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

Johannes Albrecht$^1$, Alexander Battig$^1$, Titus Mombächer$^1$

$^1$Technische Universität Dortmund

Physics at the Terascale, 27.11.2018
Motivation

LHCb datasets

Analysis strategy
- Background subtraction
- Data/Simulation differences
- Multivariate Analysis
- Optimisation

Background studies

Systematic uncertainties

Result
Motivation

- $B_s^0 \rightarrow \ell^+ \ell^-$ decays probe scalar New Physics.
- $\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-) = (3.0 \pm 0.6^{+0.3}_{-0.2}) \cdot 10^{-9}$ measured by LHCb [arXiv:1703.05747], compatible with the Standard Model (SM).
- Because of $\mu/e$ mass difference, $B_s^0 \rightarrow e^+e^-$ is helicity suppressed by a factor of $\frac{m^2_\mu}{m^2_e} \approx 4 \cdot 10^4$ compared to $B_s^0 \rightarrow \mu^+\mu^-$. 
  - Small New Physics effects, which might smaller than $B_s^0 \rightarrow \mu^+\mu^-$ uncertainty, can be visible.
- Current limit by CDF: $\mathcal{B}(B_s^0 \rightarrow e^+e^-) < 2.8 \cdot 10^{-7}$ [PRD102 (2009)201801].

![Diagram showing branching ratios and experimental limits](arXiv:1703.10160)
LHCb datasets

- LHCb has taken Data from 2011 to 2012 and 2015 to 2018
- 2011 to 2012 (Run 1) about 3 fb$^{-1}$, 2015 to 2018 (Run 2) about 6 fb$^{-1}$
- This analysis: 2011 to 2016, about 5 fb$^{-1}$
Analysis strategy

- Goal: measure upper limit for $\mathcal{B} (B_s^0 \rightarrow e^+ e^-)$ relative to $B^+ \rightarrow K^+ J/\psi (\rightarrow e^+ e^-)$
- Analysis performed on 2011, 2012, 2015 and 2016 LHCb data
- Analysis steps:
  - Cut-based preselection and trigger requirements
  - Multivariate analysis
  - Requirements on particle identification
- Analysis performed blind (exclude 90 % of signal simulation)
Branching ratio

- Measurement performed relative to $B^+ \rightarrow K^+ J/\psi (\rightarrow e^+ e^-)$ to cancel uncertainty on produced $B$ mesons

$$\mathcal{B}(B^0_s \rightarrow e^+ e^-) = \mathcal{B}(B^+ \rightarrow K^+ J/\psi) \cdot \frac{N_{B^0_s \rightarrow e^+ e^-}}{N_{B^+ \rightarrow K^+ J/\psi}} \cdot \frac{\epsilon_{B^+ \rightarrow K^+ J/\psi}}{\epsilon_{B^0_s \rightarrow e^+ e^-}} \cdot \frac{f_u}{f_s} = \alpha \cdot N_{B^0_s \rightarrow e^+ e^-}$$

- This way, $f_q$ and $\mathcal{B}(B^+ \rightarrow K^+ J/\psi)$ are the only external inputs

- Emission of bremsstrahlung relevant for electrons
  - Changes invariant mass shape and kinematics
  - Efficiency of selection might differ

  $\rightarrow$ Efficiencies calculated separately in 12 categories:
  - 4 years
  - 3 categories of bremsstrahlung, where none, one or both electrons have corrections applied
Background subtraction on control channel \( B^+ \rightarrow K^+ J/\psi (\rightarrow e^+ e^-) \)

- Split into two parts
  - Remove partially reconstructed background
- Partially reconstructed background (i.e. \( B^0 \rightarrow K^* (\rightarrow K^+ \pi^-) J/\psi (\rightarrow e^+ e^-) \)) can be removed by constraining the \( e^+ e^- \) pair to the nominal \( J/\psi \) mass

![Graph showing m(K^+J/\psi) distribution](image-url)
sFit of control channel

- After removal of partially reconstructed background the control channel can be fitted
- Fit performed in the categories of bremsstrahlung
- Each category described by two Crystal Ball functions [DESY-F31-86-02], with tail parameters fixed from fits to simulation

Below: Fits to the $B^+ \rightarrow K^+ J/\psi (\rightarrow e^+ e^-)$ data. 2012 data is shown on the left, 2016 data on the right. The background level is found to be negligible.
Data/Simulation differences

- Some variables are not perfectly reproduced in simulation
- For training of the multivariate classifier, good agreement is needed.
  → reweight simulation
- Multivariate approach (GBReweighter, [arxiv: 1608.05806]):
  - Train Boosted Decision Tree (BDT) [Nucl.Instrum.Meth. A543 (2005) 577-584]
  - Weight simulation to lower the BDTs separation power
  - Repeat until no separation is possible anymore
  - Weights are then applied to signal and control channel

![Graphs showing data and simulation comparisons for 2012 and 2016]
Multivariate Analysis - Training

- Train BDT to suppress combinatorial background (random combinations of electrons passing preselection)
- Simulated $B^0_s \rightarrow e^+ e^-$ events are used as signal proxy, the upper mass sideband as background proxy
- Input variables contain
  - isolation variables for the $B$ meson and $e$ tracks
  - topological variables
  - kinematic variables
- Training is done using k-folding
  - Every event is classified by an independent BDT
BDT classifier compared between simulation and data sidebands. 2011/2012 is shown on the left, 2015/2016 on the right.
Optimisation

- BDT selection suppresses combinatorial background
- Requirements on the particle identification (PID) of the detector suppress misidentification backgrounds

→ To ensure optimal selection, requirements on BDT and PID are optimised simultaneously by maximising figure of merit [arXiv: physics:0308063]

\[
FOM = \frac{\epsilon_{B_s^0\to e^+e^-}}{\sqrt{N_{bkg} + 3/2}}
\]

\begin{align*}
\text{ProbNe}(e^\pm) & \quad \text{BDT classifier} \\
FOM & \quad \text{ProbNe}(e^\pm)
\end{align*}

Alexander Battig, TU Dortmund
Background studies

- Physical backgrounds are studied using simulated samples
- Three main categories:
  1. Double misidentification ($B \rightarrow hh'$) strongly suppressed by PID: $5 \cdot 10^{-4}$ events exp.
  2. Partially reconstructed decays ($B \rightarrow he^+ e^-$) ≈ 3 events exp.
  3. Partial reconstruction + misidentification (e.g. $B^0 \rightarrow \pi e\nu$) ≈ 0.4 events exp.

→ Combined pollution by physical backgrounds ($\approx 3.4$) is found to be small in the signal region compared to remaining combinatorial background ($\approx 45$).

![Graphs showing mass distributions for $B^0 \rightarrow K\pi$ and $B_s \rightarrow K_s e^+ e^-$ events before PID requirements.](image-url)
### Systematic uncertainties

- Different sources of systematic uncertainty taken into account
- Uncertainty is still statistically dominated

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Trigger $B^0_s \rightarrow e^+ e^-$</td>
<td>0</td>
<td>2.16</td>
<td>1.67</td>
<td>1.86</td>
<td>2.27</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.49</td>
<td>1.07</td>
<td>0.17</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4.62</td>
<td>5.15</td>
<td>4.16</td>
<td>5.17</td>
</tr>
<tr>
<td>Trigger $B^+ \rightarrow K^+ J/\psi$</td>
<td>0</td>
<td>1.22</td>
<td>2.99</td>
<td>5.40</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>4.82</td>
<td>4.69</td>
<td>2.51</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3.43</td>
<td>3.60</td>
<td>2.08</td>
<td>0.52</td>
</tr>
<tr>
<td>BDT</td>
<td>0</td>
<td>2.22</td>
<td>3.03</td>
<td>4.06</td>
<td>1.37</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1.86</td>
<td>1.87</td>
<td>3.23</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2.29</td>
<td>1.30</td>
<td>0.97</td>
<td>0.39</td>
</tr>
<tr>
<td>PID (binning)</td>
<td>0</td>
<td>3.14</td>
<td>3.21</td>
<td>2.06</td>
<td>2.23</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2.48</td>
<td>2.48</td>
<td>0.82</td>
<td>1.15</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.55</td>
<td>0.38</td>
<td>0.99</td>
<td>1.28</td>
</tr>
<tr>
<td>PID (sWeights)</td>
<td>0</td>
<td>6.09</td>
<td>5.81</td>
<td>1.91</td>
<td>2.48</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>5.11</td>
<td>5.91</td>
<td>2.02</td>
<td>1.81</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4.07</td>
<td>4.65</td>
<td>2.14</td>
<td>1.15</td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
<td>7.53</td>
<td>8.08</td>
<td>7.55</td>
<td>4.22</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>7.69</td>
<td>8.19</td>
<td>4.65</td>
<td>2.33</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>7.37</td>
<td>8.02</td>
<td>5.27</td>
<td>5.44</td>
</tr>
</tbody>
</table>
Analysed dataset corresponds to 5 fb$^{-1}$

Expected upper limit on $\mathcal{B}(B_s^0 \rightarrow e^+ e^-)$ using CLs method \cite{CERN-OPEN-2000-205}

$$\mathcal{B}(B_s^0 \rightarrow e^+ e^-) < (5 - 10) \cdot 10^{-9} @95\% \text{ CL}$$

Reaches interesting sensitivity for theory