Search for $B_s \rightarrow \mu^+ \mu^-$ Decays

Walter Hopkins

Cornell University

Lattice Meets Experiment 2010
Motivation

- $B_s \to \mu^+\mu^-$ can only occur through higher order FCNC diagrams in Standard Model (SM)
- This decay is not only suppressed by the GIM Mechanism but also by helicity
- SM predicts very low rate with little SM background ($B R(B_s \to \mu^+\mu^-) = (3.86 \pm 0.57) \times 10^{-9}$, M. Artuso et al, Eur. Phys. J. C57)
- Super symmetry (SUSY) models predict enhancement of this decay by $\tan\beta^6$
- Clean experimental signature $\to \tau$‘s would have stronger coupling but experimentally difficult
Detector

- Reporting on $3.7 \text{ fb}^{-1}$ CDF result, first shown in Fall 2009
- Secondary vertex ID with excellent Silicon tracker: $\sigma_{p_t/p_t^2} \sim 0.15\%$ and $\sigma_{vtx} = 30 \mu m$
- Muon System
Experimental Challenges

- Large background at hadron collider
  - Must reduce large background around dimuon mass of $m_{B_s} = 5.37$ GeV
  - Analysis requirements: Design an effective discriminant, determine the efficiency for signal, and estimating the background level
Central-Central (CMU) and Central-Forward (CMX) Di-muon Trigger

- **Central**: $p_T > 2.0$ GeV and $|\eta| < 0.6$ – **Forward**: $p_T > 2.2$ GeV and $0.6 < |\eta| < 1.0$
- $p_T$ cuts restrict us to well understood trigger regions

Basic Quality Cuts

- Tracker tracks with hits in 3 silicon layers
- Likelihood and dE/dx based muon Id
- Vertex Quality
- Loose preselection and analysis cuts
  - $p_T(\mu^+\mu^-) > 4.0$ GeV; 3D Decay length significance $> 2$
  - Loose Isolation and opening angle (pointing) cuts

Still background dominated after a reduction of events of 4 orders of magnitude
Analysis Method

- Measure rate of $B_s \rightarrow \mu^+ \mu^-$ relative to $B^+ \rightarrow J/\psi K^+$, $J/\psi \rightarrow \mu^+ \mu^-$
- Apply same selection to find $B^+ \rightarrow J/\psi K^+$
- Systematic uncertainties will cancel in ratio $\Rightarrow$ e.g. dimuon trigger efficiency is the same for both modes
- D0 total $B^+$ yield: $5728 \pm 85$ (with $5 \text{ fb}^{-1}$)

$$BR(B_s \rightarrow \mu^+ \mu^-) = \frac{N_{B_s} \epsilon_{B_s}^{\text{trig}} \epsilon_{B_s}^{\text{reco}}}{N_{B^+} \epsilon_{B^+}^{\text{trig}} \epsilon_{B^+}^{\text{reco}}} \frac{\alpha_{B^+}}{\alpha_{B_s}} \frac{1}{\epsilon_{NN}} \frac{f_\mu}{f_s} \cdot BR(B^+ \rightarrow J/\psi K^+ \rightarrow \mu^+ \mu^- K^+)$$

From Data, From MC, From PDG
Analysis Method

- Estimate acceptances and efficiencies
- Identify variables that discriminate signal and background
- Make multivariate discriminant, for background rejection
  - Optimized with Pythia signal MC and data mass sideband
  - Validate in $B^+$ sample
- Estimate Background
  - Combinatoric background
  - Peaking background: $B \rightarrow hh$
Discriminating Variables (CDF)

- Invariant mass of muons with $2.5\sigma$ window, $\sigma=24$ MeV
- 3 Secondary vertex related variables
  - $\lambda = c\tau$, proper decay time
  - $\frac{\lambda}{\sigma_\lambda}$
  - $\Delta\alpha = |\phi_B - \phi_{vtx}|$
- Isolation: $\frac{p_T(B)}{\sum p_T(trks)+p_T(B)}$
- Transverse momentum of B and lower momentum muon

\[ \Delta \alpha^{3d} [\text{rad}] \]
\[ \lambda^{3d} \text{ significance} \]
Discriminating Variables: Neural Network (CDF)

- Combined all variables except mass in neural network
- Unbiased optimization based on MC signal and sideband data
- Extensively tested for mass bias

CDF Preliminary 3.7 fb\(^{-1}\)

- **Signal MC**
- **Data Sideband**

**Run IIa**

**DØ Run II Preliminary**

Fraction per 0.02

- **B_\text{s} \rightarrow \mu^+ \mu^- Signal**
- **Background (Sidebands)**

Arbitrary Unit

BDT Output
Control Regions (CDF)

- Test background estimates in blinded signal region with independent data samples
- Compare predicted vs. observer background events for multiple NN events

Regions

- OS-: Opposite sign muons with negative proper decay length
- SS+ and SS-: Same sign muons, positive and negative decay length
- FM: OS- & OS+ with one $\mu$ failing muon id and loose vertex cuts

<table>
<thead>
<tr>
<th>sample</th>
<th>NN cut</th>
<th>CMU-CMU</th>
<th>CMU-CMX</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>pred</td>
<td>obsv</td>
</tr>
<tr>
<td>OS-</td>
<td>$0.80 &lt; \nu_{NN} &lt; 0.95$</td>
<td>$275 \pm (9)$</td>
<td>$287$</td>
</tr>
<tr>
<td></td>
<td>$0.95 &lt; \nu_{NN} &lt; 0.995$</td>
<td>$122 \pm (6)$</td>
<td>$121$</td>
</tr>
<tr>
<td></td>
<td>$0.995 &lt; \nu_{NN} &lt; 1.0$</td>
<td>$44 \pm (4)$</td>
<td>$41$</td>
</tr>
<tr>
<td>SS+</td>
<td>$0.80 &lt; \nu_{NN} &lt; 0.95$</td>
<td>$2.7 \pm (0.9)$</td>
<td>$1$</td>
</tr>
<tr>
<td></td>
<td>$0.95 &lt; \nu_{NN} &lt; 0.995$</td>
<td>$1.2 \pm (0.6)$</td>
<td>$0$</td>
</tr>
<tr>
<td></td>
<td>$0.995 &lt; \nu_{NN} &lt; 1.0$</td>
<td>$0.6 \pm (0.4)$</td>
<td>$0$</td>
</tr>
<tr>
<td>SS-</td>
<td>$0.80 &lt; \nu_{NN} &lt; 0.95$</td>
<td>$8.7 \pm (1.6)$</td>
<td>$9$</td>
</tr>
<tr>
<td></td>
<td>$0.95 &lt; \nu_{NN} &lt; 0.995$</td>
<td>$3.0 \pm (1.0)$</td>
<td>$4$</td>
</tr>
<tr>
<td></td>
<td>$0.995 &lt; \nu_{NN} &lt; 1.0$</td>
<td>$0.9 \pm (0.5)$</td>
<td>$0$</td>
</tr>
<tr>
<td>FM+</td>
<td>$0.80 &lt; \nu_{NN} &lt; 0.95$</td>
<td>$169 \pm (7)$</td>
<td>$169$</td>
</tr>
<tr>
<td></td>
<td>$0.95 &lt; \nu_{NN} &lt; 0.995$</td>
<td>$55 \pm (4)$</td>
<td>$43$</td>
</tr>
<tr>
<td></td>
<td>$0.995 &lt; \nu_{NN} &lt; 1.0$</td>
<td>$20 \pm (2)$</td>
<td>$20$</td>
</tr>
</tbody>
</table>
Expected Sensitivities (CDF)

- Single event sensitivity is at SM level \( (= 3.86 \times 10^{-9}) \)
- Largest uncertainty from \( \frac{f_u}{f_s} \)

<table>
<thead>
<tr>
<th></th>
<th>CMU-CMU</th>
<th>CMU-CMX</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{\alpha_{B^+}}{\alpha_{B_s}} )</td>
<td>0.300 ( \pm 0.018 ) ( (\pm 6%) )</td>
<td>0.196 ( \pm 0.0014 ) ( (\pm 7%) )</td>
</tr>
<tr>
<td>( \frac{\epsilon_{B^+}^{trig}}{\epsilon_{B_s}^{trig}} )</td>
<td>0.99935 ( \pm 0.00012 ) ( (-) )</td>
<td>0.97974 ( \pm 0.00016 ) ( (-) )</td>
</tr>
<tr>
<td>( \frac{\epsilon_{B_s}^{reco}}{\epsilon_{B_s}^{reco}} )</td>
<td>0.82 ( \pm 0.03 ) ( (\pm 4%) )</td>
<td>0.83 ( \pm 0.03 ) ( (\pm 4%) )</td>
</tr>
<tr>
<td>( \epsilon_{B_s}^{NN} (NN &gt; 0.80) )</td>
<td>0.776 ( \pm 0.047 ) ( (\pm 6%) )</td>
<td>0.789 ( \pm 0.047 ) ( (\pm 6%) )</td>
</tr>
<tr>
<td>( N_{B^+} )</td>
<td>14300 ( \pm 170 ) ( (\pm 1%) )</td>
<td>5460 ( \pm 110 ) ( (\pm 2%) )</td>
</tr>
<tr>
<td>( f_u/f_s )</td>
<td>3.86 ( \pm 0.59 ) ( (\pm 15%) )</td>
<td>3.86 ( \pm 0.59 ) ( (\pm 15%) )</td>
</tr>
<tr>
<td>( BR(B^+ \to J/\psi K^+ \to \mu^+\mu^- K^+) )</td>
<td>( 5.94 \pm 0.21 \times 10^{-5} ) ( (\pm 4%) )</td>
<td>( 5.94 \pm 0.21 \times 10^{-5} ) ( (\pm 4%) )</td>
</tr>
<tr>
<td>SES (All bins)</td>
<td>( 5.1 \times 10^{-9} ) ( (\pm 18%) )</td>
<td>( 8.5 \times 10^{-9} ) ( (\pm 19%) )</td>
</tr>
<tr>
<td>SES (Combined)</td>
<td>( 3.2 \times 10^{-9} ) ( (\pm 18%) )</td>
<td></td>
</tr>
</tbody>
</table>

**Neural Network**

- 3 NN bins, majority of sensitivity comes from highest bin
- Treated separately \( \rightarrow \) Different Signal/Background
- Lower NN bins added \( \rightarrow \) 50% increase in efficiency and improved sensitivity
- Expected Signal: \( NN>0.8 \) \( \rightarrow \) 1.2 events
Background

- Combinatoric Background
  - Estimated with linear fit to sideband
  - Use p0 and exp fit in highest NN bin for syst. error estimation
- $B \rightarrow hh$
  - Peaks in signal region
  - Use $B_{s(d)} \rightarrow hh$ MC to estimate acceptance and convolute with muon fake rate from data using $D^*$ tagged to $D \rightarrow K\pi$
  - Order of magnitude larger for $B_d$ vs. $B_s$
  - For NN>0.995 in $B_d$ mass window 0.81 events

Walter Hopkins (Cornell University)  
Search for $B_s \rightarrow \mu^+ \mu^-$ Decays  
Lattice Meets Experiment 2010
Dimuon Mass vs NN

CDF at 3.7 fb$^{-1}$

D0 at 2 fb$^{-1}$

Search for $B_s \rightarrow \mu^+ \mu^-$ Decays

Lattice Meets Experiment 2010
• Systematic uncertainties included
• CDF has worlds best limit at $4.3 \times 10^{-8}$ @ 95% CL with 3.7 fb$^{-1}$
• D0 expected sensitivity with 5 fb$^{-1}$: $5.3 \times 10^{-8}$ @ 95% CL
• Last published D0 limit with 2 fb$^{-1}$: $9.3 \times 10^{-8}$ @ 95% CL

<table>
<thead>
<tr>
<th></th>
<th>$\mathcal{B}(B_s \rightarrow \mu^+\mu^-)$</th>
<th>$\mathcal{B}(B_d \rightarrow \mu^+\mu^-)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected $\mathcal{B}$</td>
<td>$2.7 \times 10^{-8}$</td>
<td>$7.2 \times 10^{-9}$</td>
</tr>
<tr>
<td>Observed $\mathcal{B}$</td>
<td>$3.6 \times 10^{-8}$</td>
<td>$6.0 \times 10^{-9}$</td>
</tr>
</tbody>
</table>
CDF working to update and improve analysis for later this year

- More data, up to 6.7 fb$^{-1}$
- Apply improved dE/dx calibration
- Increase acceptance by introducing more detector regions, now better understood
Future: LHCb

- B Focused forward experiment → many boosted B’s
- Will reach SM limits quickly with less luminosity
- Similar discriminating variables

Plot shows for $E_{cm}=14$ TeV
**Conclusion**

**CDF Preliminary Results with 3.7 fb\(^{-1}\)**

\[
\begin{align*}
\mathcal{BR}(B_s \rightarrow \mu^+\mu^-) &= 4.3 \times 10^{-8} \text{ at } 95\% \text{ CL} \\
\mathcal{BR}(B_d \rightarrow \mu^+\mu^-) &= 7.6 \times 10^{-9} \text{ at } 95\% \text{ CL}
\end{align*}
\]

- Reached sensitivity at the \(3.2 \times 10^{-9}\) level
- Set the world’s best limits for both \(B_s\) and \(B_d\) in these modes
- Probing new parameter space across a variety of New Physics models
- D0 updating their analysis with 5 fb\(^{-1}\)
- LHCb projects Tevatron limit with 0.15 fb\(^{-1}\) at \(E_{cm} = 14\) TeV