Analysis Presentation: V_{cb}

Part III: From Signal Selection to the Result

Kilian Lieret, Markus Prim

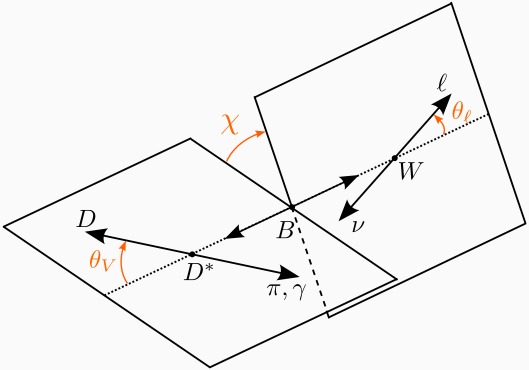
25th March 2021

Universität Bonn - Physikalisches Institut

Signal Reconstruction

$B \to D\ell\nu$
Signal variable: w

$B \to D^*\ell\nu$
Signal variable: w, $\cos\theta_\ell$, $\cos\theta_V$, χ



Signal Reconstruction

Analysis Strategy

Signal Reconstruction

$B \rightarrow D\ell\nu$

$B \rightarrow D^{*}\ell\nu$

$B \rightarrow (D^{(*)}\pi)\ell\nu$

$+$

Full Event Interpretation

Monte Carlo Corrections

BFs + Form Factors for

$B \rightarrow D\ell\nu$

$B \rightarrow D^{*}\ell\nu$

$B \rightarrow D^{**}\ell\nu$

New Gap Description

pID, Tracking, slow pion and K^0_S Eff. Corrections

Signal Extraction

2D Likelihood Fit

in m^2_{Miss} **and** p^*_ℓ

Separating **Signal** and **Normalisation** Modes

$+$ Constraint of D^{**} Modes

Data for other Analyses:

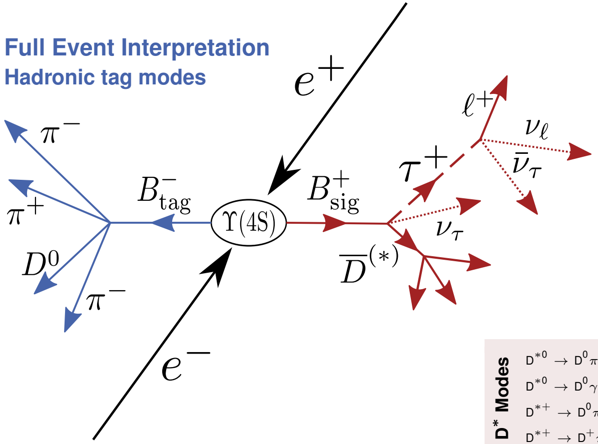
$B \rightarrow D^{**}\ell\nu$ Study

V_{cb} Measurement

$\mathcal{R}(D^{(*)})$ Determination

Reconstruction

Full Event Interpretation
Hadronic tag modes



D^- Modes

$D^+ \to K^- \pi^+ \pi^+$
$D^+ \to K^- \pi^+ \pi^+ \pi^0$
$D^+ \to K^- \pi^+ \pi^+ \pi^+ \pi^-$
$D^+ \to K^0_S \pi^+$
$D^+ \to K^0_S \pi^+ \pi^0$
$D^+ \to K^0_S \pi^+ \pi^+ \pi^-$
$D^+ \to K^0_S K^+$
$D^+ \to K^+ K^- \pi^+$

D^0 Modes

$D^0 \to K^- \pi^+$
$D^0 \to K^- \pi^+ \pi^0$
$D^0 \to K^- \pi^+ \pi^+ \pi^-$
$D^0 \to K^- \pi^+ \pi^+ \pi^- \pi^0$
$D^0 \to K^0_S \pi^0$
$D^0 \to K^0_S \pi^+ \pi^-$
$D^0 \to K^0_S \pi^+ \pi^- \pi^0$
$D^0 \to K^- K^+$

D^* Modes

$D^{*0} \to D^0 \pi^0$
$D^{*0} \to D^0 \gamma$
$D^{*+} \to D^0 \pi^+$
$D^{*+} \to D^+ \pi^0$

Signal Reconstruction

B_{tag}–Reconstruction

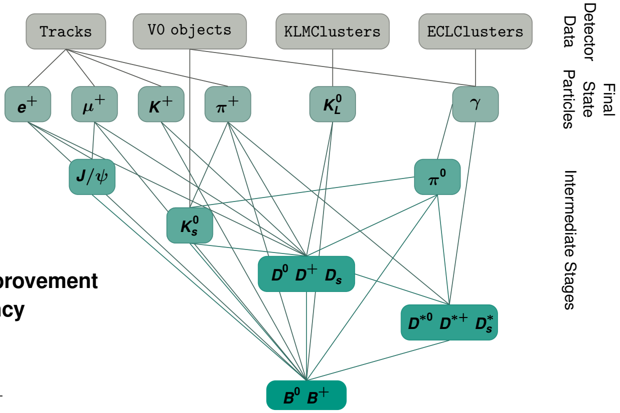
KIT Karlsruhe Institute of Technology

Full Event Interpretation[1]

Fully automated hierarchical reconstruction of tag-side B-mesons in

- **hadronic** and/or
- semi-leptonic

decay modes.



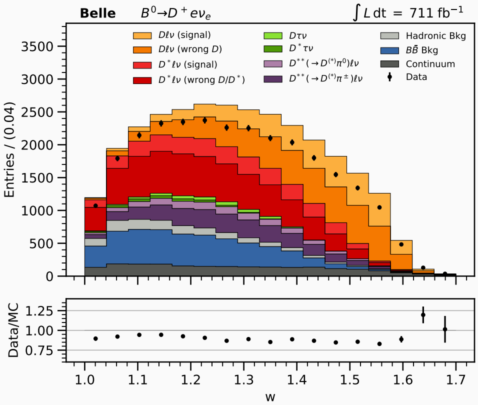
$\Rightarrow$ **Expecting significant improvement of reconstruction efficiency**

1: Computing and Software for Big Science 3:6 (2019)

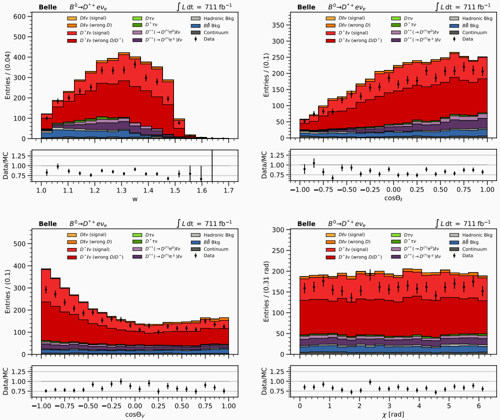
Signal Reconstruction (only $B^0 \to D^{(*)+} e \nu_e$)

- After very loose selection criteria applied, we find:

$B^0 \to D^+ e \nu_e$



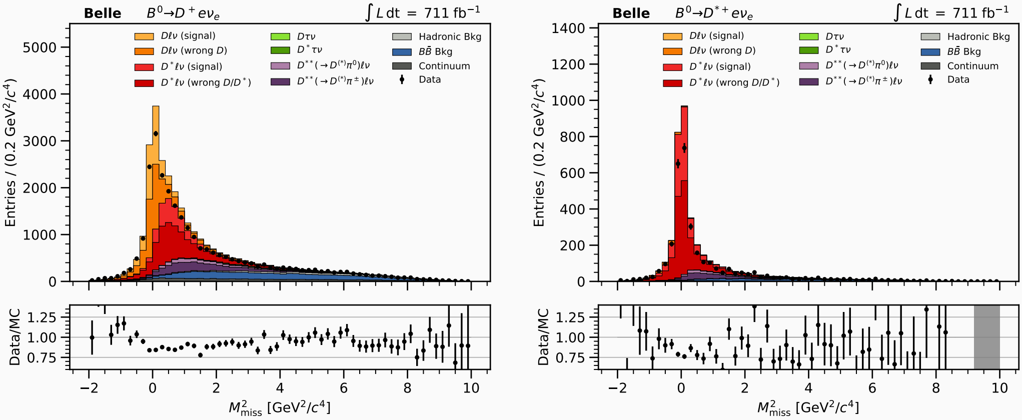
$B^0 \to D^{*+} e \nu_e$



To fit the distributions, we need to remove our background.

Background Subtraction

- Find a variable which is model independent to separate signal from background.
- Candidate for this analysis: M^2_{miss} (Alternative: $E_{\mathrm{miss}} - p_{\mathrm{miss}}$)



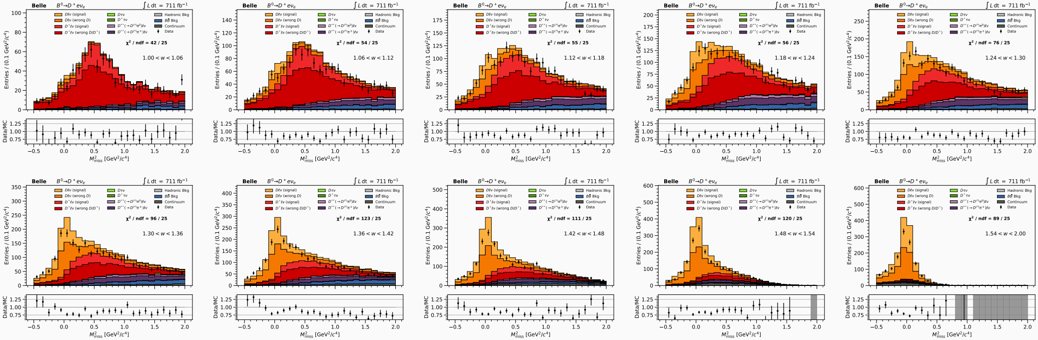
Background Subtraction

From now on we focus on $B^0 \to D^+ e\nu_e$, but everything applies to $B^0 \to D^{*+} e\nu_e$.

- We are interest in the background-free differential distribution of w.
- Extract signal yield in each w bin by fitting M^2_{miss}.

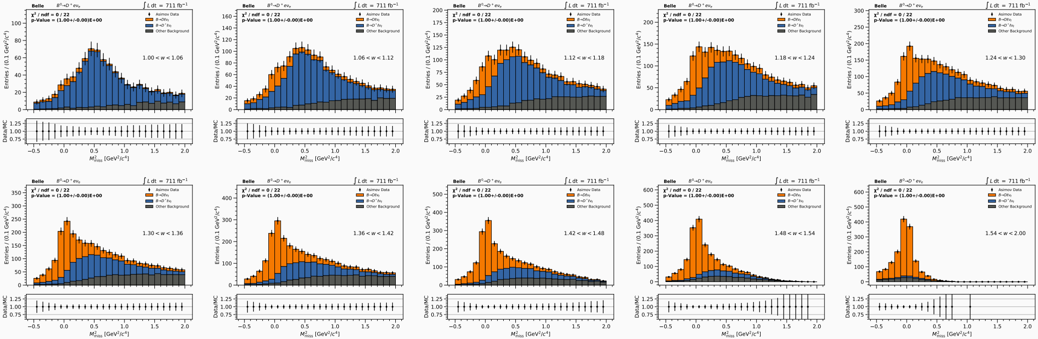
Background Subtraction

- Split the M^2_{miss} into the desired w-bins.
- Group sensible categories together.



Background Subtraction

- Next step: Fit signal yields, but we see that the MC does not describe data well.
- $\Rightarrow$ Develop method that handles that.



Sometimes you hit a roadblock.

Asimov Data

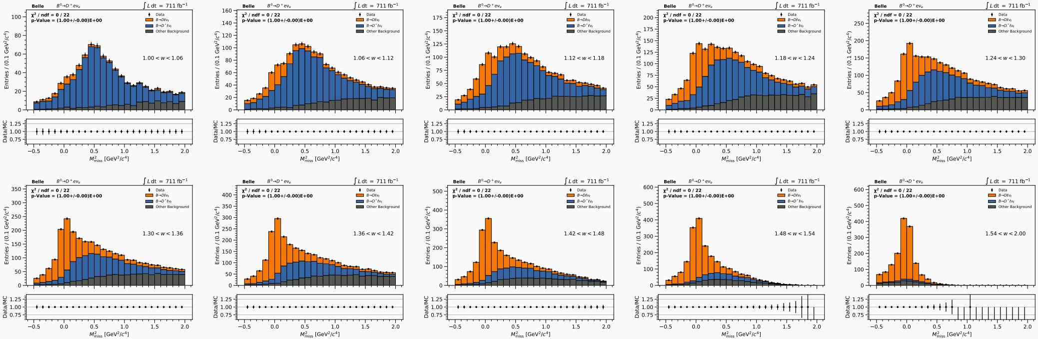
Definition

- The Asimov data set is defined as such that when one uses it to evaluate the estimators for all parameters, one obtains the true parameter values.
- In practice this means creating data histograms with entries equivalent to MC.

The advantages of working with Asimov data are:

- Development of the analysis not on data – the 'box' is closed.
- Closure tests are available, e.g. for your signal extraction.
- In pratice it is very useful to make sure what you developed works properly, but your method always has to survive the application on data – check your methods on (sideband) data early and often.

Asimov Data – Example



The Fit Model for Background Subtraction

Fit Model

Straightforward fit model to extract signal yields: $\mathcal{L}(\nu) = \prod_i^{\text{bins}} \mathcal{P}(n_i | \nu_i(\theta))$

- The Poisson distribution $\mathcal{P}$, the number of measured events n_i in bin i and the number of expected events ν_i in bin i.
- The expectation $\nu_i = \sum_k^{\text{templates}} f_{ik}\eta_k$, where η_k is the total number of expected events for template k and f_{ik} is the fraction of events expected in bin i of template k.
- The fractions are defined as $f_{ik} = \frac{\eta_{ik}}{\sum_{j=1}^{\text{bins}} \eta_{jk}}$

$\Rightarrow$ See "Fitting: A guided overview" by Angelo Di Canto for a more thorough and general discussion on fitting.

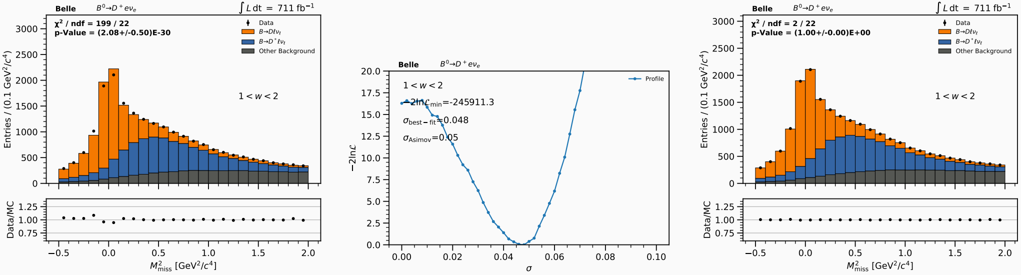
The Fit Model for Background Subtraction

Modified Fit Model
Instead of creating the templates once before the fit, at each Likelihood evaluation we apply a smearing to the unbinned data:

$M_{\mathrm{miss}}^2 \rightarrow M_{\mathrm{miss}}^2 \times \mathcal{S}$

- Draw a random number for each event from $\mathcal{S}$ and smear the distribution by doing that.
- Example: Gaussian Smearing $\mathcal{S} = \mathcal{G}(\mu = 0, \sigma)$ with
- Develop method on a well understood sample: Smear the Asimov data

Fit Model on Smeared Asimov Data



Our model can extract the resolution from the Asimov data.

The Roadblock

End of this analysis

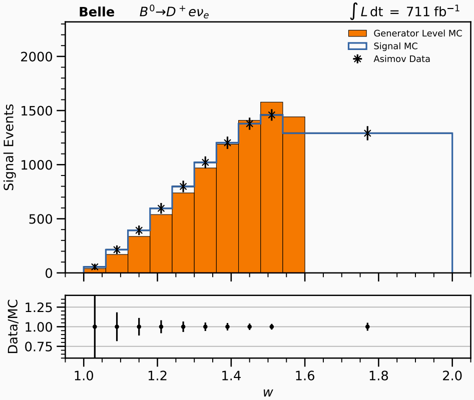
- Our story on Belle data ends now, as everything shown is work in progress,
- ... and the resolution effect can not be described by a simple Gaussian.

Cooking a solution

- Fit the resolution in each channel / bin individually, link the resolution, or fit on the unbinned distribution?
- Smear only signal or all components?
- Which function actually describes the resolution effect?

Let as assume for a moment, that we were able to create the background subtracted spectrum and find ...

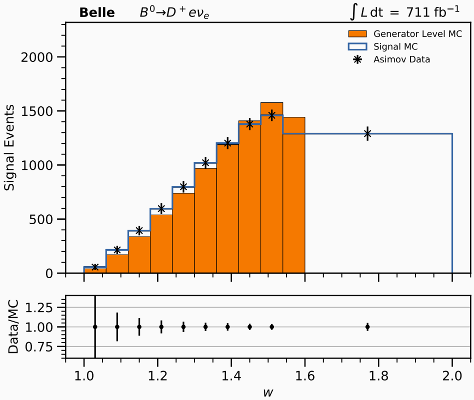
Background Subtracted Spectrum



Comparing Theory with Experimental Data

- Either ‘forward-fold’ the theory or unfold the experimental data.
- The ‘forward-folding’ can only be done by the experimentalist (you), because it requires access to the full data. For semileptonic decays there is the HAMMER framework to do that [2002.00020].
- Unfolding experimental data makes the result available for everyone and can be revisited in the future and/or compared to other experiments. But unfolding has its own problems.
- You can apply both methods and check if you can consistently extract your parameter(s) of interest.

Comparing Theory with Experimental Data



Forward-Folding

Vary parameters of interest in the orange distribution, apply detector effects again, compare to experimental data.

Unfolding

Invert that process, by 'removing' the detector effects from the experimental data.

The mathematical problem: $$\underbrace{f_{\mathrm{measured}}(x)}_{\text{Simulated Data}} = \int \underbrace{R(x|y)}_{\text{Detector}} \quad \underbrace{f_{\mathrm{true}}(y)}_{\text{Event Generator}} \quad \mathrm{d}y$$

Unfolding [A SURVEY OF UNFOLDING METHODS FOR PARTICLE PHYSICS by G. Cowan]

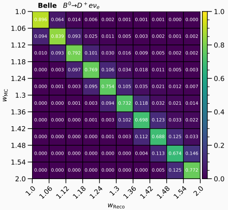
Unfolding Methods

There are several unfolding methods ‘on the market’, e.g.

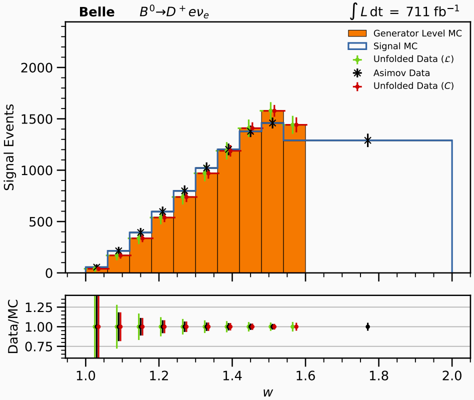
- Bin-by-bin unfolding, with a correction factor per bin $C_i = \mu_i^{\mathrm{MC}} / \nu_i^{\mathrm{MC}}$
  - $\hat{\mu}_i = C_i n_i$
- Inverting the response matrix R
  - $\hat{\mu}_i = R_{ij}^{-1} n_j$
  - χ^2-based unfolding: $\chi^2 = (\mathsf{w}_{\mathrm{Reco}} - \mathsf{R}\mathsf{w}_{\mathrm{MC}})\mathsf{C}^{-1}(\mathsf{w}_{\mathrm{Reco}} - \mathsf{R}\mathsf{w}_{\mathrm{MC}})^{\mathsf{T}}$
  - Similar approach to inverting the response matrix, without actually inverting it.

Migration Matrix

- $R_{ij} = P(\text{observed in bin i} | \text{true value in bin j})$
- With a perfect detector and reconstruction this would be diagonal.



Unfolded Spectrum



As our analysis is WIP, we fall back now on an already unfolded spectrum from a previous measurement.

Interpretation of the Result

After the hard work of the analysis is done, use time to interpret it!

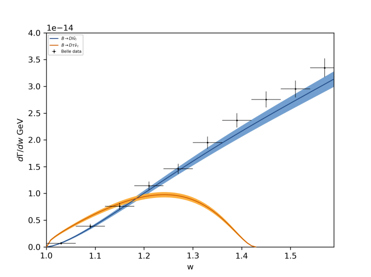
- Analysis often done with single purpose in mind.
- Can your data be interpreted in other contexts
- ... or can you find a new / better ways to interpret your results?

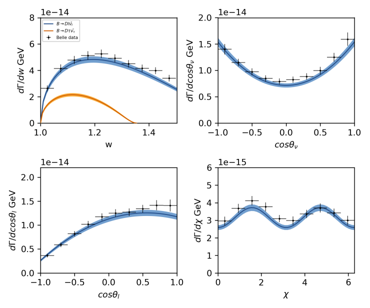
Recap: Why do we look at the Belle data again?

- V_{cb} extraction was already performed using the full data set, but we have now improved algorithms.
- We analyze $B \to D\ell\nu$ and $B \to D^{*}\ell\nu$ simultaneously. Increases complexity, but could potentially improve the result and yield new insights by correlating the two measurements, e.g. coupling yields.

Linking $B \to D\ell\nu$ and $B \to D^*\ell\nu$ [1703.05330]

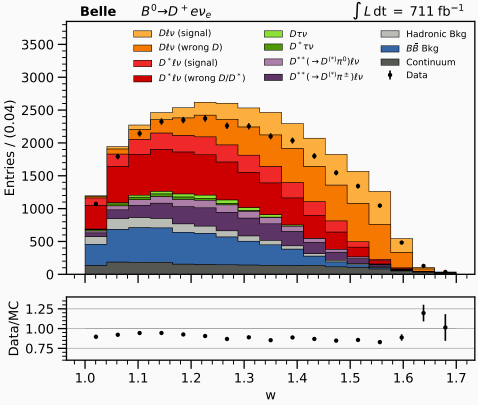
- Historically only CLN, then also BGL form factors fitted.
- There exists also a form factors parametrization which links the form factors of $B \to D\ell\nu$ and $B \to D^*\ell\nu$: BLPR.
- Additional new result over 'only' repeating the $|V_{\mathrm{cb}}|$ fits with BGL and CLN.





Interpretation of the Result

$$B^0 \to D^+ e \nu_e$$



Utilize the result to the fullest

- There is a statistical independent D^* sample in the $B \to D\ell\nu$ channels.
- Perform form factor fit and V_{cb} extraction on that.
- Caveat: Only w available, not the helicity angles.
- Statistical independent new result.

Take-Away Message for your own Analysis

- Develop a method on Asimov data, test it on Toy MC, and run it on data. This gives you a lot of control and understanding what is working and what is not.
- Check your sidebands early and often, to catch problems early.
- Do not get discouraged if you hit a roadblock. There is always a solution.
- Take time and develop new (and better) methods for old problems.
- Squeeze your analysis for more results, don't let someone else pick the low hanging fruits.