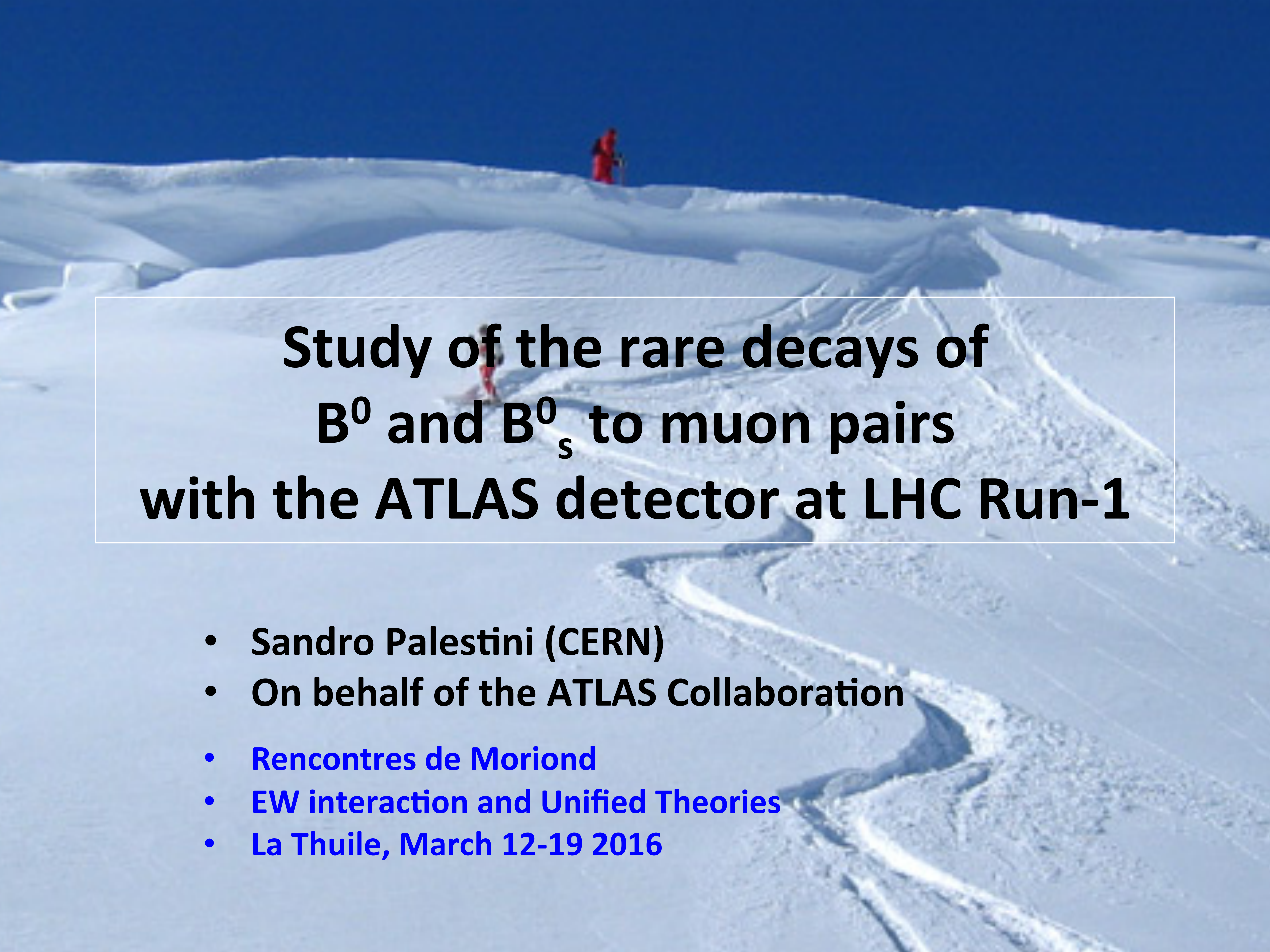


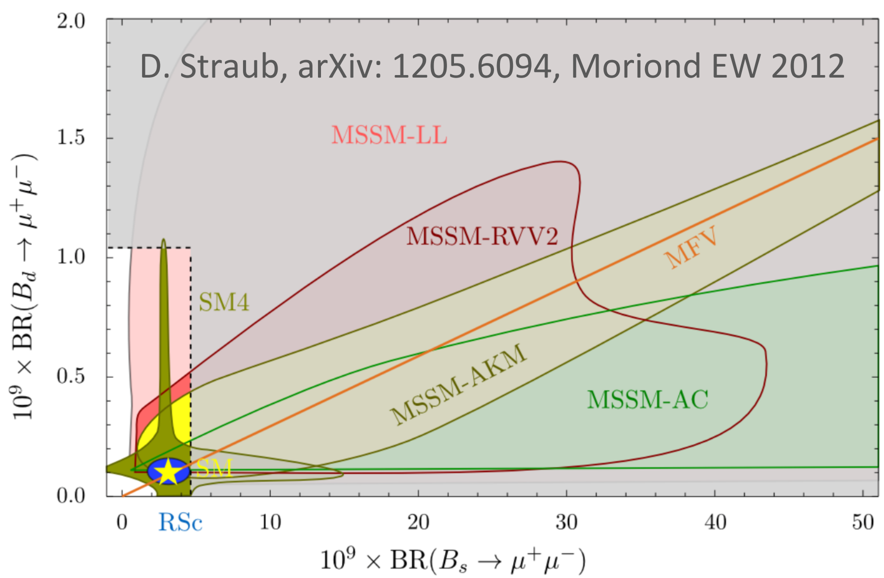
Study of the rare decays of B^0 and B^0_s to muon pairs with the ATLAS detector at LHC Run-1

- Sandro Palestini (CERN)
- On behalf of the ATLAS Collaboration

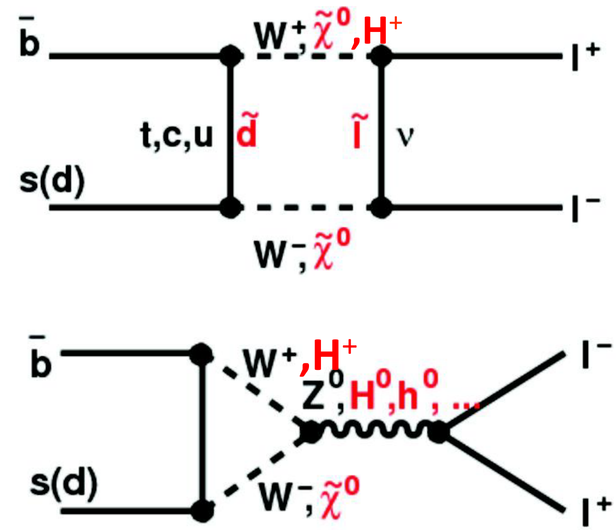
- Rencontres de Moriond
- EW interaction and Unified Theories
- La Thuile, March 12-19 2016

Motivation and available results

- FCNC process, further affected by helicity suppression, and predicted accurately in the SM:
  - BR($B^0_s \to \mu^+\mu^-$)=(3.65+-0.23)×10^{-9},
  - BR($B^0 \to \mu^+\mu^-$) =(1.06+-0.09)×10^{-10} [C. Bobeth et al., PRL 112 (2104) 101801]
- Hence, sensitive to physics BSM



Examples of diagrams with BSM contributions



CMS and LHCb combined result:

- $BR(B^0_s) = (2.8{+}0.7{-}0.6)\times10^{-9}$
- $BR(B^0) = (3.9{+}1.6{-}1.4)\times10^{-10}$

  [Nature 522 (2015) 68-72]

(Talk by Sanjay Kumar Swain in this conference)

SM prediction:

- $BR(B^0_s)=(3.65{+-}0.23)\times10^{-9}$,
- $BR(B^0)=(1.06{+-}0.09)\times10^{-10}$

  [C. Bobeth et al., PRL 112 (2104) 101801]

New results from ATLAS are presented today,
based on 5+20 fb^{-1} collected in LHC Run 1,
with improved analysis techniques over previous results from ATLAS.

Analysis flow

- Trigger: muon pairs: $p_T(\mu)$ > 4 GeV in 2011, > 4 GeV or 6 GeV in 2012
- Require track reconstruction in both the inner detector and the muon spectrometer (*combined muons*)
- Reconstruct signal $B^0_{(s)} \to \mu^+\mu^-$

  and reference channels: $B^+ \to J/\psi K^+$, $B^0_s \to J/\psi \phi$ ($J/\psi \to \mu^+\mu^-$, $\phi \to K^+K^-$)
- m(μ+μ-) range for signal: 4766 to 5966 MeV, interval 5166 to 5526 MeV blinded until the selection and the entire analysis chain is defined.
- $p_T(B)$>8 GeV, $|\eta(B)|$<2.5.
- Enhance signal to background ratio by means of two MVA classifiers:
  - continuum-BDT, for the combinatorial background
  - fakes-BDT against the background due to hadron misidentification
- Signal extraction with likelihood fit over mass distribution, in different intervals of continuum-BDT output
- Normalization with B^+ yield and efficiency ratio $B^+/B^0_{(s)}$.

Normalization

The yield N_s , N_d of B^0_s->μ+μ- , B^0 ->μ+μ- events is normalized to the channel:

$$B^+ \to J/\psi(\to \mu^+\mu^-)K^+$$

using the equation:

$$\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-) = \frac{N_{d(s)}}{\varepsilon_{\mu^+\mu^-}} \times [\mathcal{B}(B^+ \to J/\psi K^+) \times \mathcal{B}(J/\psi \to \mu^+\mu^-)] \frac{\varepsilon_{J/\psi K^+}}{N_{J/\psi K^+}} \times \frac{f_u}{f_{d(s)}}$$

In more detail, the normalization is performed separately in four classes corresponding to three trigger selections used in 2012 and one in 2011:

$$\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-) = N_{d(s)} \times [\mathcal{B}(B^+ \to J/\psi K^+) \times \mathcal{B}(J/\psi \to \mu^+\mu^-)] \times \frac{f_u}{f_{d(s)}} \times \frac{1}{\mathcal{D}_{\text{norm}}}$$

with the normalization term including efficiency ratios and pre-scaling factors α_k common to the two channels.

$$\mathcal{D}_{\text{norm}} = \sum_k N^k_{J/\psi K^+} \alpha_k \left(\frac{\varepsilon_{\mu^+\mu^-}}{\varepsilon_{J/\psi K^\pm}}\right)_k$$

Systematic uncertainties related to prescaling factors, trigger and reconstruction efficiency are minimized with this procedure

Background

The key-word in these kind of studies.

In order of relative amplitude:

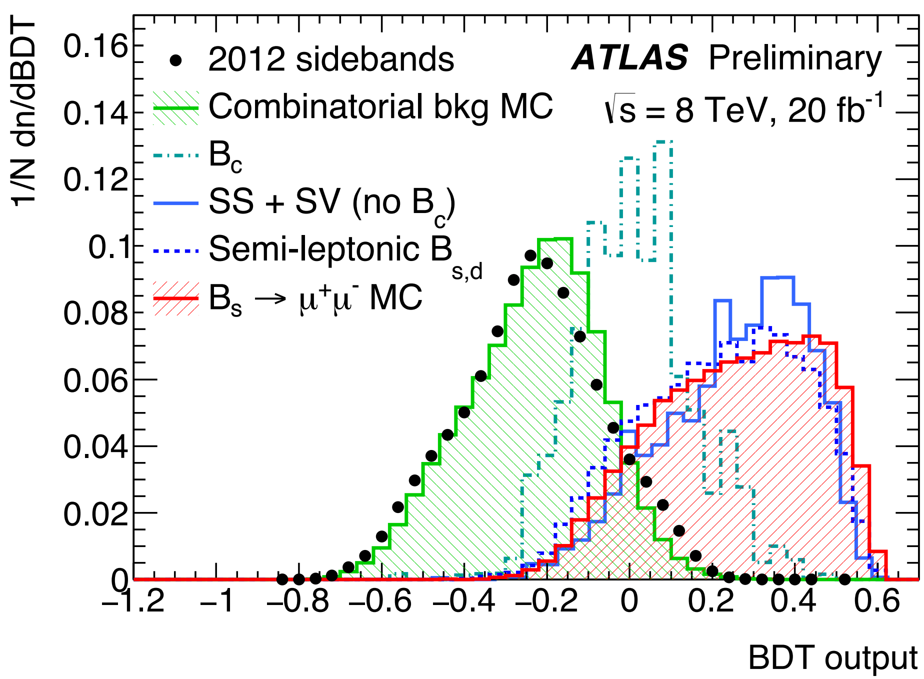
1. Combinatorial background: dominant component, small mass dependence
2. Partially reconstructed *B*-> μμ*X* decays: low mass, the tail extend to the signal region (Named Same Vertex / Same Side background SS+SV)
3. Peaking background from double-hadrons misidentification in $B^0_{(s)}$-> hh′ (smaller component, but overlaid with signal)

Combinatorial background and *continuum-BDT*

The dominant background due muon pair originate from the uncorrelated decays of hadrons produced in the hadronization of a *b* and a $\bar{b}$ quark (or *c*, $\bar{c}$ quark).

Signal and combinatorial background are separated with a MVA classifier:

- 15 variables related to the B candidate, to the muons forming the candidate, to the other tracks from the same collision and to pile-up primary vertices



Normalization of combinatorial bkg. In data is ≈3000 higher than SS+SV component,
Semileptonic (with one fake-muon) account for ≅30% of SV, B_c component is lower

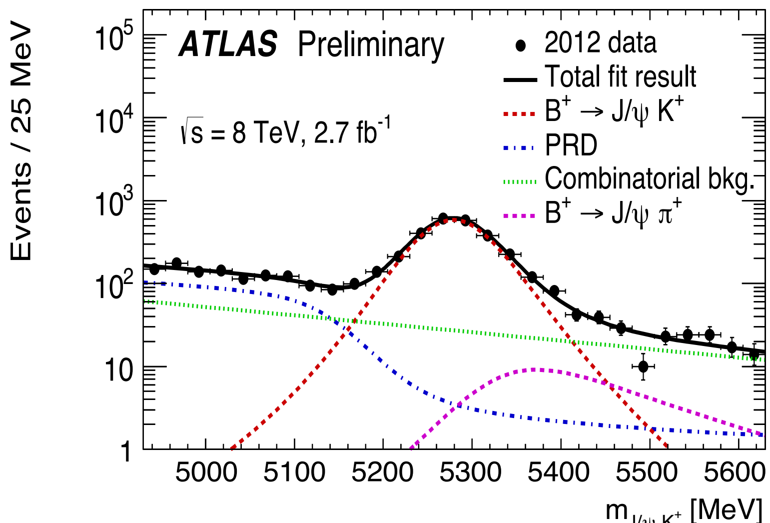
Training of the BDT on a large MC sample (1.4 G events) of uncorrelated *b-(c-)* hadrons and *$\bar{b}$-($\bar{c}$-)* hadrons with forced decays into final states containing muons.

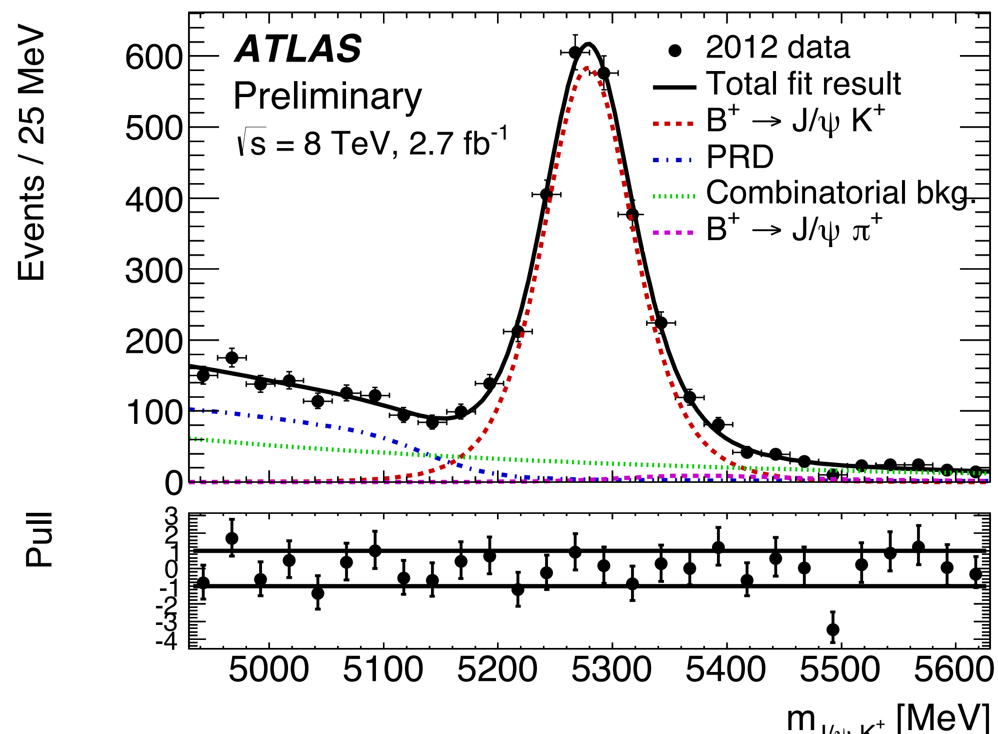
B^+ yield

$$\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-) = N_{d(s)} \times [\mathcal{B}(B^+ \to J/\psi K^+) \times \mathcal{B}(J/\psi \to \mu^+\mu^-)] \times \frac{f_u}{f_{d(s)}} \times \frac{1}{\mathcal{D}_{\text{norm}}}$$

$$\mathcal{D}_{\text{norm}} = \sum_k N^k_{J/\psi K^+} \alpha_k \left(\frac{\varepsilon_{\mu^+\mu^-}}{\varepsilon_{J/\psi K^\pm}} \right)_k$$

- Extracted after fakes-BDT and continuum-BDT selection
- Extracted separately in the 4 categories
- Unbinned fit in mass distribution maximum likelihood fit.



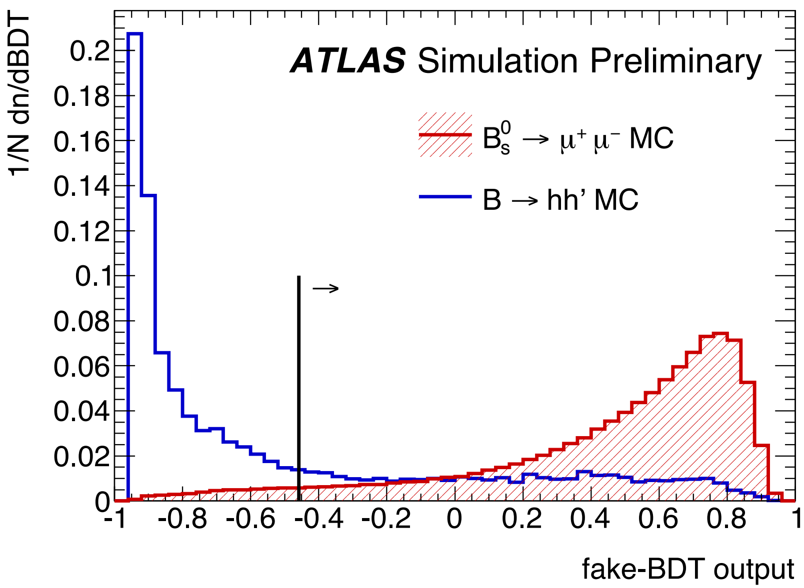


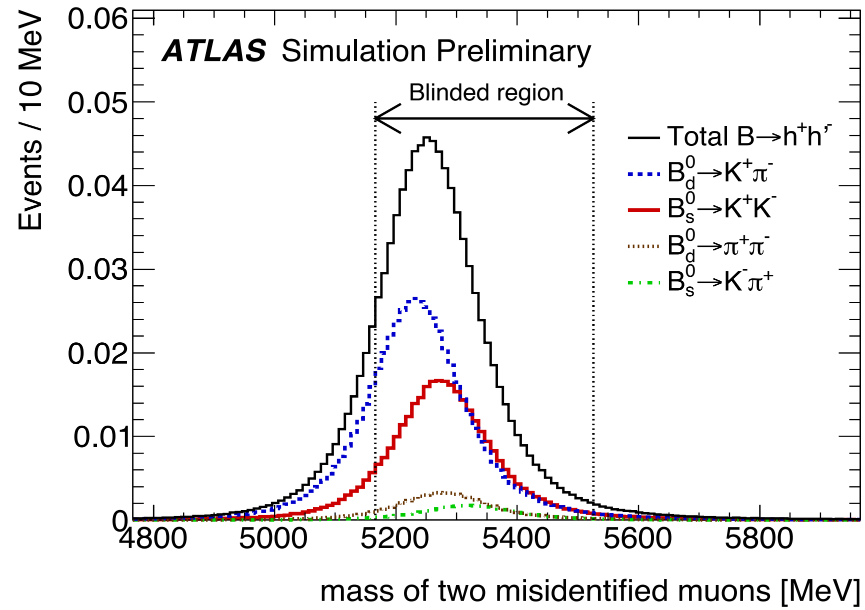
B^+ signal in the main category of the 2012 sample

Hadron-to-muon misidentification and peaking background

- Studied with full simulation of of the B-> hh′ , Λ_b -> ph and of the detector response,
- validated with data from ϕ -> KK and B^+ -> J/ψ K^+ decay.
- The probability of misidentification as muon is low for *combined muons*: about 0.28% for kaons and 0.12% for pions.

This fraction is reduced by a factor 0.4 with a dedicated classifier (*fakes-BDT)* with an efficiency of prompt muons set at 95%.





Use B^+ -> J/ψ K^+ yield and efficiency ratio from master formula: the total number of peaking-background events is 1.0±0.4

Efficiency ratio $B^+/B^0_{(s)}$

$$\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-) = N_{d(s)} \times [\mathcal{B}(B^+ \to J/\psi K^+) \times \mathcal{B}(J/\psi \to \mu^+\mu^-)] \times \frac{f_u}{f_{d(s)}} \times \frac{1}{\mathcal{D}_{\text{norm}}}$$

$$\mathcal{D}_{\text{norm}} = \sum_k N^k_{J/\psi K^+} \alpha_k \left(\frac{\varepsilon_{\mu^+\mu^-}}{\varepsilon_{J/\psi K^\pm}} \right)_k$$

In each category the efficiency ratio is obtained from MC.

- p_T and η spectra are tuned on the reference channels.
- Trigger efficiencies are extracted from *tag* & *probe* studies based on J/ψ and Υ.

MC-to-data comparison done also on the discriminating variables used in the continuum BDT:

- only Isolation (based on p_T of tracks within a cone of ΔR<0.7) requires tuning, in the B^+ mode.

For B^+/B^0_s :

- additional correction due to difference between mean B^0_s lifetime and $B_H^{(s)}$ lifetime (SM prediction for $\mu^+\mu^-$ decay)

  - Total correction to efficiency ratio: +3.4% for B^0, -0.6% for B^0_s
  - Total systematic uncertainty ±5.9%

Extraction of $B^0_{(s)} \to \mu^+\mu^-$ signal

$$\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-) = N_{d(s)} \times [\mathcal{B}(B^+ \to J/\psi K^+) \times \mathcal{B}(J/\psi \to \mu^+\mu^-)] \times \frac{f_u}{f_{d(s)}} \times \frac{1}{\mathcal{D}_{\text{norm}}}$$

$$\mathcal{D}_{\text{norm}} = \sum_k N^k_{J/\psi K^+} \alpha_k \left(\frac{\varepsilon_{\mu^+\mu^-}}{\varepsilon_{J/\psi K^\pm}} \right)_k$$

N_d and N_s are extracted with a likelihood fit to signal and background, in the m(μ+μ-) mass distribution, performed in three intervals of continuum-BDT output

Signal and background model:

- Signal: two superimposed Gaussians (average width: 80 MeV, independent of continuum-BDT).
- Low mass background: exponential dependence on m(μ+μ-) , determined on data in mass sideband, independent of continuum-BDT),
- Continuum background: linear dependence on mass
  (very small correlation between mass and continuum-BDT; sideband data consistent with MC)
- Peaking background: Gaussian mass dependence, equal amplitude in the BDT bins, normalization 1.0±0.4 events in total.

Systematic uncertainties in the likelihood fit

For the signal model:

- Mass scale and resolution uncertainties
- Uncertainties in the relative efficiencies of the continuum BDT-intervals
  - Studied on the distributions of the BDT-output and of the discriminating variables in the reference channels (calibration and systematic uncertainty)

For the background model:

- Explicit modeling of 3-body semileptonic decays
- Mass dependence of combinatorial and SS+SV components
- Correlation between mass dependences and BDT dependences

The systematic uncertainties are smaller than the statistical uncertainties (by factor ≈5 for B^0_s, ≈ 2.5 to 5 for B^0 - depending on signal amplitude)

Likelihood fit to branching fractions

$$\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-) = N_{d(s)} \times [\mathcal{B}(B^+ \to J/\psi K^+) \times \mathcal{B}(J/\psi \to \mu^+\mu^-)] \times \frac{f_u}{f_{d(s)}} \times \frac{1}{\mathcal{D}_{\text{norm}}}$$

$$\mathcal{D}_{\text{norm}} = \sum_k N^k_{J/\psi K^+} \alpha_k \left(\frac{\varepsilon_{\mu^+\mu^-}}{\varepsilon_{J/\psi K^\pm}}\right)_k$$

The normalization includes:

- B^+ branching fraction (world averages)
- The production ratio for f_u/f_s from the ATLAS measurement of f_s/f_d performed in the same p_T, η range [PRL 115(2015)262001]
- The efficiency ratios and B^+ yields in the $\mathcal{D}_{\text{norm}}$ term

The total uncertainty in the normalisation is

- ±11% for BR($B^0_s \to \mu^+\mu^-$)
- ± 7% for BR($B^0 \to \mu^+\mu^-$)

Interpolated background and expected signal

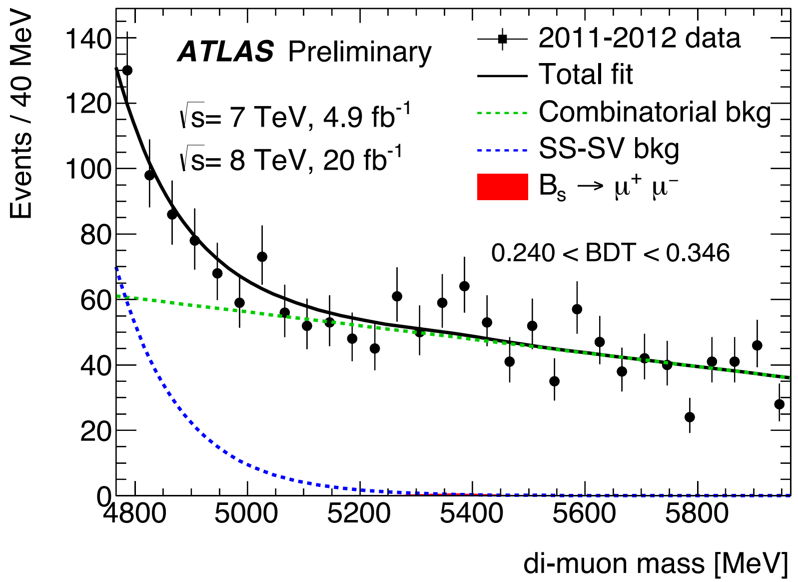
Number of background events interpolated in the blinded region, for each of the three bins in the continuum-BDT output:

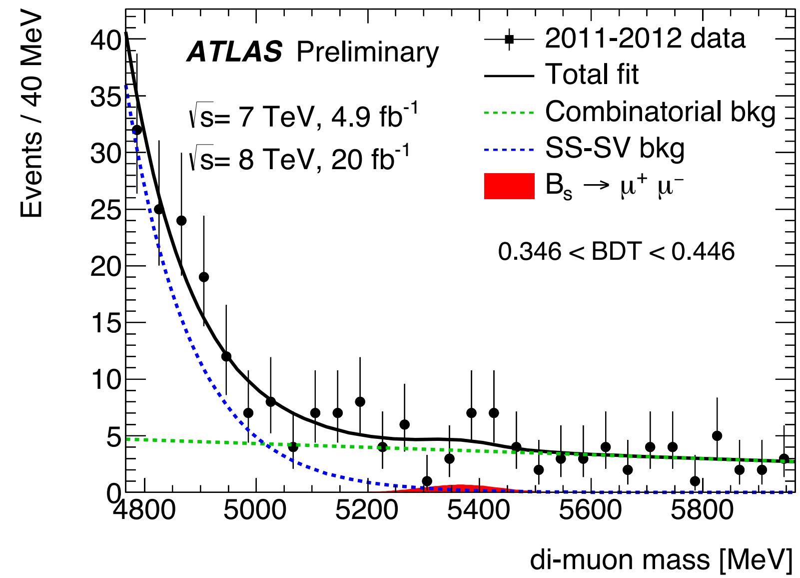
- in bin-1, 0.240 to 0.346: 510±29 background events
- in bin-2, 0.346 to 0.446: 32±6 background events
- In bin-3, 0.446 to 0.7 5±2 background events

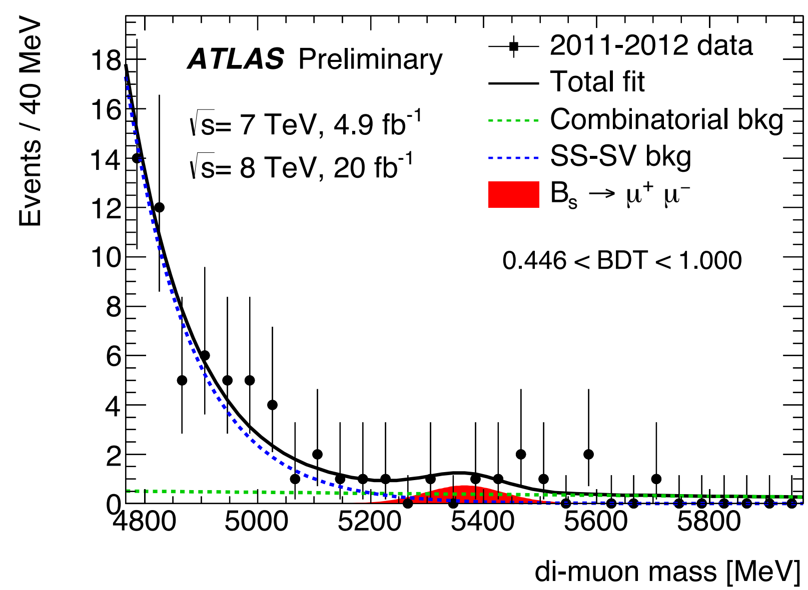
Expected signal from SM prediction:
41 $B^0_{(s)}$ and 5 B^0 events
equally distributed in the three bins.

And now we are ready to look into the signal region.

$B^0_s \rightarrow \mu^+\mu^-$ and $B^0 \rightarrow \mu^+\mu^-$ event yield

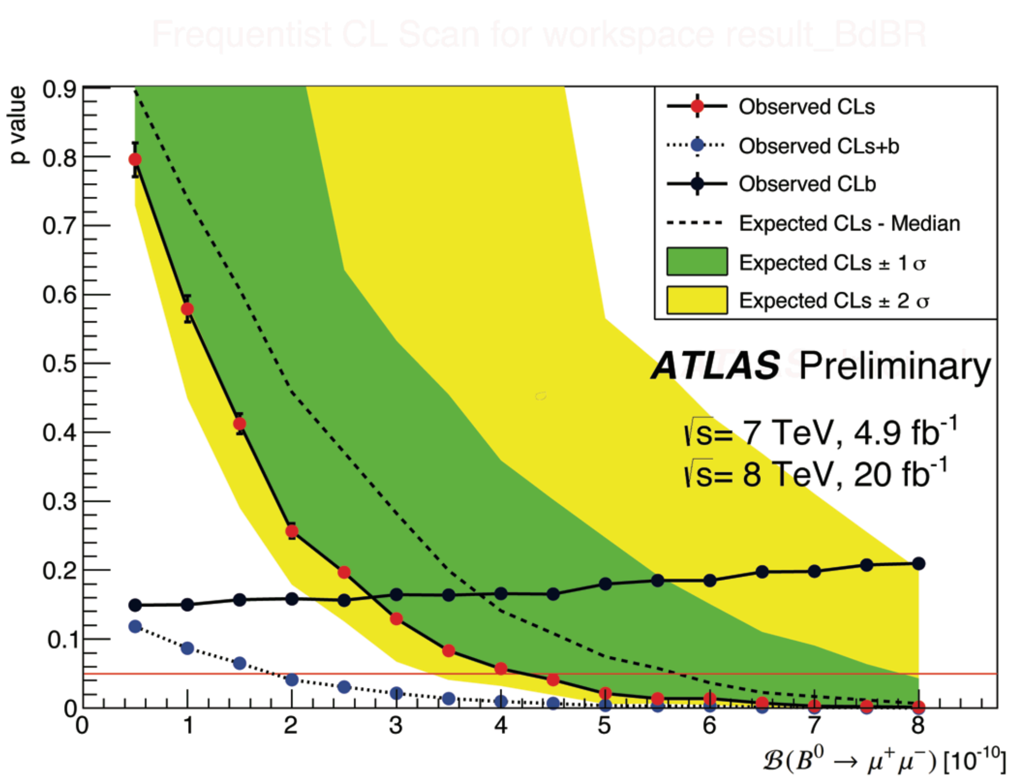






Fewer B^0_s events than expected, no B^0 events

Likelihood fit to branching fractions: Results - 1



Result for BR($B^0 \to \mu^+\mu^-$) :

upper limit set using CL_s technique, with pseudo-MC experiments

- No signal, BR($B^0_s \to \mu^+\mu^-$) left free to be determined in the fit
- BR($B^0 \to \mu^+\mu^-$) $< 4.2 \times 10^{-10}$ at 95% CL
- CL_b is ≈ 0.15 for BR($B^0 \to \mu^+\mu^-$) near 0: -1σ fluctuation of background, expected limit $< 5.7^{+2.1}_{-1.2} \times 10^{-10}$

The limit is higher than the SM prediction BR(B^0)$_{SM}$=(1.06+-0.09)×10^{-10}

The expected significance for BR($B^0 \to \mu^+\mu^-$), assuming the SM branching fraction, is 0.2 σ.

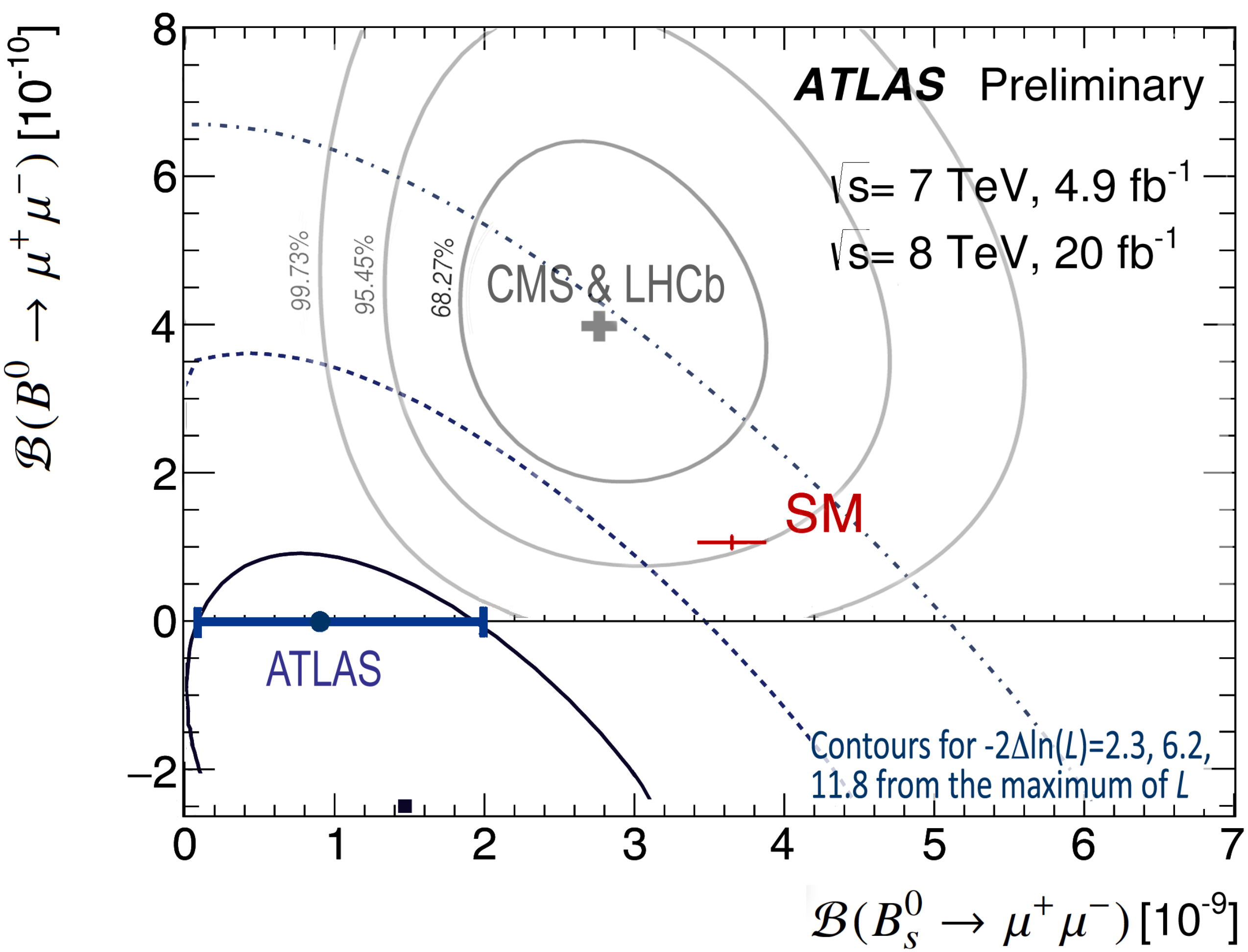
Likelihood fit to branching fractions: results - 2

Result for $BR(B^0_s \to \mu^+\mu^-)$: $BR(B^0_s \to \mu^+\mu^-) = 0.9^{+1.1}_{-0.8} \times 10^{-9}$

- The central value is obtained within the boundary of non-negative branching fractions
- The errors are obtained by means of a frequentist belt, using pseudo-MC experiments, and include both statistic and systematic error. The systematic uncertainty is $\sigma_{syst} = \pm 0.3 \times 10^{-9}$.
- The upper limit from CL_s is $BR(B^0_s \to \mu^+\mu^-) < 3.0 \times 10^{-9}$ at 95% CL

  The expected limit (no signal, $BR(B^0 \to \mu^+\mu^-)$ determined in the fit) is $< 1.8^{+0.7}_{-0.4} \times 10^{-9}$.
- The observed compatibility with the null hypothesis (no signal, $BR(B^0 \to \mu^+\mu^-)$ determined in the fit) corresponds to $p = 0.08$ (1.4 σ). The expected significance assuming the SM branching fraction is 3.1 σ.
- Compatibility with SM ($BR(B^0_s)_{SM} = (3.65 \pm 0.23) \times 10^{-9}$, $BR(B^0)_{SM} = (1.06 \pm 0.09) \times 10^{-10}$) : *p*-value = 0.048 (2.0 σ) obtained from pseudo-MC experiments, using the likelihood ratio, for the simultaneous fit to $BR(B^0_s \to \mu^+\mu^-)$ and $BR(B^0 \to \mu^+\mu^-)$.

Likelihood contours
without imposing natural boundaries



The contours corresponding to -2Δln(*L*)=2.3, 6.2, 11.8 are shown relative to the absolute maximum of *L*, regardless of its position outside of the natural boundary.

The minimum within the boundary of non-negative branching fraction is shown with the error bar for the frequentist 68% confidence range for BR($B^0_s \to \mu^+\mu^-$)

Also shown are the contours from the combination of CMS and LHCb [Nature 522 (2015) 68-72]

Conclusions:

Using the data collected in Run-1, ATLAS has today preliminary results on the rare decays of B^0_s and B^0 into muon pairs.

For B^0_s:

- $BR(B^0_s \to \mu^+\mu^-) = 0.9^{+1.1}_{-0.8} \times 10^{-9}$
- $< 3.0 \times 10^{-9}$ at 95% CL (from CL_s)
- The limit is lower that the SM prediction $(BR(B^0_s)_{SM}=(3.65+-0.23)\times10^{-9})$
- The result is lower than the central value of the CMS & LHCb combination, but the difference to the central value is smaller $(BR(B^0_s)_{CMS\ \&\ LHCb} = (2.8^{+0.7}_{-0.6})\times10^{-9})$

For B^0:

- $BR(B^0 \to \mu^+\mu^-) < 4.2 \times 10^{-10}$ at 95% CL (from CL_s)
- The limit is above the SM prediction
- and reaches the central value of the CMS & LHCb combination $BR(B^0)_{CMS\&LHCB} = (3.9^{+1.6}_{-1.4})\times10^{-10}$.

The compatibility with the SM, for the simultaneous fit, is 2.0 σ .