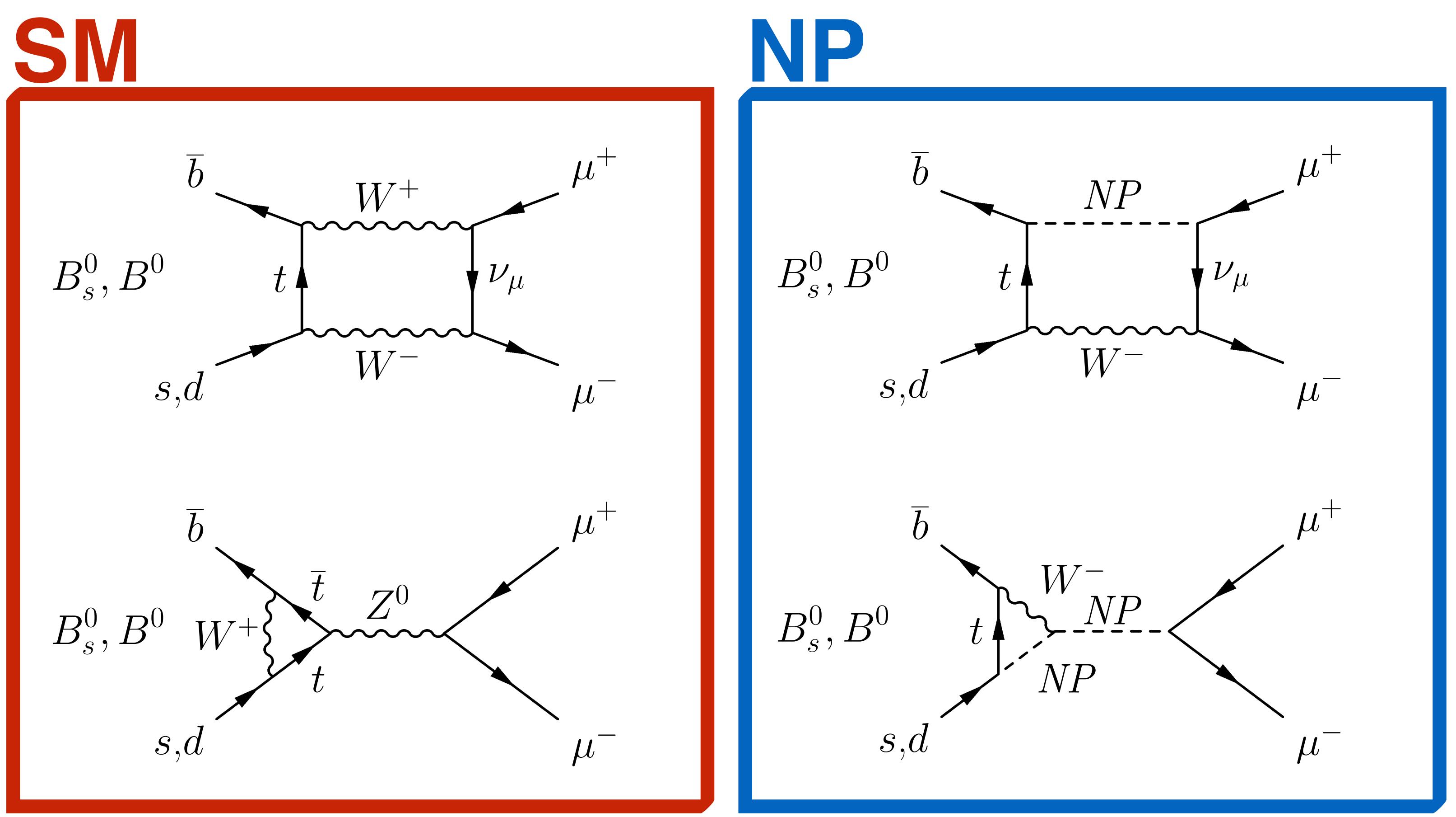
Overview of $B_s \to \mu\mu$ decays at LHCb

Flavio Archilli
(CERN)

Novel aspects of b to s transitions: investigating new channels
5-7 October 2015
St Charles Campus, Université Aix-Maseille

theory

- 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 $\tan^6\beta/m_A{}^4$



theory

The branching fraction of $B^0_{(s)} \rightarrow \mu^+\mu^-$ can be written in the SM as:

$$\mathcal{B}^{t=0}(B_s^0 \to \mu^+\mu^-) = \frac{G_F^4 M_W^4}{\pi^2} \tau_{B_s} f_{B_s}^2 m_{B_s} m_\mu^2 \sqrt{1 - \frac{4m_\mu^2}{m_{B_s}^2}} |V_{tb}V_{ts}^*|^2 |C_{10}|^2 + \dots$$

where:

- f_{Bs}^2 contains the non-perturbative QCD of the meson decay Bs obtained through lattice calculations
- $V_{tb}V_{ts}^*$ are the elements of the CKM matrix
- $|C_{10}|^2$ contain all information about short-distance physics

theory

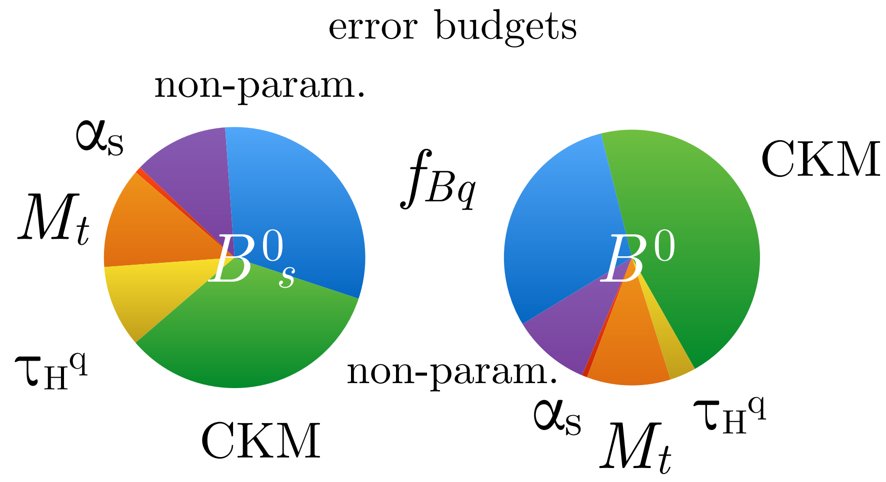
- Untagged time integrated branching fraction predictions:

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

updated with the latest top mass measurement (Tevatron+LHC combination)

[hep-ex/1403.4427]

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



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

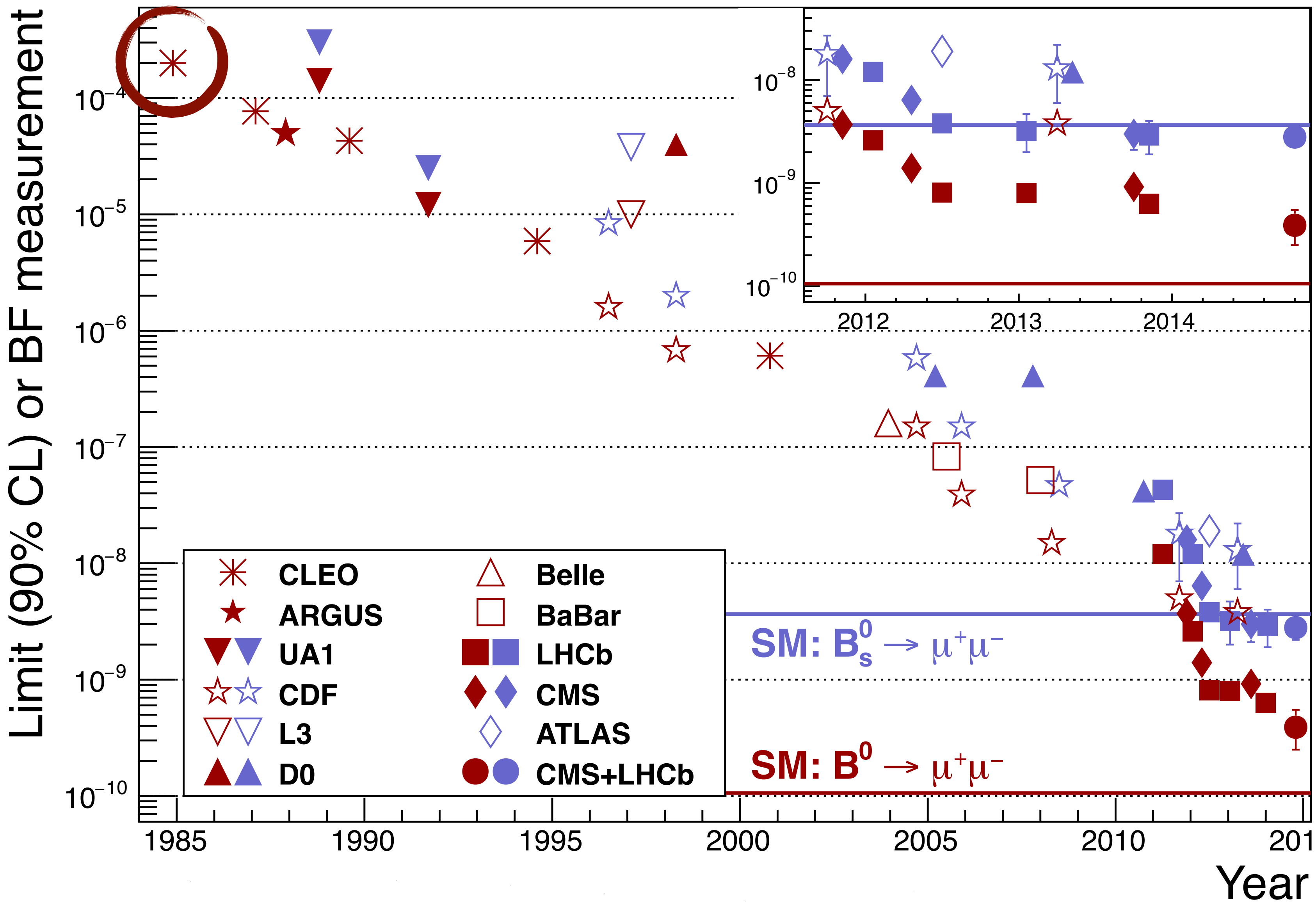
$$
\mathcal{R} = \frac{\mathcal{B}(B^0 \to \mu^+\mu^-)}{\mathcal{B}(B_s^0 \to \mu^+\mu^-)} = \frac{\tau_{B_d}}{1/\Gamma_H^s} \left(\frac{f_{B_d}}{f_{B_s}}\right)^2 \left|\frac{V_{td}}{V_{ts}}\right|^2 \frac{M_{B_d}\sqrt{1-\frac{4m_\mu^2}{M_{B_d}^2}}}{M_{B_s}\sqrt{1-\frac{4m_\mu^2}{M_{B_s}^2}}} = 0.0295^{+0.0028}_{-0.0025}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

➡ stringent test of Minimal Flavour Violation hypothesis

the adventure begins

- **1984**: Cleo sets the first upper limit on B^0

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 $\bar{B}^0 \to \mu^+\mu^-$, 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 $\bar{B}^0 \to e^+e^-$ 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 $\bar{B}^0 \to \mu^\pm e^\mp$, 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-



PHYSICAL REVIEW D VOLUME 30, NUMBER 11 1 DECEMBER 1984

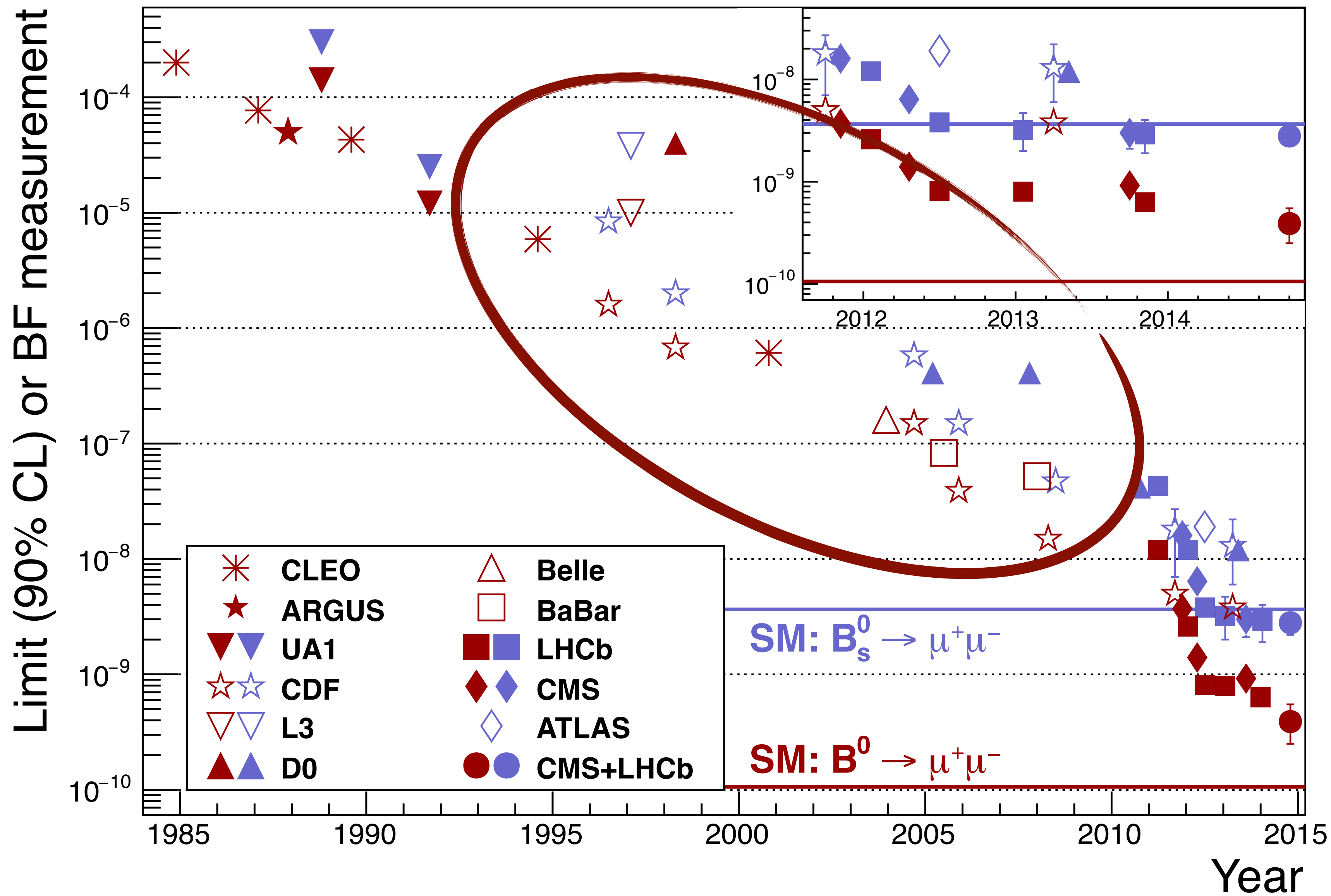
Two-body decays of *B* mesons

(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.

the adventure begins

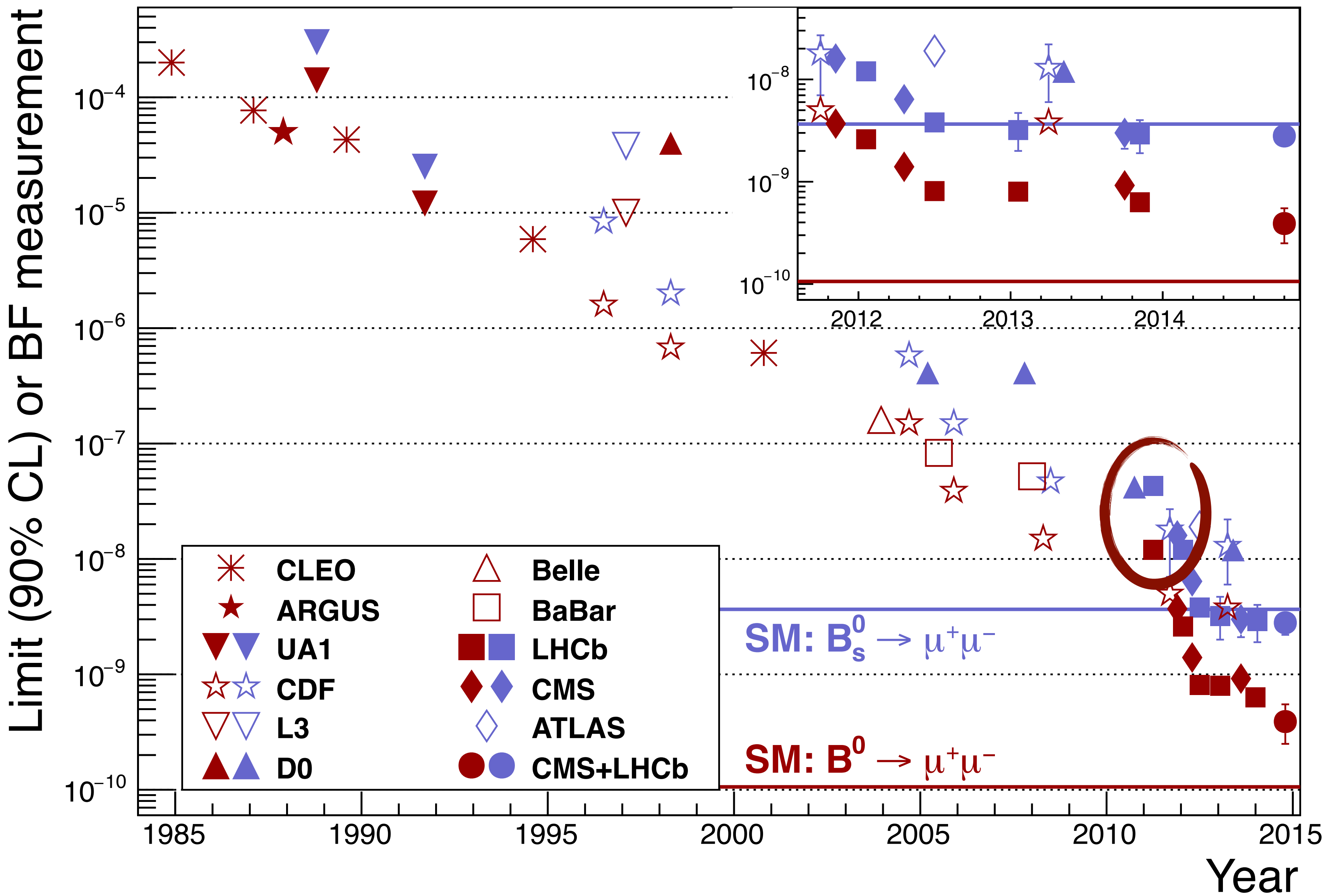
- **1984**: Cleo sets the first upper limit on B^0
- **1998-2010**: Impressive effort from the B-factories and Tevatron to push down the limit



the adventure begins

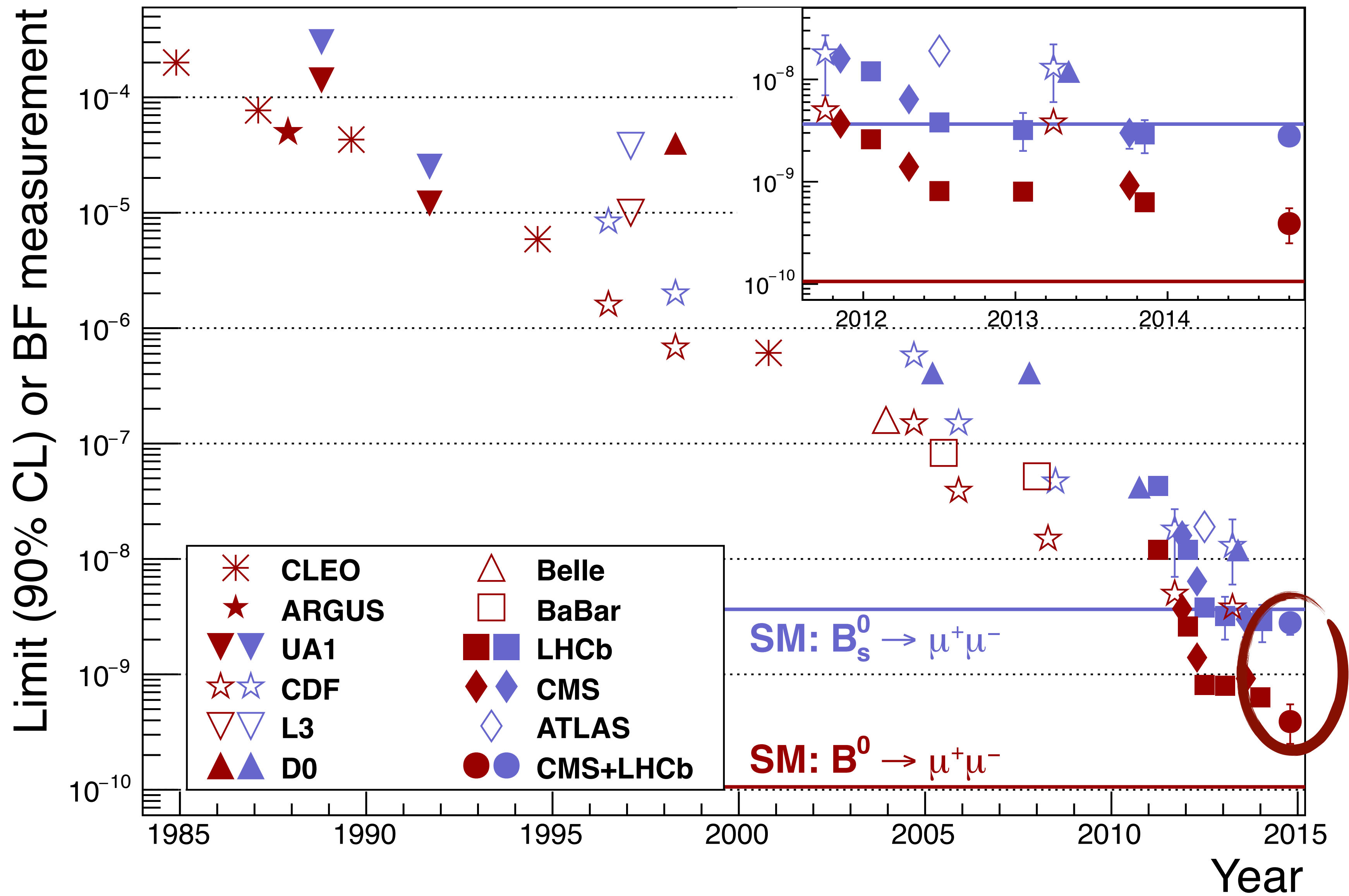
- **1984**: Cleo sets the first upper limit on B^0
- **1998-2010**: Impressive effort from the B-factories and Tevatron to push down the limit
- **2011**: LHC era begins. LHCb first measurement on $0.37 fb^{-1}$ is already competitive with CDF.
- **2012**: LHCb found the first evidence $B^0_s \to \mu^+\mu^-$ with $2.1 fb^{-1}$

[PRL 110, 021801 (2013)]



the adventure begins

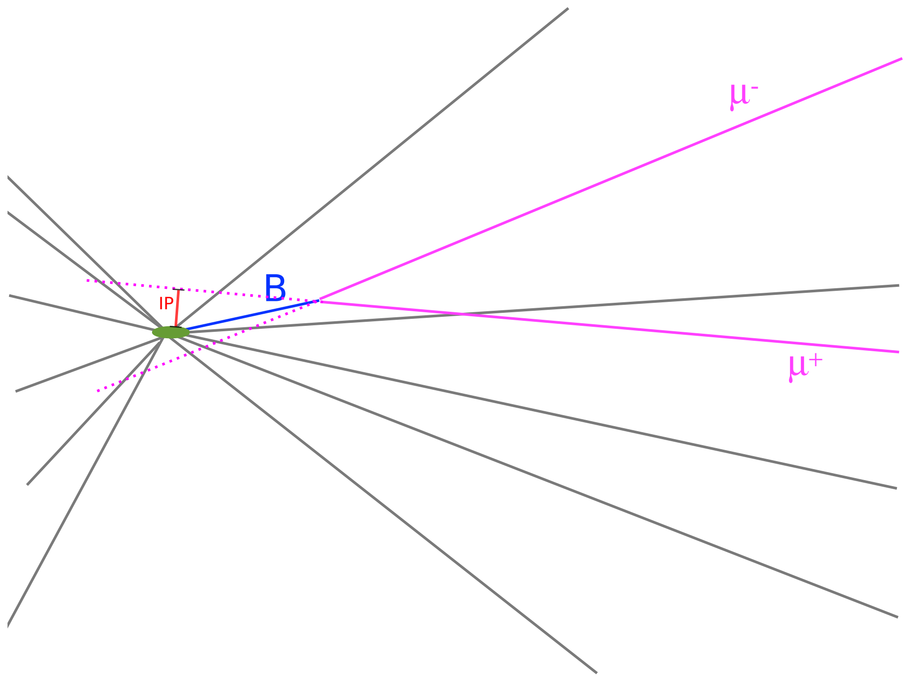
- **1984**: Cleo sets the first upper limit on B^0
- **1998-2010**: Impressive effort from the B-factories and Tevatron to push down the limit
- **2011**: LHC era begins. LHCb first measurement on $0.37 fb^{-1}$ is already competitive with CDF.
- **2015**: LHCb+CMS analysis combination. First observation of $B_s \to \mu^+\mu^-$

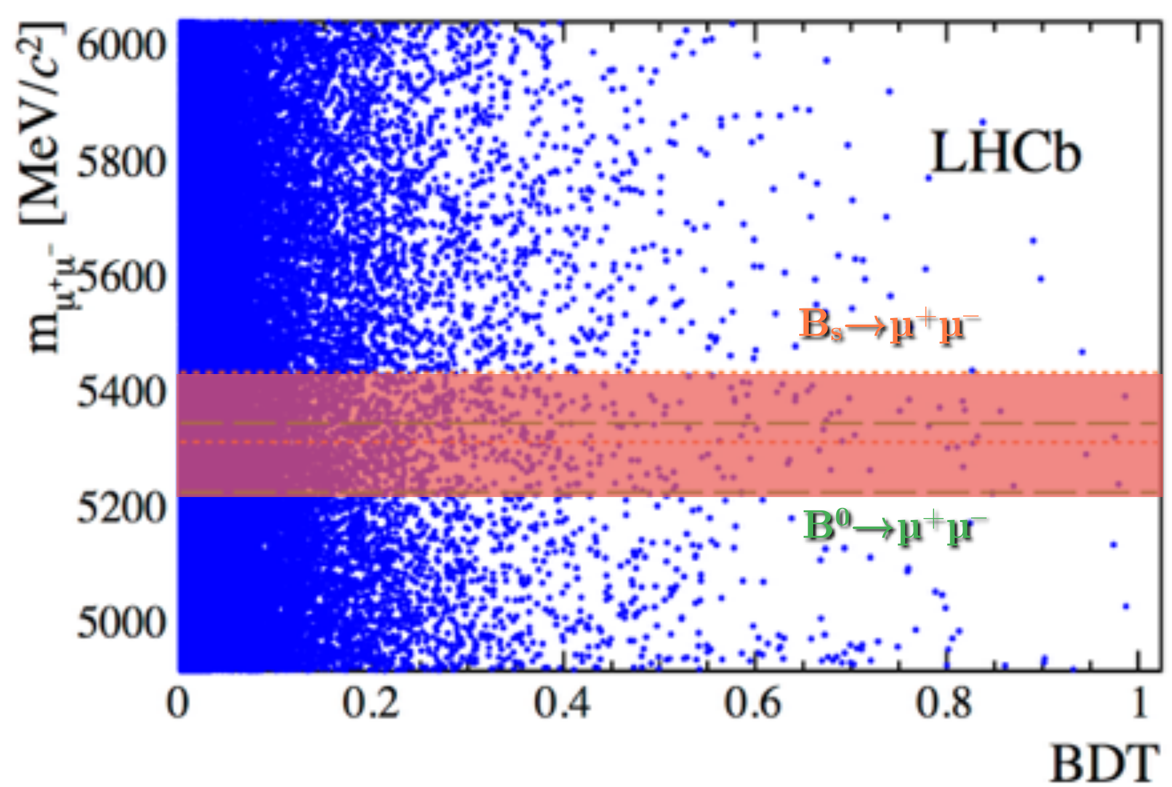


[Nature 522, 68-72 (04 June 2015)]

$B^0_{(s)} \to \mu^+\mu^-$ analysis strategy @LHCb

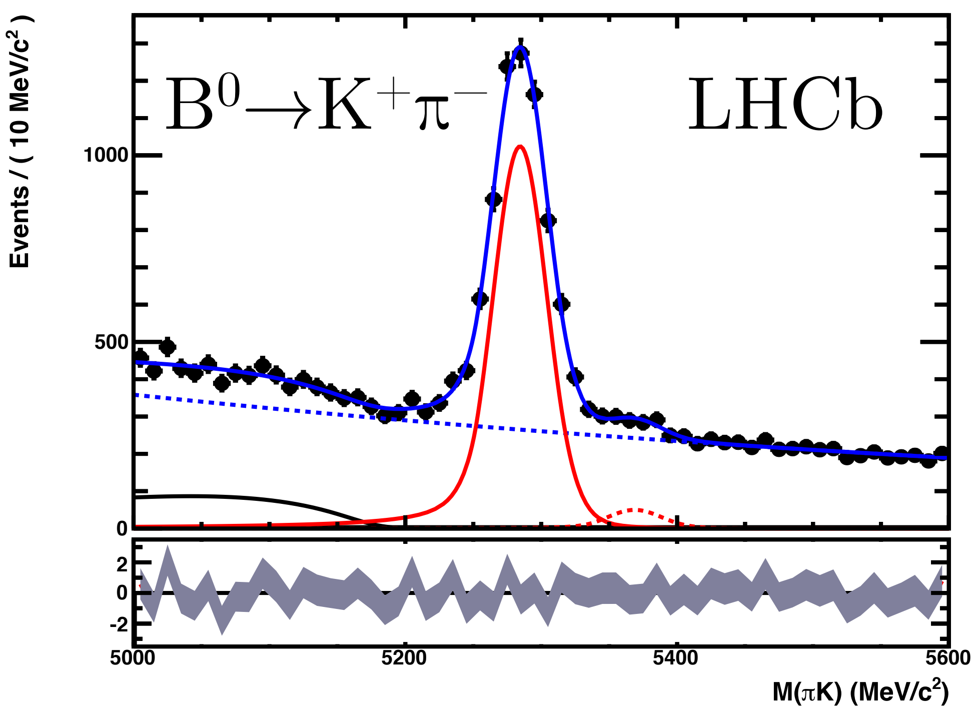
- **Loose selection**:
  - A pair of opposite charged muons with and $m_{\mu\mu} \in [4900,6000]$ MeV/c^2
  - Good quality vertex displaced respect to the interaction point
- **Sig. and bkg. classification in $m_{\mu\mu}$ vs MVA classifier (BDT) plane**:
  - BDT based on kinematic and geometrical variables, trained with MC calibrated on data
  - Search window kept blind until analysis optimised
- **Data driven calibration**
- **Comparison between expectation and observed events**:
  - Branching fraction from unbinned likelihood fit
  - Upper limit from CLs method

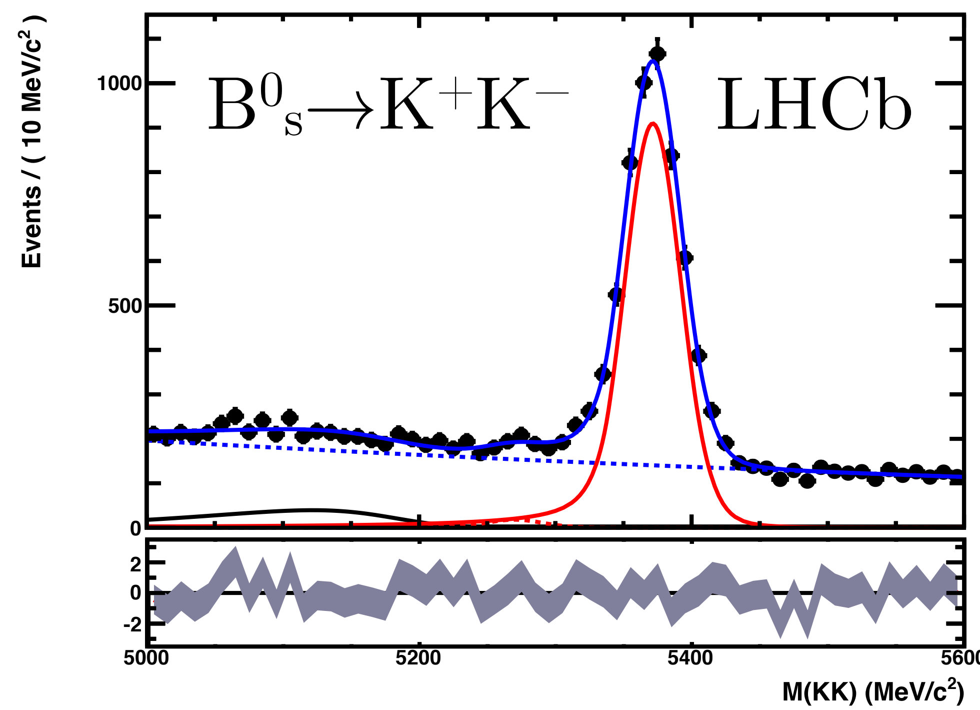




Where is the signal peak?

- Determination of mass peak position and resolution essential for this search
  - mean value calibrated with well visible exclusive B→hh’ decays
  - resolution from di-μ resonance and exclusive B→hh’

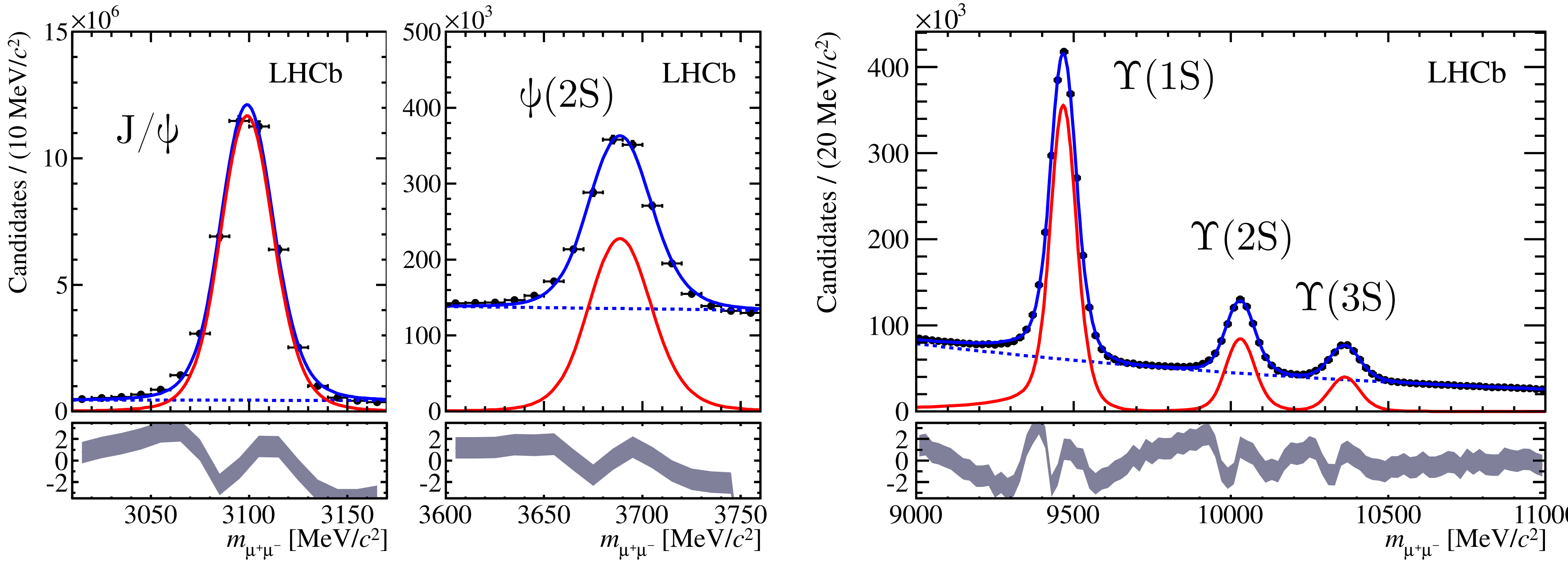




m_{B^0}	$(5284.36 \pm 0.26_{\text{stat}} \pm 0.13_{\text{syst}})$ MeV$/c^2$
$m_{B^0_s}$	$(5371.55 \pm 0.41_{\text{stat}} \pm 0.16_{\text{syst}})$ MeV$/c^2$

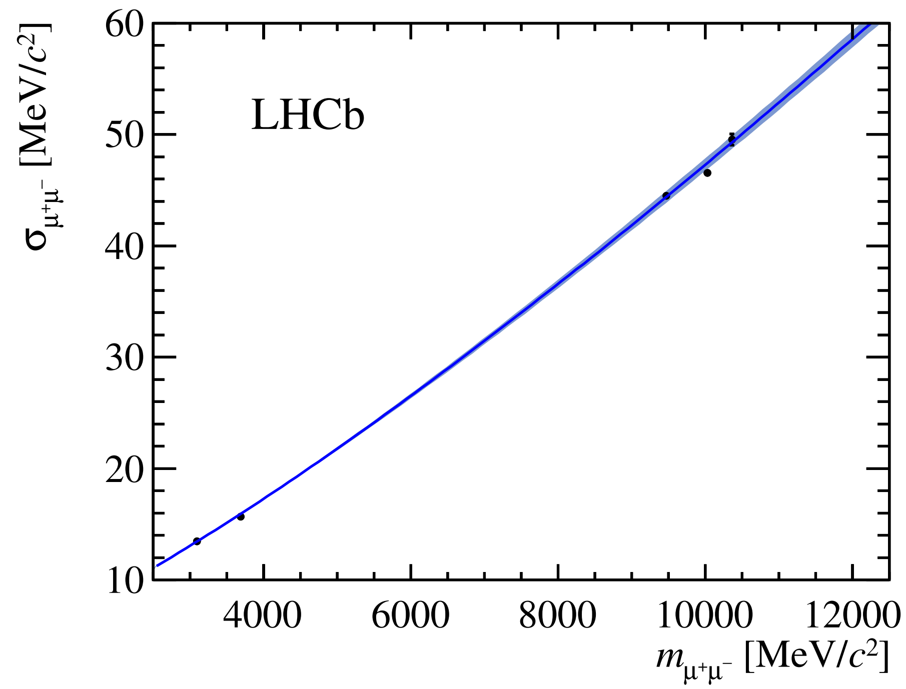
Peak position determinations for 7 TeV and 8 TeV data agree at better than $5{\times}10^{-4}$

Where is the signal peak?



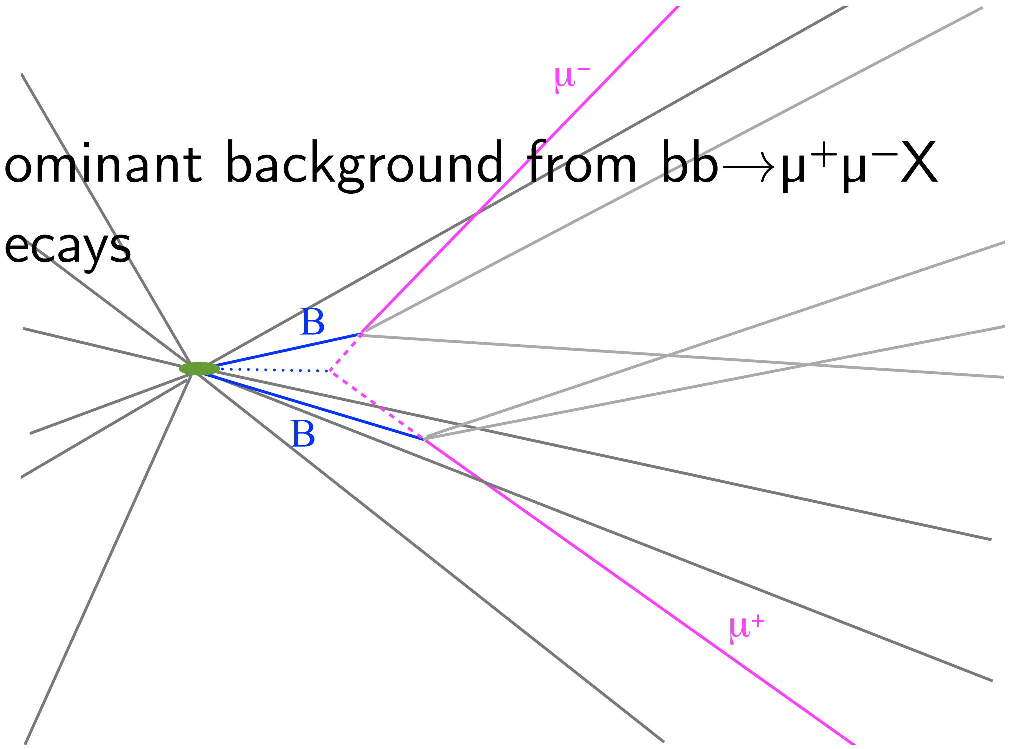
- fit to di-muon resonances from J/ψ up to ϒ(3S) and interpolate the resolution at B0 and Bs mass values.
- Compare with the resolution obtained with hadronic 2 body B decays

$$\sigma_{B^0} = (24.6 \pm 0.4)\,\mathrm{MeV/c^2}$$
$$\sigma_{B^0_s} = (25.0 \pm 0.4)\,\mathrm{MeV/c^2}$$

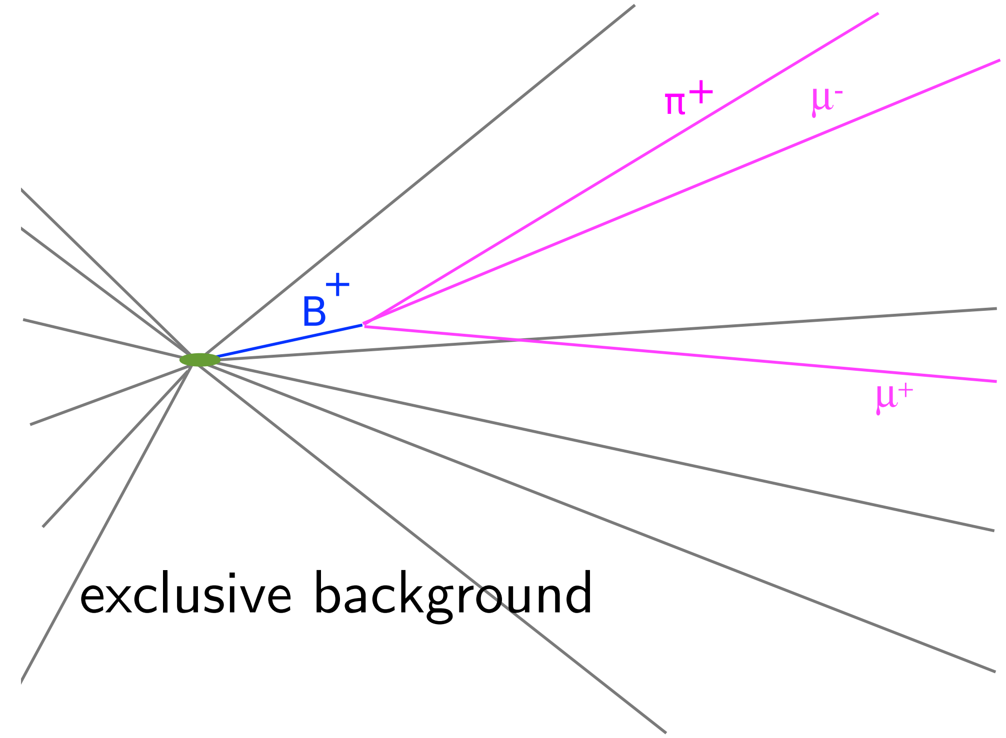


backgrounds

dominant background from $b\bar{b} \to \mu^+\mu^- X$ decays



exclusive background



These background sources can affect the result in two ways:

1) non negligible contribution in the signal mass window, $m(B^0_{(s)}) \pm 60$ MeV/c^2

2) mass shape different from exponential → bias in the combinatorial background interpolation from mass sidebands

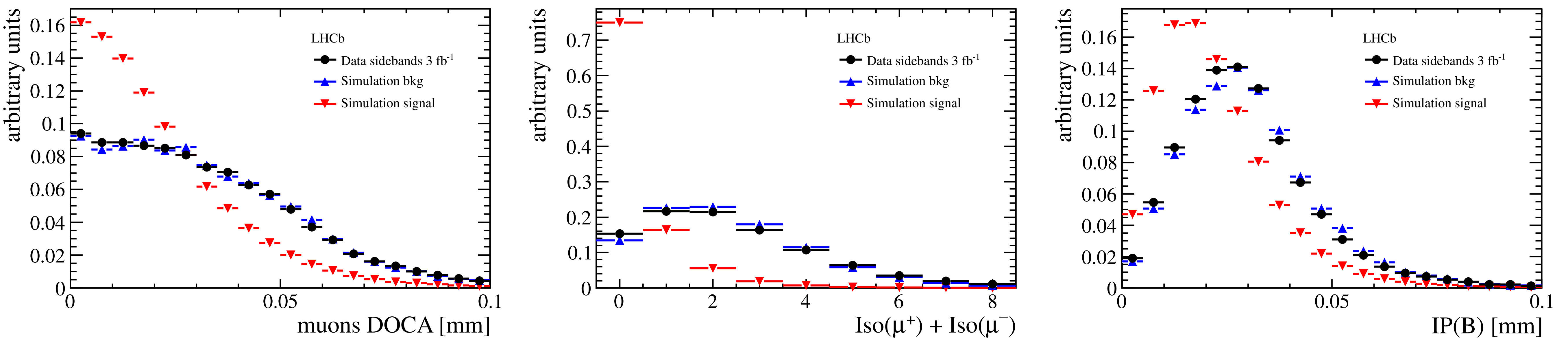
Hadronic	$B \to h^+ h'^-$ with $h = \pi, K$
Semileptonic	$B^0 \to \pi^- \mu^+ \nu$ $B^0_s \to K^- \mu^+ \nu$ $\Lambda^0_b \to \bar{p} \mu^+ \nu$ $B^+_c \to J/\psi \mu^+ \nu_\mu$ $B^0_s \to D^-_s (\to \mu^- \nu) \mu^+ \nu$
Rare decays	$B^{+(0)} \to \pi^{+(0)} \mu^+ \mu^-$ $B^0_s \to \mu^+ \mu^- \gamma$

in red the ones added as a separated component in the final fit

BDT classifier

- A multivariate classifier is trained using 12 variables to discriminate between signal and dominate background

- Optimised and trained on MC and calibrated using Data

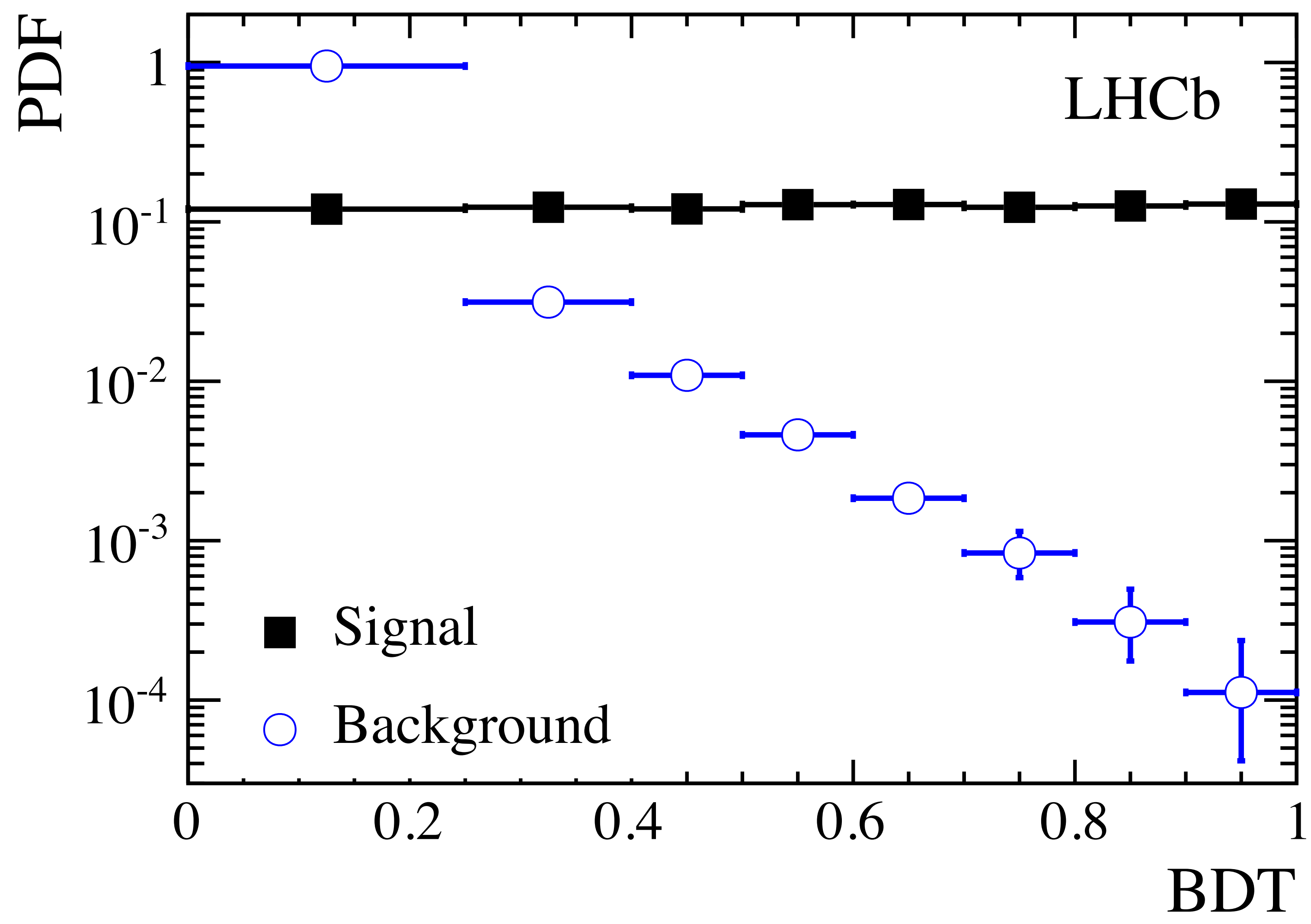


BDT calibration

- BDT output defined to be **flat for signal**, and peaked at zero for background

- Signal BDT shape from exclusive $B^0_{(s)} \to h^+ h'^-$ ($h = K, \pi$) events, which have same topology as the signal

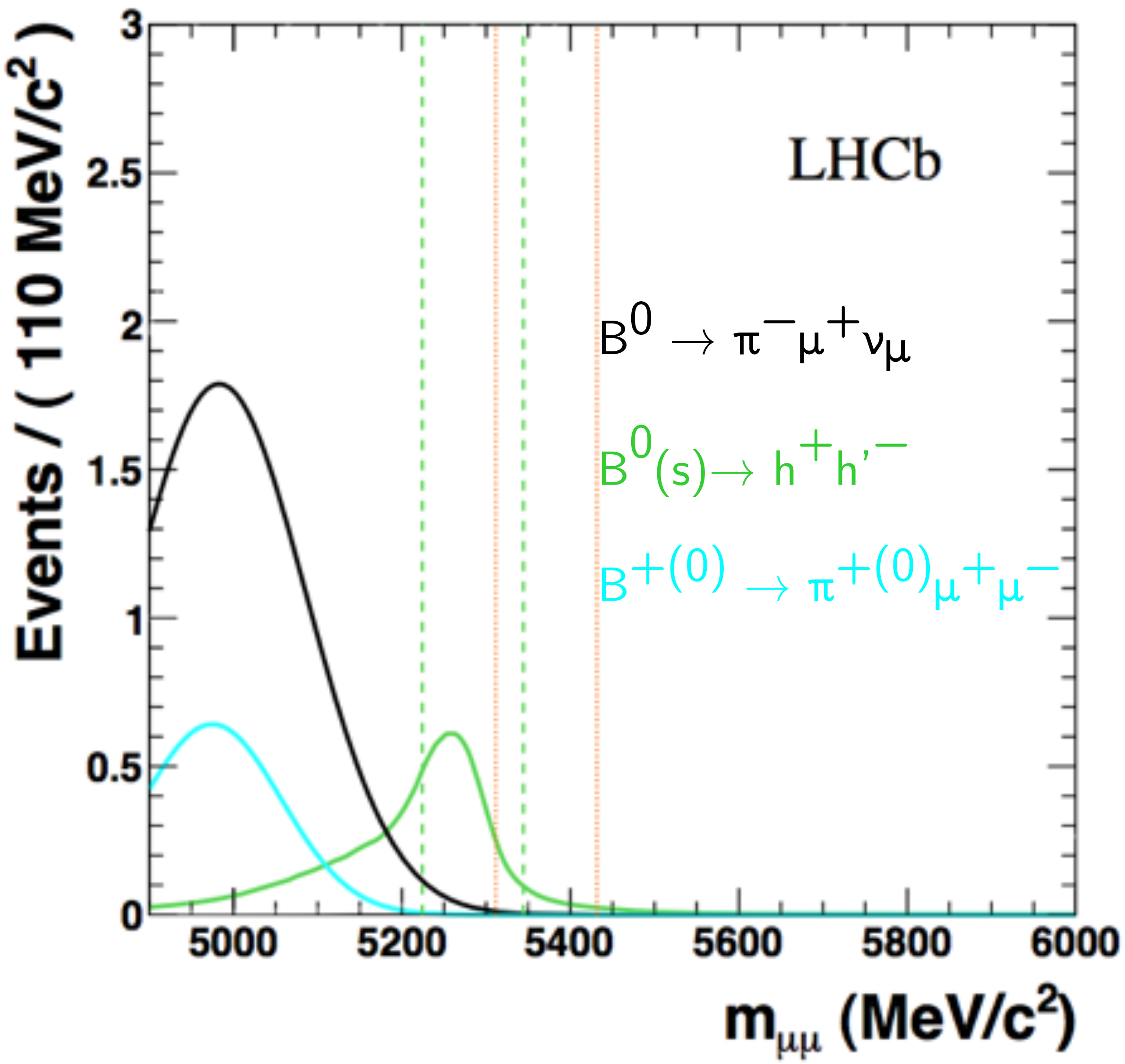
- Background BDT shape is evaluated on the di-muon mass sidebands



exclusive backgrounds

Several exclusive backgrounds pollute the lower mass sideband. Extremely important to constraint these sources to improve the sensitivity.

- Invariant mass and BDT distributions from high statistics MC samples, weighted by misID probabilities measured on data
- For **semileptonic** and **rare decays** expected yields evaluated by normalising to $B^{\pm} \to J/\psi K^{\pm}$
- **B→ h^{+}h'$^{-}$** yield fitted on data

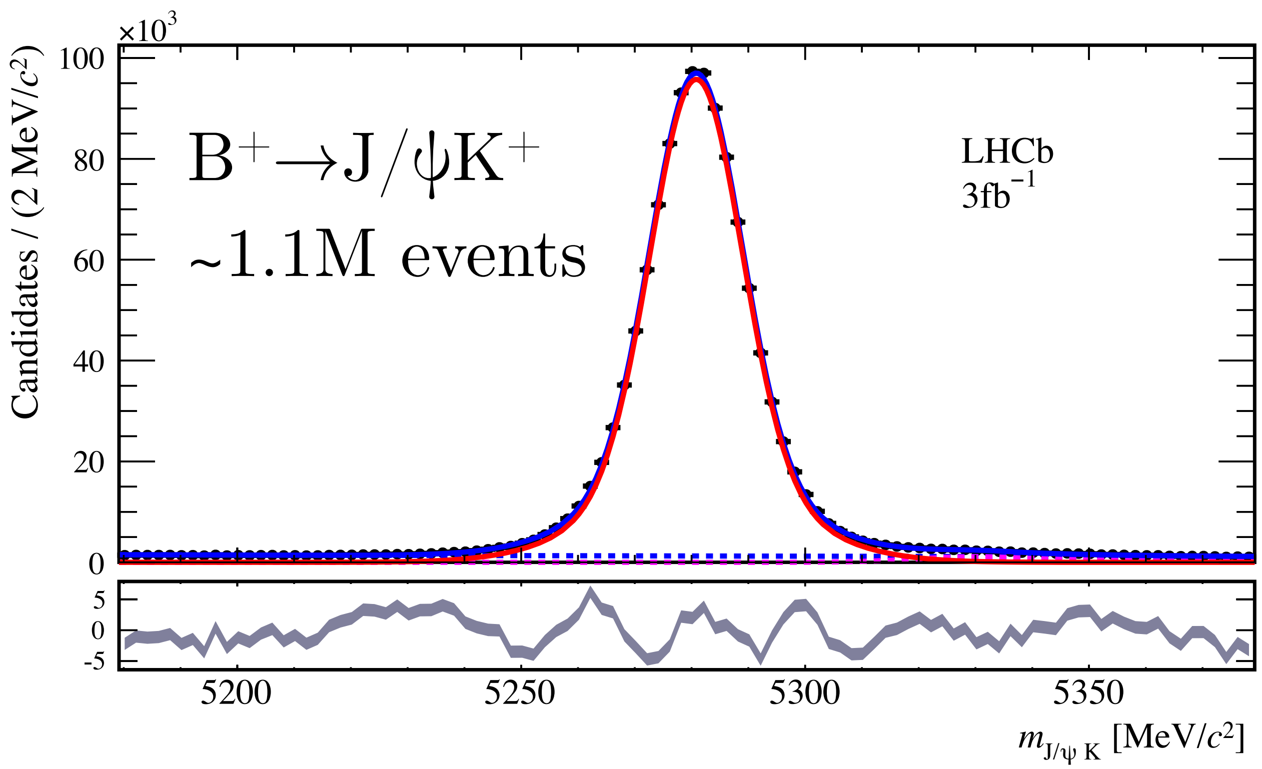


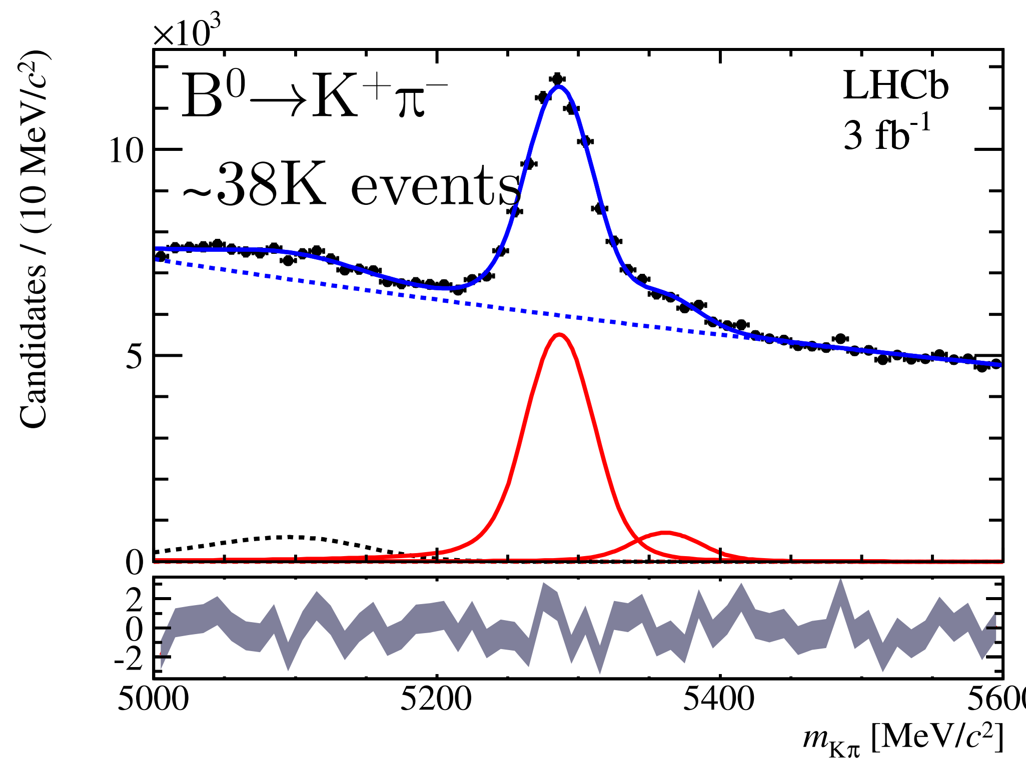
Normalisation

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

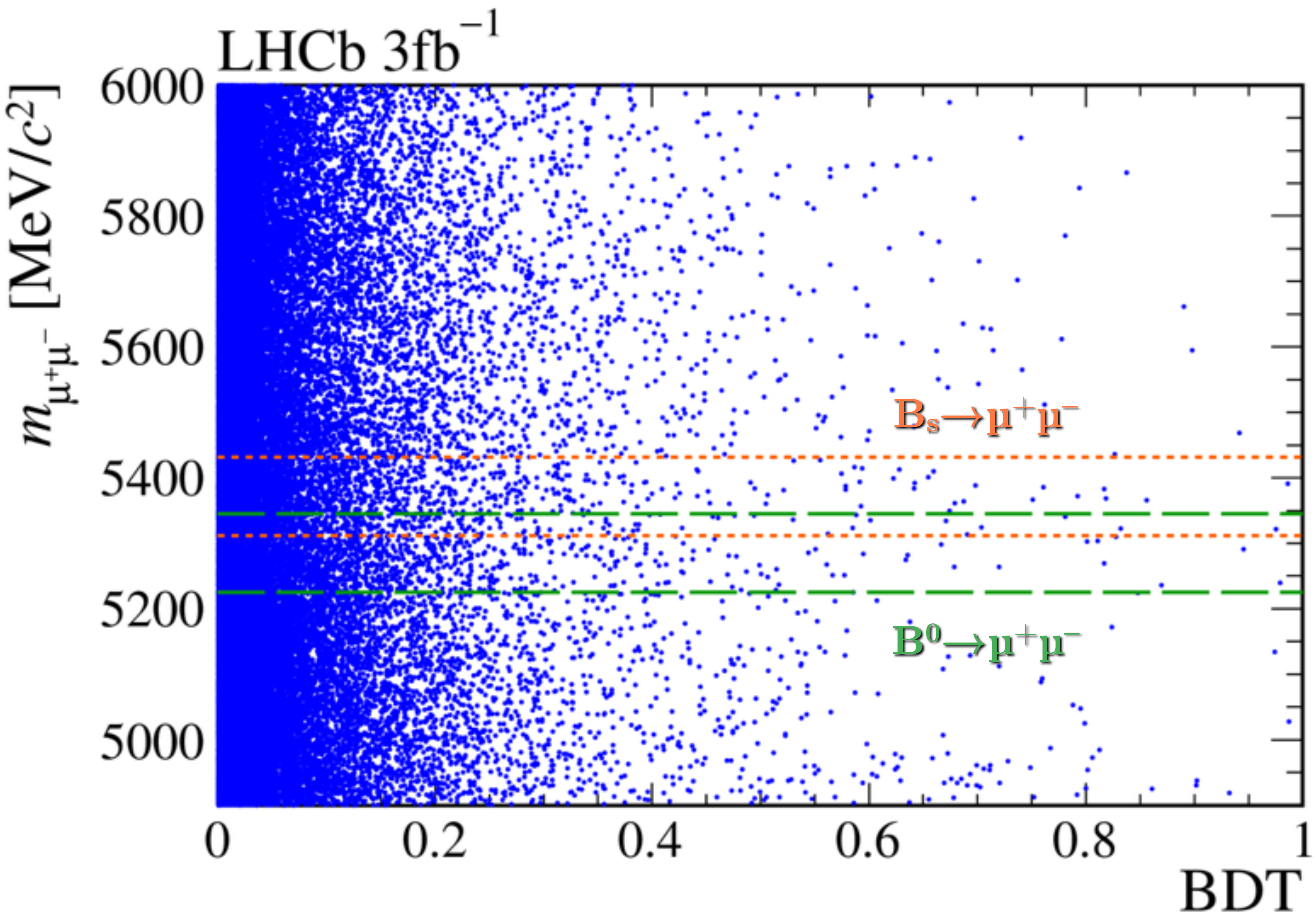
- Two control channels used for the normalization: B+→J/ψK+ and B0→K+π−
- From MC and x-checked on data
- Trigger efficiency from J/ψ→μμ data
- Hadronisation fraction from LHCb measurement $f_s / f_d = 0.259 \pm 0.015$
- Using the SM signal we expect 39.5±4.3 Bs→μ+μ− and 4.5±0.4 B0→μ+μ−

$\alpha_s = (9.41 \pm 0.65) \cdot 10^{-11}$
$\alpha = (2.40 \pm 0.09) \cdot 10^{-11}$

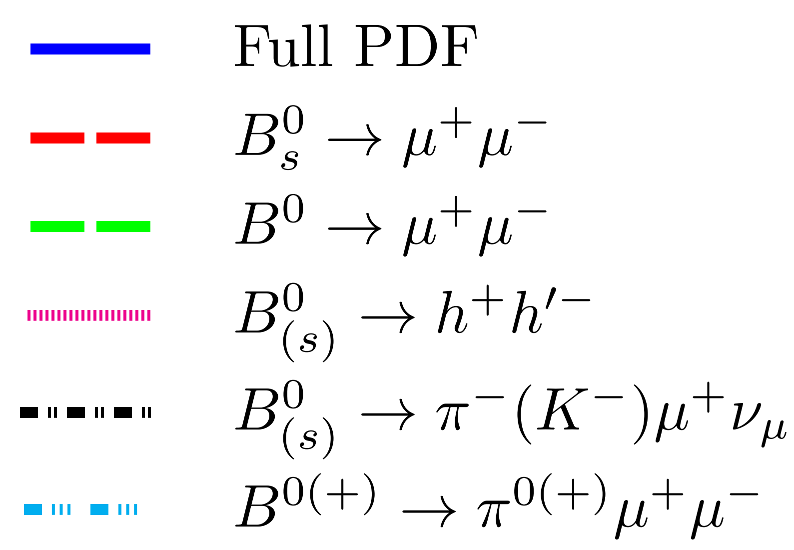




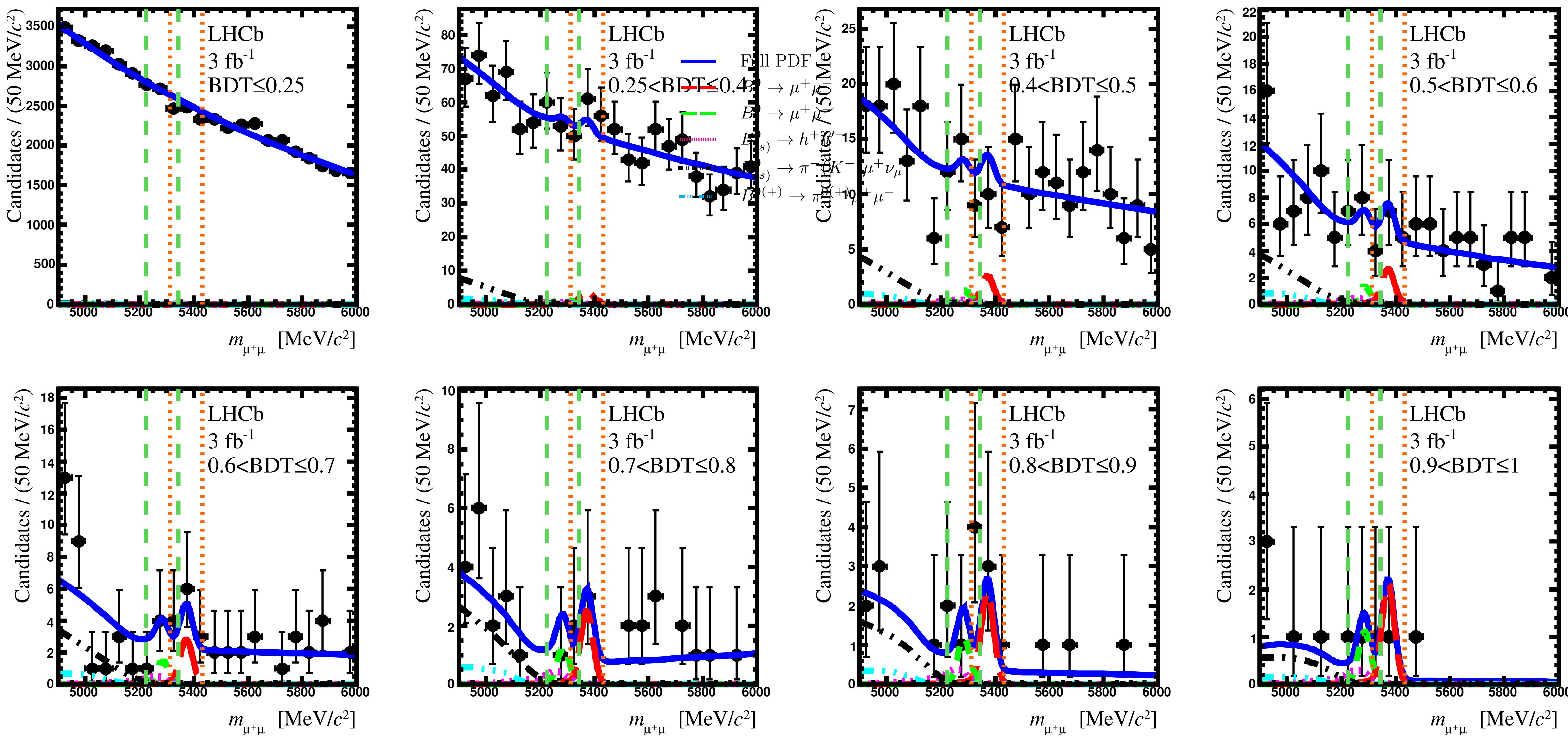
opening the box



Full fit



Simultaneous unbinned maximum likelihood fit to the mass spectra in 8 BDT bins



Results

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

[PRL 111 (2013) 101805]

- The full fit gives:

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (2.9^{+1.1}_{-1.0}(\text{stat})^{+0.3}_{-0.1}(\text{syst})) \times 10^{-9}$$

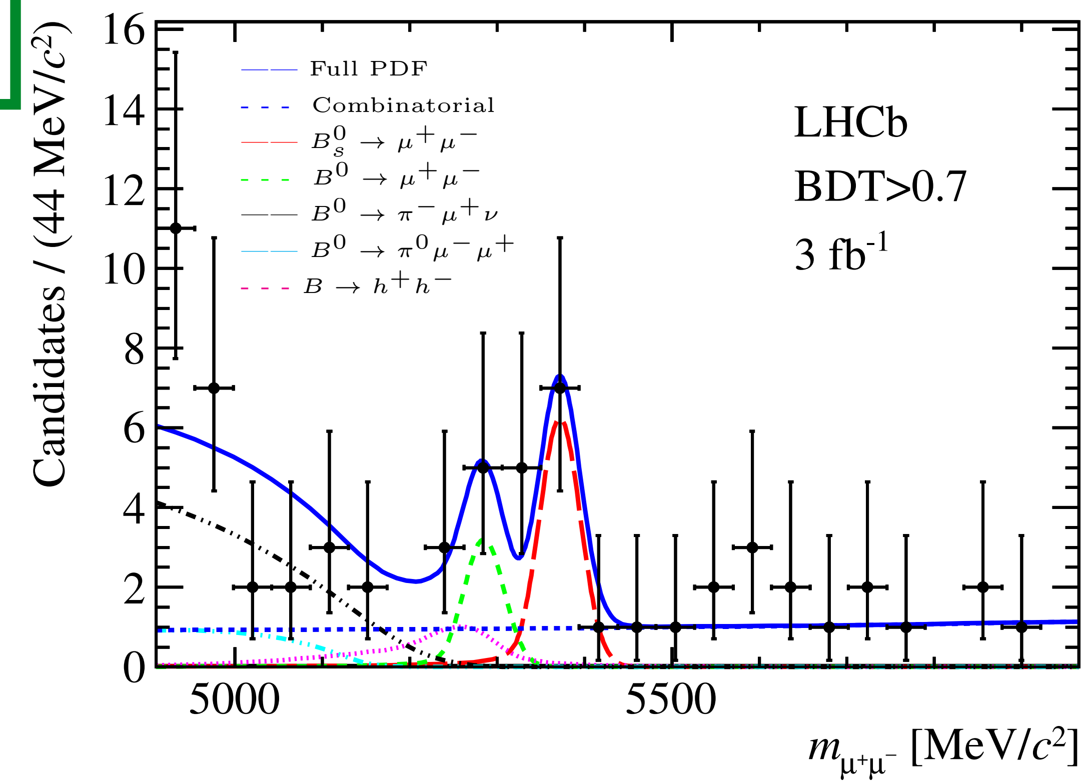
with a p-value of $6.8 \cdot 10^{-5}$ equivalent to a significance of 4.0 σ

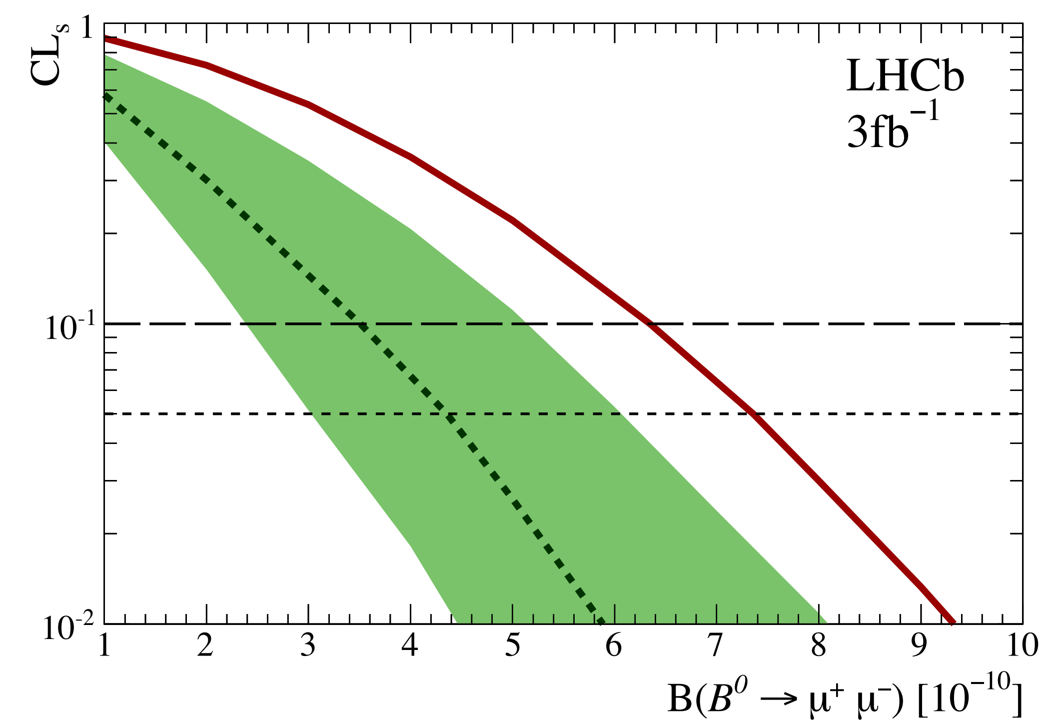
- For the B^0:

$$\mathcal{B}(B^0 \to \mu^+\mu^-) = (3.7^{+2.4}_{-2.1}(\text{stat})^{+0.6}_{-0.4}(\text{syst})) \times 10^{-10}$$

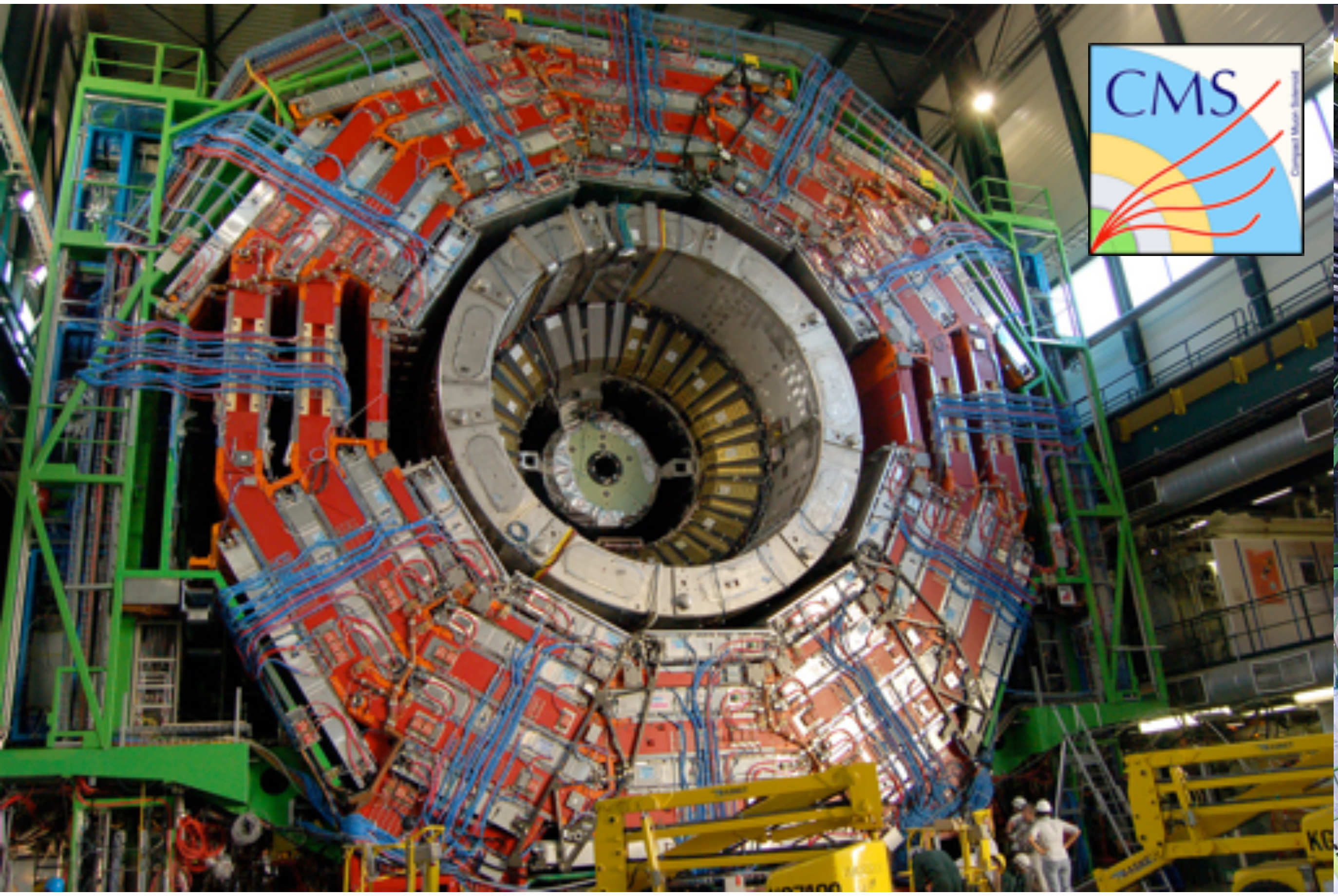
- with a p-value of $4.4 \cdot 10^{-2}$ equivalent to a significance of 2.0 σ
- Systematics from nuisance parameters and background model
- Given no evidence of $B^0 \to \mu^+\mu^-$ the following upper limit has been evaluated:

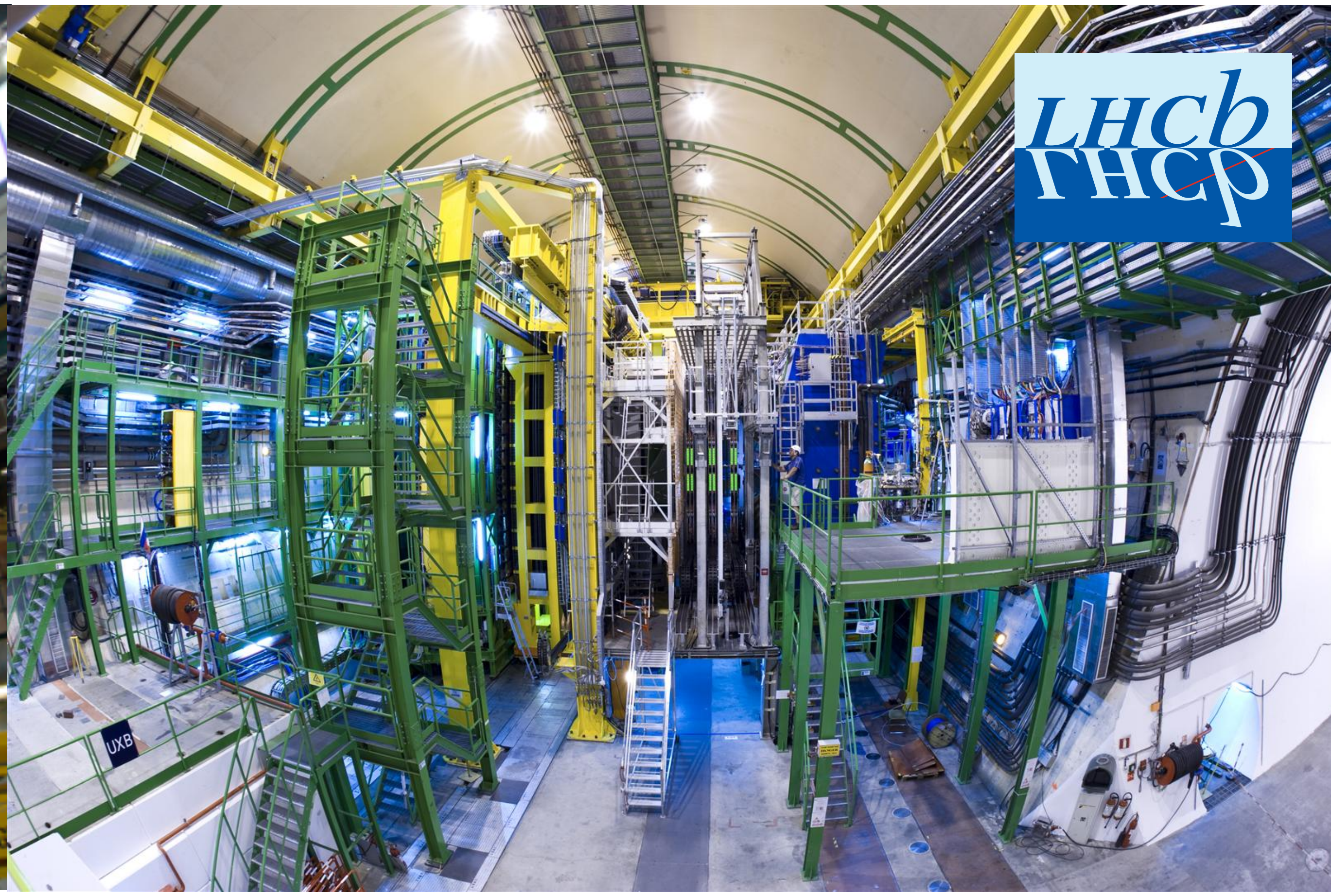
$$\mathcal{B}(B^0 \to \mu^+\mu^-) < 6.3(7.4) \cdot 10^{-10} \text{ @ } 90(95)\% \text{ CL}$$





Combination with CMS





CMS characteristics:

- 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 depending on $|\eta|$

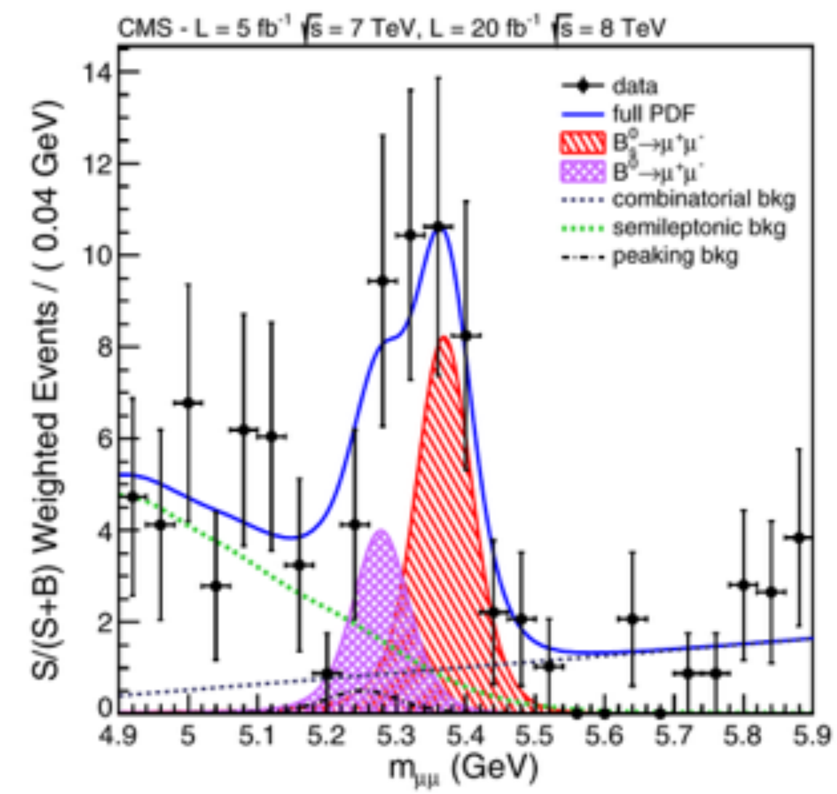
CMS results

[PRL 111 (2013) 101804]

Very similar analysis strategies between CMS and LHCb.

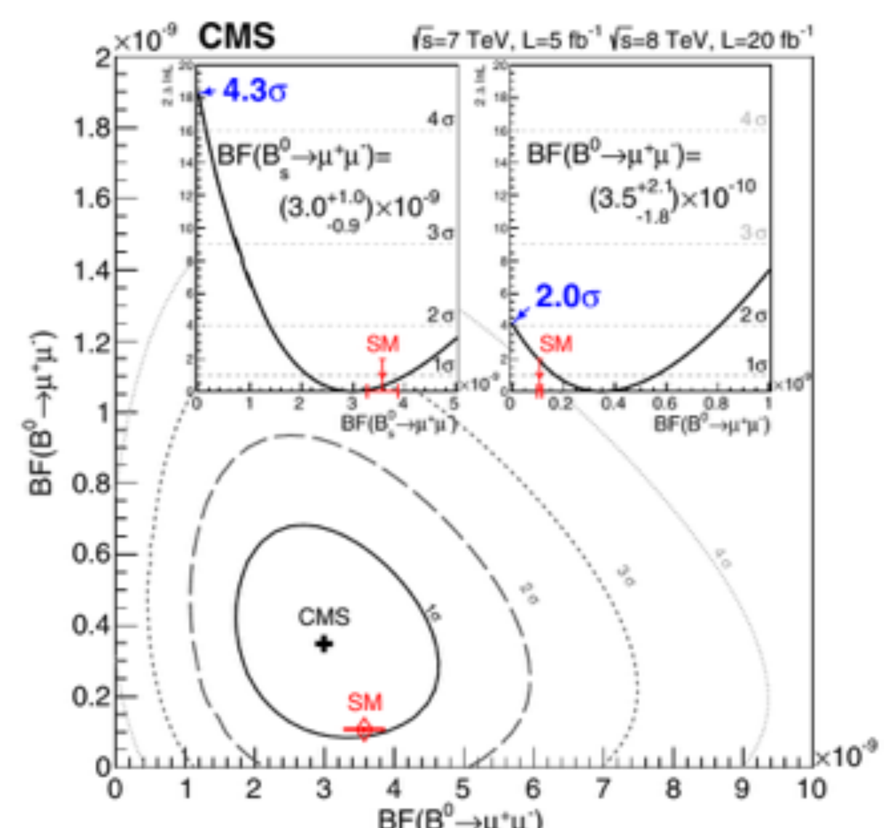
Analysis based on full Run I data 5+20 fb^{-1} for CMS at 7 and 8 TeV

- Soft preselection
- Search divided in "barrel" (both muons $|\eta| < 2.5$) and "endcap" (at least one muon with $|\eta| > 2.5$)
- Multivariate classifier (BDT) for signal/background discrimination
- Search in the invariant mass distribution and 12 BDT categories



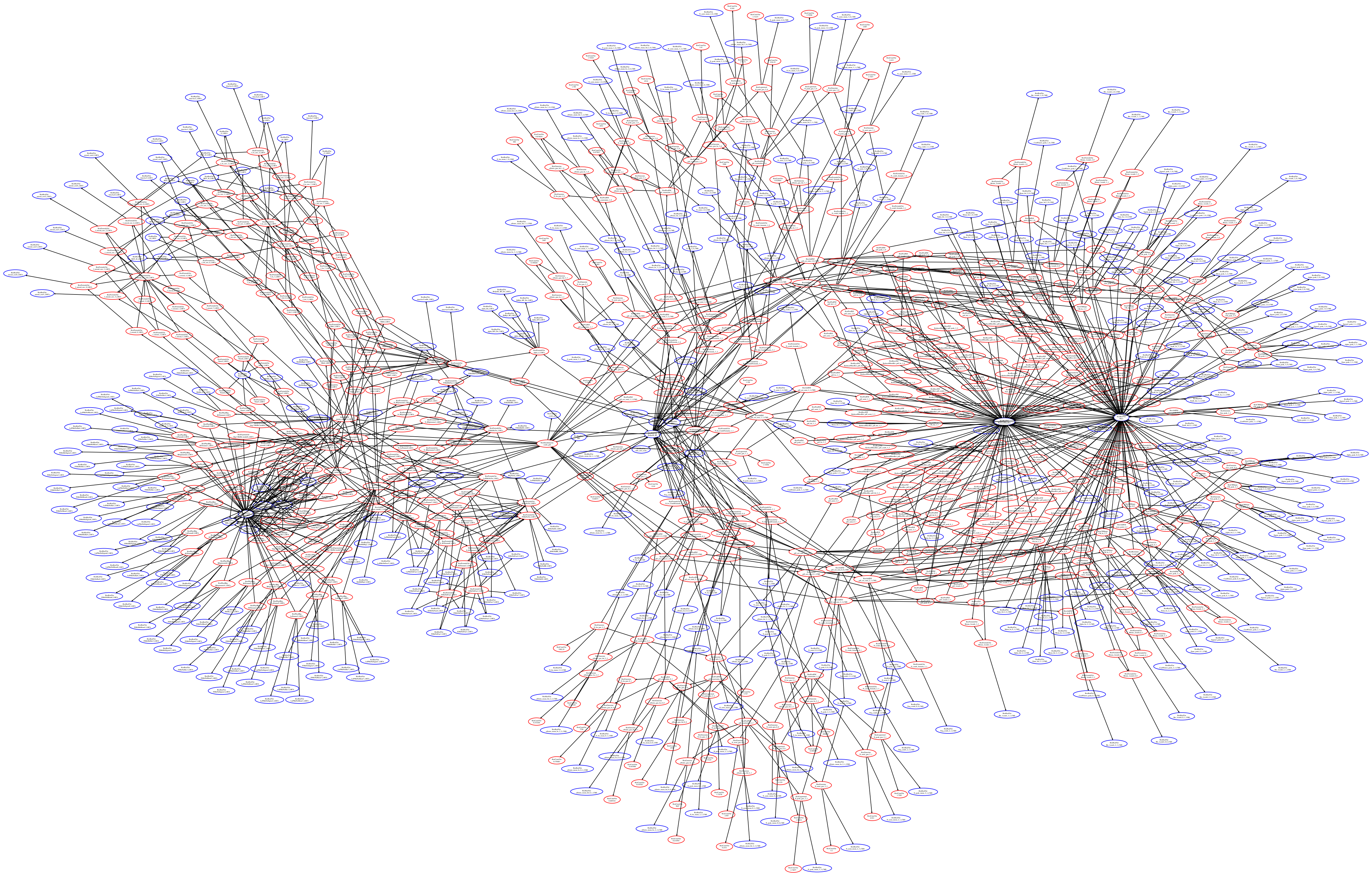
$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = 3.0^{+1.0}_{-0.9} \times 10^{-9} \ \ (4.3\sigma)$$

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



~1fb⁻¹ at LHCb is equivalent to ~10fb⁻¹ at CMS

Fit framework

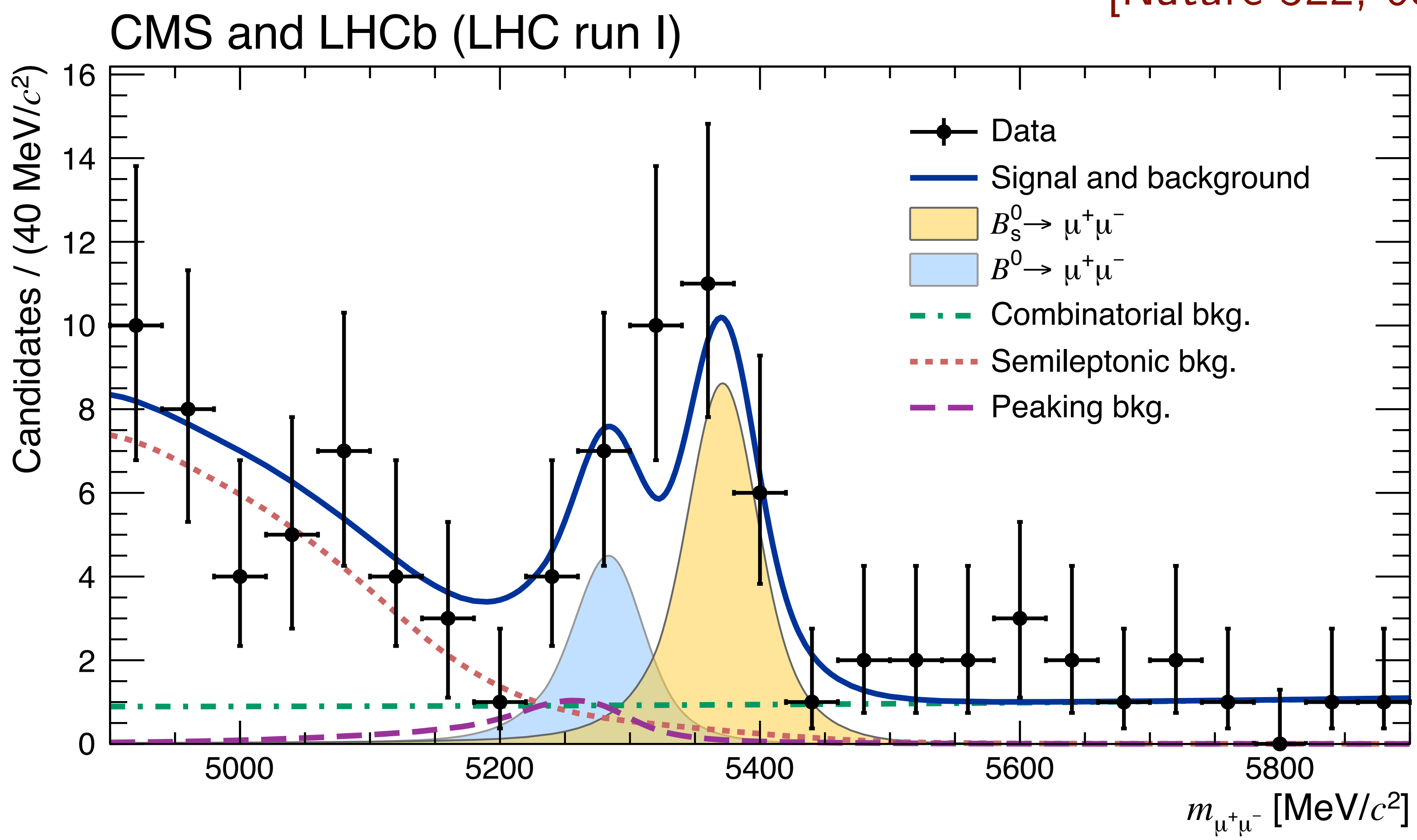


Fit framework

- 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
- Full Feldman-Cousins procedure for the interval on the B($B^0 \to \mu^+\mu^-$) branching fraction
- 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

Combination results

[Nature 522, 68-72 (04 June 2015)]



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

6.2 σ significance observed

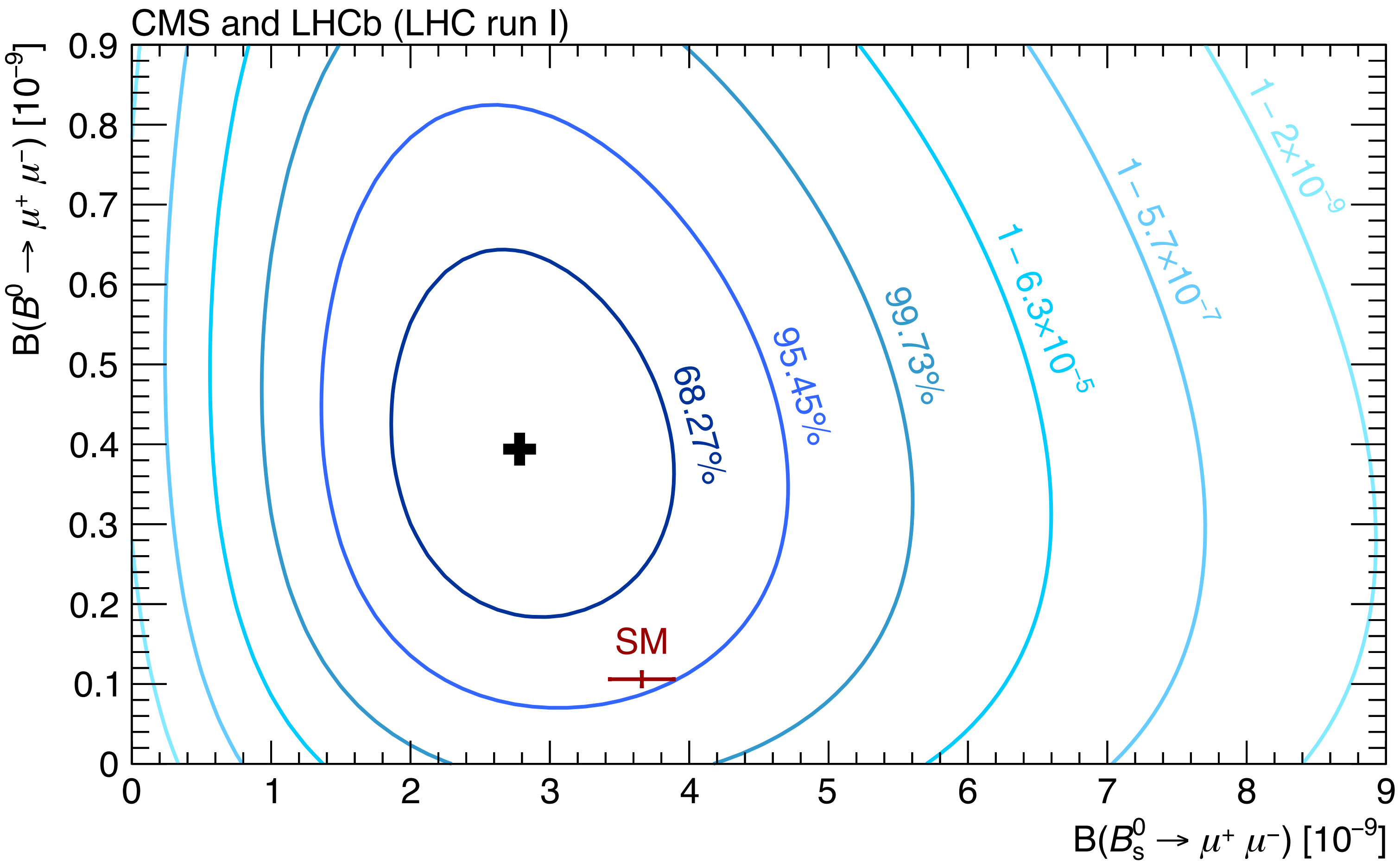
$$\mathcal{B}(B^0 \to \mu^+\mu^-) = 3.9^{+1.6}_{-1.4} \times 10^{-10}$$

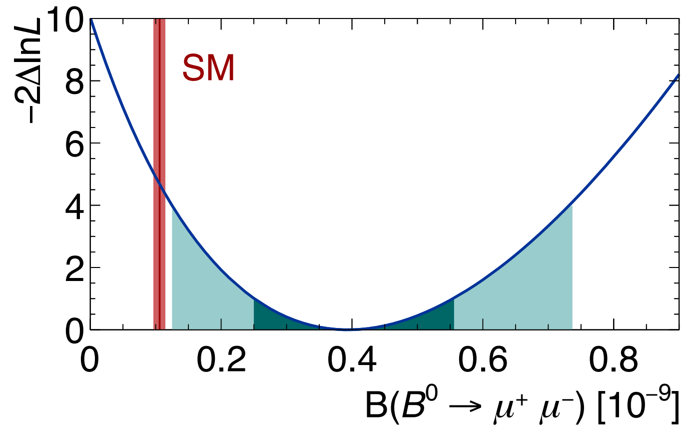
3.0 σ observed (from FC method)

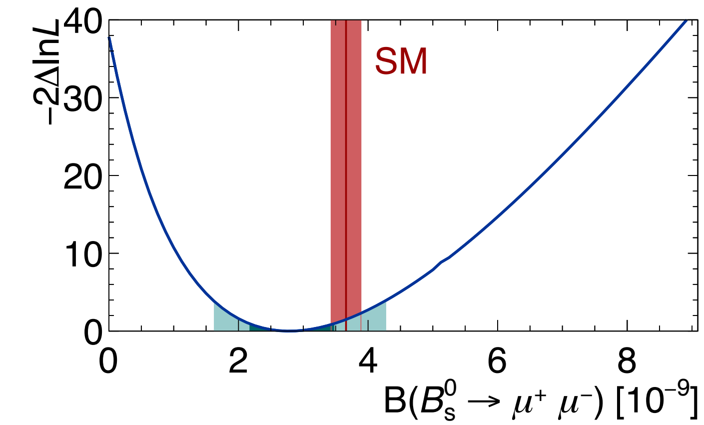
- Which represent the first observation of the $B^0_s \to \mu^+\mu^-$ decay and a first evidence for the $B^0 \to \mu^+\mu^-$ decay.
- Compatibility with the SM predictions: 2.2 σ for B0 and 1.2 σ for Bs

Combination results

[Nature 522, 68-72 (04 June 2015)]



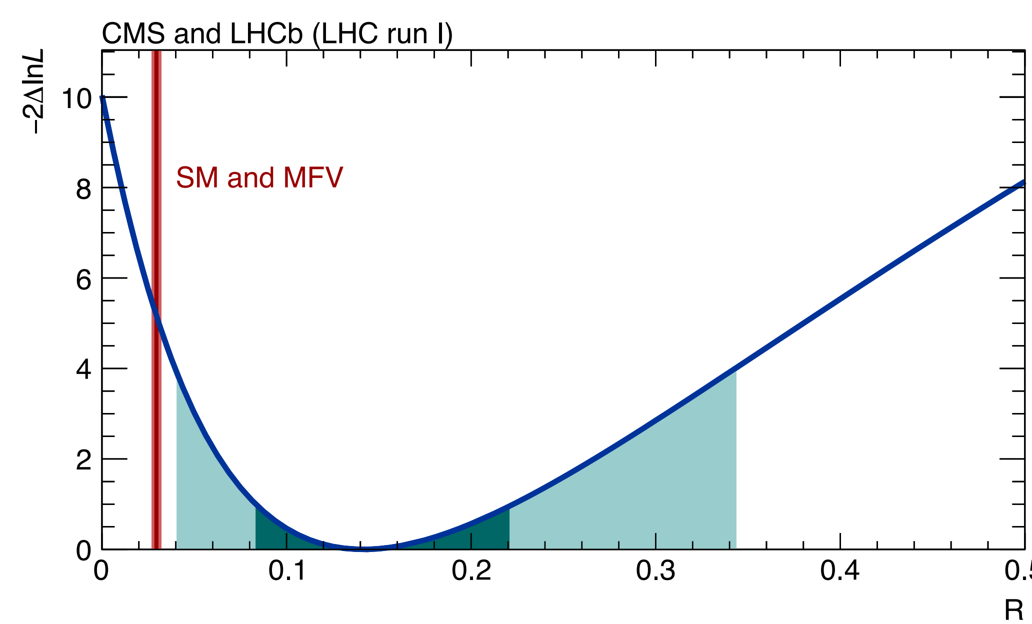




- Ratio of the 2 modes useful to constraint MFV models with same flavour structure as SM
- Fit result:

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

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



implications for NP models

The general effective hamiltonian for $B^0_s \rightarrow \mu^+\mu^-$ can be written as:

$$\mathcal{H}_{eff} = -\frac{G_F}{\sqrt{2}\pi} V_{ts}^* V_{tb} \alpha \left[C_{10} O_{10} + C_S O_S + C_P O_P + C'_{10} O'_{10} + C'_S O'_S + C'_P O'_P \right]$$

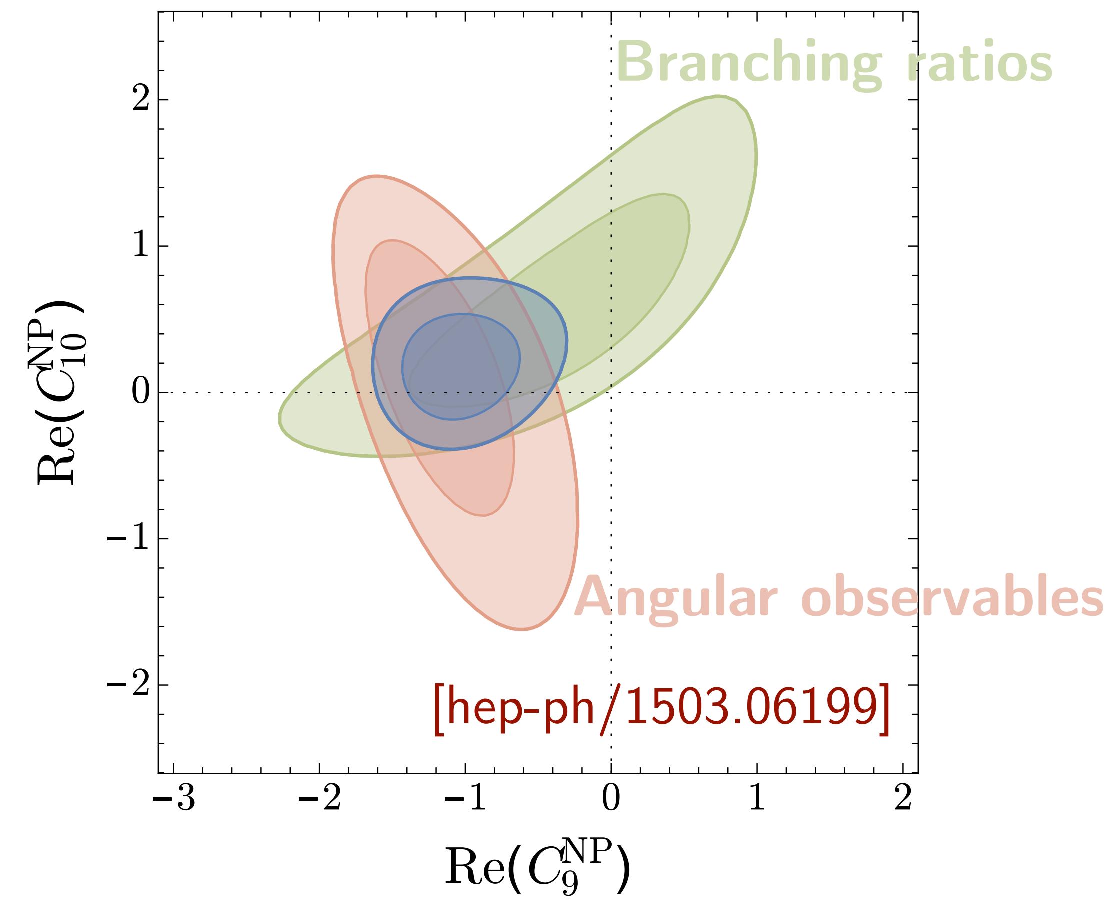
where:

- $O_i^{(\prime)}$ are the for fermion operator
- $C_i^{(\prime)}$ are the Wilson's coefficient

In the SM only C_{10} appears. Any deviation of these coefficient or additional operator can be constrained with the BF

implications for NP models

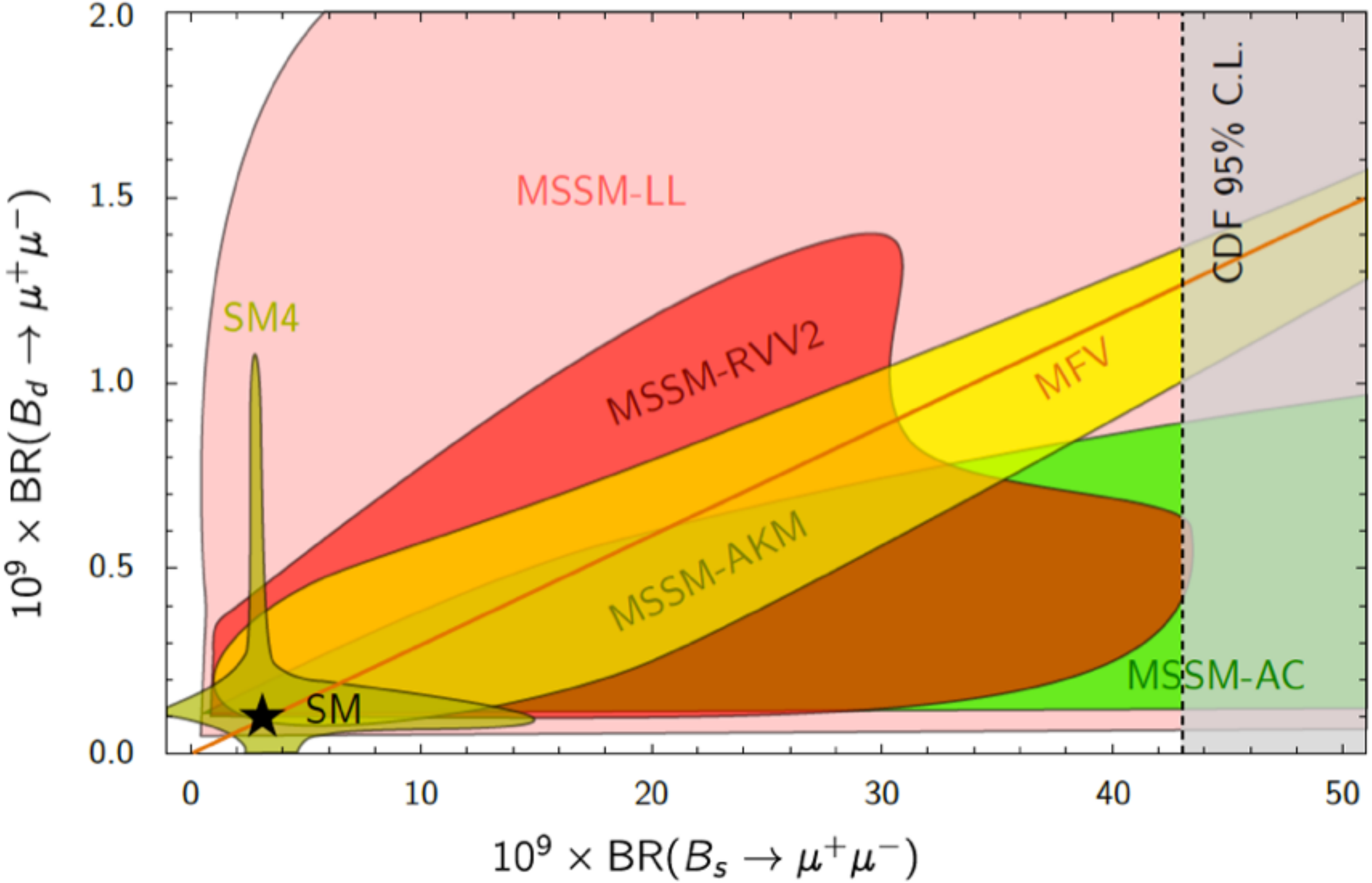
- A recent global fit to $b \to sll$ observables shows a large discrepancy with SM (3.7σ level) mainly due to measurements of lepton flavour universality and angular distributions in $B \to K^{(*)}ll$ decays at LHCb



- No evidence of large NP contribution to C_{10}

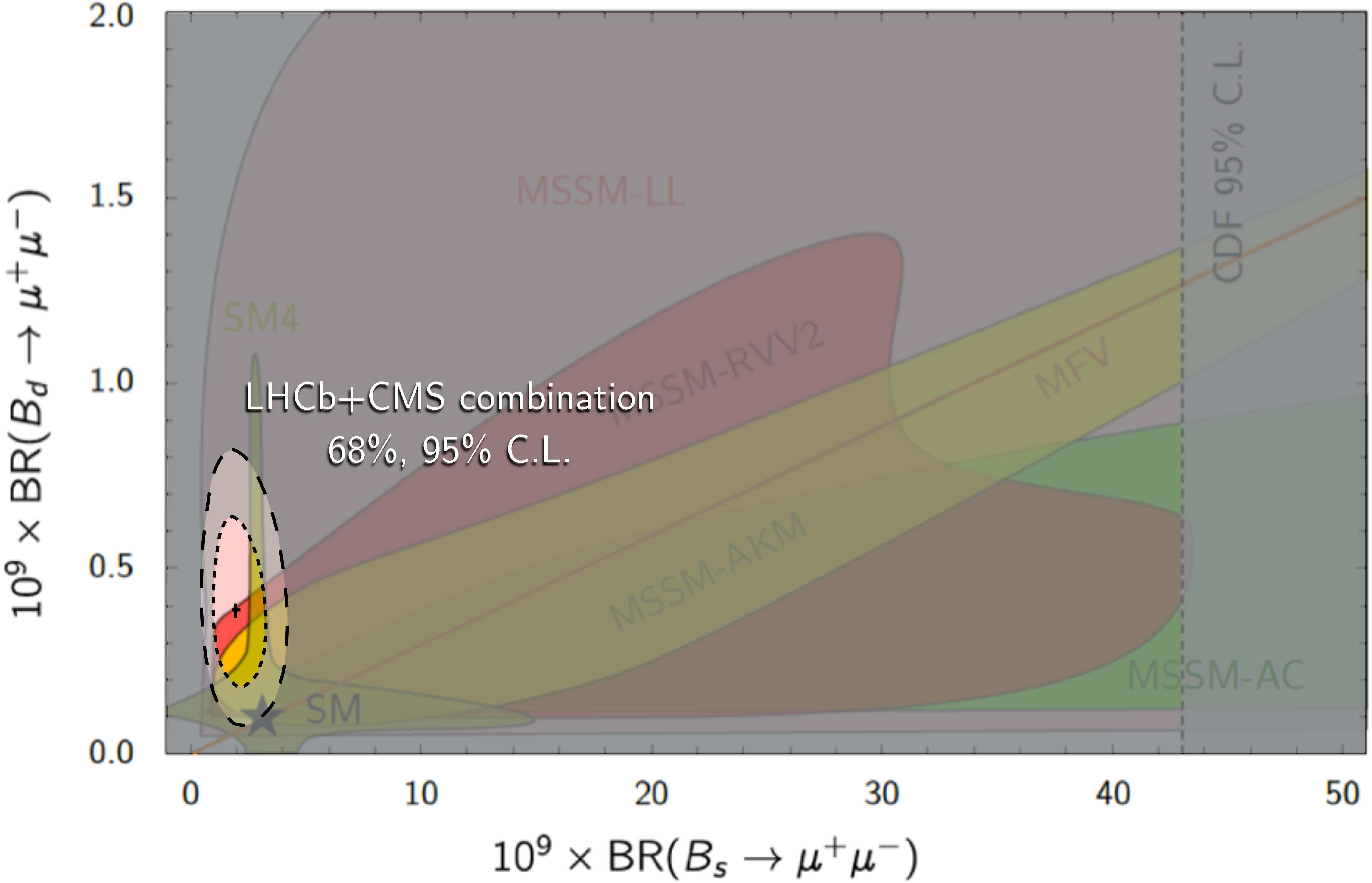
implications for NP models

[D. Straub http://arxiv.org/abs/1205.6094]



implications for NP models

[D. Straub http://arxiv.org/abs/1205.6094]



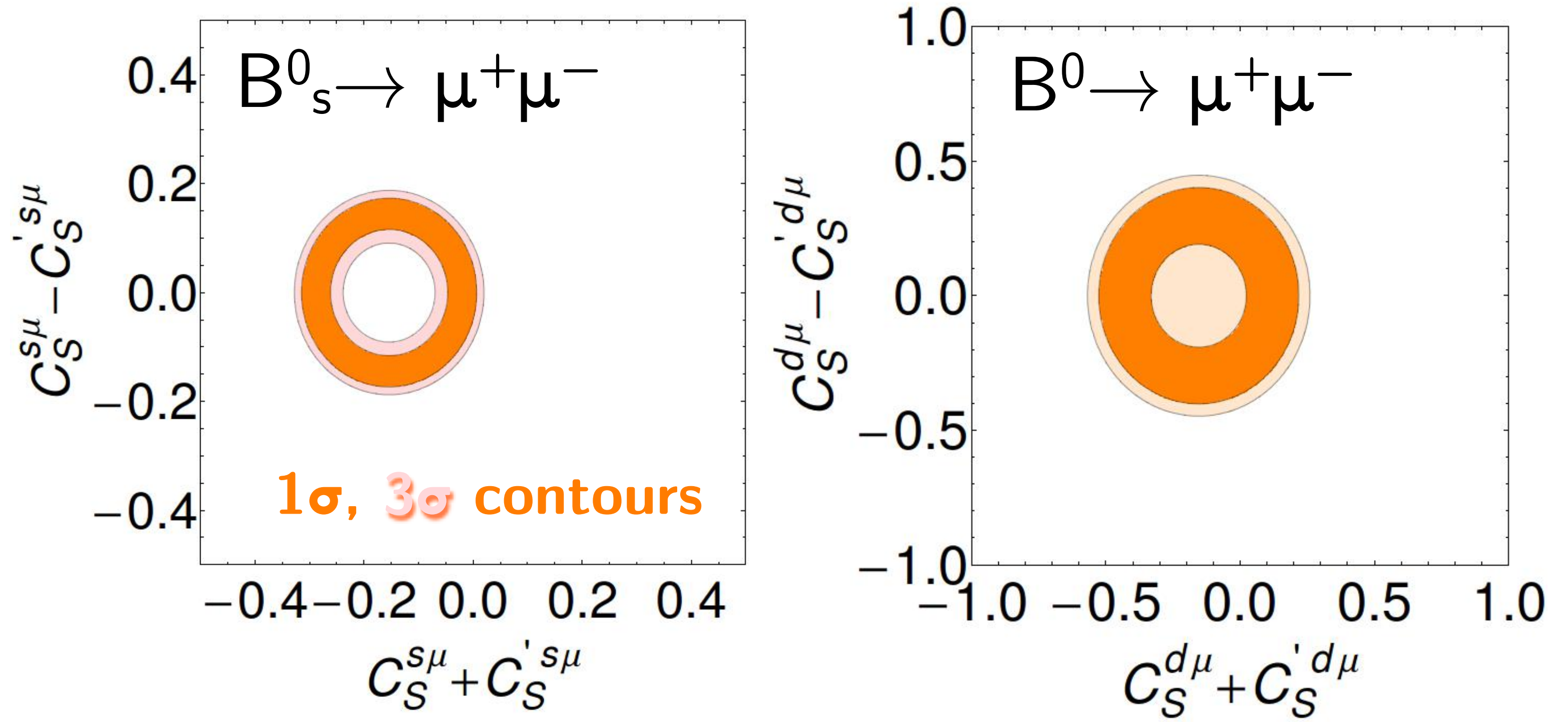
implications for NP models

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) \propto m_\mu^2 \left(\left| (C_{10}^{\rm SM} + \cancel{C_{10}^{\rm NP}} - \cancel{C'_{10}}) - \frac{M_{B_s}}{2m_\mu}(C_{SP} + C'_{SP}) \right|^2 + \left| \frac{M_{B_s}}{2m_\mu}(C_{SP} - C'_{SP}) \right|^2 \right)$$

- If we consider C_{10}' negligible the BF is proportional to the sum and the difference between C_S and C_P.

[Phys. Rev. Lett. 113, 241802 (2014)]

- **radius** proportional to the branching fraction while the **width of the rings** the experimental accuracy
- **Breaking the degeneracy will require other observables!**



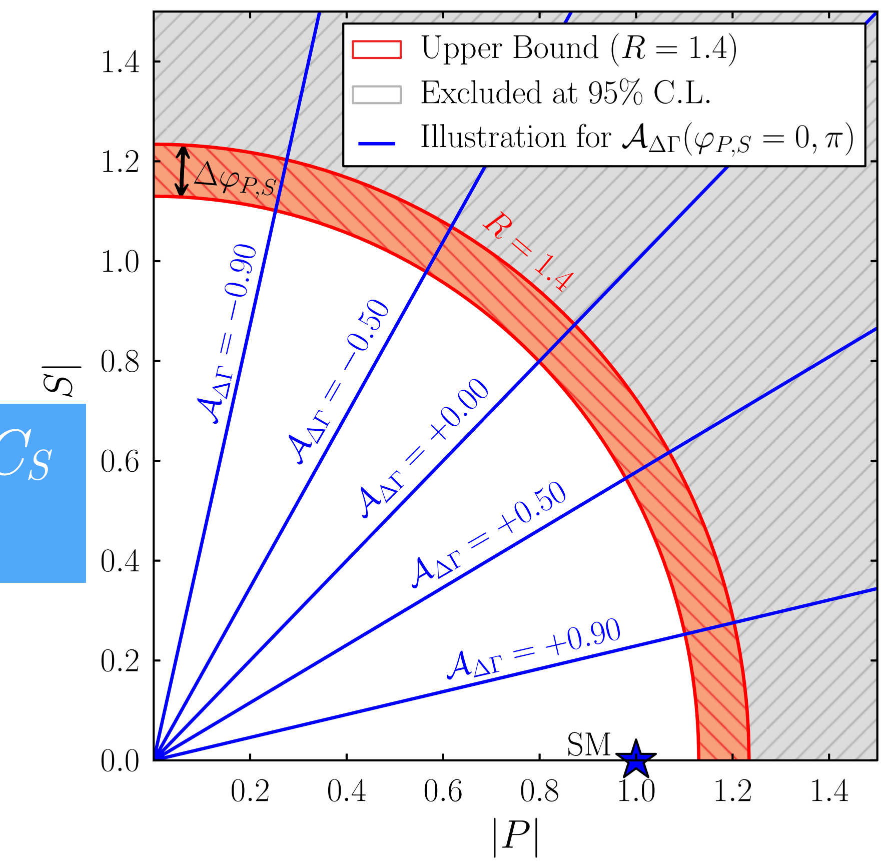
New observable

[K. De Bruyn et al. Phys.Rev.Lett. 109 (2012) 041801]

- Ratio with SM expectation not sufficient

$$\overline{R}_{ql} = \frac{\overline{\mathcal{B}}_{ql}}{(\overline{\mathcal{B}}_{ql})_{\rm SM}} = \frac{1+\mathcal{A}^{ll}_{\Delta\Gamma}\, y_q}{1+y_q}\left(|S|^2+|P|^2\right)$$

proportional to C_S and C_P

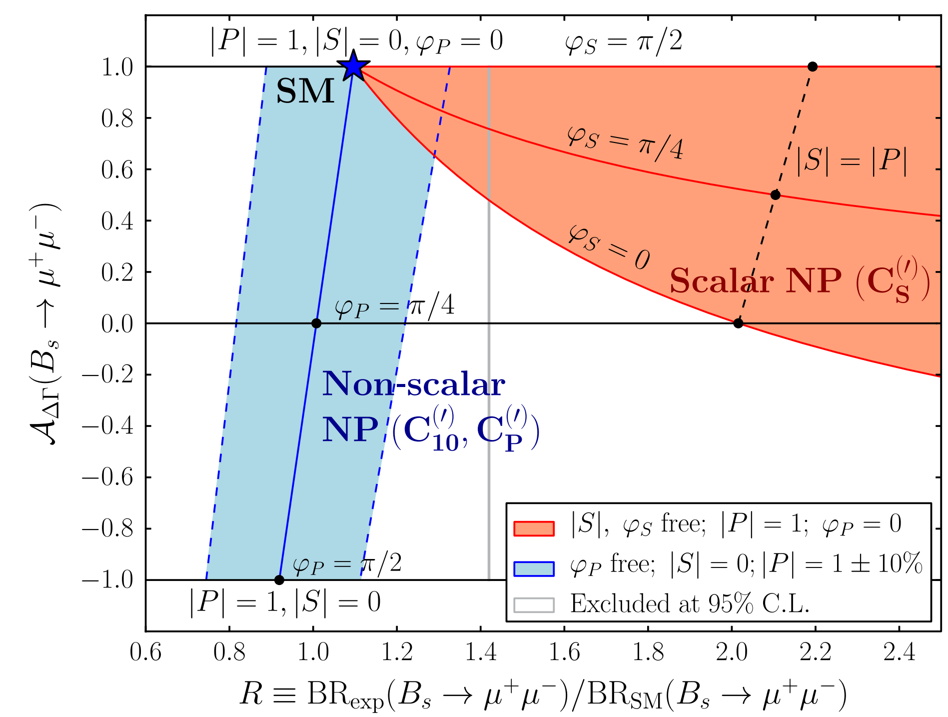


- $A_{\Delta\Gamma}$ proportional to the effective lifetime

$$\tau_{\mu\mu} = \frac{\tau_{B_s}}{(1-y_s^2)}\frac{1+2\mathcal{A}_{\Delta\Gamma}y_s+y_s^2}{1+\mathcal{A}_{\Delta\Gamma}y_s}$$

- Effective lifetime offers theoretically clean probe of NP complementary to branching fraction.

- LHCb could reach a 5% uncertainty on the effective lifetime with 46fb^{-1}.

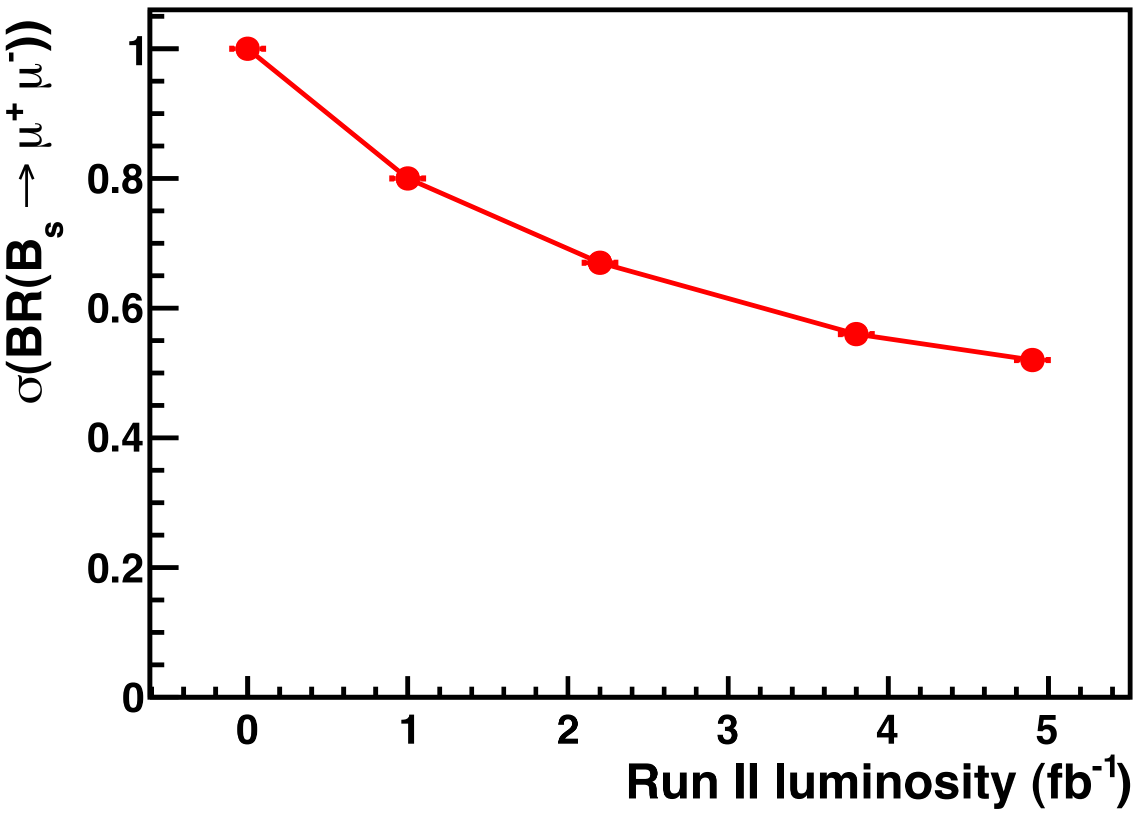


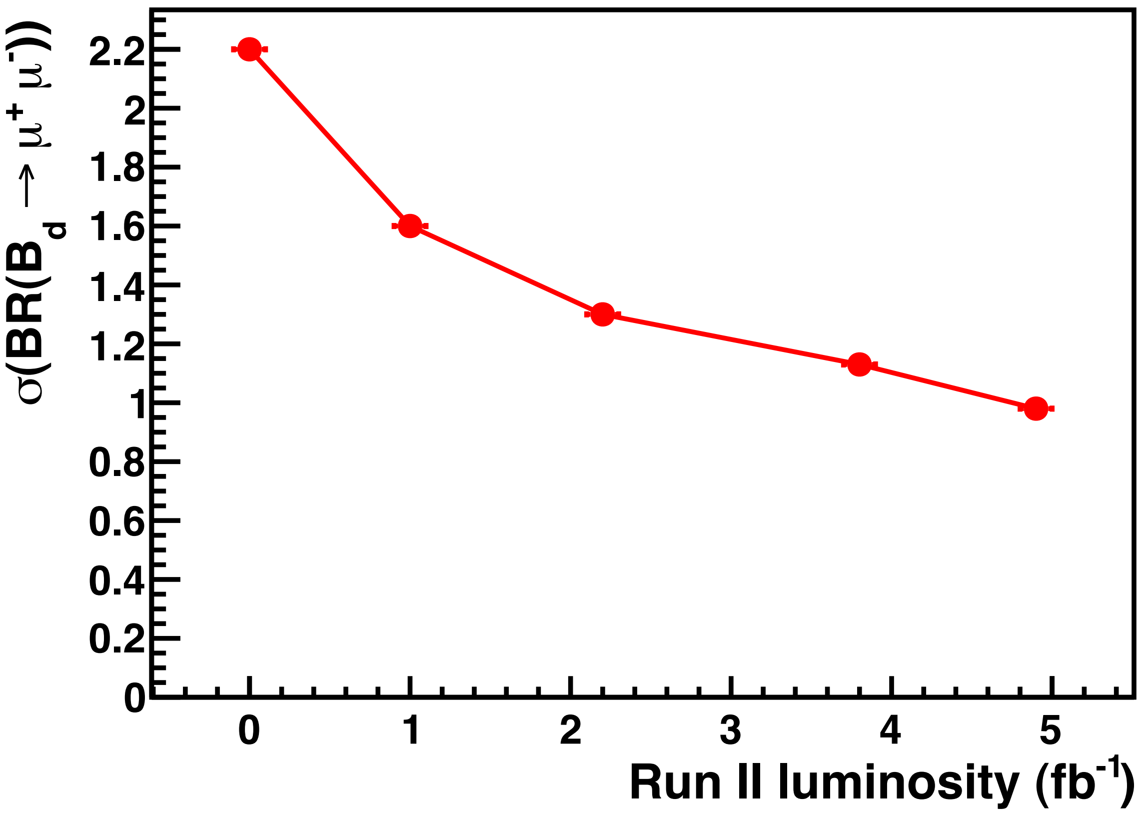
Conclusions

- Rare decays, and $B^0_s \to \mu^+\mu^-$ in particular, are fundamental probes of NP at high energy scales.
- After a long search $B^0_s \to \mu^+\mu^-$ was observed with rate compatible with the SM level; excess at 3 σ level observed for the $B^0 \to \mu^+\mu^-$ hypothesis with respect to the background-only hypothesis, compatibility with SM at 2.2 σ level.
- Both decays are excellent probes for the extended Higgs sector and can be used to set stringent limits to BSM scenarios.
- These studies are still core program of LHCb.
- New observables reachable at high luminosity might constrain alternative theories even more.
- Also one of the best probe among the flavour observable for top quark mass [CERN-PH-TH-2015-191]

Backup

Projections





- Projection based on toys based on current published analysis
- Still room for improvement

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:

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

- 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