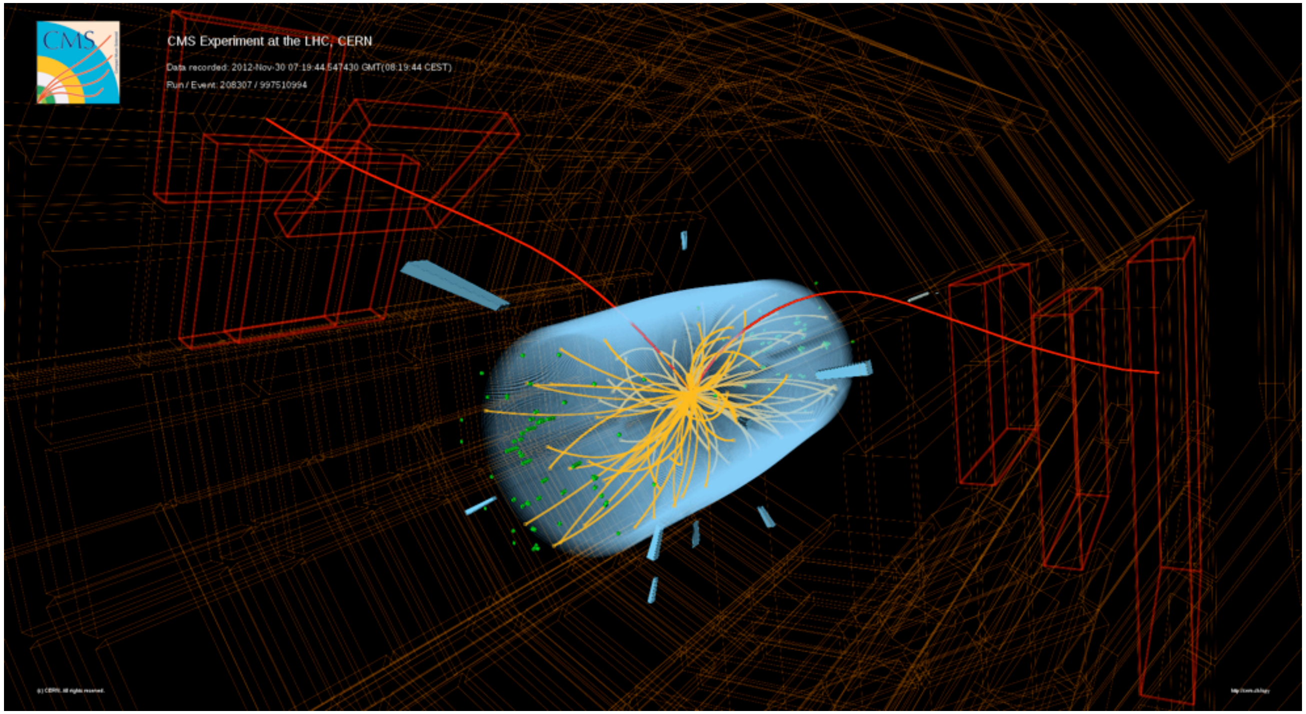
The Undefeated and Triumphant Standard Model of Particle Physics

Observation of Rare $B_s \rightarrow \mu\,\mu$ Decay



M. Baarmand

Oct. 10, 2013

A Standard Model

- **A SM is a theoretical framework built from and consistent with all relevant available observations, which predicts new testable phenomena…**
  - must be self-consistent and consistent with all relevant data
  - must have predictive power
  - must be "simple" and "elegant"

. **Examples:**

Ptolemy's epicycles for planets: predicted Paths of the Planets

Keppler's mechanics of planets: simplified/improved dramatically on the above!

Newton's mechanics: gave a deep understanding in terms of Gravity

Mendeleev's periodic table: deeper understanding of Chemistry

Einstein's STR: an "improvement" on Newtonian relativity!

Bohr model of atom

Synapse/neuron structure of brain

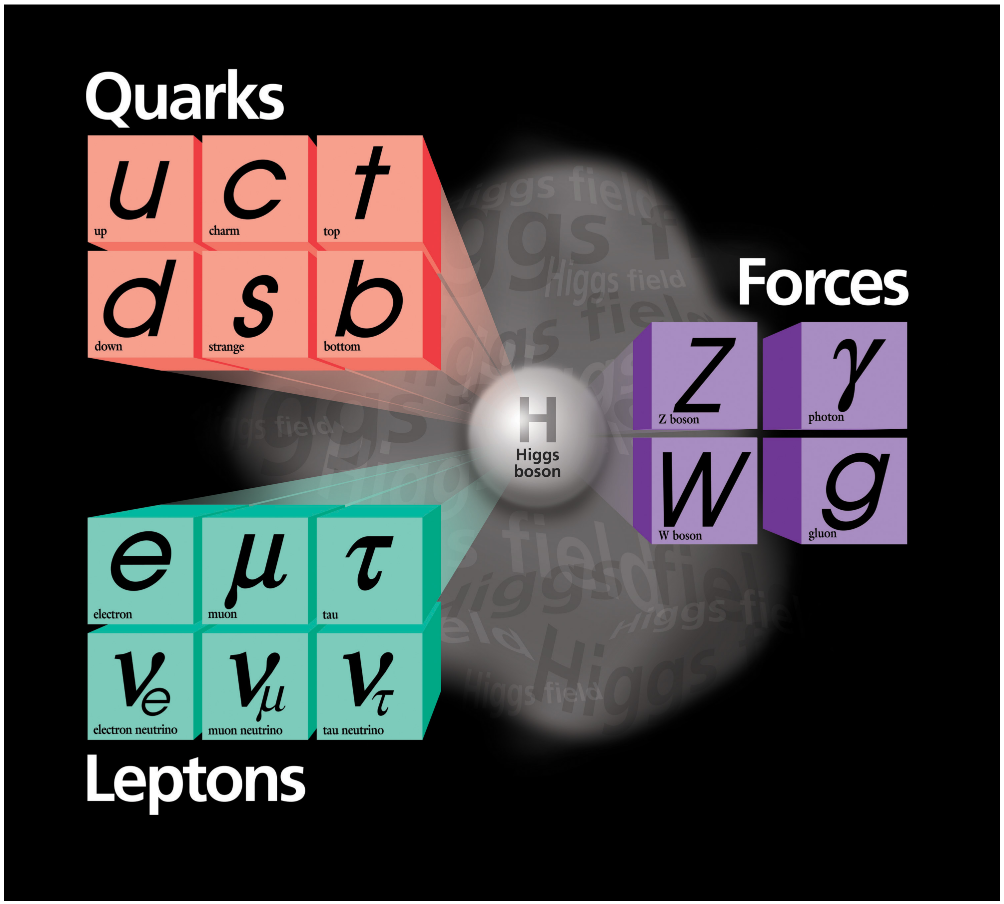
- **As a larger realm is explored, a SM may need revision**

The Standard Model of Particle Physics

A Crowning Achievement of 20th Century Science

Quantum Mechanics and Special Theory of relativity along with elementary particles discovered have led to the Standard Model of Particle Physics: Modern-day "Periodic Table" of fundamental particles and their interactions.

Matter particles



Force particles

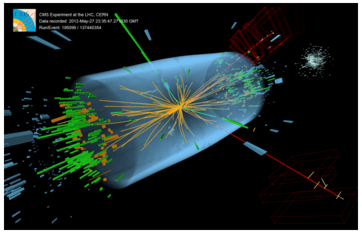
Over the past 50 years the SM has been tested to excellent precision.

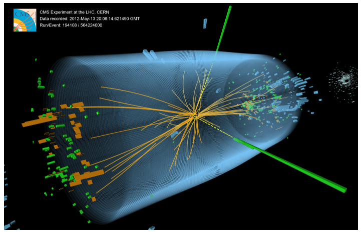
NB: Gravity not included!

Experimental Evidence for SM

- DIS experiments - existence of quarks
  SLAC 1972-73
- Observation of charm and bottom quarks
  SLAC, BNL, FNAL 1974-80
- Observation of neutral currents (Z boson exchange)
  CERN 1973
- Observation of jets and 3 jet final states (gluon radiation)
  DESY 1979-80
- Direct observation of W and Z bosons
  CERN 1983
- Direct observation of top quark
  FNAL 1995
- Direct observation of tau lepton
  FNAL 2000
- Observation of Higgs boson
  CERN 2012

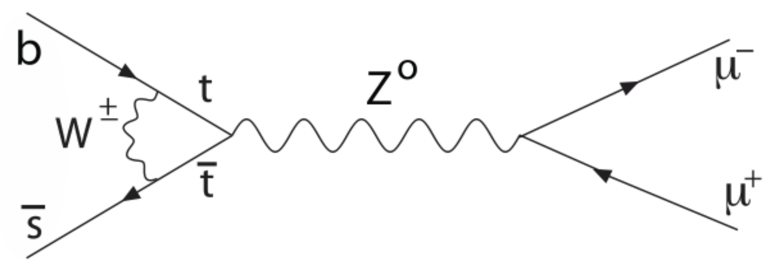
Several Nobel prizes!

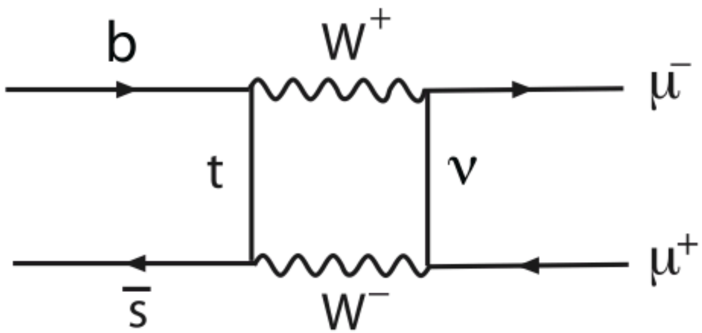




Most Recent Test of SM

- SM predicts a tiny branching fraction for $B_s \to \mu^+ \mu^-$
  - $BR^{t=0} (B_s \to \mu\mu) = (3.25 \pm 0.17) \times 10^{-9}$ [A. Buras et al. arXiv:1303.3820]
  - $BR^{<t>} (B_s \to \mu\mu) = (3.56 \pm 0.18) \times 10^{-9}$ time-integrated measured [De Bruyn et al. (PRL 109, 041801)] [A. Buras et al. arXiv:1303.3820]
  - forbidden at tree level, only through higher-order loop diagrams
  - helicity suppressed

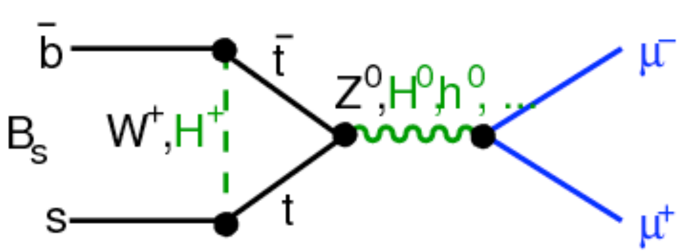


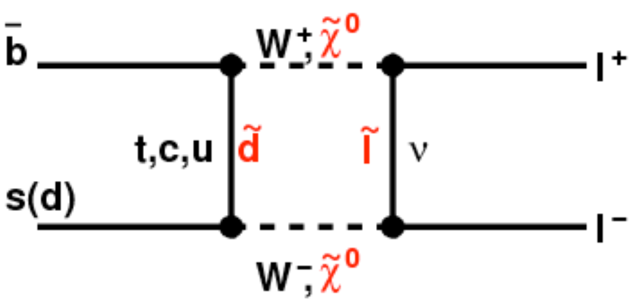


- Even smaller branching fraction for $B_d \to \mu^+ \mu^-$
  - $BR^{t=0}(Bd \to \mu\mu) = (1.07 \pm 0.10) \times 10^{-10}$
  - Cabibbo suppressed due to |Vtd|<|Vts|

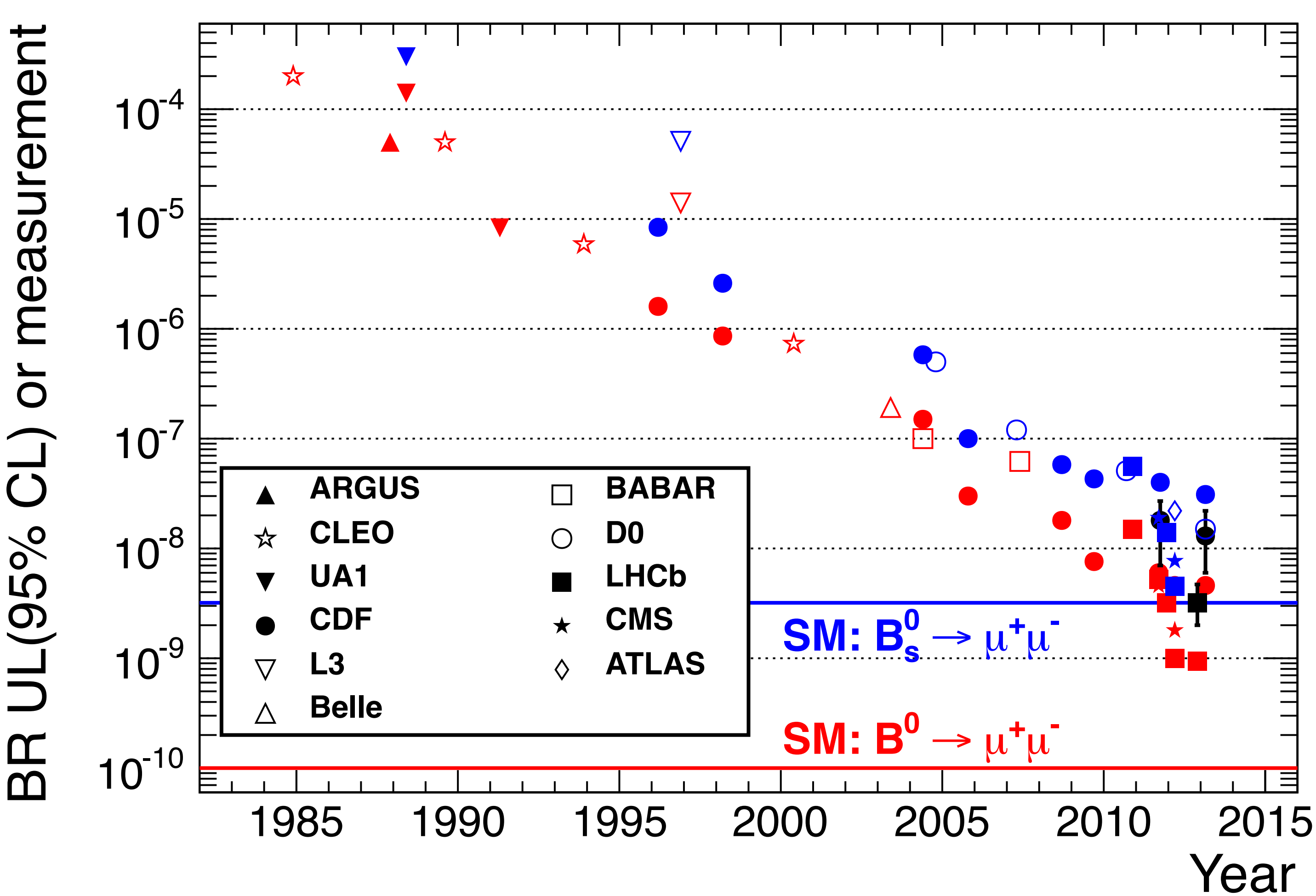
Motivation: Physics Beyond SM

- $B_s \to \mu^+ \mu^-$ and $B_d \to \mu^+ \mu^-$ sensitive probes for BSM physics
  - 2HDM: $BR(B_{s/d} \to \mu\mu) \propto \tan^4 \beta$ and $m(H^+)$
    J. R. Ellis et al, JHEP 05 (2006) 063
  - MSSM: $BR(B_{s/d} \to \mu\mu) \propto \tan^6 \beta$
    J.Parry, Nucl. Phys. B 760 (2007) 38
  - Leptoquarks
    S. Davidson and S. Descotes-Genon, JHEP 11 (2010) 073
  - 4th generation top
    - Wei-Shu Hou, Masaya Kohda, Fanrong Xu, Phys. Rev. D87, 094005 (2013)





25-year Quest for Rare B_s Decay



LHCb 3.5 σ evidence: $BR(B_S \to \mu\mu) = \left(3.2^{+1.4}_{-1.2}\ (stat)^{+0.5}_{-0.3}\ (syst)\right) \times 10^{-9}$

Best until now

ATLAS+CMS+LHCb: $BR(B_d \to \mu\mu) < 8.4 \times 10^{-10}$ @95% CL

Muon Trigger and Reconstruction

Muon Detectors

DT, CSC, RPC

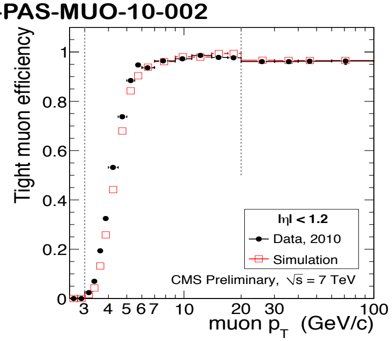
Large coverage $|\eta_\mu|<2.4$

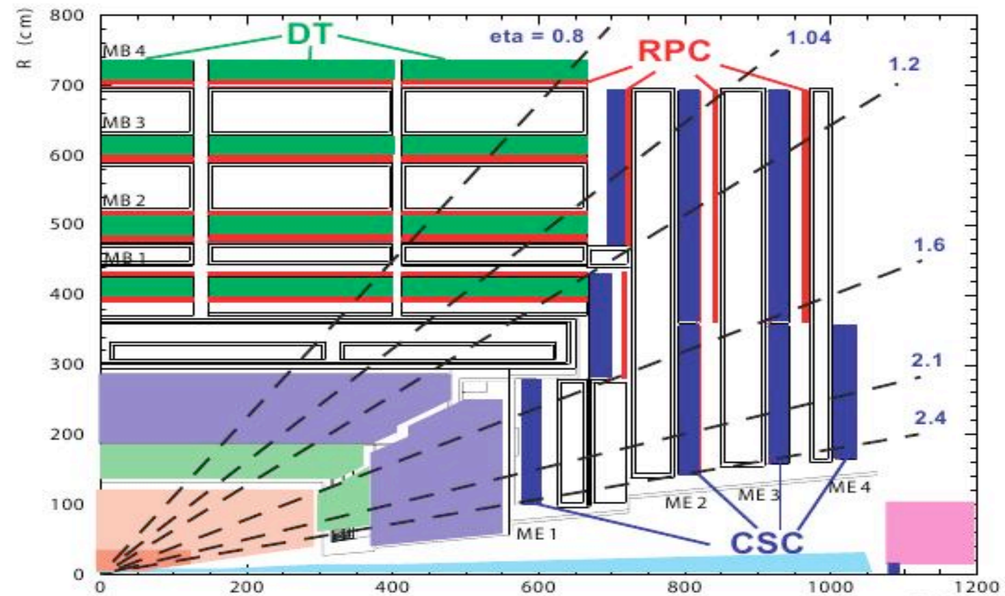
Excellent p_T resolution ≈ 1%

μ candidate: match between muon segments and silicon track

μ reconstruction efficiency ≈99%

CMS-PAS-MUO-10-002





L1

μ p_T> 3 GeV

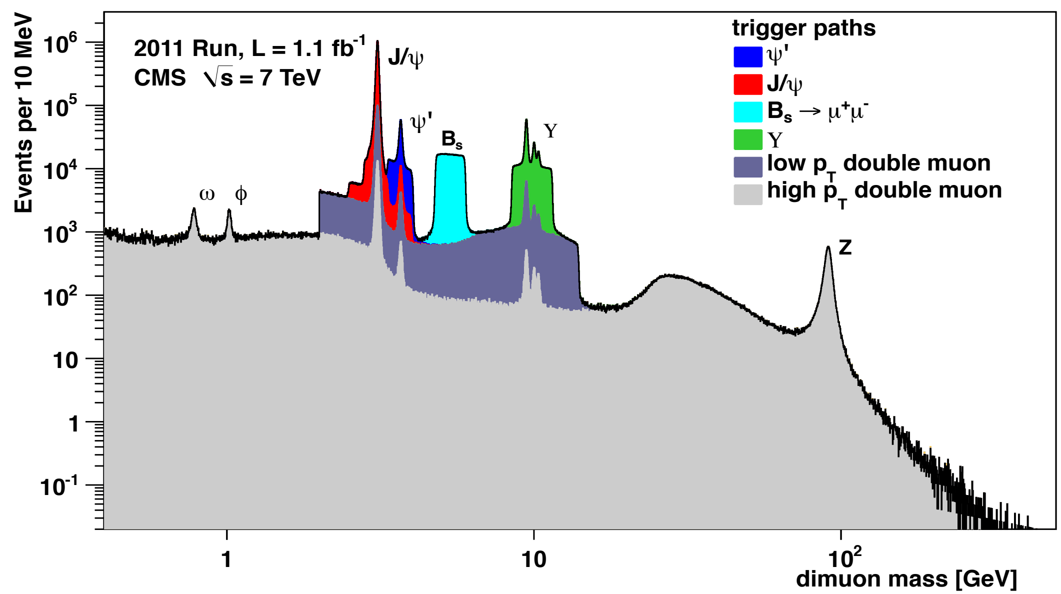
HLT $B_s \to \mu\mu$

μ,μ p_T> 3,4 GeV $|\eta_{\mu\mu}|<1.8$

μ,μ p_T> 4,4 GeV $1.8<|\eta_{\mu\mu}|<2.2$

p_T ($\mu\mu$) > 5 GeV

4.8 < m(μμ) < 6.0 GeV



M. Baarmand
Oct. 10, 2013

Analysis Overview

- Blind analysis of data samples
  - 5 fb^{-1} at $\sqrt{s}$=7 TeV in 2011
  - 20 fb^{-1} at $\sqrt{s}$=8 TeV in 2012

Region definitions	Invariant mass (GeV)
overall window	4.90 mμ1μ2 5.90
blind window	5.20 mμ1μ2 5.45
$B^0 \rightarrow \mu^+\mu^-$ window	5.20 mμ1μ2 5.30
$B_s \rightarrow \mu^+\mu^-$ window	5.30 mμ1μ2 5.45

- Unbinned maximum likelihood fit to $\mu\mu$ mass and discriminant

- Normalization sample: $B^\pm \rightarrow J/\psi K^\pm \rightarrow (\mu^+\mu^-) K^\pm$
  - avoid uncertainties in b production cross section
  - eliminate need for luminosity measurement
  - mitigate effects of uncertainties in efficiencies

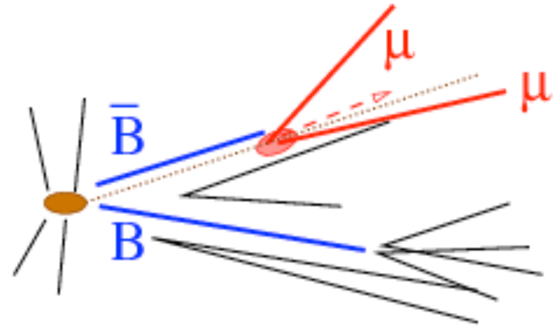
$$Br\left(B_s^0 \rightarrow \mu^+\mu^-\right) = \frac{N_S}{N_{obs}^{B^+}} \frac{f_u}{f_s} \frac{\varepsilon_{tot}^{B^+}}{\varepsilon_{tot}} Br\left(B^+\right)$$

- Control sample: $B_s \rightarrow J/\psi\, \varphi \rightarrow (\mu^+\mu^-)(K^+K^-)$ validate B_s in data and simulations

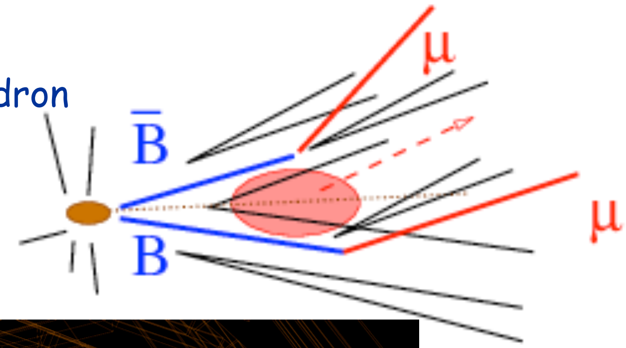
- Divide the data sample in two main categories for each year:
  - $\mu\mu$ in the barrel ($|\eta|<1.4$) ➨better sensitivity, B_s mass resolution ≈ 40 MeV
  - ≥1 μ in the endcap ➨more events but B_s mass resolution ≈ 60 MeV

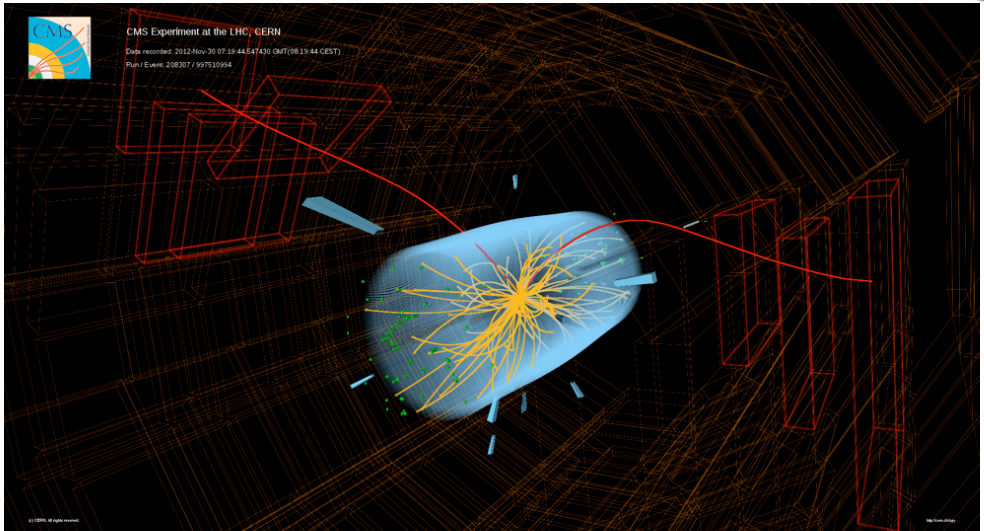
Event Characteristics

- Signal $B_{s/d} \to \mu^+\mu^-$:
  - two reconstructed muons
  - invariant mass around $B_{s/d}$ mass
  - long lived B: well reconstructed secondary vertex and momentum aligned with flight direction

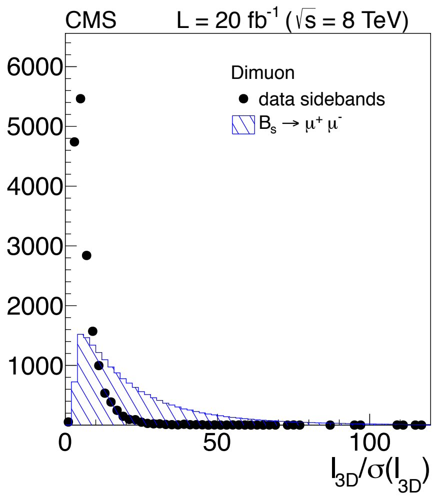


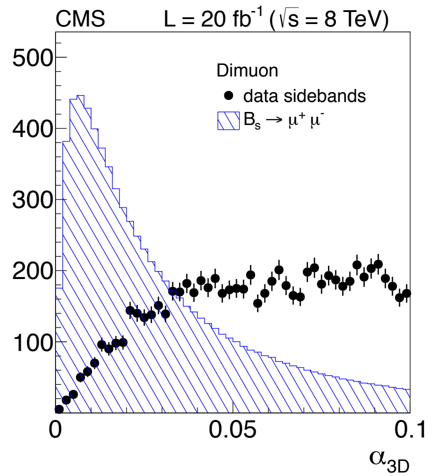
- Backgrounds:
  - two semileptonic B decays
  - one semileptonic B decay and one misidentified hadron
  - Single B decays:
    - peaking (ex. $B_s \to K^- K^+$)
    - rare semileptonic (ex. $\Lambda_b \to p\mu\nu$)

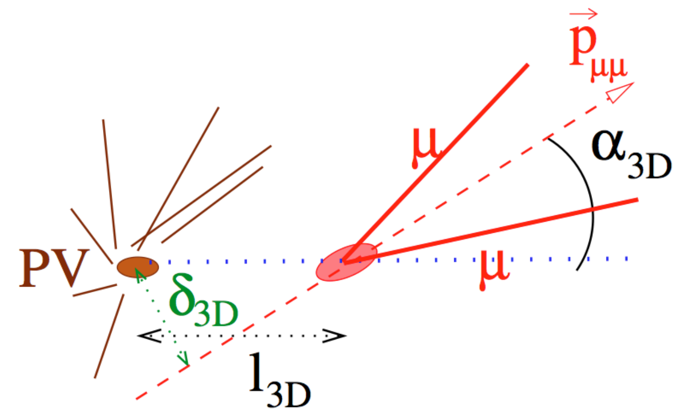


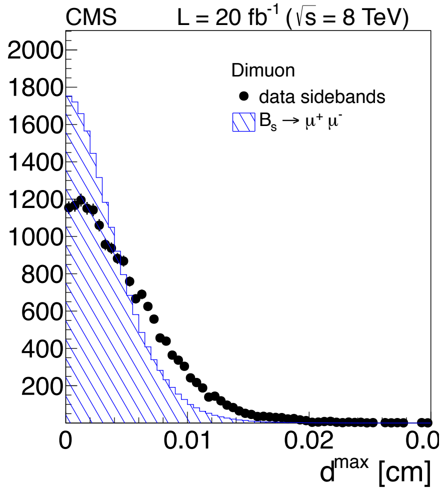


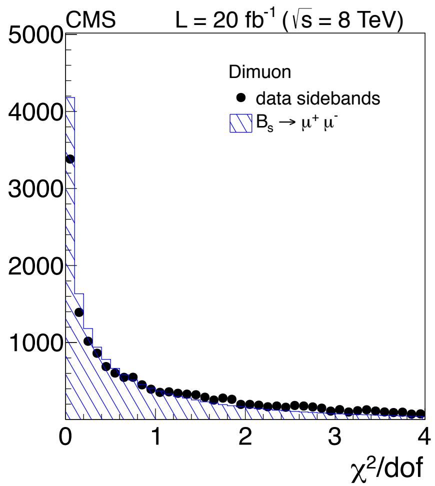
Vertex Variables

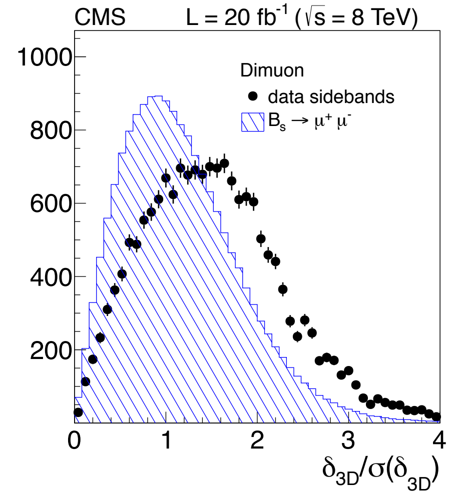


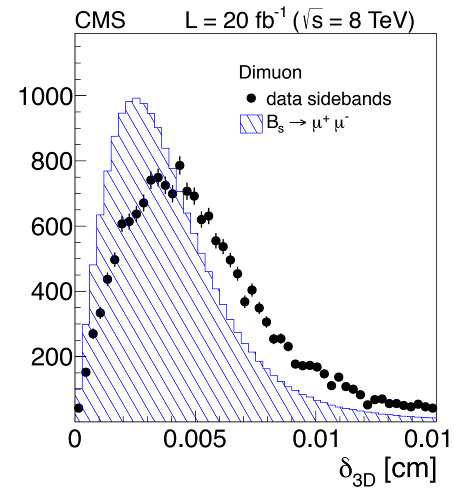










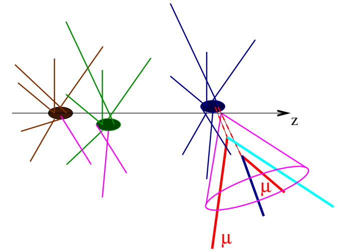


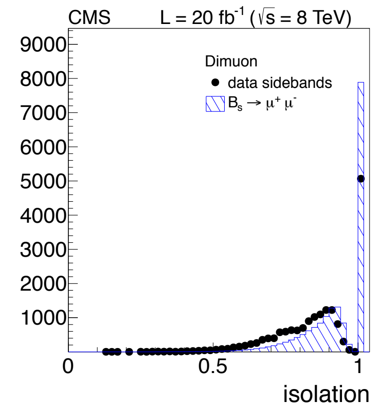
M. Baarmand
Oct. 10, 2013

Event Selection - Isolation

- Primary vertex isolation: relative $\mu^+\mu^-$ isolation

$$I = \frac{p_T(B)}{p_T(B) + \sum_{\Delta R<0.7, p_T>0.9GeV} p_T}$$

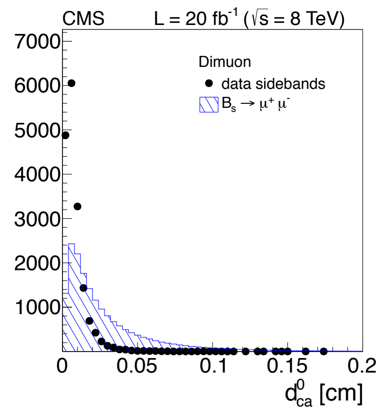


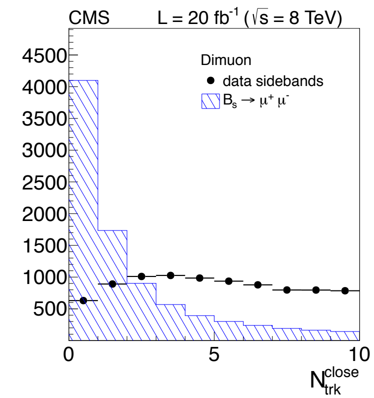


- B vertex isolation

Either tracks not associated to any P.V.
or tracks associated to same B candidate

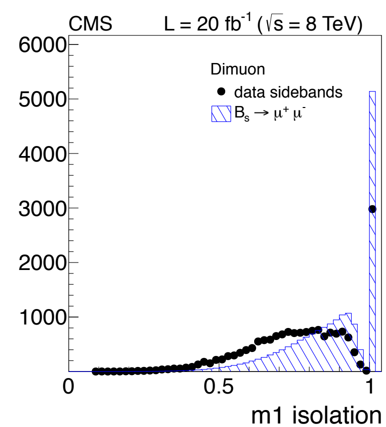
Distance of closest track to SV (d_{ca})
Number of close tracks in dca < 300 μm
and pT > 0.5 GeV





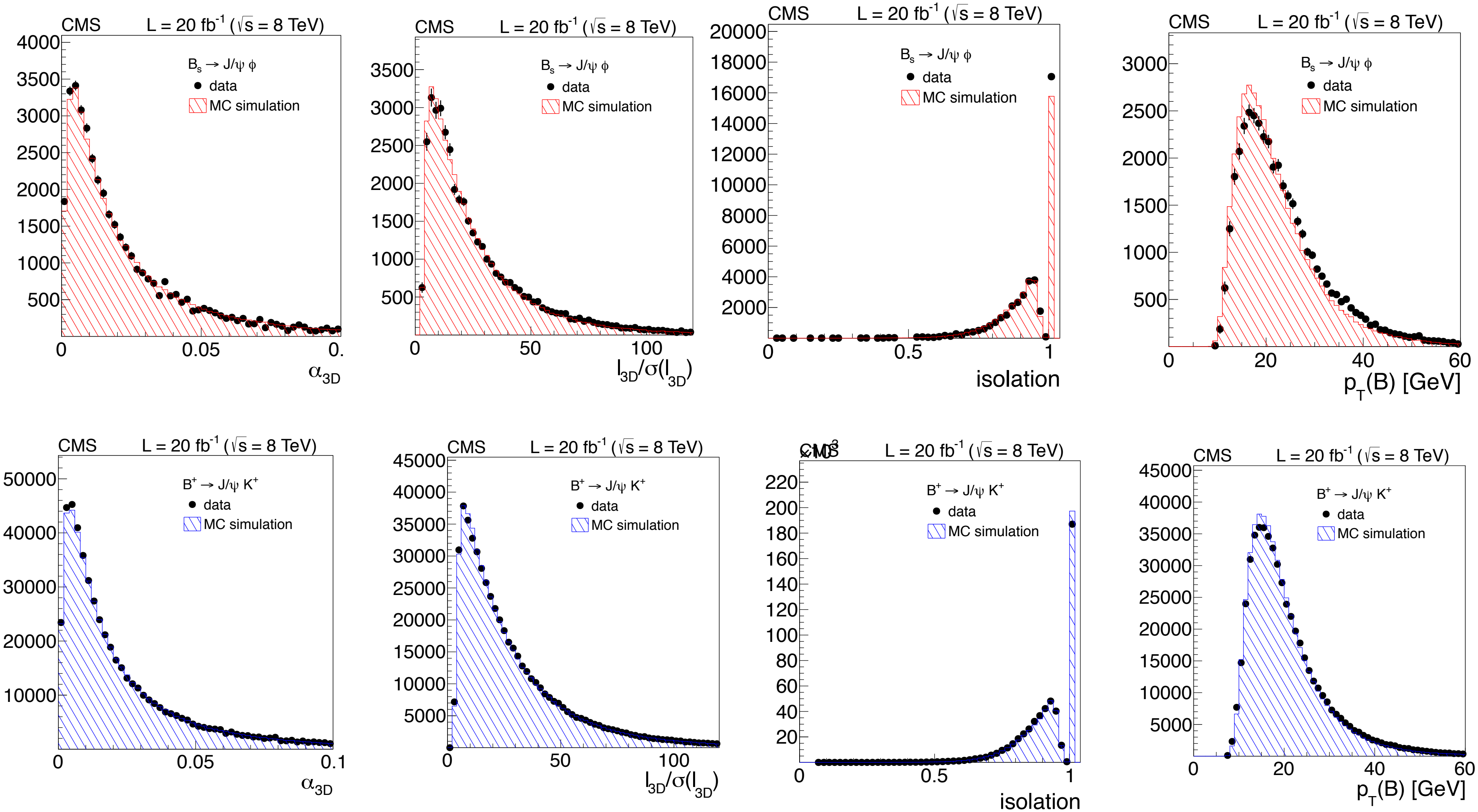
- Muon isolation

$$I_\mu = \frac{p_\mu}{p_\mu + \sum_{\Delta R<0.5; p_T>0.5GeV} p}$$



Data-MC Comparison

- Good agreement between sideband-subtracted distributions and MC



Boosted Decision Tree Selection

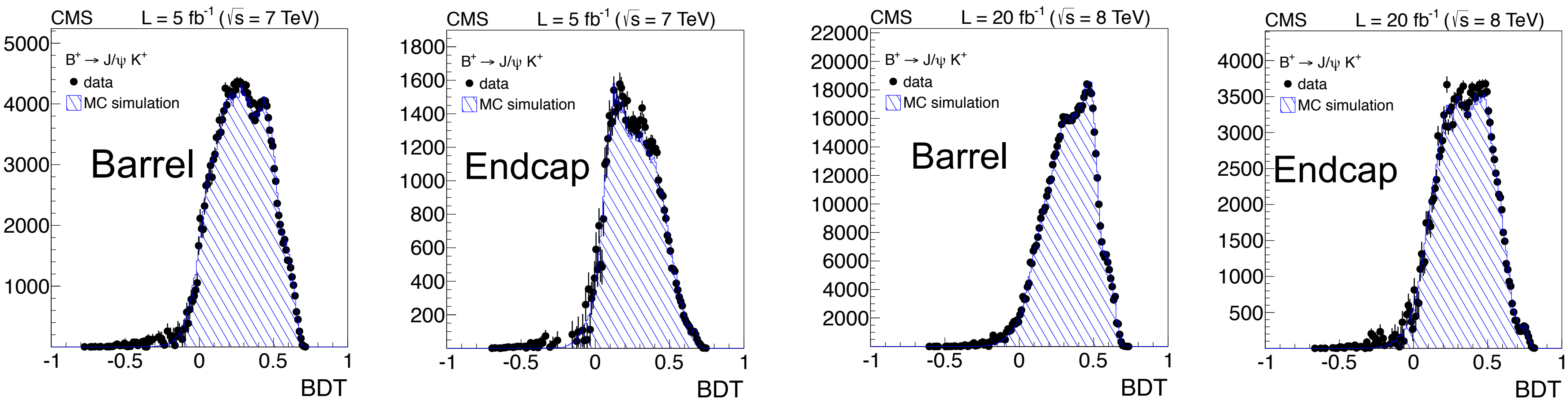
- BDT training – TMVA framework
  - MultiVariate Analysis: involves observation and analysis of more than one statistical outcome variable at a time, the technique is used to perform studies across multiple dimensions while taking into account the effects of all variables
  - signal: Bs MC simulation
  - background: dimuon data sidebands
  - to avoid bias, a given BDT used for training on "1st" event, tested on 2nd" and applied on "3rd", and then rotate

- Checks and studies
  - BDT output insensitive to mass using MC signal with shifted mass
  - BDT output shows no difference for high- and low-mass sidebands
  - BDT output insensitive to pileup

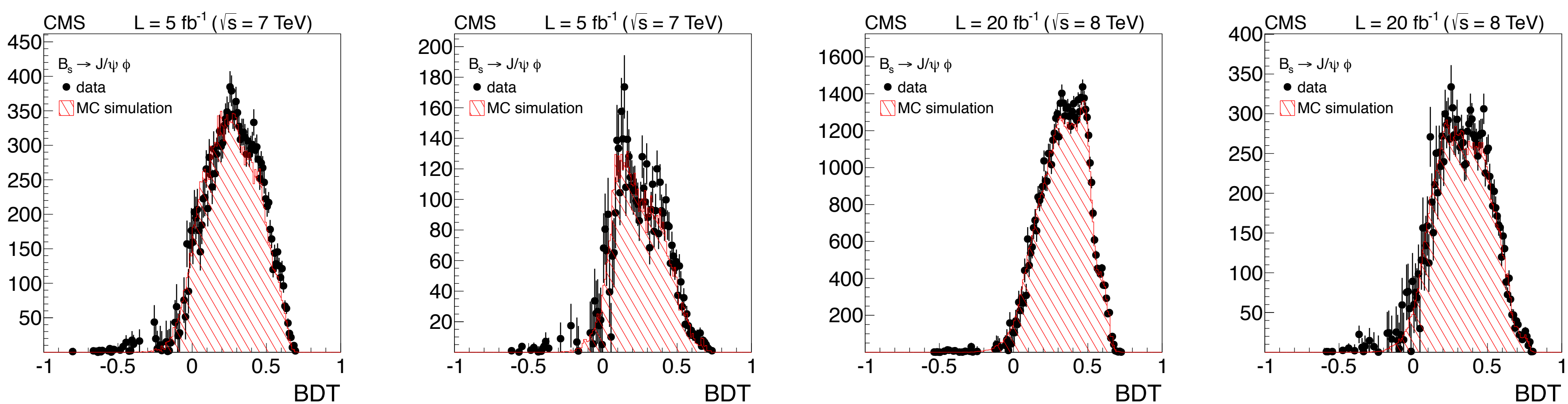
- Use the same BDT for normalization (J/ψ K^+) and control (J/ψ φ) samples

Simulation vs Data

$B^{\pm} \rightarrow J/\psi\, K^{\pm}$: difference ➔ 3% systematic error



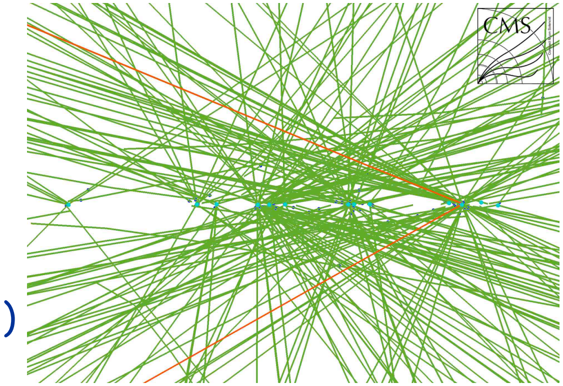
$B_s \rightarrow J/\psi\, \varphi$: 9.5% (2011) and 3.5% (2012)

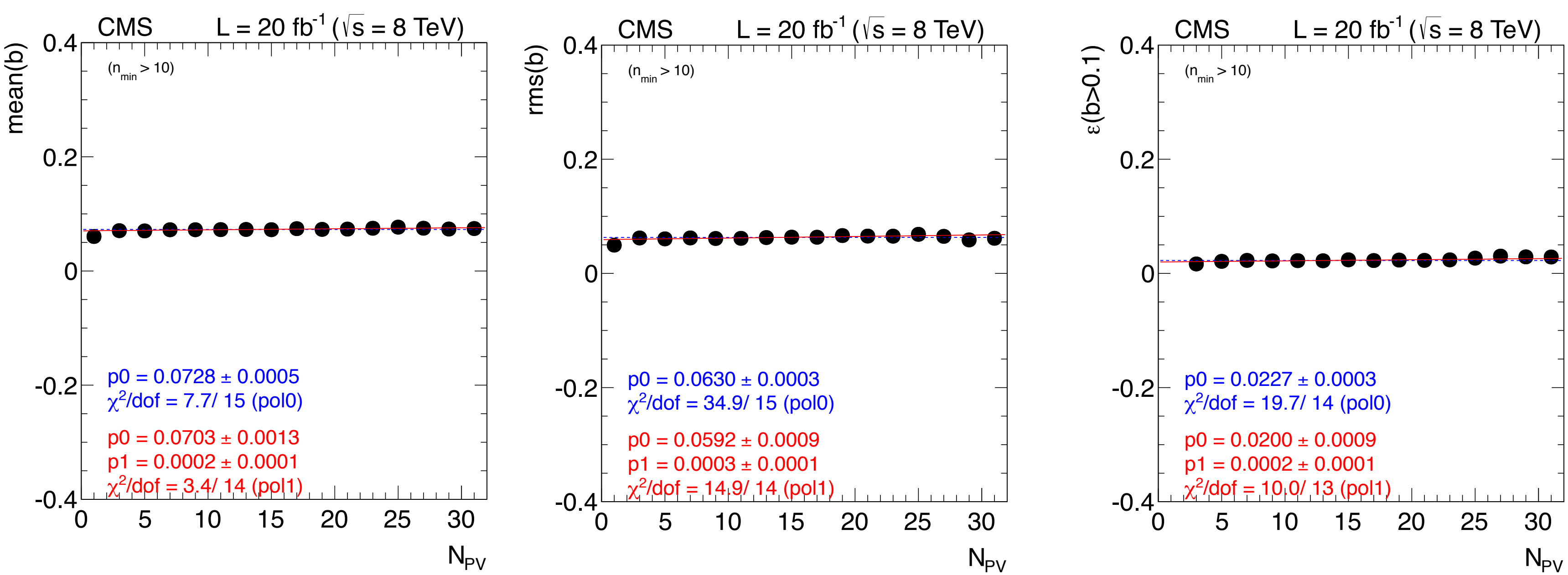


M. Baarmand
Oct. 10, 2013

Multiple pp interactions - pileup

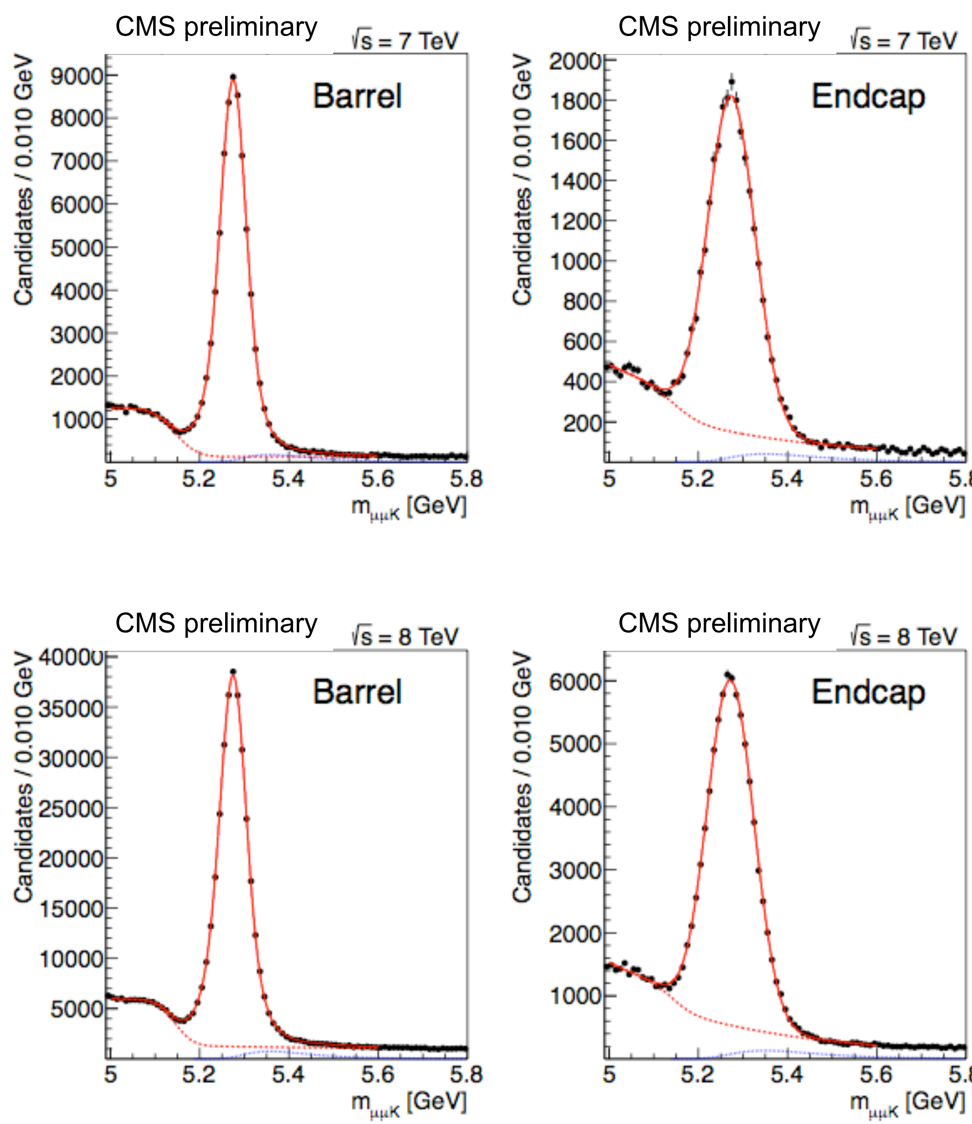
- Number of PU~ 9 (2011) and ~21 (2012)
- event selection tuned to be pileup independent
  - e.g. isolation searches for tracks coming from the same primary vertex or not associated with any
- input variables insensitive to the number of primary vertices
- selection compatible with constant efficiency up to 30 PV (~40 PU)





Normalization Channel: $B^{\pm} \rightarrow J/\psi\, K^{\pm}$

- Same selections as for signal, plus
  - $3.0 < m(\mu\mu) < 3.2$ GeV
  - $p_T(\mu\mu) > 7$ GeV
  - $p_T(K) > 0.5$ GeV
  - all tracks used in vertexing
- Yield extraction
  - signal: double (single) Gaussian in barrel (endcap)
  - background: Error function for $B_d \rightarrow J/\psi\, K^{*} \rightarrow \mu+\mu\text{-}K\text{-}(\pi+)$ decays
  - background: Landau function for $B^{\pm} \rightarrow J/\psi \pi^{\pm}$ decays
- Estimated systematic error on the event yield: 5%

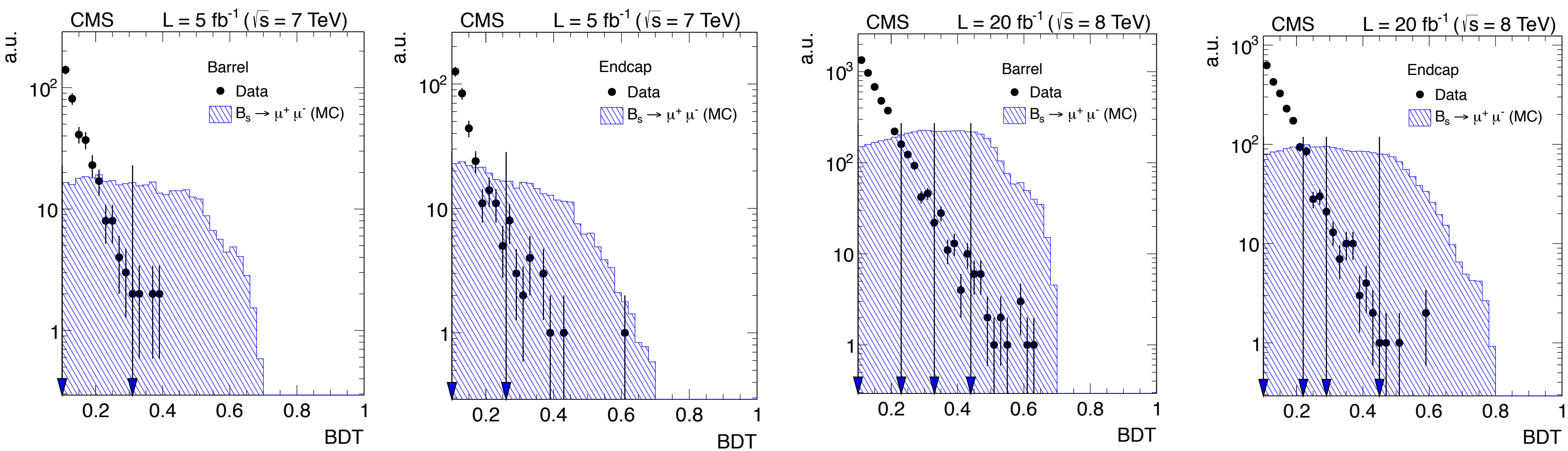


Determination of Branching Fraction

- UML fit to 12 mass distributions in BDT bins split in Barrel/Endcap:

min. bin edges	1	2	3	4
2011 barrel	0.10	0.31	-	-
2011 endcap	0.10	0.29	-	-
2012 barrel	0.10	0.23	0.33	0.44
2012 endcap	0.10	0.22	0.29	0.45

  - BDT binning chosen to equalize the expected number of signal events
  - systematic errors: branching fractions and f_s/f_u



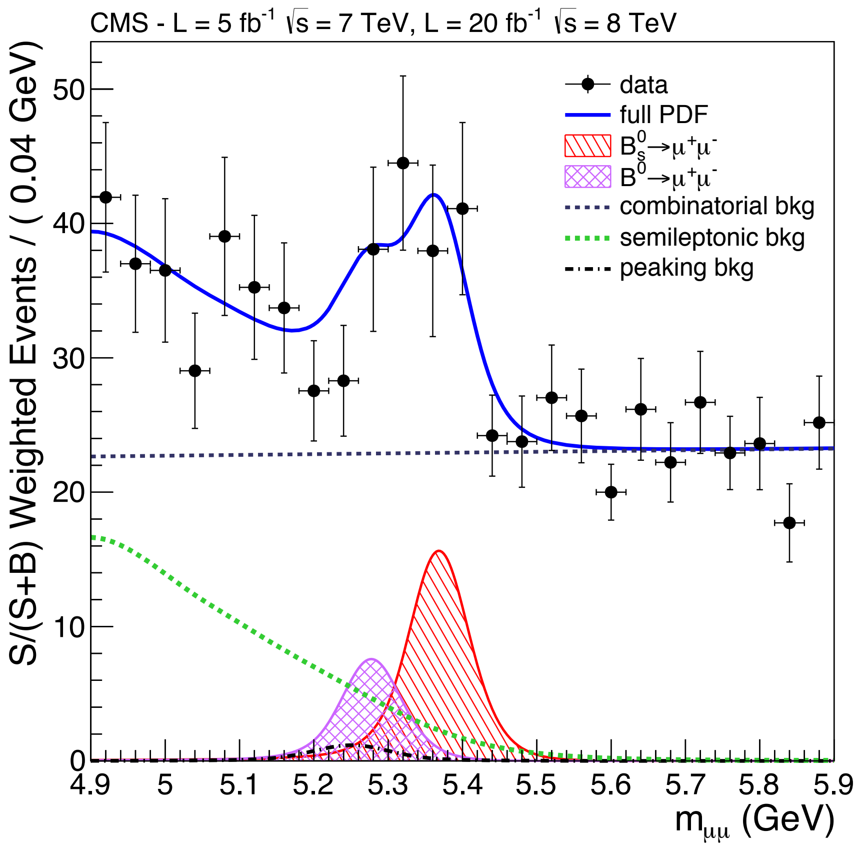
- To extract CL_S limits on $BR(B_d \rightarrow \mu\mu)$ use 1D-BDT
- Optimized cut on BDT output and event counting in mass windows

$b >$	barrel	endcap
2011	0.29	0.29
2012	0.38	0.39

M. Baarmand
Oct. 10, 2013

Unbinned Maximum Likelihood Fit

- Fit B_s and B_d simultaneously
- Signals: Crystal Ball, normalization floating
- Peaking background:
  - sum of Gaussian and Crystal Ball (same mean)
  - constrained (Log-Normal) to expectation and normalized to the measured B^+ yield
  - yield cross checked on independent dataset
- Rare semileptonic background:
  - fixed shape, normalization floating constrained within 75% of nominal value
  - constrained Gaussian kernels from MC
- Combinatorial background:
  - first degree polynomial
  - validated with independent data set
- Per-event mass resolution included



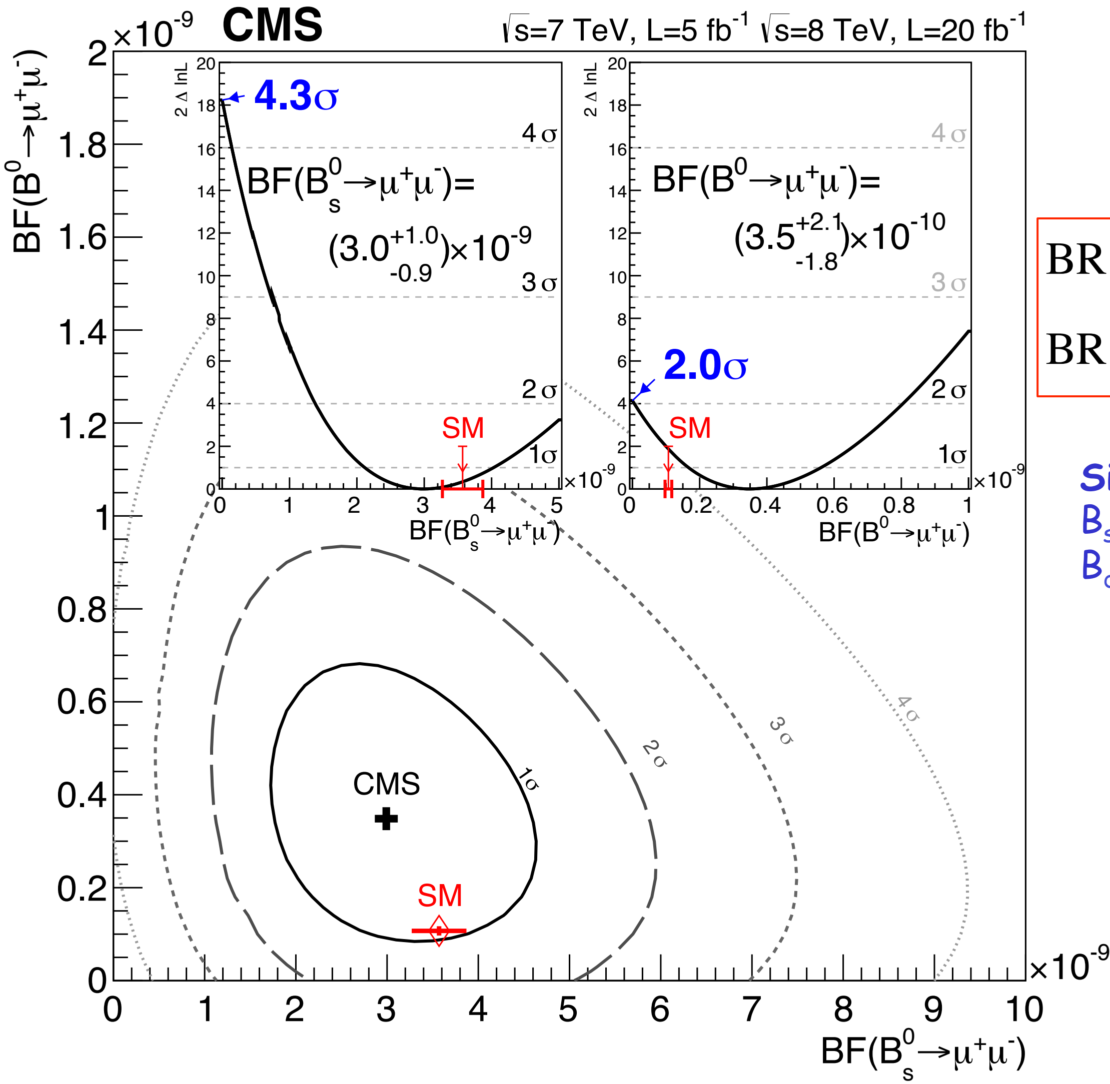
$$\mathrm{BR}(B_S \to \mu\mu) = \frac{N_S^i}{N_{B^+}^i} \times \frac{f_u}{f_s} \times \left(\frac{\varepsilon_s^i}{\varepsilon_u^i}\right) \times \mathrm{BR}(B_d \to J/\psi K^\pm) \times \mathrm{BR}(J/\psi \to \mu\mu)$$

$$\mathrm{BR}(B_d \to \mu\mu) = \frac{N_d^i}{N_{B^+}^i} \times \left(\frac{\varepsilon_s^i}{\varepsilon_u^i}\right) \times \mathrm{BR}(B_d \to J/\psi K^\pm) \times \mathrm{BR}(J/\psi \to \mu\mu)$$

Systematic Errors

- Implemented as Gaussian pdf constraints in the UML fit
  - hadron to muon misidentification probability
    - studied with $D^* \to D^0 \pi$, $D^0 \to K\pi$, $K_S \to \pi\pi$, $\Lambda \to p\pi$
    - 50% uncertainty, conservatively assumed to be uncorrelated
  - BR uncertainties
    - dominated by $\Lambda_b \to p\mu\nu$ (6.5×10^{-4}) with 100% uncertainty
  - $f_s/f_u = 0.256 \pm 0.020$ from LHCb
    - additional 5% to account for possible p_T and η dependence
    - in situ studies show no p_T dependence from ratios of $B^\pm \to J/\psi K^\pm$ vs $B_s \to J/\psi\, \varphi$
  - Normalization channel
    - yields 5%
    - $\mathrm{BR}(B_d \to J/\psi K^\pm) \times \mathrm{BR}(J/\psi \to \mu\mu) = (6.0 \pm 0.2) \times 10^{-5}$

Branching Fraction Results



$$BR(B_S \to \mu\mu) = \left(3.0^{+0.9}_{-0.8}\ (\text{stat})^{+0.6}_{-0.4}\ (\text{syst})\right) \times 10^{-9}$$

$$BR(B_d \to \mu\mu) = \left(3.5^{+2.1}_{-1.8}\ (\text{stat+syst})\right) \times 10^{-10}$$

Significances

$B_s \to \mu\mu$ 4.3 σ (expected 4.8 σ)

$B_d \to \mu\mu$ 2.0 σ

Upper Limits on $B_d \to \mu\mu$

- No significant excess is observed for $B_d \to \mu\mu$
  - Upper limit computed using CL_s method, based on observed events in the signal and sideband regions with the 1D-BDT method

Expected and observed events in signal regions

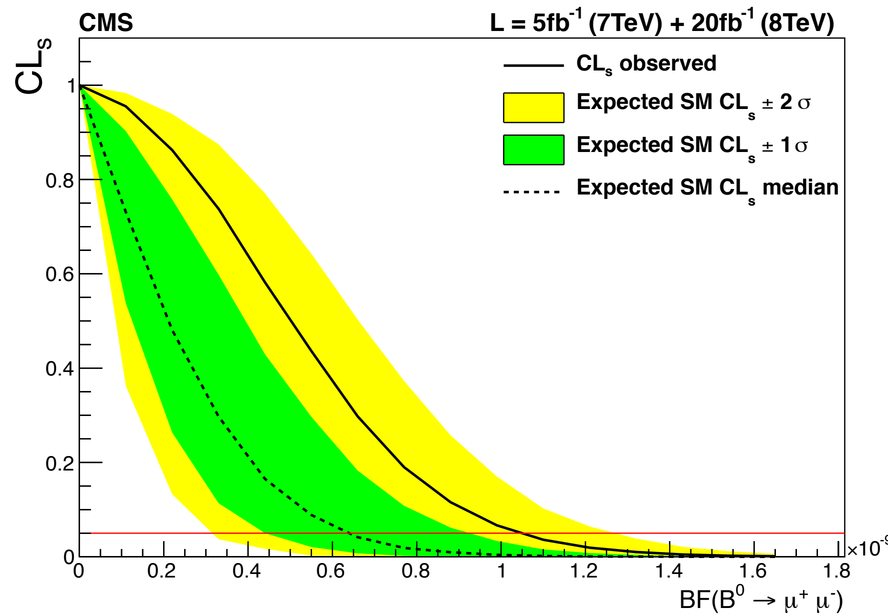
	2011 barrel		2012 barrel	
	$B^0 \to \mu^+\mu^-$	$B^0_s \to \mu^+\mu^-$	$B^0 \to \mu^+\mu^-$	$B^0_s \to \mu^+\mu^-$
$\varepsilon_{tot}[\%]$	0.33 ± 0.03	0.30 ± 0.04	0.24 ± 0.02	0.23 ± 0.03
N^{exp}_{signal}	0.27 ± 0.03	2.97 ± 0.44	1.00 ± 0.10	11.46 ± 1.72
N^{exp}_{total}	1.3 ± 0.8	3.6 ± 0.6	7.9 ± 3.0	17.9 ± 2.8
N_{obs}	3	4	11	16

	2011 endcap		2012 endcap	
	$B^0 \to \mu^+\mu^-$	$B^0_s \to \mu^+\mu^-$	$B^0 \to \mu^+\mu^-$	$B^0_s \to \mu^+\mu^-$
$\varepsilon_{tot}[\%]$	0.20 ± 0.02	0.20 ± 0.02	0.10 ± 0.01	0.09 ± 0.01
N^{exp}_{signal}	0.11 ± 0.01	1.28 ± 0.19	0.30 ± 0.03	3.56 ± 0.53
N^{exp}_{total}	1.5 ± 0.6	2.6 ± 0.5	2.2 ± 0.8	5.1 ± 0.7
N_{obs}	1	4	3	4

$$\mathrm{BR}(B_d \to \mu\mu) < 1.1 \times 10^{-9} \ @95\% \ \mathrm{CL}$$

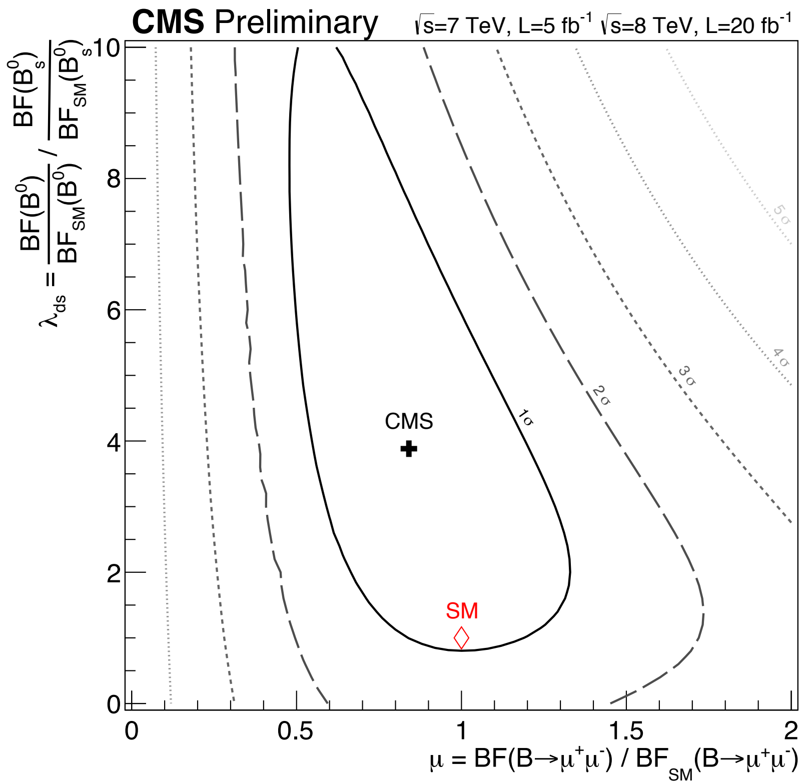
(expected 6.3×10^{-10} in presence of SM+background)

$$\mathrm{BR}(B_d \to \mu\mu) < 9.2 \times 10^{-10} \ @90\% \ \mathrm{CL}$$



SM Compatibility Check

- $BR_{SM}(B_s \to \mu\mu) = (3.56 \pm 0.18) \times 10^{-9}$
- $BR_{SM}(B_d \to \mu\mu) = (1.07 \pm 0.10) \times 10^{-10}$



$$\mu = \frac{\mathrm{BR}(B_S \to \mu\mu)}{\mathrm{BR}_{SM}(B_S \to \mu\mu)}$$

$$\lambda_{ds} = \frac{\mathrm{BR}(B_d \to \mu\mu)}{\mathrm{BR}_{SM}(B_d \to \mu\mu)} \Big/ \frac{\mathrm{BR}(B_S \to \mu\mu)}{\mathrm{BR}_{SM}(B_S \to \mu\mu)}$$

Simultaneous fit

$$\mu = 0.84^{+0.31}_{-0.25};\ \lambda_{ds} = 3.9^{+3.7}_{-2.2}$$

Fit for μ (fix Λ_{ds} to SM)

$$\mu = 1.01^{+0.31}_{-0.26}$$

Fit for Λ_{ds} (fix μ to SM)

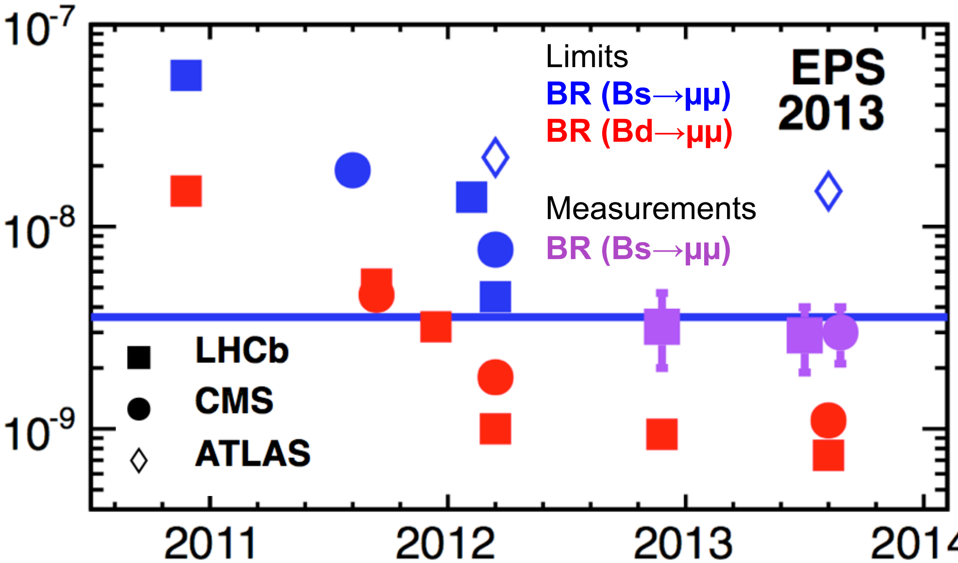
$$\lambda_{ds} = 3.1^{+2.0}_{-1.7}$$

Summary

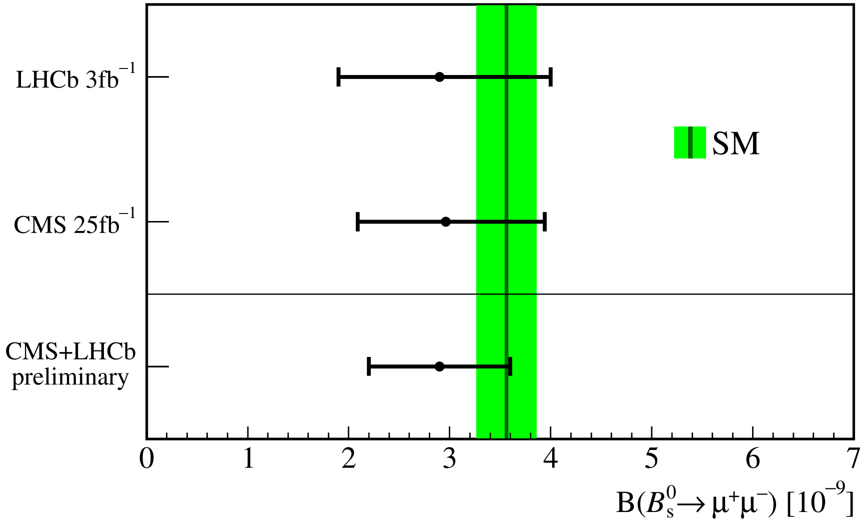
- CMS measurements using 25 fb^{-1} data

$$BR(B_S \to \mu\mu) = \left(3.0^{+1.0}_{-0.9}\right) \times 10^{-9}$$ statistics dominated, 4.3 σ significance

$$BR(B_d \to \mu\mu) < 1.1 \times 10^{-9} \; @95\% \; CL$$



- LHCb + CMS combined measurement: accounting for correlations



$BR^{<t>}(B_s \to \mu^+\mu^-) = (2.9 \pm 0.7) \times 10^{-9}$

Consistent with SM expectations!

Is SM a Complete Theory?!

- SM successfully explains and predicts many observed phenomena
  - its predictions verified at 10^{-3} level up to TeV energies

- SM pasting together of strong and electroweak interactions

  SM = SU(3) x SU(2) x U(1)

  - 19 (26 with $m_\nu \neq 0$) arbitrary parameters that can only be determined from experimental measurements
  - Unanswered questions such as
    - origin of flavor
    - number of generations
    - fermion masses ($m_t/m_u \sim 3\times10^4$; $m_\tau/m_e \sim 4\times10^3$)
    - matter – antimatter asymmetry
    - dark matter / dark energy
  - Lack of grand unification of fundamental forces

- SM merely an effective (low energy) theory valid up to some scale, where new physics appears!

Beyond the Standard Model

Supersymmetry:

Extension of Poincare group to include boson-fermion symmetry

New mirror spectrum of particles

Large number of new parameters (105 in minimal SUSY SM)

Theoretically nice:

- additional particles cancel divergences in m_H - can naturally be of order EW scale
- SUSY closely approximates SM at low energies
- allows unification of forces at higher energies
- provides a path to incorporation of gravity and string theory
- lightest neutralino is a cosmic dark matter candidate

Extra dimensions:

Large-scale compactification of extra dimensions

String theory motivated but with observed effects at EW scale O(TeV)

Theoretically nice:

- solves hierarchy problem by reducing GUT scale
- gravity may propagate in 4+n dimensions, would see effects only at very small distances, perhaps reachable in pp LHC Collisions e.g. Kaluza-Klein gravitons and Z-like particles

Conclusions

- The SM passes yet another test – still victorious!
- But its demise will come – it is logically incomplete
- The defeater will have to address

    number of quark/lepton generations, origin of flavor

    matter – antimatter asymmetry

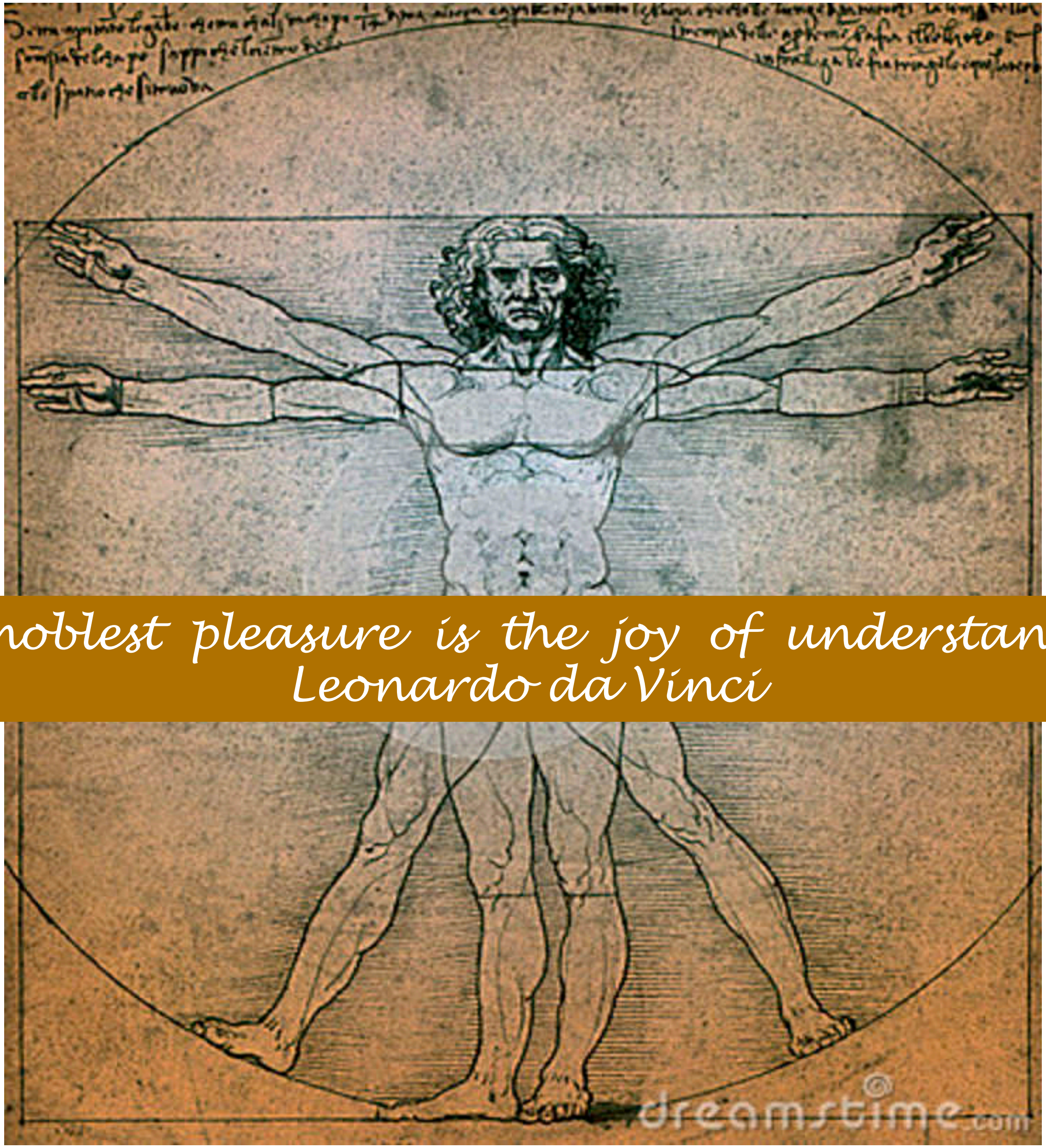
    unification of forces, including gravity

    dark matter / dark energy, etc.

Flammarion engraving
unknown artist, circa 1800

The Quest for Knowledge and Understanding Continues!

M. Baarmand
Oct. 10, 2013



"The noblest pleasure is the joy of understanding."
Leonardo da Vinci

Vitruvian Man
Leonardo da Vinci
circa 1490

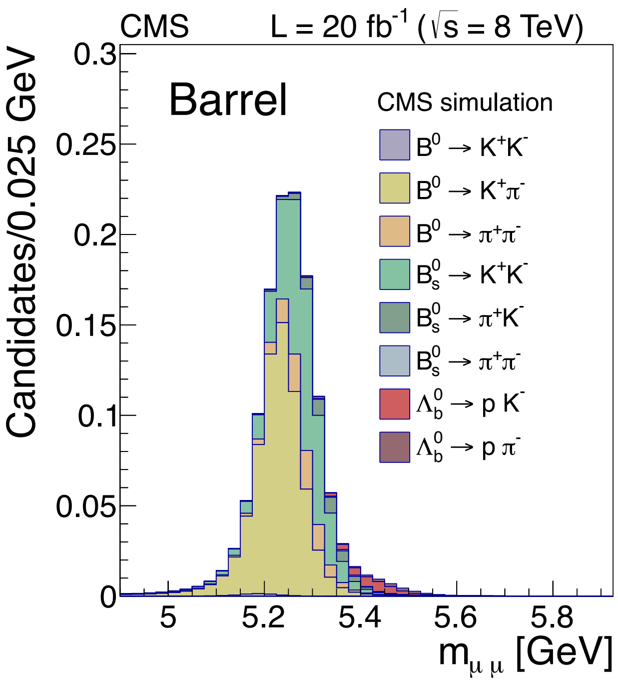
Backup Slides

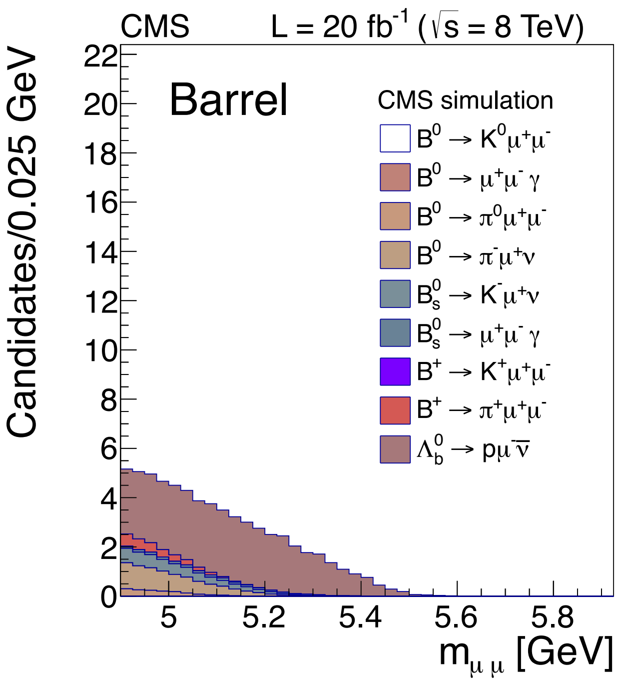
Rare Backgrounds

- Expected number of events in each channel normalized to $B^{\pm}$ in data:

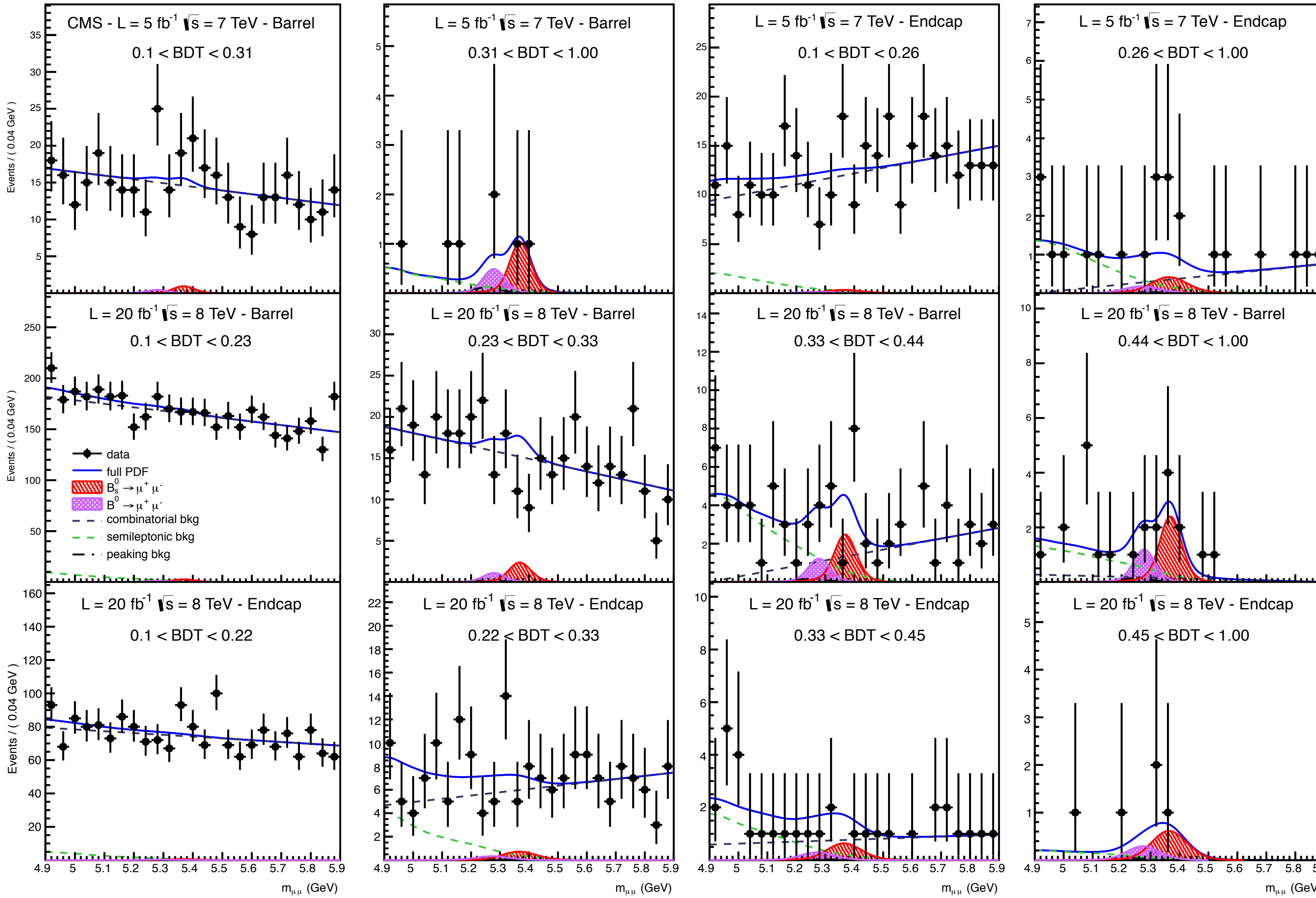
$$N(X) = \frac{Br(Y \to X)}{Br(B^{\pm} \to J/\psi K^{\pm})} \frac{f_Y}{f_u} \frac{\varepsilon_{tot}(X)}{\varepsilon_{tot}(B^{\pm})} N_{obs}(B^{\pm})$$

  - weighted with muon-misid evaluated from data
  - systematic errors: branching fractions and f_s/f_u





Categorized-BDT Fits Results



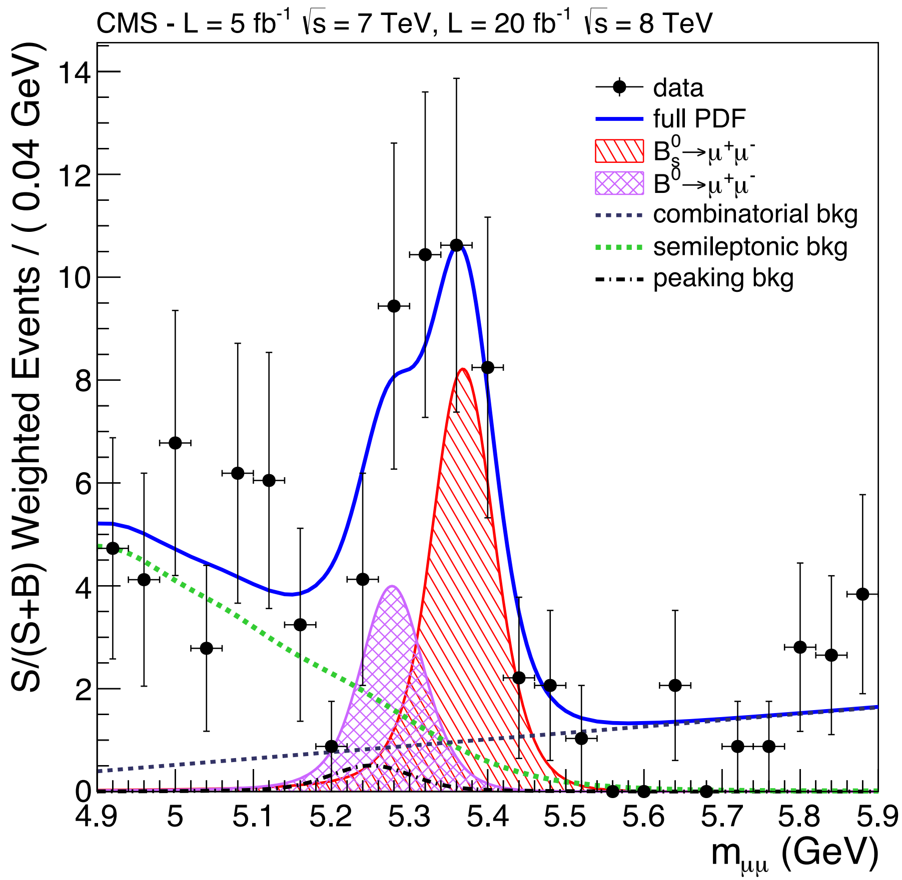
M. Baarmand
Oct. 10, 2013

1D BDT Results - Cross Check

- Significance

  $B_s \rightarrow \mu\mu$ 4.8 σ (expected 4.7 σ)

- Less sensitive wrt categorized-BDT
- used as a cross check



M. Baarmand
Oct. 10, 2013