XIIIth International Conference on Heavy Quarks and Leptons

| Center for Neutrino Physics | Blacksburg, Virginia | 22th - 27th May 2016 **WirginiaTech**



 $B_{D/s} \rightarrow \mu^+ \mu^- \text{ in ATLAS}$

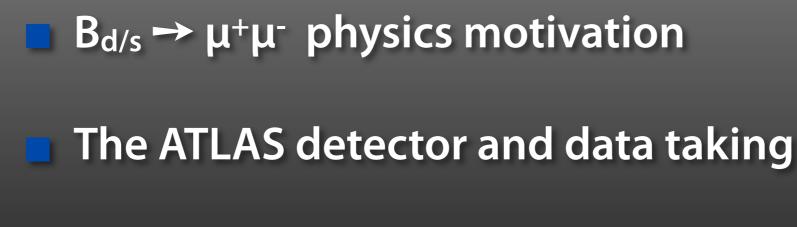
Jaroslav Günther

on behalf of the ATLAS Collaboration



Leopold-Franzens-Universität Innsbruck, Austria

Outline



B_{d/s} \rightarrow $\mu^+\mu^-$ analysis overview

Backgrounds

B[±] \rightarrow J/ Ψ K[±] yield extraction

(J/ $\Psi \pi^{\pm}$) / (J/ ΨK^{\pm}) ratio measurement

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

Results of BR($B_{d/s} \rightarrow \mu^+\mu^-$) measurement

$B_{D/s} \rightarrow \mu^+ \mu^-$ decay

Motivation

accurate SM prediction - helicity suppressed FCNC

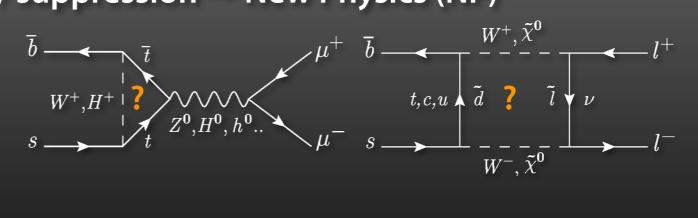
observation of decay rate enhancement / suppression → New Physics (NP)

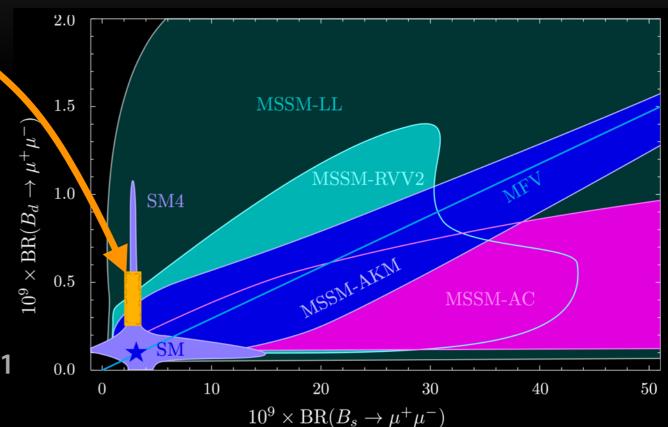
strong QCD-free constraint on NP
genuine probe of Yukawa interactions
EW precision test (wrt. Z penguin)

Evidences of CMS+LHCb combined [1] BR(B_s $\rightarrow \mu^{+}\mu^{-}) = (2.8 + 0.7 - 0.6) \times 10^{-9}$ BR(B_d $\rightarrow \mu^{+}\mu^{-}) = (3.9 + 1.6 - 1.4) \times 10^{-10}$

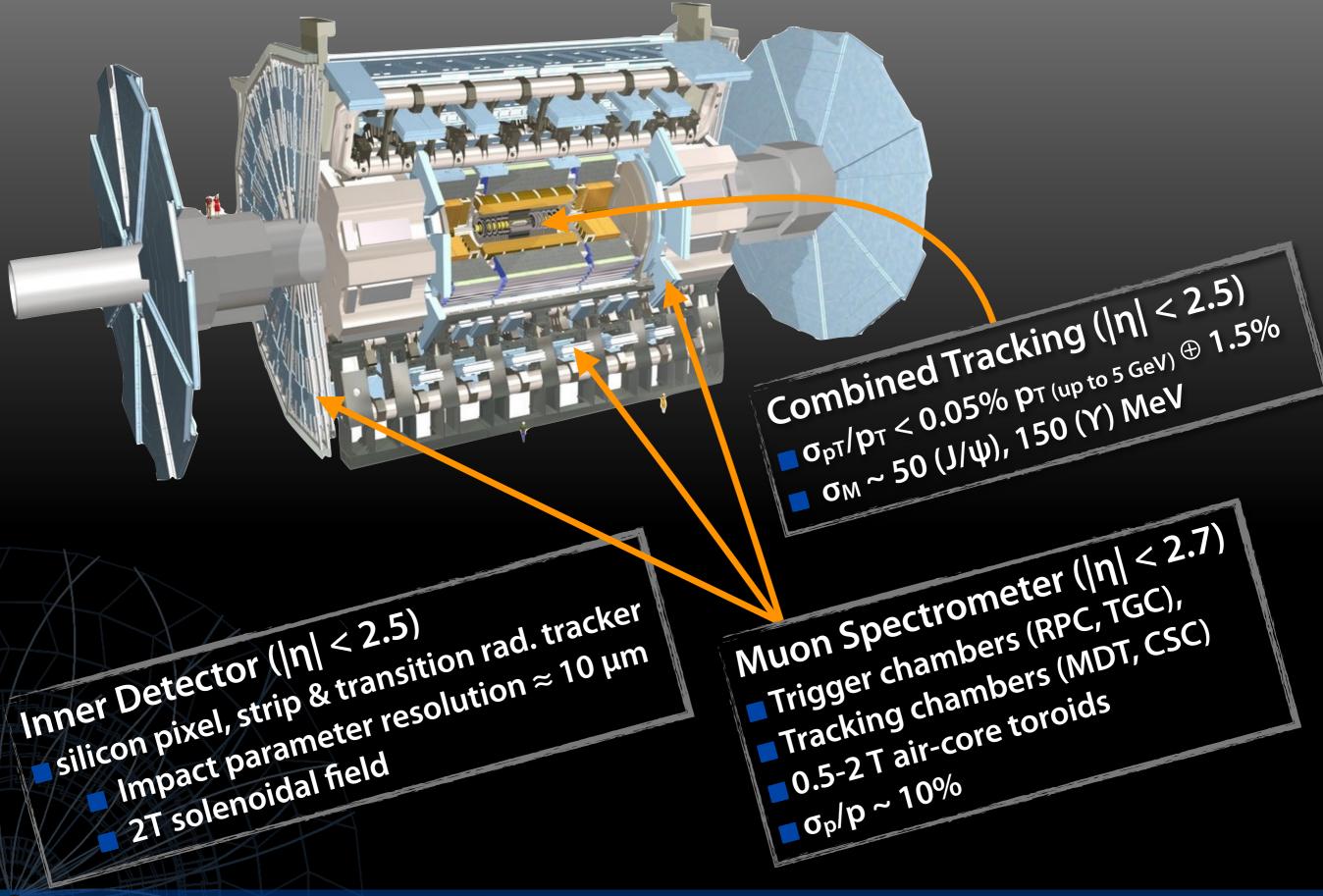
Standard Model [2] BR_{SM}(B_s → $\mu^+\mu^-$) = (3.65 ± 0.23) x 10⁻⁹ BR_{SM}(B_d → $\mu^+\mu^-$) = (1.06 ± 0.09) x 10⁻¹⁰

[1] Nature 522 (2015) 68-72
[2] C. Bobeth et al., Phys. Rev. Lett. 112 (2014) 101801
[3] Buras et al., Eur.Phys.J. C72 (2012) 2172





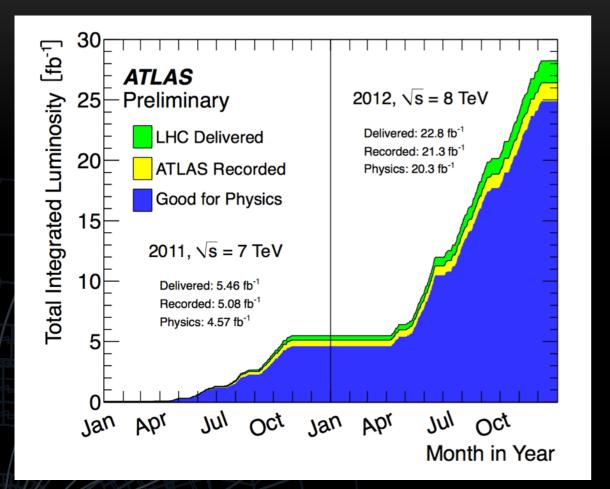
THE ATLAS DETECTOR

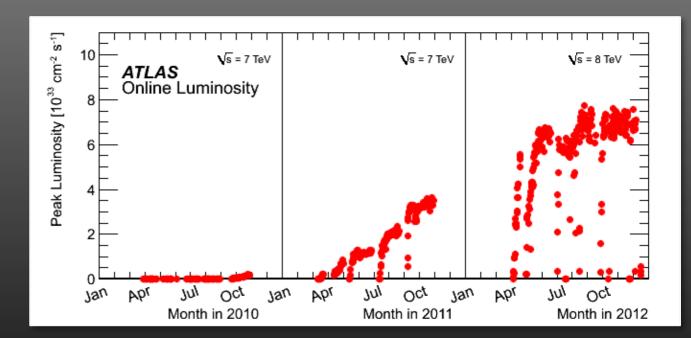


Run I Data Taking

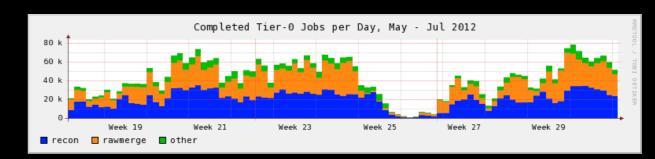
2011 ~ 4.6 fb⁻¹
instantaneous luminosity & pile-up steadily increasing

2012 ~ 20.4 fb⁻¹
Flatter instantaneous luminosity profile
Challenging pile-up conditions !





di-muon trigger event selection Tier-0 processing ~24 PB of RAW and derived data Up-to 80k Tier-0 jobs completed/day



$B_{D/s} \rightarrow \mu + \mu$ - strategy @ ATLAS

Analysis Features

Blind analysis technique: exclude mass region 5166 < m(μ+μ-) < 5526 MeV Multivariate analysis - 2 BDTs for background suppression

data control samples used for x-checks & bkg understanding

Relative branching ratio measurement

reference signal decay $B^{\pm} \rightarrow J/\Psi K^{\pm} \rightarrow \mu + \mu - K^{\pm}$ (large stat.)

partial cancelation of systematics (on luminosity, cross-sec., efficiencies, trigger)

BR (B[±] \rightarrow J/ $\Psi\pi^{\pm}$)/BR(B[±] \rightarrow J/ Ψ K[±]) = natural parallel measurement

$$\mathcal{BR}(B_{d(s)} o \mu^+ \mu^-) = ~~ \mathcal{BR}(B^\pm o J/\psi K^\pm o \mu^+ \mu^- K^\pm) imes igg| rac{f_u}{f_{d(s)}}$$

Hadronisation probabilities Yields from likelihood mass fits Trigger and luminosity weight factors (Acceptance x Efficiency) ratio of signal/ref. channel \mathcal{D}_{norm} Normalisation term $j \in (4 \text{ event categories})$

$$\times N_{\mu+\mu^{-}} \times \left[\sum_{j} \left(N_{J/\Psi K}^{j} + \times \alpha_{j}\right)\right]^{-1} \times \left(\frac{\epsilon_{\mu+\mu^{-}}}{\epsilon_{J/\Psi K}}\right)_{j}^{-1}$$

$B_{D/s} \rightarrow \mu + \mu$ - strategy @ ATLAS

Analysis Features

Blind analysis technique: exclude mass region 5166 < m($\mu^+\mu^-$) < 5526 MeV

Multivariate analysis - 2 BDTs for background suppression

data control samples used for x-checks & bkg understanding

Relative branching ratio measurement

reference signal decay $B^{\pm} \rightarrow J/\Psi K^{\pm} \rightarrow \mu + \mu - K^{\pm}$ (large stat.)

partial cancelation of systematics (on luminosity, cross-sec., efficiencies, trigger)

BR (B[±] \rightarrow J/ $\Psi\pi^{\pm}$)/BR(B[±] \rightarrow J/ Ψ K[±]) = natural parallel measurement

$$\mathcal{BR}(B_{d(s)} \to \mu^+ \mu^-) = \frac{\mathcal{BR}(B^{\pm} \to J/\psi K^{\pm} \to \mu^+ \mu^- K^{\pm}) \times \frac{f_u}{f_{d(s)}}}{\frac{\mathsf{PDG}, \mathsf{ATLAS}}{\mathsf{PDG}, \mathsf{ATLAS}}} \, (\mathsf{f}_{\mathsf{s}}/\mathsf{f}_{\mathsf{d}}; \mathsf{assuming}} \, \mathsf{f}_{\mathsf{u}}/\mathsf{f}_{\mathsf{d}} = 1)$$

<u>Data</u>

$$imes \ N_{\mu^+\mu^-} imes \Big[\sum_{j} (N^j_{J/\Psi K^\pm} imes lpha_j]$$

N $_{\mu+\mu-}$ - unbinned extended ML fit simultaneously in 3 bins of c-BDT classifier N $_{J/\Psi K\pm}$ - 4 unbinned extended ML fits (data + MC fit simultaneously)

Simulation - (efficiency x acceptance) terms ε_x

"calibrated" on data
 systematics from Data/MC discrepancies

$$imes \Big(rac{\epsilon_{\mu^+\mu^-}}{\epsilon_{J/\Psi K^\pm}}\Big)_{m j}\Big]^{-1}$$

Event Selection

<u>4 mutually exclusive trigger cate</u>	<u>gories</u>	
(to avoid efficiency loss)		
2011 data		
 2011: 2 μ, each with p_T > 4 GeV 		(22%)
2012 data		
2. T1 : 2 μ, with p _{T,µ1} > 6 GeV	′& p _{T,μ2} > 4 GeV	(68%)
3. T2 : 2 μ , each with p_T > 4 GeV		5 (6%)
4. T3 : 2 μ , each with p _T > 4 GeV	& both in η > 1.05	(4%)
<u>Cuts</u>	Additional cuts for all	<u>channels</u>
<mark>_</mark> m(μμ) ∈ [4766,5966]GeV	(~5% loss in signal, reduce	e bkg. by a factor ~0.4)
Kaons, Muons, B mesons : η < 2.5	ΔR < 1.5	
p _T (μ) > 4 GeV	α2D <1.0	
■ p _T (K) > 1 GeV	■ Lxy > 0	
■ p _T (B) > 8 GeV		
TTT I KA		
ID+MS combined tracking		

mass resolution in end-caps better by 30% wrt. 7 TeV analysis Improved Primary Vertex (PV) identification → 99.8 % efficient selected PV has min. distance (in z) to a point of closest approach of p(B) direction (3D) extrapolated to beam line.

$B_{D/s}$ → µ+µ- Backgrounds

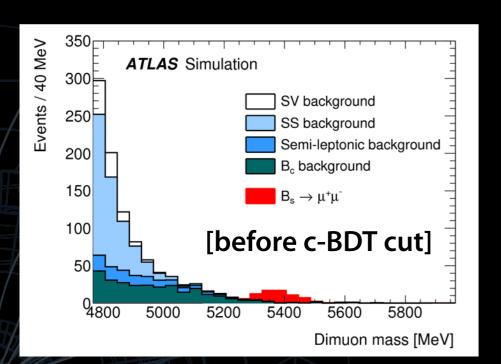
Combinatorial Background

■b**5**→µ⁺µ⁻reduced by <u>c-BDT trained on MC</u> in B_s fit: 1st order polynomial

Partially reconstructed decays

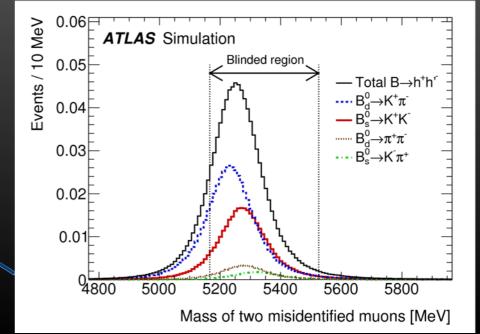
Same-vertex (SV)
e.g.: B⁰_d → K μ⁺μ⁻, B_c→J/ψ(μ⁺μ⁻)μ⁺√
Same-side (SS) different vertices
b → μ⁻c(→ μ⁺X')X
Semi-leptonic

e.g.: $B \rightarrow \mu hv$, h mis-identified as μ^{ℓ} in B_s fit: exponential



Peaking background

B → hh', h = π, K, p mis-identified as μ dangerously mimicks the signal in B_s fit: signal-like PDF



<u>f-BDT trained on MC</u>

reduced bkg. from h mis-identified as μ
performance validated on data (fake K from Φ → K+K⁻, B[±] → J/ΨK[±])
rejection 7 times > wrt. 7 TeV analysis
B_{d/s}→μμ signal eff. set to 95% : expect 1.0 ± 0.4 (MC) fake events [0.2 ± 1.0 (data)] (x-check on signal data w reverse f-BDT cut)

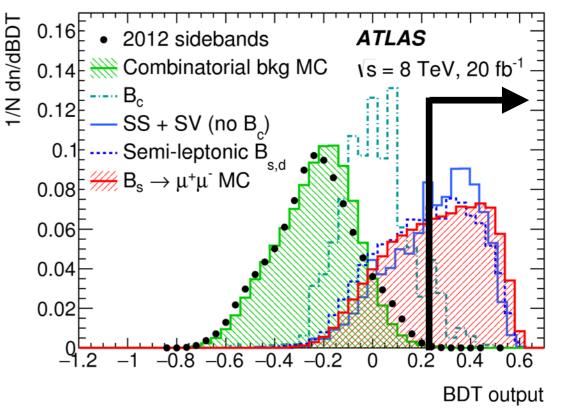
Combinatorial Background Discrimination

 Largest Background component (2000 x higher in data than SS+SV)
 Boosted Decision Tree classifier applied for suppression (c-BDT)
 15 input variables exploiting : B-meson kinematic Good PV-SV separation Δx (cτ ~ 450µm) Collinearity between Δx and P(µµ) None or few additional tracks around the SV µ prop., X² wrt. other vertices, B isolation...

Trained on 1.4G MC

events with µ from b/c quark decays MC has 8 x > stat. than in sideband data left sideband used for c-BDT validation

Combinatorial bkg reduced ~1000 x Signal efficiency 54%



Data/MC validated on control samples

B[±] → J/ΨK[±], B_s → J/ΨΦ
Data/MC discrepancies accounted for as systematic error.

Reference Channel Yield

Unbinned extended maximum likelihood (ML) fits

- $m(\mu^+\mu^-K^\pm)$ Data mass distribution fitted (simultaneously with MC)
- Extract yields of $B^{\pm} \rightarrow J/\Psi \pi^{\pm} \& B^{\pm} \rightarrow J/\Psi K^{\pm}$ at the same time ("+" sign plot notation)
- Several parameters are free to vary
- Normalisations mass scale & resolution slope of combinatorial bkg.

Main systematics

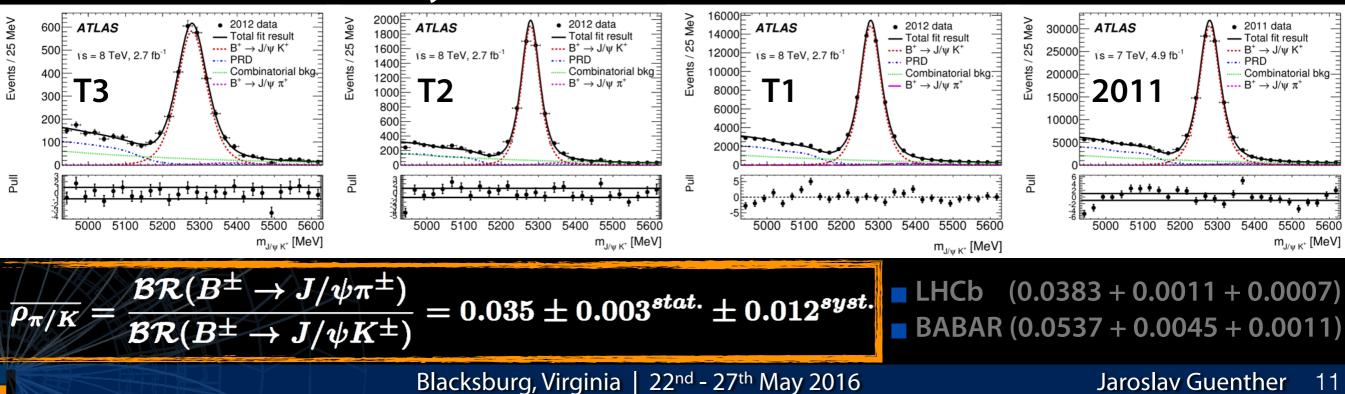
Combinatorial model

MC modelling of PRD

4- component fit partially reconstructed decays (PRD) combinatorial bkg,
J/Ψ π[±] & J/Ψ K[±]
shapes sim. constrained by MC

Signal charge asymmetry (K⁻ reconstructed less efficiently)

Total contribution to $\mathcal{D}_{\mathsf{norm}}$ systematics ± 0.8%



(Acceptance x Efficiency) Ratio of signal/Ref. channel

Extracted from simulation (MC samples)

separately for each channel and category (p_T(B) > 8 GeV, [η(B)] < 2.5) corrections from B⁺ or B_s → J/Ψ Φ data on : p_T(B) and η(B) spectra B_s → μ+μ- lifetime and B isolation trigger efficiencies corrected using J/ψ and Y data (tag & probe studies)

Main Systematic Uncertainty

vertexing (B_s vs B[±]) & track reconstruction K⁺

Residual Data/MC discrepancy on the c-BDT variables (±3.2%)

■ Total systematic uncertainty contribution from B[±] yield and B_s/B[±] acc. x eff. ratio ± 5.9%

Statistical uncertainty in simulation	0.5%
$p_{\rm T}, \eta$ reweighting and trigger efficiency	1.3%
Data to MC discrepancy in discriminating variables	4.2%
K^+ and B^+ reconstruction	3.6%
Residual trigger efficiency systematic uncertainty	1.5%
B^+ yield	0.8%
Total uncertainty	5.9%

Expected Signal Yield

$$\sum_{k} N_{\mu\mu}^{k} = \frac{f_{s}}{f_{u}} \times \frac{\mathcal{BR}(B_{(s)}^{0} \to \mu^{+}\mu^{-})}{\mathcal{BR}(B^{\pm} \to J/\psi K^{\pm} \to \mu^{+}\mu^{-}K^{\pm})} \times \sum_{k} \frac{N_{J/\Psi K}^{k} \alpha_{k}}{R_{A\epsilon}^{k}}$$

Estimate expected $B_s \rightarrow \mu^+\mu^-$ signal yield

Assuming :

- SM branching ratios
- c-BDT and f-BDT was applied
- reference channel yield measured
- R^kA^ε was evaluated

then all the inputs are known :

Data Category (k)	$R^{k}_{A^{\epsilon}} = \epsilon_{J/\Psi K^{\pm}} / \epsilon_{\mu^{+}\mu^{-}}$	Ν ^k J/Ψ K±	α_k	f_s/f_d ($f_{u/d}=1$)	N _{Bs} → µ+µ-		
T1	$0.180 \pm 0.001 \pm 0.009$	46 860 ± 290 ± 280	7.23			27	
T2	$0.226 \pm 0.004 \pm 0.014$	5 200 ± 84 ± 100	7.28	0.240 ± 0.020		2.4	
T3	$0.189 \pm 0.005 \pm 0.022$	2 512 ± 91 ± 42	7.29	ATLAS		1.4	
2011	0.156 ± 0.002 ±0.009	95 900 ± 420 ± 1 100	1	arXiv:1507.08925		8.8	

Expect 41 B_s and 5 B_d events in the signal region

 N_{Bs} contains +4% correction on $B_s \rightarrow \mu + \mu$ - efficiency (due to diff. lifetime of B_s and $B_H^{(s)}$)

$B_{D/s} \rightarrow \mu + \mu - signal fit$

Unbinned extended maximum likelihood (ML) fit

m(μ+μ-) Data mass distribution fitted simultaneously in 3 c-BDT bins

w different S/B ratio to constrain the bkg.

Extract N_{Bs} and N_{Bd}

Each bin has 18% signal efficiency

Basic fit configuration

Background models: independent bkg yields and rel. fractions in each bin

Combinatorial bkg - 1st order Chebychev

slopes extracted independently in each bin

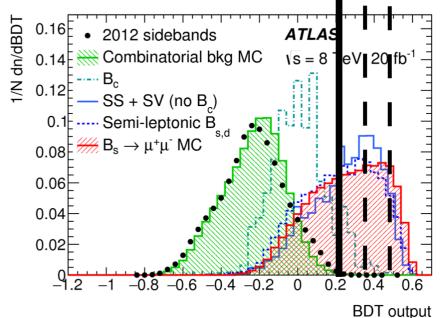
Same side/vertex bkg - Exponential

constrained to same shape in all bins

Peaking bkg - signal-like PDF w fixed shape (MC), yield 1.0 \pm 0.4, and rel. fractions Semi-leptonic bkg: no additional PDF (Gaussian tail in low-mass added only for systematics)

Signal:

2 Double Gaussian shapes from MC B_s / B_d sample signal fractions in each bin constrained to exp. equal signal efficiency 2 normalisation parameters N_{Bs} and N_{Bd}



B_{D/s} → μ + μ - **Fit Sensitivity & Systematics**

MC Toy Experiments

SM expected significance:

 $s_{Bs \rightarrow \mu+\mu^-} = 3.1 \sigma, s_{Bd \rightarrow \mu+\mu^-} = 0.2 \sigma$

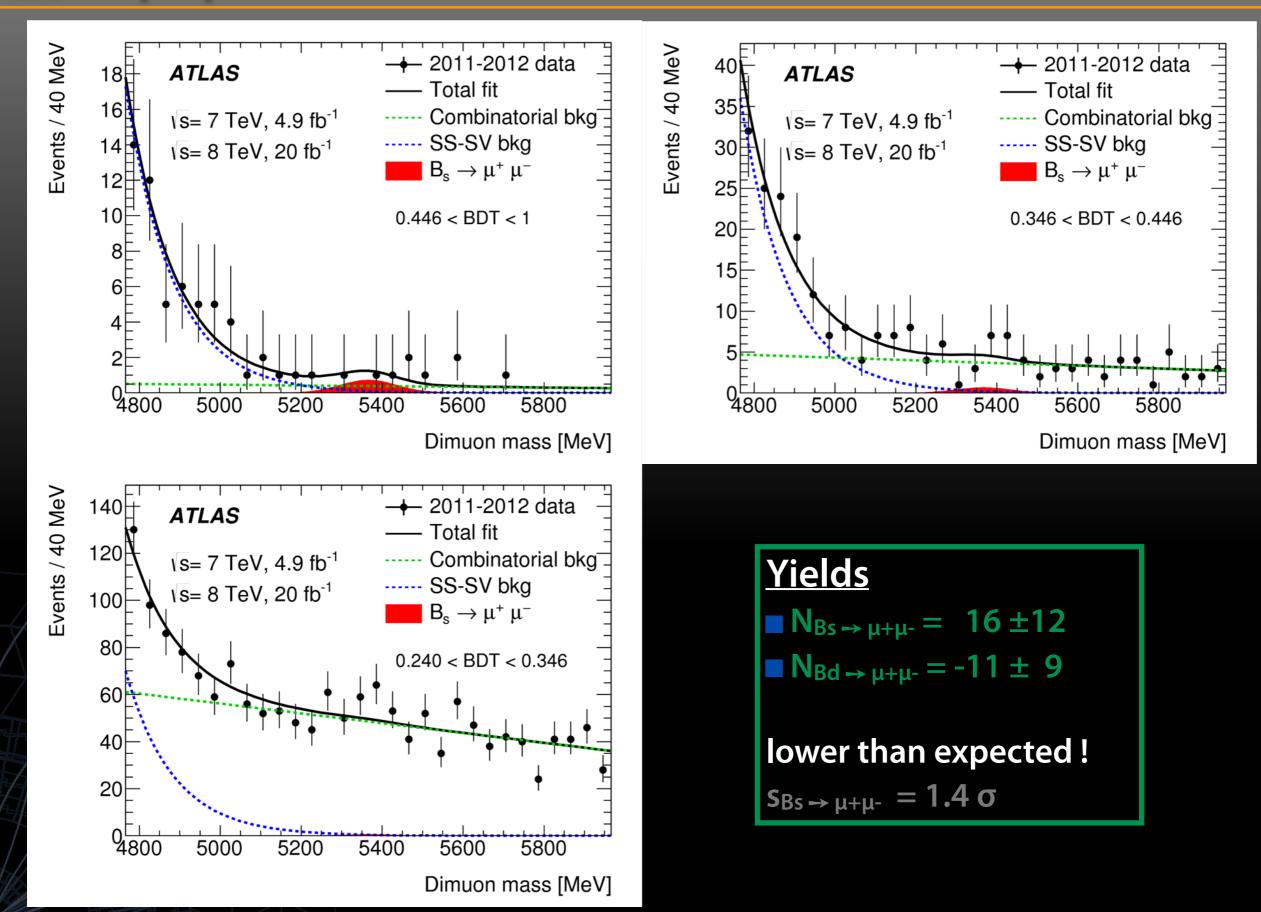
(Toys w mean $N_{Bs \rightarrow \mu+\mu} = 41$, $N_{Bd \rightarrow \mu+\mu} = 5$ & imposing $N_{Bs/d \rightarrow \mu+\mu} > 0$)

Dominant Systematics: Uncertainty in the relative efficiencies of the 3 c-BDT bins Alternative signal and background models

	$\mathcal{B}(B^0_s \to \mu^+ \mu^-)$	$\mathcal{B}(B^0 \to \mu^+ \mu^-)$
Scale uncertainties		
$\mathcal{B}(B^+ \to J/\psi K^+) \times \mathcal{B}(J/\psi \to \mu\mu)$ branching fractions $B^0_{(s)}/B^+$ production ratio B^+ yield and $B^0_{(s)}/B^+$ efficiency ratio Relative efficiency of continuum-BDT intervals Signal and background model	3.1% 8.3% 5.9% 9% 6%	$3.1\% \ 0 \ 5.9\% \ 9\% \ 0$
Total scale uncertainty	16%	11%
Offset uncertainties		
Signal and background model	0.2×10^{-9}	0.7×10^{-10}

Statistically limited measurement

$B_{D/s} \rightarrow \mu + \mu - FIT RESULT$



$B_{D/s} \rightarrow \mu + \mu$ - Branching Ratios

Branching ratio extraction

Central BR value: looking at minimum of the likelihood imposing $N_{Bs/d \rightarrow \mu+\mu-} > 0$ ($N_{Bd} = 0$; $N_{Bs} = 11$) $BR(B_d \rightarrow \mu+\mu) = 0$, $BR(B_s \rightarrow \mu+\mu) = 0.9 \times 10^{-9}$

Uncertainties:

Neyman construction of frequentist confidence belt (MC Toy experiments) include both statistical (dominant) and systematic uncertainties $\sigma_{syst.} = \pm 0.3 \times 10^{-9}$

Upper limits set @ 95 % CL using CLs technique

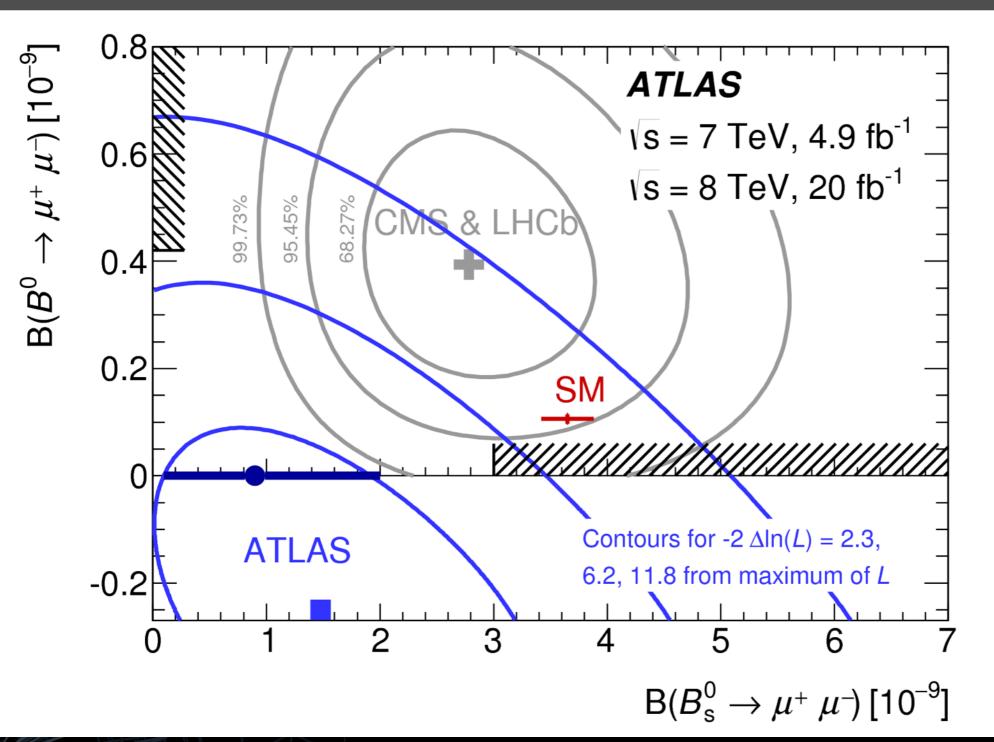
Results $BR(B_s) = (0.9 + 1.1 - 0.8) \times 10^{-9}$ $BR(B_s) < 3.0 \times 10^{-9}$ @ 95 % CL	$\frac{SM Expected}{BR_{SM}(B_s)} = (3.65 \pm 0.23) \times 10^{-9}$
BR(B _d) < 4.2 x 10 ⁻¹⁰ @ 95 % CL	$BR_{SM}(B_d) = (1.06 \pm 0.09) \times 10^{-10}$

ATLAS precision comparable with CMS and LHCb despite their (by construction) better m($\mu^+\mu^-$) resolution (~1.5x for CMS; ~3x for LHCb)

Result

Result compatible (lower than) SM p-value = 0.048 (2.0σ) (obtained by profile likelihood ratio)

Possibility of destructive interference of NP with the SM



CERN-EP-2016-064 arXiv:1604.04263 Submitted to EPJC

Summary

- Presented results of the ATLAS search for $B_{d/s} \rightarrow \mu^+ \mu^-$ decays on full Run I data from LHC
 - Significant improvement in analysis techniques
 - Results compatible with (lower than) CMS and LHCb
- 2.0 σ compatibility with the SM prediction (SM higher) w room for destructive interference of NP with SM

Planning analysis of Run2 (~100 fb⁻¹) data

BR(B_s $\rightarrow \mu^+\mu^-) = (0.9 + 1.1 - 0.8) \times 10^{-9}$

BR(B_s $\rightarrow \mu^+\mu^-$) < 3.0 x10⁻⁹ @ 95 % CL

 $BR(B_d \rightarrow \mu^+\mu^-) < 4.2 \times 10^{-10} @ 95 \% CL$

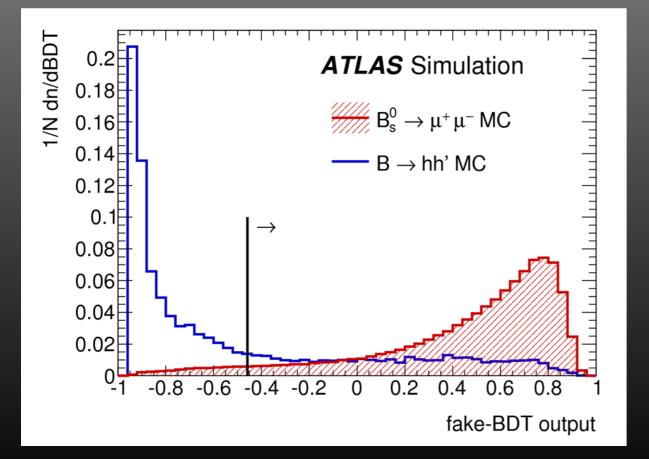
CERN-EP-2016-064, arXiv:1604.04263, Submitted to EPJC



BACKUP



F-BDT

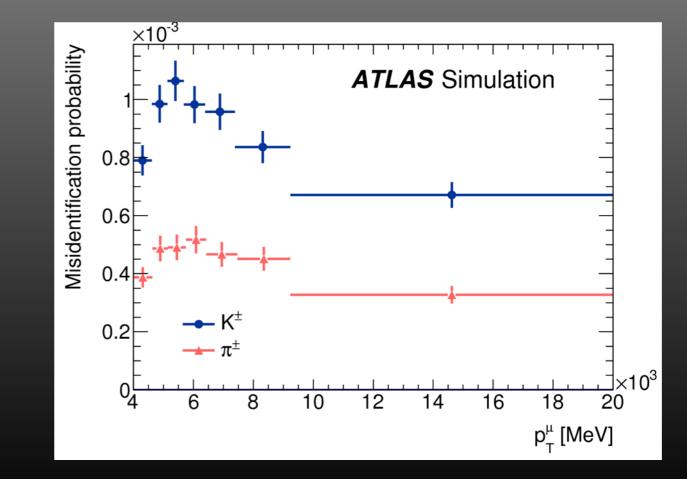


Critical aspect of the analysis

hadrons mis-identified as muons fake signal and can destroy sensitivity to NP

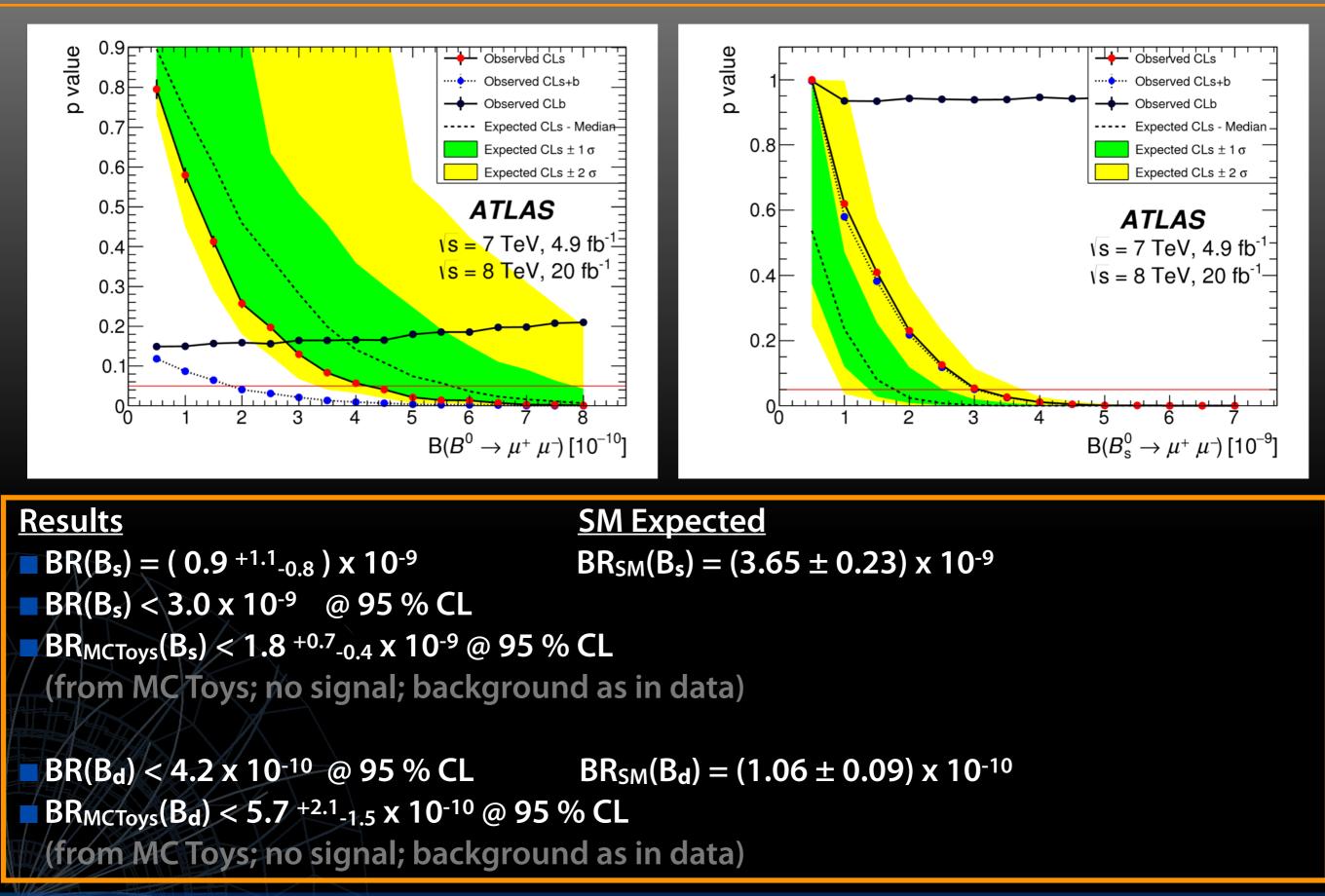
f-BDT reduces mis-identification probability to 0.09%/0.04%/<0.01% for K/π/p.

Punch through 3% (8%) level for K (π).



- 1. Absolute value of the track rapidity measured in the ID.
- 2. Ratio q/p (charge over momentum) measured in the MS.
- 3. Scattering curvature significance: maximum variation of the track curvature between adjacent layers of the ID.
- 4. χ^2 of the track reconstruction in the MS.
- 5. Number of hits used to reconstruct the track in the MS.
- 6. Ratio of the values of q/p measured in the ID and in the MS, corrected for the average energy loss in the calorimeter.
- 7. χ^2 of the match between the tracks reconstructed in the ID and MS.
- 8. Energy deposited in the calorimeters along the muon trajectory obtained by combining ID and MS tracks.

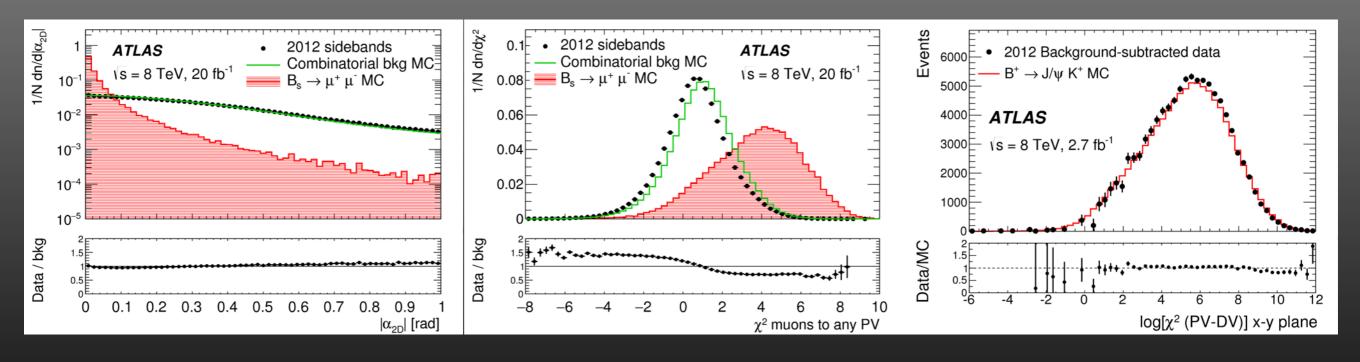
CLS LIMITS

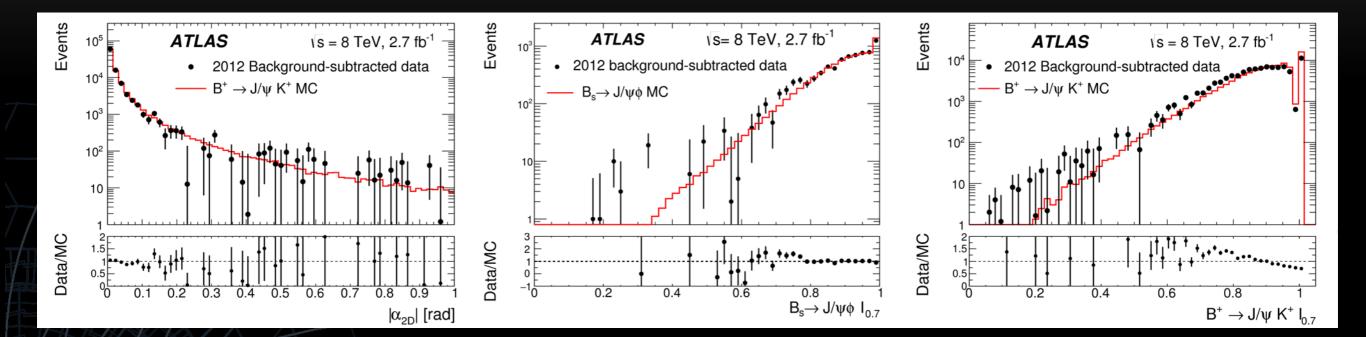


c-BDT

Variable	Description
p_{T}^B	Magnitude of the <i>B</i> candidate transverse momentum $\overrightarrow{p_{\mathrm{T}}}^{B}$.
$\chi^2_{\rm PV,DV} _{xy}$	Significance of the separation $\overrightarrow{\Delta x}$ between production (<i>i.e.</i> associated PV) and decay (DV) vertices in the transverse projection: $\overrightarrow{\Delta x}_{\mathrm{T}} \cdot \Sigma_{\overrightarrow{\Delta x}_{\mathrm{T}}}^{-1} \cdot \overrightarrow{\Delta x}_{\mathrm{T}}$, where $\Sigma_{\overrightarrow{\Delta x}_{\mathrm{T}}}$ is the covariance matrix.
ΔR	three-dimensional opening between \overrightarrow{p}^B and $\overrightarrow{\Delta x}$: $\sqrt{\alpha_{2D}^2 + \Delta \eta^2}$
$ \alpha_{2\mathrm{D}} $	Absolute value of the angle between $\overrightarrow{p_{\mathrm{T}}}^B$ and $\overrightarrow{\Delta x_{\mathrm{T}}}$ (transverse projection).
L_{xy}	Projection of $\overrightarrow{\Delta x}_{T}$ along the direction of $\overrightarrow{p}_{T}^{B}$: $(\overrightarrow{\Delta x}_{T} \cdot \overrightarrow{p}_{T}^{B})/ \overrightarrow{p}_{T}^{B} $.
$\mathrm{IP}_B^{\mathrm{3D}}$	three-dimensional impact parameter of the B candidate to the associated PV.
$\mathrm{DOCA}_{\mu\mu}$	Distance of closest approach (DOCA) of the two tracks forming the B candidate (three-dimensional).
$\Delta \phi_{\mu\mu}$	Difference in azimuthal angle between the momenta of the two tracks forming the B candidate.
$ d_0 ^{\max}$ -sig.	Significance of the larger absolute value of the impact parameters to the PV of the tracks forming the B candidate, in the transverse plane.
$ d_0 ^{\min}$ -sig.	Significance of the smaller absolute value of the impact parameters to the PV of the tracks forming the B candidate, in the transverse plane.
$P_{\mathrm{L}}^{\mathrm{min}}$	Value of the smaller projection of the momenta of the muon candidates along $\overrightarrow{p_{\mathrm{T}}}^B$.
$I_{0.7}$	Isolation variable defined as ratio of $ \vec{p_T}^B $ to the sum of $ \vec{p_T}^B $ and of the transverse momenta of all additional tracks contained within a cone of size $\Delta R < 0.7$ around the <i>B</i> direction. Only tracks with $p_T > 0.5$ GeV and matched to the same PV as the <i>B</i> candidate are included in the sum.
$\mathrm{DOCA}_{\mathrm{xtrk}}$	DOCA of the closest additional track to the decay vertex of the B candidate. Tracks matched to a PV different from the B candidate are excluded.
$N_{ m xtrk}^{ m close}$	Number of additional tracks compatible with the decay vertex (DV) of the <i>B</i> candidate with $\ln(\chi^2_{\text{xtrk,DV}}) < 1$. The tracks matched to a PV different from the <i>B</i> candidate are excluded.
$\chi^2_{\mu,\mathrm{xPV}}$	Minimum χ^2 for the compatibility of a muon in the <i>B</i> candidate with a PV different from the one associated with the <i>B</i> candidate.

Data/MC comparison





Blacksburg, Virginia | 22nd - 27th May 2016

Jaroslav Guenther 24

Miscellauneous

