Status of $B_{d,s}\rightarrow\mu^{+}\mu^{-}$ in LHCb

III CPAN DAYS, Barcelona
November 2nd, 2011

Xabier Cid Vidal

Outline

- Introduction
- Analysis in LHCb
  - Overview
  - Calibration and normalization
- Results
- Conclusions

Introduction

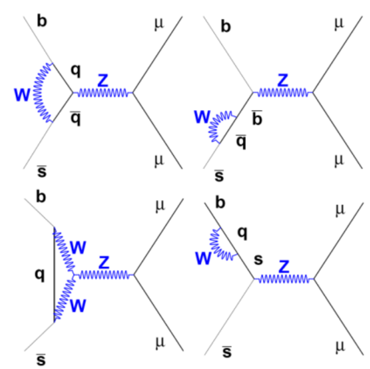
Introduction

- $B_{d,s} \to \mu^+\mu^-$ can access NP through new virtual particles entering in the loop → indirect search of NP, accesing higher energy scales!

- These decays are very supressed in the SM:
  - **BR($B_s \to \mu\mu$) = (3.35 ± 0.32)·10^{-9}**
  - **BR ($B_d \to \mu\mu$) = (1.03 ± 0.09)·10^{-10}**

M.Blanke et al., JHEP 10 003,2006

Main SM diagrams



(q = u, c, t)

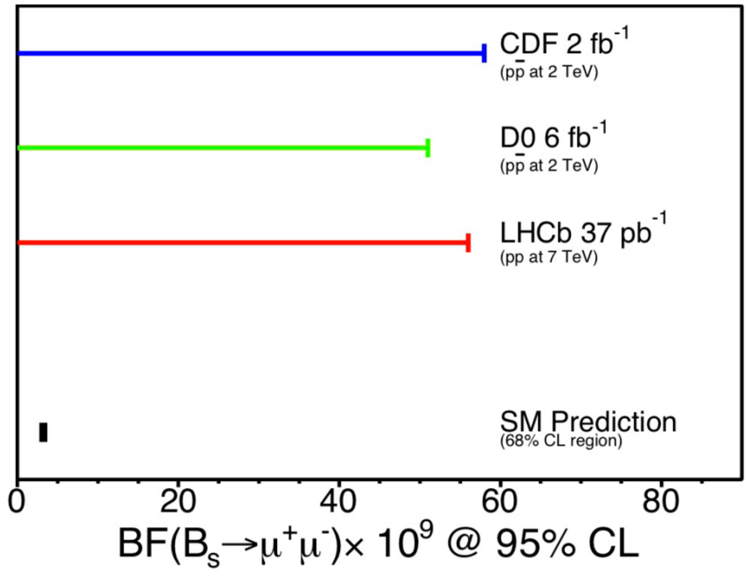
- Theory details of the decay will be explained by **Diego Martinez** on his talk tomorrow:

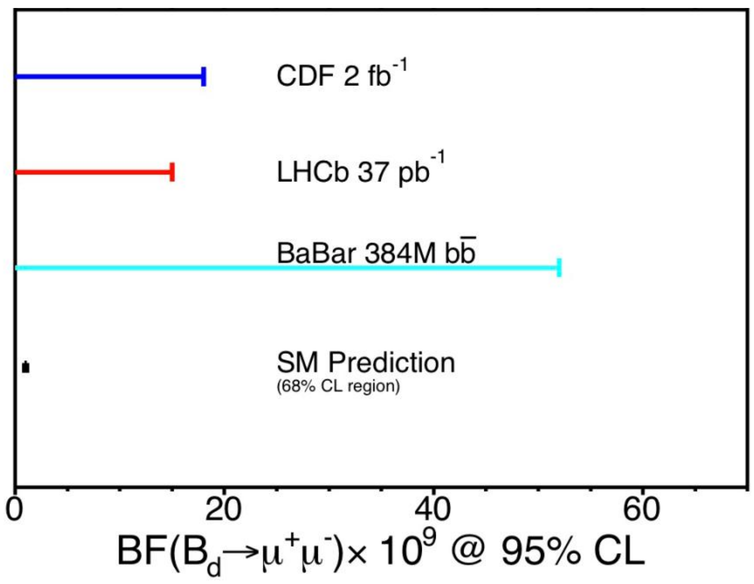
  Theory implications of $B_s \to \mu\mu$ recent measurements by LHCb and CMS

Current experimental status (I)

- LHCb already published one analysis based on 37 pb^{-1} from 2010 data [*Physics Letter B 699 (2011)330-340]*
  - Observed **$BR(B_s \rightarrow \mu^+\mu^-) < 4.3 \times 10^{-8}$** ($5.6 \times 10^{-8}$) @ 90 (95)% [CL Expected: 5.1 (6.5)]
  - Observed **$BR(B^0 \rightarrow \mu^+\mu^-) < 1.2 \times 10^{-8}$** ($1.5 \times 10^{-8}$) @ 90 (95)% [CL Expected: 1.4 (1.8)]

- Experimental status before summer





Current experimental status (II)

- **CDF annoucement** (July 2011), With 7 fb^{-1} taken at 1.96 TeV:
  - $BR(B_s \to \mu\mu) < 4.0 \times 10^{-8}$ @ 95% CL
  - $BR(B \to \mu\mu) < 6.0 \times 10^{-9}$ @ 95% CL

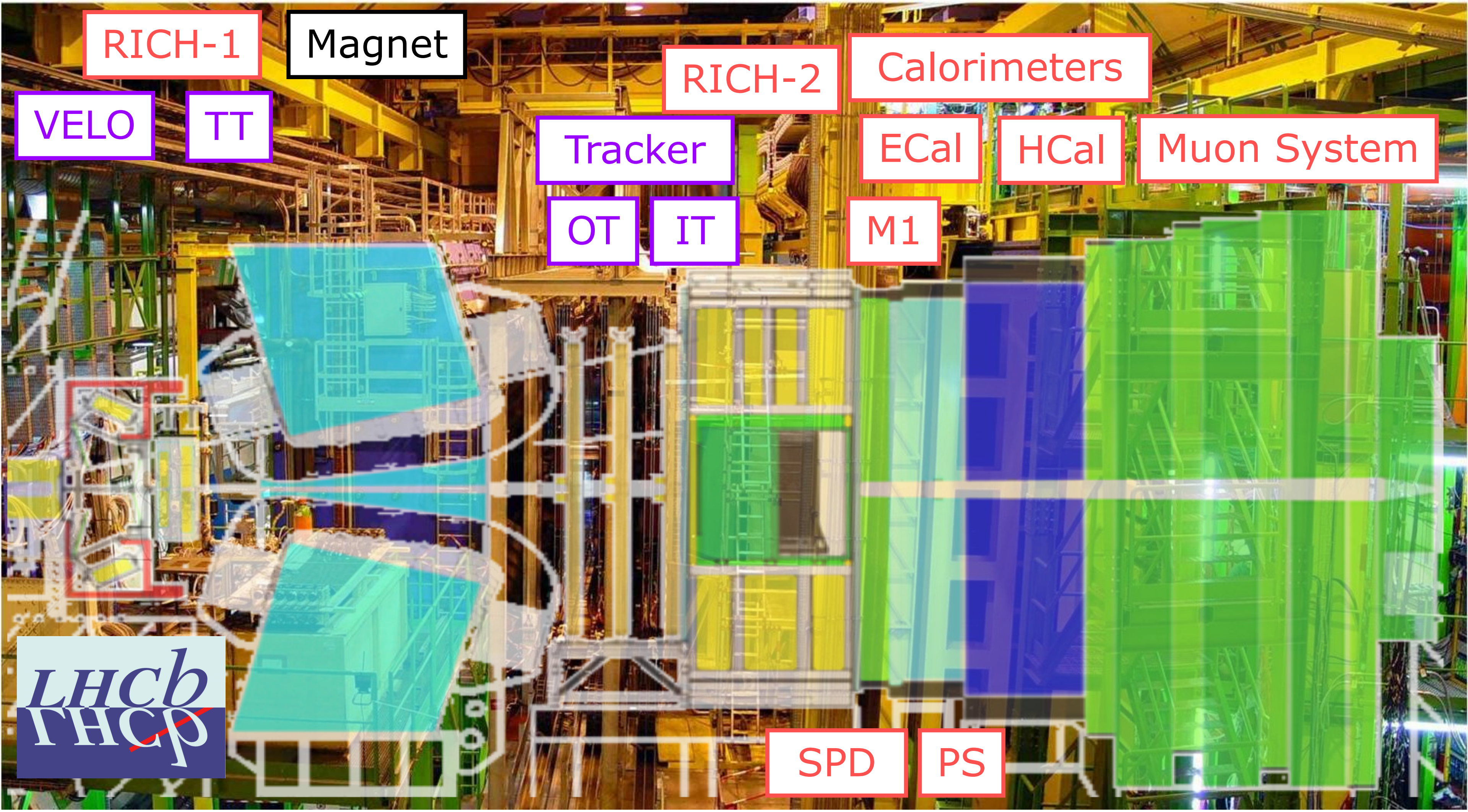
  Excess of B_s candidates: p-value (to be bkg) 0.27%,
  $BR(B_s \to \mu\mu) = 1.8^{+1.1}_{-0.9} \times 10^{-8}$

- An update based on LHCb data taken before July 2011 (**370 pb^{-1}**) is presented here.
  - Assuming SM, **4.00 B_s and 0.39 B_d events** are expected in these data after selection.
  - Assuming CDF measurement, 21^{+14}_{-10} B_s are expected
  - Paper with these data is under review in the Collaboration and will be submitted to *Phys. Lett. B* at the end of November.
  - LHCb has already collected 1 fb^{-1}

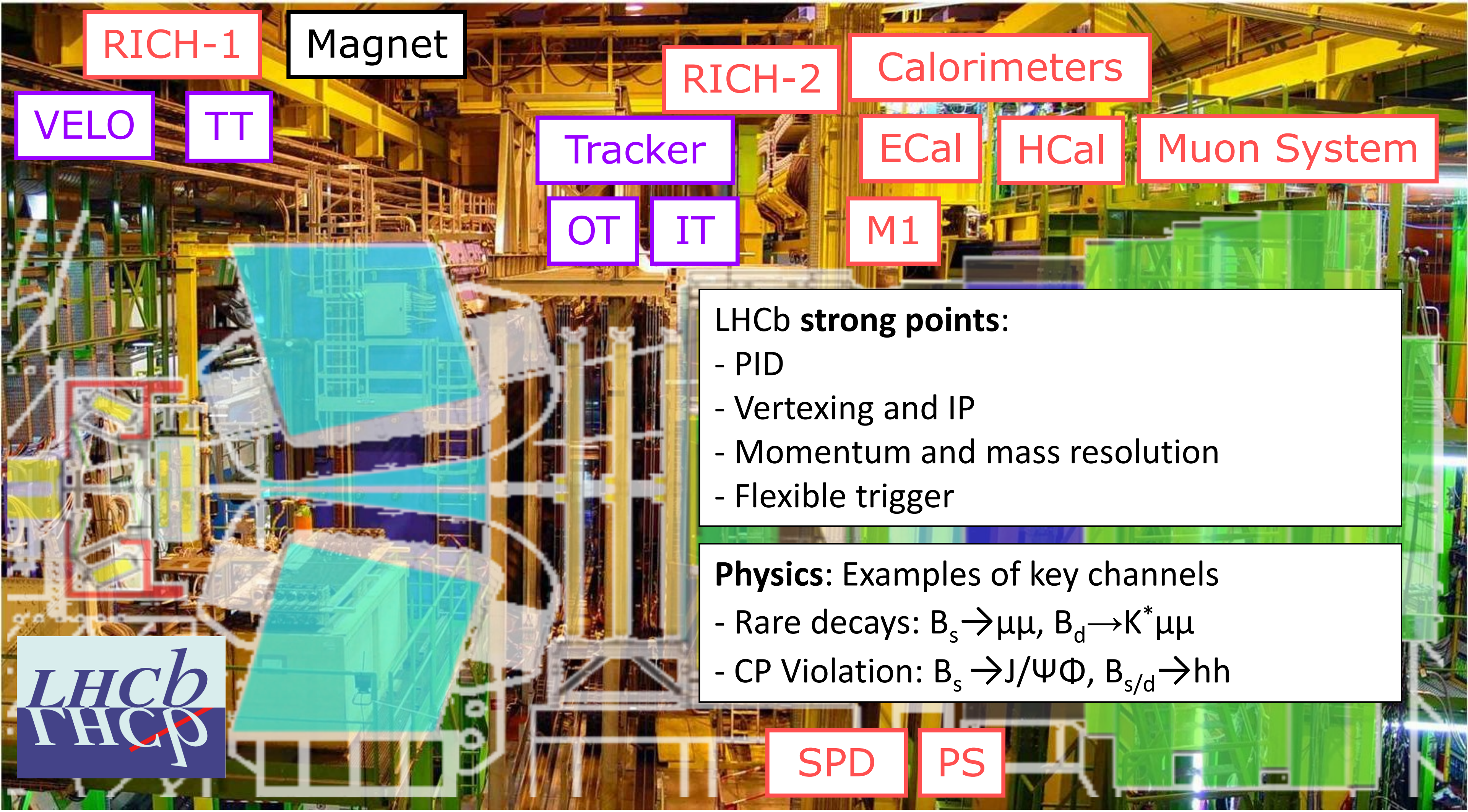
Analysis in LHCb

⟶ Overview

LHCb overview



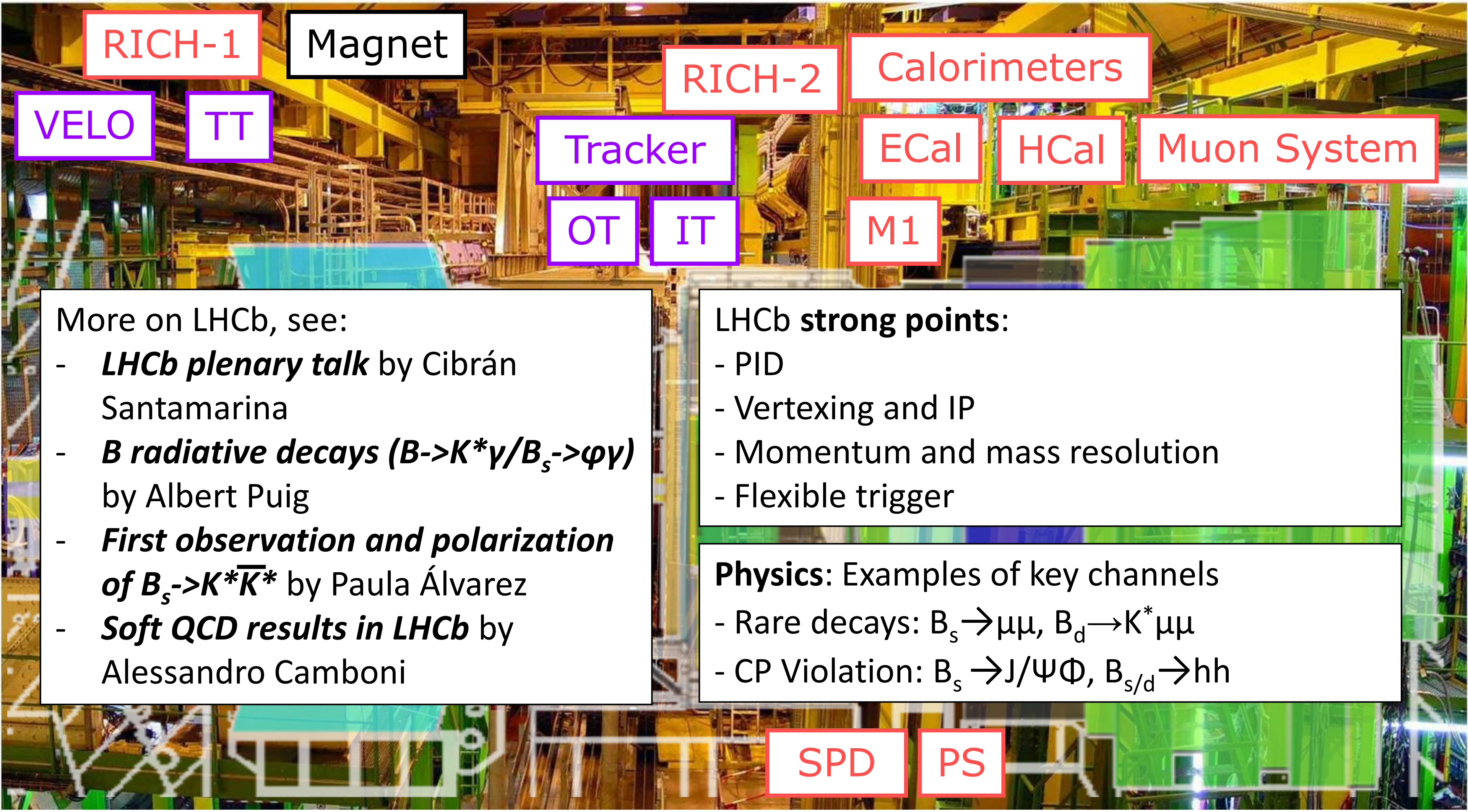
LHCb overview



LHCb **strong points**:
- PID
- Vertexing and IP
- Momentum and mass resolution
- Flexible trigger

Physics: Examples of key channels
- Rare decays: $B_s \rightarrow \mu\mu$, $B_d \rightarrow K^* \mu\mu$
- CP Violation: $B_s \rightarrow J/\Psi\Phi$, $B_{s/d} \rightarrow hh$

LHCb overview



More on LHCb, see:

- ***LHCb plenary talk*** by Cibrán Santamarina
- ***B radiative decays ($B\to K^*\gamma/B_s\to\phi\gamma$)*** by Albert Puig
- ***First observation and polarization of $B_s\to K^*\overline{K}^*$*** by Paula Álvarez
- ***Soft QCD results in LHCb*** by Alessandro Camboni

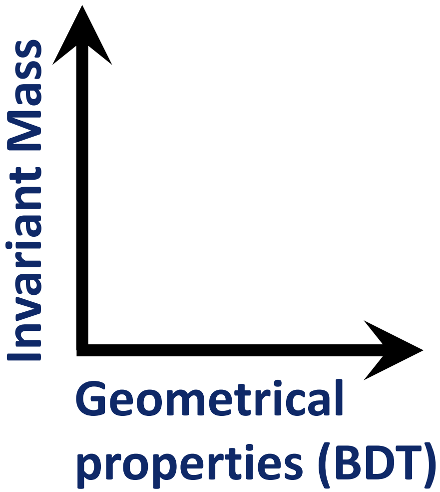
LHCb **strong points**:
- PID
- Vertexing and IP
- Momentum and mass resolution
- Flexible trigger

Physics: Examples of key channels
- Rare decays: $B_s\to\mu\mu$, $B_d\to K^*\mu\mu$
- CP Violation: $B_s\to J/\Psi\Phi$, $B_{s/d}\to hh$

Analysis overview

- Selection: apply some cuts on all μμ candidates to remove most of the background.

- Classify each event using two variables (bins in a 2D parameter space):
  - **Geometrical properties** (combined in Boosted Decision Tree)
  - **Invariant Mass**



- 2D space is binned, so that each bin is treated as an independent experiment. Results combined using **Modified Frequentist Approach**

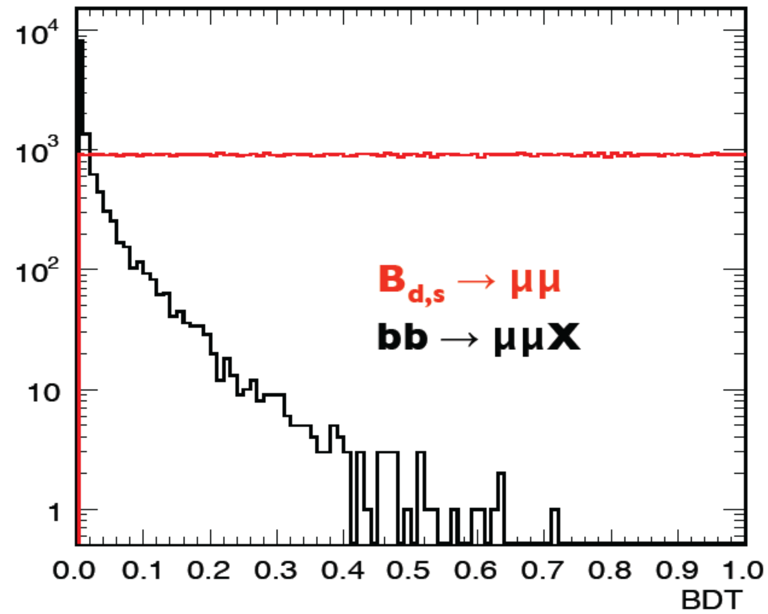
see T. Junk NIM A434, 435,1999

- Use of **control channels** to avoid dependence on simulation:
  - **Calibration of relevant variables**
  - **Normalization**

- The procedure is **blinded**: the signal region is not looked at until all the analysis is considered to be completed

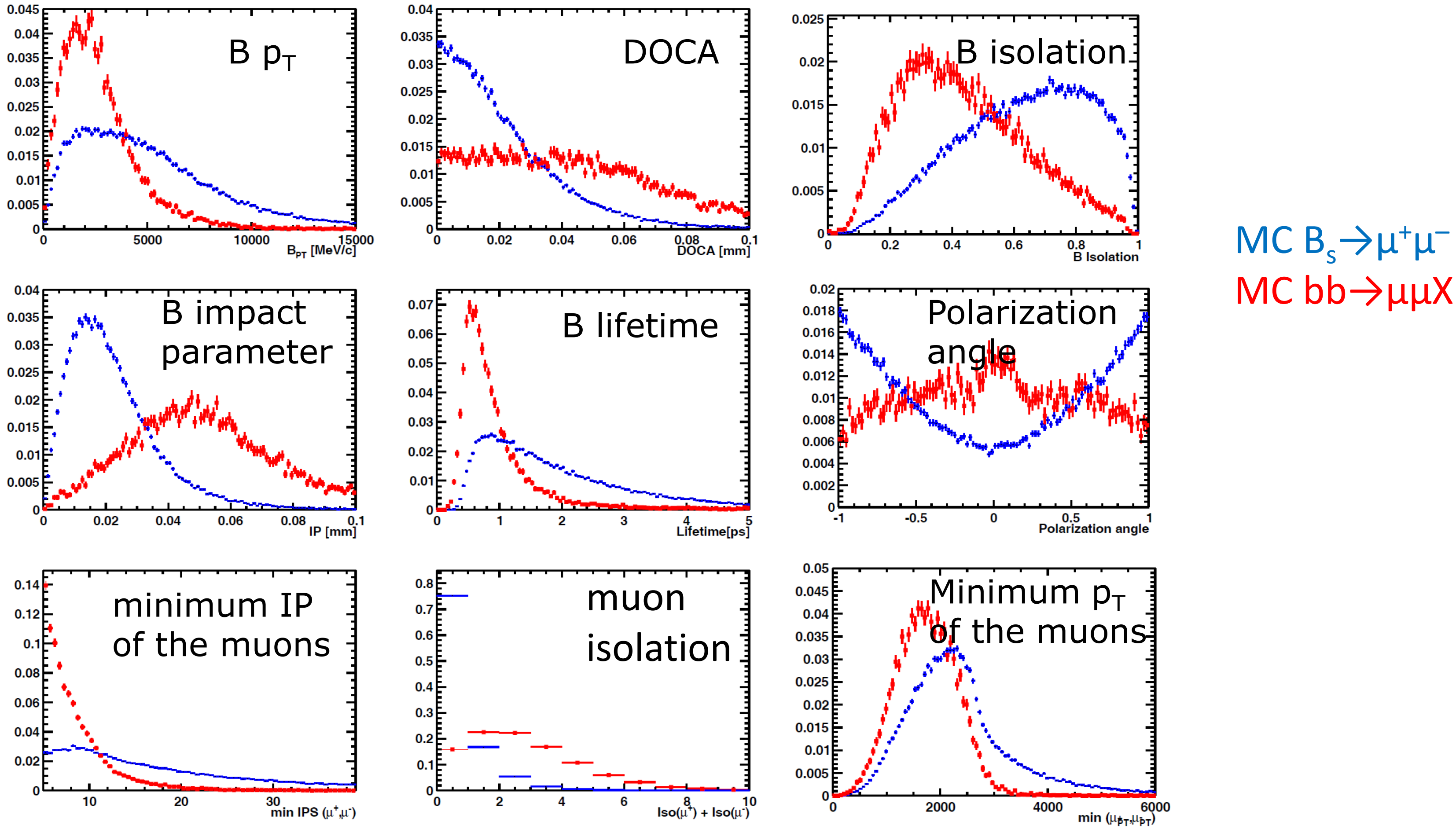
Boosted decision tree

- Main background of the analysis is combinatorial, from two real muons
  - reduce it by using MVA classifier built using 9 variables related to the **geometry** and **kinematics** of the event
    - B impact parameter, B lifetime, muon isolation, DOCA, B p_T, minimum impact parameter of the muons
    - B isolation
    - Polarization angle
    - Minimum p_T of the muons
  - Choice of variables to avoid correlation with invariant mass
  - Optimization and training on MC, using $B_s \to \mu^+\mu^-$ and $bb \to \mu\mu X$



Boosted decision tree variables

- Discrimination power of the variables combined in the BDT:



Analysis in LHCb

→ Calibration and normalization

Calibration and normalization

- **Signal** is distributed in several bins of a 2D space, so the fraction of signal in each bin is needed. That means calibrating both:
  - **Invariant mass:** Can be calibrated, e.g., with fit of $B \rightarrow hh$ line shape or from charmonium and bottonium resonances
  - **Boosted decision tree**: $B \rightarrow hh$ triggered independently of signal (event triggered by the other B)

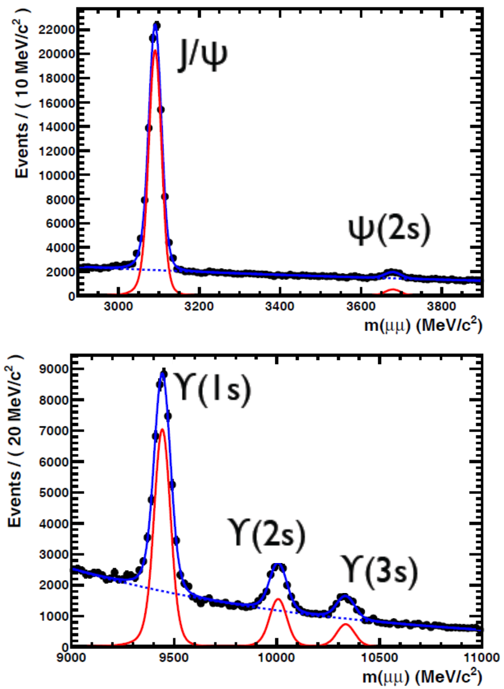
- The amount of **background** expected in each bin of the 2D space is required as well. Obtained from data!

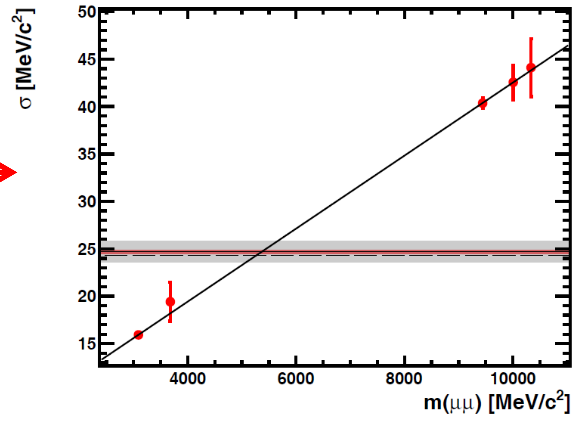
- **Overall normalization** to control channels also crucial to convert the amount of signal into a branching ratio

Invariant mass calibration

- The invariant mass is modeled with a Crystal Ball
  - Mean: obtained from exclusive $B_s \rightarrow K^+K^-$ and $B \rightarrow K^+\pi^-$
  - Resolution: obtained from interpolation of the σ's of dimuon resonances (J/ψ, ψ(2s), Υ's), crosschecked with inclusive and exclusive $B_{d,s} \rightarrow h^+h^-$

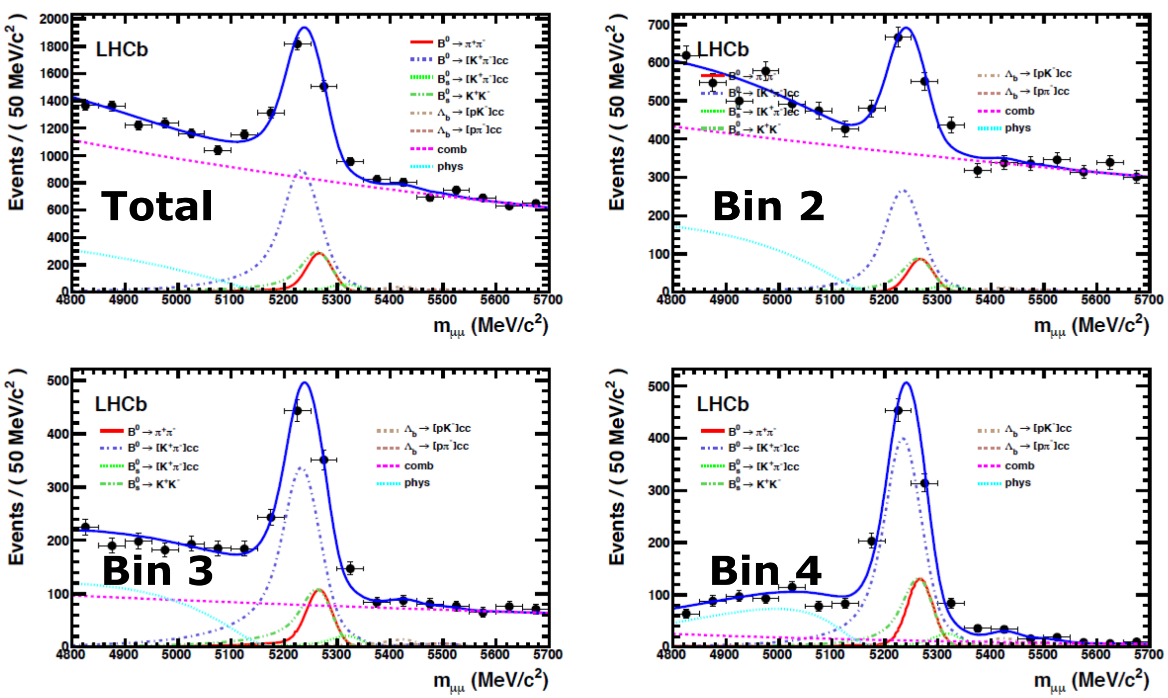
	B_d	B_s
σ (MeV/c^2)	24.3±0.2(sta.)±1.0(sys.)	24.6±0.2(sta.)±1.0(sys.)





BDT signal calibration (I)

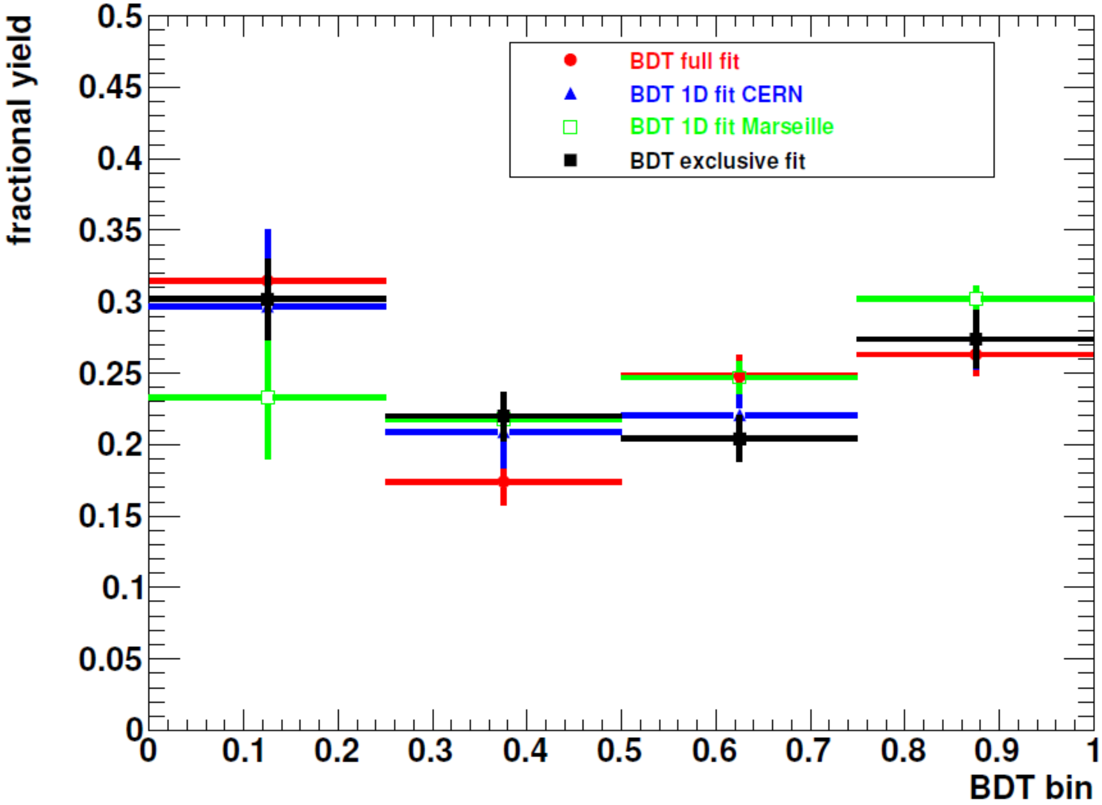
- B → hh mimics the signal, but the trigger biases the sample: use events in which B → hh is not responsible for the trigger
  - Complexity of the fit: different channels: $B_{d,s} \rightarrow hh'$ h=K,π



	$N(B_{d,s} \rightarrow hh)$
Total	4386 ± 198
Bin 2 (0.25<BDT<0.50)	1193 ± 92
Bin 3 (0.50<BDT<0.75)	1697 ± 72
Bin 4 (0.75<BDT<1.00)	1803 ± 63

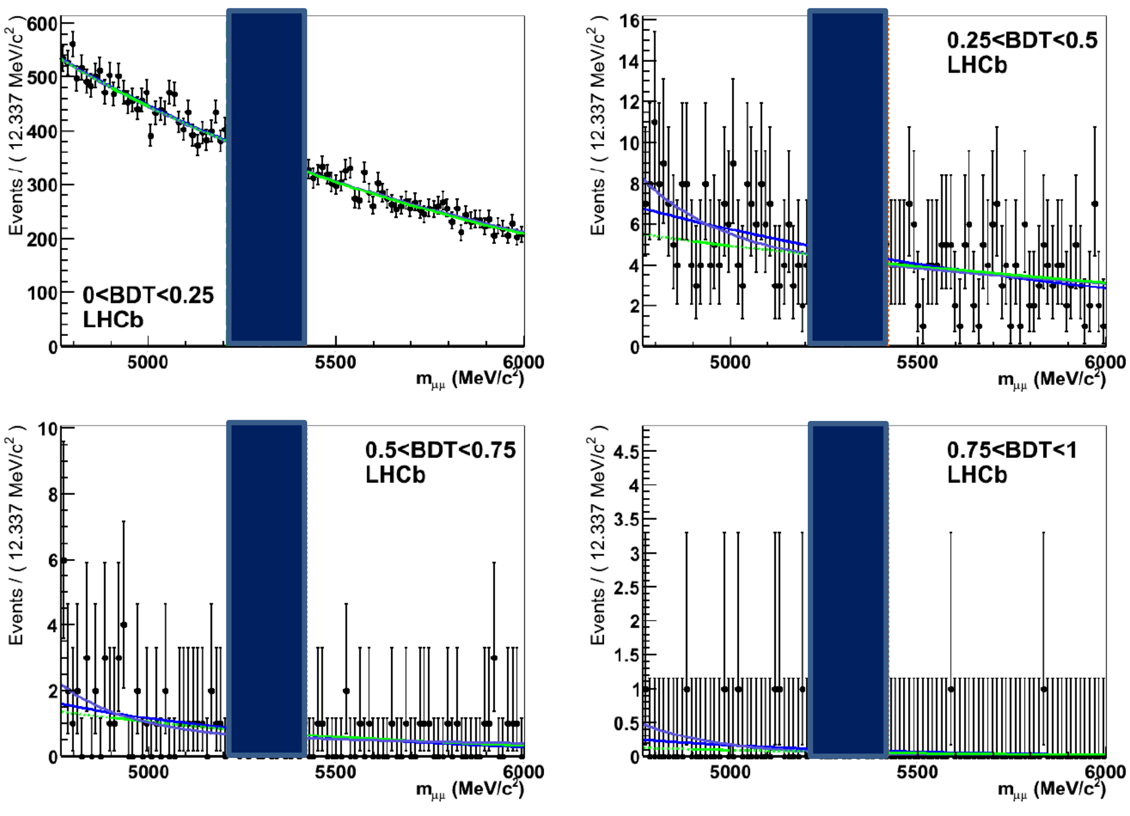
BDT signal calibration (II)

- Systematic errors: difference between different fitting models



BDT and mass background calibration

- Obtain from mass sidebands, divided in 4 equal BDT bins:
  - Systematic evaluated using different fit functions and ranges



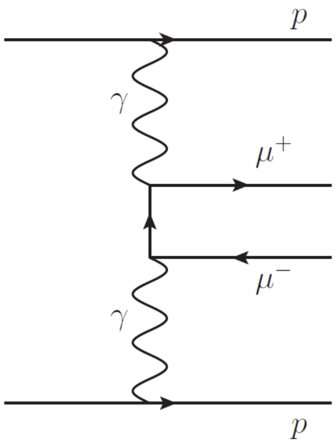
	B^0	B_s
Bin 1	3552^{+21}_{-21}	3328^{+20}_{-20}
Bin 2	$46.6^{+2.4}_{-2.3}$	$43.9^{+2.3}_{-2.2}$
Bin 3	$7.97^{+1.02}_{-0.94}$	$7.08^{+0.99}_{-0.91}$
Bin 4	$1.06^{+0.39}_{-0.32}$	$0.93^{+0.39}_{-0.31}$

Reminder:

- 4 B_s expected assuming SM (~around 1/bin)
- 21^{+14}_{-10} B_s expected assuming CDF's result (~ around $5.2^{+3.5}_{-2.5}$/bin)

Other background sources

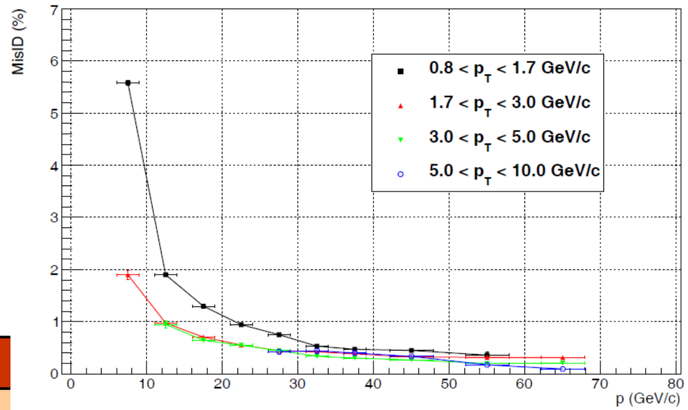
- **Dimuon from diphoton** production, in combination with other primary vertex. Features:
  - B mass: combinatorial, can reach large masses
  - BDT response ~flat, as the fake flight distance can be large
  - $B(p_T)$ is soft: killed with cut $B(p_T) > 0.5$ GeV/c



- **Peaking background** $B \rightarrow hh \rightarrow \mu\mu$ (both hadrons misidentified as muons)
  - Measure $K \rightarrow \mu$ and $\pi \rightarrow \mu$ misID using calibration sample $D^0 \rightarrow K\pi$
  - Convolute misID with the hadron p, p_T phase space of $B \rightarrow hh'$, obtained from MC
  - Total number expected in both mass windows:

B^0	B_s
5.0±1.0	1.0±0.4

Example: $\pi \rightarrow \mu$ misID rate vs p,p_T



Normalization (I)

- Needed to convert the number of events into a BR without relying on knowledge of σ_{bb}, integrated luminosity or absolute efficiencies. Uses channel with known BR!

$$\mathrm{BR} = \mathrm{BR}_{\mathrm{cal}} \times \frac{\epsilon^{\mathrm{REC}}_{\mathrm{cal}}\epsilon^{\mathrm{SEL|REC}}_{\mathrm{cal}}\epsilon^{\mathrm{TRIG|SEL}}_{\mathrm{cal}}}{\epsilon^{\mathrm{REC}}_{\mathrm{sig}}\epsilon^{\mathrm{SEL|REC}}_{\mathrm{sig}}\epsilon^{\mathrm{TRIG|SEL}}_{\mathrm{sig}}} \times \frac{f_{\mathrm{cal}}}{f_{B^0_q}} \times \frac{N_{B^0_q\to\mu^+\mu^-}}{N_{\mathrm{cal}}} = \alpha_{\mathrm{cal}} \times N_{B^0_q\to\mu^+\mu^-}$$

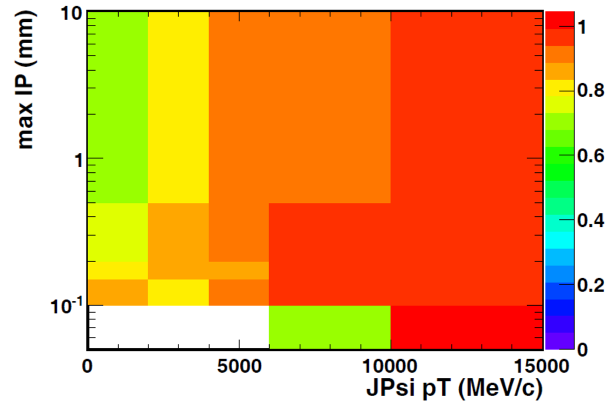
- Normalization channels:
  - $B^+\to J/\psi K^+$: same trigger, one track more, f_d/f_s (best option!)
  - $Bs\to J/\psi\phi$: same trigger, two tracks more, but large BR error
  - $B\to K^+\pi^-$: different trigger, same number of tracks, f_d/f_s

- Factors:
  - ϵ^{REC}, $\epsilon^{SEL|REC}$: estimated from MC, differences data/MC are considered as uncertainties
  - $\epsilon^{TRIG|SEL}$: estimated from data, uncertainties cancel in the ratio
  - f_s/f_d: production fraction measured at LHCb

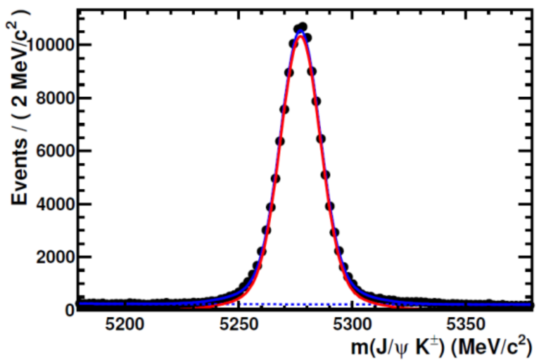
$f_s/f_d = 0.267 \pm 0.021$

Normalization (II)

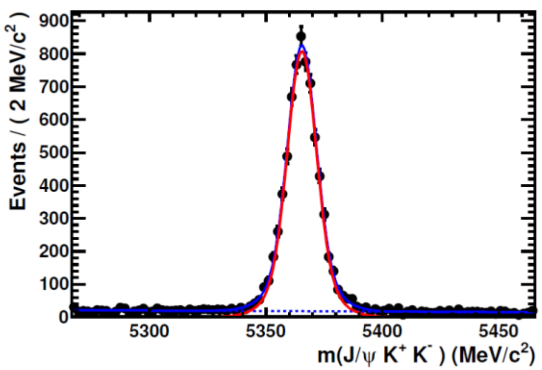
- Trigger efficiency computation:
  - Obtain $\varepsilon^{TRIG|SEL}$ from data J/ψ, and parametrize vs J/ψ p_T, max IP. Then apply MC p_T IP of $B_s \to \mu\mu$ candidates



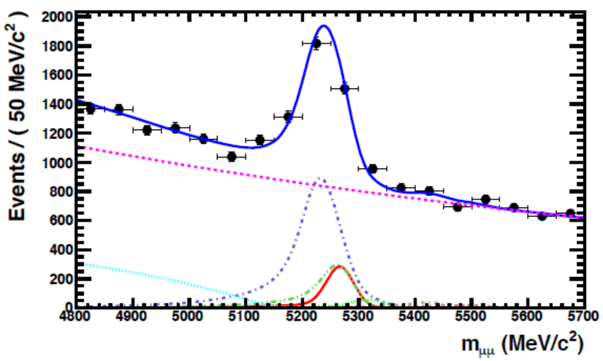
- Number of candidates



$B^+ \to J/\psi K^+$



$B_s \to J/\psi\phi$



$B \to hh$
(from which $B \to K^+\pi^-$ is obtained)

Normalization (III)

- Summary of normalization:

$$\mathrm{BR} = \mathrm{BR_{cal}} \times \frac{\epsilon^{\mathrm{REC}}_{\mathrm{cal}}\epsilon^{\mathrm{SEL|REC}}_{\mathrm{cal}}\epsilon^{\mathrm{TRIG|SEL}}_{\mathrm{cal}}}{\epsilon^{\mathrm{REC}}_{\mathrm{sig}}\epsilon^{\mathrm{SEL|REC}}_{\mathrm{sig}}\epsilon^{\mathrm{TRIG|SEL}}_{\mathrm{sig}}} \times \frac{f_{\mathrm{cal}}}{f_{B^0_q}} \times \frac{N_{B^0_q\to\mu^+\mu^-}}{N_{\mathrm{cal}}} = \boxed{\alpha_{\mathrm{cal}}} \times N_{B^0_q\to\mu^+\mu^-}$$

	$\mathcal{B}$ $(\times 10^{-5})$	$\frac{\epsilon^{REC}_{cal}\epsilon^{SEL\|REC}_{cal}}{\epsilon^{REC}_{sig}\epsilon^{SEL\|REC}_{sig}}$	$\frac{\epsilon^{TRIG\|SEL}_{cal}}{\epsilon^{TRIG\|SEL}_{sig}}$	N_{cal}	$\alpha^{cal}_{B_d\to\mu^+\mu^-}$ $(\times 10^{-10})$	$\alpha^{cal}_{B_s\to\mu^+\mu^-}$ $(\times 10^{-9})$
$B^+\to J/\psi K^+$	6.01±0.21	0.480±0.015	0.95±0.01	124518±2025	2.22±0.11	0.83±0.08
$B_s\to J/\psi\phi$	3.4±0.9	0.240±0.014	0.95±0.01	6940±93	2.95±0.84	1.10±0.30
$B\to K^+\pi^-$	1.94±0.06	0.83±0.03	0.049±0.004	6853±957	1.93±0.33	0.72±0.14

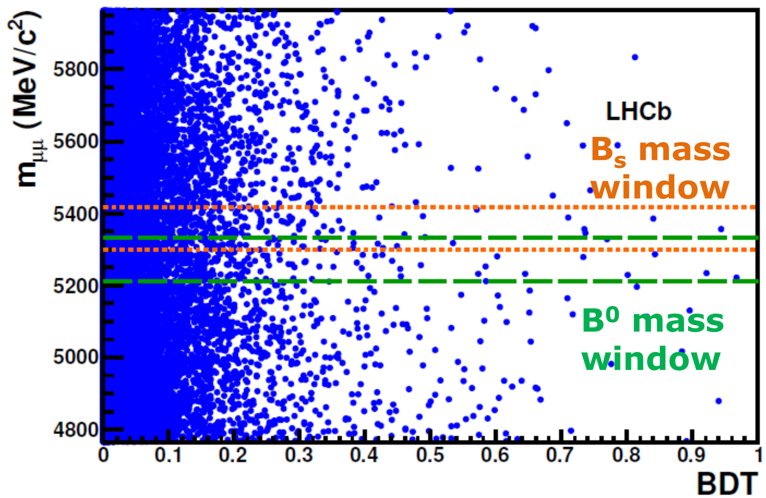
Combination of the 3 α factors →

$\alpha^{cal}_{B_d\to\mu^+\mu^-}$	$\alpha^{cal}_{B_s\to\mu^+\mu^-}$
$(2.20 \pm 0.11) \times 10^{-10}$	$(8.38 \pm 0.74) \times 10^{-10}$

Results

Limit computation

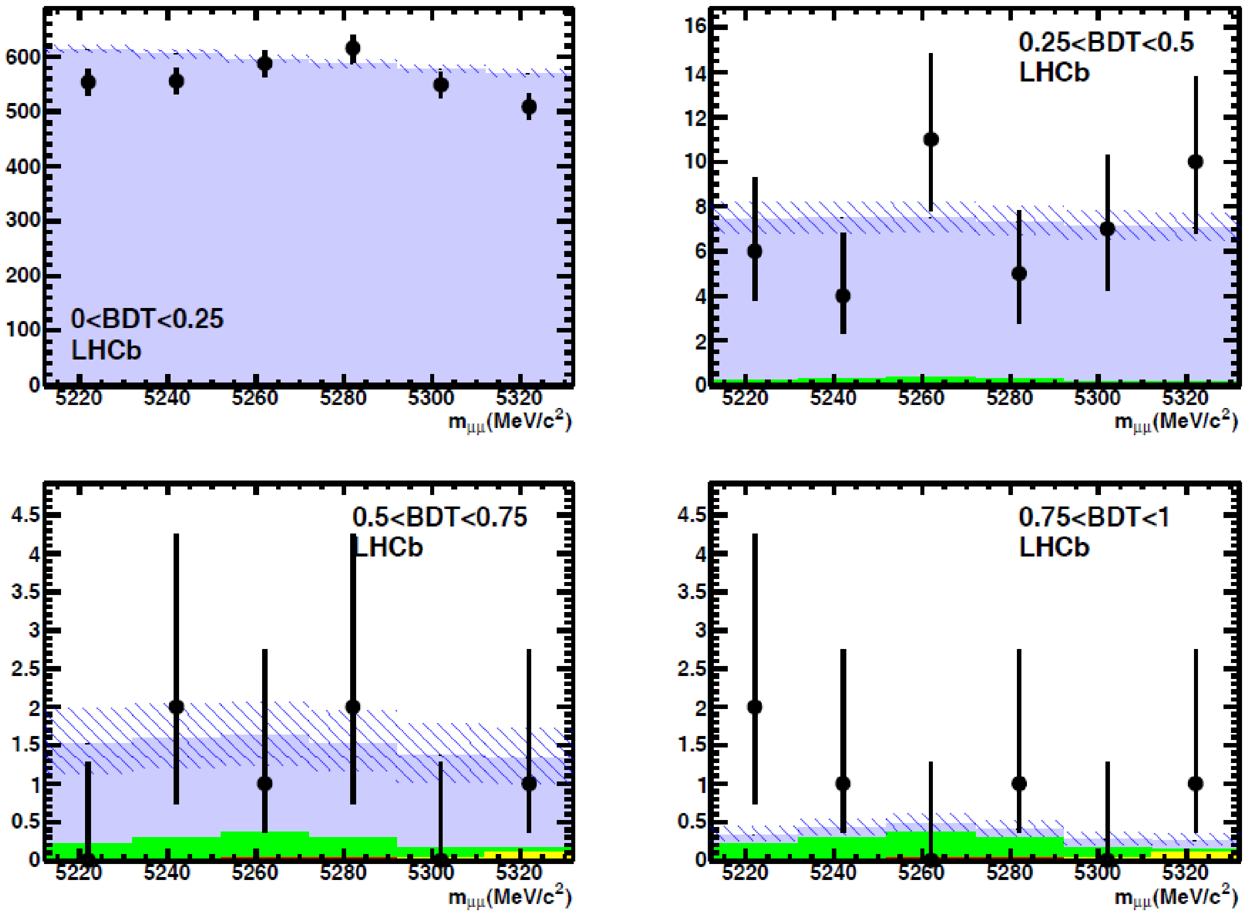
- How do we extract a limit?
  - Count the events in 4 BDT and 6 $m_{\mu\mu}$ bins
  - For each bin compute the expected signal and background yields
  - Evaluate compatibility between observed and expected with:
    - **S+Bkg hypothesis [CL_{S+B}]**
    - **Bkg only hypothesis [CL_B]**
  - **CL_S** = CL_{S+B}/CL_B compatibility with the signal hypothesis. **Used to compute the exclusion**



Distribution of selected dimuon events in the invariant mass vs BDT

Results: $B^0 \to \mu\mu$ (I)

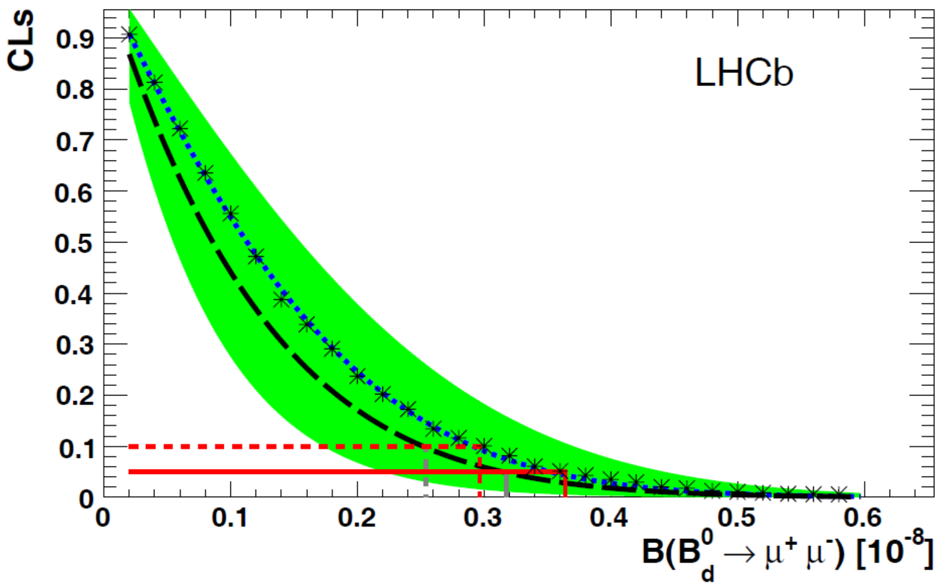
- B^0 mass distribution in BDT bins



- Combinatorial background
- $B \to hh \to \mu\mu$ (misID)
- SM $B \to \mu\mu$
- SM $B_s \to \mu\mu$

Results: $B^0 \to \mu\mu$ (II)

- CL_s vs BR
  - dashed: expected bkg
  - green band: expected within ±1σ
  - stars: observation

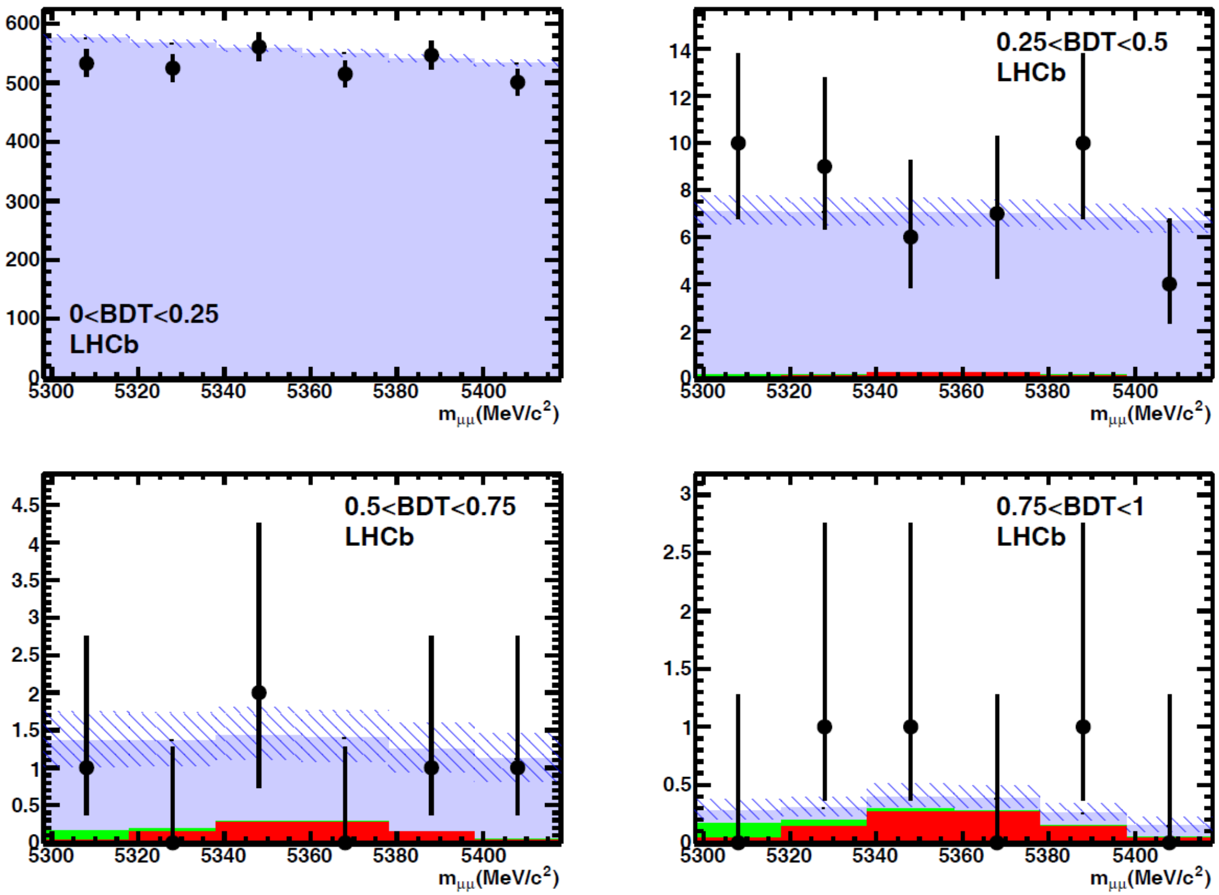


World best measurement!

	at 90% CL	at 95% CL	CL_B
Expected BR limit (bkg. only hypothesis)	2.5×10^{-9}	3.2×10^{-9}	-
Observed BR limit	3.0×10^{-9}	3.6×10^{-9}	0.68

Results: $B_s \to \mu\mu$ (I)

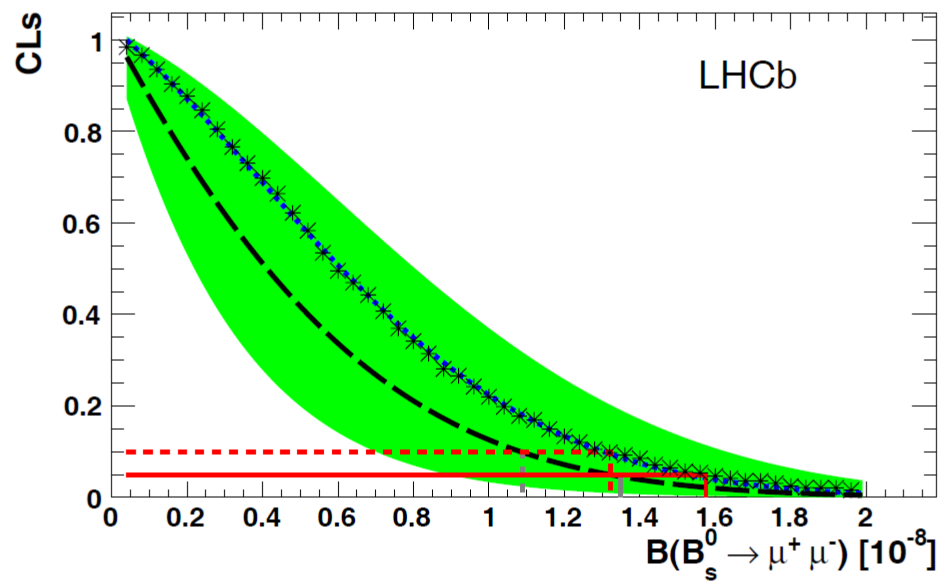
- B_s mass distribution in BDT bins



- Combinatorial background
- $B \to hh \to \mu\mu$ (misID)
- SM $B_s \to \mu\mu$

Results: $B_s \to \mu\mu$ (II)

- CL_s vs BR
  - dashed: expected bkg+SM
  - green band: expected within ±1σ
  - stars: observation



Compatibility with the background-only hypothesis: **p-value = 1-CL_B = 5 %**
Observing the ~SM BR at ~2σ level

World best measurement!

	at 90% CL	at 95% CL	CL_B
Expected BR limit (bkg. + SM hypothesis)	1.09×10^{-8}	1.35×10^{-8}	-
Observed BR limit	1.32×10^{-8}	1.58×10^{-8}	0.95

Results: 2010+2011

- Results can be further improved by adding the 37 pb^{-1} of the 2010 analysis:
  - 2010+2011 combined data result:

	at 90% CL	at 95% CL
Observed BR($B^0 \rightarrow \mu\mu$) limit	2.6×10^{-9}	3.2×10^{-9}
Observed BR($B_s \rightarrow \mu\mu$) limit	1.2×10^{-8}	1.4×10^{-8}

Conclusions

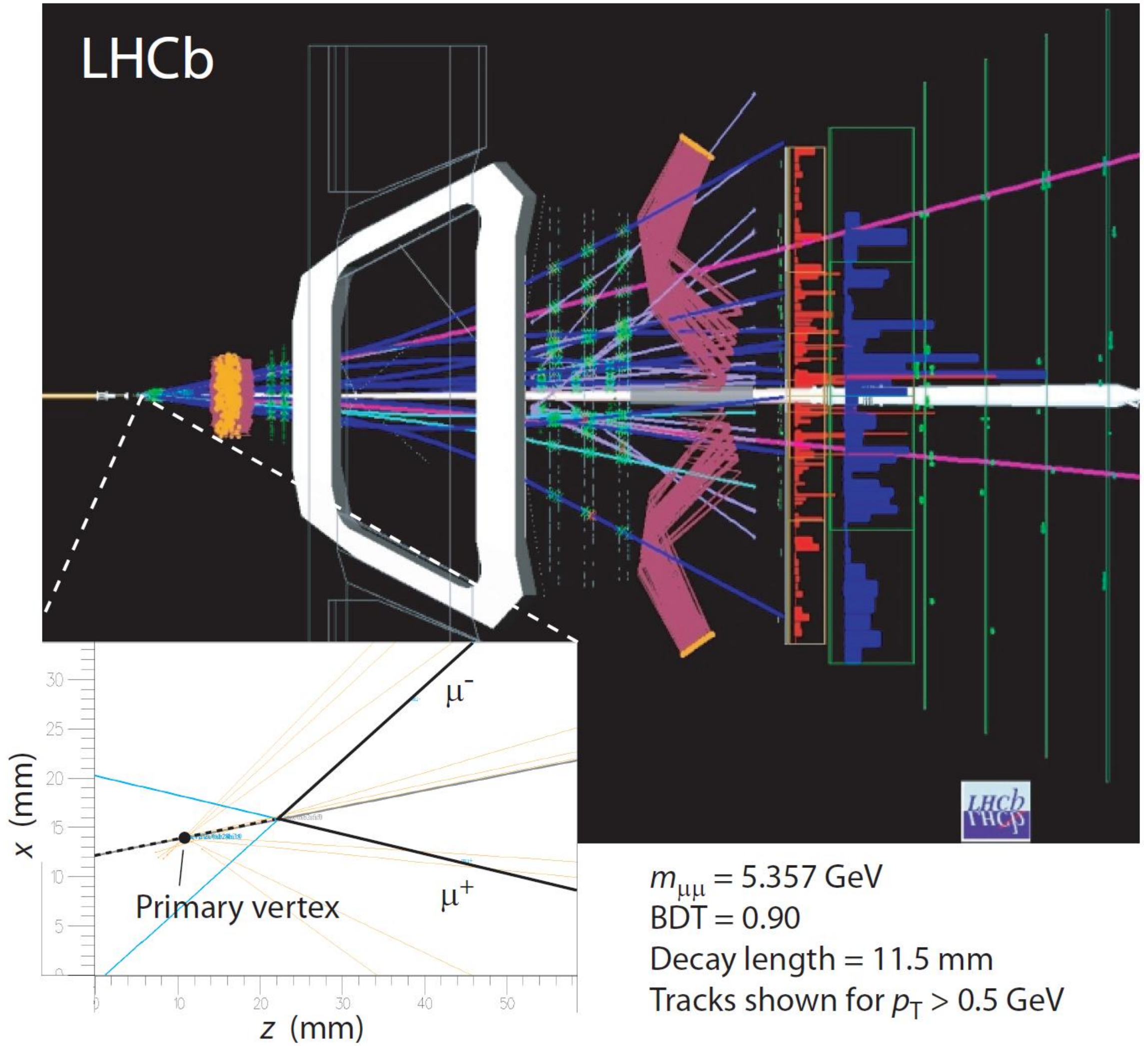
Conclusions

- LHCb has measured **world best upper limits** in the branching ratios of both $B^0 \rightarrow \mu^+\mu^-$ and $B_s \rightarrow \mu^+\mu^-$
- For $\mathbf{B_s \rightarrow \mu^+\mu^-}$, our results are compatible with background+SM expectance, and **do not confirm the excess** claimed at the beginning of summer **by CDF**.
- When combining 2010 and 2011 data LHCb observes, at 95% CL:

$$BR(B^0 \rightarrow \mu\mu) < 3.2 \times 10^{-9}$$
$$BR(B_s \rightarrow \mu\mu) < 1.4 \times 10^{-8}$$

- These results have lots of **theory implications**, that will be discussed in Diego Martinez talk.

The event

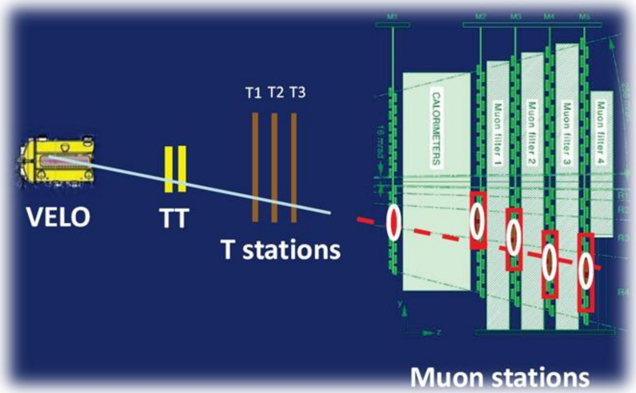


LHCb
μ⁻
μ⁺
Primary vertex
x (mm)
z (mm)
$m_{\mu\mu}$ = 5.357 GeV
BDT = 0.90
Decay length = 11.5 mm
Tracks shown for p_T > 0.5 GeV

Backup

Muon PID

- Particles with associated hits after extrapolation to the muon chambers are flagged as muons. Some of them might not be actual muons (misidentification).

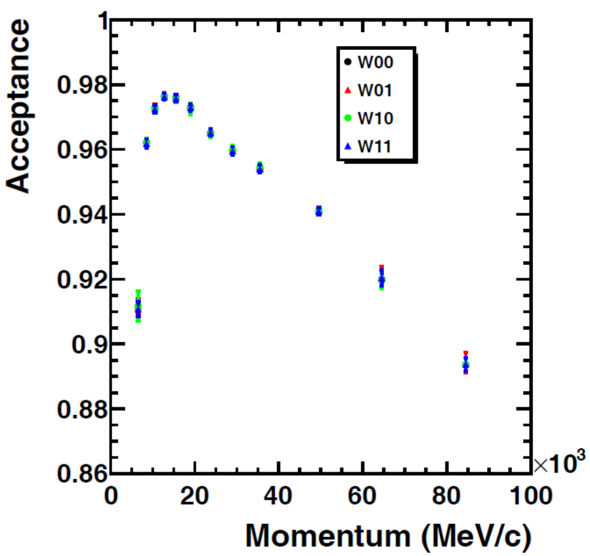


- Important for selection efficiency and also to study the presence of possible peaking backgrounds:

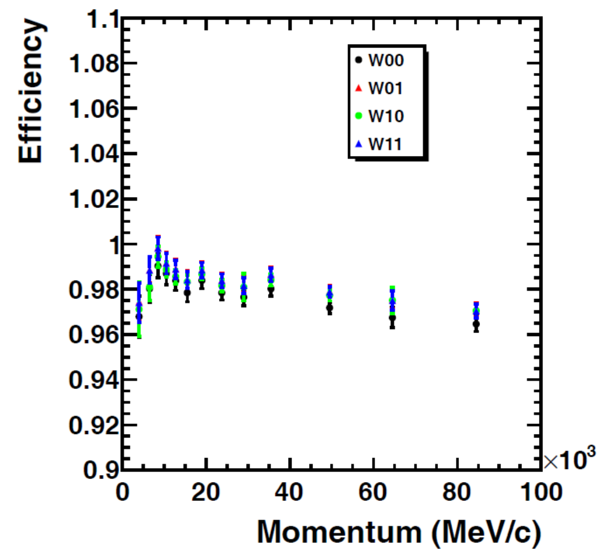
$B \to hh \to \mu\mu$

MuonID acceptance and efficiency

- Muon ID acceptance and efficiency
  - Computed from data: use a probe-μ from J/ψ from B, $B^+ \to J/\psi K^+$
    - probe-μ (a track without muon, calorimeter information and not selected by the trigger)



$\alpha = 0.94 \pm 0.01$ (sys)



$\varepsilon = 0.984 \pm 0.001$ (stat) ± 0.005 (sys)

Normalization systematics

- Normalization systematic uncertainties summary

	$B^+ \to J/\psi K^+$	$B_s \to J/\psi\phi$	$B\to K^+\pi^-$
BR	3%	26%	3%
Reco. eff	2%	4%	3%
muonID	1%	1%	2%
Sel. eff	1%	3%	1%
Trigger eff.	1%	1%	8%
f_d/f_s	8%	-	8%
Yield	2%	6%	7%

Systematic: reconstruction efficiencies
→ 3/4 tracks ratio data/MC

$$\frac{\epsilon^{\text{REC,data}}_{B+\to J/\psi K+} / \epsilon^{\text{REC,data}}_{B^0_d\to J/\psi K^{*0}}}{\epsilon^{\text{REC,MC}}_{B+\to J/\psi K+} / \epsilon^{\text{REC,MC}}_{B^0_d\to J/\psi K^{*0}}} = 1.035 \pm 0.04 \pm 0.06$$

Systematic: selection efficiencies
→ Smeared MC/normal MC

channel	normal	smeared
$B^0_s \to \mu^+\mu^-$	51.8%	50.8%
$B^0 \to K^+\pi^-$	57.3%	56.5%
$B^+ \to J/\psi K^+$	43.4%	42.6%
$B^0_s \to J/\psi(\mu^+\mu^-)\phi(KK)$	34.1%	33.4%

Normalization – B → Kπ

- Normalize to B→K⁺π is equivalent to normalize to B → hh TIS (trigger independent of signal)!

using

$$\epsilon_{hh}^{TRIG/SEL} = \epsilon_{hh}^{TIS/SEL} \frac{N_{hh}}{N_{hh}^{TIS}}$$

$$\alpha = \mathcal{B}(B^0 \to K^+\pi^-) \times \frac{f_N}{f_{sig}} \frac{\epsilon_{hh}^{REC}}{\epsilon_s^{REC}} \frac{\epsilon_{hh}^{SEL/REC}}{\epsilon_s^{SEL/REC}} \frac{\epsilon_{hh}^{TIS/SEL}}{\epsilon_s^{TRIG/SEL}} \frac{1}{f_{B^0\to K^+\pi^-} N_{hh}^{TIS}}$$

  - The values obtained with the trigger studies and the full fit to the B → hh mass:

$$\frac{\epsilon_{hh}^{TIS/SEL}}{\epsilon_{B_s^0\to\mu^+\mu^-}^{TRIG|SEL}} = (5.1 \pm 0.04)\%$$

$N^{TIS}_{Bhh} = 6853 \pm 957$

$f_{B\to K\pi} = 0.605 \pm 0.027$

f_s/f_d at LHCb

- Previous result used the HFAG average from LEP/Tevatron.
- This ratio is now evaluated at LHCb
  - f_s/f_d is measured at LHCb with hadronic decays $B^0 \to D^\pm K^\mp$ or $B^0 \to D^\pm \pi^\mp$ and $B_s \to D_s{}^\pm \pi^\mp$

$$f_s/f_d = 0.253 \pm 0.017^{\text{stat}} \pm 0.017^{\text{syst}} \pm 0.020^{\text{theor}}$$

LHCb-PAPER-2011-006

  - And semileptonic decays

$$\left(\frac{f_s}{f_d}\right)_{\text{sl}} = 0.268 \pm 0.008(\text{stat})^{+0.022}_{-0.020}(\text{syst}),$$

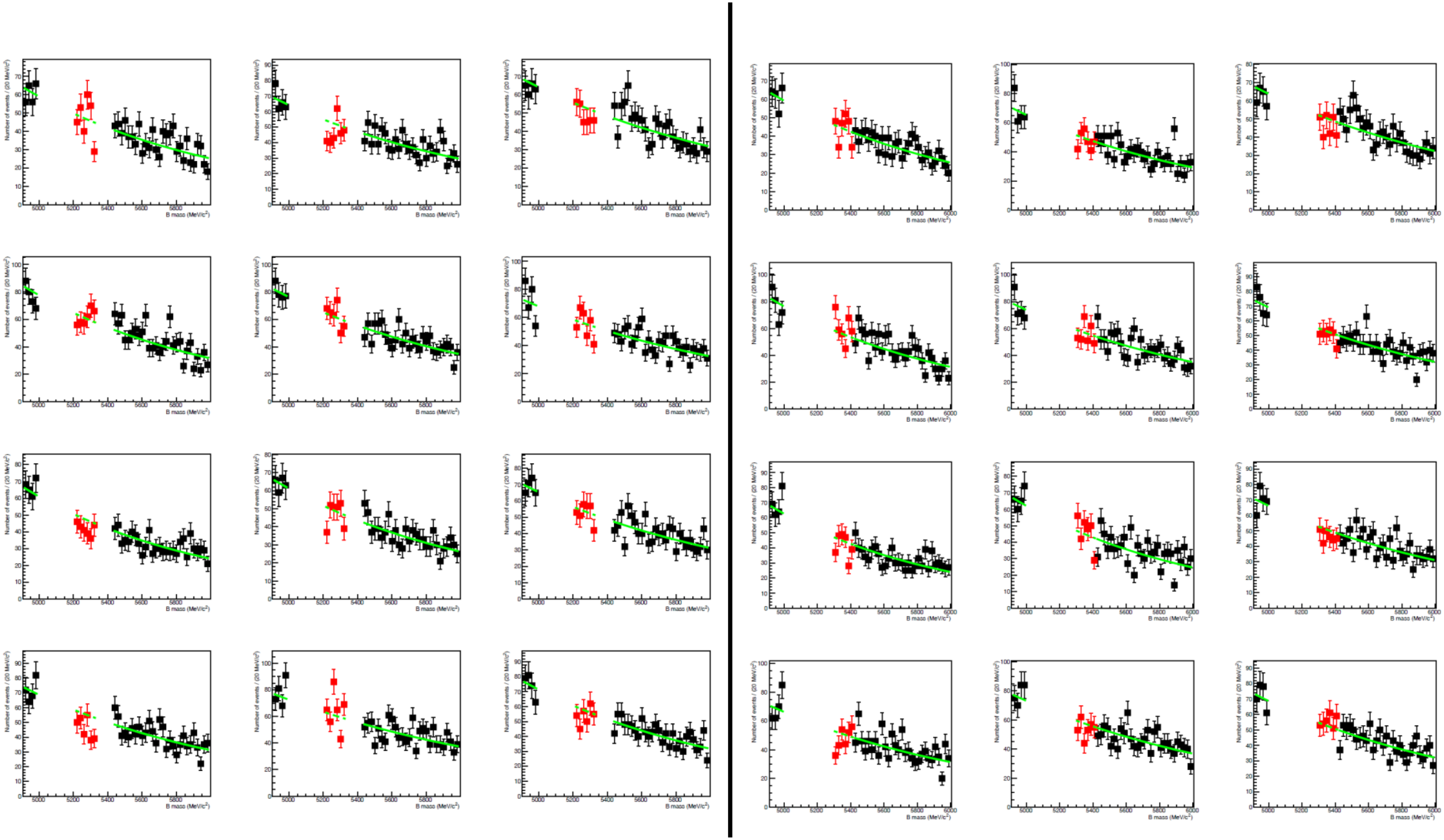
LHCb-CONF-2011-028

  - The average is computed:

$$\langle f_s/f_d \rangle = 0.267^{+0.021}_{-0.020}$$

LHCb-CONF-2011-034

First BDT Bin



CMS result

- Result features:
  - 1.1 fb^{-1}, larger pile-up but not affecting the search
  - trigger p_T > 2 GeV/c, σ ~ 24 MeV/c^2
  - cut based analysis, use B^+ as normalization channel

Table 1: The event selection efficiencies for signal events ε_{tot}, the SM-predicted number of signal events N^{exp}_{signal}, the expected number of combinatorial background events N^{exp}_{comb} and peaking background events N^{exp}_{peak}, and the number of observed events N_{obs} in the barrel and endcap channels for $B^0_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$.

	Barrel		Endcap	
	$B^0 \to \mu^+\mu^-$	$B^0_s \to \mu^+\mu^-$	$B^0 \to \mu^+\mu^-$	$B^0_s \to \mu^+\mu^-$
ε_{tot}	$(3.6 \pm 0.4) \times 10^{-3}$	$(3.6 \pm 0.4) \times 10^{-3}$	$(2.1 \pm 0.2) \times 10^{-3}$	$(2.1 \pm 0.2) \times 10^{-3}$
N^{exp}_{signal}	0.065 ± 0.011	0.80 ± 0.16	0.025 ± 0.004	0.36 ± 0.07
N^{exp}_{comb}	0.40 ± 0.23	0.60 ± 0.35	0.53 ± 0.27	0.80 ± 0.40
N^{exp}_{peak}	0.25 ± 0.06	0.07 ± 0.02	0.16 ± 0.04	0.04 ± 0.01
N_{obs}	0	2	1	1

CMS BR(B→μμ) < 4.6 10^{-9} 95% CL
LHCb BR(B→μμ) < 3.2 10^{-9} 95% CL

CMS BR(B_s→μμ) < 1.9 10^{-8} 95% CL
LHCb BR(B_s→μμ) < 1.4 10^{-8} 95% CL