IL NUOVO CIMENTO **47 C** (2024) 207
DOI 10.1393/ncc/i2024-24207-8

Colloquia: HADRON2023

# Machine learning exotic hadrons

L. Ng([1]), Ł. Bibrzycki([2])(*), J. Nys([3]), C. Fernández-Ramírez([4]), A. Pilloni([5])([6]), V. Mathieu([7])([8]), A. J. Rasmusson([11]) and A. P. Szczepaniak([9])([10])([11])

([1]) *Department of Physics, Florida State University - Tallahassee, FL 32306, USA*
([2]) *Faculty of Physics and Applied Computer Science, AGH University of Krakow al. A. Mickiewicza 30, 30-059 Kraków, Poland*
([3]) *Institute of Physics, Ecole Polytechnique Fédérale de Lausanne (EPFL) CH-1015 Lausanne, Switzerland*
([4]) *Departamento de Física Interdisciplinar, Universidad Nacional de Educación a Distancia (UNED) - Madrid E-28040, Spain*
([5]) *Dipartimento di Scienze Matematiche e Informatiche, Scienze Fisiche e Scienze della Terra, Università degli Studi di Messina - Messina I-98166, Italy*
([6]) *INFN, Sezione di Catania - Catania I-95123, Italy*
([7]) *Departament de Física Quántica i Astrofísica and Institut de Ciències del Cosmos, Universitat de Barcelona - E-08028 Barcelona, Spain*
([8]) *Departamento de Física Teórica, Universidad Complutense de Madrid and IPARCOS E-28040 Madrid, Spain*
([9]) *Theory Center, Thomas Jefferson National Accelerator Facility - Newport News, VA 23606, USA*
([10]) *Center for Exploration of Energy and Matter, Indiana University - Bloomington, IN 47403, USA*
([11]) *Department of Physics, Indiana University - Bloomington, IN 47405, USA*

received 21 December 2023

**Summary.** — We show that a neural network trained with synthetic differential intensities calculated with scattering length approximated amplitudes classifies the $P_c(4312)^+$ signal as a virtual state located at the 4th Riemann sheet with very high certainty. This is in line with the results of other analyses but surpasses them by providing a simultaneous evaluation of probabilities of competing scenarios, like, *e.g.*, the interpretation as a bound state. Using the Shapley Additive Explanations we identified the energy bins which are key for the physical interpretation.

## 1. – Introduction

Since the observation of the first tetraquark candidate $\chi_{c1}(3872)$ at Belle [1], the analalysis of the structure and production mechanisms of exotic hadrons has been at the center of both experimental and theoretical effort of hadron physics community. The $P_c(4312)^+$ state observed in the $\Lambda_b^0 \to J/\psi p K^-$ decay at LHCb [2] as a peak in the

(*) E-mail: lukasz.bibrzycki@agh.edu.pl

 

$J/\psi p$ invariant mass distribution drew a lot of attention. If confirmed as a resonance, $P_c(4312)^+$ would necessarily consist of five valence quarks. The state's proximity to the $\Sigma_c^+\overline{D}^0$ threshold pointed towards hadronic molecule interpretation but other interpretations followed. In [3] a bottom-up approach was taken, where a general analytical structure of the unitary amplitude was exploited. The minimally biased model pointed towards interpretation of the $P_c(4312)$ as a virtual state. Here we also exploit the bottom-up approach but rather than constraining the model parameters with experimental data, we use an artificial neural network to map the structure of the amplitude poles onto physically interpretable classes.

## 2. – Physics basis

We consider a two-channel model with channels $J/\psi p$ and $\Sigma_c^+\overline{D}^0$ labelled 1 and 2, respectively. The starting point of the model construction is the coupled channel form of the inverse amplitude which reads

$$T_{ij}^{-1} = M_{ij} - ik_i\delta_{ij}, \tag{1}$$

where $k_i$ are channel momenta and $M_{ij}$ is a matrix which was shown in [4] to be singularity free in the energy region close to threshold, and can thus be Taylor expanded there. Furher we use the first term of this expansion, *i.e.*, a constant matrix $M_{ij}$. The channel momenta compatible with this approximation are defined as $k_i = \sqrt{s - s_i}$, where

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

are threshold energies in respective channels. The explicit form of the elastic amplitude for the $pJ/\psi$ channel reads

$$T_{11} = \frac{M_{22} - ik_2}{(M_{11} - ik_1)(M_{22} - ik_2) - M_{12}^2}, \tag{3}$$

where $M_{11}$, $M_{22}$ and $M_{12}$ are model parameters. It is assumed that the state under study has a well-definded spin and a contribution of other partial waves boils down to the incoherently added background to be parametrized as a first-order polynomial in energy squared $B(s) = b_0 + b_1 s$. In this study we are going to model the $J/\psi p$ invariant mass distribution observed in the $\Lambda_b^0 \to J/\psi p K^-$ decay. The $\Lambda_b^0$ decay vertex function similarly to the backgound can be parametrized as the first degree polynomial $P_1(s) = p_0 + p_1 s$. So, the differential energy-dependent intensity reads

$$\frac{\mathrm{d}N}{\mathrm{d}\sqrt{s}} = \rho(s)\left[|P_1(s)T_{11}(s)|^2 + B(s)\right], \tag{4}$$

with $\rho(s)$ being the phase space factor. In principle, eq. (4) contains also the term describing a transition from $\Sigma_c^+\overline{D}^0$ to $J/\psi p$ channel. Here, however, we are mostly concerned with the pole structure of the amplitude on the 4-sheeted Riemann surface, which for the $T_{21}$ is the same as for $T_{11}$. So, the effect of $T_{12}$ can be absorbed into $P_1(s)$ and $B(s)$ functions (see [3] for detailed discussion). To account for the finite mass resolution of the experiment, eq. (4) was convolved with the Gaussian with mass-dependent width, see Supplementary Material of [5] for details. By construction, the

amplitude of eq. (3) has 4 poles but only 2 of them, being complex conjugate, lie close to the physical region on either 2nd or 4th Riemann sheet and are relevant for the shape of the observed signal, see fig. 1. With decreasing $M_{12}$, *i.e.*, when two channels decouple, these poles approach the imaginary $k_2$-axis with positive value of $\mathrm{Im}k_2$ for the bound state and negative one for the virtual state. In the scattering length approximation this sign is controlled by the sign of $M_{22}$ parameter (positive for the virtual state, negative for the bound state). Thus our classification problem consists in mapping the intensity shape into 4 combinations of bound *vs.* virtual states on either 2nd or 4th sheet. In what follows we label these classes $b|2$, $b|4$, $v|2$ and $v|4$, respectively.

## 3. – Neural network model

The intensity function of eq. (4) depends on 7 parameters —$M_{11}$, $M_{22}$, $M_{12}$, $p_0$, $p_1$, $b_0$ and $b_1$. The training data set of $10^5$ samples was produced by randomly sampling the model parameters within suitably chosen intervals, see [6] for details on training data generation and ANN architecture. The artificial neural network (ANN) classifier has been trained by feeding model intensities split into 65 energy bins to the input layer and matching them with respective class assignments using the categorical cross-entropy as a loss function. To make the ANN model more robust against experimental uncertainties, some portion of the Gaussian noise was added to the simulated data. In fig. 2 we show how the classification accuracy depends on the level of noise (5% noise was used in actual prediction, which is our estimate of the evarage error percentage as releted to the maximum intensity) as well as the model accuracy on the validation data set, which was held out during training. Noteworthy, the accuracy exceeds 90% for all classes.

## 4. – Results

Here we discuss the results of the analysis applied to the $J/\psi p$ invariant mass distribution weighted with $\cos\theta_{P_c}$, *i.e.*, the value of the cosine of the angle between final state kaon and the $J/\psi$ as measured in the $P_c^+$ rest frame. See [6] for results obtained for other data sets. To obtain statistically significant predictions from the ANN classification model a distribution of class assignments has been produced in two ways. First, a

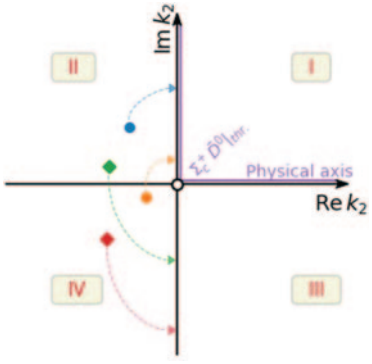


Fig. 1. – Pole structure in the complex $k_2$-plane. Poles marked with circles evolve to bound states while those marked with diamonds to virtual states with decreasing inter-channel coupling (figure taken from [6]).

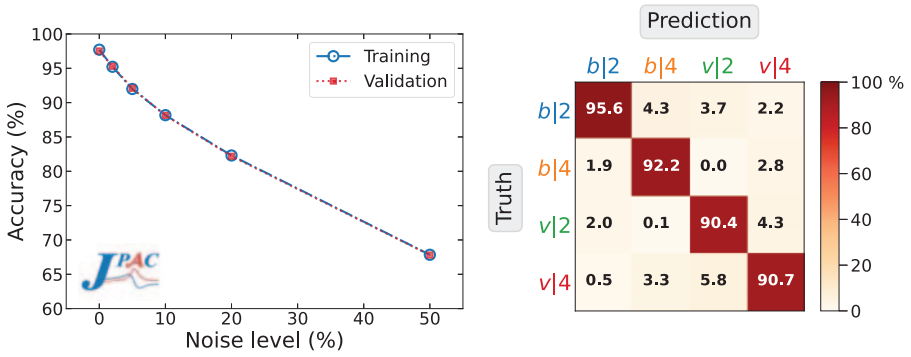


Fig. 2. – ANN classification accuracy as a function of the noise level (left panel) and model accuracy on the validation data set (right panel).

bootstrap method was used [7], where a set of $10^4$ examples was produced by sampling with Gaussian distribution around experimental central values of intensity with standard deviation equal to experimental error. This sample was fed to the trained model and produced the 4-class distribution. Secondly, a dropout method was used [8], where the Gaussian noise was simulated by randomly switching off some weights in the trained model. Histograms of class assignment probabilities based on two methods are shown in fig. 3. Both of them indicate that $v|4$, *i.e.*, a virtual state corresponding to the pole on the 4th Riemann sheet is by far a most probable class assignment. Finally the SHapley Additive exPlanations (SHAP) [9] were used to identify energy bins which are most important for the ANN model to choose a particular class. As shown in fig. 4 the energy bins close to the $\Sigma_c^+\overline{D}^0$ threshold are decisive for $v|4$ class assignment, which provides an *ex post* justification of the scattering length approximation of the amplitude. It is worth mentioning that this analysis can be extended in several directions like taking into account the interference with the background, microscopically motivated description of the production mechanism or reducing the bias related to limited intervals of parameters used for the training data set generation.

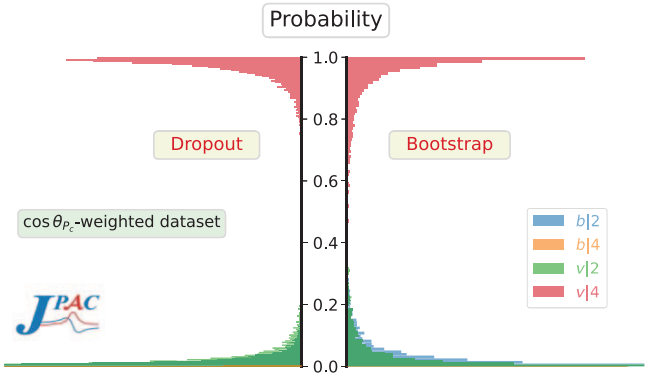


Fig. 3. – Histograms reflecting the class assignment probabilities obtained with dropout and bootstrap methods.

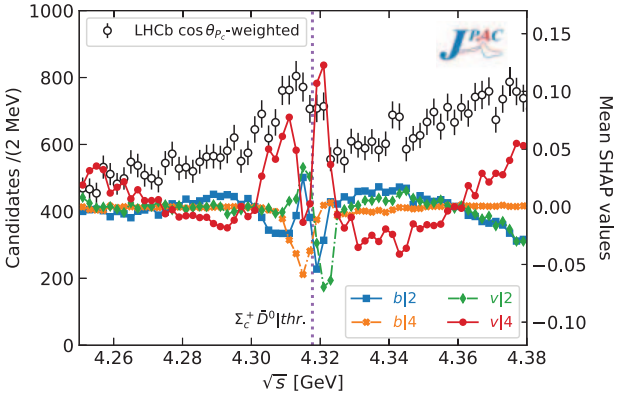


Fig. 4. – SHAP values for four possible class assignments and energy bins around $P_c(4312)$ maximum.

* * *

This work was supported by Polish Science Center Grant No. 2018/29/B/ST2/02576, and U.S. Department of Energy Grants Nos. DE-AC05-06OR23177, DE-FG02-87ER40365, and DE-FG02-92ER40735. AP has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 754496. CFR is supported by Spanish Ministerio de Educación y Formación Profesional under Grant No. BG20/00133. VM is a Serra Húnter fellow and acknowledges support from the Spanish national Grants No. PID2019–106080 GB-C21 and No. PID2020-118758GB-I00.

REFERENCES

[1] Choi S.-K. *et al.*, *Phys. Rev. Lett.*, **91** (2003) 262001.
[2] Aaij R. *et al.*, *Phys. Rev. Lett.*, **122** (2019) 222001.
[3] Fernández-Ramírez C., Pilloni A., Albaladejo M., Jackura A., Mathieu V., Mikhasenko M., Silva-Castro J. A. and Szczepaniak A. P., *Phys. Rev. Lett.*, **123** (2019) 092001.
[4] Frazer W. R. and Hendry A. W., *Phys. Rev.*, **134** (1964) B1307.
[5] Aaij R. *et al.*, *Phys. Rev. Lett.*, **115** (2015) 072001.
[6] Ng L., Bibrzycki L., Nys J., Fernández-Ramírez C., Pilloni A., Mathieu V., Rasmusson A. J. and Szczepaniak A. P., *Phys. Rev. D*, **105** (2022) L091501.
[7] Efron B. and Tibshirani R., *An Introduction to the Bootstrap*, in *Chapman & Hall/CRC Monographs on Statistics & Applied Probability* (Taylor & Francis) 1994.
[8] Gal Y. and Ghahramani Z., *Dropout as a Bayesian approximation: Representing model uncertainty in deep learning*, in *Proceedings of the 33rd International Conference on Machine Learning*, edited by Balcan M. F. and Weinberger K. Q., Vol. **48** of *Proceedings of Machine Learning Research* (New York, USA) 2016, pp. 1050–1059.
[9] Lundberg S. M. and Lee S.-I., *A unified approach to interpreting model predictions*, in *Advances in Neural Information Processing Systems*, edited by Guyon I., Luxburg U. V., Bengio S., Wallach H., Fergus R., Vishwanathan S. and Garnett R., Vol. **30** (Curran Associates, Inc.) 2017.