A Discussion of CDF's Recent $B_s \rightarrow \mu^+\mu^-$ Result

D.Glenzinski
Fermilab

27-Sep-2011

To be clear…

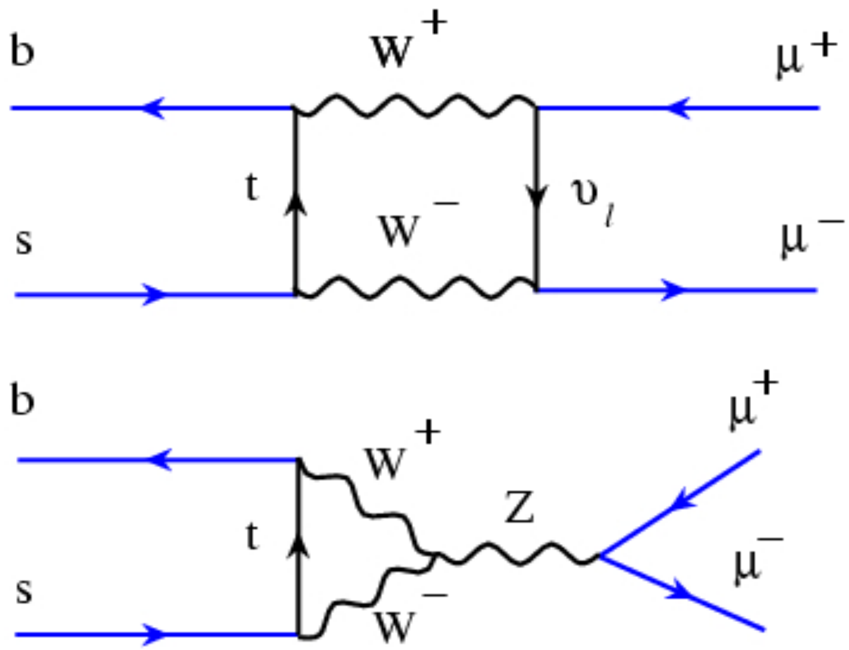
- "Search for $B_s \rightarrow \mu^+\mu^-$ and $B_d \rightarrow \mu^+\mu^-$ Decays with CDF II"
  - arXiv:1107.2304
  - accepted to PRL
- Public web page

  http://www-cdf.fnal.gov/physics/new/bottom/110707.blessed-Bsd2mumu/

Why?

Introduction

- In the SM $B_s \rightarrow \mu^+\mu^-$ is an FCNC… only possible at the loop level

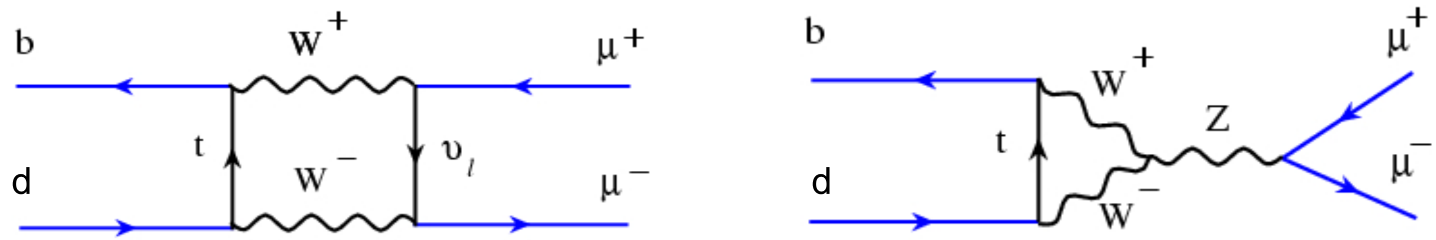


$$BF(B_s \rightarrow \mu^+\mu^-) = (3.2 \pm 0.2) \times 10^{-9}$$

(E.Gamiz et al. (HPQCD Collaboration), A.J. Buras et al.)

Introduction

- All this also true for $B_d \rightarrow \mu^+\mu^-$ decays too

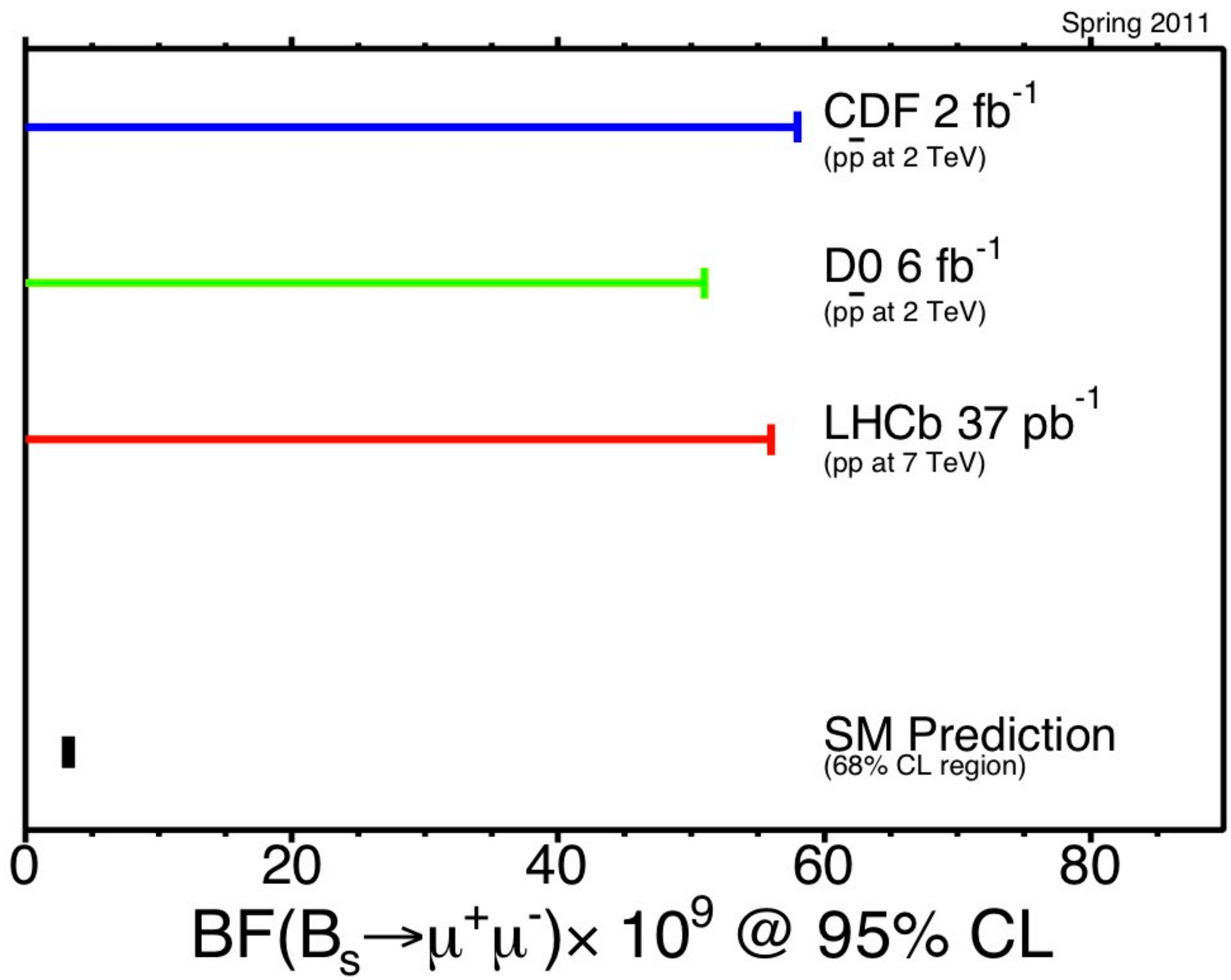


$$BF(B_d \rightarrow \mu^+\mu^-) = (1.0 \pm 0.1) \times 10^{-10}$$

(E.Gamiz et al. (HPQCD Collaboration), A.J. Buras et al.)

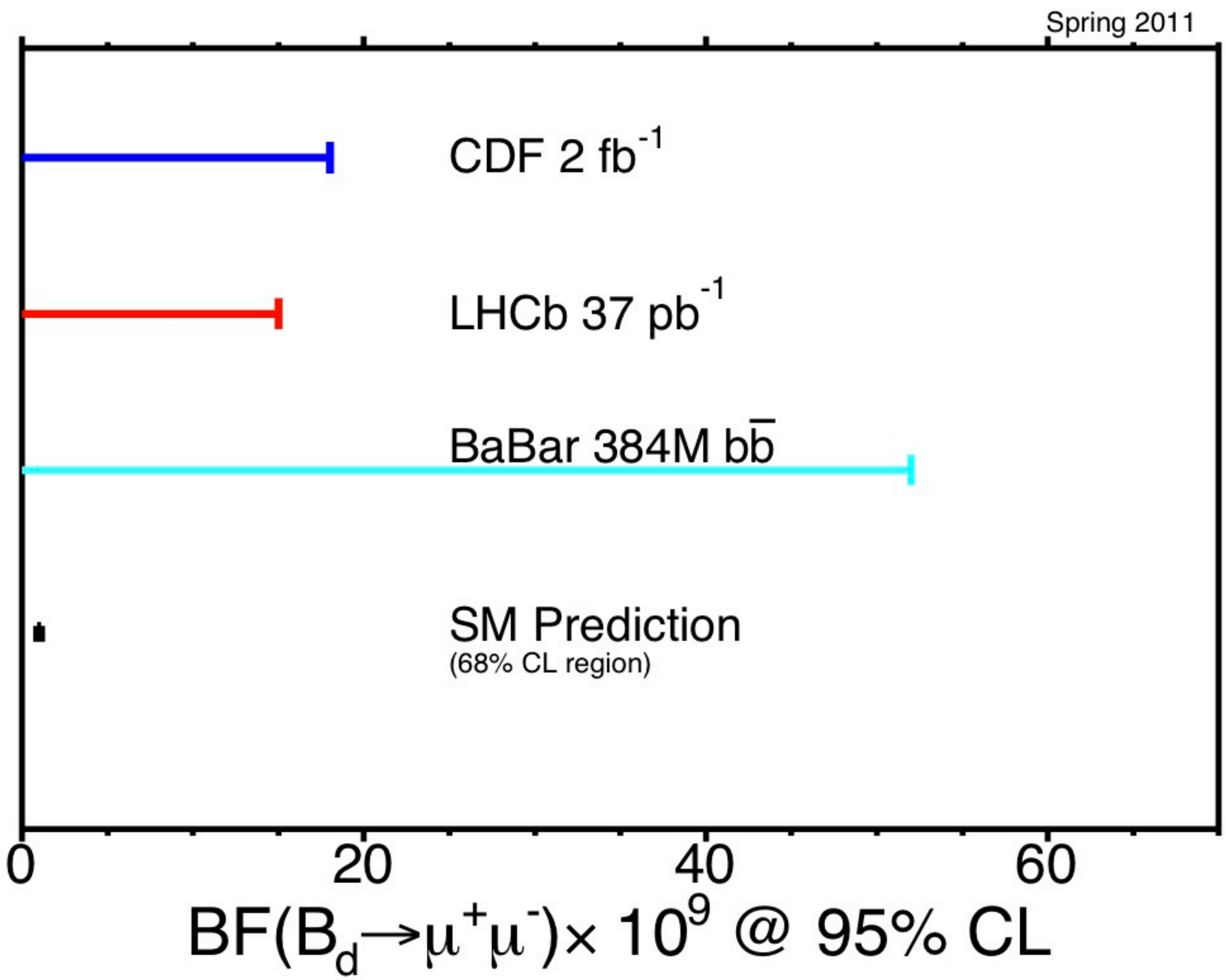
- BF relative to $BF(B_s \rightarrow \mu^+\mu^-)$ model dependent
  - measurements of both sensitive to flavor structure of underlying physics model
  - In MFV models, $BF(B_d/B_s) \sim |V_{td}|^2/|V_{ts}|^2 \sim 1/20$

Experimental Status: Spring 2011



- Has not yet been experimentally observed

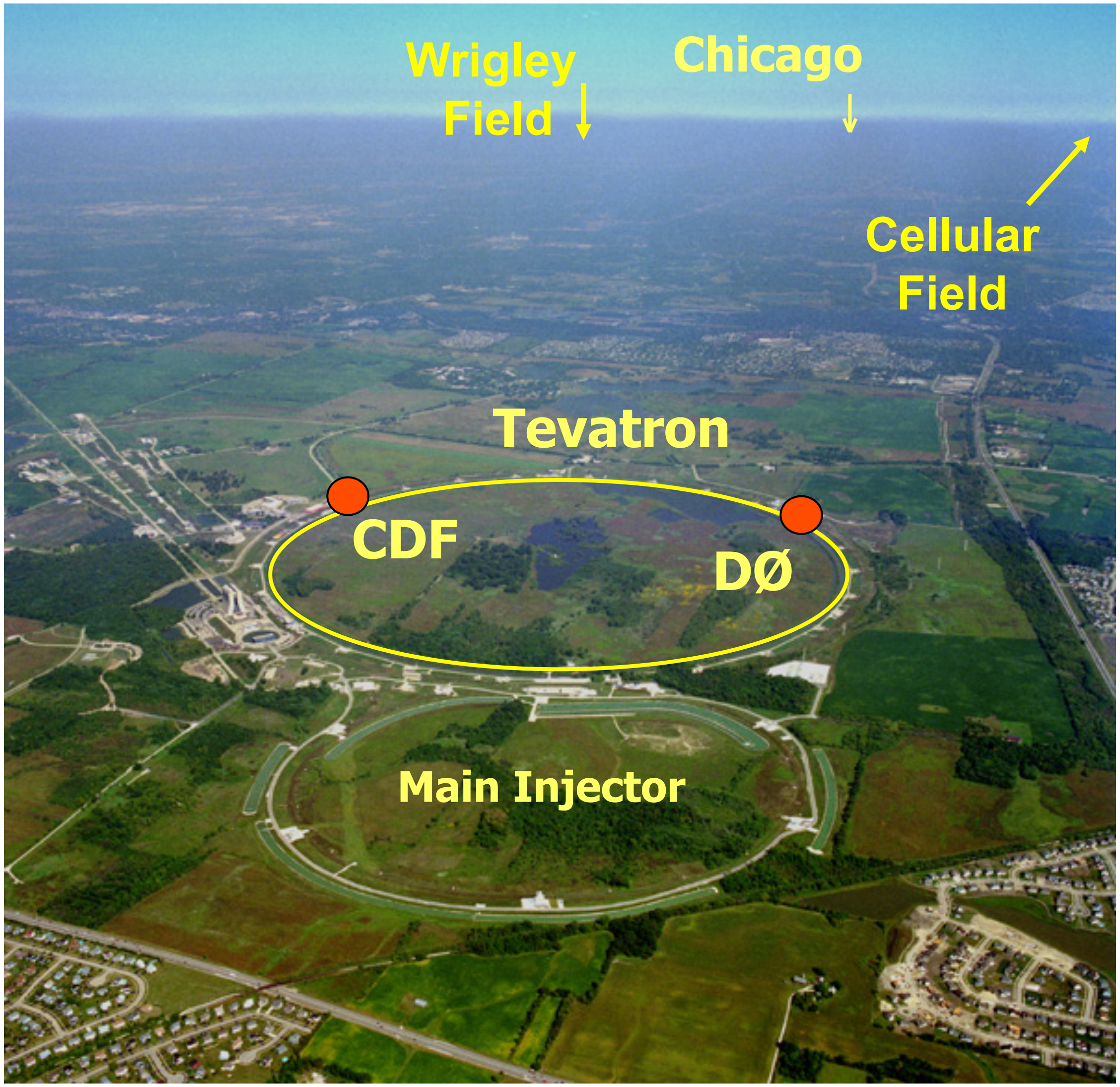
Experimental Status: Spring 2011



- Has not yet been experimentally observed

Where?

Fermilab Tevatron



- $p\bar{p}$ collider at

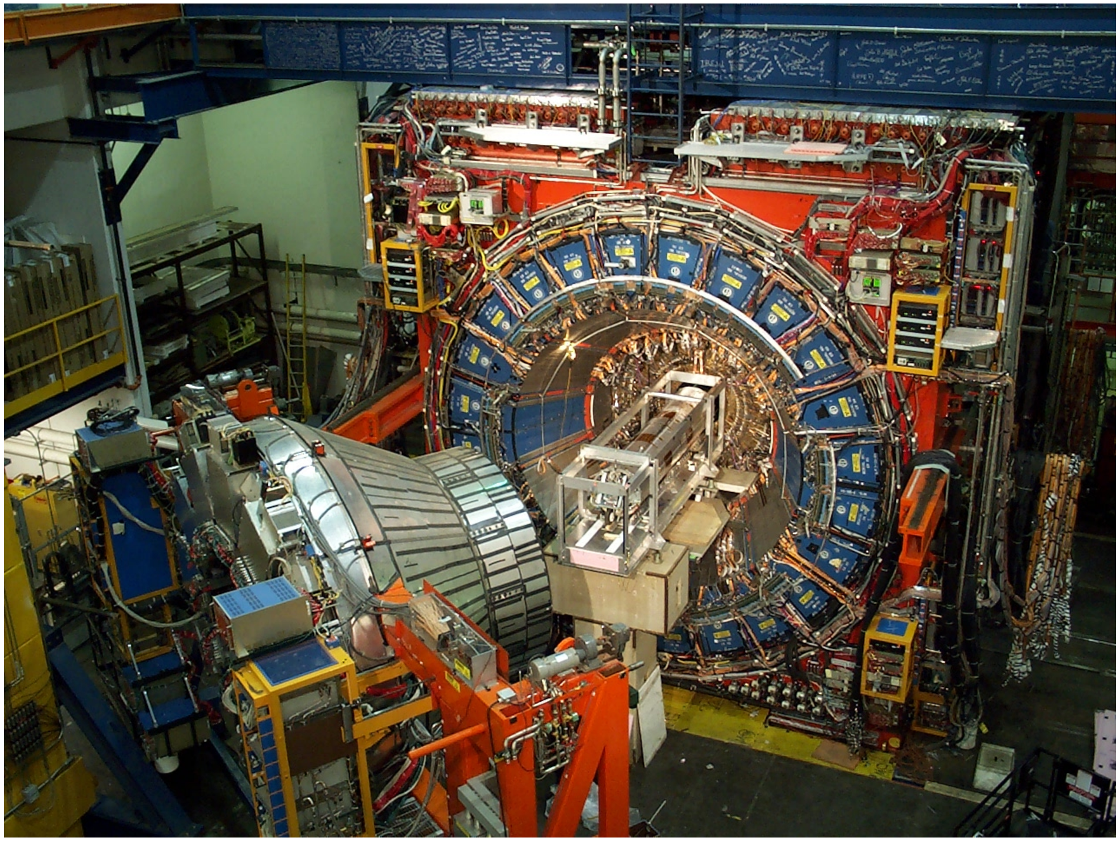
  $E_{cm} = 2$ TeV

- Run-II 2001-2011
  (12 fb^{-1} / exp delivered)

- Performing excellently

- All B-hadron species copiously produced

The CDF Experiment

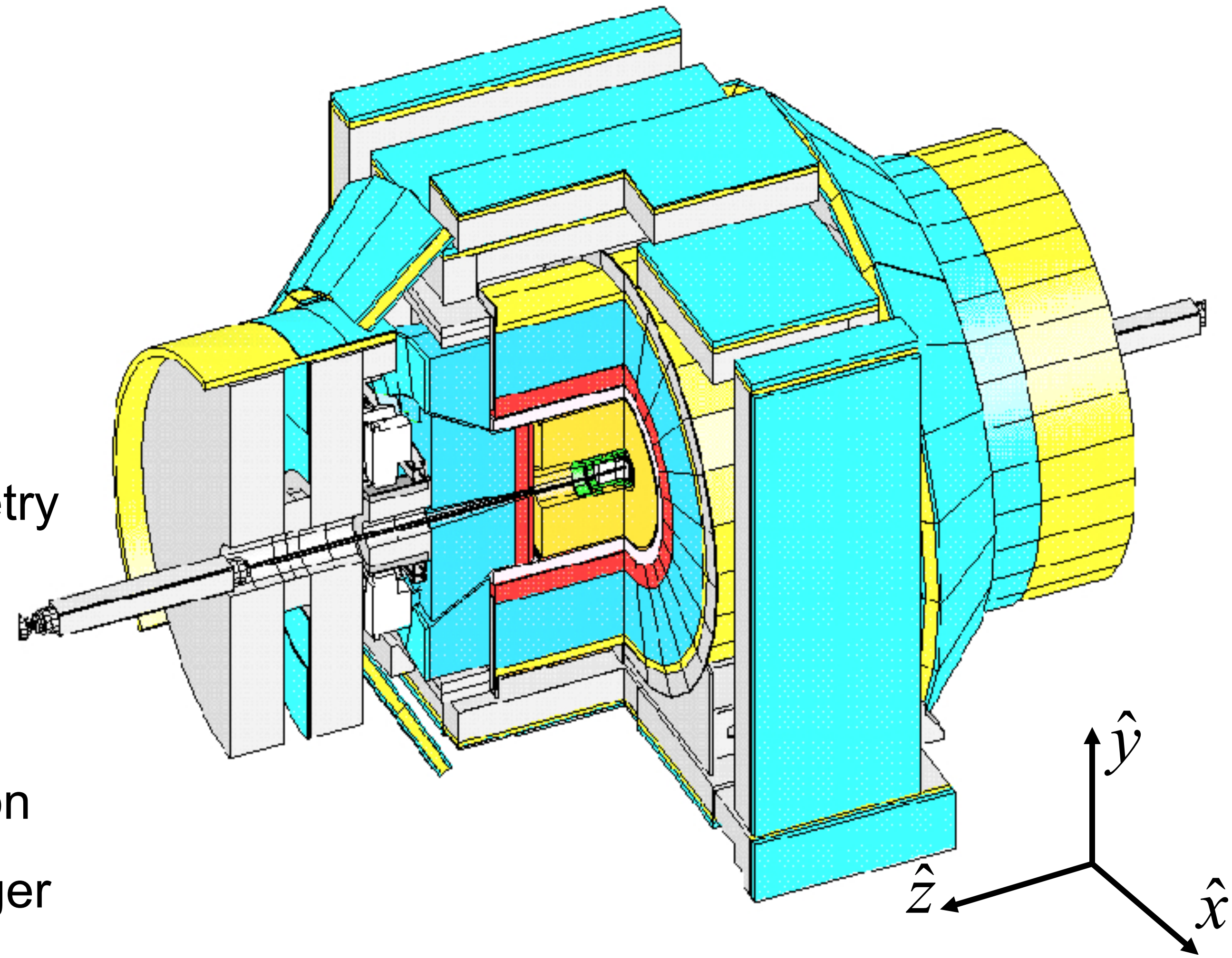


- Multipurpose collider detector
- Pioneered silicon detectors at a hadron collider
- International collaboration, 600+ members

CDF Detector

Features:

- Precision silicon vertexing
- Large radius drift chamber (r=1.4m)
- 1.4 T solenoid
- projective calorimetry ($|\eta| < 3.5$)
- muon chambers ($|\eta| < 1.0$)
- Particle identification
- Silicon Vertex Trigger



$$\eta = -\ln(\tan(\theta/2))$$

How?

Analysis Description

- This is a simple analysis
  1) Find events with 2 muons in them
  2) Identify means to suppress background while keeping as much signal as possible
  3) Look for a bump in the $m_{\mu\mu}$ distribution

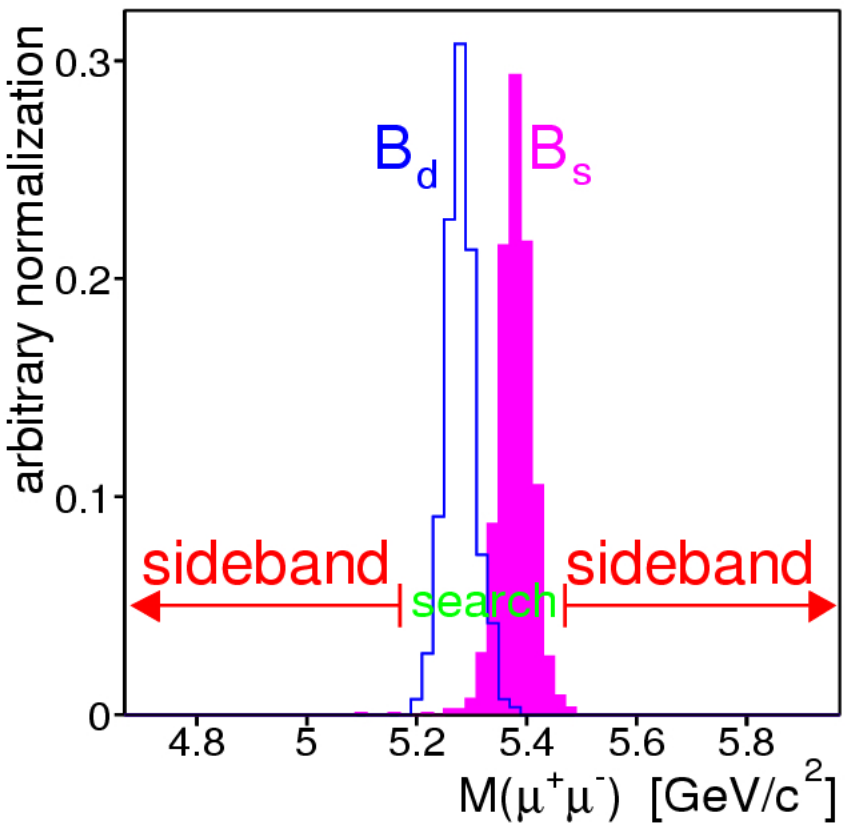
Analysis Strategy

- Our strategy is simple
  - "blind" ourselves to an extended mass signal region
  - Use data in the mass sidebands to estimate the dominant backgnd contribution to the signal region
  - Employ an a priori optimization to choose our final selection criteria
  - Build confidence in background estimate using control regions prior to "opening the box"

Emphasis on being robust and unbiased

Definition of Signal / Sideband Regions

- We "blind" the data in an extended signal region



Search Region:

- $5.169 < m_{\mu\mu} < 5.469$ GeV/c^2
- corresponds to +/- $4\sigma(m_{\mu\mu})$
- final region +/- $2.5\sigma(m_{\mu\mu})$

Sideband Regions:

- additional 0.5 GeV on either side of search region
- used to understand Bkgd

Some Preliminaries

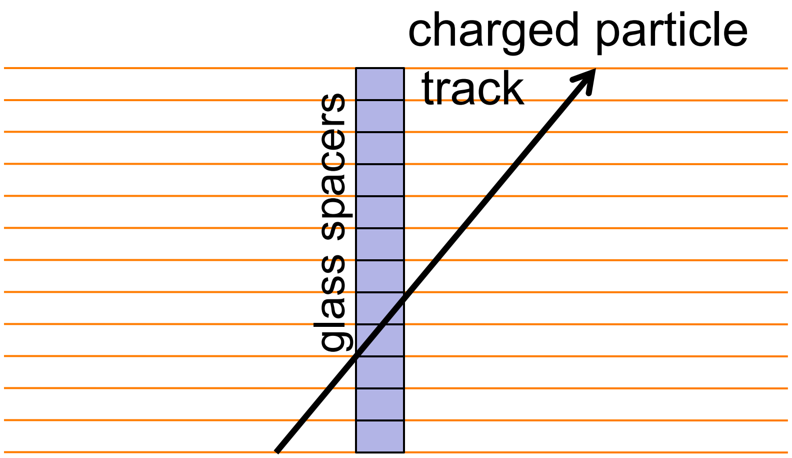
CDF Run-II $B_s \rightarrow \mu^+\mu^-$ Publications

1) 170 pb^{-1} PRL 93, 032001 (2004). 78 citations
   – Sensitivity x3 improvement over Run I
2) 350 pb^{-1} PRL 95, 221805 (2005). 50 citations
   – Sensitivity x4 improvement over 1)
3) 2 fb^{-1} PRL 100, 101802 (2008). 197 citations
   – Sensitivity x4 improvement over 2)
4) 7fb^{-1} accepted PRL, arXiv:1107.2304
   – Sensitivity x3 improvement over 3)

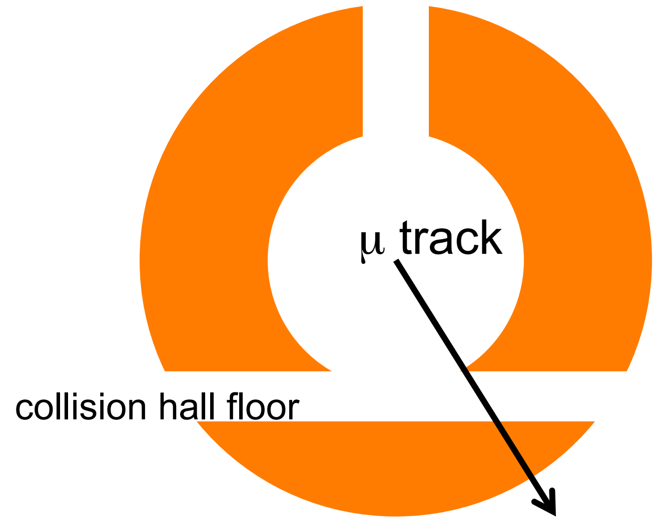
Sensitivity Improvements

- Added acceptance
  - Include events that pass near COT “spacer”
  - Include events in CMX “mini-skirts”
- Improved background discrimination
  - Improved dE/dX calibrations
  - Improved performance of multi-variate discriminant used in final selection criteria

Additional Acceptance



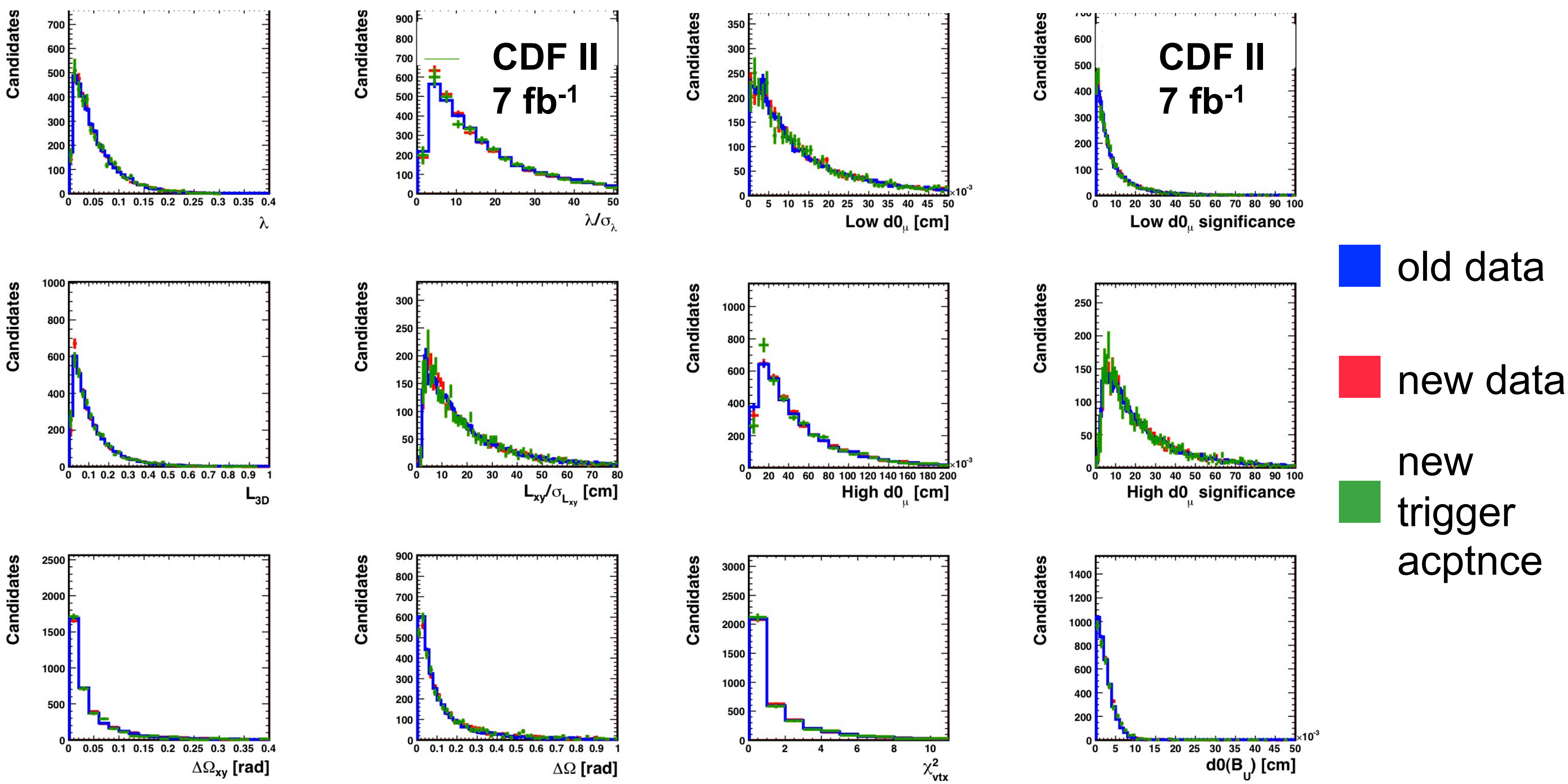
CDF Tracking Drift Chamber - COT
(not to scale)



CDF Central Muon Extension – CMX
(cartoon – not to scale)

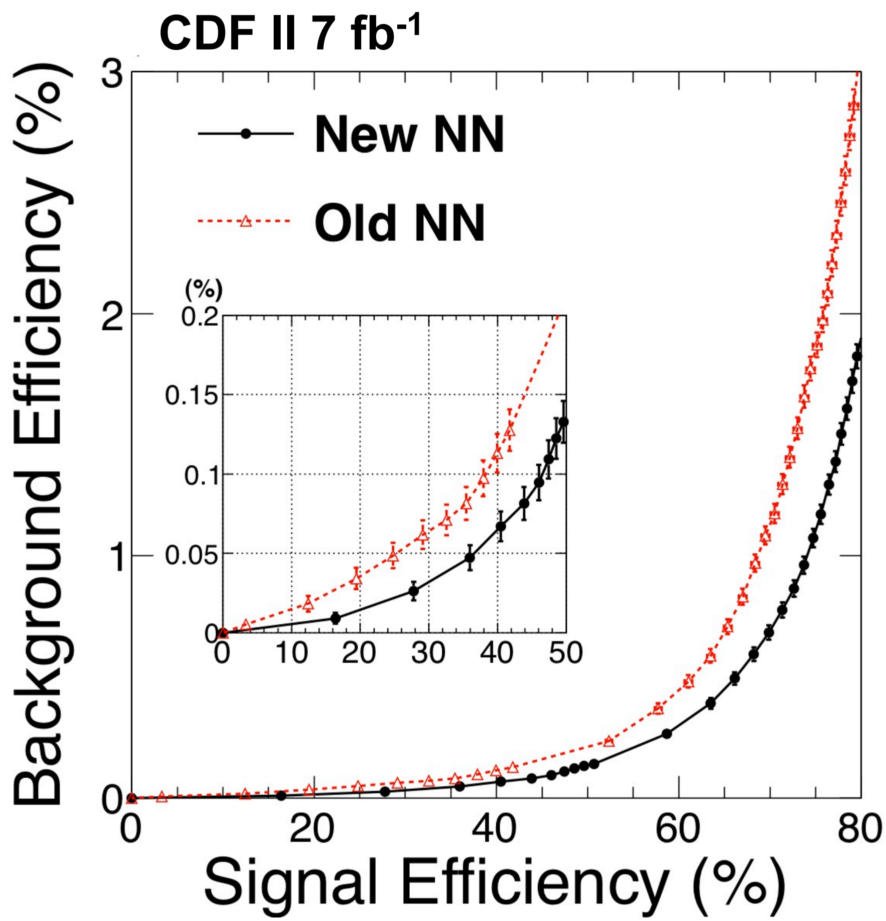
- We've always had these events on tape
- Just needed to understand the trigger efficiencies in these regions

Additional Acceptance



- Kinematics unaffected

Improved Background Discrimination



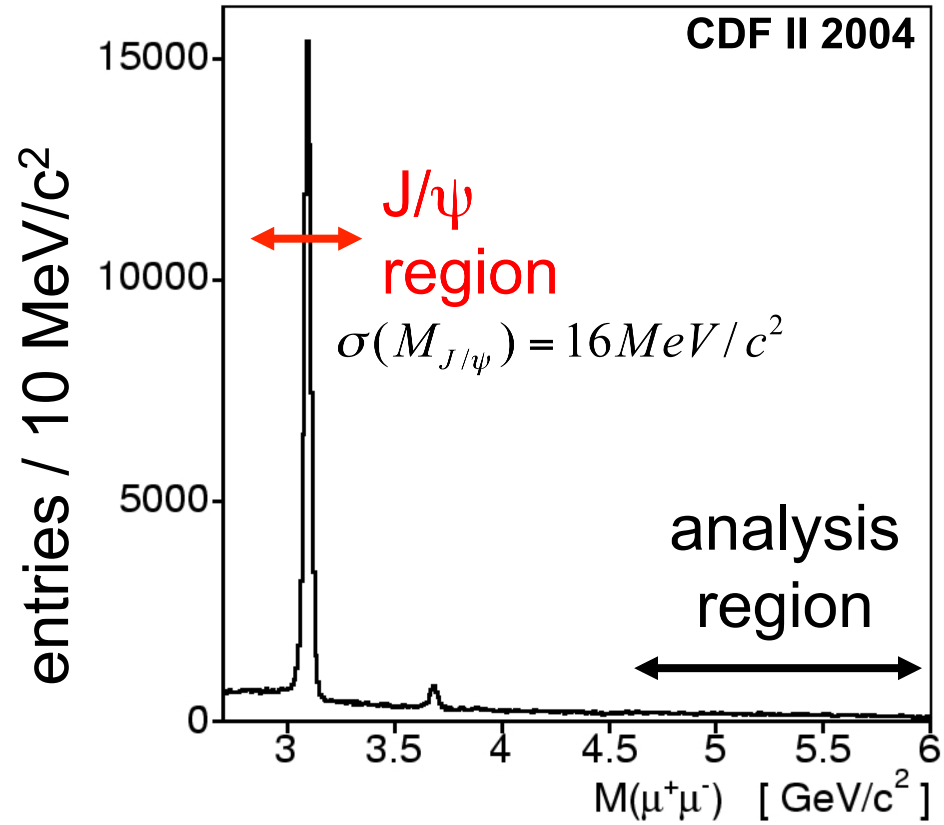
- Improved Neural Net (NN) performance

Normalization

- We employ a relative normalization
  - Using $B^+ \rightarrow J/\psi K^+ \rightarrow \mu^+\mu^- K^+$ events
  - Collect B^+ and signal events with same trigger
  - Many exp. uncertainties significantly reduced

$$BF(B_{s,d} \rightarrow \mu^+\mu^-) = \left(\frac{N_{Bs,d}}{N_{B+}}\right)\left(\frac{\alpha_{B+}\varepsilon_{B+}}{\alpha_{Bs,d}\varepsilon_{Bs,d}}\right)\left(\frac{f_u}{f_s}\right) BF(B^+ \rightarrow J/\psi K^+)$$

Di-muon Mass Distribution from Trigger



- The trigger paths used for this analysis
  - Collect signal sample: $B \rightarrow \mu^+\mu^-$
  - Collect control sample: $J/\psi \rightarrow \mu^+\mu^-$, $B^+ \rightarrow J/\psi\ K^+$

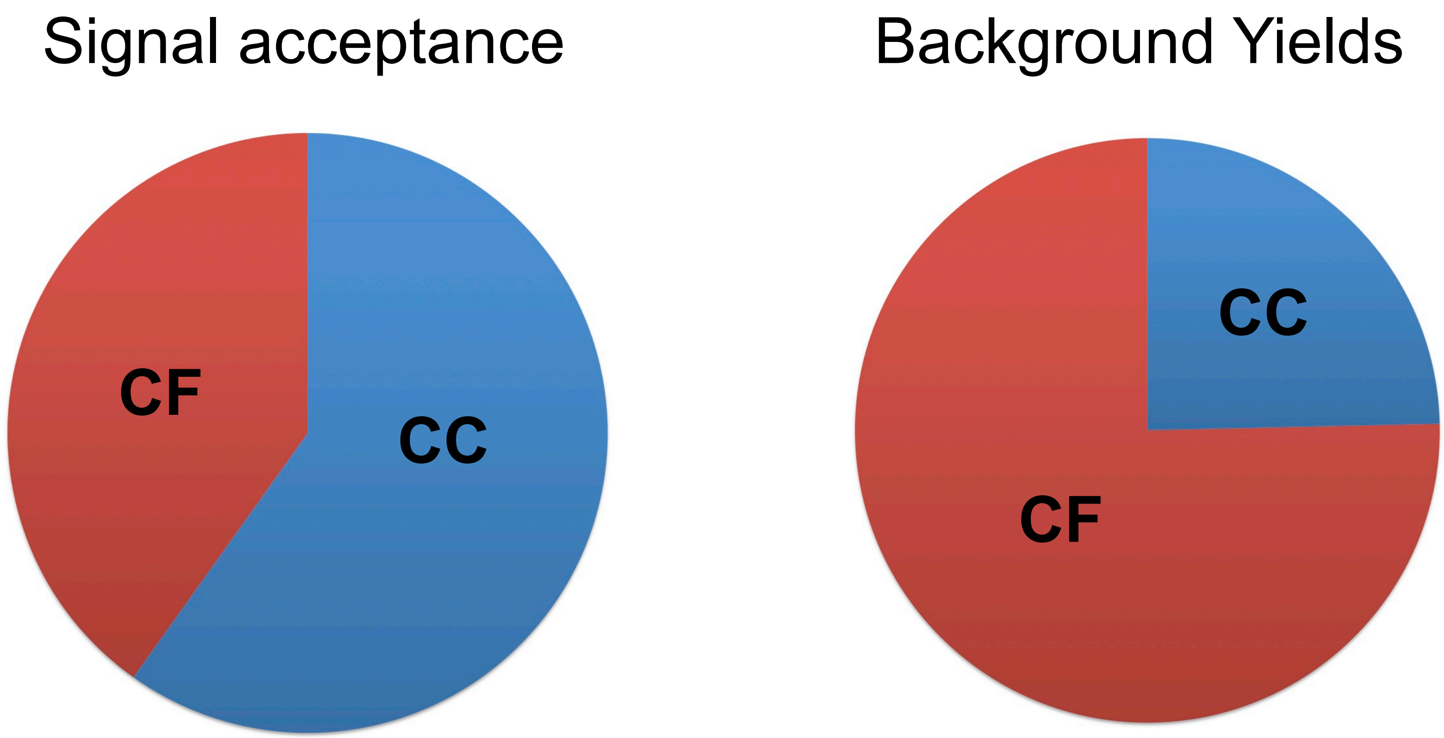
Our Trigger Paths

- Collect data using two separate trigger paths corresponding to two separate topologies:

- "Central-Central" (CC)
  - both muons $|\eta| < 0.6$
  - $P_T(\mu) > 1.5$ GeV/c
  - $2.7 < m_{\mu\mu} < 6.0$ GeV/c^2
  - $\Delta\phi(\mu\mu) < 2.25$ rad
  - $P_T(\mu^+) + P_T(\mu^-) > 5$ GeV/c

- "Central-Forward" (CF)
  - $|\eta_{\mu 1}|<0.6$, $0.6<|\eta_{\mu 2}|<1$
  - $P_T(C) > 1.5$ GeV/c
  - $P_T(F) > 2.0$ GeV/c
  - $2.7 < m_{\mu\mu} < 6.0$ GeV/c^2
  - $\Delta\phi(\mu\mu) < 2.25$ rad
  - $P_T(\mu^+) + P_T(\mu^-) > 5$ GeV/c

CC vs CF Channels



- Treat each channel separately, combine at end

Normalization

- We employ a relative normalization

$$BF(B_{s,d} \to \mu^+\mu^-) = \left(\frac{N_{Bs,d}}{N_{B+}}\right)\left(\frac{\alpha_{B+}\varepsilon_{B+}}{\alpha_{Bs,d}\varepsilon_{Bs,d}}\right)\left(\frac{f_u}{f_s}\right)BF(B^+ \to J/\psi K^+)$$

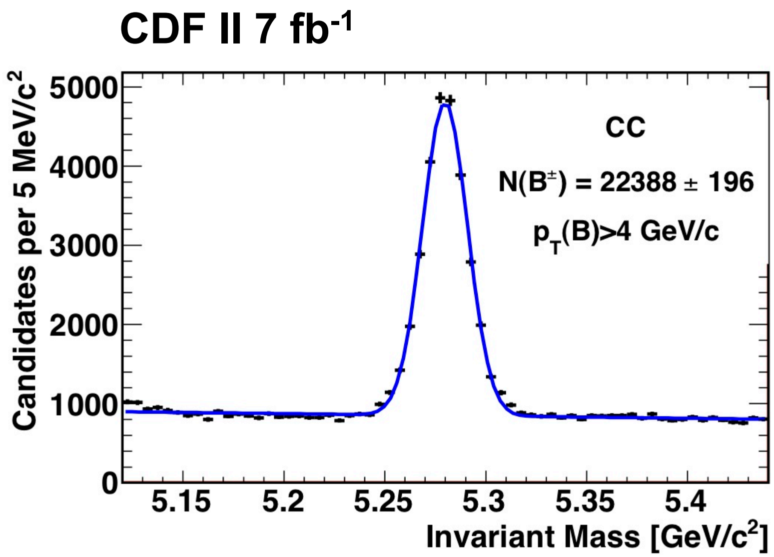
From fits to the data.

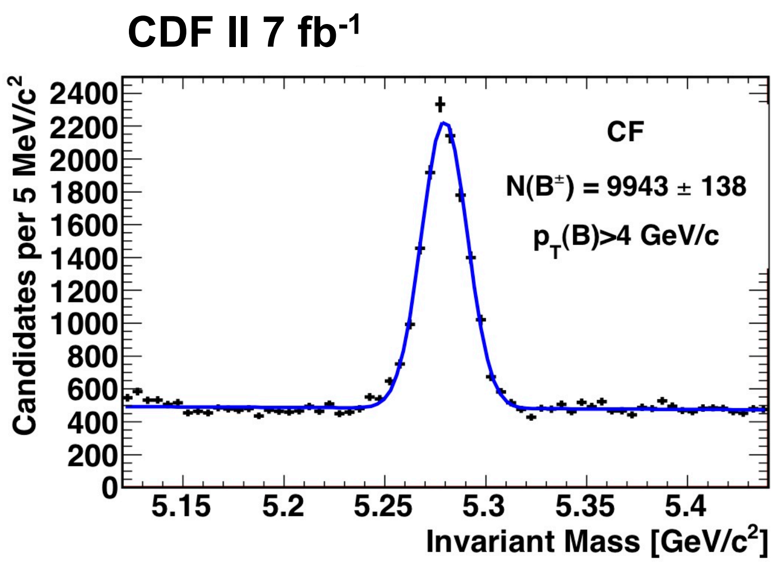
From the PDG 2010:

$$\frac{f_u}{f_s} = 3.55 \pm 0.47$$

$$BF(B^+ \to J/\psi K^+)BF(J/\psi \to \mu^+\mu^-) = (6.01 \pm 0.21) \times 10^{-5}$$

Normalization: B+ Yield





- Use sideband subtracted signal yields
- $B^+ \rightarrow J/\psi\pi^+$ contamination <1%

Normalization

- We employ a relative normalization

$$BF(B_{s,d} \to \mu^+\mu^-) = \left(\frac{N_{Bs,d}}{N_{B+}}\right)\left(\frac{\alpha_{B+}\varepsilon_{B+}}{\alpha_{Bs,d}\varepsilon_{Bs,d}}\right)\left(\frac{f_u}{f_s}\right)BF(B^+ \to J/\psi K^+)$$

$\alpha_B \equiv$ geometric and kinematic acceptance of trigger
(from MC simulation)

$$\varepsilon_B \equiv \varepsilon_{\mathrm{reco}} \cdot \varepsilon_{\mathrm{NN}} \cdot \varepsilon_{\mathrm{mass}} = \left(\varepsilon_{\mathrm{track}} \cdot \varepsilon_{\mu\text{-ID}} \cdot \varepsilon_{\mathrm{vertex}}\right) \cdot \varepsilon_{\mathrm{NN}} \cdot \varepsilon_{\mathrm{mass}}$$

From data using "Tag and Probe"

From MC, checked with B^+ and J/ψ data

Normalization

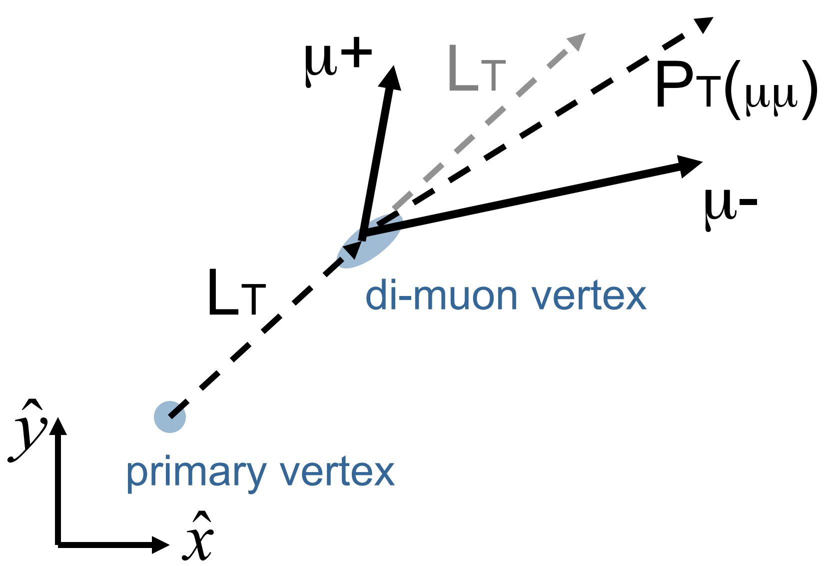
$$BF(B_{s,d} \to \mu^+\mu^-) = \mathrm{N}_{\mathrm{Bs,d}} \cdot ses$$

	CC		CF	
$(\alpha_{B^+}/\alpha_{B_s})$	0.307 ± 0.018	$(\pm 6\%)$	0.197 ± 0.014	$(\pm 7\%)$
$(\epsilon^{trig}_{B^+}/\epsilon^{trig}_{B_s})$	0.99935 ± 0.00012	$(< 1\%)$	0.97974 ± 0.00016	$(< 1\%)$
$(\epsilon^{reco}_{B^+}/\epsilon^{reco}_{B_s})$	0.85 ± 0.06	$(\pm 8\%)$	0.84 ± 0.06	$(\pm 9\%)$
$\epsilon^{NN}_{B_s}(NN > 0.70)$	0.915 ± 0.042	$(\pm 4\%)$	0.864 ± 0.040	$(\pm 4\%)$
$\epsilon^{NN}_{B_s}(NN > 0.995)$	0.461 ± 0.021	$(\pm 5\%)$	0.468 ± 0.022	$(\pm 5\%)$
N_{B^+}	22388 ± 196	$(\pm 1\%)$	9943 ± 138	$(\pm 1\%)$
f_u/f_s	3.55 +/- 0.47	$(\pm 13\%)$	3.55 +/- 0.47	$(\pm 13\%)$
$BR(B^+ \to J/\psi K^+ \to \mu^+\mu^- K^+)$	$(6.01 \pm 0.21) \times 10^{-5}$	$(\pm 4\%)$	$(6.01 \pm 0.21) \times 10^{-5}$	$(\pm 4\%)$
SES (All bins)	$(2.9 \pm 0.5) \times 10^{-9}$	$(\pm 18\%)$	$(4.0 \pm 0.7) \times 10^{-9}$	$(\pm 18\%)$

ses(CC+CF) = 1.7×10^{-9}

- Uncertainty includes: variations in the $p_T(B)$ spectrum, kinematic differences between J/ψ and $B_s \to \mu\mu$, variations in simulation parameters, differences between B^+ data and MC

Some Definitions



3D and 2D versions of variables
2D denoted with subscript "T"

- P(B) = momentum of B

$$P_B = P_{\mu\mu} = \vec{P}_{\mu+} + \vec{P}_{\mu-}$$

- L: decay length

$$L = \vec{L} \cdot \vec{P}_{\mu\mu} / \left|\vec{P}_{\mu\mu}\right|$$

- λ: proper decay time

$$\lambda = \mathrm{cLm}_{\mu\mu} / P_{\mu\mu}$$

- ΔΩ = pointing angle

$$\Delta\Omega = \angle(\vec{L}, \vec{P}_{\mu\mu})$$

The Details…

Analysis Description

- This is a simple analysis
  1) ✓ Find events with 2 muons in them
  2) Identify means to suppress background while keeping as much signal as possible
  3) Look for a bump in the $m_{\mu\mu}$ distribution

Suppress Background, Keep Signal

- We start with some simple “baseline” requirements to ensure two good muons that originate from a common vertex

- Then we exploit features of our signal events to discriminate signal from background

Baseline Requirements

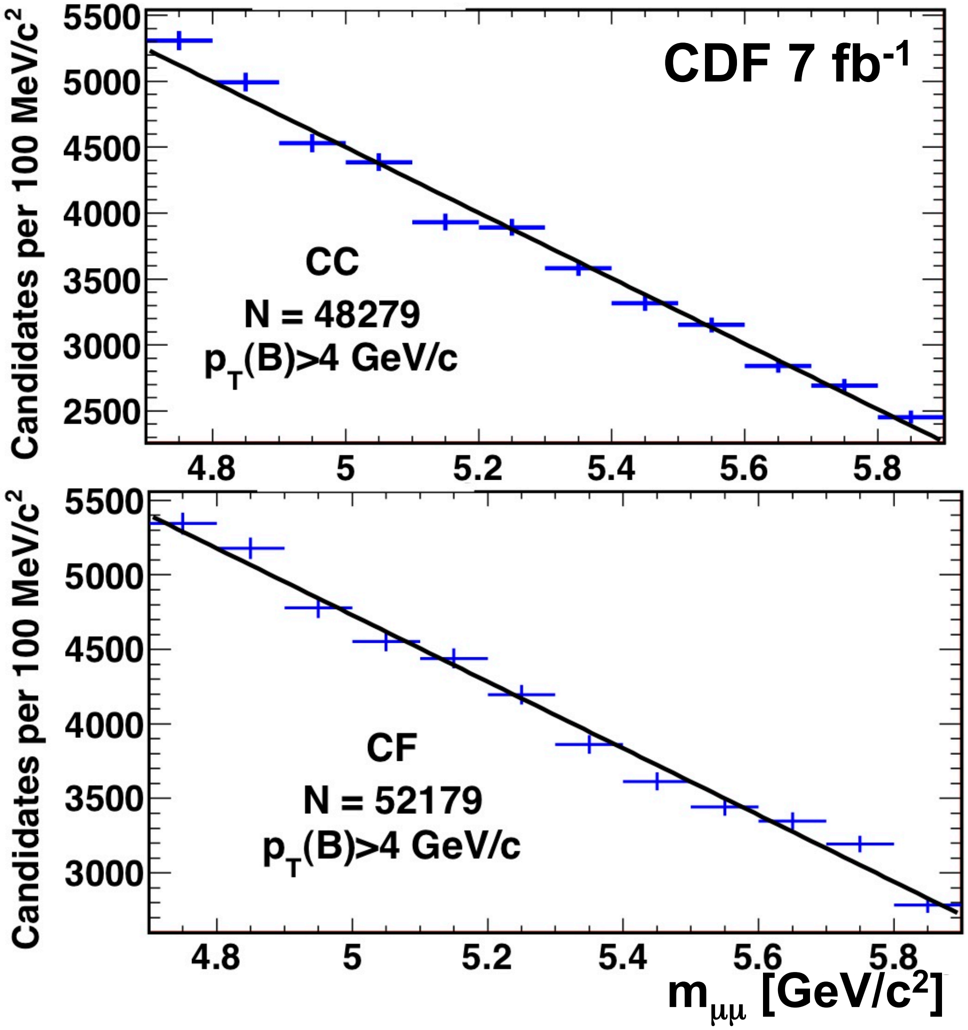
We require:

- "good" COT tracks and μ track-stubs
- >=3 silicon r-ϕ hits
- $4.669 < m_{\mu\mu} < 5.969$ GeV/c^2
- "good" vertex
  - $\sigma(L)<150$ μm
  - $\chi^2 < 15$
  - $L_T < 1$ cm

- $P_T(C)>2.0$, $P_T(F)>2.2$ GeV/c
- $P_T(\mu\mu) > 4$ GeV/c
- $\lambda < 0.3$ cm
- $\lambda/ \sigma_\lambda > 2$
- $\Delta\Omega < 0.70$ rad
- Isolation > 0.50

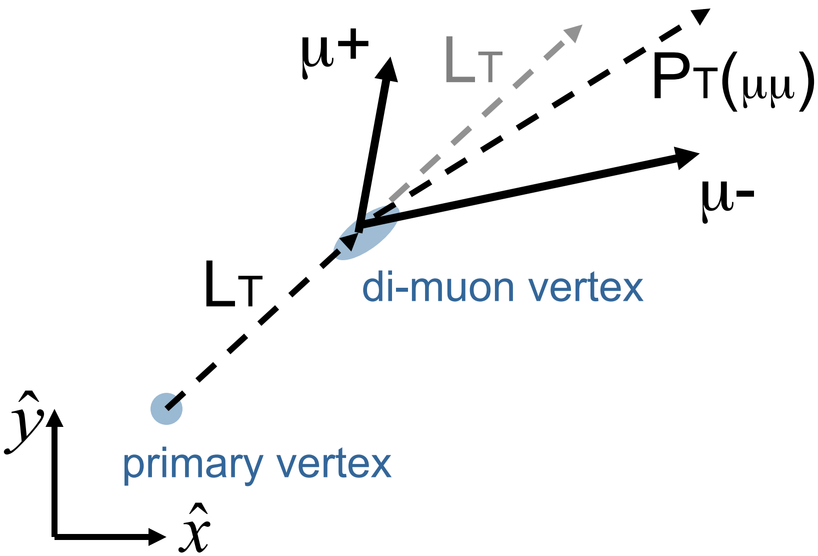
maintain most the signal while significantly reducing bgd

Baseline sample



- Completely background dominated

Discriminate Signal from Background



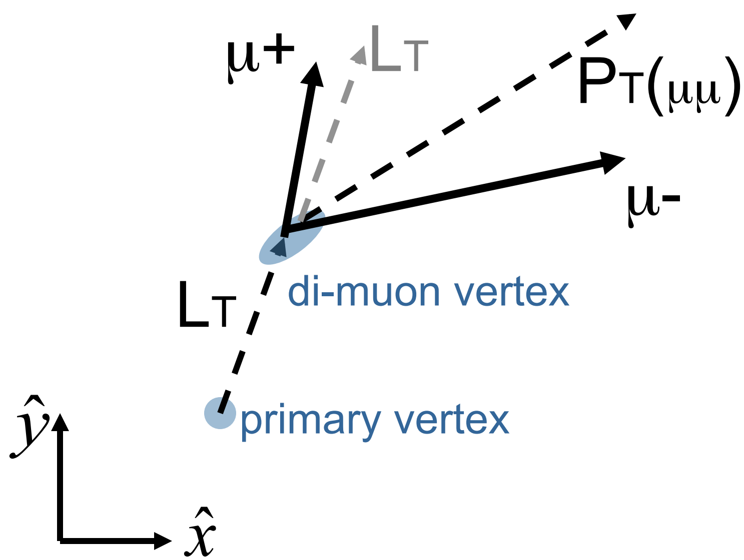
Signal characteristics

- final state is fully reconstructed
- B_s has long lifetime ($c\tau$ = 440 μm)
- B fragmentation is hard

For real $B_s \rightarrow \mu^+\mu^-$ expect:

- $m_{\mu\mu} = m(B_s)$
- $\lambda = cL_T\, m_{\mu\mu}/P_T(\mu\mu)$ to be large
- L_T and $P_T(\mu\mu)$ to be co-linear (ie. small $\Delta\Omega$)
- few additional tracks

Discriminate Signal from Background



Contributing Backgrounds

– sequencial semi-leptonic decay, $b \rightarrow \mu^- cX \rightarrow \mu^+\mu^- X$

– double semi-leptonic decay, $g \rightarrow bb \rightarrow \mu^+\mu^- X$

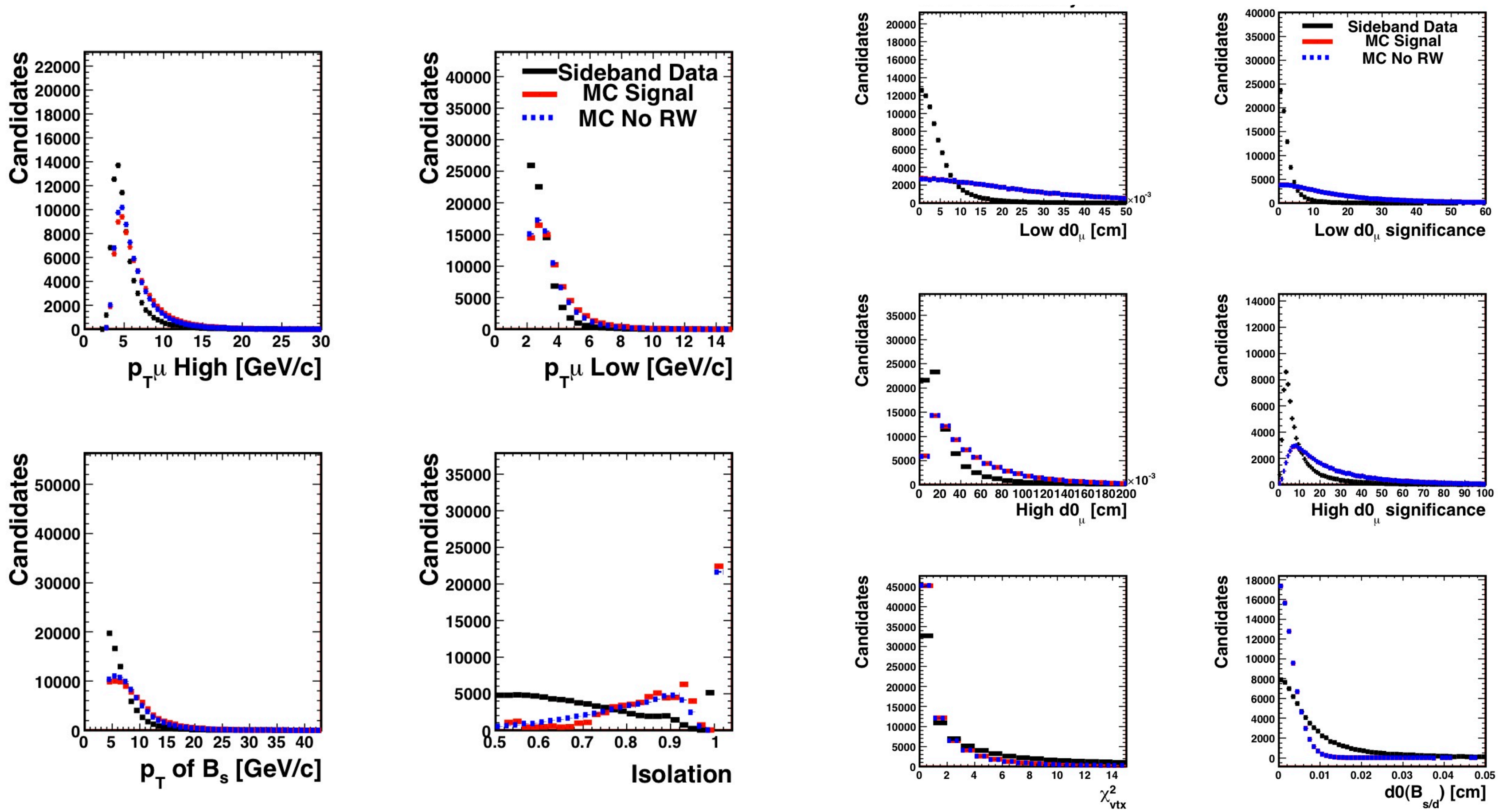
– continuum $\mu^+\mu^-$, μ + fake fake+fake

In general:

- $m_{\mu\mu} \neq m(B_s)$
- $\lambda = cL_T\, m_{\mu\mu}/P_T(\mu\mu)$ will be smaller

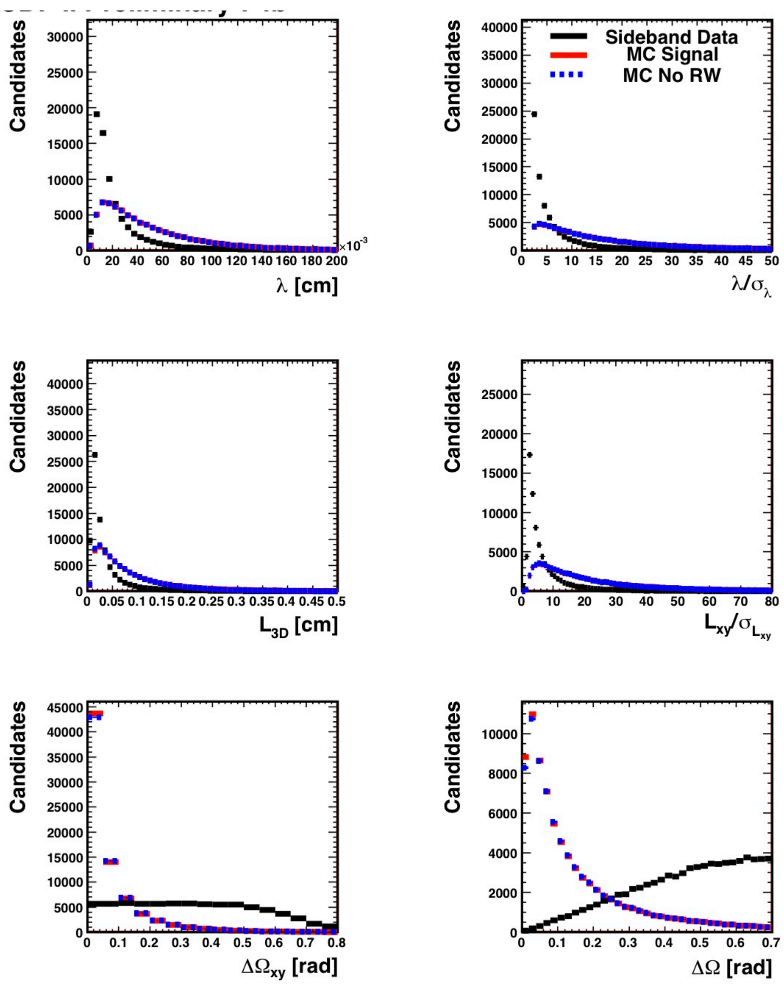
- L_T and $P_T(\mu\mu)$ will not be co-linear (large $\Delta\Omega$)
- more additional tracks

Discriminating Variables



- Some variables that take advantage of these distinguishing characteristics

Discriminating Variables

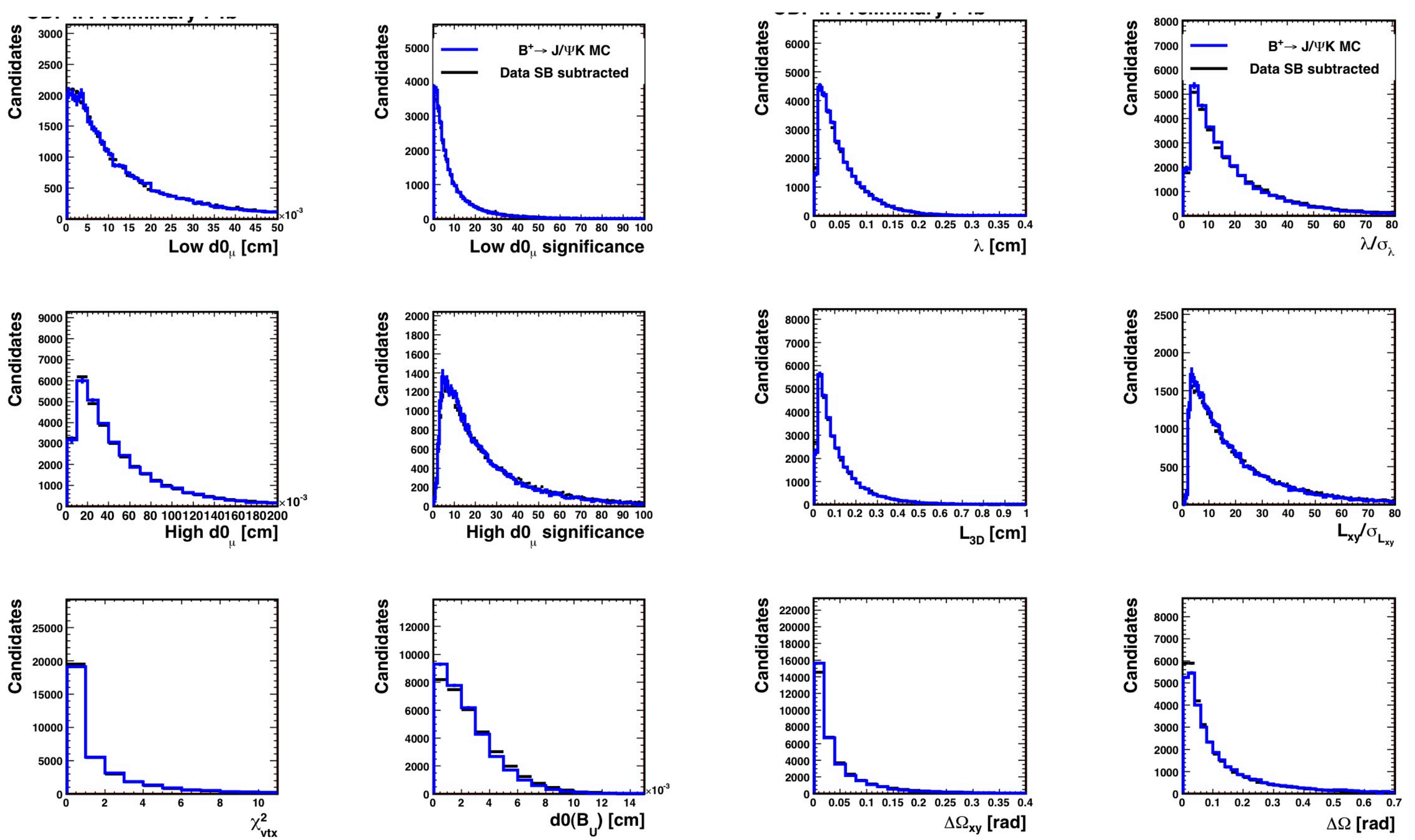


- more variables that take advantage of these distinguishing characteristics

Discriminate Signal from Background

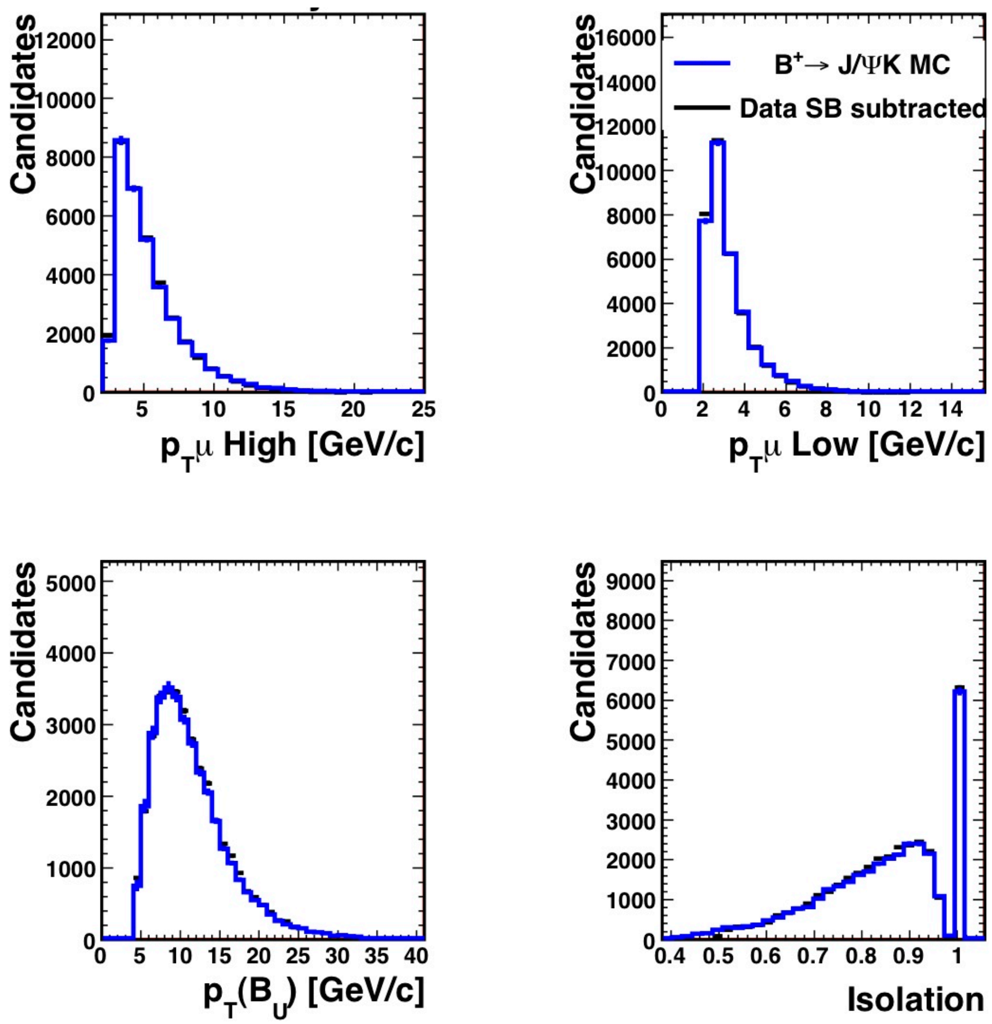
- Employ a Neural Net to optimally combine the information from these variables
  - We **exclude** mass information from the NN
  - M.Feindt and U.Kerzel, NIM A 559, 190 (2006)
- Training
  - Signal: $B_s \rightarrow \mu^+\mu^-$ MC
  - Background: mass sideband regions
  - Some fraction of each sample set aside to test for bias and overtraining

MC modeling



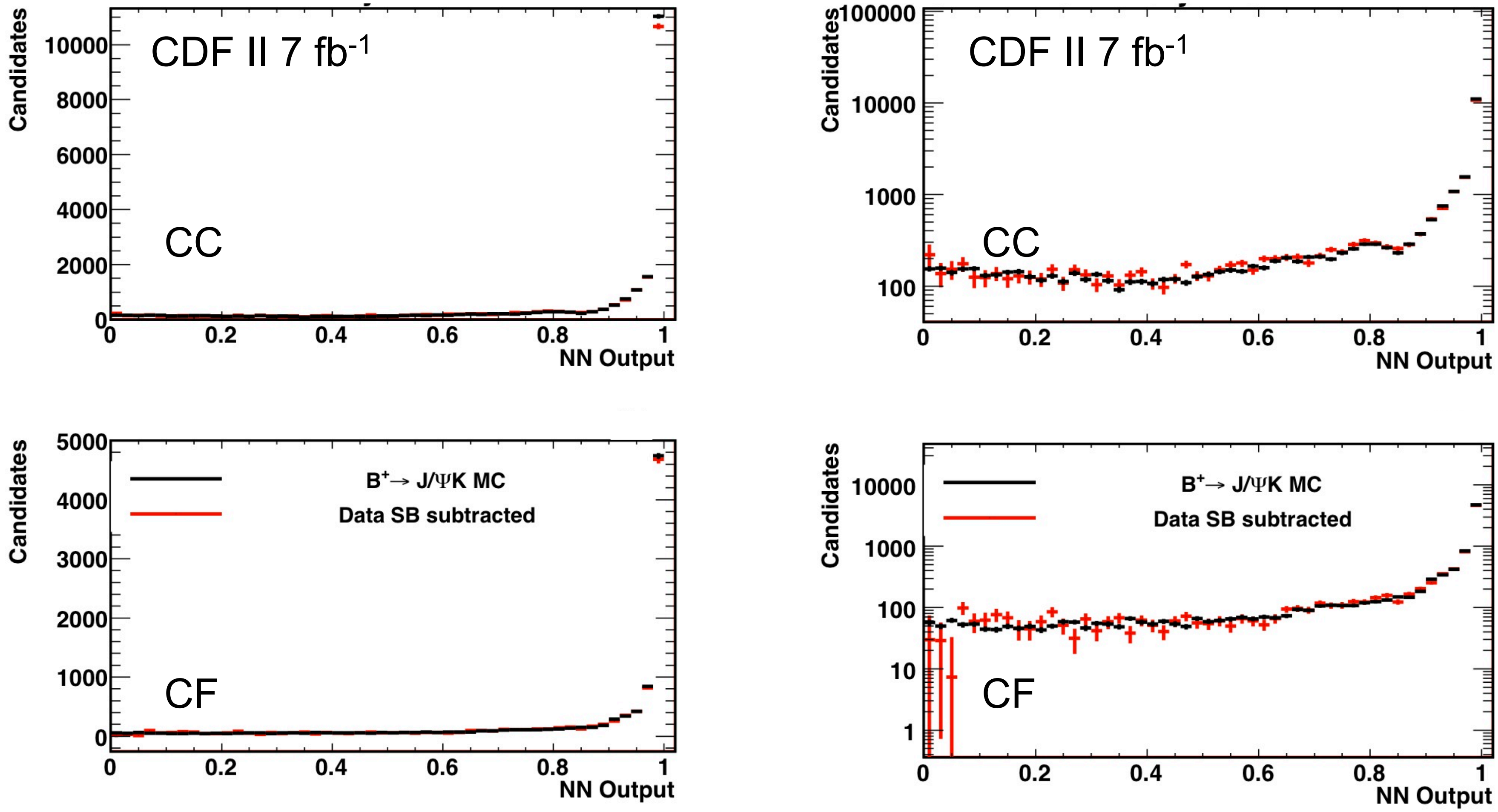
- Verify modeling of signal MC using B^+ events

MC modeling



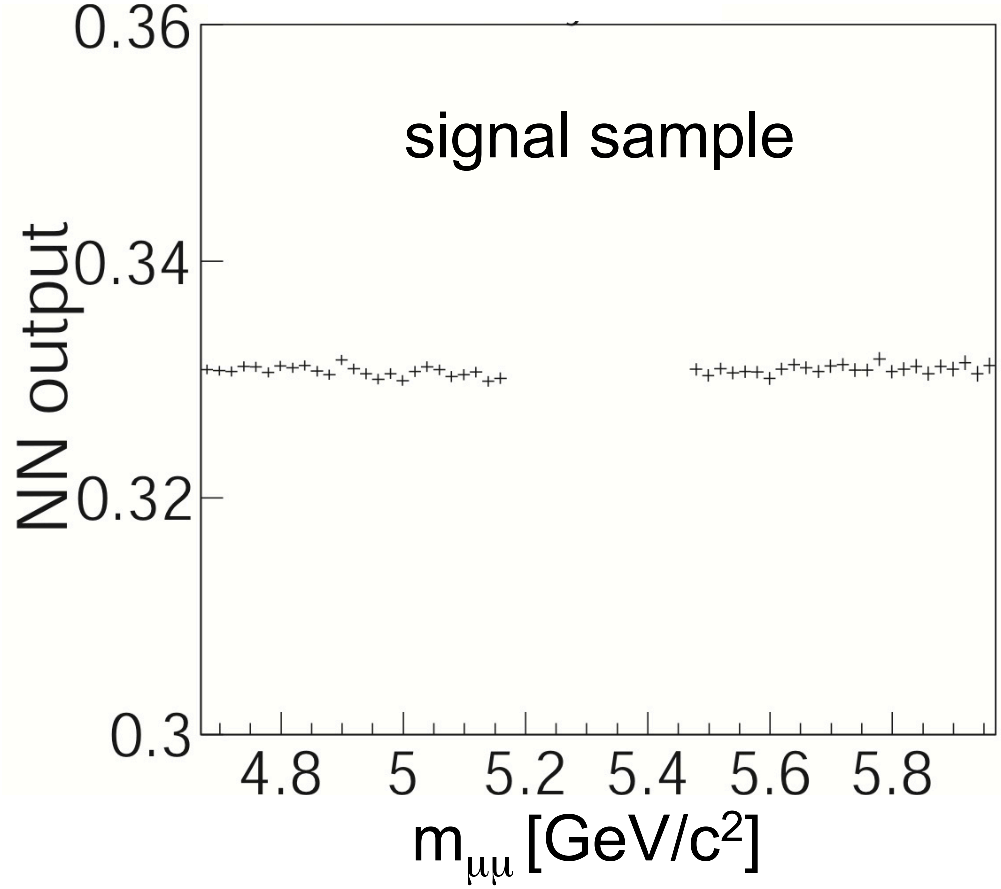
- Verify modeling of signal MC using B^+ events

MC modeling



- Verify modeling of signal MC using B^+ events

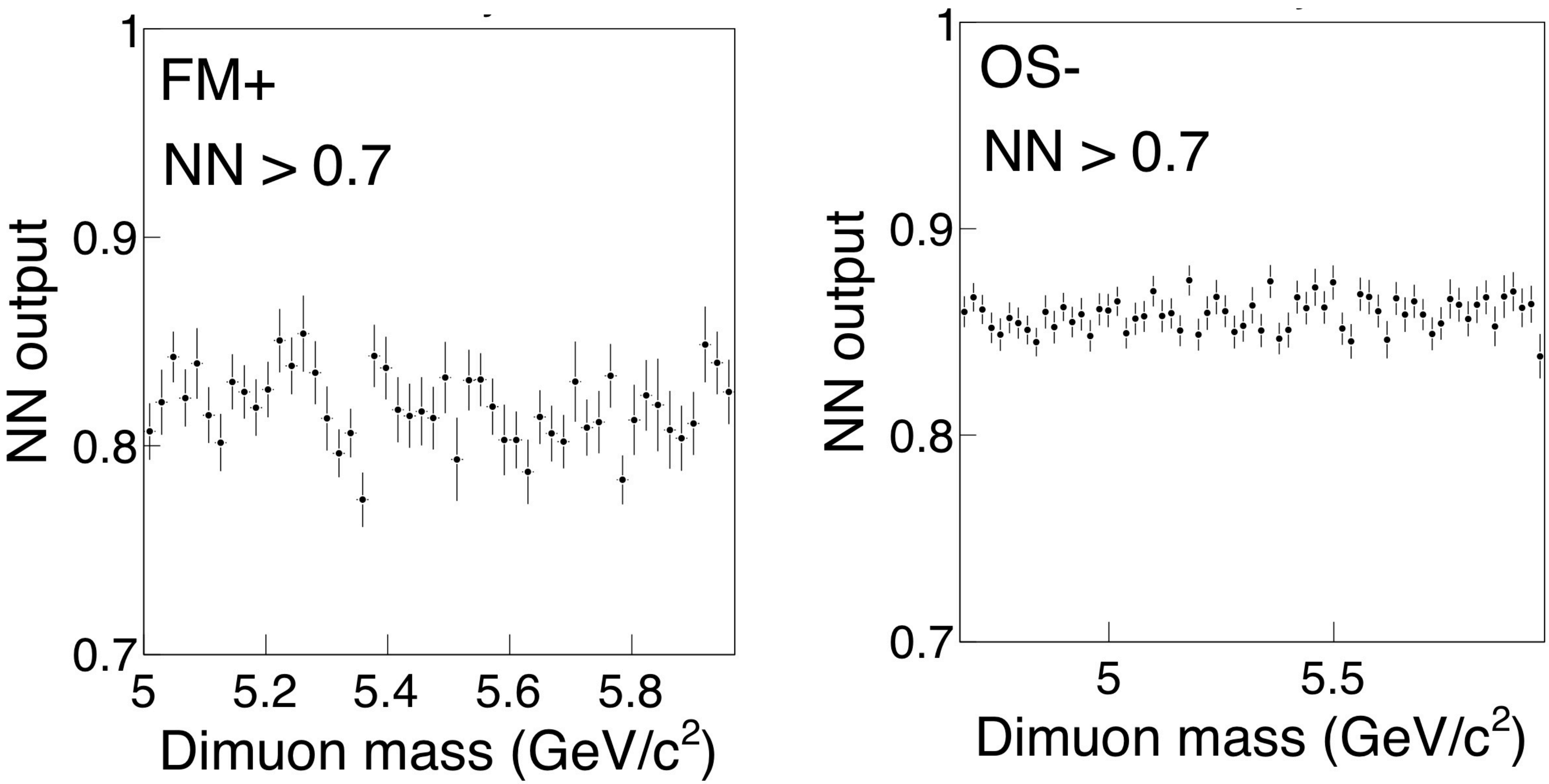
NN correlation with $m_{\mu\mu}$



- Important to verify ν_{NN} is independent of $m_{\mu\mu}$

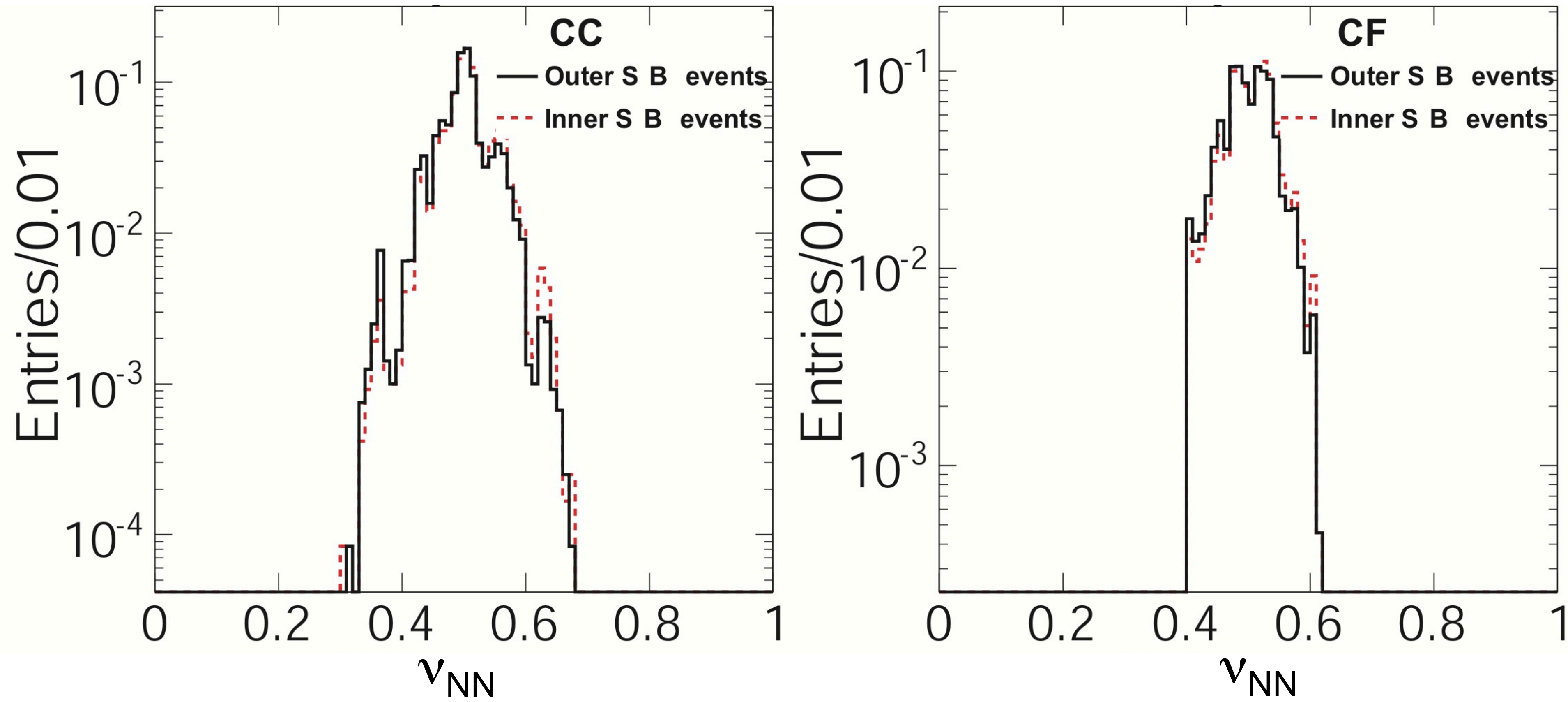
NN correlation with $m_{\mu\mu}$

background dominated control samples



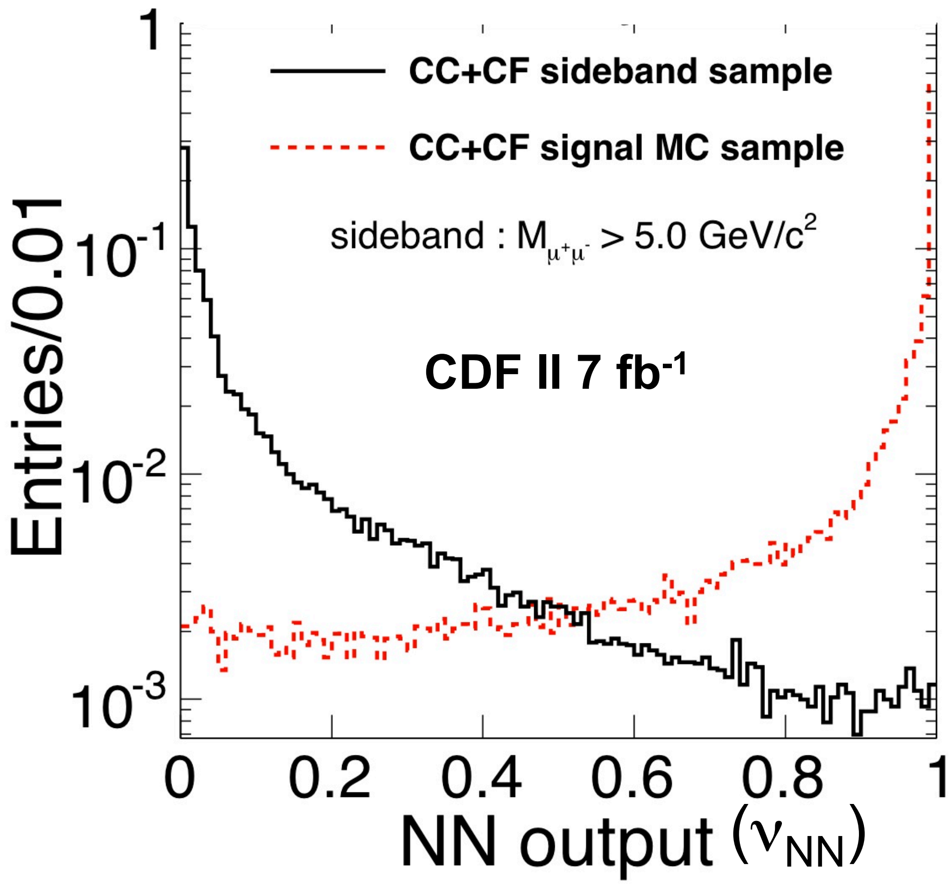
- Important to verify ν_{NN} is independent of $m_{\mu\mu}$

NN correlation with $m_{\mu\mu}$



- Important to verify ν_{NN} is independent of $m_{\mu\mu}$

NN Separation



- achieves powerful background discrimination

NN Variables

variable	description
$\Delta\Omega$	angle btwn L and p(B) (3D)
Isolation	B candidate isolation
$\|d_0(\mu_1)\|$	muon i.p. where $\|d_0(\mu_1)\| > \|d_0(\mu_2)\|$
$\|d_0(B)\|$	B candidate i.p.
L_T/σ_{LT}	decay length significance in xy plane
$\chi^{2(}vtx)$	vertex chi-squared vertex
L	decay length (3D)
$\min(p_T(\mu_1),p_T(\mu_2))$	minimum muon p_T
$\|d_0(\mu_2)\|/\sigma_{d0}$	muon i.p. significance
λ/σ_λ	proper time significance
λ	proper time
$\|d_0(\mu_2)\|$	muon i.p.
$\Delta\Omega_T$	angle btwn L_T and $p_T(B)$ (2D)
$\|d_0(\mu_1)\|/\sigma_{d0}$	muon i.p. significance

- A ranked list of the 14 variables used in the NN with the most significant variables at the top

Optimization of NN Requirements

- Figure-of-merit: expected limit

$$\left\langle BF(B_s \to \mu^+\mu^-)\right\rangle = \left(\frac{\left\langle N_{Bs}^{90\%CL}\right\rangle}{N_{B+}}\right)\left(\frac{f_u}{f_s}\right)\left(\frac{\alpha_{B+}\varepsilon_{B+}}{\alpha_{Bs}\varepsilon_{Bs}}\right) BF(B^+ \to J/\psi K^+)$$

$$\left\langle N_{Bs}^{90\%CL}\right\rangle = \sum_{n_{obs}=0}^{\infty} \mathrm{P}(n_{obs} \mid n_{bg}) \cdot N_{Bs}^{90\%CL}(n_{bg}, \Delta_{bg}, \Delta_{\alpha\cdot\varepsilon})$$

- Exploit S/B differences in ν_{NN} and $m_{\mu\mu}$
  - Bin in (ν_{NN}, $m_{\mu\mu}$) and optimize in 2D
- Broad minimum observed
  - Move away from regions with very few SB events
  - Choose something ~middle of minimum

Final NN Requirements

ν_{NN} bins
0.700 – 0.760
0.760 – 0.850
0.850 – 0.900
0.900 – 0.940
0.940 – 0.970
0.970 – 0.987
0.987 – 0.995
0.995 <

X

B_s mass bins	B_d mass bins
5310 – 5334	5219 – 5243
5334 – 5358	5243 – 5267
5358 – 5382	5267 – 5291
5382 – 5406	5291 – 5315
5406 – 5430	5315 – 5339

- Require $0.70 < \nu_{NN}$ ($\varepsilon_S \sim 90\%$, $\varepsilon_B \sim x\%$)
- Use 40 ($\nu_{NN}, m_{\mu\mu}$) bins
  - Each for CC and CF channels

Analysis Description

- This is a simple analysis

  1) ✓ Find events with 2 muons in them

  2) ✓ Identify means to suppress background while keeping as much signal as possible

  3) Look for a bump in the $m_{\mu\mu}$ distribution

Analysis Description

- This is a simple analysis
  1) ✓ Find events with 2 muons in them
  2) ✓ Identify means to suppress background while keeping as much signal as possible
  3) Look for a bump in the $m_{\mu\mu}$ distribution
     - Understand signal distributions
     - Understand background yields

Estimating Signal Yield

- Signal yield estimated for each (ν_{NN}, $m_{\mu\mu}$) bin using relative normalization

$$N_{Bs,d} = \left(\frac{N_{B+}}{BF(B^+ \to J/\psi K^+)}\right)\left(\frac{f_s}{f_u}\right)\left(\frac{\alpha_{Bs,d}\varepsilon_{Bs,d}}{\alpha_{B+}\varepsilon_{B+}}\right)BF(B_{s,d} \to \mu^+\mu^-)$$

$$\varepsilon_B \equiv \varepsilon_{\text{reco}} \cdot \varepsilon_{\text{NN}} \cdot \varepsilon_{\text{mass}} = \left(\varepsilon_{\text{track}} \cdot \varepsilon_{\mu\text{-ID}} \cdot \varepsilon_{\text{vertex}}\right) \cdot \varepsilon_{\text{NN}} \cdot \varepsilon_{\text{mass}}$$

Varies bin-by-bin

Estimates of SM $B_s \to \mu^+\mu^-$ Yields

CC channel:

NN Bin/Mass Bin	5.310-5.334	5.334-5.358	5.358-5.382	5.382-5.406	5.406-5.430
0.700-0.760	0.002 ± 0.000	0.007 ± 0.001	0.011 ± 0.002	0.006 ± 0.001	0.001 ± 0.000
0.760-0.850	0.004 ± 0.001	0.015 ± 0.003	0.020 ± 0.004	0.011 ± 0.002	0.003 ± 0.001
0.850-0.900	0.004 ± 0.001	0.010 ± 0.002	0.014 ± 0.003	0.008 ± 0.001	0.002 ± 0.000
0.900-0.940	0.005 ± 0.001	0.016 ± 0.003	0.023 ± 0.004	0.012 ± 0.002	0.002 ± 0.000
0.940-0.970	0.008 ± 0.001	0.022 ± 0.004	0.032 ± 0.006	0.016 ± 0.003	0.003 ± 0.001
0.970-0.987	0.010 ± 0.002	0.029 ± 0.005	0.041 ± 0.007	0.022 ± 0.004	0.005 ± 0.001
0.987-0.995	0.013 ± 0.002	0.046 ± 0.008	0.062 ± 0.011	0.031 ± 0.006	0.007 ± 0.001
0.995-1.000	0.052 ± 0.009	0.167 ± 0.030	0.227 ± 0.040	0.119 ± 0.021	0.029 ± 0.005

Σ=1.1evt

CF channel:

NN Bin/Mass Bin	5.310-5.334	5.334-5.358	5.358-5.382	5.382-5.406	5.406-5.430
0.700-0.760	0.002 ± 0.000	0.006 ± 0.001	0.007 ± 0.001	0.005 ± 0.001	0.001 ± 0.000
0.760-0.850	0.003 ± 0.001	0.012 ± 0.002	0.015 ± 0.003	0.009 ± 0.002	0.002 ± 0.000
0.850-0.900	0.003 ± 0.001	0.009 ± 0.002	0.012 ± 0.002	0.006 ± 0.001	0.001 ± 0.000
0.900-0.940	0.004 ± 0.001	0.012 ± 0.002	0.017 ± 0.003	0.009 ± 0.002	0.002 ± 0.000
0.940-0.970	0.005 ± 0.001	0.015 ± 0.003	0.021 ± 0.004	0.013 ± 0.002	0.003 ± 0.001
0.970-0.987	0.008 ± 0.002	0.026 ± 0.005	0.036 ± 0.007	0.019 ± 0.003	0.005 ± 0.001
0.987-0.995	0.007 ± 0.001	0.021 ± 0.004	0.029 ± 0.005	0.017 ± 0.003	0.004 ± 0.001
0.995-1.000	0.039 ± 0.007	0.116 ± 0.021	0.159 ± 0.029	0.090 ± 0.016	0.023 ± 0.004

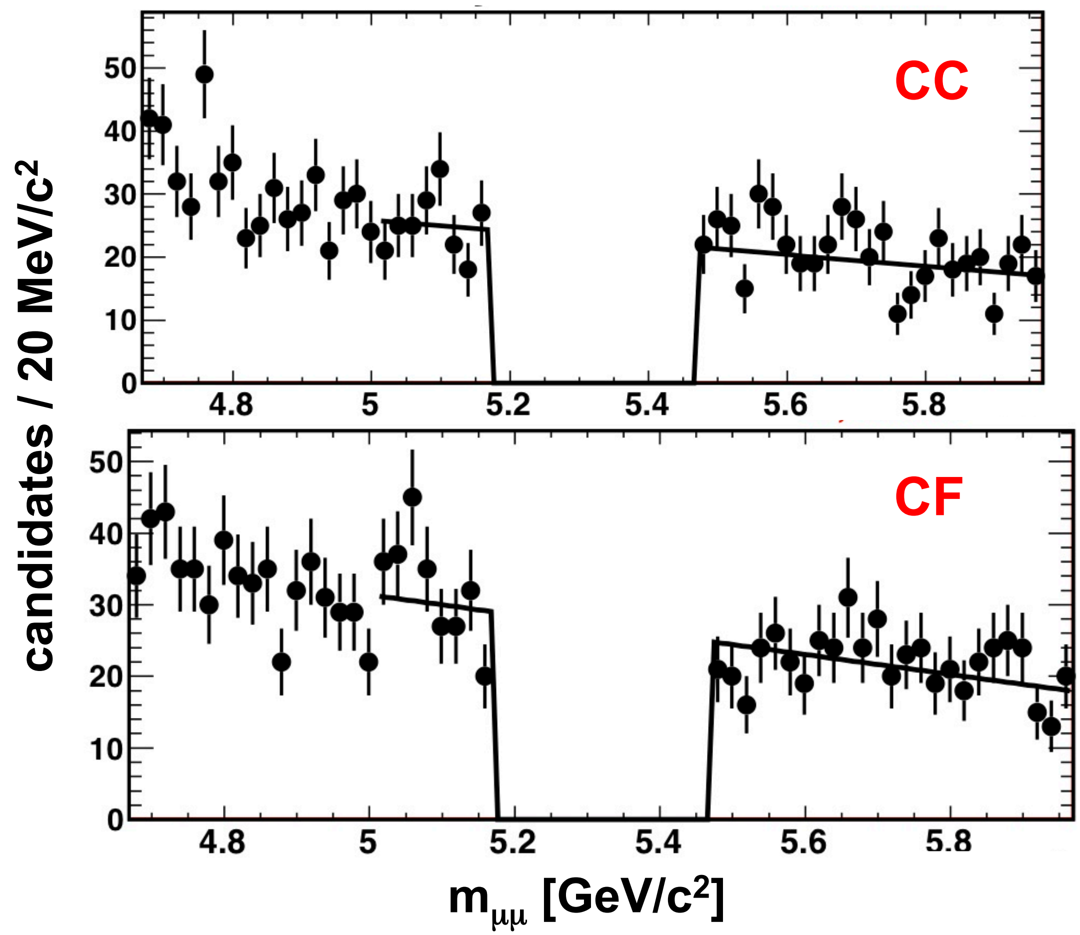
Σ=0.8evt

- Number of B_s signal events per bin, BF=SM

Estimating Background Yield

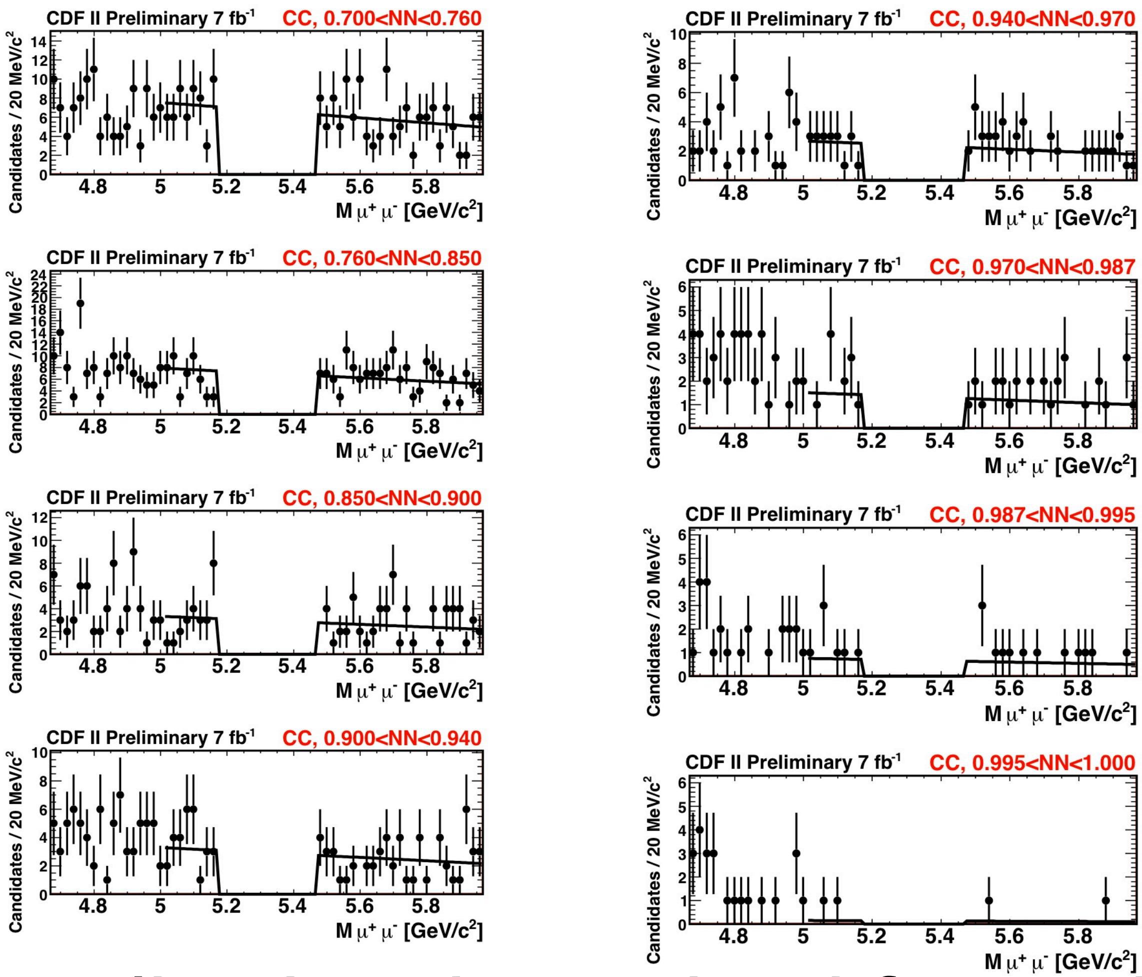
- Only 2 components to the background
  1) Combinatoric
     - Estimated using mass sidebands
  2) Peaking
     - Only source from $B \rightarrow h^+ h'^-$ ($h = \pi$, or K)
     - Kinematics taken from dedicated MC samples
     - Probability that π,K survive muon ID criteria is taken from D* tagged $D \rightarrow \pi K$ sample
- Verify accuracy of estimates using background control samples

Estimating Combinatoric Background



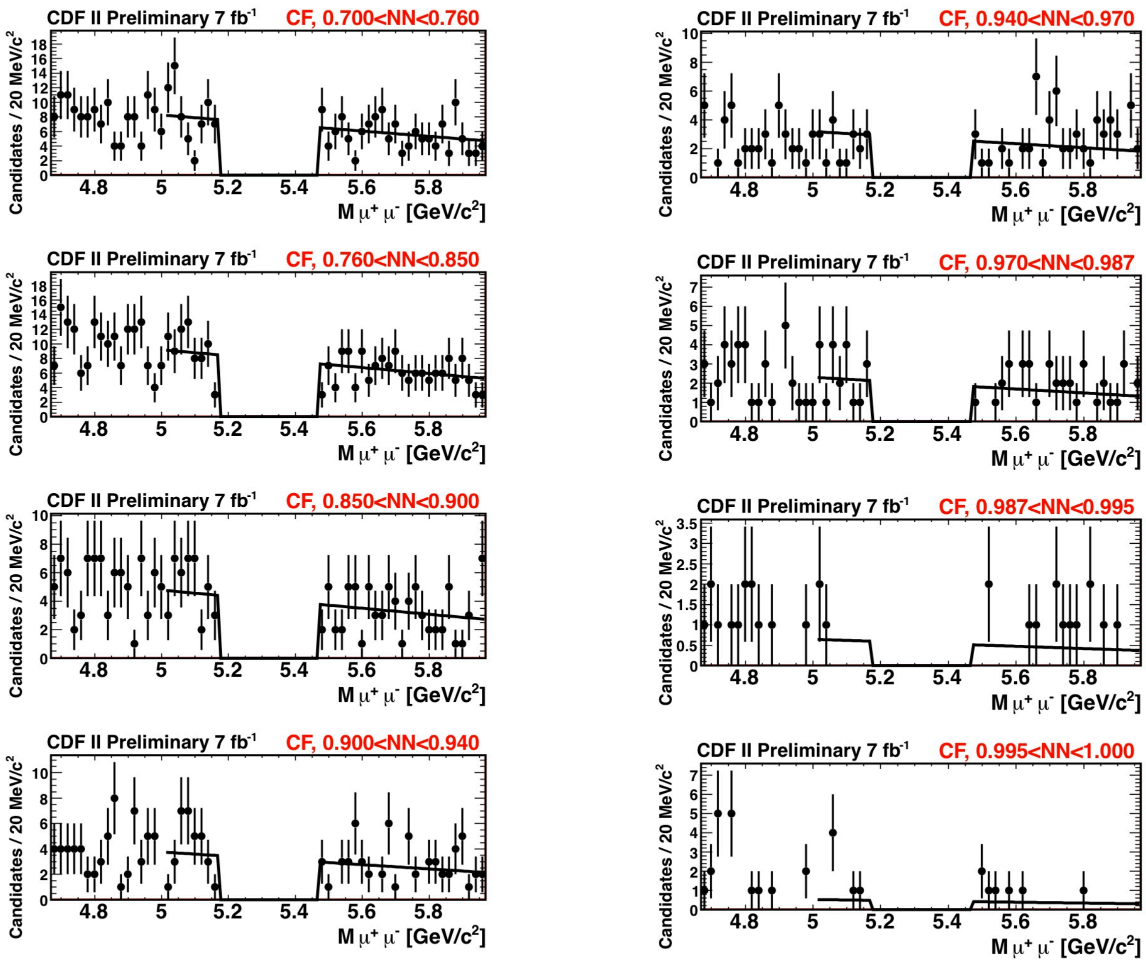
- Slope: from fit to $m_{\mu\mu}$>5 GeV/c^2 for 0.70 < ν_{NN}
  – CC and CF channels separately

Estimating Combinatoric Background



- Normalization determined for each ν_{NN} bin separately (for CC and CF separately)

Estimating Combinatoric Background



- Normalization determined for each ν_{NN} bin separately (for CC and CF separately)

Combinatoric Background Estimate: B_s

Combinatoric background: B_s region

NN Bin \ Mass Bin (GeV/c^2)	5.310−5.334	5.334−5.358	5.358−5.382	5.382−5.406	5.406−5.430
CC					
$0.700<NN<0.760$	8.02±0.62	7.94±0.61	7.87±0.61	7.79±0.60	7.71±0.59
$0.760<NN<0.850$	8.42±0.64	8.34±0.63	8.26±0.62	8.18±0.62	8.10±0.61
$0.850<NN<0.900$	3.55±0.39	3.51±0.39	3.48±0.39	3.44±0.38	3.41±0.38
$0.900<NN<0.940$	3.51±0.39	3.47±0.39	3.44±0.38	3.41±0.38	3.37±0.38
$0.940<NN<0.970$	2.86±0.35	2.83±0.35	2.81±0.34	2.78±0.34	2.75±0.34
$0.970<NN<0.987$	1.61±0.39	1.6±0.39	1.58±0.38	1.57±0.38	1.55±0.37
$0.987<NN<0.995$	0.81±0.23	0.80±0.23	0.79±0.22	0.78±0.22	0.78±0.22
$0.995<NN<1.000$	0.16±0.11	0.16±0.10	0.16±0.10	0.16±0.10	0.16±0.10
CF					
$0.700<NN<0.760$	8.49±0.65	8.39±0.64	8.28±0.63	8.17±0.62	8.07±0.61
$0.760<NN<0.850$	9.45±0.69	9.33±0.68	9.21±0.67	9.1±0.66	8.98±0.65
$0.850<NN<0.900$	4.91±0.48	4.85±0.47	4.79±0.46	4.73±0.46	4.67±0.45
$0.900<NN<0.940$	3.87±0.42	3.82±0.41	3.77±0.41	3.73±0.40	3.68±0.40
$0.940<NN<0.970$	3.29±0.38	3.25±0.38	3.21±0.37	3.17±0.37	3.12±0.36
$0.970<NN<0.987$	2.37±0.53	2.34±0.53	2.31±0.52	2.28±0.52	2.25±0.51
$0.987<NN<0.995$	0.67±0.20	0.66±0.20	0.65±0.20	0.64±0.19	0.63±0.19
$0.995<NN<1.000$	0.54±0.27	0.53±0.27	0.53±0.27	0.52±0.26	0.51±0.26

- uncertainty includes: slope & normalization uncertainties as well as variations in fit function and range

Combinatoric Background Estimate: B_d

Combinatoric background: B_d region

NN Bin \ Mass Bin (GeV/c^2)	5.219—5.243	5.243—5.267	5.267—5.291	5.291—5.315	5.315—5.339
CC					
0.700<NN<0.760	8.31±0.64	8.24±0.63	8.16±0.63	8.08±0.62	8.00±0.62
0.760<NN<0.850	8.73±0.66	8.65±0.65	8.57±0.65	8.49±0.64	8.41±0.63
0.850<NN<0.900	3.68±0.41	3.64±0.40	3.61±0.40	3.57±0.40	3.54±0.39
0.900<NN<0.940	3.63±0.40	3.60±0.40	3.57±0.40	3.53±0.39	3.50±0.39
0.940<NN<0.970	2.97±0.36	2.94±0.36	2.91±0.36	2.88±0.35	2.86±0.35
0.970<NN<0.987	1.67±0.40	1.66±0.40	1.64±0.40	1.62±0.39	1.61±0.39
0.987<NN<0.995	0.84±0.24	0.83±0.23	0.82±0.23	0.81±0.23	0.80±0.23
0.995<NN<1.000	0.17±0.11	0.17±0.11	0.16±0.11	0.16±0.11	0.16±0.11
CF					
0.700<NN<0.760	8.89±0.68	8.78±0.67	8.68±0.66	8.57±0.65	8.47±0.65
0.760<NN<0.850	9.89±0.72	9.78±0.71	9.66±0.70	9.54±0.69	9.42±0.69
0.850<NN<0.900	5.14±0.50	5.08±0.49	5.02±0.49	4.96±0.48	4.90±0.47
0.900<NN<0.940	4.05±0.44	4.00±0.43	3.96±0.43	3.91±0.42	3.86±0.42
0.940<NN<0.970	3.44±0.40	3.40±0.40	3.36±0.39	3.32±0.39	3.28±0.38
0.970<NN<0.987	2.48±0.56	2.45±0.55	2.43±0.55	2.40±0.54	2.37±0.53
0.987<NN<0.995	0.70±0.21	0.69±0.21	0.68±0.21	0.67±0.20	0.66±0.20
0.995<NN<1.000	0.57±0.29	0.56±0.28	0.55±0.28	0.55±0.28	0.54±0.27

- uncertainty includes: slope & normalization uncertainties as well as variations in fit function and range

Estimating Peaking Backgrounds

- Backgrounds which peak near the mass signal region will not be included in the combinatoric background estimates
- Only relevant sources of such events:
  - $B_d \rightarrow K^+\pi^-$, $\pi^+\pi^-$, K^+K^-
  - $B_s \rightarrow K^+K^-$, π^+K, $\pi^+\pi^-$
- These are suppressed because:
  - BF are small (10^{-5} to $<10^{-7}$)
  - $m_{\mu\mu}$ calculated assuming muon mass
  - Probability(π/K $\rightarrow$ fake μ) is small ($<1\times10^{-2}$)

Estimating Peaking Backgrounds

- To estimate yield, solve for $N_{B_{s,d}}$:

$$\frac{BF(B_{s,d} \to h^+ h'^-)}{BF(B^+ \to J/\psi K^+)} = \left(\frac{N_{Bhh}}{N_{B+}}\right)\left(\frac{f_u}{f_{s,d}}\right)\left(\frac{\alpha_{B+}\varepsilon_{B+}}{\alpha_{Bs,d}\varepsilon_{Bs,d}}\right)$$

- Obtain α*ε :

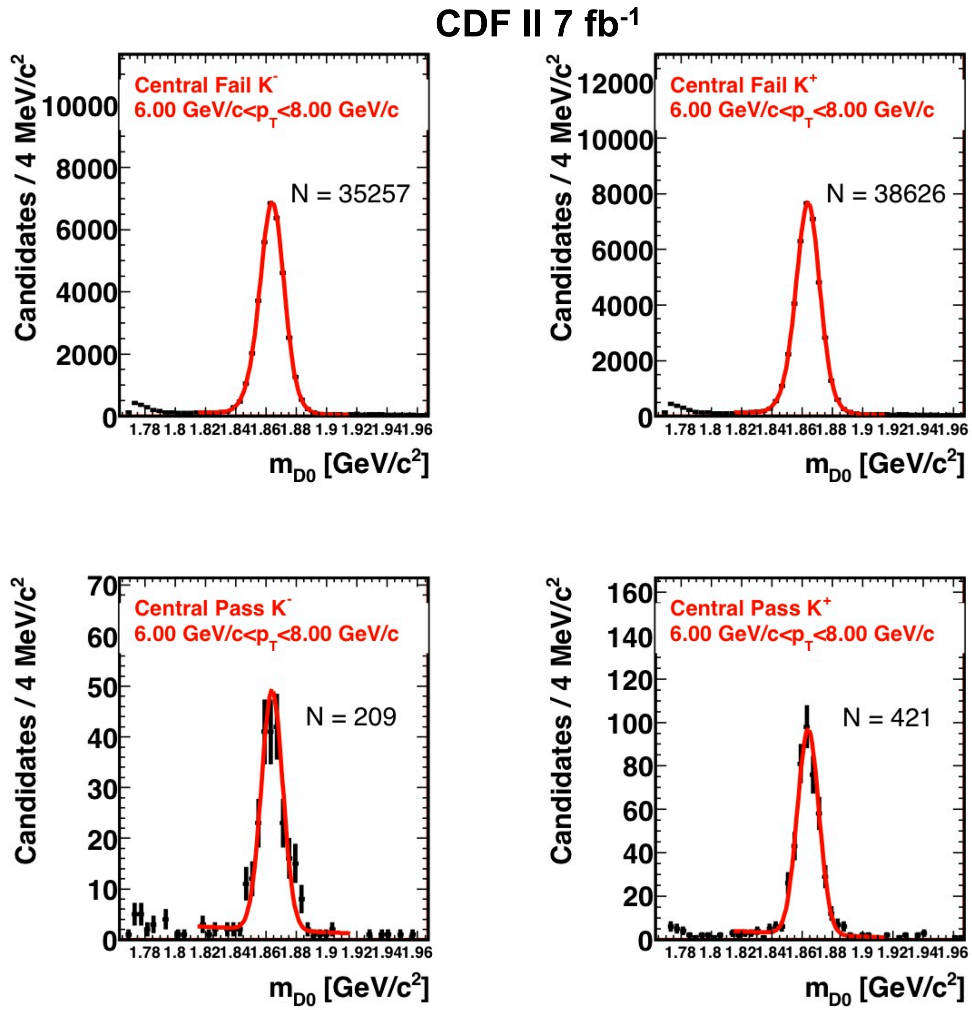
$$\alpha_B \equiv \text{ geometric and kinematic acceptance of trigger}$$

$$\varepsilon_B \equiv \varepsilon_{\text{reco}} \cdot \varepsilon_{\text{NN}} \cdot \varepsilon_{\text{mass}} = \left(\varepsilon_{\text{track}} \cdot \varepsilon_{\mu\text{-ID}} \cdot \varepsilon_{\text{vertex}}\right) \cdot \varepsilon_{\text{NN}} \cdot \varepsilon_{\text{mass}}$$

(blue) same as $B \to \mu^+\mu^-$

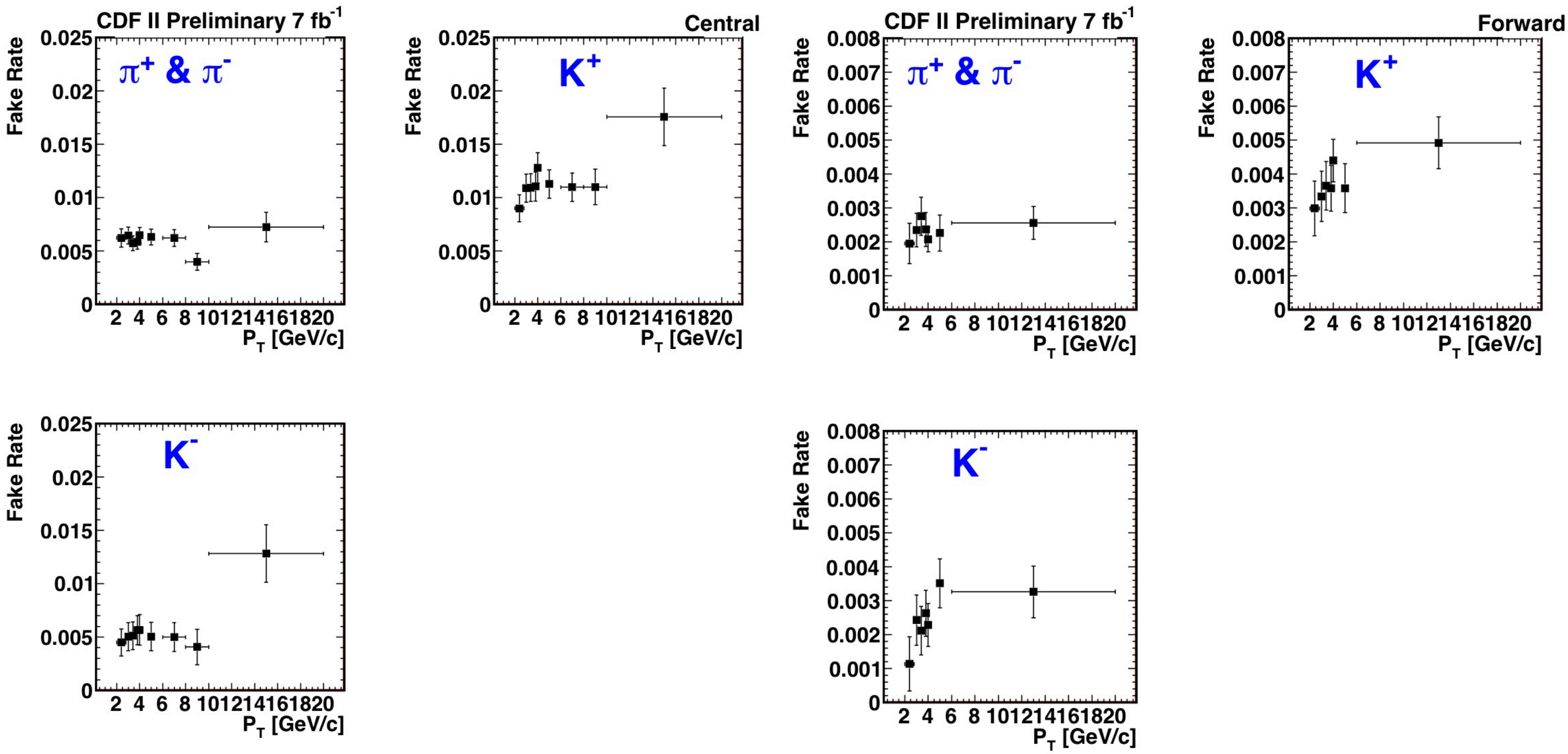
(yellow) requires special treatment

Estimating Peaking Backgrounds



- $\varepsilon_{\mu-fake}$ is taken from D* tagged $D^+ \rightarrow \pi^+ K^-$ data

Estimating Peaking Backgrounds



- $\varepsilon_{\mu-\text{fake}}$ parameterized (p_T, $\mathcal{L}_{inst}$)
  - Separately for $\pi^{+/-}$, K^+, and K^-

B→hh Background Estimate: B_s

B→hh background: B_s region

NN Bin \ Mass Bin (GeV/c^2)	5.310—5.334	5.334—5.358	5.358—5.382	5.382—5.406	5.406—5.430
CC					
0.700<NN<0.760	0.002±0.001	0.001± <0.001	–	–	–
0.760<NN<0.850	0.004±0.001	0.002± <0.001	0.001± <0.001	–	–
0.850<NN<0.900	0.004±0.001	0.001± <0.001	–	–	–
0.900<NN<0.940	0.005±0.001	0.002± <0.001	0.001± <0.001	–	–
0.940<NN<0.970	0.008±0.002	0.002±0.001	0.001± <0.001	–	–
0.970<NN<0.987	0.010±0.002	0.003±0.001	0.001± <0.001	–	–
0.987<NN<0.995	0.013±0.003	0.005±0.001	0.002±0.001	0.001± <0.001	–
0.995<NN<1.000	0.052±0.012	0.019±0.005	0.006±0.003	0.002±0.001	0.001± <0.001
CF					
0.700<NN<0.760	0.001± <0.001	–	–	–	–
0.760<NN<0.850	0.001± <0.001	0.001± <0.001	–	–	–
0.850<NN<0.900	0.001± <0.001	–	–	–	–
0.900<NN<0.940	0.002± <0.001	0.001± <0.001	–	–	–
0.940<NN<0.970	0.002±0.001	0.001± <0.001	–	–	–
0.970<NN<0.987	0.003±0.001	0.001± <0.001	0.001± <0.001	–	–
0.987<NN<0.995	0.003±0.001	0.001± <0.001	–	–	–
0.995<NN<1.000	0.015±0.004	0.006±0.002	0.002±0.001	0.001± <0.001	–

- Uncertainty includes: BF and fake-μ rate uncertainties (statistics of D* sub-samples, D^0 fits, residual luminosity dependence)

B→hh Background Estimate: B_d

B→hh background: B_d region

NN Bin \ Mass Bin (GeV/c^2)	5.219—5.243	5.243—5.267	5.267—5.291	5.291—5.315	5.315—5.339
CC					
0.700<NN<0.760	0.011±0.003	0.010±0.002	0.008±0.002	0.004±0.001	0.002± <0.001
0.760<NN<0.850	0.019±0.006	0.019±0.004	0.014±0.003	0.008±0.002	0.003±0.001
0.850<NN<0.900	0.016±0.005	0.013±0.003	0.010±0.002	0.006±0.001	0.002± <0.001
0.900<NN<0.940	0.022±0.006	0.021±0.005	0.016±0.004	0.009±0.002	0.003±0.001
0.940<NN<0.970	0.034±0.010	0.028±0.006	0.022±0.005	0.012±0.003	0.004±0.001
0.970<NN<0.987	0.042±0.013	0.037±0.009	0.029±0.007	0.016±0.004	0.006±0.001
0.987<NN<0.995	0.060±0.018	0.059±0.014	0.043±0.010	0.024±0.006	0.008±0.002
0.995<NN<1.000	0.231±0.068	0.211±0.049	0.157±0.036	0.090±0.022	0.035±0.008
CF					
0.700<NN<0.760	0.003±0.001	0.003±0.001	0.002±0.001	0.001± <0.001	–
0.760<NN<0.850	0.005±0.002	0.006±0.002	0.004±0.001	0.003±0.001	0.001± <0.001
0.850<NN<0.900	0.004±0.001	0.004±0.001	0.004±0.001	0.002±0.001	0.001± <0.001
0.900<NN<0.940	0.006±0.002	0.006±0.002	0.005±0.001	0.003±0.001	0.001± <0.001
0.940<NN<0.970	0.008±0.003	0.008±0.002	0.006±0.002	0.004±0.001	0.001± <0.001
0.970<NN<0.987	0.012±0.004	0.013±0.003	0.011±0.003	0.006±0.002	0.002±0.001
0.987<NN<0.995	0.010±0.003	0.011±0.003	0.009±0.002	0.005±0.001	0.002± <0.001
0.995<NN<1.000	0.057±0.018	0.061±0.015	0.048±0.012	0.028±0.007	0.011±0.003

- Uncertainty includes: BF and fake-μ rate uncertainties (statistics of D* sub-samples, D^0 fits, residual luminosity dependence)

Background Summary: B_s Search

Combinatoric:

NN Bin	CC	CF
$0.700 < NN < 0.970$	129.2±6.5	146.3±7.0
$0.970 < NN < 0.987$	7.9±1.9	11.6±1.8
$0.987 < NN < 0.995$	4.0±1.1	3.3±1.0
$0.995 < NN < 1.000$	0.79±0.52	2.6±1.5

B → hh:

NN Bin	CC	CF
$0.700 < NN < 0.970$	0.03±0.01	0.01± <0.01
$0.970 < NN < 0.987$	0.01± <0.01	0.01± <0.01
$0.987 < NN < 0.995$	0.02± <0.01	0.01± <0.01
$0.995 < NN < 1.000$	0.08±0.02	0.03±0.01

- Focus on 3 most sensitive ν_{NN} bins
  - integrating over $m_{\mu\mu}$ bins, first 5 ν_{NN} bins

Background Summary: B_d Search

Combinatoric:

NN Bin	CC	CF
$0.700 < NN < 0.970$	134.0±6.6	153.4±7.3
$0.970 < NN < 0.987$	8.2±2.0	12.1±1.9
$0.987 < NN < 0.995$	4.1±1.2	3.4±1.1
$0.995 < NN < 1.000$	0.8±0.5	2.8±1.6

B → hh:

NN Bin	CC	CF
$0.700 < NN < 0.970$	0.31±0.08	0.09±0.02
$0.970 < NN < 0.987$	0.13±0.03	0.05±0.01
$0.987 < NN < 0.995$	0.19±0.05	0.04±0.01
$0.995 < NN < 1.000$	0.72±0.20	0.20±0.05

- Focus on 3 most sensitive ν_{NN} bins
  - integrating over $m_{\mu\mu}$ bins, first 5 ν_{NN} bins

Cross-check Background Methodology

- We employ these data samples:
  - Opposite sign $\mu\mu$
    - L > 0 (OS+) this is our signal sample
    - L < 0 (OS-) bgd control sample
      - Dominated by combinatoric background
      - Kinematics very similar to signal sample
  - Same sign $\mu\mu$ (SS) bgd control sample
      - Dominated by combinatoric background
      - Different kinematics from signal sample
  - Fake-μ enhanced sample (FM) bgd control sample (require >=1 muon to fail μ-ID requirements)
      - Large B→hh contribution
      - Different kinematics from signal sample

Cross-check Background Methodology

sample	*NN* cut	CC pred	obsv	prob(%)
OS-	0.700<NN<0.760	217.4±(12.5)	203	77.7
	0.760<NN<0.850	262.0±(14.1)	213	99.1
	0.850<NN<0.900	117.9±(8.6)	120	44.7
	0.900<NN<0.940	112.1±(8.4)	116	39.4
	0.940<NN<0.970	112.7±(8.4)	108	64.2
	0.970<NN<0.987	80.2±(6.9)	75	68.3
	0.987<NN<0.995	67.6±(6.3)	41	99.8
	0.995<NN<1.000	32.5±(4.2)	35	37.5
SS+	0.700<NN<0.760	3.0±(0.9)	3	55.0
	0.760<NN<0.850	3.3±(1.0)	5	25.4
	0.850<NN<0.900	1.5±(0.7)	2	43.2
	0.900<NN<0.940	0.9±(0.5)	1	56.8
	0.940<NN<0.970	1.2±(0.6)	1	65.9
	0.970<NN<0.987	1.5±(0.7)	2	43.2
	0.987<NN<0.995	0.3±(0.3)	0	74.1
	0.995<NN<1.000	0.3±(0.3)	0	74.1
SS-	0.700<NN<0.760	5.7±(1.3)	8	23.7
	0.760<NN<0.850	8.4±(1.6)	7	69.8
	0.850<NN<0.900	3.3±(1.0)	6	14.3
	0.900<NN<0.940	2.4±(0.8)	4	24.0
	0.940<NN<0.970	2.4±(0.8)	4	24.0
	0.970<NN<0.987	2.1±(0.8)	0	12.2
	0.987<NN<0.995	1.5±(0.7)	0	22.3
	0.995<NN<1.000	0.3±(0.3)	1	30.0
FM+	0.700<NN<0.760	118.3±(8.6)	136	11.1
	0.760<NN<0.850	110.5±(8.3)	121	22.3
	0.850<NN<0.900	52.0±(5.4)	37	96.3
	0.900<NN<0.940	37.3±(4.5)	37	53.0
	0.940<NN<0.970	20.1±(3.3)	20	52.3
	0.970<NN<0.987	8.3±(2.0)	6	77.1
	0.987<NN<0.995	8.7±(2.0)	3	97.5
	0.995<NN<1.000	20.8±(3.5)	24	30.7

- Compare #observed to #predicted in all 80 (ν_{NN}, $m_{\mu\mu}$)bins across all background dominated control samples

Cross-check Background Methodology

sample	*NN* cut	CF pred	CF obsv	CF prob(%)
OS-	0.700<NN<0.760	209.3±(12.0)	187	88.8
	0.760<NN<0.850	332.3±(16.3)	325	62.0
	0.850<NN<0.900	146.7±(9.7)	144	57.7
	0.900<NN<0.940	144.2±(9.6)	139	63.9
	0.940<NN<0.970	128.6±(8.9)	112	88.4
	0.970<NN<0.987	92.8±(7.4)	89	63.0
	0.987<NN<0.995	45.4±(5.0)	55	14.0
	0.995<NN<1.000	38.3±(4.5)	37	58.2
SS+	0.700<NN<0.760	0.3±(0.3)	1	30.0
	0.760<NN<0.850	4.2±(1.1)	4	57.8
	0.850<NN<0.900	0.3±(0.3)	3	1.3
	0.900<NN<0.940	0.6±(0.4)	1	45.4
	0.940<NN<0.970	0.9±(0.5)	1	56.8
	0.970<NN<0.987	0.6±(0.4)	0	54.9
	0.987<NN<0.995	0.5±(0.4)	0	60.1
	0.995<NN<1.000	0.3±(0.3)	1	30.0
SS-	0.700<NN<0.760	4.2±(1.1)	4	57.8
	0.760<NN<0.850	5.1±(1.2)	7	27.1
	0.850<NN<0.900	2.7±(0.9)	2	71.0
	0.900<NN<0.940	0.9±(0.5)	4	2.8
	0.940<NN<0.970	3.0±(0.9)	1	92.3
	0.970<NN<0.987	2.4±(0.8)	5	12.2
	0.987<NN<0.995	0.6±(0.4)	0	54.9
	0.995<NN<1.000	1.8±(0.7)	0	16.5
FM+	0.700<NN<0.760	54.8±(5.6)	66	12.7
	0.760<NN<0.850	66.3±(6.2)	57	83.1
	0.850<NN<0.900	33.7±(4.3)	25	90.3
	0.900<NN<0.940	17.4±(3.1)	26	6.6
	0.940<NN<0.970	9.5±(2.2)	15	10.2
	0.970<NN<0.987	5.3±(1.7)	9	13.4
	0.987<NN<0.995	2.7±(1.2)	3	49.3
	0.995<NN<1.000	2.1±(1.0)	8	0.7

- Compare #observed to #predicted in all 80 (ν_{NN}, $m_{\mu\mu}$)bins across all background dominated control samples

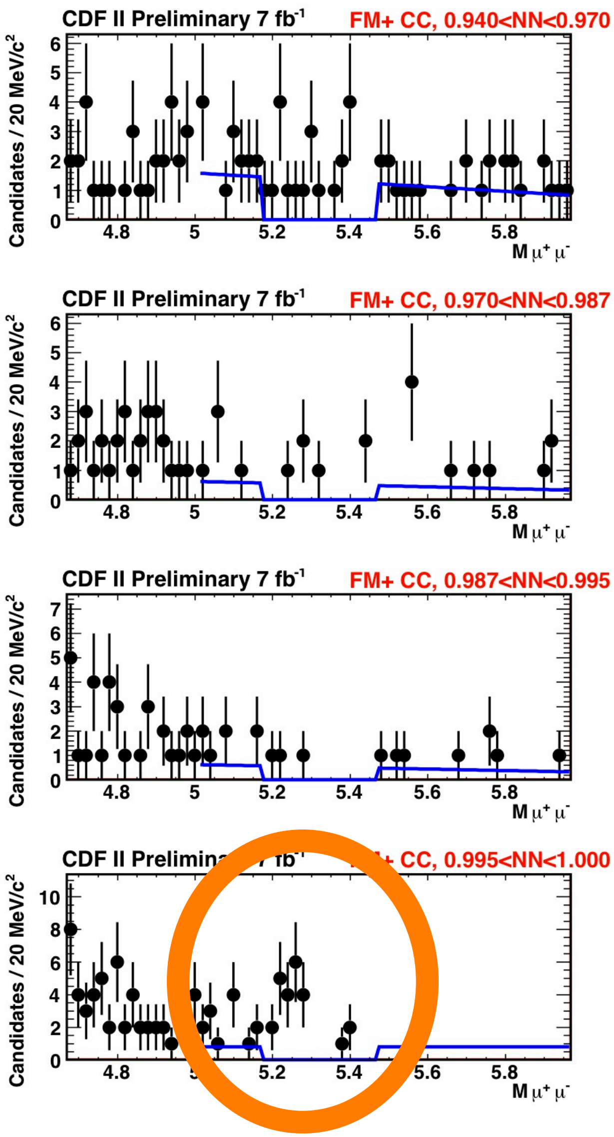
Cross-check Background Methodology

Control Sample	Prediction	Nobs	Prob(N>=Nobs)
OS-	2140.0±53.9	1999	98%
SS+	19.7±3.4	25	19%
SS-	46.8±5.3	53	25%
FM+	567.8±25.4	593	24%
Sum	2774.3±59.9	2670	91%

Table: A comparison of the predicted and observed number of events in an extended signal mass region for all NN cuts for all the control samples. This is used as a cross check of the background estimates.

- Integrating over all bins in each sample

Cross-check Background Methodology



- Observe B→hh in predicted place at predicted rate

What?

Sensitivity

- CDF expected sensitivity BF($B \rightarrow \mu\mu$):

B_d : 4.6 E-9 @ 95% CL
3.6 E-9 @ 90% CL

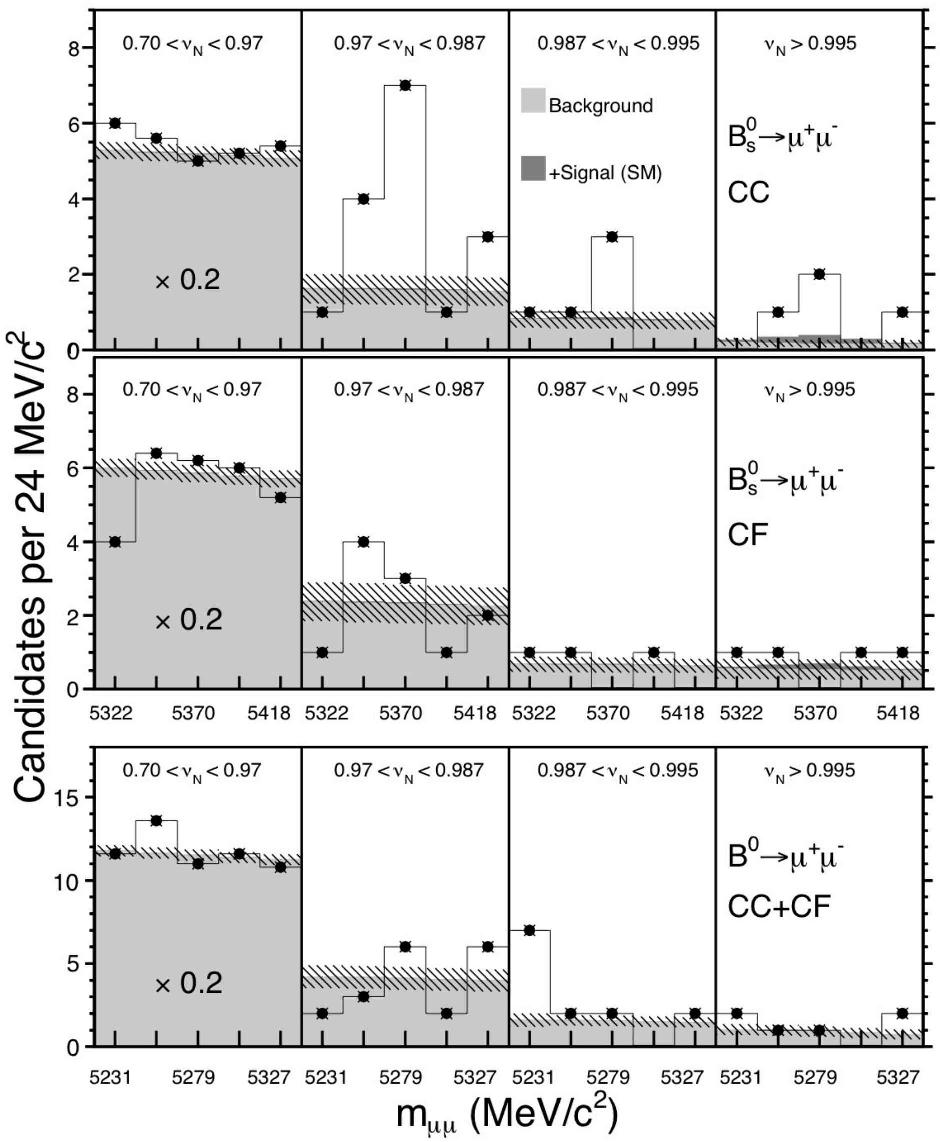
B_s : 1.5 E-8 @ 95% CL
1.1 E-8 @ 90% CL

- Among world's best in both channels
  - CMS: B_d = 4.8E-9 B_s = 1.8E-8
  - LHCb: B_d = 3.1E-9 B_s = 1.0E-8

  @ 95% CL

(all of these derived assuming background-only)

Result



- Comparison of data to background prediction in the (ν_{NN}, $m_{\mu\mu}$) bins from the optimization
- Only showing systematic uncertainties

Likelihood ratio

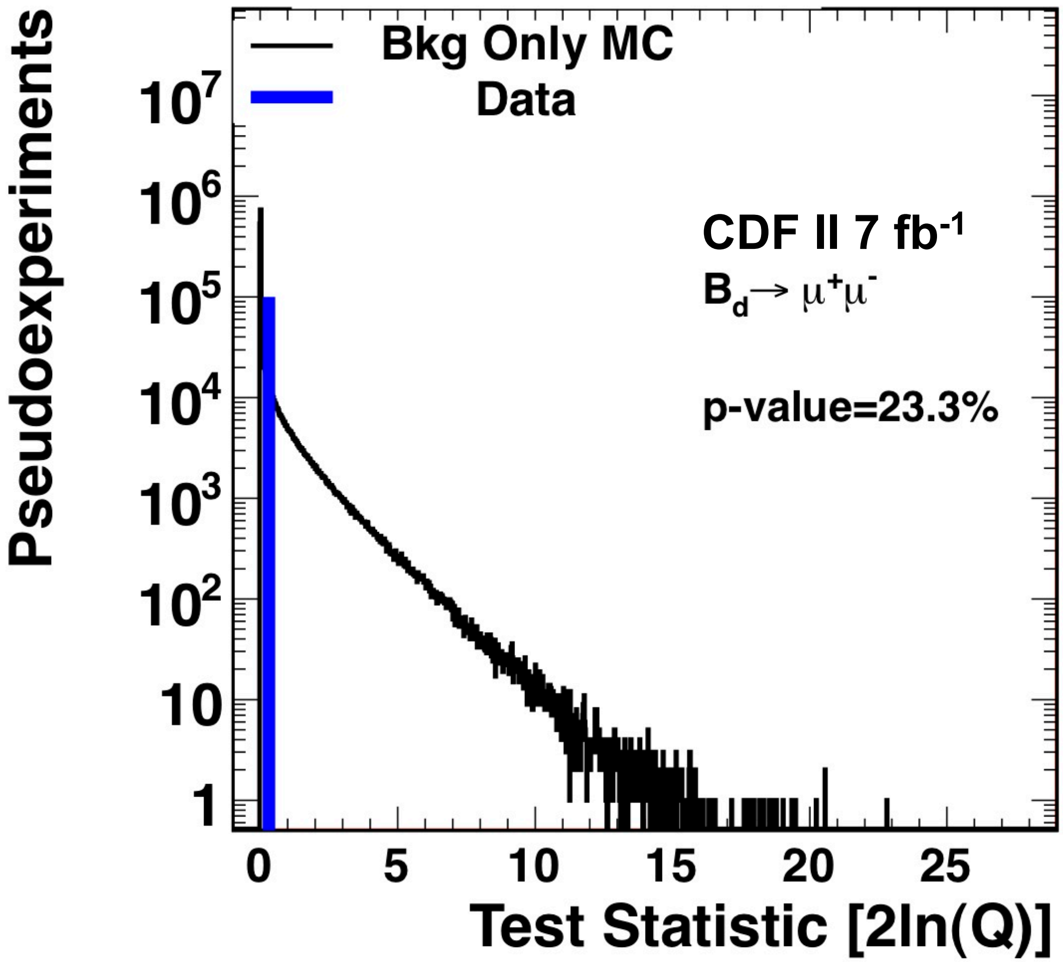
- We fit the data twice
  1) Assuming signal = 0
  2) Leave signal BF floating
  – Then take ratio: Q = L(s+b) / L(b)

- The likelihood:

$$\mathrm{L} = \left[\prod_{i=1}^{\mathrm{Nbins}} \mathrm{P}(n_{obs}^{i} \mid s_i + b_i)\right] \prod_{\mathrm{j=1}}^{\mathrm{Nsyst}} \mathrm{G}\left(x_j \mid \sigma_j\right)$$

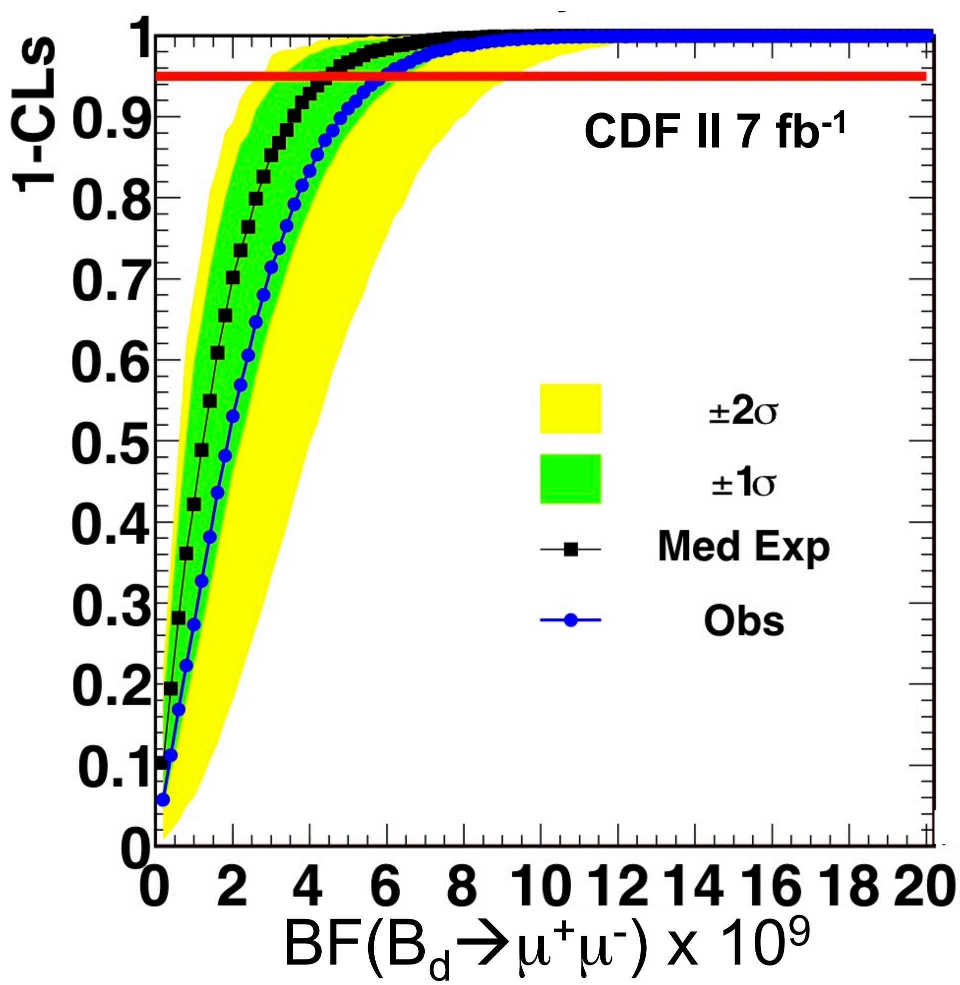
$$s_i = F(BF(B_s \to \mu^+\mu^-), x_j), \quad b_i = F(x_j)$$

Result: B_d



- p-value using background-only pseudo-exp.

Result: B_d



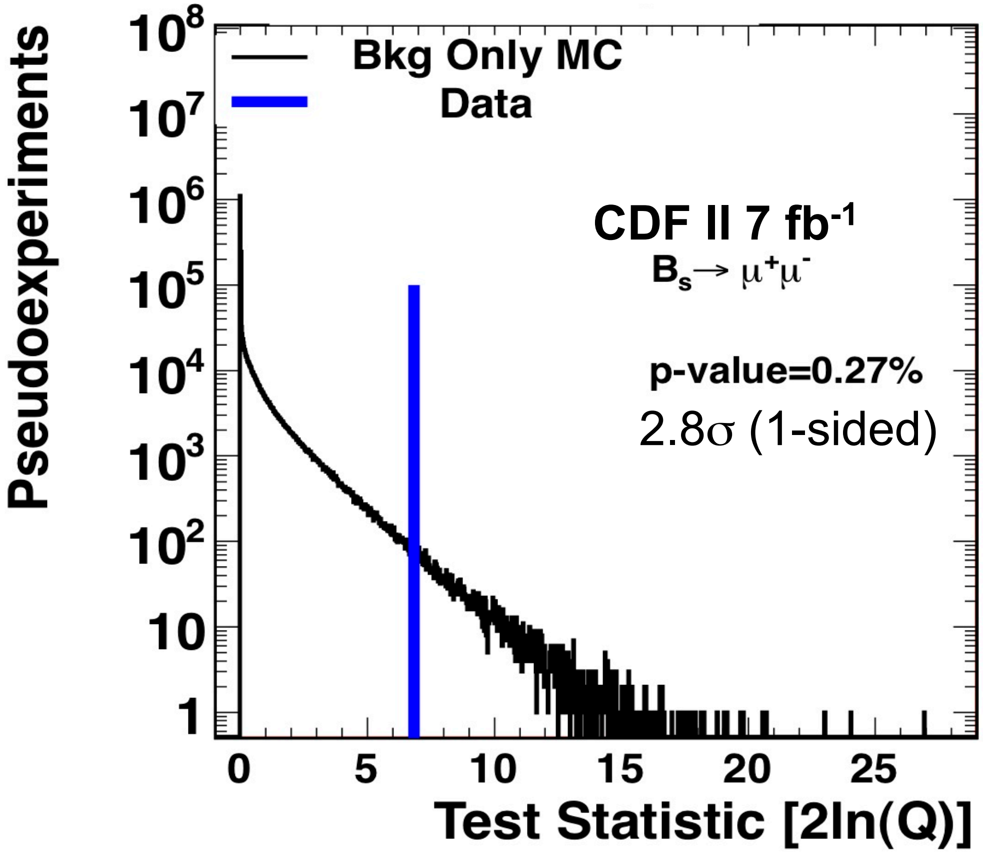
- Expected:

$$BF(B_d \rightarrow \mu^+\mu^-) \begin{array}{l} < 4.6 \times 10^{-9} \text{ (95\% CL)} \\ < 3.6 \times 10^{-9} \text{ (90\% CL)} \end{array}$$

- Observed:

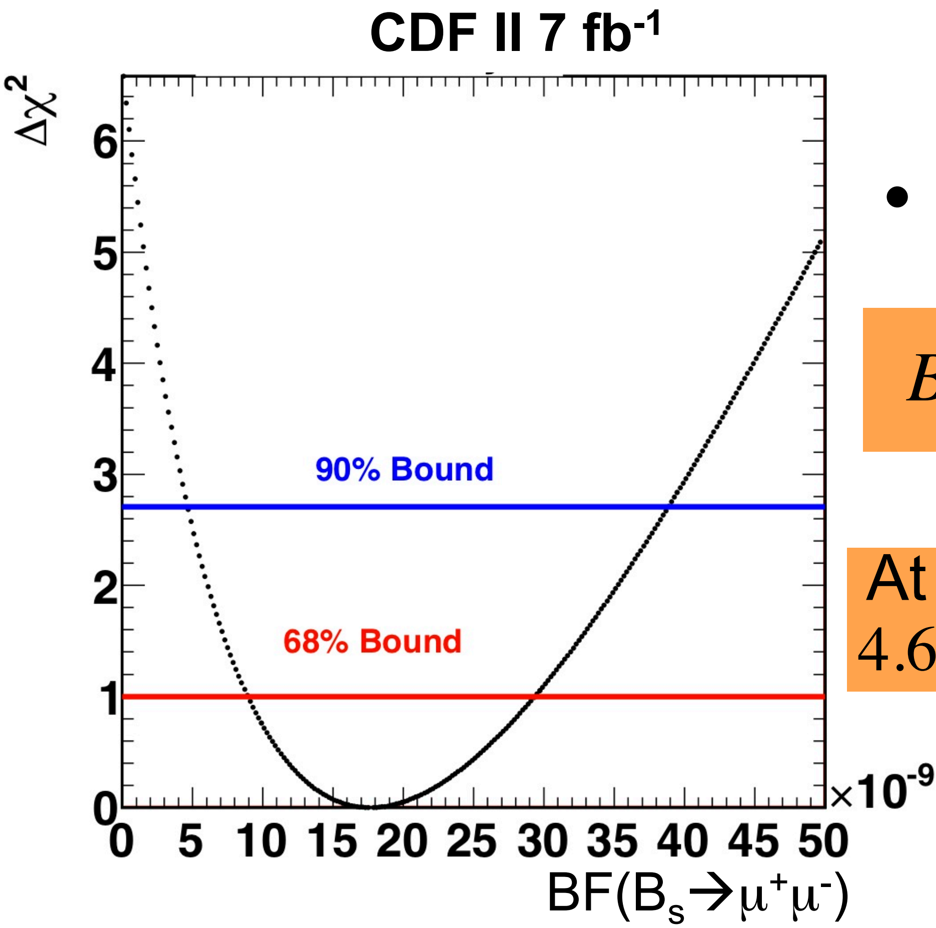
$$BF(B_d \rightarrow \mu^+\mu^-) \begin{array}{l} < 6.0 \times 10^{-9} \text{ (95\% CL)} \\ < 5.0 \times 10^{-9} \text{ (90\% CL)} \end{array}$$

Result: B_s



- p-value using background-only pseudo-exp.
  - If we include SM signal, p-value → 1.9% (2.1σ)

Result: B_s



- Fit to determine BF

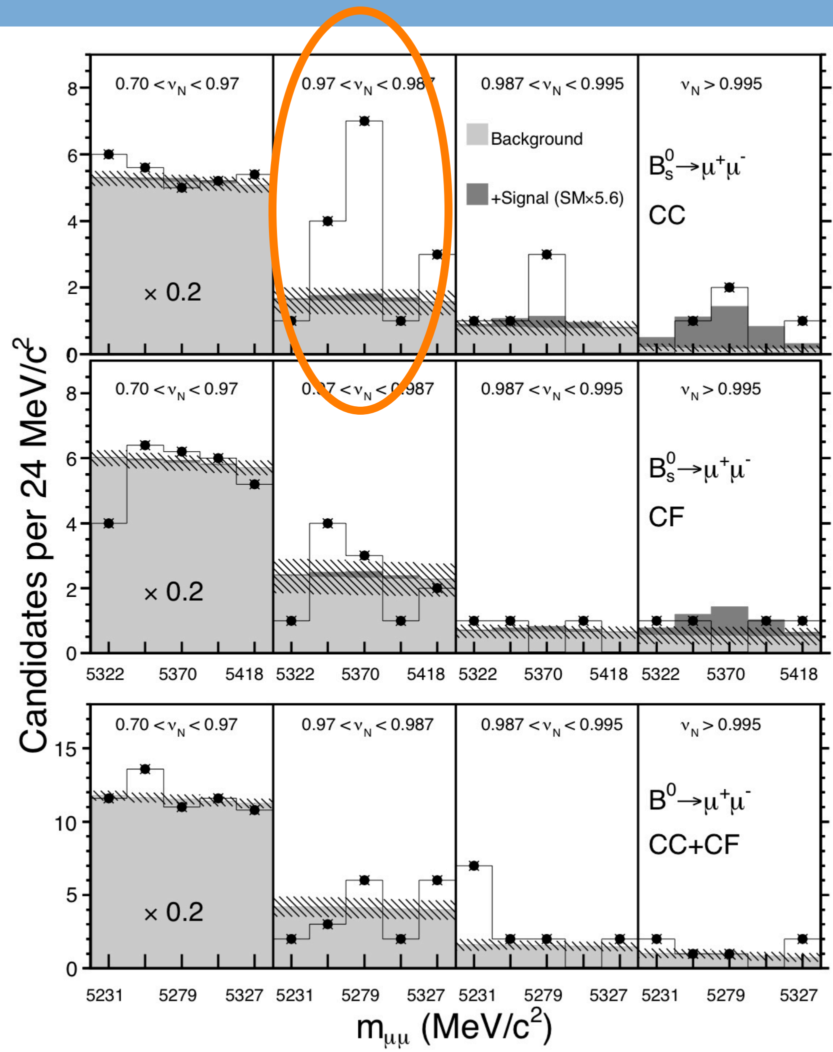
$$BF(B_s \to \mu^+\mu^-) = \left(1.8^{+1.1}_{-0.9}\right) \times 10^{-8}$$

At 90% CL:

$$4.6 \times 10^{-9} < BF(B_s \to \mu^+\mu^-) < 3.9 \times 10^{-8}$$

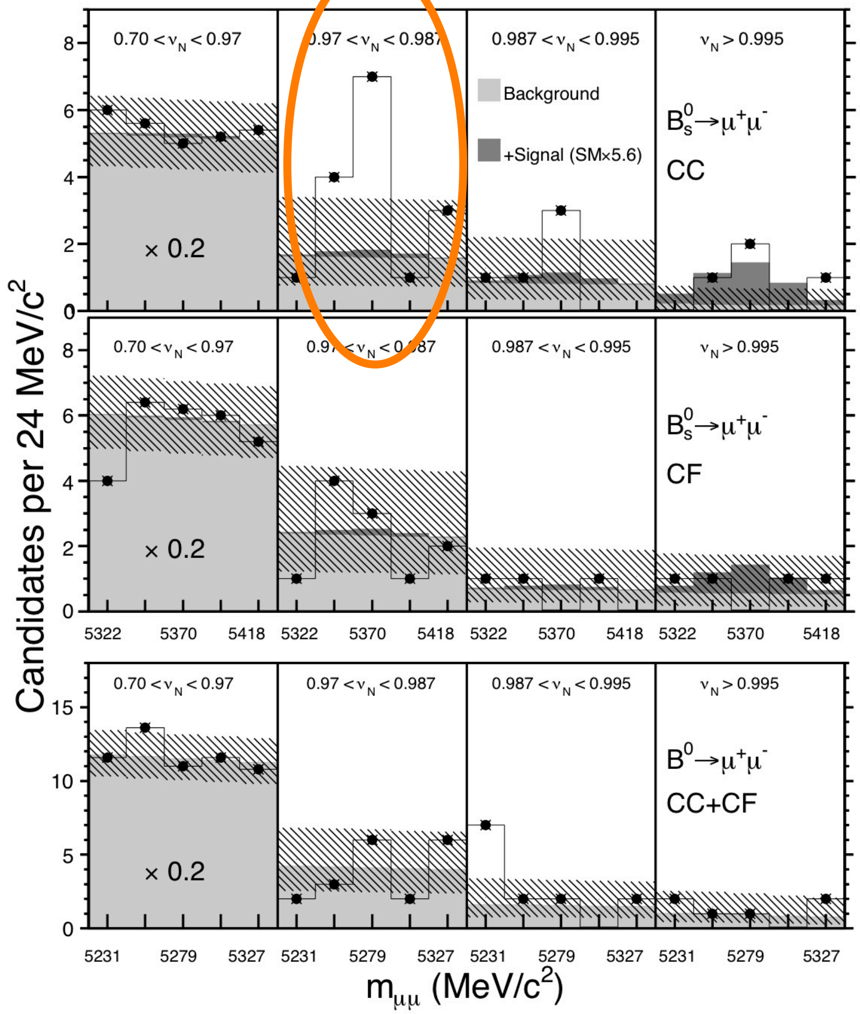
But…

Not so fast…



- What is this?

Not so fast…



Uncertainty:

syst ⊕ Poisn.

- What is this?

Possibilities

- Only two possible problems to consider (this is a simple analysis):

  1) Problem with background estimate
     - e.g. your π/K fake rates are wrong

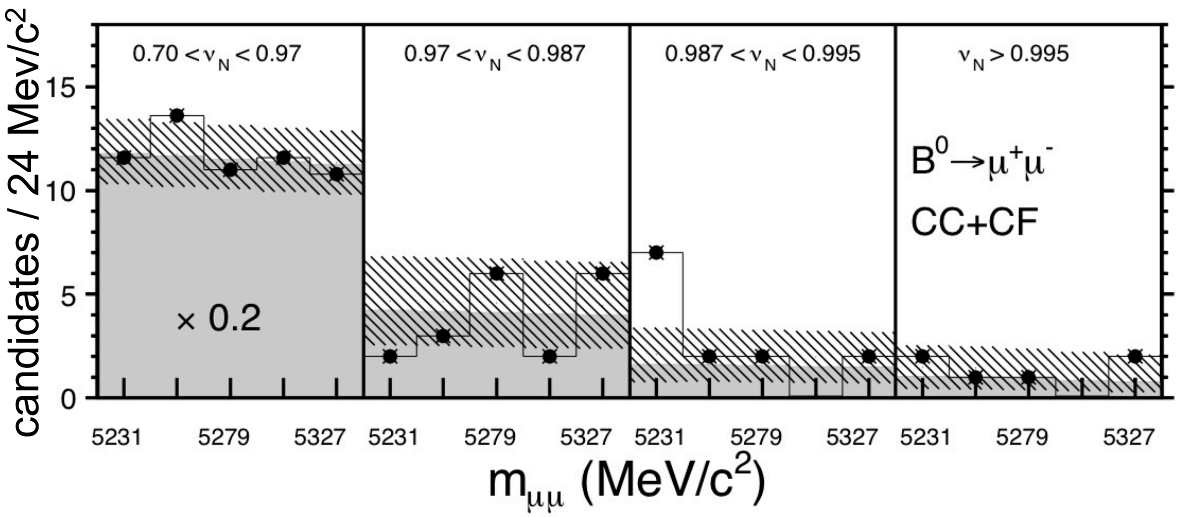
  2) Problem with NN
     - e.g. NN is over trained or mis-modeled

Recall

- Background estimate in B_d region
  - Uses exact same sideband events
  - Uses exact same sideband fits for slope and normalization
  - Uses exact same π/K → "μ" fake rates

Recall

- Background estimate in B_d region
  - Uses exact same sideband events
  - Uses exact same sideband fits for slope and normalization
  - Uses exact same π/K → "μ" fake rates



  - Accurately predicts data in signal region

Recall

- The yield of B→hh events in B_d region is about a factor of 10 larger than in the B_s region
  - If there were a problem with the π/K fake rates, it would show-up much more significantly in B_d
- In order to account for the observed excess, fake rates would have to be off by x10
  - They have a systematic uncertainty of 20%
  - Would generate much larger excesses in other bins

Possibilities

- Only two possible problems to consider (this is a simple analysis):

  ✓1) Problem with background estimate
  - e.g. your π/K fake rates are wrong

  2) Problem with NN
  - e.g. NN is over trained or mis-modeled

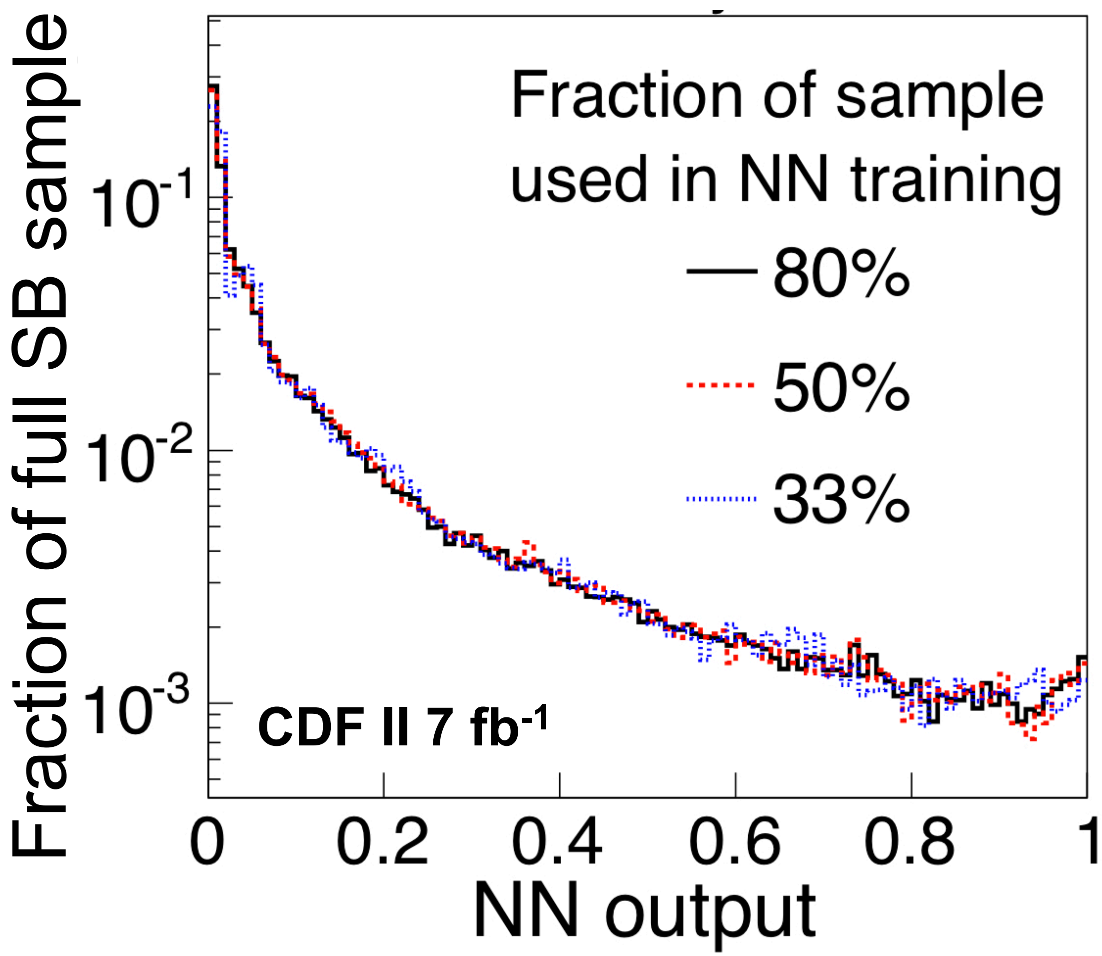
Possibilities

- Only two possible problems to consider (this is a simple analysis):

✓ 1) Problem with background estimate
   - e.g. your π/K fake rates are wrong

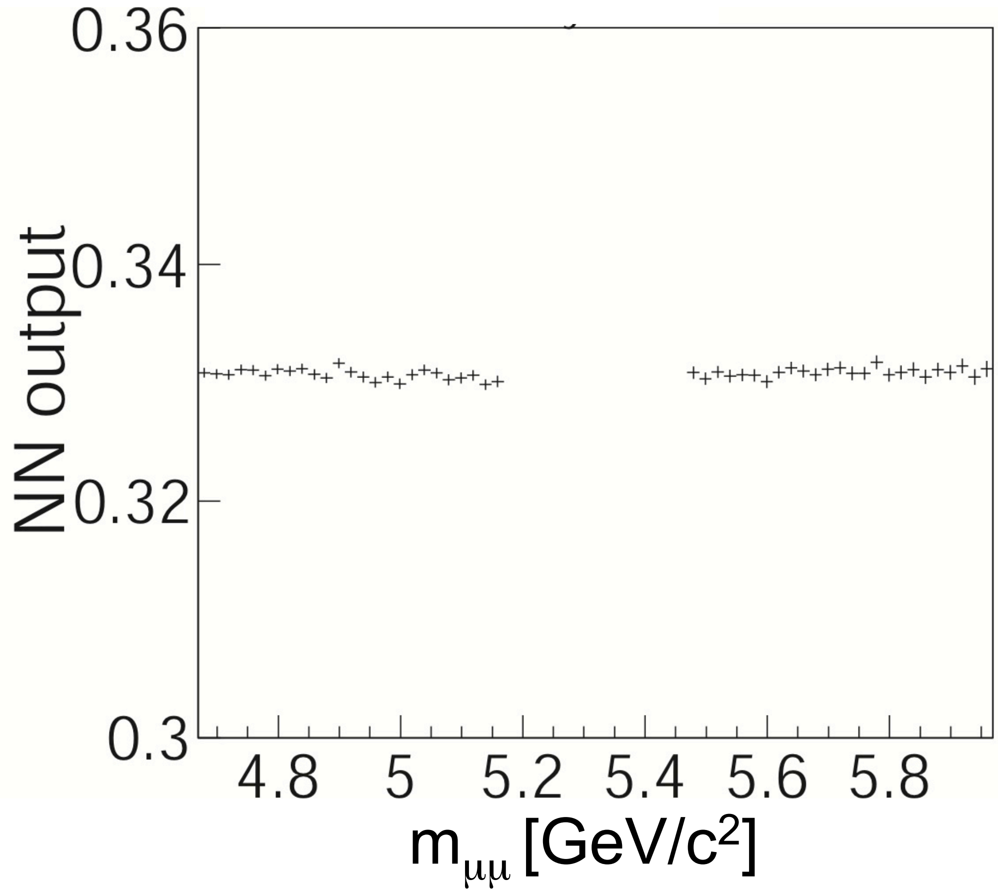
2) Problem with NN
   - NN over trained and biases comb. bgd. low
   - NN has mass bias suppressing B_d events
   - Shape of ν_{NN} distribution poorly modeled

Problems with NN: Overtraining?



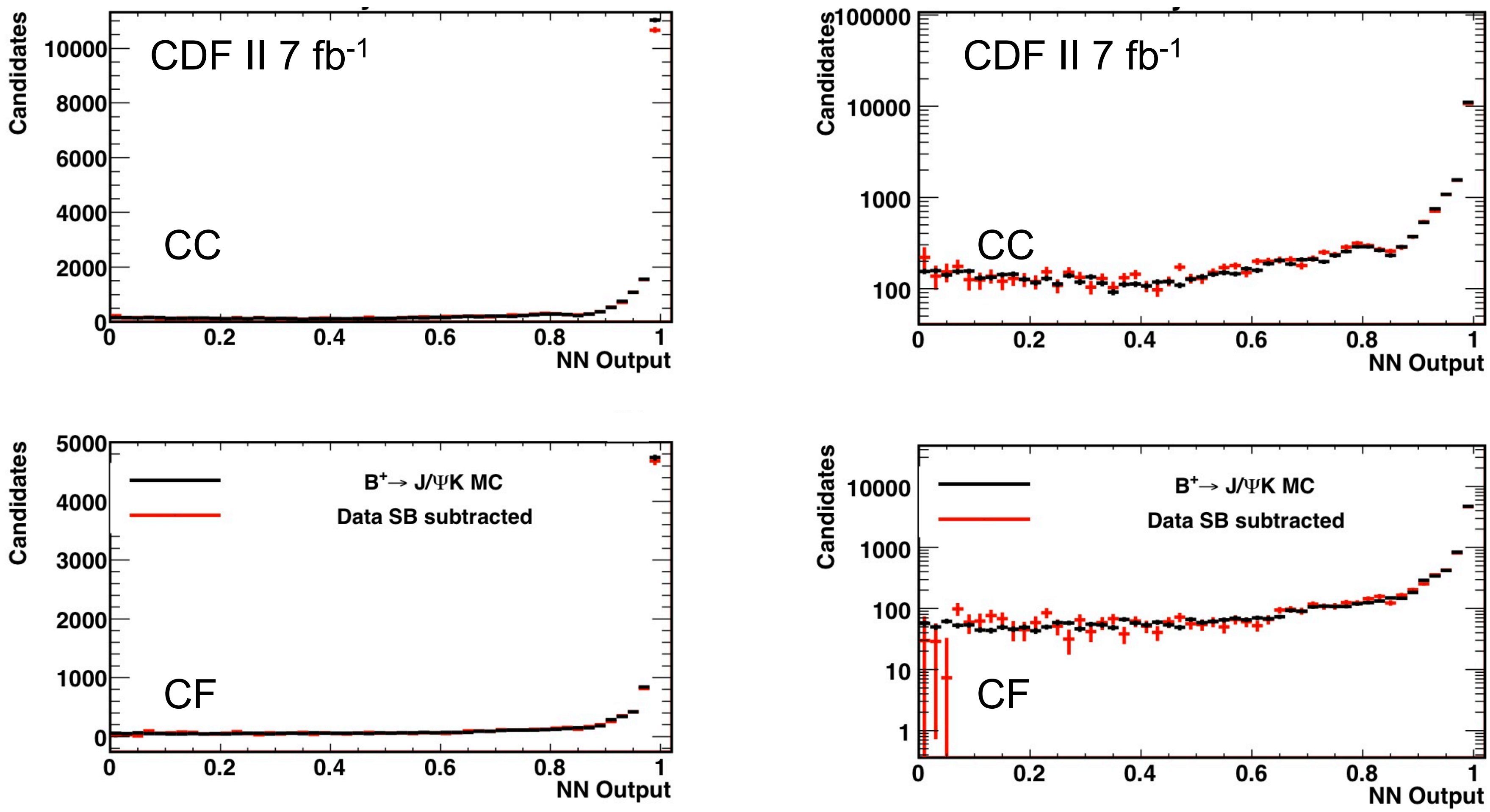
- No evidence of overtraining or bias

Recall



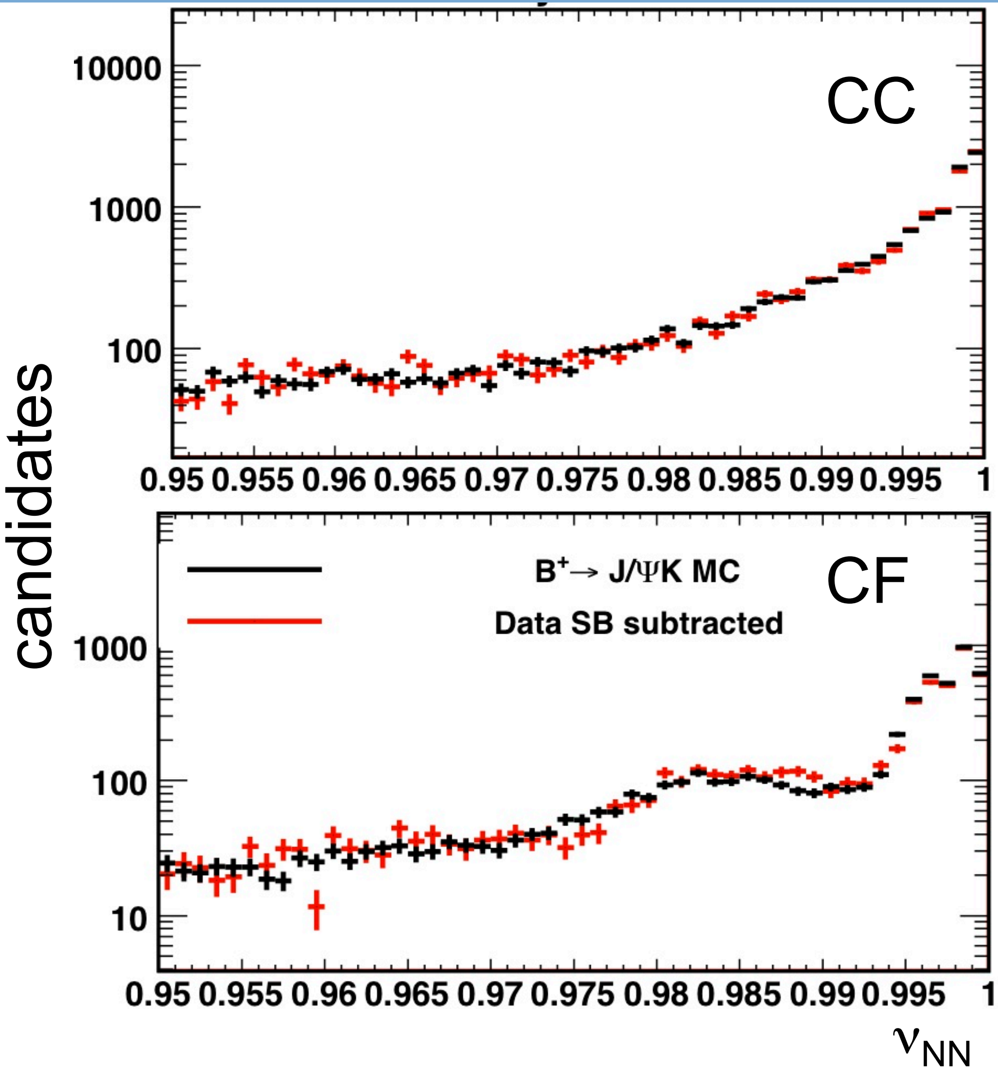
- No evidence that ν_{NN} is correlated with $m_{\mu\mu}$ (cf. pages 36-38)

Recall



- No evidence of a significant MC mis-modeling of ν_{NN} distribution for real B-decays

In addition



- Even in the steeply falling region above 0.99

Possibilities

- Only two possible problems to consider (this is a simple analysis):

✓1) Problem with background estimate
  - e.g. your π/K fake rates are wrong

✓2) Problem with NN
  - ✓NN over trained and biases comb. bgd. Low
  - ✓NN has mass bias suppressing B_d events
  - ✓Shape of ν_{NN} distribution poorly modeled

So?

Our conclusion

From the PRL:

"The source of the data excess in the $0.970 < \nu_{NN} < 0.987$ bin of the B_s signal region is investigated. ... Because the data in the B_d search region shows no excess, problems with the background estimates are ruled out. … Problems with the NN are ruled out … [since] studies find no evidence of a ν_{NN} - $m_{\mu\mu}$ correlation, no evidence of overtraining, and no evidence of a significant mis-modeling of the ν_{NN} shape…. In short, there is no evidence that the excess in this bin is caused by a mistake or systematic error in our background estimates or our modeling of the ν_{NN} performance and distribution. The most plausible remaining explanation is that this is a statistical fluctuation."

Our conclusion

"For our central result we use the full set of bins that had been established a priori since this represents an unbiased choice."

$$p - \text{value(b-only)} \;=\; 0.27\%$$
$$p - \text{value(b+SM)} \;=\; 1.9\%$$

$$BF(B_s \to \mu^+\mu^-) = \left(1.8^{+1.1}_{-0.9}\right) \times 10^{-8}$$

$$4.6 \times 10^{-9} < BF(B_s \to \mu^+\mu^-) < 3.9 \times 10^{-8} \,@\, 90\% CL$$

FYI

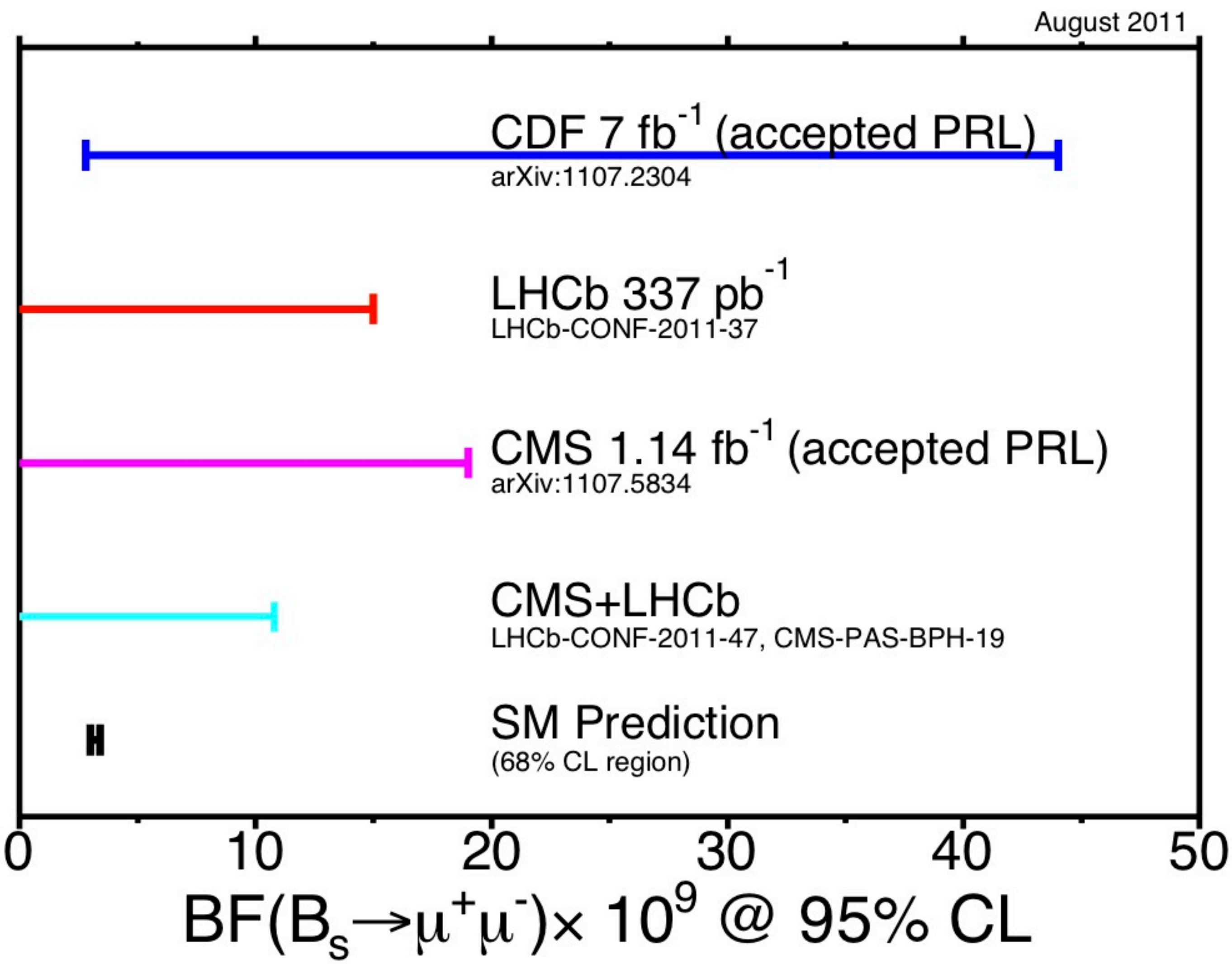
"…if we remove the 0.970 < ν_{NN} < 0.987 bin the results are not significantly affected."

All bins ($0.70 < \nu_{NN}$)	2 Highest Bins ($0.987 < \nu_{NN}$)
$p-\text{value(b-only)} = 0.27\%$	$p-\text{value(b-only)} = 0.66\%$
$p-\text{value(b+SM)} = 1.9\%$	$p-\text{value(b+SM)} = 4.1\%$
$BF(B_s \to \mu^+\mu^-) = \left(1.8^{+1.1}_{-0.9}\right) \times 10^{-8}$	$BF(B_s \to \mu^+\mu^-) = \left(1.4^{+1.0}_{-0.8}\right) \times 10^{-8}$
$(4.6-39)\times 10^{-9}\ @90\%CL$	$(3.3-33)\times 10^{-9}\ @90\%CL$

Closing Remarks

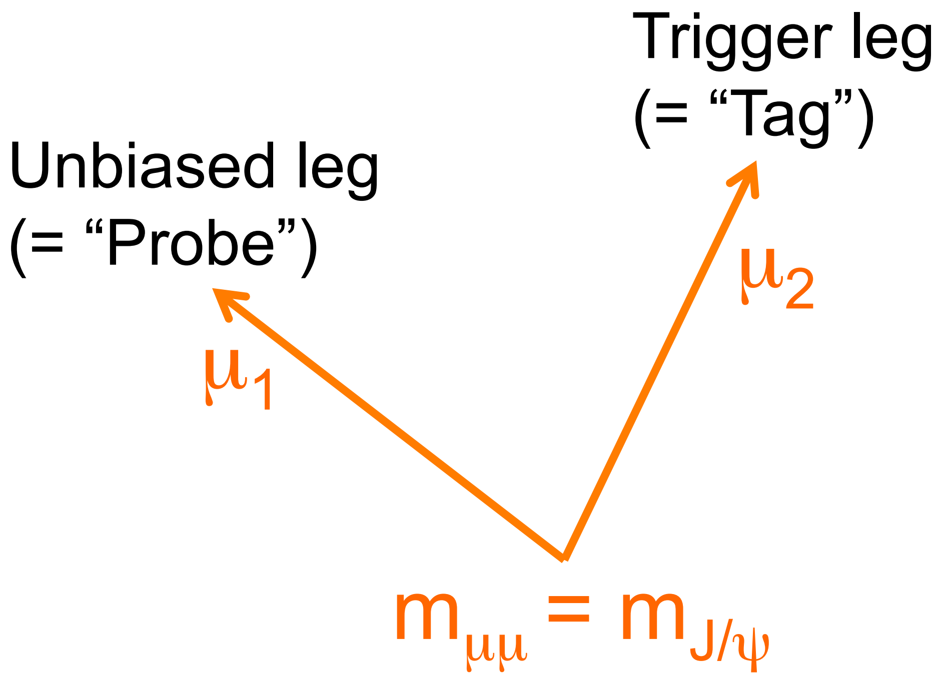
- CDF has an excess of $B_s \rightarrow \mu^+\mu^-$ events at the level of >2.7σ relative to bgd-only
- The fitted BF is compatible with the results from other experiments and the SM
- CDF will increase the data set by another 40% and publish a PRD

Closing Remarks



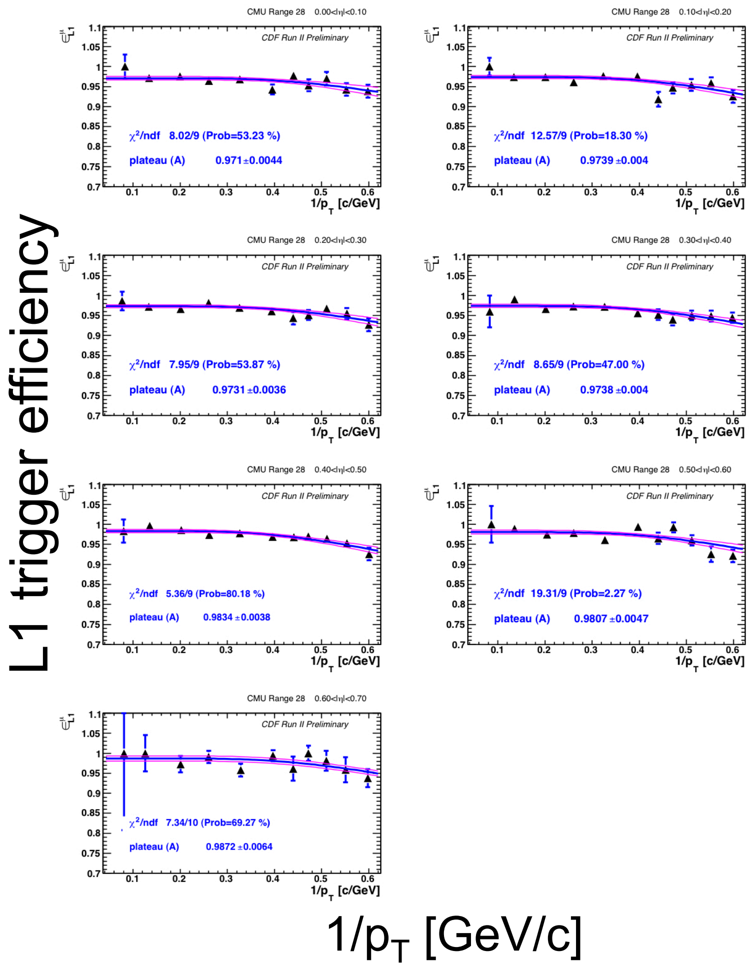
Backup Slides

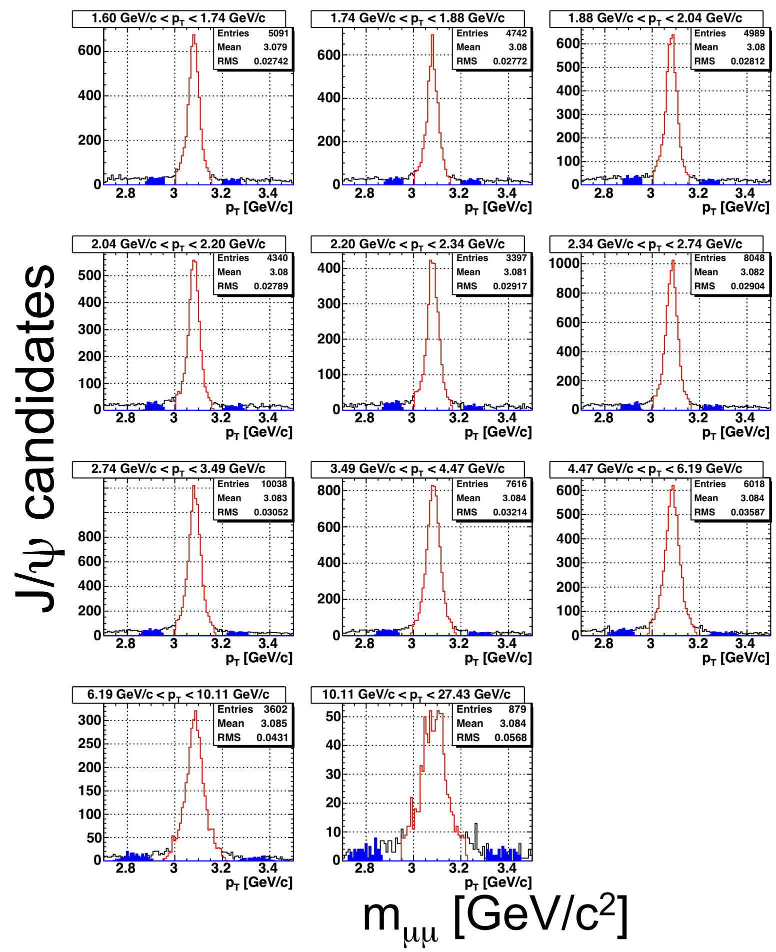
Trigger efficiency



- “Tag-and-Probe” method using $J/\psi \rightarrow \mu^+\mu^-$ events collected with a single-leg μ trigger

Trigger Efficiency





- $\varepsilon_{trig} = \mathcal{F}(p_T, \phi, \eta, \text{run\#}, |z|^{COT}_{min})$

Expected Limit

We used the set of requirements which yielded the minimum *a priori* expected BR Limit:

$$\left\langle BR(B_s \rightarrow \mu^+\mu^-)\right\rangle = \frac{\left\langle N_{signal}^{90\%CL}\right\rangle}{\alpha \cdot \varepsilon_{total} \cdot \sigma_{B_s} \int L dt}$$

where we've summed over all possible n_{obs}:

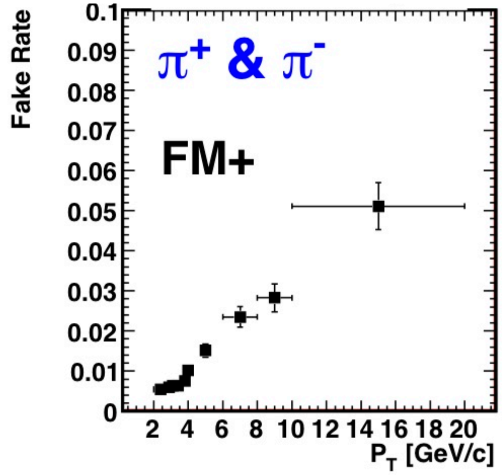
$$\left\langle N_{signal}^{90\%CL}\right\rangle = \sum_{n_{obs}=0}^{\infty} \mathrm{P}(n_{obs} \mid n_{bg}) \cdot N_{signal}^{90\%CL}(n_{bg}, \Delta_{bg}, \Delta_{\alpha\cdot\varepsilon})$$

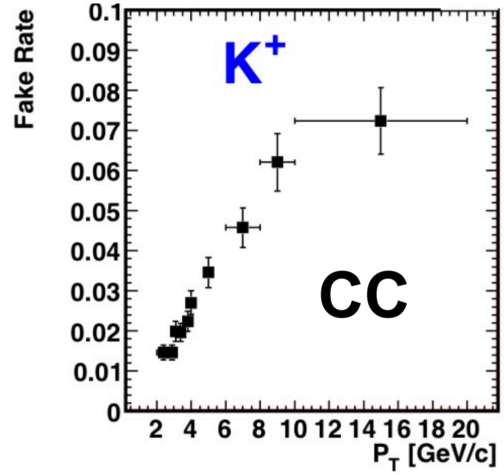
Poisson prob of observing n_{obs} when expecting n_{bg}

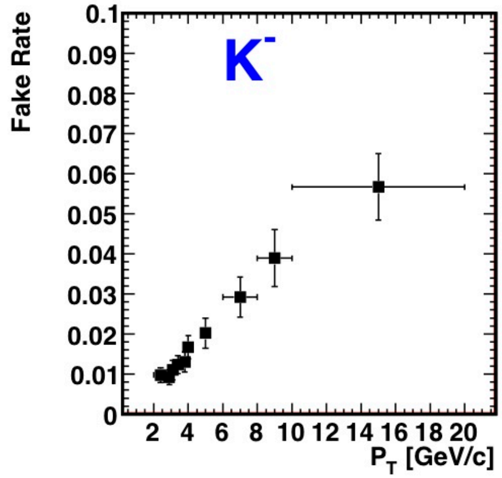
90% CL UL on N_{signal} when expecting n_{bg} bkgd evts using Bayesian Method and including uncertainties

Hadron to muon Fake Rates: FM+ Sample

CDF II 7 fb^{-1}



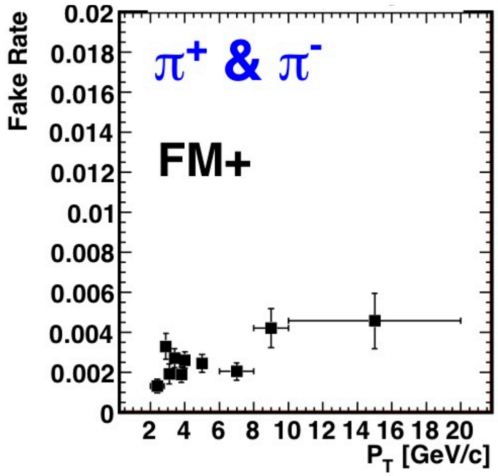


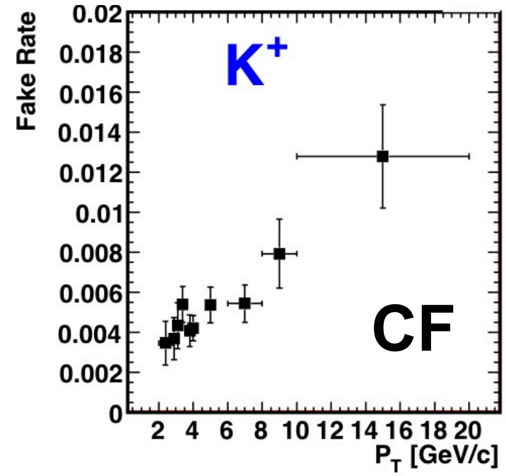


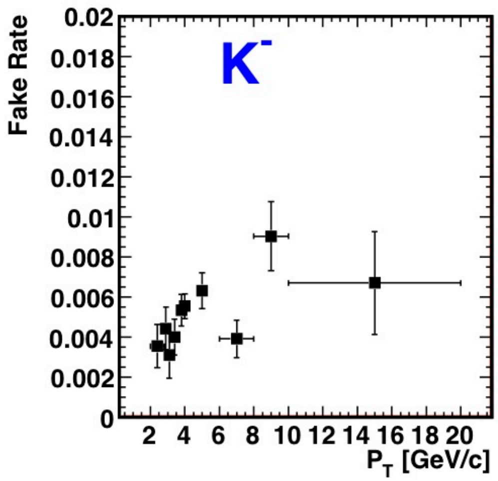
- These rates are about x5 larger than signal event sample

Hadron to muon Fake Rates: FM+ Sample

CDF II 7 fb^{-1}

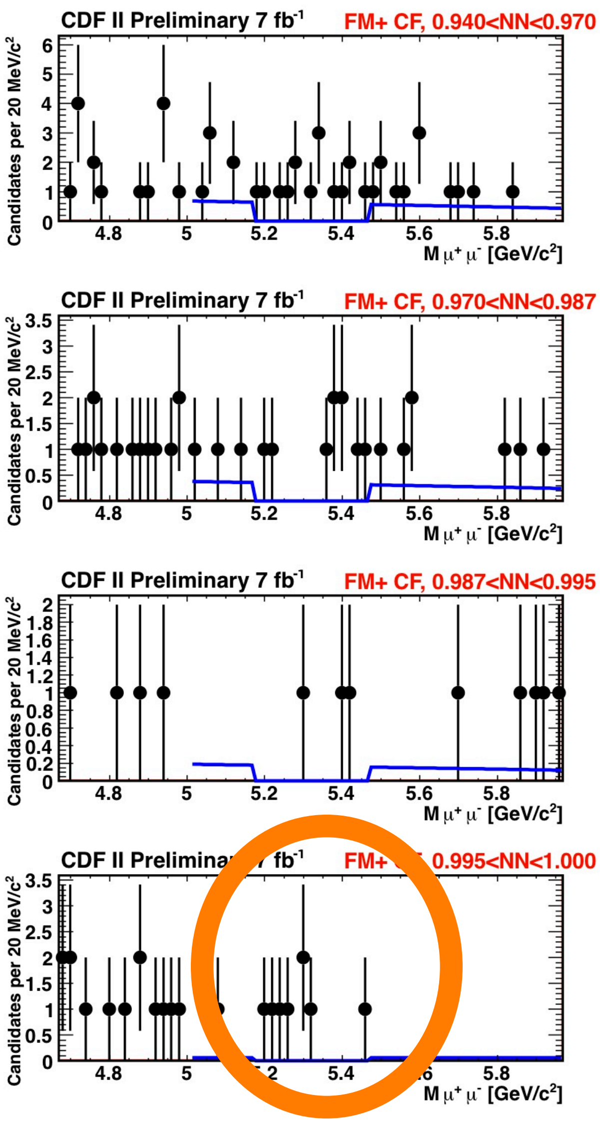






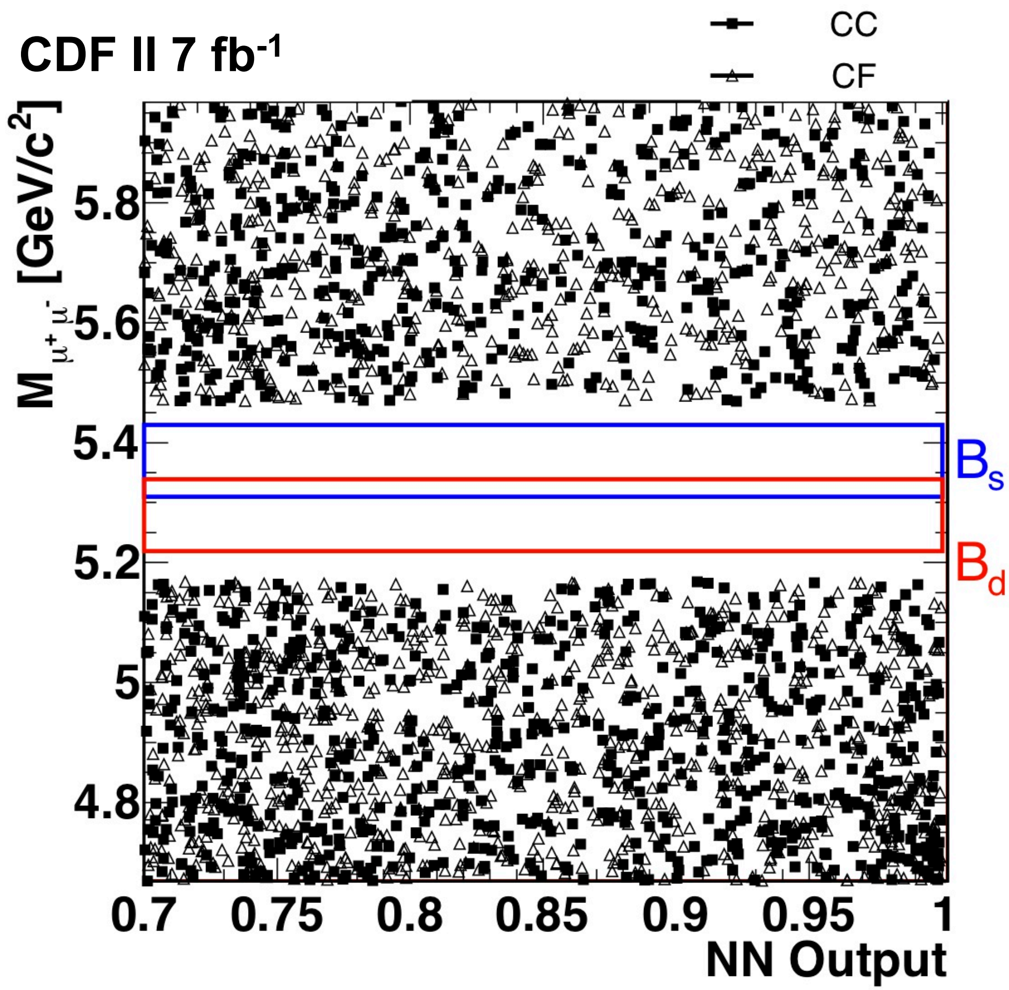
- These rates are about x2 larger than signal event sample

Cross-check Background Methodology



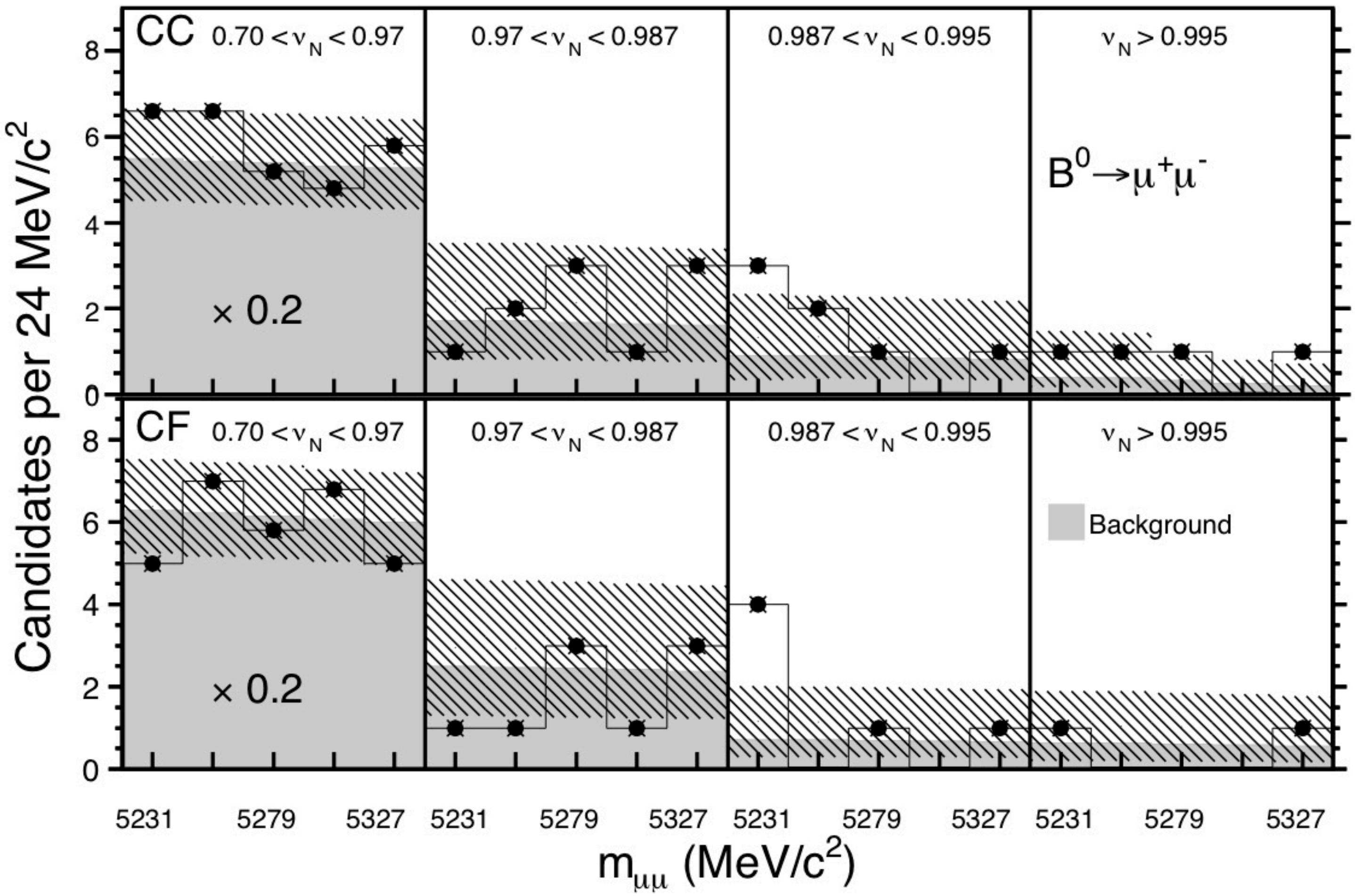
- Excess in this bin looks more consistent with combinatoric than B→hh

Search sample: ν_{NN} vs $m_{\mu\mu}$ distribution



- Extended signal region blinded

Results in B_d Region



- B_d results for CC and CF separately

Results in B_d Region

	Mass Bin (GeV/c^2)	5.219-5.243	5.243-5.267	5.267-5.291	5.291-5.315	5.315-5.339	Total
CC NN bin	Exp Bkg	8.32±0.64	8.25±0.63	8.17±0.63	8.09±0.62	8.01±0.62	40.83
0.7-0.76	Obs	11	10	6	5	7	39
CC NN bin	Exp Bkg	8.75±0.66	8.67±0.65	8.58±0.65	8.5±0.64	8.41±0.63	42.91
0.76-0.85	Obs	8	10	5	6	9	38
CC NN bin	Exp Bkg	3.69±0.41	3.66±0.4	3.62±0.4	3.58±0.4	3.54±0.39	18.09
0.85-0.9	Obs	7	2	6	5	4	24
CC NN bin	Exp Bkg	3.66±0.4	3.62±0.4	3.58±0.4	3.54±0.39	3.5±0.39	17.9
0.9-0.94	Obs	5	8	5	5	5	28
CC NN bin	Exp Bkg	3.0±0.36	2.97±0.36	2.93±0.36	2.9±0.35	2.86±0.35	14.65
0.94-0.97	Obs	2	3	4	3	4	16
CC NN bin	Exp Bkg	1.71±0.50	1.69±0.50	1.67±0.50	1.64±0.49	1.62±0.49	8.33
0.97-0.987	Obs	1	2	3	1	3	10
CC NN bin	Exp Bkg	0.90±0.28	0.89±0.28	0.86±0.27	0.84±0.27	0.81±0.27	4.29
0.987-0.995	Obs	3	2	1	0	1	7
CC NN bin	Exp Bkg	0.40±0.21	0.38±0.20	0.32±0.17	0.25±0.15	0.20±0.14	1.54
0.995-1	Obs	1	1	1	0	1	4
CF NN bin	Exp Bkg	8.89±0.68	8.79±0.67	8.68±0.66	8.58±0.65	8.47±0.65	43.41
0.7-0.76	Obs	7	10	10	12	9	48
CF NN bin	Exp Bkg	9.9±0.72	9.78±0.71	9.66±0.7	9.54±0.69	9.42±0.69	48.31
0.76-0.85	Obs	7	10	11	13	10	51
CF NN bin	Exp Bkg	5.15±0.5	5.09±0.49	5.02±0.49	4.96±0.48	4.9±0.47	25.12
0.85-0.9	Obs	3	4	1	2	1	11
CF NN bin	Exp Bkg	4.06±0.44	4.01±0.43	3.96±0.43	3.91±0.42	3.86±0.42	19.8
0.9-0.94	Obs	3	5	5	6	4	23
CF NN bin	Exp Bkg	3.45±0.4	3.41±0.4	3.37±0.39	3.32±0.39	3.28±0.38	16.83
0.94-0.97	Obs	5	6	2	1	1	15
CF NN bin	Exp Bkg	2.50±0.59	2.47±0.58	2.44±0.58	2.40±0.57	2.37±0.56	12.17
0.97-0.987	Obs	1	1	3	1	3	9
CF NN bin	Exp Bkg	0.71±0.25	0.70±0.25	0.69±0.25	0.68±0.24	0.67±0.24	3.44
0.987-0.995	Obs	4	0	1	0	1	6
CF NN bin	Exp Bkg	0.62±0.42	0.62±0.42	0.60±0.41	0.57±0.40	0.55±0.39	2.97
0.995-1	Obs	1	0	0	0	1	2

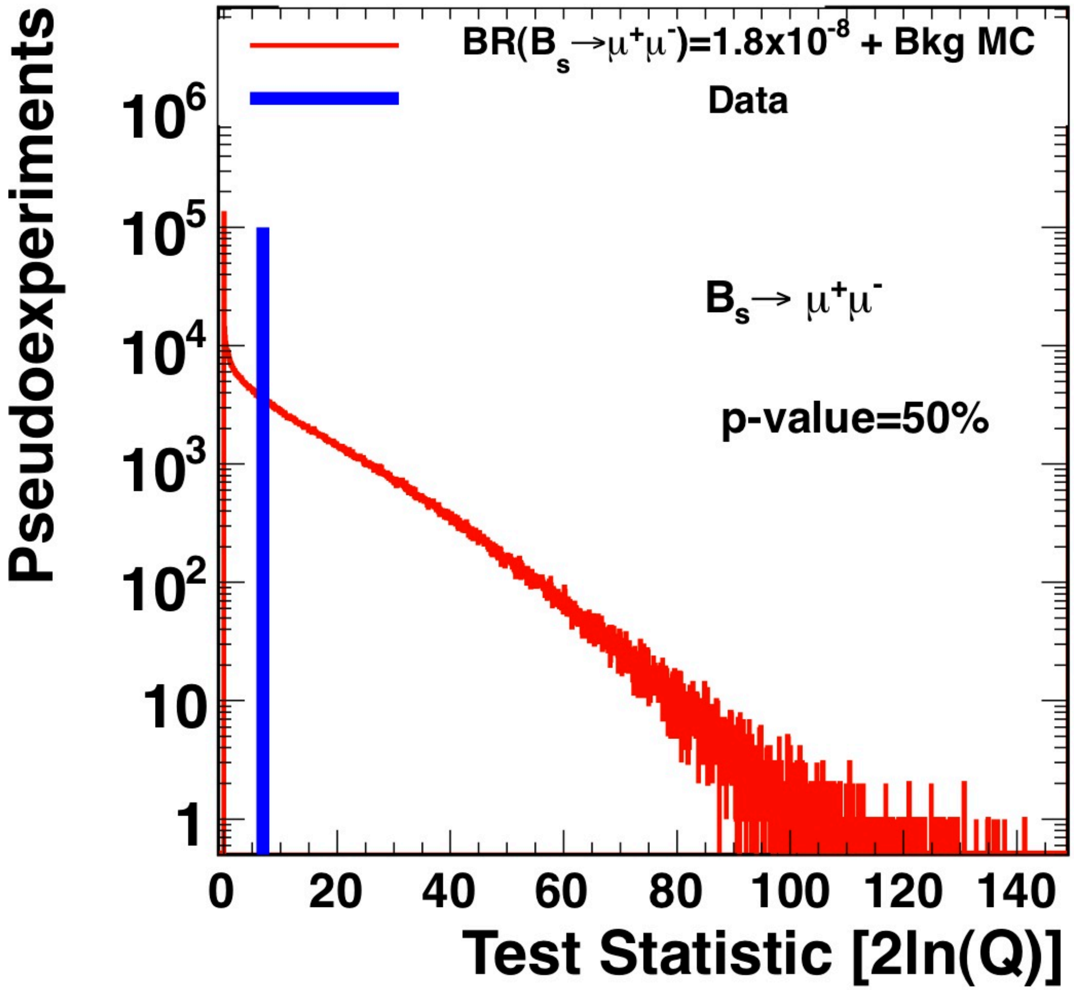
Table: B_d signal window for CC(top) and CF(bottom): Expected backgrounds, including $B \to hh$, and number of observed events.

Results in B_s Region

	Mass Bin (GeV/c^2)	5.31-5.334	5.334-5.358	5.358-5.382	5.382-5.406	5.406-5.43	Total
CC NN bin	Exp Bkg	8.02±0.62	7.94±0.61	7.87±0.61	7.79±0.6	7.71±0.59	39.34
0.7-0.76	Obs	9	6	6	2	5	28
CC NN bin	Exp Bkg	8.43±0.64	8.34±0.63	8.26±0.62	8.18±0.62	8.1±0.61	41.32
0.76-0.85	Obs	8	6	11	11	7	43
CC NN bin	Exp Bkg	3.55±0.39	3.51±0.39	3.48±0.39	3.44±0.38	3.41±0.38	17.4
0.85-0.9	Obs	5	6	2	5	4	22
CC NN bin	Exp Bkg	3.51±0.39	3.47±0.39	3.44±0.38	3.41±0.38	3.37±0.38	17.2
0.9-0.94	Obs	4	5	4	5	7	25
CC NN bin	Exp Bkg	2.87±0.35	2.84±0.35	2.81±0.34	2.78±0.34	2.75±0.34	14.04
0.94-0.97	Obs	4	5	2	3	4	18
CC NN bin	Exp Bkg	1.62±0.49	1.60±0.48	1.58±0.47	1.57±0.47	1.55±0.46	7.92
0.97-0.987	Obs	1	4	7	1	3	16
CC NN bin	Exp Bkg	0.82±0.27	0.80±0.27	0.79±0.26	0.78±0.26	0.78±0.26	3.97
0.987-0.995	Obs	1	1	3	0	0	5
CC NN bin	Exp Bkg	0.21±0.14	0.18±0.13	0.16±0.12	0.16±0.12	0.16±0.12	0.87
0.995-1	Obs	0	1	2	0	1	4
CF NN bin	Exp Bkg	8.49±0.65	8.39±0.64	8.28±0.63	8.17±0.62	8.07±0.61	41.4
0.7-0.76	Obs	8	13	9	9	9	48
CF NN bin	Exp Bkg	9.45±0.69	9.33±0.68	9.21±0.67	9.1±0.66	8.98±0.65	46.07
0.76-0.85	Obs	7	8	7	11	4	37
CF NN bin	Exp Bkg	4.91±0.48	4.85±0.47	4.79±0.46	4.73±0.46	4.67±0.45	23.95
0.85-0.9	Obs	1	5	6	3	5	20
CF NN bin	Exp Bkg	3.87±0.42	3.82±0.41	3.77±0.41	3.73±0.4	3.68±0.4	18.88
0.9-0.94	Obs	4	1	6	3	3	17
CF NN bin	Exp Bkg	3.29±0.38	3.25±0.38	3.21±0.37	3.17±0.37	3.12±0.36	16.04
0.94-0.97	Obs	0	5	3	4	5	17
CF NN bin	Exp Bkg	2.38±0.56	2.34±0.55	2.31±0.54	2.28±0.54	2.25±0.53	11.57
0.97-0.987	Obs	1	4	3	1	2	11
CF NN bin	Exp Bkg	0.67±0.24	0.66±0.24	0.65±0.24	0.64±0.23	0.63±0.22	3.25
0.987-0.995	Obs	1	1	0	1	0	3
CF NN bin	Exp Bkg	0.56±0.39	0.54±0.38	0.53±0.38	0.52±0.37	0.51±0.36	2.66
0.995-1	Obs	1	1	0	1	1	4

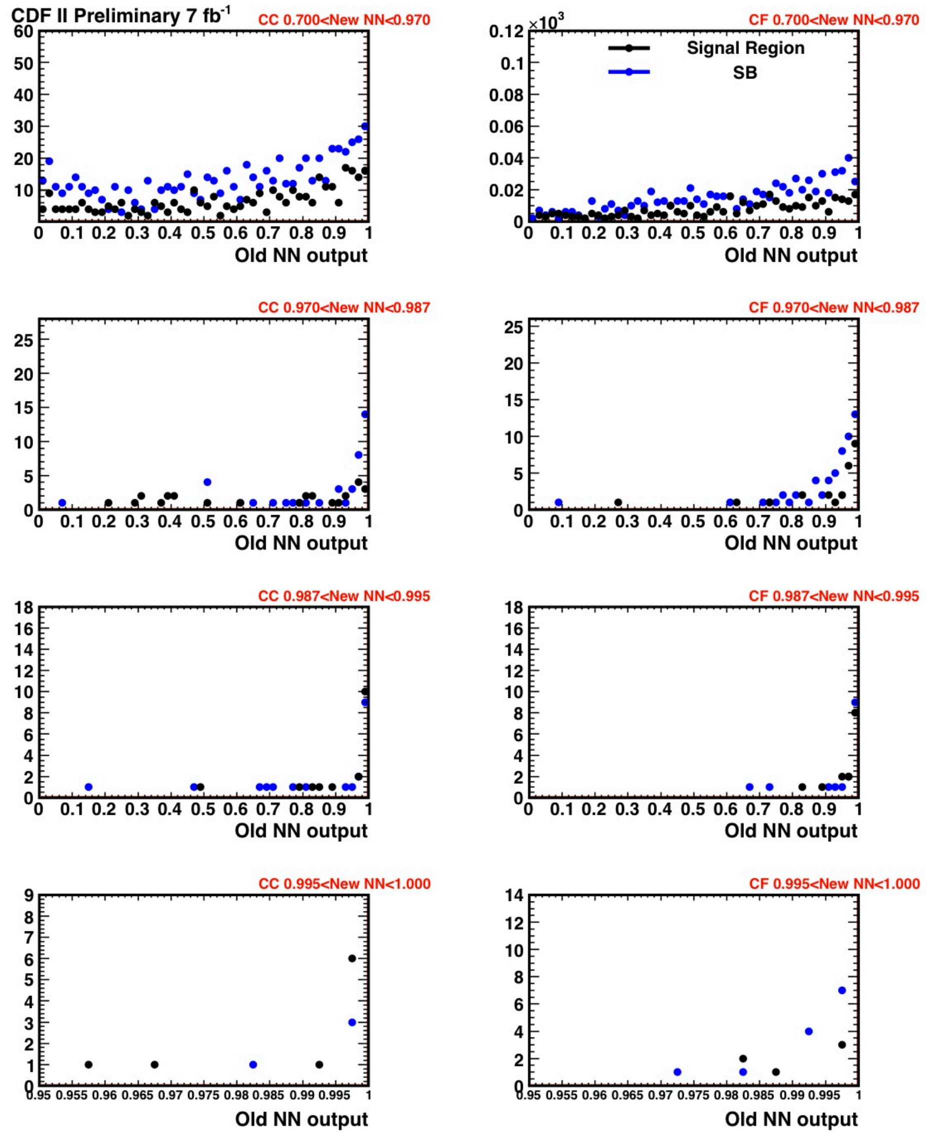
Table: B_s signal window for CC(top) and CF(bottom): Expected backgrounds, including $B \rightarrow hh$, and number of observed events.

p-value using best fit BF



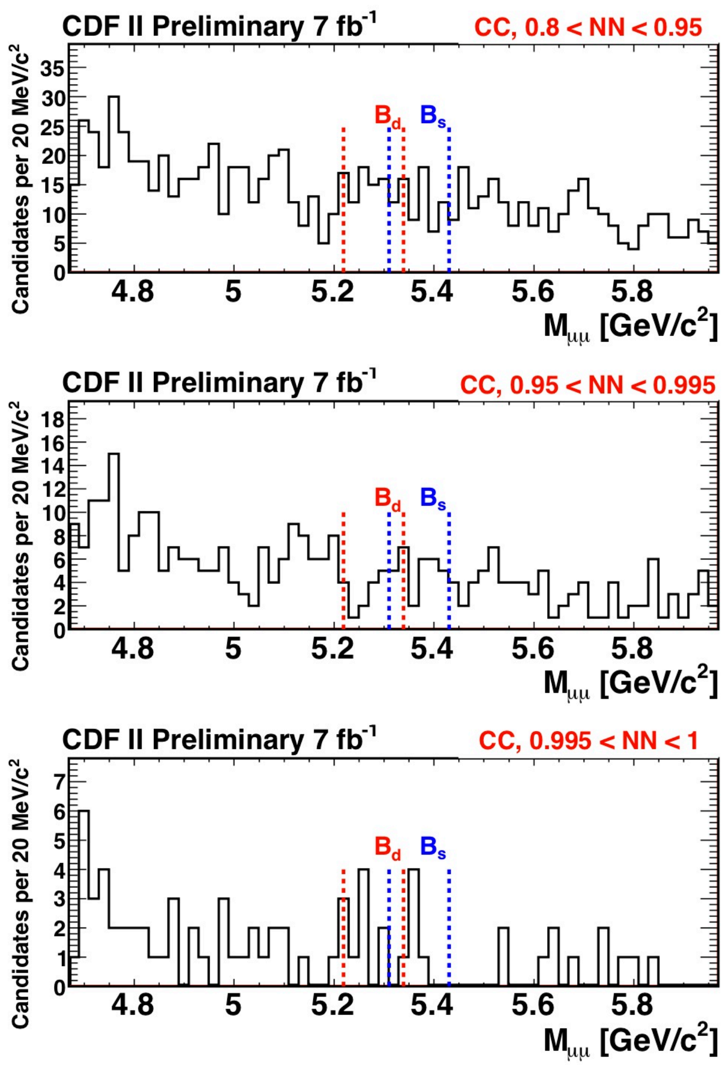
- pseudo-experiments used BF=best fit BF=5.6*SM

Comparisons with Old NN



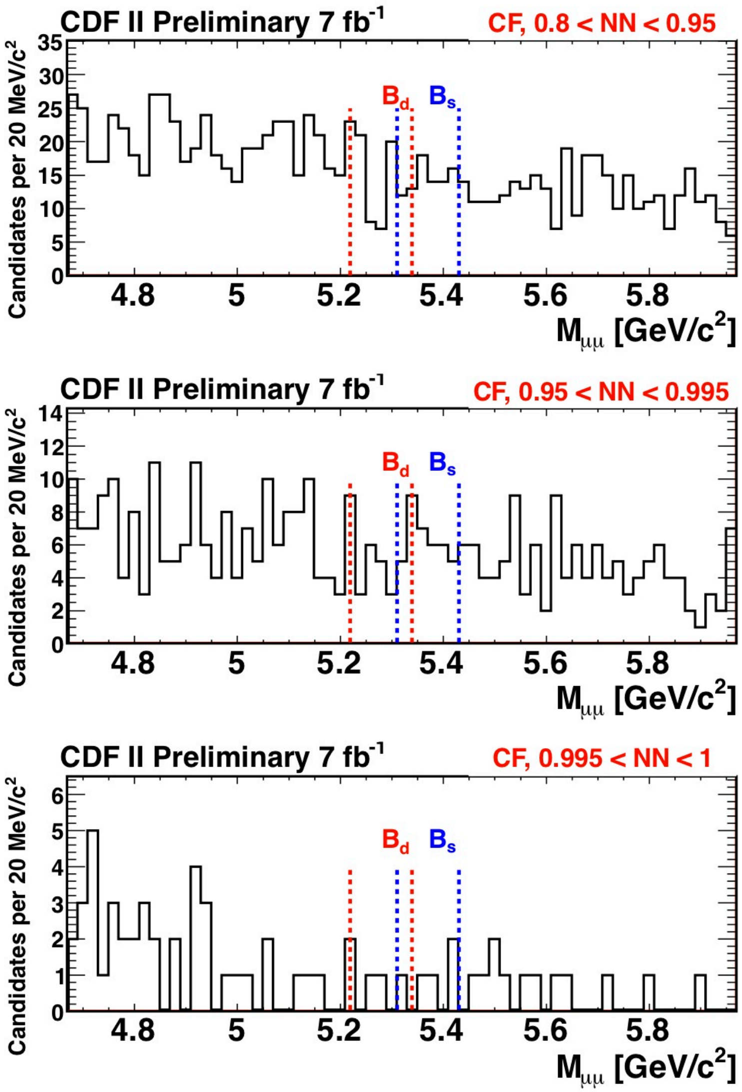
- The high score newNN events also high score in the oldNN

Comparisons with Old NN



- $m_{\mu\mu}$ distributions using oldNN and binning optimized for oldNN in 2 fb^{-1} PRL

Comparisons with Old NN



- $m_{\mu\mu}$ distributions using oldNN and binning optimized for oldNN in 2 fb^{-1} PRL