Analysis presentation \( V_{cb} \)

Part 1: Overview and tag side

Florian Bernlochner\(^1\), Thomas Kuhr\(^2,3\), Kilian Lieret\(^2,3\), Felix Metzner\(^4\), Markus Prim\(^1\)

\(^1\)Universität Bonn
\(^2\)Ludwig-Maximilian University
\(^3\)Excellence Cluster Origins
\(^4\)Karlsruher Institut für Technologie

March 16, 2021
Overview

- Part 2 (Kilian): Tagging calibration for $|V_{cb}|$ measurement with $B \rightarrow D^* \ell \bar{\nu}_\ell$ (but the calibration looks at $B \rightarrow X \ell \nu$)
- Part 3 (Markus): $|V_{cb}|$ measurement with $B \rightarrow D^* \ell \bar{\nu}_\ell$

Caveats:
- Not a physics talk: We’ll cut corners and go on tangents about doing analysis
- I’m a bit out of the loop with recent $|V_{cb}|$ papers...
- Not a rockstar presentation :/
Part 1

1 Why?

2 How?
   - Measuring $|V_{cb}|$
   - Exclusive vs Inclusive Analyses
   - Tagging
Why $|V_{cb}|$?

- The CKM matrix $V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cb} & V_{cs} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}$ is unitary in the SM, i.e.

$$VV^\dagger = 1 = V^\dagger V$$

- Thus $\sum_k |V_{ik}|^2 = 1$ for all $k$ (weak unitarity) and
- $\sum_k V_{ik} V_{jk}^* = 0$ for all $i,j$
- Most popular: $V_{ub}^* V_{ud} + V_{cb}^* V_{cd} + V_{tb}^* V_{td} = 0$
- Powerful test of the SM
- Can also be visualized as a triangle “unitarity triangle”
Why $|V_{cb}|$

![Diagram showing the $\rho$-$\eta$ plane with various exclusions and solutions](attachment:image.png)
Measuring $|V_{cb}|$

- Pure leptonic: $B_c \rightarrow \ell \bar{\nu}$. Unavailable @ $B$-factories
- Pure hadronic: Hard theoretically
- Semileptonic:
  - Theory: Small EW corrections; QCD uncertainties under control
  - Experiment: Only one neutrino missing, good BRs ($\approx 10\%$)
  - $\Rightarrow$ Best opportunity to measure $|V_{cb}|$
    - Exclusively: Either $B \rightarrow D^* \ell \bar{\nu}_\ell$ (this analysis) or $B \rightarrow D \ell \bar{\nu}_\ell$
    - Inclusively: $B \rightarrow X_c \ell \nu$

Exclusive decays
Exclusive vs Inclusive Analyses

- **Tension** between inclusive \(B \rightarrow X_c \ell \nu\) and exclusive \(B \rightarrow D^{(*)} \ell \bar{\nu}_\ell\) determination of \(V_{cb} > 3\sigma\) (2018)

![inclusive vs exclusive analysis graph]

- Speculation about tension being related to FF parametrization in exclusive mode (CLN vs BGL)
How to extract $|V_{cb}|$ (exclusive)

- Say we want to extract $|V_{cb}|$ from $B \rightarrow D^* \ell \bar{\nu}_\ell$

$$\frac{d\Gamma(B \rightarrow D^* \ell \bar{\nu}_\ell)}{dw} = \text{stuff we know} \cdot \text{FF}^2(w) \cdot |V_{cb}|^2,$$

- $w = \vec{v}_B \vec{v}_D(\ast) = \frac{m_B^2 + m_{D(\ast)}^2 - q^2}{2m_B m_{D(\ast)}}, \quad q^2 = (p_B - p_{D(\ast)})^2$

- $\text{FF}^2(w)$: form factors
- In order to get $|V_{cb}|^2$ we need to know something about $\text{FF}^2(w)$
- Lattice QCD can only confidently tell us $\text{FF}^2(1)$ (but this corresponds to $q^2 = 0$ and we don't have data there)

- Only solution: Use parameterized model for $\text{FF}^2(w)$ (model dependency! CLN vs BGL etc.), fit parameters, extrapolate to $w = 1$ and then use lattice result to get $|V_{cb}|^2$

- In other words: need to fit form factor parameters $\Rightarrow$ measure kinematic distributions
- $B \rightarrow D \ell \bar{\nu}_\ell$ is similar
Tagged vs Untagged Analyses

**Tagged Analysis**

+ High purity
  - Low efficiency
    
    (0.3% @Belle → 0.55% @Belle II)
  
+ Get $B$ meson momentum from tag side!

**Untagged Analysis**

- Low purity
+ High efficiency
  
  (11% @Belle → 20% @Belle II)
Part 2: $|V_{cb}|$ from hadronically tagged $B \rightarrow D^* \ell \bar{\nu}_\ell$ at Belle: Tag side & tagging calibration
Goals

Reanalysis of hadronically tagged $B \rightarrow D^* \ell \bar{\nu}_\ell$ at Belle

- Previous measurement: [1702.01521] (preliminary)
- Extract $|V_{cb}|$, form factors, kinematic distributions (separately for $e/\mu$, potentially also investigating model dependencies)
- Belle dataset still exceeds Belle II dataset by large margin
- Expect around $2 \times$ efficiency when using Belle II software (FEI)

⇒ Use Belle II software on Belle data! (b2bii)
Improved hadronic tagging @Belle II

Hadronic tagging @Belle II:
Around 5000 channels!
2.5× efficiency!

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>MVA</th>
<th>Efficiency</th>
<th>Purity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belle v1</td>
<td>Cut</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Belle v3</td>
<td>Cut</td>
<td>0.1%</td>
<td>0.25%</td>
</tr>
<tr>
<td>Belle NB</td>
<td>NB</td>
<td>0.2%</td>
<td>0.25%</td>
</tr>
<tr>
<td>Belle II FEI</td>
<td>BDT</td>
<td>0.5%</td>
<td>0.25%</td>
</tr>
</tbody>
</table>

Calibration for excl. $B \rightarrow D^{(*)}l\bar{\nu}_l$
$B \rightarrow Xl\nu$ (systematics limited)

$B^+ \rightarrow D^0\pi^+$
$B^+ \rightarrow D^0\pi^+\pi^0$
$B^+ \rightarrow D^0\pi^+\pi^0\pi^0$
$B^+ \rightarrow D^0\pi^+\pi^0\pi^0\pi^0$
$B^+ \rightarrow D^0\pi^+\pi^0\pi^0\pi^0\pi^0$
$B^+ \rightarrow D_s^+D^0$
$B^+ \rightarrow D_s^+D^0$
$B^+ \rightarrow D_s^+D^0$
$B^+ \rightarrow D_s^+D^0$
$B^+ \rightarrow D_s^+D^0$
$B^+ \rightarrow D_s^+D^0$
$B^+ \rightarrow D_s^+D^0$
$B^+ \rightarrow J/\psi K^+$
$B^+ \rightarrow J/\psi K^+$
$B^+ \rightarrow J/\psi K^+$
$B^+ \rightarrow J/\psi K^+$
$B^+ \rightarrow J/\psi K^+$
$B^+ \rightarrow J/\psi K^+$
$B^+ \rightarrow J/\psi K^+$
$B^+ \rightarrow J/\psi K^+$

$B^0 \rightarrow D^0\pi^+$
$B^0 \rightarrow D^0\pi^+\pi^0$
$B^0 \rightarrow D^0\pi^+\pi^0\pi^0$
$B^0 \rightarrow D^0\pi^+\pi^0\pi^0\pi^0$
$B^0 \rightarrow D^0\pi^+\pi^0\pi^0\pi^0\pi^0$

$D^+ \rightarrow D^0\pi^+\pi^0$
$D^+ \rightarrow D^0\pi^+\pi^0\pi^0$
$D^+ \rightarrow D^0\pi^+\pi^0\pi^0\pi^0$
$D^+ \rightarrow D^0\pi^+\pi^0\pi^0\pi^0\pi^0$
$D^+ \rightarrow D^0\pi^+\pi^0\pi^0\pi^0\pi^0\pi^0$

$D^+ \rightarrow D^0\pi^+\pi^0\pi^0$
$D^+ \rightarrow D^0\pi^+\pi^0\pi^0\pi^0$
$D^+ \rightarrow D^0\pi^+\pi^0\pi^0\pi^0\pi^0$
$D^+ \rightarrow D^0\pi^+\pi^0\pi^0\pi^0\pi^0\pi^0$

$D^+ \rightarrow D^0\pi^+\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0$

$D^+ \rightarrow D^0\pi^+\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0$

$D^+ \rightarrow D^0\pi^+\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0$

$D^+ \rightarrow D^0\pi^+\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0$

$D^+ \rightarrow D^0\pi^+\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0$

$D^+ \rightarrow D^0\pi^+\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0$

$D^+ \rightarrow D^0\pi^+\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0$

$D^+ \rightarrow D^0\pi^+\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0\pi^0$

$D^+ \rightarrow D^0\pi^+\pi^0$
Calibration of hadronic tagging

Tagging

- The FEI returns a list of $B_{\text{tag}}$ candidates with a likelihood and the decay channel.
- Often we need to know the efficiency of this selection.
- Unfortunately this efficiency can be different between MC and data $\Longrightarrow$ calibration!

How?

- Reconstruct $\Upsilon(4S) \rightarrow B_{\text{tag}}B_{\text{sig}}$ on MC and data.
- Idea: Find $B_{\text{sig}}$ channel with large BF that we trust.
- Compare yields and blame all differences on $B_{\text{tag}}$ reconstruction, i.e. the FEI.
- Here: Inclusive $B_{\text{sig}} \rightarrow X\ell\nu$.
- Need to do this in bins of the decay channel and the likelihood output.
Dataset & MC fixes

The dataset

- Full Belle (1!) dataset \((710 \text{ fb}^{-1}, 771 \times 10^6 B\bar{B}) + 1 \text{ stream of MC}\)
- In total there are 10 streams of Belle MC \((1 \text{ stream} = \text{MC equivalent of total data taken})\)
- “Generic” MC types: Charged \((B^+B^-)\), Mixed \((B^0\bar{B}^0)\), Charm \((c\bar{c})\), UDS \((q\bar{q})\)
- There is also special/rare/signal MC for special purposes

MC updates

- Belle MC was produced long ago \(\implies\) need to update to match better understanding
- Most simple updates: branching fractions (BFs)
- Only need to do this for most relevant BF: \(B \rightarrow D(*)\)
In practice: MC reweighting

Simplest case

- Let’s say our MC was generated with $BF(B \rightarrow XYZ) = 0.5$ (this is written in the decay file), but it’s actually 0.25!
- Do we need to regenerate our MC? 😱
- Easy fix: Simply add weight of 0.5 to all $B \rightarrow XYZ$ events

Uncertainties

- What if we only know $BF(B \rightarrow XYZ) = 0.5 \pm 0.1$?
- Uncertainty needs to be propagated to later results
- Add varied weight columns, i.e. `__weight__`, `__weight_xyz_up__`, `__weight_xyz_down__`
- Either repeat analysis for different weights or have other way to incorporate them (→ fitting)

Form factors (FF)

- Can also set weights to adjust *differential* cross sections! → Update FFs and more!
Dataset & MC fixes

- Lepton PID corrections
- Updated BFs for $B \rightarrow D^{(*)}$, FF reweighted for $B \rightarrow D, D^*, D^{**}$
- Generate dedicated $D^{**}$ samples (family of poorly known particles)
- Big issue: The gap
  - Sum up all exclusive $B \rightarrow X\ell\nu$ modes in MC, e.g. $B \rightarrow D^{(*)}\ell\bar{\nu}_\ell, \ldots$
  - Much smaller than direct measurement of $B \rightarrow X\ell\nu \rightarrow$ Gap between sum of exclusive and inclusive
  - We need to fill this gap somehow $\rightarrow$ educated guesses + large uncertainties
- Update Belle gap filling with dedicated samples $D_1(\rightarrow D_{\pi\pi})\ell^+\nu_\ell$, $D^*\eta\ell^+\nu_\ell$ and $D^{(*)}_{\pi\pi}\ell^+\nu_\ell$ (currently 100% scaling uncertainty for total gap component)
In practice: Generating signal MC

- You normally shouldn’t generate MC yourself → data production liaison
- MC generation is governed by decay files → read them to understand your MC!
- There’s a online book lesson about it

Define dm 0.507e12

Alias myB0 B0
Alias myAnti-B0 anti-B0
ChargeConj myB0 myAnti-B0

Decay Upsilon(4S)
1.000000 B0 anti-B0 myB0 myAnti-B0 VSS_BMIX dm;
Enddecay

Alias myD_1- D_1-
Alias myD_1+ D_1+
ChargeConj myD_1- myD_1+

Decay myB0
1.000 myD_1- e+ nu_e PHOTOS ISGW2;
Enddecay
CDecay myAnti-B0

Decay myD_1-
1.000 anti-D*0 pi- PHOTOS VVS_PWAVE 0.0 0.0 0.0 0.0 1.0 0.0;
Enddecay
CDecay myD_1+

End
In practice: Cuts & Best candidate selection

- Cuts & best candidate selection drastically reduce memory/disk requirements
- **Dilemma**: Do them “online” (using basf2) or “offline” (e.g. pandas)
- **Online**: Change cut value needs reprocessing with basf2 to (takes a lot of time)
- **Offline**: Dealing with a lot of data (memory & disk space issues for Belle dataset – Belle 2 data is smaller, so you might be fine) ⟷ consider looking into dask and friends for an intermediate step
- Makes sense to use soft online cuts and harder offline cuts
- Similar problem: Which **columns** to write out? ⟷ columnar data formats might at least help with fast reading of partial column set
- No golden rule, but **deliberate choices** are better
Example fits

$$B_{\text{sig}} \rightarrow X\ell\nu, \ B_{\text{tag}} \rightarrow \bar{D}^0\pi^+\pi^+\pi^-$$

Before fitting

Fitted
In practice: Fitting frameworks

- Many fitting frameworks: ROOFit, zFit, pyHF, ... → choose wisely!
  - Vary a lot in usability and quality of documentation!
  - Good to think ahead! (Features? How are uncertainties handled? Integration in your code? Help?)

- Different ways to incorporate uncertainties: Either
  - fits with fixed templates/weights → repeat fits with varied templates/weights → collect results and calculate uncertainty from differences
  - have more complicated likelihood where uncertainties are nuisance parameters

\[
\mathcal{L} = \prod_i \mathcal{P}(n_i; \nu_i) \times \prod_k \mathcal{G}_k(\vec{0}, \vec{\theta}_k, \Sigma_k)
\]

\[
\nu_i = \sum_k \eta_k \prod_s (1 + \tilde{\theta}_{ks}) \frac{\eta_{ik}(1 + \theta_{ik})}{\sum_j \eta_{jk}(1 + \theta_{jk})},
\]
The calibration factors
CFs and uncertainties
In practice: Validating fits

How do we know to trust our fits?

- Expect to repeat fits very, very, often...
- Simple validation: Linearity check
  - Generate artificial dataset from MC “bootstrapping” with different signal strengths
  - Fit this
  - Expect to get back signal strength that we put in
- Can also look at $p$-value distribution (should be flat!)

![Graph showing $p$-value distribution](image-url)
| $V_{cb}$ |

- $|V_{cb}|$ needed to test unitarity of CKM matrix
- $\text{BR}(B \rightarrow D^{(*)}\ell\bar{\nu}_\ell) = \text{stuff we know} \cdot \text{FF}^2(w) \cdot |V_{cb}|^2 \rightarrow \text{exclusive measurement}$
- Form factors $\text{FF}^2(w)$ make trouble
- Other method of extraction: Inclusive measurement; in tension with exclusive measurements

This analysis: $|V_{cb}|$ from hadronically tagged $B \rightarrow D^*\ell\bar{\nu}_\ell$ at Belle

- $B_{\text{tag}}$ reconstructed by the FEI
- The FEI has different efficiencies on data and MC $\rightarrow$ Calibration
- For calibration need something we trust: $B_{\text{sig}} \rightarrow X\ell\nu$
- Reconstruct $\Upsilon(4S) \rightarrow B_{\text{tag}}B_{\text{sig}}$; blame efficiency difference on FEI $\rightarrow$ calibration factors
Backup
Cuts & best candidate selection

Electrons, Muons

\[
\begin{array}{c|c}
\text{e} & \mu \\
\hline
dr < 2 \text{ cm} & \\
|dz| < 4 \text{ cm} & \\
|p_T^*| > 0.1 \text{ GeV/c} & \\
p > 0.3 \text{ GeV/c} & p > 0.6 \text{ GeV/c} \\
eIDBelle > 0.6 & muIDBelle > 0.9 \\
\text{muIDBelleQuality} = 1 &
\end{array}
\]

Photon

goodBelleGamma = 1

Event

Number of tracks with \( dr < 2 \text{ cm} \) and
\(|dz| < 4 \text{ cm} \): \( \leq 17 \)
Number of FS photons: \( < 18 \)

\( B_{\text{tag}} \)

\( M_{bc} > 5.27 \text{ GeV/c}^2 \)
signal probability \( > 0.001 \)
CS output \( > 0.2 \) (CS via BDT)

Best Candidate Selection

- Pick \( B_{\text{tag}} \) with smallest \( \Delta E \)
- Pick lepton with highest \( |p_{\ell}^*| \)
- Still multiple candidates \( \rightarrow \) pick random
Fit strategies

- Decay channels vary significantly in number of events $\implies$ number of bins in FEI classifier depends on decay channel (currently aiming for $> 4500$ events per bin)
- Perform a fit to extract yields for signal
- Calibration factor $= \frac{\text{fitted yield}}{\text{MC yield}}$
- Binned maximum likelihood fit (systematics included in likelihood $\implies$ pull on the shape of the distributions $\implies$ Backup)

Current strategy: 2 component fit

- Fake leptons + secondary leptons + continuum are fitted as one component (background) $\implies$ Requires additional systematic uncertainty $\implies$ todo
- Fitted with 2 components: Signal + background
- Number of bins fixed to 9 (will be made dynamic later)
Fitting and systematics

Maximize binned likelihood, allowing systematics to pull on the shapes:

\[
\mathcal{L} = \prod_i P(n_i; \nu_i) \times \prod_k G_k(\tilde{\theta}, \theta, \Sigma_k)
\]

\[
\nu_i = \sum_k \eta_k \prod_s (1 + \tilde{\theta}_{ks}) \frac{\eta_{ik}(1 + \theta_{ik})}{\sum_j \eta_{jk}(1 + \theta_{jk})},
\]

where

- \(\mathcal{P}\): Poisson distribution
- \(n_i\): Measured events in bin \(i\)
- \(\nu_i\): Expected events in bin \(i\)
- \(k\): Component of fit
- \(\eta_k\): Yield of fit template \(k\)
- \(\eta_{ik}\): Bin content of fit template \(k\), bin \(i\)
- \(\theta_{ik}\): Additive nuisance parameter bin \(i\), fit component \(k\)
- \(\tilde{\theta}_{is}\): Multiplicative nuisance parameter from source \(s\) for fit component \(k\)
- \(G\): Multivariate Gaussian with covariance matrix \(\Sigma_k\)

Fitted: \(\eta_k, \theta_{ik}, \tilde{\theta}_{ks}\)
Reconstruction signal side

π, K meson
Identification via PID likelihood ratio, impact parameters.
π^0: From γ candidates (clusters in calorimeter not matched to any track)

<table>
<thead>
<tr>
<th></th>
<th>arXiv</th>
<th>Signal</th>
<th>Tag</th>
<th>D^0 modes</th>
<th>D^+ modes</th>
<th>D^*- modes</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1010.5620]</td>
<td>D^*ℓν_ℓ</td>
<td>No</td>
<td>K^-π^+</td>
<td></td>
<td></td>
<td>D^-π^-</td>
</tr>
<tr>
<td>[1809.03290]</td>
<td>D^*ℓν_ℓ</td>
<td>Hadr.</td>
<td>K^-π^+(π)(π)</td>
<td>K^-π^+π^+</td>
<td></td>
<td>D^0π^-, D^-π^0</td>
</tr>
<tr>
<td>[1702.01521]</td>
<td>D^*ℓν_ℓ</td>
<td>Hadr.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[1510.03657]</td>
<td>Dℓν_ℓ</td>
<td>Hadr.</td>
<td>K^-π^+(π)(π), K^0π^+π^-(π^0), K^0π^0, K^+K^-, π^+π^-(π^0), K^0K^0, π^0π^0, K^0π^0, K^-π^+π^+π^-</td>
<td>K^-π^+π^+(π^0), K^0π^+π^+π^-, π^+π^+(π)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Results and Outlook @Belle

<table>
<thead>
<tr>
<th>Link</th>
<th>Channel</th>
<th>Tag</th>
<th>CLN</th>
<th>BGL</th>
<th>Unfold</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1010.5620]</td>
<td>$D^* \ell \bar{\nu}_\ell$</td>
<td>No</td>
<td>38.4 ± 0.9</td>
<td>42.5 ± 1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[1809.03290]</td>
<td>$D^* \ell \bar{\nu}_\ell$</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[1702.01521]</td>
<td>$D^* \ell \bar{\nu}_\ell$</td>
<td>Had.</td>
<td>37.4 ± 1.3</td>
<td></td>
<td></td>
<td>Prelim. Soon: Separate results $\ell = e$ and $\ell = \mu$</td>
</tr>
<tr>
<td>[1510.03657]</td>
<td>$D \ell \bar{\nu}_\ell$</td>
<td>Had.</td>
<td>39.86 ± 1.33</td>
<td>40.83 ± 1.13</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

HFLAV
Summer 2016
/dof = 30.2/23 (CL = 14.40 %)

HFLAV
Summer 2016
/dof = 4.7/8 (CL = 79.30 %)

Kilian Lieret
Ludwig-Maximilian University
Projection in bins of kinematic variable

\[ B \rightarrow D\ell \bar{\nu}_\ell \]

10 equal-size bins in \( w \).
Good resolution (0.005) vs bin width (0.06) \( \Rightarrow \) Bin migration neglected

\[ B \rightarrow D^*\ell\bar{\nu}_\ell \]

- 10 equal size bins in \( w, \chi, \cos \theta_\ell, \cos \theta_{D^*} \) (Projections)
- Correlation between the 4 distributions (\( \rightarrow \) toy experiments)
- Finite resolution \( \Rightarrow \) Migration!
  \( \rightarrow \) Mig. matrix from truth vs reco MC.
  \( \rightarrow \) Fold theory (easy) or unfold measurement (hard)
Fit variables

Tagged analyses

Fit variable: \( m_{\text{miss}}^2 := (p_B - p_{D(*)} - p_\ell)^2 \) with \( p_B = p_{\text{LER}} + p_{\text{HER}} - p_{\text{tag}} \)

Correct reco \( \Rightarrow \) Peak at 0; Missed particles \( \Rightarrow \) Peak > 0; Particles from tag side \( \Rightarrow \) Peak < 0

Untagged analyses

Fit variables:

\[
\cos \theta_{B,D^*\ell} := \frac{2E_B^*E_{D^*\ell} - m_B^2 - m_{D^*\ell}^2}{2|p_B^*||p_{D^*\ell}^*|}
\]

Correct reco \( \Rightarrow \) \(-1 \leq \cos \theta_{B,D^*\ell} \leq 1\)

\( \Delta m = m_{D^*} - m_D \)

\( p_\ell \)
Fit strategies

Belle tagged $B \rightarrow D \ell \bar{\nu}_\ell$ [1510.03657]

$B = B^0$, $\ell = e$

Fit

Binned extended likelihood fit (Barlow, Beeston, 1993)

Fit variable

$m^2_{\text{miss}} := (p_B - p_D - p_\ell)^2$

Templates

From MC

Fixed Norm.

“other” background from MC

Float. Norm.

2: $B \rightarrow D \ell \bar{\nu}_\ell$ and $B \rightarrow D^* \ell \bar{\nu}_\ell$ normalization
Fit strategies

Belle tagged $B \rightarrow D^* \ell \bar{\nu}_\ell$ [1702.01521]

Fit
- Unbinned likelihood fit

Fit variable
- $m_{\text{miss}}^2 := (p_B - p_{D^*} - p_\ell)^2$

Templates
- From MC

Fixed Norm.
- Ratios of background normalizations from MC

Float. Norm.
- 2: Correctly reco. sig. + sig. with $D^{(*)}$ wrongly reco. norm.; total background normalization
Fit strategies

Belle untagged $B \rightarrow D^* \ell \bar{\nu}_\ell$ $[1010.5620]$

Fit
Binned likelihood fit

Fit variable
$\cos \theta_{B,D^*\ell}$, $\Delta m$, $p_\ell$

Templates
Continuum from off-resonance, rest from MC\(^1\)

Fixed Norm.
Continuum from off-resonance (corrected for $1/s$ dependency)

Float. Norm.
6: Normalizations for signal and backgrounds

\(^1\)For $\ell = \mu$: Shape of fake $\ell$ corr. with data from $K_s^0 \rightarrow \pi^+ \pi^-$; $\ell$ PID eff. corr. with data from $2\gamma \rightarrow e^+ e^- / \mu^+ \mu^-$
**Fit strategies**

Belle untagged $B \rightarrow D^* \ell \bar{\nu}_\ell$ (new) $^{[1809.03290]}$

**Fit**
- Binned likelihood fit

**Fit variable**
- $\cos \theta_{B,D^*\ell}$, $\Delta m$, $p_\ell$

**Templates**
- Continuum from off-resonance, rest from MC$^2$

**Fixed Norm.**
- Continuum from off-resonance (corrected for $1/s$ dependency, kinematics)

**Float. Norm.**
- 6: Normalizations for signal and backgrounds

$^2$Shape of fake $\ell$ corr. with data from $D^* \rightarrow D^0 \pi$, $D^0 \rightarrow K \pi$; lepton PID eff. corr. with data from $ee \rightarrow ee\gamma$, $ee \rightarrow \mu\mu(\gamma)$ and $J/\psi \rightarrow \ell^+\ell^-$; low momentum track reco. eff. corr. with control sample of $B \rightarrow D^* \ell \bar{\nu}_\ell$
$V_{cb}$ and form factors fit

$\chi^2$ fit

$$\chi^2 = \left( \bar{\nu}_{\text{sig}} - \bar{\nu}_{\text{sig}}^{\text{pred}} \right) C^{-1} \left( \bar{\nu}_{\text{sig}} - \bar{\nu}_{\text{sig}}^{\text{pred}} \right) + \chi^2_{\text{nuisance}},$$

where:

- $\bar{\nu}_{\text{sig}}$ yields in bins of kinematic variables $(D\ell\bar{\nu}_{\ell}: w, D^*\ell\bar{\nu}_{\ell}: w, \chi, \cos \theta_{\ell}, \cos \theta_{D^*})$
- $\bar{\nu}_{\text{sig}}^{\text{pred}} = (\epsilon_{\text{reco}}\epsilon_{\text{tag}}) M_{\text{mig}} \Delta \Gamma$
  $\Delta \Gamma_i$: theory expectation diff. CS in bin $i$
  (depends on FF param, $|V_{cb}|$),
  $M_{\text{mig}}$: migration matrix
- $C$: Covariance matrix
- $\chi^2_{\text{nuisance}}$: Account for multiplicative factors degenerate with $|V_{cb}|$