Machine learning on exotic hadrons

王倩
South China Normal University
qianwang@m.scnu.edu.cn

强子物理在线论坛100期特别活动，

合肥，中科大，2024年6月30日-7月4日

Outline

- Motivation
- The status of pentaquarks
- Framework
- Results
- Summary and outlook

Motivation

- 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……

- High energy nuclear physics

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

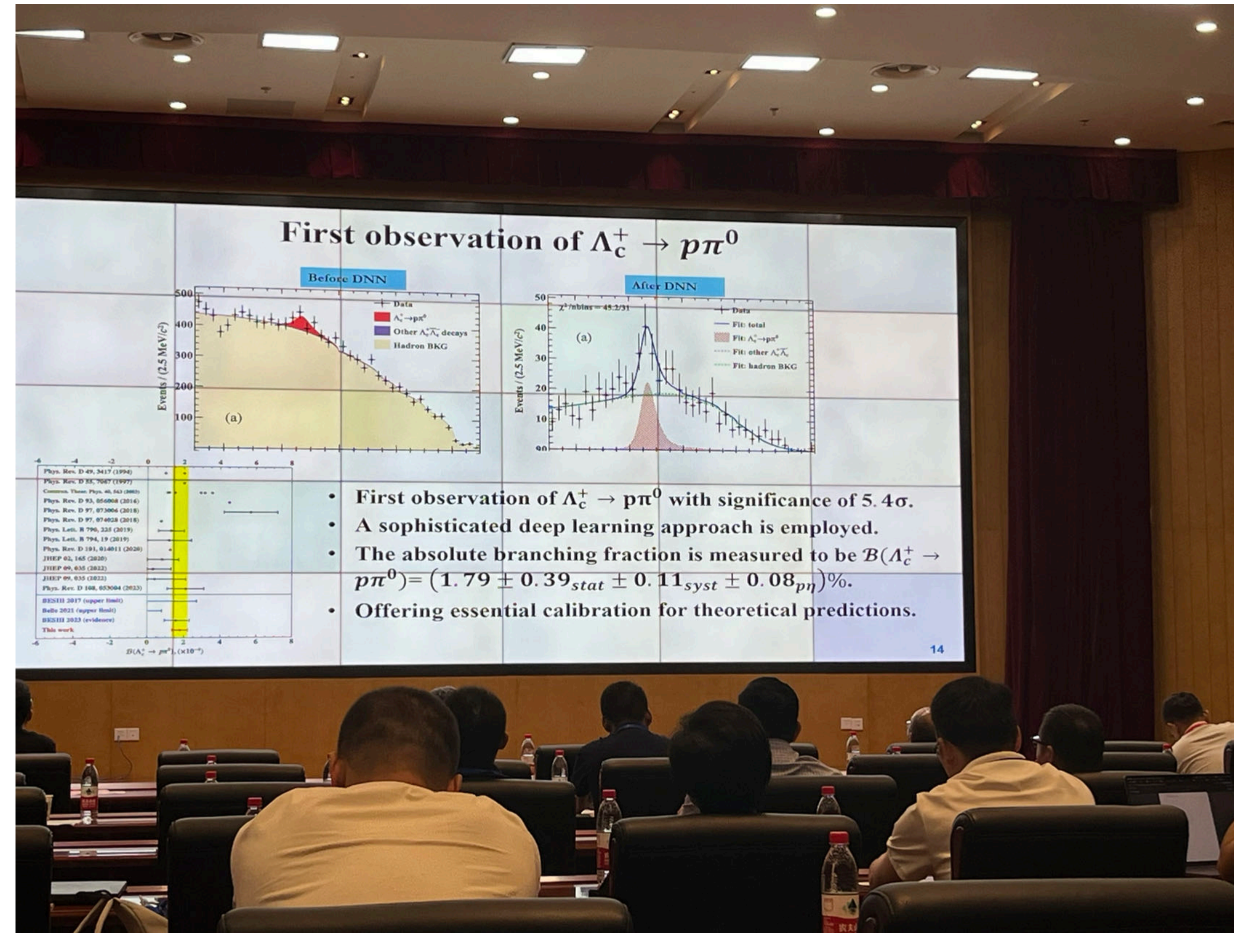
Boehnlein et.al., RMP94(2022)031003……

- Experimental data analysis

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

- Theoretical physics

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

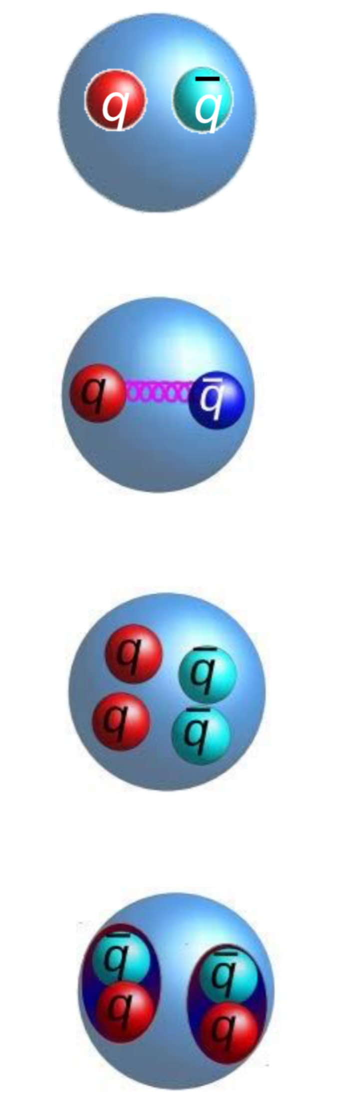


Yang-Heng Zheng's talk

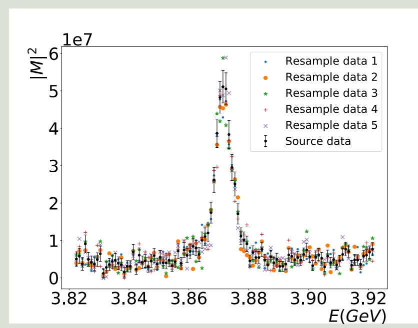
Motivation

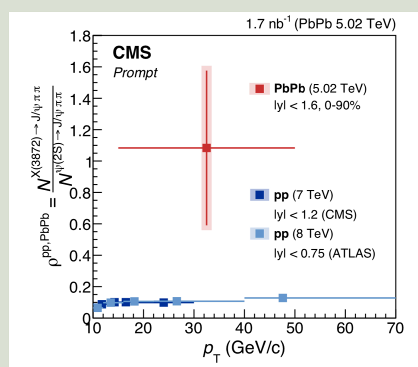
Standard approach to analyze experimental data

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

The status of pentaquarks

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

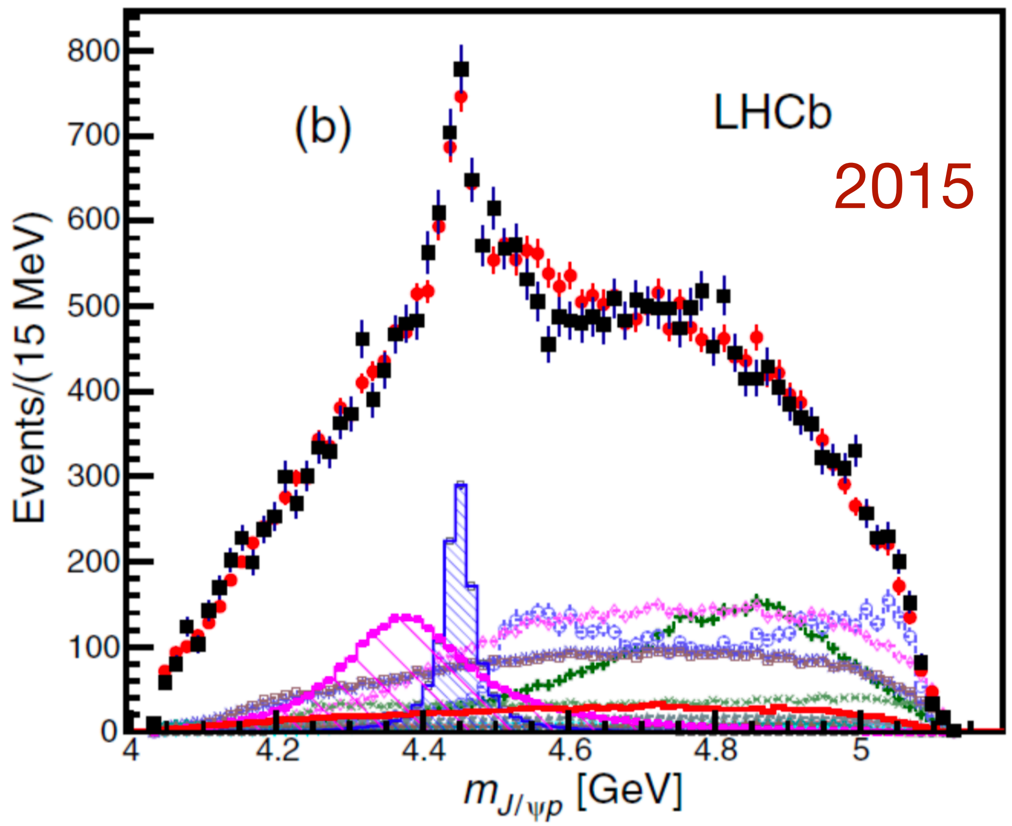
Wu, Molina, Oset, **Zou**, PRL105(2010)232001

(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

The status of pentaquarks

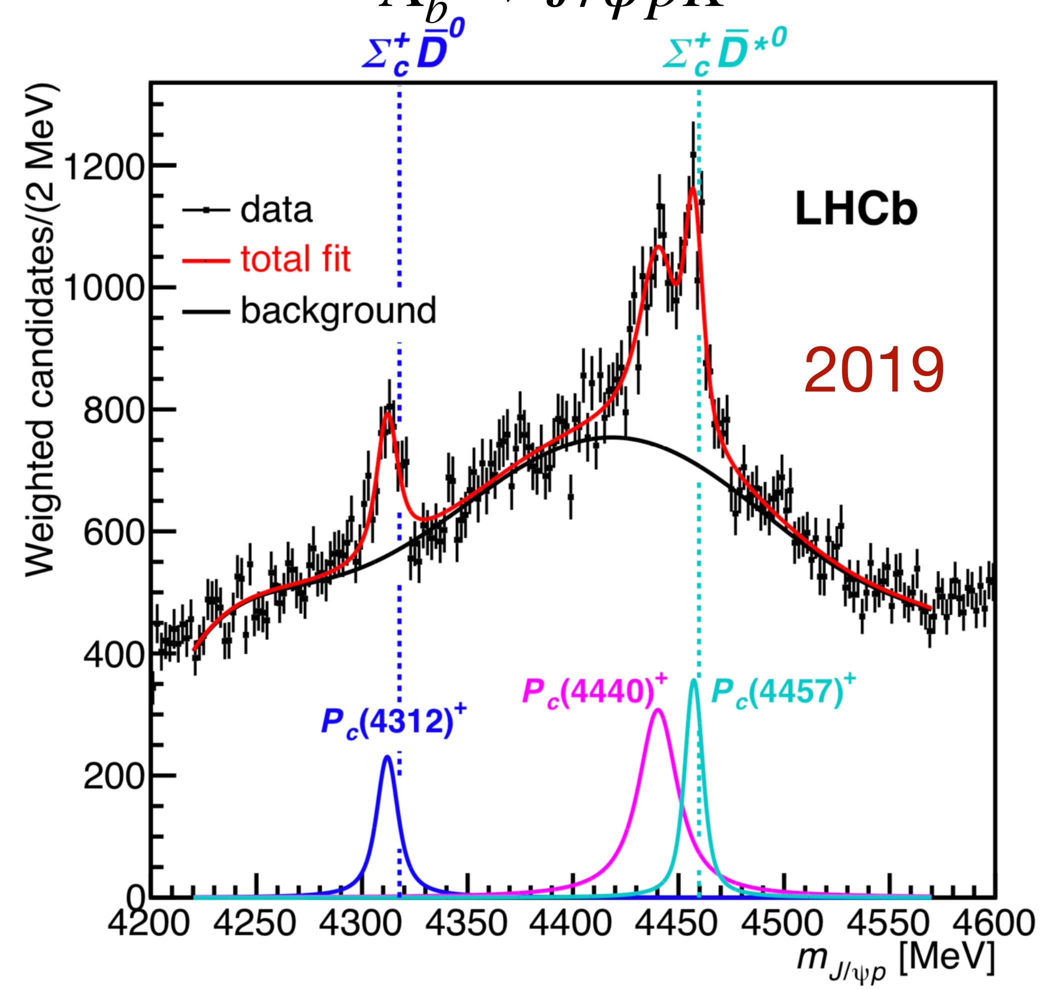
The observation of hidden charm pentaquarks



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

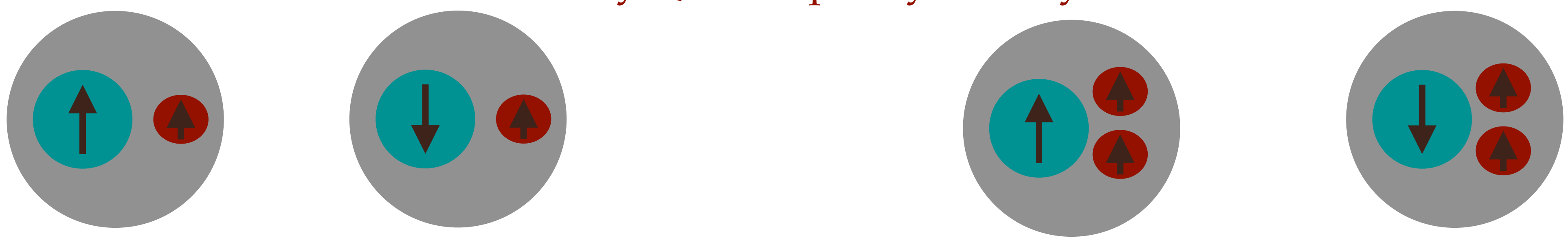
The status of pentaquarks

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



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

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

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

Spin rearrangement

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

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

Two LECs to LO

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

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

The status of pentaquarks

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

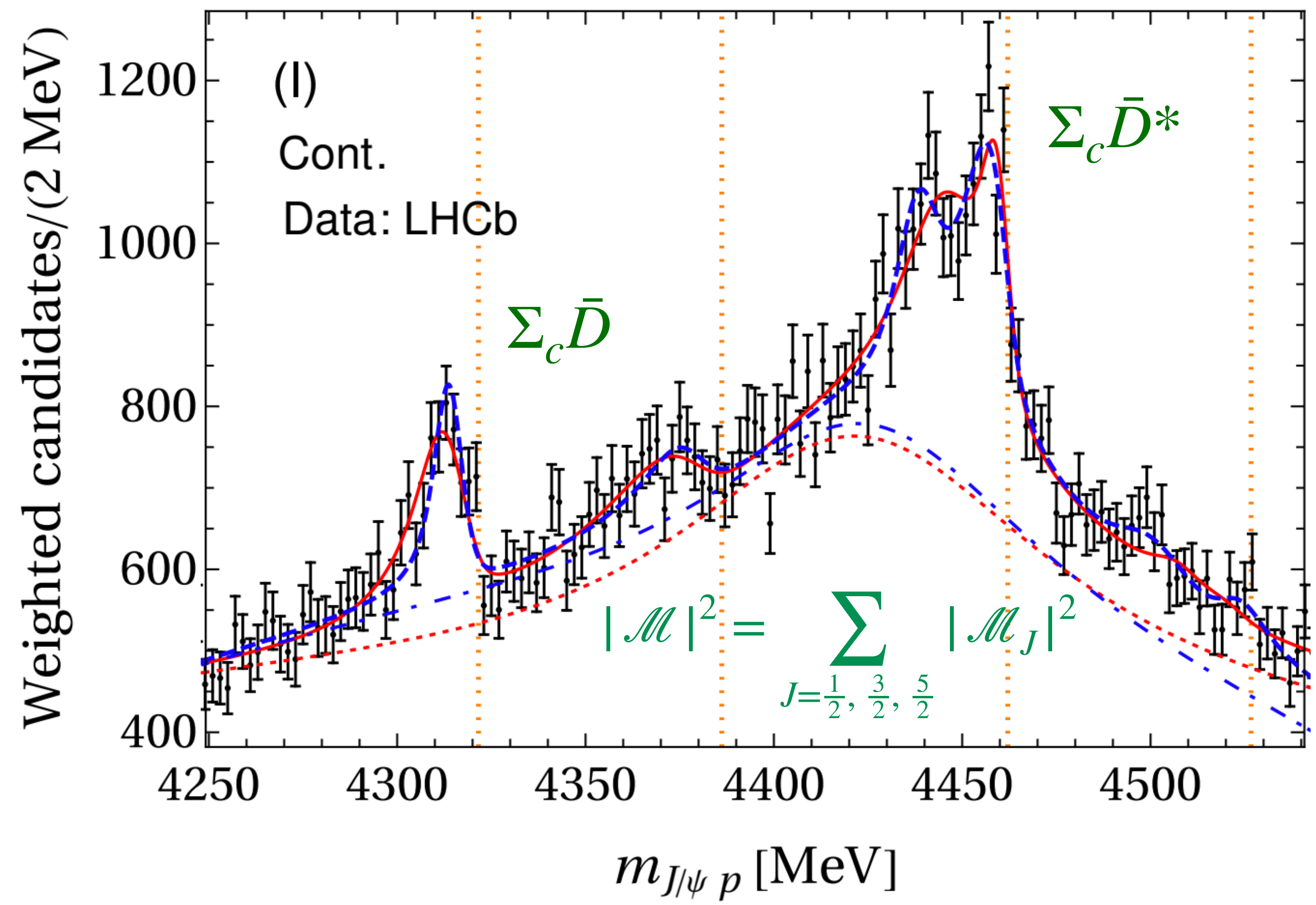
- Two parameters determined by

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

- 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
- $\chi^2_A < \chi^2_B$
- The effect of each data point is different

The status of pentaquarks

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

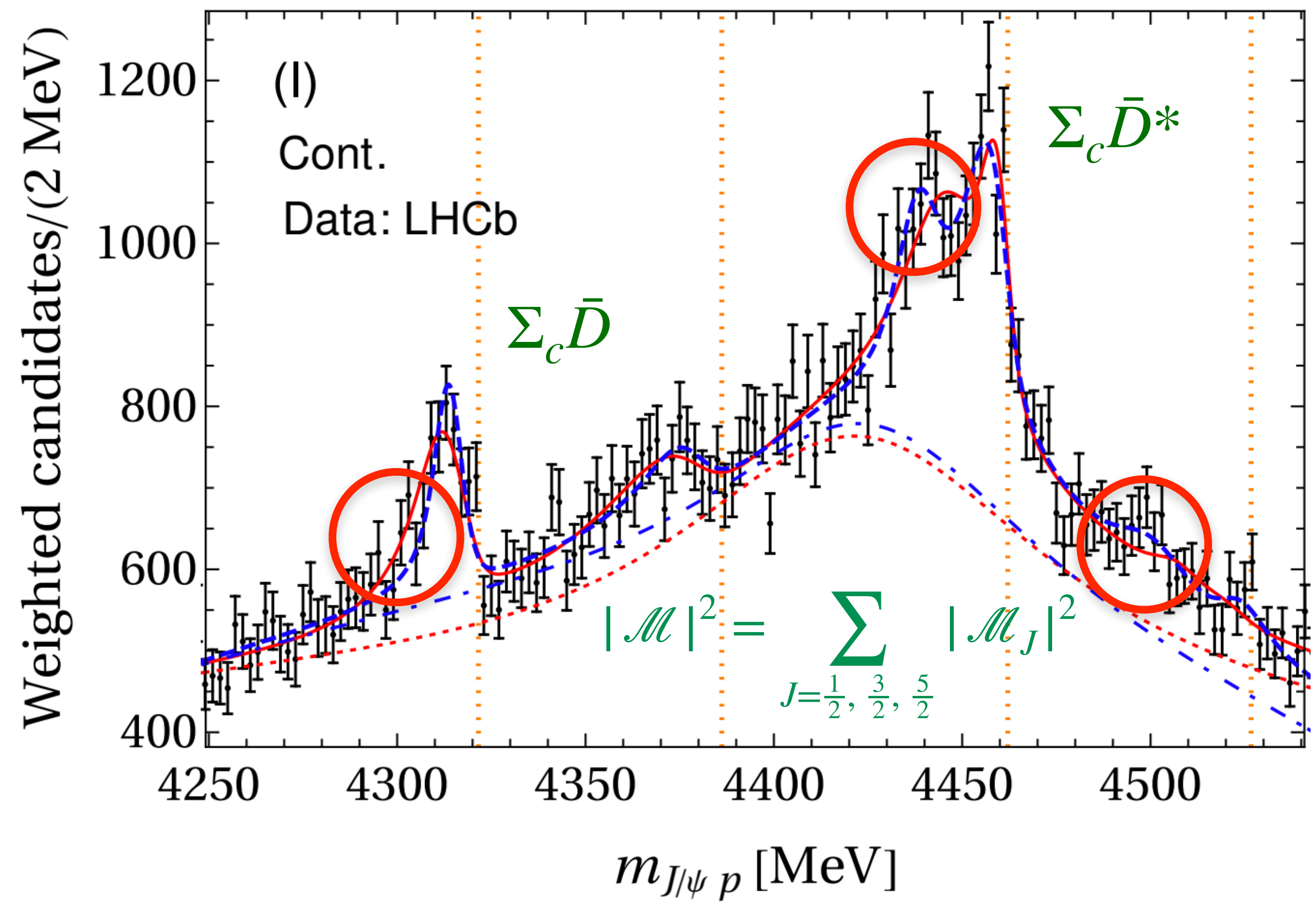
- Two parameters determined by

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

- Two solutions

Solution A (χ^2/d.o.f. = 1.01)

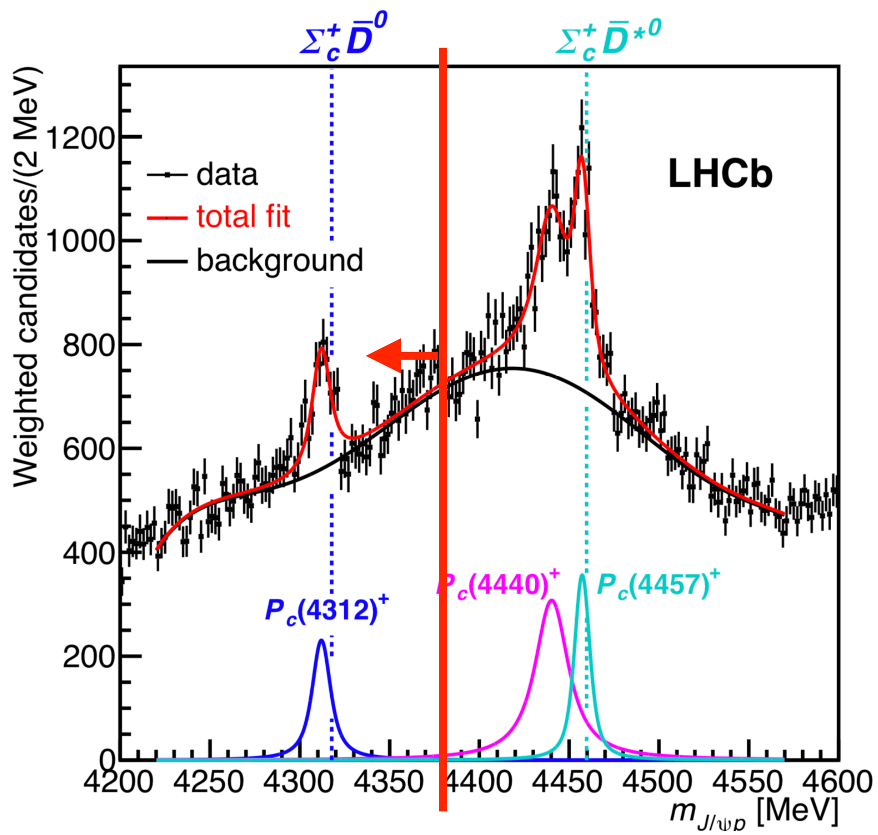
Solution B (χ^2/d.o.f. = 1.03)

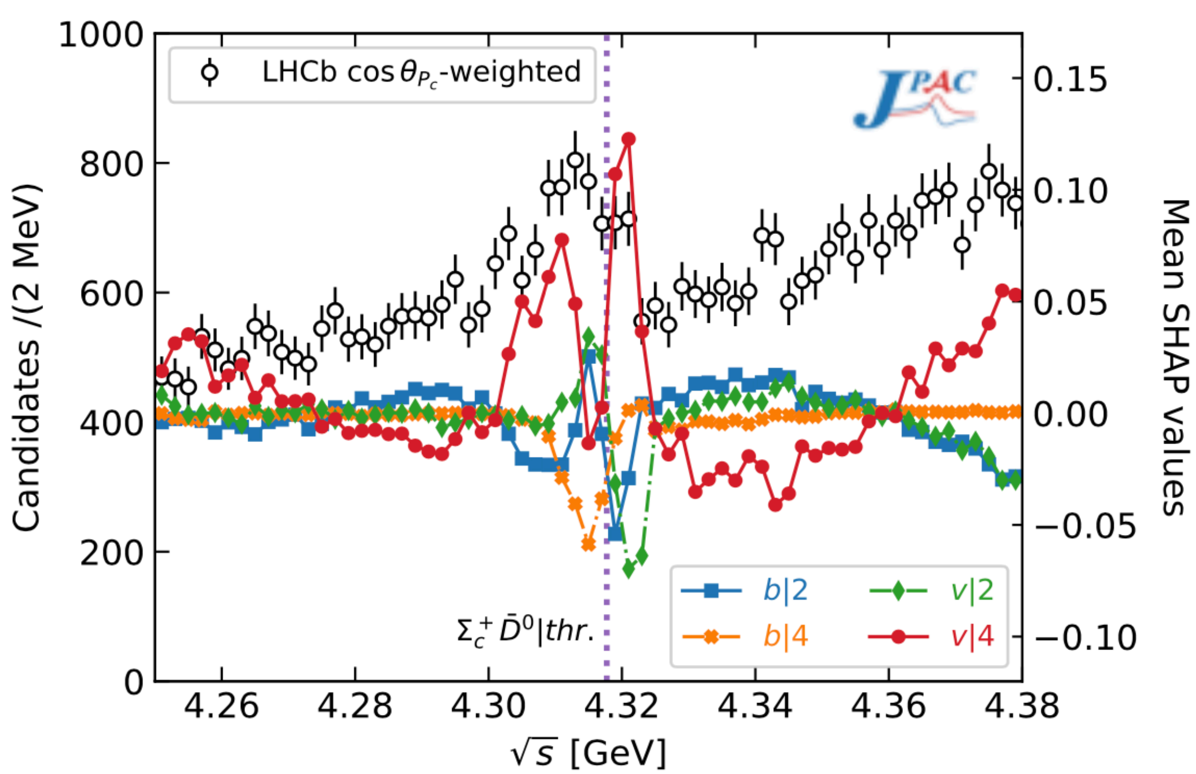


- Two parameters g_S, g_D for $J/\psi p$, $\eta_c p$
- Predict pole positions accurately
- $\chi_A^2 < \chi_B^2$
- The effect of each data point is different

The status of pentaquarks

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

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

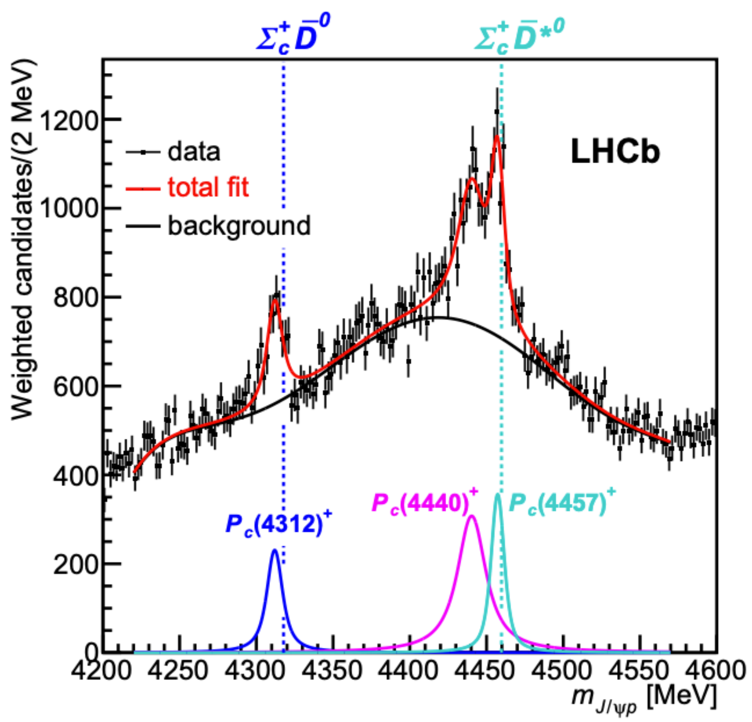
$$T(s) = \frac{m_{22} - ik_2}{(m_{11} - ik_1)(m_{22} - 9k_2) - m_{12}^2}$$

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

JPAC, PRD105(2022)L091501

Framework

LHCb, PRL122(2019)222001

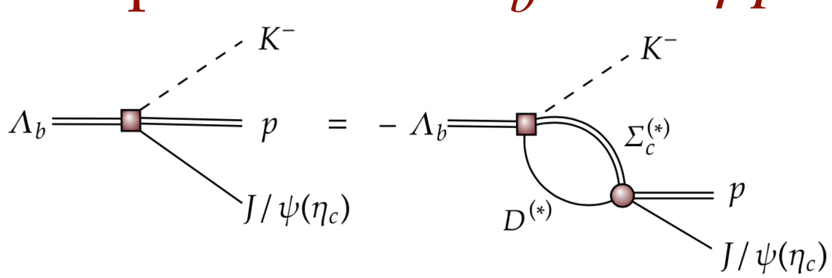


The $\Sigma_c^{(*)}\bar{D}^{(*)}$ molecular picture

- $P_c(4312)$ 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?

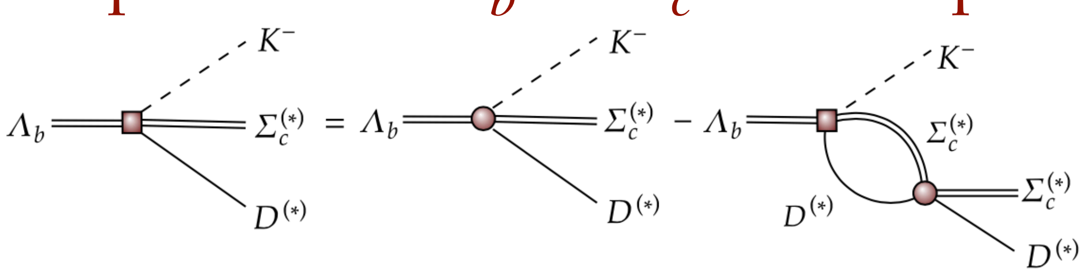
LO HQEFT, **Du**, **Baru, Guo, Hanhart, Meißner**, Oller, **QW,** PRL124(2020)072001

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



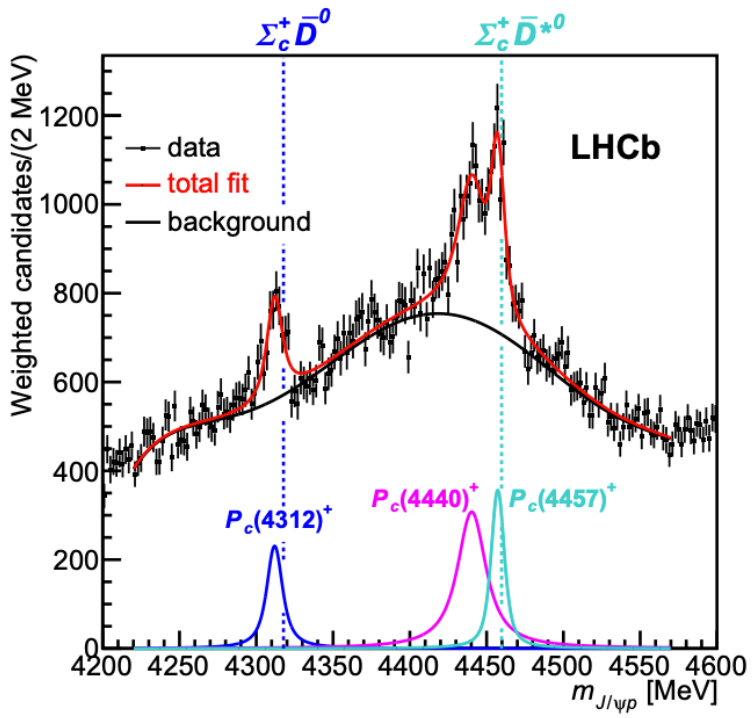
Zhang, Liu, Hu, **QW, Meißner**, Sci.Bull.68(2023)981-989

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



Framework

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

Framework

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

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

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

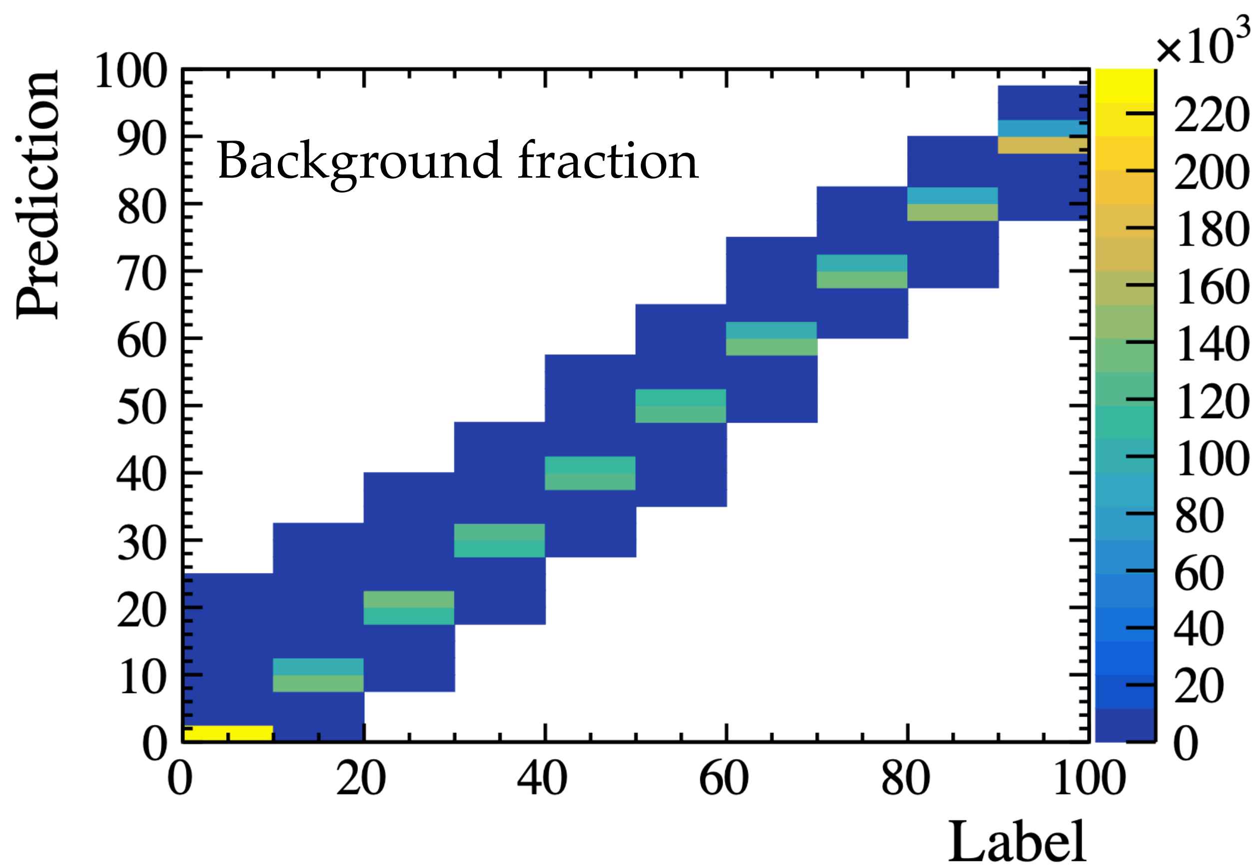
$$\mathscr{F}_1^{\frac{3}{2}} \in [-3900,-3600], \quad \mathscr{F}_2^{\frac{3}{2}} \in [-1900,-1600], \quad \mathscr{F}_3^{\frac{3}{2}} \in [-4800,-4500],$$

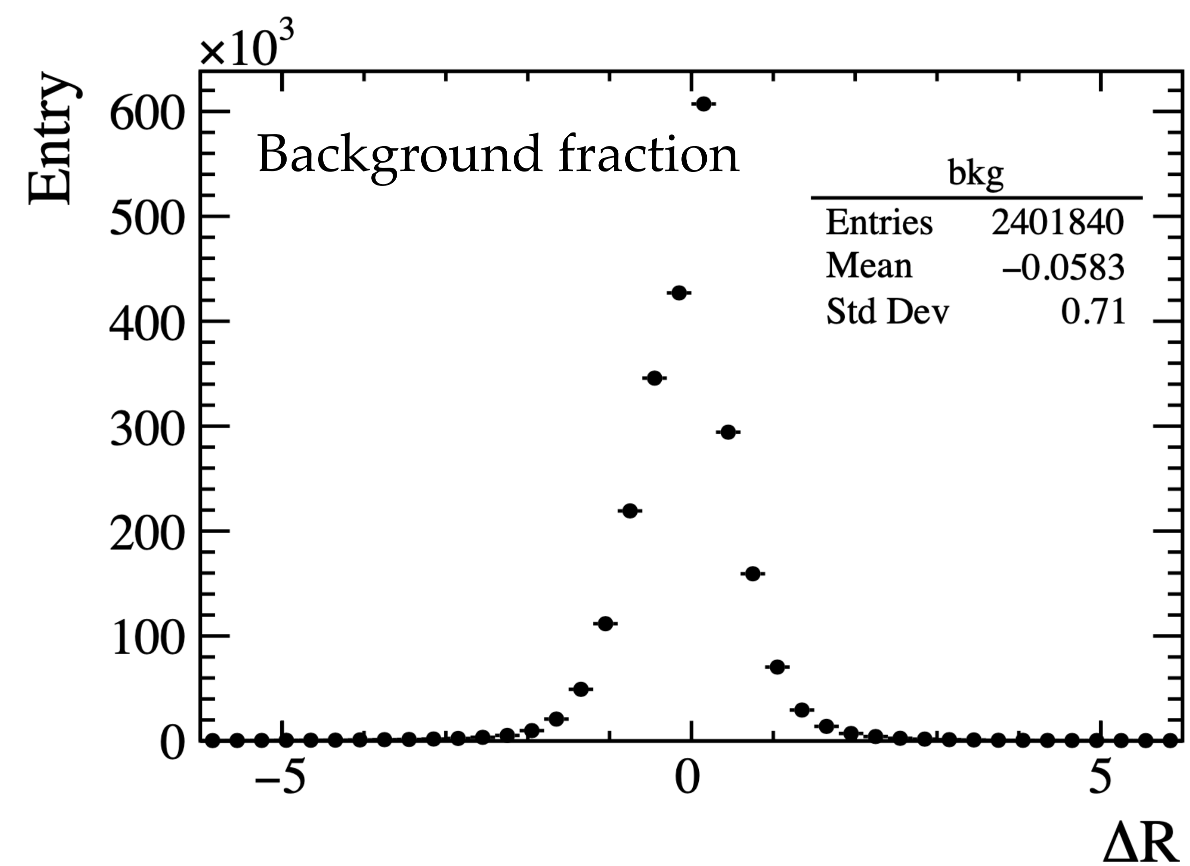
Framework

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

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

- 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



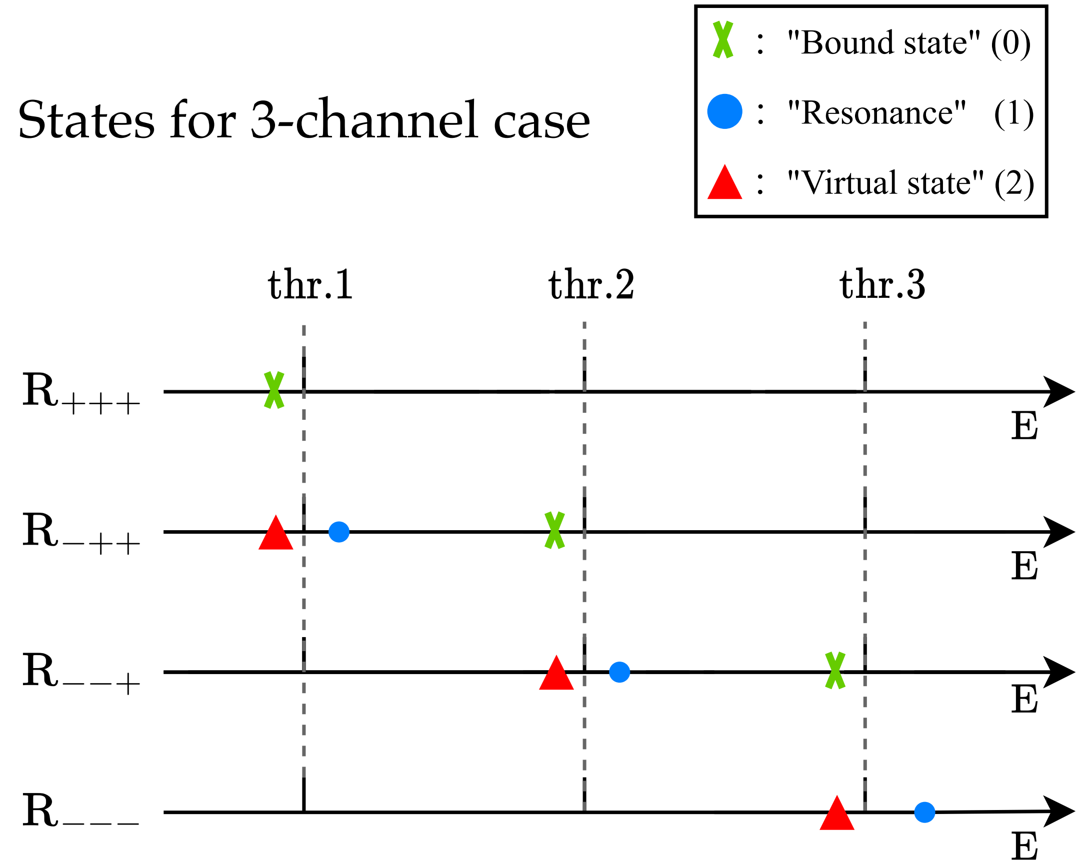


Framework

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

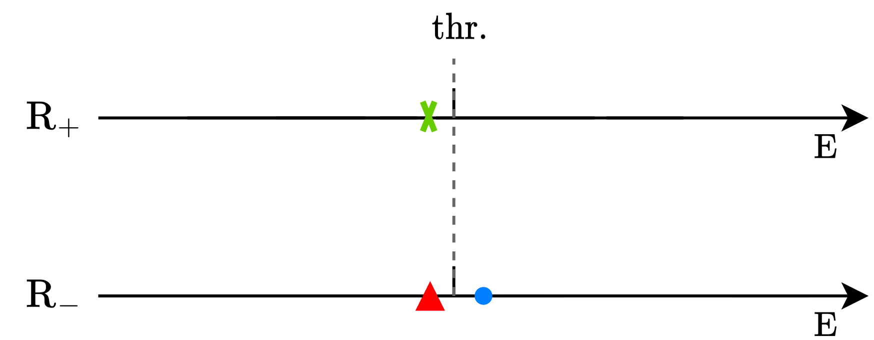


LHCb, PRL122(2019)222001

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

Solution B 0000

States for 1-channel case



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

Results

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

Predicted probability

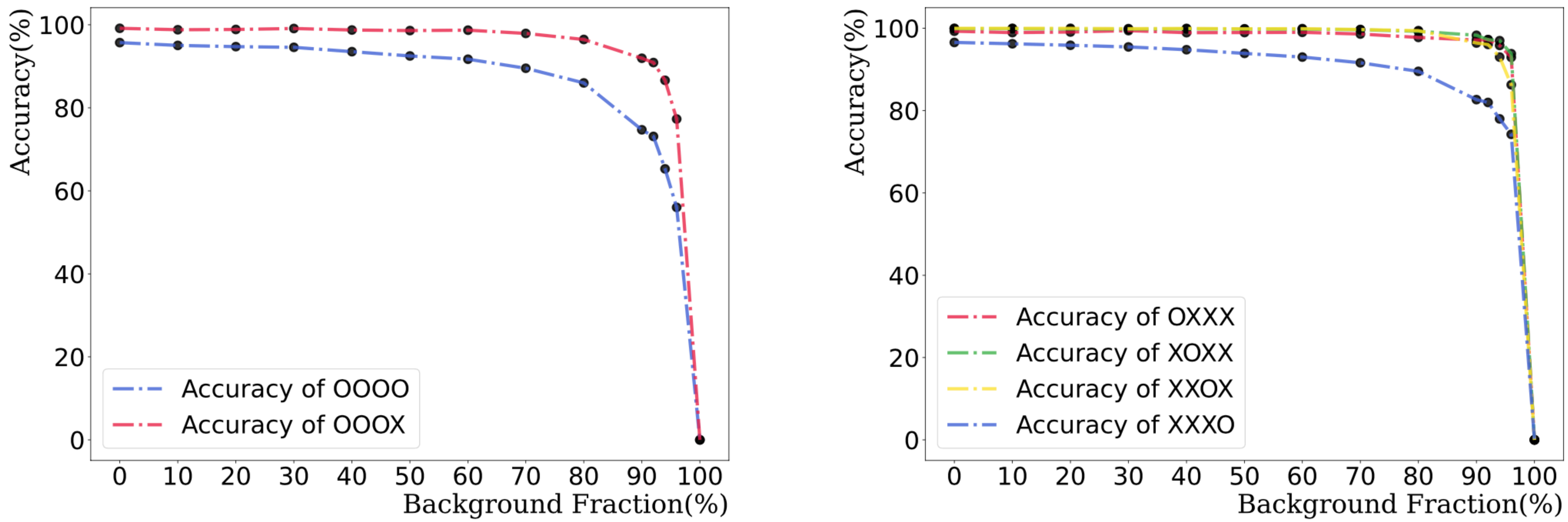
Output(%) Label / NN	0000	1000	1001	1002	100X	others
prediction of NN trained with $\{S^{90}\}$ samples.						
NN 1	0.69	89.13	1.42	8.75	99.30	0.01
NN 2	0.03	5.83	38.47	55.30	99.60	0.37
NN 3	0.03	5.39	15.79	78.41	99.59	0.11
NN 4	0.01	1.9	27.01	70.95	99.86	0.13
NN 5	2.40	94.45	0.15	2.99	97.59	0.01
5 NNs 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

Results

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

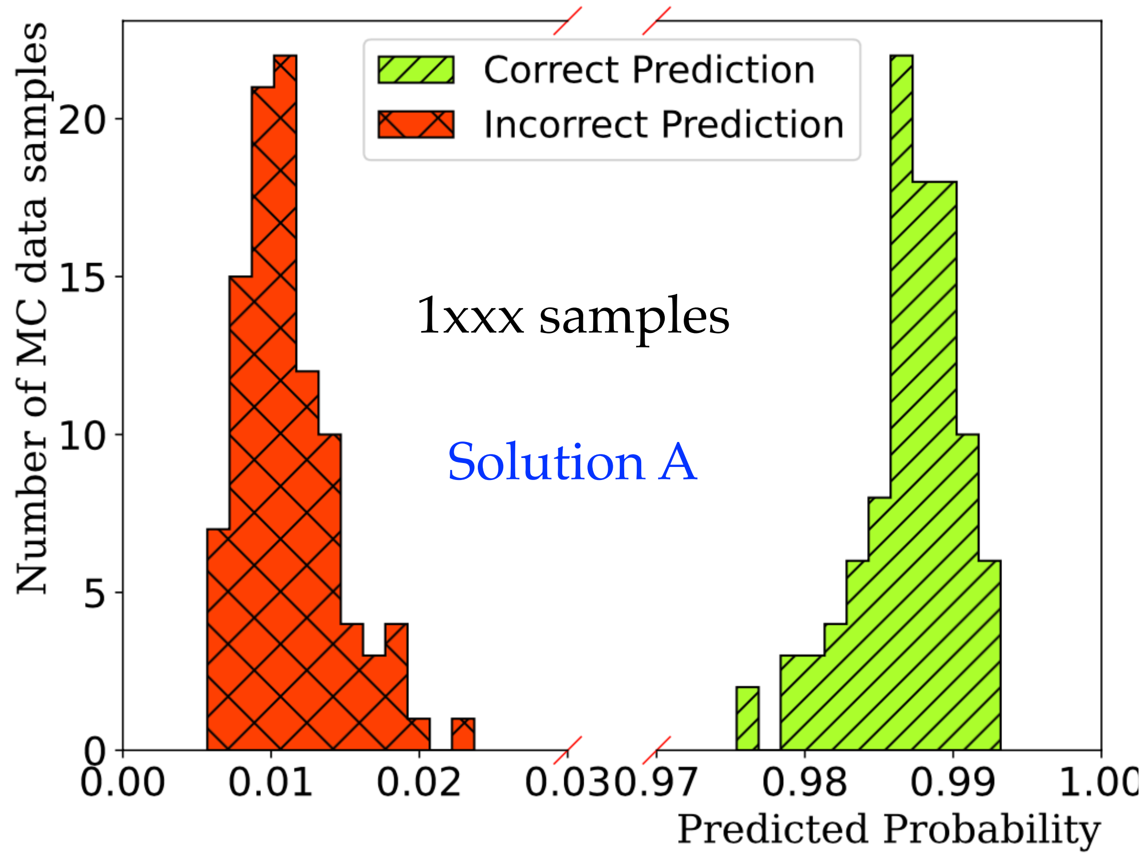
Zhang, Liu, Hu, **QW, Meißner**, Sci.Bull.68(2023)981-989

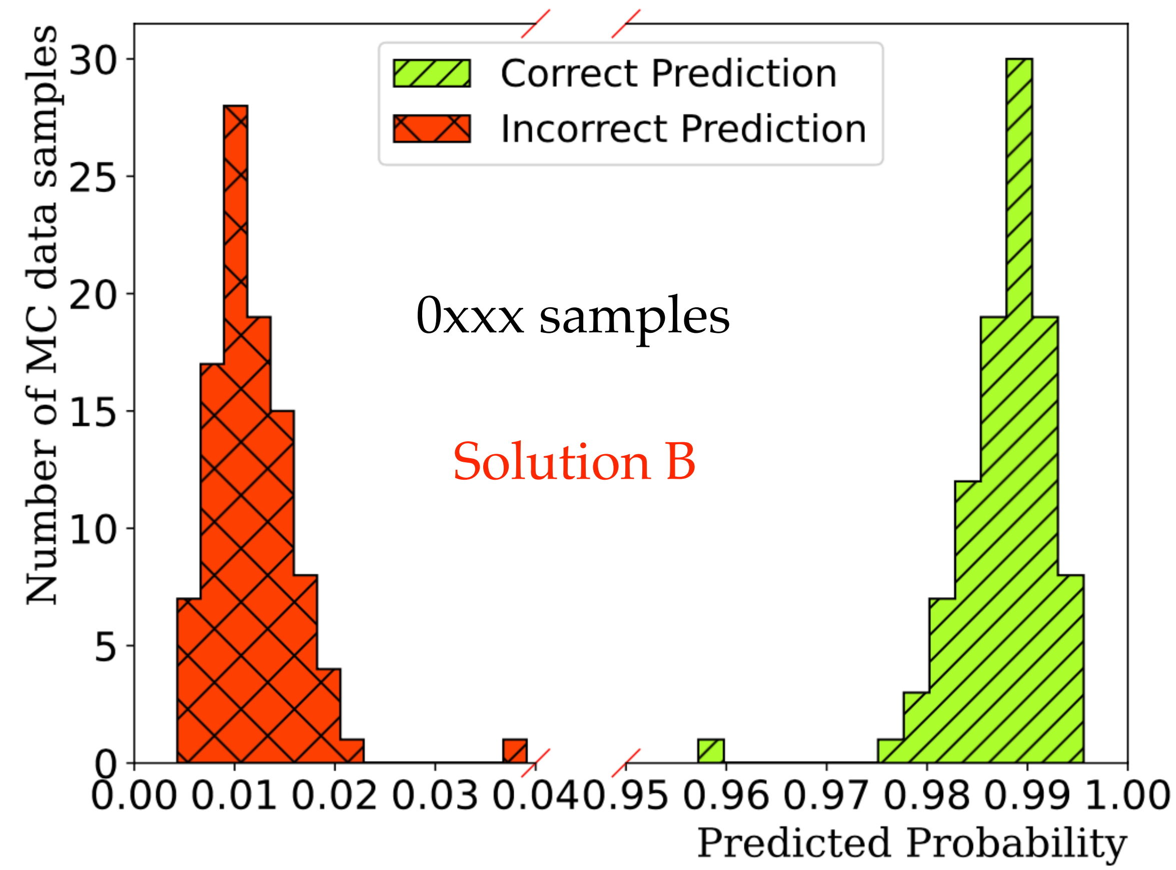
Results

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

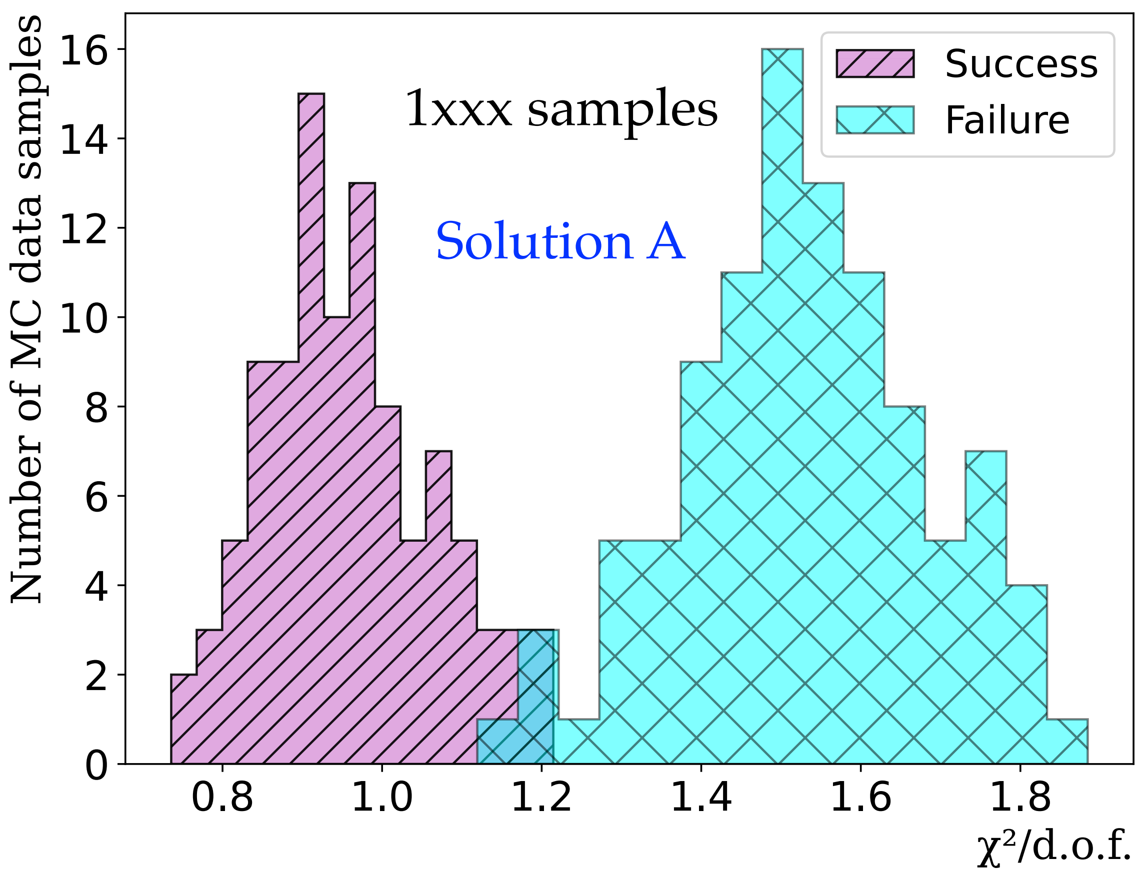
Zhang, Liu, Hu, **QW, Meißner**, Sci.Bull.68(2023)981-989

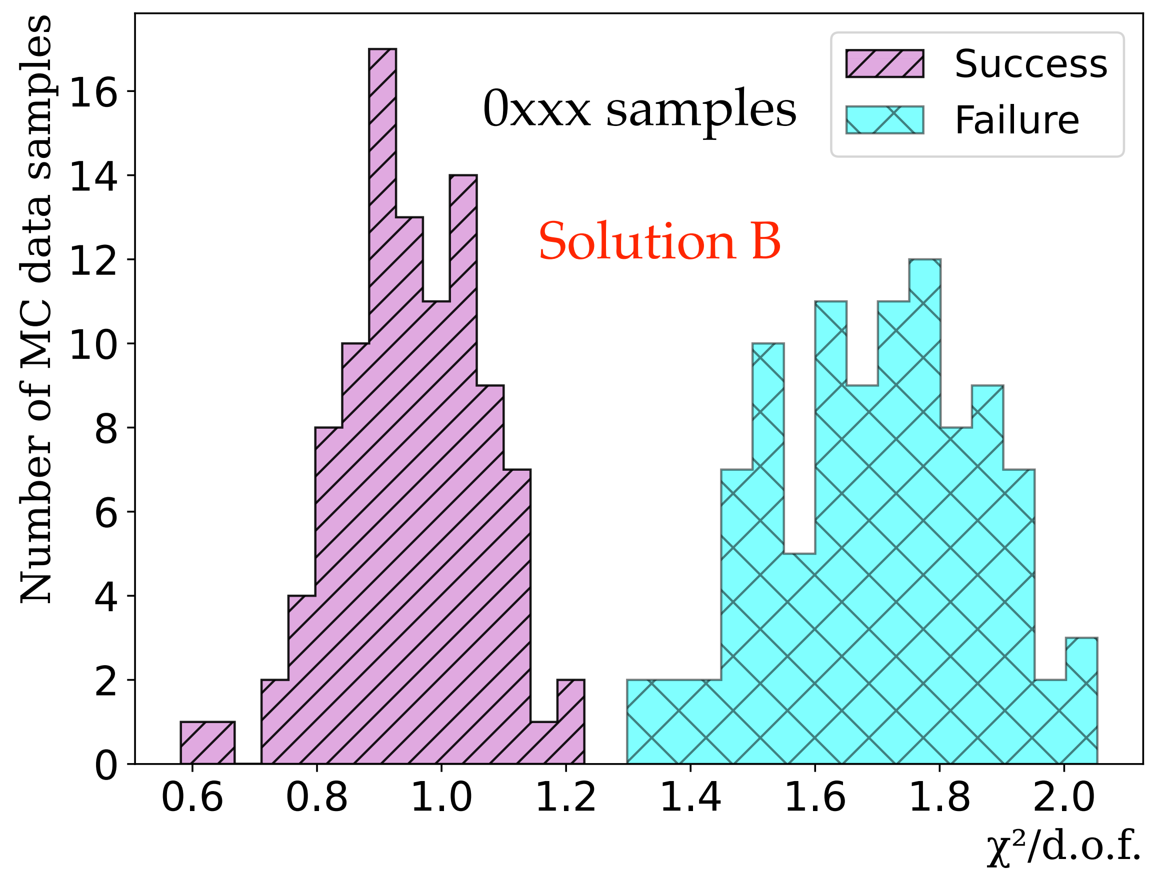
Results

Why NN favors Solution A?

Generate 100 1xxx samples and 100 0xxx samples

Reduced chisq from the normal fitting



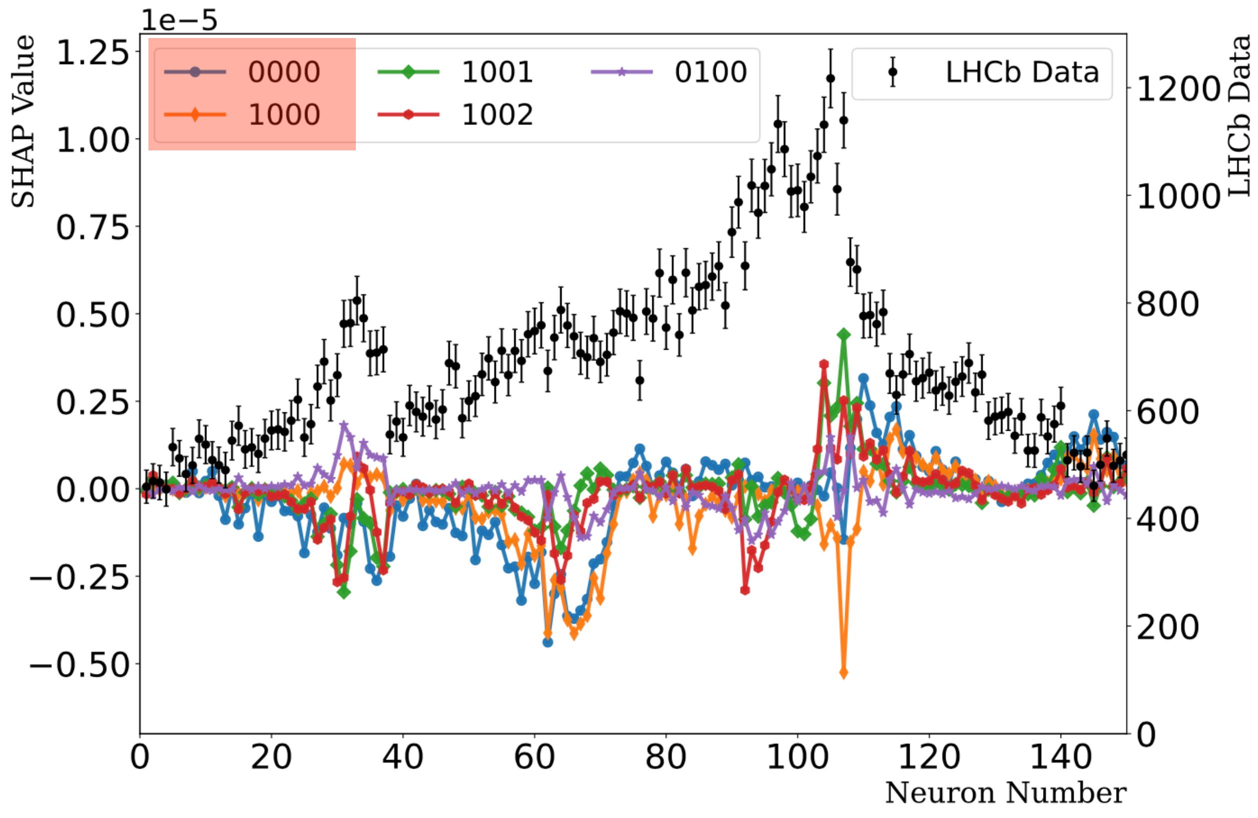


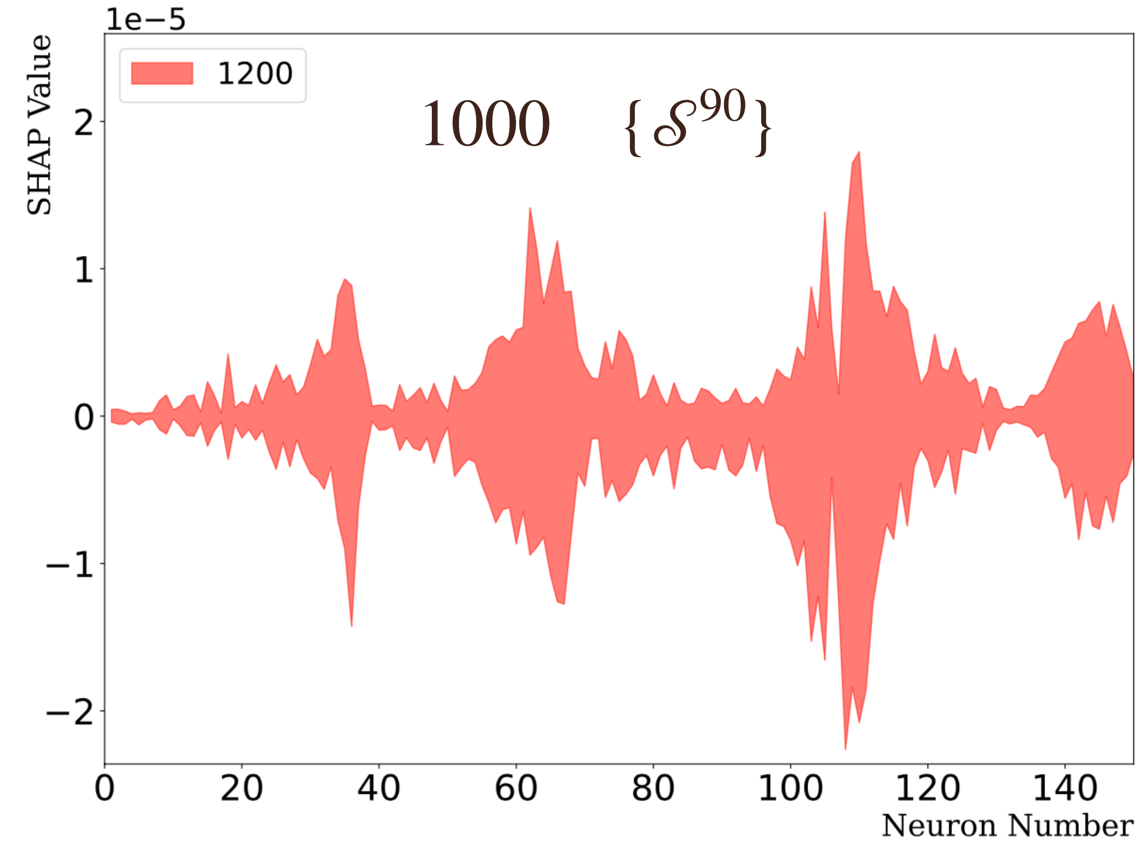
- A 3% misidentification for 1xxx samples

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

Results

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 peaks in the mass spectrum have a greater impact.

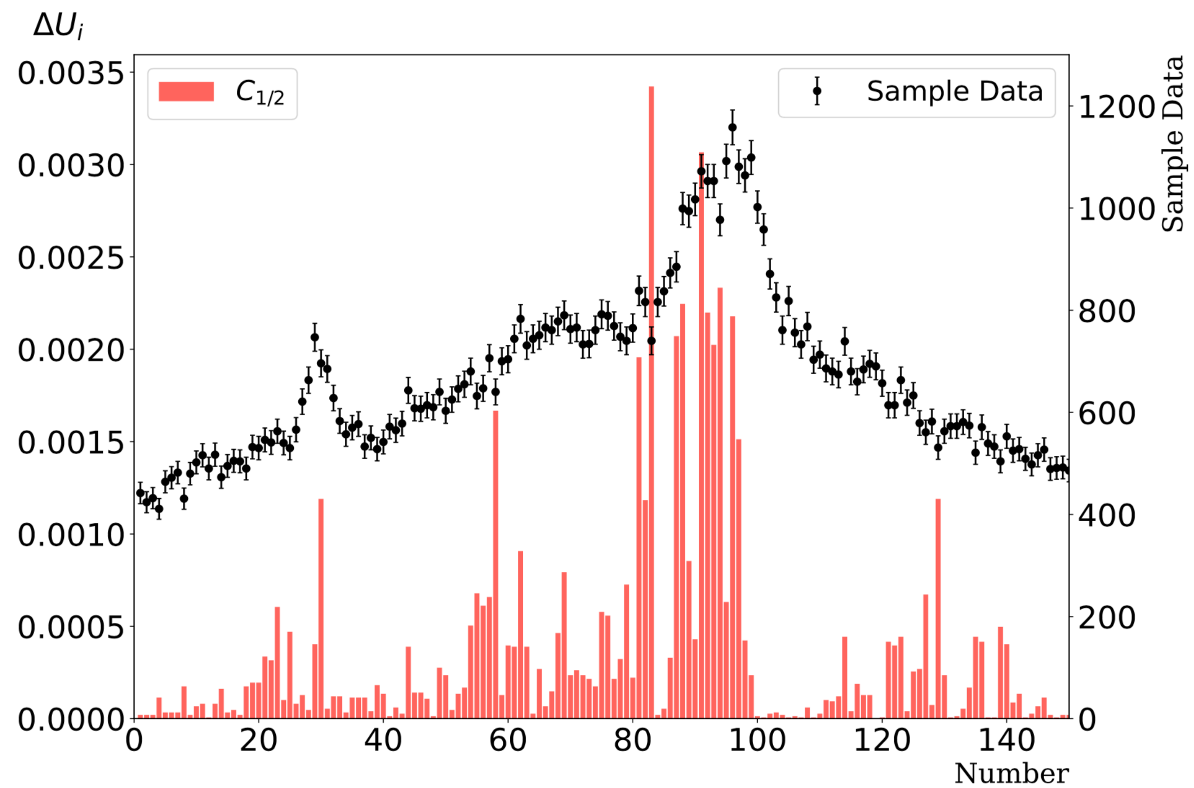
Zhang, Liu, Hu, **QW, Meißner,** Sci.Bull.68(2023)981-989

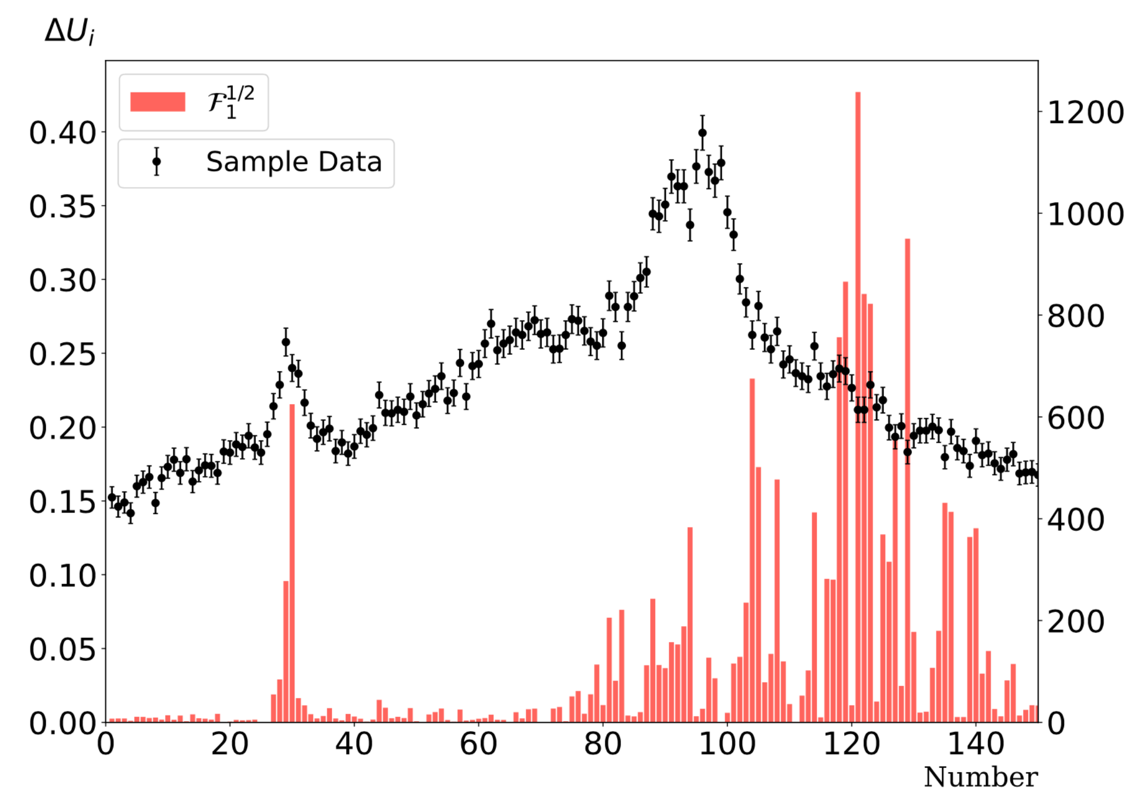
Results

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

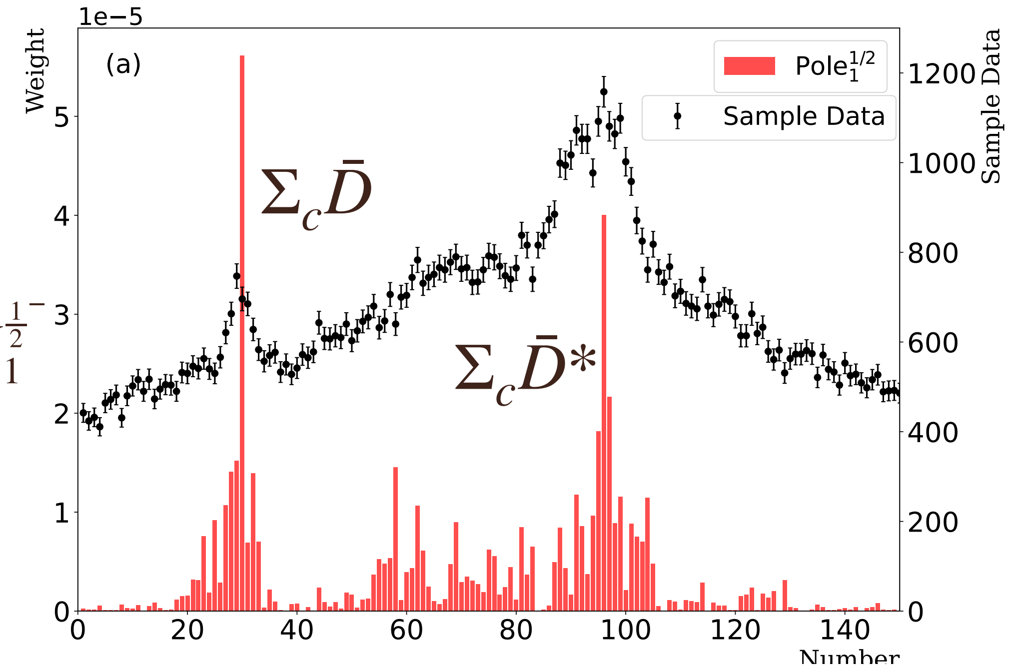
Results

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

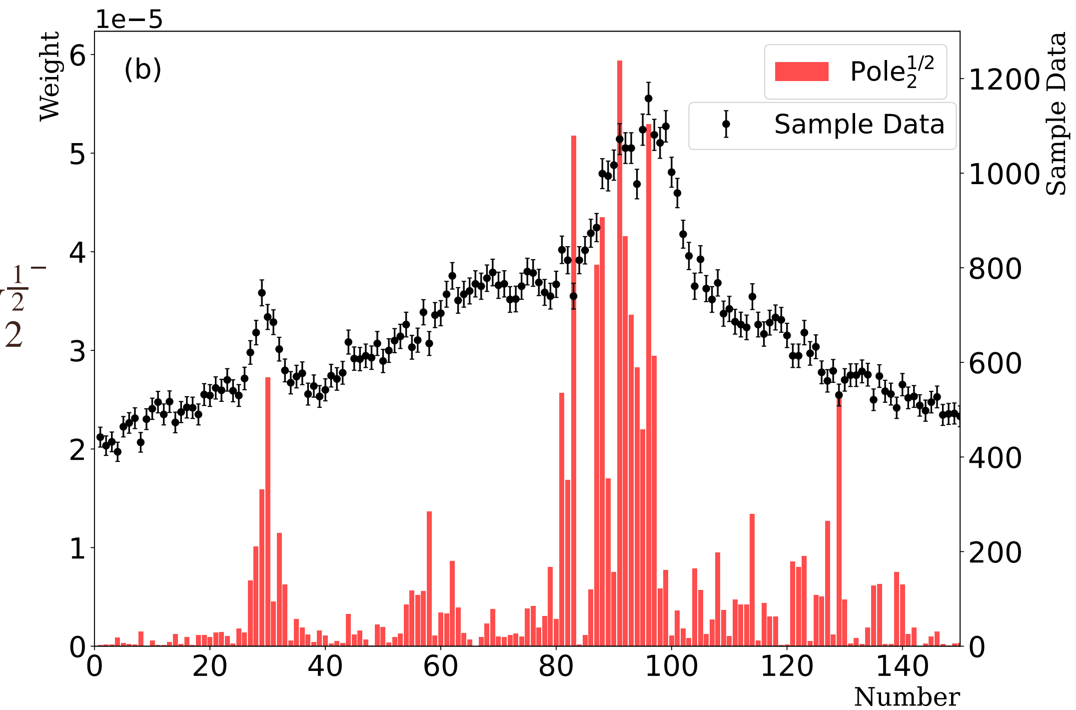
The impact of each experimental data point in normal fitting

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

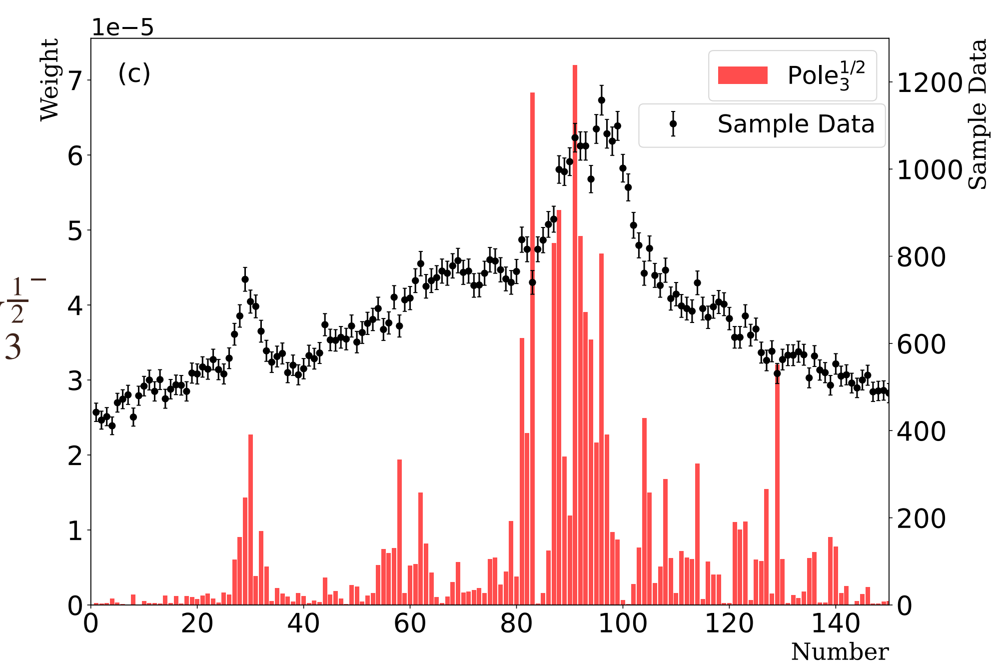
$\Delta U_1^{\frac{1}{2}^-}$



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



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



- 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

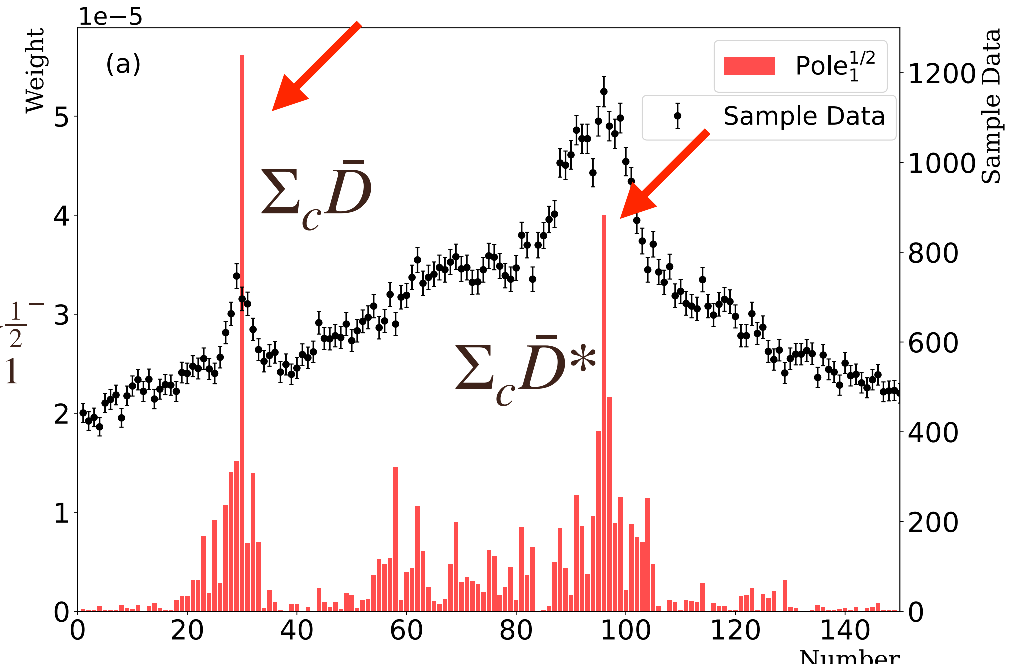
Results

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

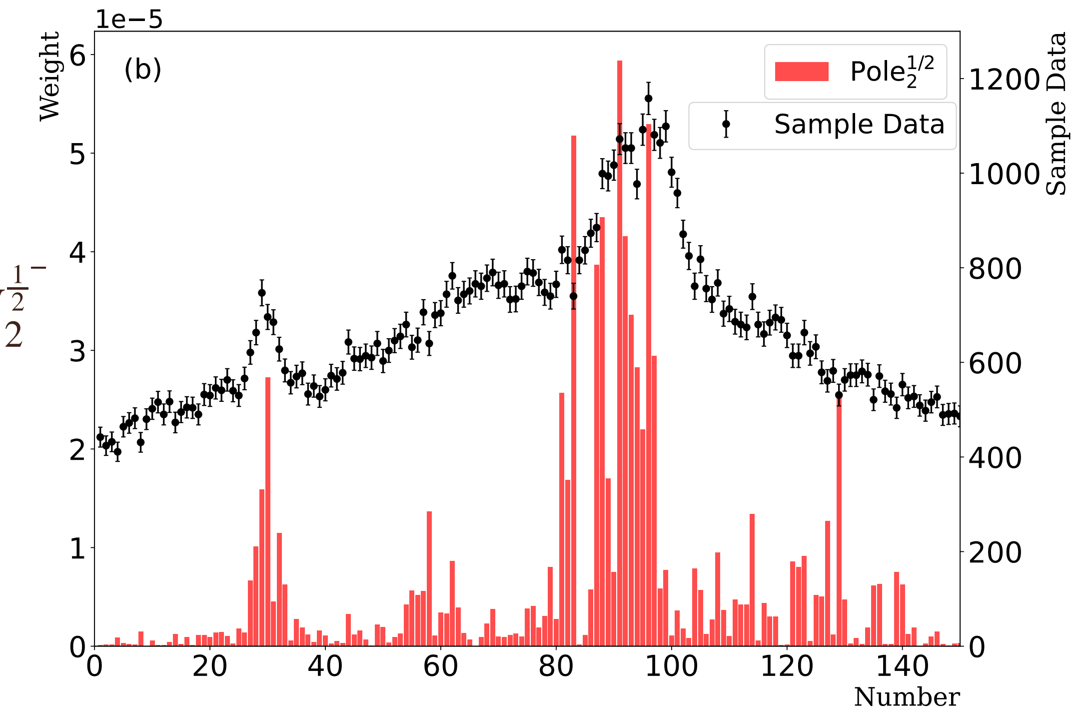
The impact of each experimental data point in normal fitting

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

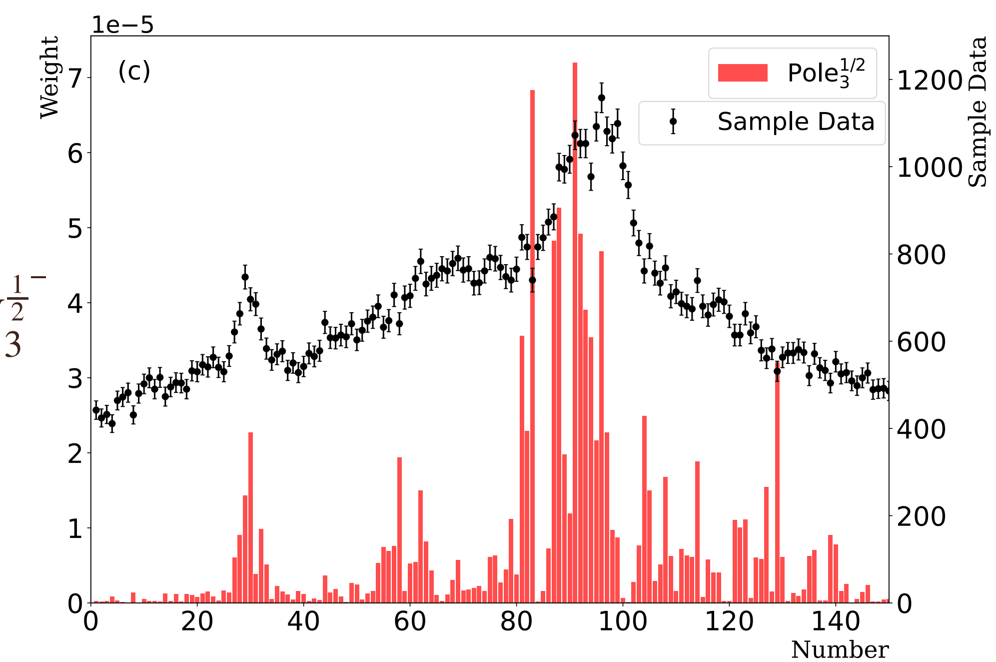
$\Delta U_1^{\frac{1}{2}^-}$



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



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



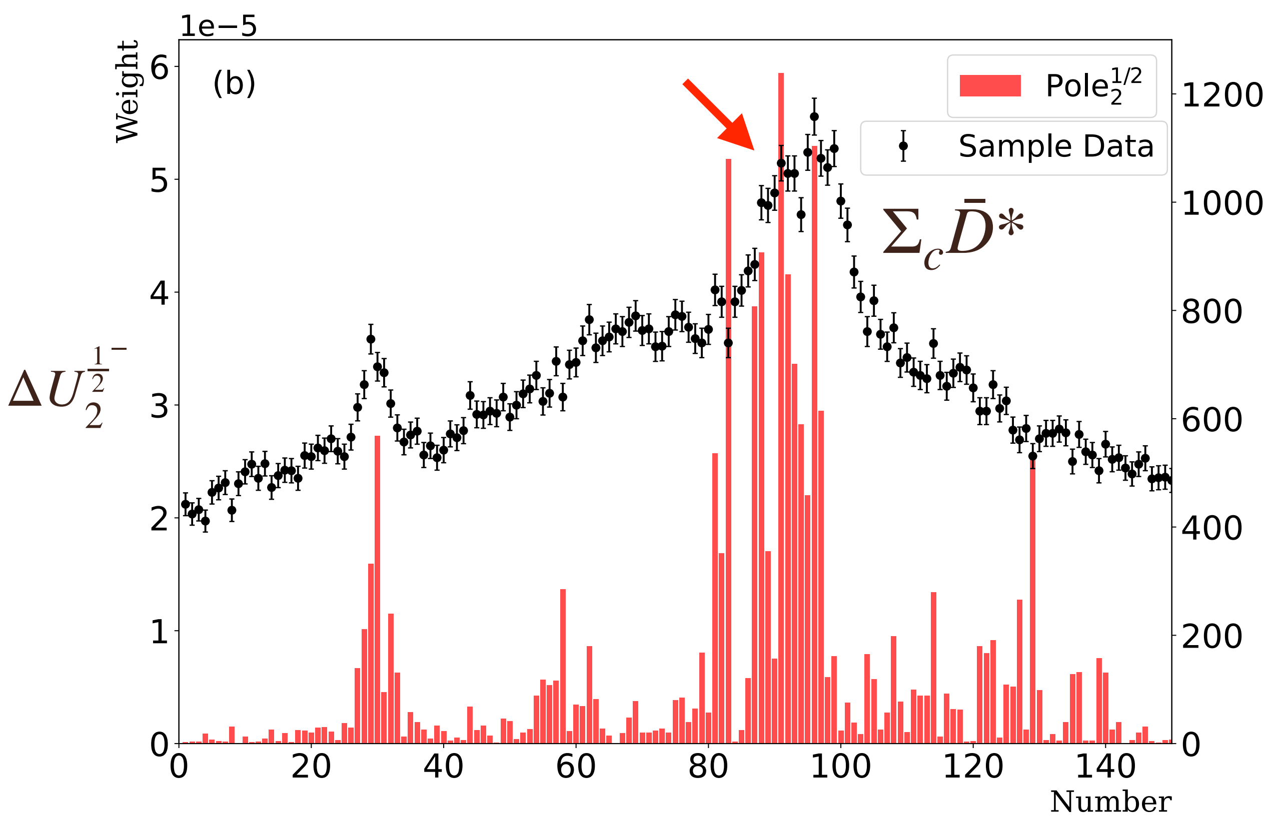
- The experimental data around the coupled-channels still have strong constraints on the physics

Results

The impact of each experimental data point in normal fitting

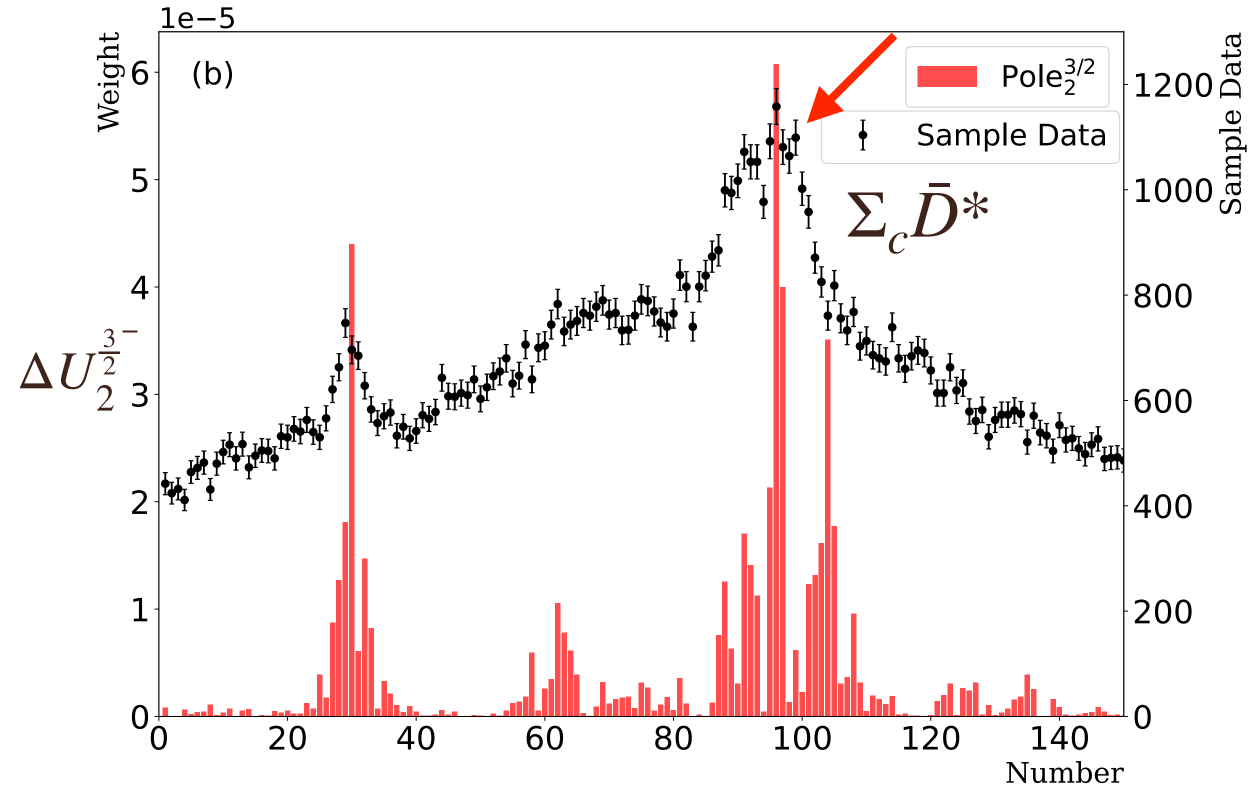
Another analogous quantity $\Delta U_i^{J^P} \equiv \left| \frac{\mathrm{Re[Pole}_i\mathrm{](on)} - \mathrm{Re[Pole}_i\mathrm{](off)}}{\mathrm{Re[Pole}_i\mathrm{](on)}} \right|$ for the ith pole

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

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

Summary and outlook

- Apply the ML to hadron physics
- 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!