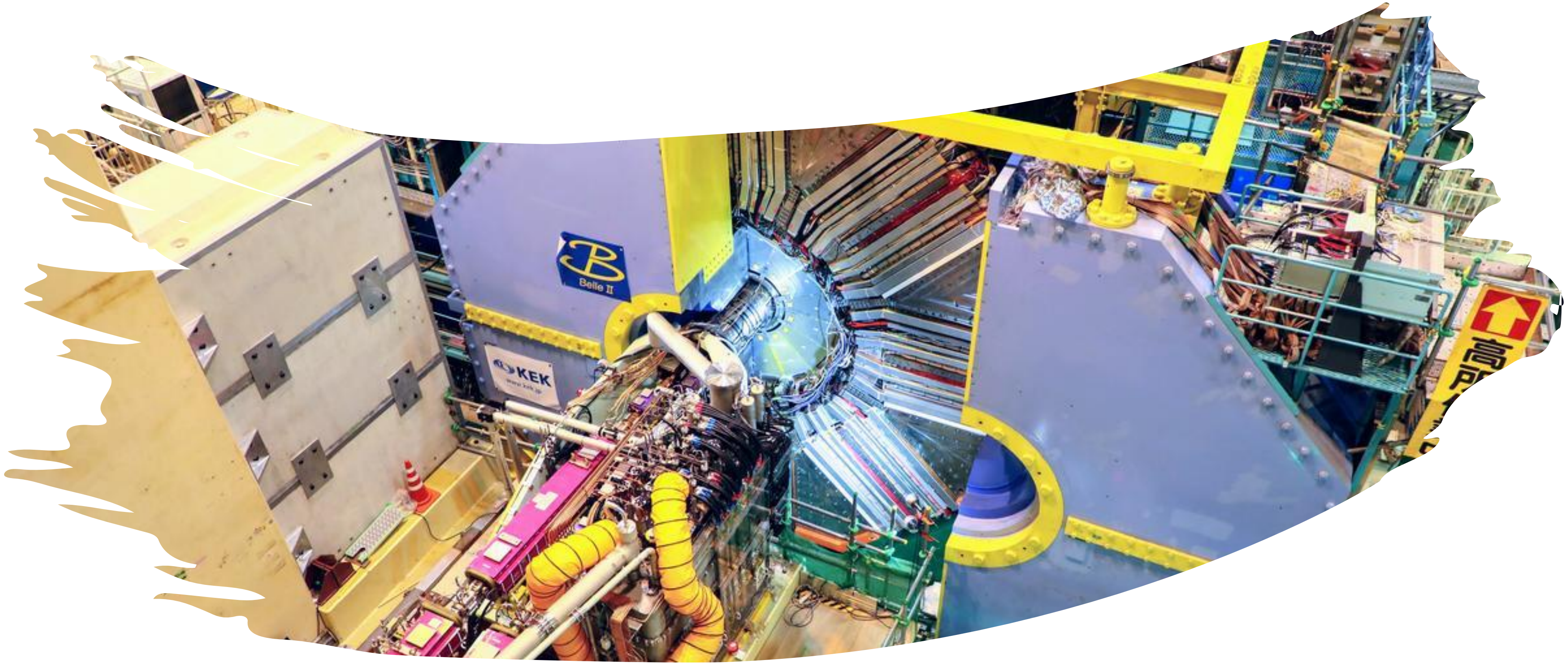



# A tour on semileptonic measurements at Belle (II)

Markus Prim – University of Bonn

# Today's Tour

Exclusive $B \to D^{(*,**)}\ell\nu_\ell$

Measurements with $\ell = e, \mu$

Inclusive $B \to X_c\ell\nu_\ell$

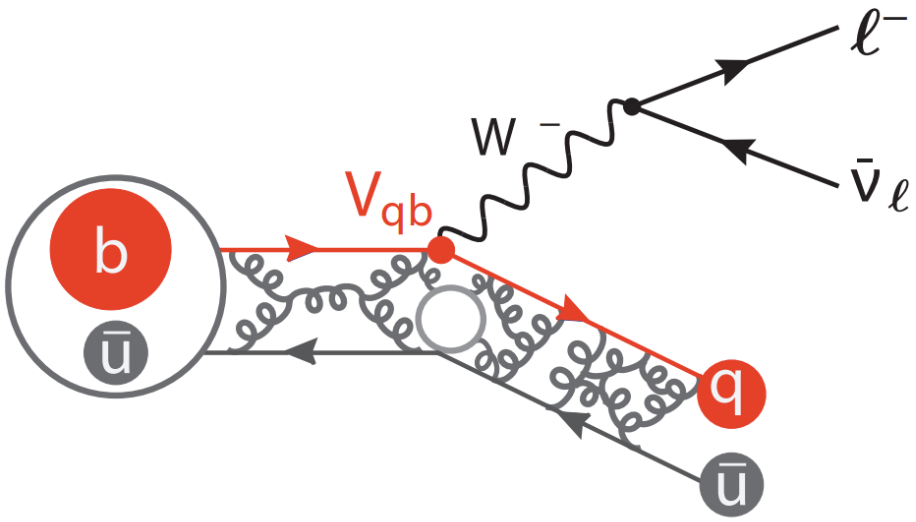
Measurements with $\ell = e, \mu$

$|V_{cb}|$ Extraction

Ratio Measurements (with $\tau$)

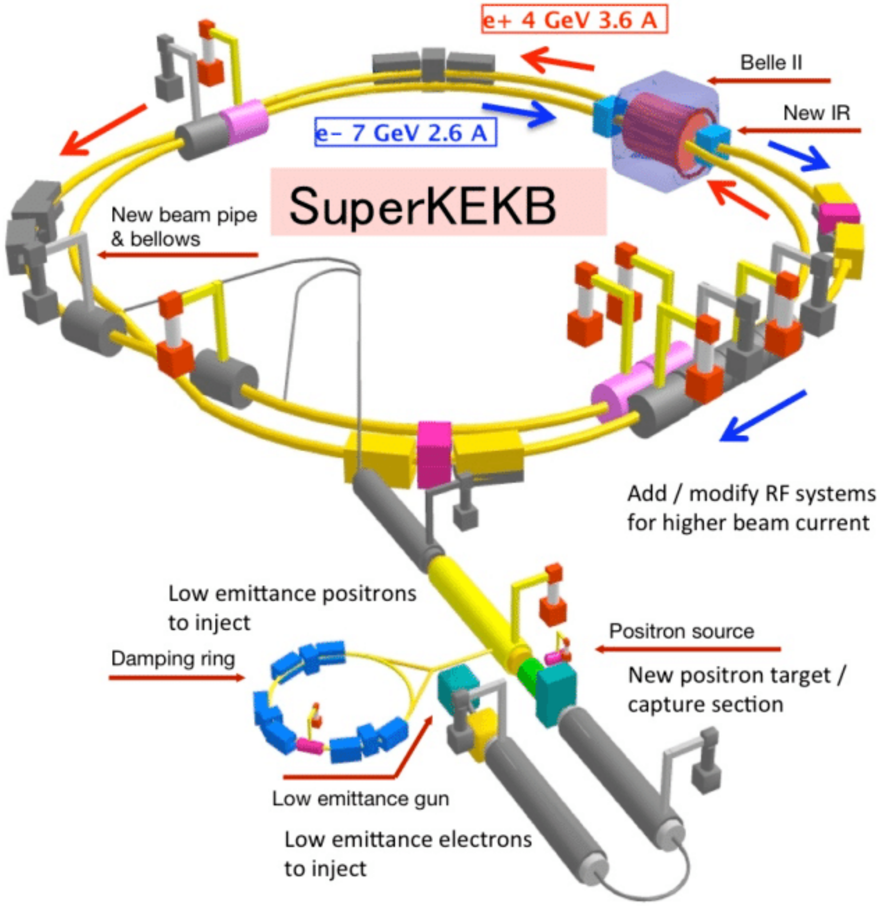
$$R = \frac{b \to q\tau\nu_\tau}{b \to q\ell\nu_\ell}$$

LFU Test



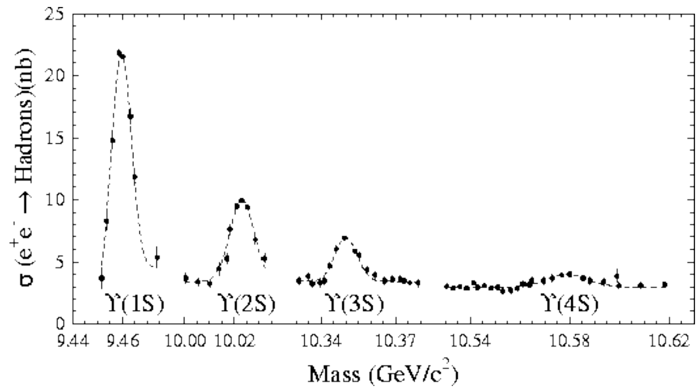

- The experimental techniques also apply to $b \to u\ell\nu_\ell$
  Additional challenge here is suppressing the abundant $b \to c\ell\nu_\ell$
- I present concepts, this means that some numbers or plots are possibly outdated

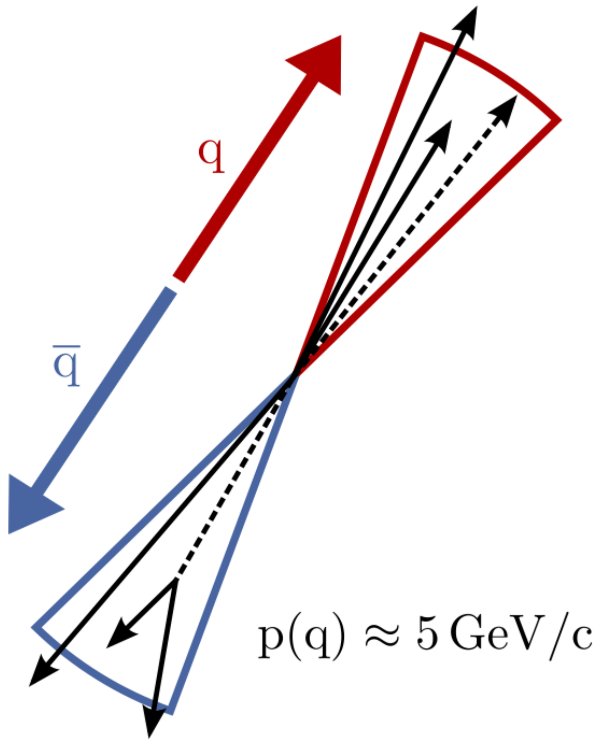
# SuperKEKB Accelerator

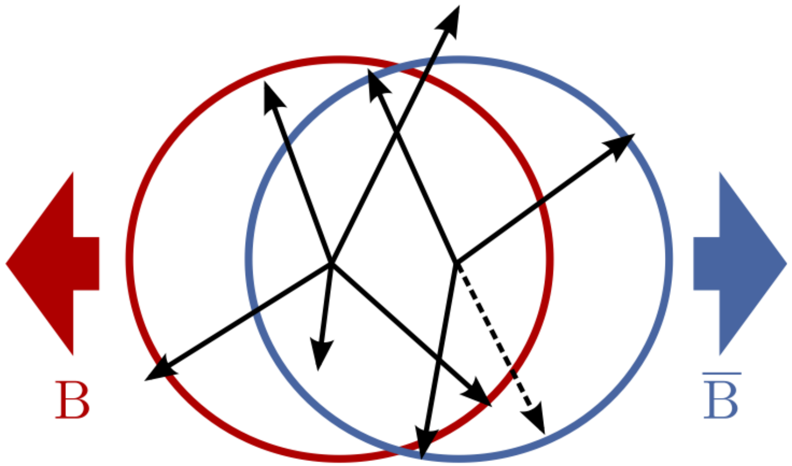



- Asymmetric energy $e^+e^-$ collider on the $\Upsilon(4S)$ resonance
- Clean environment, production of the $\Upsilon(4S) \to B\bar{B}$
  - no additional particles
  - no underlying event
- $B$ mesons produced (almost) at rest
- Small cross-section $\sim$1.1nb

# Event Topology at Belle II



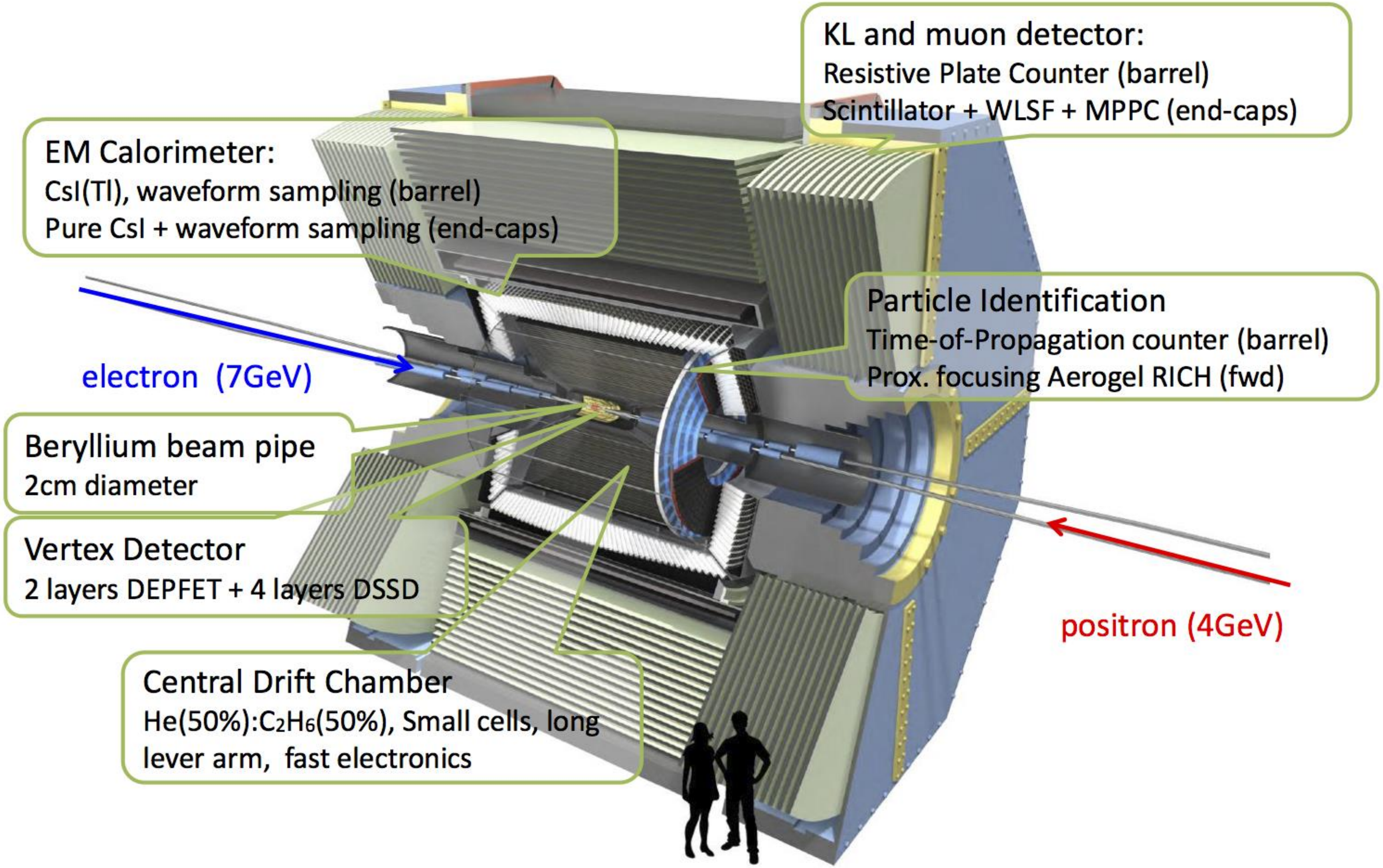




$e^+e^- \to q\bar{q} \quad (q \in \{u, d, s, c\})$

$e^+e^- \to \Upsilon(4S) \to B\bar{B}$

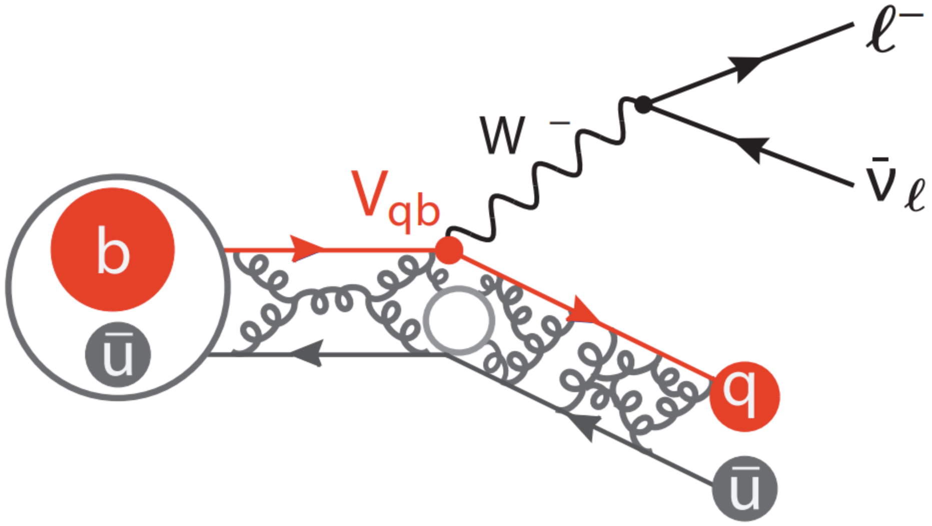
Continuum suppression utilizes the difference in event topology:
Fox-Wolfram moments, Cleo cones, thrust variables, etc

# Belle II Detector



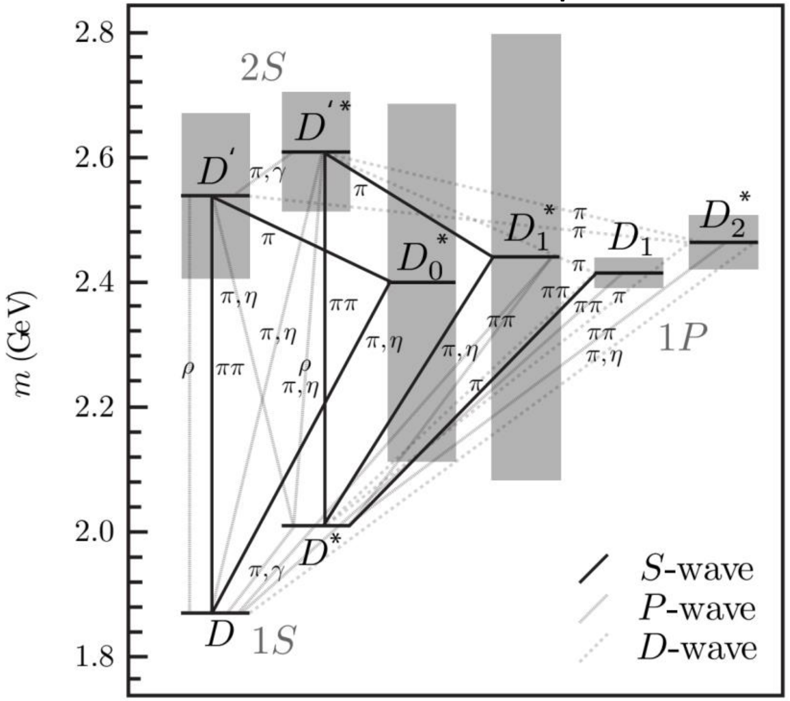
# Exclusive Measurements

# Exclusive Measurements




- Exclusive measurements focus on explicit, resonant, final states
- For $b \to c\ell\nu$ transitions, these are
  - 2 L=0 states $D, D^*$
    These saturate ~75% of the inclusive $B \to X_c\ell\nu$ rate
    and are the principal route to extract $|V_{cb}|$
  - 4 L=1 states: $D_0, D_1', D_1, D_2$ (or $D_0^*, D_1^*, D_1, D_2^*$ , simply $D^{**}$)
    These saturate ~15% of the inclusive $B \to X_c\ell\nu$ rate
    and mostly a source of background
- What makes up the last ~10% of the inclusive branching fraction?

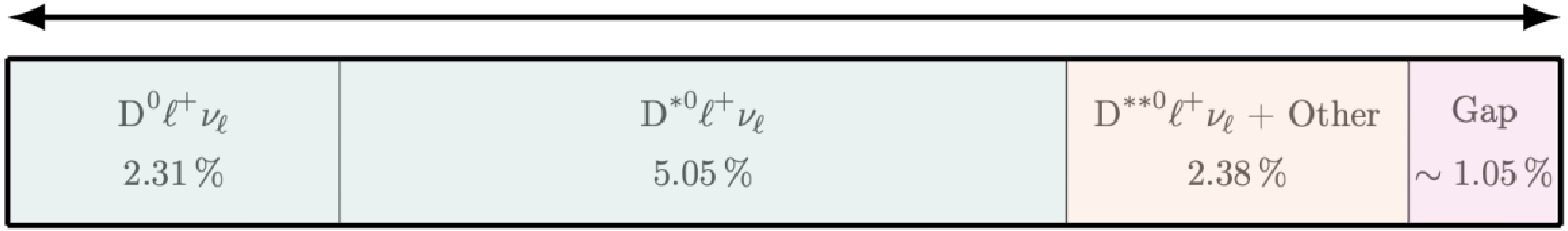
Charmed Meson Spectrum



# Non-Resonant Decays and the „Gap“

$$\mathcal{B}(\mathrm{B}^+ \to X_c^0 \ell^+ \nu_\ell) \approx 10.79\,\%$$

F. Metzner



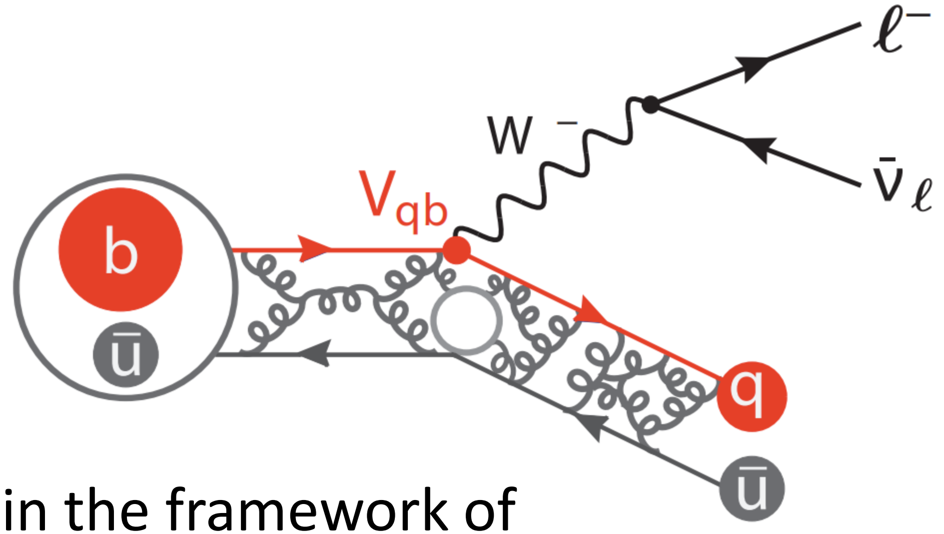

| Decay | $\mathcal{B}(B^+)$ | $\mathcal{B}(B^0)$ |
|---|---|---|
| $B \to D\ell^+\nu_\ell$ | $(2.4098 \pm 0.0709) \cdot 10^{-2}$ | $(2.2396 \pm 0.0664) \cdot 10^{-2}$ |
| $B \to D^*\ell^+\nu_\ell$ | $(5.5023 \pm 0.1146) \cdot 10^{-2}$ | $(5.1137 \pm 0.1082) \cdot 10^{-2}$ |
| $B \to D_1\ell^+\nu_\ell$ | $(6.6322 \pm 1.0894) \cdot 10^{-3}$ | $(6.1638 \pm 1.0127) \cdot 10^{-3}$ |
| $B \to D_0^*\ell^+\nu_\ell$ | $(4.2000 \pm 0.7500) \cdot 10^{-3}$ | $(3.9033 \pm 0.6972) \cdot 10^{-3}$ |
| $B \to D_1'\ell^+\nu_\ell$ | $(4.2000 \pm 0.9000) \cdot 10^{-3}$ | $(3.9033 \pm 0.8366) \cdot 10^{-3}$ |
| $B \to D_2^*\ell^+\nu_\ell$ | $(2.9337 \pm 0.3248) \cdot 10^{-3}$ | $(2.7265 \pm 0.3020) \cdot 10^{-3}$ |
| $B \to D\pi\pi\ell^+\nu_\ell$ | $(0.6228 \pm 0.8857) \cdot 10^{-3}$ | $(0.5788 \pm 0.8232) \cdot 10^{-3}$ |
| $B \to D^*\pi\pi\ell^+\nu_\ell$ | $(2.1600 \pm 1.0247) \cdot 10^{-3}$ | $(2.0074 \pm 0.9523) \cdot 10^{-3}$ |
| $B \to D_sK\ell^+\nu_\ell$ | $(0.3000 \pm 0.1421) \cdot 10^{-3}$ | - |
| $B \to D_s^*K\ell^+\nu_\ell$ | $(0.2900 \pm 0.1942) \cdot 10^{-3}$ | - |
| ? | | |
| $B \to X_c\ell^+\nu_\ell$ | $(10.8 \pm 0.4) \cdot 10^{-2}$ | $(10.1 \pm 0.4) \cdot 10^{-2}$ |

Well known
Some tension when comparing isospin modes

Broad states based on measurements by BaBar, Belle, and DELPHI

Measurements by BaBar and Belle

# Exclusive Measurements



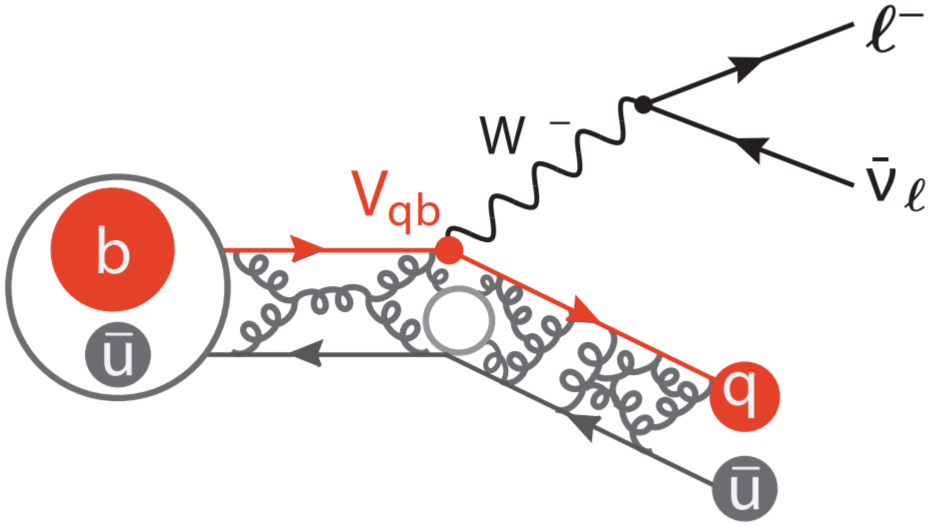

- Exclusive measurements focus on explicit, resonant, final states
- Hadronic matrix element can not be calculated within the framework of perturbation theory. It is parametrized by form factors, e.g. for $D^*$
- In the past:
  - Functional form of the form factors unknown, must be derived from data
  - Normalization of the form factors from Lattice QCD
- Since last/this year:
  - Beyond zero-recoil lattice predictions for the functional form of the form factors

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

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

Heavy quark symmetry basis

# Exclusive Measurements




- Access to more than form factors & $|V_{cb}|$
- Forward-backward asymmetries $(\Delta)A_{FB}$

$$A_{\rm FB} = \frac{\int_0^1 {\rm d}\cos_\ell\, {\rm d}\Gamma/{\rm d}\cos_\ell - \int_{-1}^0 {\rm d}\cos_\ell\, {\rm d}\Gamma/{\rm d}\cos_\ell}{\int_0^1 {\rm d}\cos_\ell\, {\rm d}\Gamma/{\rm d}\cos_\ell + \int_{-1}^0 {\rm d}\cos_\ell\, {\rm d}\Gamma/{\rm d}\cos_\ell}$$

- Longitudinal polarization fraction $(\Delta)F_L$

$$\frac{1}{\Gamma}\frac{{\rm d}\Gamma}{{\rm d}\cos\theta_V} = \frac{3}{2}\left(F_L\cos^2\theta_V + \frac{1-F_L}{2}\sin^2\theta_V\right)$$

- Lepton flavor universality ratio $R_{e\mu}$

$$R_{e\mu} = \frac{\mathcal{B}(B\to D^* e\bar{\nu}_e)}{\mathcal{B}(B\to D^* \mu\bar{\nu}_\mu)}$$

These are probes for new physics

# Exclusive Measurement Strategies

## Untagged Measurements

+ Very high efficiency

+ Absolute branching fraction straightforward

- Less experimental control, e.g., more backgrounds

- Signal B rest frame not *directly* accessible

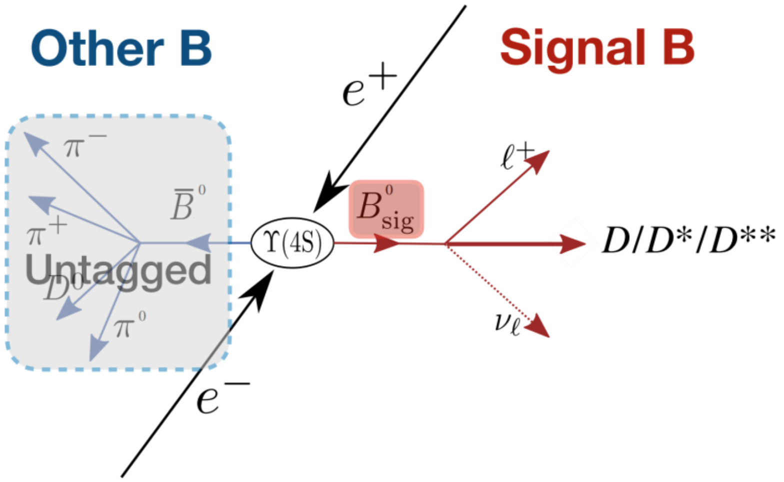
## Tagged Measurements

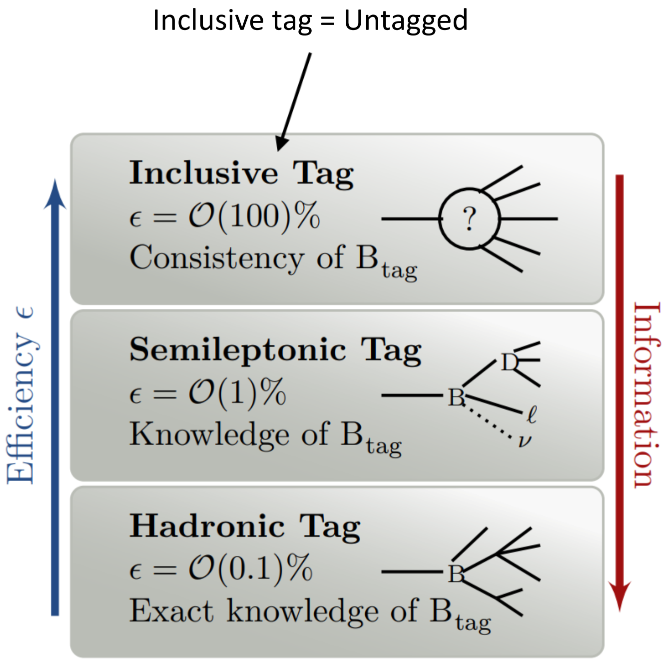
+ High degree of experimental control

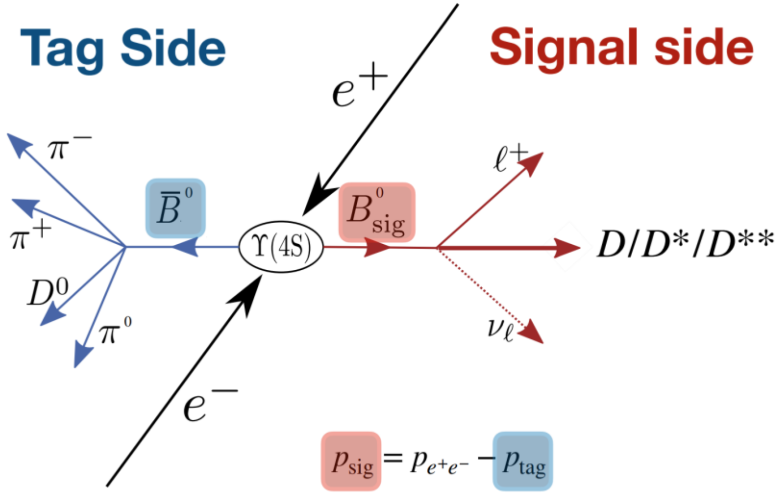
+ Hadronic tagging gives access to the signal B rest frame

- Understanding efficiencies is difficult

- Tagging efficiency reduces effective statistical power



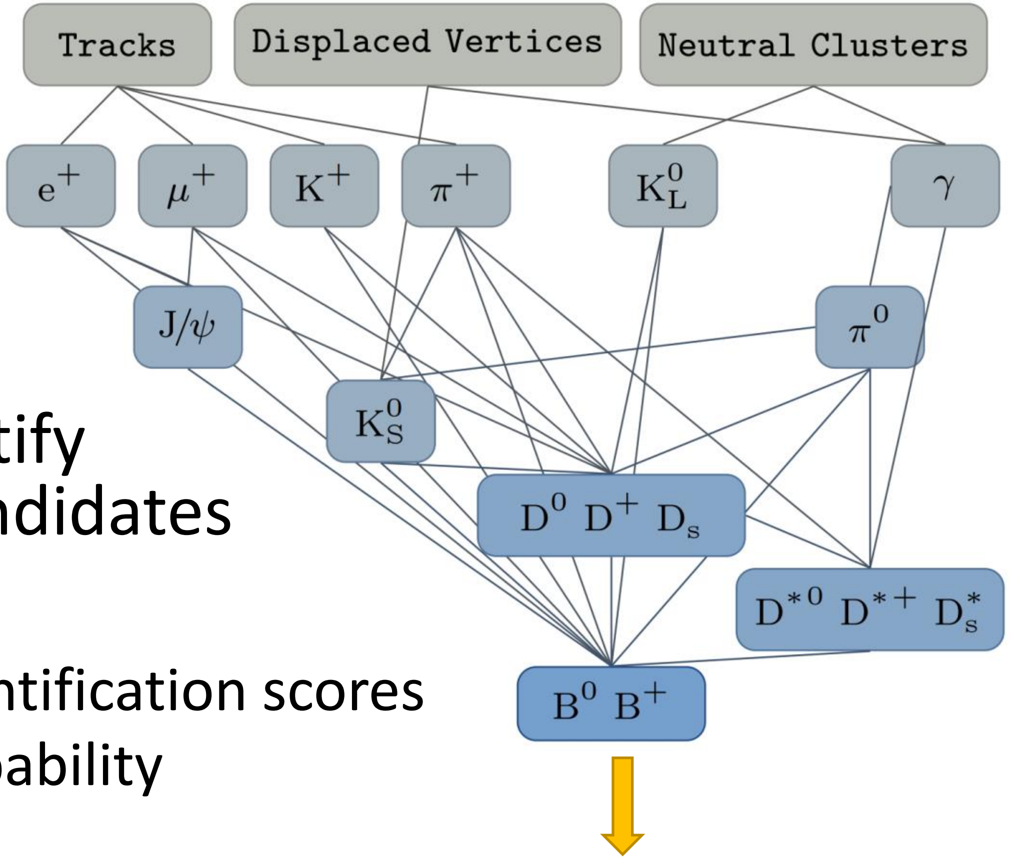



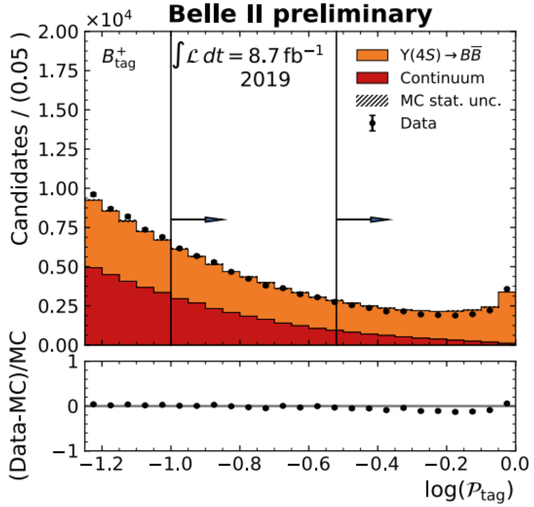



# Tagging at B-factories

# Full Event Interpretation

- Hierarchical bottom-up approach
- Classifiers (BDT or NN) are trained to identify correctly reconstructed (intermediate) candidates
- At each step:
  - Input variables: four-momenta & particle identification scores
  - Output: score that can be interpreted as probability
  - Mild selection on the output score
- Over 10’000 decay cascades are automatically reconstructed
- E.g., Hadronic tagging efficiency is ~0.3%




arXiv:2008.06096
arXiv:1807.08680

Belle‘s Full Reconstruction works conceptually the same

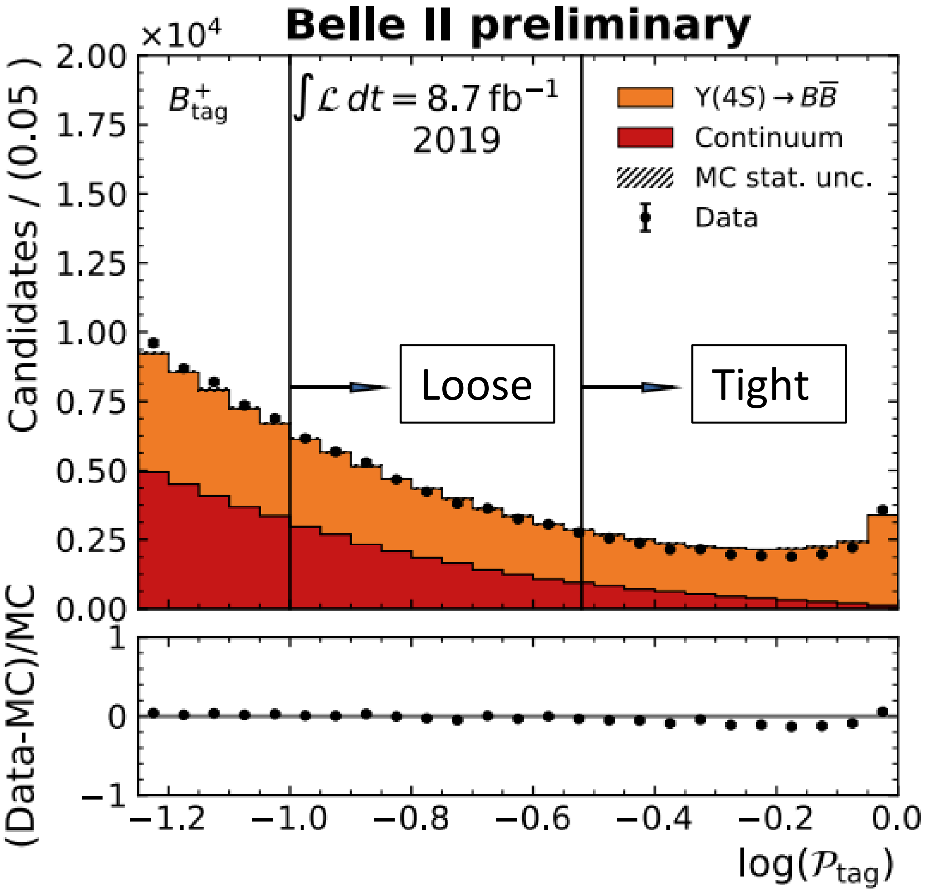
# Full Event Interpretation

beam-constrained mass

B-Meson mass

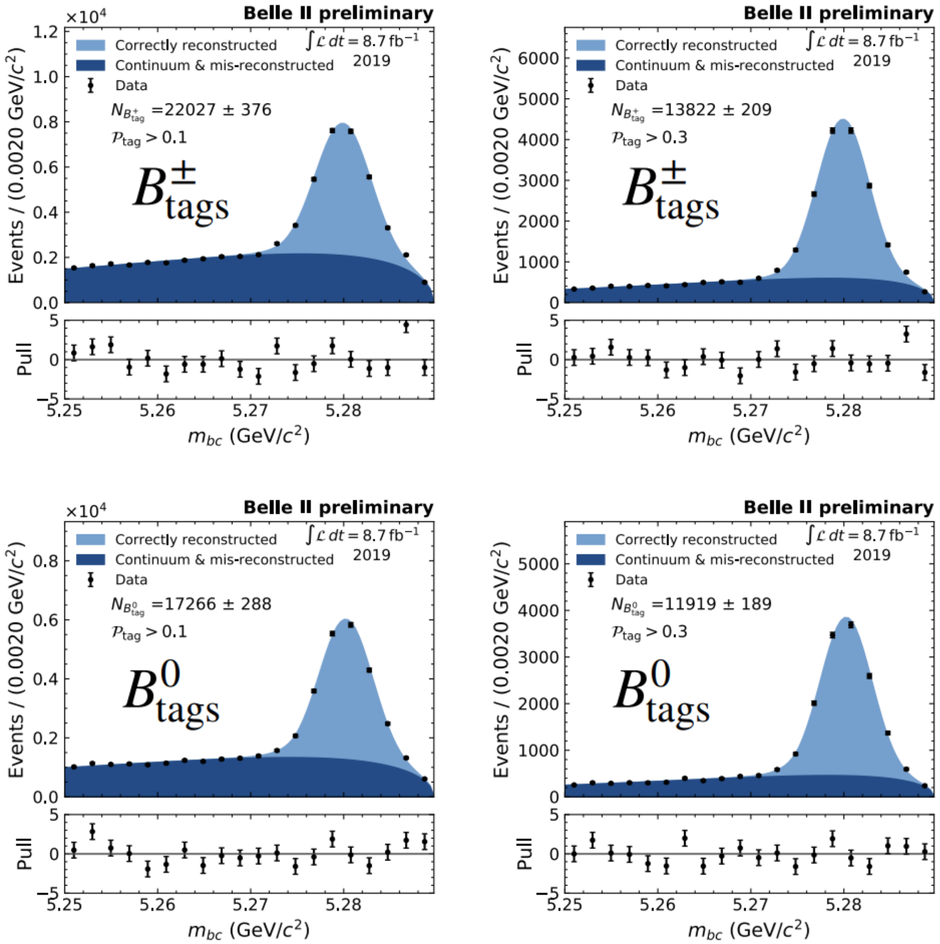
$$m_{bc} = \sqrt{\frac{E_{beam}^2}{4} - \left|\vec{p}_{B_{tag}}\right|^2} \approx m_B$$

$P_{tag}$ = Output Classifier: Measure/Probability of how well the B-Meson is reconstructed



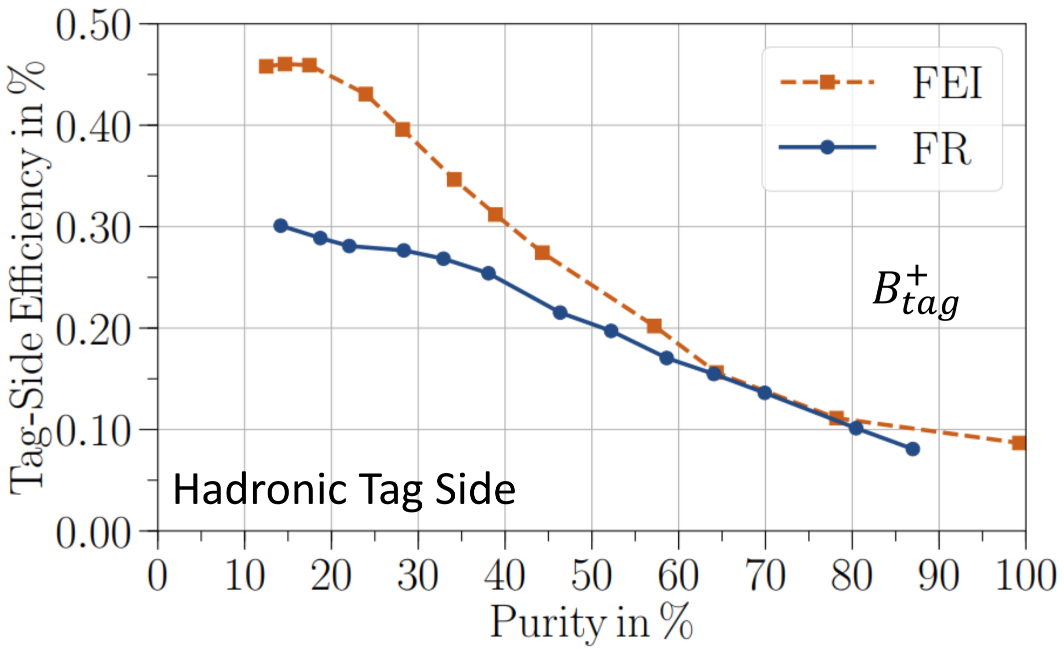

Loose

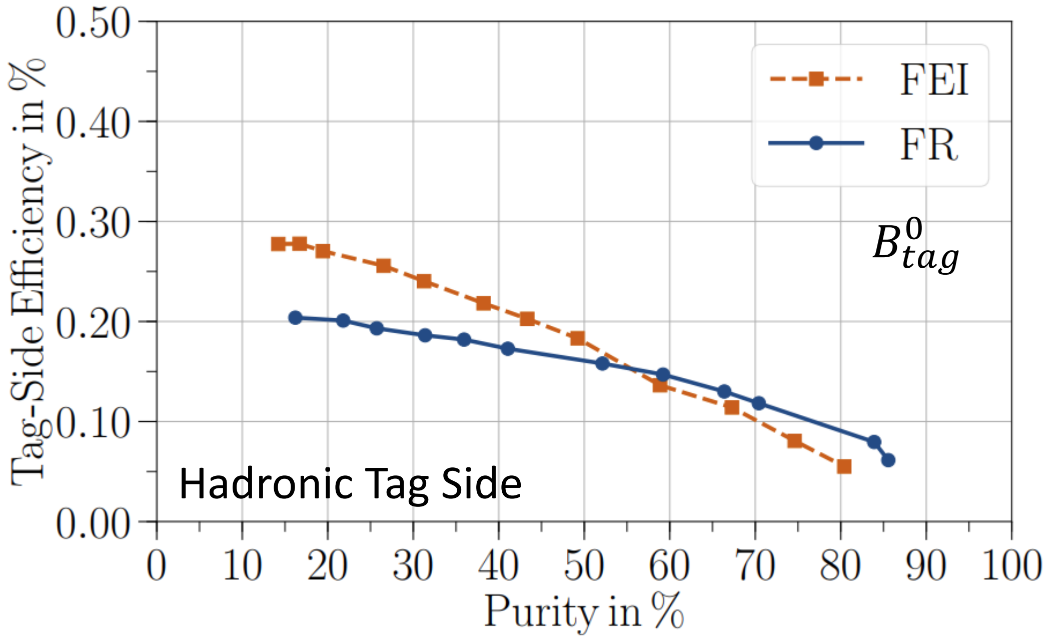
Tight



# Full Event Interpretation

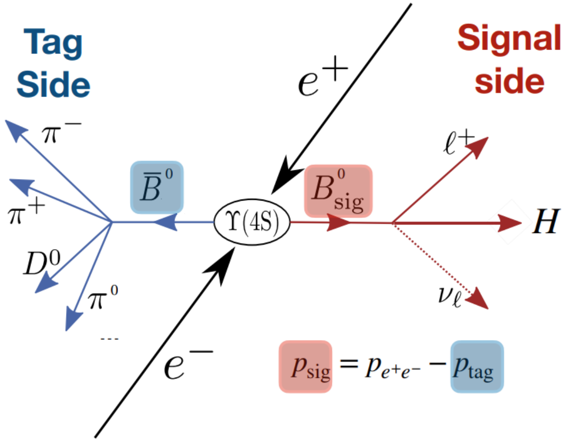
- Algorithm output is tunable based on the receiver operating characteristic
- Trade-off between efficiency and purity
- Calibration of the algorithm required depending on the working point



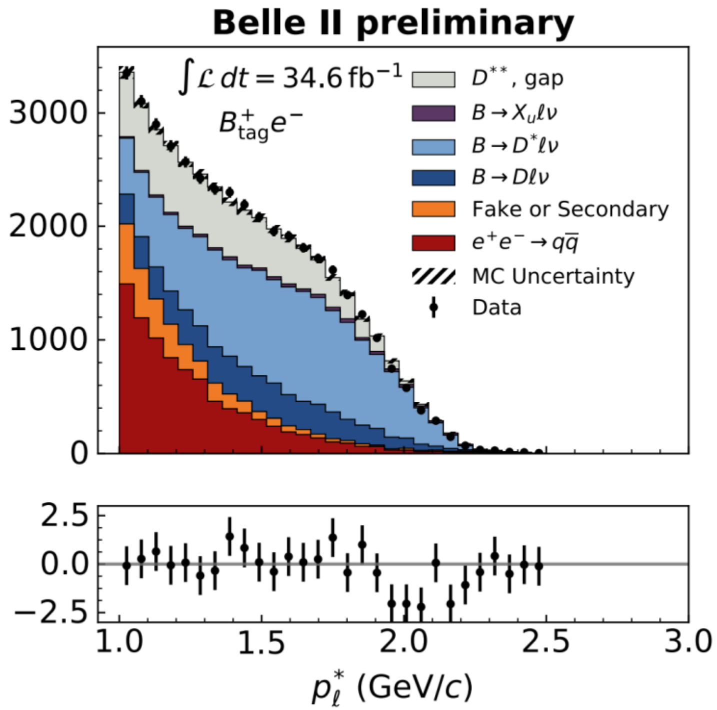



# Full Event Interpretation - Calibration

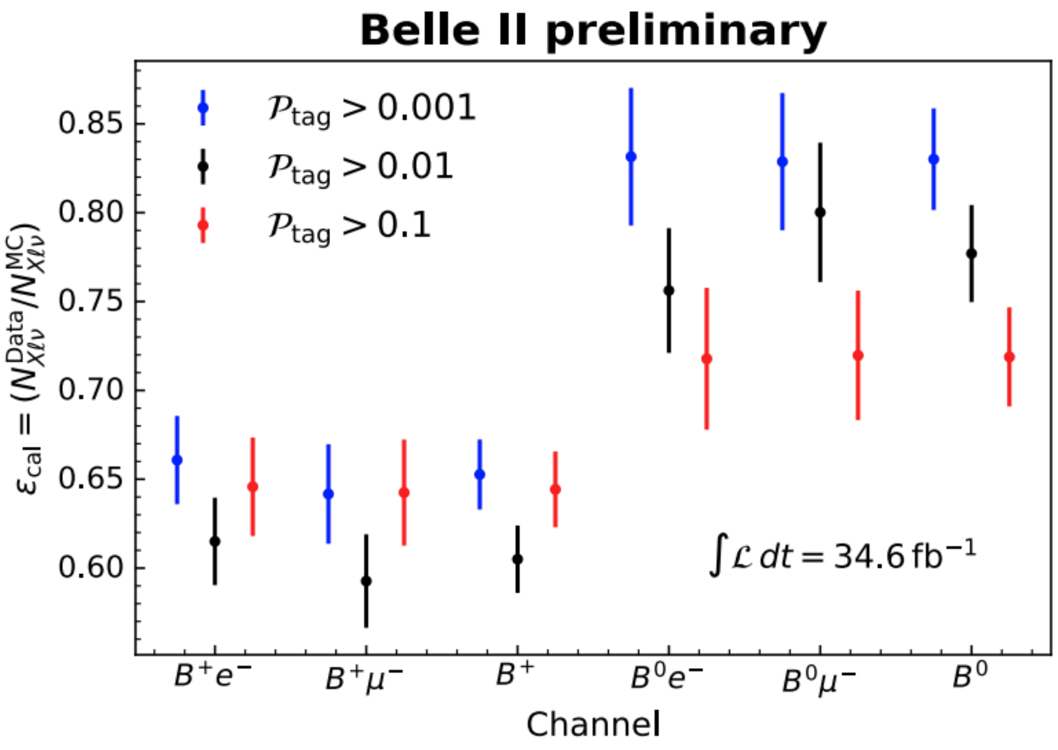
- The algorithm uses:
  - uncalibrated detector information
  - possibly outdated simulation (branching ratios, line shapes)
- Aggregates into the output score
- Use a well-measured independent process to calibrate the efficiency
- Assumption: Signal- and Tag-Side factorize (are independent)



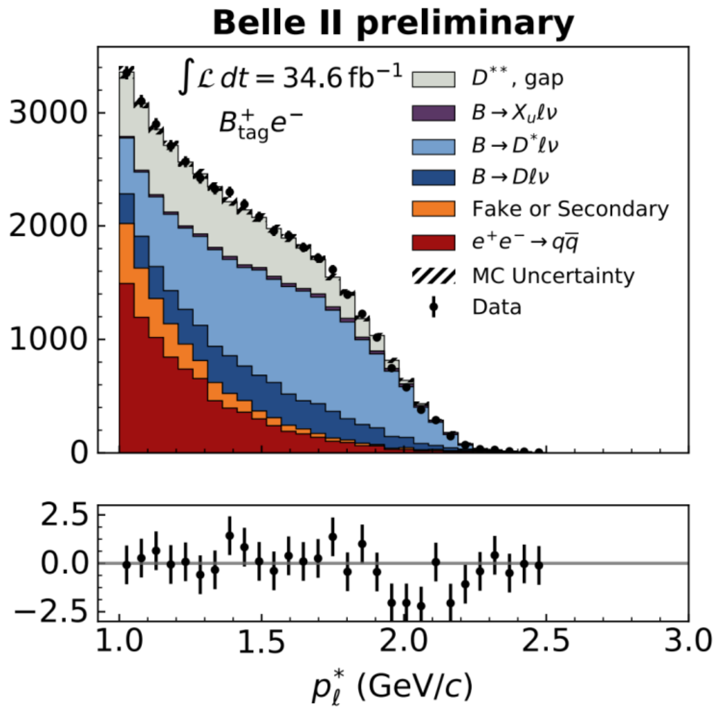

$$\epsilon = \frac{N^{Data}_{X_c \ell \nu}}{N^{MC}_{X_c \ell \nu}}$$



# Full Event Interpretation - Calibration



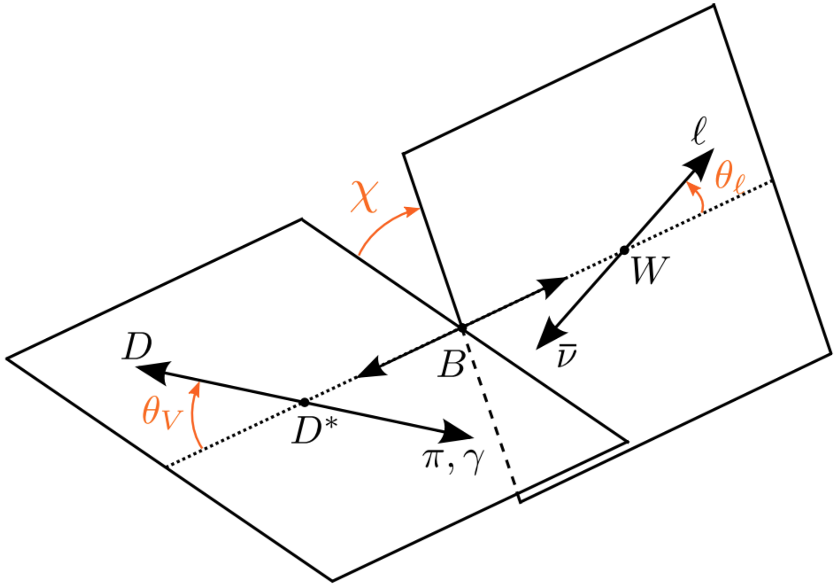

$$\epsilon = \frac{N^{Data}_{X_c\ell\nu}}{N^{MC}_{X_c\ell\nu}}$$




The tagging algorithm can be calibrated, but this introduces an additional systematic uncertainty (~3%) to the analysis

# Tagged Exclusive

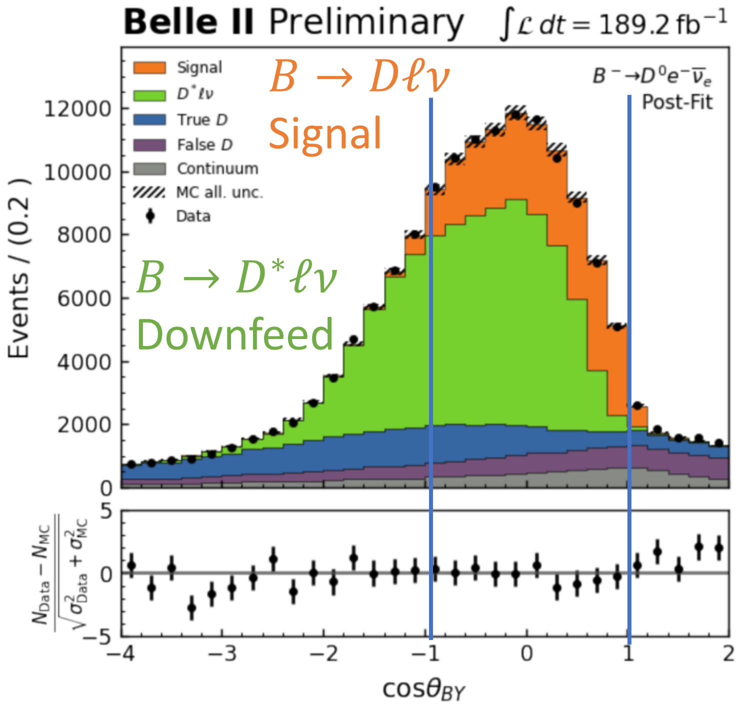
$$w = v_B \cdot v_{D^{(*)}} = \frac{m_B^2 + m_{D^{(*)}}^2 - q^2}{2 m_B m_{D^{(*)}}}$$




# Tagged $B \to D^{(*)}\ell\nu$

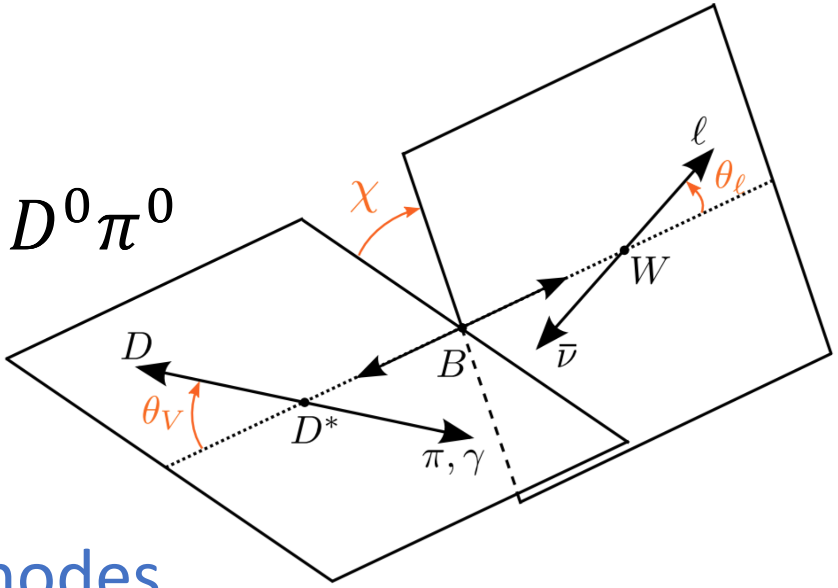
How are they different?

- $B \to D^*\ell\nu$: measure $\{w, \cos\theta_\ell \cos\theta_V, \chi\}$
  - Factor of ~3 larger branching fraction
  - $D^* \to D\pi_s$ slow pion efficiency needs to be understood
  - $D^*$ more challenging on the lattice
- $B \to D\ell\nu$: measure $\{w, \cos\theta_\ell\}$
  - Easier to reconstruct, but challenging large background component from $B \to D^*\ell\nu$ downfeed
- Future: Measure both decays simultaneously
  - link $B \to D^*\ell\nu$ signal and downfeed
  - Use that their form factors are not independent in the framework of HQET

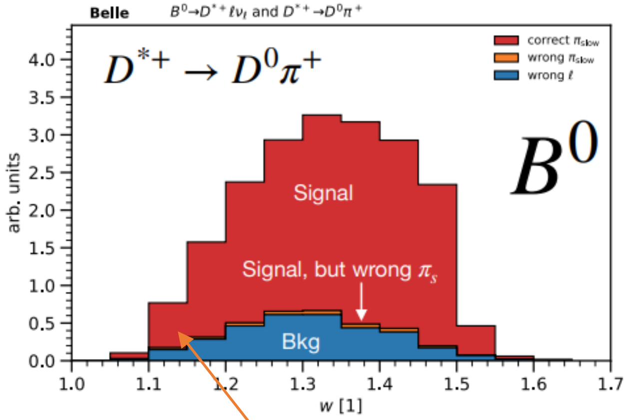


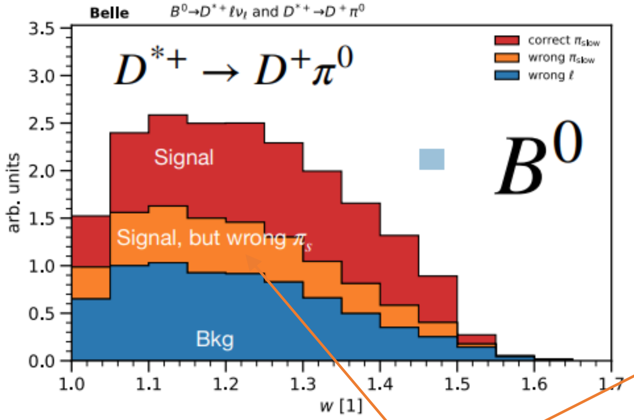
$$w = v_B \cdot v_{D^{(*)}} = \frac{m_B^2 + m_{D^{(*)}}^2 - q^2}{2 m_B m_{D^{(*)}}}$$

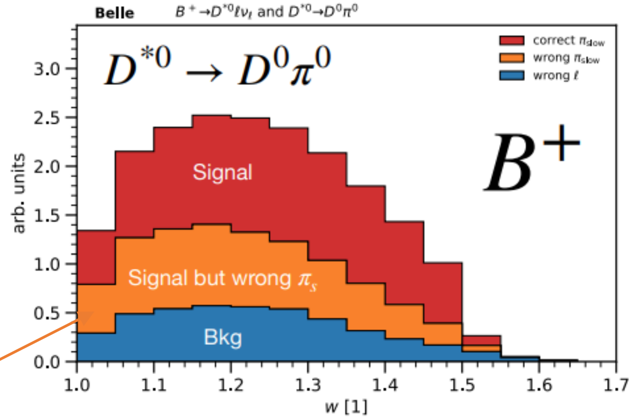
# Tagged $B \to D^* \ell \nu$




- Reconstruct $D^{*+} \to D^0\pi^+, D^{*+} \to D^+\pi^0, D^{*0} \to D^0\pi^0$
  $D^* \to D\gamma$ has a 30% branching fraction, why not add it in as well?
- B rest-frame can be directly reconstructed from the tag-side: Access to $w, \theta_l, \theta_V, \chi$
- But low effective statistics, reconstruct many D modes



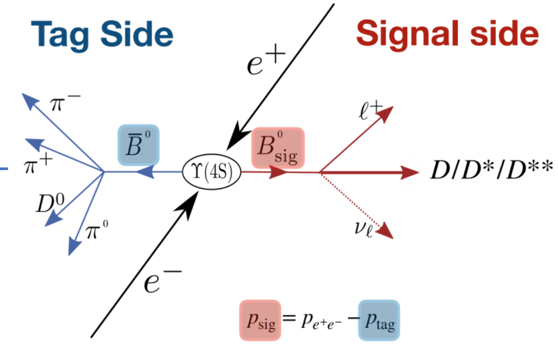




Where is this turn on coming from?

Why so many wrong $\pi_s$?

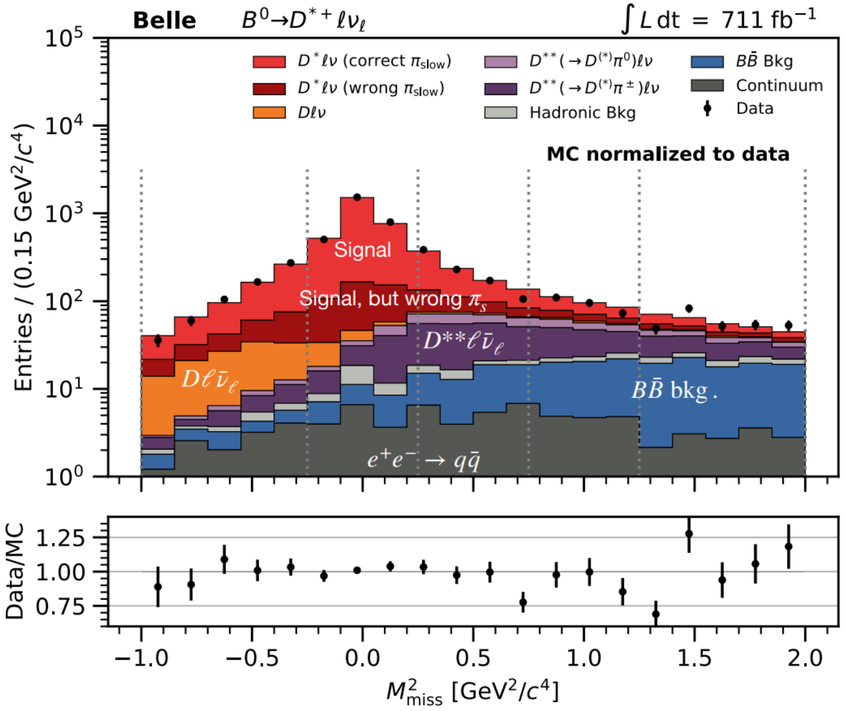
arXiv:2301.07529

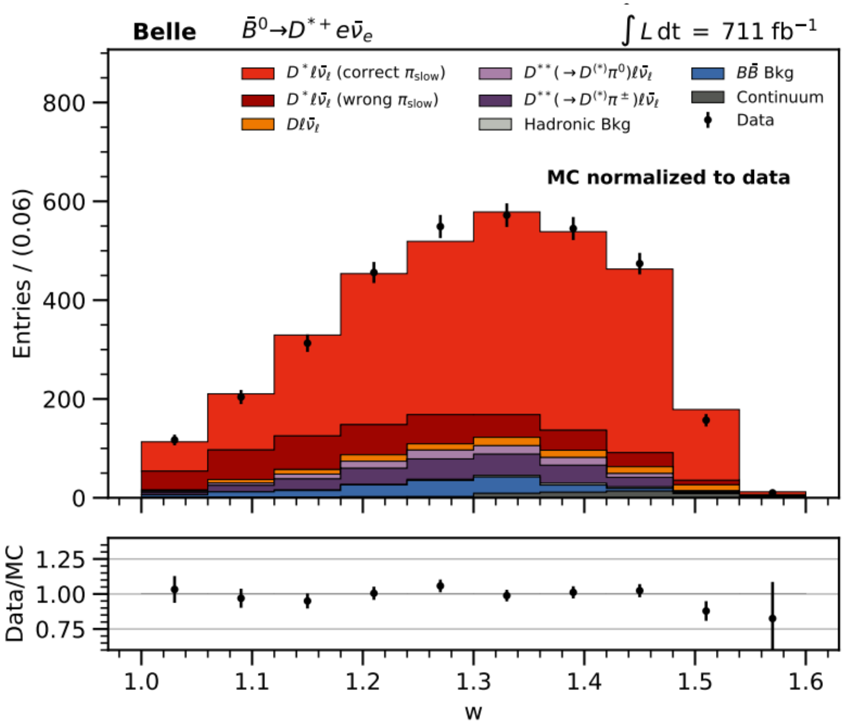
# Tagged $B \to D^* \ell \nu$ – Background Subtraction




- Need to subtract residual background contributions from
  - Other semileptonic decays ($B \to D\ell\nu$, $B \to D^{**}\ell\nu$)
  - Other B decays (fake or real leptons)
  - From continuum ($e^+e^- \to q\bar{q}$)

$$0 = m_\nu^2 = \mathrm{M}_{\mathrm{miss}}^2 = (\mathrm{E}_{\mathrm{miss}}, \mathrm{p}_{\mathrm{miss}})^2 = (\mathrm{p}_\mathrm{B} - \mathrm{p}_{\mathrm{D}^*} - \mathrm{p}_\ell)^2$$

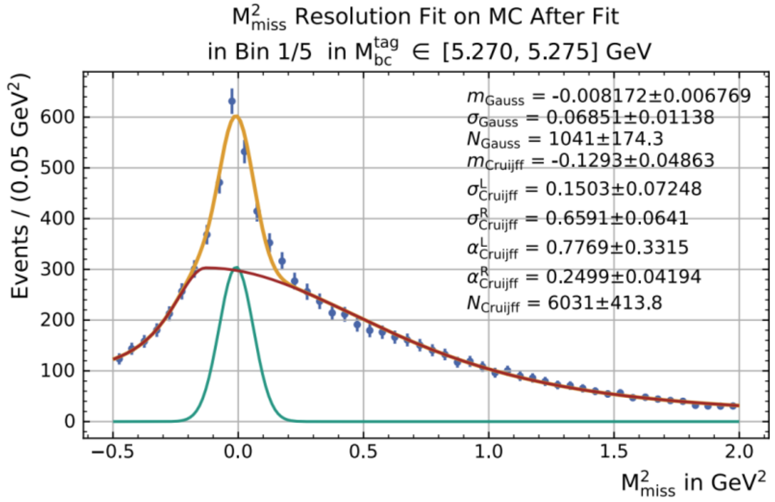


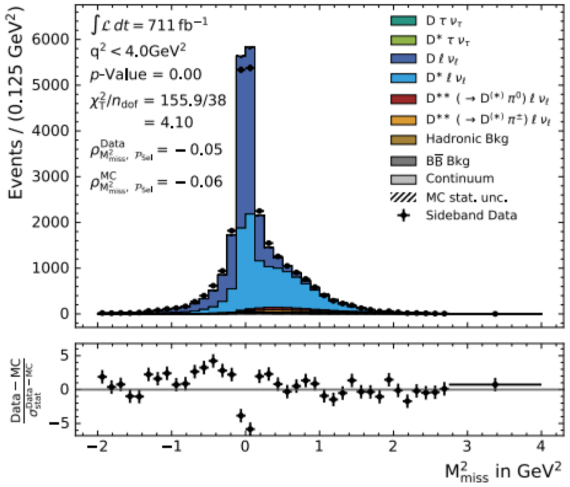



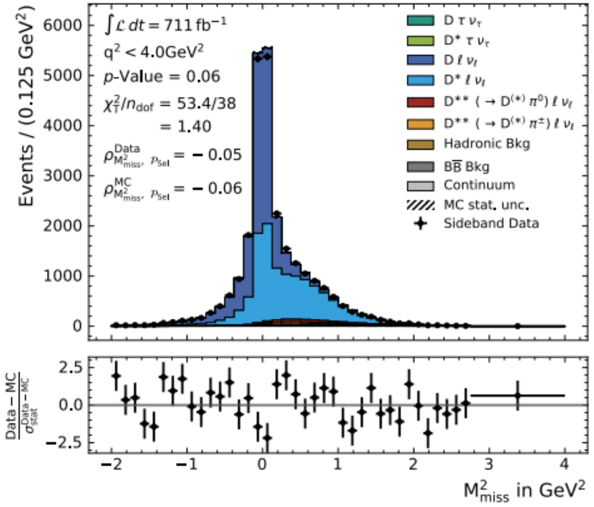

Alternatively, but same principle:
$U = E_{miss} - |p_{miss}|$

# Tagged $B \to D^* \ell \nu$ – Background Subtraction

- $\mathrm{M}^2_{\mathrm{miss}}$ is model independent, low impact of e.g., FF uncertainties
- But: MC modelling of $\mathrm{M}^2_{\mathrm{miss}}$ is challenging, non-trivial resolution effects due to the convolution of many variables
- Smearing function $f_{AL}$ derived from data



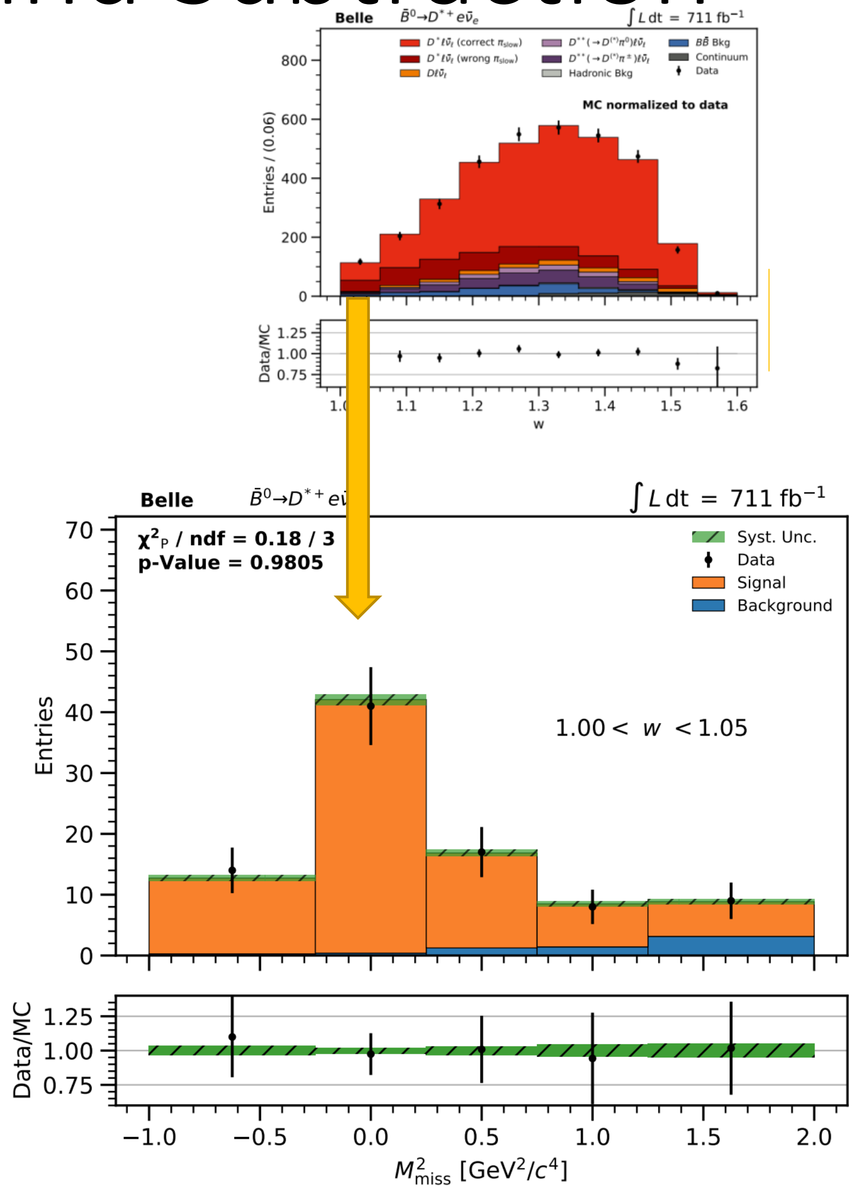

Asymmetric Laplace distribution

$$f_{\mathrm{AL}}(x; m, \lambda, \kappa) = \frac{\lambda}{\kappa + 1/\kappa} \begin{cases} \exp\left((\lambda/\kappa)(x-m)\right) & \text{if } x < m, \\ \exp(-\lambda\kappa(x-m)) & \text{if } x \geq m, \end{cases}$$

# Tagged $B \to D^* \ell \nu$ – Background Subtraction

- Different strategies available:
  - Binned likelihood fits to 1D projections; coarse binning reduces modelling dependence on e.g., background shape and resolution
  - Fit to the 4D distribution; binned approach suffers from curse of dimensionality; unbinned approach needs to deal with efficiency & migration
  - Measure angular coefficients

$$\frac{d^4\Gamma}{dq^2\, d\cos\theta^*\, d\cos\theta_\ell\, d\chi} = \frac{9}{32\pi} \left[ \left( I_1^s \sin^2\theta^* + I_1^c \cos^2\theta^* \right) + \left( I_2^s \sin^2\theta^* + I_2^c \cos^2\theta^* \right) \cos 2\theta_\ell \right.$$
$$+ I_3 \sin^2\theta^* \sin^2\theta_\ell \cos 2\chi + I_4 \sin 2\theta^* \sin 2\theta_\ell \cos\chi + I_5 \sin 2\theta^* \sin\theta_\ell \cos\chi$$
$$+ \left( I_6^c \cos^2\theta^* + I_6^s \sin^2\theta^* \right) \cos\theta_\ell + I_7 \sin 2\theta^* \sin\theta_\ell \sin\chi$$
$$\left. + I_8 \sin 2\theta^* \sin 2\theta_\ell \sin\chi + I_9 \sin^2\theta^* \sin^2\theta_\ell \sin 2\chi \right],$$



# Tagged $B \to D^* \ell \nu$ – Detector Migrations

- Resolution caused by detector effects and mis-reconstructions causes migration of events into neighboring bins
- Parametrized as a migration matrix

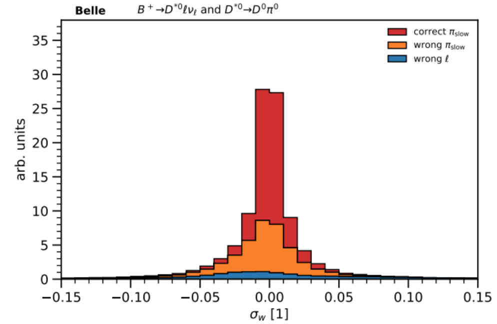
$$M_{ij} = P(\text{reco. in bin i} \mid \text{true value in bin j})$$

- Recover "true" values by this mapping of reconstructed → true
- "Simplest" method: Matrix inversion

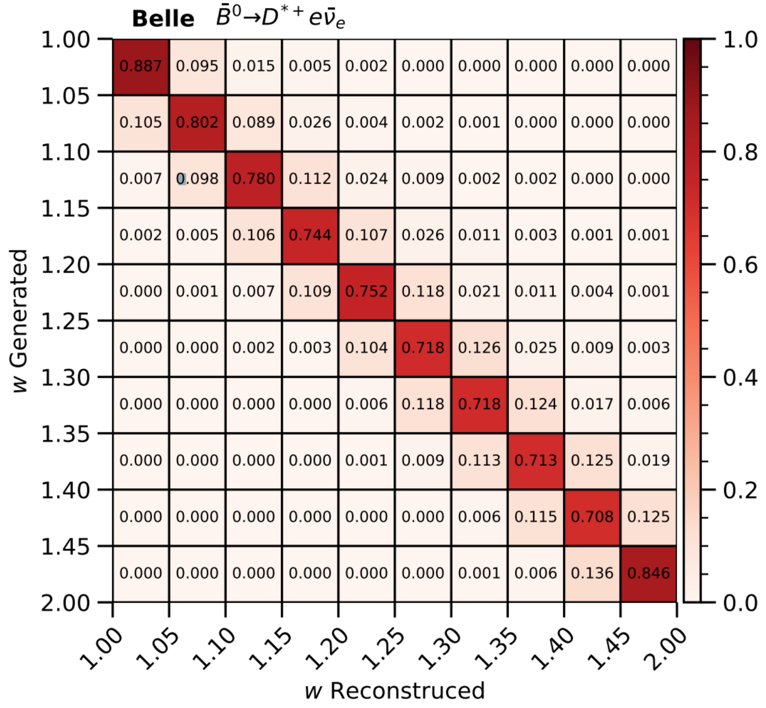
$$x_{true} = M_{ij}^{-1} x_{reco}$$

Unfolding is a whole topic on its own:

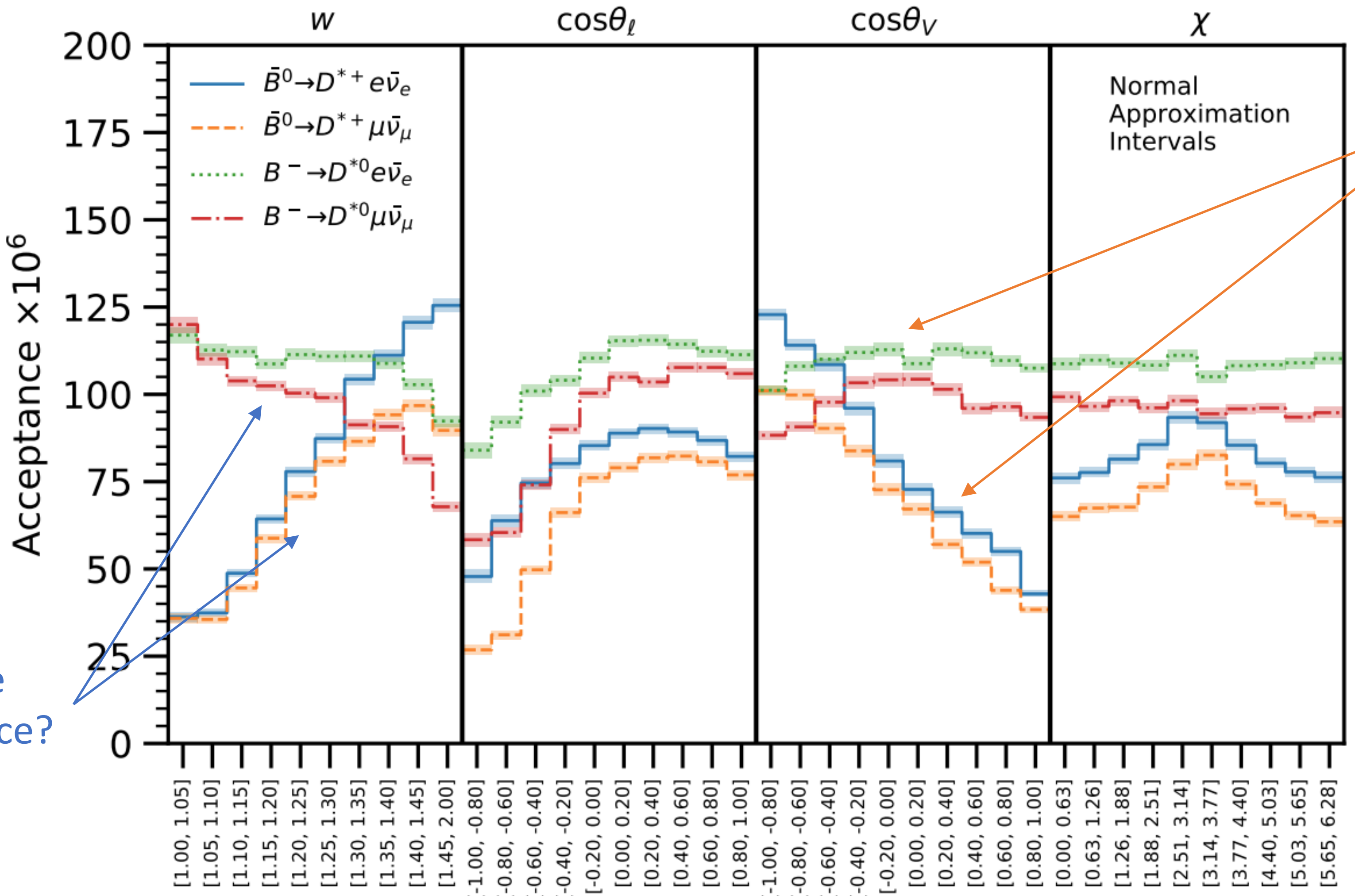
- Treatment of the variance-bias tradeoff
- Unbinned unfolding




causes

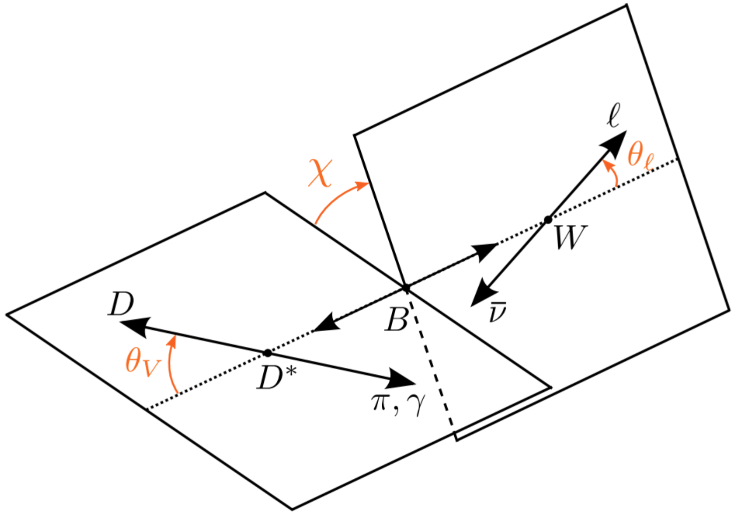


# Tagged $B \to D^* \ell \nu$ – Acceptance x Efficiency

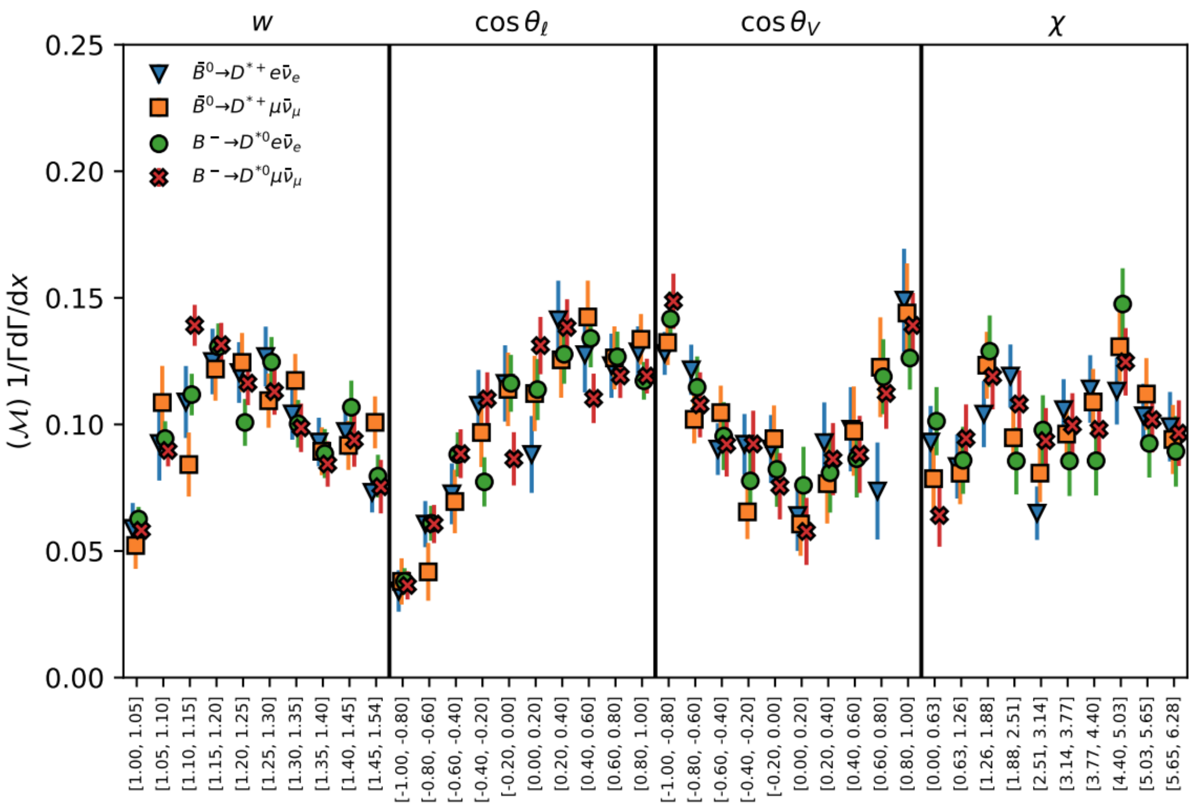



Why the difference?

Why the difference?



# Tagged $B \to D^* \ell \nu$ – Result




The „true“ 1D projections of the 4D decay rate after:

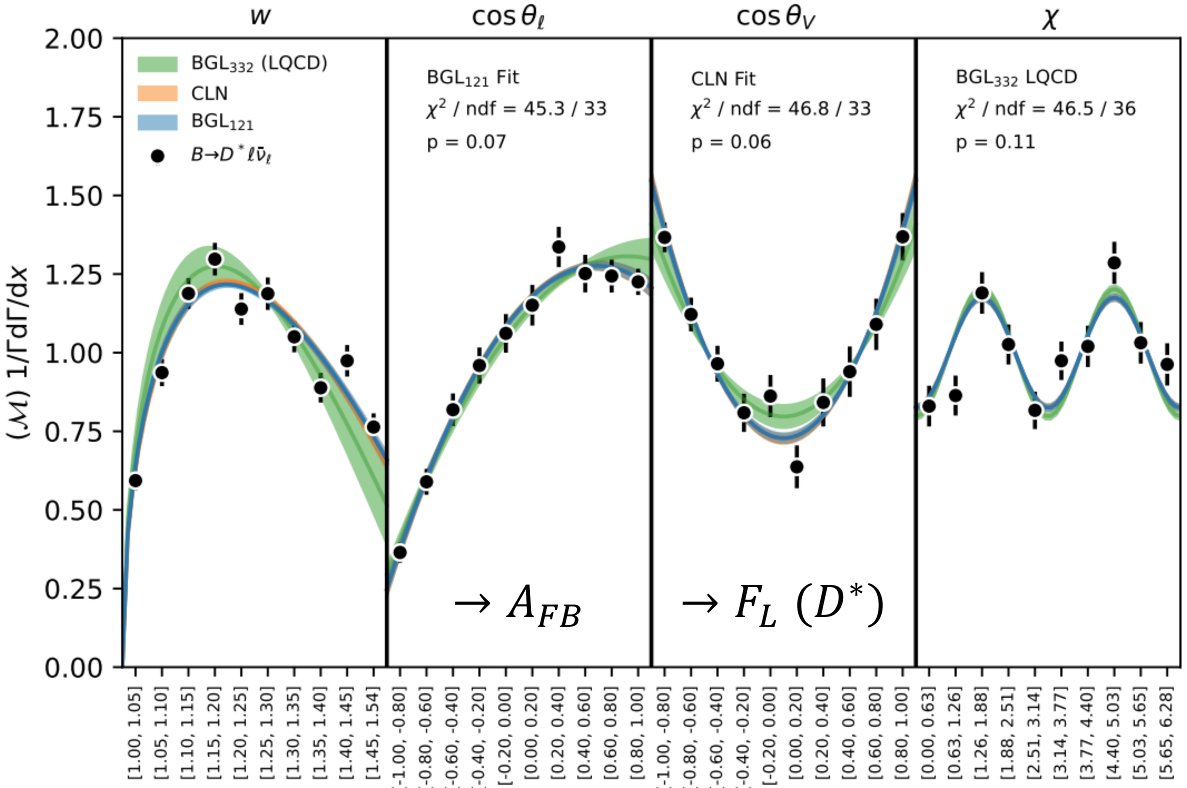
- Background subtraction
- Unfolding
- Correcting for acceptance and efficiency

Each 1D projection shows the same data!

- Determine correlations between different projections with bootstrapping
- Replicate the data by sampling with replacement and repeat analysis N times
- N depends on the
  - required precision on
  - true value of

  the correlation coefficients

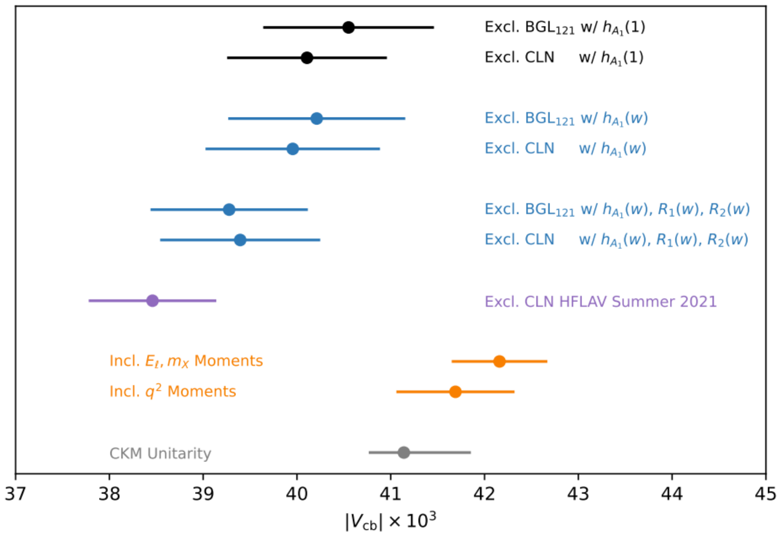
# Tagged $B \to D^* \ell \nu$ – Result




Both BGL and CLN can describe the data
Caveat using BGL: Truncation of the series

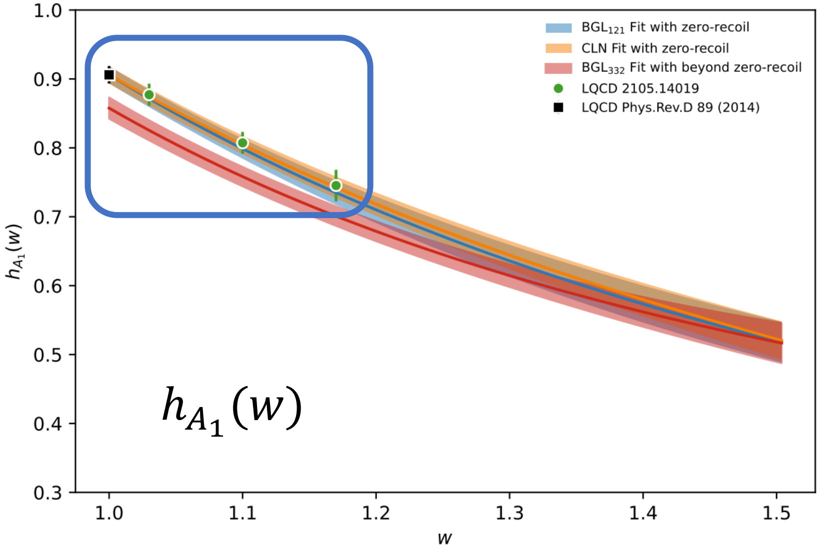
## Extract physics!

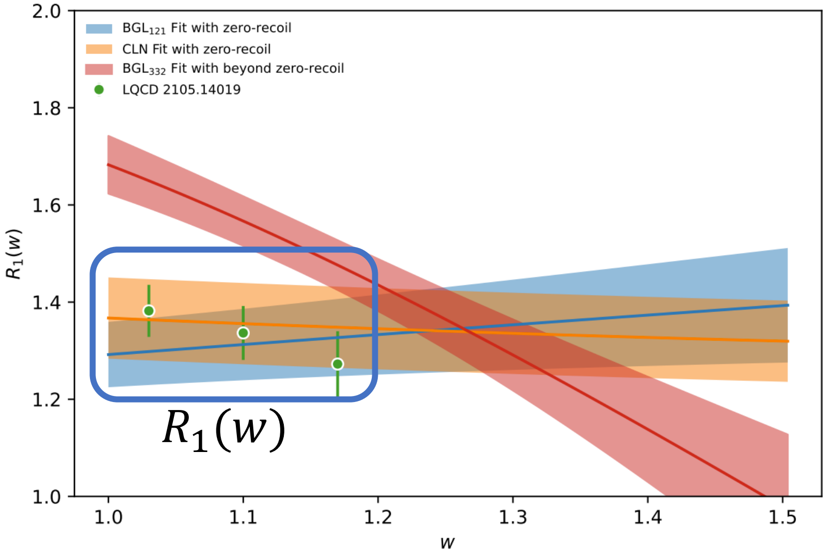
- Fit the 4D shapes with the model
- Choose the form factor parameterization
  - BGL, CLN, BLPR(XP)
- Extract form factors and $|V_{cb}|$ with the help from lattice QCD

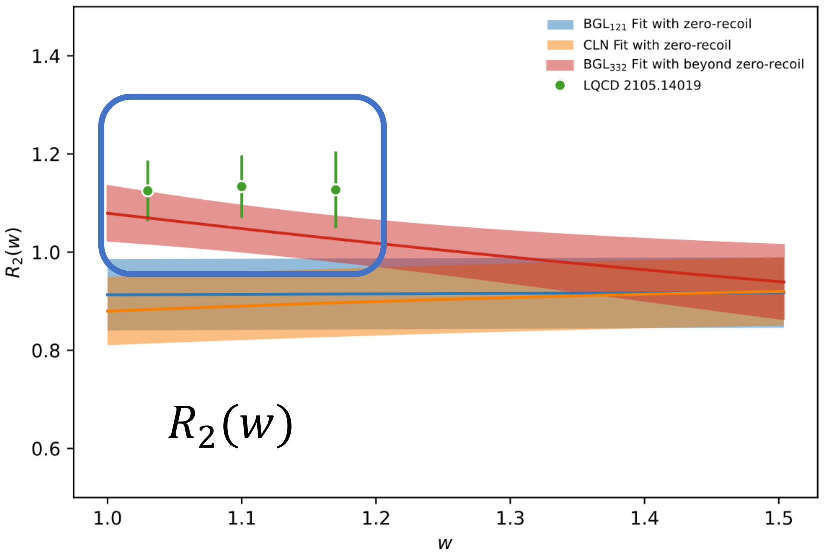


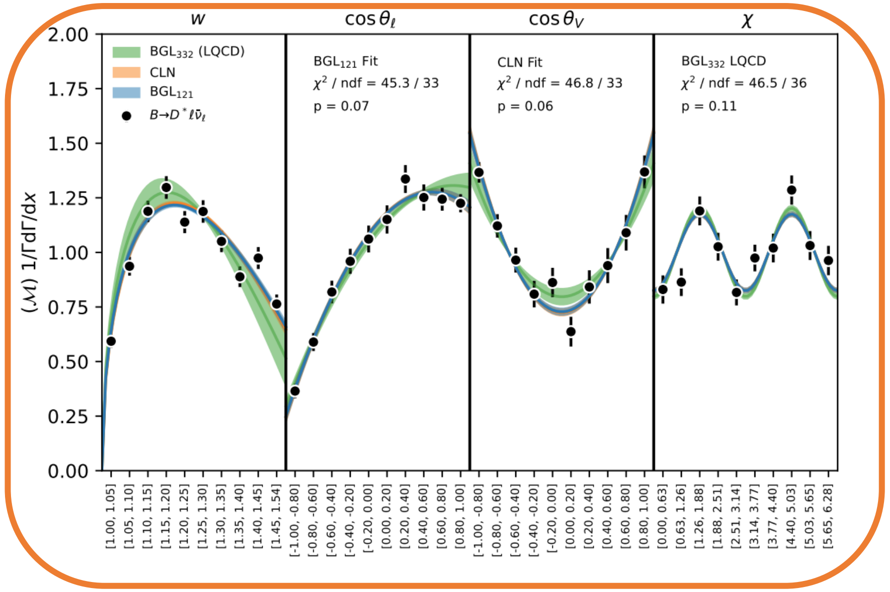

Extracted $|V_{cb}|$ depends on the lattice input

# Tagged $B \to D^* \ell \nu$ – Lattice Inputs

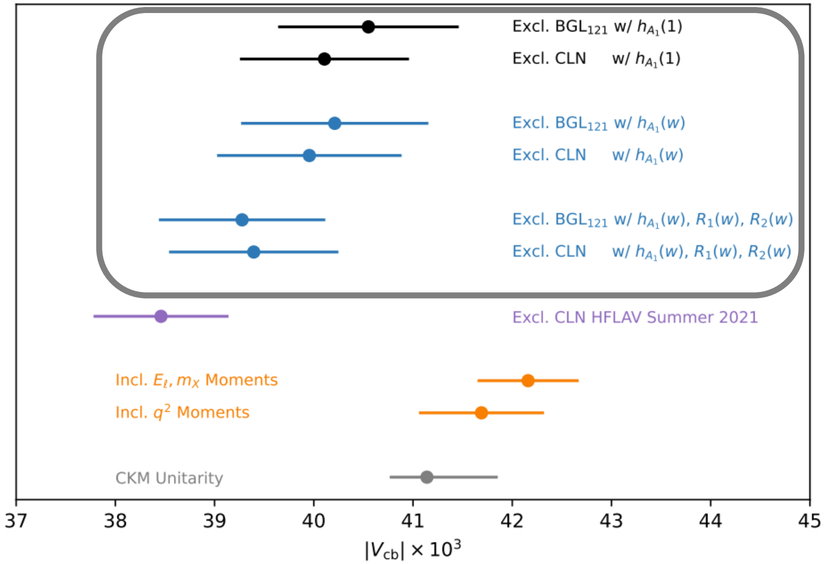



$$
\chi^2 = \left(\frac{\Delta \vec{\Gamma}^{\mathrm{m}}}{\Gamma^{\mathrm{m}}} - \frac{\Delta \vec{\Gamma}^{\mathrm{p}}(\vec{x})}{\Gamma^{\mathrm{p}}(\vec{x})}\right) C_{\mathrm{exp}}^{-1} \left(\frac{\Delta \vec{\Gamma}^{\mathrm{m}}}{\Gamma^{\mathrm{m}}} - \frac{\Delta \vec{\Gamma}^{\mathrm{p}}(\vec{x})}{\Gamma^{\mathrm{p}}(\vec{x})}\right)^T + \left(\Gamma^{\mathrm{ext}} - \Gamma^{\mathrm{p}}(\vec{x})\right)^2 / \sigma(\Gamma^{\mathrm{ext}})^2 + \left(h_X - h_X^{\mathrm{LQCD}}\right) C_{\mathrm{LQCD}}^{-1} \left(h_X - h_X^{\mathrm{LQCD}}\right), \quad (28)
$$



# Tagged $B \to D^{*}\ell\nu$ – Truncation

- One model-independent way to parameterize are BGL form factors
- How to truncate the series?
  - Truncate to soon: Introduces model dependence
  - Truncate to late: Increase variance of the result
- BGL form factors:

$$g(z) = \frac{1}{P_g(z)\phi_g(z)} \sum_{n=0}^{N} a_n z^n, \qquad f(z) = \frac{1}{P_f(z)\phi_f(z)} \sum_{n=0}^{N} b_n z^n, \qquad \mathcal{F}_1(z) = \frac{1}{P_{\mathcal{F}_1}(z)\phi_{\mathcal{F}_1}(z)} \sum_{n=0}^{N} c_n z^n,$$

# Tagged $B \to D^* \ell \nu$ – Truncation

## Nested hypothesis test

Bernlochner, Ligeti, Robinson 1902.09553

**Challenge nested fits**

$\mathrm{BGL}_{n_a,n_b,n_c}$ ↔ $\mathrm{BGL}_{n_a+1,n_b,n_c}$

$\mathrm{BGL}_{n_a,n_b,n_c}$ ↔ $\mathrm{BGL}_{n_a,n_b+1,n_c}$

$\mathrm{BGL}_{n_a,n_b,n_c}$ ↔ $\mathrm{BGL}_{n_a,n_b,n_c+1}$

**Test statistics & Decision boundary**

$\Delta\chi^2 = \chi^2_N - \chi^2_{N+1}$  $\Delta\chi^2 > 1$

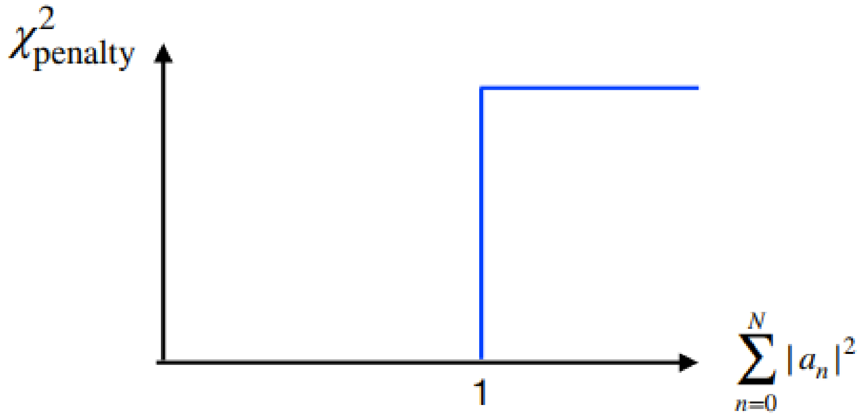
Distributed like a $\chi^2$-distribution with 1 dof (Wilk's theorem)

## Unitarity bounds

e.g. Gambino, Jung, Schacht 1905.08209

$$\sum_{n=0}^{N} |a_n|^2 \leq 1 \qquad \sum_{n=0}^{N} \left( |b_n|^2 + |c_n|^2 \right) \leq 1$$

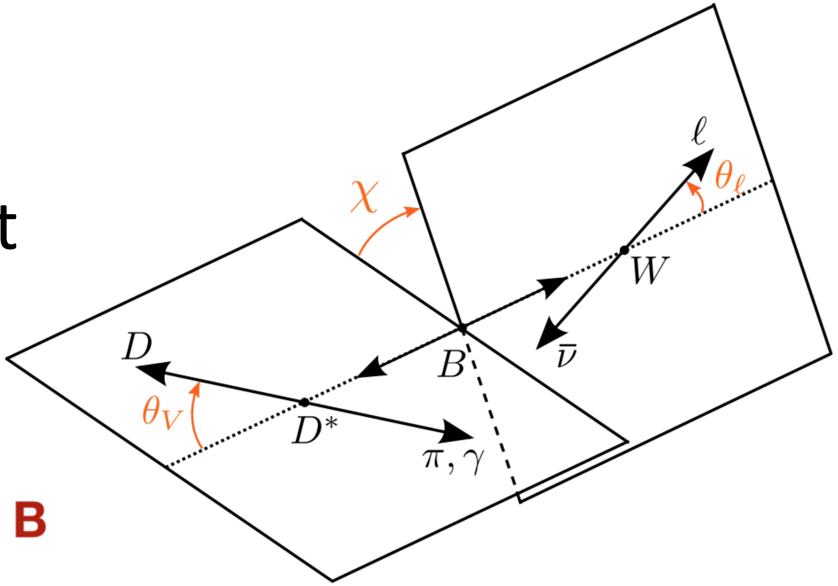
e.g. $\chi^2 \to \chi^2 + \chi^2_{\mathrm{penalty}}$



# Untagged Exclusive

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

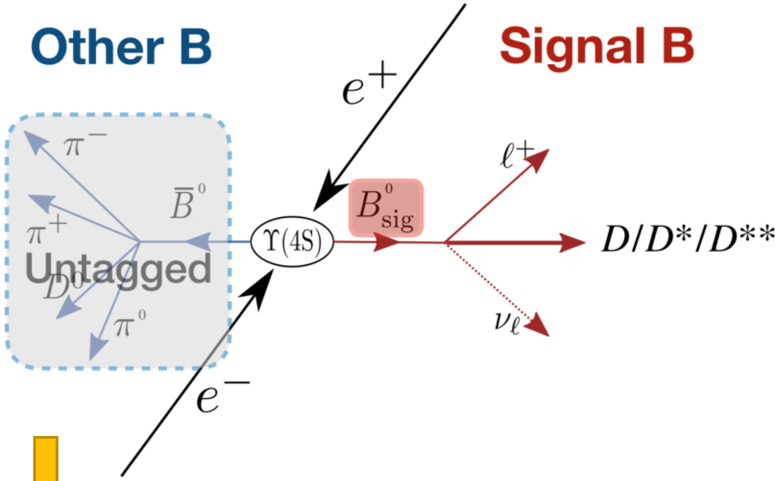
# Untagged $B \to D^* \ell \nu$




- Abundant statistics; reconstruct only the cleanest mode

$$D^{*+} \to D^0[\to K^+\pi^-]\pi^+$$

- Reconstruct signal side, everything else is assigned to the other $B$ meson
- Event kinematics: ROE method

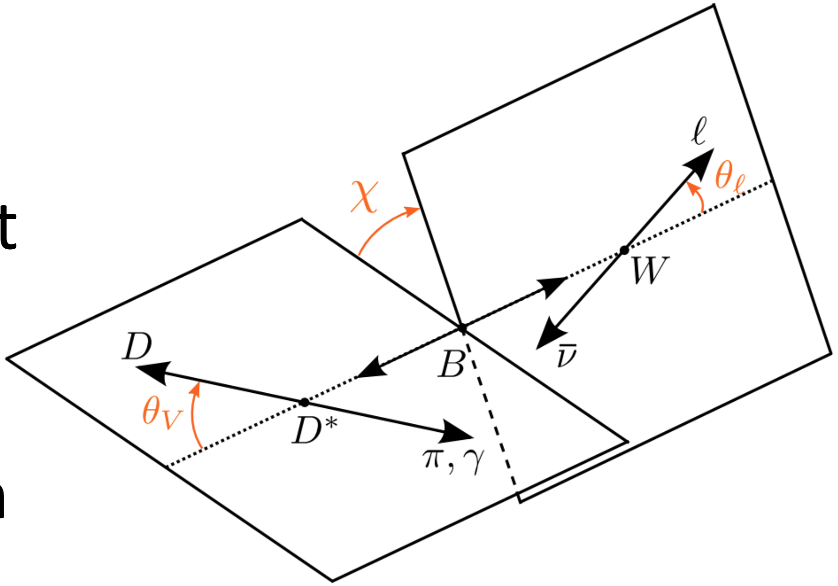



Why can we do this?

$$\vec{p}_{incl} = \sum_i p_i \quad \Rightarrow \quad \vec{p}_{B_{sig}} = -\vec{p}_{incl}$$

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

# Untagged $B \to D^* \ell \nu$

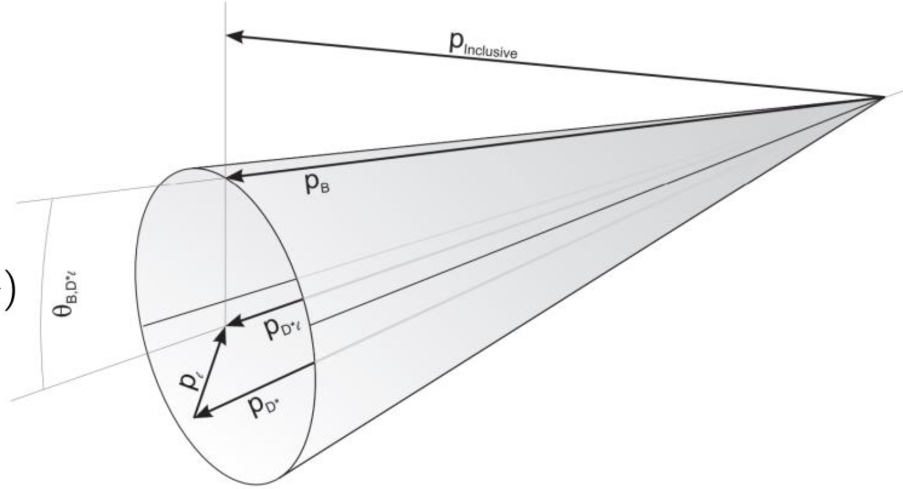



- Abundant statistics; reconstruct only the cleanest mode

$$D^{*+} \to D^0[\to K^+\pi^-]\pi^+$$

- Exploit that B meson lies on a cone, which has an opening angle defined by the visible particles

$$\cos\theta_{B,D^*\ell} = \frac{2E_B E_{D^*\ell} - m_B^2 - m_{D^*\ell}^2}{2|\vec{p}_B||\vec{p}_{D^*\ell}|}$$



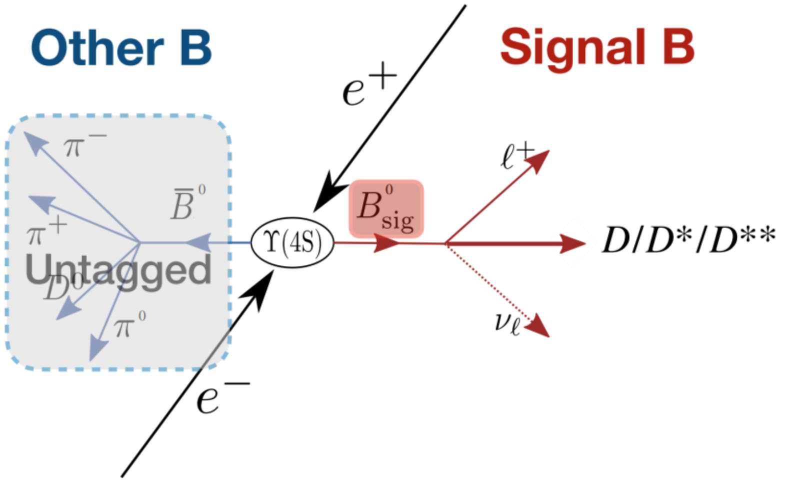

- Calculate for 10 points on the cone

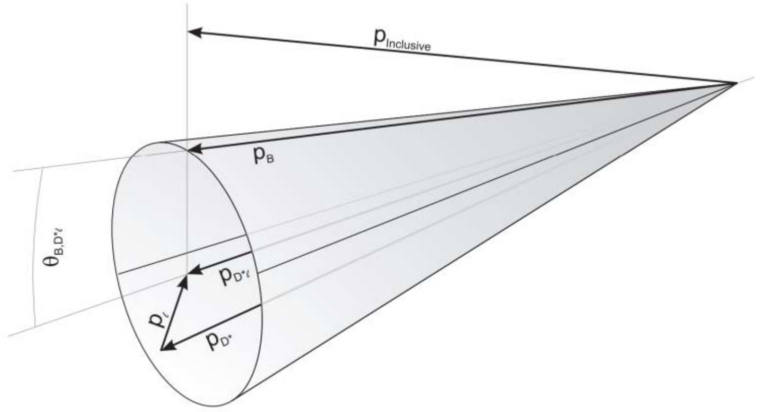
$$(E^B, p_x^B, p_y^B, p_z^B) = (E_{\mathrm{Beam}}^{\mathrm{CM}}/2, |\vec{p}_B^{\mathrm{CM}}| \sin\theta_{BY}\cos\phi, |\vec{p}_B^{\mathrm{CM}}| \sin\theta_{BY}\sin\phi, |\vec{p}_B^{\mathrm{CM}}| \cos\theta_{BY})$$

- Utilize that the angular distribution of $\Upsilon(4S) \to B\bar{B}$ is $\sin^2\theta_B$
  Weighted average over the 10 points $w_i = \sin^2\theta_B$

polar angle

# Untagged $B \to D^* \ell \nu$



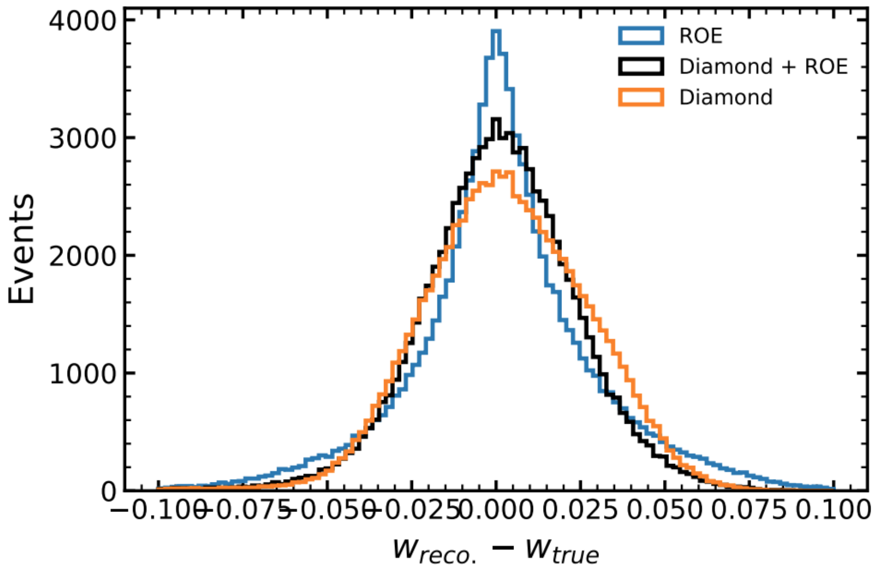




$$\cos\theta_{B,D^*\ell} = \frac{2E_B E_{D^*\ell} - m_B^2 - m_{D^*\ell}^2}{2|\vec{p}_B||\vec{p}_{D^*\ell}|}$$

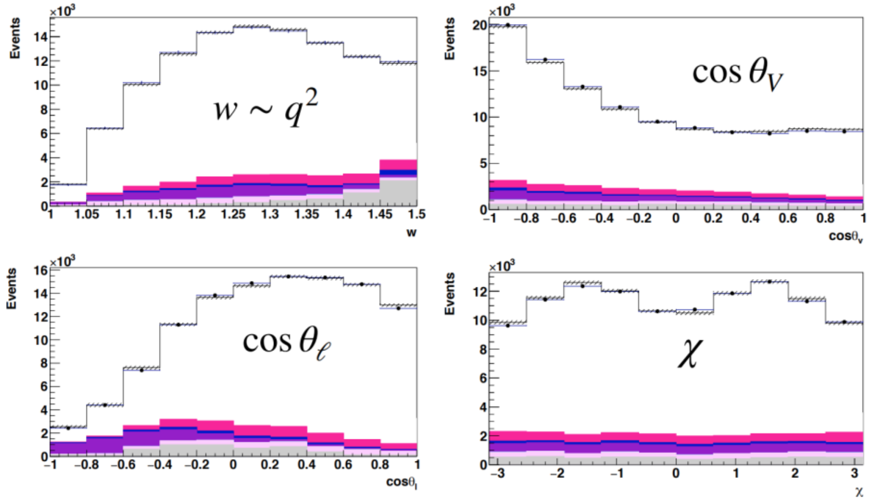
$$(E^B, p_x^B, p_y^B, p_z^B) = (E^{CM}_{Beam}/2, |\vec{p}_B^{\,CM}|\sin\theta_{BY}\cos\phi, |\vec{p}_B^{\,CM}|\sin\theta_{BY}\sin\phi, |\vec{p}_B^{\,CM}|\cos\theta_{BY})$$

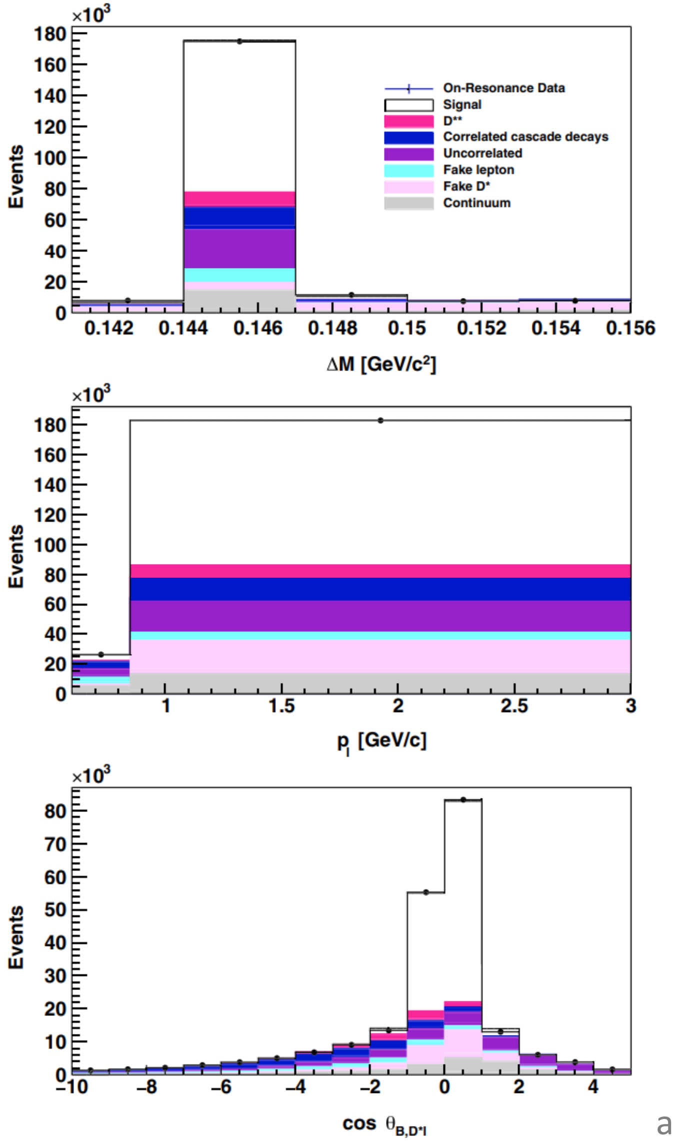
$$\vec{p}_{incl} = \sum_i p_i \Rightarrow \vec{p}_{B_{sig}} = -\vec{p}_{incl}$$

Both methods can be combined!



# Untagged $B \to D^* \ell \nu$ – Background Subtraction




Subtract residual backgrounds using

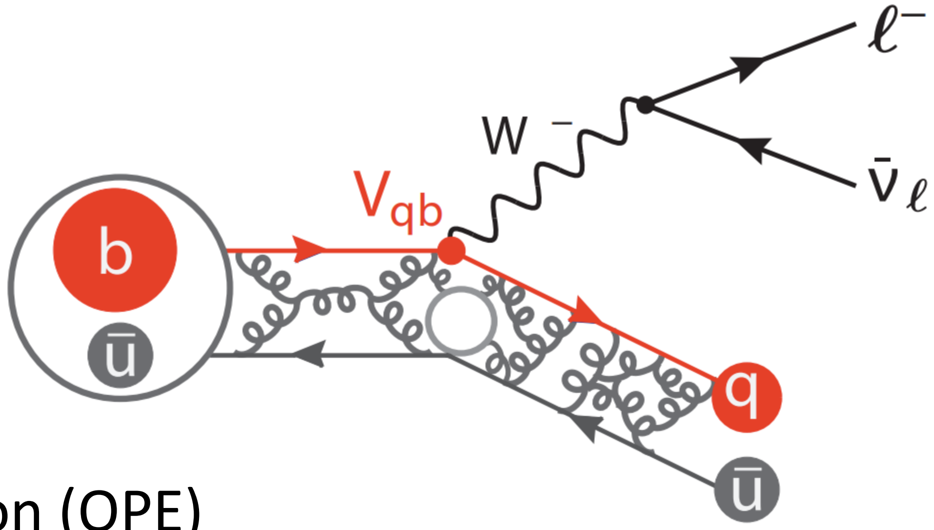
- $\Delta M = m_{D^*} - m_D$ discriminates fake and true $D^*$
- $\cos\theta_{B,D^*\ell} = \frac{2E_B E_{D^*\ell} - m_B^2 - m_{D^*\ell}^2}{2|\vec{p}_B||\vec{p}_{D^*\ell}|}$ discriminates signal and background
- $p_\ell$ to control fake leptons

From here proceed same as for the tagged analysis

arXiv:1809.03290

# Tagged Inclusive

# Inclusive Measurements




- Inclusive measurements stay agnostic with respect to the hadronic system
- Theoretical framework is Operator Product Expansion (OPE)

$$d\Gamma = d\Gamma_0 + d\Gamma_{\mu_\pi}\frac{\mu_\pi^2}{m_b^2} + d\Gamma_{\mu_G}\frac{\mu_G^2}{m_b^2} + d\Gamma_{\rho_D}\frac{\rho_D^3}{m_b^3} + d\Gamma_{\rho_{LS}}\frac{\rho_{LS}^3}{m_b^3} + \ldots$$

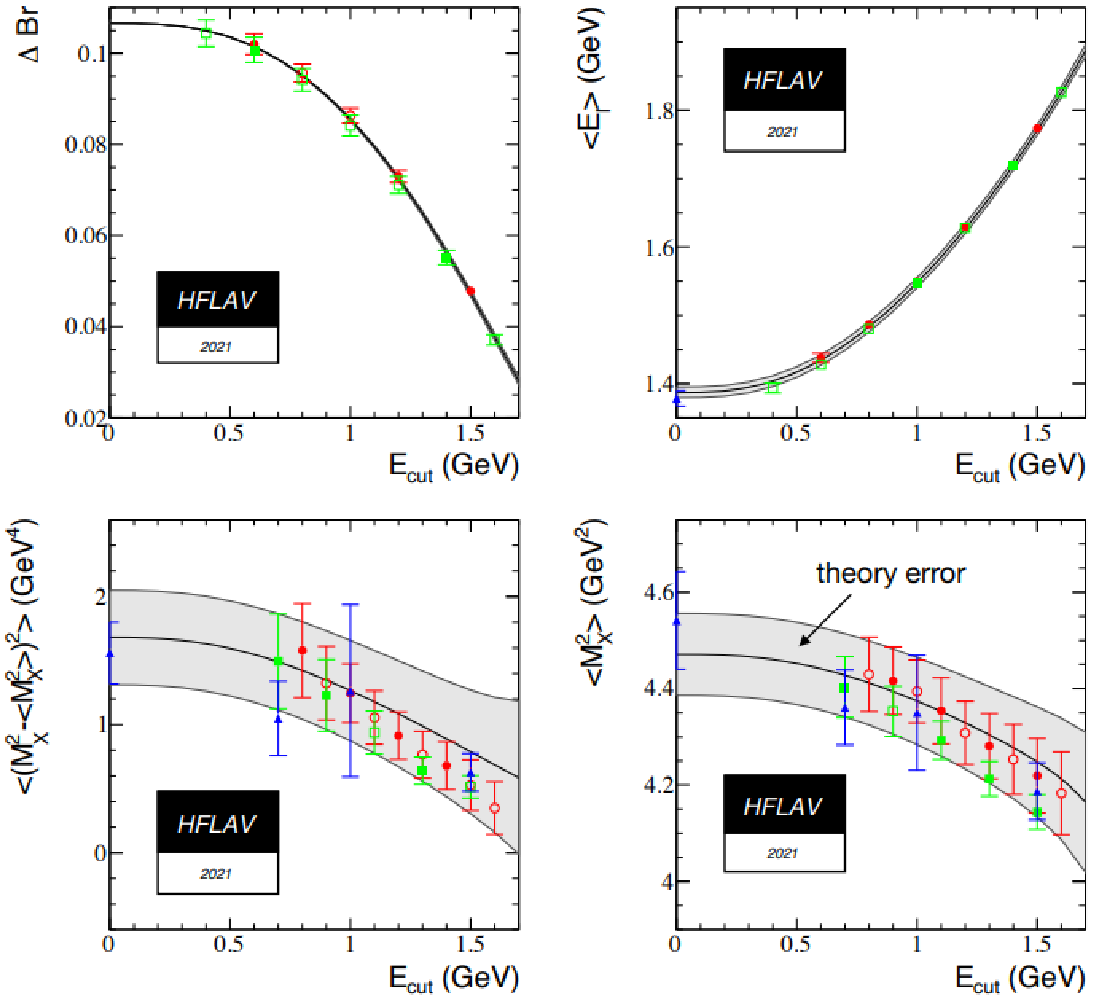
- $d\Gamma$ are calculated perturbatively
- Non-perturbative dynamics encapsulated in the HQE parameters $\mu_\pi, \mu_G, \rho_D, \rho_{LS}$

→ Extract HQE parameters from data (similar to the form factors)

→ Measure spectral moments: hadronic mass, lepton energy, momentum transfer, …

# Inclusive $B \to X_c \ell \nu \langle m_X, E_l \rangle$

| Experiment | Hadron moments $\langle M^n_X \rangle$ | Lepton moments $\langle E^n_l \rangle$ | References |
|---|---|---|---|
| BaBar | n=2 c=0.9,1.1,1.3,1.5<br>n=4 c=0.8,1.0,1.2,1.4<br>n=6 c=0.9,1.3 [1] | n=0 c=0.6,1.2,1.5<br>n=1 c=0.6,0.8,1.0,1.2,1.5<br>n=2 c=0.6,1.0,1.5<br>n=3 c=0.8,1.2 [1,2] | [1] Phys.Rev. D81 (2010) 032003<br>[2] Phys.Rev. D69 (2004) 111104 |
| Belle | n=2 c=0.7,1.1,1.3,1.5<br>n=4 c=0.7,0.9,1.3 [3] | n=0 c=0.6,1.4<br>n=1 c=1.0,1.4<br>n=2 c=0.6,1.4<br>n=3 c=0.8,1.2 [4] | [3] Phys.Rev. D75 (2007) 032005<br>[4] Phys.Rev. D75 (2007) 032001 |
| CDF | n=2 c=0.7<br>n=4 c=0.7 [5] | . | [5] Phys.Rev. D71 (2005) 051103 |
| CLEO | n=2 c=1.0,1.5<br>n=4 c=1.0,1.5 [6] | . | [6] Phys.Rev. D70 (2004) 032002 |
| DELPHI | n=2 c=0.0<br>n=4 c=0.0<br>n=6 c=0.0 [7] | n=1 c=0.0<br>n=2 c=0.0<br>n=3 c=0.0 [7] | [7] Eur.Phys.J. C45 (2006) 35-59 |




| Br(B -> X$_c$lnu) (%) | $\|V_{cb}\|$ ($10^{-3}$) | $m_b^{kin}$ (GeV) | $mu^2_{pi}$ (GeV$^2$) | |
|---|---|---|---|---|
| 10.65 +/- 0.16 | 42.19 +/- 0.78 | 4.554 +/- 0.018 | 0.464 +/- 0.076 | details |

- HQE parameters extracted from the measured moments
- Semileptonic rate from theory

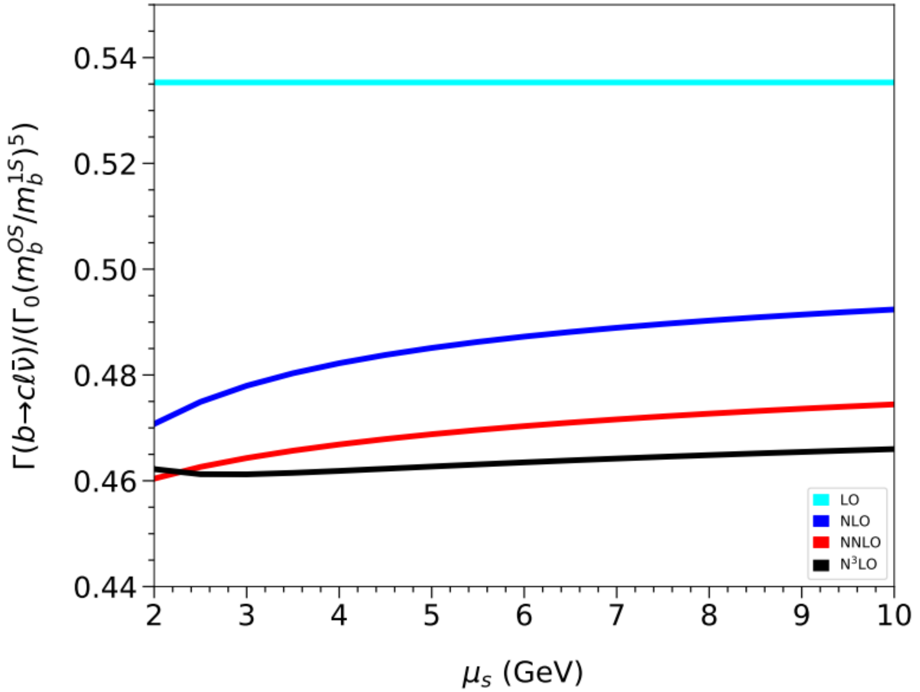
# State-of-the-Art

- Relatively old measurements, but recent progress on the theory side!
  Semileptonic decay rate at N3LO

  M. Fael, K. Schönwald, M. Steinhauser
  Phys.Rev.D 104 (2021) 1, 016003, arXiv:2011.13654

- Updated inclusive fit to $\langle M_X \rangle, \langle E_\ell \rangle$
  $|V_{cb}| = 42.16 \times 10^{-3}$ with 1.2% precision

  M. Bordone, B. Capdevila, P. Gambino
  Phys.Lett.B 822 (2021) 136679, arXiv:2107.00604




Renormalization Scale

# Inclusive $B \to X_c \ell \nu \langle q^2 \rangle$

- Number of matrix elements increase at higher orders

$$d\Gamma = d\Gamma_0 + d\Gamma_{\mu_\pi} \frac{\mu_\pi^2}{m_b^2} + d\Gamma_{\mu_G} \frac{\mu_G^2}{m_b^2} + d\Gamma_{\rho_D} \frac{\rho_D^3}{m_b^3} + d\Gamma_{\rho_{LS}} \frac{\rho_{LS}^3}{m_b^3} + \dots$$

- New idea: Exploit reparameterization invariance

M. Fael, T. Mannel, K. Vos
JHEP 02 (2019) 177, arXiv:1812.07472

- Spectral moments

$$v = p_B/m_B$$

$$\langle M^n[w] \rangle = \int d\Phi \, w^n(v, p_\ell, p_\nu) \, W^{\mu\nu} L_{\mu\nu}$$

$w = (m_B v - q)^2 \Rightarrow \langle M_X^n \rangle$ Moments — not RPI (depends on $v$)

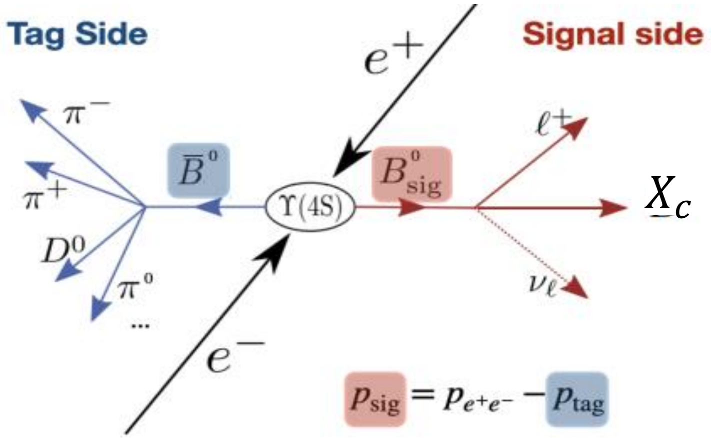
$w = v \cdot p_\ell \Rightarrow \langle E_\ell^n \rangle$ Moments — not RPI (depends on $v$)

$w = q^2 \Rightarrow \langle (q^2)^n \rangle$ Moments — RPI! (does not depend on $v$)

$\langle q^2 \rangle$ moments measured by Belle and Belle II

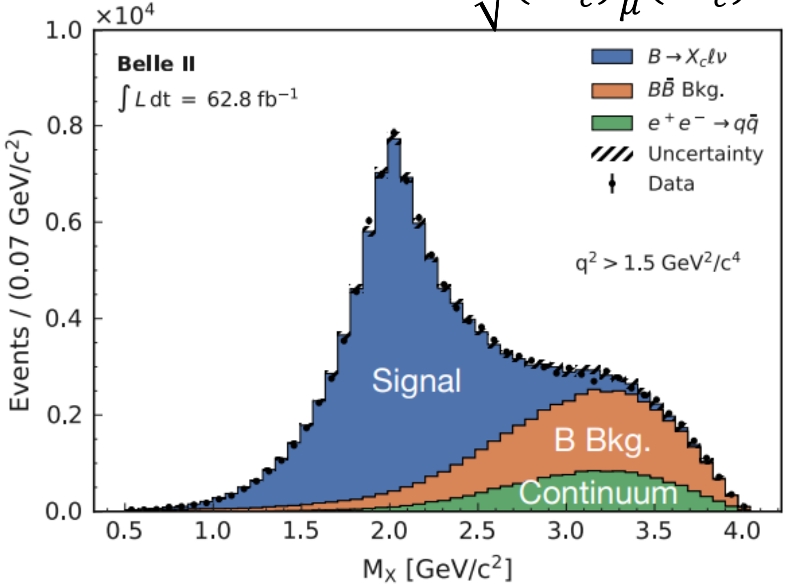
PRD 104, 112011 (2021), arXiv:2109.01685
Submitted to PRD, arXiv:2205.06372

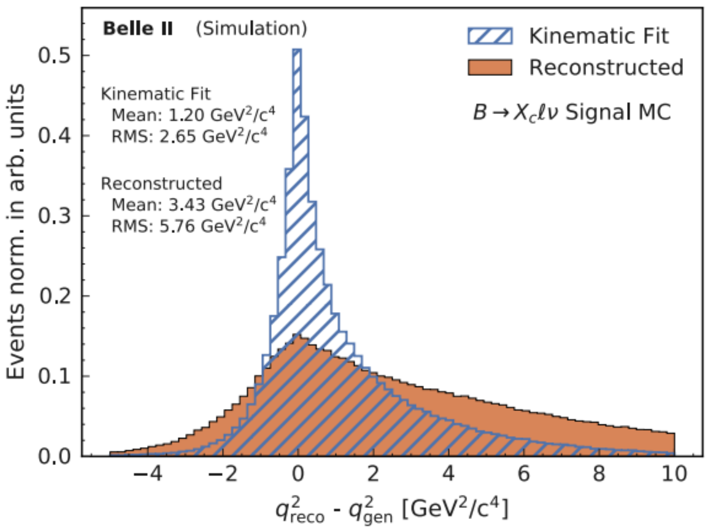
# Inclusive $B \to X_c \ell \nu \; \langle q^2 \rangle$



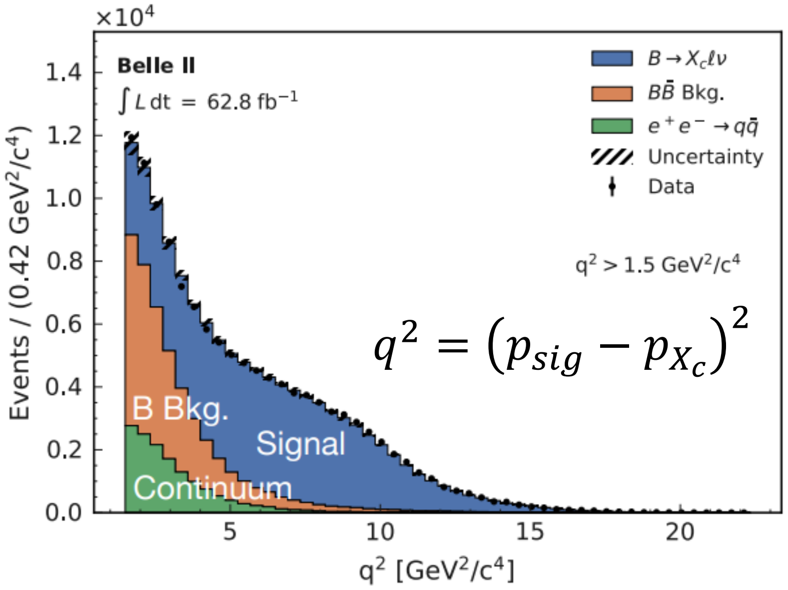

Access to full event kinematics via hadronic tagging

$$M_X = \sqrt{(p_{X_c})_\mu (p_{X_c})^\mu}$$




Kinematic fit drastically improves resolution

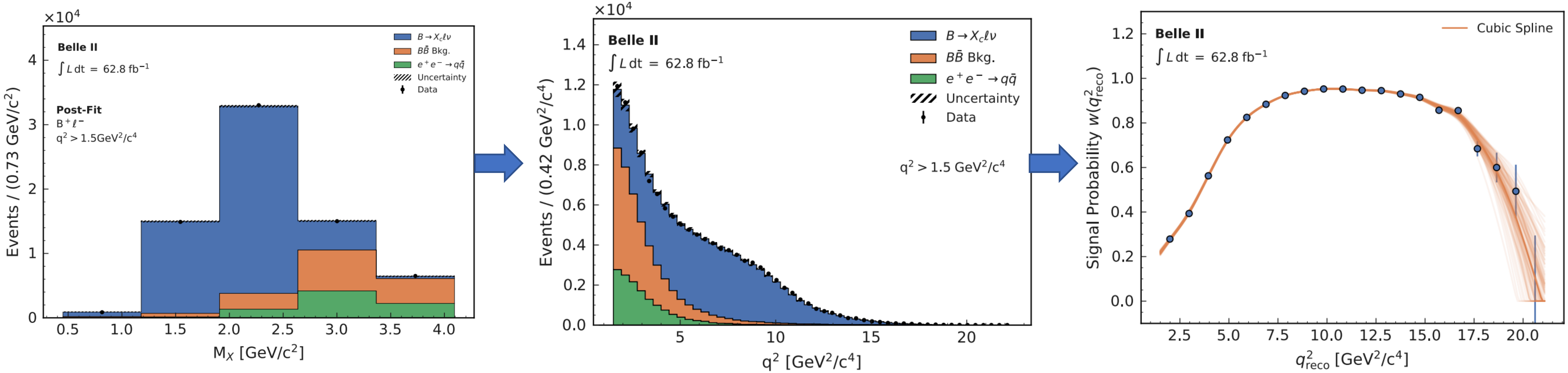


# Inclusive $B \to X_c \ell \nu$ $\langle q^2 \rangle$

Event-wise Master-formula

$$\langle q^{2n} \rangle = \frac{\sum_i^{N_{data}} w(q^2_{i,reco}) \times q^{2n}_{i,calib}}{\sum_j^{N_{data}} w(q^2_{j,reco})} \times C_{calib} \times C_{gen}$$

- Step 1: Subtract Background




Determine background normalization by fitting $M_X$ and determine event weights

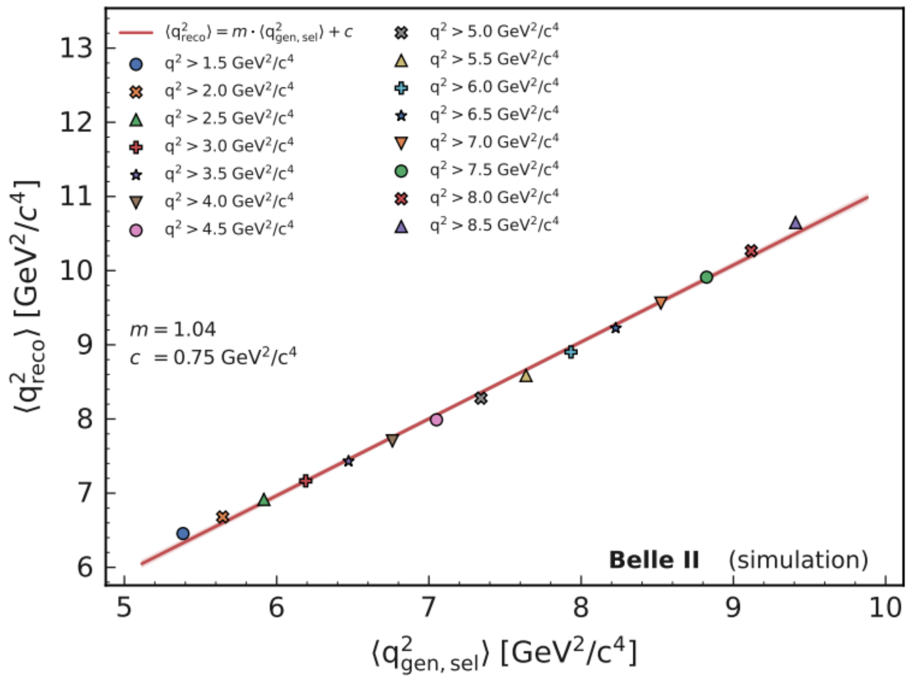
# Inclusive $B \to X_c \ell \nu \; \langle q^2 \rangle$

Event-wise Master-formula

$$\langle q^{2n} \rangle = \frac{\sum_i^{N_{data}} w(q^2_{i,reco}) \times q^{2n}_{i,calib}}{\sum_j^{N_{data}} w(q^2_{j,reco})} \times C_{calib} \times C_{gen}$$

- Step 2: Calibrate Moments
- Exploit linear dependence between reconstructed and true moments

$$q^{2m}_{i,cal} = (q^{2m}_{i,reco} - c)/m$$



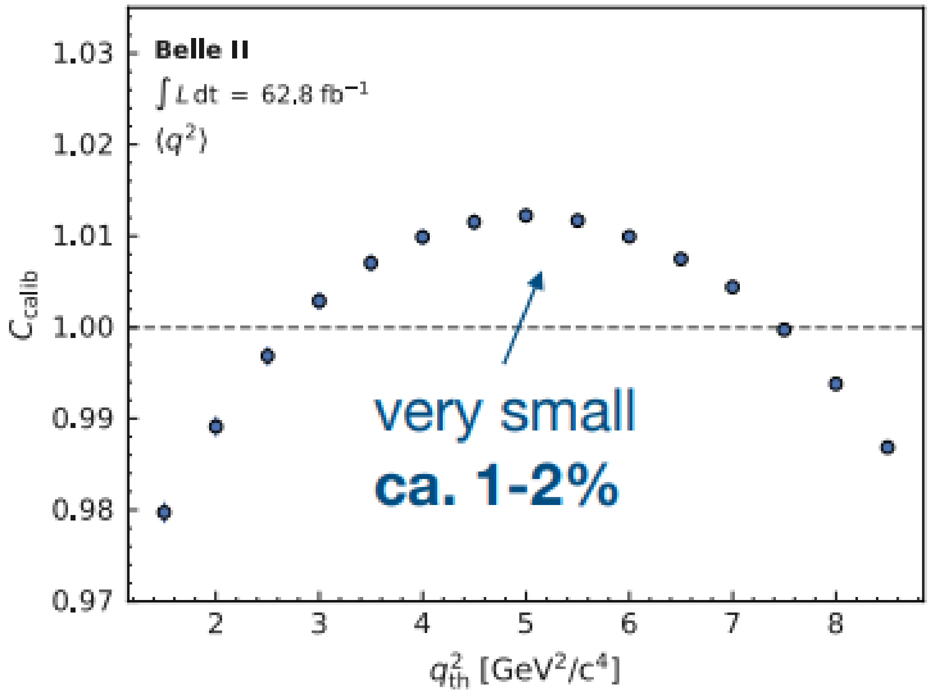
# Inclusive $B \to X_c \ell \nu \ \langle q^2 \rangle$

Event-wise Master-formula

$$\langle q^{2n} \rangle = \frac{\sum_i^{N_{data}} w(q^2_{i,reco}) \times q^{2n}_{i,calib}}{\sum_j^{N_{data}} w(q^2_{j,reco})} \times C_{calib} \times C_{gen}$$

- Step 3: Refine calibration
- Correct for small deviations from the linear behavior

$$C_{calib} = \langle q^{2n}_{gen,sel} \rangle / \langle q^{2n}_{calib} \rangle$$



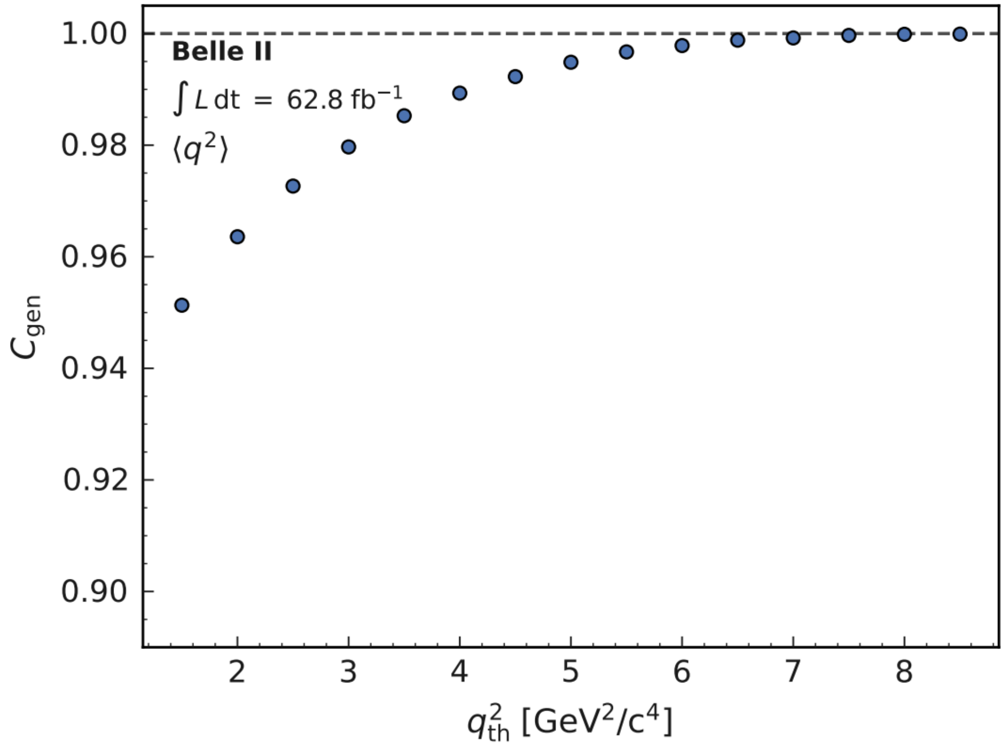
Event-wise Master-formula

$$\langle q^{2n}\rangle = \frac{\sum_i^{N_{data}} w(q^2_{i,reco}) \times q^{2n}_{i,calib}}{\sum_j^{N_{data}} w(q^2_{j,reco})} \times C_{calib} \times C_{gen}$$

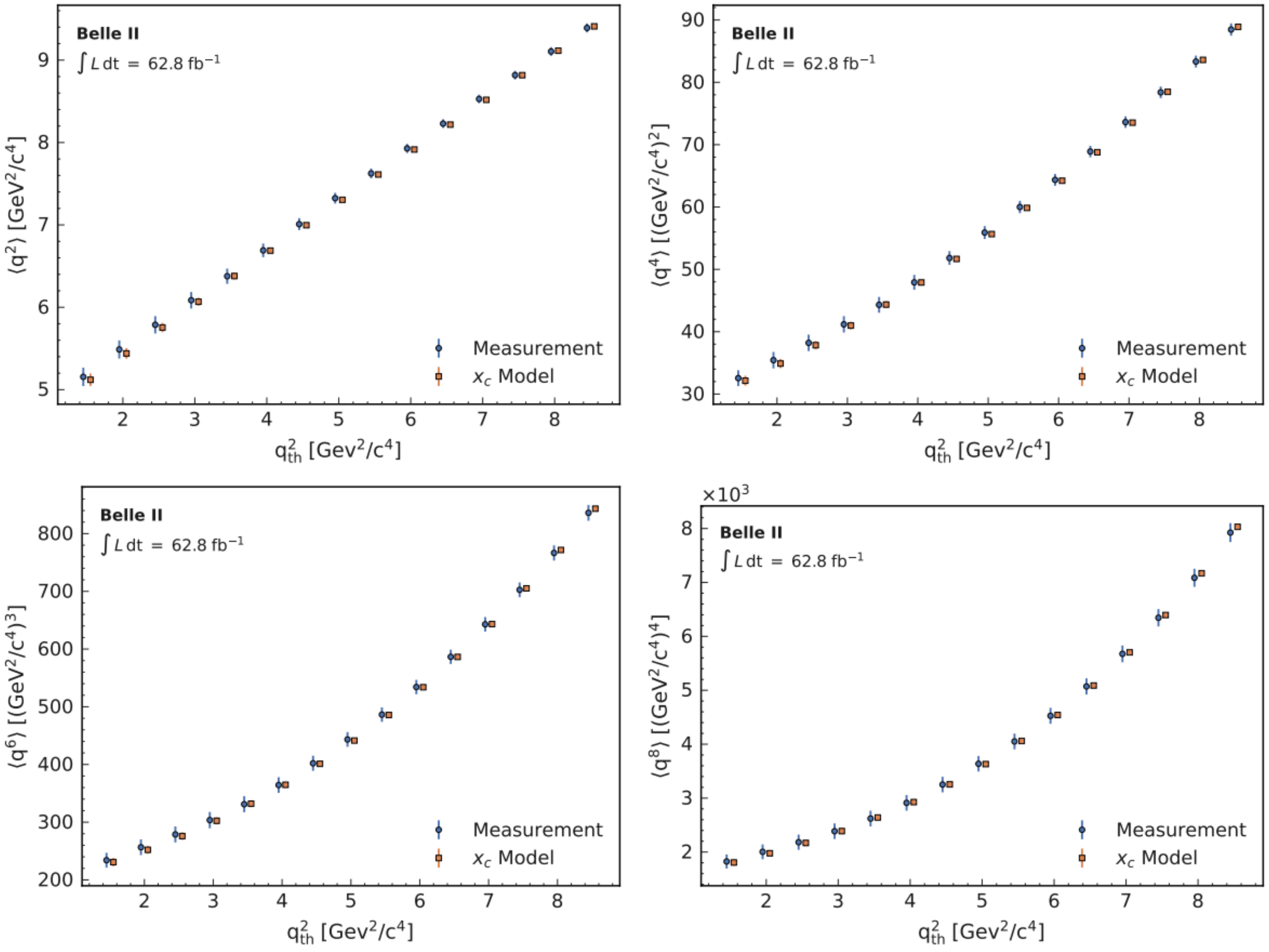
# Inclusive $B \to X_c \ell \nu \, \langle q^2 \rangle$

- Step 4: Correct for selection efficiencies
- Dominant effect:
  lepton reconstruction efficiency

$$C_{gen} = \langle q^{2n}_{gen}\rangle / \langle q^{2n}_{gen,sel}\rangle$$



# Inclusive $B \to X_c \ell \nu$ $\langle q^2 \rangle$

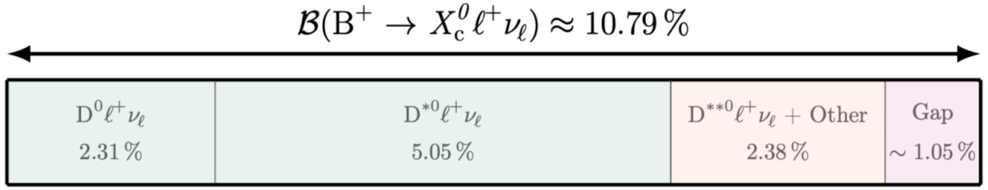


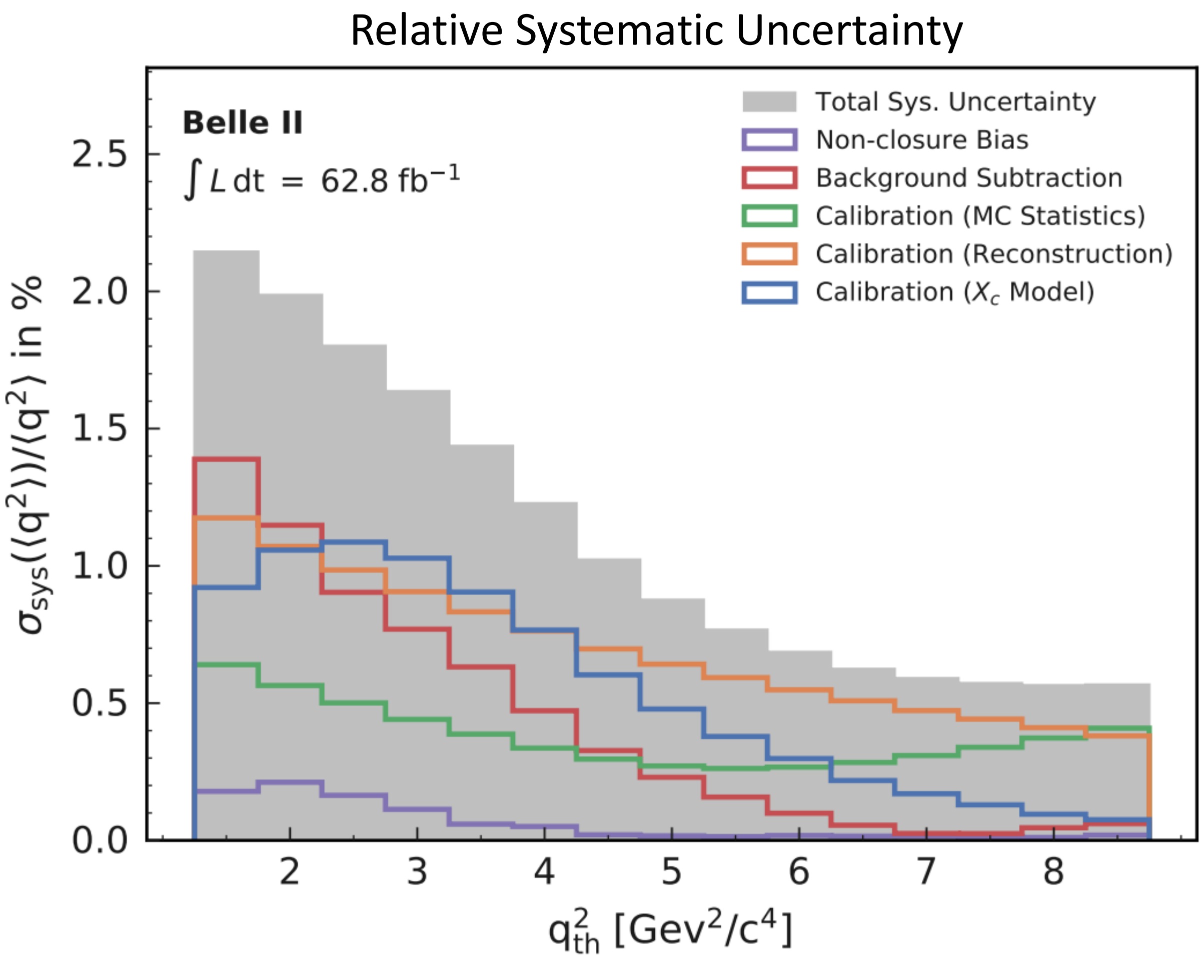

Perform analysis with different thresholds of $q^2$

… and extract $|V_{cb}|$

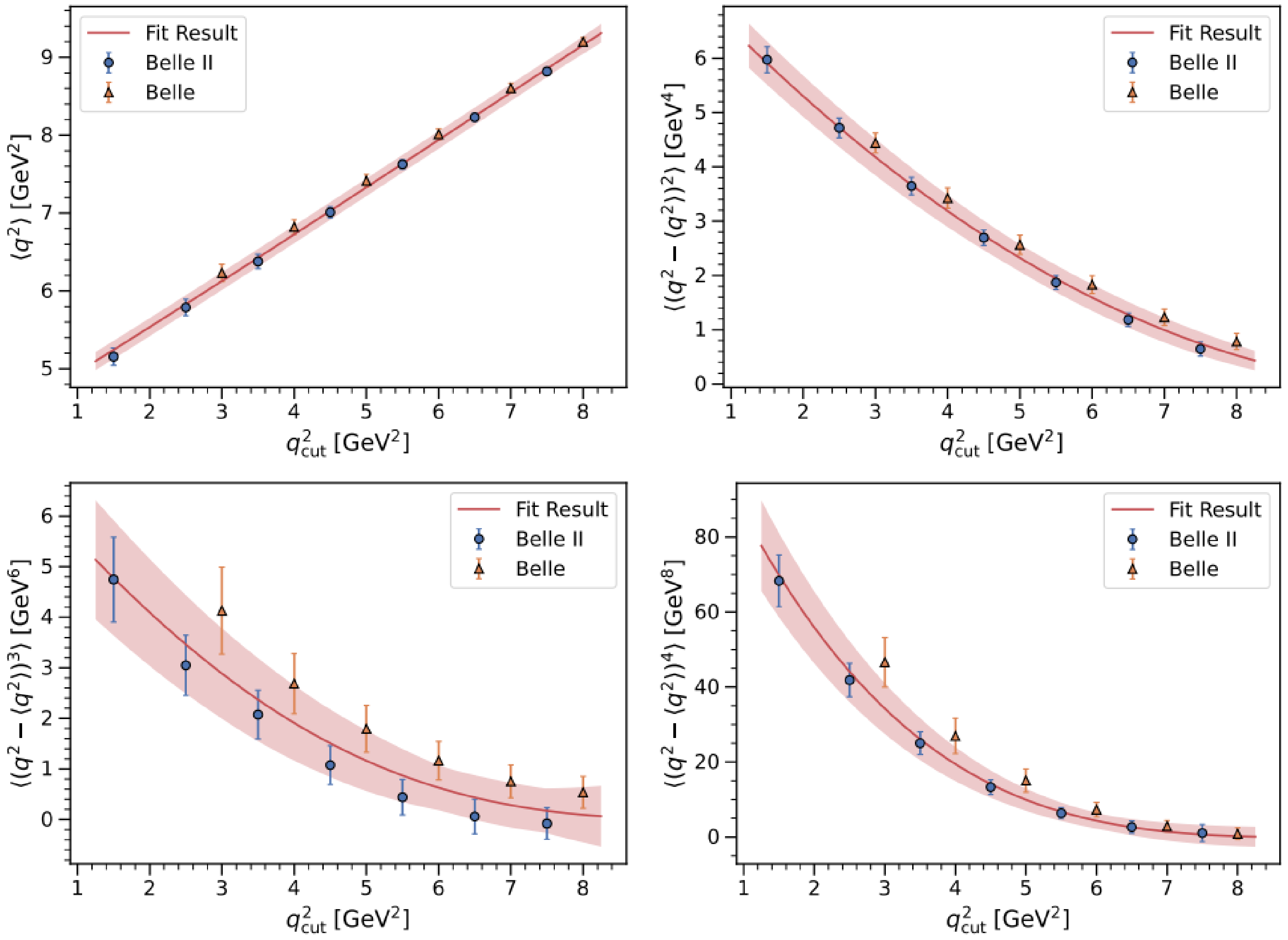
# Inclusive $B \to X_c \ell \nu \ \langle q^2 \rangle$

- Correlations can be extracted with bootstrapping
- Leading uncertainties are from
  - Reconstruction
  - Background subtraction
  - $X_c$ model



# $|V_{cb}|$ from Inclusive $B \to X_c \ell \nu$ $\langle q^2 \rangle$




F. Bernlochner, M. Fael, K. Olschwesky, E. Persson, R. Van Tonder, K. Vos, M. Welsch [*JHEP* 10 (2022) 068, [arXiv:2205.10274]

- Inclusive fit $\langle q^2 \rangle$
  $|V_{cb}| = 41.69$ with 1.5% precision!

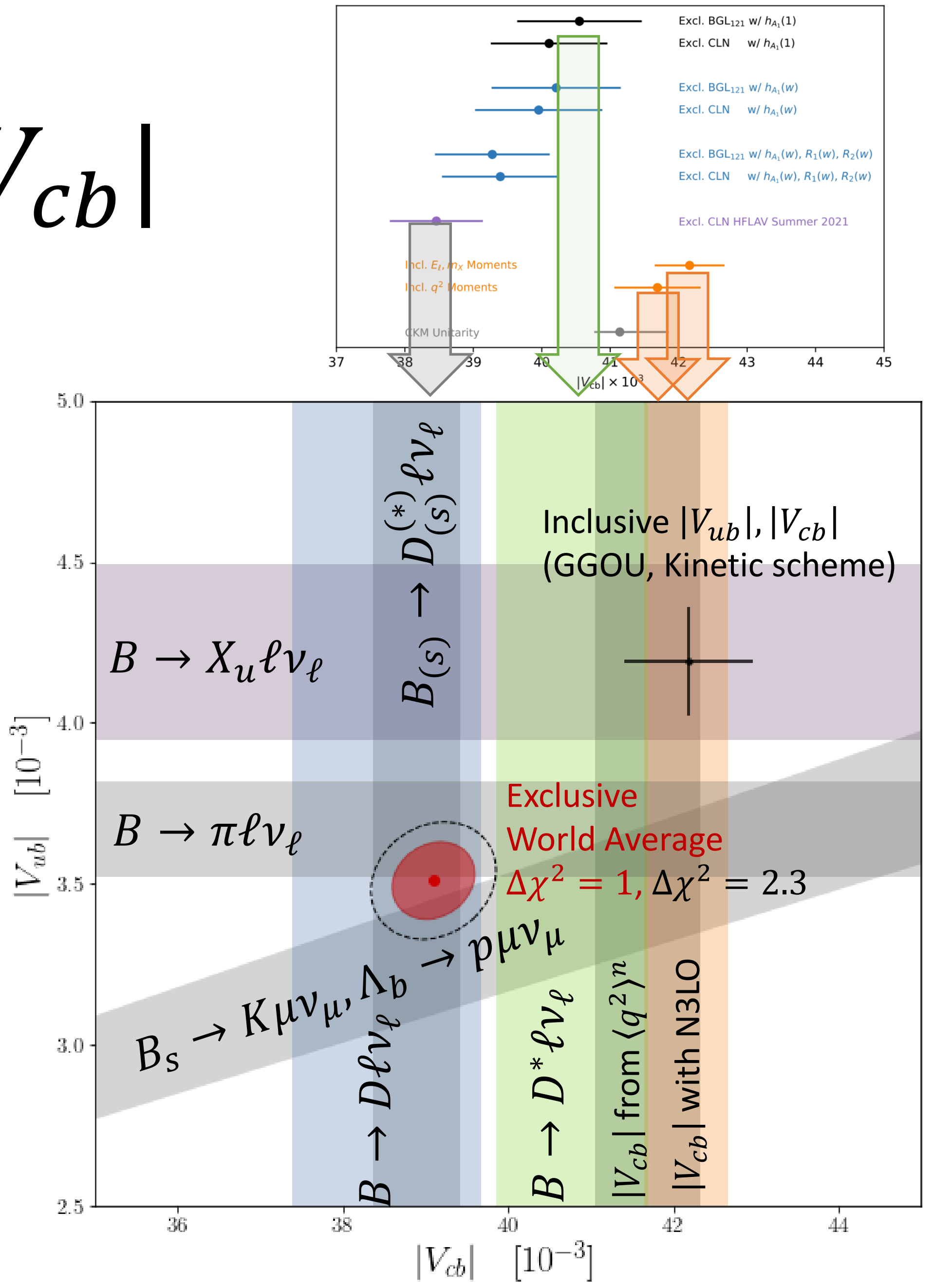
# Summary on $b \to c\ell\nu_\ell$ and $|V_{cb}|$

# Summary on $b \to c\ell\nu_\ell$ and $|V_{cb}|$

- Belle II can measure $b \to c\ell\nu_\ell$ transitions
  - exclusively (tagged and untagged)
  - inclusively (tagged)
- Different experimental techniques to recover the event kinematics
- Different theoretical frameworks to extract $|V_{cb}|$

# Summary on $b \to c\ell\nu_\ell$ and $|V_{cb}|$

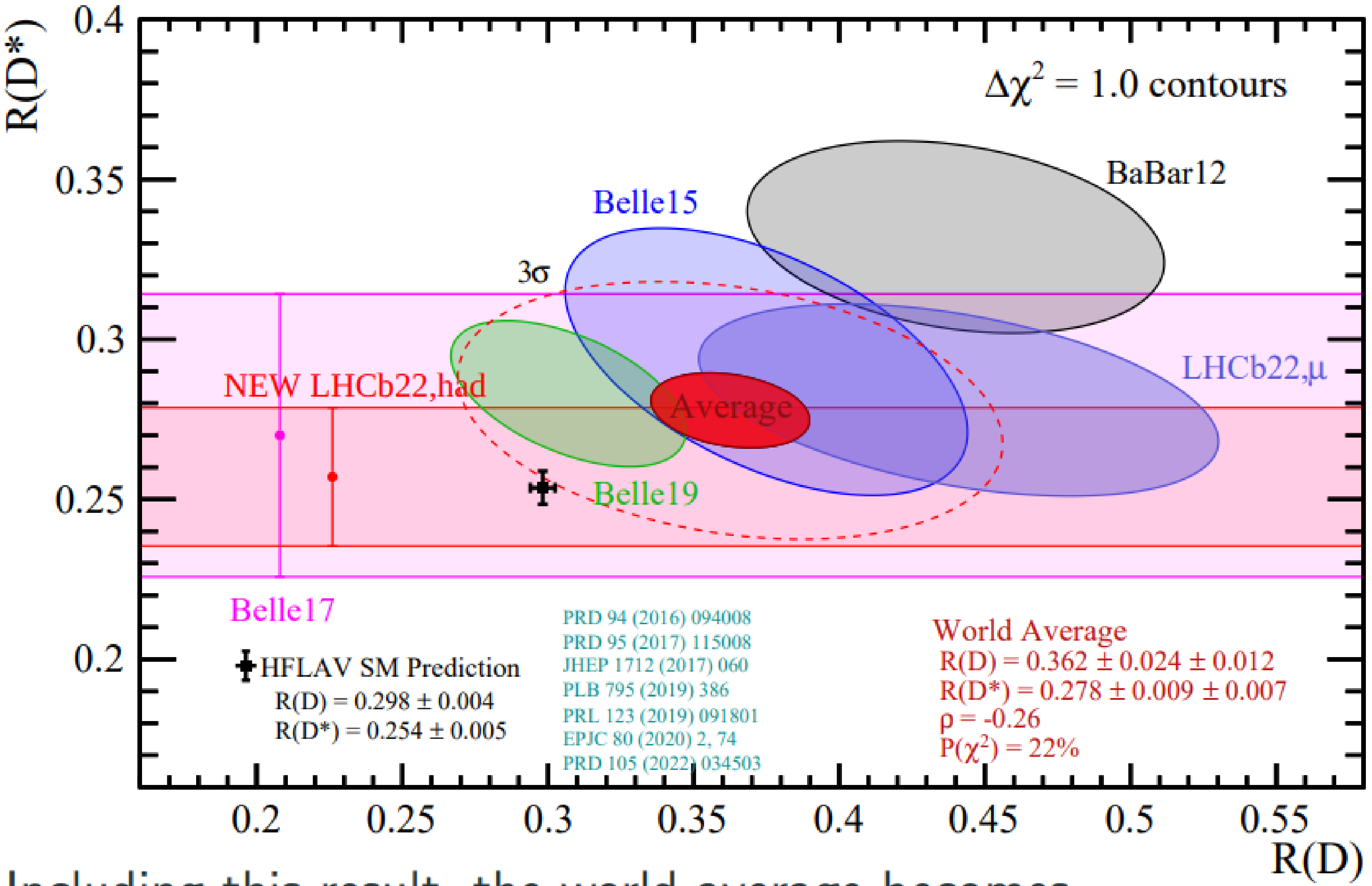
- Belle II can measure $b \to c\ell\nu_\ell$ transitions
  - exclusively (tagged and untagged)
  - inclusively (tagged)
- Different experimental techniques to recover the event kinematics
- Different theoretical frameworks to extract $|V_{cb}|$
- Different results!
  - This is a decade old tension and yet to be understood



## $R(D^{(*)})$

# $R(D^*)$ with hadronic $\tau$ decays

**NEW!**



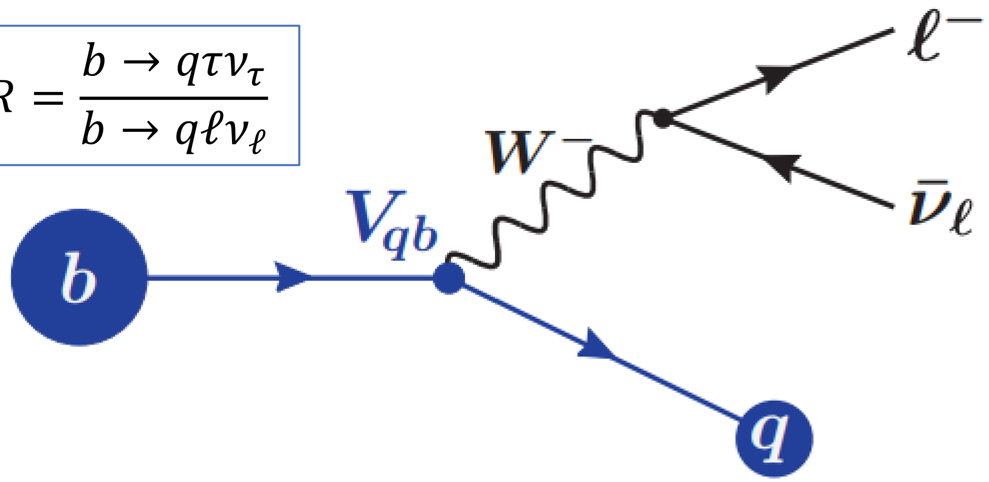

**HFLAV PRELIMINARY**

[LHCb-PAPER-2022-052]
(In preparation)

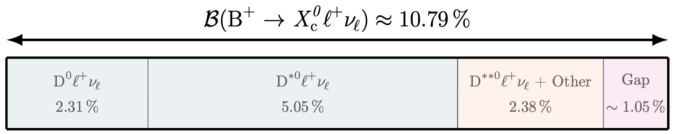
- Including this result, the world average becomes $R(D^*) = 0.278 \pm 0.011$; $R(D) = 0.362 \pm 0.027$
- The deviation w.r.t. the SM stays at **$3.0\sigma$** level for the combination of $R(D)$-$R(D^*)$

# Measurement Strategies

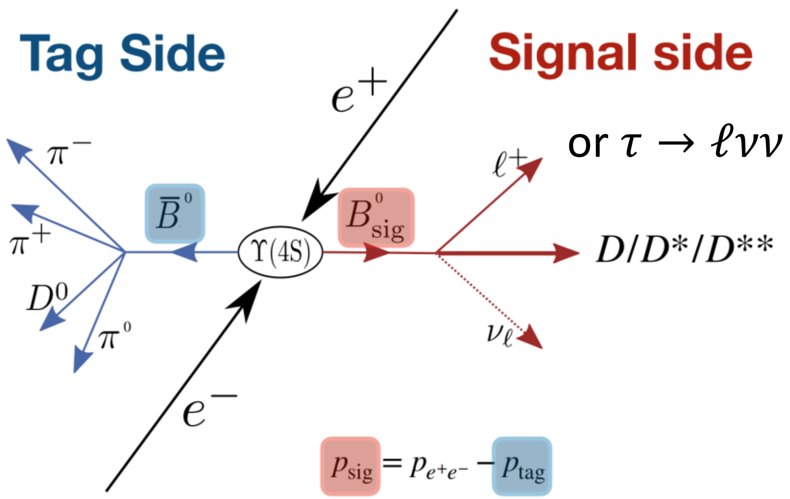
$$R = \frac{b \to q\tau\nu_\tau}{b \to q\ell\nu_\ell}$$




- Leptonic or hadronic $\tau$ decays?
  - Leptonic is cleaner (less background)
  - Hadronic allows to measure more properties (e.g., $\tau$ polarization)
- Exclusive or inclusive approach on the hadronic system?
  - $R(D^{(*)})$
  - $R(X)$ (challenging due to $X_c$ modelling)
- How to split signal from normalization?
  - Tagging, matching topology, kinematics



# Measurement Strategies




- $e^+e^- \to \Upsilon(4S) \to B\bar{B}$
  No additional particles in the event
- Fully reconstruct signal and tag side

→Each measured track/cluster has to be assigned

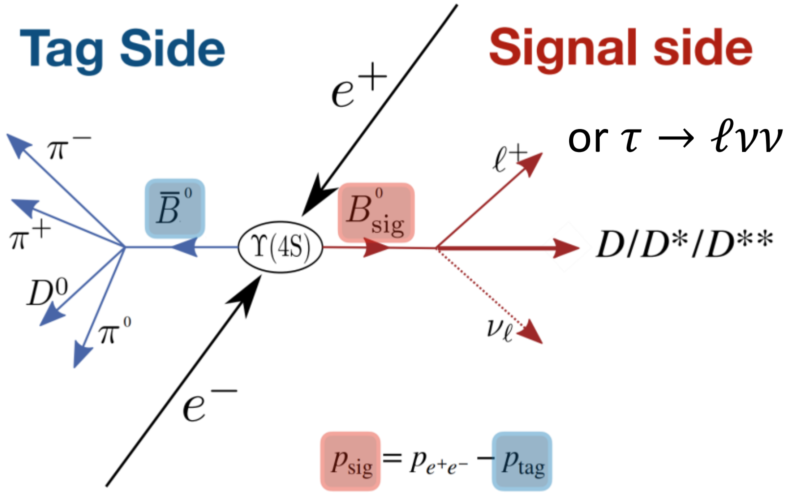
- Missing 4-momentum can be reconstructed

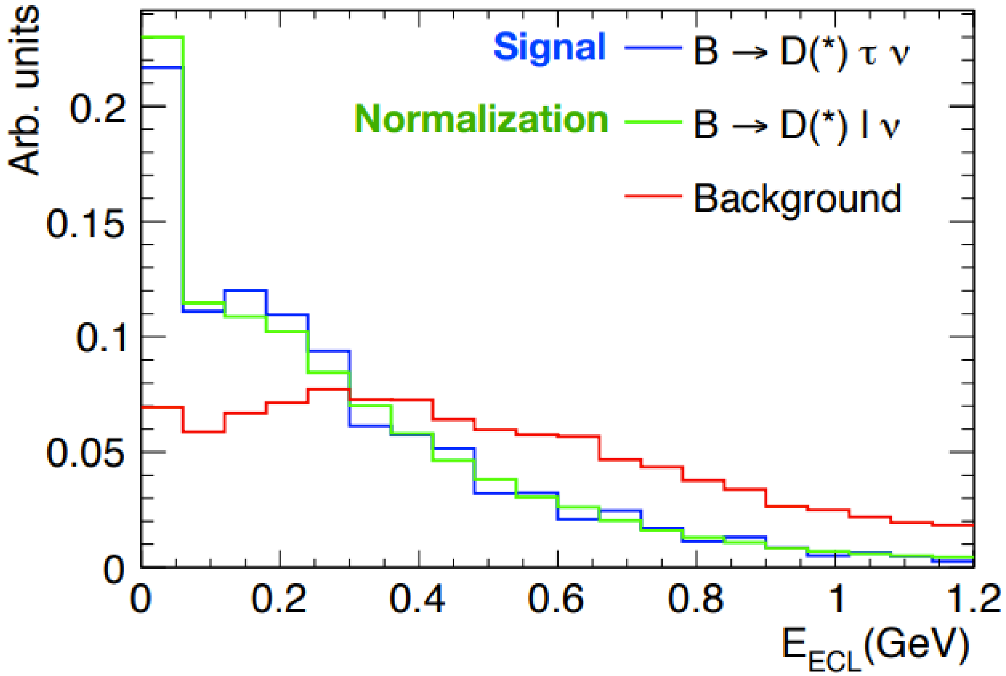
$$p_{miss} = (p_{beam} - p_{B_{tag}} - \mathrm{p}_{\mathrm{D}^{(*)}} - \mathrm{p}_\ell)$$

- Small tagging efficiency compensated by large data sample

# (one of) Belle's $R(D^{(*)})$

- with leptonic $\tau$ decays
- with semileptonic tagging
- Key variable: $E_{ECL} = \sum_i E_i^{\gamma} = E_{extra}$



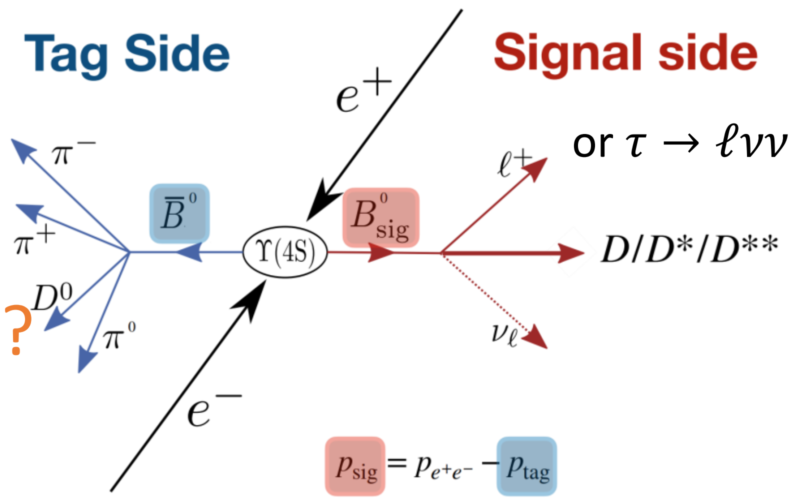



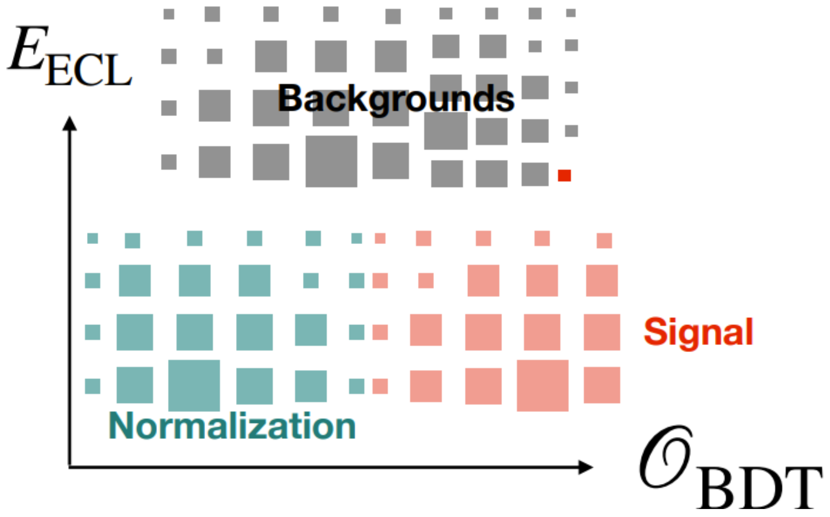

- Require no additional tracks in the event
- Signal and normalization peak at $E_{ECL} = 0$
- How to discriminate signal from normalization?

# (one of) Belle‘s $R(D^{(*)})$

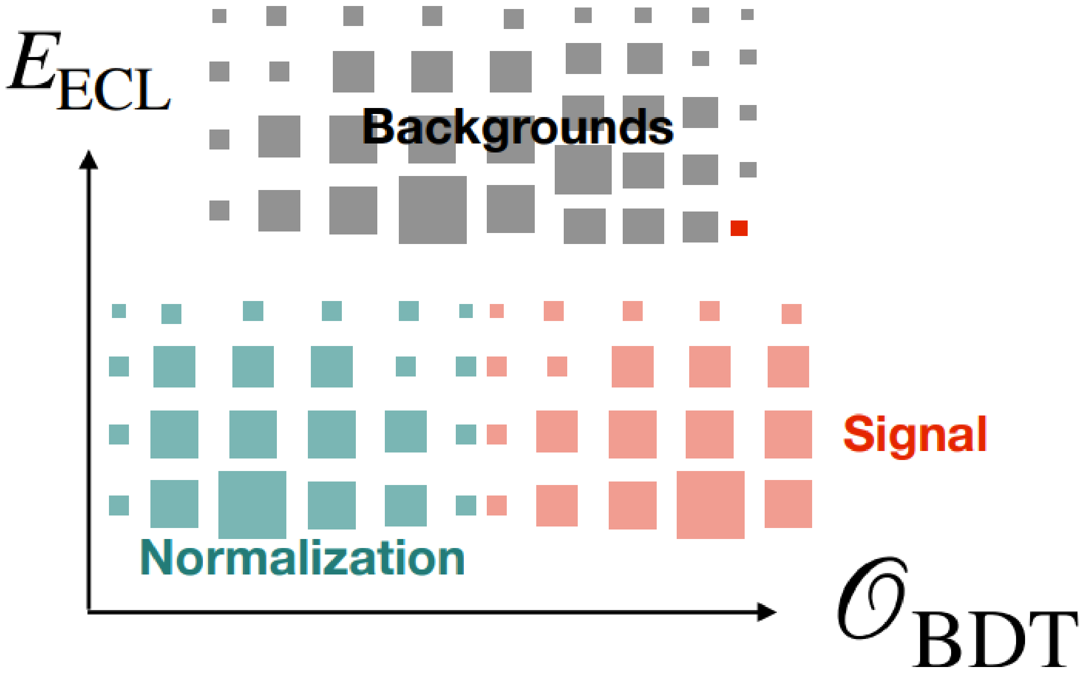
- How to discriminate signal from normalization?
- Use difference in event kinematics
  - $\cos\theta_{B,D^{(*)}\ell} = \frac{2E_B E_{D^{(*)}\ell} - m_B^2 - m_{D^{(*)}\ell}^2}{2|\vec{p}_B||\vec{p}_{D^{(*)}\ell}|}$
  - $\mathrm{M}_{\mathrm{miss}}^2 = (\mathrm{E}_{\mathrm{miss}}, \mathrm{p}_{\mathrm{miss}})^2 = (\mathrm{p_B} - \mathrm{p}_{\mathrm{D}^{(*)}} - \mathrm{p}_\ell)^2$
  - $E_{vis} = \sum_i E_i$ (visible Energy)

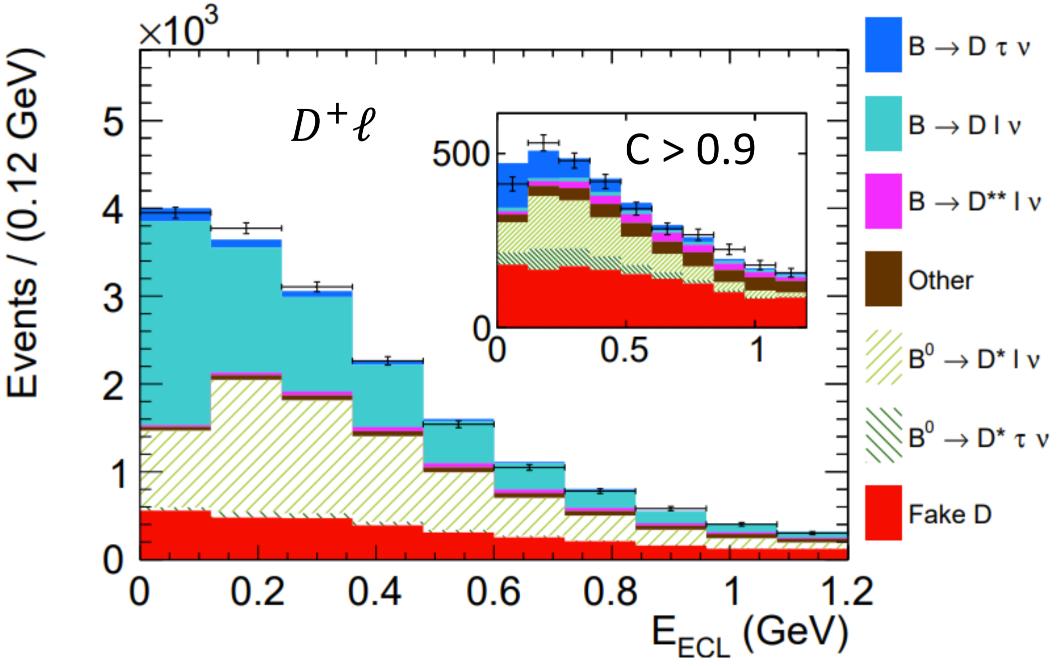
→ Construct a MVA classifier with these inputs



# (one of) Belle's $R(D^{(*)})$

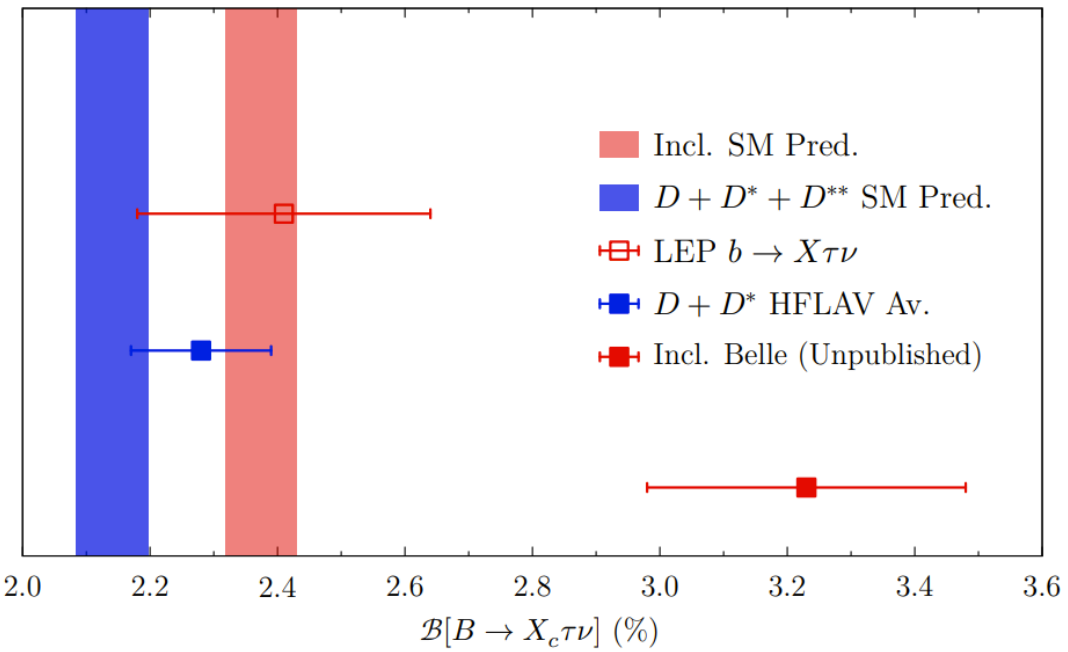


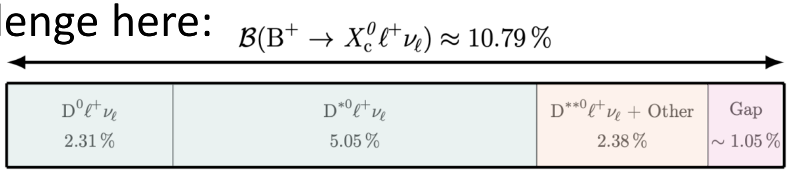

$$\mathcal{R}(D) = 0.307 \pm 0.037 \pm 0.016$$
$$\mathcal{R}(D^*) = 0.283 \pm 0.018 \pm 0.014,$$

# Other LFU Measurements at Belle (II)

**$R(X)$**




The challenge here:




**$R(\Upsilon)$**