Milano Bicocca University

TEST OF LEPTON FLAVOUR UNIVERSALITY IN SEMILEPTONIC DECAYS AT THE LHCB EXPERIMENT

Second Year PhD final seminar, Cycle XXXIV

September 3, 2019

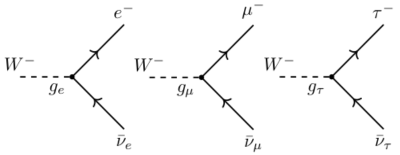
Supervisor: Prof. Marta Calvi

Simone Meloni, 763674
s.meloni1@campus.unimib.it

Introduction and motivations of the analysis

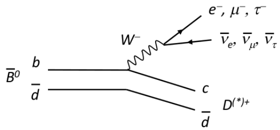
Lepton Flavour Universality in the SM

- Lepton Flavour Universality (LFU) is an accidental symmetry in the Standard Model



- Equality of the couplings of gauge bosons to leptons ($g_e = g_\mu = g_\tau$)
- LFU can be violated in New Physics Models with mass dependent coupling

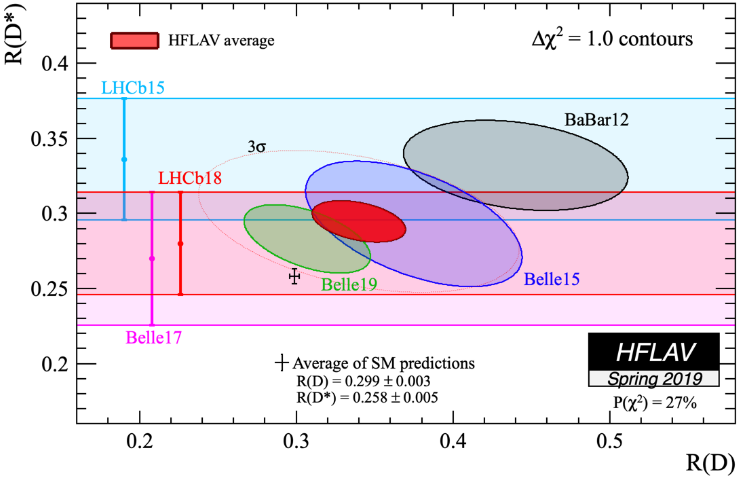
- The hypothesis can be tested in $b \to c l \nu$
  - Relatively simple description in the Standard Model via Tree Level Processes
  - High Transition rate



- Differences for decays with e, μ, τ should originate only from mass differences
- Test variables are ratios of Branching Fractions

$$\mathcal{R}(D^{(*)}) = \frac{\mathcal{B}(B \to D^{(*)}\tau\nu)}{\mathcal{B}(B \to D^{(*)}\mu\nu)}$$

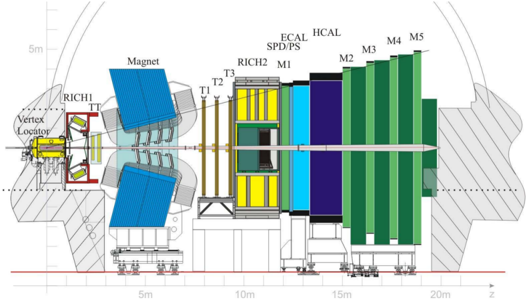
Previous measurements



- Various measurements of $\mathcal{R}(D^{(*)})$ combined
- Tension at the 3.1 σ level wrt SM predictions
- No measurement of $\mathcal{R}(D^{+,0})$ performed at an hadron collider so far.

The LHCb experiment

- LHCb is a single arm spectrometer with angular coverage $2 < \eta < 5$



Excellent performances in

- Primary and secondary vertices reconstruction (VELO)
- Resolution on tracks momentum (Tracking Stations)
- Photons, Electrons, Muons and Hadrons identification (ECAL, HCAL, Muon Stations)
- $\pi / K / p$ identification (RICH1 and RICH2)

- $5.9\ fb^{-1}$ collected at $\sqrt{s} = 13\,TeV$
- $2\ fb^{-1}$ used for this analysis (2015+2016)

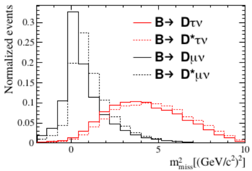
- **All the subdetectors are used for the analysis...**
- **...the response of some of them must be emulated offline in our fast simulations.**

$\mathcal{R}(D^+)$ with $\tau \to \mu\nu\nu$ at LHCb, analysis strategy

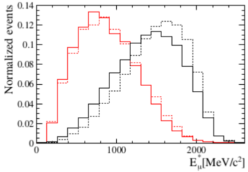
- Measuring $\mathcal{R}(D^{+,*}) = \frac{\mathcal{B}(B \to D^{(+,*)}\tau\nu)}{\mathcal{B}(B \to D^{(+,*)}\mu\nu)}$
  - $\tau \to \mu\nu\nu$
  - $D^+ \to K^-\pi^+\pi^+$

- Simultaneous measurement of $\mathcal{R}(D^+)$ and $\mathcal{R}(D^*)$ with $D^* \to D^+\pi^0$, with un-reconstructed π^0
- Signal ($B \to D^{(*)}\tau\nu$) and normalization ($B \to D^{(*)}\mu\nu$) have the same final state
  - We separate them through a fit to variables evaluated in an approximated rest frame

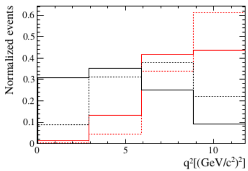
m^2_{miss}



E^*_μ



$q^2 = (p_B - p_D)^2$



Physical backgrounds estimated with MC simulations

- $B \to DH_cX,\ H_c \to \mu\nu X'$
- $B \to D^{**}\mu\nu X$

Other backgrounds in a data driven way

- Fake-D
- Combinatorial
- μ MisID

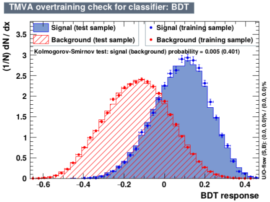
Analysis strategy

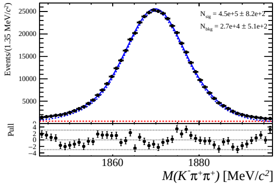
Selections and data driven background

- The data events used are triggered solely on the hadronic part of the event
- The selections require well vertexed, high-p_T μ and $D(\rightarrow K\pi\pi)$ candidates with opposite charge
- Background of prompt-charm from PV removed by requiring a big impact parameter of the D.

Fake-D background

- Particle Identification criteria on the daughters of the D to suppress Fake-D contributions
- Fake-D further suppressed by means of a BDT trained on:
  - **signal**: MC of $B \rightarrow D\mu\nu$
  - **background**: mass sidebands
- Remaining background is statistically subtracted by means of a mass fit





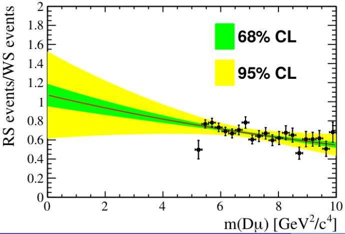
Selections and data driven background

MisID background

- $B \to Dh(\mu)X$
- Estimated in a region enriched of hadrons
- Both in the Right-Sign ($D^{\pm}\mu^{\mp}$) and Wrong-Sign ($D^{\pm}\mu^{\pm}$) sample
- Divided in reconstructed hadron categories
- The contributions from hadron species evaluated by deconvolving the MisID matrix
- The true number of events from each hadron specie is then converted in $N(\hat{\mu}|h)$

$$\begin{pmatrix} N_{\hat{\pi}} \\ N_{\hat{K}} \\ \vdots \\ N_{\hat{g}} \end{pmatrix} = \begin{pmatrix} P(\hat{\pi} \mid \pi) & P(\hat{\pi} \mid K) & \cdots & P(\hat{\pi} \mid g) \\ P(\hat{K} \mid \pi) & P(\hat{K} \mid K) & \cdots & P(\hat{K} \mid g) \\ \vdots & \vdots & \ddots & \vdots \\ P(\hat{g} \mid \pi) & P(\hat{g} \mid K) & \cdots & P(\hat{g} \mid g) \end{pmatrix} \begin{pmatrix} N_{\pi} \\ N_{K} \\ \vdots \\ N_{g} \end{pmatrix}$$

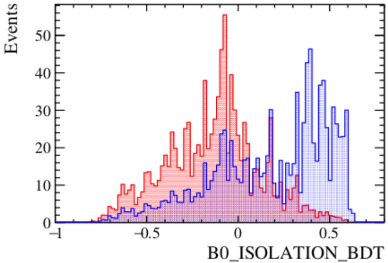
Combinatorial background



- The shape of the combinatorial is estimated from WS combinations
- The MisID is subtracted
- The normalization is corrected from RS/WS ratio, as a function of B mass

Isolation and control regions

- The physical backgrounds are suppressed with a **charged particle isolation BDT**
- It assigns to each non-signal particle a probability of coming from the decay vertex
- We cut on the maximum BDT value in each event

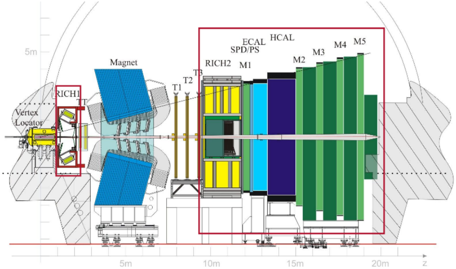


- By inverting this cut we have defined control regions to help us understanding better the background compositions
- We want to perform a simultaneous fit to signal and control regions, with common shape and normalization nuisance parameters

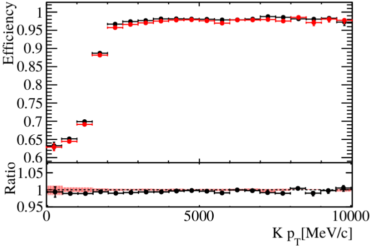
	high-B^0 mass	1OS	2OS	DD
Anti-isolated track(s)		1 π-like	2π-like	1K-like
Charge requirements		D^+h-	$\pi^+\pi^-$	
Purpose	Combinatorial MisID	$B \to D^{**}\mu\nu$	$B \to D_J^{**}\mu\nu$	$B \to DH_cX$

MC simulation

- in 2015 + 2016 dataset (2fb^{-1}) we have $\approx$ 2.8M events
- We need lots of MC, simulating the full detector is unfeasable
- We are using a tracker-only sample of 3B events
- simulate everything which is not in the red boxes



- We emulate the hadron trigger efficiency offline, using tracker information
- I have been in charge of the emulation of the first software level trigger (HLT1)
- **This year I have finished implementing the emulation on 2016 MC**
- **We have published an internal note to document the achievements**

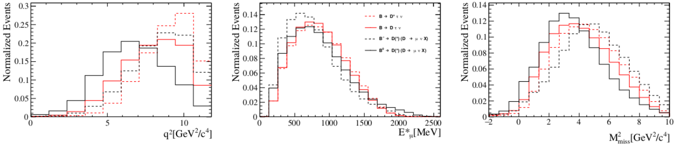


Backgrounds and main systematic uncertainties

The Double Charm background

- The most dangerous background is Double-Charm
- for each control region, 4 templates, dividing the sample by
  - charge of the B mother (B^0, B^+)
  - decay topology (Two body, Multi body)

Two Body templates



$$w(\alpha_1) = 1 + 2\alpha_1 \left(\sqrt{\frac{m^2_{D_1 D_2} - (m_{D_1} + m_{D_2})^2}{(m_B - m_K)^2 - (m_{D_1} + m_{D_2})^2}} - \frac{1}{2} \right)$$

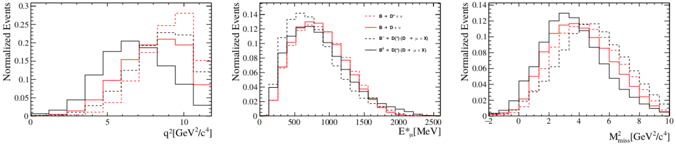
$$w(\alpha_2) = (1 - 2\alpha_2) + 8\alpha_2 \left(\sqrt{\frac{m^2_{D_1 D_2} - (m_{D_1} + m_{D_2})^2}{(m_B - m_K)^2 - (m_{D_1} + m_{D_2})^2}} - \frac{1}{2} \right)^2$$

- The **Multi body** decays are not well known
- Their shape is reweighted and fitted from data

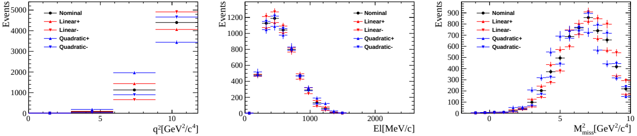
The Double Charm background

- The most dangerous background is Double-Charm
- for each control region, 4 templates, dividing the sample by
  - charge of the B mother (B^0, B^+)
  - decay topology (Two body, Multi body)

Two Body templates

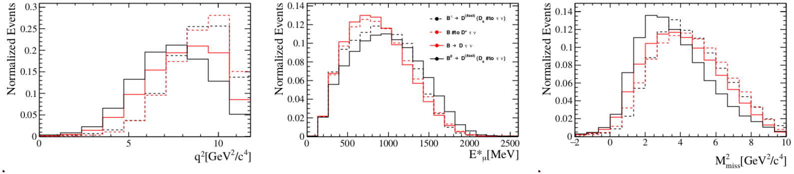


- The **Multi body** decays are not well known
- Their shape is reweighted and fitted from data



Double charm with τ

- If the additional mesons are D_s, the μ can also come from a τ, through the decay of $D_s \to \tau\nu$
- $\mathcal{B}(D_s \to \tau\nu) = 5.5\%$
- We have a dedicated MC sample for this contribution
- This background is very dangerous since it is very similar to the signal in the fit variables



- We have seen in toys that the fit gets unstable when leaving this contribution float freely
- **We constrain it relative to the μ double charm component**

External constraints on the Double charm with τ

- Common normalization parameters as the μ component
- Common shape systematic uncertainty
-

Component	Shape	Normalization
$B^0 \to D^-(D_s^+ \to \tau^+\nu_\tau)$	MC	$N(DD) \times (1. - f_{Bu}) \times f_{DD}^{B_d} \times f_{\tau/\mu} \times \hat{\mathcal{B}}_\mu^\tau$
$B^0 \to D^-(D_s^+ \to \tau^+\nu_\tau)X$	MC + Shape Var.	$N(DD) \times (1 - f_{Bu}) \times (1. - f_{DD}^{B_d}) \times f_{\tau/\mu} \times \hat{\mathcal{B}}_\mu^\tau$
$B^+ \to D^-(D_s^+ \to \tau^+\nu_\tau)$	MC	$N(DD) \times f_{Bu} \times f_{DD}^{B_u} \times f_{\tau/\mu} \times \hat{\mathcal{B}}_\mu^\tau$
$B^+ \to D^-(D_s^+ \to \tau^+\nu_\tau)X$	MC + Shape Var.	$N(DD) \times f_{Bu} \times (1 - f_{DD}^{B_d}) \times f_{\tau/\mu} \times \hat{\mathcal{B}}_\mu^\tau$

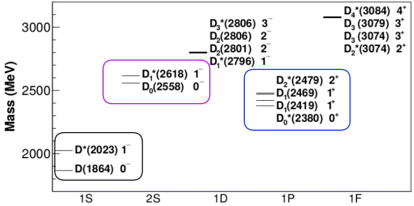
- Two additional normalization factors are included for this component

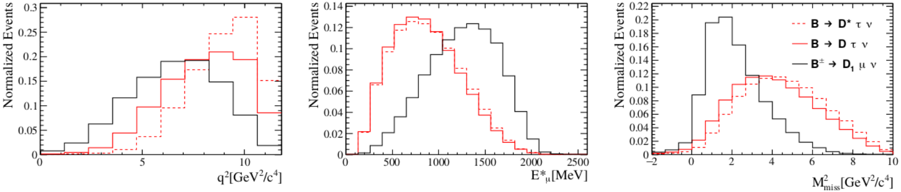
- $f_{\tau/\mu}$ (fixed) contains:
  - the fraction of D_s modes in the muonic sample (in each template, from MC)
  - the ratio of efficiencies (in each template, from MC)

- $\mathcal{B}_\mu^\tau$ is:
  - the ratio between $\mathcal{B}(D_s \to (\tau \to \mu\nu\nu)\nu)$ and $\mathcal{B}(D_s \to \mu\nu X)$, from PDG
  - constrained with a 30% gaussian uncertainty

Feed Down background

- The D^+ and $D^{*,+}$ mesons are the ground states formed by $c-d$ pairs (L=0, S=0)
- Other excited states we consider correspond to $S=1$ and $L=1$.
- They usually decay as $D^{**}\to D^{(*)}\pi$
- We have one MC dataset for each of the $B\to D^{**}$ contributions

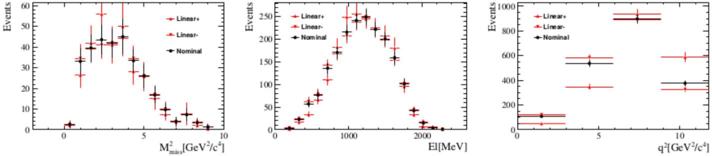




High mass states

- The 2S states are not well known
- We follow the same phenomenological approach we followed for the Double Charm component
- Their shape is reweighted, and we let the fitter interpolate between the alternative templates

$$w(\alpha) = 1+2\alpha\left(\frac{(p_\mu + p_\nu)^2 - m_\mu^2}{8\mathrm{GeV}^2} - 0.5\right)$$



Low mass states

- By combining two spin-1/2 particles in a 1P state, you end up with 4 possible states
- Two braoad states and two narrow states
- We split the simulation in 8 templates according to
  - charge of the B mother
  - D^{**} state

Resonance	Mass (MeV)	Width (MeV)	Decay modes
$D_0^*(2300)^0$	2300 ± 19	274 ± 40	$D^+\pi^-$(seen)
$D_1(2420)^0$	2420.8 ± 0.5	31.7 ± 2.5	$D^*(2010)^+\pi^-$(seen) $D^0\pi^+\pi^-$(seen) $D^+\pi^-$(not seen) $D^{*0}\pi^+\pi^-$(not seen)
$D_2^*(2460)^0$	2460.7 ± 0.4	47.5 ± 1.1	$D^+\pi^-$(seen) $D^*(2010)^+\pi^-$(seen) $D^0\pi^+\pi^-$(not seen) $D^{*0}\pi^+\pi^-$(not seen)
$D_2^*(2465)^\pm$	2465.5 ± 1.3	46.7 ± 1.2	$D^0\pi^+$(seen) $D^{*0}\pi^+$(seen) $D^+\pi^+\pi^-$(not seen) $D^{*+}\pi^+\pi^-$(not seen)

- In each analysis region we fit the yield of one of the 8 components
- We normalize the others using the ratio of efficiencies and ratio of $\mathcal{B}$.

Component	Shape	Normalization
$B^0 \to (D_1 \to D^- X)\mu^+\nu_\mu$	MC + Hammer LLSW	$N(D_1^\pm)$
$B^0 \to (D_0^* \to D^- X)\mu^+\nu_\mu$	MC + Hammer LLSW	$N(D_1^\pm) \times \varepsilon^{D_0^{*\pm}}_{D_1^\pm} \times \hat{\mathcal{B}}^{D_0^{*\pm}}_{D_1^\pm}$
$B^0 \to (D_1' \to D^- X)\mu^+\nu_\mu$	MC +Hammer LLSW	$N(D_1^\pm) \times \varepsilon^{D_1'^\pm}_{D_1^\pm} \times \hat{\mathcal{B}}^{D_1'^\pm}_{D_1^\pm}$
$B^0 \to (D_2^* \to D^- X)\mu^+\nu_\mu$	MC + Hammer LLSW	$N(D_1^\pm) \times \varepsilon^{D_2^{*\pm}}_{D_1^\pm} \times \hat{\mathcal{B}}^{D_2^{*\pm}}_{D_1^\pm}$

Low mass states: external measurements

- We take into account both decay paths to arrive to a $D^+\mu$ final state
  - $B \to D^{**} \to D$
  - $B \to D^{**} \to D^* \to D$
- Some of the decays have only been observed, no measured branching fraction available
- Use Isospin conservation to generalize from measured $\mathcal{B}$

$$\frac{\mathcal{B}(D^{**\pm} \to D^{(*)0}\pi^{\pm})}{\mathcal{B}(D^{**\pm} \to D^{(*)}\pi)} = \frac{\mathcal{B}(D^{**0} \to D^{(*)\pm}\pi^{\mp})}{\mathcal{B}(D^{**0} \to D^{(*)}\pi)} = \frac{2}{3}$$

- All the states widths are saturated by $D^*\pi$ or/and $D\pi$ decays.
- The only exception is D_1^0, which has been lately seen decay to $D\pi\pi$
- I enlarge the error on $\mathcal{B}$ by 10% to include this

Resonance	Mass (MeV)	Width (MeV)	Decay modes
$D_0^*(2300)^0$	2300 ± 19	274 ± 40	$D^+\pi^-$(seen)
$D_1(2420)^0$	2420.8 ± 0.5	31.7 ± 2.5	$D^*(2010)^+\pi^-$(seen)
			$D^0\pi^+\pi^-$(seen)
			$D^+\pi^-$(not seen)
			$D^{*0}\pi^+\pi^-$(not seen)
$D_2^*(2460)^0$	2460.7 ± 0.4	47.5 ± 1.1	$D^+\pi^-$(seen)
			$D^*(2010)^+\pi^-$(seen)
			$D^0\pi^+\pi^-$(not seen)
			$D^{*0}\pi^+\pi^-$(not seen)
$D_2^*(2465)^\pm$	2465.5 ± 1.3	46.7 ± 1.2	$D^0\pi^+$(seen)
			$D^{*0}\pi^+$(seen)
			$D^+\pi^+\pi^-$(not seen)
			$D^{*+}\pi^+\pi^-$(not seen)

$$\begin{aligned} \mathcal{B}(D_1 \to D^*\pi) &= 1, \\ \mathcal{B}(D_1' \to D^*\pi) &= 1, \\ \mathcal{B}(D_0^* \to D\pi) &= 1, \\ \mathcal{B}(D_2^* \to D^*\pi) + \mathcal{B}(D_2^* \to D\pi) &= 1. \end{aligned}$$

- **All the errors and correlation of ratios of $\mathcal{B}$ are put in the fit**

Form Factors: definitions and parameterization

- We want to include the systematic uncertainty that comes from the choice of the model used to generate the $B \to D^{(*)}\ell\nu$ decays.
- The hadronic matrix elements cannot be evaluated from first principles
- They are expressed through **Form Factors**, which can be then measured

$$\langle D|\,\bar{c}\gamma_\mu b\,|\bar{B}\rangle = f_+(q^2)(p_B^\mu + p_D)^\mu + \left[f_0(q^2) - f_+(q^2)\right]\frac{m_B^2 - m_D^2}{q^2}q^\mu,$$
$$\langle D^*|\,\bar{c}\gamma^\mu b\,|B\rangle = -ig(q^2)\varepsilon^{\mu\nu\rho\sigma}\varepsilon_\nu^*(p_B + p_{D^*})_\rho q_\sigma$$
$$\langle D^*|\,\bar{c}\gamma^\mu\gamma^5 b\,|B\rangle = \varepsilon^{*\mu}f(q^2) + a_+(q^2)\varepsilon^*\cdot p_B(p_B + p_{D^*})^\mu + a_-(q^2)\varepsilon^*\cdot p_B$$

- Various parameterizations for them in the literature in terms of $z = \frac{\sqrt{w+1}-\sqrt{2}}{\sqrt{w-1}-\sqrt{2}}$, where $w = v_B \cdot v_{D^{(*)}}$

CLN

- uses dispersion relations, unitarity and HQET
- all form factors are expressed using a universal Isgur-Wise function
- a single tunable parameter ρ

$$f(z) \approx [1 - 8\rho^2 z + (51\rho^2 - 10)z^2 - (252\rho^2 - 84)^3]$$

BGL

- More general, does not use HQET assumptions
- More parameters, expansion series (truncated at finite order)

$$f(z) = \frac{1}{P(z)\phi(z)}\sum_{i=0}^{\infty} a_i z^i$$

- Considered more reliable in the community

External measurements constraints on the form factors

$B \to D$

- FF expanded at third order in BGL
- Parameters constrained using results from a paper which fits Belle, BaBar, FNAL, HPQCD data.
- One of the parameters fixed using a maximum recoil relation

$$f_+(q^2 = 0) = f_0(q^2 = 0)$$

$B \to D^*$

- FF expanded at second order in BGL
- Parameters constrained using results from a paper which fits Belle unfolded data.
- One of the parameters fixed using a zero-recoil relation

$$F_1(z = 0) = \text{constant} \times P_1(z = 0)$$

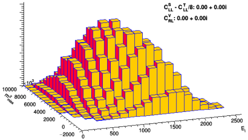
- The helicity suppressed $B \to D^*$ form factor is not measured, and it is being fixed in the fit
- All the errors and correlations between the parameters are taken into account in the fit

Hammer as a forward folding tool

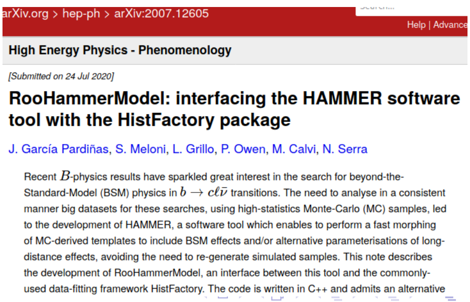
- How do we include the shape variations due to the change in the form factor parameters?
- We forward-fold the variations into the MC simulation (templates morphing)

- We use the **Hammer** tool, which is able to reweight distributions to change FF parameterizations
- It is fast enough to be able to be used at each step of the minimization

Helicity Amplitude Module for Matrix Element Reweighting



- In collaboration with Hammer, **we developed an interface to insert the tool in our fitters**
- We tested the interface, released the code and published the documentation
- The tool can be used also to extract NP Wilson coefficients directly from data, in model independent analyses


arXiv.org > hep-ph > arXiv:2007.12605

Help | Advance

High Energy Physics - Phenomenology

[Submitted on 24 Jul 2020]

RooHammerModel: interfacing the HAMMER software tool with the HistFactory package

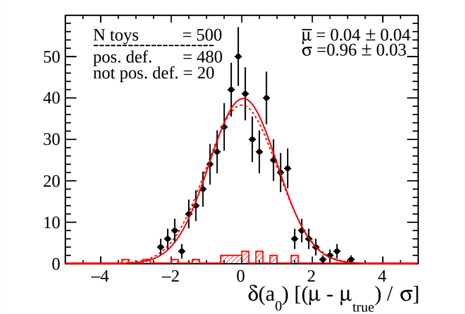
J. García Pardiñas, S. Meloni, L. Grillo, P. Owen, M. Calvi, N. Serra

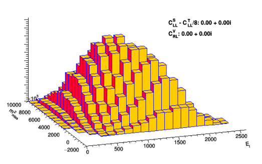
Recent B-physics results have sparkled great interest in the search for beyond-the-Standard-Model (BSM) physics in $b \to c\ell\bar{\nu}$ transitions. The need to analyse in a consistent manner big datasets for these searches, using high-statistics Monte-Carlo (MC) samples, led to the development of HAMMER, a software tool which enables to perform a fast morphing of MC-derived templates to include BSM effects and/or alternative parameterisations of long-distance effects, avoiding the need to re-generate simulated samples. This note describes the development of RooHammerModel, an interface between this tool and the commonly-used data-fitting framework HistFactory. The code is written in C++ and admits an alternative

Hammer as a forward folding tool

- How do we include the shape variations due to the change in the form factor parameters?
- We f
- We us distrib
- It is f of the
- In col interf
- We te publis
- The t coefficients directly from data, in model independent analyses

- Now using this tool into our analysis
- example for a pull of one FF parameter





lp | Advance

tware

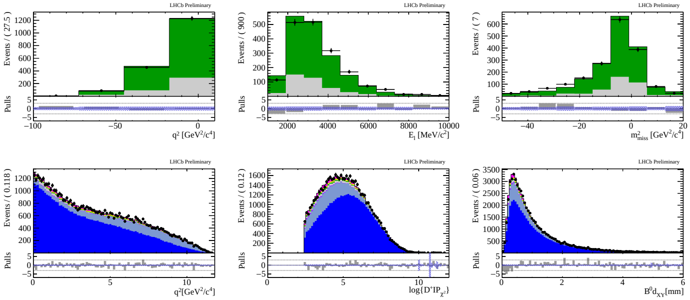
sistent es, led phing of MC-derived templates to include BSM effects and/or alternative parameterisations of long-distance effects, avoiding the need to re-generate simulated samples. This note describes the development of RooHammerModel, an interface between this tool and the commonly-used data-fitting framework HistFactory. The code is written in C++ and admits an alternative

Data/MC comparisons

Data/MC agreement

- with the model we have developed, we are comparing the data and the MC in some validation regions

- Region of $m_{D\mu} > m_B$
- Only non physical backgrounds contribute to this region
  - Combinatorial
  - MisID
- Normalization enriched region: $m^2_{\text{miss}} < 0$
  - $B \to D\mu\nu$
  - $B \to D^*\mu\nu$
- Fit and topological variables



- After having performed the fit to the data, we plan to do a final data/MC agreement check, projecting the fit result in all regions and various variables

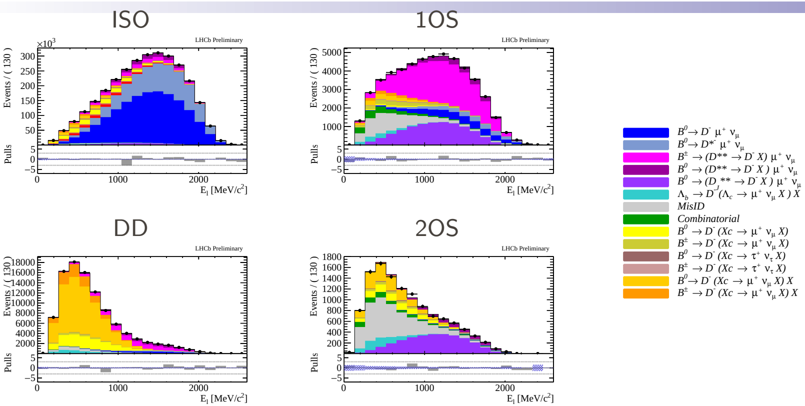
Toy studies

Toy studies

Component	Shape	Normalization
$B^0 \to D^- \mu^+ \nu_\mu$	MC + Hammer BGLVar	$N(D\mu) \times \mathrm{TF}$
$B^0 \to D^- \tau^+ \nu_\mu$	MC + Hammer BGLVar	$N(D\mu) \times \mathrm{TF} \times \mathcal{R}_{\mathrm{raw}}(D^+)$
$B^0 \to D^{*-} \mu^+ \nu_\mu$	MC + Hammer BGLVar	$N(D^*\mu) \times \mathrm{TF}$
$B^0 \to D^{*-} \tau^+ \nu_\mu$	MC + Hammer BGLVar	$N(D^*\mu) \times \mathrm{TF} \times \mathcal{R}_{\mathrm{raw}}(D^*)$
$B^0 \to D^-(X_c \to \mu^+ \nu_\mu X)$	MC	$N(DD) \times (1. - f_{Bu}) \times f_{DD}^{B_d}$
$B^0 \to D^-(X_c \to \mu^+ \nu_\mu X)X$	MC + Shape Var.	$N(DD) \times (1 - f_{Bu}) \times (1. - f_{DD}^{B_d})$
$B^+ \to D^-(X_c \to \mu^+ \nu_\mu X)$	MC	$N(DD) \times f_{Bu} \times f_{DD}^{B_u}$
$B^+ \to D^-(X_c \to \mu^+ \nu_\mu X)X$	MC + Shape Var.	$N(DD) \times f_{Bu} \times (1 - f_{DD}^{B_d})$
$B^0 \to D^-(D_s^+ \to \tau^+ \nu_\tau)$	MC	$N(DD) \times (1. - f_{Bu}) \times f_{DD}^{B_d} \times f_{\tau/\mu} \times \hat{\mathcal{B}}_\mu^\tau$
$B^0 \to D^-(D_s^+ \to \tau^+ \nu_\tau)X$	MC + Shape Var.	$N(DD) \times (1 - f_{Bu}) \times (1. - f_{DD}^{B_d}) \times f_{\tau/\mu} \times \hat{\mathcal{B}}_\mu^\tau$
$B^+ \to D^-(D_s^+ \to \tau^+ \nu_\tau)$	MC	$N(DD) \times f_{Bu} \times f_{DD}^{B_u} \times f_{\tau/\mu} \times \hat{\mathcal{B}}_\mu^\tau$
$B^+ \to D^-(D_s^+ \to \tau^+ \nu_\tau)X$	MC + Shape Var.	$N(DD) \times f_{Bu} \times (1 - f_{DD}^{B_d}) \times f_{\tau/\mu} \times \hat{\mathcal{B}}_\mu^\tau$
$\Lambda_b \to (\Lambda_c \to \mu\nu X)DX'$	MC	$N(\Lambda_b)$
$B^0 \to (D_1 \to D^- X)\mu^+ \nu_\mu$	MC + Hammer LLSW	$N(D_1^\pm)$
$B^0 \to (D_0^* \to D^- X)\mu^+ \nu_\mu$	MC + Hammer LLSW	$N(D_1^\pm) \times \varepsilon_{D_1^\pm}^{D_0^{*\pm}} \times \hat{\mathcal{B}}_{D_1^\pm}^{D_0^{*\pm}}$
$B^0 \to (D_1' \to D^- X)\mu^+ \nu_\mu$	MC +Hammer LLSW	$N(D_1^\pm) \times \varepsilon_{D_1^\pm}^{D_1'^\pm} \times \hat{\mathcal{B}}_{D_1^\pm}^{D_1'^\pm}$
$B^0 \to (D_2^* \to D^- X)\mu^+ \nu_\mu$	MC + Hammer LLSW	$N(D_1^\pm) \times \varepsilon_{D_1^\pm}^{D_2^{*\pm}} \times \hat{\mathcal{B}}_{D_1^\pm}^{D_2^{*\pm}}$
$B^\pm \to (D_1 \to D^- X)\mu^+ \nu_\mu$	MC + Hammer LLSW	$N(D_1^\pm) \times \varepsilon_{D_1^\pm}^{D_1^0} \times \hat{\mathcal{B}}_{D_1^\pm}^{D_1^0}$
$B^\pm \to (D_0^* \to D^- X)\mu^+ \nu_\mu$	MC + Hammer LLSW	$N(D_1^\pm) \times \varepsilon_{D_1^\pm}^{D_0^{*0}} \times \hat{\mathcal{B}}_{D_1^\pm}^{D_0^{*0}}$
$B^\pm \to (D_1' \to D^- X)\mu^+ \nu_\mu$	MC + Hammer LLSW	$N(D_1^\pm) \times \varepsilon_{D_1^\pm}^{D_1'^{*0}} \times \hat{\mathcal{B}}_{D_1^\pm}^{D_1'^{*0}}$
$B^\pm \to (D_2^* \to D^- X)\mu^+ \nu_\mu$	MC + Hammer LLSW	$N(D_1^\pm) \times \varepsilon_{D_1^\pm}^{D_2^{*0}} \times \hat{\mathcal{B}}_{D_1^\pm}^{D_2^{*0}}$
$B^0 \to (D_J^{**} \to D^- X)\mu^+ \nu_\mu$	MC + Shape Var.	$N(D_J^{**})$
MisID	data	$N(\mathrm{MisID})$
Combinatorial	data	$N(\mathrm{Comb})$

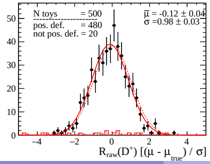
- The fit model is very complicated
- many constrained parameters and a lot more free parameters.
- **I spent a lot of time this year developing a stable and reliable fit model.**
- **I have tested the model against fit bias and coverage issues.**

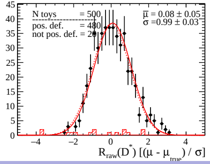
Toy studies



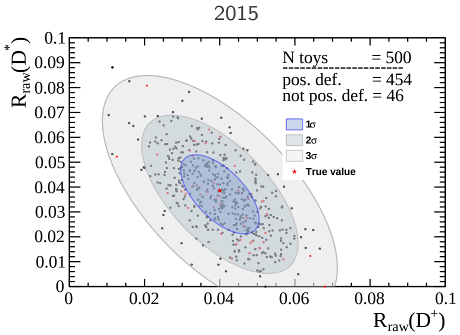
- The fits are simultaneous in all signal and control regions used in the nominal fit
- The datasets are generated taking the nominal model, smeared with a Poisson uncertainty in each bin

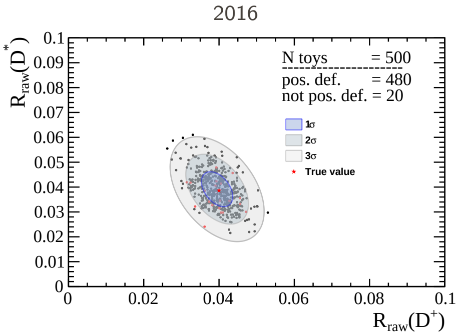
- No bias is observed in any of the parameters





Toy studies: results





parameter	true value	uncertainty	rel. uncertainty
$N(D^+\mu\nu)$	2.54e+05	1.3e+03	0.5%
$N(D^*\mu\nu)$	1.36e+05	2.0e+03	1.5%
$R_{raw}(D^+)$	4.00e-02	1.0e-02	$2.6e+01$%
$R_{raw}(D^*)$	3.86e-02	1.6e-02	$4.1e+01$%

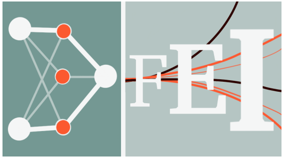
parameter	true value	uncertainty	rel. uncertainty
$N(D^+\mu\nu)$	1.52e+06	2.8e+03	0.18%
$N(D^*\mu\nu)$	8.14e+05	3.4e+03	0.42%
$R_{raw}(D^+)$	4.00e-02	4.2e-03	$1e+01$%
$R_{raw}(D^*)$	3.86e-02	7.0e-03	$1.8e+01$%

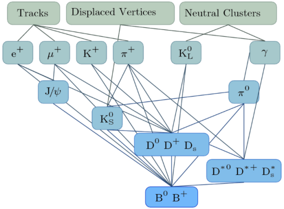
- Raw numbers have to be converted into measured $\mathcal{R}(D^{+,*})$, but **it will be a very competitive result**

A new project: DFEI

DFEI

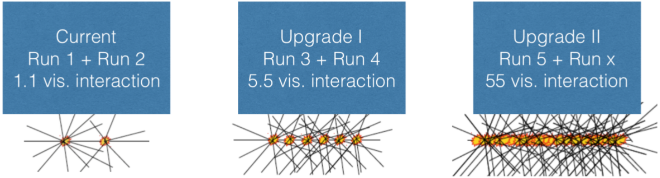
- I have lately joined a new project, linked to Juliàn's Marie Curie
- **DFEI: Deep Full event interpretation in LHCb**
- At the moment the signal reconstruction is done based on a **signal-hypothesis** approach:
  - You reconstruct the signal particles, the rest is considered background
- Some other experiments try to reconstruct all the decays in the event
  - Belle II: Full Event Interpretation (Decision Tree)
- **Aim**: Try to reconstruct all (reconstructible?) decays





DFEI: Why?

- The main background to be modelled in many key analyses is the **Combinatorial**
  - Decay of the other b-hadron in the event
  - Tracks from the rest of the event
  - The situation will significantly worsen with the LHCb upgrades

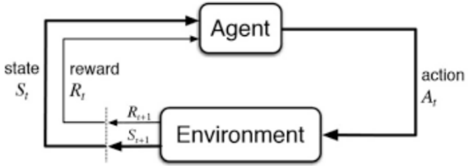


- Why do we want to try Deep Learning?
- The increase in luminosity poses computational challenges for the trigger
  - **One can try to enhance the information in the trigger with DL**
  - E.g.: can we avoid trying out *all* particle combinations in the online reconstruction?
- Limited available storage:
  - Can we compress the information somehow with DL?

The first approach: Reinforcement Learning

- The idea is to have an agent which will have the role to combine particles, assigning PID hypothesis etc.
- How do you train this agent?
- **Reinforcement Learning:** neither Supervised nor unsupervised learning
  - training data: experiences of the agent
  - training signal: reward from the environment

- It is all about the interaction with the environment
- The agent senses the state of the environment and decides upon an action
- The environment gives a reward signal to the agent
- It presents the agent with a new state



- This techniques are used AIs to beat games. To make a parallel with chess:
  - Move your pawns → combine particles
  - Board → list of particles you can combine

The reinforcement learning problem

- The aim is to maximize the expected total reward, G
- The rewards are discounted by a factor $0 \leq \gamma \leq 1$
- $\pi(a|s)$ is the policy with which the action are chosen.
- You have to find the best policy

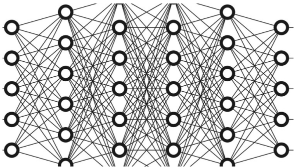
$$G = \sum_{t=0}^{T} \gamma^t E_\pi[r_t]$$

- For each state-action pair, you can assign a number Q_π, telling you how much you value that combination
- The best policy is the one that choses, for each state, the action with maximum value

$$Q_\pi(s, a) = E_\pi[G|s, a]$$

- **Solving the reinforcement learning problem is equivalent to find the optimal policy, or equivalently finding the best value function**

- Some algorithms use tables of states and actions to approximate the best Q function
- This is impossible in our case, since the number of possible states is vastly large
- We use neural networks as function approximators for the Q value



Deep-Q learning

- The optimal Q function is not known, but under the otpimal policy it follows some recursion relations: **Bellman Optimality equations:**

$$Q^*(s_t, a_t) = E[r_{t+1} + \gamma \max_{a_{t+1}} Q^*(s_{t+1}, a_{t+1})]$$

- Recursive problem: at each time step you formulate a minimization problem to minimize the difference between the left hand side and the one-sample approximation of the right hand side

$$Err(w) = 0.5|Q_w(s,a) - r - \gamma max_{a'} Q_{\bar{w}}(s', a')|^2$$

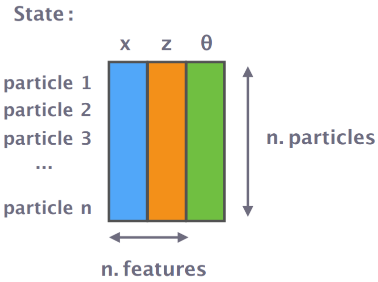
DQN algorithm

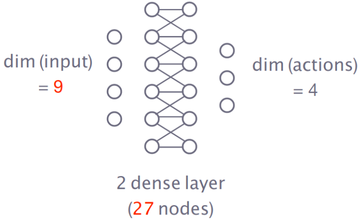
- Observe s, select and execute a
- observe s' and get reward r
- Gradient $\frac{\partial Err}{\partial w} = [Q_w(s,a) - r - \gamma max_{a'} Q_w(s', a')] \frac{\partial Q_w}{\partial w}$
- update weights $w \leftarrow w - \alpha \frac{\partial Err}{\partial w}$

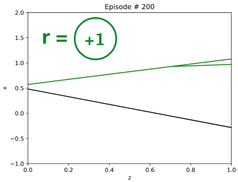
- Deep Learning APIs give you the tools to evaluate automatically the gradients
- **I have implemented this algorithm and some other tools that are needed for the reinforcement learning problem**

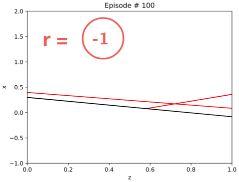
First test of the algorithm in a mockup environment

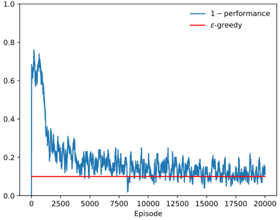
- We tested this in a simple, 2D world with just 3 particles, 2 with the same mother











- Project at its very start
- I am in charge of looking into more advanced algorithms

Conclusions

- Testing lepton flavour universality in semileptonic decays at the LHCb experiment
- Interesting discrepancies are being observed in similar analyses
- Our analysis will report the simultaneous measurement of $\mathcal{R}(D^+)$ and $\mathcal{R}(D^*)$.

- This year I concentrated on including many systematic uncertainties in the fit
- Very difficul analysis with lots of nasty background: many external measurements are needed to constrain better the fit to data
- I have tested the model against bias and coverage, in all the control regions of the analysis
- We are now fitting the data and assessing the data/MC agreement in validation region

- I have joined a more technical project
- Aims at studying if a full event interpretation is feasable at LHCb and if it can bring some advantages
- Involves usage of state-of-the-art Machine Learning techniques
- Starting from scratch, in a field very different from our expertise

Backup

Low mass states

- With the previous assumptions and external measurements I have evaluated the $\mathcal{B}$ for all the states
- I use these numbers to evaluate the constraint on $\mathcal{B}$ in the fit

Mode	$\mathcal{B}_{\to D^* \to D^+}(10^{-3})$	$\mathcal{B}_{\to D^+}(10^{-3})$	$\sigma/\mathcal{B}$
$\bar{B}^0 \to D_1^+ \mu^- \nu$	0.45	0.0	11.1%
$\bar{B}^0 \to D_2^{*+} \mu^- \nu$	0.11	0.61	23.1%
$\bar{B}^0 \to D_1^{'+} \mu^- \nu$	0.5	0.0	29.1%
$\bar{B}^0 \to D_0^{*+} \mu^- \nu$	0.0	1.5	40.0%
$B^- \to D_1^0 \mu^- \nu$	0.98	0.0	7.5%
$B^- \to D_2^{*0} \mu^- \nu$	0.33	1.53	9.6%
$B^- \to D_1^{'0} \mu^- \nu$	0.87	0.0	22.3%
$B^- \to D_0^{*0} \mu^- \nu$	0.0	2.5	20.0%

- All the $\mathcal{B}$ ratios have the same denominator, so they are correlated with each other
- The constraints, with the full correlation matrix, are put in the fit to include systematic uncertainties for the D^{**} composition

External measurements on $B \to D$ form factors

- We expand the FF at third order in BGL
- We constrain the parameters using results from a paper which fits Belle, BaBar, FNAL, HPQCD data:

parameter	value		error
a_{+0}	0.01566	$\pm$	0.00011
a_{+1}	-0.0342	$\pm$	0.0031
a_{+2}	-0.090	$\pm$	0.022
a_{00}	0.07935	$\pm$	0.00058
a_{01}	-0.205	$\pm$	0.014
a_{02}	-0.23	$\pm$	0.10

- a_{00} is fixed to the value of other parameters from a maximum recoil relation

$$f_+(q^2 = 0) = f_0(q^2 = 0)$$

- All the errors and correlations are taken into account in the fit
- In order to avoid numerical problems in the minimization, the covariance matrix of this result is diagonalized and the fit is performed on its principal components

External measurements on $B \to D^*$ form factors

- We expand all the parameters at second order, except one which is expanded at third order in BGL
- We constrain the parameters using results from a paper which fits Belle unfolded data

parameter	value		error
a_0	0.000379	$\pm$	0.000249
a_1	0.026954	$\pm$	0.009320
b_0	0.000550	$\pm$	0.000023
b_1	-0.002040	$\pm$	0.001064
c_1	-0.000433	$\pm$	0.000264
c_2	0.005350	$\pm$	0.004606
d_0	0.002		
d_1	-0.013		

- The helicity suppressed form factor parameters d have never been measured and are fixed in the fit.
- c_0 is fixed to other parameters values through the zero-recoil relation

$$F_1(z=0) = \text{constant} \times P_1(z=0) \tag{1}$$

- All the errors and correlations are taken into account in the fit
- In order to avoid numerical problems in the minimization, the covariance matrix of this result is diagonalized and the fit is performed on its principal components

The idea of Hammer

- The Hammer package takes the moves from the observation that the matrix elements for a semileptonic decay is linear in the form factors (or can be written in a linear form by a first order expansion)

$$\mathcal{M} = FF^{\alpha}\mathcal{M}^{\alpha} \tag{2}$$

- A given vector (FF) corresponds to a given choice of the form factors parameters used to evaluate the rate

$$\Gamma \approx |\mathcal{M}|^2 = |\mathcal{M}^{\alpha}\mathcal{M}^{\alpha}|^2 \tag{3}$$

- Instead of filling histograms with events, they can be filled with tensors $\mathcal{W}^{\alpha\beta} = \mathcal{M}^{\alpha}\mathcal{M}^{\beta}$
- When one needs the number of events in a given bin, the tensors can be contracted

$$\Gamma \approx FF^{T} \cdot \mathcal{W} \cdot FF \tag{4}$$

How can this be used?

- Knowing the tensors and having generated a MC sample with one choice of Form Factor parameters, one can reweight the Reco-Level histograms (one weight factor per histogram bin).

$$r_i = \frac{\Gamma_{\mathrm{new}}}{\Gamma_{\mathrm{old}}} \tag{5}$$

- It is quick to evaluate the weights, only linear operations involved
- A change in the model is *convolved* inside the full simulation, instead of deconvolving data from experimental resolutions

The Hammer architecture

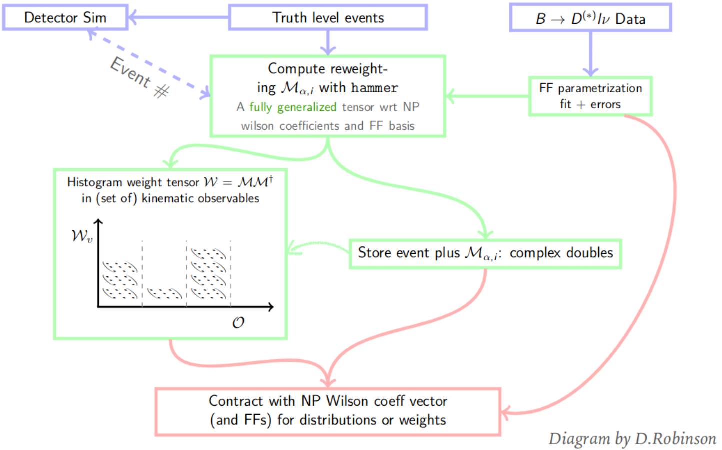
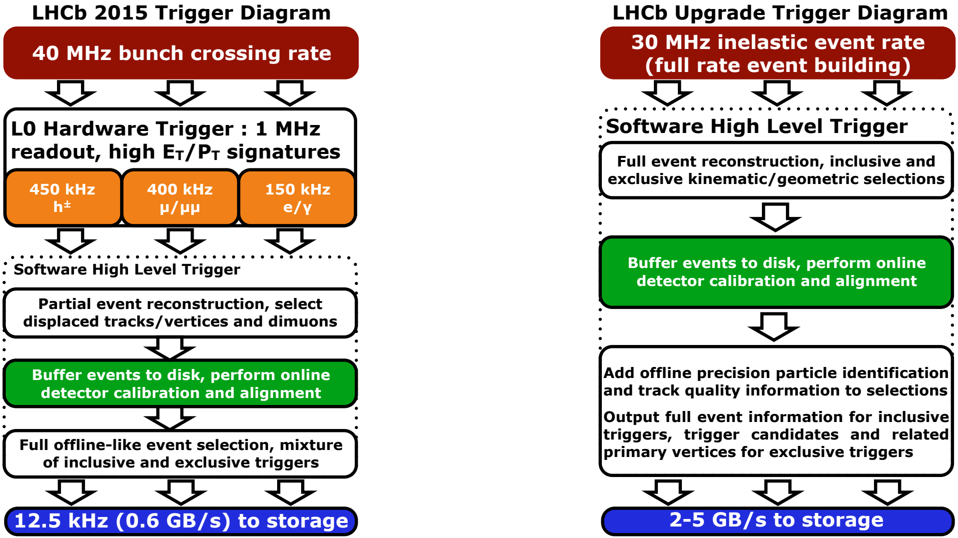


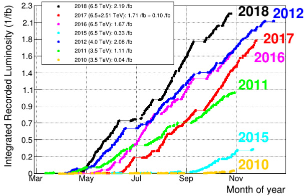
Diagram by D.Robinson

(C++ library w/ python bindings + optional histogram interface to ROOT → can be integrated easily with existing software)

Trigger configurations



LHCb performance numbers



Resolutions	momentum resolution: Δ p / p = 0.5 % at low momentum to 1.0% at 200 GeV/c (see the detector performance paper for a plot)
	ECAL resolution (nominal): 1 % + 10 %/ √(E[GeV])
	impact parameter resolution: (15 +29/pT[GeV]) μm
	invariant mass resolution: ~8 MeV/c^2 for B → J/ψ X decays with constraint on J/ψ mass ~22 MeV/c^2 for two-body B decays ~100 MeV/c^2 for B_s → φ γ, dominated by photon contribution
	decay time resolution: ~45 fs for B_s → J/ψ φ and for B_s → D_s π
Efficiencies	percentage of working detector channels: ~99 % for all sub-detectors
	data taking efficiency:90 % (good for analyses: 99%)
	trigger efficiencies: ~ 90 % for dimuon channels ~ 30 % for multi-body hadronic final states
	track reconstruction efficiency: ~ 96 % for long tracks
	Particle ID efficiency: Electron ID ~ 90 % for ~ 5 % e→h mis-id probability Kaon ID ~ 95 % for ~ 5 % π→K mis-id probability Muon ID ~ 97 % for 1-3 % π→μ mis-id probability

Detailed description of the selections

Trigger

L0 Any of:	
	L0Global TIS
	L0Hadron TOS
Hlt1 Any of:	
	Hlt1TrackMVA TOS
	Hlt1TwoTrackMVA TOS
Hlt2:	
	Hlt2XcMuXForTauB2XcMu TOS

L0 Trigger

L0 trigger	E_T/p_T threshold 2015	2016	2017	SPD threshold
Hadron	> 3.6 GeV	> 3.7 GeV	> 3.46 GeV	< 450
Photon	> 2.7 GeV	> 2.78 GeV	> 2.47 GeV	< 450
Electron	> 2.7 GeV	> 2.4 GeV	> 2.11 GeV	< 450
Muon	> 2.8 GeV	> 1.8 GeV	> 1.35 GeV	< 450
Muon high p_T	> 6.0 GeV	> 6.0 GeV	> 6.0 GeV	none
Dimuon	> 1.69 GeV^2	> 2.25 GeV^2	> 1.69 GeV^2	< 900

HLT1Track and HLT1TwoTrackMVA

Hlt1TrackMVa
Input tracks selections (after Track-Fit)
$p_T > 500$ MeV $p > 3$ GeV Track χ^2/d.o.f. < 4.0
$\chi^2/n.d.f < 2.5$ $(p_T > 25\ \text{GeV}/c \wedge IP_{\chi^2} > 7.4) \vee [(1\ \text{GeV}/c < p_T < 25\ \text{GeV}/c) \wedge$ $\log(IP_{\chi^2}) > \left(\frac{1}{p_T[\text{GeV/c}] - 1}\right)^2 + \left(\frac{1.1}{25\ \text{GeV}/c}\right)(25\ \text{GeV}/c - p_T) + \log(7.4)]$

Table 14: Requirements of the Hlt1TrackMVa trigger line during 2015 data taking

	Hlt1TwoTrackMVa
Requirement on the single tracks	$p_T > 0.5$ GeV/c $p > 5.0$ GeV/ $\chi^2/n.d.f < 2.5$
Requirements on the track pair before vertexing	$(p_1 + p_2)_T > 2$ GeV/c $DOCA(1,2) < 10$
Requirements on the track pair combination	Vertex $\chi^2 < 10$ $m_{corr} > 1$ GeV/c^2 $2 < \eta < 5$ $DIRA > 0$
MVA requirements	MVA Output > 0.95
MVA training variables	Vertex χ^2 Vertex distance χ^2 $p_{T,1} + p_{T,2}$ How many tracks with $IP_{\chi^2} < 16$.

HLT2, Stripping, Preselection, Filtering

Particle	Variable	Hlt2 cuts	Stripping cuts	Filtering cuts
K, π	K $PIDK$	> 2	> 4	–
	π $PIDK$	< 4	< 2	–
	Track IP_{χ^2}	> 9	> 9	> 9
	Track p_T [MeV/c]	> 200	> 300	> 300
	Track p [MeV/c]	> 5000	> 2000	> 2000
	≥ 1 track p_T [MeV/c]	> 800	–	–
	$\sum$ track p_T [MeV/c]	> 2500	> 2500	> 2500
	Track GhostProb	–	< 0.5	< 0.5
D	D mass interval [MeV/c^2]	1830 – 1910	1790 – 1950	1770 – 1970
	D p_T [MeV/c]	> 2000	–	–
	D child pair DOCA [mm]	< 0.10	–	–
	D χ^2_{vertex}/ndf	< 10	< 4	< 4
	D DIRA	> 0.999	> 0.999	> 0.999
	D FD_{χ^2}	> 25	> 25	> 25
μ	μ IP_{χ^2}	> 16	> 16	> 16
	μ $PID\mu$	–	> -200	–
	μ GhostProb	–	< 0.5	< 0.5
	μ p [MeV/c]	–	> 3000	> 3000
$D\mu$	$D\mu$ χ^2_{vertex}/ndf	< 15	< 6	< 6
	$D\mu$ DIRA	> 0.999	> 0.999	0.999
	$D\mu$ DOCA [mm]	< 0.50	–	–
	$D\mu$ FD_{χ^2}	> 50	–	–
	$D\mu$ mass interval [MeV/c^2]	< 10500	0 – 10000	–
	$D\mu$ mass interval [MeV/c^2] (Before vert.)	< 11000	< 10200	–

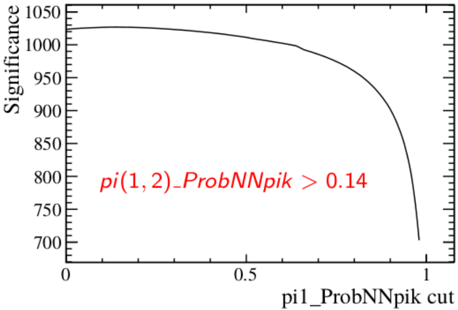
Generator level selections

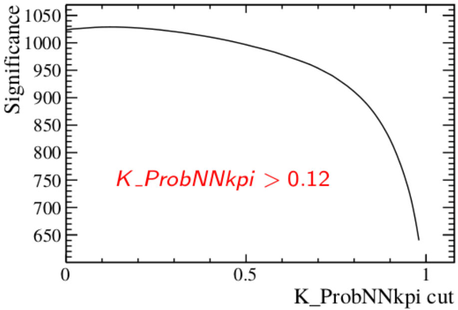
Variable	Requirement
$\theta(\pi)$	$\in [0.01, 0.4]$ rad
$\theta(K)$	$\in [0.01, 0.4]$ rad
$\theta(\mu)$	$\in [0.01, 0.4]$ rad
$p_T(\pi)$	> 150 MeV/c
$p_T(K)$	> 150 MeV/c
$p(\mu)$	> 2500 MeV/c
$p(K^+) + p(\pi^-) + p(\pi^-)$	> 15000 MeV/c
$p_T(K^+) + p_T(\pi^-) + p_T(\pi^-)$	> 2300 MeV/c

Table 3: List of generator level selections

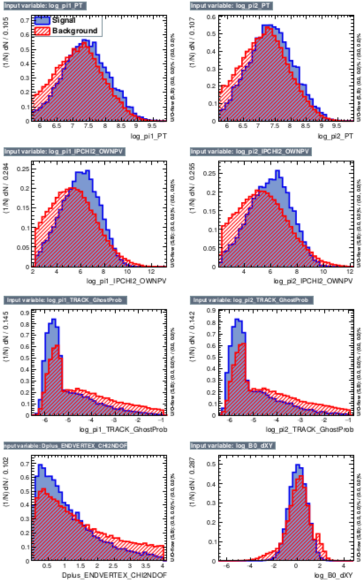
D daughters PID selection

- Already existing PID cuts: $pi1\_DLLK < 2$, $pi2\_DLLK < 2$ and $K\_DLLK > 4$.
- **New PID variables**: **ProbNNpi** $*$ (**1** $-$ **ProbNNk**) for pions and **ProbNNk** $*$ (**1** $-$ **ProbNNpi**) for kaons.
- Cut on each variable optimised on data (to avoid using PID info from MC), through fits to the 3-body mass distribution, taking $S/\sqrt{S+B}$ as FoM.





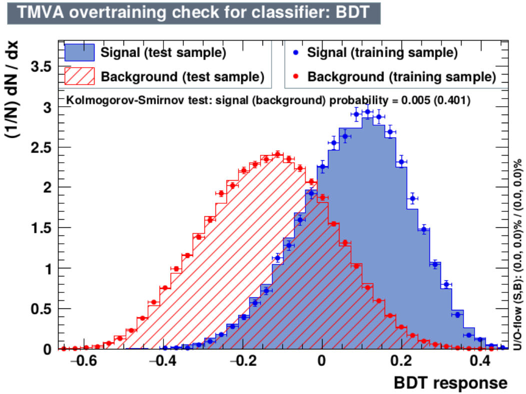
BDT against non D background- Input variables

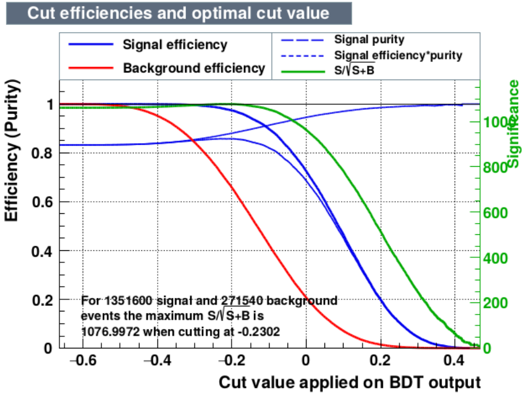


- log(pi1_PT)
- log(pi2_PT)
- log(K_PT)
- log(pi1_IPCHI2_OWNPV)
- log(pi2_IPCHI2_OWNPV)
- log(K_IPCHI2_OWNPV)
- log(pi1_TRACK_GhostProb)
- og(pi2_TRACK_GhostProb)
- og(K_TRACK_GhostProb)
- Dplus_ENDVERTEX_CHI2NDOF
- log(B0_dXY)

BDT against non D background- Input variables

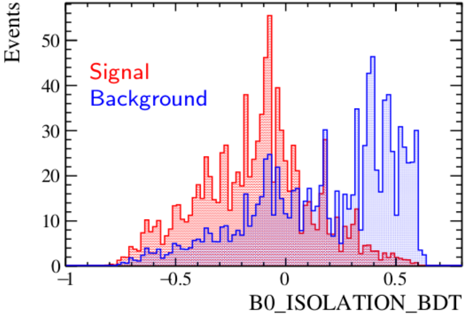
- **Signal**$B^0 \to D^+ \mu\nu$
- **Background:**D sidebands
- Cut optimized on $\frac{S}{\sqrt{S+B}}$, > -0.23

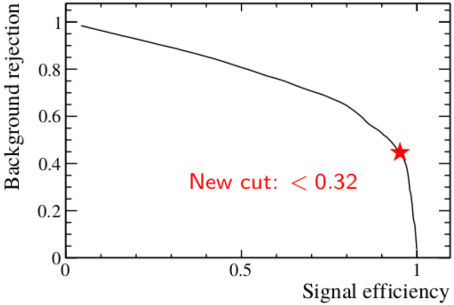




Charged Isolation

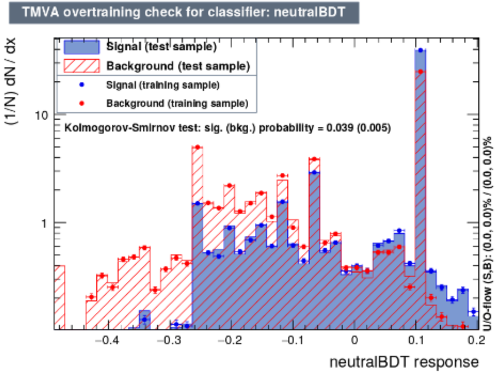
- Using the BDT trained for the $R(D^*)$ measurement.
- Old cut of < 0.15 reoptimised for this analysis.
- **Signal sample:** Bd2Dpmunu MC sample (11574061).
- **Background sample:** Bd2DD, DD cocktail, MC sample (11995203).





Neutral Isolation

- Two independent methods trained to suppress additional neutral particles
  - The two methods are then combined in a single Neutral isolation output
- Signal: $B \to D\mu\nu$
- Background: $B \to (D^* \to D\pi^0)\mu\nu$

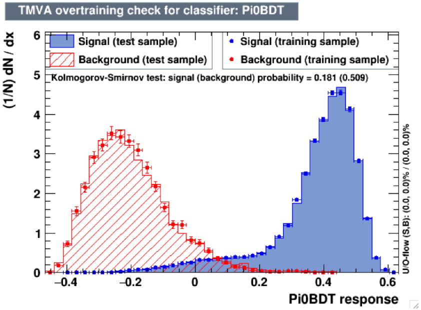


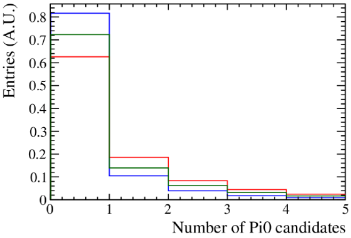
- Signal effficiency 0.9
- Background rejection 0.3
- Cut > -0.16

- The two BDTs used in input to this one are explained in the followng slides

Resolved Pions BDT

- For each π^0 in the event evaluate a BDT trained on
  - π^0s from $B \to D\mu\nu$ as signal
  - Truth matched π^0s from $B \to (D^* \to D\pi^0)\mu\nu$ as background
- Evaluate a per event quantity by counting how many π^0s with BDT< 0

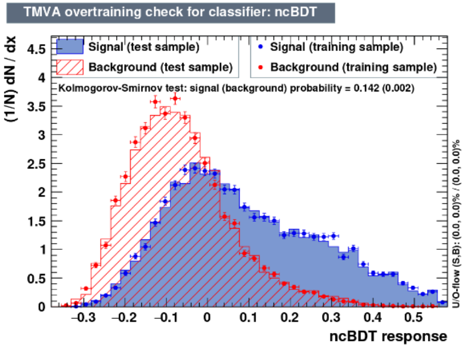




$Bd \to D^+ munu$, $Bd \to D^{**} munu$, $Bd \to D^{*+}(\to D^+\pi^0) munu$

Neutral Cones BDT

- In each event construct a cone around the D^+ flight direction
- Evaluate a BDT trained using variables related to activity inside the cone



$\mathcal{R}(D^+)$ with $\tau \to \mu\nu\nu$ at LHCb, analysis strategy

- **Aim** of the analysis is to measure

$$\mathcal{R}(D^+) = \frac{\mathcal{BF}(\bar{B}^0 \to D^+\tau^-\bar{\nu}_\tau)}{\mathcal{BF}(\bar{B}^0 \to D^+\mu^-\bar{\nu}_\mu)}$$

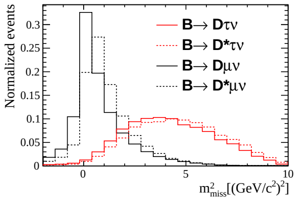
$$\mathcal{BF}(\tau^- \to \mu^-\nu_\tau\bar{\nu}_\mu) = (17.39 \pm 0.04)\%$$
$$\mathcal{BF}(D^+ \to K^+\pi^-\pi^+) = (8.98 \pm 0.28)\%$$

- **Theoretical point of view:** clean because $|V_{cb}|$ and hadronic form factors uncertainties cancel in the ratio

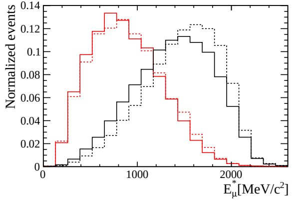
$$\mathcal{R}(D^+)_{\rm SM} = 0.300 \pm 0.008$$

- **Experimental point of view**: Signal and normalization channels have the same final state
  - Most of uncertainties due to efficiency and reconstruction cancel
  - The two channels are separated using **3 kinematical variables**, computed in the B rest frame

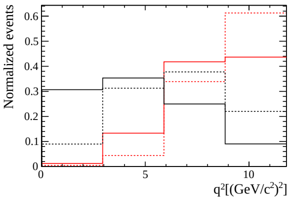
m^2_{miss}



E^*_μ



$q^2 = (p_B - p_D)^2$



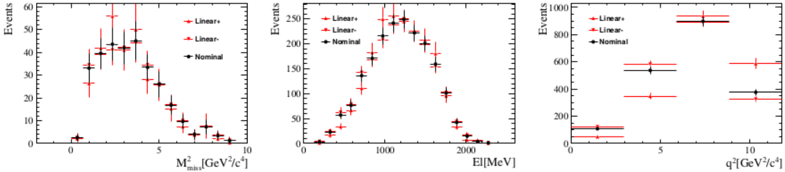
Shape systematics

- Some backgrounds are modelled by `cocktails` of poorly known B decays.
- The assumptions about their composition can induce biases in the measurement.

- Varying all the assumed branching ratios inside the cocktails would be a titanic work
- The control samples can actually be used to check the data MC agreement
- The idea is to let the fit have enough variation to adjust the MC shape in the control regions.
- **Solution:** Include some phenomenological shape variation as systematics

Reweight to the $B \to D_J^{**}\mu\nu$ sample

$$w(\alpha) = 1 + 2\alpha\left(\frac{(p_\mu + p_\nu)^2 - m_\mu^2}{8\,\mathrm{GeV}^2} - 0.5\right)$$



(h) ISO sample

Shape systematics

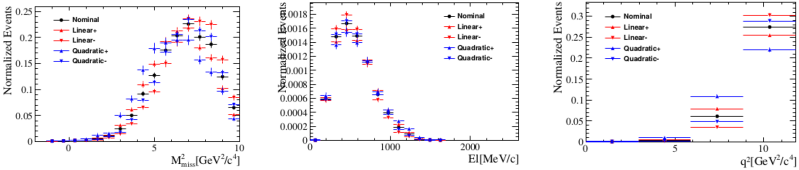
- Some backgrounds are modelled by `cocktails` of poorly known B decays.
- The assumptions about their composition can induce biases in the measurement.

- Varying all the assumed branching ratios inside the cocktails would be a titanic work
- The control samples can actually be used to check the data MC agreement
- The idea is to let the fit have enough variation to adjust the MC shape in the control regions.
- **Solution:** Include some phenomenological shape variation as systematics

Reweight to the $B \to DD\mu\nu$ sample

$$w(\alpha_1) = 1 + 2\alpha_1 \left(\sqrt{\left(\frac{m^2_{D_1D_2} - (m_{D_1} + m_{D_2})^2}{(m_B - m_K)^2 - (m_{D_1} + m_{D_2})^2} \right)} - \frac{1}{2} \right)$$

$$w(\alpha_2) = (1 - \alpha_2) + 8\alpha_2 \left(\sqrt{\left(\frac{m^2_{D_1D_2} - (m_{D_1} + m_{D_2})^2}{(m_B - m_K)^2 - (m_{D_1} + m_{D_2})^2} \right)} - \frac{1}{2} \right)^2$$



MC samples

Sample	Event type
$\bar{B}^0 \to D^+\tau(\to \mu\bar{\nu}_\mu\nu_\tau)\bar{\nu}_\tau$	11574060
$\bar{B}^0 \to D^{*+}(\to D^+\pi^0)\tau(\to \mu\bar{\nu}_\mu\nu_\tau)\bar{\nu}_\tau$	11574401
$\bar{B}^0 \to D^+\mu\bar{\nu}_\mu$	11574061
$\bar{B}^0 \to D^{*+}(\to D^+\pi^0)\mu\bar{\nu}_\mu$	11574402
$\bar{B}^0 \to D^{**+}(\to D^+X)\mu\bar{\nu}_\mu$	11574403
$\bar{B}^0 \to D^{**+}(\to D^+X)\mu\bar{\nu}_\mu$, high mass	11574070
$B^- \to D^{**0}(\to D^+X)\mu\bar{\nu}_\mu$	12874050
$\bar{B}^0 \to D^+H_c(\to \mu\bar{\nu}_\mu X')X$	11995204
$B^- \to D^+H_c(\to \mu\bar{\nu}_\mu X')X$	12995604
$\Lambda_b \to (\Lambda_c \to \mu\nu X)DX'$	15976000
$B^0 \to D^\pm(D_s \to \tau\nu)X$	11995214
$B^\pm \to D^\pm(D_s \to \tau\nu)X$	12995615

Table 2: List of Monte Carlo samples used.

Full MC, how many?

Sample	Generated events (2015)
$\bar{B}^0 \to D^+\tau(\to \mu\bar{\nu}_\mu\nu_\tau)\bar{\nu}_\tau$	2M (0.1M)
$\bar{B}^0 \to D^{*+}(\to D^+\pi^0)\tau(\to \mu\bar{\nu}_\mu\nu_\tau)\bar{\nu}_\tau$	2M (0.1M)
$\bar{B}^0 \to D^+\mu\bar{\nu}_\mu$	10M (0.5M)
$\bar{B}^0 \to D^{*+}(\to D^+\pi^0)\mu\bar{\nu}_\mu$	10M (0.5M)
$\bar{B}^0 \to D^{**+}(\to D^+X)\mu\bar{\nu}_\mu$	2M (0.14M)
$\bar{B}^0 \to D^{**+}(\to D^+X)\mu\bar{\nu}_\mu$, high mass	1M (0.06M)
$B^- \to D^{**0}(\to D^+X)\mu\bar{\nu}_\mu$	2M (0.14M)
$\bar{B}^0 \to D^+H_c(\to \mu\bar{\nu}_\mu X')X$	6M (0.34M)
$B^- \to D^+H_c(\to \mu\bar{\nu}_\mu X')X$	3M (0.15M)
$\Lambda_b \to (\Lambda_c \to \mu\nu X)DX'$	–
$B^0 \to D^\pm(D_s \to \tau\nu)X$	–
$B^\pm \to D^\pm(D_s \to \tau\nu)X$	–

Table 5: Generated full-MC samples. In parenthesis, the number of events after filtering is indicated.

Tracker Only, how many?

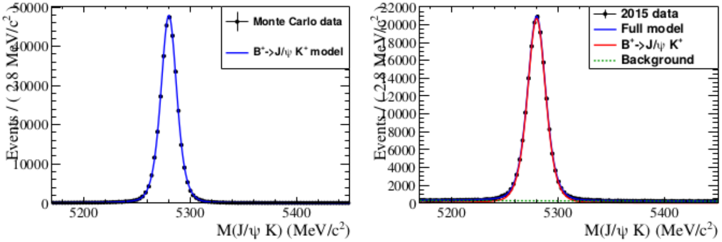
Sample	Generated events (2015)	Generated events (2016)
$\bar{B}^0 \to D^+\tau(\to \mu\bar{\nu}_\mu\nu_\tau)\bar{\nu}_\tau$	20M (1M)	120M (6M)
$\bar{B}^0 \to D^{*+}(\to D^+\pi^0)\tau(\to \mu\bar{\nu}_\mu\nu_\tau)\bar{\nu}_\tau$	20M (1M)	120M (6M)
$\bar{B}^0 \to D^+\mu\bar{\nu}_\mu$	100M (5M)	600M (30M)
$\bar{B}^0 \to D^{*+}(\to D^+\pi^0)\mu\bar{\nu}_\mu$	100M (5M)	600M (30M)
$\bar{B}^0 \to D^{**+}(\to D^+X)\mu\bar{\nu}_\mu$	20M (1M)	120M (6M)
$\bar{B}^0 \to D^{**+}(\to D^+X)\mu\bar{\nu}_\mu$, high mass	10M (0.5M)	60M (3M)
$B^- \to D^{**0}(\to D^+X)\mu\bar{\nu}_\mu$	20M (1M)	120M (6M)
$\bar{B}^0 \to D^+H_c(\to \mu\bar{\nu}_\mu X')X$	60M (3.4M)	360M (18M)
$B^- \to D^+H_c(\to \mu\bar{\nu}_\mu X')X$	30M (1.5M)	180M (9M)
$\Lambda_b \to (\Lambda_c \to \mu\nu X)DX'$	2M (0.1M)	12M (0.6M)
$B^0 \to D^\pm(D_s \to \tau\nu)X$	4M (0.2M)	24M (1.2M)
$B^\pm \to D^\pm(D_s \to \tau\nu)X$	4M (0.2M)	24M (1.2M)

Table 6: Generated tracker-only-MC samples. In parenthesis, the number of events after filtering is indicated.

Kinematic reweighting 1

Control sample: $B^+ \to J/\psi K^+$ (2015 so far)

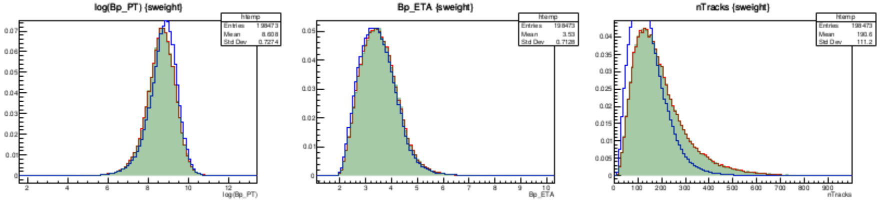
- **Stripping:** BetaSBu2JpsiKDetachedLine.
- **Trigger:** (L0Muon || L0DiMuon) && Hlt1TrackMuon && Hlt2DimuonDetatchedHeavy.
- Using DTF with constraint on the PV and the Jpsi mass.
- **Preselection:** similar to the $R(D^*)$ analysis, rectangular cuts and sWeights (using the Bp mass as variable, signal + comb. bkg.).



Kinematic reweighting 2

Weights obtained from **GBreweighter**, trained on the **3D distribution** of:

- $\log(Bp\_PT)$.
- Pseudorapidity (ETA) of the Bp.
- Number of tracks in the event.



Blue line: MC

Red line: sWeighted signal data

Green filled area: reweighted MC

Checked that the reweight does not negatively affect the J/ψ and K kinematic distributions (they actually improve).

Correction to the Double Charm control sample

- Reweight $B^0 \to D_1 D_2 X$ and $B^\pm \to D_1 D_2 X$ events with two (common) weight functions

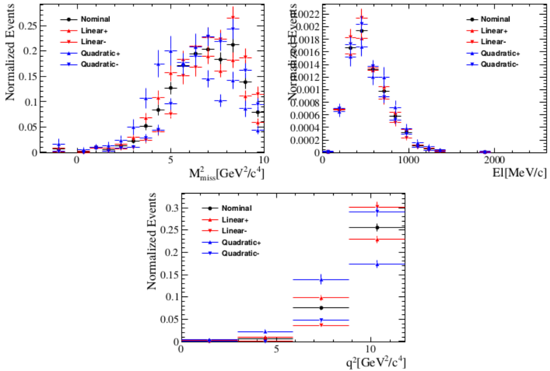
$$w(\alpha_1) = 1 + 2\alpha_1 \left(\sqrt{\left(\frac{m^2_{D_1 D_2} - (m_{D_1} + m_{D_2})^2}{(m_B - m_K)^2 - (m_{D_1} + m_{D_2})^2} \right)} - \frac{1}{2} \right)$$

$$w(\alpha_2) = (1 - \alpha_2) + 8\alpha_2 \left(\sqrt{\left(\frac{m^2_{D_1 D_2} - (m_{D_1} + m_{D_2})^2}{(m_B - m_K)^2 - (m_{D_1} + m_{D_2})^2} \right)} - \frac{1}{2} \right)^2$$

- Evaluate the templates at $\alpha_i = \pm 1$, and include them in the fit as systematic variations

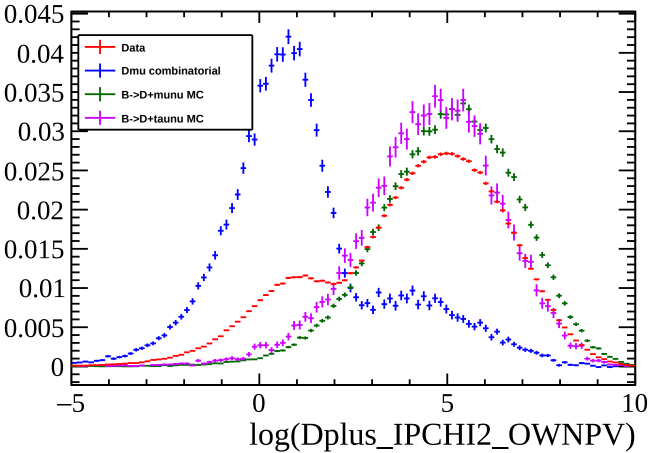
- Interpolate between them and fit for α_i

$B^\pm \to D_1 D_2 X$



Combinatorial background suppression

- The combinatorial fraction seemed a little bit too high at the beginning
- We think to have tracked down the problem... We miss a IP_{χ^2} cut for the D^+ candidate
  - This cut would reject most of the Combinatorial from prompt D candidates



- Thanks a lot to Greg for the suggestion!!

Data driven backgrounds walkthrough

- The measurement relies on the correct evaluation of the backgrounds, that must be tackled in order
- Analysis chain:
  1. μ-**MisID**: Unfold its distribution from real data using weights extracted from prescaled `!isMuon` sample, both in $\{D^+\mu^-\}cc$ and $\{D^+\mu^+\}cc$ samples.
  2. **Non** D **background:** Extract sWeights from the D-MassFit to the $\mu-$PID weighted sample
  3. **Combinatorial**: taken from the sWeighted, μ-MisID subtracted $\{D^+\mu^+\}cc$ sample
  4. **Physical backgrounds**: estimated from MC
     - Eventually extracting corrections using data driven studies in dedicated control samples
- Before...
  - sWeights were extracted before evaluating PID weights ($1 \longleftrightarrow 2$)
  - The weights were extracted for the whole sample, and then some isolation categories were defined
- Now...
  - We first define all the isolation categories (And never touch selections again!)
  - ...The whole analysis chain is repeated for all the isolation categories we are defining

$\mathcal{R}(D^*)$ with $\tau \to \mu\nu\nu$

(PRL 115, 11804)

Model uncertainties	Absolute size ($\times 10^{-2}$)
Simulated sample size	2.0
Misidentified μ template shape	1.6
$\overline{B}^0 \to D^{*+}(\tau^-/\mu^-)\overline{\nu}$ form factors	0.6
$\overline{B} \to D^{*+}H_c(\to \mu\nu X')X$ shape corrections	0.5
$\mathcal{B}(\overline{B} \to D^{**}\tau^-\overline{\nu}_\tau)/\mathcal{B}(\overline{B} \to D^{**}\mu^-\overline{\nu}_\mu)$	0.5
$\overline{B} \to D^{**}(\to D^*\pi\pi)\mu\nu$ shape corrections	0.4
Corrections to simulation	0.4
Combinatorial background shape	0.3
$\overline{B} \to D^{**}(\to D^{*+}\pi)\mu^-\overline{\nu}_\mu$ form factors	0.3
$\overline{B} \to D^{*+}(D_s \to \tau\nu)X$ fraction	0.1
Total model uncertainty	**2.8**
Normalization uncertainties	**Absolute size ($\times 10^{-2}$)**
Simulated sample size	0.6
Hardware trigger efficiency	0.6
Particle identification efficiencies	0.3
Form-factors	0.2
$\mathcal{B}(\tau^- \to \mu^-\overline{\nu}_\mu\nu_\tau)$	< 0.1
Total normalization uncertainty	**0.9**
Total systematic uncertainty	**3.0**

$$\mathcal{R}(D^*) = 0.336 \pm 0.027(\text{stat.}) \pm 0.030(\text{syst.})$$

2.1 σ higher than the Standard Model

systematic uncertainties

- MC statistics
- Shape of the Mis-ID background
- Shape of the MC derived background models
  - Depend on the statistics in the control regions
  - They will be reduced in the measurements performed with the RunII data
- Hadronic form factors