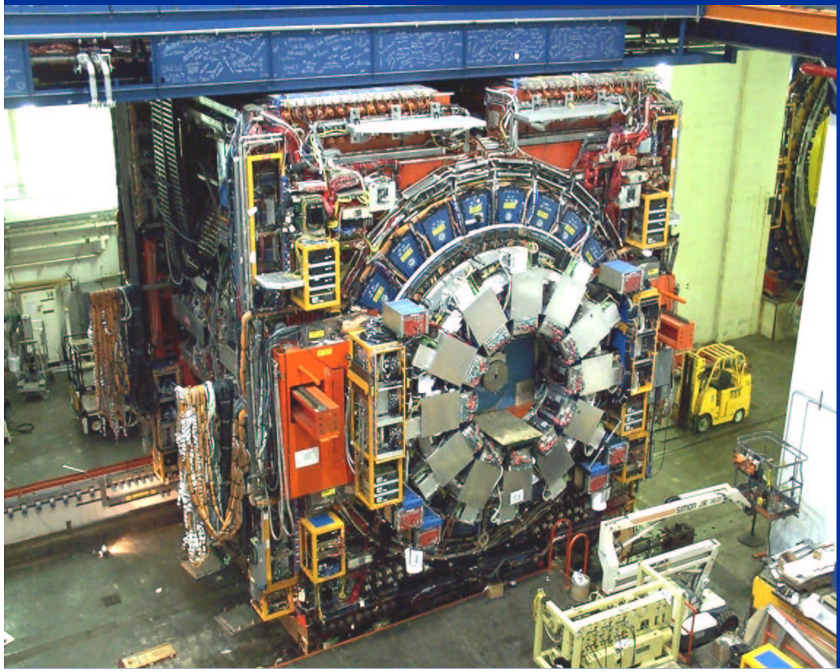
Search for $B_s \rightarrow \mu^+\mu^-$ and $B_d \rightarrow \mu^+\mu^-$ Decays at CDF

Cheng-Ju S. Lin
(Fermilab/CDF)

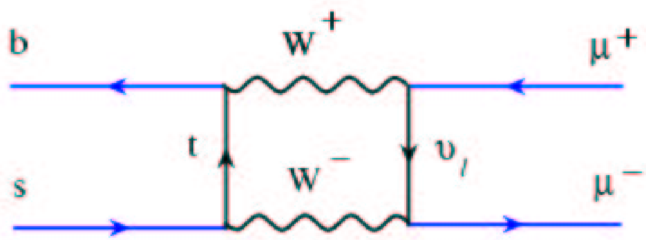


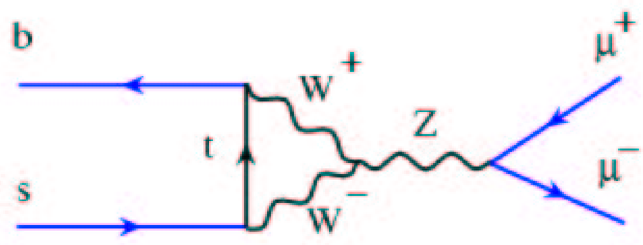
"Wine & Cheese" Seminar

June 17, 2005

Introduction

- In the Standard Model, the FCNC decay of B $\rightarrow \mu^+\mu^-$ is heavily suppressed





SM prediction → $BR(B_s \rightarrow \mu^+\mu^-) = (3.5 \pm 0.9) \times 10^{-9}$

(Buchalla & Buras, Misiak & Urban)

- $B_d \rightarrow \mu\mu$ is further suppressed by CKM couping $(v_{td}/v_{ts})^2$

- SM prediction is below the sensitivity of current experiments (CDF+D0): SM → Expect to see 0 events at the Tevatron

Any signal would indicate new physics!!

BEYOND STANDARD MODEL

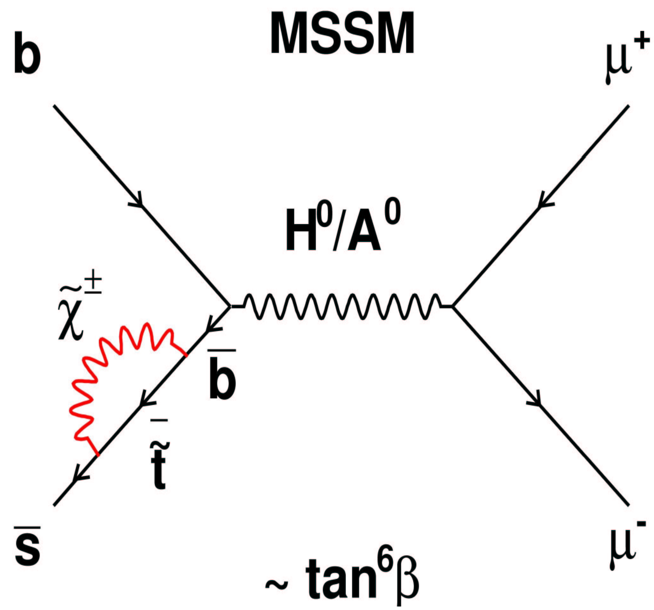
- In many SUSY models, the BR could be enhanced by many orders of magnitude:

For example:

- MSSM: Br($B \rightarrow \mu\mu$) is proportional to $\tan^6\beta$

- GUT SO(10) models prefer high $\tan\beta$ (Yukawa coupling unification)

– BR could be as large as 10-1000 times the SM prediction



Would be observable at the Tevatron

BEYOND STANDARD MODEL

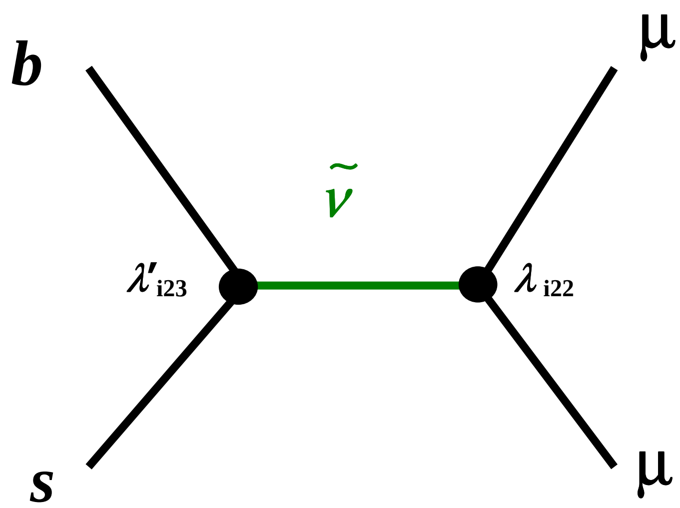
- Another example:
  R-Parity violating SUSY

  - Tree level diagram is allowed in R-parity violating (RPV) SUSY models.

  - Possible to observe decay even for low value of tanβ

  - Enhancement depends strongly on coupling constants (λ , λ')

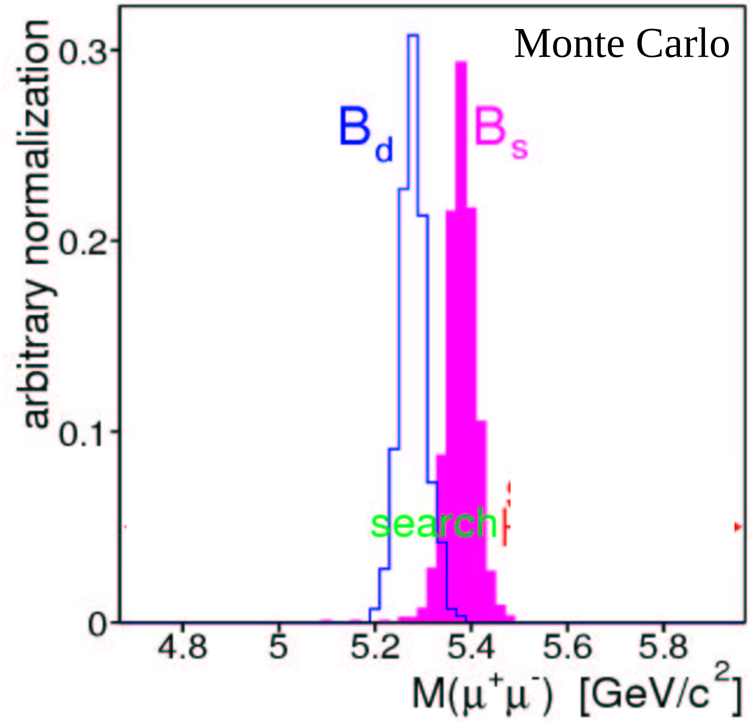
R-parity violating SUSY



Would also be observable at the Tevatron

PROBE OF NEW PHYSICS

- New physics may enhance B_s and $B_d \rightarrow \mu\mu$ differently
- Minimal-flavor-violation (MFV) assumption in SUSY yields SM relations between B_s and $B_d \rightarrow \mu\mu$ decays
- Can observe both B_s and B_d: unique to Tevatron
- CDF has the mass resolution to distinguish two decays, $\sigma(M_{\mu\mu})$~23MeV : unique to CDF



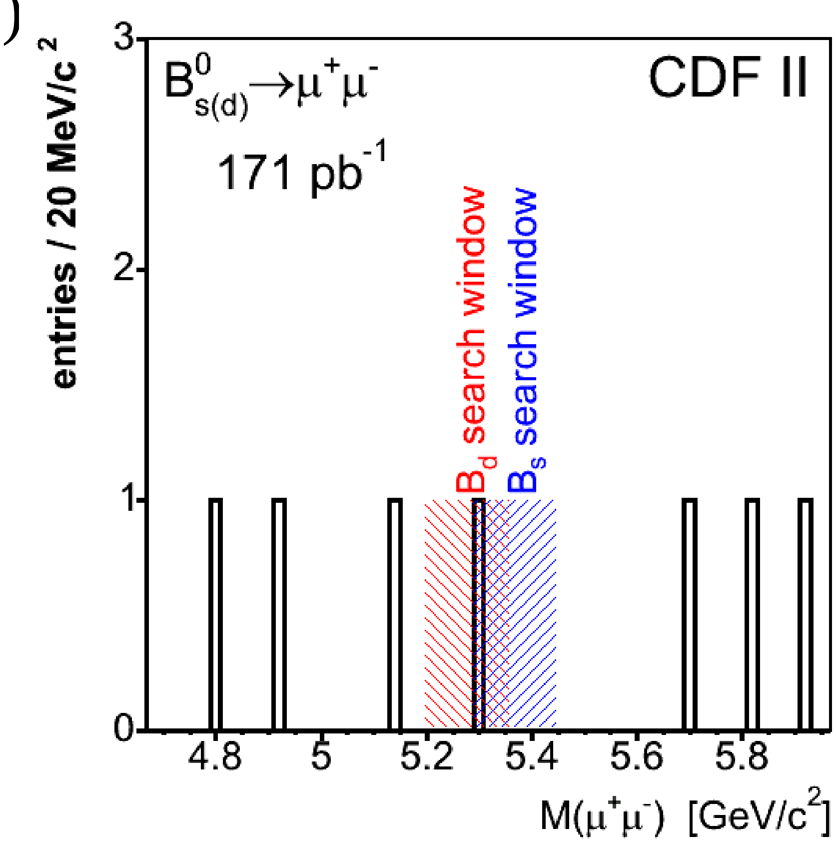
$M(B_s)$-$M(B_d)$~90MeV

- Either observation or null search, results will provide important clues about possible scenarios of new physics beyond SM

A Brief Recap of Analysis History

- CDF PRL 93, 032001, 2004 ($171 pb^{-1}$)
  Expected 1.1 background
  Observed 1 event

  2004 CDF limit BR($B_s \rightarrow \mu^+\mu^-$)
  $< 5.8 \times 10^{-7}$ @90% CL
  $< 7.5 \times 10^{-7}$ @95% CL



- Since then:
  - Using x2 data sample
  - Using extended muon coverage (increased acceptance 50%)
  - Lowered p_T threshold on B candidate
  - Improve signal bkg separation using a likelihood discriminant
  - Significantly improved the sensitivity of the analysis

CDF II Detector

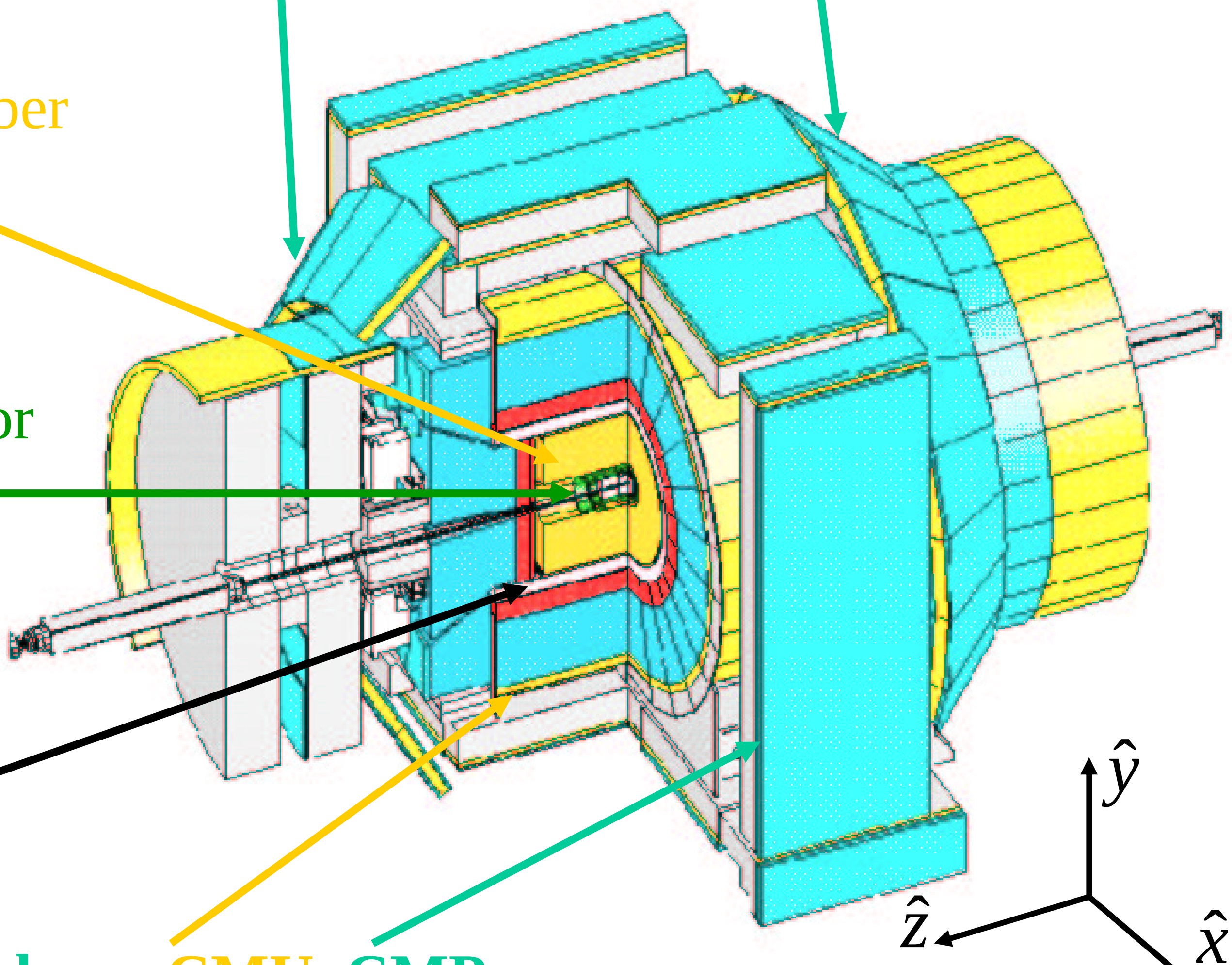
Important components relevant to the analysis highlighted:

Central Muon Extension: CMX
($0.6 < |\eta| < 1.0$)

Central Drift Chamber (COT)

Silicon Vertex Detector (L00, SVXII, ISL)

Superconducting Solenoid (1.4T)

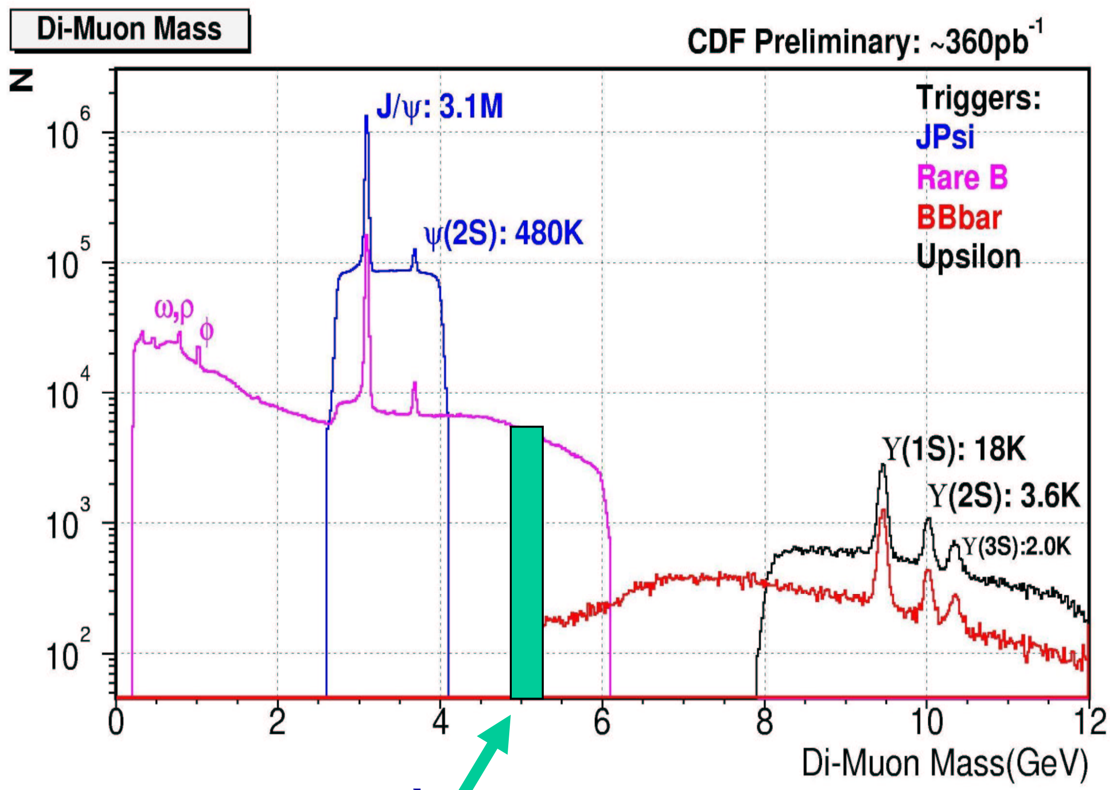


Central Muon Chambers: CMU, CMP
($|\eta| < 0.6$)

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

Data Sample

- Using 364pb^{-1} of data (Feb 02 – Aug 04) from di-muon triggers:
  - CMU(P) + CMU (central-central)
  - CMU(P) + CMX (central-extended)
- Central-central and central-extended channels treated independently in this analysis (background and efficiencies are different)



Search region

Rare B di-muon triggers requires additional cuts to reduce background relative to inclusive **J/ψ di-muon** trigger

Ingredients of the Analysis

Overall picture:

- Reconstructing di-muon events in the B mass window
- Measure the branching ratio or set a limit

  Normalized to B→J/ψ K decays

$$BR(B_s \to \mu^+\mu^-) = \frac{N_{Bs}}{\alpha_{Bs} \cdot \varepsilon_{Bs}^{total}} \frac{\alpha_{B+} \cdot \varepsilon_{B+}^{total}}{N_{B+}} \frac{f_u}{f_s} BR(B^+ \to \psi K^+) BR(\psi \to \mu^+\mu^-)$$

Key elements in the analysis:

- Construct discriminant to select B_s signal and suppress bkg
- understanding the background
- accurately measure the acceptance and efficiency ratios

Analysis optimization:

Figure of merit → expected 90% C.L. upper limit

Performed unbiased optimization

Step 1: Reconstruct "normalization mode"

Count the # of $B^+ \rightarrow \mu^+\mu^-K^+$ candidates

Selection Requires:

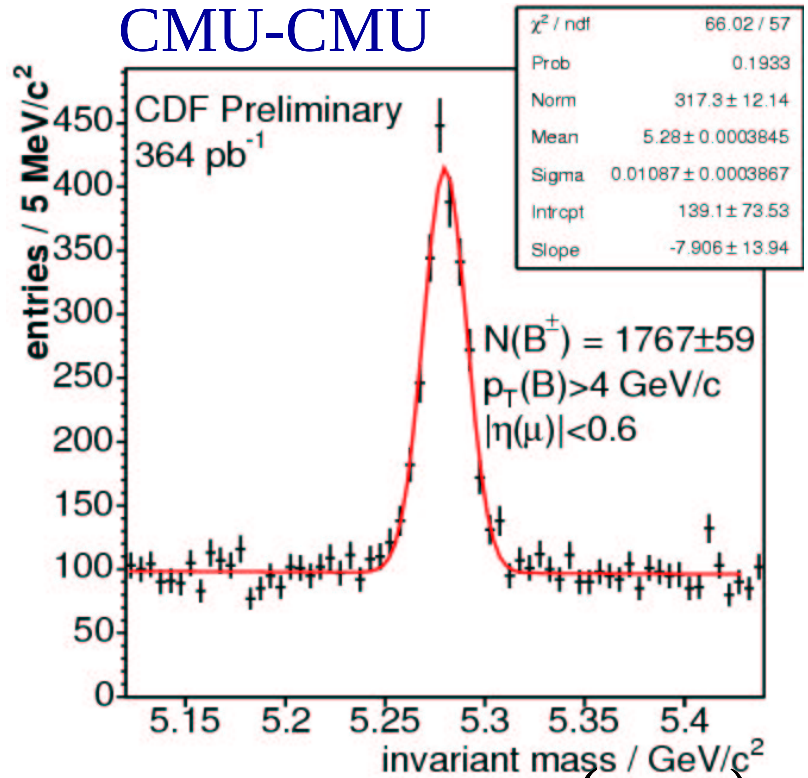
- $p_T(B)$>4 GeV && $|y(B)|<1$
- $p_T(K^+)$>1GeV
- good vertex fit quality
- $\lambda/\sigma(\lambda) > 2$

$$\lambda = \frac{cL_{3D}M_{\mu\mu}}{\left|\vec{p}(B)\right|}$$

λ = proper decay length

[λ (B^+) = ~502μm]

CMU-CMU



($m_{\mu\mu K}$)

$B^+ \rightarrow \mu\mu K^+$ in rare B trigger Sample:

N(CMU/P-CMU) = 1767±59

N(CMU/P-CMX) = 698±39

Step 2: B_s Sample Selection "baseline cuts"

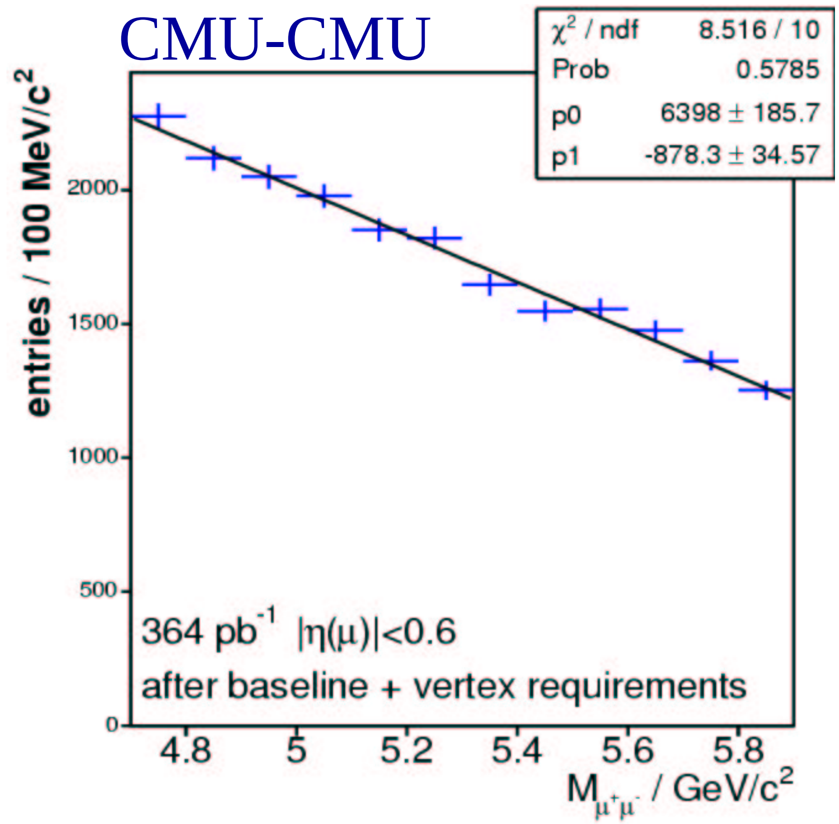
For $B_s \rightarrow \mu^+\mu^-$:

Pre-selection requires:

- $p_T(B)>4$ GeV && $|y(B)|<1$
- $\lambda/\sigma(\lambda) > 2$
- good vertex fit quality

[λ (B_s) = ~502mm]

$B_s \rightarrow \mu\mu$ Search Sample:
N(CMU-CMU) = 22459
N(CMU-CMX) = 14305
(completely Bgd dominated)



Background shapes are linear for both channels

"Baseline" cuts are loose cuts to reject events that are clearly background

Step 3: Enhance Bs,d→μμ and Suppress Background

We have explored various discriminating variables for selecting B→μμ events and suppress bkg. The chosen ones are:

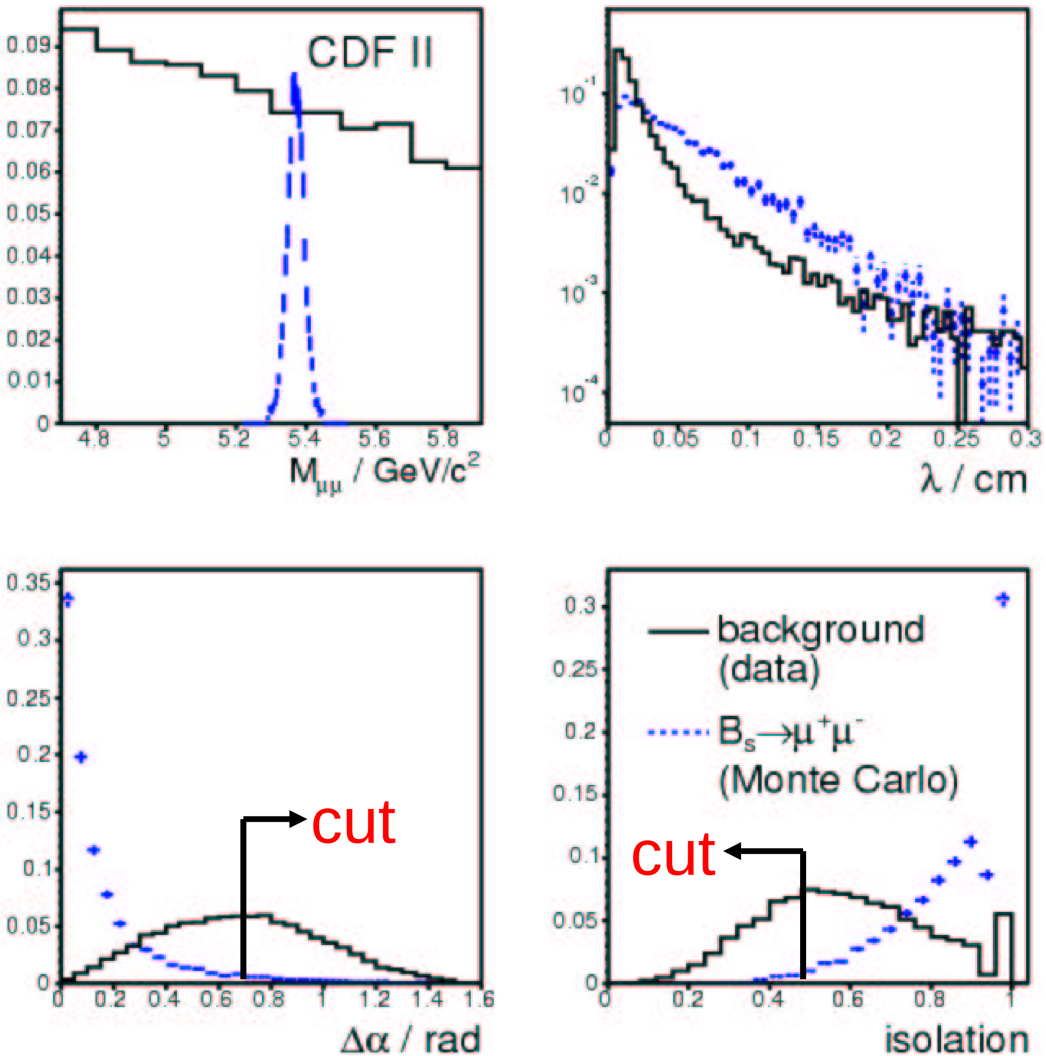
- Invariant μ+μ- mass, $M_{vtx.}$: within +/- 60MeV (2.5σ)
- Proper decay-length (λ): $\lambda = \frac{cL_{3D}M_{vtx}}{|\vec{p}(B)|}$
- Isolation (Iso): $Iso = \frac{p_T(B)}{p_T(B) + \sum_i p_T^i(\Delta R_i < 1.0)}$

  (fraction of p_T from B→μμ within $\Delta R=(\eta^2+\phi^2)^{1/2}$ cone of 1)

- "pointing (Δα)": $\Delta\alpha = \angle(\vec{p}(B) - \vec{L}_{3D})$

  (3D opening angle between B_s momentum and decay axis)

Discriminating Variables (Sig vs BKG)



To further reduce Bgd, we apply the additional cuts:

$\Delta\alpha < 0.70$ rad
&&
Iso > 0.50

eff(signal) ~ 92%

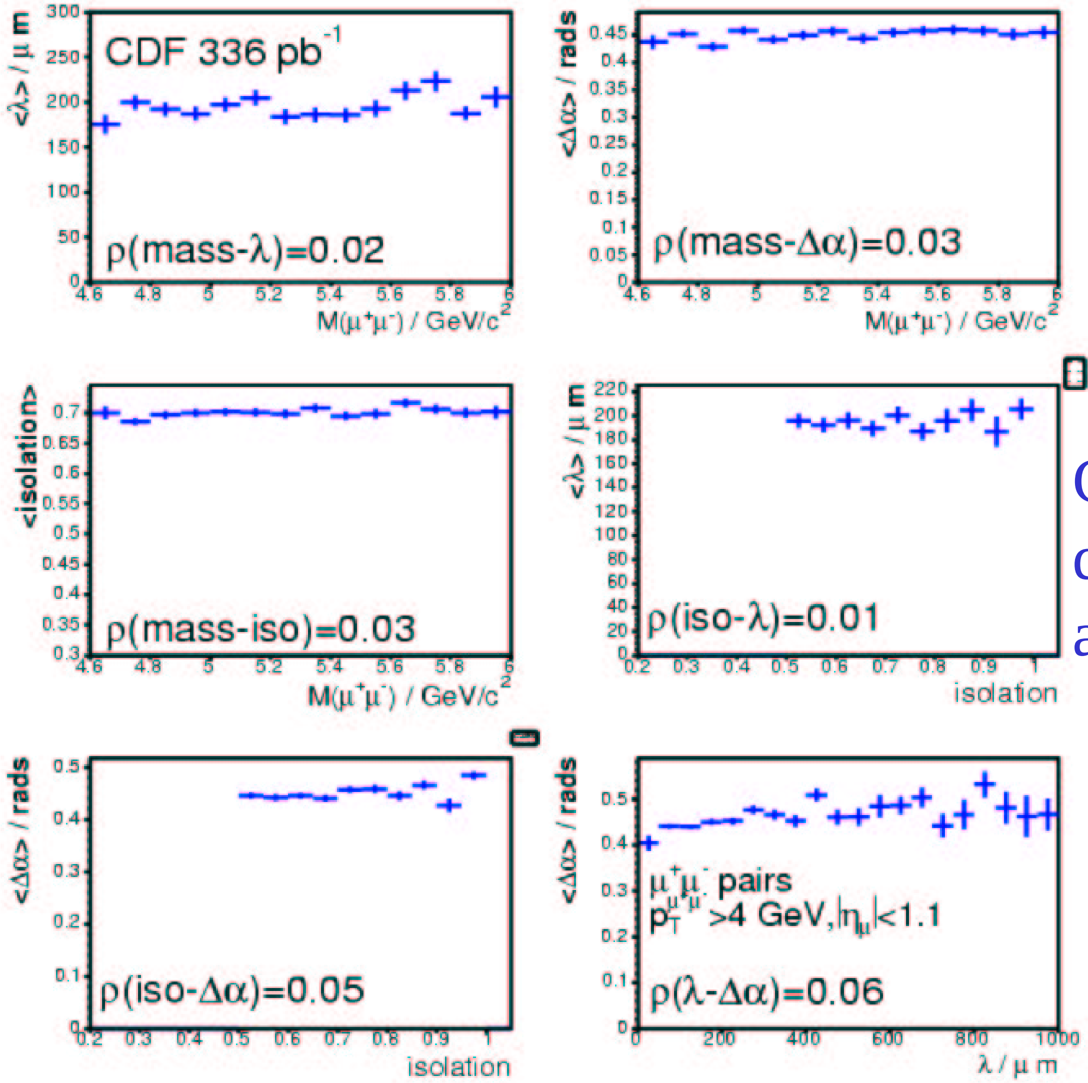
This leaves in $\mu^+\mu^-$ data:

N(CMU-CMU) = 6242

N(CMU-CMX) = 4908

~x3 down in bkg but still…

Correlations Between Variables



$$\rho_{xy} = \frac{1}{N-1} \cdot \frac{\sum_{i=1}^{N} (x_i - \hat{x})(y_i - \hat{y})}{\sigma_x \sigma_y}$$

Correlations between discriminating variables are negligible:

→ straighforward to construct likelihood discriminant (see next slide)

Likelihood Ratio Discriminant

- We use a likelihood ratio method: $$LH = \frac{\prod_i P_s(x_i)}{\prod_i P_s(x_i) + \prod_i P_b(x_i)}$$

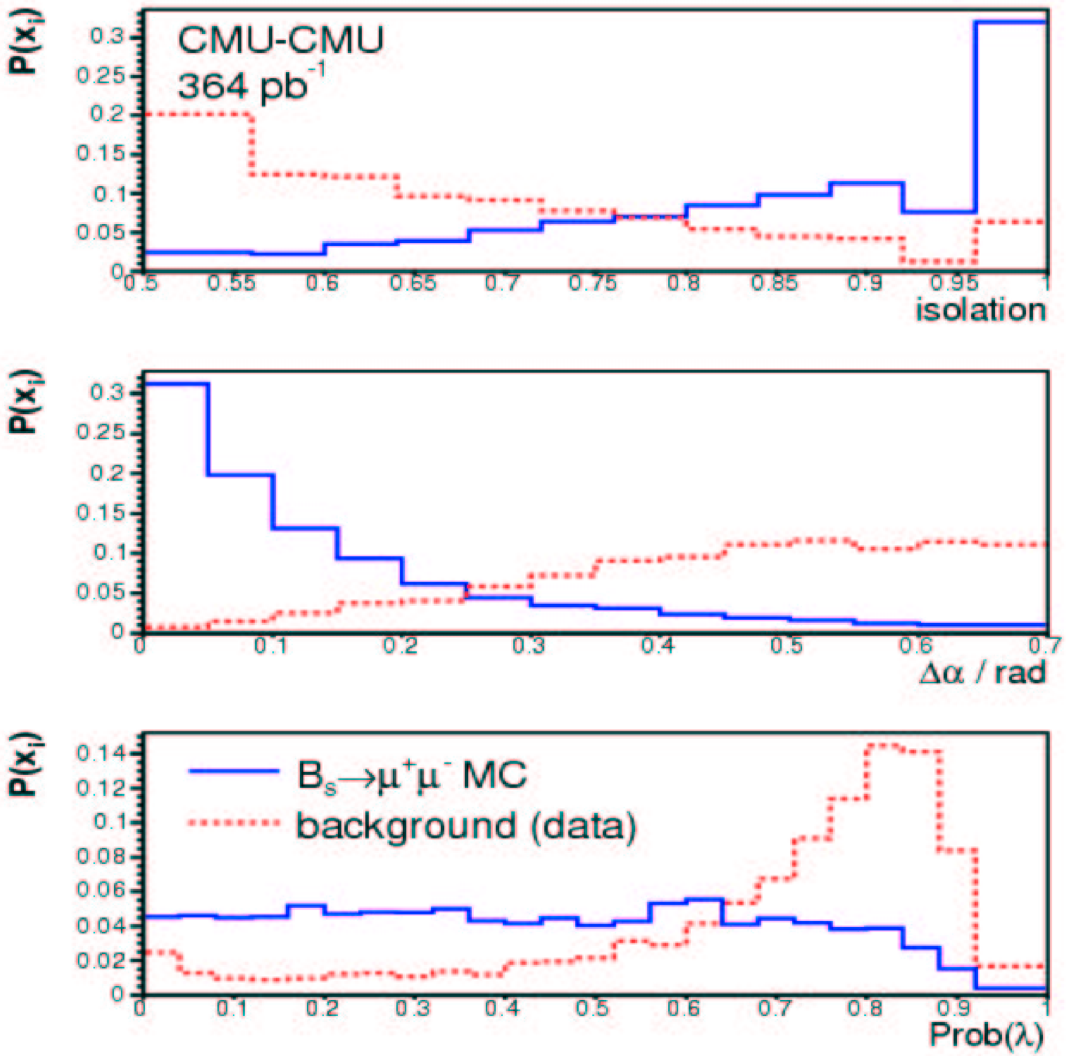
  $P_{s/b}$ is the probability for a given sig/bkg to have a value of x, where i runs over all discriminating variables.

- The chosen variables are:
  - isolation (iso)
  - 3D pointing ($\Delta\alpha$),
  - proper decay length probability [$P(\lambda)=\exp(-\lambda/\lambda_{Bs})$]

- PDF for the individual likelihood is reconstructed from the data sideband for background and Pythia MC for signal

Likelihood PDFs

CMU-CMU



Signal and background PDFs for:

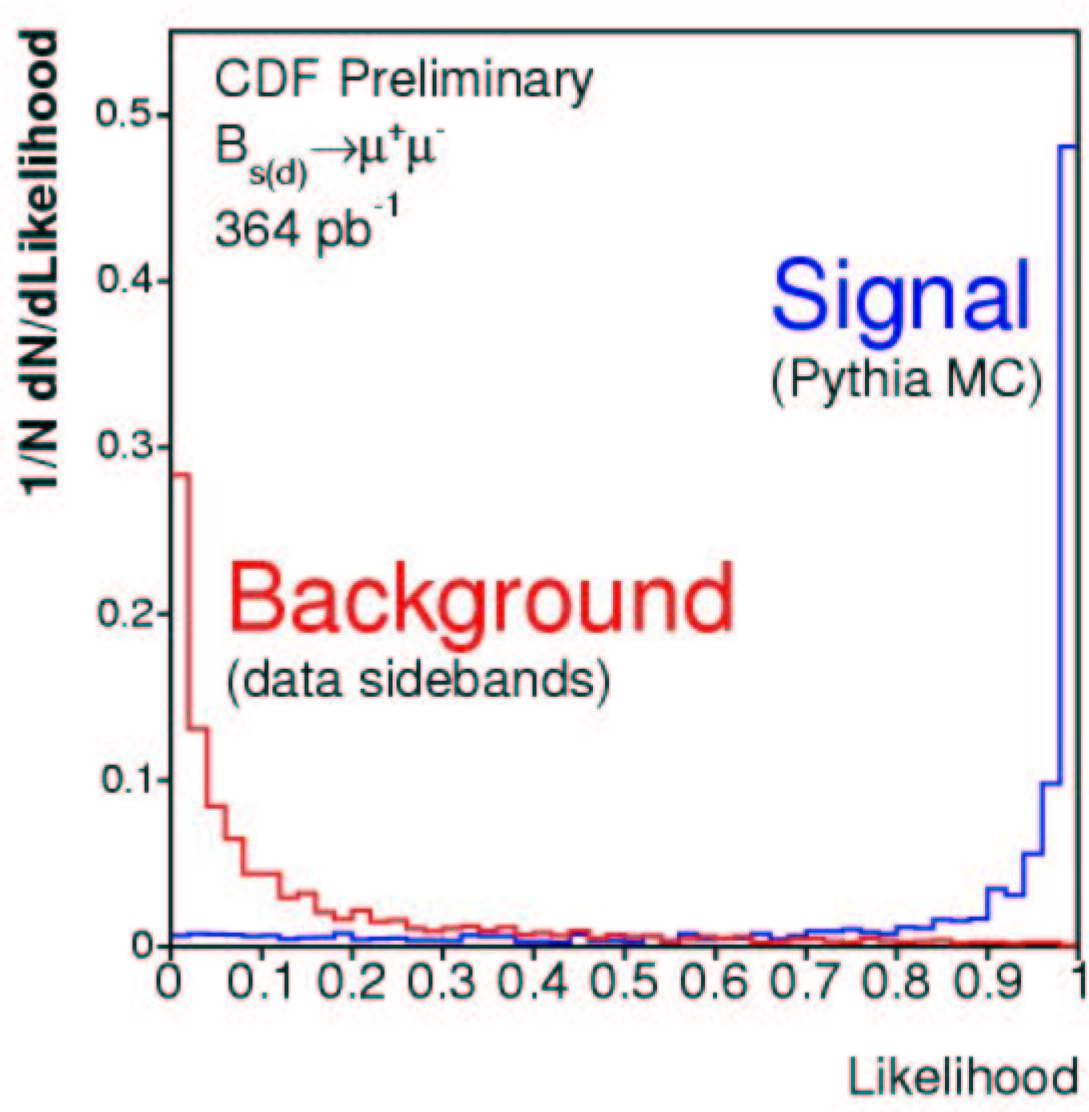
Isolation

Pointing angle

Proper decay length probability

* Similar distributions for CMU-CMX

Likelihood Ratio Signal vs Bkg



Likelihood ratio has strong discriminating power between signal and background!!

Step 4: Understand Background

Dominant bkg is combinatorial

- Events with two tracks identified as muons (could be one of more fakes)
- Mass distribution linear
- Estimate from upper and lower mass side-bands
- Also investigated possible backgrounds using a generic bbbar MC

$B \rightarrow \pi\pi$, $K\pi$, KK (with K, π misidentified as muons):

- Measured π and K μ-fake rates using $D^0 \rightarrow K\pi$ (D* tagged) events from data
- Convolute muon fake rates with $B \rightarrow hh$ BR and p_T spectra

→ Background contribution from two body hadronic B decay is below our expected sensitivity

Use data control-samples to cross-check background estimates!!!

Step 5: Estimate Background

We extrapolate number of events in the side-bands to the signal region to estimate expected background

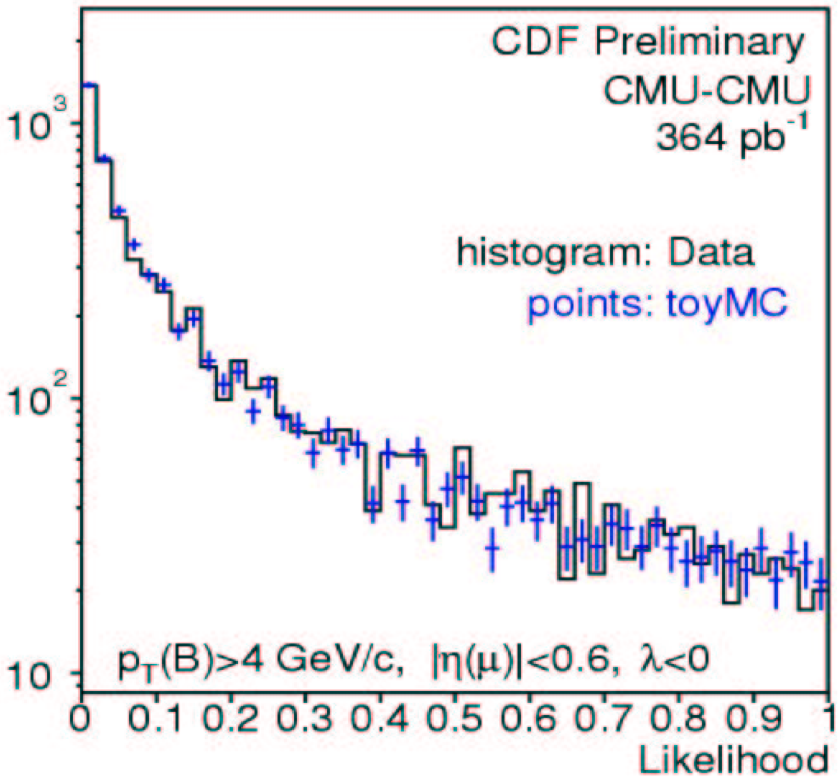
$$B = N_{SB} \cdot R_{mass} \cdot R_{LH}$$

- B = #events in signal region surviving all requirements
- N_{SB} = #events in mass sidebands surviving baseline+vtx cuts
- R_{mass} = ratio of the width of the side-band to signal mass windows
- R_{LH} = fraction of bgd events expected to survive LH cut

Estimate LH Bkg Rejection (R_{LH})

Since discriminating variables are uncorrelated, use toy MC to estimate R_{LH} based on input distributions from data SB

Likelihood Ratio Data vs Toy MC



KS-Prob(CMU)=11%
KS-Prob(CMX)=5%

Likelihood Ratio Rejection from Toy MC

	RLH	
Cut	CMU-CMU	CMU-CMX
LH>0.85	0.0245+/-0.0005	0.0226+/-0.0005
LH>0.92	0.0130+/-0.0004	0.0120+/-0.0003
LH>0.99	0.0014+/-0.0001	0.0015+/-0.0001
		(Errors are stat only)

LH strongly suppresses bkg

Cross-Check Bkg Estimate Using Control Samples

1.) OS- : opposite-charge dimuon, $\lambda < 0$

2.) SS+ : same-charge dimuon, $\lambda > 0$

3.) SS- : same-charge dimuon, $\lambda < 0$

4.) FM : fake muon sample (at least one leg failed muon stub matching cut)

	LH cut	CMU-CMU pred	CMU-CMU obsv	CMU-CMX pred	CMU-CMX obsv
	>0.50	236+/-4	235	172+/-3	168
OS-	>0.90	37+/-1	32	33+/-1	36
	>0.99	2.8+/-0.2	2	3.6+/-0.2	3
	>0.50	2.3+/-0.2	0	2.8+/-0.3	3
SS+	>0.90	0.25+/-0.03	0	0.44+/-0.04	0
	>0.99	<0.10	0	<0.10	0
	>0.50	2.7+/-0.2	1	3.7+/-0.3	4
SS-	>0.90	0.35+/-0.03	0	0.63+/-0.06	0
	>0.99	<0.10	0	<0.10	0
	>0.50	84+/-2	84	21+/-1	19
FM+	>0.90	14.2+/-0.4	10	3.9+/-0.2	3
	>0.99	1.0+/-0.1	2	0.41+/-0.03	0

Step 6: Compute Acceptance and Efficiencies

$$\left(\frac{\alpha_{B+}}{\alpha_{Bs}}\right)\cdot\left(\frac{\varepsilon_{B+}^{trig}}{\varepsilon_{Bs}^{trig}}\right)\cdot\left(\frac{\varepsilon_{B+}^{reco-\mu\mu}}{\varepsilon_{Bs}^{reco-\mu\mu}}\right)\cdot\left(\frac{\varepsilon_{B+}^{vtx}}{\varepsilon_{Bs}^{vtx}}\right)\cdot\varepsilon_{B+}^{reco-K}\cdot\frac{1}{\varepsilon_{Bs}^{LH}}$$

Red = From Pythia MC   Green = From J/ψ Data
Blue = combination of MC and Data

- $\alpha(B^+/B_s)$ = 0.297 +/- 0.008 (CMU-CMU)
  = 0.191 +/- 0.006 (CMU-CMX)
- $\varepsilon^{trig}(B^+/B_s)$ = 0.9997 +/- 0.0016 (CMU-CMU)
  = 0.9986 +/- 0.0014 (CMU-CMX)
- $\varepsilon^{reco-\mu\mu}(B^+/B_s)$ = 1.00 +/- 0.03 (CMU-CMU/X)
- $\varepsilon^{vtx}(B^+/B_s)$ = 0.986 +/- 0.013 (CMU-CMU/X)
- $\varepsilon^{reco-K}(B^+)$ = 0.938 +/- 0.016 (CMU-CMU/X)

Likelihood Ratio Efficiency for B_s Signal

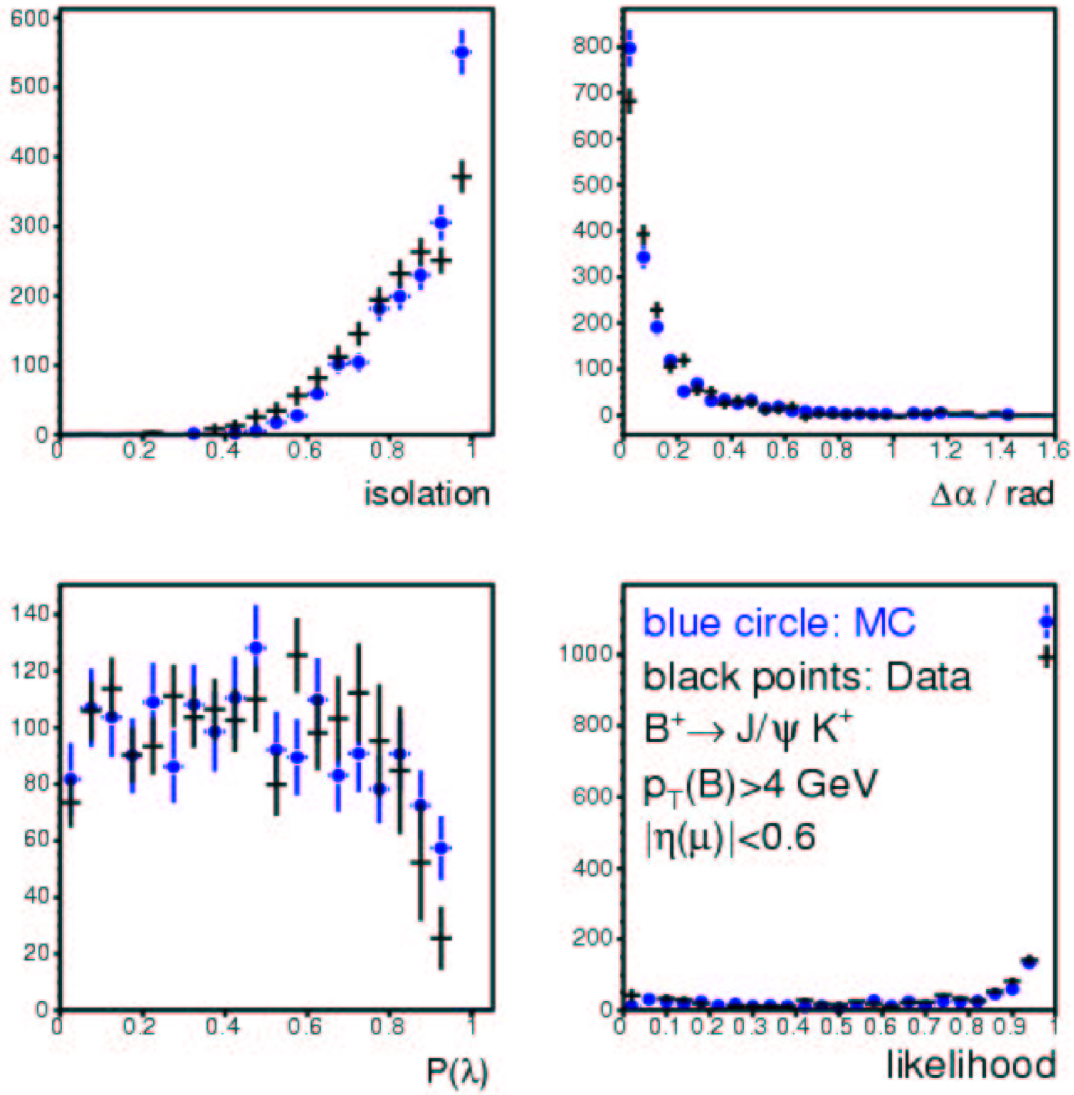
- determined from $B_s \rightarrow \mu\mu$ MC
- MC modeling checked by comparing $\varepsilon_{LH}(B^+)$ between MC and sideband subtracted Data

cut	$\varepsilon_{LH}(B_s)$ CMU-CMU	$\varepsilon_{LH}(B_s)$ CMU-CMX
LH>0.90	(70+/-1)%	(66+/-1)%
LH>0.92	(67+/-1)%	(65+/-1)%
LH>0.95	(61+/-1)%	(60+/-1)%
LH>0.98	(48+/-1)%	(48+/-1)%
LH>0.99	(38+/-1)%	(39+/-1)%

(stat uncertainties only)

Checking MC Modeling of Signal LH

Compare B^+ Data and MC



For CMU-CMU:

- MC reproduces Data efficiency vs LHood cut to 10% or better
- Assign 10% (relative) systematic
- CMU-CMX MC vs Data agreement is better

Step 7: Analysis Optimization

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

$$\langle BR(B_s \to \mu^+\mu^-)\rangle = \frac{\langle N^{90\% CL}_{signal}\rangle}{N_{B^+\to J/\psi K}} \cdot \frac{1}{\varepsilon^{Bs}_{LH}} \cdot \frac{(\alpha\varepsilon)^{tot}_{B^+}}{(\alpha\varepsilon)^{tot}_{Bs}} \cdot \frac{f_d}{f_s} \cdot Br(B^+ \to J/\psi K)$$

Quantities to be optimized

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

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 (w/ flat prior) and including uncertainties

Optimization Result

- Stat and syst uncertainties are included in the limit calculation

    Dominant uncertainty is f_s/f_u from PDG ~ 15%(rel)

    Fragmentation ratio (B_s/B_d)

- For optimization, we scan: LH>0.90-0.99, p_T(B)>4-6 GeV
- Assume 1 fb^{-1} of data

→ Optimal cuts: LH>0.99 and p_T(B)>4GeV

More Cross-Checks Prior to Looking in the Box

- We "opened" the box in successive steps. Check bkg estimate at each step to look for surprises

- Estimating # of bkg in the data signal region:

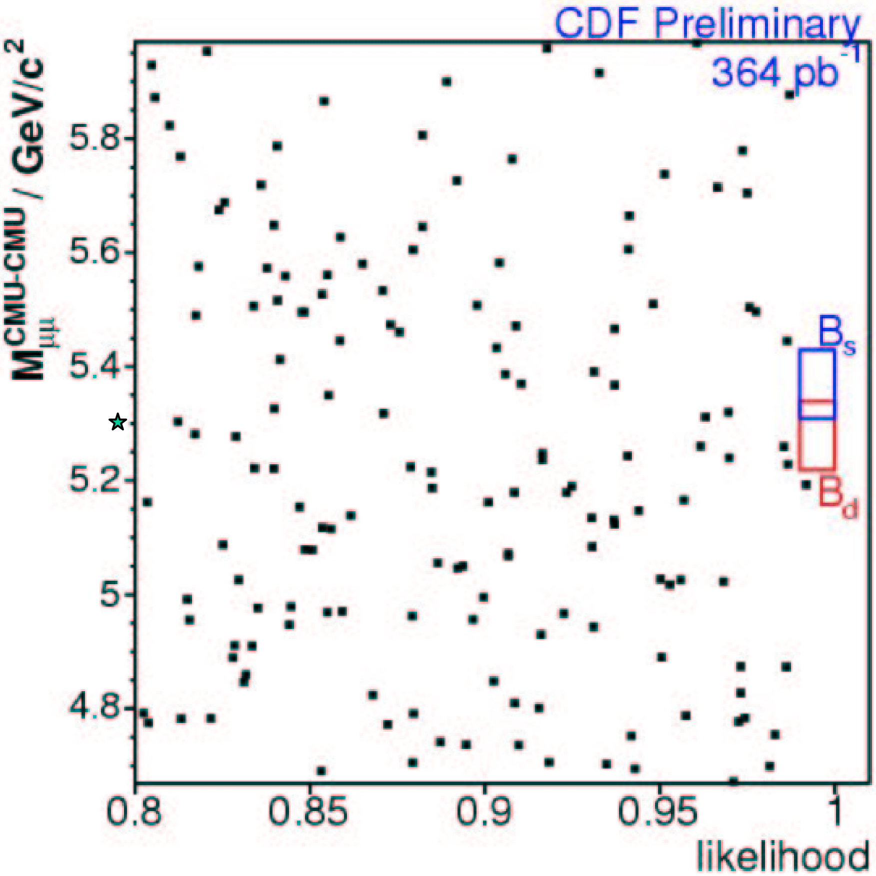
	CMU-CMU		CMU-CMX		
	Npred	Nobsv	Npred	Nobsv	
LH>0.50:	146 +/- 22	136	99 +/- 20	99	
LH>0.90:	24 +/- 4	20	17 +/- 3	9	(Poisson Prob=2.6%)

- We have also compared the likelihood distribution between sideband and signal region (dominated by bkg). The resulting KS probabilities are: 66% (CMU-CMU) and 76% (CMU-CMX)

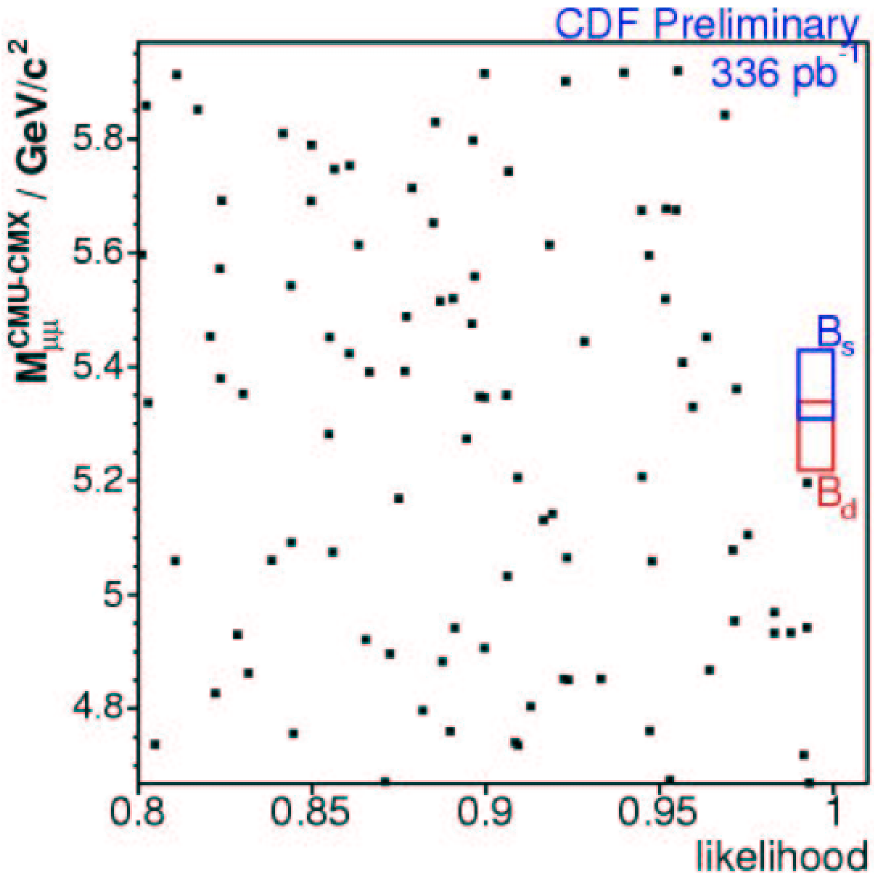
Step 8: Looking at Results

For optimized cuts of LH >0.99 and $p_T(B)$ > 4GeV
and a ± 60 MeV window around world avg B mass

CMU-CMU Channel



CMU-CMX Channel



We observed 0 event in the signal region!

$B_s \to \mu\mu$ Summary

For optimized cuts of LH >0.99 and $p_T(B)$ > 4GeV

CMU-CMU:

Single event sensitivity = $(1.0\pm0.2)\times10^{-7}$

Expected # bkg ($364pb^{-1}$) = 0.81 ± 0.12

CMU-CMX:

Single event sensitivity = $(1.6\pm0.3)\times10^{-7}$

Expected # bkg ($336pb^{-1}$) = 0.66 ± 0.13

Comparing sensitivity with previous analysis:

Expected limit for this new analysis

$< \mathbf{2.0}\times10^{-7}$ **@ 90% CL**

Old analysis using this same data-set:

$< \mathbf{3.3}\times10^{-7}$ **@ 90% CL**

Limits Summary

Bs: we observed 0 events which yields a combined limit of:

1.6×10^{-7} @ 90% CL

2.1×10^{-7} @ 95% CL

Bd: we observed 0 events which yields a combined limit of:

3.9×10^{-8} @ 90% CL

5.1×10^{-8} @ 95% CL

$Br(B_s \to \mu\mu) < 4.1\times10^{-7}$ @ 90% CL ; D0 PRL 94 (2005) 042001 ($240pb^{-1}$)

$Br(B_s \to \mu\mu) < 5.8\times10^{-7}$ @ 90% CL ; CDF PRL 93 (2003) 032001 ($171pb^{-1}$)

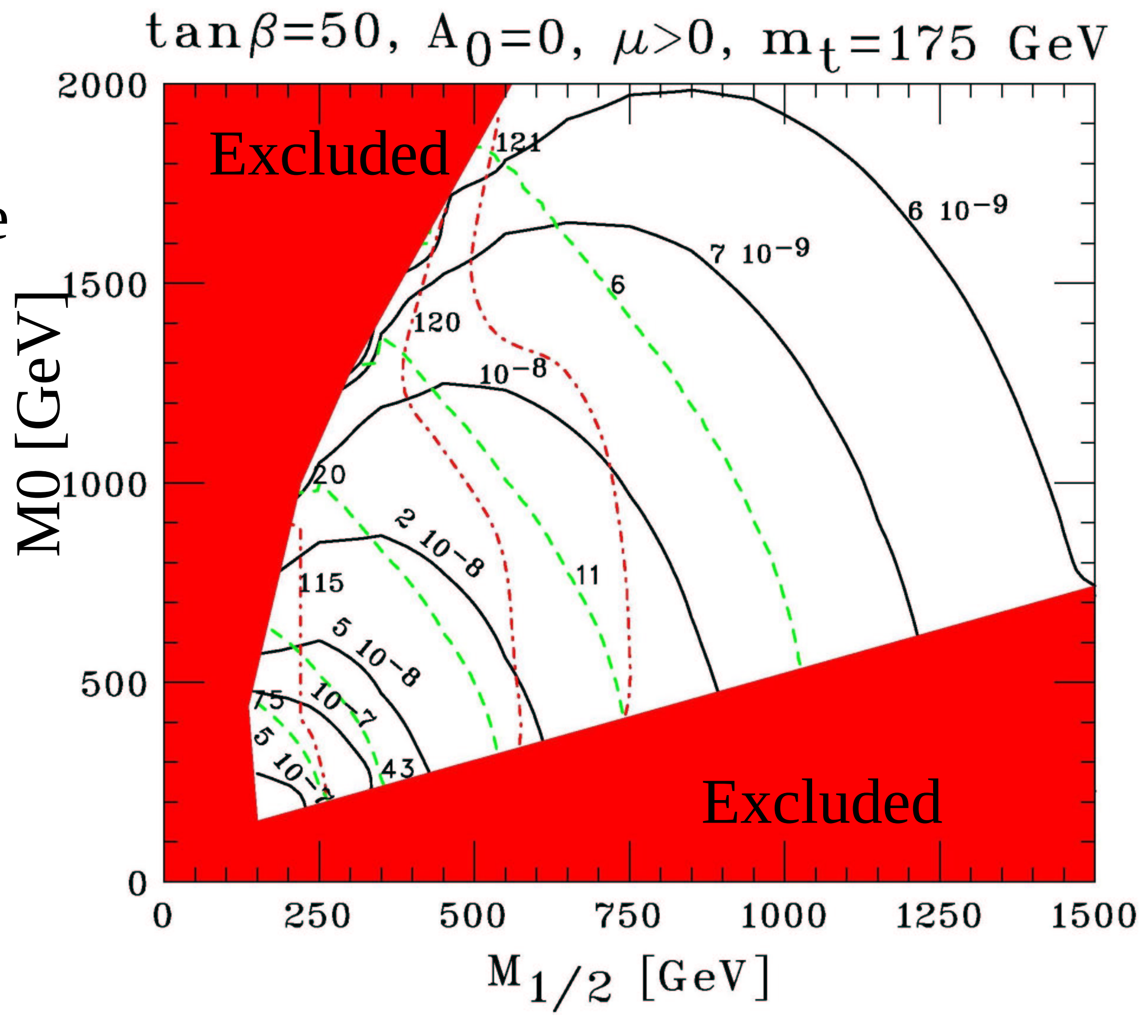
$Br(B_d \to \mu\mu) < 8.0\times10^{-8}$ @ 90% CL ; BaBar PRL 94 (2005) 221803 ($111fb^{-1}$)

Both CDF B_s and B_d results are x2 better than the best published result!!!

mSUGRA M0 vs $M_{1/2}$

Dedes, Dreiner, Nierste, PRL 87(2001) 251804

- We are beginning to carve into mSUGRA space
- For m_h~115GeV implies $10^{-8} < Br(B_s \to \mu\mu) < 3\times10^{-7}$



Solid red = excluded by theory or experiment

Dashed red line = light Higgs mass (m_h)

Dashed green line = (δa_μ)susy (in units of 10^{-10})

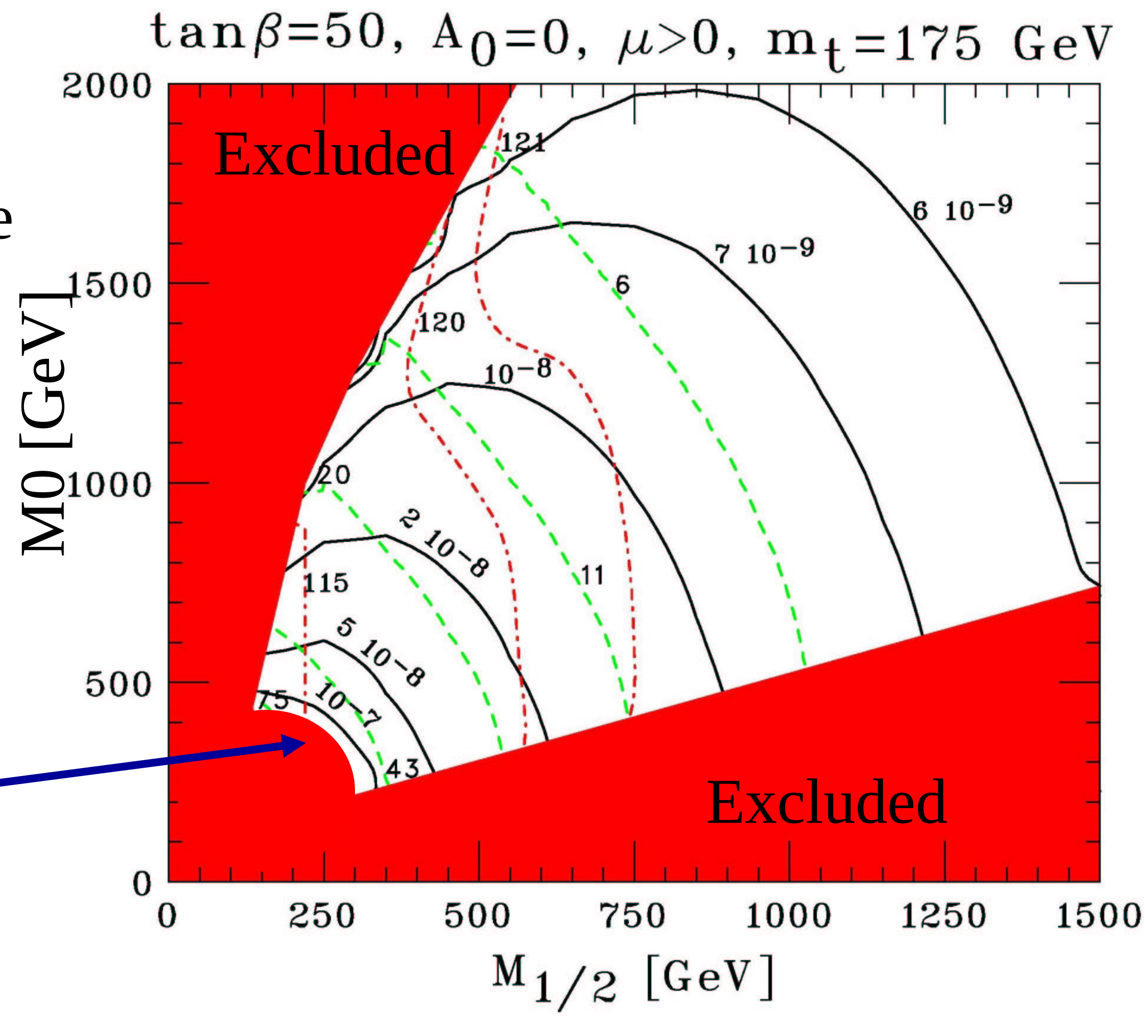
Black line = $Br(B_s \to \mu\mu)$

mSUGRA M0 vs $M_{1/2}$

Dedes, Dreiner, Nierste, PRL 87(2001) 251804

- We are beginning to carve into mSUGRA space
- For m_h~115GeV implies 10^{-8}<Br($B_s \rightarrow \mu\mu$)<3×10^{-7}

Excluded by this new result



Solid red = excluded by theory or experiment

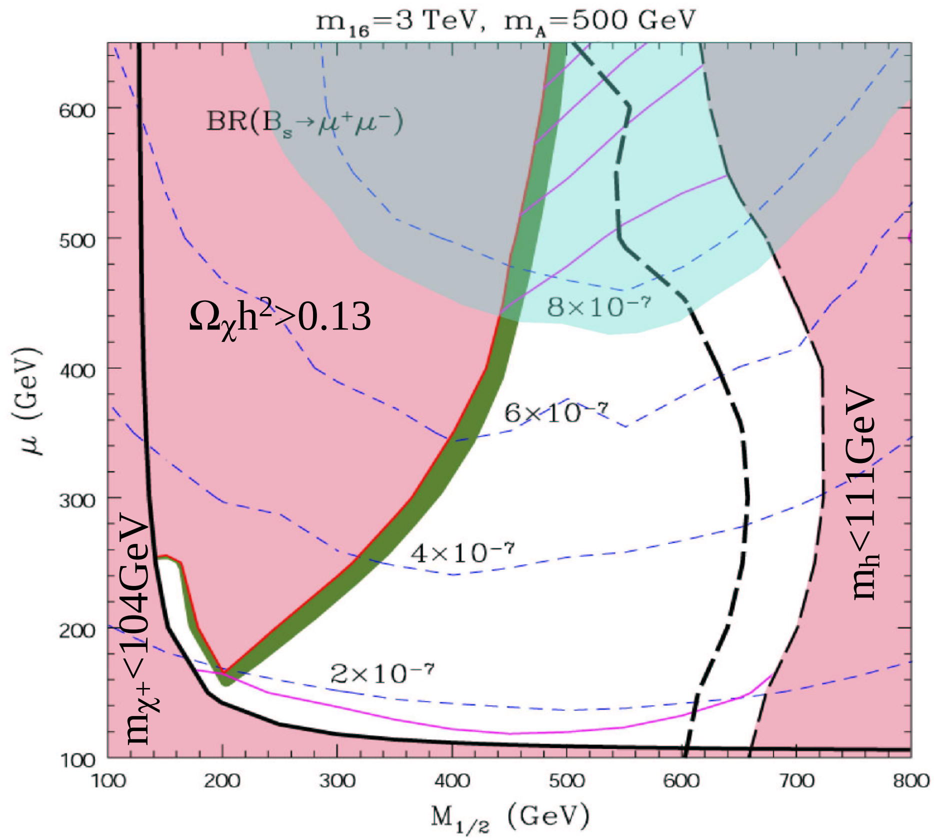
Dashed red line = light Higgs mass (m_h)

Dashed green line = (δa_μ)susy (in units of 10^{-10})

Black line = Br($B_s \rightarrow \mu\mu$)

SO(10) Unification Model

R. Dermisek *et al.*, hep-ph/0304101

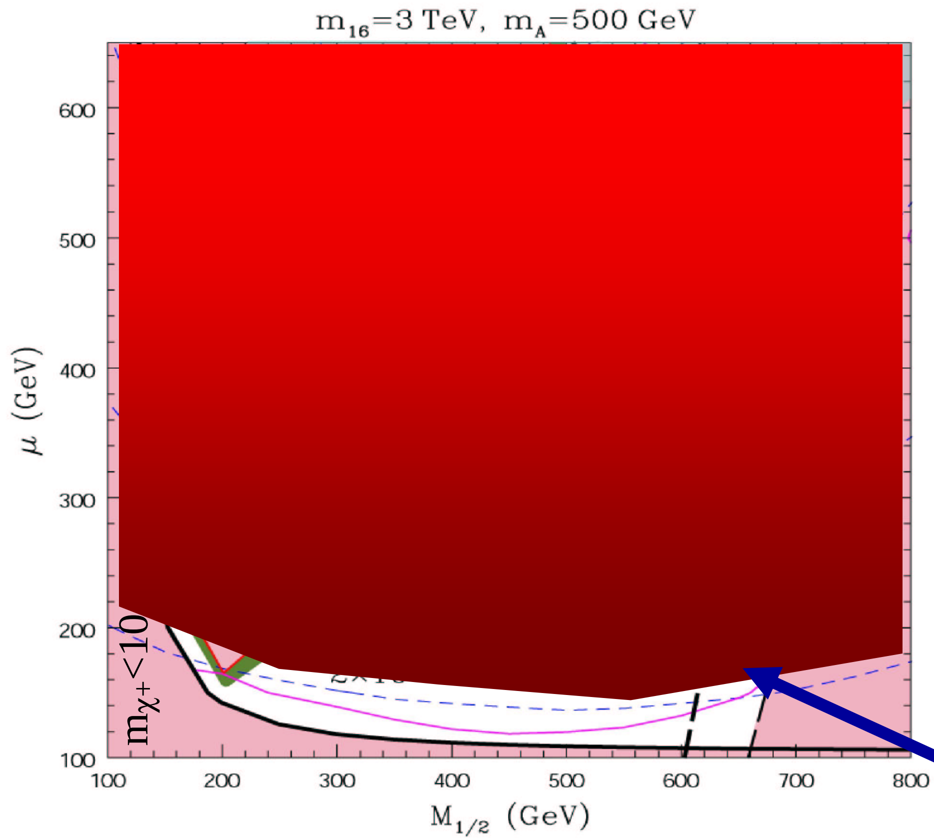


- $\tan(\beta)\sim 50$ constrained by unification of Yukawa coupling
- White region is not excluded
- Unification valid for small $M_{1/2}$ (~500GeV)

Red regions are excluded by either theory or experiments
Green region is the WMAP preferred region
Blue dashed line is the Br(Bs→μμ) contour
Light blue region excluded by old Bs→μμ analysis

SO(10) Unification Model

R. Dermisek *et al.*,
hep-ph/0304101



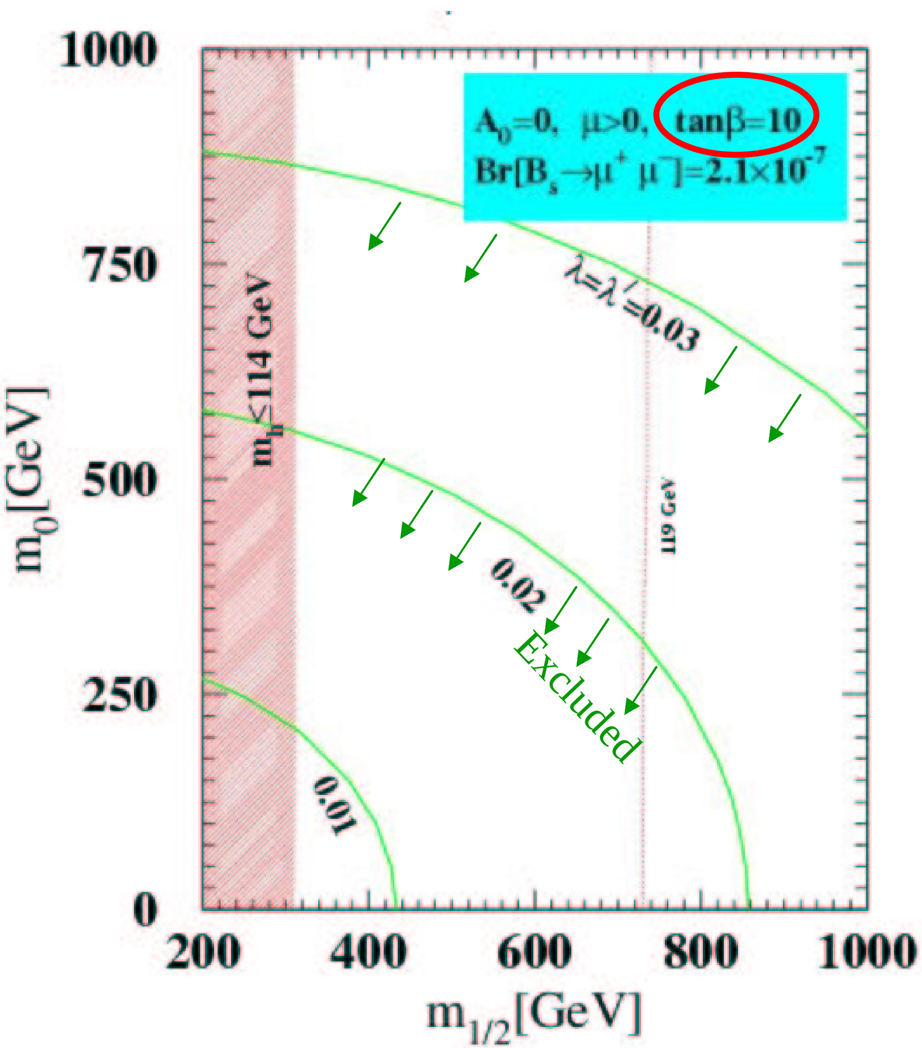
- New Br($B_s \rightarrow \mu\mu$) limit strongly disfavors this solution for m_A= 500 GeV

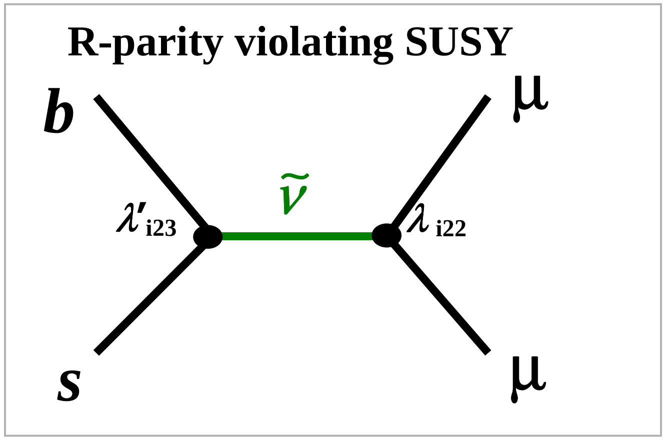
Excluded by this new result

Red regions are excluded by either theory or experiments
Green region is the WMAP preferred region
Blue dashed line is the Br(Bs→μμ) contour
Light blue region excluded by old Bs→μμ analysis

RPV SUSY EXCLUSION

B. Dutta et al, PLB 538 (2002) 121

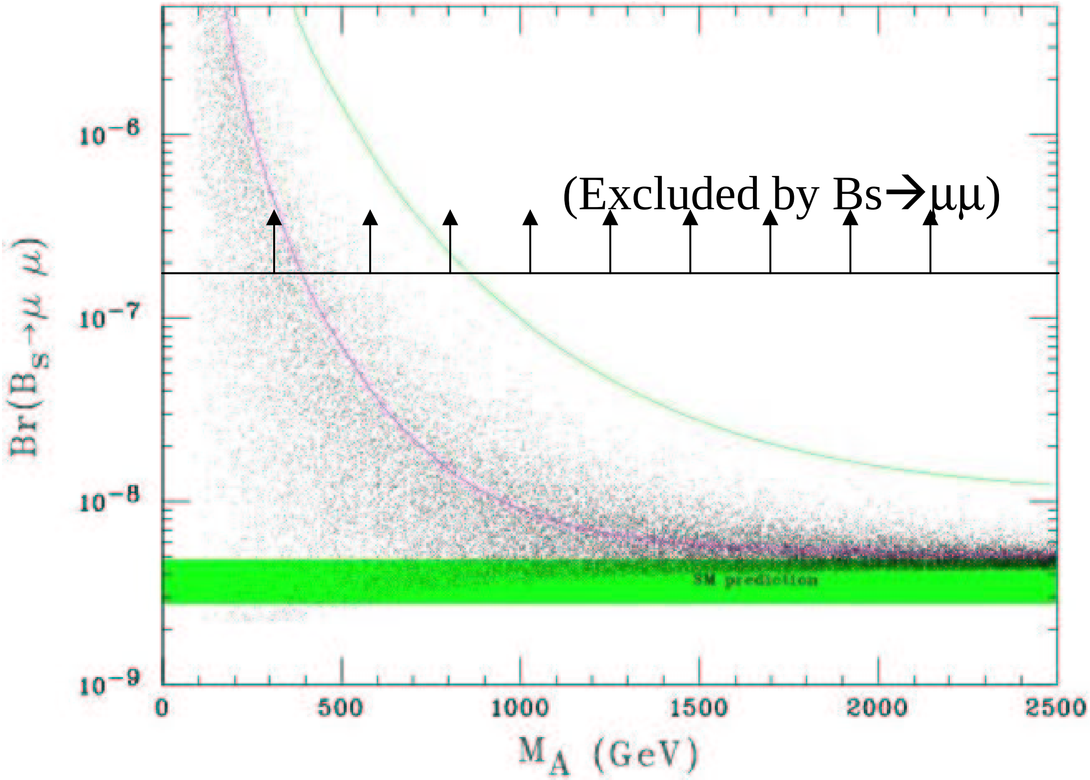




- Possible to exclude phase space even for small tan(β)
- Exclusion strongly depends on the coupling.

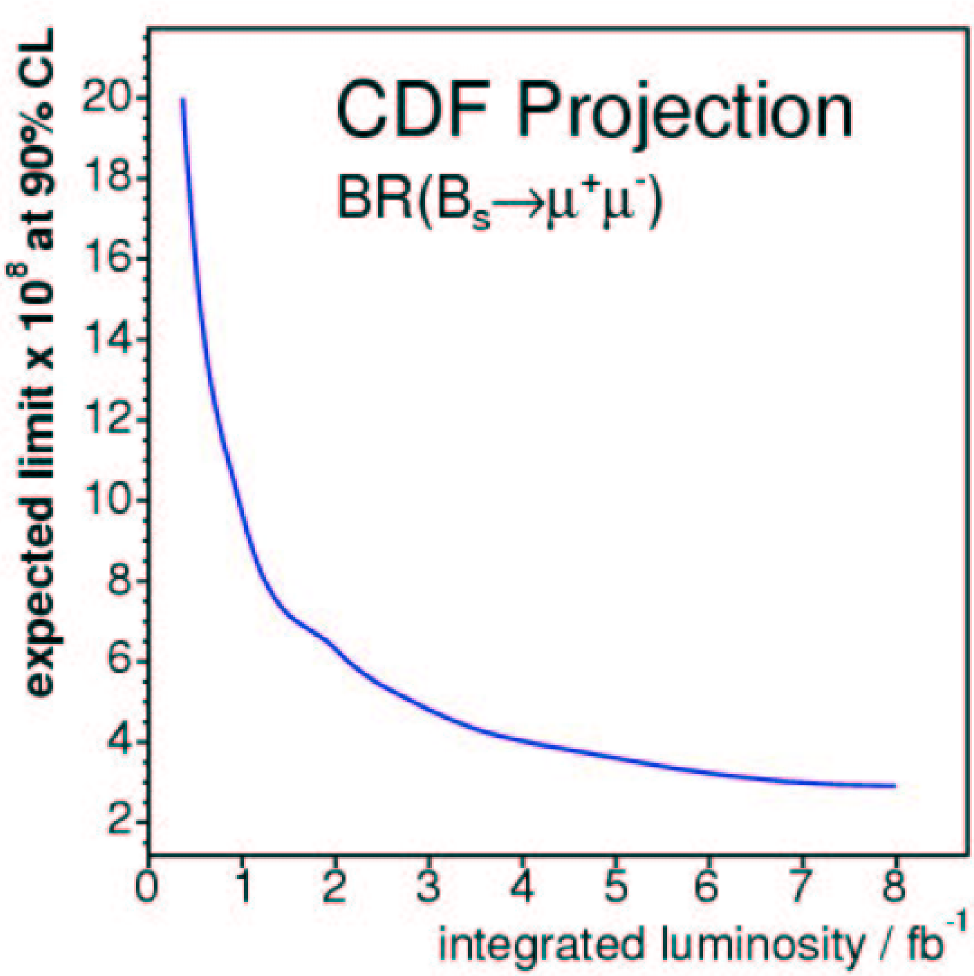
B→μμ Sensitivity To Heavy Higgs

MFV MSSM (tanβ=50)
(A. Dedes et al, hep-ph/0407285)



- Br(Bs→μμ) is sensitive to the mass of heavy Higgs
- If the branching ratio is measured → sets an upper limit on the mass of the heavy Higgs
- m_A mass limit is where BR crosses the green curve (same argument hold of any tanβ≤ 50)
- The mass limit is fairly model independent

Bs→μμ Limit Projection



- Extrapolate based on the current analysis which was optimized for 1/fb
- Assume background and single-event-sensitivity scale linearly with luminosity
- Will need to re-optimize the analysis for > 3/fb

CDF and D0 Working Group

- CDF-D0 working group is formed to combine the $B \rightarrow \mu\mu$ limits from both experiments:
  D0 Preliminary : $Br(B_s \rightarrow \mu\mu) < 3.0\times10^{-7}$ @ 90% CL (D0 note 4733, ~300pb^{-1})
  CDF Preliminary : $Br(B_s \rightarrow \mu\mu) < 1.6\times10^{-7}$ @ 90% CL
- Two independent groups cross-checking each other's combined results. Aim to release preliminary combined results for LP05
- Combined CDF and D0 results is expected to improve the limit by ~20%

Summary

- $B_s \rightarrow \mu\mu$ is a powerful probe of new physics. Could potentially provide the first hint of SUSY at the Tevatron
- Using 364/pb of data, CDF has obtained world best limits on B_s and B_d channels (working on PRL draft):

$$Br(B_s \rightarrow \mu\mu) < 1.6\times10^{-7} \text{ @ 90\% CL}$$
$$< 2.1\times10^{-7} \text{ @ 95\% CL}$$

$$Br(B_d \rightarrow \mu\mu) < 4.2\times10^{-8} \text{ @ 90\% CL}$$
$$< 5.5\times10^{-8} \text{ @ 95\% CL}$$

- The limits are now starting to constrain interesting regions of SUSY parameter space
- We have covered an order of magnitude since RunI result. Will cover at least another order of magnitude before the end of RunII

Hint of SUSY may just be around the corner!!