LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

Constructing model-agnostic likelihoods

A method for the reinterpretation of particle physics results

Lorenz Gärtner[1], Nikolai Hartmann[1], Lukas Heinrich[2], Malin Horstmann[2], Thomas Kuhr[1], Méril Reboud[3], Slavomira Stefkova[4], Danny van Dyk[5]

[1]LMU Munich, [2]TU Munich, [3]Universität Siegen, [4]KIT, [5]IPPP Durham

16.05.2024

ORIGINS Excellence Cluster

Belle II

PUNCH 4NFDI

DFG

Reusability in HEP

Publishing statistical models: Getting the most out of particle physics experiments

Kyle Cranmer[1*], Sabine Kraml[2‡], Harrison B. Prosper[3§] (editors), Philip Bechtle[4], Florian U. Bernlochner[4], Itay M. Bloch[5], Enzo Canonero[6], Marcin Chrzaszcz[7], Andrea Coccaro[8], Jan Conrad[9], Glen Cowan[10], Matthew Feickert[11], Nahuel Ferreiro Iachellini[12,13] Andrew Fowlie[14], Lukas Heinrich[15], Alexander Held[1], Thomas Kuhr[13,16], Anders Kvellestad[17], Maeve Madigan[18], Farvah Mahmoudi[15,19], Knut Dundas Morå[20], Mark S. Neubauer[11], Maurizio Pierini[15], Juan Rojo[8], Sezen Sekmen[22], Luca Silvestrini[23], Veronica Sanz[24,25], Giordon Stark[26], Riccardo Torre[8], Robert Thorne[27], Wolfgang Waltenberger[28], Nicholas Wardle[29], Jonas Wittbrodt[30]

Forum [8], with the current status and updated recommendations presented in Ref. [6]. This paper takes these decade-long efforts to what we argue is the logical conclusion: if we wish to maximize the scientific impact of particle physics experiments, decades into the future, we should make the publication of full statistical models, together with the data to convert them into likelihood functions, standard practice. A statistical model provides the complete mathematical description of an experimental analysis and is, therefore, the appropriate starting

arXiv:2109.04981 [hep-ph]

- Publishing likelihoods is good, but likelihood $\neq$ statistical model.
- Model dependent likelihood ➜ **limited interpretability**

Increase interpretability? Reinterpretation!

What result would we get if we replace signal model A with signal model B?

- Analysis results are based on a statistical model

$$p(\text{data}|\text{model A})$$

- Unfortunately, we cannot say much about

$$p(\text{data}|\text{model B})$$

- Kinematic shape differences affect acceptance / efficiency.
- **What do we need to probe new physics?**

redist

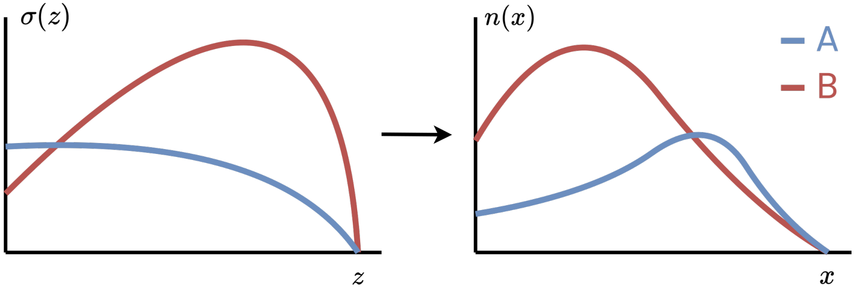
A novel shape-respecting reinterpretation method

arXiv:2402.08417 [hep-ph]

github.com/lorenzennio/redist

Shape-respecting reinterpretation

Kinematic to reconstructed number density

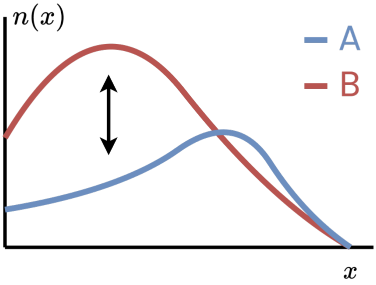


z – kinematic d.o.f.

x – reconstruction / fitting variable(s)

$$n(x) = \int dz\, L\, \varepsilon(x|z)\, \sigma(z) = \int dz\, n(x,z)$$

A new signal distribution



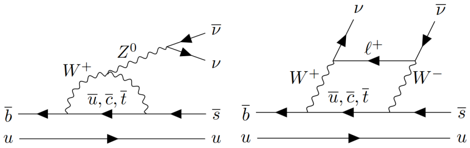
$$n_B(x) = \int dz\, L\, \varepsilon(x|z)\, \sigma_B(z) = \int dz\, L\, \varepsilon(x|z)\, \sigma_A(z) \frac{\sigma_B(z)}{\sigma_A(z)} = \int dz\, \underbrace{n_A(x,z)}_{\text{main object}} w(z).$$

Likelihood + $n(x, z)$ = model-agnostic likelihood

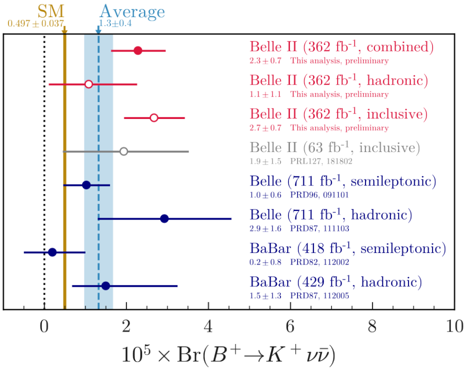
$B^+ \to K^+ \nu\bar{\nu}$

A flavor physics showcase

Why reinterpret $B^+ \to K^+ \nu\bar{\nu}$?



- Suppression of FCNCs in the SM



arXiv:2311.14647 [hep-ex]

➜ BSM effects could substantially affect observables.

Ingredients for reinterpretation

1. Likelihood $L(\boldsymbol{n}, \boldsymbol{a} \mid \boldsymbol{\eta}, \chi)$
2. Joint number density $n(x, z)$
3. Decay kinematics ➜ $w(z)$

1. pyhf statistical model

A statistical model for multi-bin histogram-based analysis and its interval estimation.

Likelihood function for observed event counts $\boldsymbol{n}$ is

$$L(\boldsymbol{n}, \boldsymbol{a} \mid \boldsymbol{\eta}, \boldsymbol{\chi}) = \underbrace{\text{Pois}\left(\boldsymbol{n} \mid \boldsymbol{\nu}(\boldsymbol{\eta}, \boldsymbol{\chi})\right)}_{\text{data likelihood}} \quad \underbrace{c\left(\boldsymbol{a} \mid \boldsymbol{\chi}\right)}_{\text{constraint likelihood}}$$

Expected number of events are

$$\boldsymbol{\nu}(\boldsymbol{\eta}, \boldsymbol{\chi}) = \boldsymbol{\kappa}(\boldsymbol{\eta}, \boldsymbol{\chi}) \left(\boldsymbol{\nu}^0(\boldsymbol{\eta}, \boldsymbol{\chi}) + \boldsymbol{\Delta}(\boldsymbol{\eta}, \boldsymbol{\chi})\right).$$

Use redist for bin weights $\boldsymbol{\kappa}(\boldsymbol{\eta}, \boldsymbol{\chi}) = n_B \, / \, n_A$.

1. Bayesian pyhf statistical model

A statistical model for multi-bin histogram-based analysis and its interval estimation.

Posterior function for observed event counts $\boldsymbol{n}$ is

$$p(\boldsymbol{\eta}, \boldsymbol{\chi} | \boldsymbol{n}, \boldsymbol{a}) \propto \underbrace{\text{Pois}\left(\boldsymbol{n} | \boldsymbol{\nu}(\boldsymbol{\eta}, \boldsymbol{\chi})\right)}_{\text{data likelihood}} \underbrace{p(\boldsymbol{\chi} | \boldsymbol{a})}_{\text{constraint prior}} \underbrace{p(\boldsymbol{\eta})}_{\text{unconstraint prior}}$$

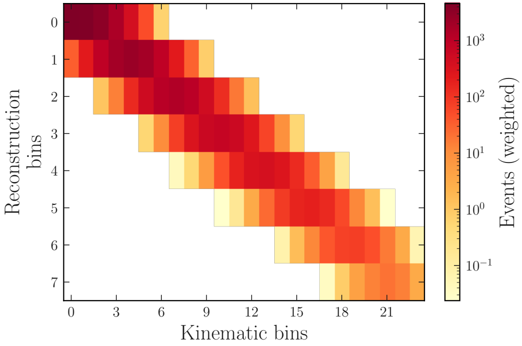
Expected number of events are

$$\boldsymbol{\nu}(\boldsymbol{\eta}, \boldsymbol{\chi}) = \boldsymbol{\kappa}(\boldsymbol{\eta}, \boldsymbol{\chi}) \left(\boldsymbol{\nu}^0(\boldsymbol{\eta}, \boldsymbol{\chi}) + \boldsymbol{\Delta}(\boldsymbol{\eta}, \boldsymbol{\chi})\right).$$

Use redist for bin weights $\kappa(\boldsymbol{\eta}, \boldsymbol{\chi}) = n_B \,/\, n_A$.

2. Joint number density

The main object for reinterpretation, $n_A(x, z)$.
This should be provided by the experimental analysts.



In this example the reconstruction variable is $x = q^2_{rec.}$ and the kinematic variable is $z = q^2_{gen.}$

3. Decay kinematics

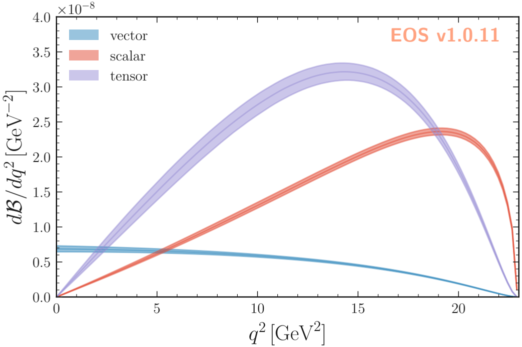
Weak effective theory differential branching ratio dependence on the Wilson coefficients.

$$\frac{d\mathcal{B}(B \to K\nu\bar{\nu})}{dq^2} = \frac{3G_F^2\alpha^2\tau_B}{32\pi^5 m_B^3} \left|V_{ts}^* V_{tb}\right|^2 \sqrt{\lambda_{BK}}\, q^2 \left[\frac{\lambda_{BK}}{24q^2} \left| f_+(q^2) \right|^2 \left|C_{\mathrm{VL}} + C_{\mathrm{VR}}\right|^2 + \frac{\left(m_B^2 - m_K^2\right)^2}{8\left(m_b - m_s\right)^2} \left| f_0(q^2) \right|^2 \left|C_{\mathrm{SL}} + C_{\mathrm{SR}}\right|^2 + \frac{2\lambda_{BK}}{3\left(m_B + m_K\right)^2} \left| f_T(q^2) \right|^2 \left|C_{\mathrm{TL}}\right|^2 \right]$$

for $J^P = 0^-$ kaon states.

[arXiv:2111.04327 [hep-ph]]

Weak effective theory contributions



SM contains only *vectorial* contribution.

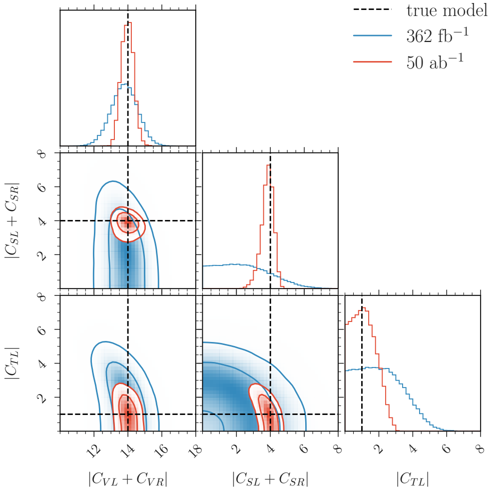
Reinterpretation results
(insights from toy study)

Results from toy study

- Analysis assumes SM (model A)
  - $C_{\mathrm{VL}} \simeq 6.6\,, C_i = 0 \quad \forall \quad i \neq \mathrm{VL}$
- Weak effective theory (models B)
- "Data" contains new physics
- Uniform priors:

$$\begin{aligned} 5 \leq &|C_{\mathrm{VL}} + C_{\mathrm{VR}}| &\leq 20 \\ 0 \leq &|C_{\mathrm{SL}} + C_{\mathrm{SR}}| &\leq 15 \\ 0 \leq &|C_{\mathrm{TL}}| &\leq 15 \end{aligned}$$

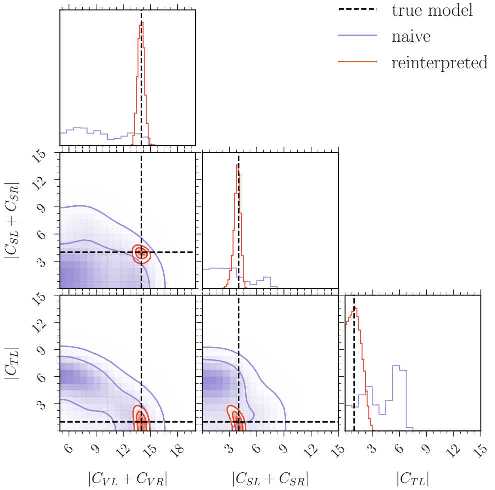
arXiv:2402.08417 [hep-ph]



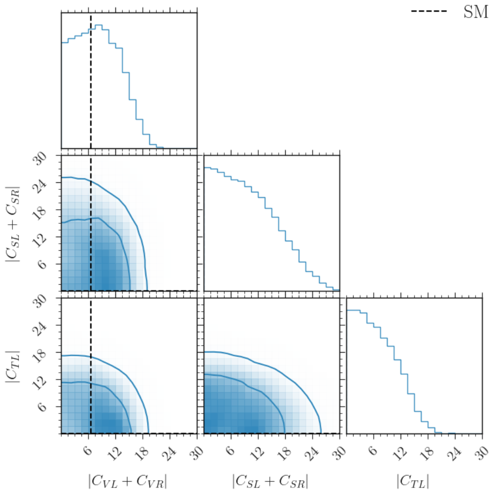
The necessity for reinterpretation

Simplified model reinterpretation, assuming acceptances /efficiencies are independent of kinematic shape differences (e.g. simple BR rescaling).
➜ **biased posterior**

arXiv:2402.08417 [hep-ph]



$B^+ \to K^+ \nu\bar{\nu}$ @ Belle II 63fb^{-1}



Application to the first $B^+ \to K^+ \nu\bar{\nu}$ analysis of 63fb^{-1} Belle II data, Phys.Rev.Lett.127.181802.

Summary

- **Shape-respecting reinterpretation through reweighting**
  - Model-agnostic likelihoods
  - Only requires **likelihood** + $\mathbf{n}(\mathbf{x}, \mathbf{z})$.
  - **FAST**
- **Scientific benefits**
  - **Bias-free inference** on BSM parameters.
  - **Combinations** with other channels and/or experiments.

arXiv:2402.08417 [hep-ph]

Publishing model-agnostic likelihoods is crucial for a bias-free exploitation of experimental results.

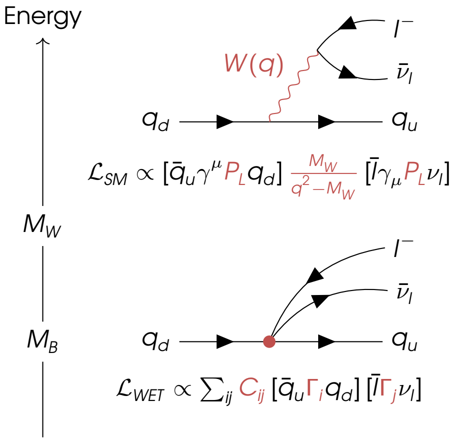
lorenz.gaertner@physik.uni-muenchen.de

Effective field theory

- **High energy collisions**
  W boson potentially on-shell (massive)
- **Lower energy quark decays**
  W boson always off-shell
- **Weak effective theory**
  W is integrated out

$$\mathcal{L}_{SM} \to \mathcal{L}_{WET} = \sum C_i \, \mathcal{O}_i$$

➜ **Parametrization** i.t.o.
Wilson coefficients $\psi = \{C_i\}$.



Weak Effective Theory for $B \to K\nu\bar{\nu}$

Contribution operators

The effective Lagrangian is

$$\mathcal{L}^{WET} = \sum_{X=L,R} C_{\mathrm{VX}}\, \mathcal{O}_{\mathrm{VX}} + \sum_{X=L,R} C_{\mathrm{SX}}\, \mathcal{O}_{\mathrm{SX}} + C_{\mathrm{TL}}\, \mathcal{O}_{\mathrm{TL}} + \text{h.c.}$$

The $d = 6$ contributing operators in and beyond the SM are given by

$$\mathcal{O}_{\mathrm{VL}} = (\overline{\nu_L}\gamma_\mu \nu_L)(\overline{s_L}\gamma^\mu b_L) \qquad \mathcal{O}_{\mathrm{VR}} = (\overline{\nu_L}\gamma_\mu \nu_L)(\overline{s_R}\gamma^\mu b_R)$$

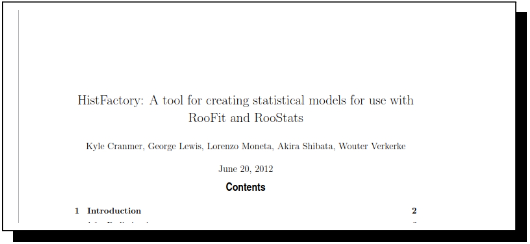
$$\mathcal{O}_{\mathrm{SL}} = \left(\overline{\nu_L^C}\nu_L\right)(\overline{s_R} b_L) \qquad \mathcal{O}_{\mathrm{SR}} = \left(\overline{\nu_L^C}\nu_L\right)(\overline{s_L} b_R)$$

$$\mathcal{O}_{\mathrm{TL}} = \left(\overline{\nu_L^C}\sigma_{\mu\nu}\nu_L\right)(\overline{s_R}\sigma^{\mu\nu} b_L)$$

[arXiv:2111.04327 [hep-ph]]

A statistical model for multi-bin histogram-based analysis and its interval estimation.

pyhf = **py**thonic **H**ist**F**actory



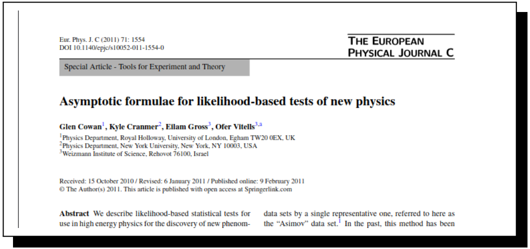
HistFactory: A tool for creating statistical models for use with RooFit and RooStats

Kyle Cranmer, George Lewis, Lorenzo Moneta, Akira Shibata, Wouter Verkerke

June 20, 2012

Contents

Interval estimation based on



Eur. Phys. J. C (2011) 71: 1554
DOI 10.1140/epjc/s10052-011-1554-0

THE EUROPEAN PHYSICAL JOURNAL C

Special Article - Tools for Experiment and Theory

Asymptotic formulae for likelihood-based tests of new physics

Glen Cowan[1], Kyle Cranmer[2], Eilam Gross[3], Ofer Vitells[3,a]
[1]Physics Department, Royal Holloway, University of London, Egham TW20 0EX, UK
[2]Physics Department, New York University, New York, NY 10003, USA
[3]Weizmann Institute of Science, Rehovot 76100, Israel

Received: 15 October 2010 / Revised: 6 January 2011 / Published online: 9 February 2011

Abstract We describe likelihood-based statistical tests for use in high energy physics for the discovery of new phenom- data sets by a single representative one, referred to here as the "Asimov" data set.[1] In the past, this method has been

HistFactory / pyhf statistical model

A statistical model for multi-bin histogram-based analysis and its interval estimation.

Likelihood function for observed event counts $\boldsymbol{n}$ is

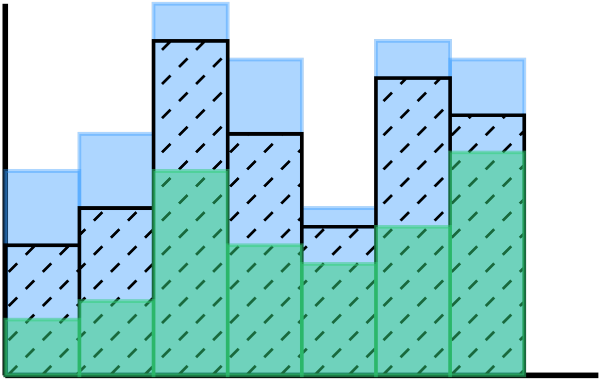
$$L(\boldsymbol{n}, \boldsymbol{a} \mid \boldsymbol{\eta}, \boldsymbol{\chi}) = \underbrace{\prod_{c \in \text{channels}} \prod_{b \in \text{bins}} \text{Pois}\left(n_{cb} \mid \nu_{cb}(\boldsymbol{\eta}, \boldsymbol{\chi})\right)}_{\text{multiple channels}} \underbrace{\prod_{\chi \in \boldsymbol{\chi}} c_\chi \left(a_\chi \mid \chi\right)}_{\text{constraint terms}}$$

Expected number of events per channel per bin are

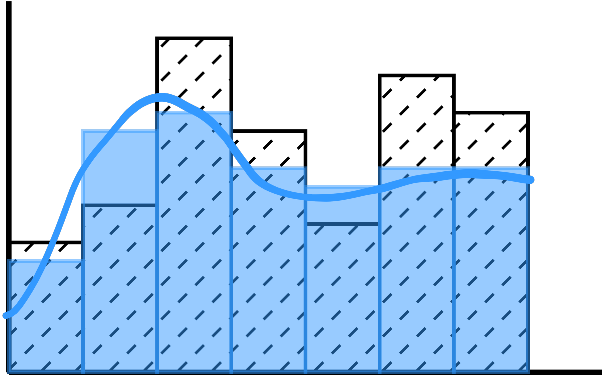
$$\nu_{cb}(\boldsymbol{\eta}, \boldsymbol{\chi}) = \sum_{s \in \text{samples}} \underbrace{\prod_{\kappa \in \boldsymbol{\kappa}} \kappa_{scb}(\boldsymbol{\eta}, \boldsymbol{\chi})}_{\text{multiplicative modifiers}} \left(\nu^0_{scb}(\boldsymbol{\eta}, \boldsymbol{\chi}) + \underbrace{\sum_{\Delta \in \boldsymbol{\Delta}} \Delta_{scb}(\boldsymbol{\eta}, \boldsymbol{\chi})}_{\text{additive modifiers}}\right).$$

Use redist for bin weights $\kappa(\boldsymbol{\eta}, \boldsymbol{\chi}) = n_B \,/\, n_A$.

Modifiers



Custom modifiers



Modifiers and constraints

Description	Modification	Constraint Term c_χ	Input
Uncorrelated Shape	$\kappa_{scb}(\gamma_b) = \gamma_b$	$\prod_b \mathrm{Pois}\left(r_b = \sigma_b^{-2} \middle| \rho_b = \sigma_b^{-2}\gamma_b\right)$	σ_b
Correlated Shape	$\Delta_{scb}(\alpha) = f_p\left(\alpha \middle| \Delta_{scb,\alpha=-1}, \Delta_{scb,\alpha=1}\right)$	$\mathrm{Gaus}\left(a = 0 \middle| \alpha, \sigma = 1\right)$	$\Delta_{scb,\alpha=\pm1}$
Normalisation Unc.	$\kappa_{scb}(\alpha) = g_p\left(\alpha \middle| \kappa_{scb,\alpha=-1}, \kappa_{scb,\alpha=1}\right)$	$\mathrm{Gaus}\left(a = 0 \middle| \alpha, \sigma = 1\right)$	$\kappa_{scb,\alpha=\pm1}$
MC Stat. Uncertainty	$\kappa_{scb}(\gamma_b) = \gamma_b$	$\prod_b \mathrm{Gaus}\left(a_{\gamma_b} = 1 \middle| \gamma_b, \delta_b\right)$	$\delta_b^2 = \sum_s \delta_{sb}^2$
Luminosity	$\kappa_{scb}(\lambda) = \lambda$	$\mathrm{Gaus}\left(l = \lambda_0 \middle| \lambda, \sigma_\lambda\right)$	$\lambda_0, \sigma_\lambda$
Normalisation	$\kappa_{scb}(\mu_b) = \mu_b$		
Data-driven Shape	$\kappa_{scb}(\gamma_b) = \gamma_b$		