Search for $B_s^0 \rightarrow \mu^+\mu^-$ and $B^0 \rightarrow \mu^+\mu^-$ in CMS (2011 dataset)

Urs Langenegger
(PSI)

2012/05/30

▷ Introduction
▷ Candidate(s) selection
▷ Data/MC validation
▷ Pileup independence
▷ Search Analysis
▷ Results
Motivation: Search for New Physics

- **Decays highly suppressed in Standard Model** (Buras 2010)
  - effective FCNC, helicity suppression
  - SM expectation:
    \[
    \begin{align*}
    \mathcal{B}(B_{s}^{0} \rightarrow \mu^{+}\mu^{-}) &= (3.2 \pm 0.2) \times 10^{-9} \\
    \mathcal{B}(B^{0} \rightarrow \mu^{+}\mu^{-}) &= (1.0 \pm 0.1) \times 10^{-10}
    \end{align*}
    \]
  - Cabibbo-enhancement \(|V_{ts}| > |V_{td}|\) of \(B_{s}^{0} \rightarrow \mu^{+}\mu^{-}\) over \(B^{0} \rightarrow \mu^{+}\mu^{-}\)
    - only in MFV models

- **Sensitivity to new physics**
  - 2HDM: \(\mathcal{B} \propto (\tan \beta)^{4}m_{H^{+}}\); MSSM: \(\mathcal{B} \propto (\tan \beta)^{6}\)
    - sensitivity to extended Higgs boson sectors
    - Constraints on parameter regions
  - \(B_{s}^{0} \rightarrow \mu^{+}\mu^{-}\) (and \(B^{0} \rightarrow \mu^{+}\mu^{-}\)) considered as golden channel(s)
    - high sensitivity to new physics
    - (very) small theoretical uncertainties
    - comparable in sensitivity to \(\mu \rightarrow e\gamma, B \rightarrow X\nu\bar{\nu}\)

Urs Langenegger
Search for \(B_{s}^{0} \rightarrow \mu^{+}\mu^{-}\) and \(B^{0} \rightarrow \mu^{+}\mu^{-}\) in CMS (2012/05/30)
State of the art (upper limit at 95%CL)

- At the Tevatron

<table>
<thead>
<tr>
<th>Upper limit</th>
<th>$B_s^0 \rightarrow \mu^+\mu^-$</th>
<th>$B^0 \rightarrow \mu^+\mu^-$</th>
</tr>
</thead>
<tbody>
<tr>
<td>D0$^1$</td>
<td>$5.1(4.0) \times 10^{-8}$</td>
<td>n/a</td>
</tr>
<tr>
<td>CDF$^2$</td>
<td>$4.0(1.5) \times 10^{-8}$</td>
<td>$6.0 \times 10^{-9}$</td>
</tr>
</tbody>
</table>

$^1$) 6.1 fb$^{-1}$, PL, B693, 539

$^2$) 7 fb$^{-1}$, PRL, 107, 191801

- At the LHC:

<table>
<thead>
<tr>
<th>Upper limit</th>
<th>$B_s^0 \rightarrow \mu^+\mu^-$</th>
<th>$B^0 \rightarrow \mu^+\mu^-$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMS$^3$</td>
<td>$1.9(1.8) \times 10^{-8}$</td>
<td>$3.6 \times 10^{-9}$</td>
</tr>
<tr>
<td>LHCb$^4$</td>
<td>$1.4(1.3) \times 10^{-8}$</td>
<td>$3.2 \times 10^{-9}$</td>
</tr>
<tr>
<td>ATLAS$^5$</td>
<td>$2.2(2.3) \times 10^{-8}$</td>
<td>n/a</td>
</tr>
<tr>
<td>CMS$^6$</td>
<td>$7.7(8.4) \times 10^{-9}$</td>
<td>$1.8 \times 10^{-9}$</td>
</tr>
<tr>
<td>LHCb$^7$</td>
<td>$4.5(7.2) \times 10^{-9}$</td>
<td>$1.0 \times 10^{-9}$</td>
</tr>
</tbody>
</table>

$^3$) 1.1 fb$^{-1}$, PRL, 107, 191802
$^4$) 0.4 fb$^{-1}$, PL, B708, 55
$^5$) 2.4 fb$^{-1}$, arXiv:1204.0735 [hep-ex]
$^6$) 5 fb$^{-1}$, JHEP 04(2012), 033
$^7$) 1 fb$^{-1}$, arXiv:1203.4493 [hep-ex]

- CDF$^2$ also has $B(B_s^0 \rightarrow \mu^+\mu^-) = (1.8^{+1.1}_{-0.9}) \times 10^{-8}$
Analysis overview

• **Signal** \( B_s^0 \rightarrow \mu^+ \mu^- \)
  - two muons from one decay vertex
  - well reconstructed secondary vertex
  - momentum aligned with flight direction
  - mass around \( m_{B_s^0} \)

• **Background**
  - two semileptonic \((B)\) decays (gluon splitting)
  - one semileptonic \((B)\) decay and one misidentified hadron
  - rare single \(B\) decays
    - peaking, e.g. \( B_s^0 \rightarrow K^+ K^- \)
    - non-peaking, e.g. \( B_s^0 \rightarrow K^- \mu^+ \nu \)
Methodology

- **Measurement of** $B_s^0 \rightarrow \mu^+\mu^-$ **relative to normalization channel:**
  - similar trigger and selection to reduce systematic uncertainties

$$B(B_s^0 \rightarrow \mu^+\mu^-; 95\%C.L.) = \frac{N(n_{obs}, n_B, n_S; 95\%C.L.)}{\varepsilon_{B_s^0} N_{B_s^0}} = \frac{N(n_{obs}, n_B, n_S)}{\varepsilon_{B_s^0} \mathcal{L} \sigma(pp \rightarrow B_s^0)}$$

$$= \frac{N(n_{obs}, n_B, n_S)}{N(B^\pm \rightarrow J/\psi K^\mp)} \frac{A_{B^+}}{A_{B_s^0}} \frac{\varepsilon^{\text{ana}}_{B^+} \varepsilon_{B_s^0}^{\mu} \varepsilon^{\text{trig}}_{B_s^0}}{f_u} B(B^+ \rightarrow J/\psi [\mu^+\mu^-] K)$$

- **Calibration/validation of MC:**
  - $B^\pm \rightarrow J/\psi K^\mp$ normalization with high statistics
  - $B_s^0 \rightarrow J/\psi \phi$ $B_s^0$ signal MC ($p_\perp$, isolation, . . . )

- **Analysis in two channels**
  - **barrel** (both muons $|\eta| < 1.4$):
    - better signal/background ratio
    - good mass resolution (36 MeV)
  - **endcap** (at least one muon with $|\eta| > 1.4$):
    - add more statistics [$\sigma(m) \approx 70$ MeV]

⇒ **Blind analysis**

Urs Langenegger  
Search for $B_s^0 \rightarrow \mu^+\mu^-$ and $B^0 \rightarrow \mu^+\mu^-$ in CMS (2012/05/30)
The CMS detector

- Design prioritization
  - lepton ID $\rightarrow$ muons
  - $b/\tau$ tagging $\rightarrow$ tracking
  - jets and $E_T$

<table>
<thead>
<tr>
<th>Component</th>
<th>Characteristics</th>
<th>Resolutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pixel</td>
<td>3/2 Si layers</td>
<td>$\delta z \approx 20 , \mu m$, $\delta \phi \approx 10 , \mu m$</td>
</tr>
<tr>
<td>Tracker</td>
<td>10/12 Si strips</td>
<td>$\delta (p_\perp)/p_\perp \approx 1%$</td>
</tr>
<tr>
<td>ECAL</td>
<td>PbWO$_4$</td>
<td>$\delta E/E \approx 3%/\sqrt{E} \oplus 0.5%$</td>
</tr>
<tr>
<td>HCAL (B)</td>
<td>Brass/Sc, $&gt; 7.2 \lambda$</td>
<td>$\delta E/E \approx 100\sqrt{E}%$</td>
</tr>
<tr>
<td>HCAL (F)</td>
<td>Fe/Quartz</td>
<td>$\delta (E_T) \approx 0.98\sqrt{\sum E_T}$</td>
</tr>
<tr>
<td>Magnet</td>
<td>3.8 T solenoid</td>
<td>$\delta (p_\perp)/p_\perp \approx 10%$ (STA)</td>
</tr>
<tr>
<td>Muons</td>
<td>DT/CSC + RPC</td>
<td></td>
</tr>
</tbody>
</table>

Weight 12'500 t
Length 21.6 m
Diameter 15 m
Magnetic field 3.8 T

Tracking resolution:
impact parameter $\approx 15 \, \mu m$
Muon reconstruction

- Large muon acceptance $|\eta| < 2.4$
  - drift tubes
  - cathode strip chambers
  - resistive plate chambers
- 3 muon reconstruction algorithms
  - standalone muon: in muon system (trigger ingredient)
  - global muon (‘GM’): outside-in standalone muon $\rightarrow$ to inner track
  - tracker muon (‘TM’): inside-out inner track $\rightarrow$ muon detector

Muon misidentification

$\varepsilon(\mu|\pi) \leq 0.1\%$

$\varepsilon(\mu|K) \leq 0.1\%$

$\varepsilon(\mu|p) \leq 0.05\%$

measured in data:

$D^* \rightarrow D^0 \pi^+_s \rightarrow K^- \pi^+ \pi^+_s$

$\Lambda \rightarrow p \pi^-$

CMS-PAS-MUO-10-002
Trigger

- Dimuon trigger
  - L1 (hardware) trigger
  - High-level trigger
    full tracking and vertexing
  - requirements tightened over time
- HLT $B^0_s \rightarrow \mu^+ \mu^-$
  - inv. mass $4.8 < m_{\mu^+ \mu^-} < 6.0$ GeV
  - dimuon vertex $\mathcal{P}(\chi^2, \text{dof}) > 0.5\%$
  - distance of closest approach $d_{ca} < 0.5$ cm
  - single muon $p_\perp > 4$ GeV, dimuon $p_\perp > 3.9(5.9)$ GeV in barrel (endcap)
- HLT $B^\pm \rightarrow J/\psi K^\pm$ and $B^0_s \rightarrow J/\psi \phi$
  - single muon $p_\perp > 4$ GeV, dimuon $p_\perp > 6.9$ GeV
  - distance of closest approach among muons $d_{ca} < 0.5$ cm
  - invariant dimuon mass $2.9 < m_{\mu^+ \mu^-} < 3.3$ GeV
  - pointing angle $\cos \alpha_{xy} > 0.9$ and dimuon vertex $\mathcal{P}(\chi^2/\text{dof}) > 15\%$
  - ‘displaced’ $J/\psi$: flight length significance $\ell/\sigma(\ell) > 3$
3D vertexing

- All silicon tracker
  - high granularity, low occupancy
  - very well described by MC simulation

- Pixel detector
  - $100 \times 150 \, \mu \text{m}^2$ pixel size
  - substantial charge sharing
  - excellent resolution

⇒ Essential in high-pileup environment!
Candidate selection
Two analyses

1. Search analysis $B \to \mu^+ \mu^-$ in two channels
   - **barrel** (both muons $|\eta| < 1.4$):
   - **endcap** ($\geq 1$ muon with $|\eta| > 1.4$):

2. Validation analysis in one channel
   
   $B^\pm \to J/\psi K^\pm$ and $B^0_s \to J/\psi \phi$ (and dimuons)

Overlays of data and MC simulation (selection summary on p. 20)
   - ‘all other’ selection criteria are applied

MC signal

data background in sidebands ($4.9 < m < 5.2$ GeV and $5.45 < m < 5.9$ GeV)
Signal selection: vertexing

• Choose one primary vertex
  ▶ longitudinal impact parameter (z position)
  ▶ refit without signal tracks

• Discriminating variables
  ▶ pointing angle $\alpha_{3D}$
  ▶ $B$ vertex fit quality $\chi^2/dof$
  ▶ flight length significance $\ell_{3d}/\sigma(\ell_{3d})$
  ▶ 3D impact parameter $\delta_{3D}$ and significance $\delta_{3D}/\sigma(\delta_{3D})$

Urs Langenegger
Search for $B^0_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ in CMS (2012/05/30)
Isolation I

- **Primary vertex isolation**: Relative dimuon isolation
  - ‘classic’ variable
    - in cone around dimuon momentum
    - for tracks in cone with $\Delta R < 0.7$
      - with $p_\perp > 0.9$ GeV
      - either associated to same PV as candidate
        or with $d_{ca} < 500 \mu$m and not associated to another PV

  $$I = \frac{p_\perp (\mu^+\mu^-)}{p_\perp (\mu^+\mu^-) + \sum_{\Delta R < 0.7} p_\perp}$$

Parameters tuned to minimize data/MC discrepancy ($B^\pm \rightarrow J/\psi K^\pm$) and maximize dimuon bg rejection.

---

Urs Langenegger

Search for $B_s^0 \rightarrow \mu^+\mu^-$ and $B^0 \rightarrow \mu^+\mu^-$ in CMS (2012/05/30)
Isolation II

- **$B$ vertex isolation:**
  - based on tracks reconstructed in the proximity of the secondary $B$ vertex
  - avoid pileup dependence:
    - either tracks associated to no primary vertex
    - or tracks associated to same vertex as $B$ candidate
  - $d_{ca}^0$: distance of closest track to $B$ vertex
  - $N_{trk}^{\text{close}}$: number of close tracks
    - $d_{ca} < 300 \, \mu$m
    - $p_{\perp} > 0.5 \, \text{GeV}$

- **Validation of $B_s^0$ MC:**
  - $B_s^0 \rightarrow J/\psi \phi$
    - (see below)
Normalization and control samples

- **Normalization sample**
  - $B^\pm \rightarrow J/\psi K^\pm$
  - validation of $B^+ \text{ MC}$

- **Control sample**
  - $B_s^0 \rightarrow J/\psi \phi$
  - validation of $B_s^0 \text{ signal MC}$

- **Combine $J/\psi$ with 1 or 2 ‘kaons’**
  - $3.0 < m(\mu\mu) < 3.2 \text{ GeV}$
  - $p_\perp(\mu\mu) > 7 \text{ GeV}$
  - $p_\perp(K) > 0.5 \text{ GeV}$
  - additional selection for $\phi$:
    - $0.995 < m(KK) < 1.045 \text{ GeV}$
    - $\Delta R(K, K) < 0.25$
  - all 3 (4) tracks used in vertexing

- **Comparison of (sideband-subtracted) data and MC simulation**
  - MC simulation normalized to data
Kinematics

- Bridges to CMS, 5 fb$^1$ \( \sqrt{s} = 7 \text{ TeV} \)

\[ B^+ \rightarrow J/\psi K^+ \]
- Data
- MC simulation

\[ B^+ \rightarrow J/\psi K^+ \]
- Data
- MC simulation

- **Leading muon** \( p_\perp \)

- **Sub-leading muon** \( p_\perp \)

\[ B \eta \]

\[ B \ p_\perp \]

- **Barr Langenegger**

Search for \( B^0_s \rightarrow \mu^+ \mu^- \) and \( B^0 \rightarrow \mu^+ \mu^- \) in CMS (2012/05/30)
Isolation

Candidate distributions for $B^+ \to J/\psi K^+$ and $B^0 \to J/\psi K^+$ in CMS (2012/05/30)
Selection efficiency (uncertainty)

- Determine selection efficiency in data
  - MC simulation

  with respect to 'all other selection requirements', e.g. for $B^{\pm} \to J/\psi K^{\pm}$:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Selection</th>
<th>MC</th>
<th>Data</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>muon $p_{\perp}$</td>
<td>$p_{\perp} &gt; 4.0$ GeV</td>
<td>0.927 ± 0.001</td>
<td>0.926 ± 0.001</td>
<td>-0.002 ± 0.001</td>
</tr>
<tr>
<td>pointing angle</td>
<td>$\alpha_{3D} &lt; 0.050$ rad</td>
<td>0.994 ± 0.000</td>
<td>0.995 ± 0.000</td>
<td>+0.000 ± 0.000</td>
</tr>
<tr>
<td>vertex fit</td>
<td>$\chi^{2}/dof &lt; 2.0$</td>
<td>0.936 ± 0.001</td>
<td>0.928 ± 0.001</td>
<td>-0.009 ± 0.001</td>
</tr>
<tr>
<td>impact parameter</td>
<td>$\delta_{3D} &lt; 0.008$</td>
<td>0.972 ± 0.001</td>
<td>0.972 ± 0.001</td>
<td>+0.001 ± 0.001</td>
</tr>
<tr>
<td>impact param. sign.</td>
<td>$\delta_{3D}/\sigma(\delta_{3D}) &lt; 2.000$</td>
<td>0.959 ± 0.001</td>
<td>0.944 ± 0.001</td>
<td>-0.015 ± 0.001</td>
</tr>
<tr>
<td>flight length sig.</td>
<td>$\ell_{3d}/\sigma(\ell_{3d}) &gt; 15.0$</td>
<td>0.923 ± 0.001</td>
<td>0.926 ± 0.001</td>
<td>+0.004 ± 0.001</td>
</tr>
<tr>
<td>isolation</td>
<td>$I &gt; 0.80$</td>
<td>0.893 ± 0.001</td>
<td>0.871 ± 0.001</td>
<td>-0.025 ± 0.002</td>
</tr>
<tr>
<td>close tracks</td>
<td>$N_{trk} &lt; 2$</td>
<td>0.978 ± 0.000</td>
<td>0.975 ± 0.000</td>
<td>-0.003 ± 0.001</td>
</tr>
<tr>
<td>$d_{ca}^{0}$</td>
<td>$d_{ca}^{0} &gt; 0.015$ cm</td>
<td>0.917 ± 0.001</td>
<td>0.929 ± 0.001</td>
<td>+0.013 ± 0.001</td>
</tr>
</tbody>
</table>

⇒ Systematic uncertainty from (quadr.) sum of relative differences

- $B^{\pm} \to J/\psi K^{\pm}$: 4%
  (largest single deviation: 2.5% from isolation)
- $B_{s}^{0} \to J/\psi \phi$: 3%
  (largest single deviation: 1.5% from $B$ vertex $\chi^{2}/dof$)
- idem for signal selection efficiency uncertainty
Pileup dependence?
Pileup independence

- Determine selection efficiency vs $N_{\text{PV}}$ in data
  - 2011 dataset:
    - $\langle N_{\text{PV}} \rangle \approx 8$
    - $\text{RMS}(z) \approx 5.6\,\text{cm}$

$B^{\pm} \rightarrow J/\psi K^{\pm}$

$B_{s}^{0} \rightarrow J/\psi \phi$

- MC: also checked $\varepsilon$ for
  - $N_{\text{PV}} < 6$
  - $N_{\text{PV}} > 10$

$\Rightarrow$ no pileup dependence

Urs Langenegger

Search for $B_{s}^{0} \rightarrow \mu^{+}\mu^{-}$ and $B^{0} \rightarrow \mu^{+}\mu^{-}$ in CMS (2012/05/30)
Search Analysis

Search for $B_s^0 \rightarrow \mu^+\mu^-$ and $B^0 \rightarrow \mu^+\mu^-$ in CMS (2012/05/30)
Selection for search analysis

- Random grid optimization
  - 14 variables included in $1.4 \times 10^6$ runs
  - ‘rounding’ of best parameters for final selection

<table>
<thead>
<tr>
<th>Variable</th>
<th>Barrel</th>
<th>Endcap</th>
<th>units</th>
<th>comparison to old analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_{\perp \mu, 1}$</td>
<td>4.5</td>
<td>4.5</td>
<td>GeV</td>
<td>same</td>
</tr>
<tr>
<td>$p_{\perp \mu, 2}$</td>
<td>4.0</td>
<td>4.2</td>
<td>GeV</td>
<td>tighter in endcap</td>
</tr>
<tr>
<td>$p_{\perp B}$</td>
<td>6.5</td>
<td>8.5</td>
<td>GeV</td>
<td>tighter in endcap</td>
</tr>
<tr>
<td>$\ell_{3d}$</td>
<td>1.5</td>
<td>1.5</td>
<td>cm</td>
<td>tighter</td>
</tr>
<tr>
<td>$\alpha &lt;$</td>
<td>0.050</td>
<td>0.030</td>
<td>rad</td>
<td>looser</td>
</tr>
<tr>
<td>$\chi^2/dof &lt;$</td>
<td>2.2</td>
<td>1.8</td>
<td></td>
<td>looser</td>
</tr>
<tr>
<td>$\ell_{3d}/\sigma(\ell_{3d}) &gt;$</td>
<td>13.0</td>
<td>15.0</td>
<td></td>
<td>looser</td>
</tr>
<tr>
<td>$I &gt;$</td>
<td>0.80</td>
<td>0.80</td>
<td>cm</td>
<td>redefined</td>
</tr>
<tr>
<td>$d_{ca}^0 &gt;$</td>
<td>0.015</td>
<td>0.015</td>
<td>cm</td>
<td>redefined</td>
</tr>
<tr>
<td>$\delta_{3D} &lt;$</td>
<td>0.008</td>
<td>0.008</td>
<td>cm</td>
<td>new</td>
</tr>
<tr>
<td>$\delta_{3D}/\sigma(\delta_{3D}) &lt;$</td>
<td>2.000</td>
<td>2.000</td>
<td>cm</td>
<td>new</td>
</tr>
<tr>
<td>$N_{trk} &lt;$</td>
<td>2</td>
<td>2</td>
<td>cm</td>
<td>new</td>
</tr>
</tbody>
</table>

- Total efficiency $\times$ acceptance

<table>
<thead>
<tr>
<th>Efficiency</th>
<th>Barrel</th>
<th>Endcap</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B_s^0 \rightarrow \mu^+ \mu^-$</td>
<td>0.0029 ± 0.0002</td>
<td>0.0016 ± 0.0002</td>
</tr>
<tr>
<td>$B^\pm \rightarrow J/\psi K^\pm$</td>
<td>0.00110 ± 0.00009</td>
<td>0.00032 ± 0.00004</td>
</tr>
</tbody>
</table>
Dimuon mass distribution (blinded)

- Low background (sidebands shown only)

Urs Langenegger

Search for $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ in CMS (2012/05/30)
Measurement of $B^\pm \rightarrow J/\psi K^\pm$

- Use identical selection as for dimuon, plus
  - $3.0 < m(\mu\mu) < 3.2 \text{ GeV}$
  - $p_\perp(\mu\mu) > 7 \text{ GeV}$, $p_\perp(K) > 0.5 \text{ GeV}$
  - all tracks used in vertexing

- Fit function
  - signal: double Gaussian
  - background: exponential + error function
    partially reconstructed $B$ decays
    $B^0 \rightarrow J/\psi K^* \rightarrow \mu^+\mu^- K^- (\pi^+)$

<table>
<thead>
<tr>
<th></th>
<th>Barrel</th>
<th>Endcap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptance</td>
<td>$0.162 \pm 0.006$</td>
<td>$0.111 \pm 0.006$</td>
</tr>
<tr>
<td>$\varepsilon_{\text{tot}}$</td>
<td>$0.00110 \pm 0.00009$</td>
<td>$0.00032 \pm 0.00004$</td>
</tr>
<tr>
<td>$N_{\text{obs}}$</td>
<td>$82712 \pm 4146$</td>
<td>$23809 \pm 1203$</td>
</tr>
</tbody>
</table>

- Systematic error on yield: 5%
  - variation of
    - background pdf
    - vary signal pdf
  - mass-constrain dimuons to $J/\psi$ (better resolution)
Rare backgrounds

- Rare backgrounds
  - CKM-suppressed semileptonic decays
e.g. \( B_s^0 \rightarrow K^- \mu^+ \nu \), one fake muon
    large \( B \), but mostly at low masses
  - ‘peaking’ hadronic decays
    e.g. \( B_s^0 \rightarrow K^- K^+ \), two fake muons
  - Normalization to \( B^+ \) yield in data
    \[
    N(X) = \frac{\mathcal{B}(Y \rightarrow X)}{\mathcal{B}(B^+ \rightarrow J/\psi K^\pm)} \frac{f_Y}{f_u} \frac{\varepsilon_{\text{tot}}(X)}{\varepsilon_{\text{tot}}(B^+)} N_{\text{obs}}(B^+)
    \]
    weighting with misid rate \( f \) (or \( \varepsilon_\mu \)) and \( \varepsilon_{\text{trig}} \)

- Note
  - \( B^0 \) more affected than \( B_s^0 \)
  - endcap more diluted than barrel
    lower efficiency
- Systematic error varies
  - branching fraction uncertainties
  - \( f_s/f_u = 0.267 \pm 0.021 \) [LHCb, arxiv:1111.2357]
## Systematic uncertainties

- **Acceptance:**
  - mixture of production processes
    - gluon fusion
    - flavor excitation
    - gluon splitting
  - half of acceptance variation
  - Studied variables sensitive to mixture
    - muon vs $B$ candidate:
      - $\Delta R(B, \mu)$
      - $p_\perp(\mu)$

- **Selection efficiency**
  - from data/MC comparisons
  - quadratic sum for all selection criteria

- **Muon trigger and efficiency**
  - full variation, for thresholds $4 < p_\perp < 8$ GeV
  - efficiency difference between data and MC

### Table: Systematic Uncertainties

<table>
<thead>
<tr>
<th>Category</th>
<th>Barrel</th>
<th>Endcap</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon_{\text{tot}}$ (signal)</td>
<td>3%</td>
<td>3%</td>
</tr>
<tr>
<td>$\varepsilon_{\text{tot}}$ (normalization)</td>
<td>4%</td>
<td>4%</td>
</tr>
<tr>
<td>kaon tracking</td>
<td>4%</td>
<td>4%</td>
</tr>
</tbody>
</table>

### Table: Additional Systematic Uncertainties

<table>
<thead>
<tr>
<th>Category</th>
<th>Barrel</th>
<th>Endcap</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$ trigger</td>
<td>3%</td>
<td>6%</td>
</tr>
<tr>
<td>$\mu$ ID</td>
<td>4%</td>
<td>8%</td>
</tr>
</tbody>
</table>
Summary of systematic errors

- Systematic uncertainties propagated into upper limit calculation
  all errors below in %

<table>
<thead>
<tr>
<th>Category</th>
<th>Uncertainty</th>
<th>Barrel</th>
<th>Endcap</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_s/f_u$</td>
<td>production ratio of $u$ and $s$ quarks</td>
<td>8.0</td>
<td>8.0</td>
</tr>
<tr>
<td>acceptance</td>
<td>production processes</td>
<td>3.5</td>
<td>5.0</td>
</tr>
<tr>
<td>$P_{ij}^B$</td>
<td>mass scale and resolution</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>efficiency (signal)</td>
<td>discrepancies data/MC simulation</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>efficiency (normalization)</td>
<td>discrepancies data/MC simulation</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td>efficiency (normalization)</td>
<td>kaon track efficiency</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td>efficiency</td>
<td>trigger</td>
<td>3.0</td>
<td>6.0</td>
</tr>
<tr>
<td>efficiency</td>
<td>muon identification</td>
<td>4.0</td>
<td>8.0</td>
</tr>
<tr>
<td>normalization</td>
<td>fit pdf</td>
<td>5.0</td>
<td>5.0</td>
</tr>
<tr>
<td>background</td>
<td>shape of combinatorial background</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td>background</td>
<td>rare decays</td>
<td>20.0</td>
<td>20.0</td>
</tr>
</tbody>
</table>
Cross checks

- Determination of \( \mathcal{B}(B_s^0 \to J/\psi \phi)/\mathcal{B}(B^\pm \to J/\psi K^\pm) \)
  - barrel vs. endcap
  - \( B^+ \) fitting
  - consistent definitions
    - acceptance
    - efficiency
    - (different number of daughters)
  - \( f_s/f_u \) from LHCb

- Inverted isolation sample \((I < 0.7, \text{not blinded})\)
  - comparison of prediction vs. observation
  - validation of rare backgrounds
  - background interpolation

- Stability vs. time (HLT changes)
  - yields (dimuons, normalization and control sample)
  - yield ratios
Results
Upper limit calculation

**Methodology**

- CL$_S$
- Feldman-Cousins
- statistical model:

\[
\begin{align*}
N^B_s &\sim\text{Pois}(\tau^B_s \nu^B_b + \nu^B_{s,\text{rare}} + P^B_{ss} \mu^B_s \nu^B_s + P^B_{sd} \mu^B_d \nu^B_d) \\
N^B_d &\sim\text{Pois}(\tau^B_d \nu^B_b + \nu^B_{d,\text{rare}} + P^B_{ds} \mu^B_s \nu^B_s + P^B_{dd} \mu^B_d \nu^B_d)
\end{align*}
\]

with \((i = s, d)\)

- \(\tau^B_i\): Ratio of \((B^0_i \to \mu\mu)\)-signal window size to size of background window
- \(\nu^B_{i,\text{rare}}\): Expected number of rare background in \((B^0_i \to \mu\mu)\)-signal window.
- \(\nu^B_i\): Expected number of reconstructed \((B^0_i \to \mu\mu)\) decays in barrel region assuming the SM
- \(P^B_{ij}\): Probability for a reconstructed \(B^0_j \to \mu\mu\) decay to be in \((B^0_i \to \mu\mu)\)-signal window.
- \(\mu_i\): Signal strength of \(B^0_i \to \mu\mu\), that is the ratio of true branching ratio to SM branching ratio.

**Systematic error on cross feed** \(P^B_{i,j}\)

- mass scale and resolution
- measure \(J/\psi \to \mu^+\mu^-\), \(B^0_s \to \mu^+\mu^-\), \(\Upsilon(1S) \to \mu^+\mu^-\)
- compare MC resolution (and position) with prediction (interpolation)
**Expectations and observation**

<table>
<thead>
<tr>
<th>Variable</th>
<th>$B^0 \rightarrow \mu^+\mu^-$ Barrel</th>
<th>$B_s^0 \rightarrow \mu^+\mu^-$ Barrel</th>
<th>$B^0 \rightarrow \mu^+\mu^-$ Endcap</th>
<th>$B_s^0 \rightarrow \mu^+\mu^-$ Endcap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal</td>
<td>0.24 $\pm$ 0.02</td>
<td>2.70 $\pm$ 0.41</td>
<td>0.10 $\pm$ 0.01</td>
<td>1.23 $\pm$ 0.18</td>
</tr>
<tr>
<td>Combinatorial bg</td>
<td>0.40 $\pm$ 0.34</td>
<td>0.59 $\pm$ 0.50</td>
<td>0.76 $\pm$ 0.35</td>
<td>1.14 $\pm$ 0.53</td>
</tr>
<tr>
<td>Peaking bg</td>
<td>0.33 $\pm$ 0.07</td>
<td>0.18 $\pm$ 0.06</td>
<td>0.15 $\pm$ 0.03</td>
<td>0.08 $\pm$ 0.02</td>
</tr>
<tr>
<td>Sum</td>
<td>0.97 $\pm$ 0.35</td>
<td>3.47 $\pm$ 0.65</td>
<td>1.01 $\pm$ 0.35</td>
<td>2.45 $\pm$ 0.56</td>
</tr>
<tr>
<td>Observed</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>

**Signal regions**

$B^0 : 5.20 < m < 5.30 \text{ GeV}$

$B_s^0 : 5.30 < m < 5.45 \text{ GeV}$
Results: upper limits

- Upper limit on $\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-)$ and $\mathcal{B}(B^0 \rightarrow \mu^+\mu^-)$

<table>
<thead>
<tr>
<th>upper limit (95%CL)</th>
<th>observed</th>
<th>(median) expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-)$</td>
<td>$7.7 \times 10^{-9}$</td>
<td>$8.4 \times 10^{-9}$</td>
</tr>
<tr>
<td>$\mathcal{B}(B^0 \rightarrow \mu^+\mu^-)$</td>
<td>$1.8 \times 10^{-9}$</td>
<td>$1.6 \times 10^{-9}$</td>
</tr>
</tbody>
</table>

- $p$-values (for background-only hypotheses)

<table>
<thead>
<tr>
<th>$p$-values</th>
<th>background only</th>
<th>SM cross feed</th>
<th>floating cross feed</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-)$</td>
<td>0.06 (1.5σ)</td>
<td>0.07 (1.5σ)</td>
<td>0.11 (1.2σ)</td>
</tr>
<tr>
<td>$\mathcal{B}(B^0 \rightarrow \mu^+\mu^-)$</td>
<td>0.11 (1.2σ)</td>
<td>0.29 (0.6σ)</td>
<td>0.24 (0.7σ)</td>
</tr>
</tbody>
</table>
Interpretation examples

- Empty region due to previous upper limit and other published data

$\Rightarrow$ strongest impact at large $\tan \beta$
Conclusions

• Search for $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ in 2011 dataset

<table>
<thead>
<tr>
<th>upper limit (95%CL)</th>
<th>observed</th>
<th>expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)$</td>
<td>$7.7 \times 10^{-9}$</td>
<td>$8.4 \times 10^{-9}$</td>
</tr>
<tr>
<td>$\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-)$</td>
<td>$1.8 \times 10^{-9}$</td>
<td>$1.6 \times 10^{-9}$</td>
</tr>
</tbody>
</table>

• Significant improvement
  ▶ at EPS 2011: $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) < 1.9 \times 10^{-8}$ (at 95%CL)
  ▶ more/changed variables, e.g., better $B$ vertex isolation
  ▶ improved sensitivity
  ▶ higher signal/background ratio
  ▶ publ. in JHEP 04(2012), 033

• Well prepared for 2012 data
  ▶ higher instantaneous lumi (trigger thresholds are looser)
  ▶ high pileup (tracking and vertexing)
Backup
A ‘new’ analysis

- Analysis was performed **blind**
  - reblinded old data \(1.1 \text{ fb}^{-1}\)
  - total amount of data: \(4.9 \text{ fb}^{-1}\)

- Significant analysis modifications
  - tighter muon identification \((3 \times \text{ smaller fake rate})\)
  - isolation variables
    - primary vertex isolation (redefined)
    - \(B\) vertex isolation: distance of closest track (redefined)
    - \(B\) vertex isolation: track counting (new)
  - non-monotonous changes
  - 3D impact parameter and its significance (new)

⇒ Better analysis
  - pileup independence up to \(N_{PV} \approx 30\)
  - higher sensitivity
  - larger signal/background

Analysis is (still) cut-n-count
Unblinding post-mortem

- Comparing signal events of this analysis to the EPS-2011 analysis
  - no overlap between signal events
    - 2 barrel $B^0$ candidates: killed by (changed) $d_{ca}^0$
    - 1 endcap $B^0$ candidate: killed by (tightened) $\ell_{3D}$ selection
    - 1 endcap $B_s^0$ candidate: killed by (new) 3D impact parameter (and its significance)
  - one signal event is from EPS-2011 dataset
    - $\chi^2$/dof requirement is now looser
  - one common event in the sideband
    - kinematical variables are identical
    - isolation variables changed (to be expected)

- Summary of signal box migrations
  - Barrel
    - $B^0$: 0 become 2
    - $B_s^0$: 2 disappear, 2 appear
  - Endcap
    - $B^0$: 1 disappears, 0 appear
    - $B_s^0$: 1 disappears, 4 appear