Machine learning on exotic hadrons

Qian Wang

South China Normal University

YITP long-term and Nishinomiya-Yukawa memorial workshop

Hadrons and Hadron Interactions in QCD 2024 (HHIQCD 2024)

-- Experiments, Effective theories, and Lattice --

14th Oct. - 15th Nov., 2024

Yukawa Institute for Theoretical Physics, Kyoto University, Japan

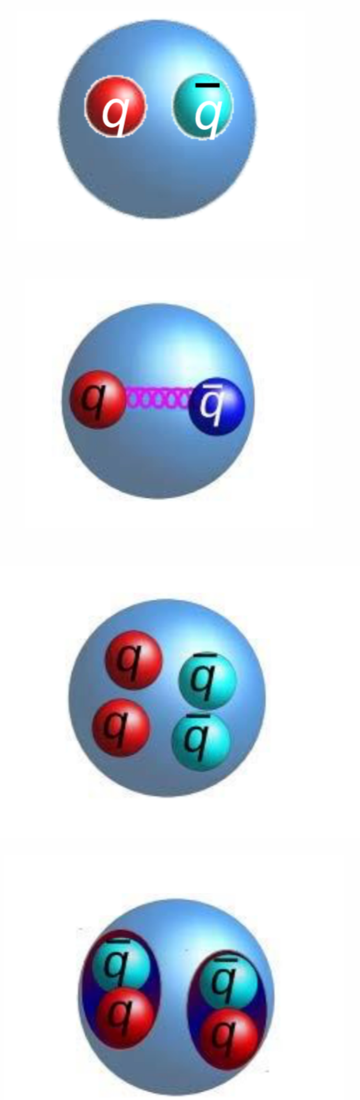
Outline

- Motivation
- One-channel case, $X(3872)$ and T_{cc}^{+}
- Multi-channel case, P_c
- Summary and outlook

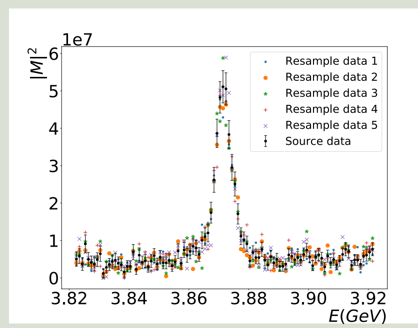
Motivation

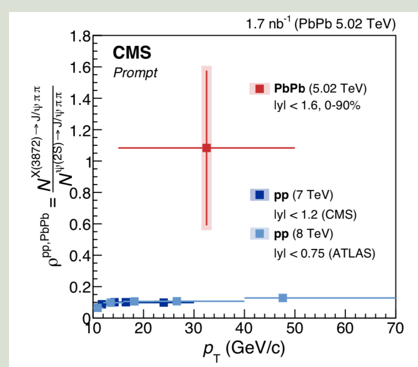
Standard approach to analyze experimental data

A model



Experimental data





Decay widths

Production rates

……

Extract parameters

Scattering length

Effective range

LECs

……

Conclusions

Yes: 100%

Pole positions

Effective couplings

Compositeness

……

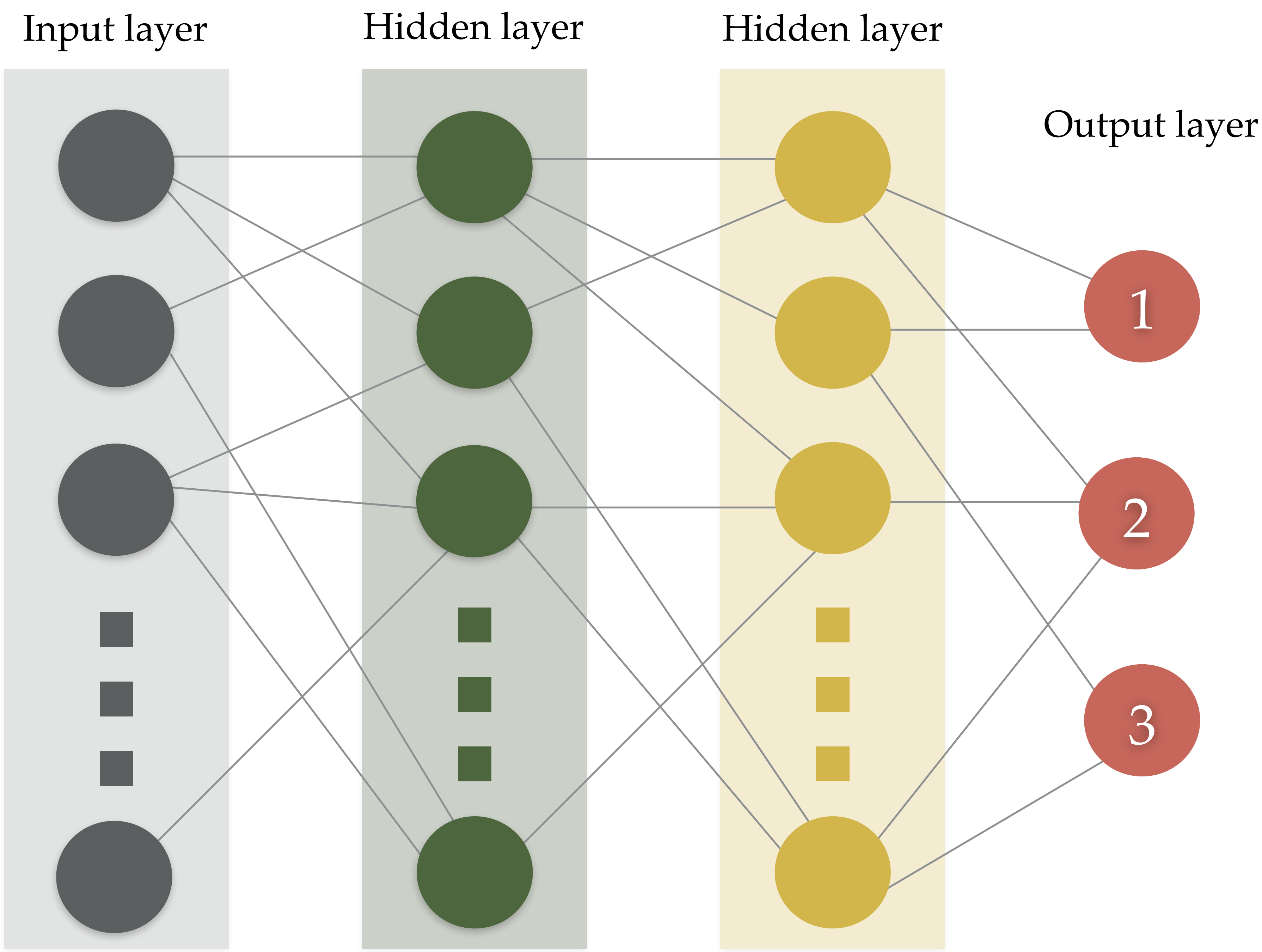
No: 0%

Top-down

JPAC, PRD105(2022)L091501

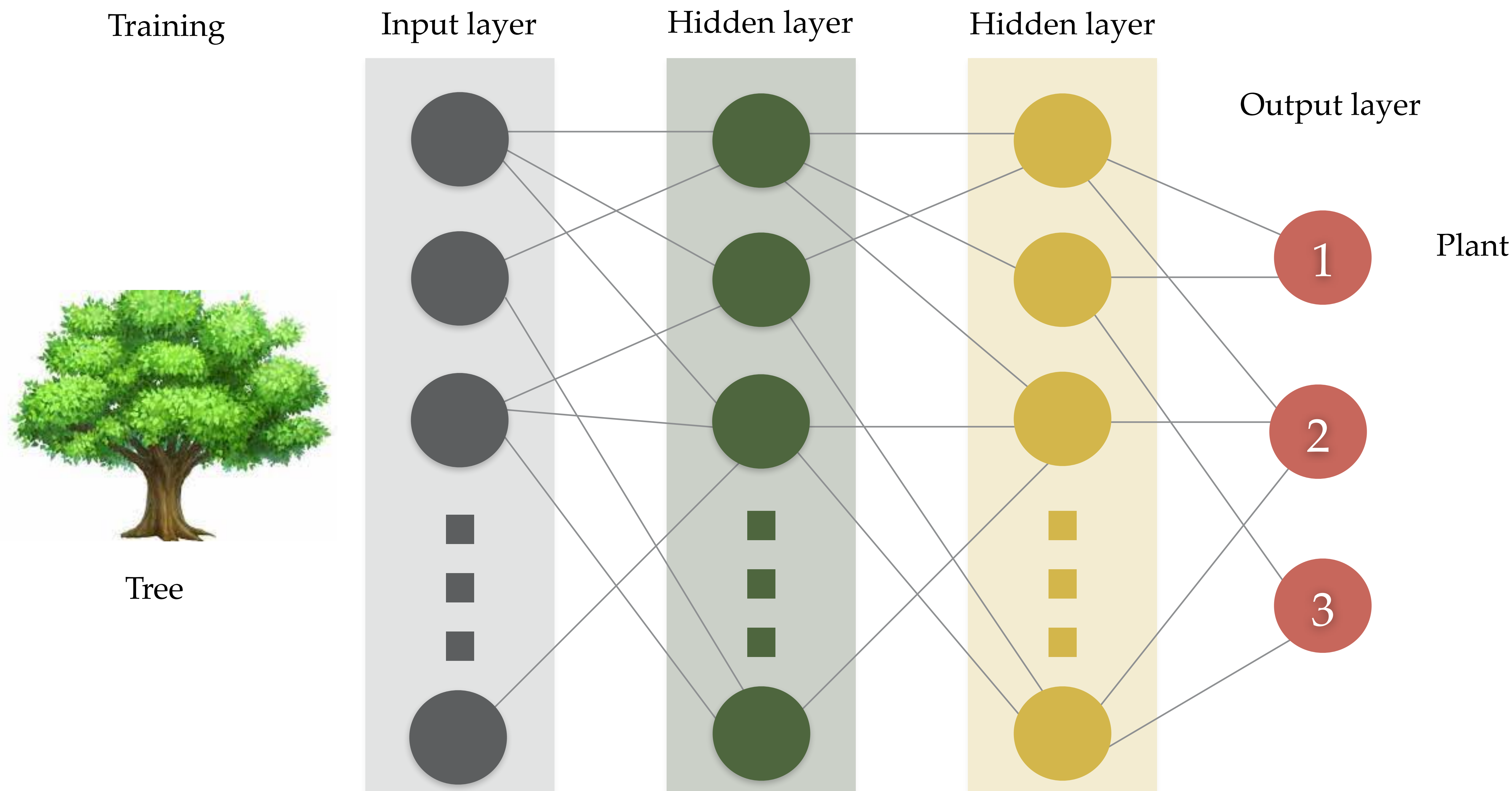
Motivation

Training



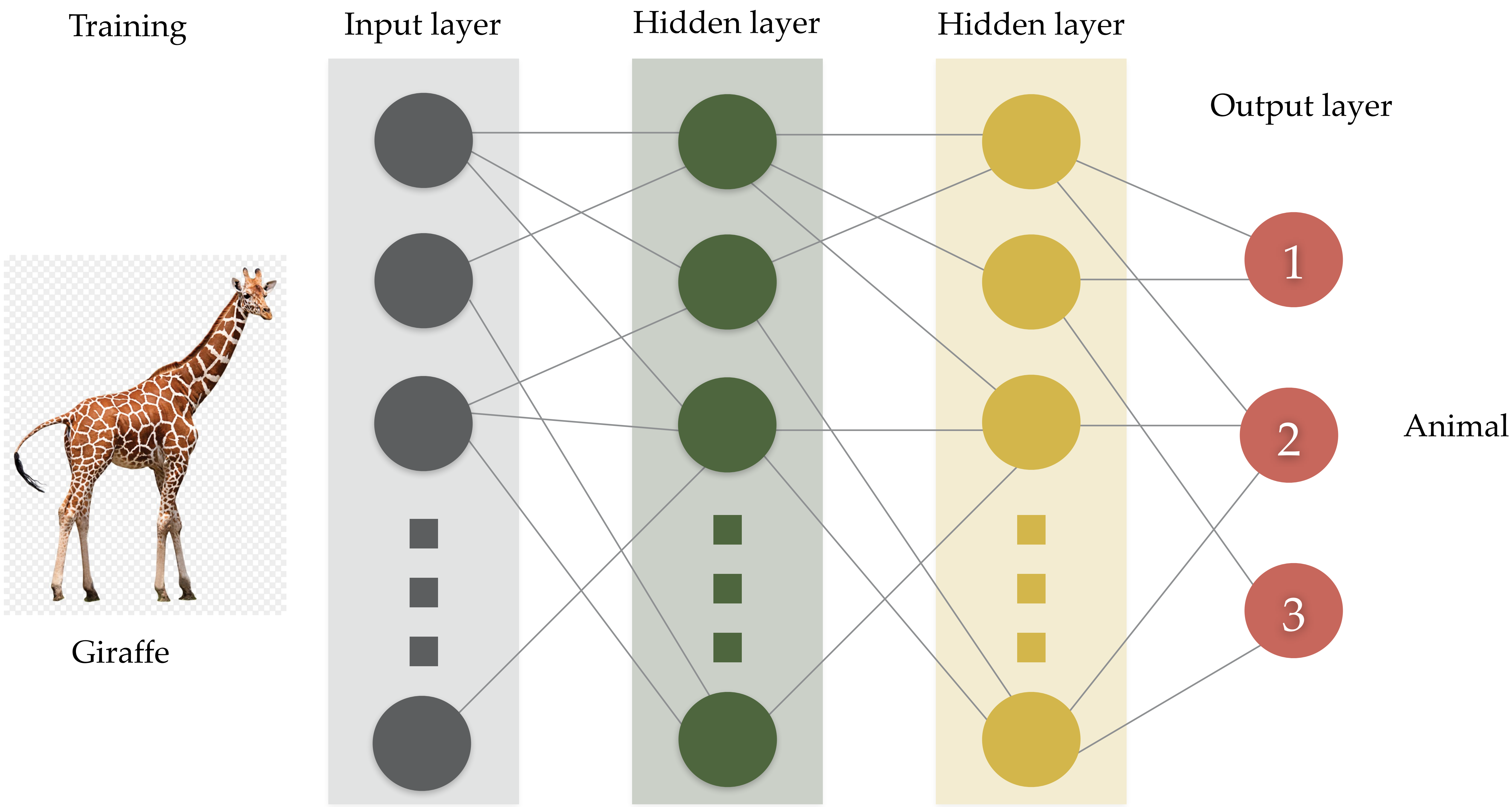
Sombillo, et al., PRD104(2021)036001, PRD102(2020)016024, PRD104(2021)036001

Motivation



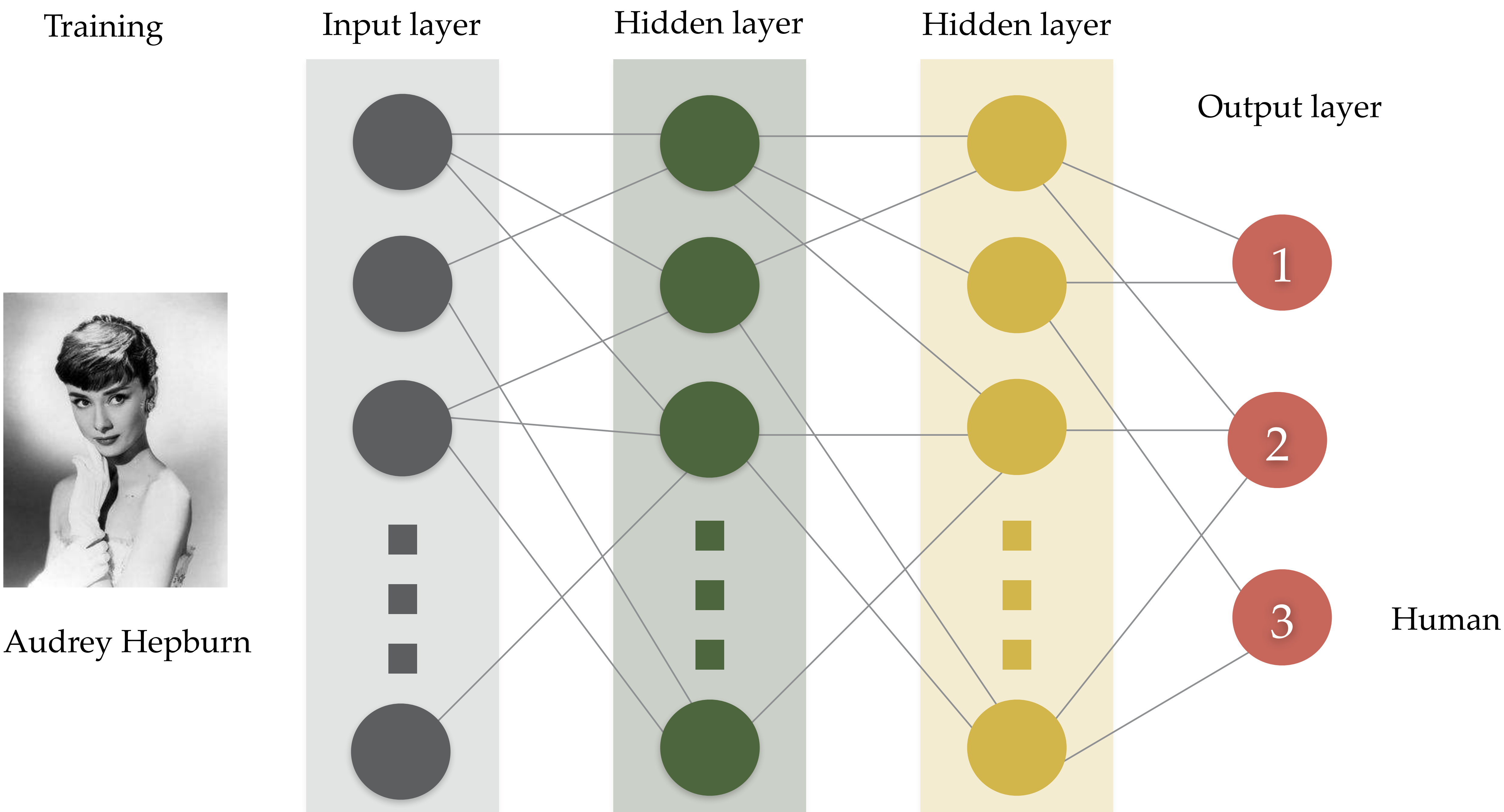
Sombillo, et al., PRD104(2021)036001, PRD102(2020)016024, PRD104(2021)036001

Motivation



Sombillo, et al., PRD104(2021)036001, PRD102(2020)016024, PRD104(2021)036001

Motivation



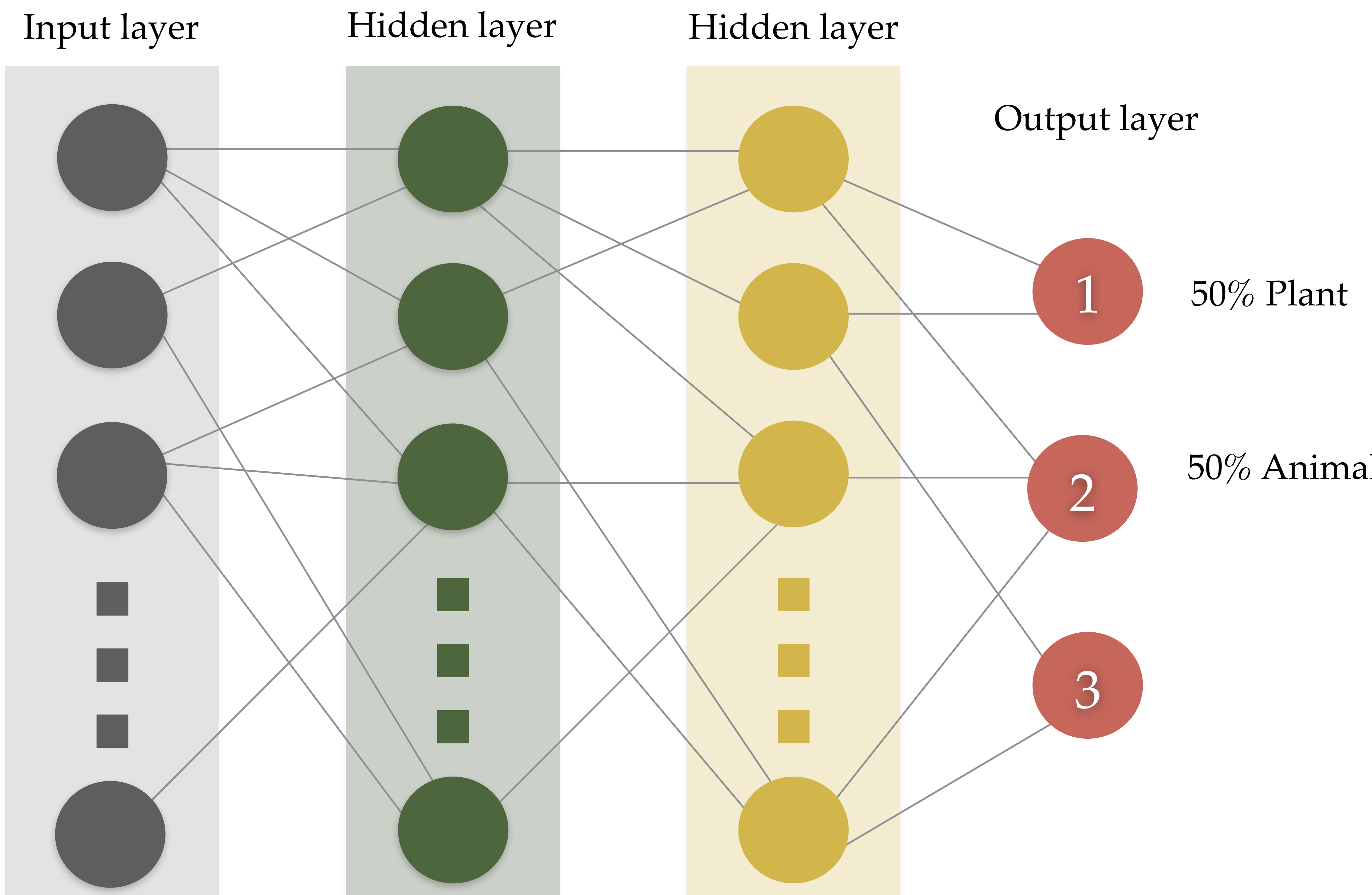
Sombillo, et al., PRD104(2021)036001, PRD102(2020)016024, PRD104(2021)036001

Motivation

Training

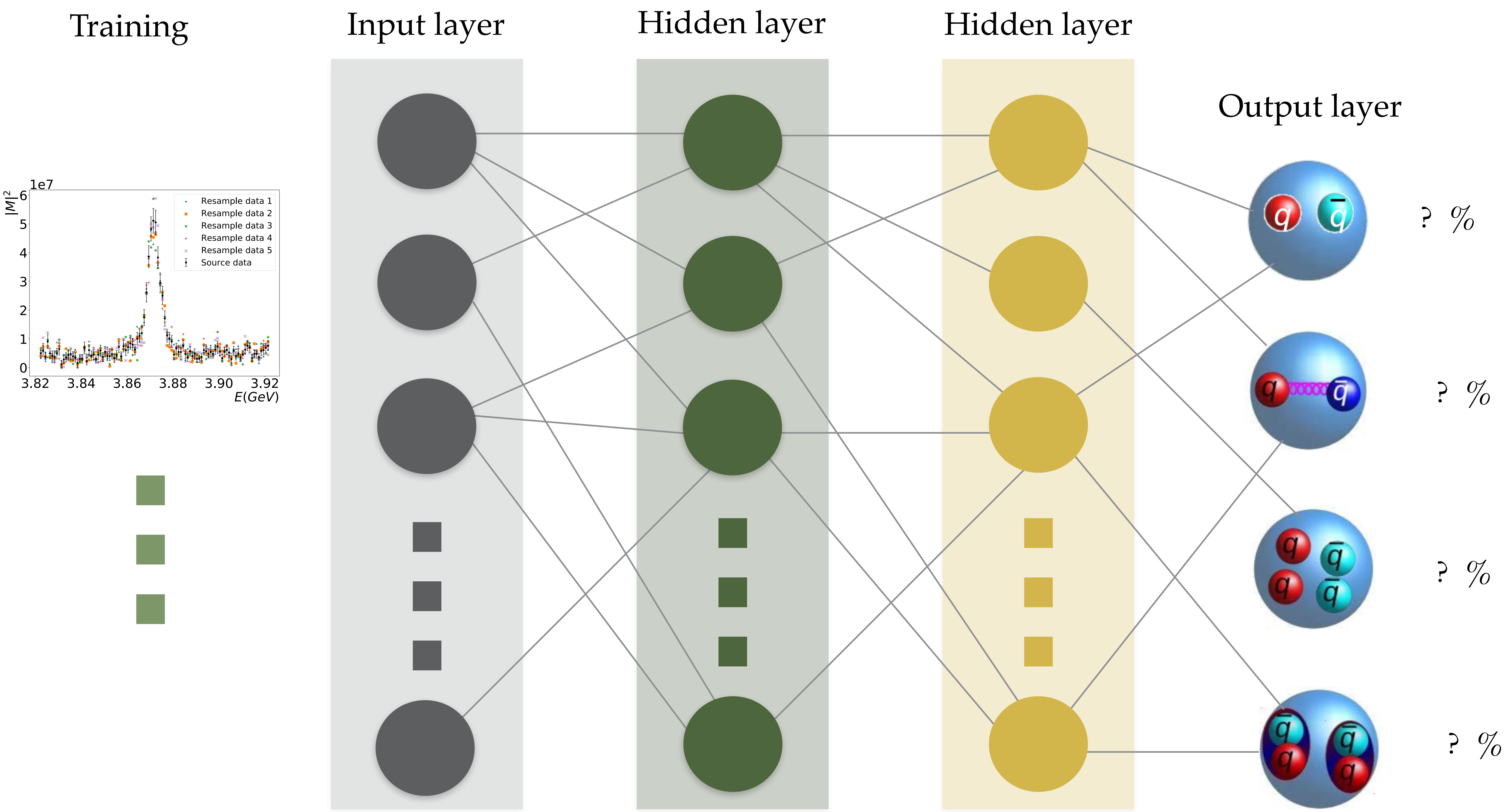
winterworm

summerherb



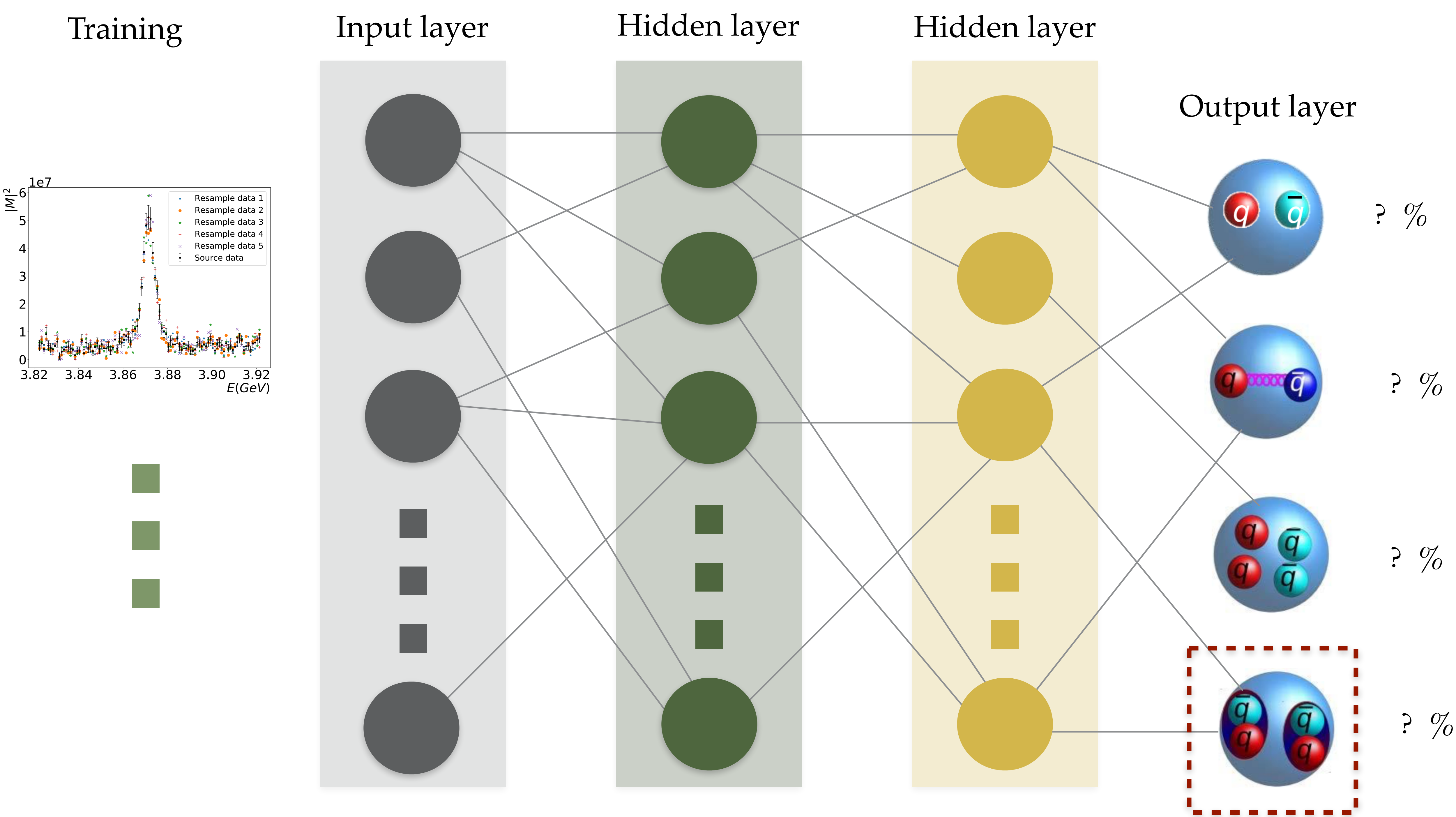
Sombillo, et al., PRD104(2021)036001, PRD102(2020)016024, PRD104(2021)036001

Motivation



Sombillo, et al., PRD104(2021)036001, PRD102(2020)016024, PRD104(2021)036001

Motivation



Sombillo, et al., PRD104(2021)036001, PRD102(2020)016024, PRD104(2021)036001

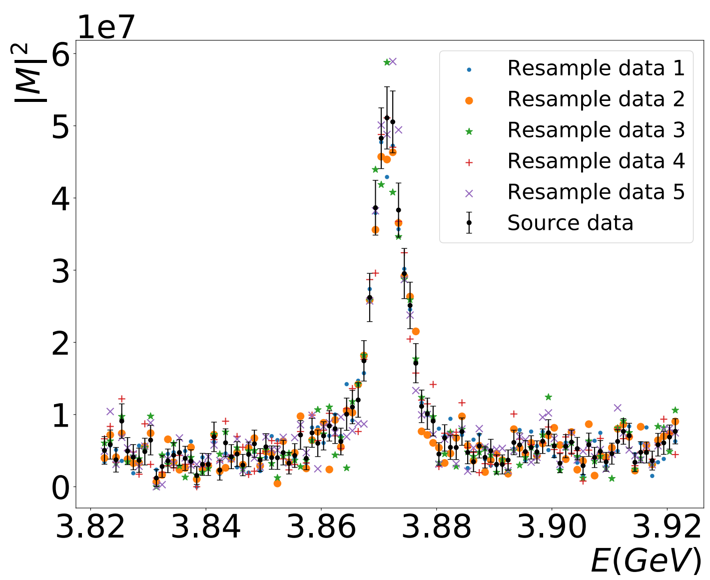
One-channel case

The first step: One-channel line shape in HM picture

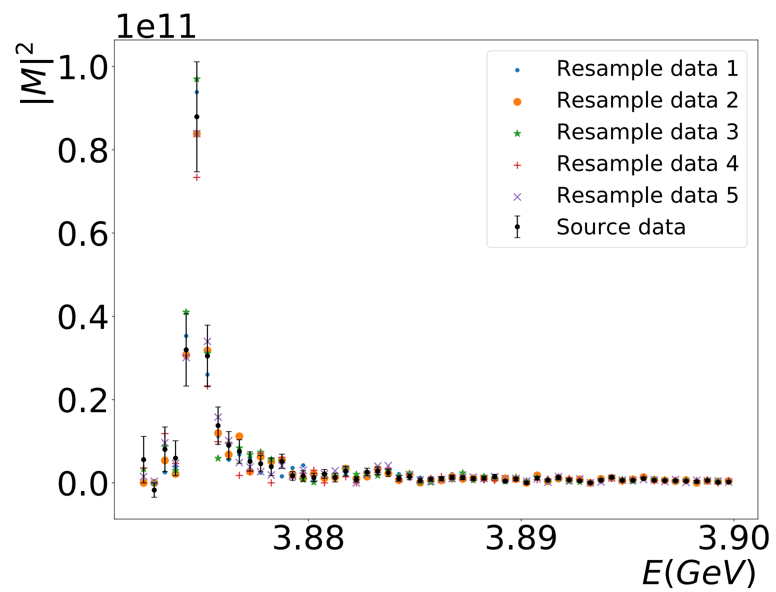
Why one-channel case?

- The most simple case
- A given structure has one foremost channel
- In the isospin limit, sometimes the requirement can be satisfied

$X(3872)$ $D\bar{D}^* + c.c.$



T_{cc}^+ DD^*



Liu, Zhang, Hu, **QW**, PRD105(2022)076013

One-channel case

Generate samples

Effective range expansion

$$T_{NR}(E) = -\frac{2\pi}{\mu}\frac{1}{1/a + (r/2)k^2 - ik}$$

The structure is driven by

Dong, Guo, Zou, PRL126(2021)152001

$$\mathrm{PDF}(E; a, r, \mathrm{threshold}, \sigma) = \int |T_{NR}(E)|^2 G(E' - E)dE'$$

if the elastic channels domain in the production vertex.

Gaussian function
$$G(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma^2}}$$

Liu, Zhang, Hu, **QW**, PRD105(2022)076013

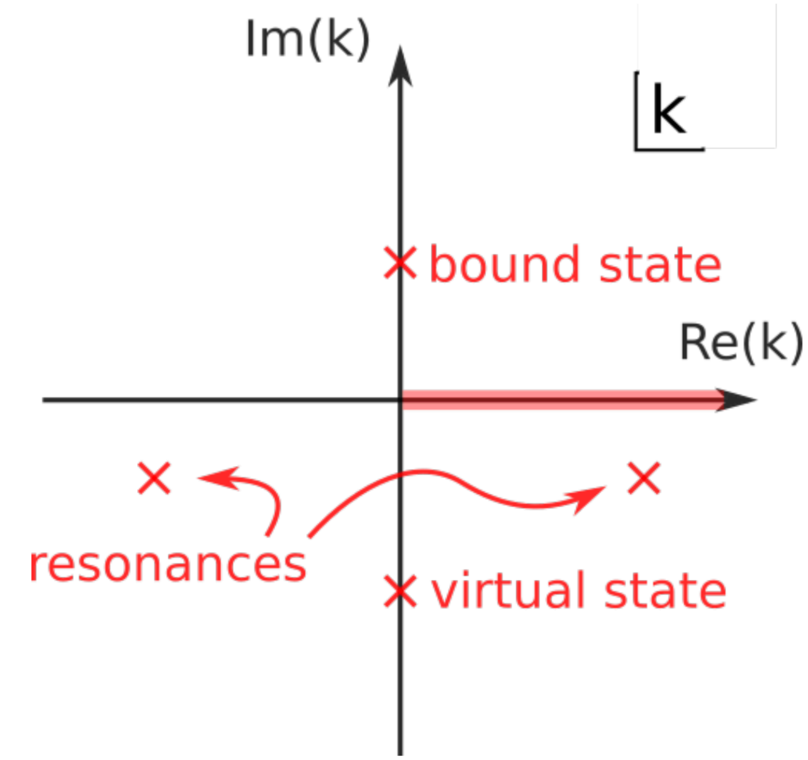
One-channel case

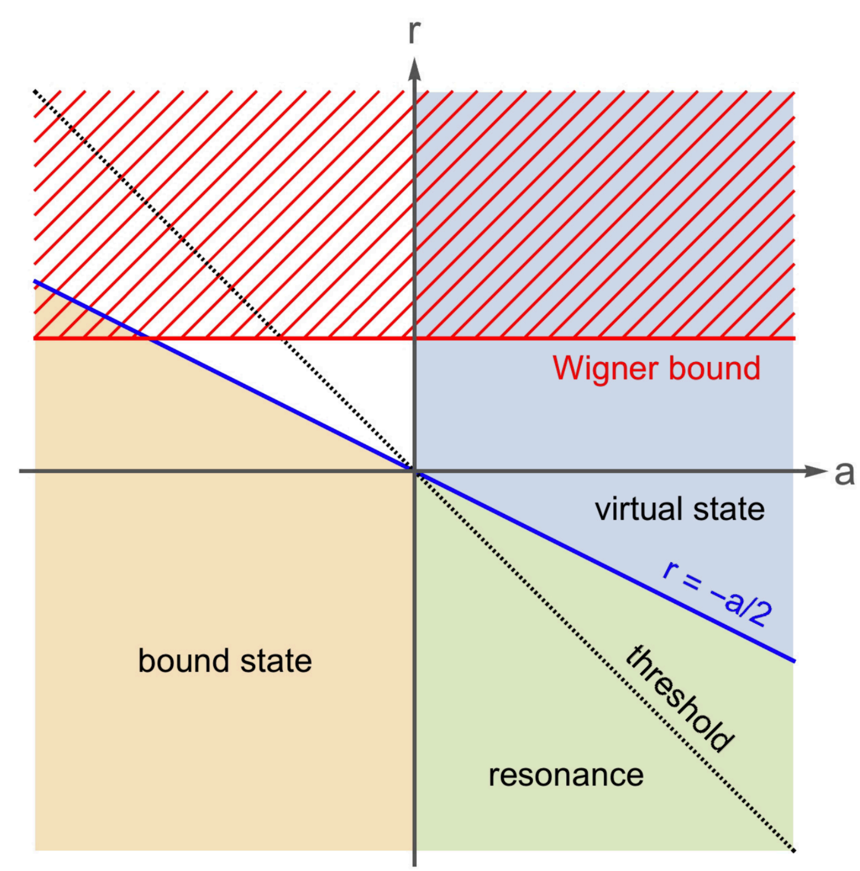
Generate samples

Choose parameter region

$$
\begin{aligned}
a &\in [4.93, 14.80]\ \text{fm},\\
r &\in [0.49, 0.99] \cup [-9.87, -0.49]\ \text{fm},\\
m_1 + m_2 &\in [2.8, 3.9]\ \text{GeV},\\
\sigma &\in [0.5, 10]\ \text{MeV}.
\end{aligned}
$$

- Allow for bound state, virtual state and resonances
- Cover charmonium(-like) region
- Resolution is set to cover usual experimental values
- Generate 150000 samples (100 data points)
- 45000 samples for testing



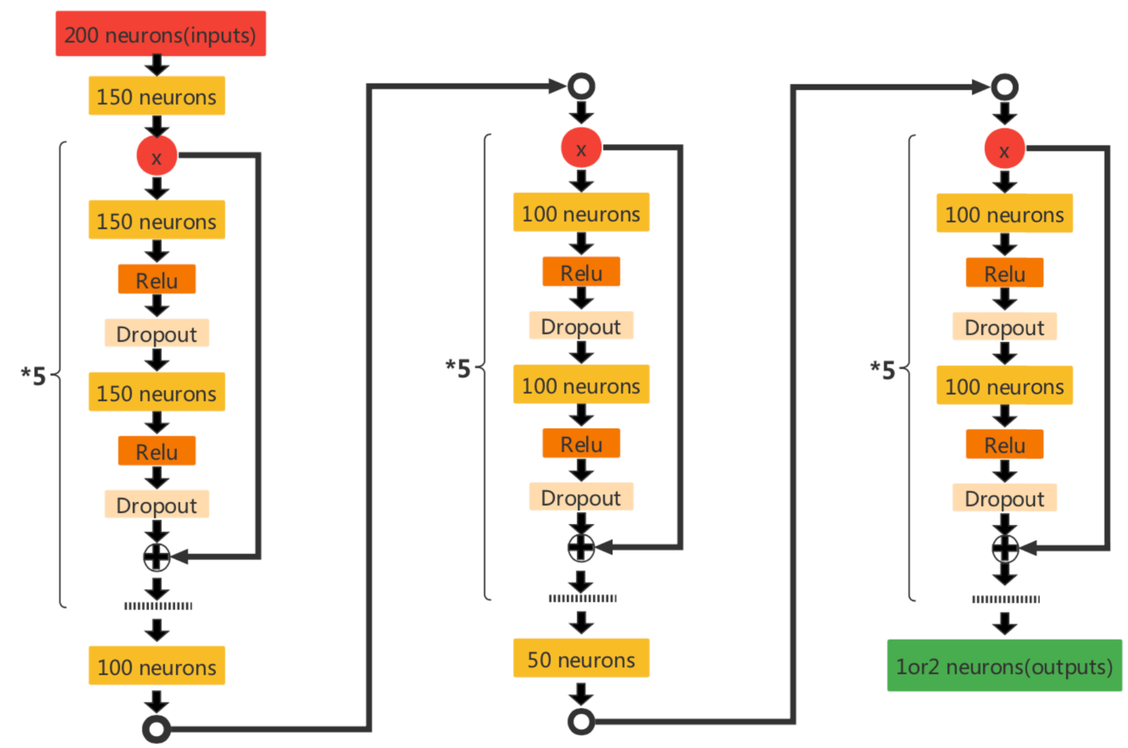


Matuschek, Baru, Guo, Hanhart, EPJA57(2021)101

One-channel case

Training

- ResNet: stable, avoid degradation
- PyTorch: an open source machine learning framework
- Relu: activation function, nonlinearly
- Dropout: drop neurons, avoid overfitting



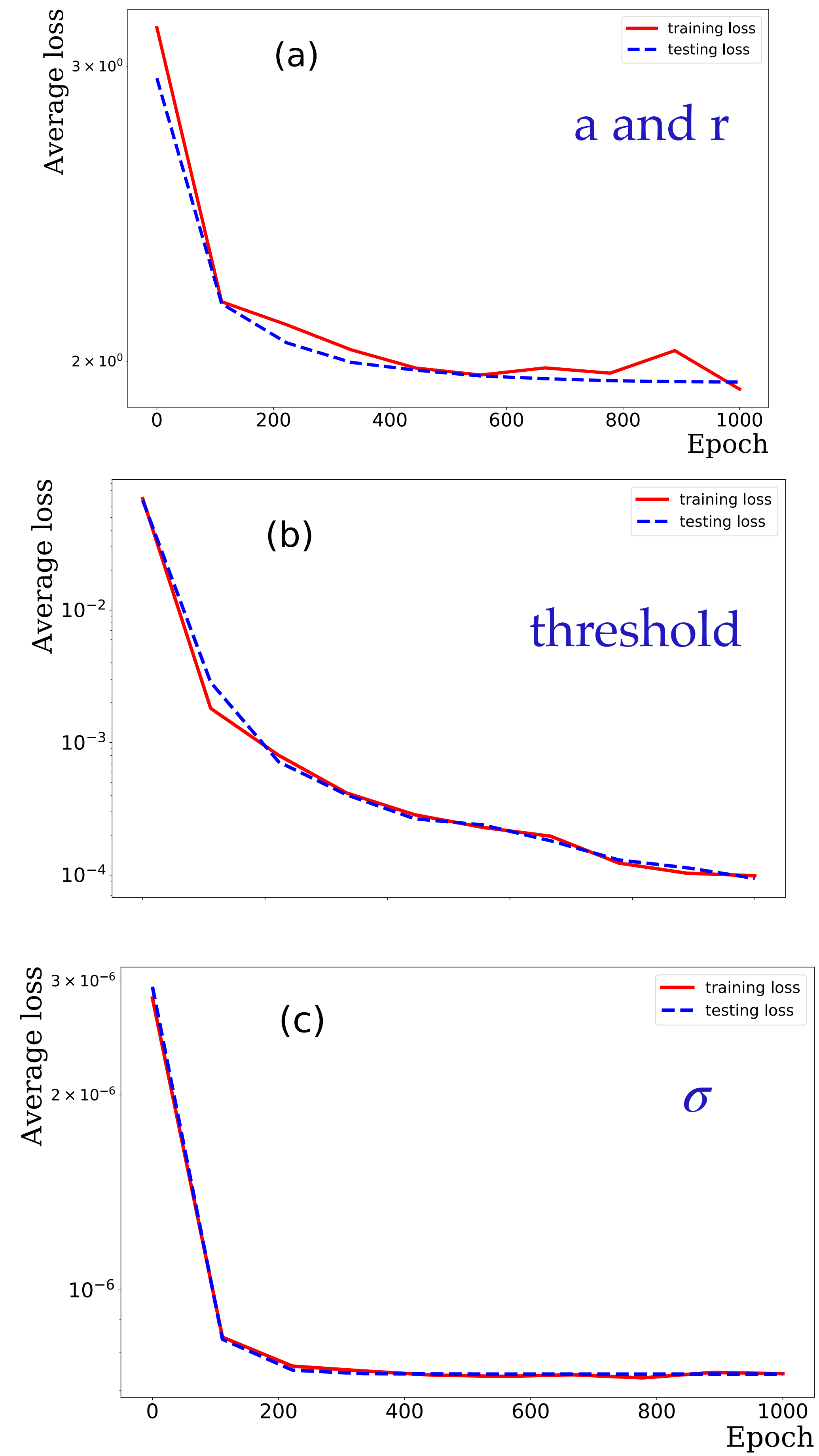
- A multi-layer-perception based ResNet is implemented with PyTorch.
- The parameters a and r are regressed simultaneously.
- The parameters threshold and σ are regressed individually.

Liu, Zhang, Hu, **QW**, PRD105(2022)076013

One-channel case

Training

- MSELoss function: Euclidean distance between the predicted values and the label values.
- MSELoss functions converge rapidly.
- MSELoss function converge after 200 training.
- 1000 training epochs can give a good result.

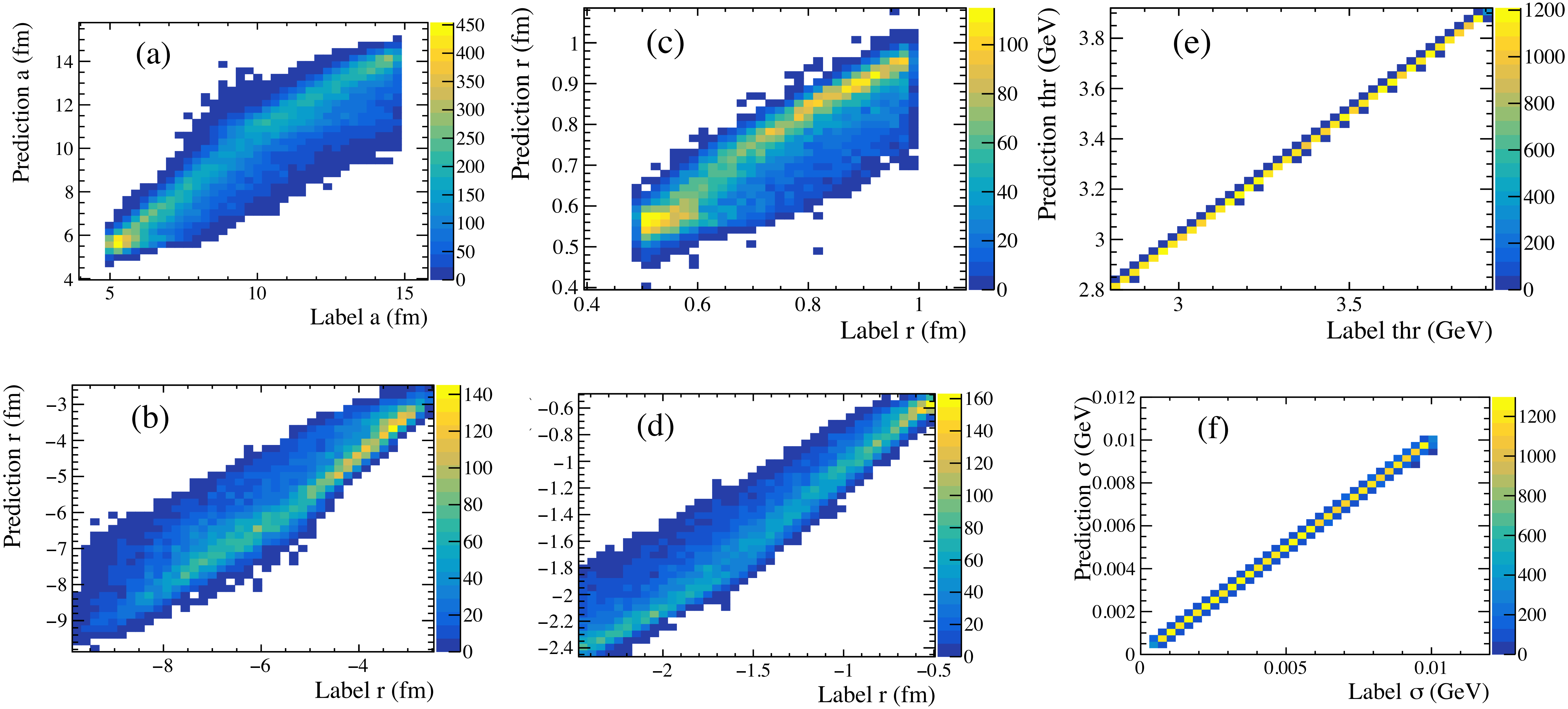


Liu, Zhang, Hu, **QW**, PRD105(2022)076013

One-channel case

Training

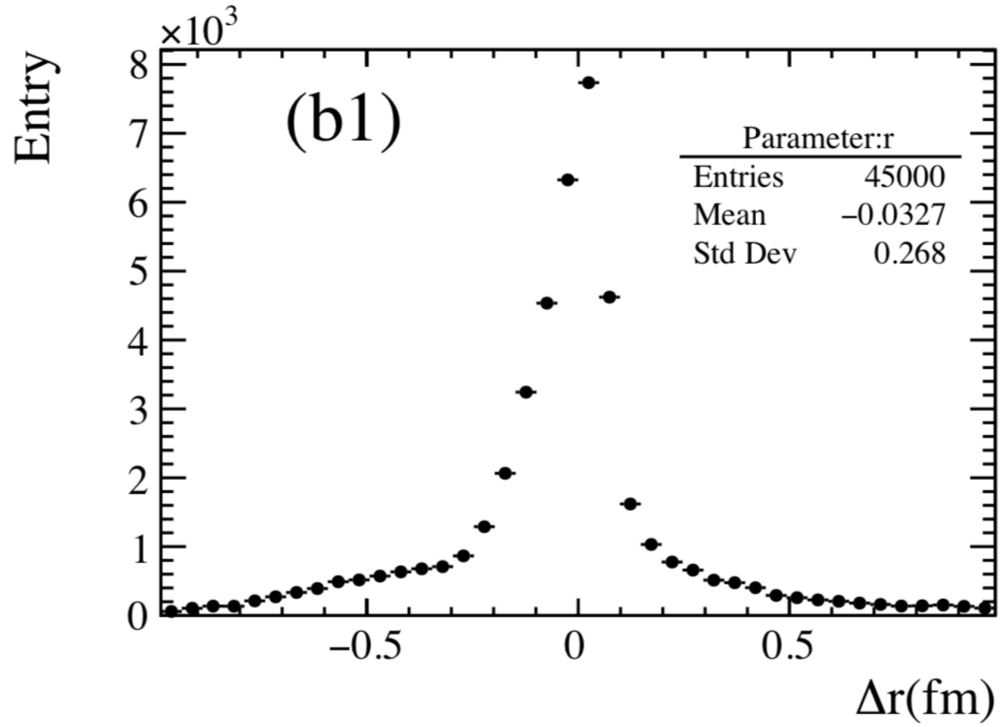
The correlation of the predicted and label values (almost to 1)

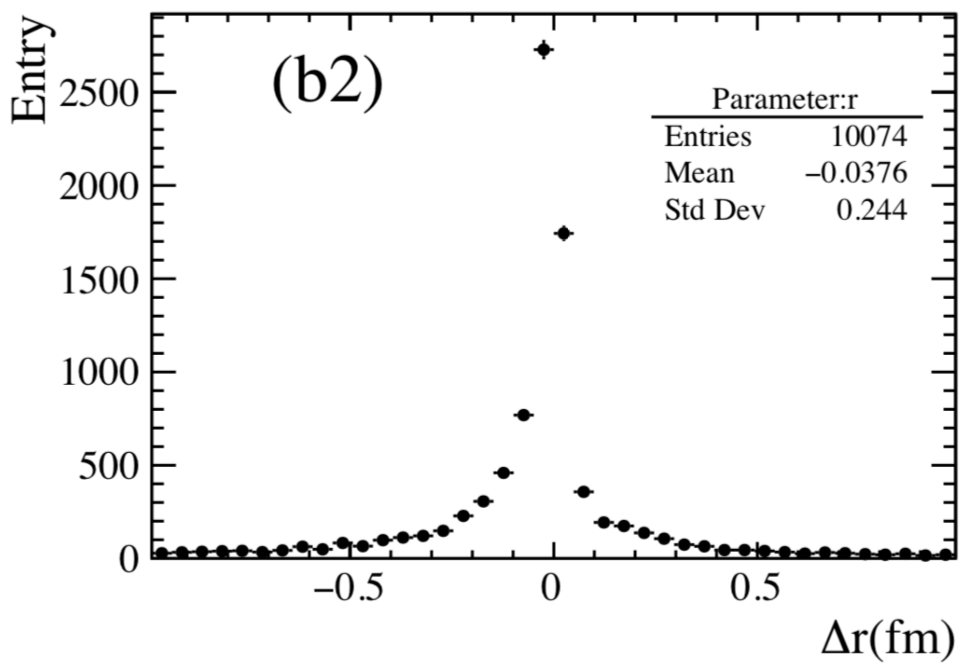


Liu, Zhang, Hu, **QW**, PRD105(2022)076013

One-channel case

Evaluation





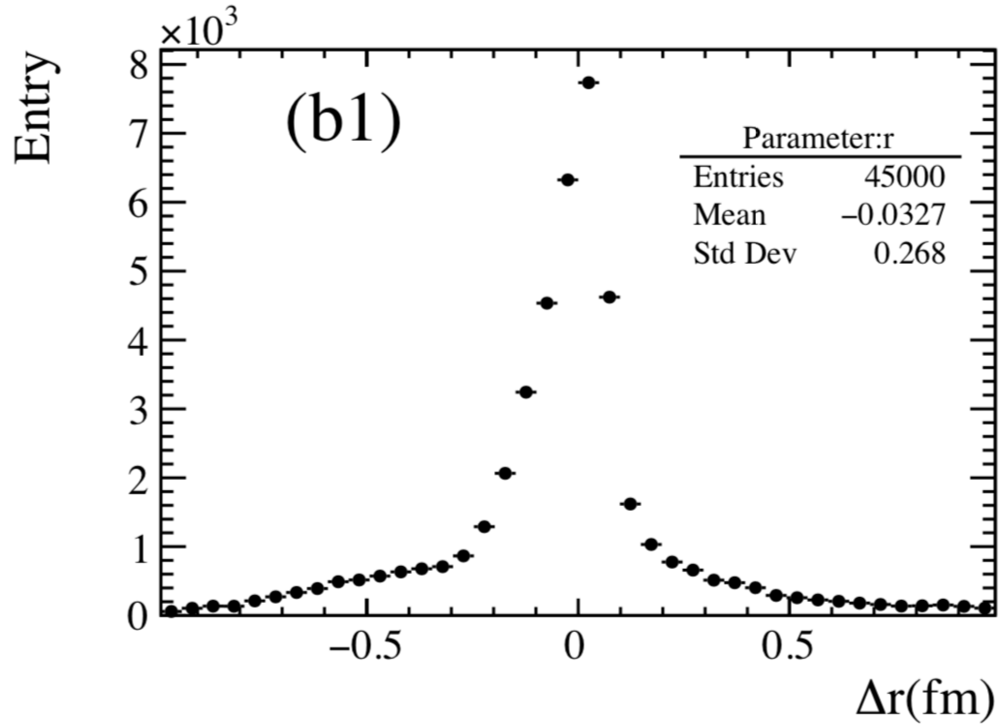
- The difference between the predicted values and the label values.
- The left distribution is obtained by testing 45000 samples.
- The means measure the deviation of the predicted values from the label ones.
- The root-of-mean-square (RMS) measures the intrinsic uncertainties.
- Compare to the 10074/45000 fitting results.

Liu, Zhang, Hu, **QW**, PRD105(2022)076013

One-channel case

Evaluation

Liu, Zhang, Hu, **QW**, PRD105(2022)076013



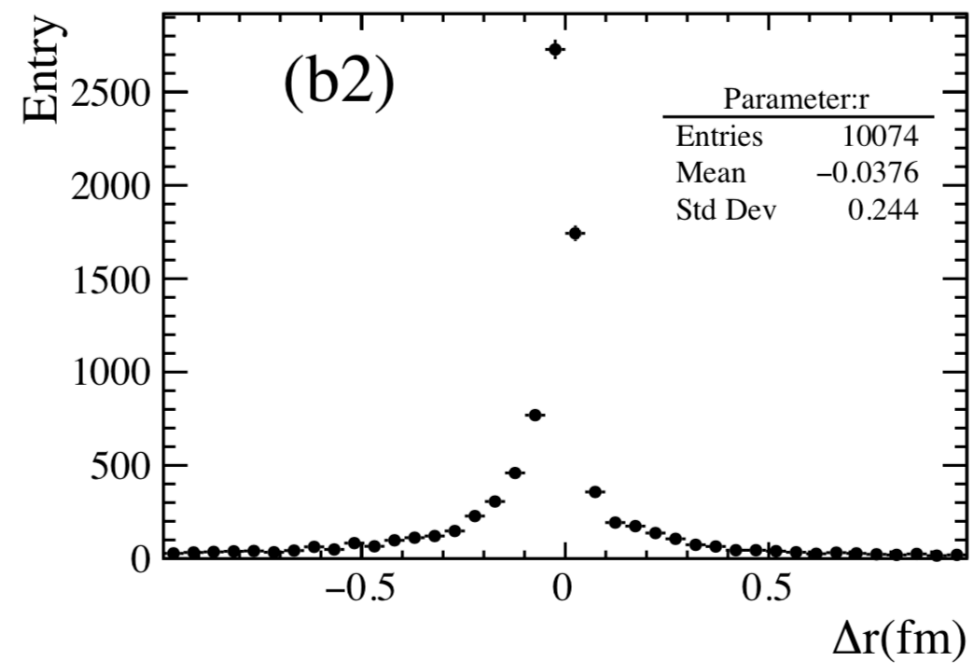


Table I. The biases and errors information of models

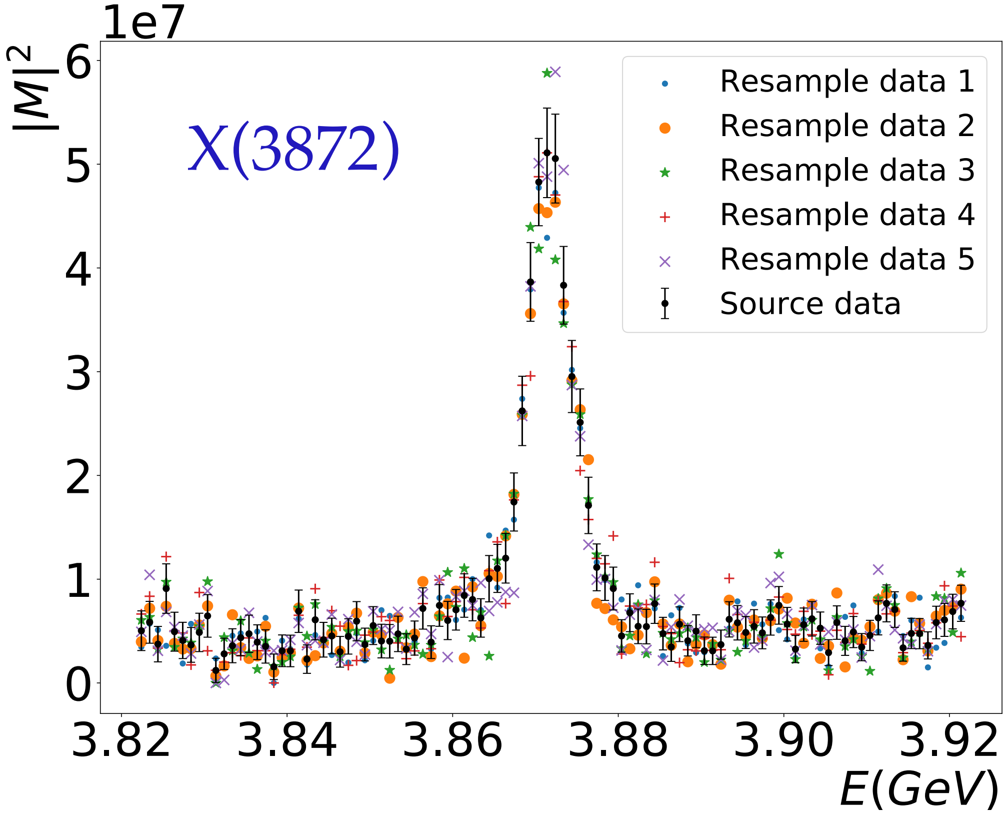
methods→	Deep learning		Fitting	
parameters↓	bias	uncertainty	bias	uncertainty
a (fm)	-0.010	1.040	-1.67	2.740
r (fm)	-0.033	0.268	-0.038	0.244
threshold (MeV)	0.75	0.52	-0.16	0.31
σ (MeV)	-0.0001	0.06	-0.0098	0.10

The biases can be neglected.

One-channel case

Apply to the *X*(3872)

Liu, Zhang, Hu, **QW**, PRD105(2022)076013



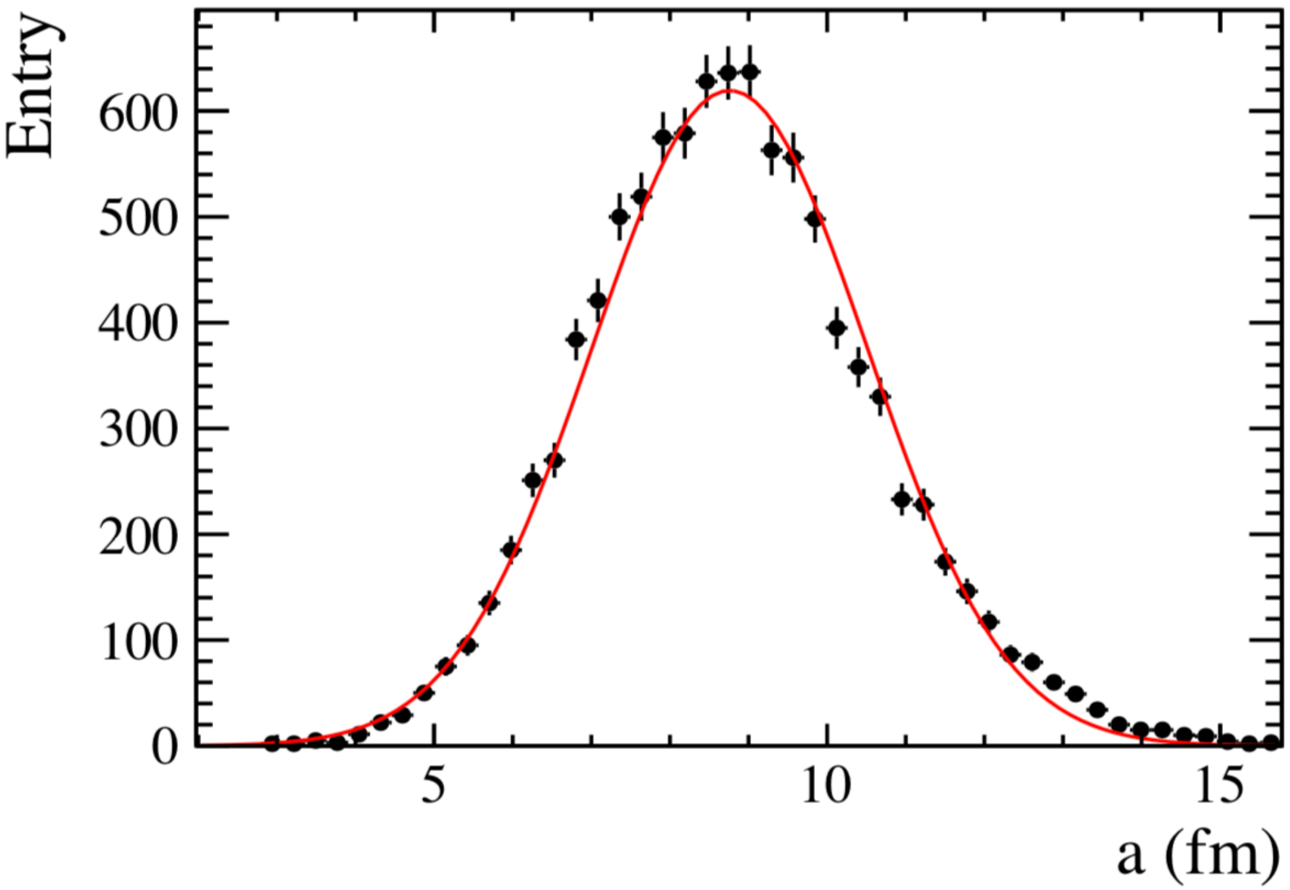
- $D\bar{D}^* + c.c.$ channel
- a, r and threshold are consistent with those from fitting within 1σ
- The errors are obtained in bootstrap (10000 data sets)

$\boldsymbol{X(3872)}$ **parameters**	Deep Learning	Fit
parameter a (fm)	8.76 ± 1.75	9.95 ± 0.34
parameter r (fm)	0.56 ± 0.55	0.32 ± 0.08
parameter threshold (MeV)	3871.30 ± 0.52	3871.20 ± 0.01
parameter σ (MeV)	1.20 ± 0.15	1.70 ± 0.16

One-channel case

Apply to $X(3872)$

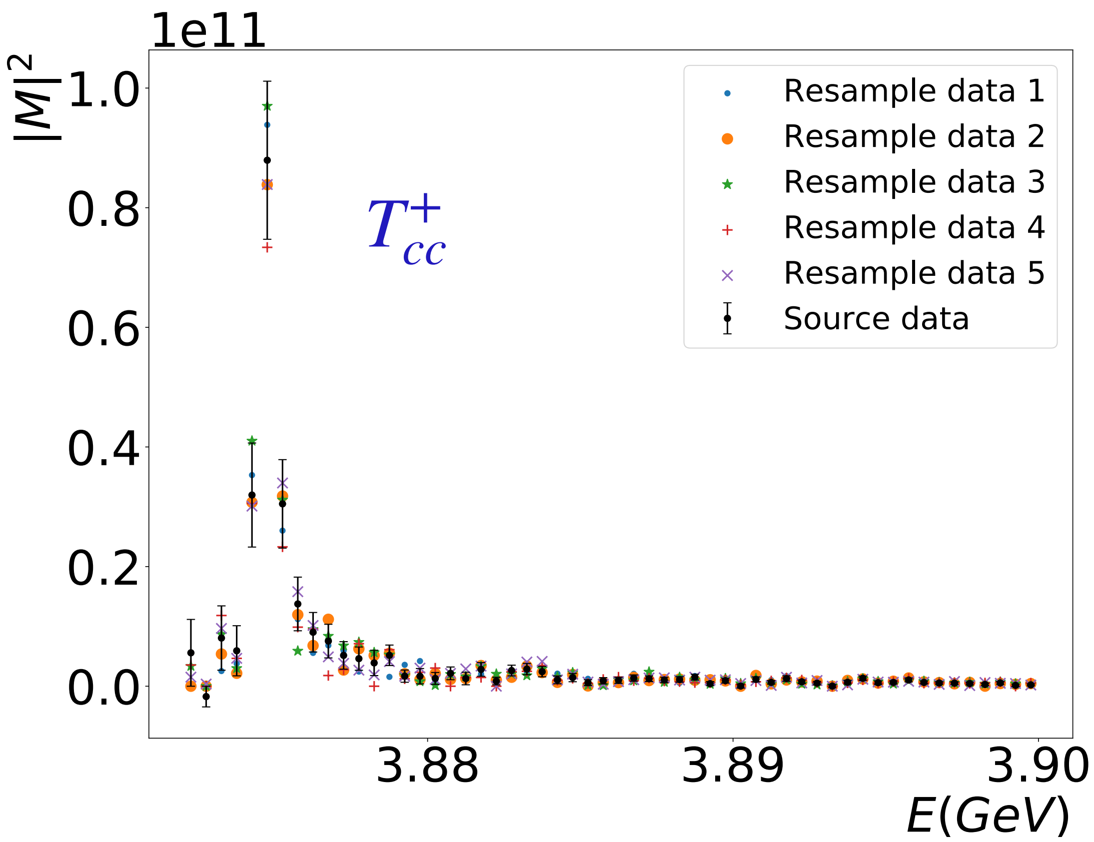
- The distributions of the predicted parameters
- Gaussian-like distribution
- Red lines: fit with Gaussian function
- The mean and RMS are the center value and error of the parameter



Liu, Zhang, Hu, **QW**, PRD105(2022)076013

One-channel case

Apply to the T_{cc}^{+}



- DD^* channel
- a, r and threshold are consistent with those from fitting within 1σ
- The errors are obtained in bootstrap (10000 data sets)

$\boldsymbol{T_{cc}^{+}}$ **parameters**	Deep Learning	Fit
parameter a (fm)	8.23 ± 1.04	13.74 ± 4.77
parameter r (fm)	-2.79 ± 0.27	-2.15 ± 0.21
parameter threshold (MeV)	3874.83 ± 0.51	3874.53 ± 0.13
parameter σ (MeV)	1.10 ± 0.06	0.11 ± 0.12

Liu, Zhang, Hu, **QW**, PRD105(2022)076013

Multi-channel case

The history of pentaquarks — Bing-Song Zou, Sci.Bull.66(2021)1258

$\Lambda(1405)$ predicted by Dalitz and Tuan in 1959 — Dalitz and Tuan, PRL2(1959)425

- An excited state of a three-quark (uds) system
- $\bar{K}N$ hadronic molecule with $udsq\bar{q}$

A similar situation for $N^\star(1535)$

- An excited state of a three-quark (uds) system
- $\bar{K}\Sigma - \bar{K}\Lambda$ dynamical generated state with $qqqs\bar{s}$ — Kaiser, Siegel, Weise, NPA594(1995)325

Pentaquark in hidden charm sector

Liu, Zou, PRL96(2006)042002

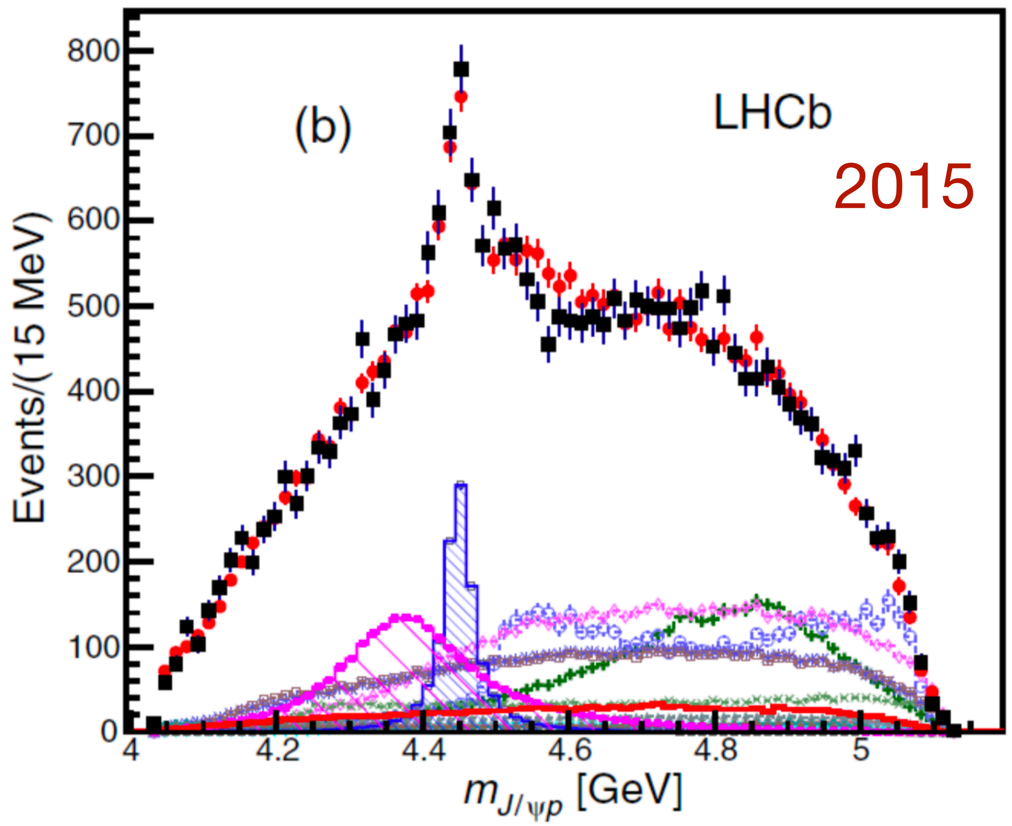
Wu, Molina, Oset, Zou, PRL105(2010)232001

(I,S)	z_R (MeV)	g_a		
$(1/2,0)$		$\bar{D}\Sigma_c$	$\bar{D}\Lambda_c^+$	
	4269	2.85	0	
$(0,-1)$		$\bar{D}_s\Lambda_c^+$	$\bar{D}\Xi_c$	$\bar{D}\Xi_c'$
	4213	1.37	3.25	0
	4403	0	0	2.64

(I,S)	z_R (MeV)	g_a		
$(1/2,0)$		$\bar{D}^*\Sigma_c$	$\bar{D}^*\Lambda_c^+$	
	4418	2.75	0	
$(0,-1)$		$\bar{D}_s^*\Lambda_c^+$	$\bar{D}^*\Xi_c$	$\bar{D}^*\Xi_c'$
	4370	1.23	3.14	0
	4550	0	0	2.53

Multi-channel case

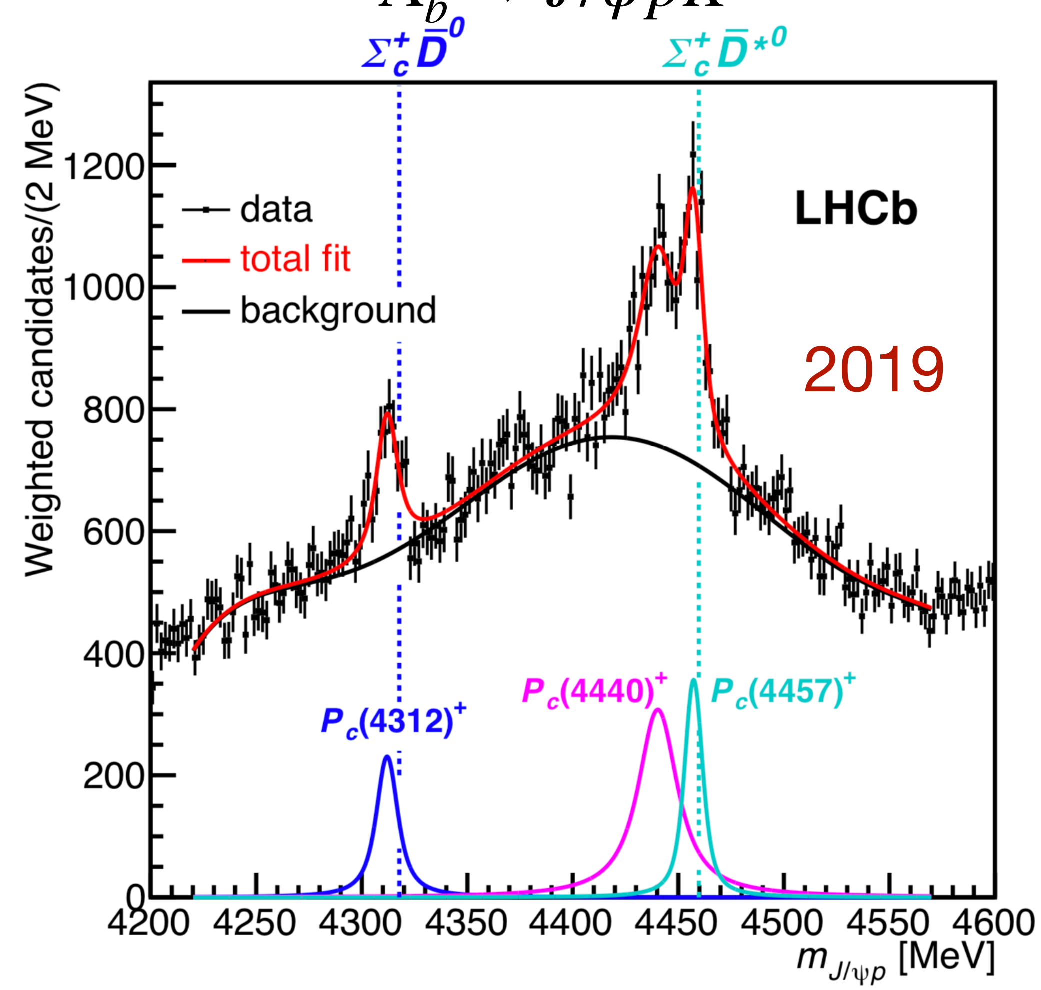
The observation of hidden charm pentaquarks



$P_c(4380): 4380 \pm 8 \pm 29$ MeV

$P_c(4450)^+: 4449.8 \pm 1.7 \pm 2.5$ MeV

$\Lambda_b \to J/\psi p K^-$



$P_c(4312)^+: 4311.9 \pm 0.7^{+6.8}_{-0.6}$ MeV

$P_c(4440)^+: 4440.3 \pm 1.3^{+4.1}_{-4.7}$ MeV

$P_c(4457)^+: 4457.3 \pm 0.6^{+4.1}_{-1.7}$ MeV

"spin puzzle"

$J(\Sigma_c) = \frac{1}{2}$ $\quad J(\bar{D}^*) = 1$

$J(P_c(4440)) = ?$ $\quad J(P_c(4457)) = ?$

Wang, Huang, Zhang, Zou, PRC84(2011)015203,

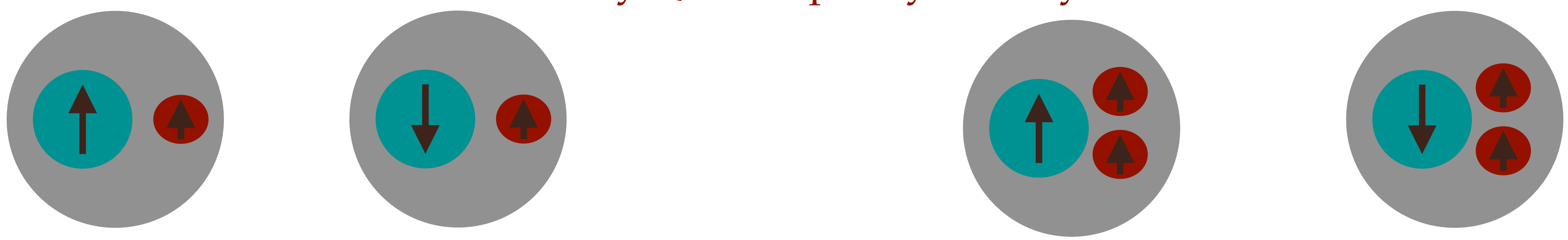
Wu, Lee, Zou, PRC85(2012)044002

Multi-channel case

The $\Sigma_c^{(*)}\bar{D}^{(*)}$ molecular picture   Du, Baru, Guo, Hanhart, Meißner, Oller, **QW,** PRL124(2020)072001

$m_Q \to \infty$ the strong interaction independent of the spin of heavy quark

Heavy Quark Spin Symmetry



$J = s_l + \frac{1}{2}$ $J = s_l - \frac{1}{2}$ $J = s_l + \frac{1}{2}$ $J = s_l - \frac{1}{2}$

$m_{D^*} - m_D = 142\ \mathrm{MeV}$ $m_{\Sigma_c^*} - m_{\Sigma_c} = 64\ \mathrm{MeV}$

$s_l = \frac{1}{2}^-$ doublet $s_l = 1^+$ doublet

Spin rearrangement

$$\left([\bar{Q}q_{J_{l_1}}]_{j_1}[Q(qq)_{J_{l_2}}]_{j_2}\right)_J \sim \sum_{HL} \mathscr{C}^{j_1 j_2 J}_{j_{l_1} j_{l_2} HL} \left((\bar{Q}Q)_H (qqq)_L\right)_J$$

$\bar{D}^{(*)}$ $\Sigma_c^{(*)}$

Two LECs to LO

$$C_{\frac{1}{2}} \equiv \langle H \otimes \frac{1}{2} | \hat{H} | H \otimes \frac{1}{2} \rangle$$

$$C_{\frac{3}{2}} \equiv \langle H \otimes \frac{3}{2} | \hat{H} | H \otimes \frac{3}{2} \rangle$$

Multi-channel case

Du, Baru, Guo, Hanhart, Meißner, Oller, **QW,** PRL124(2020)072001

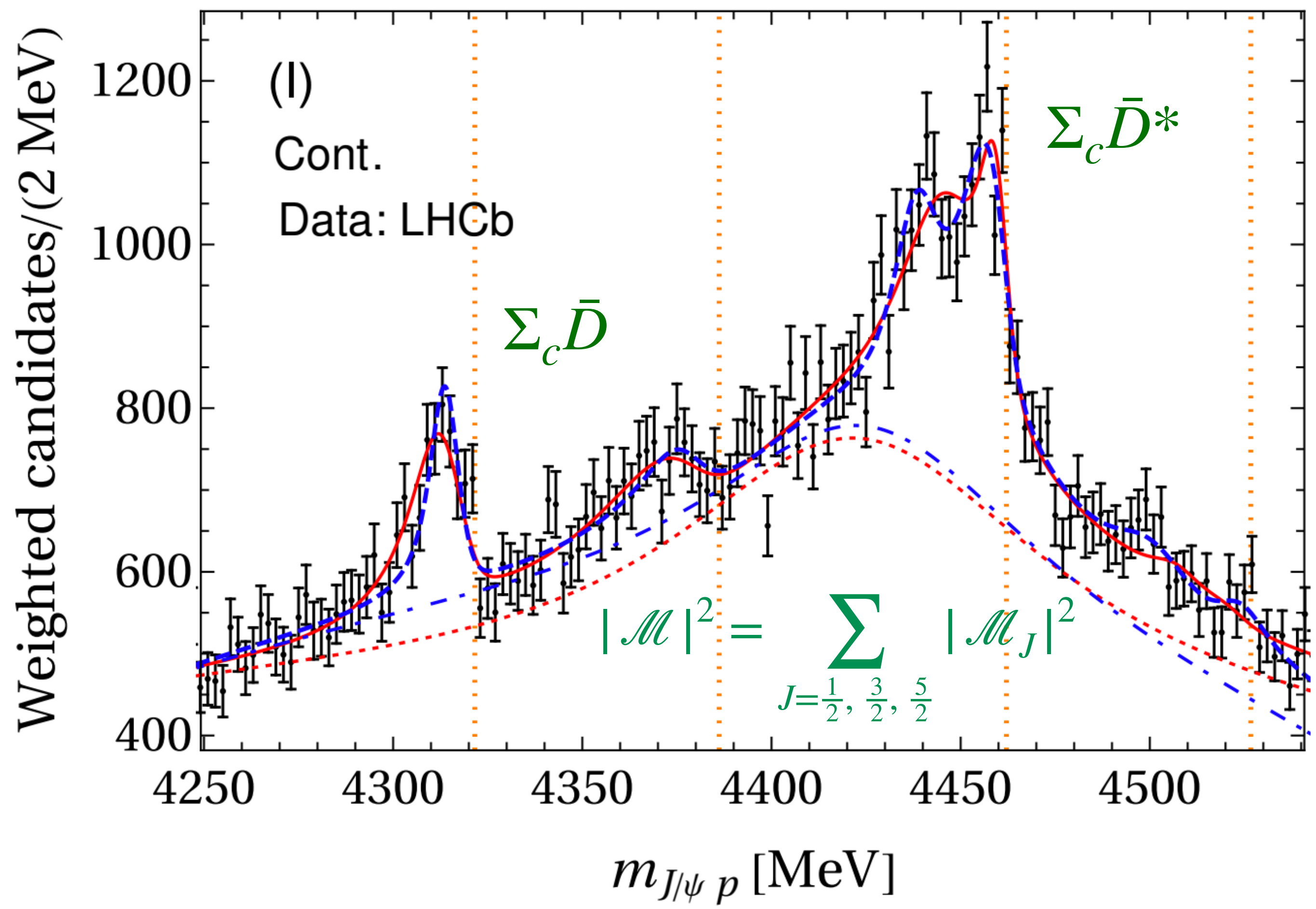
The $\Sigma_c^{(*)}\bar{D}^{(*)}$ molecular picture

Liu et.al., PRL122(2019)242001

Solution A (χ^2/d.o.f. = 1.01)

Solution B (χ^2/d.o.f. = 1.03)

Scenario	Molecule	J^P	B (MeV)	M (MeV)
A	$\bar{D}\Sigma_c$	$\frac{1}{2}^-$	7.8 – 9.0	4311.8 – 4313.0
A	$\bar{D}\Sigma_c^*$	$\frac{3}{2}^-$	8.3 – 9.2	4376.1 – 4377.0
A	$\bar{D}^*\Sigma_c$	$\frac{1}{2}^-$	Input	4440.3
A	$\bar{D}^*\Sigma_c$	$\frac{3}{2}^-$	Input	4457.3
A	$\bar{D}^*\Sigma_c^*$	$\frac{1}{2}^-$	25.7 – 26.5	4500.2 – 4501.0
A	$\bar{D}^*\Sigma_c^*$	$\frac{3}{2}^-$	15.9 – 16.1	4510.6 – 4510.8
A	$\bar{D}^*\Sigma_c^*$	$\frac{5}{2}^-$	3.2 – 3.5	4523.3 – 4523.6
B	$\bar{D}\Sigma_c$	$\frac{1}{2}^-$	13.1 – 14.5	4306.3 – 4307.7
B	$\bar{D}\Sigma_c^*$	$\frac{3}{2}^-$	13.6 – 14.8	4370.5 – 4371.7
B	$\bar{D}^*\Sigma_c$	$\frac{1}{2}^-$	Input	4457.3
B	$\bar{D}^*\Sigma_c$	$\frac{3}{2}^-$	Input	4440.3
B	$\bar{D}^*\Sigma_c^*$	$\frac{1}{2}^-$	3.1 – 3.5	4523.2 – 4523.6
B	$\bar{D}^*\Sigma_c^*$	$\frac{3}{2}^-$	10.1 – 10.2	4516.5 – 4516.6
B	$\bar{D}^*\Sigma_c^*$	$\frac{5}{2}^-$	25.7 – 26.5	4500.2 – 4501.0



- Two parameters determined by

$P_c(4440)$, $P_c(4457)$

- Two solutions

- Two parameters g_S, g_D for $J/\psi p$, $\eta_c p$
- Predict pole positions accurately
- $\chi^2_A < \chi^2_B$
- The effect of each data point is different

Multi-channel case

Du, Baru, Guo, Hanhart, Meißner, Oller, **QW,** PRL124(2020)072001

The $\Sigma_c^{(*)}\bar{D}^{(*)}$ molecular picture

Liu et.al., PRL122(2019)242001

Scenario	Molecule	J^P	B (MeV)	M (MeV)
A	$\bar{D}\Sigma_c$	$\frac{1}{2}^-$	7.8 – 9.0	4311.8 – 4313.0
A	$\bar{D}\Sigma_c^*$	$\frac{3}{2}^-$	8.3 – 9.2	4376.1 – 4377.0
A	$\bar{D}^*\Sigma_c$	$\frac{1}{2}^-$	Input	4440.3
A	$\bar{D}^*\Sigma_c$	$\frac{3}{2}^-$	Input	4457.3
A	$\bar{D}^*\Sigma_c^*$	$\frac{1}{2}^-$	25.7 – 26.5	4500.2 – 4501.0
A	$\bar{D}^*\Sigma_c^*$	$\frac{3}{2}^-$	15.9 – 16.1	4510.6 – 4510.8
A	$\bar{D}^*\Sigma_c^*$	$\frac{5}{2}^-$	3.2 – 3.5	4523.3 – 4523.6
B	$\bar{D}\Sigma_c$	$\frac{1}{2}^-$	13.1 – 14.5	4306.3 – 4307.7
B	$\bar{D}\Sigma_c^*$	$\frac{3}{2}^-$	13.6 – 14.8	4370.5 – 4371.7
B	$\bar{D}^*\Sigma_c$	$\frac{1}{2}^-$	Input	4457.3
B	$\bar{D}^*\Sigma_c$	$\frac{3}{2}^-$	Input	4440.3
B	$\bar{D}^*\Sigma_c^*$	$\frac{1}{2}^-$	3.1 – 3.5	4523.2 – 4523.6
B	$\bar{D}^*\Sigma_c^*$	$\frac{3}{2}^-$	10.1 – 10.2	4516.5 – 4516.6
B	$\bar{D}^*\Sigma_c^*$	$\frac{5}{2}^-$	25.7 – 26.5	4500.2 – 4501.0

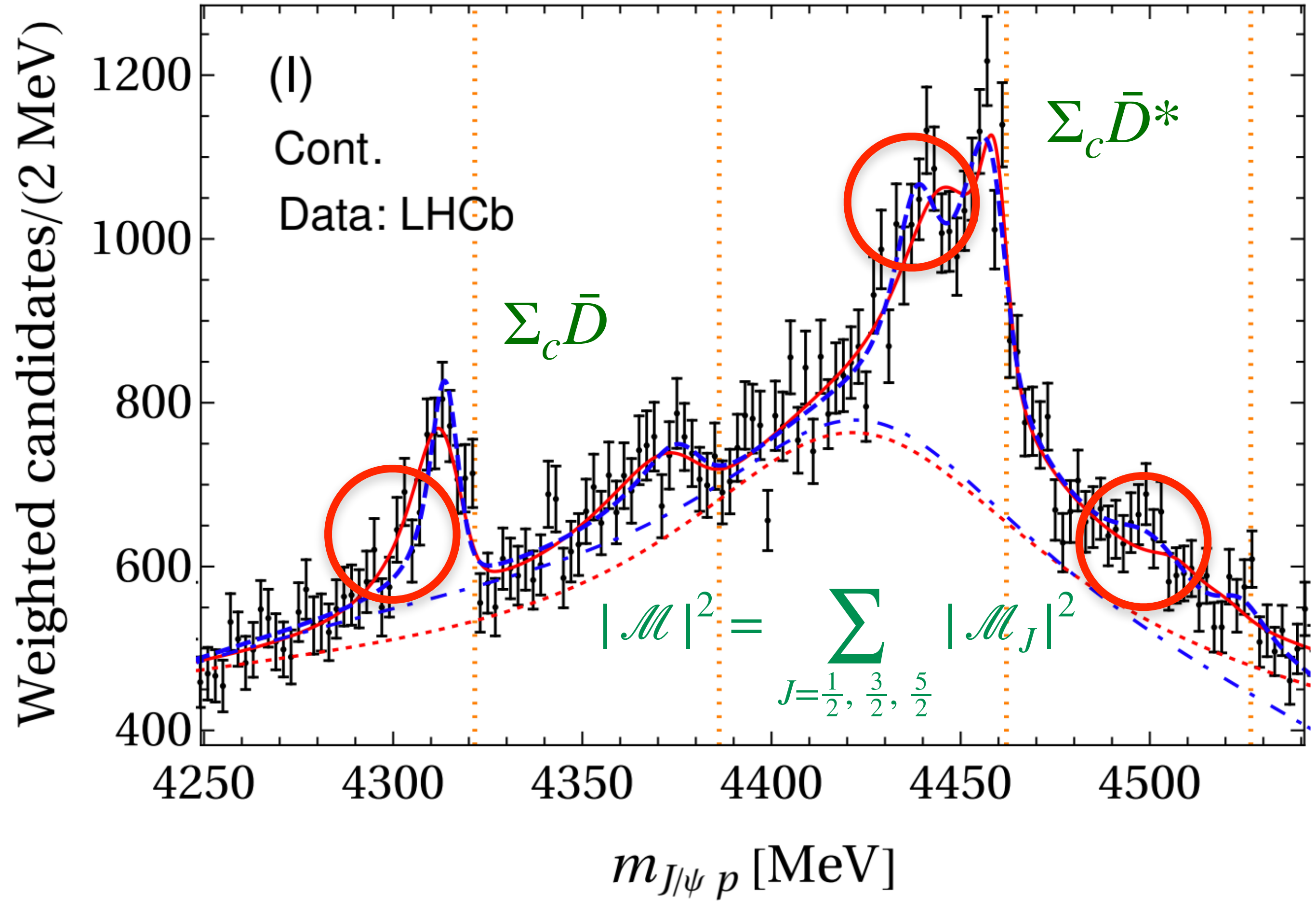
- Two parameters determined by

$P_c(4440),\ P_c(4457)$

- Two solutions

Solution A (χ^2/d.o.f. = 1.01)

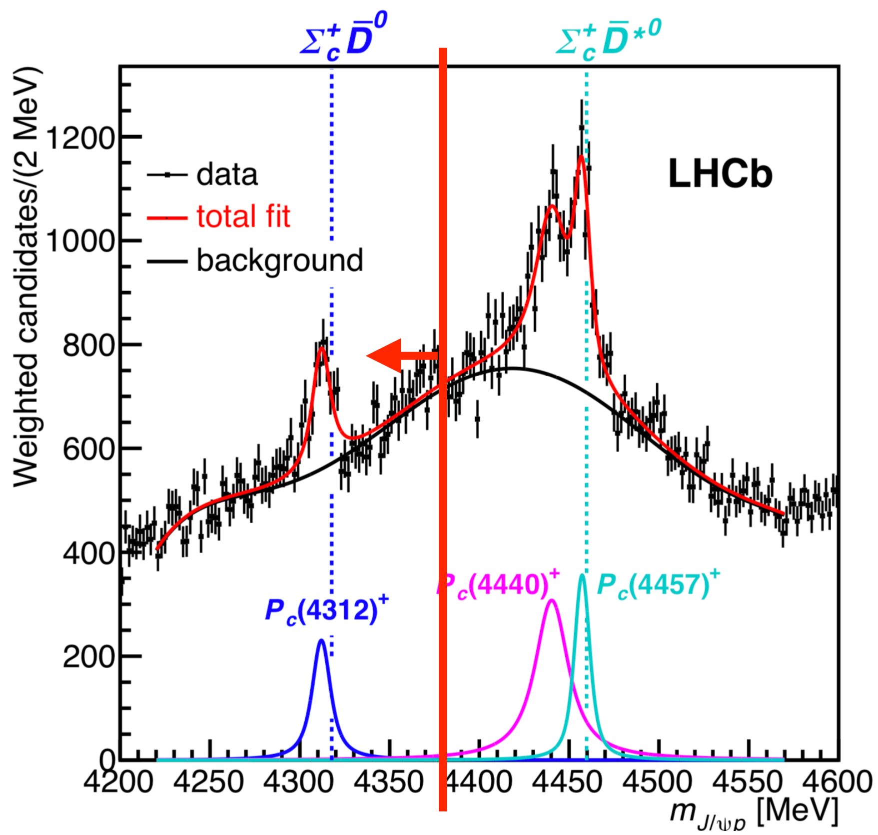
Solution B (χ^2/d.o.f. = 1.03)

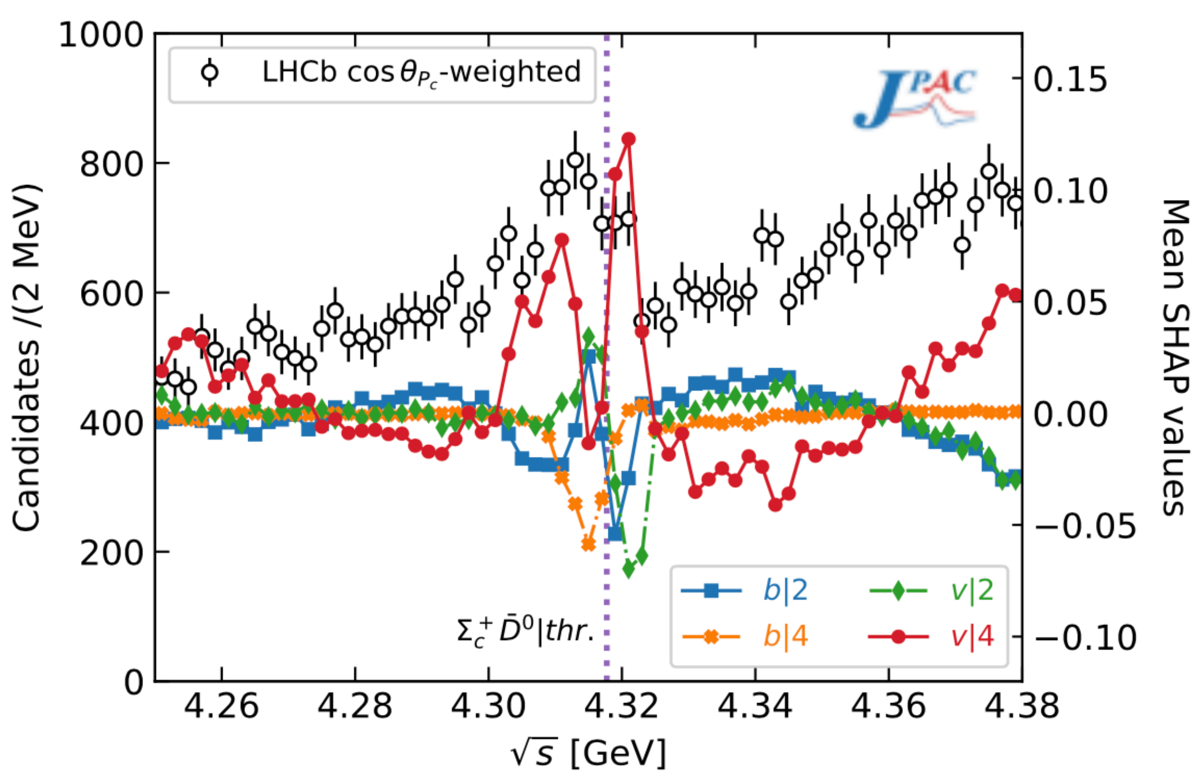


- Two parameters $g_S,\ g_D$ for $J/\psi p,\ \eta_c p$
- Predict pole positions accurately
- $\chi_A^2 < \chi_B^2$
- The effect of each data point is different

Multi-channel case

Hidden charm pentaquarks in machine learning





- Focus on the region below 4.375 GeV
- Two channel case: $J/\psi p$, $\Sigma_c\bar{D}$
- Do not respect HQSS
- Parametrization

$$I(s) = \rho(s)[\,|P(s)T(s)|^2 + B(s)]$$

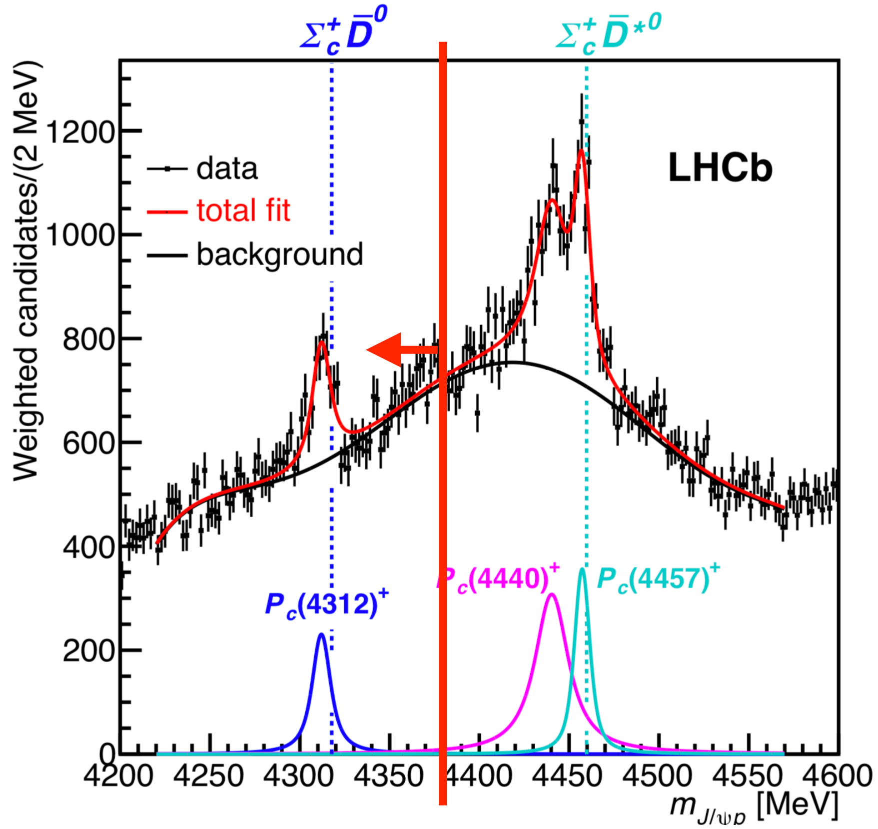
$$T(s) = \frac{m_{22} - ik_2}{(m_{11} - ik_1)(m_{22} - 9k_2) - m_{12}^2}$$

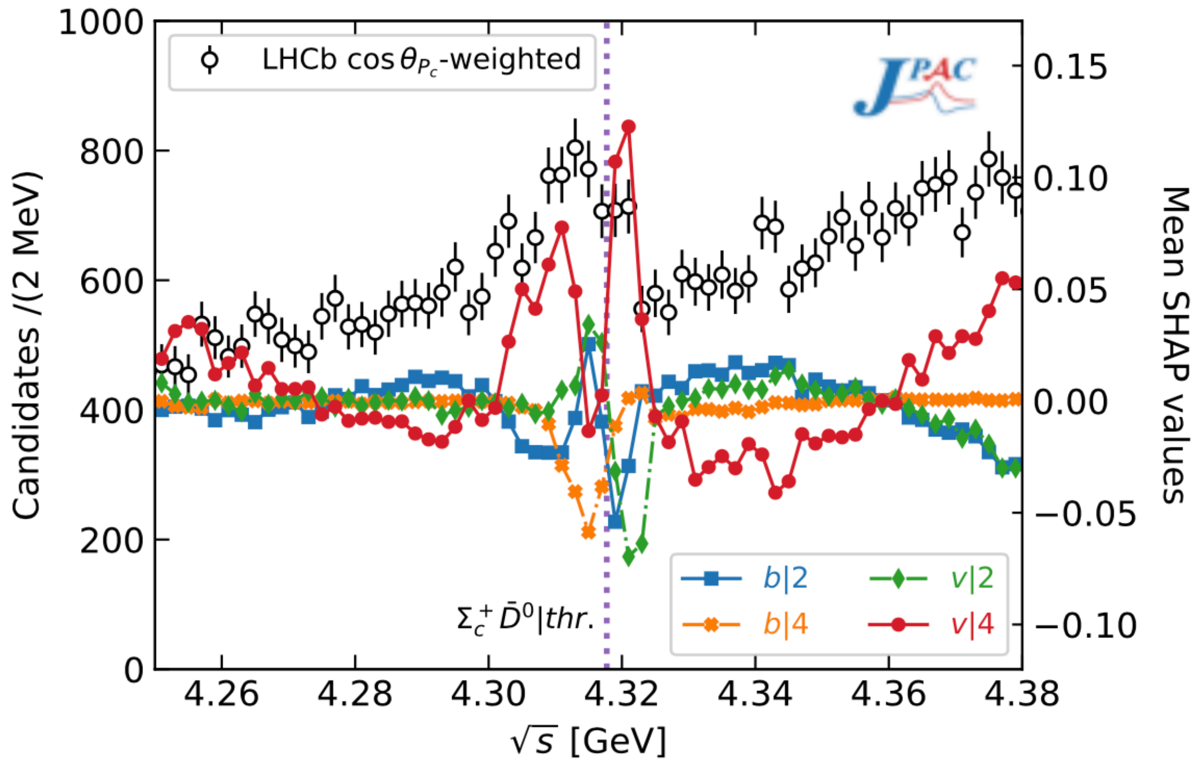
- $P_c(4312)$ is a virtual state
- SHAP analysis indicates the role of each bin

JPAC, PRD105(2022)L091501

Multi-channel case

Hidden charm pentaquarks in machine learning





- Nuclear physics

Niu et.al., PLB778(2018)48, Niu et.al., PRC99(2019)064307,

Ma et.al., CPC44(2020)014104,

Bedaque et al., EPJA3(2021)025003,

Sombillo, et al., PRD104(2021)036001, PRD102(2020)016024,

PRD104(2021)036001,……

- High energy nuclear physics

Balidi et.al., PRD93(2016)094034,

Boehnlein et.al., RMP94(2022)031003……

- Experimental data analysis

Guest et.al., Annu. Rev. Nucl. Part. Sci68(2018)161……

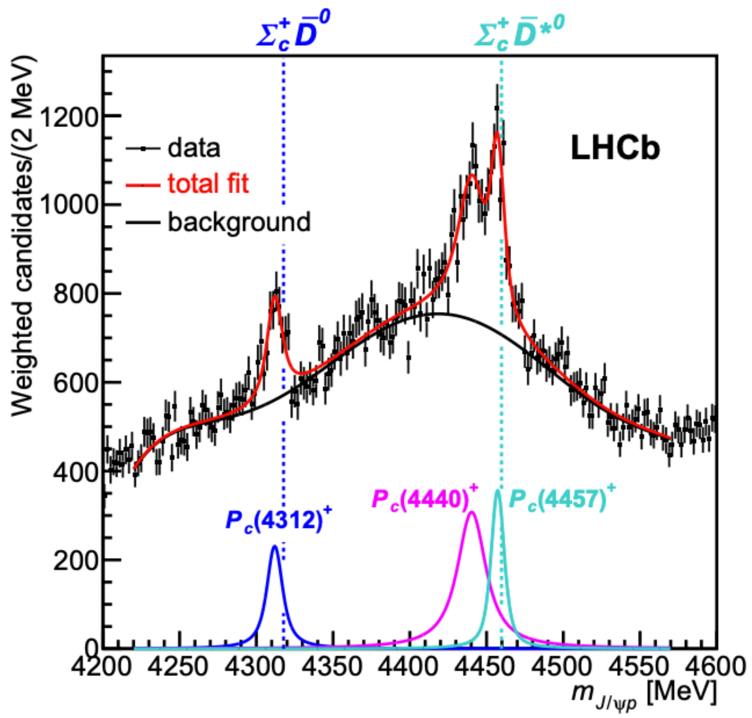
- Theoretical physics

Carleo et.al., Science 355(2017) 602……

JPAC, PRD105(2022)L091501

Multi-channel case

LHCb, PRL122(2019)222001

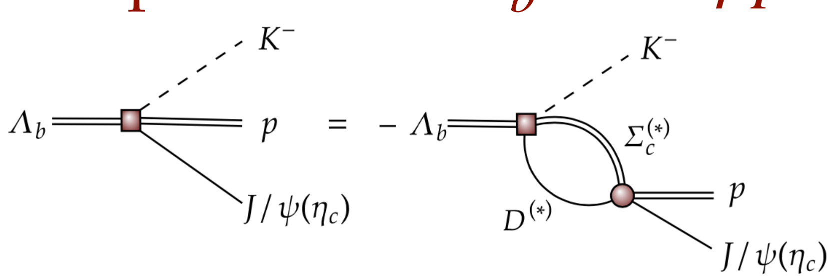


The $\Sigma_c^{(*)}\bar{D}^{(*)}$ molecular picture

- $P_c(4312)$ bound state or virtual state?
- Spin assignment of $P_c(4440)$ and $P_c(4457)$?
- The pole situations for all the P_c states?
- Whether NN approach obtains more than the normal fitting approach?

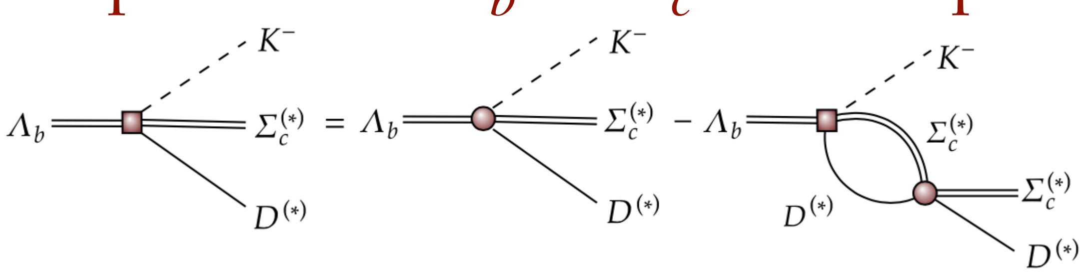
LO HQEFT, Du, Baru, Guo, Hanhart, Meißner, Oller, **QW,** PRL124(2020)072001

The decay amplitude for $\Lambda_b \to J/\psi p K^-$ process



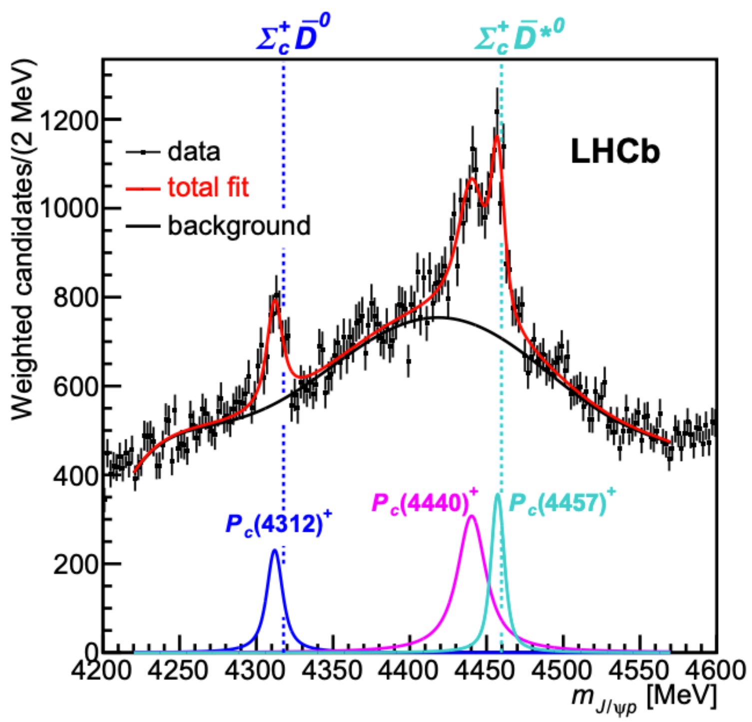
Zhang, Liu, Hu, **QW,** Meißner, Sci.Bull.68(2023)981-989

The decay amplitude for $\Lambda_b \to \Sigma_c^{(*)}\bar{D}^{(*)}K^-$ process



Multi-channel case

LHCb, PRL122(2019)222001



The $\Sigma_c^{(*)}\bar{D}^{(*)}$ molecular picture

- $P_c(4312)$ bound state or virtual state?
- Spin assignment of $P_c(4440)$ and $P_c(4457)$?
- The pole situations for all the P_c states?
- Whether NN approach obtains more than the normal fitting approach?

LO HQEFT, Du, Baru, Guo, Hanhart, Meißner, Oller, **QW,** PRL124(2020)072001

The decay amplitude for $\Lambda_b \to J/\psi p K^-$ process

$$U_i^J(E,k) = -\sum_\alpha \int \frac{d^3\mathbf{q}}{(2\pi)^3} \mathscr{V}_{i\alpha}^J(k) G_\alpha(E,q) U_\alpha^J(E,q)$$

$\alpha,\ \beta,\ \ldots$ for $\Sigma_c^{(*)}\bar{D}^{(*)}$ channels

The decay amplitude for $\Lambda_b \to \Sigma_c^{(*)}\bar{D}^{(*)}K^-$ process

$i,\ j,\ \ldots$ for $J/\psi p,\ \eta_c p$ channels

$$U_\alpha^J(E,p) = P_\alpha^J - \sum_\beta \int \frac{d^3\mathbf{q}}{(2\pi)^3} V_{\alpha\beta}^J(E,p,q) G_\beta(E,q) U_\beta^J(E,q)$$

Zhang, Liu, Hu, **QW,** Meißner, Sci.Bull.68(2023)981-989

Multi-channel case

The Probability Distribution Function

Zhang, Liu, Hu, **QW,** Meißner, Sci.Bull.68(2023)981-989

$$\mathrm{PDF}(E;\mathscr{P}) = \alpha \sum_J \int |U^J|^2 \mathrm{p.s.}(E) G(E'-E) dE' + (1-\alpha)\mathrm{Chebyshev}_6(E)$$

- U^J the production amplitude of $\Lambda_b \to J/\psi p K^-$ process with $J = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}$
- p.s.(E) the phase space
- $G(E'-E)$ Gaussian function representing experimental resolution
- $\mathrm{Chebyshev}_6(E)$ the 6th order Chebyshev polynomial for background contribution
- $1-\alpha$ the background fraction with $\alpha \in (0,1]$
- Parameter regions

$$g_S \in [0,10]\ \mathrm{GeV}^{-2} \qquad g_D \in [0.5,1.5] \times g_S \qquad C_{3/2} \in [0.5,1.5] \times C_{1/2} \qquad \mathscr{F}_1^{\frac{5}{2}} \in [600,900]$$

$$C_{1/2} \in [-20,0]\ \mathrm{GeV}^{-2} \qquad \mathscr{F}_1^{\frac{1}{2}} \in [0,300] \qquad \mathscr{F}_2^{\frac{1}{2}} \in [700,1000] \qquad \mathscr{F}_3^{\frac{1}{2}} \in [-3600,-3300]$$

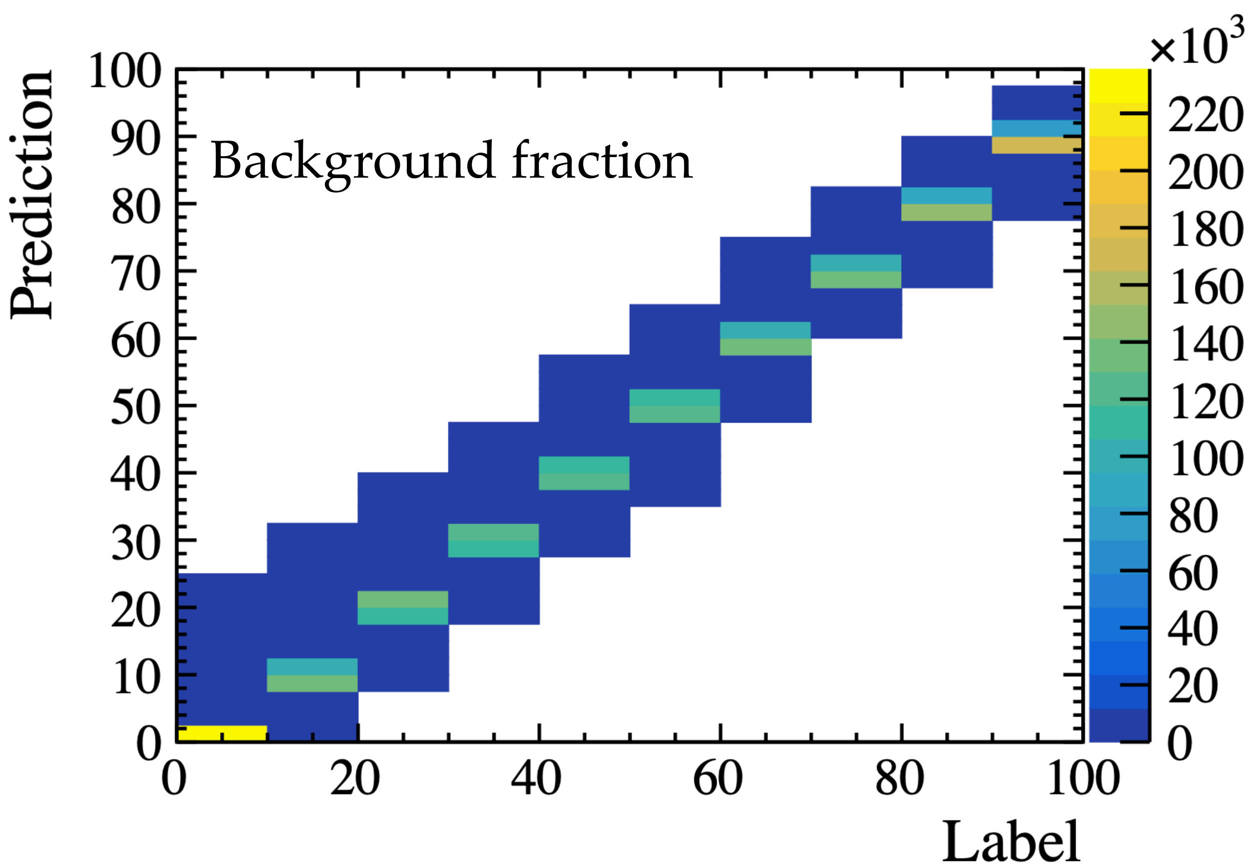
$$\mathscr{F}_1^{\frac{3}{2}} \in [-3900,-3600], \quad \mathscr{F}_2^{\frac{3}{2}} \in [-1900,-1600], \quad \mathscr{F}_3^{\frac{3}{2}} \in [-4800,-4500],$$

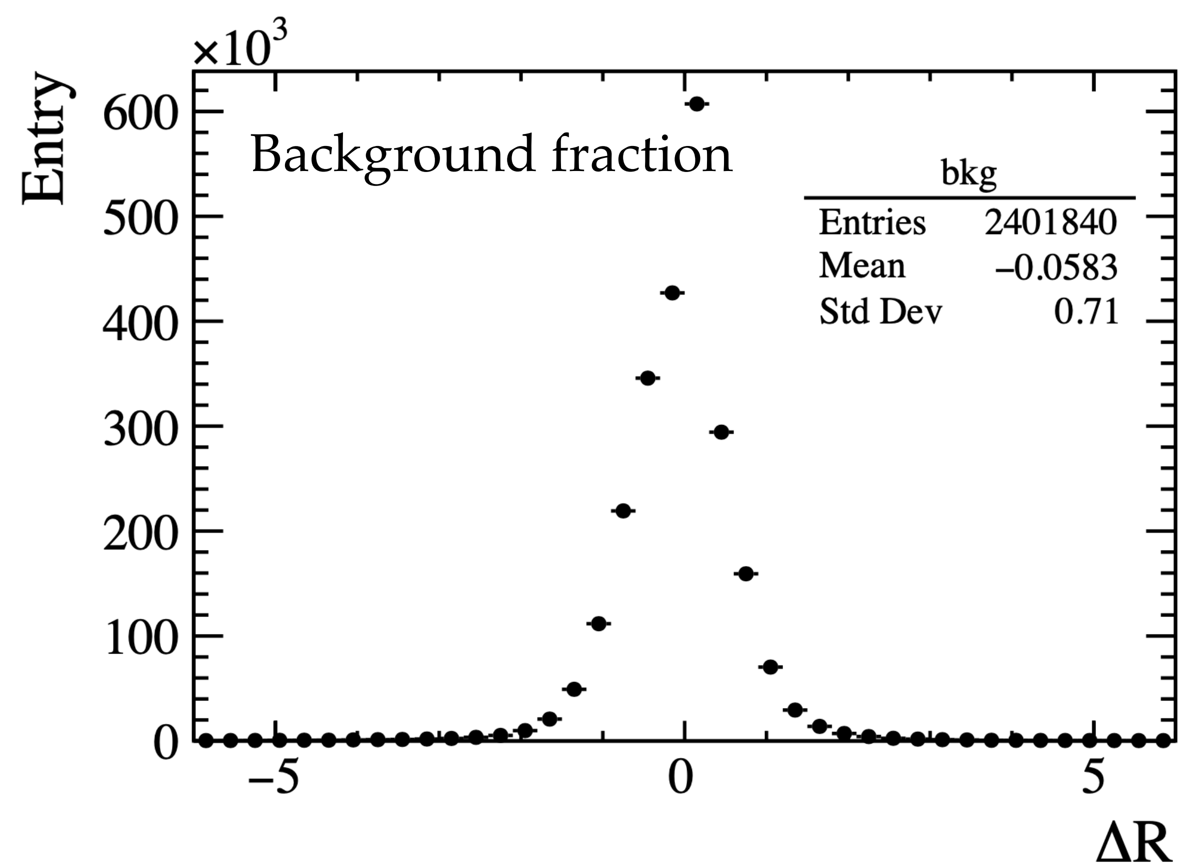
Multi-channel case

The Probability Distribution Function   Zhang, Liu, Hu, **QW,** Meißner, Sci.Bull.68(2023)981-989

$$\mathrm{PDF}(E;\mathscr{P}) = \alpha \sum_J \int |U^J|^2 \mathrm{p.s.}(E) G(E'-E) dE' + (1-\alpha)\mathrm{Chebyshev}_6(E)$$

- The samples are produced by ROOT and GSL
- Various background samples denoted as S^{90}, i.e. $1-\alpha = 90\,\%$
- $1-\alpha = (96.0 \pm 0.8)\,\%$ from a ResNet-based NN



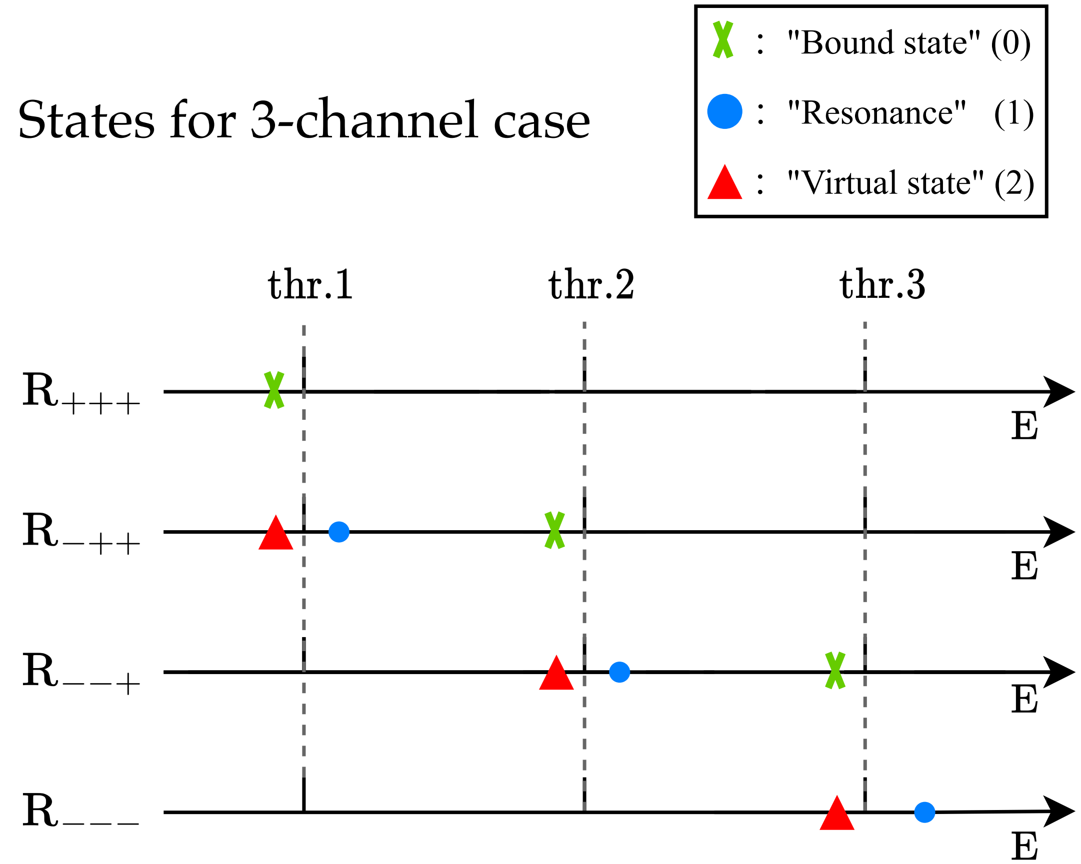


Multi-channel case

States and labels

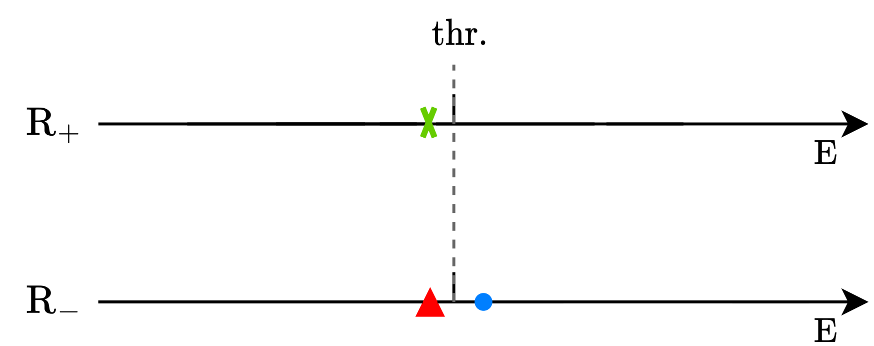
- "+" and "-" for phy. and unphy. sheets
- $\frac{1}{2}^-$ dyn. Channels: $\Sigma_c\bar{D}$, $\Sigma_c\bar{D}^*$, $\Sigma_c^*\bar{D}^*$
- $\frac{3}{2}^-$ dyn. Channels: $\Sigma_c^*\bar{D}$, $\Sigma_c\bar{D}^*$, $\Sigma_c^*\bar{D}^*$
- $\frac{5}{2}^-$ dyn. Channel: $\Sigma_c^*\bar{D}^*$

States for 3-channel case



LHCb, PRL122(2019)222001

States for 1-channel case



Bound state for $J = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}$ channels

Solution B 0000

Zhang, Liu, Hu, **QW,** Meißner, Sci.Bull.68(2023)981-989

- Mass label 1 and 0 for $J_{P_c(4440)} = \frac{1}{2}$, $J_{P_c(4457)} = \frac{3}{2}$ and $J_{P_c(4440)} = \frac{3}{2}$, $J_{P_c(4457)} = \frac{1}{2}$,

i.e. solution A and B in PRL122(2019)242001, PRL124(2020)072001,JHEP08(2021)157

Multi-channel case

Predicted probability

Training and verification

240184 samples

Mass Relation Label	State Label	Number of Samples
0	000	46951
1	000	4283
1	001	1260
1	002	4360
0	100	3740
0	110	4320
0	111	7520
1	111	360
0	200	9590
1	200	280
1	210	3980
1	211	2690
1	220	50240
1	221	50512
1	222	50098

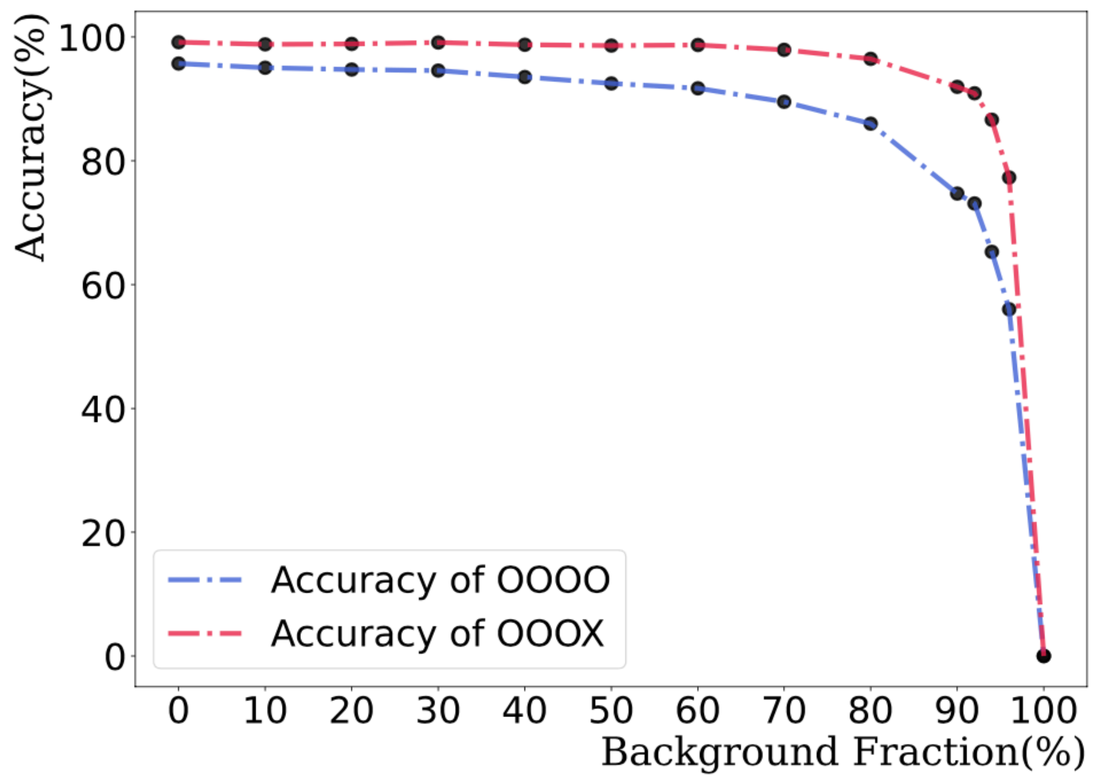
Output(%) NN \ Label	0000	1000	1001	1002	100X	others
prediction of NN trained with $\{S^{90}\}$ samples.						
NN 1	0.69	89.13	1.42	8.75	99.30	0.01
NN 2	0.03	5.83	38.47	55.30	99.60	0.37
NN 3	0.03	5.39	15.79	78.41	99.59	0.11
NN 4	0.01	1.9	27.01	70.95	99.86	0.13
NN 5	2.40	94.45	0.15	2.99	97.59	0.01
5 NNs Average	0.63(1.03)	39.34	16.57	43.28	99.19(0.91)	0.13(0.15)
10 NNs Average	0.36(0.74)	21.16	20.69	57.62	99.47(0.68)	0.12(0.13)
prediction of NN trained with $\{S^{92}\}$ samples.						
NN 1	0.00	0.15	5.37	94.47	99.99	0.00
NN 2	0.00	0.07	4.11	95.81	99.99	0.00
NN 3	0.00	0.78	13.57	85.61	99.96	0.03
NN 4	0.00	0.81	19.02	80.16	99.99	0.00
NN 5	0.14	15.13	16.91	67.80	99.84	0.00
5 NNs Average	0.03(0.06)	3.39	11.80	84.77	99.95(0.06)	0.01(0.01)
10 NNs Average	0.01(0.04)	1.78	9.50	88.70	99.97(0.04)	0.00(0.01)

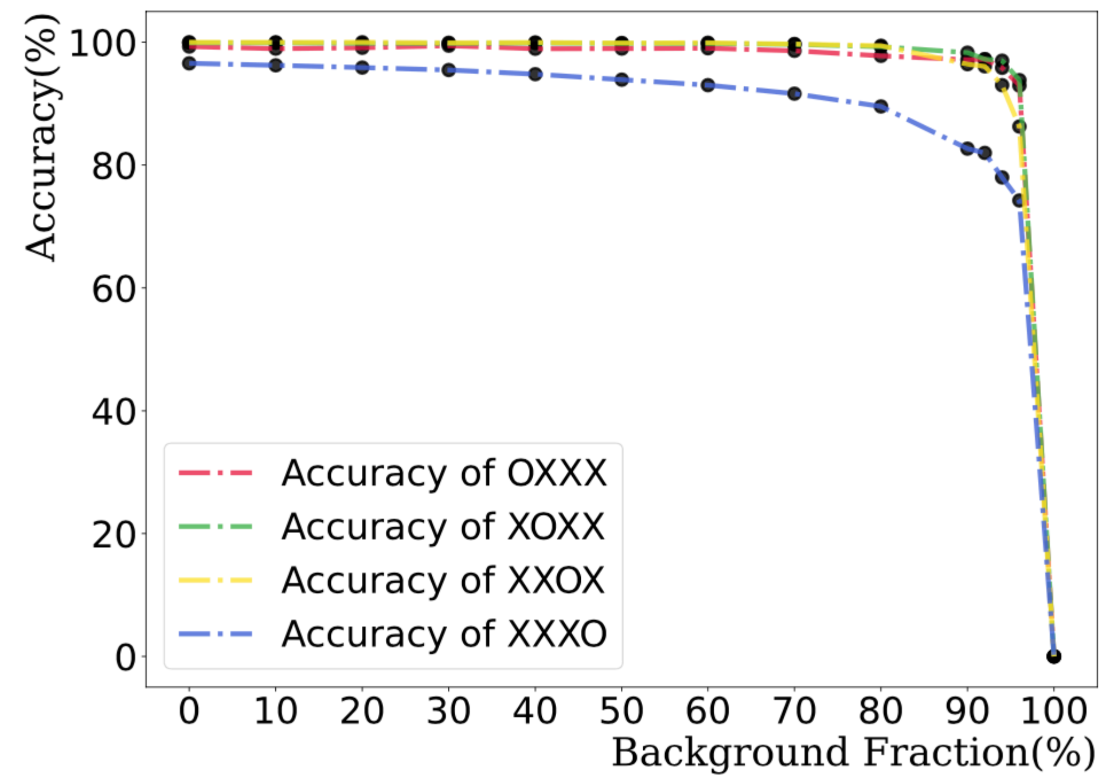
- 5 and 10 NN models with an identical structure under different initialization
- The uncertainties decrease with the increasing number of NNs
- Top 3 probabilities, 1000,1001,1002 favor solution A
- Bound states in $J^P = \frac{1}{2}^-, \frac{3}{2}^-$ channels, Undetermined for $J^P = \frac{5}{2}^-$ channel
- The NNs successfully retrieve the state label with an accuracy (standard deviation) of $75.91(1.18)\,\%$,$73.14(1.05)\,\%$,$65.25(1.80)\,\%$,$54.35(2.32)\,\%$ for the samples $\{\mathcal{S}^{90}\}, \{\mathcal{S}^{92}\}, \{\mathcal{S}^{94}\}, \{\mathcal{S}^{96}\}$

Zhang, Liu, Hu, **QW,** Meißner, Sci.Bull.68(2023)981-989

Multi-channel case

The accuracy of NNs





- "O" opens the label. "X" close the label.
- Accuracy decreases with the increasing background fraction
- The lower accuracy is also because of the $\frac{5}{2}^-$ channel

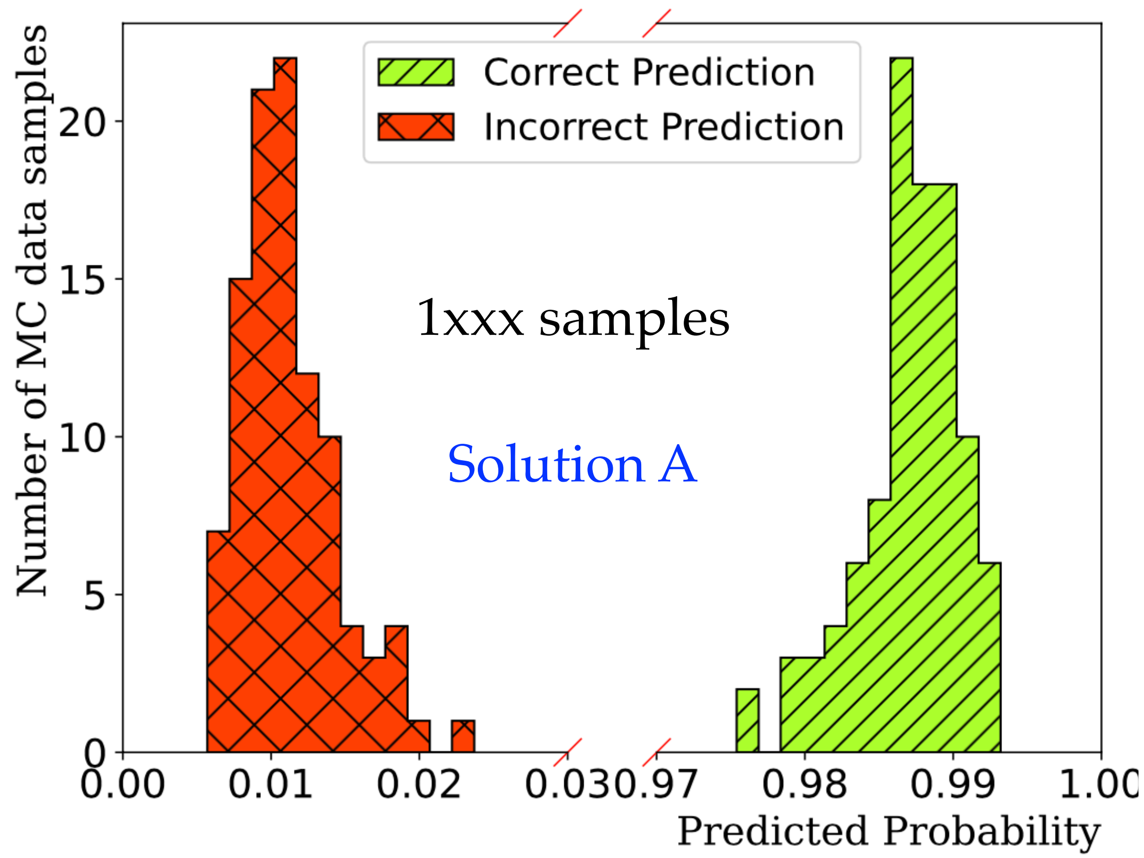
Zhang, Liu, Hu, **QW,** Meißner, Sci.Bull.68(2023)981-989

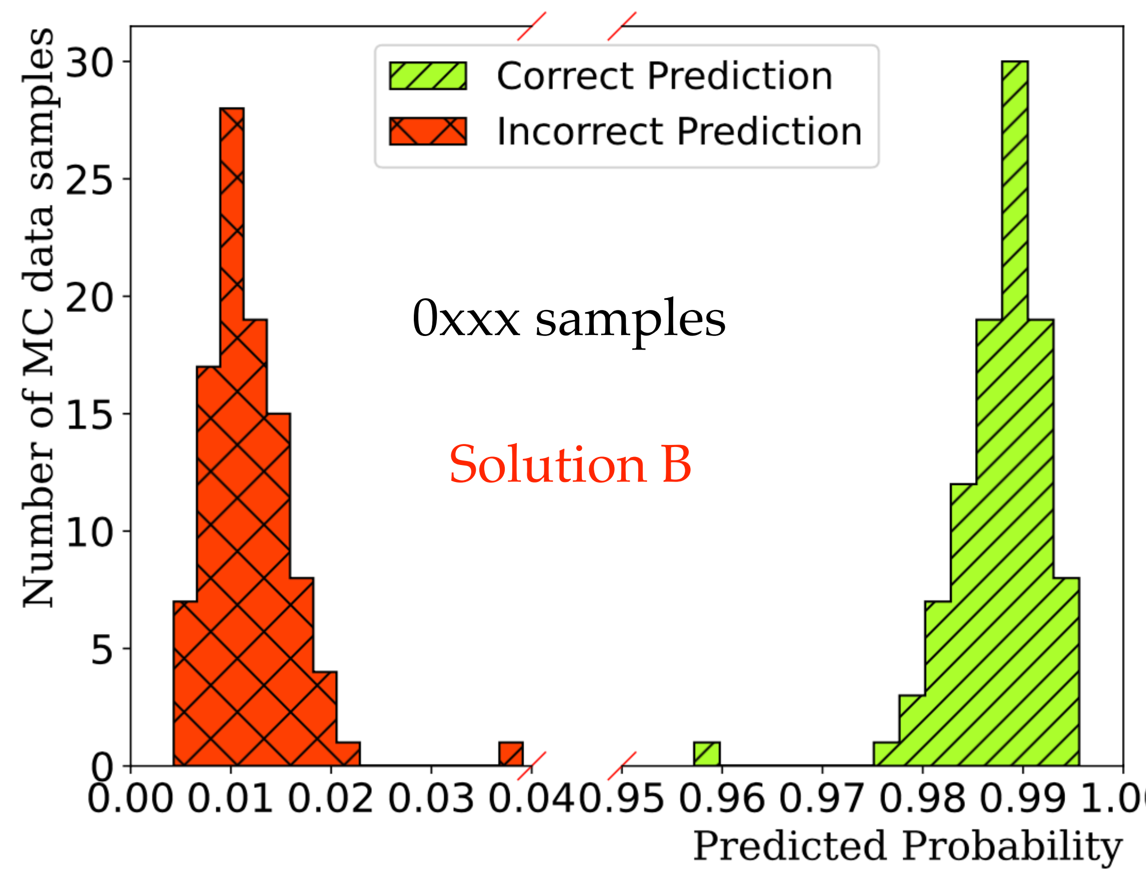
Multi-channel case

Why NN favors Solution A?

Generate 100 1xxx samples and 100 0xxx samples

Probabilities from NN





- The NN can make a good prediction
- The two solutions are well distinguished for both samples

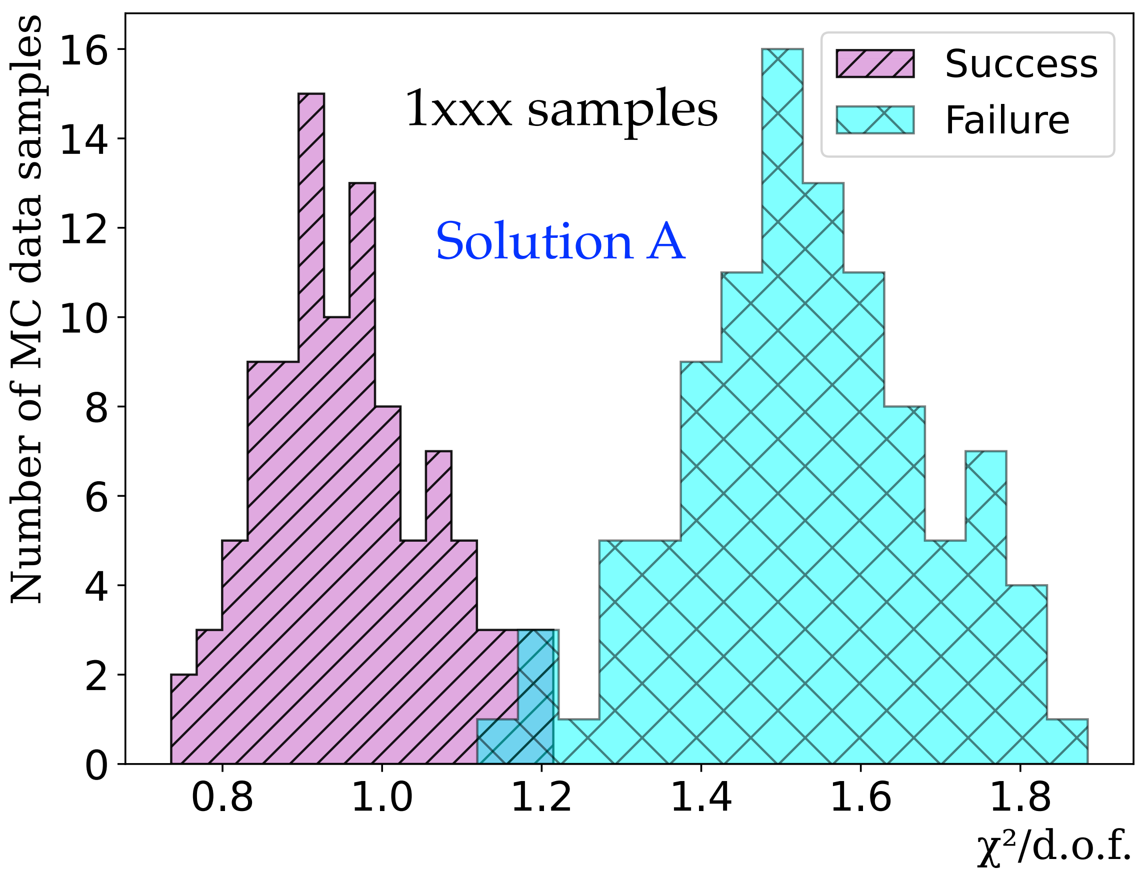
Zhang, Liu, Hu, QW, Meißner, Sci.Bull.68(2023)981-989

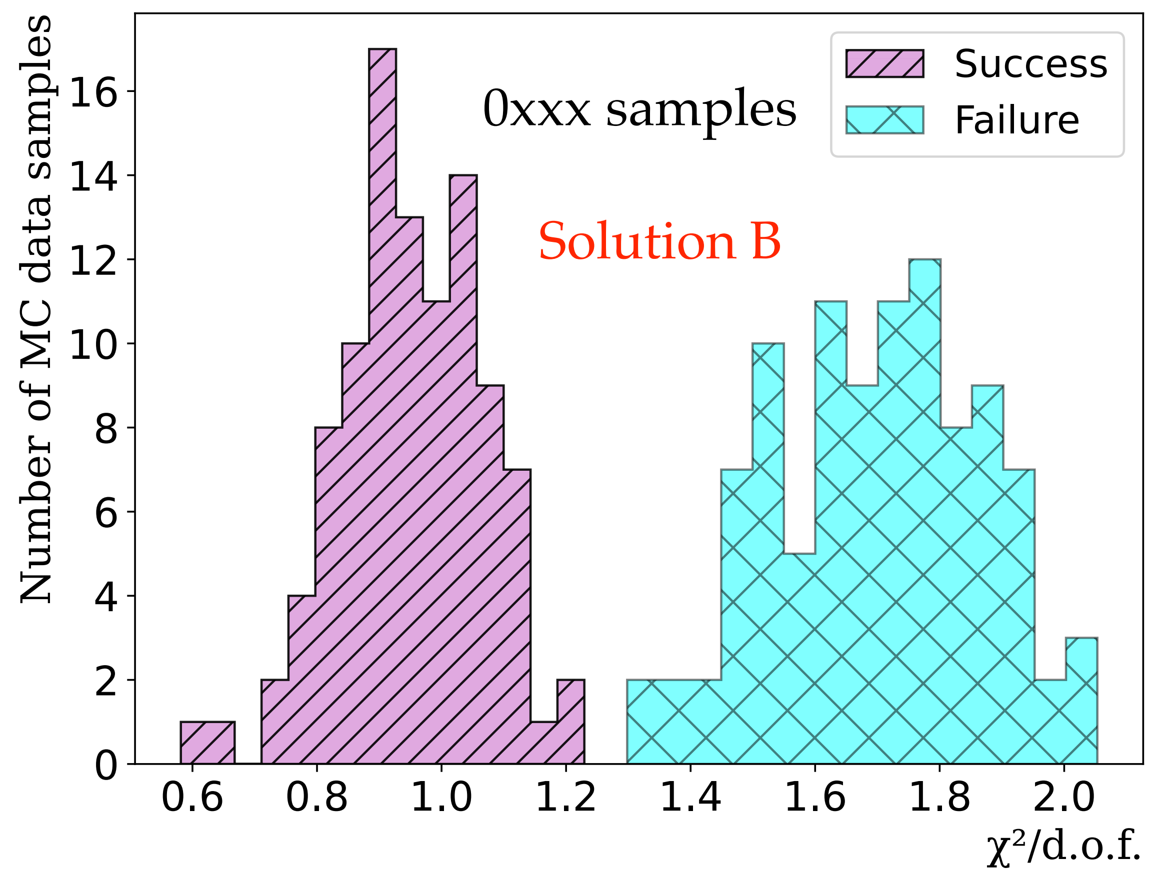
Multi-channel case

Why NN favors Solution A?

Generate 100 1xxx samples and 100 0xxx samples

Reduced chisq from the normal fitting



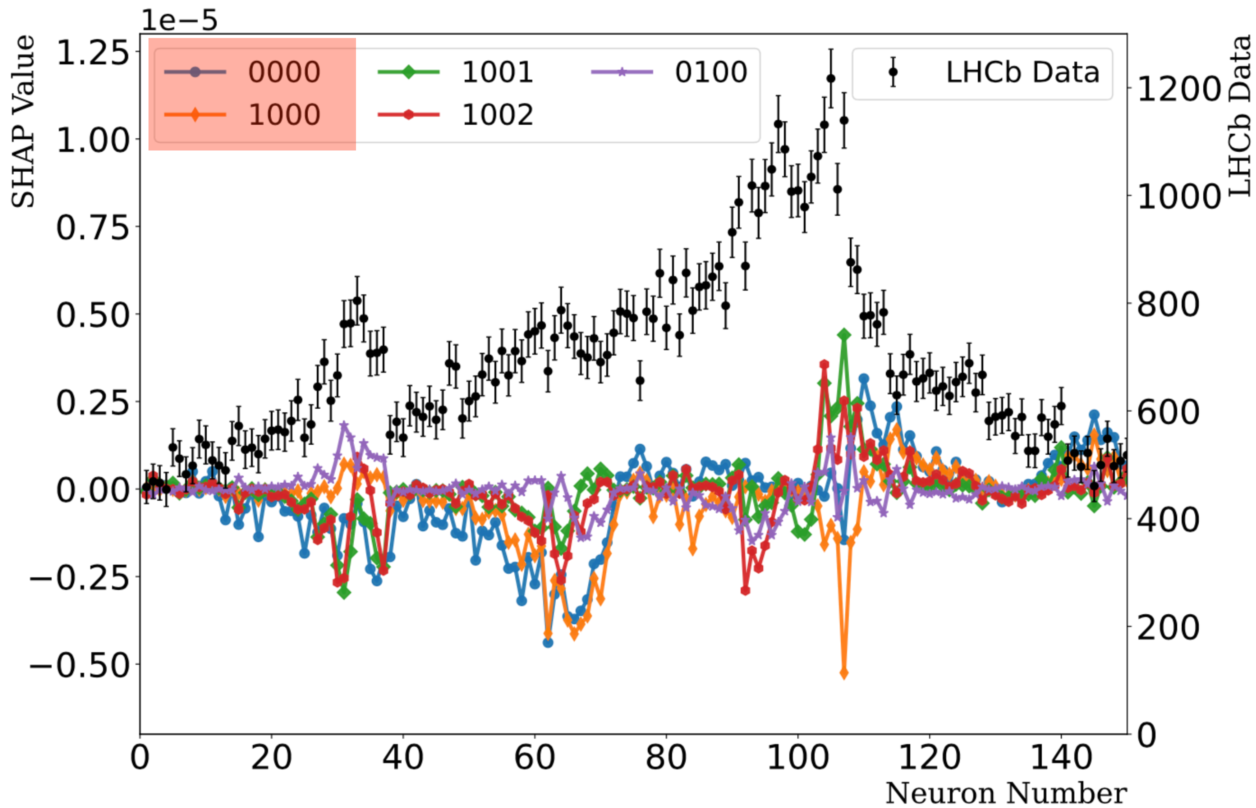


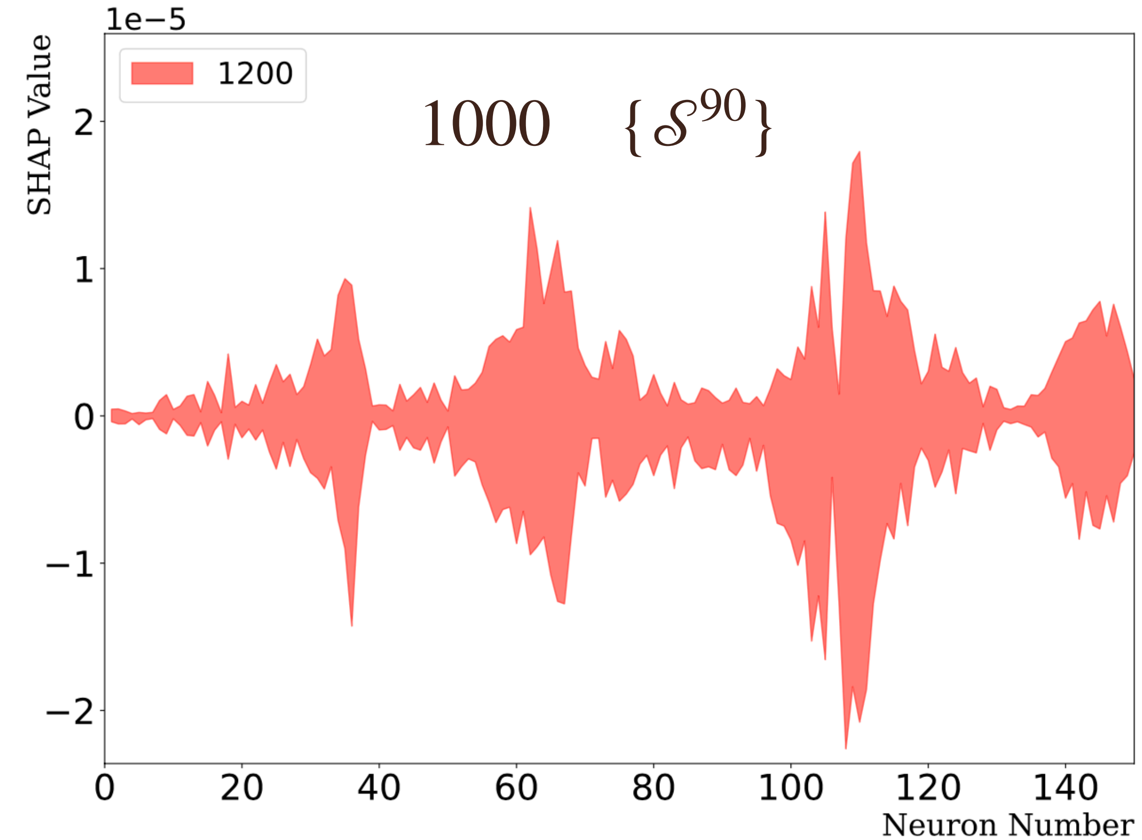
- A 3% misidentification for 1xxx samples

Zhang, Liu, Hu, **QW,** Meißner, Sci.Bull.68(2023)981-989

Multi-channel case

The impact of each experimental data point in NN





- The Shapley Additive exPlanation (SHAP) is investigated.
- A positive (negative) SHAP value indicates that a given data point is pushing the NN classification in favor of (against) a given class.
- The data points around the thresholds in the mass spectrum have a greater impact.

Zhang, Liu, Hu, **QW,** Meißner, Sci.Bull.68(2023)981-989

Multi-channel case

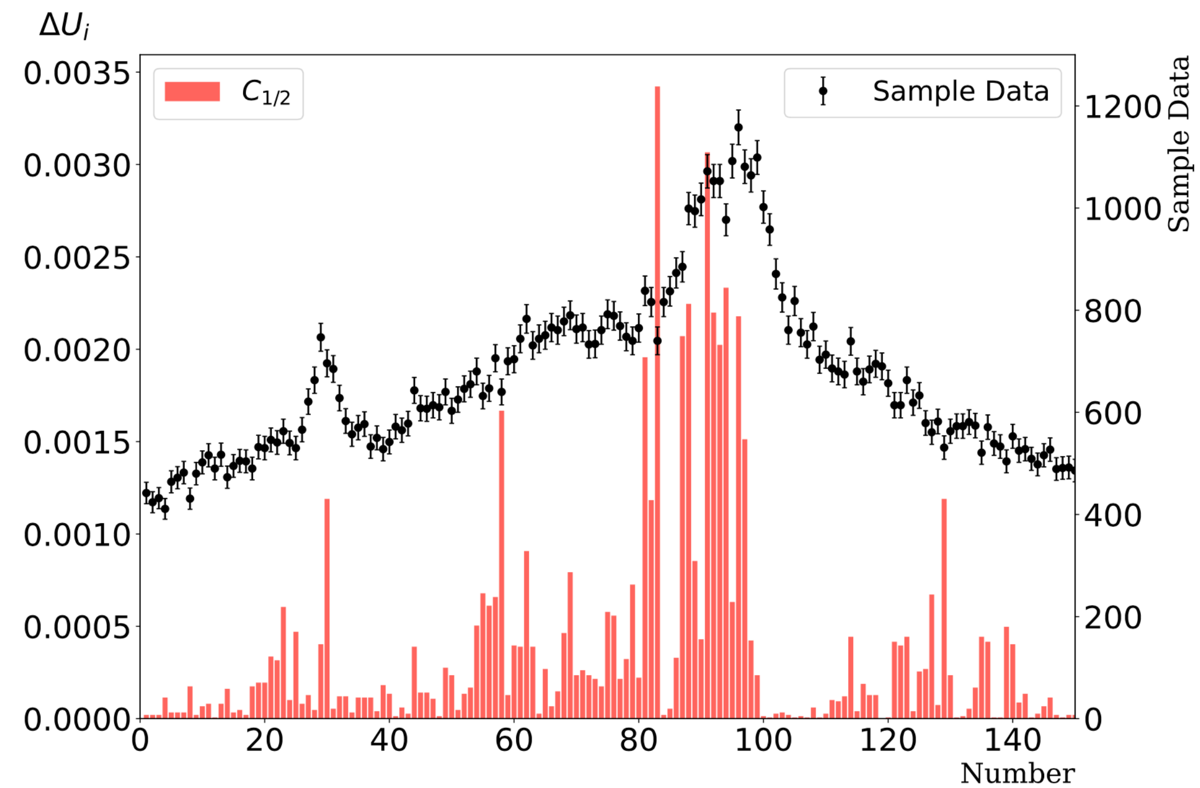
The impact of each experimental data point

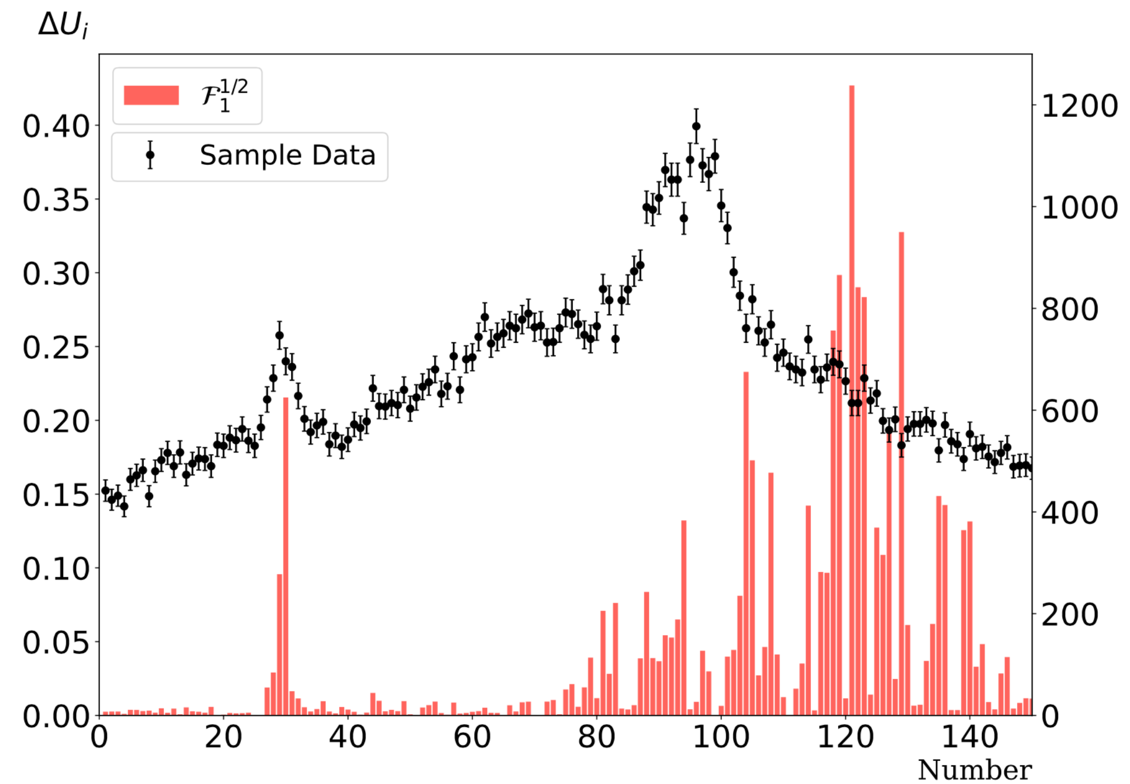
in normal fitting

A analogous quantity

$$\Delta U_i \equiv \left| \frac{\mathscr{P}_i(\text{on}) - \mathscr{P}_i(\text{off})}{\mathscr{P}_i(\text{on})} \right| \text{for the ith para.}$$

- The bins near threshold do not show strong constraints on parameters due to the large correlation among the parameters
- Data at higher energy have large constraints on the production parameters





Zhang, Liu, Hu, **QW,** Meißner, Sci.Bull.68(2023)981-989

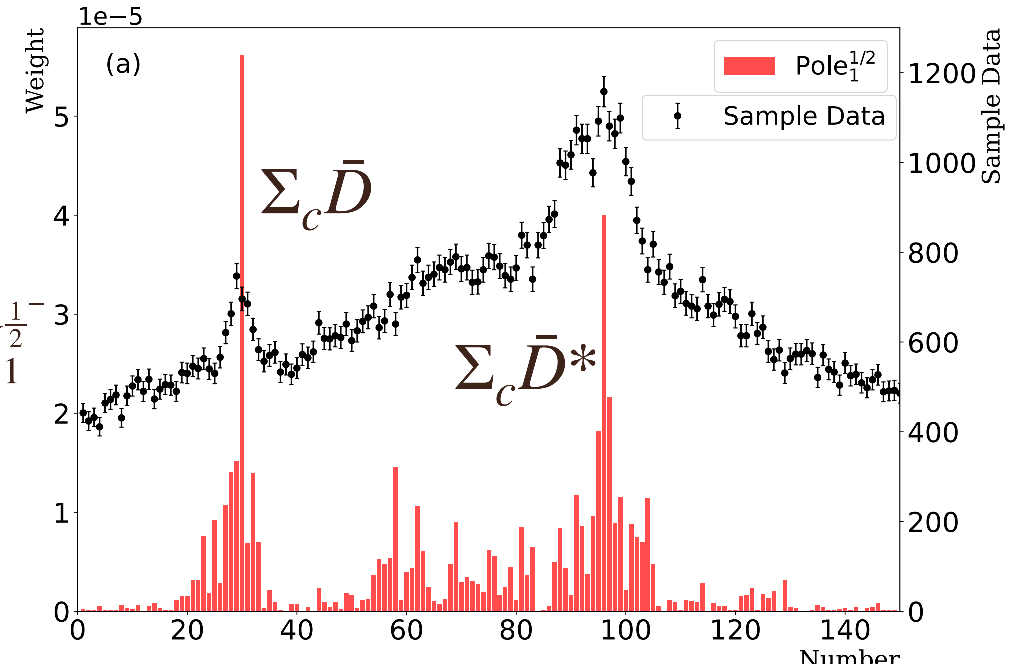
Multi-channel case

Zhang, Liu, Hu, **QW,** Meißner, Sci.Bull.68(2023)981-989

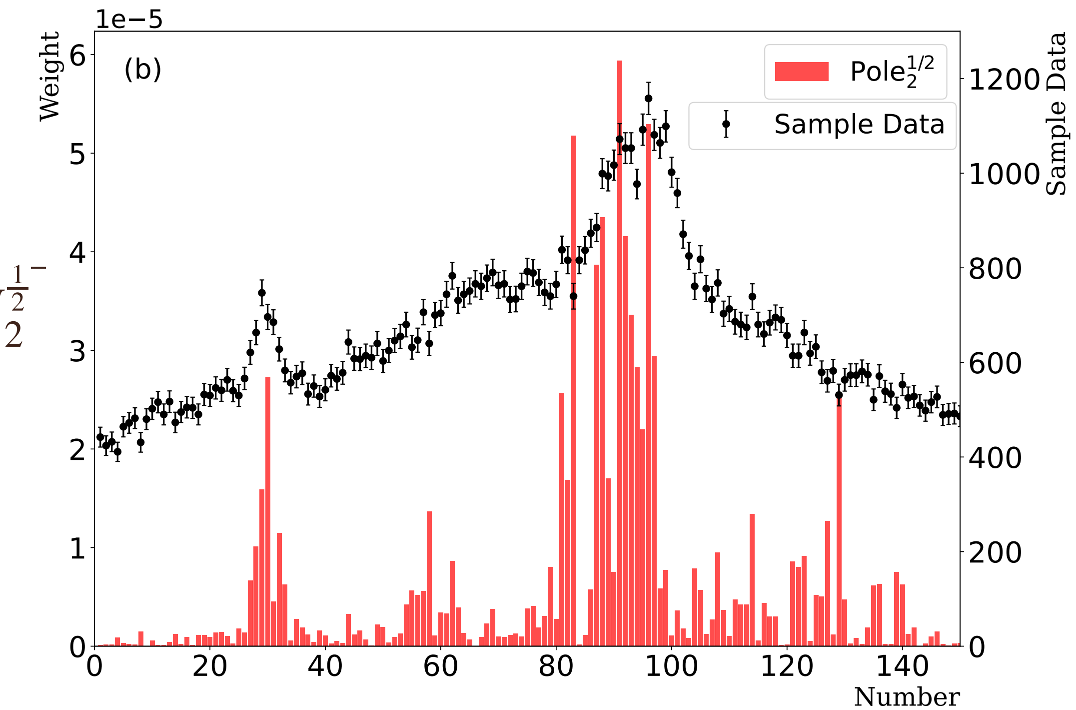
The impact of each experimental data point in normal fitting

Another analogous quantity $\Delta U_i^{J^P} \equiv \left| \dfrac{\mathrm{Re[Pole}_i](\mathrm{on}) - \mathrm{Re[Pole}_i](\mathrm{off})}{\mathrm{Re[Pole}_i](\mathrm{on})} \right|$ for the ith pole

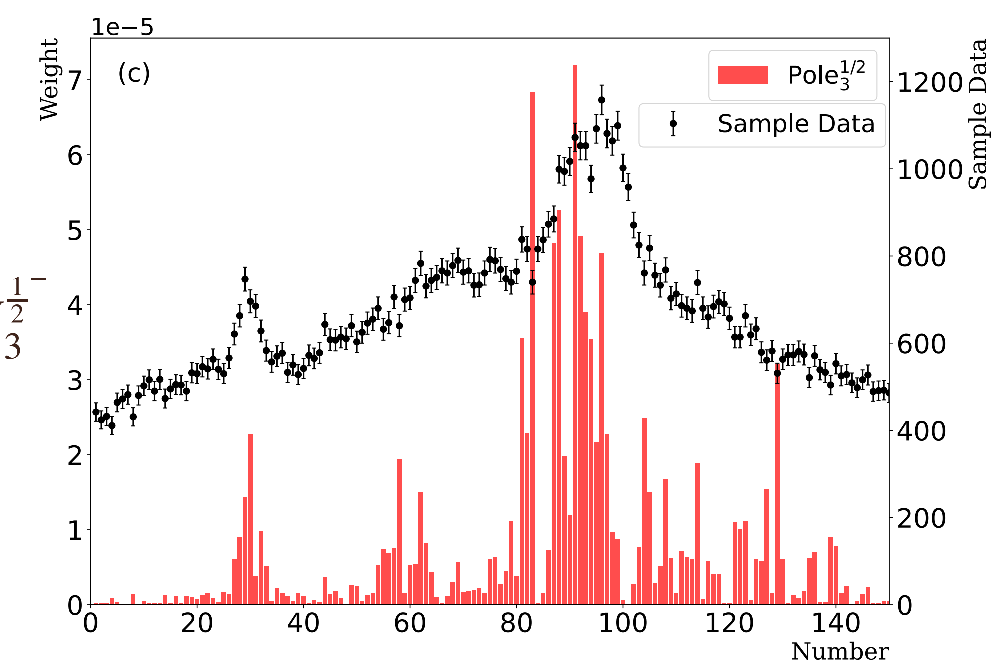
$\Delta U_1^{\frac{1}{2}^-}$



$\Delta U_2^{\frac{1}{2}^-}$



$\Delta U_3^{\frac{1}{2}^-}$



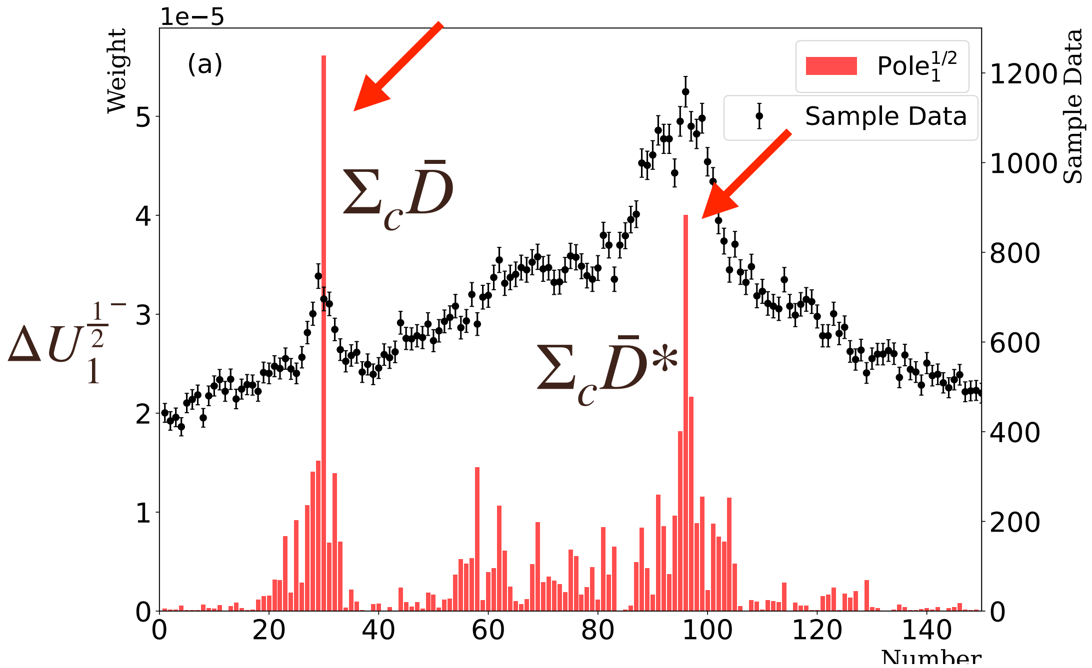
- The data around the $\Sigma_c\bar{D}$, $\Sigma_c\bar{D}^*$ thresholds are more important
- The data around the $\Sigma_c^*\bar{D}^*$ threshold are not important, due to its small production rate

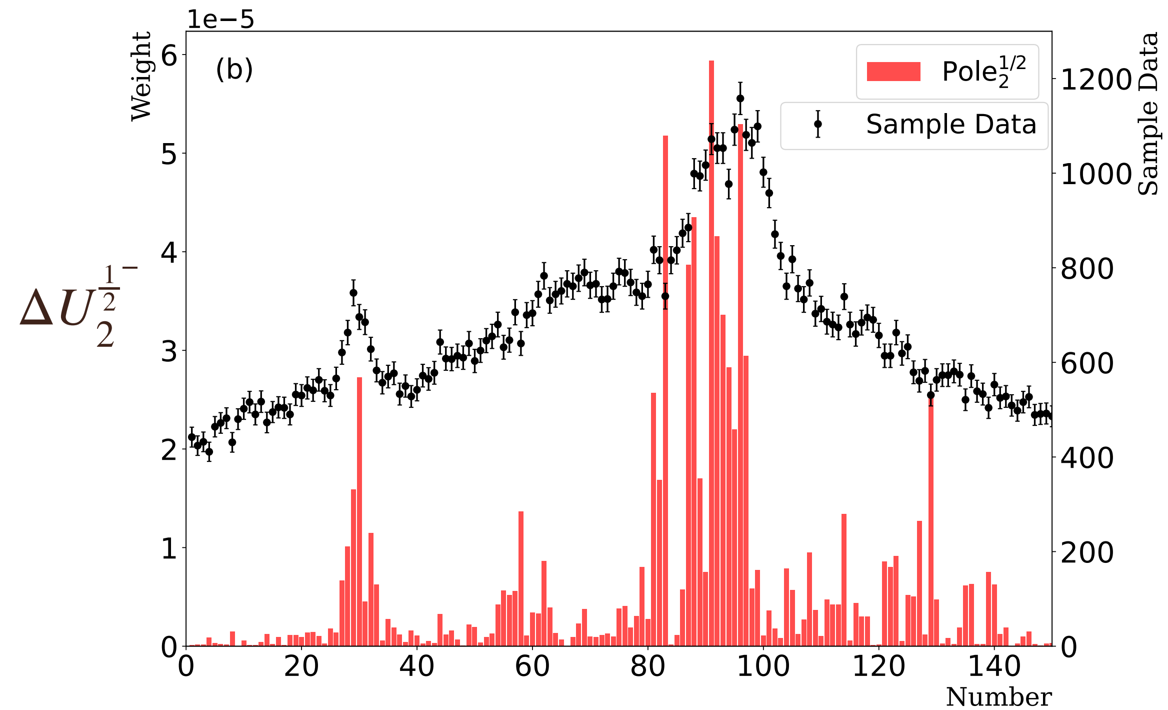
Multi-channel case

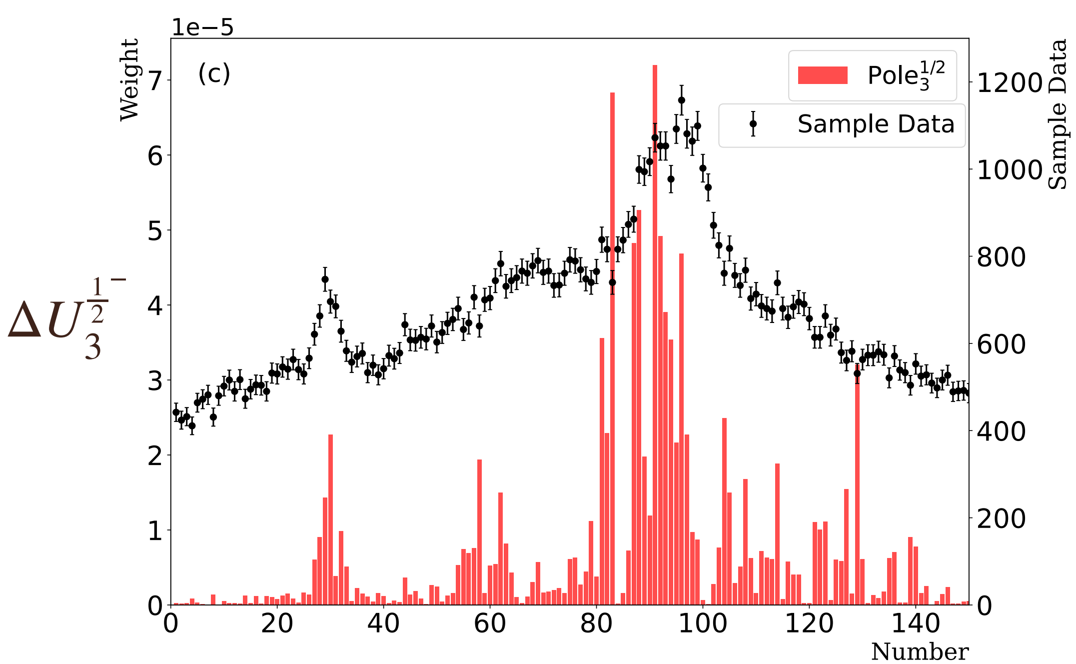
Zhang, Liu, Hu, **QW,** Meißner, Sci.Bull.68(2023)981-989

The impact of each experimental data point in normal fitting

Another analogous quantity $\Delta U_i^{J^P} \equiv \left| \dfrac{\mathrm{Re[Pole}_i\mathrm{](on)} - \mathrm{Re[Pole}_i\mathrm{](off)}}{\mathrm{Re[Pole}_i\mathrm{](on)}} \right|$ for the ith pole







- The experimental data around the coupled-channels still have strong constraints on the physics

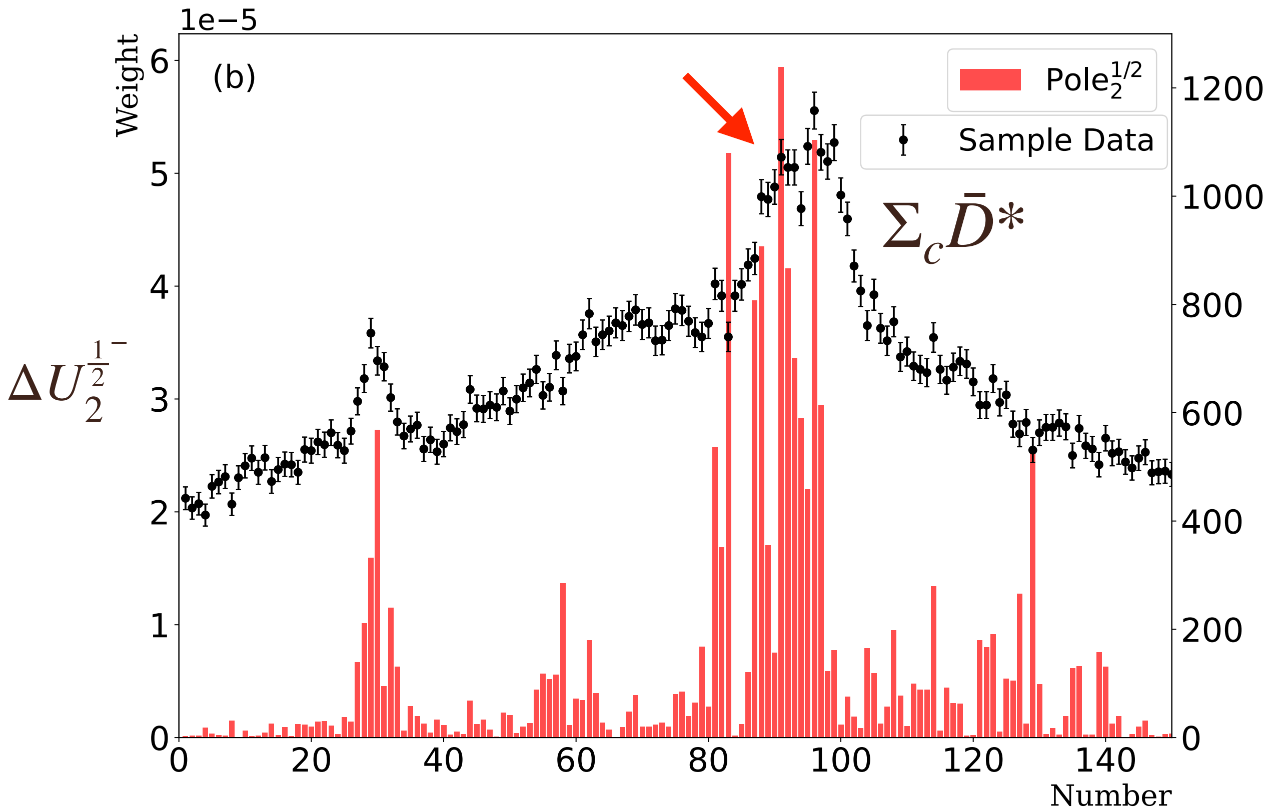
Multi-channel case

Zhang, Liu, Hu, **QW,** Meißner, Sci.Bull.68(2023)981-989

The impact of each experimental data point in normal fitting

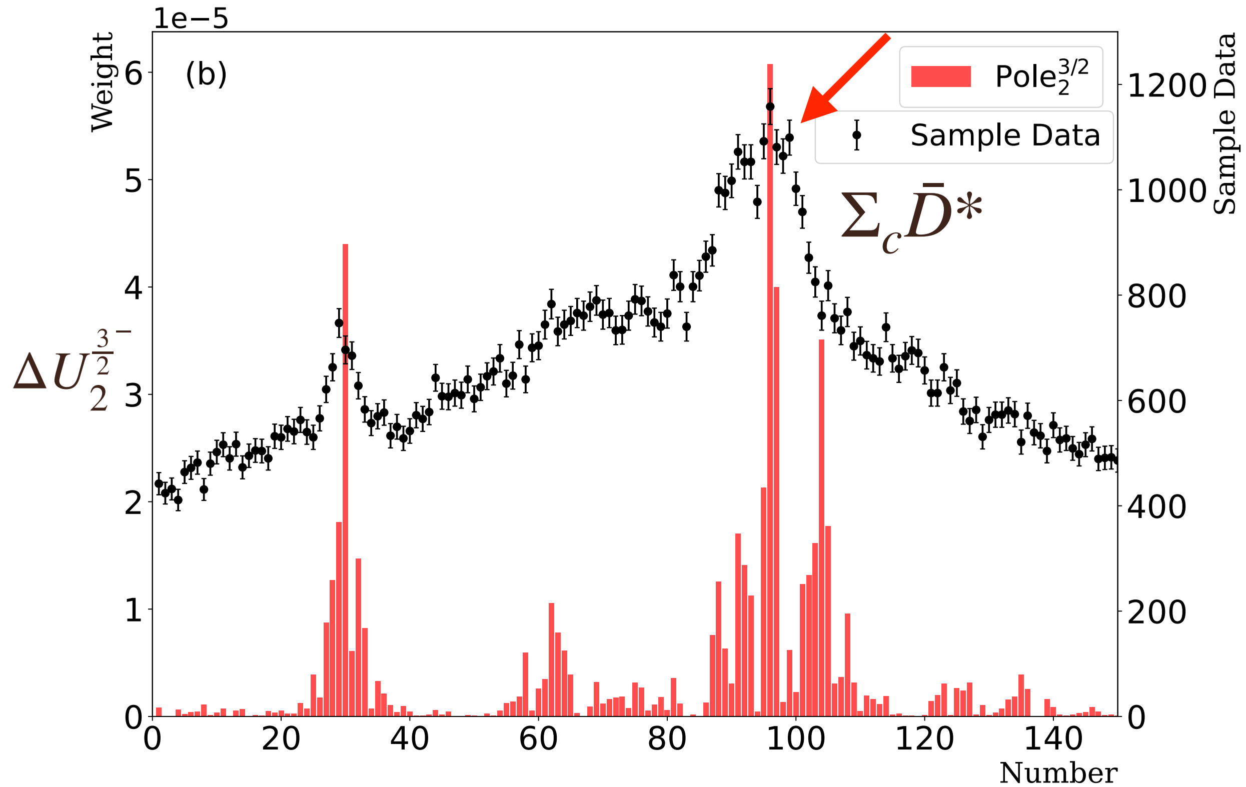
Another analogous quantity $\Delta U_i^{J^P} \equiv \left| \frac{\mathrm{Re[Pole}_i](\mathrm{on}) - \mathrm{Re[Pole}_i](\mathrm{off})}{\mathrm{Re[Pole}_i](\mathrm{on})} \right|$ for the ith pole

The 2nd pole in $J^P = \frac{1}{2}^-$



data around Pc(4440) have large constraints

The 2nd pole in $J^P = \frac{3}{2}^-$



data around Pc(4457) have large constraints

- The sample corresponds to Solution A

Summary and outlook

- We provide more insights about how the NN-based approach predicts the nature of exotics from the mass spectrum.
- Regress parameters for one-channel case
- Our NN-based approach favors Solution A in LO HQEFT
- Poles in the $J = \frac{1}{2}, \frac{3}{2}$ channels behave as bound states
- In the NN-based approach, the role of each data bin on the underlying physics is well reflected by the SHAP value. For the normal fitting, such a direct relation does not exist.

Thank you very much for your attention!