Search for $B_s^0 \rightarrow \mu^+\mu^-$ at D0

EPS-HEP Conference
19th July 2013

Mark Williams, Indiana University
Searching for the SM?

$B_s^0 \rightarrow \mu^+\mu^-$ decay heavily suppressed in the standard model

- Flavor Changing Neutral Current
- + Helicity suppressed

SM expectation:
$\text{Br}(B_s^0 \rightarrow \mu^+\mu^-) = (3.5 \pm 0.2) \times 10^{-9}$

BSM physics (e.g. SUSY) can significantly affect this decay $\Rightarrow$ new particles in loops

BR can be enhanced by several orders of magnitude, or even suppressed.
Searching for the SM?

$B_s^0 \rightarrow \mu^+\mu^-$ decay heavily suppressed in the standard model

- Flavor Changing Neutral Current
- $+$ Helicity suppressed

SM expectation:

$\text{Br}(B_s^0 \rightarrow \mu^+\mu^-) = (3.5 \pm 0.2) \times 10^{-9}$

BSM physics (e.g. SUSY) can significantly affect this decay $\Rightarrow$ new particles in loops

BR can be enhanced by several orders of magnitude, or even suppressed.
Experimental Status

- Long history of measurements from Tevatron and LHC
- Analyses progressively more advanced

LHCb reported $3\sigma$ evidence, consistent with SM rate (HCP, Nov 2012).

**PRL 110 021801**

---

Getting the most from your data
Ingredients for Success

Aim 1: Maximise expected signal significance / sensitivity

• Understand backgrounds
Ingredients for Success

**Aim 1: Maximise expected signal significance / sensitivity**

- Understand backgrounds

Thick **shielding** before muon system, & **independent magnetized tracking** – heavily suppresses hadronic punchthrough and decay-in-flight

⇒ Backgrounds dominated by real muons from b and c decay

$p\bar{p} \ @ \ \sqrt{s} = 1.96 \ TeV$
Ingredients for Success

Aim 1: Maximise expected signal significance / sensitivity

- Understand backgrounds

Design variables, and train MVA discriminants, based on known backgrounds
Ingredients for Success

Aim 1: Maximise expected signal significance / sensitivity

- Understand backgrounds

Design variables, and train MVA discriminants, based on known backgrounds

Signal

Muons form vertices with other tracks
Ingredients for Success

Aim 1: Maximise expected signal significance / sensitivity

- Understand backgrounds

Design variables, and train MVA discriminants, based on known backgrounds
Ingredients for Success

Aim 1: Maximise expected signal significance / sensitivity

• Understand backgrounds
• Use best possible signal model (MC)
Ingredients for Success

Aim 1: Maximise expected signal significance / sensitivity

• Understand backgrounds
• Use best possible signal model (MC)

Aim 2: Convert search results into BR limit/measurement

• Determine trigger/selection efficiency for signal
• Validate performance in data
Ingredients for Success

**Aim 1: Maximise expected signal significance / sensitivity**
- Understand backgrounds
- Use best possible signal model (MC)

**Aim 2: Convert search results into BR limit/measurement**
- Determine trigger/selection efficiency for signal
- Validate performance in data

Determine BR($B_s^0 \rightarrow \mu\mu$) by comparing search results with $B^+$ yield

Many efficiency uncertainties cancel in the ratio (luminosity, trigger, muon ID…)

Use normalisation mode $B^+ \rightarrow J/\psi K^+ \rightarrow \mu^+\mu^- K^+$
Event Selection

- Single and dimuon triggers
- Loose preselection (same muon requirements for signal and normalisation mode)
- Final selection uses two BDTs, one trained for each main background (30 variables)
- Innovative variables, including e.g. tertiary vertex information
Event Selection

- Single and dimuon triggers
- Loose preselection (same muon requirements for signal and normalisation mode)
- Final selection uses two BDTs, one trained for each main background (30 variables)
- Innovative variables, including e.g. tertiary vertex information
- Verify MC performance using B\(^+\) channel
Single Event Sensitivity

\[ \text{SES} = \text{BR}(B_s^0 \rightarrow \mu\mu) \] which would give 1 signal event in the data, at preselection

Needed to **optimise BDT cuts**, and **convert results into BR limit**

\[
\text{SES} = \frac{1}{N(B^\pm)} \times \frac{\varepsilon(B^\pm)}{\varepsilon(B_s^0)} \times \frac{f(b \rightarrow B^\pm)}{f(b \rightarrow B_s^0)} \times \text{BR}(B^\pm \rightarrow J/\psi K^\pm, J/\psi \rightarrow \mu^+\mu^-)
\]

Yield in normalisation channel

Ratio of efficiencies (from MC)

External constraints (PDG)
Single Event Sensitivity

\[ \text{SES} = \text{BR}(B_s^0 \rightarrow \mu\mu) \text{ which would give 1 signal event in the data, at preselection} \]

Needed to **optimise BDT cuts**, and **convert results into BR limit**

\[
\text{SES} = \frac{1}{N(B^\pm)} \times \frac{\varepsilon(B^\pm)}{\varepsilon(B_s^0)} \times \frac{f(b \rightarrow B^\pm)}{f(b \rightarrow B_s^0)} \times \text{BR}(B^\pm \rightarrow J/\psi K^\pm, J/\psi \rightarrow \mu^+ \mu^-)
\]

- **Yield in normalisation channel**
- **Ratio of efficiencies (from MC)**
- **External constraints (PDG)**

\[ \Rightarrow \text{Crucial to trust signal model in MC} \]
\[ \text{All variables must be examined} \]
Single Event Sensitivity

$SES = \text{BR}(B_s^0 \rightarrow \mu\mu)$ which would give 1 signal event in the data, at preselection

Needed to **optimise BDT cuts**, and convert results into BR limit

$$SES = \frac{1}{N(B^\pm)} \times \frac{\varepsilon(B^\pm)}{\varepsilon(B_s^0)} \times \frac{f(b \rightarrow B^\pm)}{f(b \rightarrow B_s^0)} \times \text{BR}(B^\pm \rightarrow J/\psi K^\pm, J/\psi \rightarrow \mu^+\mu^-)$$

Yield in normalisation channel

Ratio of efficiencies (from MC)

External constraints (PDG)

$SES = 3.36 \times 10^{-10}$

$\Rightarrow 10.4 \pm 1.1$ signal events expected, assuming SM BR (before BDT cuts)
Optimisation

Divide data into three orthogonal samples to train, optimise, and evaluate BDT performance

Training/testing samples further divided into high/low mass regions

Train BDTs

Test + optimise BDTs

Use to determine expected BG yield in signal region

Blinded region: $4.9 < M_{\mu\mu} < 5.8$ GeV

High-mass sideband: dominated by double-b decay – use to train dedicated BDT

Low-mass sideband: dominated by sequential decay – use to train second BDT

$\sigma(\mu\mu) = 125$ MeV
Optimisation

1) Coarse optimisation to maximise expected signal significance

2) Final cuts chosen to give best expected limit

Apply cuts to both BDT output values – rectangle close to optimal shape in 2D distribution.
Expectations…

After applying all cuts, expected yields in signal region:

- \( N(BG, \mu\mu) = 4.0 \pm 1.5 \)
- \( N(BG, B_s^0 \rightarrow K^+K^-) = 0.3 \pm 1.5 \)

(determined using \( K \rightarrow \mu \) ‘fake rate’ from \( B \rightarrow \mu D^0X \), suppressed by shielding)

- \( N(SM B_s^0 \rightarrow \mu^+\mu^-) = 1.23 \pm 0.13 \)

\( \Rightarrow \) Expected limit: \( BR(B_s^0 \rightarrow \mu^+\mu^-) < 23 \times 10^{-9} \)

Results: Control Region

Blinded mass range contains *control regions* either side of signal window

Signal expectation negligible – allows test of method to estimate BG

Before full unblinding, compare expected and observed yields:

Expected: $6.7 \pm 2.6$ events

Observed: 9 events

⇒ BG consistent with expectations

Also tested high/low mass control regions individually
Results: Signal Region

In signal region:

Expected: 4.3 ± 1.6 BG events
1.2 ± 0.1 SM signal events

Observed: 3 events

Observed limit (95% C.L.):

$$\text{BR}(B_s^0 \rightarrow \mu^+\mu^-) < 15 \times 10^{-9}$$
Results: Signal Region

In signal region:

Expected: $4.3 \pm 1.6$ BG events
1.2 ± 0.1 SM signal events

Observed: 3 events

Observed limit (95% C.L.):

$\text{BR}(B_s^0 \rightarrow \mu^+\mu^-) < 15 \times 10^{-9}$

Best observed limit from Tevatron
Results: Signal Region

In signal region:

Expected: \(4.3 \pm 1.6\) BG events

\(1.2 \pm 0.1\) SM signal events

Observed: 3 events

**Observed limit (95\% C.L.):**

\[
\text{BR}(B_s^0 \rightarrow \mu^+\mu^-) < 15 \times 10^{-9}
\]

Best observed limit from Tevatron

Competitive, and consistent, with LHC results
Summary

Final word from D0 on this decay

Significant effort to make the most of our data:

• Targeted BDTs for main backgrounds
• Novel tertiary vertexing variables

Set best Tevatron limit on the BR, in line with results from LHC.

\[ \text{BR}(B_s^0 \rightarrow \mu^+\mu^-) < 15 \times 10^{-9} \]
@ 95% C.L.

PRD 87, 072006 (2013)
Extra Material
Third Track Efficiency

Normalisation channel:

\[ \text{B}^+ \rightarrow \text{J/}\psi \text{ K}^+ \rightarrow \mu\mu\text{K}^+ \]

has third track in addition to two muons;

Important to compare tracking efficiency in MC and data, since this effect won’t cancel in the ratio

Use \( \text{B}^+ \rightarrow \text{J/}\psi \text{K}^*\text{0} \rightarrow \mu\mu\text{K}\pi \)

to compare ‘4th track’ efficiency in data/MC

Data/MC ratio is \(0.88\pm0.06\)

Correction applied for five separate data epochs
Expected vs Observed Yields

In signal region:

Expected: $4.3 \pm 1.6$ BG events

$1.2 \pm 0.1$ SM signal events

Observed: 3 events

Observed limit (95% C.L.):

$\text{BR}(B_s^0 \rightarrow \mu^+\mu^-) < 15 \times 10^{-9}$

Observed yield consistent with expectations over broad range of BDT cuts – no systematic bias