$B_{s,d} \to \mu^+\mu^-$ AT THE LHC

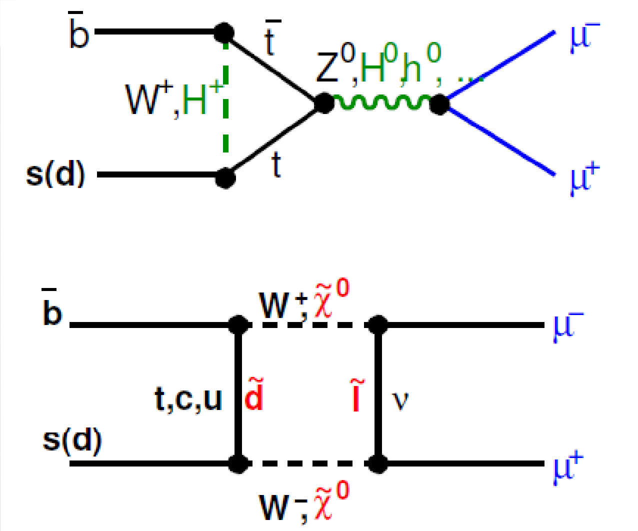
ROBERTO COVARELLI (*UNIV. OF ROCHESTER*)

VIII MEETING ON B-PHYSICS
GENOVA, 6-7 FEBRUARY 2012

OUTLINE

- Physics motivations
- Detectors and data samples
- Data analysis
  - Common analysis techniques
  - $B_{s,d} \to \mu^+\mu^-$ at CMS
  - $B_{s,d} \to \mu^+\mu^-$ at LHCb
- Combination of the two LHC limits
- Future perspectives

THEORETICAL MOTIVATIONS

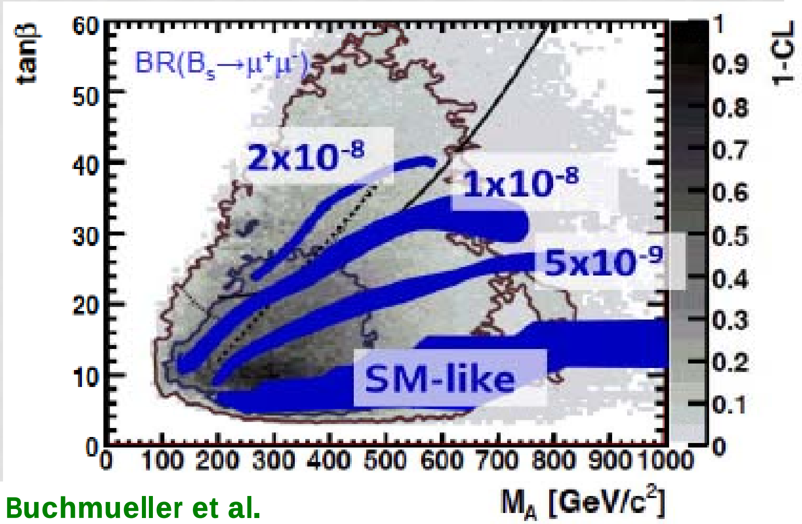


- Decays highly suppressed in SM
  - Effective FCNC decay
  - Helicity suppression
  - Overall Cabibbo suppression

$$\begin{aligned}\mathcal{B}(B_s^0 \to \mu^+\mu^-) &= (3.2 \pm 0.2) \times 10^{-9} \\ \mathcal{B}(B^0 \to \mu^+\mu^-) &= (1.0 \pm 0.1) \times 10^{-10}\end{aligned}$$

Buras, Isidori & Paradisi Phys.Lett. B694 402 (2010)

- Enhancements in several BSM physics scenarios:
  - CMSSM $\frac{\mathcal{B}_{MSSM}}{\mathcal{B}_{SM}} \propto \frac{(\tan\beta)^6}{M_A{}^4}$ Isidori & Paradisi Phys.Lett. B639 499 (2006)
  - NUHM1 →



Buchmueller et al. Eur.Phys.J. C64, 391 (2009)

EXPERIMENTAL MOTIVATIONS

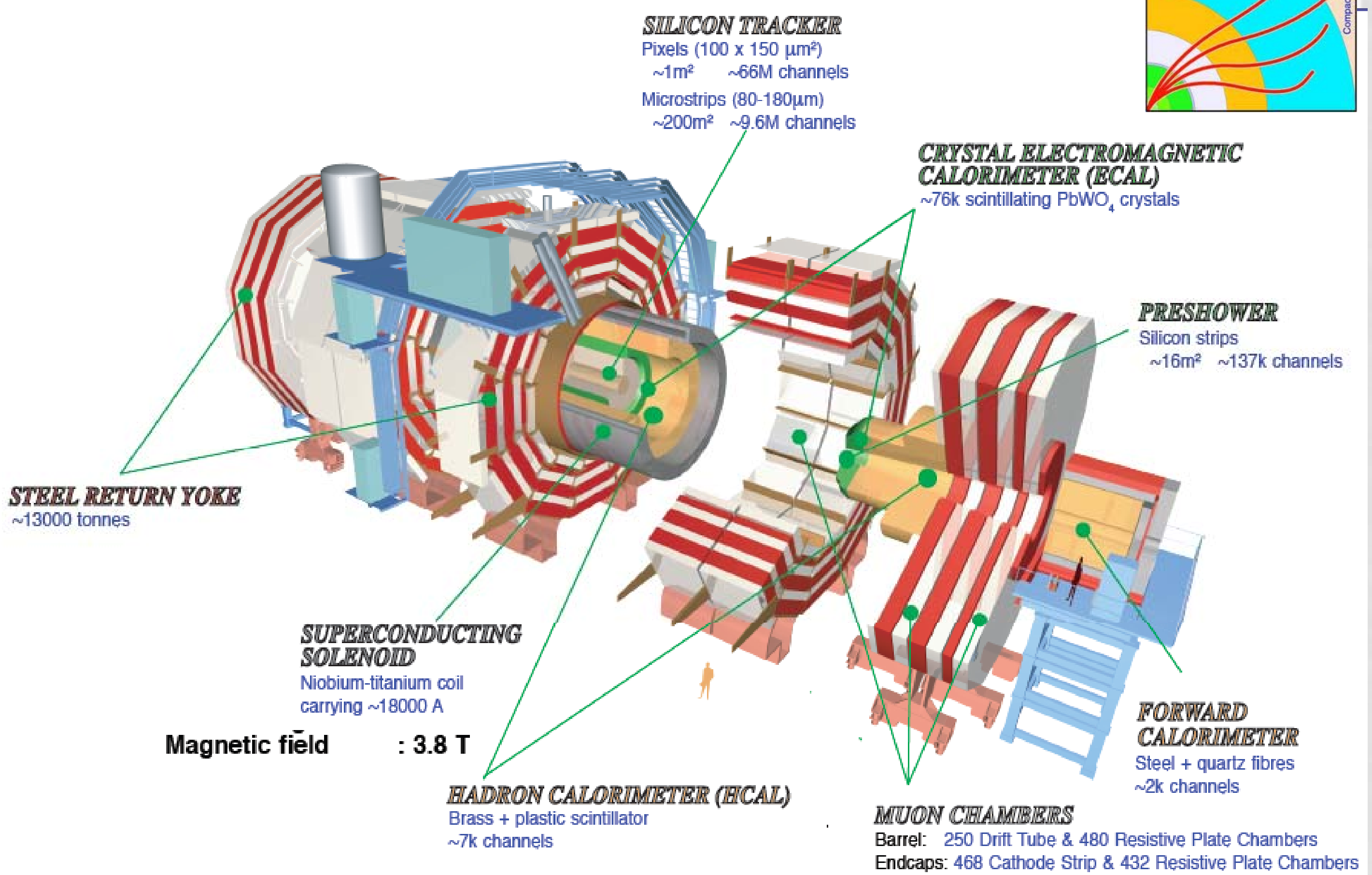
- Ultra-clean experimental signature
- Small BRs → needing huge $b\bar{b}$ samples
- CKM-hierarchy favoring B_s over B^0 ($\boxed{\mathrm{BR}(B_d \to \mu^+\mu^-)/\mathrm{BR}(B_s \to \mu^+\mu^-) \approx |V_{td}/V_{ts}|^2}$ under generic MFV assumption)

→ Hadron machines outperforming B-factories

- Before LHC data, world best results from Tevatron:
  - D0: $\mathrm{BR}(B_s \to \mu\mu) < 5.1 \times 10^{-8}$ @ 95% CL ($6.1\ \mathrm{fb}^{-1}$)
  - CDF: $\mathrm{BR}(B_s \to \mu\mu) < 4.0 \times 10^{-8}$ @ 95% CL ($6.9\ \mathrm{fb}^{-1}$) but observing an excess in data compatible with a measurement of $\mathrm{BR}(B_s \to \mu\mu) = (1.8_{-0.9}^{+1.1}) \times 10^{-8}$ [5.6 x SM]

THE CMS DETECTOR

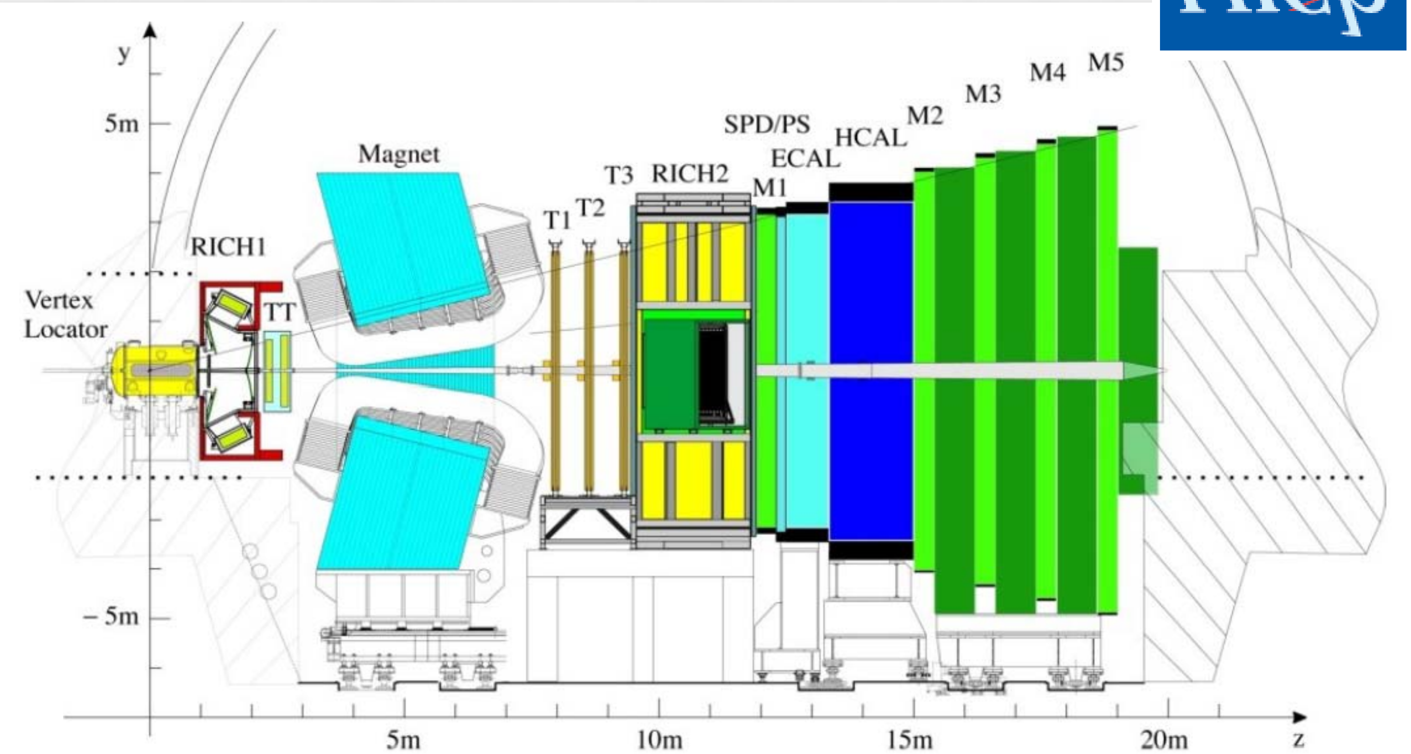
Muon reconstruction from combined information of silicon tracker and external detectors (DT, CSC, RPC)



- Muon triggers:
  - segments found in muon detectors (Level 1)
  - fast regional track reconstruction (High-Level Trigger)

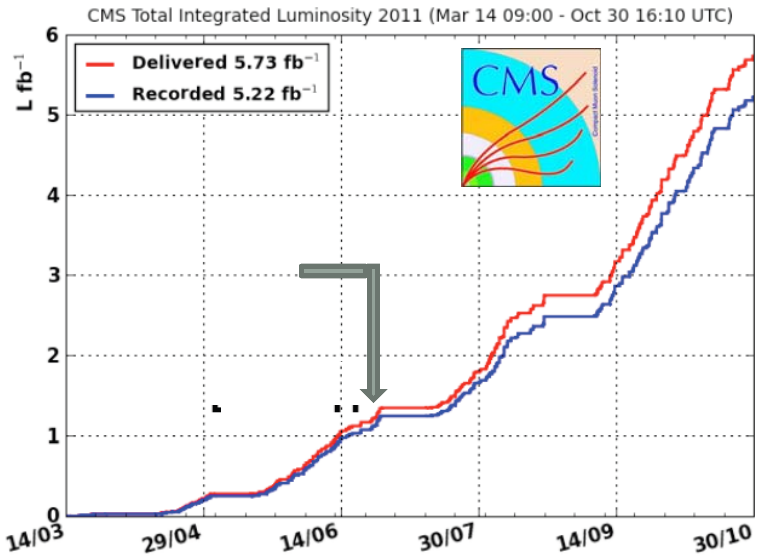
THE LHC-B DETECTOR

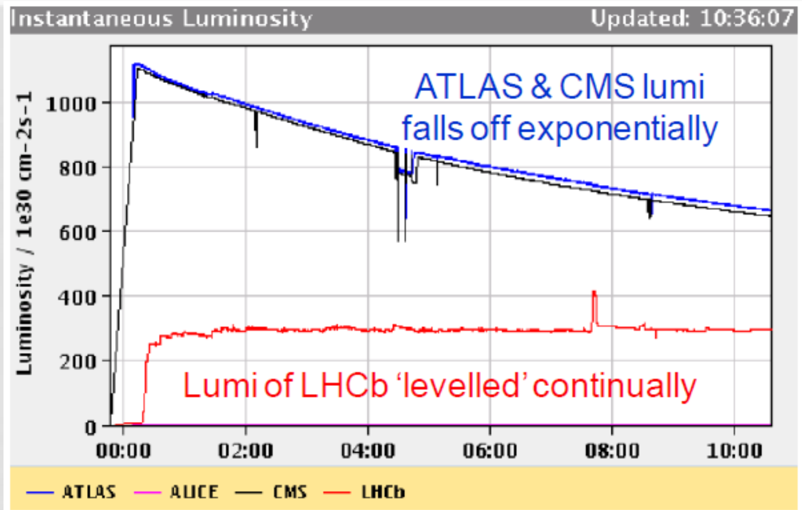
Muon reconstruction from tracking stations (silicon + straw tubes) and muon chambers (MWPC)

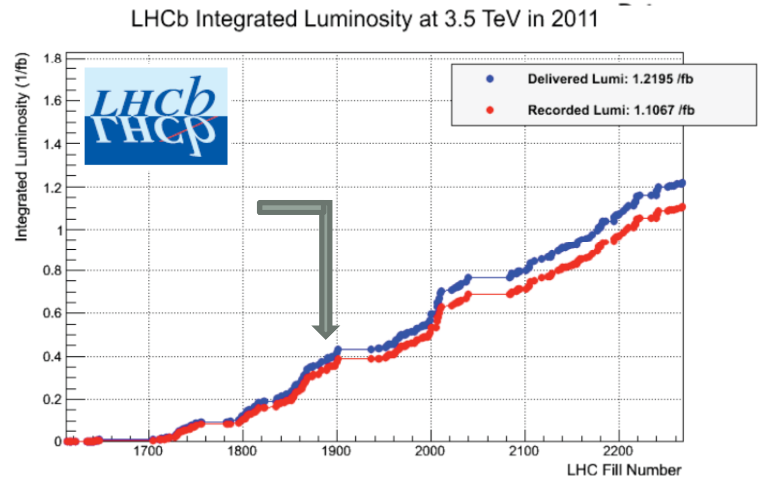


- Muon and hadron triggers (L0 + 2 HLT steps):
  - Low p_T reach thanks to forward geometry and high output rate (3 kHz)
  - Vertexing, displacement and/or invariant mass requirements to select J/ψ and $B_{s,d}$

INTEGRATED LUMINOSITY







- Used for analyses presented here:
  - 1.14 fb^{-1} (CMS)
  - 0.37 fb^{-1} (LHCb)

    plus combination with independent 2010 analysis (37 pb^{-1})

COMMON ANALYSIS TECHNIQUES

- Blind analyses
- Use of "normalization channels" to
  - remove large uncertainties on luminosity and σ_{bb}
  - reduce efficiency systematic uncertainties in the ratio

$$\mathcal{B}(B_q^0 \to \mu^+\mu^-) = \mathcal{B}_{\rm norm} \times \frac{\epsilon_{\rm norm}^{\rm REC}\epsilon_{\rm norm}^{\rm SEL|REC}\epsilon_{\rm norm}^{\rm TRIG|SEL}}{\epsilon_{\rm sig}^{\rm REC}\epsilon_{\rm sig}^{\rm SEL|REC}\epsilon_{\rm sig}^{\rm TRIG|SEL}} \times \frac{f_{\rm norm}}{f_{B_q^0}} \times \frac{N_{B_q^0\to\mu^+\mu^-}}{N_{\rm norm}} = \alpha_{B_q^0\to\mu^+\mu^-} \times N_{B_q^0\to\mu^+\mu^-}$$

  - LHCb: normalization factor (α) determined on 3 channels and averaged
    - $B^+ \to J/\psi\ (\mu^+\mu^-)\ K^+$
    - $B_s \to J/\psi\ (\mu^+\mu^-)\ \phi\ (K^+K^-)$
    - $B^0 \to K^-\pi^+$

    same muon selection and trigger, slightly different kinematics (extra tracks) (for the first two channels)

    same kinematics, different trigger (for $B^0 \to K^-\pi^+$)
  - CMS: only uses $B^+ \to J/\psi\ K^+$ ($B_s \to J/\psi\ \phi$ used as control sample to estimate systematics on analysis selection, impossible to trigger on hadron modes)

COMMON UNCERTAINTIES

$$\mathcal{B}(B_q^0 \to \mu^+\mu^-) = \mathcal{B}_{\rm norm} \times \frac{\epsilon_{\rm norm}^{\rm REC}\epsilon_{\rm norm}^{\rm SEL|REC}\epsilon_{\rm norm}^{\rm TRIG|SEL}}{\epsilon_{\rm sig}^{\rm REC}\epsilon_{\rm sig}^{\rm SEL|REC}\epsilon_{\rm sig}^{\rm TRIG|SEL}} \times \frac{f_{\rm norm}}{f_{B_q^0}} \times \frac{N_{B_q^0\to\mu^+\mu^-}}{N_{\rm norm}} = \alpha_{B_q^0\to\mu^+\mu^-} \times N_{B_q^0\to\mu^+\mu^-}$$

	$\mathcal{B}$ $(\times 10^{-5})$
$B^+ \to J/\psi(\mu^+\mu^-)K^+$	5.98 ± 0.22
$B_s^0 \to J/\psi(\mu^+\mu^-)\phi(K^+K^-)$	3.4 ± 0.9
$B^0 \to K^+\pi^-$	1.94 ± 0.06

common, but small

- CMS used a less recent value from PDG to quote the limits → updated to LHCb f_s/f_d for the combination

- LHCb result uses its own measurements of f_s/f_d
  - $B^0 \to D^-K^+$ and $B_s \to D_s^-\pi^+$

$$f_s/f_d = 0.250 \pm 0.024_{stat} \pm 0.017_{syst} \pm 0.017_{theo}$$

  - $B^0 \to D^-\pi^+$ and $B_s \to D_s^-\pi^+$

$$f_s/f_d = 0.256 \pm 0.014_{stat} \pm 0.019_{syst} \pm 0.026_{theo}$$

  - semileptonic decays $\frac{f_s}{f_d + f_u} = 0.134 \pm 0.004^{+0.011}_{-0.010}$

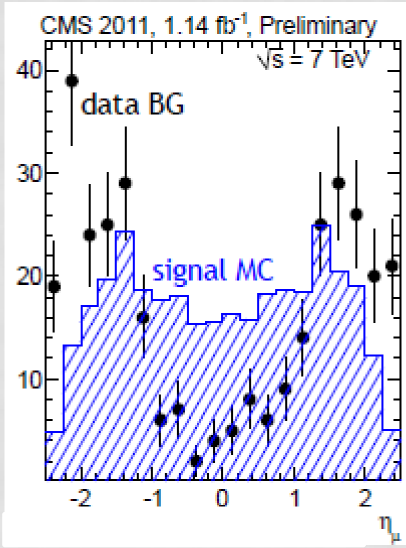
- Average: $\boxed{f_s/f_d = 0.267^{+0.021}_{-0.020}}$

CMS ANALYSIS

CMS Collaboration
Phys. Rev. Lett. 107, 191802 (2011)

ANALYSIS PRINCIPLES

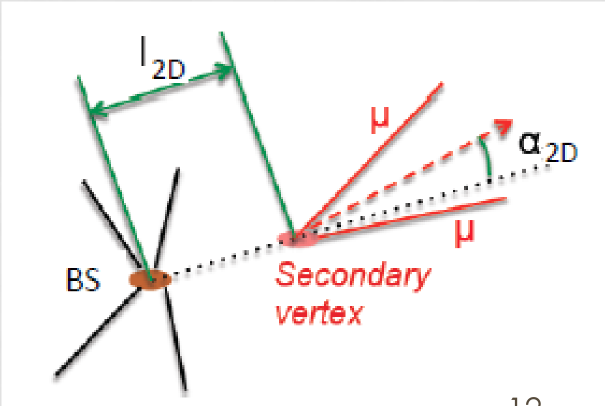
- Main backgrounds:
  - Collimated muons from two semileptonic B decays (→ gluon splitting)
  - One muon from semileptonic B decay plus one misidentified hadron
  - Rare decays
    - Peaking (e.g. $B_s \to K^+ K^-$)
    - Non-peaking (e.g. $B_s \to K^+ \mu^- \nu$)



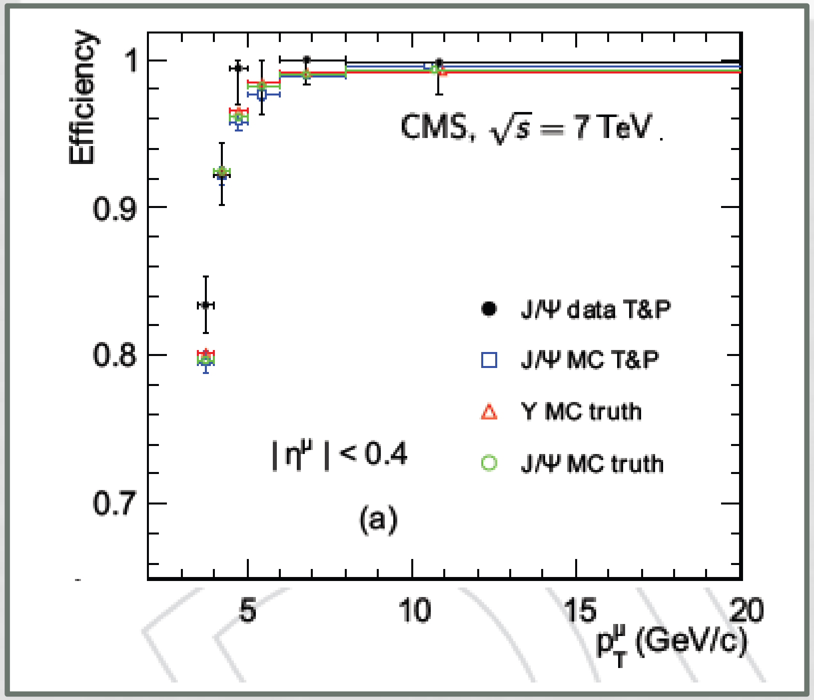
- Separated analysis in barrel/endcap
- Cut optimization and count in B^0 and B_s mass windows
  - check robustness against large pile-up variations
  - mass sidebands for expected background estimation
- Efficiency ratios from MC and checked in data
  - Muon efficiencies: using the "tag-and-probe" method
  - Selection efficiency: using the $B_s \to J/\psi\ \phi$ control sample

SELECTION: TRIGGER

- Essential ingredient → HLT output must be a few Hz!
- Di-muon trigger
  - L1 (hardware) with no p_T requirement ← few kHz at current peak luminosities
  - HLT for $B_s \to \mu\mu$:
    - Opposite charge
    - Single muon $p_T^{\mu} > 2$ GeV/c, dimuon $p_T^{\mu\mu} > 4$ GeV/c
    - Invariant mass $4.8 < m_{\mu\mu} < 6.0$ GeV/c^2
    - Distance of closest approach (DCA) < 5 mm
  - HLT for $B^+ \to J/\psi K^+$ and $B_s \to J/\psi \phi$:
    - J/ψ invariant mass cut
    - $p_T^{\mu} > 3$ GeV/c, $p_T^{\mu\mu} > 7$ GeV/c
    - "Displaced vertex": cuts on vertex χ^2 probability, 2D pointing angle ($\cos\alpha_{2D}$)



MUON EFFICIENCY RATIO

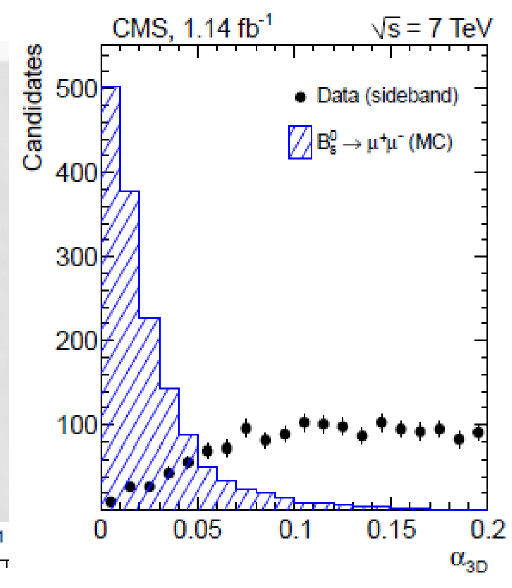


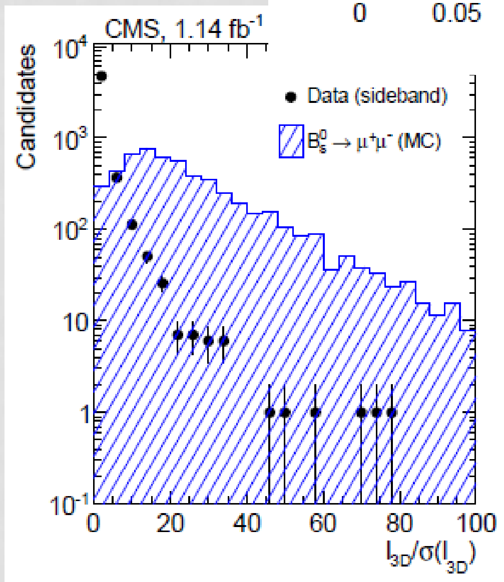
- Systematic uncertainties:
  - Obtained from differences in MC and tag-and-probe

- Determined from MC and cross-checked with data ("tag-and-probe" method)
  - Use inclusive J/ψ with no (or partial) trigger bias on one muon
  - In events with a J/ψ candidate, ask for one well-identified muon ("tag")
  - The other muon ("probe") can pass or not pass the selection *S* under investigation
  - Invariant mass plots separate for the two cases
  - The fitted $N_{pass\text{-}S}/N_{all}$ gives an unbiased estimate of the efficiency ε_S

SELECTION: CUT OPTIMIZATION (1)

- Signal resolution from MC
  - 36 to 77 MeV/c^2 depending on rapidity → (blinded region 5.20 < $m_{\mu\mu}$ < 5.45 GeV/c^2)
- Background from mass sidebands (4.9 < $m_{\mu\mu}$ < 5.9 GeV/c^2 excluding blinded)
- Grid search optimization for best upper limit
- Selection variables:
  1) Muon and dimuon p_T
  2) vertex χ^2 probability
  3) 3D pointing angle ($\cos\alpha_{3D}$)
  4) Flight length significance ($\ell_{3D}/\sigma(\ell_{3D})$)





SELECTION: CUT OPTIMIZATION (2)

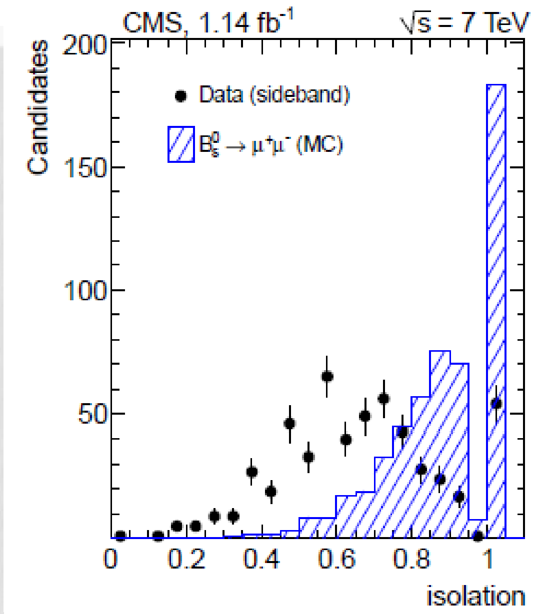
5) Relative di-muon isolation in a cone around the B direction

$$I = \frac{p_\perp(\mu^+\mu^-)}{p_\perp(\mu^+\mu^-) + \sum_{\Delta R<1} p_\perp}$$

- Tracks are counted in the isolation if
  1) $p_T > 0.9$ GeV/c
  2) are associated to the same PV as the B candidate or have a DCA < 0.5 mm with respect to the SV

Essential for pile-up independence of the efficiency

6) DCA of the closest track to the SV

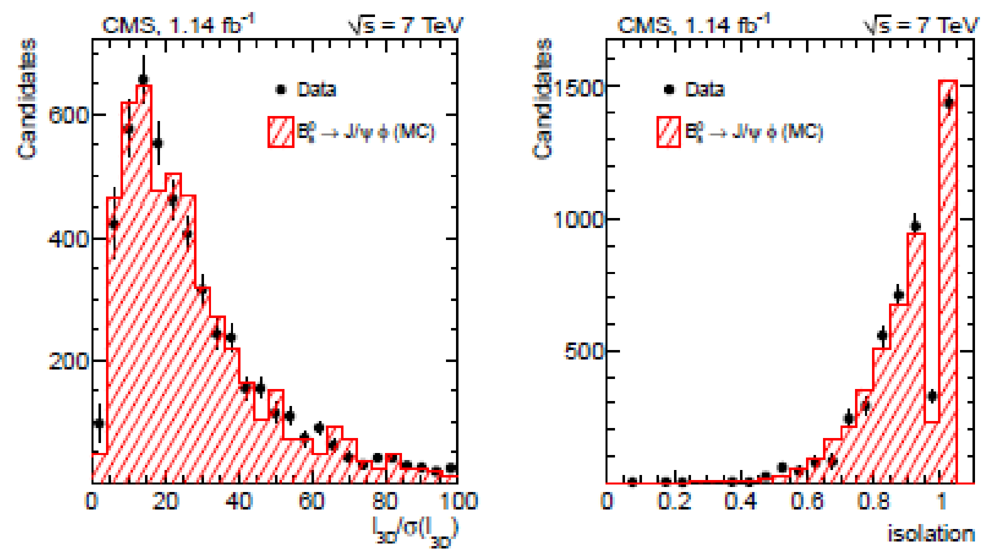


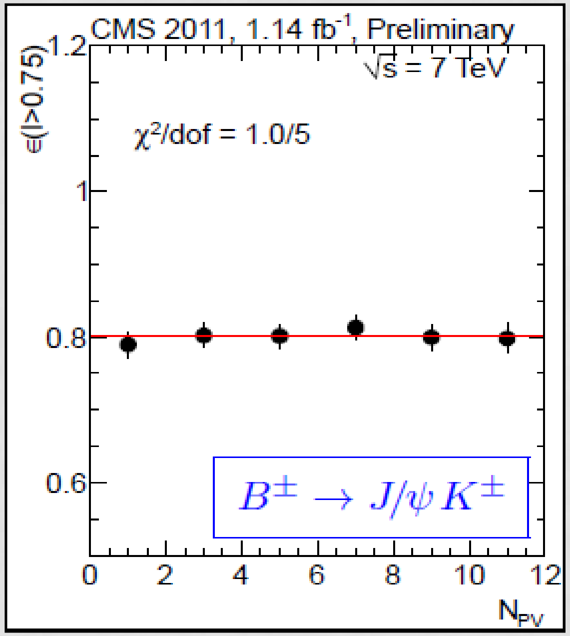
Optimization result

Variable		Barrel	Endcap	units
$p_{\perp\mu_1}$	>	4.5	4.5	GeV
$p_{\perp\mu_2}$	>	4.0	4.0	GeV
$p_{\perp B}$	>	6.5	6.5	GeV
χ^2/dof	<	1.6	1.6	
α	<	0.050	0.025	rad
$\ell_{3d}/\sigma(\ell_{3d})$	>	15.0	20.0	
I	>	0.75	0.75	
d^0_{ca}	>	n/a	0.015	cm

CUT SYSTEMATICS AND PILE-UP

- Cut efficiency systematics determined from data/MC difference in the $B_s \to J/\psi\ \phi$ control sample
  - determined separately for each cut then added in quadrature

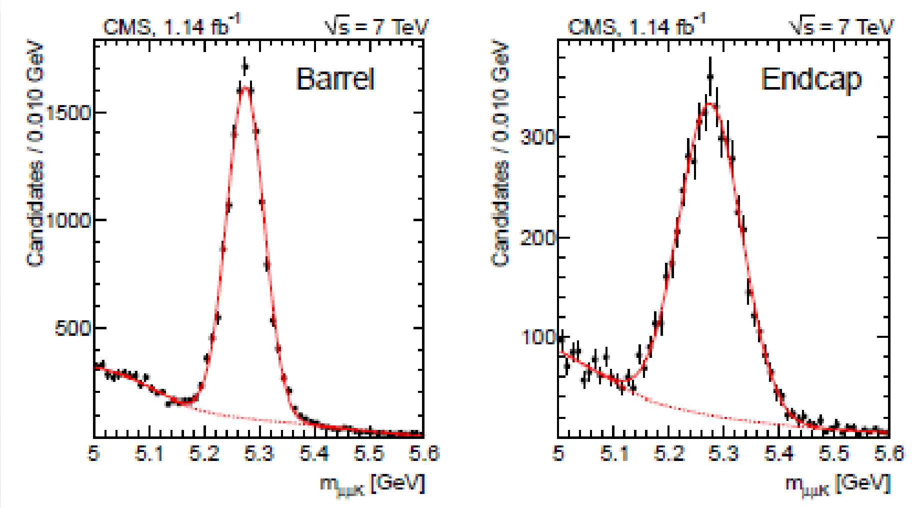


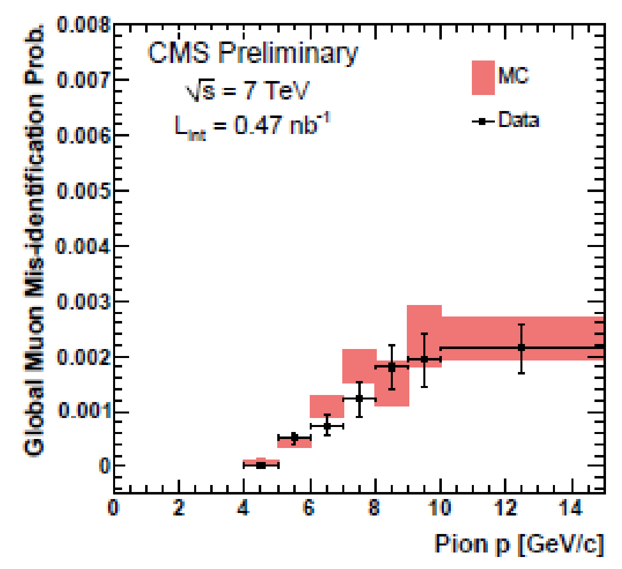


- Pile-up independence of selection checked
  - OK for the sample considered ($<N_{PV}>^{max}$ ~ 12)

NORMALIZATION AND BACKGROUND

- N_{norm} from invariant mass fit to the $B^+ \to J/\psi\ K^+$ sample
  - Barrel: 13050 ± 650
  - Endcap: 4450 ± 220
  - Uncertainties include statistics and systematics (mainly from background shape)





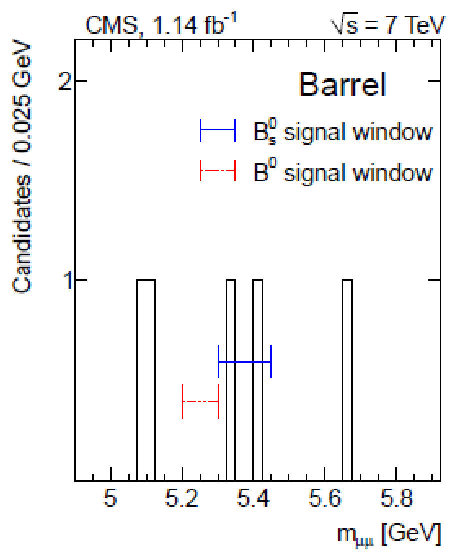
- Combinatorial background from sideband interpolation assuming flat distribution
- Peaking background shapes from MC → sum of many exclusive decays
  - Contribution in the signal region estimated by weighting with measured K, π and p muon mistag probabilities (from $K_s \to \pi^+ \pi^-$, $\phi \to K^+ K^-$ and $\Lambda \to p\, \pi^-$)

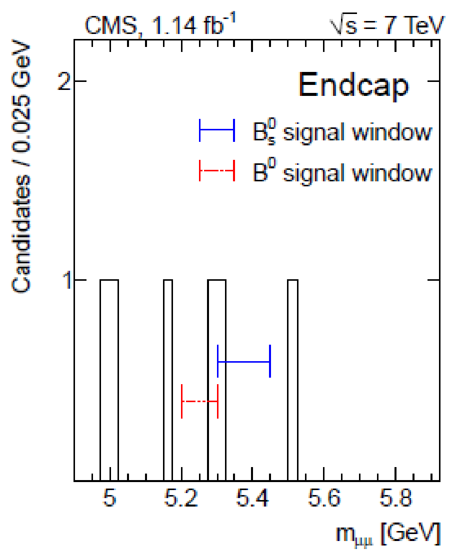
SYSTEMATICS AND CROSS-CHECKS

- Signal estimation
  - Analysis efficiency: data/MC comparison in $B_s \rightarrow J/\psi\ \phi$ — 8%
  - Resolution: data/MC comparison in J/ψ and $\Upsilon(1S)$ — 3%
- Normalization sample
  - Kaon tracking efficiency — 4%
  - Analysis efficiency: data/MC comparison — 4%
  - Fitting procedure — 5%
- Sideband background estimation
  - From sample with reversed isolation cut — 4%
- Acceptance/efficiency ratios
  - Acceptance: variation of production processes — 4%
  - Efficiency ratio for muon reconstruction and identification — 5%
  - Efficiency ratio for muon trigger — 3%
- Other cross-checks
  - Yields vs. data-taking period, measurement of $BR(B_s \rightarrow J/\psi\ \phi)$

RESULTS

	Barrel		Endcap	
	$B^0 \to \mu^+\mu^-$	$B_s^0 \to \mu^+\mu^-$	$B^0 \to \mu^+\mu^-$	$B_s^0 \to \mu^+\mu^-$
Acceptance	$(24.62 \pm 0.99) \times 10^{-2}$	$(24.72 \pm 0.99) \times 10^{-2}$	$(22.61 \pm 0.91) \times 10^{-2}$	$(23.14 \pm 0.93) \times 10^{-2}$
ε_{tot}	$(0.36 \pm 0.04) \times 10^{-2}$	$(0.36 \pm 0.04) \times 10^{-2}$	$(0.21 \pm 0.02) \times 10^{-2}$	$(0.21 \pm 0.02) \times 10^{-2}$
N_{signal}^{exp}	0.065 ± 0.011	0.80 ± 0.16	0.025 ± 0.004	0.36 ± 0.07
N_{bg}^{exp}	0.40 ± 0.23	0.60 ± 0.35	0.53 ± 0.27	0.80 ± 0.40
N_{peak}^{exp}	0.25 ± 0.06	0.07 ± 0.02	0.16 ± 0.04	0.04 ± 0.01
N_{obs}	0	2	1	1





- Expected ULs

$BR(B_s \to \mu\mu) < 1.8 \times 10^{-8}$ @ 95% CL

$BR(B^0 \to \mu\mu) < 4.8 \times 10^{-9}$ @ 95% CL

- Observed ULs

$BR(B_s \to \mu\mu) < 1.9 \times 10^{-8}$ @ 95% CL

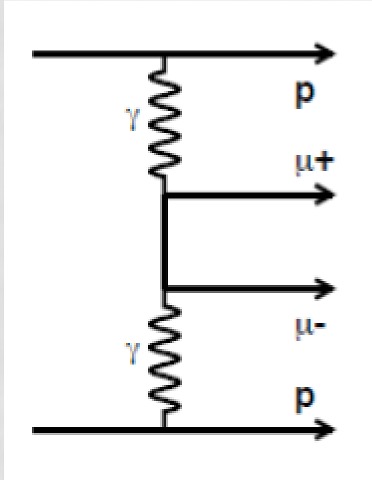
$BR(B^0 \to \mu\mu) < 4.6 \times 10^{-9}$ @ 95% CL

LHC-B ANALYSIS

LHCb Collaboration, Phys. Lett. B699, 330 (2011)
LHCb Collaboration, arXiv 1112.1600 (2012)
accepted by Phys. Lett. B

ANALYSIS PRINCIPLES

- Main backgrounds:
  - Muons from two semileptonic B decays
  - Contribution from photoproduction → effectively removed with a cut on $p_T^{\mu\mu}$
  - Muon from semileptonic B decay plus misidentified hadron and rare decays → sub-dominant



- Relying on MC as little as possible
  - Pre-selection
  - Use of a multivariate analysis to discriminate signal from background (likelihood → Boosted Decision Tree (BDT) in 2011)
  - Signal extraction in 2D binning of invariant mass / BDT (6x4)
    - Signal mass, BDT likelihoods from data control samples
- Efficiency ratios from data or MC cross-checked with data

EFFICIENCY RATIOS

- Reconstruction efficiency ratio from MC
  - Main uncertainty from extra tracks in normalization channels (4%)
  - Cross-checked comparing efficiencies of $B^0 \to J/\psi\, K^{*0}$ / $B^+ \to J/\psi\, K^+$
- Selection efficiency ratio from MC
  - Verified by comparing data/MC control samples → only significant discrepancy in muon IP distributions but negligible effect on ratio
- Trigger efficiency from data, by defining:
  - "TIS" events: events with at least one triggering object not matched to the signal candidate → mostly coming from decay products of the other B in the event
  - "TOS" events: events with signal matching the trigger

$$\epsilon^{\mathrm{TRIG|SEL}} = \frac{N^{\mathrm{TRIG}}}{N^{\mathrm{SEL}}} = \frac{N^{\mathrm{TIS}}}{N^{\mathrm{SEL}}}\frac{N^{\mathrm{TRIG}}}{N^{\mathrm{TIS}}} = \epsilon^{\mathrm{TIS}}\frac{N^{\mathrm{TRIG}}}{N^{\mathrm{TIS}}} \approx N^{\mathrm{TIS\&TOS}}/N^{\mathrm{TOS}} \times \frac{N^{\mathrm{TRIG}}}{N^{\mathrm{TIS}}}$$

  - Determined on J/ψ as a function of the muon p_T and IP, and "applied" to the B_s sample (1-9%)

BDT INPUTS

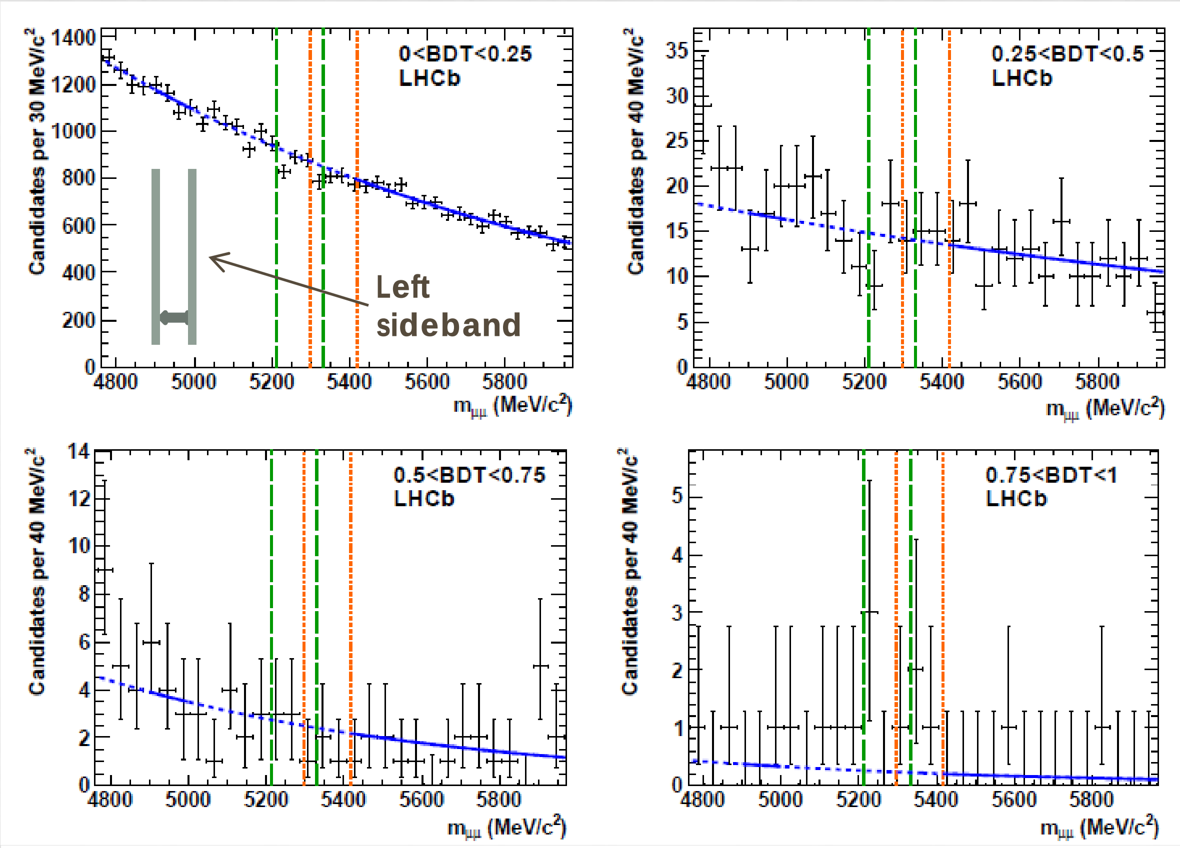
- Variables
  1) Dimuon proper decay time
  2) Dimuon p_T
  3) Dimuon impact parameter
  4) DCA between the 2 muons
  5) Single-muon isolation (number of tracks forming a good vertex with the muon)
  6) B-candidate isolation (defined as in CMS)
  7) Minimum p_T of the 2 muons
  8) Minimum IP of the 2 muons
  9) Cosine of the lowest-p_T muon polarization angle

- Training on MC
- Calibration with data (see next slides)

N.B. No muon-ID requirements
→ can be calibrated with $B_{s,d} \to h^+h^-$ decays

$$\cos P = \frac{p_{y,\mu 1}\, p_{x,B} - p_{x,\mu 1}\, p_{y,B}}{p_{T,B}\,(m_{\mu\mu}/2)}$$

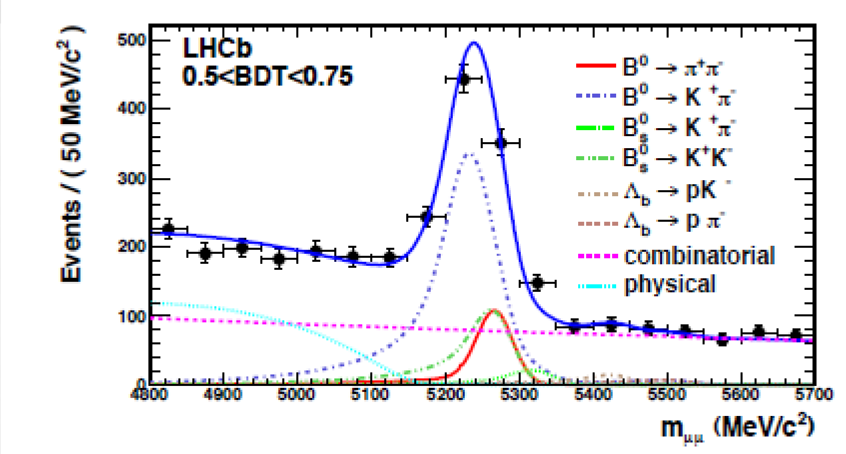
COMBINATORIAL BDT PROBABILITY

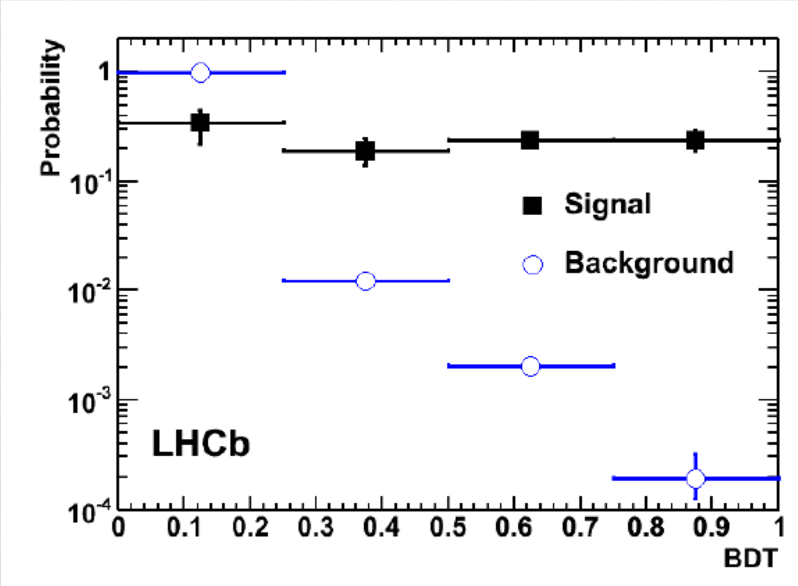


- Background yields estimated using sidebands only and extrapolating to signal region in the 4 BDT bins
- Result checked using different fit functions (1, 2 exponential(s), linear… etc.)
  - Nominal fit only uses [4.9, 5.0] GeV/c^2 as lower sideband to avoid cascade and peaking decays

SIGNAL BDT PROBABILITY

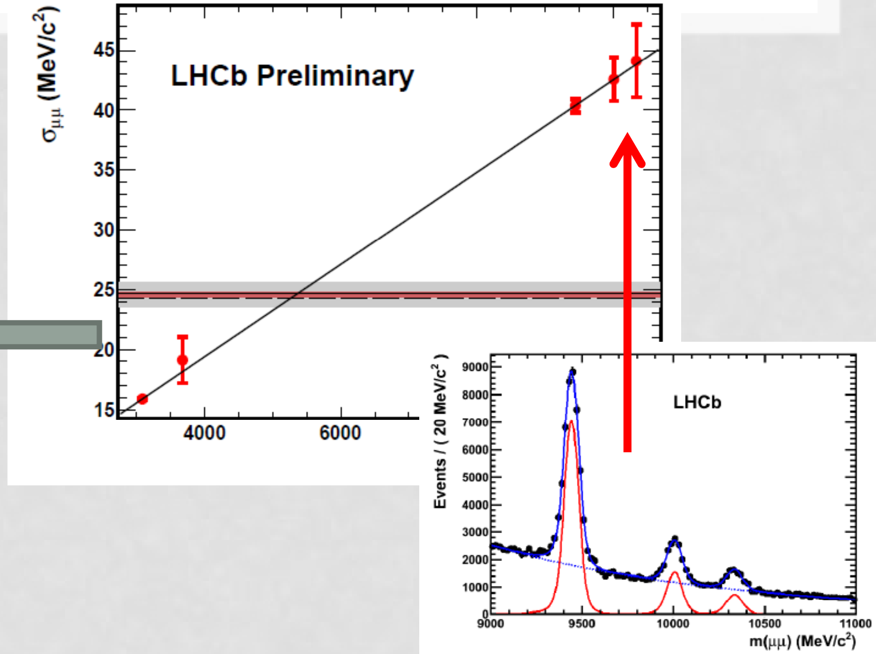
- Fit in the 4 BDT bins to control samples
  1) $B_{s,d} \rightarrow h^+h^-$ inclusive decays
  2) $B_{s,d} \rightarrow KK, K\pi, \pi\pi$ exclusive decays (PID used)
- Only use "TIS" events to remove bias from hadronic triggers
- Differences between 1) and 2) used as systematics on signal likelihood (3-10%)





SIGNAL MASS SHAPE

- Determined from data in two alternative ways:
  1) Interpolating linearly charm-onium and bottomonium mass resolutions
  2) Using the hadronic samples in the previous slide ("TIS" events of $B_{s,d} \to h^+h^-$ inclusive and $B_{s,d} \to KK, K\pi, \pi\pi$)



method	$\sigma(B_s^0)$ (MeV/c^2)
interpolation	$24.6 \pm 0.2_{\text{stat}} \pm 1.0_{\text{syst}}$
$B^0_{(s)} \to h^+h^-$ inclusive	$23.7 \pm 0.4_{\text{stat}} \pm 1.5_{\text{syst}}$
$B^0_{(s)} \to h^+h^-$ exclusive	$23.5 \pm 0.2_{\text{stat}} \pm 1.3_{\text{syst}}$

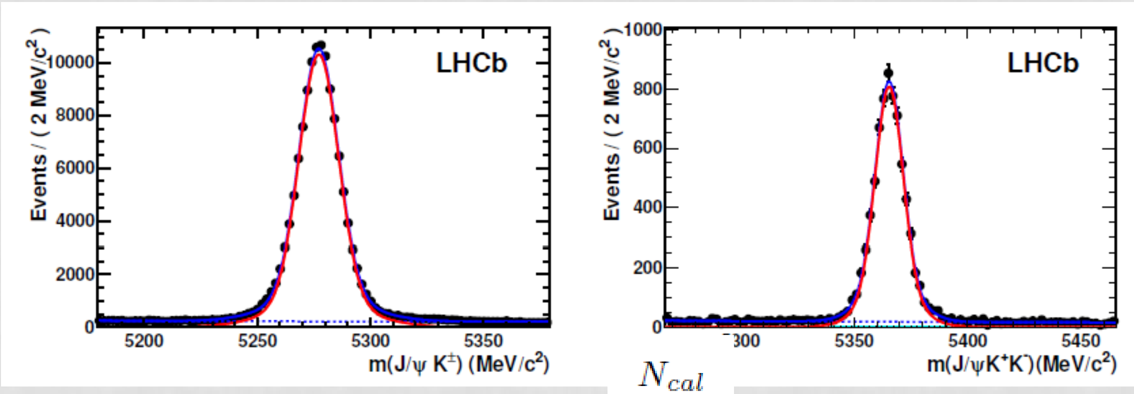
- Muon momentum scale
- p_T reweighting

- Variation of "physical" background (3-body B decays)

PEAKING BG AND NORMALIZATION

- $B_{s,d} \to h^+h^-$ decays are also used to estimate peaking background from double muon mistag
  - $N_{B^0} = 5.0 \pm 0.9$ $N_{B^0_s} = 1.0 \pm 0.4$ in the whole BDT range
  - BDT distribution assumed to be the same as the signal
- Normalization factors averaged over the 3 control samples

$$\alpha^{\rm norm}_{B^0_s\to\mu^+\mu^-} = (8.38 \pm 0.74) \times 10^{-10}$$
$$\alpha^{\rm norm}_{B^0\to\mu^+\mu^-} = (2.20 \pm 0.11) \times 10^{-10}$$



	N_{cal}
$B^+ \to J/\psi K^+$	$124\,518 \pm 2\,025$
$B^0_s \to J/\psi\phi$	$6\,940 \pm 93$
$B^0 \to K^+\pi^-$	$4\,146 \pm 608$

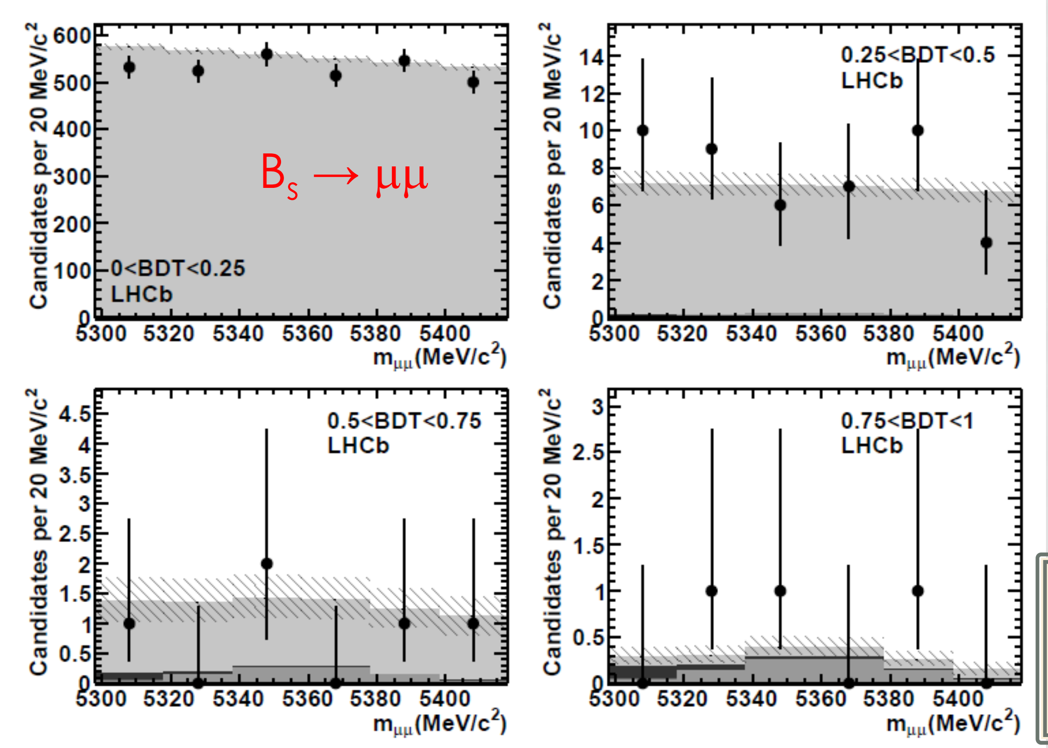
- Including uncertainties on:
  - Fitted yields
  - BRs
  - Reconstruction, selection and trigger efficiency ratios

RESULTS

Expectations

■ = signal (SM BR) □ = combinatorial bkg

■ = peaking bkg hatched = bkg uncertainty



- Expected ULs (SM BR hypotheses)

$BR(B_s \rightarrow \mu\mu) < 1.4 \times 10^{-8}$ @ 95% CL

$BR(B^0 \rightarrow \mu\mu) < 3.2 \times 10^{-9}$ @ 95% CL

- Observed ULs

$BR(B_s \rightarrow \mu\mu) < 1.6 \times 10^{-8}$ @ 95% CL

$BR(B^0 \rightarrow \mu\mu) < 3.6 \times 10^{-9}$ @ 95% CL

- Combined with 2010 result to obtain:

$BR(B_s \rightarrow \mu\mu) < 1.4 \times 10^{-8}$ @ 95% CL

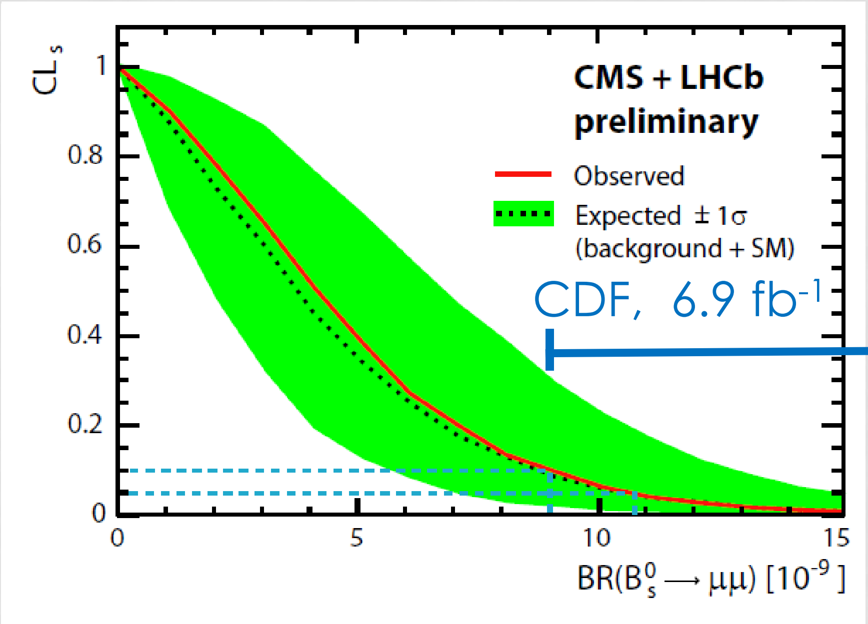
$BR(B^0 \rightarrow \mu\mu) < 3.2 \times 10^{-9}$ @ 95% CL

B_S: CMS/LHC-B COMBINATION

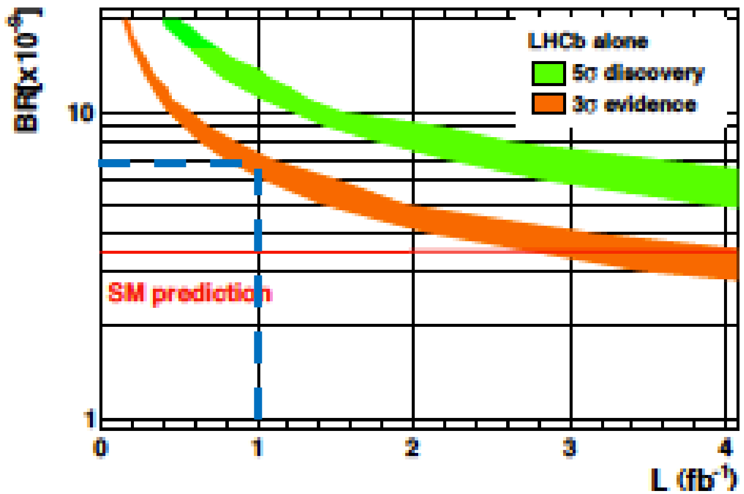
CMS and LHCb Collaborations
LHCb-CONF-2011-047
CMS-PAS-BPH-11-019 (2011)

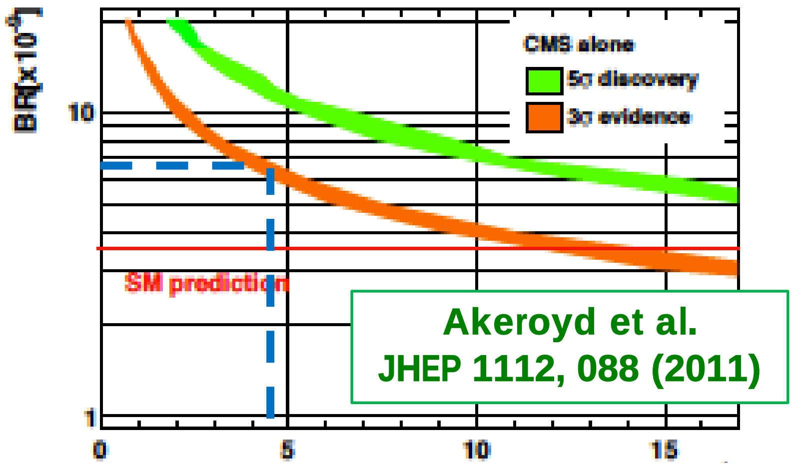
- Observed limits:
  - CMS: $BR(B_s \to \mu\mu) < 1.9 \times 10^{-8}$ @ 95% CL
  - LHCb: $BR(B_s \to \mu\mu) < 1.5 \times 10^{-8}$ @ 95% CL (0.30 fb^{-1}, preliminary)
- Combination using CL_s method (only significant common systematics: f_s / f_d)

- Combined limit:
  - $BR(B_s \to \mu\mu) < 1.1 \times 10^{-8}$ @ 95% CL
  - p-values for hypotheses of:
    - Background only: 8%
    - Background + SM signal: 55%
    - Background + 5.6 x (SM signal) : 0.3%

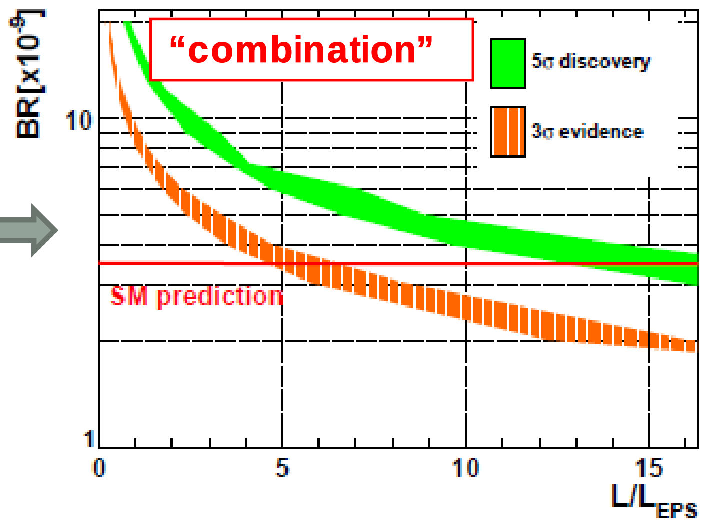


(CLOSE) FUTURE PERSPECTIVES





- Dotted blue lines indicate current integrated luminosity
- Two experiments expected to provide similar ULs (2 ÷ 3 x SM, assuming no analysis improvements) → big advantage from combination
- Results of both likely to be presented at winter conferences



(LESS CLOSE) FUTURE PERSPECTIVES

2012 run

- LHC possibly running at 8 TeV (~15% increase in $b\bar{b}$ cross-section) → Chamonix workshop taking place now
- <u>LHCb:</u> maintaining absolute priority on $B_{s,d} \to \mu^+\mu^-$
- <u>CMS</u>: also considered as a "benchmark channel"
  - Possible improvements in analysis (e.g. move to multivariate techniques)
  - Detailed studies ongoing, especially for trigger
    - Double-muon HLT with DCA cut and offline analysis checked to be almost insensitive to pile-up
    - "Displacement" requirements verified to be very efficient on control samples → room for application to $B_s \to \mu\mu$ before raising p_T thresholds

CONCLUSIONS

- The LHCb (CMS) experiment has searched for the rare decays $B_{s,d} \rightarrow \mu^+\mu^-$ in data samples of 0.37 (1.14) fb^{-1}
  - Analogies in analysis technique: normalize to modes with similar signature and well-known BRs
  - Differences in:
    - Background composition: due to different rapidity regions and pile-up conditions
    - Signal selection: cut-and-count vs. multivariate analysis
- Single best world limits from LHCb
- Statistical combination of the results yields:

$$BR(B_s \rightarrow \mu\mu) < 1.1 \times 10^{-8} \text{ @ } 95\% \text{ CL}$$

  - **<u>Results with full 2011 statistics will be the first thorough test of SM</u>**
  - For B_d we are not quite yet there…