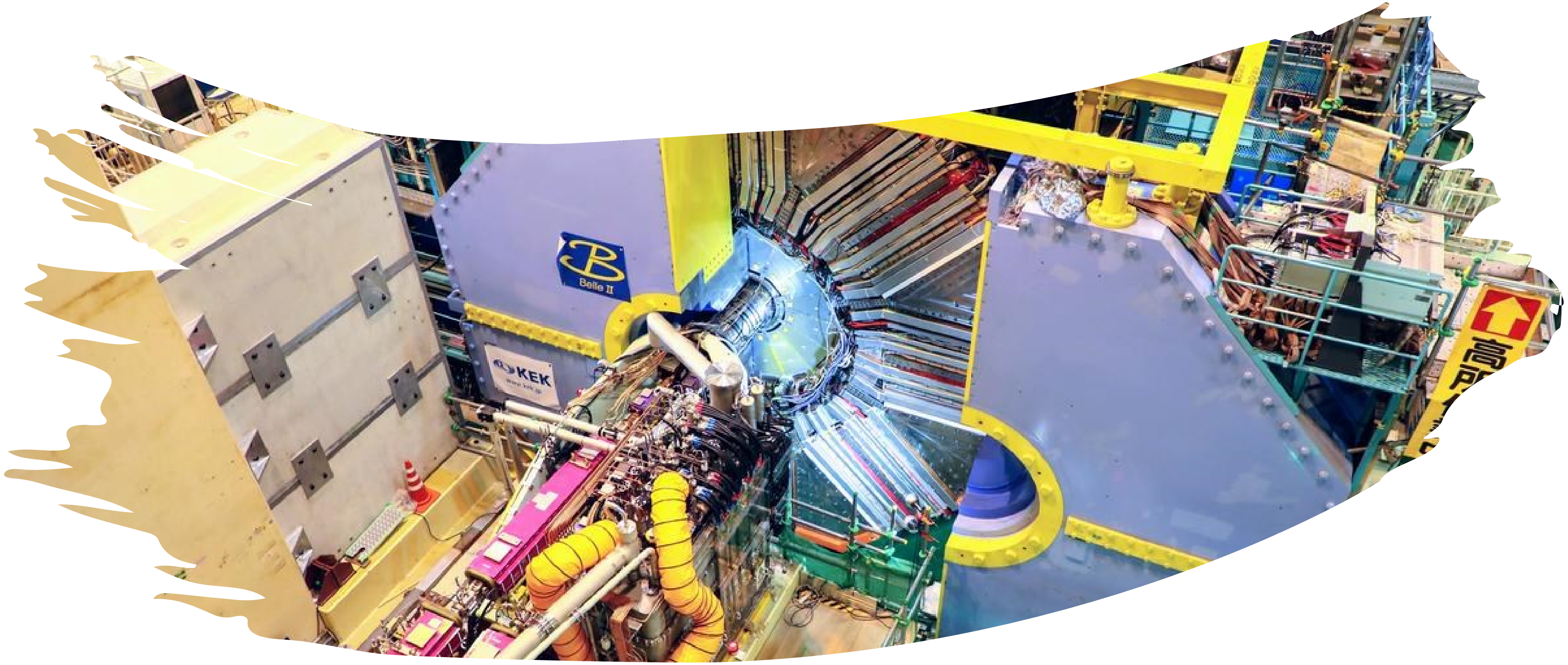


Beyond Unfolded Distributions

Which measurements can we provide to improve knowledge on form factors? What are the shortcomings of unfolded distributions? What can we provide beyond/along-with them? What can angular analyses add?

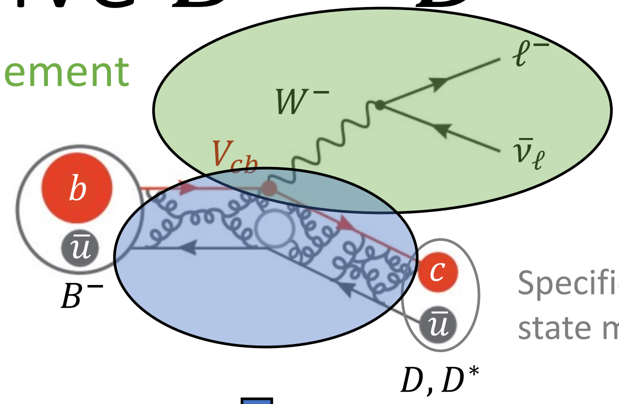
Markus Prim
markus.prim@uni-bonn.de

Bundesministerium für Bildung und Forschung

UNIVERSITÄT BONN

Exclusive $B \to D^{(*)}\ell\bar{\nu}_\ell$

Leptonic Matrix Element



Specific final state meson

$$|V_{cb}|^2 \propto \frac{\Gamma(B \to D\ell\bar{\nu}_\ell)}{\mathcal{G}(1)} \propto \frac{\Gamma(B \to D^*\ell\bar{\nu}_\ell)}{\mathcal{F}(1)}$$

$$\mathcal{G}(1) = h_+(1) \quad \mathcal{F}(1) = h_{A_1}(1)$$

Hadronic Matrix Elements can not be calculated from first principles

→ Can be parameterized with form factors $h_X = h_X(w)$ and extracted from data

→ Lattice QCD must provide (at least) inputs on their normalization

$$\frac{\langle D(p')|\bar{c}\gamma^\mu b|B(p)\rangle}{\sqrt{m_B m_D}} = h_+(v+v')^\mu + h_-(v-v')^\mu$$

$$\frac{\langle D^*(p')|\bar{c}\gamma^\mu b|B(p)\rangle}{\sqrt{m_B m_{D^*}}} = h_V \epsilon^{\mu\nu\alpha\beta}\epsilon^*_\nu v'_\alpha v_\beta$$

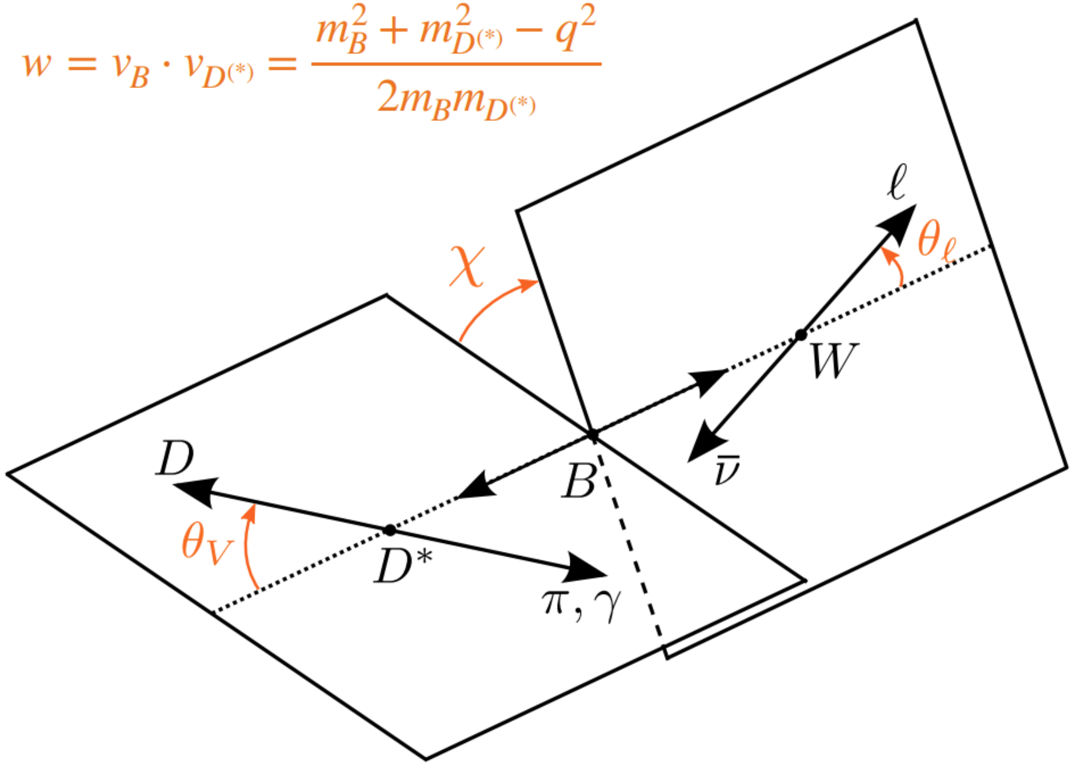
$$\frac{\langle D^*(p')|\bar{c}\gamma^\mu\gamma^5 b|B(p)\rangle}{\sqrt{m_B m_{D^*}}} = h_{A_1}(w+1)\epsilon^{*\mu} - h_{A_2}(\epsilon^*\cdot v)v^\mu - h_{A_3}(\epsilon^*\cdot v)v'^\mu$$

Heavy Quark Symmetry Basis

Common parameterizations for the form factors:
BGL, CLN, BLPRXP

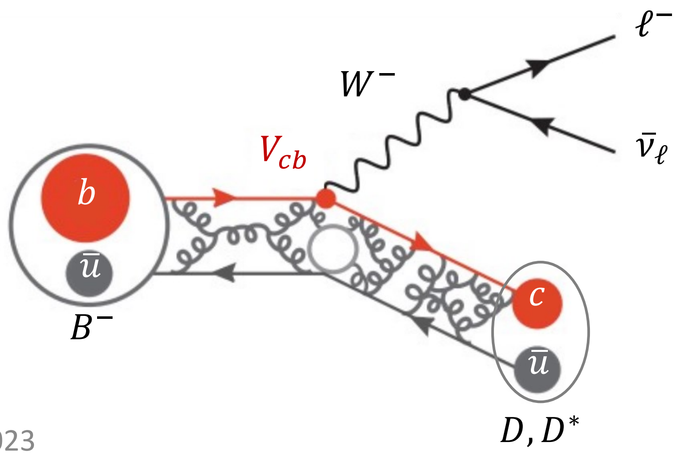
Exclusive $B \to D^{(*)}\ell\bar{\nu}_\ell$

$$w = v_B \cdot v_{D^{(*)}} = \frac{m_B^2 + m_{D^{(*)}}^2 - q^2}{2m_B m_{D^{(*)}}}$$

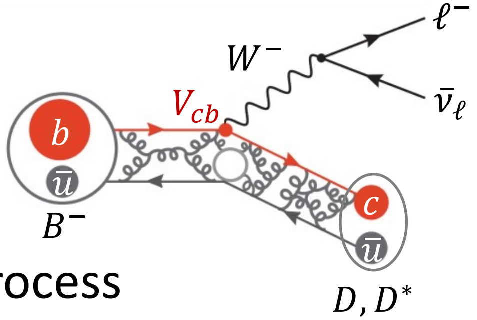


- Form factors are a function of w only, and measuring w is sufficient to determine them
- Angles provide information on, e.g.
  - Decorrelation of form factors
  - Forward-backward asymmetry
  - Longitudinal polarization fraction
  - Angular asymmetries sensitive to new physics
- Measure the marginal distributions for
  - $B \to D^*\ell\bar{\nu}_\ell \to \mathrm{w}, \cos\theta_\ell, \cos\theta_V, \chi$
  - $B \to D\ell\bar{\nu}_\ell \to \mathrm{w}, \cos\theta_\ell$

Combined $B \to D^{(*)} \ell \bar{\nu}_\ell$



Combined $B \to D^{(*)} \ell \bar{\nu}_\ell$ measurements

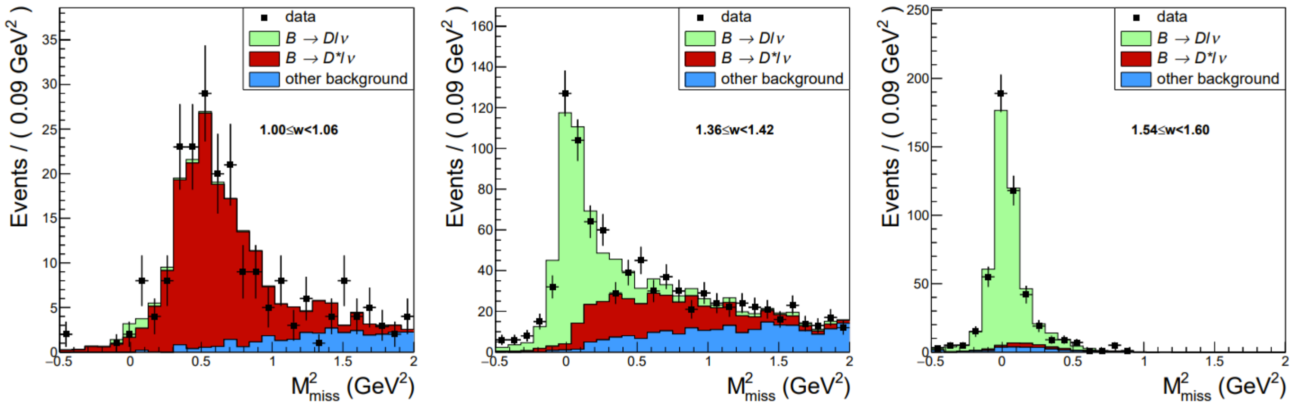


Theoretical Reasons

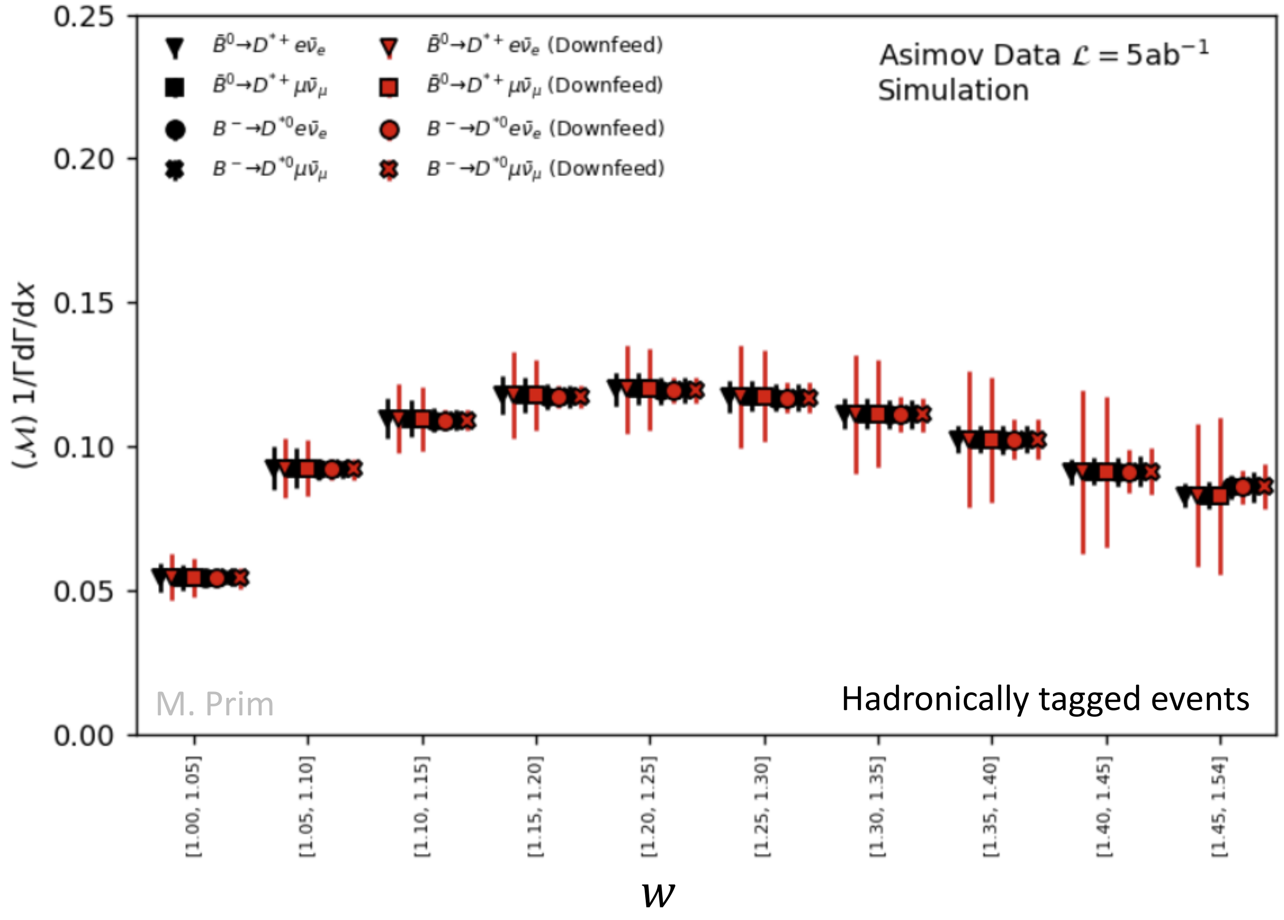
- Identical parton level process
- Combined analysis in e.g., HQET allows stronger constraints on form factors

Experimental Reasons

- Analysis strategies very similar
- Only we (experimentalists) can provide the correct correlations between the two decays
- Constraining the D^* downfeed background and treating it as signal improves
  - $B \to D\ell\bar{\nu}_\ell$ signal extraction
  - Recovers events with missed slow pions in $B \to D^*\ell\bar{\nu}_\ell$ yielding more events at low w (recovering "all" D^*)



Combined $B \to D^{(*)}\ell\bar{\nu}_\ell$ measurements

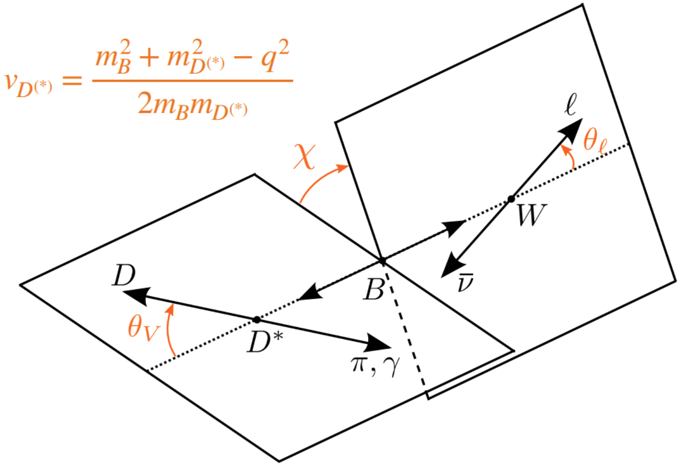


"Regular" $B \to D^*\ell\bar{\nu}_\ell$ precision
"Downfeed" $B \to D^*\ell\bar{\nu}_\ell$ precision
→ "all D^* recovered"

Claim 1: Combined $B \to D^{(*)}\ell\bar{\nu}_\ell$ will provide better precision on $|V_{cb}|$ than individual results.

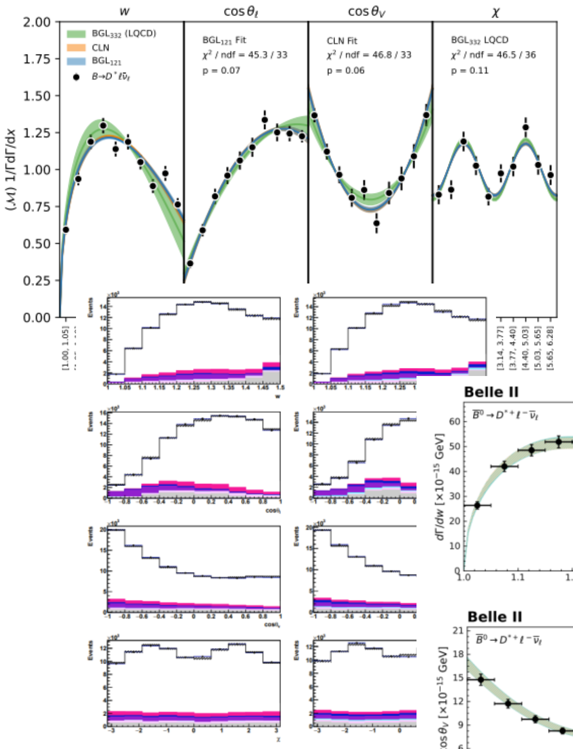
How to measure $B \to D^{(*)} \ell \bar{\nu}_\ell$

$$w = v_B \cdot v_{D^{(*)}} = \frac{m_B^2 + m_{D^{(*)}}^2 - q^2}{2 m_B m_{D^{(*)}}}$$



How to measure $B \to D^{(*)}\ell\bar{\nu}_\ell$

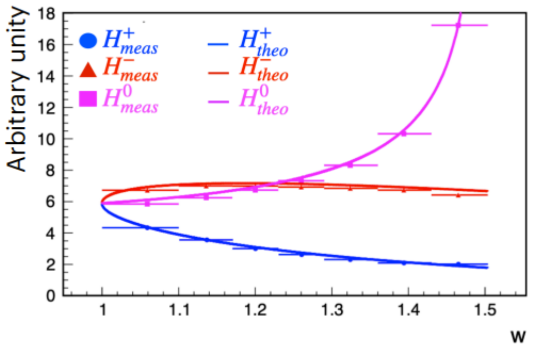
Marginal distributions
The "classic" way to determine form factors from data

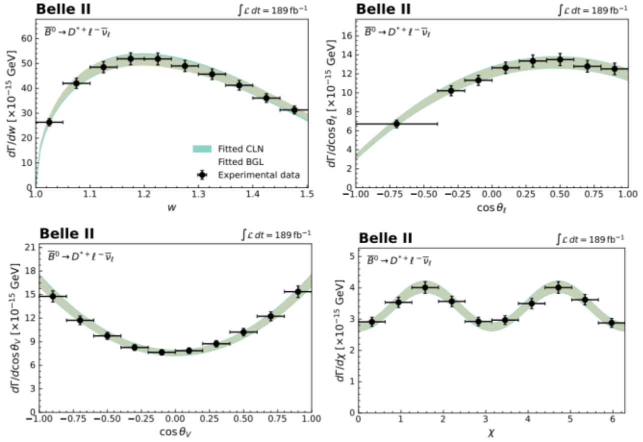


N-dimensional distributions?

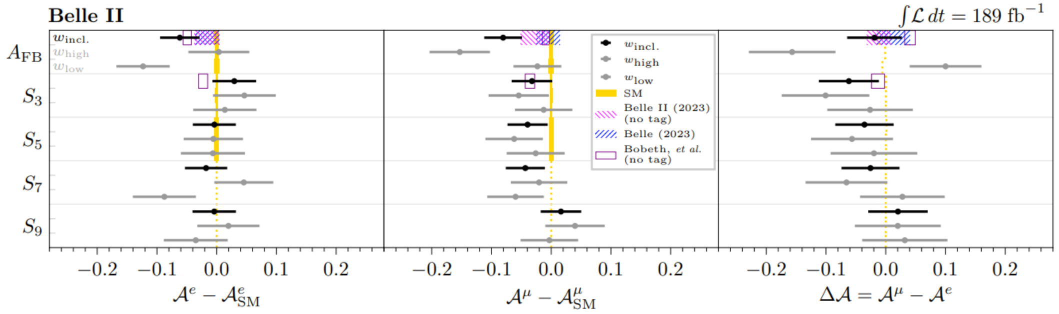
?

Helicity Amplitudes

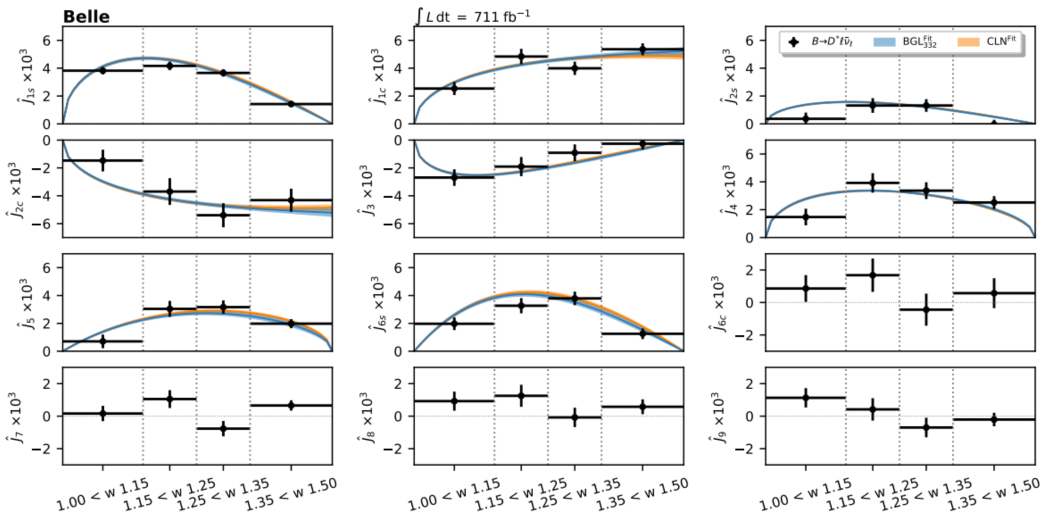




LFU Observables
Differential information on angular asymmetries



Angular Coefficients



Angular Coefficients

$$\frac{\mathrm{d}\Gamma(B\to D^*\ell\nu_\ell)}{\mathrm{d}w\,\mathrm{d}\cos\theta_\ell\,\mathrm{d}\cos\theta_\mathrm{V}\,\mathrm{d}\chi} = \frac{6 m_\mathrm{B} m_{\mathrm{D}^*}^2}{8(4\pi)^4}\sqrt{w^2-1}(1-2wr+r^2)G_\mathrm{F}^2\eta_\mathrm{EW}^2|V_\mathrm{cb}|^2$$
$$\times\Big((1-\cos\theta_\ell)^2\sin^2\theta_\mathrm{V}H_+^2 + (1+\cos\theta_\ell)^2\sin^2\theta_\mathrm{V}H_-^2$$
$$+4\sin^2\theta_\ell\cos^2\theta_\mathrm{V}H_0^2 - 2\sin^2\theta_\ell\sin^2\theta_\mathrm{V}\cos 2\chi H_+H_-$$
$$-4\sin\theta_\ell(1-\cos\theta_\ell)\sin\theta_\mathrm{V}\cos\theta_\mathrm{V}\cos\chi H_+H_0$$
$$+4\sin\theta_\ell(1+\cos\theta_\ell)\sin\theta_\mathrm{V}\cos\theta_\mathrm{V}\cos\chi H_-H_0\Big),$$

$$\frac{\mathrm{d}\Gamma(B\to D^*\ell\nu_\ell)}{\mathrm{d}w\,\mathrm{d}\cos\theta_\ell\,\mathrm{d}\cos\theta_\mathrm{V}\,\mathrm{d}\chi} = \frac{2G_\mathrm{F}^2\eta_\mathrm{EW}^2|V_\mathrm{cb}|^2 m_B^4 m_{\mathrm{D}^*}}{2\pi^4}\times\Big(J_{1s}\sin^2\theta_\mathrm{V} + J_{1c}\cos^2\theta_\mathrm{V}$$
$$+(J_{2s}\sin^2\theta_\mathrm{V} + J_{2c}\cos^2\theta_\mathrm{V})\cos 2\theta_\ell + J_3\sin^2\theta_\mathrm{V}\sin^2\theta_\ell\cos 2\chi$$
$$+J_4\sin 2\theta_\mathrm{V}\sin 2\theta_\ell\cos\chi + J_5\sin 2\theta_\mathrm{V}\sin\theta_\ell\cos\chi + (J_{6s}\sin^2\theta_\mathrm{V} + J_{6c}\cos^2\theta_\mathrm{V})\cos\theta_\ell$$
$$+J_7\sin 2\theta_\mathrm{V}\sin\theta_\ell\sin\chi + J_8\sin 2\theta_\mathrm{V}\sin 2\theta_\ell\sin\chi + J_9\sin^2\theta_\mathrm{V}\sin^2\theta_\ell\sin 2\chi\Big).$$

Helicity amplitudes are linear combinations of angular coefficients and vice versa, but …

… the angular coefficients provide additional differential information on angular asymmetries sensitive to new physics wrt. to marginal distribution!

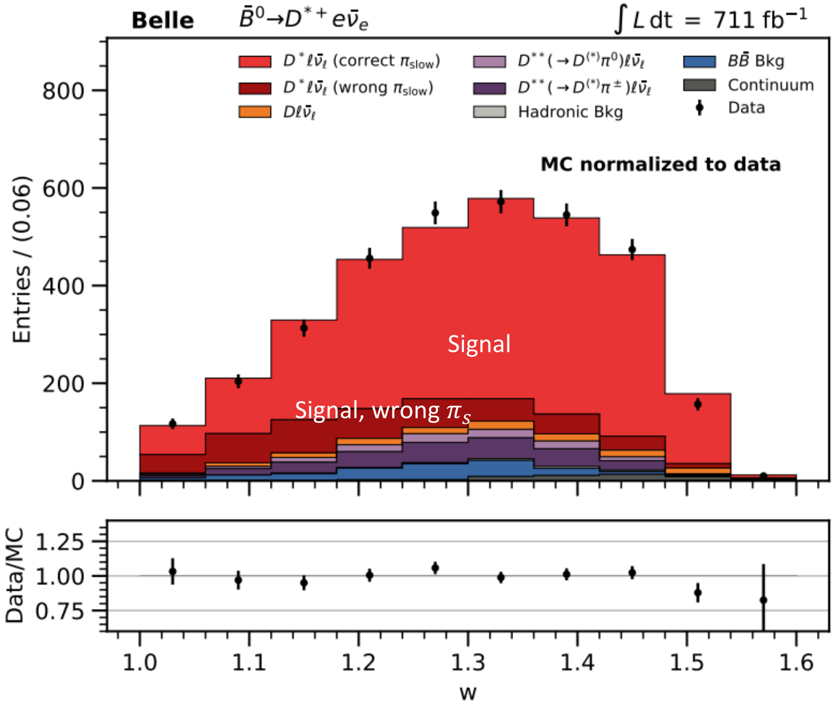
Similar idea:
Measurement of the helicity amplitudes

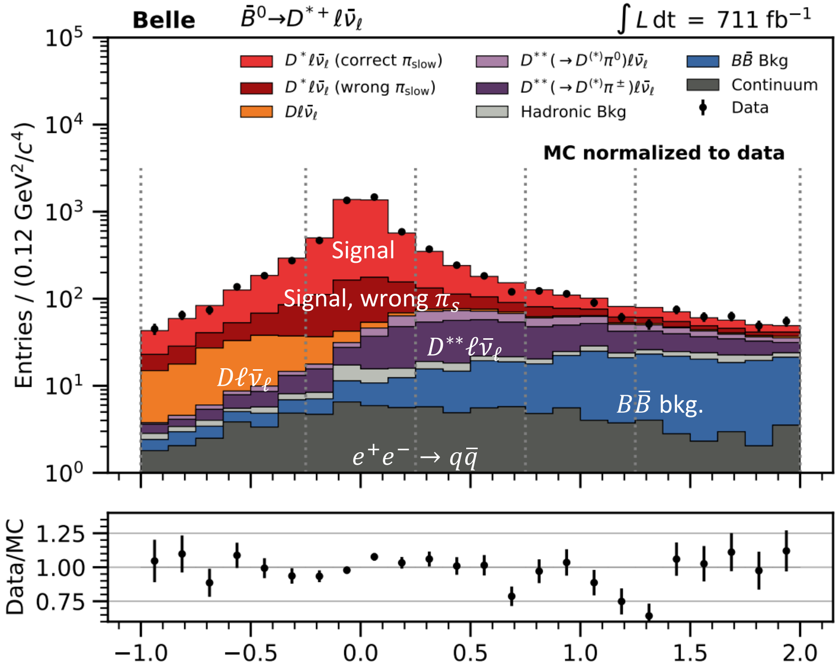
$$\frac{d^2\Gamma}{dw\,d\cos\theta_\ell} = \frac{1}{2}\Gamma_0(w)|V_{cb}|^2\Big\{\underbrace{[H_+^2(w)+H_-^2(w)+2H_0^2(w)]}_{a(w)} + 2\underbrace{[H_-^2(w)-H_+^2(w)]}_{b(w)}\cos\theta_\ell + \underbrace{[H_+^2(w)+H_-^2(w)-2H_0^2(w)]}_{c(w)}\cos^2\theta_\ell\Big\}.$$

Differential Distributions of $B \to D^* \ell \bar{\nu}_\ell$

$\bar{B}^0 \to D^{*+}(\to D^0\pi_s^+, D^+\pi_s^0)\ell\bar{\nu}_\ell$

$B^- \to D^{*0}(\to D^0\pi_s^0)\ell\bar{\nu}_\ell$





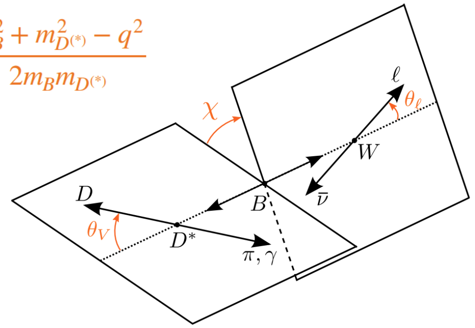
Background subtraction in independent variable to reduce model dependency.

Extraction Method: Missing Mass Squared

$$0 = m_\nu^2 = \mathrm{M}_{\mathrm{miss}}^2 = (p_{e^+e^-} - \mathrm{p_B} - \mathrm{p_{D^*}} - \mathrm{p}_\ell)^2$$

+ Unfolding and Acceptance Correction

$$w = v_B \cdot v_{D^{(*)}} = \frac{m_B^2 + m_{D^{(*)}}^2 - q^2}{2 m_B m_{D^{(*)}}}$$



Angular Coefficients of $B \to D^* \ell \bar{\nu}_\ell$

Instead of binning in $w, \cos\theta_\ell, \cos\theta_V, \chi$, we now bin the data to determine the angular coefficients in bins of w and:

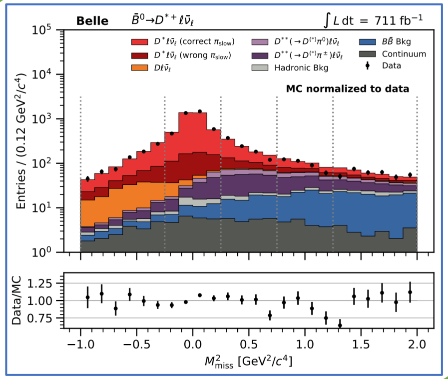
Phys.Rev.D 90 (2014) 9, 094003

$$\bar{J}_i = \frac{1}{N_i} \sum_{j=1}^{8} \sum_{k,l=1}^{4} \eta_{i,j}^{\chi} \eta_{i,k}^{\theta_\ell} \eta_{i,l}^{\theta_V} \left[\chi^{(j)} \otimes \chi^{(k)} \otimes \chi^{(l)} \right]$$

$J_i = J_i(w)$

Normalization Weights Unfolded Yields

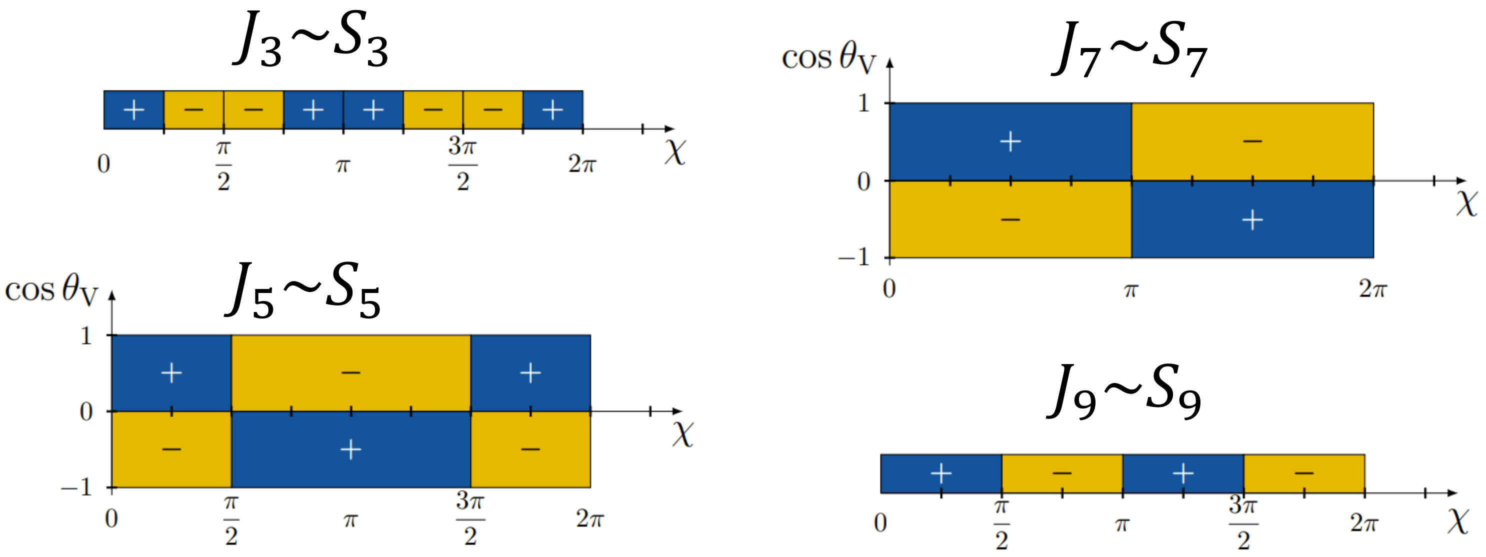
J_i	η_i^{χ}	$\eta_i^{\theta_\ell}$	$\eta_i^{\theta_V}$	normalization N_i
J_{1s}	$\{+\}$	$\{+,a,a,+\}$	$\{-,c,c,-\}$	$2\pi(1)2$
J_{1c}	$\{+\}$	$\{+,a,a,+\}$	$\{+,d,d,+\}$	$2\pi(1)(2/5)$
J_{2s}	$\{+\}$	$\{-,b,b,-\}$	$\{-,c,c,-\}$	$2\pi(-2/3)2$
J_{2c}	$\{+\}$	$\{-,b,b,-\}$	$\{+,d,d,+\}$	$2\pi(-2/3)(2/5)$
J_3	$\{+,-,-,+,+,-,-,+\}$	$\{+\}$	$\{+\}$	$4(4/3)^2$
J_4	$\{+,+,-,-,-,-,+,+\}$	$\{+,+,-,-\}$	$\{+,+,-,-\}$	$4(4/3)^2$
J_5	$\{+,+,-,-,-,-,+,+\}$	$\{+\}$	$\{+,+,-,-\}$	$4(\pi/2)(4/3)$
J_{6s}	$\{+\}$	$\{+,+,-,-\}$	$\{-,c,c,-\}$	$2\pi(1)2$
J_{6c}	$\{+\}$	$\{+,+,-,-\}$	$\{+,d,d,+\}$	$2\pi(1)(2/5)$
J_7	$\{+,+,+,+,-,-,-,-\}$	$\{+\}$	$\{+,+,-,-\}$	$4(\pi/2)(4/3)$
J_8	$\{+,+,+,+,-,-,-,-\}$	$\{+,+,-,-\}$	$\{+,+,-,-\}$	$4(4/3)^2$
J_9	$\{+,+,-,-,+,+,-,-\}$	$\{+\}$	$\{+\}$	$4(4/3)^2$



Same signal extraction through model independent variable

$$\frac{d\Gamma(B \to D^* \ell \nu_\ell)}{dw\, d\cos\theta_\ell\, d\cos\theta_V\, d\chi} = \frac{2 G_F^2 \eta_{EW}^2 |V_{cb}|^2 m_B^4 m_{D^*}}{2\pi^4} \times \Big(J_{1s} \sin^2\theta_V + J_{1c} \cos^2\theta_V$$
$$+ (J_{2s} \sin^2\theta_V + J_{2c} \cos^2\theta_V) \cos 2\theta_\ell + J_3 \sin^2\theta_V \sin^2\theta_\ell \cos 2\chi$$
$$+ J_4 \sin 2\theta_V \sin 2\theta_\ell \cos\chi + J_5 \sin 2\theta_V \sin\theta_\ell \cos\chi + (J_{6s} \sin^2\theta_V + J_{6c} \cos^2\theta_V) \cos\theta_\ell$$
$$+ J_7 \sin 2\theta_V \sin\theta_\ell \sin\chi + J_8 \sin 2\theta_V \sin 2\theta_\ell \sin\chi + J_9 \sin^2\theta_V \sin^2\theta_\ell \sin 2\chi \Big) .$$

Angular asymmetries in $B \to D^* \ell \bar{\nu}_\ell$



$$\mathcal{A}_x(w) = \frac{N_x^+(w) - N_x^-(w)}{N_x^+(w) + N_x^-(w)}.$$

The angular coefficients provide

- access to the form factors to determine $|V_{cb}|$
- additional differential information on angular asymmetries sensitive to new physics!
- Helicity amplitudes and angular coefficients are interchangeable (linear combinations)

Claim 2: Measuring angular coefficients provides the maximum amount of information in a single measurement.

Unfolding

Unfolding – A quick intro

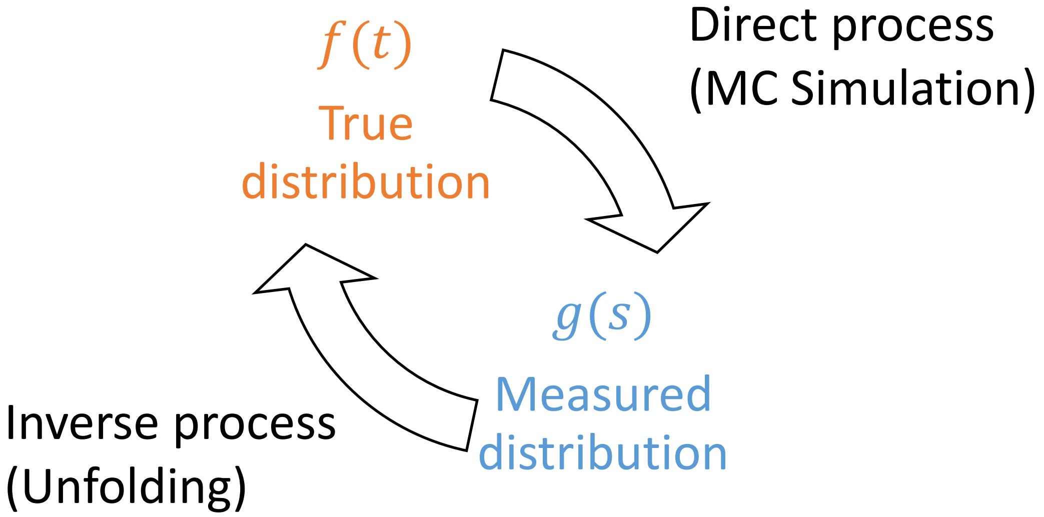
Direct and Inverse processes

Fredholm integral equation of the first kind

$$\int_{\Omega} K(s,t)\, f(t)\, dt = g(s)$$

- Kernel function $K(s,t)$ describes the physical measurement process
- Implicitly known from MC Simulation assuming a model $f(t)^{\text{model}}$
- The inverse problem is an ill-posed problem

Access to true distribution required to test / fit our theories!



Unfolding – A quick intro

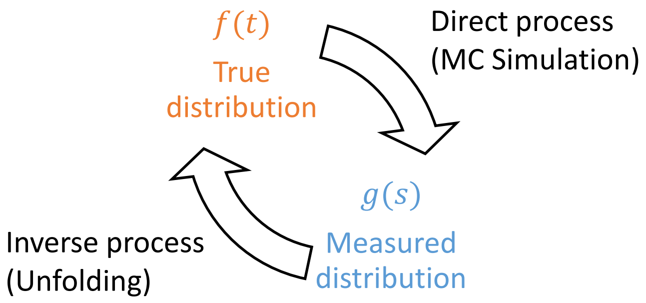
Discretization and linear solution

$$A\,x = y \Rightarrow x = A^{-1}\,y$$

True distribution $f(t) \rightarrow x$

Measured distribution $g(s) \rightarrow y$

Kernel $K(s,t) \rightarrow A$



- Matrix inversion provides unbiased results (no assumptions)
- Regularization sometimes necessary, because of e.g., statistical fluctuations

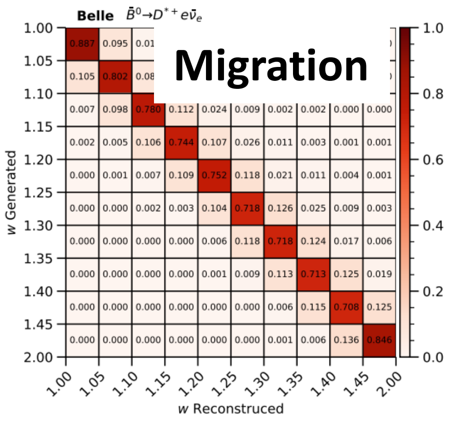
$$y = y_{\text{signal}} + \varepsilon_{\text{stat}} + b$$

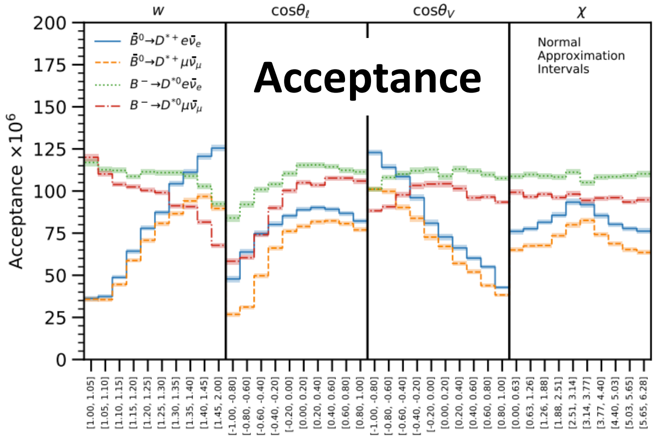
- Various methods exists, e.g., Tikhonov regularization, to deal with this if necessary
  - Choice of regularization is subjective
  - Regularization introduces constraints and for that reason reduces the uncertainties

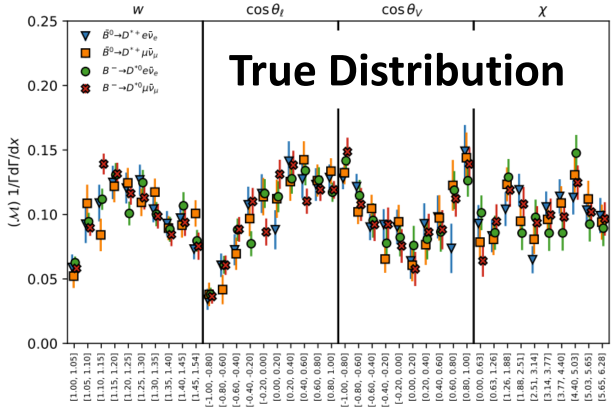
In praxis

Unfolding

Background Subtracted Yields

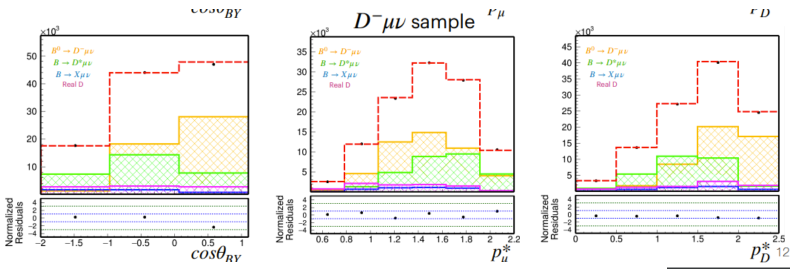




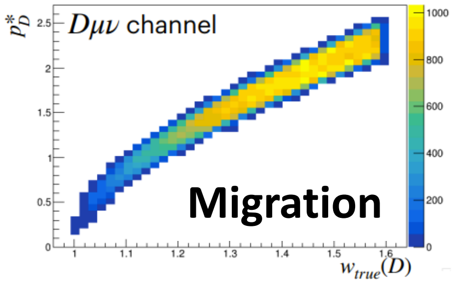


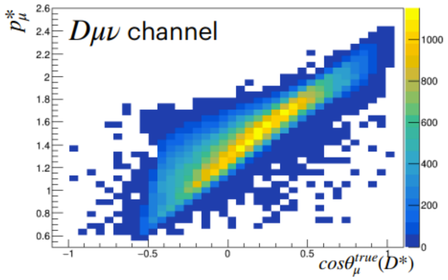
Theory

Forward Folding



Background Subtracted Yields





Acceptance

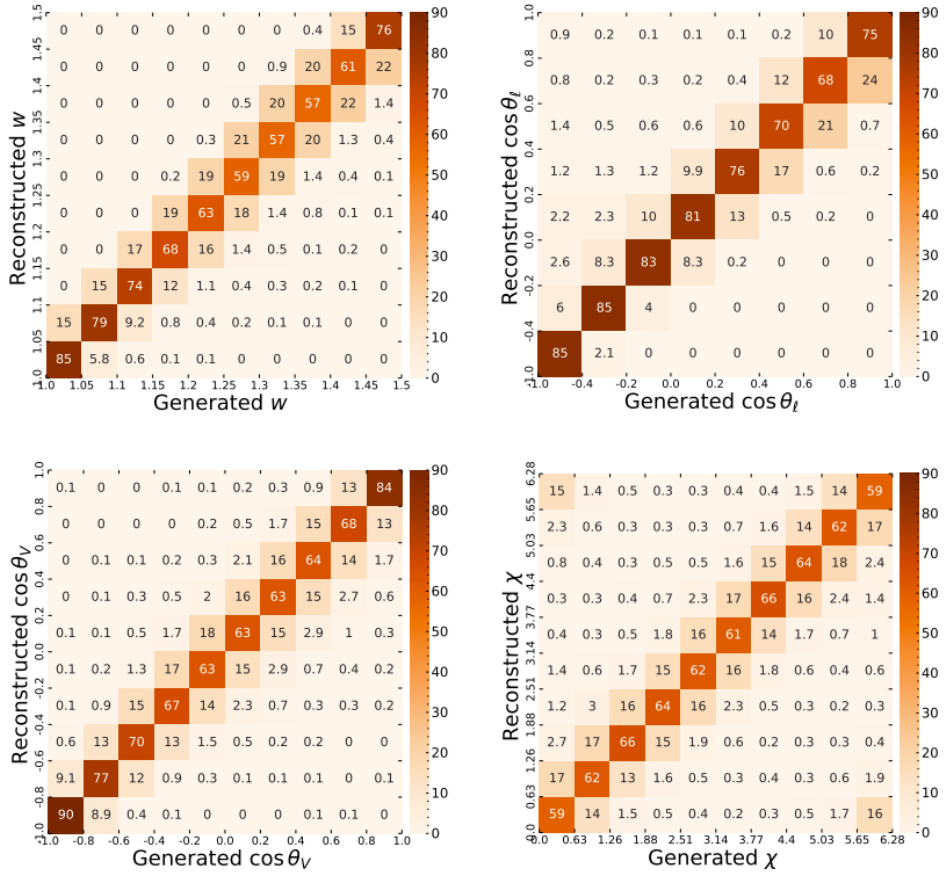
True Distribution

Theory

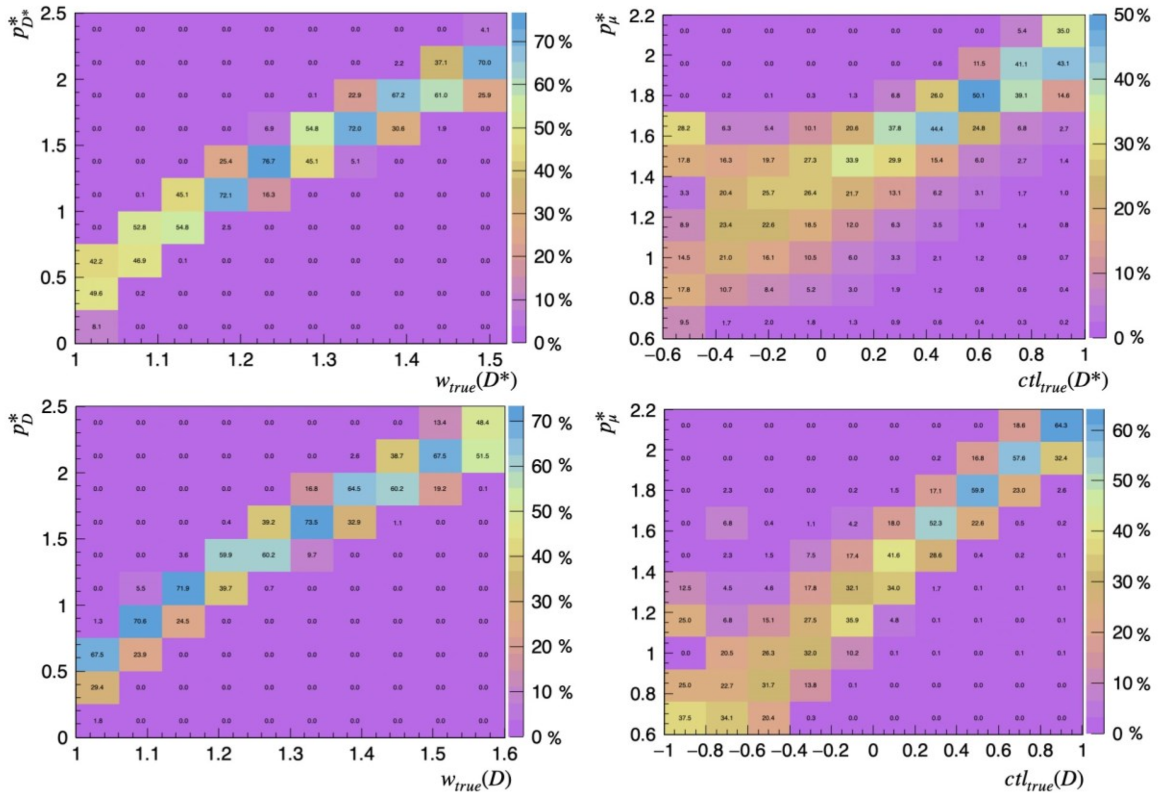
$$p_i = \frac{\Gamma_{new}}{\Gamma_{ref}} \frac{d\Gamma_{new}/d\Omega}{d\Gamma_{ref}/d\Omega}$$

Choice of unfolding variables

Dominated by Detector Resolution
(reconstruction of w is complicated)



Dominated by Kinematics
(momentum resolution is good)



Unfolding – Quo vadis?

Preserving our results

- To preserve the data, we need to provide access to the true distributions
- Experimentally we have the choice between forward folding and unfolding

Claim 3: Traditional binned methods for solving the Fredholm Equation will limit our future sensitivity.

Can we overcome the inherent limitations?

Methods are inherently binned

- Binning fixed and can not be changed
- Performance sensitive to binning

Response matrix depends on auxiliary features

- Detector-level quantity might not capture full detector effect

Limited number of observables

- Curse of dimensionality

Probably Yes!

4D (2D) analysis of $B \rightarrow D^{(*)}\ell\bar{\nu}_\ell$

4D (2D) analysis of $B \to D^{(*)}\ell\bar{\nu}_\ell$

Deep Learning to the rescue!

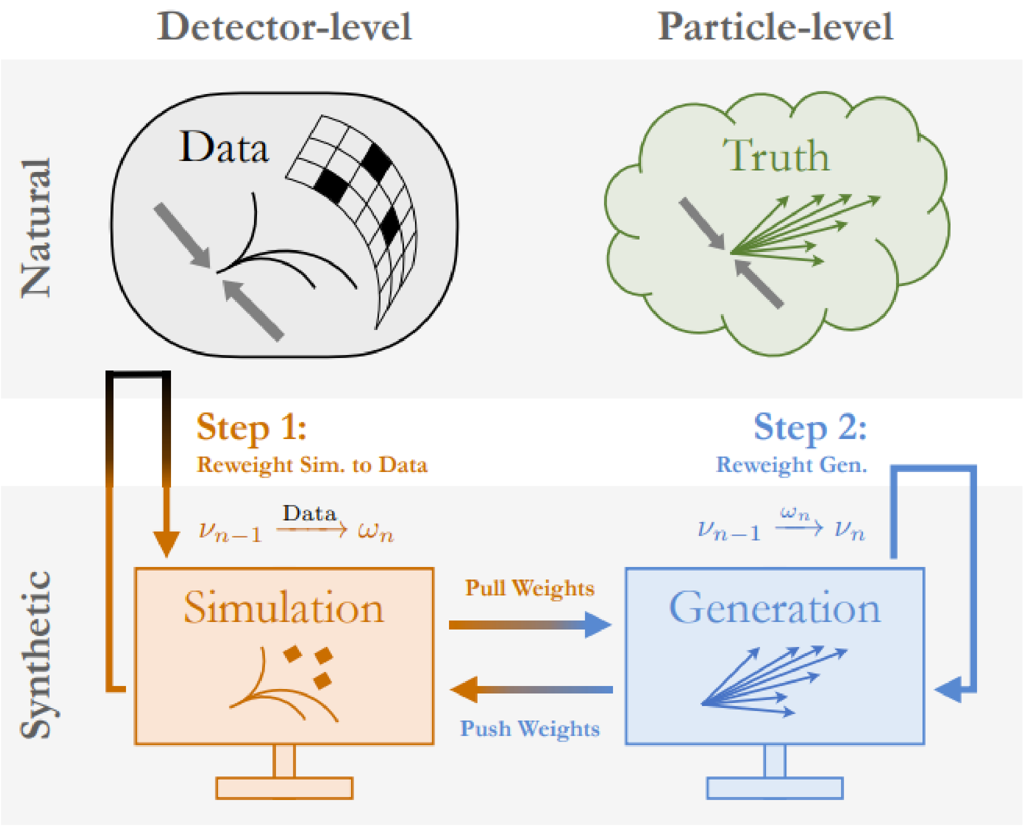
OmniFold: A Method to Simultaneously Unfold All Observables

Anders Andreassen, Patrick T. Komiske, Eric M. Metodiev, Benjamin Nachman, Jesse Thaler

Collider data must be corrected for detector effects ("unfolded") to be compared with many theoretical calculations and measurements from other experiments. Unfolding is traditionally done for individual, binned observables without including all information relevant for characterizing the detector response. We introduce OmniFold, an unfolding method that iteratively reweights a simulated dataset, using machine learning to capitalize on all available information. Our approach is unbinned, works for arbitrarily high-dimensional data, and naturally incorporates information from the full phase space. We illustrate this technique on a realistic jet substructure example from the Large Hadron Collider and compare it to standard binned unfolding methods. This new paradigm enables the simultaneous measurement of all observables, including those not yet invented at the time of the analysis.

Key Idea: Train a neural network to distinguish between data and simulation → turns density estimation into classification

We [Belle (II)] already employ similar concept: Continuum Reweighting



Claim 4: Unbinned methods are the future.

Summary

Claim 1: Combined $B \to D^{(*)}\ell\bar{\nu}_\ell$ will provide better precision on $|V_{cb}|$ than individual results.

Claim 2: Measuring angular coefficients provides the maximum amount of information in a single measurement.

Claim 3: Traditional binned methods for solving the Fredholm Equation will limit our future sensitivity.

Claim 4: Unbinned methods are the future.

Backup

Where do we need to gain precision?

Hadronically tagged measurement with full
LQCD input from FNAL, HPQCD, JLCQCD:

$$|V_{cb}| = (41.0 \pm 0.3 \pm 0.4 \pm 0.5) \times 10^{-3}$$

Uncertainties from shape, normalization, LQCD