Model selection in electromagnetic production of kaons

Dalibor Skoupil

(in collaboration with P. Bydžovský, A. Cieplý, D. Petrellis, and D. Trnková)

Phys. Rev. C **93**, 025204 (2016)
Phys. Rev. C **97**, 025202 (2018)
Phys. Rev. C **104**, 065202 (2021)
Phys. Rev. C **107**, 045206 (2023)

Workshop of Electro- and Photoproduction of Hypernuclei and Related Topics 2024

Nuclear Physics Institute, CAS, Řež, Czech Republic

October 17, 2024

Motivation for the work on kaon photo/electroproduction

- We aim at understanding the baryon spectrum and production dynamics of particles with strangeness at low energies.

- Constituent Quark Model predicts a lot more N^* states than was observed in pion production experiments $\rightarrow$ "missing" resonance problem.

- Models for the description of elementary hyperon electroproduction are a suitable tool for hypernuclear physics calculations (PR C 106, 044609 (2022), PR C 108, 024615 (2023)).

- New good-quality photoproduction data from BGOOD, LEPS, GRAAL, MAMI and (particularly) CLAS collaborations allow us to tune free parameters of the models.

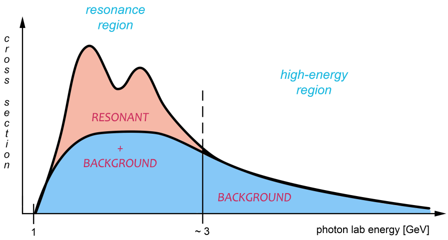
- As the α_s increases with decreasing energy, we cannot use perturbative QCD at low energies $\rightarrow$ need for introducing effective theories and models.

Introduction

Photoproduction process

$$p(p) + \gamma(k) \to K^+(p_K) + \Lambda(p_\Lambda)$$

- Threshold: $E_\gamma^{lab} = 0.911\,\text{GeV}$, $W = 1.609\,\text{GeV}$
- In the lowest order, the reaction is described by the exchange of hadrons.
    - *The 3rd nucleon-resonance region:*
      many resonant states and no dominant one in the kaon photoproduction
      $\to$ need to assume a large number of nucleon resonances with mass $< 2.5\,\text{GeV}$



- Resonance region:
  resonance contributions dominate (N^*)
- Background:
  a plenty of nonresonant contributions
  (p, K, Λ; K^* and Y^*)

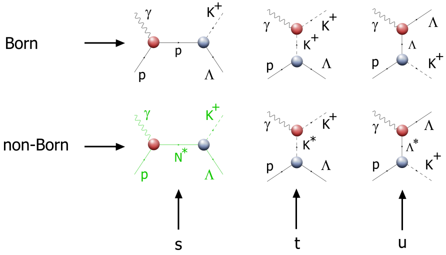
Isobar model

Single-channel approximation

- higher-order contributions (rescattering, FSI) included, to some extent, by means of effective values of the coupling constants

Use of effective hadron Lagrangian

- hadrons either in their ground or excited states
- amplitude constructed as a sum of tree-level Feynman diagrams
  - background and resonant part



- hadronic form factors account for the inner structure of hadrons

Free parameters (couplings, hff's cutoffs) adjusted to experimental data.

Satisfactory agreement with the data in the energy range $W = 1.6 - 2.5\,\text{GeV}$.

Isobar model

Calculation procedure

- Reaction amplitude: sum of s-, t-, and u-channel (non) Born amplitudes

$$\mathbb{M} = \sum_x \mathbb{M}_x, \text{ where } x \equiv s,\ t,\ u,\ N^*,\ K^*,\ Y^*$$

- Each contribution can be rewritten in a compact form

$$\mathbb{M}(p, p_\Lambda, k) = \bar{u}(p_\Lambda)\gamma_5 \left(\sum_{j=1}^{6} \mathcal{A}_j(s,t,u)\mathcal{M}_j \right) u(p),$$

where $\mathcal{A}_j$ are scalar amplitudes and $\mathcal{M}_j$ are gauge-invariant operators, *i.e.* $k_\mu \mathcal{M}_j^\mu = 0$,

$$\begin{aligned}
\mathcal{M}_1 &= \tfrac{1}{2}\left[\not{k}\,\not{\varepsilon} - \not{\varepsilon}\,\not{k}\right], & \mathcal{M}_2 &= (p\cdot\varepsilon) - (k\cdot p)\tfrac{(k\cdot\varepsilon)}{k^2}, \\
\mathcal{M}_3 &= (p_\Lambda\cdot\varepsilon) - (k\cdot p_\Lambda)\tfrac{(k\cdot\varepsilon)}{k^2}, & \mathcal{M}_4 &= \not{\varepsilon}(k\cdot p) - \not{k}(p\cdot\varepsilon), \\
\mathcal{M}_5 &= \not{\varepsilon}(k\cdot p_\Lambda) - \not{k}(p_\Lambda\cdot\varepsilon), & \mathcal{M}_6 &= \not{k}(k\cdot\varepsilon) - \not{\varepsilon}k^2.
\end{aligned}$$

Isobar model
Calculation procedure

- CGLN amplitudes $f_i(k^2, s, t)$

$$\mathbb{M} = \chi_\Lambda^\dagger \mathcal{F} \chi_p; \quad \mathcal{F} = f_1(\vec{\sigma}\cdot\vec{\varepsilon}) - \mathrm{i} f_2(\vec{\sigma}\cdot\hat{p}_K)[\vec{\sigma}\cdot(\hat{k}\times\vec{\varepsilon})] + f_3(\vec{\sigma}\cdot\hat{k})(\hat{p}_K\cdot\vec{\varepsilon}) + f_4(\vec{\sigma}\cdot\hat{p}_K)(\hat{p}_K\cdot\vec{\varepsilon}) + f_5(\vec{\sigma}\cdot\hat{k})(\hat{k}\cdot\vec{\varepsilon}) + f_6(\vec{\sigma}\cdot\hat{p}_K)(\hat{k}\cdot\vec{\varepsilon})$$

where *e.g.*

$$f_1 = N^*[-(W-m_p)\mathcal{A}_1 + (k\cdot p)\mathcal{A}_4 + (k\cdot p_\Lambda)\mathcal{A}_5 - k^2\mathcal{A}_6]$$

- Response functions, *e.g.* transverse cross section

$$\frac{d\sigma}{d\Omega} = \sigma_T = C\left\{|f_1|^2 + |f_2|^2 - 2\,\mathrm{Re}\, f_1 f_2^* \cos\theta_K + \sin^2\theta_K\left[\frac{1}{2}(|f_3|^2 + |f_4|^2) + \mathrm{Re}\,(f_1 f_4^* + f_2 f_3^* + f_3 f_4^* \cos\theta_K)\right]\right\},$$

(for other response functions see Z. Phys. A **352** (1995) 327)

Isobar model

Novel features of our isobar model

Exchanges of high-spin resonant states

- non physical lower-spin components removed by appropriate choice of $\mathcal{L}_{int}$

$$V_S^{\mu}\,\mathcal{P}^{(1/2)}_{ij,\mu\nu}\,V_{EM}^{\nu} = 0$$

Energy-dependent decay widths of nucleon resonances $\rightarrow$ restoration of unitarity

$$\Gamma(\vec{q}) = \Gamma_{N^*}\,\frac{\sqrt{s}}{m_{N^*}}\,\sum_i x_i \left(\frac{|\vec{q}_i|}{|\vec{q}_i^{\,N^*}|}\right)^{2l+1}\,\frac{D(|\vec{q}_i|)}{D(|\vec{q}_i^{\,N^*}|)},$$

Extension from photoproduction to electroproduction

- Phenomenological form factors in the electromagnetic vertex
- Longitudinal couplings of N^*'s to γ^* (crucial at small Q^2)

$$V^{EM}(N^*_{1/2}p\gamma) = -\,i\,\frac{g_3^{EM}}{(m_R+m_p)^2}\Gamma_{\mp}\,\gamma_\beta\,\mathcal{F}^{\beta},$$

$$V^{EM}_{\mu}(N^*_{3/2}p\gamma) = -\,i\,\frac{g_3^{EM}}{m_R(m_R+m_p)^2}\gamma_5\Gamma_{\mp}\,(\not{q}\,g_{\mu\beta} - q_\beta\gamma_\mu)\,\mathcal{F}^{\beta},$$

$$V^{EM}_{\mu\nu}(N^*_{5/2}p\gamma) = -\,i\,\frac{g_3^{EM}}{(2m_p)^5}\Gamma_{\mp}(q_\alpha q_\beta g_{\mu\nu} + q^2 g_{\alpha\mu}g_{\beta\nu} - q_\alpha q_\nu g_{\beta\mu} - q_\beta q_\nu g_{\alpha\mu})\,p^{\alpha}\,\mathcal{F}^{\beta}.$$

Fitting the data in the $K^+\Lambda$ channel

Minimization of $\chi^2/\text{n.d.f.}$ with help of MINUIT library

Resonance selection

- *s* channel: spin-1/2, 3/2, and 5/2 N^* with mass $< 2.5\,\text{GeV}$;
- *t* channel: $K^*(892)$, $K_1(1272)$
- *u* channel: $Y^*(1/2)$ and $Y^*(3/2)$

Free parameters ($\approx 30 + 10$):

- SU(3)_f : $-4.4 \leq g_{K\Lambda N}/\sqrt{4\pi} \leq -3.0$,
  $0.8 \leq g_{K\Sigma N}/\sqrt{4\pi} \leq 1.3$
- K^*'s have vector and tensor couplings
- spin-1/2 resonance $\to$ 1 parameter;
  spin-3/2 and 5/2 resonance
  $\to$ 2 parameters
- 2 cut-off parameters for the hff
- 1 longitudinal coupling for each N^*
- 2 cut-off parameters for the emff of K^* and K_1

Experimental data

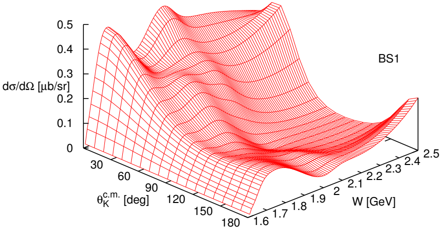
3383 $p(\gamma, K^+)\Lambda$ data

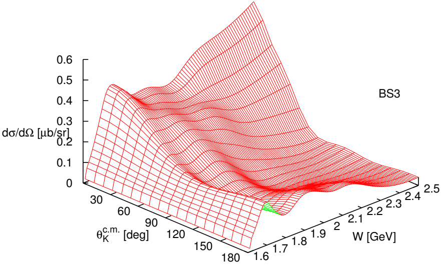
- cross section for $W < 2.355\,\text{GeV}$
  (CLAS 2005 & 2010; LEPS, Adelseck-Saghai)
- hyperon polarisation for $W < 2.225\,\text{GeV}$
  (CLAS 2010)
- beam asymmetry (LEPS)

171 $p(e, e'K^+)\Lambda$ data

- $\sigma_U, \sigma_T, \sigma_L, \sigma_{LT'}, \sigma_K$

Resulting models for the $K^+\Lambda$ photo- and electroproduction





BS1 model ($\chi^2/n.d.f. = 1.64$)

- $S_{11}(1535)$, $S_{11}(1650)$, $F_{15}(1680)$, $P_{13}(1720)$, $F_{15}(1860)$, $D_{13}(1875)$, $F_{15}(2000)$;
- $K^*(892)$, $K_1(1272)$;
- $\Lambda(1520)$, $\Lambda(1800)$, $\Lambda(1890)$, $\Sigma(1660)$, $\Sigma(1750)$, $\Sigma(1940)$;
- multidipole form factor:

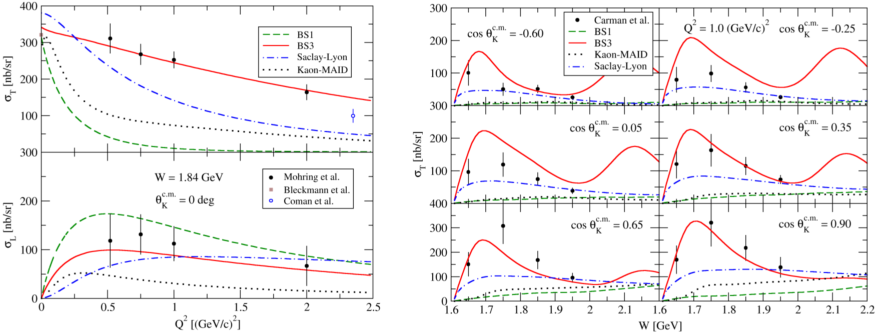
  $\Lambda_{bgr} = 1.88\,\text{GeV}$, $\Lambda_{res} = 2.74\,\text{GeV}$

BS3 model ($\chi^2/n.d.f. = 1.74$)

- $S_{11}(1535)$, $S_{11}(1650)$, $F_{15}(1680)$, $P_{11}(1710)$, $P_{13}(1720)$, $F_{15}(1860)$, $D_{13}(1875)$, $P_{13}(1900)$, $F_{15}(2000)$, $D_{13}(2120)$;
- $K^*(892)$, $K_1(1272)$;
- $\Lambda(1405)$, $\Lambda(1600)$, $\Lambda(1890)$, $\Sigma(1670)$;
- dipole form factor:

  $\Lambda_{bgr} = 1.24\,\text{GeV}$, $\Lambda_{res} = 0.89\,\text{GeV}$

Transverse, σ_T, and longitudinal, σ_L, cross sections of $p(e, e'K^+)\Lambda$



Extension from photo- to electroproduction

- BS1: naive extension by adding em. form factors only
- BS3: em. form factors and longitudinal couplings of N^*'s to γ^* added

New fits for $K^+\Sigma^-$ channel

χ^2 minimization and overfitting

Fitting procedure with MINUIT library: **minimizing the** χ^2

$$\chi^2 = \sum_{i=1}^{N} \frac{[d_i - p_i(c_1, \ldots, c_n)]^2}{\sigma_{d_i}^2},$$

$(c_1, \ldots, c_n)$ - set of free parameters, $(d_1, \ldots, d_N)$ - set of data points, p_i - theory, σ_{d_i} - error

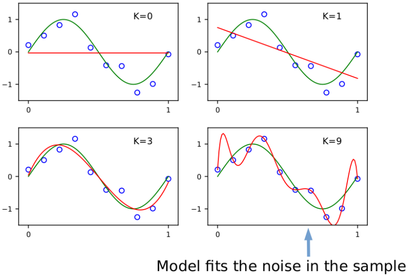
Problem: χ^2 minimization cannot prevent overfitting

Example: polynomial curve fitting

- $f(x, \mathbf{w}) = w_0 + w_1 x + w_2 x^2 + \cdots + w_k x^k$
- increasing order of polynomial k fits the data well...

  ...but gives only poor description of the function which generated them...

  ...and may fail to generalize to new data
- **Occam's razor** (law of parsimony): simpler models should be preferred



New fits of $K^+\Sigma^-$ channel

Least Absolute Shrinkage and Selection Operator (LASSO)

Remedy to the overfitting issue: regularization

- introduce a penalty term to the $\chi^2 \rightarrow$ penalization of large parameter values

$$\chi^2_P = \chi^2 + P(\lambda)$$

- penalty term: $P(\lambda) = \lambda^4 \sum_{i=1}^{N_{res}} |g_i|$

  λ - regularization parameter, g_i - resonances' couplings

- LASSO forces some of the parameters to zero
  $\rightarrow$ selection of a subset of the fit parameters

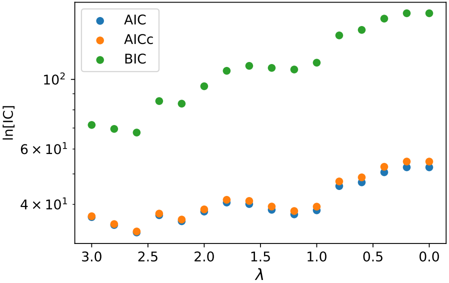
- λ controls the strength of the penalty and thus the complexity of the model
  $\rightarrow$ higher powers of λ allow fine sampling of the region of small λ

New fits of $K^+\Sigma^-$ channel

Information criteria:

- Akaike information criterion
  $\mathrm{AIC} = 2n_i + \chi^2_P$
- Corrected Akaike information criterion
  $\mathrm{AICc} = \mathrm{AIC} + \frac{2n_i(n_i+1)}{N-n_i-1}$
- Bayesian information criterion
  $\mathrm{BIC} = n_i \ln(N) + \chi^2_P$

  n_i - no. of parameters corresponding to λ_i
  N - number of data points



Applying the information criteria – forward selection

1. start with the full model: parameters initialized within $\langle -1; +1 \rangle$; use λ_{max}
2. perform LASSO χ^2_P minimization and compute IC
3. in each run reduce λ and run LASSO with the values of the previous run as starting values
4. repeat until λ_{min} is reached

Optimal λ occurs at the minimum of the IC.

New fits of $K^+\Sigma^-$ channel

Fitting procedure

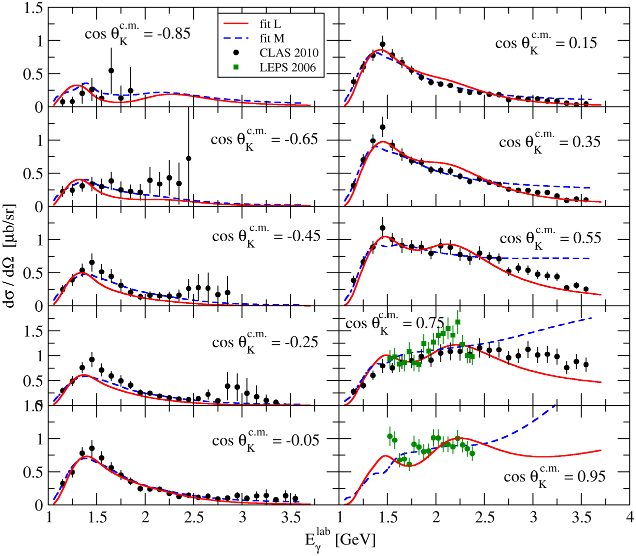
- resonance selection: motivation from previous analysis of $K^+\Lambda$ channel
- non resonant part: Born terms and exchanges of K^* and K_1 and Σ^*'s
- resonant part: exchanges of N^*'s and Δ^*'s in the s channel
- around 600 data utilized to fit $\leq$ 25 parameters
  (new CLAS data on Σ $\rightarrow$ main motivation for this study)
- result with the smallest $\chi^2/\text{ndf} = 2.3 \rightarrow$ **fit M** (25 parameters, 14 resonances)
- **LASSO** applied at fit M: $\chi^2_P/\text{ndf} = 3.4 \rightarrow$ **fit L** (17 parameters, 9 resonances)

Characteristics of models

- only one Δ resonance introduced
- no hyperon resonances needed for reliable data description
- results in very good agreement with the cross-section and beam-asymmetry data
- fit L is very economical

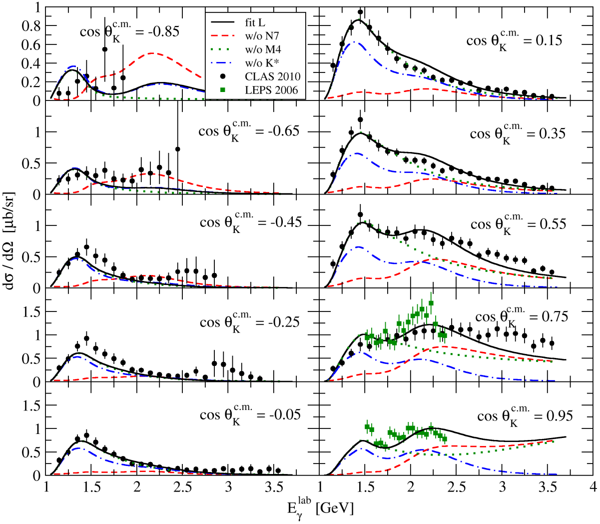
New fits of $K^+\Sigma^-$ channel

Differential cross section in dependence on the photon lab energy



New fits of $K^+\Sigma^-$ channel

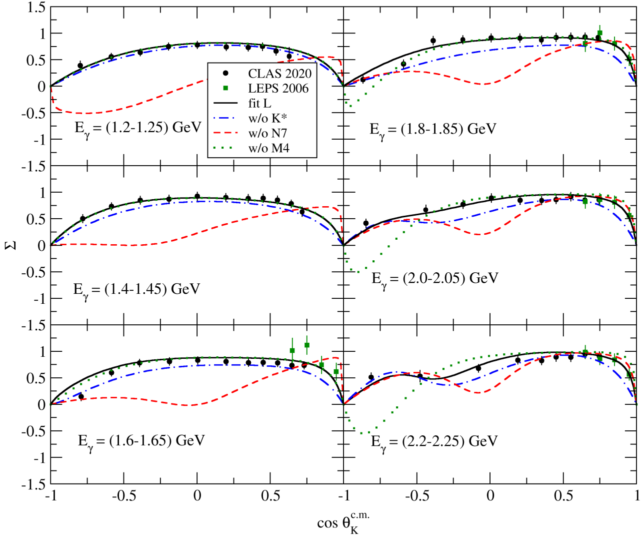
Differential cross section in dependence on the photon lab energy - fit L w/o individual resonances



Notation: N7: N(1720)$3/2^+$, M4: N(2060)$5/2^-$

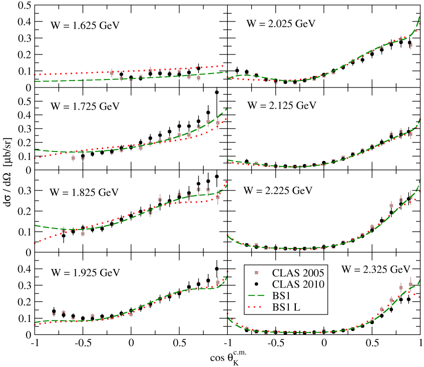
New fits of $K^+\Sigma^-$ channel

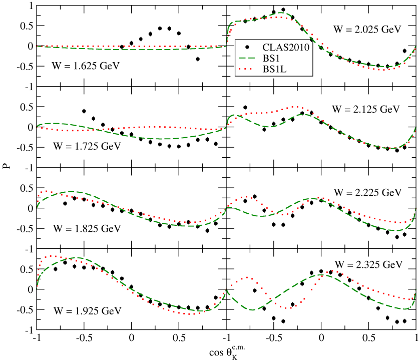
Beam asymmetry in dependence on the kaon center-of-mass angle - fit L w/o individual resonances



LASSO in the $K^+\Lambda$ channel

Selecting a subset of resonances (SOČ – D. Trnková)





Resonances in BS1 and BS1L:

- $S_{11}(1535)$, $S_{11}(1650)$, $F_{15}(1680)$, $P_{13}(1720)$, $F_{15}(1860)$, $D_{13}(1875)$, $F_{15}(2000)$;
- $K^*(892)$, $K_1(1272)$;
- $\Lambda(1520)$, $\Lambda(1800)$, $\Lambda(1890)$, $\Sigma(1660)$, $\Sigma(1750)$, $\Sigma(1940)$

Model	no. of resonances	no. of parameters	χ^2/n.d.f.
BS1	16	31	1.6
BS1L	9	20	3.6

Σ photoproduction channels

A systematic study

Data base (Prog. Part. Nucl. Phys. **111** (2020) 103752):

channel	$d\sigma/d\Omega$	$P, \Sigma, T, \ldots$	%
$\gamma + p \rightarrow K^+ + \Sigma^0$	6681	1465	92.67
$\gamma + n \rightarrow K^+ + \Sigma^-$	429	36	5.29
$\gamma + p \rightarrow K^0 + \Sigma^+$	116	57	1.97
$\gamma + n \rightarrow K^0 + \Sigma^0$	0	6	0.07

Strategy:

1. fit parameters in the $K^+\Sigma^0$ channel $\rightarrow$ select a basic set of resonances
2. do fits in other channels $\rightarrow$ modify the resonance set (if necessary)
3. **OR** relate as many couplings as possible among the channels (using isospin symmetry, helicity amplitudes)

... $K^+\Sigma^0$ channel comprises more than 90% of the available data, shouldn't we consider all the other channels as predictions?

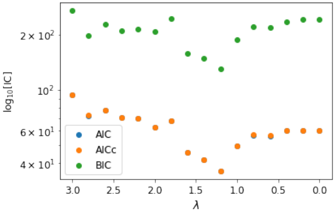
New fits of $K^+\Sigma^0$ channel

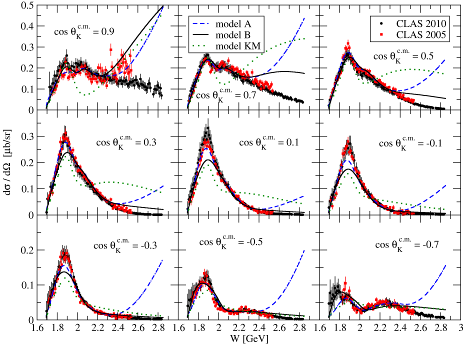
Models A (w/o LASSO) and B (w/ LASSO) ... submitted to PR C

Resonances in models A and B:

- $N3(1535)1/2^-$, $N4(1650)1/2^-$, $N9(1680)5/2^+$, $N7(1720)3/2^+$, $P5(1820)5/2^+$, $M4(1860)1/2^-$, $P4(1875)3/2^-$, $P2(1900)3/2^+$, $\Delta(1900)3/2^+$, $P3(2000)5/2^+$, $M2(2300)1/2^+$;
- $K^*(892)$, $K_1(1272)$;
- $\Lambda(1890)$, $\Sigma(1660)$.

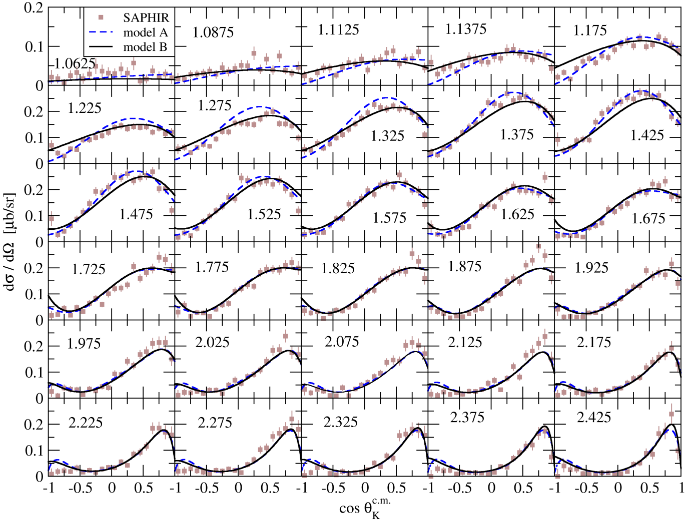
Model	no. of resonances	no. of parameters	χ^2/n.d.f.
A	15	29	2.02
B	7	14	4.67





New fits of $K^+\Sigma^0$ channel

Models A (w/o LASSO) and B (w/ LASSO) – comparison with SAPHIR data



Refitting the model's parameters in the $K^+\Lambda$ channel

Ridge regression and cross validation for suppressing hyperon couplings

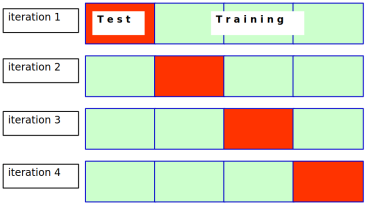
Why refit?

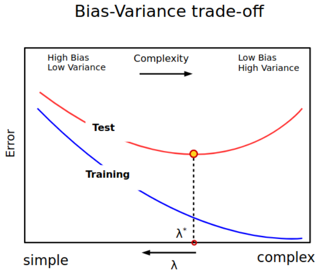
- include recent measurements of polarization observables (PR C **93**, (2016) 065201)
- need to investigate more the role of hyperon resonances in *KY* photoproduction
- large values of hyperon couplings: ridge regression to suppress them during the fitting procedure

Ridge regularization

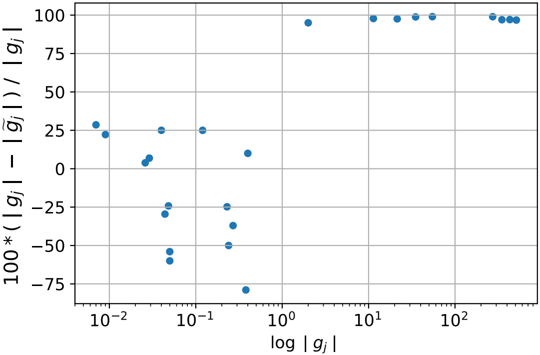
- penalized χ^2_P: $\chi^2_P = \chi^2 + \lambda^4 \sum_{i=1}^{n_\Lambda} g_i^2$, ($n_\Lambda$ = no. of Y couplings)
- parameter values reduced but they are *not* reduced to zero

Cross validation (here 4-fold)





Relative percentage reduction of the resonance couplings



BS2

Tag	Resonance	g_1	g_2
L1	$\Lambda(1405)\ 1/2$	9.67	–
S1	$\Sigma(1660)\ 1/2^{+}$	-8.09	–
L4	$\Lambda(1800)\ 1/2^{-}$	-11.55	–
S4	$\Sigma(1940)\ 3/2^{-}$	-0.86	0.18

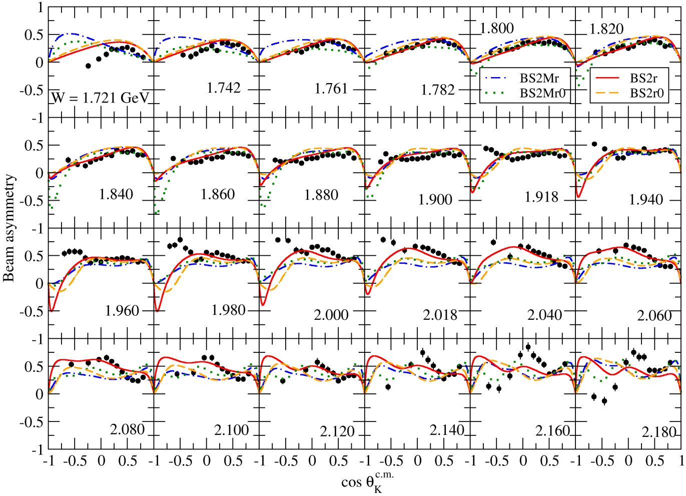
BS2r

2.624	
−5.925	
−1.409	
−0.685	0.079

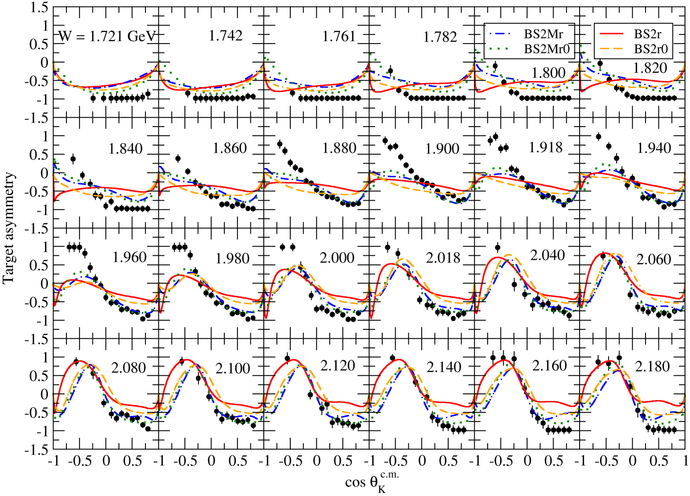
g_j - values from the unregularized fitting

$\tilde{g}_j$ - values after performing Ridge regularization

$K^+\Lambda$ channel: beam asymmetry Σ



$K^+\Lambda$ channel: target asymmetry T



Summary

New version of isobar model for the $K^+\Lambda$ channel

- available for calculations online at:

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

  ...sadly, this is no longer true

Description extended from the $K^+\Lambda$ channel to the $K^+\Sigma^-$ channel.

First fits on the $K^+\Sigma^0$ channel; analysis of other Σ photoproduction channels soon...

Regularization methods introduced as a remedy for overfitting and as model selection tools.

Outlook

- testing the models in the calculations for hypernucleus production
  (PR C 106, 044609 (2022), PR C 108, 024615 (2023))
- performing an analysis of Σ photoproduction channels
- extending the analysis of electroproduction beyond $Q^2 = 1$ GeV^2
- studying the production of Ξ hypernuclei
- preparing amplitude for η' photoproduction

Thank you for your attention!