DeepQuark: deep-neural-network approach to multiquark bound states

Wei-Lin Wu [1,*] Lu Meng [2,3,†] and Shi-Lin Zhu [4,‡]

[1]School of Physics, Peking University, Beijing 100871, China
[2]Institut für Theoretische Physik II, Ruhr-Universität Bochum, D-44780 Bochum, Germany
[3]School of Physics, Southeast University, Nanjing 210094, China
[4]School of Physics and Center of High Energy Physics, Peking University, Beijing 100871, China

For the first time, we implement the deep-neural-network-based variational Monte Carlo approach for the multiquark bound states, whose complexity surpasses that of electron or nucleon systems due to strong SU(3) color interactions. We design a novel and high-efficiency architecture, DeepQuark, to address the unique challenges in multiquark systems such as stronger correlations, extra discrete quantum numbers, and intractable confinement interaction. Our method demonstrates competitive performance with state-of-the-art approaches, including diffusion Monte Carlo and Gaussian expansion method, in the nucleon, doubly heavy tetraquark, and fully heavy tetraquark systems. Notably, it outperforms existing calculations for pentaquarks, exemplified by the triply heavy pentaquark. For the nucleon, we successfully incorporate three-body flux-tube confinement interactions without additional computational costs. In tetraquark systems, we consistently describe hadronic molecule T_{cc} and compact tetraquark T_{bb} with an unbiased form of wave function ansatz. In the pentaquark sector, we obtain weakly bound $\bar{D}^*\Xi_{cc}^*$ molecule $P_{cc\bar{c}}(5715)$ with $S=\frac{5}{2}$ and its bottom partner $P_{bb\bar{b}}(15569)$. They can be viewed as the analogs of the molecular T_{cc}. We recommend experimental search of $P_{cc\bar{c}}(5715)$ in the D-wave $J/\psi\Lambda_c$ channel. DeepQuark holds great promise for extension to larger multiquark systems, overcoming the computational barriers in conventional methods. It also serves as a powerful framework for exploring confining mechanism beyond two-body interactions in multiquark states, which may offer valuable insights into nonperturbative QCD and general many-body physics.

I. INTRODUCTION

The quark model proposed by Gell-Mann [1] and Zweig [2] provides a remarkably successful framework for understanding the conventional hadron spectrum, classifying hadrons into mesons ($q\bar{q}$) and baryons (qqq). Although multiquark states such as tetraquarks ($qq\bar{q}\bar{q}$) and pentaquarks ($qqqq\bar{q}$) were proposed concurrently [1, 2], their existence remained elusive until recent years. Since the discovery of the $X(3872)$ in 2003 [3], numerous multiquark candidates have been observed, including many manifestly exotic states like $T_{c\bar{c}1}(3900)$ [4, 5], $T_{cc}(3875)^+$ [6, 7] , $P_{c\bar{c}}$ states [8, 9], and $T_{cc\bar{c}\bar{c}}(6900)^0$ [10], establishing the existence of multiquark hadrons (see [11–18] for comprehensive reviews). Elucidating the configurations of these multiquark states poses new challenges to the quark model, and also offers a unique window into the nonperturbative regime of quantum chromodynamics (QCD)—the fundamental theory of strong interactions. The investigation of clustering behaviors in multiquark systems, particularly in distinguishing compact configurations from loosely bound hadronic molecules, offers an unprecedented opportunity to explore quantum many-body dynamics governed by SU(3) color interactions, which is fundamentally different from systems dominated by electromagnetic or nuclear forces.

Solving the quantum many-body problem of multiquark states in quark models can be quite challenging. In addition to the exponential scaling of wave function dimensionality with the particle number, the multiquark systems involve an extra SU(3) color degree of freedom compared to electron and nucleon systems, causing even greater complexity. Moreover, constructing multiquark systems of interest often requires imposing constraints from various quantum numbers—such as spin, parity, flavor, and color—that arise from underlying symmetry principles. These constraints significantly complicate the theoretical treatment. For example, in calculating the ground states of electron systems, spins are usually not explicitly constrained, as the ground state naturally converges to a certain spin configuration. However, in the case of doubly charmed tetraquark T_{cc} [6, 7], the total spin should be constrained to 1 to avoid falling into the lower scalar DD threshold. Furthermore, the strong color interaction between (anti)quarks leads to significant correlations, rendering the single-particle approximation—effective in nuclear and atomic systems—invalid. As a result, the shell structure observed in those systems is absent in multiquark states. Without such a simplified approximation as a guidance or starting point, the full dynamical multichannel treatments of multiquark systems become computationally intensive. Additionally, progress in lattice QCD [19, 20] seems to favor the few-body confinement mechanisms over pairwise interactions in multiquarks, introducing new challenges to their quantum many-body descriptions [21]. The aforementioned complexities make comprehensive calculations of multiquark states highly challenging, even for five-body systems. Existing approaches, such as basis expansion methods exemplified by the Gaussian expansion method (GEM) [22], are hindered from complete calculations due to the exponential growth of basis states. Meanwhile, diffusion Monte Carlo (DMC) suffers from the notorious sign problem [23], which severely limits its applicability to such strongly correlated systems. Previous studies on pentaquarks [24–30] made various approximations in the spatial configurations [25–29] or color degree of

* wlwu@pku.edu.cn
† lu.meng@rub.de
‡ zhusl@pku.edu.cn

freedom [24, 25, 30] to simplify the calculations, which may result in unknown systematical errors and unreliable conclusions.

Recently, the development of machine learning techniques provides a new approach to solving the quantum many-body problem. Motivated by the exceptional capacity of deep neural networks (DNNs) to approximate high-dimensional functions [31], DNNs are anticipated to be a flexible and effective wave function representation capable of capturing complex many-body correlations efficiently [32, 33]. DNN-based variational Monte Carlo (VMC) method has been successfully applied to quantum spin systems [34], atomic and molecular physics [35–37], condensed matter [38, 39] and nuclear physics [40–46], achieving high accuracies and demonstrating potential for scaling to larger systems. However, its application in multiquark remains unexplored. DNN-based wave functions, which are free from *a priori* assumptions, can unbiasedly and consistently describe various multiquark configurations, including hadronic molecules and compact multiquark states. Moreover, VMC circumvents the sign problem in imaginary time evolution of DMC, making it capable of handling strongly correlated systems. Meanwhile, unlike basis expansion methods, VMC enables the treatments of complicated interactions that go beyond two-body force even without extra computational costs, paving the way for investigations of confining mechanisms and many-body forces in multiquark states.

In this work, we develop a DNN-based VMC approach, DeepQuark, to calculate multiquark bound states in the quark model. A major difference between DeepQuark and previous DNN-based studies is the architecture of many-body wave function. Most of the previous works use a determinant type wave function ansatz [36–38, 43], which originates from the idea of single-particle orbitals in electron and nucleon systems, and consider the correlations by multiple determinants, Jastrow factors and backflow transformation. However, such constructions may not apply to the strongly correlated hadron systems. Instead, we construct the DeepQuark wave function in the coupled color-spin-isospin bases, which represents the correlations in the most general way while automatically enforcing symmetry requirements. We first benchmark our results against GEM and DMC for the nucleon in two different confinement interactions and tetraquark bound states. We further demonstrate DeepQuark's ability for larger systems by investigating triply heavy pentaquarks $QQqq\bar{Q}$ ($Q = b, c; q = u, d$), where bound states that are analogs of the doubly heavy tetraquarks $QQ\bar{q}\bar{q}$ are obtained.

II. HAMILTONIAN

The nonrelativistic Hamiltonian in the quark model reads

$$H = \sum_i (m_i + \frac{p_i^2}{2m_i}) + \sum_{i<j} V_{ij}, \tag{1}$$

where the first two terms are the mass and kinetic energy of the (anti)quark i, respectively. V_{ij} represents the two-body interaction between the (anti)quark pair (ij). In this work, we adopt a minimal quark model—the AL1 model introduced in Refs. [47, 48]—which includes the one-gluon-exchange interaction V_{OGE} and a two-body linear confinement term V_{conf},

$$\begin{aligned} V_{\text{OGE},ij} &= -\frac{3}{16}\boldsymbol{\lambda}_i \cdot \boldsymbol{\lambda}_j \left(-\frac{\kappa}{r_{ij}} - \Lambda + \frac{8\pi\kappa'}{3m_i m_j} \frac{e^{-r_{ij}^2/r_0^2}}{\pi^{3/2} r_0^3} \boldsymbol{s}_i \cdot \boldsymbol{s}_j \right), \\ V_{\text{conf},ij} &= -\frac{3}{16}\boldsymbol{\lambda}_i \cdot \boldsymbol{\lambda}_j \lambda r_{ij}. \end{aligned} \tag{2}$$

Here, $\boldsymbol{\lambda}_i$ is the SU(3) Gell-Mann color matrix acting on the i-th quark (replaced by -$\boldsymbol{\lambda}^*$ for antiquark), and $\boldsymbol{s}_i$ is the spin-operator. The parameters of the model were determined by fitting the meson spectra across all flavor sectors. This interaction structure closely resembles that of the well-known Cornell model [49, 50] but takes the light quark sector into consideration. For baryon systems [48], an additional phenomenological term $V_{123} = -\frac{C}{m_1 m_2 m_3}$, which is inversely proportional to the quark mass, is introduced to mimic the three-body interaction effect. For hadron systems with heavy quarks, this term only gives small corrections and therefore is neglected in calculation.

For the nucleon, we also test the flux-tube confinement interaction $V_{\text{conf}}^{\text{ft}}$, which is proportional to L_{min}, the minimal length of the color flux tubes connecting three quarks at a junction point [51–53],

$$V_{\text{conf}}^{\text{ft}} = \sigma L_{\text{min}}. \tag{3}$$

The parameters of the potential and the masses of heavy mesons and baryons are given in Appendix A. In the absence of spin-orbit and tensor operators, the total angular momentum J, total orbital angular momentum L and total spin S are all good quantum numbers. Since ground states are expected to be S-wave, we have $J = S$.

III. NEURAL-NETWORK WAVE FUNCTION

The architecture of the DeepQuark wave function is illustrated in Fig. 1. The core of the framework is to construct a multiquark wave function with the desired symmetry, namely, a color-singlet state with definite total spin S, isospin I and parity π that obeys Fermi-Dirac statistics. To that end, we work in the coupled bases. Taking isoscalar vector doubly heavy tetraquark as an example, the color-spin-isospin degrees of freedom can be coupled into the following independent bases,

$$\begin{aligned} \chi_{\bar{3}_c \otimes 3_c} \phi^{s_a, s_b} \xi^{I=0} &= \left[(QQ)_{\bar{3}_c}^{s_a} (\bar{q}\bar{q})_{3_c}^{s_b, I=0} \right]_{1_c}^{S=1}, \\ \chi_{6_c \otimes \bar{6}_c} \phi^{s_a, s_b} \xi^{I=0} &= \left[(QQ)_{6_c}^{s_a} (\bar{q}\bar{q})_{\bar{6}_c}^{s_b, I=0} \right]_{1_c}^{S=1}, \end{aligned} \tag{4}$$

where s_a and s_b take all possible combinations. χ, ϕ, and ξ represent the color, spin, and isospin coupled bases, which are further mapped into vectors α_c, α_s and α_t, respectively. Here, α is a standard basis vector in $\mathbb{R}^n$, with n being the number of independent bases. The mapping into standard basis vectors rather than integers from 1 to n ensures that no prejudiced

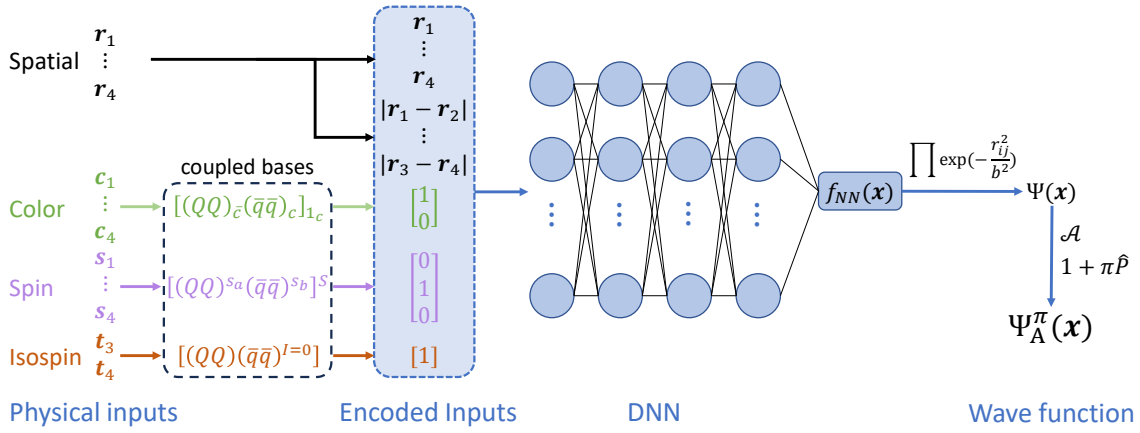

FIG. 1. Architecture of the DeepQuark wave function, taking isoscalar doubly heavy tetraquark as an example. The physical inputs in spatial, color, spin and isospin degrees of freedom are transformed as the encoded inputs before fed into the DNN. Boundary condition, fermionic antisymmetrization ($\mathcal{A}$) and parity projection ($1 + \pi\hat{P}$) are imposed on the scalar output of the DNN $f_{NN}(\boldsymbol{x})$ to obtain the multiquark wave function $\Psi_{\rm A}^{\pi}(\boldsymbol{x})$.

correlation between different bases is introduced. By taking the vectors α as inputs of the DNN, the symmetry information is encoded into the DeepQuark wave function.

For the spatial degree of freedom, we include the (anti)quark coordinates in the center of mass frame $\boldsymbol{r}_i$ and the distances between two (anti)quarks $|\boldsymbol{r}_i - \boldsymbol{r}_j|$ as input features. The inclusion of the distances is redundant in principle but can improve the performance of neural-network wave function [36]. The magnitude of interparticle distance is a non-smooth function at zero, which could better describe the wave function cusps [36] arising from short-range color Coulomb interaction in Eq. (2). It is also an important variable with direct physical meaning, which can encapsulate interparticle correlations effectively. It is worth mentioning that as the wave function is a function of three-dimensional coordinates $\boldsymbol{r}_i$, it includes contributions from all orbital angular momentum L. The wave function should naturally converge to the S-wave ground state in the optimization process.

The input features $\boldsymbol{x} = (\boldsymbol{r}_i, |\boldsymbol{r}_i - \boldsymbol{r}_j|, \alpha_c, \alpha_s, \alpha_t)$ are fed into a DNN with four fully connected hidden layers. Each layer takes the outputs of the previous layer as inputs and carries out the following mapping,

$$\boldsymbol{x}_{\rm out} = \sigma(\boldsymbol{W}\boldsymbol{x}_{\rm in} + \boldsymbol{c}), \tag{5}$$

where $\boldsymbol{W}$, $\boldsymbol{c}$ are variational parameters, and $\sigma = \tanh$ is the activation function. Since the parameters are initialized randomly, the scalar output of the network f_{NN} are multiplied by a boundary condition $\prod_{i<j} \exp(-r_{ij}^2/b^2)$ to confine the system in a localized space. We take $b = 2 \sim 4$ fm, which is on the order of the typical range of color confinement, $\Lambda_{\rm QCD}^{-1} \sim 1$ fm. Finally, the fermionic antisymmetry and parity projection is enforced,

$$\Psi_{\rm A}^{\pi}(\boldsymbol{x}) = (1 + \pi\hat{P})\mathcal{A}\left[f_{NN}(\boldsymbol{x}) \prod_{i<j} \exp\left(-\frac{r_{ij}^2}{b^2}\right)\right], \tag{6}$$

where $\hat{P}$ and $\mathcal{A}$ are the spatial inversion operator and antisymmetric operator of identical particles, respectively. Enforcing antisymmetry by explicitly summing over all possible permutations leads to a factorial complexity. However, given the current experimental and theoretical progress, we will only focus on multiquark systems with at most 3 to 4 identical particles, and such a complexity is manageable.

The DNN is trained in an unsupervised way using the variational principle. The parameters are optimized by minimizing the energy expectation value:

$$E_{\boldsymbol{\theta}} = \frac{\langle\psi_{\boldsymbol{\theta}}|H|\psi_{\boldsymbol{\theta}}\rangle}{\langle\psi_{\boldsymbol{\theta}}|\psi_{\boldsymbol{\theta}}\rangle} \geq E_0, \tag{7}$$

where $\boldsymbol{\theta}$ denotes the parameters and E_0 is the exact ground-state energy. A good enough wave function ansatz converges to the ground state in the optimization process. In each iteration, the Metropolis-Hastings Monte Carlo method [54, 55] is used to evaluate the energy expectation and its gradient with respect to the parameters $\nabla_{\boldsymbol{\theta}} E_{\boldsymbol{\theta}}$. The parameters are then updated using the stochastic reconfiguration [56], a commonly used optimization method in VMC [41, 43],

$$\boldsymbol{\theta}^{i+1} = \boldsymbol{\theta}^i - \eta(S + \epsilon I)^{-1}\nabla_{\boldsymbol{\theta}^i} E_{\boldsymbol{\theta}^i}, \tag{8}$$

where i is the iteration step, η is the learning rate, $\epsilon = 10^{-3}$ is taken for numerical stability, and S is the Quantum Fisher information matrix,

$$S_{ab} = \frac{\langle\partial_{\theta_a}\psi_{\boldsymbol{\theta}}|\partial_{\theta_b}\psi_{\boldsymbol{\theta}}\rangle}{\langle\psi_{\boldsymbol{\theta}}|\psi_{\boldsymbol{\theta}}\rangle} - \frac{\langle\partial_{\theta_a}\psi_{\boldsymbol{\theta}}|\psi_{\boldsymbol{\theta}}\rangle}{\langle\psi_{\boldsymbol{\theta}}|\psi_{\boldsymbol{\theta}}\rangle}\frac{\langle\psi_{\boldsymbol{\theta}}|\partial_{\theta_b}\psi_{\boldsymbol{\theta}}\rangle}{\langle\psi_{\boldsymbol{\theta}}|\psi_{\boldsymbol{\theta}}\rangle}. \tag{9}$$

More details of the optimization process can be found in the Appendix B.

IV. RESULTS AND DISCUSSIONS

As a warm-up exercise, we test DeepQuark on few-electron systems, including e^+e^- (Ps), $e^+e^-e^-$ (Ps^-) and $e^+e^+e^-e^-$ (Ps_2). They are dominated by simple Coulomb interaction and can be regarded as the quantum electrodynamics counterparts of multiquark systems [57]. The DeepQuark results can reach a high accuracy with less than 1‰ relative difference compared to the benchmark energies (see Appendix C).

The optimization performance of DeepQuark in hadron systems is shown in Fig. 2. We take four systems as examples, including the nucleon, doubly charmed and fully charmed tetraquarks, and triply bottomed pentaquark. It can be seen that DeepQuark achieves good convergence within a few thousand iterations for all these systems. The statistical errors of the DeepQuark energy results are less than 0.1 MeV, which are significantly smaller than the model uncertainty and can be neglected. In the nucleon system, DeepQuark quickly converges to the benchmark results from GEM and DMC [58] for both two-body AL1 interaction in Eq. (2) and flux-tube confinement interaction in Eq. (3). Calculating flux-tube interactions is computationally intractable in basis expansion methods, whereas DeepQuark handles them efficiently through Monte Carlo evaluation. This highlights DeepQuark's ability in managing complicate potentials of various forms.

For tetraquark states, we investigate two systems that are of most concern, the isoscalar doubly heavy tetraquark $QQ\bar{q}\bar{q}$ (T_{QQ}) and fully heavy tetraquark $QQ\bar{Q}\bar{Q}$ (T_{4Q}). The existence of a deeply bound T_{QQ} state for large enough heavy quark mass has long been anticipated [60, 61] and was also predicted by lattice QCD study [62, 63]. Moreover, great experimental progress has been made in the discovery of T_{cc} [6, 7]. The T_{4Q} system serves as a clear platform to investigate the short-range gluon exchange and confinement interaction, as it is less affected by the chiral dynamics. Bound states in such a system are inconclusive but a family of T_{4c} resonances have been discovered experimentally [10, 64, 65]. Recently, the CMS collaboration determined three T_{4c} resonances with high significances and analyzed their quantum numbers to be $J^{PC} = 2^{++}$ [66]. In Table I, we present the ground-state properties of these tetraquark states and compare them with the results from GEM, which was shown to be a superior numerical approach to solving tetraquark bound states [59]. For T_{QQ} systems, the DeepQuark ground-state energies are consistent with the GEM results in general, and lower the energies of T_{cc} states by $\simeq 1$ MeV. This shows that DeepQuark has a stronger expressive power for the multiquark wave function than GEM. The ground state of T_{bb} is dominated by the $\chi_{\bar{3}_c\otimes 3_c}$ color configuration, whereas T_{cc} exhibits sizable contributions from both $\chi_{\bar{3}_c\otimes 3_c}$ and $\chi_{6_c\otimes\bar{6}_c}$ configurations. Starting from a randomly initialized trial wave function, DeepQuark naturally converges to these two ground states with distinct mixing effects, demonstrating its strong capability in handling coupled-channel problems. On the other hand, no bound state solution is found in T_{4Q} systems. The ground-state energies lie above the lowest meson-meson thresholds, and the color proportions $\chi_{\bar{3}_c\otimes 3_c} : \chi_{6_c\otimes\bar{6}_c} \simeq 1:2$ are in accordance with the meson-meson scattering states [67]. However, since the DeepQuark wave function is constrained in a localized space, it cannot converge to an ideal scattering state with zero relative momentum. Therefore, the ground-state energies are ~ 10 MeV above the thresholds and the color proportions slightly deviates from $\chi_{\bar{3}_c\otimes 3_c} : \chi_{6_c\otimes\bar{6}_c} \simeq 1:2$.

The DeepQuark wave function can consistently describe multiquark states with various configurations, including compact tetraquark and meson molecule. The root-mean-square (rms) radii of T_{bb} and T_{cc} are shown in Table I. T_{bb} is a distinct compact tetraquark, where two bottom quarks form a compact diquark cluster of 0.33 fm and the light quarks orbit around the heavy diquark. In contrast to the compact T_{bb}, the size of T_{cc} is significantly larger, with $r_{cc}, r_{\bar{q}\bar{q}} > r_{c\bar{q}}$, suggesting a molecular structure of two charmed mesons [68]. A loosely bound molecular T_{cc} is in agreement with the experimental results [6, 7]. The separation between two charmed mesons becomes larger when the binding energy is tuned to be as small as the experimental one [69].

Inspired by the existence of $\bar{Q}\bar{Q}qq$ bound states, we further employ DeepQuark to investigate the isoscalar triply heavy pentaquark ($QQqq\bar{Q}$), which is the partner of the doubly heavy tetraquark if we consider the diquark-antiquark symmetry [70, 71], namely replacing a heavy antiquark $\bar{Q}$ by a heavy diquark QQ. As the two heavy quarks may form a compact object with the same color representation $\bar{3}_c$ as the antiquark, it is expected that the interactions in these two systems are similar, and $\bar{Q}\bar{Q}qq$ bound states may imply the existence of $QQqq\bar{Q}$ bound states. DeepQuark has great advantages in extending to such pentaquark systems, since the computational complexity is almost the same as solving the tetraquark systems. A complete set of color basis for the pentaquark is given by,

$$\begin{aligned}
\chi_{\bar{3}_c\otimes\bar{3}_c} &= \left\{\left[(QQ)_{\bar{3}_c}(qq)_{\bar{3}_c}\right]_{3_c}\bar{Q}\right\}_{1_c},\\
\chi_{\bar{3}_c\otimes 6_c} &= \left\{\left[(QQ)_{\bar{3}_c}(qq)_{6_c}\right]_{3_c}\bar{Q}\right\}_{1_c},\\
\chi_{6_c\otimes\bar{3}_c} &= \left\{\left[(QQ)_{6_c}(qq)_{\bar{3}_c}\right]_{3_c}\bar{Q}\right\}_{1_c}.
\end{aligned} \tag{10}$$

We show the ground-state properties in Table II. In contrast to the expectations from diquark-antiquark symmetry, we find no bound state in the $S = \frac{1}{2}, \frac{3}{2}$ $QQqq\bar{Q}$ systems. The reason is that the existence of $\bar{Q}$ breaks the diquark-antiquark symmetry by taking one of the heavy quarks to form a quarkonium $(Q\bar{Q})_{1_c}$, which is more stable and has lower energy than the configuration with $(QQ)_{\bar{3}_c}$. The ground states in these systems are scattering states of a heavy quarkonium and a singly heavy baryon. However, in $S = \frac{5}{2}$ system, an S-wave $S = \frac{3}{2}$ singly heavy isoscalar baryon is not allowed by the antisymmetry of light quarks. As a result, the lowest threshold is $\bar{D}^*\Xi_{cc}^*$ for $ccqq\bar{c}$ and $B^*\Xi_{bb}^*$ for $bbqq\bar{b}$. We find bound state solutions $P_{cc\bar{c}}(5715)$ and $P_{bb\bar{b}}(15569)$, whose binding energies are 3 MeV and 14 MeV, respectively. They may decay to meson-baryon channels with lower total spin if spin-orbit coupling is considered. For example, $P_{cc\bar{c}}(5715)$ may be searched for in the $J/\psi\Lambda_c$ channel, but its width is expected to be suppressed by the D-wave decay. From the rms

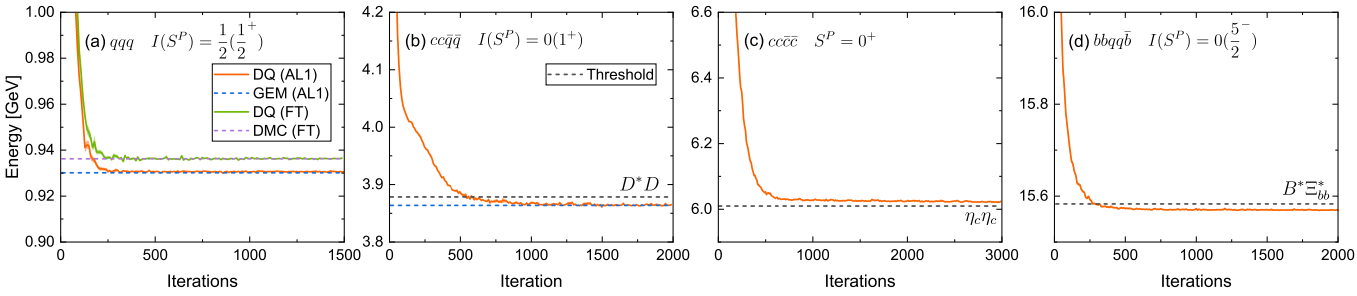

FIG. 2. The energy estimate as a function of iteration steps for the (a) nucleon in the AL1 potential and flux-tube confinement interaction (FT), (b) isoscalar vector doubly charmed tetraquark, (c) scalar fully charmed tetraquark, and (d) isoscalar triply bottomed tetraquark systems in the optimization progress of DeepQuark (DQ). The Monte Carlo standard errors of the energies are shown by the shaded area, which are very tiny. The lowest two-body dissociation thresholds are represented by the black dashed lines. The ground-state energies given by the Gaussian expansion method (GEM) [58, 59] and diffusion Monte Carlo (DMC) [58] are respectively displayed by the blue and purple dashed lines for comparison.

TABLE I. Ground-state energies, color proportions and rms radii (in fm) of the isoscalar doubly heavy ($QQ\bar{q}\bar{q}$) and fully heavy ($QQ\bar{Q}\bar{Q}$) tetraquark systems. The binding energies ΔE (in MeV) are with respect to the lowest dihadron thresholds listed in the third column. "NB" indicates no bound state solution. The statistical errors of the energies from DeepQuark (DQ) are less than 0.1 MeV. The ground-state energies from GEM [59] are listed for comparison.

			DQ						GEM [59]
	S^P	Thresholds	ΔE	$\chi_{\bar{3}_c\otimes 3_c}$	$\chi_{6_c\otimes\bar{6}_c}$	$r_{Q\bar{q}}$	r_{QQ}	$r_{\bar{q}\bar{q}}$	ΔE
$cc\bar{q}\bar{q}$	1^+	D^*D	-15	55%	45%	1.06	1.24	1.41	-14
$bb\bar{q}\bar{q}$	1^+	$\bar{B}^*B$	-153	97%	3%	0.69	0.33	0.78	-153
$cc\bar{c}\bar{c}$	$0^+, 1^+, 2^+$	$\eta_c\eta_c, \eta_cJ/\psi, J/\psi J/\psi$	NB	∼34%	∼66%				NB
$bb\bar{b}\bar{b}$	$0^+, 1^+, 2^+$	$\eta_b\eta_b, \eta_b\Upsilon, \Upsilon\Upsilon$	NB	∼34%	∼66%				NB

TABLE II. Ground-state energies, color proportions and rms radii (in fm) of the isoscalar triply heavy pentaquark ($QQqq\bar{Q}$) systems. The remaining notations follow the same conventions as in Table I.

	S^P	Thresholds	ΔE	$\chi_{\bar{3}_c\otimes\bar{3}_c}$	$\chi_{\bar{3}_c\otimes 6_c}$	$\chi_{6_c\otimes\bar{3}_c}$	r_{QQ}	r_{Qq}	r_{qq}	$r_{Q\bar{Q}}$	$r_{q\bar{Q}}$
$ccqq\bar{c}$	$\frac{1}{2}^-, \frac{3}{2}^-$	$\eta_c\Lambda_c, J/\psi\Lambda_c$	NB	∼35%	0%	∼65%					
	$\frac{5}{2}^-$	$\bar{D}^*\Xi^*_{cc}$	-3	27%	73%	0%	0.50	1.39	1.90	1.73	1.38
$bbqq\bar{b}$	$\frac{1}{2}^-, \frac{3}{2}^-$	$\eta_b\Lambda_b, \Upsilon\Lambda_b$	NB	∼35%	0%	∼65%					
	$\frac{5}{2}^-$	$B^*\Xi^*_{bb}$	-14	19%	80%	1%	0.30	0.89	1.22	0.88	0.88

radii of $P_{cc\bar{c}}(5715)$, the distance $r_{cc} = 0.5$ fm is consistent with the compact size of Ξ^*_{cc}, whereas $r_{c\bar{c}} = 1.73$ fm is substantially larger. The spatial separation suggests a molecular configuration composed of loosely bound $\bar{D}^*$ and Ξ^*_{cc}, which is analogous to the molecular T_{cc}. $P_{bb\bar{b}}(15569)$ exhibits a similar configuration, but its size is compacted by the heavy quark mass.

V. CONCLUSIONS AND OUTLOOKS

For the first time, we implement the DNN-based VMC approach for the multiquark bound states, whose complexity surpasses that of electron or nucleon systems due to strong SU(3) color interactions. We design a novel and high-efficiency architecture, DeepQuark, to address the unique challenges in multiquark systems like stronger correlations, extra discrete quantum numbers, and intractable confinement interaction. We have shown that DeepQuark is competitive with state-of-the-art approaches including DMC and GEM in baryon and tetraquark systems, reaching high accuracies on ground-state energies with low computational costs. In the nucleon, DeepQuark easily adapts both two-body and flux-tube interactions. In tetraquark systems, it consistently describe meson molecule T_{cc} and compact tetraquark T_{bb} with an unbiased form of wave function ansatz. Moreover, we use DeepQuark to investigate triply heavy pentaquark ($QQqq\bar{Q}$) systems. We obtain weakly bound $\bar{D}^*\Xi^*_{cc}$ molecule $P_{cc\bar{c}}(5715)$ with $S = \frac{5}{2}$ and its bottom partner $P_{bb\bar{b}}(15569)$. They can be viewed as the analogs of the molecular T_{cc}. We recommend experimental search of $P_{cc\bar{c}}(5715)$ in the D-wave $J/\psi\Lambda_c$ channel.

DeepQuark holds great promise for extension to diverse pentaquark and even hexaquark systems, overcoming the computational barriers that limit conventional methods. Such

investigations can provide forward-looking predictions for future experiments. Furthermore, DeepQuark serves as a powerful framework for exploring confining mechanism beyond two-body interactions in tetraquark states, which may offer valuable insights into multiquark inner structures and nonperturbative QCD. The techniques employed in DeepQuark could enrich the toolkit of deep learning approaches integrated with physics, particularly shedding new light on quantum many-body phenomena.

ACKNOWLEDGMENT

We thank Yao Ma, Yan-Ke Chen and Liang-Zhen Wen for the helpful discussions. L.M. also thank Yilong Yang and Pengwei Zhao for discussions on DNN and VMC. This project was supported by the National Natural Science Foundation of China (No. 12475137), and ERC NuclearTheory (Grant No. 885150). The codes are developed based on the NetKet package [72]. The computational resources are supported by High-performance Computing Platform of Peking University.

DATA AVAILABILITY

The data supporting this study's findings are available within the article.

[1] M. Gell-Mann, A Schematic Model of Baryons and Mesons, Phys. Lett. **8**, 214 (1964).
[2] G. Zweig, An SU(3) model for strong interaction symmetry and its breaking. Version 1 10.17181/CERN-TH-401 (1964).
[3] S. K. Choi *et al.* (Belle), Observation of a narrow charmonium-like state in exclusive $B^{\pm} \to K^{\pm}\pi^{+}\pi^{-}J/\psi$ decays, Phys. Rev. Lett. **91**, 262001 (2003), arXiv:hep-ex/0309032.
[4] M. Ablikim *et al.* (BESIII), Observation of a Charged Charmoniumlike Structure in $e^{+}e^{-} \to \pi^{+}\pi^{-}J/\psi$ at $\sqrt{s}$ =4.26 GeV, Phys. Rev. Lett. **110**, 252001 (2013), arXiv:1303.5949 [hep-ex].
[5] Z. Q. Liu *et al.* (Belle), Study of $e^{+}e^{-}ß\pi^{+}\pi^{-}J/\psi$ and Observation of a Charged Charmoniumlike State at Belle, Phys. Rev. Lett. **110**, 252002 (2013), [Erratum: Phys.Rev.Lett. 111, 019901 (2013)], arXiv:1304.0121 [hep-ex].
[6] R. Aaij *et al.* (LHCb), Observation of an exotic narrow doubly charmed tetraquark, Nature Phys. **18**, 751 (2022), arXiv:2109.01038 [hep-ex].
[7] R. Aaij *et al.* (LHCb), Study of the doubly charmed tetraquark T_{cc}^{+}, Nature Commun. **13**, 3351 (2022), arXiv:2109.01056 [hep-ex].
[8] R. Aaij *et al.* (LHCb), Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^{-} p$ Decays, Phys. Rev. Lett. **115**, 072001 (2015), arXiv:1507.03414 [hep-ex].
[9] R. Aaij *et al.* (LHCb), Observation of a narrow pentaquark state, $P_c(4312)^{+}$, and of two-peak structure of the $P_c(4450)^{+}$, Phys. Rev. Lett. **122**, 222001 (2019), arXiv:1904.03947 [hep-ex].
[10] R. Aaij *et al.* (LHCb), Observation of structure in the J/ψ -pair mass spectrum, Sci. Bull. **65**, 1983 (2020), arXiv:2006.16957 [hep-ex].
[11] H.-X. Chen, W. Chen, X. Liu, and S.-L. Zhu, The hidden-charm pentaquark and tetraquark states, Phys. Rept. **639**, 1 (2016), arXiv:1601.02092 [hep-ph].
[12] A. Hosaka, T. Iijima, K. Miyabayashi, Y. Sakai, and S. Yasui, Exotic hadrons with heavy flavors: X, Y, Z, and related states, PTEP **2016**, 062C01 (2016), arXiv:1603.09229 [hep-ph].
[13] R. F. Lebed, R. E. Mitchell, and E. S. Swanson, Heavy-Quark QCD Exotica, Prog. Part. Nucl. Phys. **93**, 143 (2017), arXiv:1610.04528 [hep-ph].
[14] F.-K. Guo, C. Hanhart, U.-G. Meißner, Q. Wang, Q. Zhao, and B.-S. Zou, Hadronic molecules, Rev. Mod. Phys. **90**, 015004 (2018), [Erratum: Rev.Mod.Phys. 94, 029901 (2022)], arXiv:1705.00141 [hep-ph].
[15] Y.-R. Liu, H.-X. Chen, W. Chen, X. Liu, and S.-L. Zhu, Pentaquark and Tetraquark states, Prog. Part. Nucl. Phys. **107**, 237 (2019), arXiv:1903.11976 [hep-ph].
[16] N. Brambilla, S. Eidelman, C. Hanhart, A. Nefediev, C.-P. Shen, C. E. Thomas, A. Vairo, and C.-Z. Yuan, The XYZ states: experimental and theoretical status and perspectives, Phys. Rept. **873**, 1 (2020), arXiv:1907.07583 [hep-ex].
[17] H.-X. Chen, W. Chen, X. Liu, Y.-R. Liu, and S.-L. Zhu, An updated review of the new hadron states, Rept. Prog. Phys. **86**, 026201 (2023), arXiv:2204.02649 [hep-ph].
[18] L. Meng, B. Wang, G.-J. Wang, and S.-L. Zhu, Chiral perturbation theory for heavy hadrons and chiral effective field theory for heavy hadronic molecules, Phys. Rept. **1019**, 1 (2023), arXiv:2204.08716 [hep-ph].
[19] F. Okiharu, H. Suganuma, and T. T. Takahashi, Detailed analysis of the tetraquark potential and flip-flop in SU(3) lattice QCD, Phys. Rev. D **72**, 014505 (2005), arXiv:hep-lat/0412012.
[20] M. Cardoso, N. Cardoso, and P. Bicudo, Variational study of the flux tube recombination in the two quarks and two quarks system in Lattice QCD, Phys. Rev. D **86**, 014503 (2012), arXiv:1204.5131 [hep-lat].
[21] P. Bicudo and M. Cardoso, Tetraquark bound states and resonances in the unitary and microscopic triple string flip-flop quark model, the light-light-antiheavy-antiheavy $qq\bar{Q}\bar{Q}$ case study, Phys. Rev. D **94**, 094032 (2016), arXiv:1509.04943 [hep-ph].
[22] E. Hiyama, Y. Kino, and M. Kamimura, Gaussian expansion method for few-body systems, Prog. Part. Nucl. Phys. **51**, 223 (2003).
[23] M. Troyer and U.-J. Wiese, Computational complexity and fundamental limitations to fermionic quantum Monte Carlo simulations, Phys. Rev. Lett. **94**, 170201 (2005), arXiv:cond-mat/0408370.
[24] E. Hiyama, A. Hosaka, M. Oka, and J.-M. Richard, Quark model estimate of hidden-charm pentaquark resonances, Phys. Rev. C **98**, 045208 (2018), arXiv:1803.11369 [nucl-th].

[25] J. F. Giron and R. F. Lebed, Fine structure of pentaquark multiplets in the dynamical diquark model, Phys. Rev. D **104**, 114028 (2021), arXiv:2110.05557 [hep-ph].

[26] Y. Yan, Y. Wu, X. Hu, H. Huang, and J. Ping, Fully heavy pentaquarks in quark models, Phys. Rev. D **105**, 014027 (2022), arXiv:2110.10853 [hep-ph].

[27] G. Yang, J. Ping, and J. Segovia, Fully charm and bottom pentaquarks in a lattice-QCD inspired quark model, Phys. Rev. D **106**, 014005 (2022), arXiv:2205.11548 [hep-ph].

[28] H.-T. An, S.-Q. Luo, Z.-W. Liu, and X. Liu, Fully heavy pentaquark states in constituent quark model, Phys. Rev. D **105**, 074032 (2022), arXiv:2203.03448 [hep-ph].

[29] Z.-B. Liang, F.-X. Liu, and X.-H. Zhong, All-heavy pentaquarks, Phys. Rev. D **111**, 056013 (2025), arXiv:2402.17974 [hep-ph].

[30] M. C. Gordillo, J. Segovia, and J. M. Alcaraz-Pelegrina, Diffusion Monte Carlo calculation of fully heavy pentaquarks, Phys. Rev. D **110**, 094024 (2024), arXiv:2409.04130 [hep-ph].

[31] Y. LeCun, Y. Bengio, and G. Hinton, Deep learning, Nature **521**, 436 (2015).

[32] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, Machine learning and the physical sciences, Rev. Mod. Phys. **91**, 045002 (2019), arXiv:1903.10563 [physics.comp-ph].

[33] X. Zhang, L. Wang, J. Helwig, Y. Luo, C. Fu, Y. Xie, M. Liu, Y. Lin, Z. Xu, K. Yan, K. Adams, M. Weiler, X. Li, T. Fu, Y. Wang, A. Strasser, H. Yu, Y. Xie, X. Fu, S. Xu, Y. Liu, Y. Du, A. Saxton, H. Ling, H. Lawrence, H. Stärk, S. Gui, C. Edwards, N. Gao, A. Ladera, T. Wu, E. F. Hofgard, A. M. Tehrani, R. Wang, A. Daigavane, M. Bohde, J. Kurtin, Q. Huang, T. Phung, M. Xu, C. K. Joshi, S. V. Mathis, K. Azizzadenesheli, A. Fang, A. Aspuru-Guzik, E. Bekkers, M. Bronstein, M. Zitnik, A. Anandkumar, S. Ermon, P. Liò, R. Yu, S. Günnemann, J. Leskovec, H. Ji, J. Sun, R. Barzilay, T. Jaakkola, C. W. Coley, X. Qian, X. Qian, T. Smidt, and S. Ji, Artificial intelligence for science in quantum, atomistic, and continuum systems (2025), arXiv:2307.08423 [cs.LG].

[34] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, Science **355**, 602 (2017).

[35] J. Han, L. Zhang, and W. E, Solving many-electron schrödinger equation using deep neural networks, Journal of Computational Physics **399**, 108929 (2019).

[36] D. Pfau, J. S. Spencer, A. G. D. G. Matthews, and W. M. C. Foulkes, Ab initio solution of the many-electron schrödinger equation with deep neural networks, Phys. Rev. Res. **2**, 033429 (2020).

[37] J. Hermann, Z. Schätzle, and F. Noé, Deep-neural-network solution of the electronic schrödinger equation, Nature Chemistry **12**, 891 (2020).

[38] X. Li, Z. Li, and J. Chen, Ab initio calculation of real solids via neural network ansatz, Nature Communications **13**, 7895 (2022).

[39] J. Kim, G. Pescia, B. Fore, J. Nys, G. Carleo, S. Gandolfi, M. Hjorth-Jensen, and A. Lovato, Neural-network quantum states for ultra-cold Fermi gases, Commun. Phys. **7**, 148 (2024), arXiv:2305.08831 [cond-mat.quant-gas].

[40] J. Keeble and A. Rios, Machine learning the deuteron, Physics Letters B **809**, 135743 (2020).

[41] C. Adams, G. Carleo, A. Lovato, and N. Rocco, Variational Monte Carlo Calculations of A$\leq$4 Nuclei with an Artificial Neural-Network Correlator Ansatz, Phys. Rev. Lett. **127**, 022502 (2021), arXiv:2007.14282 [nucl-th].

[42] Y. L. Yang and P. W. Zhao, A consistent description of the relativistic effects and three-body interactions in atomic nuclei, Phys. Lett. B **835**, 137587 (2022), arXiv:2206.13208 [nucl-th].

[43] Y. Yang and P. Zhao, Deep-neural-network approach to solving the ab initio nuclear structure problem, Phys. Rev. C **107**, 034320 (2023), arXiv:2211.13998 [nucl-th].

[44] Y.-L. Yang and P.-W. Zhao, Reconciling Light Nuclei and Nuclear Matter: Relativistic ab initio Calculations, Chin. Phys. Lett. **42**, 051201 (2025), arXiv:2405.04203 [nucl-th].

[45] B. Fore, J. Kim, M. Hjorth-Jensen, and A. Lovato, Investigating the crust of neutron stars with neural-network quantum states, Commun. Phys. **8**, 108 (2025), arXiv:2407.21207 [nucl-th].

[46] Y. Yang, E. Epelbaum, J. Meng, L. Meng, and P. Zhao, Chiral symmetry and peripheral neutron-α scattering, (2025), arXiv:2502.09961 [nucl-th].

[47] C. Semay and B. Silvestre-Brac, Diquonia and potential models, Z. Phys. C **61**, 271 (1994).

[48] B. Silvestre-Brac, Spectrum and static properties of heavy baryons, Few Body Syst. **20**, 1 (1996).

[49] E. Eichten, K. Gottfried, T. Kinoshita, K. D. Lane, and T.-M. Yan, Charmonium: The Model, Phys. Rev. D **17**, 3090 (1978), [Erratum: Phys.Rev.D 21, 313 (1980)].

[50] E. Eichten, K. Gottfried, T. Kinoshita, K. D. Lane, and T.-M. Yan, Charmonium: Comparison with Experiment, Phys. Rev. D **21**, 203 (1980).

[51] X. Artru, String Model with Baryons: Topology, Classical Motion, Nucl. Phys. B **85**, 442 (1975).

[52] T. T. Takahashi, H. Matsufuru, Y. Nemoto, and H. Suganuma, The Three quark potential in the SU(3) lattice QCD, Phys. Rev. Lett. **86**, 18 (2001), arXiv:hep-lat/0006005.

[53] T. T. Takahashi, H. Suganuma, Y. Nemoto, and H. Matsufuru, Detailed analysis of the three quark potential in SU(3) lattice QCD, Phys. Rev. D **65**, 114509 (2002), arXiv:hep-lat/0204011.

[54] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, Equation of state calculations by fast computing machines, The Journal of Chemical Physics **21**, 1087 (1953), https://pubs.aip.org/aip/jcp/article-pdf/21/6/1087/18802390/1087_1_online.pdf.

[55] W. K. Hastings, Monte carlo sampling methods using markov chains and their applications, Biometrika **57**, 97 (1970), https://academic.oup.com/biomet/article-pdf/57/1/97/23940249/57-1-97.pdf.

[56] S. Sorella, Wave function optimization in the variational monte carlo method, Phys. Rev. B **71**, 241103 (2005).

[57] Y. Ma, L. Meng, L.-Z. Wen, and S.-L. Zhu, Trilepton and tetralepton bound and resonant states: The QED counterpart of multiquark states, Phys. Rev. D **111**, 073001 (2025), arXiv:2501.00871 [hep-ph].

[58] Y. Ma, L. Meng, Y.-K. Chen, and S.-L. Zhu, Ground state baryons in the flux-tube three-body confinement model using diffusion Monte Carlo, Phys. Rev. D **107**, 054035 (2023), arXiv:2211.09021 [hep-ph].

[59] L. Meng, Y.-K. Chen, Y. Ma, and S.-L. Zhu, Tetraquark bound states in constituent quark models: Benchmark test calculations, Phys. Rev. D **108**, 114016 (2023), arXiv:2310.13354 [hep-ph].

[60] S. Zouzou, B. Silvestre-Brac, C. Gignoux, and J. M. Richard, FOUR QUARK BOUND STATES, Z. Phys. C **30**, 457 (1986).

[61] A. V. Manohar and M. B. Wise, Exotic Q Q anti-q anti-q states in QCD, Nucl. Phys. B **399**, 17 (1993), arXiv:hep-ph/9212236.

[62] A. Francis, R. J. Hudspith, R. Lewis, and K. Maltman, Lattice Prediction for Deeply Bound Doubly Heavy Tetraquarks, Phys. Rev. Lett. **118**, 142001 (2017), arXiv:1607.05214 [hep-lat].

[63] P. Junnarkar, N. Mathur, and M. Padmanath, Study of doubly heavy tetraquarks in Lattice QCD, Phys. Rev. D **99**, 034507 (2019), arXiv:1810.12285 [hep-lat].

[64] A. Hayrapetyan *et al.* (CMS), New Structures in the J/ψJ/ψ Mass Spectrum in Proton-Proton Collisions at s=13 TeV, Phys. Rev. Lett. **132**, 111901 (2024), arXiv:2306.07164 [hep-ex].

[65] G. Aad *et al.* (ATLAS), Observation of an Excess of Dicharmonium Events in the Four-Muon Final State with the ATLAS Detector, Phys. Rev. Lett. **131**, 151902 (2023), arXiv:2304.08962 [hep-ex].

[66] A. Hayrapetyan *et al.* (CMS), Determination of the spin and parity of all-charm tetraquarks, (2025), arXiv:2506.07944 [hep-ex].

[67] W.-L. Wu, Y.-K. Chen, L. Meng, and S.-L. Zhu, Benchmark calculations of fully heavy compact and molecular tetraquark states, Phys. Rev. D **109**, 054034 (2024), arXiv:2401.14899 [hep-ph].

[68] N. Li, Z.-F. Sun, X. Liu, and S.-L. Zhu, Coupled-channel analysis of the possible $D^{(*)}D^{(*)}$, $\overline{B}^{(*)}\overline{B}^{(*)}$ and $D^{(*)}\overline{B}^{(*)}$ molecular states, Phys. Rev. D **88**, 114008 (2013), arXiv:1211.5007 [hep-ph].

[69] W.-L. Wu, Y. Ma, Y.-K. Chen, L. Meng, and S.-L. Zhu, Doubly heavy tetraquark bound and resonant states, Phys. Rev. D **110**, 094041 (2024), arXiv:2409.03373 [hep-ph].

[70] F.-K. Guo, C. Hidalgo-Duque, J. Nieves, and M. P. Valderrama, Heavy-antiquark–diquark symmetry and heavy hadron molecules: Are there triply heavy pentaquarks?, Phys. Rev. D **88**, 054014 (2013), arXiv:1305.4052 [hep-ph].

[71] Z.-Y. Wang, C.-W. Xiao, Z.-F. Sun, and X. Liu, Possible molecules of triple-heavy pentaquarks within the extended local hidden gauge formalism, Phys. Rev. D **110**, 076014 (2024), arXiv:2407.13319 [hep-ph].

[72] F. Vicentini, D. Hofmann, A. Szabó, D. Wu, C. Roth, C. Giuliani, G. Pescia, J. Nys, V. Vargas-Calderón, N. Astrakhantsev, and G. Carleo, NetKet 3: Machine Learning Toolbox for Many-Body Quantum Systems, SciPost Phys. Codebases , 7 (2022).

[73] S. Navas *et al.* (Particle Data Group), Review of particle physics, Phys. Rev. D **110**, 030001 (2024).

[74] Y. K. Ho, Variational calculation of ground-state energy of positronium negative ions, Phys. Rev. A **48**, 4780 (1993).

[75] D. B. Kinghorn and R. D. Poshusta, Nonadiabatic variational calculations on dipositronium using explicitly correlated Gaussian basis functions, Phys. Rev. A **47**, 3671 (1993).

[76] S. Bubin and L. Adamowicz, Nonrelativistic variational calculations of the positronium molecule and the positronium hydride, Phys. Rev. A **74**, 052502 (2006).

Appendix A: Quark Model

The parameters of the AL1 potential [47, 48] adopted in this work are taken from Ref. [48] and listed in Table III. The masses of heavy mesons and baryons in the model are given in Table IV.

For the nucleon, we also test the flux-tube confinement model as shown in Fig. 3, where the linear pairwise confinement interactions $V_{\text{conf},ij}$ in Eq. (2) are replaced by the flux-tube interaction in Eq. (3). The string tension σ in Eq. (3) is given as $\sigma = 0.9204\lambda$ [58], which was determined by fitting the experimental mass of $\Omega^-(sss)$ and is consistent with the best fitting parameters in lattice QCD simulation [52]. In baryon systems, the minimal length of color flux tubes can be solved analytically [53] and is given as

$$L_{\text{min}} = \begin{cases} \left[\frac{1}{2}(a^2+b^2+c^2) + \frac{\sqrt{3}}{2}\sqrt{d(d-2a)(d-2b)(d-2c)}\right]^{1/2} & \max(\theta_a,\theta_b,\theta_c) < \frac{2\pi}{3}, \\ d - \max(a,b,c) & \max(\theta_a,\theta_b,\theta_c) > \frac{2\pi}{3}, \end{cases} \tag{A1}$$

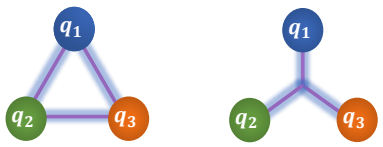

FIG. 3. Two confinement scenarios for the baryons. The left and right panels represent the pairwise confinement mechanism and the flux-tube confinement mechanism, respectively.

where $\theta_a, \theta_b, \theta_c$ are the interior angles of the baryon triangles, a, b, c are the lengths of sides, and $d = a + b + c$.

Appendix B: Details of the optimization process

The deep neural network (DNN) in the DeepQuark wave function consists of the input features $\boldsymbol{x} = (\boldsymbol{r}_i, |\boldsymbol{r}_i - \boldsymbol{r}_j|, \alpha_c, \alpha_s, \alpha_t)$, four hidden fully connected layers and a one-dimensional output $f_{NN}(\boldsymbol{x})$. Here $\boldsymbol{r}_i, \alpha_c, \alpha_s, \alpha_t$ are the spatial three-dimensional coordinates in the center of mass frame and the encoded vectors of color, spin and isospin coupled bases, respectively. The number of nodes in each hidden layer and the total number of variational parameters in the DNN for different systems are given in Table V. We use larger DNNs for larger systems to ensure the flexibility of the wave function.

In each iteration step during the optimization process, we use the Metropolis-Hastings Monte Carlo algorithm [54, 55] to generate sample points that are distributed according to the probability

$$P(R,\alpha) = \frac{|\psi_{\boldsymbol{\theta}}(R,\alpha)|^2}{\sum_\alpha \int dR\, |\psi_{\boldsymbol{\theta}}(R,\alpha)|^2}, \tag{B1}$$

where $\psi_{\boldsymbol{\theta}}$ is the antisymmetric wave function, $R = (\boldsymbol{r}_1, \cdots, \boldsymbol{r}_N)$ and $\alpha = (\alpha_c, \alpha_s, \alpha_t)$ denote the continuous and discrete degrees of freedom, respectively. Each Metropo-

TABLE III. The parameters in the AL1 quark potential model.

κ	λ[GeV^2]	Λ[GeV]	κ'	m_b[GeV]	m_c[GeV]	m_q[GeV]	r_0[GeV^{-1}]	A[GeV^{B-1}]	B	C[GeV^4]
0.5069	0.1653	0.8321	1.8609	5.227	1.836	0.315	$A\left(\frac{2m_im_j}{m_i+m_j}\right)^{-B}$	1.6553	0.2204	2.02×10^{-3}

TABLE IV. The masses (in MeV) of heavy mesons and baryons in the AL1 quark potential model. The results from Gaussian expansion method and the experimental values are shown for comparison.

	J^P	DQ	GEM [58, 59]	EXP [73]
$D(c\bar{q})$	0^-	1862	1862	1867
$\bar{B}(b\bar{q})$		5294	5294	5279
$\eta_c(c\bar{c})$		3005	3005	2984
$\eta_b(b\bar{b})$		9424	9424	9399
$D^*(c\bar{q})$	1^-	2016	2016	2009
$\bar{B}^*(b\bar{q})$		5350	5350	5325
$J/\psi(c\bar{c})$		3101	3101	3097
$\Upsilon(b\bar{b})$		9462	9462	9460
$\Lambda_c^+(cqq)$	$\frac{1}{2}^+$	2290	2291	2286
$\Lambda_b^0(bqq)$		5636	5636	5620
$\Xi_{cc}^*(ccq)$	$\frac{3}{2}^+$	3702	3702	...
$\Xi_{bb}^*(bbq)$		10232	10232	...

TABLE V. The number of nodes in each hidden layer and the total number of variational parameters in the DNN for different systems.

Systems	S^P	Nodes	Parameters
e^+e^-	0^+	$(16,16,16,16)$	961
$e^+e^-e^-$	0^+	$(16,16,16,16)$	1041
$e^+e^+e^-e^-$	0^+	$(16,16,16,16)$	1137
qqq	$\frac{1}{2}^+$	$(16,16,16,16)$	1105
$QQ\bar{q}\bar{q}$	1^+	$(32,16,16,16)$	1889
$QQ\bar{Q}\bar{Q}$	0^+	$(32,16,16,16)$	1825
$QQ\bar{Q}\bar{Q}$	1^+	$(32,16,16,16)$	1857
$QQ\bar{Q}\bar{Q}$	2^+	$(32,16,16,16)$	1793
$QQqq\bar{Q}$	$\frac{1}{2}^-$	$(40,20,20,20)$	3081
$QQqq\bar{Q}$	$\frac{3}{2}^-$	$(40,20,20,20)$	3041
$QQqq\bar{Q}$	$\frac{5}{2}^-$	$(40,20,20,20)$	2921

lis step consists of a Gaussian kick for the spatial coordinates and a random flip of the coupled channels α [41]. The newly proposed configuration (R', α') is accepted with probabilities

$$P = \frac{|\psi_{\boldsymbol{\theta}}(R', \alpha')|^2}{|\psi_{\boldsymbol{\theta}}(R, \alpha)|^2}. \tag{B2}$$

The energy expectation value $E_{\boldsymbol{\theta}}$ and its gradient with respect to the parameters $\nabla_{\boldsymbol{\theta}} E_{\boldsymbol{\theta}}$ can be estimated using these sample points.

$$\begin{aligned} E_{\boldsymbol{\theta}} &= \frac{\sum_{\alpha\alpha'} \int dR\, \psi^*_{\boldsymbol{\theta}}(R,\alpha) H_{\alpha\alpha'}(R) \psi_{\boldsymbol{\theta}}(R,\alpha')}{\sum_\alpha \int dR\, |\psi_{\boldsymbol{\theta}}(R,\alpha)|^2} \\ &= \sum_{\alpha\alpha'} \int dR\, P(R,\alpha) \psi_{\boldsymbol{\theta}}^{-1}(R,\alpha) H_{\alpha\alpha'}(R) \psi_{\boldsymbol{\theta}}(R,\alpha') \\ &= \frac{1}{N} \sum_n E_L(R_n, \alpha_n), \end{aligned} \tag{B3}$$

where N is the number of sample points (R_n, α_n), and the local energy E_L is given by,

$$E_L(R,\alpha) = \sum_{\alpha'} \psi_{\boldsymbol{\theta}}(R,\alpha)^{-1} H_{\alpha\alpha'}(R) \psi_{\boldsymbol{\theta}}(R,\alpha'). \tag{B4}$$

Similarly, the gradient can be estimated as,

$$\begin{aligned} \nabla_{\boldsymbol{\theta}} E_{\boldsymbol{\theta}} = 2 \Bigg[&\frac{1}{N} \left(\sum_n E_L(R_n,\alpha_n) \nabla_{\boldsymbol{\theta}} \log \psi_{\boldsymbol{\theta}}(R_n,\alpha_n) \right) - \\ &\frac{1}{N^2} \left(\sum_n E_L(R_n,\alpha_n) \right) \left(\sum_n \nabla_{\boldsymbol{\theta}} \log \psi_{\boldsymbol{\theta}}(R_n,\alpha_n) \right) \Bigg]. \end{aligned} \tag{B5}$$

The parameters are then updated using the stochastic reconfiguration [56], a commonly used optimization method in VMC [41, 43],

$$\boldsymbol{\theta}^{i+1} = \boldsymbol{\theta}^i - \eta (S + \epsilon I)^{-1} \nabla_{\boldsymbol{\theta}^i} E_{\boldsymbol{\theta}^i}, \tag{B6}$$

where i is the iteration step, η is the learning rate, $\epsilon = 10^{-3}$ is taken for numerical stability, and S is the Quantum Fisher information matrix,

$$S_{ab} = \frac{\langle \partial_{\theta_a} \psi_{\boldsymbol{\theta}} | \partial_{\theta_b} \psi_{\boldsymbol{\theta}} \rangle}{\langle \psi_{\boldsymbol{\theta}} | \psi_{\boldsymbol{\theta}} \rangle} - \frac{\langle \partial_{\theta_a} \psi_{\boldsymbol{\theta}} | \psi_{\boldsymbol{\theta}} \rangle}{\langle \psi_{\boldsymbol{\theta}} | \psi_{\boldsymbol{\theta}} \rangle} \frac{\langle \psi_{\boldsymbol{\theta}} | \partial_{\theta_b} \psi_{\boldsymbol{\theta}} \rangle}{\langle \psi_{\boldsymbol{\theta}} | \psi_{\boldsymbol{\theta}} \rangle}. \tag{B7}$$

The NetKet package [72] provides efficient implementation of the stochastic reconfiguration algorithm.

During initial optimization, the trial wave function resides far from the ground state in parameter space. We generate a small sample of $N = 2 \times 10^4$ points to rapidly estimate $E_{\boldsymbol{\theta}}$ and $\nabla_{\boldsymbol{\theta}} E_{\boldsymbol{\theta}}$ with deliberate coarseness, enabling efficient early-stage convergence. As the DeepQuark wave function approaches convergence, we increase the sample size to $N = 4-8 \times 10^4$ to stabilize the optimization process. During the final optimization stage, we save 10 distinct sets of wave function parameters at a 50-step interval. The energy and other physical observables of these wave function configurations are subsequently evaluated using $N \sim 10^6$ sample points. The parameter set yielding the lowest energy expectation value is selected as the final ground-state wave function.

TABLE VI. Ground-state energies (in eV) of the few-electron systems. The statistical errors of the DeepQuark results are shown in the parentheses. Results from other methods are shown for comparison. Ps can be solved exactly while Ps^- and Ps_2 are solved using variational method.

	DQ	Other Methods
e^+e^-	-6.80301(16)	-6.803[a]
$e^+e^-e^-$	-7.12882(16)	-7.130 [74]
$e^+e^+e^-e^-$	-14.0347(7)	-14.04 [75, 76]

[a] exact

Appendix C: Results of electron systems

The ground-state energy results of few-electron systems are listed in Table VI. For comparative purposes, we include well-established benchmark results such as the exact solution of Ps, and high-precision variational results for Ps^- using Hylleraas-type wave functions [74] and Ps_2 using explicitly correlated Gaussian functions [75, 76].