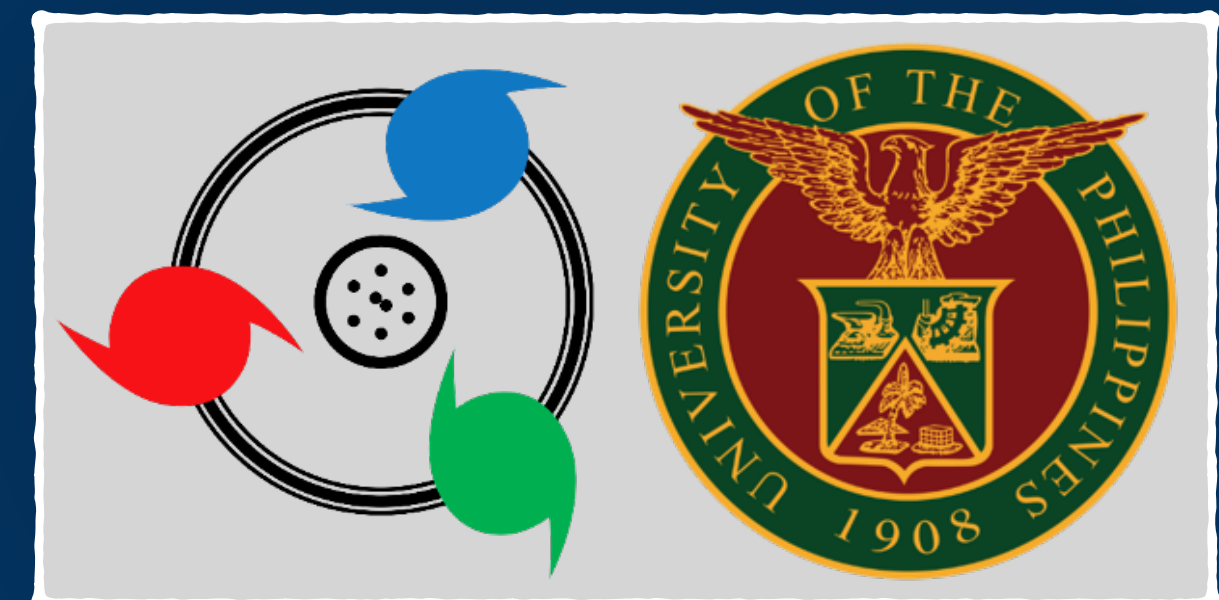


Development of Deep Learning Techniques for Hadron Spectroscopy

Denny Lane Sombillo

National Institute of Physics, University of the Philippines Diliman



Some updates from the Philippines

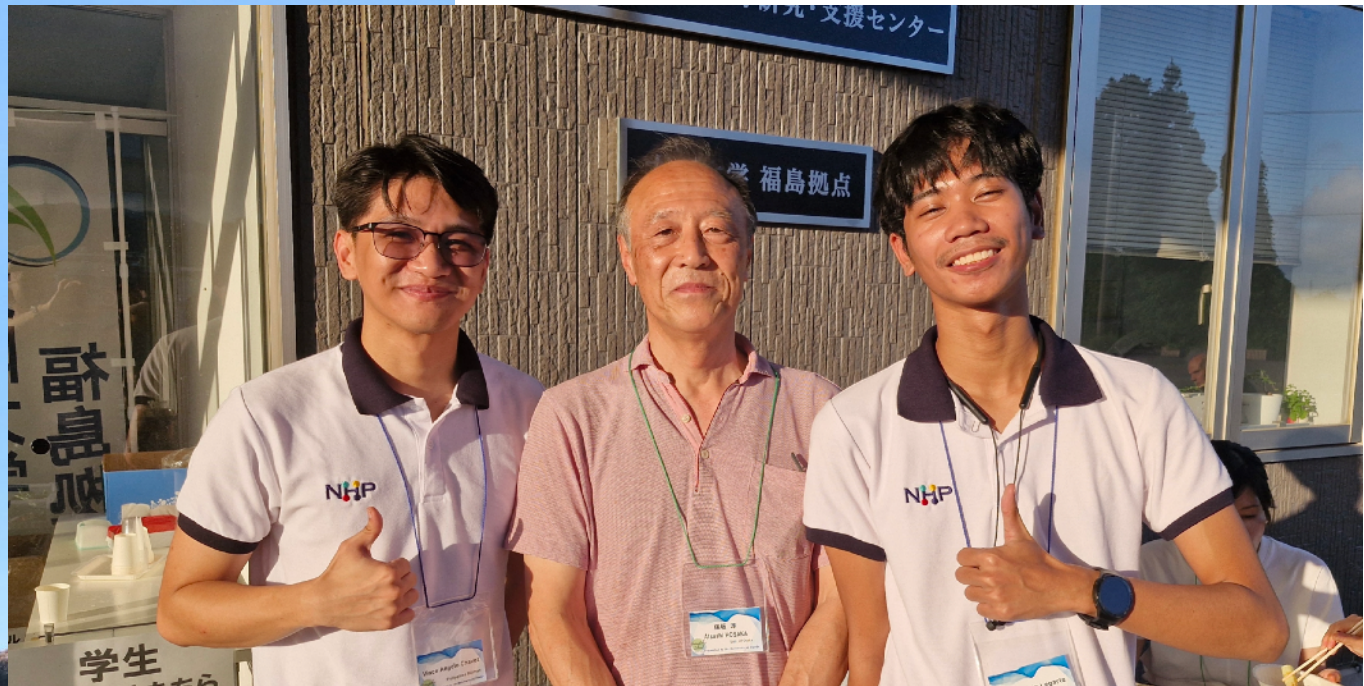
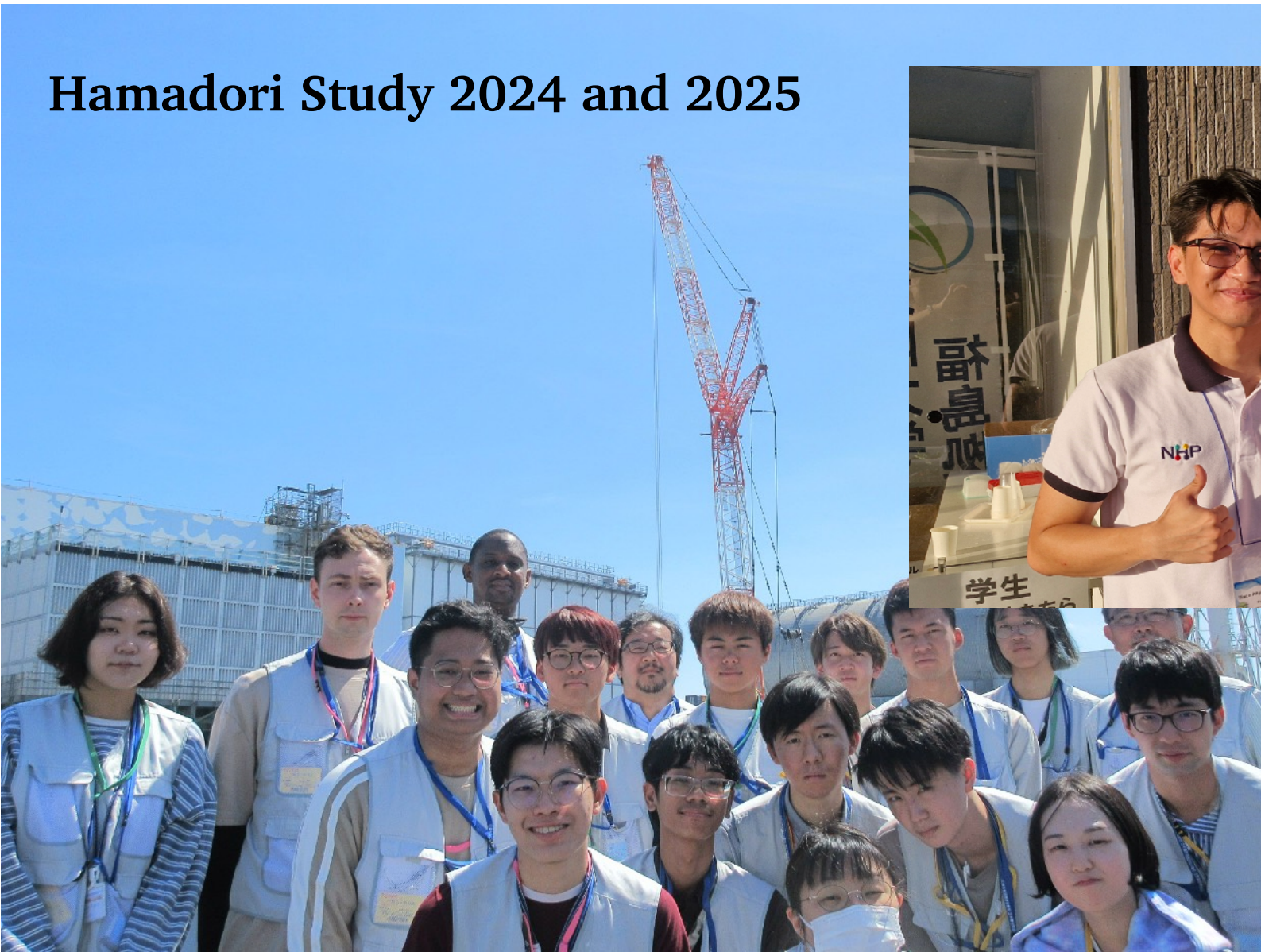
Hashimoto Prize & ANPhA 1st (gold)

Darwin Alexander Co (Univ. of Philippines Diliman)
Understanding triangle singularity using deep neural network



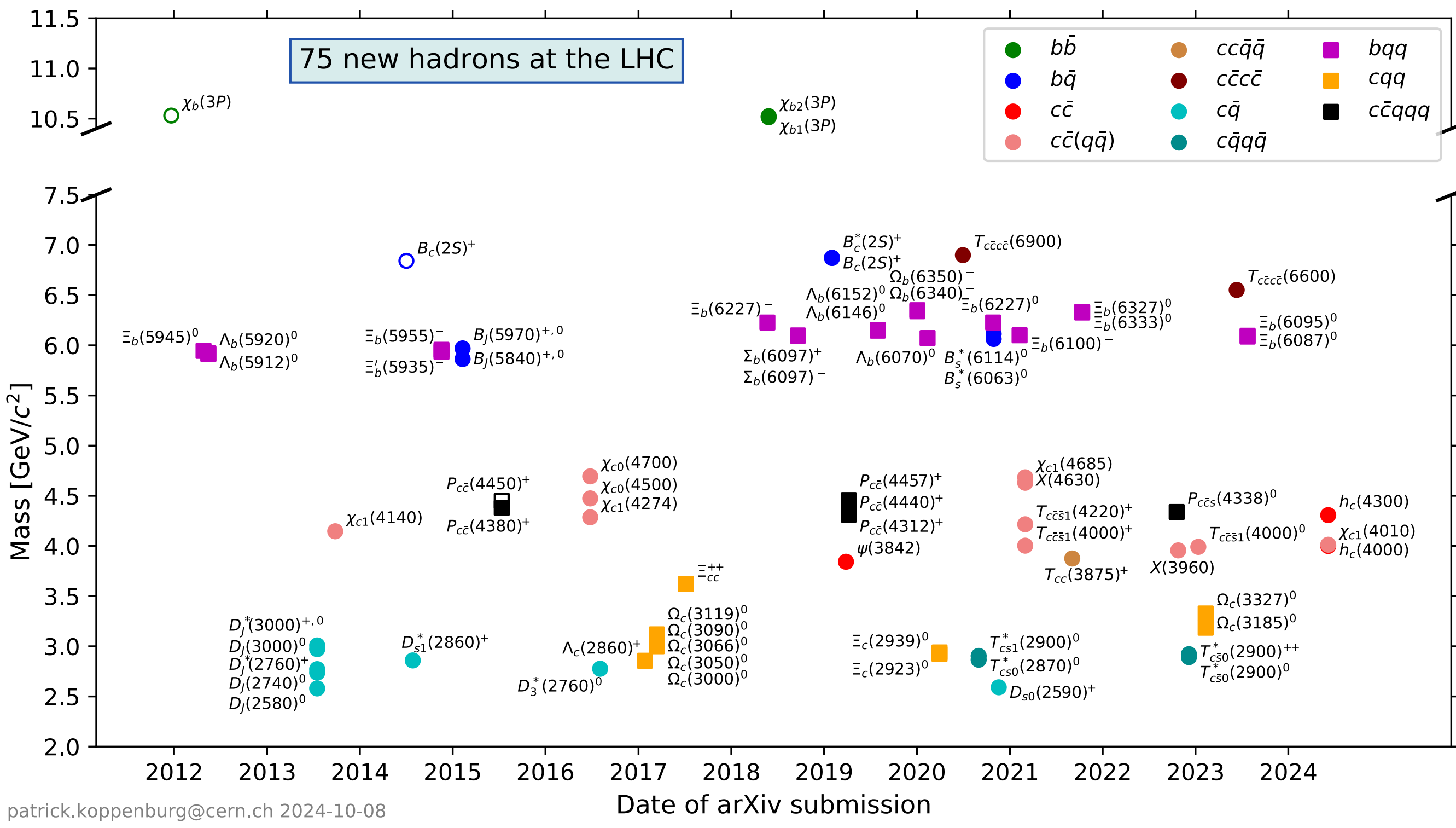
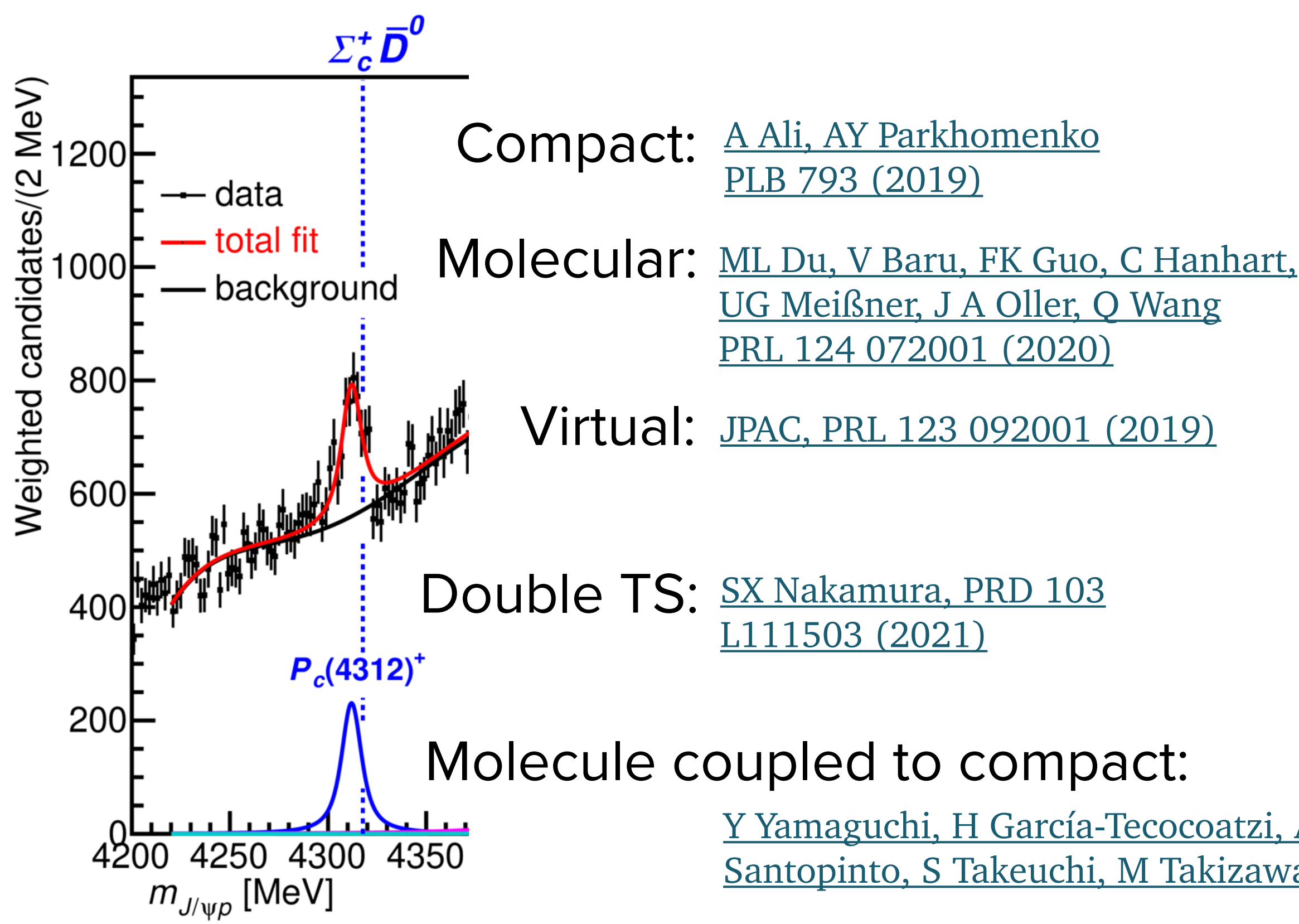
SNP CNS Summer School Incentive Prize

Vince Angelo Chavez "Analysis of f0(980) using Deep Neural Network and Uniformized S-Matrix"



Hadron spectroscopy - line shape interpretation

- Many observations of possible new states.
- Some are near hadron-hadron thresholds.
- Experiments give us the line shape.
- How do we interpret/ classify them?



Hadron spectroscopy - line shape interpretation

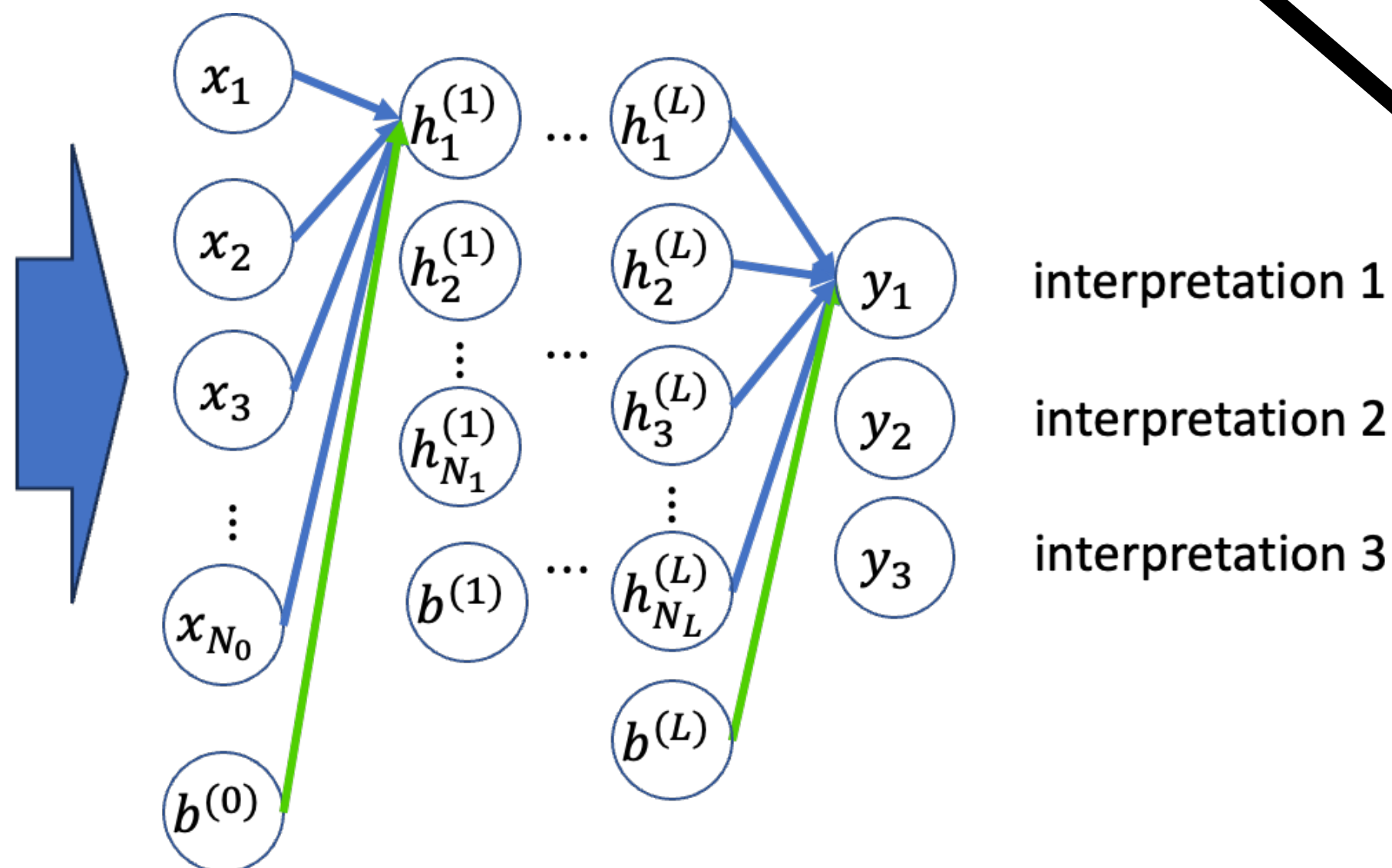
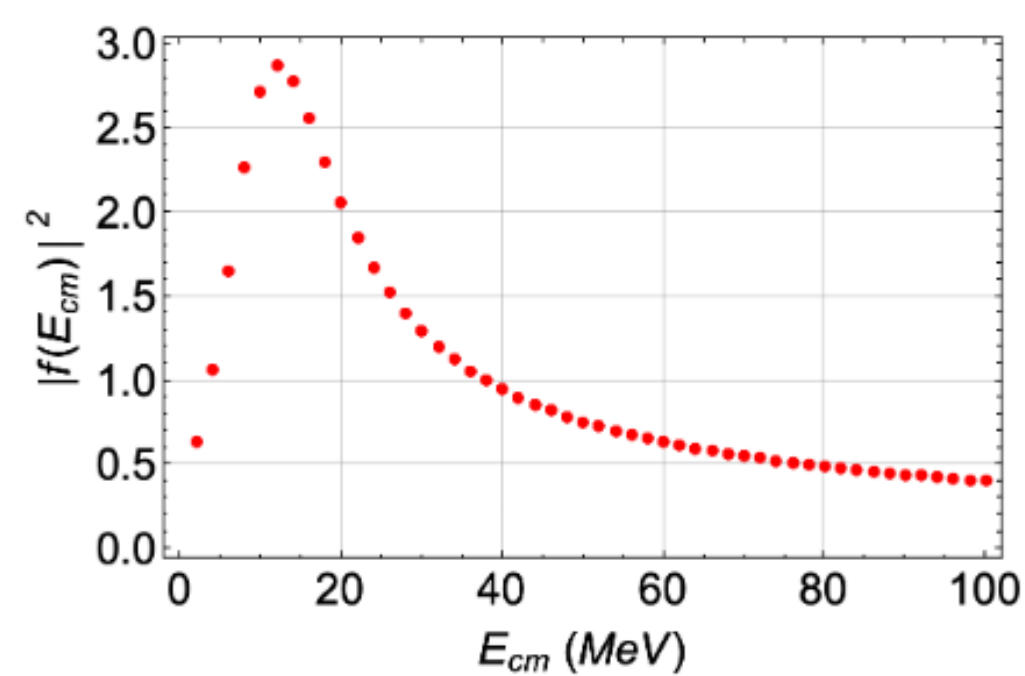
Why use machine learning in line shape analysis?

Line shape interpretation - classification problem

Deep neural network (DNN) as a universal approximator (universal map).

[K Hornik, M Stinchcomber, H White, Neural Net., 2 5 \(1989\)](#)

DNN can be trained to map input line shape space into output interpretation space.



Linear transformation + non-linear activations and squashing

$$z_j^{(0)} = W_{ij}^{(0)} x_i + b_j^{(0)}$$

$$h_j^{(1)} = \text{ReLU}(z_j^{(1)})$$

$\vdots$

$$z_j^{(n)} = W_{ij}^{(n)} h_i^{(n)} + b_j^{(n)}$$

$$h_j^{(n+1)} = \text{ReLU}(z_j^{(n)})$$

$\vdots$

$$z_j^{(L)} = W_{ij}^{(L)} h_i^{(L)} + b_j^{(L)}$$

$$y_j = \frac{\exp(z_j^{(L)})}{\sum_i \exp(z_i^{(L)})}$$

Backpropagation to minimize the cost:

$$C(\hat{W}, \vec{b}) = \frac{1}{X} \sum_{\vec{x}} \vec{a}(\vec{x}) \cdot \log [\vec{y}_{W,b}(\vec{x})]$$

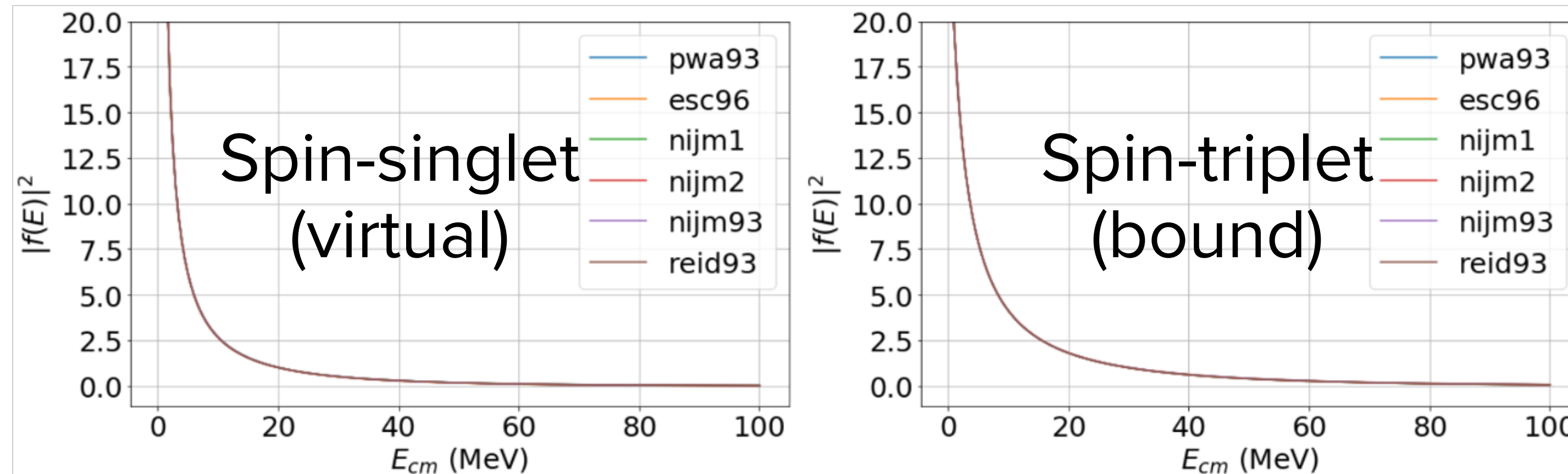
[DE Rumelhart, GE Hinton, RJ Williams, Nature 323, 533-536 \(1986\)](#)

Deep learning: proof of principle

DLBS, YI, TS, AH PRD 102 016024 (2020)
DLBS, YI, TS, AH Few-Body Syst. 62, 52 (2021)

Benchmarked on the known nucleon-nucleon bound state

Given only the s-wave cross section, the origin of enhancement can be unambiguously identified.



For near-threshold pole:
 $k \cot \delta \sim -1/a$ (constant)

$$|f(k)|^{-2} = |k \cot \delta - ik|^2 \sim \frac{1}{a^2} + k^2$$

Not possible to distinguish bound vs virtual pole enhancements.

S-matrix can have distant singularities on the unphysical sheet.

Use different (unitary, analytic) backgrounds to help DNN distinguish bound and virtual enhancements.

$$S(k) = \exp \left[2i\delta_{bg}(k) \right] \frac{k + i\gamma}{k - i\gamma}; \quad \delta_{bg} = \alpha \tan^{-1} \left(\frac{k}{\beta} \right)$$

Deep learning: proof of principle

DLBS, YI, TS, AH PRD 102 016024 (2020)
DLBS, YI, TS, AH Few-Body Syst. 62, 52 (2021)

Optimize parameters of DNN using mock amplitudes:

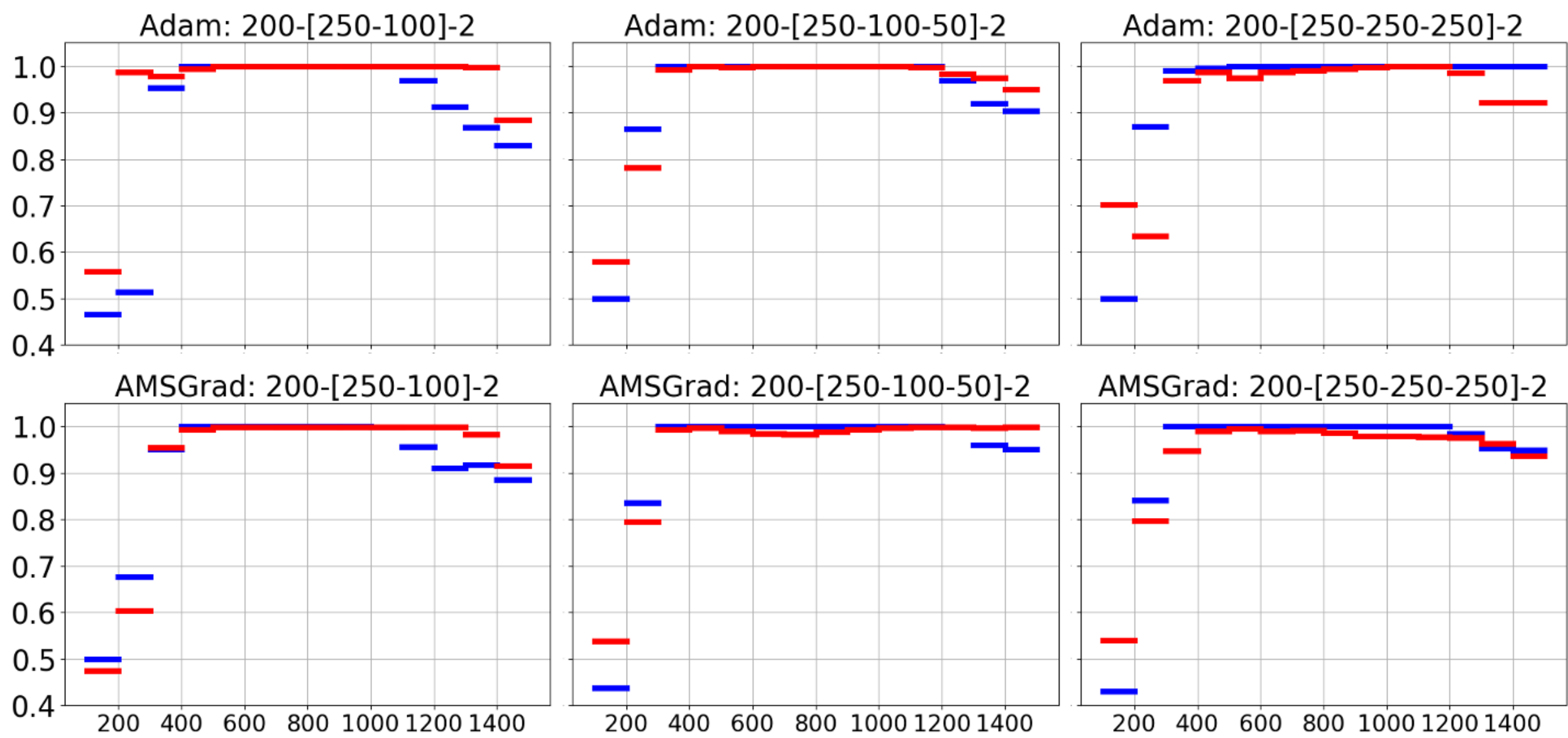
$$S(k) = \exp \left[2i\delta_{bg}(k) \right] \frac{k + i\gamma}{k - i\gamma}$$

3,200,000 training data

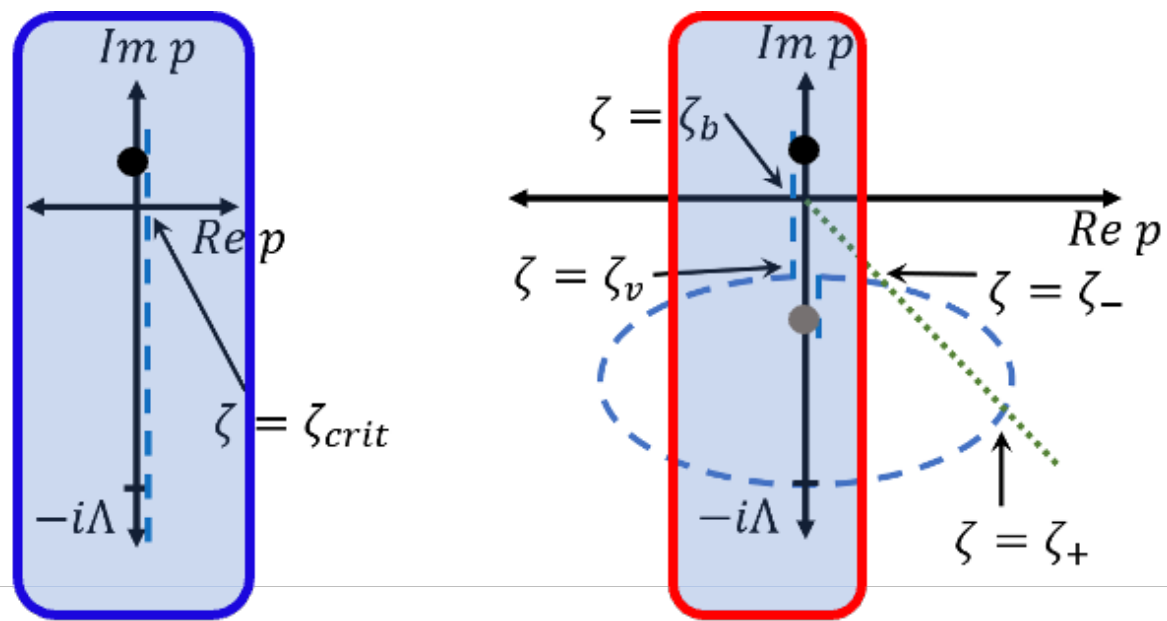
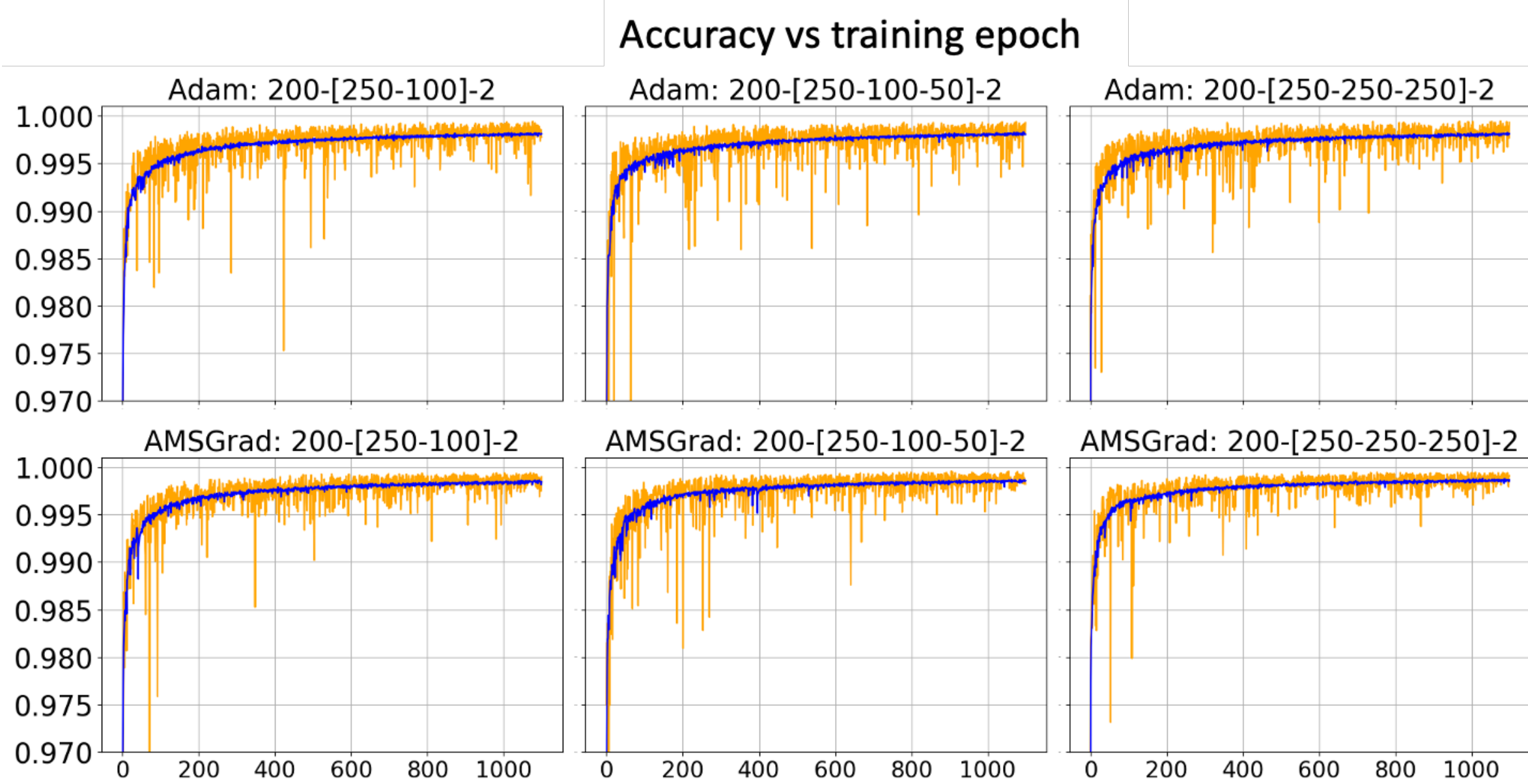
800,000 testing data

Validation using separable potential model

$$v(p, p') = \zeta \left(\frac{\Lambda^2}{p^2 + \Lambda^2} \right) \left(\frac{\Lambda^2}{p'^2 + \Lambda^2} \right)$$



cut-off parameter Λ



Energy independent coupling
 $\zeta \rightarrow \text{constant}$

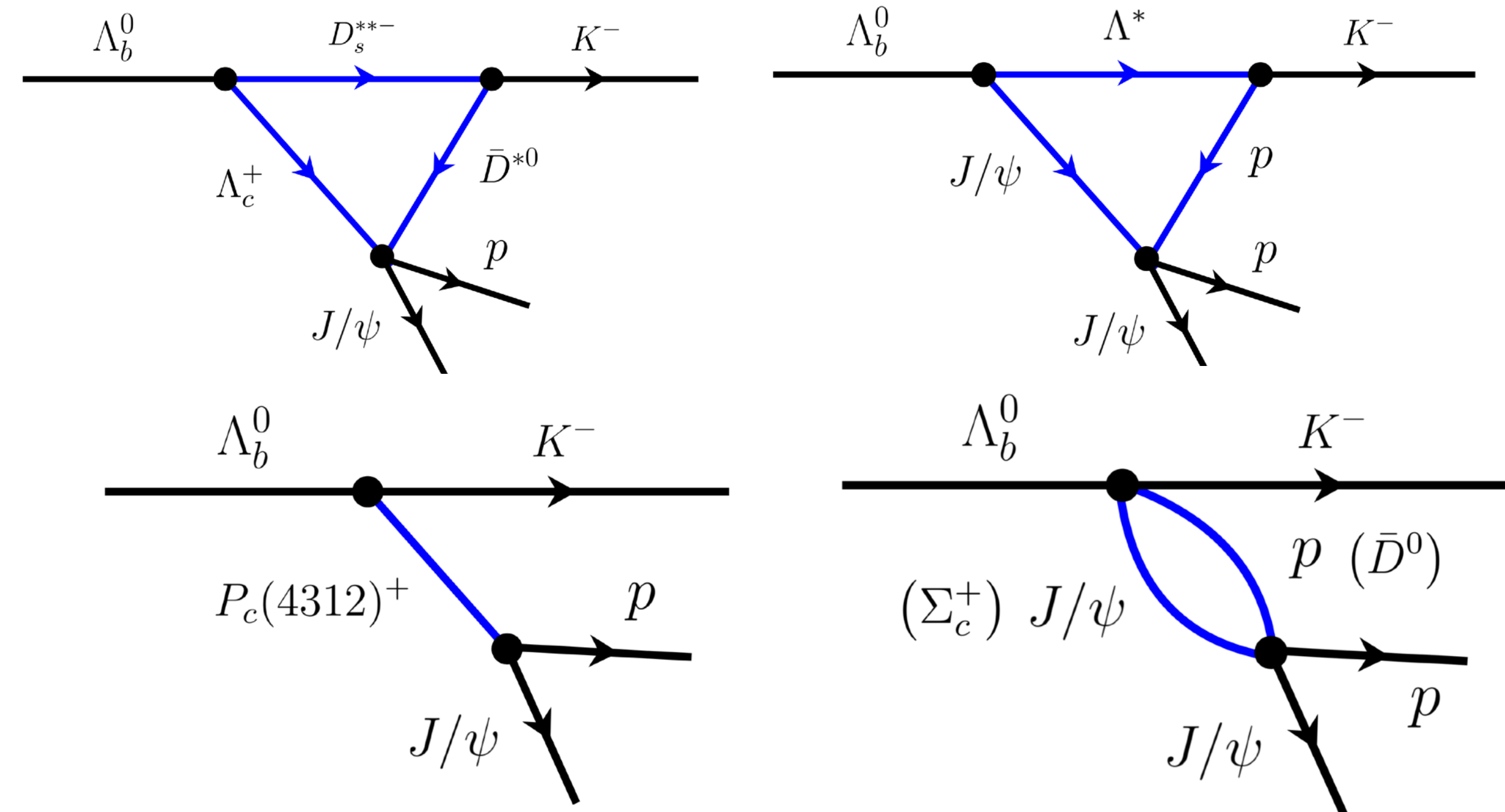
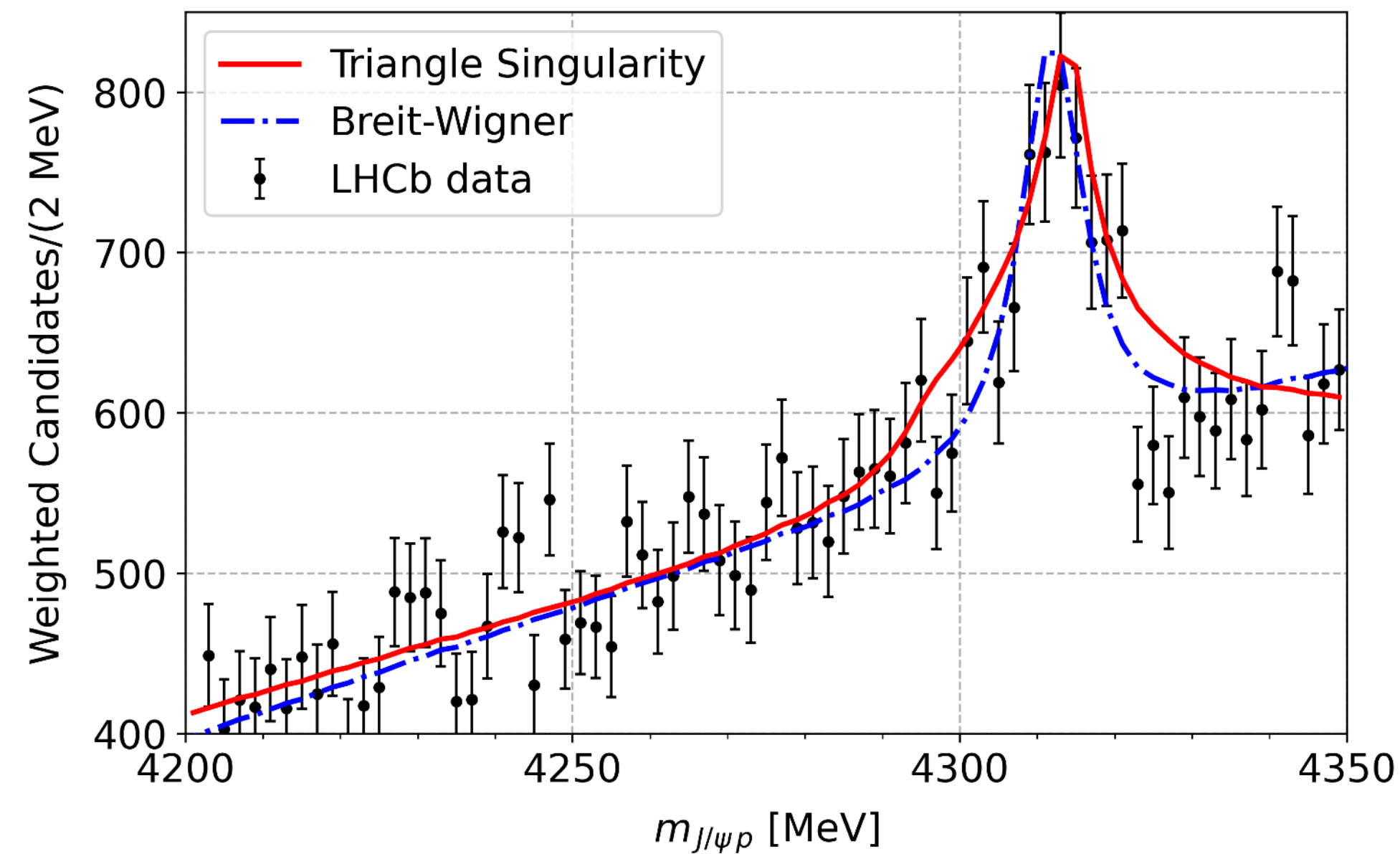
Energy independent coupling
 $\zeta \rightarrow (E - M_{sep})\zeta$

Trained and validated DNN deployed to probe the nucleon-nucleon

	PWA93	ECS96	NijmI	NijmII	Nijm93	Reid93
1S_0	virtual	virtual	virtual	virtual	virtual	virtual
3S_1	bound	bound	bound	bound	bound	bound

Can we train a DNN to discriminate different origin of enhancements?

- Triangle mechanism
- Coupled-channel effects
- True resonance



If yes, then DL can complement our analysis.

If no, then TS and pole enhancements are inherently ambiguous.

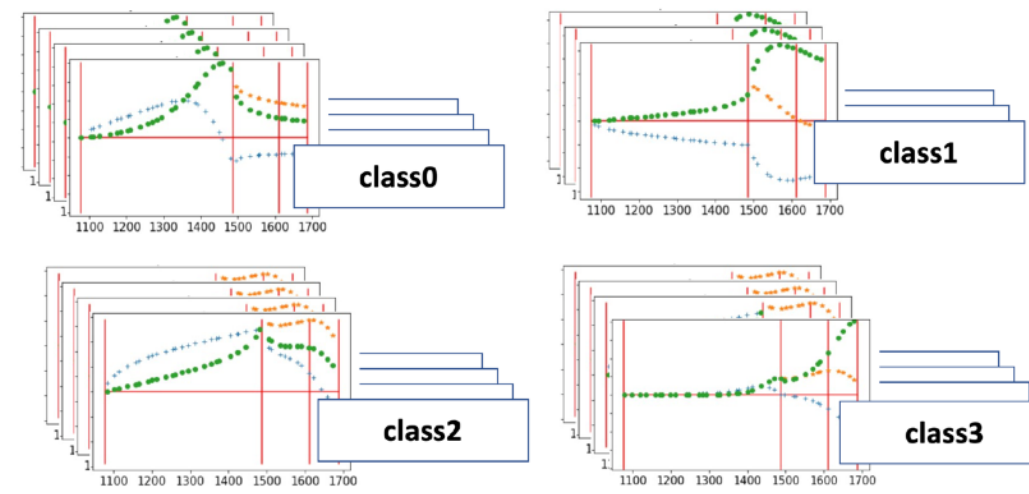
DL as a Model-selection framework

DAO Co, VAA Chavez, DLBS,
PRD 110 (2024) 114034

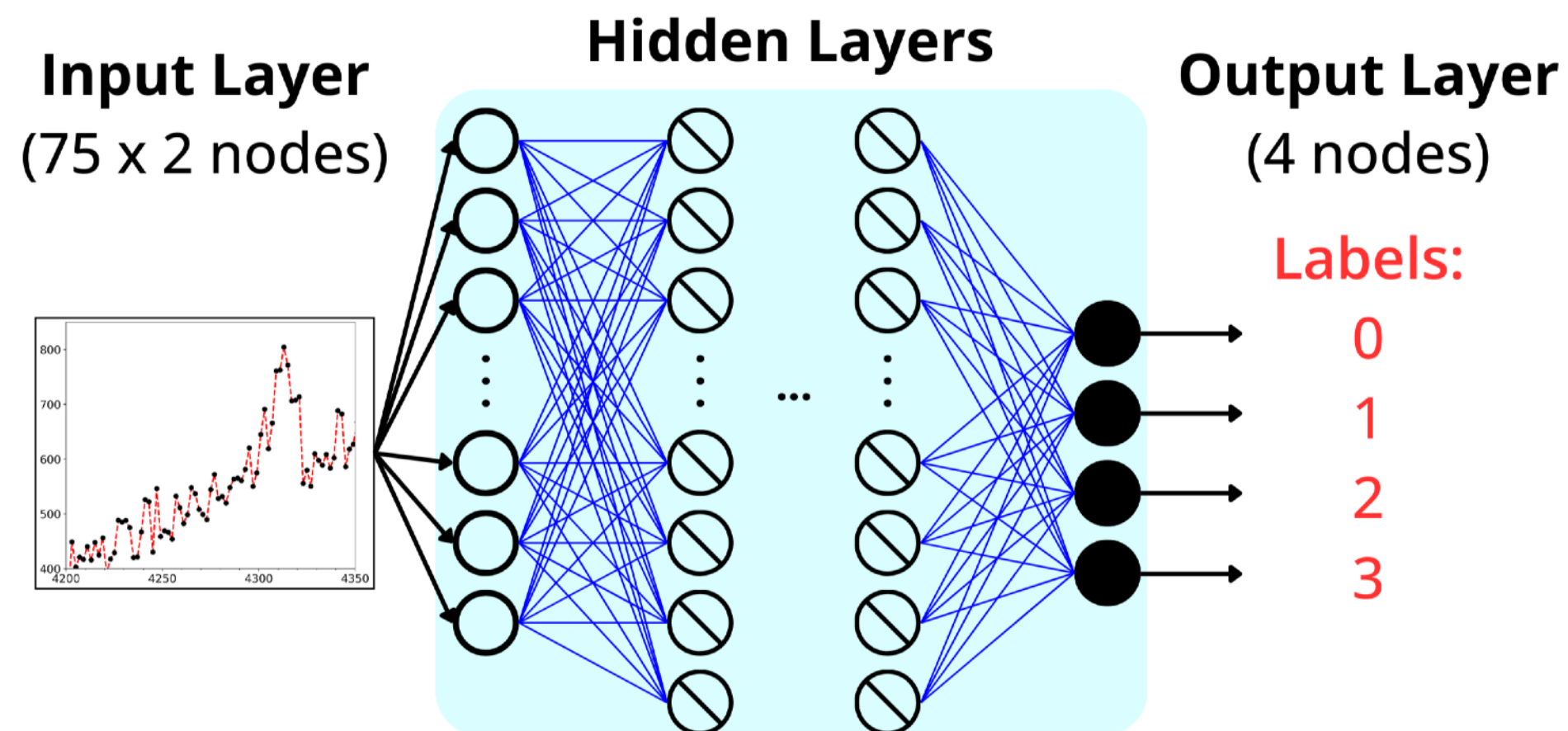
1. Choose the region of interest: [4200 MeV - 4350 MeV]

2. Generate the training dataset

- Triangle singularity
- Pole of S-matrix
 - 1 pole in 2nd RS
 - 1 pole in 4th RS
 - 1 pole each in 2nd and 3rd RS

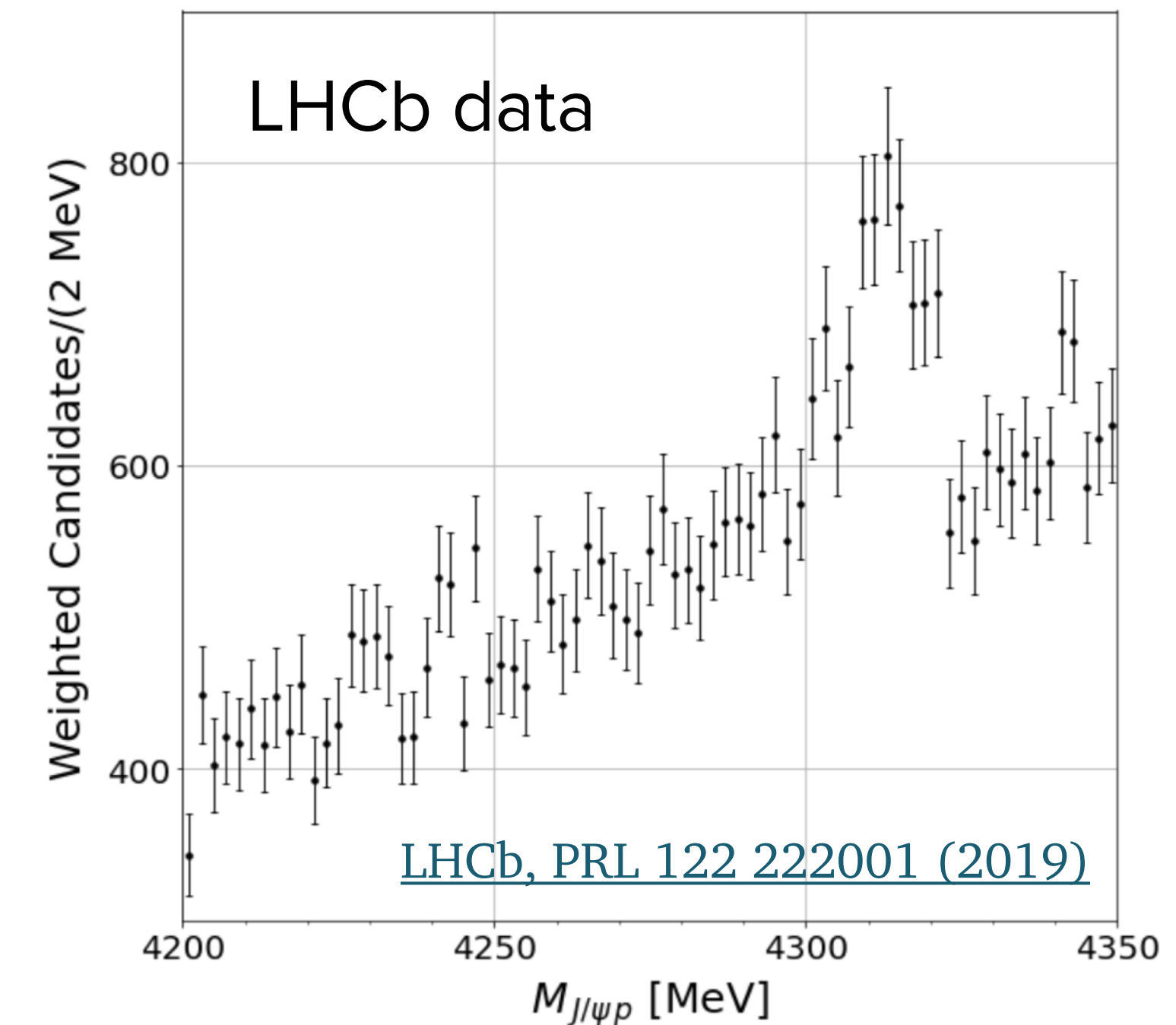


3. Design a set of DNN to solve the classification problem



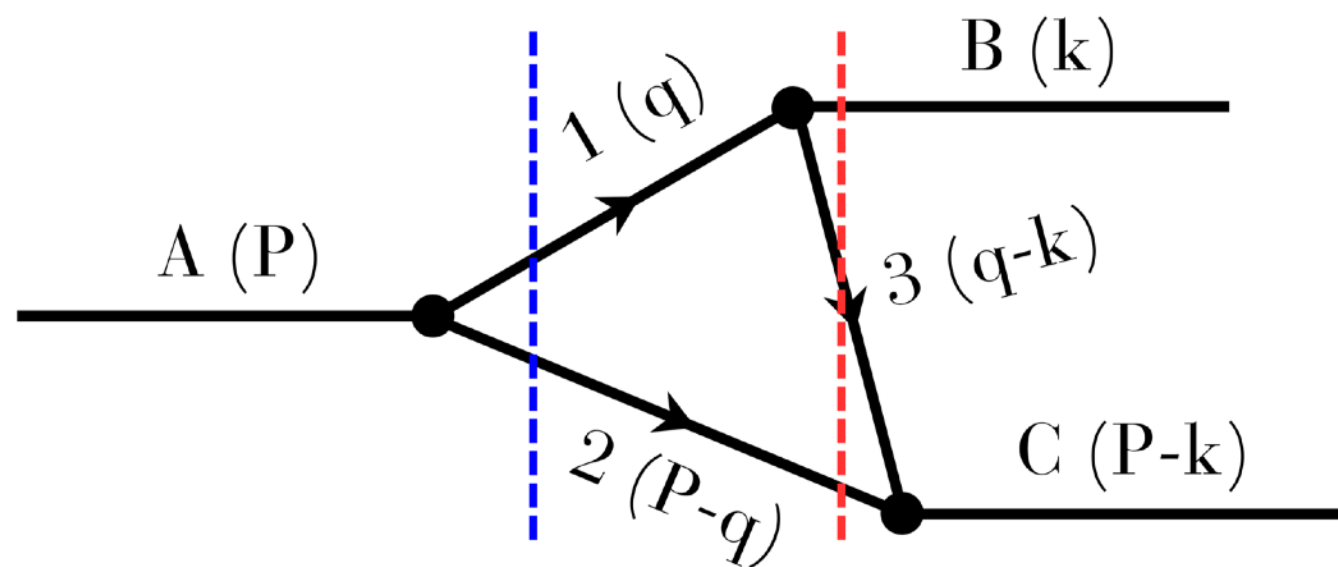
4. Train, test, and validate the DNN

5. Use the trained DNN to interpret the experimental data



TS vs Pole-based enhancements

DAO Co, VAA Chavez, DLBS,
PRD 110 (2024) 114034



The amplitude has no singularity (no pole)
No need to introduce new hadronic state to explain the enhancement.

$$I(k) \propto \int_0^\infty \frac{q^2 f(q) dq}{P^0 - \sqrt{m_1^2 + q^2} - \sqrt{m_2^2 + q^2} + i\epsilon}$$

M Bayar, F Aceti, FK Guo, E Oset, PRD 126 074039 (2016)

$$f(q) = \int_{-1}^1 \frac{dz}{E_C - \omega(q) - \sqrt{m_3^2 + q^2 + k^2 - 2qkz} + i\epsilon}$$

FK Guo, XH Liu, S Sakai, PPNP 122 103757 (2020)

$$m_1^2 \in \left[\frac{M_A^2 m_3 + M_B^2 m_2}{m_2 + m_3} - m_2 m_3, (M_A - m_2)^2 \right] \qquad m_C^2 \in \left[(m_2 + m_3)^2, \frac{M_A m_3^2 - M_B^2 m_2}{M_A - m_2} + M_A m_2 \right]$$

Mass condition - crucial in generating mock dataset

$$S_{11}(p_1,p_2) = \prod_m \frac{D_m(-p_1,p_2)}{D_m(p_1,p_2)}$$

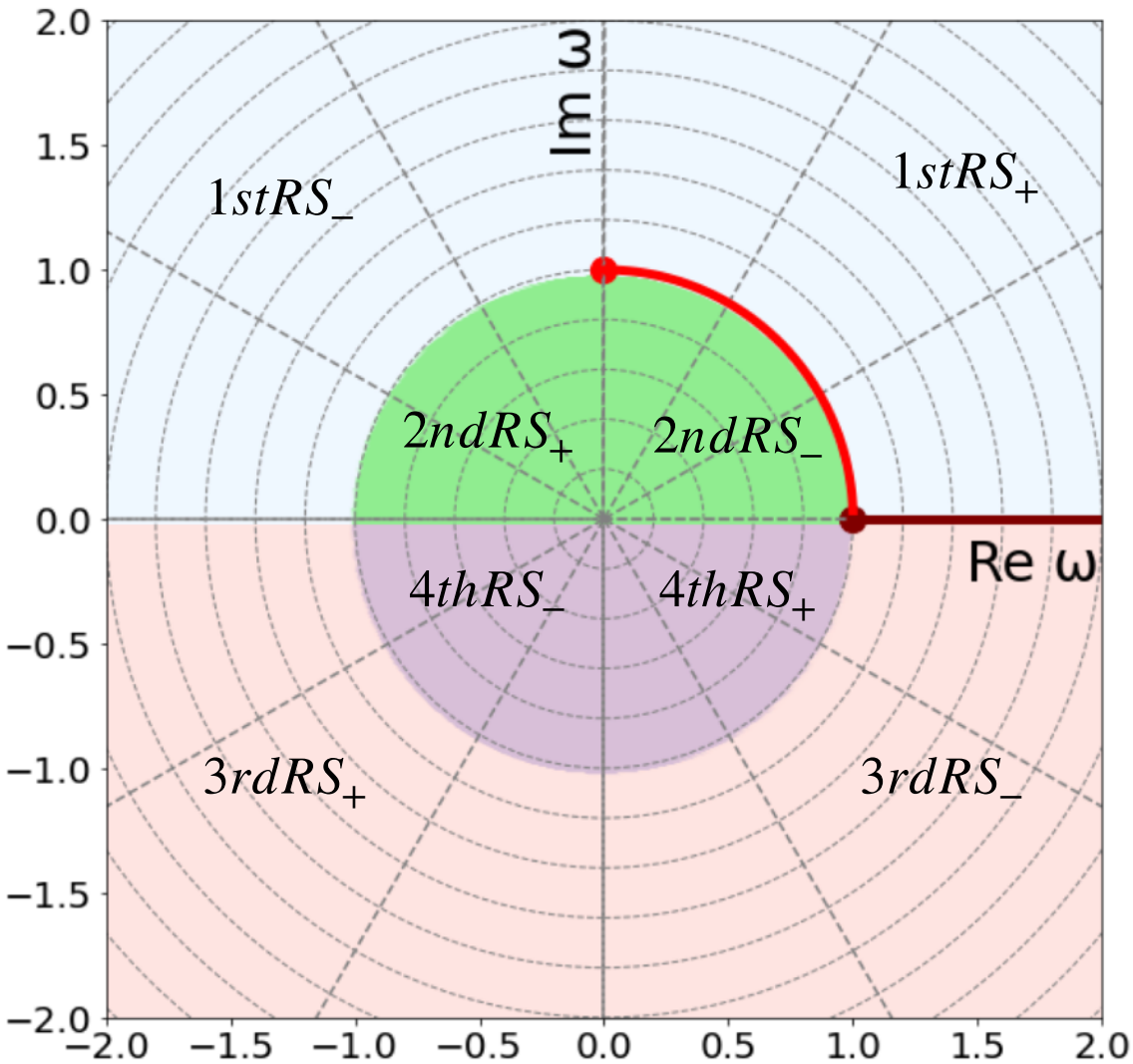
KJ Le Couteur, Proc. Roy. Soc (London) A256 (1960)
RG Newton J. Math. Phys. 2, 188 (1961)

$$p_k \rightarrow q_k; \quad s = q_k^2 + \epsilon_k^2; \quad \omega = \frac{q_1 + q_2}{\sqrt{\epsilon_2^2 - \epsilon_1^2}} \quad \frac{1}{\omega} = \frac{q_1 - q_2}{\sqrt{\epsilon_2^2 - \epsilon_1^2}}$$

One pole per $D_m(q_1, q_2)$: ω_m

M Kato, Ann. Phys., 31 1 (1965)

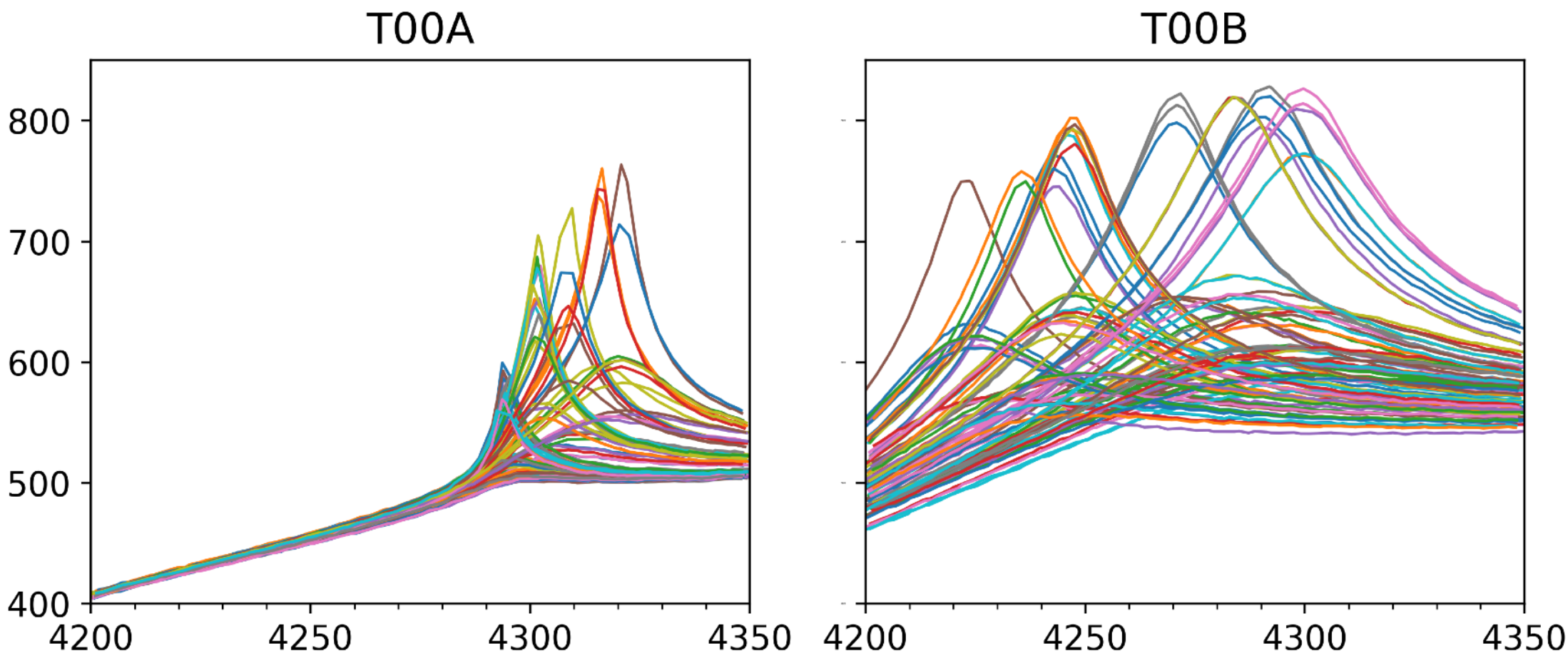
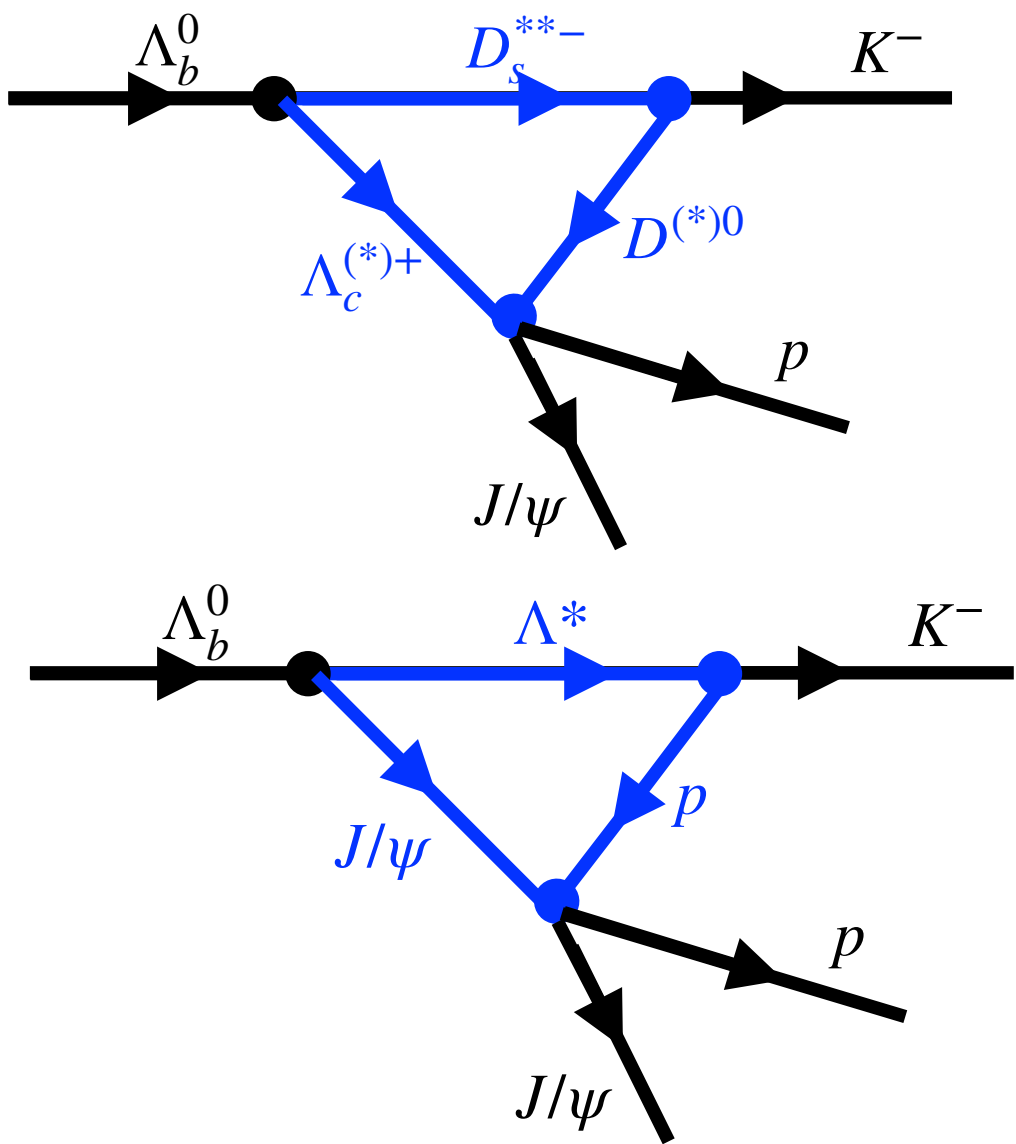
$$D_m(q_1, q_2) = D_m(\omega) = \frac{1}{\omega^2} (\omega - \omega_m) (\omega + \omega_m^*) (\omega - \omega_{\bar{m}}) (\omega + \omega_{\bar{m}}^*)$$



Sample training datasets

Triangle singularity

Parameter	Range of values [MeV]
$m_{\Lambda_b^0}$	5619.60 ± 0.17
m_{K^-}	493.677 ± 0.016
$m_{\Lambda_c^+}$	2286.46 ± 0.14
$m_{\bar{D}^{*0}}$	2006.85 ± 0.05
$m_{J/\psi}$	3096.9 ± 0.006
$m_{D_s^{*-}}$	$[3209.80, 3315.00]$
m_{Λ^*}	$[2490.00, 2522.70]$
Λ	$[2000.0, 2500.0]$
ϵ	$[1.0, 10.0]$

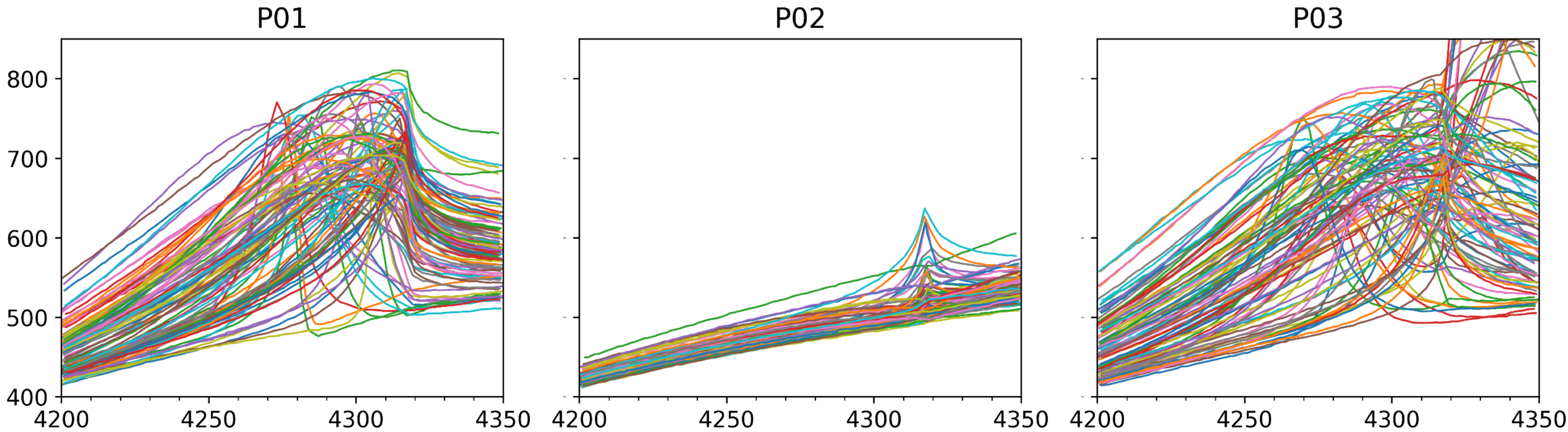


Pole-based enhancements

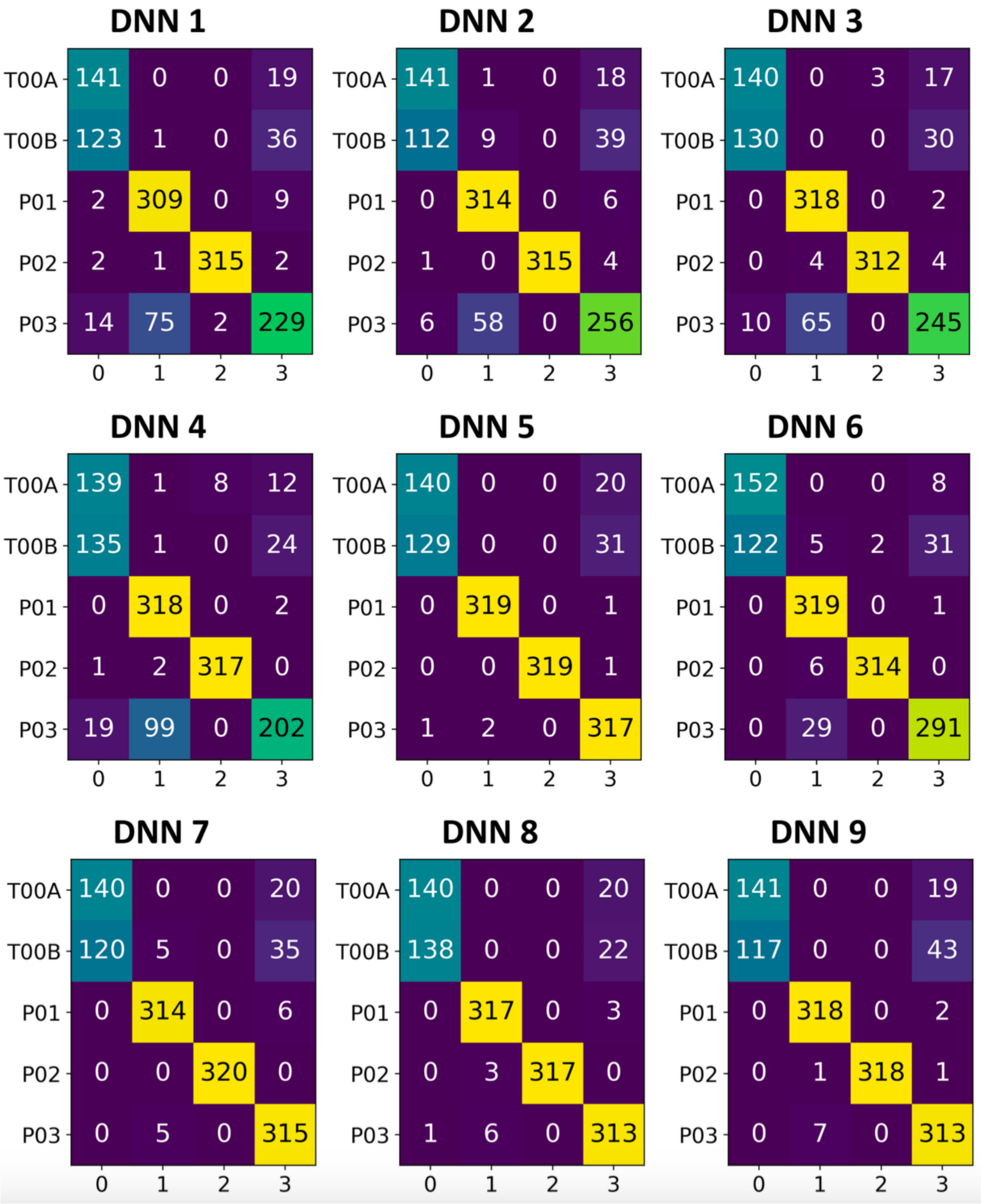
$$\begin{cases} T_2 - 50 \leq \text{Re } E_{\text{pole}} \leq 4350 & \text{all RS} \\ -100 \leq \text{Im } E_{\text{pole}} < 0 & [bt] \text{ \& } [bb] \\ 0 < \text{Im } E_{\text{pole}} \leq 100 & [tb] \end{cases}$$

$4 \times 10,000$
Training dataset generated

4×320
Testing dataset generated

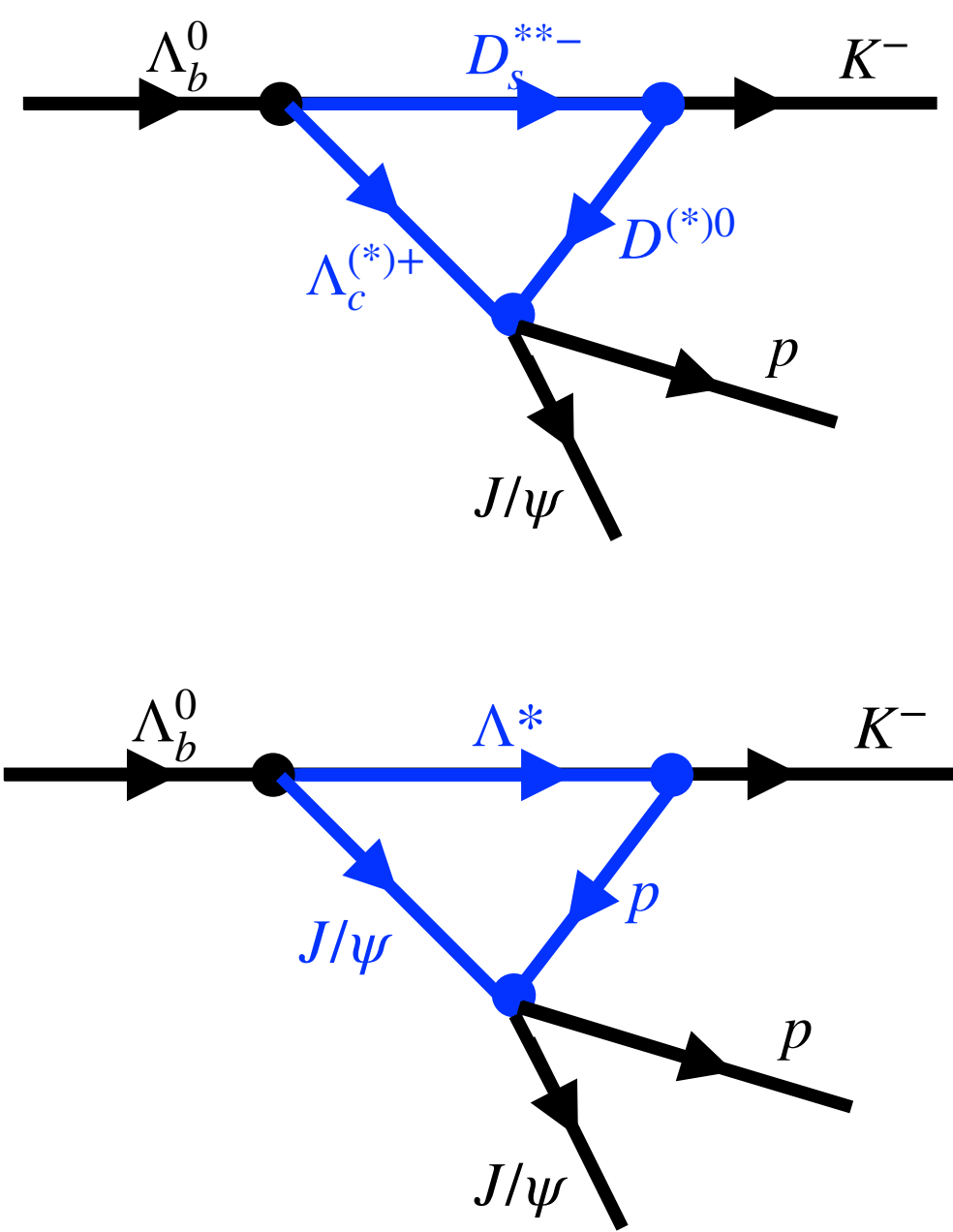


Confusion Matrix



Generate a different set of 320 validation dataset per classification.

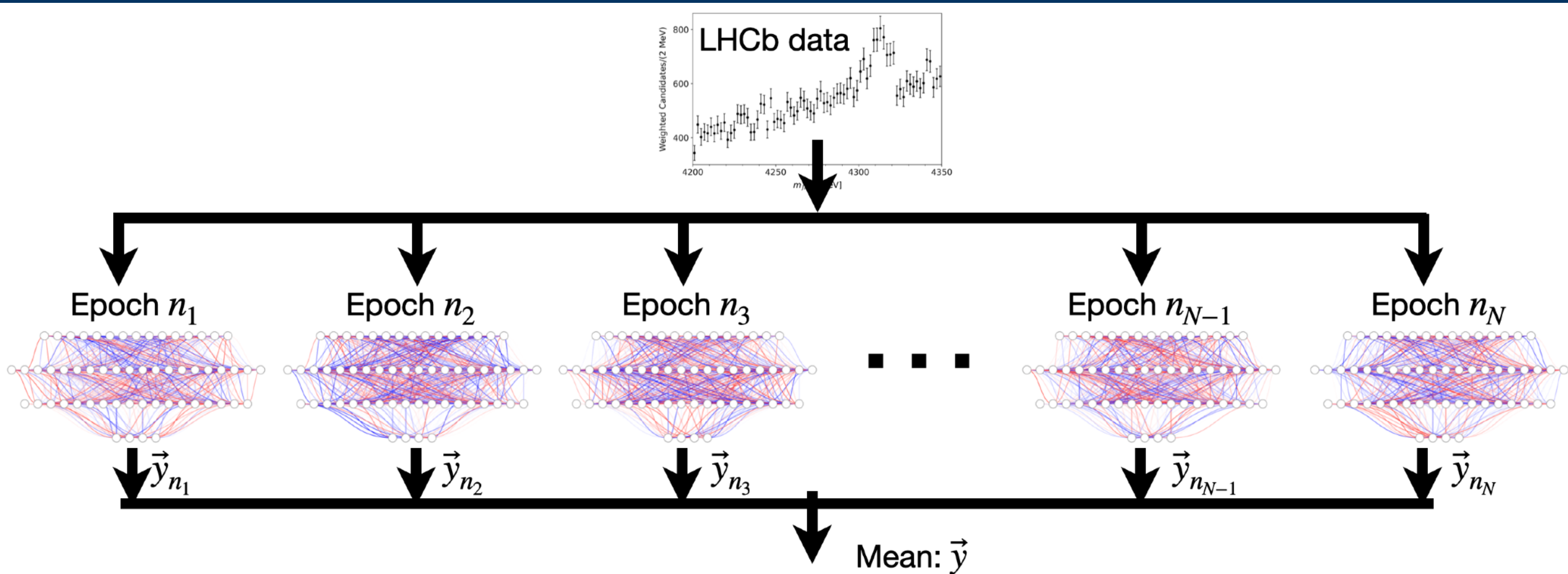
Label	Description
0	Triangle mechanism
1	1 pole in 2nd RS
2	1 pole in 4th RS
3	1 pole in 2nd and 1 pole in 3rd RSs



Slight confusion between TS and BW-like line shapes.

If the experimental data will favor either the TS or class 3, further analysis must be done.

Inference stage (snapshot ensemble)



- Bootstrapped 3000 line shapes from the experimental data. (Uniform distribution)
- Feed to the DNN state at epoch n
- Get the mean count

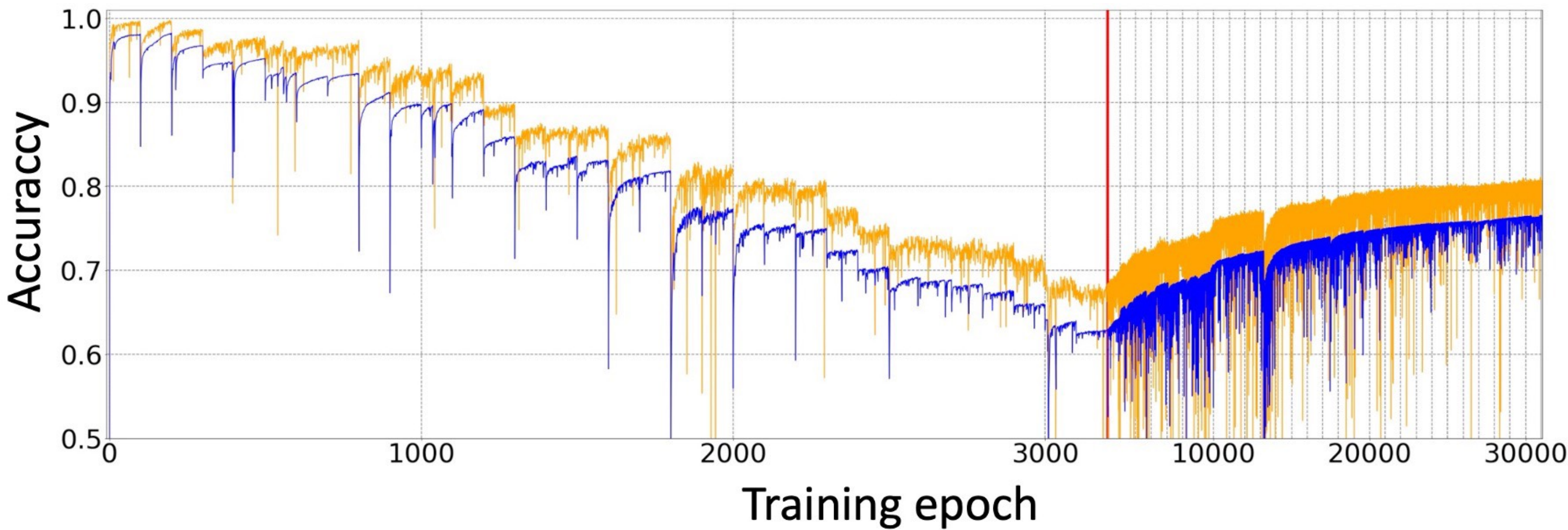
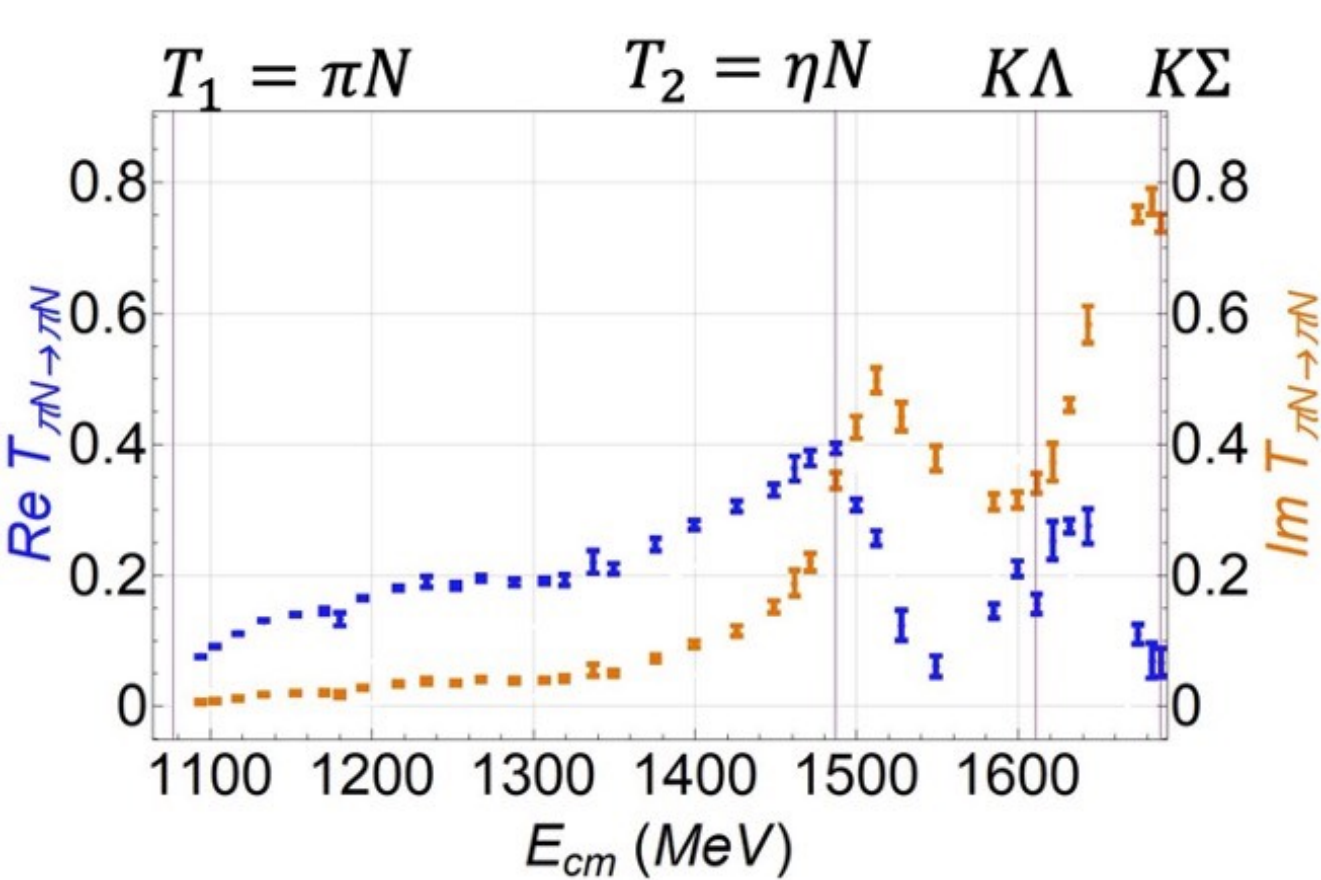
- TS is ruled out by pure line shape analysis!
- It is possible to distinguish kinematical cusp vs pole despite the presence of experimental uncertainty.

The data favored the pole-based interpretation: 1 pole in 2nd RS

Model	TS	1 pole in RS2	1 pole in RS4	1 pole each in RS2 & RS4
DNN 1	261	1930	0.248	804
DNN 2	18.4	1998	969	14.5
DNN 3	0.653	2960	0	36
DNN 4	0	2990	0	14.2
DNN 5	0	2660	0	337
DNN 6	0.436	2999	0	0.891
DNN 7	1	2590	0	414
DNN 8	0	2999	0	1.06
DNN 9	0	2970	0	127

Probe deeper: Pole structure of $P_{c\bar{c}}(4312)^+$

DNN designed to probe the pole structure is difficult to train.



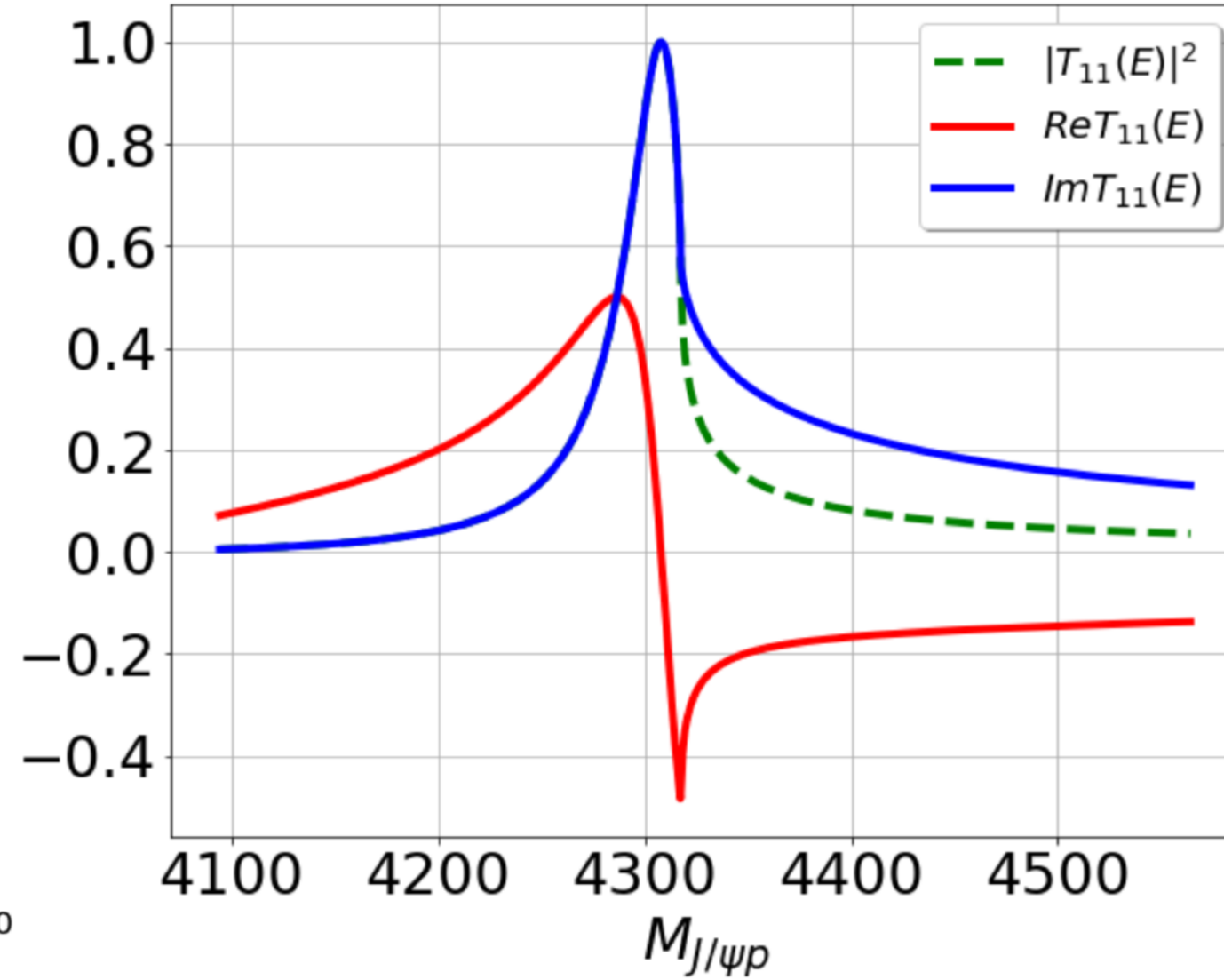
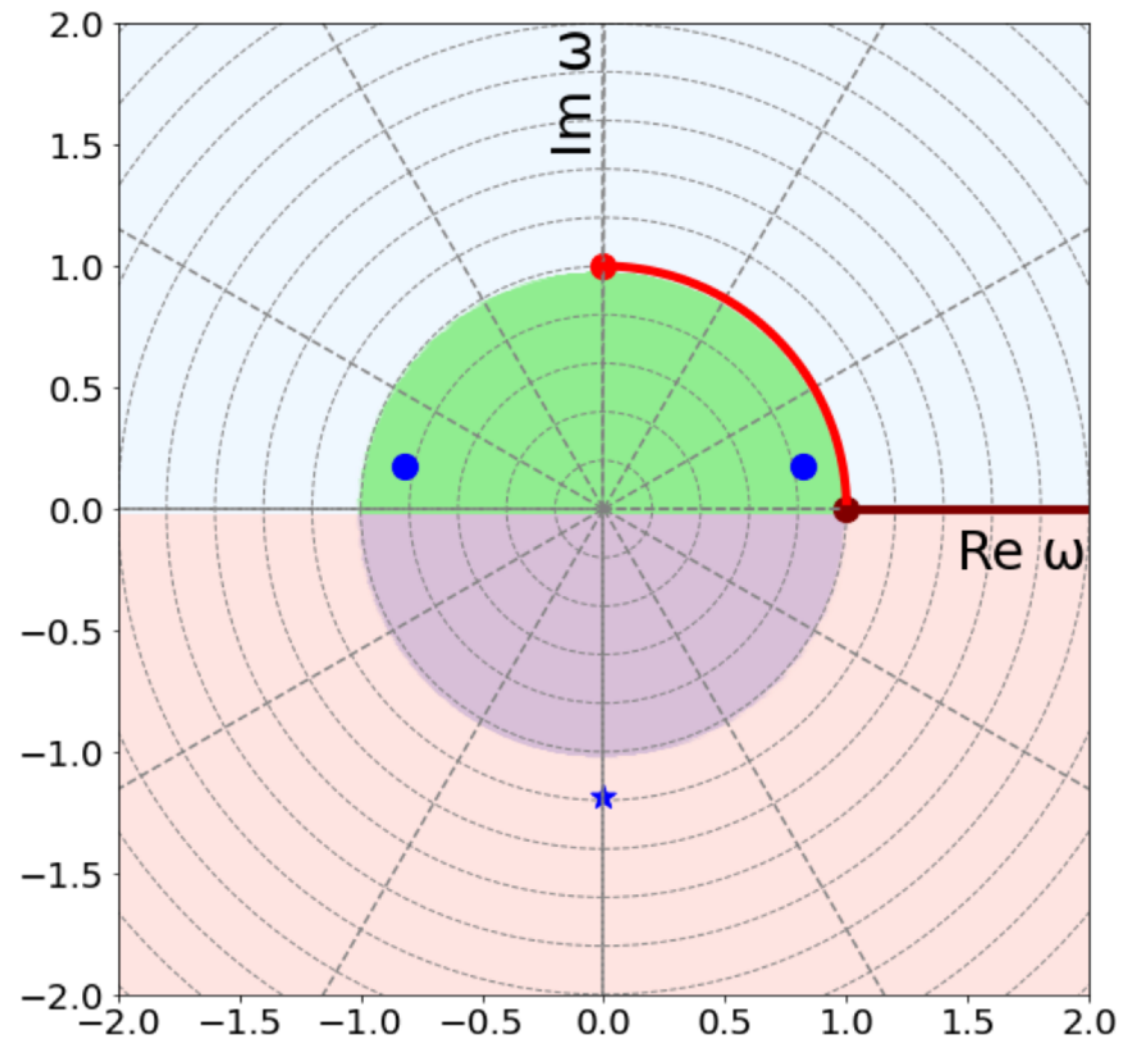
Label	S-matrix pole configuration
0	no nearby pole
1	1 pole in $[bt]$
2	2 poles in $[bt]$
$\vdots$	$\vdots$
32	1 pole in $[bt]$, 2 poles in $[bb]$ and 1 pole in $[tb]$
33	1 pole in $[bt]$, 1 pole in $[bb]$ and 2 poles in $[tb]$
34	1 pole in $[bt]$, 1 pole in $[bb]$ and 1 pole in $[tb]$

[DLBS, YI, TS, AH PRD 104 3 036001 \(2021\)](#)

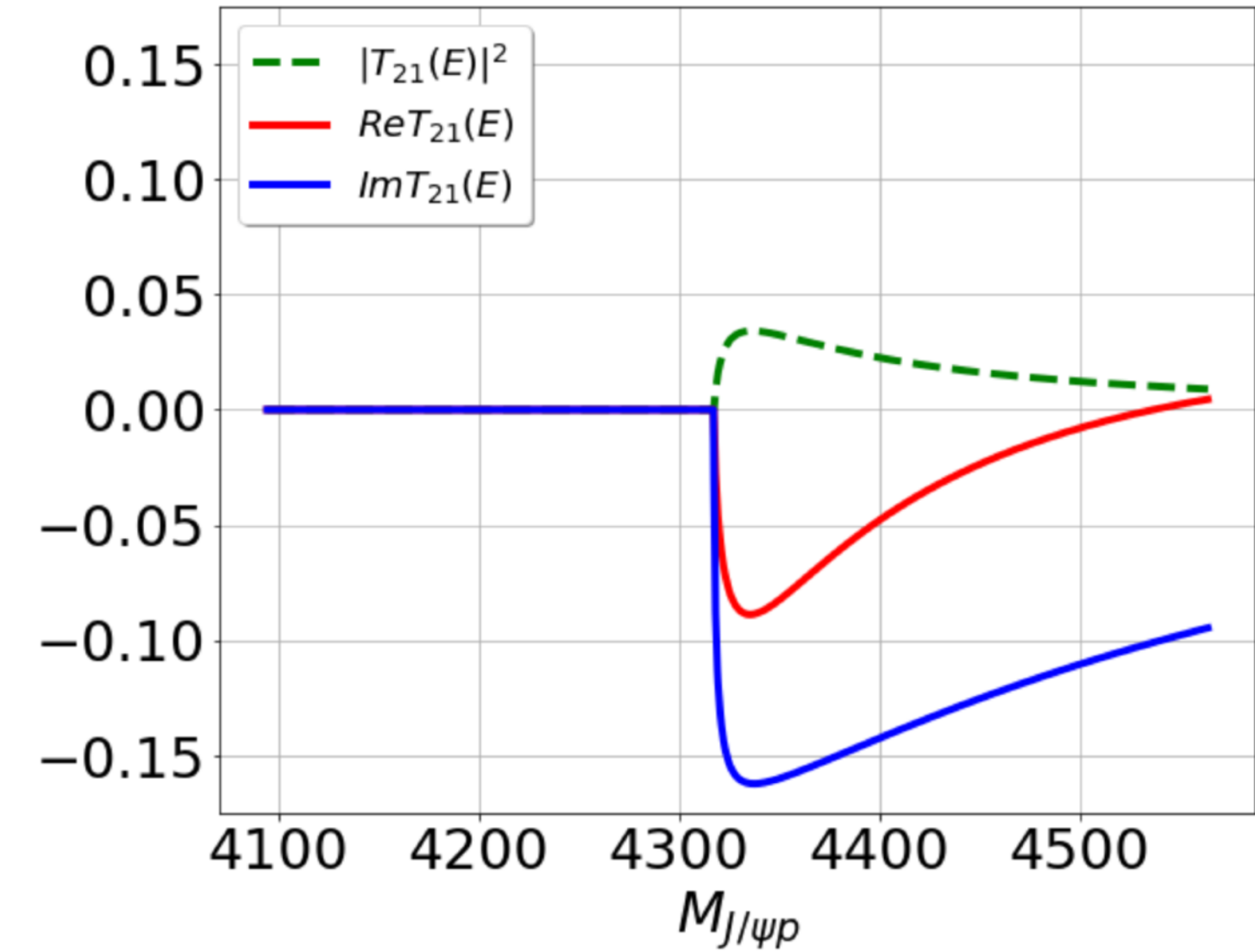
Line shape - pole structure map, intrinsically ambiguous.
Complications introduced by pole-shadow pair.

Probe deeper: Pole structure of $P_{c\bar{c}}(4312)^+$

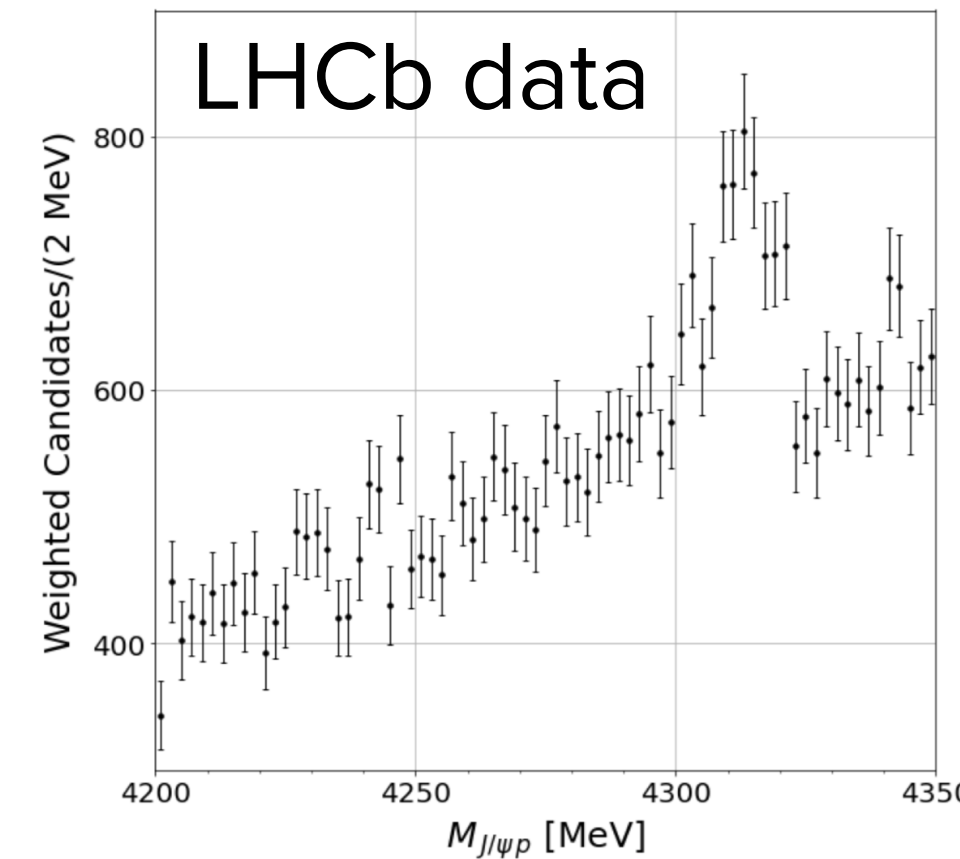
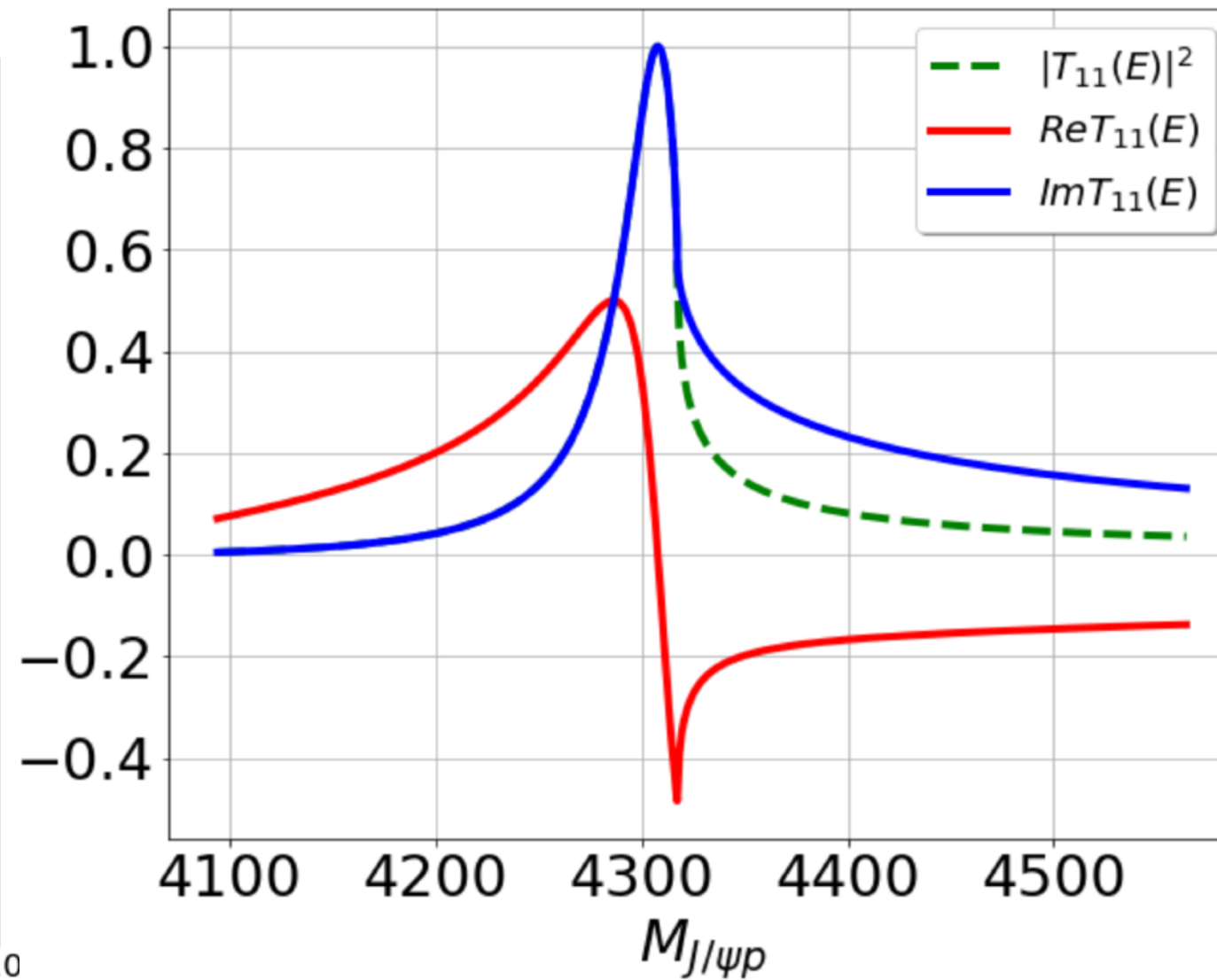
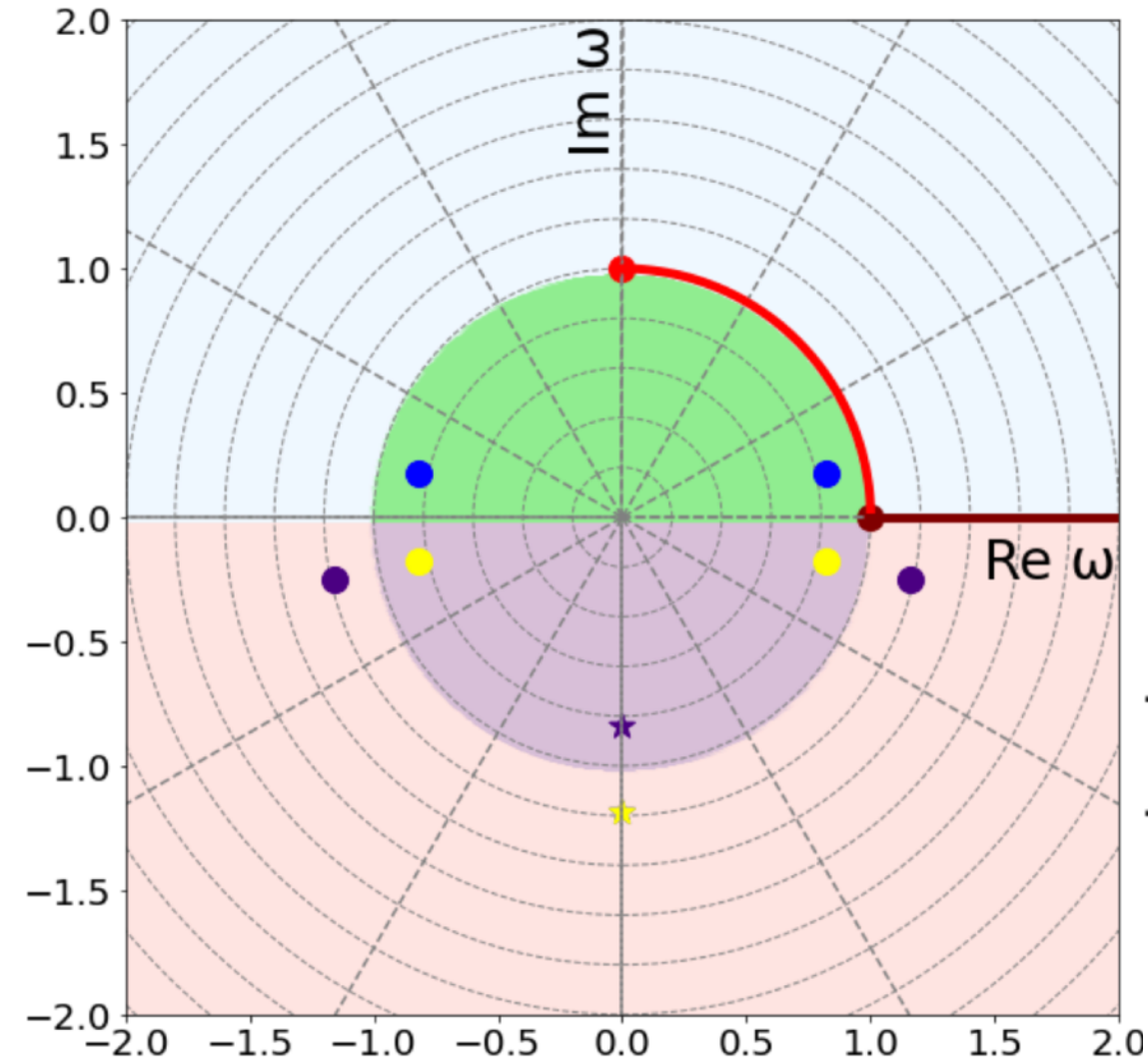
1 pole in the 2nd RS



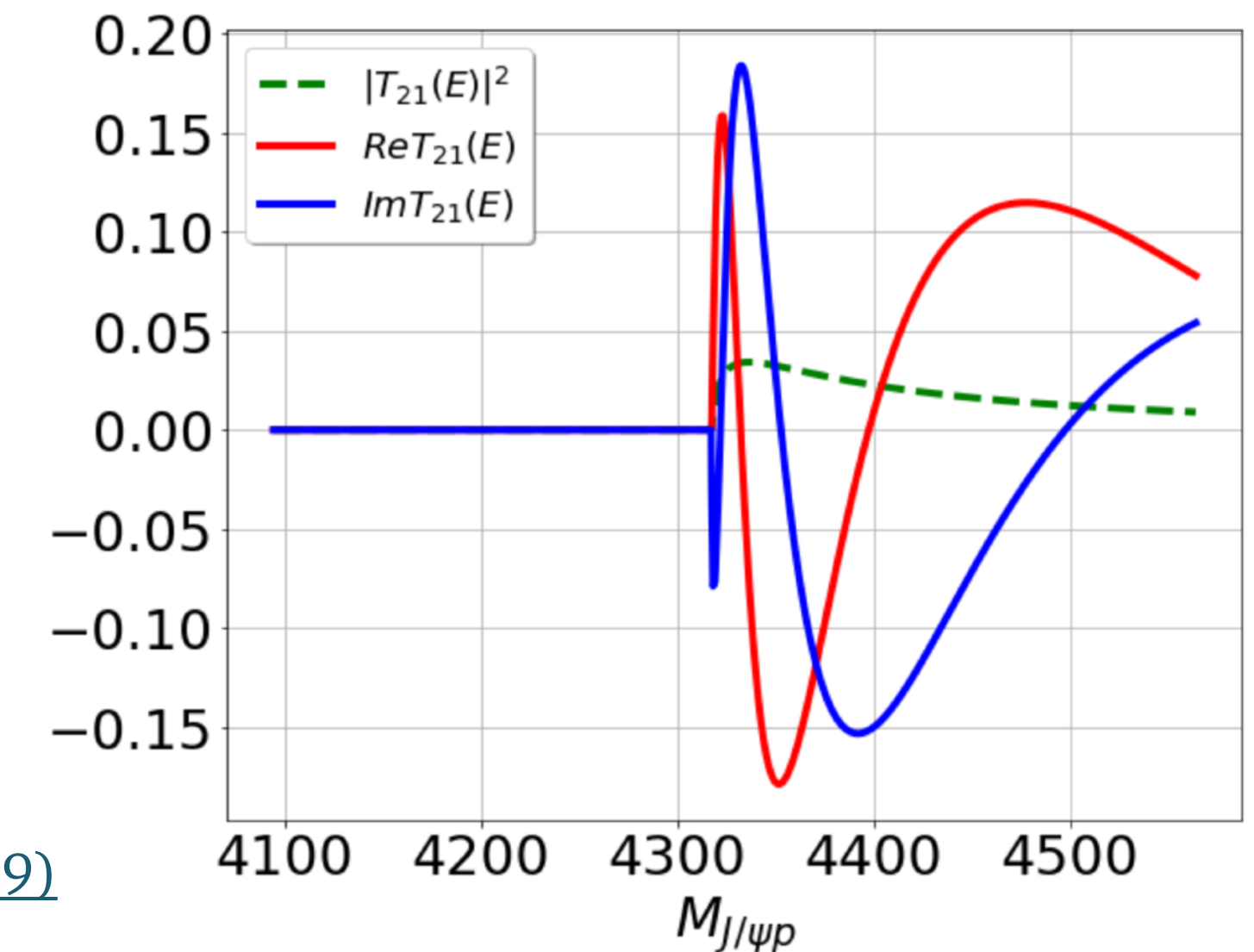
It is not enough
to use one
element of the
T-matrix.
We need to
include the off-
diagonal element.



1 pole in each RS

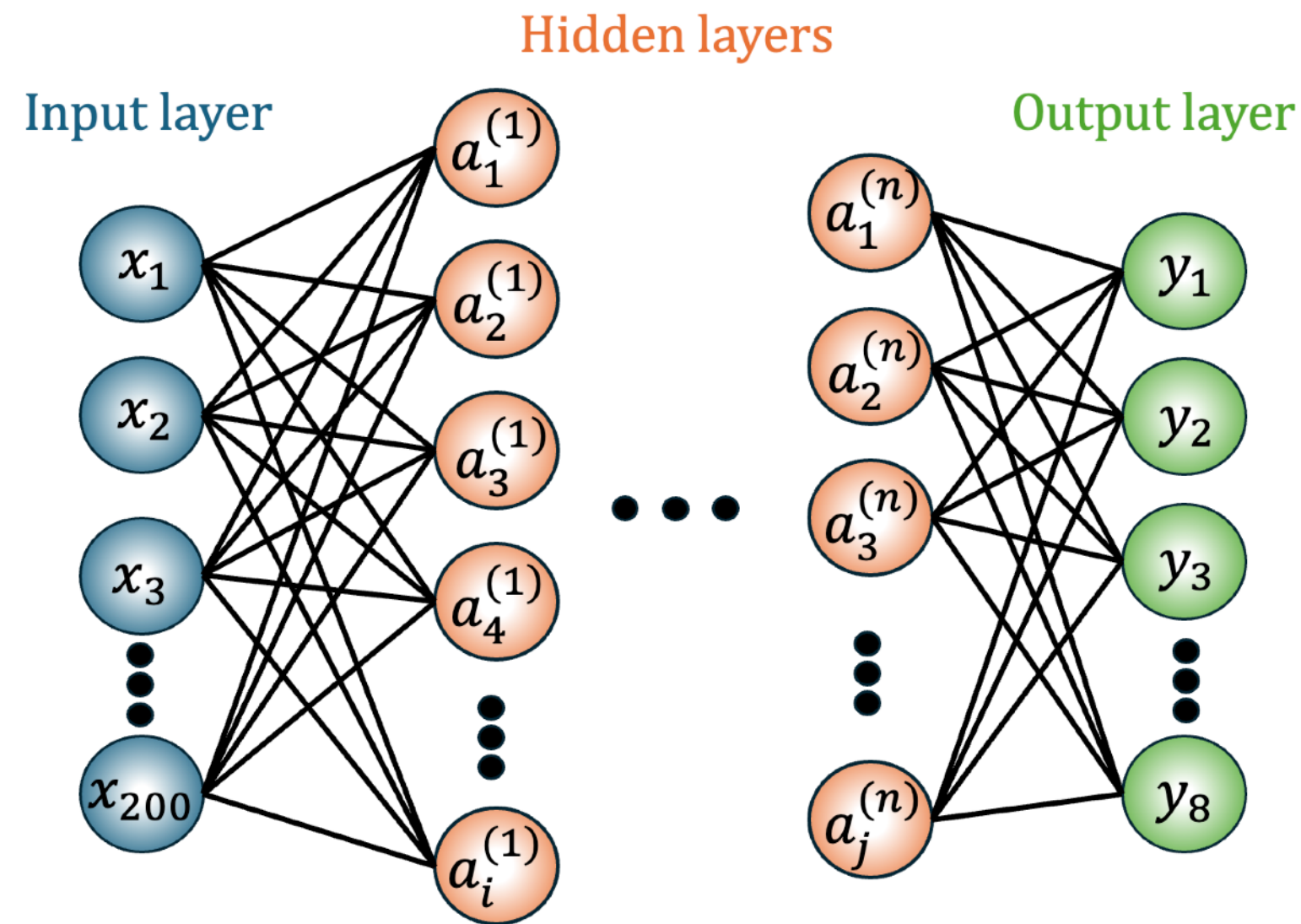
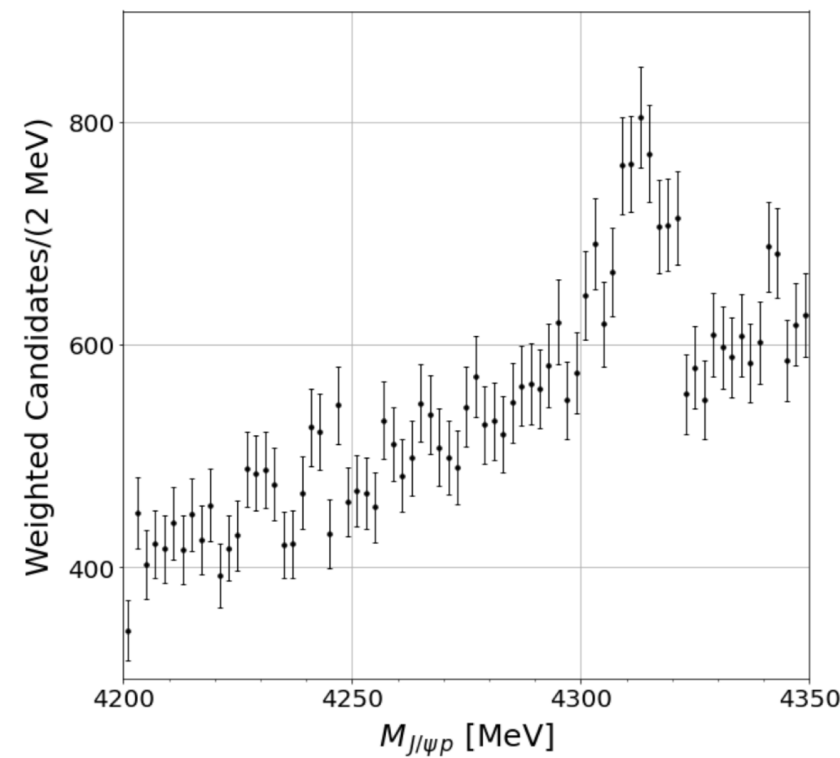


[LHCb, PRL 122 222001 \(2019\)](#)



[LMS, DLBS PRC 108 045204 \(2023\)](#)

Probe deeper: Pole structure of $P_{c\bar{c}}(4312)^+$



Class label	S -matrix pole configuration
0	1 pole on $[bt]$
1	1 pole on $[bb]$
2	1 pole on $[tb]$
3	1 pole on $[bt]$ and 1 pole on $[bb]$
4	1 pole on $[bb]$ and 1 pole on $[tb]$
5	1 pole on $[bb]$, 1 pole on $[tb]$, 1 pole on $[bt]$
6	2 poles on $[bb]$, 1 pole on $[tb]$
7	1 pole on $[bb]$, 2 poles on $[tb]$

$$\frac{dN}{d\sqrt{s}} = \rho(s) \left[|F(s)|^2 + B(s) \right];$$

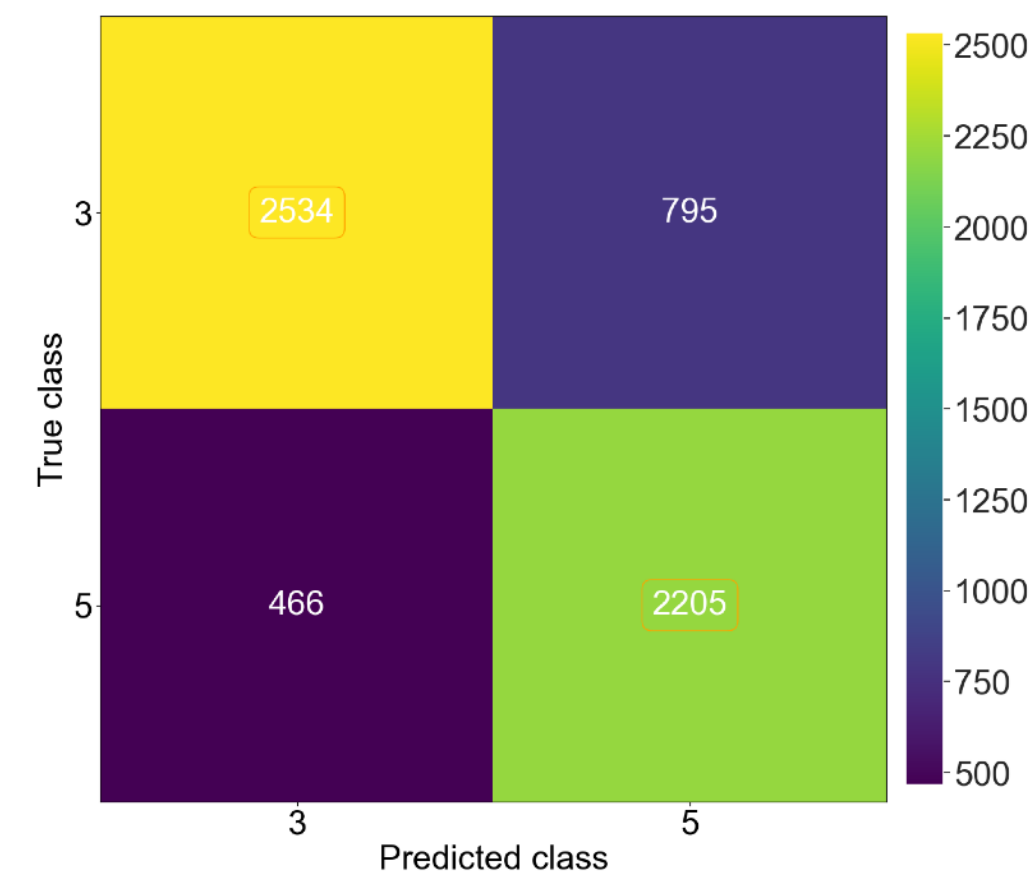
$$F(s) = \alpha_1 T_{11}(s) + \alpha_2 T_{21}(s)$$

$$S_{11}(\omega) = \prod_m \frac{D_m(-1/\omega)}{D_m(\omega)} \quad S_{22}(\omega) = \prod_m \frac{D_m(1/\omega)}{D_m(\omega)}$$

$$S_{11}S_{22} - S_{12}S_{21} = \prod_m \frac{D_m(-\omega)}{D_m(\omega)} \quad \hat{S} = \hat{1} + 2i\hat{T}$$

Class 3:
1 pole in 2nd RS and
1 pole in the 3rd

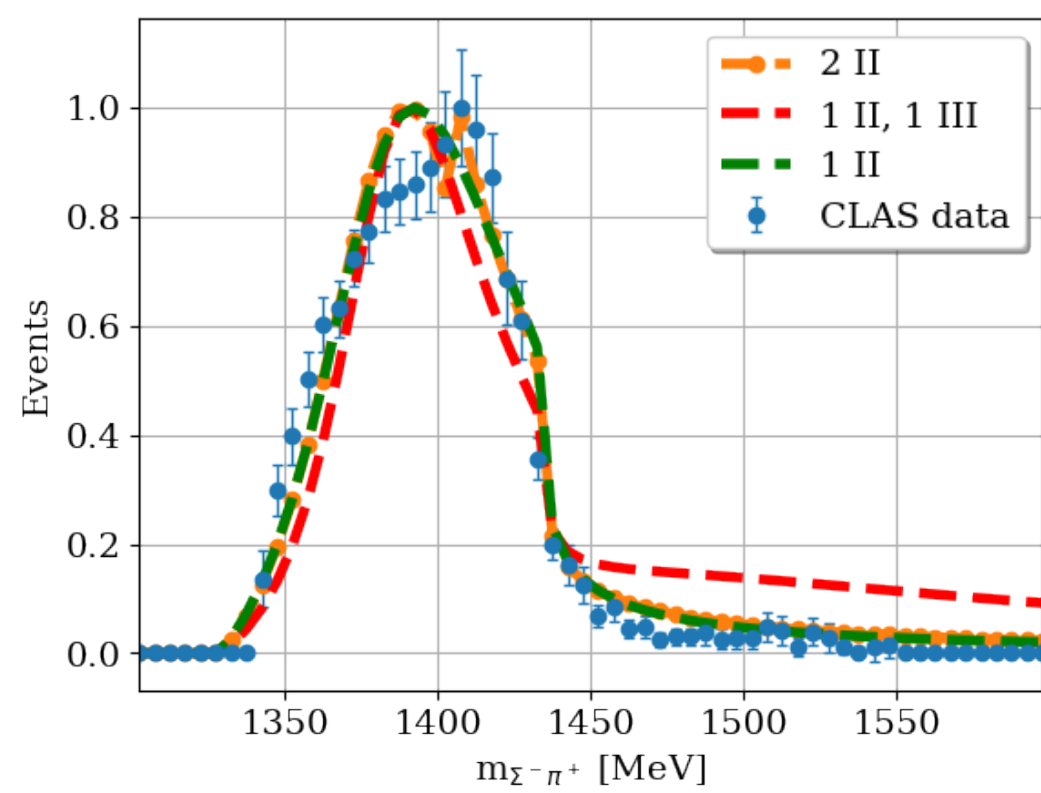
Class 5:
1 pole in each
unphysical RS



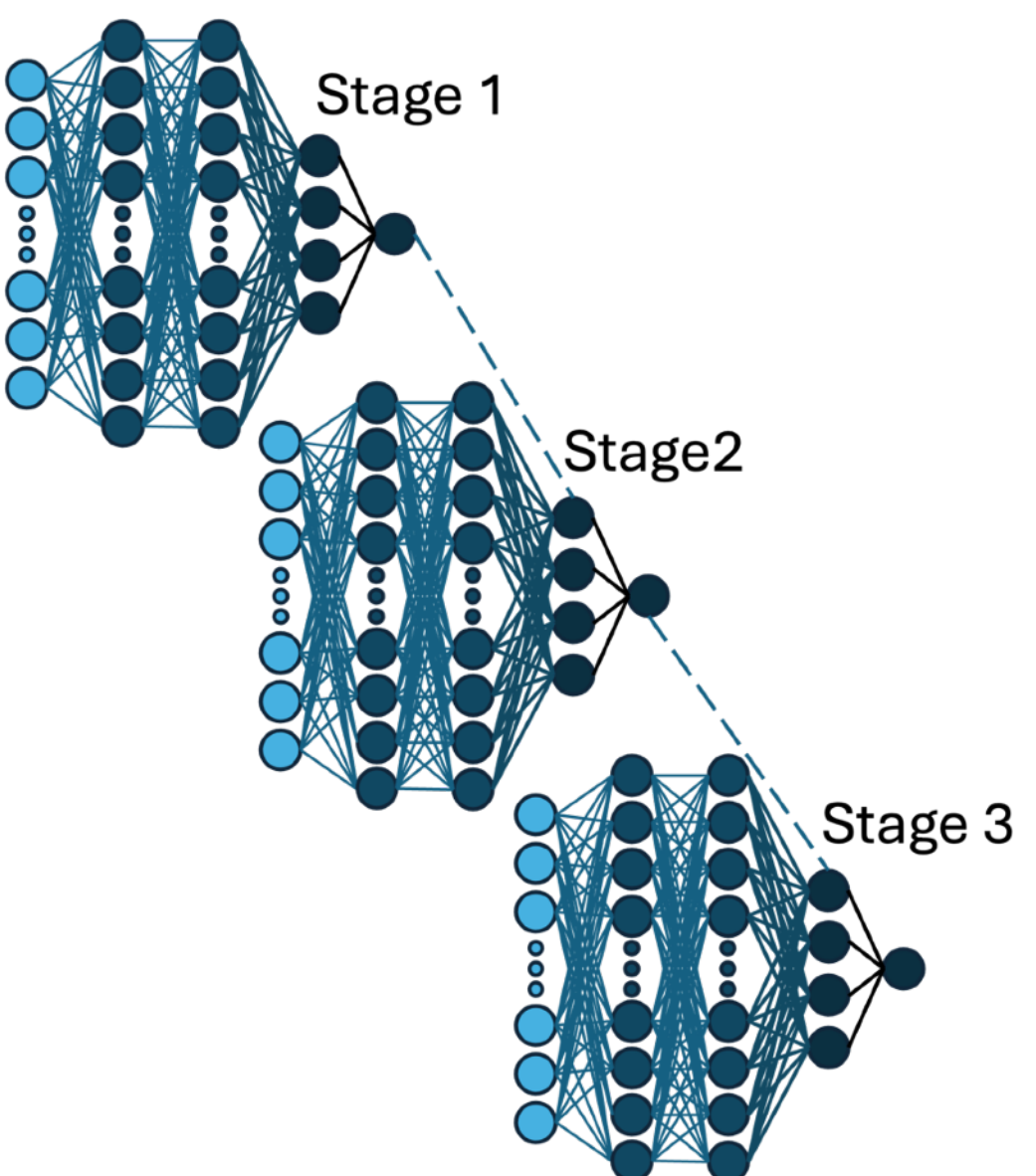
Pole in 2nd and 3rd RS
-resemble pole-shadow pair
-possible true resonance
decaying into $J/\psi p$

Pole in 4th RS
-Virtual state of $\Sigma_c \bar{D}$
coupled to $J/\psi p$

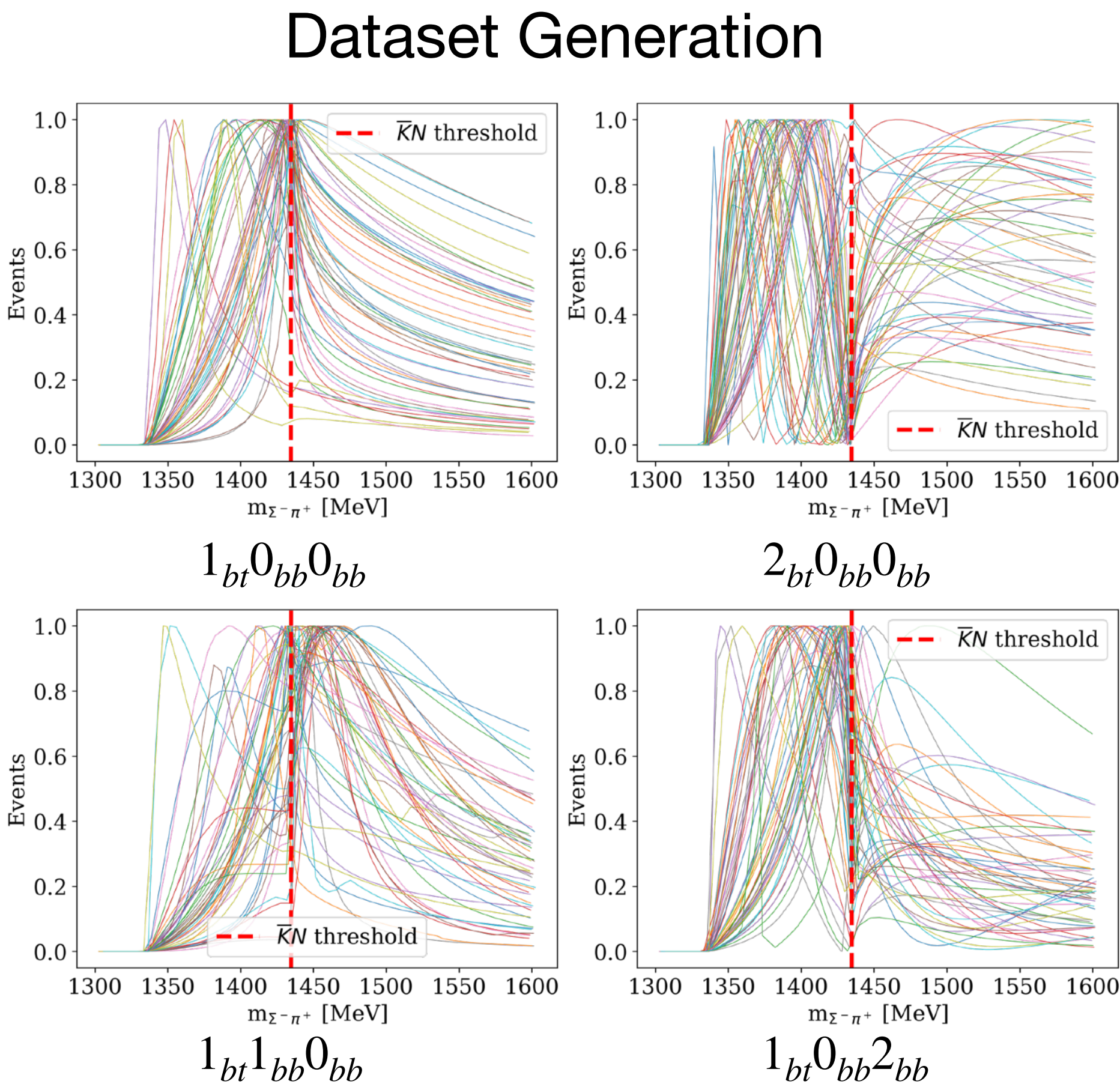
Pole structure of $\Lambda(1405)$: Preliminary Results



Ambiguous Pole Structures

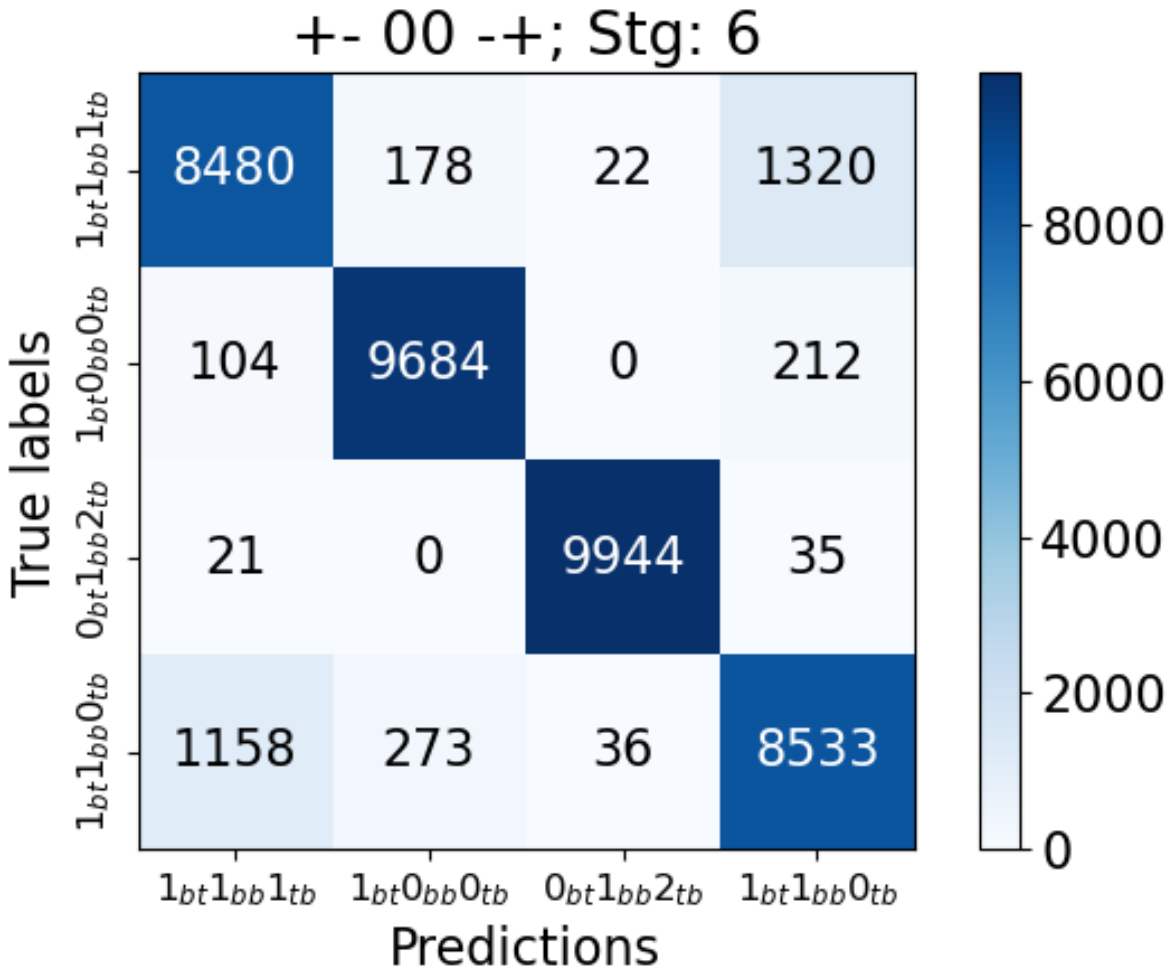


Stage Training

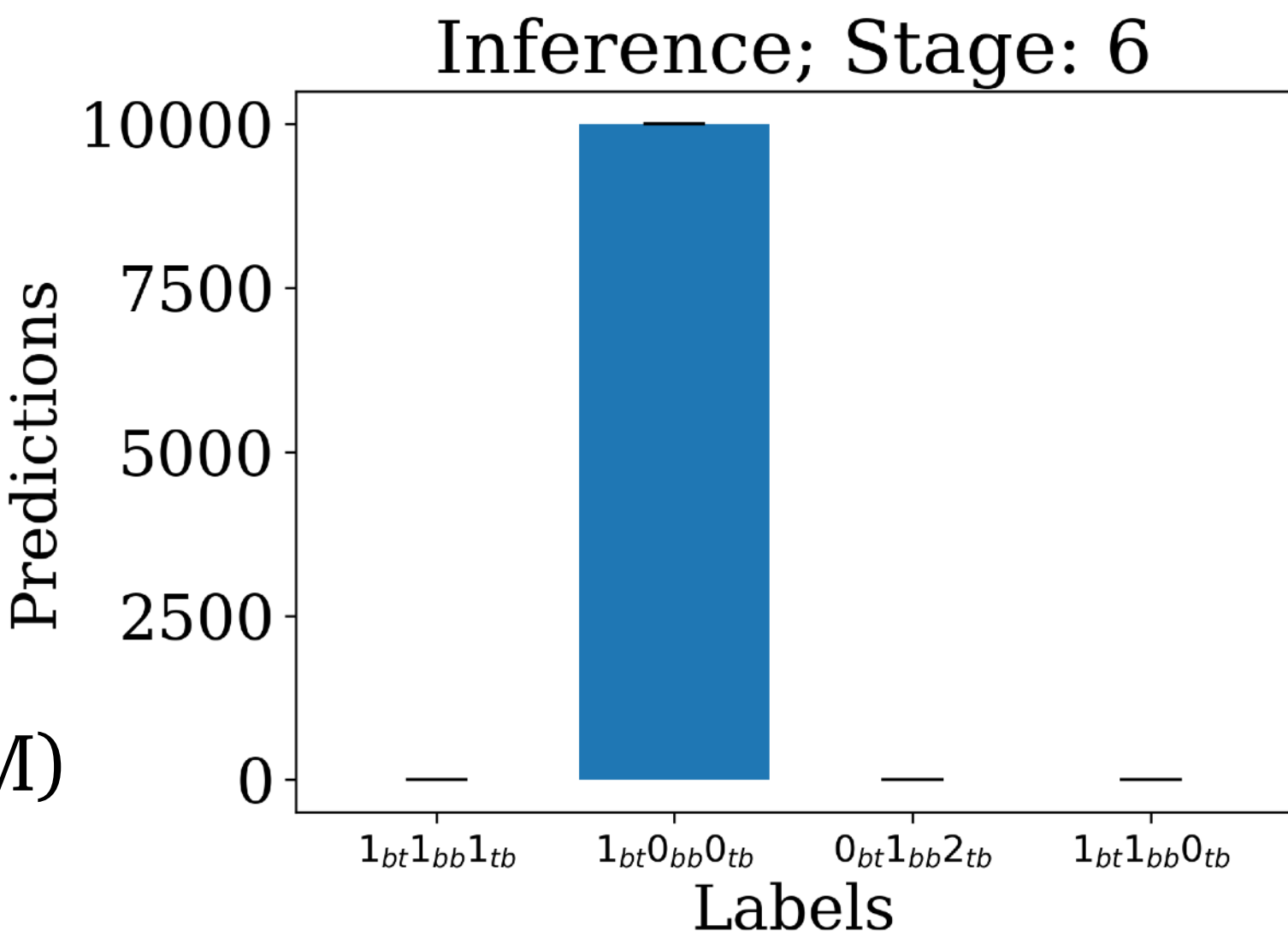


Two-pole (on the same RM) cannot be confirmed.

[VAA Chavez, DLBS, In preparation](#)

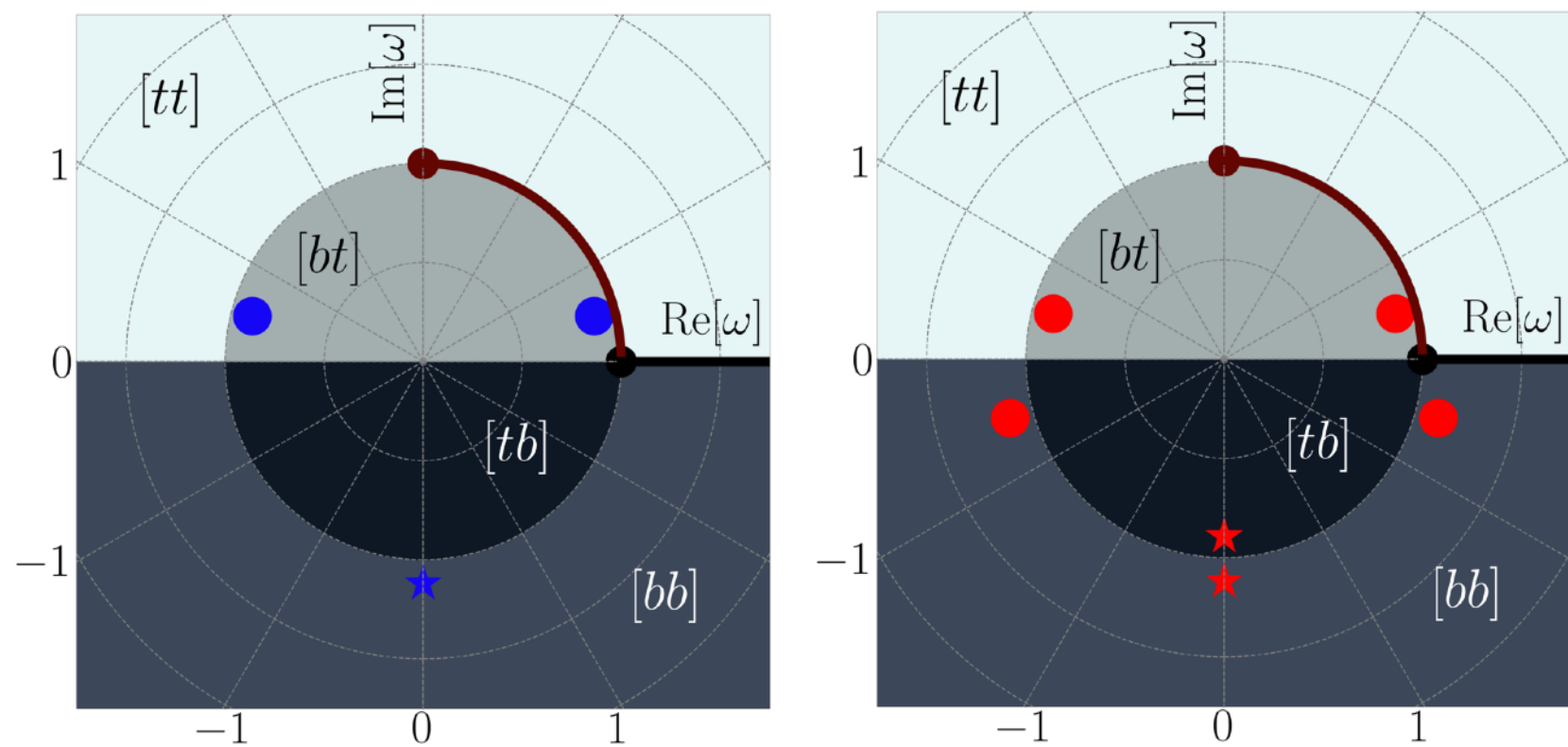
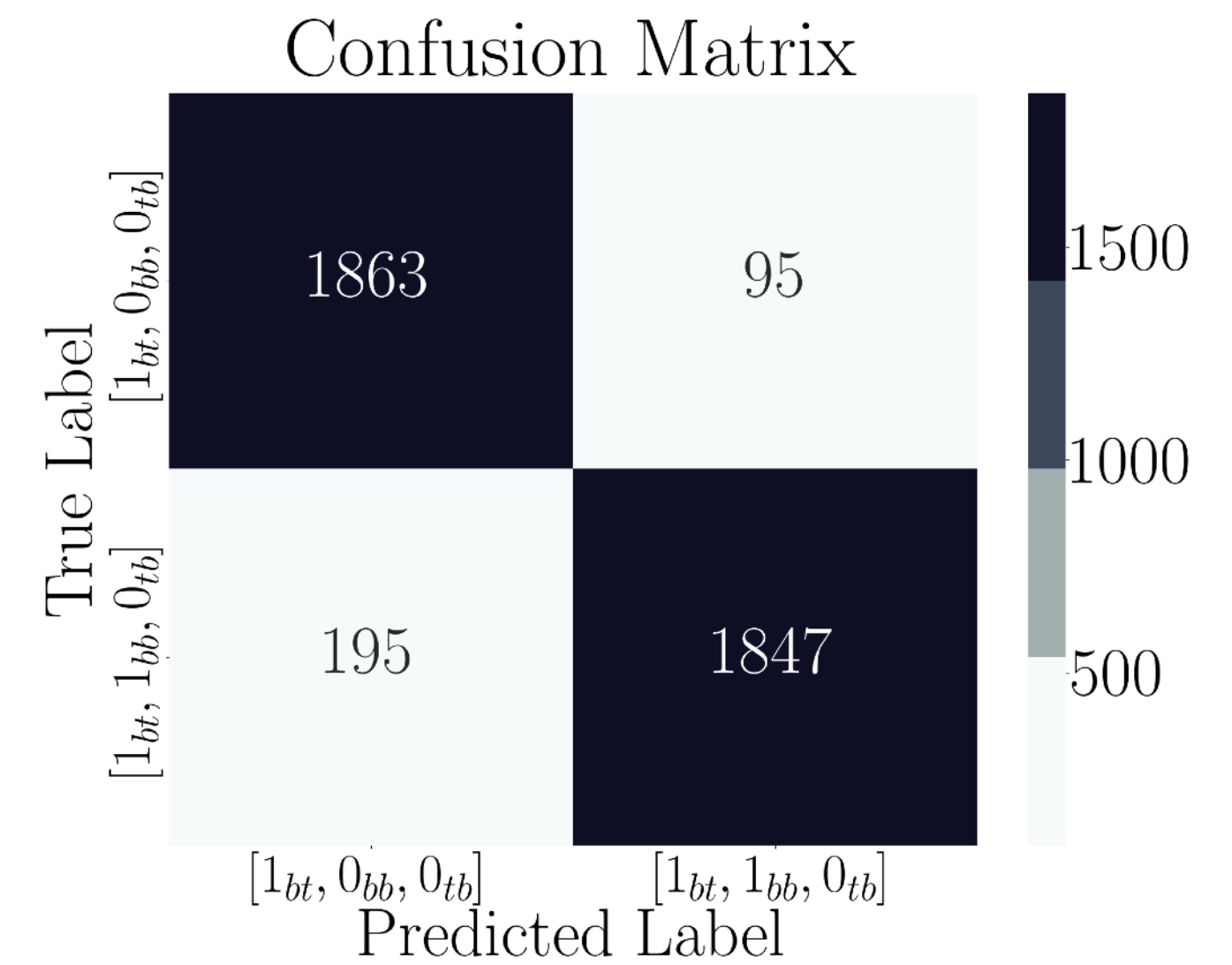
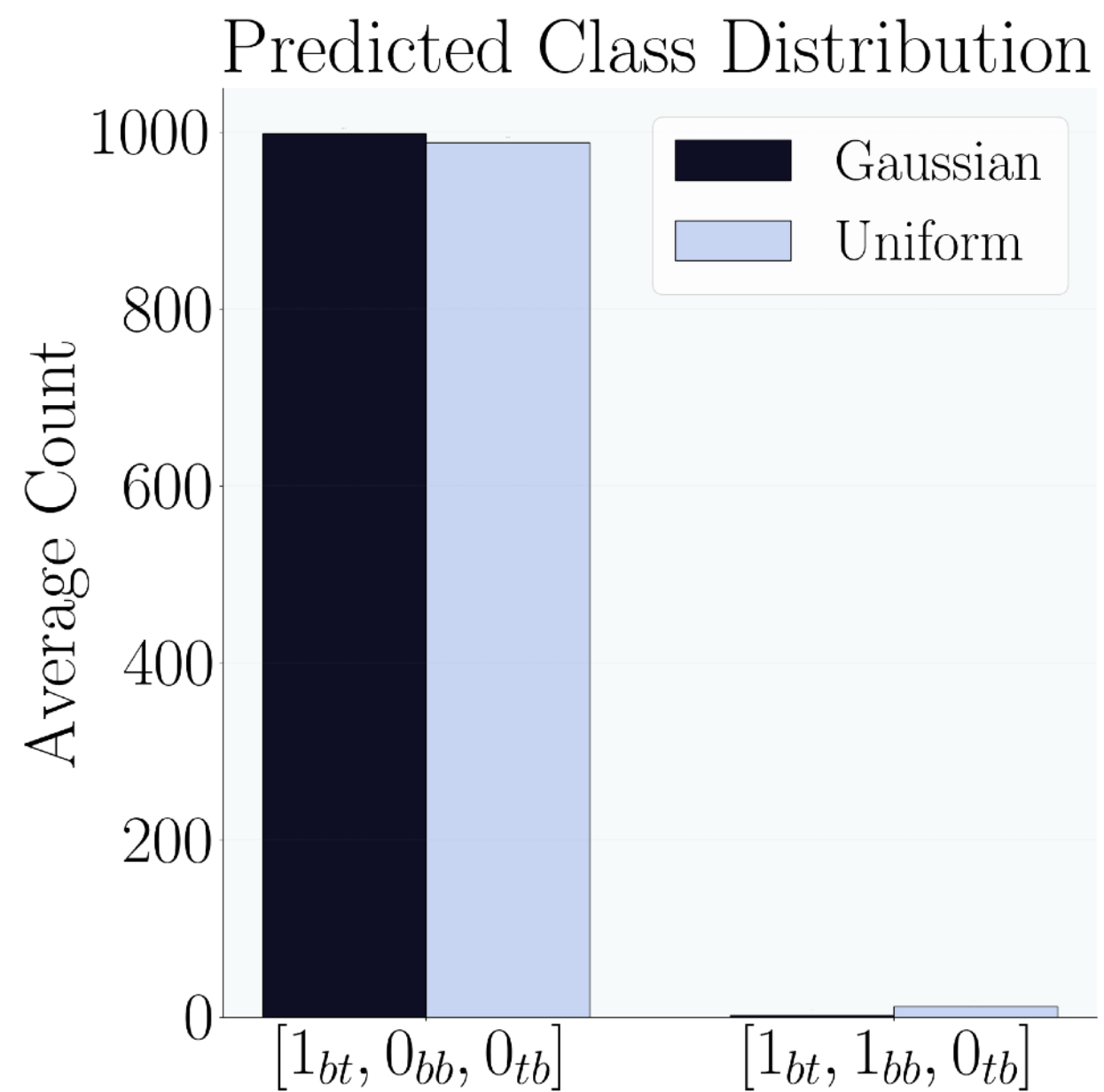
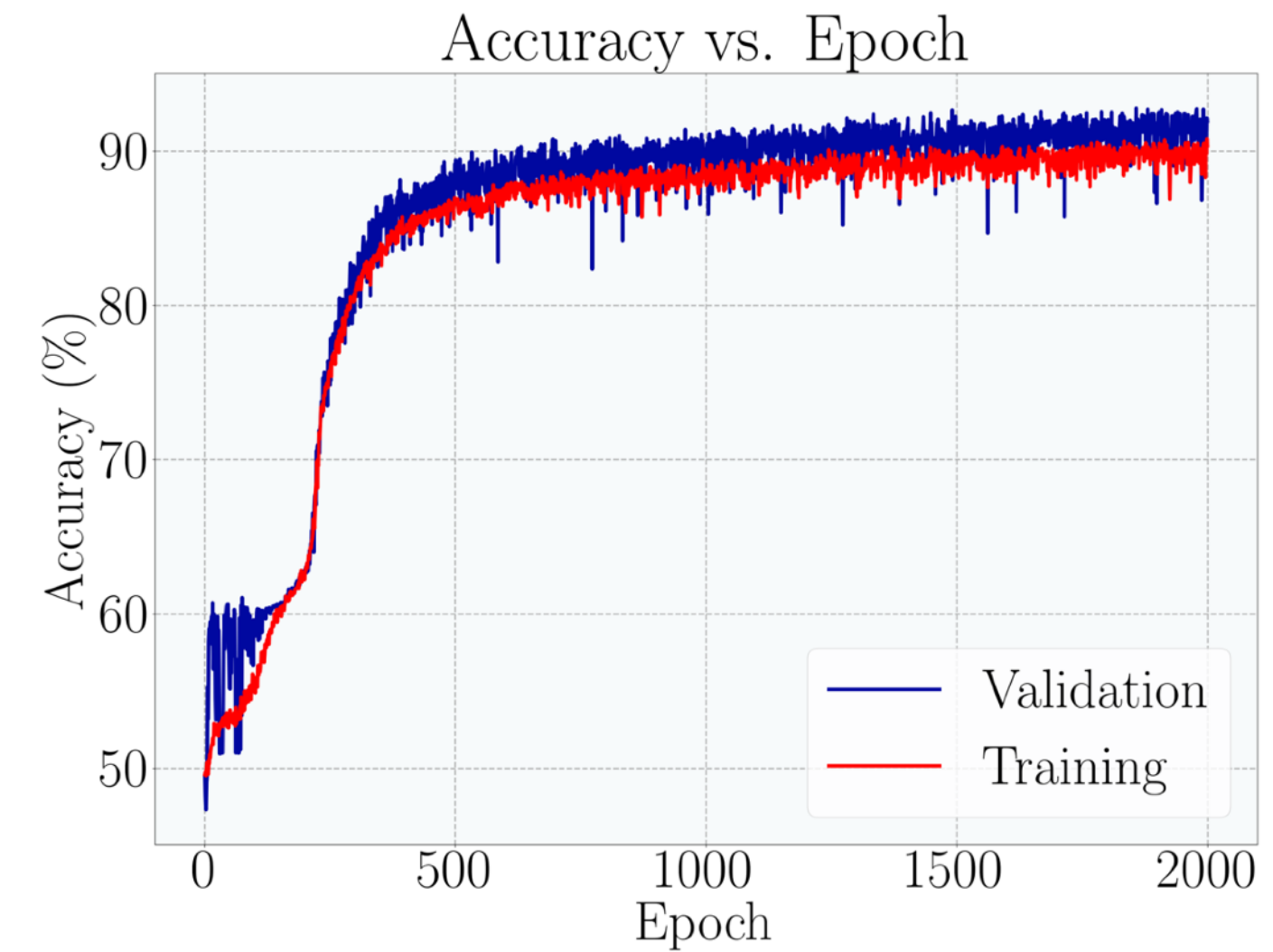
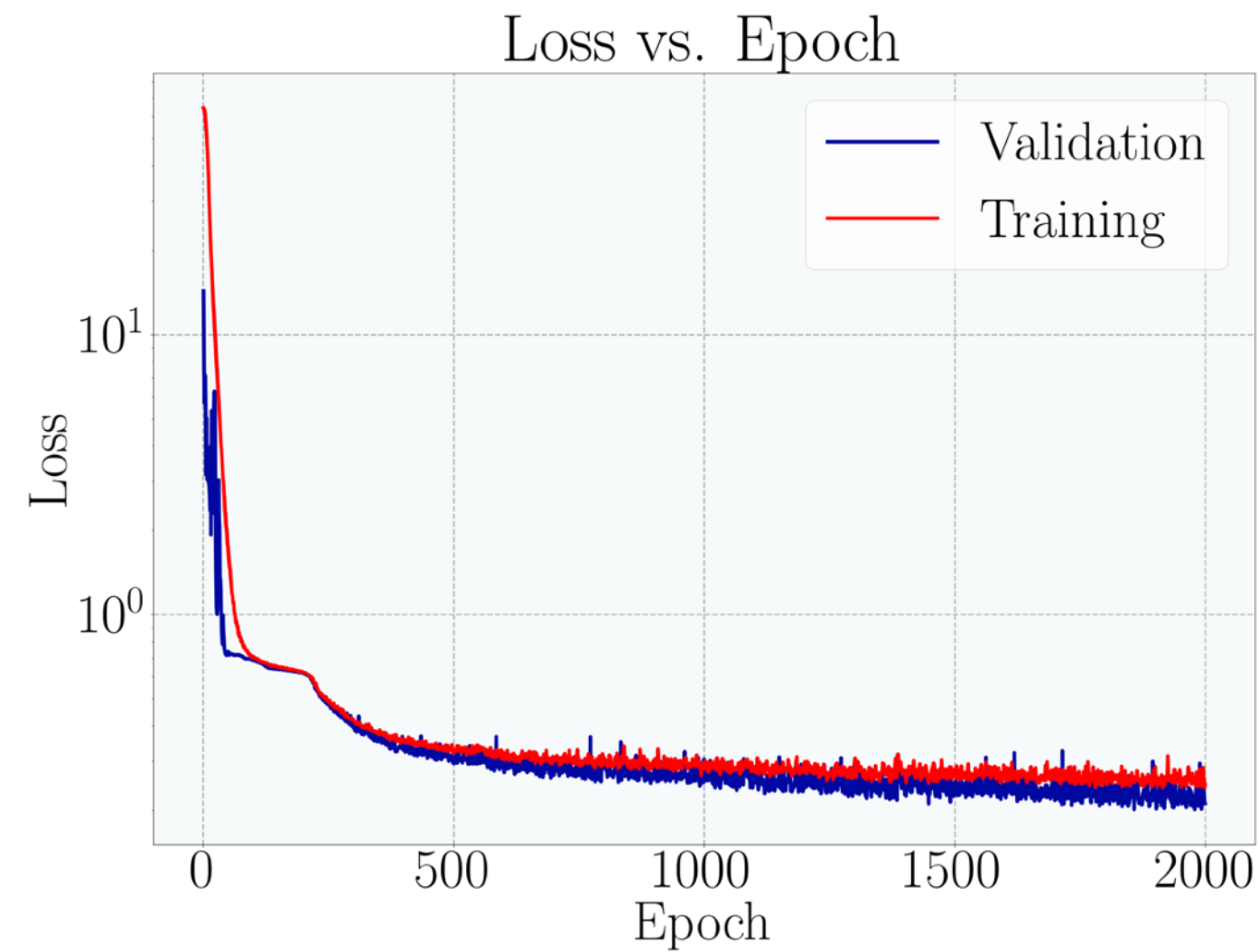
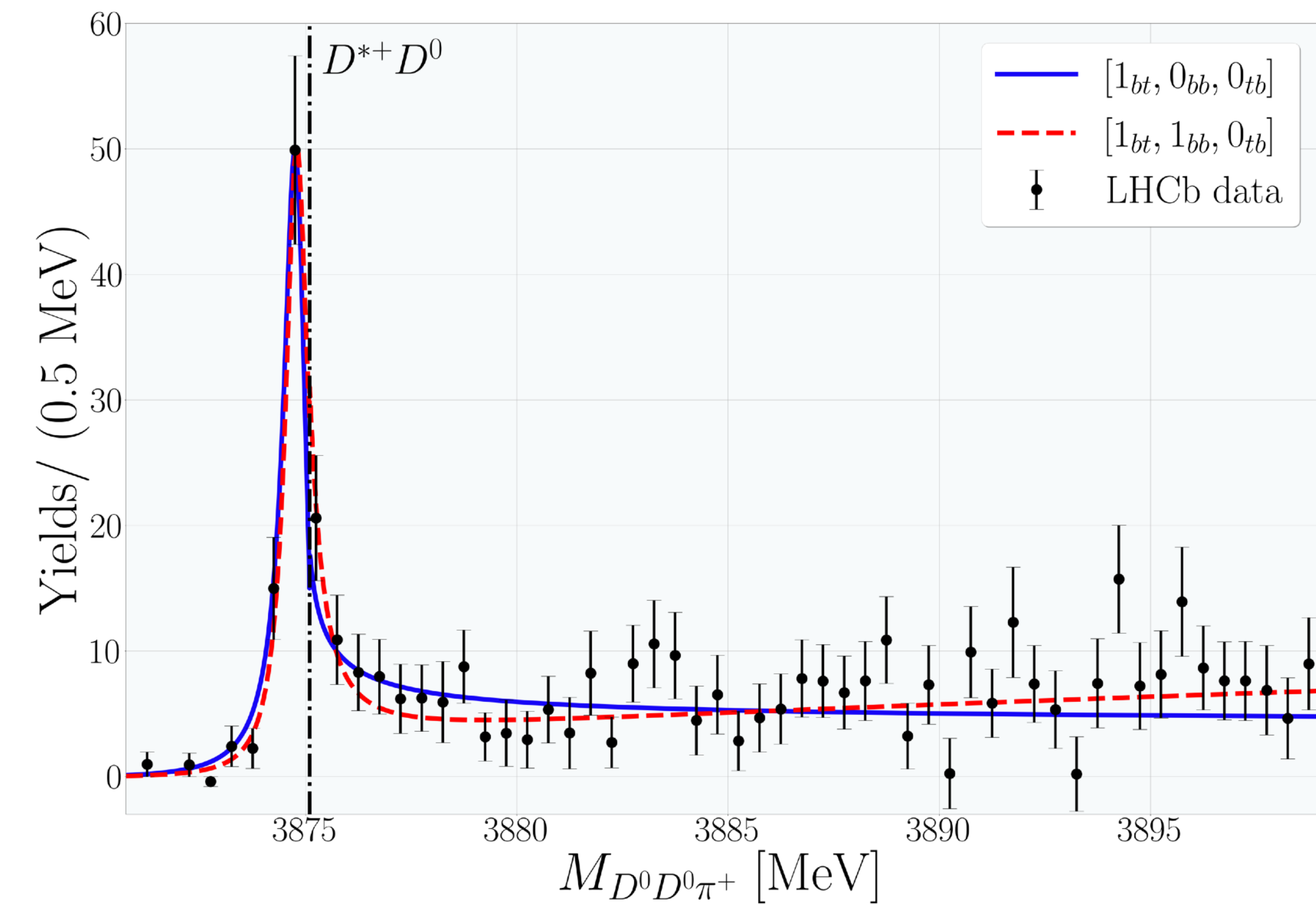


Confusion Matrix



Inference; Stage: 6

Pole structure of $T_{cc}(3875)^+$: Preliminary Results



[JB Pagayon, DLBS, In preparation](#)

Conclusion and Outlook

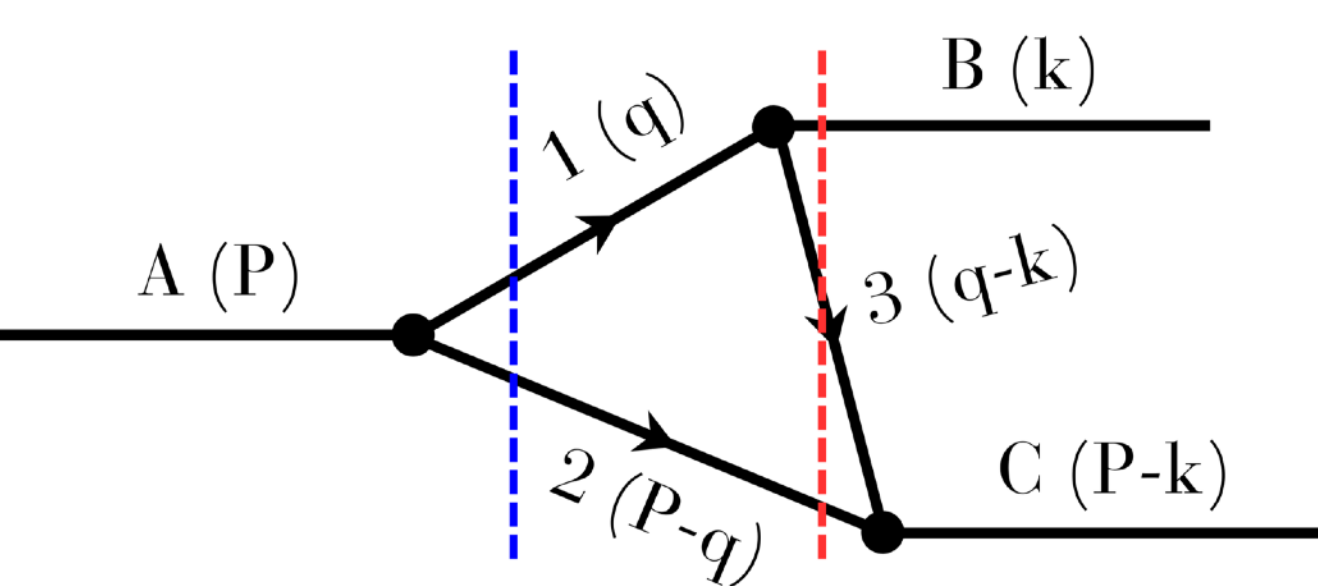
- It is possible to distinguish kinematical enhancements with dynamical pole-based enhancements despite the presence of experimental uncertainty.
- To fully utilized the power of ML, use it a as a model-selection framework.
 - Multi-parameter of a DNN can be used to cover a wider model space.
- Using the ML approach, we have shown that
 - $P_{c\bar{c}}(4312)^+$ is NOT due to (single) triangle singularity
 - $P_{c\bar{c}}(4312)^+$ is a possible true resonance that is contaminated by the coupled-channel interaction of $\Sigma_c \bar{D}$ (having a virtual state) with $J/\psi p$.

Outlook

- Apply the method to other near-threshold phenomena.
- Apply the method to correlation function.

Thank you for your attention.

Back up: TS Ruled out



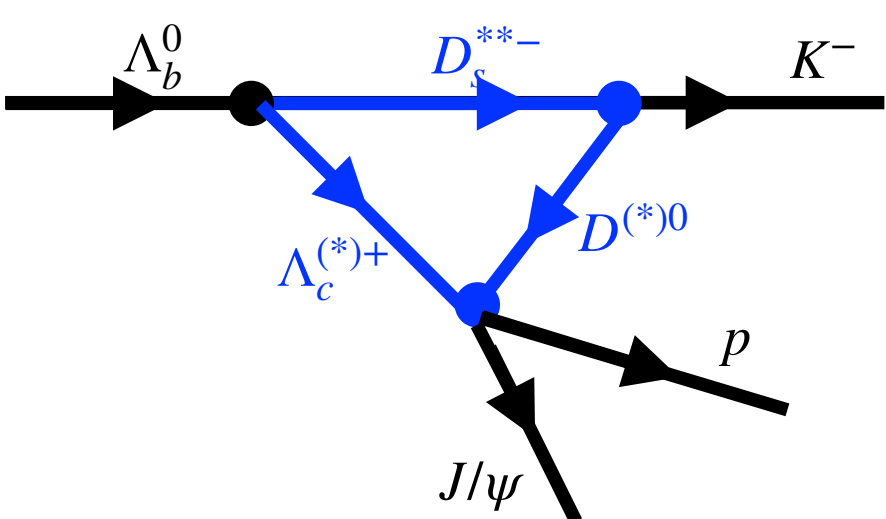
$$I(k) \propto \int_0^\infty \frac{q^2 f(q) dq}{P^0 - \sqrt{m_1^2 + q^2} - \sqrt{m_2^2 + q^2} + i\epsilon}$$
$$f(q) = \int_{-1}^1 \frac{dz}{E_C - \omega(q) - \sqrt{m_3^2 + q^2 + k^2 - 2qkz} + i\epsilon}$$

M Bayar, F Aceti, FK Guo, E Oset, PRD 126 074039 (2016)
FK Guo, XH Liu, S Sakai, PPNP 122 103757 (2020)

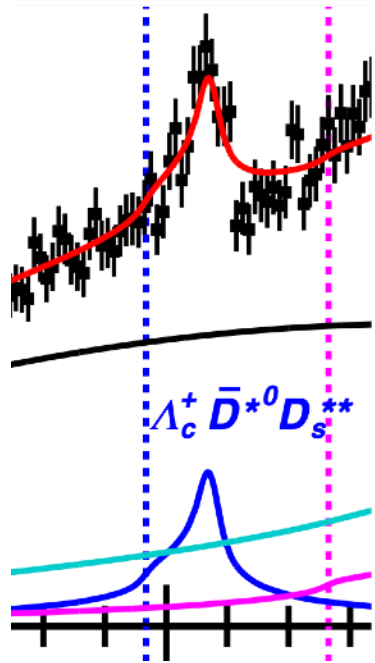
$$m_1^2 \in \left[\frac{M_A^2 m_3 + M_B^2 m_2}{m_2 + m_3} - m_2 m_3, (M_A - m_2)^2 \right]$$
$$m_C^2 \in \left[(m_2 + m_3)^2, \frac{M_A m_3^2 - M_B^2 m_2}{M_A - m_2} + M_A m_2 \right]$$

The amplitude has no singularity (no pole)
No need to introduce new hadronic state to explain the enhancement.

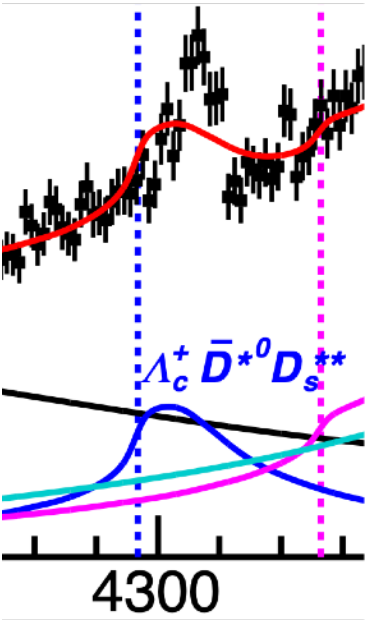
Mass condition - crucial in
generating mock dataset



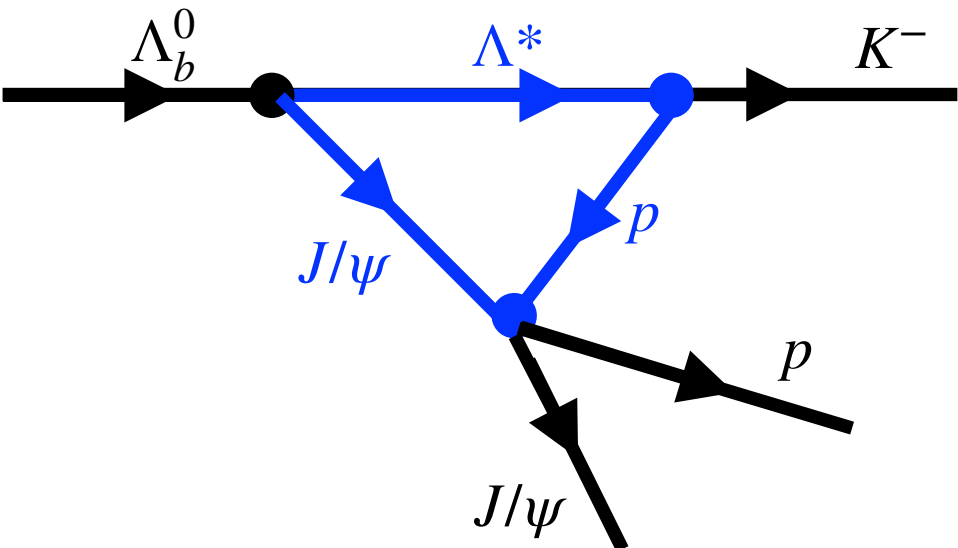
Good fit but with
unrealistic
 $\Gamma \sim 1\text{MeV}$ for D_s^{**}
 $D_s^{*-}(3288)$
Not in PDG



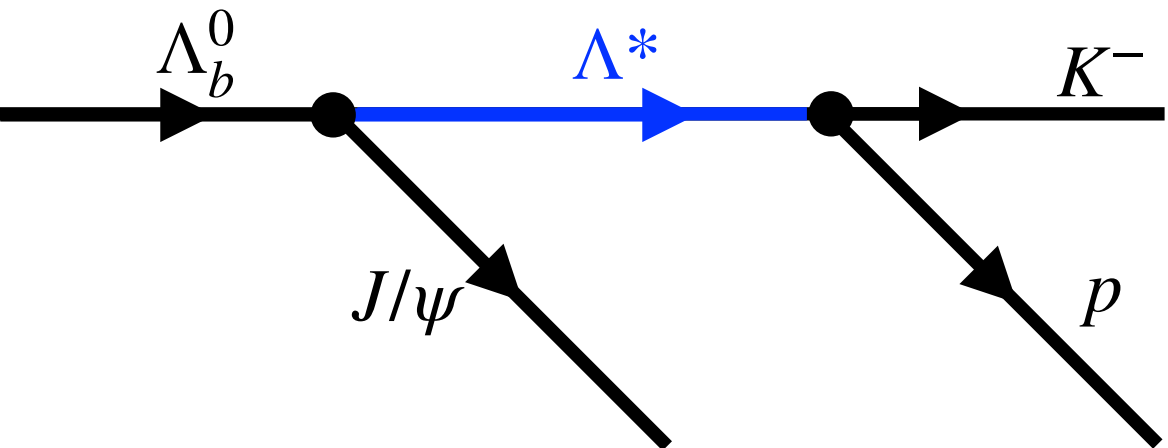
Bad fit with
realistic
 $\Gamma \sim 50\text{MeV}$ for
 D_s^{**}



TS interpretation for the
 $P_{c\bar{c}}(4312)^+$ - ruled out
LHCb, PRL 122 222001 (2019)



TS enhancement is drowned by the
tree-level decay. (Schmid theorem)
VR Debastiani, S Sakai, E Oset, EPJC, 79, 69 (2019)



General S-matrix parametrization

$$S_{11}(p_1, p_2) = \prod_m \frac{D_m(-p_1, p_2)}{D_m(p_1, p_2)}$$

[KJ Le Couteur, Proc. Roy. Soc \(Lo
RG Newton J. Math. Phys. 2, 188](#)

$$p_k \rightarrow q_k; \quad s = q_k^2 + \epsilon_k^2; \quad \omega = \frac{q_1 + q_2}{\sqrt{\epsilon_2^2 - \epsilon_1^2}}$$

One pole per $D_m(q_1, q_2)$: ω_m

$$D_m(q_1, q_2) = D_m(\omega) \quad \text{M Kato, Ann. Phys., 31 1 (1965)}$$

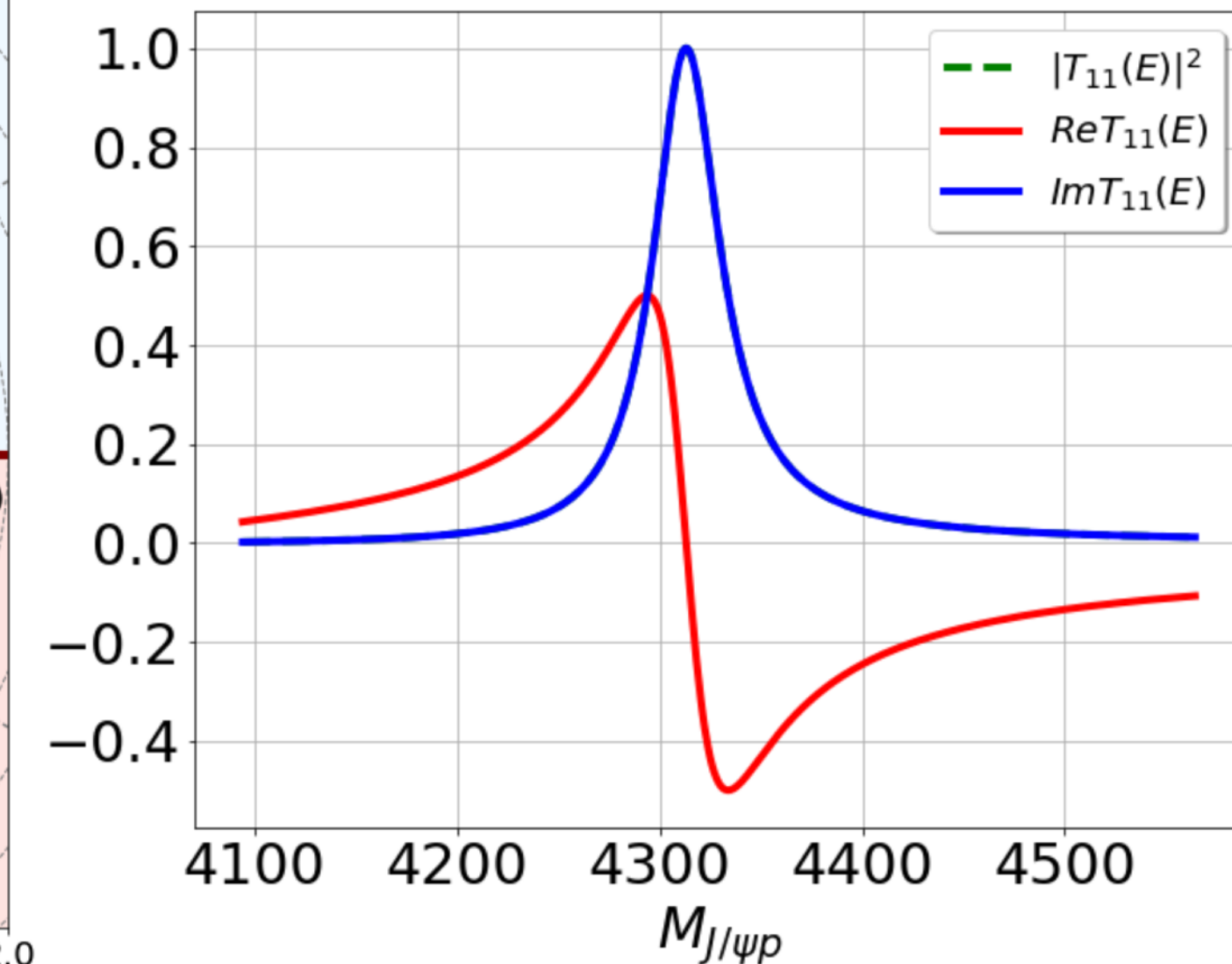
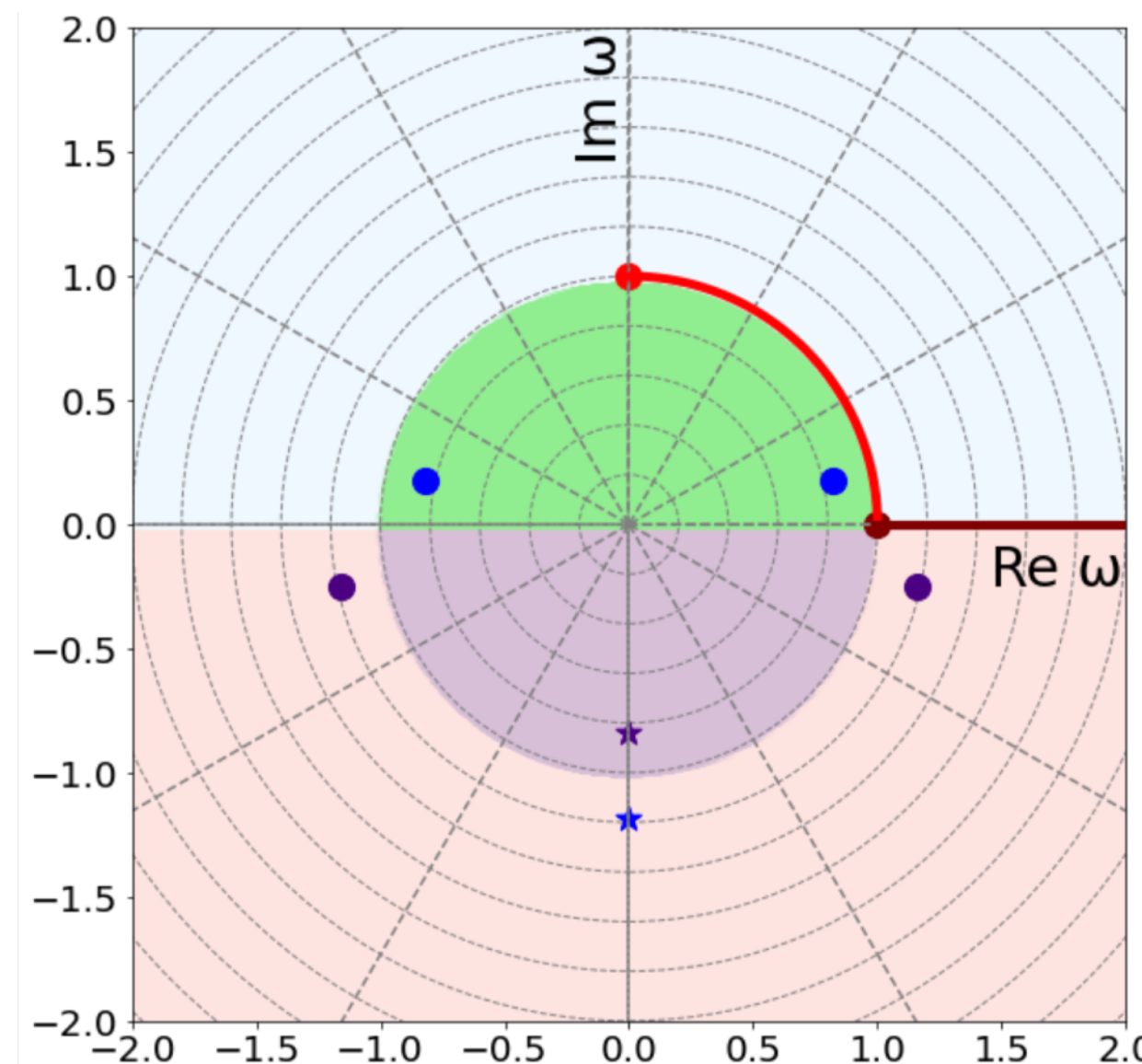
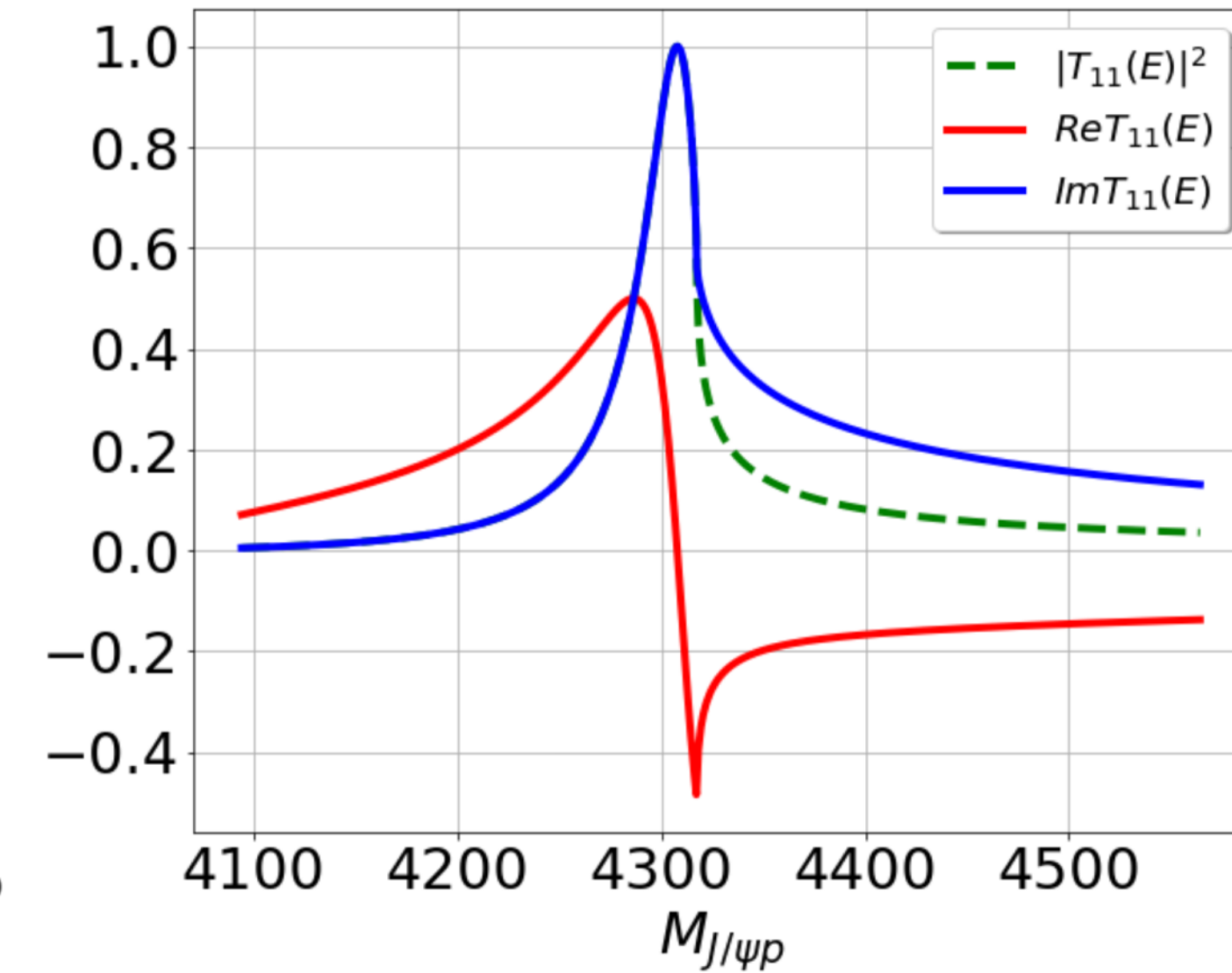
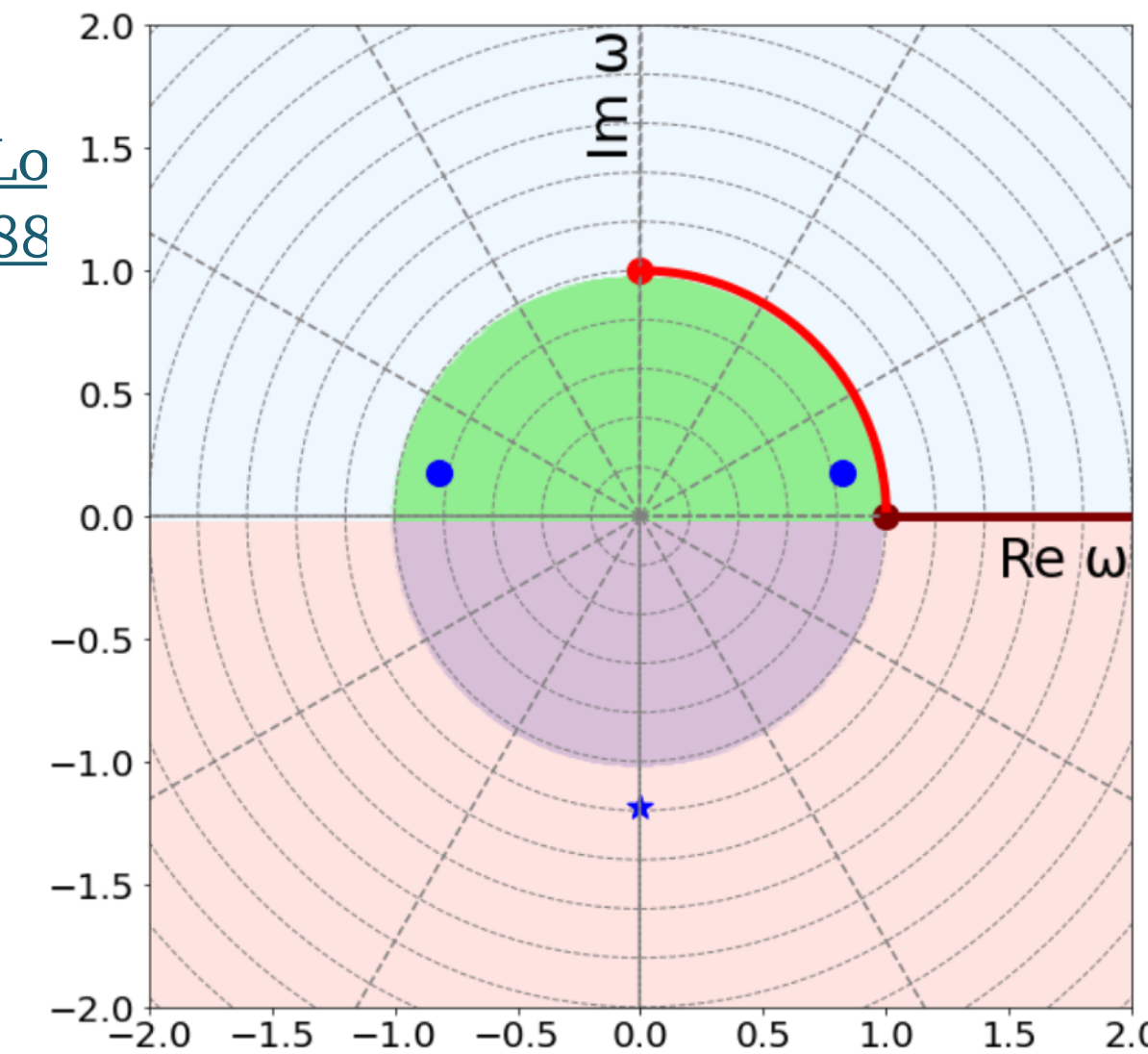
$$= \frac{1}{\omega^2} (\omega - \omega_m) (\omega + \omega_m^*) (\omega - \omega_{\bar{m}}) (\omega + \omega_{\bar{m}}^*)$$

$\omega_{\bar{m}}$ needed only to ensure $\lim_{\omega \rightarrow \infty} S_{11} \rightarrow 1$;

$$|\omega_m \omega_{\bar{m}}| = 1$$

- Poles are introduced independently.
- Cusp at the threshold can be controlled via pole placement.

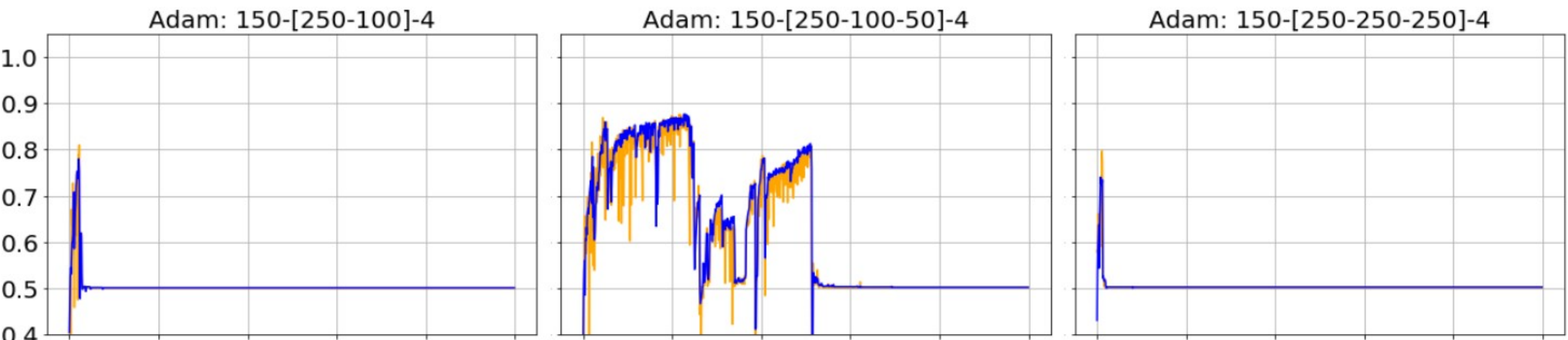
[LMS, DLBS PRC 108 045204 \(2023\)](#)



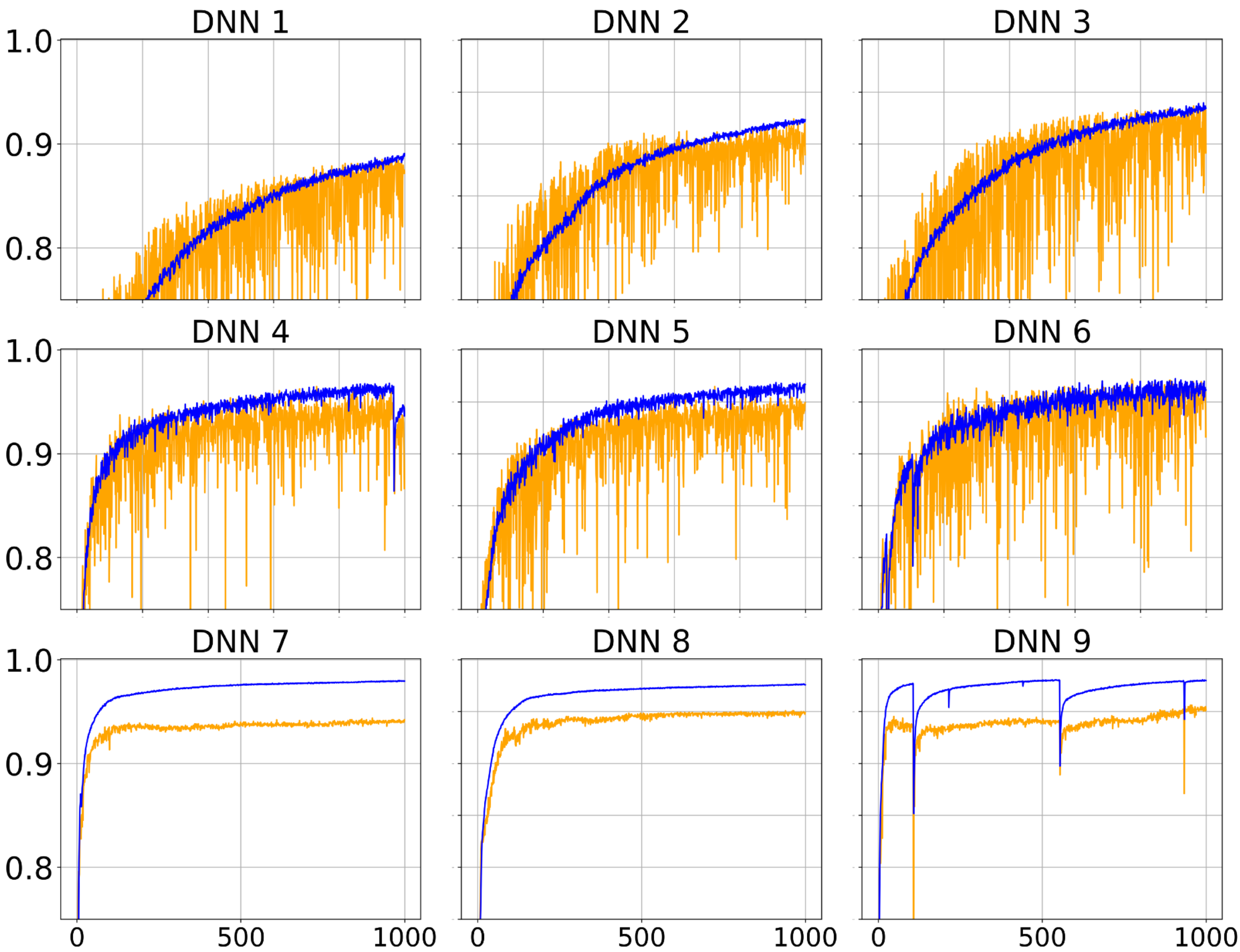
Back-up: Training performance

Model	Optimizer and Architecture
DNN 1	AdaGrad: 150-[250-100]-4
DNN 2	AdaGrad: 150-[250-100-50]-4
DNN 3	AdaGrad: 150-[250-250-250]-4
DNN 4	AMSGrad: 150-[250-100]-4
DNN 5	AMSGrad: 150-[250-100-50]-4
DNN 6	AMSGrad: 150-[250-250-250]-4
DNN 7	SMORMS3: 150-[250-100]-4
DNN 8	SMORMS3: 150-[250-100-50]-4
DNN 9	SMORMS3: 150-[250-250-250]-4

Other optimizers cannot even learn the classification task.

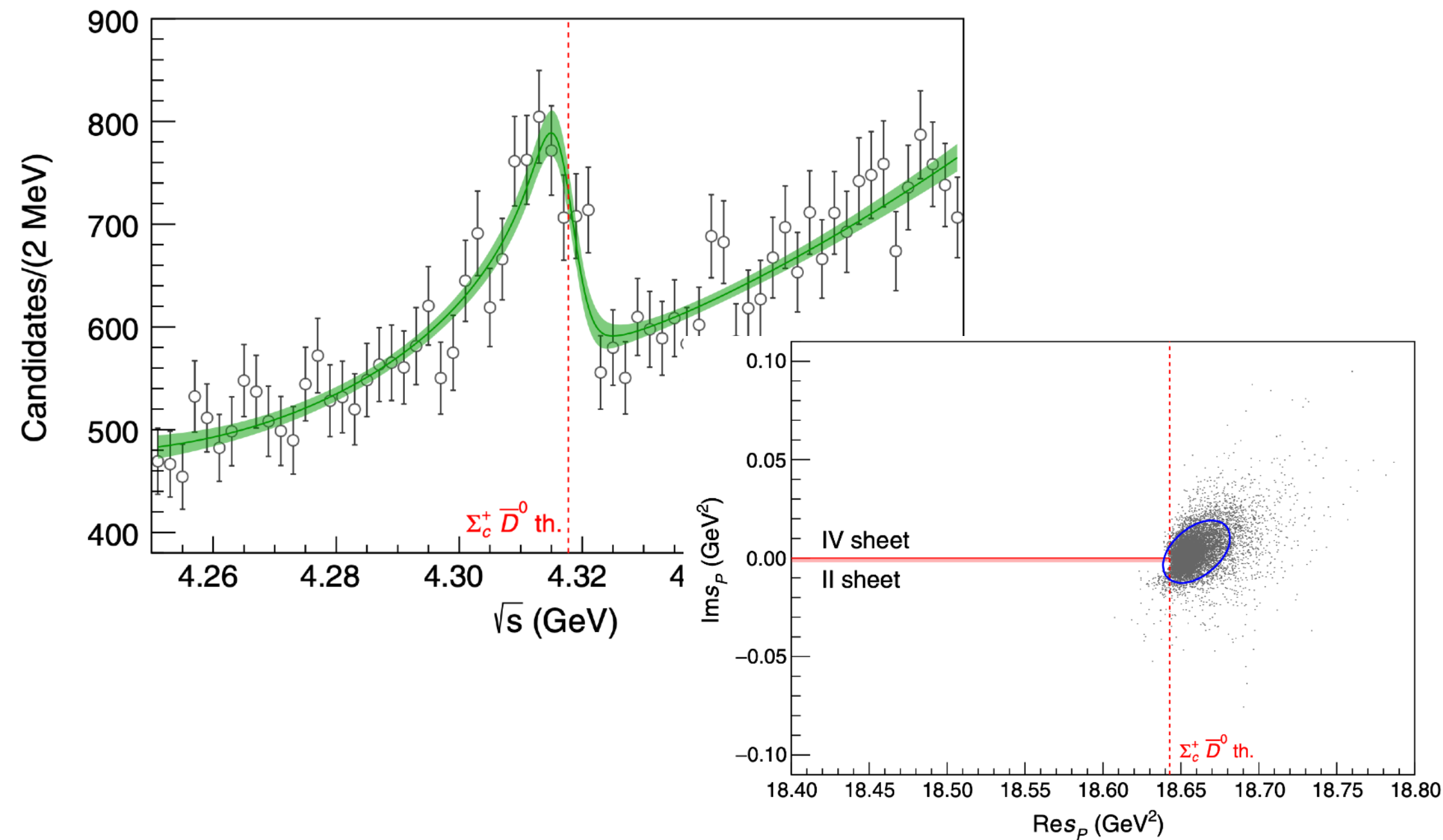


The classification task is non-trivial.



Training epoch vs accuracy

Conventional fitting vs Deep learning



[JPAC, PRL 123 092001 \(2019\)](#)

Fitting:

Constrain the parameter space of ERE:

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

Result:

$P_{c\bar{c}}(4312)^+$ is a virtual state of $\Sigma_c^+ \bar{D}^0$

TABLE I. Softmax output probabilities [26] for the three experimental datasets by LHCb [8].

	$b 2$	$b 4$	$v 2$	$v 4$
$\cos \theta_{P_c}$ -weighted	0.6%	<0.01%	1.1%	98.3%
$m_{Kp} > 1.9$ GeV	1.4%	<0.1%	1.6%	97.0%
m_{Kp} all	5.4%	<0.1%	21.0%	73.6%

[JPAC, PRD 105 L091501 \(2022\)](#)

Train DNN using ERE: $T(s) = \frac{m_{22} - ik_2}{(m_{11} - ik_1)(m_{22} - ik_2) - m_{12}^2}$

Result/ interpretation:

$P_{c\bar{c}}(4312)^+$ is a virtual state of $\Sigma_c^+ \bar{D}^0$

DL result is consistent with standard fitting method.

Is DL just a “glorified” alternative to fitting?

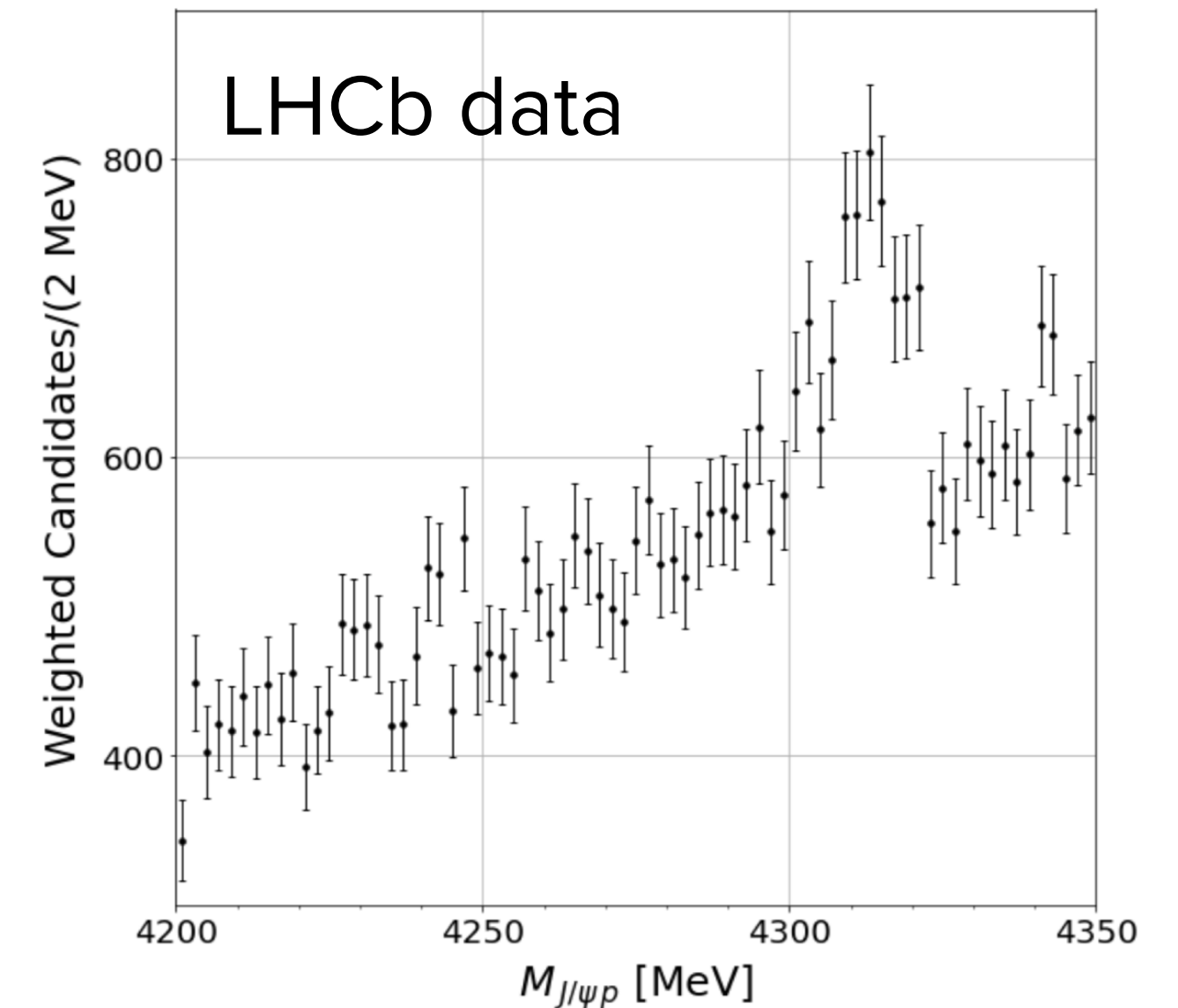
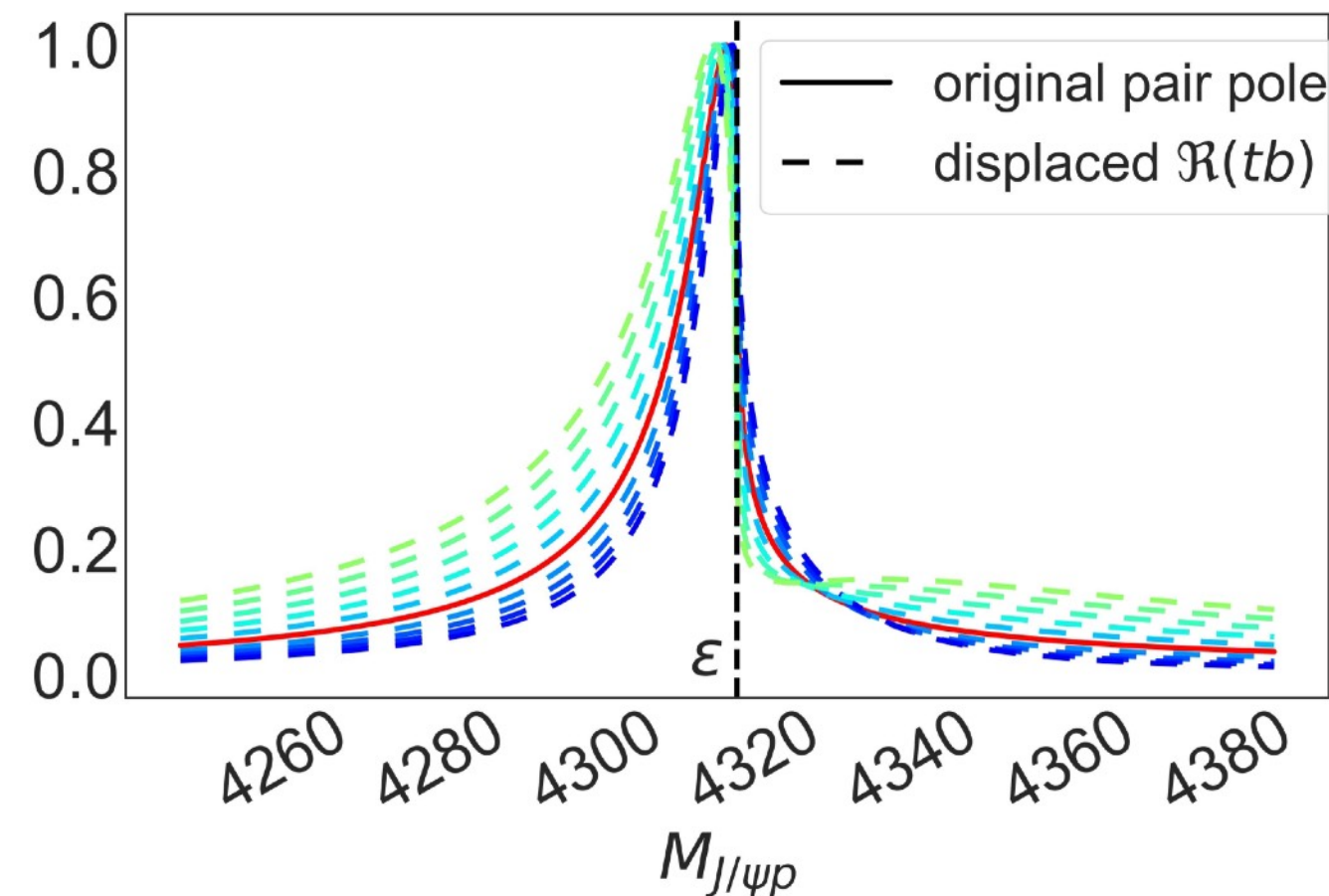
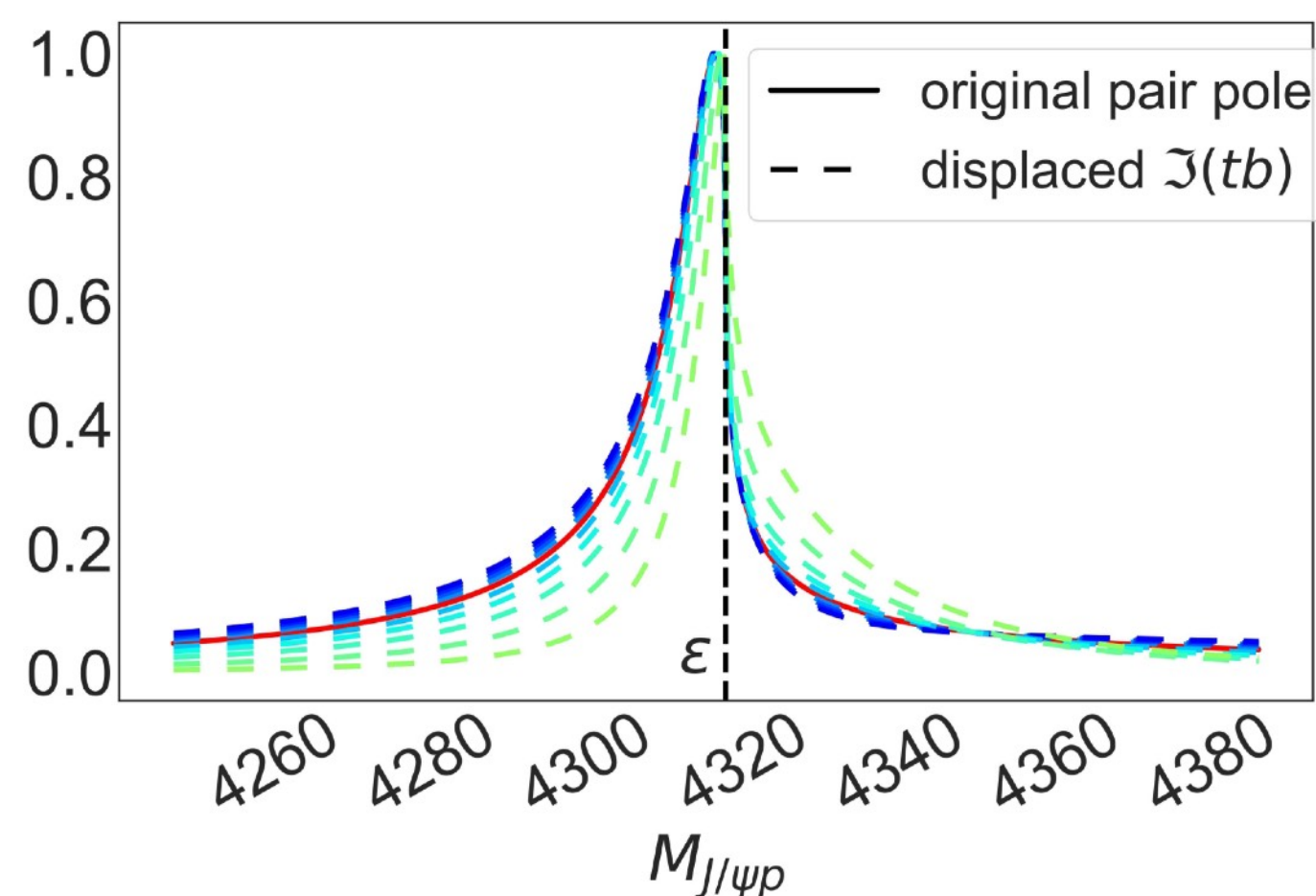
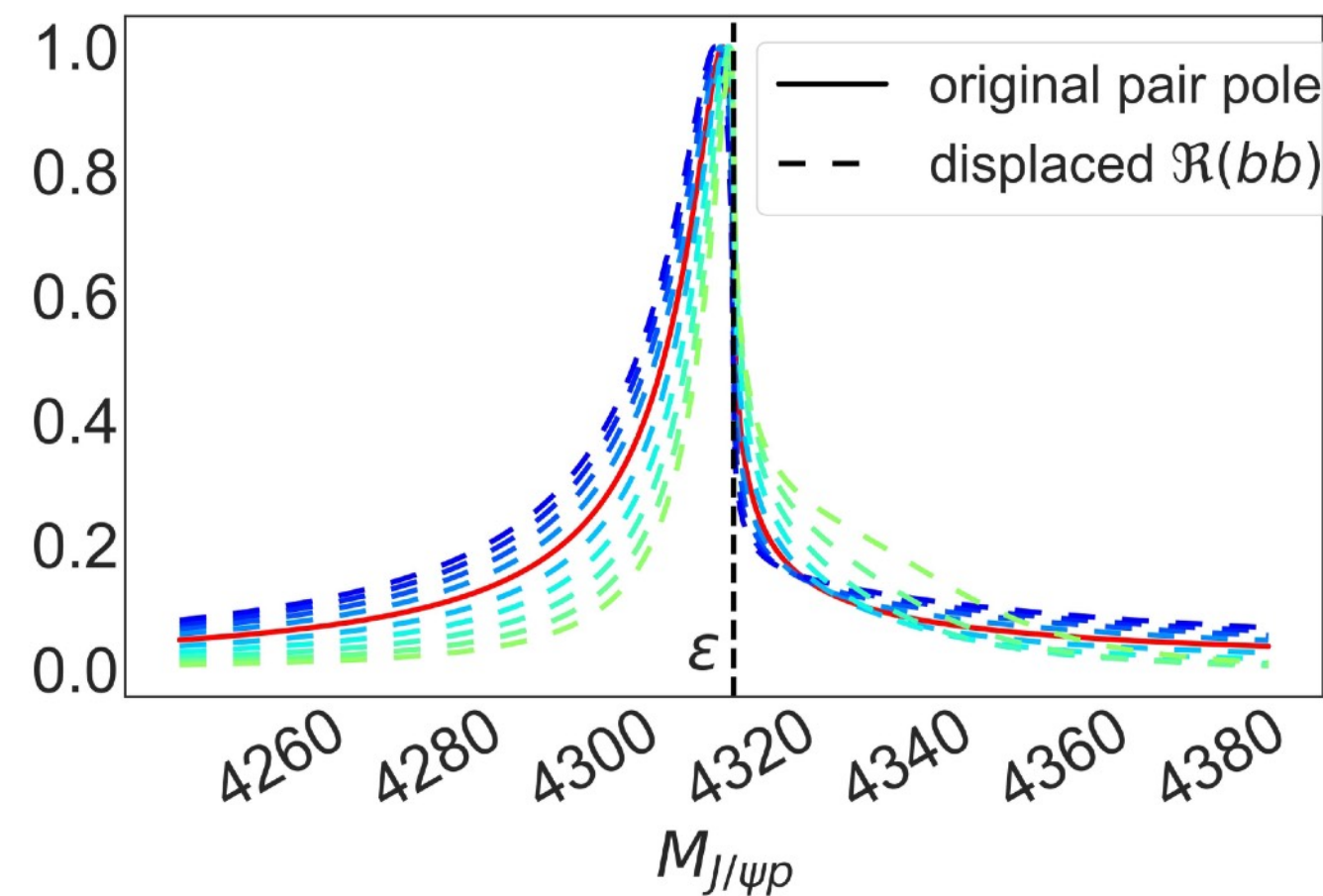
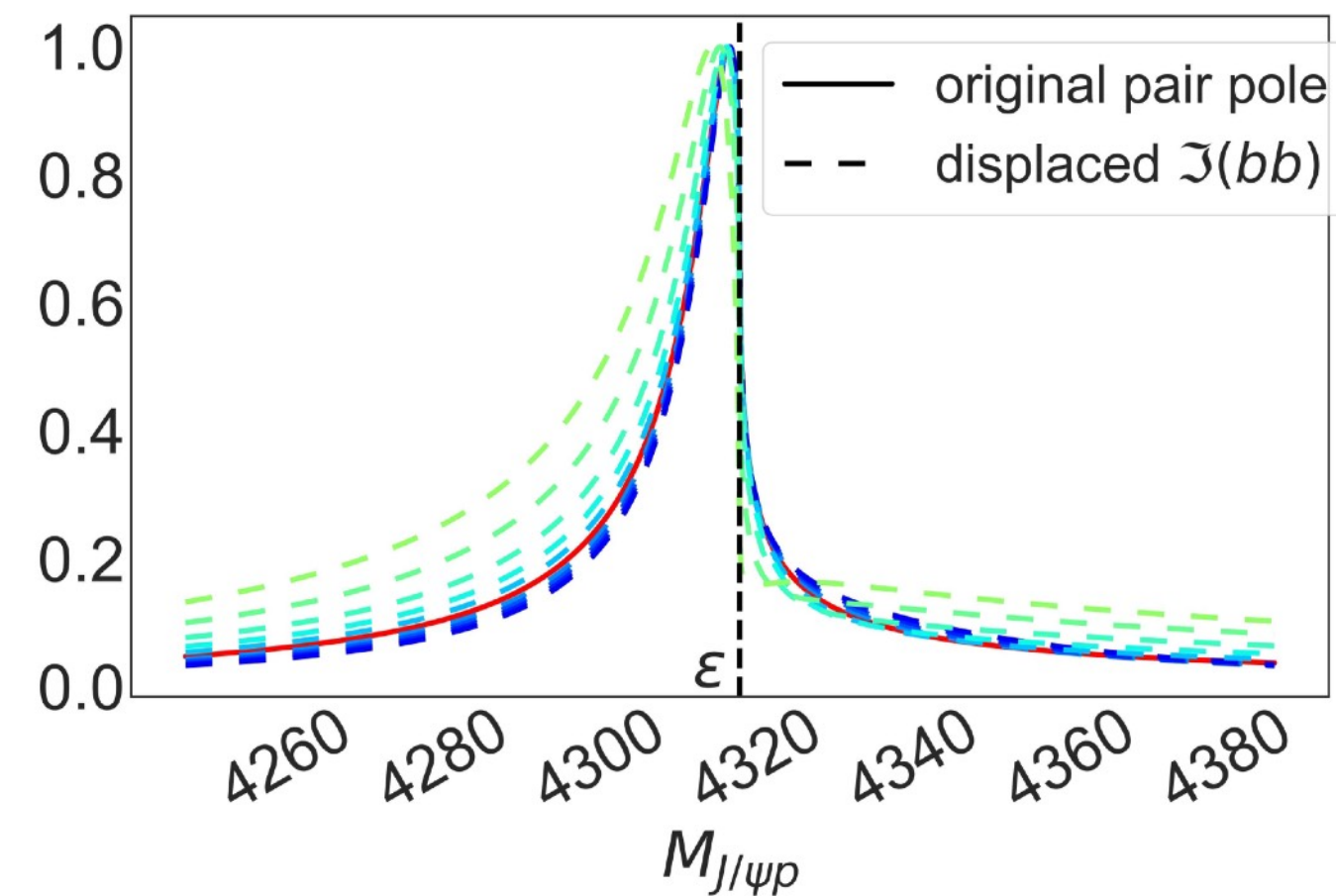
Maybe we can do more.

Proposal:

- Use DL to constrain the model space.
- Use fitting when constraining the parameters of a model.

Back Up: $P_{c\bar{c}}(4312)^+$

Maybe unlikely to happen -
to produce ambiguous pole structure, the poles must have
the same position.



[LHCb, PRL 122 222001 \(2019\)](#)

Slight variation in the line shape

Experimental uncertainty might
hide them.

[LMS, DLBS PRC 108 045204 \(2023\)](#)

Back Up: General parametrization

$$S_{11}(p_1, p_2) = \prod_m \frac{D_m(-p_1, p_2)}{D_m(p_1, p_2)}$$

[KJ Le Couteur, Proc. Roy. Soc \(London\) A256 \(1960\)](#)

[RG Newton J. Math. Phys. 2, 188 \(1961\)](#)

ERE - Scattering length approx.

$$D(p_1, p_2) = (M_{11} - ip_1)(M_{22} - ip_2) - M_{12}^2$$

$D(p_1, p_2)$ can only vanish when the $\bar{p}_1$ and $\bar{p}_2$ have opposite imaginary parts.

[W Frazer and A Hendry, Phys. Rev., 134, B1307 \(1964\)](#)

Flatté-like parametrization

$$D(p_1, p_2) = E - M + ip_1 + i\gamma_2 p_2$$

Shadow poles may appear on the physical sheet.

[RJ Eden, JR Taylor, Phys. Rev. 133, B1575 \(1974\)](#)

$$\left[(p_1 - i\beta_1)^2 - \alpha_1^2 \right] + \lambda \left[(p_2 - i\beta_2)^2 - \alpha_2^2 \right] = 0$$

We need a general parametrization:

- Pole position can be controlled and RS can be assigned.
- Poles are independent of each other.

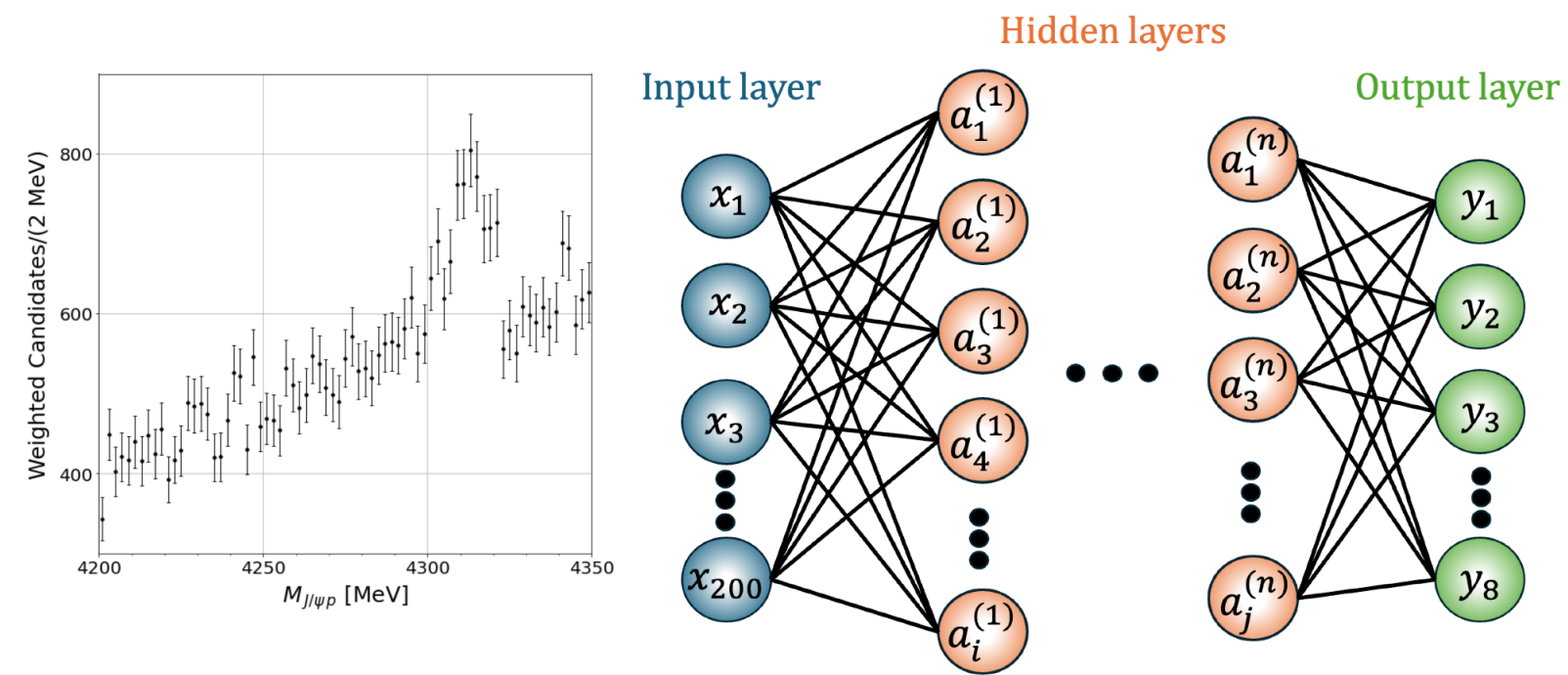
$$\left[(p_1 - i\beta_1)^2 - \alpha_1^2 \right] \left[\left(p_1 - i\beta_1 \frac{1-\lambda}{1+\lambda} \right)^2 - \left(\alpha_1^2 + \frac{4\lambda\beta_2^2}{(1+\lambda)^2} \right) \right] = 0$$

$$\left[(p_2 - i\beta_2)^2 - \alpha_2^2 \right] \left[\left(p_2 + i\beta_2 \frac{1-\lambda}{1+\lambda} \right)^2 - \left(\alpha_2^2 + \frac{4\lambda\beta_1^2}{(1+\lambda)^2} \right) \right] = 0$$

Main pole

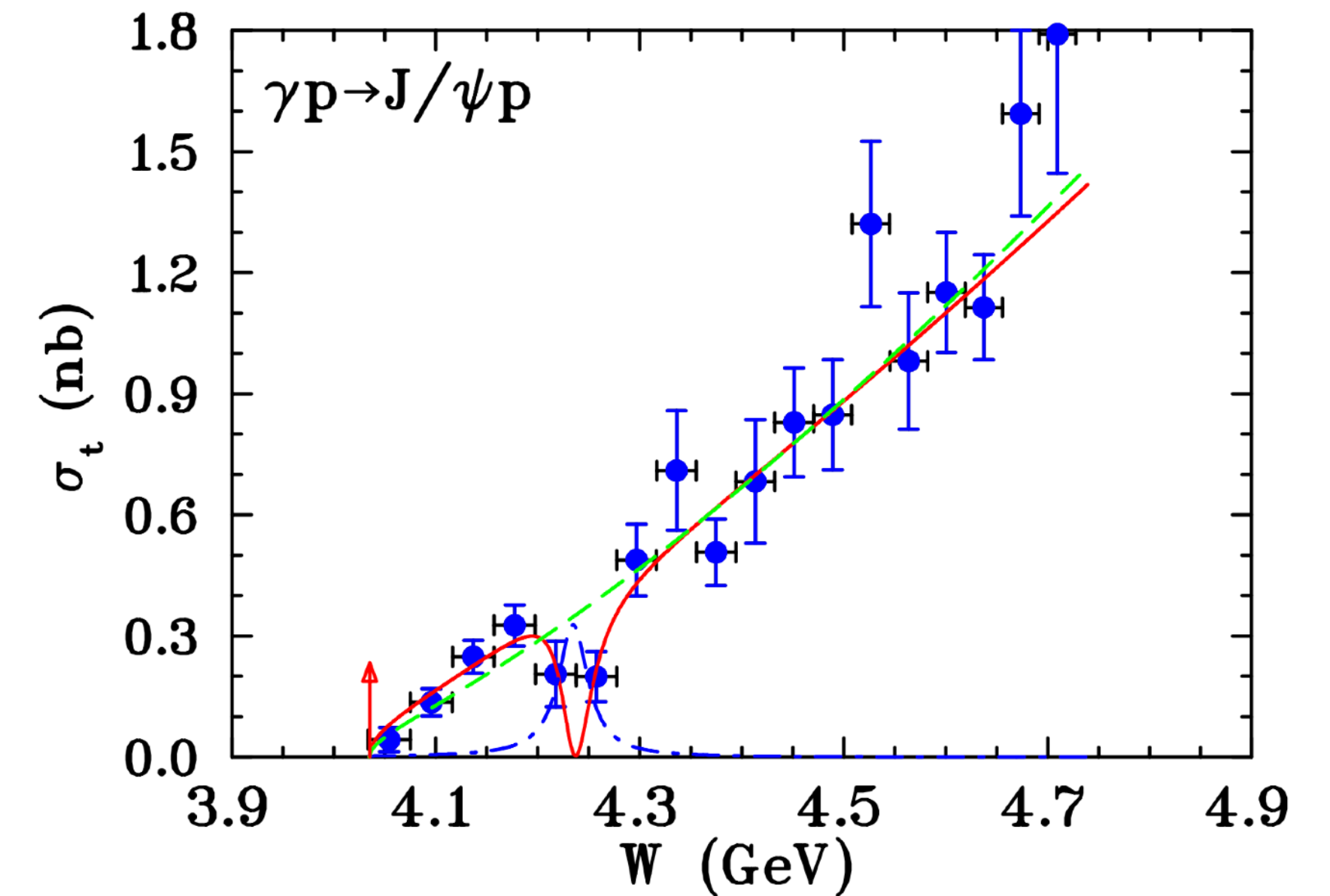
Shadow poles

Back Up: $P_{c\bar{c}}(4312)^+$



Pole in 2nd and 3rd RS
 -resemble pole-shadow pair
 -possible true resonance
 decaying into $J/\psi p$

Pole in 4th RS
 -Virtual state of $\Sigma_c \bar{D}$
 coupled to $J/\psi p$



In contrast with the analysis of JPAC in 2019

[JPAC, PRL 123 092001 \(2019\)](#)

GlueX: J/ψ photo-production 2023 result - structure found in the J/ψ - p cross-section

[GlueX, PRC 108 025201 \(2023\)](#)

The dip structure can be interpreted as a resonance that interfere with the non-resonant background.

[I Strakovsky, et al, PRC 108 015202 \(2023\)](#)

No contribution from the $\Sigma_c \bar{D}$ - cannot be interpreted as molecular.