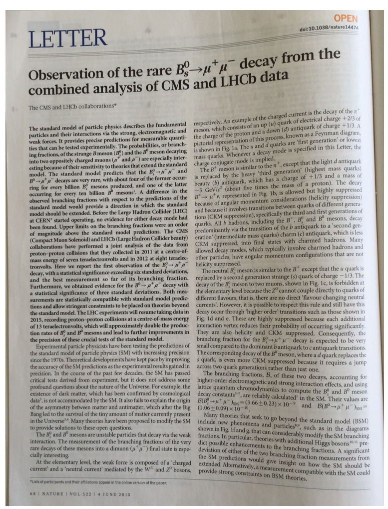
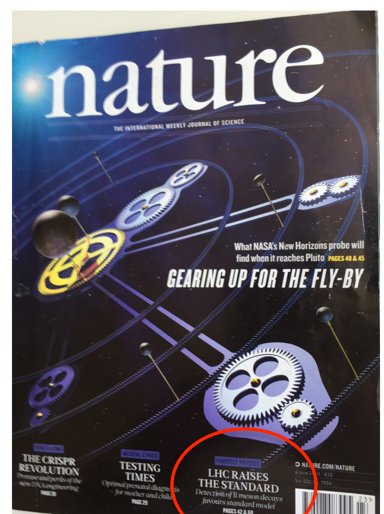
CMS and LHCb: Combining to Corner $B_s \rightarrow \mu^+\mu^-$

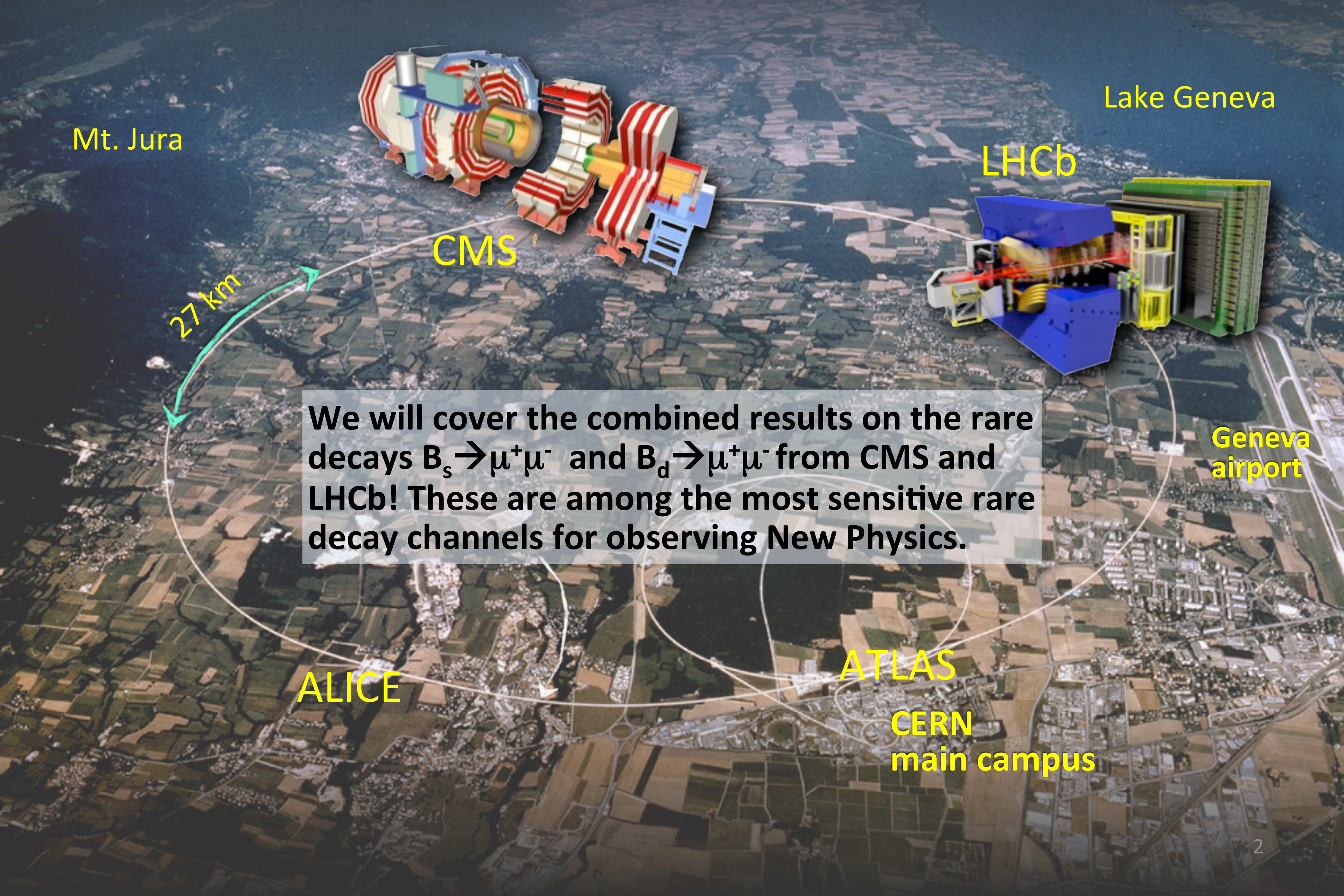
Joel Butler, Fermilab, July 31, 2015

Nature Science 522, 4 June 2015 (doi:10.1038/nature14474)!!!


OPEN
doi:10.1038/nature14474
LETTER
Observation of the rare $B^0_s \rightarrow \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data
The CMS and LHCb collaborations*



THE LARGE HADRON COLLIDER

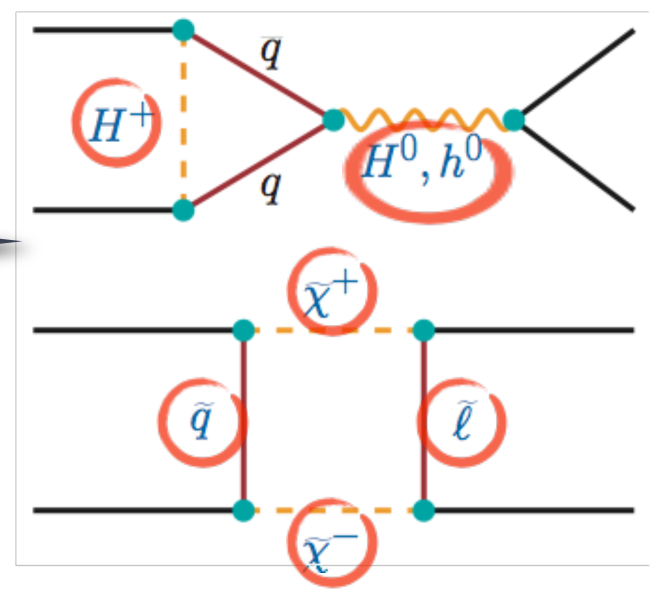


We will cover the combined results on the rare decays $B_s \rightarrow \mu^+\mu^-$ and $B_d \rightarrow \mu^+\mu^-$ from CMS and LHCb! These are among the most sensitive rare decay channels for observing New Physics.

TWO WAYS OF LOOKING FOR NEW PHYSICS

- **Directly through** production and study of new high mass particles at **HIGH ENERGY**
- **Indirectly through HIGH PRECISION measurements:**
  - **"Virtual" manifestation** of new particles that participate in the loop processes and can be discovered by seeing deviations from the Standard Model prediction.

If these particles cannot be observed in the direct searches, this is another way one can still look for them!



Both direct and indirect searches are necessary and complement each other!

CMS and LHCb exemplify these two approaches to discovering New Physics

- CMS is designed to search directly for new particles and interactions with mass scales typical of EW symmetry breaking, i.e. 100 GeV/c^2 – a few TeV/c^2. It has a very broad program, by no means optimized for B physics!
  - However, since many of the hypothesized new particles would decay into b-quarks (b-jets) and/or muons the detection of B hadron decays and muons was a key element in its design
- LHCb is optimized to study CP asymmetries and rare decays of particles containing b(and c-)-quarks to detect deviations from precise SM predictions that would indicate the presence of New Physics
- These lead the two experiments to
  - instrument complementary angular regions with respect to the colliding beams,
  - operate at different proton-proton collision rates, and
  - select b-quark events with different efficiencies
- CMS operates at higher collision rates than LHCb but with lower efficiency for such "low mass" states, resulting in similar sensitivity to this decay.
  - **In this case, CMS is an "ENERGY frontier" experiment doing CROSSOVER "INTENSITY frontier" physics.**

CMS and LHCb results in 2013 triggered a decision to combine them

- LHCb Collaboration, Aaij, R. *et al.*, First **evidence** for the decay $B^0_s \rightarrow \mu^+\mu^-$. *Phys. Rev. Lett.* **110**, 021801 (2013) (Jan. 7, 2013)
  - The probability that the background could produce such an excess or larger is 5.3×10^{-4} corresponding to a signal significance of **3.5 standard deviations**.
- CMS Collaboration, Chatrchyan S. *et al.*, **Measurement** of the $B^0_s \rightarrow \mu^+\mu^-$ branching fraction and search for $B^0_s \rightarrow \mu^+\mu^-$ with the CMS experiment. *Phys. Rev. Lett.* **111**, 101804 (2013). (Sept. 5, 2013)
  - An excess of $Bs \rightarrow \mu+\mu-$ events with respect to background of **4.3 standard deviations**.
- LHCb Collaboration, Aaij, R. *et al.*, **Measurement** of the $B^0_s \rightarrow \mu^+\mu^{--}$ branching fraction and search for $B^0_s \rightarrow \mu^+\mu^-$ at the LHCb experiment . *Phys. Rev. Lett.* **111**, 101805 (2013) (Sept. 5, 2013)
  - An excess of $Bs \rightarrow \mu+\mu-$ with respect to the background of **4.0 standard deviations**.

The two ~4 σ results from two "complementary and independent" experiments should constitute an "observation" but we decided combine the results at the FITTER LEVEL to resolve all small differences in corrections and take all correlations from theory and experiment into account, to get the best branching fraction possible, expecting it would easily exceed the 5 σ line in the sand for "observation".

Outline

- The characteristics that make a decay a promising one for observing New Physics
- $B_{(s,d)} \rightarrow \mu^+\mu^-$ meet the requirements
- A high level view of the two experiments and their analyses
- The combination effort and results
- The future of these measurements
- Conclusion

While we focus on the OBSERVATION of $B_s \rightarrow \mu^+\mu^-$ at >5 σ, we also have a combined limits for $B_d \rightarrow \mu^+\mu^-$ and discuss the prospects of eventually observing it. **An important aspect of this effort is that it tells us the work that has to be done to take the next steps in investigation of these states.**

Characteristics of a Good Decay for Searching for New Physics

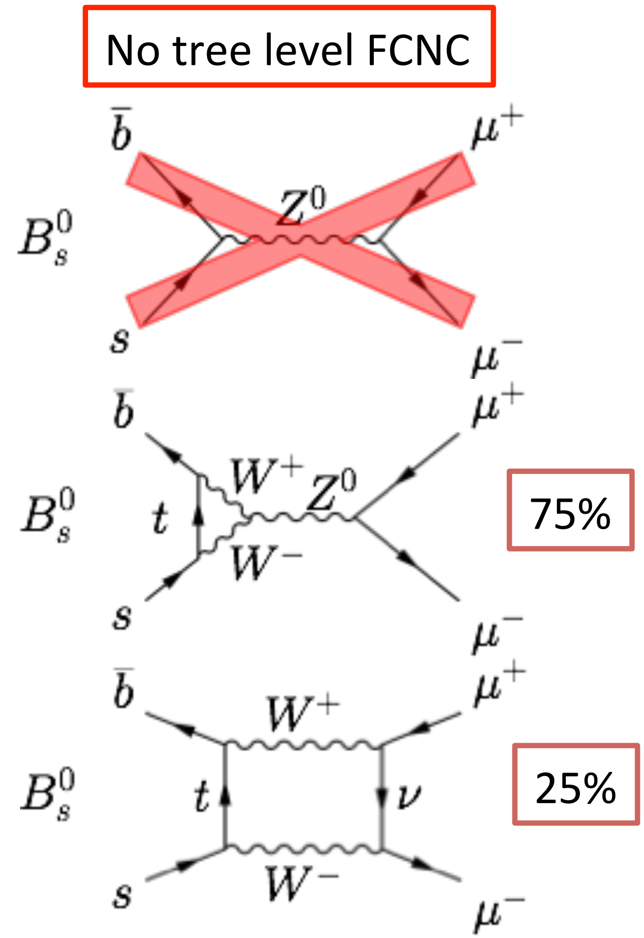
- Must be highly suppressed in SM to see New Physics (NP)
  - Figure of Merit = $N_{NP}/\sqrt{(N_{NP} + [N_{SM} + N_{bkgnd}])}$, since N_{SM} is part of the "background" want it to be small
- Must have a reliable theoretical prediction for the branching fraction of the rare decay
- Must be produced at an interesting level
  - "Gap" between existing limit and SM
  - "Desirable": sensitivity down to SM level
- Must be able to trigger and reconstruct the state with high efficiency and low experimental background
- "Desirable, but not required": there should be models of New Physics that predict BR within the unexplored "gap"

Some B_s, B_d meson properties

- The B_s meson is a $\bar{b}s$ bound state; the B_d meson is a $\bar{b}d$ bound state
- The Mass of the B_s is 5366.7 MeV/c^2 and the B_d is 5279.55 MeV/c^2
    - $M_{B_s} - M_{B_d} = \sim 87$ MeV/c^2
- B_s^0 is a flavor eigenstate, not a mass eigenstate, and oscillates rapidly between B_s and $\overline{B}_s$
- The interactions that produce mixing also can produce a difference in lifetimes between the two mass eigenstates B_{sH} and B_{sL} of about 10%
- The B_d^0 has weaker mixing, oscillates more slowly and there is almost no difference in the lifetimes of its two mass eigenstates
- Both B_d and B_s have mean lifetimes of 1.5ps, corresponding to $c\tau$ of ~450μ m
- The distance from the production (primary) vertex to the B decay (secondary) vertex can be measured and used to eliminate most prompt backgrounds

$B_{s,d} \to \mu^+\mu^-$ in THE STANDARD MODEL

- In the Standard Model, $B_{s,d} \to \mu^+\mu^-$ decays are **highly suppressed**:
  - Flavor Changing Neutral Current (FCNC) processes in SM are forbidden at tree level but can proceed through Z-pengiun, and box diagrams (suppressed by $[m_W/m_t]^2$).
  - CKM suppressed: $|V_{tq}|^2$
  - Helicity suppressed: $[m_\mu/m_B]^2$
- Resulting **tiny branching fractions**, but rather *robust theory predictions are available*



Two Consequences

- $BR(B_d \rightarrow \mu^+\mu^-) << BR(B_s \rightarrow \mu^+\mu^-)$, therefore difficult
  - CKM suppressed, since $V_{td} << V_{ts}$ (factor of 20)
  - Slightly compensated (factor of ~2.5) at LHC since B_d's are produced with higher cross section than B_s's
- $BR(B_{(s,d)} \rightarrow e^+e^-)$ inaccessible
  - Further helicity suppressed by $(m_e/m_\mu)^2$

Standard Model Prediction

Decay constant

Wilson coefficient

$$\overline{\mathcal{B}}_{q\ell} = \frac{|N|^2 M_{B_q}^3 f_{B_q}^2}{8\pi\Gamma_H^q} \beta_{q\ell} r_{q\ell}^2 |C_A(\mu_b)|^2 + \mathcal{O}(\alpha_{em})$$

$$N = V_{tb}^{\star} V_{tq} G_F^2 M_W^2 / \pi^2$$

$$r_{q\ell} = 2m_\ell / M_{B_q} \qquad \beta_{q\ell} = \sqrt{1 - r_{q\ell}^2}$$

Ref: Bobeth et al, PRL 112, 101801 (2014)
All input numbers and formulae are in a backup slide.

Results

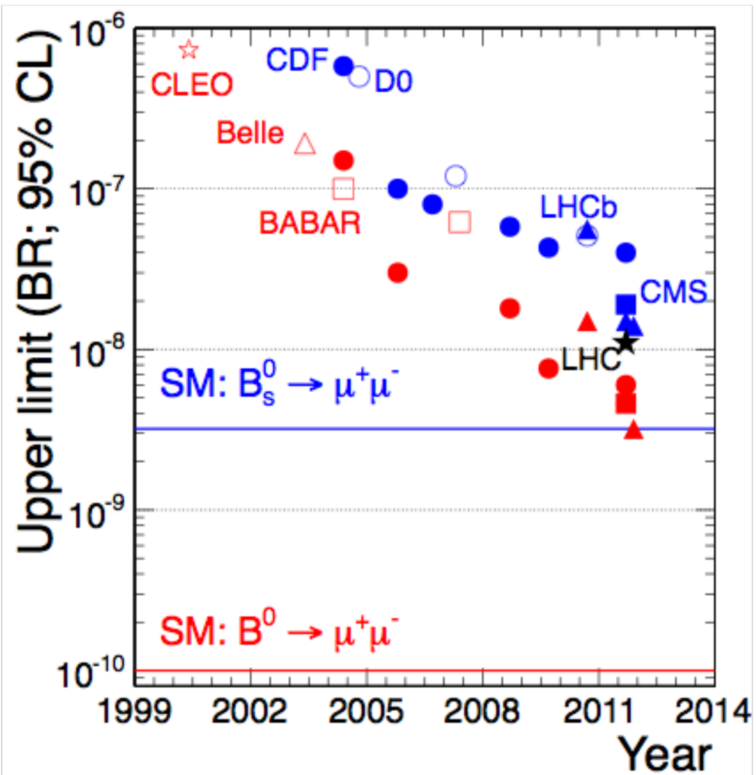
$B(B_s \to \mu^+\mu^-) = (3.65 \pm 0.23) \times 10^{-9}$
$B(B_d \to \mu^+\mu^-) = (1.06 \pm 0.09) \times 10^{-10}$

TABLE II. Relative uncertainties from various sources in $\overline{\mathcal{B}}_{s\ell}$ and $\overline{\mathcal{B}}_{d\ell}$. In the last column they are added in quadrature.

	f_{B_q}	CKM	τ_H^q	M_t	α_s	Other parameters	Nonparametric	$\sum$
$\overline{\mathcal{B}}_{s\ell}$	4.0%	4.3%	1.3%	1.6%	0.1%	< 0.1%	1.5%	6.4%
$\overline{\mathcal{B}}_{d\ell}$	4.5%	6.9%	0.5%	1.6%	0.1%	< 0.1%	1.5%	8.5%

Vcb

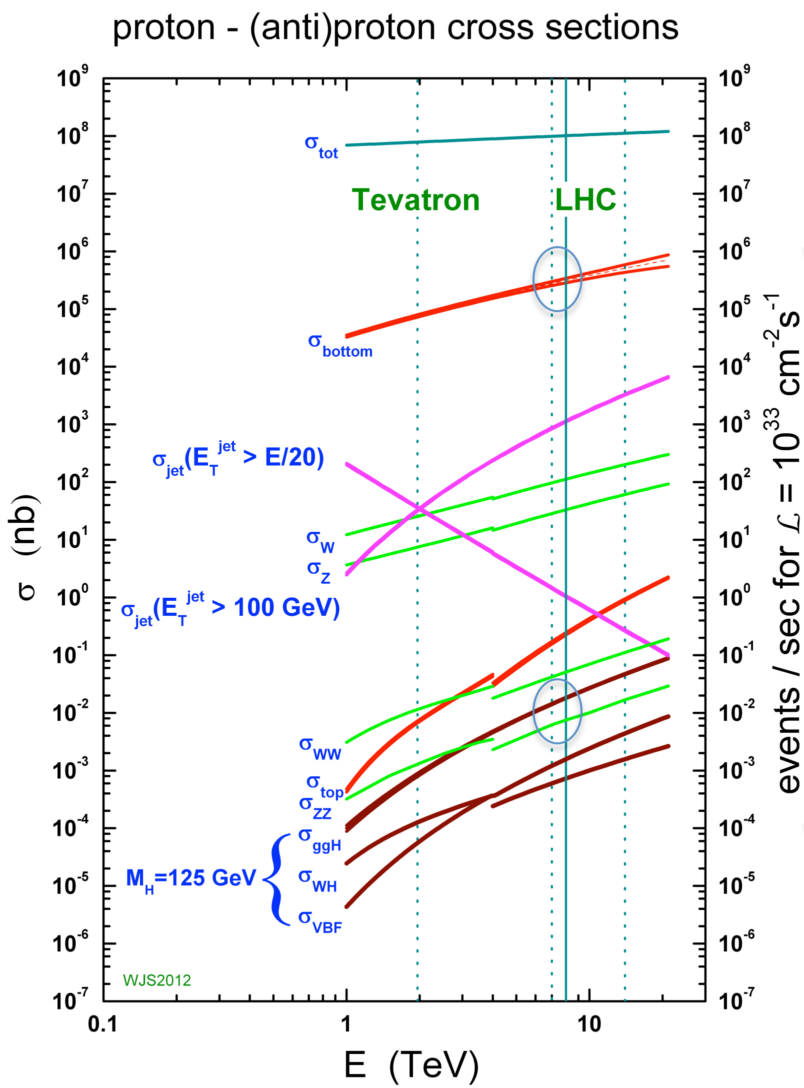
STATUS IN EARLY 2012



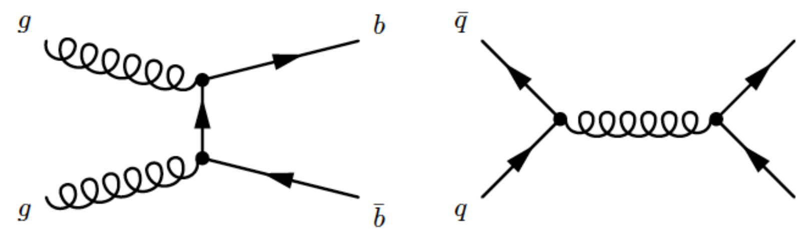
Just one order of magnitude shy from the SM predictions!

- Many experiments have looked for this decay and pushed the limits as much as possible since ~1985.
- The **B-factory** experiment BELLE at the KEKB electron-position collider cannot do better than the LHC.
  - This is due to the fact that the production cross sections are much larger (>3 orders) in high energy **proton collisions**.
- **At the startup to the LHC, there was still room for the B_s decay to be significantly larger than the SM prediction**

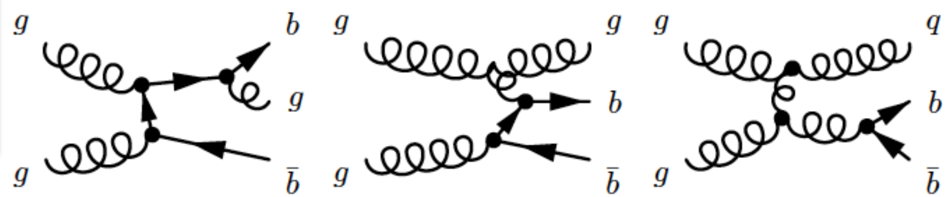
B Production at the LHC is large



LO – Pair creation



NLO –
pair creation, Flavor Excitation, Gluon splitting



B hadron production rates at LHC

Note that both charge states are included in these results.

Expt	Observable (p_T in GeV)	$\sigma^{\rm exp}$	$\sigma^{\rm FONLL}$	Comments
1: LHCb [56]	$\sigma(H_b, 2 \le \eta \le 6)$	$75.3 \pm 11.4\ \mu$b	$70.8\ ^{+33.3}_{-24.4}\ \mu$b	average $b + \bar{b}$
2: LHCb [57]	$\sigma(B^{\pm}, p_T < 40, 2 < y < 4.5)$	$41.4 \pm 3.4\ \mu$b	$40.1\ ^{+19.0}_{-14.5}\ \mu$b	$f(b \to B^-) = 0.403$
3: CMS [55]	$\sigma(B^0,\ p_T^B > 5,\ \lvert y^B\rvert < 2.2)$	$33.2 \pm 4.3\ \mu$b	$25.5\ ^{+10.5}_{-7.1}\ \mu$b	$f(b \to B^0) = 0.403$
4: CMS [54]	$\sigma(B^+,\ p_T^B > 5,\ \lvert y^B\rvert < 2.4)$	$28.1 \pm 4.4\ \mu$b	$27.2\ ^{+11.2}_{-7.5}\ \mu$b	$f(b \to B^-) = 0.403$
5: CMS [58]	$\sigma(B_s^0,\ 8 < p_T^B < 50,\ \lvert y^B\rvert < 2.4)$ $\times{\rm BR}(B_s^0 \to J/\psi\, \phi)$	6.9 ± 0.8 nb	$4.5\ ^{+2.3}_{-1.9}$ nb (includes BR uncertainty)	$f(b \to B_s^0) = 0.11$ ${\rm BR}(B_s^0 \to J/\psi\, \phi) = (1.4 \pm 0.5) \times 10^{-3}$
6: LHCb [64]	$\sigma(H_b \to J/\psi, p_T^{\psi} < 14, 2 < y_{\psi} < 4.5)$	$1.14 \pm 0.16\ \mu$b	$1.16\ ^{+0.55}_{-0.42}\ \mu$b	${\rm BR}(b \to J/\psi) = 0.0116$
7: ALICE [66]	$\sigma(H_b \to J/\psi, p_T^{\psi} > 1.3, \lvert y_{\psi}\rvert < 0.9)$	$1.26 \pm 0.16\ \mu$b	$1.33\ ^{+0.59}_{-0.48}\ \mu$b	${\rm BR}(b \to J/\psi) = 0.0116$
8: CMS [73]	$\sigma(H_b \to \mu,\ p_T^{\mu} > 6,\ \lvert y^{\mu}\rvert < 2.1)$	$1.32 \pm 0.34\ \mu$b	$0.855\ ^{+0.28}_{-0.19}\ \mu$b	${\rm BR}(b \to \ell) = 0.0108$ ${\rm BR}(b \to c \to \ell) = 0.096$

Table 1: Summary of various cross-section measurements, compared against the FONLL predictions. The numbers labeling each measurement refer to the entries in Fig. 10.

JHEP 10 (2012) 137

For B_u, 30μb x 10^{34} cm^{-2}s^{-1} x 5x 10^6 s/year = 1.5 x 10^{12} $B_u$year B_s is maybe a factor of two lower, so it should be possible to detect a BR of ~few X 10^{-9} or even 10^{-10}!

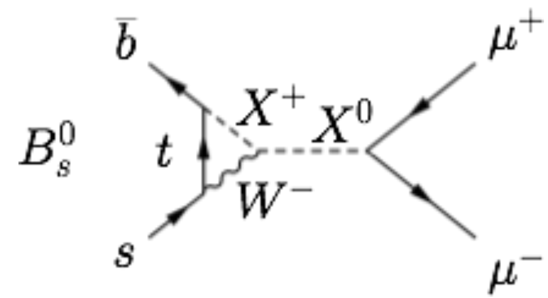
Basic Sensitivity to New Physics

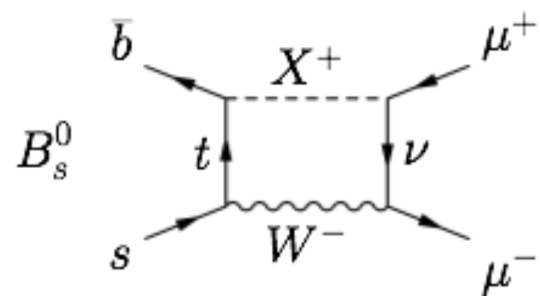
- Quite general and comes from the unique SM aspects that cause these decays to be suppressed in the first place.
- These constraints can be broken if the NP has
  - Different helicity structure (scalar particles, particles with different coupling)
  - Different flavor structure

Builders of models of new physics have long had to face the constraint of the strong suppression in nature of FCNC decays, which grows stronger as experimental limits on these decays are tightened.

$B_{s,d} \rightarrow \mu^+\mu^-$: THE POTENTIAL FOR NEW PHYSICS

- Loop diagram + Suppressed SM + Theoretically clean =
  An excellent place to look for new physics.
- Sensitive to extended Higgs sectors
  ⇒ Constrains NP parameter spaces.
- A few NP examples:
  - **2HDM**: $\mathcal{B} \propto \tan^4\beta$, and $m(H^+)$
    H. E. Logan & U. Nierste, Nucl. Phys. B 586 (2000) 39
  - **CMSSM/mSUGRA**: $\mathcal{B} \propto \tan^6\beta$
    C. S. Huang et al. PRD 63 (2001) 114021
    K. S. Babu & C. Kolda, PRL 84 (2000) 228
    A. Dedes, et al, PRL 87 (2001) 251804
  - **Leptoquarks**
    S. Davidson and S. Descotes-Genon, JHEP 11 (2010) 073

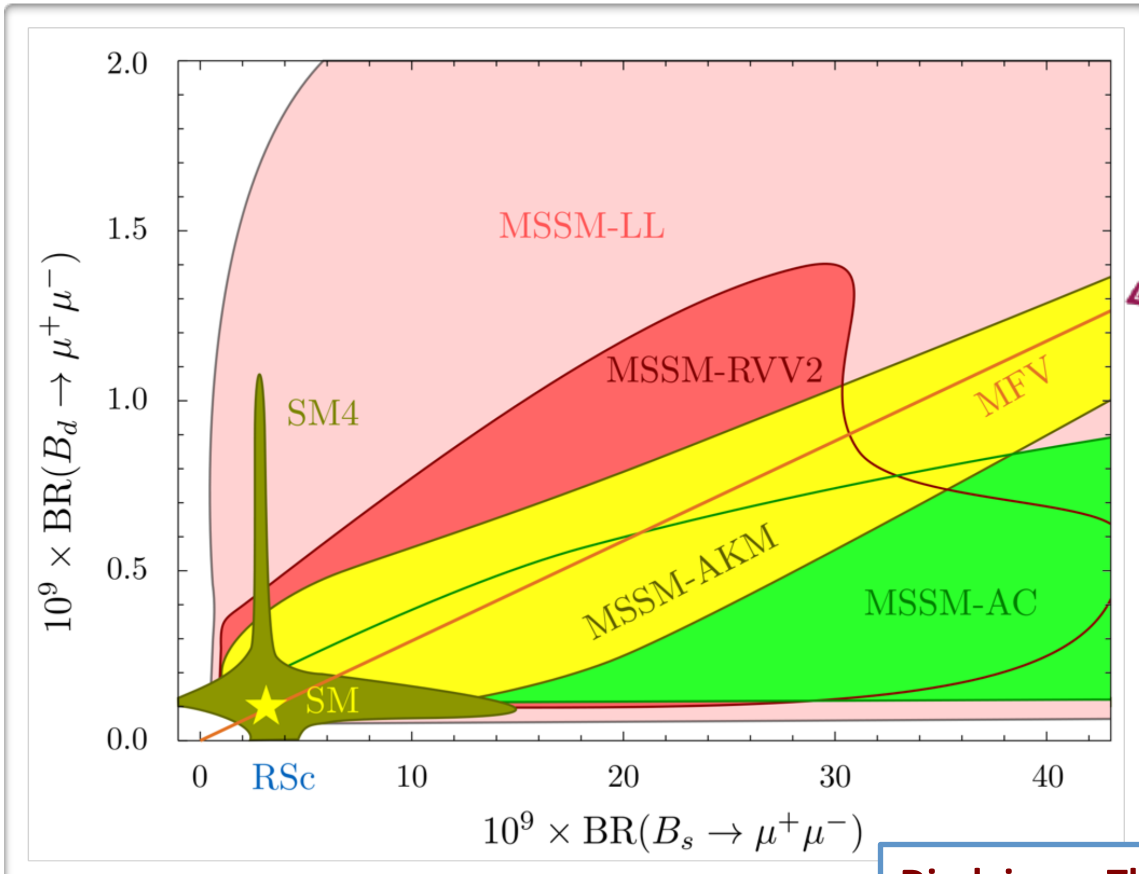




Any difference in branching fraction from SM could provide a smoking gun signal of new physics.

New physics Example

Ref: D. M. Straub, arXiv: 1012.3893



Larger $B_d \to \mu^+\mu^-$

Larger $B_s \to \mu^+\mu^-$

A crucial test of minimal flavor violation models.

When the LHC Started New Physics scenarios that that could boost the BBs→μμ decay rates by **10~20x** were still possible

Disclaimer: This is just an example of A FEW MODELS – 4 particular SUSY/FLAVOR models, SM4 (SM with 4th generation), and MFV hypothesis. Lots of other models

B Physics at Hadron Colliders

- **The Opportunity**
  - **The LHC has very high B cross sections, σ(pp) ~ 500 μb (14GeV)**
  - **It is a "Broadband, High Luminosity B Factory", giving access to B_d, B_u, B_s, all b-baryon, and B_c states.**
  - **The LHC is intrinsically an asymmetric gluon collider**
- **The challenge in general (** for this analysis)**
  - **Charged particle tracking and momentum (mass) resolution ****
  - **Muon identification and reconstruction ****
  - **Triggering (** on muons)**
  - **Primary and secondary vertexing (* to identify muons from B decays)**
  - **Charged hadron identification (* to pick out hadronic two body decays)**
  - **High speed/throughput data acquisition ****
  - **At LHC, multiple interactions in each crossing**
    - **At luminosity of $10^{34}cm^{-2}s^{-1}$, ~25 interactions/crossing with spacing of 25 ns**

While LHCb is optimized for B physics, CMS must optimize for a wide range of Higgs, SUSY, and Beyond-the Standard Model searches for particles up to a few TeV!

CMS-LHCb Complementarity

Goal	New Physics by direct observation of High Mass particles	New Physics by observation of new virtual heavy particles in loops
Detecting region	Central	Forward
Dataset	high luminosity $>10^{34}$/ high pile-up / high trigger thresholds: **-Muons only for B physics triggers/** high efficiency for heavy objects/ low efficiency for light objects/	low luminosity $<10^{33}$/ low pile-up / low trigger thresholds : - **Muons, electrons, and all-hadronic triggers/** high efficiency for light objects/ low efficiency for heavy objects/
Expected performance	**Very similar on states containing muons. LHCb superior on all hadronic states, where CMS can't trigger well and lacks particle id .**	

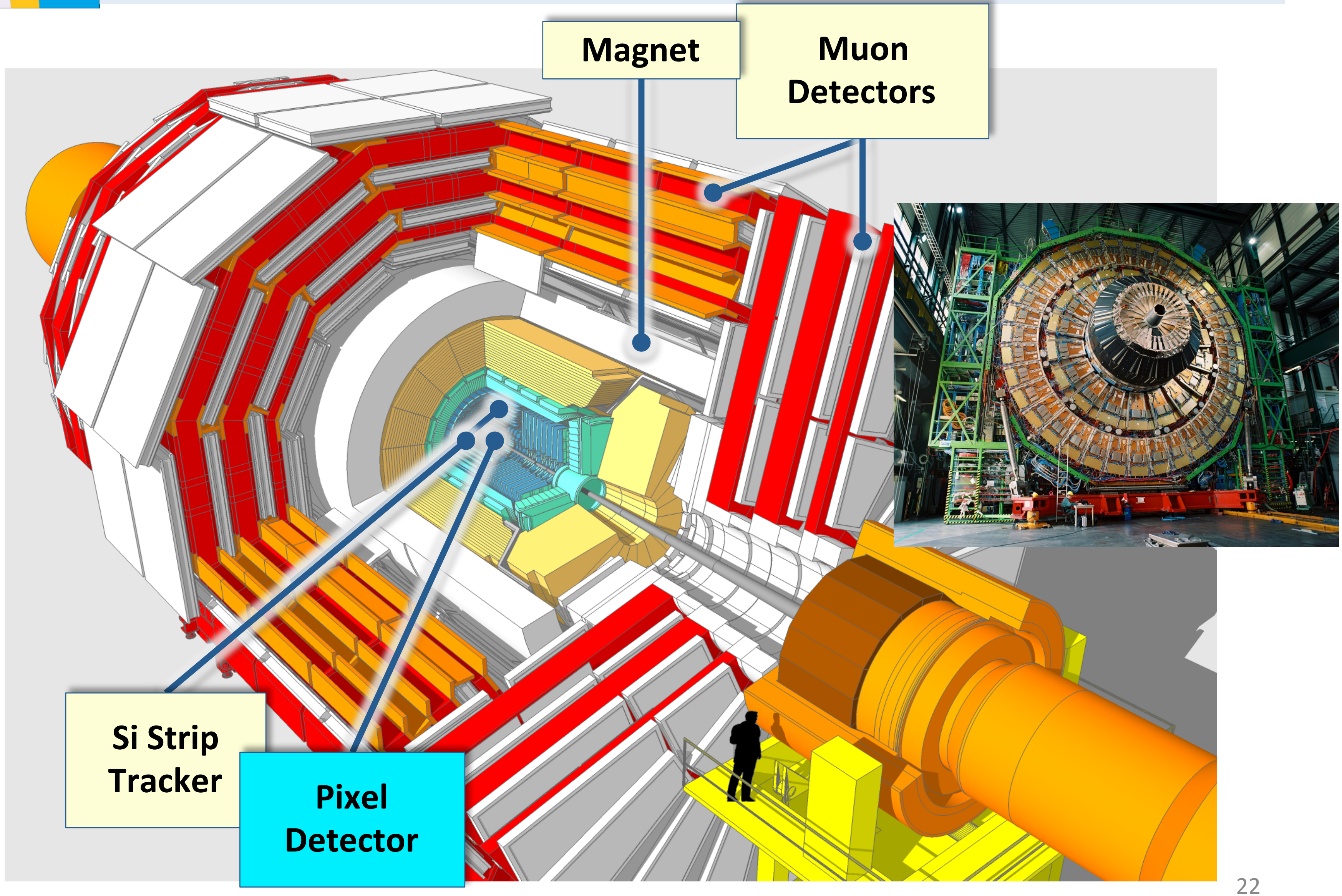
CMS Design Concept

- Very large solenoid – **6m diameter x 13 m long centered on IR (angular coverage**
- Tracking and calorimetry fits inside the solenoid
  - particles measured before they pass through the solenoid coil and cryostat, which would degrade their resolution
- **Very strong field – 3.8 T**
  - **Excellent momentum resolution**
  - Coils up soft charged particles
- **Tracking chambers in the return iron to trigger on, track, and identify muons**
  - This makes the system very compact
  - Weight of CMS is dominated by all the steel and is 14,000 Tonnes
- A lead tungstate crystal calorimeter (~76K crystals) for photon and electron reconstruction
- Hadron calorimeters for jet and missing E_t reconstruction (provides coverage to η~5)

- Charged Particle Tracking is based on all-silicon components
  - **A silicon pixel detector out to radius ~ 12 cm**
  - **A silicon microstrip detector from 25 cm to 110 cm**
    - Small pitch gives CMS excellent charged particle tracking and primary and secondary vertex reconstruction
    - High segmentation results in very low occupancy
    - Silicon detectors are very radiation hard

Muon momentum is measured in the muon system but the best resolution comes from associating a silicon track, which has excellent momentum resolution, with the muon track and doing a full fit. Challenge is to do this with high pileup → fine pitch→ low occupancy outer tracker

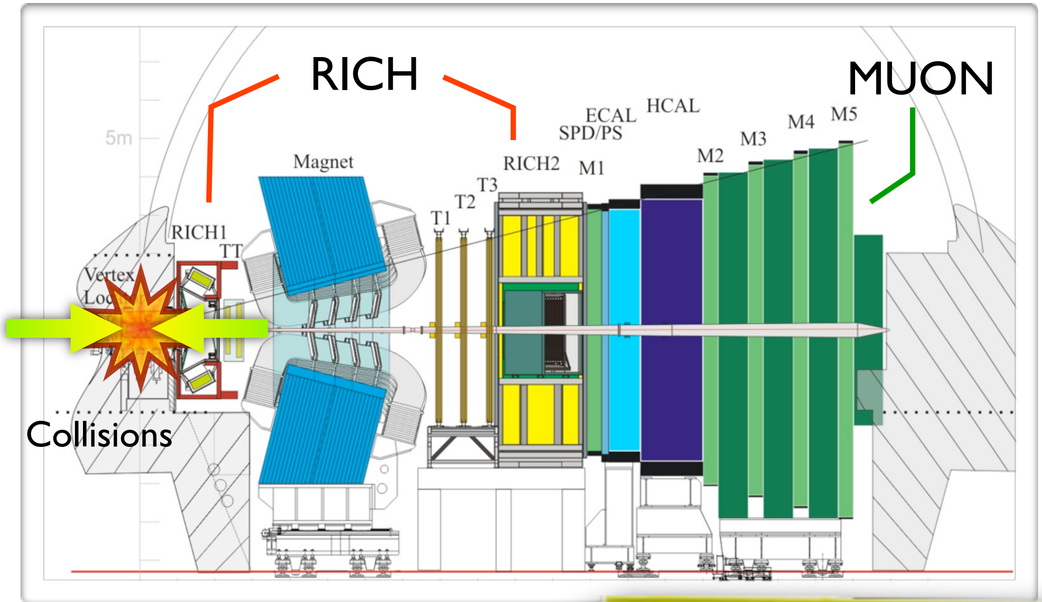
CMS Schematic

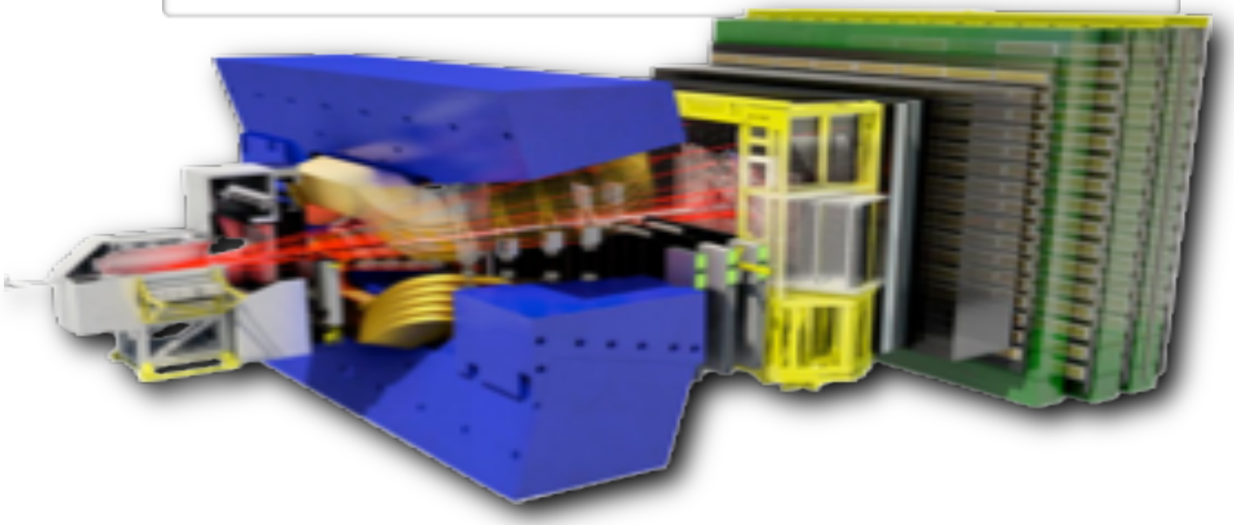


LHCb DESIGN CONCEPT

- **Forward spectrometer** only, optimized for heavy flavor physics.
  - A large physics cross section in a limited transverse area
  - Based on a 4 T-m Dipole magnet
  - Particles are typically few-100 GeV/c,
    - Particle identification, which requires length along the beam
    - Photon and neutral pion reconstruction benefits from high energy
- Excellent track resolution in the target momentum region, σ_p/p : **0.4-0.6%**.
- "Vertex Locator (VELO) " for primary, secondary vertices 8mm from beam
  - Small pitch gives LHCb excellent charged particle tracking and primary and secondary vertex reconstruction
- Muons identified by ability of tracks to penetrate steel at the end of the system and register hits in embedded detectors
- Dimuon triggers are very efficient for the candidates within the acceptance, ~90% for $B_{s,d}\rightarrow\mu\mu$.

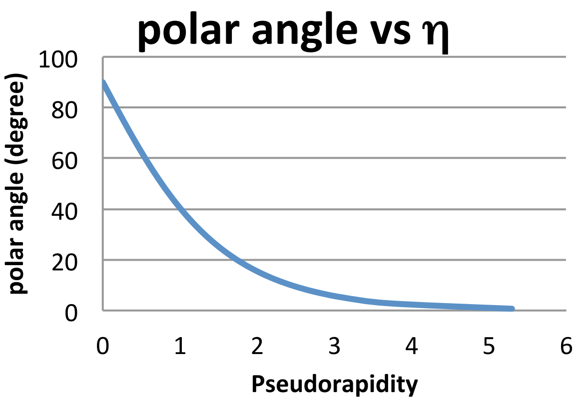
LHCb Schematic

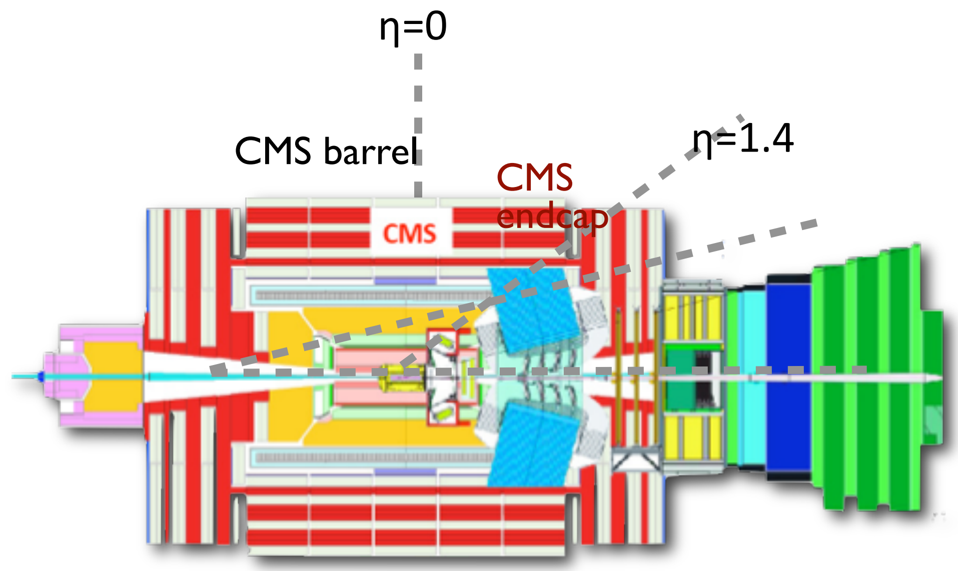


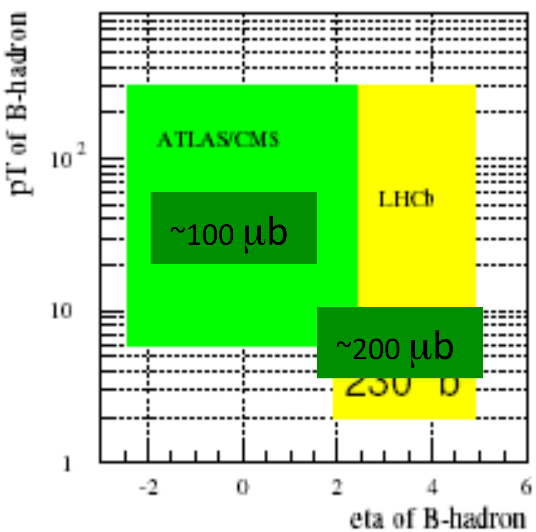


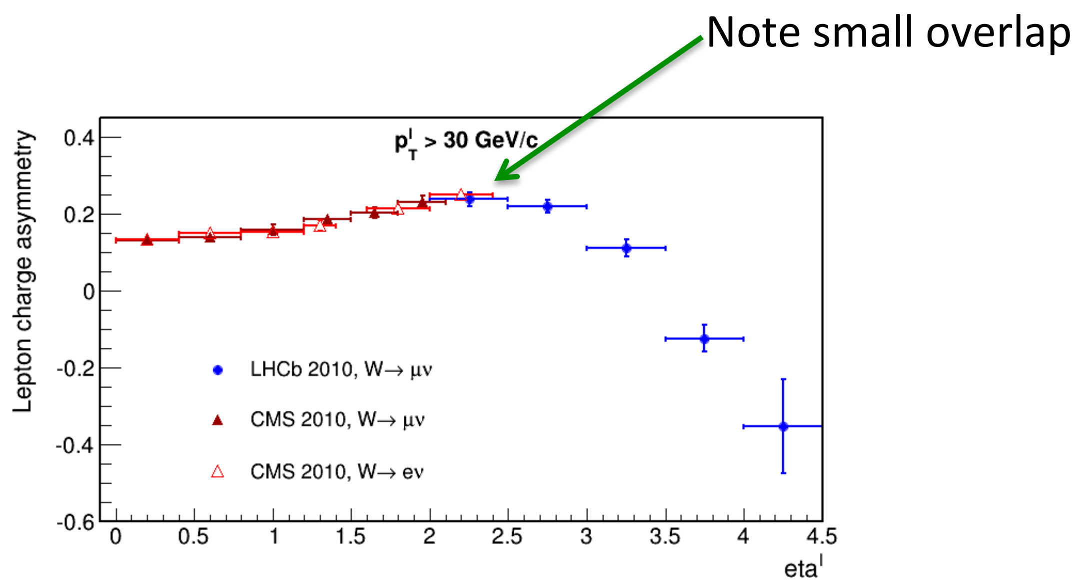
Complementary η coverage

$$\eta = -\ln(\tan(\frac{\Theta}{2}))$$





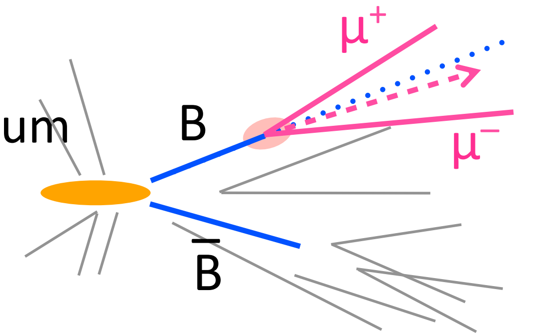




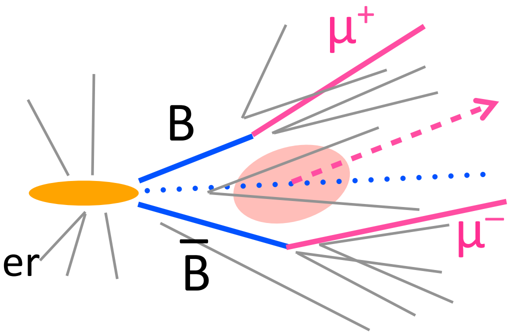
muon charge asymmetry from W

ANALYSIS ASPECTS: EVENT SIGNATURE

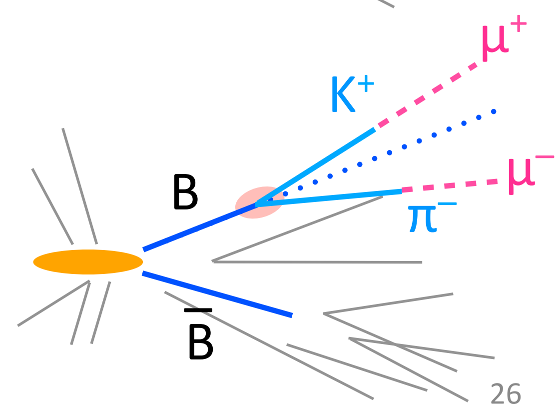
- **The $B_{s,d} \to \mu^+\mu^-$ signal**
  - two muons from one displaced vertex; momentum aligned with its flight direction; invariant mass peaking at $M(B_{s,d})$.



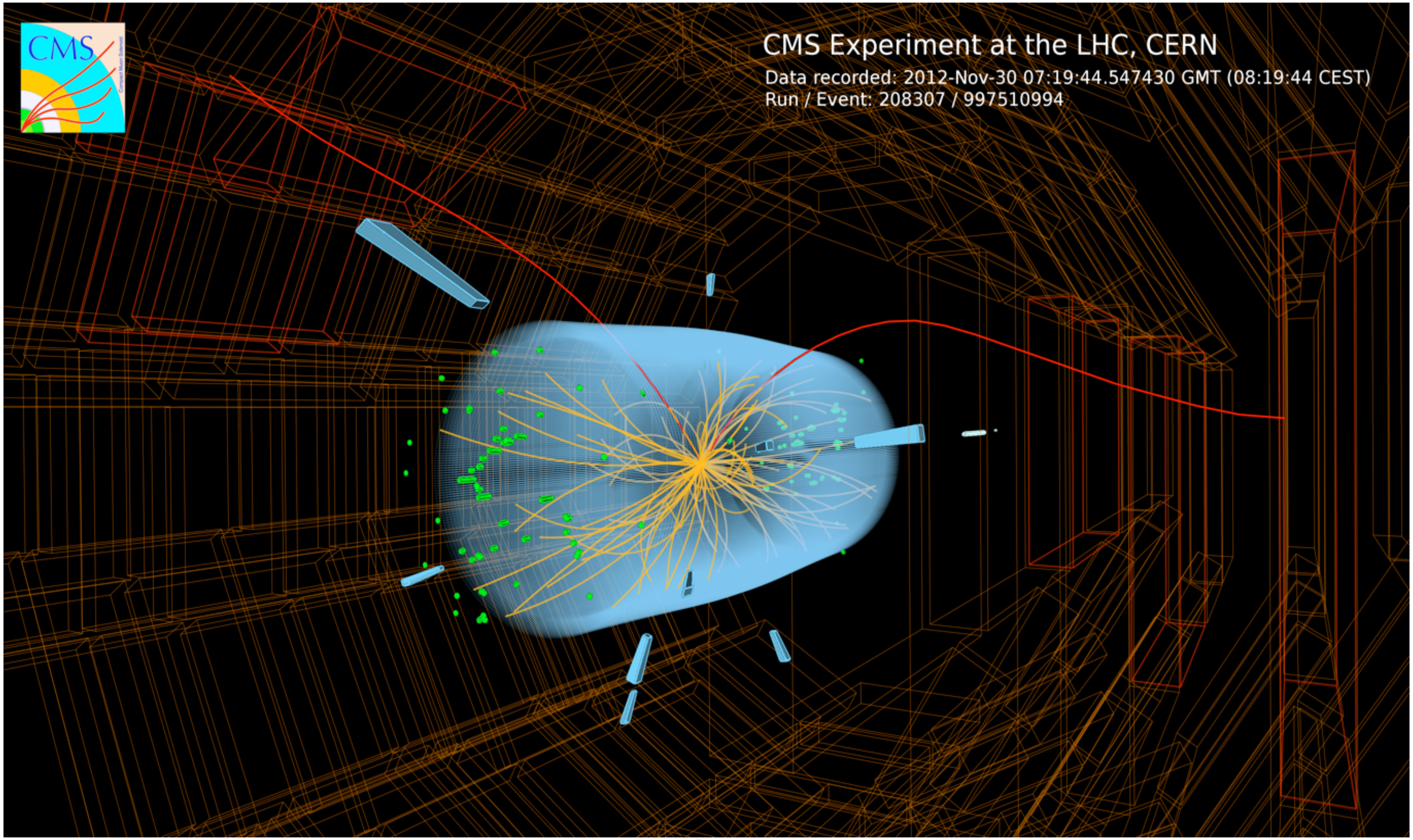
- **Background sources**
  - two semileptonic B decays
  - one semileptonic B + a misidentified hadron
  - rare background from single B meson decays: e.g. **$B \to K\pi/KK$** (*peaking*), **$B_s \to K^-\mu^+\nu$, $\Lambda_b \to p\mu\nu$** (*not peaking*), where hadrons either appear to be muons through decay or "punch-through"



Powerful background suppression reached by muon quality, well-reconstructed secondary vertex, isolation, pointing angle, and $M(\mu\mu)$ resolution.

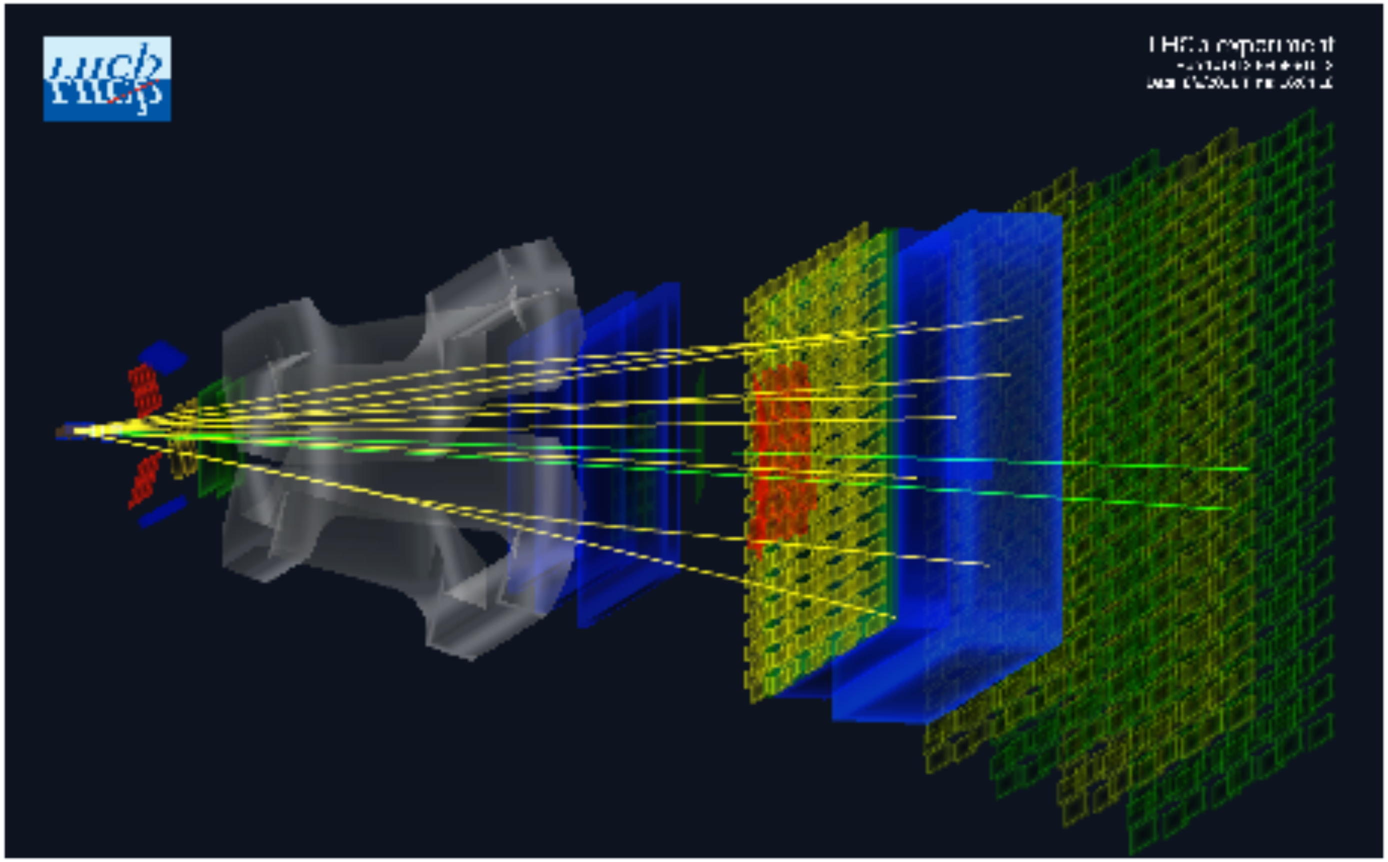


CMS: Representative Event Display



Extended Data Figure 3 | A candidate $B_s^0 \to \mu^+\mu^-$ event recorded in the CMS detector in 2012 produced in proton-proton collisions at 8 TeV. The red arched curves represent the trajectories of the muons from B_s^0 decay.

LHCb: Representative Event Display



Extended Data Figure 4 | A candidate $B_s^0 \to \mu^+\mu^-$ decay recorded in the LHCb detector in 2011 produced in proton-proton collisions at 7 TeV. The proton-proton collision occur on the left-hand side, at the origin of the trajectories depicted with the yellow curves. The green curves represent the trajectories of the muons from B_s^0 candidate decay.

Analysis Strategy

Analysis strategy	Reconstruct the dimuon mass and select events within broad mass window
Background discrimination	Use a Multivariate analysis of vertex toplogy, isolation, track and vertex quality with Boosted Decision Tree (BDT) to remove backgrounds
Signal extraction	**Unbinned maximum likelihood (UML) fit to the mass distribution, w/ the events in BDT bins.** Normalized by $B^+ \rightarrow J/\psi K^+$ decays to fit directly for B_s and B_d branching fractions.
Expected performance	**Very similar on these final states**

Despite many differences in detectors and interaction rates, analysis strategies are similar!
LHCb takes advantage of its ability to collect large samples of $B_{s,d} \rightarrow \pi\pi$, πK, KK to check many aspects of the analysis.

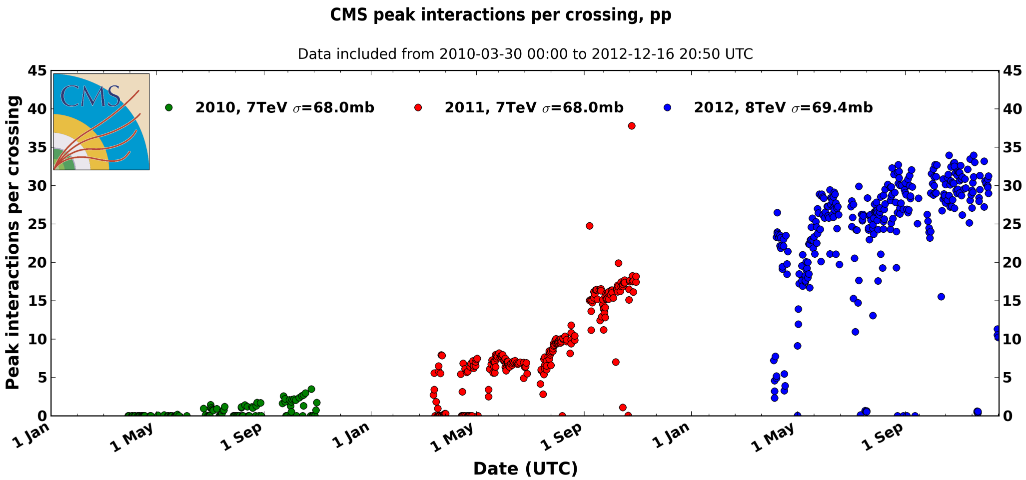
ONE VERY BIG DIFFERENCE
PILE-UP IN CMS

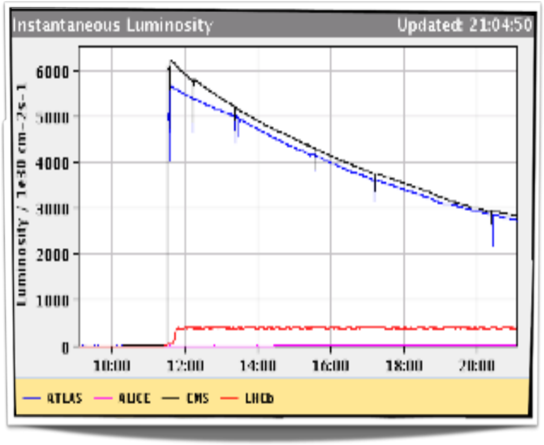
Due to requirement of high-*mass* physics searches, CMS needs to run at a much higher luminosity, but then has an issue with event overlap — the so-called pile-up!

CMS has in average ~25 collisions in each beam crossing; while LHCb has only 1.7 collisions per crossing. LHCb can only take few x 10^{32} because readout electronics saturates. Rate is kept steady at LHCb by "luminosity" leveling. CMS can correctly associate charged tracks to the true interaction vertex because of highly segmented pixel detector and all silicon strip outer tracker (100 Mega-pixel 3D camera)

CMS has shown that all variables used in this (mainly) ALL CHARGED PARTICLE analysis are "pileup-safe" using data and simulation.

Data Samples, Luminosity, and Pileup in CMS and LHCb





CMS has more variation in running conditions than LHCb. This complicates the analysis (more categories that have to be handled separately). Note for CMS luminosity and pileup much higher at 8 TeV (2012) than at 7 TeV (2011).

Luminosity	CMS (fb^{-1})	LHCb (fb^{-1})
7 TeV	5 (av pileup 8)	1
8 TeV	19.7(av pileup 21)	2

CMS DIMUON TRIGGERS

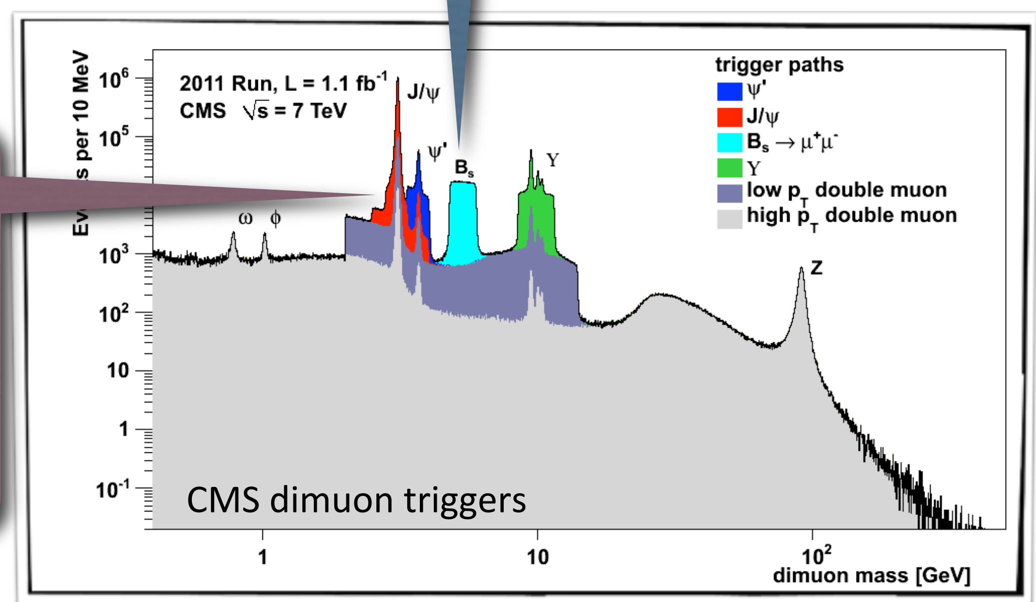
- The analysis starts from the dimuon triggers:
  - Hardware trigger (L1) – tracks formed from tracks from hits in muon chambers (P_T)
  - Software trigger (HLT) with full tracking & vertex reconstruction.
  - Criteria are tightened when luminosity increases.

B→μμ trigger

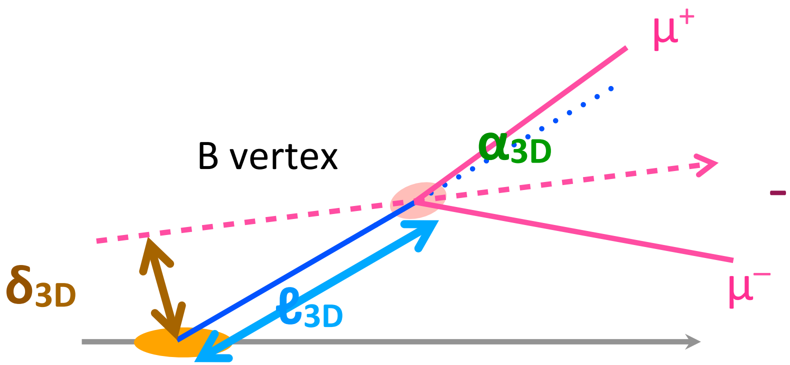
M(μμ) window
good μμ vertexing quality
minimal μ and μμ p_T thresholds
maximal distance of closest approach for μμ

Displaced J/ψ trigger

M(μμ) window
good μμ vertexing quality
minimal μ and μμ p_T thresholds
maximal distance of closest approach for μμ
significant transverse flight distance
collinear pointing angle



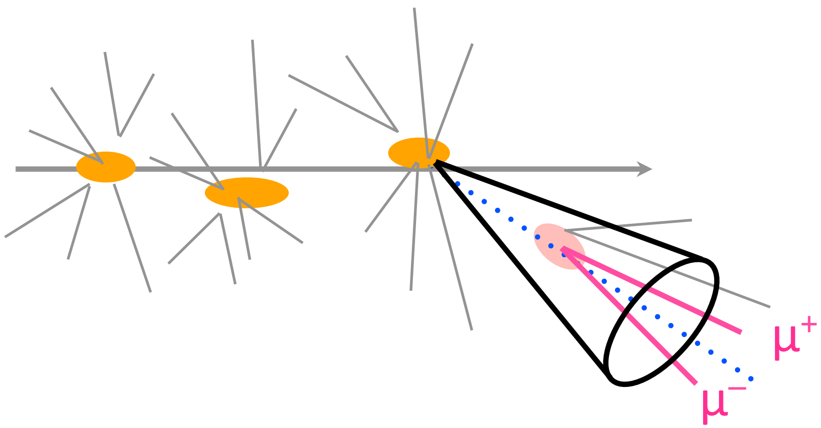
BACKGROUND SUPPRESSION USING VERTEX VARIABLES



- **α_{3D}**: pointing angle – the angle between candidate momentum and flight direction in 3D.

- **δ_{3D} & $\delta_{3D}/\sigma(\delta_{3D})$**: the 3D impact parameter of the B with respect to the primary vertex.
- **ℓ_{3D} & $\ell_{3D}/\sigma(\ell_{3D})$**: B flight length respect to the primary vertex, measured in 3D.
- **χ^2/d.o.f.** : vertex fit χ^2 of the B candidate.

BACKGROUND SUPPRESSION USING ISOLATION



- **Iso**: isolation of the B candidate, defined in cone around μμ momentum, loop over the charged tracks.
- $\mathbf{d^0_{ca}}$: minimum distance of closest approach of a track in the event to the decay vertex
- $\mathbf{N_{trk}}$: number of tracks in the vicinity of the B decay vertex.

Put everything into a BDT to further enhance background suppression power.

ANALYSIS : NORMALIZATION

- The absolute number of B mesons determined from production cross section has a hugh uncertainty. Use a well-measured decay channel to normalize the branching fractions of $B_{s,d} \rightarrow \mu^+\mu^-$.
- Normalization channel **$B^+ \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\, K^+$**
  (LHCb also take $B^0 \rightarrow K\pi$ as normalization / cross check).
- Control channel for Calibrations/validations with $B^+ \rightarrow J/\psi K^+$, $B^0 \rightarrow K^-\pi$ (LHCb) $^+$and $B_s \rightarrow J/\psi\phi(K^+K^-)$.

$$BR(B_{s,d} \rightarrow \mu^+\mu^-) = \frac{N_S}{N(B^\pm \rightarrow J/\psi K^\pm)} \times BR(B^\pm \rightarrow J/\psi K^\pm) \times \frac{A(B^\pm)}{A(B_s)} \frac{\varepsilon^{ana}(B^\pm)}{\varepsilon^{ana}(B_s)} \frac{\varepsilon^{\mu}(B^\pm)}{\varepsilon^{\mu}(B_s)} \frac{\varepsilon^{trig}(B^\pm)}{\varepsilon^{trig}(B_s)} \frac{f_u}{f_s}$$

Acceptance

Selection efficiency

muon identification

Trigger efficiency

B-hadronization composition, for B_s only
(LHCb measured value: 0.259±0.015)

Similar trigger & selection for reducing systematics

Issues in Comparing Theory and Experiment

- **Absolute normalization** of branching fraction for B_s given lack of accurate absolute branching fraction from B factories
  - The B factories, running on the Υ(4S), have provided absolute branching fractions for B_d and B_u
  - KEKB/BELLE has not done enough running, and may never, on the the Υ(5S) to provide sufficiently precise absolute branching fractions for the B_s for our purposes
- **Radiative corrections** to the branching fractions in theory and experiment
- **Lack of knowledge of some** branching fractions and production cross sections needed for modeling the **background distributions**
  - Λ_b absolute branching fractions and production cross sections were not established
    - Especially, $\Lambda_b \rightarrow p\mu\nu$.
- Alignment of **measured quantity** with SM computation of the branching fraction, given the complex nature of the B_s time dependence
  - Experiment: time-integrated branching fraction
  - Theory: CP average branching fraction at production (t=0)

Absolute Normalization

- Normalize to $B^+ \rightarrow J/\psi K^+$
  - decay mode with final state similar to $\mu^+\mu^-$ (reduce dependence on knowledge of trigger efficiency, etc.); and
  - absolute BR is known from B factories
- To get the BR for B_s, we need the fragmentation probability of a b-quark to an s-quark vs to a u(or d)-quark, a.k.a f_s/f_u.
  - LHCb derives this from the ratios of decays such as $B_s \rightarrow D_s \pi^+$ to $B^0 \rightarrow D^- K^+$.
  - There is almost no variation of this in η or P_T within LHCb's acceptance

$$\frac{f_s}{f_u} = 0.259 \pm 0.015$$

LHCb, J. High Energy Phys. **4,** 1(2013)band LHCb-CONF-2013-011

- CMS demonstrated (to 5%) that it is also valid at the lower average values of η and higher values of P_T in CMS by studying the kinematic dependence of the ratio of $B^+ \rightarrow J/\psi$ K+ to $B_s \rightarrow J/\psi$, $\phi(1020)$, $\phi \rightarrow K^+K^-$.

Radiative Corrections

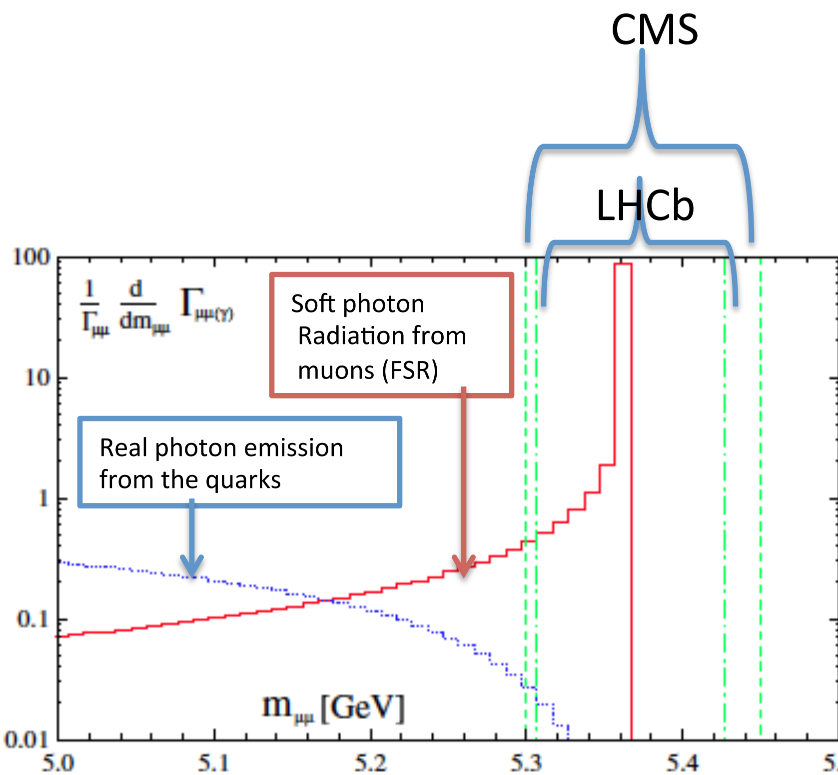


FIG. 1 (color online). Contributions to the dimuon invariant-mass spectrum in $B_s \to \mu^+\mu^-(n\gamma)$ with $n = 0, 1, 2, \ldots$ (see the text). Both of them are displayed in bins of 0.01 GeV width.

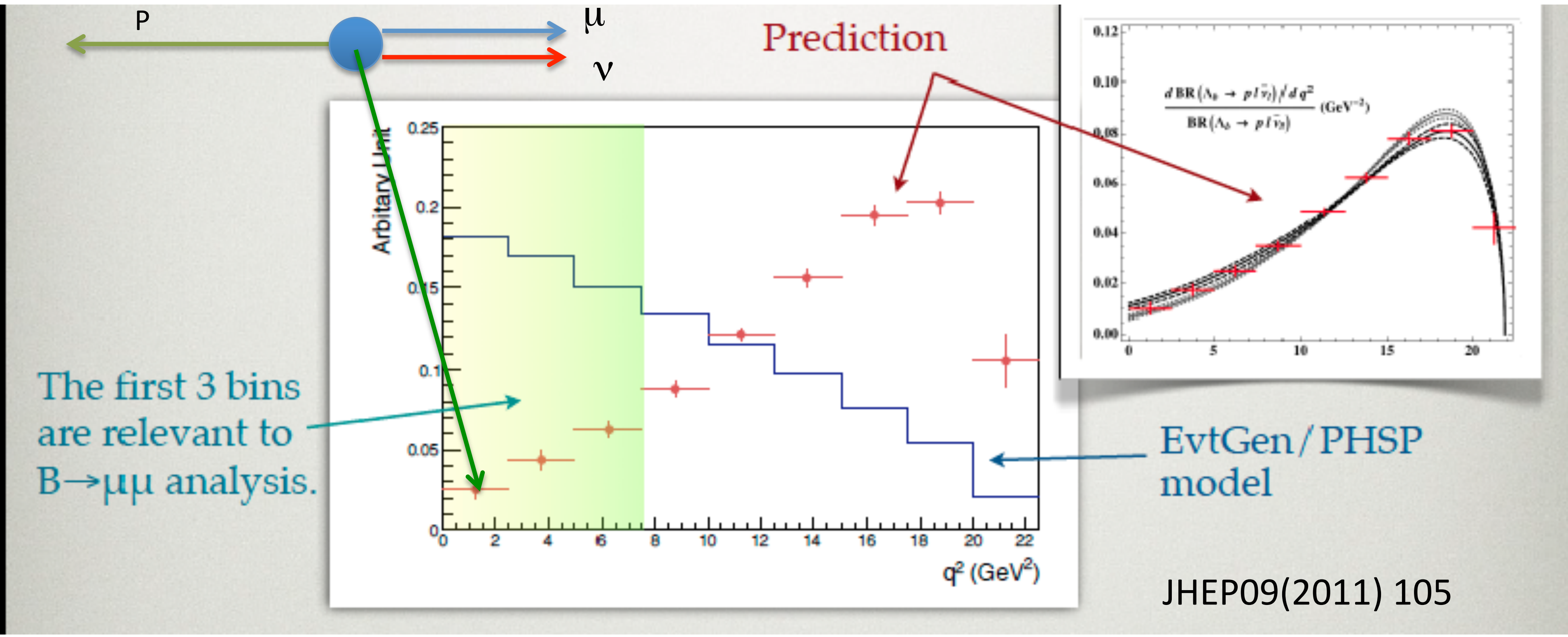
FROM BOBETH

- **Are radiative corrections handled consistently between experiment and theory**?
  - Experiments compensate for FSR (red tail) by using PHOTOS in their simulation so loss is restored in the efficiency correction
  - Experiments ignore the real emission from the quark lines but it is a negligible addition in their mass windows
  - Theorists do not include the real emission in their calculation of the BR
- **Answer is yes,** with any residual effect small compared to overall theory uncertainty

$\Lambda_b \rightarrow p\mu\nu$ Model

- This is the dominant semileptonic B decay background in the analysis.
  - This was very uncertain because branching fraction was unknown at time of the Nature article
  - **Since measured by LHCb!!! (prelim from talk [3.9± 0.8]x10^{-4})**
- Branching fraction used in CMS publication
  - BF = 6.5 x 10^{-4}, with ± 100% uncertainty
  - Decay was modeled as phase space from EvtGen
- Updated BF and decay model to agree with LHCb
  - BF = (4.94 ± 2.19) X $10^{-4.}$
  - Decay model based on reference shown in next slide, is calculated and given as a function of q^2.

$\Lambda_b \rightarrow p\mu\nu$ decay Model



$q^2 = M^2(\mu\nu)$: lower $q^2 \rightarrow$ higher $M(p\mu) \rightarrow$ closer to $B_d \rightarrow \mu\mu$ signal
While the level of the BR turned out to not as significant as feared, the new q^2 model reduces the $\Lambda_b \rightarrow p\mu\nu$ contribution by a large factor.

Aligning Theory and Experiment

- **Theoretical predictions are for "branching fractions" at production (t=0)**
- **Experiments present "time averaged" results.**
  - **Unless the state observed is a QM superposition of states with significantly different lifetimes, these should agree. If the lifetimes are different, then the experimental average is not the same as the value at t=0**
- The B_d is a superposition of B^0_H and B^0_L, whose lifetime is the same within 10^{-3}.
- **The B_s is a superposition of B_{sH} and B_{sL}, whose lifetimes are now measured to be different by ~10-15%**
  - From helicity considerations, $\mu^+\mu^-$ final state is CP eigenstate and corresponds to high degree to B_{sH}, since CP is nearly conserved in the SM in this decay
  - The experimental result and the theoretical calculation must be brought into "alignment"
- The mechanism for aligning theory and experiment is given in
  - Bobeth and De Bruyn et al., Probing New Physics via the Bs→ $\mu^+\mu^-$ effective lifetime, Phys. Rev. Lett. **109,** 041801
  - Note that the correction scheme requires that one assumes no NP effect in the time evolution and is, therefore, model dependent

$$\mathrm{BR}(B_s \Rightarrow f)_{theo} = \left[\frac{1 - y_s^2}{1 + A^f_{\Delta\Gamma}\, y_s}\right] \mathrm{BR}(B_s \Rightarrow f)_{exp}$$

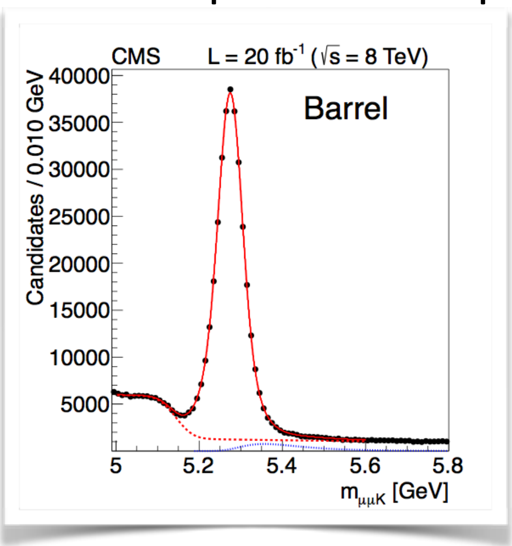
$A^f_{\Delta\Gamma}$ =1 in SM

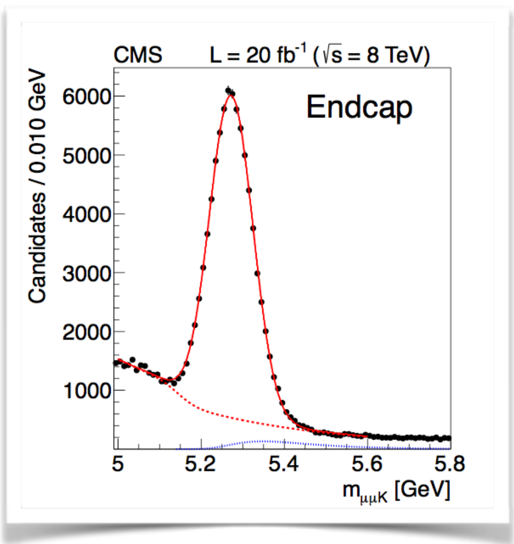
$$\text{where } y_s = \frac{\Delta\Gamma_s}{2\Gamma_s},\quad A^f_{\Delta\Gamma} = \frac{R^f_H - R^f_L}{R^f_H + R^f_L},$$

$$\text{and } \tau^{-1}_{B_s} = \Gamma_{(s)} = \frac{\left[\Gamma^{(s)}_L + \Gamma^{(s)}_H\right]}{2}$$

CMS Normalization & Control channels

- The **"normalization" channel $B^+ \rightarrow J/\psi(\rightarrow\mu^+\mu^-)\ K^+$** is fairly easy to measure ("big" branching fraction).
- Similar reconstruction efficiency and analysis selection criteria cancels some of the systematic uncertainties.
- Remaining issues are: **5% from the yield extraction** and the **f_s/f_u parameter (~9%)**.
- The **"control" channel $B_s \rightarrow J/\psi(\rightarrow\mu^+\mu^-)\ \phi(1020),\ \phi\rightarrow K^+K^-$** is used to show MC is a good representation of the data and validate other aspects of the procedure with a large (>50,000) sample of B_s mesons



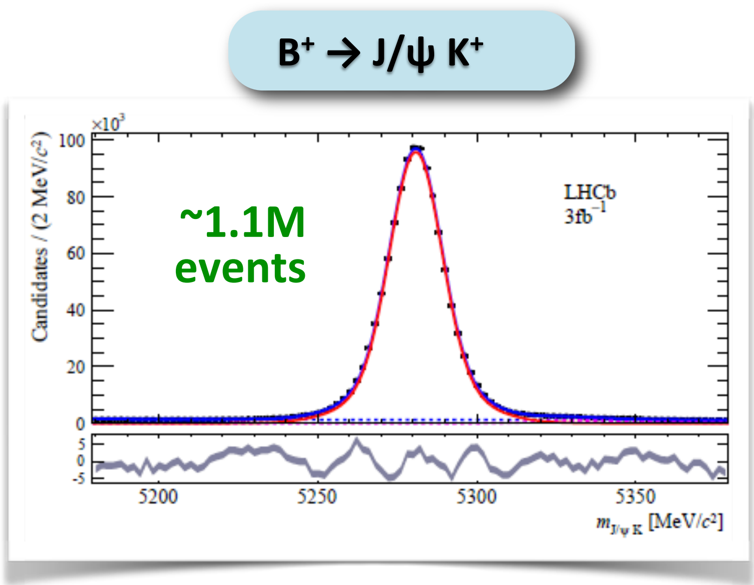


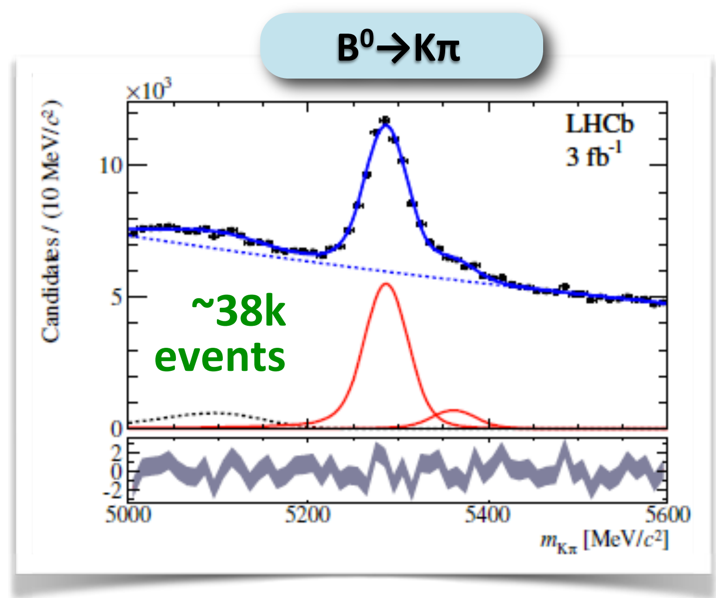
Channel	7 TeV	8 TeV
Barrel	$(71.2\pm4.1)\times10^3$	$(309\pm16)\times10^3$
Endcap	$(21.4\pm1.1)\times10^3$	$(69.3\pm3.5)\times10^3$

Note poorer endcap resoluton

LHCb Normalization & Control channels

- LHCb analysis uses two **normalization** channels: **$B^+ \rightarrow J/\psi(\rightarrow\mu^+\mu^-) K^+$** (similar trigger as the signal) and **$B^0 \rightarrow K\pi$** (same event topology with 2 tracks, but with different trigger).
- The only remaining factor is the hadronization fraction **f_s/f_u parameter (~6% precision)**, measured from the ratio between $B_s \rightarrow D_s\mu X$ and $B \rightarrow D\mu X$ decays.

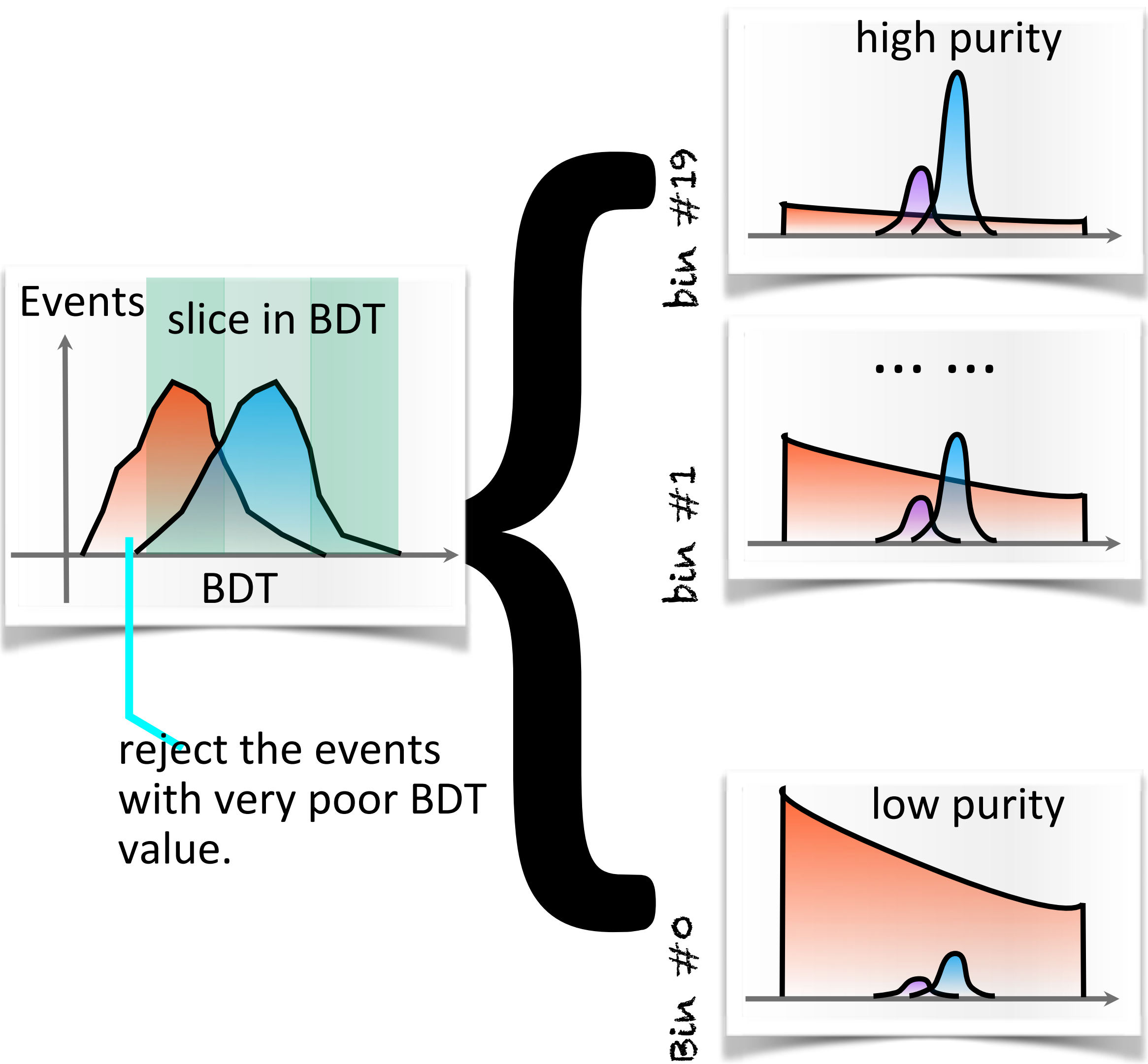




Boosted Decision Tree

- An ensemble (~few hundred) of "related" decision trees each placing different selection requirements on the individual variables (typically ~10 for this analysis) of an event to achieve the best classification of events as 'signal-like' or 'background-like', expressed as a sum of weights from the individual trees.
- BDTs are "trained" on samples of known signal and background events (from MC or data) to get the weights
- Trees are connected by "boosting" algorithm, which gives training events misclassified by the preceding tree higher weights for the next tree
  - This gives a single variable that takes into account all the variables used in the tree to give a number that classifies the event. The bigger the number, usually the more "signal-like" is the event.
- For training
  - CMS uses simulated signal and data from the mass sidebands for background (with a lot of care)
  - LHCb uses both simulated data and background (major computational effort)
- Events from data are each passed through ALL trees and a composite "score" or "weight" is obtained that indicates how signal-like the event is.
- BDTs are used elsewhere in the analyses, e.g. for optimal muon id and rejection of decays and hadronic punch-through in CMS.

Fitting Mass Spectra in BDT Categories



$$L_{\rm tot} = \prod_{i=0}^{19} L_i L_i^{\rm constr}$$

Events are characterized by BDT value and dimuon mass (and a per-event mass error in CMS).

The events are sorted into "bins" of the BDT variable

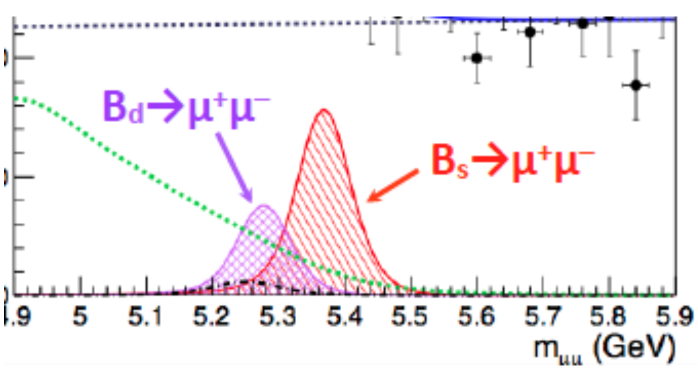
Unbinned maximum likelihood fit to the dimuon mass spectrum simultaneously in all BDT bins is used to get the total number of signal events, captured in the fit by the two branching fractions

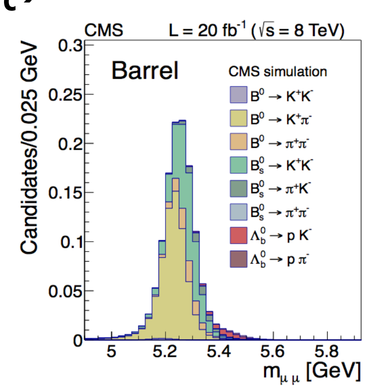
Variables entering BDTs

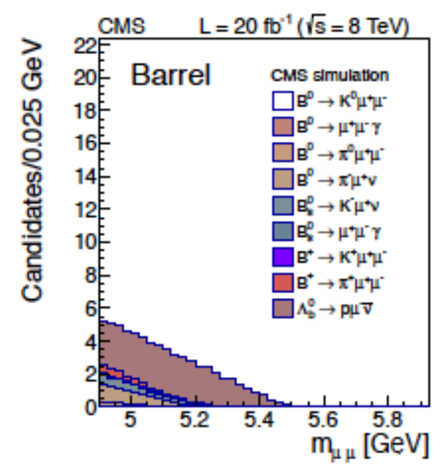
- CMS and LHCb each use ~10 selected variables
  - Quality of muon tracks/reconstruction
  - Kinematic variables of muons and $B_{s,d}$
  - Primary and secondary vertex topology and fit quality
  - Isolation variables
- Training
  - CMS uses simulated signal and background from sidebands of dimuon mass plot
  - LHCb uses simulated signal and background events

Background Contributions

- Combinatorial: caused by combinations of muons coming from two different b or c weak decays, b,c→Xμν. Studied on sidebands
- Rare Semileptonic or missing neutral: like B_s→ $K^-\mu^+\nu_\mu$ or Λ_b→$p\mu^-\nu_\mu$, where the hadron is mistaken for a muon, or Bd→$\pi^0\mu^+\mu^-$ .Studied on data and MC
- Peaking: Two body decays like B_s→K^+K^-, where both hadrons are misidentified as muons. Studied on data and MC
- Cross- feed: between B_s→$\mu^+\mu^-$ and B_d→ $\mu^+\mu^-$, because of mass resolution. Mainly a problem for Bd (Bs is a background!). Studied on MC
  - **CMS mass resolution varies from 32 MeV in barrel to 75 MeV at highest eta in endcap**
  - **LHCb mass resolution almost constant at 25 MeV/c2**
  - **B_s – B_d Mass difference is 87 MeV/c2**







Aligning the Analyses

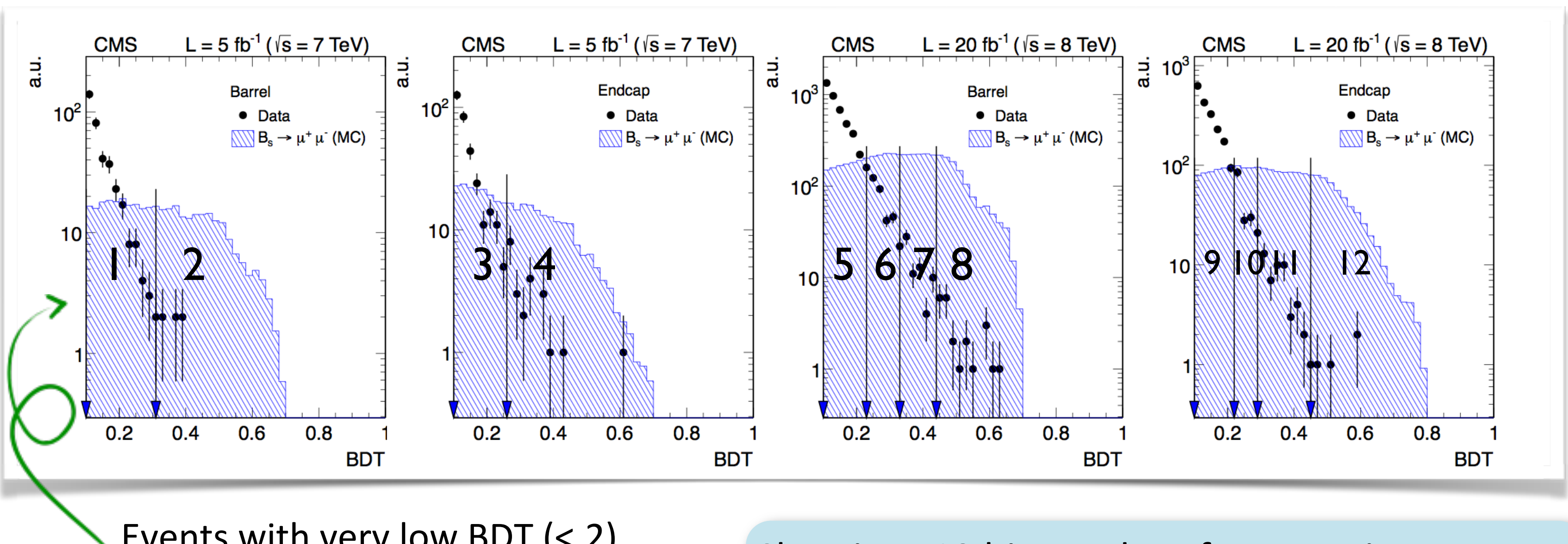
- Maintain the same analyses/fitting strategy as in the separate analyses (fortunately quite similar) but there are differences
- Adopt the same external inputs(physics parameters, branching fractions (e.g. for background estimates) , models) to the extent possible
- Apply small fixes and corrections and make sure, at the end, the experiments are as consistent as possible
- Share the events with all variables needed to compute unbinned likelihoods in categories of the final BDT variable used by each experiment
  - The BDT variables, their distributions and the number of BDT categories differs between the two experiments
- Carry out EXTENDED Unbinned Maximum Likelihood fit for a common value of BR(B_s) and BR(B_d) while sharing two parameters, constrained within their uncertainties:
  - BR(**$B^+ \to J/\psi(\to\mu^+\mu^-) K^{+}$)**
  - fs/fu, with additional systematic uncertainty of 5% for CMS

Fit Categories

- **CMS has many variations**
  - Luminosity rose steadily in 2011 and much of 2012, leading to trigger changes and change in pileup
    - Split into 2011 (7 TeV) and 2012 (8 TeV)
  - CMS has very different mass resolution in barrel and endcap
    - Split into Central (both muons in barrel) or Endcap (>0 muons in endcap)
  - Some BDT bins (at very low values) had so much background that they could only harm the result
    - Cut applied, BDT>0.2
    - For 2011, use 2 bins for barrel & 2 bins for endcap = 4 bins
    - For 2012, use 4 bins for barrel & 4 for and endcap =8 bins
    - In each year, bins are chosen to have ~= expected signal based on MC
  - CMS used a "per-event" resolution in its likelihood to account for the variation in its resolution with η.
- **LHCb was very stable**
  - Ran at essentially same luminosity – saturated readout rate
    - Stable trigger conditions
  - Very uniform mass resolution
  - Chose 8 bins in BDT, split based on sophisticated optimization algorithm using toy MC
- **Summary: 20 bins in BDT – 12 CMS, 8 LHCb each with a mass distribution including potential signal and backgrounds described in slide 47.**

CMS BDT Categorizing

- In order to extract the maximal statistical power out of the analysis, the events are categorized according to the **beam energy** and **detector region**, as well as the **combined BDT value**.
- An average efficiency for each bin is determined from simulation



Events with very low BDT (<.2) are not used due to the very large background level.

Chop into 12 bins and perform maximum likelihood fits simultaneously.

The Fit Function

- Extended likelihood has 5 contributions

$$L = N_{B_s^0} F_{B_s^0} + N_{B^0} F_{B^0} + N_{\rm comb} F_{\rm comb} + N_{\rm peak} F_{\rm peak} + N_{\rm semi} F_{\rm semi}$$

- Systematic uncertainties are added as Gaussian constrained (or in some cases log normal) nuisance parameters
- Efficiencies in each BDT bin
  - CMS - from simulation
  - LHC - from analysis of data on two body hadronic states $B_{s,d} \rightarrow h^+h'^-$, where h,h' = π or K.
- Total likelihood is the product of all independent categories L_i and of all constraints

$$L_{\rm tot} = \prod_{i=0}^{19} L_i L_i^{\rm constr}$$

The Combined Fit

- RooWorkspace containing data and all PDFs
- Construct a Global likelihood
  - B_s and B_d branching fractions are common and are the main results
  - Shared (correlated) parameters
    - fs/fu: CMS value is constrained to LHCb value with an "extra" 5% Gaussian error to account for possible P_T-η dependence
    - Branching fractions for $B^+ \rightarrow J/\psi\ K^+$ and $J/\psi \rightarrow \mu^+\mu^-$
- Do fit using same compiler, gcc 4.8, and root, 5.34.10 with 1 core only (round off issues with multi-core)
- Total number of parameters needed to describe the signals, the backgrounds, the resolution, and all systematics was 155.

Fit Results

COMBINED

$$BF(B_s) = (2.8\ {+0.7}/{-0.6}) \times 10^{-9}$$
$$BF(B_d) = (3.9\ {+1.6}/{-1.4}) \times 10^{-10}$$

RATIO TO SM

$$S(B_s) = 0.76\ {+0.20}/{-0.18}$$
$$S(B_d) = 3.7\ {+1.6}/{-1.4}$$

BF RATIO

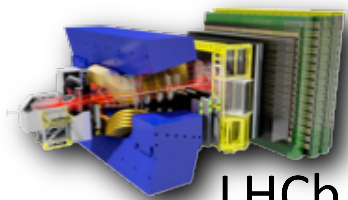
$$R = BF(B_d)/BF(B_s) = 0.14\ {+0.08}/{-0.06}$$

TH

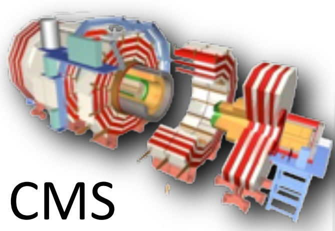
$$BF(B_s) = (3.65 \pm 0.23) \times 10^{-9}$$
$$BF(B_d) = (1.06 \pm 0.09) \times 10^{-10}$$

PRL 112, 101801 (2014)

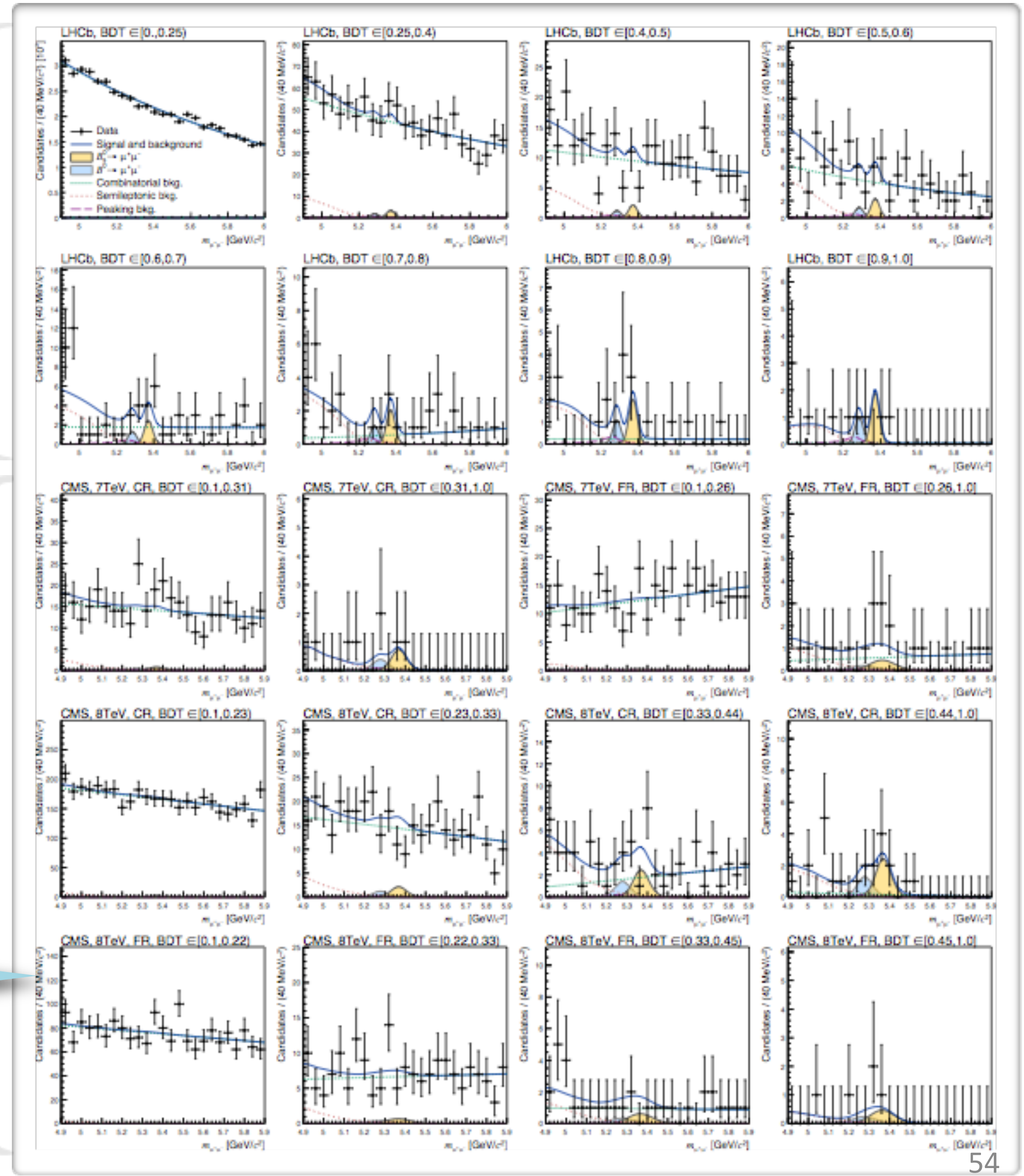
The Combined Fit



LHCb



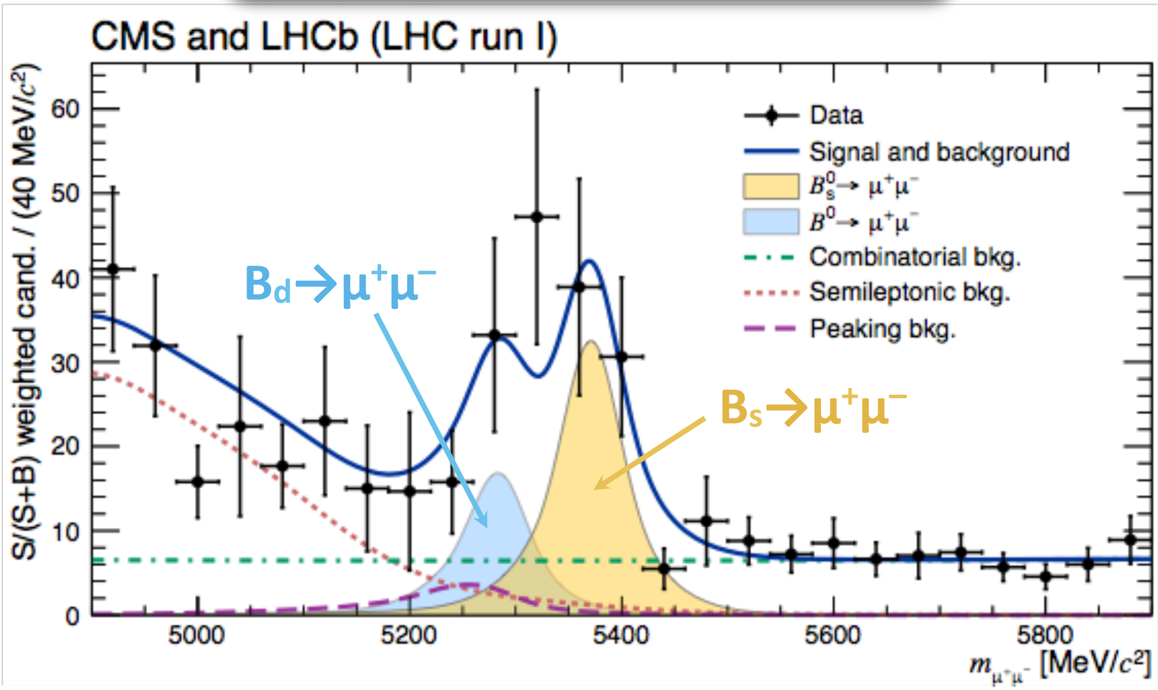
CMS



A giant fitter (w/ 155 free parameters) has been prepared to fit 20 BDT-categories simultaneously.
A single fit takes 6~12 hours to converge!

COMBINED RESULTS

S/(S+B) weighted mass from 20 bins



Channel	Branching fraction
$B_s \to \mu^+\mu^-$	$(2.8\ ^{+0.7}_{-0.6})\ \times 10^{-9}$
$B_d \to \mu^+\mu^-$	$(3.9\ ^{+1.6}_{-1.4})\ \times 10^{-10}$

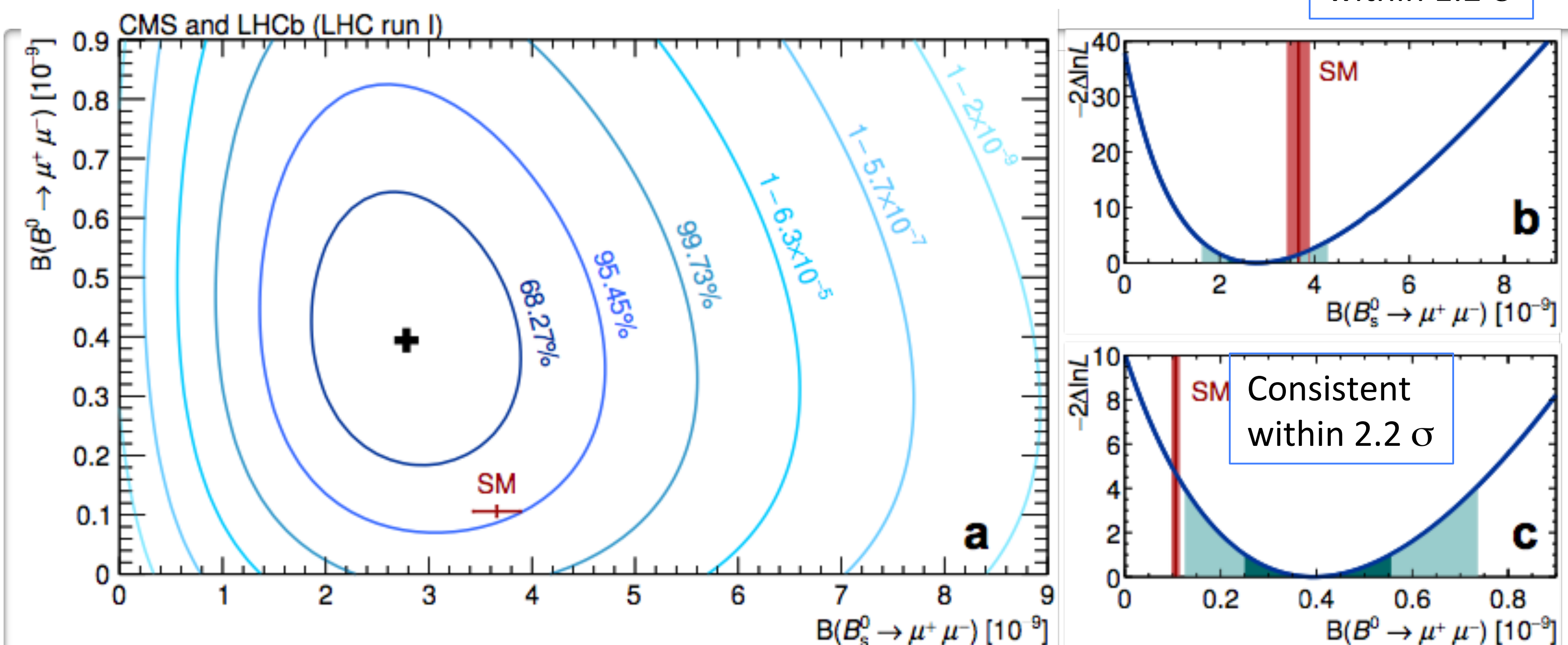
The uncertainties for both channels are reduced dramatically with the combined fit!

- Using Wilks' theorem, the statistical significance is computed to be **6.2σ** (7.4σ expected) for the $B_s \to \mu^+\mu^-$ decay.
- For the $B_d \to \mu^+\mu^-$ mode, a 3.2σ is obtained (0.8σ expected).

LIKELIHOOD SCANS

- Likelihood scans have been carried out in 2D and 1D.
- All other nuisance parameters are profiled in the scans.

Consistent within 1.2 σ



Consistent within 2.2 σ

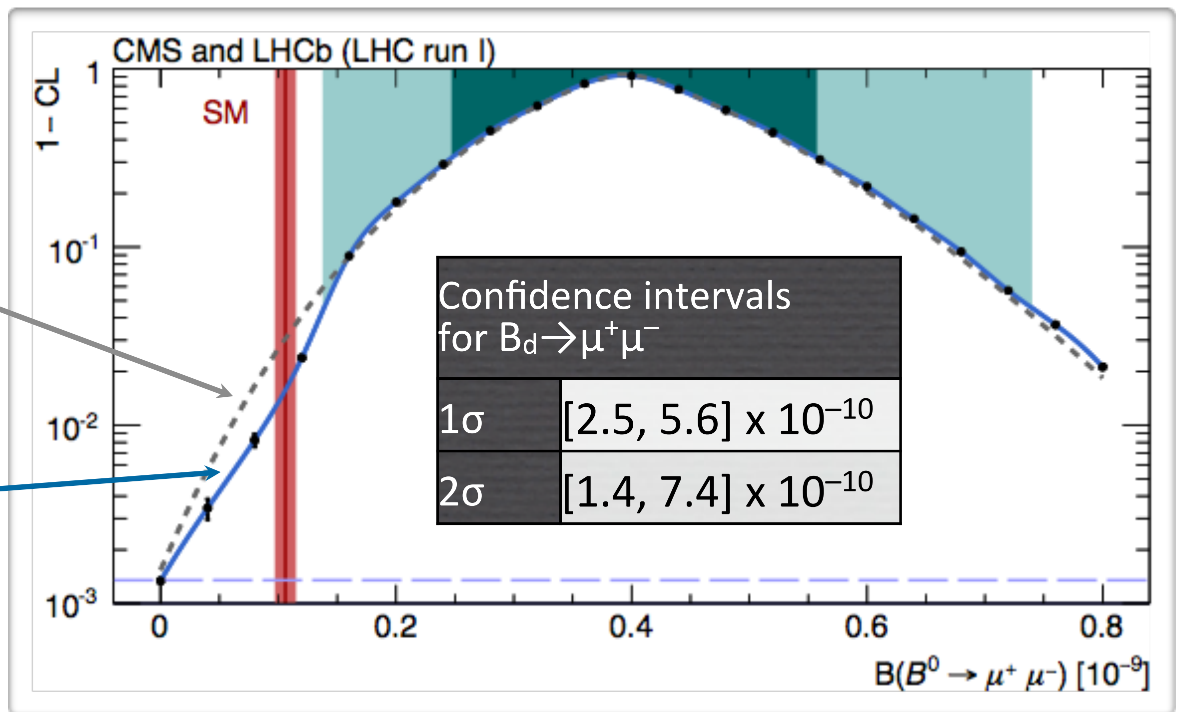
81 X 81 points in range $BR(B_s \to \mu^+\mu^-) \in [0,10^{-8}]$ and $BR(B_d \to \mu^+\mu^-) \in [0,10^{-9}]$
In single projections, unscanned BR is floated in fit.
These scans actually took 4-5 days with 240 CPU cores to calculate!

$B_d \rightarrow \mu^+\mu^-$ SIGNIFICANCE

- The B_d channel is very close to its physical boundary, the estimation of significance might not be fully sufficient with the Wilks' theorem.
- A full **Feldman-Cousins** procedure (based on pseudo experiments) has been introduced to provide a better estimation. The calculated statistical significance is **3.0σ** (consistent with **3.2σ** from Wilks' theorem).

Asymptotic approach (Wilks' theorem)

Feldman-Cousins procedure

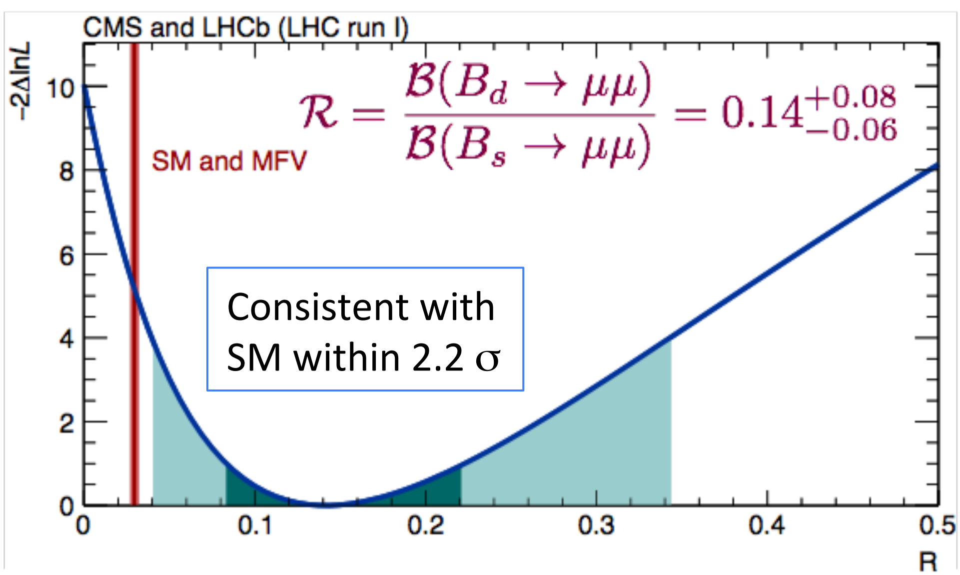


Confidence intervals for $B_d \rightarrow \mu^+\mu^-$	
1σ	[2.5, 5.6] x 10^{-10}
2σ	[1.4, 7.4] x 10^{-10}

RATIOS OF BRANCHING FRACTIONS

- The branching fractions relative to their SM predictions are also measured, which provide a direct comparison to the theory:

Channel	Ratio to SM branching fraction
$B_s \rightarrow \mu^+\mu^-$	$0.76^{+0.20}_{-0.18}$
$B_d \rightarrow \mu^+\mu^-$	$3.7^{+1.6}_{-1.4}$

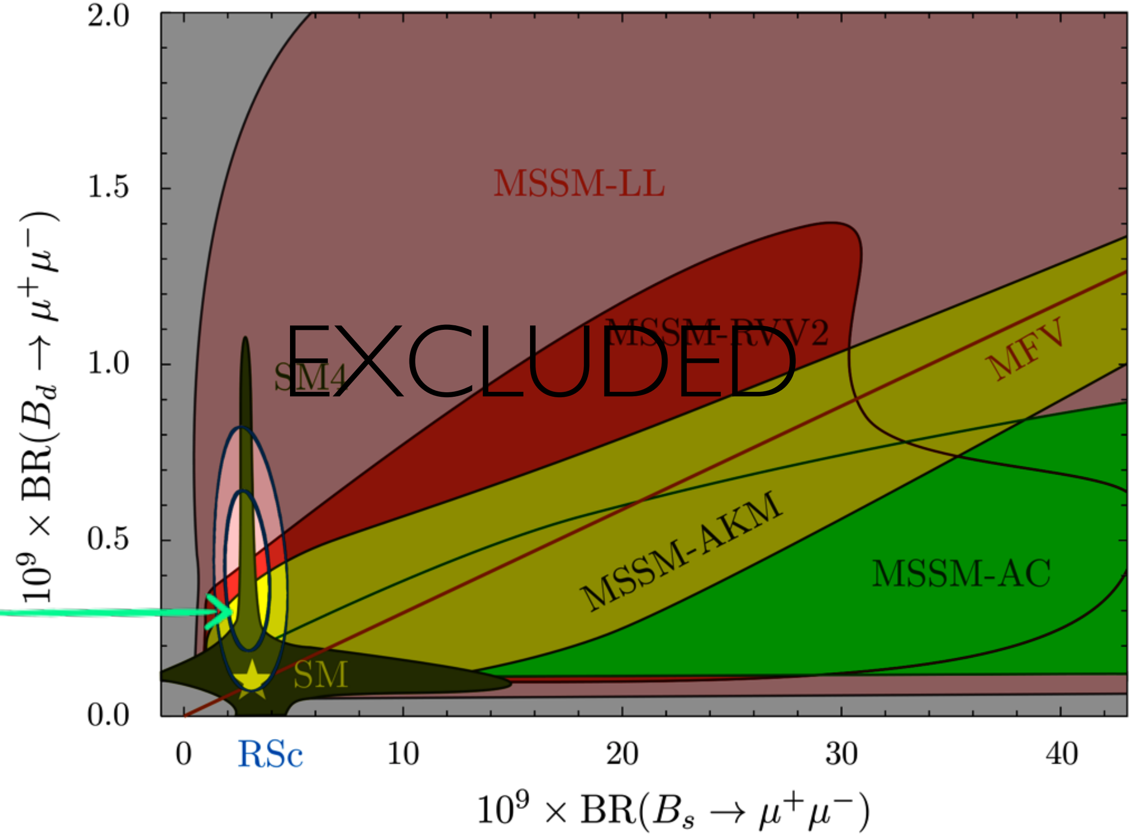


- Alternatively one can also measure the ratios between B_d and $B_s \rightarrow \mu^+\mu^-$ branching fractions, where the major part of theoretical uncertainty cancels.

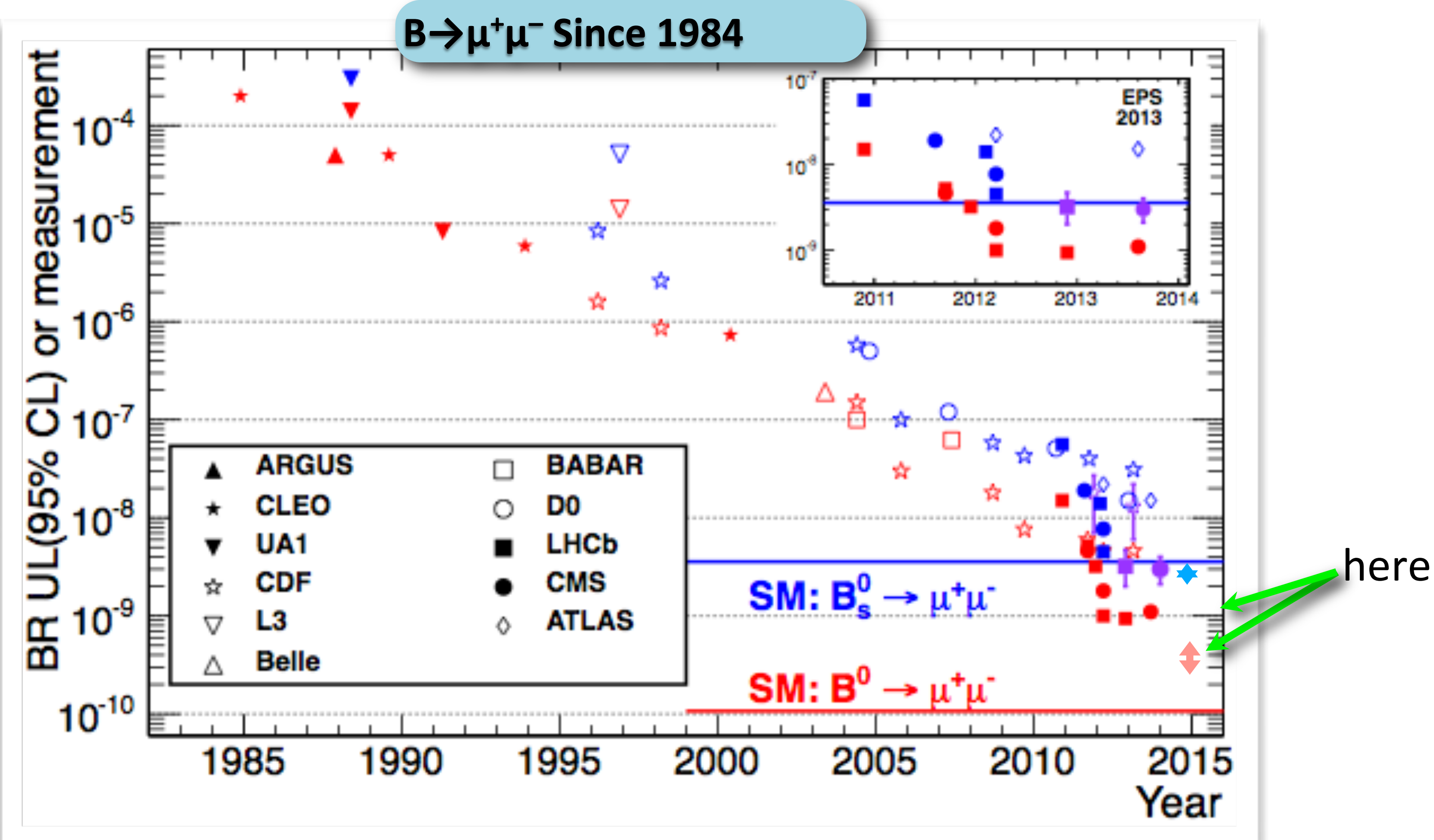
Revisiting our "limited" NP Space

- **Remember our DISCLAIMER** – this is just an exercise
- These models all have their allowed parameter space severely limited

Remaining space from the CMS+LHCb combined fit for these models.



Historical Summary



It took 30 years to finally measure the $B_s \to \mu^+\mu^-$ decay; The result turns out to be very close to the prediction and gives a stringent limit on the physics beyond the Standard Model. **There is still a possibility of ~50% deviation from the SM, which will be resolved by more statistics in the next few years.**

Forecast

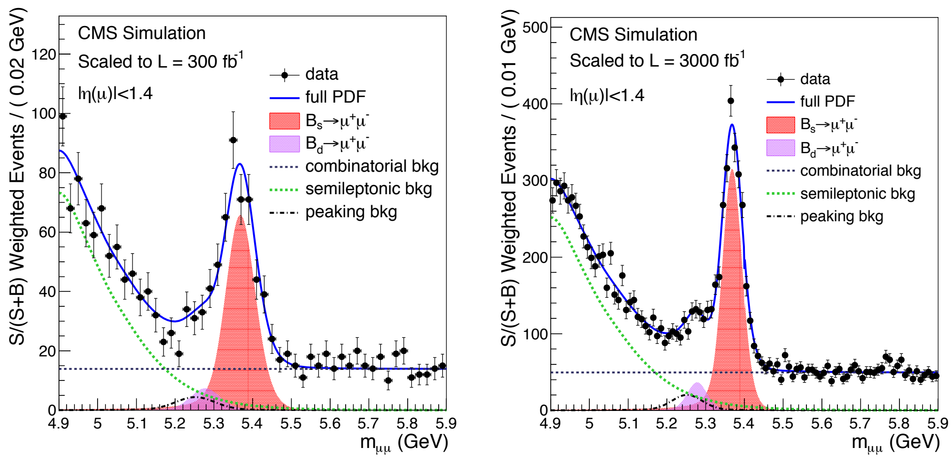


Figure 3: Projections of the mass fits to 300 fb^{-1} (left) and 3000 fb^{-1} (right) of integrated luminosity (L), respectively assuming the expected performances of Phase-I and Phase-II CMS detectors.

➡ Next target is $B_d \to \mu^+\mu^-$

Estimate of analysis sensitivity

$\mathcal{L}$ (fb^{-1})	$N(B^0_s)$	$N(B^0)$	$\delta\mathcal{B}(B^0_s \to \mu^+\mu^-)$	$\delta\mathcal{B}(B^0 \to \mu^+\mu^-)$	B^0 sign.	$\delta\frac{\mathcal{B}(B^0\to\mu^+\mu^-)}{\mathcal{B}(B^0_s\to\mu^+\mu^-)}$
20	18.2	2.2	35%	$>$ 100%	$0.0 - 1.5\,\sigma$	$>$ 100%
100	159	19	14%	63%	$0.6 - 2.5\,\sigma$	66%
300	478	57	12%	41%	$1.5 - 3.5\,\sigma$	43%
300 (barrel)	346	42	13%	48%	$1.2 - 3.3\,\sigma$	50%
3000 (barrel)	2250	271	11%	18%	$5.6 - 8.0\,\sigma$	21%

CMS-PAS-FTR-14-015

Physics Summary

- **A full "fitter level" combination** for CMS and LHCb analyses for $B_{s,d} \to \mu^+\mu^-$ at likelihood level has been carried out.
- The best fitted branching fractions are

  $$BF(B_s) = (2.8\ +0.7/-0.6) \times 10^{-9}$$
  $$BF(B_d) = (3.9\ +1.6/-1.4) \times 10^{-10}$$

  The observed significance for $B_s \to \mu^+\mu^-$ is 6.2σ.
- **$B_s \to \mu^+\mu^-$ is "observed"**
- Based on the combined fitter, the 1D/2D likelihood scans, Feldman-Cousins scans, the combined mass plots have been produced.
- The result is consistent with the SM, but more data will resolve whether the small offsets that exist will disappear
- Published in Nature Science, June 4, 2015

Conclusions

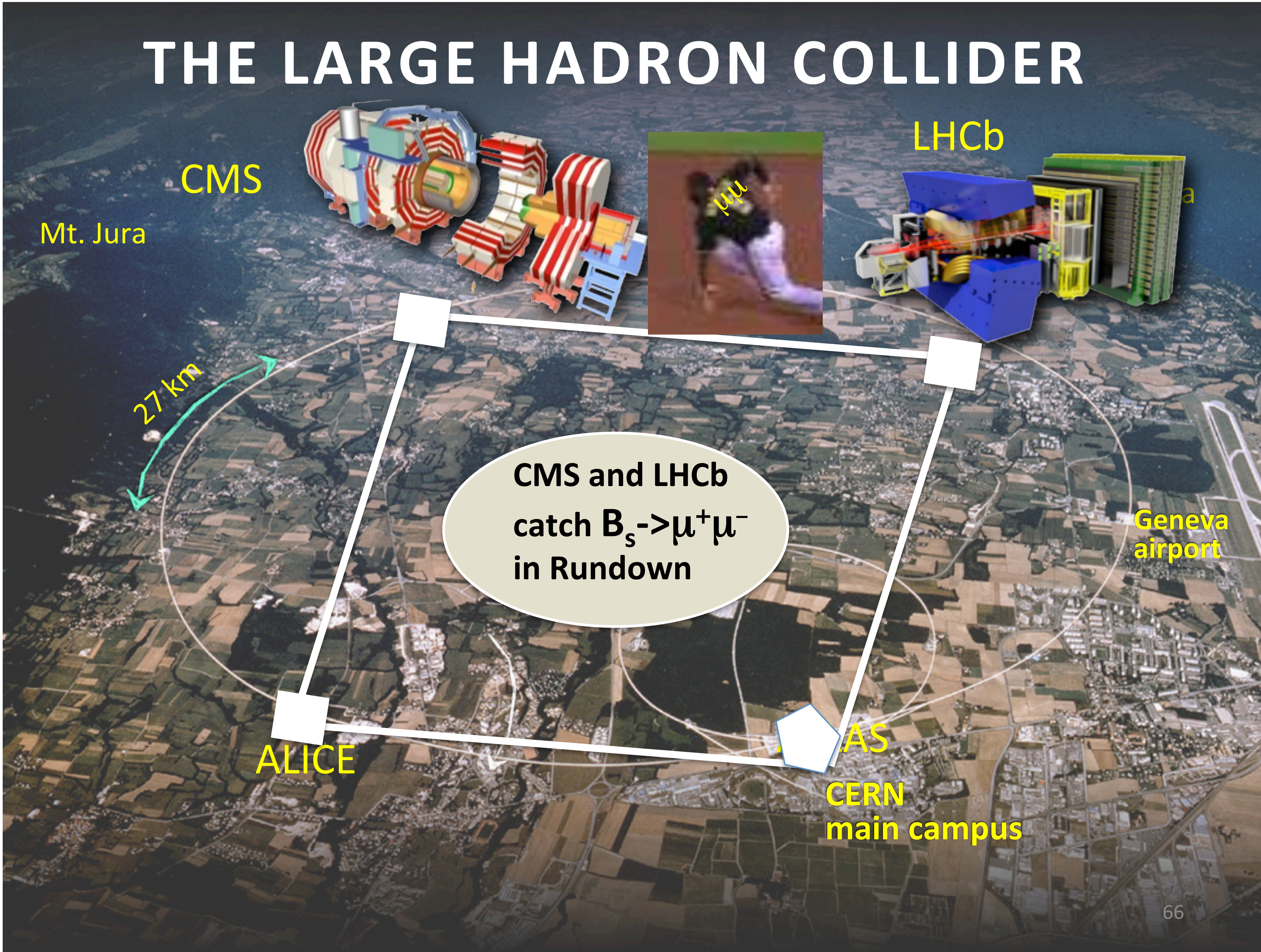
- The combined analysis produced results not too dissimilar from an earlier naïve averaging and within expectations
- The combination process produced refinements to both analyses ***
- The process also exposed the work that needs to be done to guarantee that both analyses can continue to take advantage of the more than a factor of 100 in luminosity that will become available of the next two decades and to combine results to help corner the B_d.
- Projected sensitivities tell the theory community that they will need to sharpen their calculations as well
- The observation of $B_d \rightarrow \mu^+\mu^-$ is difficult, but achievable, quest at the LHC
  - It will require a targeted analysis
  - The B_s is a significant background that will emphasize good mass resolution

The quest continues

Modes of Collaboration

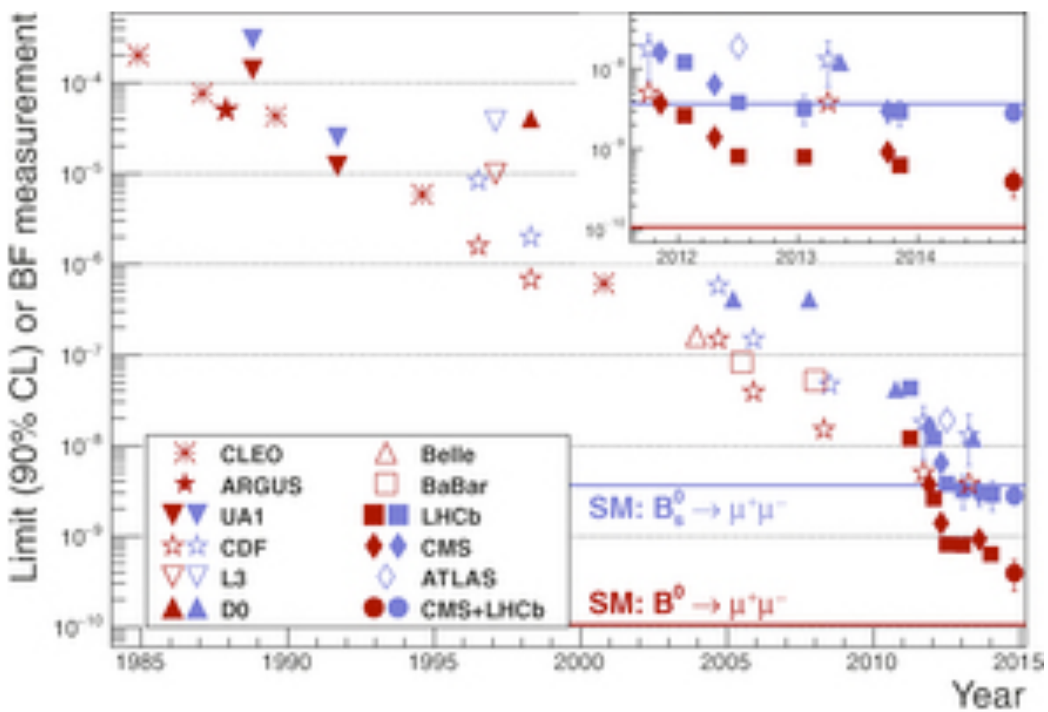
There are various ways that two disparate groups of people with different approaches to a problem can behave. I am pleased to report that our (CMS') interactions with LHCb were friendly, constructive and "harmonious". And we each learned from the other. So thanks to LHCb!

Backup Slides


THE LARGE HADRON COLLIDER
CMS
Mt. Jura
LHCb
μμ
27 km
CMS and LHCb catch $B_s \to \mu^+\mu^-$ in Rundown
Geneva airport
ALICE
CERN main campus

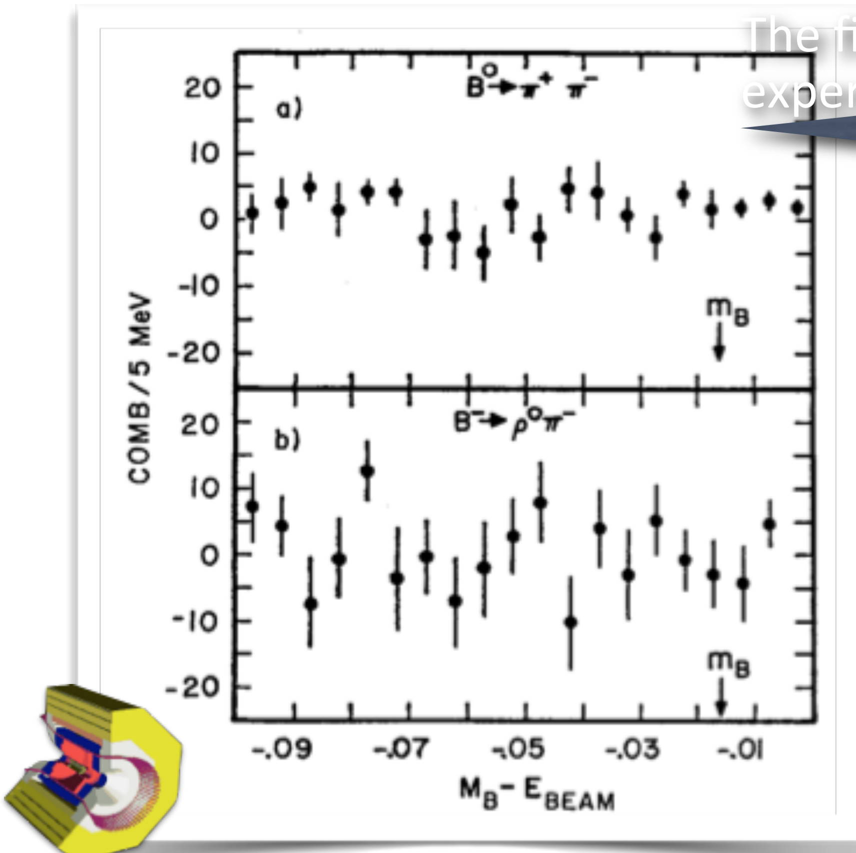
Historical Perspective – 25 years

- CMS and LHCb are only the most recent collaborations to have searched for and $B^0\ \mu^+\ \mu^-$ decays. Over three decades, a total of eleven collaborations have taken part in this search. This plot gathers the results from CLEO, ARGUS, UA1, CDF, L3, DØ, Belle, Babar, LHCb, CMS, and ATLAS.



HISTORY OF MEASUREMENTS

(Old CLEO cannot distinguish π/μ)

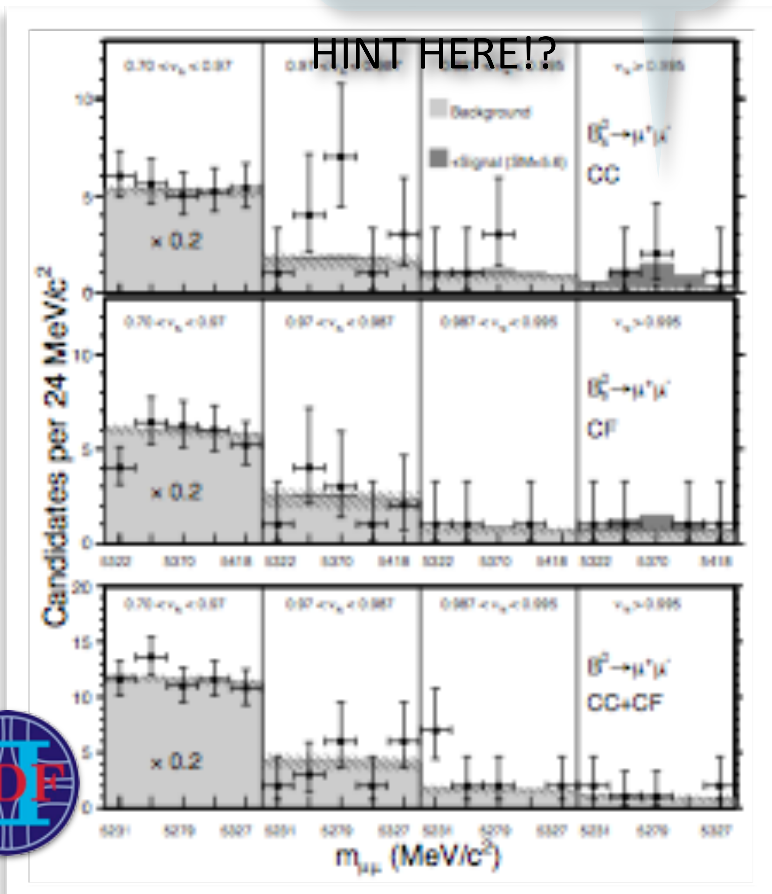


The first limit (**<0.05%**) was set by the CLEO experiment in 1984 w/ 42K **B_d** mesons.

- If one wants to measure a decay rate of $O(10^{-9})$ with a 20% precision, one needs to produce **10^{10}~10^{11}** B mesons, even with a 100% efficient detector.
- This is very difficult until recent colliders, ie. the LHC.
- Only upper limits can be calculated if seeing no signal.

Ref: CLEO Coll., PRD 30, 2279 (1984)

THE LAST HINT BEFORE LHC?



Ref: CDF Coll., PRL 107, 191801 (2011)

- The last CDF result shows a kind of hint that the signal is much stronger than the prediction from the model:

$$B(B_s \to \mu^+\mu^-) = (1.8\,^{+1.1}_{-0.9}) \times 10^{-8}$$

- The observed hint is **5 times** larger than the SM but with low statistical precision!

  Hint of new physics?

  Just a fluctuation?

- This is the end of Tevtron experiments, they cannot do better anymore! **Time for LHC!**

Inputs to Theory Calculation

TABLE I. Numerical inputs.

Parameter	Value	Unit	Ref.
G_F	1.166379×10^{-5}	GeV^{-2}	[13]
$\alpha_s^{(5)}(M_Z)$	0.1184(7)		[13]
$\alpha_{em}^{(5)}(M_Z)$	1/127.944(14)		[13]
$\Delta\alpha_{em,\text{hadr}}^{(5)}(M_Z)$	0.02772(10)		[13]
M_Z	91.1876(21)	GeV	[13]
M_t	173.1(9)	GeV	[13]
M_H	125.9(4)	GeV	[13]
M_{B_s}	5366.77(24)	MeV	[13]
M_{B_d}	5279.58(17)	MeV	[13]
f_{B_s}	227.7(4.5)	MeV	[14]
f_{B_d}	190.5(4.2)	MeV	[14]
$1/\Gamma_H^s$	1.615(21)	ps	[15]
$2/(\Gamma_H^d + \Gamma_L^d)$	1.519(7)	ps	[15]
$\lvert V_{cb}\rvert$	0.0424(9)		[16]
$\lvert V_{tb}^{\star} V_{ts}/V_{cb}\rvert$	0.980(1)		[17,18]
$\lvert V_{tb}^{\star} V_{td}\rvert$	0.0088(3)		[17,18]

Wilson Coefficients

$$C_A(\mu_b) = 0.4802\, R_t^{1.52}\, R_\alpha^{-0.09} - 0.0112\, R_t^{0.89}\, R_\alpha^{-0.09} = 0.4690\, R_t^{1.53}\, R_\alpha^{-0.09}, \quad (4)$$

$$C_A(\mu_b) = 0.4802 \tilde{R}_t^{1.50}\, R_\alpha^{0.015} - 0.0112\, \tilde{R}_t^{0.86}\, R_\alpha^{-0.031} = 0.4690\, \tilde{R}_t^{1.51}\, R_\alpha^{0.016}, \quad (5)$$

where $R_\alpha = \alpha_s(M_Z)/0.1184$, $R_t = M_t/(173.1\ \text{GeV})$ and $\tilde{R}_t = m_t/(163.5\ \text{GeV})$. The fits are accurate to better than

Bobeth et al.

Theory Results

B($B_s \to \mu^+\mu^-$) = (3.65 ± 0.23) × 10^{-9}
B($B_d \to \mu^+\mu^-$) = (1.06 ± 0.09) × 10^{-10}

$$\overline{\mathcal{B}}_{se} \times 10^{14} = (8.54 \pm 0.13) R_{t\alpha} R_s = 8.54 \pm 0.55,$$

$$\overline{\mathcal{B}}_{s\tau} \times 10^{7} = (7.73 \pm 0.12) R_{t\alpha} R_s = 7.73 \pm 0.49,$$

$$\overline{\mathcal{B}}_{de} \times 10^{15} = (2.48 \pm 0.04) R_{t\alpha} R_d = 2.48 \pm 0.21,$$

$$\overline{\mathcal{B}}_{d\mu} \times 10^{10} = (1.06 \pm 0.02) R_{t\alpha} R_d = 1.06 \pm 0.09,$$

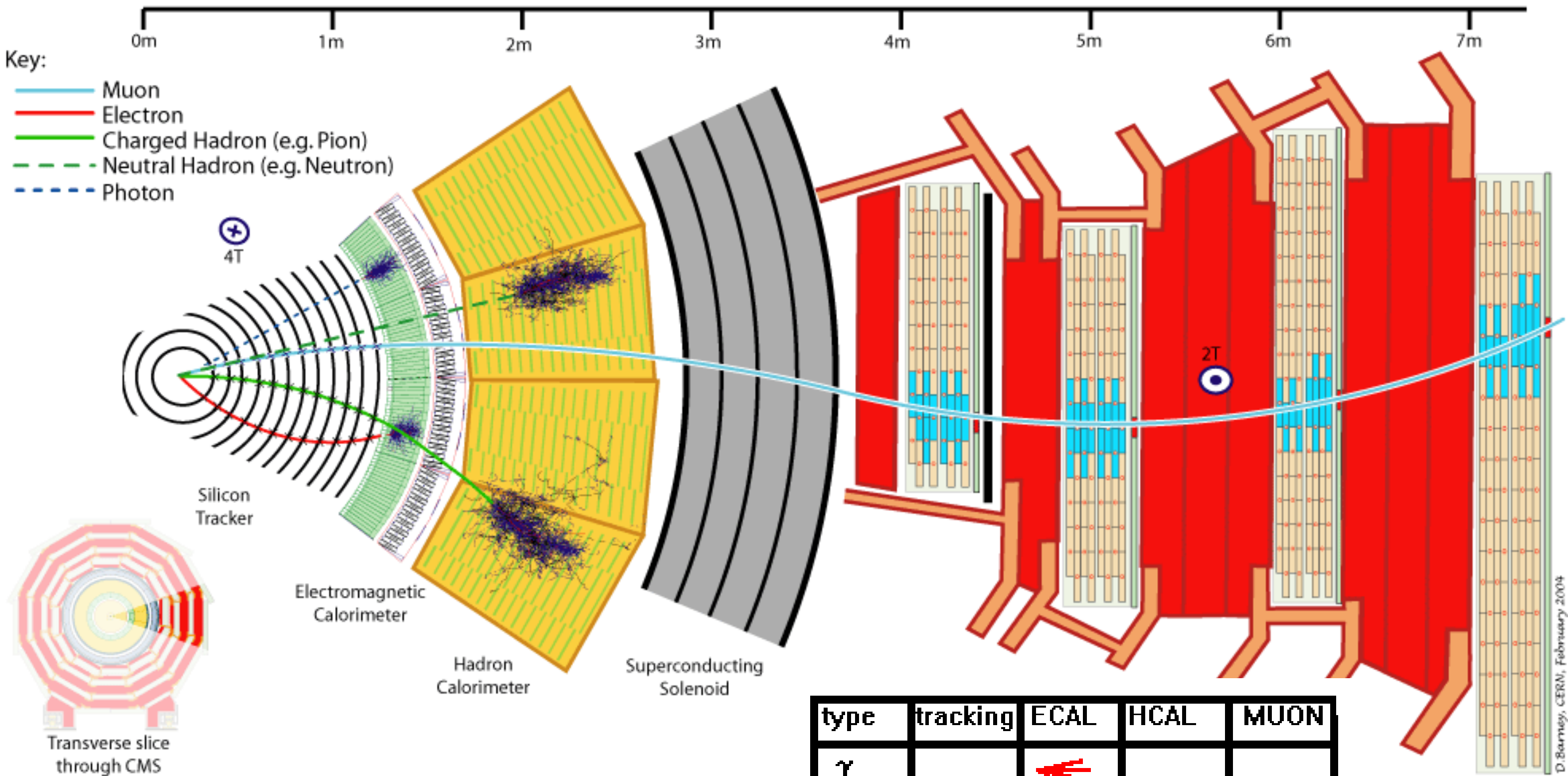
$$\overline{\mathcal{B}}_{d\tau} \times 10^{8} = (2.22 \pm 0.04) R_{t\alpha} R_d = 2.22 \pm 0.19,$$

TABLE II. Relative uncertainties from various sources in $\overline{\mathcal{B}}_{s\ell}$ and $\overline{\mathcal{B}}_{d\ell}$. In the last column they are added in quadrature.

	f_{B_q}	CKM	τ_H^q	M_t	α_s	Other parameters	Nonparametric	$\sum$
$\overline{\mathcal{B}}_{s\ell}$	4.0%	4.3%	1.3%	1.6%	0.1%	< 0.1%	1.5%	6.4%
$\overline{\mathcal{B}}_{d\ell}$	4.5%	6.9%	0.5%	1.6%	0.1%	< 0.1%	1.5%	8.5%

Bobeth et al.

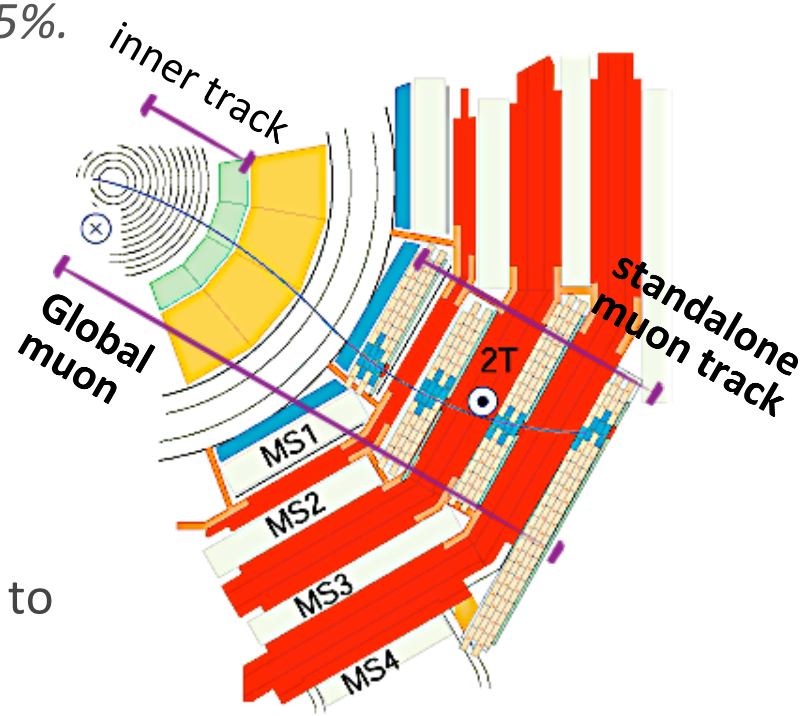
CMS Slice



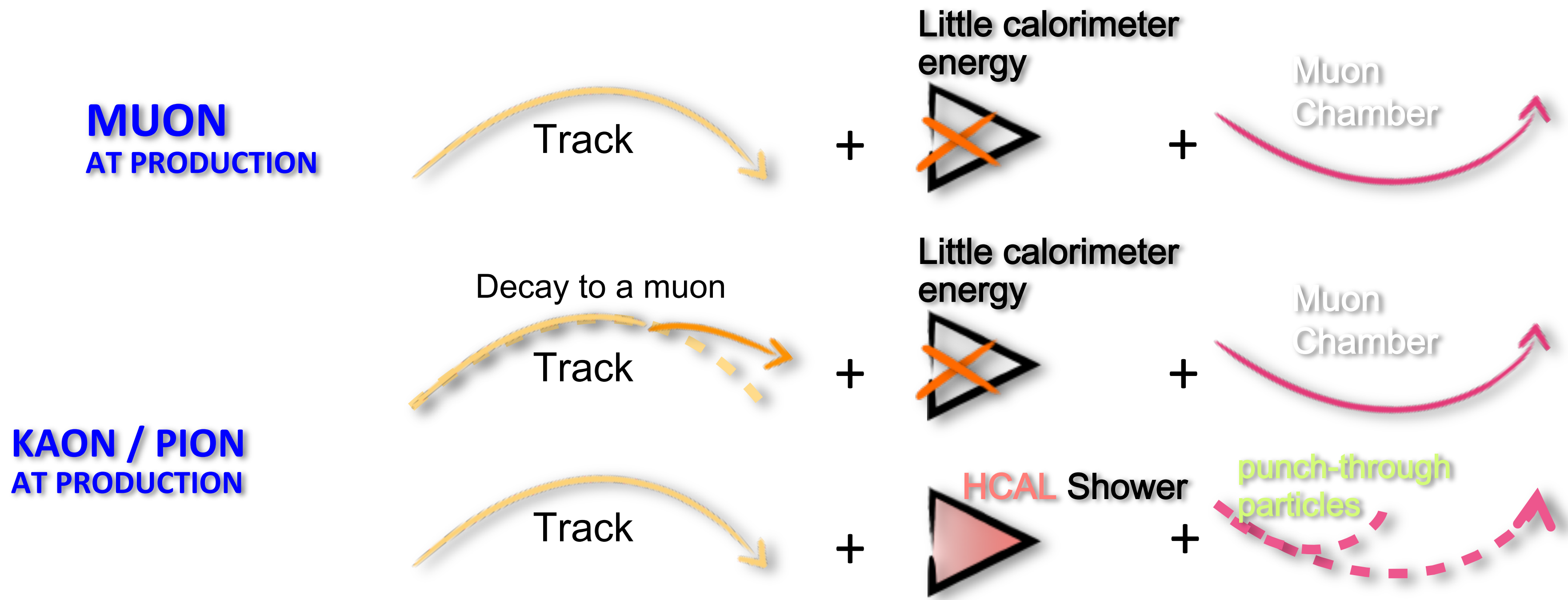
type	tracking	ECAL	HCAL	MUON
γ				
e				
μ				
Jet				
Et miss				

CMS MUON RECONSTRUCTION

- **Muon system:**
  - 3 different devices installed with a large coverage.
    - 2 redundant systems in barrel and 2 in ebdcap
  - Good dimuon mass resolution: *0.6-1.5%.*
- **Reconstruction algorithms:**
  - **standalone muon**:
    reconstructed in muon system only
  - **global muon**:
    standalone muon ⇒ inner track
  - **tracker muon**:
    inner track ⇒ muon system
- **Muon misidentification**
  - An essential piece of the analysis due to the **fake muons from kaons/pions.**



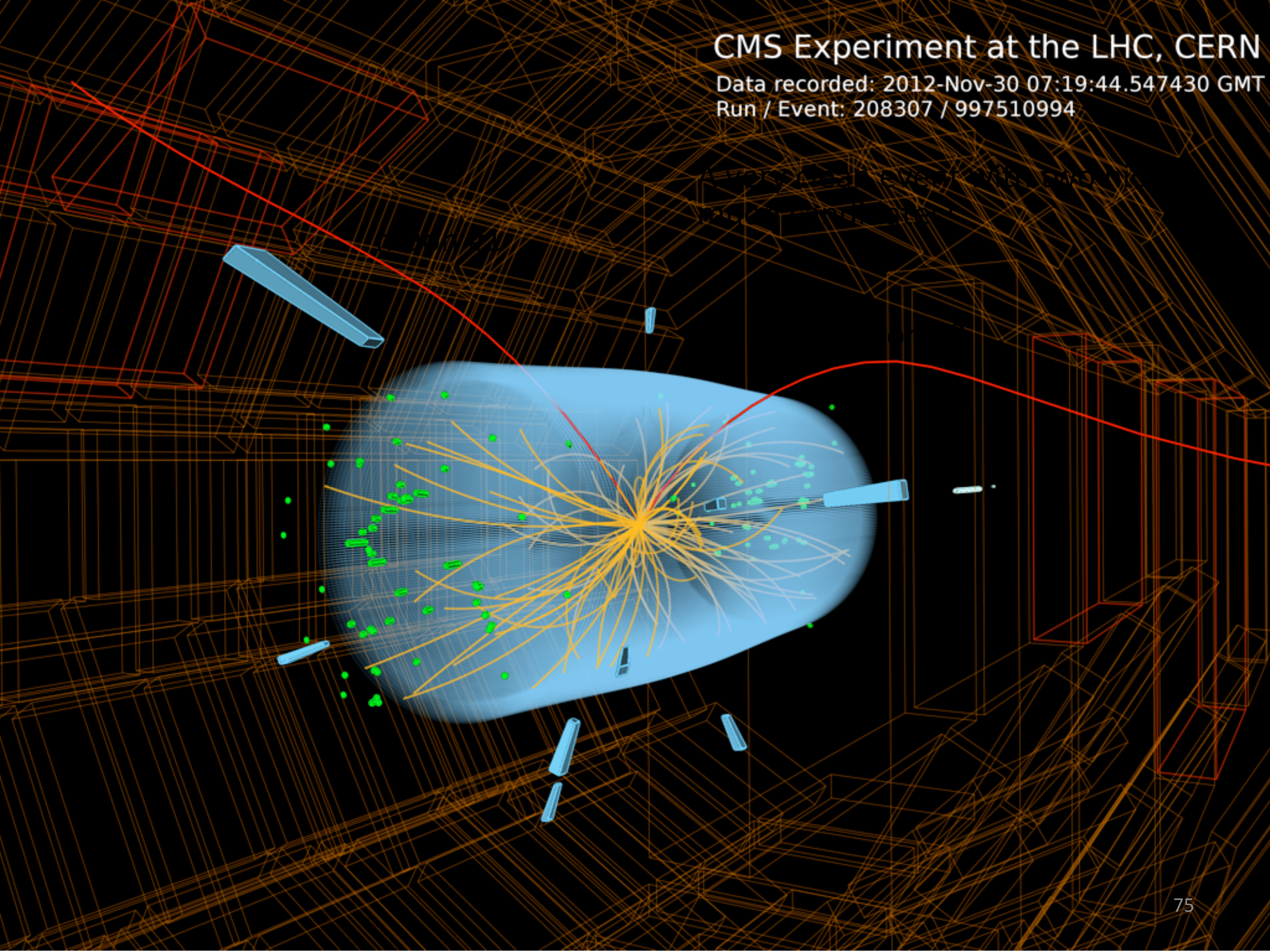
Muon Identification



1) Tracker hits association
2) Muon chamber segments
3) Adopting active "kink" finder
4) Track fitting quality
5) Kinematical distributions

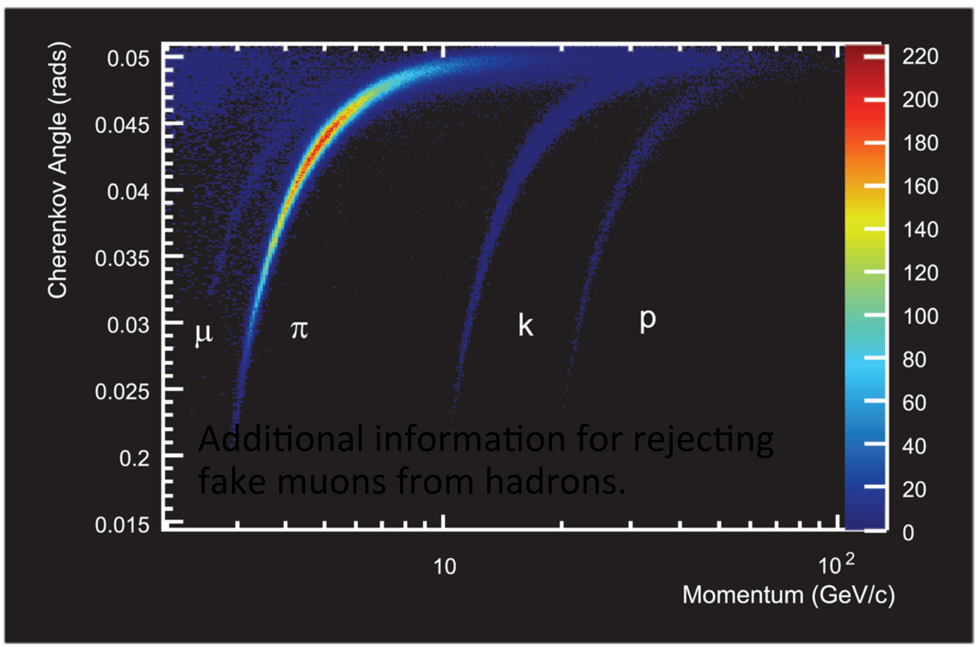
Inject into **BDT** **and reject poor muons.**

Powerful fake-muon rejection rate:
$\varepsilon(\pi \Rightarrow \mu) < 0.15\%$
$\varepsilon(K \Rightarrow \mu) < 0.20\%$
$\varepsilon(p \Rightarrow \mu) < 0.10\%$


CMS Experiment at the LHC, CERN
Data recorded: 2012-Nov-30 07:19:44.547430 GMT
Run / Event: 208307 / 997510994

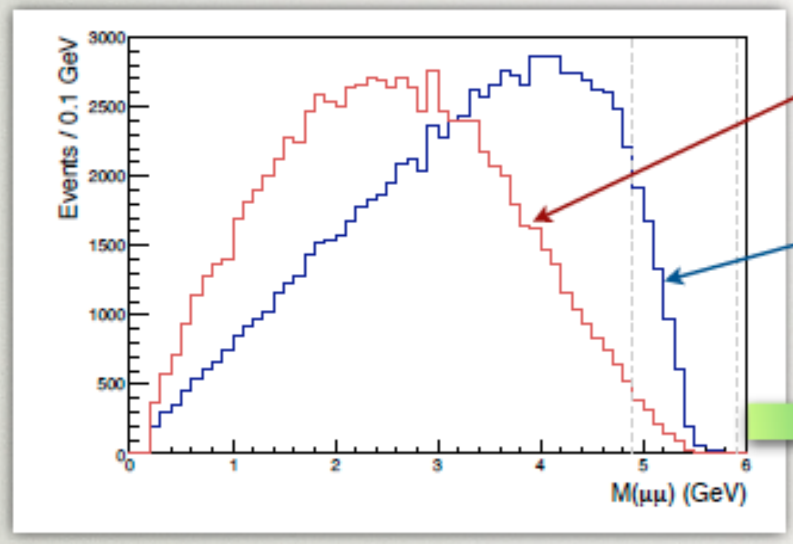
LHCb PARTICLE ID

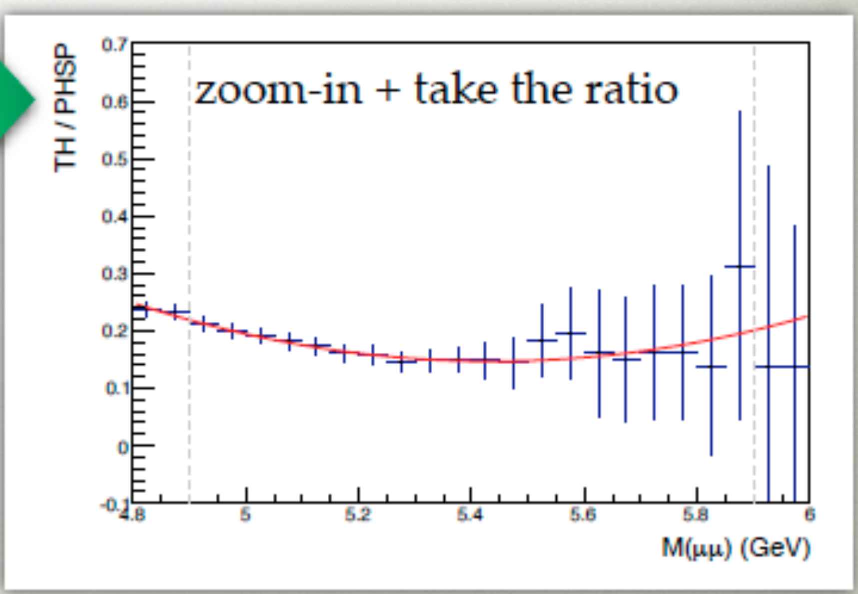
- LHCb has a dedicated (active) particle identification device: RICH (Ring Imaging Cherenkov) detector.
- **A global particle ID likelihood** is constructed based on the information from the **RICH** detectors, calorimeters (**CALO**), and **MUON** system.



Powerful muon identification with high (~98%) efficiency:
Based on muon chambers information + the global PID likelihood:
$\varepsilon(\pi \rightarrow \mu)$~0.6%
$\varepsilon(K \rightarrow \mu)$~0.4%
$\varepsilon(p \rightarrow \mu)$~0.3%

$\Lambda_b \to p\mu\nu$ MODEL





- The average contribution in the signal region should be scaled by ~0.179 (from model) and by 0.76 (from BF), resulting the factor of ~7.5 reduction.
- Inject this correction (the curve) into the PDF construction.

$$\langle \Gamma(B_s(t) \Rightarrow f)\rangle = \Gamma(B_0^s(t) \Rightarrow f) + \Gamma(\overline{B}_0^s(t) \Rightarrow f)$$

$$= R_H^f e^{-\Gamma_H^{(s)} t} + R_L^f e^{-\Gamma_L^{(s)} t}$$

$$\langle \Gamma(B_s(t) \Rightarrow f) \rangle = (R_H^f + R_L^f) e^{-\Gamma_s t}$$

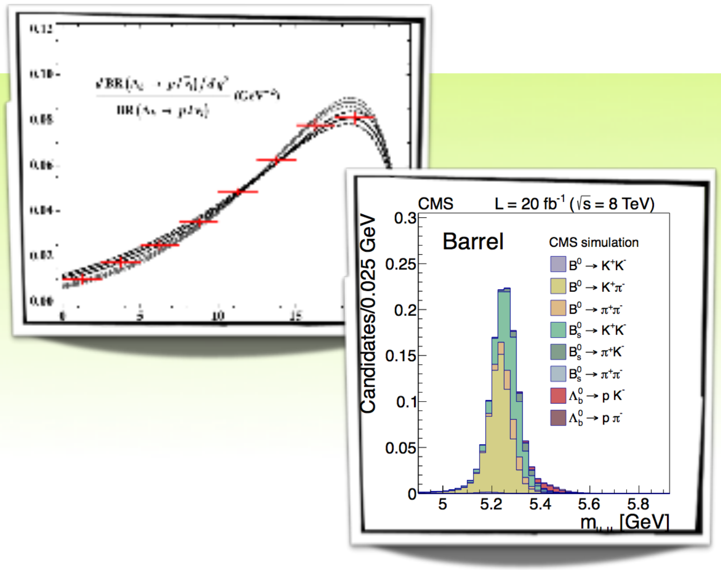
$$\times \left[\cosh\left(\frac{y_s t}{\tau_{B_s}}\right) + A_{\Delta\Gamma}^f \sinh\left(\frac{y_s t}{\tau_{B_s}}\right) \right]$$

$$\text{where } y_s = \frac{\Delta \Gamma_s}{2\Gamma_s}, \; A_{\Delta\Gamma}^f = \frac{R_H^f - R_L^f}{R_H^f + R_L^f},$$

$$\text{and } \tau_{B_s}^{-1} = \Gamma_{(s)} = \frac{\left[\Gamma_L^{(s)} + \Gamma_H^{(s)}\right]}{2}$$

Analysis Synchronization

The analysis external inputs have been synchronized as well.



Common theoretical models and branching fractions are introduced for the backgrounds from B decays.

$$\Gamma(B_s \to \mu\mu) \propto e^{-t/\tau_{\mathrm{gen}}}$$

Predicted lifetime distribution

$$e^{-t/\tau_{B_s}} \left[\cosh\left(\frac{y_s t}{\tau_{B_s}}\right) + A_{\Delta\Gamma} \sinh\left(\frac{y_s t}{\tau_{B_s}}\right)\right]$$

$$\tau_{B_s} = 1.516 \pm 0.011 \text{ ps}$$

Common (corrected) B_s lifetime distribution is also introduced in the simulated samples.

Common Branching Fractions

Table 2: Summary of the branching fraction numbers used as input by the two analyses and corresponding references.

Channel	Branching fraction	Ref.
$B^+ \to J/\psi K^+$	$(1.028 \pm 0.031) \cdot 10^{-3} \times (5.93 \pm 0.06) \cdot 10^{-2} = (6.10 \pm 0.19) \cdot 10^{-5}$	[5]
$B^0 \to K^+\pi^-$	$(1.96 \pm 0.05) \cdot 10^{-5}$	[5]
$B_s^0 \to K^-\mu^+\nu$	$(1.27 \pm 0.59) \cdot 10^{-4}$	[25]
$B^0 \to \pi^-\mu^+\nu$	$(1.44 \pm 0.05) \cdot 10^{-4}$	[5]
$\Lambda_b \to p\mu^-\bar{\nu}$	$(4.94 \pm 2.19) \cdot 10^{-4}$	[14] and updates
$B^{+,0} \to \pi^{+,0}\mu^+\mu^-$	$(2.3 \pm 0.6) \cdot 10^{-8} \times (1.47 \pm 0.20) = (3.38 \pm 0.99) \cdot 10^{-8}$	[25, 26]