\boldsymbol{h}_b^{(L)}),$$

$$\phi_{i\mu}^{(n)} = O_N^{(n)}(\boldsymbol{f}_i)_\mu \exp(-\xi^{(n)^2} \bar{r}_i^2),$$
$$\varphi_M^{(n)} = O_M^{(n)}(\boldsymbol{f}_M) \exp\Big(-\sum_{\alpha=1}^{3} \eta_\alpha^{(n)^2} \bar{r}_{M,\alpha}^2\Big),$$

Deep Sets



Deep Sets： a permutation-invariant form by summing over all index

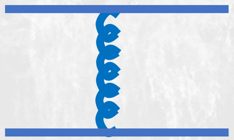
$$J = \rho\left(\sum_{i\neq j} u_{NN}(\boldsymbol{x}_{ij}),\ \sum_{i} u_{MN}(\boldsymbol{x}_{iM})\right)$$

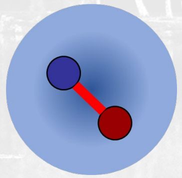
- u_{NN}, u_{MN}: maps x_{ij}, x_{iM} into latent space R^F
- ρ: $R^{2\mathrm{F}} \rightarrow R$

AL1 model

$$H = \sum_i \frac{p_i^2}{2m_i} + \sum_{i<j} V_{\mathrm{OGE},ij} + V_{\mathrm{conf},ij}$$

- Minimal nonrelativistic model extended to multiquark sectors
- Pairwise interaction:
  - one-gluon-exchange $$V_{\mathrm{OGE},ij} = \frac{\lambda_i \cdot \lambda_j}{4}\left[\frac{\alpha_s}{r} - \frac{8\pi\alpha_s'}{3m_i m_j}\tilde{\delta}(\sigma;r)\boldsymbol{S_i}\cdot\boldsymbol{S_j}\right]$$
  - linear confinement $$V_{\mathrm{conf},ij} = \frac{\lambda_i \cdot \lambda_j}{4}\left(-\frac{3}{4}br + \Lambda\right)$$



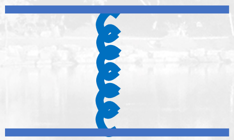


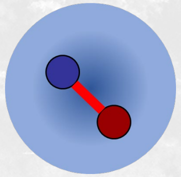
- Parameters fitted by meson spectra across all flavor sectors

More models

TABLE SM-IV. Different ingredients in five different quark models.

	One-gluon-exchange		Confinement		meson-exchange
	asymptotic freedom	Hyperfine	form	screening effect	
AL1	No	Gaussian	linear	No	No
AL2	Yes	Gaussian	linear	No	No
AP1	No	Gaussian	$p = 2/3$	No	No
AP2	Yes	Gaussian	$p = 2/3$	No	No
SLM	No	Yukawa	linear	Yes	Yes



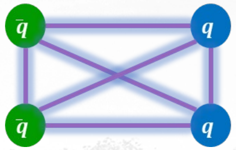


Potential	DQ		GEM [4]	
	E	ΔE	E	ΔE
AL1	10491	-153	10493	-152
AP1	10502	-176	10504	-175
AL2	10473	-151	10474	-151
AP2	10471	-174	10472	-174
SLM	10230	-362	10232	-360

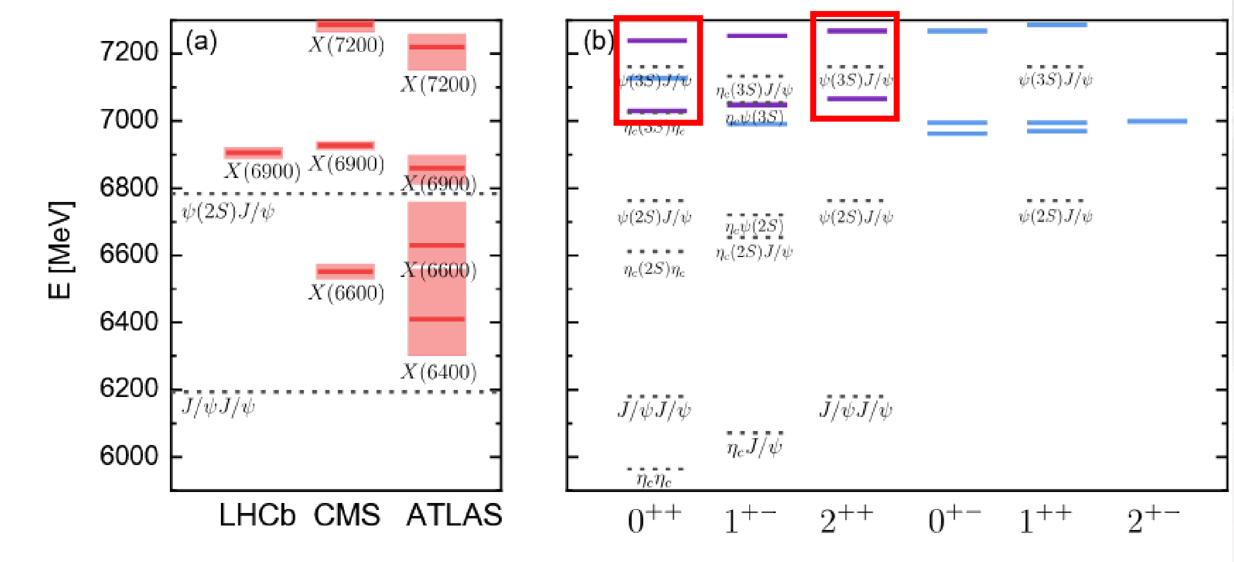
- Accommodate a wide variety of interactions
- Taking $bb\bar{q}\bar{q}$ as a testing system, all in agreement with GEM
- A flexible framework for spectroscopy studies

QM for tetraquark states

- Does QM work in multiquark sector?
- Previous tetraquark study:
  - Pairwise confinement interaction model



  - Gaussian expansion method (GEM) to solve 4-body problem PPNP 51, 223 (2003)
  - Bound and resonant states
    - ✓ $T_{cc}(3875)^+$ PLB 814, 136095 (2021)
    - ✓ $T_{cc\bar{c}\bar{c}}(6900), T_{cc\bar{c}\bar{c}}(7200)$ PRD 109, 054034 (2024)
    - ✓ $T^*_{cs0}(2870)^0$ PRD 109, 014010 (2024)



Experiment Theory