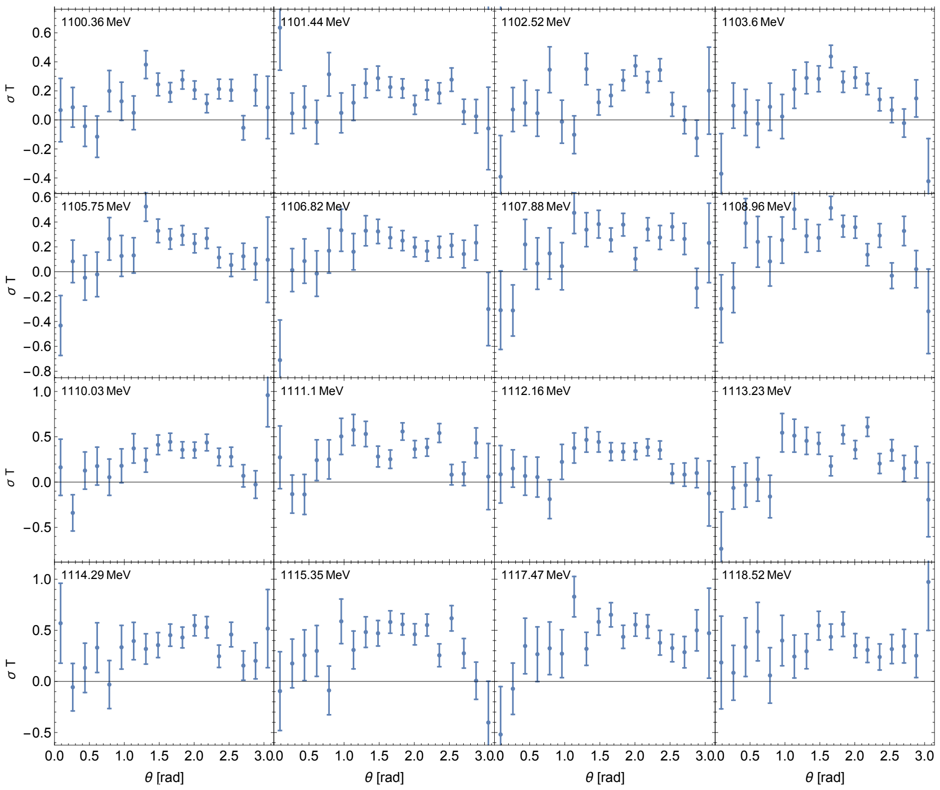
Minimal Models for Partial-Wave Analysis

J. LANDAY, M. DÖRING, C. FERNÁNDEZ RAMÍREZ, B. HU, AND R. MOLINA

Supported by NSF (PIF No. 1415459), DOE No. DE-AC05-06OR23177 and
Grant No. DE-SC0016582, CONACYT (Mexico) No. 251817, PAPIIT-DGAPA (UNAM)
No. IA101717, Red FAE, Mexico

Problem: How to fit all of the data?



S. Schumann et al. [A2 Collaboration], Phys. Lett. B 750, 252 (2015).

Multipole Parameterization

- No angular dependence
- Energy-dependent parameterization
- All observables can be expressed as bilinear superposition of multipoles
- Parameterization includes S, P, and D waves

$$
\begin{aligned}
\frac{d\sigma}{d\Omega}(\theta) &= \frac{q}{k}(A + B\cos(\theta) + C\cos^2(\theta)) \\
\Sigma &= \frac{q}{2k}(P_3^2 - P_2^2)\sin^2(\theta) / \frac{d\sigma}{d\Omega}(\theta) \\
P_{23}^2 &= \frac{1}{2}(P_2^2 + P_3^2) \\
A &= E_{0+}^2 + P_{23}^2 \\
B &= 2Re(E_{0+}P1^*) \\
C &= P_1^2 - P_{23}^2
\end{aligned}
$$

$$P_1 = \frac{1}{1000 m_{\pi+}} \cdot \frac{q(W, m_{\pi 0})}{m_{\pi+}} \cdot \sum_{i=1}^{n} \frac{p_{1i}}{10^{-i}} \cdot \left(\frac{(\omega(W) - m_{\pi 0})}{m_{\pi+}}\right)^i$$

Multipole Parameterization

- No angular dependence
- Energy-dependent parameterization
- All observables can be expressed as bilinear superposition of multipoles
- Parameterization includes S, P, and D waves

$$\begin{aligned}
\frac{d\sigma}{d\Omega}(\theta) &= \frac{q}{k}(A + B\cos(\theta) + C\cos^2(\theta)) \\
\Sigma &= \frac{q}{2k}(P_3^2 - P_2^2)\sin^2(\theta)/\frac{d\sigma}{d\Omega}(\theta) \\
P_{23}^2 &= \frac{1}{2}(P_2^2 + P_3^2) \\
A &= E_{0+}^2 + P_{23}^2 \\
B &= 2Re(E_{0+}P1^*) \\
C &= P_1^2 - P_{23}^2
\end{aligned}$$

$$P_1 = \frac{1}{1000 m_{\pi+}} \cdot \underbrace{\frac{q(W,m_{\pi 0})}{m_{\pi+}}}_{\text{Threshold behavior}} \cdot \sum_{i=1}^{n} \frac{p_{1i}}{10^{-i}} \cdot \left(\frac{(\omega(W) - m_{\pi 0})}{m_{\pi+}}\right)^i$$

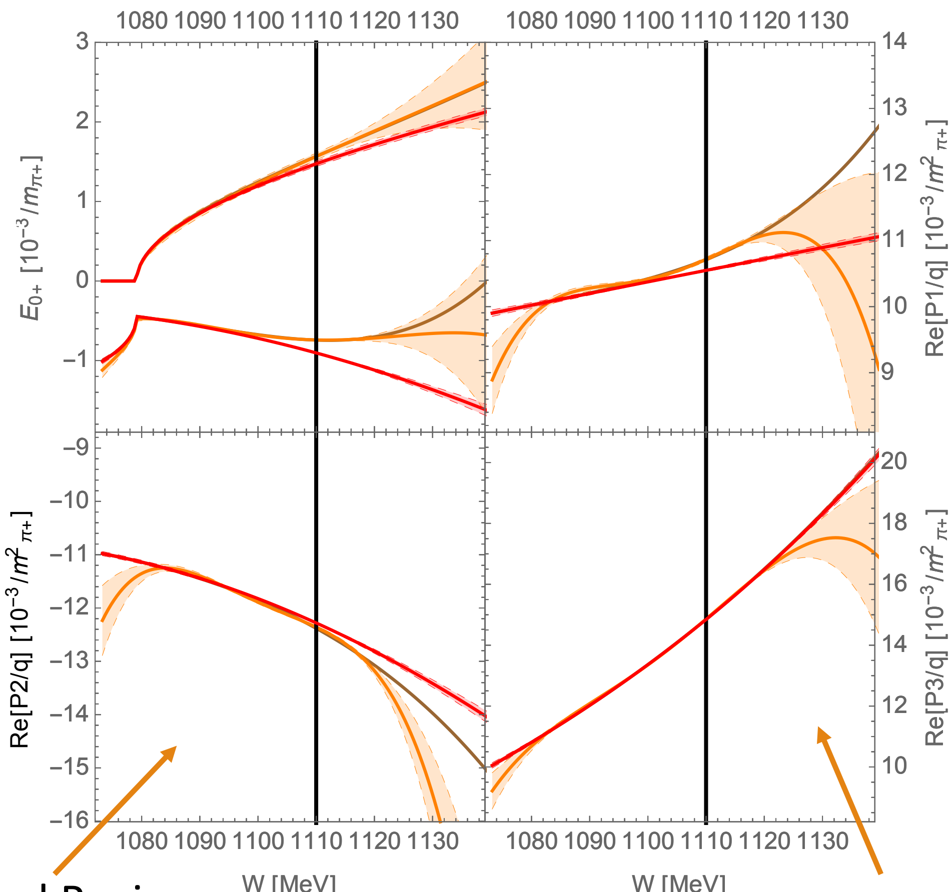
Multipole Parameterization

- No angular dependence
- Energy-dependent parameterization
- All observables can be expressed as bilinear superposition of multipoles
- Parameterization includes S, P, and D waves

$$\begin{aligned}
\frac{d\sigma}{d\Omega}(\theta) &= \frac{q}{k}(A + B\cos(\theta) + C\cos^2(\theta)) \\
\Sigma &= \frac{q}{2k}(P_3^2 - P_2^2)\sin^2(\theta)/\frac{d\sigma}{d\Omega}(\theta) \\
P_{23}^2 &= \frac{1}{2}(P_2^2 + P_3^2) \\
A &= E_{0+}^2 + P_{23}^2 \\
B &= 2Re(E_{0+}P1^*) \\
C &= P_1^2 - P_{23}^2
\end{aligned}$$

$$P_1 = \frac{1}{1000 m_{\pi+}} \cdot \underbrace{\frac{q(W, m_{\pi 0})}{m_{\pi+}}}_{\text{Threshold behavior}} \cdot \underbrace{\sum_{i=1}^{n} \frac{p_{1i}}{10^{-i}} \cdot \left(\frac{(\omega(W) - m_{\pi 0})}{m_{\pi+}}\right)^i}_{\text{Simple Taylor Series}}$$

Different models can give satisfactory fits. How do we determine the optimal one?



Fitted Region

Predicted Region

All solutions pass Pearson's Chi-Squared test.

Orange Solution- 23 parameters

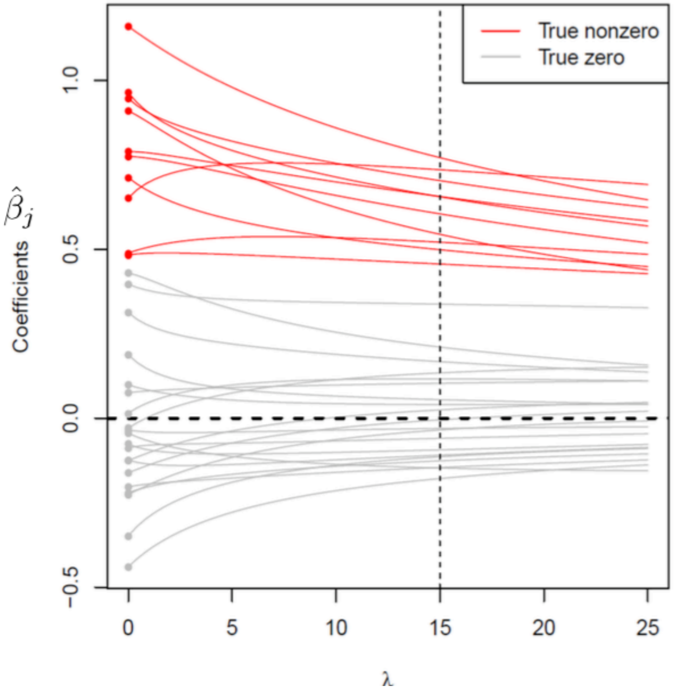
Red Solution – 13 parameters

Least Absolute Shrinkage and Selection Operator (LASSO)

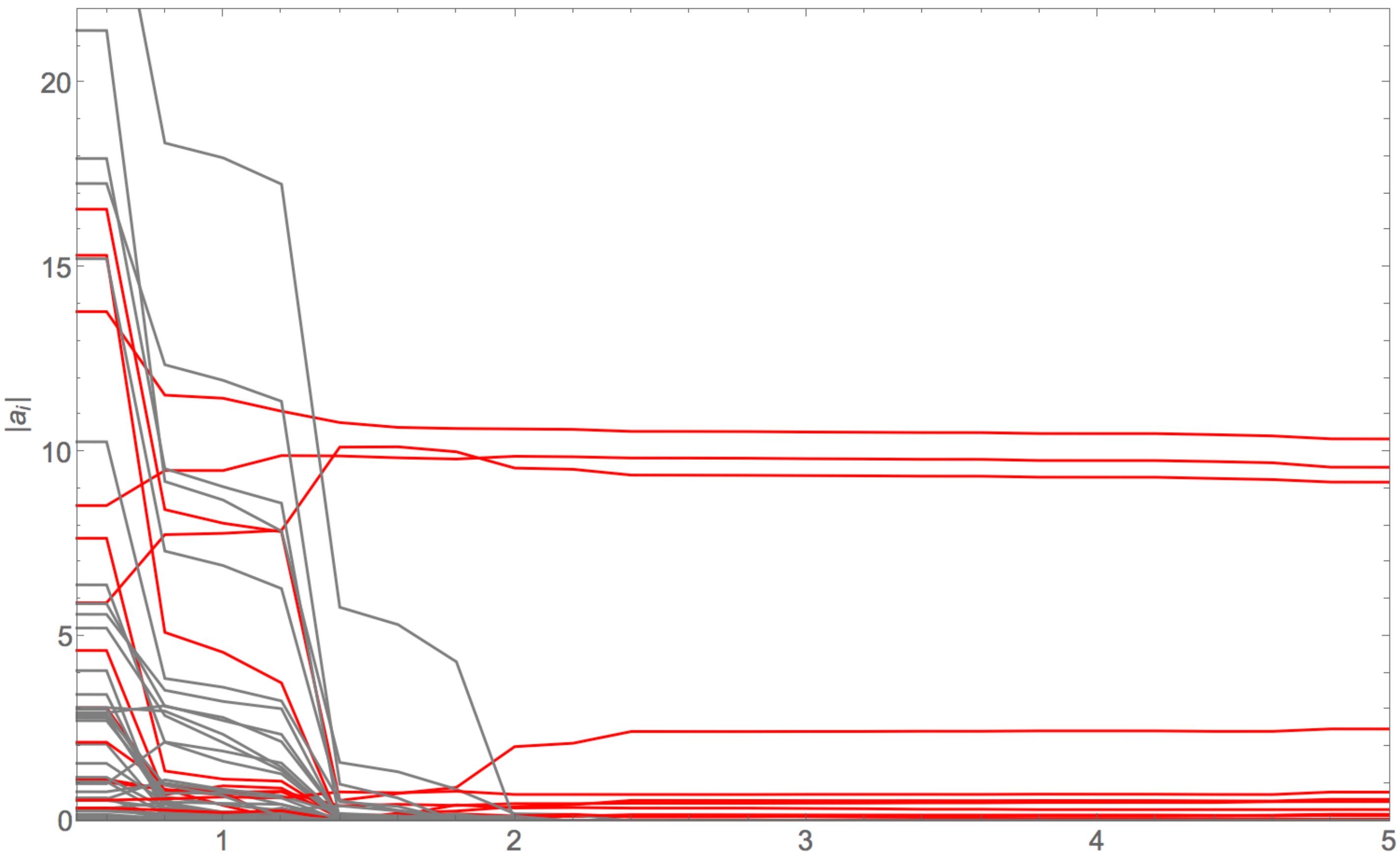
- Consider adding the moduli of fit parameters β_j to the chi-square, to some power ℓ :

$$\chi^2(\beta_j) = \sum_{i=1}^{n} \frac{(y_i - f(x_i, \beta_j))^2}{\sigma_i^2} \rightarrow \chi^2(\beta_j) = \sum_{i=1}^{n} \frac{(y_i - f(x_i, \beta_j))^2}{\sigma_i^2} + \lambda \sum_{j=1}^{m} |\beta_j|^\ell \quad \underline{\text{LASSO}}: \ell = 1$$

- What happens with the best parameters $\hat{\beta}_j$ of the minimized chi-square, as a function of the penalty parameter λ ?
- Consider "Ridge regression", that is a " ℓ_2 penalty", i.e., $\ell_2 = 2$.
- Result: Truly non-zero parameters stay finite, despite the finite penalty (as long as the penalty is not too large). Truly zero parameters become smaller, but do not go strictly to zero.
- Can we force truly-zero parameters to vanish?



Lasso Example: Fit to data from toy model with known best parameters

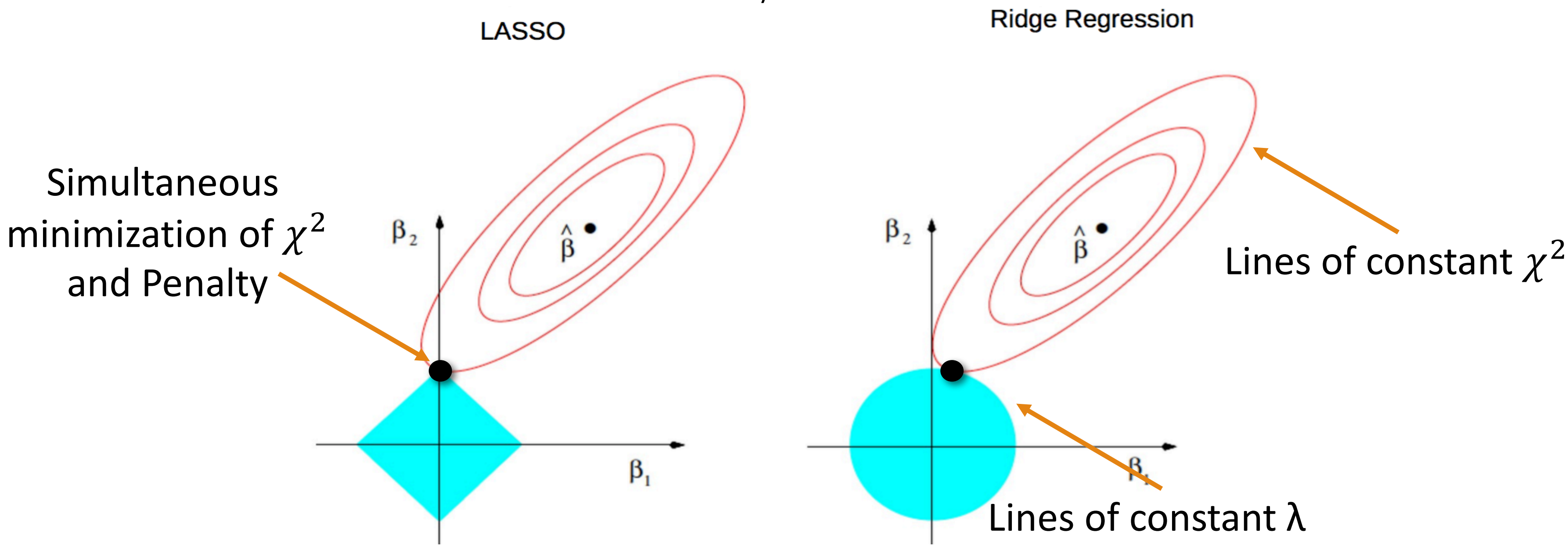


Red- Non-Zero Parameters λ Gray- Zero-Parameters

LASSO is capable of setting coefficients exactly to zero

$$\underbrace{\sum_{i=1}^{n} \frac{(y_i - f(x_i, \beta_j))^2}{\sigma_i^2}}_{\text{Normal } X^2} + \underbrace{\lambda \sum_{j=1}^{m} |\beta_j|}_{\text{Penalty Term}}$$

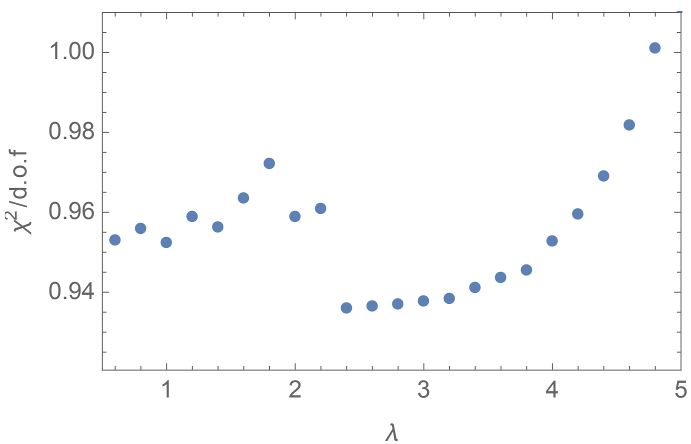
$\hat{\beta}_i$: Best parameters without penalty
$\beta_i = 0$: Best parameters only penalty

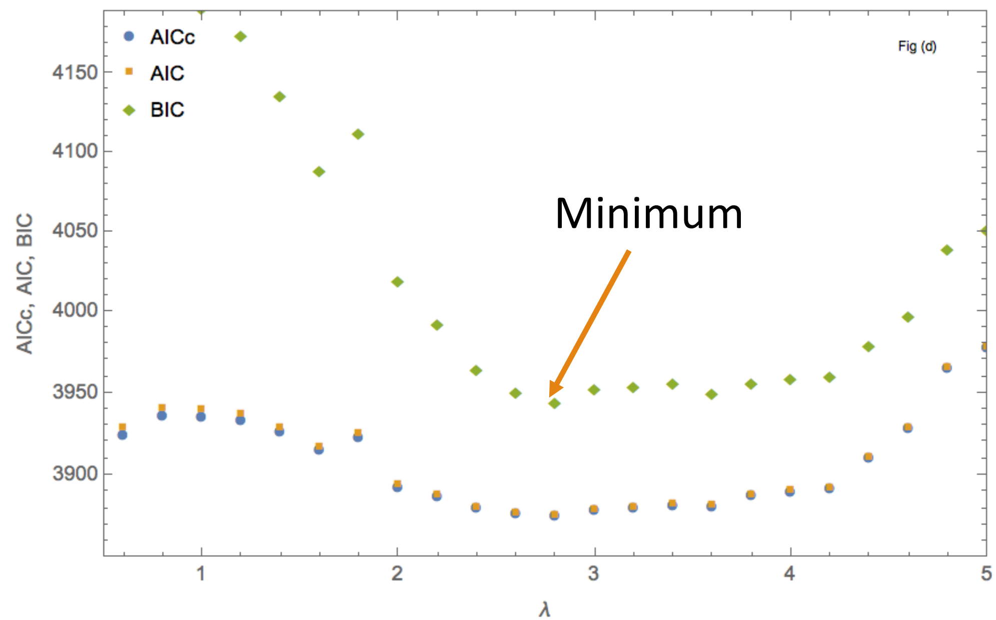


How to decide best value of λ?

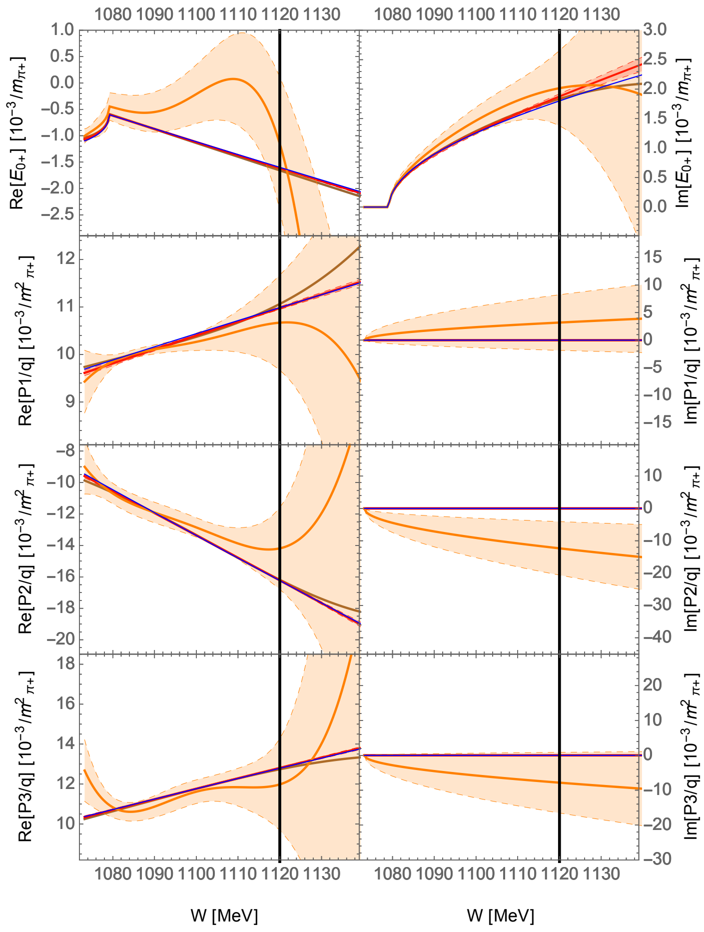
$$
\begin{aligned}
AIC &= 2k + \chi^2 \\
AICc &= AIC + \frac{2k(k+1)}{(n-k+1)} \\
BIC &= k \log n + \chi^2
\end{aligned}
$$

k : Number of parameters
n : Number of data points



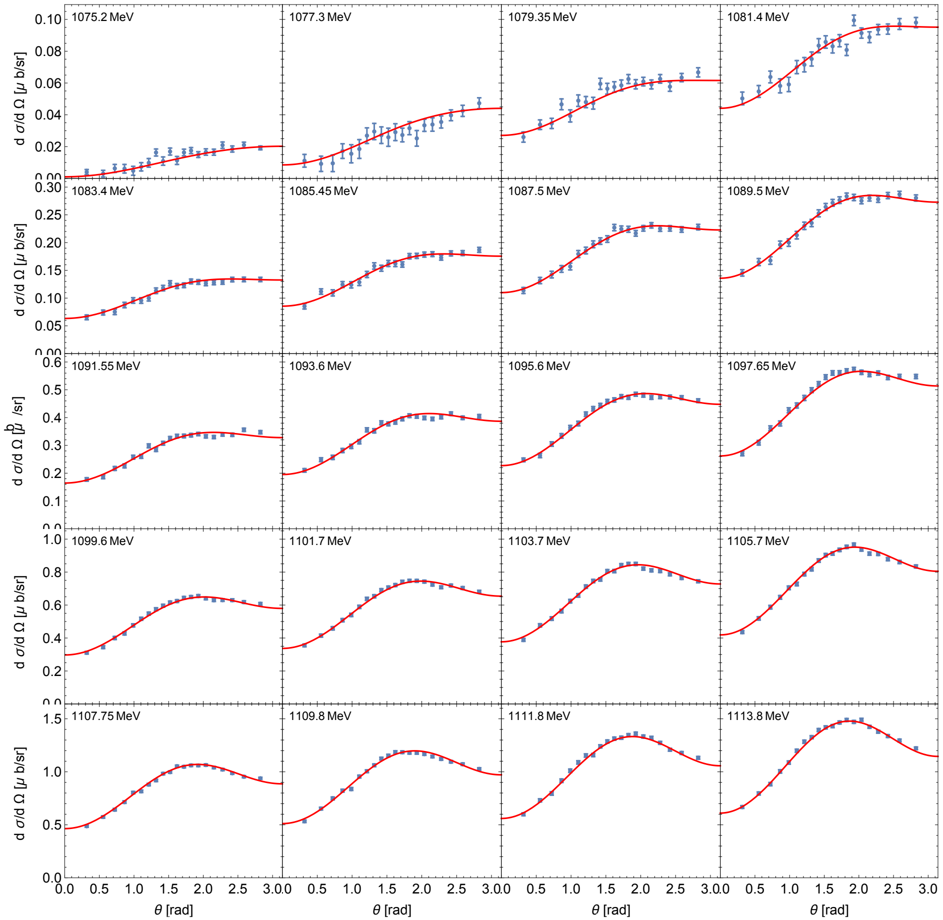


Toy Model Results

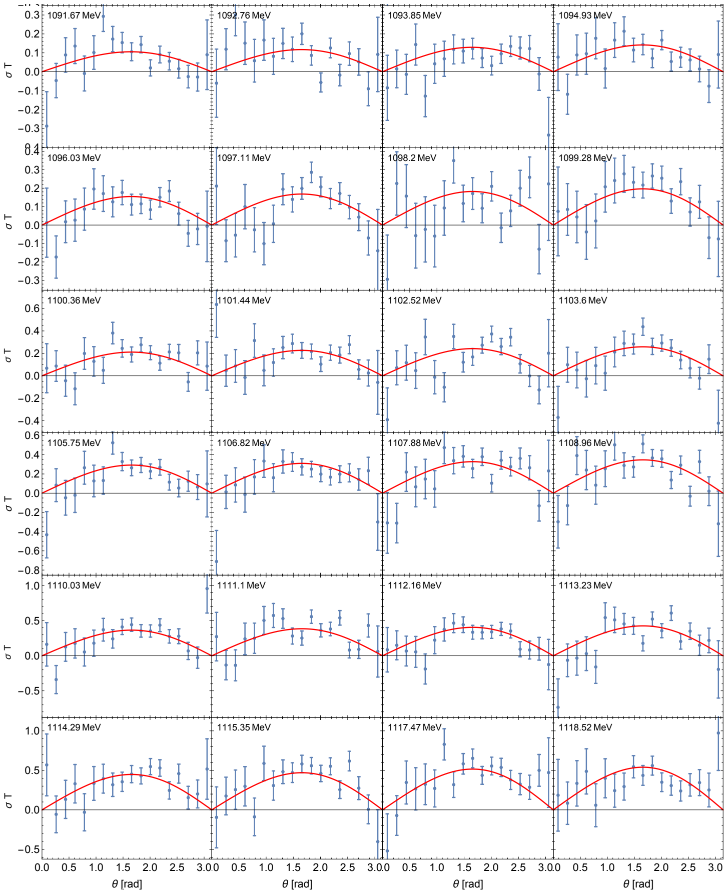


- Generate data from a toy model using a 9 parameter model (2 real S-waves, 1 imaginary S-wave, and 2 real $P_{1,2,3}$ –waves shown in blue
- LASSO (red) eliminates 36 parameters from a 46 parameter fit (orange) and reconstructs the true solution (blue) quite accurately
- LASSO sets all imaginary parts of P-waves and D- waves correctly to 0
- LASSO solution predicts true solution quite accurately beyond the fitted W_{max} =1120 MeV

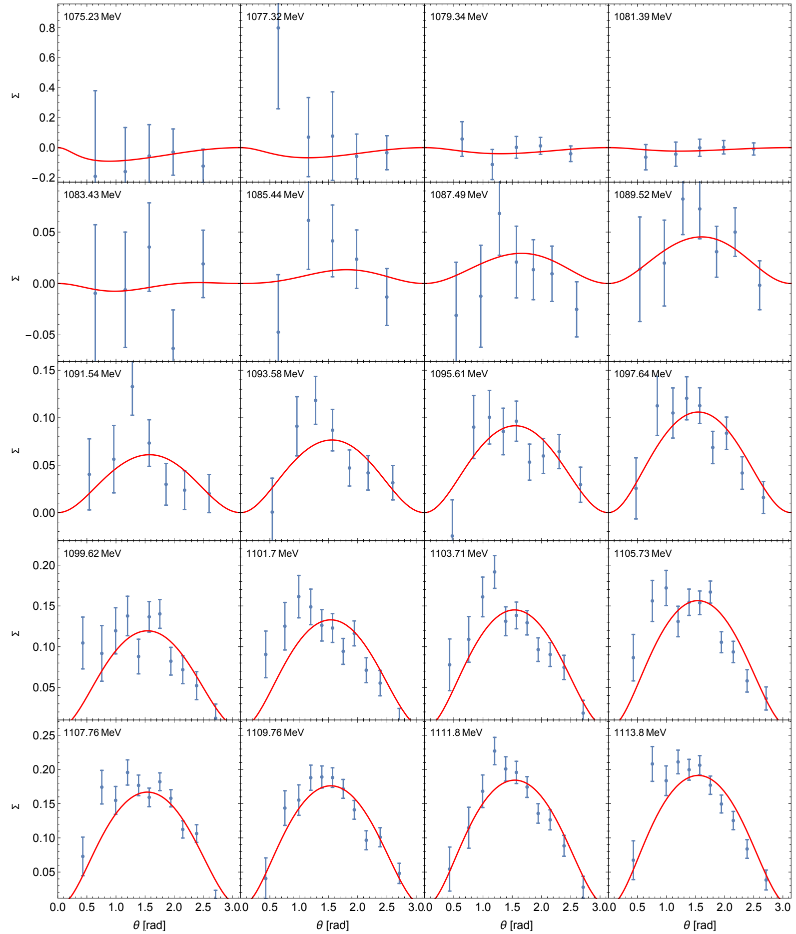
Analysis of Real Data (Differential Cross Section)



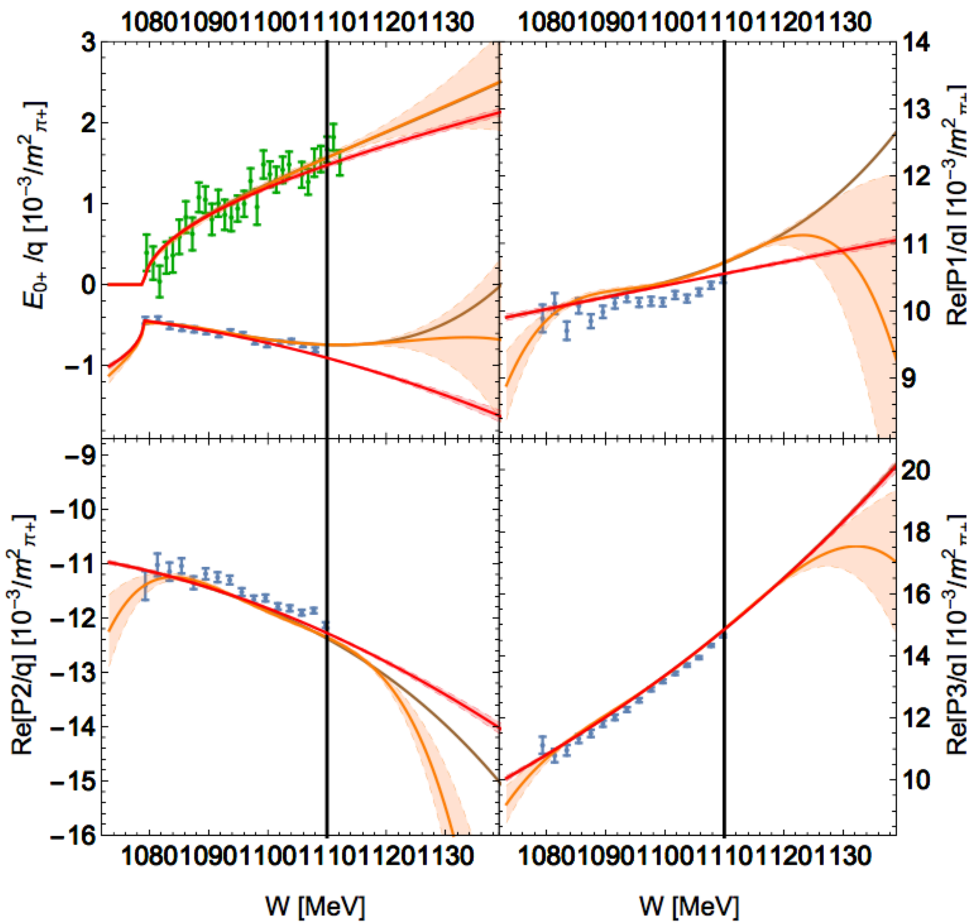
Analysis of Real Data (Target Polarization)



Analysis of Real Data (Beam Assymetry)



Extraction of Multipoles



46 parameter fit

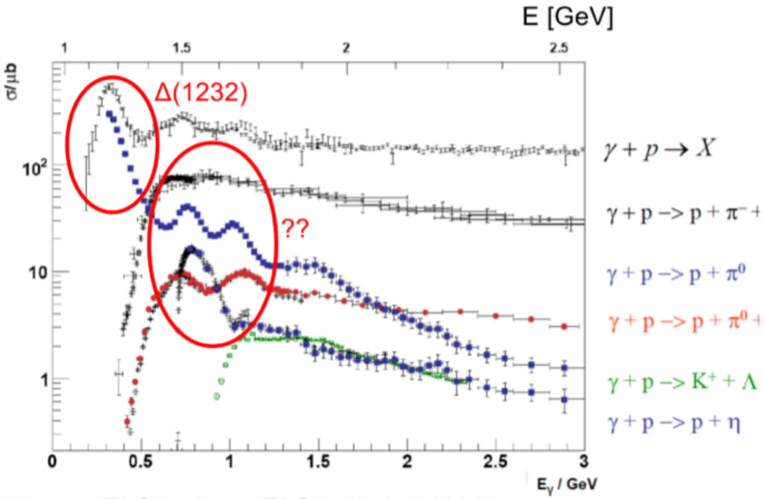
10 parameter fit

SE Extraction: D. Hornidge et al. Phys. Rev. Lett. 111, 062004(2013)

SE Extraction: S. Schumann et al, Phys. Lett. B 750, 252 (2015).

Outlook: The missing resonance problem

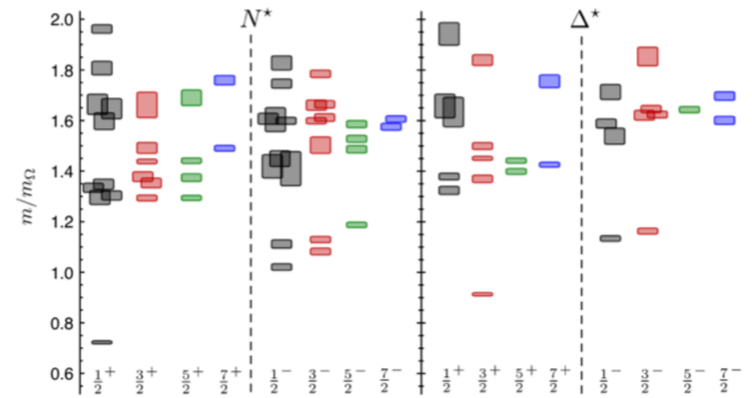
Experimental study of hadronic reactions



source: ELSA; data: ELSA, JLab, MAMI

$\Longleftrightarrow$

Theoretical predictions of excited hadrons e.g. from lattice calculations:



$m_\pi = 396$ MeV [Edwards *et al.*, Phys.Rev. D84 (2011)]

Much less resonances experimentally seen (even after multipole decomposition) than theoretically predicted. $\Longrightarrow$ Missing resonance problem

- In energy dependent fits of multipoles to data, many times resonances are used as easy fix to describe data → Problematic overpopulation of resonance spectrum
- Solution: Penalize the occurrence of resonances with LASSO and only allow them if really needed by data.
- In practical terms: Provide very flexible background (so the fit can use it to describe data)+ overpopulated resonance spectrum. Then, LASSO will eliminate all but the significant resonances

Conclusions

LASSO provides automatized and blindfolded method to scan larges classes of models and select the simplest one

In the case of pion photo-production LASSO fit results provide optimal accuracy and predictability.

LASSO has great potential to be used with resonance data and address missing resonance problem

Solution— The LASSO

- <u>L</u>east <u>A</u>bsolute <u>S</u>election and <u>S</u>hrinkage <u>O</u>perator
- Invented by Tibshirani 1996 http://statweb.stanford.edu/~tibs/
- Free online books and lectures:
  - Web page with 2013 *Data Mining* lecture by Ryan Tibshirani, http://www.stat.cmu.edu/~ryantibs/datamining/
  - Free online books:
  - *The Elements of Statistical Learning: Data Mining, Inference, and Prediction,* T. Hasti, R. Tibshirani, J. Friedman, Springer 2009 second ed.; E-book available at: http://statweb.stanford.edu/~tibs/lasso.html
  - A bit easier to follow:

    An Introduction to Statistical Learning, Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, Springer 2015; 6th printing; E-book available at http://www-bcf.usc.edu/~gareth/ISL/.

The F-Test

We implemented an F-test to compare the two models:

$$y = \frac{(\chi_1^2 - \chi_2^2)/k}{\chi_2^2/(n-m-k)}$$

k: Number of parameters in model 1
m+k: Number of parameters in model 2
n : Number of of data points

y = 1.64 < y = 2.63, indicating that the over fit is indeed not significantly better than the simplest fit.

Other statistical tests lead to similar conclusions

Shapiro Wilk

Kolmogorov Smirnov

Anderson Darling