Studies of rare B meson decays with the CMS detector

Jacopo Pazzini[1]
on behalf of the CMS collaboration

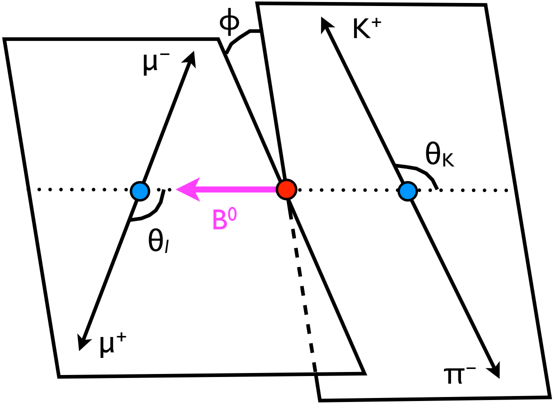
[1] Padova University & INFN

Rare *B* mesons decays are a crucial probe of physics Beyond the Standard Model

- FCNC, higly suppressed decays
- Virtual New Physics particles can contribute in loops
- Tiny effects so NP can be at the same level of SM
- High precision predictions of SM observables
- If observed, deviations from SM will be a indirect evidence of NP
  - Rare *B* decays $\Rightarrow$ potential smoking gun for physics BSM

In this talk:

- Angular analysis and branching ratio measurement of the decay $B^0 \rightarrow K^{*0}\mu^+\mu^-$
- Measurement of the $B_s^0 \rightarrow \mu^+\mu^-$ branching fraction and search for $B^0 \rightarrow \mu^+\mu^-$

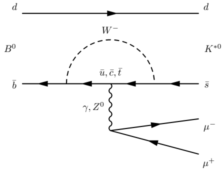


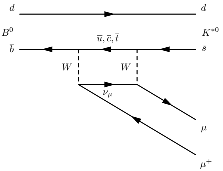
Accepted for Publication in Phys. Lett. B

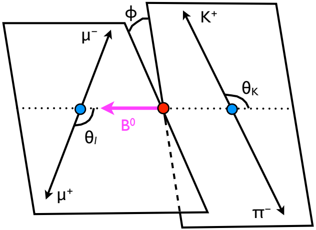
CMS-PAS-BPH-11-009

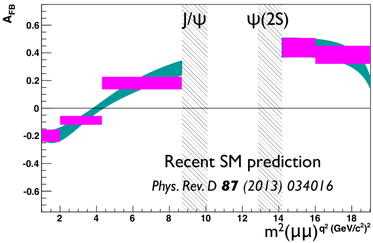
arXiv:1308.3409

- $b \to s\ell^+\ell^-$ transition, suppressed in the standard model:
  - ▷ FCNC
    - Tree-level suppression
    - Box/penguin diagrams
  - ▷ Robust calculations, especially at low q$^2 = m^2_{\mu\mu}$
- Four-particles final states ($K\pi\mu\mu$) $\to$ angular observables
  - ▷ Decay fully described by q^2 and three angles (θ_K,θ_L,ϕ)
- CMS measurements:
  - ▷ Differential branching fraction, d$\mathcal{B}$/dq^2
  - ▷ Dimuon forward-backward asymmetry, A_{FB}
  - ▷ Longitudinal K^{*0} polarization, F_L









- Full 2011 dataset: 5 fb^{-1}, $\sqrt{s} = 7$ TeV
- Unbinned extended maximum likelihood fits in q^2 bins:
  - 6 signal bins ($1 - 19$ GeV^2)
  - 1 normalization bin ($B^0 \to K^{*0} J/\psi$)
  - 1 control bin ($B^0 \to K^{*0} \psi'$)

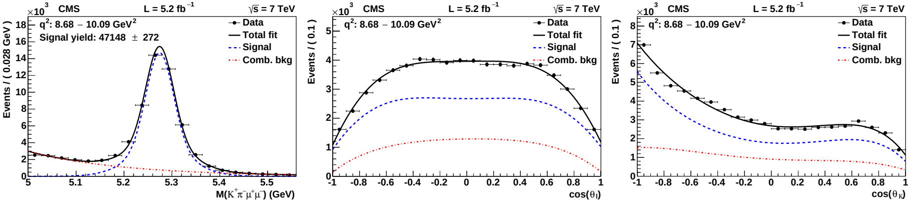
  } Signal and control channels selected/treated identically
- $\mathbf{A_{FB}}$ and $\mathbf{F_L}$ are extracted from fits to the two angular variables: $\cos\theta_K$, and $\cos\theta_L$.

$$\frac{1}{\Gamma}\frac{d^3\Gamma}{d\cos\theta_K\, d\cos\theta_L\, dq^2} = \frac{9}{16}\left\{\left[\frac{2}{3}\mathbf{F_S} + \frac{4}{3}\mathbf{A_S}\cos\theta_K\right]\left(1-\cos^2\theta_L\right)\right.$$
$$+ (1-\mathbf{F_S})\left[2\mathbf{F_L}\cos^2\theta_K\left(1-\cos^2\theta_L\right)\right.$$
$$+ \frac{1}{2}(1-\mathbf{F_L})\left(1-\cos^2\theta_K\right)\left(1+\cos^2\theta_L\right)$$
$$\left.\left. + \frac{4}{3}\mathbf{A_{FB}}\left(1-\cos^2\theta_K\right)\cos\theta_L\right]\right\}$$

- **Residual contributions** determined by the fit to the normalization sample $B^0 \to K^{*0} J/\psi$:
  - Spinless (S-wave) contribution to the $K\pi$ combination, $\mathbf{F_S}$
  - S-wave/P-wave interference, $\mathbf{A_S}$
- $d\mathcal{B}/dq^2$ is measured relative to the normalization channel $B^0 \to K^{*0} J/\psi$ from fit to the B^0 cand. invariant mass:

$$\frac{d\mathcal{B}\left(B^0 \to K^{*0}\mu^+\mu^-\right)}{dq^2} = \frac{Y_S}{Y_N}\frac{\epsilon_N}{\epsilon_S}\frac{\mathcal{B}\left(B^0 \to K^{*0} J/\psi\right)}{dq^2}$$

- Fit to the $B^0 \rightarrow K^{*0} J/\psi$ channel ($8.68 < q^2 < 10.09$ GeV^2 bin)



- Vector (P-wave) and scalar (S-wave) contributions obtained from the fit
  - ▷ Direct estimation of the non-resonant $K\pi$ component from data

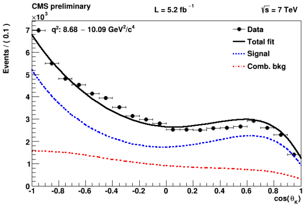
CMS results

- $F_L = 0.554 \pm 0.004$
- $A_{FB} = -0.004 \pm 0.004$

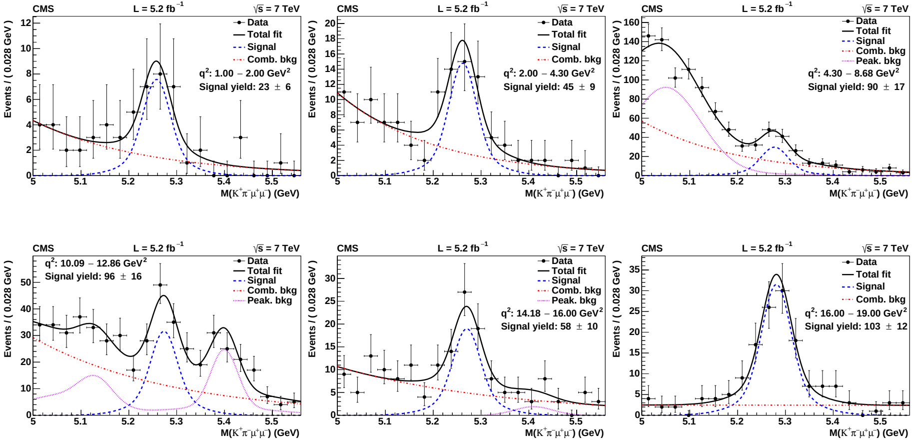
Expectations

- $F_L = 0.570 \pm 0.008$ [PDG]
- A_{FB} compatible with zero

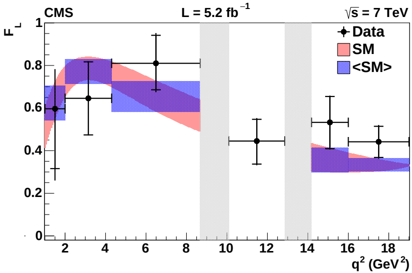
Fit performed neglecting the scalar $K\pi$ contributions (F_S,A_S) →

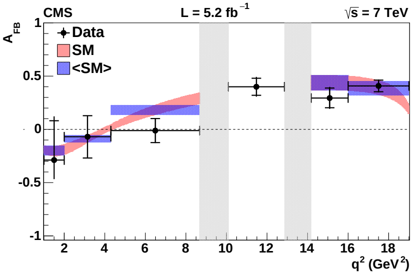


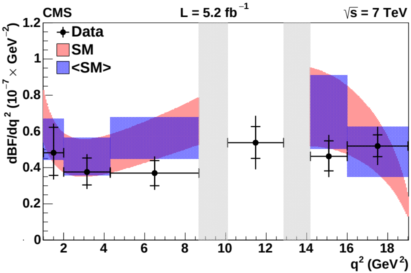
- A total of more than 400 signal events
- Clear yields from all the q^2 bins



- Signal
- Combinatorial background: randomly associated hadron-muon pairs
- Peaking background: feed-through from dimuon resonant channels

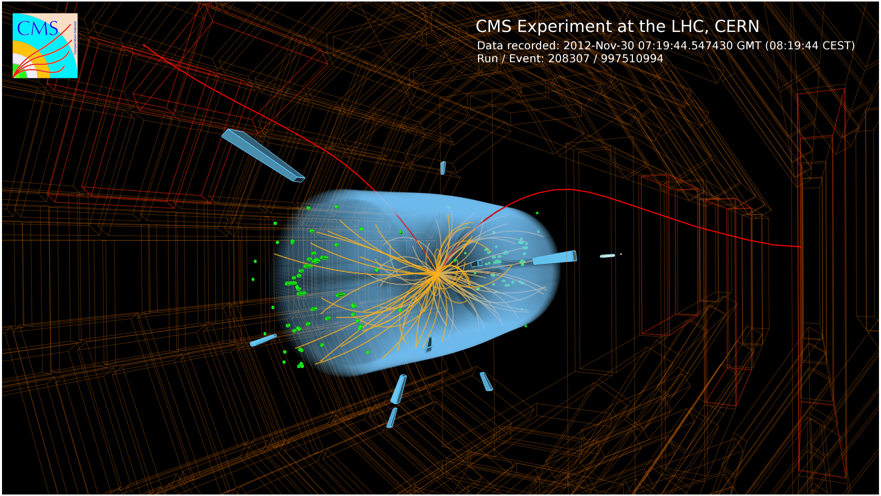






- Short error bars: statistical uncertainty
- Long error bars: total uncertainty
- Red band: standard model predictions
- Blue boxes: SM predictions averaged over the q^2 bin

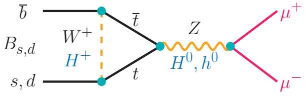
- No deviations from the standard model are found
- Results still dominated by the statistical uncertainty

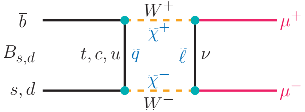


Phys. Rev. Lett. 111,101804 (2013)
arXiv:1307.5025

THE $B_s^0\,(B^0) \to \mu^+\mu^-$ DECAY

- $B_s^0\,(B^0) \to \mu^+\mu^-$ decays highly suppressed in standard model:
  - ▷ FCNC
    - Tree-level suppression
    - Box/penguin diagrams
  - ▷ Elicity suppression (m_μ^2 / m_B^2)
  - ▷ Internal quark annihilation (f_B^2 / m_B^2)





SM predictions (time-integrated)

$$\mathcal{B}\left(B_s^0 \to \mu^+\mu^-\right) = (3.57 \pm 0.30) \times 10^{-9}$$

$$\mathcal{B}\left(B^0 \to \mu^+\mu^-\right) = (1.07 \pm 0.10) \times 10^{-10}$$

[De Bruyn et al. PRL 109, 041801]

[A. Buras et al. arXiv:1303.3820]

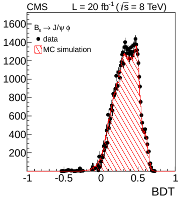
- $\mathcal{B}\left(B_s^0\,(B^0) \to \mu^+\mu^-\right)$ probe for physics BSM
  - ▷ $\mathcal{B} \sim (\tan\beta)^6$ in MSSM
  - ▷ $\mathcal{B} \sim (\tan\beta)^4$, m_{H^+} in 2HDM
- (Very) Small theoretical uncertainties
- High sensitivity to new physics
  - ▷ Comparable to $\mu \to e\gamma$, $B \to X\nu\nu$

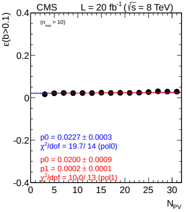
- Complete re-analysis of entire 2011 (5 fb^{-1}, $\sqrt{s} = 7\,\text{TeV}$) + 2012 (20 fb^{-1}, $\sqrt{s} = 8\,\text{TeV}$) datasets
  - Blind analysis (2011 dataset re-blinded)
  - Better muonID and fake suppression (BDT based)
  - New (and improved) analysis variables
  - BDT based discriminant
  - Simultaneous maximum likelihood fit to dimuon mass in several BDT bins and two rapidity ranges

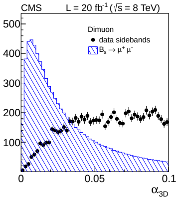
$$\mathcal{B}\left(B_s^0 \to \mu^+\mu^-\right) = \frac{N_s}{N_{obs}^{B^+}} \frac{f_u}{f_s} \frac{\varepsilon_{tot}^{B^+}}{\varepsilon_{tot}} \mathcal{B}\left(B^+ \to J/\psi\, \mathrm{K}^+\right)$$

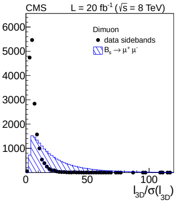
- $\mathcal{B}\left(B_s^0\left(B^0\right) \to \mu^+\mu^-\right)$ is measured relative to the $B^+ \to J/\psi\,\mathrm{K}^+$ normalization channel
  - $\mathcal{B}$ independent from the luminosity measurement
  - Avoid uncertainties related to b quark production cross section
  - Mitigate the effects of uncertainties in the efficiencies
- Control sample $B_s^0 \to J/\psi\,\phi\,(KK)$ to compare and validate B_s^0 mesons in data and MC simulations

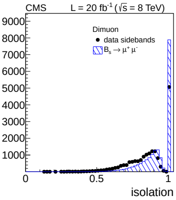
- The events are selected by a Boosted Decision Tree (BDT)
  - ▷ Signal: $B_s^0 \to \mu^+\mu^-$ MC simulation
  - ▷ Background: dimuon data mass sidebands
  - ▷ Data randomly split in 3 subset:
    - Train on 1, test on 2, apply on 3
    - Avoid selection bias
- Studies on BDT output:
  - ▷ insensitive to invariant mass
  - ▷ insensitive to pileup
- Use the same BDT for normalization and control samples
  - ▷ differences between Data and MC BDT output as systematic uncertainties
- A (very short) selection of the most significant variables:

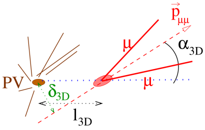








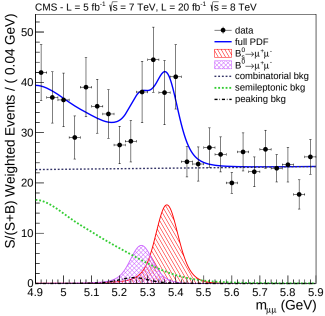


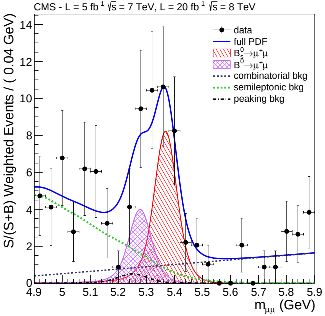


- 2 approaches:
  - Categorized-BDT to extract $\mathcal{B}\left(B_s^0 \to \mu^+\mu^-\right)$
    - UML fit to 12 mass distributions in BDT bins split in Barrel/Endcap
    - BDT binning chosen to equalize the expected number of signal events
  - 1D-BDT to extract CL_s limits on $\mathcal{B}\left(B^0 \to \mu^+\mu^-\right)$
    - Optimized cut on BDT output and event counting in mass windows

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

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





- Observation of $B_s^0 \to \mu^+\mu^-$ decay

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

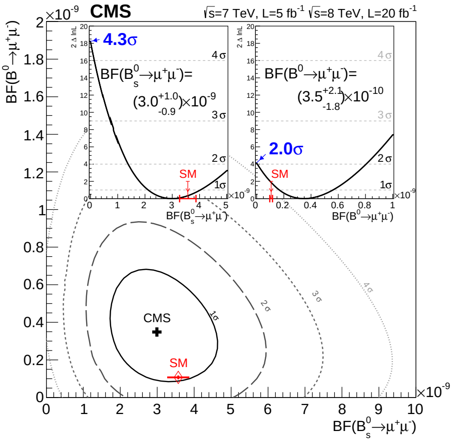
4.3 σ significance (expected 4.8 σ median)

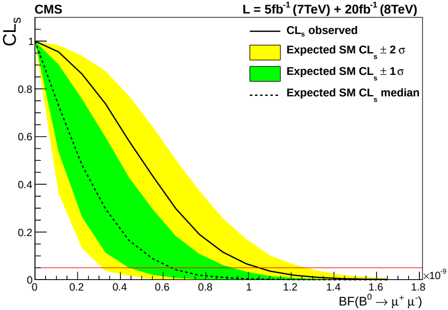
- No significant excess of $B^0 \to \mu^+\mu^-$

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

2.0 σ significance

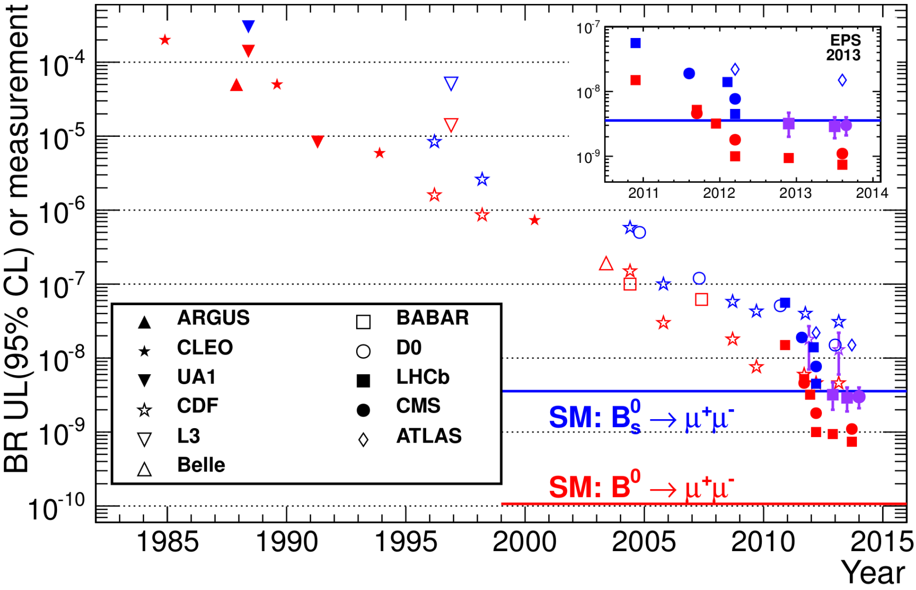
$$\mathcal{B}\left(B^0 \to \mu^+\mu^-\right) < 1.1 \times 10^{-9},\ 95\%\ \text{CL}$$





Conclusions and Outlook

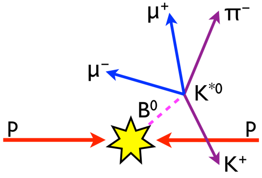
- First CMS angular analysis of $B^0 \to K^{*0}\mu^+\mu^-$, using 2011 dataset (5 fb^{-1})
  - More than 400 signal decays
  - No deviations from the SM predictions are found
  - Results with 2011 data are statistically limited
    - Analysis of the 2012 data is starting
- $\mathcal{B}\left(B_s^0\left(B^0\right) \to \mu^+\mu^-\right)$ analysis, using 2011+2012 datasets (25 fb^{-1})
  - Substantial improvements to previous analysis
    - Muon identification with BDT
    - Analysis selection with BDT
    - UML fit to mass distributions
  - 4.3 σ significance of $B_s^0 \to \mu^+\mu^-$ observation
    - Consistent with SM predictions
  - No significant excess is observed for $B^0 \to \mu^+\mu^-$ decay
    - Upper limit using CLs method
  - Future plans
    - Looking forward for LHC Run2
- All public BPH results from the CMS collaboration
  - `https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsBPH`


BR UL(95% CL) or measurement
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{-8}
10^{-9}
10^{-10}
1985
1990
1995
2000
2005
2010
2015
Year
ARGUS
CLEO
UA1
CDF
L3
Belle
BABAR
D0
LHCb
CMS
ATLAS
SM: $B_s^0 \to \mu^+\mu^-$
SM: $B^0 \to \mu^+\mu^-$
EPS
2013
10^{-7}
10^{-8}
10^{-9}
2011
2012
2013
2014

BACKUP

Signal:

- $B^0 \to K^{*0}\mu^+\mu^-$
  - ▷ Two well reconstructed/identified muons with invariant mass far from the J/ψ and ψ' (q^2 bins)
  - ▷ $K^{*0}(892)$ reconstructed through the decay in $K\pi$ (both tracks must fail muonID)
  - ▷ 4-track vertex candidate identified as B^0 $(\overline{B}^0)$ if $K^+\pi^-$ $(K^-\pi^+)$ is the closest to the nominal K^{*0} mass



Background:

- Peaking
  - ▷ Feed-through from $B^0 \to K^{*0}J/\psi$ and $B^0 \to K^{*0}\psi'$ decays not removed by the q^2 cut
  - ▷ Dimuon mass reconstructed far from the true J/ψ (ψ') mass $\to$ $K\pi\mu\mu$ mass shifted from the true B^0 mass
  - ▷ Shape obtained from simulation.
- Combinatorial
  - ▷ Unavoidable fraction of mis-reconstructed B^0 from randomly associated muons to hadrons
  - ▷ Shape described by a single exponential in mass, and by polynomials for the angular variables.

Muon selection

- events recorded with dimuon low-mass, vertex displaced trigger
- $\mu^+\mu^-$ vertex transverse flight length significance cut (in trigger)
- $1 < m_{\mu\mu} < 4.8$ GeV (in trigger)
- $p_T(\mu) > 3, 4, 4.5, 5$ GeV (depending on trigger)

Hadron selection

- hadron must fail muon ID
- $p_T(h) > 0.75$ GeV
- cut on transverse distance of closest approach
- $|m(K\pi) - m_{PDG}(K^{*0})| < 80$ MeV

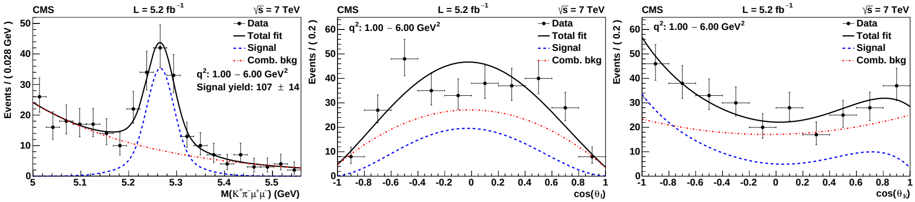
B^0 selection

- B^0 vertex CL $> 9\%$
- B^0 vertex transverse flight length significance cut
- $\cos(\alpha) > 0.9994$, α being the angle in transverse plane between B^0 momentum and line-of-flight

CP state assignment

- tag the CP state based on closest distance from K^{*0} PDG mass
- reject event if both K^{*0} and $\overline{K^{*0}}$ masses are within 50MeV of PDG mass $(\sim 1\Gamma)$
- mistag rate : 8% on MC; 10% on data

- "Special" bin $\to$ very precise theoretical predictions

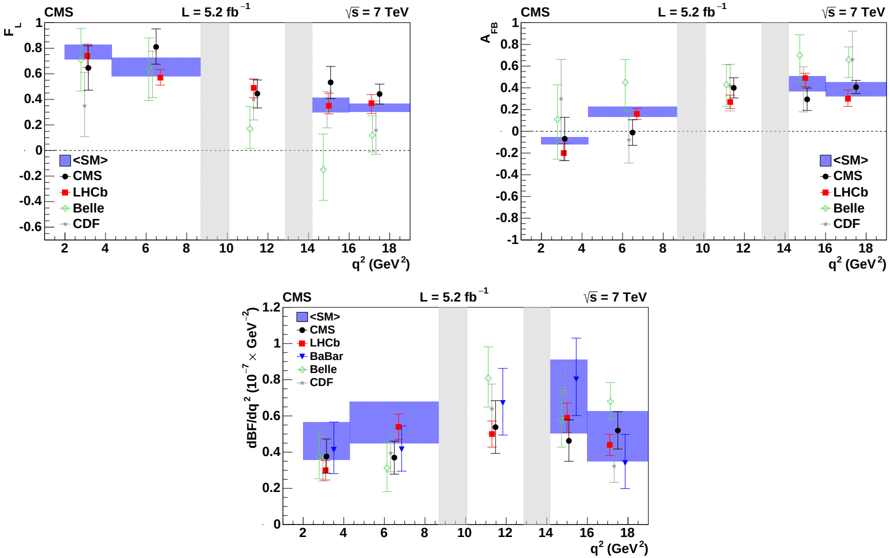


CMS results

- $F_L = 0.68 \pm 0.10$ (stat.) ± 0.02 (syst.)
- $A_{FB} = -0.07 \pm 0.12$ (stat.) ± 0.01 (syst.)
- $\mathrm{d}\mathcal{B}/\mathrm{dq}^2 = (4.4 \pm 0.6 \text{ (stat.)} \pm 0.7 \text{ (syst.)}) \times 10^{-8}$ GeV^{-2}

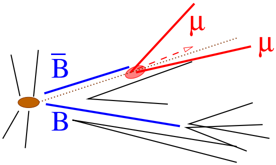
SM predictions (Bobeth et al. PRD 87 034016)

- $F_L = 0.74^{+0.06}_{-0.07}$
- $A_{FB} = -0.04 \pm 0.03$
- $\mathrm{d}\mathcal{B}/\mathrm{dq}^2 = \left(4.9^{+1.0}_{-1.1}\right) \times 10^{-8}$ GeV^{-2}


CMS
L = 5.2 fb^{-1}
$\sqrt{s}$ = 7 TeV
F_L
A_{FB}
dBF/dq^2 ($10^{-7} \times$ GeV^{-2})
q^2 (GeV^2)
<SM>
CMS
LHCb
BaBar
Belle
CDF

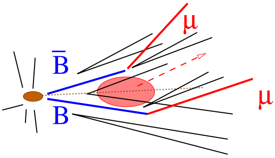
Signal

- Two muons from one decay (secondary) vertex
  - ▷ Well reconstructed secondary vertex
  - ▷ Dimuon momentum aligned with SV flight direction
  - ▷ $m_{\mu\mu}$ close to the nominal $m_{B_s^0\,(B^0)}$ value
  - ▷ PV/SV/muon isolation



Background

- Combinatorial (estimated from sidebands)
  - ▷ One semileptonic B decay and one hadron (K/π) misidentified as a muon
  - ▷ Two semileptonic B decays
- Rare single B decays (estimated from MC simulation)
  - ▷ Peaking: $B_s^0 \to K\pi$, $B_s^0 \to KK$, $B^0 \to \pi\pi$, ...
  - ▷ Non-peaking: $B_s^0 \to K\mu\nu$, $\Lambda_b \to p\mu\nu$, ...

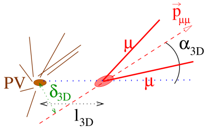


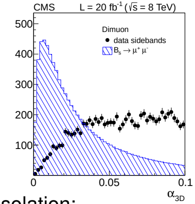
Crucial aspects of the analysis

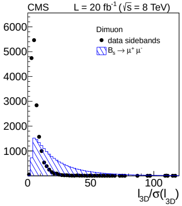
- muonID and $h \to \mu$ suppression
- vertexing variables
- isolation variables

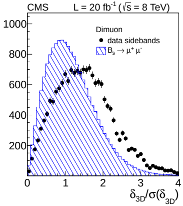
- Improved muon identification to cope with peaking and semileptonic backgrounds
  - ▷ Use a MVA based on Silicon Tracker and Muon Detector information
    - distribution of hits in the tracker as compared to expected
    - kink identification
    - muon segment compatibility
    - χ^2/ndof
    - muon p_T and η
  - ▷ Muon MC efficiency validated with Tag&Probe methods in data
  - ▷ Muon fake rate reduced by a factor of $\sim$2 compared to previous analysis, at the expenses of $\sim$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)
    - $D^{*0} \to D^0\pi$, $K_s \to \pi\pi$, $\Lambda \to p\pi$

- Vertexing:
  - ▷ Angle between PV ad SV
  - ▷ SV flight length significance
  - ▷ Dimuon impact parameter significance

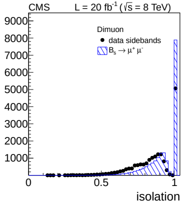


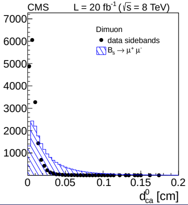


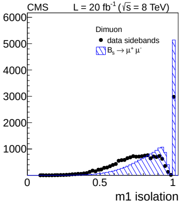




- Isolation:
  - ▷ Primary Vertex isolation (dimuon isolation)
  - ▷ Secondary Vertex isolation (B-cand. isolation)
  - ▷ Muon isolation

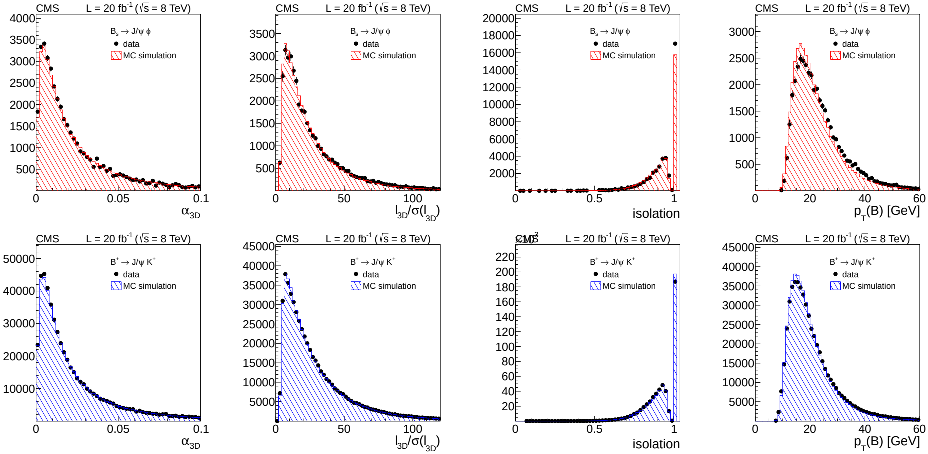




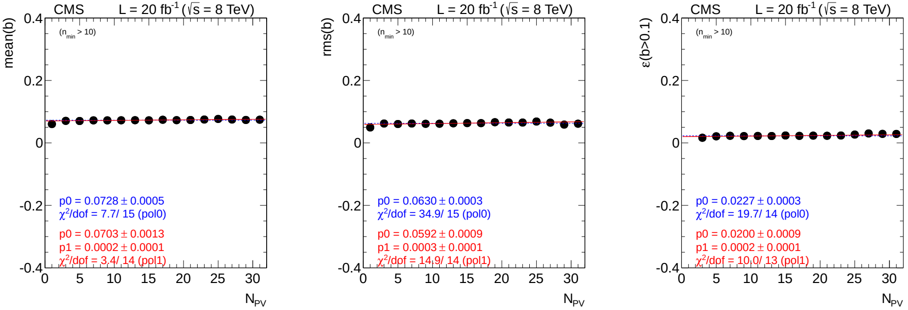


$B_s^0\,(B^0) \to \mu^+\mu^-$ – Data vs MC comparison

- Comparison of sidebands subtracted distributions
  - ▷ Good agreement

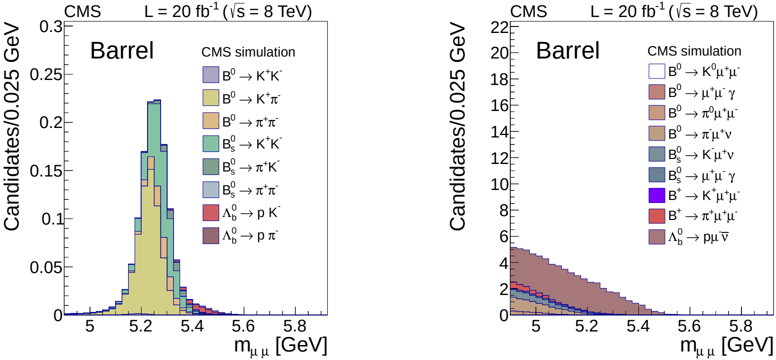


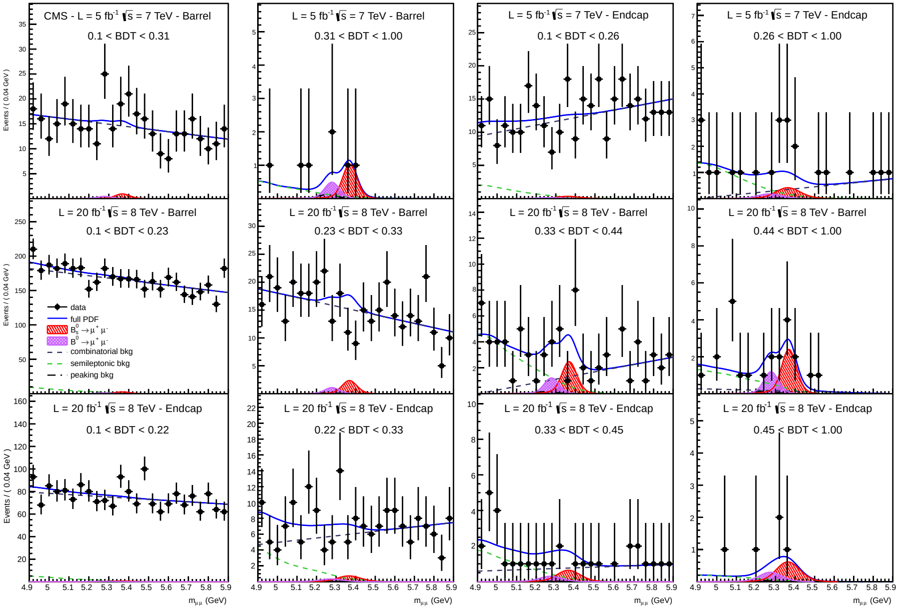
- Average number of interactions per bunch crossing
  - 2011: ≈ 9
  - 2012: ≈ 21
- Pileup independence checked
  - Signal MC event samples with pileup
    - every single variable used in BDT is shown to be pileup independent
  - Data studies with BDT output distribution vs. N_{PV}
    - mean and RMS, efficiency of BDT requirement
- No significant pileup dependence observed



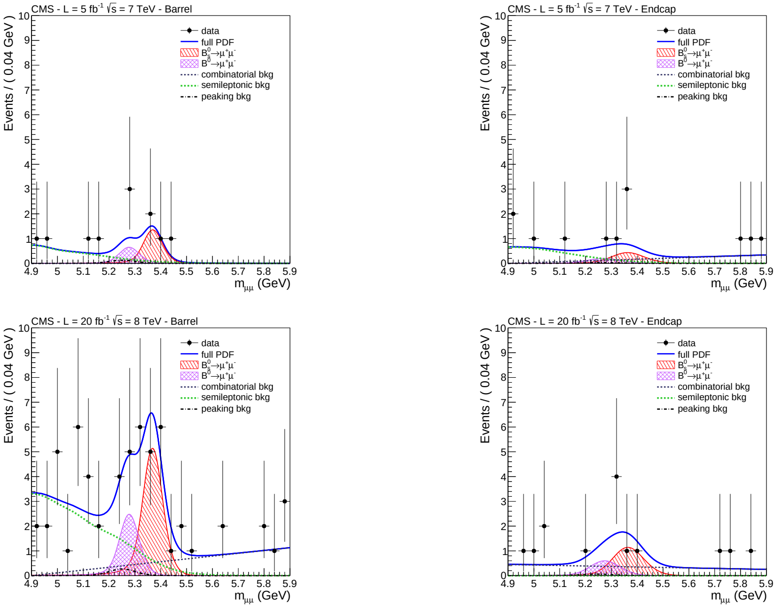
- Normalized to $B^+ \to J/\psi\,\mathrm{K}^+$ yield in data

$$N(X) = \frac{\mathcal{B}\,(Y \to X)}{\mathcal{B}\,\left(B^+ \to J/\psi\,\mathrm{K}^+\right)}\,\frac{f_Y}{f_u}\,\frac{\varepsilon^X_{tot}}{\varepsilon^{B^+}_{tot}}\,N_{obs}(B^+)$$

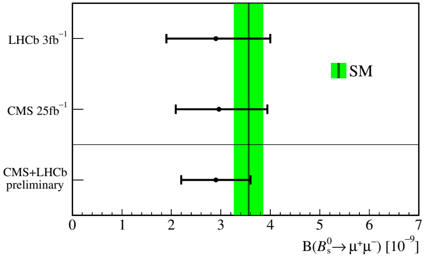


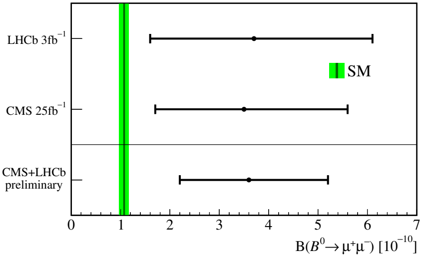

CMS - L = 5 fb^{-1} $\sqrt{s}$ = 7 TeV - Barrel
0.1 < BDT < 0.31
L = 5 fb^{-1} $\sqrt{s}$ = 7 TeV - Barrel
0.31 < BDT < 1.00
L = 5 fb^{-1} $\sqrt{s}$ = 7 TeV - Endcap
0.1 < BDT < 0.26
L = 5 fb^{-1} $\sqrt{s}$ = 7 TeV - Endcap
0.26 < BDT < 1.00
L = 20 fb^{-1} $\sqrt{s}$ = 8 TeV - Barrel
0.1 < BDT < 0.23
L = 20 fb^{-1} $\sqrt{s}$ = 8 TeV - Barrel
0.23 < BDT < 0.33
L = 20 fb^{-1} $\sqrt{s}$ = 8 TeV - Barrel
0.33 < BDT < 0.44
L = 20 fb^{-1} $\sqrt{s}$ = 8 TeV - Barrel
0.44 < BDT < 1.00
L = 20 fb^{-1} $\sqrt{s}$ = 8 TeV - Endcap
0.1 < BDT < 0.22
L = 20 fb^{-1} $\sqrt{s}$ = 8 TeV - Endcap
0.22 < BDT < 0.33
L = 20 fb^{-1} $\sqrt{s}$ = 8 TeV - Endcap
0.33 < BDT < 0.45
L = 20 fb^{-1} $\sqrt{s}$ = 8 TeV - Endcap
0.45 < BDT < 1.00
data
full PDF
$B_s^0 \to \mu^+\mu^-$
$B^0 \to \mu^+\mu^-$
combinatorial bkg
semileptonic bkg
peaking bkg
Events / (0.04 GeV)
$m_{\mu\mu}$ (GeV)

$B_s^0\ (B^0) \to \mu^+\mu^-$ – 1D-BDT FITS



- CMS-PAS-BPH-13-007
- CERN-LHCb-CONF-2013-012





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

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