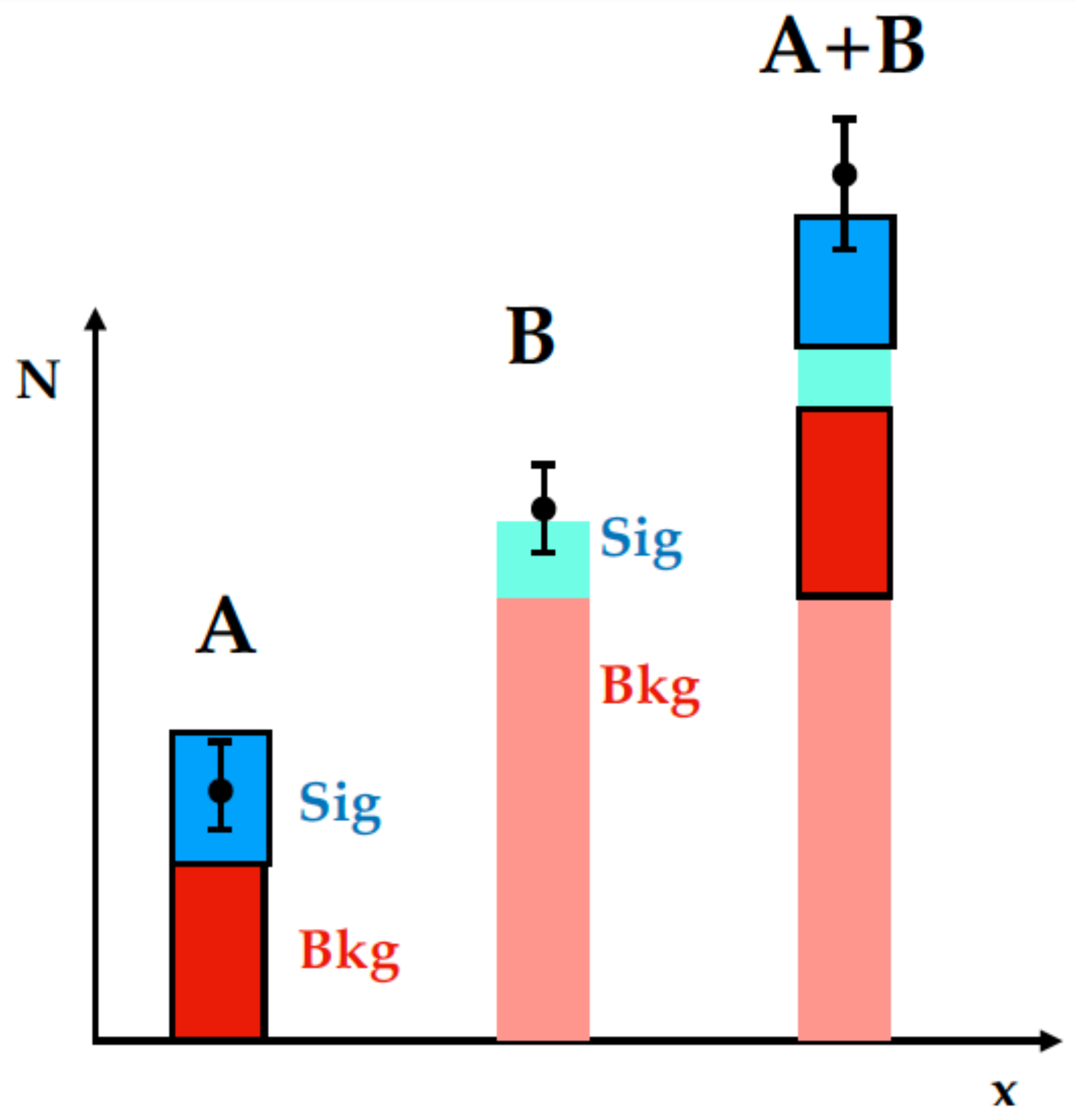
Introduction

I thought that one topic that can be shared among different experiments and phenomenologists is the statistical treatment and interpretation of the results.

Here I am sharing some random thoughts and mainly questions

- **Low stats vs high stats**:
  - Low stats (almost no background): transfer of knowledge neutrino physics -> others
- **Sharing information**: HEPdata
- Frequentist vs Bayesian
- Global fits and **combinations**
  - Global fit in the SM exist, no combination for dark sector searches as far as I know (maybe GAMBIT?)
- Counting vs **shape analysis** (unbinned analysis)
  - See next slide

Binned—> unbinned



The combination of A and B is better (more significant) than considering all the events in the same bin

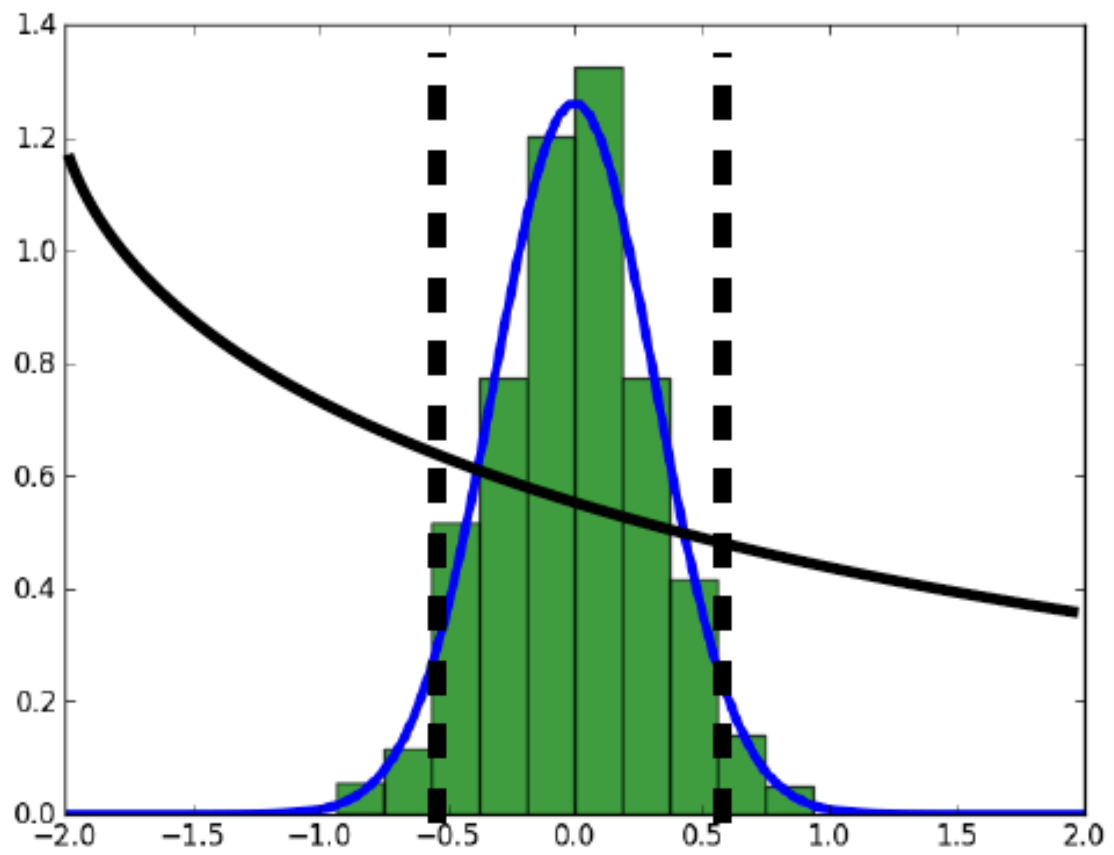
$$P_{s+b}(X \le X_{obs}) = \sum_{X(\{d'_i\}) \le X(\{d_i\})} \prod_{i=1}^{n} \frac{e^{-(s_i+b_i)}(s_i+b_i)^{d'_i}}{d'_i!}$$

Thomas Junk, Nuclear Instruments and Methods in Physics Research A 434 (1999) 435-443

- Combining categories with different S/B increases the sensitivity

- Shape = using many bins, each one with a different S/B. Each bin is a sort of "channel" or category

- Shape analysis can be
  - binned (many categories)
  - or unbinned



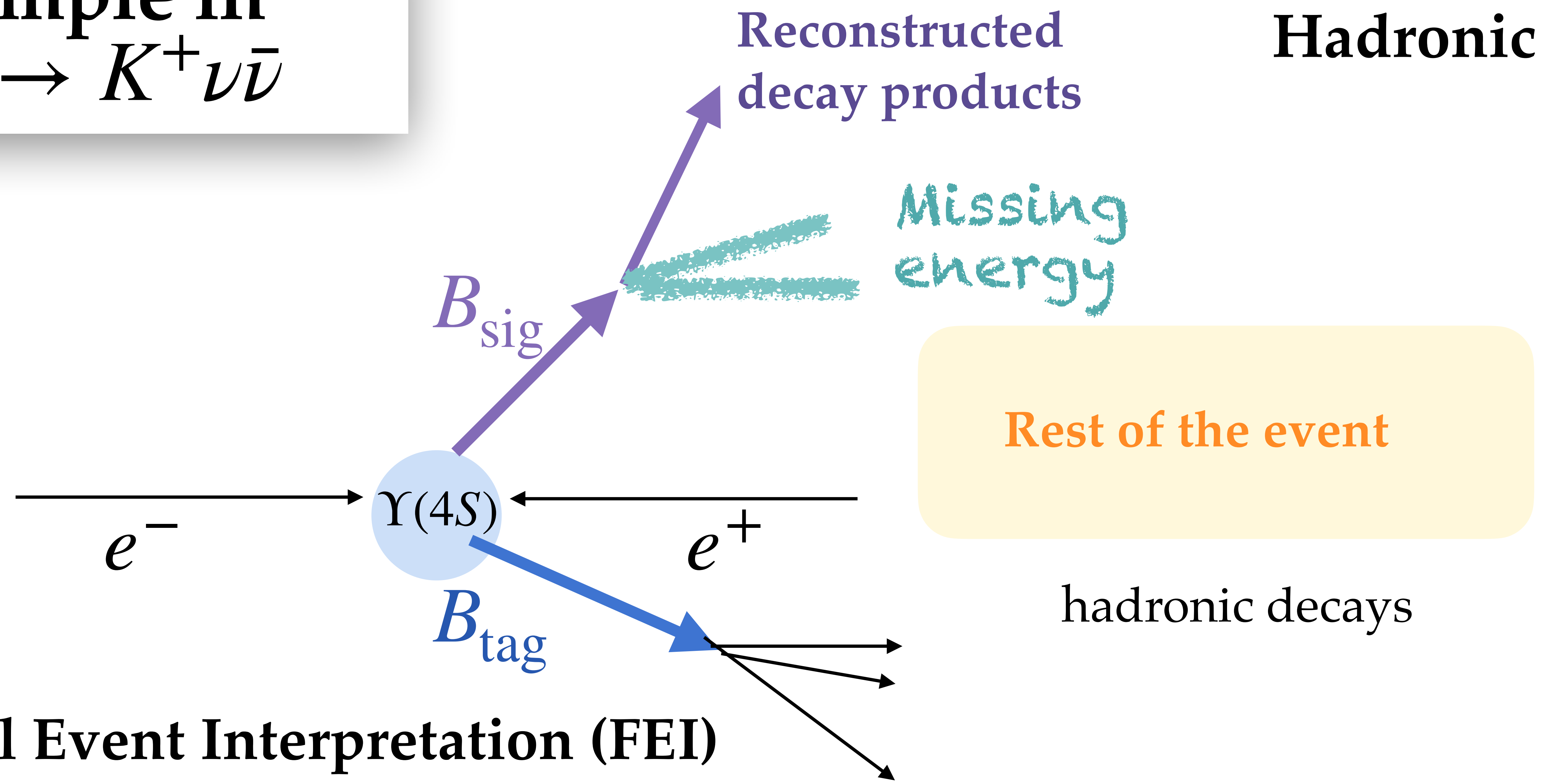
$$k^{-1}\prod_i\left(\mu S \boxed{f_s(x_i)} + B \boxed{f_b(x_i)}\right) \cdot e^{-(\mu S+B)}$$

Shapes

By cutting a window:
- Loose the tails of the signal
- Loose the possibility to combine channels with different S/B

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

Hadronic tag



Full Event Interpretation (FEI)

[Keck, T. et al. Comput Softw Big Sci 3, 6 (2019)]

MVA based B-tagging algorithm with hierarchical approach to reconstruct $\mathcal{O}(10^4)$ decay chains

$$q^2_{rec} = \frac{s}{4} + m_K^2 - \sqrt{s}E_K^* - 2\mathbf{p}_{tag} \cdot \mathbf{p}_K$$

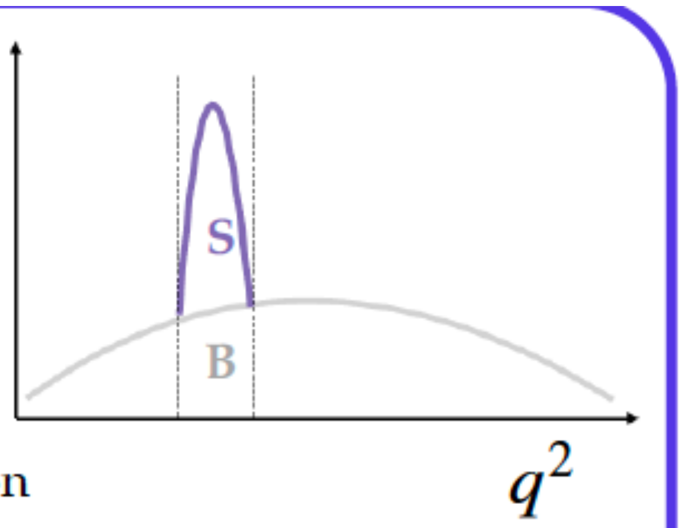
Resolution can be as low as 0.1 GeV2

2-body decay: $B \to KX_{inv}$

Bump hunting in q^2 distribution

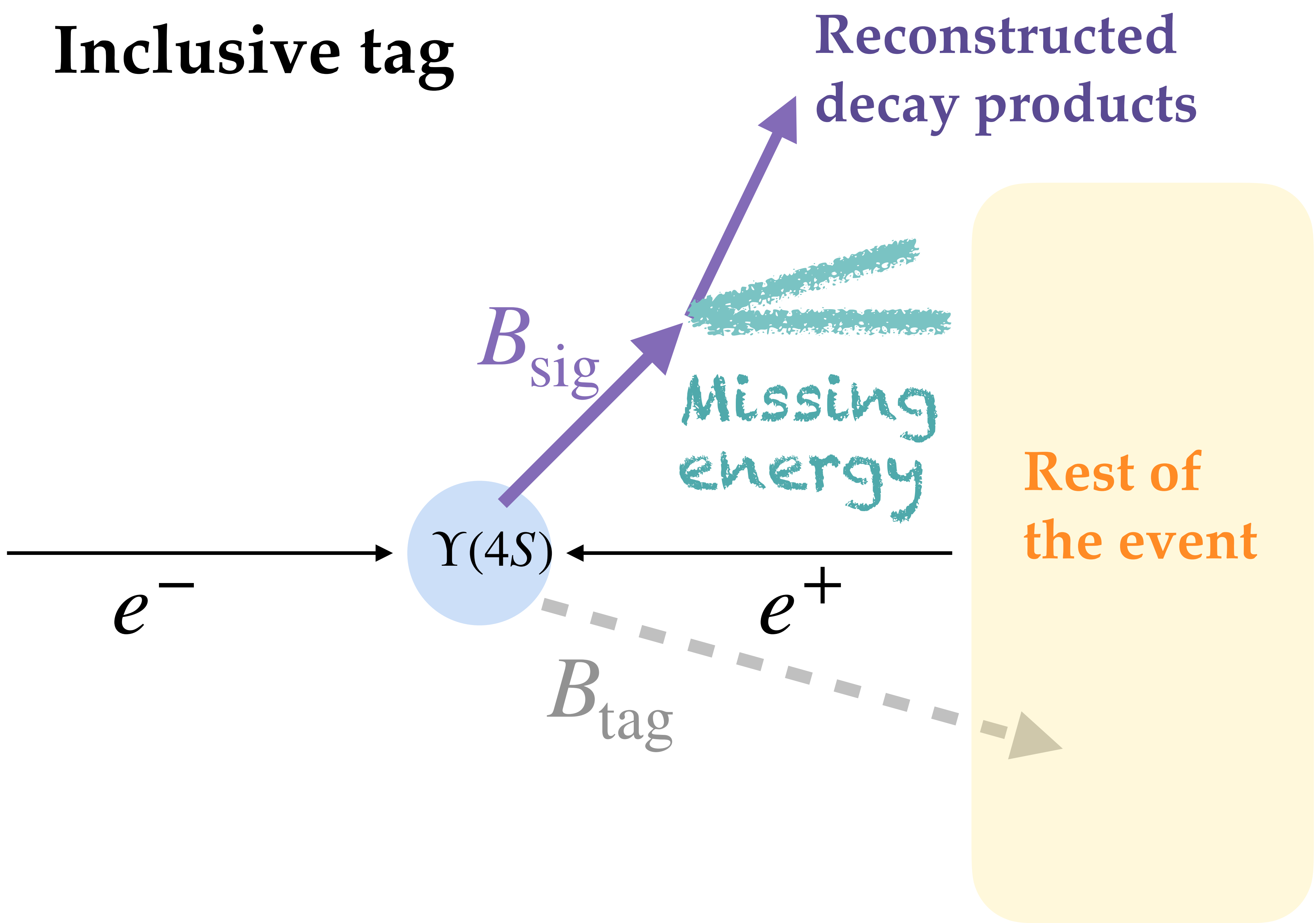
- The sensitivity for BR($B \to KX_{inv}$) will be improved wrt the 3-body decay
- Shape analysis on the basic SM search or optimization of the selection (m_X dependent)
- **The HTA will gain importance due to the better resolution in q^2**

Several dark sector scenarios could be constrained: Higgs-mixing, ALP, QCD axion, axiflavon



Sensitive to low mass scale (~ GeV) new physics

Inclusive tag



$$q^2_{\text{rec}} = s/4 + M_K^2 - \sqrt{s}E_K^*$$

Resolution is ~ 0.1 GeV2

The tools

ROOT

RooFit

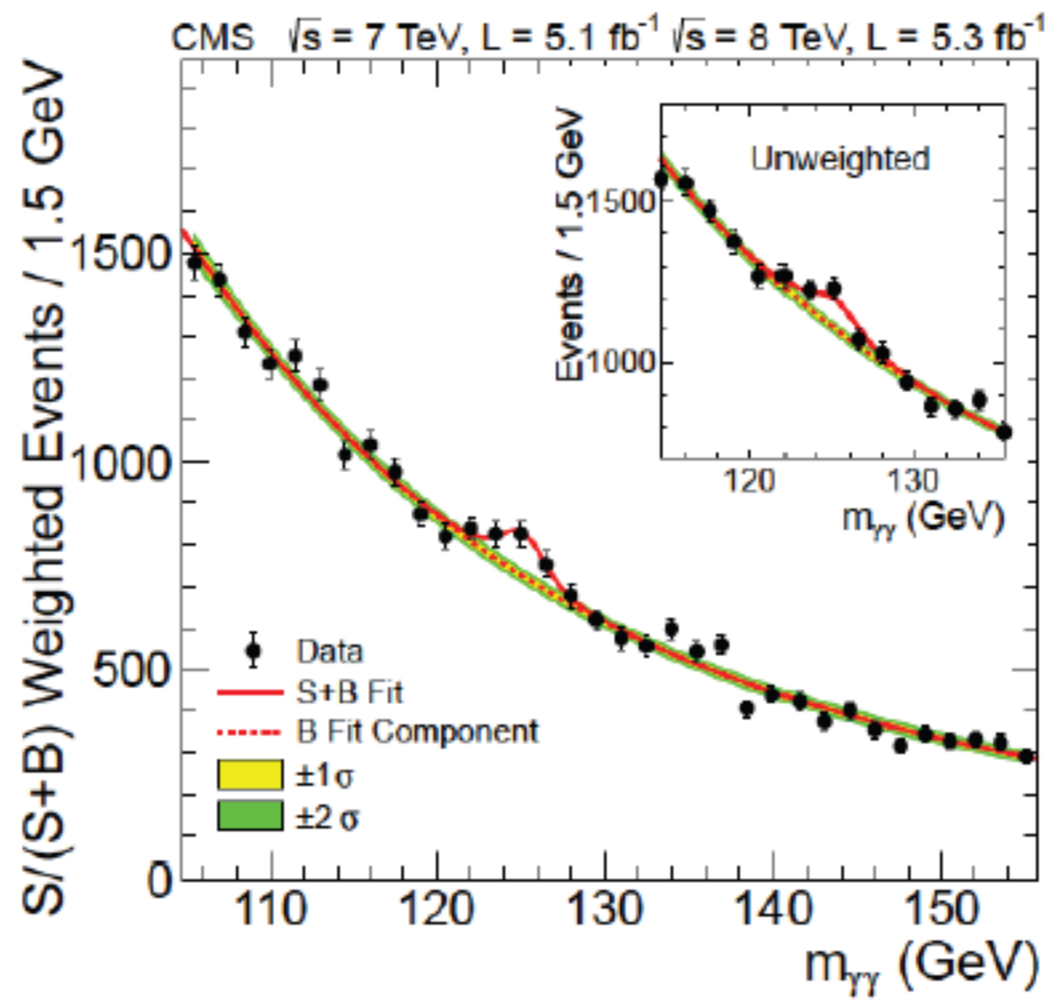
RooStats (binned and **unbinned**)
https://root.cern/doc/v630/group__Roostats.html

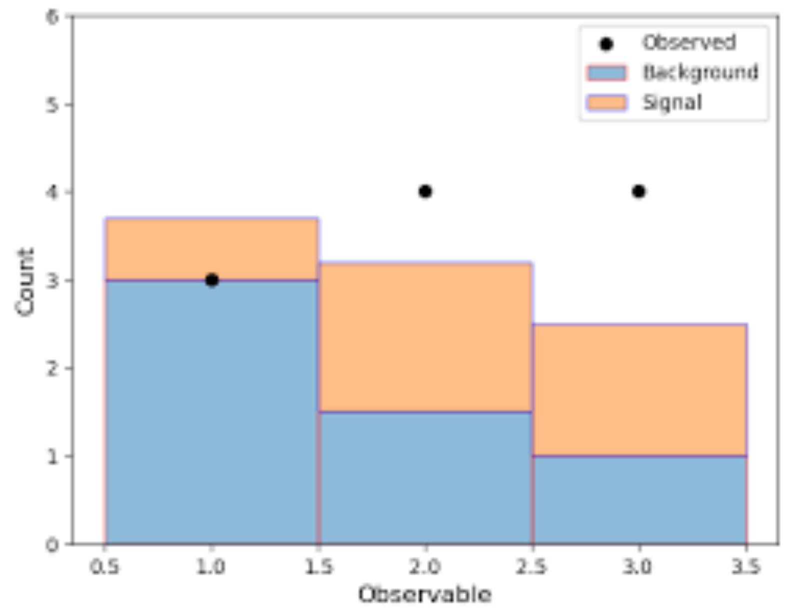
Binned analysis tool (workspace,..)

HistFactory [CERN-OPEN-2012-016]
https://root.cern/doc/v630/group__HistFactory.html

In Python

PyHF
https://scikit-hep.org/pyhf/





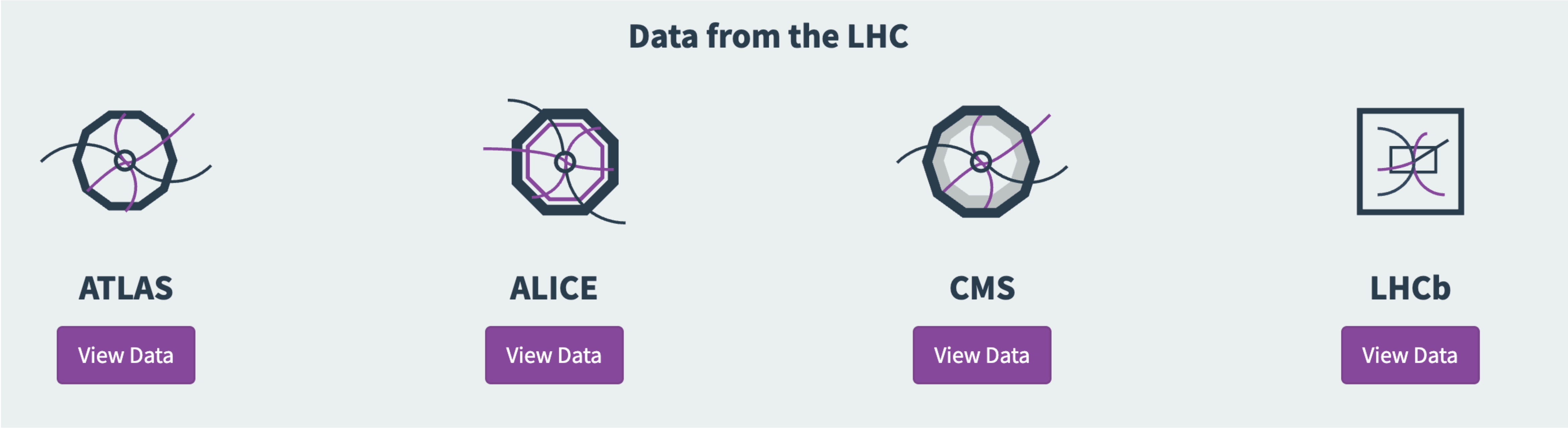
pyhf
∂ifferentiable
ℒikelihoods

Do we have an analogous for shape analysis?

Maybe Zfit? [zfit.loss.UnbinnedNLL]

Sharing results

HEPData: https://www.hepdata.net/



BelleII
https://www.hepdata.net/search/?q=belle&page=1&collaboration=Belle-II

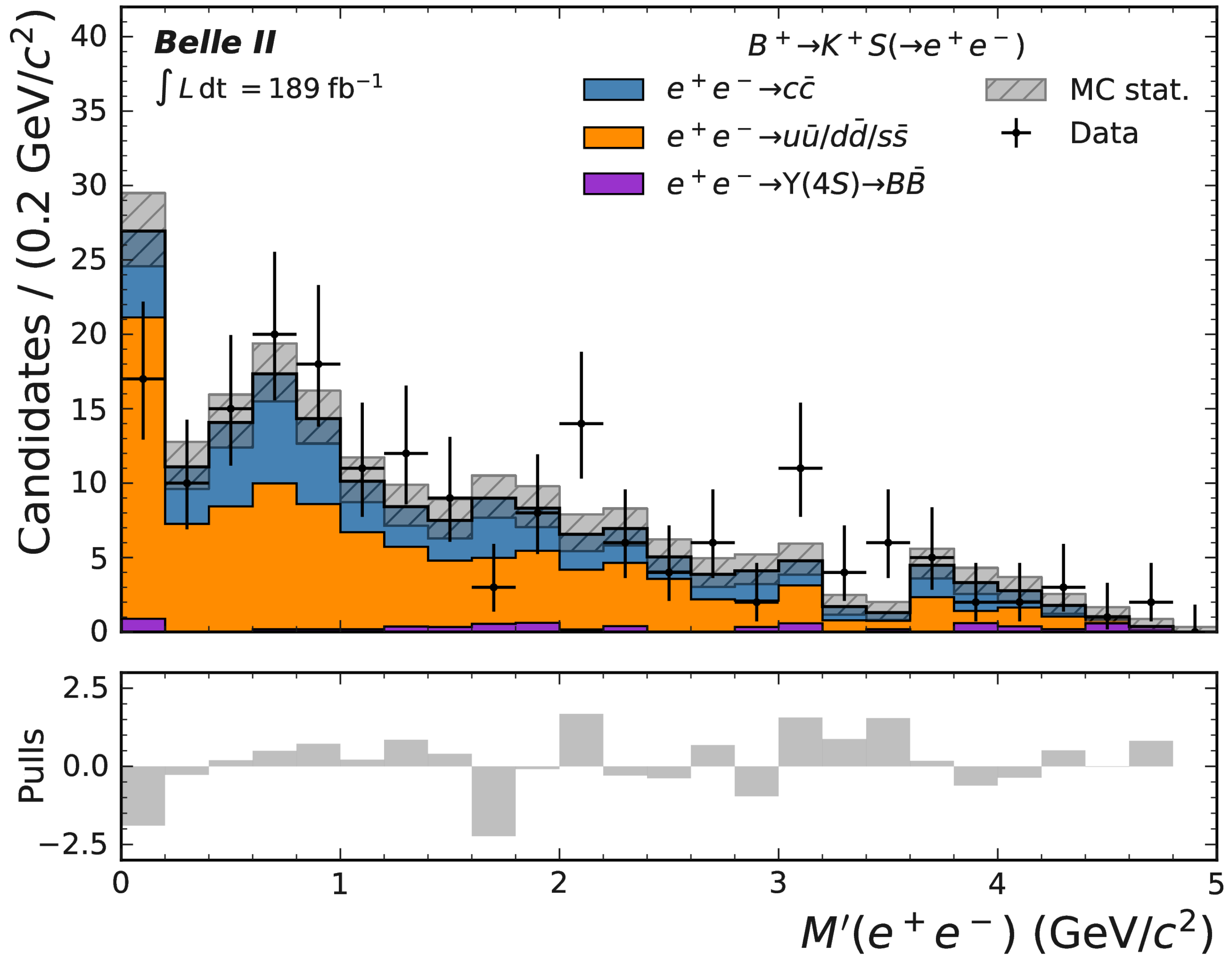
Example:

Version 2 **Search for a long-lived spin-0 mediator in $b \to s$ transitions at the belle II experiment**

The Belle-II collaboration Adachi, I. ; Adamczyk, K. ; Aggarwal, L. ; *et al.*

Phys.Rev.D 108 (2023) L111104, 2023.

Inspire Record 2665757 DOI 10.17182/hepdata.147283

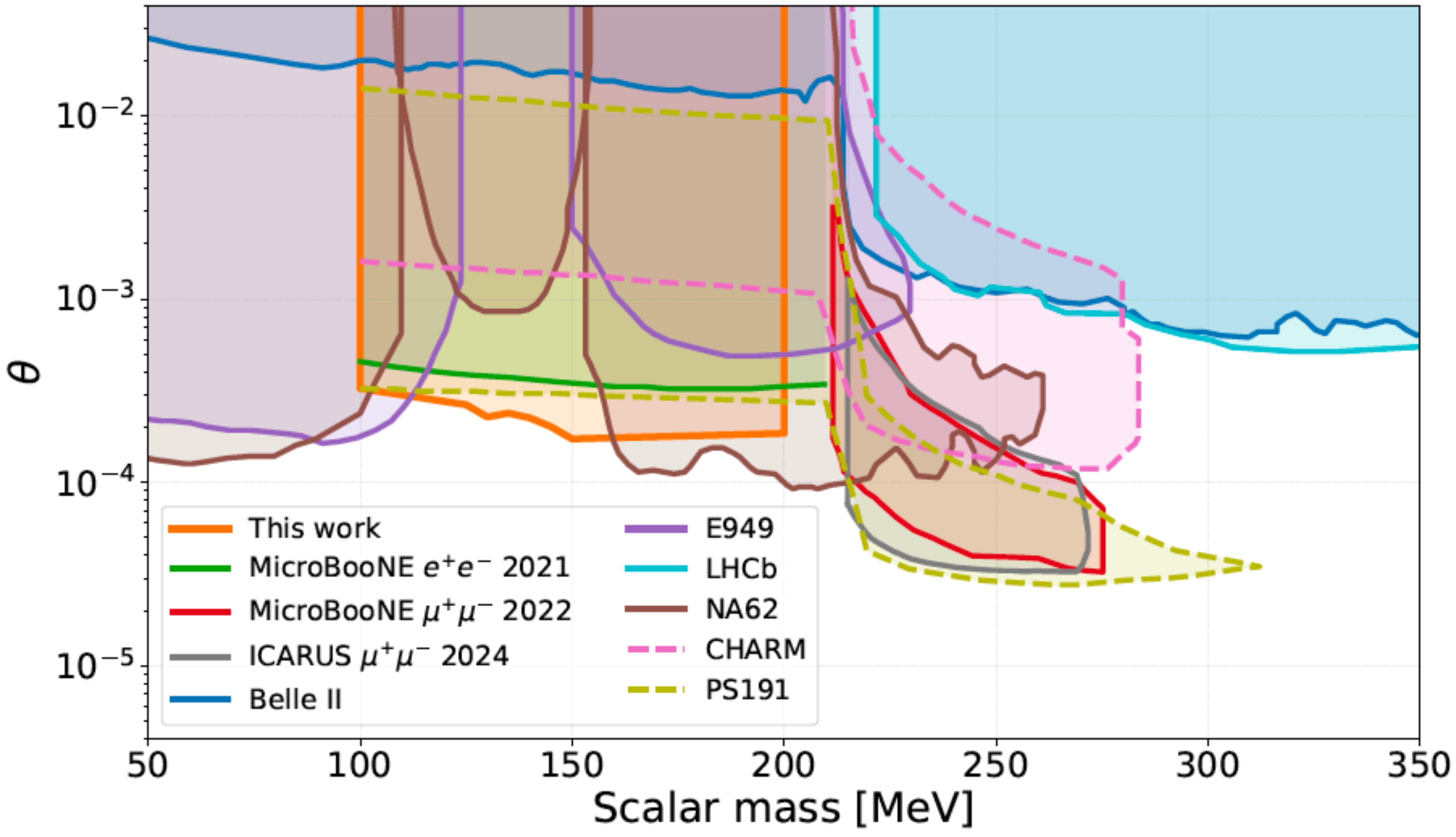


Combinations in dark sector?

We have many strong limits to several benchmarks of dark sector portals, we usually show **comparisons among different experiments and analyses** (for example Physics Beyond Colliders).

Can we do a combination of different experiments? With the share of the likelihood it could be possible

For example (learned about this result by MicroBooNE yesterday):



Many very different experiments (neutrino, B-factories, hadron colliders, Kaon factories….)

That’s all….

What do other experiments use?

Back-up

HistFactory [CERN-OPEN-2012-016]

$$\mathcal{P}(n_b|\mu) = \mathrm{Pois}(n_{\mathrm{tot}}|\mu S + B)\left[\prod_{b\in\mathrm{bins}} \frac{\mu\nu_b^{\mathrm{sig}} + \nu_b^{\mathrm{bkg}}}{\mu S + B}\right] = \mathcal{N}_{\mathrm{comb}} \prod_{b\in\mathrm{bins}} \mathrm{Pois}(n_b|\mu\nu_b^{\mathrm{sig}} + \nu_b^{\mathrm{bkg}})$$

i: bin index,
n_i: number of events in bin "i"

Counting, bins

$$\prod_i \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-\mu s_i - b_i}$$

Binned

Unbinned

$$k^{-1} \prod_i (\mu S \boxed{f_s(x_i)} + B \boxed{f_b(x_i)}) \cdot e^{-(\mu S + B)}$$

i: index of the event
k: total number of events
x: a discriminant variable

S, B, expected number of signal and background, they are treated as in the counting experiment

They are not fixed ==> extended fit in the profiling

Unbinned analysis in RooStats

$$\mathrm{f}_{\mathrm{tot}}(\mathcal{D}_{\mathrm{sim}}, \mathcal{G}|\alpha) = \prod_{c \in \mathrm{channels}} \left[\mathrm{Pois}(n_c|\nu_c(\alpha)) \boxed{\prod_{e=1}^{n_c} f_c(x_{ce}|\alpha)} \right] \cdot \boxed{\prod_{p \in \mathbb{S}} f_p(a_p|\alpha_p)}$$

Kyle Cranmer
arXiv:1503.07622v1

Distribution of the discriminant variable

Constraint terms
Information about nuisance parameters

<u>Global observable:</u> value of the nuisance parameter estimated with an auxiliary measurement
Pseudo-experiments are tossed for several values of the global observable

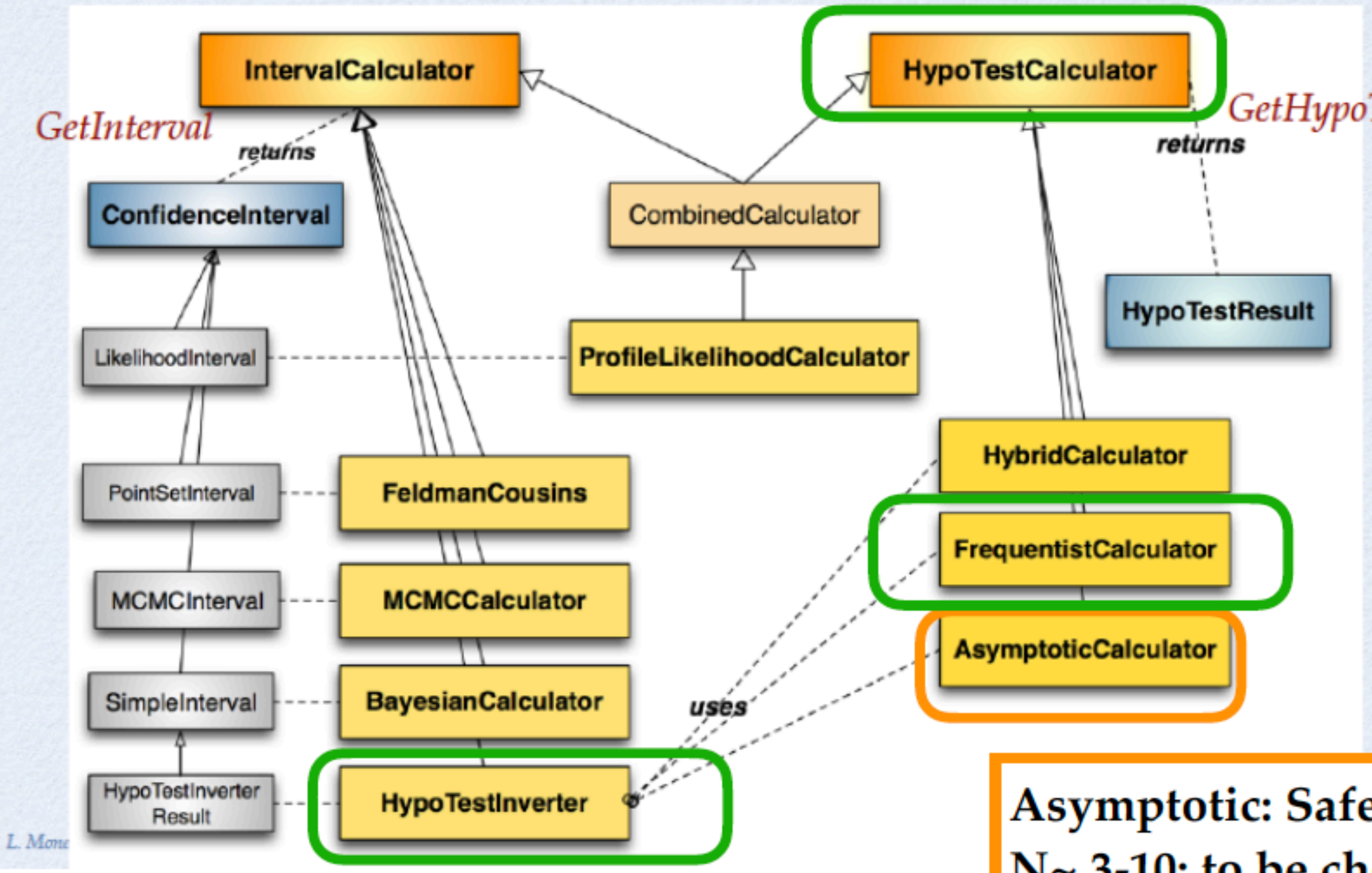
Global observable

Nuisance parameter

$$\prod_{p \in \mathbb{S}} f_p(\boxed{a_p}|\boxed{\alpha_p})$$

RooStats Design

- C++ interfaces and classes mapping to real statistical concepts



Asymptotic: Safe for N>~ 10
N~ 3-10: to be checked

From official ROOT training

Data means
- real data for observed upper limit
- pseudo data for expected upper limit

FrequentistCalculator

- Generate toys using nuisance parameter at their conditional ML estimate ($\theta = \hat{\theta}_{\mu}$) by fitting them to the observed data
- Treat constraint terms in the likelihood (e.g. systematic errors) as auxiliary measurements
  - introduce global observables which will be varied (tossed) for each pseudo-experiment
  - $L = \text{Poisson}(n_{obs} \mid \mu + b)\,\text{Gaussian}(b_0 \mid b, \sigma_b)$
    - b_0 is a global observables, varied for each toys but it needs to be considered constant when fitting
    - n_{obs} is the observable which is part of the data set
    - μ is the parameter of interest (poi)
    - b is the nuisance parameter

This is an example with gaussian bkg!

(and Gaussian is symmetric…)

Constraints for nuisance parameters

Gamma distribution $n_p^0 = \tau_p = (1/\sigma_p^{\text{rel}})^2$

(obtained from the ON/OFF problem)

$$\boxed{\rho(n) = \frac{1}{\alpha}\,\frac{(n/\alpha)^N}{N!}\,\exp(-n/\alpha)}$$

Used for bkg yield

PDF	Likelihood $\propto$	Prior π_0	Posterior π
$G(a_p \vert \alpha_p, \sigma_p)$	$G(\alpha_p \vert a_p, \sigma_p)$	$\pi_0(\alpha_p) \propto \text{const}$	$G(\alpha_p \vert a_p, \sigma_p)$
$\text{Pois}(n_p \vert \tau_p \beta_p)$	$P_\Gamma(\beta_p \vert A = \tau_p; B = 1 + n_p)$	$\pi_0(\beta_p) \propto \text{const}$	$P_\Gamma(\beta_p \vert A = \tau_p; B = 1 + n_p)$
$P_{\text{LN}}(n_p \vert \beta_p, \sigma_p)$	$\beta_p \cdot P_{\text{LN}}(\beta_p \vert n_p, \sigma_p)$	$\pi_0(\beta_p) \propto \text{const}$	$P_{\text{LN}}(\beta_p \vert n_p, \sigma_p)$
$P_{\text{LN}}(n_p \vert \beta_p, \sigma_p)$	$\beta_p \cdot P_{\text{LN}}(\beta_p \vert n_p, \sigma_p)$	$\pi_0(\beta_p) \propto 1/\beta_p$	$P_{\text{LN}}(\beta_p \vert n_p, \sigma_p)$

Lognormal distribution

Used for normalization (multiplicative) factors:

$$\boxed{\rho(\theta) = \frac{1}{\sqrt{2\pi}\ln(\kappa)} \exp\left(-\frac{(\ln(\theta/\tilde{\theta}))^2}{2(\ln\kappa)^2}\right)\frac{1}{\theta}}$$