DEEP LEARNING AS A UNIFIED MODEL-SELECTION TOOL

DENNY SOMBILLO

RCNP, Osaka University and NIP, University of the Philippines Diliman

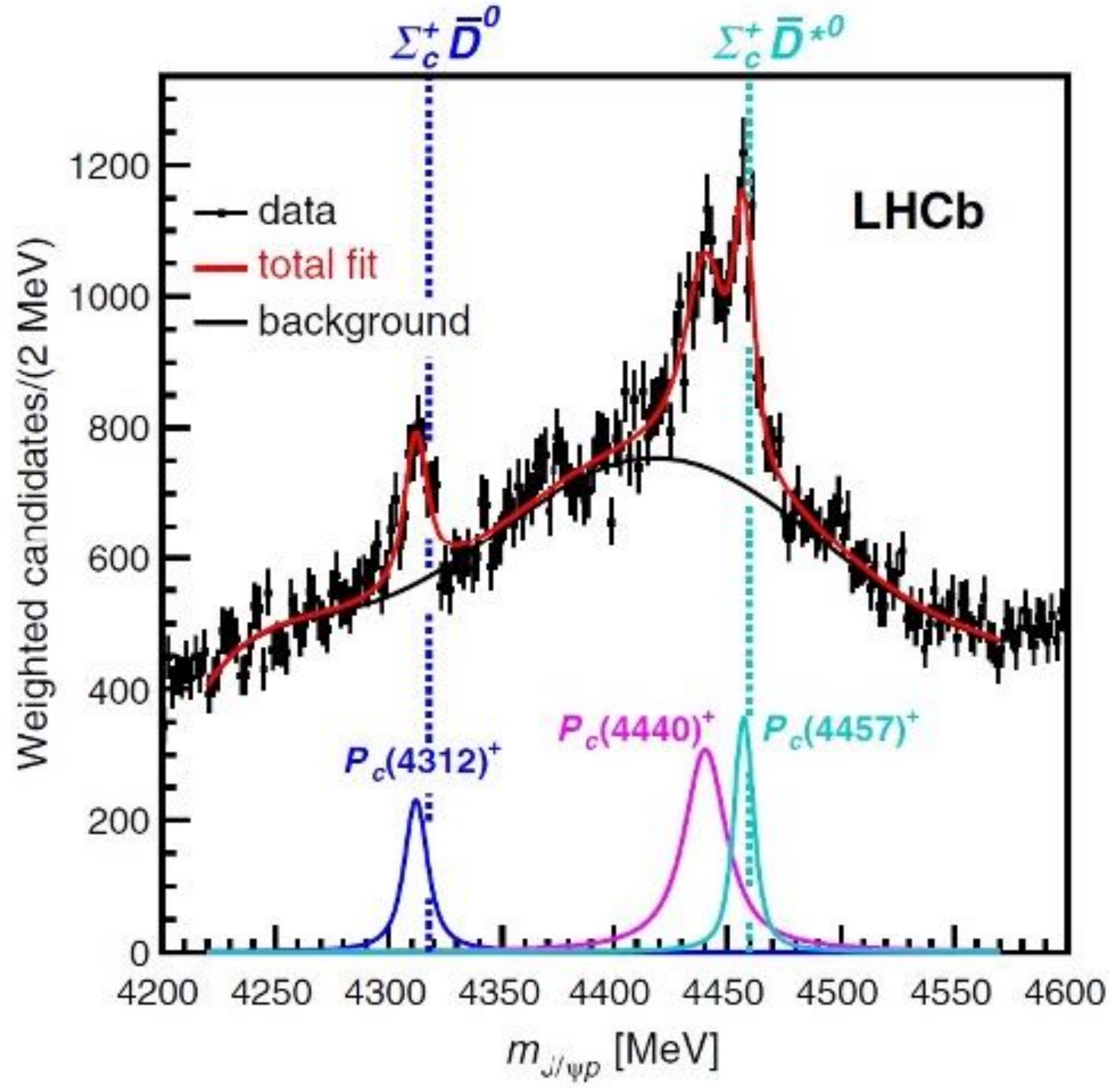
DLBS, YI, TS, AH PRD 102 (2020)
DLBS, YI, TS, AH Few-Body Syst. 62, 52 (2021)
DLBS, YI, TS, AH PRD 104 (2021)

In collaboration with:
YOICHI IKEDA
TORU SATO
ATSUSHI HOSAKA

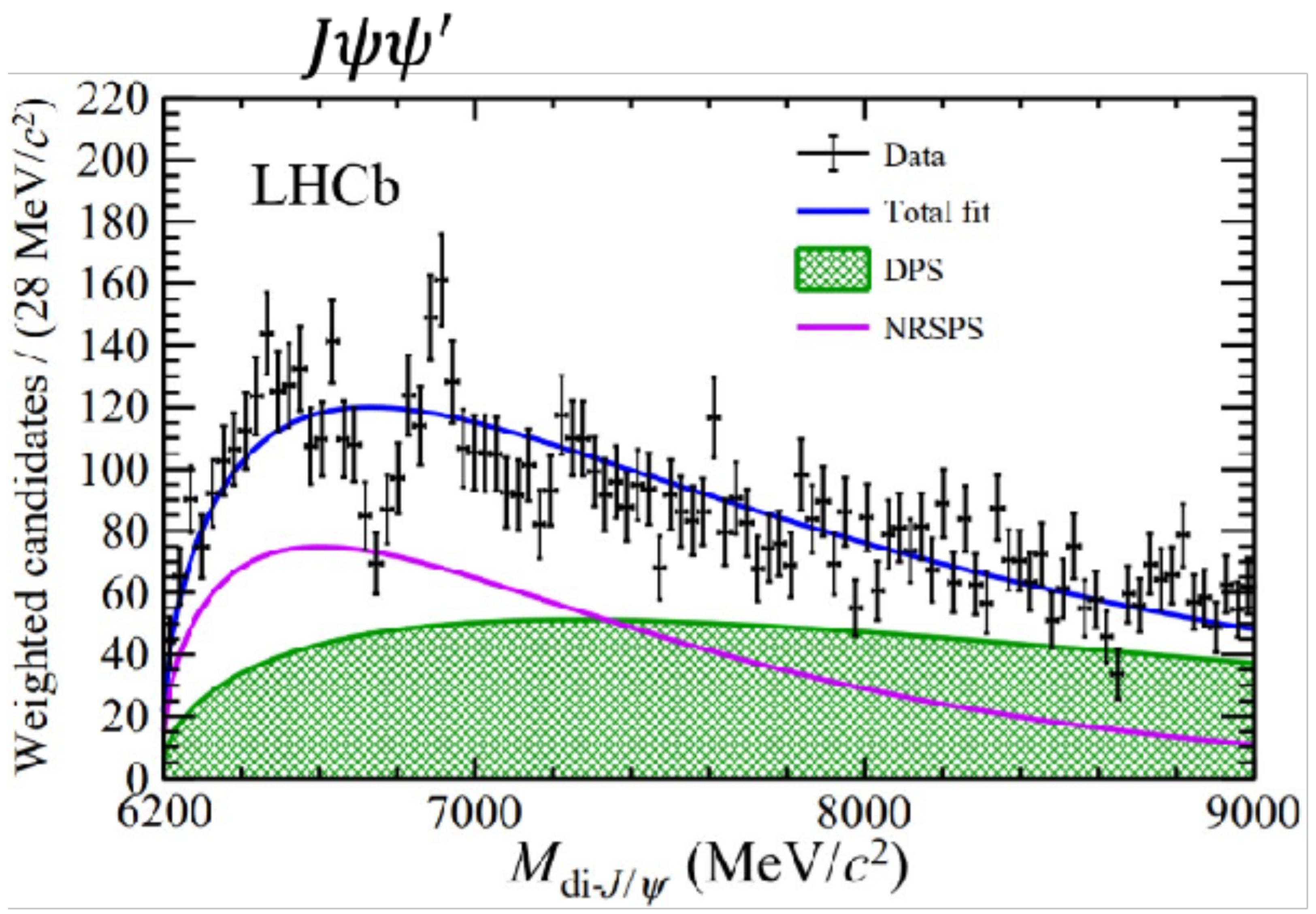
大阪大学 OSAKA UNIVERSITY
RCNP

Motivation

Objective:
Deduce the structure of observed resonances using only the experimental data.



10.1103/PhysRevLett.122.222001



10.1016/j.scib.2020.08.032

Possibilities:

- Compact state
- Molecular
- Virtual
- Cusp, Triangle singularity, etc.

$\boxed{\text{Pole counting argument.}}$

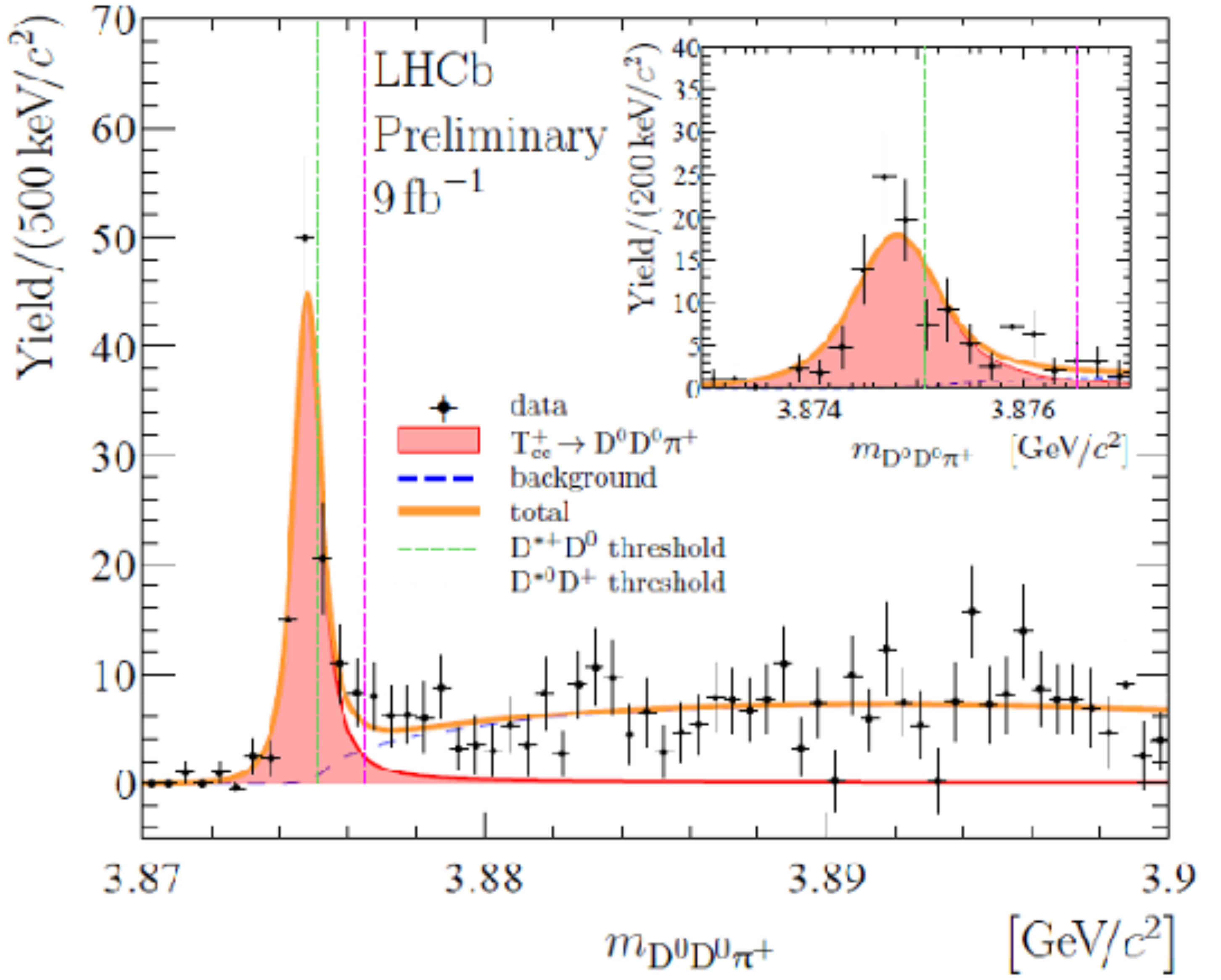
D. Morgan, Nucl.Phys.A 543, 4 (1992)

- Number of poles is related to the nature of resonance.

  D. Morgan, M.R. Pennington, Phys.Lett.B 258 (1991)
  D. Morgan, Nucl.Phys.A 543, 4 (1992)
  D. Morgan, M. R. Pennington, Phys. Rev. D 48 (1993)

- Connected to compositeness

  Phys. Lett. B 586 (2004)

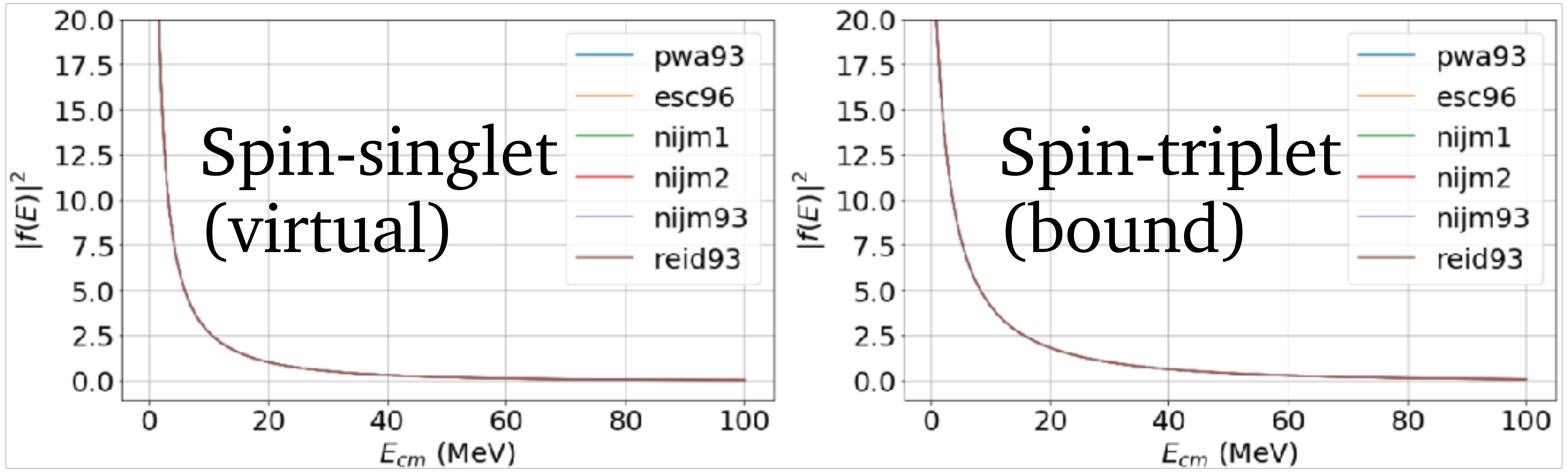


arXiv:2109.01038

Deep learning: alternative analysis tool

Benchmarked on the known nucleon-nucleon bound state

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



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

Applied on the pion-nucleon coupled channel scattering

Identify the number of poles in each Riemann sheet to reproduce the amplitude.

DLBS, YI, TS, AH PRD 104 (2021)

Recently applied on Pc(4312)

Interpret the signal with the help of DNN JPAC Collaboration arXiv:2110.13742 (2021)

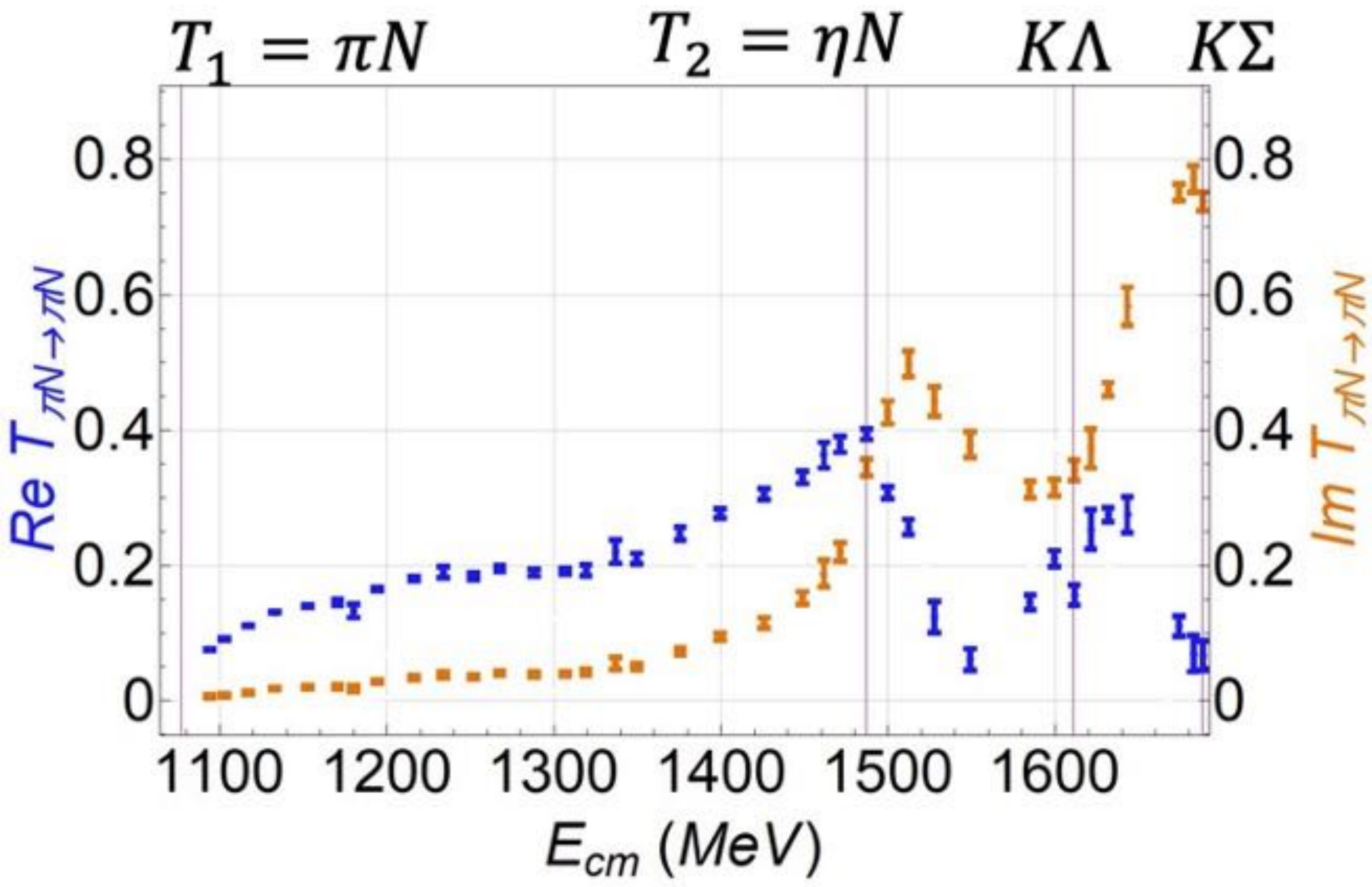
DL as a unified model selection framework

Select a pole-based model:
How many nearby poles in each Riemann sheet are needed to reproduce the experimental data.

Model space restriction:
Maximum of 4 poles, distributed in any of the unphysical sheets.

Two-channel case: 35 pole-based models

Label	S-matrix pole configuration
0	no nearby pole
1	1 pole in $[bt]$
2	2 poles in $[bt]$
⋮	⋮ ⋮ ⋮ ⋮
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 $[tb]$



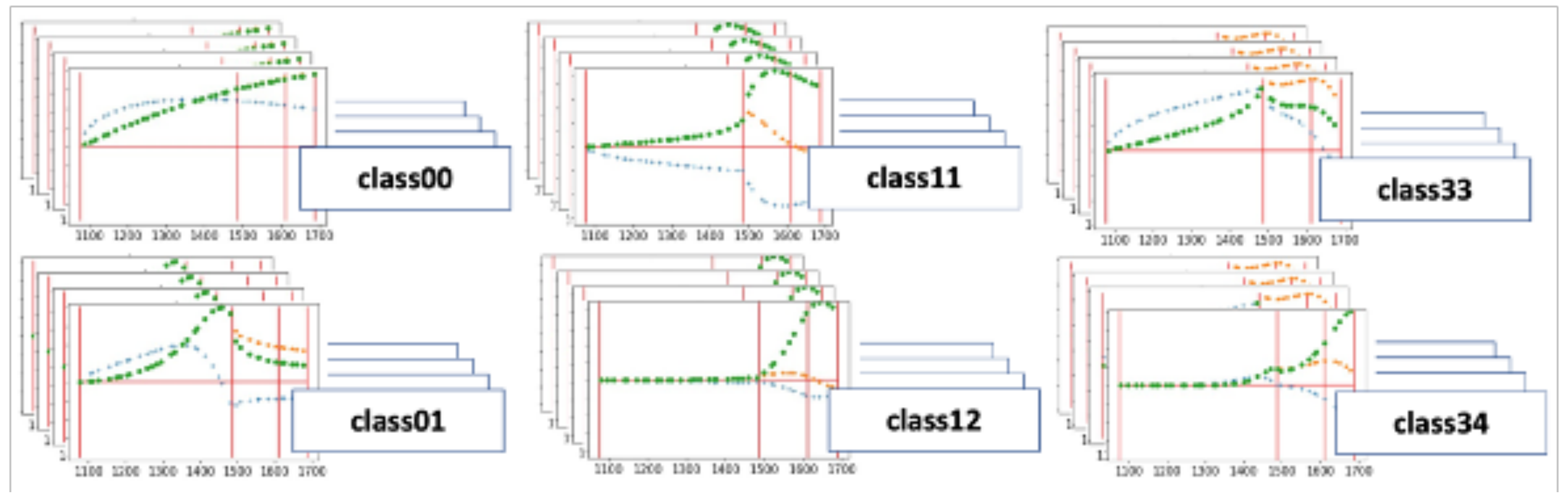
It is reasonable to expect that more than one pole-based model can describe the data due to the error bars.

DL as a unified model selection framework

DL approach:

Generate the training dataset

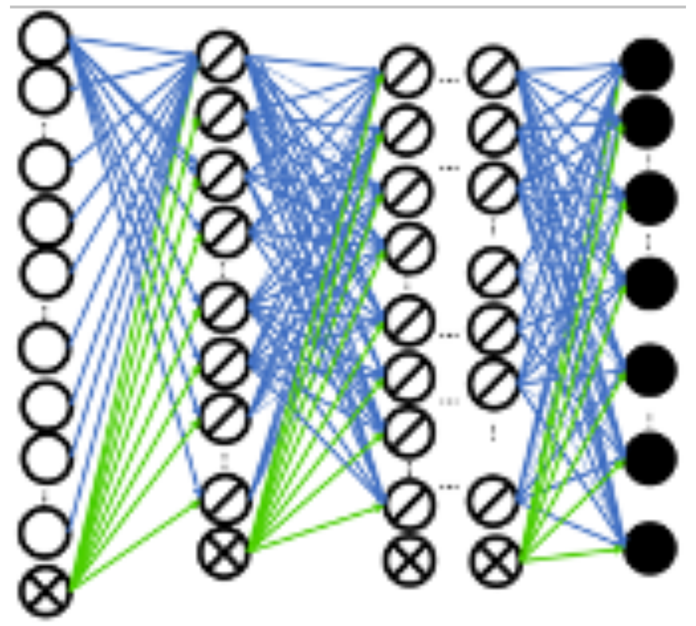
- Use only the general properties of S-matrix
- Include the energy uncertainty



Optimize the parameters of the deep neural network

Input layer:

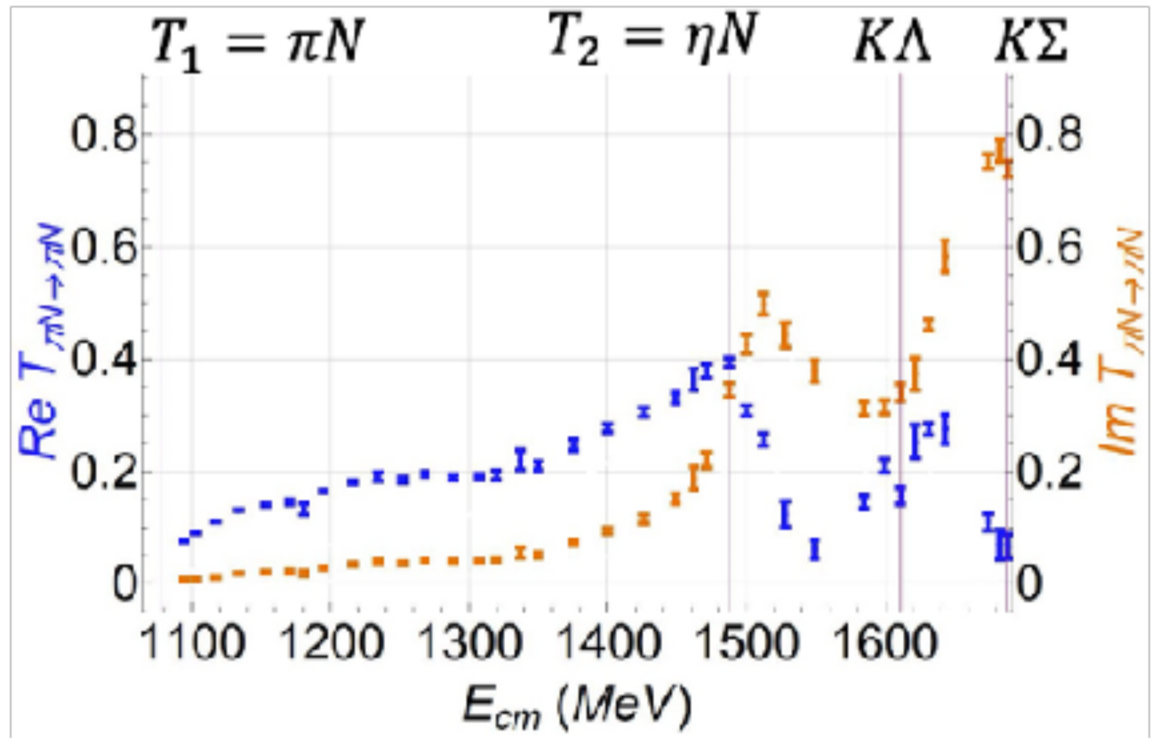
- Energy points
- Real part of amplitude
- Imaginary part of amplitude

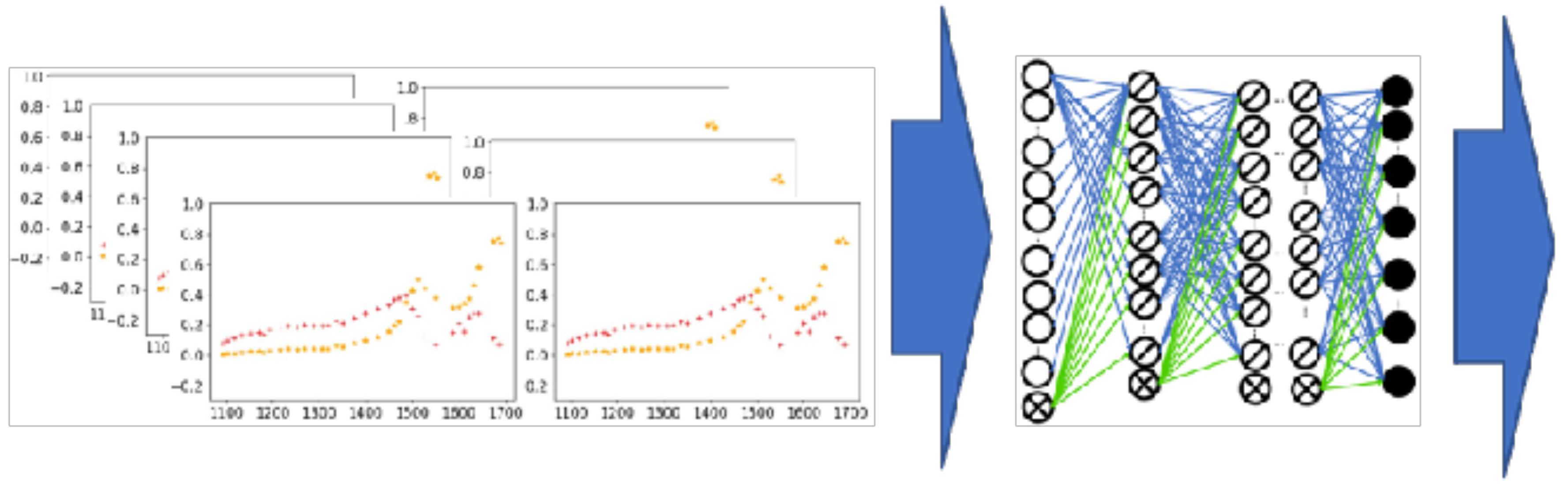


Output layer:

- Pole model 0
- Pole model 1
- Pole model 2
- …
- …
- Pole model 34

Deploy the trained DNN to extract model from the





Sample result:

- X% 2[bt]-0[bb]-0[tb]
- Y% 0[bt]-2[bb]-0[tb]
- Z% 0[bt]-1[bb]-3[tb]
- …

Training dataset (generation of model space)

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

General form of S-matrix:

- Hermiticity below the lowest threshold
- Unitarity
- Analyticity

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

$$S_{22}(p_1,p_2) = \prod_m \frac{D_m(p_1,-p_2)}{D_m(p_1,p_2)}$$

$$S_{11}S_{22} - S_{12}^2 = \prod_m \frac{D_m(-p_1,-p_2)}{D_m(p_1,p_2)}$$

The available experimental data will determine the relevant matrix element.

$$S_{11} = 1 + 2iT_{11}$$

Select a convenient representation of $D_m(p_1,p_2)$ to control the pole and RS.

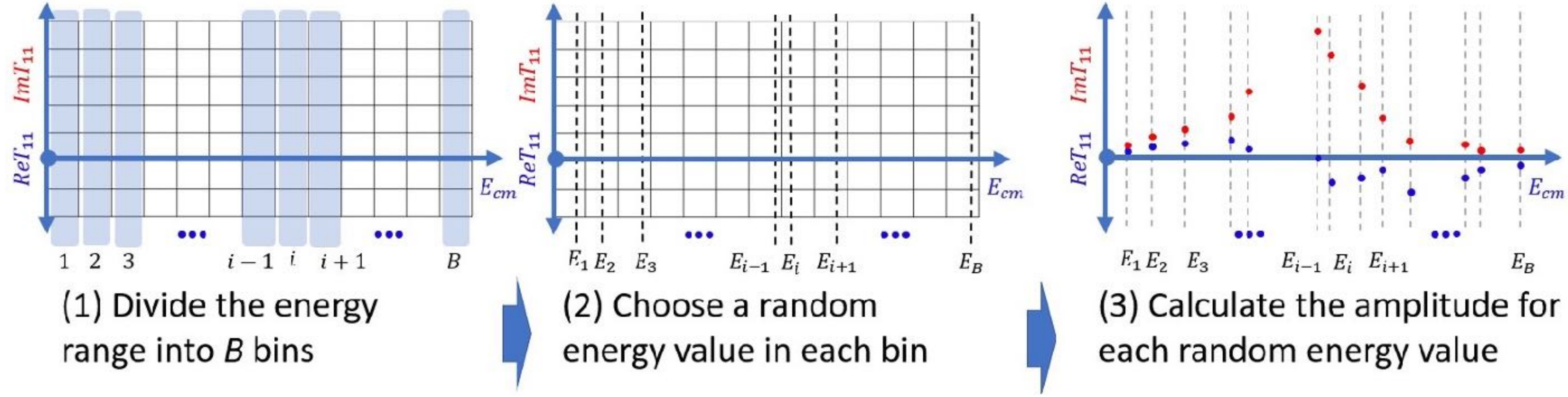
$$D_m(p_1,p_2) = \left[\left(p_1 - i\beta_{1m}\right)^2 - \alpha_{1m}^2\right] + \lambda_m\left[\left(p_2 - i\beta_{2m}\right)^2 - \alpha_{2m}^2\right]$$

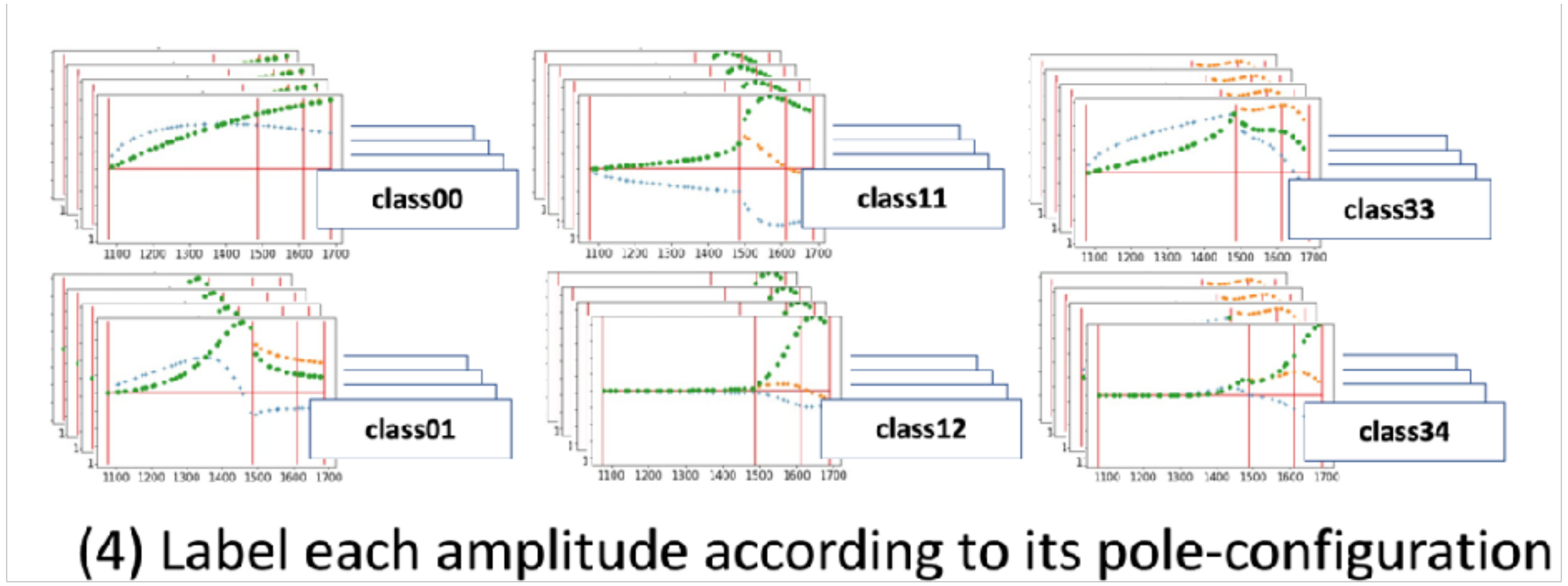
A more convenient way is the use of uniformization:
W. Yamada and O. Morimatsu, PRC 103 (2021)

All the poles used to form one amplitude are independent of each other.
No *a priori* assumption on how the poles are related in the uncoupled limit.

Training dataset (generation of model space)

Incorporate uncertainty in the energy:





Label	S-matrix pole configuration
0	no nearby pole
1	1 pole in [bt]
2	2 poles in [bt]
⋮	⋮ ⋮ ⋮ ⋮
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]

Optimization of DNN model

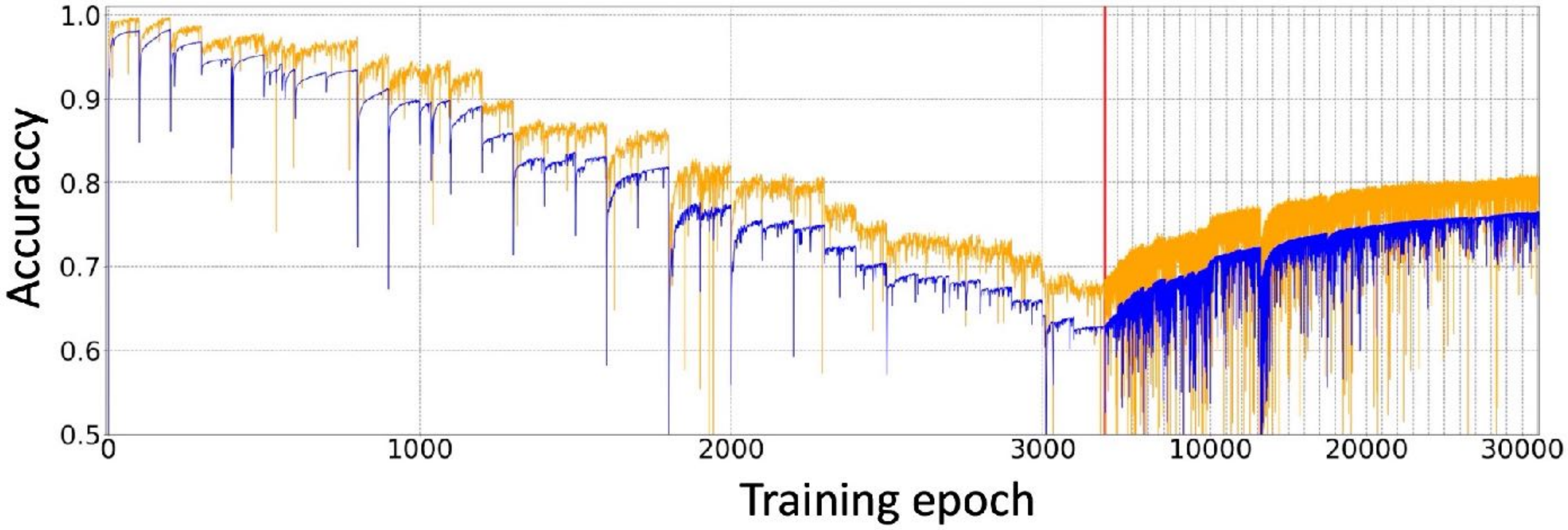
Chosen DNN architecture

Layer	Number of nodes	Activation Function
Input	111+1	
1st	200+1	ReLU
2nd	200+1	ReLU
3rd	200+1	ReLU
Output	35	Softmax

We adopted the **curriculum method** to train the DNN using the noisy dataset.

After $\sim$ 31,000 epochs the final training and testing accuracies are 76.5 % and 80.4 % , respectively.

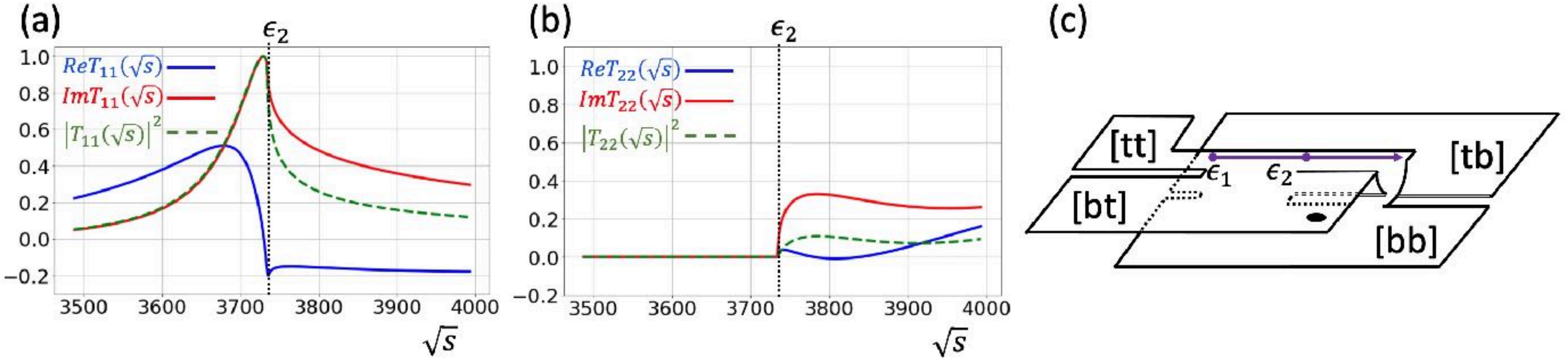
Performance in curriculum training



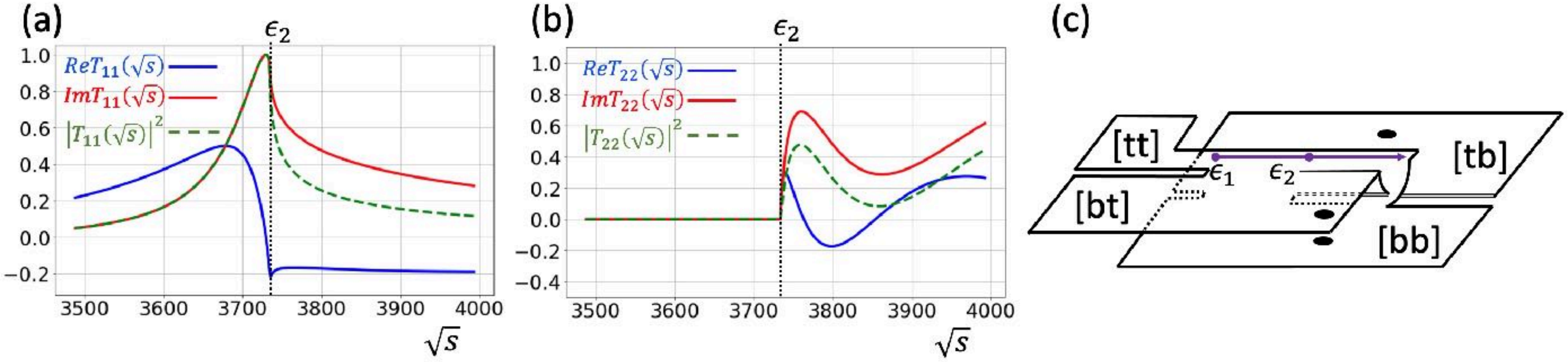
Noticeable saturation

Can this be improved?

Intrinsic ambiguity in the lineshape



Amplitude with one pole in [bt] sheet.



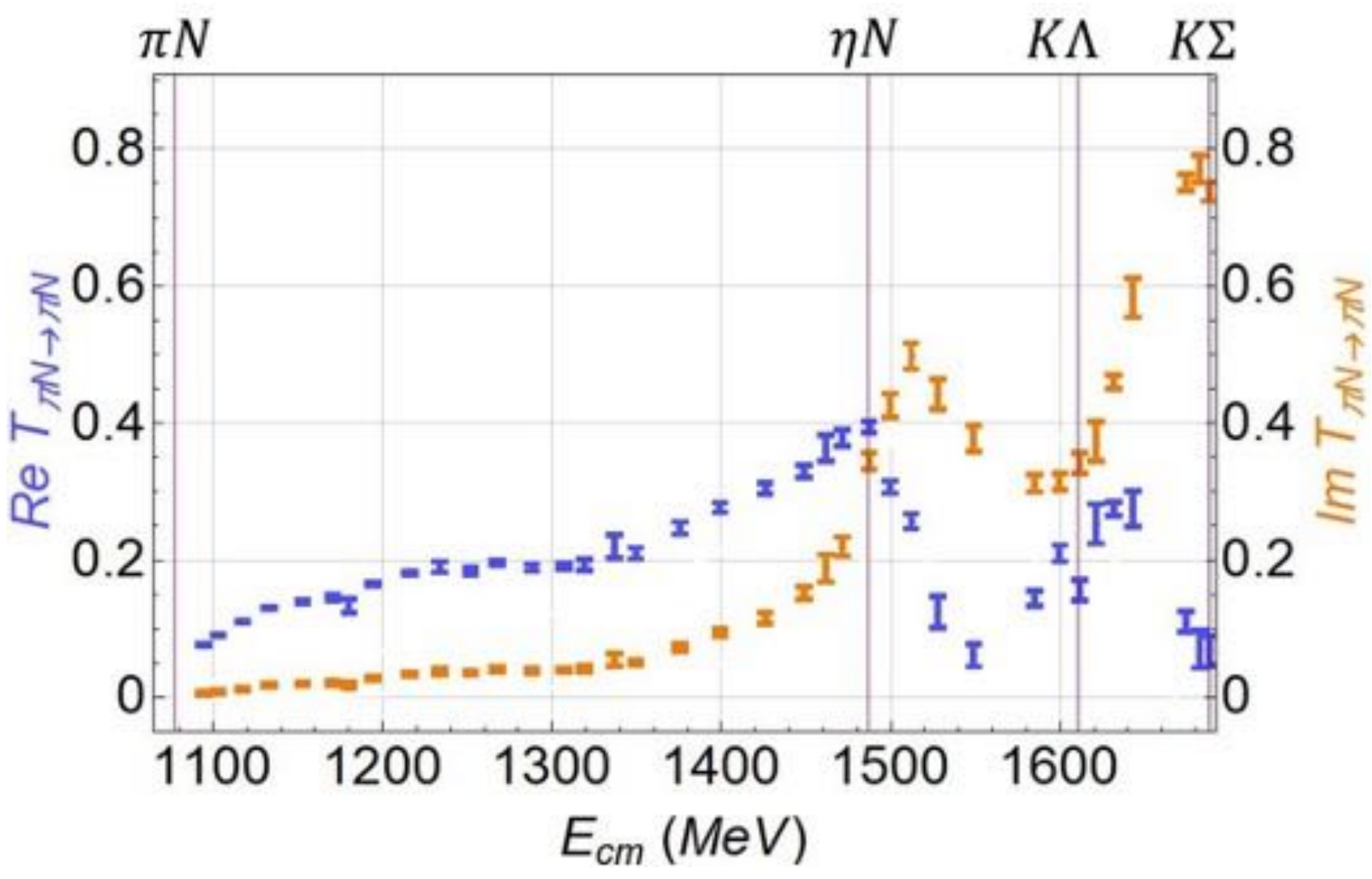
Amplitude with one pole in each unphysical sheet. All poles with the same real and imaginary parts.

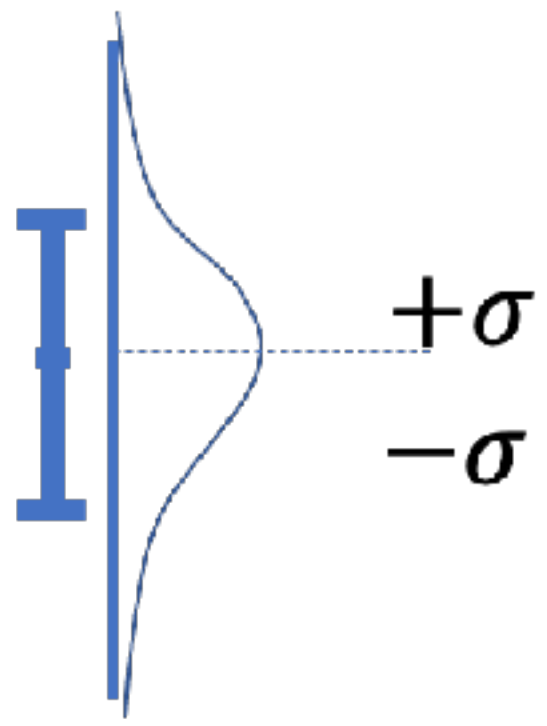
Identical lower channel amplitude.

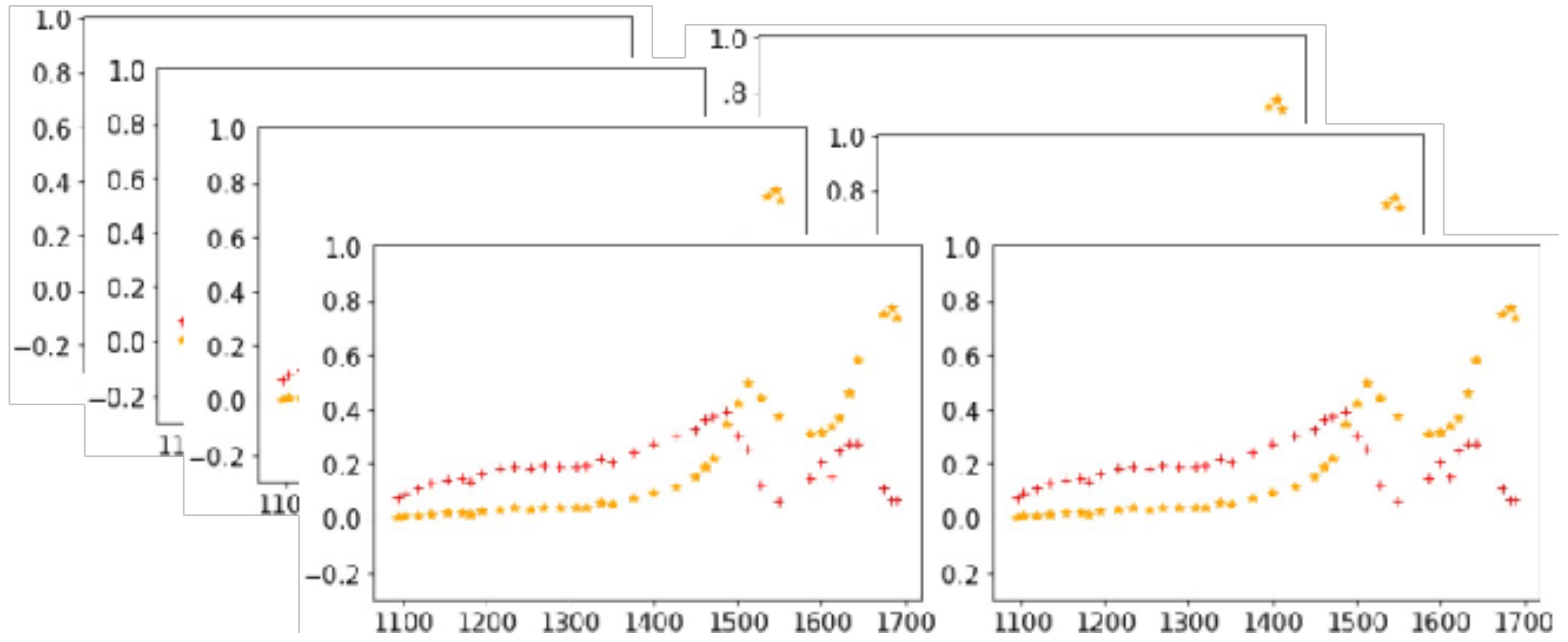
Higher channel amplitude can be distinguished.

The only way to improve the DNN performance is to include the higher (or cross) channel amplitude.

Inference stage: application





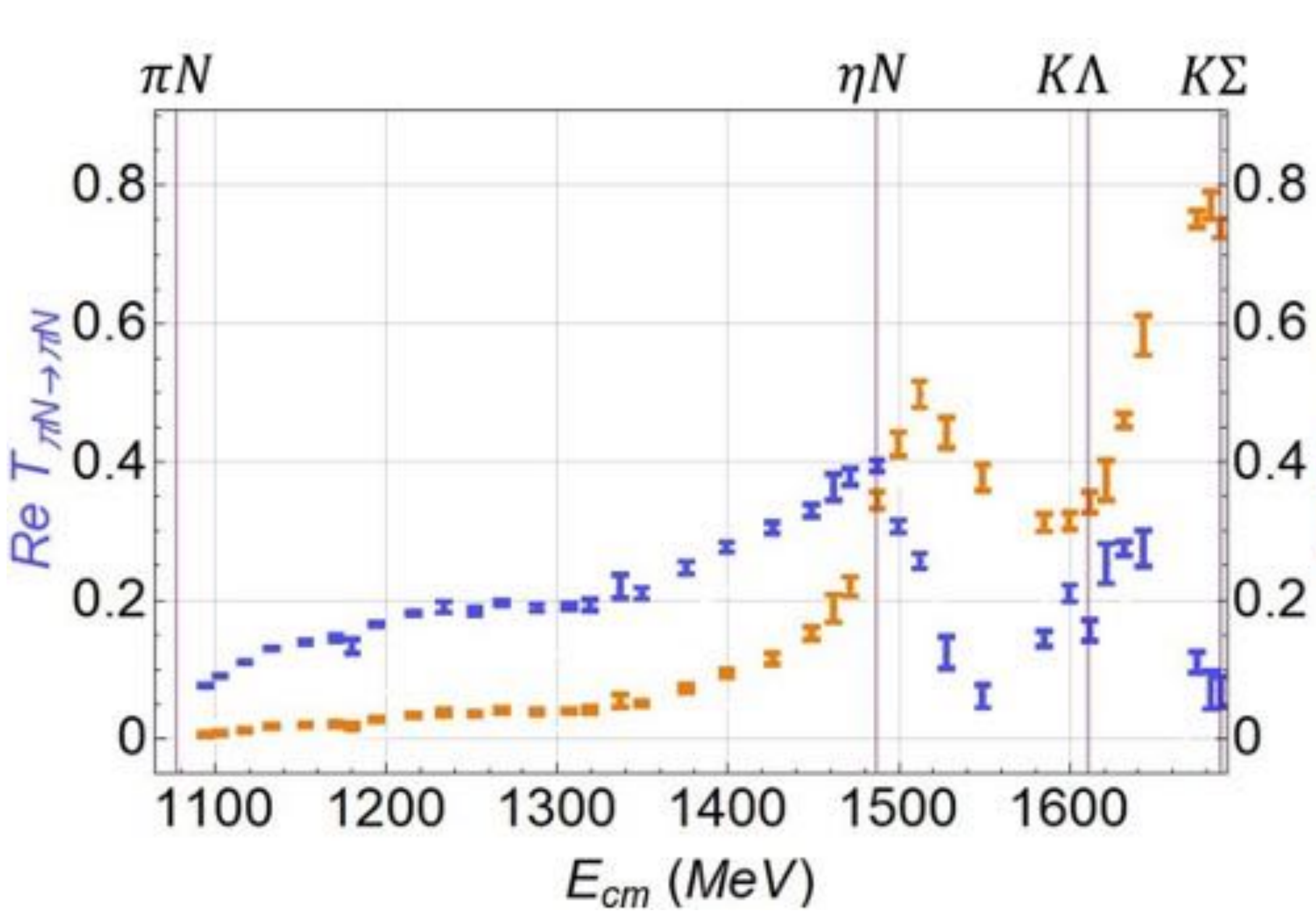


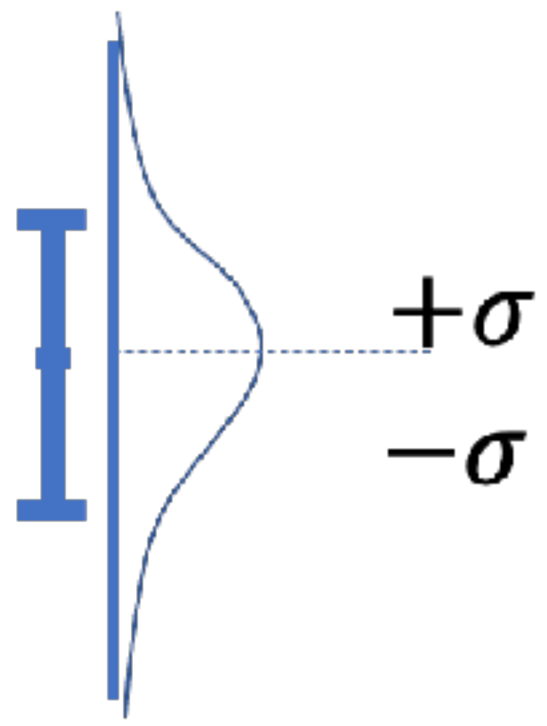
- Draw points from each error bar using a Gaussian distribution.
- Construct inference amplitudes from the experimental data using the drawn points.
- Feed the inference amplitudes to the trained DNN.

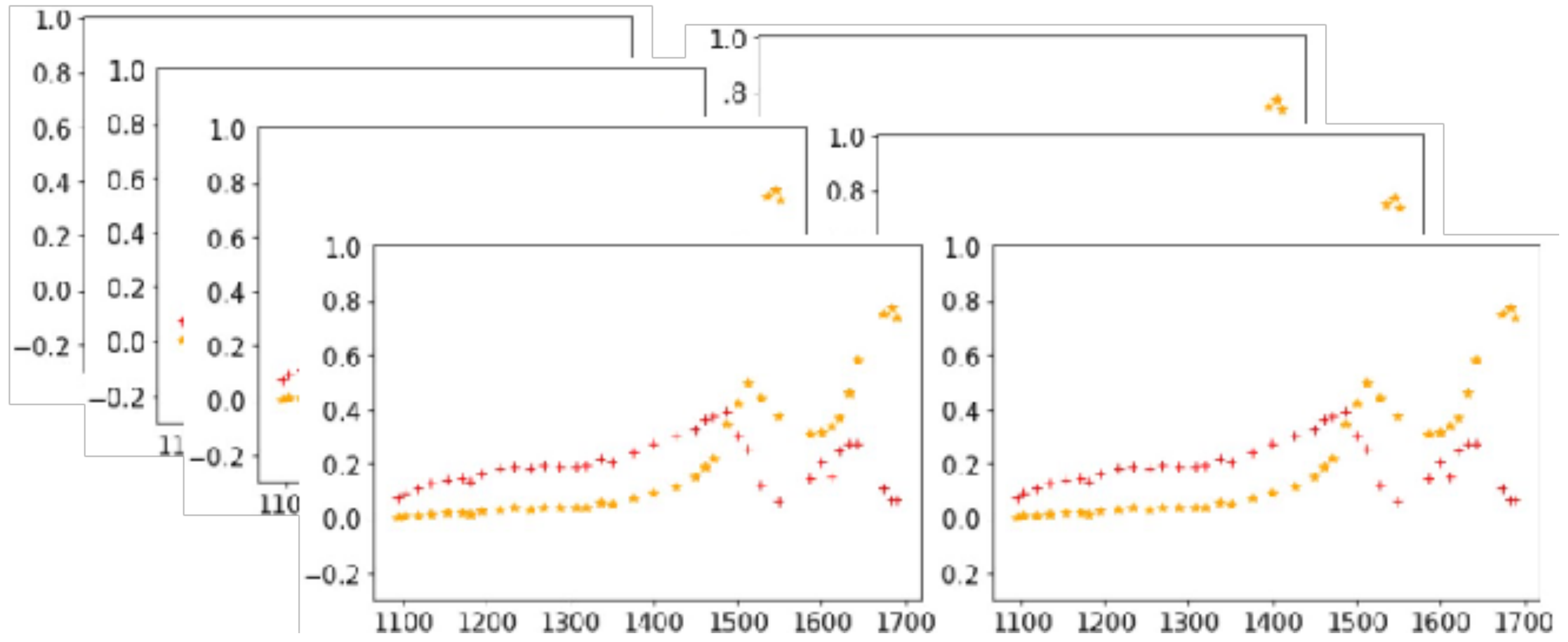
Interference on 10^6 amplitudes

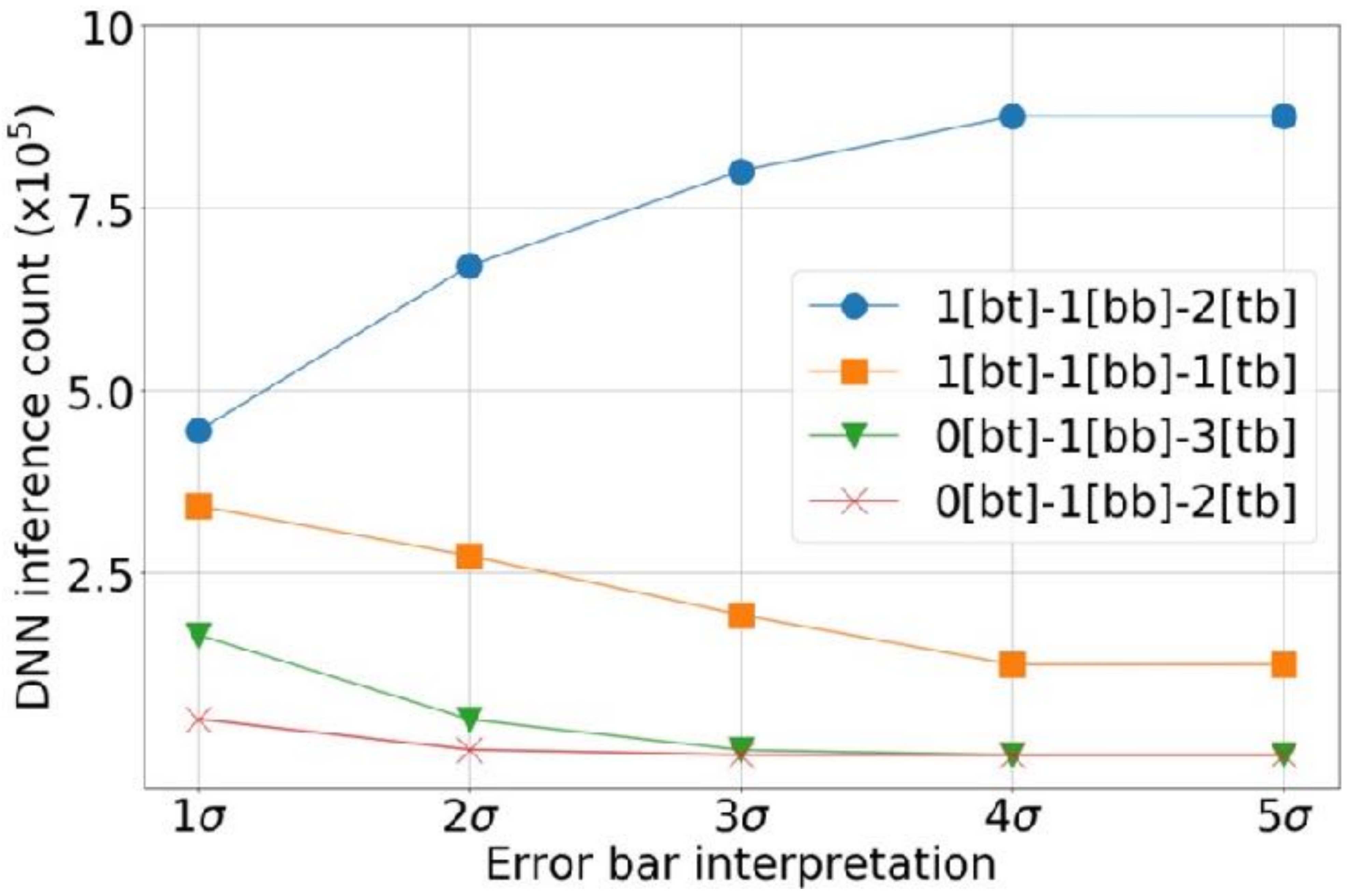
- 44.6% 1[bt]-1[bb]-2[tb]
- 34.1% 1[bt]-1[bb]-1[tb]
- 16.4% 0[bt]-1[bb]-3[tb]
- 04.9% 0[bt]-1[bb]-2[tb]

Inference stage: application





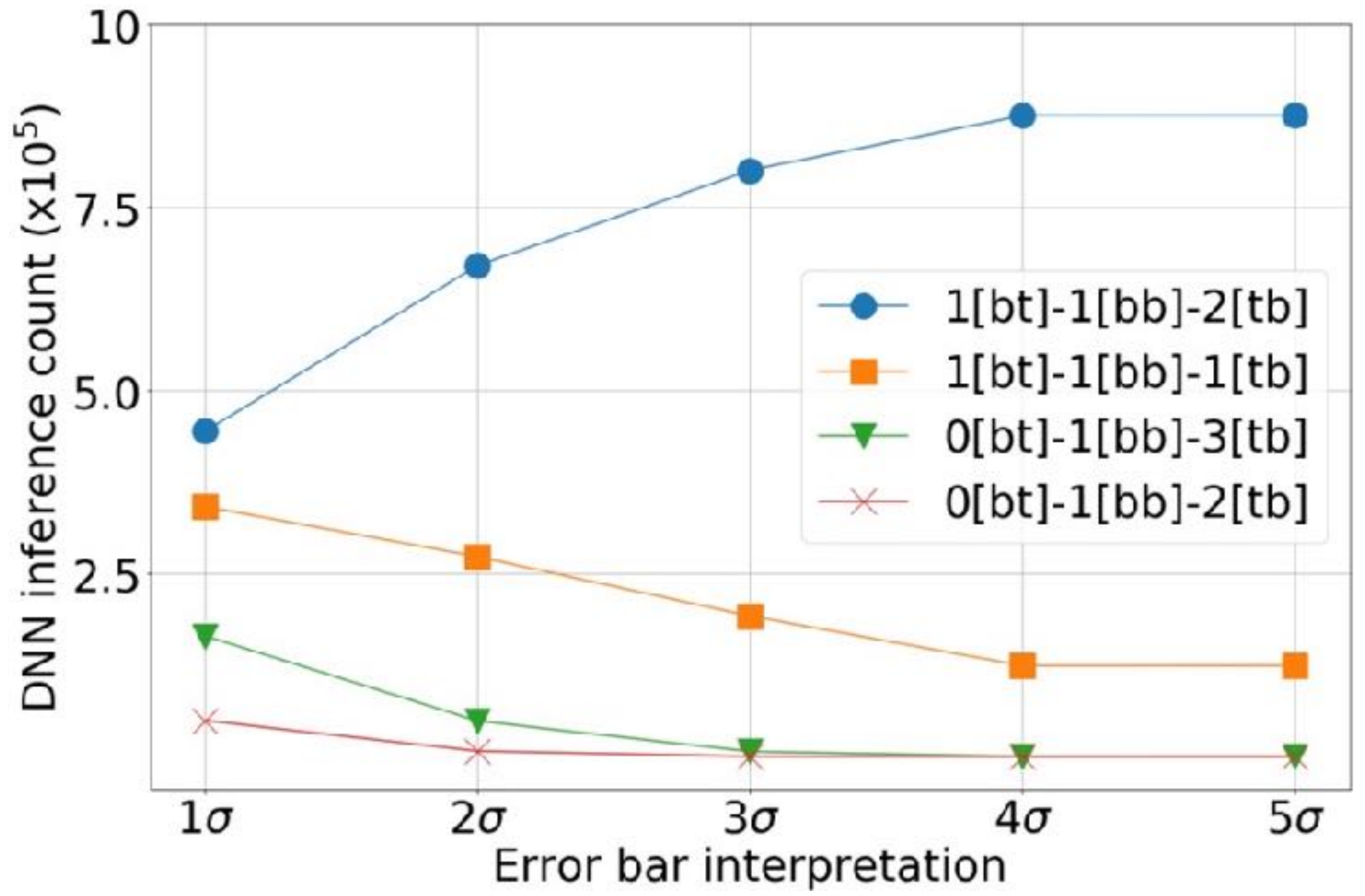




Interference on 10 ^ 6 amplitudes
Using uniform distribution

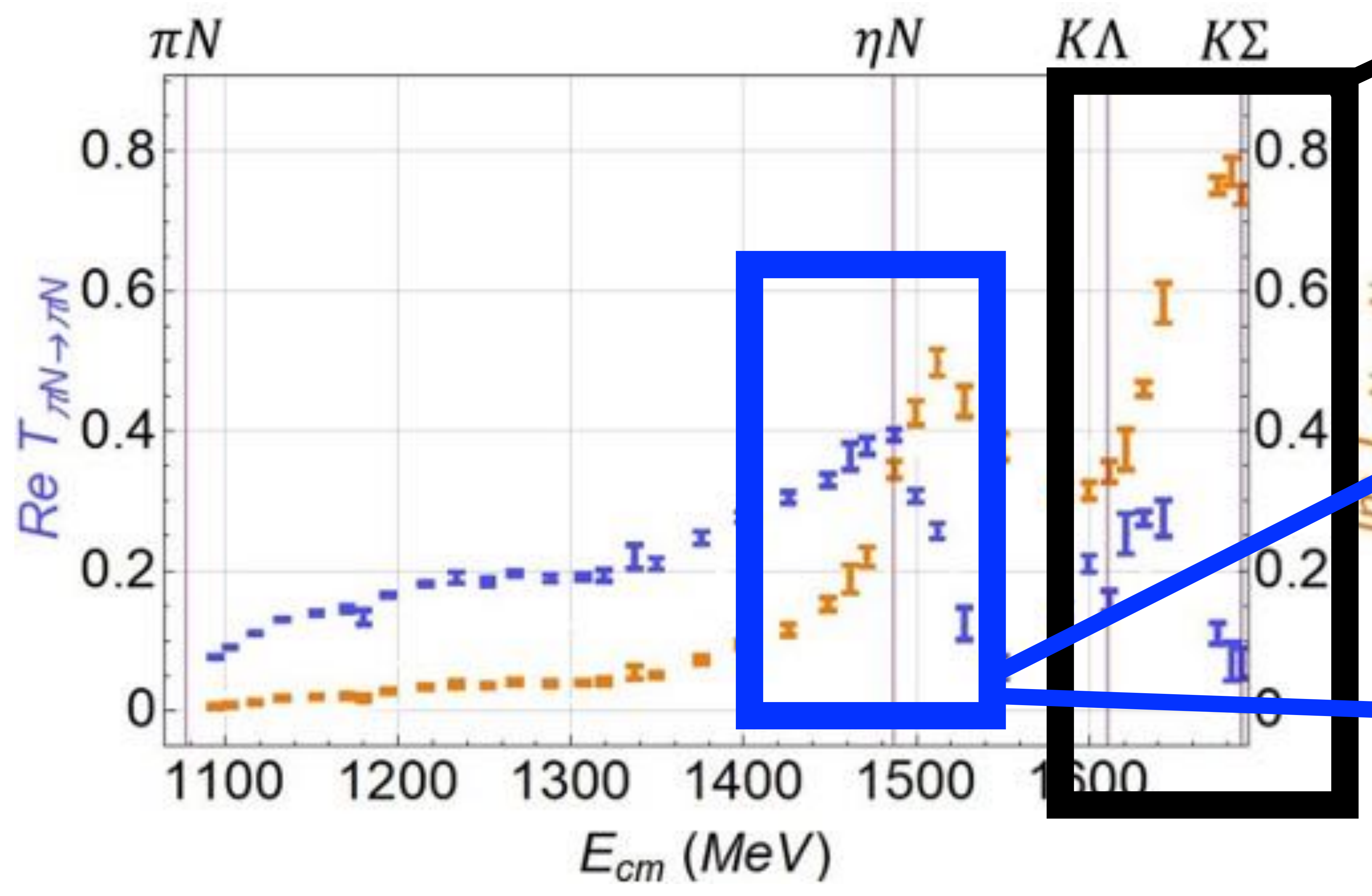
- 60.3% 1[bt]-1[bb]-2[tb]
- 30.9% 1[bt]-1[bb]-1[tb]
- 07.5% 0[bt]-1[bb]-3[tb]
- 01.3% 0[bt]-1[bb]-2[tb]

Interpretation of results



Dominant pole model: 1[bt], 1[bb] and 2[tb]

The detected [bb] and [bt] poles are the closest RS to the scattering region.



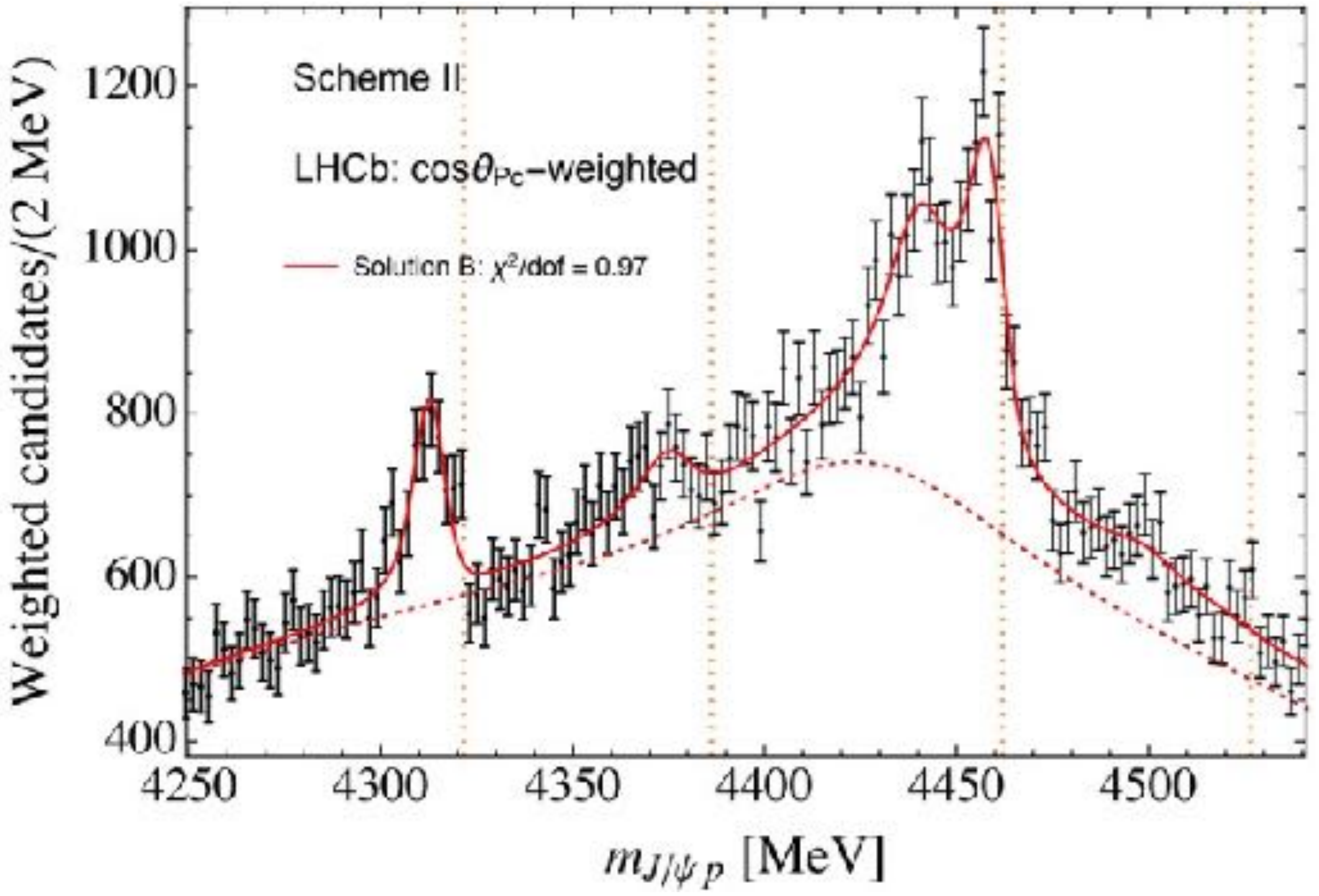
Peak is above the second threshold (ηN). Can only be caused by a [bb] pole

If [bt] pole is close to ηN threshold, the peak should reach the unitarity limit. Thus, [bt] pole is NOT the cause of ηN enhancement.

Can be attributed to the detected [tb] pole.

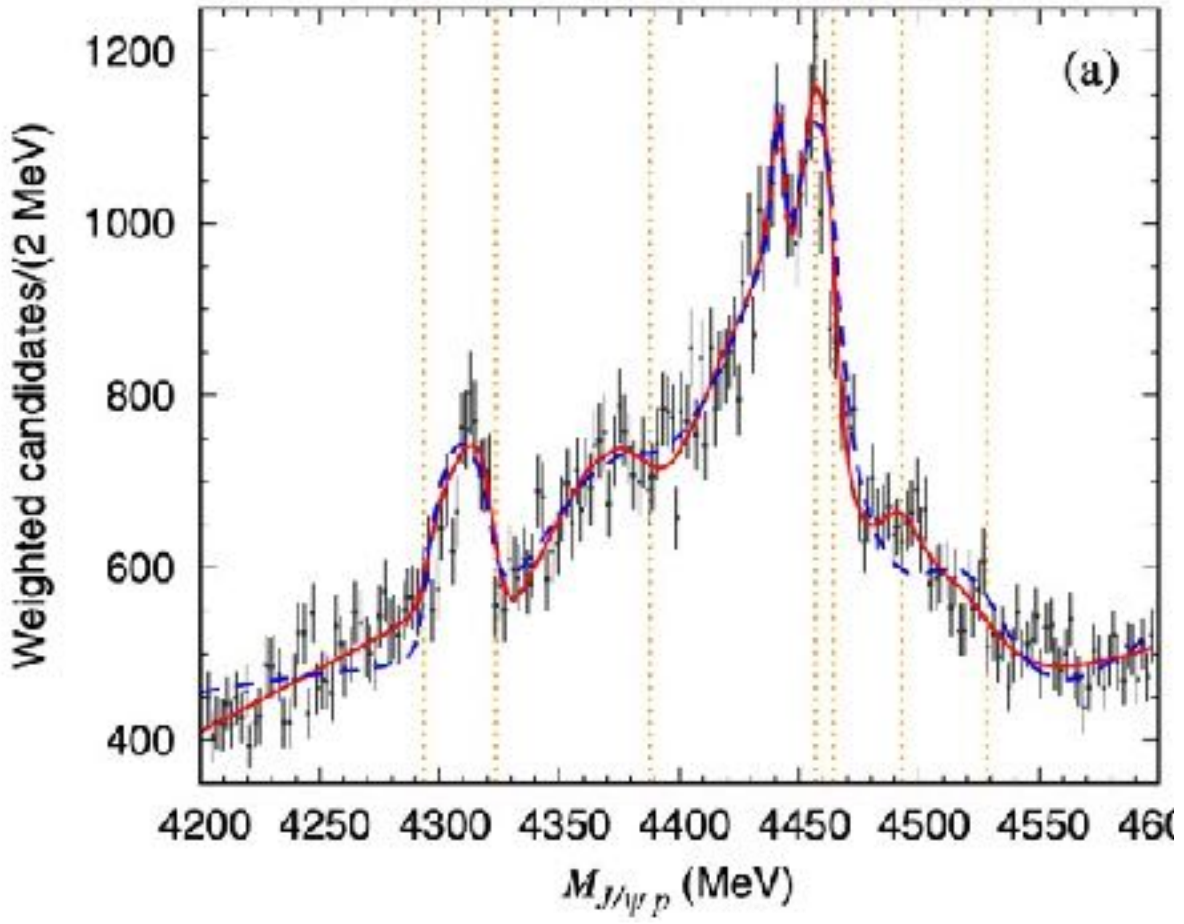
Summary and Outlook

- We can teach DNN to recognize the pole structure of a given amplitude.
- Deep learning can be used as unified model selection framework.
- No *a priori* assumptions is made on the detected poles since they are produced independently in the training dataset.



Molecular picture of Pc states

Phys. Rev. Lett. 124 (2020)



Double triangle singularities

Phys. Rev. D. 103 (2021)

- Same data, almost the same quality of fit but two conflicting models.
- Which is a better description of the data?

- Maybe DNN can give an unbiased answer. (Stay tuned!)

Thank you for listening!