Scattering Amplitude Analysis Using Neural Networks

Denny Sombillo

National Institute of Physics, University of the Philippines Diliman

Yoichi Ikeda
Toru Sato
Atsushi Hosaka

DLBS, YI, TS, AH PRD 102 016024 (2020)
DLBS, YI, TS, AH Few-Body Syst. 62, 52 (2021)
DLBS, YI, TS, AH PRD 104 036001 (2021)

Approaches in hadron spectroscopy

JPAC
10.1016/j.ppnp.2022.103981
10.48550/arXiv.2112.13436

Top-down approach

Theory/ Model

↓

Generate hadron spectrum

↓ Is the theory a good description of experiment?

Compare with experimental data

Bottom-up strategy

Start with experimental data

↓

Use a generic amplitude ansatz

↓

Extract physics by fitting

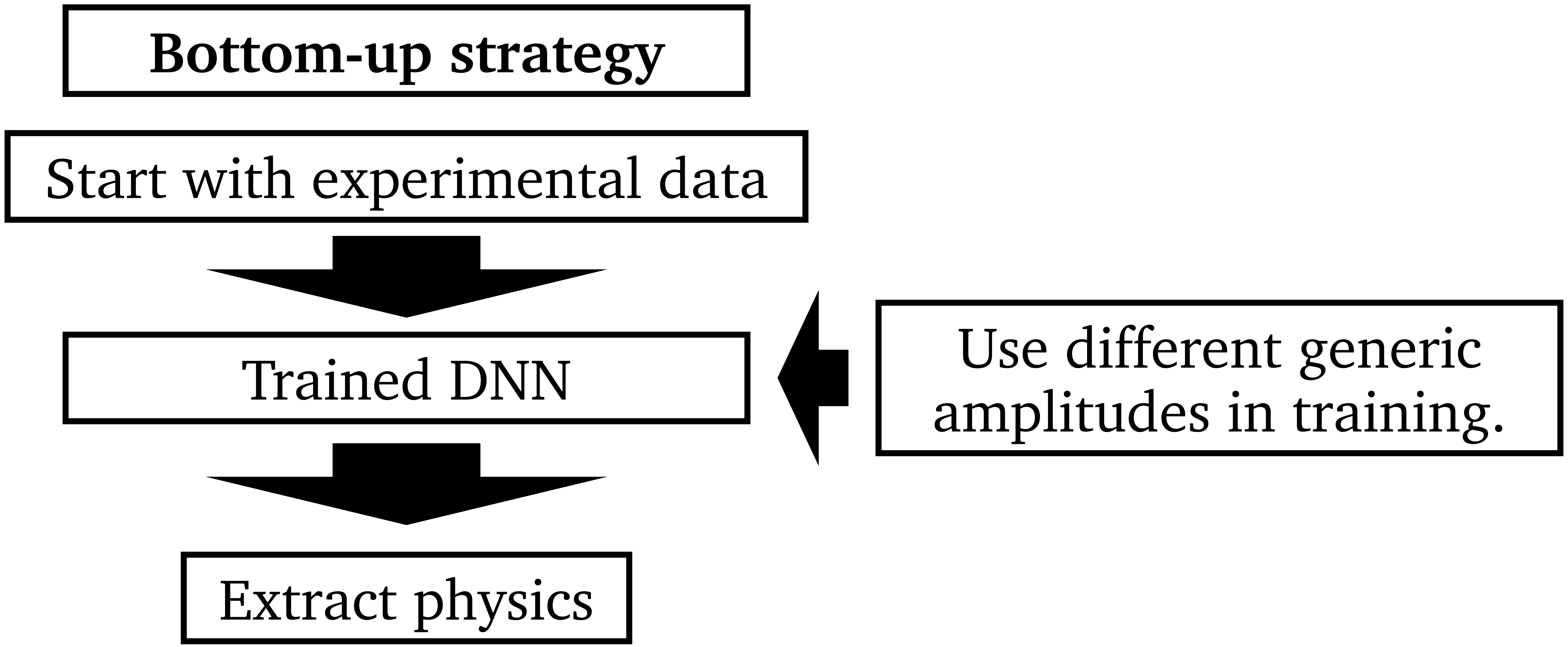
In the bottom-up strategy, we are not interested in explaining everything.

What is the experiment telling us?

Approaches in hadron spectroscopy

JPAC
10.1016/j.ppnp.2022.103981
10.48550/arXiv.2112.13436

The bottom-up strategy can be augmented by the deep learning approach.



Classifier-type DNN already in use:

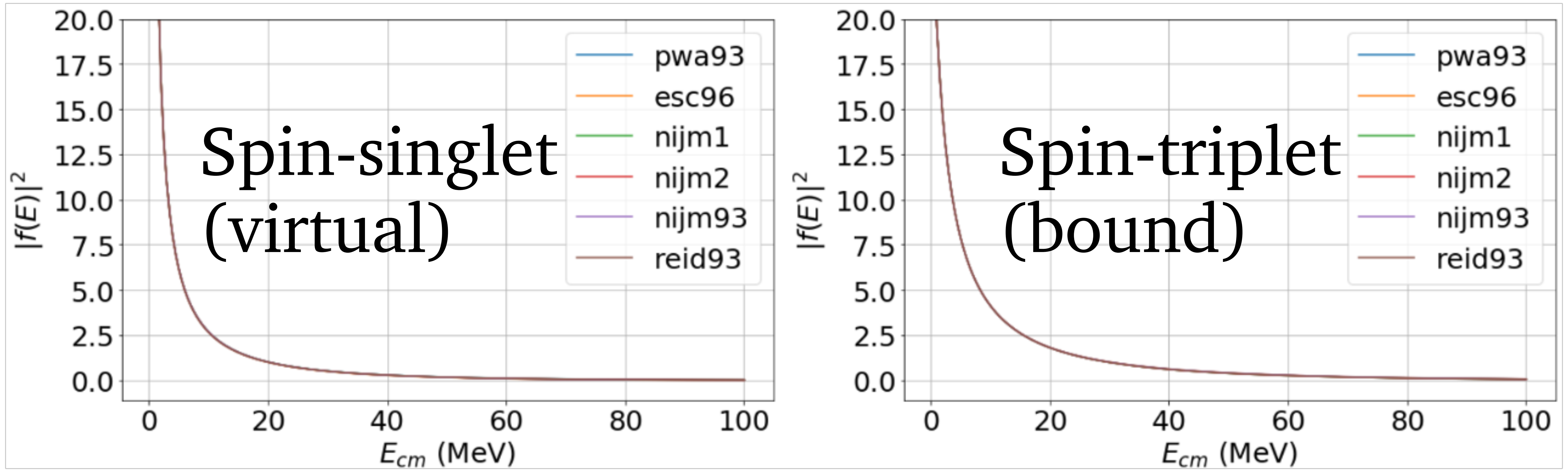
- Bound-virtual for single channel — DLBS, YI, TS, AH PRD 102 016024 (2020); DLBS, YI, TS, AH Few-Body Syst. 62, 52 (2021)
- Pc(4312) — JPAC Collaboration PRD 105 L091501 (2020)
- Pole configuration — DLBS, YI, TS, AH PRD 104 036001 (2021)

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



For near-threshold pole:

$$k \cot\delta \sim -1/a \text{ (constant)}$$

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

There is no way to discriminate a bound state pole enhancement with a virtual enhancement using only $|f(k)|^2$ on the scattering region.

In addition to the near-threshold pole, the S-matrix can have distant singularities on the unphysical sheet.

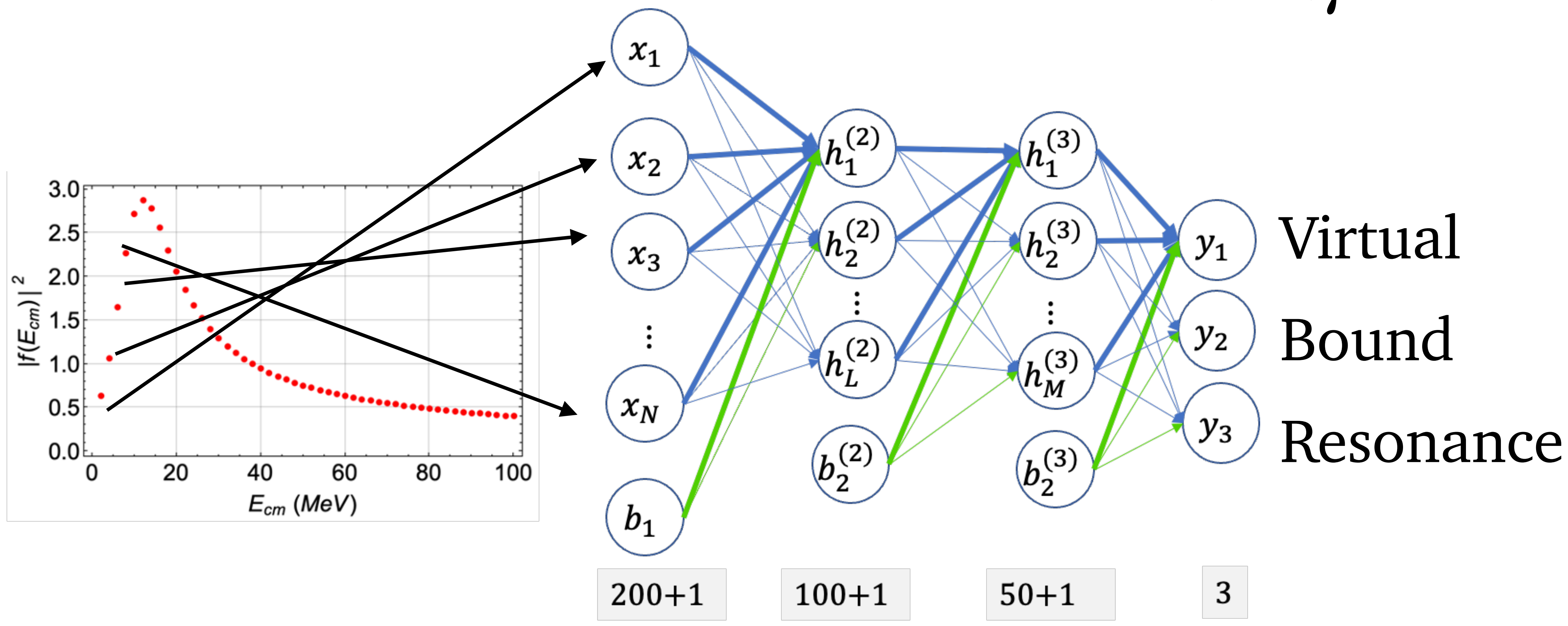
Use different (unitary, analytic) backgrounds to help DNN distinguish bound and virtual enhancements.

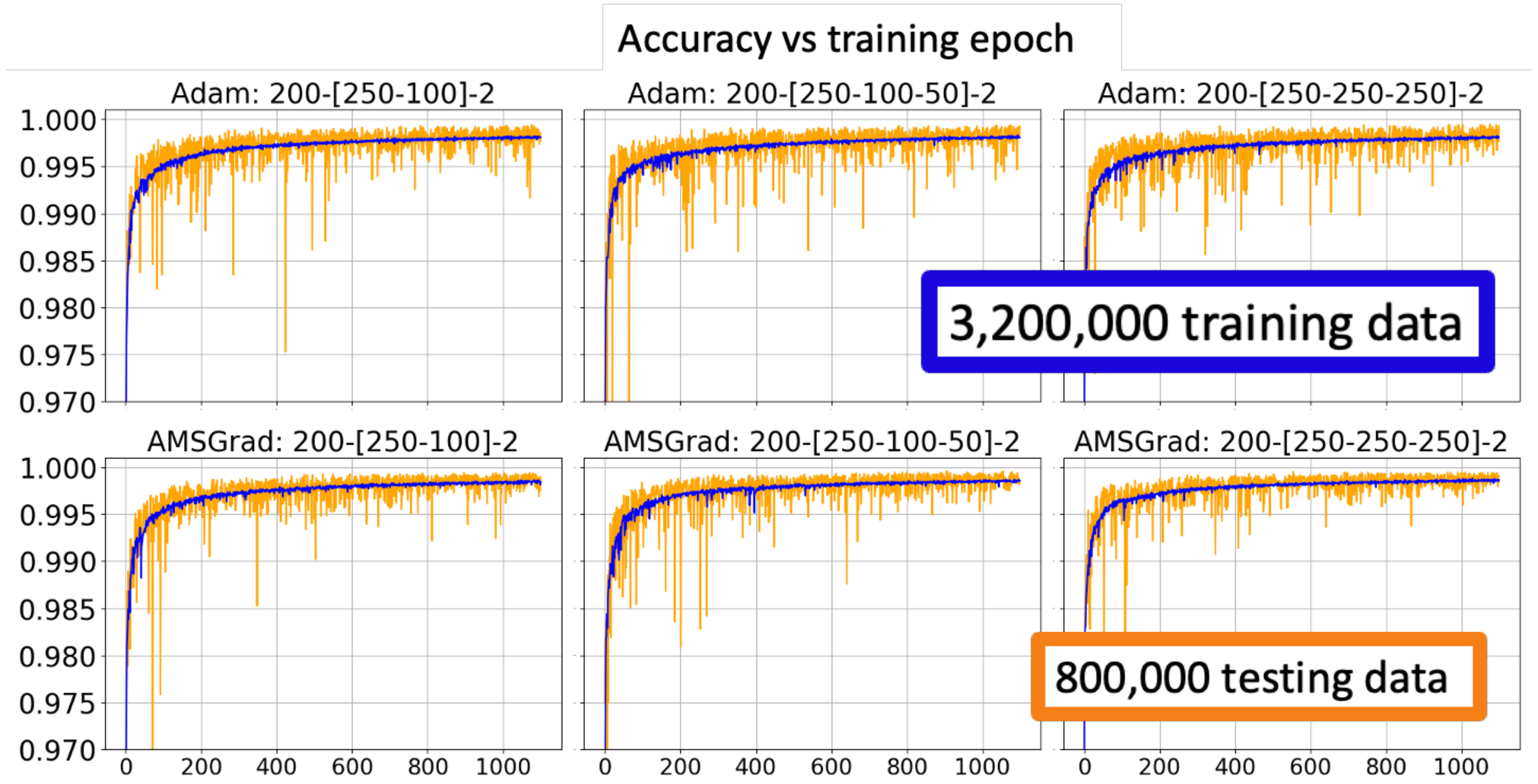
$$S(k) = \exp\left[2i\delta_{bg}(k)\right]\frac{k+i\gamma}{k-i\gamma}$$

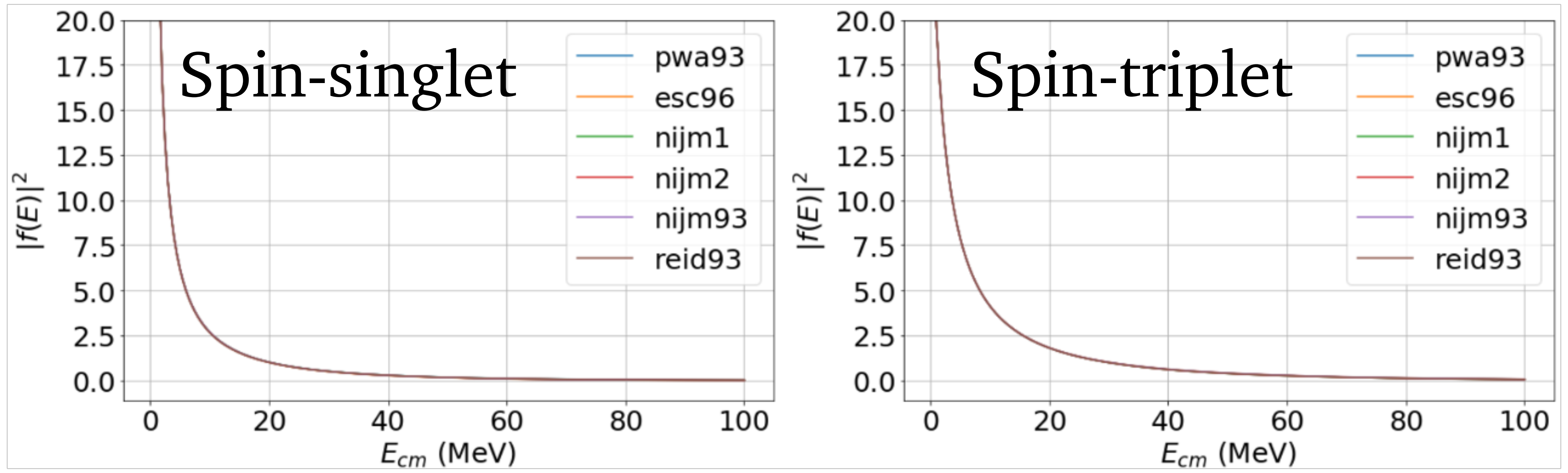
Deep learning: alternative analysis tool

Optimize parameters of DNN using mock amplitudes

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







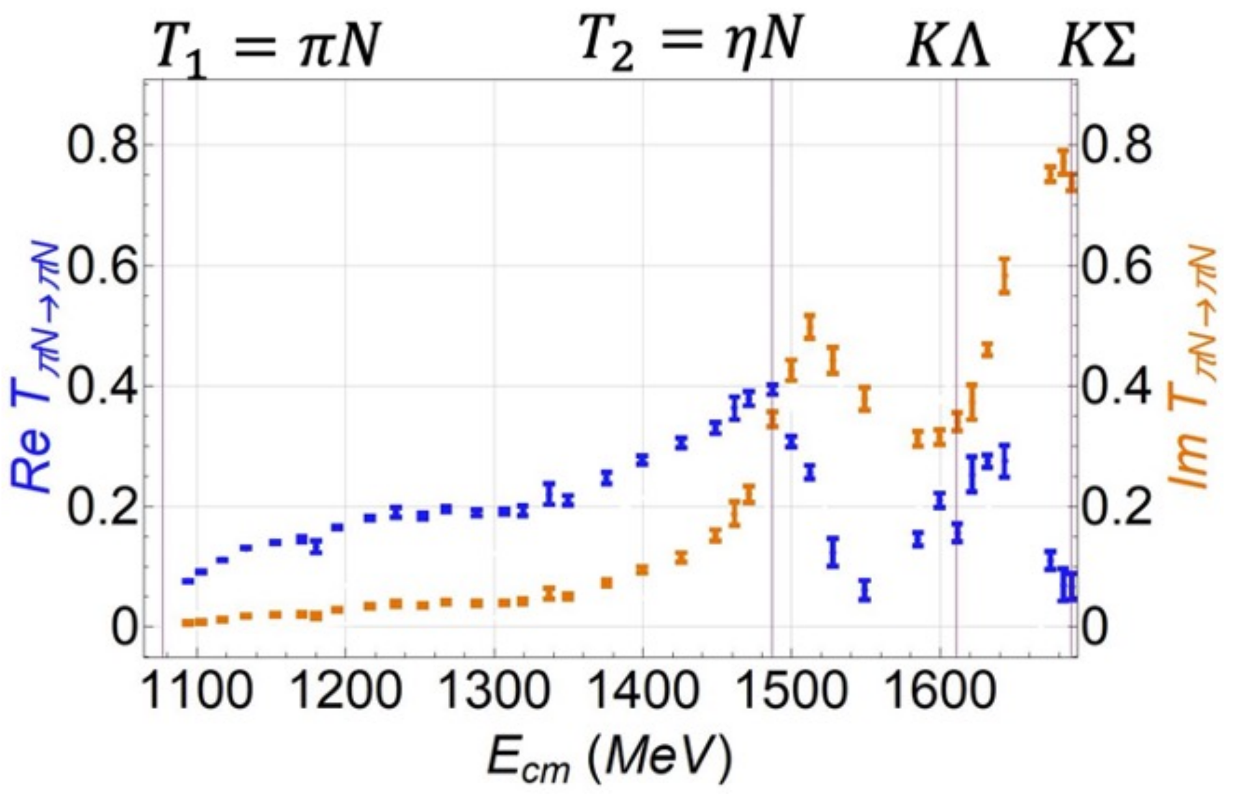
Trained and validated DNN
deployed to probe the nucleon-nucleon

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

Pole structure of coupled channel system

The nature of enhancement can only be deduced in a model-dependent way.

We can at least extract the pole configuration in a model-independent way.



How many nearby poles in each Riemann sheet are needed to reproduce the experimental data?

Inspired by:

Two-pole structure of $\Lambda(1405)$

T. Hyodo and U. -G. Meißner PDG 2021 review

Pole-counting argument - nature of $f_0(S^*)$ (aka $f_0(980)$)

D. Morgan and M. R. Pennington PRD 48 1185 (1993)
D. Morgan Nuc. Phys. A 543 632-644 (1992)

Badalyan, *et. al.*, Phys. Rep. 82, 2, 31-177 (1982)

Table 1.1
Classification of the two-channel poles. In the column "Origin" we indicate the type of one-channel pole which obtains by switching off the coupling to the second channel (with lower threshold)

Label	Origin	Location: sheet, energy half plane	Name
UBS	BS	II, lower	unstable bound state
IVS	VS	IV, upper	inelastic virtual state
BW	BW	III, lower	Breit–Wigner main
BW*	BW*	III, upper	Breit–Wigner conjugate main
BW_1	BW	II or IV, lower	Breit–Wigner shadow
BW_1^*	BW*	II or IV, upper	Breit–Wigner conjugate shadow
CC	infinity or distant singularity	II, III, or IV, lower and upper	coupled-channel pole

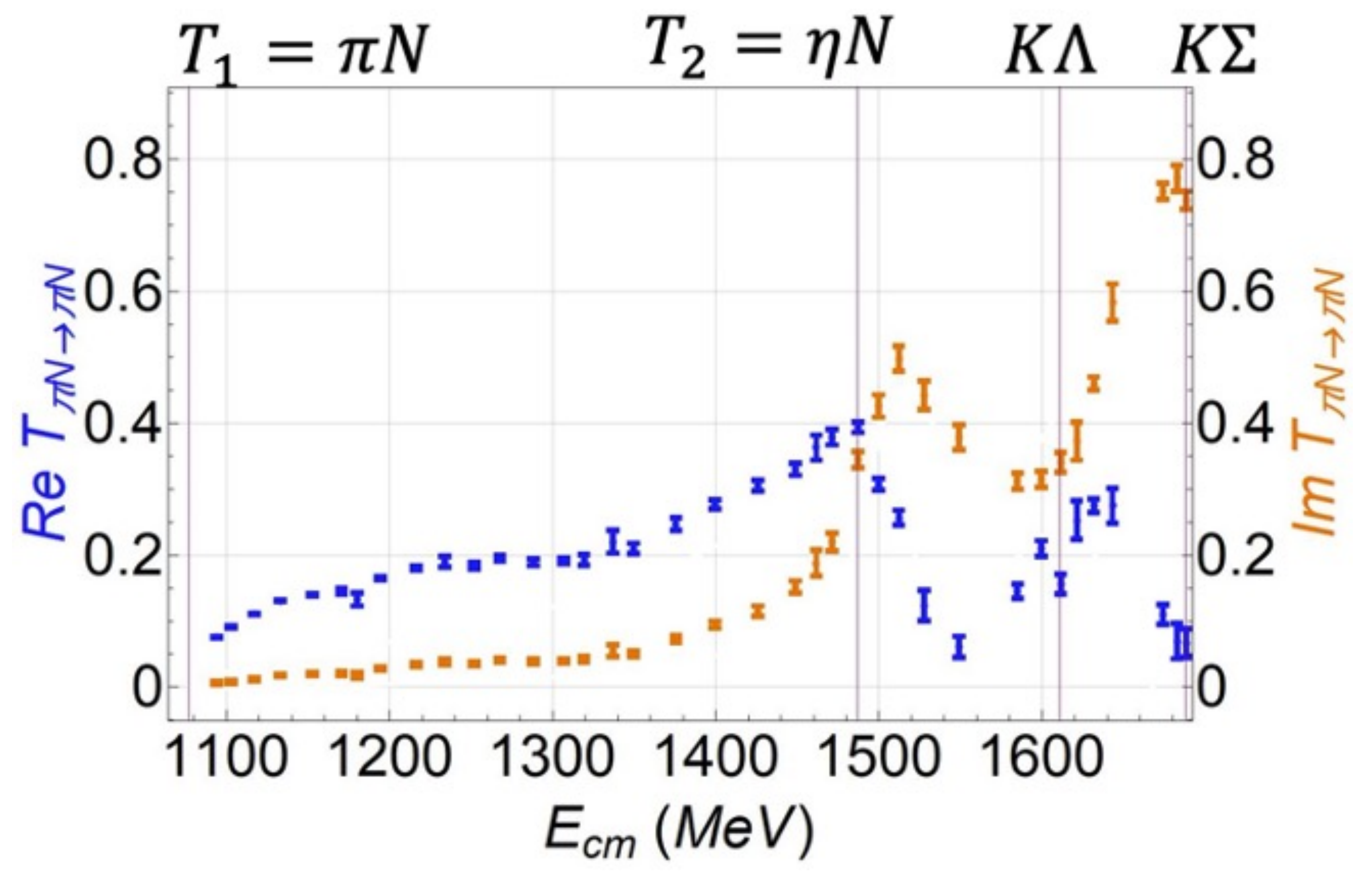
Pole structure of coupled channel system

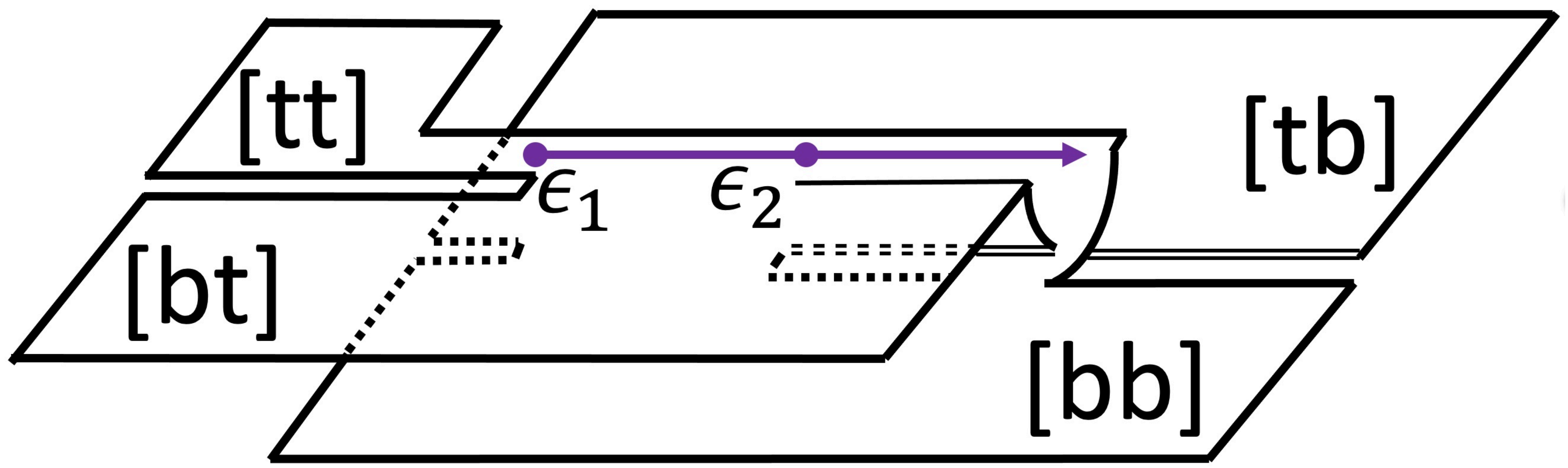
DNN task:
Count the number of nearby poles in each Riemann sheet using only the scattering data.

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

Two-channel case: 35 possible pole configurations

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 $[tb]$





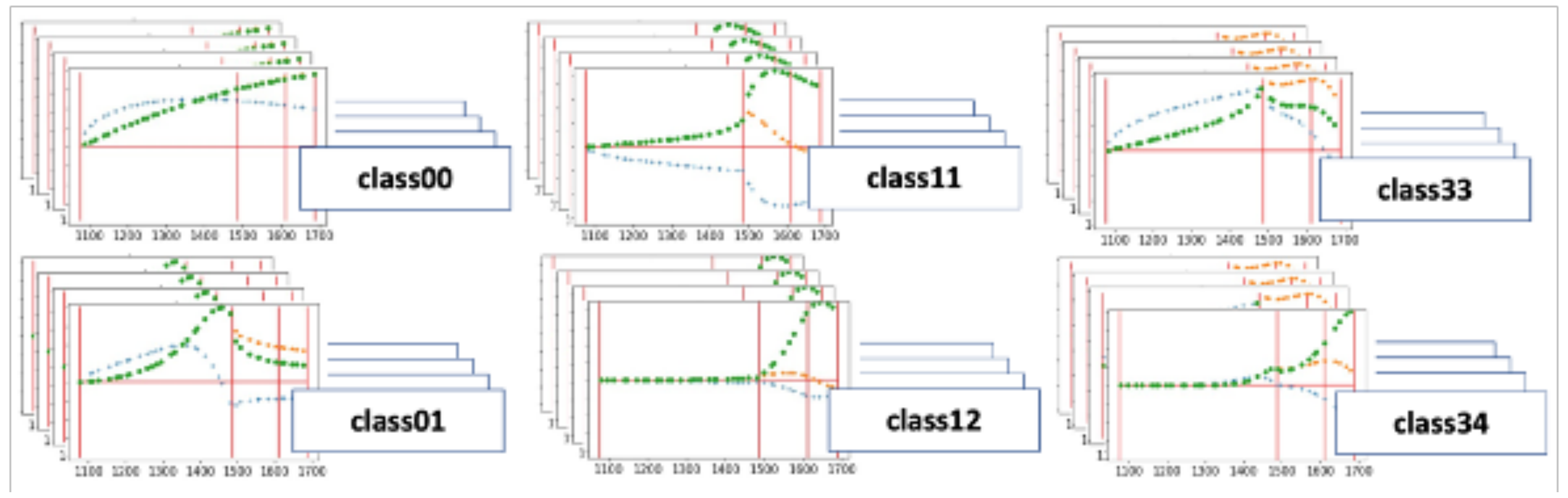
It is reasonable to expect that more than one pole-based model can describe the data due to the error bars.

Pole structure of coupled channel system

DL approach:

Generate the training dataset

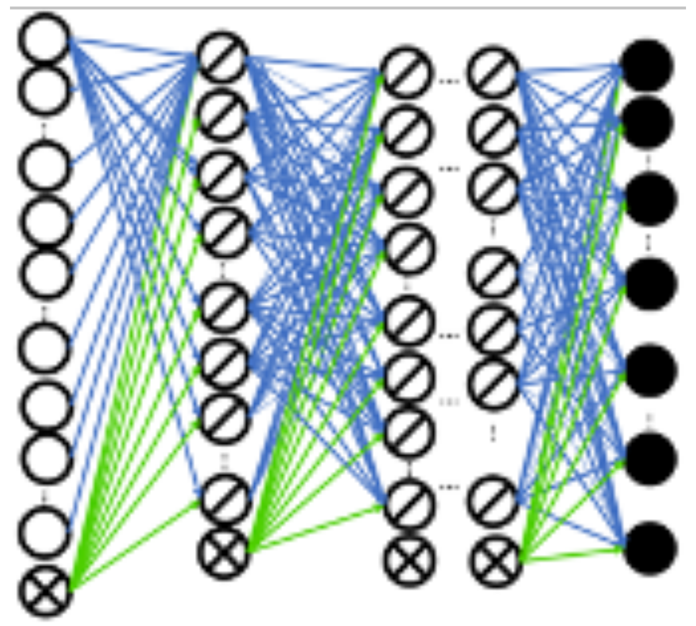
- Use only the general properties of S-matrix
- Include the energy uncertainty



Optimize the parameters of the deep neural network

Input layer:

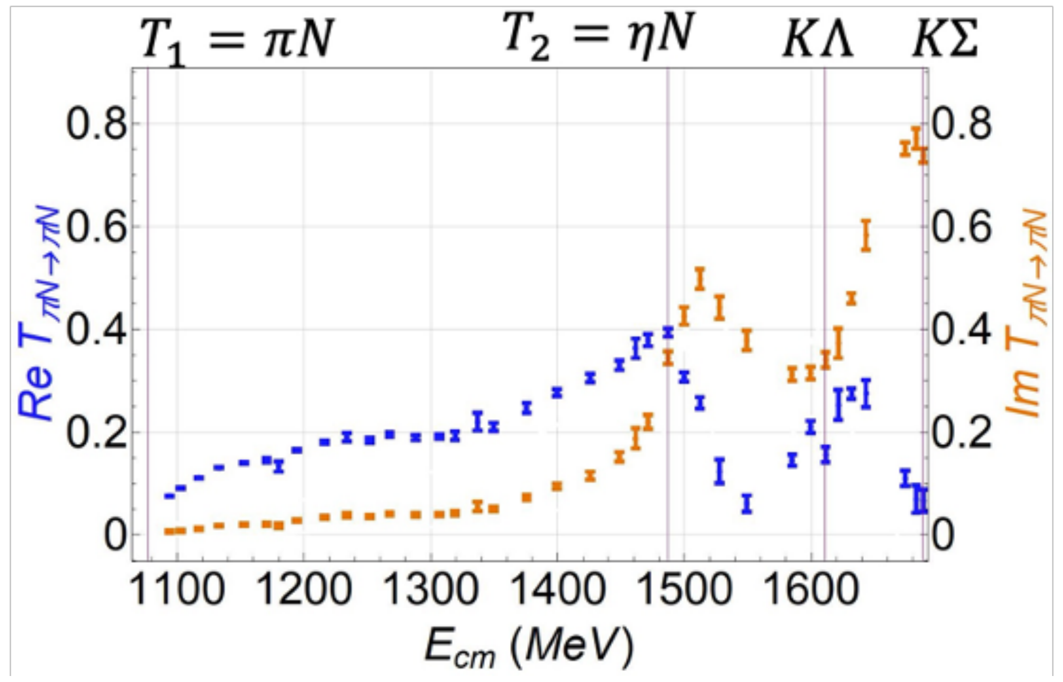
- Energy points
- Real part of amplitude
- Imaginary part of amplitude

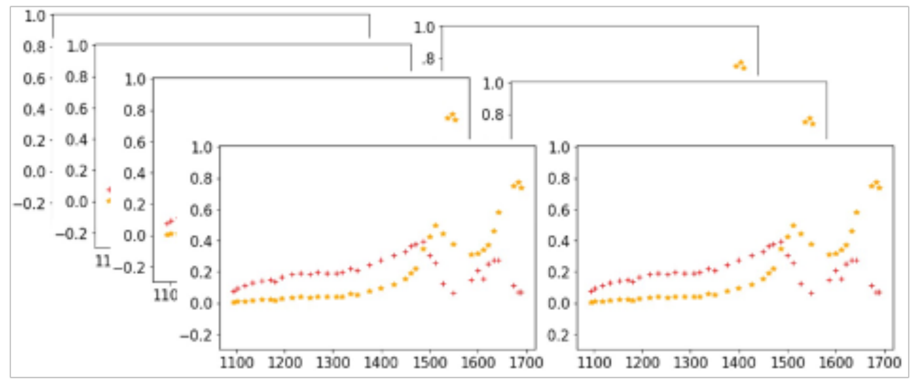


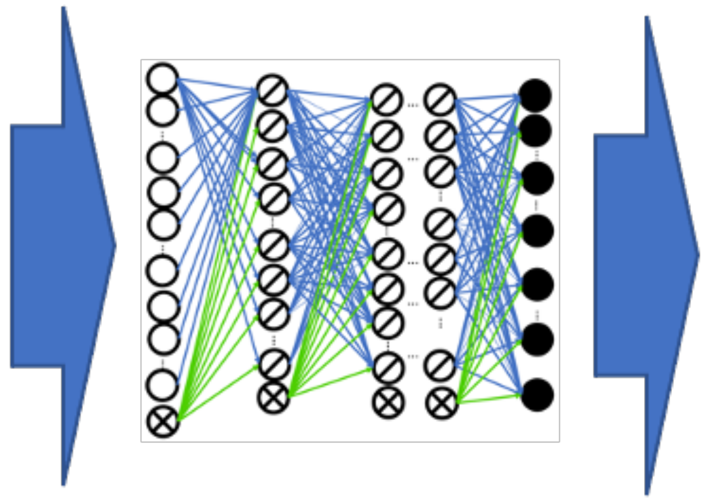
Output layer:

- Pole config. 0
- Pole config. 1
- Pole config. 2
- …
- …
- Pole config. 34

Deploy the trained DNN to extract model from the experimental data.







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)

General form of S-matrix:

- Hermiticity below the lowest threshold
- Unitarity
- Analyticity

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

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

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

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

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

The available experimental data will determine the relevant matrix element.

Ensure that only one nearby pole E_m is generated by each $D_m(p_1, p_2)$.

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

$\alpha_{1m}, \alpha_{2m}, \beta_{1m},$ and β_{2m} are related to E_m.

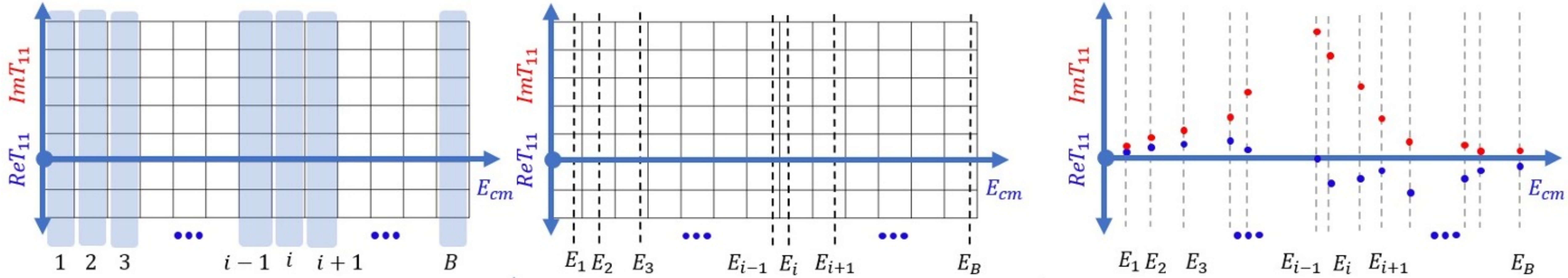
Do not let β_{1m} and β_{2m} be both positive. (Analyticity)

The RS is determined by the signs of β_{1m} and β_{2m}.

Extra parameter λ_m to push the other pole far from the scattering region without violating analyticity.

Training dataset (generation of model space)

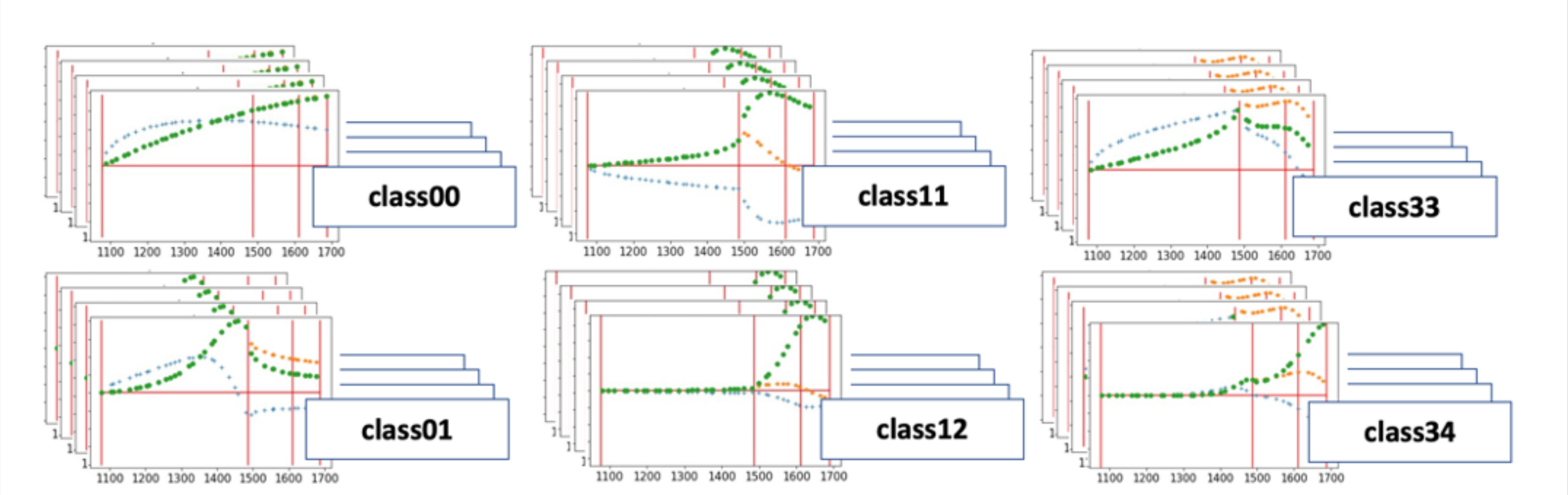
Incorporate uncertainty in the energy:



(1) Divide the energy range into B bins

(2) Choose a random energy value in each bin

(3) Calculate the amplitude for each random energy value



(4) Label each amplitude according to its pole-configuration

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

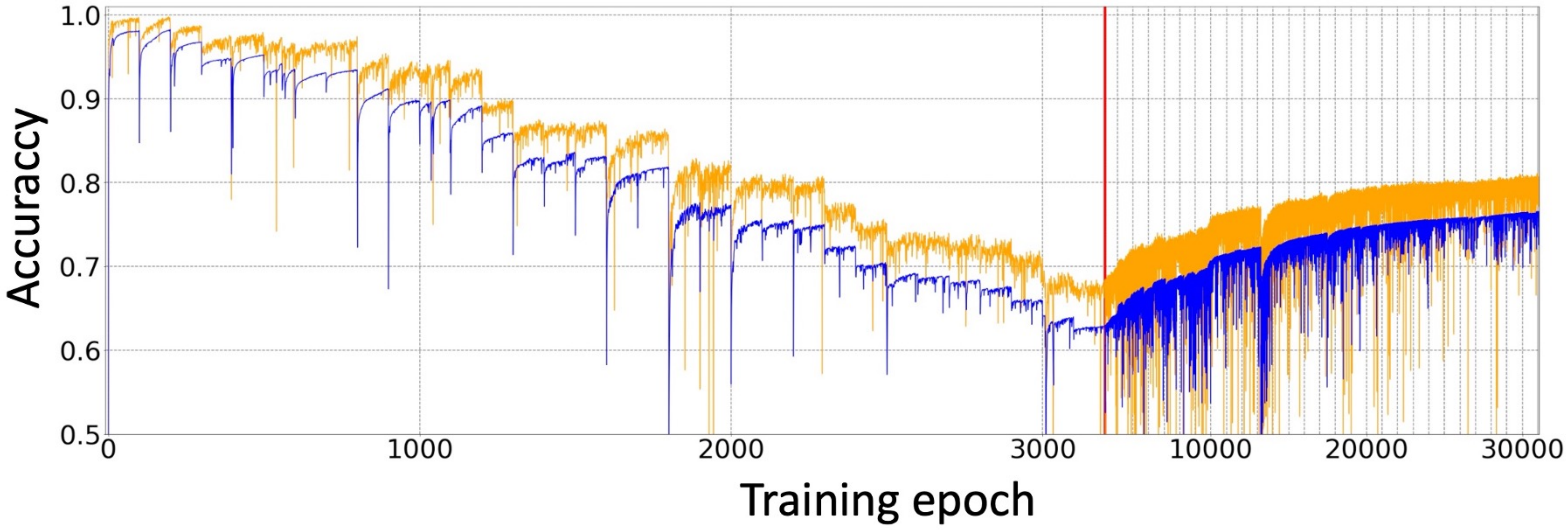
Chosen DNN architecture

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

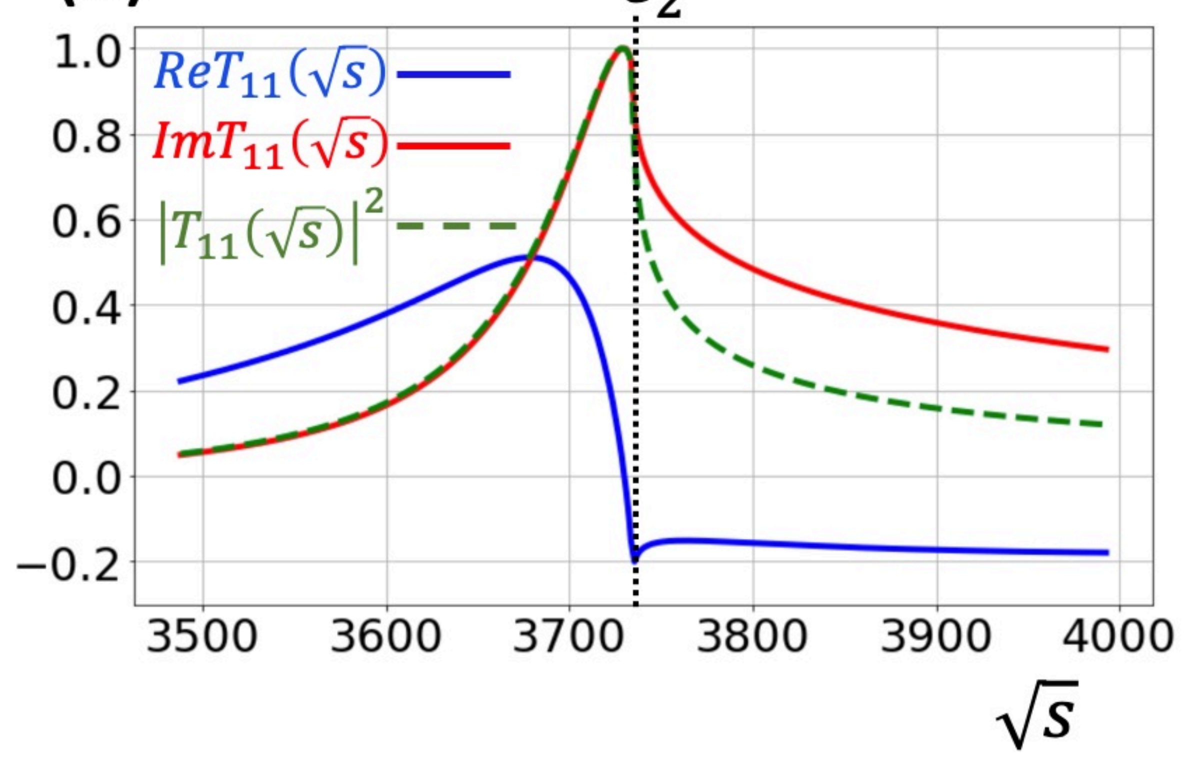


Noticeable saturation

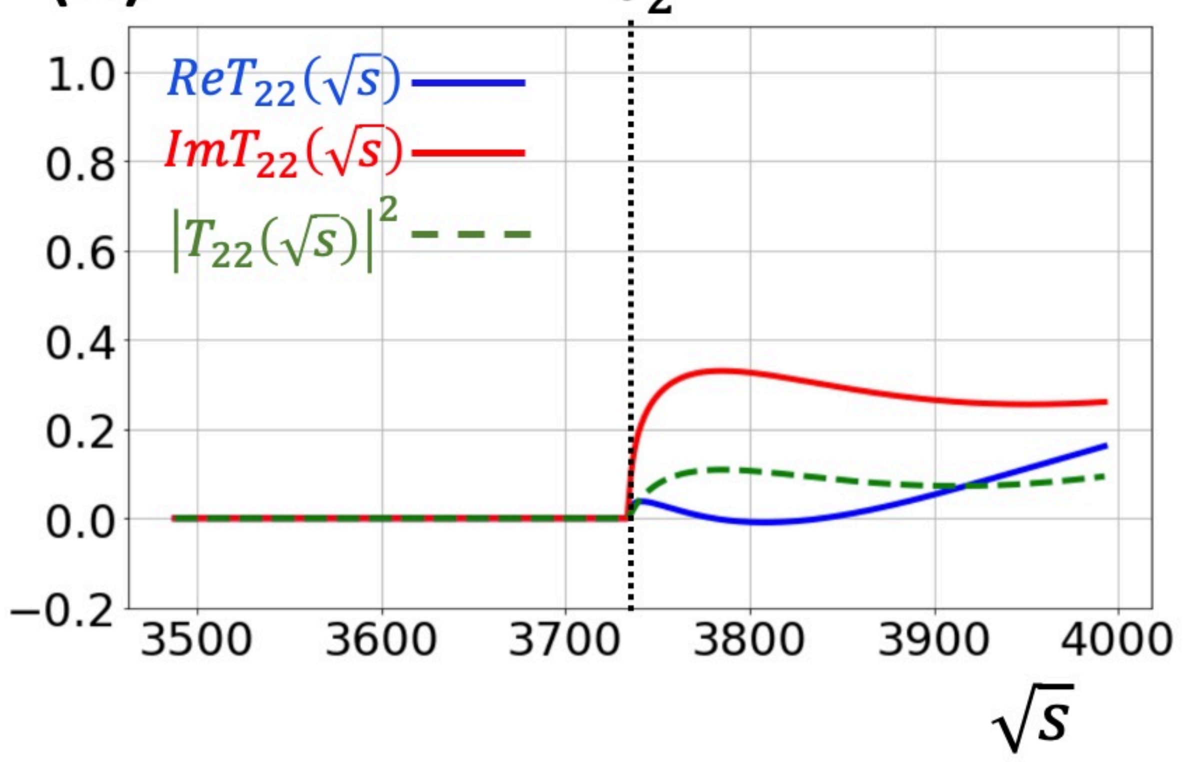
Can this be improved?

Intrinsic ambiguity in the lineshape

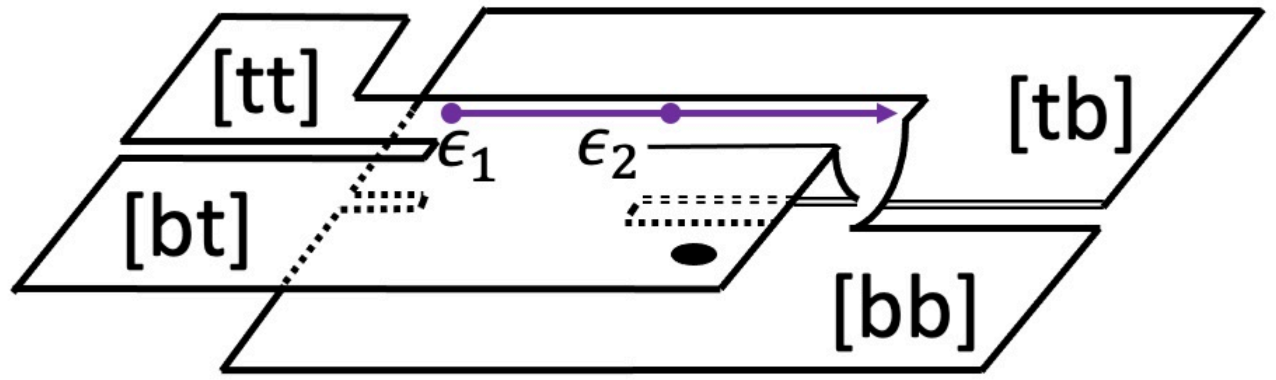
(a)



(b)

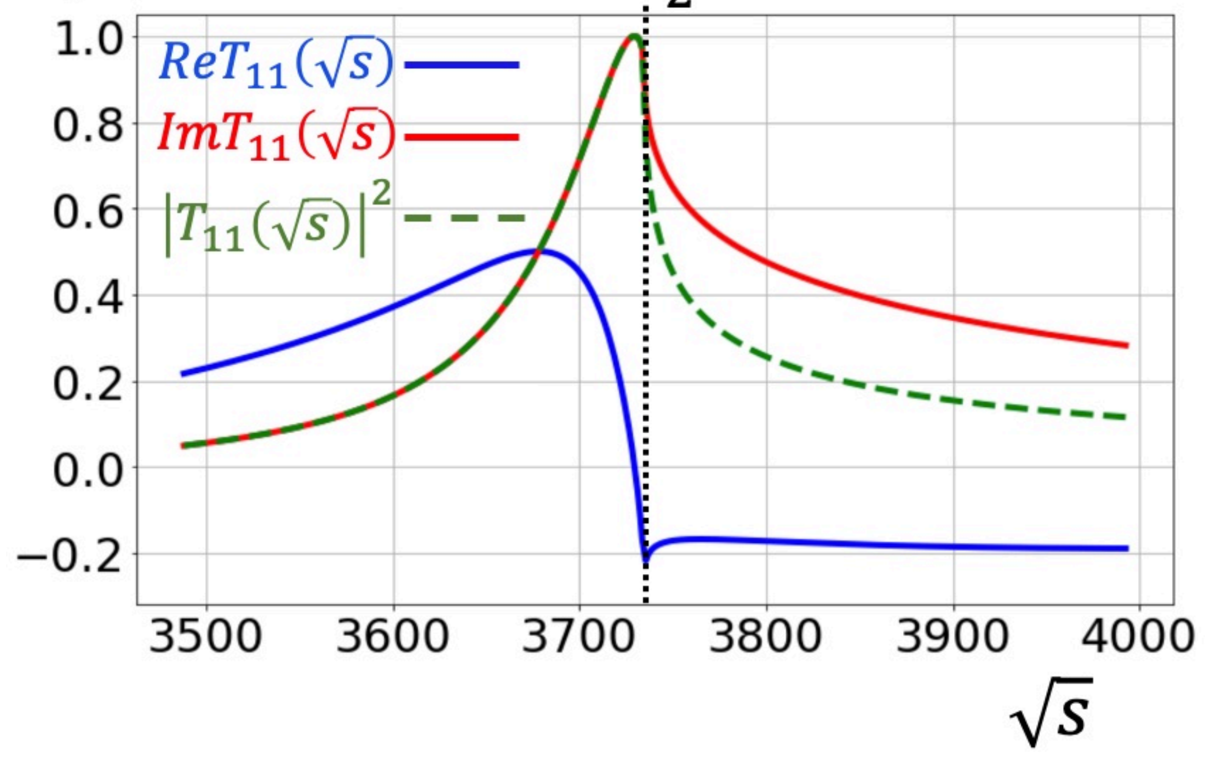


(c)

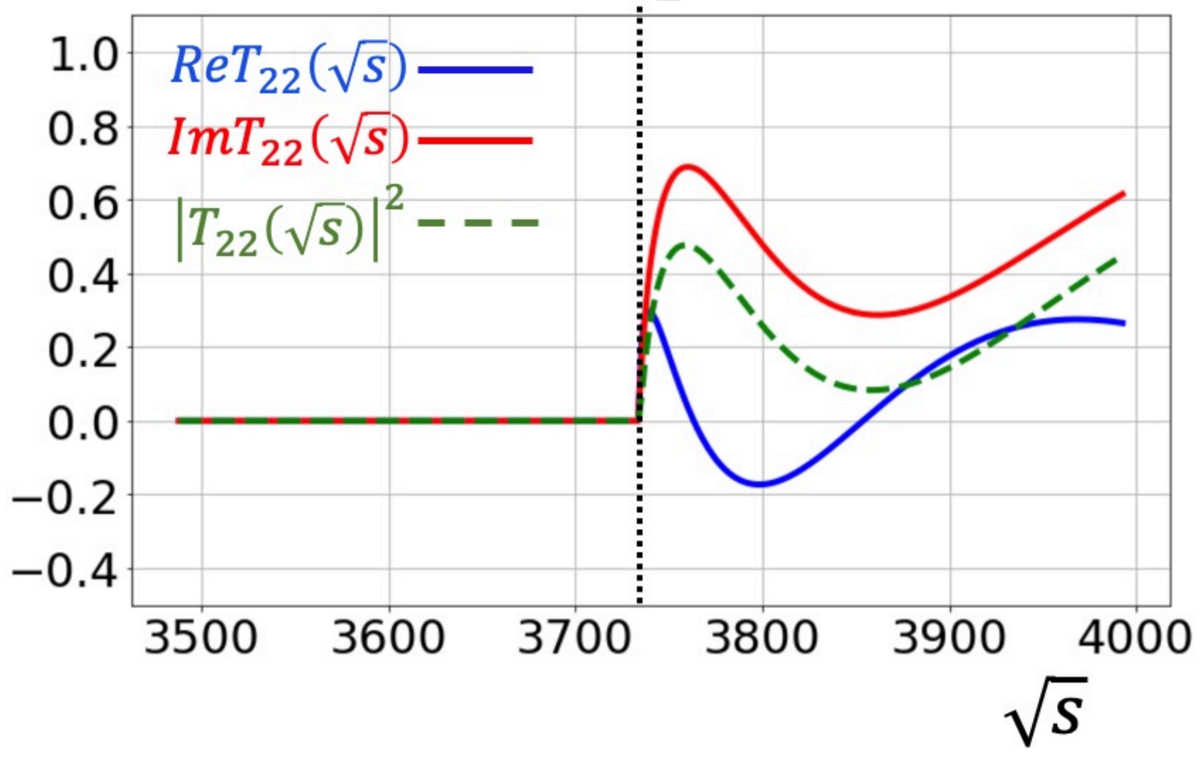


Amplitude with one pole in [bt] sheet.

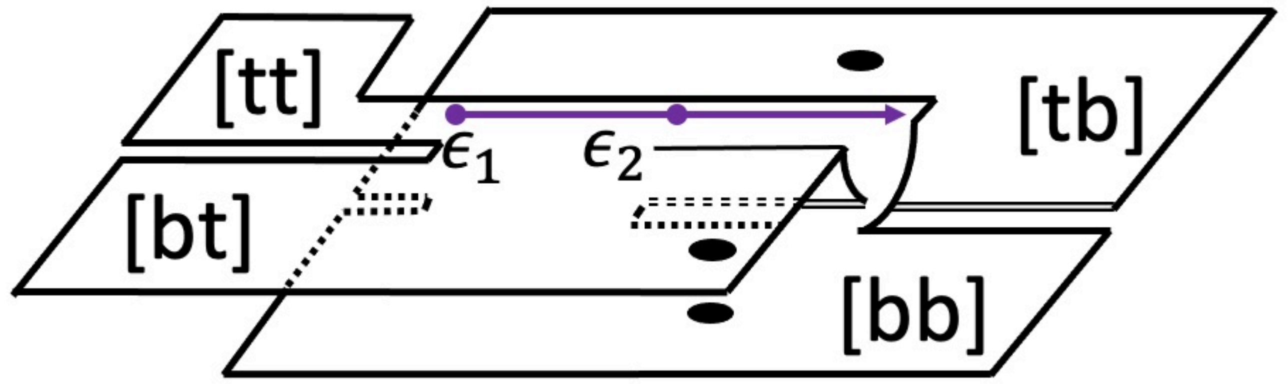
(a)



(b)



(c)



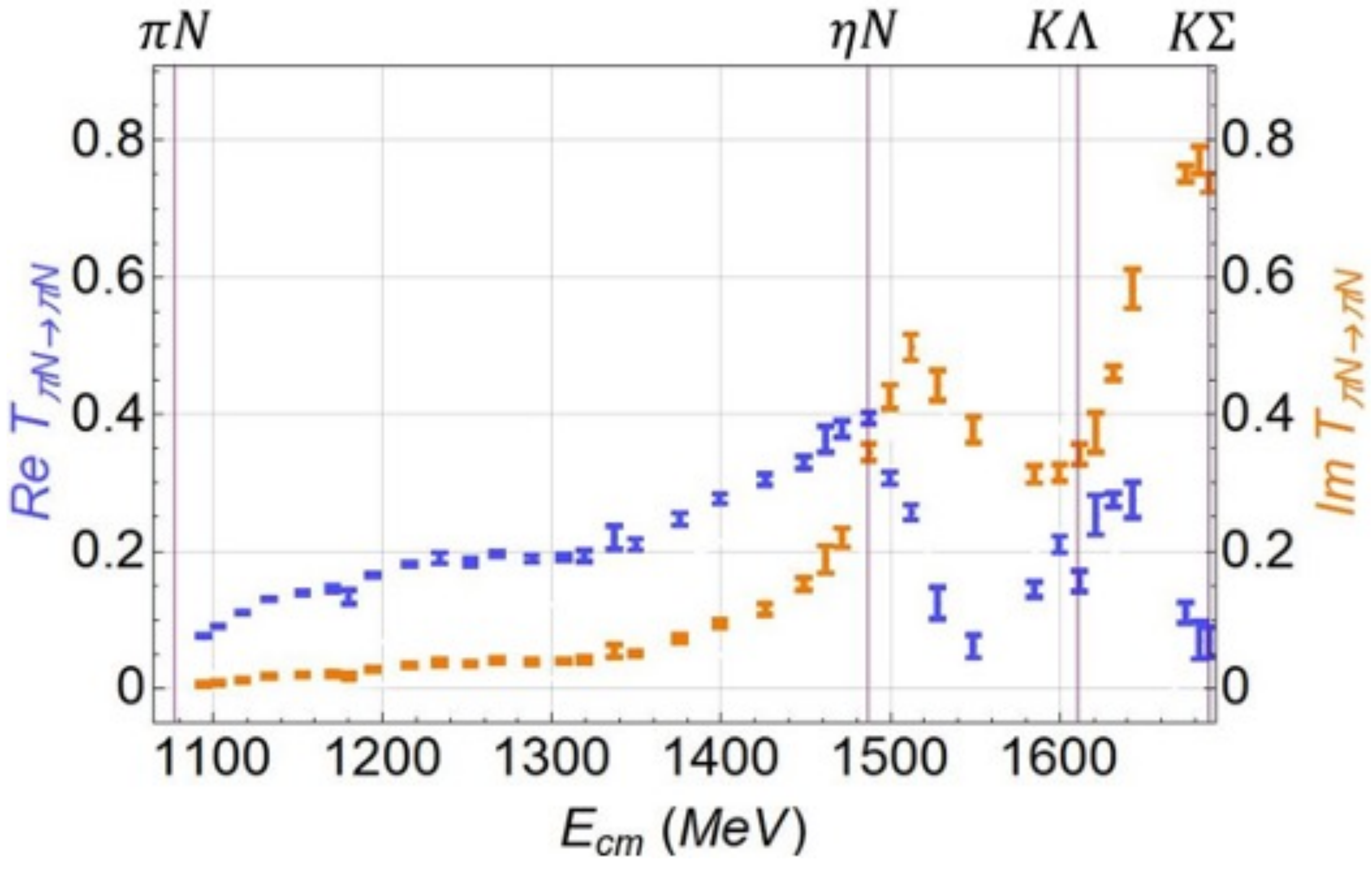
Amplitude with one pole in each unphysical sheet. The two extra poles have the same energy values.

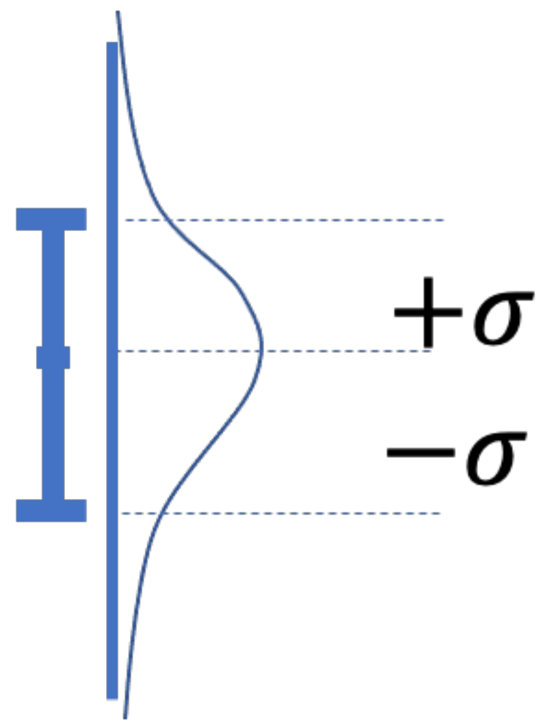
Identical lower channel amplitude.

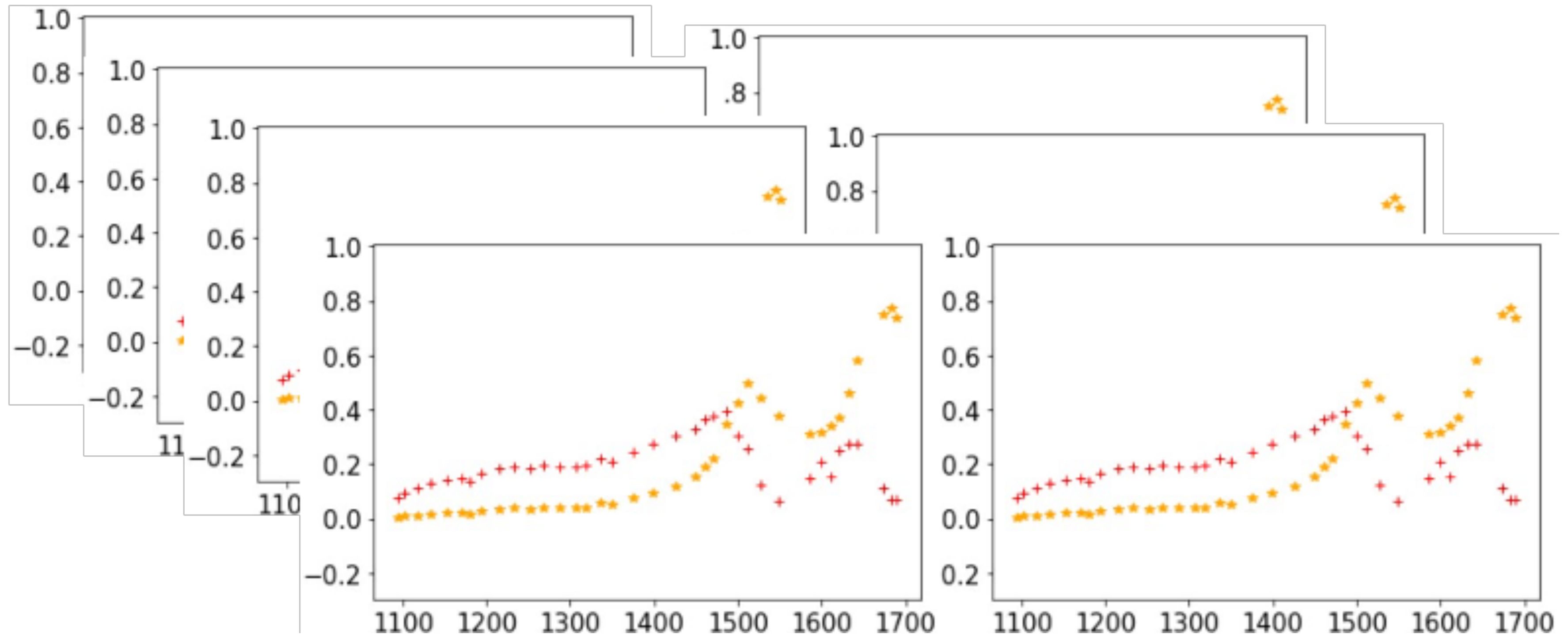
Higher channel amplitude can be distinguished.

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

Inference stage: application





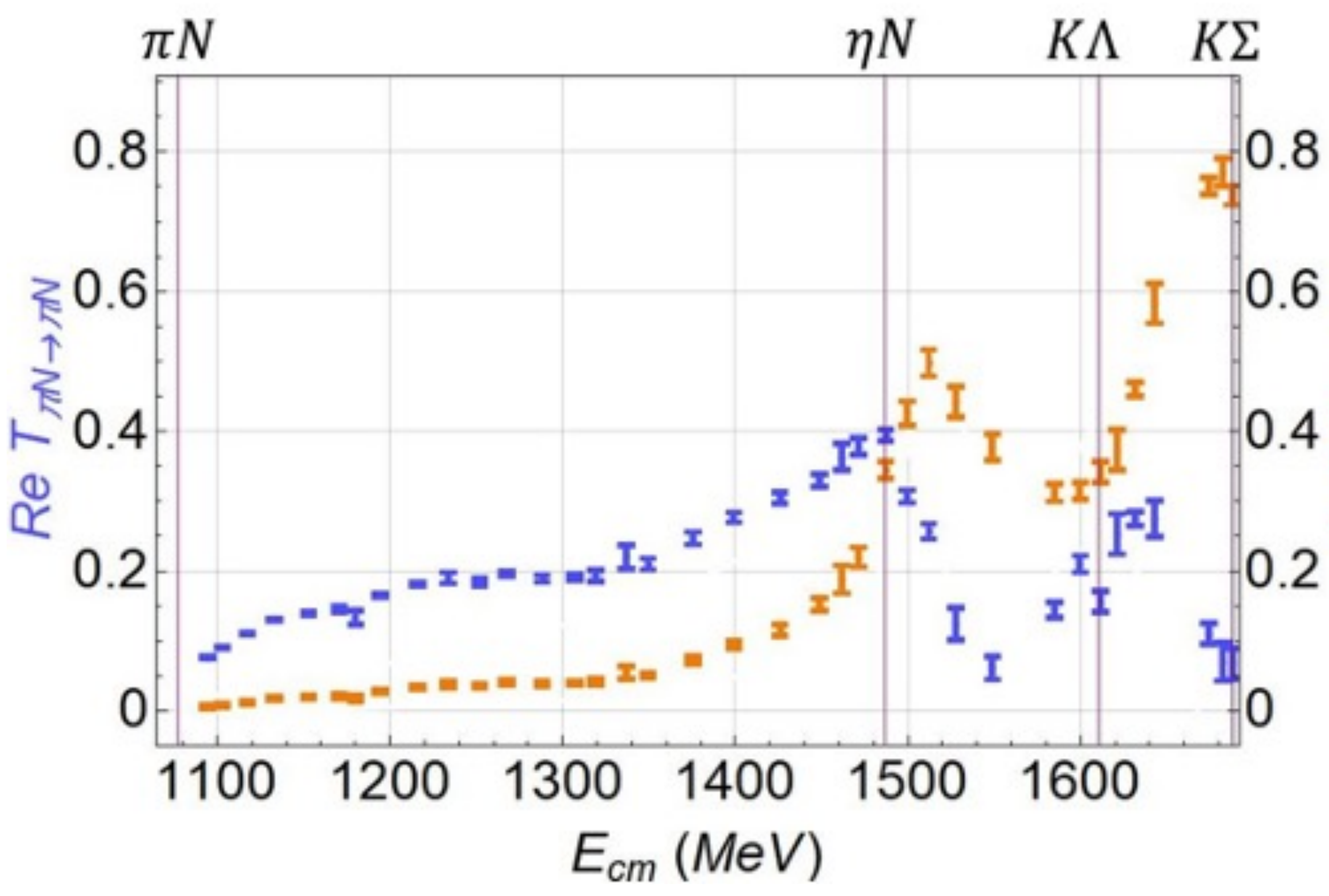


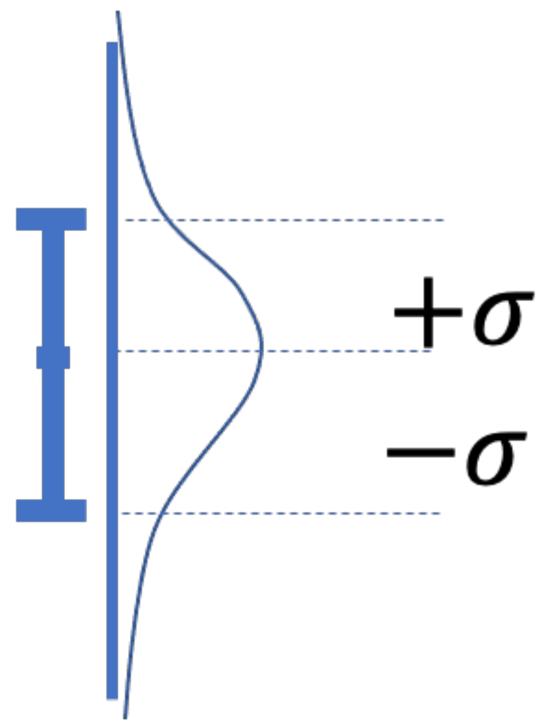
- Draw points from each error bar using a Gaussian distribution.
- Construct inference amplitudes from the experimental data using the drawn points.
- Feed the inference amplitudes to the trained DNN.

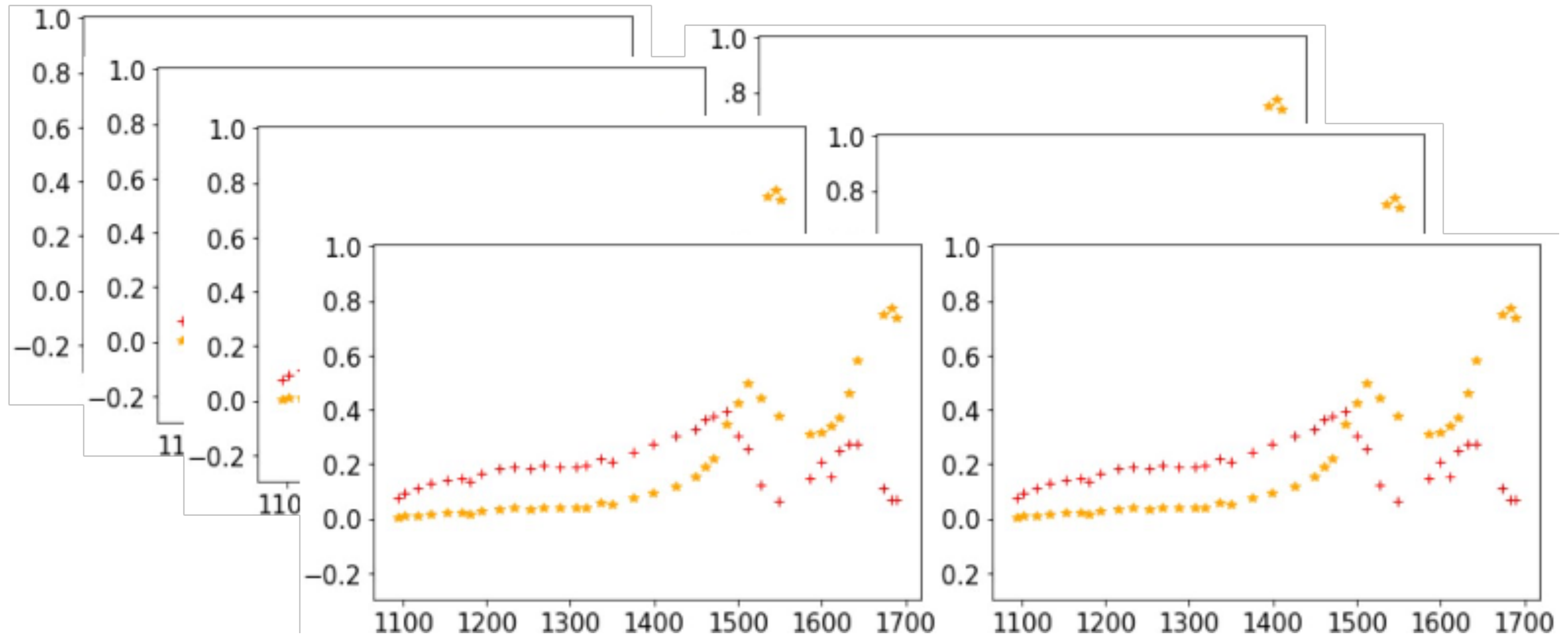
Interference on 10^6 amplitudes

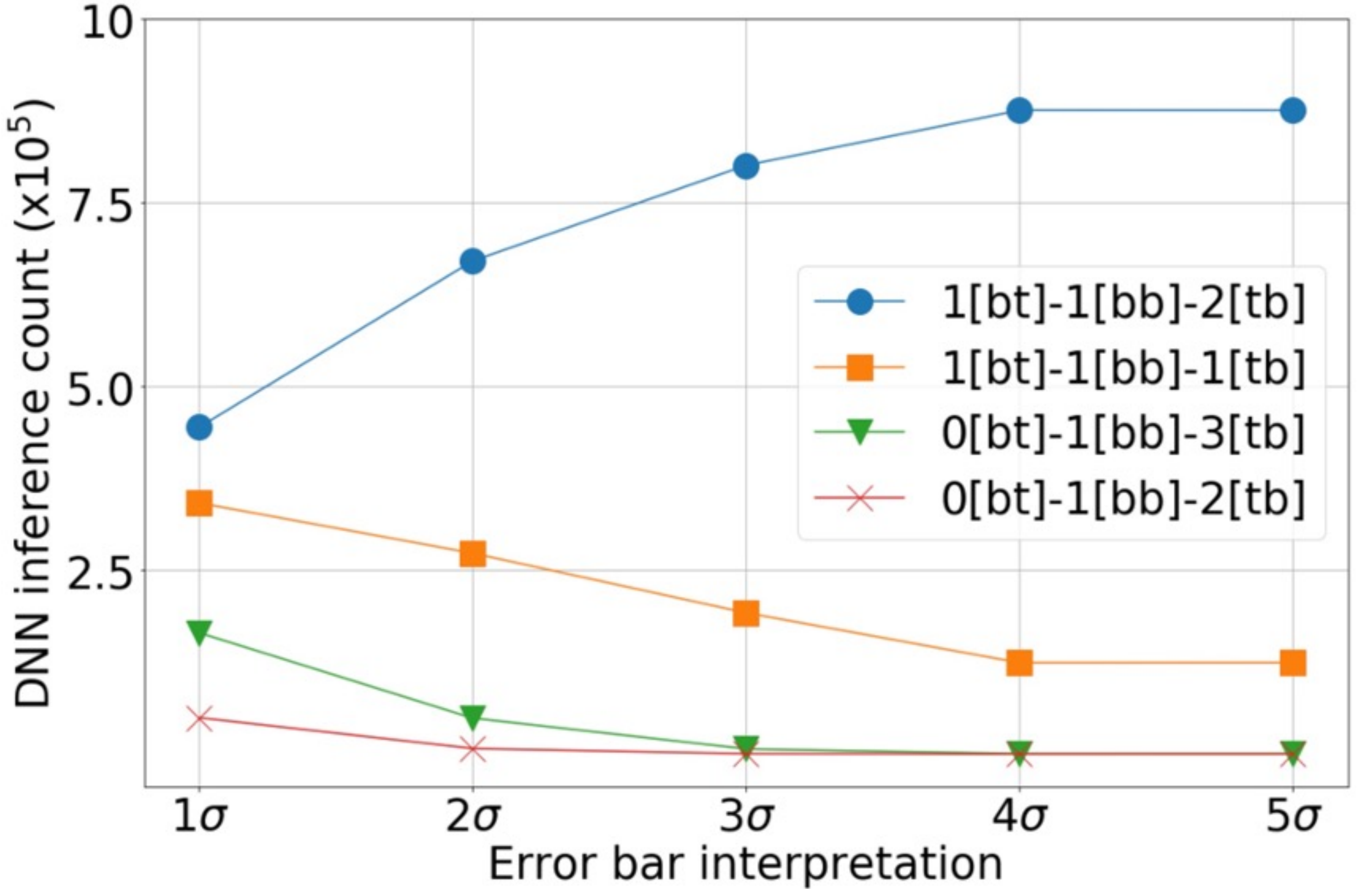
- 44.6% 1[bt]-1[bb]-2[tb]
- 34.1% 1[bt]-1[bb]-1[tb]
- 16.4% 0[bt]-1[bb]-3[tb]
- 04.9% 0[bt]-1[bb]-2[tb]

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]

Summary

- We can teach DNN:
  - to distinguish threshold enhancement in single-channel scattering
  - to determine the pole configuration of a coupled-channel amplitude

- Deep learning approach can be used as a model-selection framework.

- The results of deep learning method can be used to design an amplitude ansatz appropriate for a given experimental data.

Acknowledgement

We would like to acknowledge The Mathematical Analysis Unit of the Center for Infectious Disease Education and Research (CiDER), Osaka University for the conference and travel support.