Reinterpretation with model-agnostic likelihoods

Constraining new physics with $B^+ \to K^+ \nu\bar{\nu}$ @ Belle II

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] IJCLab, [4]KIT, [5]IPPP Durham

October 16, 2024

What we dream of...

What we dream of...

... the PDF of observables x given theory parameters θ

$$p(x|\theta) = \int dz_d \int dz_p \underbrace{p(x|z_d)}_{\text{observables}} \underbrace{p(z_d|z_p)}_{\text{detector response}} \underbrace{p(z_p|\theta)}_{\textit{prediction}}$$

What we dream of...

... the PDF of observables x given theory parameters θ

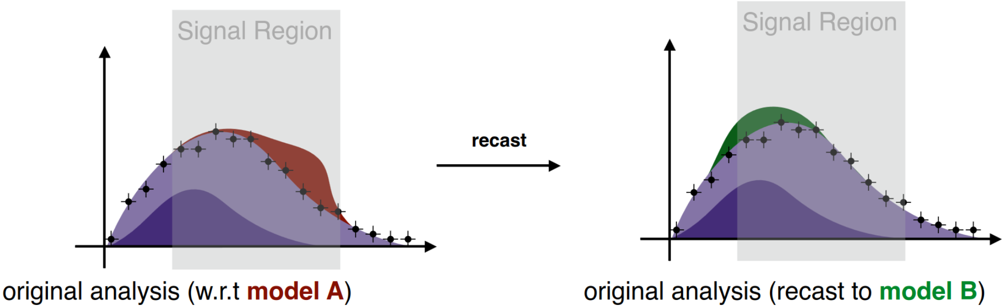
$$p(x|\theta) = \int dz_d \int dz_p \underbrace{p(x|z_d)}_{\text{observables}} \underbrace{p(z_d|z_p)}_{\text{detector response}} \underbrace{p(z_p|\theta)}_{\textit{prediction}}$$

Intractability forces us to approximate

$$\text{MC} \sim p(x|\theta)$$

➜ **MC data is model dependent**

What result would we get if we replace the signal model?



[source]

Approaches to reinterpretation

Recasting / simulation based reinterpretation

- Produce and analyse new MC samples
- Requires full analysis strategy
- Relatively accurate (analysis still optimized for original model)

🔥 Resource-heavy

Approaches to reinterpretation

Recasting / simulation based reinterpretation

- Produce and analyse new MC samples
- Requires full analysis strategy
- Relatively accurate (analysis still optimized for original model)

🔥 Resource-heavy

Simplified model reinterpretation

- Assumption: acceptances / efficiencies unaffected by kinematic shape differences

⚠️ Potentially biased results

Approaches to reinterpretation

Recasting / simulation based reinterpretation

- Produce and analyse new MC samples
- Requires full analysis strategy
- Relatively accurate (analysis still optimized for original model)

🔥 Resource-heavy

Simplified model reinterpretation

- Assumption: acceptances / efficiencies unaffected by kinematic shape differences

⚠️ Potentially biased results

Is accurate reinterpretation without new MC samples possible?
♻️ **Yes, we can reweight samples or even histograms directly!**

redist

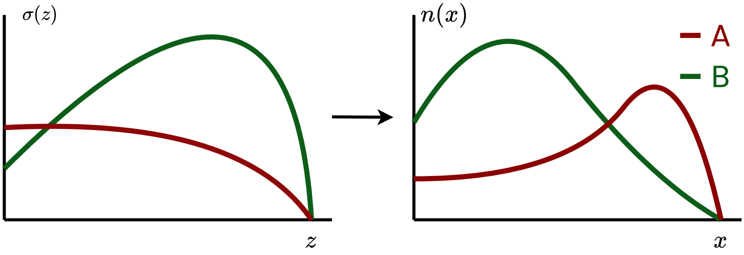
A novel reinterpretation method

[EPJC]

github.com/lorenzennio/redist

Templates from kinematic predictions

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



z – kinematic d.o.f.

x – reconstruction / fitting variable(s)

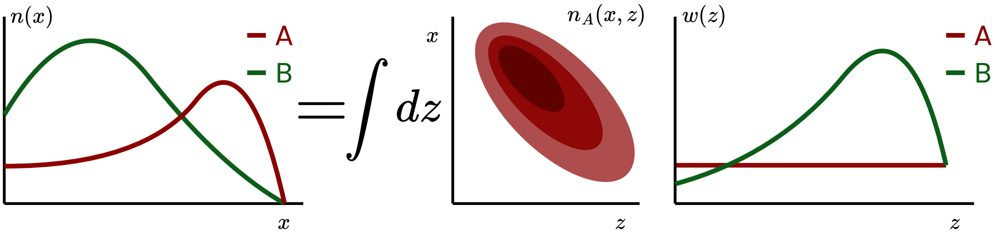
L – luminosity

$\varepsilon(x|z)$ – conditional efficiency

$n(x,z)$ – joint number density

Reweight to new model

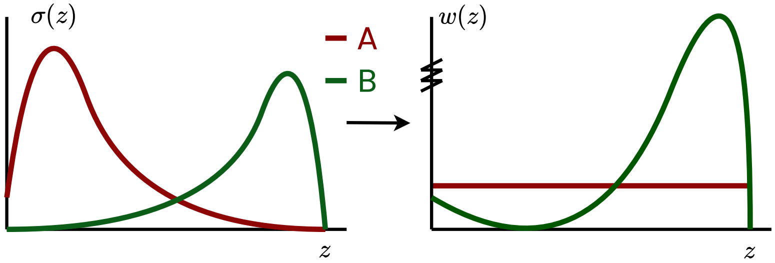
$$n_B(x) = \int dz\, L\, \varepsilon(x|z)\, \sigma_B(z) = \int dz\, L\, \varepsilon(x|z)\, \sigma_A(z) \frac{\sigma_B(z)}{\sigma_A(z)} = \int dz\, \underbrace{n_A(x,z)}_{\text{main object}}\, w(z).$$



$p(x|\theta)$ & $n(x,z)$ = model-agnostic likelihood

Method limitations

Substantial model changes ➜ large weights



Minimal requirement:

$$\text{supp}(\sigma_B) \in \text{supp}(\sigma_A)$$

Always possible to compare only in $\text{supp}(\sigma_A)$

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

Evidence for $B^+ \to K^+\nu\bar{\nu}$ – What next?

Belle II has found "**Evidence for $B^+ \to K^+\nu\bar{\nu}$ Decays**"
arXiv:2311.14647 [hep-ex]

What does this really mean?

- This statement is based on the model

$$p(x|\mathrm{SM})$$

- However, we cannot say much about

$$p(x|\mathrm{NP})$$

Benefit of combination

Total of 3 analyses

- **ITA**

$$\underbrace{p(x_{ITA}|\theta_{ITA})}_{\mathrm{BDT}_2\times q^2_{rec}}$$

- **HTA**

$$\underbrace{p(x_{HTA}|\theta_{HTA})}_{\mathrm{BDTh}}$$

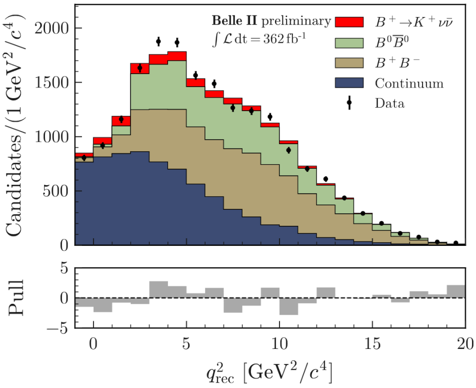
- **ITA+HTA**

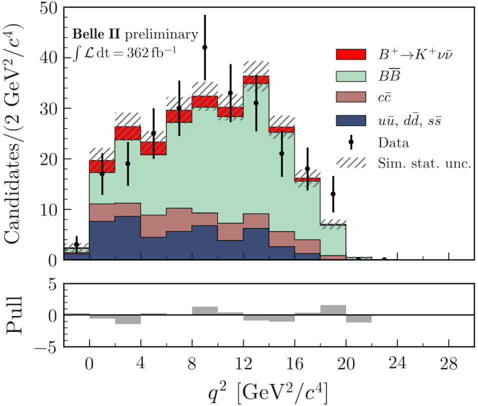
$$p(x|\theta) = \underbrace{p(x_{ITA}|\theta'_{ITA})}_{\mathrm{BDT}_2\times q^2_{rec}} \cdot \underbrace{p(x_{HTA}|\theta'_{HTA})}_{\mathrm{BDTh}}$$

➜ 10% increase in precision over ITA result

arXiv:2311.14647 [hep-ex]

There are always open questions...





Excess around 3 $\mathrm{GeV}^2 < q^2_{rec} < 7\ \mathrm{GeV}^2$ for ITA best-fit projection.

arXiv:2311.14647 [hep-ex]

What if the $B^+ \to K^+ \nu\bar{\nu}$ signal includes $B^+ \to K^+ X$?

X – light ($m_X \leq m_B - m_K$) resonant boson, decaying invisibly or escaping undetected

See arXiv:2311.14629 [hep-ph], arXiv:2312.12507 [hep-ph] for first estimates.

Ingredients for reinterpretation

1. Likelihood $p(x|\theta)$
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 n is

$$p(x|\theta) = p(n, a|\eta, \chi) = \underbrace{\mathrm{Pois}\left(n|\nu(\eta,\chi)\right)}_{\text{data likelihood}} \quad \underbrace{c\left(a|\chi\right)}_{\text{constraint likelihood}}$$

Expected number of events are

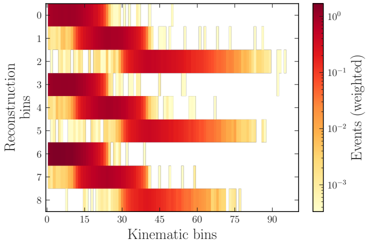
$$\nu(\eta,\chi) = \kappa(\eta,\chi)\left(\nu^0(\eta,\chi) + \Delta(\eta,\chi)\right).$$

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

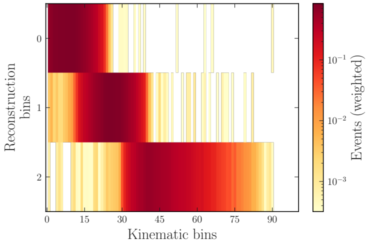
2. Joint number density

The main object for reinterpretation, $n(x, z)$.
➜ **This should be published by collaborations.**

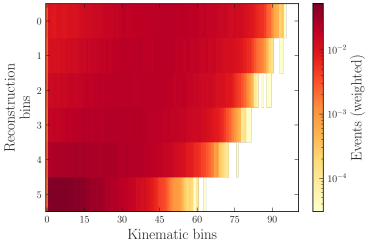
ITA 1



ITA 2



HTA



Kinematic binning: $q^2_{gen.} = [-1, (0, 22.885, 100)]$ GeV^2

3. Decay kinematics

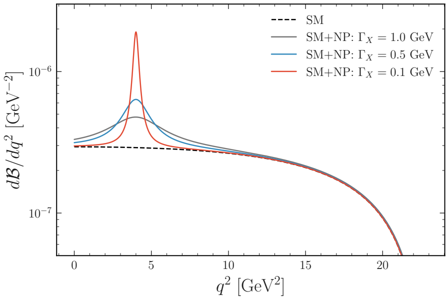
- Model as sharp resonance peak,

$$\frac{d\mathcal{B}}{dq^2} = \frac{d\mathcal{B}_{SM}}{dq^2} + \mu_X \, p_{RBW}(q^2; m_X, \Gamma_X) \cdot 10^{-6},$$

with the PDF

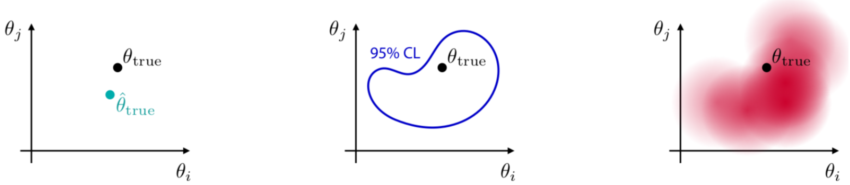
$$p_{RBW}(q^2; m_X, \Gamma_X) = \frac{k}{(q^2 - m_X^2)^2 - m_X^2 \Gamma_X^2}$$

- **Free parameters** are μ_X, m_X (and Γ_X)
- Bound by resolution $\Gamma_X > 0.1$ GeV



Model shown for $\mu_X = 1$, $m_X = 2$ GeV.

What do we want to learn?



Gilles Louppe @ PHYSTAT-SBI 2024

- **Frequentist**: Find $\hat{\theta}$ maximising $p(x|\theta)$. Build confidence interval.
- **Bayesian**: Compute posterior $p(\theta|x) \propto p(x|\theta)p(\theta)$. Obtain credible intervals.

Bayesian inference more insightful for multiple, correlated POIs.
Posterior model simply from pyhf likelihood:

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

$B^+ \to K^+X$ marginal posterior

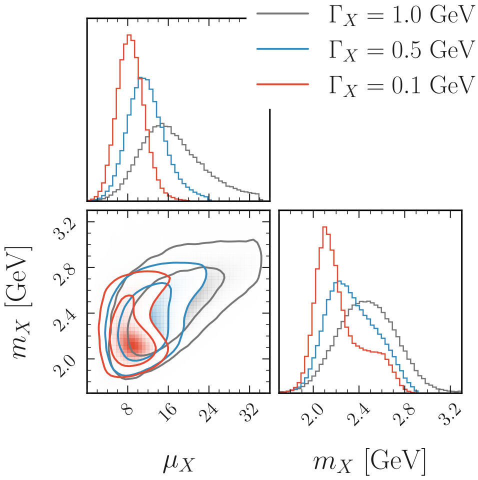
- 1d and 2d *marginalized* posterior

$$p(\eta|n,a) = \int d\chi \; p(\eta,\chi|n,a)$$

- Uniform priors for POIs
- **Credible intervals** (95%) from $\Gamma_X = 0.1$ GeV

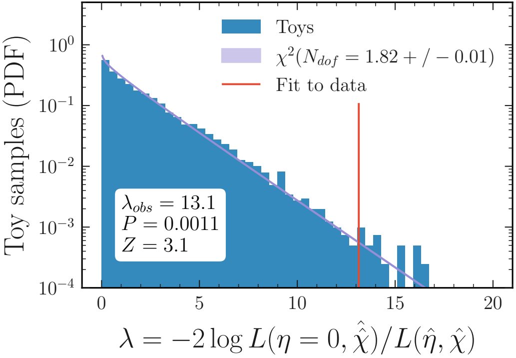
$$\mathcal{B}(B^+ \to K^+X) \in [0.34, 1.4] \times 10^{-5}$$
$$m_X \in [1.9, 2.7]\ \text{GeV}$$

- Consistent with results of `arXiv:2311.14629 [hep-ph]`, `arXiv:2312.12507 [hep-ph]`



Model comparison

Pointwise: P-value



- Toy study for $\Gamma_X = 0.1$ GeV model, 10k toys

➜ **Significance of $Z = 3.1$ over SM+background**

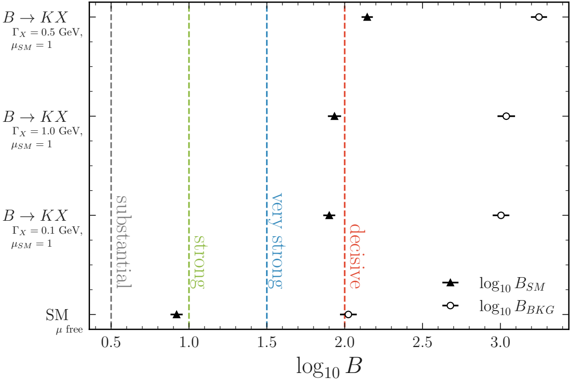
Averaged: Bayes factor

Compare the probabilities of the observed data being produced by a given model.

$$p(\theta|x, M) = \frac{p(x|\theta, M)\, p(\theta|M)}{p(x|M)}$$

$$p(x|M) = \int d^n\theta \; p(x|\theta, M)\, p(\theta|M)$$

$$B_M = \frac{p(x|M')}{p(x|M)}$$



➜ $B \to KX$ models dominate over SM.

⚠ *Average* likelihood in the *full* parameter space ➜ prior dependent

Summary

- ♻ **Reinterpretability is important**
  - Necessary for **bias-free inference** on BSM parameters.
  - Enables **combinations** with other channels and/or experiments.
  - ➜ Increases **impact** of results.
- ⚖ **Reinterpretation can be easy (redist)**
  - Only requires **likelihood** and $\boldsymbol{n(x, z)}$ (or MC samples with kinematic information).
  - **FAST**
- $\boldsymbol{B \to K \nu \bar{\nu}}$ **reinterpretation i.t.o. B → KX**
  - High significance model ($Z = 3.1$) with potential explanation for the observed excess in data.
  - Credible intervals (95%)

$$BR(B^+ \to K^+ X) \in [0.34, 1.4] \times 10^{-5} \qquad m_X \in [1.9, 2.7] \text{ GeV}$$

- pyhf model and $n(x, z)$ to be published by Belle II (under internal review)

lorenz.gaertner@physik.uni-muenchen.de
[B2Documents] [method paper]

$B \to KX$ posterior

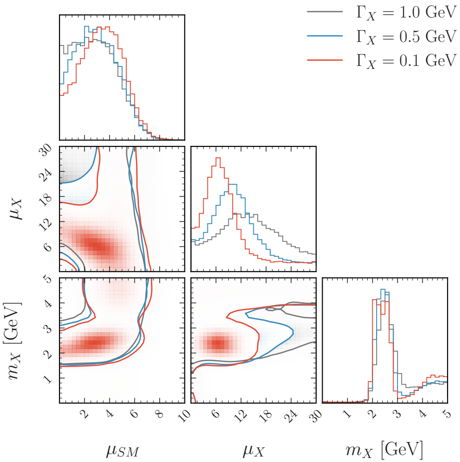
- Based on ITA+HTA likelihood
- Priors are uniform with support

$$\begin{aligned} 0 \leq \mu_{SM} &\leq 10 \\ 0 \leq \mu_X &\leq 20 \\ 0 \leq m_X &\leq 4.8 \text{ GeV} \end{aligned}$$

- Contours encompass 68% and 95% of samples
- 3 different Γ_X cases

➜ bound by resolution $\Gamma_X > 0.1$ GeV

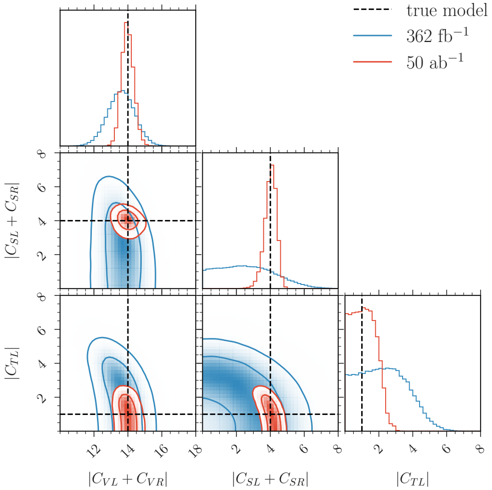
- consistent with results of arXiv:2311.14629 [hep-ph]



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

- Weak effective theory $\sigma(z|\psi)$, $\psi = \{C_i\}$ (models B)
- Analysis assumes SM (model A)
    - $C_{\text{VL}} \simeq 6.6\,, C_i = 0 \quad \forall \quad i \neq \text{VL}$
- "Data" contains new physics

arXiv:2402.08417 [hep-ph]

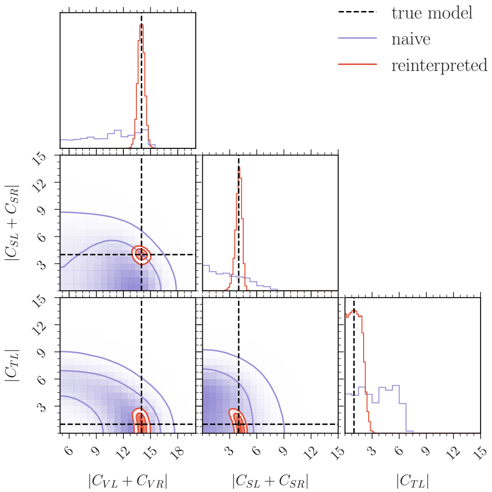


The necessity for reinterpretation

- naive reinterpretation
  = simple BR rescaling

➜ biased posterior

arXiv:2402.08417 [hep-ph]



HistFactory model

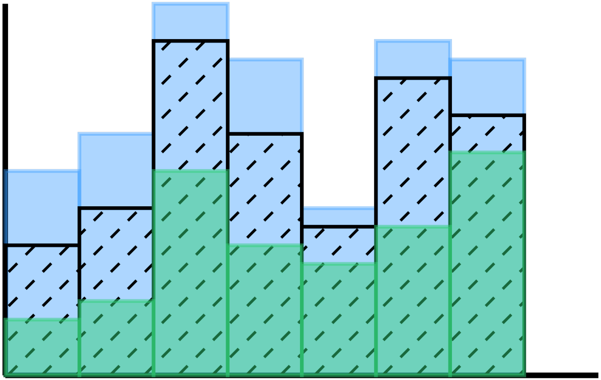
Likelihood function for observed event counts n is

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

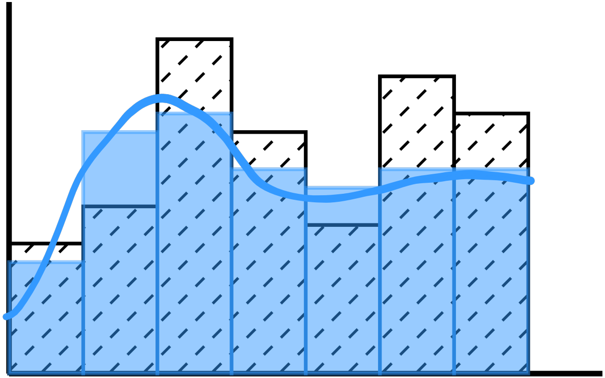
Expected number of events per channel per bin are

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

Modifiers



Custom modifiers



Modifiers and constraints

Description	Modification	Constraint Term c_χ	Input
Uncorrelated Shape	$\kappa_{scb}(\gamma_b) = \gamma_b$	$\prod_b \mathrm{Pois}\left(r_b = \sigma_b^{-2} \middle| \rho_b = \sigma_b^{-2}\gamma_b\right)$	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	$\kappa_{scb}(\gamma_b) = \gamma_b$		