$B^0_s \rightarrow \mu^+\mu^-$ at the LHC

Flavio Archilli

on behalf of CMS and LHCb collaborations

8th International Workshop on the CKM Unitarity Triangle

8-12 September 2014

Vienna, Austria

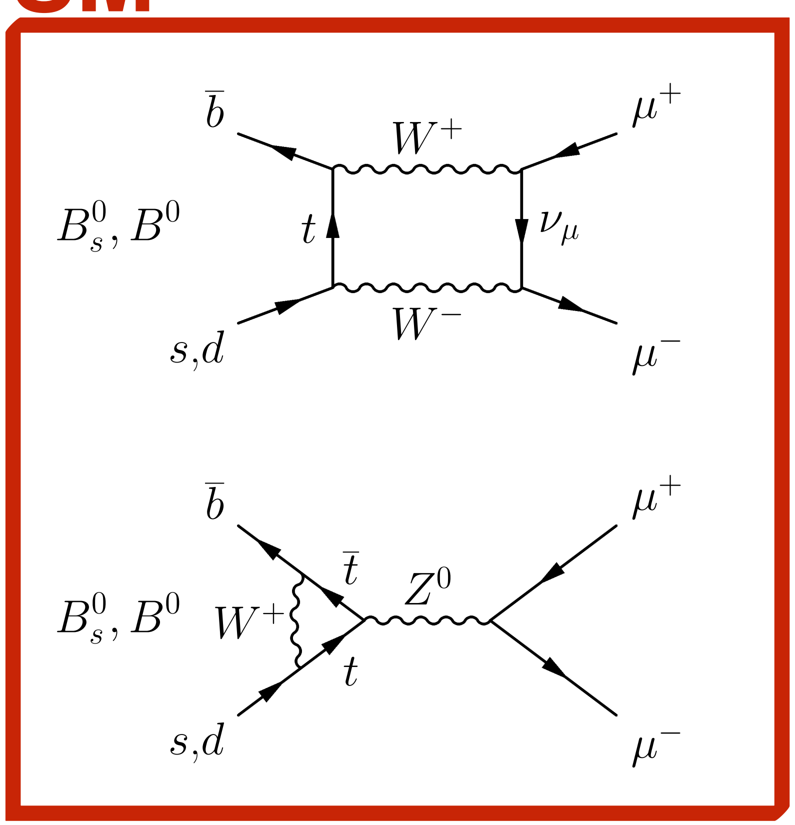
Outline

- Theoretical introduction
- Experimental status
- Combination of CMS and LHCb results
- Results and conclusions

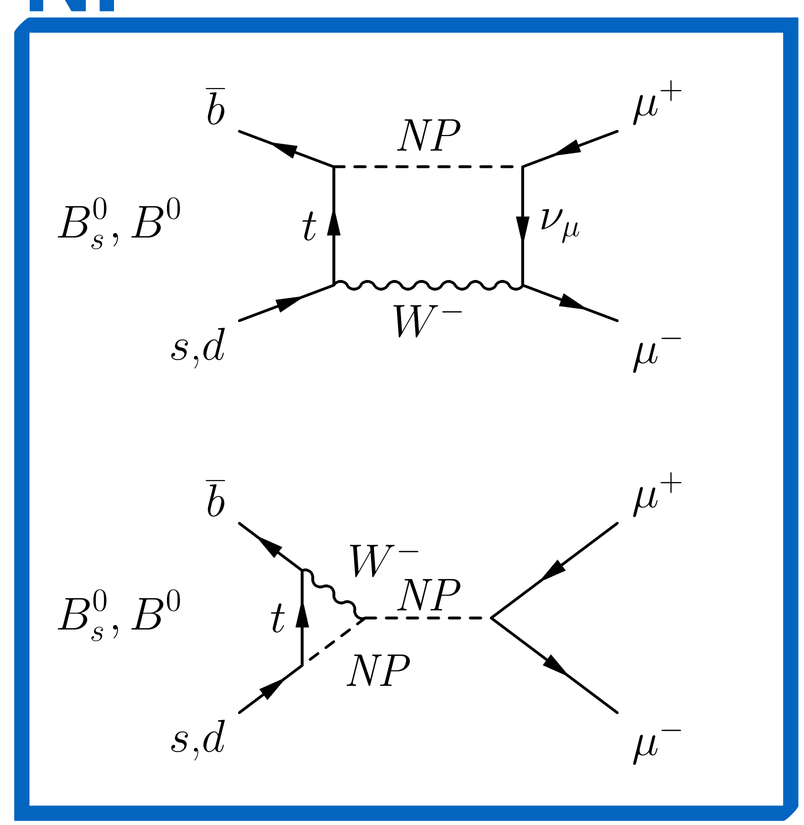
Theory (1)

- Highly suppressed in the SM: FCNC and helicity suppressed, proceeding via Z penguin and W-box
- The helicity suppression of vector(-axial) terms make these decays particularly sensitive to NP (pseudo-)scalar contribution, such as extra Higgs doublets (MSSM), can raise their BFs
- e.g. in MSSM the BF is proportional to $\tan^6\beta / {m_A}^4$

SM



NP



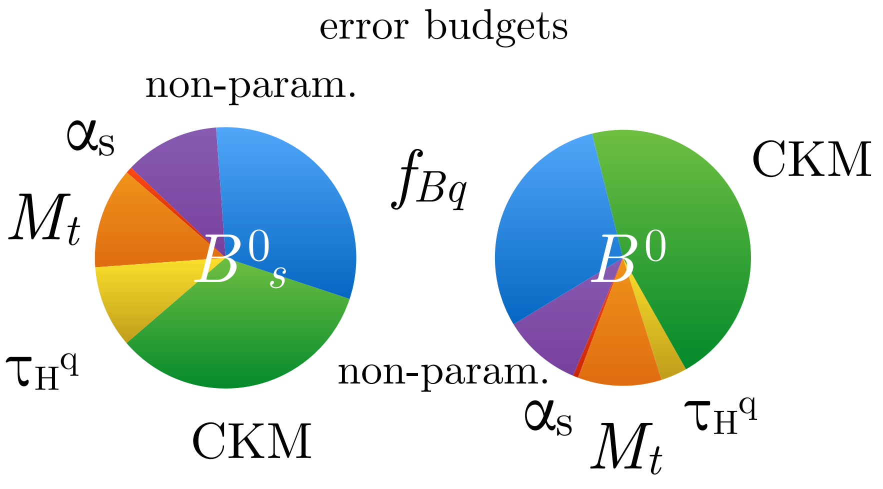
Theory (2)

Bobeth et al. [PRL 112 (2014) 101801]

- Untagged time integrated branching fraction predictions:

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (3.66 \pm 0.23) \times 10^{-9}$$
$$\mathcal{B}(B^0 \to \mu^+\mu^-) = (1.06 \pm 0.09) \times 10^{-10}$$

updated with the latest top mass measurement (Tevatron+LHC combination) [hep-ex/1403.4427]

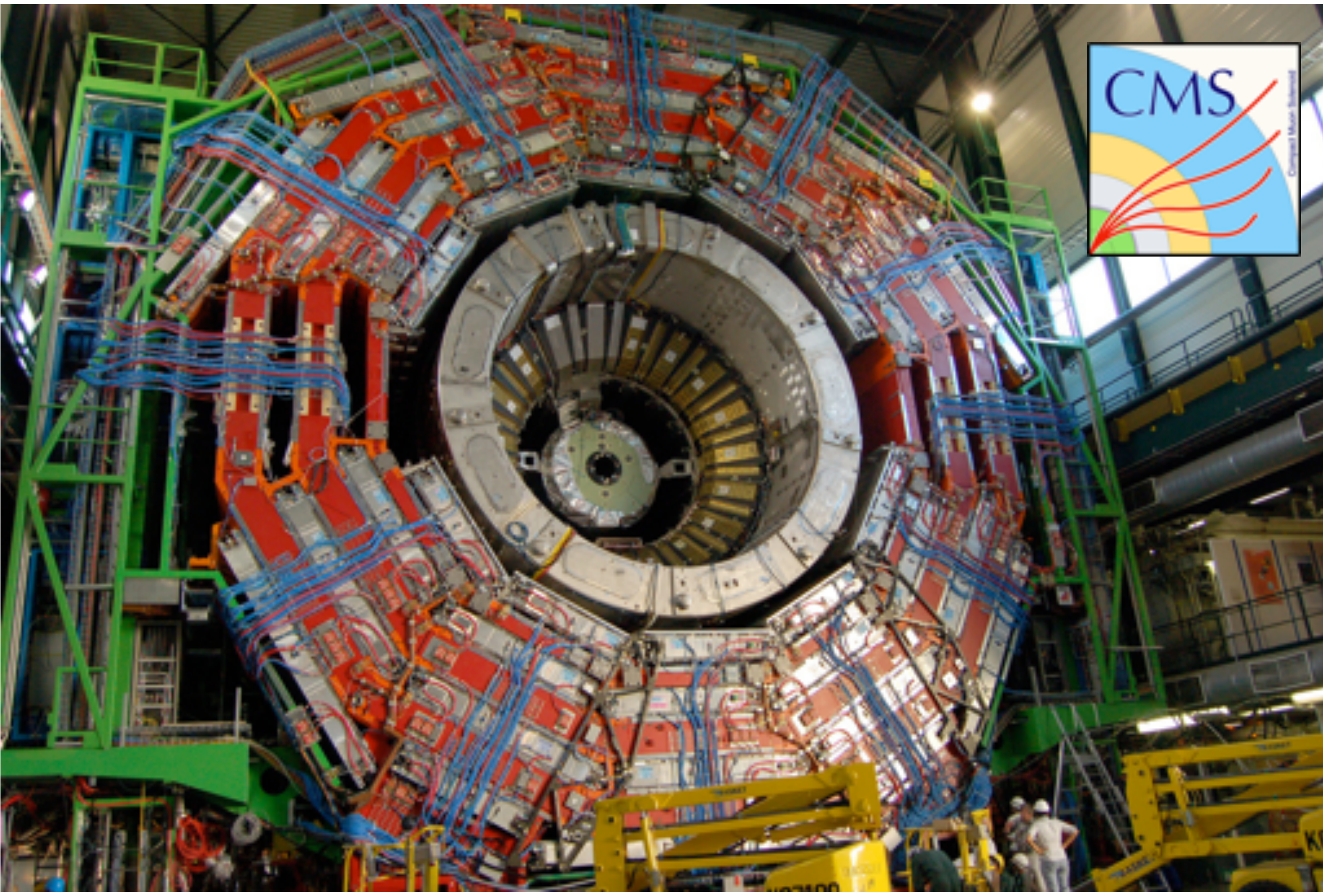


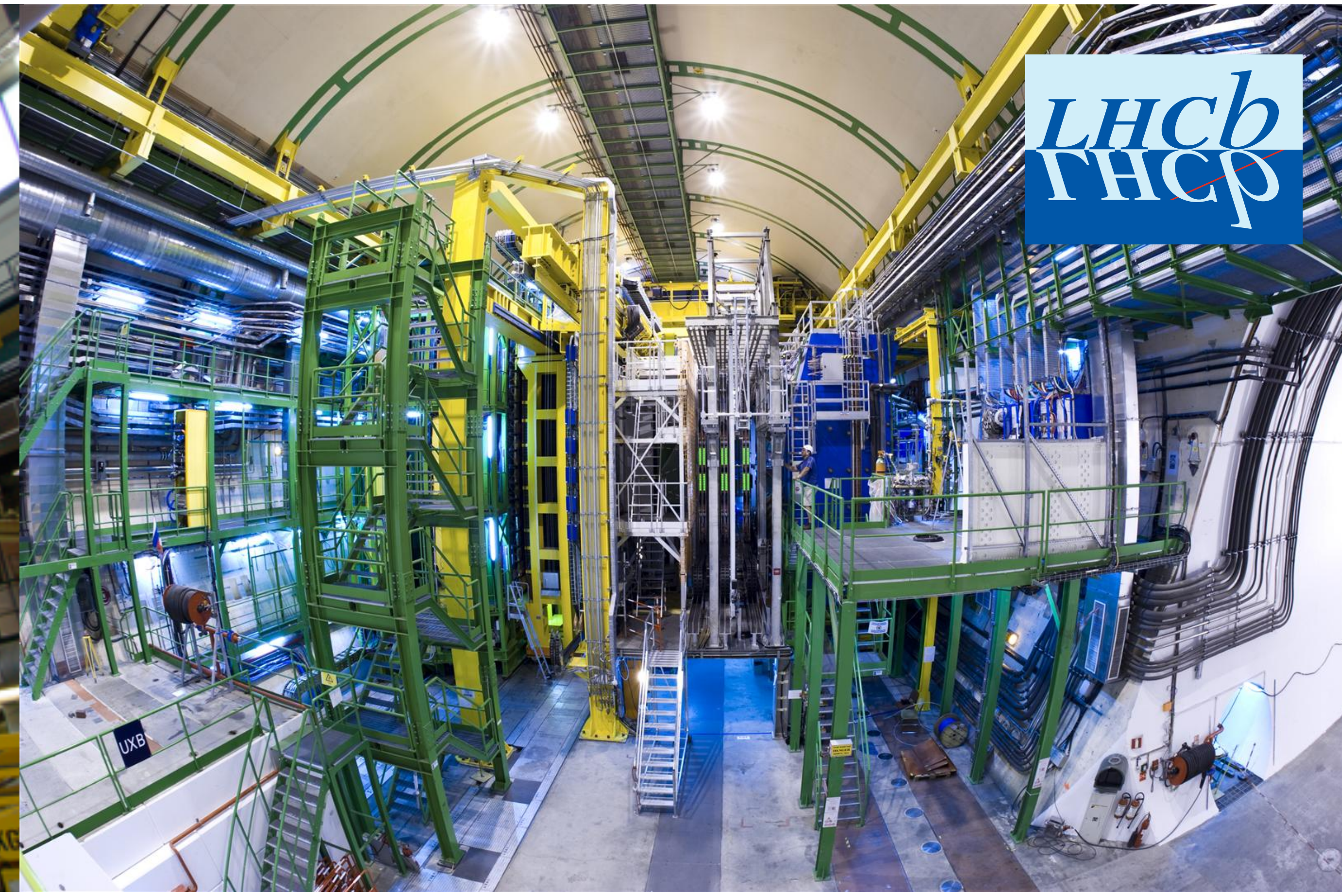
- Ratio of branching fractions of two modes powerful to discriminate among models beyond the SM. Precisely predicted in SM:

$$\mathcal{R} = \frac{\mathcal{B}(B^0 \to \mu^+\mu^-)}{\mathcal{B}(B_s^0 \to \mu^+\mu^-)} = \frac{\tau_{B_d}}{1/\Gamma_H^s}\left(\frac{f_{B_d}}{f_{B_s}}\right)^2 \left|\frac{V_{td}}{V_{ts}}\right|^2 \frac{M_{B_d}\sqrt{1-\frac{4m_\mu^2}{M_{B_d}^2}}}{M_{B_s}\sqrt{1-\frac{4m_\mu^2}{M_{B_s}^2}}} = 0.0295^{+0.0028}_{-0.0025}$$

➡ stringent test of Minimal Flavour Violation hypothesis

Experiments





- Good trigger and muon PID
- No hadron PID
- Silicon tracker excellently working to resolve the signal decays in the high pileup environment
- Di-muon mass resolution 32-75 MeV/c^2

- Efficient muon trigger
- Good muon and hadron PID
- Luminosity levelling at 4×10^{32} cm^{-2}s^{-1}
- Di-muon mass resolution 25 MeV/c^2

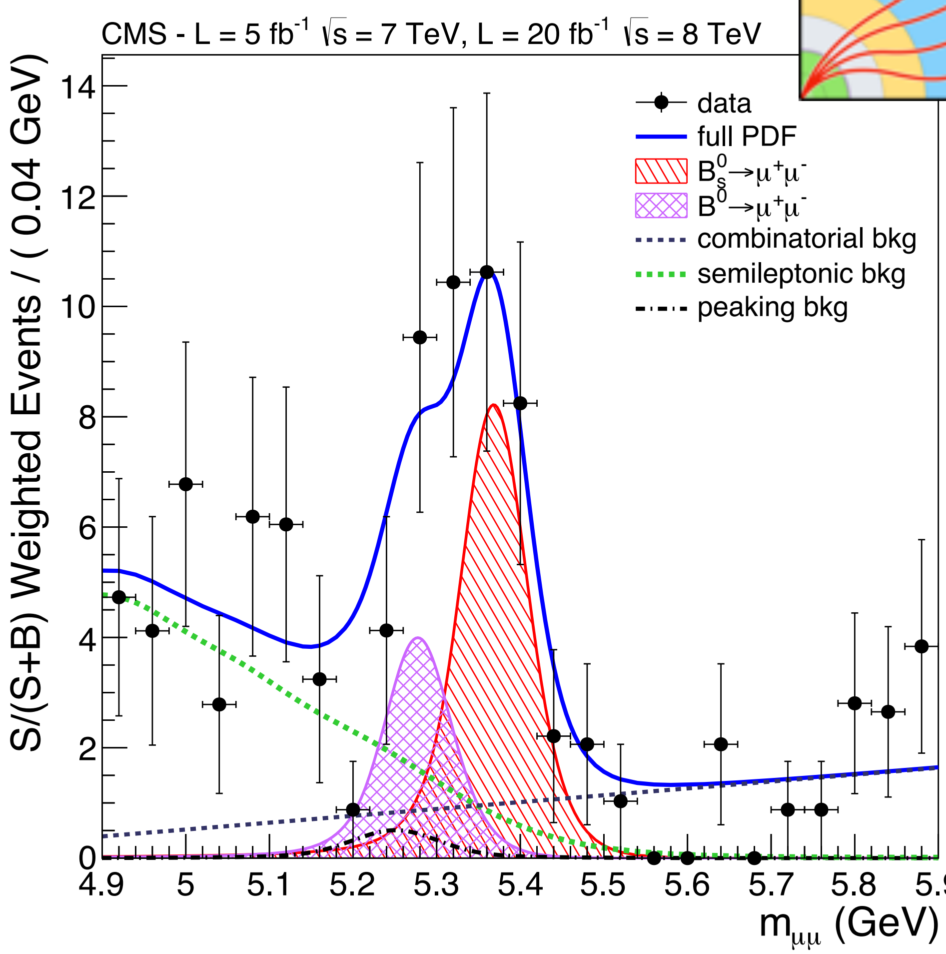
Analysis description

- Very similar analysis strategies between CMS and LHCb
  - Soft preselection
  - Multivariate classifier (BDT) for signal/background discrimination
  - Search in the invariant mass distribution and BDT categories
  - Calibration of di-muon mass ($m_{\mu\mu}$) with Quarkonia ($\rightarrow\mu\mu$) (also exclusive $B\rightarrow hh'$ for LHCb)
  - Background: combinatorial ($bb\rightarrow\mu\mu X$), peaking from double misID ($B\rightarrow hh'$) and semileptonic decays ($B^0\rightarrow\pi^-\mu^+\nu$, $B^0_s\rightarrow K^-\mu^+\nu$, $B^{0(+)}\rightarrow\pi^{0(+)}\mu^+\mu^-$, $\Lambda^0_b\rightarrow p\mu^-\overline{\nu}$)
  - Normalisation with $B^+\rightarrow J/\psi K^+$ (and $B^+\rightarrow K^+\pi^-$ for LHCb)
- 2+1 fb^{-1} for LHCb and 5+20 fb^{-1} for CMS at 7 and 8 TeV
- ~1fb^{-1} at LHCb is equivalent to ~10fb^{-1} at CMS

EPS-HEP 2013

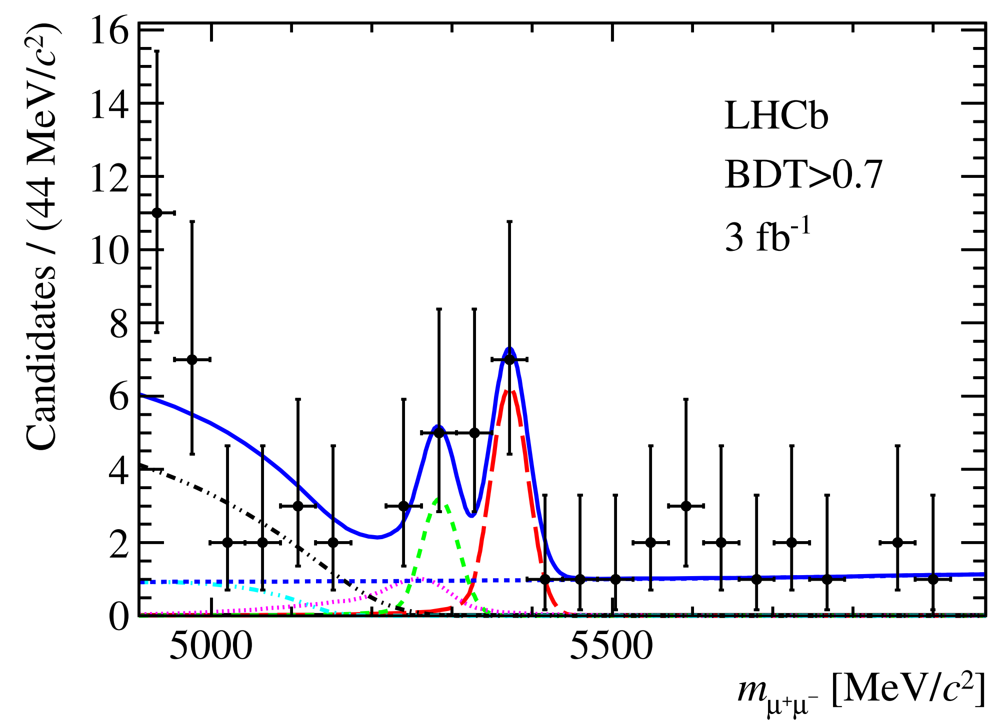
- During EPS-HEP 2013 CMS and LHCb collaborations presented their results based on $25 fb^{-1}$ and $3 fb^{-1}$ respectively

[PRL 111 (2013) 101804]



Nov. 2012: LHCb found the first evidence of the $\mathcal{B}(B^0{}_s \to \mu^+\mu^-)$ with $2.1 fb^{-1}$

[PRL 111 (2013) 101805]



$$\mathcal{B}(B^0_s \to \mu^+\mu^-) = 3.0^{+1.0}_{-0.9} \times 10^{-9} \ \ (4.3\sigma)$$

$$\mathcal{B}(B^0 \to \mu^+\mu^-) = 3.5^{+2.1}_{-1.8} \times 10^{-10} \ \ (2.0\sigma)$$

$$\mathcal{B}(B^0_s \to \mu^+\mu^-) = 2.9^{+1.1}_{-1.0} \times 10^{-9} \ \ (4.0\sigma)$$

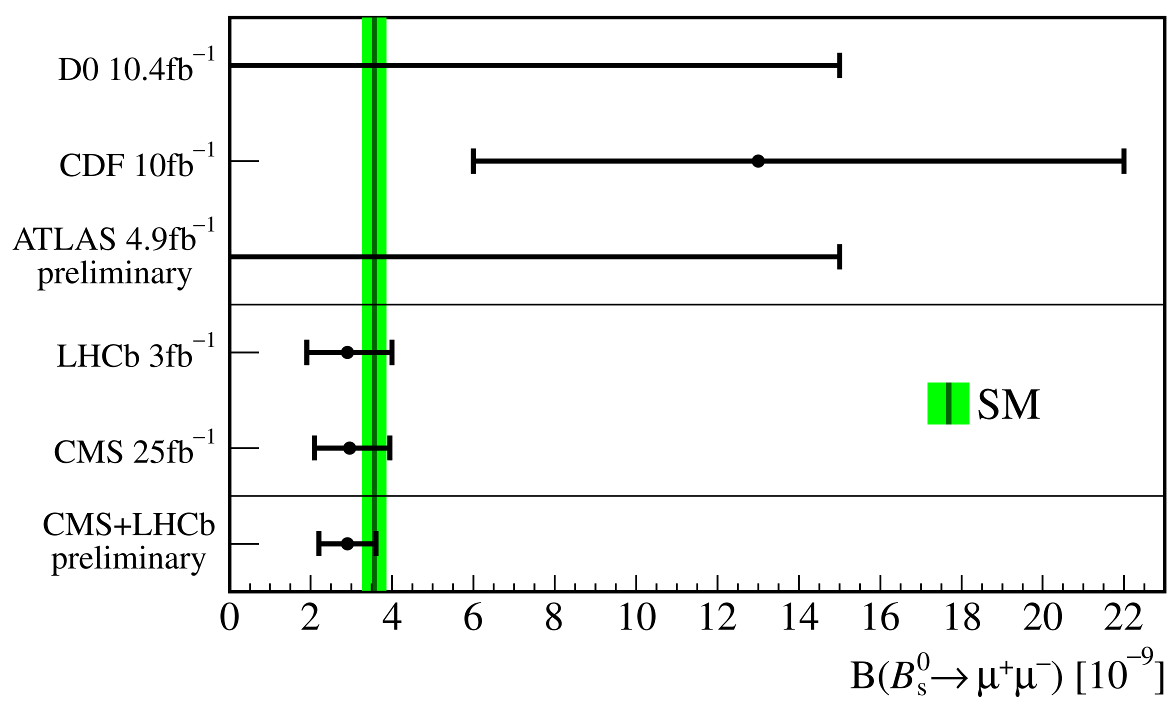
$$\mathcal{B}(B^0 \to \mu^+\mu^-) = 3.7^{+2.4}_{-2.1} \times 10^{-10} \ \ (2.0\sigma)$$

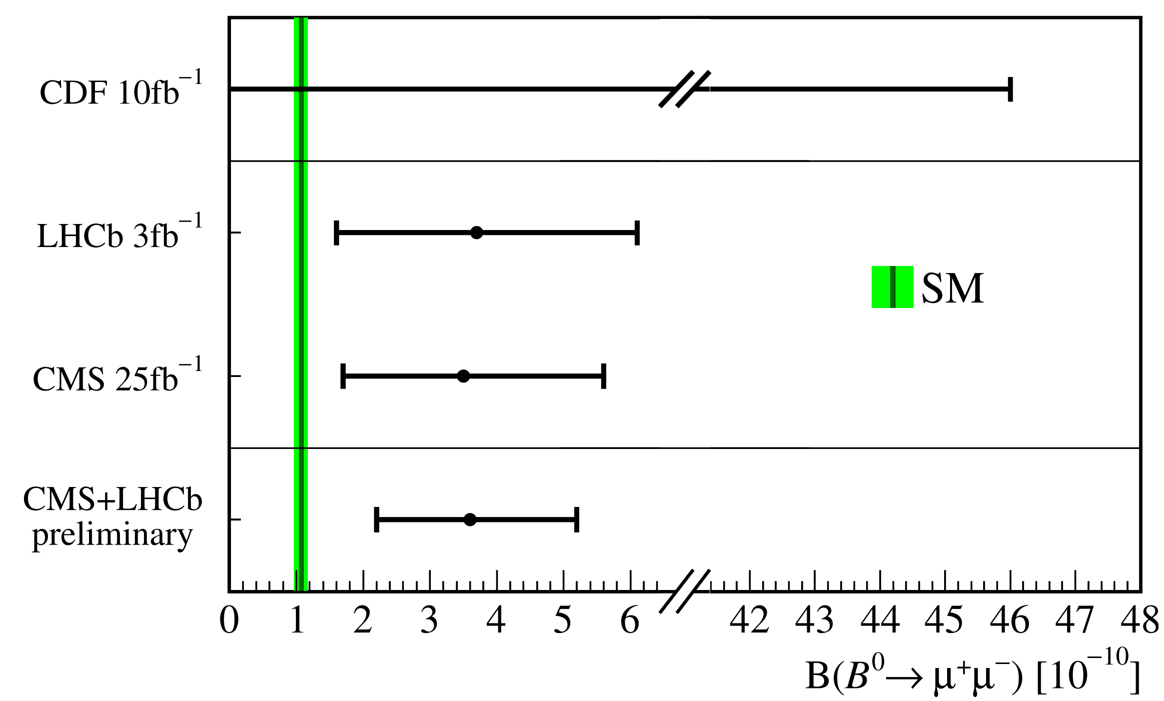
Preliminary combination

A preliminary combination already presented at EPS-2013

- Correlation on f_d/f_s taken into account
- No estimate of significance was provided

[CMS-PAS-BPH-13-007]
[LHCb-CONF-2013-012]





$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = 2.9 \pm 0.7 \times 10^{-9}$$

$$\mathcal{B}(B^0 \to \mu^+\mu^-) = 3.6^{+1.6}_{-1.4} \times 10^{-10}$$

Full likelihood combination will be presented today

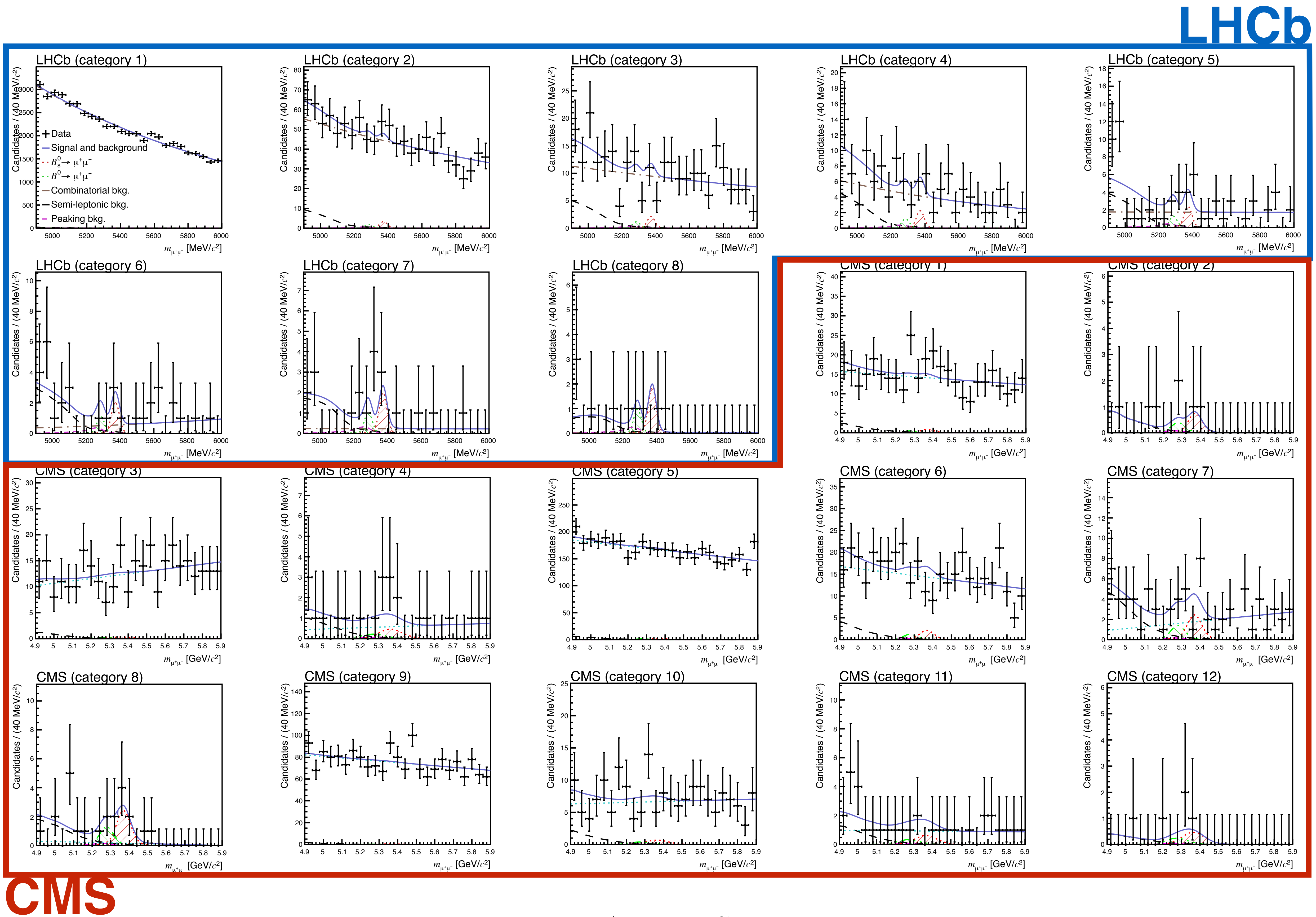
Full likelihood combination

- Full likelihood combination of CMS and LHCb results awaited:
  - Estimate significance
  - Take into account correlations
- Some changes were made to harmonise the analysis
  - $\Lambda^0_b \to p\mu^-\overline{\nu}$ background source was included in the nominal fit, with updated BF and MC simulation changed to include a more realistic model for the properties of the decay
  - Lifetime bias correction calculated and applied (already included in LHCb analysis)
  - Updated value of the ratio of hadronisation fractions f_d/f_s=3.86±0.22

Simultaneous Fit

- the two datasets are used together in a single combined experiment
- simultaneous unbinned extended maximum likelihood fit to the mass spectra in 8 BDT bins for LHCb and 12 categories for CMS
- Shared parameters:
  - the branching fraction of the two signals $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ and $\mathcal{B}(B^0 \to \mu^+\mu^-)$
  - the branching fraction of the normalisation channel $\mathcal{B}(B^+ \to J/\psi K^+)$
  - the ratio of the hadronisation fractions f_d/f_s
- Assuming SM BF 94±7 $B^0_s \to \mu^+\mu^-$ and 10.5±0.6 $B^0 \to \mu^+\mu^-$ expected for the two together

Simultaneous Fit



Fit result

from the simultaneous fit we get:

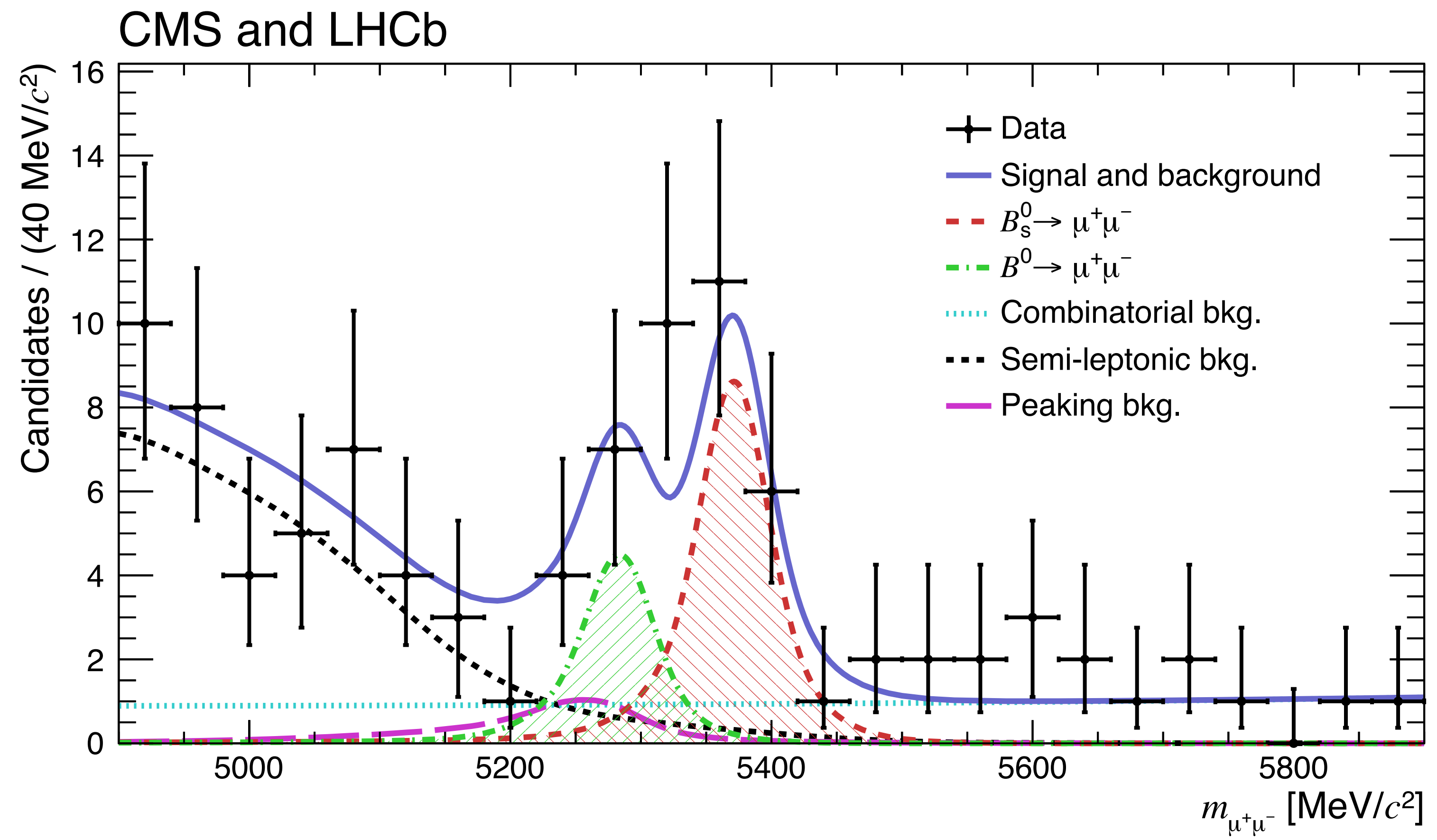
$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = 2.8^{+0.7}_{-0.6} \times 10^{-9}$$

$$\mathcal{B}(B^0 \to \mu^+\mu^-) = 3.9^{+1.6}_{-1.4} \times 10^{-10}$$

Using the Wilks' theorem the statistical significance from the likelihood is:

- 6.2 σ for the $B^0{}_s \to \mu^+\mu^-$ (Expected SM 7.6 σ)
  - ✦ First **observation**

- 3.2 σ for the $B^0 \to \mu^+\mu^-$ (Expected SM 0.8 σ)

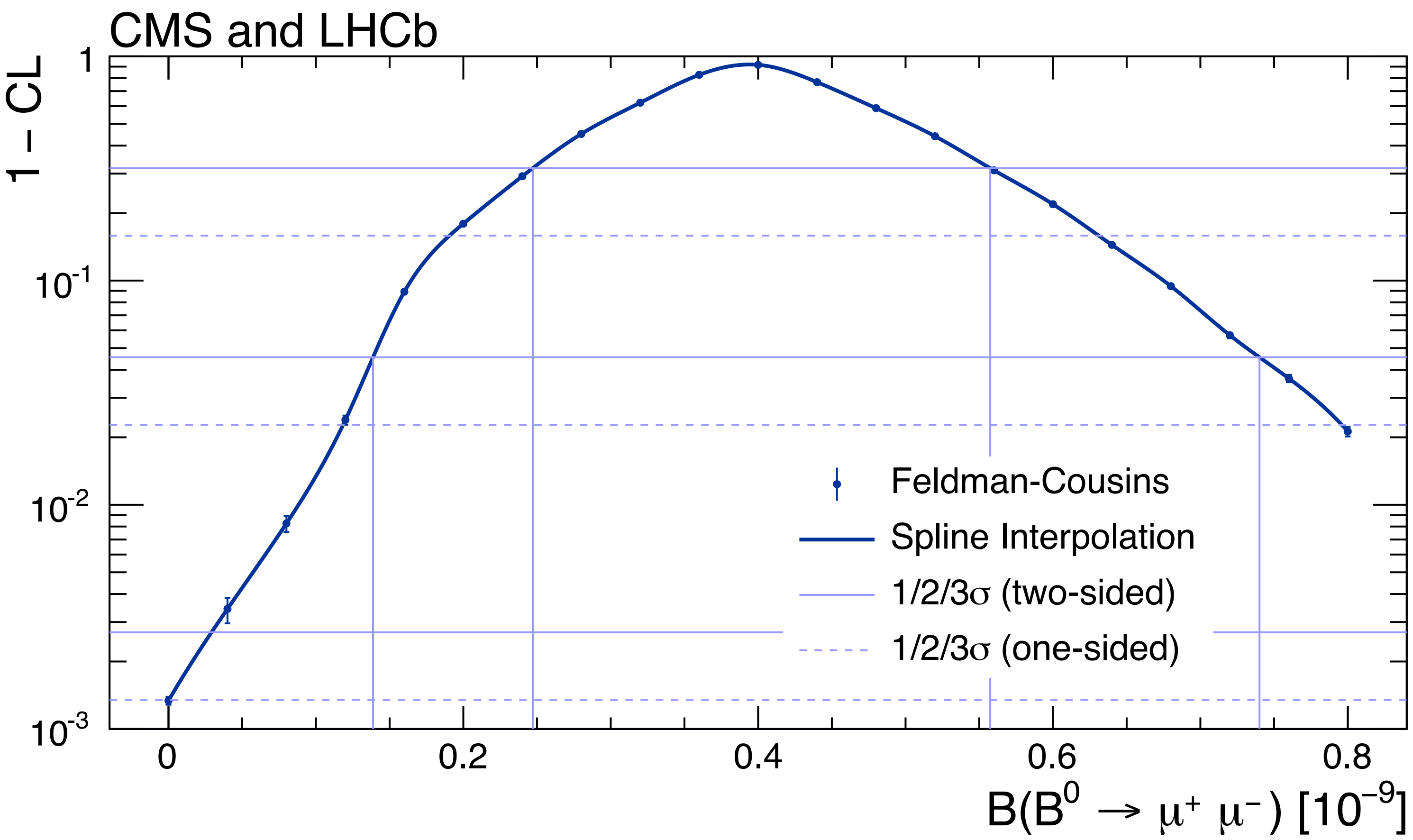


projection of invariant mass of best 6 categories selected through S/(S+B) value

Wilks' theorem assumes asymptotic behaviour, Feldman-Cousin approach is used for $B^0 \to \mu^+\mu^-$

Feldman-Cousins method for B^0 significance

- Since the Wilks' theorem shows $B^0 \to \mu^+\mu^-$ signal significance slightly above 3 σ level, a Feldman-Cousins based method is performed.

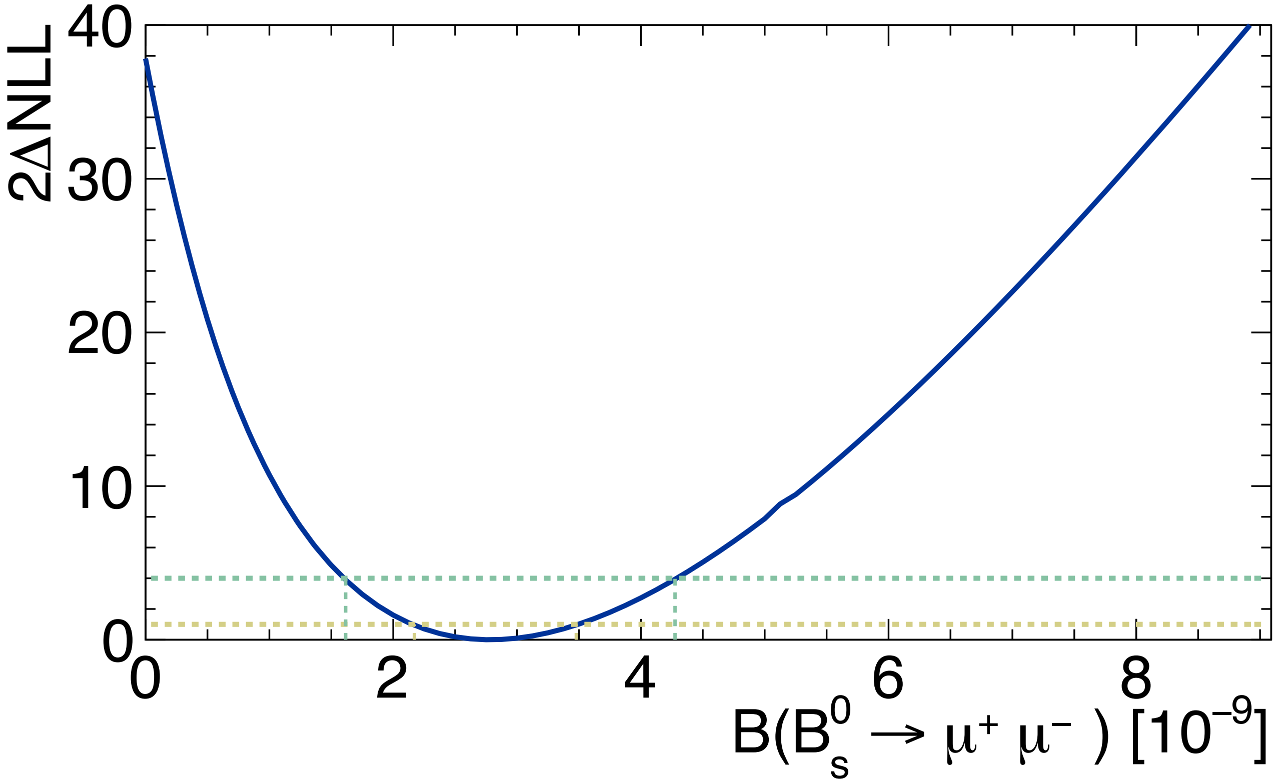


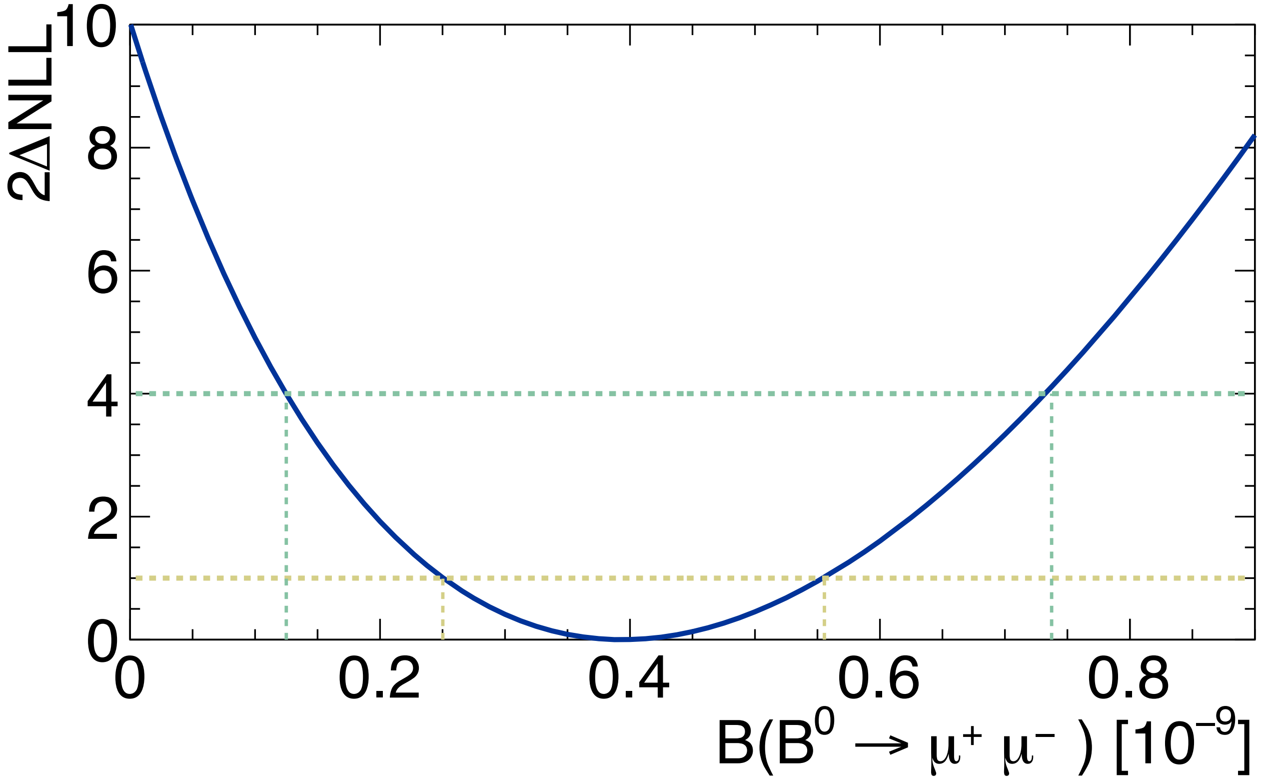
p-value = $(1.34 \pm 0.05) \times 10^{-3}$ corresponding to 3.0 σ

1 σ $\mathcal{B}(B^0 \to \mu^+\mu^-) \in [2.50, 5.55] \times 10^{-10}$

2 σ $\mathcal{B}(B^0 \to \mu^+\mu^-) \in [1.38, 7.40] \times 10^{-10}$

Likelihood profile





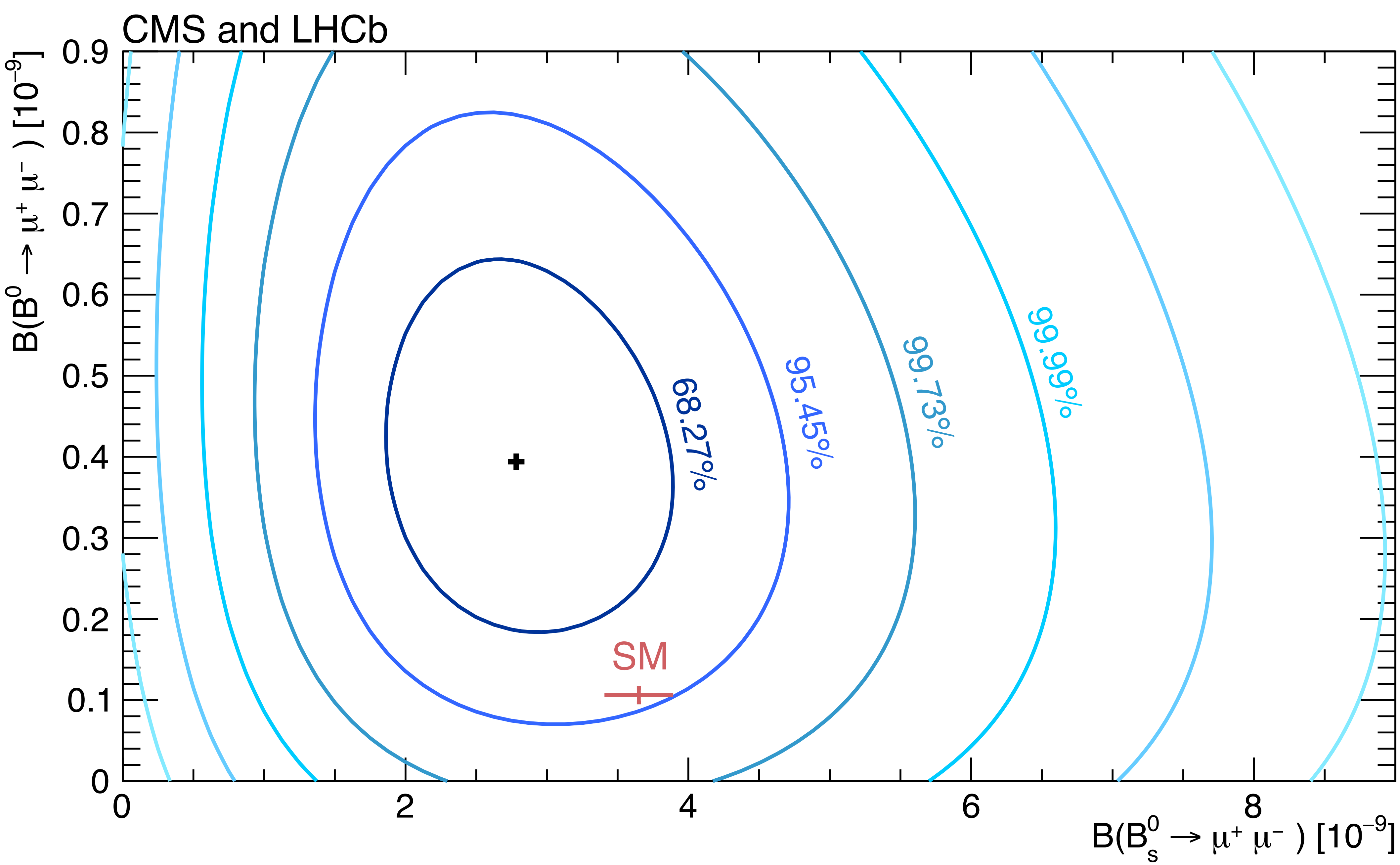
1 σ $\mathcal{B}(B_s^0 \to \mu^+\mu^-) \in [2.2, 3.5] \times 10^{-9}$

2 σ $\mathcal{B}(B_s^0 \to \mu^+\mu^-) \in [1.6, 4.3] \times 10^{-9}$

1 σ $\mathcal{B}(B^0 \to \mu^+\mu^-) \in [2.5, 5.6] \times 10^{-10}$

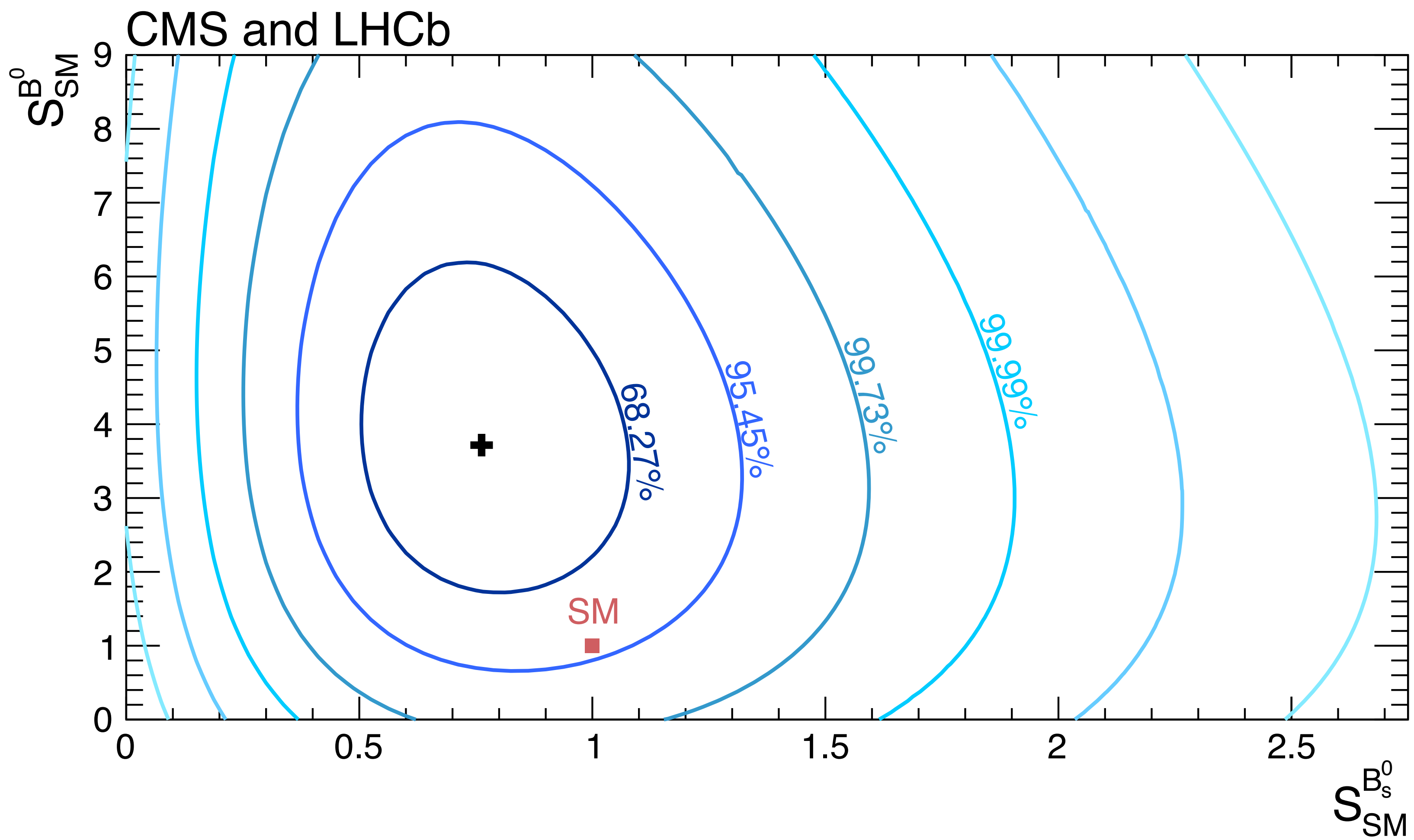
2 σ $\mathcal{B}(B^0 \to \mu^+\mu^-) \in [1.3, 7.4] \times 10^{-10}$

2D likelihood scan of Branching Fractions



Probability contours plot from 2D likelihood scan

2D likelihood scan of $\mathcal{B}/\mathcal{B}^{\mathrm{SM}}$



- fit and 2D LL scan also performed with the ratio of BF's over their SM expectation value:

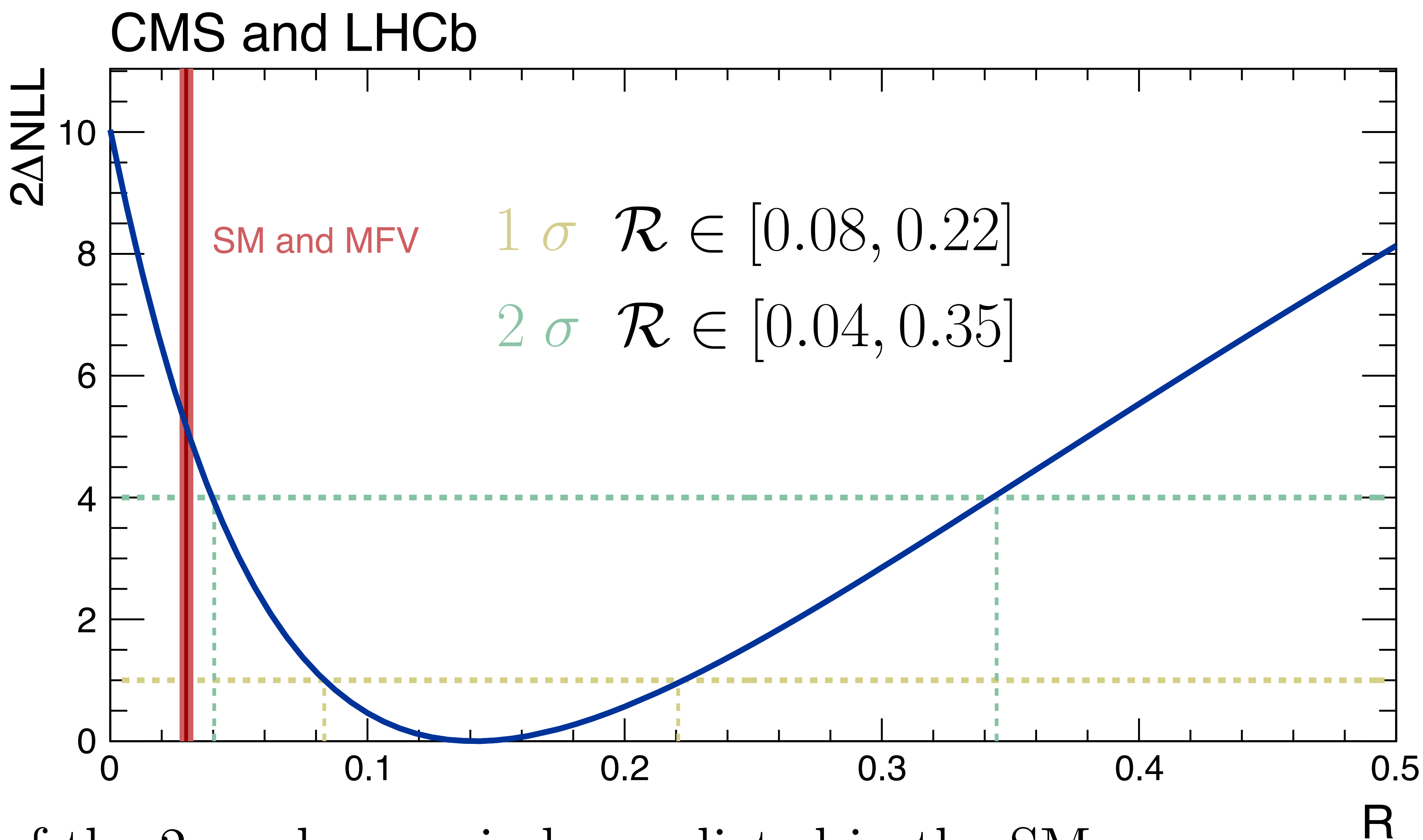
$$\mathcal{S}_{\mathrm{SM}}^{B^0_{(s)}} = \frac{\mathcal{B}(B^0_{(s)}\to\mu^+\mu^-)}{\mathcal{B}(B^0_{(s)}\to\mu^+\mu^-)^{\mathrm{SM}}}$$

- theoretical error included in the fit

$$\mathcal{S}_{\mathrm{SM}}^{B^0_s} = 0.76^{+0.20}_{-0.18} \qquad \mathcal{S}_{\mathrm{SM}}^{B^0} = 3.7^{+1.6}_{-1.4}$$

✶Compatibility with the SM: 2.2 σ for B^0 and 1.2 σ for $B^0{}_s$

1D likelihood profile of $\mathcal{B}(B^0_s \to \mu\mu)/\mathcal{B}(B^0 \to \mu\mu)$

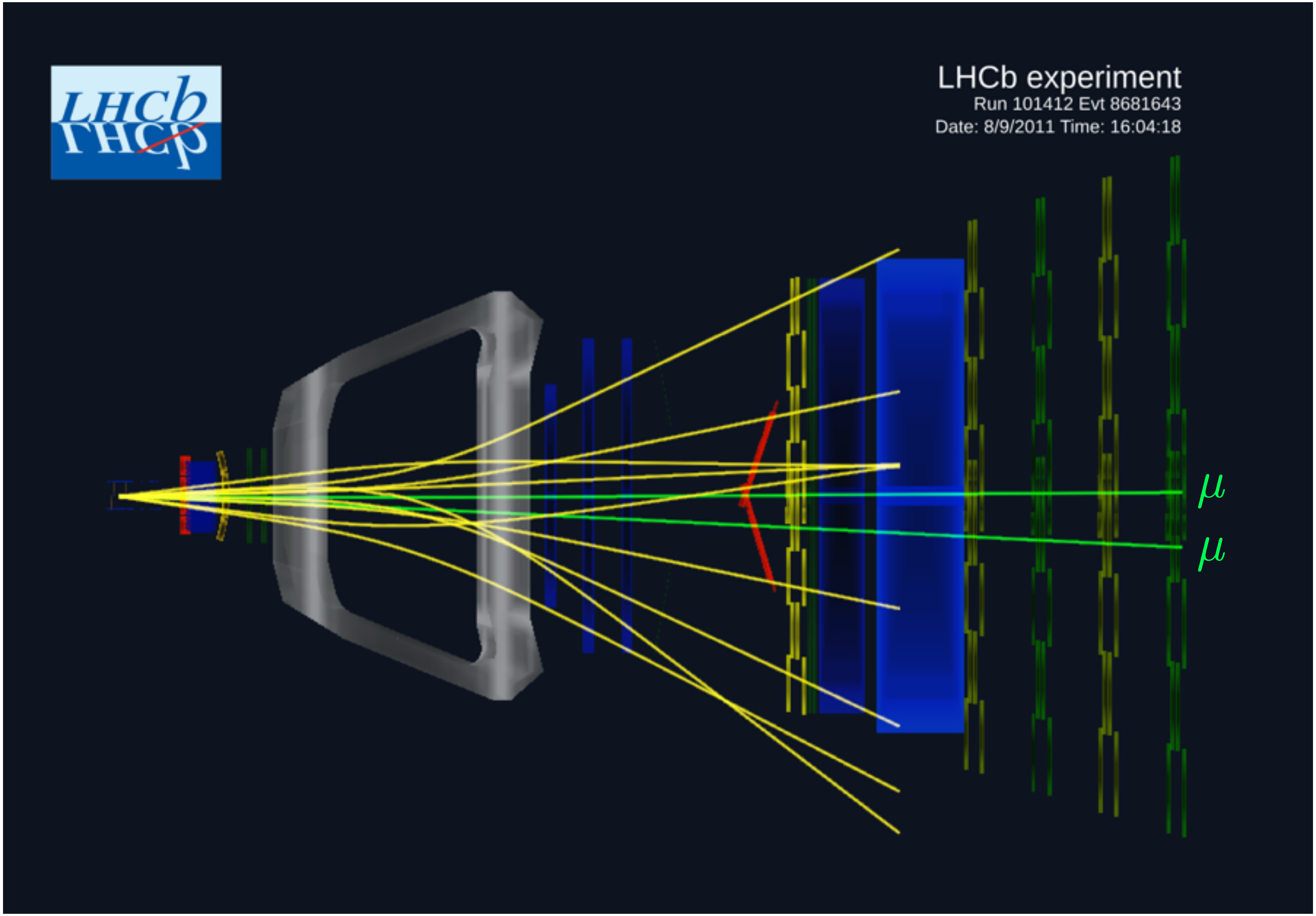


- Ratio of the 2 modes precisely predicted in the SM
- Useful to constraint MFV models with same flavour structure as SM
- Fit result:

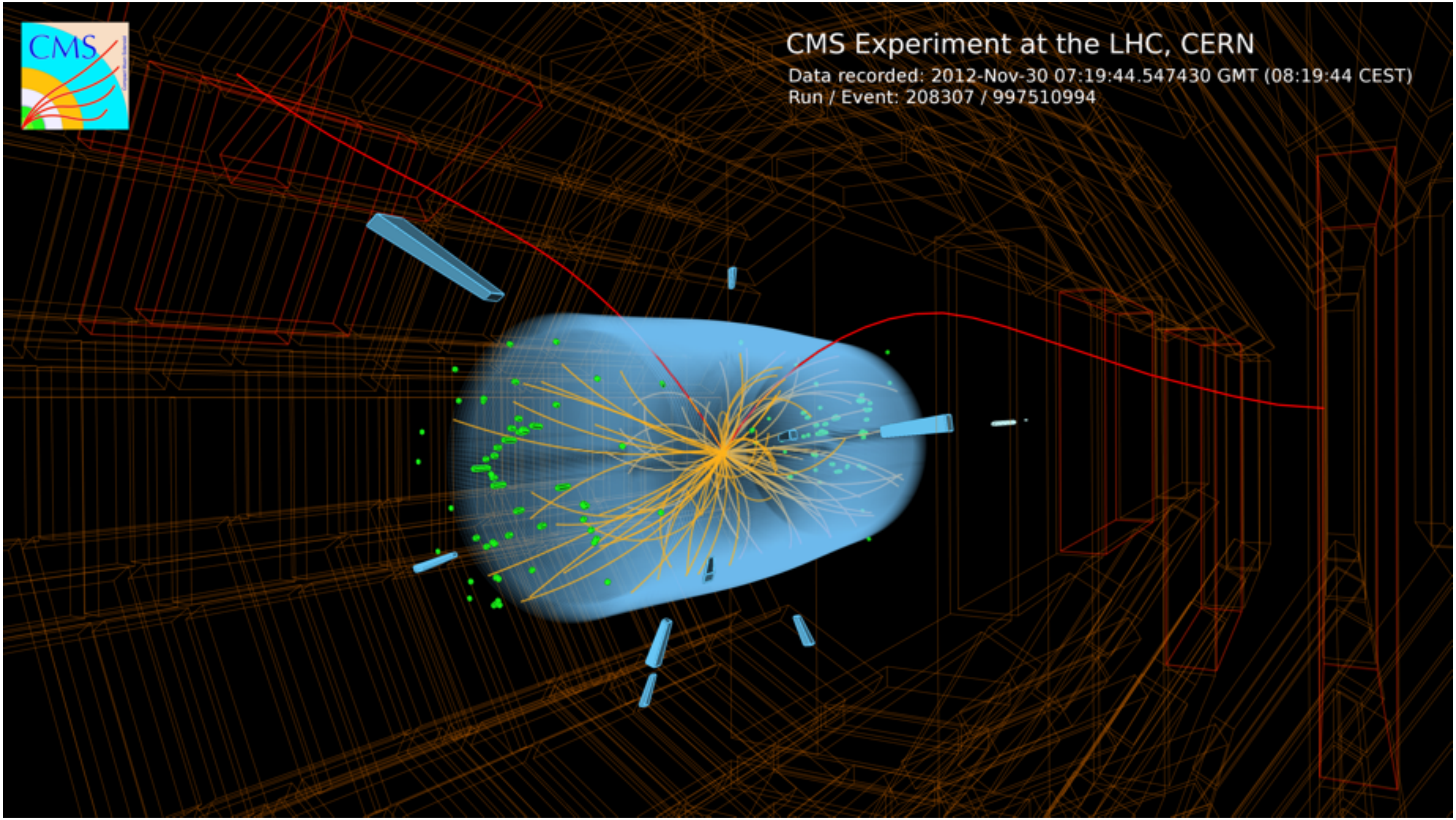
$$\mathcal{R} = 0.14^{+0.08}_{-0.06}$$

- Compatibility with the SM at 2.3 σ (including theoretical uncertainty)

LHCb event display



CMS event display



30 years ago...

PHYSICAL REVIEW D VOLUME 30, NUMBER 11 1 DECEMBER 1984

Two-body decays of *B* mesons

(Received 8 June 1984; revised manuscript received 10 September 1984)

Various exclusive and inclusive decays of B mesons have been studied using data taken with the CLEO detector at the Cornell Electron Storage Ring. The exclusive modes examined are mostly decays into two hadrons. The branching ratio for a B meson to decay into a charmed meson and a charged pion is found to be about 2%. Upper limits are quoted for other final states ψK^-, $\pi^+\pi^-$, $\rho^0\pi^-$, $\mu^+\mu^-$, e^+e^-, and $\mu^\pm e^\mp$. We also give an upper limit on inclusive ψ production and improved charged multiplicity measurements.

state particles with a pion mass or less. When the final-state particles are leptons the limits are improved by using the lepton identification capabilities of the CLEO detector.[14] For the decay $\overline{B}{}^0 \to \mu^+\mu^-$, we improve our limit by requiring that both muons penetrate the iron and produce signals in drift chambers. We find no such events. After correcting for detection efficiency (33%), we set an upper limit of 0.02% at 90% confidence for this decay. We improve our limit for $\overline{B}{}^0 \to e^+e^-$ by requiring that only one of the electrons be positively identified in the dE/dx and shower-chamber systems. One found candidate, coupled with a detection efficiency of 33%, gives a 90%-confidence-level upper limit of 0.03%. Finally, for the decay $\overline{B}{}^0 \to \mu^\pm e^\mp$, we require the muon to be identified but the electron needs to be positively identified if it is in the fiducial volume of the electron detectors. This pro-

$$\mathcal{B}(B^0 \to \mu^+\mu^-) < 2 \times 10^{-4} \text{ at } 90\% \text{ C.L.}$$

Conclusions

- Combined LHCb and CMS searches for the $B^0_{(s)} \to \mu^+ \mu^-$ decays with full RunI statistics.
- With a significance of 6.2 σ today we presented the first observation of $B^0_s \to \mu^+ \mu^-$ decay, compatibility with the SM at 1.2 σ.
- Excess at 3 σ level observed for the $B^0 \to \mu^+ \mu^-$ hypothesis with respect to the background-only hypothesis. This is compatibile with SM at 2.2 σ.
- Compatibility of the ratio of branching fractions with SM at level of 2.3 σ.
- ATLAS results not mentioned as 2012 analysis is still ongoing. It is anticipated that they will join the party soon.

Spares

Lifetime bias correction

- $B_s \to \mu^+\mu^-$ time dependent width:

$$\Gamma(B_s^0(t) \to \mu^+\mu^-) = (R_H + R_L)e^{-\Gamma_s t}\left[\cosh\frac{y_s t}{\tau_{B_s}} + \mathcal{A}_{\Delta\Gamma}\sinh\frac{y_s t}{\tau_{B_s}}\right]$$

where:

$$y_s = \frac{\Gamma_L - \Gamma_H}{\Gamma_L + \Gamma_H} = 0.01615 \pm 0.0085 \qquad \mathcal{A}_{\Delta\Gamma} = \frac{\Gamma_{B^0_{s,H}\to\mu\mu} - \Gamma_{B^0_{s,L}\to\mu\mu}}{\Gamma_{B^0_{s,H}\to\mu\mu} + \Gamma_{B^0_{s,L}\to\mu\mu}} \stackrel{SM}{=} 1$$

- So the time integrated efficiency is model dependent: $\varepsilon = \frac{\int \varepsilon(t)\Gamma_{\mathcal{A},y_s}(t)dt}{\int \Gamma_{\mathcal{A},y_s}(t)dt}$

- Normalization to be corrected to take into account this effect:

$$\delta_\epsilon = \frac{\epsilon^{\mathcal{A}_{\Delta\Gamma},y_s}}{\epsilon^{MC}} = \frac{\int_0^\infty \Gamma(B_s^0(t) \to \mu^+\mu^-, \mathcal{A}_{\Delta\Gamma}, y_s)\epsilon(t)\mathrm{d}t}{\int_0^\infty \Gamma(B_s^0(t) \to \mu^+\mu^-, \mathcal{A}_{\Delta\Gamma}, y_s)\mathrm{d}t} \cdot \frac{\int_0^\infty e^{-\Gamma_{MC}t}\mathrm{d}t}{\int_0^\infty e^{-\Gamma_{MC}t}\epsilon(t)\mathrm{d}t}$$

Correction for B_S = 4.50±0.03%
Correction for B^0 = 1.48±0.01%

- also corrected because time dependent