Observation of the rare $B_s^0 \to \mu^+\mu^-$ decay

The combined analysis of CMS and LHCb data

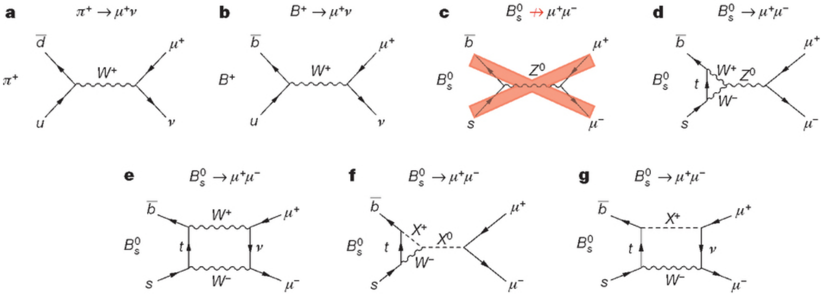
Luke Pritchett

April 22, 2016

Table of Contents

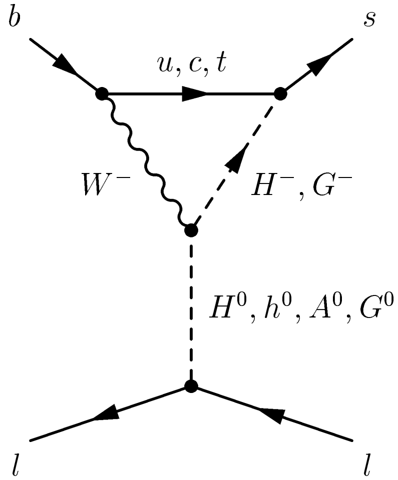
Flavor physics and B decays

- $B^0_{(s)} \to \mu^+\mu^-$ is a flavor-changing neutral process.
- In the SM, FCNCs are suppressed by the GIM mechanism.
- The Standard Model prediction is
  $Br(B_s^0 \to \mu\mu) = (3.66 \pm 0.23) \times 10^{-9}$ and
  $Br(B^0 \to \mu\mu) = (1.06 \pm 0.09) \times 10^{-10}$



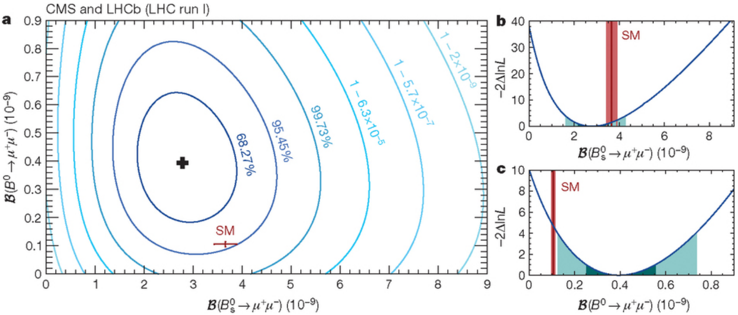
BSM theories

- BSM theories often predict flavor changing physics, especially in the third generation.
- E.g., 2HDM, SUSY, topcolor, etc.
- FCNC measurements are a powerful constraint on BSM physics.



Results Overview

- Observation of decay $B_s^0 \to \mu^+\mu^-$ with significance $> 6\sigma$
- Evidence for decay $B^0 \to \mu^+\mu^-$ with significance approximately 3σ.
- All measured branching fractions and ratios are compatible with the SM within 2.3σ.



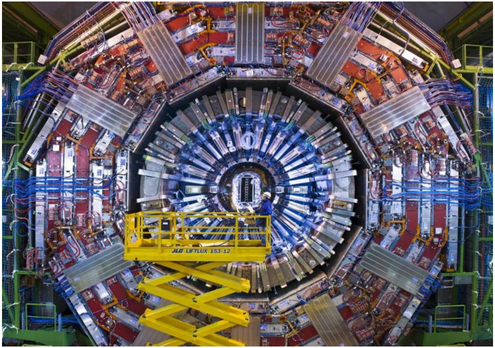
Basic Search Strategy

1. Tag $B^0_{(s)} \to \mu^+\mu^-$ events.
2. Distinguish signal versus background
3. Distinguish B^0 from B^0_s
4. Normalize signal to get branching ratios
5. Combine statistics

Table of Contents

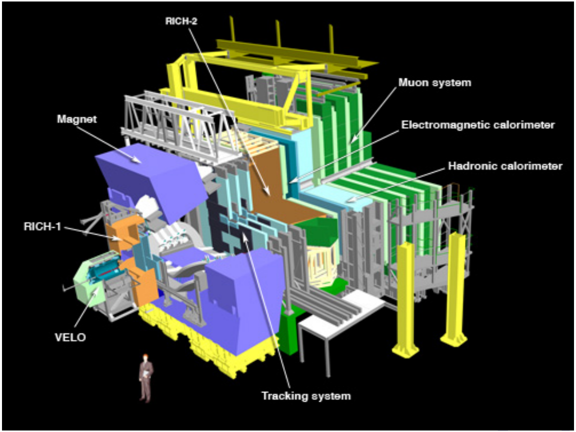
CMS Detector

- CMS is great, but we won't cover its instrumentation here.



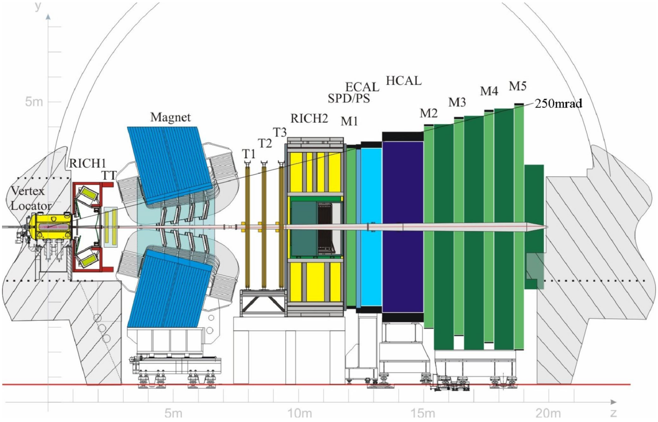
LHCb Experiment

- Occupies pit 8, previous home of DELPHI.
- Purpose: to study precision flavor physics, CP violation, matter / anti-matter asymmetry.
- b physics is a great portal for this mission.



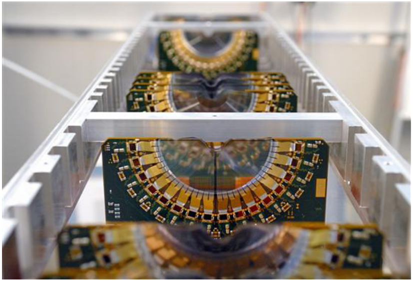
LHCb Detector

- Designed for precision b physics, distinguishing B mesons.
- Forward detector to capture B mesons.
- Excellent vertex and momentum resolution.



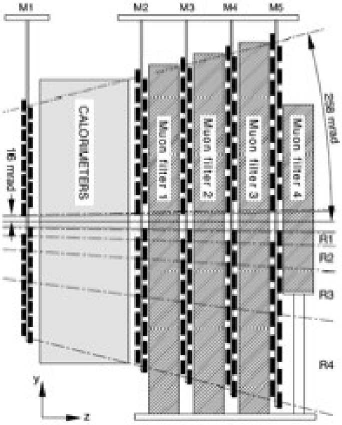
LHCb Detector: VELO

- Vertex Locator system built around the interaction point.
- Reproduces tracks in an $r - \phi$ coordinate system.
- Resolution is $\sim 8\,\mu m$.



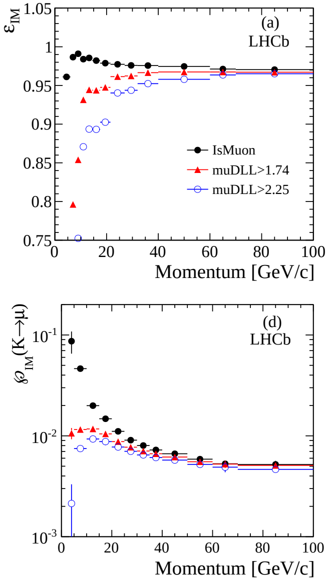
LHCb Detector: Muon system

- One section before the calorimeters.
  - Triple-GEM gas detector
- Four sections behind the calorimeters.
  - Multiwire Proportional Chambers (MWPCs)
  - Designed for 99% efficiency.



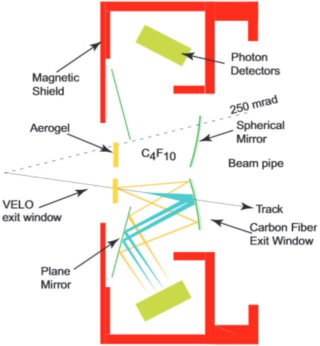
LHCb Detector: Muon selection

1. Loose binary selection based on penetration.
2. Log likelihood cuts using tracking information.
3. Combined likelihood to further discriminate pions versus muons.



LHCb Detector: Other components

- Trackers
  - Trigger tracker, silicon microstrip detector.
  - Inner tracker, silicon microstrip behind magnet.
  - Outer tracker, Kapton / Al straws.
- Photon detectors
  - For extra π vs. K discrimination
  - Ring Imaging Cherenkov counters (RICH 1 and RICH 2).
    - Specialized Hybrid Photon Detectors



LHCb Detector: Other components

- Calorimeters
  - Reconstruction of π^0 and prompt photons is essential for flavor tagging and B-meson decays.
  - ECAL
    - Electron detection must reject charged π's and π^0's.
    - Uses a preshower detector before for charged, and a scintillator pad for π^0.
    - Scintillator / lead structure.
    - Energy resolution $\sigma_E/E = 10\%/\sqrt{E}$ (R in GeV).
  - HCAL
    - Iron and scintillating tiles.

LHCb: Triggering

- L0 Trigger
  - Hardware triggering
  - Reduces 40 MHz crossing rate to 1 MHz readout.
  - Reconstructs highest E_T hadron, electron and photon clusters, two highest p_T muons.
  - VELO estimates the number of primary pp interactions in each bunch crossing.
- HLT
  - Executed asynchronously on a processor farm.
  - Reduces event rate from 1 MHz to 2 kHz.

Table of Contents

Backgrounds: Combinatorial background

- Muons from other processes, mainly semileptonic decays of other B mesons.
- Evaluated by extrapolating data from nearby mass sidebands, [4.9 GeV, m_{B^0} - (60 MeV)] and [$m_{B^0_s}$+(60 MeV), 60 GeV].
- Modeled with a first-degree polynomial.
- Can be reduced to a certain extent via tracking and vertex analysis.

Backgrounds: Muon misidentification

	Yield in full BDT range	Fraction with BDT > 0.7 [%]
$B^0_{(s)} \to h^+ h'^-$	15 ± 1	28
$B^0 \to \pi^- \mu^+ \nu_\mu$	115 ± 6	15
$B^0_s \to K^- \mu^+ \nu_\mu$	10 ± 4	21
$B^{0(+)} \to \pi^{0(+)} \mu^+ \mu^-$	28 ± 8	15
$\Lambda^0_b \to p \mu^- \overline{\nu}_\mu$	70 ± 30	11

- Pions or Kaons from B decays misidentified as muons.
  - $B^0 \to \pi^- \mu^+ \nu,\ B^0_s \to K^- \mu^+ \nu,\ \Lambda^0_b \to p \mu^- \overline{\nu}$
- Invisible pions
  - $B^+ \to \pi^+ \mu^+ \mu^-,\ B^0 \to \pi^0 \mu^+ \mu^-$
- These have lower $m_{\mu\mu}$ invariant masses than signal, except:
- $B^0_{(s)} \to h^+ h'^-$, where $h^{(')} = \pi$ or K.
  - Estimated by normalizing to the observed $B^+ \to J/\psi K^+$ yield.
  - $B^0_s \to \pi^+ K^-$ has known branching fraction $(1.91 \pm 0.31) \times 10^{-5}$.

Event selection

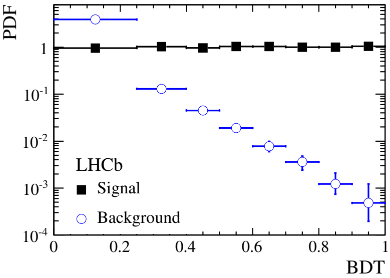
- Signal candidates are chosen primarily by muon and dimuon triggers.
- Important preliminary variables are p_T cuts and vertex properties.
- LHCb: $0.25 < p_T < 40$ GeV and $p < 500$ GeV. CMS: $p_T > 4.0$ GeV for individual muons, $4.8 < m_{\mu\mu} < 6.0$ GeV.
- Muon tracks should form a secondary vertex, displaced from a primary vertex.
  - Time-of-flight significance > 15 between SV and most significant PV.
  - Allow B candidates with $p_T > 0.5$ GeV, decay time less than $9\times$lifetime.
- Average trigger efficiency (for older CMS experiment) for events in signal samples from MC is $(39 - 85)\% \pm (3 - 6)\%$.

Boosted Decision Tree analysis

- BDT analysis further reduces backgrounds, misIDs
  - CMS: Hadron-to-muon misID below 2.2×10^{-3} for π,K, and p, as determined from well-identified hadrons in data.
- Both BDTs use 12 variables:
  - B candidate decay time, impact parameter, and p_T
  - Minimum χ^2_{IP} of the two muons with respect to any PV.
  - Closest approach of the two muons.
  - A 3D pointing angle
  - Flight length significance between SV and PV.
  - A few others

Boosted Decision Tree training

- BDTs were trained on simulated signal. For background LHCb used simulations of $b\bar{b} \to \mu^+\mu^- + X$, CMS used the mass sidebands.
    - Data background split into three sets, BDT training is independent of its application.



- 20 BDT discriminant bins, 8 from $LHCb$ and 12 from CMS. Bins have roughly equal expected signal yield.
- BDT dependence on $m_{\mu\mu}$ is linear and small.

Table of Contents

Normalizing the Branching Ratios

- Count the number of $B^+ \to J/\psi K^+$ decays
- Use measured branching fraction to count B^+ production.
- Assume B^0 and B^+ are produced at the same rate.
- Use measured ratio of B^+ to B_s^0.
- This procedure introduces uncertainty into the data, correlated between LHCb and CMS. The statistical analysis recognizes this.

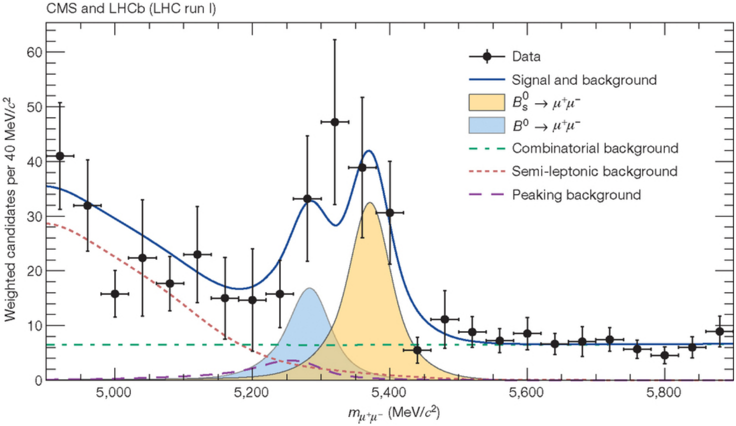
Normalizing the Branching Ratios

$$Br(B_s^0 \to \mu^+\mu^-) = \frac{N_{B_s^0 \to \mu^+\mu^-}}{N_{\text{norm}}} \times \frac{f_d}{f_s} \times \frac{\epsilon_{\text{norm}}}{\epsilon_{B_s^0 \to \mu^+\mu^-}} \times Br_{\text{norm}}$$

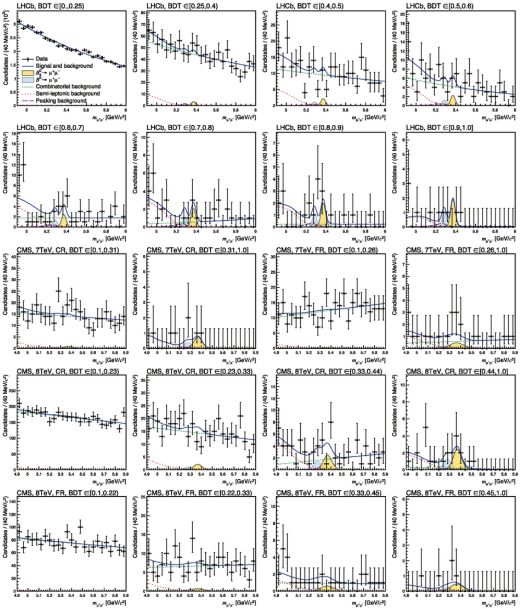
- f_d/f_s is the ratio of probablilities for a b quark to hadronize into a B^0 versus a B_s^0.
- Theory says that the probablities for B^+ and B^0 are the same, also checked on the data.
- $f_d/f_s = 3.86 \pm 0.22$ as measured by LHCb previously, confirmed within error by CMS.
- ϵ's are signal reconstruction efficiencies, measured from simulation and data.

Fits and branching ratios

- Unbinned extended maximum likelihood fit of signal function to the combined data with all its discriminants.
- Where possible, nuisance parameters are constrained to their known values with Gaussian distributions.

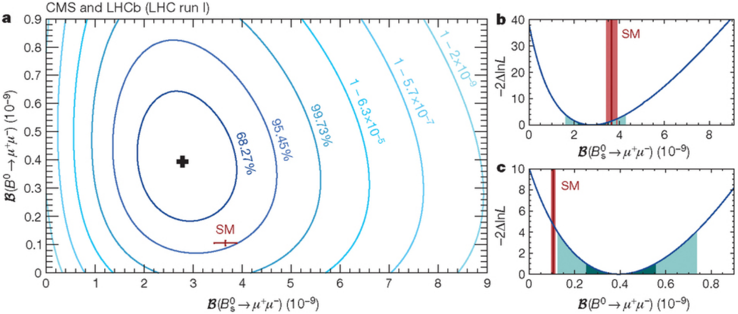


Fits and branching ratios



Fits and branching ratios

- The confidence intervals are created with the Feldman-Cousins procedure, based on maximizing log-likelihood ratios $-2\log\left(P(\text{data}|Br)/P(\text{data}|Br^*)\right)$.
- Statistical uncertainty is obtained by repeating the fit with all nuisance parameters to their fitted values.



Fits and branching ratios

- ATLAS just published their results for $B_s^0 \to \mu\mu$ decays (14 April, 2016):