Search for $B^0_{(s)} \rightarrow \mu^+ \mu^-$ at LHCb

Marc-Olivier Bettler, INFN Firenze
on behalf of the LHCb collaboration
LHCb

- Large $b\bar{b}$ cross section,
- Large acceptance for $B$ decays,
- Very efficient muon trigger,
- Good particle ID, tracking and reconstruction.

LHCb published limits $\mathcal{B}(B^0_{(s)} \rightarrow \mu^+\mu^-)$ on the 37 pb$^{-1}$ of data from 2010.


We present here an update on 300 pb$^{-1}$ collected in 3 months of 2011. Assuming SM branching fractions and after selection, we expect in these data

- 3.2 $B^0_s \rightarrow \mu^+\mu^-$ events and 0.32 $B^0 \rightarrow \mu^+\mu^-$ events.

LHCb has 630 pb$^{-1}$ on tape at the moment and expects to take $\sim$ 1 fb$^{-1}$ this year.
Selection

- **Muon** trigger
- **Soft selection**, similar to normalization channels
- **Blind** signal region $[M_{B^0} - 60 \text{ MeV}/c^2, M_{B^0_s} + 60 \text{ MeV}/c^2]$

Signal/background Discrimination

- **Boosted Decision Tree** combining topological and kinematic information
- **Invariant mass** of the muon pair

Data driven calibration through control channels to get signal and background expectations

Normalization to convert a number of observed events into a branching fraction using channels of known branching fractions

Observation/exclusion measurement confronting the experimental observation to signal and background expectations using the CL$_s$ method in bins of mass and BDT.

$B^0_{(s)} \rightarrow \mu^+\mu^-$ at LHCb

Marc-Olivier Bettler
Nine input variables such no correlation with mass:
• $B$ impact parameter, lifetime and $p_T$
• $\mu$ isolation, DOCA and minimal impact parameter

► $B$ isolation
► Polarization variable
► minimum $\mu$ $p_T$

Trained on MC, with
► $B^0_s \rightarrow \mu^+\mu^-$ signal and
► $b\bar{b} \rightarrow \mu\mu X$ background.
The BDT response is calibrated on data using:

- $B_{(s)}^0 \rightarrow h^+ h^-$ events for the signal
  
  - same topology as $B_s^0 \rightarrow \mu^+ \mu^-$
  
  - triggered very differently

- use of candidates that did not contribute to the trigger decision to avoid bias.

- events in the **mass sidebands** for the background.
The Invariant mass is modeled with a Crystal Ball.

- The **means** of the Gaussian are obtained from $B_s^0 \rightarrow K^+K^-$ and $B^0 \rightarrow K^+\pi^-$.
- The **width** of the Gaussian is obtained from data the interpolation between the mass resolution of dimuon resonances ($J/\psi$, $\psi(2S)$ and $\Upsilon$’s) cross-checked with inclusive/exclusive fits of $B_{(s)}^0 \rightarrow h^+h^-$.

$\sigma = 24.6 \pm 0.2 \pm 1.0 \text{ MeV}/c^2$

The search windows are defined as $\pm 60 \text{ MeV}/c^2$ around each $B_{(s)}^0$ mass.
The combinatorial background expectation is extracted from a fit to the mass sidebands in bins of BDT. Systematics are evaluated using different fit models and fit ranges.
The dominant background is composed of real muons from $b\bar{b} \rightarrow \mu\mu X$ events. The other sources of background stem from:

- $pp \rightarrow p\mu^+\mu^- p$ processes.
  - Isolated muons,
  - Dimuon mass possibly high,
  - Very low $p_T$ ★ Very efficiently anti-selected requiring $p_T(\text{dimuon}) > 500 \text{ MeV/c}$

- $B^{0}_{(s)} \rightarrow h^+h^-$ processes were the two final states are misidentified as muons
  - The misid probability is obtained from control channels and reweighted to the $B$ kinematics
  - The distribution of the misidentified $B^{0}_{(s)} \rightarrow h^+h^-$ in bins of mass is evaluated from reweighted MC.
Normalization

\[ \mathcal{B}_{\text{sig}} = \mathcal{B}_{\text{norm}} \times \frac{\epsilon_{\text{norm}}}{\epsilon_{\text{sig}}} \times \frac{\epsilon_{\text{TRIG|SEL}}}{\epsilon_{\text{sig}}} \times \frac{f_{\text{norm}}}{f_{\text{sig}}} \times \frac{N_{\text{sig}}}{N_{\text{norm}}} = \alpha \times N_{\text{sig}} \]

Evaluated on MC, cross-checked on data
Measured new measurement by LHCb

Three complementary normalization channels are used:

\[ B^0_s \rightarrow J/\psi \phi \]

- \[ 5919 \pm 84 \text{ (LHCb preliminary)} \]

\[ B^+ \rightarrow J/\psi K^+ \]

- \[ 107358 \pm 1759 \text{ (LHCb preliminary)} \]

\[ B^0 \rightarrow K^+ \pi^- \]

- \[ 5732 \pm 506 \text{ (LHCb)} \]

\[ \alpha_{B^0_s \rightarrow \mu^+ \mu^-} = (9.84 \pm 0.91) \times 10^{-10}, \]

\[ \alpha_{B^0 \rightarrow \mu^+ \mu^-} = (2.89 \pm 0.15) \times 10^{-10}. \]

\[ B^0_{(s)} \rightarrow \mu^+ \mu^- \text{ at LHCb} \]

Marc-Olivier Bettler
fs/fd is now measured at LHCb:

- In hadronic decays, via the relative yields of $B^0 \rightarrow D^\pm K^\mp$ or $B^0 \rightarrow D^\pm \pi^\mp$ to $B^0_s \rightarrow D^\pm \pi^\mp$.

$$f_s/f_d = 0.253 \pm 0.017^{\text{stat}} \pm 0.017^{\text{syst}} \pm 0.020^{\text{theo}}$$

See Stefania Ricciardi’s talk

- And in semileptonic decays:

$$\frac{f_s}{f_u + f_d} = 0.134 \pm 0.004^{+0.011}_{-0.010}$$

The combination of the two measurements

$$f_s/f_d = 0.267^{+0.021}_{-0.020}$$

was used in the normalization.
In the $B_s^0 \rightarrow \mu^+\mu^-$ search region

![Graphs showing BDT bins for $B_s^0 \rightarrow \mu^+\mu^-$](image)

<table>
<thead>
<tr>
<th>Preliminary</th>
<th>BDT bin</th>
<th>[0, 0.25]</th>
<th>[0.25, 0.5]</th>
<th>[0.5, 0.75]</th>
<th>[0.75, 1]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comb. bkg exp.</td>
<td>2969 ± 68</td>
<td>25 ± 2.5</td>
<td>2.89 ± 0.89</td>
<td>0.66 ± 0.40</td>
<td></td>
</tr>
<tr>
<td>SM signal exp.</td>
<td>1.26 ± 0.13</td>
<td>0.61 ± 0.06</td>
<td>0.67 ± 0.07</td>
<td>0.72 ± 0.07</td>
<td></td>
</tr>
<tr>
<td>Observed</td>
<td>2872</td>
<td>26</td>
<td>3</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

$B_s^0 \rightarrow \mu^+\mu^-$ at LHCb  
Marc-Olivier Bettler
Preliminary results from 300 pb$^{-1}$ of data collected in 2011:

$$\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-) < 1.6 \ (1.3) \times 10^{-8} \ \text{at 95\% (90\%) C.L.}$$

- expected limit, bkg-only $< 1.0 \ (0.8) \times 10^{-8}$
- expected limit, bkg&SM $< 1.5 \ (1.2) \times 10^{-8}$
In the $B^0 \rightarrow \mu^+ \mu^-$ search region

<table>
<thead>
<tr>
<th>BDT bin</th>
<th>Comb. bkg exp.</th>
<th>MisID bkg exp.</th>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preliminary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$[0, 0.25]$</td>
<td>$3176 \pm 69$</td>
<td>$0.6^{+0.4}_{-0.3}$</td>
<td>$3025$</td>
</tr>
<tr>
<td>$[0.25, 0.5]$</td>
<td>$26.6 \pm 2.5$</td>
<td>$0.6^{+0.4}_{-0.3}$</td>
<td>$31$</td>
</tr>
<tr>
<td>$[0.5, 0.75]$</td>
<td>$3.1 \pm 0.9$</td>
<td>$0.6^{+0.4}_{-0.3}$</td>
<td>$5$</td>
</tr>
<tr>
<td>$[0.75, 1]$</td>
<td>$0.74 \pm 0.40$</td>
<td>$0.6^{+0.4}_{-0.3}$</td>
<td>$4$</td>
</tr>
</tbody>
</table>

$B^0_{(s)} \rightarrow \mu^+ \mu^-$ at LHCb

Marc-Olivier Bettler
Preliminary results from 300 pb$^{-1}$ of data collected in 2011:

\[ \mathcal{B}(B^0 \rightarrow \mu^+ \mu^-) < 5.2 \ (4.2) \times 10^{-9} \text{ at } 95\% \ (90\%) \text{ C.L.} \]

expected limit, bkg-only < 3.1 \ (2.4) \times 10^{-9}
Extrapolation based on the 37 pb$^{-1}$ collected in 2010 and analysed with the 2010-analysis. This can be considered rather conservative.

The $\sim 1$ fb$^{-1}$ expected at the end of 2011 allow a considerable constraint on MSSM.
Extrapolation based on the 37 pb$^{-1}$ collected in 2010 and analysed with the 2010-analysis. This can be considered rather conservative.
• LHCb presents a preliminary result with 300 pb$^{-1}$ at $\sqrt{s} = 7$ TeV

\[ \mathcal{B}(B^0_s \rightarrow \mu^+\mu^-) < 1.6 (1.3) \times 10^{-8} \text{ at } 95\% \ (90\%) \text{ C.L.} \]

\[ \mathcal{B}(B^0 \rightarrow \mu^+\mu^-) < 5.2 (4.2) \times 10^{-9} \text{ at } 95\% \ (90\%) \text{ C.L.} \]

• Combining with our previous result on 37 pb$^{-1}$:

\[ \mathcal{B}(B^0_s \rightarrow \mu^+\mu^-) < 1.5 (1.2) \times 10^{-8} \text{ at } 95\% \ (90\%) \text{ C.L.} \]

• And combining further with CMS result:

\[ \mathcal{B}(B^0_s \rightarrow \mu^+\mu^-) < 1.1 (0.9) \times 10^{-8} \text{ at } 95\% \ (90\%) \text{ C.L.} \]

• The excess seen by CDF is not confirmed.
Backup Material
Best $B^0_s \rightarrow \mu^+ \mu^-$ candidate in terms of BDT (0.9) and mass ($5357 \text{ MeV}/c^2$). The muons’ vertex is well displaced with respect to the primary vertex ($l = 115 \text{ mm}$). The $B$ momentum is $58.16 \text{ GeV}/c$, its transverse momentum is $3.5 \text{ GeV}/c$ and its lifetime is $3.52 \text{ ps}$. 
Observed distribution

\[ B^0 \rightarrow \mu^+\mu^- \text{ at LHCb} \]

Marc-Olivier Bettler
New Variables, BDT vs GL

- $B$ isolation = $p_T(B)/p_T(B) + \sum_{\text{tracks}} p_T(\text{tracks})$
  for tracks such that $\sqrt{\delta \eta^2 + \delta \phi^2} < 1.0$.

- Polarization variable = \(\frac{(p_{y,B} \times p_{x,\mu} - p_{x,B} \times p_{y,\mu})}{2 \times p_{T,B} \times M_B}\)
  with $\mu$ is the muon with minimum $p_T$. Cosine of the angle between the muon momentum in the $B$ rest frame and the vector perpendicular to the $B$ momentum and the beam axis.

Our previous result was obtained using another MultiVariate tool: the GL. With additional variables, the best discrimination power was obtained using a BDT. Trained and tested on MC.
## Normalization uncertainties

<table>
<thead>
<tr>
<th>$B$</th>
<th>$\mathcal{B}$</th>
<th>$\epsilon_{\text{cal}}^{\text{REC}} / \epsilon_{\text{sig}}^{\text{REC}}$</th>
<th>$\mathcal{N}_{\text{cal}}$</th>
<th>$\alpha_{B_0(s) \rightarrow \mu^+\mu^-}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B^+ \rightarrow J/#psi K^+$</td>
<td>$6.01 \pm 0.21$</td>
<td>$0.48 \pm 0.02$</td>
<td>$107358 \pm 1759$</td>
<td>$2.58 \pm 0.16$</td>
</tr>
<tr>
<td>$B^0_s \rightarrow J/#psi \phi$</td>
<td>$3.4 \pm 0.9$</td>
<td>$0.23 \pm 0.02$</td>
<td>$5919 \pm 84$</td>
<td>$3.39 \pm 0.98$</td>
</tr>
<tr>
<td>$B^0 \rightarrow K^+\pi^-$</td>
<td>$1.94 \pm 0.06$</td>
<td>$0.86 \pm 0.03$</td>
<td>$5732 \pm 506$</td>
<td>$2.47 \pm 0.57$</td>
</tr>
</tbody>
</table>

Summary of the factors and their uncertainties needed to calculate the normalization factors ($\alpha_{B_0(s) \rightarrow \mu^+\mu^-}$) for the three normalization channels considered. The trigger efficiency and number of $B^0 \rightarrow K^+\pi^-$ candidates correspond to TIS events only.

<table>
<thead>
<tr>
<th>$B^+ \rightarrow J/#psi K^+$</th>
<th>$B^0_s \rightarrow J/#psi \phi$</th>
<th>$B^0 \rightarrow K^+\pi^-$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{B} = 6.01 \times 10^{-5} \pm 3.5%$</td>
<td>$\mathcal{B} = 3.4 \times 10^{-5} \pm 26%$</td>
<td>$\mathcal{B} = 1.94 \times 10^{-5} \pm 3.1%$</td>
</tr>
<tr>
<td>trigger and PID similar to signal channel</td>
<td>trigger and PID similar to signal channel</td>
<td>different trigger: use $B^0_{(s)} \rightarrow h^+h^-$ candidates that did not contribute to trigger decision</td>
</tr>
<tr>
<td>+1 track wrt signal: error on tracking $\epsilon$ dominates in $\epsilon$ ratio</td>
<td>+2 tracks wrt signal: error on tracking $\epsilon$ dominates in $\epsilon$ ratio</td>
<td>topology identical to signal</td>
</tr>
<tr>
<td>$f_s/f_d$ dominates overall</td>
<td>$\mathcal{B}$ dominates overall</td>
<td>$f_s/f_d$, trigger uncertainty</td>
</tr>
</tbody>
</table>
Full fit to the invariant mass distributions of $B^0_{(s)} \rightarrow h^+h^-$ candidates in the $\pi\pi$ mass hypothesis for the whole sample (top left) and for the samples in the bin 2,3,4 of the BDT (top right, bottom left, bottom right).
Topography look like signal when associated with another PV in the event.
Dimuon mass covers $B$ mass range.
But very low $p_T$, can be anti-selected very efficiently with a $p_T$ cut at 500 MeV (98.2% eff. on signal).