APFB2025@Ho Chi Minh City

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

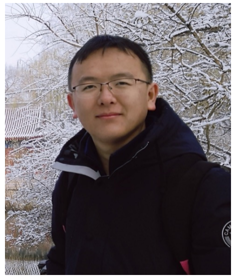
Lu Meng（孟璐）

Southeast University

Sep. 12, 2025

Shi-Lin Zhu

Wei-Lin Wu



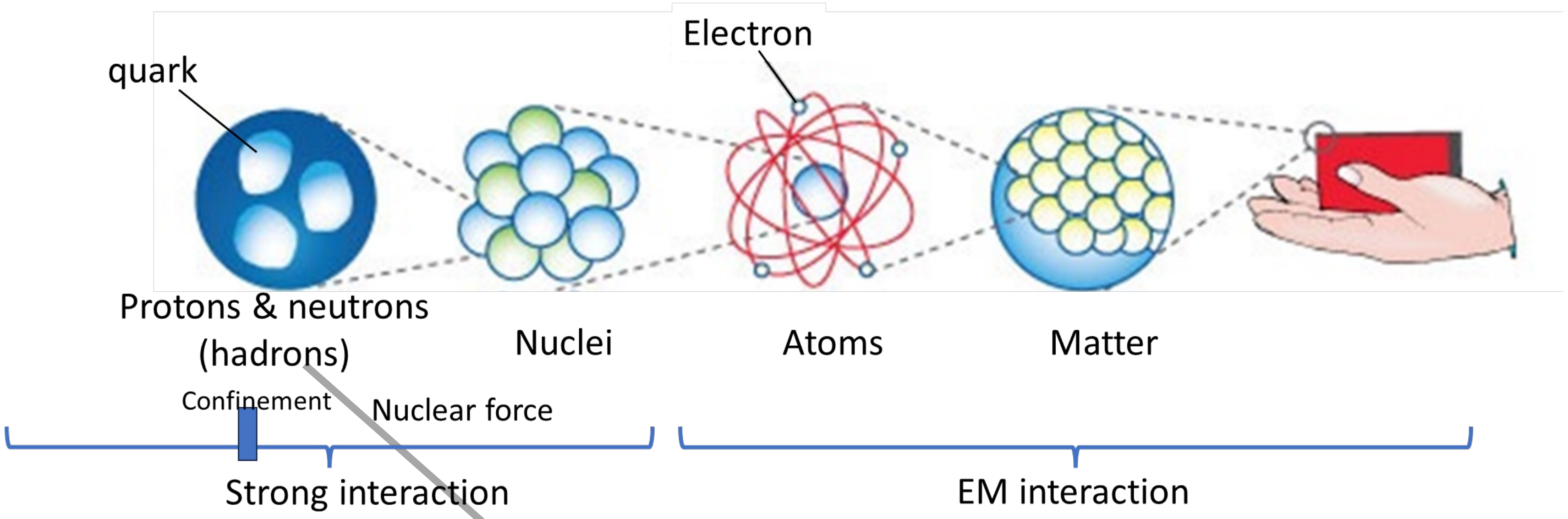
Lu Meng

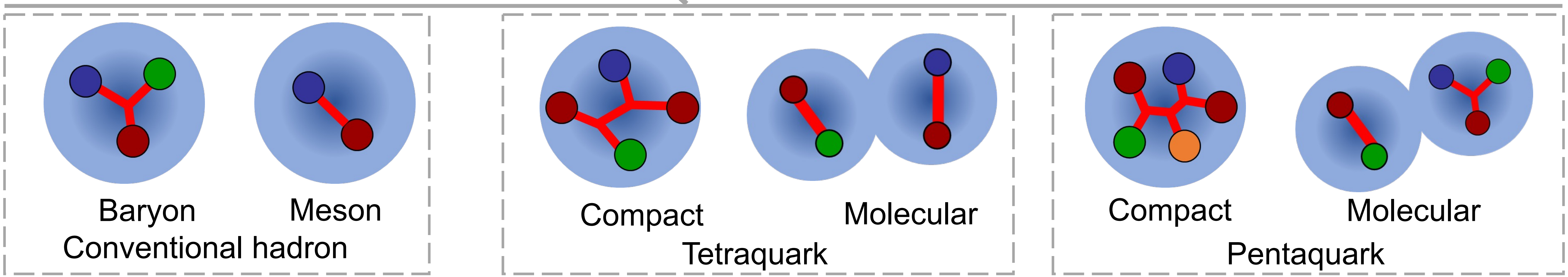
Based on arXiv: 2506.20555

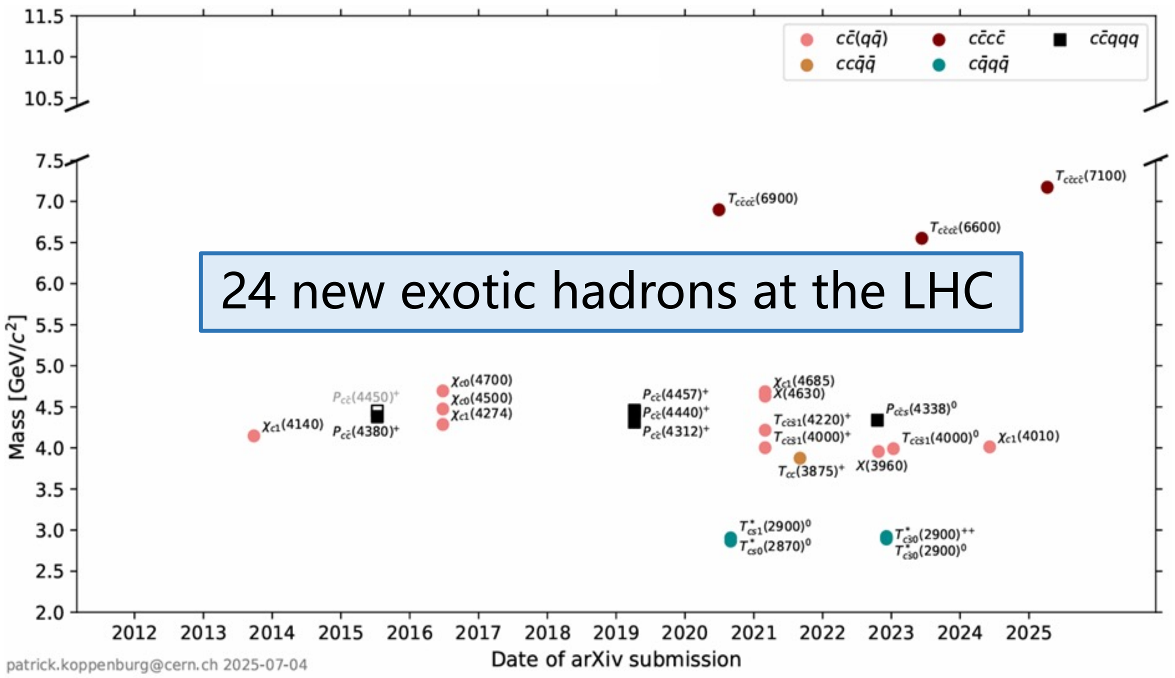
東南大學
SOUTHEAST UNIVERSITY

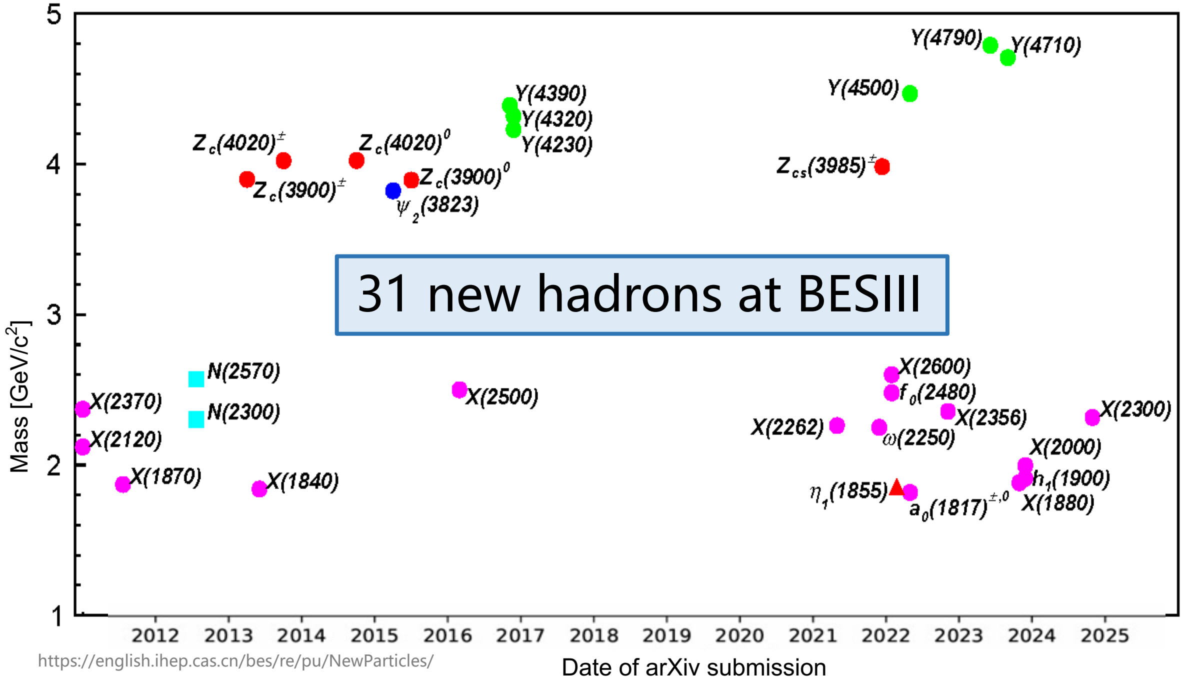
Background: Multiquark states

Multiquark states









Remaining questions

- The pattern of multiquark states
- Clustering of the quarks? Compact, or molecular or others?

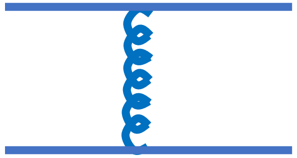
Quark potential models

- QCD becomes nonperturbative in low energy region
- Quark potential models

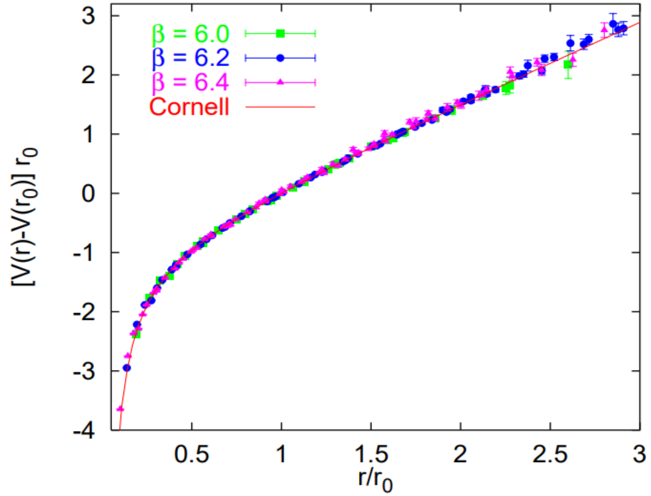
Cornell model: Eichten et al., PRD17 (1978) 3090
BGS model: Barnes et al., PRD 72 (2005) 054026
…

Minimal model :

$$V_{ij}(r) = \left[\underbrace{\frac{\alpha_s}{r} - \frac{8\pi\alpha_s}{3m_i m_j}\frac{\tau^3}{\pi^{3/2}}e^{-\tau^2 r^2}\boldsymbol{s}_i\cdot\boldsymbol{s}_j}_{\text{OGE}} + \underbrace{\left(-\frac{3b}{4}r + V_c\right)}_{\text{Confinement}}\right]\frac{\lambda_i\cdot\lambda_j}{4}$$



One-gluon-exchange



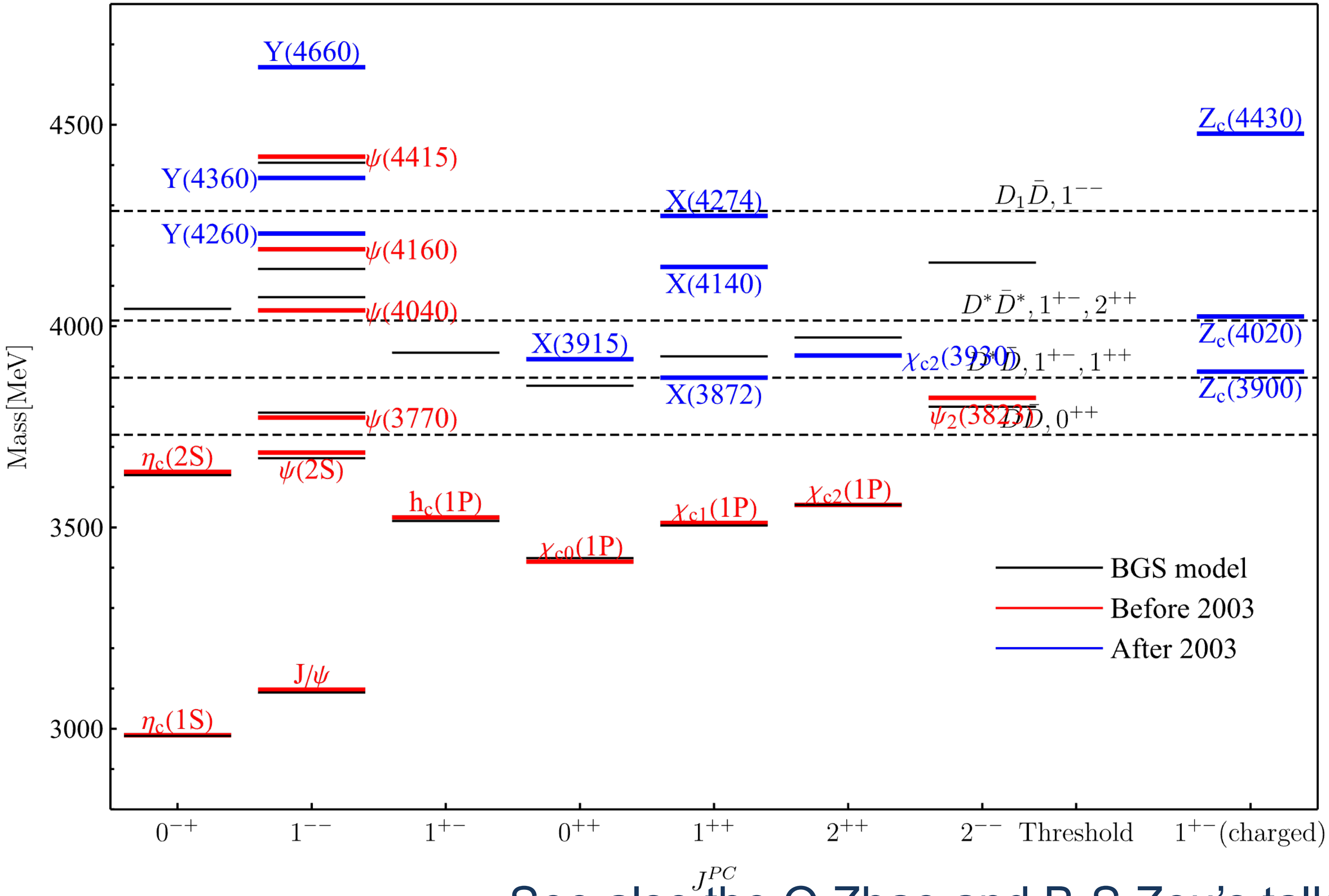
Bali et al., Phys.Rept. 343 (2001) 1-136

- Great success in conventional hadrons

Quark model for charmonia

BGS model

- Mass
- EM transitions
- Open flavor strong decays



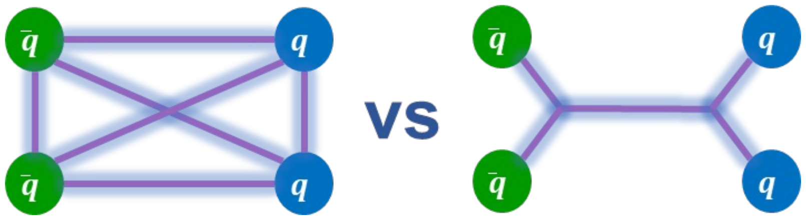
- NRQM with only 4 paras.
- Work well below open flavor thresholds
- Work better for bottomonium
- Inconsistent above the thresholds, multiquark involved

See also the Q Zhao and B-S Zou's talks

T. Barnes, S. Godfrey, and E. S. Swanson,, Phys. Rev. D 72, 054026 (2005).

Towards QMs for multiquark states

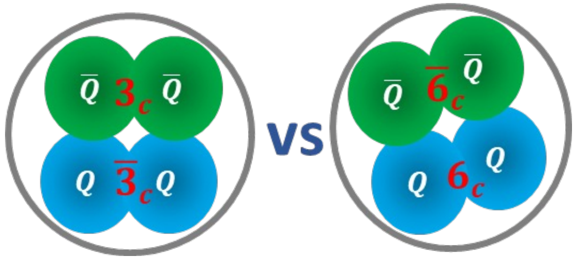
- Extend QMs to multiquark states
  - ▶ Confinement interactions? pairwise or flux-tube?
  - ▶ Lattice QCD support flux-tube picture
  - ▶ The exp. tetraquark results ⇒ confinement pictures

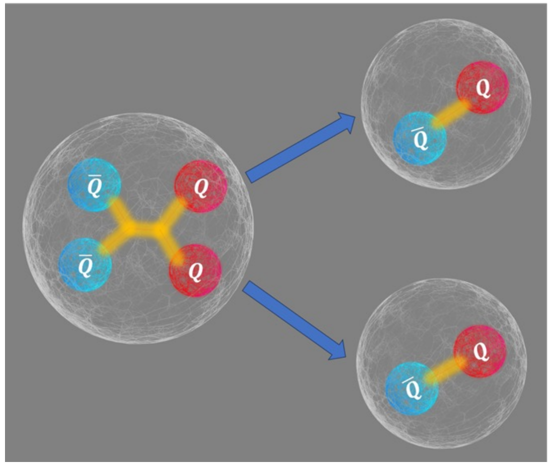


Alexandrou et al., PRD71 (2005) 014504; Okiharu et al., PRD72 (2005) 014505; Bicudo et al., PRD96 (2017), 074508

Computational Challenges

- Color structures: e.g. $3-\bar{3}$ and $6-\bar{6}$ tetraquark
- Matrix elements of double Y-type potential
- Four/five body problem
- Resonance above the di-hadron thresholds





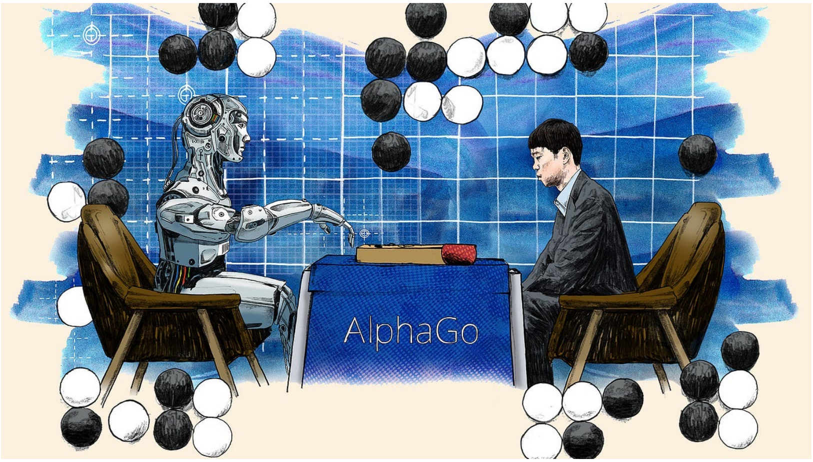
AI4Science

Background: Deep neural networks

Turning point of AI

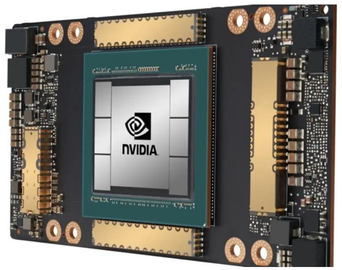
東南大學
SOUTHEAST UNIVERSITY

Year	Milestone	Who	Significance
2010	ImageNet	Fei-Fei Li	Large-scale dataset that enabled the deep learning revolution.
2012	AlexNet	Krizhevsky et al.	Proven power of deep CNNs and GPU acceleration for AI.
2014	GANs	Ian Goodfellow	Introduced framework for generative AI to create realistic data.
2015/16	TensorFlow/ Pytorch	Google/Meta	Critical open-source software that democratized AI research globally.
2016	AlphaGo	DeepMind	First AI to defeat a world champion in the game of Go.
2010s	AI Chips	NVIDIA, Google, Apple	Specialized GPUs/TPUs enabled training of massive AI models.
2017	Transformer	Google	Revolutionary architecture that made modern large language models possible.
2018	BERT	Google	Bidirectional language model that set new standards for NLP.
2020	AlphaFold 2	DeepMind	Solved the decades-old scientific challenge of protein folding.
2022	ChatGPT	OpenAI	Popularized large language models through accessible public interaction.

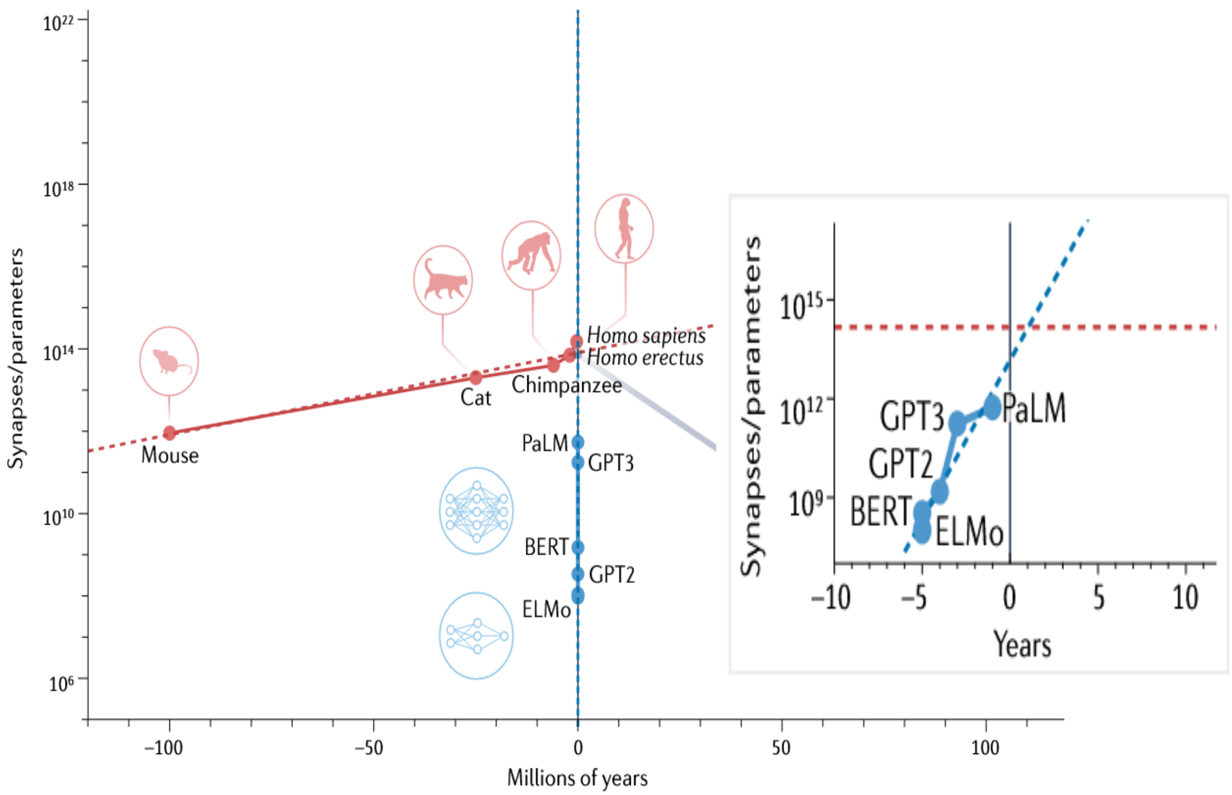


10 milestones of AI selected by DeepSeek

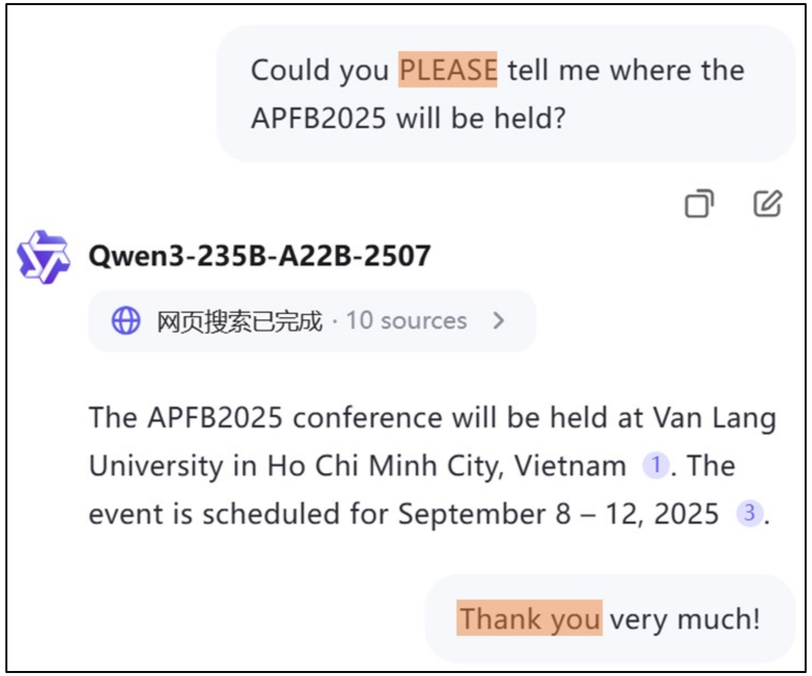
deepseek



Biological intelligence vs AI



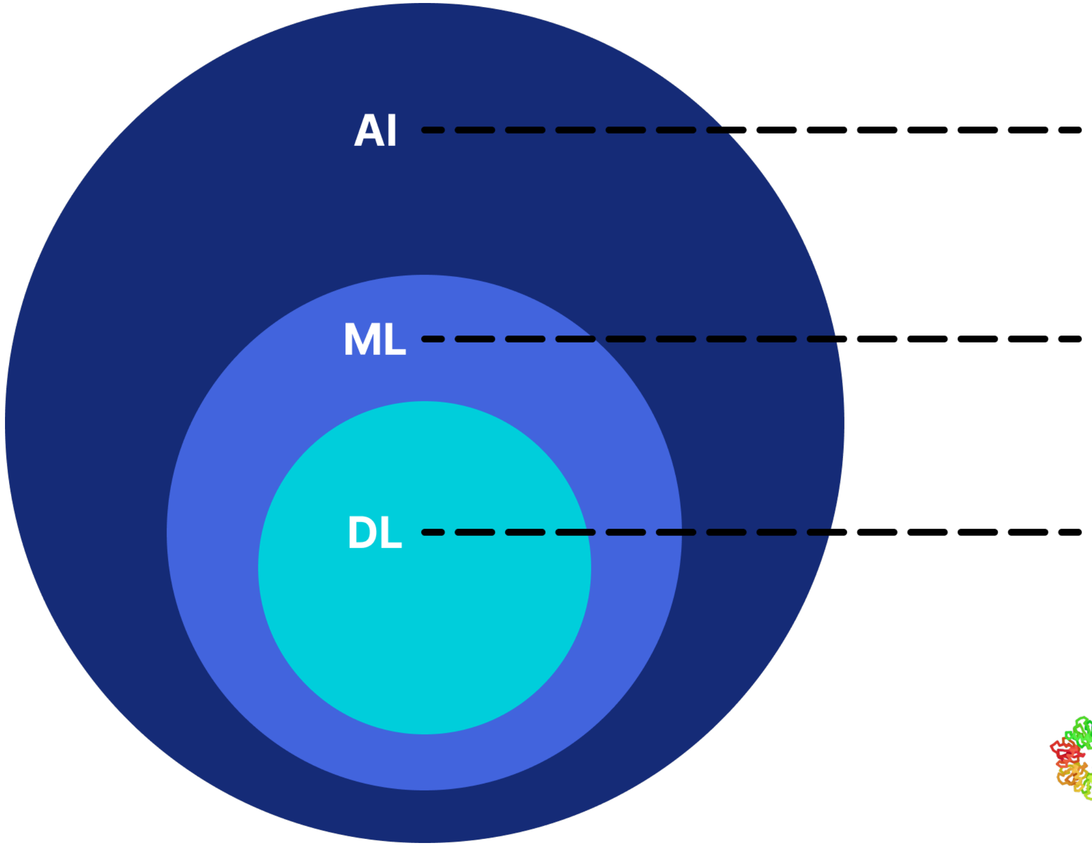
The evolution of biological and artificial intelligence



Being more polite to chatbots may increase your chances of survival when robots win

Schwartz, M. D. (2022). *Nature Reviews Physics*, *4*(12), 741–742.

AI vs ML vs DL



Artificial Intelligence
Engineering of making intelligent machines and programs.

Machine Learning
Ability to learn without being explicitly programmed.

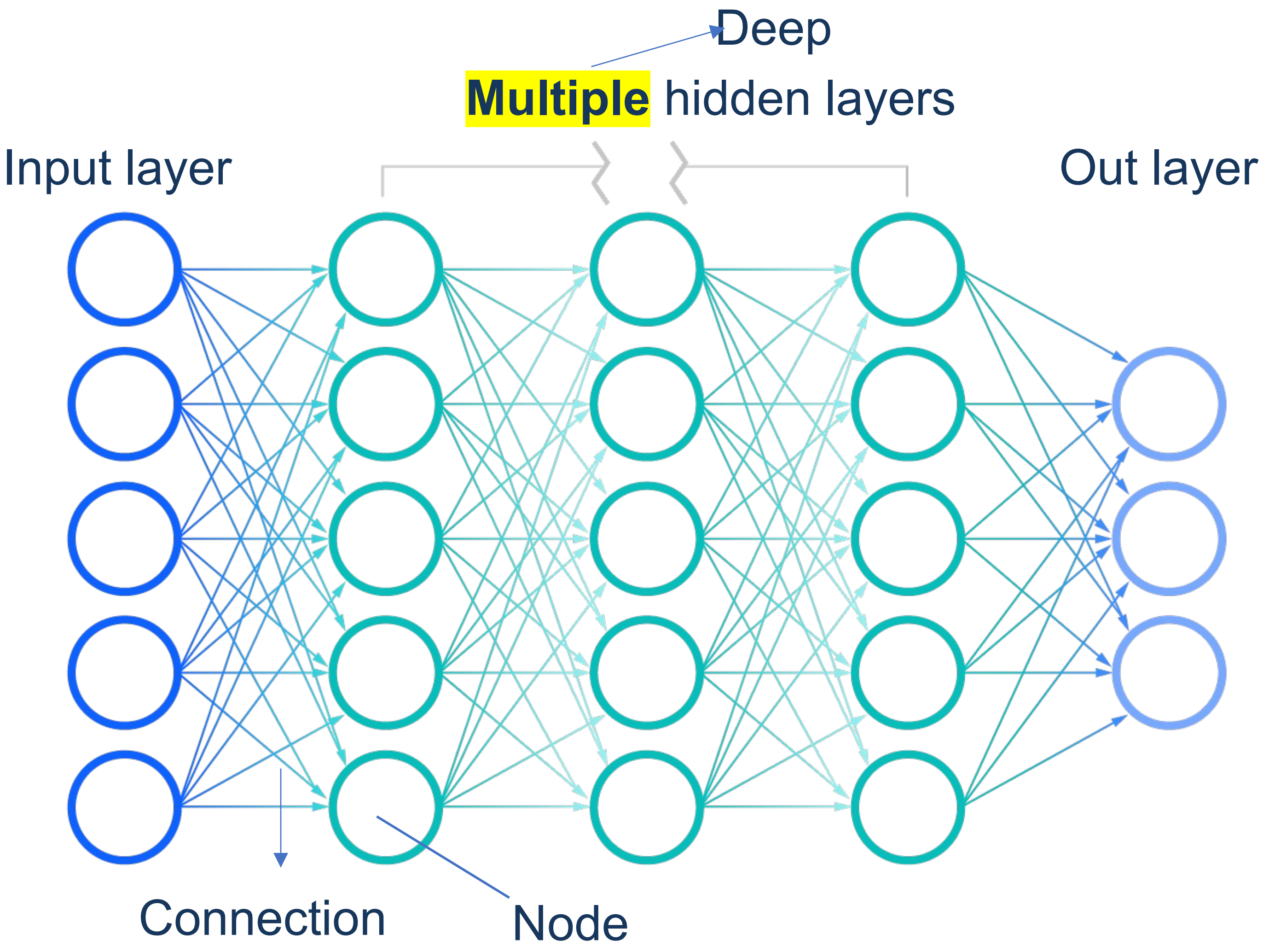
Deep Learning
Learning based on deep neural networks.

neurosnap.ai

https://neurosnap.ai/blog/post/64279cadfeb3e5ca5ba0904a

AI boom from 2010s was powered by **deep learning**

Deep neutral network



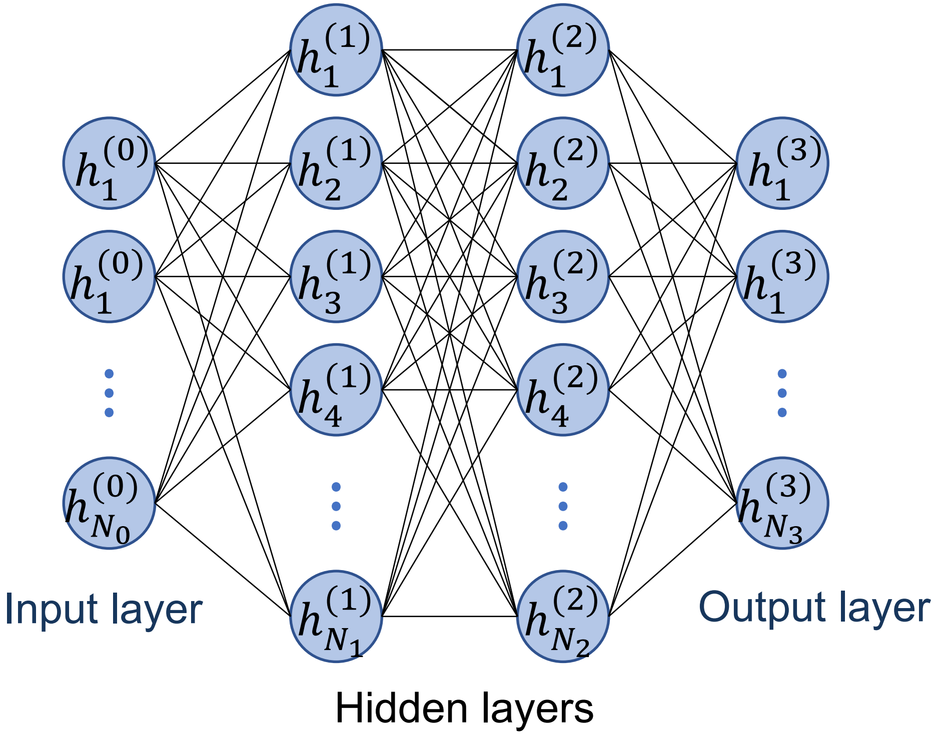
Neutral network: simple units (neurons) connected to form complex network

Deep：Enable learn complex relation in data

Connections between layers

$$\vec{h}^{(i+1)} = \sigma\left[W^{(i)}\vec{h}^{(i)} + \vec{b}^{(i)}\right]$$

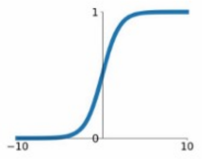
Nonlinear Linear



- Linear part: Weight matrices $W^{(i)}$ and biases $\vec{b}^{(i)}$
- Nonlinear part: Activation functions σ

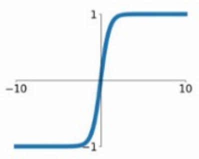
Sigmoid

$\sigma(x) = \frac{1}{1+e^{-x}}$



tanh

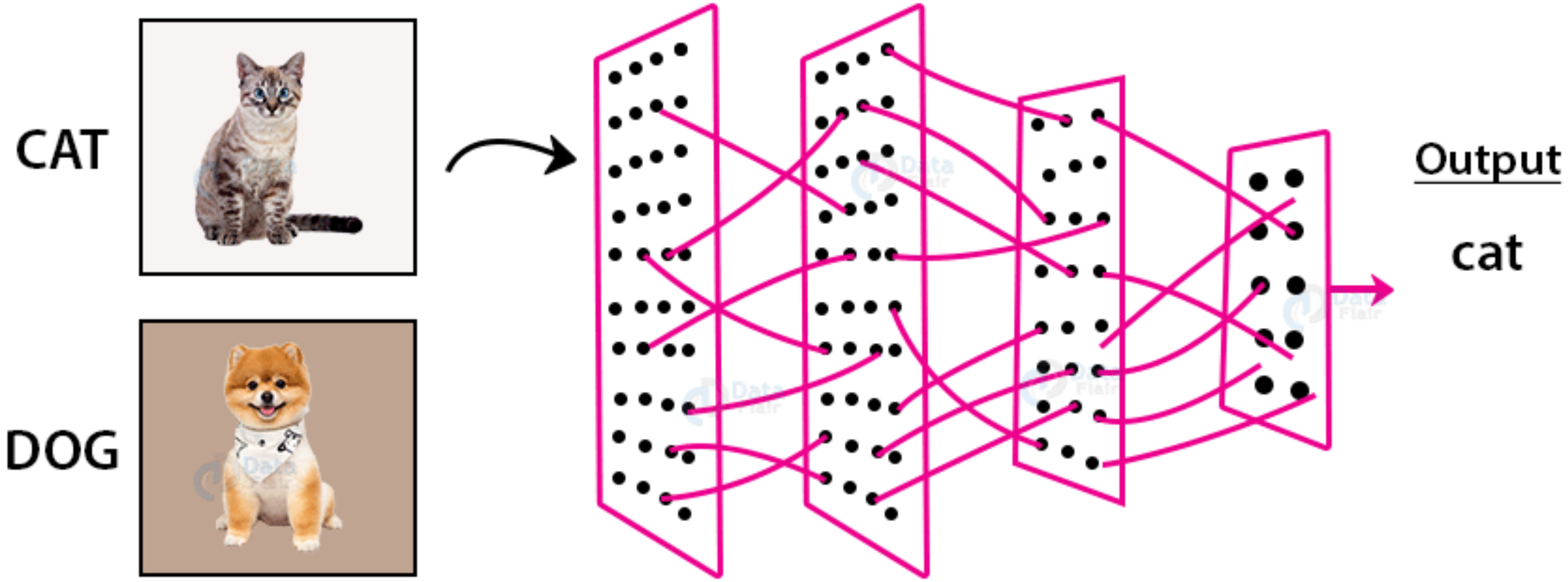
$\tanh(x)$



…

- Training: determine $W^{(i)}$ and biases $\vec{b}^{(i)}$ to minimize loss functions
  - ▶ Loss function： quantifies the difference between the expected and actual output.
- Highly efficient optimization
  - ▶ Backpropagation: automatic differentiation (AD)
  - ▶ GPUs
  - ▶ …

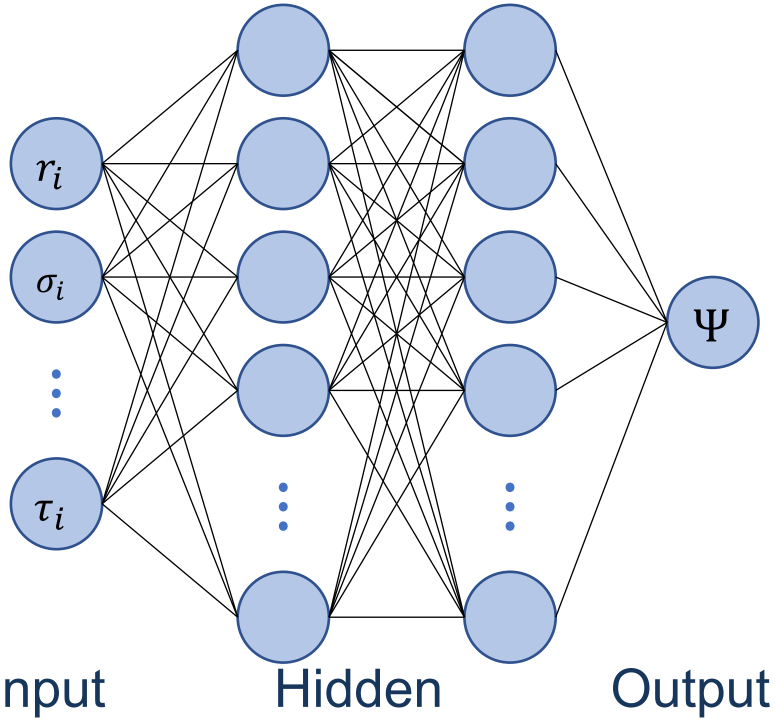
DNNs: Data-driven application



See also Julius B. Pagayon's talk about $T_{cc}(3875)$

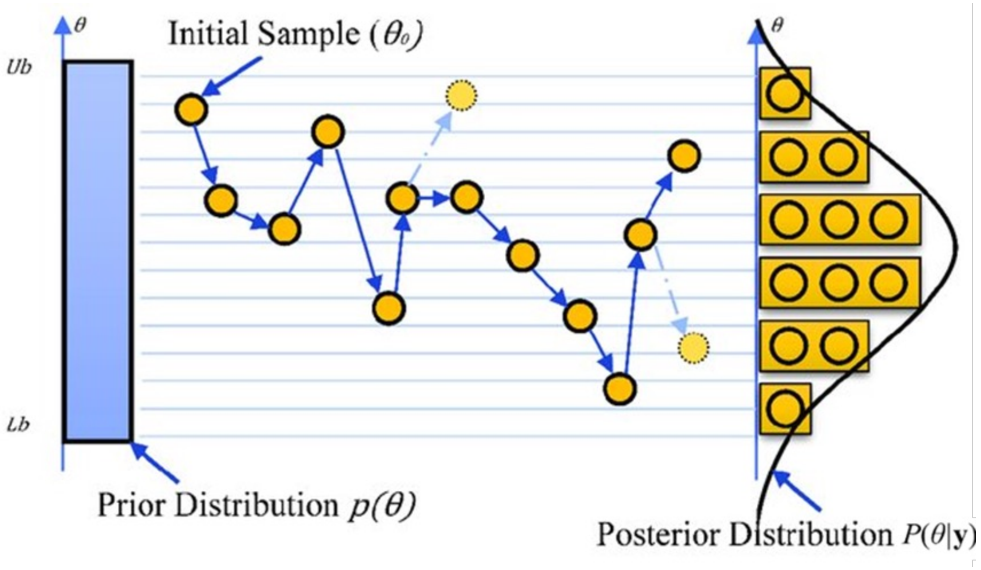
DNN-based VMC method

Neural Network Quantum States (NQSs):



Loss function: $E_{\boldsymbol{\theta}} = \dfrac{\langle \psi_{\boldsymbol{\theta}} | H | \psi_{\boldsymbol{\theta}} \rangle}{\langle \psi_{\boldsymbol{\theta}} | \psi_{\boldsymbol{\theta}} \rangle} \geq E_0$

Variational Monte Carlo (VMC)



Sample using Metropolis-Hasting algorithm

$$E_{\boldsymbol{\theta}} = \frac{\int |\psi_{\boldsymbol{\theta}}(r)|^2 \frac{H\psi(r)}{\psi(r)} dr}{\int |\psi_{\boldsymbol{\theta}}(r)|^2 dr}$$

Train the DNNs not by the data but by the physics equation: **physics-informed**

Advantages

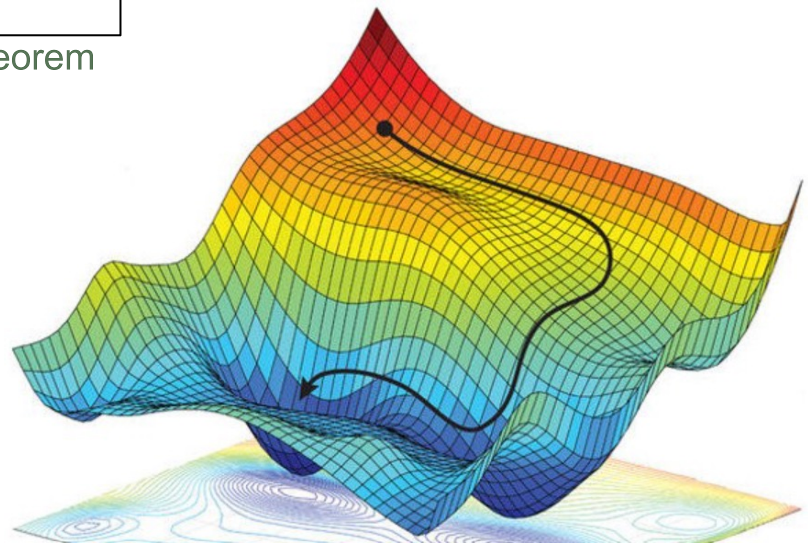
- DNNs: approximate high-dimensional functions
  - ► **Universal approximation theorems**

sufficiently large or deep

Neural networks with a certain structure can, in principle, approximate ***any continuous function*** *to* ***any desired accuracy***

https://en.wikipedia.org/wiki/Universal_approximation_theorem

  - ► Variational principle: the more general the trial function, the more accurate the solution
  - ► Unbiasedly distinguish between molecular and compact tetraquark states

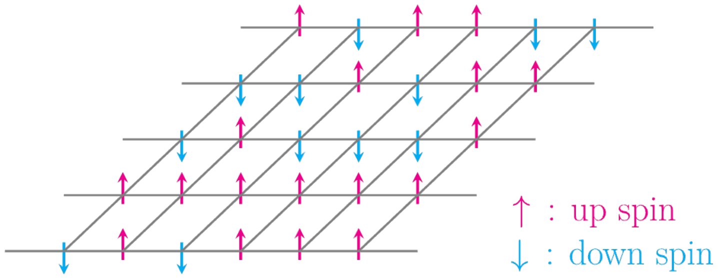


- VMC: no additional computational cost to few-body potential
  - ► Flux-tube confinement potential
- Circumvents the sign problem in imaginary time evolution of DMC
- Fast optimization: AD+GPUs
- Easy-to-use open-source libraries

TensorFlow PyTorch

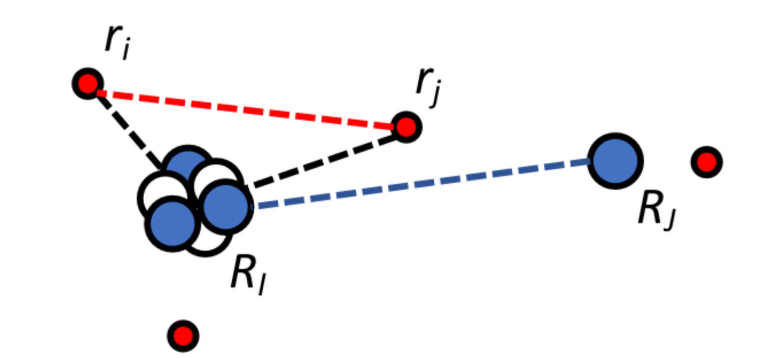
DNN in many-body problems

Spin model



G. Carleo and M. Troyer, Science 355, (2017).

Electronic systems



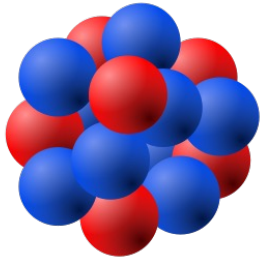
PauliNet: J. Hermann, Z. Schätzle, and F. Noé, Nat. Chem. 12, 891 (2020).

FermiNet: D. Pfau et al., Phys. Rev. Res. 2, 033429 (2020).

…

Review: J. Hermann et al., Nat. Rev. Chem.7, 692–709 (2023)

Nuclear structure



Deuteron: J. W. T. Keeble and A. Rios, Phys. Lett. B 809, 135743 (2020).

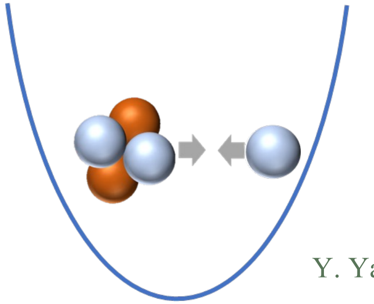
$\boldsymbol{A \leq 4}$**:** C. Adams *et al.*, Phys. Rev. Lett. **127**, 022502 (2021).

FeynmanNet: Y. Yang and P. Zhao, Phys. Rev. C **107**, 034320 (2023).

Hypernuclei: Z.-X. Zhang *et al.*, arXiv:2508.03575

…

Neutron-α scattering



Chiral Nuclear force
Trapped five-body problems
(DNNs+VMC)

Y. Yang, E. Epelbaum, J. Meng, **LM**, and P. Zhao, arXiv:2502.09961

see E. Epelbaum's talk

DeepQuark

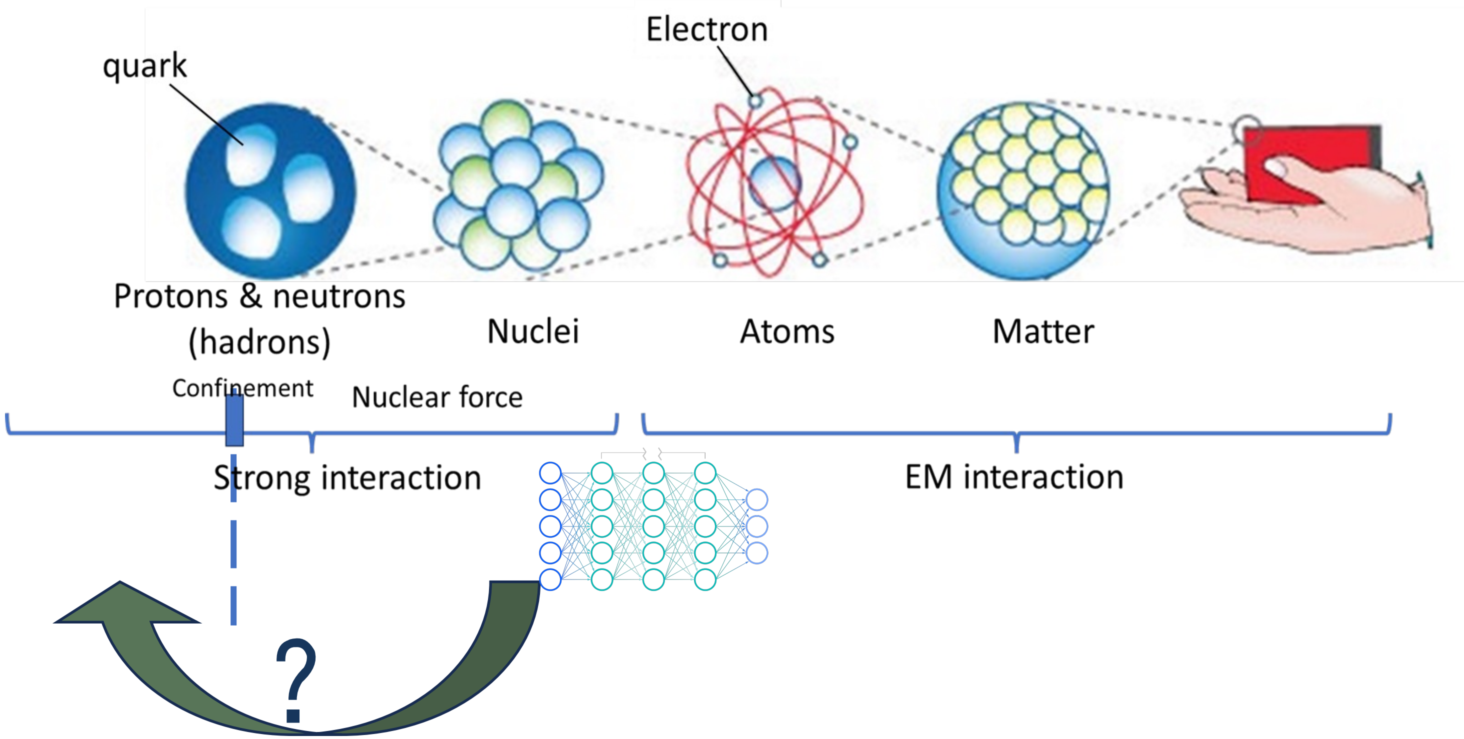
Quark potential models

$$V_{ij}(r) = V_{ij}^{OGE} + V_{ij}^{conf.}$$
$$= \left[\frac{\alpha_s}{r} - \frac{8\pi\alpha_s}{3m_Q^2}\tilde{\delta}(\sigma;r)\boldsymbol{S}_i\cdot\boldsymbol{S}_j - \frac{3}{4}br\right]\frac{\lambda_i\cdot\lambda_j}{4}$$

Quark potential model

QCD-inspired model

- quarks as the degree of freedom
- SU(3) color symmetry
- Confinement
- ….



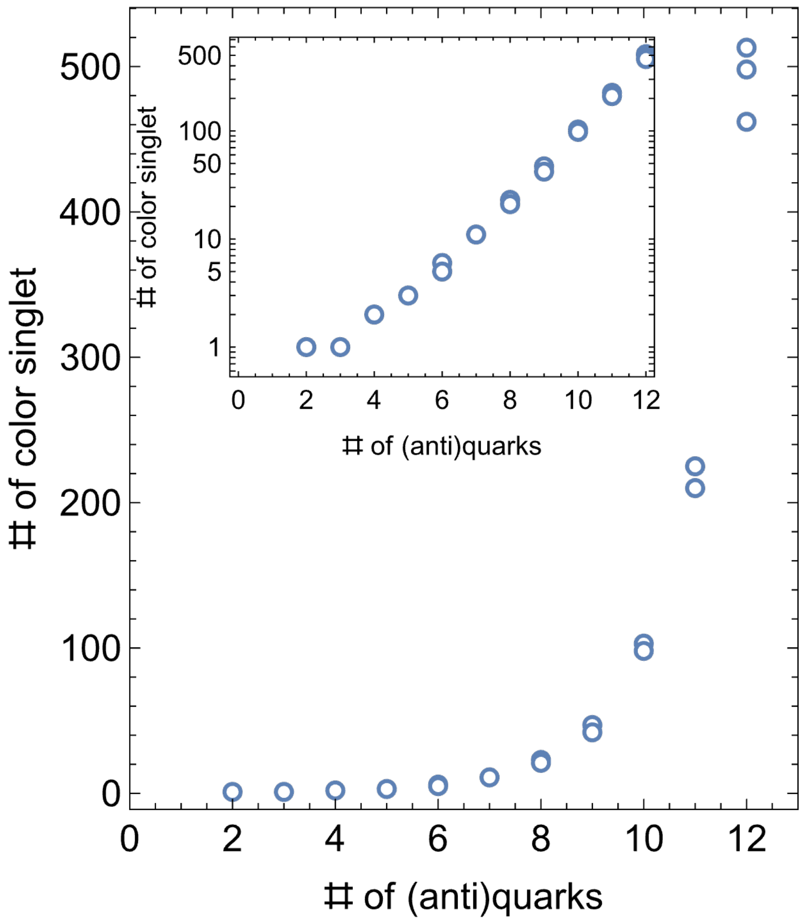
Challenge 1: Spin and color projection

- Conventional spin projection
  - ▶ Option 1： no projection, calculate the ground state
  - ▶ Option 2： penalty terms

$$f_{loss}(\theta) = \langle E \rangle_\theta + \langle S^2 \rangle_\theta$$

However, for multiquark systems

- State with higher spin could more interesting
  - ▶ $[cc\bar{q}\,\bar{q}]_{S=0}$: DD threshold; $[cc\bar{q}\,\bar{q}]_{S=1}$: T_{cc} state
- Color projection is needed
  - ▶ color singlet
  - ▶ SU(3) symmetry

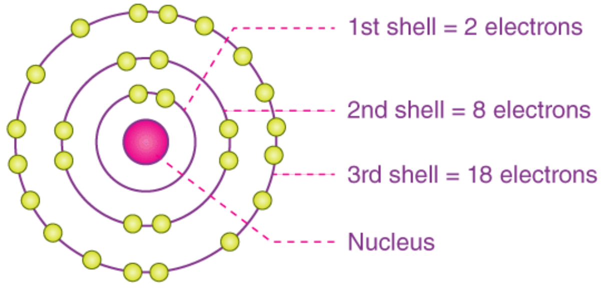


Challenge 2: Strong correlation

- Shell model: good guidance or initial point

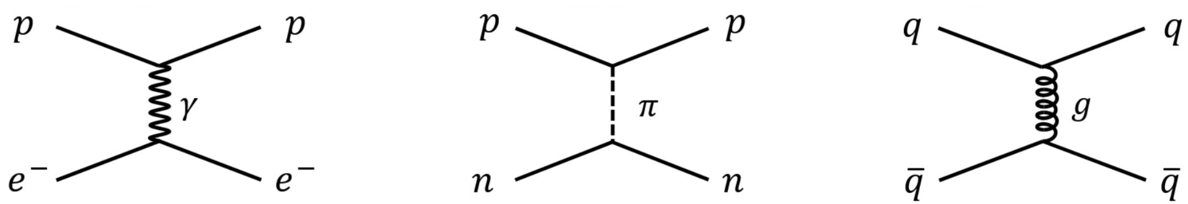
$$\Psi(\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_N)=\frac{1}{\sqrt{N!}}\begin{vmatrix}\chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_1)\\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \cdots & \chi_N(\mathbf{x}_2)\\ \vdots & \vdots & \ddots & \vdots\\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \cdots & \chi_N(\mathbf{x}_N)\end{vmatrix}$$

For electron and nucleon systems



Electron Shell

- Much stronger interaction among quarks
  - ▶ No evidence of a multiquark shell structure
  - ▶ Few-body correlation could be important

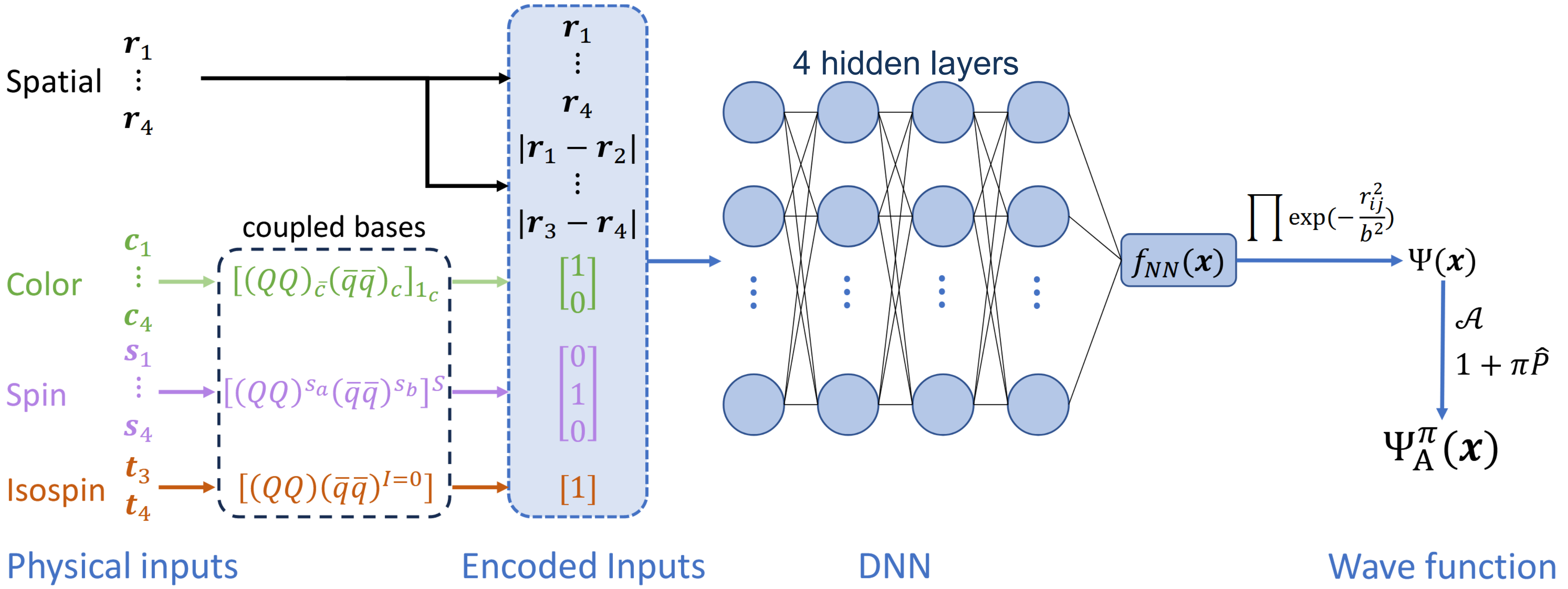


$$e^2=0.1 \qquad m_e\approx 0.5\ \mathrm{MeV}$$

$$\frac{m_\pi^2 g_A^2}{16\pi f_\pi^2}\sim 0.1 \qquad m_N=938\ \mathrm{MeV}$$

$$\frac{4}{3}\alpha_s(m_c)\sim 0.4 \qquad m_c\approx 1273\ \mathrm{MeV}$$

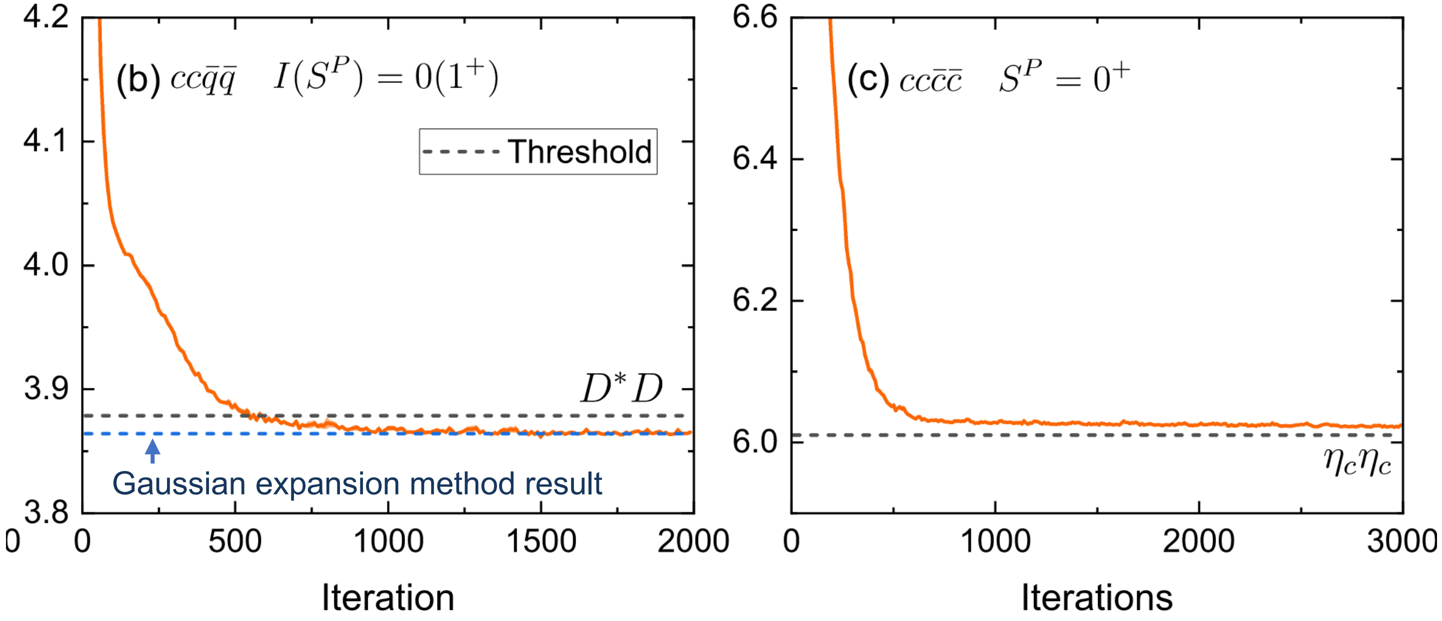
DeepQuark



Powered by NetKet
https://github.com/netket

Tetraquark states: pairwise interaction

- Convergence after 1000 iterations
- Statistical errors less than 0.1 MeV
- Competitive performance with Gaussian expansion method
- $cc\bar{q}\bar{q}$: molecular state
- $bb\bar{q}\bar{q}$: compact heavy diquark
- $QQ\bar{Q}\bar{Q}$: no bound states



	$I(S^P)$	Thresh.	ΔE	$P_{\bar{3}_c\otimes 3_c}$	$P_{6\otimes 6_c}$	r_{QQ}	$r_{\bar{q}\bar{q}}$	$r_{Q\bar{q}}$
$cc\bar{q}\bar{q}$	$\mathbf{0(1^+)}$	DD^*	-15	55%	45%	1.24	1.41	1.06
$bb\bar{q}\bar{q}$	$\mathbf{0(1^+)}$	$\bar{B}\bar{B}^*$	-153	97%	3%	0.33	0.78	0.69
$QQ\bar{Q}\bar{Q}$	$\mathbf{0(0^+)}$ $\mathbf{0(1^+)}$ $\mathbf{0(2^+)}$	$\eta_c\eta_c$ η_cJ/ψ $J/\psi J/\psi$	No bound					

ΔE in MeV, r in fm

Pentaquark states: pairwise interaction

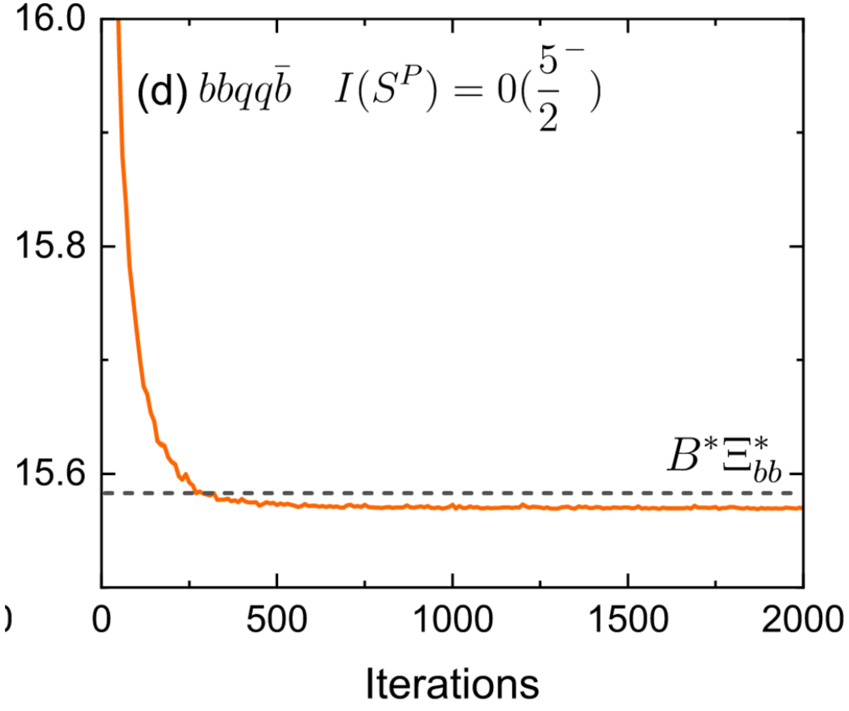
- Exact pentaquark calculations in are computationally prohibitive
  - ▶ Approximations in the spatial or color configurations
- Possible bound pentaquark systems
  - ▶ Heavy-diquark-antiquark-symmetry: $(QQ)_{\bar{3}_c} \to \bar{Q}$
  - ▶ $\bar{Q}\bar{Q}qq \to QQ\bar{Q}qq$
  - ▶ Given $M_{\bar{Q}\bar{Q}qq} < M_{\bar{Q}q} + M_{\bar{Q}q}$, $M_{QQ\bar{Q}qq} < M_{QQq} + M_{\bar{Q}q}$?
- Exact calculation via DeepQuark

$$\chi_{\bar{3}_c\otimes\bar{3}_c} = \left\{\left[(QQ)_{\bar{3}_c}(qq)_{\bar{3}_c}\right]_{3_c}\bar{Q}\right\}_{1_c},$$

$$\chi_{\bar{3}_c\otimes 6_c} = \left\{\left[(QQ)_{\bar{3}_c}(qq)_{6_c}\right]_{3_c}\bar{Q}\right\}_{1_c},$$

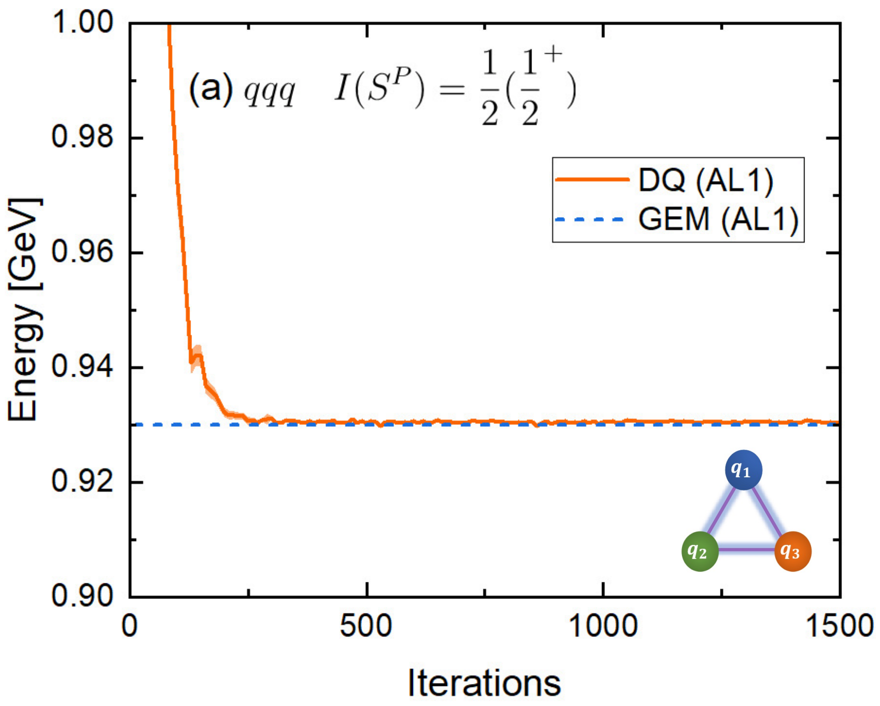
$$\chi_{6_c\otimes\bar{3}_c} = \left\{\left[(QQ)_{6_c}(qq)_{\bar{3}_c}\right]_{3_c}\bar{Q}\right\}_{1_c}.$$

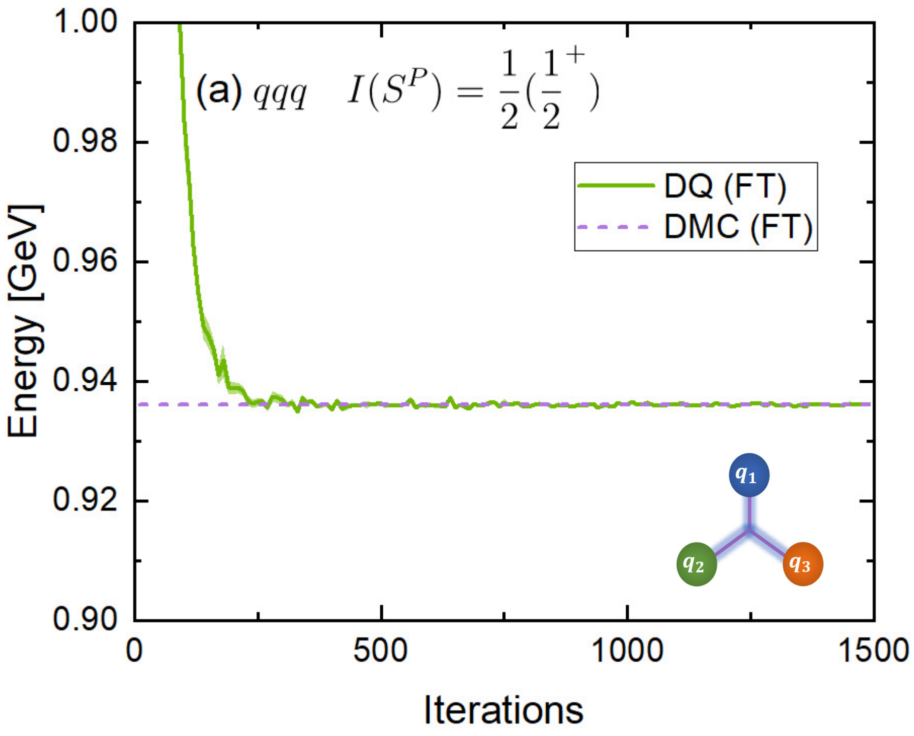
  - ▶ Moderate increase in comput. cost relative to the tetraquark



	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

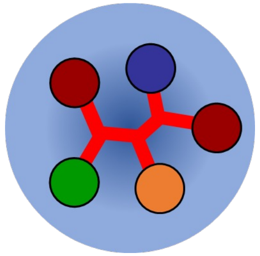
- Lattice QCD supports the flux-tube potential for qqq baryons and tetraquarks
- Baryons: Y-type interaction, junction position to minimize the total string length
  - ► Small impact on baryon spectrum, but complicate dramatically matrix element calculations
- DeepQuark: incorporate Y-type interaction without additional cost, thanks to the VMC
- Ready for tetraquark systems, where different confinement patterns lead to distinct results



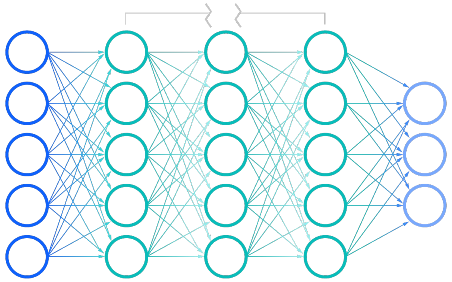


Summary and outlook

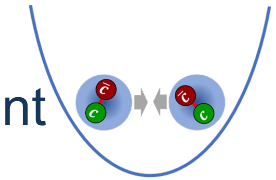
東南大學
SOUTHEAST UNIVERSITY

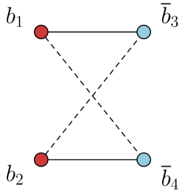


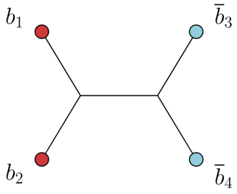
- Multiquark states
  - ▶ Challenge: few-body resonant problem
  - ▶ Opportunity: tetraquark exp. advance ⇒ confinement pictures



- DNN+VMC: high efficiency
  - ▶ General wave functions, accurate
  - ▶ Unbiased wave function ansatz： clustering behavior
  - ▶ Few-body interaction, with no extra cost
  - ▶ AD+GPUs+ML library
- DeepQuark
  - ▶ Color and spin projections
  - ▶ Strong correlation
  - ▶ Surpass traditional methods starting from pentaquark states
  - ▶ Ready for flux-tube confinements
- Outlook:
  - ▶ DNNs: excited states, scattering, resonances
  - ▶ All-charm tetraquarks with flux-tube confinement





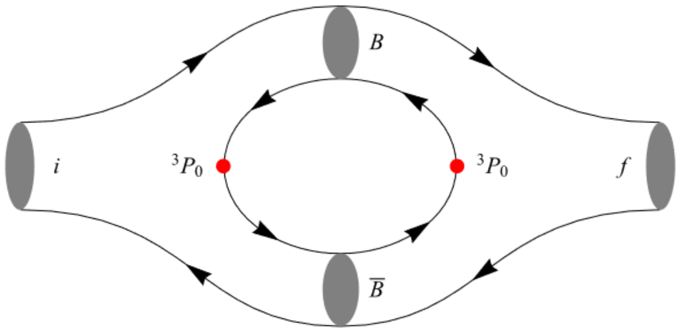


Thanks for your attention!

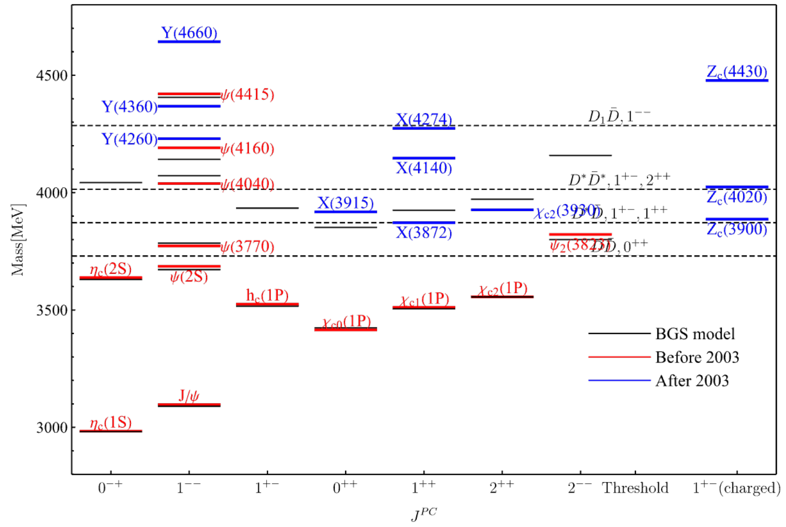
BACKUP

Quark models in new era

- Heavy-quarkonium-like states: unquenched quark model



B-S Zou's talks
https://indico.cern.ch/event/1457095/contributions/6563231/



- **Multiquark systems**

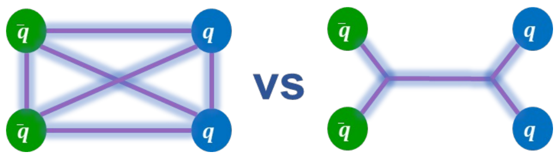
$[c\bar{c}qq]$ P_c	$[cc\bar{c}\bar{c}]$ $X(6900)$ $X(6600)$ $X(7100)$	$[cs\bar{u}\bar{d}]$ $T_{cs1}(2900)$ $T_{cs0}(2900)$	$[cs\bar{u}\bar{d}]$ $Z_{cs}(3985)$ $Z_{cs}(4000)$	$[cc\bar{u}\bar{d}]$ $T_{cc}(3875)^+$	$[c\bar{s}u\bar{d}][c\bar{s}\bar{u}d]$ $T_{c\bar{s}0}(2900)^{++}$ $T_{c\bar{s}0}(2900)^0$	$[c\bar{c}qqq]$ $P_{cs}(4338)$ $P_{cs}(4459)$
	2006.16957 2306.07164 2304.08962 2506.07944	2009.00025 2009.00026	2011.07855 2103.01803	2109.01038 2109.01056	2212.02716 2212.02717 2411.19781	2210.10346 2012.10380 2502.09951

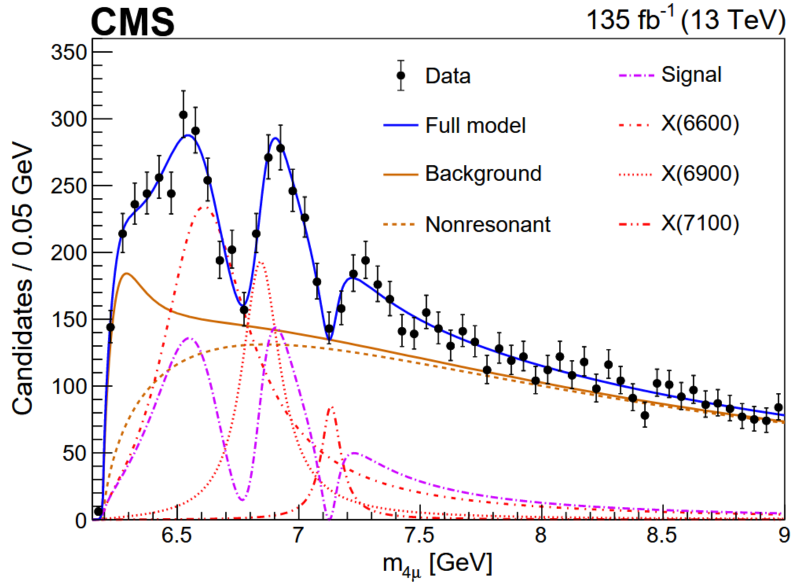
Particle Zoo 2.0

Chances and difficulties

- **Opportunities:** all-charm tetraquark **family**
  - ▶ Great experimental advances: LHCb, CMS, ATLAS
  - ▶ Simple systems
    - **Constituent quark is almost the current quark**
    - **Small relativistic effect**
    - **Unlikely exchange light mesons between (anti)quarks**
  - ▶ Different confinements would leave imprints on the mass spectrum

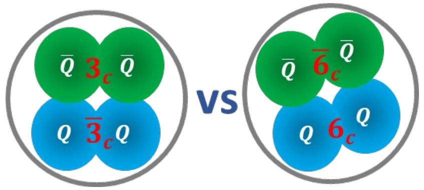
    Alexandrou:2004ak, Okiharu:2004ve, Bicudo:2017usw



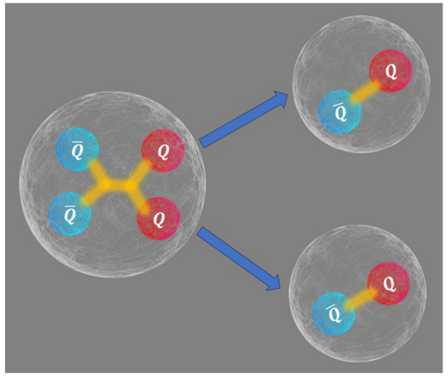


- **Challenges**
  - ▶ Color structures: e.g. $3-\bar{3}$ and $6-\bar{6}$ tetraquark
  - ▶ Matrix element of double Y-type potential
  - ▶ Four/five body problem
  - ▶ Resonance above the di-hadron thresholds
  - ▶ ….

Interactions?
few-body (resonance) problem?

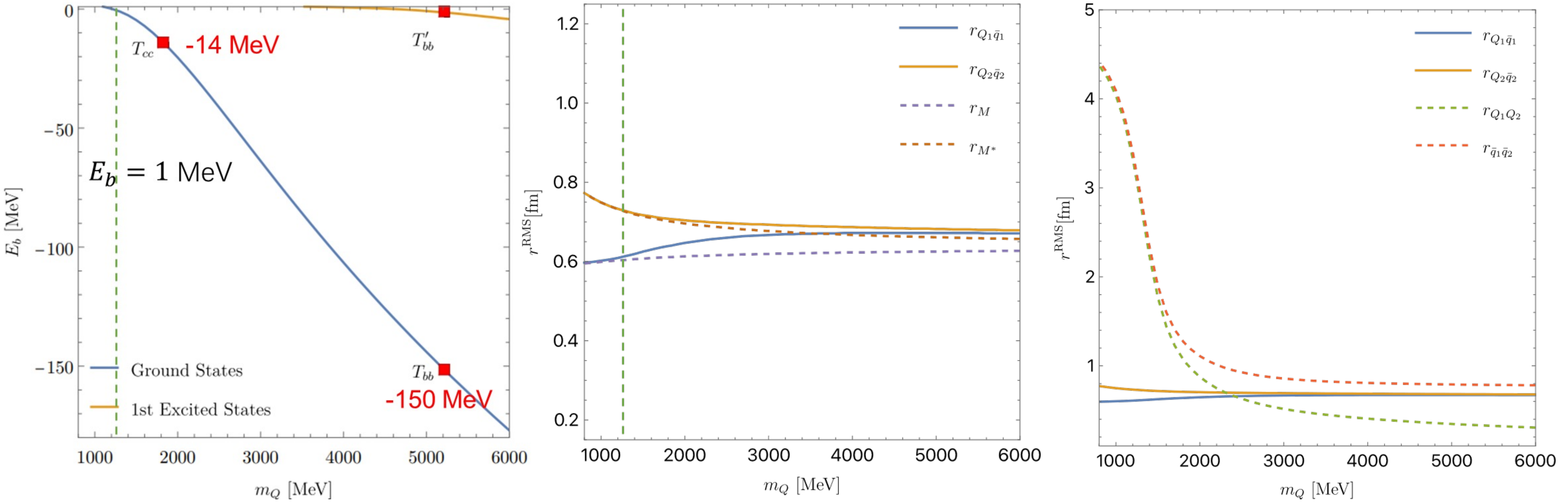


Color structures



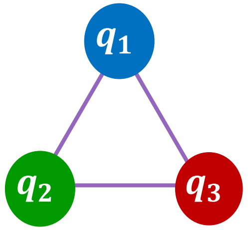
quark rearrangement

Molecular or compact ?

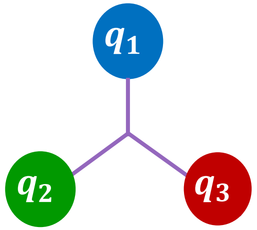


- Tuning the m_Q to m_b: (bb) compact diquark
- Tuning the m_Q to make $E_b < 1$ MeV: molecular states

Baryons



$$V^{\Delta}_{\text{conf}} = \sigma_{\Delta} \sum_{i<j} r_{ij}$$

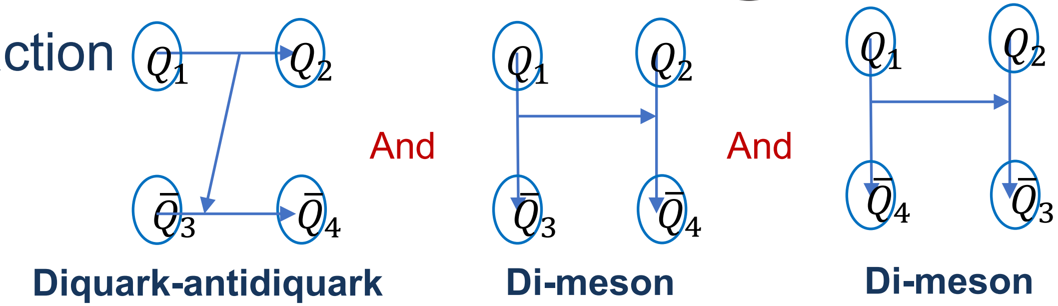


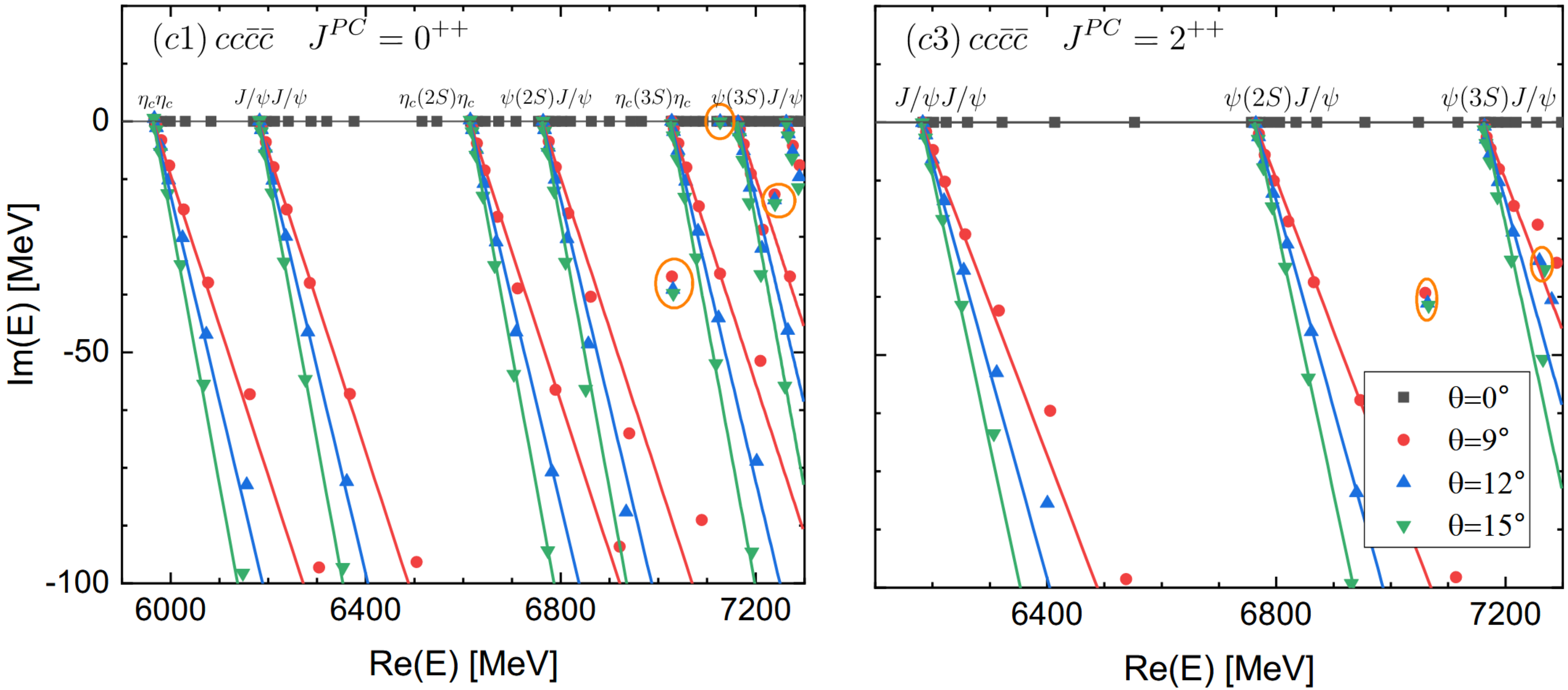
$$V^{Y}_{\text{conf}} = \sigma_{Y} L_{min}$$

$$0.5 \leq \frac{L_{min}}{L_{\Delta}} \leq 0.58$$

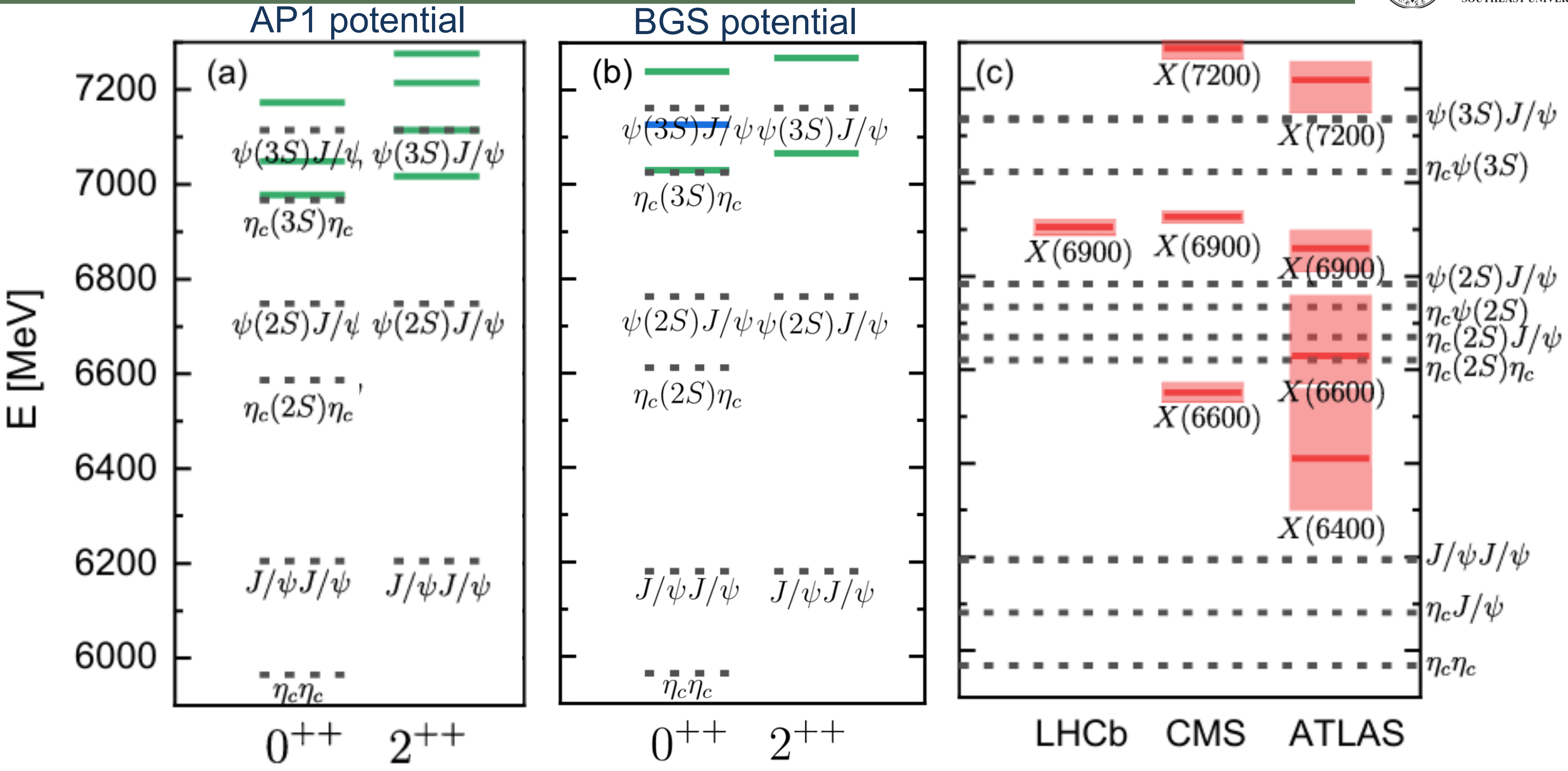
All-charm tetraquarks

- BGS, AL1, AP1 models: **pairwise** confinement interaction
- Configurations: diquark-antidiquark, di-meson×2
- Tetraquark as resonances





All-charm tetraquarks

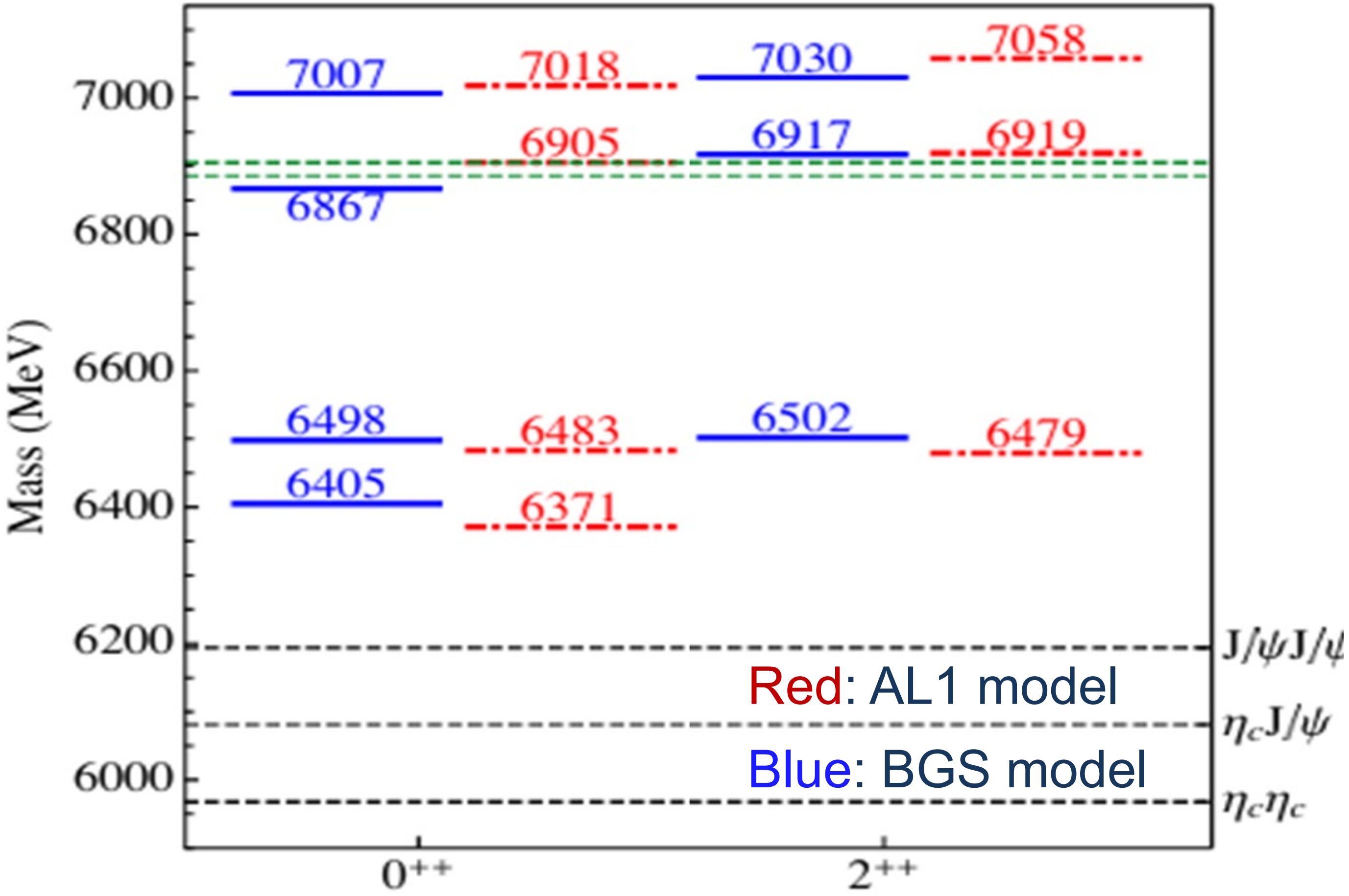


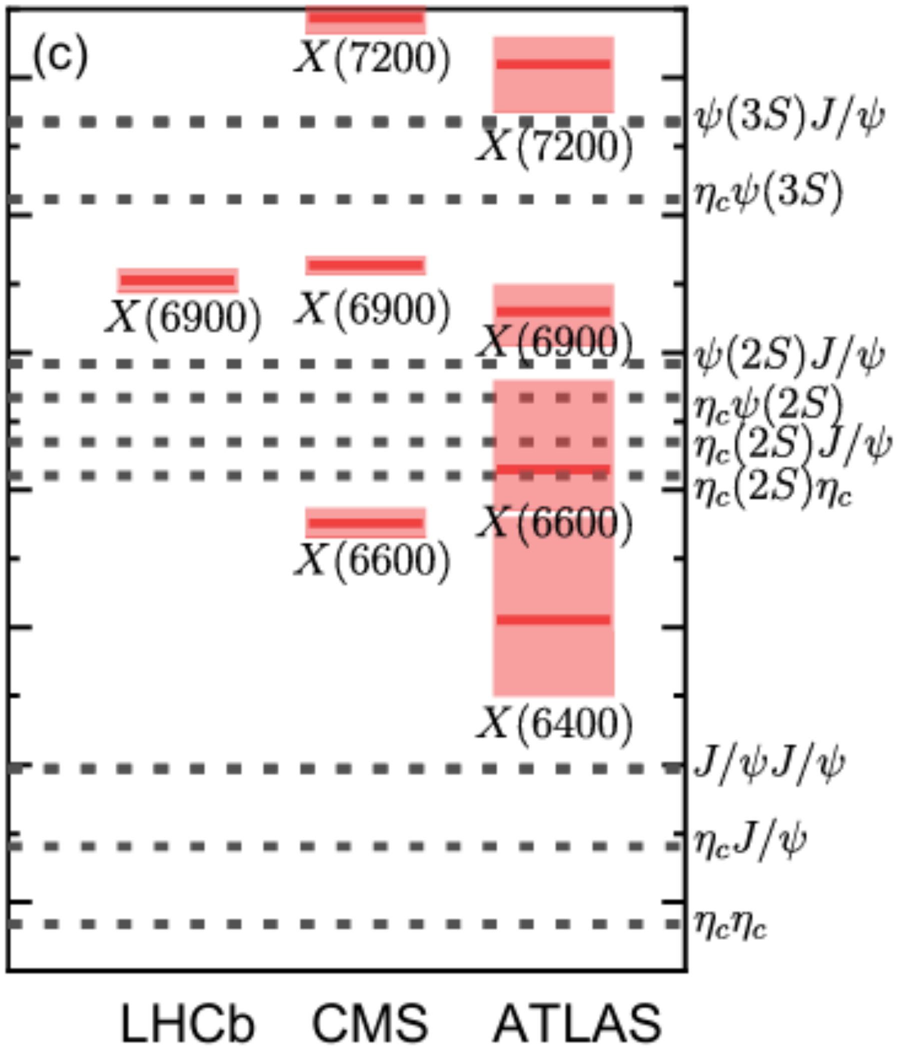
X(6600) is missing in pairwise confinement models

See also G.-J. Wang, Q. Meng, and M. Oka, PRD106, 096005 (2022).

All-charm tetraquarks

- Earlier calculation neglecting di-meson configurations



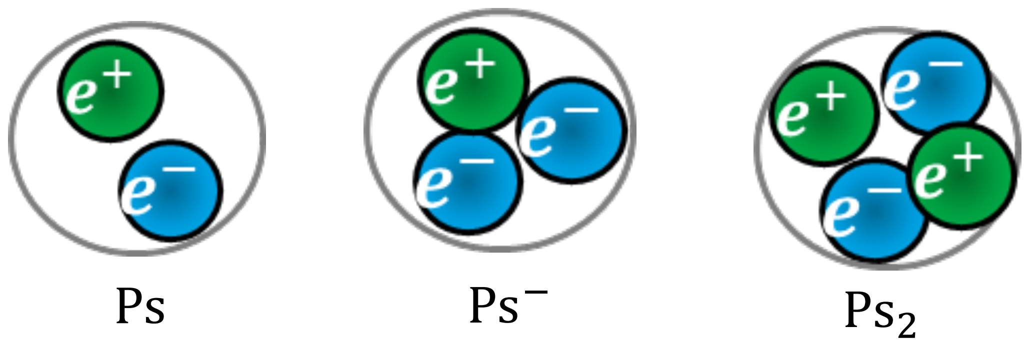


G-J Wang, LM, M Oka, S-L Zhu, PRD 104, 036016 (2021)

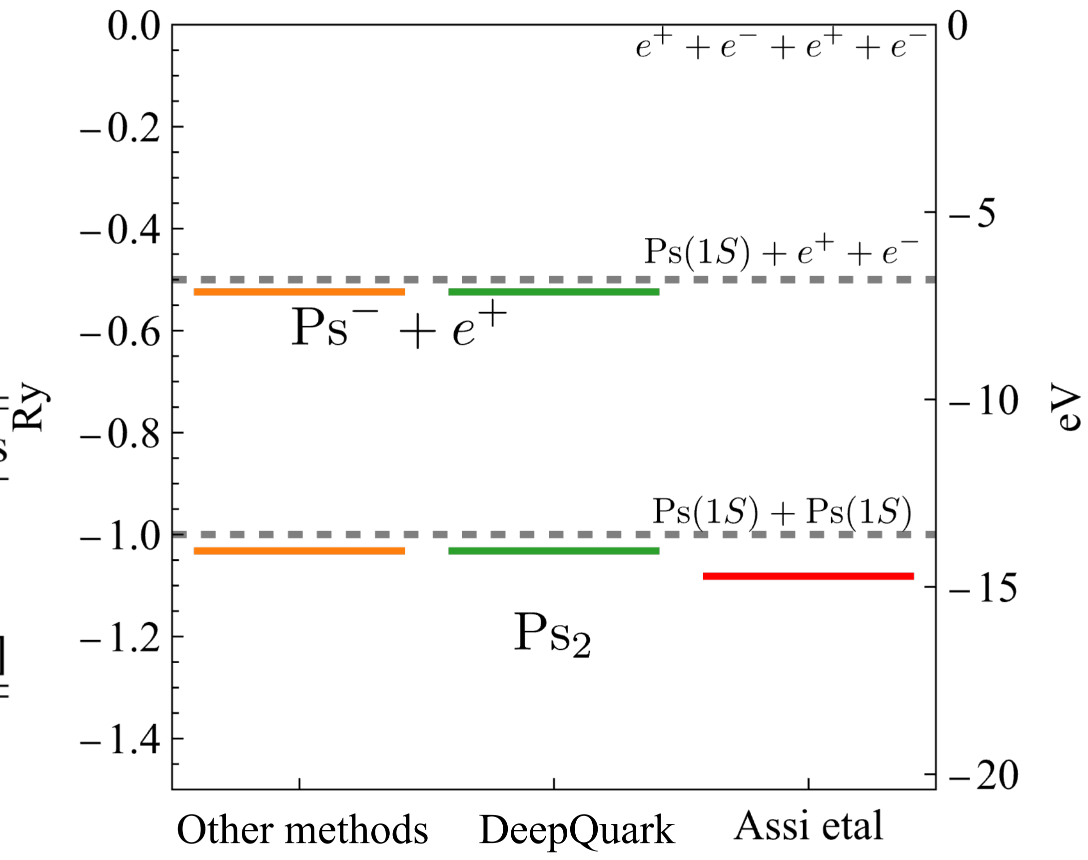
- Pairwise model: overestimate the transition between diquark-antidiquark and dimeson??

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

Electrons systems



	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]



tion, 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,

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

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