



CMS Experiment at the LHC, CERN

Data recorded: 2012-Nov-30 07:19:44.547430 GMT (08:19:44 CEST)  
Run / Event: 208307 / 997510994

# Observation of the $B_s \rightarrow \mu^+ \mu^-$ at the LHC: CMS results

<http://arxiv.org/pdf/1307.5025.pdf>

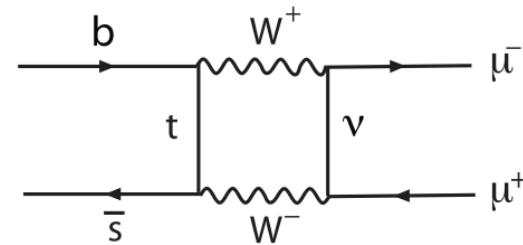
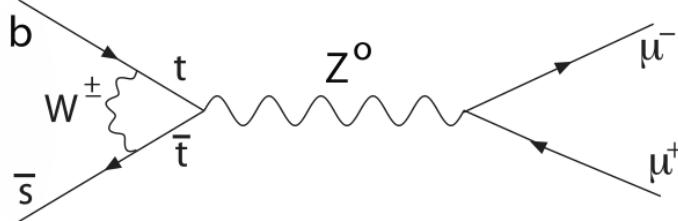
CMS-BPH-13-004-003

CERN-PH-EP-2013-129

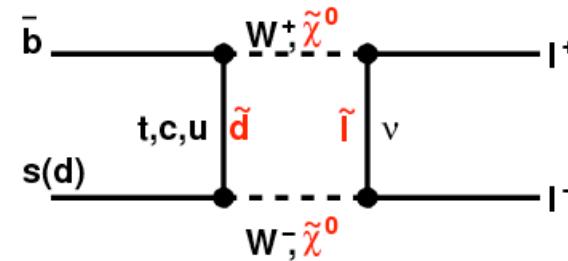
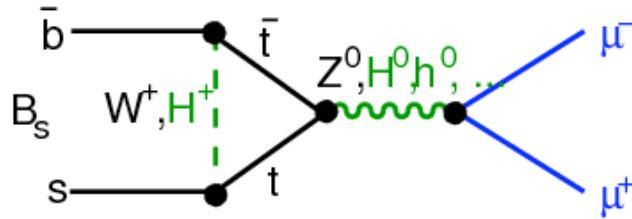
*Fabrizio Palla (INFN Pisa)*  
*on behalf of the CMS Collaboration*  
*LPCC seminar - Aug. 6<sup>th</sup> 2013 - CERN*

## In Standard Model decays highly suppressed

- forbidden at tree level, can only proceed through higher-order loop diagrams
- helicity suppressed
- Cabibbo enhancement of  $B_s \rightarrow \mu\mu$  over  $B_d \rightarrow \mu\mu$  due to  $|V_{td}| < |V_{ts}|$



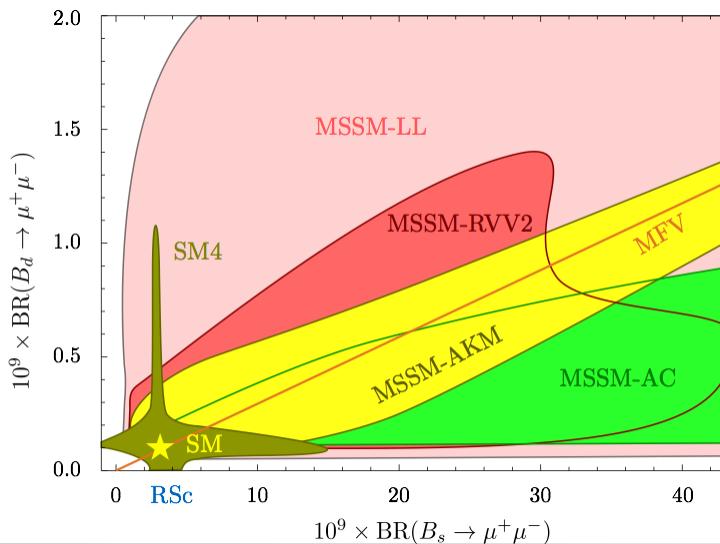
- A few examples of interfering New Physics contributions





$B_{s/d} \rightarrow \mu\mu$  are potentially sensitive probes for Physics Beyond SM, a few examples:

- 2HDM:  $\text{BR}(B_{s/d} \rightarrow \mu\mu) \propto \tan^4 \beta$  and  $m(H^+)$ 
  - ➔ J. R. Ellis et al, JHEP 05 (2006) 063
- MSSM:  $\text{BR}(B_{s/d} \rightarrow \mu\mu) \propto \tan^6 \beta$ 
  - ➔ 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).



D. M. Straub  
<http://arxiv.org/pdf/1012.3893v2.pdf>

# Theory prediction



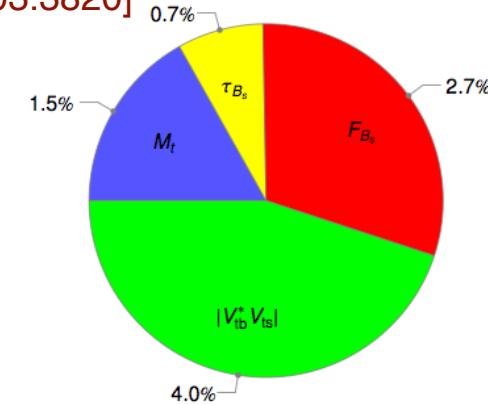
Standard Model prediction for CP averaged branching fractions at t=0

$\text{BR}^{t=0}(B_s \rightarrow \mu\mu) = (3.25 \pm 0.17) \times 10^{-9}$

[A. Buras et al. arXiv:1303.3820]

$\text{BR}^{t=0}(B_d \rightarrow \mu\mu) = (1.07 \pm 0.10) \times 10^{-10}$

$$\text{BR}(B_s \rightarrow \mu^+ \mu^-)_{\text{SM}} = 3.25 \times 10^{-9} \left( \frac{M_t}{173.2 \text{ GeV}} \right)^{3.07} \left( \frac{F_{B_s}}{225 \text{ MeV}} \right)^2 \left( \frac{\tau_{B_s}}{1.500 \text{ ps}} \right) \left| \frac{V_{tb}^* V_{ts}}{0.0405} \right|^2$$



Experimental values are time integrated

[De Bruyn et al. (PRL 109, 041801)]

[A. Buras et al. arXiv:1303.3820]

$\text{BR}^{<\leftrightarrow>} = \left[ \frac{1 - y_s^2}{1 + \mathcal{A}_{\Delta\Gamma}^{\mu\mu} y_s} \right] \text{BR}^{t=0}(B_s \rightarrow \mu\mu) = (3.56 \pm 0.18) \times 10^{-9}$

$$y_s \equiv \frac{\Gamma_L^{(s)} - \Gamma_H^{(s)}}{\Gamma_L^{(s)} + \Gamma_H^{(s)}} = \frac{\Delta\Gamma_s}{2\Gamma_s}$$

$$\begin{aligned} \tau_{B_s} &= 1.516 \pm 0.011 \text{ ps} \\ \Delta\Gamma_s / 2\Gamma_s &= 0.0615 \pm 0.0085 \end{aligned}$$

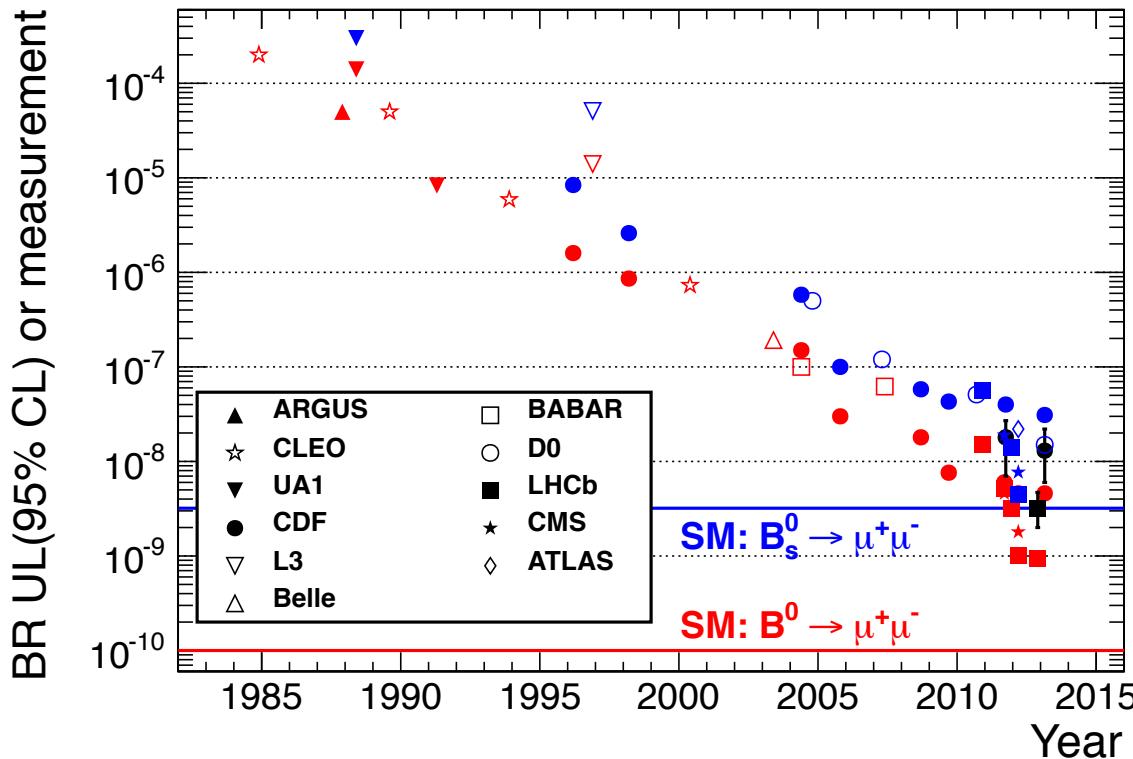
More than 25 years quest



Previous CMS publications

EPS 2011: PRL, 107, 191802

Moriond 2012: JHEP, 1204, 033



Best until recent:

$$\text{BR}(B_d \rightarrow \mu\mu) < 8.4 \times 10^{-10} \text{ @ 95% CL}$$

ATLAS+CMS+LHCb best upper limit on  $B_d \rightarrow \mu\mu$

## CMS Detector

Pixels  
Tracker  
**ECAL**  
**HCAL**  
Solenoid  
Steel Yoke  
Muons

**STEEL RETURN YOKE**  
~13000 tonnes

**SUPERCONDUCTING SOLENOID**  
Niobium-titanium coil carrying ~18000 A

**HADRON CALORIMETER (HCAL)**  
Brass + plastic scintillator  
~7k channels

Total weight : 14000 tonnes  
Overall diameter : 15.0 m  
Overall length : 28.7 m  
Magnetic field : 3.8 T

**SILICON TRACKER**  
Pixels (100 x 150  $\mu\text{m}^2$ )  
~1m<sup>2</sup> ~66M channels  
Microstrips (80-180  $\mu\text{m}$ )  
~200m<sup>2</sup> ~9.6M channels

**CRYSTAL ELECTROMAGNETIC CALORIMETER (ECAL)**  
~76k scintillating PbWO<sub>4</sub> crystals

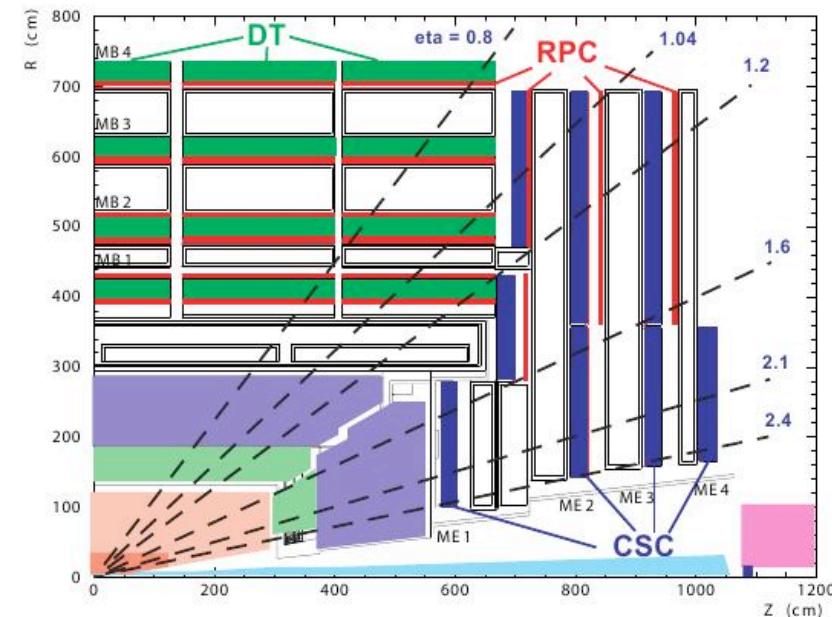
**PRESHOWER**  
Silicon strips  
~16m<sup>2</sup> ~137k channels

**FORWARD CALORIMETER**  
Steel + quartz fibres  
~2k channels

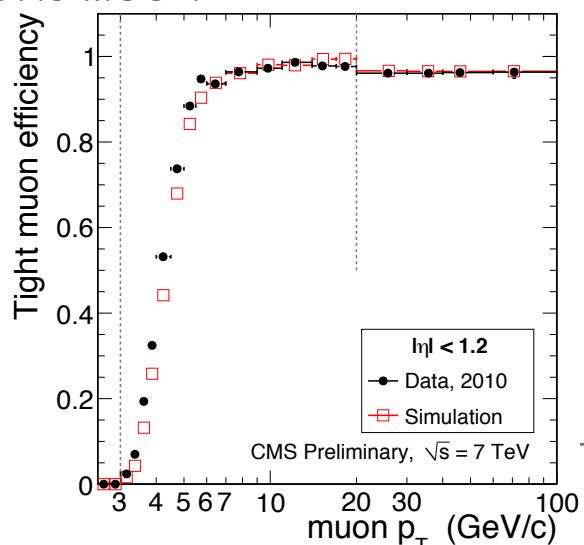
**MUON CHAMBERS**  
Barrel: 250 Drift Tube & 480 Resistive Plate Chambers  
Endcaps: 468 Cathode Strip & 432 Resistive Plate Chambers

# Muon reconstruction

- Tracks: Excellent  $p_T$  resolution  $\approx 1\%$ 
  - Tracking efficiency  $> 99\%$  for central muons
  - Excellent vertex reconstruction and impact parameter resolution ( $\approx 15 \mu\text{m}$ )
- Muon candidates:
  - Match between muon segments and a silicon track
  - Large pseudorapidity coverage:  $|\eta| < 2.4$
- Muon identification and trigger efficiencies evaluated with
  - MC methods
  - Data-driven methods: Tag & Probe
- Muon misidentification rates measured in data using
  - $D^* \rightarrow D^0 \pi$ ,  $D^0 \rightarrow K\pi$
  - $K_S \rightarrow \pi\pi$
  - $\Lambda \rightarrow p\pi$



CMS-PAS-MUO-10-002





## Dimuon Trigger

### L1 Hardware Trigger

- $p_T > 3$  GeV (few kHz)

### HLT Full tracking and vertexing

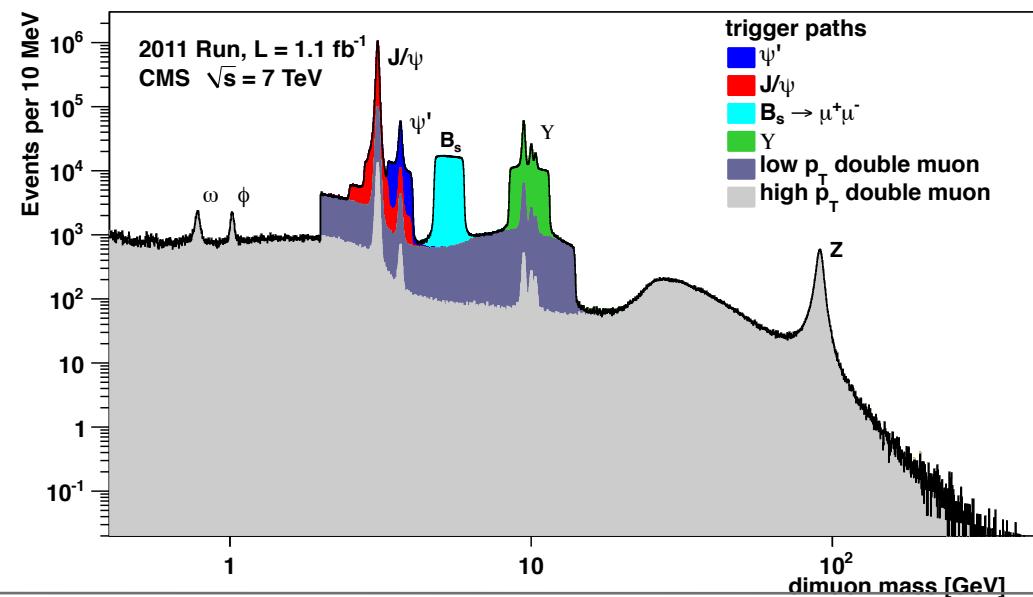
### HLT $B_s \rightarrow \mu\mu$

- Leading and sub-leading  $\mu$   $p_T > 3,4$  (4,4) GeV  $|\eta_{\mu\mu}| < 1.8$  ( $1.8 < |\eta_{\mu\mu}| < 2.2$ )
- $p_T(\mu\mu) > 5$  (4.8-6) GeV
- $4.8 < m(\mu\mu) < 6.0$  GeV
- $P(\chi^2/\text{dof}) > 0.5\%$



### HLT $B^\pm \rightarrow J/\psi K^\pm$ and $B_s \rightarrow J/\psi \phi$

- Two  $\mu$  with opposite charge, each with  $p_T > 4$  GeV,  $|\eta| < 2.2$
- $p_T(\mu\mu) > 6.9$  GeV
- $2.9 < m(\mu\mu) < 3.3$  GeV
- $\cos \alpha > 0.9$
- $L/\sigma(L) > 3$
- distance of closest approach  $< 0.5$  cm
- $P(\chi^2/\text{dof}) > 15\%$



# Analysis overview

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

- Re-blind 2011 data
- Unbinned maximum likelihood fit to dimuon mass and discriminant
- Normalization sample  $B^\pm \rightarrow J/\psi K^\pm \rightarrow (\mu^+ \mu^-) K^\pm$

- avoid uncertainties in the  $b$  production cross section
- eliminate the need for luminosity measurement
- mitigate the effects of uncertainties in efficiencies

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

Region definitions	Invariant mass (GeV)
overall window	$4.90 < m_{\mu_1 \mu_2} < 5.90$
blinding window	$5.20 < m_{\mu_1 \mu_2} < 5.45$
$B^0 \rightarrow \mu^+ \mu^-$ window	$5.20 < m_{\mu_1 \mu_2} < 5.30$
$B_s \rightarrow \mu^+ \mu^-$ window	$5.30 < m_{\mu_1 \mu_2} < 5.45$

- Control sample  $B_s \rightarrow J/\psi \phi \rightarrow (\mu^+ \mu^-)(K^+ K^-)$  to compare and validate  $B_s$  mesons in data and MC simulations
- Divide the sample in two main categories, per each year of data taking:
  - ◆ both muons in the barrel ( $|\eta| < 1.4$ ) ➔ better sensitivity,  $B_s$  mass resolution  $\approx 40 \text{ MeV}$
  - ◆  $\geq 1 \mu$  in the endcap ➔ more events but  $B_s$  mass resolution  $\approx 60 \text{ MeV}$



# What is new



5  $\text{fb}^{-1}$  data at  $\sqrt{s}=7 \text{ TeV}$  in 2011 and 20  $\text{fb}^{-1}$  at  $\sqrt{s}=8 \text{ TeV}$  in 2012

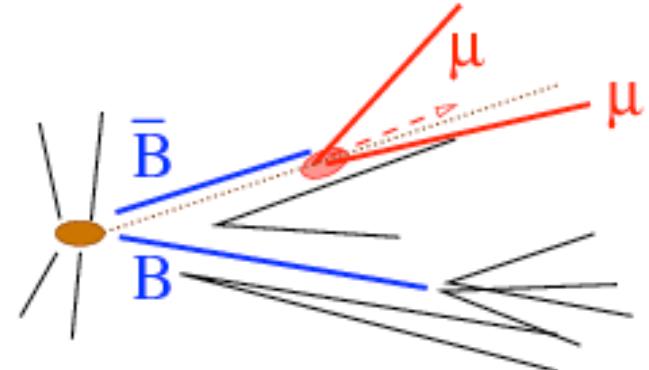
- Better muon identification and fake suppression (MVA based)
- New analysis variables
- MVA based discriminant
- Simultaneous maximum likelihood fit to dimuon mass in several discriminant bins in barrel and endcaps.

# Event characteristics



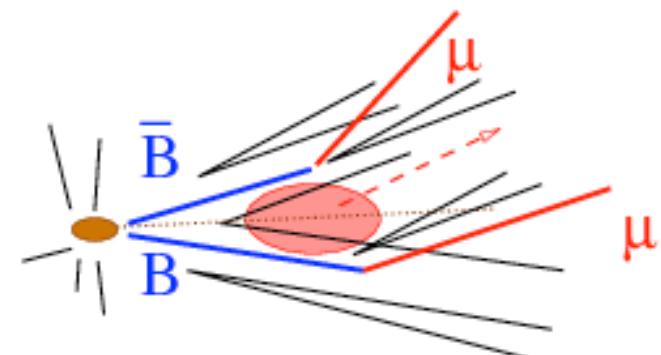
Signal  $B_{s/d} \rightarrow \mu^+ \mu^-$ :

- two reconstructed muons
- invariant mass around  $m(B_{s/d})$
- long lived  $B$ , with a well reconstructed secondary vertex and a momentum aligned with flight direction



## Backgrounds

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

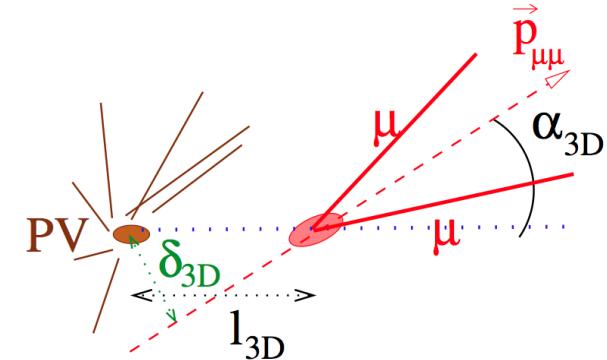
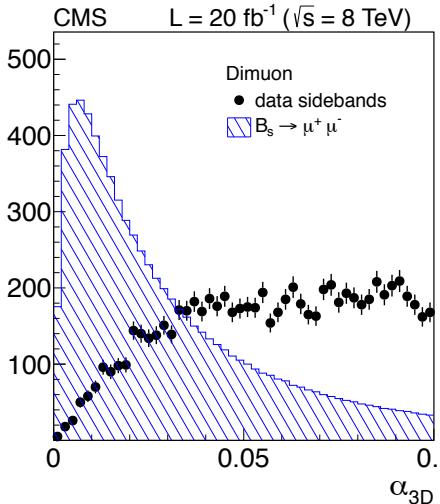
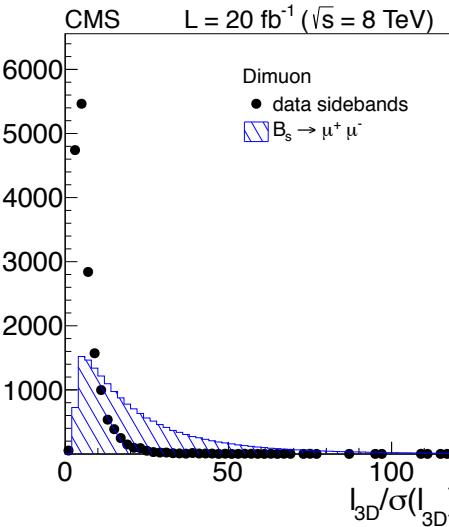




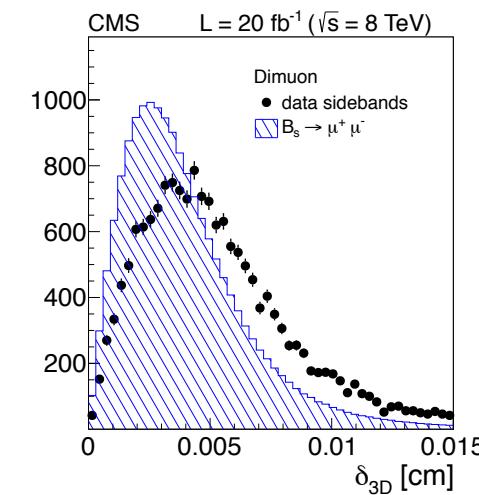
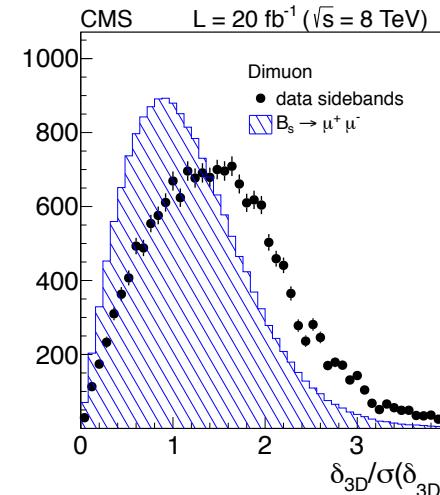
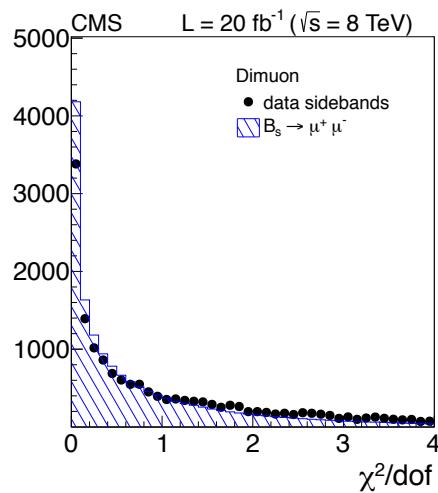
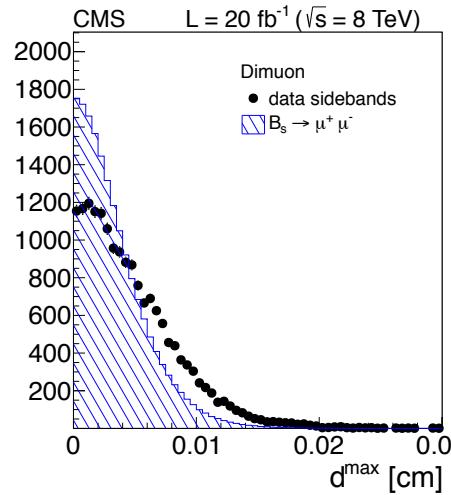
## Improved muon identification to cope with peaking and semileptonic backgrounds

- ➊ Use a MVA based on Silicon Tracker and Muon Detector information
  - distribution of hits in the tracker as compared to expected
  - kink identification
  - muon segment compatibility
  - $\chi^2/\text{ndof}$
  - muon  $p_T$  and  $\eta$
- ➋ Muon MC efficiency validated with Tag&Probe methods in data
- ➌ Muon fake rate reduced by a factor of  $\sim 2$  compared to previous analysis, at the expenses of  $\sim 10\%$  muon efficiency reduction
  - $\epsilon(\mu|\pi) < 0.15\%$
  - $\epsilon(\mu|K) < 0.20\%$
  - $\epsilon(\mu|p) < 0.10\%$
- ➍ Analysis uses MC  $p_T$ -dependent fake rate, validated with data (50% uncertainty)
  - $D^* \rightarrow D^0 \pi$  ( $D^0 \rightarrow K\pi$ ),  $K_S \rightarrow \pi\pi$ ,  $\Lambda \rightarrow p\pi$

# Vertexing variables



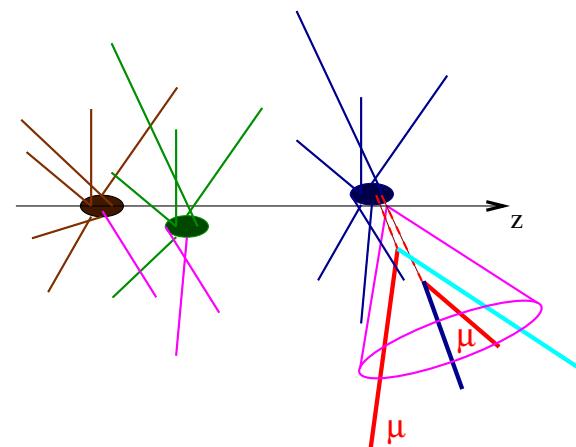
Data side-bands vs signal MC:



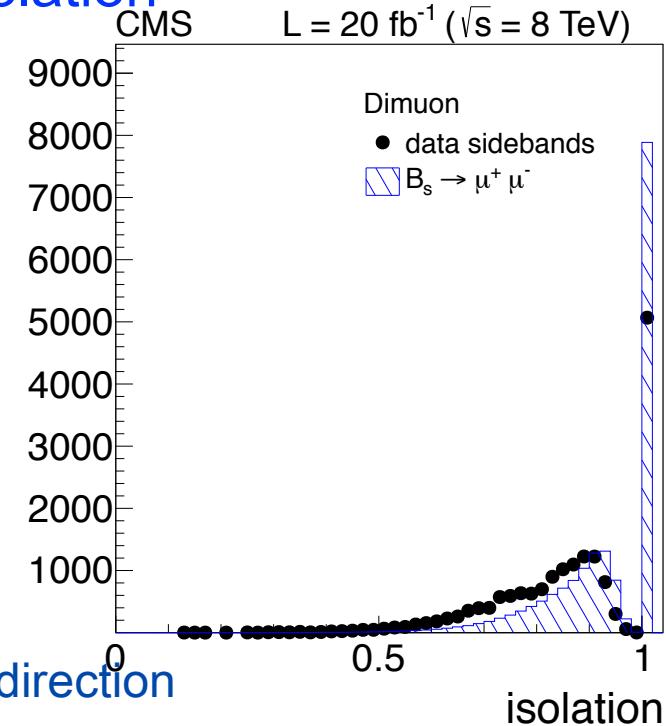
# Isolation (1)



## Primary vertex isolation: relative dimuon isolation



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



All tracks within cone of  $\Delta R=0.7$  around dimuon direction and with  $p_T>0.9 \text{ GeV}$

- either from the same PV
- or  $d_{ca}<0.5 \text{ mm}$  from B vertex.

dip at  $\sim 0.97$  from minimum track  $p_T$  requirement

# Isolation (2)



## B-vertex isolation

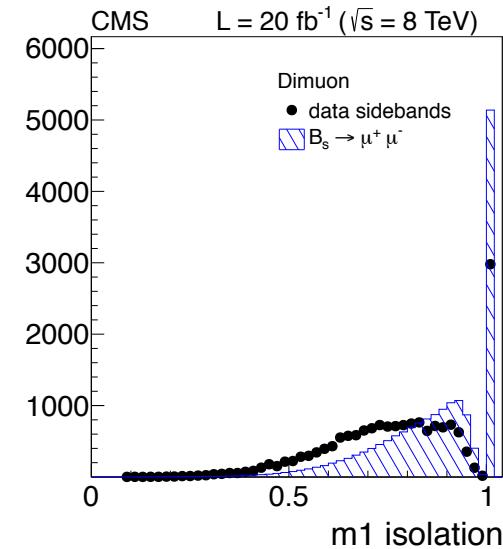
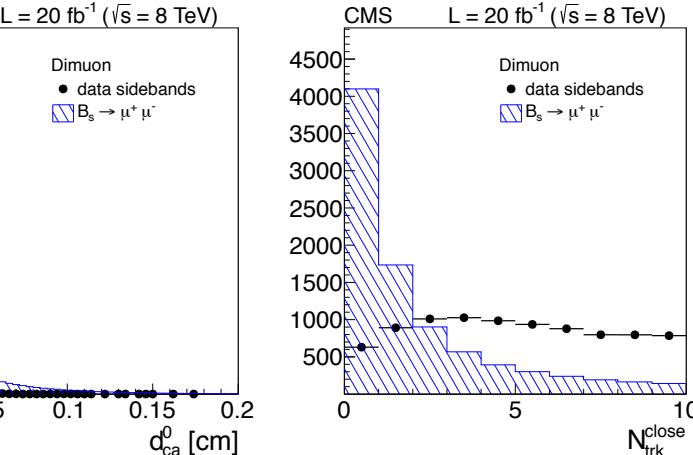
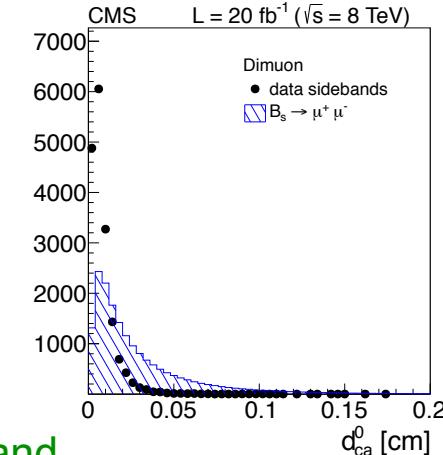
- tracks reconstructed close to the secondary vertex
  - either tracks not associated to any P.V.
  - or tracks associated to the same B candidate
  - Distance of the closest track to SV ( $d_{ca}$ )
  - Number of close tracks in  $d_{ca} < 300 \mu\text{m}$  and  $p_T > 0.5 \text{ GeV}$



## Muon isolation (new!)

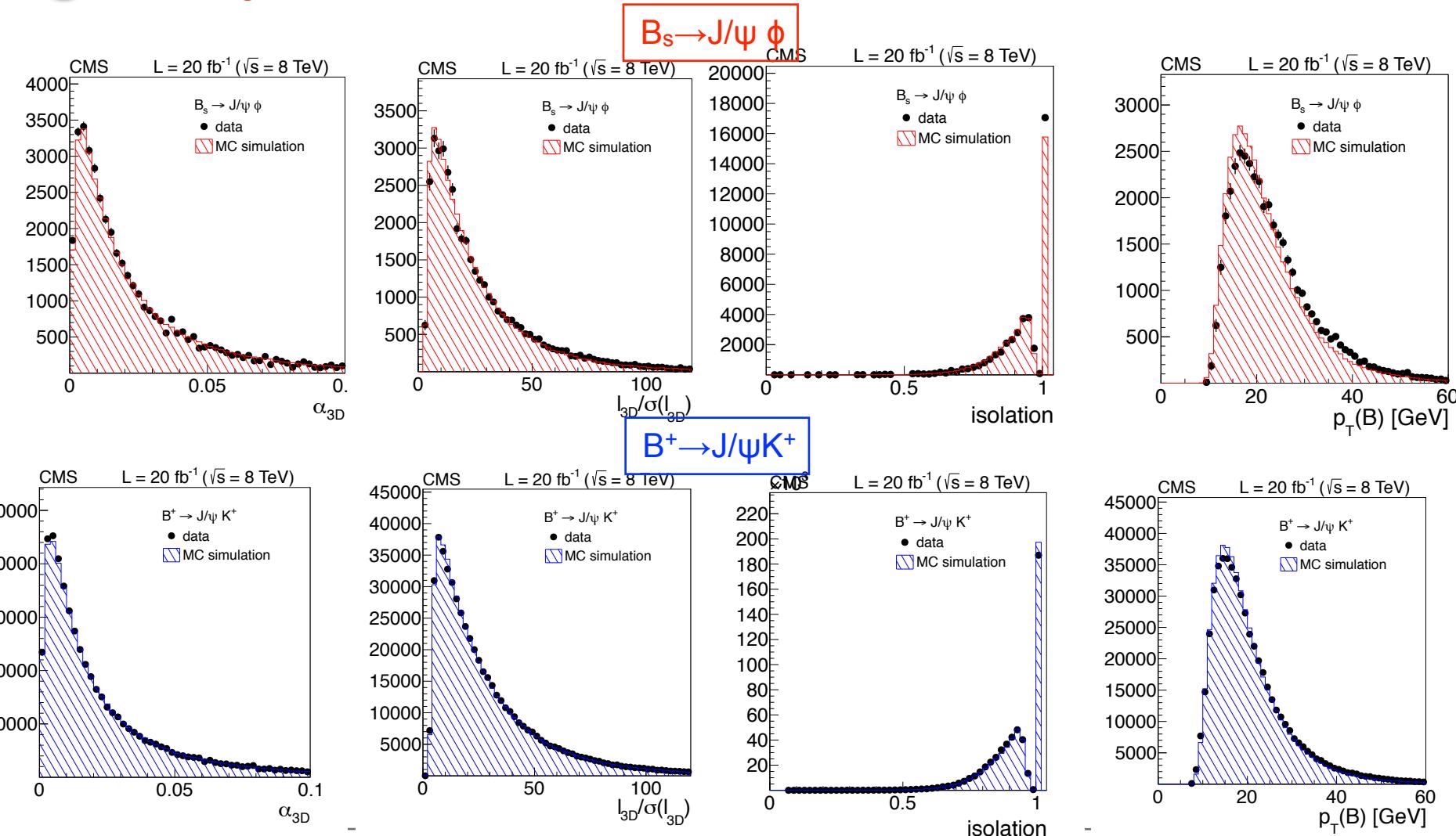
$$I_\mu = \frac{p_\mu}{p_\mu + \sum_{\Delta R < 0.5; p_T > 0.5 \text{ GeV}} p}$$

- tracks in muon cone with  $\Delta R = 0.5$ 
  - $p_T > 0.5 \text{ GeV}$
  - $d_{ca} < 1 \text{ mm}$  from the same PV or not associated to another PV



# Data - Simulation comparison

Good agreement of sideband-subtracted distributions





## BDT Training



### TMVA framework

- signal:  $B_s$  MC simulation
- background: dimuon data sidebands



In order to avoid any bias and use three different BDTs:

- A given BDT was used for training on “1<sup>st</sup>” event, tested on “2<sup>nd</sup>” and applied on “3<sup>rd</sup>”, and then rotate



## 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 ( $J/\psi K^+$ ) and control ( $J/\psi \phi$ ) samples

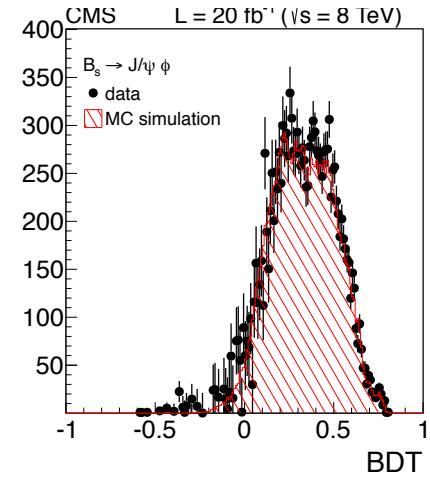
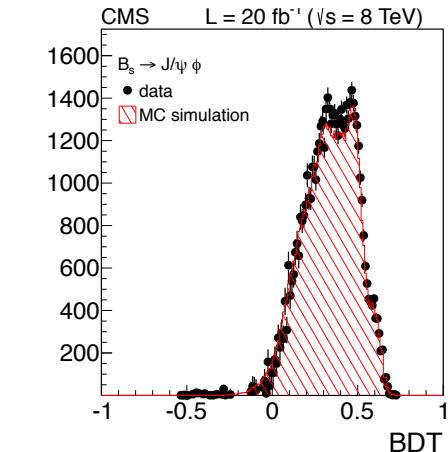
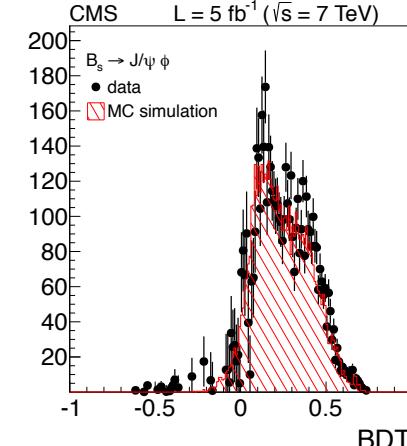
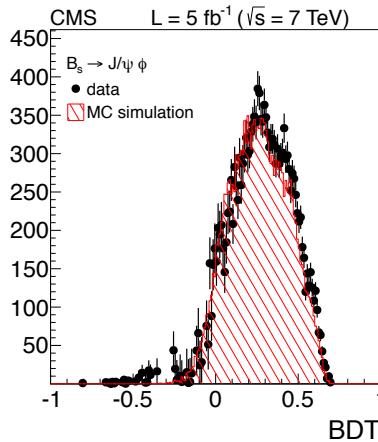
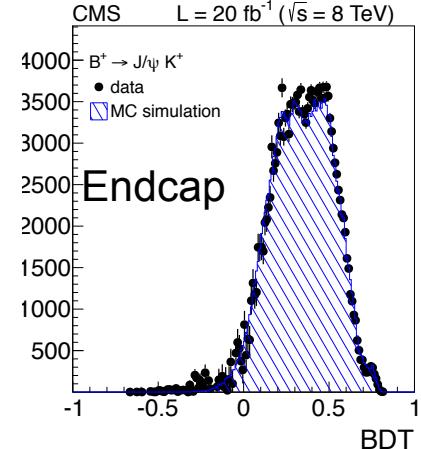
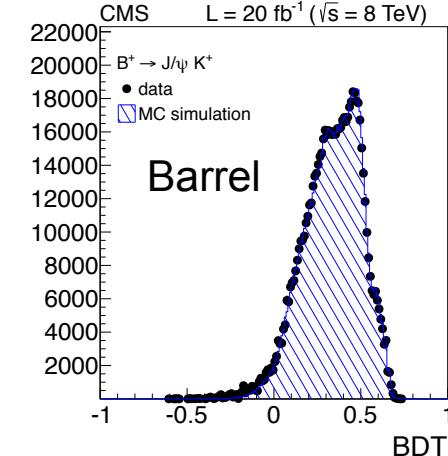
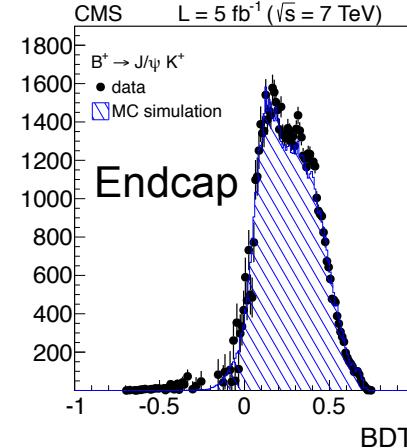
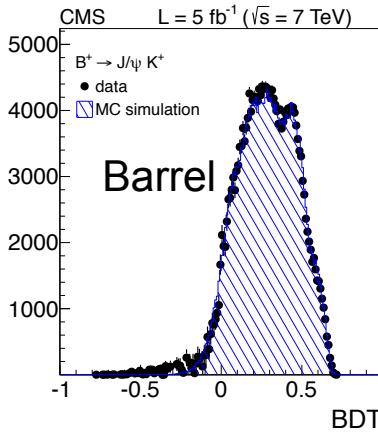
# MC vs data comparison



Use differences between data and MC for systematics

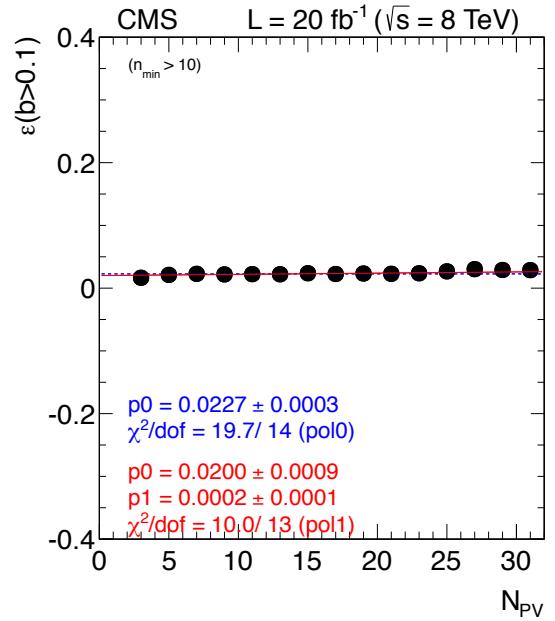
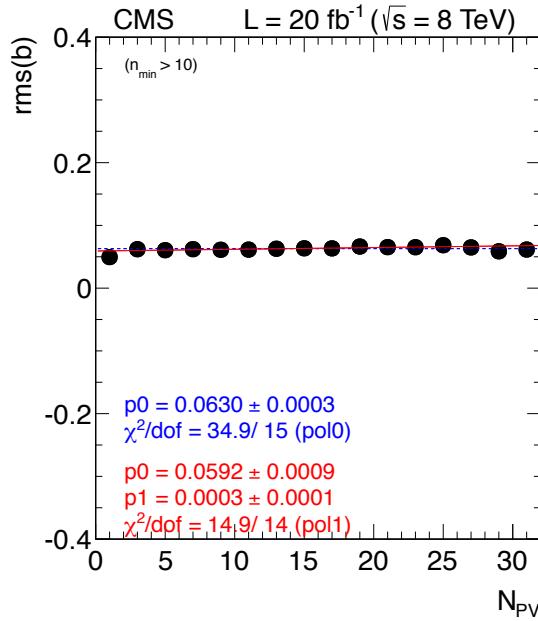
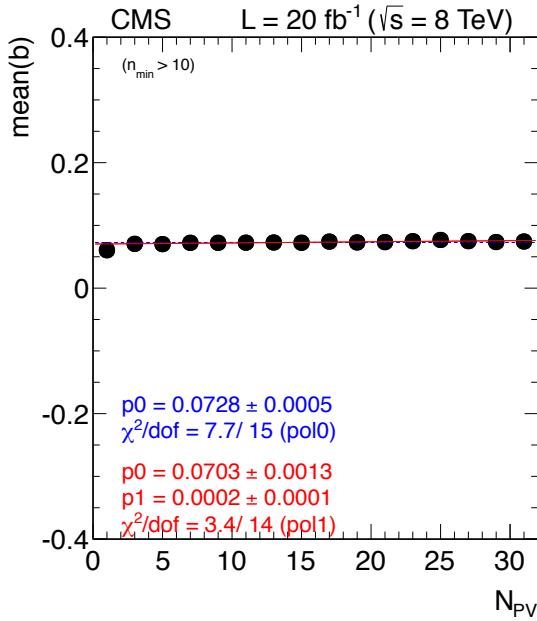
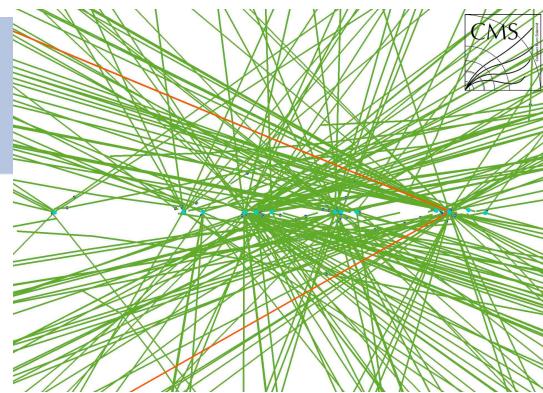


$B^\pm \rightarrow J/\psi K^\pm$  3% ;  $B_s \rightarrow J/\psi \phi$  9.5% (2011) and 3.5% (2012)



# Pile-up

- $N_{PU} \sim 9$  (2011) and  $\sim 21$  (2012)
- Selections have been tuned to be pile-up independent
  - e.g. isolation searches only for tracks coming from the same primary vertex or not associated with any
- Every input variable has been evaluated to be insensitive versus the number of reconstructed primary vertices
- All selections are compatible with a constant efficiency up to at least 30 PV ( $\sim 40$  PU)



# Normalization Channel: $B^\pm \rightarrow J/\psi K^\pm$



Same selections as for signal, plus

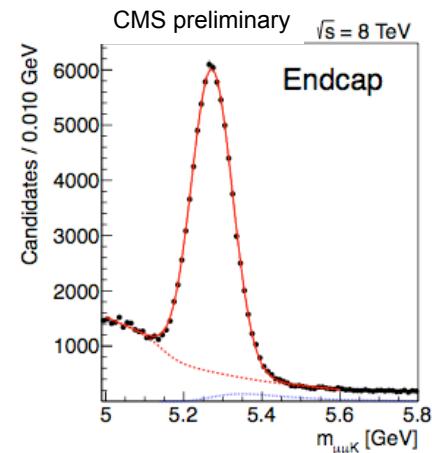
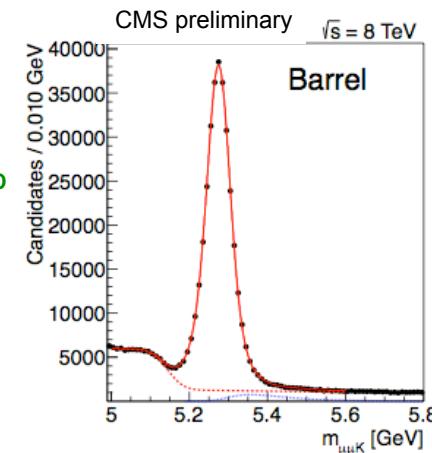
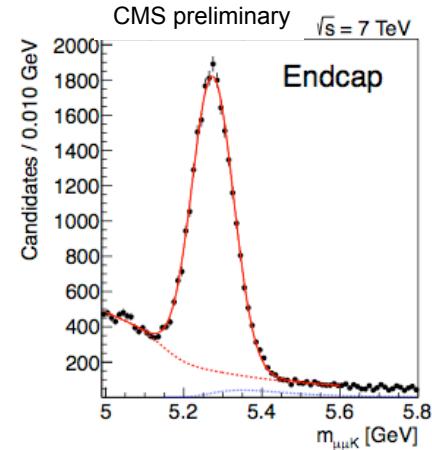
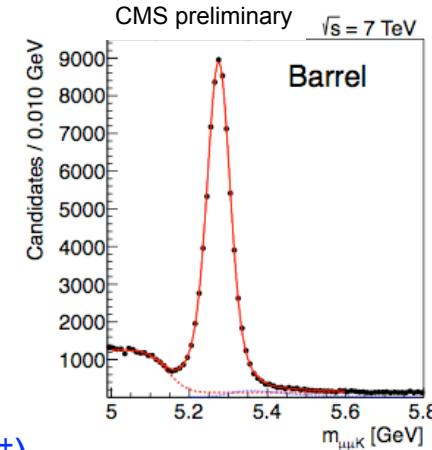
- $3.0 < m(\mu\mu) < 3.2 \text{ GeV}$
- $p_T(\mu\mu) > 7 \text{ GeV}$
- $p_T(K) > 0.5 \text{ GeV}$
- all tracks used in vertexing



**Yield extraction:**

- Signal: double (single) Gaussian in barrel (endcap)
- Backgrounds:
  - Error function for  $B_d \rightarrow J/\psi K^* \rightarrow \mu^+\mu^-K^-(\pi^+)$  decays
  - Landau function for  $B^\pm \rightarrow J/\psi \pi^\pm$  decays

- Estimated syst. error on the event yield: 5%

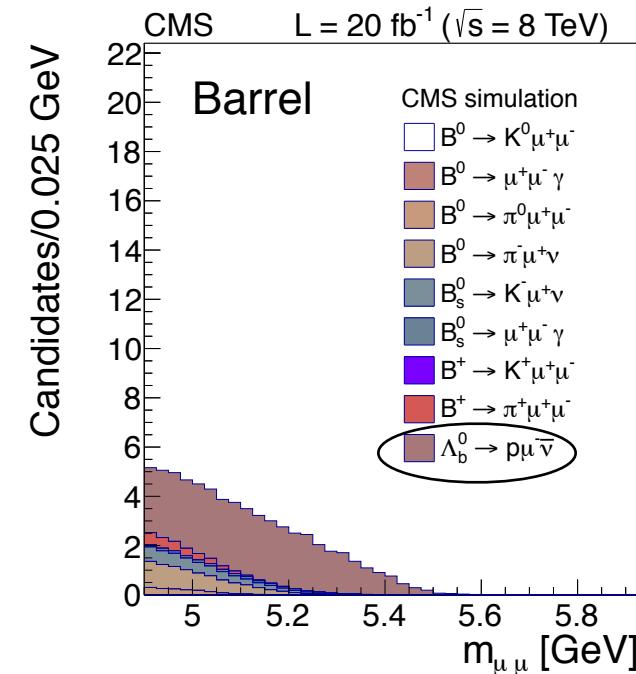
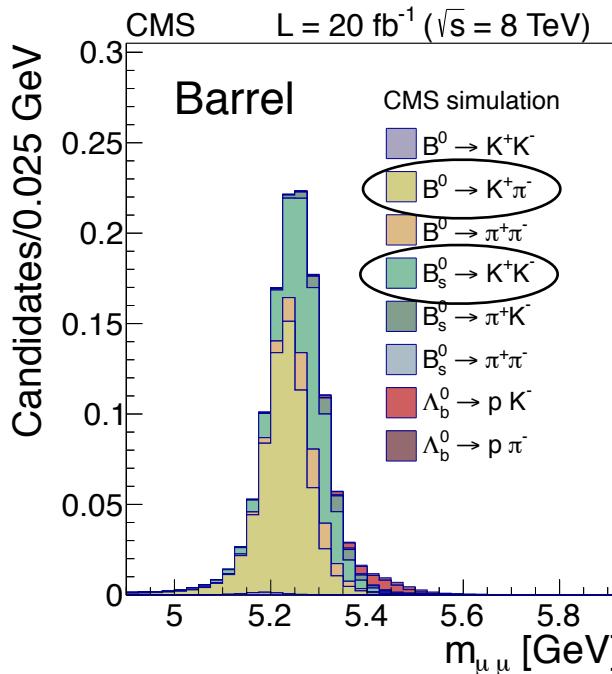


# Rare Backgrounds

- Expected no. of events in each channel normalized to  $B^\pm$  in data:

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

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



# BR extraction

To extract the  $\text{BR}(B_s \rightarrow \mu\mu)$  categorized-BDT UML mass fits

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

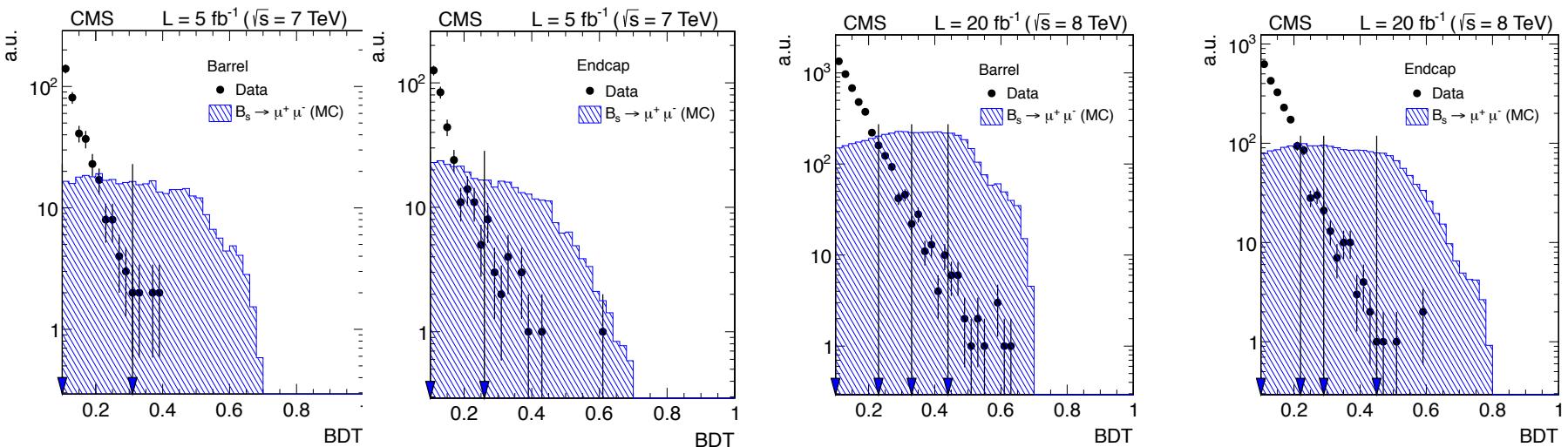
min. bin edges	1	2	3	4
2011 barrel	0.10	0.31	-	-
2011 endcap	0.10	0.29	-	-
2012 barrel	0.10	0.23	0.33	0.44
2012 endcap	0.10	0.22	0.29	0.45

■ BDT binning chosen to equalize the expected number of signal events

To extract CLs limits on  $\text{BR}(B_d \rightarrow \mu\mu)$  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



## Fit for $B_s$ 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 data-set



### Rare semileptonic background:

- Fixed shape, normalization floating constrained within 75% of nominal value
- Constrained Gaussian kernels from MC

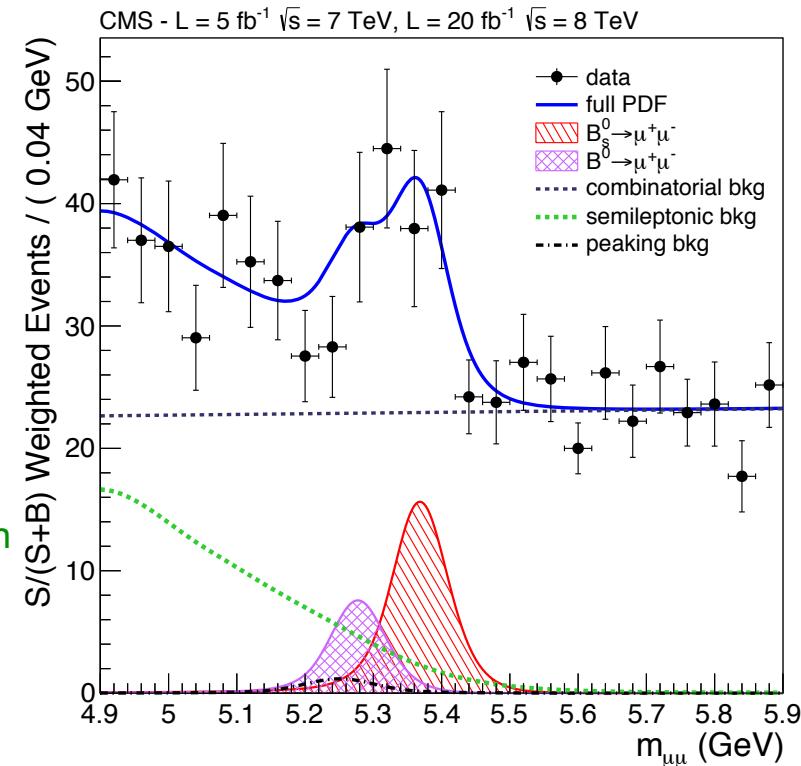


### Combinatorial background:

- First degree polynomial
- Validated with independent data set



### Per-event mass resolution included



$$\text{BR}(B_s \rightarrow \mu\mu) = \frac{N_s^i}{N_{B^+}^i} \times \frac{f_u}{f_s} \times \left( \frac{\varepsilon_s^i}{\varepsilon_u^i} \right) \times \text{BR}(B_d \rightarrow J/\psi K^\pm) \times \text{BR}(J/\psi \rightarrow \mu\mu)$$

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

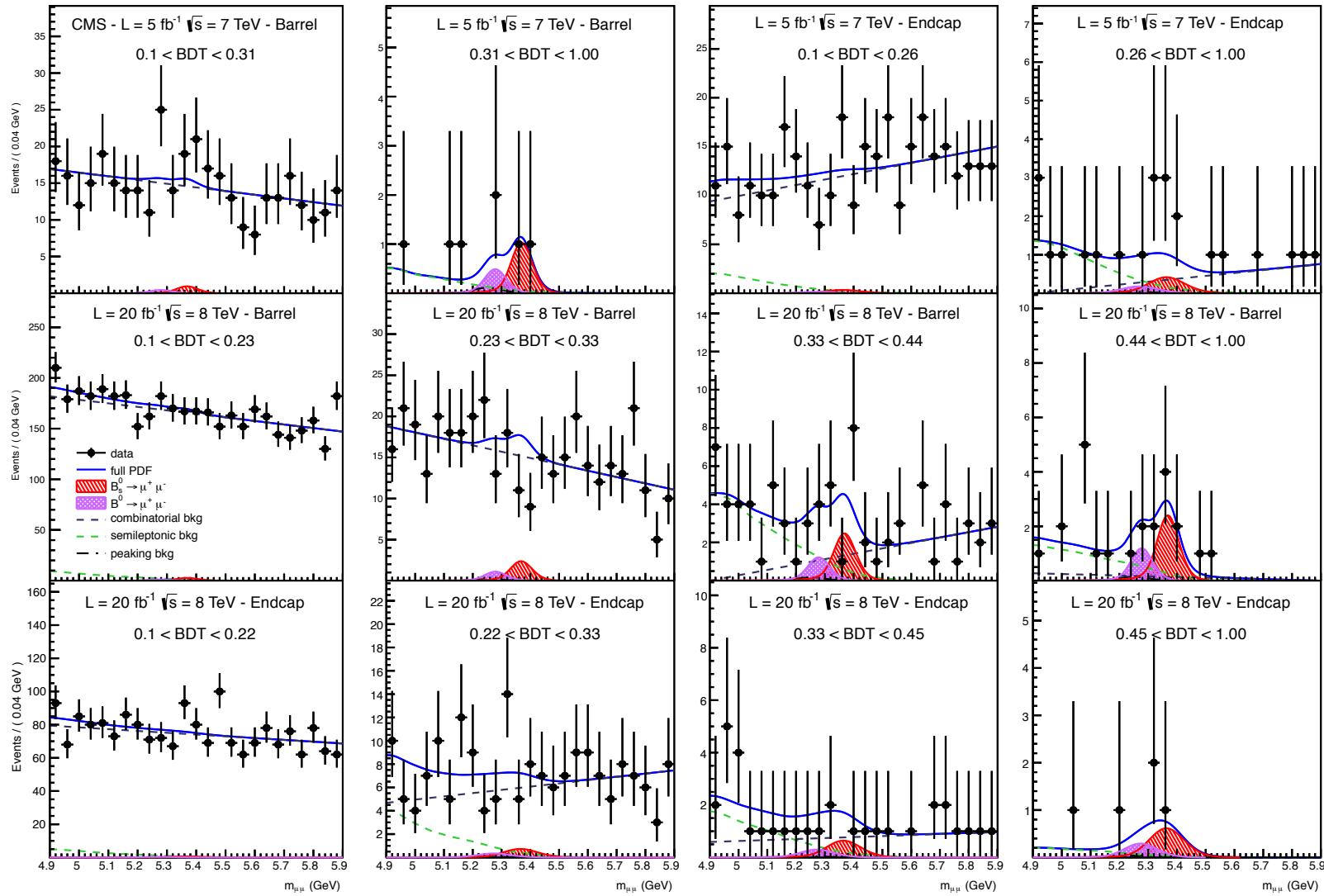
# Systematics



Implemented as Gaussian pdf constraints in the UML fit

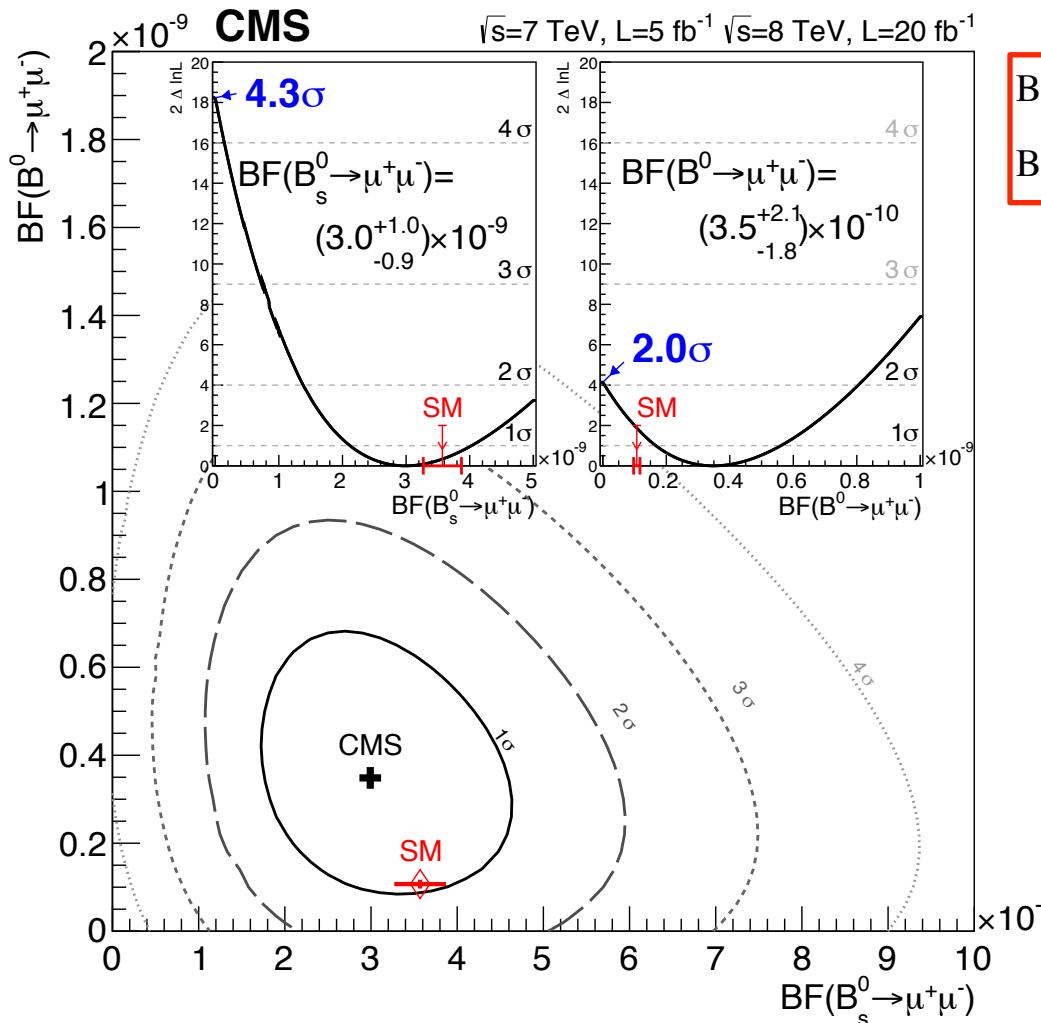
- ➊ Hadron to muon misidentification probability
  - studied with  $D^* \rightarrow D^0 \pi$ ,  $D^0 \rightarrow K\pi$ ,  $K_s \rightarrow \pi\pi$ ,  $\Lambda \rightarrow p\pi$ 
    - ➔ 50% uncertainty, conservatively assumed to be uncorrelated
- ➋ BR uncertainties
  - dominated by  $\Lambda_b \rightarrow p\mu\nu$  ( $6.5 \times 10^{-4}$ ) with 100% uncertainty
- ➌  $f_s/f_u = 0.256 \pm 0.020$  from LHCb
  - additional 5% to account for possible  $p_T$  and  $\eta$  dependence
    - ➔ in situ studies show no  $p_T$  dependence from ratios of  $B^\pm \rightarrow J/\psi K^\pm$  vs  $B_s \rightarrow J/\psi \phi$
- ➍ Normalization channel
  - yields 5%
  - $\text{BR}(B_d \rightarrow J/\psi K^\pm) \times \text{BR}(J/\psi \rightarrow \mu\mu) = (6.0 \pm 0.2) \times 10^{-5}$

# Categorized-BDT fits results





## Results of the UML fit using categorized-BDT approach



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

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

## Significances

$B_s \rightarrow \mu\mu$  4.3  $\sigma$  (expected 4.8  $\sigma$  median)

B<sub>d</sub> → μμ 2.0 σ

# Upper limits on $B_d \rightarrow \mu\mu$



No significant excess is observed for  $B_d \rightarrow \mu\mu$

- Upper limit is computed using  $CL_s$  method, based on the observed number of events in the signal and sideband regions with the 1D-BDT method.

## Expected and observed no. of events in signal regions

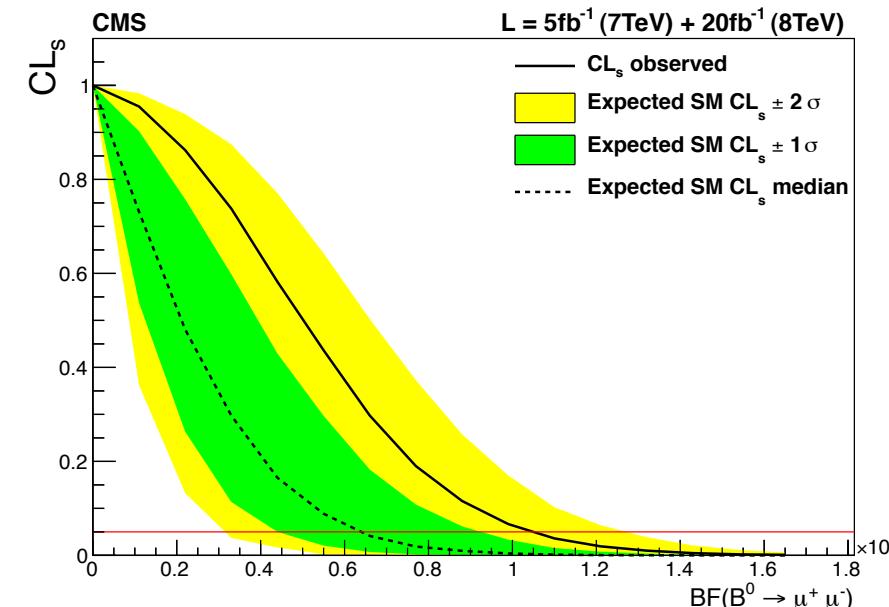
	2011 barrel		2012 barrel	
	$B^0 \rightarrow \mu^+ \mu^-$	$B_s^0 \rightarrow \mu^+ \mu^-$	$B^0 \rightarrow \mu^+ \mu^-$	$B_s^0 \rightarrow \mu^+ \mu^-$
$\epsilon_{tot} [\%]$	$0.33 \pm 0.03$	$0.30 \pm 0.04$	$0.24 \pm 0.02$	$0.23 \pm 0.03$
$N_{signal}^{exp}$	$0.27 \pm 0.03$	$2.97 \pm 0.44$	$1.00 \pm 0.10$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

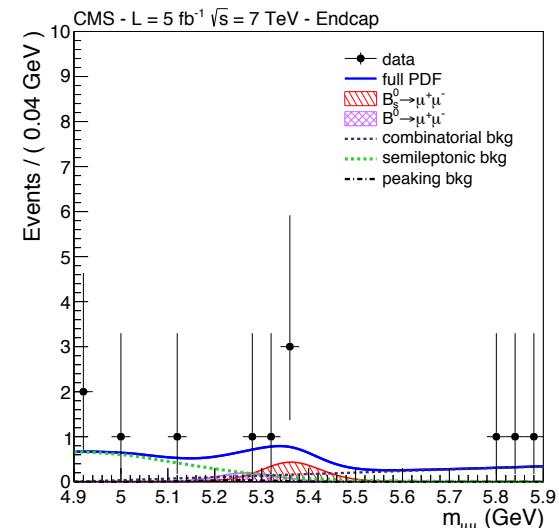
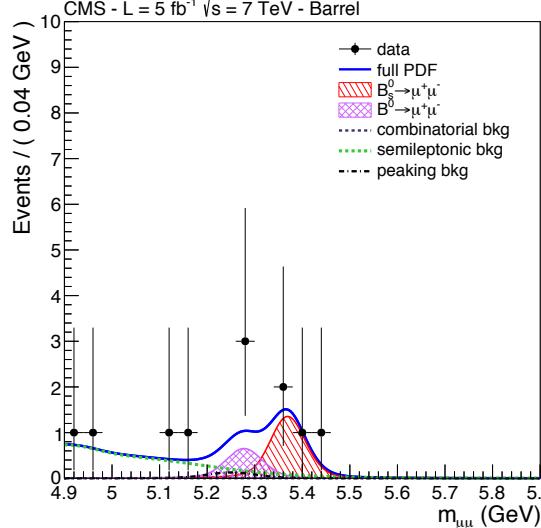
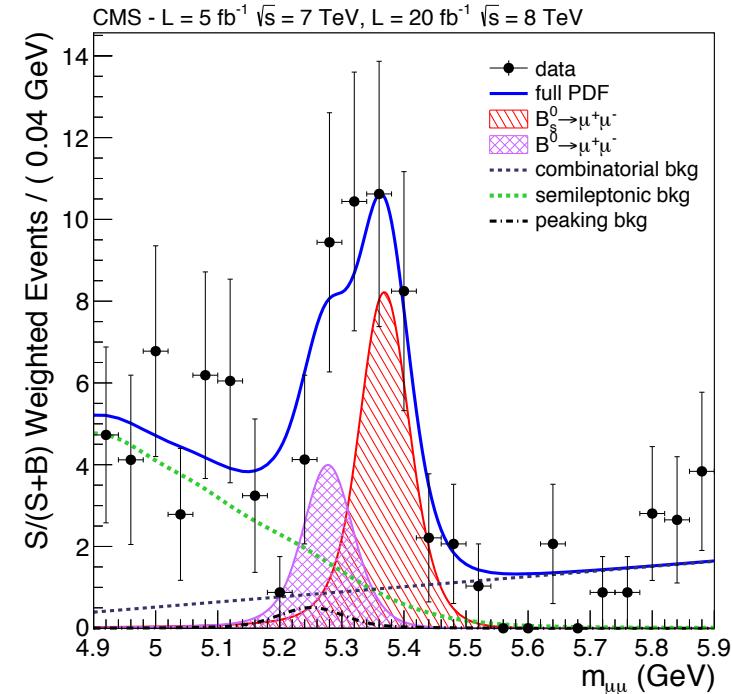
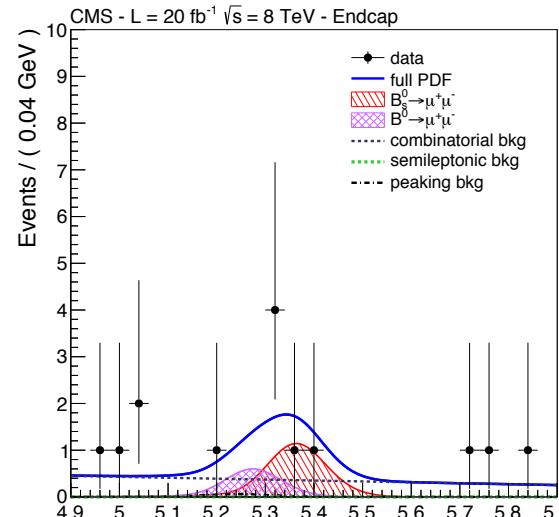
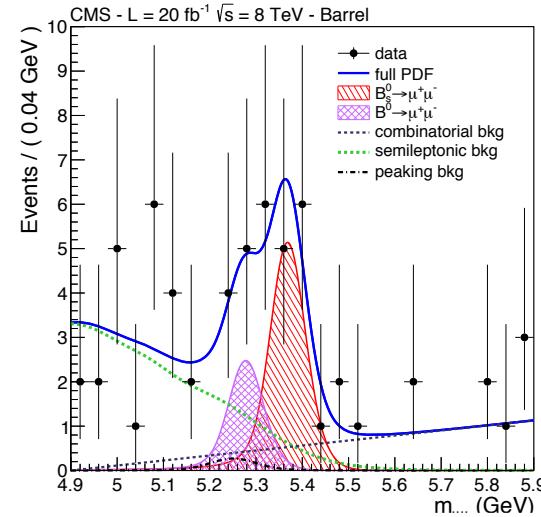
	2011 endcap		2012 endcap	
	$B^0 \rightarrow \mu^+ \mu^-$	$B_s^0 \rightarrow \mu^+ \mu^-$	$B^0 \rightarrow \mu^+ \mu^-$	$B_s^0 \rightarrow \mu^+ \mu^-$
$\epsilon_{tot} [\%]$	$0.20 \pm 0.02$	$0.20 \pm 0.02$	$0.10 \pm 0.01$	$0.09 \pm 0.01$
$N_{signal}^{exp}$	$0.11 \pm 0.01$	$1.28 \pm 0.19$	$0.30 \pm 0.03$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

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

(expected  $6.3 \times 10^{-10}$  in presence of SM+background)

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





## Significance

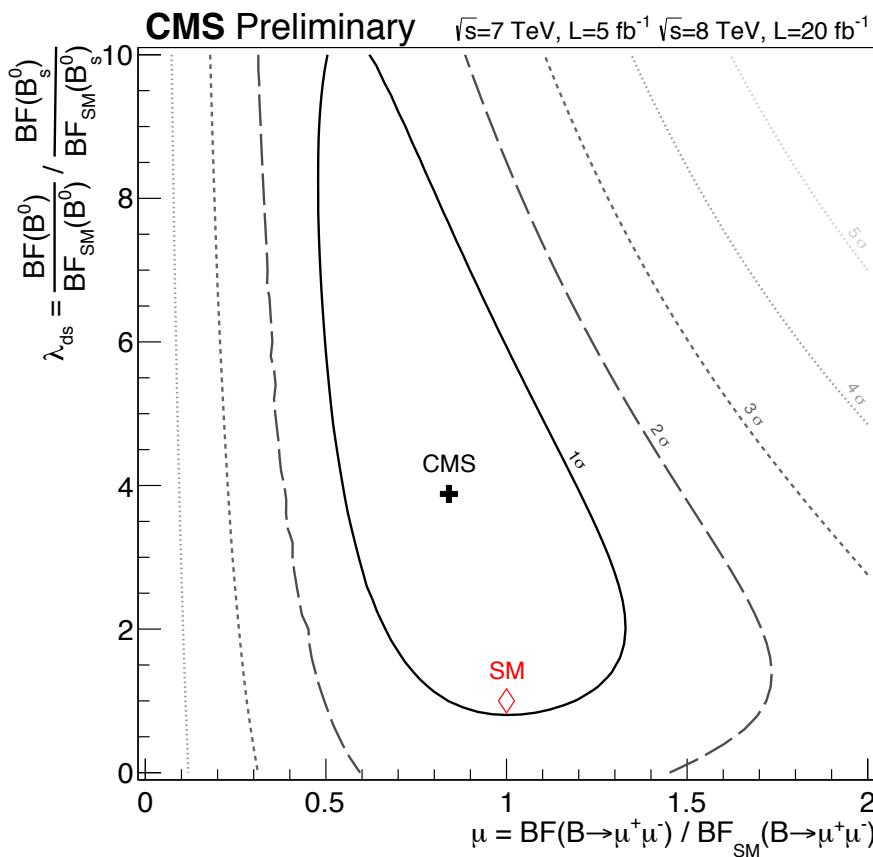
$B_s \rightarrow \mu\mu$  4.8  $\sigma$  (expected 4.7  $\sigma$  median)

Less sensitive wrt BDT-categories  
used as a cross check



Test how the BRs are compatible with the SM expectations

- $\text{BR}_{\text{SM}}(B_s \rightarrow \mu\mu) = (3.56 \pm 0.18) \times 10^{-9}$
- $\text{BR}_{\text{SM}}(B_d \rightarrow \mu\mu) = (1.07 \pm 0.10) \times 10^{-10}$



$$\mu = \frac{\text{BR}(B_s \rightarrow \mu\mu)}{\text{BR}_{\text{SM}}(B_s \rightarrow \mu\mu)}$$

$$\lambda_{ds} = \frac{\text{BR}(B_d \rightarrow \mu\mu)}{\text{BR}_{\text{SM}}(B_d \rightarrow \mu\mu)} \Bigg/ \frac{\text{BR}(B_s \rightarrow \mu\mu)}{\text{BR}_{\text{SM}}(B_s \rightarrow \mu\mu)}$$

Simultaneous fit

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

Fit for  $\mu$  (fix lambda to SM)

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

Fit for  $\lambda_{ds}$  (fix  $\mu$  to SM)

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

# Conclusions and outlook



CMS measured the  $\text{BR}(B_s \rightarrow \mu\mu)$  and put an upper limit on  $\text{BR}(B_d \rightarrow \mu\mu)$

$$\text{BR}(B_s \rightarrow \mu\mu) = (3.0^{+1.0}_{-0.9}) \times 10^{-9}$$

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

- Final LHC Run1 statistics  $\sim 25 \text{ fb}^{-1}$ , statistically dominated
- Analysis employs BDT and UML mass fit
- $4.3\sigma$  significance ( $4.8\sigma$  expected)
- Consistent with the SM expectations



## The future

- Analyze “parked data” to measure hadronic two body decays
- Optimized analysis for  $B_d \rightarrow \mu\mu$
- LHC Run 2<sup>+</sup> with increased statistics ( $300 \text{ fb}^{-1}$ )

