Search for new physics in rare decays of B-mesons at ATLAS

Darren Price (on behalf of the ATLAS experiment)

MESON 2016
14th International Workshop on Meson Production, Properties and Interactions
Krakow, Poland, June 2nd `16
Decays of $B^0_{(s)} \rightarrow \mu^+ \mu^-$ must proceed via flavour changing neutral currents:

- Loop and helicity suppressed decay
- Very sensitive to New Phenomena:
  both constructive/destructive interference possible

Purely leptonic final state:
theoretically and experimentally very clean
Decays of $B^0_{(s)} \rightarrow \mu^+ \mu^-$ must proceed via flavour changing neutral currents:

- Loop and helicity suppressed decay
- Very sensitive to New Phenomena: both constructive/destructive interference possible

Purely leptonic final state: theoretically and experimentally very clean

**SM predictions:**

$$\text{Br}(B^0_s \rightarrow \mu^+ \mu^-) = (3.65 \pm 0.23) \times 10^{-9}$$

$$\text{Br}(B^0 \rightarrow \mu^+ \mu^-) = (1.06 \pm 0.09) \times 10^{-10}$$

C. Bobeth et al., PRL 112 (2014) 101801
Decays of $B^0_{(s)} \rightarrow \mu^+ \mu^-$ must proceed via flavour changing neutral currents:

- Loop and helicity suppressed decay
- Very sensitive to New Phenomena: both constructive/destructive interference possible

Purely leptonic final state: theoretically and experimentally very clean

**SM predictions:**

- $\text{Br}(B^0_s \rightarrow \mu^+ \mu^-) = (3.65 \pm 0.23) \times 10^{-9}$
- $\text{Br}(B^0 \rightarrow \mu^+ \mu^-) = (1.06 \pm 0.09) \times 10^{-10}$

C. Bobeth et al., PRL 112 (2104) 101801

**CMS and LHCb combined result:**

- $\text{Br}(B^0_s \rightarrow \mu^+ \mu^-) = (2.8^{+0.7}_{-0.6}) \times 10^{-9}$
- $\text{Br}(B^0 \rightarrow \mu^+ \mu^-) = (3.9^{+1.6}_{-1.4}) \times 10^{-10}$

*Nature 522 (2015) 68–72*

Today, present results from ATLAS with full Run-1 dataset
Analysis overview

- Trigger events based on di-muon signature: $p_T(\mu) > 4 \ (4, \ 6) \ GeV$

- Select well-reconstructed di-muon candidate events in invariant mass range 4766–5966 MeV.
  
  *Range [5166–5526] MeV blinded until entire analysis chain defined*
  
  $m(B^{0}) = 5367 \ MeV, \ m(B^{0}) = 5280 \ MeV$

- Extract signal yield from data using unbinned maximum likelihood fit (UBML)

  *Multivariate analysis using two distinct BDTs trained for background suppression*
  
  *Data-driven control regions for cross-checks and background modelling*

- Normalise signal to $B^{\pm} \rightarrow J/\psi(\rightarrow \mu^{+}\mu^{-})K^{\pm}$ signal

  *Reference signal decay provides partial systematics cancellation*
Measure $B^0_s \rightarrow \mu^+ \mu^-$ with respect to $B^\pm \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)K^\pm$ reference channel:

$$\mathcal{B}(B^0_{(s)} \rightarrow \mu^+ \mu^-) = \frac{N_{B_d(s)}}{\varepsilon_{\mu^+ \mu^-}} \times \frac{\varepsilon_{J/\psi K^\pm}}{N_{J/\psi K^\pm}}$$

Extract yields of signal and reference from UBML fits
Correct each channel for efficiencies
Measure $B^0_s \rightarrow \mu^+ \mu^-$ with respect to $B^\pm \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)K^\pm$ reference channel:

$$
\mathcal{B}(B_{(s)}^0 \rightarrow \mu^+ \mu^-) = \frac{N_{B_d(s)}}{\varepsilon_{\mu^+ \mu^-}} \times \frac{\varepsilon_{J/\psi K^\pm}}{N_{J/\psi K^\pm}} \times \frac{f_u}{f_{d(s)}}
$$

Correct for hadronisation probability differences

*Use ATLAS measurement = 0.240±0.020 and isospin symmetry*

Measure $B^0_s \to \mu^+ \mu^-$ with respect to $B^\pm \to J/\psi (\to \mu^+ \mu^-) K^\pm$ reference channel:

$$\mathcal{B}(B^0_{(s)} \to \mu^+ \mu^-) = \frac{N_{B_d(s)}}{\varepsilon_{\mu^+ \mu^-}} \times \frac{\varepsilon_{J/\psi K^\pm}}{N_{J/\psi K^\pm}} \times \frac{f_u}{f_d(s)} \times \left[ \mathcal{B}(B^\pm \to J/\psi K^\pm) \times \mathcal{B}(J/\psi \to \mu^+ \mu^-) \right]$$

Account for reference channel branching fractions.
Measurement strategy

Modify previous formula slightly to account for changing trigger conditions:

\[
\mathcal{B}(B^0_{(s)} \rightarrow \mu^+ \mu^-) = N_{B_d(s)} \times \frac{f_u}{f_d(s)} \times \frac{1}{D_{\text{norm}}}
\]

\[
\times \left[ \mathcal{B}(B^\pm \rightarrow J/\psi K^\pm) \times \mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-) \right]
\]

\[
D_{\text{norm}} = \sum_k N_{J/\psi K^\pm}^k \alpha_k \left( \frac{\varepsilon_{\mu^+ \mu^-}}{\varepsilon_{J/\psi K^\pm}} \right)_k
\]

Above term captures changing reference channel yields and efficiency ratios for different trigger streams across 7 and 8 TeV data-taking.
Signal separated from uncorrelated $b(\rightarrow c)\rightarrow \mu$ background with MVA classifier
Background $x10^3$ larger than next largest (semi-leptonic) background
Use 15 variables related to B candidate, muons, tracks in event

Signal efficiency = 54%, for a $5x10^3$ background rejection
B-decays with two real muons look like signal but accumulate at low mass: modelled with Monte Carlo simulations

**Same-vertex (SV) backgrounds**

\[ B_d \rightarrow K^* \mu^+ \mu^-, \quad B_c \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \mu^+ \nu \]

**Same-side (SS) backgrounds**

\[ b \rightarrow \mu^- c (\rightarrow \mu^+ X') X \]

**Semi-leptonic**

\[ B \rightarrow \mu h \nu \]
Train dedicated fake BDT classifier against hadron misidentification

Studied with MC signal $B_s \rightarrow \mu^+\mu^-$, $B \rightarrow hh'$ and $\Lambda_b \rightarrow ph$ backgrounds

Validated with $\phi \rightarrow K^+K^-$, $B^+ \rightarrow J/\psi K^+$, inverted BDT $B_s \rightarrow \mu^+\mu^-$ data

$P(\text{misID}) = < 0.01\% (p), 0.4\% (K), 0.2\% (\pi)$

Background rejection 40% (x7 on 7 TeV) for 95% prompt muon efficiency working point

Total number of $B \rightarrow hh'$ contributing to signal region $1.0 \pm 0.4$
UBML fit applied to $J/\psi K^\pm$ and $J/\psi \pi^\pm$ data simultaneously

- Four-component fit for PRDs, combinatorial background, $K^\pm/\pi^\pm$ signal
- Continuum BDT and fake BDT selections applied to $B^\pm$ reference

Measure $\pi^\pm/K^\pm$ relative branching ratio:

$$
\rho_{\pi/K} = \frac{B(B^\pm \rightarrow J/\psi \pi^\pm)}{B(B^\pm \rightarrow J/\psi K^\pm)} = 0.035 \pm 0.003^{\text{stat}} \pm 0.012^{\text{syst}}
$$

LHCb: 0.0383±0.0011±0.0007
BaBar: 0.0537±0.0045±0.0011
Signal yield extracted from UBML fit to dimuon invariant mass distribution

Extracted simultaneously in three categories in three continuum BDT ranges (each with constant signal efficiency = 18%)

BDT ranges: [0.242–0.351], [0.351–0.454], [0.454–1.0]
Fit results

**Expected signal yield:**
\[ N(B_s) = 41, \quad N(B) = 5 \]
(Exp. significance \( B_s = 3.1\sigma, B_d = 0.2\sigma \))

**Observed fitted signal yield:**
\[ N_{\text{obs}}(B_s) = 16 \pm 12, \quad N_{\text{obs}}(B) = -11 \pm 9 \]

\[ N_{\text{obs}}(B_s) = 11, \quad N_{\text{obs}}(B) = 0 \] when constraining \( N \geq 0 \)
Determine branching fraction with non-negative boundary condition

Uncertainties obtained with Neyman construction of frequentist confidence belt including statistical and systematic uncertainties ($\sigma_{\text{syst}} = \pm 0.3 \times 10^{-9}$)

$$\text{Br}(B^0_s \rightarrow \mu^+ \mu^-) = (0.9^{+1.1}_{-0.8}) \times 10^{-9}$$

Upper limits from $\text{CL}_s$ approach:

$$\text{Br}(B^0_s \rightarrow \mu^+ \mu^-) < 3.0 \times 10^{-9} \text{ at } 95\% \text{ CL}$$

Observed compatibility with null (background-only) hypothesis:

$$p = 0.08 \ (1.4 \sigma)$$

Expectation for SM signal:

$$p = 0.0011 \ (3.1 \sigma)$$
Upper limit on branching fraction for $B_d$ determined using CL$_s$ technique:

$$\text{Br}(B^0_d \rightarrow \mu^+ \mu^-) < 4.2 \times 10^{-10} \text{ at } 95\% \text{ CL}$$

Expected limit < $5.7^{+2.1}_{-1.2} \times 10^{-10}$

In conjunction with $B_s$ results is compatible with the SM at: $p=0.048 \ (2.0 \sigma)$
Results: $B^0_s \rightarrow \mu^+ \mu^-$ vs. $B^0_d \rightarrow \mu^+ \mu^-$ contour


Likelihood contours without imposing natural boundaries

$|s| = 7$ TeV, $4.9$ fb$^{-1}$

$|s| = 8$ TeV, $20$ fb$^{-1}$

Contours for $-2 \Delta \ln(L) = 2.3, 6.2, 11.8$ from maximum of $L$
Using the data collected during Run-1 of the LHC, ATLAS has new results on the rare decays $B^0_s$ and $B^0$ into muon pairs

Uncertainties competitive with CMS and LHCb

Observe the following results:

$$\text{Br}(B^0_s \to \mu^+ \mu^-) = (0.9^{+1.1}_{-0.8}) \times 10^{-9}$$

$$\text{Br}(B^0_d \to \mu^+ \mu^-) < 4.2 \times 10^{-10} \text{ at } 95\% \text{ CL}$$

lower than the SM prediction

$$\text{Br}_{SM}(B^0_s \to \mu^+ \mu^-) = (3.65 \pm 0.23) \times 10^{-9}$$

$$\text{Br}_{SM}(B^0 \to \mu^+ \mu^-) = (1.06 \pm 0.09) \times 10^{-10}$$

Compatibility with the SM is 2.0 $\sigma$ in the simultaneous fit, leaving room for destructive interference from NP to the SM rate.

Analysis of Run-2 data (~100 fb$^{-1}$) expected to significantly enhance sensitivity.
Backup
Trigger and event selection

- Minimum $p_T$ thresholds on muons driven by trigger:

  7 TeV data: $p_T(\mu)>4$ GeV, $|\eta(\mu)|<2.5$

  8 TeV data split into three exclusive categories:
  - 8 TeV (T1): One muon $p_T(\mu)>6$ GeV, other $p_T(\mu)>4$ GeV, $|\eta(\mu)|<2.5$
  - 8 TeV (T2): Both $p_T(\mu)>4$ GeV and at least one in $|\eta(\mu)|<1.05$
  - 8 TeV (T3): Both $p_T(\mu)>4$ GeV and $1.05 \leq |\eta(\mu)|<2.5$

- Both selected muons must be Combined
  (Inner Detector and Muon Spectrometer muon track reconstruction)

- $B^0_{(s)} \to \mu^+\mu^-$ signal, and $B^\pm \to J/\psi(\to \mu^+\mu^-)K^\pm$, $B_s \to J/\psi(\to \mu^+\mu^-)\Phi(K^+K^-)$ reference:
  - Di-muon vertex fit
  - Association to primary vertex
  - Fiducial region: $p_T(B)>8$ GeV, $|\eta(B)|<2.5$
Background contributions to search (in decreasing size):

1. Combinatorial background
   Real muons from uncorrelated $b\to c\to \mu$ decays

2. Partially-reconstructed B decays
   Real muons coming from $B\to \mu\mu+X$ decays
   Single pion/kaon misidentified as muon (semi-leptonic B and $B_s$ decays)

3. Peaking backgrounds
   $B_s/B_d\to hh'$ ($h=\pi/K/p$). Small component, but dangerous as overlaid on signal!
Partially-reconstructed decays for $B^\pm$ reference channel
Signal yield extracted from UBML fit to dimuon invariant mass distribution

Extracted simultaneously in three categories in three continuum BDT ranges (each with constant signal efficiency = 18%)

BDT ranges: [0.242–0.351], [0.351-0.454], [0.454-1.0]

Signal:
- Two superimposed Gaussians;
- Common mean, avg. with 80 MeV, shape constrained across BDT

Low mass background:
- Exponential with mass (SS+SV)
- Shape constrained across BDT, normalisation independent

Continuum background:
- Linear with mass: small correlation with BDT interval consistent with sidebands/MC

Peaking background:
- Gaussian; equal amplitude in each BDT bin, constrained to 1.0±0.4 total
Efficiency correction

Reference-to-signal efficiency correction determined per dataset / trigger category

- $p_T-\eta$ spectra tuned on reference channels
- Trigger efficiency from data-driven tag-and-probe

MC-data comparison on discriminating variables in BDT
- only isolation needs tuning for $B^\pm$ mode
- for $B_{(s)}$ additional correction due to lifetime needed

Correction to ratio $\sim+3-4\%$ for $B^0$, $-0.6\%$ for $B^0_s$

Total systematic uncertainty from efficiency $= 5.9\%$

(includes effect of reweighting all 15 variables entering cBDT with $B^\pm$ data, trigger efficiency, $p_T-\eta$ reweighting, uncertainties on $B^\pm$ and $K^\pm$ reconstruction)
Inner B-Layer (IBL) upgrade for Run-2
Additional pixel layer, small radius: 33.25 mm (current B-layer at 50.5 mm)

Fourth pixel layer provides improved $d_0$ and $z_0$ resolution, and $\theta$ and $\phi$ resolution at low $p_T$ (~1 GeV)
Level-1 ‘topological’ triggers introduced to ATLAS for Run-2
New on-board algorithms allow trigger rate reduction of 2–5 at Level-1
Coarse topological RoI information added to di-muon signatures

Parameters such as $\Delta \phi$, $\Delta \eta$, $\Delta R$, invariant mass of muon pairs can be selected to optimise signal selection / background rejection
Rare decays of B-mesons at ATLAS – Darren Price – June 2nd 2016

**ATLAS Preliminary Simulation**

\[ \int L dt = 100 \text{ pb}^{-1} \quad \sqrt{s} = 8 \text{ TeV} \]

Run 212967 events passing 2MU4

**ATLAS Preliminary**

\[ \Delta R \]

<table>
<thead>
<tr>
<th>Level-1 muon thresholds</th>
<th>Topo cut</th>
<th>Background rejection</th>
</tr>
</thead>
<tbody>
<tr>
<td>2MU4</td>
<td>–</td>
<td>0.00 (baseline)</td>
</tr>
<tr>
<td>MU6.2MU4</td>
<td>–</td>
<td>0.40</td>
</tr>
<tr>
<td>2MU4</td>
<td>1–19</td>
<td>0.50</td>
</tr>
<tr>
<td>2MU4</td>
<td>2–8</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>7–14</td>
<td></td>
</tr>
<tr>
<td>2MU6</td>
<td>–</td>
<td>0.80</td>
</tr>
<tr>
<td>MU6.2MU4</td>
<td>2–8</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>8–13</td>
<td>0.86</td>
</tr>
<tr>
<td>2MU10</td>
<td>–</td>
<td>0.95</td>
</tr>
<tr>
<td>2MU6</td>
<td>2–9</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>8–13</td>
<td></td>
</tr>
</tbody>
</table>

**Signal efficiencies**

\[ B_s^0 \rightarrow \mu\mu, \quad B_s^0 \rightarrow J/\psi\phi, \quad J/\psi \rightarrow \mu\mu, \quad \Upsilon(1S) \rightarrow \mu\mu \]

<table>
<thead>
<tr>
<th></th>
<th>[ \int L dt = 100 \text{ pb}^{-1} \quad \sqrt{s} = 8 \text{ TeV} ]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2MU4</td>
</tr>
<tr>
<td></td>
<td>2MU6</td>
</tr>
<tr>
<td></td>
<td>2MU10</td>
</tr>
<tr>
<td></td>
<td>2MU6</td>
</tr>
</tbody>
</table>

**Level-1 m(\mu\mu) [GeV]**

**Level-1 \Delta R**

**Arbitrary units**

**Level-1 m(\mu\mu) [GeV]**

**Arbitrary units**
Significant signal gain for fixed L1 trigger rates from L1topo implementation
<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_T^B$</td>
<td>Magnitude of the $B$ candidate transverse momentum $\vec{p}_T^B$.</td>
</tr>
<tr>
<td>$\chi_{\text{PV,DV }xy}^2$</td>
<td>Significance of the separation $\Delta x$ between production (i.e. associated PV) and decay (DV) vertices in the transverse projection: $\Delta x_T \cdot \Sigma^{-1} \cdot \Delta x_T$, where $\Sigma^{-1}$ is the covariance matrix.</td>
</tr>
<tr>
<td>$\Delta R$</td>
<td>three-dimensional opening between $\vec{p}<em>T^B$ and $\Delta x$: $\sqrt{\alpha</em>{2D}^2 + \Delta y^2}$</td>
</tr>
<tr>
<td>$</td>
<td>\alpha_{2D}</td>
</tr>
<tr>
<td>$L_{xy}$</td>
<td>Projection of $\Delta x_T$ along the direction of $\vec{p}_T^B$: $(\Delta x_T \cdot \vec{p}_T^B)/</td>
</tr>
<tr>
<td>$IP_{B}^{3D}$</td>
<td>three-dimensional impact parameter of the $B$ candidate to the associated PV.</td>
</tr>
<tr>
<td>$\text{DOCA}_{\mu\mu}$</td>
<td>Distance of closest approach (DOCA) of the two tracks forming the $B$ candidate (three-dimensional).</td>
</tr>
<tr>
<td>$\Delta \phi_{\mu\mu}$</td>
<td>Difference in azimuthal angle between the momenta of the two tracks forming the $B$ candidate.</td>
</tr>
<tr>
<td>$</td>
<td>d_0</td>
</tr>
<tr>
<td>$</td>
<td>d_0</td>
</tr>
<tr>
<td>$p_L^{\text{min}}$</td>
<td>Value of the smaller projection of the momenta of the muon candidates along $\vec{p}_T^B$.</td>
</tr>
<tr>
<td>$I_{0.7}$</td>
<td>Isolation variable defined as ratio of $</td>
</tr>
<tr>
<td>$\text{DOCA}_{\text{xtrk}}$</td>
<td>DOCA of the closest additional track to the decay vertex of the $B$ candidate. Tracks matched to a PV different from the $B$ candidate are excluded.</td>
</tr>
<tr>
<td>$N_{\text{close xtrk}}$</td>
<td>Number of additional tracks compatible with the decay vertex (DV) of the $B$ candidate with $\ln(\chi^2_{\text{strk,DV}}) &lt; 1$. The tracks matched to a PV different from the $B$ candidate are excluded.</td>
</tr>
<tr>
<td>$\chi^2_{\mu,xPV}$</td>
<td>Minimum $\chi^2$ for the compatibility of a muon in the $B$ candidate with a PV different from the one associated with the $B$ candidate.</td>
</tr>
</tbody>
</table>
Correlations on discriminating variables entering the continuum BDT
### Systematic uncertainties entering branching fraction extraction

<table>
<thead>
<tr>
<th>Source of Uncertainty</th>
<th>( \mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-) )</th>
<th>( \mathcal{B}(B^0 \rightarrow \mu^+\mu^-) )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scale uncertainties</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \mathcal{B}(B^+ \rightarrow J/\psi K^+) \times \mathcal{B}(J/\psi \rightarrow \mu\mu) ) branching fractions</td>
<td>3.1%</td>
<td>3.1%</td>
</tr>
<tr>
<td>( B^0_{(s)}/B^+ ) production ratio</td>
<td>8.3%</td>
<td>0</td>
</tr>
<tr>
<td>( B^+ ) yield and ( B^0_{(s)}/B^+ ) efficiency ratio</td>
<td>5.9%</td>
<td>5.9%</td>
</tr>
<tr>
<td>Relative efficiency of continuum-BDT intervals</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>Signal and background model</td>
<td>6%</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total scale uncertainty</strong></td>
<td>16%</td>
<td>11%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Offset uncertainties</strong></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal and background model</td>
<td>( 0.2 \times 10^{-9} )</td>
<td>( 0.7 \times 10^{-10} )</td>
</tr>
</tbody>
</table>