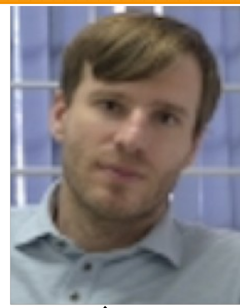
20th International Conference on Hadron Spectroscopy and Structure (HADRON 2023)
Genova, Italy, 5-9 June 2023

Machine learning exotic hadrons

Łukasz Bibrzycki
Faculty of Physics and Applied Computer Science
AGH University of Krakow

Joint Physics Analysis Center (JPAC)

Mikhail Mikasenko
TU Munich



Cesar Fernández-Ramírez
National Autonomous U. of Mexico

Daniel Winney
South China Normal U.

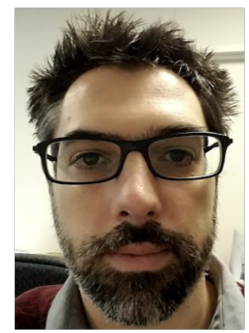
Victor Mokeev
Jefferson Lab

Sergi Gonzalez-Solis
Los Alamos National Lab.

Alessandro Pilloni
U. Messina

Łukasz Bibrzycki
AGH Kraków

Astrid Hiller Blin
U. of Heidelberg



Vincent Mathieu
U. Barcelona

Adam Szczepaniak
Indiana U.

Miguel Albaladejo
CSIC-Valencia

Arkaitz Rodas
College of William and Mary

Robert Perry
U. Barcelona

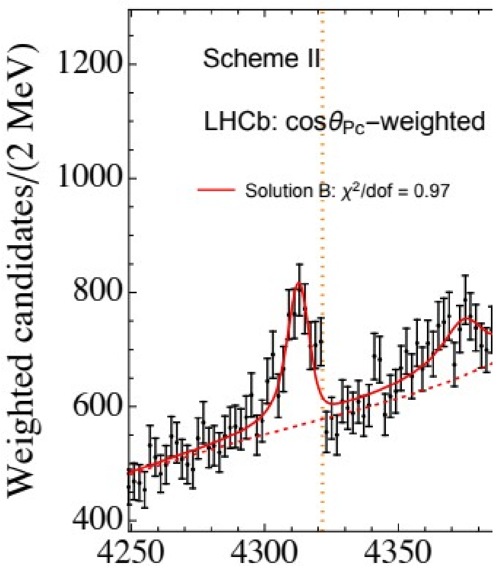
Jorge A. Silva-Castro
National Autonomous U. of Mexico

Lawrence Ng
Florida State U.

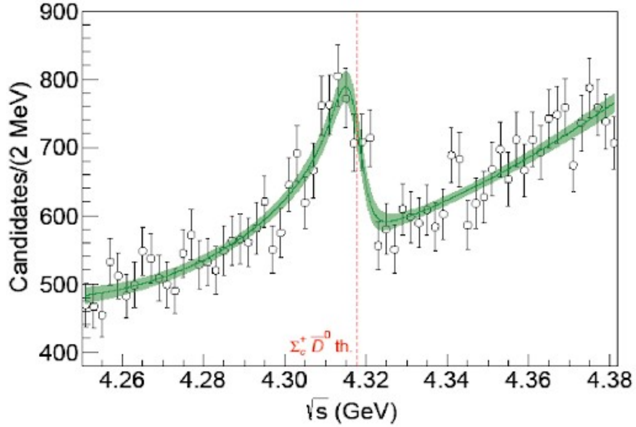
AGH

JPAC

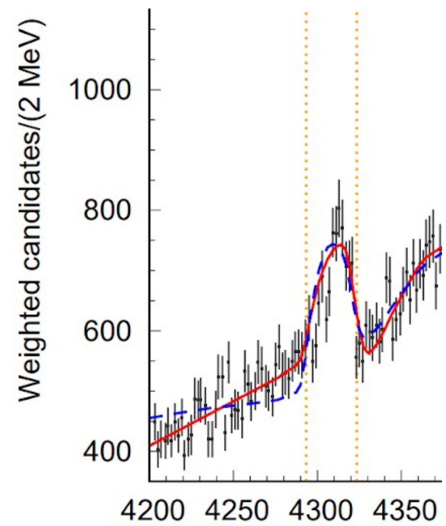
Discrepant interpretations of the $P_c(4312)$ nature



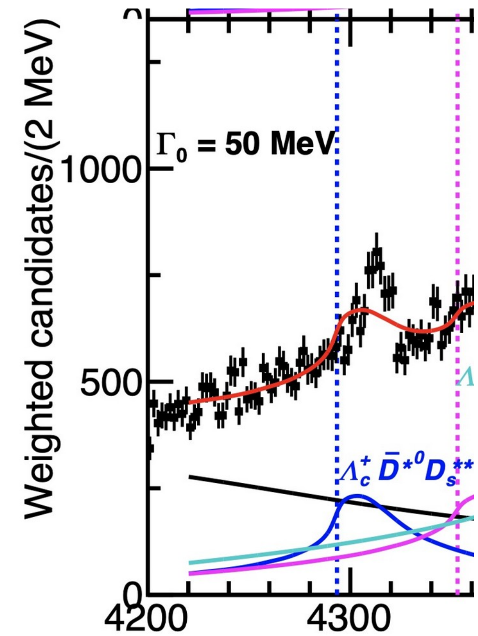
Molecule
Du et al., 2102.07159



Virtual
C. F-R et al. (JPAC), Phys. Rev. Lett. 123, 092001 (2019)



Double-triangle (w. complex coupl. in the Lagrangian)
Nakamura, Phys. Rev. D 103, 111503 (2021)



Single triangle (ruled out)
LHCb, Phys. Rev. Lett. 122, 222001 (2019)

See Cesar's slides from Wednesday session on Light Mesons for connection between the pole number/location and their physical interpretation

AGH

JPAC

Starting point - partial wave expansion of the amplitude

$$f_{ji}(s,t) = \sum_{l=0}^{\infty} (2l+1) T^l_{ji}(s) P_l(\cos\theta(s,t))$$

- Unitarity relation for partial wave amplitudes

$$\hat{T}^l - \hat{T}^{l\dagger} = 2i \hat{T}^l \hat{p} \hat{T}^{l\dagger}$$

- For fixed partial wave this can be solved in the general form:

$$\hat{T}^{-1} = \hat{M}(s) - i\hat{p}$$

- It was shown by *Frazer, Hendry, Phys. Rev. 134 (1964)* that $\hat{M}$ is symmetric and free from unitarity cuts, so can be Taylor expanded in *s*.

$$\hat{M} \approx \hat{m} - \hat{c}s$$

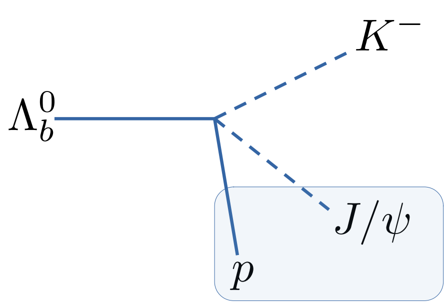
Scattering length

Effective range

AGH

JPAC

Physics model



- $P_c(4312)$ seen as a maximum in the pJ/ψ energy spectrum
- $P_c(4312)$ has a well defined spin and appears in single partial wave
- $\Sigma^+_c \bar{D}^0$ channel opens at 4.318 GeV -coupled channel problem
- Background contributes to all other waves

- Intensity

$$\frac{dN}{d\sqrt{s}} = \rho(s)\left[|P_1(s)T_{11}(s)|^2 + B(s)\right]$$

where

$$\rho(s) = pqm_{\Lambda_b} \qquad \text{phase space}$$

$$p = \lambda^{\frac{1}{2}}(s, m^2_{\Lambda_b}, m^2_K)/2m_{\Lambda_b}, \;\; q = \lambda^{\frac{1}{2}}(s, m^2_p, m^2_\psi)/2\sqrt{s}$$

$$P_1(s) = p_0 + p_1 s \qquad \text{production term}$$

$$B(s) = b_0 + b_1 s \qquad \text{background term}$$

Physics model

- 2-channel model

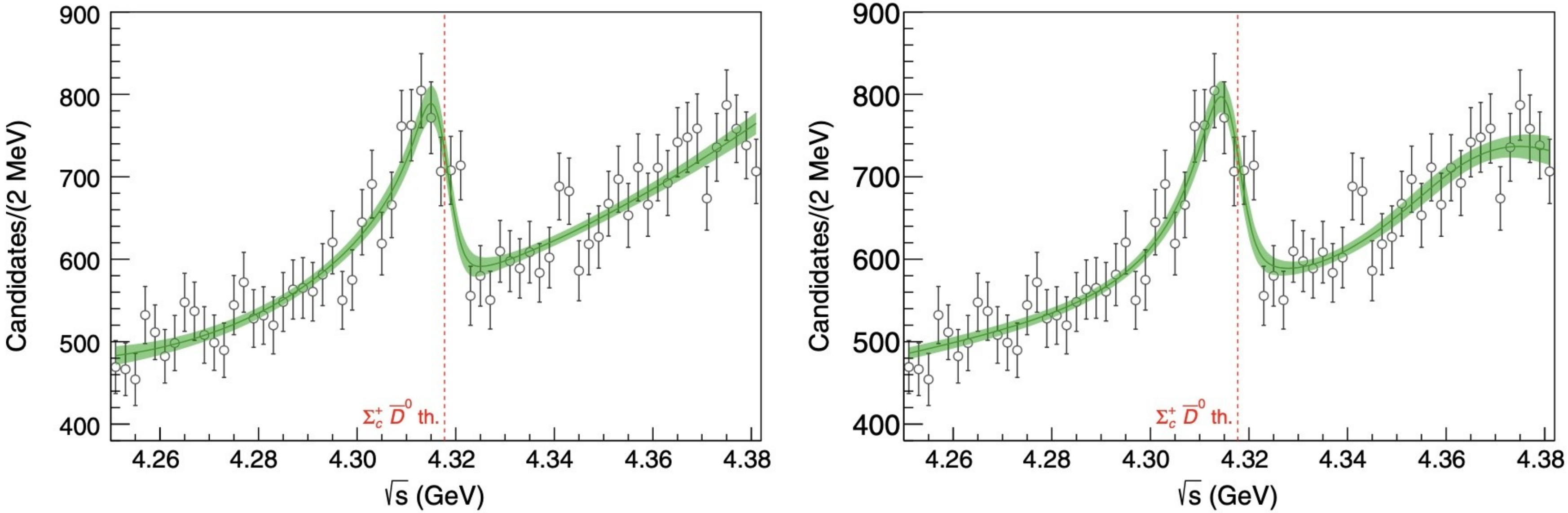
$$T_{ij}^{-1} = M_{ij} - i k_i \delta_{ij} \quad \text{where} \quad k_i = \sqrt{s - s_i}$$

$$s_1 = (m_p + m_{J/\psi})^2 \quad \text{and} \quad s_2 = (m_{\Sigma_c^+} + m_{\bar{D}^0})^2$$

- How accurate the Taylor expansion of M_{ij} has to be ?

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

Physics model – final version



Scattering length approximation   Effective range approximation

See *C. Fernandez-Ramirez Phys.Rev.Lett. 123 (2019) 9, 092001*

Finally we use the scattering length approximated amplitude as the basis for ML model

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

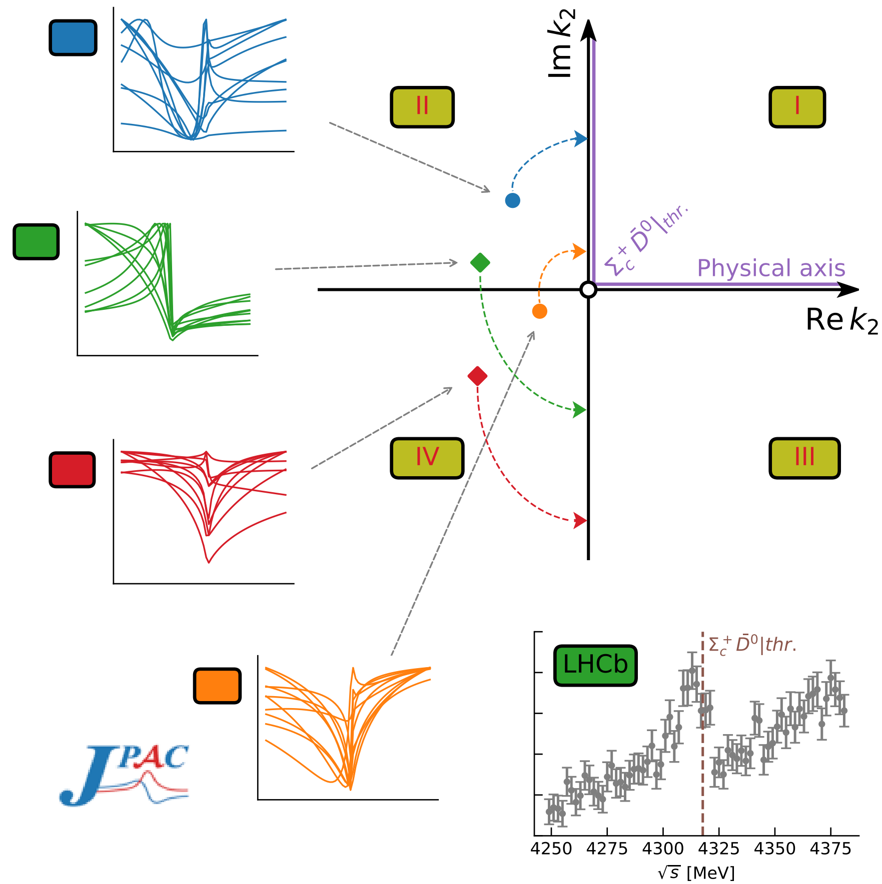
7 model parameters in total: m_{11}, m_{22}, m_{12}, p_0, p_1, b_0, b_1.

AGH

ML model – general idea

- From the physical model we produce:
  - Sample intensities (computed in 65 energy bins) – produced with randomly chosen parameter samples – **examples**
  - For each parameter sample we are able to compute the **target class** – one of the four: b|2, b|4, v|2, v|4
  - Symbolically:

$$K : \{[I_1, \ldots, I_{65}](m_{11}, m_{22}, m_{12}, p_0, p_1, b_0, b_1)\} \rightarrow \{b|2, b|4, v|2, v|4\}$$

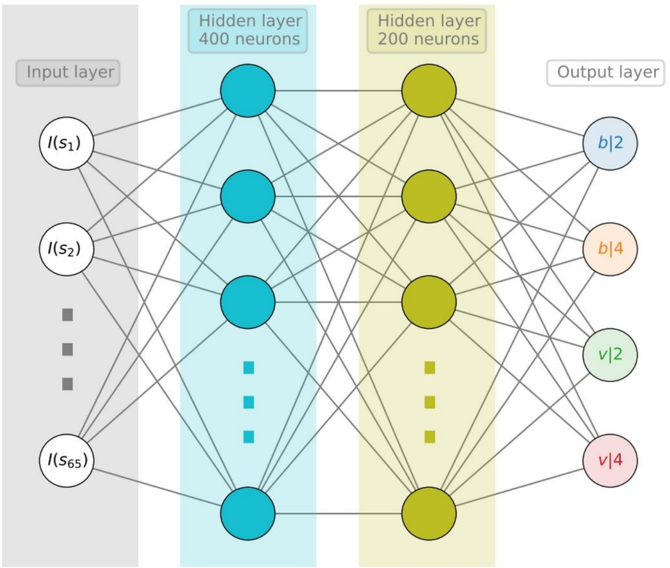


AGH

JPAC

ML model – MLP

Layer	Shape in	Shape out
Input		(B, 65)
Dense	(B, 65)	(B, 400)
Dropout(p=0.2)	(B, 400)	(B, 400)
ReLU	(B, 400)	(B, 400)
Dense	(B, 400)	(B, 200)
Dropout(p=0.5)	(B, 200)	(B, 200)
ReLU	(B, 200)	(B, 200)
Dense	(B, 200)	(B, 4)
Softmax	(B, 4)	(B, 4)



Training dataset preparation:

1. Parameters were uniformly sampled from the following ranges: b_0 = [0 ; 700], b_1 = [-40 ; 40], p_0 = [0 ; 600], p_1 = [-35 ; 35], M_{22} = [-0.4 ; 0.4], M_{11} = [-4 ; 4], M_{12}^2 = [0 ; 1.4]
2. The signal was smeared by convolving with experimental LHCb resolution:

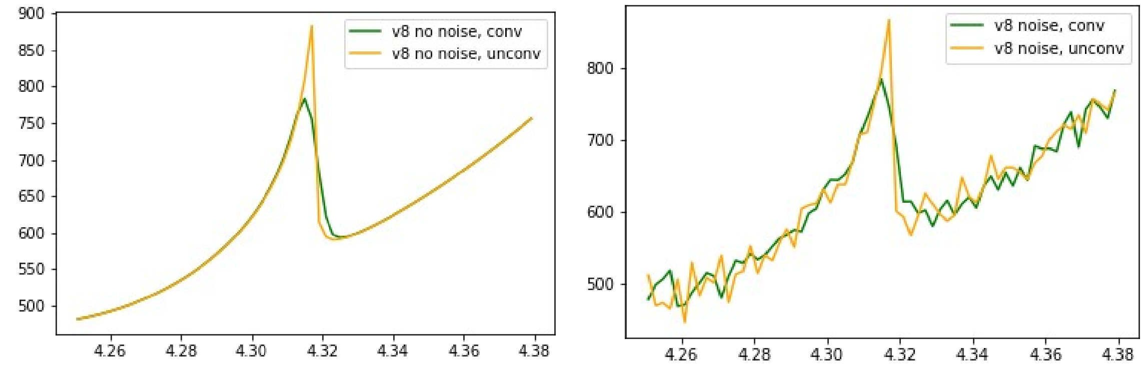
$$I(s) = \int_{m_\psi+m_p}^{m_{\Lambda_b}-m_K} I(s')_{\text{theo}} \exp\left[-\frac{(\sqrt{s}-\sqrt{s'})^2}{2R^2(s)}\right] d\sqrt{s'} \Bigg/ \int_{m_\psi+m_p}^{m_{\Lambda_b}-m_K} \exp\left[-\frac{(\sqrt{s}-\sqrt{s'})^2}{2R^2(s)}\right] d\sqrt{s'}\,,$$

$$R(s) = 2.71 - 6.56 \times 10^{-6-1} \times \left(\sqrt{s} - 4567\right)^2$$

3. To account for experimental encertainty the 5% gaussian noise was added

ML model - training

- Input examples (effect of energy smearing and noise):



- Computing target classes:
  - $m_{22}>0$ – bound state, $m_{22}<0$ – virtual state
  - To localize the poles on Riemann sheets we need to find zeros of the amplitude denominator in the momentum space:

$$p_0 + p_1\, q + p_2\, q^2 + p_3\, q^3 + q^4 = 0$$

with

$$p_0 = (s_1 - s_2)\, m_{22}^2 - \left(m_{12}^2 - m_{11} m_{22}\right)^2$$
$$p_1 = 2\,(s_1 - s_2)\, m_{22} + 2 m_{11} \left(m_{12}^2 - m_{11} m_{22}\right)$$
$$p_2 = -m_{11}^2 + m_{22}^2 + s_1 - s_2$$
$$p_3 = 2 m_{22}$$

Then poles appear on sheets defined with (η_1,η_2) pairs:

(-,+) - II sheet

(+,-) - IV sheet

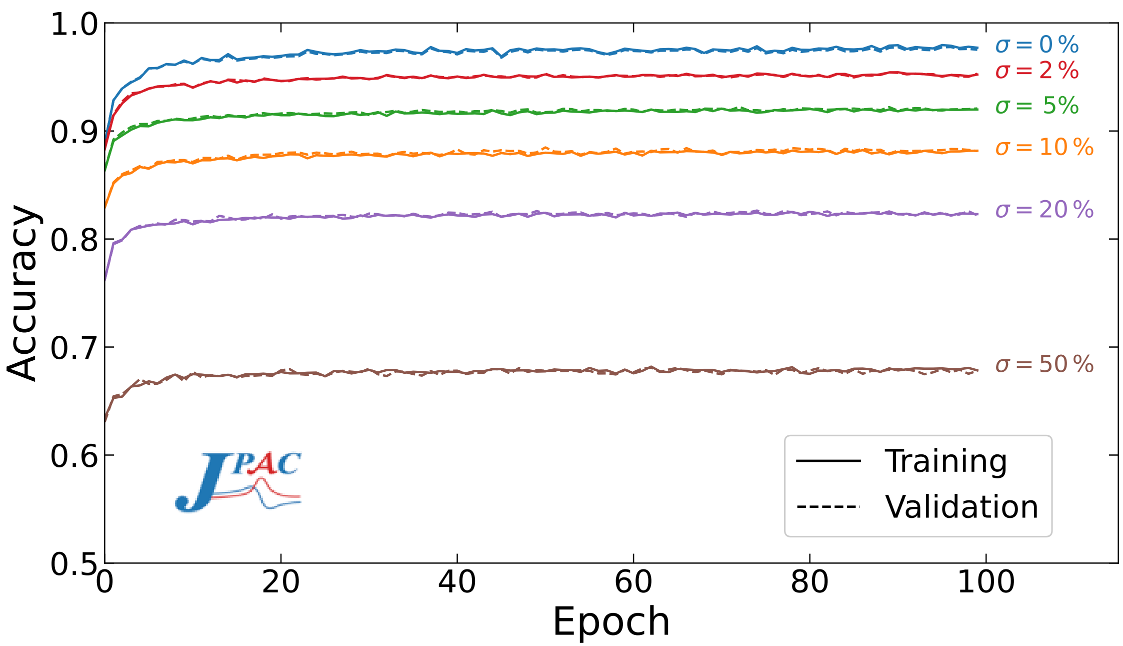
$$\eta_1 = \text{Sign Re}\left(\frac{m_{12}^2}{m_{22}+q} - m_{11}\right) \quad \eta_2 = \text{Sign Re} q$$

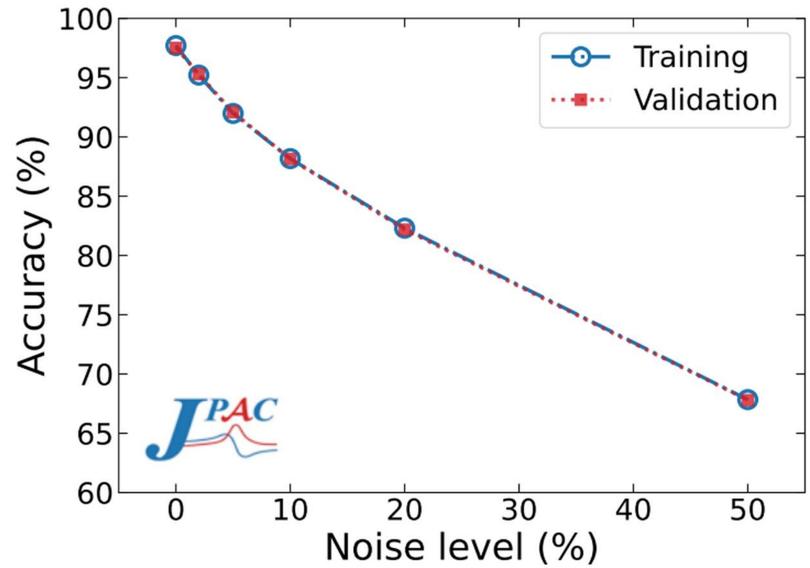
AGH

JPAC

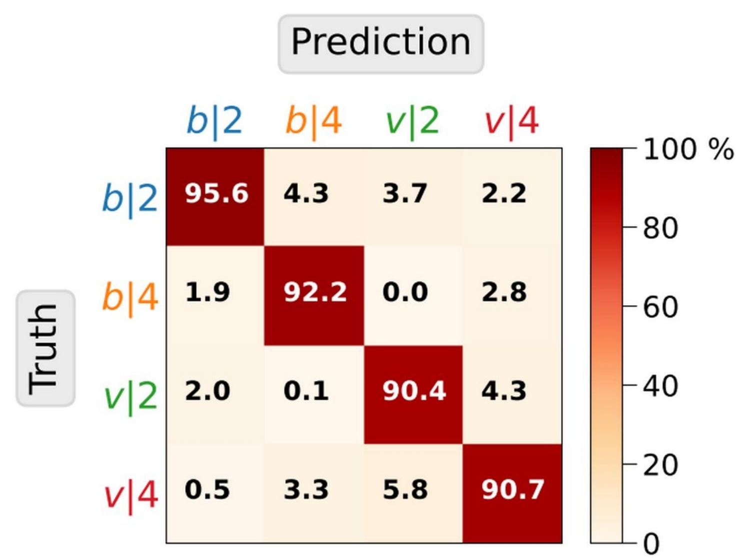
ML model – training results

Accuracy for various noise levels



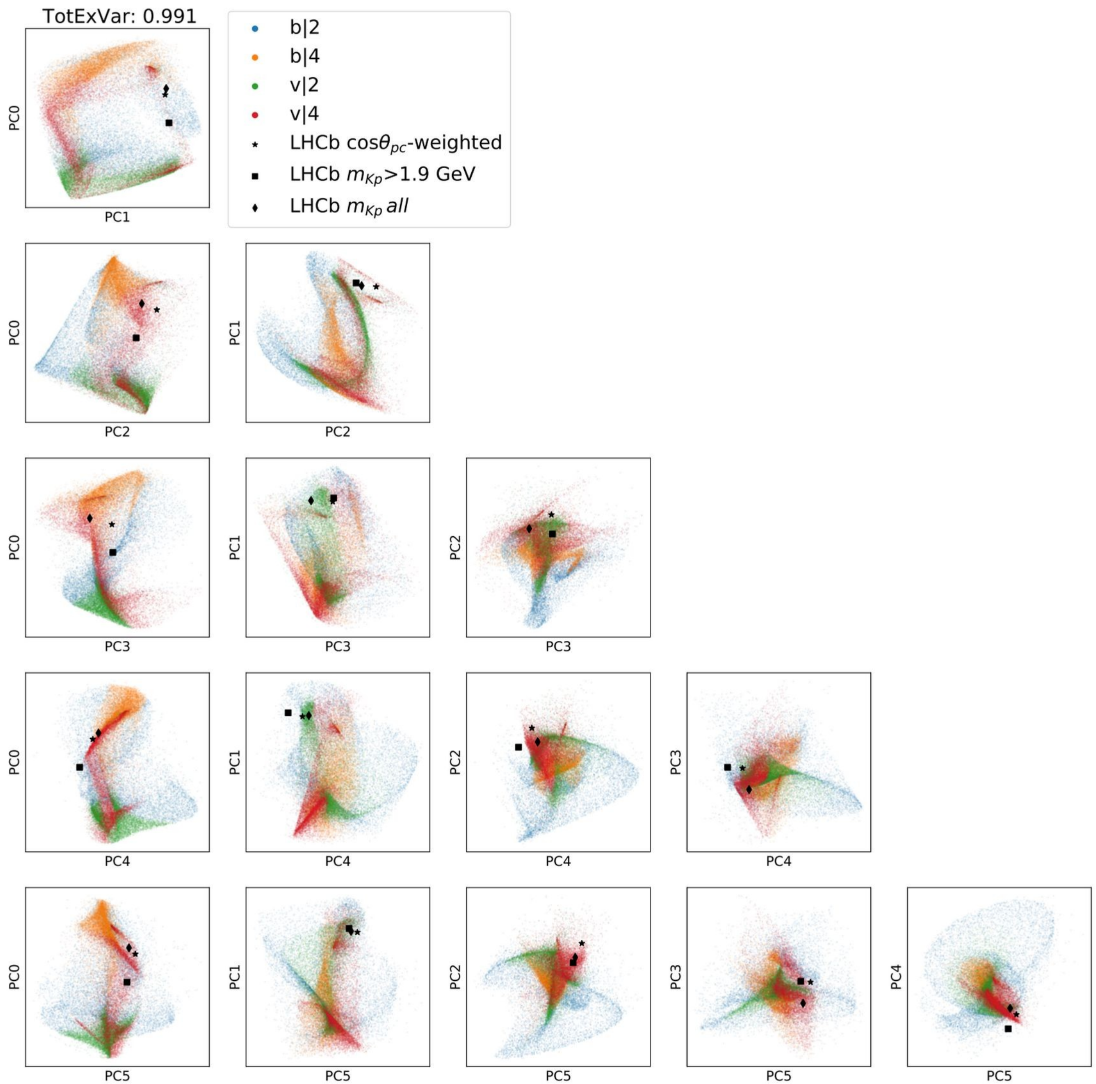


Confusion matrix for the 5% noise



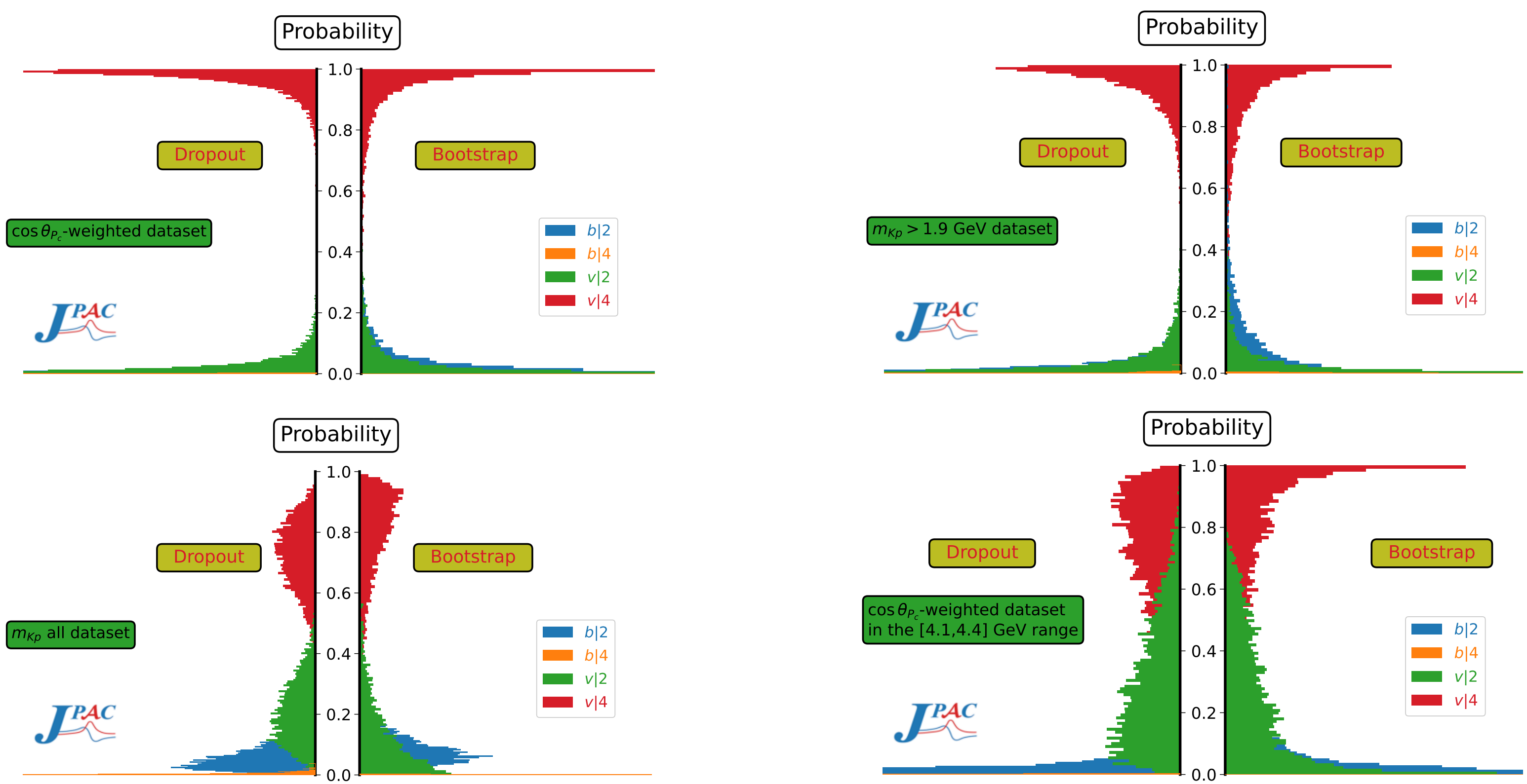
Does the training data set reflect experimental situation ?

- Dimensionality reduction - Principal Component analysis
- Over 99% of the variance can be explained with just 6 features
- Experimental data projected onto principal components are well encompassed within the training dataset



Model predictions – statistical analysis

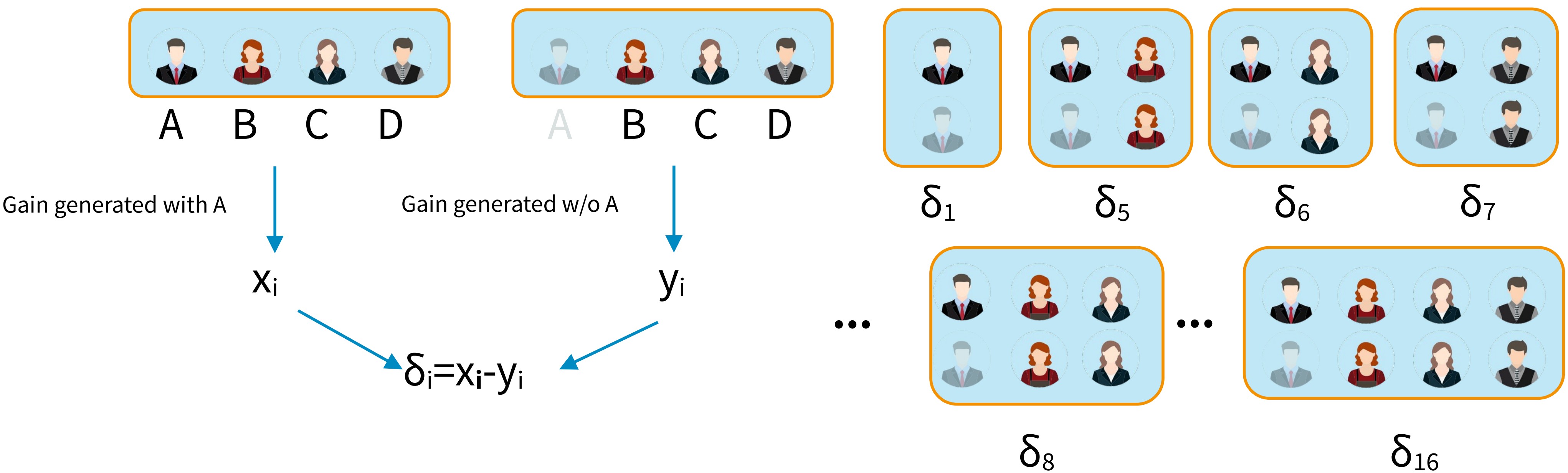
- The distribution of the target classes was evaluated with
  - the bootstrap (10 000 pseudo-data based on experimental mean values and uncertainties) and
  - dropout (10 000 predictions from the ML model with a fraction of weights randomly dropped out)



Model explanation with SHAP

- Shapley values and Shapley Additive Explanations

Shapley, Lloyd S. "Notes on the n-Person Game -- II: The Value of an n-Person Game" (1951)



Shapley value for member A: $$\phi_A = \frac{\delta_1 + \delta_2 + \cdots + \delta_{16}}{16}$$

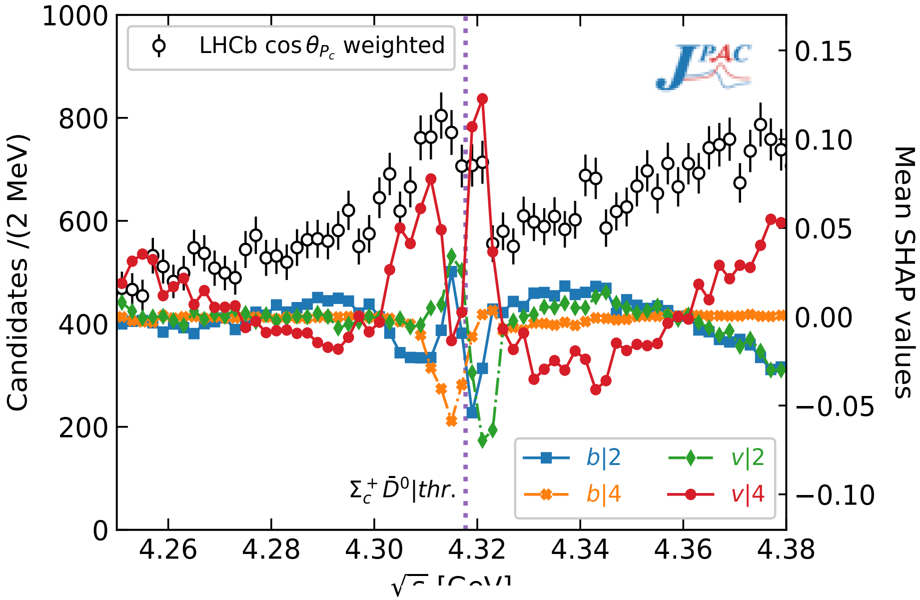
Model explanation with SHAP

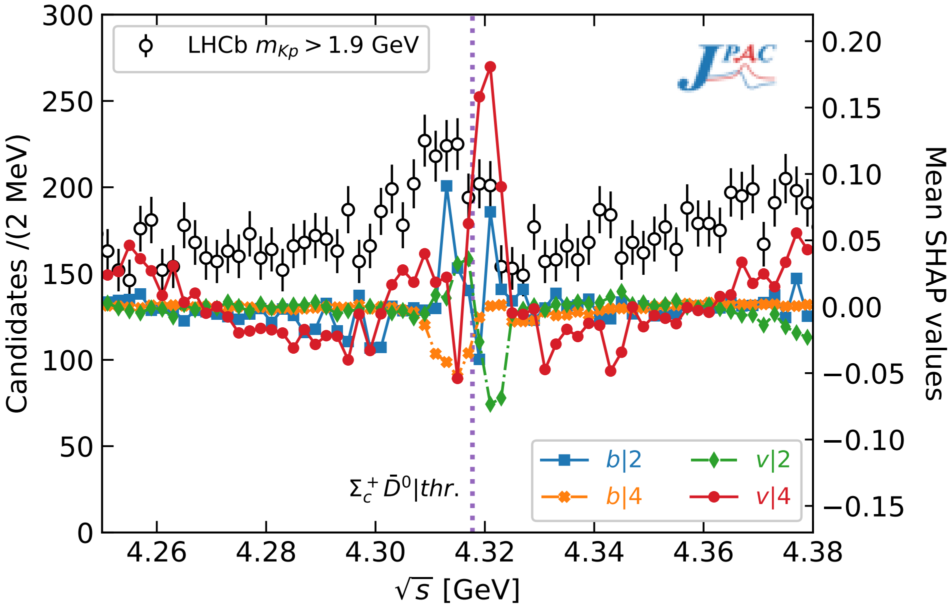
- By making an association:
  - Member of a coalition → Feature
  - Game → Function that generates classification/regression result
  - Gain → Prediction
  - We define the Shapley values for features
- Caveats:
  - A number of possible coalitions grows like 2^N
  - Prohibitively expensive computationally (NP-hard)

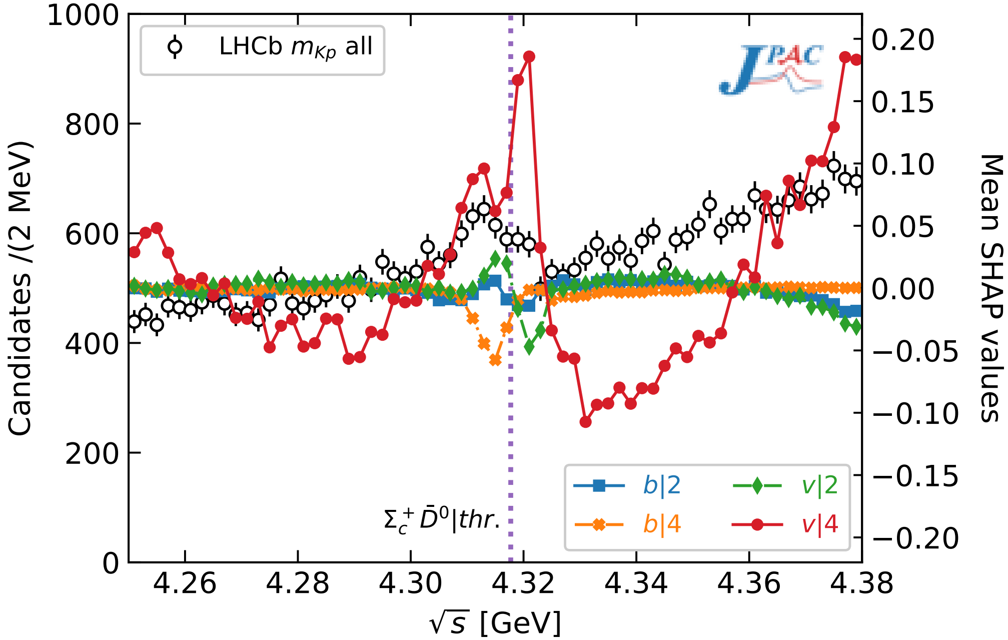
Solution: Shapley additive explanations (Lundberg, Lee, arXiv:1705.07874v2, 2017)

AGH

Model explanation with SHAP







AGH

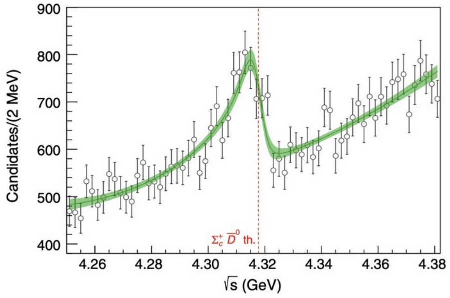
Summary

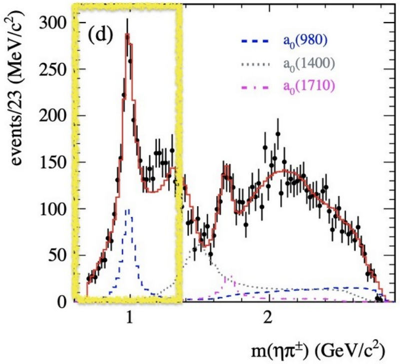
- Takeaways:
  - Standard χ^2 fit may be unstable, since small change in the input may result in large parameter fluctuations (change physics interpretation)
  - Rather than testing the single model hypothesis with χ^2, we obtained the probabilities of four competitive pole assignments for the $P_c(4312)$ state
  - The approach was model independent – meta model
  - By the analysis of the SHAP values we obtained an *ex post* justification of our scattering length approximation

AGH

Questions to be addressed

- Applying the method for larger class of resonances, described by the same physics, eg. $a_0(980)/f_0(980)$ [discussed by C. Fernadez-Ramires on Wednesday] or other resonances located near thresholds





Two situations with increasing level of generality:

- Same resonance but observed in different experiments, energy bins, channels, observables
- Various signals/resonances for which we believe they age governed by the same dynamical mechanism, eg. the interplay of several poles lying close to the physical region
- For both cases recurrent neural networks are promising tool for analysis