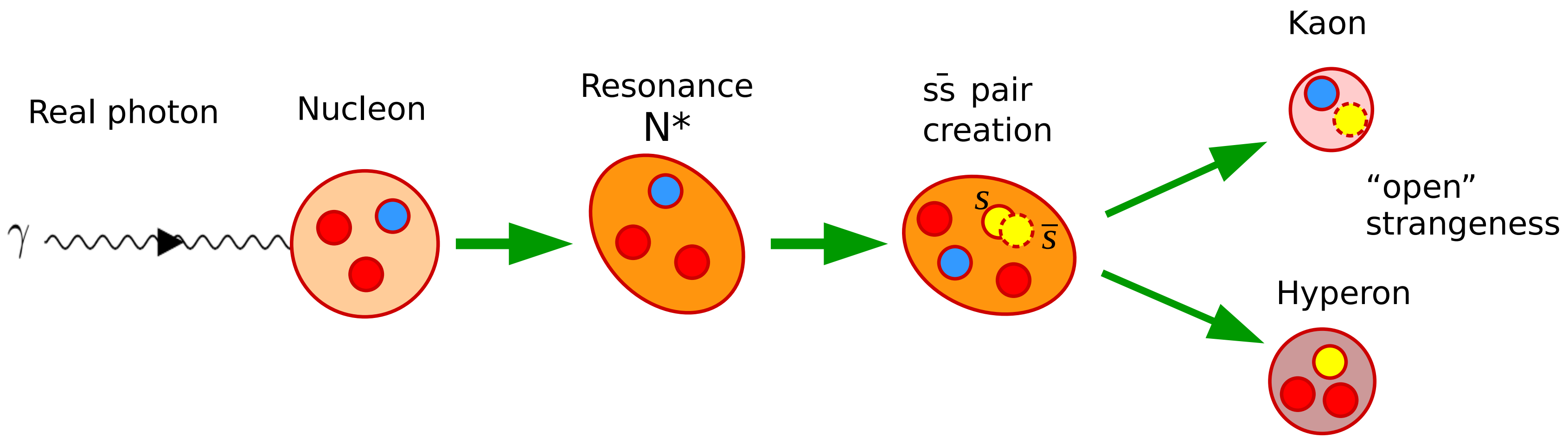
Applying Ridge regularization in the study of $K^+\Lambda$ photoproduction

D. Petrellis and D. Skoupil

Nuclear Physics Institute, Czech Academy of Sciences

Photoproduction of kaons and hyperons off nucleons



$$\gamma + p \rightarrow K^+ + \Lambda$$

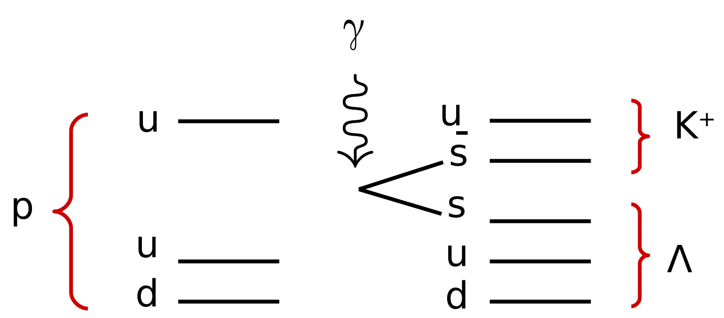
$$\gamma + p \rightarrow K^+ + \Sigma^0$$

$$\gamma + p \rightarrow K^0 + \Sigma^+$$

$$\gamma + n \rightarrow K^0 + \Lambda$$

$$\gamma + n \rightarrow K^0 + \Sigma^0$$

$$\gamma + n \rightarrow K^+ + \Sigma^-$$



Motivation

- Why $K\Lambda$ photoproduction?
  - “missing” resonances: possible decays to kaon-hyperon channels
  - Information on elementary process →
    important for predictions on production of Λ hypernuclei
- Why refit with Ridge?
  - Include recent measurements of *polarization observables* from CLAS
    (C. A. Paterson *et al*. Phys. Rev. C 93, 065201 2016)
  - Need to investigate role of hyperon resonances in $K\Lambda$ photoproduction
  - Isobar models consistently yield large values of *hyperon couplings* →
    Ridge regression to suppress them.

General features of Isobar models

- interactions described by means of effective degrees of freedom: hadrons
- The strength of the interactions → couplings = parameters to be fitted to data
- amplitude = sum of tree-level Feynman diagrams

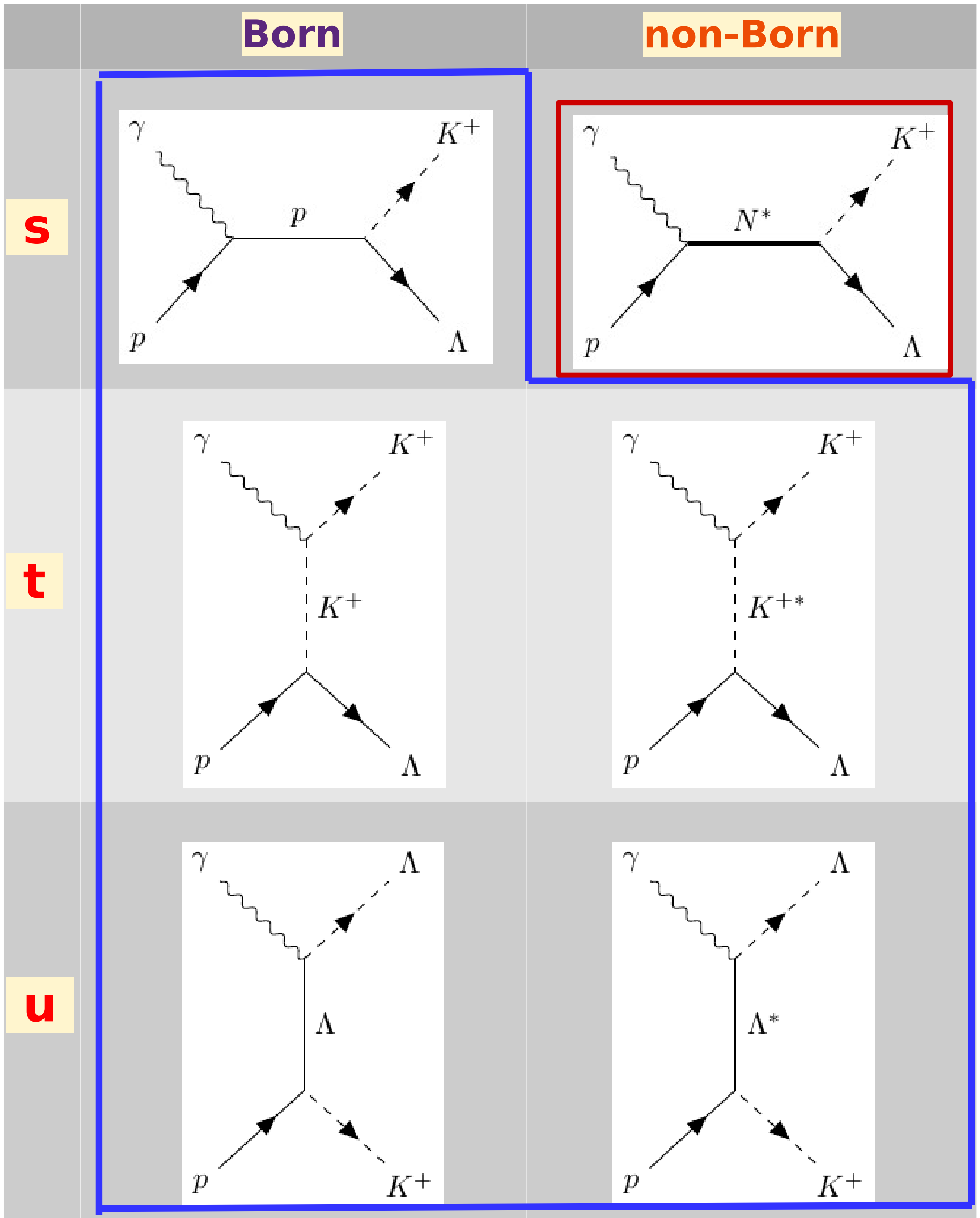
  single-channel: intermediate channels (2nd and higher orders)

  not taken into account explicitly →

  → coupling constants: *effective* values
- Saclay-Lyon, MAID & Kaon-MAID, Gent, BS1,2,3* models

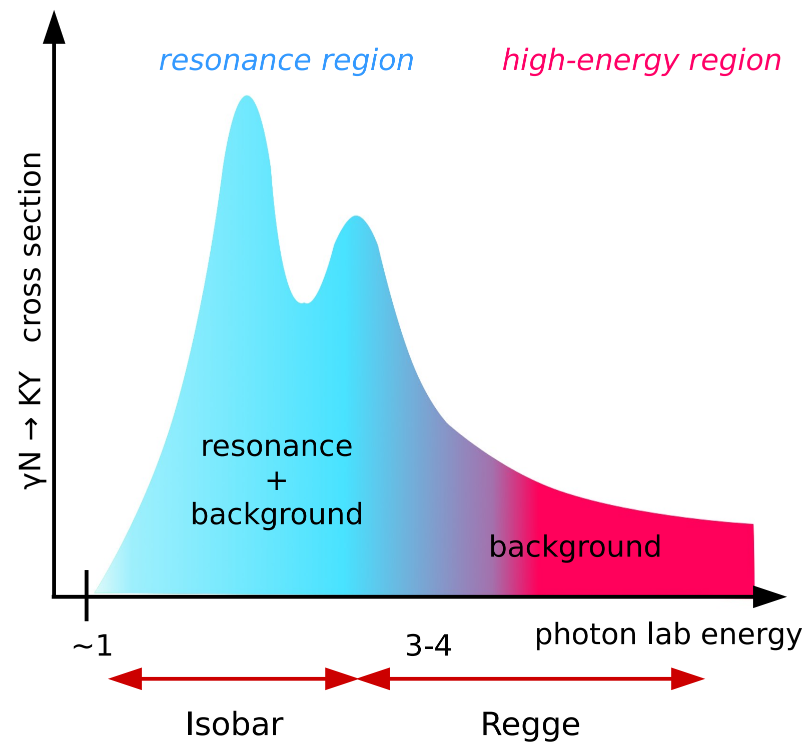
*D. Skoupil and P. Bydzovsky, Phys. Rev. C 93, 025204 (2016)
D. Skoupil and P. Bydzovsky, Phys. Rev. D 97, 025202 (2018)

Tree-level contributions to p(γ,K⁺)Λ



the only resonant diagrams

rest of diagrams
→ background



Parameters and observables

Resonances

masses, widths: from PDG

Parameters to fit

- $g_{K\Lambda p}$
- coupling constants of resonances
(= products of E/M and strong c.c.)
- hadron form factor cutoffs

4640 data points from: CLAS, LEPS

Observables

differential cross sections $d\sigma/d\Omega$
photon beam asymmetries Σ
target polarization asymmetries T
double polarizations O_x , O_z

Minimization with: MINUIT Library
Isobar code available at:
http://www.ujf.cas.cz/en/departments/department-of-theoretical-physics/isobar-model.html

Two sets of resonances (two models)

resonances included in the **BS2** model

Tag	Resonance	Mass [MeV]	Width [MeV]
K*	$K^*(892)$	891.7	50.8
K1	$K_(1272)$	1272	90
N3	$N(1535)\ 1/2^-$	1530	150
N4	$N(1650)\ 1/2^-$	1650	125
P5	$N(1860)\ 5/2^+$	1860	270
N7	$N(1720)\ 3/2^+$	1720	250
P4	$N(1875)\ 3/2^-$	1875	200
P2	$N(1900)\ 3/2^+$	1920	200
P3	$N(2050)\ 5/2^+$	2050	220
N9	$N(1685)\ 5/2^+$	1685	130
N6	$N(1710)\ 1/2^+$	1710	140
L1	$\Lambda(1405)\ 1/2^-$	1405	51
S1	$\Sigma(1660)\ 1/2^+$	1660	100
L4	$\Lambda(1800)\ 1/2^-$	1800	300
S4	$\Sigma(1940)\ 3/2^-$	1940	220

BS2M: variant of BS2

Replace $\Lambda(1800)\ 1/2^+$ (L4)

with $\Lambda(1600)\ 1/2^+$ (L2)

and $\Lambda(1810)\ 3/2^+$ (L5)

and keep the rest of BS2 resonances

based on T. Mart, and N. Nurhadiansyah, Few-Body Syst. 54, 1729–1739 (2013)

LASSO and Ridge regularization

2 ways to deal with overfitting:
- More data
- Regularization

Penalized error function

$$\chi^2 \longrightarrow \chi_P^2(\lambda) = \chi^2 + \lambda \sum_{j=1}^{K} |w_j|^q$$

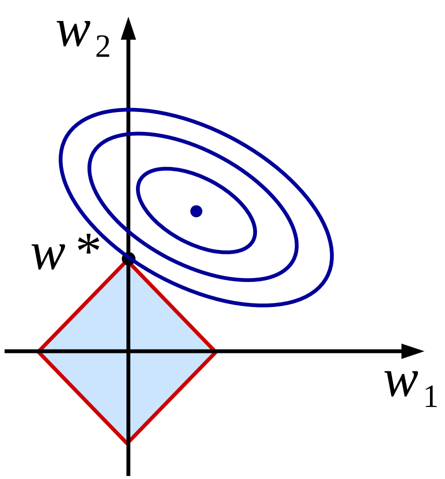
penalty term

λ : regularization parameter

w_j : model parameters → hyperon couplings

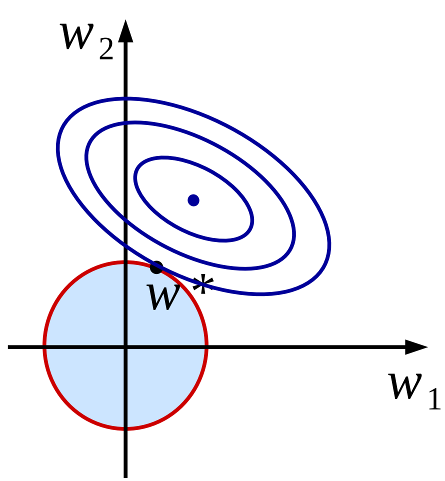
Constrained error minimization → reduces parameter values

LASSO



(q=1)

Ridge



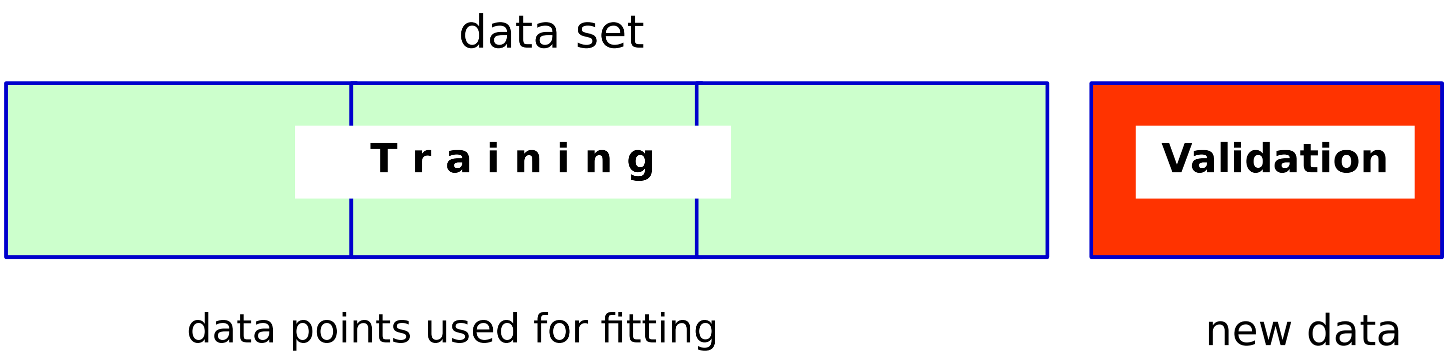
(q=2)

- q=1 → LASSO
  (Least Absolute Shrinkage and Selection Operator):
  - some parameters become zero ($w^*_1 = 0$)
    → suitable for model selection
- q=2 → Ridge:
  - parameter values reduced, but *not* to zero
    → # parameters does not change

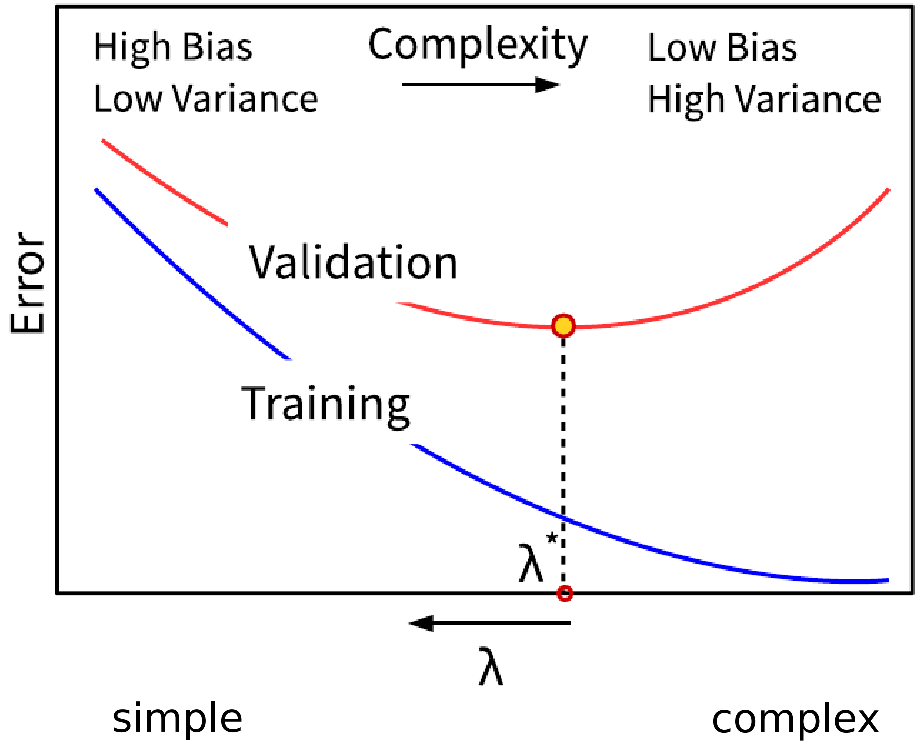
↑ λ → ↑ penalty

how to find the “right” λ?

Training & Validation



- For each λ in $\{ \lambda_{max},\ldots,\lambda_{min} \}$
- Fit model on the training set → Training Error
- Test the fitted model on the validation set → Validation Error
- Repeat, using the fitted parameter values of the last run as starting values, while decreasing λ (increasing complexity = Forward selection)
- Validation Error_{min} → λ^*: optimal λ

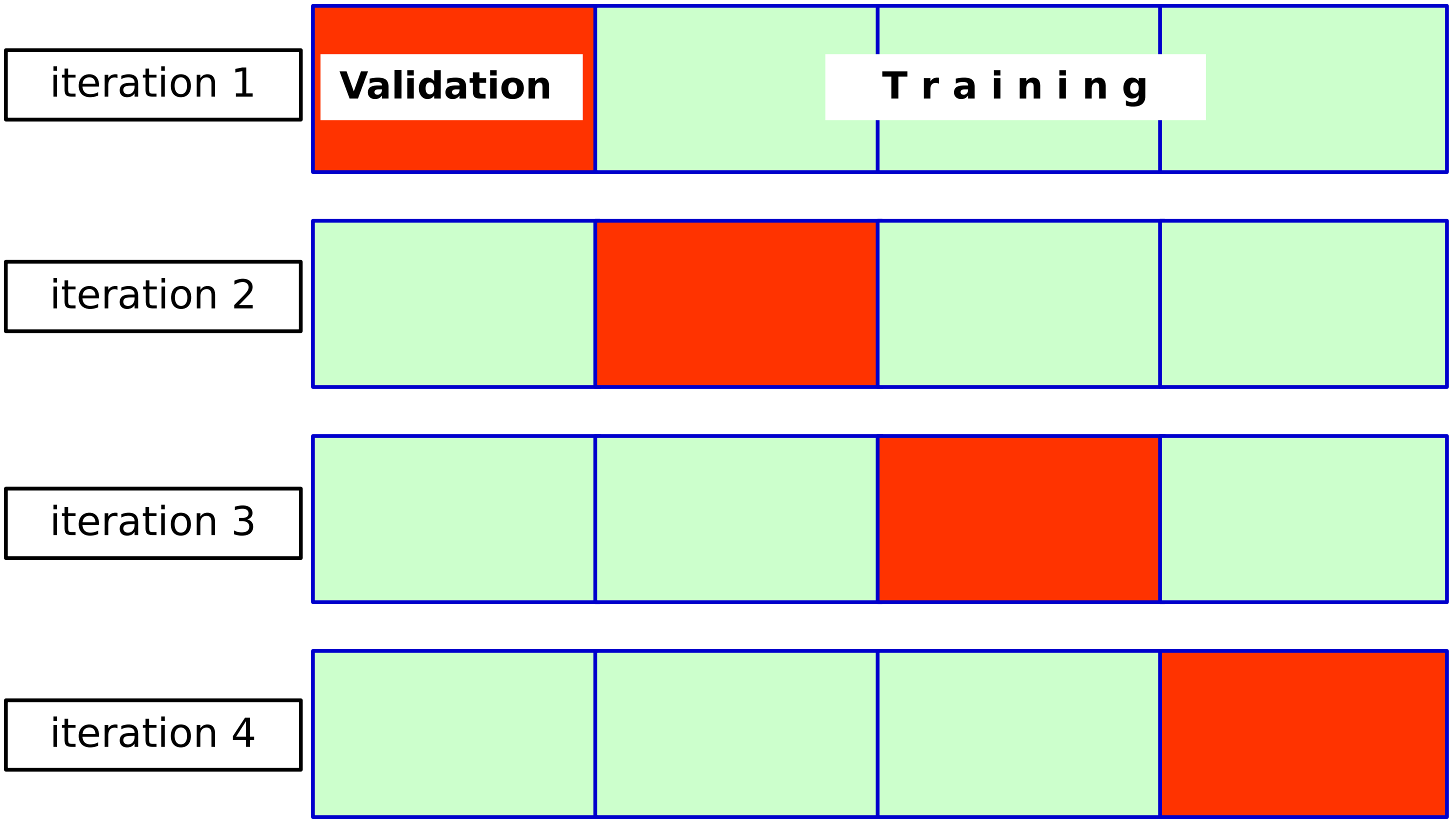


In **Ridge regression** this is the only way to determine λ^* since the number of parameters does not change

Cross-validation

- to avoid selection bias in the choice of Training / Test sets

example:
4-fold cross-validation



after **k** independent runs → average validation errors (CV) over **k** runs for each value of λ

$$\overline{CV}(\lambda) = \frac{1}{k}\sum_{l=1}^{k} CV_l(\lambda)$$

optimal λ: $\lambda^* = \underset{\lambda \in \{\lambda_{\min}, \ldots, \lambda_{\max}\}}{\arg\min} \overline{CV}(\lambda)$

Cross-validation and the “1 standard error” rule

Standard deviation of k validation errors:

$$SD(\lambda) = \sqrt{Var(CV_1(\lambda), ..., CV_k(\lambda))}$$

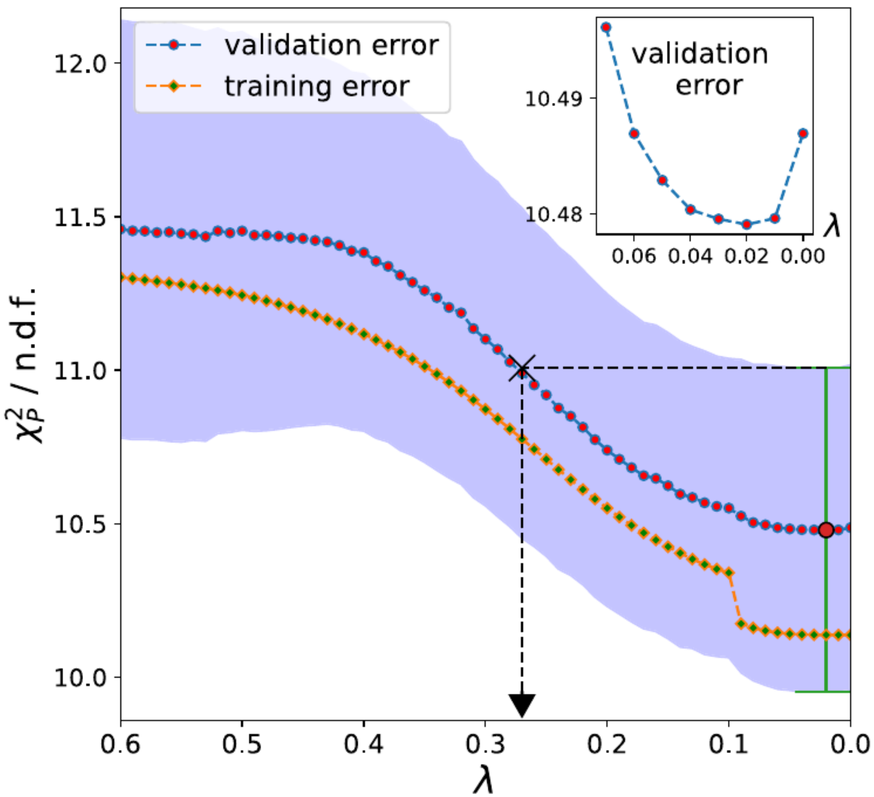
Standard error: $SE(\lambda) = SD(\lambda)/\sqrt{k}$

New optimal λ: $\overline{CV}(\tilde{\lambda}) = \overline{CV}(\lambda^*) + SE(\lambda^*)$

“...choose the most parsimonious model whose error is no more than one standard error above the error of the best model.”
T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning (2009)

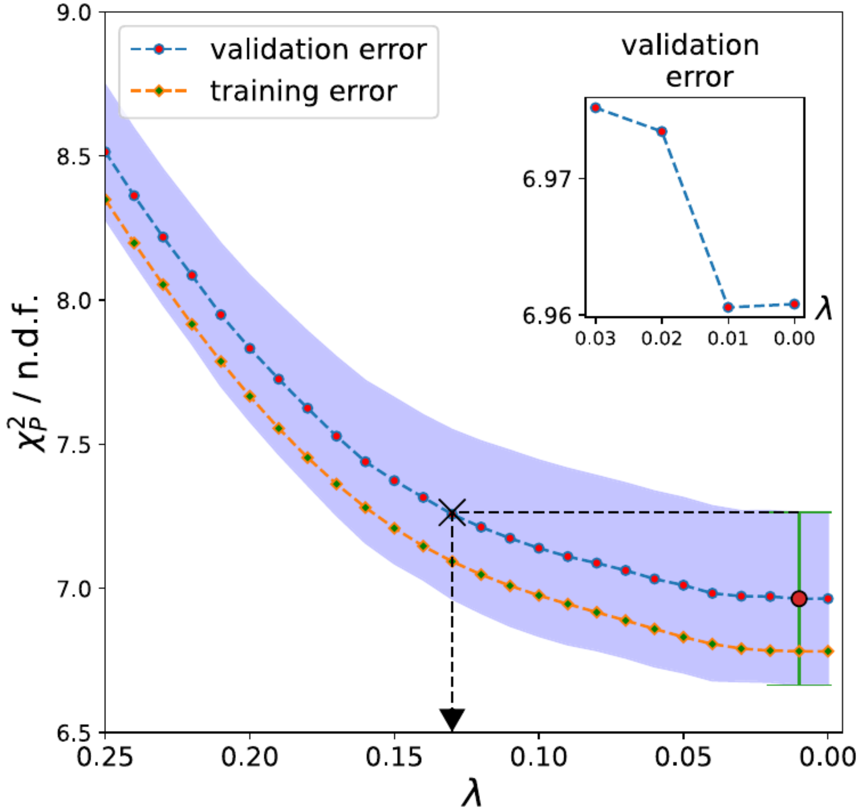
results of 3-fold C.V.

BS2r



$\lambda^* = 0.02 \rightarrow \tilde{\lambda} = 0.27$

BS2Mr



$\lambda^* = 0.01 \rightarrow \lambda = 0.13$

Parameter shrinkage due to Ridge

BS2

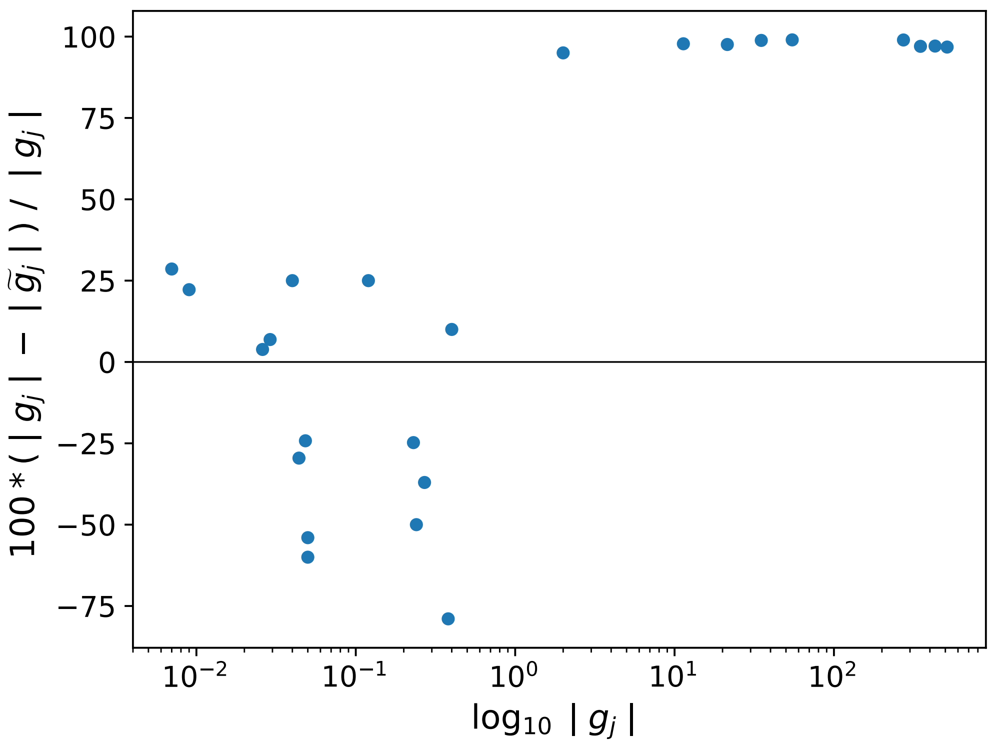
Tag	Resonance	g_1	g_2
L1	$\Lambda(1405)\ 1/2$	9.67	–
S1	$\Sigma(1660)\ 1/2^+$	-8.09	–
L4	$\Lambda(1800)\ 1/2^-$	-11.55	–
S4	$\Sigma(1940)\ 3/2^-$	-0.86	0.18

BS2r

2.624	
-5.925	
-1.409	
-0.685	0.079

hyperon couplings

relative % reduction
in absolute values
in BS2M



D.Petrellis and D. Skoupil, Phys. Rev. C 107, 045206 (2023)

Polarization observables

Single polarization asymmetry

$$A = \frac{d\sigma^{\lambda_X=+s} - d\sigma^{\lambda_X=-s}}{d\sigma^{\lambda_X=+s} + d\sigma^{\lambda_X=-s}}$$

Double polarization asymmetry

$$\frac{d\sigma^{(++)} + d\sigma^{(--)} - d\sigma^{(+-)} - d\sigma^{(-+)}}{d\sigma^{(++)} + d\sigma^{(--)} + d\sigma^{(+-)} + d\sigma^{(-+)}}$$

$$(\text{+-}) \rightarrow (\lambda_A = +s_A, \lambda_B = -s_B)$$

λ_X : polarization of particle *X*

for *X* = *N*, *Y* → λ_X : *spin projection* of nucleon, hyperon on the y(=y') axis

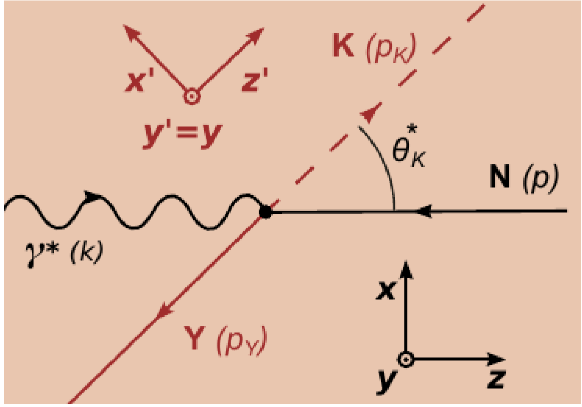
- Target asymmetry: *X* = *N* → **T**
- Recoil asymmetry: *X* = *Y*
- Photon beam asymmetry → **Σ**

- Beam-recoil: *A*,*B* = γ, *Y* → $\mathbf{O_{x'}}$ and $\mathbf{O_{z'}}$
- Beam-target: *A*,*B* = γ, *N*
- Target-recoil: *A*,*B* = *N*, *Y*

$$\Sigma^{\text{lin.}} = \frac{d\sigma^{\perp} - d\sigma^{\parallel}}{2d\sigma^{\text{unpol}}}$$

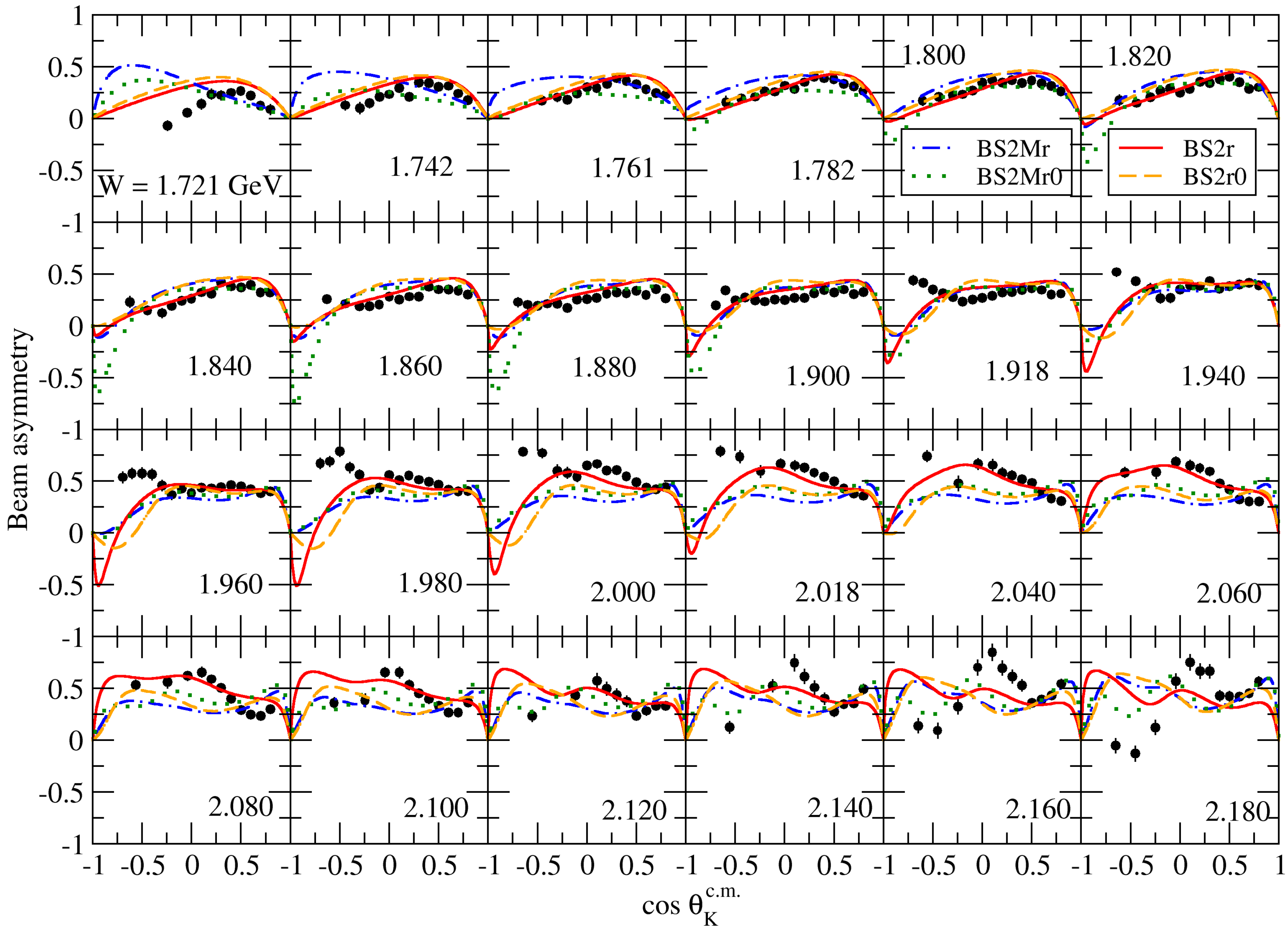
linearly polarized photons

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

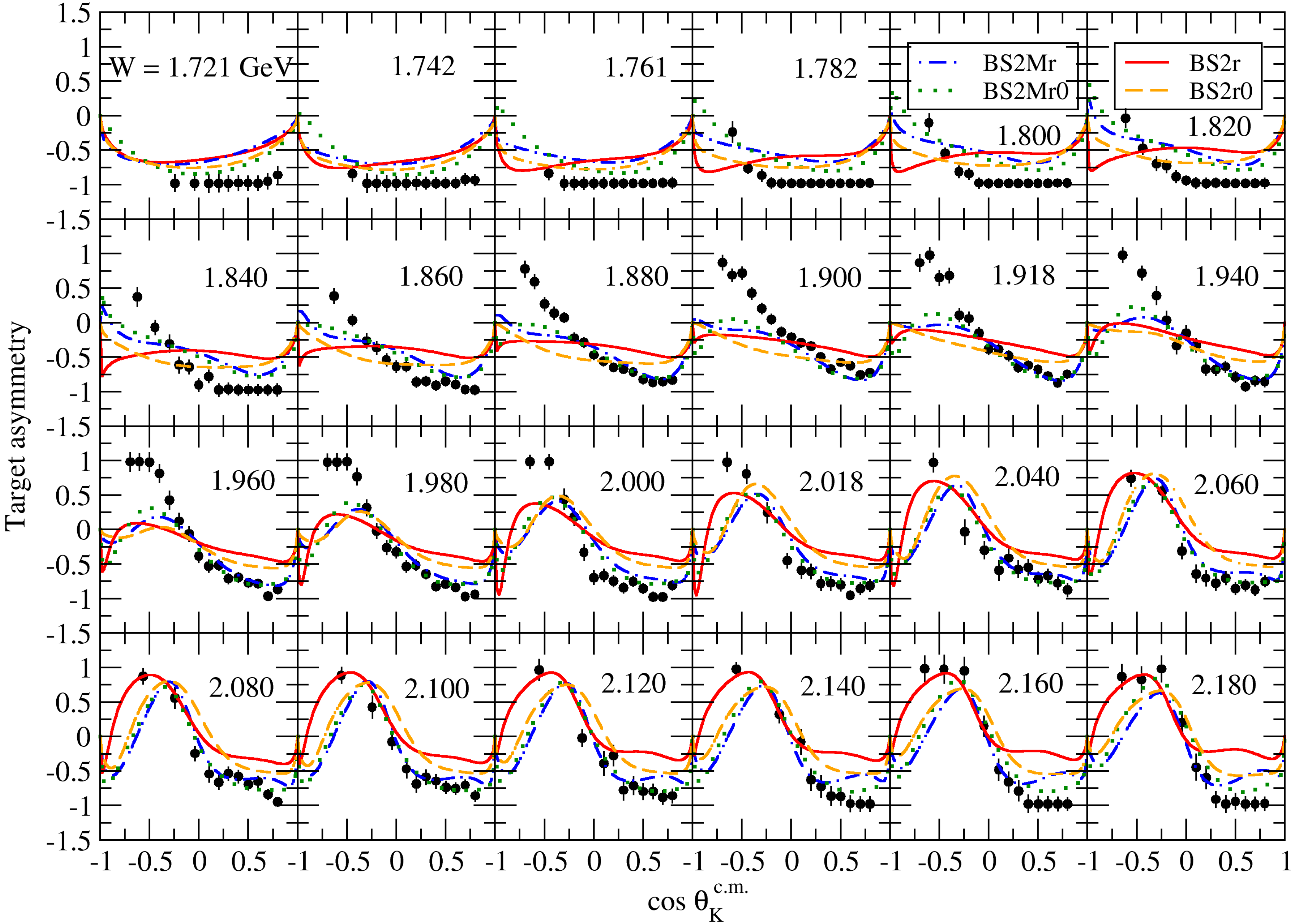


from
L. De Cruz, Bayesian model selection for electromagnetic kaon production in the Regge-plus-resonance framework, PhD Thesis, Ghent University (2012)

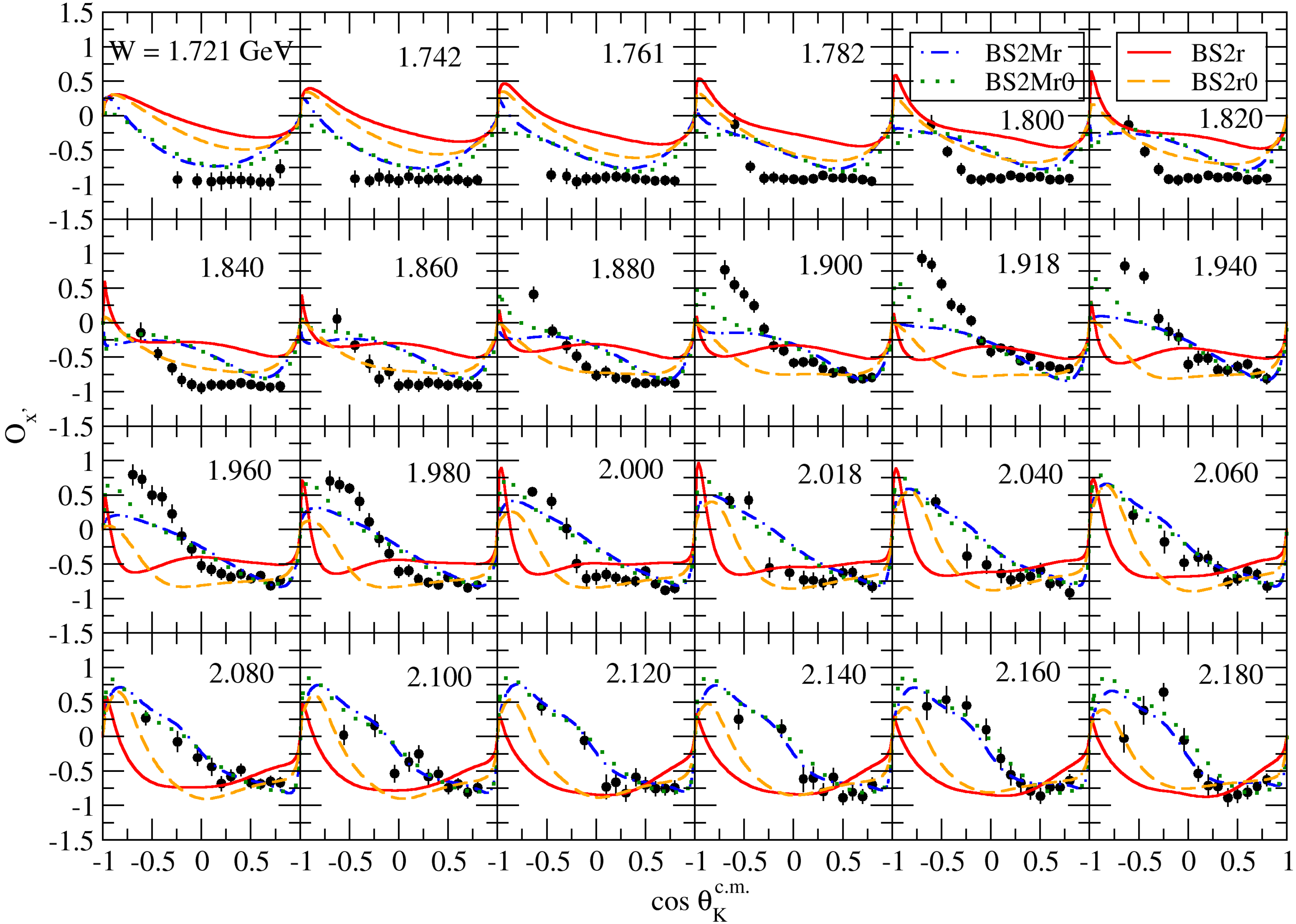
Photon-beam asymmetry



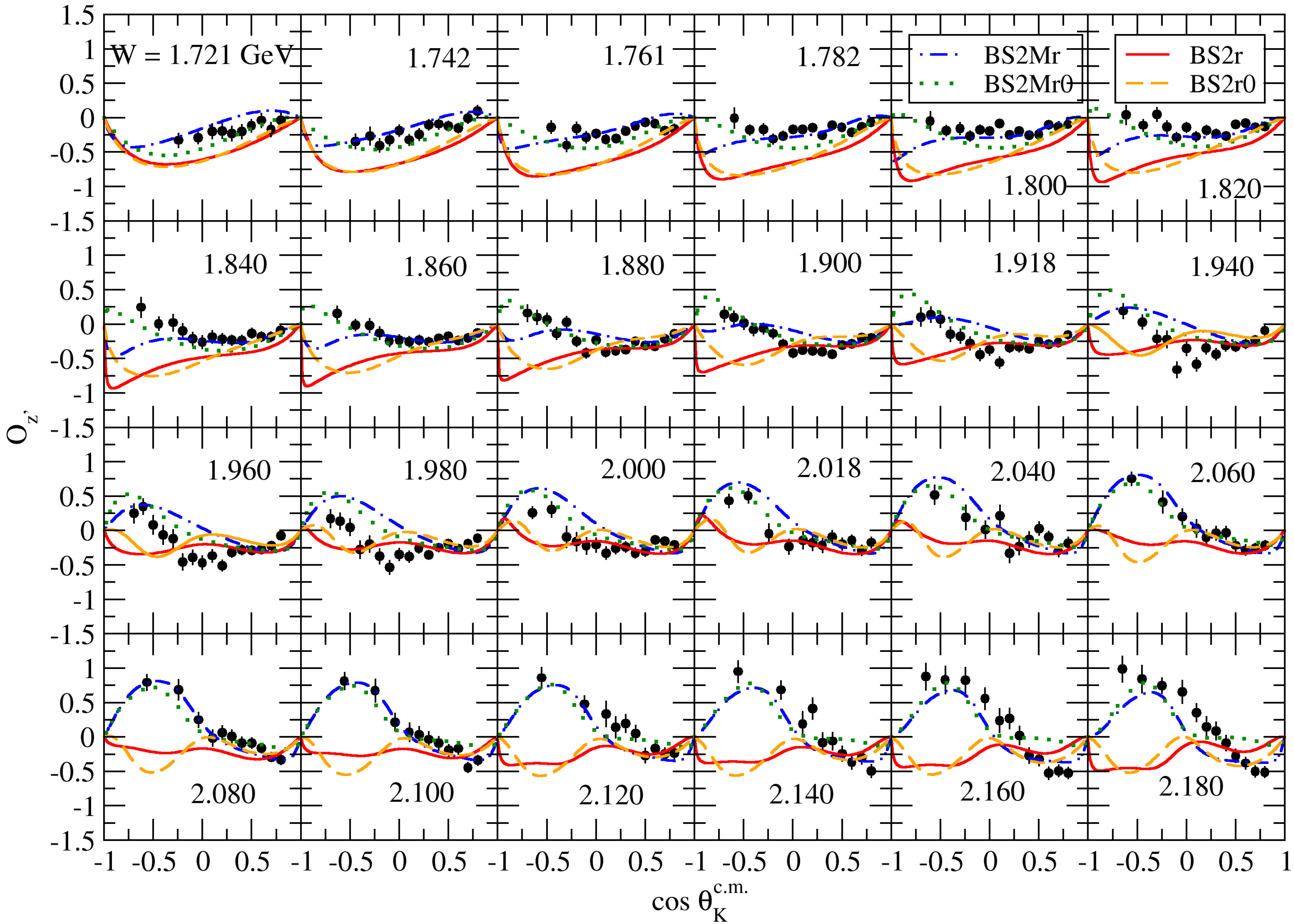
Target asymmetry



Double-polarization asymmetry $O_{x'}$

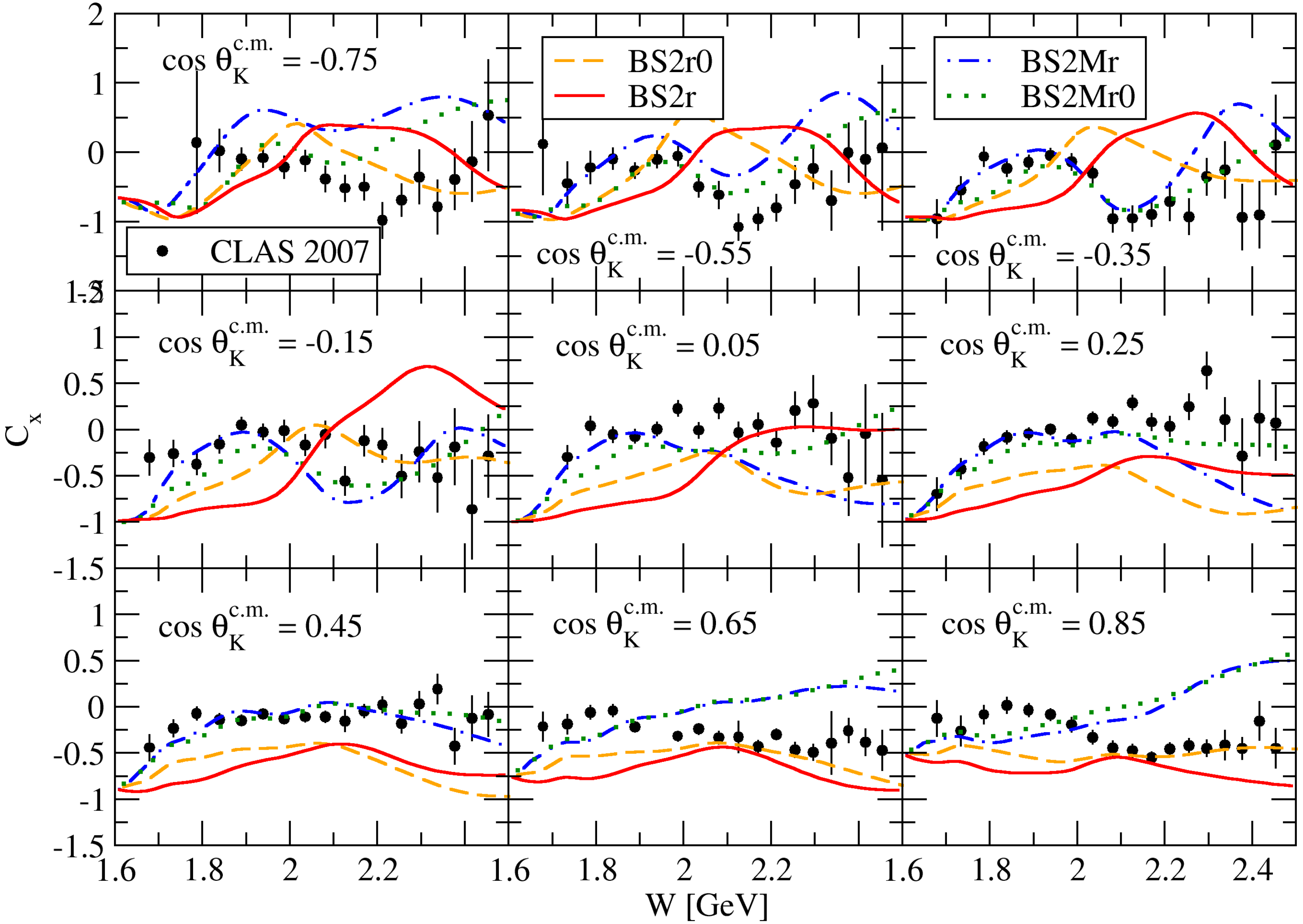


Double-polarization asymmetry $O_{z'}$



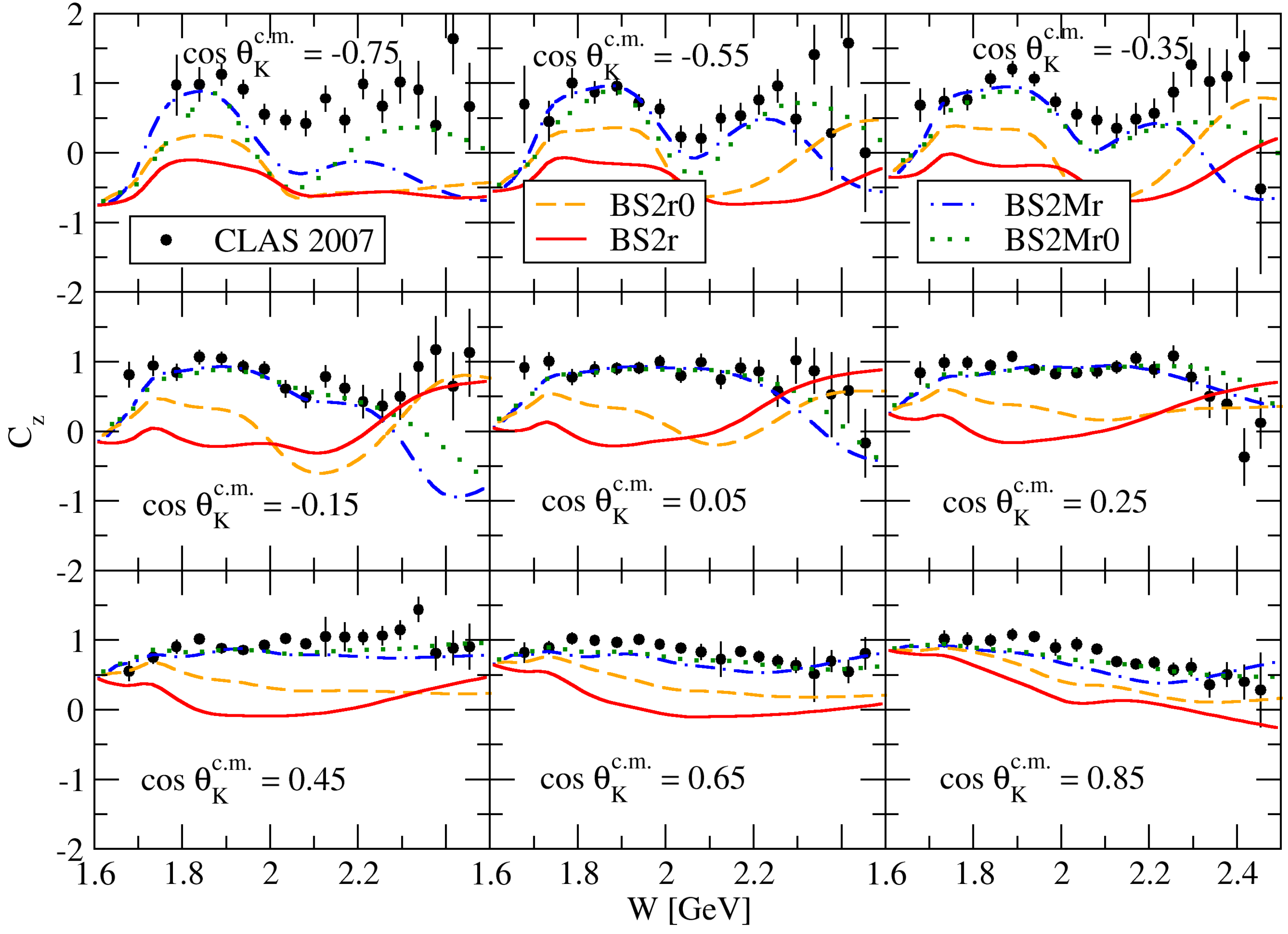
Double-polarization asymmetry C_x

Data not fitted → results = predictions



Double-polarization asymmetry C_z

Data not fitted → results = predictions



Concluding remarks

- Regularization techniques are used to deal with the problem of overfitting by making a model less sensitive to noise and thus improve its predictive power. This is achieved by not allowing the parameters of the model to take extreme values.

- We used Ridge with Cross Validation for parameter *shrinkage* in order to suppress the unphysical values of the hyperon couplings obtained in the description of $K^+\Lambda$ photoproduction.

- With Ridge regression one can work with a certain set of resonances without having to impose artificial limits on the parameters during the fitting.