Excited QCD 2022

MACHINE LEARNING FOR HADRON SPECTROSCOPY

Mount Etna

César Fernández Ramírez

JPAC

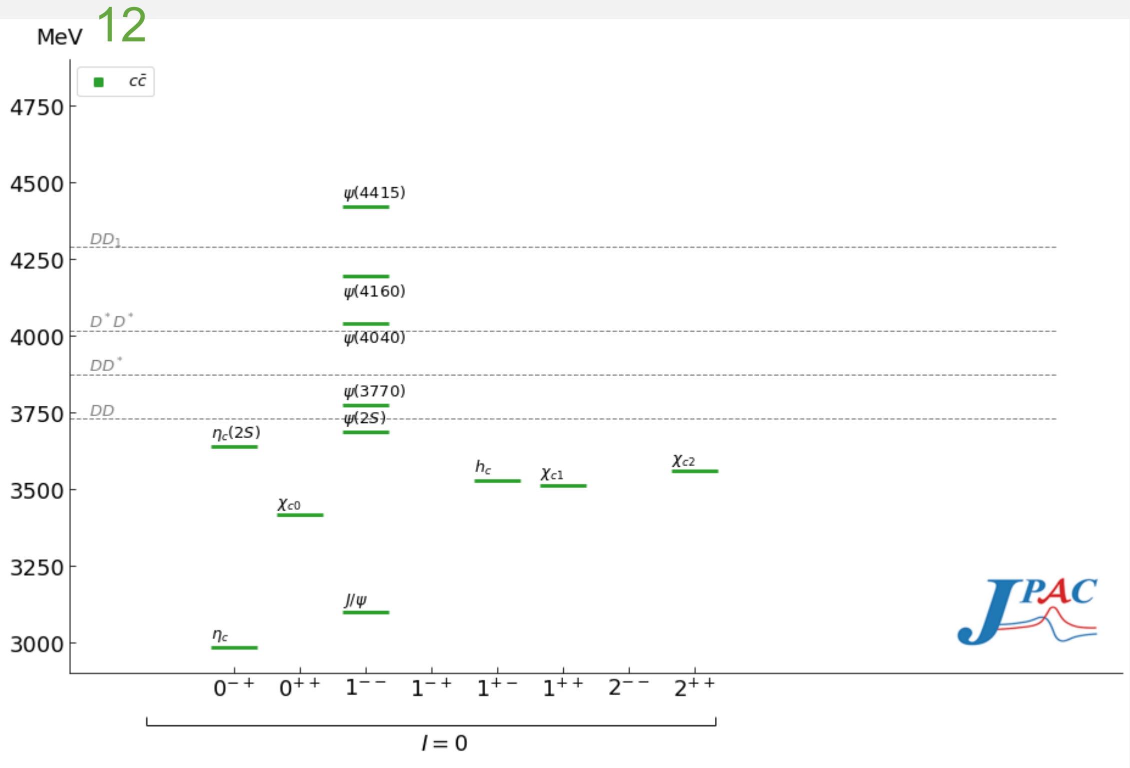
Instituto de Ciencias Nucleares UNAM

UNED

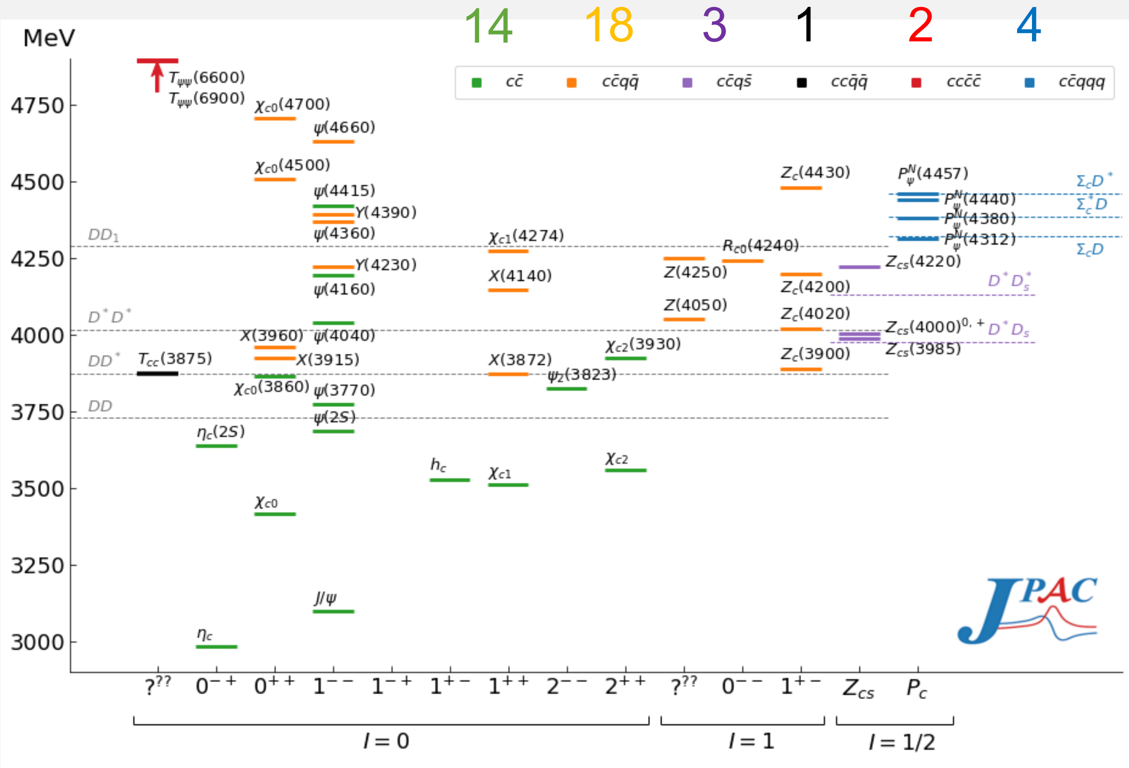
#SOMOS2030

Charmonia(-like)

2003

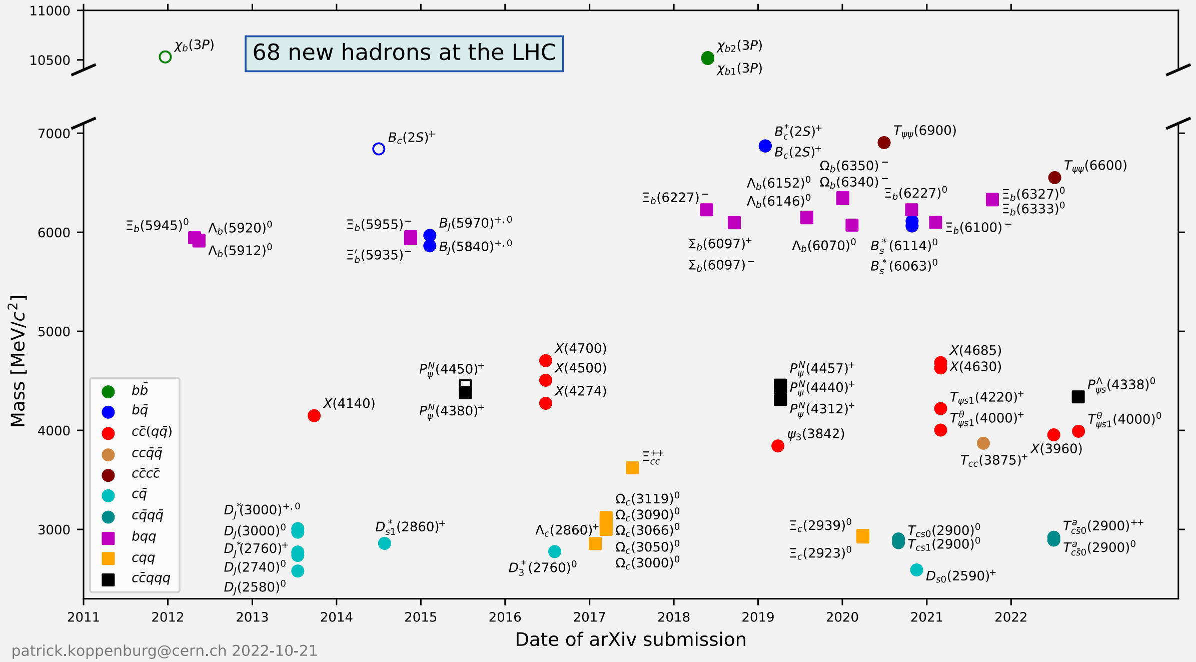


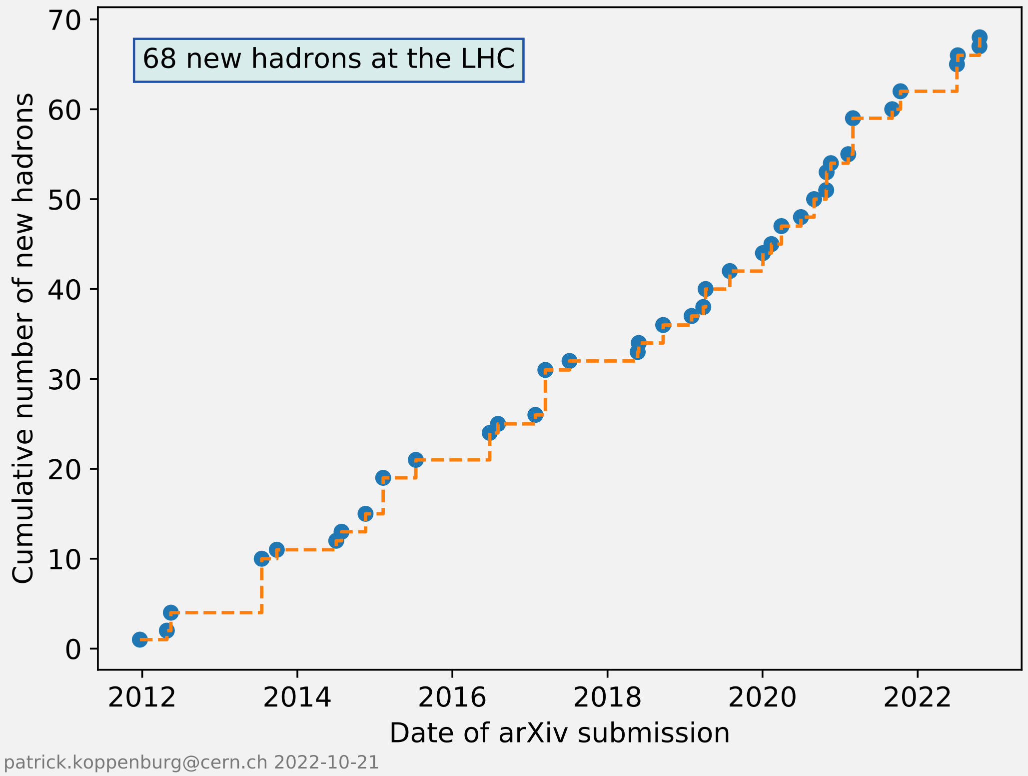
2022



And $P_{\psi s}^{\Lambda}(4338), T_{c\bar{s}0}^{a}(2900)^{0,++}$

Just LHC


68 new hadrons at the LHC
Mass [MeV/c^2]
Date of arXiv submission
11000
10500
7000
6000
5000
4000
3000
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
$b\bar{b}$
$b\bar{q}$
$c\bar{c}(q\bar{q})$
$cc\bar{q}\bar{q}$
$cc\bar{c}\bar{c}$
$c\bar{q}$
$c\bar{q}\bar{q}$
bqq
cqq
$c\bar{c}qqq$
$\chi_b(3P)$
$\chi_{b2}(3P)$
$\chi_{b1}(3P)$
$B_c(2S)^+$
$B_c^*(2S)^+$
$T_{\psi\psi}(6900)$
$T_{\psi\psi}(6600)$
$X(4140)$
$P_\psi^N(4450)^+$
$P_\psi^N(4380)^+$
$X(4700)$
$X(4500)$
$X(4274)$
$P_\psi^N(4457)^+$
$P_\psi^N(4440)^+$
$P_\psi^N(4312)^+$
$\psi_3(3842)$
$X(4685)$
$X(4630)$
$T_{cc}(3875)^+$
$X(3960)$
Ξ_{cc}^{++}
patrick.koppenburg@cern.ch 2022-10-21


68 new hadrons at the LHC
Cumulative number of new hadrons
Date of arXiv submission
0
10
20
30
40
50
60
70
2012
2014
2016
2018
2020
2022
patrick.koppenburg@cern.ch 2022-10-21

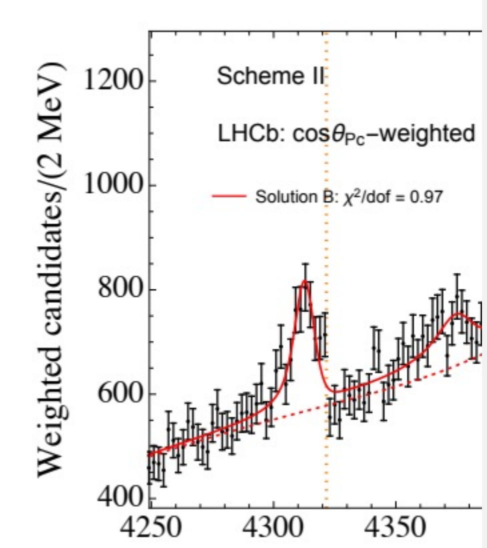

panda
NICA
NUCLOTRON BASED ION COLLIDER FACILITY
ВЭПП 2000
BESIII
mami
MAINZER MIKROTRON
J-PARC
Belle II
KLOE
Jefferson Lab
clas
GlueX
CERN
COMPASS
LHCb
ATLAS
EXPERIMENT
CMS
ALICE

Standard approach to lineshape analysis

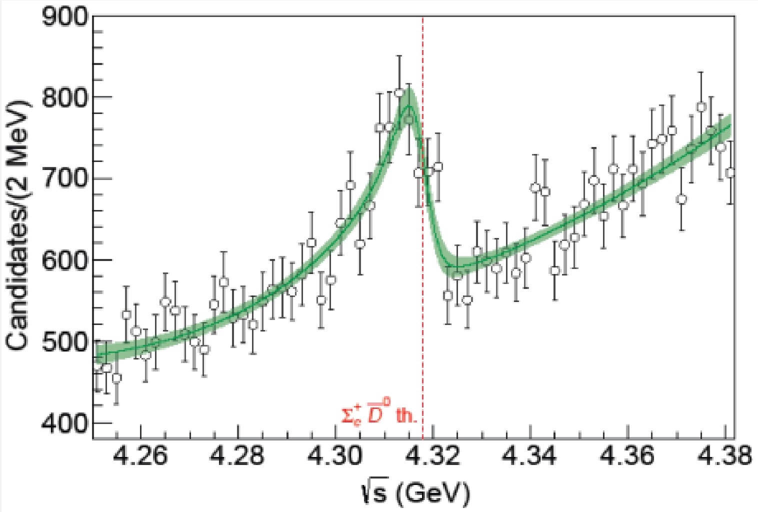
- Build a model for the amplitude and assume it is true
- Fit data using χ^2
- Extract model parameters and get pole positions and compute uncertainties
- Asses the probability that those data were generated by your model
- If everything is fine, you can claim that the interpretation embedded in the model is a possible explanation of the data
- You can do this with different models with different underlying dynamics
- Compare models? Compare dynamics?

Example $P_c(4312)$

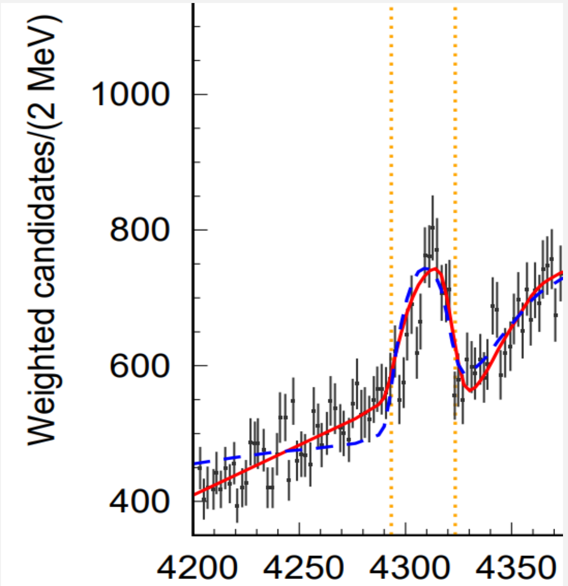
Data from



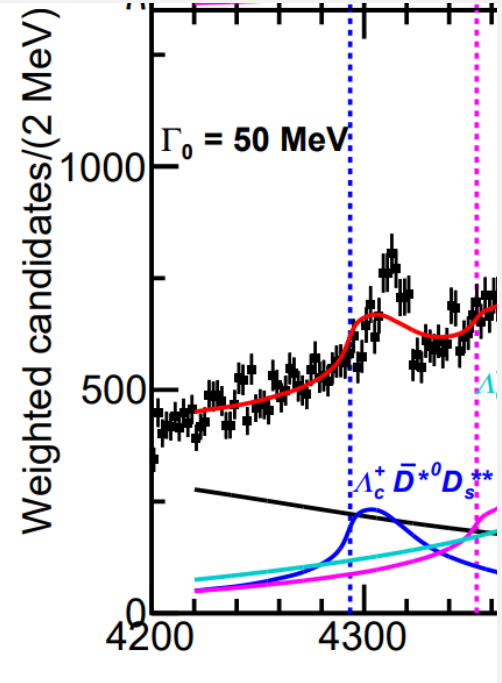
Molecule
Du et al., 2102.07159



Virtual
CFR et al. (JPAC), 1904.10021

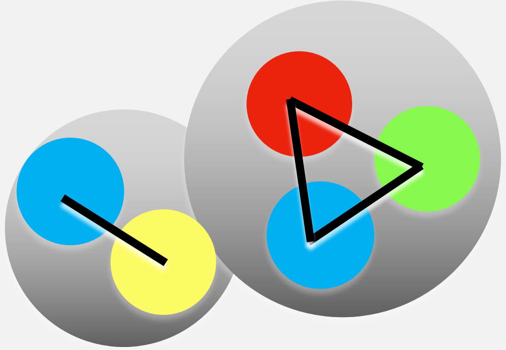


Double-triangle (w. complex coupl. in the Lagrangian)
Nakamura, 2103.06817

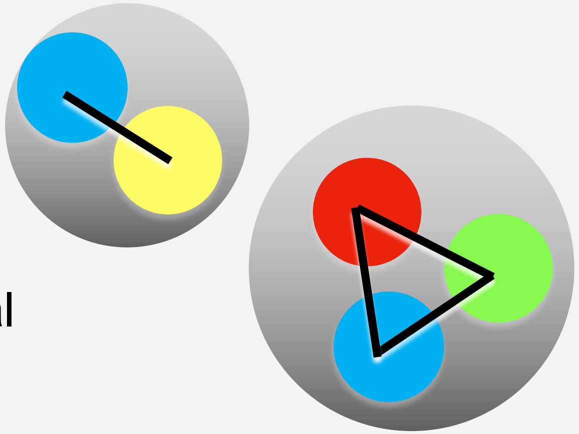


Single triangle (ruled out)
LHCb, 1904.03947

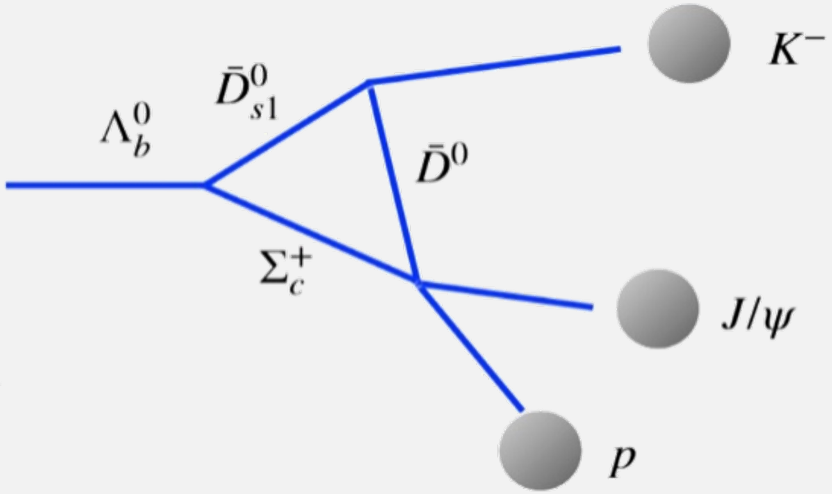
Internal dynamics



Molecule



Virtual

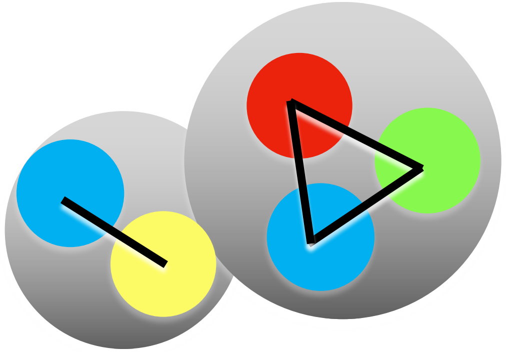


Single triangle
(double triangle=same idea but more complicated)

Threshold generated (residual interaction)

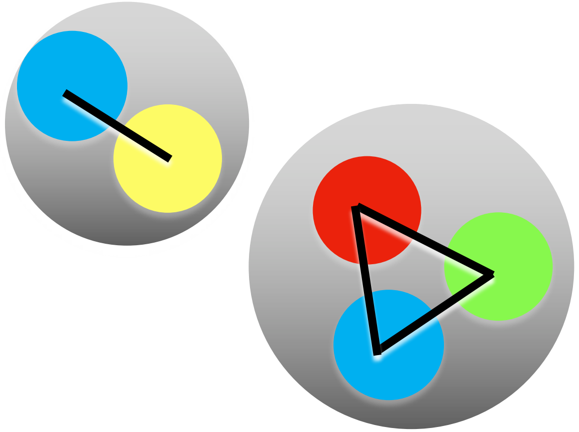
Molecular state

- Generated through the opening of a new channel
- Residual part of the strong interaction bounds the system
- Finite radius



Virtual state

- Generated through the opening of a new channel
- Residual part of the strong interaction generates the signal
- Infinite radius



Can machine learning help us?

First explorations of deep neural networks as classifiers for hadron spectroscopy:
Sombillo et al., 2003.10770, 2104.141782, 2105.04898

Can we train a neural network to analyze a lineshape and get as a result what is the probability of each posible dynamical explanation?

If posible, what other information can we gain by using machine learning techniques?

Still far away from answering those question but we are advancing

Benchmark case: $P_c(4312)$ lineshape

Outlook

Amplitude analysis of $P_c(4312)$ (benchmark case):
CFR et al. (JPAC), 1904.10021

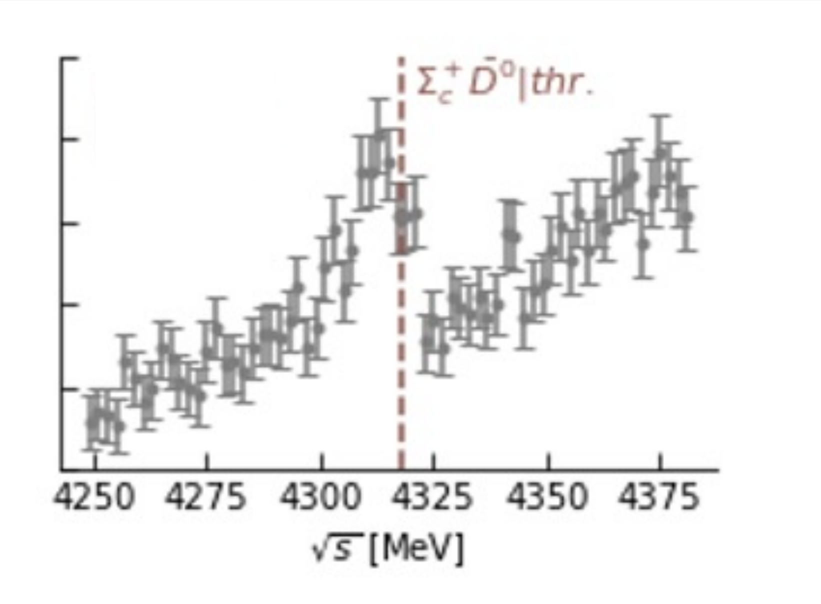
Deep neural networks: classifiers

Application to our benchmark case
Ng et al. (JPAC), 2110.13742

Additional information: SHAP values

Takeaways

$J/\psi p$ projection data



Data from LHCb, 1904.03947

We focus on $\Sigma_c^+ \bar{D}^0$ threshold region where the $P_c(4312)$ signal appears

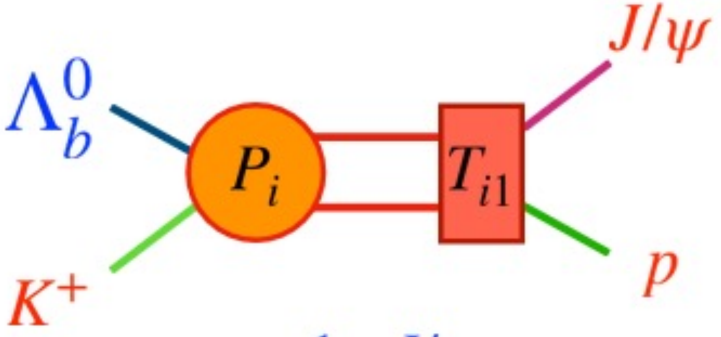
We asume only one partial wave contributes to the $P_c(4312)$

The threshold is responsible for the dynamics (no compact state interpretation)

Other singularities are irrelevant

Near- $\Sigma_C^+ \overline{D}_0$ threshold theory (two channels)

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

$F(s) =$ 

$1 : J/\psi p$
$2 : \Sigma_c^+ \bar{D}^0$

$$F(s) = P_1(s) T_{11}(s) \qquad \left(T^{-1}\right)_{ij} = M_{ij} - i k_i \delta_{ij}$$

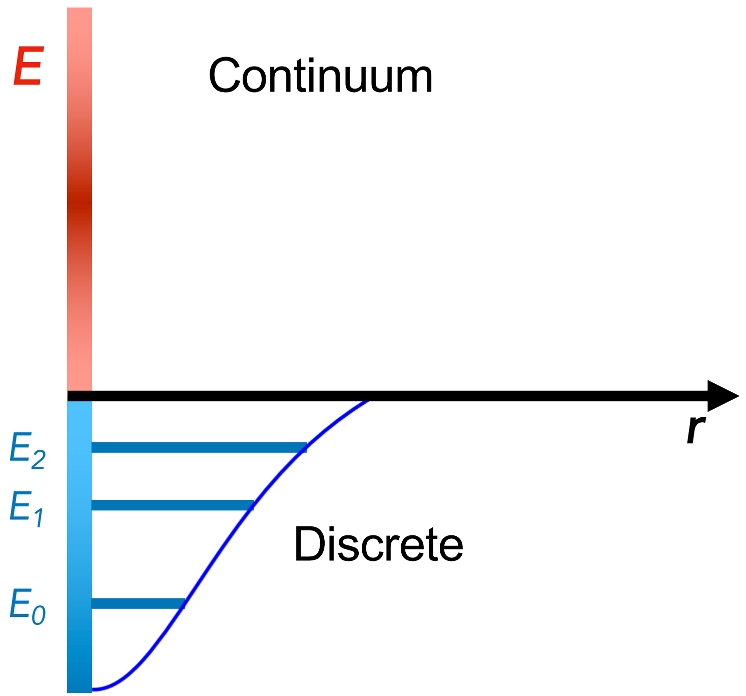
Scattering length ↓ Effective range ↓

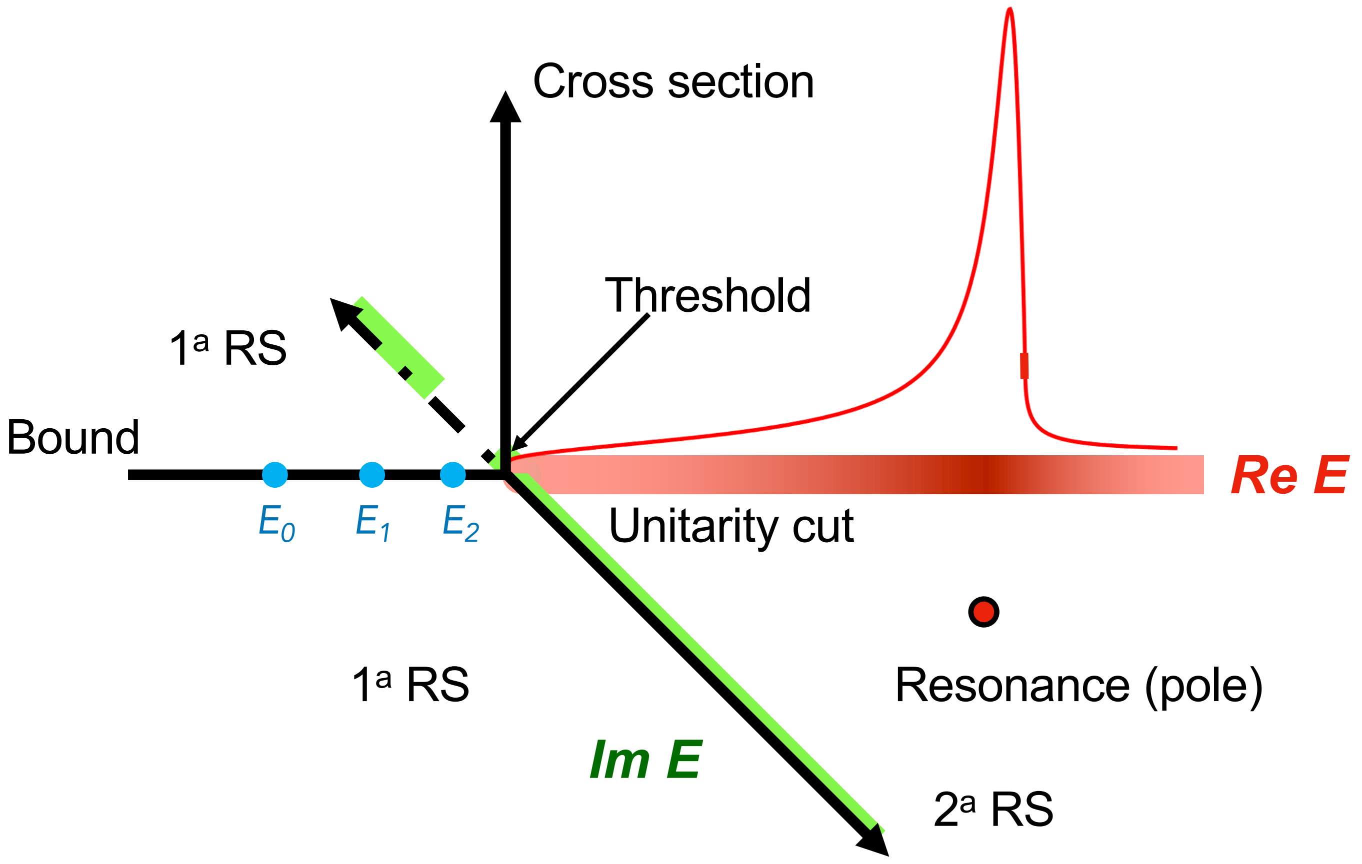
$$M_{ij}(s) = m_{ij} - c_{ij} s$$

Matrix elements M_{ij} are singularity free and can be Taylor expanded

Frazer, Hendry, PR134 (1964) B1307

Resonances (aka poles in unphysical Riemann sheets)

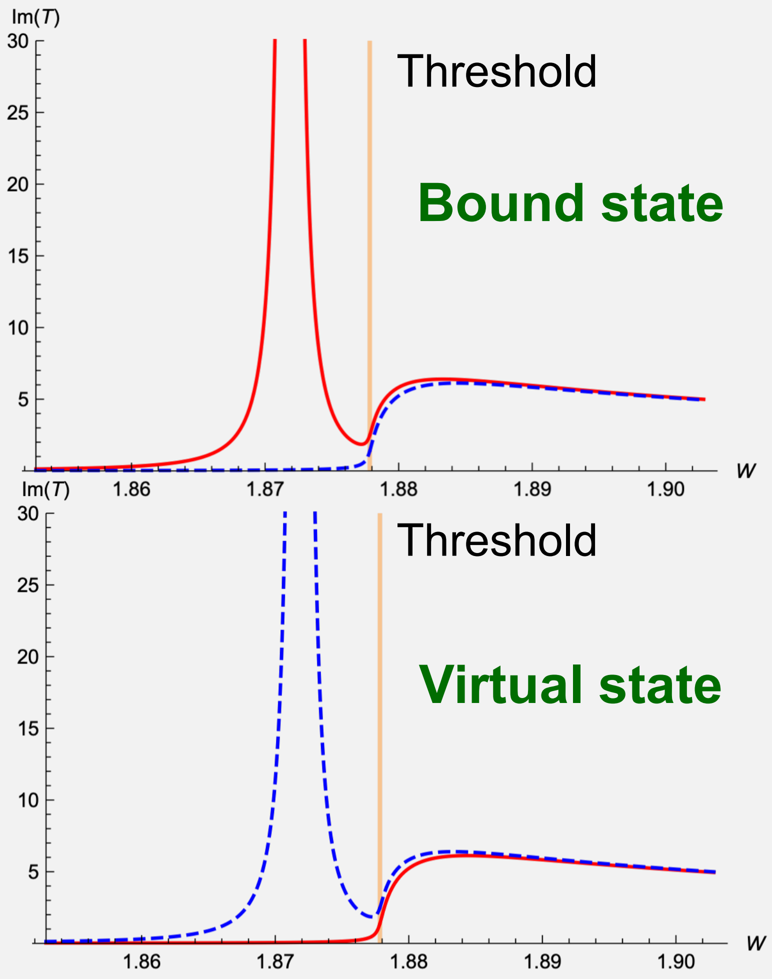


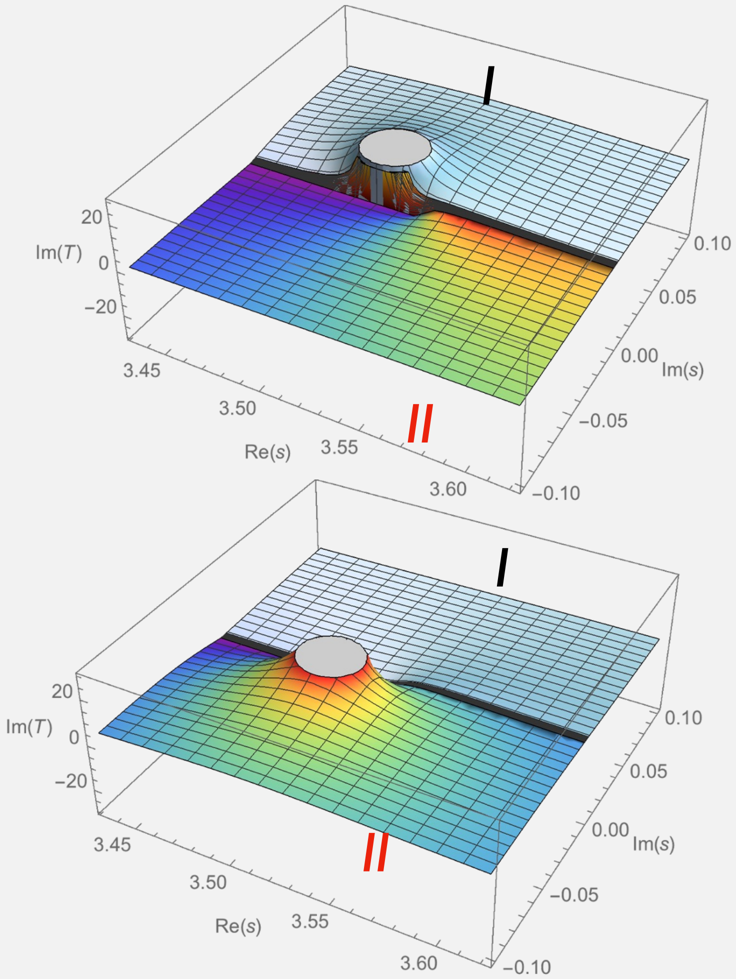


Bound vs. virtual: pole position

Amplitude close to the threshold $\rho \simeq k$

$$T = \frac{1}{-\frac{1}{a} + \frac{r_e}{2}k^2 - ik}$$

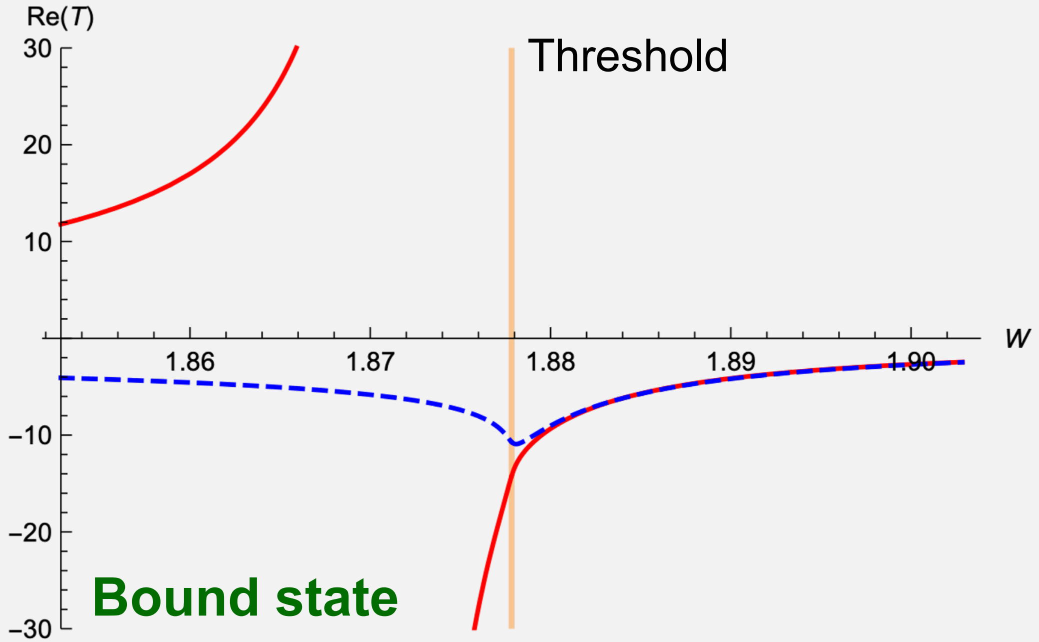


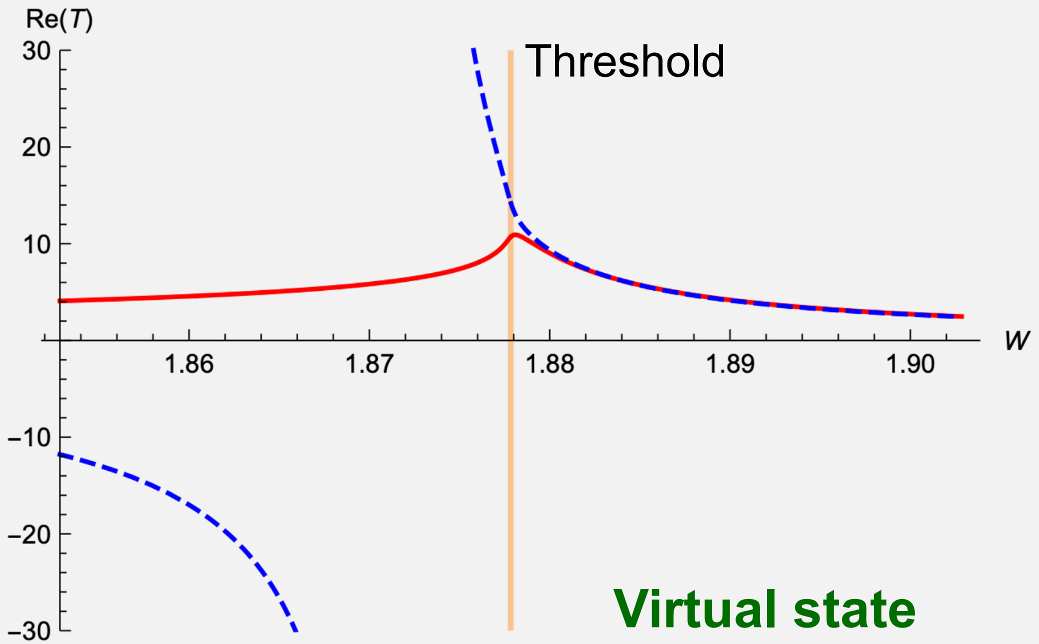


1st Riemann sheet $T(s+i\varepsilon)$

2nd Riemann sheet $T(s-i\varepsilon)$

Bound vs. virtual: scattering length





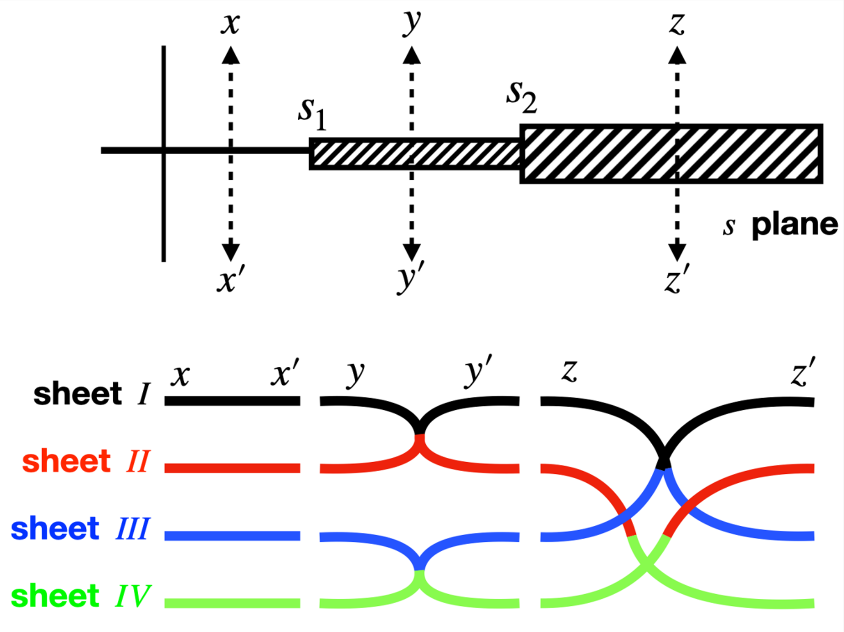
1st Riemann sheet *T(s+iε)*
2nd Riemann sheet *T(s-iε)*

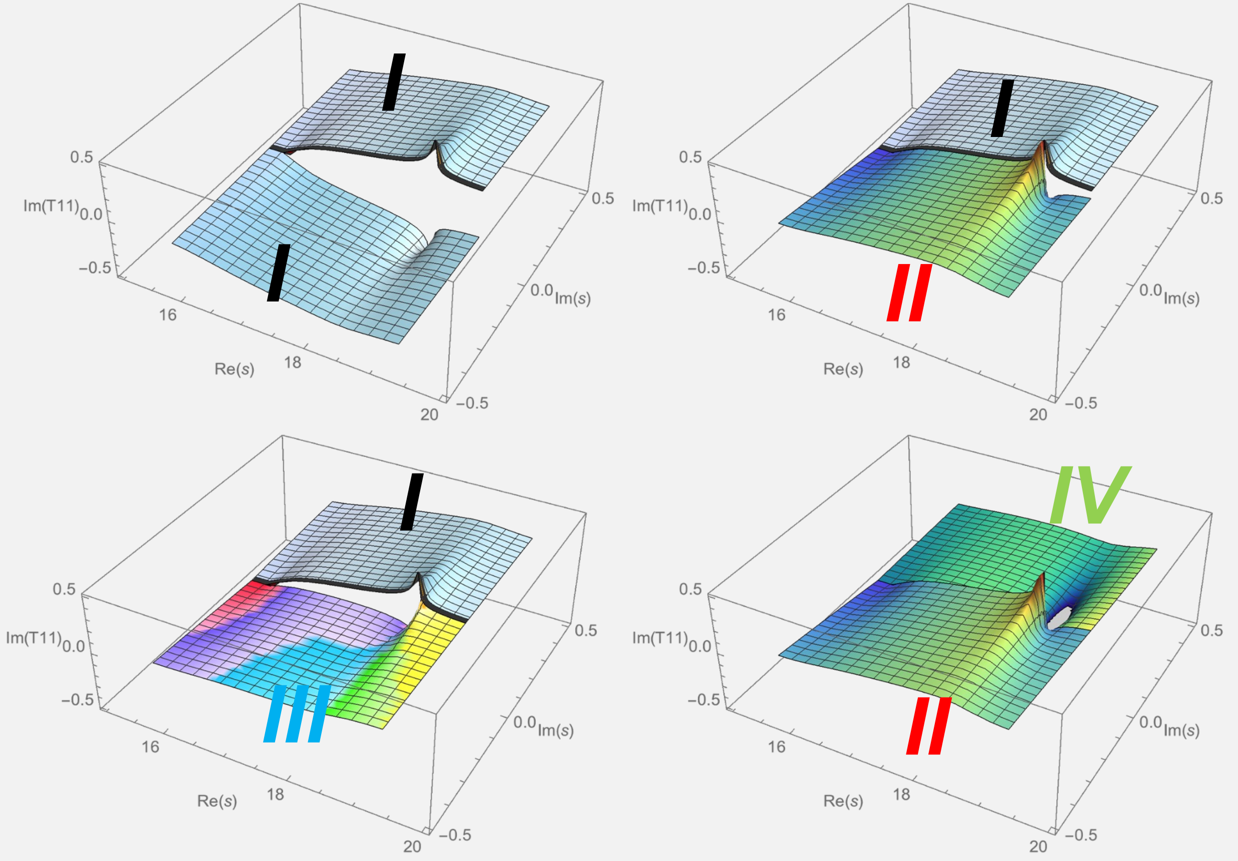
$$k \cot \delta(k) \simeq -\frac{1}{a} + \frac{r_e}{2} k^2$$

$$a_s(pn \rightarrow pn) \simeq 5.4\, fm\, (deuteron, {}^3S_1)$$
$$a_t(pn \rightarrow pn) \simeq -24\, fm\, ({}^1S_0)$$

Riemann sheets structure (two channels)

$$\begin{bmatrix} J/\psi p \to J/\psi p & J/\psi p \to \Sigma_c \bar{D}^0 \\ \Sigma_c \bar{D}^0 \to J/\psi p & \Sigma_c \bar{D}^0 \to \Sigma_c \bar{D}^0 \end{bmatrix}$$

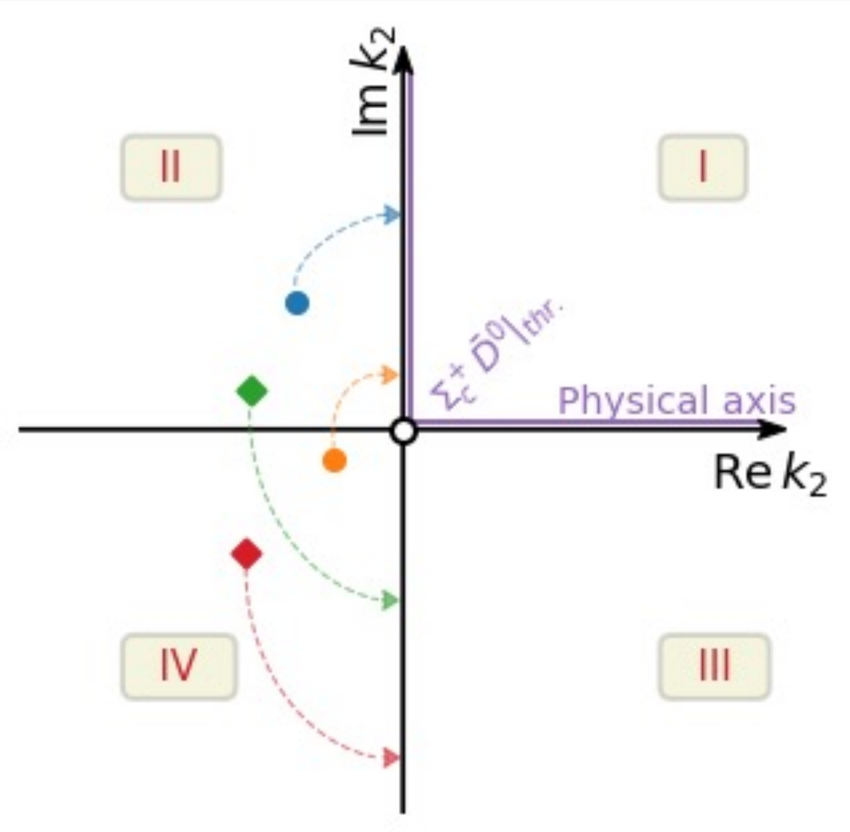




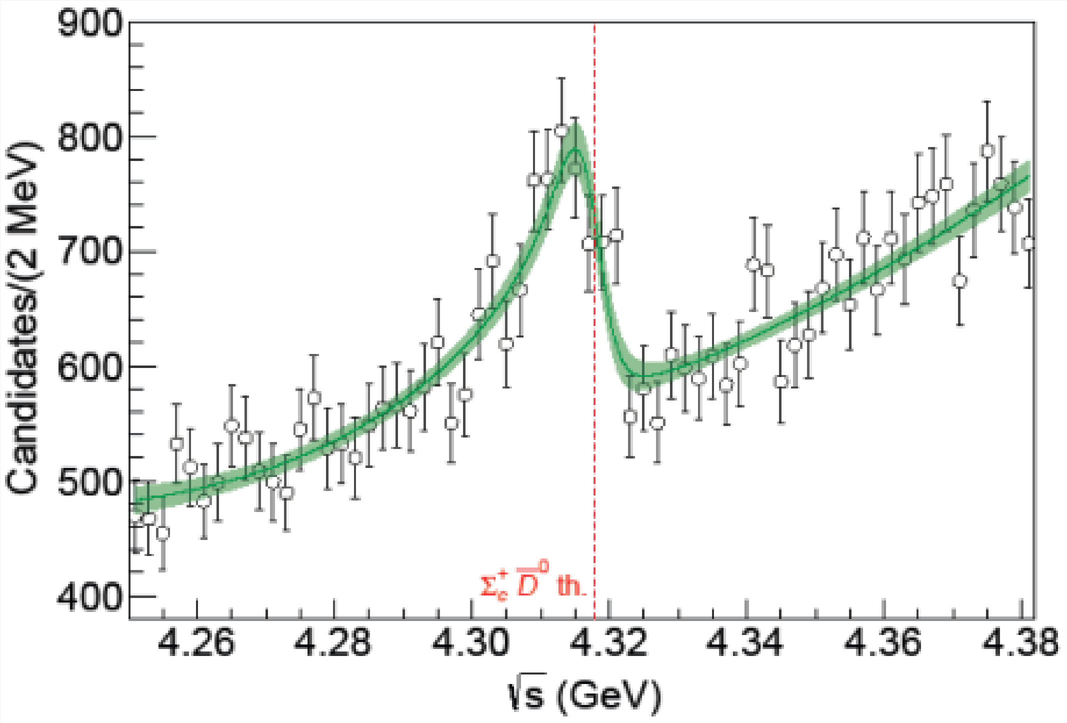
Virtual and bound states

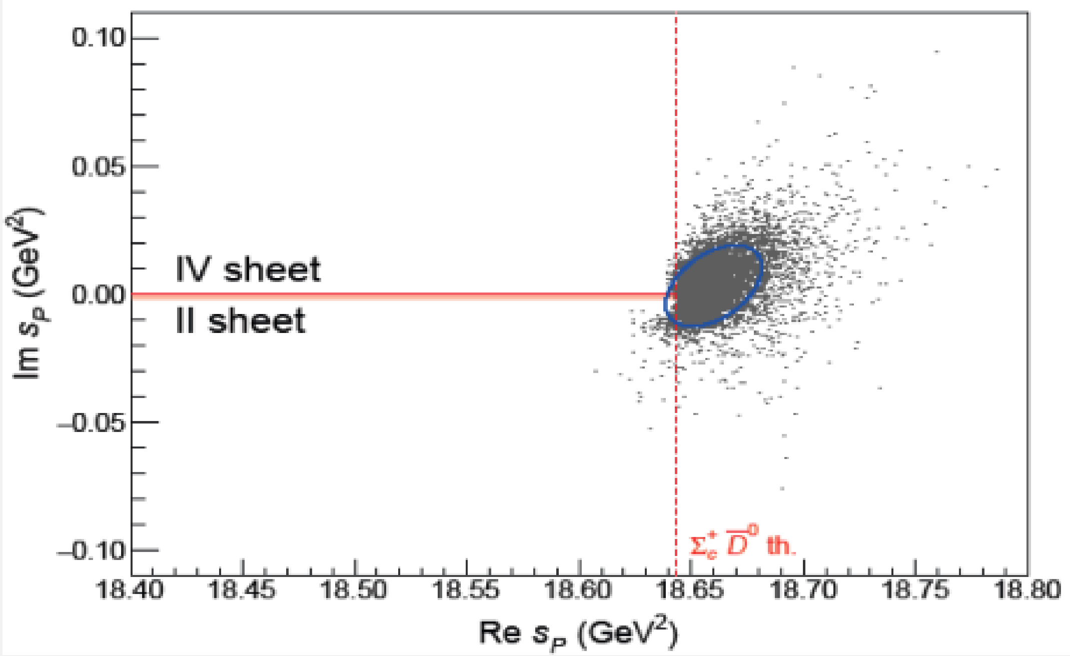
Under the the scattering length approximation the physical interpretation is given by the sign of the m_{22} parameter. Four options:

- Bound state on IV RS: b|4
- Virtual state on IV RS: v|4
- Bound state on II RS: b|2
- Virtual state on II RS: v|2



Amplitude analysis result

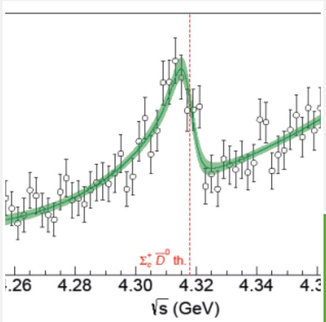




Interpretation obtained: Virtual state on IV RS (v|4)

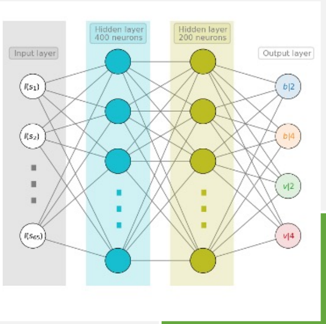
M= 4319.7±1.6 MeV	*Γ*= -0.8±2.4 MeV

Dictionary



AMPLITUDE ANALYSIS

a. Datapoints (lineshape)
b. Model with random parameters convoluted with experimental resolution
c. Experimental uncertainties
d. Physical interpretation
e. Objective (minimization) function

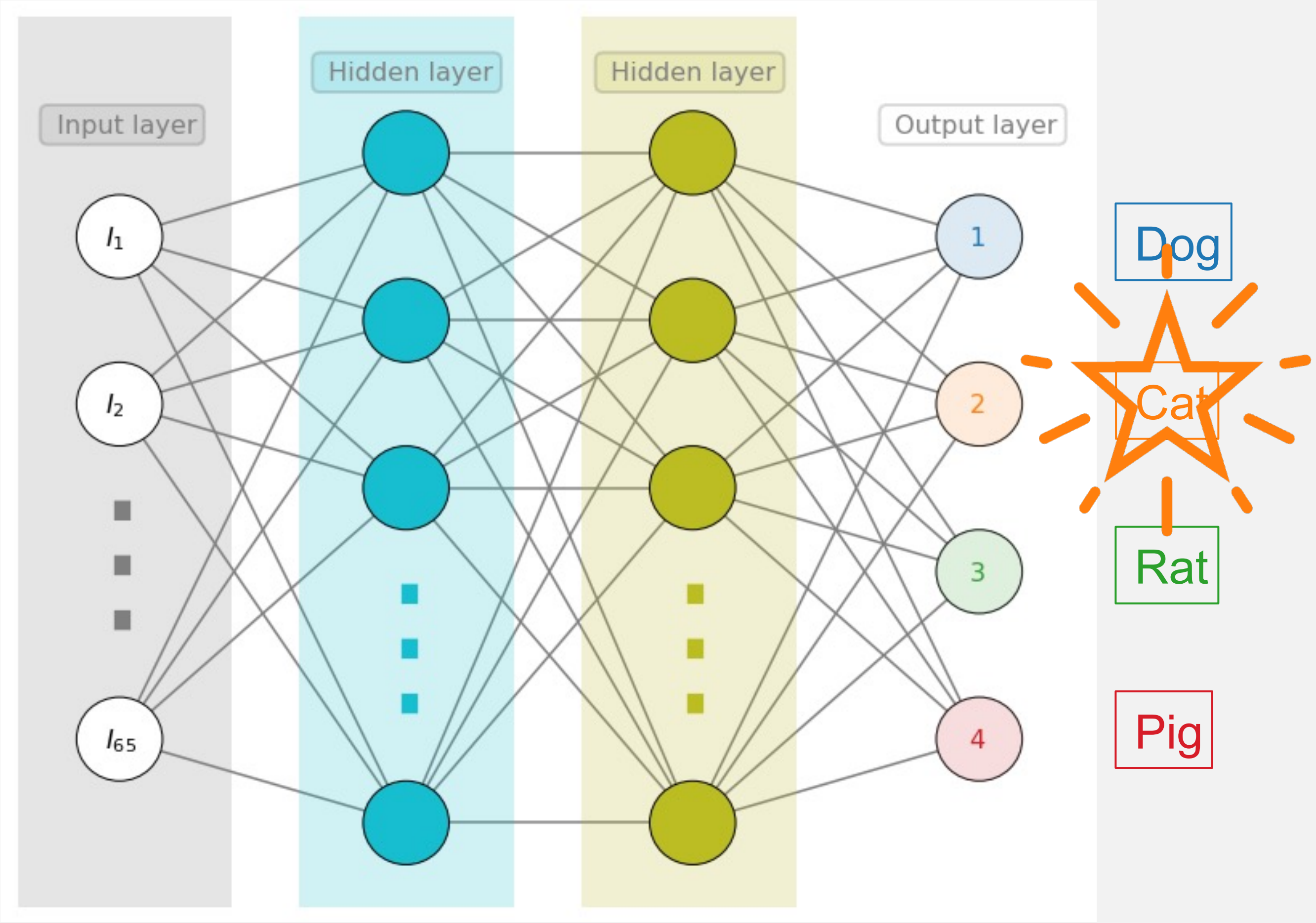


MACHINE LEARNING

a. Features
b. Training set
c. Noise
d. Classes
e. Minimize the DNN (Adam algorithm)

Neural networks as classifiers

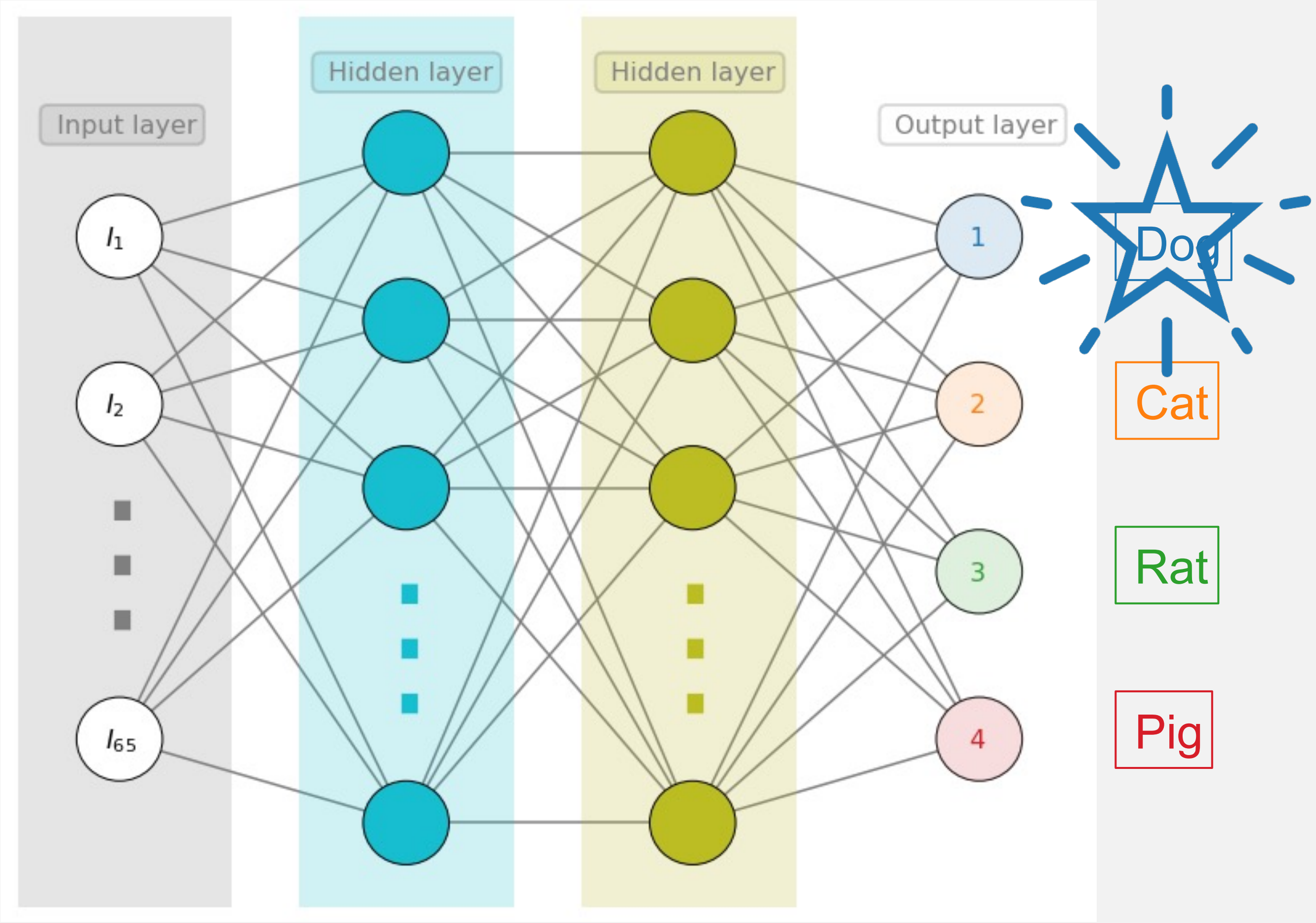
Training



20

Neural networks as classifiers

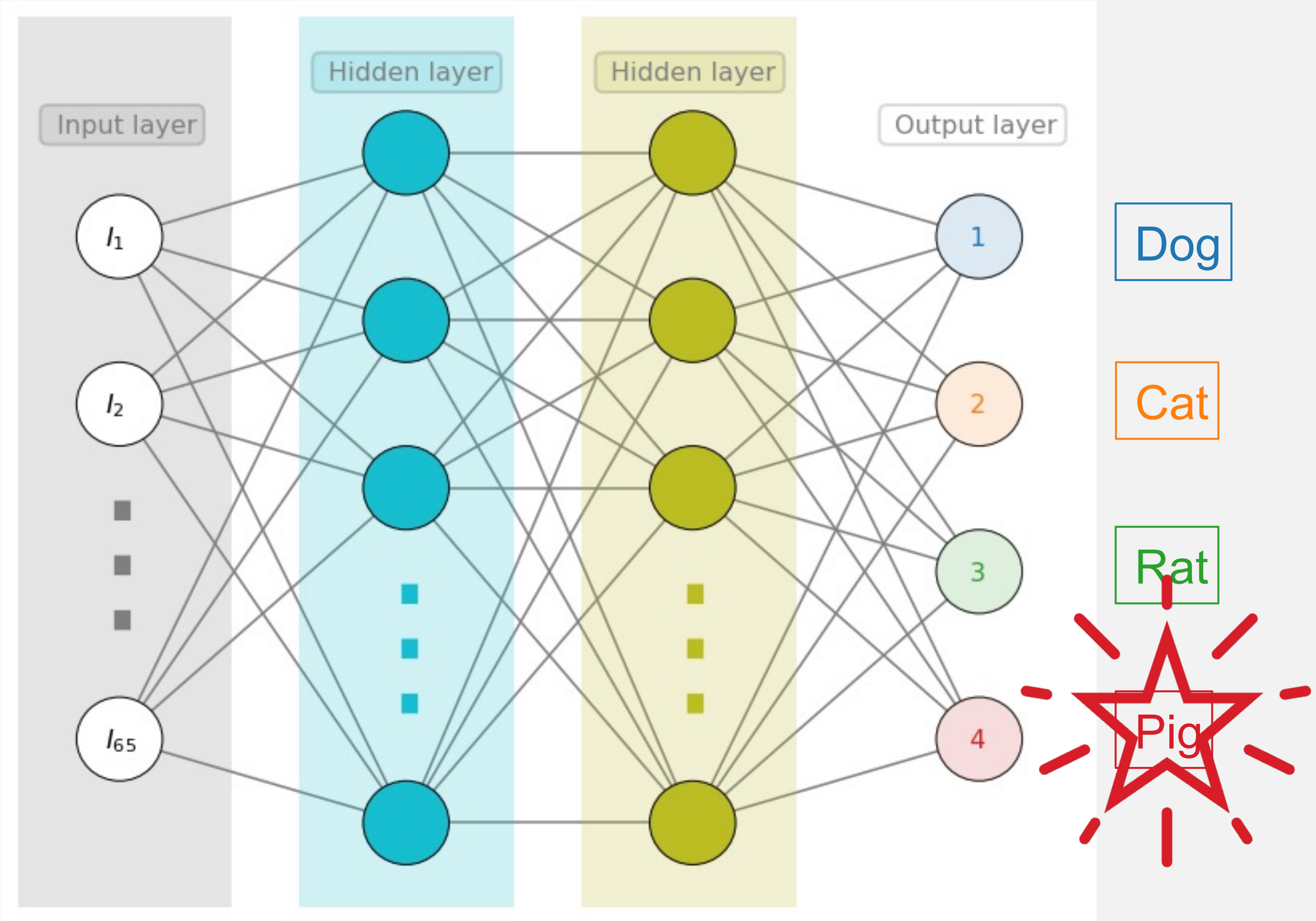
Training



21

Neural networks as classifiers

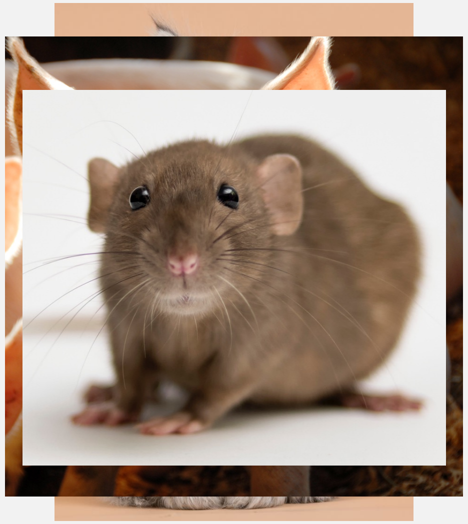
Training

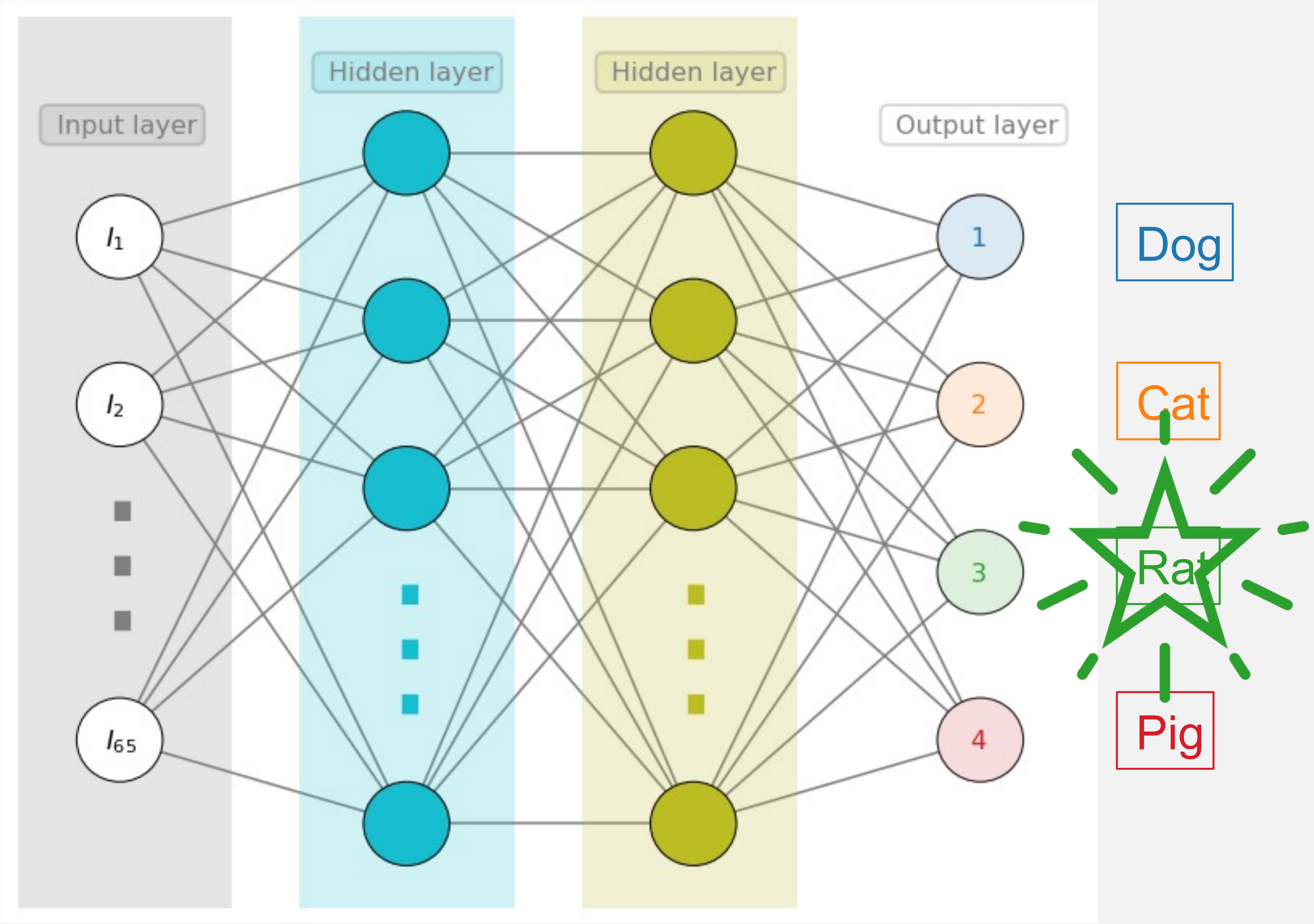


22

Neural networks as classifiers

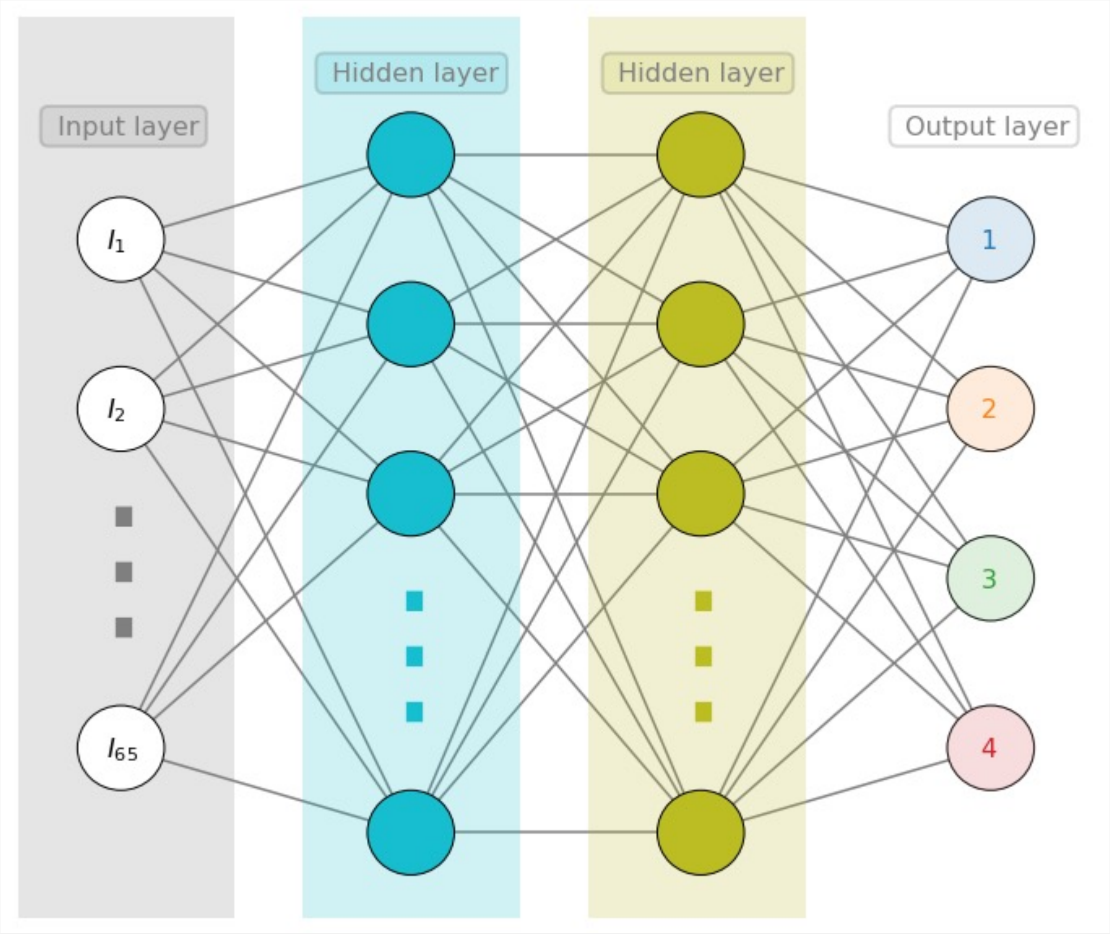
Training





Neural networks as classifiers

New picture



Dog	3%
Cat	96.8%
Rat	0.1%
Pig	0.1%

The deep neural network

The input layer contains 65 experimental datapoints

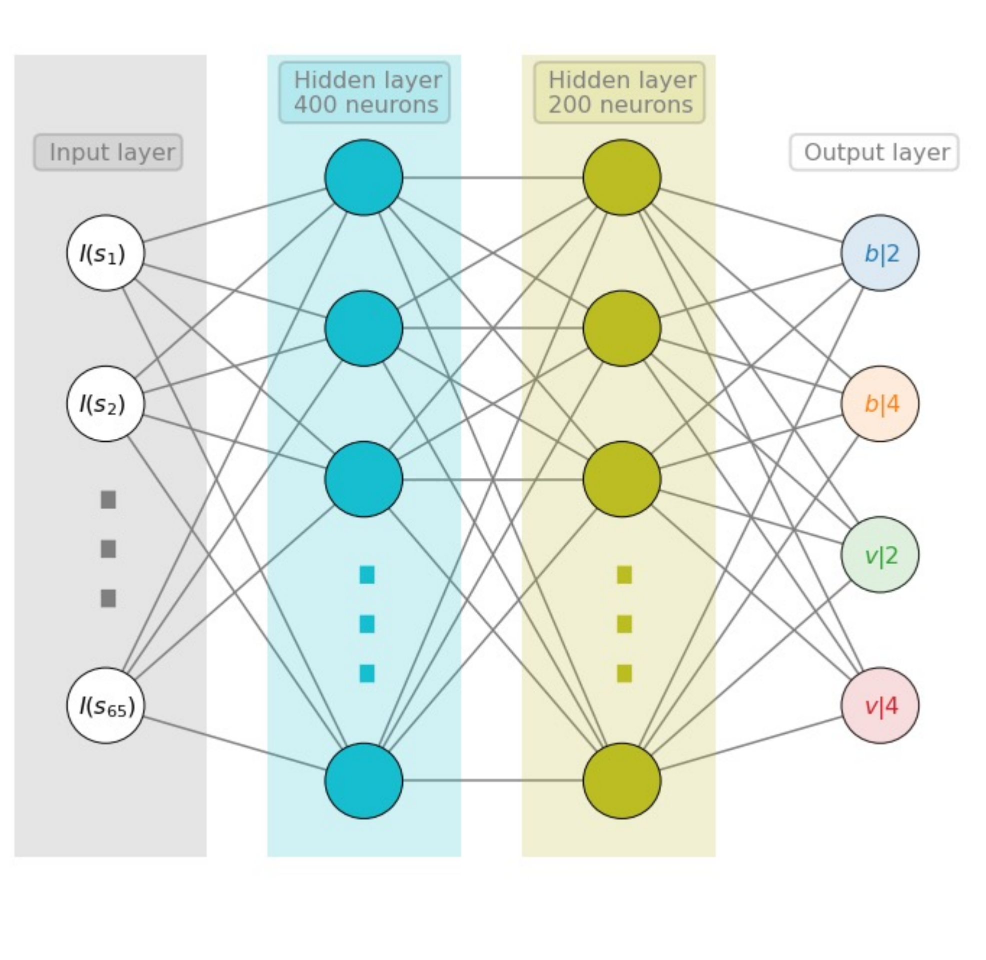
Two hidden layers with 400 and 200 neurons respectively

Output layer, four classes: v|4, b|4, b|2, v|2

Probabilities are obtained using the softmax function $p(x_i) = \frac{e^{x_i}}{\sum e^{x_j}}$

Minimizing multi-class cross-entropy loss

Implemented with PyTorch (alt-version using Keras)



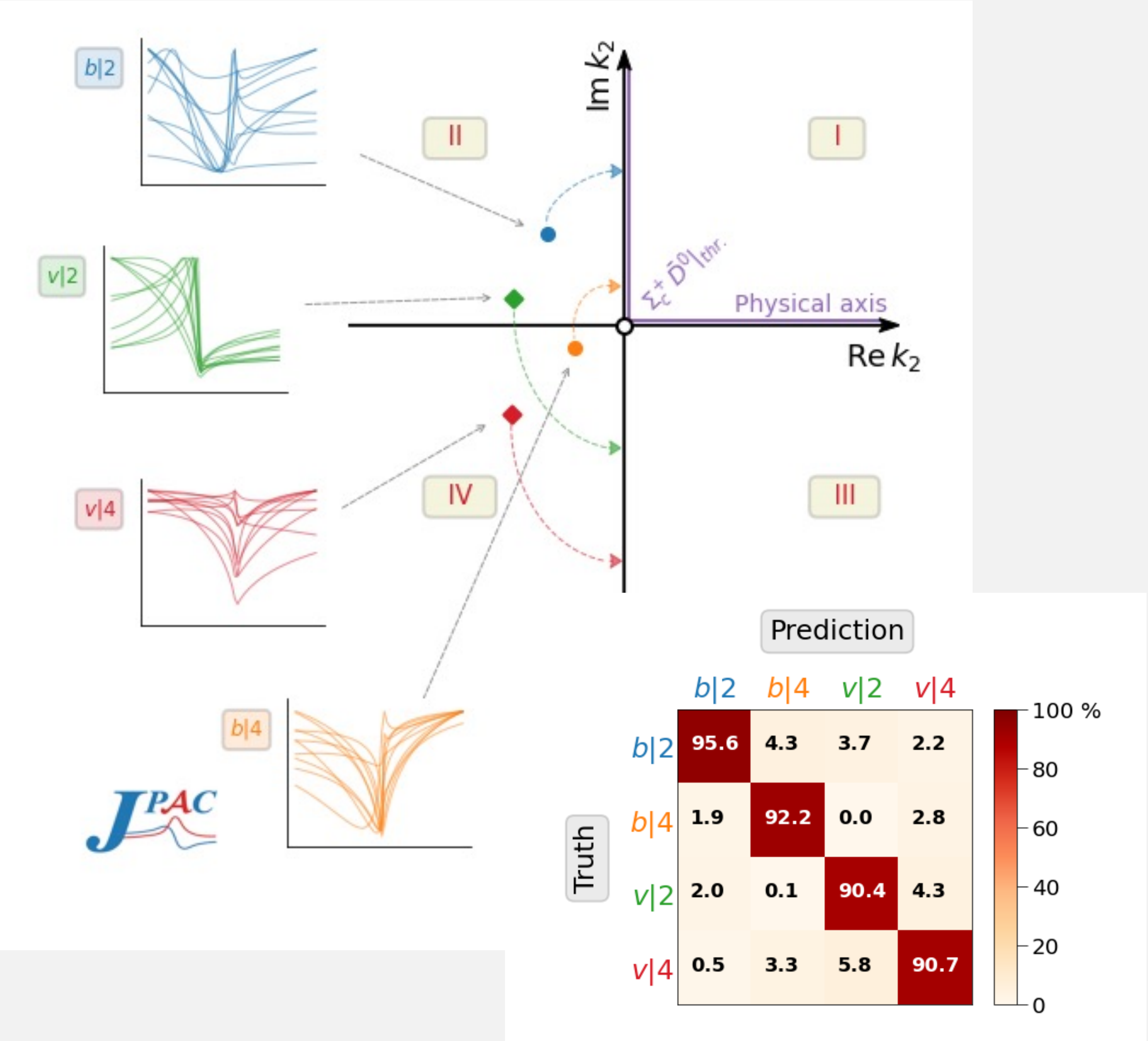
Building training set. Training the DNN

10^5 training curves generated by randomly choosing parameter values

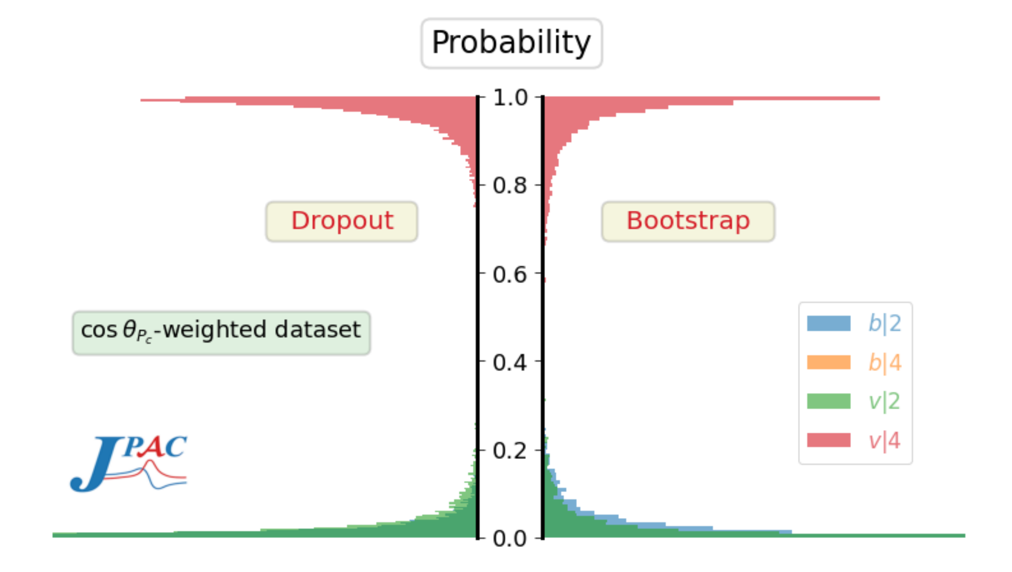
To mimic uncertainties a a 5% noise is included in the training set

The lineshapes are convoluted with the experimental resolution

Experimental uncertainty limits the accuracy of the DNN



Applying the DNN



We pass the three LHCb datasets (full, cut in m_{Kp} and weighted) through the DNN to obtain an answer. Note that we pass the three dataset through the same DNN

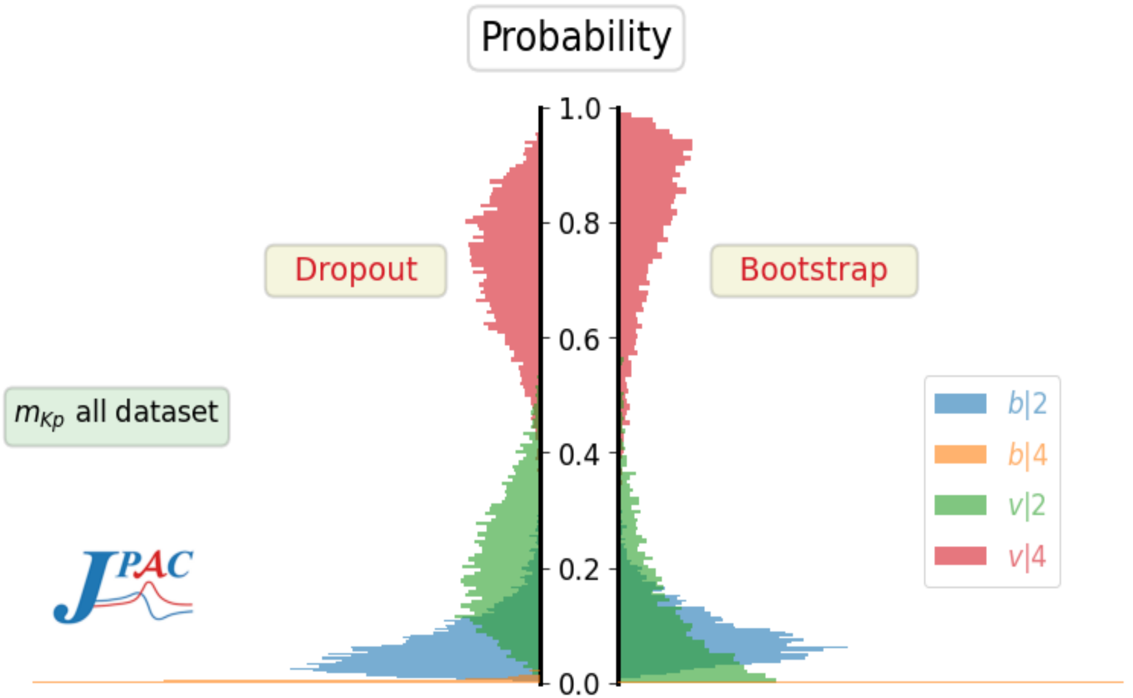
To account for uncertainties, we use two procedures: Bootstrap and Dropout

We (unsurprisingly) recover the same result as with the standard approach: v|4

But we are learning new things…

Three datasets analyzed with the same DNN

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



What we get from the DNN

The DNN targets specific regions of the parameter space (which yield stable solutions) that might be difficult to reach during optimization or might require high-resolution data

Standard χ^2fit can be indeed unstable, and a small change in the input data can induce large changes in the parameter values and therefore in the physics interpretation

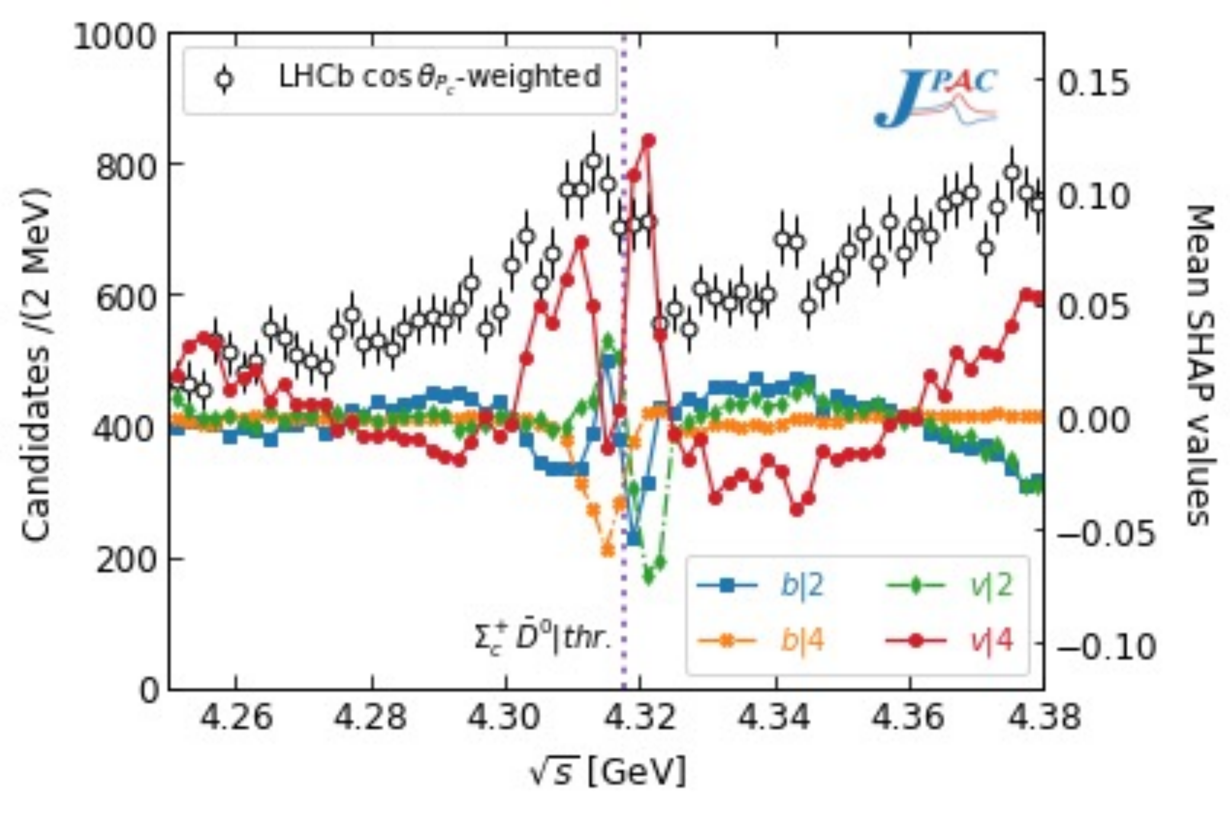
Rather than testing a single model hypothesis as a χ^2 fit would, the DNN determines the probability of each of the classes of interest, given the experimental uncertainties. The latter is possible, since the DNN learns the subtle classification boundary between the different classes

But, there's more…

SHapley Additive exPlanations (SHAP) values

Technique inherited from game theory

Allows to determine how a given feature in the input layer (in our case an experimental datapoint) impacts the decisison made by the DNN in the output layer (the classes)

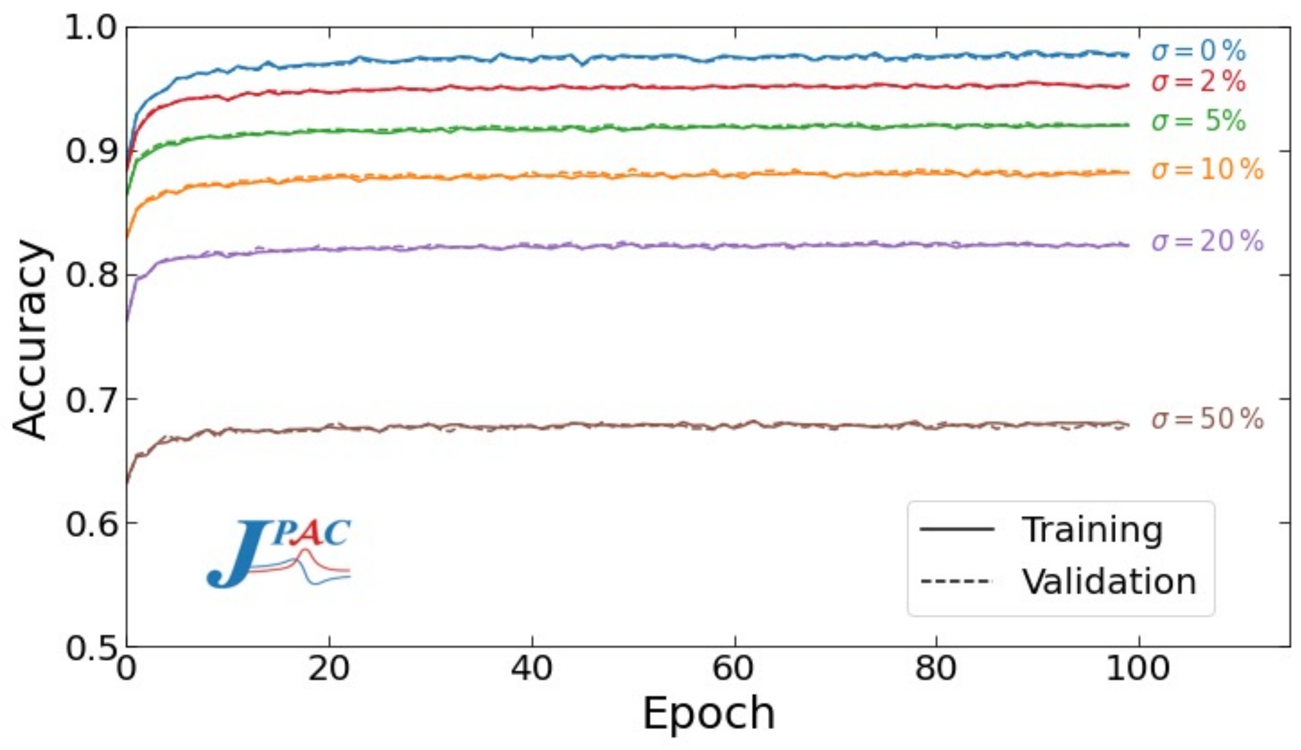


Takeaways

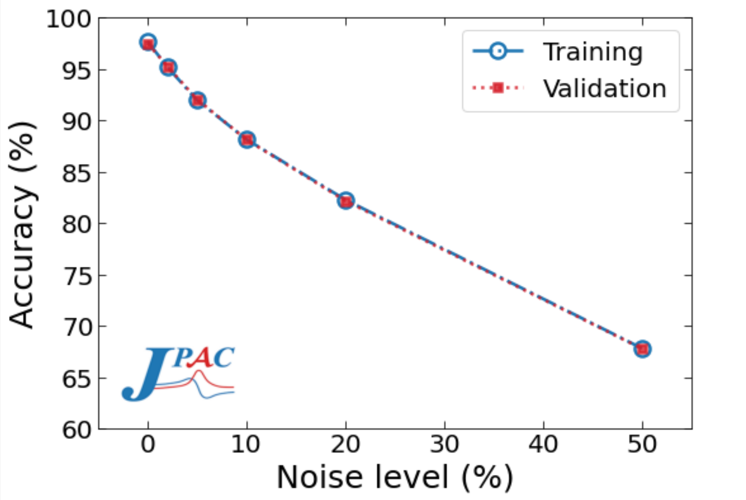
- Deep neural networks open new possibilities to answer the question on the underlying nature of a given resonance
- Possibility to gain physics insight on how the data impact the obtained interpretation
- DNN does NOT substitute the standard approach. They are complementary (we still want the amplitude and the pole position)
- DNN allows a true comparison among interpretations
- The BIG objective: Be able to train the DNN with every amplitude possibility we can devise with our twisted minds and throw the data to it, returning probabilities for each class
- Uncharted territory so we are taking baby steps: Ng et al. (JPAC), 2110.13742
- We are (hopefully) just in the begining…

Additional material

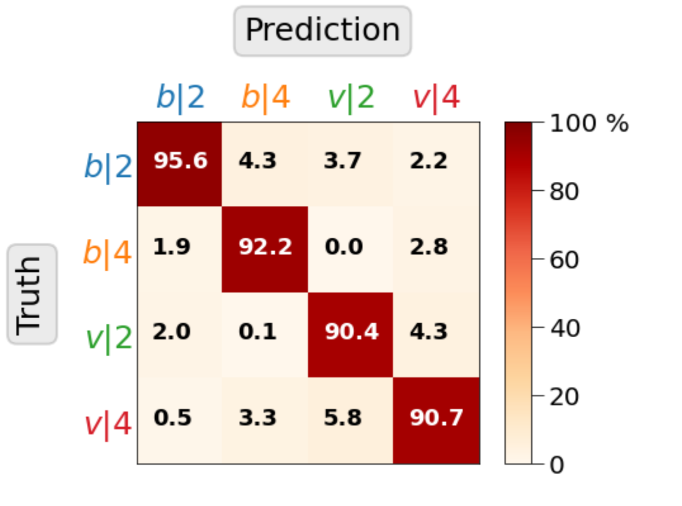
Training



σ = 0 %
σ = 2 %
σ = 5%
σ = 10 %
σ = 20 %
σ = 50 %
Training
Validation
Accuracy
Epoch
JPAC

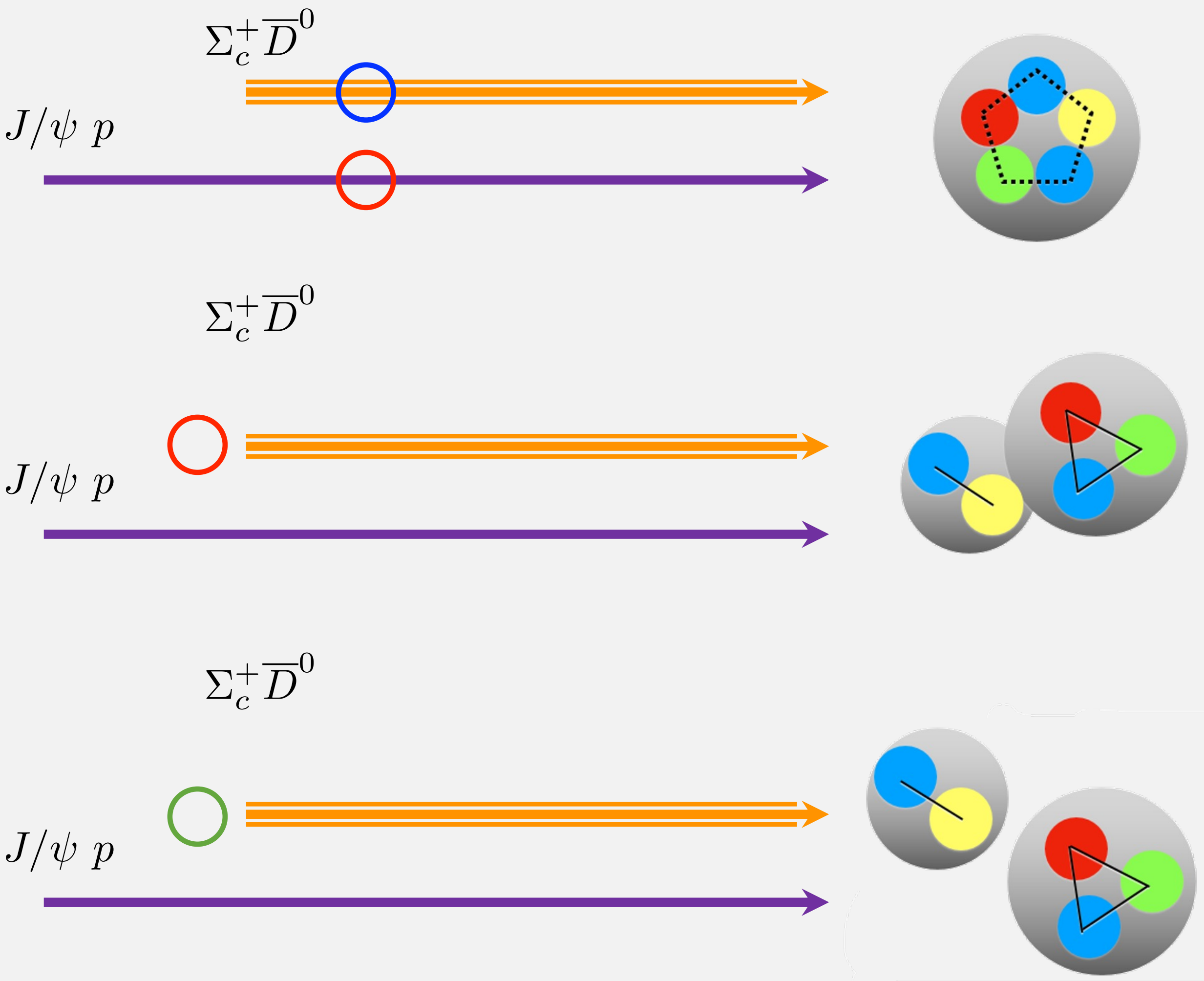


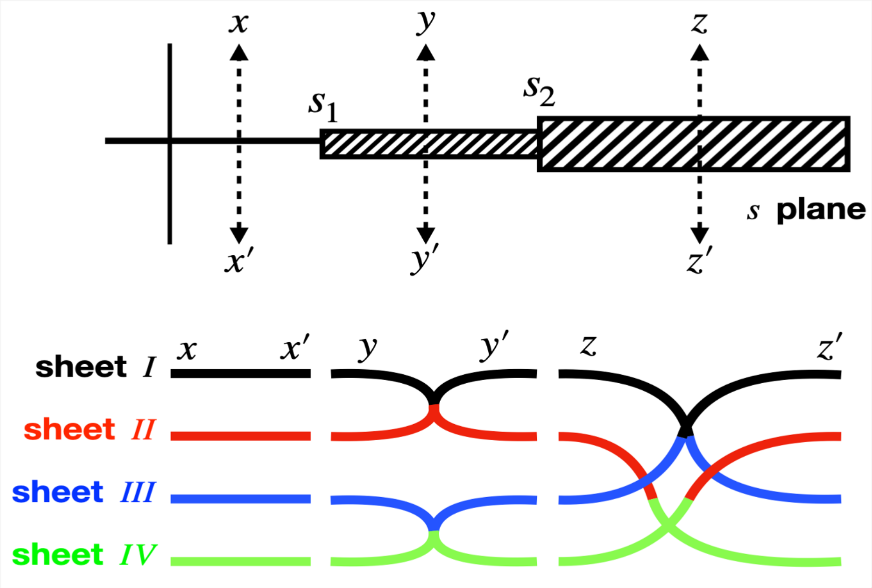
Training
Validation
Accuracy (%)
Noise level (%)
JPAC



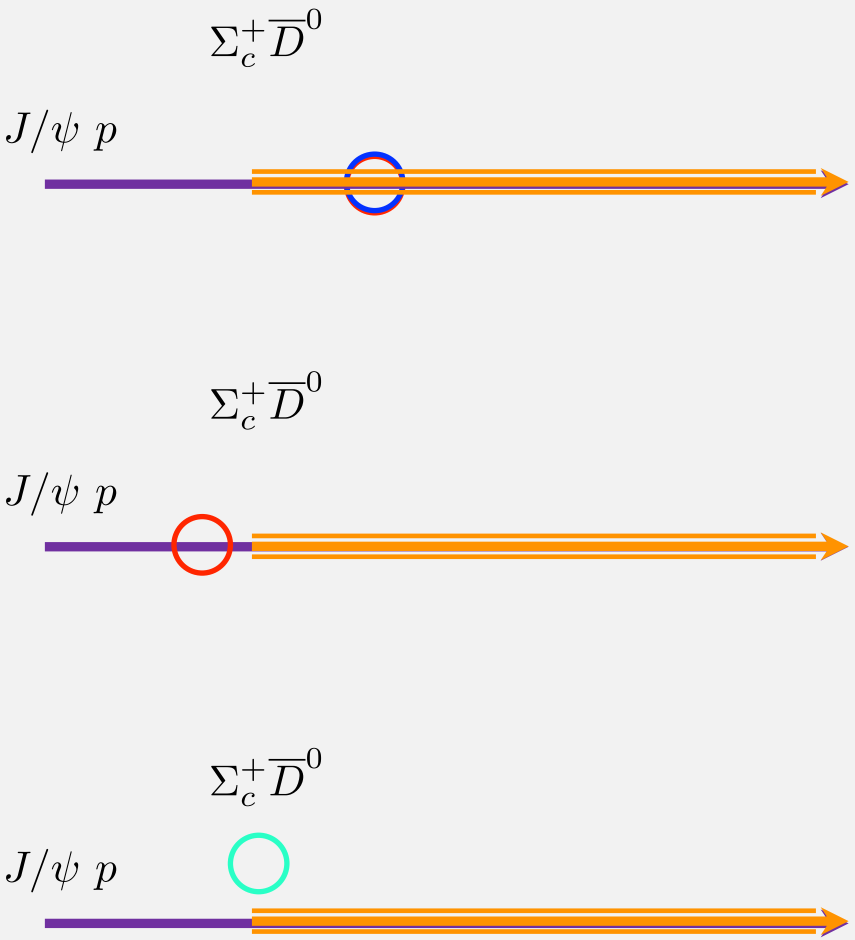
Prediction
Truth
b|2
b|4
v|2
v|4
100 %

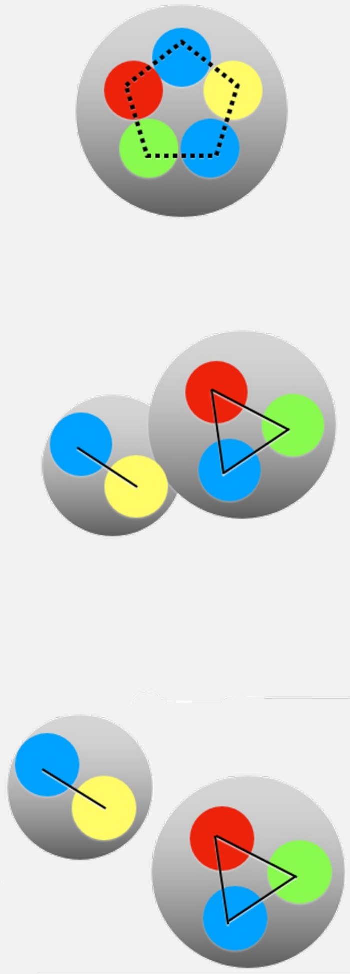
Pole movement (model dependent)

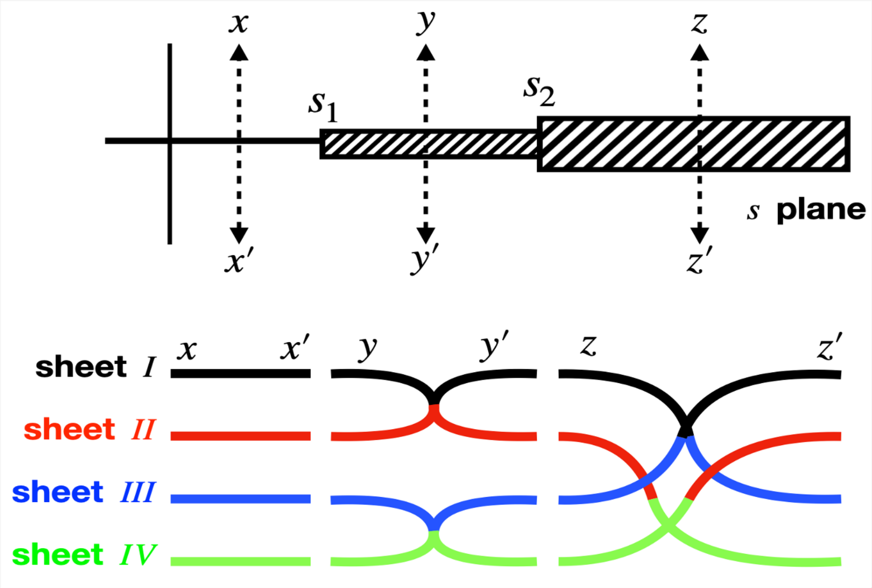




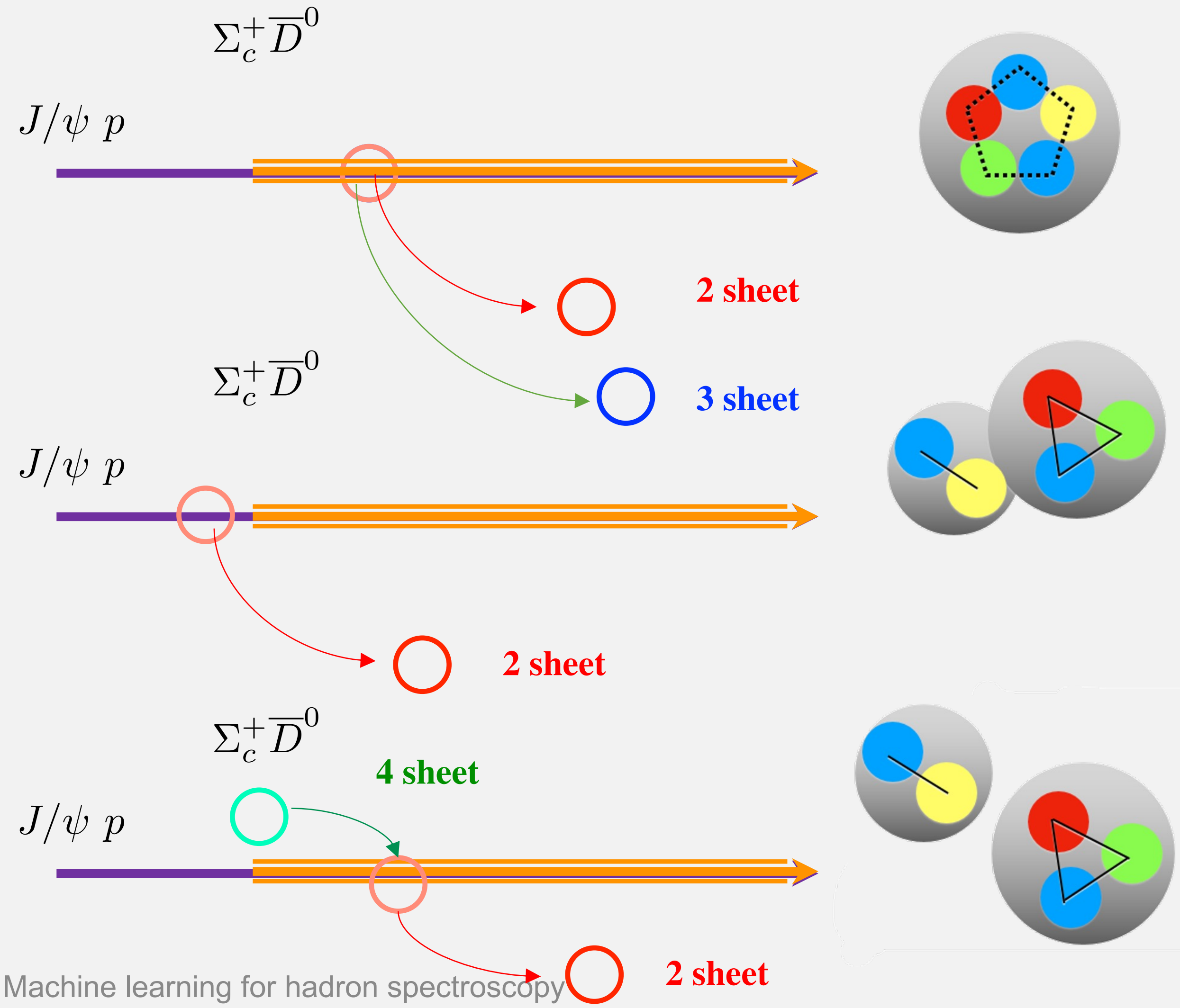
Pole movement (model dependent)

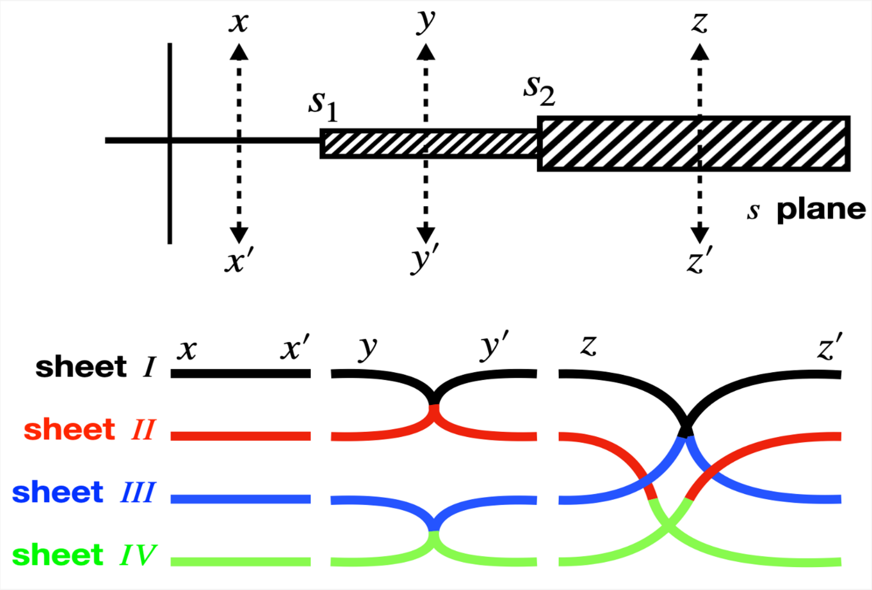






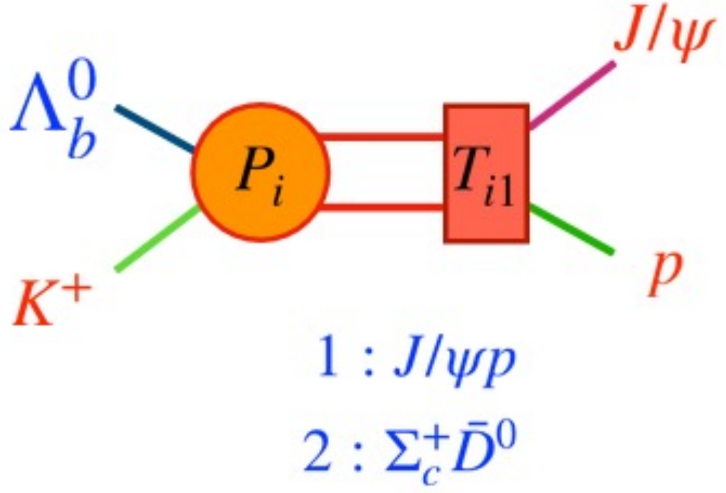
Pole movement (model dependent)





Near-threshold theory (two channels)

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

$F(s) =$ 

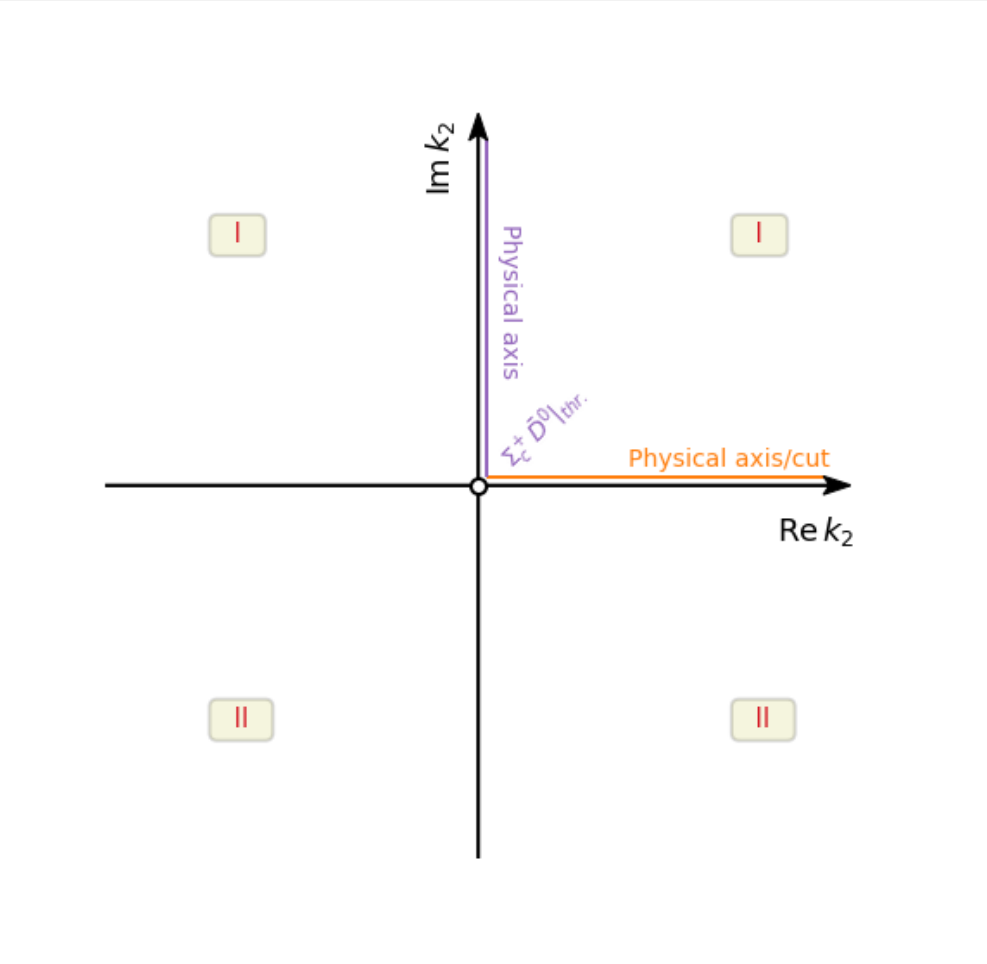
$$F(s) = P_1(s) T_{11}(s) \qquad \left(T^{-1}\right)_{ij} = M_{ij} - i k_i \delta_{ij}$$

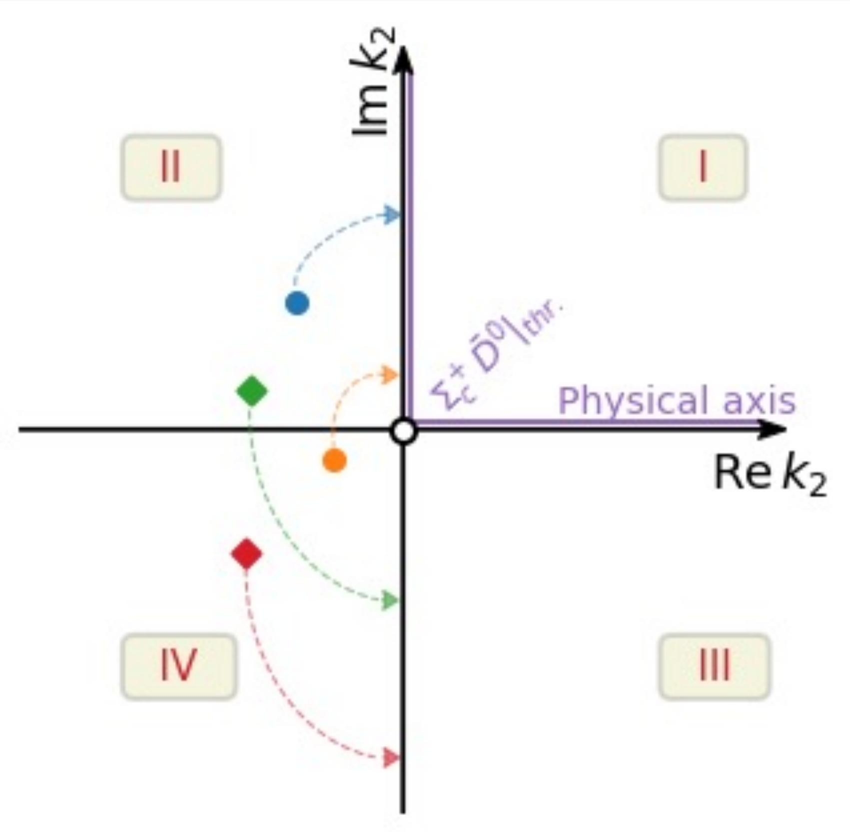
$$M_{ij}(s) = m_{ij}$$

Scattering length approximation for *T*

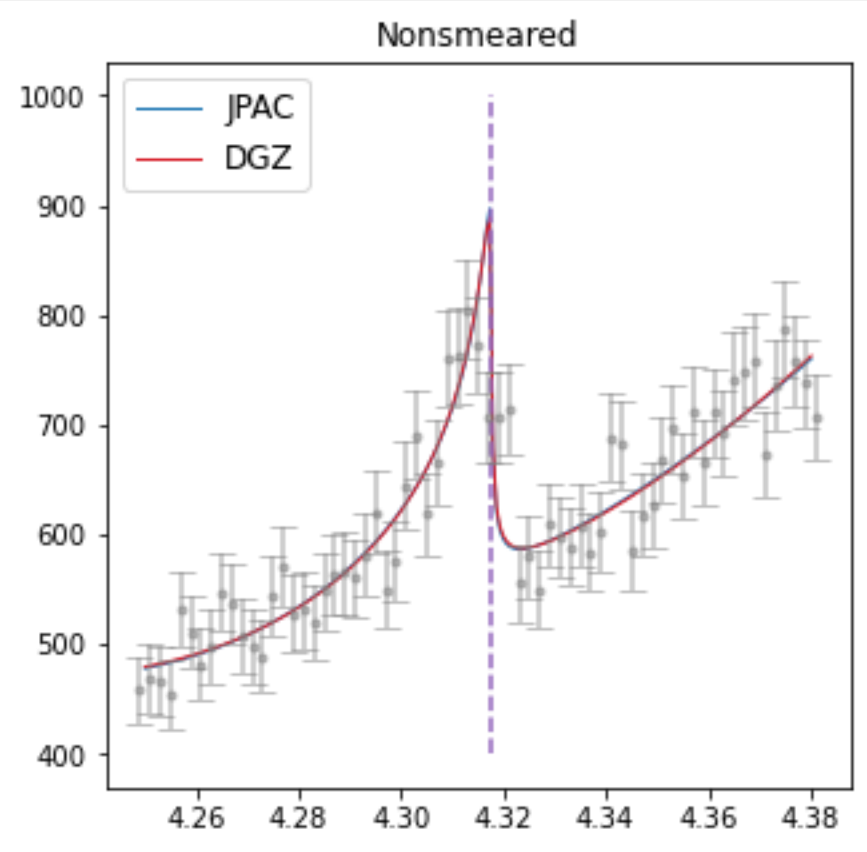
Dong, Guo & Zou (DGZ), 2011.14517
CFR et al. (JPAC), 1904.10021

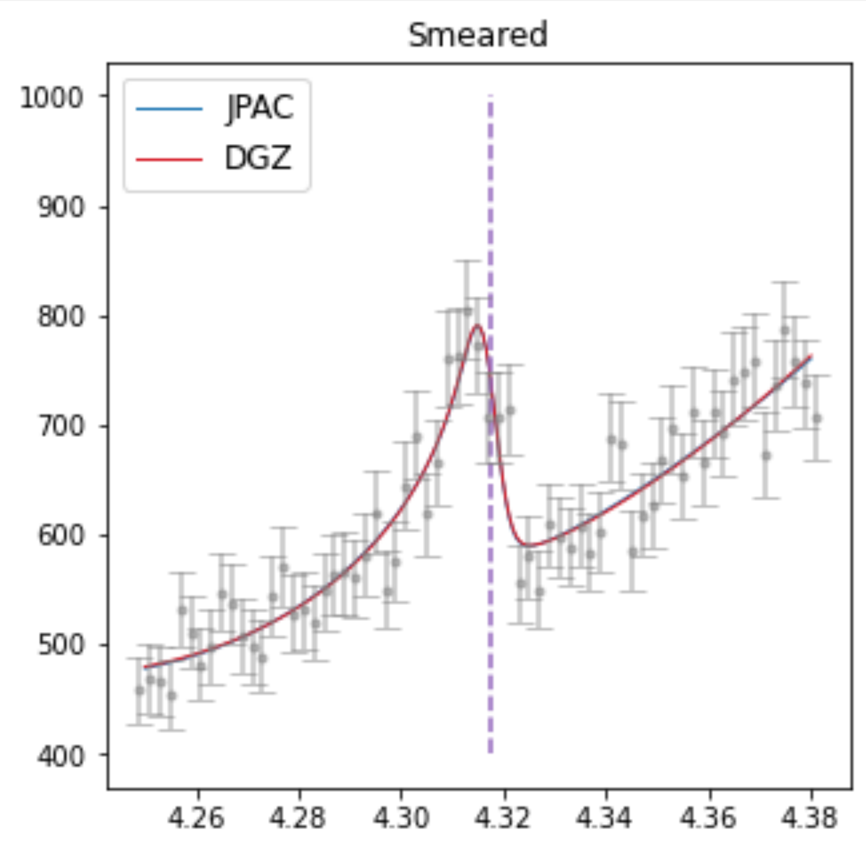
DGZ vs. JPAC (cut)





DGZ vs. JPAC amplitudes (fit to data+resolution)

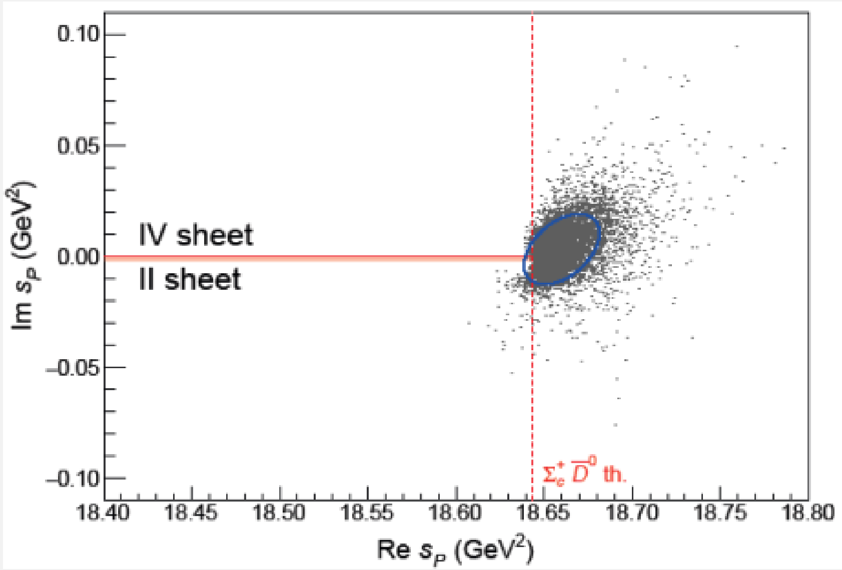




DGZ vs. JPAC (pole position)

Using JPAC amplitude for T

$$M_P = 4319 \pm 1.6 \text{ MeV}$$
$$\Gamma_P = -0.84 \pm 2.4 \text{ MeV}$$



Using DGZ amplitude for T

$$M_P = 4319.45 \text{ MeV}$$
$$\Gamma_P = 0.95 \text{ MeV}$$

Note: minus sign refers to 4th Riemann sheet

Difference in momenta (actually irrelevant)