

Model Selection for Pion Photoproduction

R. Molina*

*Instituto de Física, Universidade de São Paulo, Rua do Matão, 1371 Butantã, C.P. 05508-090
São Paulo, Brazil*
ramope@if.usp.br

J. Landay, B. Hu

The George Washington University, Washington, DC 20052, USA
jlanday@gwmail.gwu.edu, binhu@gwmail.gwu.edu

M. Döring

The George Washington University, Washington, DC 20052, USA
Theory Center, Thomas Jefferson National Accelerator Facility, Newport News, VA 23606, USA
doring@gwu.edu

C. Fernández-Ramírez

*Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, Ciudad de México
04510, Mexico*
cesar.fernandez@nucleares.unam.mx

In a precision era of hadron spectroscopy, new tools are required for the analysis of data from hadron reactions. In this talk, we show an analysis of low-energy neutral pion photoproduction data using the Least Absolute Shrinkage and Selection Operator (LASSO) in combination with criteria from information theory and K -fold cross validation. These analysis techniques will become relevant in the near future. First, we illustrate these methods with synthetic data; then, the latest available measurements of differential cross sections ($d\sigma/d\Omega$), photon-beam asymmetries (Σ), and target asymmetry differential cross sections ($d\sigma_T/d \equiv T d\sigma/d\Omega$) in the low-energy regime, are analyzed, and its feasibility for real data is demonstrated.

*XVII International Conference on Hadron Spectroscopy and Structure - Hadron2017
25-29 September, 2017
University of Salamanca, Salamanca, Spain*

*Speaker.

1. Introduction

Partial-wave analysis of hadronic reactions is a prerequisite for many theoretical approaches to access information from the experimental data, especially if the comparison to hadron resonances is the goal. Narrowing the focus to photoproduction reactions, their decomposition into partial waves (multipoles) is usually performed through an energy-dependent (ED) parametrization of the amplitude. As long as data are not abundant and precise enough, it is not yet possible to perform a (truncated partial-wave) complete experiment [1, 2] in the resonance region. A parametrization in energy is needed for the determination of resonances, or as a stabilizing starting point for single-energy (SE) solutions in which energy-binned data are fitted independently. The selection of fit parameters is a fundamental problem that we address in this study for the case of neutral pion photoproduction in the low energy region.

The parametrization of the amplitude in pion photoproduction is commonly chosen according to the considered energy range. Heavy Baryon Chiral Perturbation Theory (HBChPT) [3, 4] and Relativistic Baryon Chiral Perturbation Theory (RBChPT) [5] provide effective parametrizations for low-energy neutral pion photoproduction up to $E_\gamma \simeq 170$ MeV [6, 7]. The inclusion of the $\Delta(1232)$ resonance as an explicit degree of freedom in RBChPT allows to extend the agreement between theory and data up to $E_\gamma \simeq 200$ MeV [8]. At intermediate energies, effective field theory approaches, K -matrix parametrizations, or related approaches are used [9–12]. For higher energy, Regge parametrizations are effective [13, 14]. Partial waves can be also parameterized purely phenomenologically in terms of functions that are in agreement with basic S -matrix principles such as coupled-channel two-body unitarity, the correct threshold behavior or Fermi-Watson's theorem [15, 16], but that are otherwise left free to ensure a high degree of model independence as in the SAID approach [12].

Yet, several problems can appear in the analyses. If the amplitude is under-parameterized, the quality of the data description is not satisfactory and the quality of the extracted amplitudes is difficult to assess. Over-parametrization can result in limited predictability of the amplitude outside the fitted data range and inflated uncertainties. Furthermore, problems in the data themselves (incompatibility of data, systematics, or even statistics) may be interpreted as significant physics in over-parameterized fits. In many approaches resonances are introduced in the parametrization as explicit terms, that will unavoidably improve the fit quality at the cost of potentially false positive resonance signals [17]. The Least Absolute Shrinkage and Selection Operator (LASSO) [18–20] provides a tool to scan a plethora of different models, in particular multiple combinations of different resonances. Manually, such a scan would be impossible due to the large number of combinations, but the LASSO provides an automatized, “blind-folded” technique [21].

Having the above-mentioned extensions for future work in mind, we concentrate in this study on the question of how to select the simplest amplitude for a real-life example of photoproduction reactions.

1.1 Parametrization

An energy-dependent parametrization is formulated for both real and imaginary parts of the

three P waves, as well as for the real parts of E_{0+} and the four D -wave multipoles,

$$\text{Re, Im } \mathcal{M}_{L\pm} = \frac{q_{\pi^0}^\ell}{m_{\pi^+}^{\ell+1}} \sum_{i=0}^{i_{\max}} \frac{a_i}{10^{-i}} \left(\frac{\omega_{\pi^0} - m_{\pi^0}}{m_{\pi^+}} \right)^i \quad (1.1)$$

where q_{π^0} is the center-of-mass momentum of the neutral pion, $\omega_{\pi^0}^2 = m_{\pi^0}^2 + q_{\pi^0}^2$, and a_i are the fit parameters. This expansion is theoretically justified by ChPT calculations [3]. The quantity \mathcal{M} stands for the electric ($E_{L\pm}$) and magnetic ($M_{L\pm}$) multipoles, or alternatively, the partial waves P_1 , P_2 and P_3 for the P waves, related to each other by $E_{1+} = \frac{1}{6}(P_1 + P_2)$, $M_{1+} = \frac{1}{6}(P_1 - P_2 + 2P_3)$ and $M_{1-} = \frac{1}{3}(P_3 + P_2 - P_1)$. In Eq. (1.1) for the real parts of the multipoles, $\ell = L$, while for the imaginary parts of the P -wave multipoles, $\ell = 3L + 1 = 4$. Fit parameters are called a_i throughout, omitting the indices specifying to which partial wave they belong. For the real parts of the P -waves and D -waves, $i_{\max} = 4$ while for the imaginary parts of the P -wave, $i_{\max} = 0$. The imaginary parts of the D -waves are set to zero, fixing the overall-phase problem. For the S -wave multipole E_{0+} , a real-valued term of the form of Eq. (1.1) with $i_{\max} = 4$ is supplemented by a term of the form

$$\Delta E_{0+} = i \frac{q_{\pi^+}}{m_{\pi^+}^2} \sum_{i=0}^{i=2} \frac{a_i}{10^{-i}} \left(\frac{q_{\pi^+}}{m_{\pi^+}} \right)^{2i} \quad (1.2)$$

to take into account the π^+n threshold cusp. The term provides an imaginary part above the π^+n threshold and contributes to the real part of E_{0+} below it. In total, there are $i_{\max} = 46$ free fit parameters.

1.2 LASSO, cross validation and criteria from information theory

The χ_T^2 is defined in the LASSO method to penalize models with a large number of parameters, and thus select the *simplest* model [18–20], as follows:

$$\chi_T^2(\lambda) = \chi^2(\lambda) + P(\lambda), \quad \text{with} \quad P(\lambda) = \lambda^4 \sum_{i=1}^{i_{\max}} |a_i|. \quad (1.3)$$

The question is now how to choose the best value for λ , i. e., more accurately describes the model without overfitting the data. One of the tools which can be used is *cross validation* [18–20]. There, data are randomly divided into a training set and a validation set (usually one of five partitions). For a given λ , the penalized χ_T^2 of the training set is minimized and the χ_V^2 of the validation set is determined. For very large λ , χ_V^2 is large, because the data are underfitted. While as λ decreases a point is reached below which the data are overfitted, i.e., non-physical structures such as statistical fluctuations in the training set are described. The validation χ_V^2 becomes thus larger again. The minimum in χ_V^2 is then regarded as the point where the fit optimally describes the data without describing fluctuations. One can then further constrain the search for the simplest model by selecting the $\lambda_{\text{opt}} > \lambda_{\text{min}}$ that is compatible within errors with the minimum of χ_V^2 at $\lambda = \lambda_{\text{min}}$. This is referred to as the 1- σ rule [22].

To select the optimal λ one can also use *criteria from information theory*. The three criteria that we use to find an optimal λ are the Akaike Information Criterion (AIC) [23,24], a finite sample size corrected version of the AIC (AICc) [25], and the Bayesian Information Criterion (BIC) [26]. For all three of the criteria, the optimal value of λ is given by the respective minimum [19,20].

2. Analysis of low-energy neutral pion photoproduction data

2.1 LASSO in a Benchmark Model

We first generate synthetic data from a benchmark model (\mathcal{B} -model), in which all the a_i in Eqs. (1.1) and (1.2) are set to zero except for: a_0 and a_1 for the real parts of every S and P wave in Eq. (1.1) and a_0 in Eq. (1.2), totaling 9 parameters. All the imaginary parts of the P waves are consequently set to zero. No D waves are included. These synthetic data are then analyzed with the full 46-parameter model as defined in the previous section, minimizing the penalized $\chi_T^2 = \chi^2 + P$ for different λ according to Eq. (1.3). Using the Lasso method, in combination with cross validation or information criteria leads to a simpler model with 10 parameters, remarkably close to the true model. All parameters related with D waves turn out to be zero, as it should be. The results of this fit are described in detail in Ref. [27].

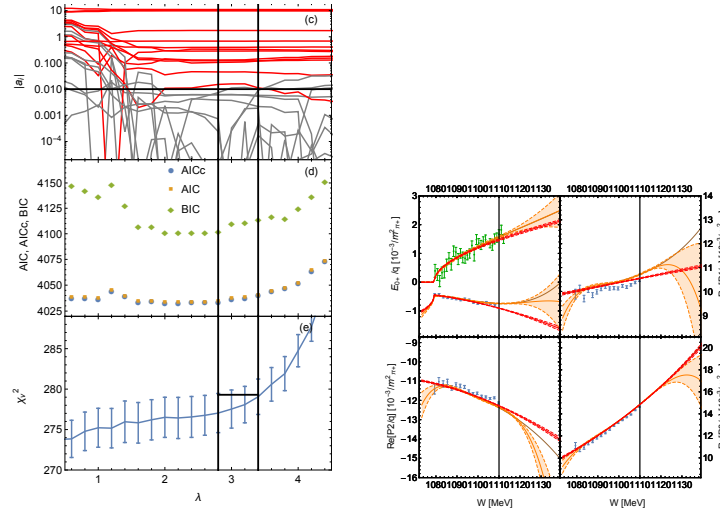


Figure 1: Left: in the top figure, the absolute value of the parameters a_i as function of λ in a logarithmic scale, are shown. The red lines indicate the finally chosen parameters, the gray lines show the unnecessary parameters. Middle, the AIC, AICc and BIC criteria, and, in the bottom, the cross-validation χ_V^2 for the analysis of real data. Right: S and P -wave partial waves from the fits to real data. The orange lines and bands show the 46-parameter fit with zero penalty $\lambda = 0$. The brown lines show the fit at optimal λ . Removing the parameters that are effectively zero and refitting produces the result with red bands. The vertical lines at $W = 1120$ MeV indicate the upper limit of fitted data. The blue circles show the single-energy solution of [6] for the real parts of S - and P -wave partial waves. The green triangles indicate the solution from Ref. [7].

2.2 LASSO for real data

We analyze here the data on $d\sigma/d\Omega$, Σ , and $d\sigma_T/d\Omega$ for the reaction $\gamma p \rightarrow \pi^0 p$ from Refs. [6, 7] for energies from the $\pi^0 p$ threshold up to $W = 1120$ MeV. Here, we set all imaginary parts of the P waves to zero. For the D waves themselves, we have kept them fixed at the real values given by the Born terms of photoproduction [6]. In total, the number of available parameters is 23. Performing the LASSO scan in combination with the information theory criteria and cross validation, we obtain the results shown in Fig. 1. The effectively non-zero parameters are shown in red. The horizontal line indicates the cut-off below which a parameter is counted as zero. It

is remarkable to observe the parameters drop by three orders of magnitude because it means that LASSO is also capable of disentangling the extreme correlations between parameters present at $\lambda = 0$, as the unnaturally large parameter values indicate. The BIC delivers the most pronounced minimum, or minimal plateau. We choose an optimal $\lambda = 2.8$ which coincides with the minimum of the BIC which is also the last point of the plateau. For the cross validation χ_V^2 there is no minimum at all in this case (as there was for \mathcal{B} -model [27]). Only an upper bound for λ can be determined continuing the upper end of the error bar at $\lambda = 2.8$ to the right as indicated. This leads to a maximal value of $\lambda = 3.4$ that is compatible within errors. Overall, the number of model parameters is reduced from 23 to 13 (in the \mathcal{B} -model: from 46 to 10).

The S and P -wave partial waves solutions are shown in Fig. 1 (right). The fit at $\lambda = 2.8$ is shown with the brown lines in Fig. 1 (right). Removing all parameters that are effectively zero and refitting the remaining 13 parameters, the final solution is indicated with the red lines and uncertainty bands. The orange lines and bands indicate the unconstrained 23-parameter fit with zero penalty $\lambda = 0$. The uncertainty bands of the 23-parameter fit are wider than the ones of the 13-parameter simplest model. We also observe largely widened error bands for the unconstrained fit at very low energies and beyond the fitted region (vertical bands) indicating the reduced predictability of the $\lambda = 0$ fit. The simplest model (red lines and bands) shows good qualitative agreement with the SE solution from [6] for the real parts of the partial waves. The imaginary part of E_{0+} of the simplest model agrees well with the SE extraction performed in [7] (green data points) as Fig. 1 (right) shows. For the cusp parameter at threshold we obtain $\beta_0 = (2.41 \pm 0.05) \cdot 10^{-3} m_{\pi^+}^{-1}$. This agrees with the value from [7]: $\beta_0 = (2.2 \pm 0.2[\text{stat.}] \pm 0.6[\text{syst.}]) \cdot 10^{-3} m_{\pi^+}^{-1}$. Note, however, the very different size of the statistical uncertainty.

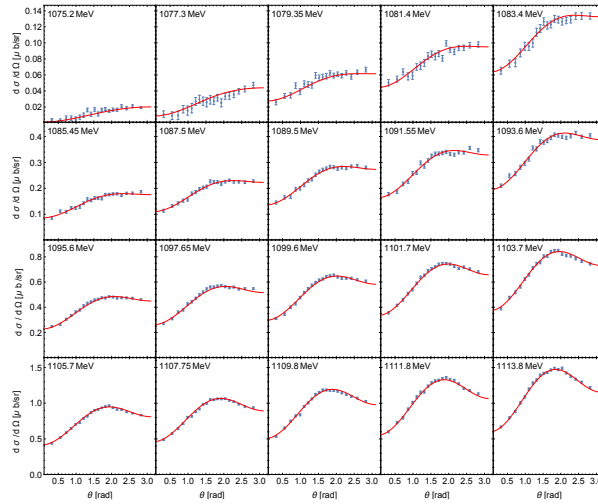


Figure 2: Differential cross section. Data from Ref. [6]. The labels in the plot indicate the scattering energy W in MeV. The red lines show the central value of the simplest model as determined in this study.

In Fig. 2, the best fit for the differential cross section data from Ref. [6] is shown, corresponding to the 13 parameter fit indicated with the red lines in Fig. 1 (right). The description of the data is consistently good. Data on polarized differential cross section $\sigma_T = d\sigma/d\Omega T$ from

Ref. [7] and beam asymmetry from Ref. [6] were also analyzed in Ref. [27]. The description of the data with the LASSO scan was remarkably good as well.

References

- [1] Y. Wunderlich, R. Beck, and L. Tiator, Phys. Rev. C **89**, 055203 (2014).
- [2] R. L. Workman, L. Tiator, Y. Wunderlich, M. Döring, and H. Haberzettl,
- [3] V. Bernard, N. Kaiser, and U.-G. Meißner, Phys. Lett. B **382**, 19 (1996). V. Bernard, B. Kubis, and U.-G. Meißner, Eur. Phys. J. A **25**, 419 (2005)
- [4] C. Fernández-Ramírez and A. M. Bernstein, Phys. Lett. B **724**, 253 (2013)
- [5] M. Hilt, S. Scherer, and L. Tiator, Phys. Rev. C **87**, 045204 (2013)
- [6] D. Hornidge *et al.* [A2 and CB-TAPS Collaborations], Phys. Rev. Lett. **111**, 062004 (2013)
- [7] S. Schumann *et al.* [A2 Collaboration], Phys. Lett. B **750**, 252 (2015).
- [8] A. N. Hiller Blin, T. Ledwig, and M. J. Vicente Vacas, Phys. Rev. D **93**, 094018 (2016)
- [9] D. Rönchen, M. Döring, F. Huang, H. Haberzettl, J. Haidenbauer, C. Hanhart, S. Krewald, U.-G. Meißner, and K. Nakayama, Eur. Phys. J. A **50**, 101 (2014); Erratum: [Eur. Phys. J. A **51**, 63 (2015)]
- [10] H. Kamano, S. X. Nakamura, T.-S. H. Lee, and T. Sato, Phys. Rev. C **88**, 035209 (2013)
- [11] C. Fernández-Ramírez, E. Moya de Guerra, and J. M. Udías, Phys. Lett. B **660**, 188 (2008)
- [12] R. L. Workman, M. W. Paris, W. J. Briscoe, and I. I. Strakovsky, Phys. Rev. C **86**, 015202 (2012)
- [13] A. Sibirtsev, J. Haidenbauer, F. Huang, S. Krewald, and U.-G. Meißner, Eur. Phys. J. A **40**, 65 (2009)
- [14] V. Mathieu, G. Fox, and A. P. Szczepaniak, Phys. Rev. D **92**, 074013 (2015)
- [15] K. M. Watson, Phys. Rev. **95**, 228 (1954).
- [16] E. Fermi, Suppl. Nuovo Cimento **2**, 17 (1955).
- [17] S. Ceci, M. Döring, C. Hanhart, S. Krewald, U.-G. Meißner and A. Švarc, Phys. Rev. C **84**, 015205 (2011)
- [18] R. Tibshirani, J. R. Stat. Soc. B **58**, 267 (1996)
- [19] T. Hasti, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer-Verlag, New York, USA, 2009).
- [20] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning* (Springer-Verlag, New York, USA, 2013).
- [21] B. Guegan, J. Hardin, J. Stevens, and M. Williams, JINST **10**, P09002 (2015)
- [22] L. Breiman and P. Spector, Int. Stat. Rev. **60**, 291 (1992).
- [23] H. Akaike, IEEE Trans. Autom. Control **19**, 716 (1974).
- [24] K. Burnham and D. Anderson, *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (Springer-Verlag, New York, 2003).
- [25] J. E. Cavanaugh, Statist. Probab. Lett. **31**, 201 (1997).
- [26] G. Schwarz, Ann. Stat. **6**, 461 (1978).
- [27] J. Landay, M. Döring, C. Fernández-Ramírez, B. Hu and R. Molina, Phys. Rev. C **95**, no. 1, 015203 (2017)