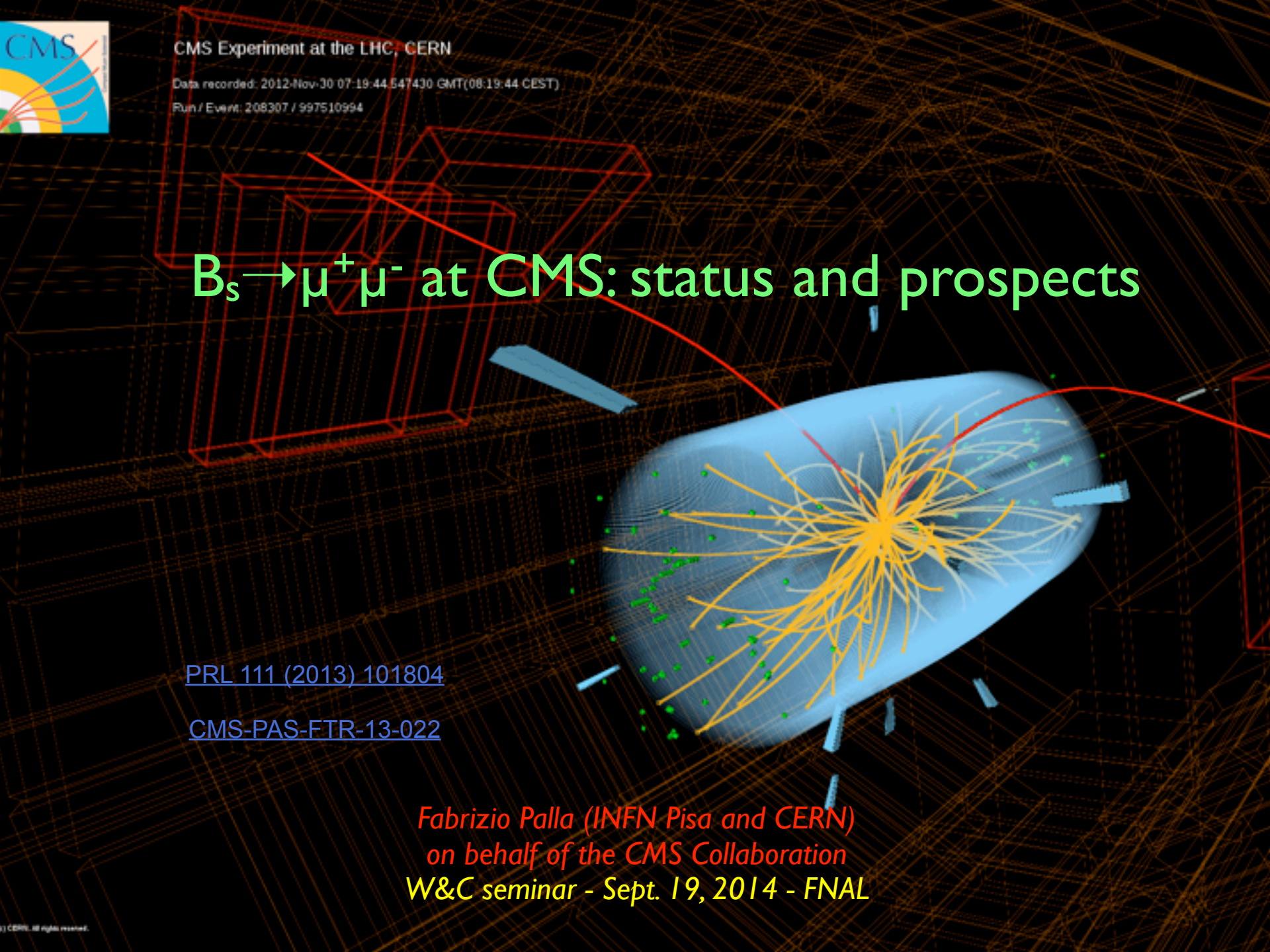


$B_s \rightarrow \mu^+\mu^-$ at CMS: status and prospects

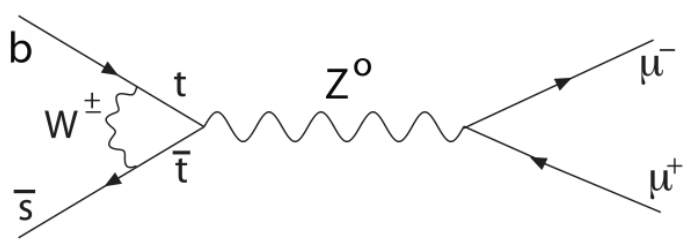
PRL 111 (2013) 101804

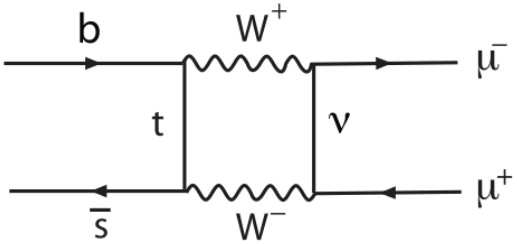
CMS-PAS-FTR-13-022

Fabrizio Palla (INFN Pisa and CERN)
on behalf of the CMS Collaboration
W&C seminar - Sept. 19, 2014 - FNAL

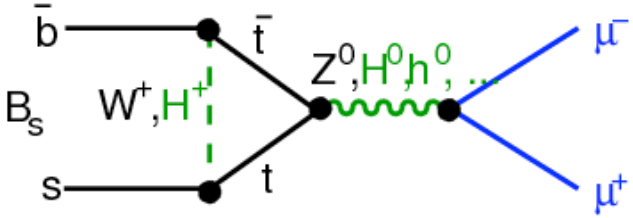
Motivations: Physics beyond SM

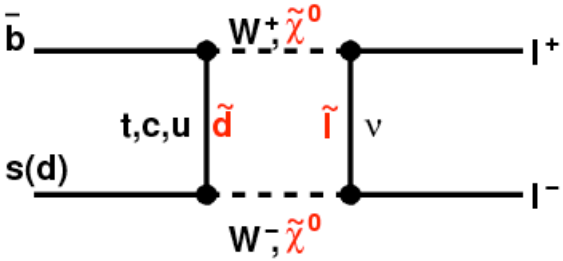
- In Standard Model decays highly suppressed
  - forbidden at tree level, can only proceed through higher-order loop diagrams
  - helicity suppressed
  - Cabibbo enhancement of $B_s \to \mu\mu$ over $B_d \to \mu\mu$ due to $|V_{td}|<|V_{ts}|$





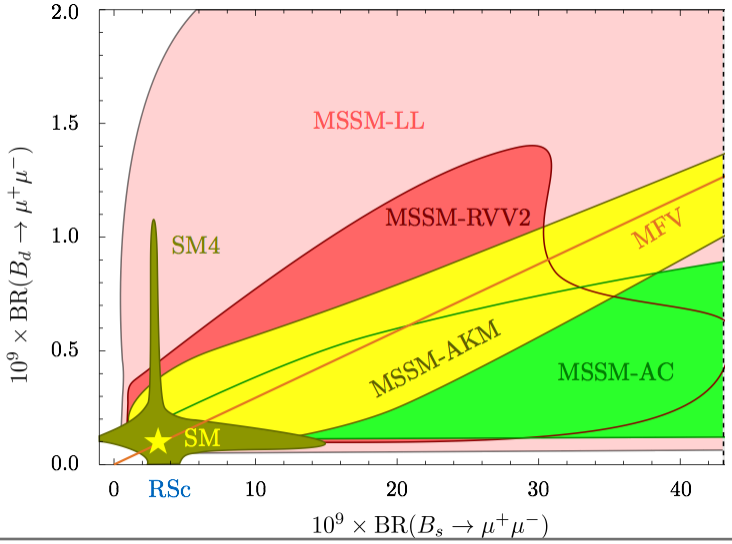
  - A few examples of interfering New Physics contributions





Physics Beyond Standard Model

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



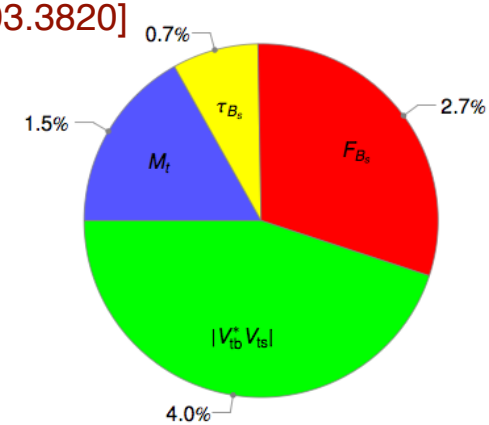
D. M. Straub
http://arxiv.org/pdf/1012.3893v2.pdf

Theory prediction

- Standard Model prediction for CP averaged branching fractions at t=0
  - $BR^{t=0}(B_s \to \mu\mu) = (3.25 \pm 0.17) \times 10^{-9}$
  - $BR^{t=0}(B_d \to \mu\mu) = (1.07 \pm 0.10) \times 10^{-10}$

[A. Buras et al. arXiv:1303.3820]

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



- Experimental values are time integrated

[De Bruyn et al. (PRL 109, 041801)]
[A. Buras et al. arXiv:1303.3820]

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

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

τ_{Bs}=1.516±0.011 ps
$\Delta\Gamma_s/2\Gamma_s$=0.0615±0.0085

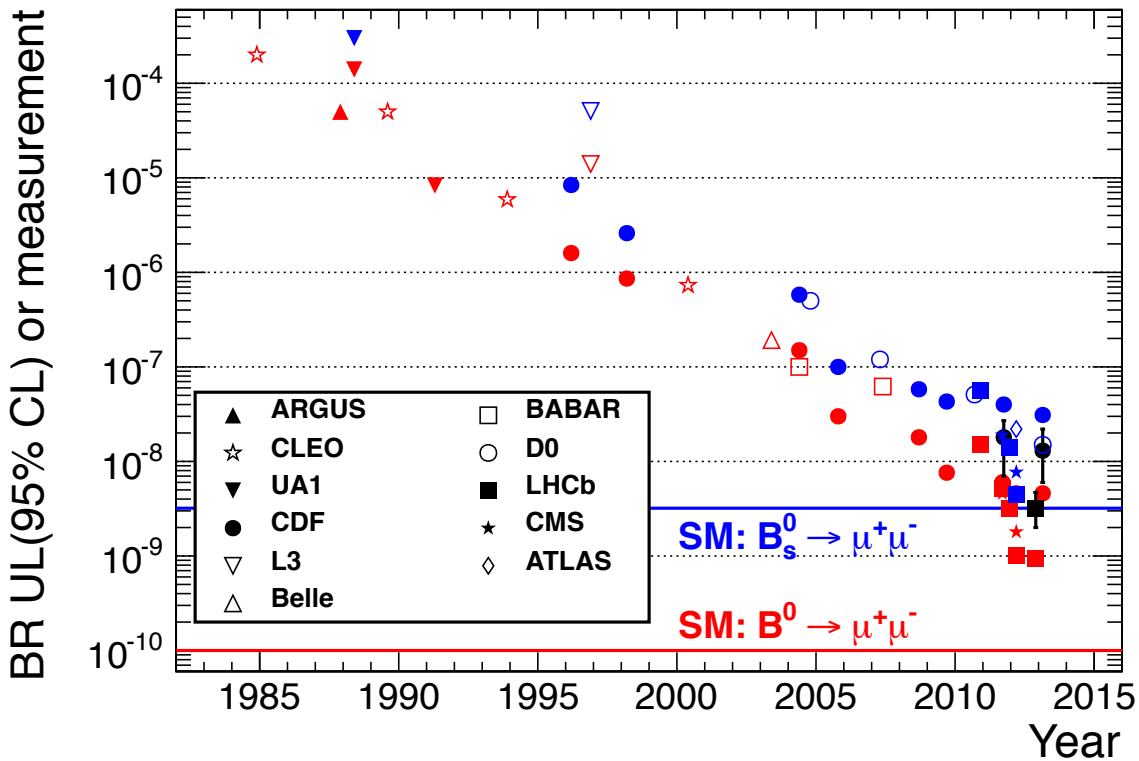
State of the art until EPS-2013

More than 25 years quest

Previous CMS publications

EPS 2011: PRL, 107, 191802

Moriond 2012: JHEP, 1204, 033



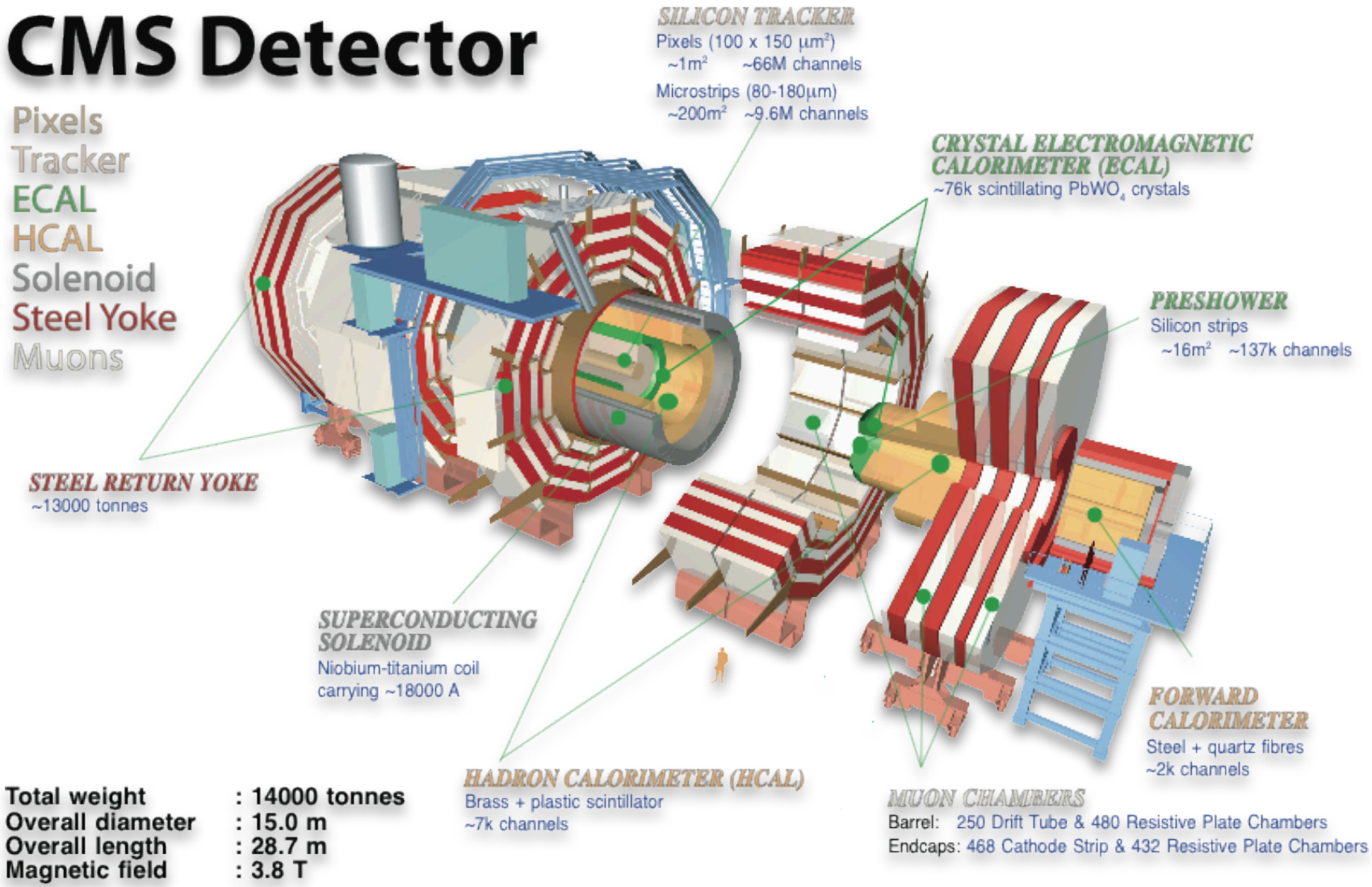
$$\mathrm{BR}(B_S \to \mu\mu) = \left(3.2^{+1.4}_{-1.2}\ (\mathrm{stat})^{+0.5}_{-0.3}\ (\mathrm{syst})\right) \times 10^{-9}$$ LHCb 3.5 σ evidence

Best before July 2013 :

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

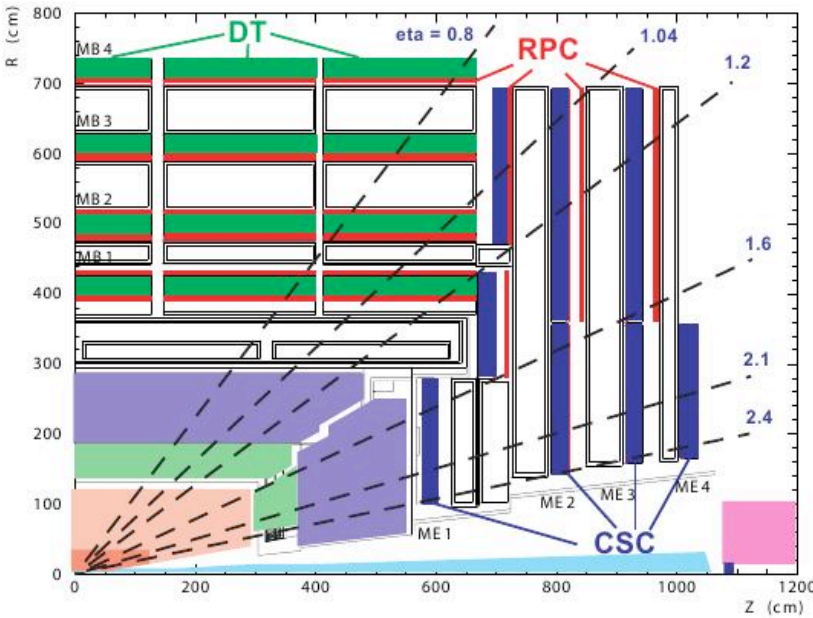
ATLAS+CMS+LHCb best upper limit on $B_d \to \mu\mu$

The CMS Detector

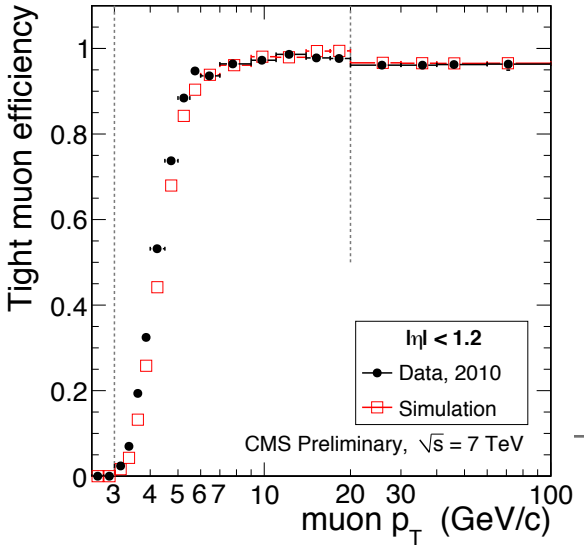


Muon reconstruction

- Tracks: Excellent p_T resolution ≈ 1%
  - Tracking efficiency > 99% for central muons
  - Excellent vertex reconstruction and impact parameter resolution (≈ 15 μm)
- 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
  - $D^* \to D^0 \pi$, $D^0 \to K\pi$
  - $K_S \to \pi\pi$
  - $\Lambda \to p\pi$

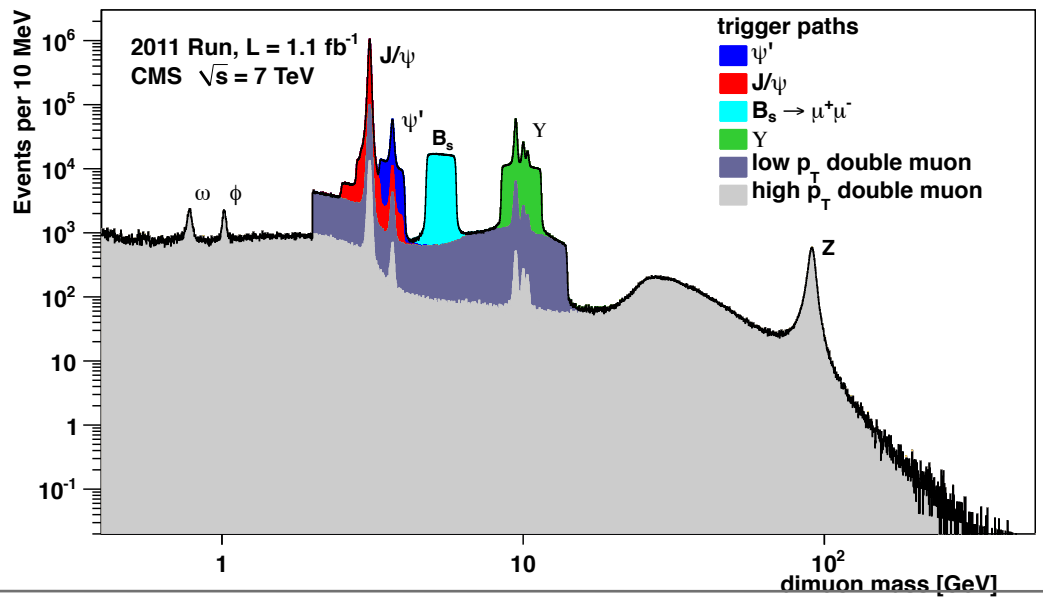


CMS-PAS-MUO-10-002



Trigger

- Dimuon Trigger
  - L1 Hardware Trigger
    - p_T>3 GeV (few kHz)
  - HLT Full tracking and vertexing
  - HLT $B_s \to \mu\mu$
    - Leading and sub-leading μ p_T>3,4 (4,4) GeV $|\eta_{\mu\mu}|$<1.8 (1.8<$|\eta_{\mu\mu}|$<2.2)
    - p_T (μμ)>5 (4.8-6) GeV
    - 4.8 <m(μμ)< 6.0 GeV
    - P(χ^2/dof)> 0.5%
- HLT $B^\pm \to J/\psi K^\pm$ and $B_s \to J/\psi\varphi$
  - $_T$4 GeV, |η|2.2
  - p_T (μμ)6.9 GeV
  - cos α 0.9
  - L/σ(L)3
  - distance of closest approach0.5 cm
  - P(χ^2/dof)15%



Analysis overview

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

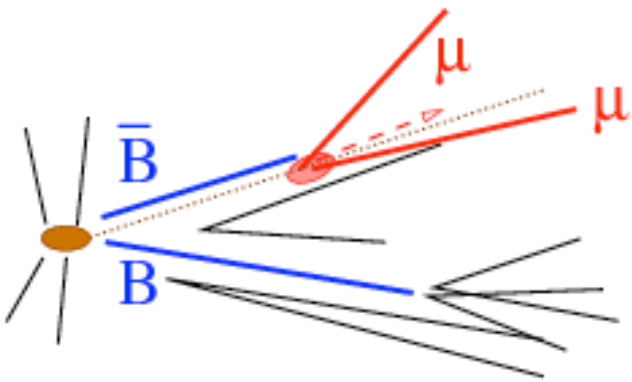
$$Br\left(B_s^0 \rightarrow \mu^+\mu^-\right) = \frac{N_S}{N_{obs}^{B^+}} \frac{f_u}{f_s} \frac{\varepsilon_{tot}^{B^+}}{\varepsilon_{tot}} Br\left(B^+\right)$$

Region definitions	Invariant mass (GeV)
overall window	4.90 $< m_{\mu1\mu2} <$ 5.90
blinding window	5.20 $< m_{\mu1\mu2} <$ 5.45
$B^0 \rightarrow \mu^+\mu^-$ window	5.20 $< m_{\mu1\mu2} <$ 5.30
$B_s \rightarrow \mu^+\mu^-$ window	5.30 $< m_{\mu1\mu2} <$ 5.45

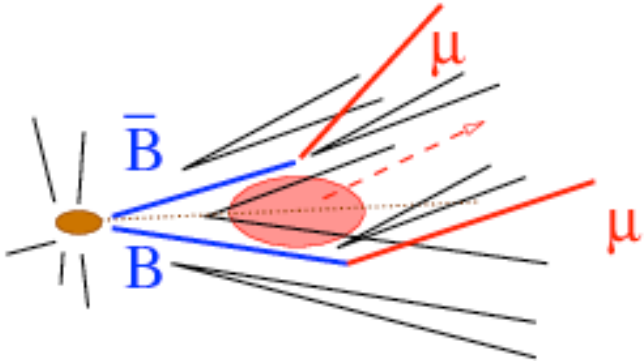
- Control sample $B_s \rightarrow J/\psi\varphi \rightarrow (\mu^+\mu^-)(K^+K^-)$ to compare and validate B_s mesons in data and MC simulations
- Divide the sample in two main categories, per each year of data taking:
  - both muons in the barrel (|η|1.4) ➡better sensitivity, B_s mass resolution ≈ 40 MeV
  - ≥1 μ in the endcap ➡more events but B_s mass resolution ≈ 60 MeV

Event characteristics

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



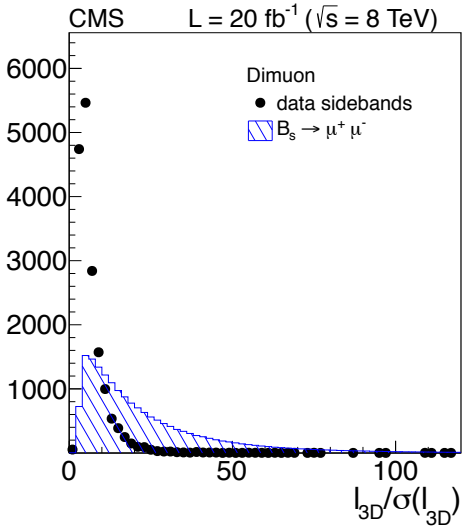
- Backgrounds
  - two semileptonic B decays
  - one semileptonic B decay and one misidentified hadron
  - Single-B decays
    - peaking (ex. $B_s \rightarrow K^- K^+$)
    - rare semileptonic (ex. $\Lambda_b \rightarrow p\mu\nu$)

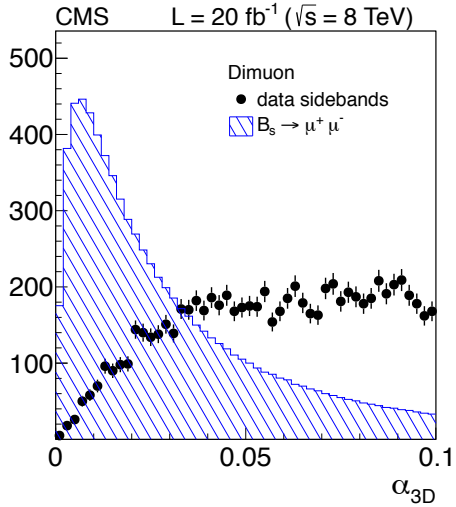


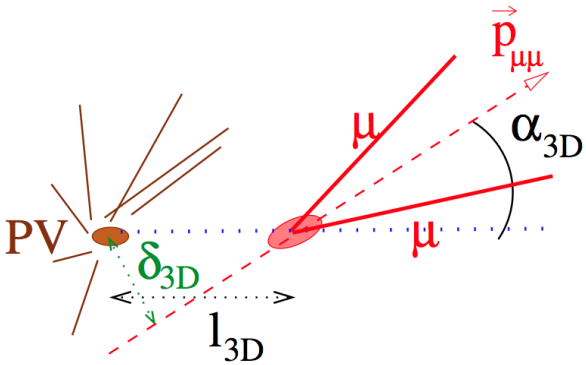
Improved muon identification

- Improved muon identification to cope with peaking and semileptonic backgrounds
  - 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
    - χ^2/ndof
    - muon p_T and η
  - Muon MC efficiency validated with TagProbe methods in data
  - Muon fake rate reduced by a factor of ~2 compared to previous analysis, at the expenses of ~10% muon efficiency reduction
    - $\varepsilon(\mu|\pi)$ ~ 0.15%
    - $\varepsilon(\mu|K)$ ~ 0.20%
    - $\varepsilon(\mu|p)$ ~ 0.10%
  - Analysis uses MC p_T-dependent fake rate, validated with data (50% uncertainty)
    - $D^* \to D^0 \pi$ ($D^0 \to K\pi$), $K_S \to \pi\pi\pi$, $\Lambda \to p\pi$

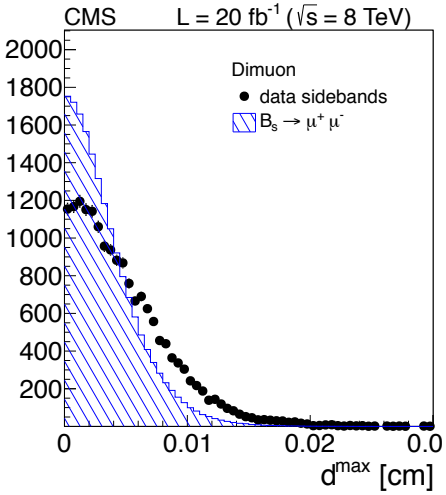
Vertexing variables

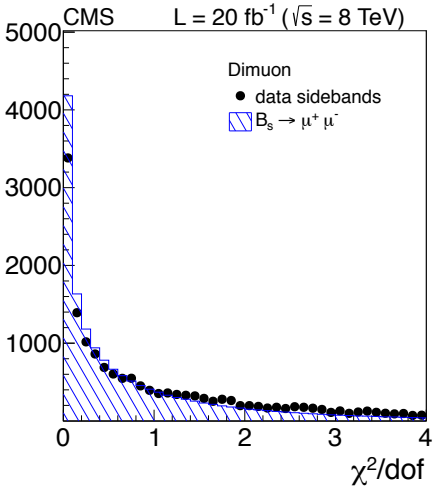


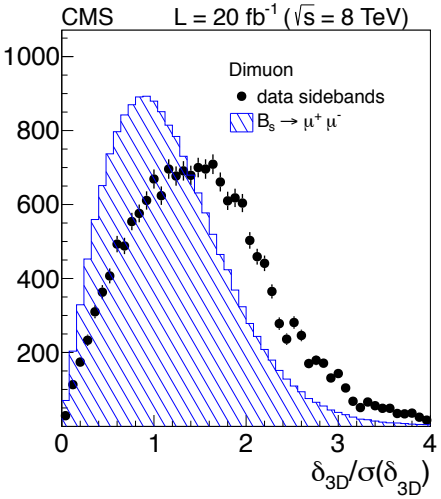


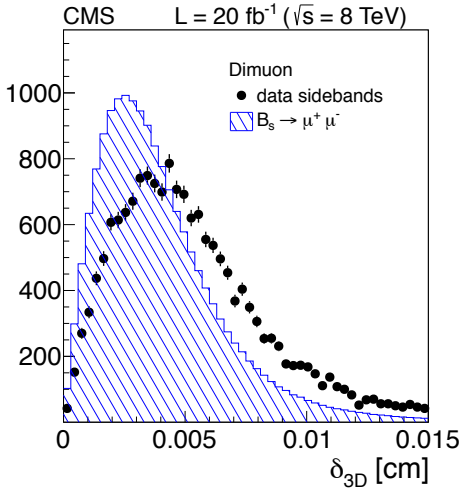


Data side-bands vs signal MC:



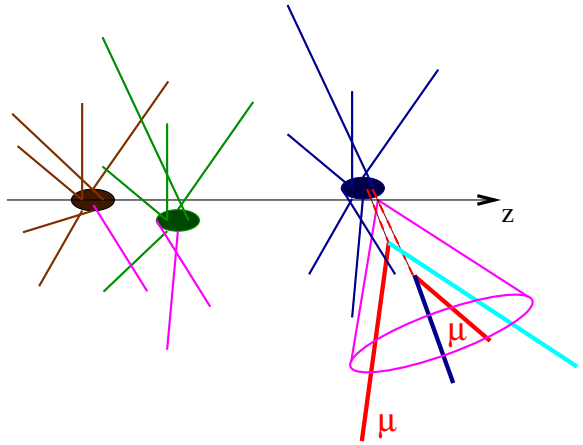




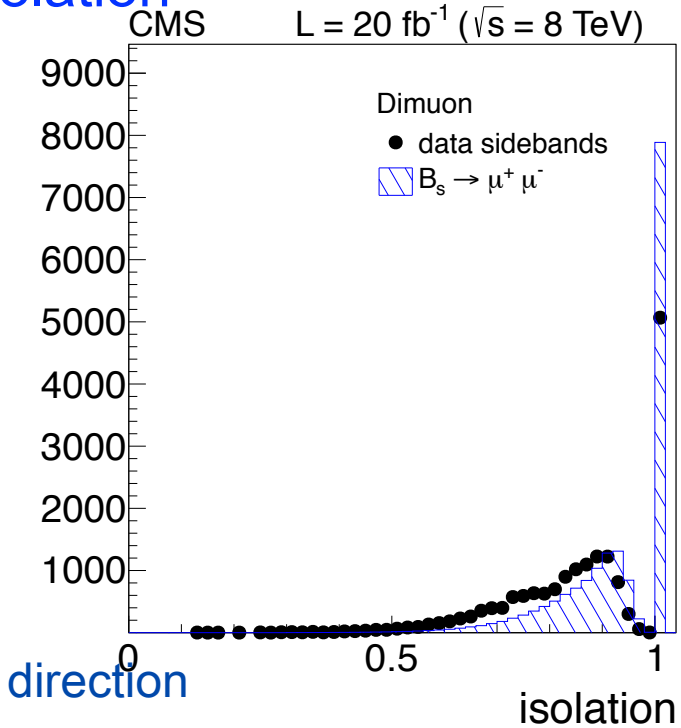


Isolation (1)

- Primary vertex isolation: relative dimuon isolation



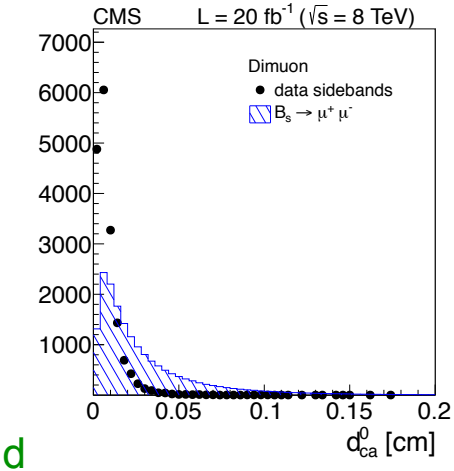
$$I = \frac{p_T(B)}{p_T(B) + \sum_{\Delta R<0.7, p_T>0.9\,GeV} p_T}$$

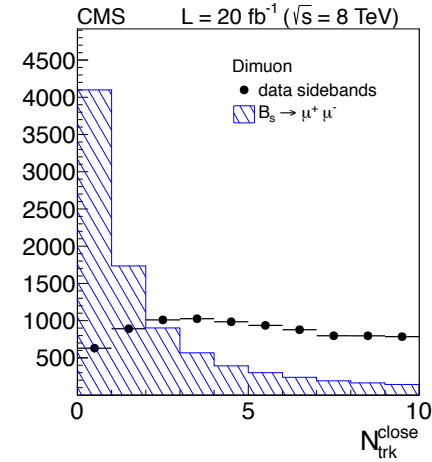


- All tracks within cone of ΔR=0.7 around dimuon direction and with p_T 0.9 GeV
  - either from the same PV
  - or d_{ca}>0.5 mm from B vertex.
- dip at ~0.97 from minimum track p_T requirement

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 μm and p_T 0.5 GeV

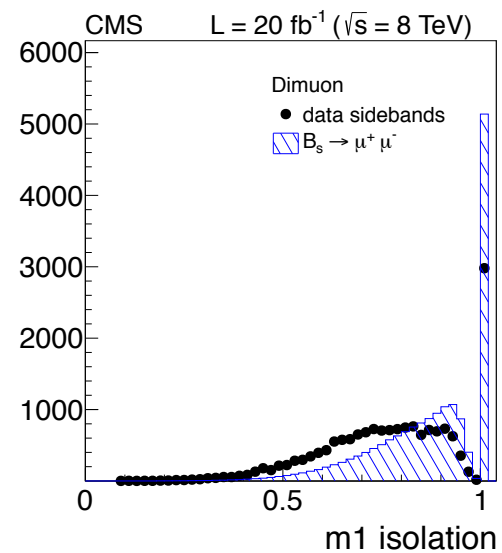




- Muon isolation (new!)

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

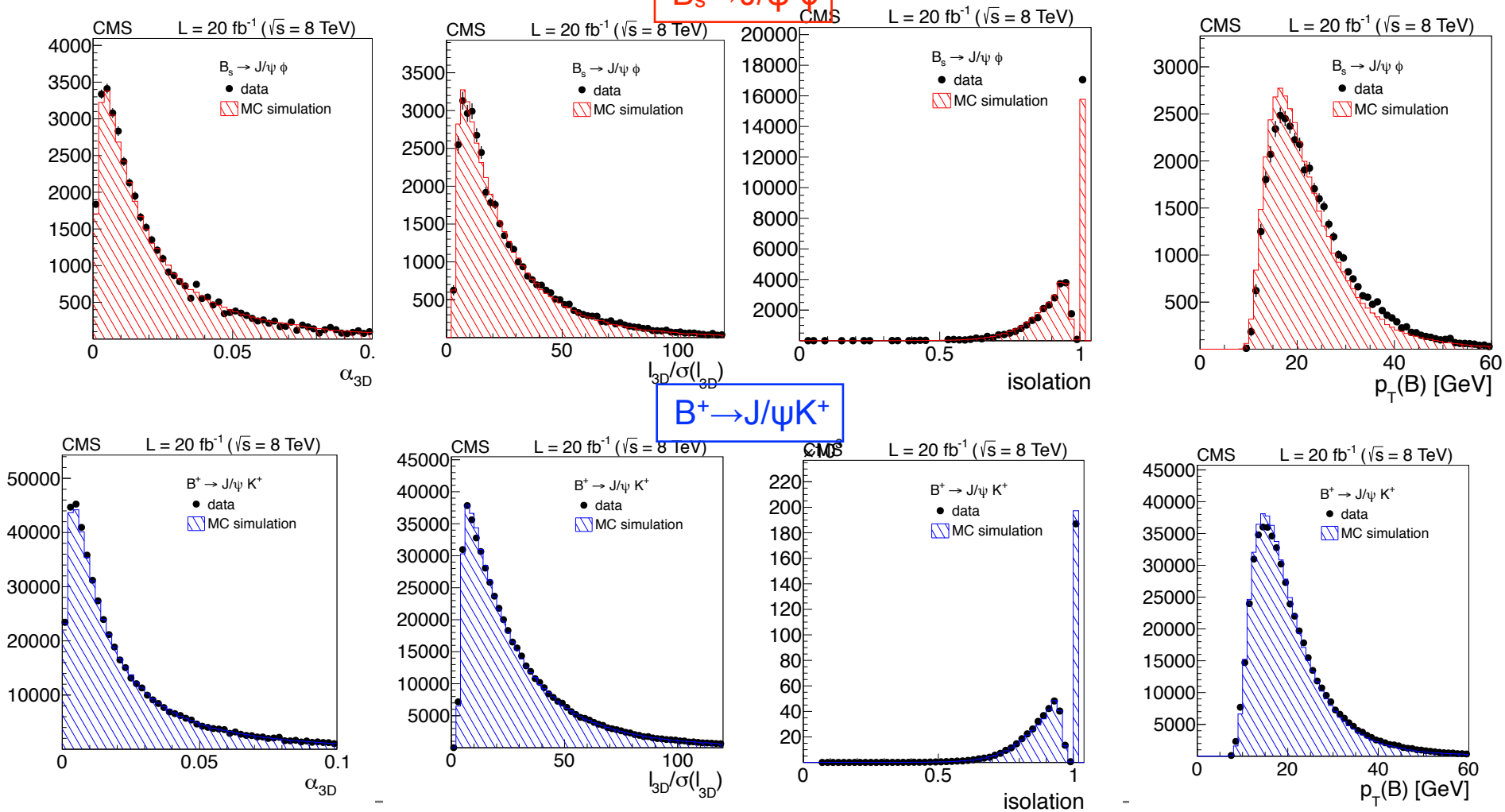
  - tracks in muon cone with ΔR=0.5
    - p_T >0.5 GeV
    - d_{ca}>1mm from the same PV or not associated to another PV



Data - Simulation comparison

Good agreement of sideband-subtracted distributions

$B_s \to J/\psi\ \phi$



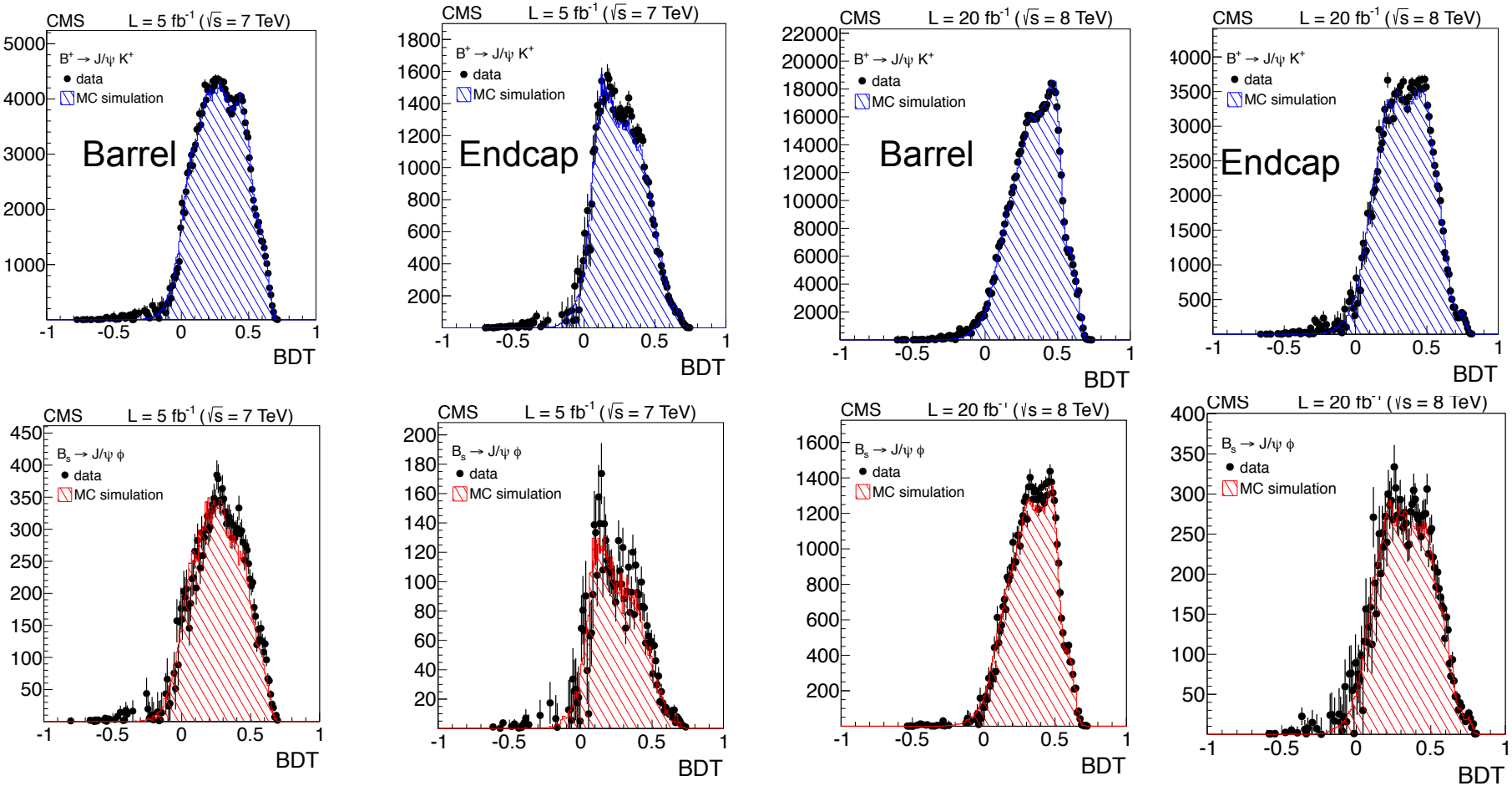
$B^+ \to J/\psi K^+$

Boosted Decision Tree (BDT) Selection

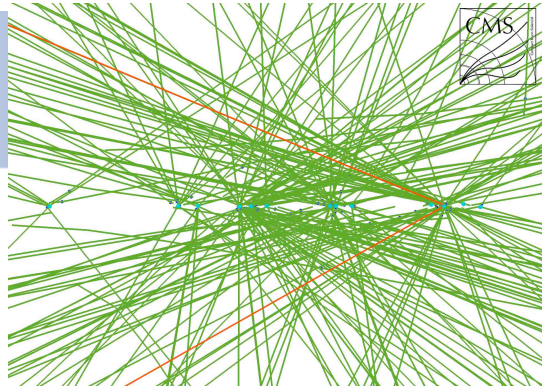
- BDT Training
  - TMVA framework
    - signal: B_s MC simulation
    - background: dimuon data sidebands
  - In order to avoid any bias and use three different BDTs:
    - A given BDT was used for training on "1st" event, tested on "2nd" and applied on "3rd", and then rotate
- Studies
  - BDT output 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

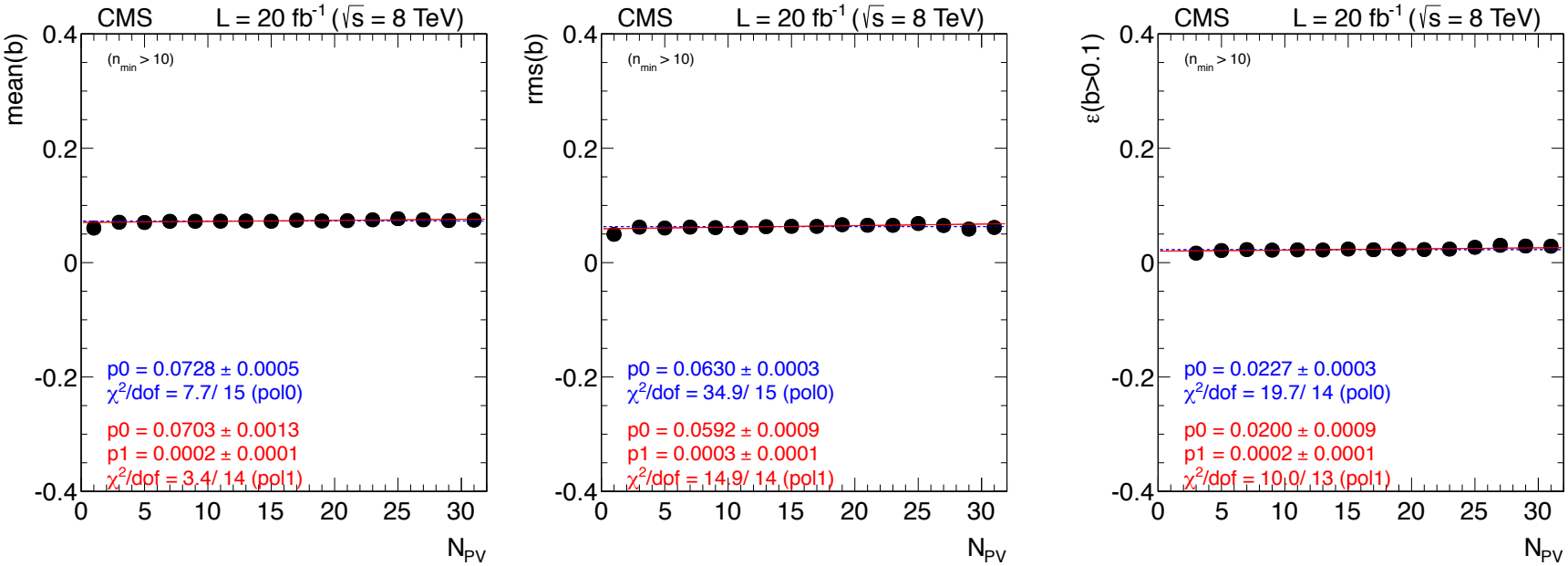
- Use differences between data and MC for systematics
  - $B^{\pm} \to J/\psi\, K^{\pm}$ 3% ; $B_s \to J/\psi\, \phi$ 9.5% (2011) and 3.5% (2012)



Pile-up

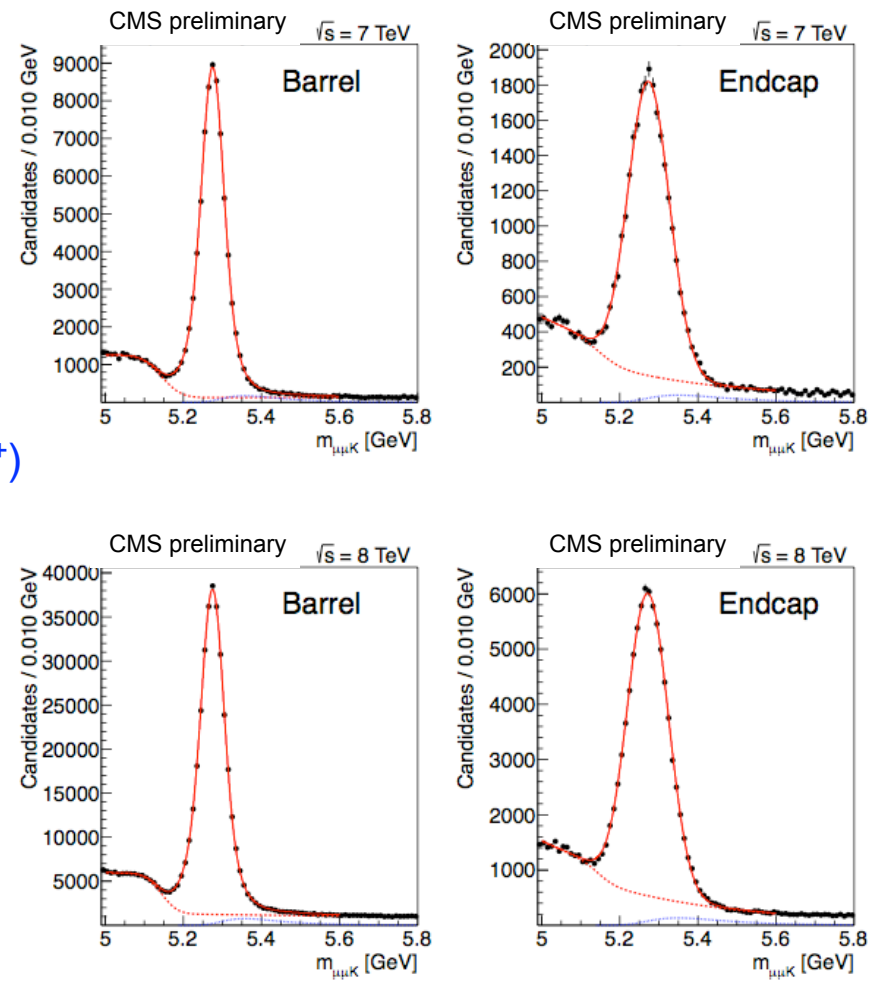


- N_{PU}~ 9 (2011) and ~21 (2012)
- Selections have been tuned to be pile-up independent
  - e.g. isolation searches only for tracks coming from the same primary vertex or not associated with any
- Every input variable has been evaluated to be insensitive versus the number of reconstructed primary vertices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



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

- Same selections as for signal, plus
  - 3.0 <m(μμ)< 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)
  - 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%

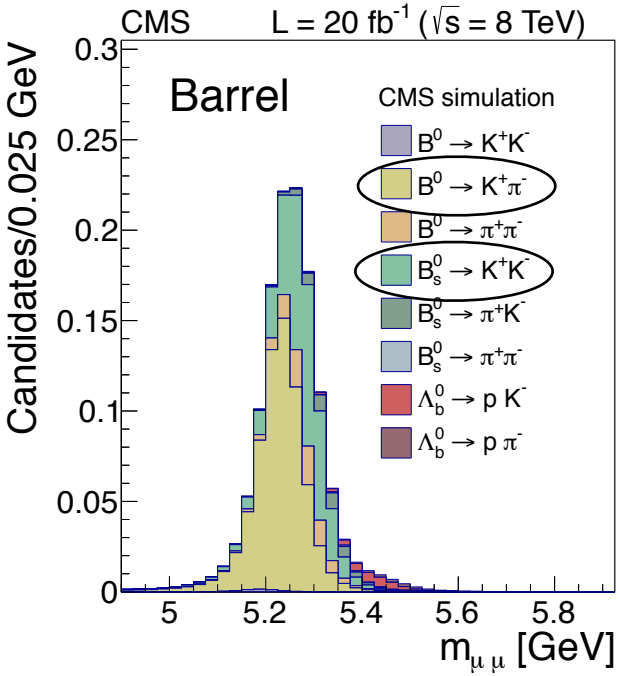


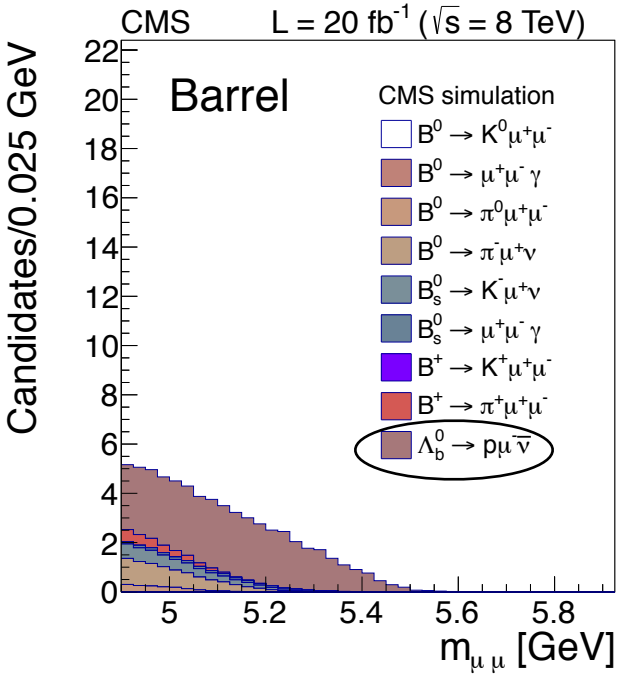
Rare Backgrounds

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

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

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





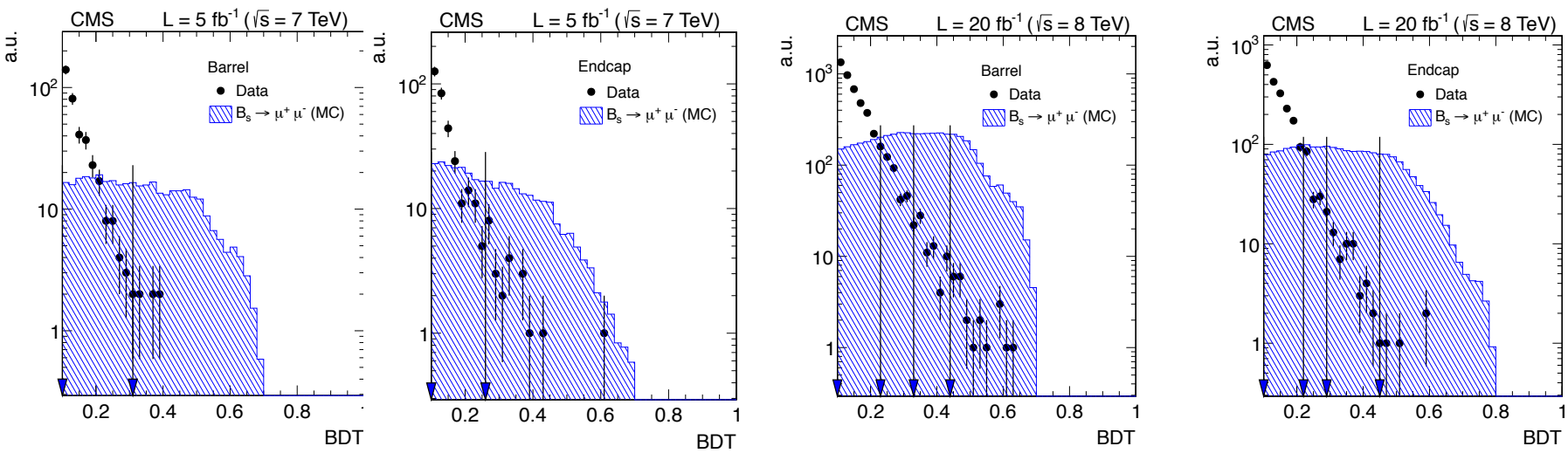
BR extraction

- To extract the BR($B_s \to \mu\mu$) 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

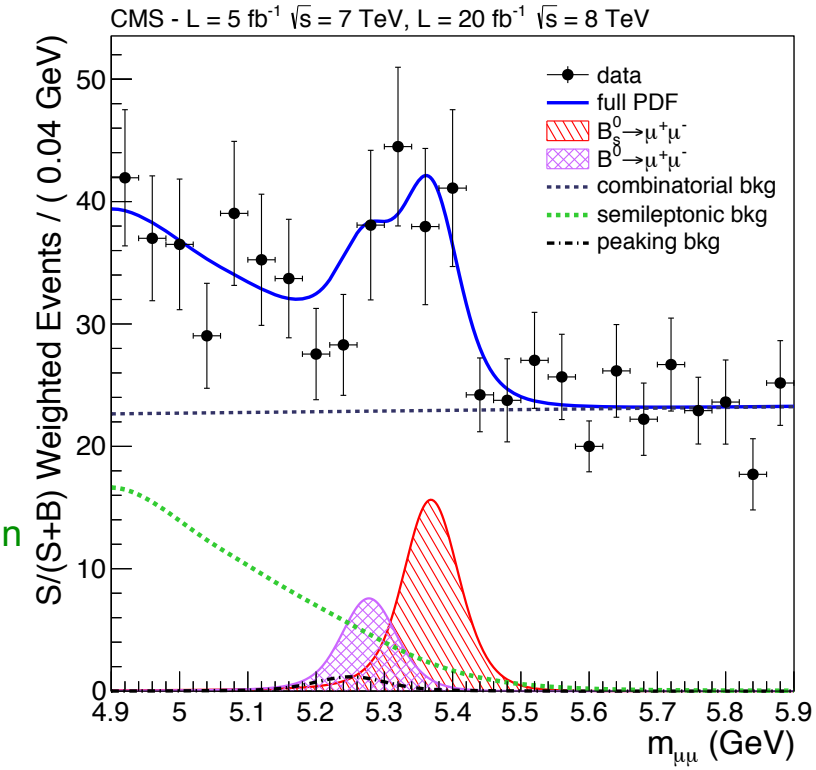
    - BDT binning chosen to equalize the expected number of signal events
- To extract CL_S limits on BR($B_d \to \mu\mu$) use 1D-BDT
  - Optimized cut on BDT output and event counting in mass windows

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



Unbinned Maximum Likelihood fit

- 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



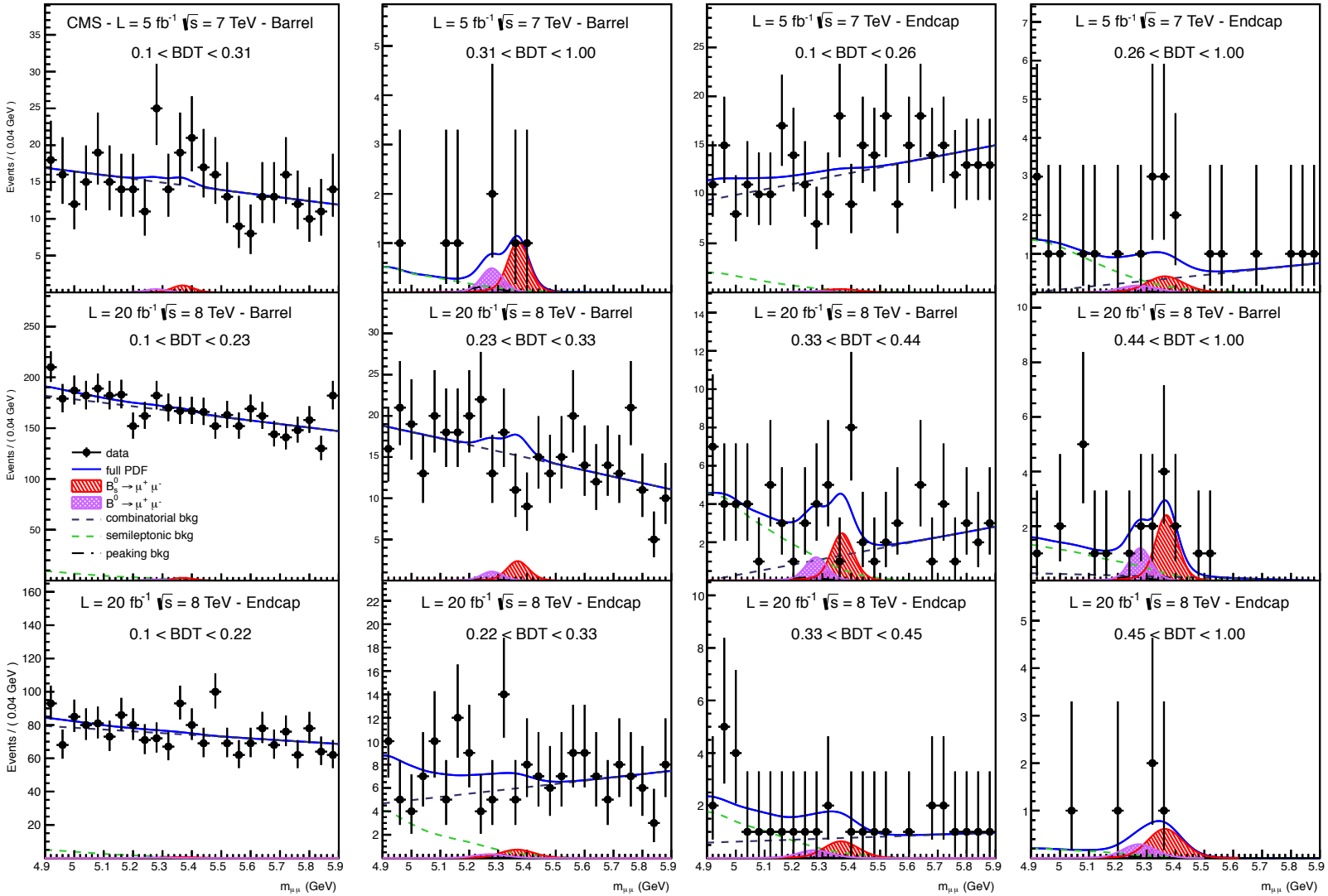
$$\mathrm{BR}(B_S \to \mu\mu) = \frac{N_S^i}{N_{B^+}^i} \times \frac{f_u}{f_s} \times \left(\frac{\varepsilon_s^i}{\varepsilon_u^i}\right) \times \mathrm{BR}(B_d \to J/\psi K^\pm) \times \mathrm{BR}(J/\psi \to \mu\mu)$$

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

Systematics

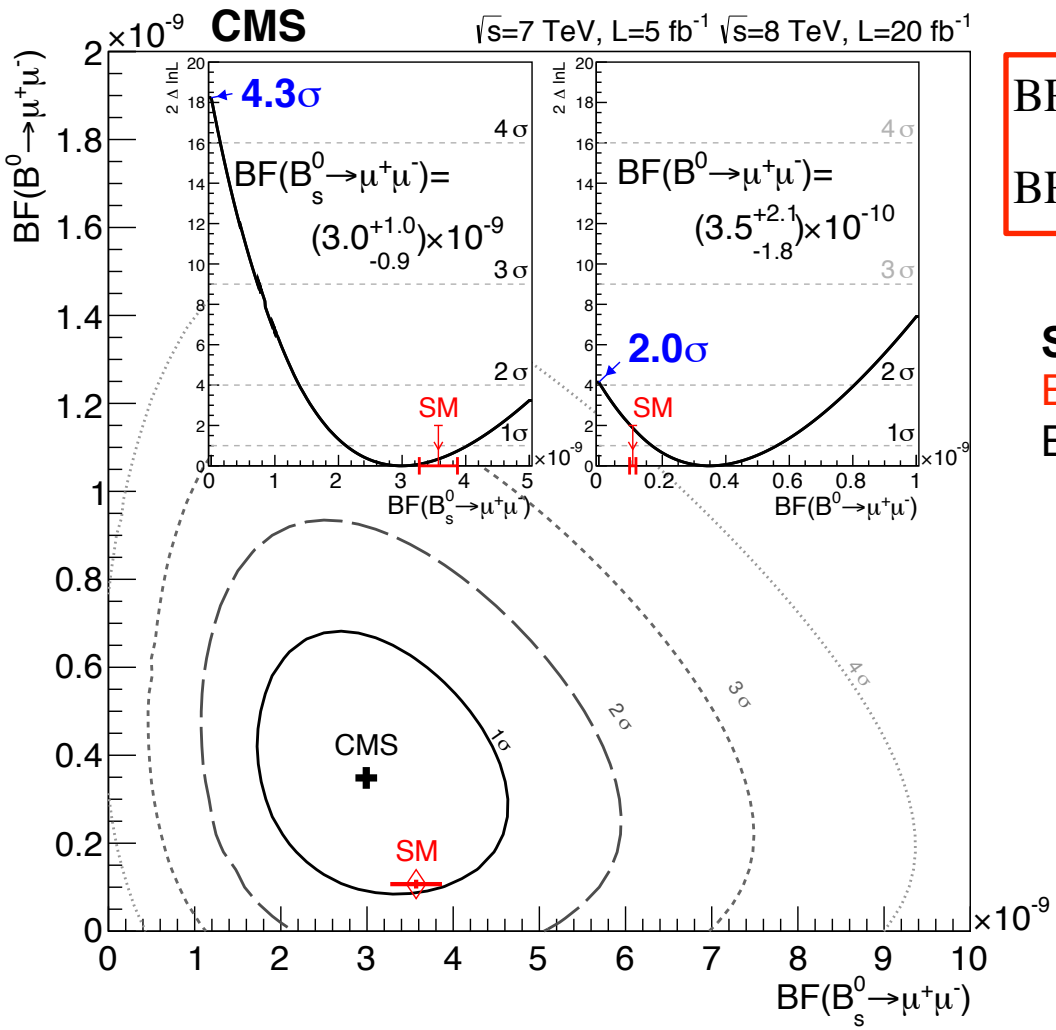
- Implemented as Gaussian pdf constraints in the UML fit
  - Hadron to muon misidentification probability
    - studied with $D^* \to D^0 \pi$, $D^0 \to K\pi$, $K_S \to \pi\pi$, $\Lambda \to p\pi$
      - ➡ 50% uncertainty, conservatively assumed to be uncorrelated
  - BR uncertainties
    - dominated 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 $Bs \to J/\psi\varphi$
  - Normalization channel
    - yields 5%
    - $\mathrm{BR}(B_d \to J/\psi K^\pm) \times \mathrm{BR}(J/\psi \to \mu\mu) = (6.0 \pm 0.2) \times 10^{-5}$

Categorized-BDT fits results



Branching Ratios results

- Results of the UML fit using categorized-BDT approach



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

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

Significances

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

$B_d \to \mu\mu$ 2.0 σ

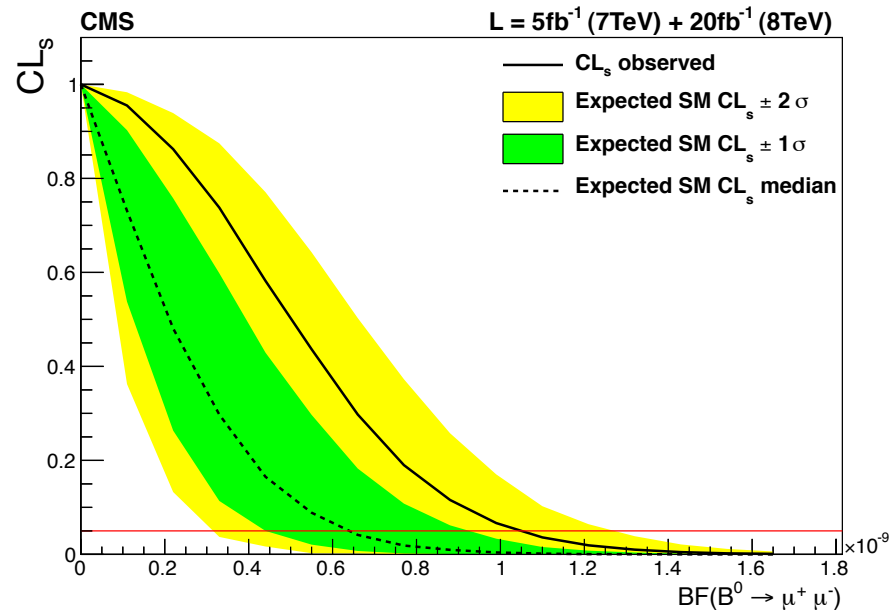
Upper limits on $B_d \to \mu\mu$

- No significant excess is observed for $B_d \to \mu\mu$
  - Upper limit 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	
	$B^0 \to \mu^+\mu^-$	$B_s^0 \to \mu^+\mu^-$	$B^0 \to \mu^+\mu^-$	$B_s^0 \to \mu^+\mu^-$
$\varepsilon_{tot}[\%]$	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	
	$B^0 \to \mu^+\mu^-$	$B_s^0 \to \mu^+\mu^-$	$B^0 \to \mu^+\mu^-$	$B_s^0 \to \mu^+\mu^-$
$\varepsilon_{tot}[\%]$	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

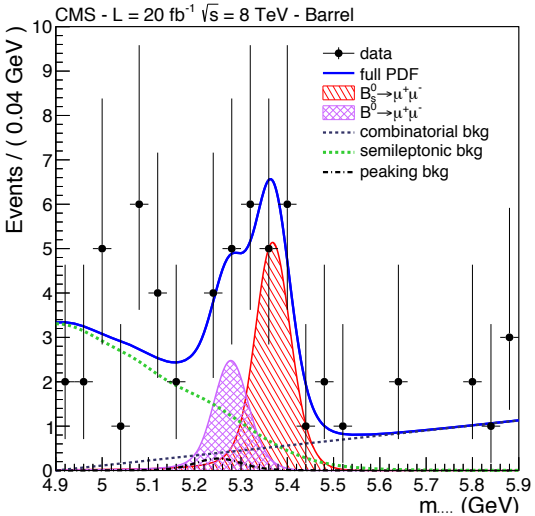


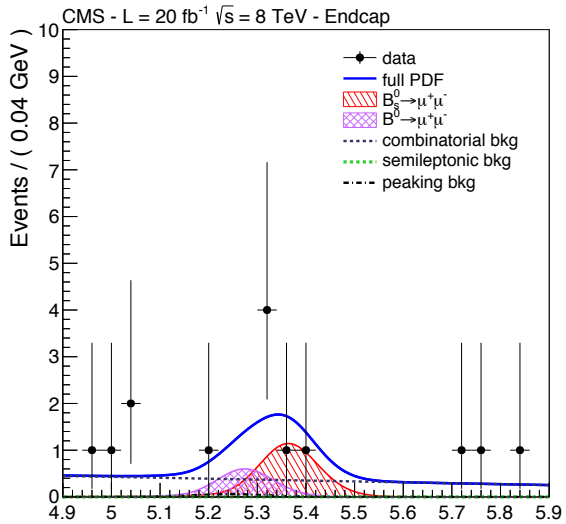
$$BR(B_d \to \mu\mu) < 1.1 \times 10^{-9} \text{ @95\% CL}$$

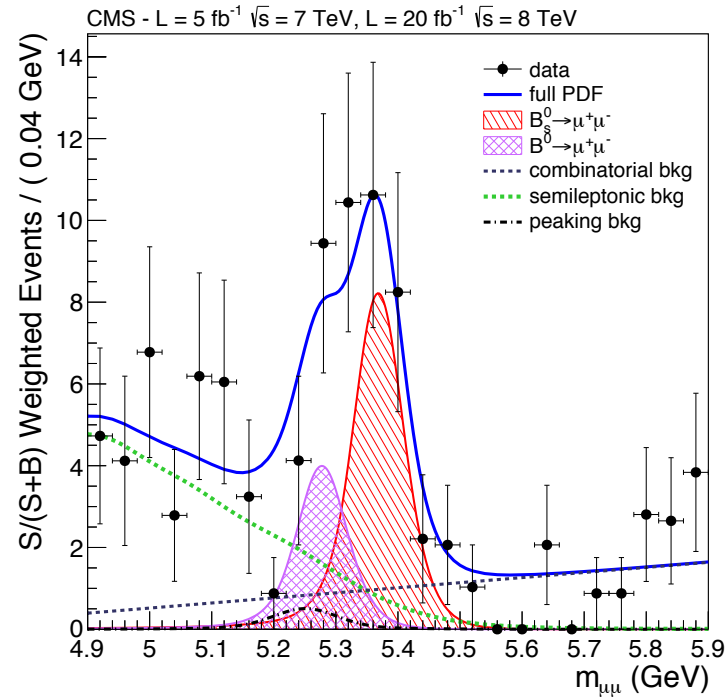
(expected 6.3×10^{-10} in presence of SM+background)

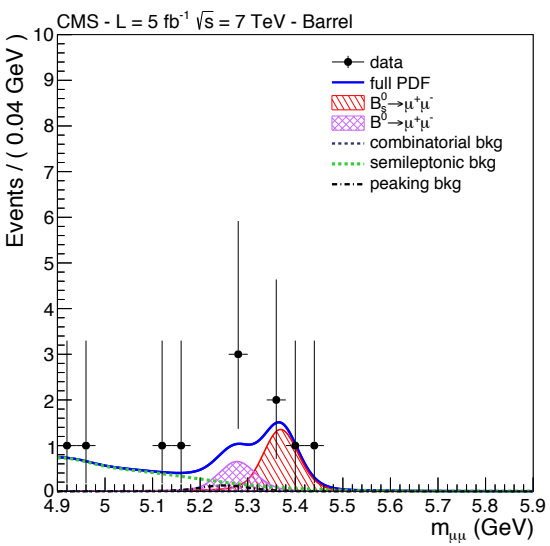
$$BR(B_d \to \mu\mu) < 9.2 \times 10^{-10} \text{ @90\% CL}$$

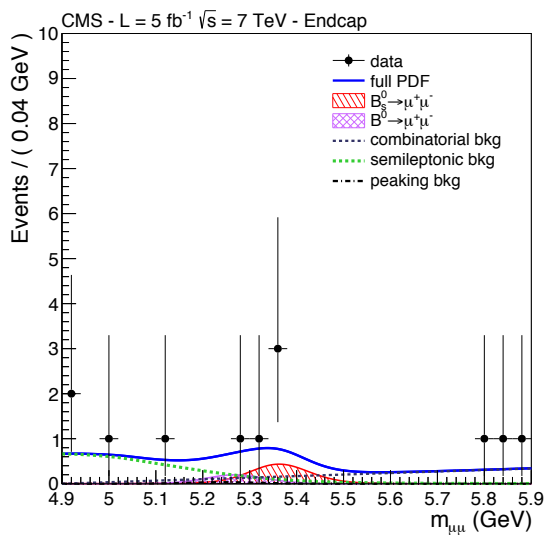
Cross-checked 1D-BDT results











Significance

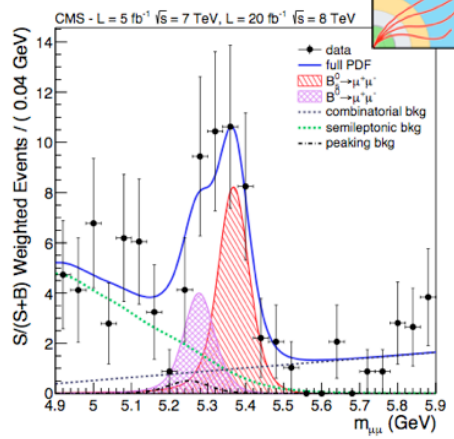
$B_s \to \mu\mu$ 4.8 σ (expected 4.7 σ median)

Less sensitive wrt BDT-categories used as a cross check

Combination with LHCb results

- At ICHEP 2013, CMS and LHCb results combined taking into account f_d/f_s correlation only

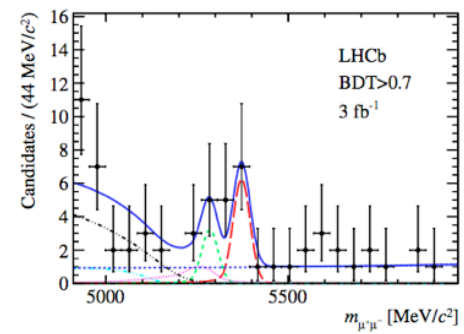
[PRL 111 (2013) 101804]



$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = 3.0^{+1.0}_{-0.9} \times 10^{-9} \quad (4.3\sigma)$$

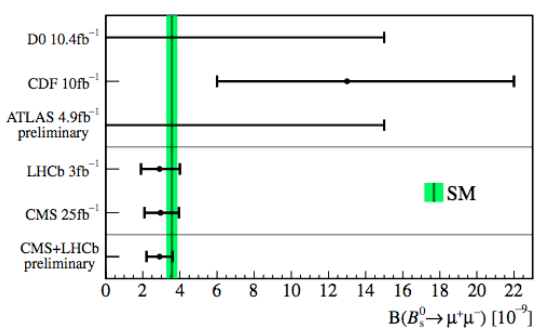
$$\mathcal{B}(B^0 \to \mu^+\mu^-) = 3.5^{+2.1}_{-1.8} \times 10^{-10} \quad (2.0\sigma)$$

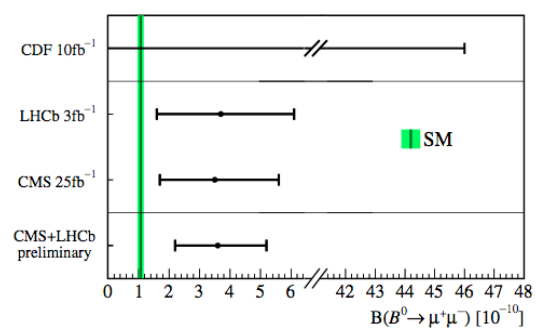
[PRL 111 (2013) 101805]



$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = 2.9^{+1.1}_{-1.0} \times 10^{-9} \quad (4.0\sigma)$$

$$\mathcal{B}(B^0 \to \mu^+\mu^-) = 3.7^{+2.4}_{-2.1} \times 10^{-10} \quad (2.0\sigma)$$





$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = 2.9 \pm 0.7 \times 10^{-9}$$

$$\mathcal{B}(B^0 \to \mu^+\mu^-) = 3.6^{+1.6}_{-1.4} \times 10^{-10}$$

[CMS-PAS-BPH-13-007]
[LHCb-CONF-2013-012]

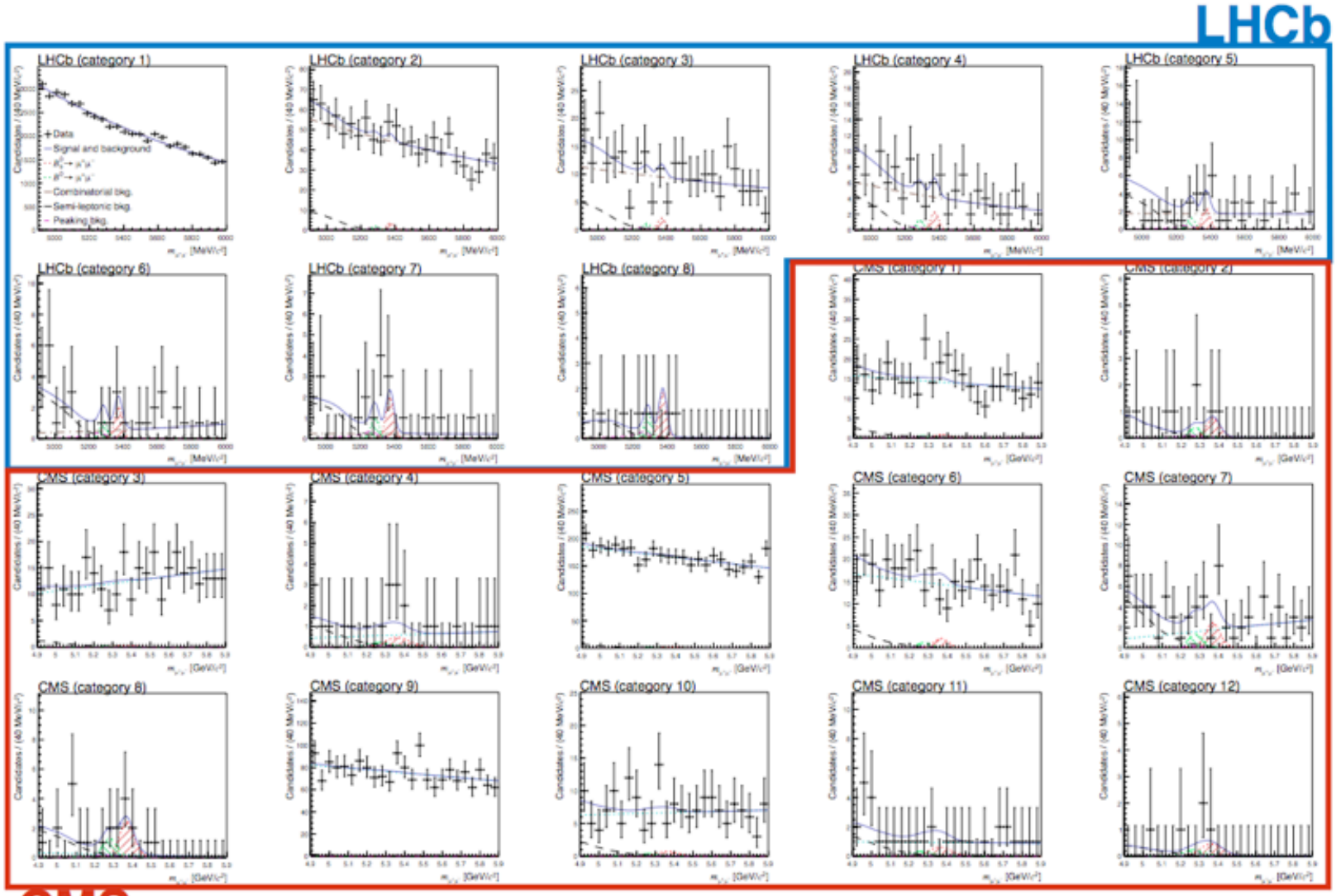
Full likelihood combination

- Full combination of CMS and LHCb results
  - Estimate significance
  - Take into account full correlations

NEW

- Changes to harmonize the analyses
  - $\Lambda^0 \rightarrow p\mu\nu$ background source was included in the nominal fit, with updated BF and MC simulation changed to include a more realistic model for the properties of the decay
  - Lifetime bias correction calculated and applied (already included in LHCb analysis)
  - Updated value of the ratio of hadronization fractions $f_d/f_s=3.86\pm0.22$
- The two datasets are used together in a single combined experiment
  - Simultaneous unbinned extended maximum likelihood fit to the mass spectra in 8 BDT bins for LHCb and 12 categories for CMS
  - Shared parameters
    - the branching fraction of the two signals $B(B_s\rightarrow\mu^+\mu^-)$ and $B(B_0\rightarrow\mu^+\mu^-)$
    - the branching fraction of the normalisation channel $B(B^+\rightarrow J/\psi K^+)$
    - the ratio of the hadronisation fractions f_d/f_s

Simultaneous fit

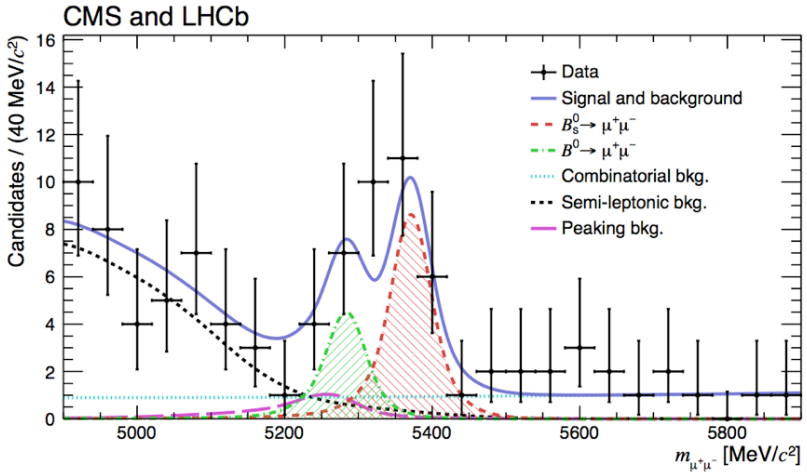


Fit results

- From the simultaneous fit we get

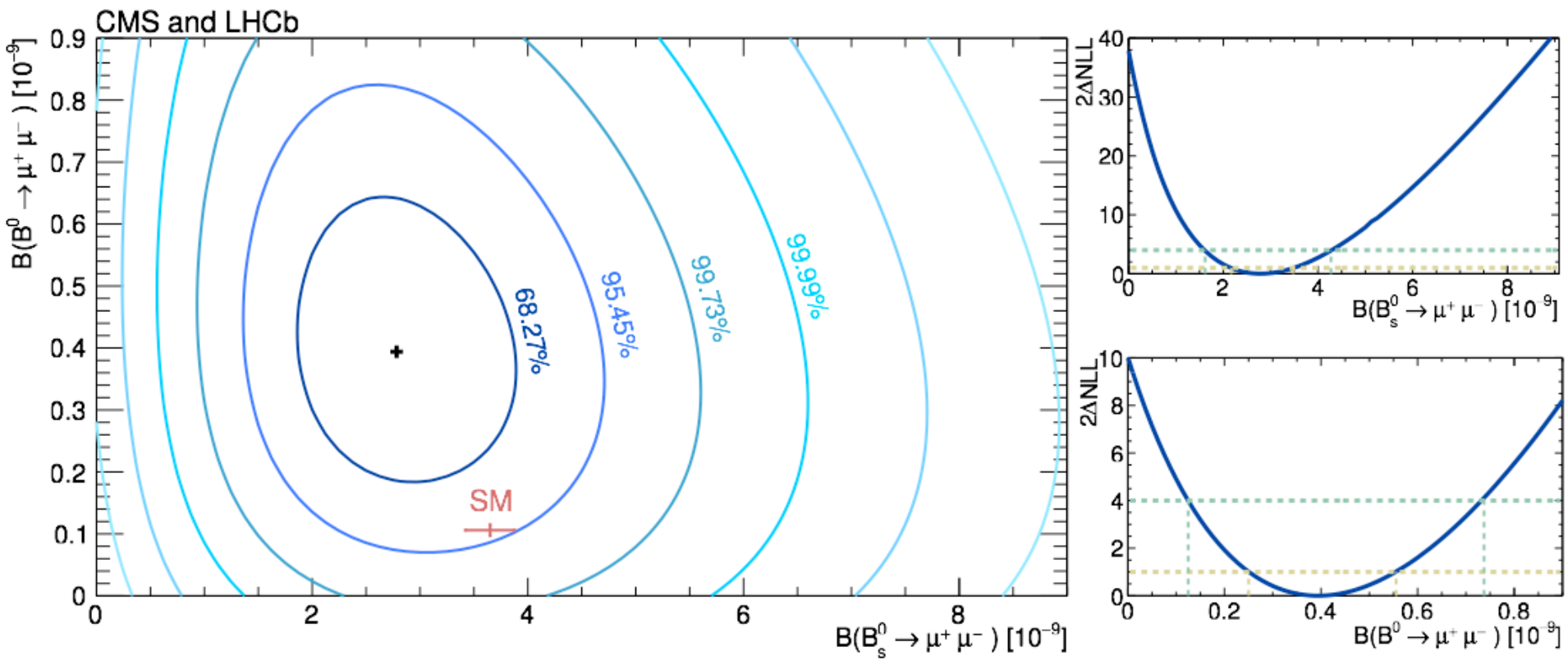
$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = 2.8^{+0.7}_{-0.6} \times 10^{-9}$$

$$\mathcal{B}(B^0 \to \mu^+\mu^-) = 3.9^{+1.6}_{-1.4} \times 10^{-10}$$

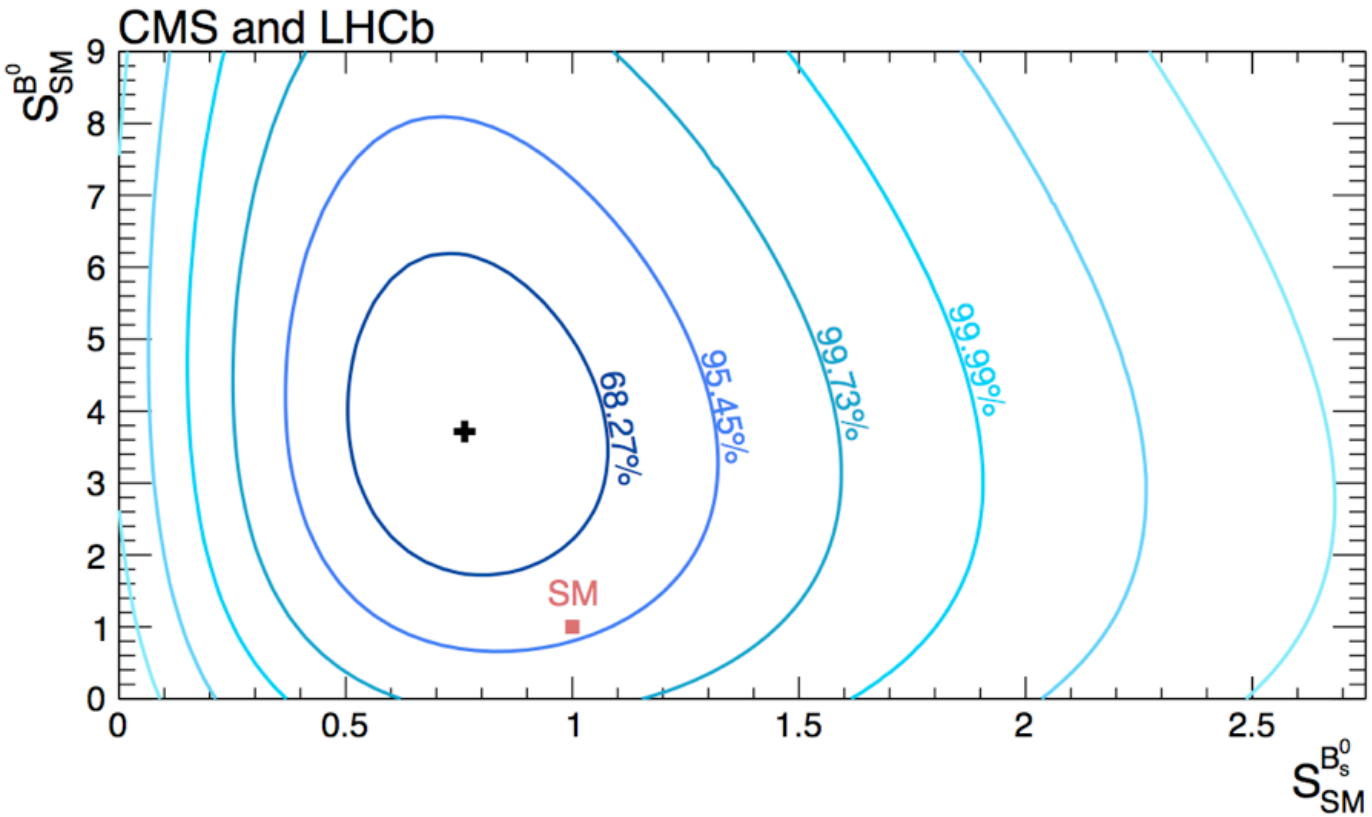


- From Wilk's theorem the statistical significances from the fit are
  - 6.2 σ (expected SM 7.6 σ) for Bs
  - 3.2 σ (expected SM 0.8 σ) for Bd
    - Since Wilk's theorem assumes asymptotic behavior, Feldman-Cousin approach used for Bd, yielding a p-value of 1.34×10^{-3} corresponding to a 3.0 σ significance

Likelihood scan of the BR



Simultaneous fit of the ratios of BR to the SM values



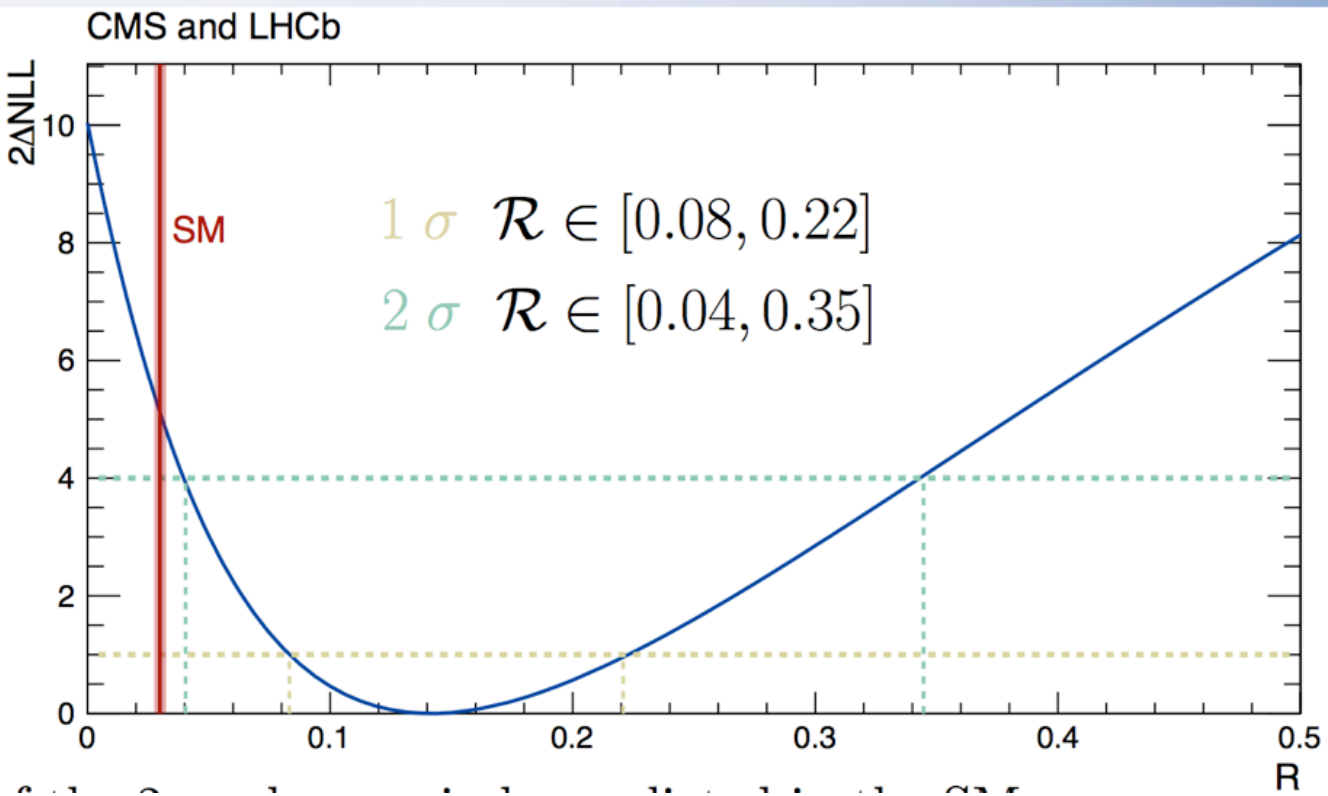
$$\mathcal{S}^{B^0_{(s)}}_{\rm SM} = \frac{\mathcal{B}(B^0_{(s)}\to\mu^+\mu^-)}{\mathcal{B}(B^0_{(s)}\to\mu^+\mu^-)^{\rm SM}}$$

$$\mathcal{S}^{B^0_s}_{\rm SM} = 0.76^{+0.20}_{-0.18}$$

$$\mathcal{S}^{B^0}_{\rm SM} = 3.7^{+1.6}_{-1.4}$$

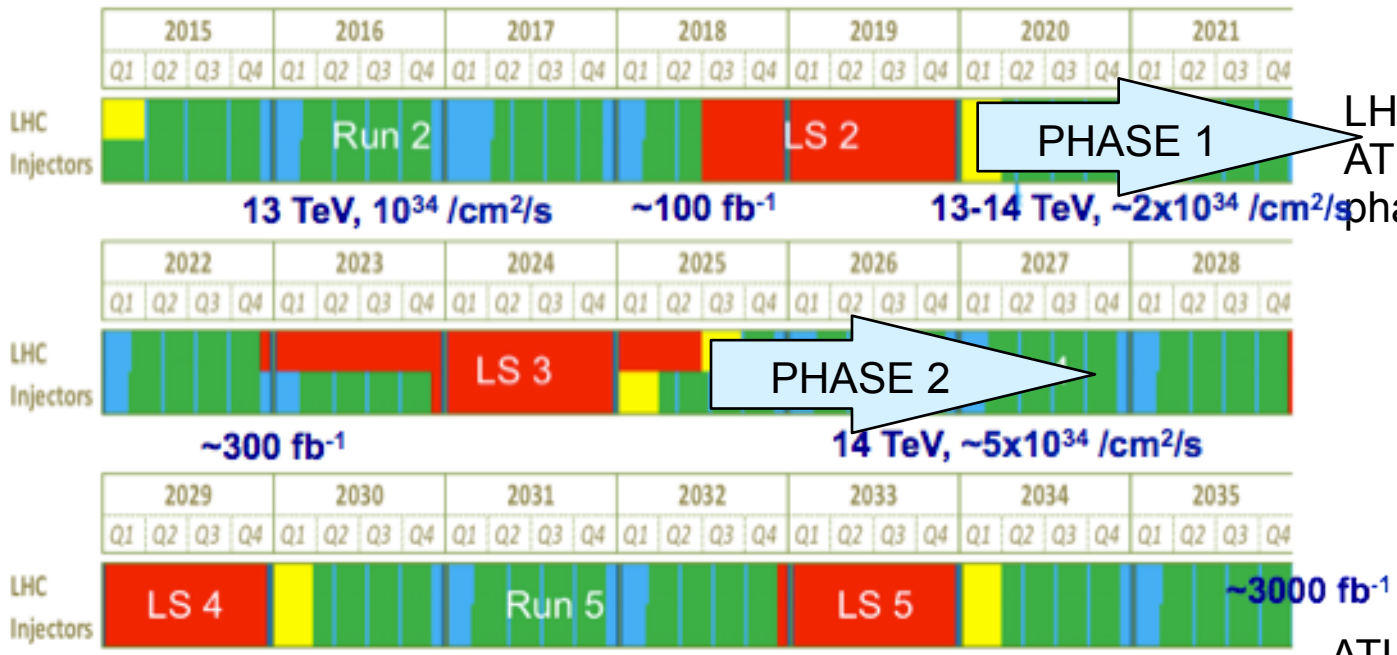
- Compatibility with the SM: 1.2 σ for Bs and 2.2 σ for Bd.
  - Including theoretical errors

Likelihood profile of B(Bs)/B(Bd)



- Ratio of the two BR precisely predicted in the SM
- Fit result R= 0.14+0.08-0.06
  - compatible with the SM at 2.3 σ (including theoretical errors)

The HL-LHC scenario



LHCb Upgrade
ATLAS/CMS Upgrades phase 1

HL-LHC
14 TeV
250 fb-1/year

ATLAS/CMS
Upgrades phase 2

At HL-LHC ~140 events/bx spread over ±15 cm (3 σ)
~3000 fb-1 expected in 10 years running

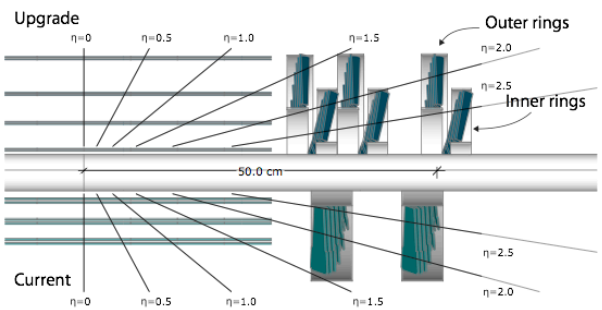
Excellent opportunity to search for (rare) and new phenomena

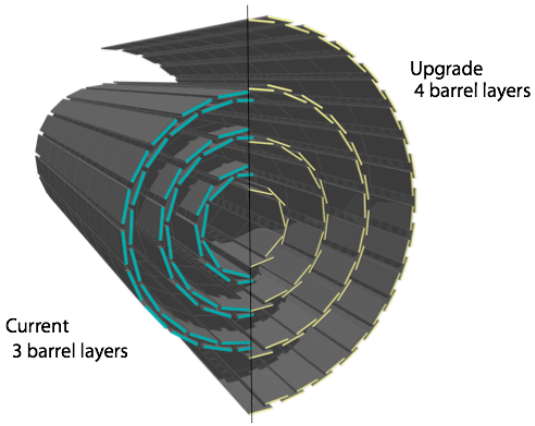
Tracker needs to be rebuilt

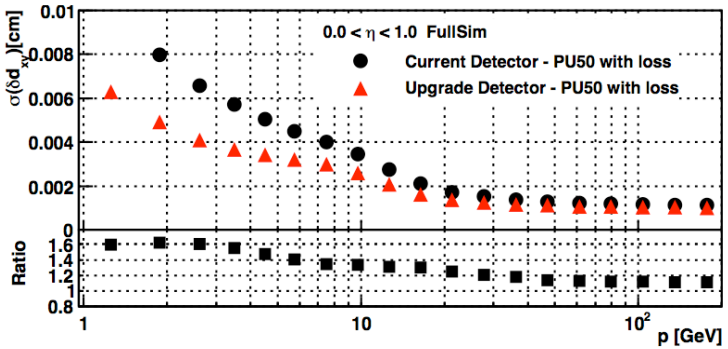
CMS Phase 1 upgrade

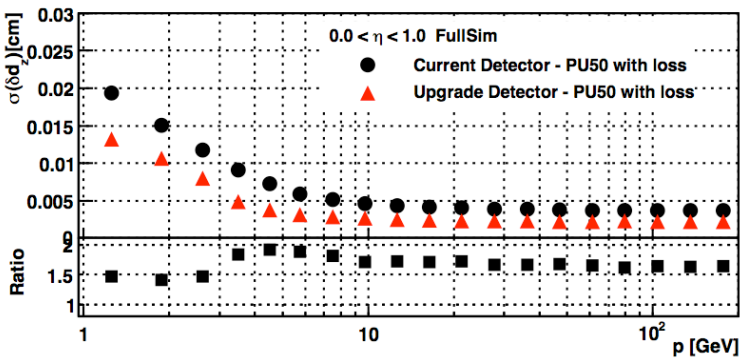
- In Phase 1 the pixel detector will be replaced
  - One more layer, closer to the beam pipe - more precise
  - Lower material budget - smaller multiple scattering
- Overall better performance for low pT tracks, used for B-physics

CERN-LHCC-2012-016
http://cds.cern.ch/record/1481838?ln=en

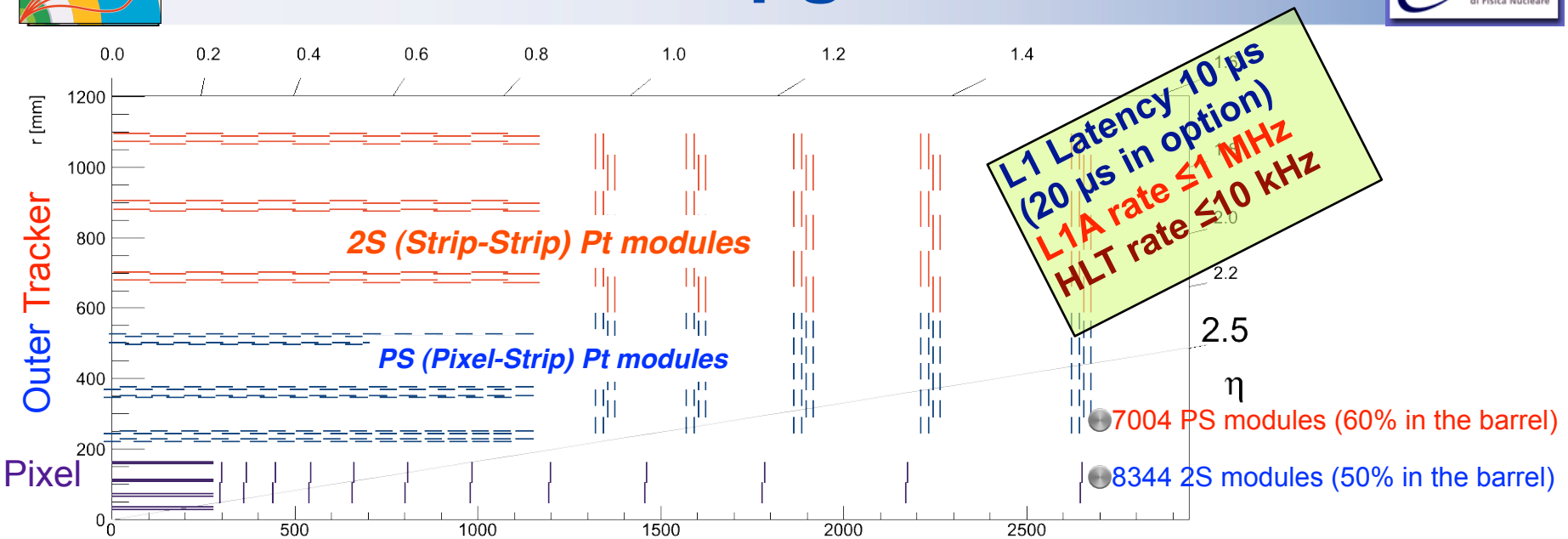








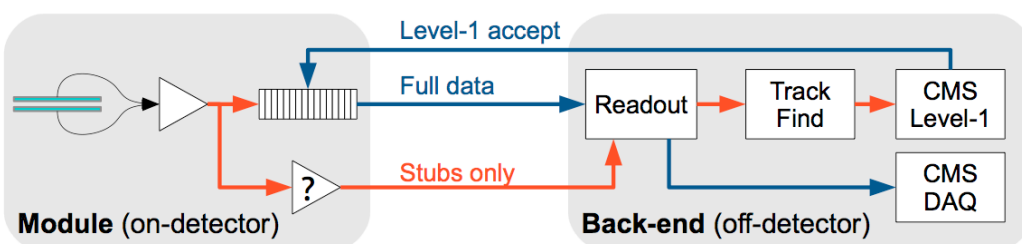
CMS Phase 2 upgraded Tracker

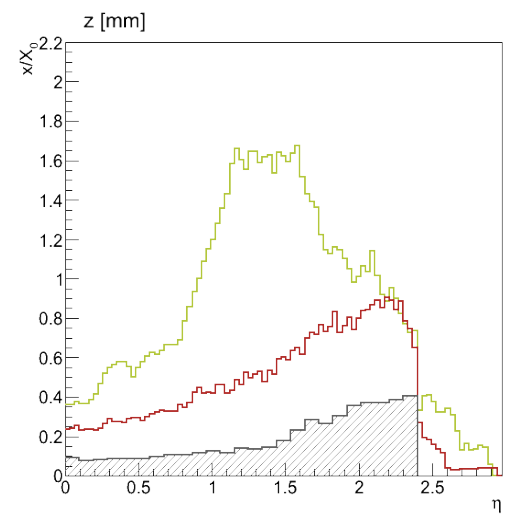


Better p_T resolution and lighter than current tracker

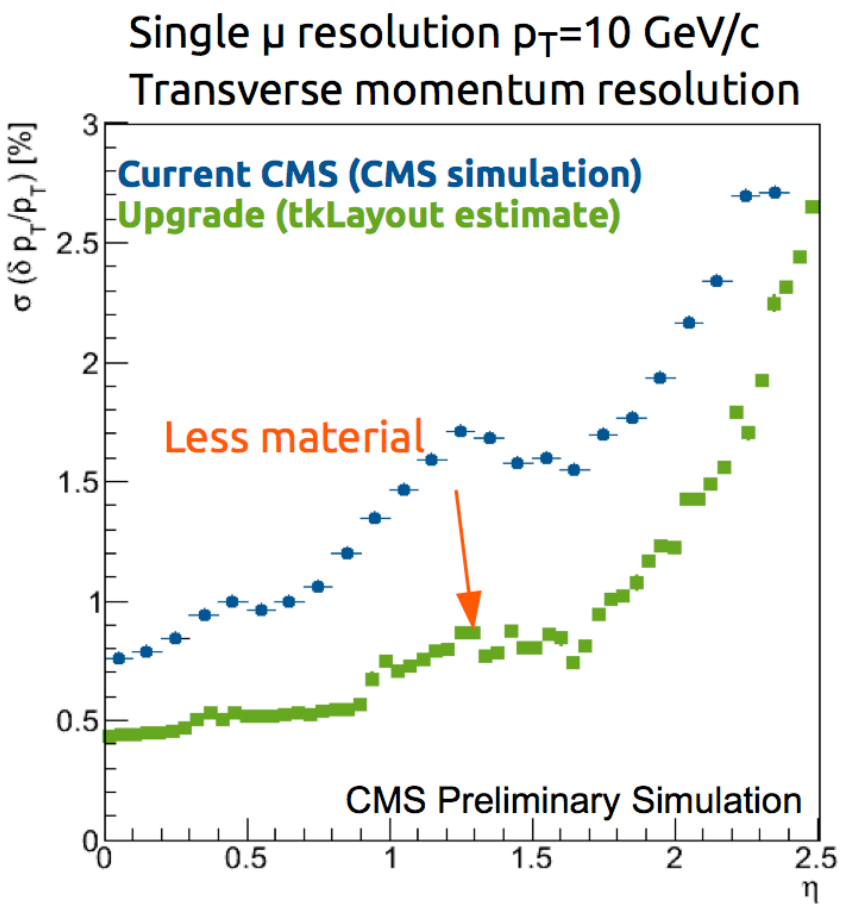
L1 Trigger functionality down to ~ 2 GeV p_T tracks

Readout and Trigger schematics





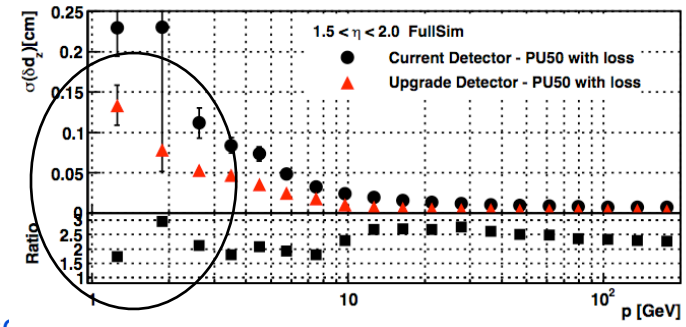
Offline tracking resolution



Clear improvement expected in the whole η range

Working scenario

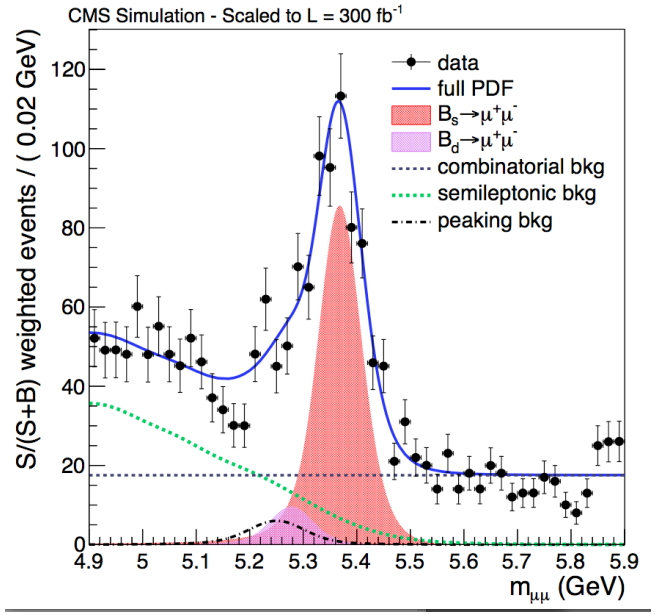
- Assumptions
  - Same trigger thresholds as the current one, in particular the $p_T(\mu)$3 GeV at L1
    - Feasible with the L1-Track Trigger
  - Reduced efficiencies due to 140 pileup events expected
    - Assume 30% loss in the isolation efficiency
    - 2.5% loss per μ reconstruction efficiency
    - The resolution in the endcap (0.5 to 1 mm) is comparable to the average vertex separation
      - Reduced impact of this region in the analysis in discerning Bd from Bs peak.



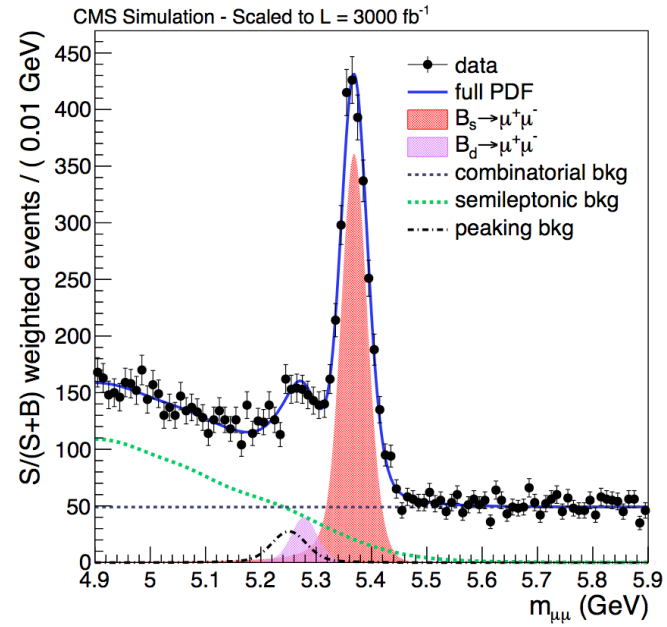
  - Systematics
    - fs/fu = 5% [now is 9%]
    - Normalization (still assume $B^{\pm}\rightarrow J/\psi K^{\pm}$) = 3% ⊕ 5%/√ (L/20 fb^{-1}) [now 5% from yields and 3% from BR]
    - Peaking background uncertainty = 10% ⊕ 50%/√(L/20 fb^{-1})
    - Semileptonic background uncertainty = 20% ⊕ 50%/√(L/20 fb^{-1})

- Detector resolutions
  - Tracker phase 2 detector has better p_T resolution
    - a factor of ~1.6 in the barrel and ~1.2 in the endcap
  - Do not take into account the better resolution in the impact parameter due to 1st layer closer to the beam-pipe and (possibly) smaller pitch

Effects on $B_{s/d} \to \mu\mu$



CMS PAS
FTR-13-022



Improved Tracker

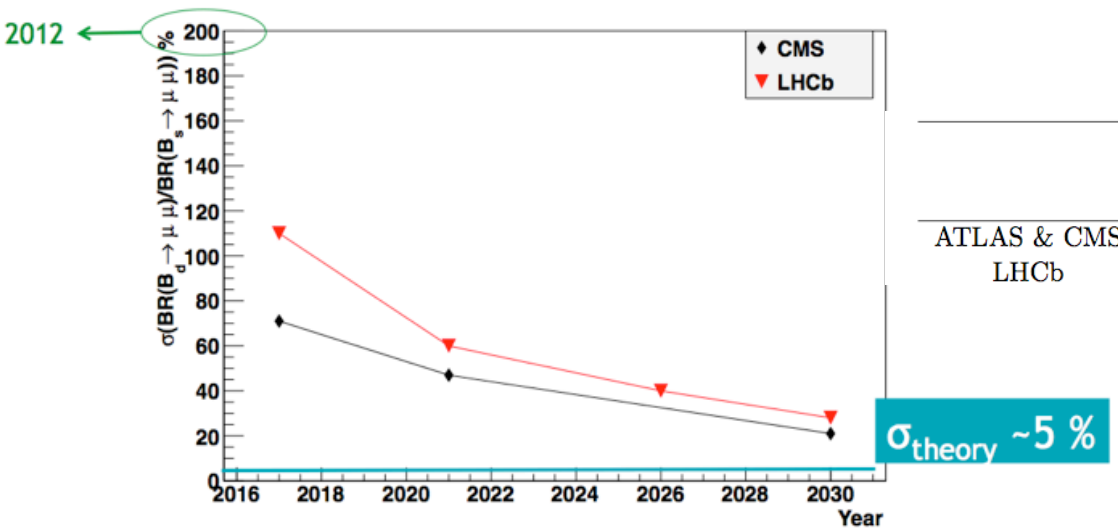
L (fb^{-1})	No. of B_s^0	No. of B^0	$\delta\mathcal{B}/\mathcal{B}(B_s{}^0 \to \mu^+\mu^-)$	$\delta\mathcal{B}/\mathcal{B}(B^0 \to \mu^+\mu^-)$	B^0 sign.	$\delta\frac{\mathcal{B}(B^0\to\mu^+\mu^-)}{\mathcal{B}(B_s^0\to\mu^+\mu^-)}$
20	16.5	2.0	35%	>100%	0.0–1.5 σ	>100%
100	144	18	15%	66%	0.5–2.4 σ	71%
300	433	54	12%	45%	1.3–3.3 σ	47%
3000	2096	256	12%	18%	5.4–7.6 σ	21%

CMS and LHCb vs time

- Conservatively assuming no changes in the analyses and SM BR
  - See ECFA Workshop Report: https://cds.cern.ch/record/1631032
  - Experiments could do better than this!

		LHC era			HL-LHC era	
		Run 1	Run 2	Run 3	Run 4	Run 5+
$\frac{\mathcal{B}(B^0\to\mu^+\mu^-)}{\mathcal{B}(B_s^0\to\mu^+\mu^-)}$	CMS	$> 100\%$	71%	47%	...	21%
	LHCb	220%	110%	60%	40%	28%

Expected precision on $BR(B_d\to\mu^+\mu^-)/BR(B_s\to\mu^+\mu^-)$



	LHC era			HL-LHC era	
	Run 1 (2010–12)	Run 2 (2015–17)	Run 3 (2019–21)	Run 4 (2024–26)	Run 5+ (2028–30+)
ATLAS & CMS	$25\,\mathrm{fb}^{-1}$	$100\,\mathrm{fb}^{-1}$	$300\,\mathrm{fb}^{-1}$	$\longrightarrow$	$3000\,\mathrm{fb}^{-1}$
LHCb	$3\,\mathrm{fb}^{-1}$	$8\,\mathrm{fb}^{-1}$	$23\,\mathrm{fb}^{-1}$	$46\,\mathrm{fb}^{-1}$	$100\,\mathrm{fb}^{-1}$

Conclusions and prospects

- CMS measured the BR($B_s \to \mu\mu$) and put an upper limit on BR($B_d \to \mu\mu$)
  - Final LHC Run1 statistics ~25 fb-1, statistically dominated
  - 4.3 σ significance (4.8 σ expected) - Consistent with the SM expectations
- Combined CMS+LHCb average results in agreement with the SM
  - BR($B_s \to \mu\mu$) with 6.2 σ significance, compatibility with the SM at 1.2 σ
  - 3 σ evidence for an excess of events in the search for $B_d \to \mu\mu$
    - this is compatible with the SM at 2.2 σ
  - Compatibility of the ratio of the BR with SM at the 2.3 σ level
- Future prospects
  - CMS, LHCb (and ATLAS) will continue to improve the precision on the BR($B_s \to \mu\mu$) and BR($B_d \to \mu\mu$)
    - Experiments are preparing the Run2 data taking and analyses ... Stay tuned!

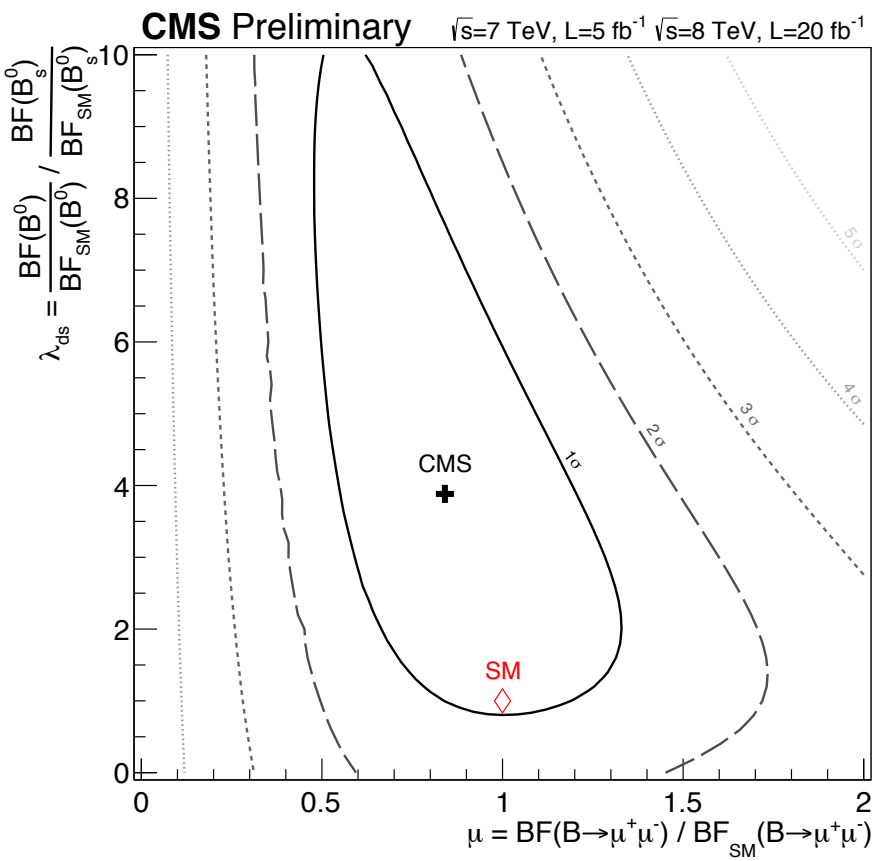
Backup

Alternative parametrization for SM compatibility check

- Test how the BRs are compatible with the SM expectations
  - $BR_{SM}(B_s \to \mu\mu) = (3.56 \pm 0.18) \times 10^{-9}$
  - $BR_{SM}(B_d \to \mu\mu) = (1.07 \pm 0.10) \times 10^{-10}$

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

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



Simultaneous fit

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

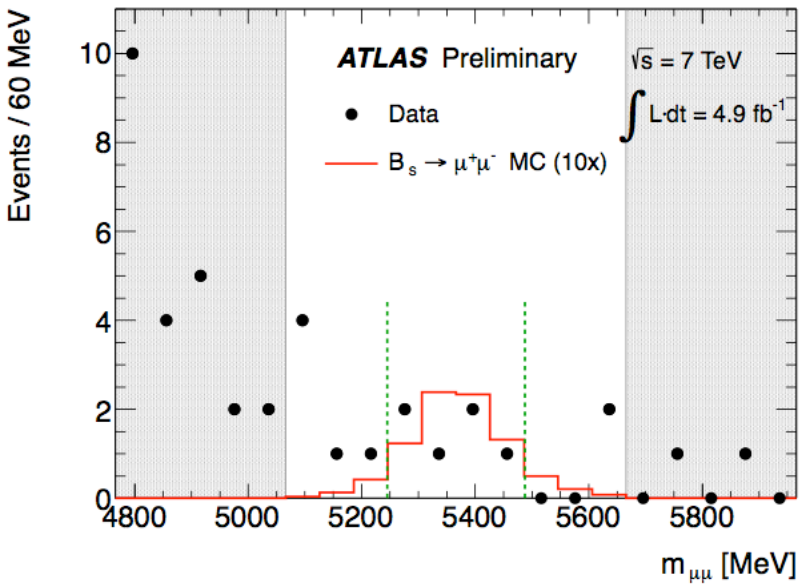
Fit for μ (fix lambda to SM)

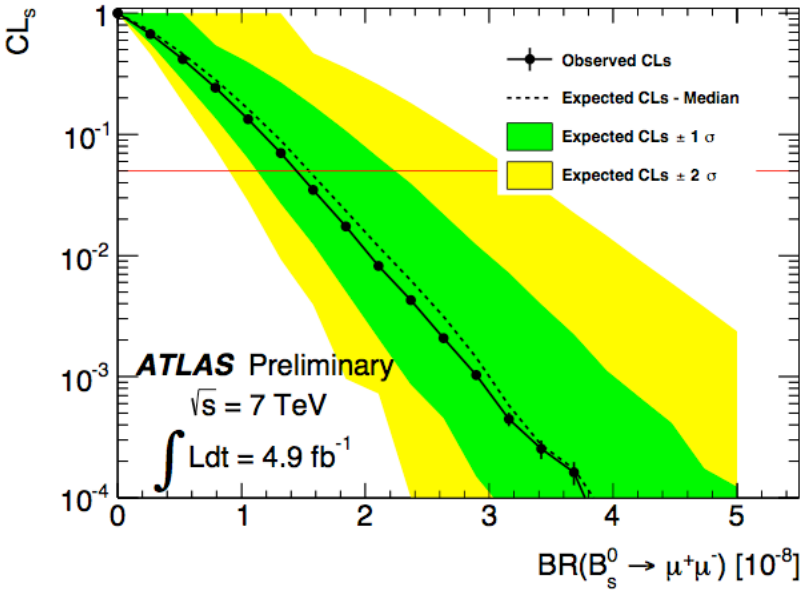
$$\mu = 1.01^{+0.31}_{-0.26}$$

Fit for λ_{ds} (fix μ to SM)

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

ATLAS

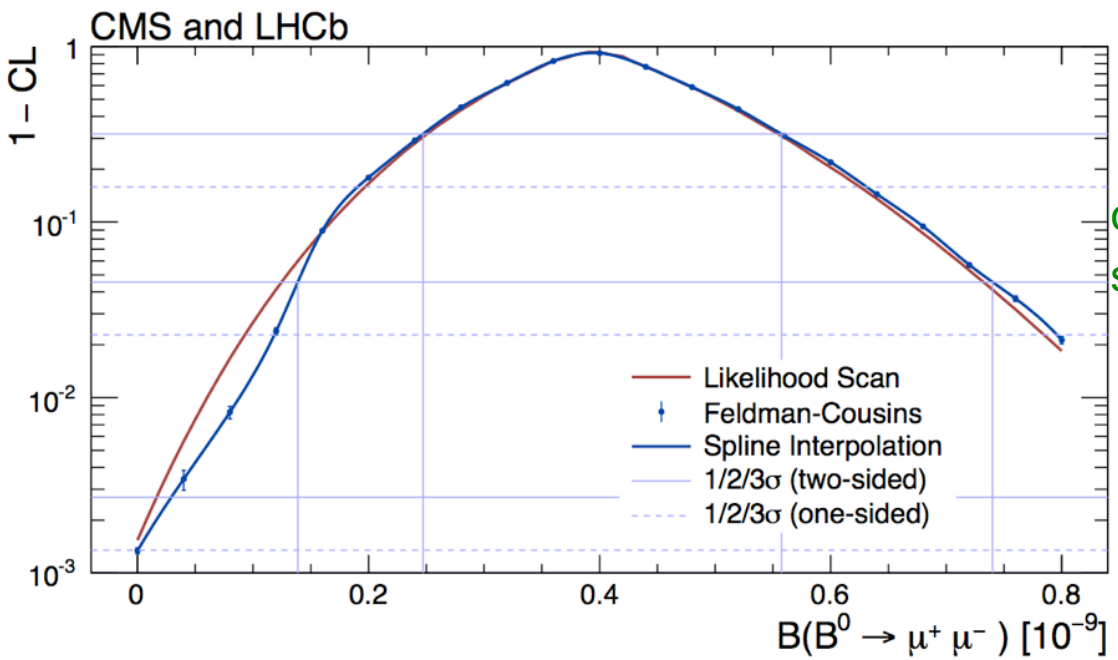

ATLAS Preliminary
√s = 7 TeV
∫ L·dt = 4.9 fb⁻¹
Data
B s → μ⁺μ⁻ MC (10x)
Events / 60 MeV
m μμ [MeV]


Observed CLs
Expected CLs - Median
Expected CLs ± 1 σ
Expected CLs ± 2 σ
ATLAS Preliminary
√s = 7 TeV
∫ Ldt = 4.9 fb⁻¹
CLs
BR(B⁰s → μ⁺μ⁻) [10⁻⁸]

CMS vs LHCb

- Muon misidentifications
  - CMS: $\varepsilon(\pi \to \mu)$~0.15%, $\varepsilon(K \to \mu)$~0.20%, $\varepsilon(p \to \mu)$~0.10%
  - LHCb: $\varepsilon(\pi \to \mu)$~0.6%, $\varepsilon(K \to \mu)$~0.4%, $\varepsilon(p \to \mu)$~0.3%
- Mass resolutions
  - CMS: 32 MeV (Barrel) 75 MeV (Endcap)
  - LHCb: 25 MeV
- Pileup
  - CMS:~9 (2011) ~21 (2012)
  - LHCb:~1.7

Feldman-Cousin for B^0 significance



- p-value of 1.34×10^{-3} corresponding to a 3.0 σ significance

$$1\,\sigma \quad \mathcal{B}(B^0 \to \mu^+\mu^-) \in [2.50, 5.55] \times 10^{-10}$$

$$2\,\sigma \quad \mathcal{B}(B^0 \to \mu^+\mu^-) \in [1.38, 7.40] \times 10^{-10}$$

Lifetime bias correction

- $B_s \to \mu^+\mu^-$ time dependent width:

$$\Gamma(B_s^0(t) \to \mu^+\mu^-) = (R_H + R_L)e^{-\Gamma_s t}\left[\cosh\frac{y_s t}{\tau_{B_s}} + \mathcal{A}_{\Delta\Gamma}\sinh\frac{y_s t}{\tau_{B_s}}\right]$$

where:

$$y_s = \frac{\Gamma_L - \Gamma_H}{\Gamma_L + \Gamma_H} = 0.01615 \pm 0.0085 \qquad \mathcal{A}_{\Delta\Gamma} = \frac{\Gamma_{B^0_{s,H}\to\mu\mu} - \Gamma_{B^0_{s,L}\to\mu\mu}}{\Gamma_{B^0_{s,H}\to\mu\mu} + \Gamma_{B^0_{s,L}\to\mu\mu}} \overset{SM}{=} 1$$

- So the time integrated efficiency is model dependent: $\varepsilon = \frac{\int \varepsilon(t)\Gamma_{\mathcal{A},y_s}(t)dt}{\int \Gamma_{\mathcal{A},y_s}(t)dt}$

- Normalization to be corrected to take into account this effect:

$$\delta_\epsilon = \frac{\epsilon^{\mathcal{A}_{\Delta\Gamma},y_s}}{\epsilon^{MC}} = \frac{\int_0^\infty \Gamma(B_s^0(t)\to\mu^+\mu^-,\mathcal{A}_{\Delta\Gamma},y_s)\epsilon(t)\mathrm{d}t}{\int_0^\infty \Gamma(B_s^0(t)\to\mu^+\mu^-,\mathcal{A}_{\Delta\Gamma},y_s)\mathrm{d}t}\cdot\frac{\int_0^\infty e^{-\Gamma_{MC}t}\mathrm{d}t}{\int_0^\infty e^{-\Gamma_{MC}t}\epsilon(t)\mathrm{d}t}$$

Correction for B_S = 4.50±0.03%
Correction for B^0 = 1.48±0.01%

- also corrected because time dependent

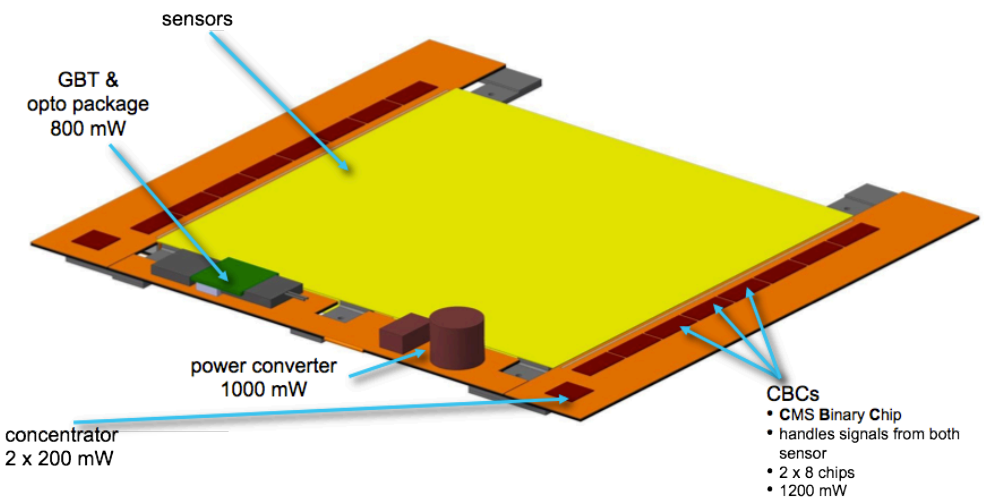
CMS p_T modules for L1 trigger

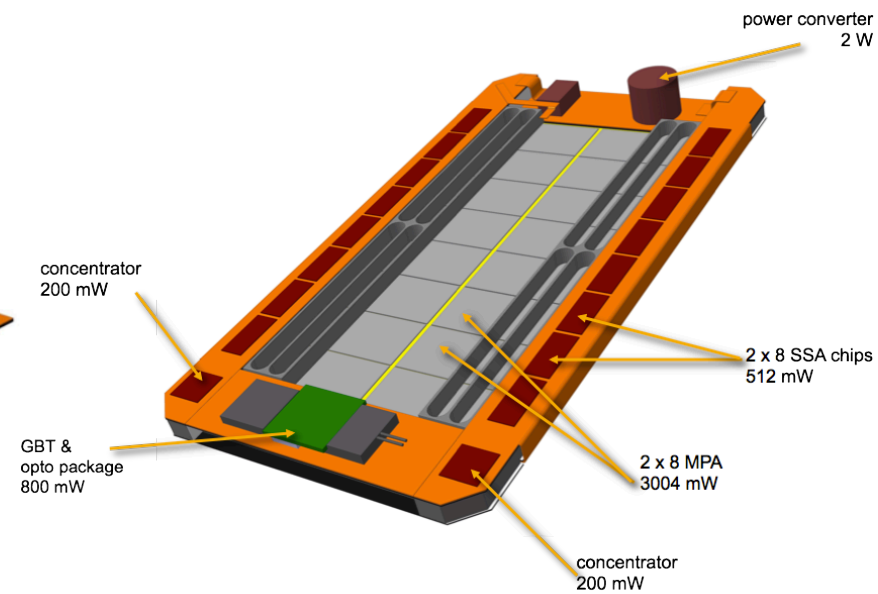
- 2S(trip) sensors modules

100 μm x 5 cm long strips on both sensors

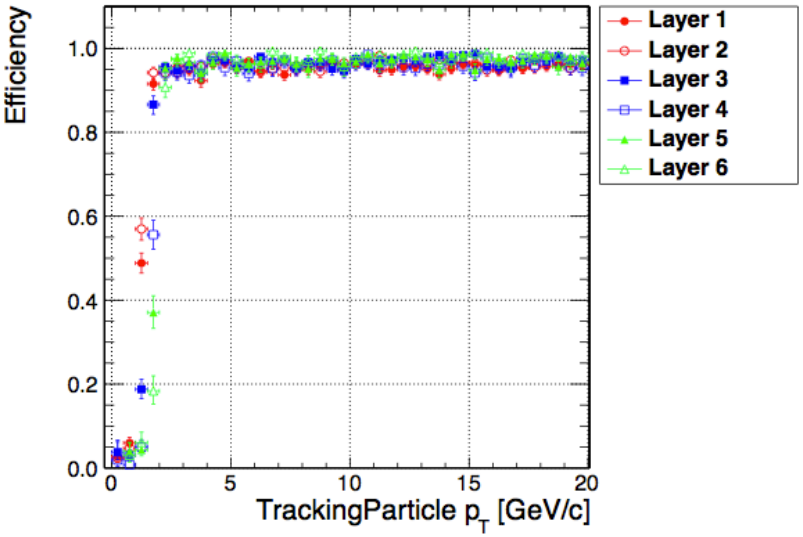
P(ixel)S(trip) module
strips = 100 μm x 2.4 cm
pixels = 100 μm x 1.5 mm
.

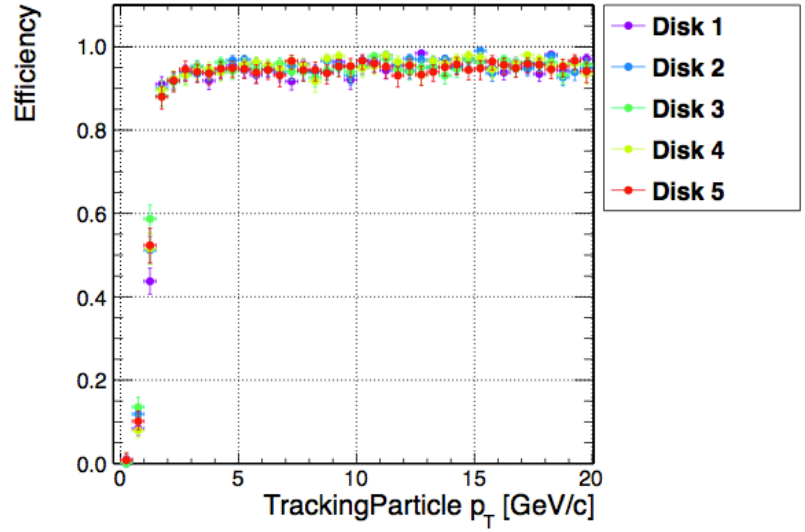
Pixels are logically OR-ed for finding coincidence in the r-ϕ plane, and the precise z-coordinate is retained in the pixel storage and provided to the trigger processors



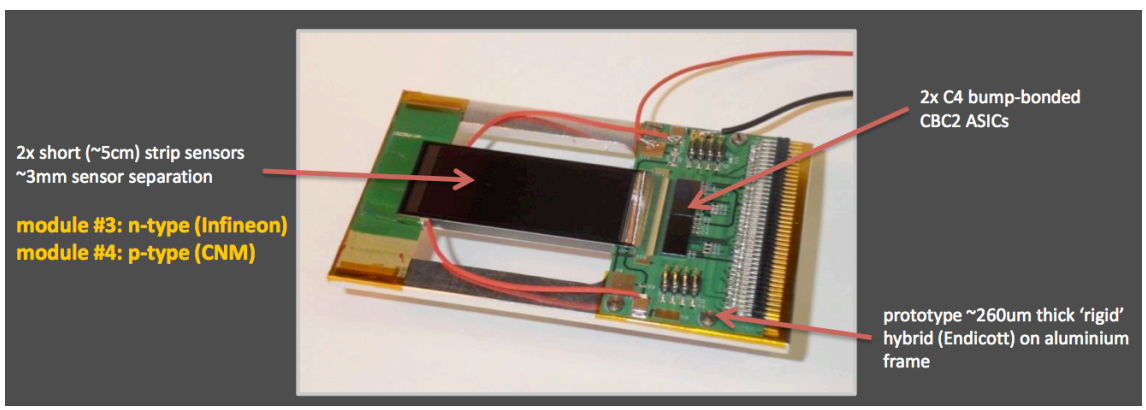


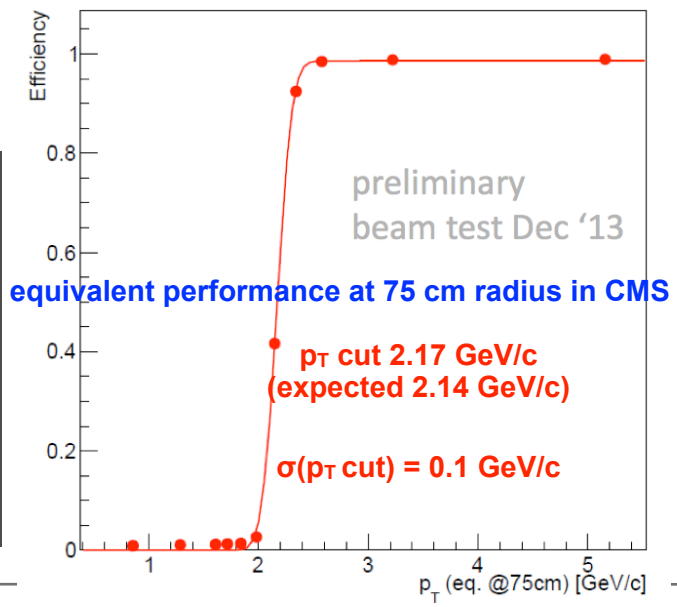
p_T modules performance



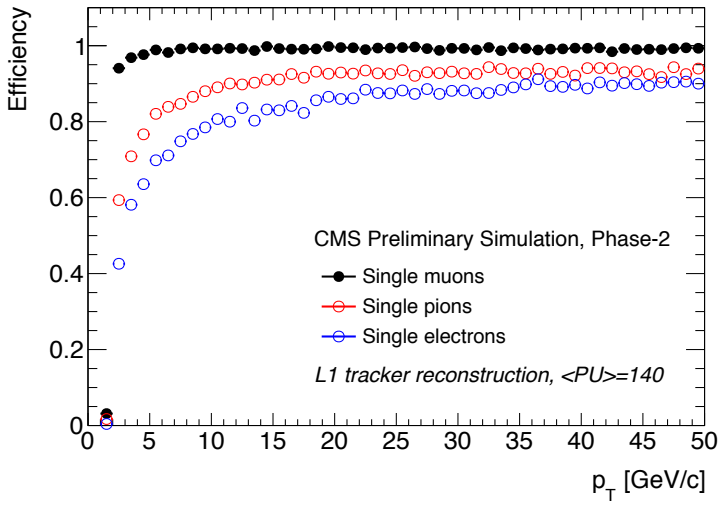


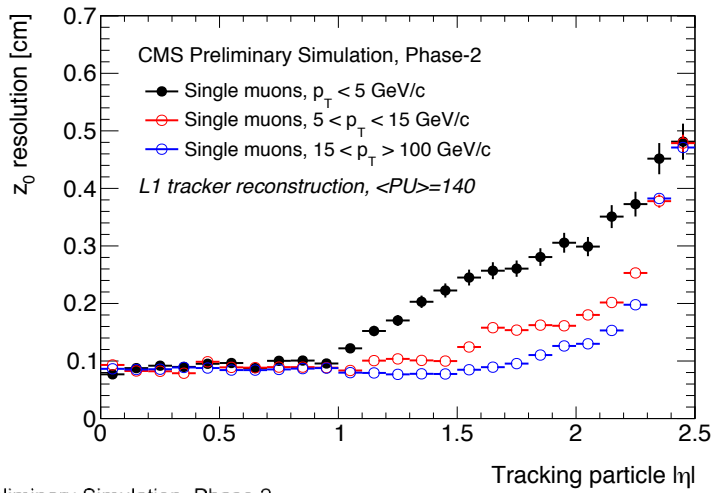
Prototype module test beam at DESY (2-4 GeV e^+)

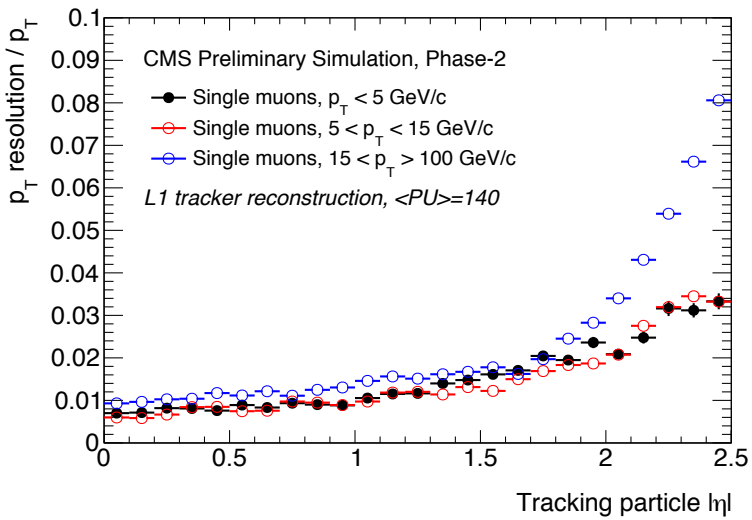


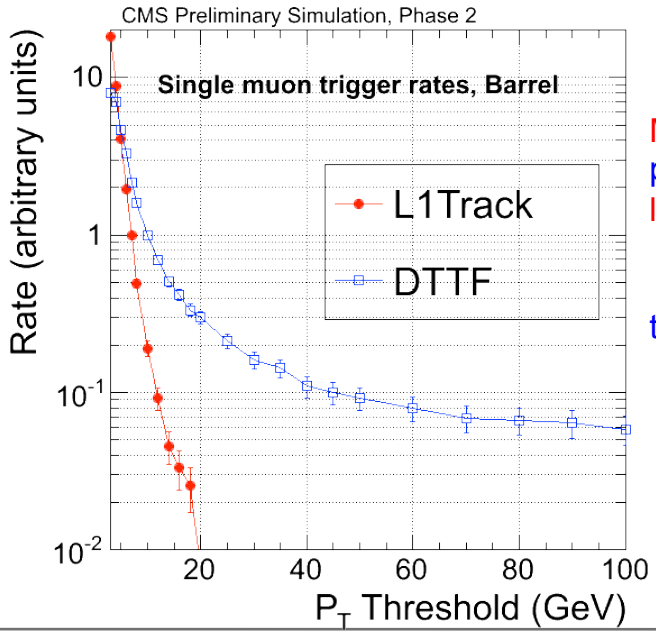


L1 tracking expected performances









Matching Drift Tube trigger primitives with L1Tracks: large rate reduction

Normalized to present trigger at 10 GeV.

Effects on $B_{s/d} \rightarrow \mu\mu$

- Chosen benchmark channel to illustrate benefits
  - Toy MC studies based on the published analysis (Phys. Rev. Lett. 111 (2013) 101804)
    - Average pileup in 2012 was 21.
  - Basic analysis relies on
    - Low μ trigger p_T thresholds (3 GeV at L1)
    - Muon and di-muon vertex isolation
    - Displaced vertex
    - Di-muon vertex aligned with flight path
    - Low muon fake rate $\varepsilon(\mu|\pi)$ 0.15%
    - All the above points will be possible with improved tracking detectors

CMS-PAS-FTR-13-022