Fitting the $K^+\Sigma^-$ photoproduction within an Isobar model using a novel approach.

Dimitrios Petrellis

Nuclear Physics Institute, Czech Academy of Sciences

Motivation

- Why KΣ photoproduction?
  - Quark models predict more resonances than observed in π-N scattering experiments (“missing” resonance problem). Indications that these states may be found in KY (Y = Λ, Σ) channels
  - E/M interaction very well understood
  - Studied at several facilities: CEBAF, MAMI, ELSA, Spring-8, GRAAL
  - New data on photon beam asymmetries from CLAS
- Isobar model: phenomenological models useful in bridging the gap between fundamental theory and experiment

Motivation

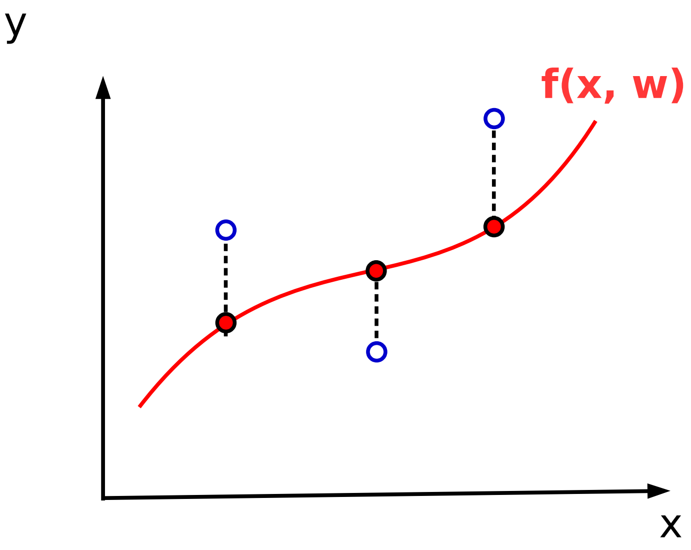
- No one single resonance dominates in $E^{lab}_{\gamma} \approx$ 1-2 GeV, but many (>20), broad and overlapping
  - extremely large number of possible combinations (**models**)
  - large number of parameters → ordinary χ^2 fitting: problematic similar minima, large *variations* in the parameter values
  - *Regularized* χ^2 fitting → penalty term constrains the number and magnitude of the parameters
    - improves the quality of the fits
    - selects the best subset of parameters → resonances evaluated as most “necessary” by the data → **model**

Ordinary Least Squares fitting

- set of data: pairs observations

$$D = \{ (x_1, y_1), (x_2, y_2), \ldots (x_N, y_N) \}$$

e.g. $x = E\,, \quad y = d\sigma / d\Omega$



- Function: $f(x, \mathbf{w})$

- Parameters: $\mathbf{w} = (w_0, w_1, \ldots w_K)$

- Goal: determine values of the parameters **w*** that minimize some error function

$$E = \sum_{i=1}^{N} \left[y_i - f(x_i, \mathbf{w}) \right]^2 \qquad \chi^2 = \sum_{i=1}^{N} \left[\frac{y_i - f(x_i, \mathbf{w})}{\sigma_i} \right]^2$$

[...this, however, is problematic as we may overfit the data]

The problem of overfitting though an example

Create artificial data by adding Gaussian noise

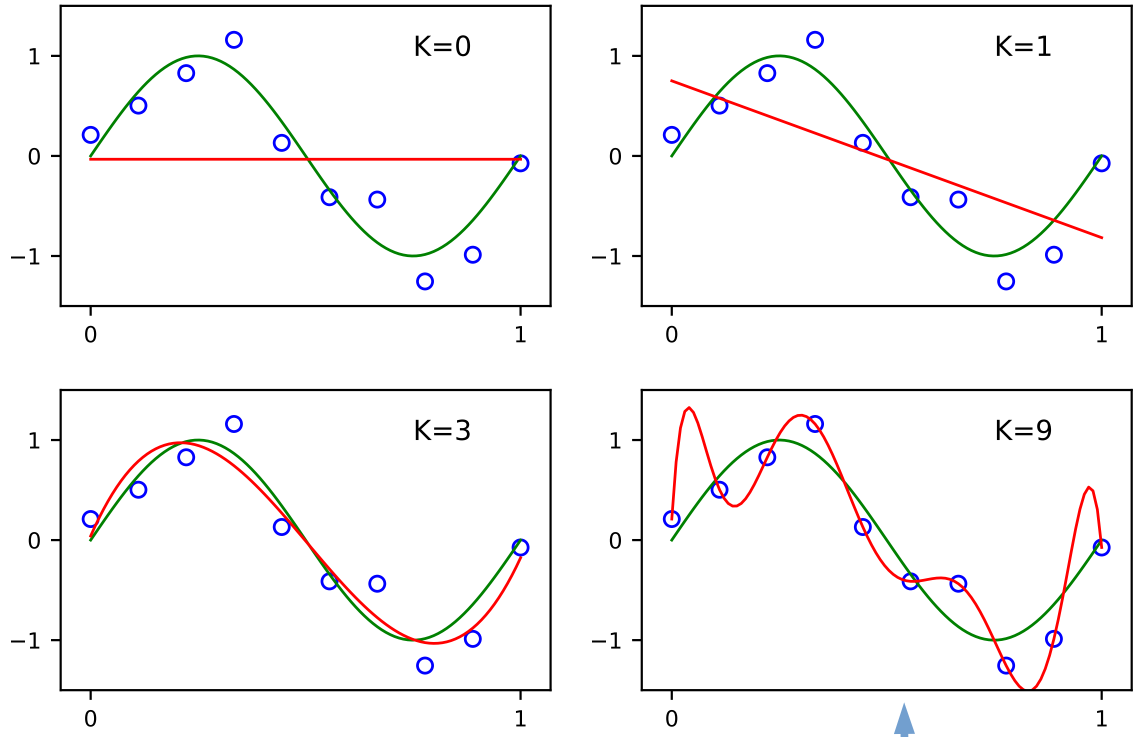
$$y = \sin(x) + \epsilon$$

fit the data with a polynomial

$$f(x, \mathbf{w}) = w_0 + w_1 x + w_2 x^2 + \ldots + w_K x^K$$

increasing order **K** of polynomial

- fits the data very well, but
- poor description of the function that generated them



Model fits the noise in the sample

Where to stop?
What is the optimal complexity of our model?
Error minimization alone, does not guarantee the quality of the fitting

{ Occam's razor
Law of parsimony }

Figure adapted from:C. Bishop Pattern Recognition and Machine Learning, Springer 2006

Regularization: a remedy for over-fitting

$$\textbf{min} \left(\chi^2 + \lambda \sum_{j=1}^{K} |w_j|^q\right)$$

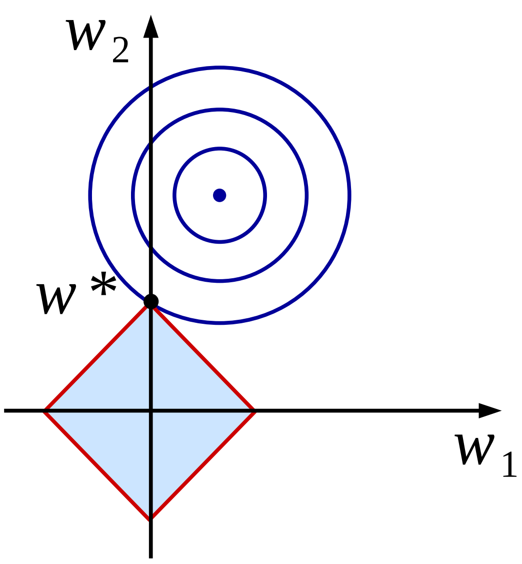
- introduction in χ^2 of a term that penalizes large values of the parameters w_j

- ~ minimize χ^2,subject to constraint:

$$\sum_{j=1}^{K} |w_j|^q \leqslant \eta$$

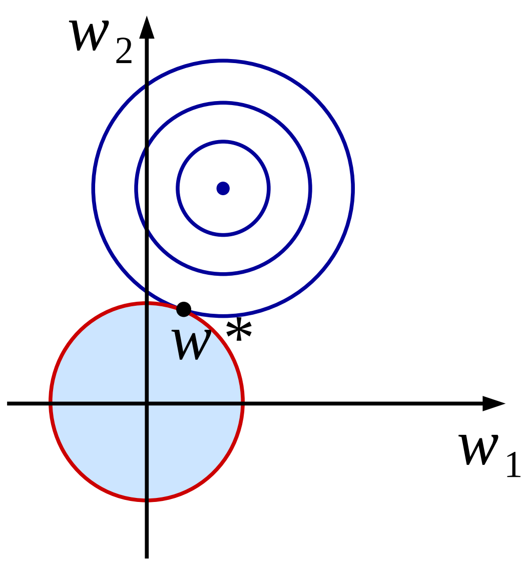
LASSO*

q = 1 → L1 norm



Ridge

q = 2 → L2 norm



w* = optimum value for **w** under the constraint

- for q = 1 (LASSO) → some parameters become zero ($w^*_1 = 0$)

* **L**east **A**bsolute **S**hrinkage and **S**election **O**perator

LASSO for variable selection

- LASSO forces some of the parameters to zero → selects a subset

- λ, regularization parameter → strength of the penalty term

- λ: controls how many parameters are switched-off and how many remain
  (λ practically selects a model)

- try several λ_1, λ_2,... values and choose the optimal λ based on:

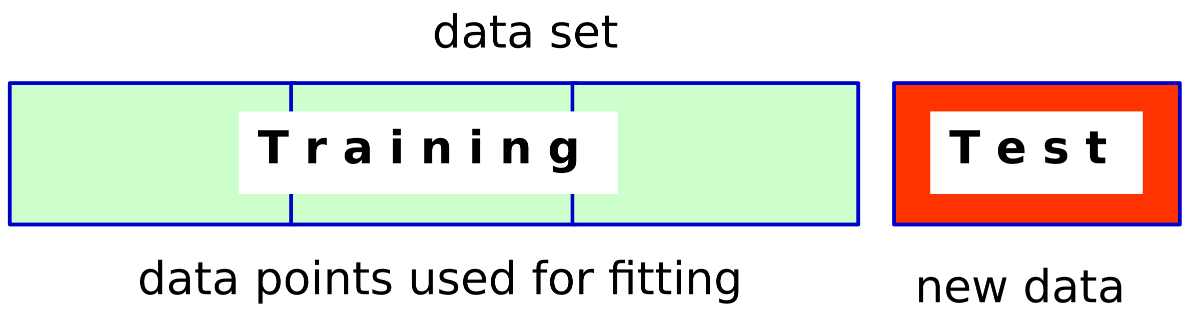
  either
  - **Validation**

  or
  - **Information criteria**

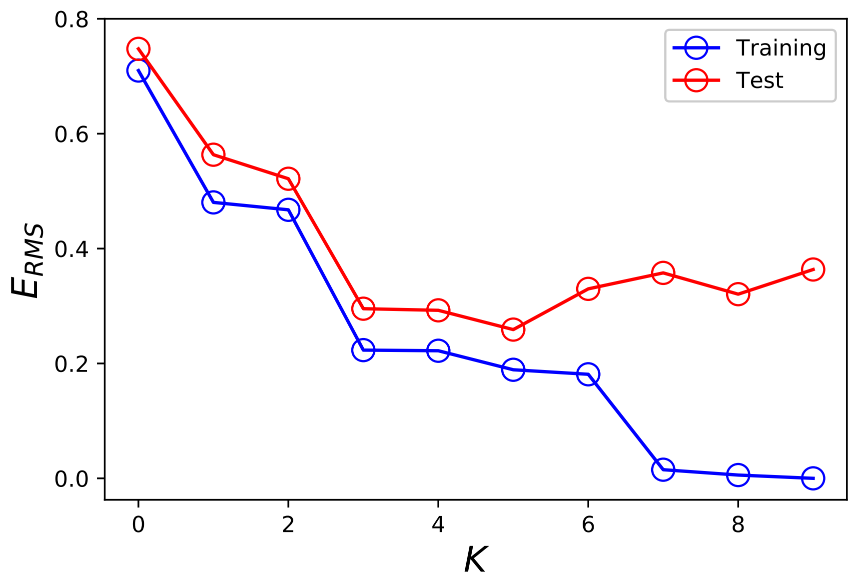
    - Akaike Information Criterion (AIC)

    - Bayesian Information Criterion (BIC)

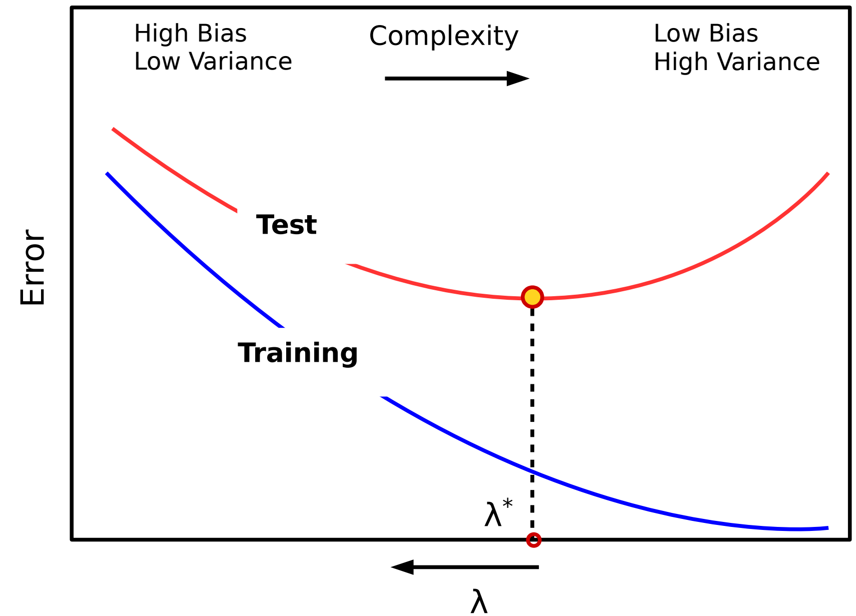
Validation: Training & Test set errors



- Fit model on the training set → Training Error
- Test the fitted model on the test set → Test Error
- Repeat while increasing complexity
  (Forward selection)



Bias-Variance trade-off



Information Criteria (IC)

Approach equivalent to validation

For a series of models *i* = 1, 2, ...m

Akaike IC: $$\textcolor{blue}{\mathrm{AIC}} = \chi^2_{\min} + 2k_i$$

Bayesian IC: $$\textcolor{red}{\mathrm{BIC}} = \chi^2_{\min} + k_i \ln(N)$$

k_i : number of parameters corresponding to model i

N : number of data points

Choose the model with the minimum AIC, BIC

[both AIC and BIC give similar results, although BIC tends to penalize complexity more]

In the case of LASSO: model i → λ_i ⟹ Choose λ_i that results in the minimum IC

Parameters and observables

Resonances
masses, widths: from PDG

Parameters
coupling constants
hadron form factor cutoffs

674 data points from: CLAS, LEPS

Observables
differential cross sections
photon beam asymmetries

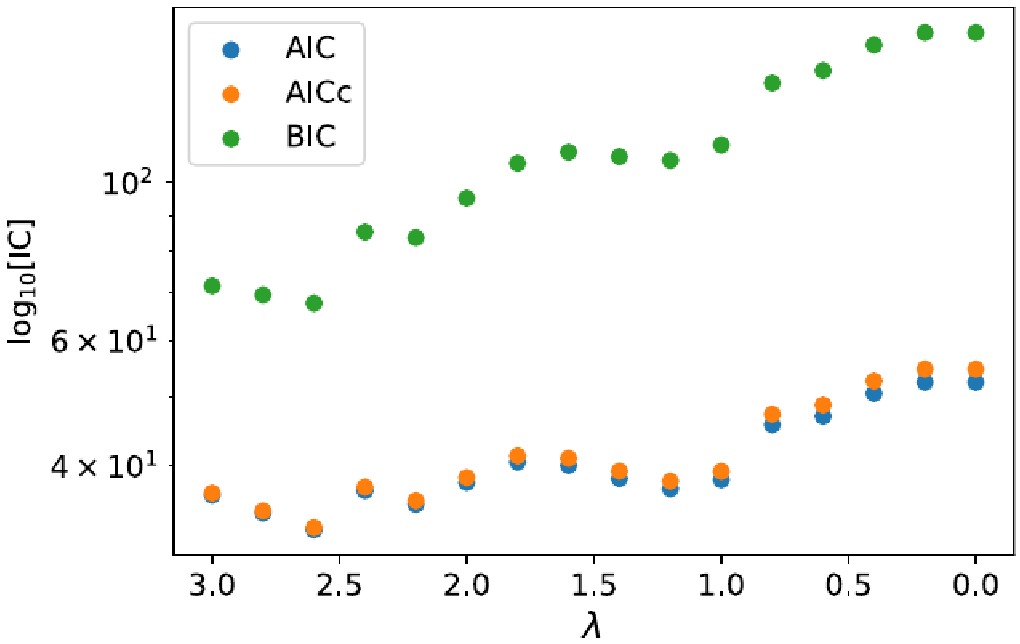
Tag	Resonance	Mass (MeV)	Width (MeV)
K*	$K^*(892)$	891.7	50.8
K1	$K_1(1270)$	1270	90
N3	$N(1535)\ 1/2^-$	1530	150
N4	$N(1650)\ 1/2^-$	1650	125
N8	$N(1675)\ 5/2^-$	1675	145
N6	$N(1710)\ 1/2^+$	1710	140
N7	$N(1720)\ 3/2^+$	1720	250
P4	$N(1875)\ 3/2^-$	1875	200
P1	$N(1880)\ 1/2^+$	1880	300
Mx	$N(1895)\ 1/2^-$	1895	120
P2	$N(1900)\ 3/2^+$	1920	200
M4	$N(2060)\ 5/2^-$	2100	400
M1	$N(2120)\ 3/2^-$	2120	300
D1	$\Delta(1900)\ 1/2^-$	1860	250

Minimization with: MINUIT Library

Isobar code available at:

http://www.ujf.cas.cz/en/departments/department-of-theoretical-physics/isobar-model.html

Applying the Information Criteria



Run: Isobar code + Minuit + LASSO

Forward selection:

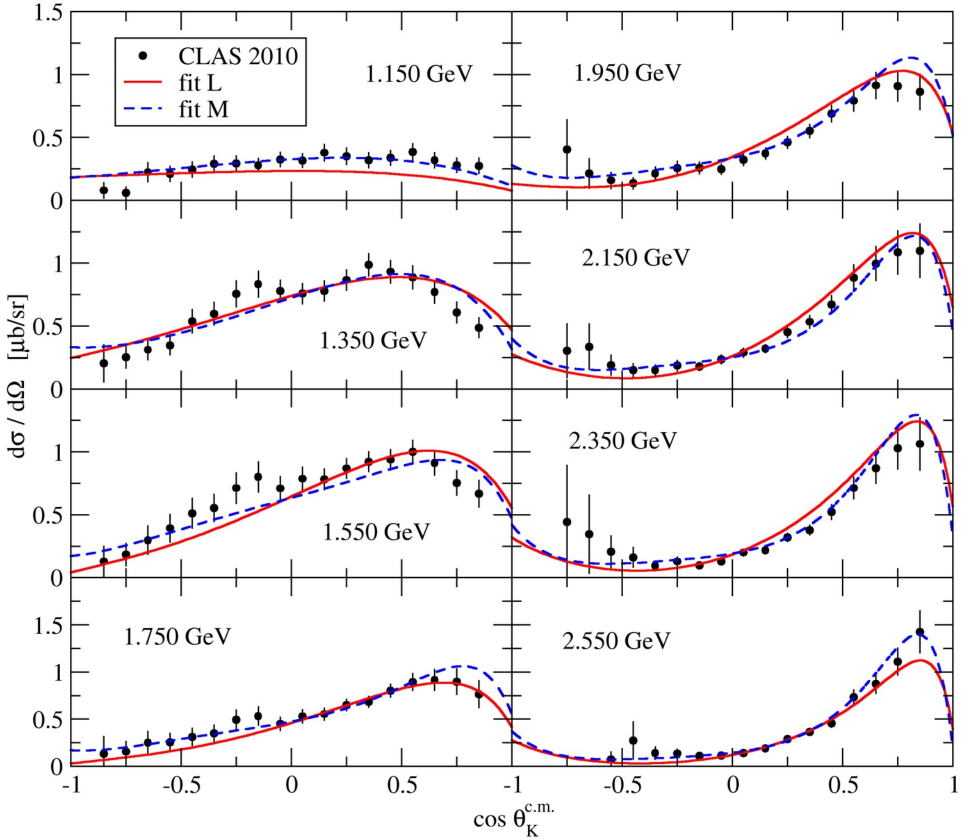
- start with the full model, all parameters initialized with random values and use some λ_{max}
- perform LASSO χ^2 minimization and compute AIC, BIC
- in each run progressively decrease λ and rerun LASSO using the fitted parameter values of the last run as starting values
- repeat until λ_{min} is reached
- optimal λ occurs at the minimum of BIC, AIC

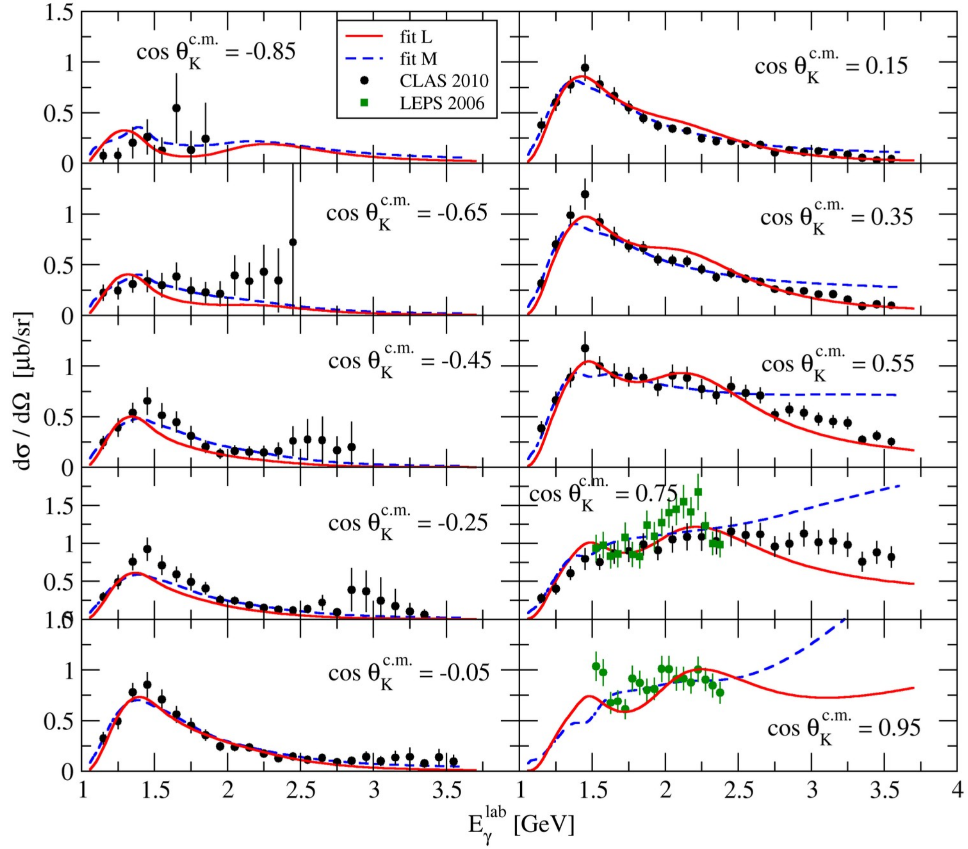
Reduction in the number of parameters

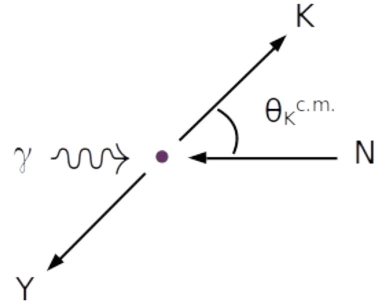
Tag	Resonance	Mass (MeV)	Width (MeV)	Branching ratio $K\Lambda$	Branching ratio $K\Sigma$	Fit M g_1	Fit M g_2	Fit L g_1	Fit L g_2
K*	$K^*(892)$	891.7	50.8			0.366 ± 0.024	1.103 ± 0.198	0.310 ± 0.019	
K1	$K_1(1270)$	1270	90			-1.448 ± 0.189	0.473 ± 0.156		
N3	$N(1535)\ 1/2^-$	1530	150			-0.709 ± 0.071			
N4	$N(1650)\ 1/2^-$	1650	125	0.07	0.00	0.314 ± 0.034		-0.085 ± 0.006	
N8	$N(1675)\ 5/2^-$	1675	145			-0.013 ± 0.001	0.022 ± 0.003	-0.010 ± 0.001	0.003 ± 0.002
N6	$N(1710)\ 1/2^+$	1710	140	0.15	0.01	-0.940 ± 0.093			
N7	$N(1720)\ 3/2^+$	1720	250	0.05	0.00	-0.098 ± 0.017	-0.082 ± 0.002	-0.187 ± 0.004	-0.126 ± 0.002
P4	$N(1875)\ 3/2^-$	1875	200	0.01	0.01	-0.220 ± 0.023	-0.223 ± 0.023	-0.042 ± 0.015	0.025 ± 0.013
P1	$N(1880)\ 1/2^+$	1880	300	0.16	0.14	-0.050 ± 0.064			
Mx	$N(1895)\ 1/2^-$	1895	120	0.18	0.13	-0.063 ± 0.005		0.019 ± 0.002	
P2	$N(1900)\ 3/2^+$	1920	200	0.11	0.05	-0.051 ± 0.005	-0.004 ± 0.001	0.027 ± 0.003	0.010 ± 0.001
M4	$N(2060)\ 5/2^-$	2100	400	0.01	0.03	-0.00001 ± 0.0001	0.003 ± 0.0003	-0.003 ± 0.0001	0.004 ± 0.0002
M1	$N(2120)\ 3/2^-$	2120	300			-0.034 ± 0.014	-0.010 ± 0.013	0.0003 ± 0.001	0.0 ± 0.0001
D1	$\Delta(1900)\ 1/2^-$	1860	250		0.01	0.298 ± 0.028			
D2	$\Delta(1930)\ 5/2^-$	1880	300						
D3	$\Delta(1920)\ 3/2^+$	1900	300						
D4	$\Delta(1940)\ 5/2^-$	1950	400						
S1	$\Sigma(1660)\ 1/2^+$	1660	100						
S2	$\Sigma(1750)\ 1/2^-$	1750	90						
S3	$\Sigma(1670)\ 3/2^-$	1670	60						
S4	$\Sigma(2010)\ 3/2^-$	1940	220						

	M full fit	L LASSO fit
no. of resonances	14	9
no. of parameters	25	17

Results: differential cross sections







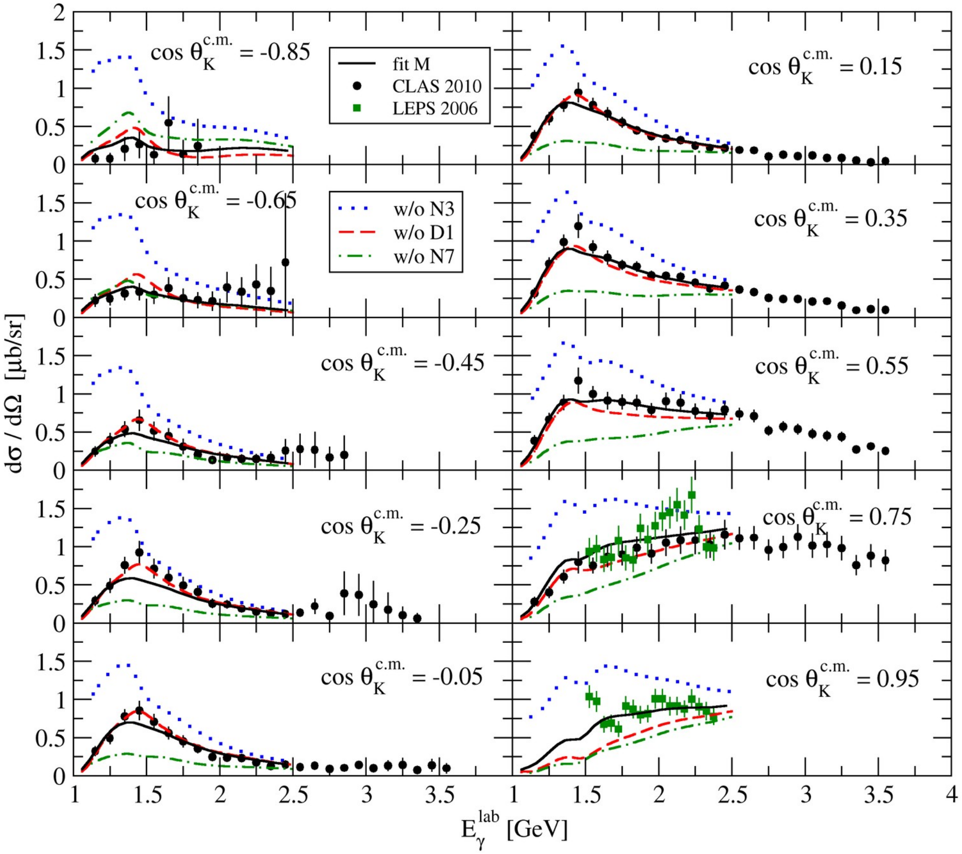
$\theta_K^{c.m.}$: Kaon center-of-mass angle

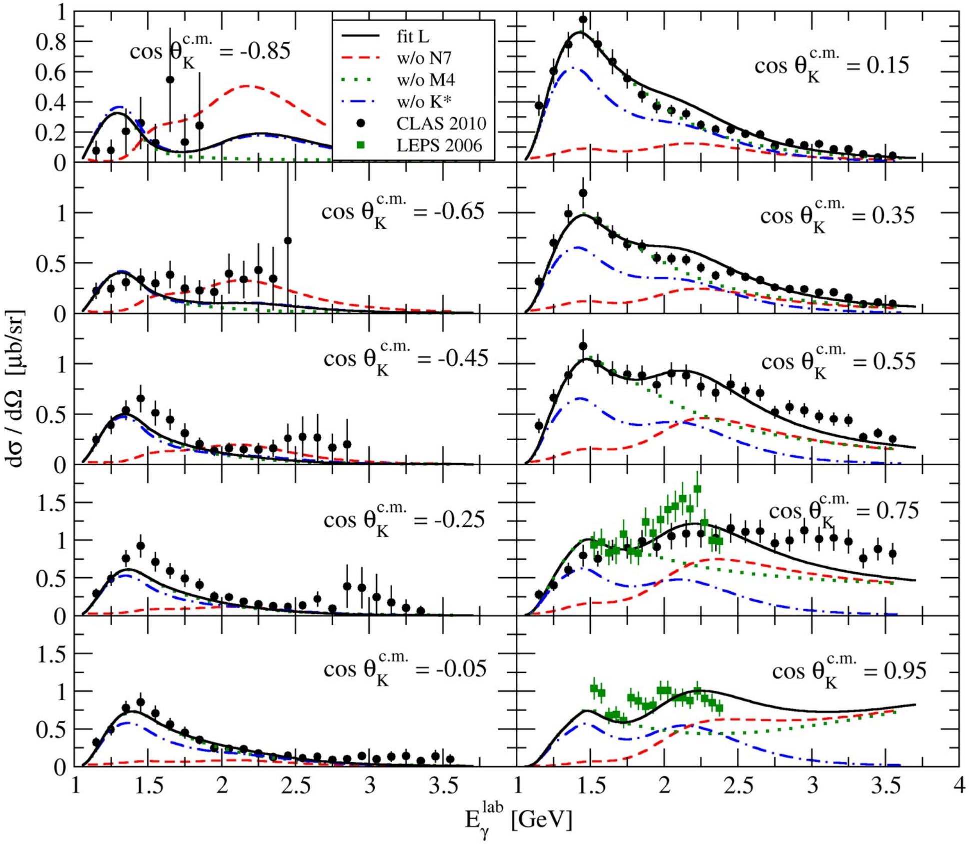
fit M: MINUIT

fit L: MINUIT + LASSO

E^{lab}_{γ}: incident photon energy

Results: differential cross sections

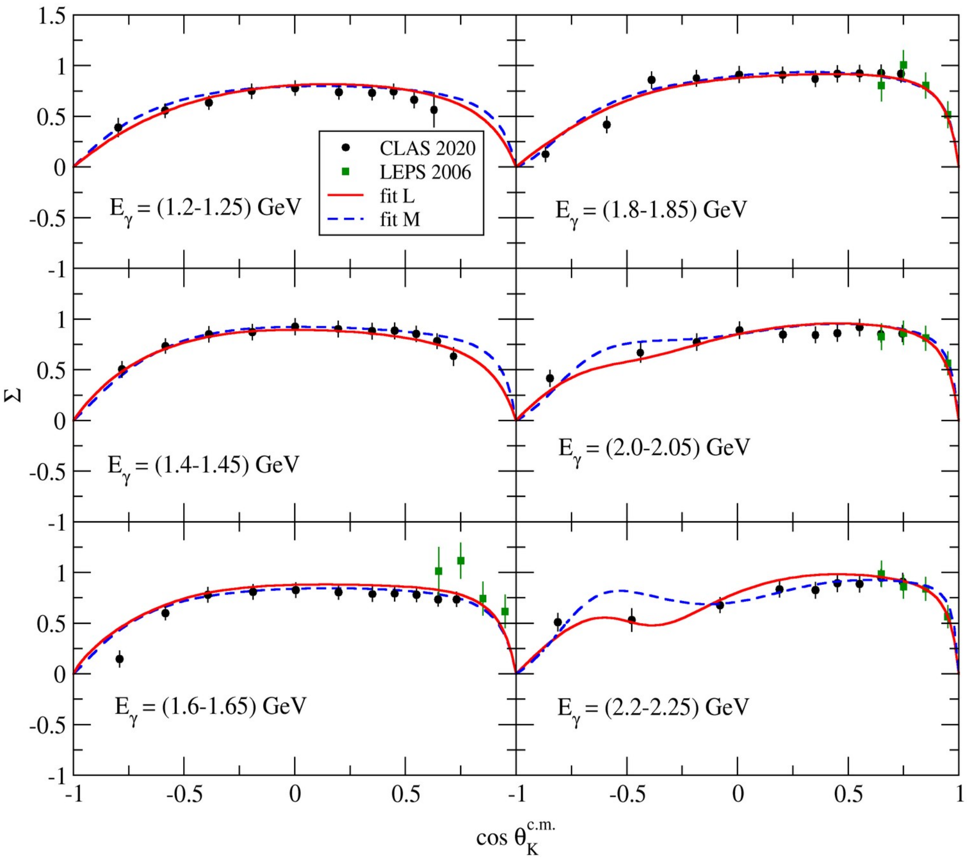


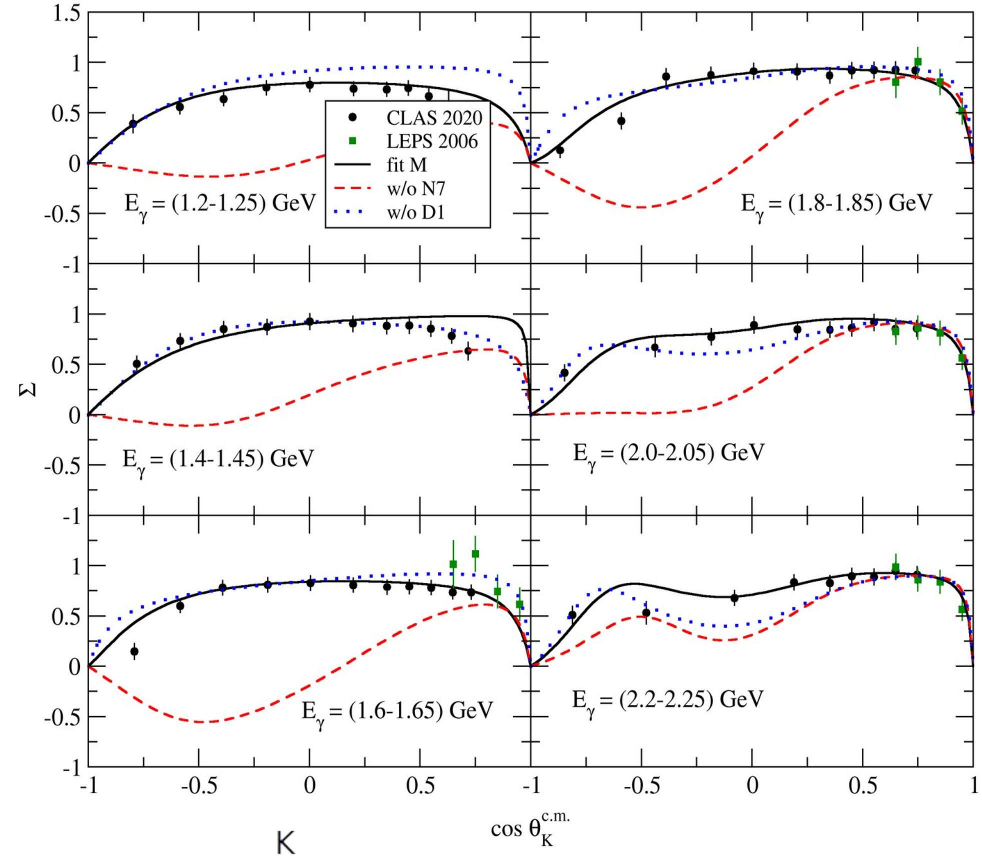


M: MINUIT fit
M fits w/o
N3 = N (1535) 1/2 −
D1 = Δ(1900) 1/2 −
N7 = N (1720) 3/2 +

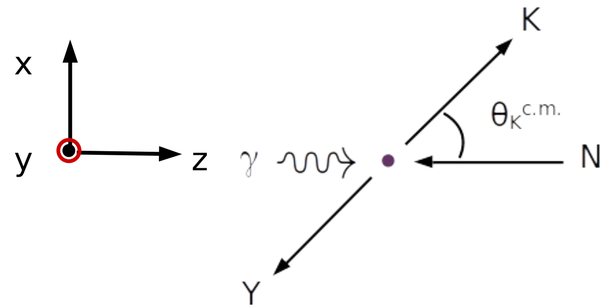
L: MINUIT + LASSO
L fits w/o
N7 = N (1720) 3/2 +
M4 = N (2060) 5/2 −
K ∗ (892)

Photon beam asymmetry Σ





Linearly polarized photons



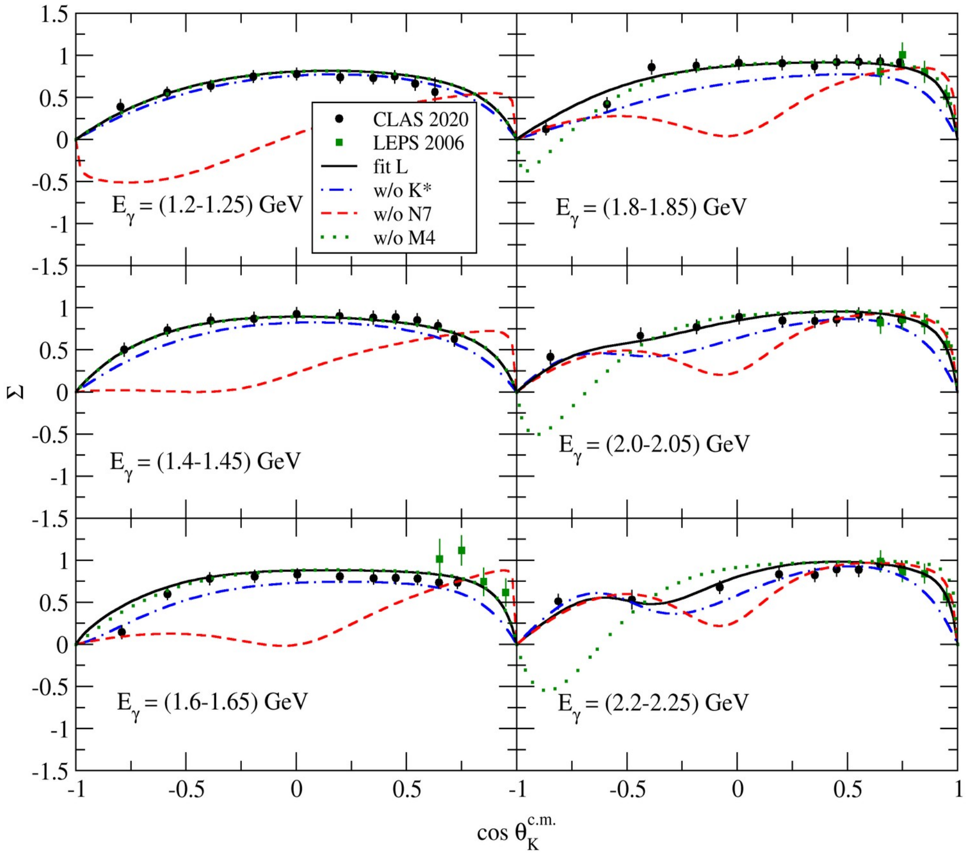
M: MINUIT fit
M fits w/o
D1 = Δ(1900) 1/2 −
N7 = N (1720) 3/2 +

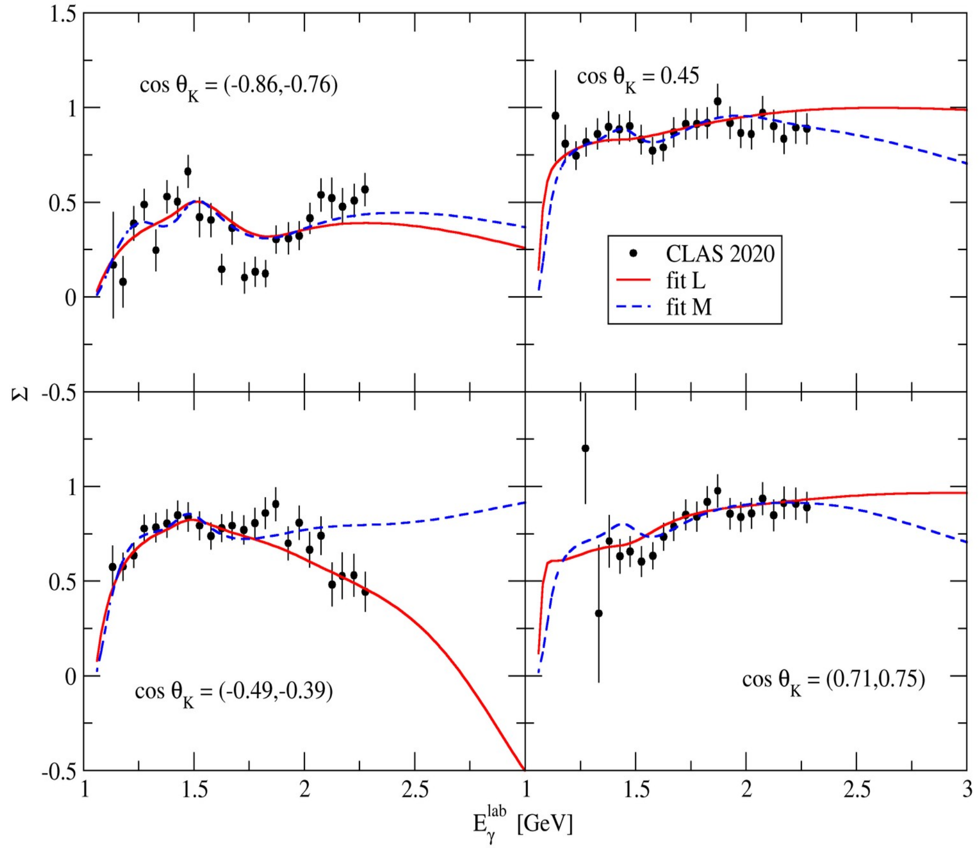
$$\epsilon^{\lambda=x} \equiv \epsilon^{\parallel} = (0,1,0,0)$$

$$\epsilon^{\lambda=y} \equiv \epsilon^{\perp} = (0,0,1,0)$$

$$\Sigma = \frac{d\sigma^{\perp} - d\sigma^{\parallel}}{d\sigma^{\perp} + d\sigma^{\parallel}}$$

Photon beam asymmetry Σ





L: MINUIT + LASSO
L fits w/o
N7 = N (1720) 3/2 +
M4 = N (2060) 5/2 −
K ∗ (892)

Summary and outlook

- In modeling $K^+\Sigma^-$ photoproduction with an Isobar model the large number of parameters makes the fitting to data problematic.
- Using regularized least squares (LASSO) we are able to reduce the number of parameters needed to describe the data.
- The combination of the LASSO method with information criteria provides a method to choose the best subset of parameters (model).
- Future plan: fit simultaneously all 4 channels of $K\Sigma$ photoproduction → relate coupling constants by SU(2) (Isospin) symmetry

Bibliographic references

- P. Bydzovsky, A. Cieply, D. Petrellis, D. Skoupil and N. Zachariou, Model selection for $K^+\Sigma^-$ photoproduction within an isobar model, Phys. Rev. C 104, 065202 (2021).
- N. Zachariou et al., Beam-spin asymmetry Σ for Σ^- hyperon photoproduction off the neutron, Phys. Lett. B 827, 136985 (2022).
- D. Skoupil and P. Bydžovský, Photo- and electroproduction of K^+ with a unitarity-restored isobar model, Phys. Rev. C 97, 025202 (2018).
- D. Skoupil and P. Bydžovský, Photoproduction of KΛ on the proton, Phys. Rev. C 93, 025204 (2016).
- H. Akaike, A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19 (6): 716–723 (1974).
- C. M. Bishop, Pattern Recognition and Machine Learning. Springer (2006).
- J. Landay M. Döring, C. Fernández-Ramírez, B. Hu and R. Molina, Model selection for pion photoproduction, Phys. Rev. C 95, 015203 (2017).
- J. Landay, M. Mai, M. Döring, H. Haberzettl and K. Nakayama, Towards the minimal spectrum of excited baryons. Phys. Rev. D 99, 016001 (2019).
- G. Schwarz, Estimating the dimension of a model, Ann. Stat. 6(2), 461-464 (1978).