

A model-independent likelihood function for the Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ analysis

Lorenz Gärtner^{1,2} on behalf of Belle II,

in collaboration with

Danny van Dyk³, Lukas Heinrich^{2,4}, MÉRIL Reboud³

¹LMU Munich, ²Excellence Cluster ORIGINS,

³IPPP Durham, ⁴TU Munich

26.10.2023



Reusability in HEP



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

Kyle Cranmer^{1*}, Sabine Kraml²¹, Harrison B. Prosper³⁶ (editors),
Philip Bechtle⁴, Florian U. Bernlochner⁵, Itay M. Bloch⁶, Enzo Canonero⁵, Marcin Chrzasczcz⁷, Andrea Cocco⁸, Jan Conrad⁹, Glen Cowan¹⁰, Matthew Feickert¹¹,
Nahuel Ferreiro Iachellini^{12,13}, Andrew Fowlie¹⁴, Lukas Heinrich¹⁵, Alexander Held^{15,19},
Thomas Kuhr^{13,16}, Anders Kvellestad¹⁷, Maeve Madigan¹⁸, Farvah Mahmoudi^{15,19},
Knut Dundas Mora²⁰, Mark S. Neubauer¹¹, Maurizio Pierini¹⁵, Juan Rojo⁸, Sezen Sekmen²², Luca Silvestrini²³, Veronica Sanz^{24,25}, Giordon Stark²⁶, Riccardo Torre⁸,
Robert Thorne²⁷, Wolfgang Waltenberger²⁸, Nicholas Wardle²⁹, Jonas Wittbrodt³⁰

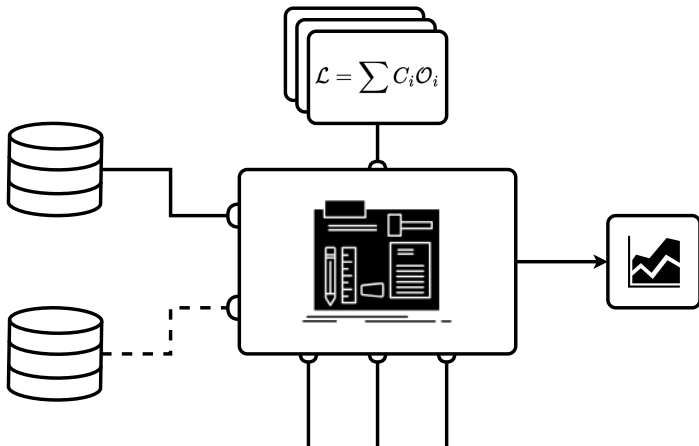
Forum [8], with the current status and updated recommendations presented in Ref. [6]. This paper takes these decade-long efforts to what we argue is the logical conclusion: if we wish to maximize the scientific impact of particle physics experiments, decades into the future, we should make the publication 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]]

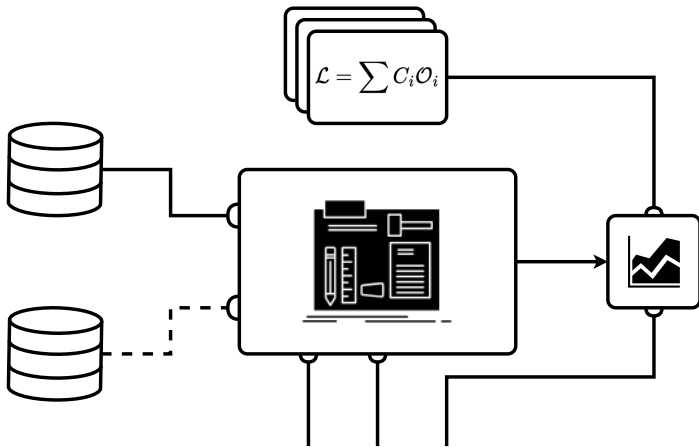
Shortcomings of model-dependencies in analyses

- **Limited interpretability** in terms of any model with different kinematic predictions.

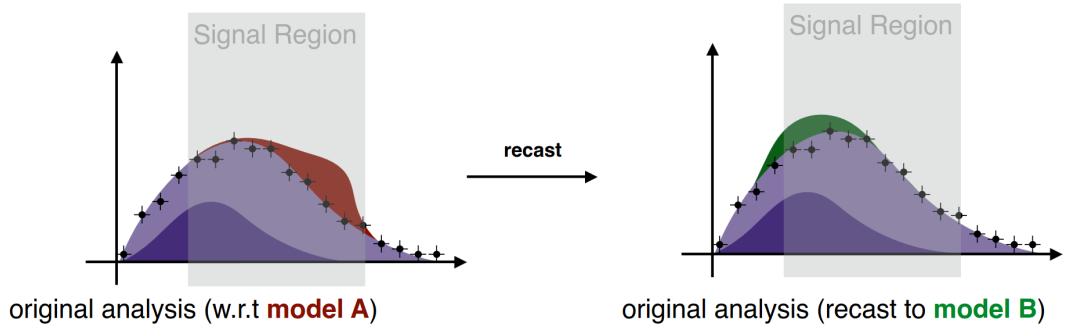
A general analysis



A better analysis



Idea behind reinterpretation



[source]

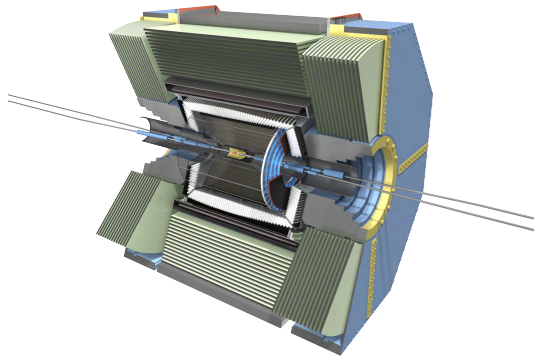


Analysis

Where is the model dependence?

SuperKEKB & Belle II introduction

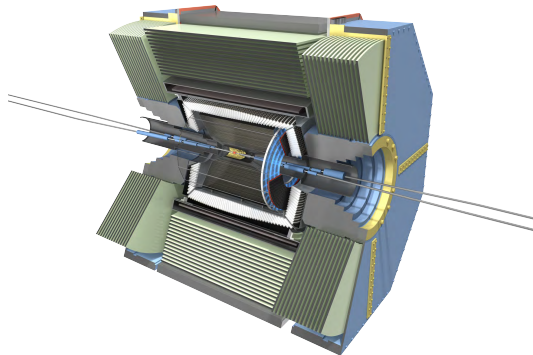
- **Luminosity** vs. energy frontier



[arXiv:1808.10567](https://arxiv.org/abs/1808.10567) [hep-ex]

SuperKEKB & Belle II introduction

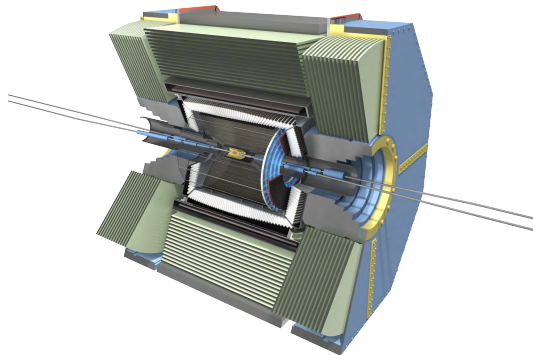
- **Luminosity** vs. energy frontier
- **SuperKEKB**
 - asymmetric e^- (7 GeV) – e^+ (4 GeV)
- **Belle II**
 - Hermetic, longitudinally asymmetric detector



[arXiv:1808.10567](https://arxiv.org/abs/1808.10567) [hep-ex]

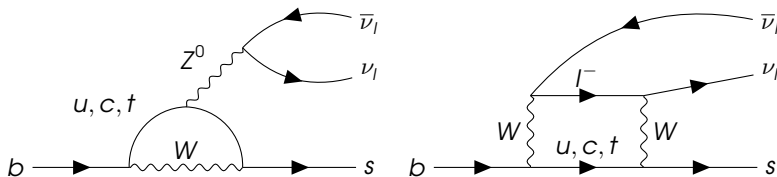
SuperKEKB & Belle II introduction

- **Luminosity** vs. energy frontier
 - **SuperKEKB**
 - asymmetric e^- (7 GeV) – e^+ (4 GeV)
 - **Belle II**
 - Hermetic, longitudinally asymmetric detector
- Missing mass analyses possible



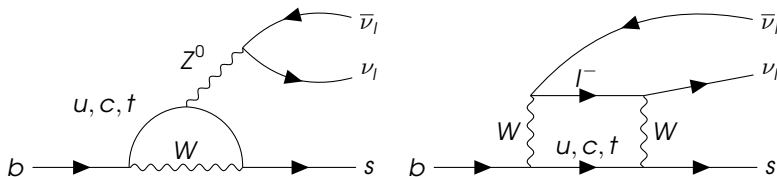
[arXiv:1808.10567](https://arxiv.org/abs/1808.10567) [hep-ex]

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



- Suppression of FCNCs in the SM.
- Tree level BSM effects could substantially affect observables.

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



- Suppression of FCNCs in the SM.
- Tree level BSM effects could substantially affect observables.
- **Benefits of reinterpretation**
 - **Sensitivity** to any current or future (B)SM prediction.
 - **Exclusion** limits in BSM parameter space inferable.
 - **Combinations** with other measurements possible.



HEPData

[[hepdata.130199](#)]

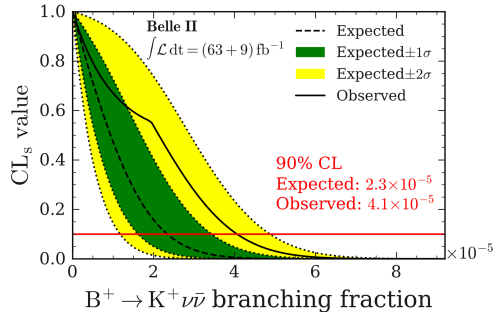
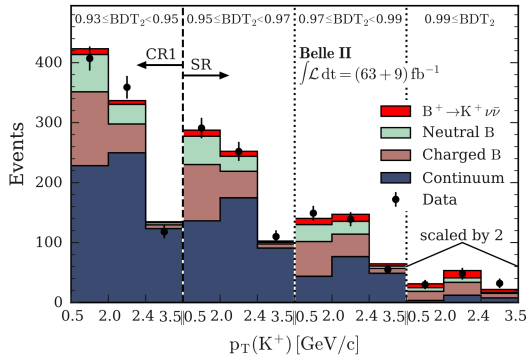
[[Phys.Rev.Lett.127.181802](#)]

The Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ analysis

1. Machine learning methods (BDTs): separate signal from background.
2. Signal MC weighted according to SM kinematic prediction
→ **model dependence**
3. Max. likelihood fit in bins of $p_T(K^+) \times BDT_2$.

The Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ analysis

1. Machine learning methods (BDTs): separate signal from background.
2. Signal MC weighted according to SM kinematic prediction
→ **model dependence**
3. Max. likelihood fit in bins of $p_T(K^+) \times BDT_2$.



$$B(B^+ \rightarrow K^+ \nu \bar{\nu}) < 4.1 \times 10^{-5} @ 90\% CL$$



Reweighting approach

How do we obtain new signal templates?

Reweighting recipe



Recipe

1. Find kinematic dependence of measured observable: $\Gamma(q^2), q^2 = (p_\nu + p_{\bar{\nu}})^2$

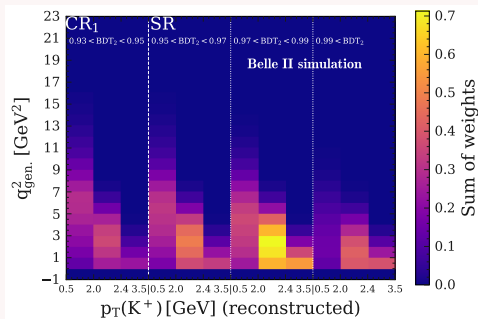
Reweighting recipe



Recipe

1. Find kinematic dependence of measured observable: $\Gamma(q^2)$, $q^2 = (p_\nu + p_{\bar{\nu}})^2$
2. Get distributions of kinematic d.o.f (q^2)

$$N_{klm} = \underbrace{p_T \times BDT_2}_{\text{analysis binning}} \times \underbrace{q_{\text{gen.}}^2}_{\text{kinematic d.o.f}}$$



Reweighting approach



Recipe

1. Find kinematic dependence of theoretical prediction: $\Gamma(q^2)$, $q^2 = (p_\nu + p_{\bar{\nu}})^2$
2. Get distributions of kinematic d.o.f (q^2)

$$N_{klm} = \underbrace{p_T \times BDT_2}_{\text{analysis binning}} \times \underbrace{q_{\text{gen.}}^2}_{\text{kinematic d.o.f}}$$

3. Apply weights in bins of kinematic d.o.f. (phase space (PHSP): $\mathcal{M} = 1$)

$$N_{kl} = \sum_{m \in q^2} N_{klm}^{\text{PHSP}} w_m = \sum_{m \in q^2} N_{klm}^{\text{PHSP}} \int_{\text{bin } m} dq^2 \frac{d\Gamma^{(B)SM}}{dq^2} \left(\frac{d\Gamma^{\text{PHSP}}}{dq^2} \right)^{-1}$$

Reweighting approach



Recipe

1. Find kinematic dependence of theoretical prediction: $\Gamma(q^2)$, $q^2 = (p_\nu + p_{\bar{\nu}})^2$
2. Get distributions of kinematic d.o.f (q^2)

$$N_{klm} = \underbrace{p_T \times BDT_2}_{\text{analysis binning}} \times \underbrace{q_{\text{gen.}}^2}_{\text{kinematic d.o.f}}$$

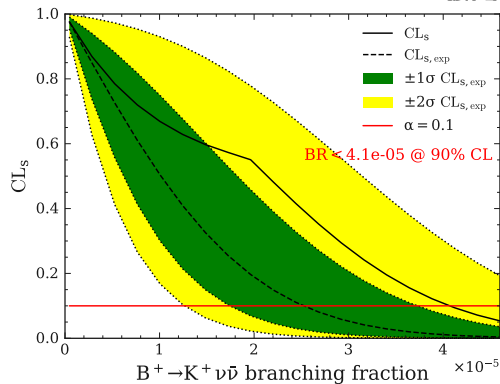
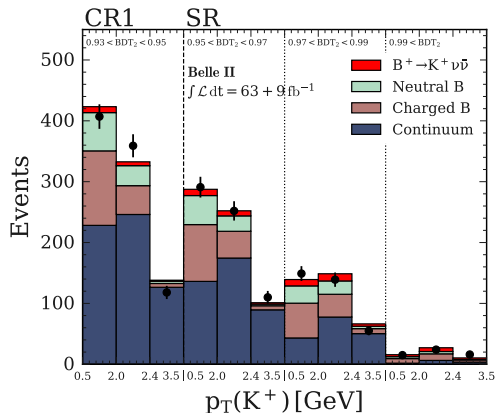
3. Apply weights in bins of kinematic d.o.f. (phase space (PHSP): $\mathcal{M} = 1$)

$$N_{kl} = \sum_{m \in q^2} N_{klm}^{\text{PHSP}} w_m = \sum_{m \in q^2} N_{klm}^{\text{PHSP}} \int_{\text{bin } m} dq^2 \frac{d\Gamma^{(B)SM}}{dq^2} \left(\frac{d\Gamma^{\text{PHSP}}}{dq^2} \right)^{-1}$$

Benefits

- + Fast
- + Versatile
- + Easily publishable

Validation: Reproducing the upper limit



$$B(B^+ \rightarrow K^+ \nu \bar{\nu}) < 4.1 \times 10^{-5} @ 90\% \text{ CL}$$



pyhf.readthedocs.io



Theory

How can we parametrize our model dependence?



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

Contribution operators

The effective Lagrangian is

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

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

$$\mathcal{O}_{VL} = (\bar{\nu}_L \gamma_\mu \nu_L) (\bar{s}_L \gamma^\mu b_L) \quad \mathcal{O}_{VR} = (\bar{\nu}_L \gamma_\mu \nu_L) (\bar{s}_R \gamma^\mu b_R)$$

$$\mathcal{O}_{SL} = (\bar{\nu}_L^c \nu_L) (\bar{s}_R b_L) \quad \mathcal{O}_{SR} = (\bar{\nu}_L^c \nu_L) (\bar{s}_L b_R)$$

$$\mathcal{O}_{TL} = (\bar{\nu}_L^c \sigma_{\mu\nu} \nu_L) (\bar{s}_R \sigma^{\mu\nu} b_L)$$

[arXiv:2111.04327 [hep-ph]]

(B)SM theory predictions

Capture BSM physics above electroweak symmetry breaking scale with

$$C_{SL} + C_{SR}$$

$$C_{VL} + C_{VR} = C_{VL}^{SM} + C_{VL}^{NP} + C_{VR}$$

$$C_{TL}$$



eos.github.io

(B)SM theory predictions



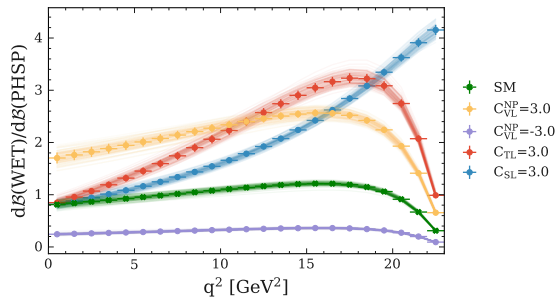
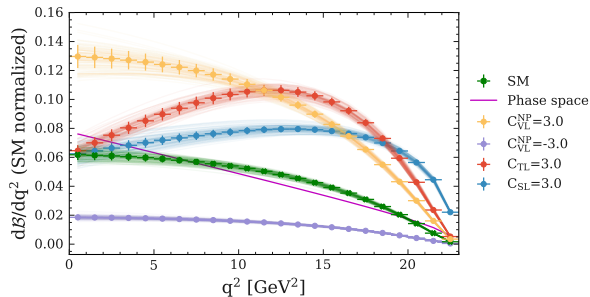
eos.github.io

Capture BSM physics above electroweak symmetry breaking scale with

$$C_{SL} + C_{SR}$$

$$C_{VL} + C_{VR} = C_{VL}^{SM} + C_{VL}^{NP} + C_{VR}$$

$$C_{TL}$$





Fruits of reinterpretation

What do we get from all this?

Wilson coefficient exclusion limits



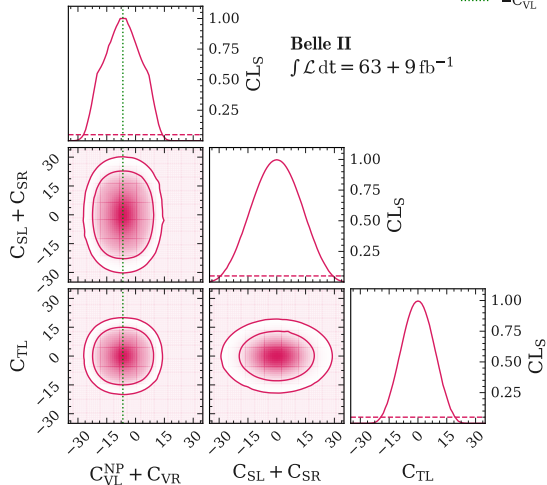
•  +  → fast & reliable

• Exclusion limits @95%CL

$$|C_{VL}^{SM} + C_{VL}^{NP} + C_{VR}| < 20.6$$

$$|C_{SL} + C_{SR}| < 29.3$$

$$|C_{TL}| < 19.4$$





Analysis update: $B^+ \rightarrow K^+ \nu \bar{\nu}$ @ 362 fb^{-1}

Inclusive $\mathcal{B} = \left(2.8^{+0.5}_{-0.5} \right) \times 10^{-5}$

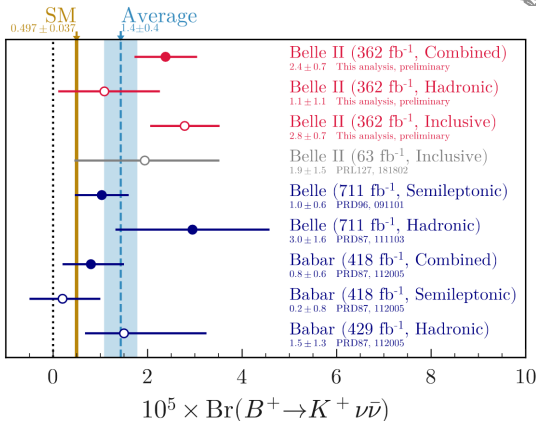
Hadronic $\mathcal{B} = \left(1.1^{+0.9}_{-0.8} \right) \times 10^{-5}$

Combined $\mathcal{B} = \left(2.4^{+0.5}_{-0.5} \right) \times 10^{-5}$

Significance of the **combined** result:

- 3.6σ wrt. null hypothesis
- 2.8σ wrt. SM

First evidence of $B^+ \rightarrow K^+ \nu \bar{\nu}$



Model-independent likelihood method will be applied and published once paper is accepted.

Presented at EPS 2023

Summary



- **Challenge**
 - Neutrino-induced experimental complexities
 - **model-dependent results**
- **Solution**
 - **Model-independent likelihood function**
Maximum likelihood fits for any given (B)SM signal prediction.
- **Tool integration (more in backup)**
 - Extend **pyhf** and interface it with **EOS** for run-time template updating.
 - Method fully applicable to other decay channels and results.
- **Scientific benefits**
 - **Exclusions in BSM parameter space.**
 - **Combinations** with other channels and/or experiments.
 - ...

Publishing such likelihoods is crucial for a full exploitation of experimental results.

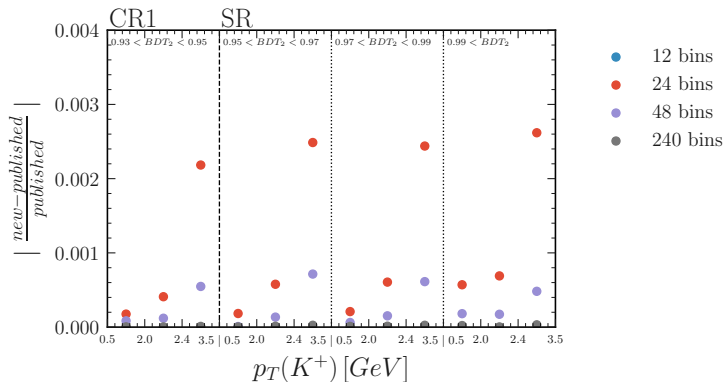


lorenz.gaertner@physik.uni-muenchen.de

Binning choice



We compared the relative accuracy of the binned weighting (new) with the event-by-event weighting (published).



Effective field theory



- **High energy collisions**

Enough energy to radiate off an on-shell (massive) W boson.

- **Lower energy quark decays**

W boson is always off-shell.

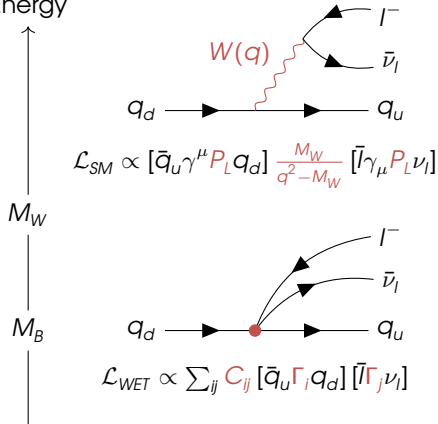
- **Weak effective theory**

W is integrated out – effects are encoded in new couplings

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

→ Model independent **parametrization**,
constrained only by Wilson coefficients C_i .

Energy





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

Decay width

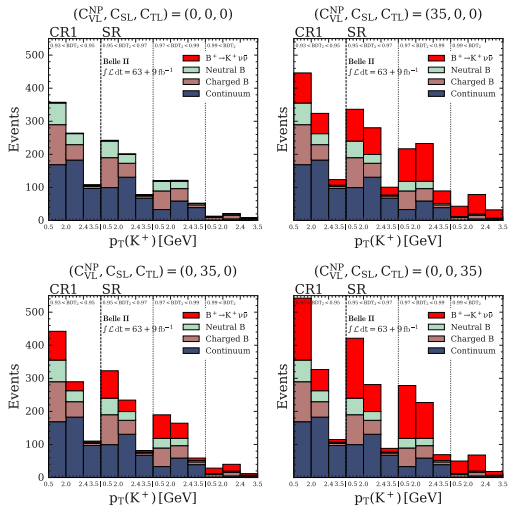
Decay width dependence on the Wilson coefficients is given by

$$\frac{d\Gamma(B \rightarrow K \nu \bar{\nu})}{dq^2} = \frac{\sqrt{\lambda_{BK}} q^2}{(4\pi)^3 m_B^3} \left[\frac{\lambda_{BK}}{24 q^2} |f_+(q^2)|^2 |C_{VL} + C_{VR}|^2 \right. \\ \left. + \frac{(m_B^2 - m_K^2)^2}{8 (m_b - m_s)^2} |f_0(q^2)|^2 |C_{SL} + C_{SR}|^2 \right. \\ \left. + \frac{2\lambda_{BK}}{3 (m_B + m_K)^2} |f_T(q^2)|^2 |C_{TL}|^2 \right]$$

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

[arXiv:2111.04327 [hep-ph]]

Expected yields



Hadronic parameters



Form factors are parameterized using the BCL parametrization

$$f_0(q^2) = \frac{\mathcal{L}}{1 - \frac{q^2}{M_{B_{s0}}^2}} \sum_{n=0}^{N-1} a_n^0 z^n$$

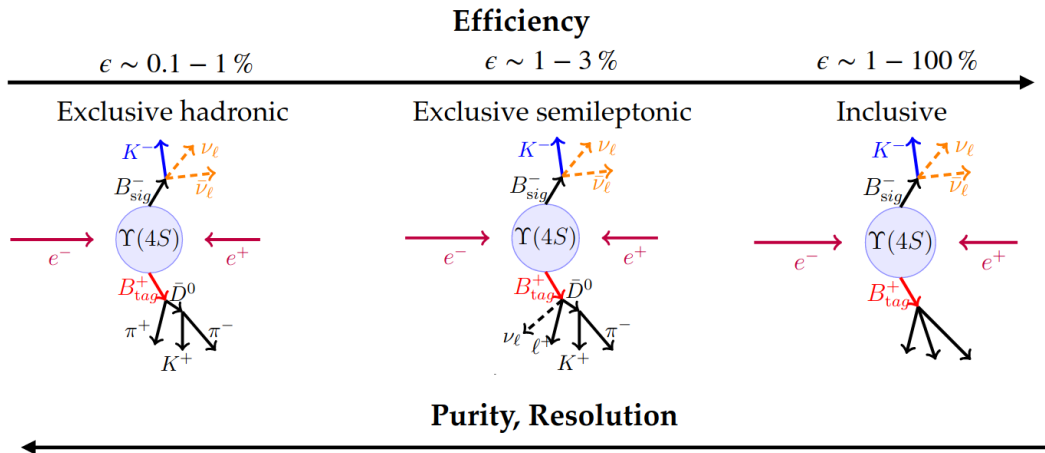
$$f_+(q^2) = \frac{\mathcal{L}}{1 - \frac{q^2}{M_{B_s^*}^2}} \sum_{n=0}^{N-1} a_n^+ \left(z^n - \frac{n}{N} (-1)^{n-N} z^N \right)$$

$$f_T(q^2) = \frac{\mathcal{L}}{1 - \frac{q^2}{M_{B_s^*}^2}} \sum_{n=0}^{N-1} a_n^T \left(z^n - \frac{n}{N} (-1)^{n-N} z^N \right),$$

The correlation matrix between the hadronic parameters.

	a_0^+	a_1^+	a_2^+	a_1^0	a_2^0	a_0^T	a_1^T
a_0^+	1.00	0.67	0.33	0.94	0.83	0.43	0.34
a_1^+	0.67	1.00	0.86	0.73	0.70	0.22	0.39
a_2^+	0.33	0.86	1.00	0.39	0.41	0.04	0.26
a_1^0	0.94	0.73	0.39	1.00	0.96	0.40	0.37
a_2^0	0.83	0.70	0.41	0.96	1.00	0.34	0.33
a_0^T	0.43	0.22	0.04	0.40	0.34	1.00	0.89
a_1^T	0.34	0.39	0.26	0.37	0.33	0.89	1.00
a_2^T	0.21	0.42	0.44	0.23	0.23	0.71	0.91

Reconstruction techniques



Different reconstruction techniques lead to nearly orthogonal data samples



Implementation

How do we realize this?

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

pyhf = **pythonic HistFactory**

Interval estimation based on

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

1 Introduction

2

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 Cowan¹, Kyle Cranmer², Eilam Gross¹, Ofer Vitells^{3,4}

¹Physics Department, Royal Holloway, University of London, Egham TW20 0EX, UK

²Physics Department, New York University, New York, NY 10003, USA

³Weizmann Institute of Science, Rehovot 76100, Israel

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

© The Author(s) 2011. This article is published with open access at Springerlink.com

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

HistFactory model



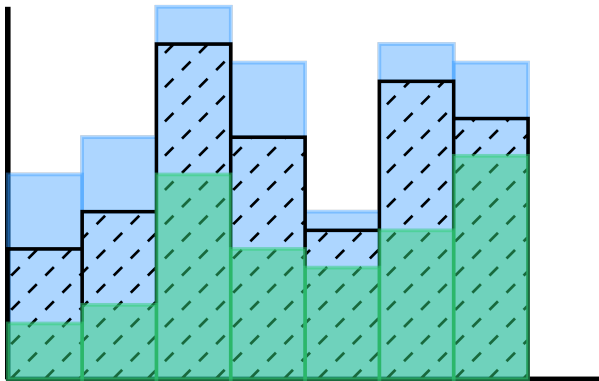
Likelihood function for observed event counts $\mathbf{n}$ is

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

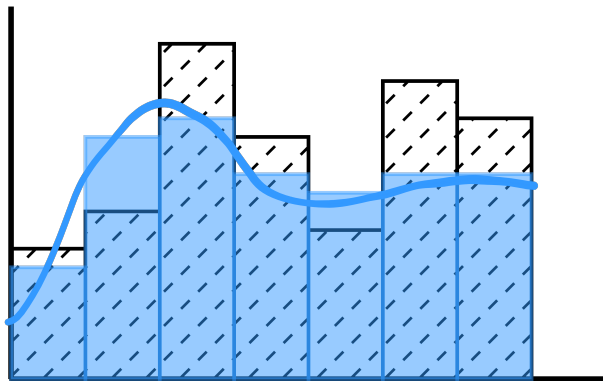
Expected number of events per channel per bin are

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

Modifiers



Custom modifiers



Modifiers and constraints



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