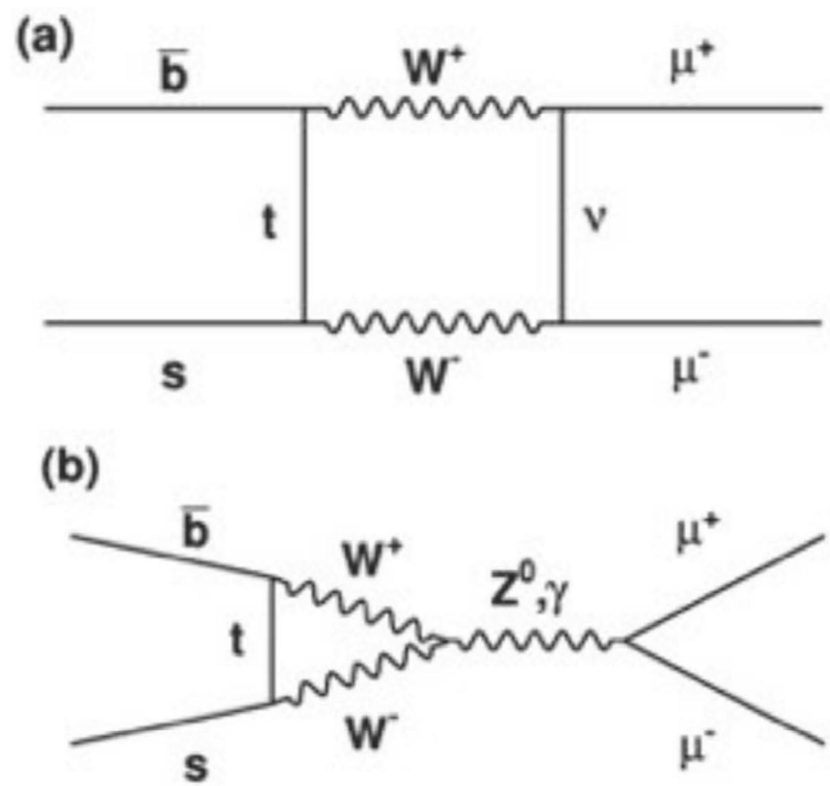
Search for $B \rightarrow \mu^+\mu^-$ at CDF

Kevin Pitts
University of Illinois

DPF 2011, Providence

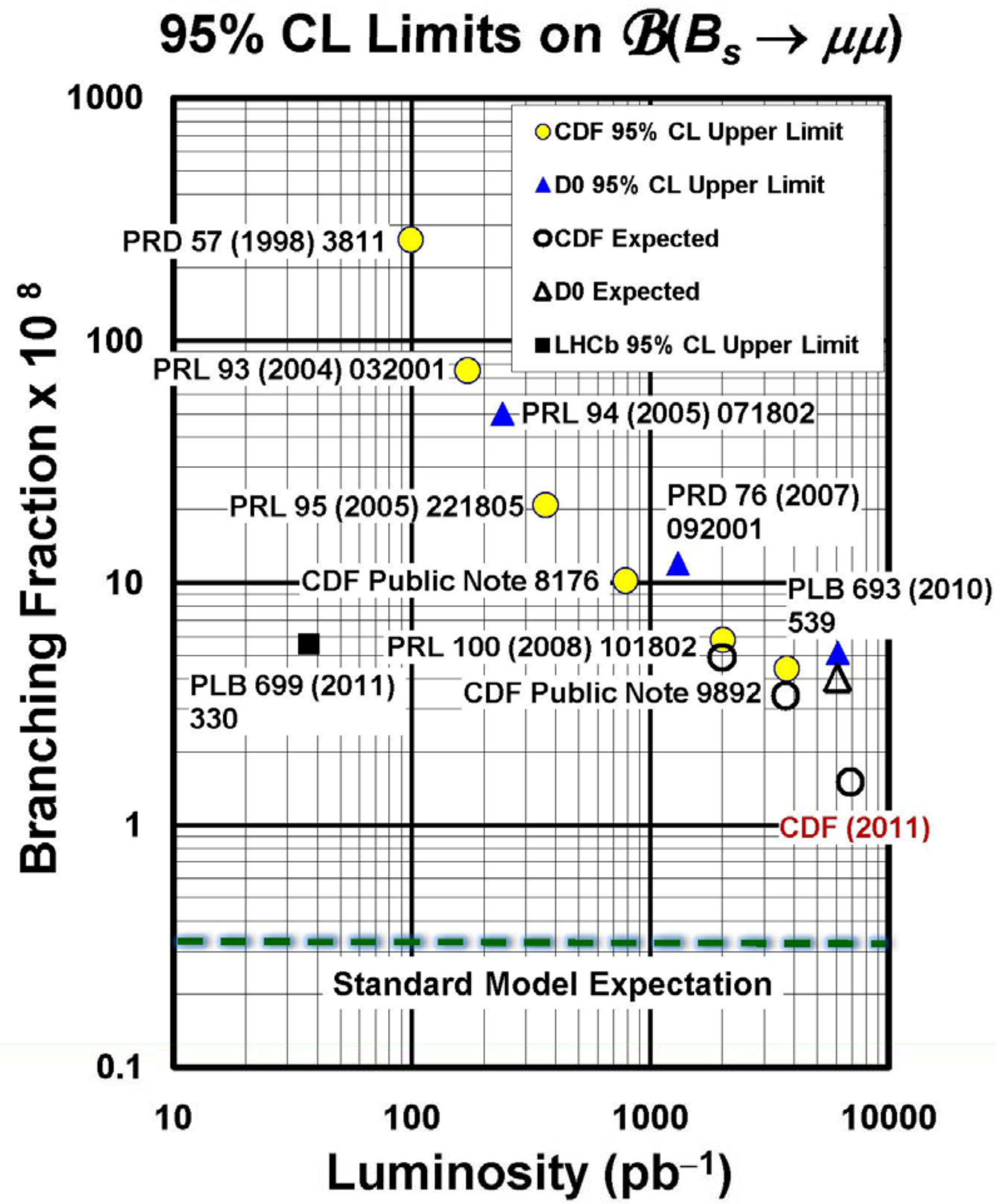
Motivation

- **SM rate well understood**
- **SM rate is small, 3.2×10^{-9}**
- **Broad class of NP models (scalar operators) enhance it by O(1-100)**
- **Clean signature**
- **Current experimental limits closing in on SM sensitivity**



History through spring 2011

- Limit has improved by more than x10 since 2004
- 2011 projected sensitivity about a factor of 5 away from standard model



Normalize to $B^+ \rightarrow J/\psi K^+$ decays

- **Number of reconstructed B_s decays:**

$$N_{B_s} = 2\mathcal{L}\sigma_{b\bar{b}} f_s \mathcal{B}\left(B_s \rightarrow \mu^+\mu^-\right)\alpha_{B_s}\varepsilon_{B_s}$$

- **Number of reconstructed B^+ decays:**

$$N_{B^+} = 2\mathcal{L}\sigma_{b\bar{b}} f_u \mathcal{B}\left(B^+ \rightarrow J/\psi K^+\right)\mathcal{B}\left(J/\psi \rightarrow \mu^+\mu^-\right)\alpha_{B^+}\varepsilon_{B^+}$$

- **σ_{bb} and $\mathcal{L}$ cancel in ratio:**

$$\mathcal{B}\left(B_s \rightarrow \mu^+\mu^-\right) = \frac{N_{B_s}}{N_{B^+}} \frac{\alpha_{B^+}\varepsilon_{B^+}}{\alpha_{B_s}\varepsilon_{B_s}} \frac{f_u}{f_s} \mathcal{B}\left(B^+ \rightarrow J/\psi K^+\right)\mathcal{B}\left(J/\psi \rightarrow \mu^+\mu^-\right)$$

$$= \frac{N_{B_s}}{N_{B^+}} \frac{\alpha_{B^+}}{\alpha_{B_s}} \frac{\varepsilon_{B^+}^{trig}}{\varepsilon_{B_s}^{trig}} \frac{\varepsilon_{B^+}^{reco}}{\varepsilon_{B_s}^{reco}} \frac{1}{\varepsilon_{B_s}^{NN}} \frac{f_u}{f_s} \mathcal{B}\left(B^+ \rightarrow J/\psi K^+\right)\mathcal{B}\left(J/\psi \rightarrow \mu^+\mu^-\right)$$

In a nutshell

$$\mathcal{B}\left(B_s \to \mu^+\mu^-\right) = \frac{N_{B_s}}{N_{B^+}} \frac{\alpha_{B^+}}{\alpha_{B_s}} \frac{\varepsilon^{trig}_{B^+}}{\varepsilon^{trig}_{B_s}} \frac{\varepsilon^{reco}_{B^+}}{\varepsilon^{reco}_{B_s}} \frac{1}{\varepsilon^{NN}_{B_s}} \frac{f_u}{f_s} \mathcal{B}\left(B^+ \to J/\psi K^+\right) \mathcal{B}\left(J/\psi \to \mu^+\mu^-\right)$$

Measure yields

Efficiencies from data & MC

Fragmentation fractions and BR from PDG

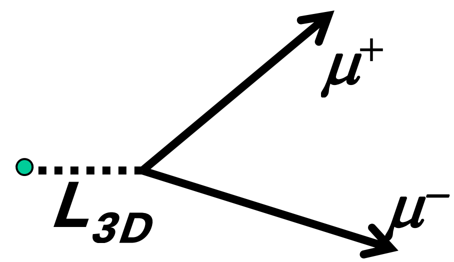
Acceptance from Monte Carlo

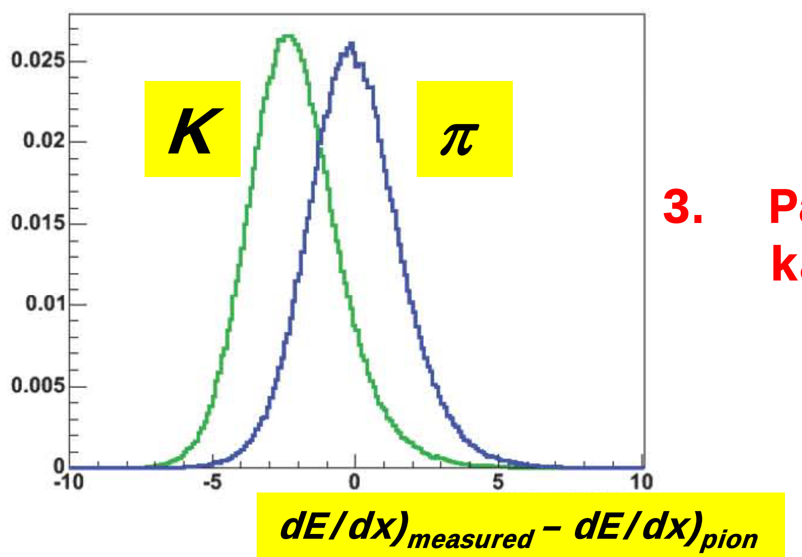
Additional aspects:

- Use neural net, optimize selection
- Extraction of *N* additionally requires evaluation of background contributions.
- Analysis is statistics limited
- Many systematics cancel in ratio, dominant syst. is f_u/f_s!

Experimental considerations

1. Want high yield, need trigger/DAQ to handle high event rates
2. Use long *B* lifetime to reject enormous prompt background, need good impact parameter resolution.

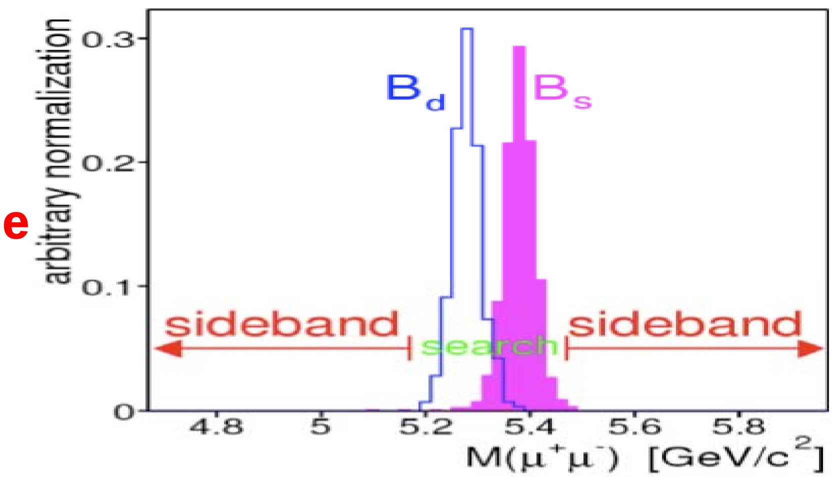


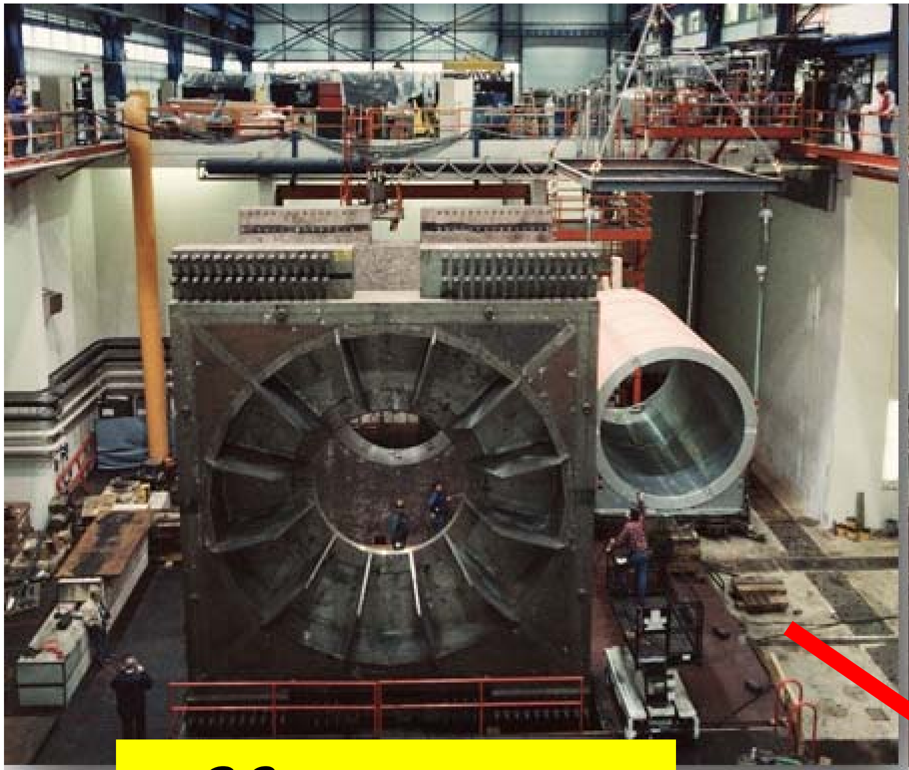


3. Particle ID reduces fake muon (particularly kaon) background

4. Good mass resolution, minimize the amount of background under peak

CDF: $\sigma_{\mu\mu}$ = 24 MeV/c^2 at 5.3 GeV



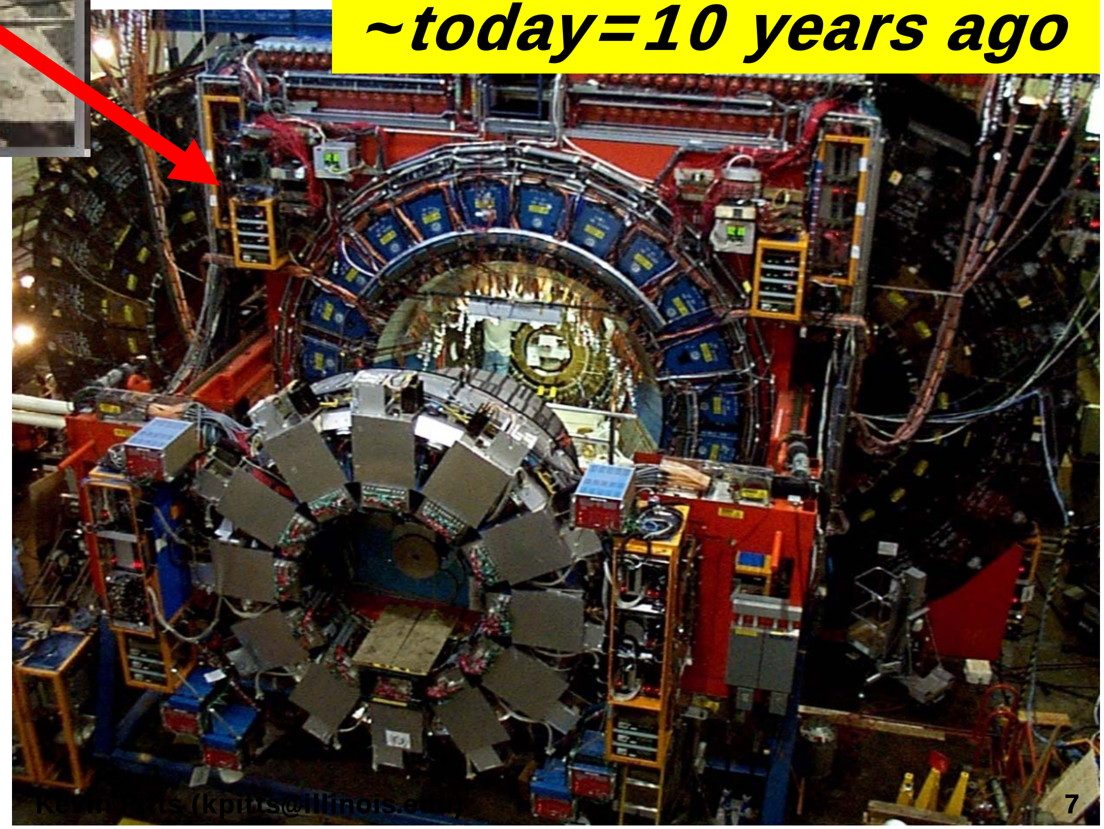


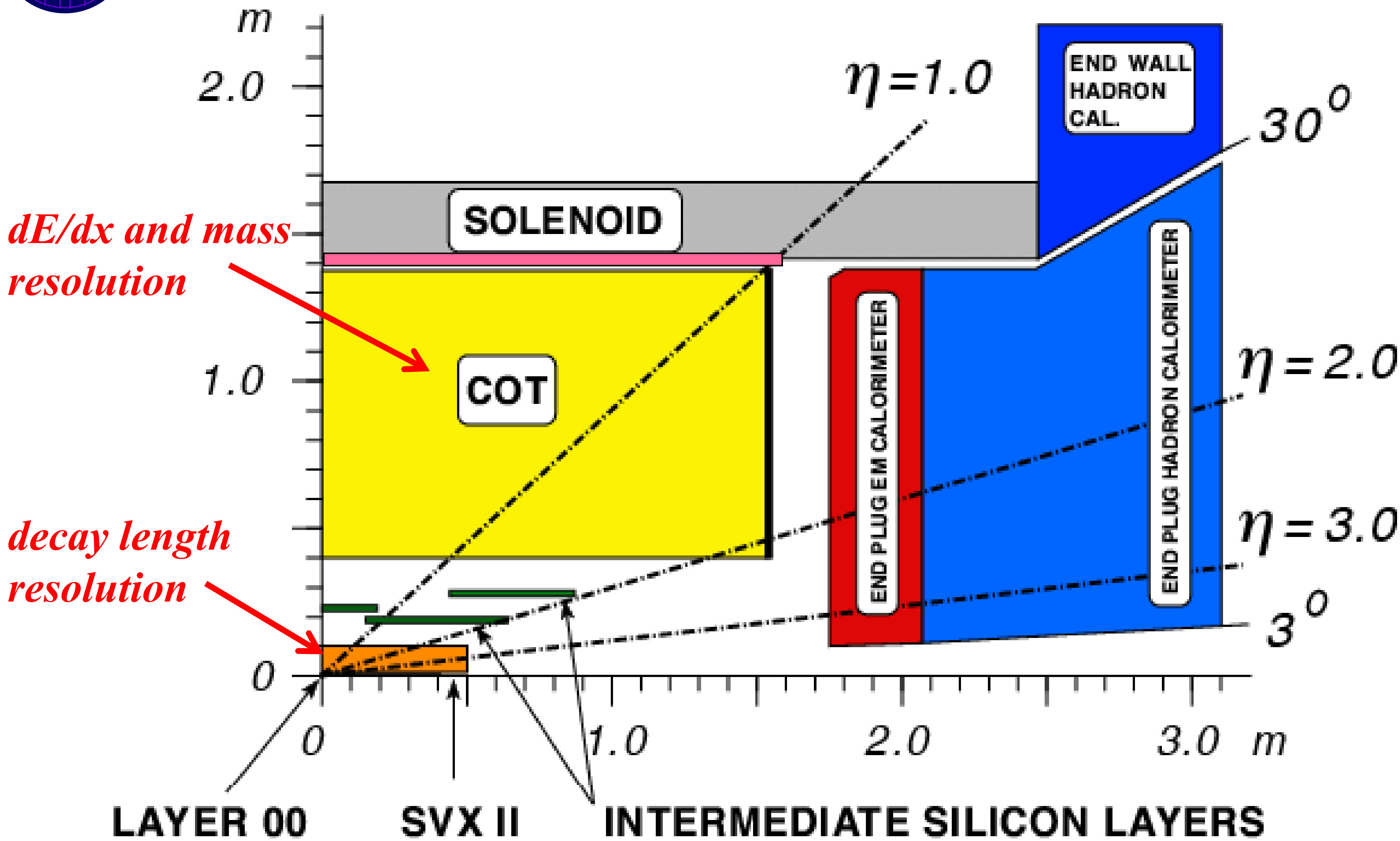
~30 years ago

The CDF Detector

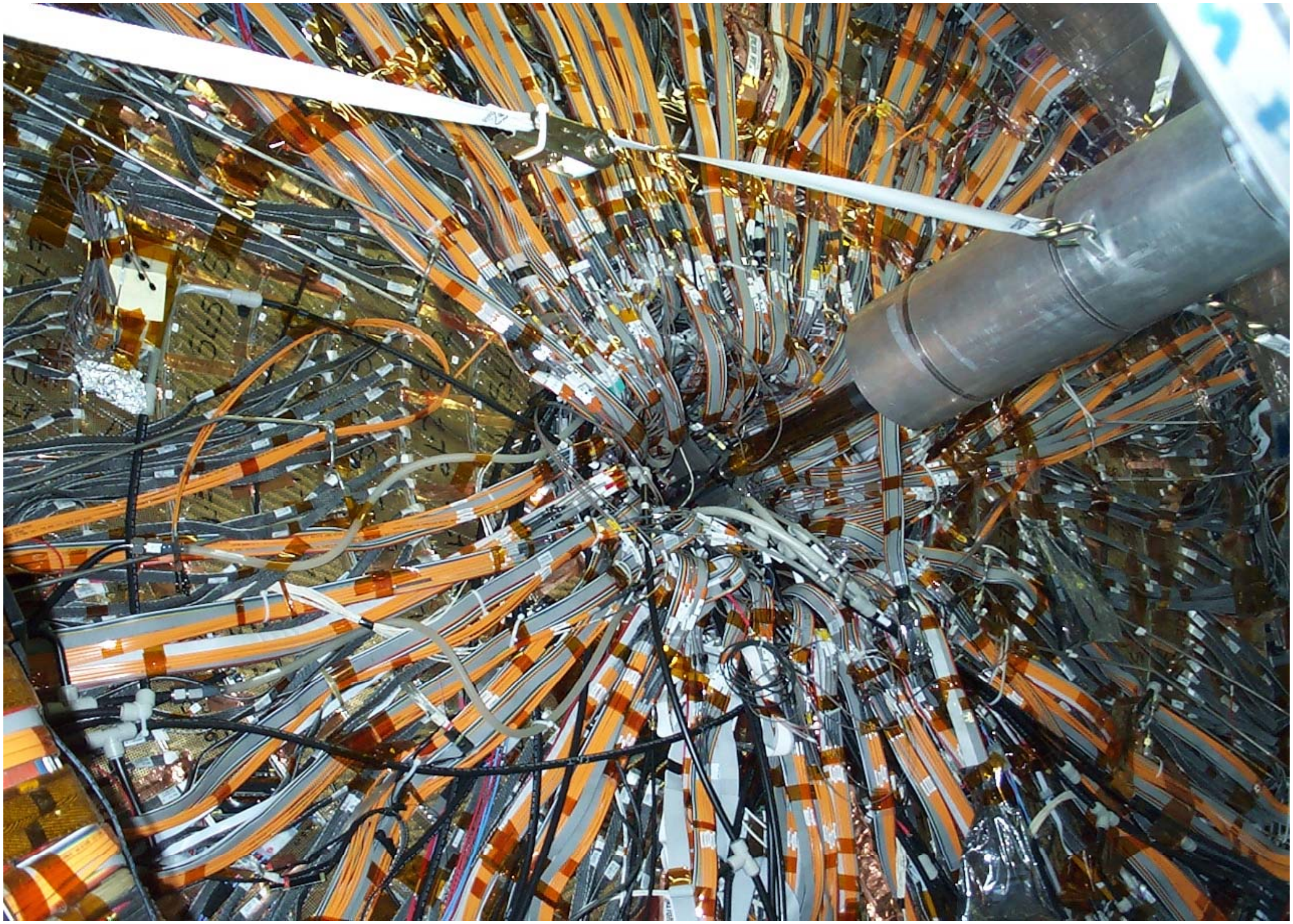
- **Central tracking**
- **Silicon vertex detector**
- **Good lepton identification**
- **Particle ID (TOF and dE/dx)**
- **High rate trigger/DAQ system**

~today=10 years ago




m
2.0
1.0
0
0
1.0
2.0
3.0 m
η=1.0
η = 2.0
η = 3.0
30°
3°
END WALL HADRON CAL.
SOLENOID
COT
END PLUG EM CALORIMETER
END PLUG HADRON CALORIMETER
dE/dx and mass resolution
decay length resolution
LAYER 00
SVX II
INTERMEDIATE SILICON LAYERS

What it looks like ~~in~~ on the end

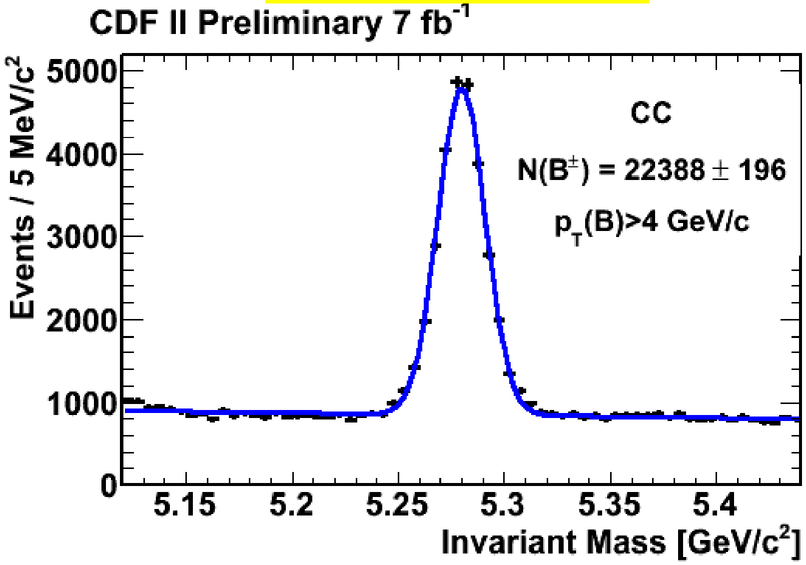


Baseline selection

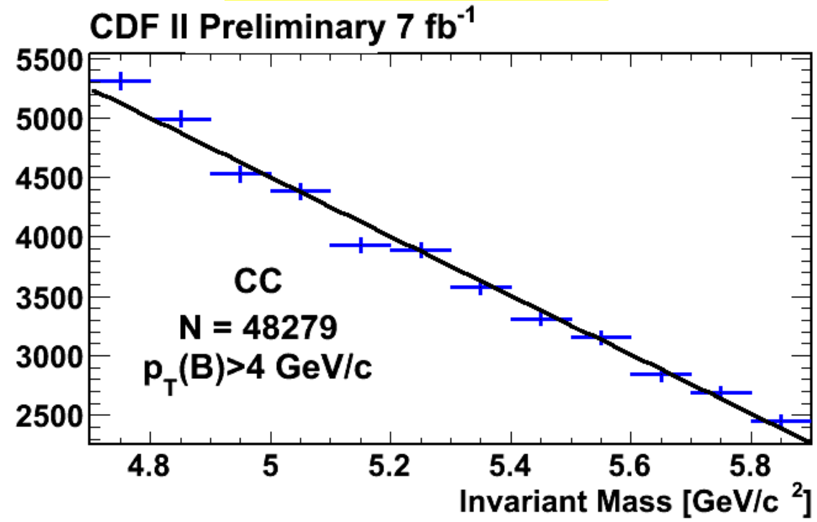
Before neural net selection.

central/central events

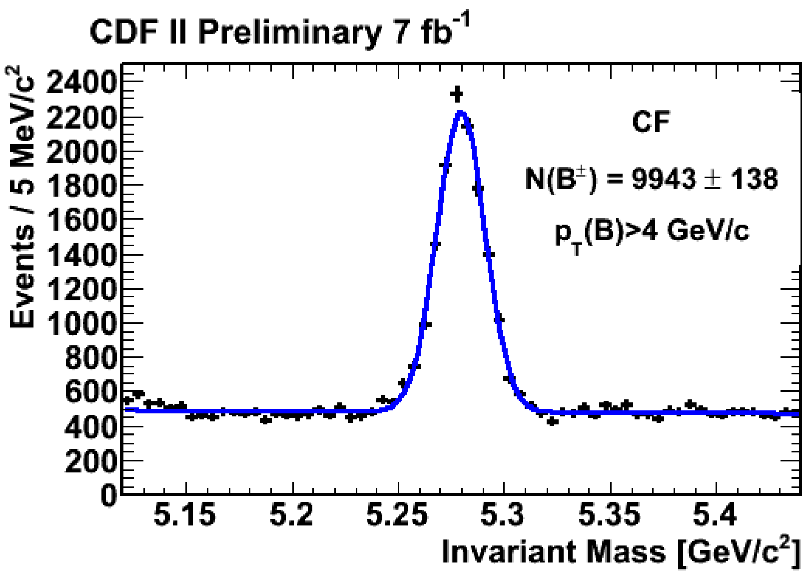
$B^+ \rightarrow J/\psi\, K^+$

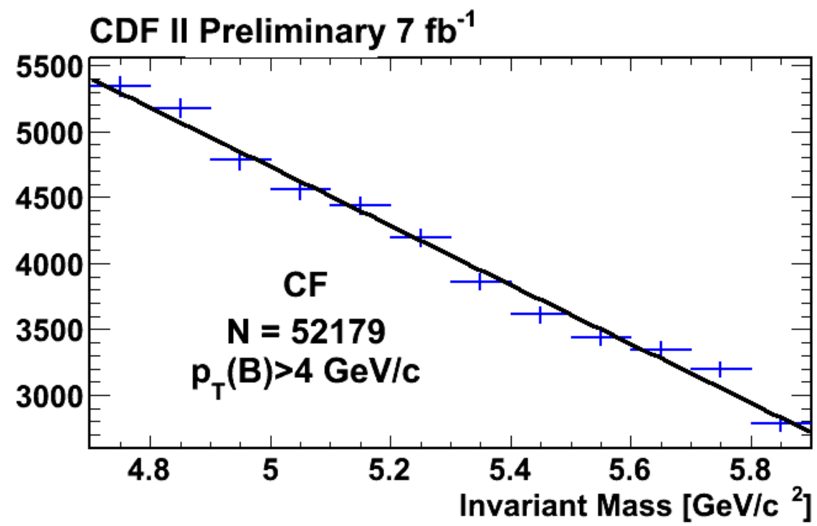


$B^0 \rightarrow \mu^+\mu^-$



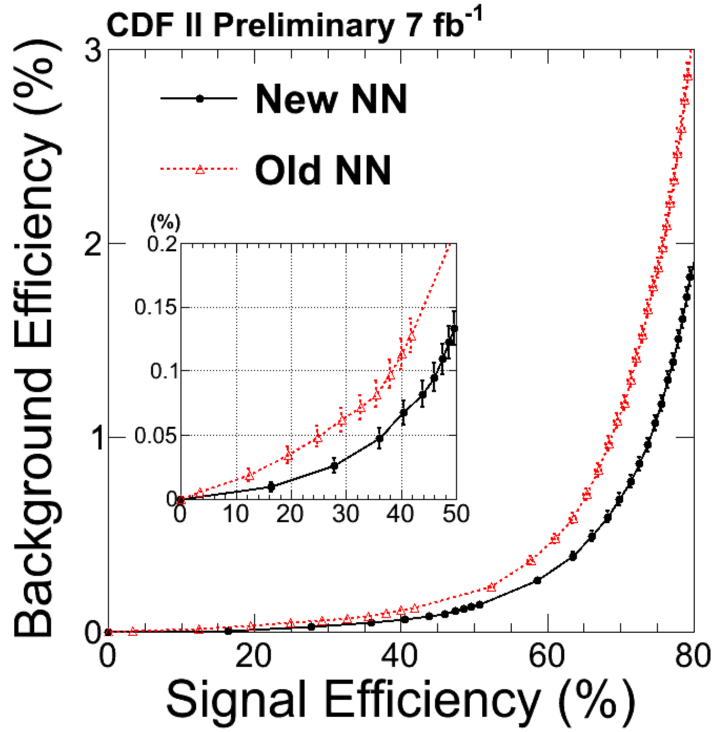
central/forward events

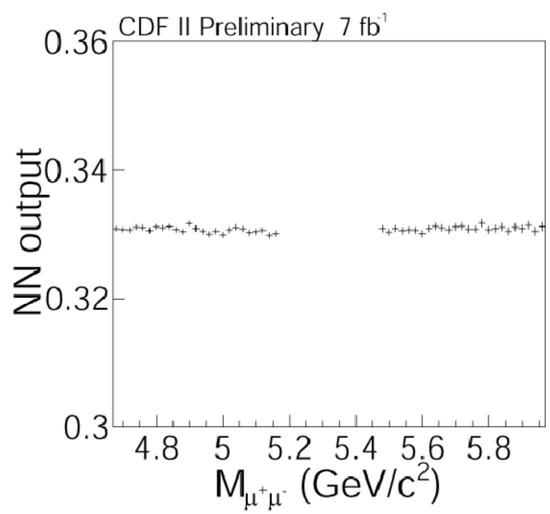




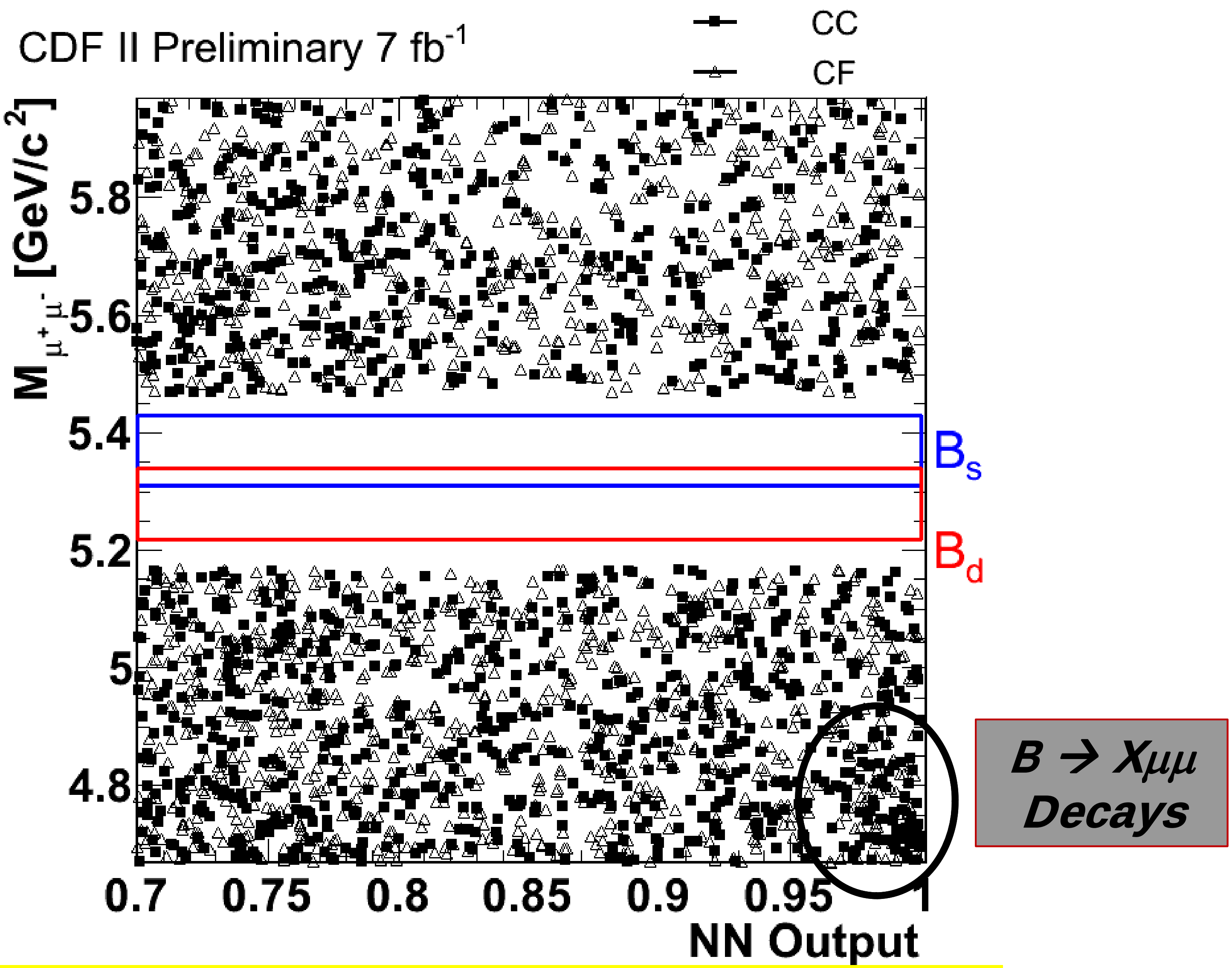
Neural Net Selection

- **Improved with respect to previous analysis**
- **Blind optimization for $B_s \to \mu\mu$ expected limit**
- **Carefully check for bias**
  - **mass dependence**
  - **overtraining**
    - **check with sideband data**





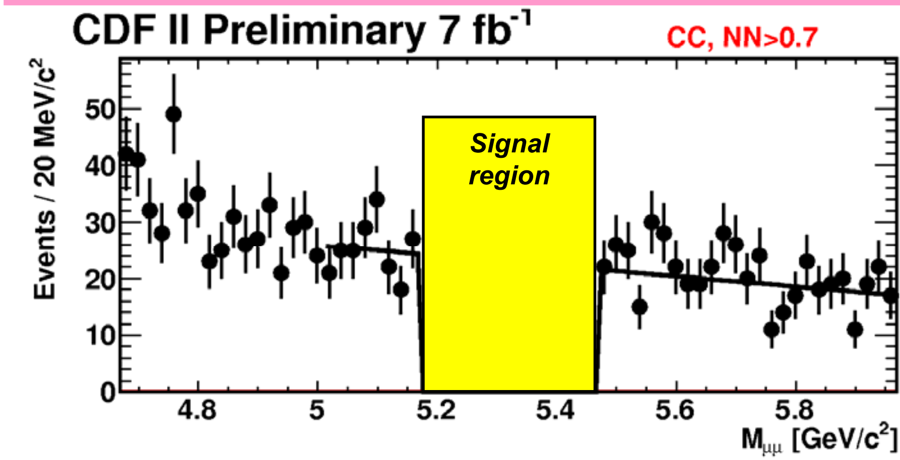
NN vs. Mass, Still Blinded

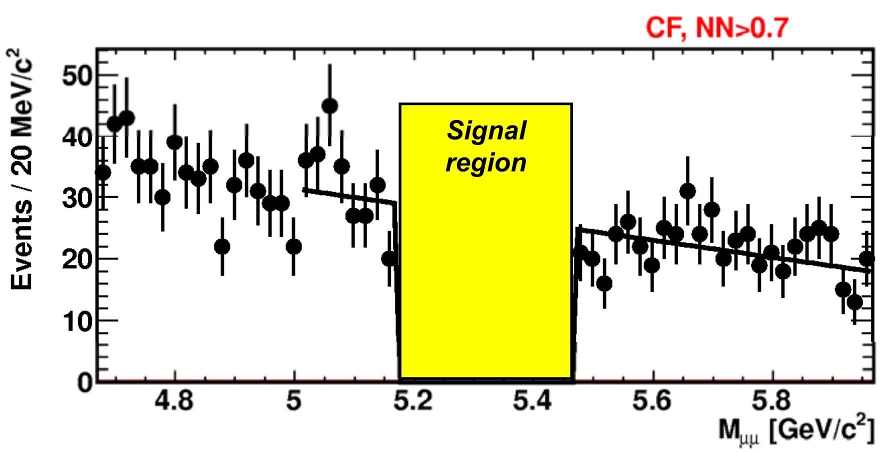


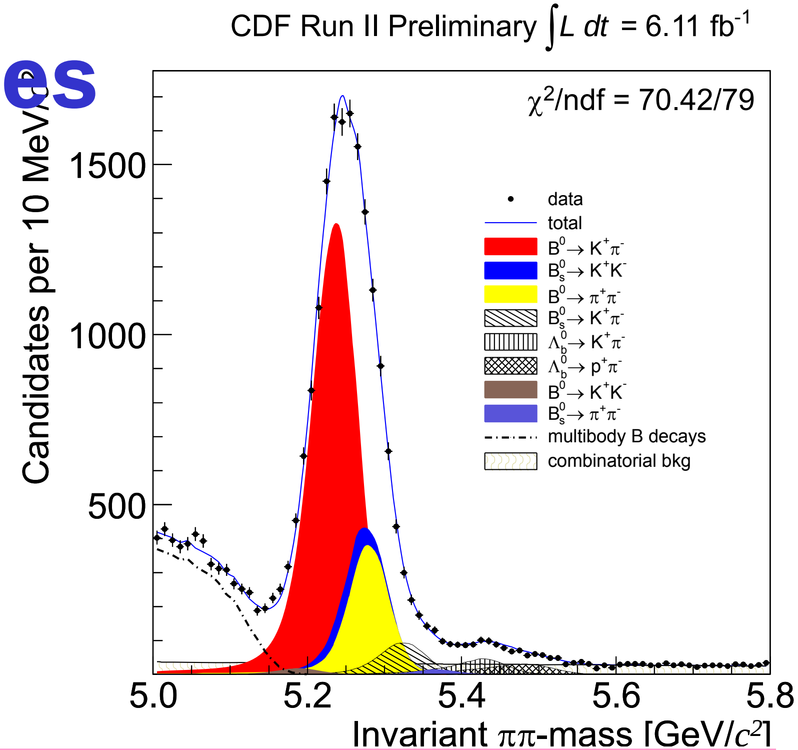
→ → higher NN = more signal-like→ →

Background Estimates

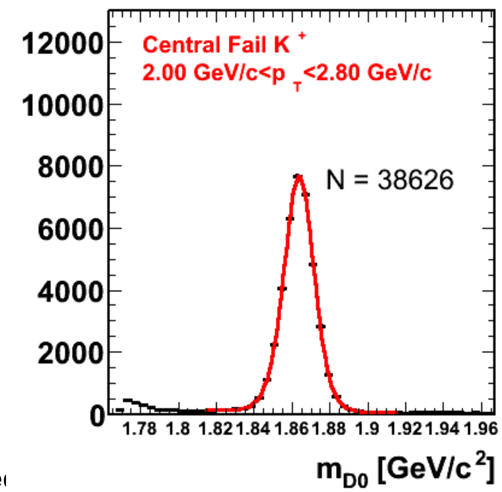
Combinatoric background:
- ***Estimated from sidebands***
- ***Exclude partial reconstructions***







Peaking background:
- ***From B → hh′***
- ***Muon misID rate from D* tagged D^0***



Checks on Control Regions

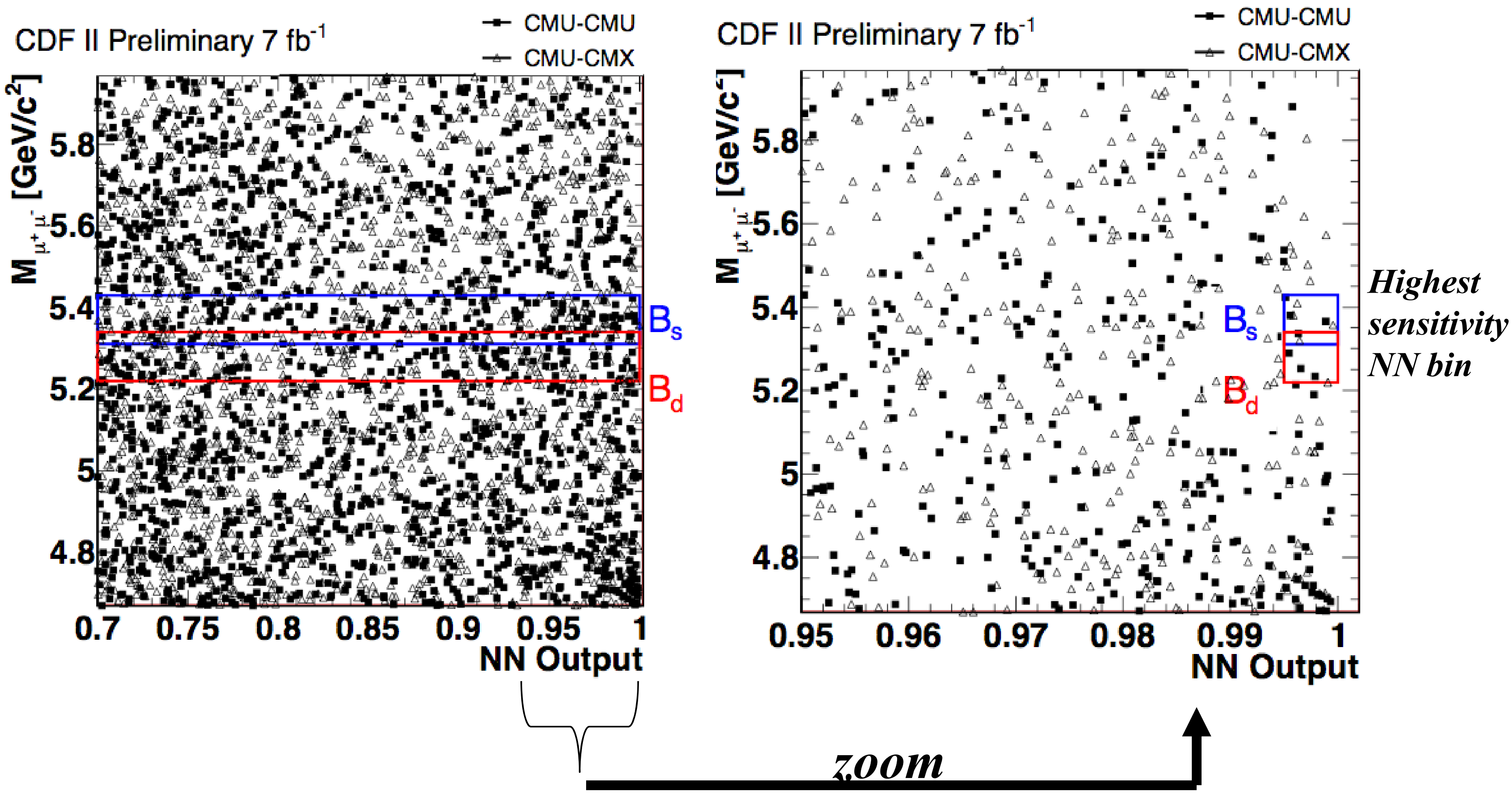
	Control Sample	Prediction	Nobs	Prob(N>=Nobs)
Negative lifetime	OS-	2140.0±53.9	1999	98%
Same-sign muons	SS+	19.7±3.4	25	19%
	SS-	46.8±5.3	53	25%
Fake muon enhanced	FM+	567.8±25.4	593	24%
	Sum	2774.3±59.9	2670	91%

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.

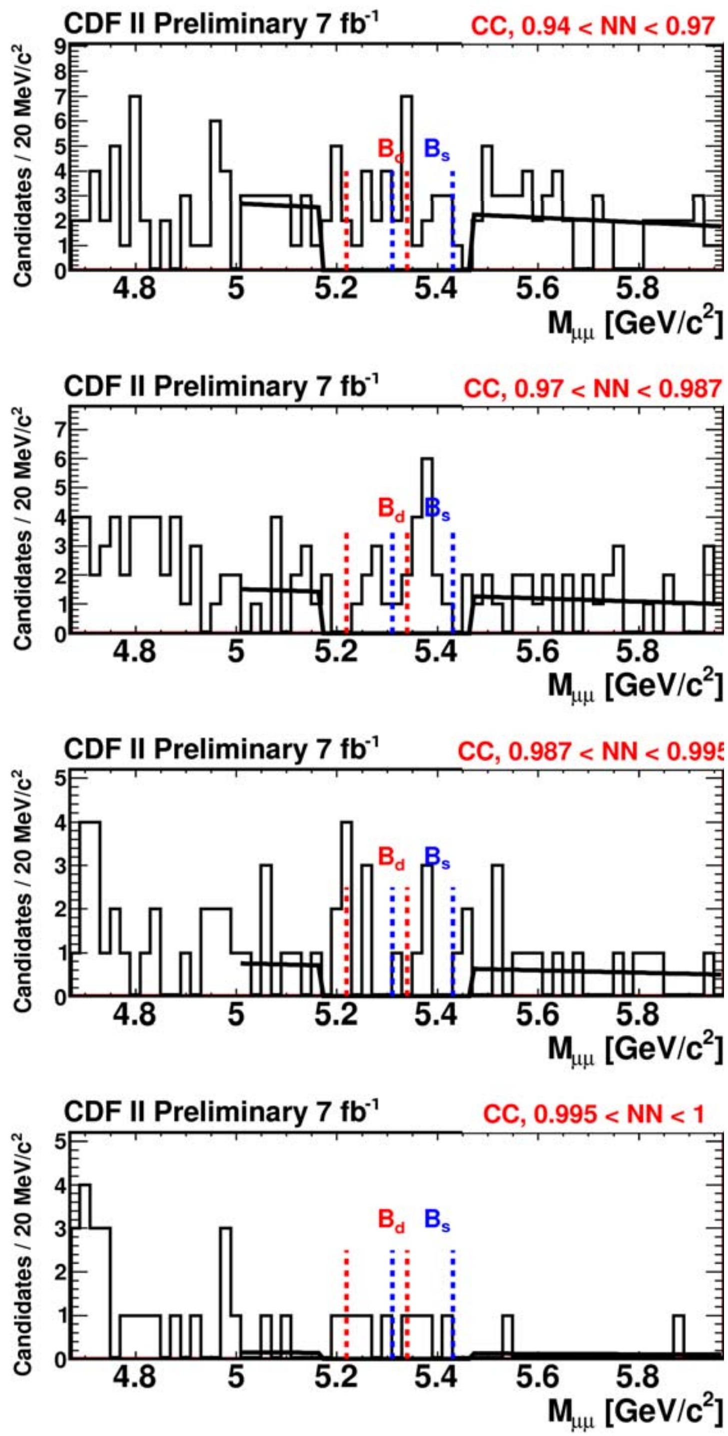
$B \to \mu^+\mu^-$ search: opening the box

Expected Sensitivity:

$\mathcal{B}(B_s \to \mu^+\mu^-) < 1.5\times10^{-8}$ @95%CL $\quad \mathcal{B}(B^0 \to \mu^+\mu^-) < 4.6\times10^{-9}$ @95%CL

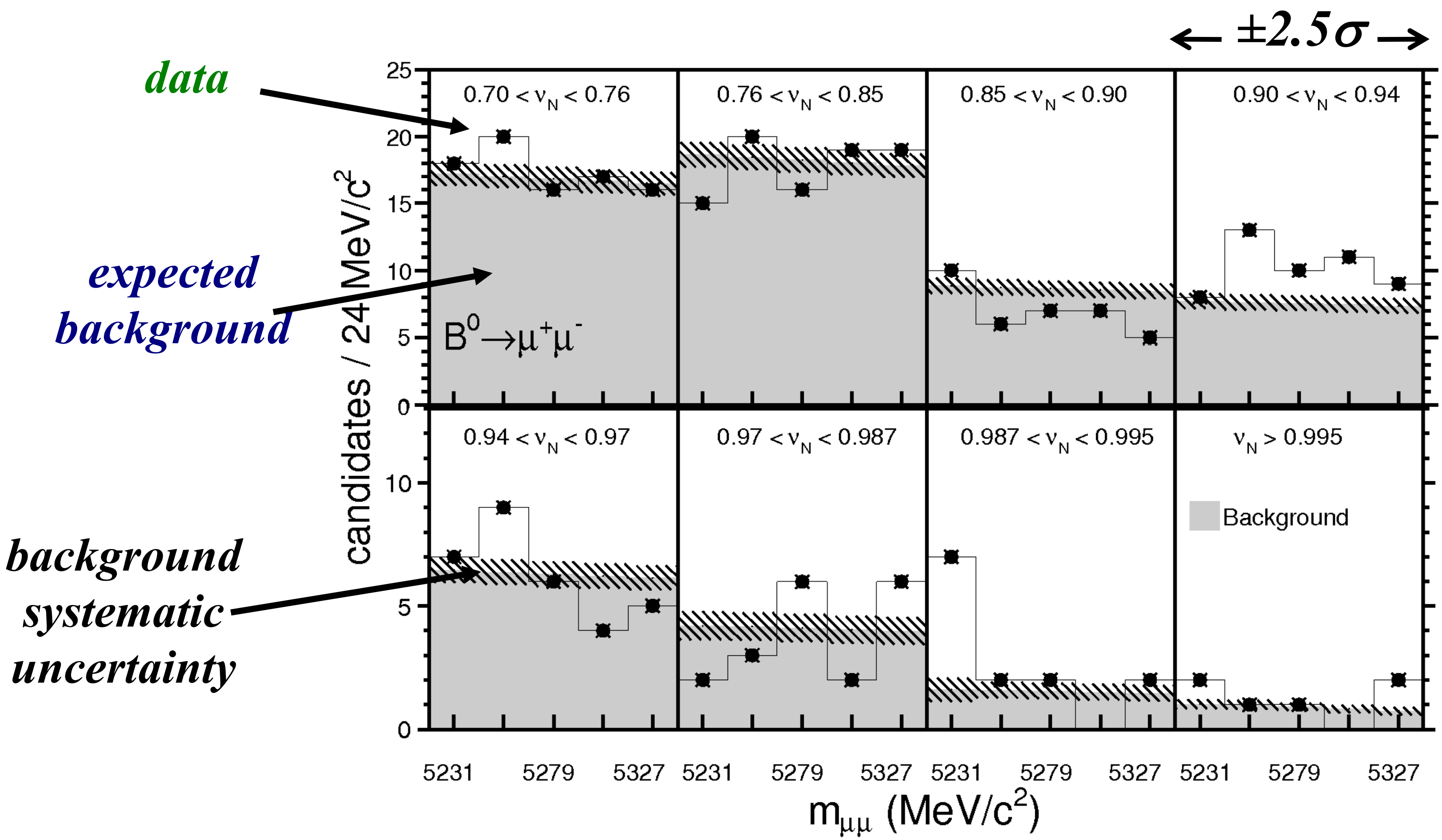


↓ ↓ **Increasing NN** ↓ ↓



Showing CC only, highest purity NN bins

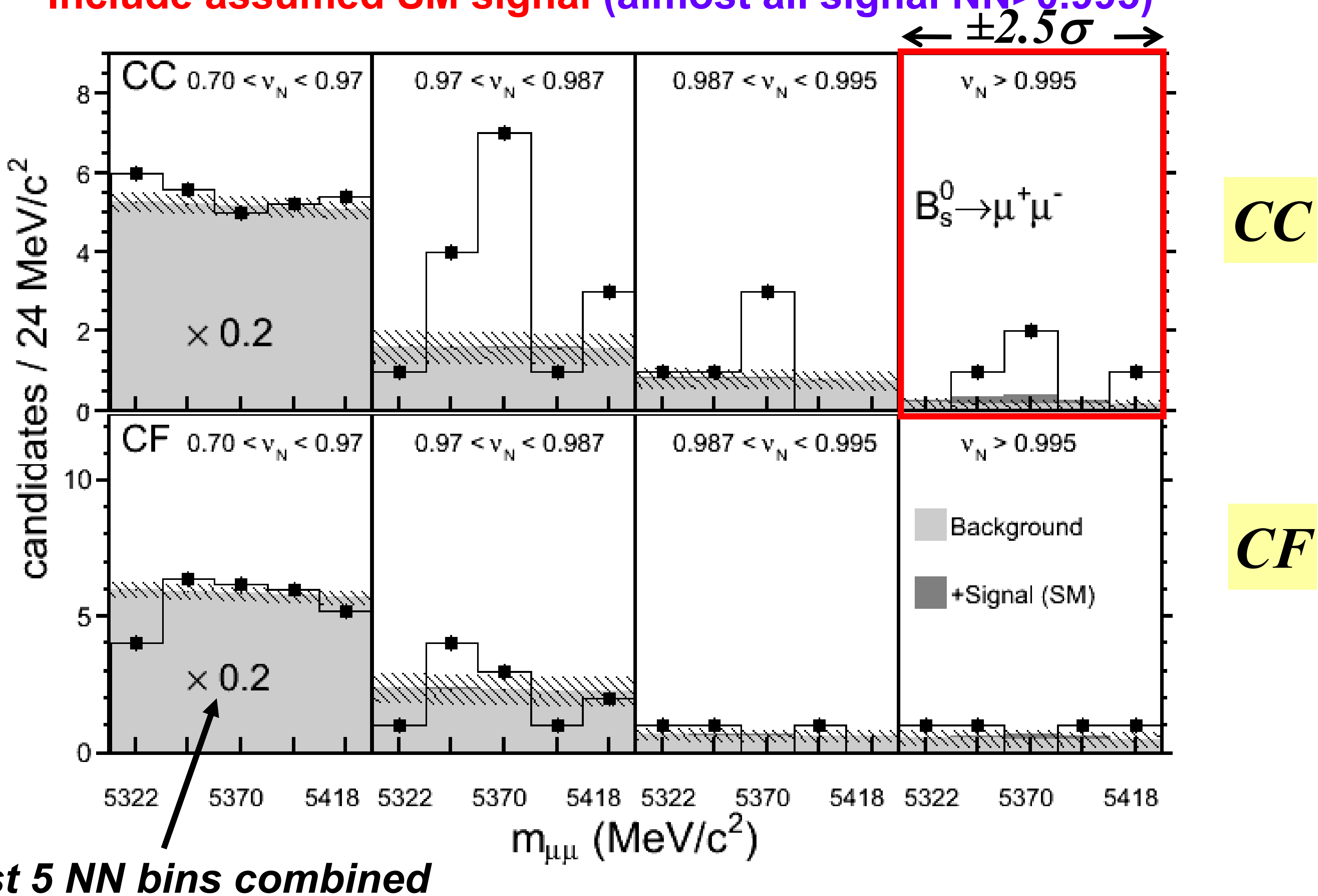
B^0 Search Window



$$\text{observed} \quad \mathcal{B}(B^0 \to \mu^+\mu^-) < 6.0 \times 10^{-9} \; 95\% \, CL$$

$B_s \rightarrow \mu^+\mu^-$ Signal Window

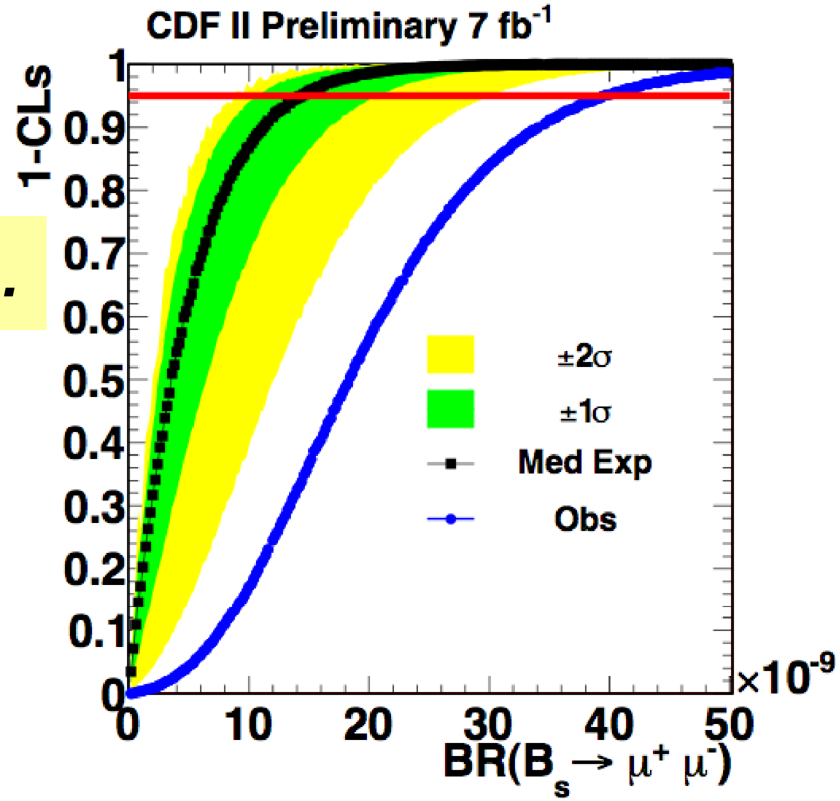
- Include assumed SM signal (almost all signal NN>0.995)



$B_s \rightarrow \mu^+\mu^-$: Observed Limit

Using the CLs method, we observe

$\mathcal{B}(B_s \rightarrow \mu^+\mu^-) < 4.0\times10^{-8}$ @ 95% C.L.



Compare to the expected limit

$\mathcal{B}(B_s \rightarrow \mu^+\mu^-) < 1.5\times10^{-8}$

well outside the 2σ consistency band

Need statistical interpretation of the observed excess:

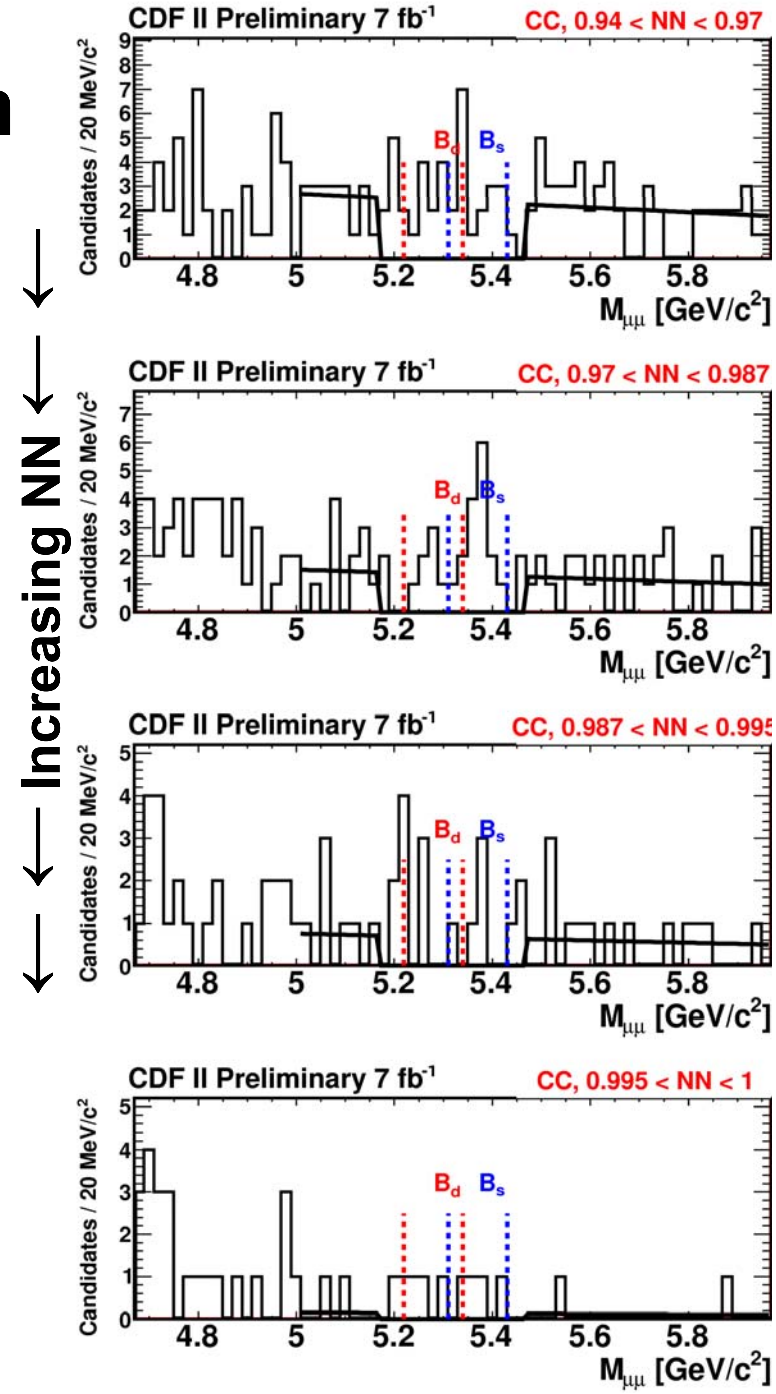
- ***what is the level of inconsistency with the background?***
- ***what does a fit to the data in the B_s search window yield?***

$B_s \rightarrow \mu^+\mu^-$ Signal Region

- **Use simulated experiments to evaluate consistency**
  - **include systematics.**

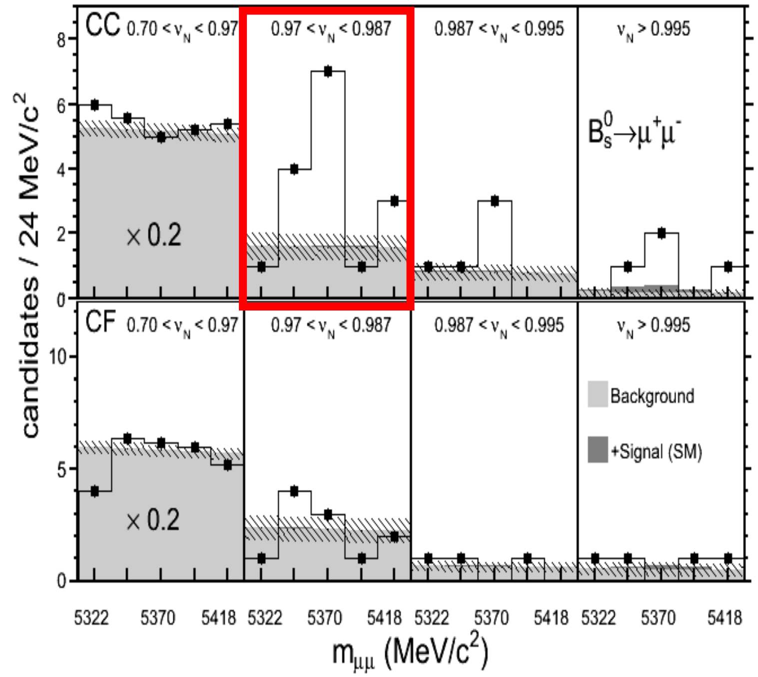
p-value=0.27%
for background only hypothesis

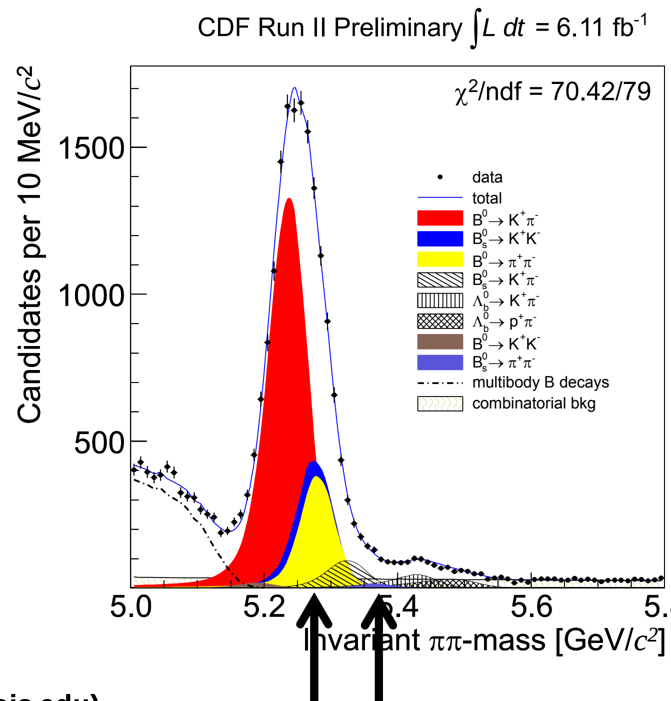
p-value=1.9% for SM signal plus background hypothesis



FAQ: Excess in 3rd Highest NN Bin

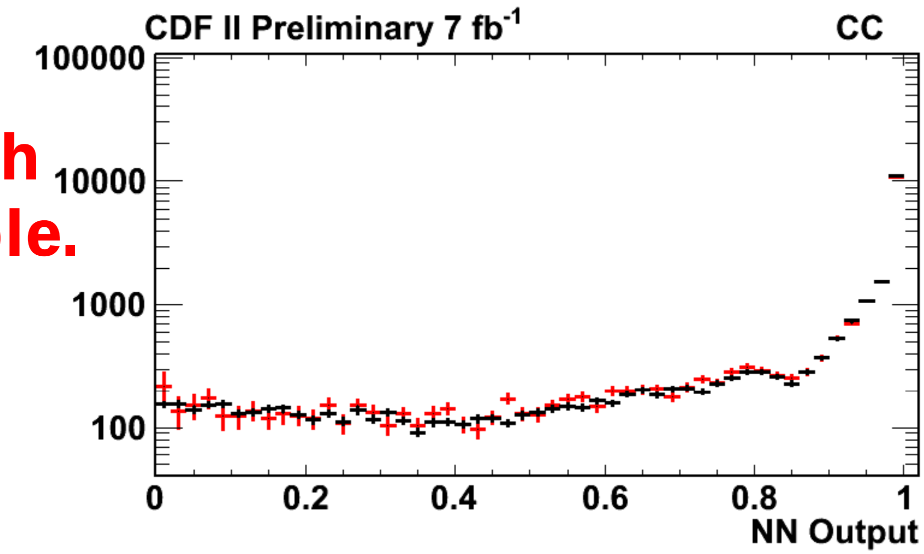
- **Reminder: B^0 analysis uses same sidebands, fits, mis-ID probabilities.**
- **Used "blinded" choices to avoid bias**
- **We also considered:**
  - **peaking bkgd. ($B \rightarrow hh'$)?**
    - Predict 0.014 $B \rightarrow hh'$ in 3rd bin.
    - Should be x10 more in B^0
    - would peak at low edge in B_s search
  - **NN biasing background?**
    - Check NN using $B^+ \rightarrow J/\psi K^+$
  - **additional combin. bkgd.?**
    - Comb. background doesn't peak
    - Prediction agrees with data in B^0 region and in lower NN bins.
  - **statistical fluctuation?**
    - Possible for 1 out of 80 bins
- **Conclude: likely a fluctuation.**
- **Little signal expected in this bin. Results not different using just 2 highest bins**

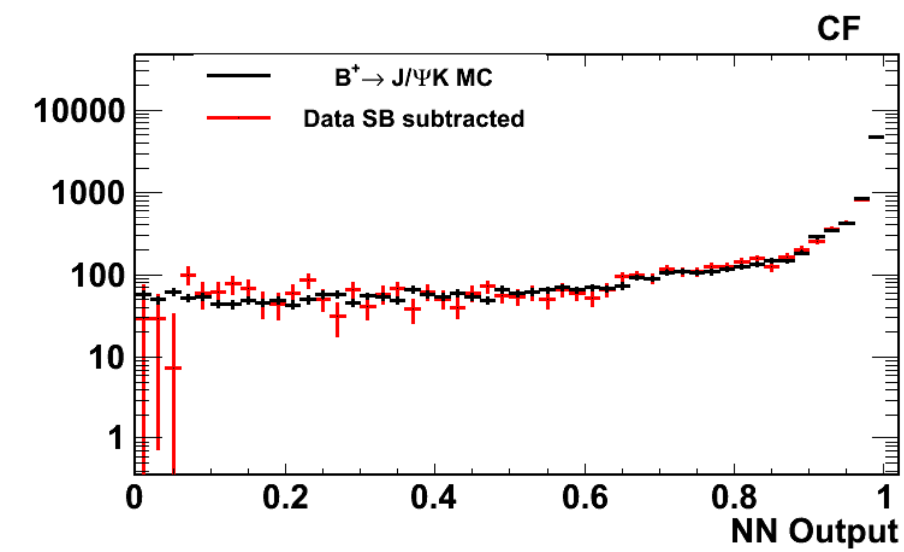




Check NN modeling

- Check NN modeling in high statistics $B^+ \rightarrow J/\psi K^+$ sample.
- Good modeling, even in highest NN bins.





Fit to the data in the B_s search window

Using the log-likelihood fit described before, we set the ***first two-sided limit for the $B_s \to \mu^+\mu^-$ branching fraction:***

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

Our central value is

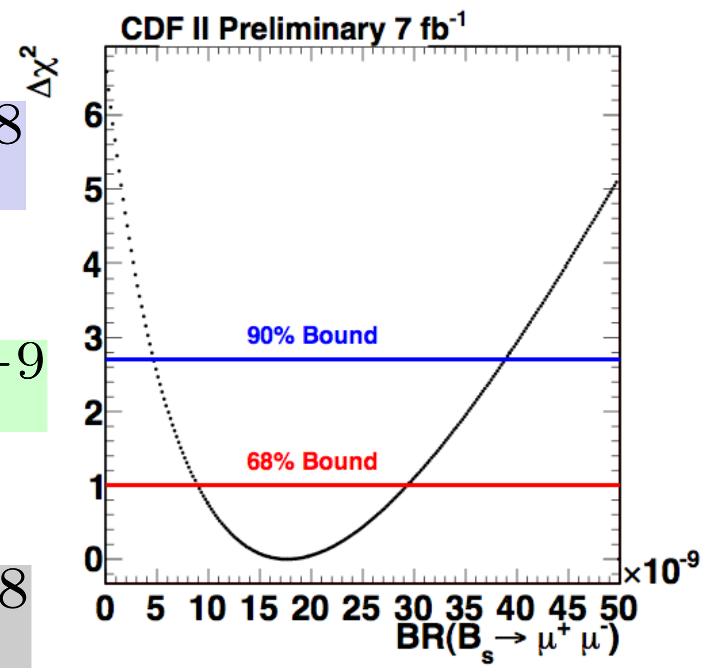
$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = 1.8^{+1.1}_{-0.9} \times 10^{-8}$$

Compare to SM expectation of:

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (3.2 \pm 0.2) \times 10^{-9}$$

Using just 2 highest NN bins

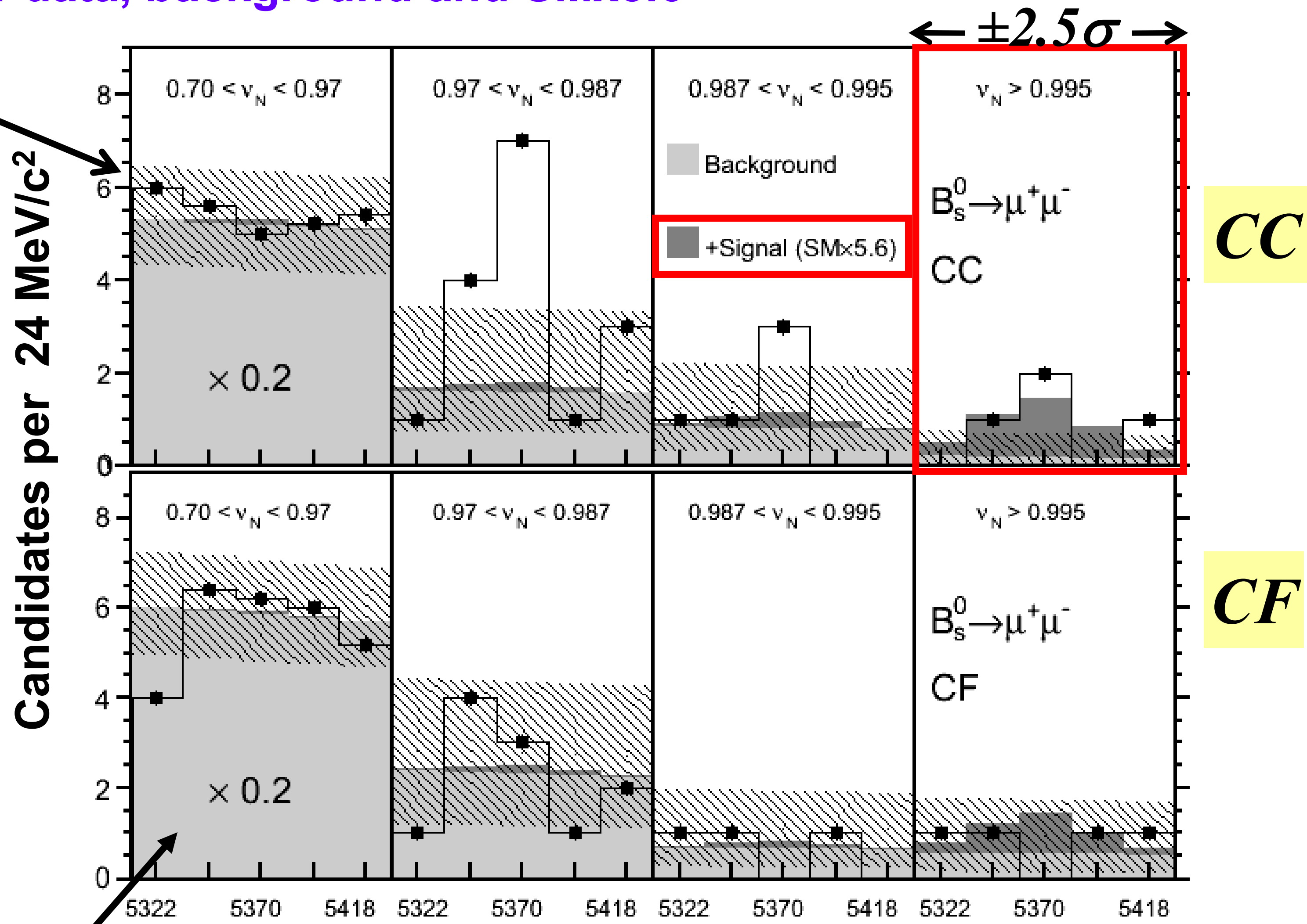
$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = 1.4^{+1.0}_{-0.8} \times 10^{-8}$$



$B_s \rightarrow \mu^+\mu^-$ Signal Window

Show data, background and SMx5.6

Plot also incorporates Poisson errors on data into uncertainty on background.



First 5 NN bins combined

Compare 90% CLs

- **CDF** $0.46\times10^{-8} < \mathcal{B}(B_s\rightarrow\mu^+\mu^-) < 3.9\times10^{-8}$
- **LHCb** $\mathcal{B}(B_s\rightarrow\mu^+\mu^-) < 1.2\times10^{-8}$
- **CMS** $\mathcal{B}(B_s\rightarrow\mu^+\mu^-) < 1.6\times10^{-8}$

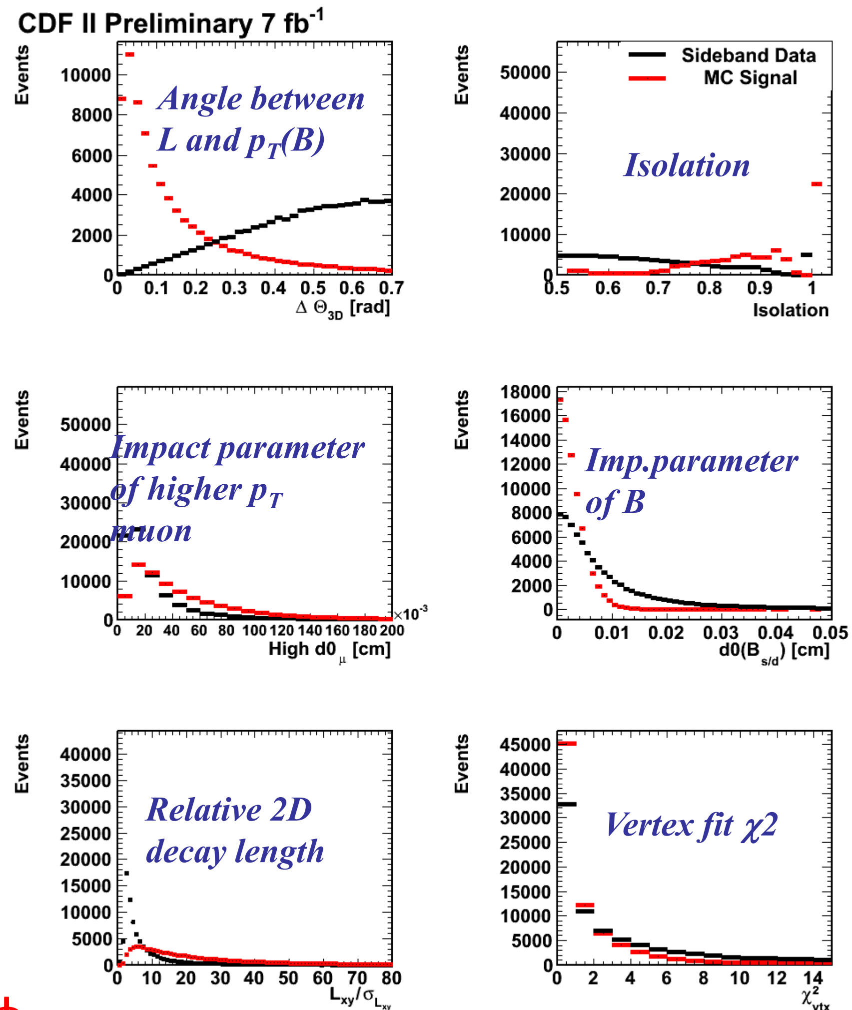
Summary

- New analysis with 7 fb^{-1} and significantly improved sensitivity
- B^0 : consistent with background
- B_s : Observe excess in search window
  - Extract two-sided confidence region on $\mathcal{B}(B_s \rightarrow \mu^+ \mu^-)$
  - Submitted to PRL, arXiv 1107.2304

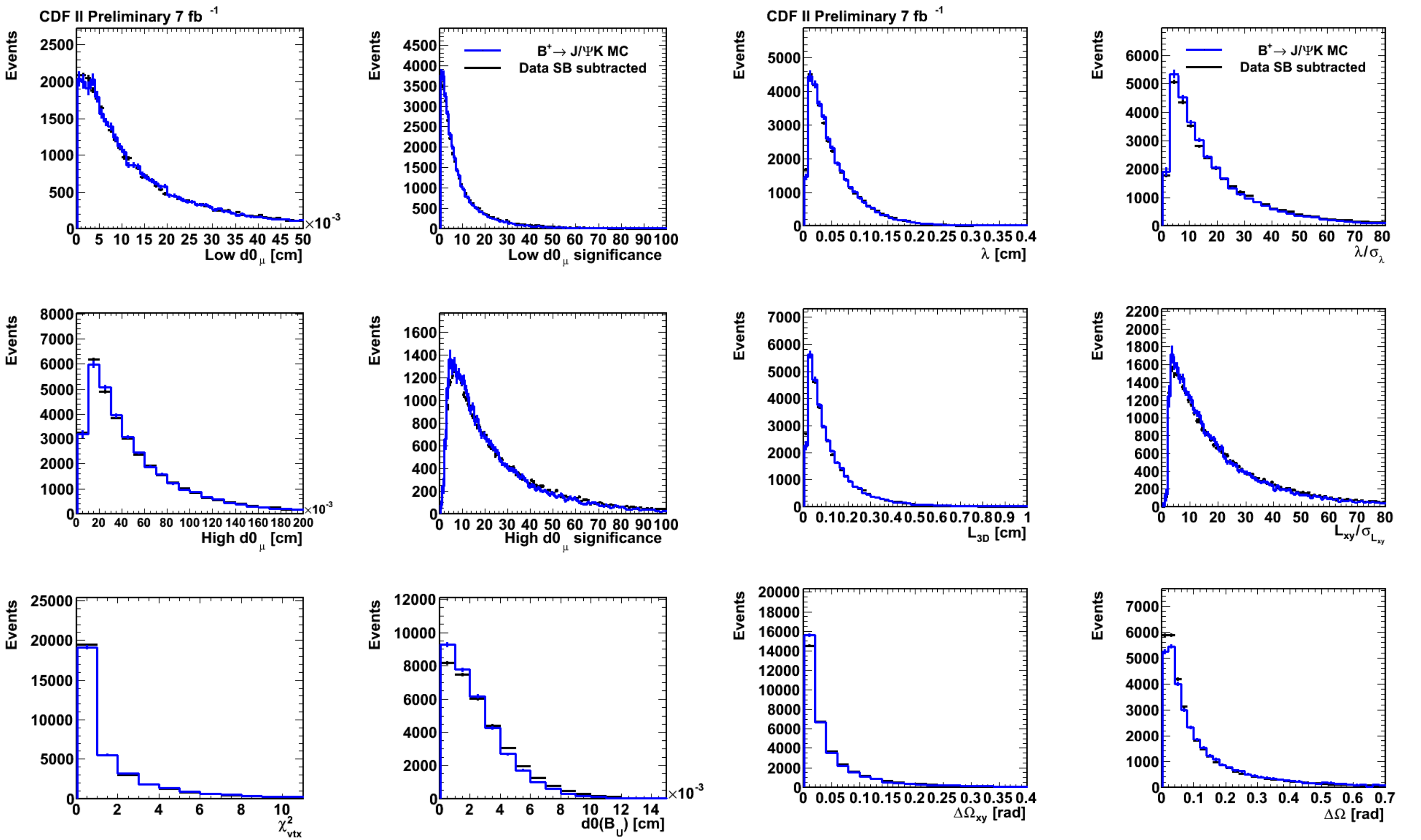
NN Variables

Rank	Variable
1	$\Delta\alpha_{3d}$
2	Isolation
3	Larger $\lvert d_0(\mu)\rvert$
4	$\lvert d_0(B_s^0)\rvert$
5	$L_{2d}/\sigma_{L_{2d}}$
6	χ^2_{vtx}
7	L_{3d}
8	Lower $p_T(\mu)$
9	Significance of smaller $\lvert d_0(\mu)\rvert$
10	$\lambda_{3d}/\sigma_{\lambda 3d}$
11	λ_{3d}
12	Smaller $\lvert d_0(\mu)\rvert$
13	$\Delta\alpha_{2d}$
14	Significance of larger $\lvert d_0(\mu)\rvert$

- p_T reweighted to $B^+ \rightarrow J/\psi\ K$
- Isolation reweighted to $B_s \rightarrow J/\psi\ \phi$



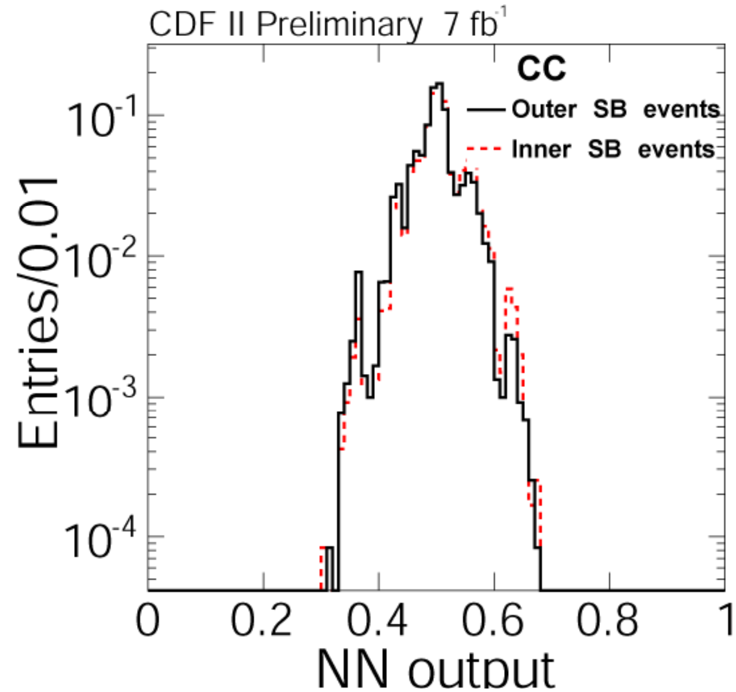
NN input variables for $B \to J/\psi K^+$ data vs. MC

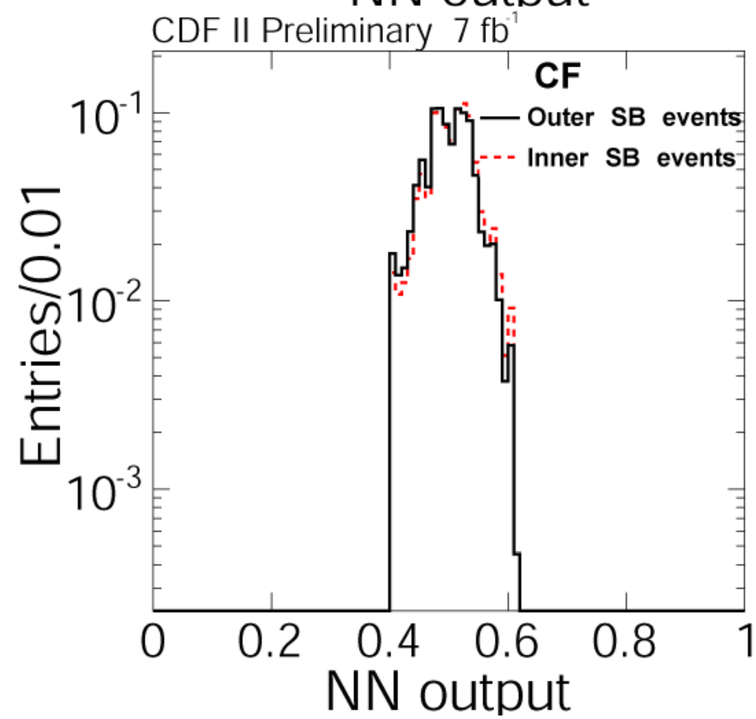


NN bias check

- **Check of possible NN mass bias.**
  - **Cut the sideband in half.**
  - **Train NN on one half as signal, the other as background.**
  - **See if NN can use input variables differentiate between samples.**

- **Answer: No.**





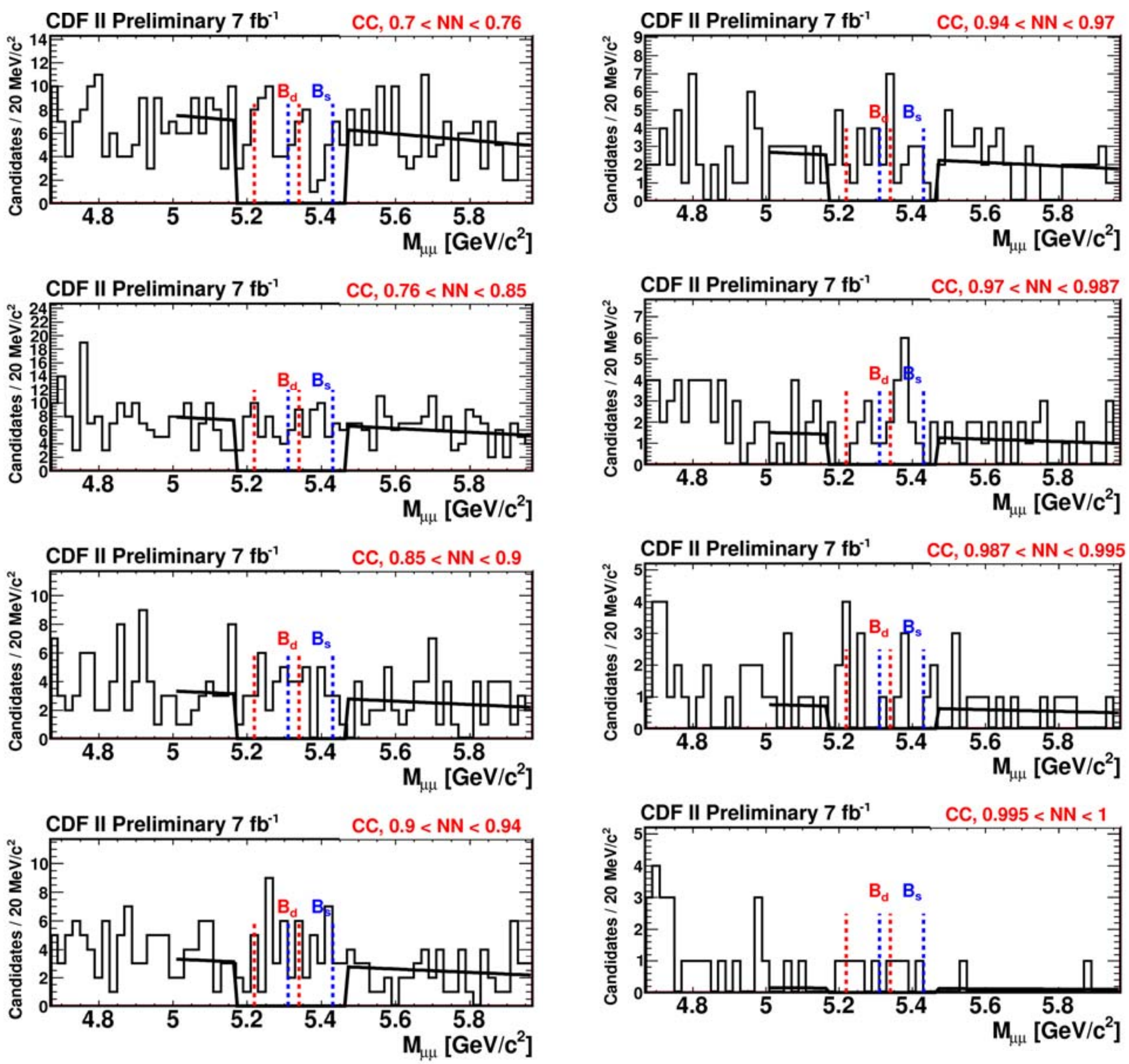
Signal Efficiency

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = \frac{N_{B_s^0}}{N_{B^+}} \frac{\alpha_{B^+}}{\alpha_{B_s^0}} \frac{\epsilon_{B^+}^{trig}}{\epsilon_{B_s^0}^{trig}} \frac{\epsilon_{B^+}^{reco}}{\epsilon_{B_s^0}^{reco}} \frac{1}{\epsilon_{B_s^0}^{NN}} \frac{f_u}{f_s} \mathcal{B}(B^+)$$

	CC		CF	
$(\alpha_{B+} / \alpha_{B_s})$	0.307 ± 0.018	$(\pm 6\%)$	0.197 ± 0.014	$(\pm 7\%)$
$(\epsilon_{B+}^{trig} / \epsilon_{B_s}^{trig})$	0.99935 ± 0.00012	$(< 1\%)$	0.97974 ± 0.00016	$(< 1\%)$
$(\epsilon_{B+}^{reco} / \epsilon_{B_s}^{reco})$	0.85 ± 0.06	$(\pm 8\%)$	0.84 ± 0.06	$(\pm 9\%)$
$\epsilon_{B_s}^{NN}(NN > 0.70)$	0.915 ± 0.042	$(\pm 4\%)$	0.864 ± 0.040	$(\pm 4\%)$
$\epsilon_{B_s}^{NN}(NN > 0.995)$	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	3.59 ± 0.37	$(\pm 13\%)$	3.59 ± 0.37	$(\pm 13\%)$
$BR(B^+ \to J/\psi K^+ \to \mu^+\mu^- K^+)$	$(6.01 \pm 0.21) \times 10^{-5}$	$(\pm 4\%)$	$(6.01 \pm 0.21) \times 10^{-5}$	$(\pm 4\%)$
SES (All bins)	$(2.9 \pm 0.5) \times 10^{-9}$	$(\pm 18\%)$	$(4.0 \pm 0.7) \times 10^{-9}$	$(\pm 18\%)$

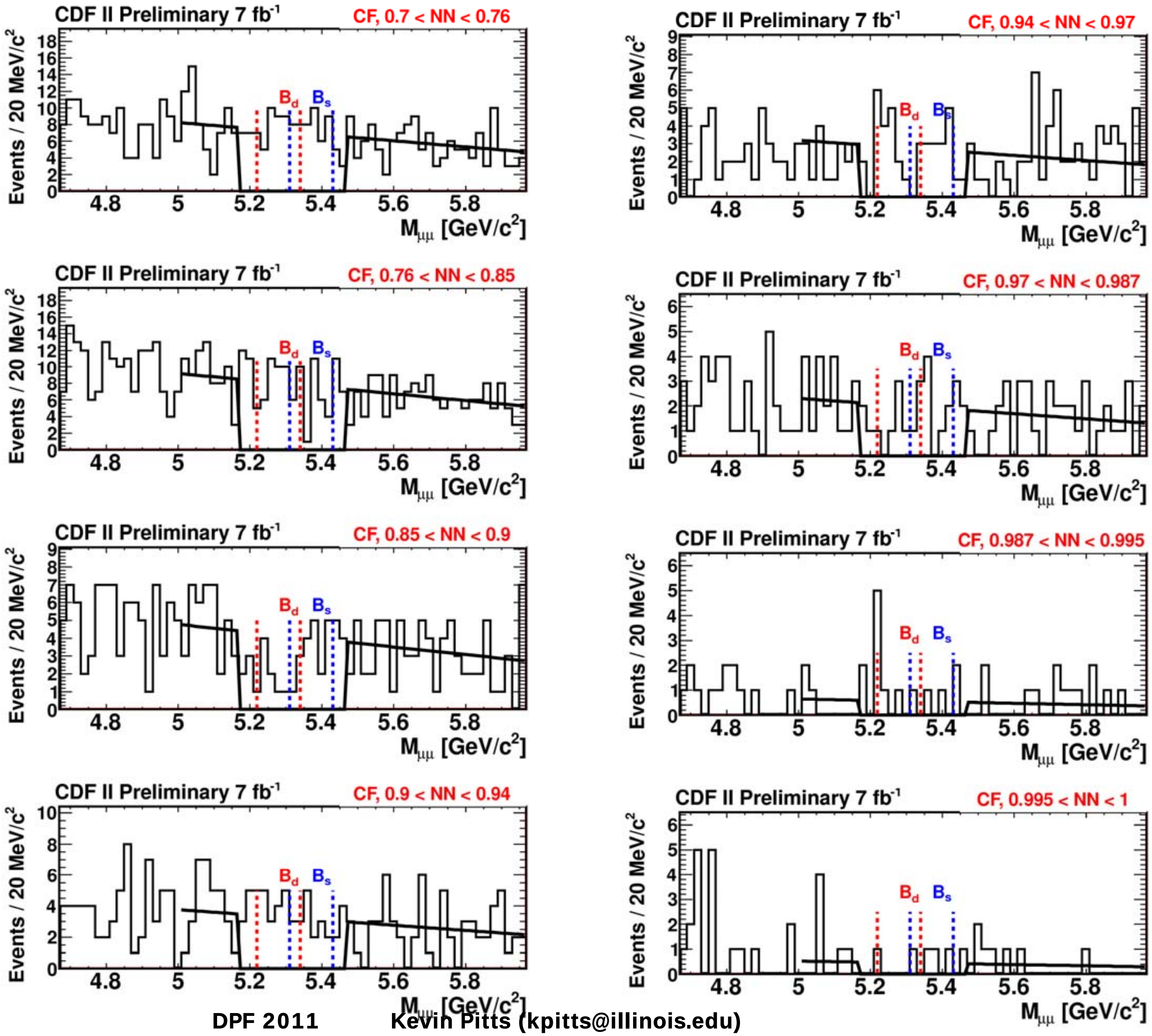
Single event sensitivity: expect to see 1.9 SM $B_s \to \mu^+\mu^-$ events

Unblinded Data



8 NN bins of CC sample

Unblinded CF Mass Plots



Full table of bkgd checks in control samples

sample	NN cut	CC pred	CC obsv	CC 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

Good agreement in most sensitive NN bins

✓ ***now have sufficient confidence in background estimation***

Control sample FM+ is rich in B->hh background.
Good agreement in highest NN bin
shows that we can accurately predict this background

**if zero events are observed, "Prob(N>=Nobs)" is the Poisson probability for observing exactly 0*

SM and Background Expectations for B_s

CC

NN Bin	ϵ_{NN}	B→hh Bkg	Total Bkg	Exp SM Signal
$0.700 < NN < 0.970$	20%	0.03	129.24±6.50	0.26±0.05
$0.970 < NN < 0.987$	8%	< 0.01	7.91±1.27	0.11±0.02
$0.987 < NN < 0.995$	12%	0.02	3.95±0.89	0.16±0.03
$0.995 < NN < 1.000$	46%	0.08	0.79±0.40	0.59±0.11

CF

NN Bin	ϵ_{NN}	B→hh Bkg	Total Bkg	Exp SM Signal
$0.700 < NN < 0.970$	21%	0.01	146.29±7.00	0.19±0.04
$0.970 < NN < 0.987$	10%	0.01	11.57±1.57	0.09±0.02
$0.987 < NN < 0.995$	8%	0.01	3.25±0.82	0.08±0.01
$0.995 < NN < 1.000$	46%	0.03	2.64±0.74	0.43±0.08

Consistency with Background and SM

- **p-value for background-only hypothesis: 0.27%**
- **p-value for background+SM hypothesis: 1.9%**