Heavy Flavor Rare Decays from ATLAS

(Search for $B_s \rightarrow \mu^+\mu^-$)

Soeren Prell

Iowa State University

(on behalf of the

ATLAS Collaboration)

BEACH 2012 – 10th International Conference on Hyperons, Charm and Beauty Hadrons

July 23-28, 2012

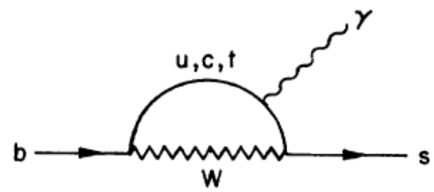
Wichita, Kansas

Outline

- *Motivation*
  - *FCNCs in B Decays*
  - *Theoretical predictions and previous results for $B_s \to \mu\mu$*
- *The ATLAS experiment*
  - *Detector components and performance*
- *Search for $B_s \to \mu\mu$*
  - *Analysis strategy*
  - *B reconstruction and event pre-selection*
  - *Backgrounds*
  - *Acceptances and Efficiencies*
  - *Reference mode yield determination*
  - *Branching ratio limit calculation*
  - *Result*
  - *Combined limit from ATLAS, CMS and LHCb*
- *Conclusions*

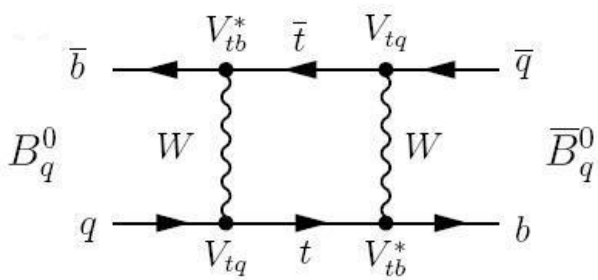
Flavor Changing Neutral Currents (FCNCs)

- FCNCs are highly suppressed in the SM

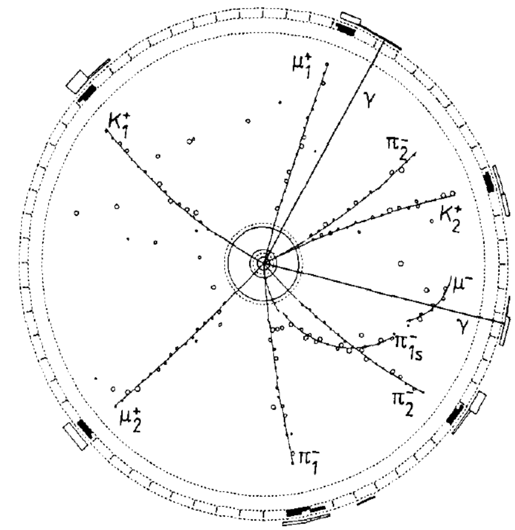


  - forbidden at tree level and FCNC loop diagrams are kinematically suppressed

- First FCNC in heavy quarks discovered in $B\overline{B}$ mixing by ARGUS in 1987



Fully-reconstructed $\Upsilon(4S) \rightarrow B^0B^0$ event



ARGUS, PLB 192, 245 (1987)

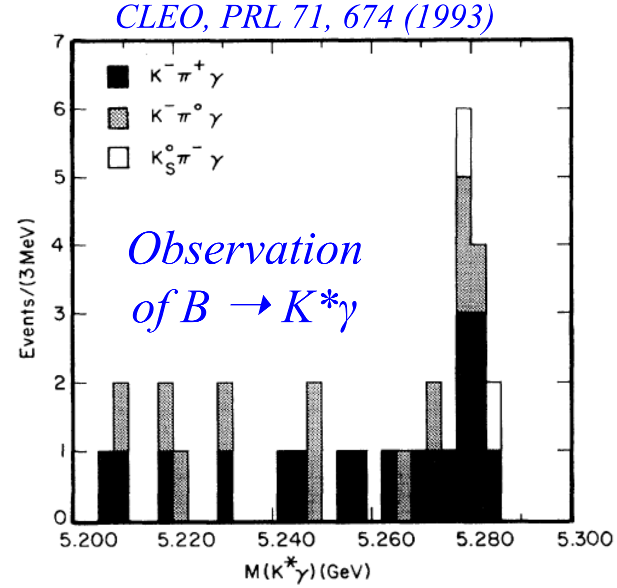
- First FCNC in exclusive radiative $b \rightarrow s\,\gamma$ decays observed by CLEO in 1993

$$\mathcal{B}(B^+ \rightarrow K^{*+}\gamma) = (4.21 \pm 0.18) \times 10^{-5}$$

$$\mathcal{B}(B^0 \rightarrow K^{*0}\gamma) = (4.33 \pm 0.15) \times 10^{-5}$$

PDG World Averages

CLEO, PRL 71, 674 (1993)



Flavor Changing Neutral Currents (FCNCs)

- *CKM-suppressed radiative $b \to d\,\gamma$ transitions discovered by BaBar and Belle*

$$\mathcal{B}(B \to \rho\gamma) = (1.39 \pm 0.25) \times 10^{-6}$$

PDG World Average

- *Semi-leptonic box diagrams rarest B decays so far measured*

$$\mathcal{B}(B \to K^{*} l^{+} l^{-}) = (1.08 \pm 0.11) \times 10^{-6}$$

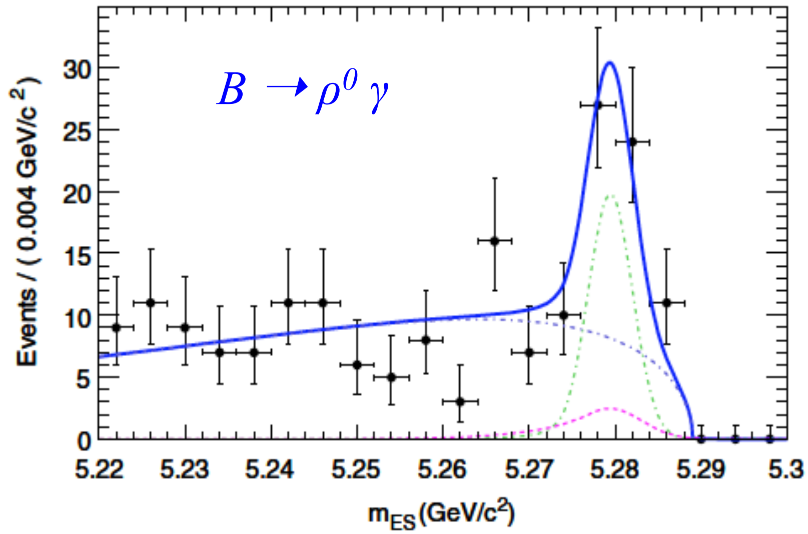
PDG World Average

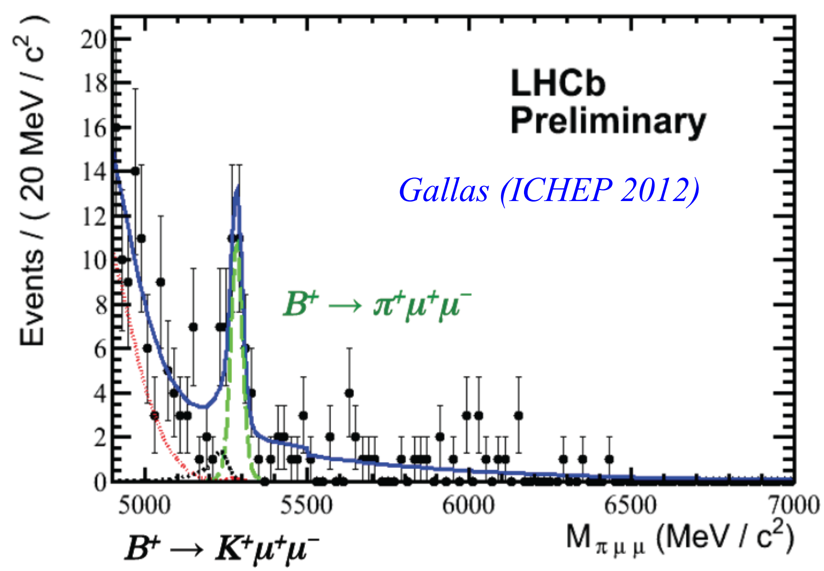
$$\mathcal{B}(B^{+} \to \pi^{+} l^{+} l^{-}) = (2.4 \pm 0.6 \pm 0.2) \times 10^{-8}$$

LHCb preliminary @ ICHEP 2012

- *All observed FCNC in neutral B decays have a "non-participating" spectator quark*
- *No observation of annihilation or exchange loop diagram in decay, yet*

BaBar, PRD 78, 112001 (2008)



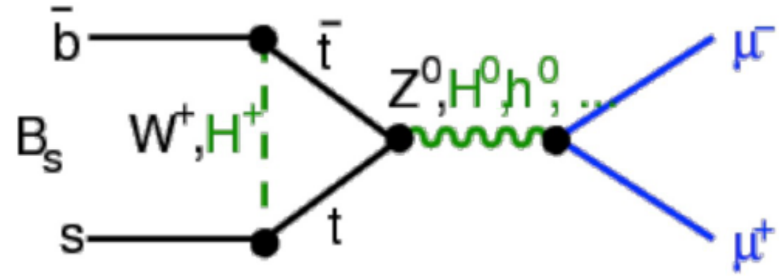


Motivation to search for $B_s \to \mu^+\mu^-$

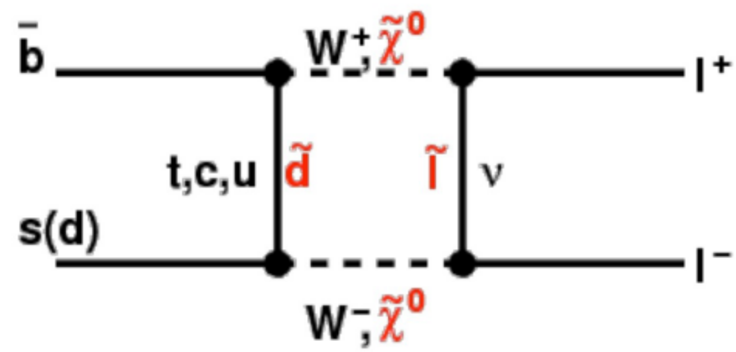
- *Standard model prediction*

$$\mathcal{B}(B_s \to \mu^+\mu^-) = (3.2 \pm 0.2) \times 10^{-9}$$

Buras et al., PLB 694, 402 (2011)

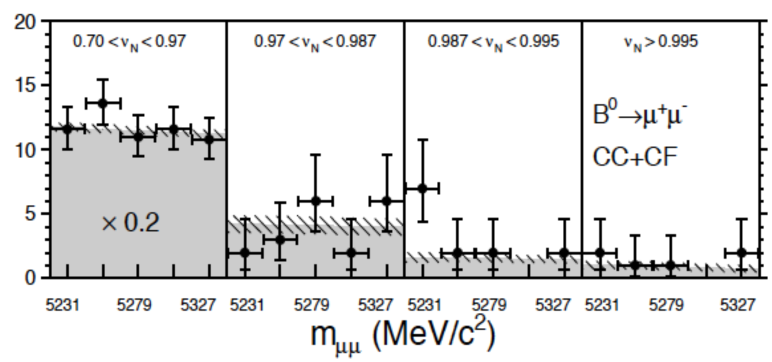


- *New Physics models*
  - *Virtual SM particles in loops could be replaced by heavy NP particles and thus significantly enhance the branching ratio*



- *Search for New Physics*
  - *Due to its small and precisely calculated branching ratio $B_s \to \mu^+\mu^-$ is a very sensitive mode for NP at very high masses*
  - *Search is complementary to direct searches at the energy frontier*
- *Best published limit on BR($B_s \to \mu^+\mu^-$) from non-LHC experiments by CDF*

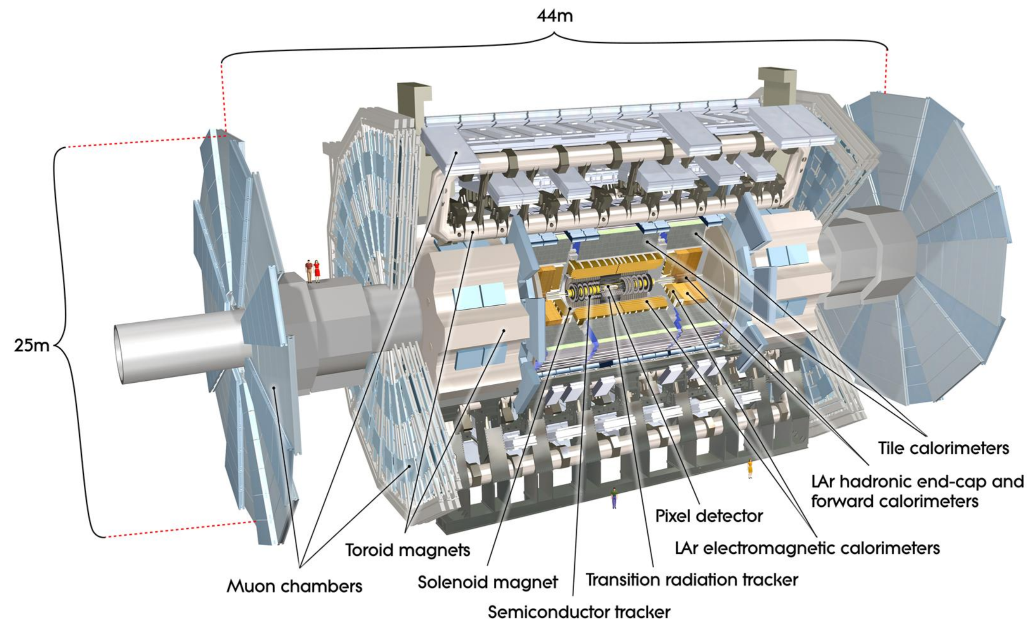
$$\mathcal{B}(B_s \to \mu^+\mu^-) < 4.0 \times 10^{-8} \text{ @ 95\% CL}$$

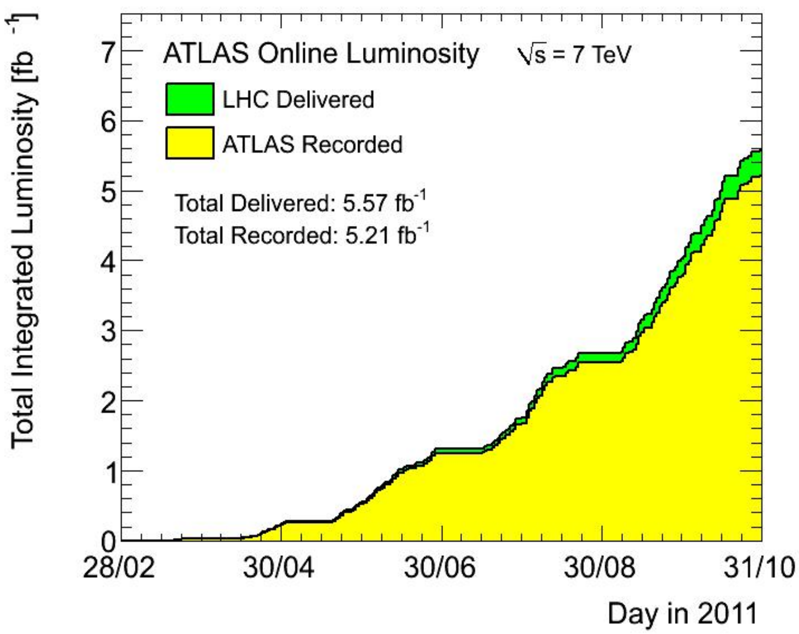


CDF, PRL 107, 191801 (2011)

The ATLAS Detector

- *Tracking*
  - *Pixel, (microstrip) Semiconductor Tracker (SCT) and Transition Radiation Tracker (TRT) in 2T solenoidal B field*
- *Muon System (MS)*
  - *Dedicated trigger and tracking chambers*
  - *0.5 – 2 T toroidal field*
- *Tracking performance*
  - *10 μm impact parameter resolution*
  - $\sigma(p_T)/p_T \sim 0.05\%\, p_T \oplus 1.5\%$
- *Data sample*
  - *5.2* fb^{-1} *recorded in 2011 (2.4* fb^{-1} *used in this analysis)*
  - *Instantaneous luminosity steadily increasing during 2011*





Analysis Strategy

- *Relative BR measurement*
  - *Partial cancelation of uncertainties (luminosity, cross-section, efficiencies,…)*
  - *Use reference decay $B^+ \to J/\psi K^+ \to \mu^+\mu^- K^+$*

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = \mathcal{B}(B^\pm \to J/\psi K^\pm \to \mu^+\mu^- K^\pm) \times \frac{f_u}{f_s} \times \frac{N_{\mu^+\mu^-}}{N_{J/\psi K^\pm}} \times \frac{A_{J/\psi K^\pm}}{A_{\mu^+\mu^-}} \times \frac{\varepsilon_{J/\psi K^\pm}}{\varepsilon_{\mu^+\mu^-}}$$

$$= N_{\mu^+\mu^-} \times \boxed{\text{SES}}$$

BR of reference channel and relative production rate

- *Taken from PDG and recent LHCb measurement*

Signal yields

- *Count $B_s \to \mu^+\mu^-$ events in signal region and "subtract" background estimated from mass sidebands*
- *Fit for $B^\pm \to \mu^+\mu^- K^\pm$ signal yield*

Acceptances and efficiencies

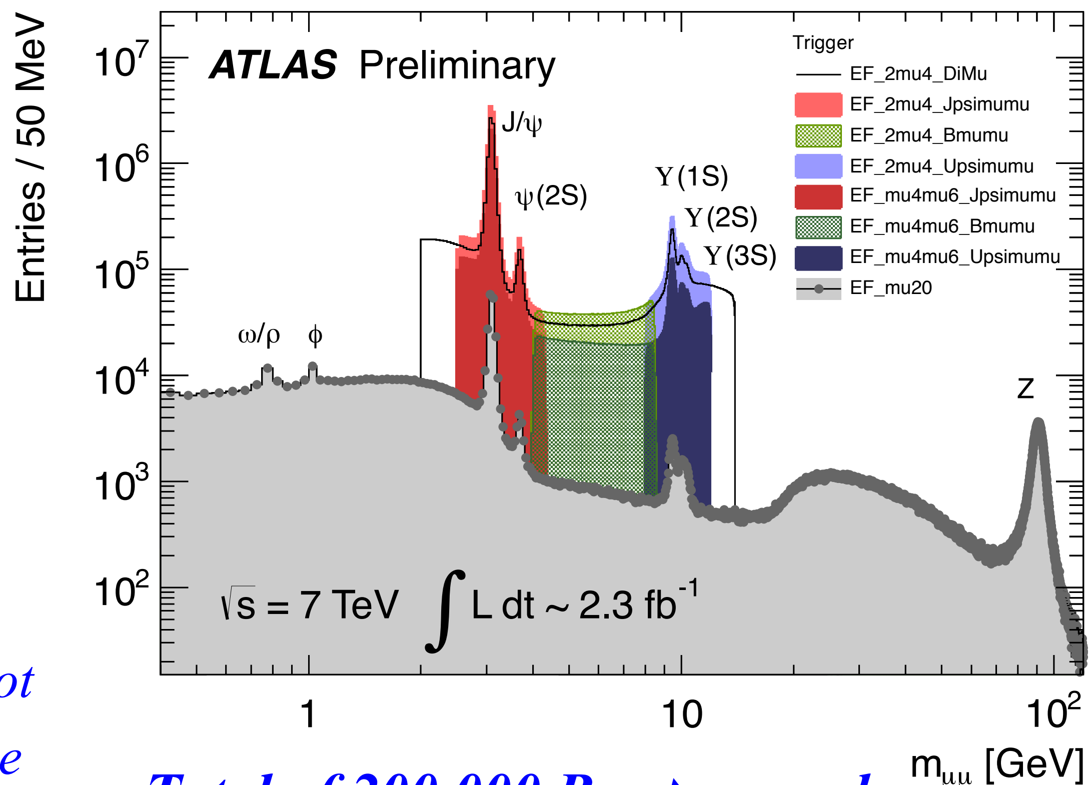
- *Determined with simulated events weighted to match data distributions*
- *Similar selection criteria for $B_s \to \mu^+\mu^-$ and reference channel*

Single event sensitivity SES

- *Corresponds to branching ratio that would yield one observed event in the data sample*

B Reconstruction and Event Pre-Selection

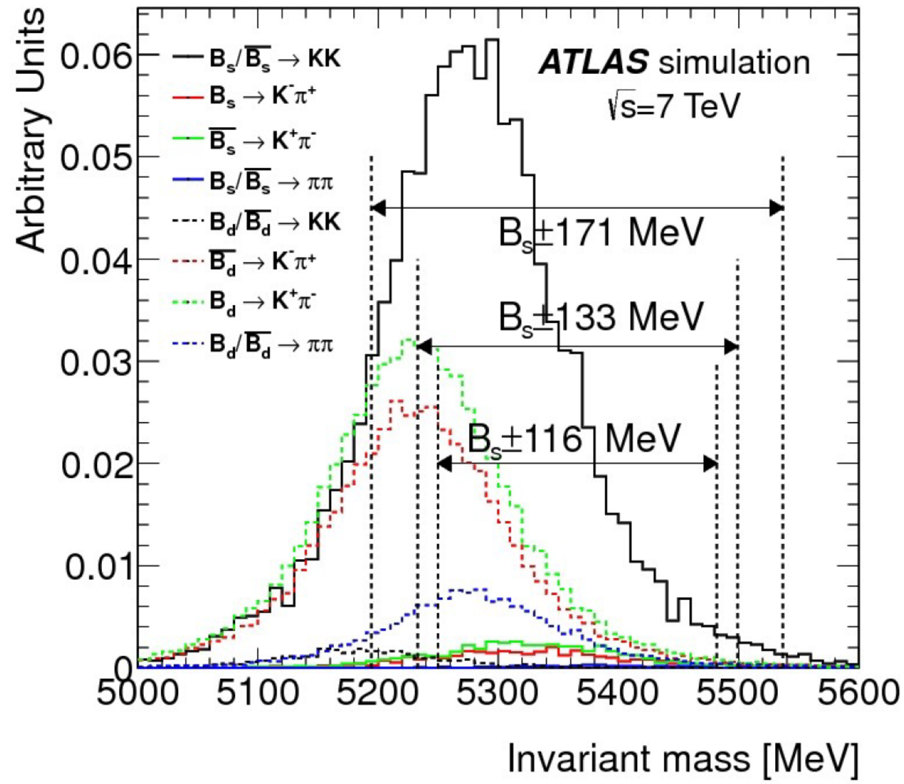
- *Trigger*
  - *Level 1 di-muon trigger, require p_T > 4 GeV for both muons*
- *Inner detector tracks*
  - *at least 1 pixel, 6 SCT and 8 TRT hits*
  - *$|\eta|$ <2.5 and p_T > 4 GeV (> 2.5 GeV) for muon (kaon) candidates*
  - *require muons to be matched to MS tracks*
- *B decay vertex*
  - *form B vertex from B daughter tracks*
  - *require vertex χ^2/DOF*
    - *< 2 for $B_s \to \mu\mu$ (85% efficient); < 6 for $B \to J/\psi K$ (99.5% efficient)*
- *B candidates*
  - *$p_T(B)$ > 8 GeV, $|\eta_B|$ < 2.5*
- *Primary vertex (PV)*
  - *determined from non-B candidate tracks and constrained to pp beam spot*
  - *if multiple PVs in an event, choose one closest in z to B decay vertex*



Total of 200,000 $B_s \to \mu\mu$ and 140,000 $B^\pm \to J/\psi K^\pm$ candidates

Backgrounds

- *Continuum background*
  - *originates from $b\bar{b} \to \mu^+ \mu^- X$*
  - *smooth variation with B mass*
    - *background yield in signal region can be interpolated from events in B mass sidebands*
- *Resonant background*
  - *B decays in which one or two hadrons are misidentified as muons*
  - *Probability for a charged K (π) to be mis-identified as muon is 0.2 (0.4) %*
  - *Main backgrounds*
    - *single-fake events (eg. $\bar{B}_s \to K^- \mu^+ \nu$)*
    - *double-fake events (eg. $B_s \to K^+ \pi^-$, $B \to \pi^+\pi^-$, $B \to K^+\pi^-$)*
  - *Estimated with simulated events to be 0.24 events total (after final selection)*

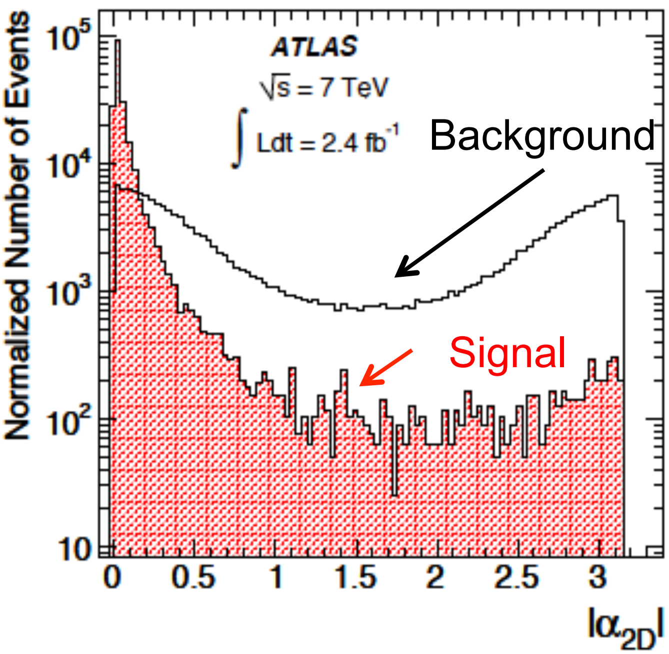


Channel	Signal Region	Sideband Regions
$B_s^0 \to \mu^+\mu^-$	[5066,5666] MeV	[4766,5066] MeV [5666,5966] MeV
$B^\pm \to J/\psi K^\pm$	[5180,5380] MeV	[4930,5130] MeV [5430,5630] MeV

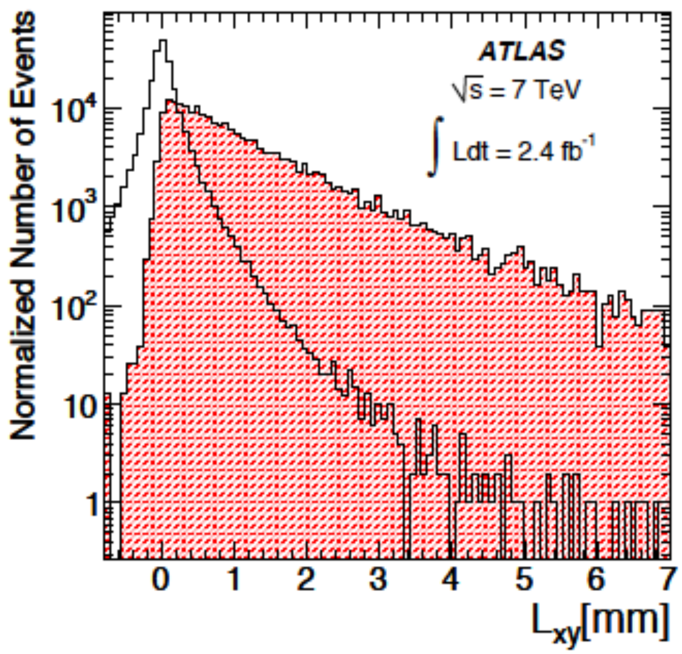
Blinded until analysis finalized

Background-discriminating Variables

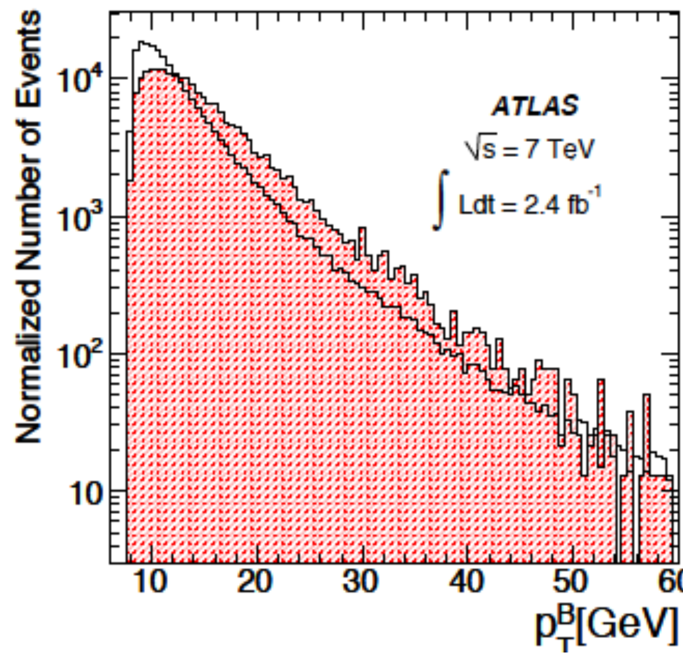
- *14 variables combined into Boosted Decision Tree (BDT)*
  - *discriminate against backgrounds from prompt tracks, muons from different b decays, secondary vertices with additional particles, and non-B processes*
  - *most of the variables exploit precisely measured displaced secondary vertex originating from the long B_s lifetime*



*α_{2D} = angle between **Δx** projection in transverse plane and B transverse momentum*



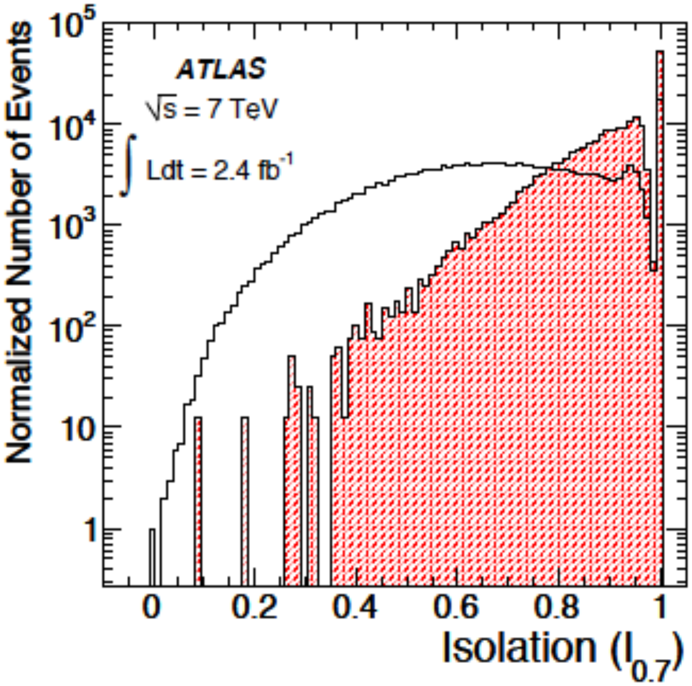
*L_{xy} = projection of **Δx** on B transverse momentum*

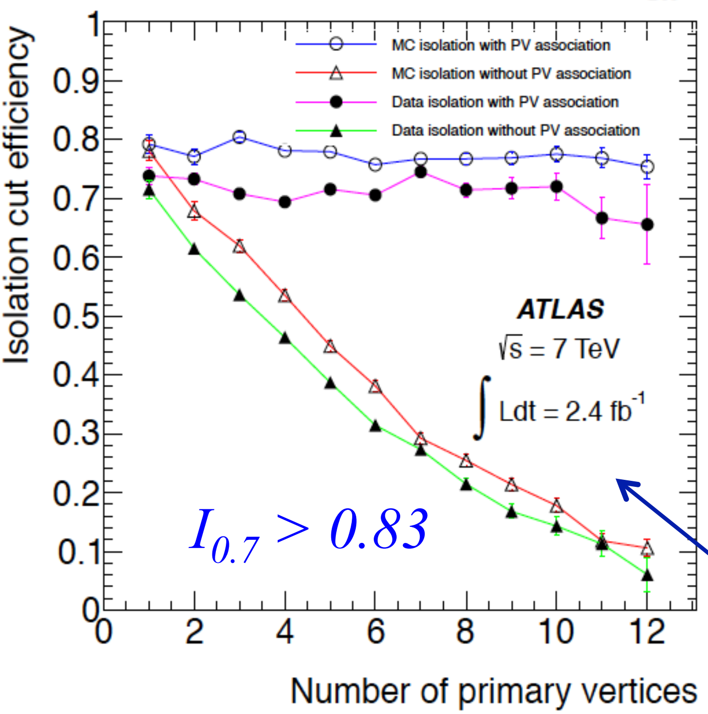


p_T^B = B transverse momentum

$\Delta\vec{x} = \vec{x}_{SV} - \vec{x}_{PV}$ = vector between *B* decay vertex and primary vertex

Sensitivity of BDT to Pile-up





- *Pile-up = multiple pp interactions per bunch crossing*
  - *Average # PVs in the data sample is 5*
- *Isolation $I_{0.7}$ is the variable in the BDT that is most sensitive to pile-up*

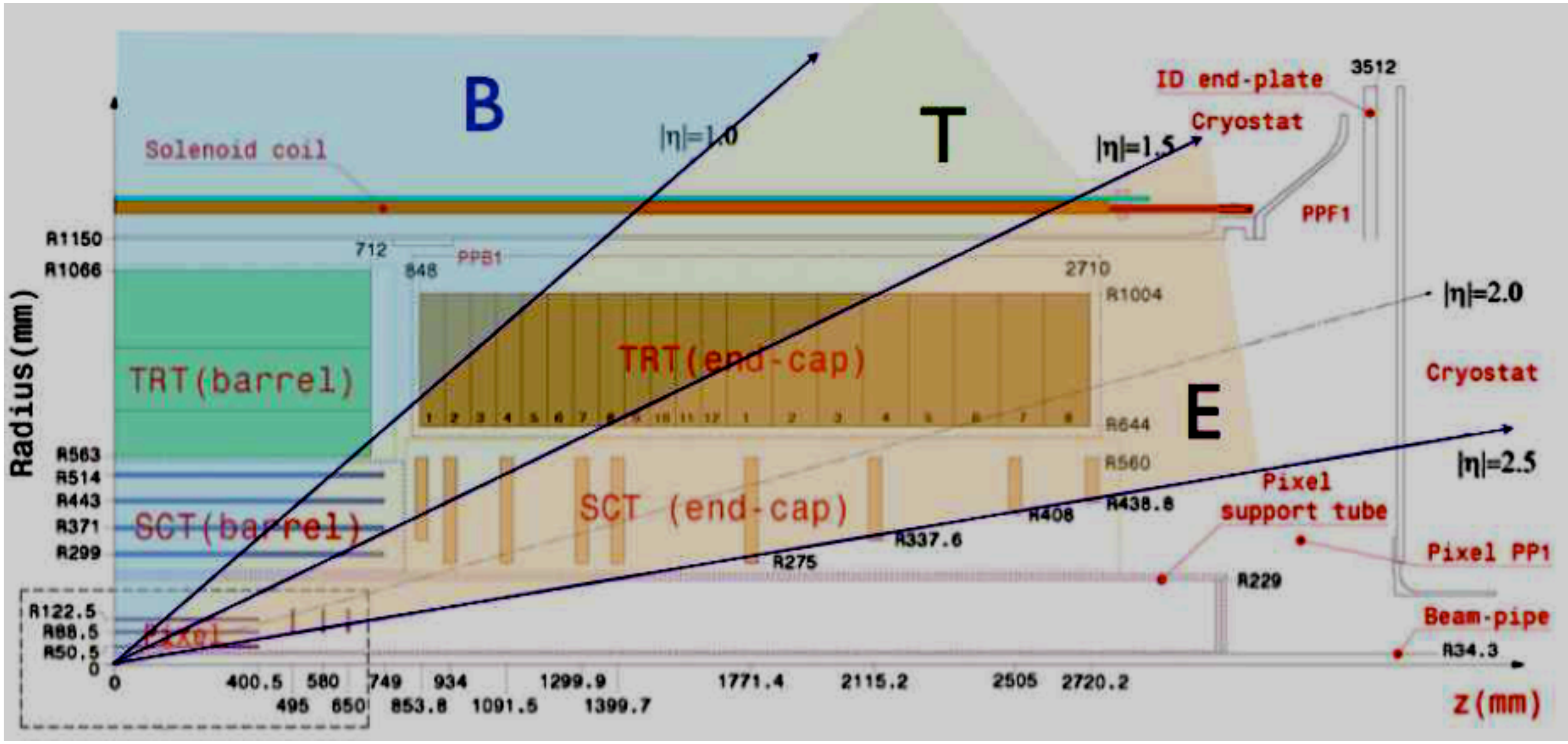
$$I_{0.7} = \frac{\left|\vec{p}_T^{\,B}\right|}{\left|\vec{p}_T^{\,B}\right| + \sum_{\Delta R<0.7} \left|\vec{p}_T^{\,track}\right|} \qquad \text{with } \Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$$

  - *$I_{0.7}$ strongly depends on # PVs, if all tracks (except the B daughters) are included in the calculation*
  - *Limiting the calculation to tracks coming from the primary vertex associated with the B candidate almost completely removes the dependence*

Data = sideband-subtracted $B \to J/\psi\, K$ events
MC = simulated $B \to J/\psi\, K$ events

Variation in B Mass Resolution

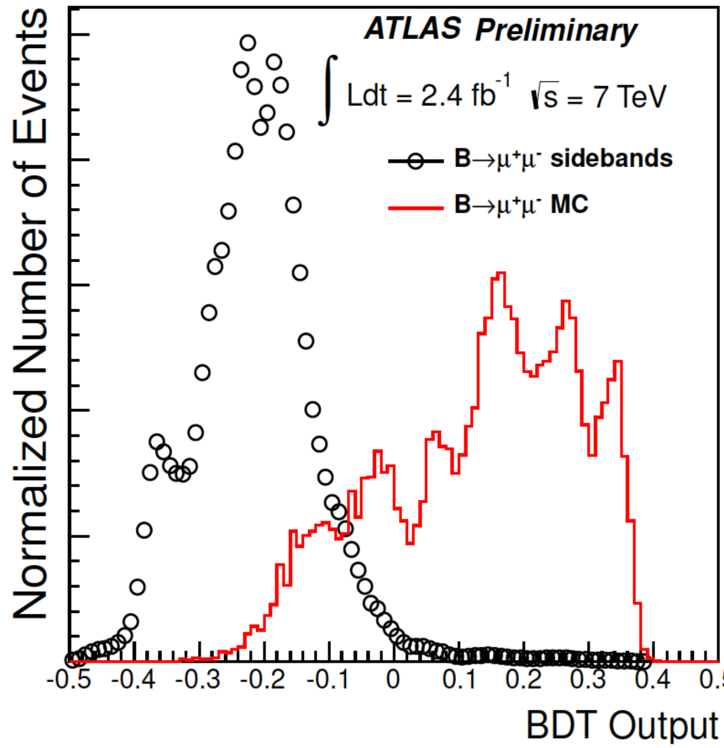
- Split signal candidates by $|\eta|$ region* and optimize subsamples separately
  - Three regions Barrel (B), Transition (T) and Endcap (E)
- B mass resolution becomes worse with increasing $|\eta|$
  - more multiple scattering and smaller magnetic field integral at large $|\eta|$
  - B mass resolutions: 60 MeV (B), 80 MeV (T), and 110 MeV (E)
  - Most B candidates are in Barrel region (51% (B), 24% (T), 25% (E))



* The maximum $|\eta|$ of the 2 muon tracks determines the subsample.

BDT Performance and Optimization of Signal Selection

- *Boosted Decision Tree (BDT)*
  - *Multivariate classifier chosen to combine 14 discriminating variables for improved signal-background separation*
  - *Trained on $B_s \to \mu\mu$ signal MC and mass sidebands in (50% of) data sample*
  - *Optimal cut between 0.23 and 0.27 depending on the $|\eta|$ subsample*



- *Maximize $\mathcal{P}$ for optimal performance of a 95% CL frequentist limit (a =2) in a counting analysis [Punzi, arXiv:physics/0308063]*
  - *Optimize BDT output cut and invariant mass search region window*

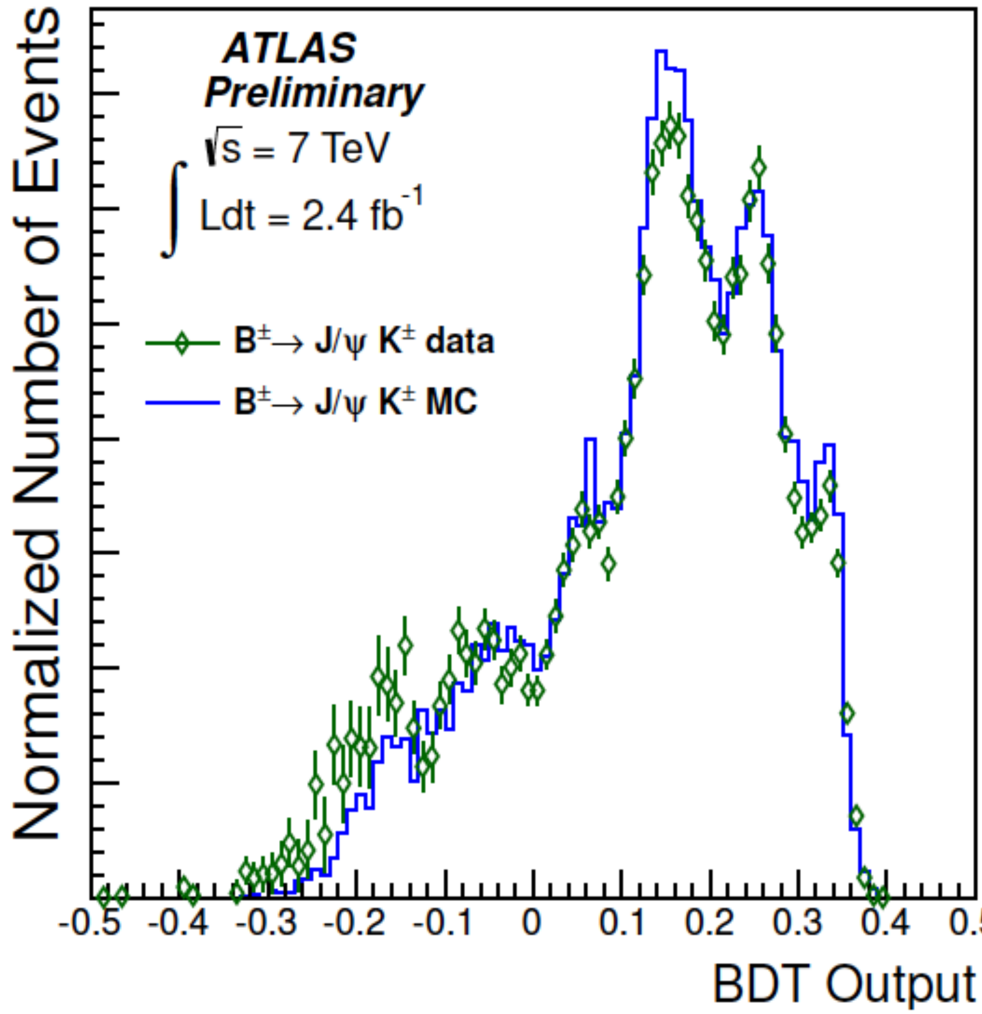
$$\mathcal{P} = \frac{\varepsilon_{Sig}}{\frac{a}{2} + \sqrt{N_{Bgd}}}$$

$|\eta_{max}|$ Range	0-1.0	1.0-1.5	1.5-2.5
invariant mass window [MeV]	±116	±133	±171
BDT output threshold	0.234	0.245	0.270

Acceptance and Efficiency Ratios

$$R_{A\varepsilon} = \frac{A_{J/\psi K^\pm}}{A_{\mu^+\mu^-}} \times \frac{\varepsilon_{J/\psi K^\pm}}{\varepsilon_{\mu^+\mu^-}}$$

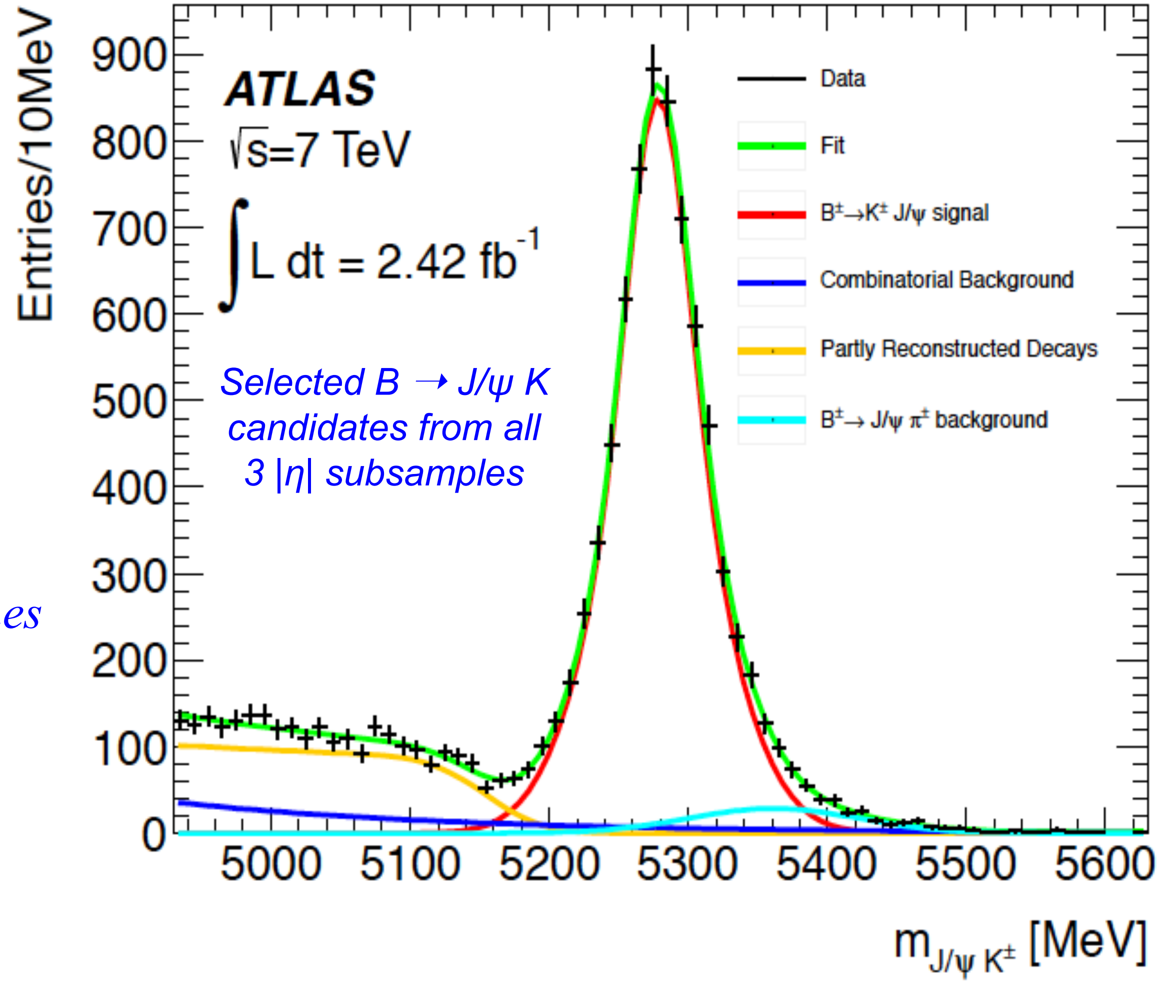
- *Dominant systmatic uncertainty in $R_{A\varepsilon}$ is due data – MC discrepancies*
  - *Re-weight events wrt most sensitive variables in BDT*
  - *Due to large correlations between decay-length sensitive variables, L_{xy} correction effectively removes differences in other variables*
  - *Differences in p_T and η of B daughters, # PV and pixel detector occupancy are accounted for in systematics*
- *Systematic change in $R_{A\varepsilon}$ is 0.6%*
  - *changes $A \times \varepsilon$ for individual samples are 10-20%, but highly correlated between $B_s \to \mu\mu$ and reference mode $B \to J/\psi K$*



$\lvert\eta_{max}\rvert$ Range	$R^i_{A\epsilon}$	Δ % Stat.	Δ % Syst.
0-1.0	0.274	3.1	3.1
1.0-1.5	0.202	4.8	5.5
1.5-2.5	0.143	5.3	5.9

$B^{\pm} \to J/\psi\ K^{\pm}$ Yield

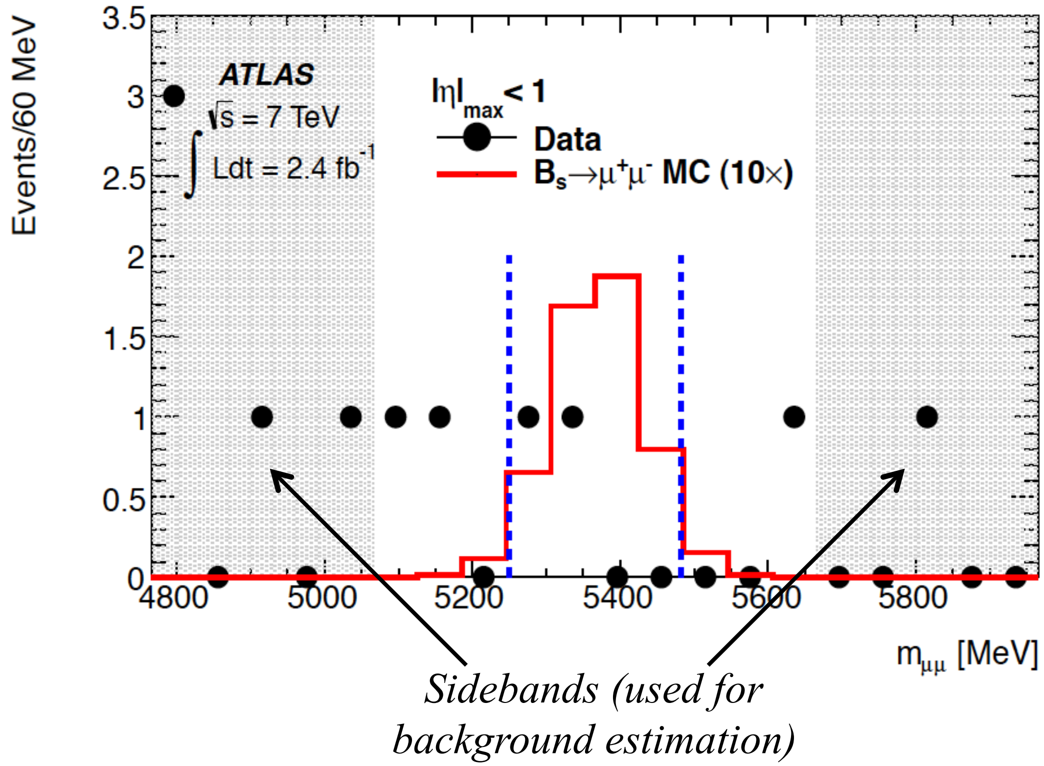
- *Extract $B^{\pm} \to J/\psi\ K^{\pm}$ yield with binned likelihood fit to B mass distribution separately for each $|\eta|$ subsample*
  - *Signal PDF: double Gaussian with common mean*
  - *Background PDFs:*
    - *Exponential (combinatorial background) + Exponential times complementary error function (partially reconstructed B decays) + Gaussian ($B \to J/\psi\ \pi$)*

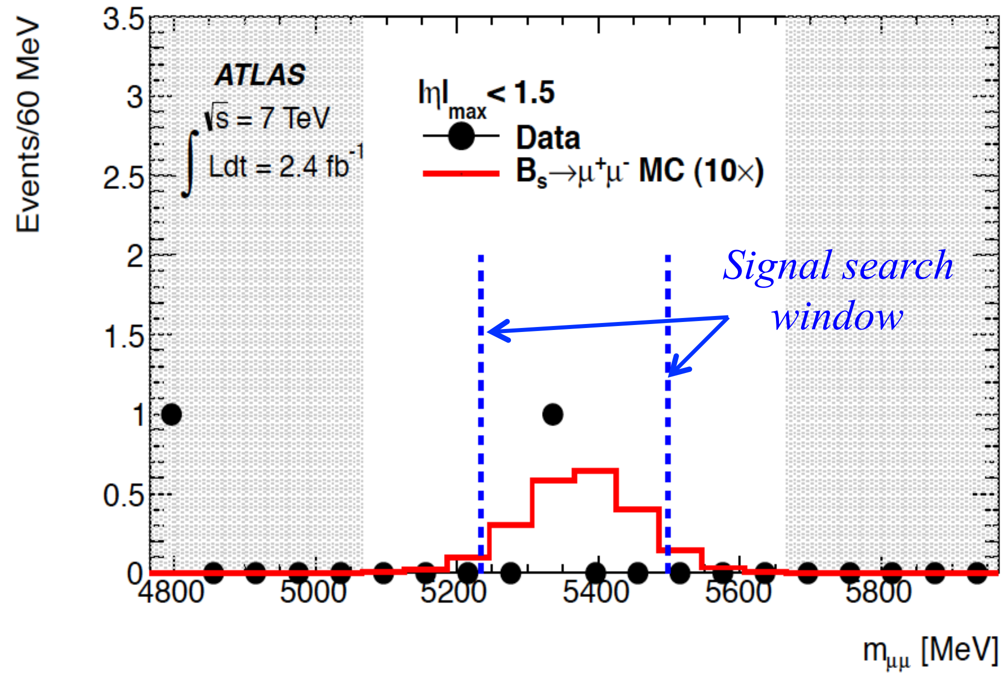


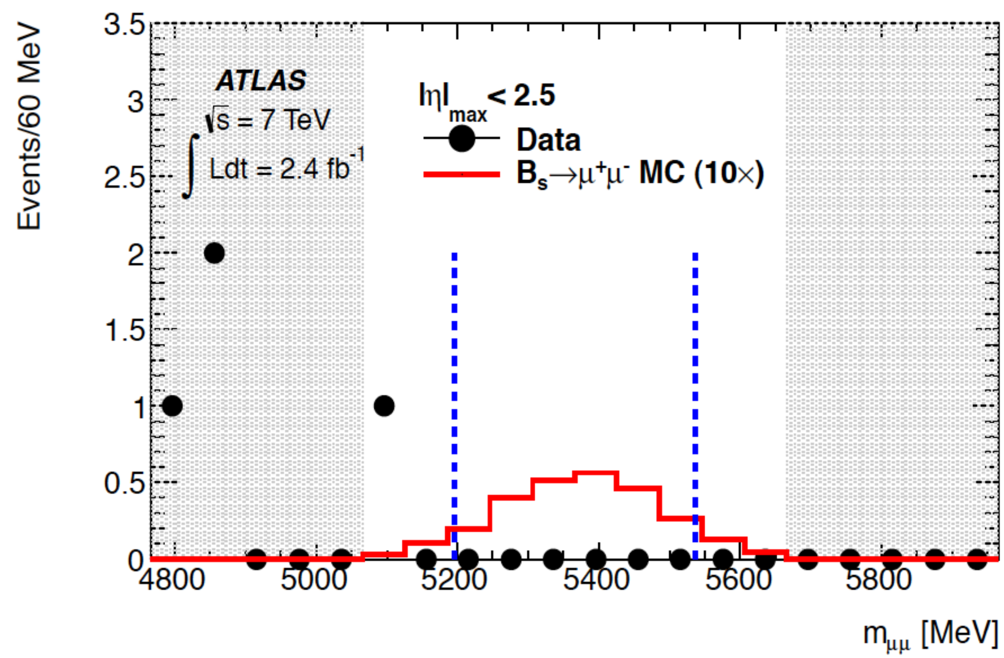
- *Systematic error estimation*
  - *vary bin size*
  - *use different signal and background PDFs*
  - *use unbinned likelihood fit with event-by-event mass resolution*

$\lvert\eta_{max}\rvert$ Range	0-1.0	1.0-1.5	1.5-2.5
$B^{\pm} \to J/\psi K^{\pm} \to \mu^{+}\mu^{-}K^{\pm}$	4300	1410	1130
statistical uncertainty	±1.6%	±2.8%	±3.0%
systematic uncertainty	±2.9%	±7.4%	±14.1%

Opening the Signal Box







- *Observe 2, 1 and 0 events in signal regions consistent with background-only expectations*
  - *MC signal distributions correspond to roughly ×10 SM prediction (ie. $BR(B_s \rightarrow \mu\mu) = 3.5 \times 10^{-8}$)*

Putting it all together…

$\lvert\eta_{max}\rvert$ Range	0-1.0	1.0-1.5	1.5-2.5
$B^\pm \to J/\psi K^\pm \to \mu^+\mu^- K^\pm$	4300	1410	1130
statistical uncertainty	±1.6%	±2.8%	±3.0%
systematic uncertainty	±2.9%	±7.4%	±14.1%

$\lvert\eta_{max}\rvert$ Range	$R^i_{A\epsilon}$	Δ % Stat.	Δ % Syst.
0-1.0	0.274	3.1	3.1
1.0-1.5	0.202	4.8	5.5
1.5-2.5	0.143	5.3	5.9

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = \mathcal{B}(B^\pm \to J/\psi K^\pm \to \mu^+\mu^- K^\pm) \times \frac{f_u}{f_s} \times \frac{N_{\mu^+\mu^-}}{N_{J/\psi K^\pm}} \times \frac{A_{J/\psi K^\pm}}{A_{\mu^+\mu^-}} \times \frac{\varepsilon_{J/\psi K^\pm}}{\varepsilon_{\mu^+\mu^-}}$$

$$= N_{\mu^+\mu^-} \times \mathrm{SES}$$

$\lvert\eta_{max}\rvert$ Range	0-1.0	1.0-1.5	1.5-2.5
SES= $(\epsilon\epsilon_i)^{-1}$ $[10^{-8}]$	0.71	1.6	1.4

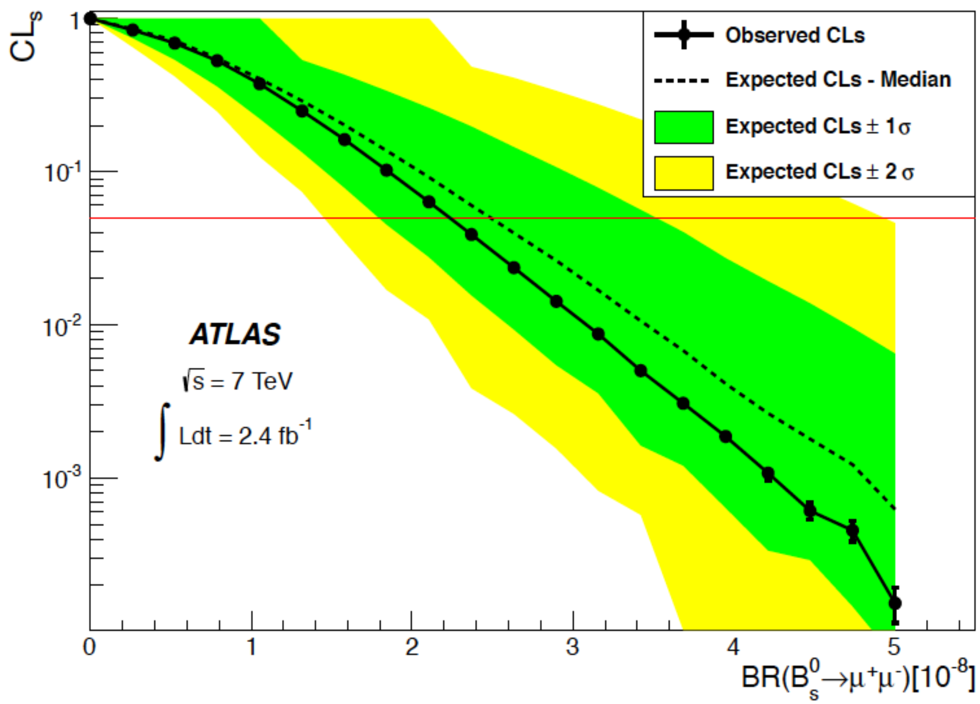
$1/(4.45 \pm 0.38) \times 10^3$

[PDG + LHCb]

BR Limit Extraction

- *Use CL_s method to extract $B_s \to \mu\mu$ branching fraction*
  - *expected 95% CL limit $2.3^{+1.0}_{-0.5} \times 10^{-8}$ (determined before unblinding)*

$$\mathcal{L} = \text{Gauss}(\epsilon_{obs}|\epsilon, \sigma_\epsilon) \times \text{Gauss}(\text{R}^{bkg}_{obs}|\text{R}^{bkg}, \sigma_{\text{R}^{bkg}}) \times \prod_{i=1}^{N_{bin}} \text{Poisson}(N_i^{obs}|\epsilon\, \epsilon_i BR + N_i^{\text{bkg}} + N_i^{B\to hh}) \times \text{Poisson}(N^{\text{bkg}}_{obs,i}|\text{R}^{bkg}\, \text{R}_i^{bkg} N_i^{\text{bkg}}) \times \text{Gauss}(\epsilon_i^{obs}|\epsilon_i, \sigma_{\epsilon_i}) .$$



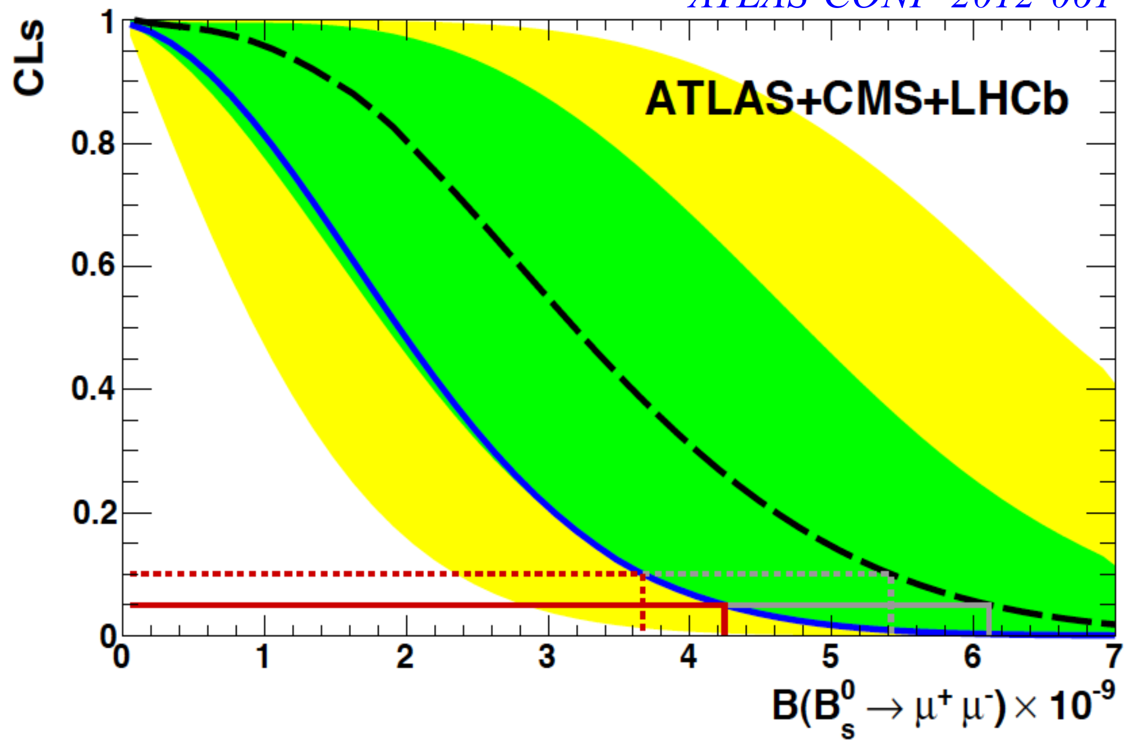
$$B(B_s \to \mu^+\mu^-) < 2.2 \times 10^{-8} \quad \text{at } 90\% \text{ CL}$$

p-value for background-only (background + SM) hypothesis is 44% (35%)

Combined ATLAS, CMS and LHCb Limit

- *Limit is compatible with background +SM signal within 1 σ ($1 - CL_{s+b} = 84\%$)*
- *p-value for background-only hypothesis ($1 - CL_b$): 5%*
- *Limit is getting close to SM prediction of $(3.2 \pm 0.2) \times 10^{-9}$*

ATLAS-CONF-2012-061



Mode	Limit	ATLAS	CMS	LHCb 2010	LHCb 2011	Combined
$B_s^0 \to \mu^+\mu^-$ (10^{-9})	Bkg Only	23	(3.6)	65	3.4	2.3
	Bkg+SM		8.4		7.2	6.1
	Obs	22	7.7 (7.2)	56	4.5	4.2

$$B(B_s \to \mu^+\mu^-) < 4.2 \times 10^{-9} \quad \text{at 95\% CL}$$

Conclusions

- *ATLAS's rich heavy flavor physics program has expanded to rare decays*

- *Other ATLAS talks on b physics at this conference*
  - *C. Hawkes: Heavy flavor cross-sections*
  - *E. Kneringer: CP violation*
  - *N. Panikashvili: Mass and lifetime of Λ_b*
  - *R. Wang: Y and χ_b production*

- *Presented branching ratio limit from ATLAS on $B_s \rightarrow \mu\mu$ which is now published in Phys.Lett. B713 (2012) 387 and an ATLAS-CMS-LHCb combined limit (ATLAS-CONF-2012-061)*
  - *95% CL limit is only 30% above Standard Model prediction*

- *More heavy flavor physics measurements from ATLAS will be coming soon*
  - *In particular an improved measurement of $B_s \rightarrow \mu\mu$ is in progress using more data and improved analysis techniques*