Observation of the rare decay $B_s^0 \to \mu^+\mu^-$ and search for $B^0 \to \mu^+\mu^-$ with the CMS experiment

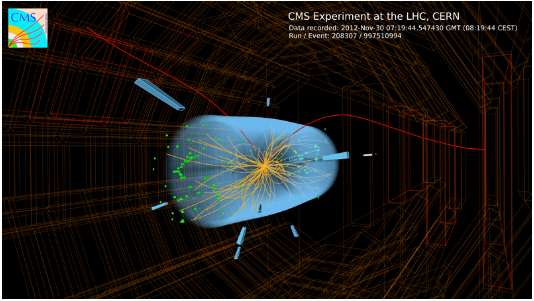
Università degli studi di Siena

Ph.D. thesis in Experimental Physics defended on December the 4th 2013

Luca Martini

Supervisor Dott. **Fabrizio Palla**

Tutor Dott.sa **Maria Agnese Ciocci**

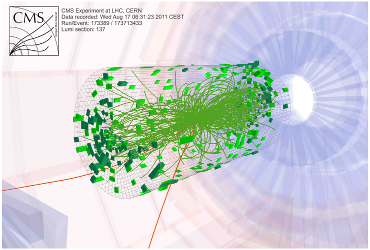


Summary

1. The $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ decays are among the most sensitive channels for the indirect search of new physics
2. CMS tracker and muon detectors allow very clean identification and high reconstruction efficiency for muons
3. LHC during 2011 and 2012 produced large samples of B_s^0 and B^0 mesons

- **In this thesis I presented**
    1. **the first measurement of $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$**
    2. **the most recent upper limit on $\mathcal{B}(B^0 \to \mu^+\mu^-)$**

$\rightarrow$ **Performed with the CMS experiment using the full 2011 and 2012 data**

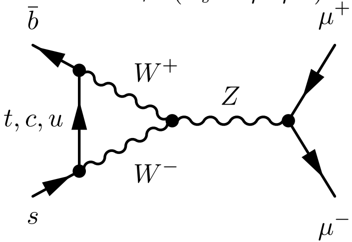


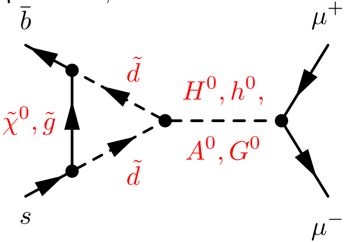
A highly suppressed decays

- $B^0_{(s)} \to \mu^+\mu^-$ decays are forbidden at tree level in the SM
  - FCNC, proceed through penguin and box diagrams, helicity suppressed by $(m_\mu/m_B)^2$, CKM suppressed
- Among the simplest and cleanest decays to evaluate [arXiv:1311.0903v3]

$$\mathcal{B}(B^0_s \to \mu^+\mu^-)_{SM} = (3.66 \pm 0.23) \times 10^{-9}$$
$$\mathcal{B}(B^0 \to \mu^+\mu^-)_{SM} = (1.06 \pm 0.09) \times 10^{-10}$$

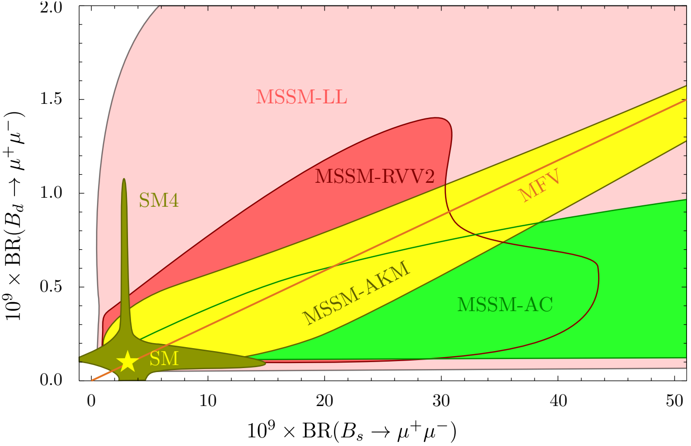
- Small values and small uncertainties (dominated by CKM and f_B), very sensitive to NP contributions
  - In SUSY, $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ is sensitive up to $\tan^6\beta$





Correlations improve sensitivity on new physics

- Different SUSY models allow different BF values

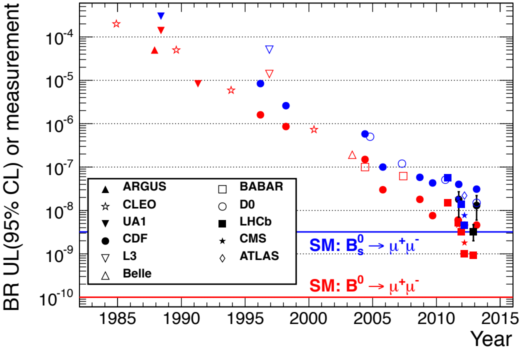


- It is crucial to measure **both** $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ and $\mathcal{B}(B^0 \to \mu^+\mu^-)$

Experimental status

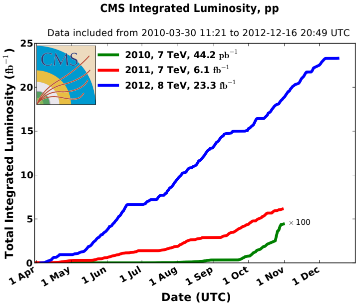
- Many HEP collaborations have published results on the search for $B_s^0 \to \mu^+\mu^-$ or $B^0 \to \mu^+\mu^-$ decays
- Over the past 30 years, significant progress has been made

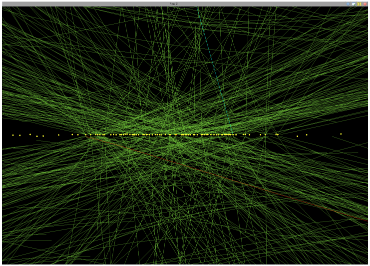
$\rightarrow$ Limiting factor is acquired statistics



LHC and CMS operating conditions

- Proton-Proton collisions, $\sqrt{s} = 7$ TeV in 2011 and 8 TeV in 2012
- LHC delivered an even-increasing instantaneous luminosity

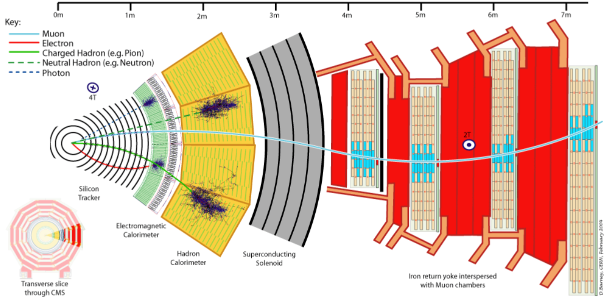




- At the cost of a challenging pile-up: $<\mu>_{2011} = 8$ PV, $<\mu>_{2012} = 21$ PV
- pile-up (PU) is the number of primary vertices per bunch crossing

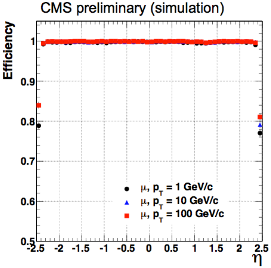
CMS main features

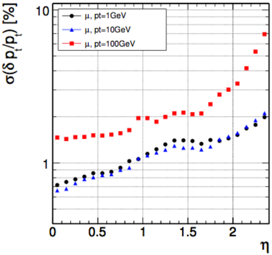
- Cylinder detector with a barrel and two endcap regions covering the region $|\eta| < 5$
- 3.8 T superconducting solenoid
- Inner full silicon tracker to reconstruct charged particle trajectories and vertices $|\eta| < 2.5$
- Muon detectors to identify and reconstruct muons with $p_T \approx$ 4 GeV-1 TeV range
- Electromagnetic and Hadronic calorimeters for energy measurements

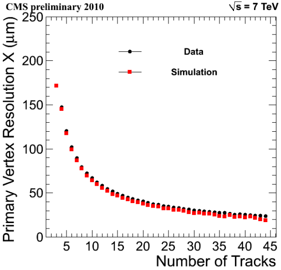


Inner Silicon Tracker performance

- High efficiency and resolution for track and vertex reconstructions

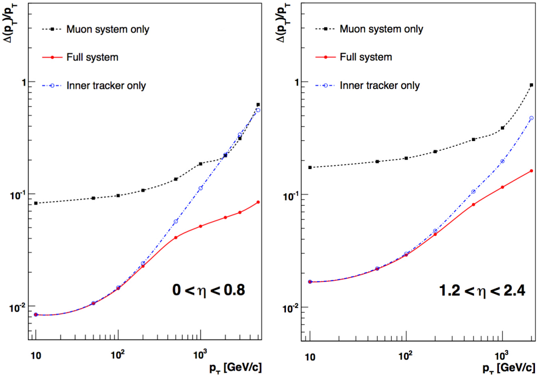






Muon detector performance

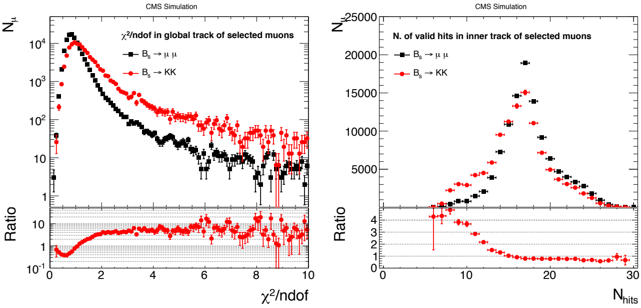
- Two reconstruction algorithms
  - Inside-out from tracker tracks extrapolating to muon system ("Tracker Muon")
  - Outside-in from muon tracks extrapolating to the tracker system ("Global Muon")



- For $B^0_{(s)}$ low p_T muons, the tracker drives the resolution

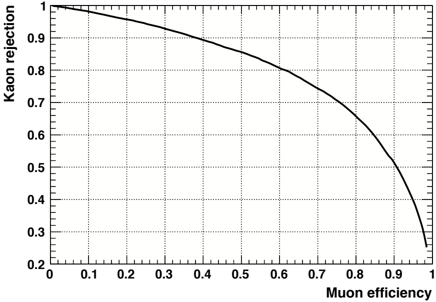
Single muon selection: muon BDT

- Rare backgrounds depend heavily on the "muon misidentification"
- **Kaon, pion and proton tracks wrongly identified as muons due to punch-through and decays-in-flight**
- Starting from well reconstructed CMS "tight" muons, improve the separation with a multivariate analysis, a Boosted Decision Tree (BDT)
- Tracker and muon track parameters used to separate true muons (from $B_s^0 \to \mu^+\mu^-$) against fake muons (from $B_s^0 \to K^+K^-$)
- Examples of input parameters: **Global track fit** χ^2, **Tracker valid hits**



Muon BDT results

- **Improvement of 50% fake rejection against a 90% muon efficiency**



Hadron	Misid. ($\times 10^{-3}$)
π	0.5–1.3
K	0.8–2.2
p	0.4–1.5

- Misidentification around 1‰
- With an uncertainty of 50%

- Muon efficiencies measured on data with $J/\psi \to \mu^+\mu^-$ with data-driven methods (*Tag & Probe*)
- Muon misidentification validated on data with:
  - $D^{*+} \to D^0\pi^+ \to K^-\pi^+\pi^+$ for kaon and pions
  - $B_s^0 \to J/\psi\phi$, $B^\pm \to J/\psi K^\pm$ for kaons
  - $\Lambda^0 \to p\pi^-$ for protons

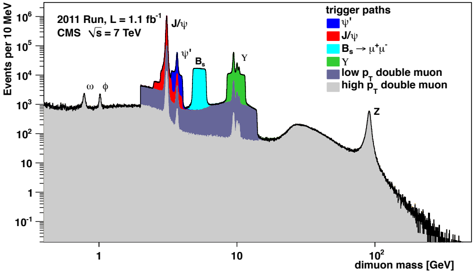
Trigger

- CMS uses a two-level trigger to save only the most interesting events from the LHC bunch crossing rate (40 MHz)

1. L1 hardware algorithms (Maximum rate < 100 kHz)
2. HLT software algorithms (Maximum rate $\lesssim 500$ Hz)

- Designed, validated and maintained optimal trigger paths
  1. B_s^0 L1 paths: 2 muons with $p_T > 3$ GeV and $|\eta| < 2.2$
  2. B_s^0 HLT paths: requests on muon and dimuon p_T and $|\eta|$, invariant mass, vertex probability and decay length

$\rightarrow$ High efficient triggers with rates of few Hz



Analysis overview

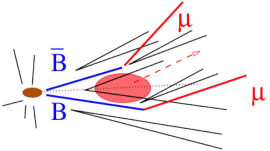
- Given the expected CMS sensitivity, results are given in two ways
  1. $\mathcal{B}(B^0 \to \mu^+\mu^-)$: counting experiment in the B^0 dimuon mass window
  2. $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$: unbinned maximum likelihood fit to the full dimuon mass
- Results extracted using a normalization to $B^\pm \to J/\psi K^\pm \to \mu^+\mu^- K^\pm$
  - to avoid uncertainties on the $b\bar{b}$ production cross-section and luminosities
  - to cancel at first order many systematic errors

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = \frac{N_{\rm obs}^{B_s^0}}{N_{\rm obs}^{B^\pm}} \times \frac{\epsilon_{B^\pm}}{\epsilon_{B_s^0}} \times \frac{f_{\rm u}}{f_{\rm s}} \times \mathcal{B}(B^\pm \to J/\psi K^\pm) \times \mathcal{B}(J/\psi \to \mu^+\mu^-)$$

- Yield ratio — Efficiency ratio — Production ratio
- A blind analysis: All selections are fixed without using events from the signal region to avoid biases
- B_s^0 signal distributions validated on data using $B_s^0 \to J/\psi\phi \to \mu^+\mu^- K^+ K^-$

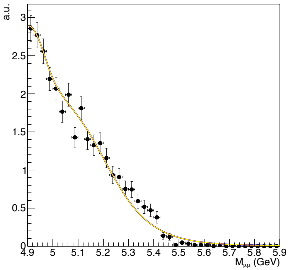
Background contributions

1. **Combinatorial**, formed by muons coming from b or c weak decays, $q \to X\mu\bar{\nu}$. **Studied on side-bands**
2. **Rare semileptonic** like $B_s^0 \to K^-\mu^+\nu_\mu$, where a hadron is misidentified as a muon. **Studied on data and MC**
3. **Rare peaking** like $B_s^0 \to K^+K^-$, where both hadrons are misidentified as muons. **Studied on data and MC**
4. **Cross-feeding** between $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ due to the detector resolution. **Studied on MC**



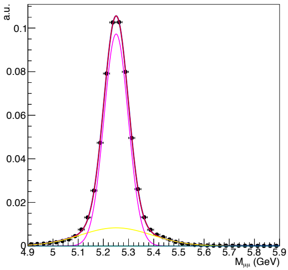
Background contributions

1. **Combinatorial**, formed by muons coming from b or c weak decays, $q \to X\mu\bar{\nu}$. **Studied on side-bands**
2. **Rare semileptonic** like $B_s^0 \to K^-\mu^+\nu_\mu$, where a hadron is misidentified as a muon. **Studied on data and MC**
3. **Rare peaking** like $B_s^0 \to K^+K^-$, where both hadrons are misidentified as muons. **Studied on data and MC**
4. **Cross-feeding** between $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ due to the detector resolution. **Studied on MC**

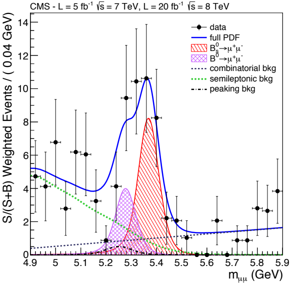


Background contributions

1. **Combinatorial**, formed by muons coming from b or c weak decays, $q \to X\mu\bar{\nu}$. **Studied on side-bands**
2. **Rare semileptonic** like $B_s^0 \to K^-\mu^+\nu_\mu$, where a hadron is misidentified as a muon. **Studied on data and MC**
3. **Rare peaking** like $B_s^0 \to K^+K^-$, where both hadrons are misidentified as muons. **Studied on data and MC**
4. **Cross-feeding** between $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ due to the detector resolution. **Studied on MC**

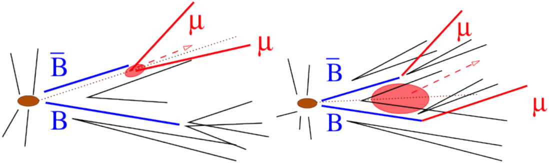


1. **Combinatorial**, formed by muons coming from b or c weak decays, $q \to X\mu\bar{\nu}$. **Studied on side-bands**
2. **Rare semileptonic** like $B_s^0 \to K^-\mu^+\nu_\mu$, where a hadron is misidentified as a muon. **Studied on data and MC**
3. **Rare peaking** like $B_s^0 \to K^+K^-$, where both hadrons are misidentified as muons. **Studied on data and MC**
4. **Cross-feeding** between $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ due to the detector resolution. **Studied on MC**

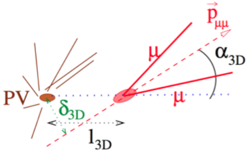


$B^0_{(s)}$ candidate selections

- Combinatorial background is made by combining two uncorrelated muons
  - their secondary vertex is badly reconstructed
  - the candidate is not compatible with coming from an existing PV
  - there are other particles around them (not isolated)

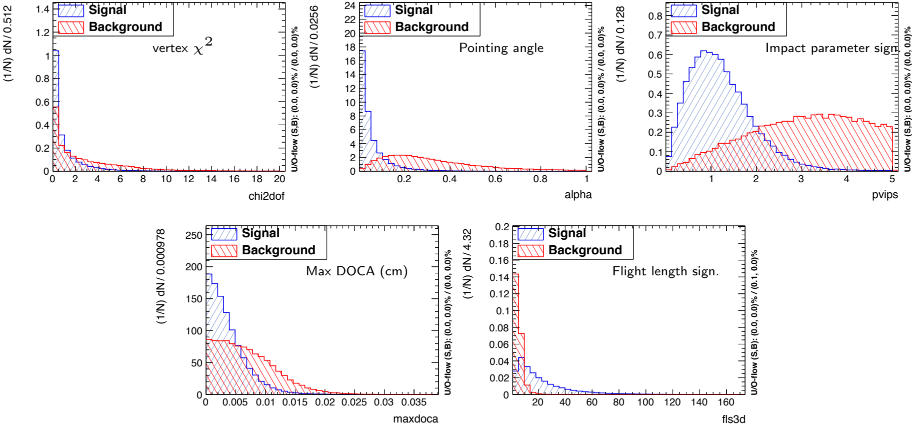


- Variables related to the Secondary Vertex
  - vertex χ^2
  - pointing angle $\alpha_{\rm 3D}$
  - 3D impact parameter δ_{3D}
  - distance of closest approach between the 2 muons $d_{\rm ca}^{\rm max}$
  - decay length $\ell_{\rm 3D}$



$B^0_{(s)}$ candidate selection: secondary vertex

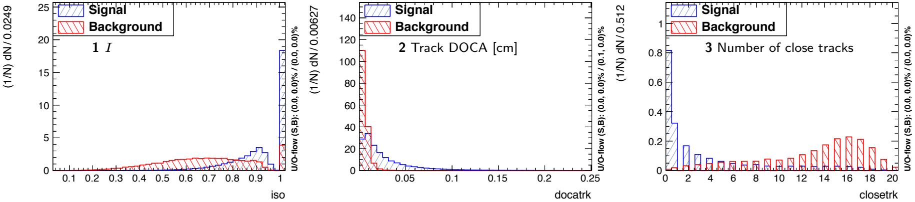
- After a loose preselection the variables are given to a BDT training:
  - **Signal =** $B^0_s \to \mu^+\mu^-$ **MC**; **Background = data side-bands**
- Large separation between signal and combinatorial background:



- Separate the signal, which is an isolated candidate, from background events containing tracks from jets
  1. Sum over all tracks in a cone around the $B^0_{(s)}$ or around the two muons:

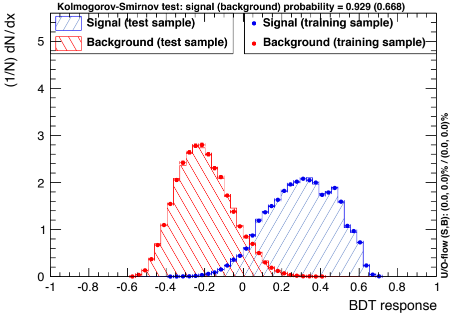
$$I = \frac{p_T(B)}{p_T(B) + \sum_{\text{trk}} p_T}$$

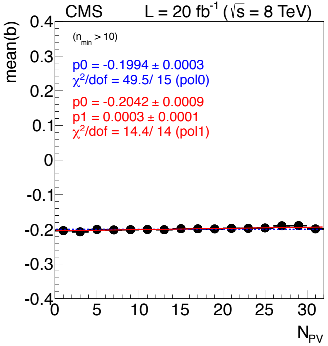
  2. Minimum d^0_{ca} of tracks
  3. Number of close tracks



$B^0_{(s)}$ BDT results

- This is a discovery analysis, figure of merit is the significance: $\frac{S}{\sqrt{S+B}}$

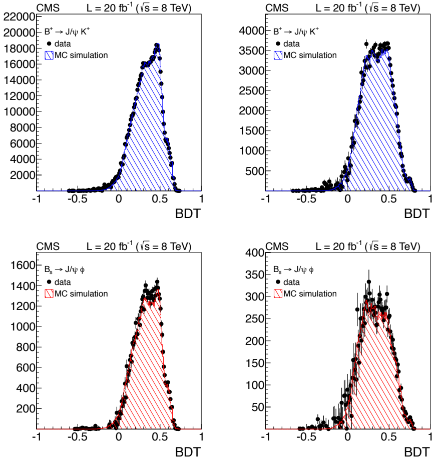




- **No systematic effect nor dependence observed on pile-up and mass**
- Systematic uncertainty taken from the difference of efficiency between data and MC in the normalization and control samples

$B^0_{(s)}$ BDT systematics

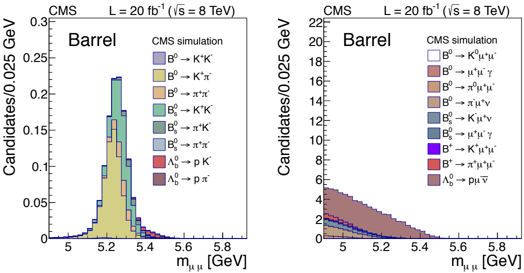
- Systematic uncertainty taken from the difference of efficiency between data and MC in the normalization and control samples



- Rare background normalized to the normalization channel $B^{\pm} \to J/\psi K^{\pm}$:

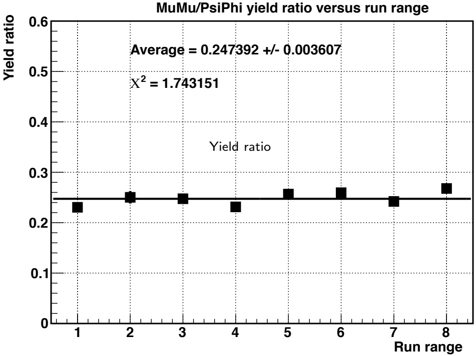
$$N(X) = \frac{\mathcal{B}(X)}{\mathcal{B}(B^{\pm} \to J/\psi K^{\pm}) \times \mathcal{B}(J/\psi \to \mu^{+}\mu^{-})} \frac{f_X}{f_u} \frac{\epsilon_X^{\text{ana}}}{\epsilon_{B^{\pm}}^{\text{ana}}} N_{\text{obs}}^{B^{\pm}}$$

- Include peaking and semileptonic decays from B_s^0, B^0, Λ_b^0
  - Uncertainties enter the likelihood as nuisance parameters



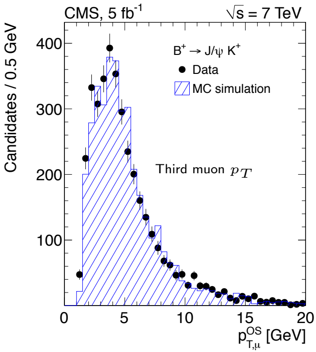
Other sytematic studies and cross-checks

1. Yield ratios are stable versus time
2. Production mechanisms, which could change acceptance and isolation, are in the right mixture in MC
3. Mass scale and resolution studied on data with J/ψ, $\psi(2S)$, $\Upsilon(nS)$



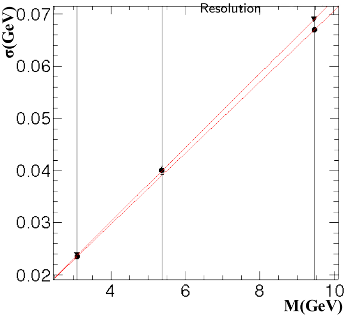
Other sytematic studies and cross-checks

1. Yield ratios are stable versus time
2. Production mechanisms, which could change acceptance and isolation, are in the right mixture in MC
3. Mass scale and resolution studied on data with J/ψ, $\psi(2S)$, $\Upsilon(nS)$



Other sytematic studies and cross-checks

1. Yield ratios are stable versus time
2. Production mechanisms, which could change acceptance and isolation, are in the right mixture in MC
3. Mass scale and resolution studied on data with J/ψ, $\psi(2S)$, $\Upsilon(nS)$



Production ratio f_s/f_u and summary

- f_s/f_u is taken from LHCb*, in a different ηp_T phase space
  - Measure $B^{\pm} \to J/\psi K^{\pm}/B_s^0 \to J/\psi\phi$ in p_T and η bins
  - Results are consistent with flat
  - $\to$ A further conservative systematic uncertainty of 5% is assigned
- Systematic summary table:

Category	Barrel (%)	Endcap (%)
f_s/f_u: production ratio of u and s quarks	8.0	8.0
acceptance: production processes	3.5	5.0
mass scale and resolution	5.0	5.0
efficiency (signal): data/MC simulation	9.5 - 3.3	7.9 - 2.3
efficiency (normalization): data/MC simulation	0.5 - 2.3	0.5 - 1.1
efficiency (normalization): kaon track efficiency	4.0	4.0
efficiency trigger	3.0	3.0
efficiency muon identification	2.0	2.0
normalization: fit on the pdf	5.0	5.0

*JHEP 1304 (2013)

Production ratio f_s/f_u and summary

- f_s/f_u is taken from LHCb*, in a different ηp_T phase space
  - Measure $B^{\pm} \to J/\psi K^{\pm} / B_s^0 \to J/\psi\phi$ in p_T and η bins
  - Results are consistent with flat
  - → A further conservative systematic uncertainty of 5% is assigned
- Systematic summary table:

Category	Barrel (%)	Endcap (%)
f_s/f_u: production ratio of u and s quarks	8.0	8.0
acceptance: production processes	3.5	5.0
mass scale and resolution	5.0	5.0
efficiency (signal): data/MC simulation	9.5 - 3.3	7.9 - 2.3
efficiency (normalization): data/MC simulation	0.5 - 2.3	0.5 - 1.1
efficiency (normalization): kaon track efficiency	4.0	4.0
efficiency trigger	3.0	3.0
efficiency muon identification	2.0	2.0
normalization: fit on the pdf	5.0	5.0

*JHEP 1304 (2013)

Unbinned maximum likelihood fit for $\mathcal{B}(B^0_s \to \mu^+\mu$

- Simultaneously fit in each data category to extract both $\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-)$
- Extended likelihood is formed by 5 contributions:

$$L = N_{B^0_s} F_{B^0_s} + N_{B^0} F_{B^0} + N_{\mathrm{comb}} F_{\mathrm{comb}} + N_{\mathrm{peak}} F_{\mathrm{peak}} + N_{\mathrm{semi}} F_{\mathrm{semi}}$$

where

  - N_i is the yield
  - F_i is the corresponding Probability Distribution Function (pdf)

- The branching fraction is

$$\mathcal{B}(B^0_s \to \mu^+\mu^-) = N_{B^0_s} \times K_{B^0_s}$$

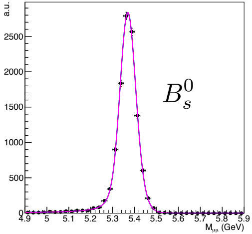
where

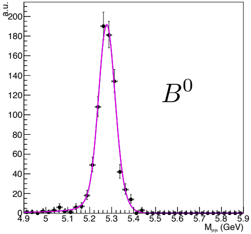
$$K_{B^0_s} = (N_{B^\pm}) \left(\frac{f_{\mathrm{s}}}{f_{\mathrm{u}}}\right) \left(\frac{\epsilon_{B^0_s}}{\epsilon_{B^\pm}}\right) \left(\frac{1}{\mathcal{B}(B^\pm \to J/\psi K^\pm) \times \mathcal{B}(J/\psi \to \mu^+\mu^-)}\right)$$

- Systematic uncertainties added as Gaussian nuisance parameters

$B^0_{(s)}$ signal modeling

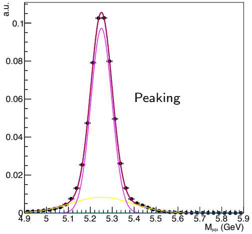
- Crystal Ball pdf
- Gaussian core portion and a power-law low-end tail, below a certain threshold
- Takes into account muon energy losses
- Width taken directly from the event mass error
  - Barrel width ≈ 50 MeV
  - Endcap width ≈ 80 MeV

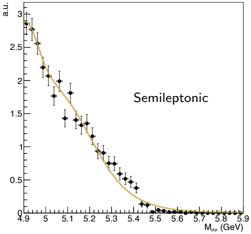




Backgrounds modeling

- Combinatorial pdf = floating first order polynomial
  - Shape studied on SB and on **non BDT muon** data sample
- Peaking pdf = Crystal Ball + Gaussian, which takes into account:
  1. Spread due to muon wrong mass assignment to kaons, pions, protons
  2. Spread due to different mass values from B_s^0, B^0, Λ_b^0
- Semileptonic pdf described with the Gaussian kernel method
  - pdf is the sum of n Gaussians where n is the number of events





“Categorized BDT” method

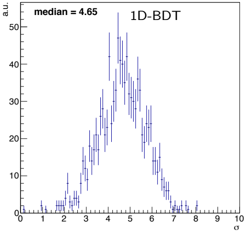
- Invariant mass distributions are sub-divided in bins of the BDT distribution
- Binning chosen to have the same event yield in each bin
- Total likelihood is the product of all independent categories L_i and of all constraints L_i^{constr}

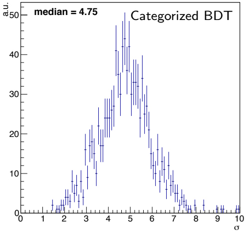
$$L_{\mathrm{tot}} = \prod_{i=0}^{11} L_i L_i^{\mathrm{constr}}$$

- Final likelihood studied with MC toy experiments
    - Show no significant bias on the BF evaluation

UML expected significance

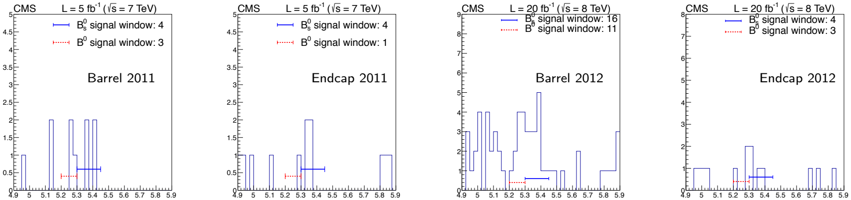
- **Null Hypothesis** $\mathcal{L}_0$: $\mathcal{B}(B_s^0 \to \mu^+\mu^-) = 0$.
- **Alternative Hypothesis** $\mathcal{L}_1$: $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ is let floating.
- The significance is evaluated with the profile likelihood ratio test





- Expected significance higher for the Categorized BDT → chosen method

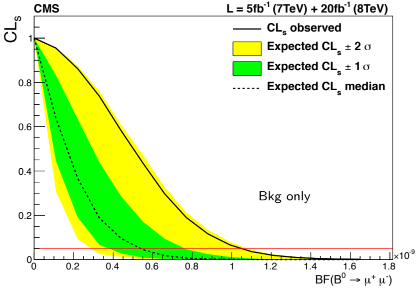
- After the final choice of all selections, the analysis is "unblinded"
- Results on the searches can be extracted

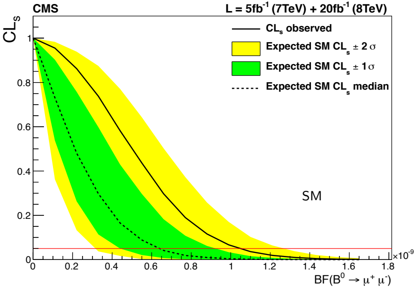


Category	$N_{B^0}^{\text{exp}}$	$N_{B^0}^{\text{obs}}$	$N_{B_s^0}^{\text{exp}}$	$N_{B_s^0}^{\text{obs}}$	$N_{B^\pm}^{\text{obs}}$
Barrel 2011	1.3 ± 0.8	3	3.6 ± 0.6	4	$(71.2 \pm 4.1) \times 10^3$
Endcap 2011	1.5 ± 0.6	1	2.6 ± 0.5	4	$(21.4 \pm 1.1) \times 10^3$
Barrel 2012	7.9 ± 3.0	11	17.9 ± 2.8	16	$(309 \pm 16) \times 10^3$
Endcap 2012	2.2 ± 0.8	3	5.1 ± 0.7	4	$(69.3 \pm 3.5) \times 10^3$

- Not significant excess in the $B^0 \to \mu^+\mu^-$ mass window
- But more than expected events, corresponding to $1.9\,\sigma$

$$\mathcal{B}(B^0 \to \mu^+\mu^-) \leq 1.1 \times 10^{-9} \qquad \text{at } 95\% \text{ CL}$$

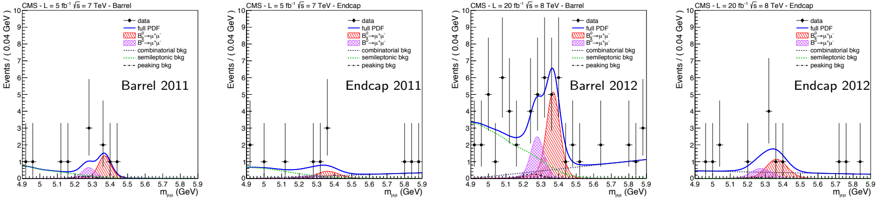




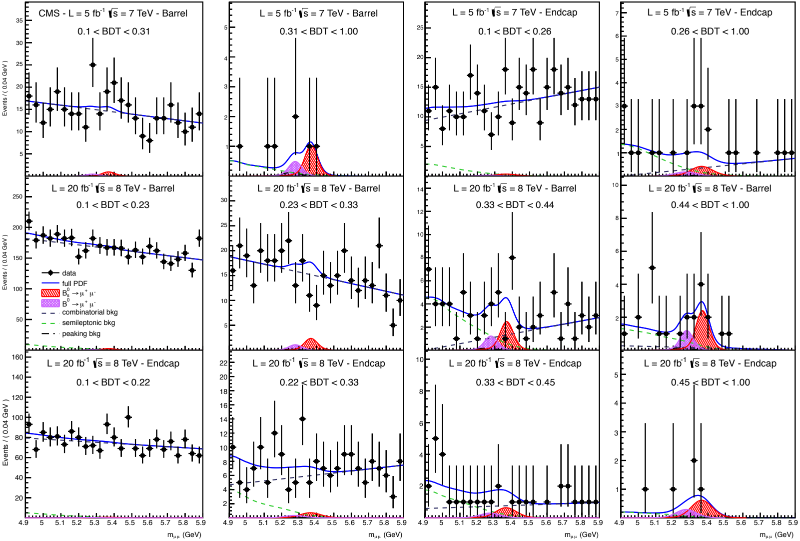
- Fits performed with both 1D-BDT and categorized BDT methods
- Categorized BDT fit results:

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (3.0^{+1.0}_{-0.9}) \times 10^{-9} \qquad [\mathrm{SM}: (3.46 \pm 0.18) \times 10^{-9}]$$
$$\mathcal{B}(B^0 \to \mu^+\mu^-) = (3.5^{+2.1}_{-1.8}) \times 10^{-10} \qquad [\mathrm{SM}: (1.07 \pm 0.10) \times 10^{-10}]$$

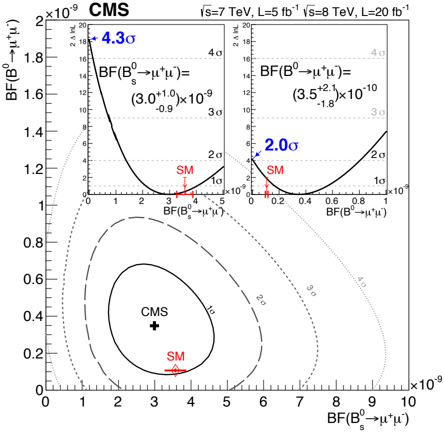
- 1D-BDT fits:



Categorized BDT fit results



Significance



$$\text{sign}(\mathcal{B}(B_s^0 \to \mu^+\mu^-)) = \mathbf{4.3}\,\sigma$$

$$\text{sign}(\mathcal{B}(B^0 \to \mu^+\mu^-)) = 2.0\,\sigma$$

Highest $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ significance up-to-date!

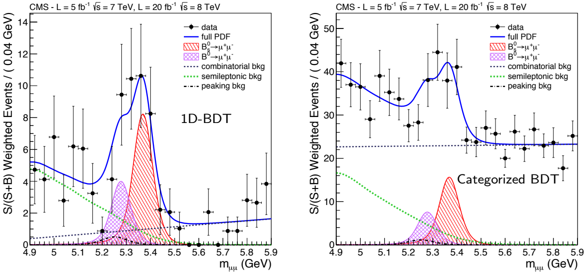
- The significance, given no $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ nor $\mathcal{B}(B^0 \to \mu^+\mu^-)$ decays:

$$\text{sign}(\mathcal{B}(B_s^0 \to \mu^+\mu^-) \cap \mathcal{B}(B^0 \to \mu^+\mu^-)) = 4.7\,\sigma$$

$\mathcal{B}(B^0_s \to \mu^+\mu^-)$ measured!

$$\mathcal{B}(B^0_s \to \mu^+\mu^-) = \left(3.0^{+0.9}_{-0.8}(\text{stat.})^{+0.6}_{-0.4}(\text{syst.})\right) \times 10^{-9}$$

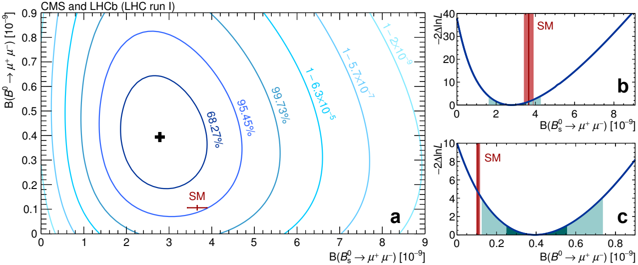
- Fully compatible with the SM expectation
- Statistically limited
- Measured significance ($4.3\,\sigma$) lower than median expected ($4.8\,\sigma$)
  - Measured BF slightly lower than SM BF
- $\mathcal{B}(B^0 \to \mu^+\mu^-)$ compatible with upper limit
  - $\mathcal{B}(B^0 \to \mu^+\mu^-) \leq 1.1 \times 10^{-9}$ at 95% CL
  - CMS upper limit still ten times SM expectation



CMS-LHCb combination of measurements

- Also the LHCb collaboration reports measurements of the $B_s^0 \to \mu^+\mu^-$ decay $(4\,\sigma)$, and the $B^0 \to \mu^+\mu^-$ decay $(2\,\sigma)$ [Phys.Rev.Lett. 111 (2013)]
- After a preliminary combination [CMS-PAS-BPH-13-007], a full combination of the likelihoods has been performed and results published on Nature [Nature 522, 68–72 (04 June 2015)]:

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = \left(2.8^{+0.7}_{-0.6}\right) \times 10^{-9} \qquad (6.2\,\sigma)$$
$$\mathcal{B}(B^0 \to \mu^+\mu^-) = \left(3.9^{+1.6}_{-1.4}\right) \times 10^{-10} \qquad (3.2\,\sigma)$$



$\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-)$ and the Run II

- LHC Run II is beginning its physics data taking this year
- We expect to get a hundred of fb^{-1} at $\sqrt{s} = 13\,\mathrm{TeV}$ with an inst. lumi. doubled w.r.t. Run I
- A challenging environment, with $<\mu>= 50\,\mathrm{PV}$ and high trigger rates
- CMS will nevertheless continue the studies of $\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-)$ with improved trigger and offline analyses

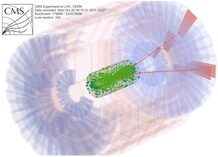
$\to$ $\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-)$ is a CMS flagship

- The main focus will move to the search for $B^0 \to \mu^+\mu^-$:

$\mathcal{L}(\mathrm{fb}^{-1})$	$\delta\mathcal{B}/\mathcal{B}(B^0_s)$	$\delta\mathcal{B}/\mathcal{B}(B^0)$	B^0 sign.	$\delta\frac{\mathcal{B}(B^0 \to \mu^+\mu^-)}{\mathcal{B}(B^0_s \to \mu^+\mu^-)}$
100 (2015–2017)	15%	66%	0.5–2.4 σ	71%

Conclusions

- In this thesis I presented the first measurement of $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ and the most recent upper limit of $\mathcal{B}(B^0 \to \mu^+\mu^-)$ using the 2011-2012 CMS data
- These results were published by the CMS collaboration on *Phys.Rev.Lett.* 111 (2013)
- $B_s^0 \to \mu^+\mu^-$ BF measurement, with a relative uncertainty of 30%, is the most precise up-to-date and it is in agreement and at the same level of precision of the LHCb result
- Both results are fully compatible with the SM predictions
- Allow to set stringent limits on new physics parameters
- High precision measurements will be the key for the indirect searches of new physics in the next future



BACK UP

B_s^0 meson proprieties

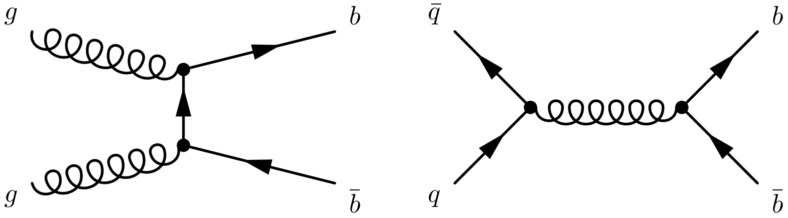
- $\bar{b}s\ 0^-$ bound-state
- B_s^0 is a flavor eigenstate, not a mass eigenstate
  - oscillates between B_s^0-$\bar{B}_s^0$ before decaying
  - important for comparing experimental and theoretical values
- mass $\approx 5.4\,\text{GeV}$
  - relatively low invariant mass to detect
- long decay length $c\tau \approx 450\,\mu\text{m}$
  - possible to measure the distance from the production (primary) vertex to the decay (secondary) vertex

SM Branching fraction

$$\mathcal{B}^0_{SM}\left(B^0_s \to \mu^+\mu^-\right) = \tau\left(B^0_s\right) \frac{G_F^2}{\pi}\left(\frac{\alpha}{4\pi \sin^2\theta_w}\right)^2 F^2_{B^0_s}\, m^2_\mu \times$$
$$\times\, m_{B^0_s}\sqrt{1-4\frac{m^2_\mu}{m^2_{B^0_s}}}\left|V^*_{tb}V_{ts}\right|^2 Y^2(x_t)$$

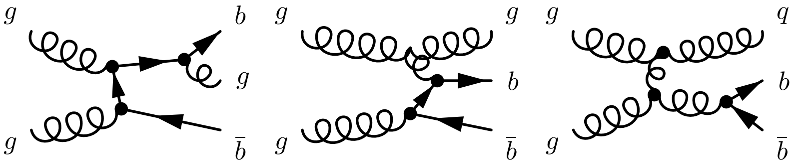
- life time
- Gauge and CKM constants and short-distance function
  - with QCD NLO corrections
- Hadronic matrix element
- chiral suppression

- Pair creation



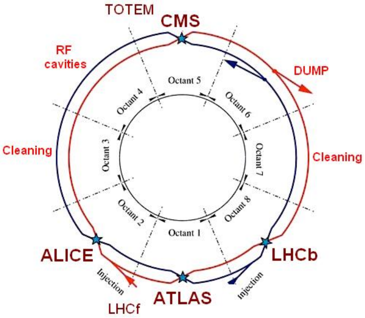
b production
NLO

- Pair creation, Flavor Excitation, Gluon Splitting



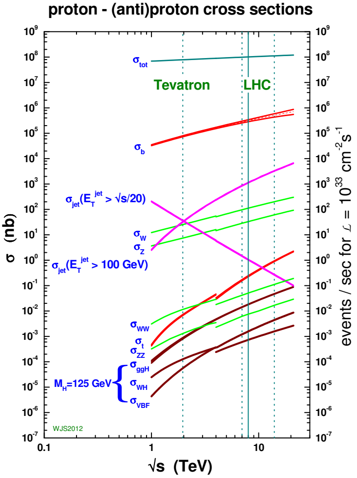
The LHC

- the LHC is a two-ring superconducting accelerator and collider installed at CERN
- 4 main detectors installed
  - CMC (general purpose)
  - Atlas (general purpose)
  - LHCb (B-Physics)
  - Alice (Heavy ions)



Luminosity and cross-section

- the rate of produced events is
$$n_{\text{event}} = \mathcal{L}\sigma_{\text{event}}$$
- σ_{event} is the production cross-section
- $\mathcal{L}$ depends only on collider parameters: number of particles per bunch, number of bunches, revolution frequency, bunch widths, ...
- at regime $\mathcal{L} = 1.0 \times 10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$



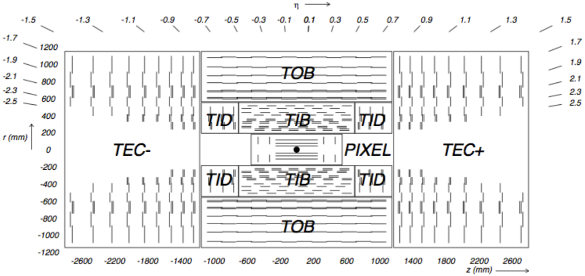
CMS coordinate system

CMS uses a right-handed coordinate system centered at the nominal IP:

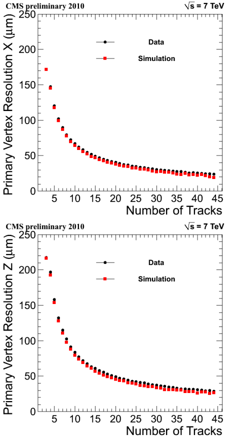
- the y-axis points vertically upward, orthogonal to the LHC plane;
- the x-axis points radially inward toward the center of the LHC;
- the z-axis points along the beam direction toward the Jura mountains.

The radial distance from the IP is defined as $r = \sqrt{x^2 + y^2}$. The azimuthal angle ϕ is measured from the x-axis in the xy plane. The polar angle θ starts from the z-axis, and the pseudorapidity variable is related to θ as $\eta = -\ln\tan(\theta/2)$.

- Full silicon technology
- 3 layers of pixel
  - 100 μm × 150 μm pixel cell size
- 10 layers of strips
  - Thickness 320–500 μm, pitch 80–180 μm
- High redundancy, low occupancy
- Allow to have at least 9 hits ensuring high efficiency and good resolution



- PV: vertex finding (in z) and vertex fitting (in 3D)
  - iterative Kalman filter
  - minimizes the squared distances of all tracks from vertex position
- SV: least mean squared minimization with Lagrange multipliers for the hypothesis under study
  - e.g. $B_s^0 \to J/\psi\phi \to \mu^+\mu^-K^+K^-$:
  - 4 final states
  - two muons coming from J/ψ
  - two kaons coming from ϕ
  - B_s^0 momentum pointing towards the PV



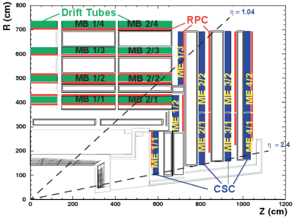
Track reconstruction

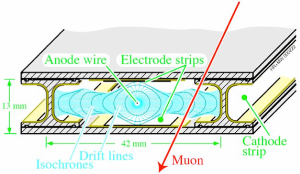
- Starting from the reconstructed hits:

1. seed generation
    - triplet of hits or two hits plus a vertex constraint
2. pattern recognition
    - combinatorial Kalman filter, iterative, takes successive detection layers, updating the track parameters for each added hit
3. ambiguity resolution
    - to avoid double counting, remove tracks with shared hits
4. final track fit
    - final construction with all hits.
    - cuts on χ^2, impact parameter, hits
    - tracks passing the tightest selection are labeled *highPurity*

Muon detectors

- Muon identification and reconstruction
- 3 gas detectors:
  1. Drift tubes (barrel)
     - 4 muon stations containing 12 chambers to measure $r\phi$ and z
  2. Cathode strip chambers (endcaps)
     - In a higher flux environment, high segmentation and radiation resistance, 4 layers to measure the $r\phi$ plane
  3. Resistive plate chambers ($|\eta| < 1.6$)
     - Lower spatial resolution but better timing, crucial for triggering on the right bunch crossing



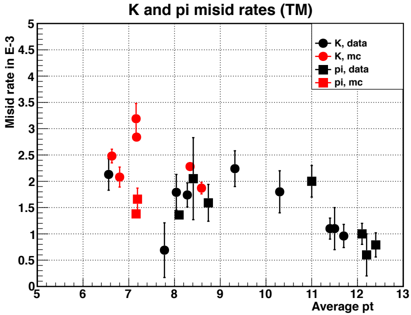


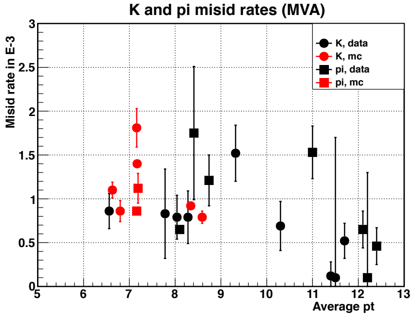
A **blind** analysis

- All selections are fixed without using events from the signal region to avoid biases
- B_s^0 signal distributions validated on data using the Control Sample (CS) $B_s^0 \to J/\psi\phi \to \mu^+\mu^- K^+ K^-$
- Background level and resolution depend strongly on muon η
  $\rightarrow$ data are split in "barrel" and "endcap" categories:
    - barrel = both muons with $|\eta| < 1.4$
    - endcap = otherwise

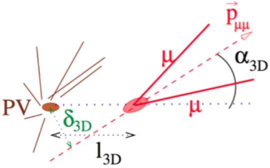
Definition	Invariant mass range (GeV)
overall mass window	$[4.9, 5.9]$
$B^0 \to \mu^+\mu^-$ signal window	$[5.2, 5.3]$
$B_s^0 \to \mu^+\mu^-$ signal window	$[5.3, 5.45]$
blind window	$[5.2, 5.45]$
side-band windows	$[4.9, 5.2] \cup [5.45, 5.9]$

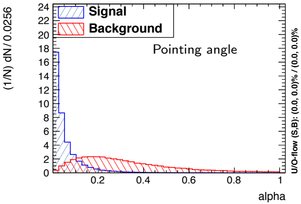
Muon BDT validation

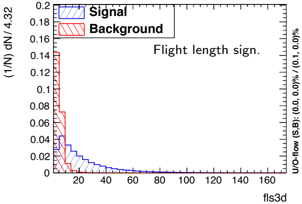


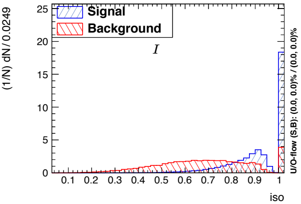


- Combinatorial background is made by combining two uncorrelated muons
  - their secondary vertex is badly reconstructed
  - not compatible with coming from an existing primary vertex
  - there are other particles around them (not isolated)
- These features are all used to separate signal from bkg with a multivariate analysis (BDT) with 10 variables



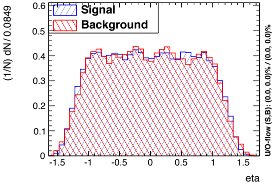


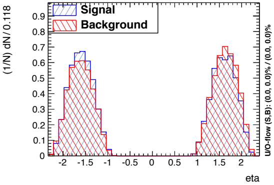




- The two muon tracks are combined to form the B candidate

Definition	Invariant mass range (GeV)
overall mass window	$[4.9, 5.9]$
$B^0 \to \mu^+\mu^-$ signal window	$[5.2, 5.3]$
$B^0_s \to \mu^+\mu^-$ signal window	$[5.3, 5.45]$
blind window	$[5.2, 5.45]$
side-band windows	$[4.9, 5.2] \cup [5.45, 5.9]$





- The number of entries in each mass range is a random variable satisfying Poissonian statistics
  - mass range: $[4.9, 5.2, 5.3, 5.45, 5.9]$ GeV
- total yield is a sum of Poissonian variables
  - comb + rare + signals
- the expected number of reconstructed decays is

$$\nu_i = \frac{\mathcal{B}_{\rm SM}(B^0_{(s)} \to \mu^+\mu^-)}{\mathcal{B}(B^\pm \to J/\psi K^\pm) \times \mathcal{B}(J/\psi \to \mu^+\mu^-)} \frac{f_{\rm s}}{f_{\rm u}} \frac{\epsilon_{B^0_s}}{\epsilon_{B^\pm}} N^{B^\pm}_{\rm obs}$$

- the likelihood is the product of all Poissonians times the constraints (bifurcated Gaussians) for systematics

Binned analysis for the $\mathcal{B}(B^0 \to \mu^+\mu^-)$ upper limit

- Yield is a sum of Poissonian variables: comb. + rare + signals
- Likelihood is the product of all Poissonians times Gaussian constraints
- Upper limit extracted with the $\mathrm{CL_S}$ method

		$N^{\text{exp}}_{\text{signal}}$	$N^{\text{exp}}_{\text{total}}$
2011	B^0 Barrel	0.27 ± 0.03	1.3 ± 0.8
	B^0_s Barrel	2.97 ± 0.44	3.6 ± 0.6
	B^0 Endcap	0.11 ± 0.01	1.5 ± 0.6
	B^0_s Endcap	1.28 ± 0.19	2.6 ± 0.5
2012	B^0 Barrel	1.00 ± 0.10	7.9 ± 3.0
	B^0_s Barrel	11.46 ± 1.72	17.9 ± 2.8
	B^0 Endcap	0.30 ± 0.03	2.2 ± 0.8
	B^0_s Endcap	3.56 ± 0.53	5.1 ± 0.7

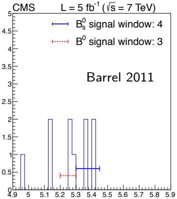
- $B^0 \to \mu^+\mu^-$ BF expected UL:

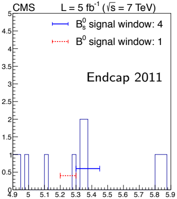
$6.3^{+2.7}_{-2.0} \times 10^{-10}$ SM

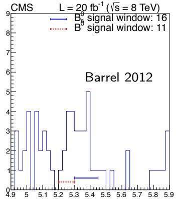
$5.4^{+2.3}_{-1.6} \times 10^{-10}$ Background only

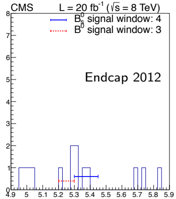
Results: the unblinding

- After the final choice of all selections, the analysis is "unblinded"
- Results on the searches can be extracted









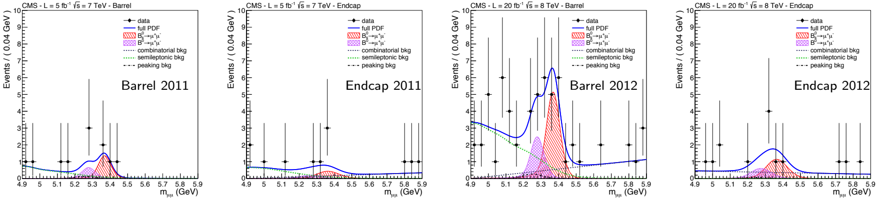
Category	$N_{B^0}^{\mathrm{exp}}$	$N_{B^0}^{\mathrm{obs}}$	$N_{B_s^0}^{\mathrm{exp}}$	$N_{B_s^0}^{\mathrm{obs}}$	$N_{B^\pm}^{\mathrm{obs}}$
Barrel 2011	1.3 ± 0.8	3	3.6 ± 0.6	4	$(71.2 \pm 4.1) \times 10^3$
Endcap 2011	1.5 ± 0.6	1	2.6 ± 0.5	4	$(21.4 \pm 1.1) \times 10^3$
Barrel 2012	7.9 ± 3.0	11	17.9 ± 2.8	16	$(309 \pm 16) \times 10^3$
Endcap 2012	2.2 ± 0.8	3	5.1 ± 0.7	4	$(69.3 \pm 3.5) \times 10^3$

- Fits performed with both 1D-BDT and categorized BDT methods
- Categorized BDT fit results:

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (3.0^{+1.0}_{-0.9}) \times 10^{-9} \qquad [\mathrm{SM}: (3.66 \pm 0.23) \times 10^{-9}]$$

$$\mathcal{B}(B^0 \to \mu^+\mu^-) = (3.5^{+2.1}_{-1.8}) \times 10^{-10} \qquad [\mathrm{SM}: (1.06 \pm 0.09) \times 10^{-10}]$$

- 1D-BDT fits:



Future of the measurements

- $\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-)$ are dominated by statistical uncertainty
- More data will improve the measurements
  - the LHC will deliver hundred and thousand of fb^{-1} to CMS at $\sqrt{s} \approx 14$ TeV

$\mathcal{L}$(fb^{-1})	$\delta\mathcal{B}/\mathcal{B}(B^0_s)$	$\delta\mathcal{B}/\mathcal{B}(B^0)$	B^0 sign.	$\delta\frac{\mathcal{B}(B^0 \to \mu^+\mu^-)}{\mathcal{B}(B^0_s \to \mu^+\mu^-)}$
100 (2015–2017)	15%	66%	0.5–2.4 σ	71%
300 (2019–2021)	12%	45%	1.3–3.3 σ	47%
3000 (2023–)	12%	18%	5.4–7.6 σ	21%

- CMS upgrades on improved muon and tracker trigger capabilities are essential
- The most challenging aspects will be the trigger rates and the pile-up

  $\rightarrow$ cope with up to $< \mu >= 140$ PVs