Deep learning the analytical structure of scattering amplitudes

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

For: JPAC and A(i)DAPT AI for Data Analysis and PreservaTion

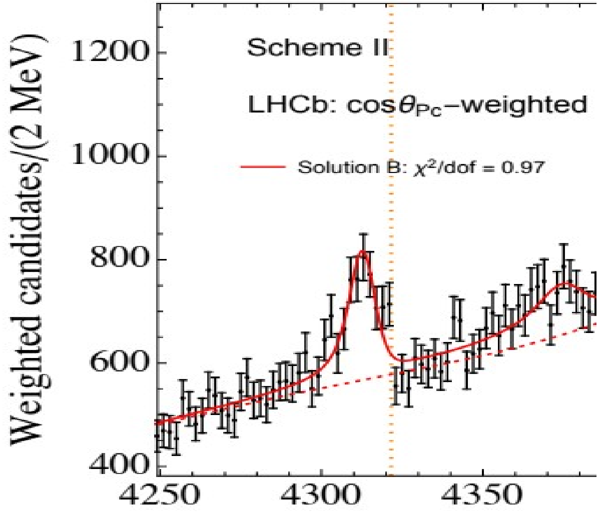
Part I

Deep learning the $P_c(4312)$ state

L. Ng, Ł. Bibrzycki, J. Nys, C. Fernandez-Ramirez, A. Pilloni, V. Mathieu,
A.J.Rasmusson, A.P. Szczepaniak

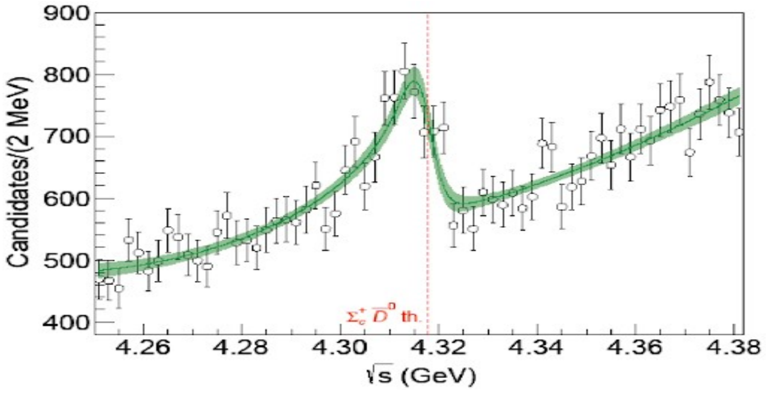
Objective

Determining the nature of the $P_c(4312)$ pentaquark candidate



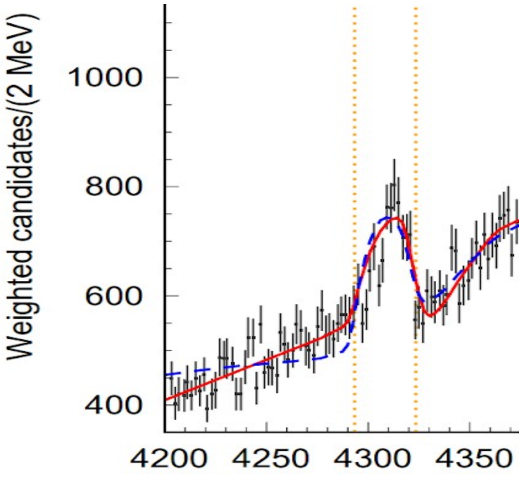
Molecule

Du et al., 2102.07159



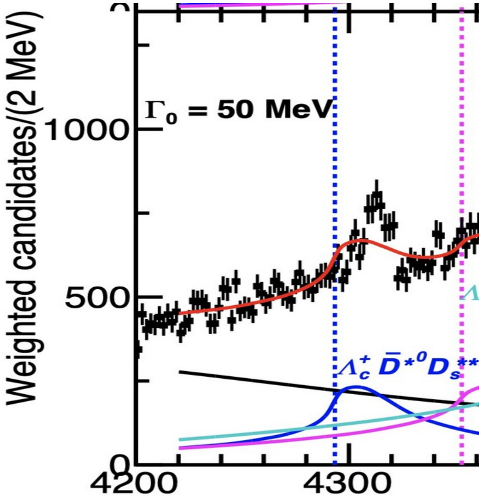
Virtual

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



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

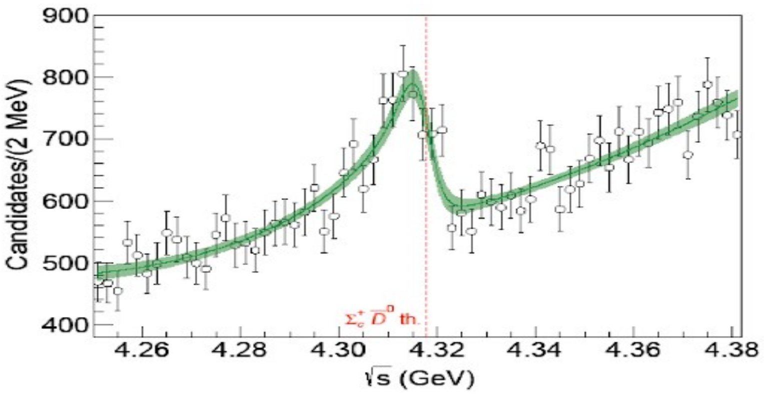
Nakamura, Phys. Rev. D 103, 111503 (2021)

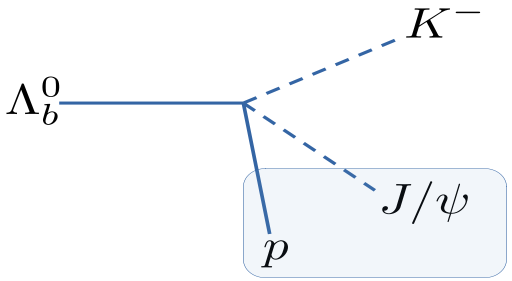


Single triangle (ruled out)

LHCb, Phys. Rev. Lett. 122, 222001 (2019)

Physics model





Experimental situation:

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

Physics model

- Starting point: Fixed partial wave (inverse) amplitude: $\hat{T}^{-1} = \hat{M}(s) - i\hat{p}$
- With *M* expandable in Taylor series (see *Frazer, Hendry, Phys. Rev. 134 (1964)*) one gets: $\hat{M} \approx \hat{m} - \hat{c}s$

Scattering length ($\hat{m}$)

Effective range ($\hat{c}$)

- Scattering amplitude for two coupled channels in scattering length approx (k_1, k_2 are channel momenta):

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

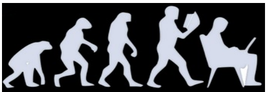
- Differential intensity:

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

where $\rho(s) = pqm_{\Lambda_b}$ is a phase space, with $p = \lambda^{\frac{1}{2}}(s, m_{\Lambda_b}^2, m_K^2)/2m_{\Lambda_b}$, $q = \lambda^{\frac{1}{2}}(s, m_p^2, m_\psi^2)/2\sqrt{s}$

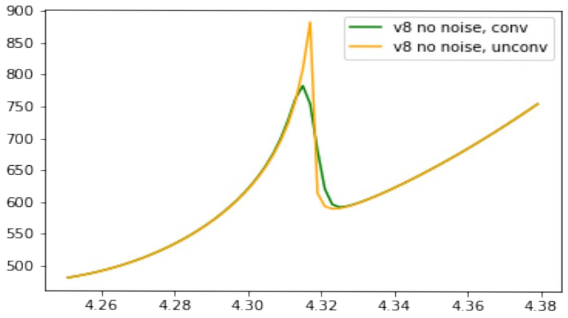
$P_1(s) = p_0 + p_1 s$ (production term) $\quad B(s) = b_0 + b_1 s$ (background term)

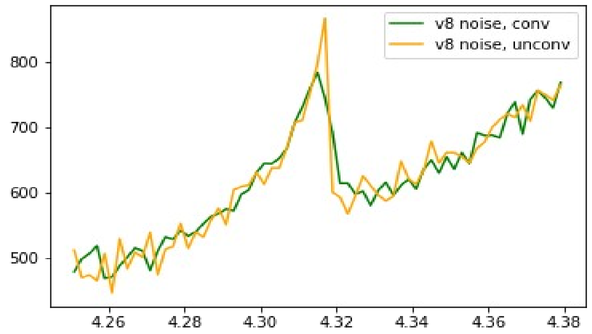
Altogether 7 model parameters: m_{11}, m_{22}, m_{12}, p_0, p_1, b_0, b_1.



ML model – training data - input

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





  - Sample intensities (computed in 65 energy bins) – produced with randomly chosen parameter samples – **examples**

  - 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]

  - For each parameter sample the **target class** was computed– one of the four:

b|2, b|4, v|2, v|4

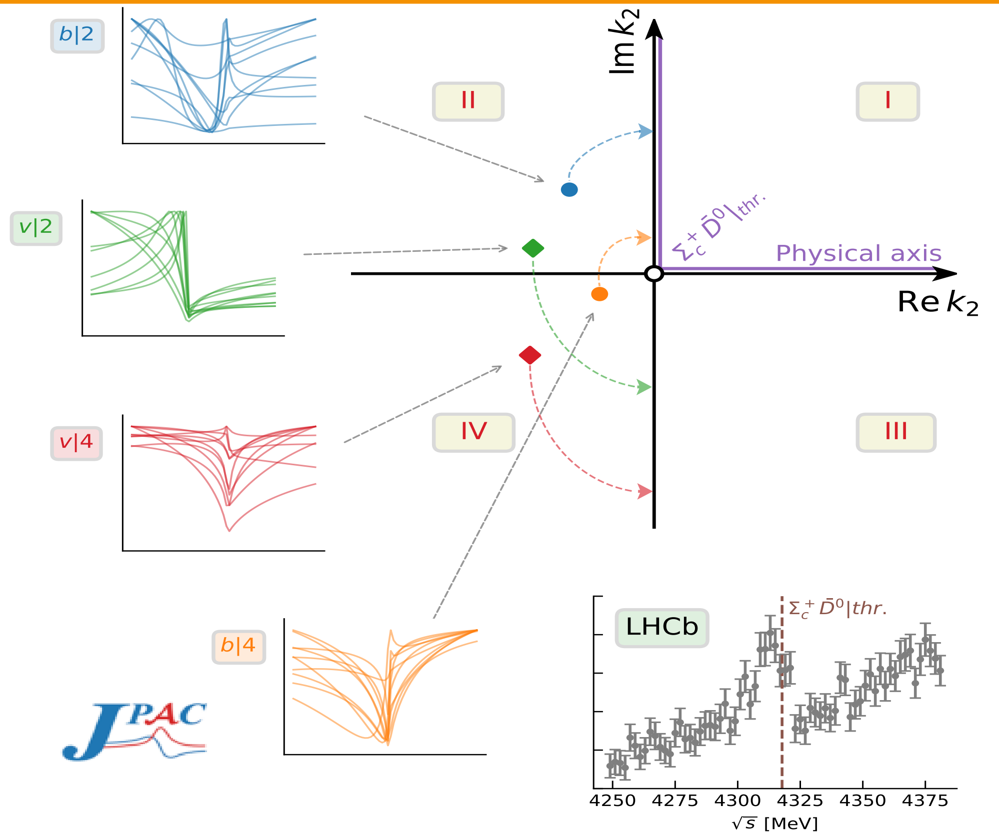
ML model - training data - targets

- Targets 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

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

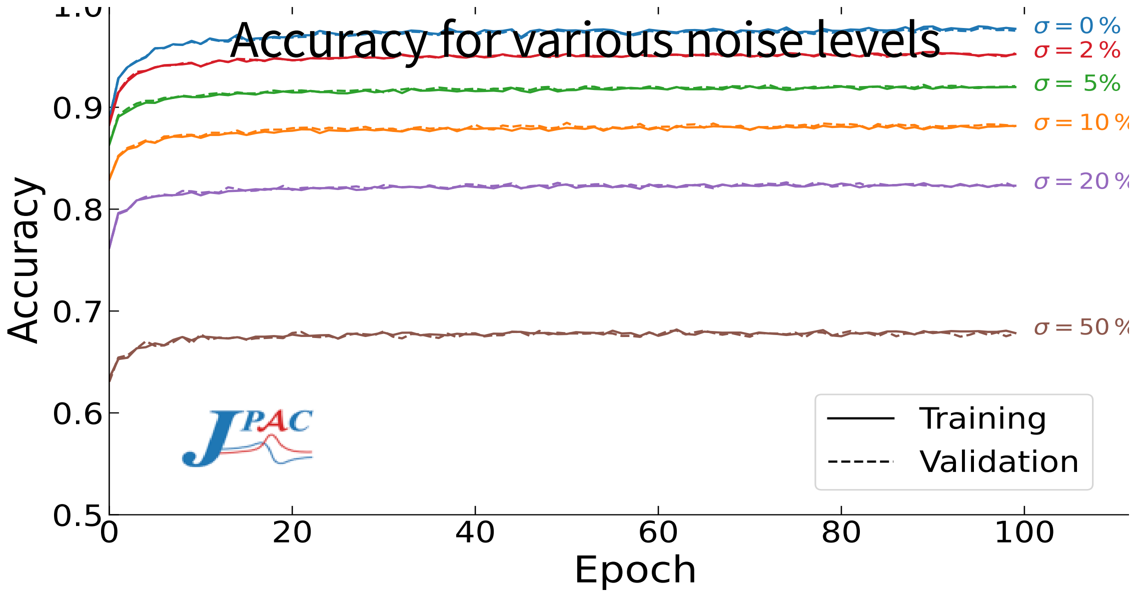


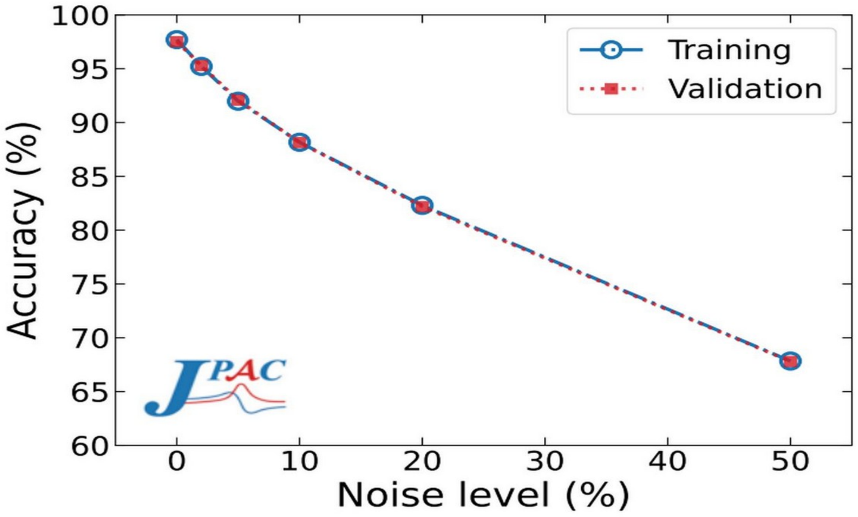
Then poles appear on sheets defined with (η_1,η_2) pairs: (-,+) - II sheet, (+,-) - IV sheet

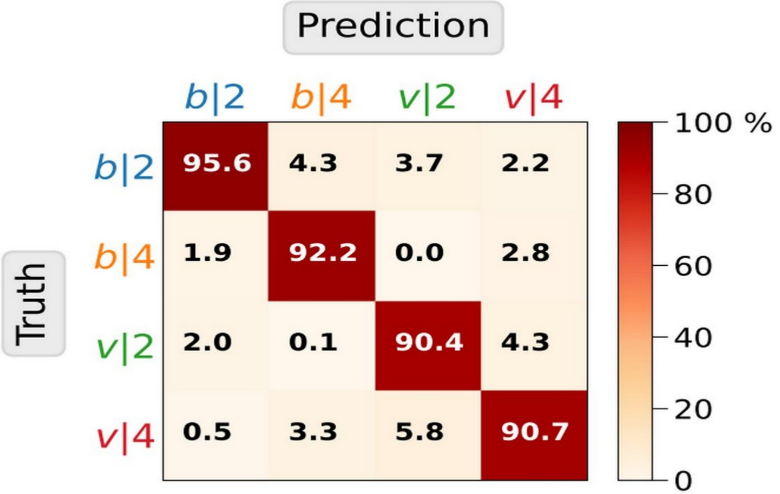
Symbolically: $K : \{[I_1, ..., I_{65}](m_{11}, m_{22}, m_{12}, p_0, p_1, b_0, b_1)\} \to \{b|2, b|4, v|2, v|4\}$

ML model – training results

Accuracy for various noise levels



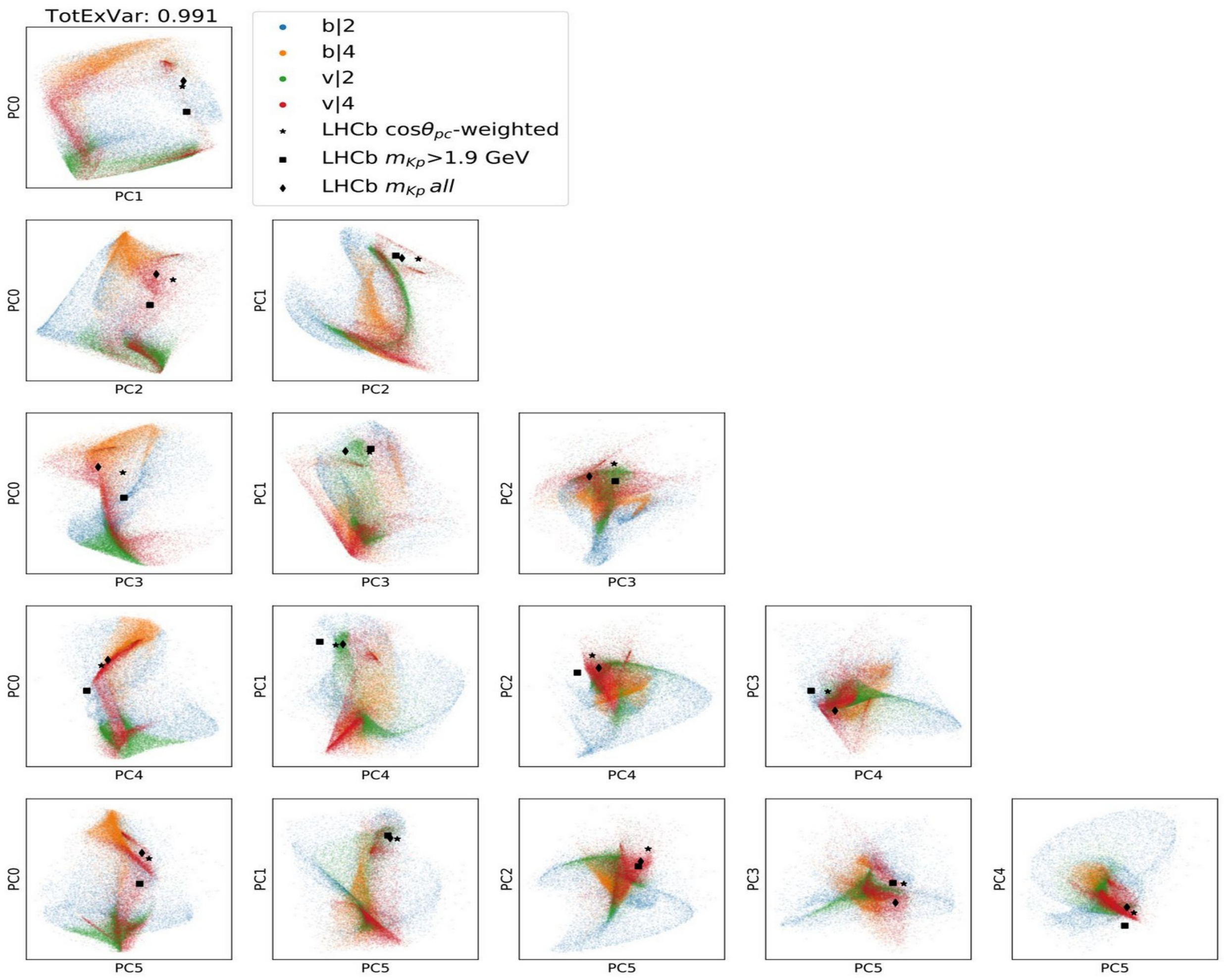




Confusion matrix for the 5% noise

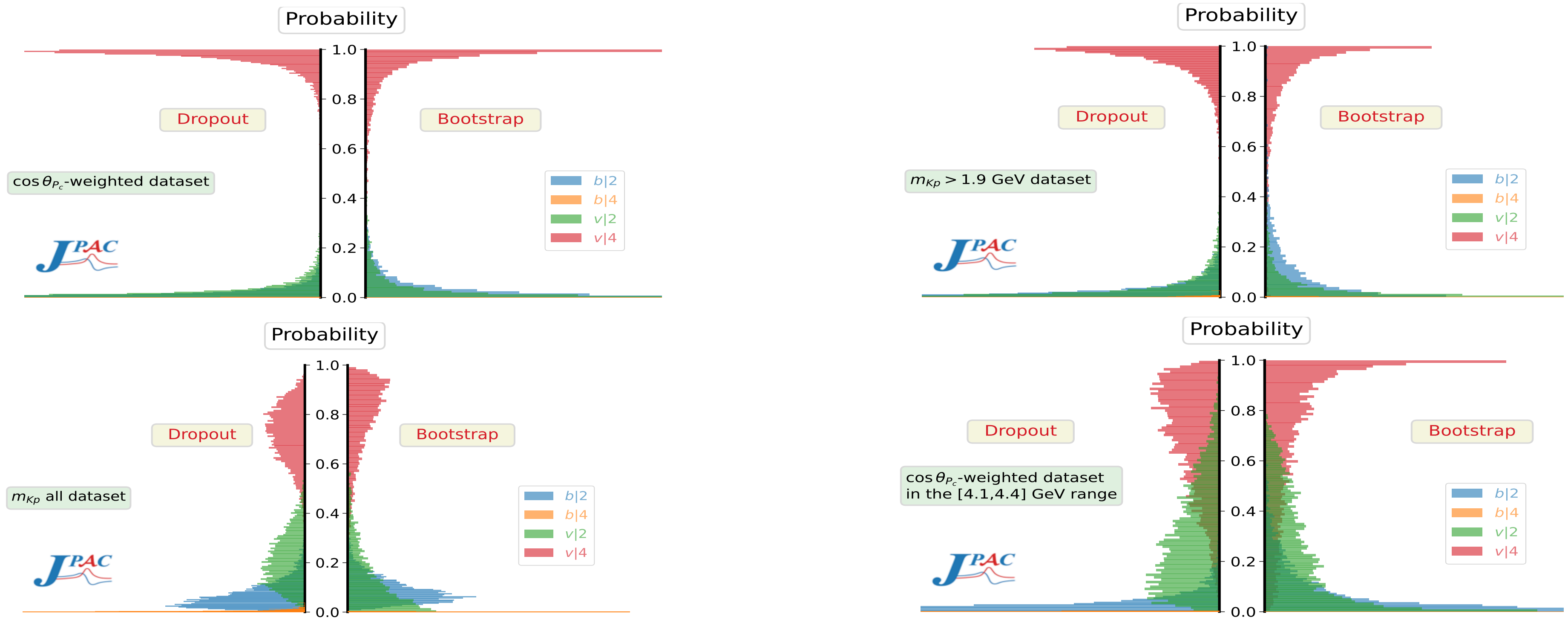
Does the training data set reflect experimental situation ?

- 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)



Summary I

- Takeaways:
  - Robust indication towards virtual state obtained.
  - 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 – various microscopical pictures can be attached to target classes – meta model.

Part II

Amplitude extraction with GANs

Glòria Montaña, A. Pilloni, Y. Li, L. Bibrzycki, M. Battaglieri
and others

Physics model

- Elastic scattering $\pi^+\pi^- \to \pi^+\pi^-$

$$A(s,\cos\theta) = \sum_{\ell=0}^{n}(2\ell+1)f_\ell(s)P_\ell(\cos\theta)$$

- Breit-Wigner type partial waves $f_\ell(s),\ \ell = 0, 1$

$$f_0(s) = \frac{m_\sigma \Gamma_\sigma}{m_\sigma^2 - s - i\Gamma_\sigma m_\sigma} \qquad m_\sigma \approx 0.475 \text{ GeV} \ , \ \Gamma_\sigma \approx 0.55 \text{ GeV}$$

$$f_1(s) = \frac{m_\rho \Gamma_\rho}{m_\rho^2 - s - i\Gamma_\rho m_\rho} \qquad m_\rho = 0.775 \text{ GeV} \ , \ \Gamma_\rho = 0.147 \text{ GeV}$$

$$\longrightarrow \quad A(s,\cos\theta) = f_0(s) + 3f_1(s)\cos\theta$$

- Differential cross section

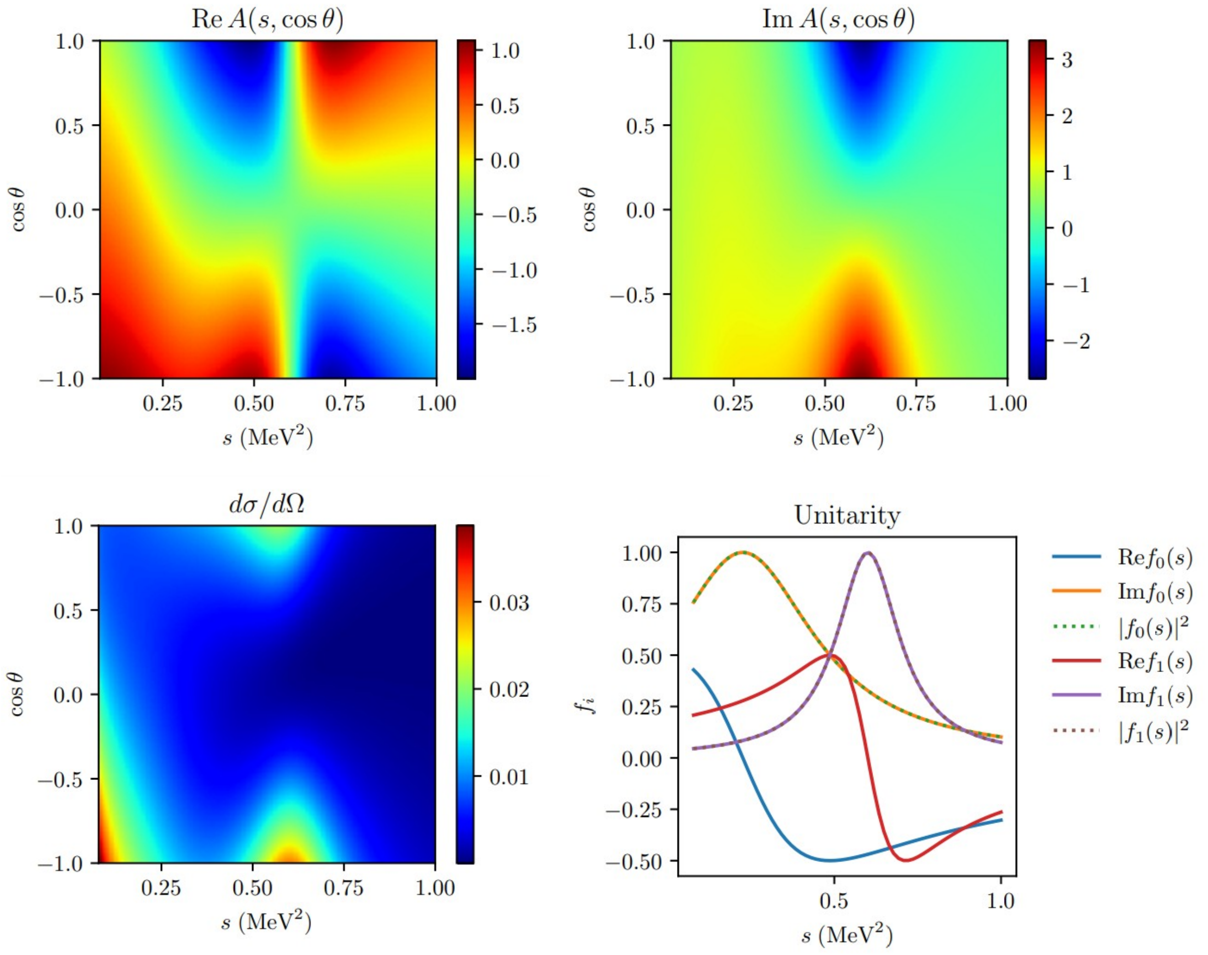
$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2}\frac{1}{s}|A(s,\theta)|^2$$

- Physics constraint: Unitarity of the partial waves

$$\text{Im} f_0(s) = |f_0(s)|^2$$

$$\text{Im} f_1(s) = |f_1(s)|^2$$

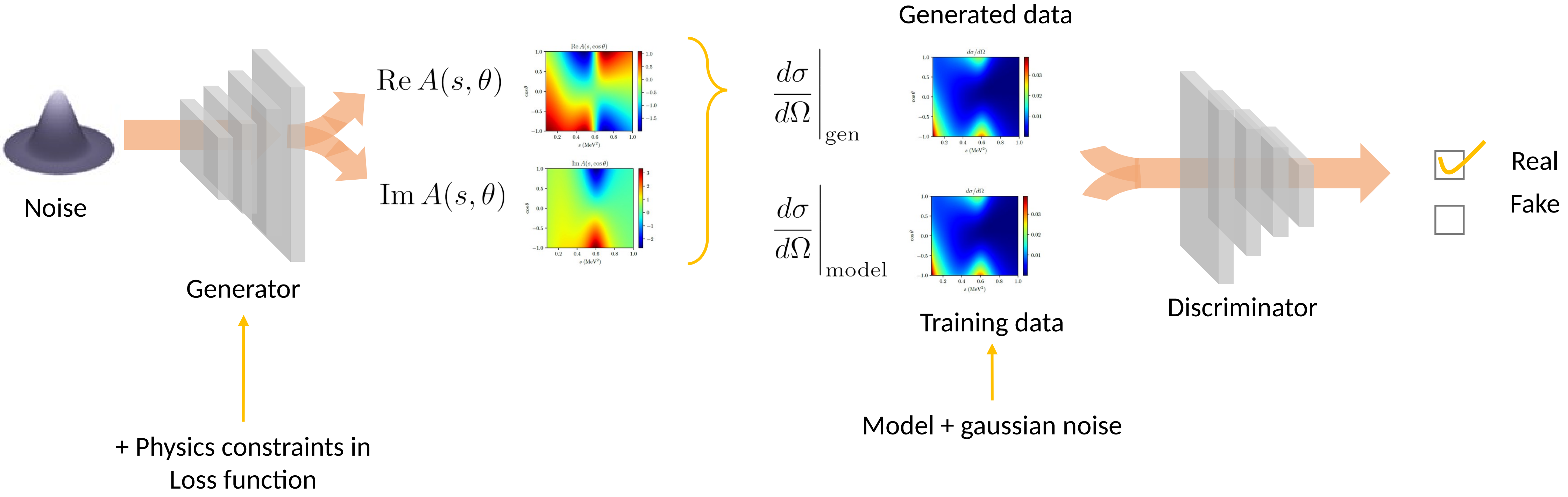
$$f_l(s) = \frac{1}{2}\int_{-1}^{+1} d(\cos\theta)P_\ell(\cos\theta)A(s,\theta)$$



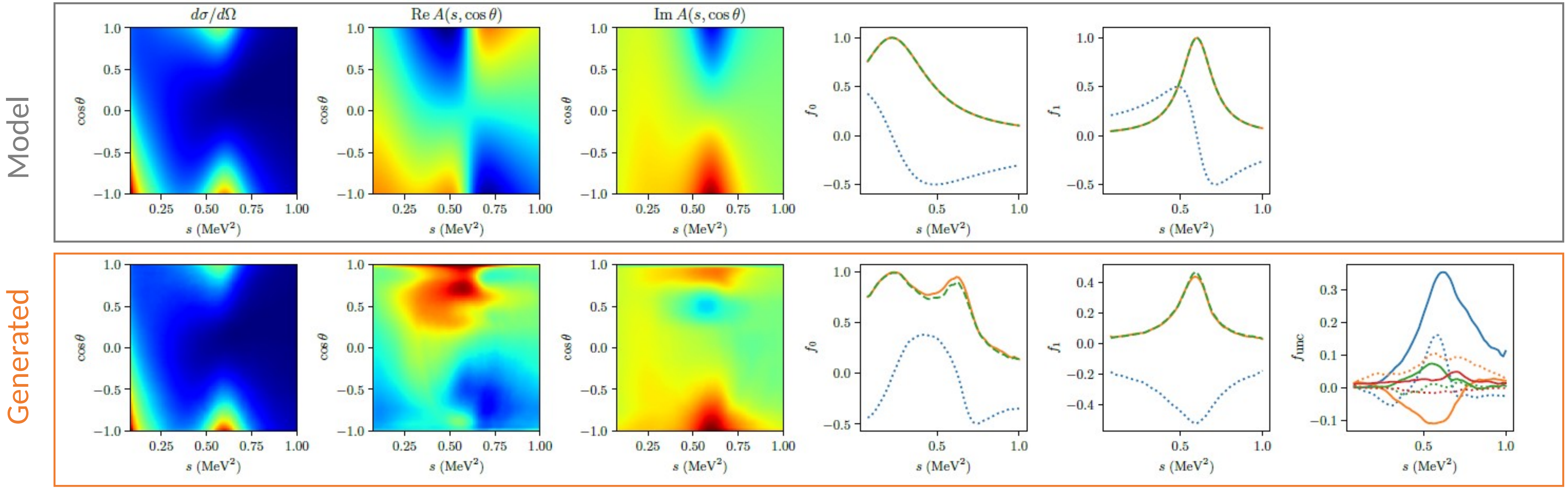
Generative Adversarial Network (GAN) with constraints

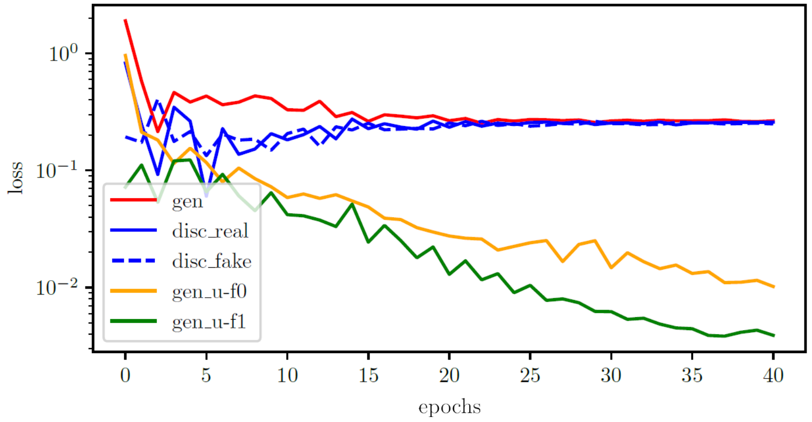
Two neural networks, the **generator** and the **discriminator**:

- The **generator** needs to capture the data distribution
- The **discriminator** estimates the probability that a sample comes from the training data rather than from the generator



Preliminary results (i)





- Cross section is reproduced qualitatively
- Unitarity constraint is satisfied
- Partial waves $\ell \geq 2$ are large

More physics constraints

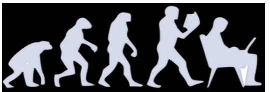
- Unitarity of the partial waves $f_\ell(s),\ \ell = 0, 1$

$$\mathrm{Im} f_0(s) = |f_0(s)|^2$$

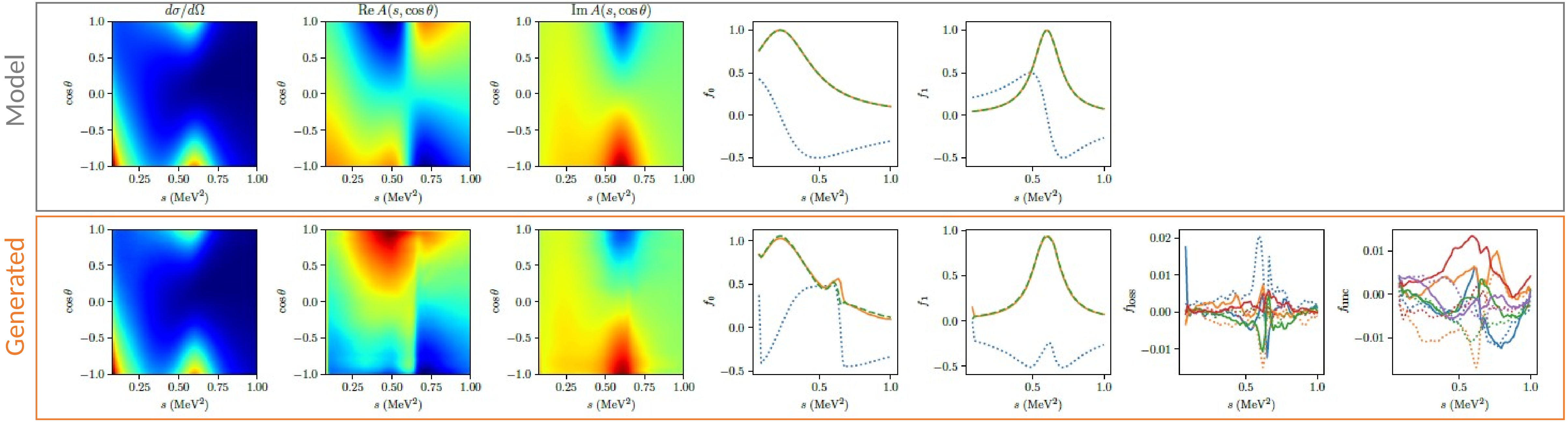
$$\mathrm{Im} f_1(s) = |f_1(s)|^2$$

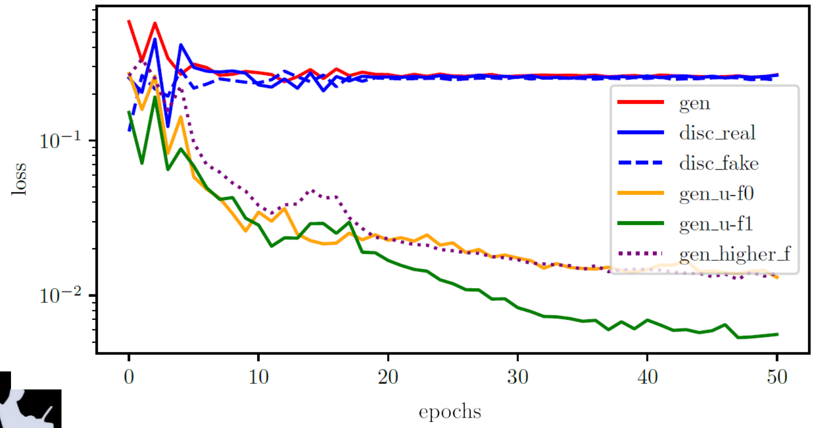
- Suppression of higher partial waves

$$f_\ell(s) = 0,\ \ \ell \geq 2$$

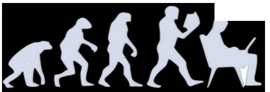


Preliminary results (ii)





- Cross section is reproduced qualitatively
- Unitarity constraint is satisfied
- Partial waves $\ell \geq 2$ are suppressed
- Ambiguity in the sign of the real part



Even more physics constraints

- Unitarity of the partial waves $f_\ell(s),\ \ell = 0, 1$

$$\mathrm{Im} f_0(s) = |f_0(s)|^2$$

$$\mathrm{Im} f_1(s) = |f_1(s)|^2$$

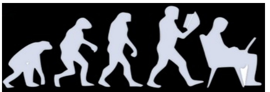
- Suppression of higher partial waves

$$f_\ell(s) = 0,\ \ \ell \geq 2$$

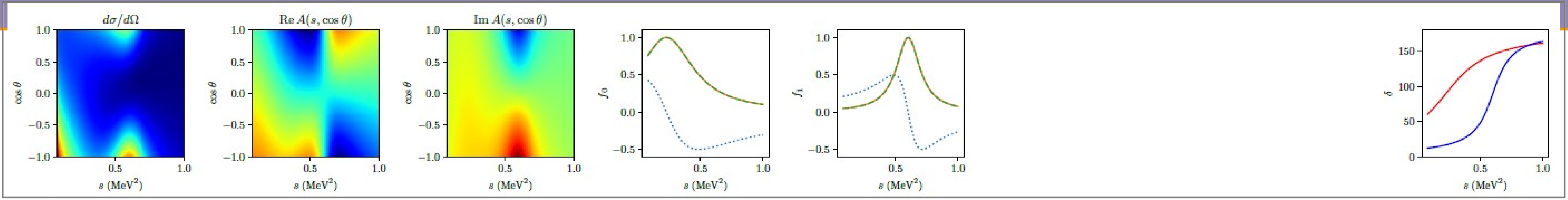
- Positive derivative of the phase shift $\delta_\ell(s) = \mathrm{atan}\left(\frac{\mathrm{Im} f_\ell(s)}{\mathrm{Re} f_\ell(s)}\right)$

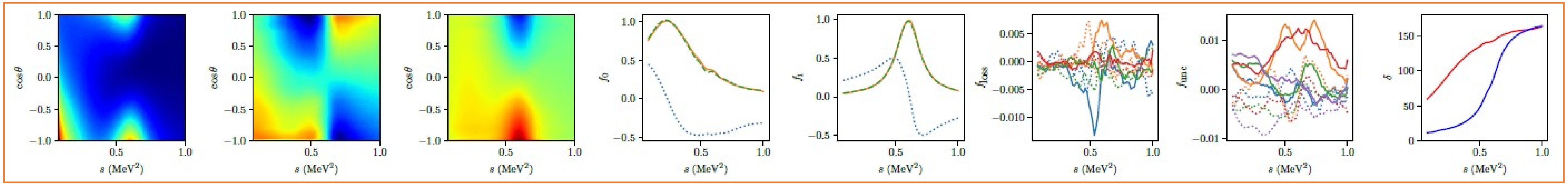
$$\frac{\mathrm{d}}{\mathrm{d}s}\ \delta_0(s) \geq 0$$

$$\frac{\mathrm{d}}{\mathrm{d}s}\ \delta_1(s) \geq 0$$

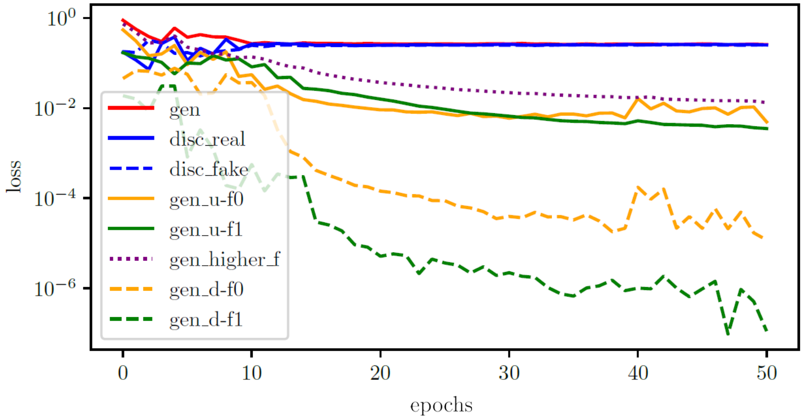


Preliminary results (iii)

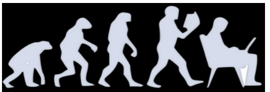




Generated

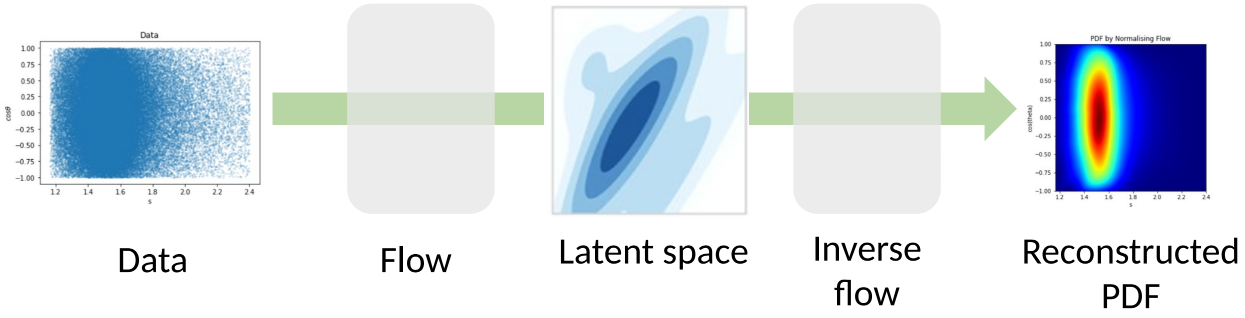


- Cross section is reproduced qualitatively
- Unitarity constraint is satisfied
- Partial waves $\ell \geq 2$ are suppressed
- The real part takes the right sign

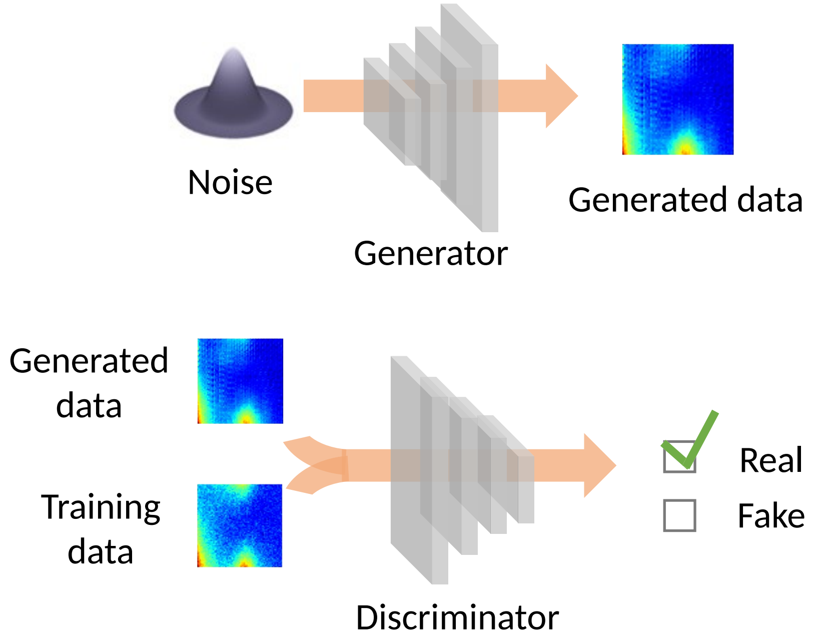


Events → Cross section → Amplitude

A. **Normalizing Flows:**
extract differential cross section (∝ Probability Density) from events distribution



B. **Generative Adversarial Networks (GANs):**
extract amplitude from differential cross sections, using unitarity



Summary II

Preliminary status, but the results of using GANs to extract amplitudes from cross sections employing physics constraints are promising.

Next steps:

- Increase gaussian noise of the training pseudodata set (currently 0.1%)
- Adjust the generator and/or discriminator models and hyperparameters for convergence
- Determine quantitative agreement between generated and model
- Extension to the event level using normalizing flows
- Extension to more complicated processes
- Generalization of the physics constraints