Search for $B_s \to \mu^+\mu^-$ Decays at CDF

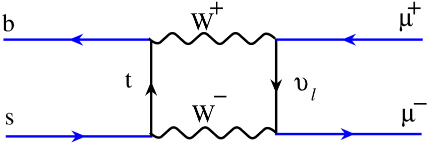
Walter Hopkins

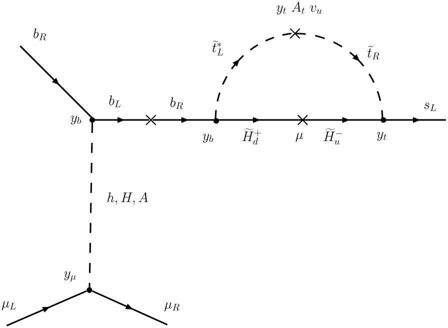
Cornell University

Lattice Meets Experiment 2010

Motivation

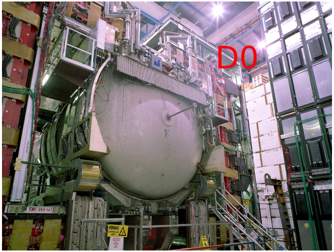
- $B_s \to \mu^+\mu^-$ can only occur through higher order FCNC diagrams in Standard Model (SM)
- This decay is not only suppressed by the GIM Mechanism but also by helicity
- SM predicts very low rate with little SM background ($\mathcal{BR}(B_s \to \mu^+\mu^-) = (3.86 \pm 0.57) \times 10^{-9}$, M.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

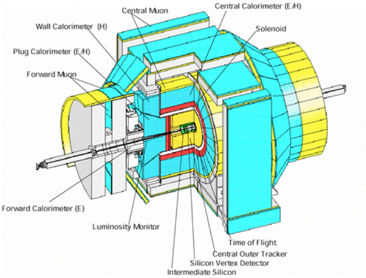


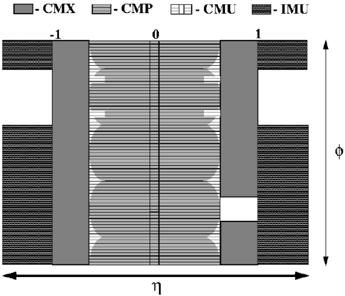


Detector

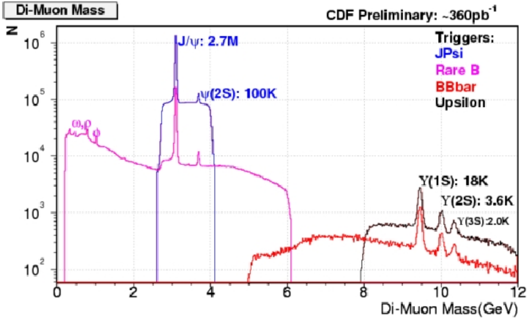
- Reporting on 3.7 fb^{-1} CDF result, first shown in Fall 2009
- Secondary vertex ID with excellent Silicon tracker: $\sigma_{p_t}/p_t^2 \sim 0.15\%$ and $\sigma_{vtx} = 30\mu m$
- Muon System







- 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



Data Sample

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

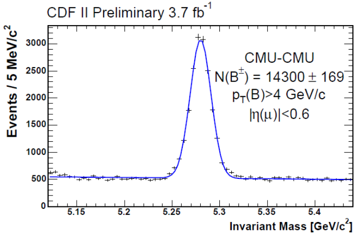
Basic Quality Cuts

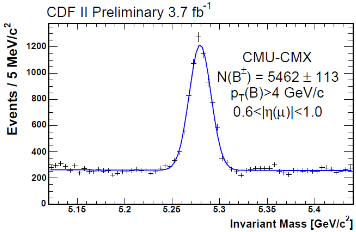
- Tracker tracks with hits in 3 silicon layers
- Likelihood and dE/dx based muon Id
- Vertex Quality
- Loose preselection and analysis cuts
  - $p_T(\mu^+\mu-) > 4.0$ GeV; 3D Decay length significance > 2
  - Loose Isolation and opening angle (pointing) cuts

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

Analysis Method

- Measure rate of $B_s \to \mu^+\mu^-$ relative to $B^+ \to J/\Psi K^+$, $J/\Psi \to \mu^+\mu^-$
- Apply same selection to find $B^+ \to J/\Psi K^+$
- Systematic uncertainties will cancel in ratio ⇒ e.g. dimuon trigger efficiency is the same for both modes
- D0 total B+ yield: 5728±85





$$\mathcal{BR}(B_s \to \mu^+\mu^-) = \boxed{\frac{N_{B_s}}{N_{B^+}} \frac{\epsilon^{trig}_{B^+}}{\epsilon^{trig}_{B_s}} \frac{\epsilon^{reco}_{B^+}}{\epsilon^{reco}_{B_s}}} \boxed{\frac{\alpha_{B^+}}{\alpha_{B_s}} \frac{1}{\epsilon^{NN}_{B_s}}} \boxed{\frac{f_u}{f_s} \cdot \mathcal{BR}(B^+ \to J/\Psi K^+ \to \mu^+\mu^- K^+)}$$

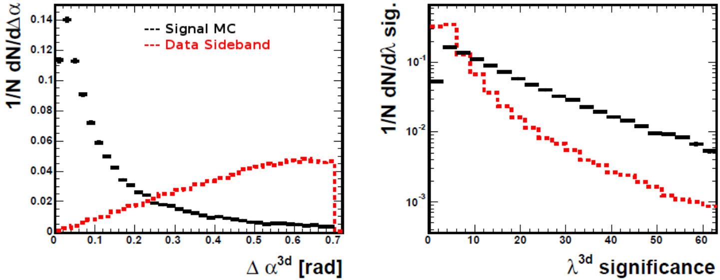
From Data, From MC, From PDG

Analysis Method

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

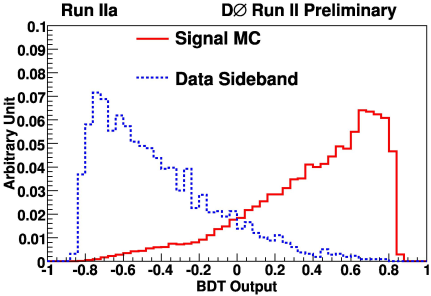
Discriminating Variables

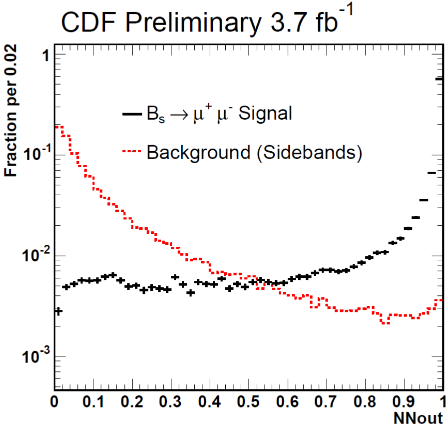
- Invariant mass of muons with 2.5σ window, σ=24 MeV
- 3 Secondary vertex related variables
  - $\lambda = c\tau$, proper decay time
  - $\frac{\lambda}{\sigma_\lambda}$
  - $\Delta\alpha = |\phi_B - \phi_{vtx}|$
- Isolation: $\frac{p_T(B)}{\Sigma p_T(trks)+p_T(B)}$
- Transverse momentum of B and lower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



Discriminating Variables: Neural Network

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





Control Regions

- 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 μ failing muon id and loose vertex cuts

		CMU-CMU			CMU-CMX		
sample	NN cut	pred	obsv	prob(%)	pred	obsv	prob(%)
OS-	$0.80 < \nu_{NN} < 0.95$	$275 \pm (9)$	287	26	$310 \pm (10)$	304	39
	$0.95 < \nu_{NN} < 0.995$	$122 \pm (6)$	121	46	$124 \pm (6)$	148	3.2
	$0.995 < \nu_{NN} < 1.0$	$44 \pm (4)$	41	36	$31 \pm (3)$	50	0.4
SS+	$0.80 < \nu_{NN} < 0.95$	$2.7 \pm (0.9)$	1	29	$2.7 \pm (0.9)$	0	10
	$0.95 < \nu_{NN} < 0.995$	$1.2 \pm (0.6)$	0	34	$1.2 \pm (0.6)$	1	66
	$0.995 < \nu_{NN} < 1.0$	$0.6 \pm (0.4)$	0	55	$0.0 \pm (0.0)$	0	-
SS-	$0.80 < \nu_{NN} < 0.95$	$8.7 \pm (1.6)$	9	49	$5.7 \pm (1.6)$	2	11
	$0.95 < \nu_{NN} < 0.995$	$3.0 \pm (1.0)$	4	36	$3.6 \pm (1.0)$	2	34
	$0.995 < \nu_{NN} < 1.0$	$0.9 \pm (0.5)$	0	43	$0.3 \pm (0.3)$	0	70
FM+	$0.80 < \nu_{NN} < 0.95$	$169 \pm (7)$	169	50	$73 \pm (5)$	64	19
	$0.95 < \nu_{NN} < 0.995$	$55 \pm (4)$	43	9	$19 \pm (2)$	18	49
	$0.995 < \nu_{NN} < 1.0$	$20 \pm (2)$	20	48	$3.6 \pm (1.0)$	3	53

Expected Sensitivities

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

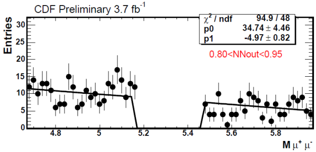
	CMU-CMU		CMU-CMX	
$(\alpha_{B+}/\alpha_{B_s})$	0.300 ± 0.018	$(\pm 6\%)$	0.196 ± 0.0014	$(\pm 7\%)$
$(\epsilon^{trig}_{B+}/\epsilon^{trig}_{B_s})$	0.99935 ± 0.00012	(-)	0.97974 ± 0.00016	(-)
$(\epsilon^{reco}_{B+}/\epsilon^{reco}_{B_s})$	0.82 ± 0.03	$(\pm 4\%)$	0.83 ± 0.03	$(\pm 4\%)$
$\epsilon^{NN}_{B_s}(NN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

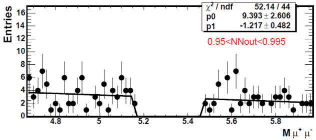
Neural Network

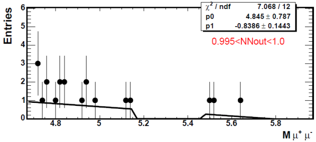
- 3 NN bins, majority of sensitivity comes from highest bin
- Treated separately → Different Signal/Background
- Lower NN bins added → 50% increase in efficiency and improved sensitivity
- Expected Signal: NN>0.8 → 1.2 events

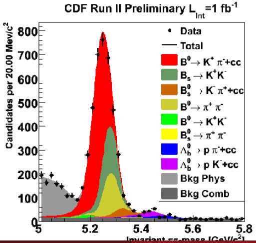
Background

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

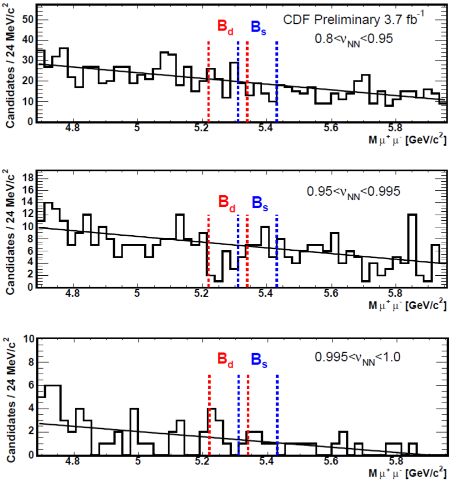


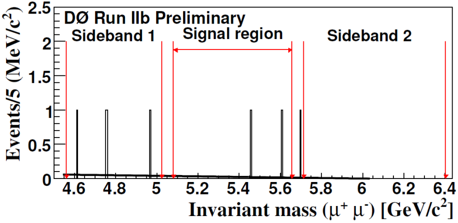






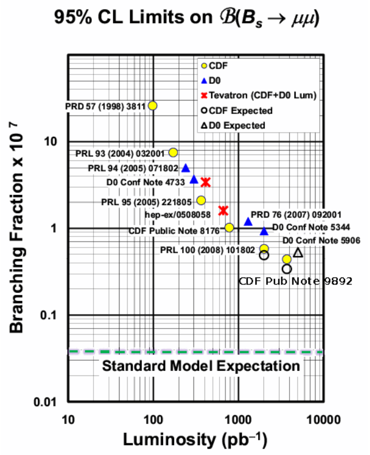
Dimuon Mass vs NN





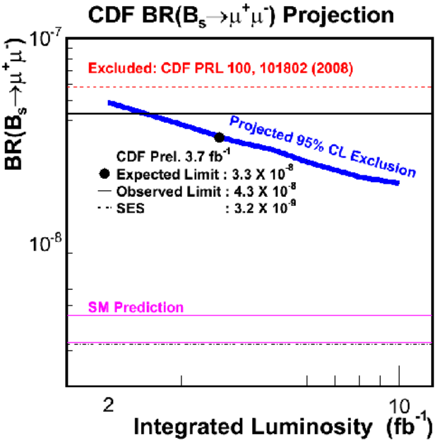
Limits

- Set limits using CLs method
- Cross checked with Bayesian method $\rightarrow$ consistent at 5% level
- Systematic uncertainties included
- CDF has worlds best limit at 4.3×10^{-8} @ 95% CL with 3.7 fb^{-1}
- D0 expected sensitivity with 5 fb^{-1}: 5.3×10^{-8} @ 95% CL
- Last published D0 limit with 2 fb^{-1}: 9.3×10^{-8} @ 95% CL



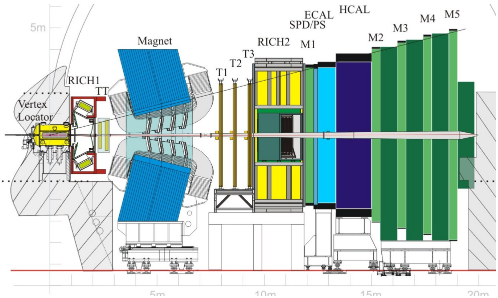
	$\mathcal{B}(B_s \rightarrow \mu^+\mu^-)$ 90%	95%	$\mathcal{B}(B_d \rightarrow \mu^+\mu^-)$ 90%	95%
Expected $\mathcal{B}$	2.7×10^{-8}	3.3×10^{-8}	7.2×10^{-9}	9.1×10^{-9}
Observed $\mathcal{B}$	3.6×10^{-8}	4.3×10^{-8}	6.0×10^{-9}	7.6×10^{-9}

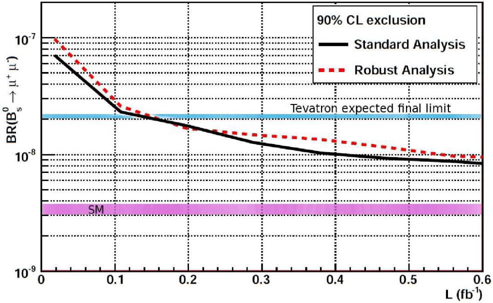
Future: CDF



- 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 $\rightarrow$ many boosted B's
- Will reach SM limits quickly with less luminosity
- Similar discriminating variables

Conclusion

CDF Results with 3.7 fb^{-1}

$\mathcal{BR}(B_s \to \mu^+\mu^-) = 4.3\text{x}10^{-8}$ at 95% CL
$\mathcal{BR}(B_d \to \mu^+\mu^-) = 7.6\text{x}10^{-9}$ at 95% CL

- Reached sensitivity at the 3.2×10^{-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