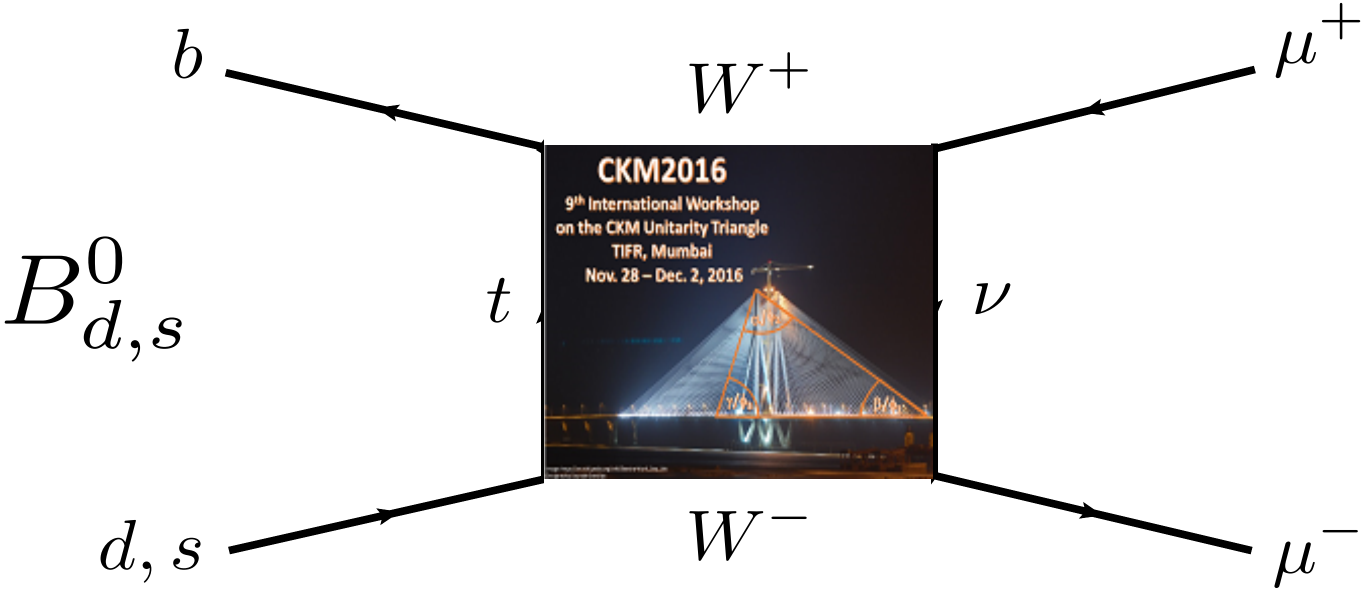
Rare B decays at ATLAS



Adrian Bevan
(on the behalf of the ATLAS Collaboration)

Queen Mary
University of London

Outline

- Physics motivation: $B_{s,d} \to \mu^+\mu^-$
- ATLAS experiment and data samples
- Analysis:
  - Overview
  - Trigger
  - Event Selection
  - Backgrounds (and suppression)
  - Signal Fit Result
  - Normalisation mode: $B^\pm \to J/\psi K^\pm$
  - Branching Fraction Calculation
  - Systematic Errors
  - Result
- Summary

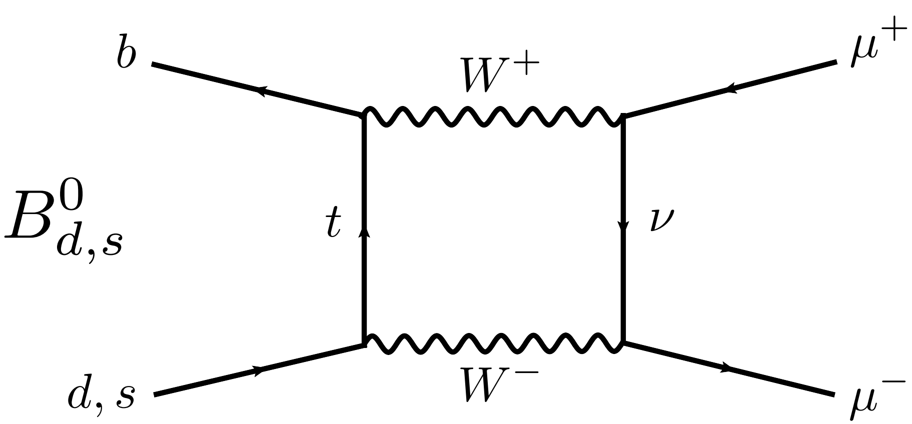
See the following paper for more details:

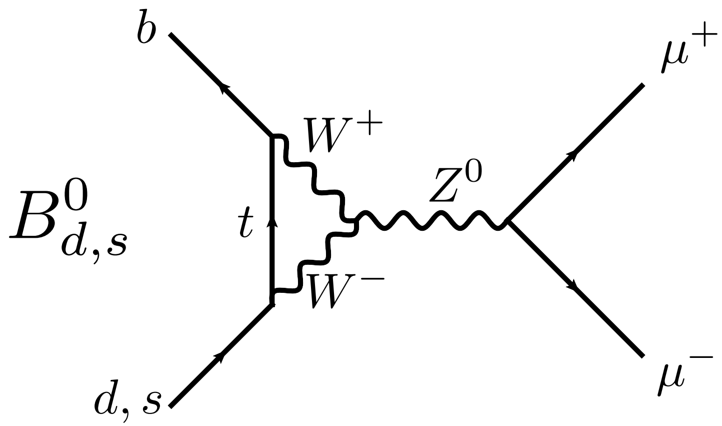
CERN-EP-2016-064
Eur. Phys. J. C76 (2016) no.9, 513

PHYSICS MOTIVATION: $B_{s,d} \to \mu^+\mu^-$

$B_{s,d} \to \mu^+\mu^-$

- Rare decays: excellent probes for new physics (NP).
- Suppressed standard model (SM) decays like FCNCs can exhibit large effects beating against small NP amplitudes.





- SM rate is theoretically well known:

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (3.65 \pm 0.23) \times 10^{-9}$$

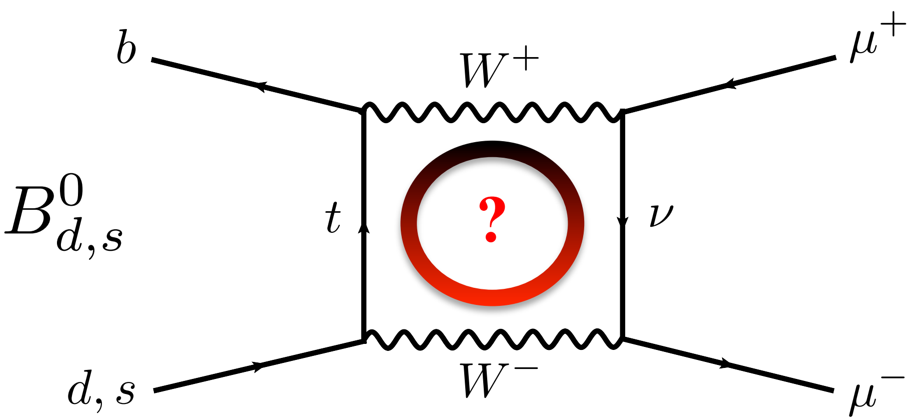
$$\mathcal{B}(B_d^0 \to \mu^+\mu^-) = (1.06 \pm 0.09) \times 10^{-10}$$

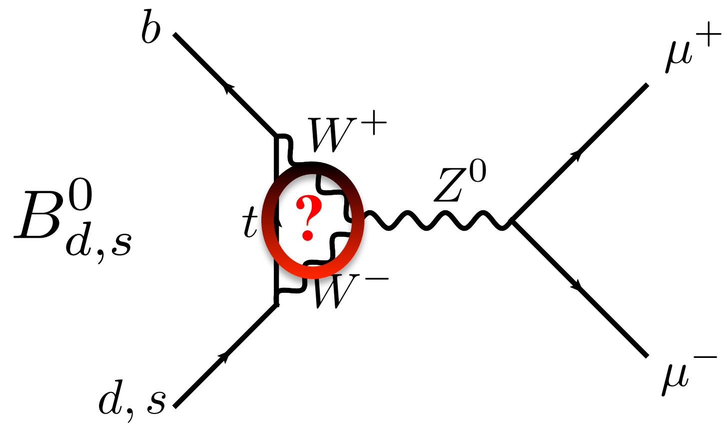
Bobeth et al., PRL **112** (2014) 101801 (arXiv:1311.0903)

- NP can enhance or suppress this rate.

$B_{s,d} \to \mu^+\mu^-$

- Rare decays: excellent probes for new physics (NP).
- Suppressed standard model (SM) decays like FCNCs can exhibit large effects beating against small NP amplitudes.





- SM rate is theoretically well known:

$$\mathcal{B}(B^0_s \to \mu^+\mu^-) = (3.65 \pm 0.23) \times 10^{-9}$$

$$\mathcal{B}(B^0_d \to \mu^+\mu^-) = (1.06 \pm 0.09) \times 10^{-10}$$

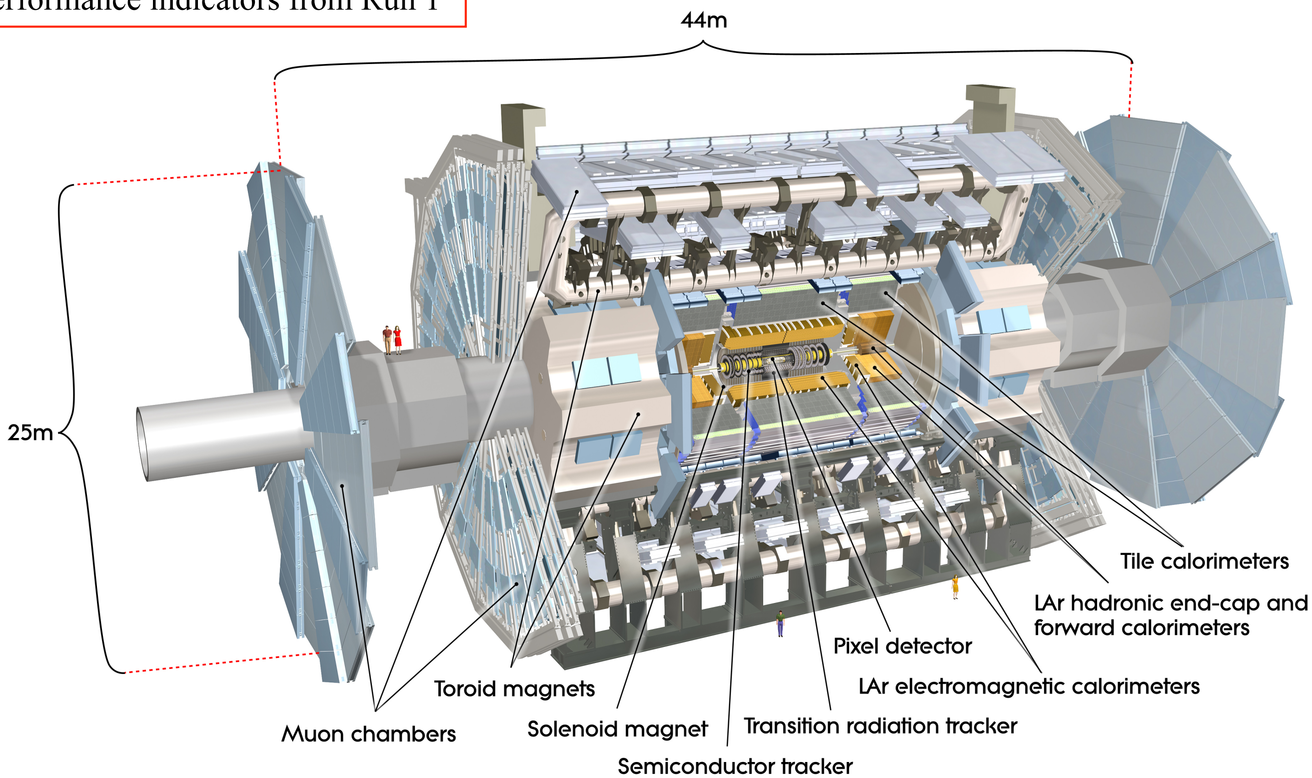
Bobeth et al., PRL **112** (2014) 101801 (arXiv:1311.0903)

- NP can enhance or suppress this rate.

ATLAS EXPERIMENT AND DATA SAMPLES

The ATLAS Detector

Performance indicators from Run 1



Trigger:
- 3-levels
- Reduces rate from 40MHz to 400Hz

Hadronic Calorimeter:
- Coverage: $|\eta| < 5$
- Fe/Scintillator tiles (central)
- Cu/W-LAr (forward)
- Trigger, Jet and MET measurement.

$$\frac{\sigma(E)}{E} \sim \frac{50\%}{\sqrt{E}} \oplus 0.03$$

Muon spectrometer: (MS)
- Coverage: $|\eta| < 2.7$
- Air core toroids (<B> 0.5T)
- Gas-based muon chambers
- Provides muon trigger
- $\sigma(p)/p \sim 10\%$

Inner detector: (ID)
- Coverage: $|\eta| < 2.5$
- Solenoid B = 2T
- Si Pixels, microstrips, and TRT straw tracker system.

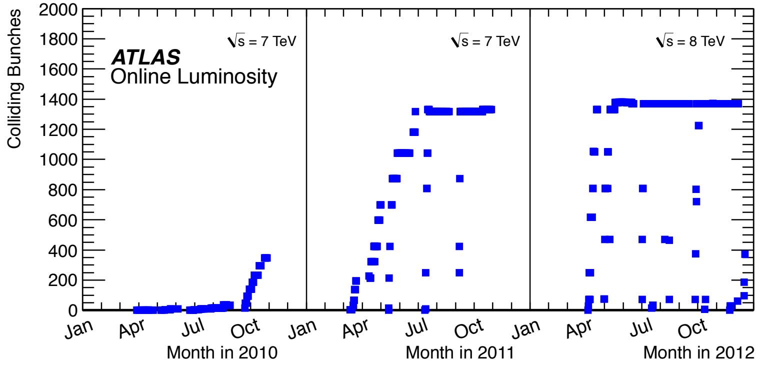
$$\frac{\sigma(p_T)}{p_T} \sim 3.8 \times 10^{-4} p_T(GeV) \oplus 0.015$$

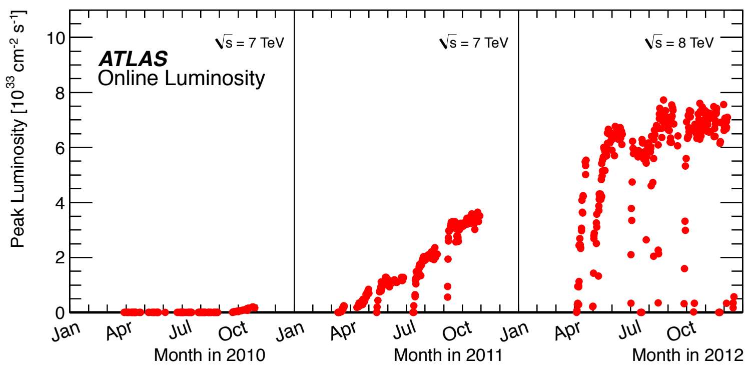
LAr Calorimeter:
- Pb-LAr accordion structure
- e/γ trigger, identification and measurement.

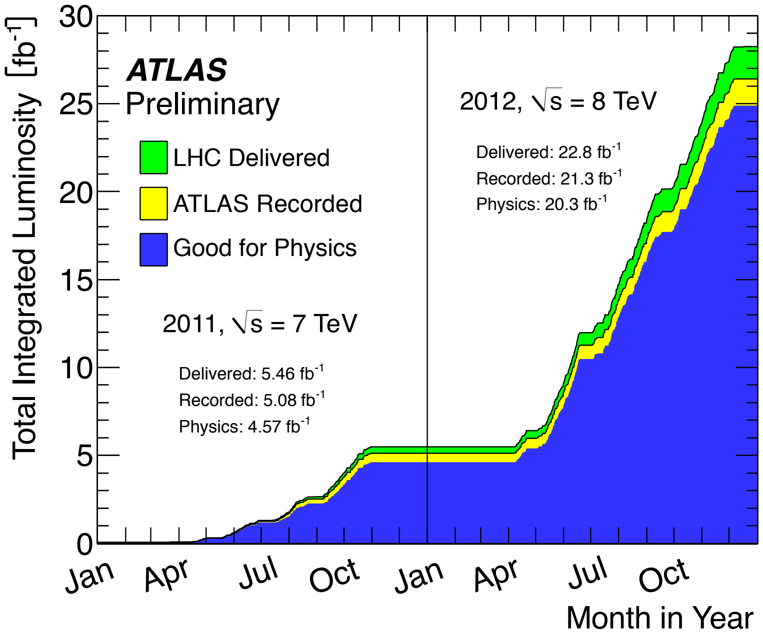
$$\frac{\sigma(E)}{E} \sim \frac{10\%}{\sqrt{E}}$$

Data samples

- Excellent LHC (& ATLAS) performance during Run 1 data sample: 4.57 fb^{-1} at 7 TeV; 20.4 fb^{-1} at 8 TeV







This result uses the full Run 1 data sample & supersedes the previous ATLAS result.

Run 2 at the LHC is continuing.

ATLAS has been upgraded with an inner pixel layer (IBL) during the shutdown: will improve background suppression for Run 2 update of this analysis.

ANALYSIS

Overview

- Signal:
  - Select signal di-muon events from data.
  - Extract yield using an un-binned extended maximum-likelihood fit to the data.
  - Use control samples to understand background suppression BDT and other crosschecks.

- Normalise signal to $B^{\pm} \to J/\psi K^{\pm}$.
  - Requires knowledge of decay constants f_s/f_d and f_u/f_d.
  - Use the ATLAS result for f_s/f_d and assume Isospin:

$$f_s/f_d = 0.240 \pm 0.020$$

ATLAS Collaboration, PRL **115** (2015) 262001 (arXiv:1507.08925)

  - Check normalisation mode against $B^{\pm} \to J/\psi \pi^{\pm}$.

Trigger

- Use a di-muon signature in the MS to form triggers for these decays.

- 2011 data: $p_T(\mu) > 4\,GeV$

- 2012 data:
  - The $p_T(\mu) > 4\,GeV$ line was pre-scaled.
  - Secondary trigger lines were used to retain signal efficiency; using tighter P_T or $|\eta|$ constraints.

T_1: "higher threshold" trigger with $p_T > 6$ GeV and > 4 GeV respectively for the two muons;
T_2: "barrel" trigger with $p_T > 4$ GeV for both muon candidates and at least one of them with $|\eta| < 1.05$ (and T_1 requirement not satisfied);
T_3: basic di-muon trigger with $p_T > 4$ GeV for both muon candidates (and T_1, T_2 requirements not satisfied).

Event Selection

- μ$^{\pm}$: Tracks using ID and MS information

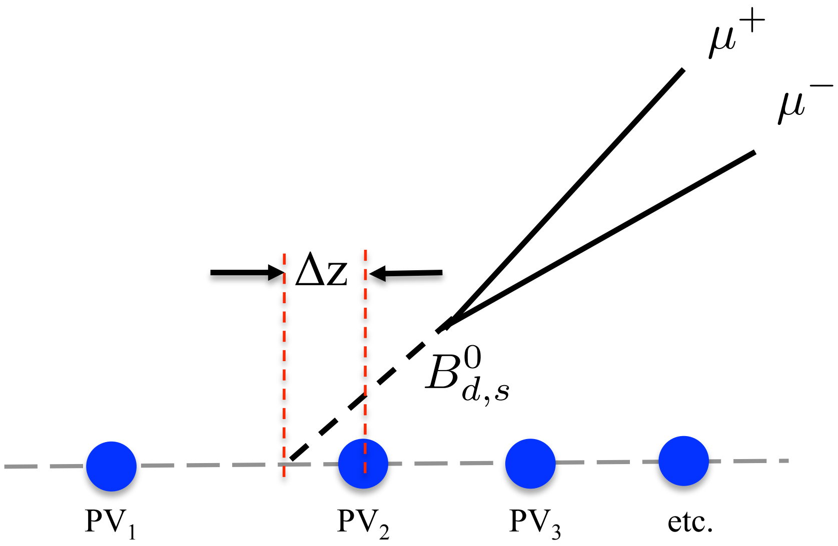
$$p_T(\mu) > 4\,GeV$$

$$|\eta| < 2.5$$

- *B*: $P_T(B) > 8\,GeV$

$$|\eta| < 2.5$$

$$m_{\mu\mu} \in [4766, 5966]\,GeV$$



- Primary Vertex (PV):
  - Reconstructed using tracks not associated with secondary vertex.
  - Project *B* 3-vector back to collision axis.
  - Minimise (in *z*) POCA to PV_i.

Backgrounds

- Extraction of limit or branching fraction ($\mathcal{B}$) depends critically on our understanding of the background...
  - Combinatoric events:
    - Pairs of independent muons selected from the event that pass the reconstruction.
  - Partially Reconstructed decays (PRD):

    $$B^0_{(s)} \to \mu^+\mu^- X$$

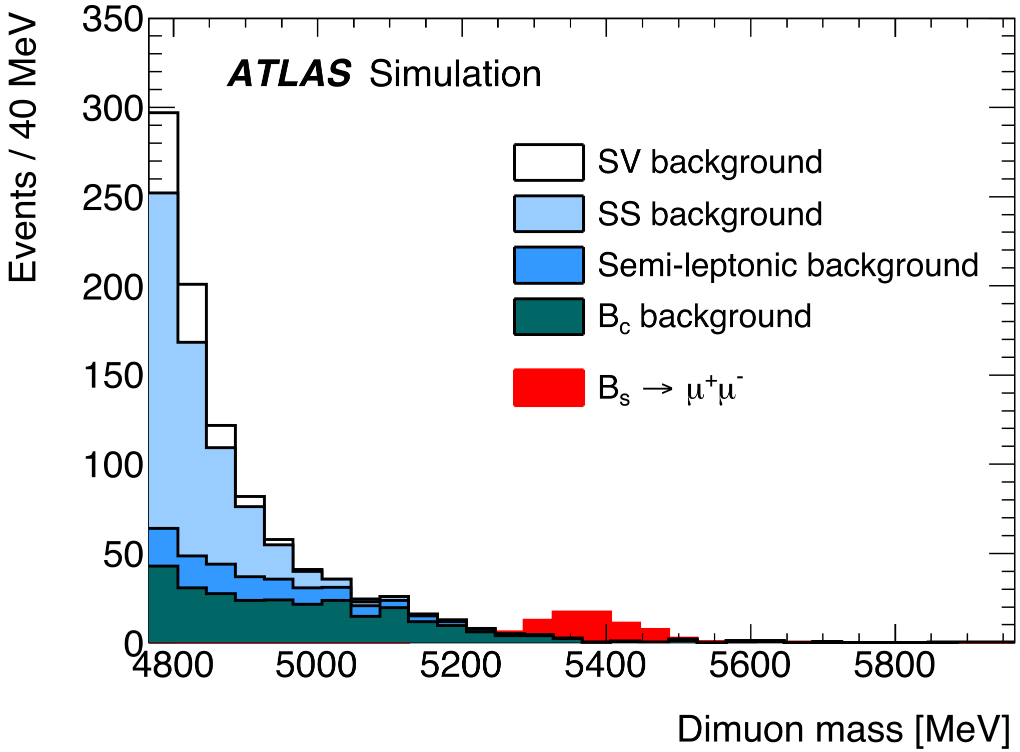
    The system *X* is not reconstructed; accumulates on the low mass sideband.
  - Peaking background:

    $$B^0_{(s)} \to hh'$$

    - Mis-id the hadron (*h*'=*π*, *K*) as a *μ* pair.

Backgrounds

- Partially Reconstructed Decays:



Same Vertex (SV):

μ pair from same vertex; e.g.

$$b \to s\mu^+\mu^-, qJ/\psi$$

Same Side (SS):

from decay cascades; e.g.

$$b \to c\mu\nu \to s(d)\mu\mu\nu\overline{\nu}$$

Semi-leptonic:

with hadron mis-id as μ; e.g.

$$B \to h\mu\nu,\ \Lambda_b \to p\mu\nu$$

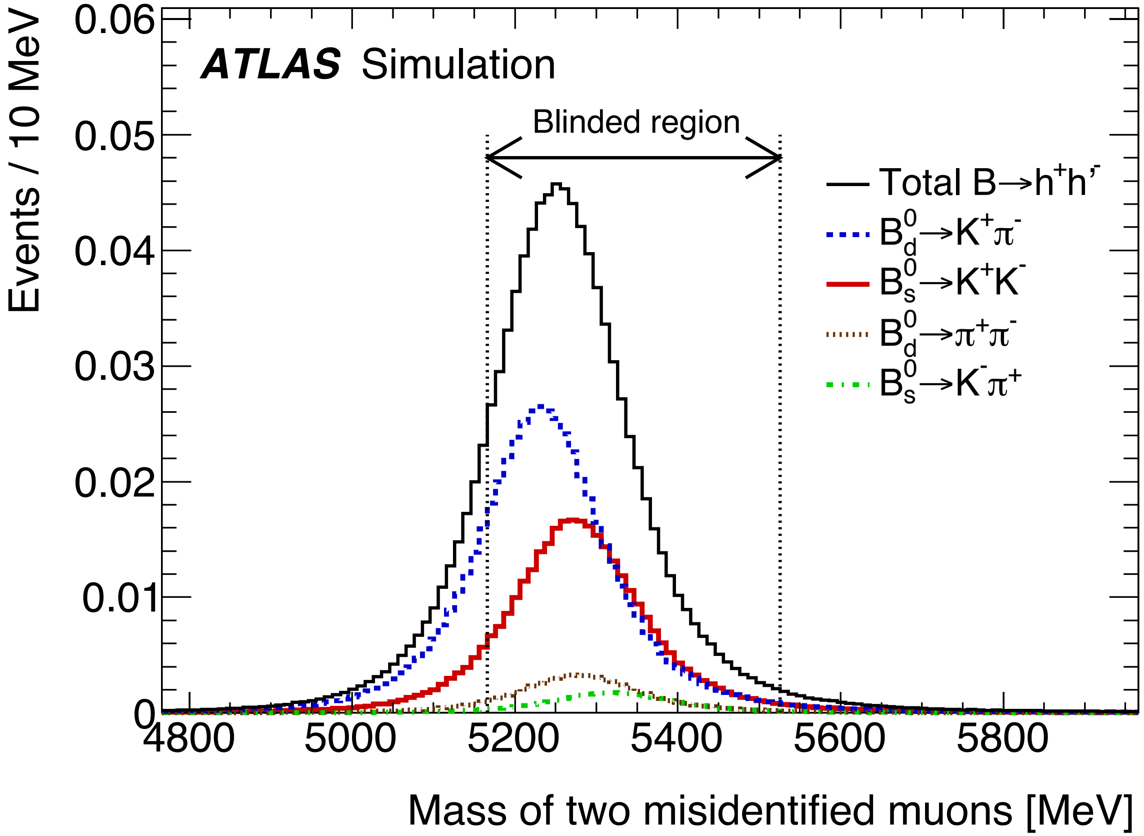
B_c:

e.g. $B_c \to J/\psi\mu\nu \to 3\mu\nu$

- Several competing contributions:
  - Dominated by same side (SS) background.
  - Feeds into the signal region from low mass side.

Backgrounds

- Two-body decays: $B^0_{(s)} \to hh'$



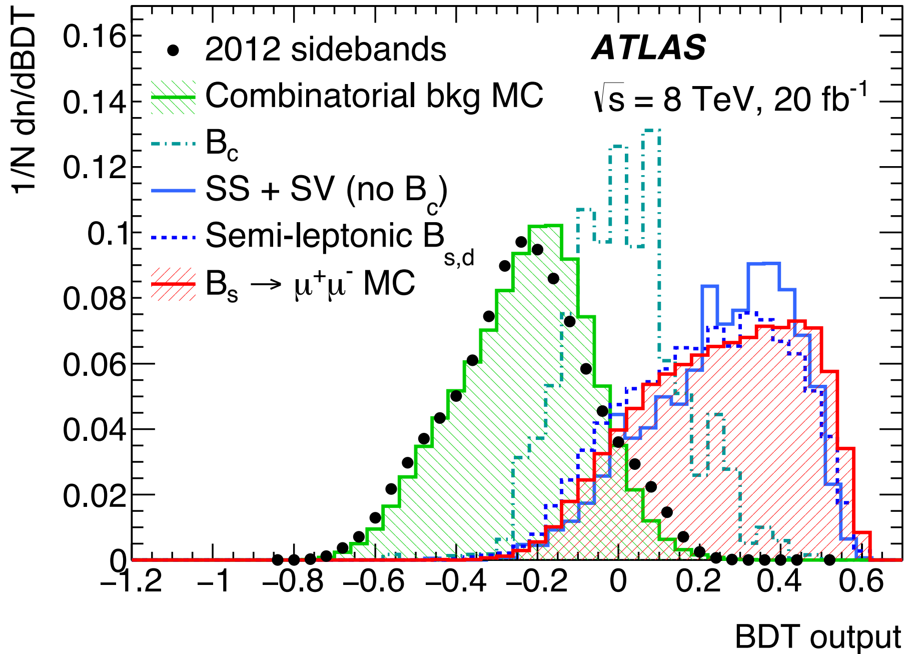
- Dangerous background that can mimic signal.
- Use fake μ BDT to suppress this contribution.
  - μ fake rate sub-per-mille
- Negligible contribution from

$$\Lambda_b \to p\pi$$

- Poor separation between B_d and B_s mesons.
- Expect 1.0±0.4 events of *hh'* background.

Backgrounds

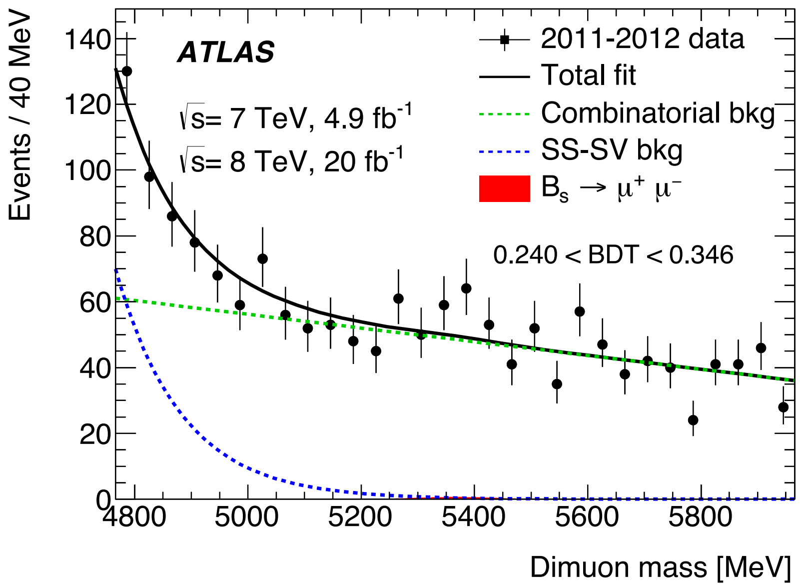
- Use a boosted decision tree (BDT) using 15 variables for combinatoric background suppression: (see backup for details)

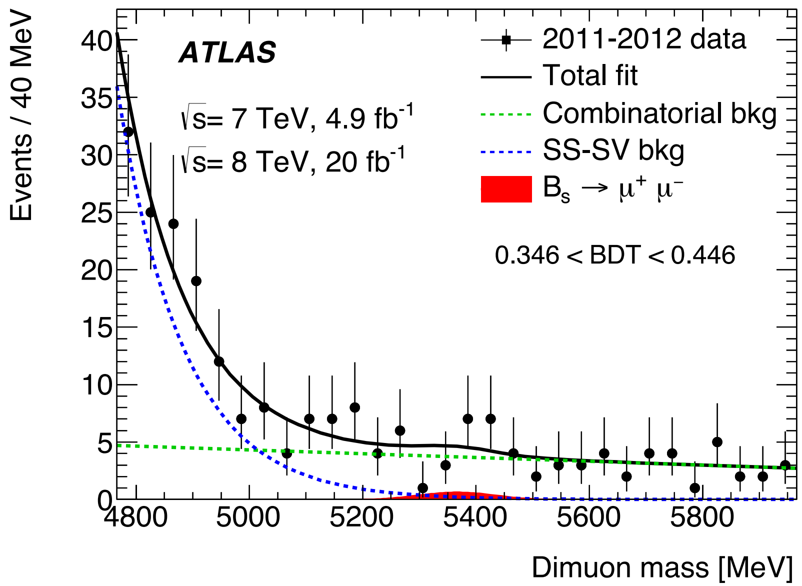


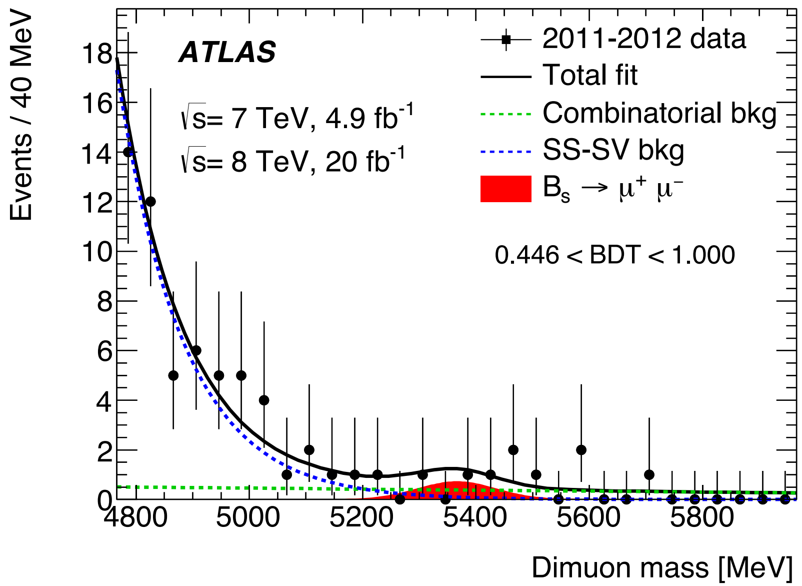
- Use $B^{\pm} \to J/\psi K^{\pm}$ and $B_s \to J/\psi\phi$ control samples to validate data/MC agreement.
  - Data/MC difference accounted for as systematic error.
- Split the data into bins of equal expected signal yield (increasing purity).

Signal Fit Result

- Fitted signal yields are N(B_s)=16±12, N(B_d)=−11±9 events.



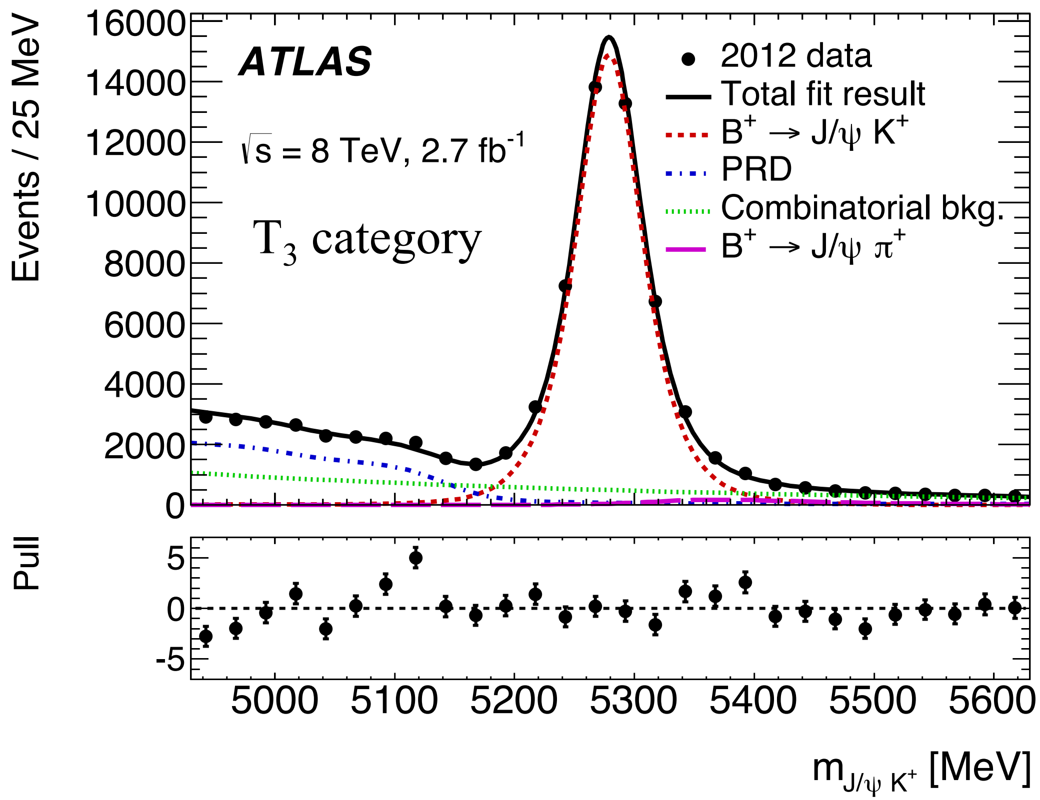




Normalisation Mode: $B^{\pm} \to J/\psi K^{\pm}$

- 4 component fit:
  - $B^{\pm} \to J/\psi K^{\pm}$
  - $B^{\pm} \to J/\psi \pi^{\pm}$
  - Combinatoric background
  - Partially reconstructed decays

Category	$N_{J/\psi K^{+}}$
T_1	$46\,860 \pm 290 \pm 280$
T_2	$5\,200 \pm 84 \pm 100$
T_3	$2\,512 \pm 91 \pm 42$
2011	$95\,900 \pm 420 \pm 1\,100$



- Validated by computing the ratio:

$$\rho_{\pi/K} = \frac{\mathcal{B}(B^+ \to J/\psi\pi^+)}{\mathcal{B}(B^+ \to J/\psi K^+)} = 0.035 \pm 0.003 \pm 0.012$$

- c.f. PDG average of 0.040±0.004.

Branching Fraction Calculation

- Several trigger lines are used for this analysis; these are accounted for via:

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

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

  $f_s/f_d = 0.240 \pm 0.020$ and $f_u/f_d = 1$

  ε_i are efficiencies for signal/normalisation

  α_k are trigger/luminosity weight factors

- and the N_i are yields obtained from the fits.

Branching Fraction Calculation

- Several trigger lines are used for this analysis; these are accounted for via:

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

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

  $f_s/f_d = 0.240 \pm 0.020$ and $f_u/f_d = 1$

  ε_i are efficiencies for signal/normalisation

  α_k are trigger/luminosity weight factors

- and the N_i are yields obtained from the fits.

Branching Fraction Calculation

- Several trigger lines are used for this analysis; these are accounted for via:

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

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

  $f_s/f_d = 0.240 \pm 0.020$ and $f_u/f_d = 1$

  ε_i are efficiencies for signal/normalisation

  α_k are trigger/luminosity weight factors

- and the N_i are yields obtained from the fits.

Branching Fraction Calculation

- Several trigger lines are used for this analysis; these are accounted for via:

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

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

  $f_s/f_d = 0.240 \pm 0.020$ and $f_u/f_d = 1$

  ε_i are efficiencies for signal/normalisation

  α_k are trigger/luminosity weight factors

- and the N_i are yields obtained from the fits.

Recall that the ATLAS f_s/f_d analysis matches the fiducial selection for this search.

Branching Fraction Calculation

- Several trigger lines are used for this analysis; these are accounted for via:

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

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

  $f_s/f_d = 0.240 \pm 0.020$ and $f_u/f_d = 1$

  ε_i are efficiencies for signal/normalisation

  α_k are trigger/luminosity weight factors

- and the N_i are yields obtained from the fits.

$$\mathcal{B}(B_d \to \mu^+\mu^-) < 4.2 \times 10^{-10}$$
$$\mathcal{B}(B_s \to \mu^+\mu^-) < 3.0 \times 10^{-9}$$

(@ 95% C.L.)

$$\mathcal{B}(B_s \to \mu^+\mu^-) = (0.9^{+1.1}_{-0.8}) \times 10^{-9}$$

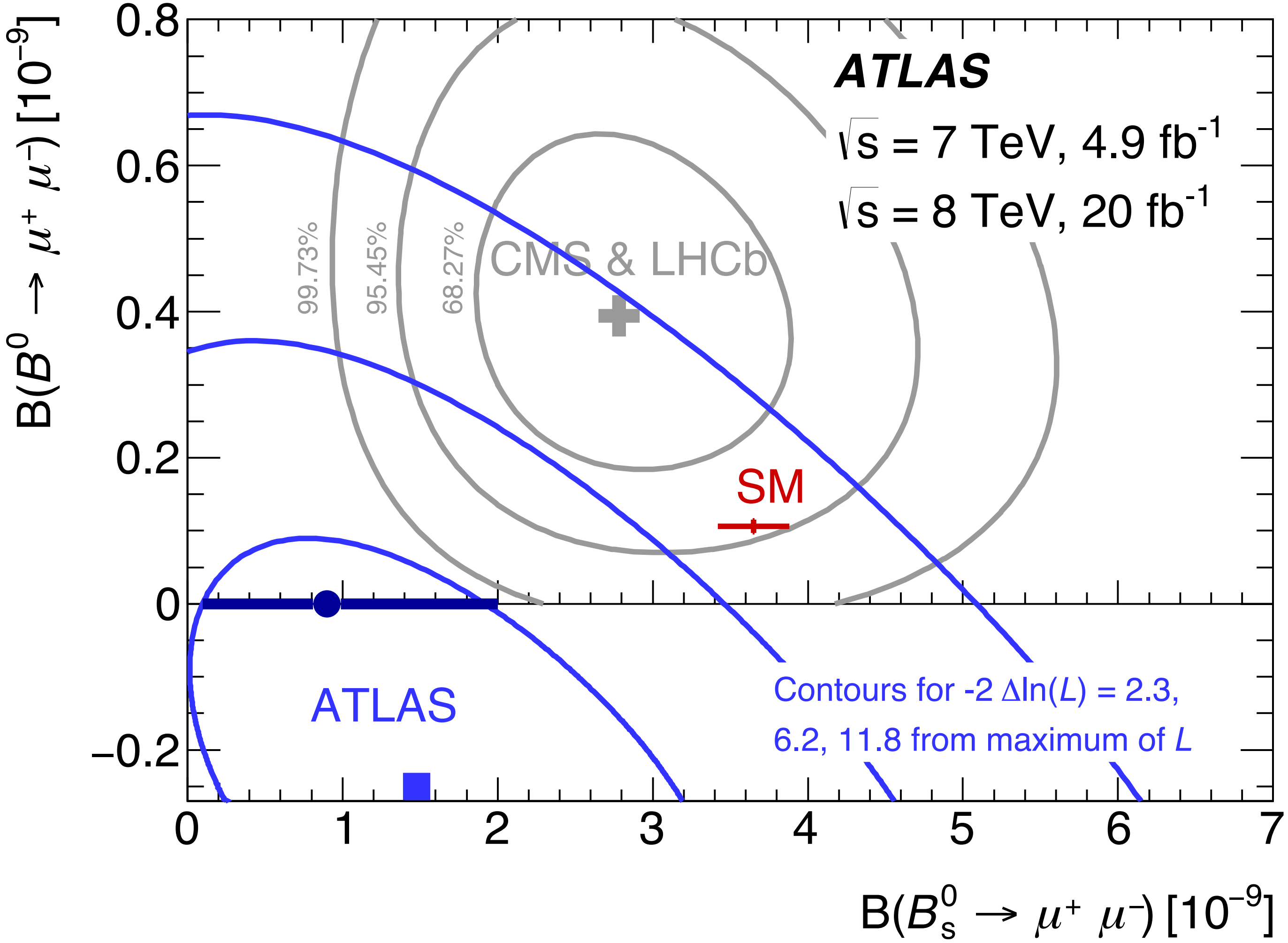
Systematic Errors

- Statistically limited measurement.
- Main systematic contribution from continuum BDT.

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

Result

- ATLAS is consistent with the SM, LHCb and CMS.



- Room for NP destructively interfering with the SM.

Summary

- Presented results for the ATLAS search for $B_{s,d} \to \mu^+\mu^-$ using Run 1 data from the LHC.

- We obtain:

$$\mathcal{B}(B_d \to \mu^+\mu^-) < 4.2 \times 10^{-10}$$
$$\mathcal{B}(B_s \to \mu^+\mu^-) < 3.0 \times 10^{-9}$$

(@ 95% C.L.)

$$\mathcal{B}(B_s \to \mu^+\mu^-) = (0.9^{+1.1}_{-0.8}) \times 10^{-9}$$

- Compatible with (lower than) the SM (p-value of 0.048).

- Compatible with (lower than) the other LHC experiments.

See the following paper for more details:
CERN-EP-2016-064, Eur. Phys. J. C76 (2016) no.9, 513.

BACKUP

Fake μ BDT

- Critical aspect of the analysis:
  - Get this wrong and you could fake a signal/destroy your sensitivity to NP.

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.

Use the following control samples to validate the performance of the fake μ BDT:

$$B^{\pm} \to J/\psi K^{\pm}$$
$$\phi \to K^{+}K^{-}$$

  - Fake rate @ 0.09% / 0.04% / <0.01% for *K* / *π* / *p*.
  - Punch through @ 3% (8%) level for *K* (*π*).

Continuum BDT

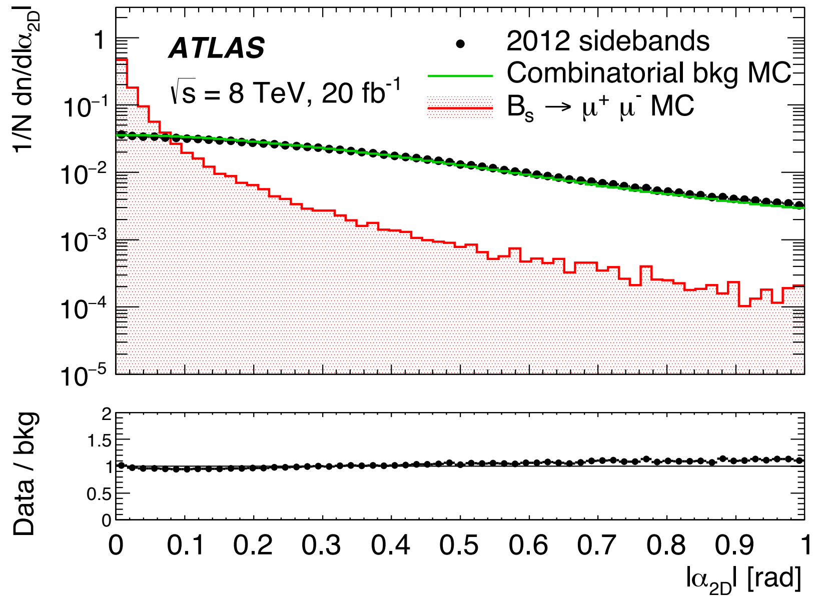
Variables ranked by order of importance in the MVA

Variable	Description
p_T^B	Magnitude of the B candidate transverse momentum $\vec{p}_T^B$.
$\chi^2_{\mathrm{PV,DV}\,xy}$	Significance of the separation $\vec{\Delta x}$ between production (PV) and decay (DV) vertices in the transverse projection: $\vec{\Delta x}_T \cdot \Sigma^{-1}_{\vec{\Delta x}_T} \cdot \vec{\Delta x}_T$, where $\Sigma_{\vec{\Delta x}_T}$ is the covariance matrix.
ΔR	3-dimensional opening between $\vec{p}^B$ and $\vec{\Delta x}$: $\sqrt{\alpha_{2D}{}^2 + \Delta\eta^2}$
$\lvert\alpha_{2D}\rvert$	Absolute value of the angle between $\vec{p}_T^B$ and $\vec{\Delta x}_T$ (transverse projection).
L_{xy}	Projection of $\vec{\Delta x}_T$ along the direction of $\vec{p}_T^B$: $(\vec{\Delta x}_T \cdot \vec{p}_T^B)/\lvert\vec{p}_T^B\rvert$.
IP_B^{3D}	3-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 (3-dimensional).
$\Delta\phi_{\mu\mu}$	Difference in azimuthal angle between the momenta of the two tracks forming the B candidate.
$\lvert d_0\rvert^{max}$ 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.
$\lvert d_0\rvert^{min}$ 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_L^{min}	Value of the smaller projection of the momenta of the muon candidates along $\vec{p}_T^B$.
$I_{0.7}$	Isolation variable defined as ratio of $\lvert\vec{p}_T^B\rvert$ to the sum of $\lvert\vec{p}_T^B\rvert$ and of the transverse momenta of all additional tracks contained within a cone $\Delta R < 0.7$ from the B direction. Only tracks with $p_T > 0.5$ GeV and associated to the same PV as the B candidate are included in the sum.
DOCA_{xtrk}	DOCA of the closest additional track to the decay vertex of the B candidate. Tracks associated to a PV different from the B candidate are excluded.
N_{xtrk}^{close}	Number of additional tracks compatible with the decay vertex (DV) of the B candidate with $\ln(\chi^2_{\mathrm{xtrk,DV}}) < 1$. The tracks associated to a PV different from the B candidate are excluded.
$\chi^2_{\mu,xPV}$	Minimum χ^2 for the compatibility of a muon in the B candidate with a PV different from the one associated to the B candidate.

Validated using:

$$B^\pm \to J/\psi K^\pm$$
$$B_s \to J/\psi \phi$$

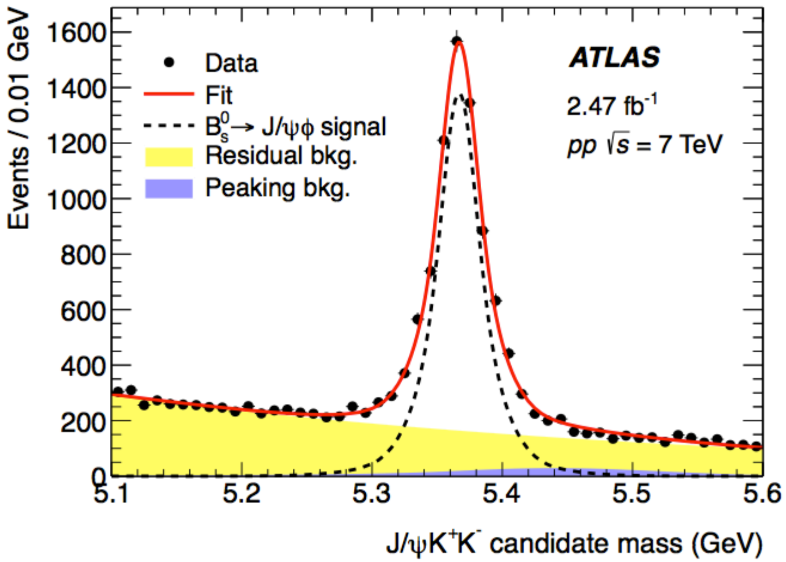
control modes and sidebands.

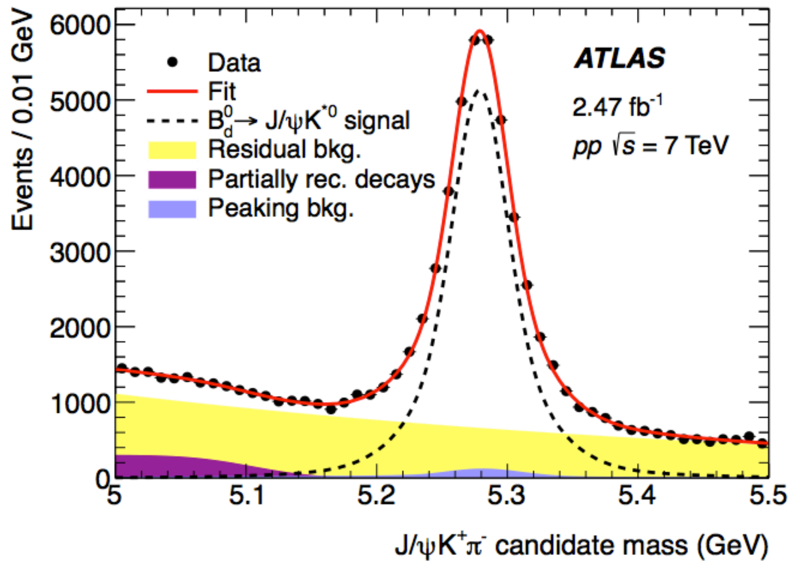


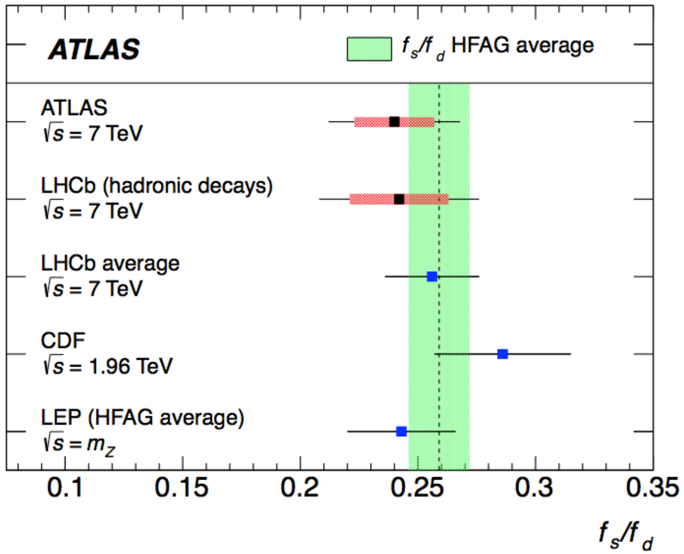
Data/MC discrepancies do not significantly alter the BDT performance and are accounted for in the systematic error reported.

$$\frac{f_s}{f_d} = \frac{N_{B_s^0}}{N_{B_d^0}} \frac{\mathcal{B}(B_d^0 \to J/\psi K^{*0})}{\mathcal{B}(B_s^0 \to J/\psi \phi)} \frac{\mathcal{B}(K^{*0} \to K^+\pi^-)}{\mathcal{B}(\phi \to K^+K^-)} \mathcal{R}_{\text{eff}},$$

ATLAS Collaboration, PRL **115** (2015) 262001 (arXiv:1507.08925)



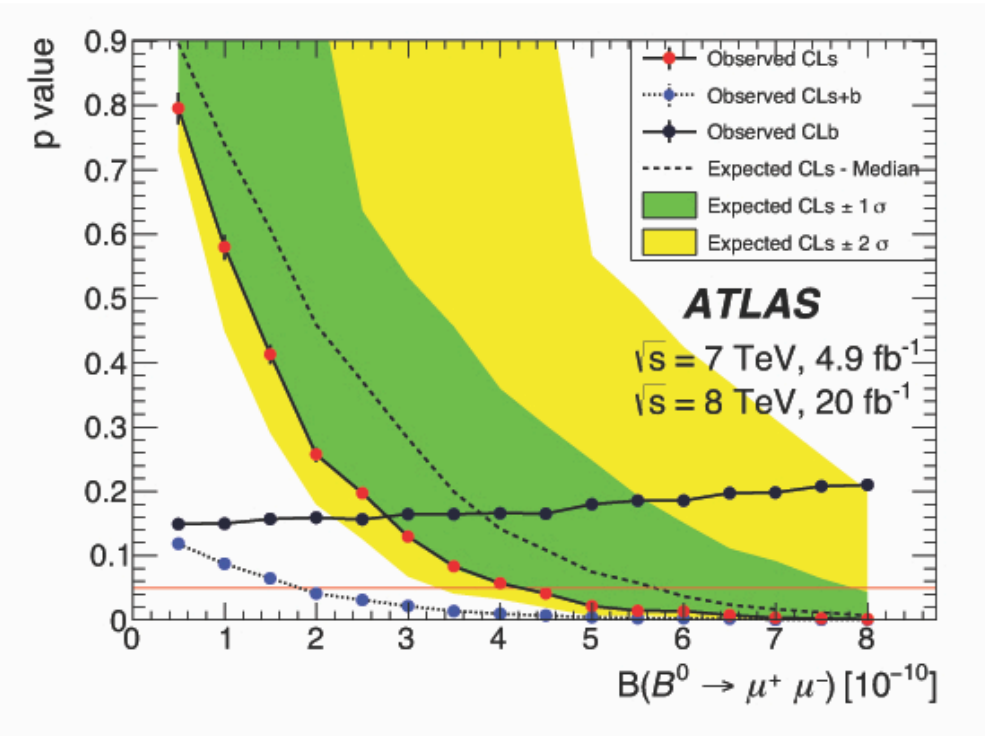


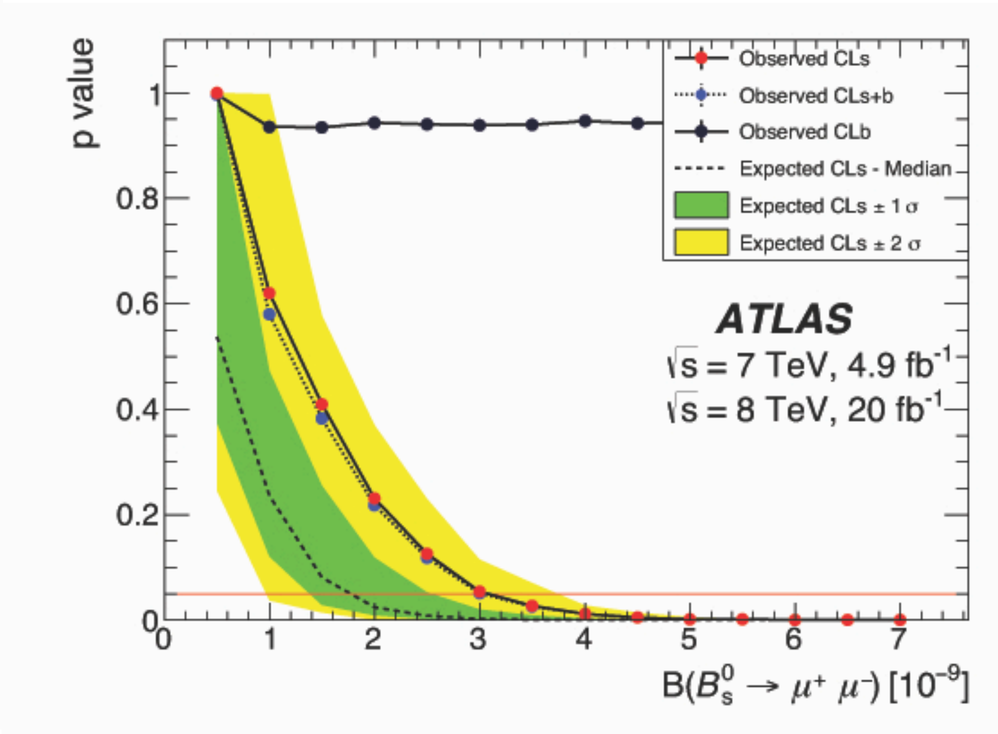


- ATLAS uses a complementary method to LHCb in order to measure f_s/f_d.
- Use $B_d \to J/\psi K^*$ & $B_s \to J/\psi\phi$.
- Results are in good agreement with each other.

$$f_s/f_d = 0.240 \pm 0.020$$

CL_s limits

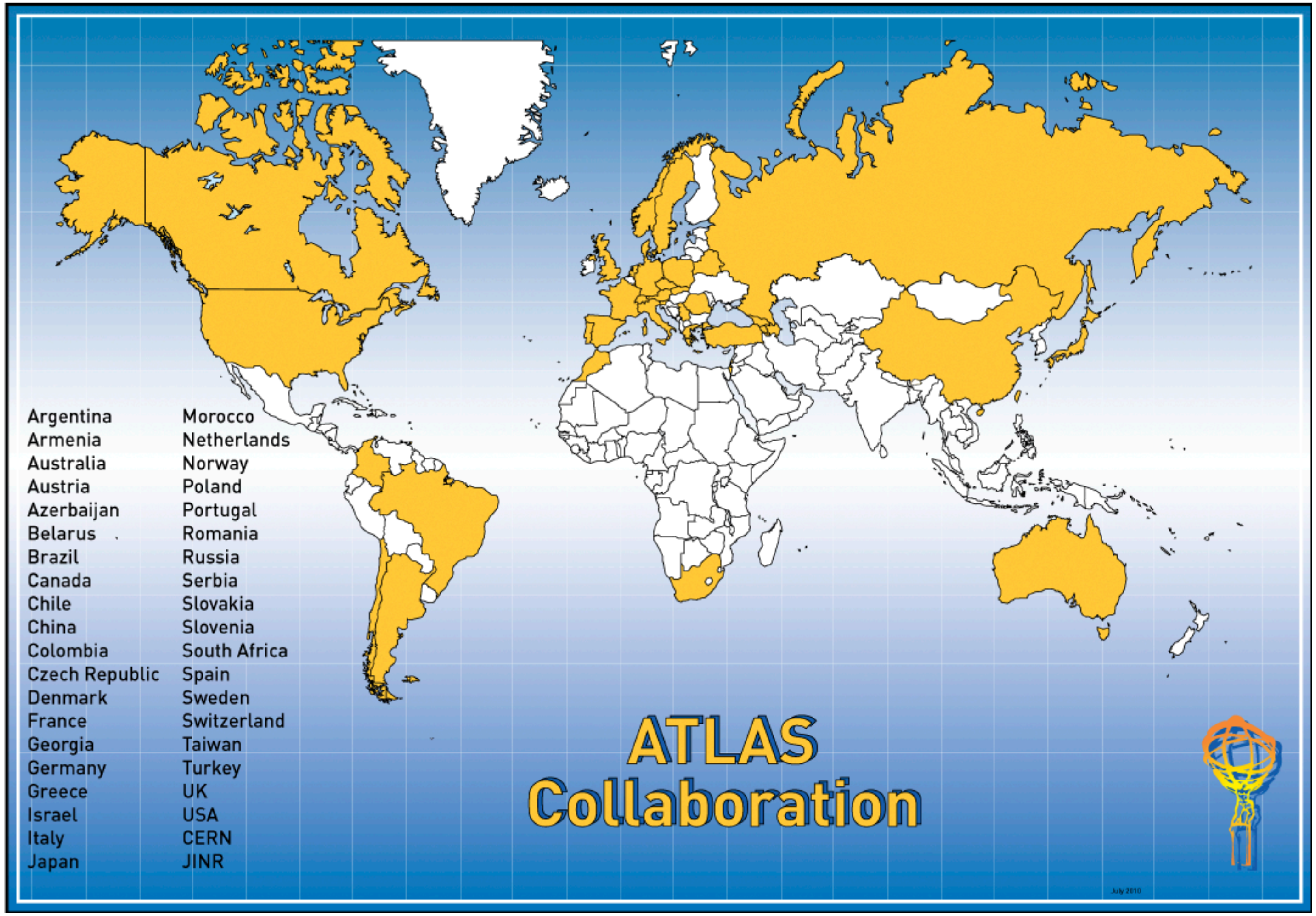




$$\mathcal{B}(B_d \to \mu^+\mu^-) < 4.2 \times 10^{-10}$$
$$\mathcal{B}(B_s \to \mu^+\mu^-) < 3.0 \times 10^{-9}$$
(@ 95% C.L.)

ATLAS Experiment

- Over 3000 Physicists from 177 institutes in 38 countries



ATLAS Detector

Inner Detector (ID) consists of:

Pixel detectors

Semiconductor Tracker (SCT)

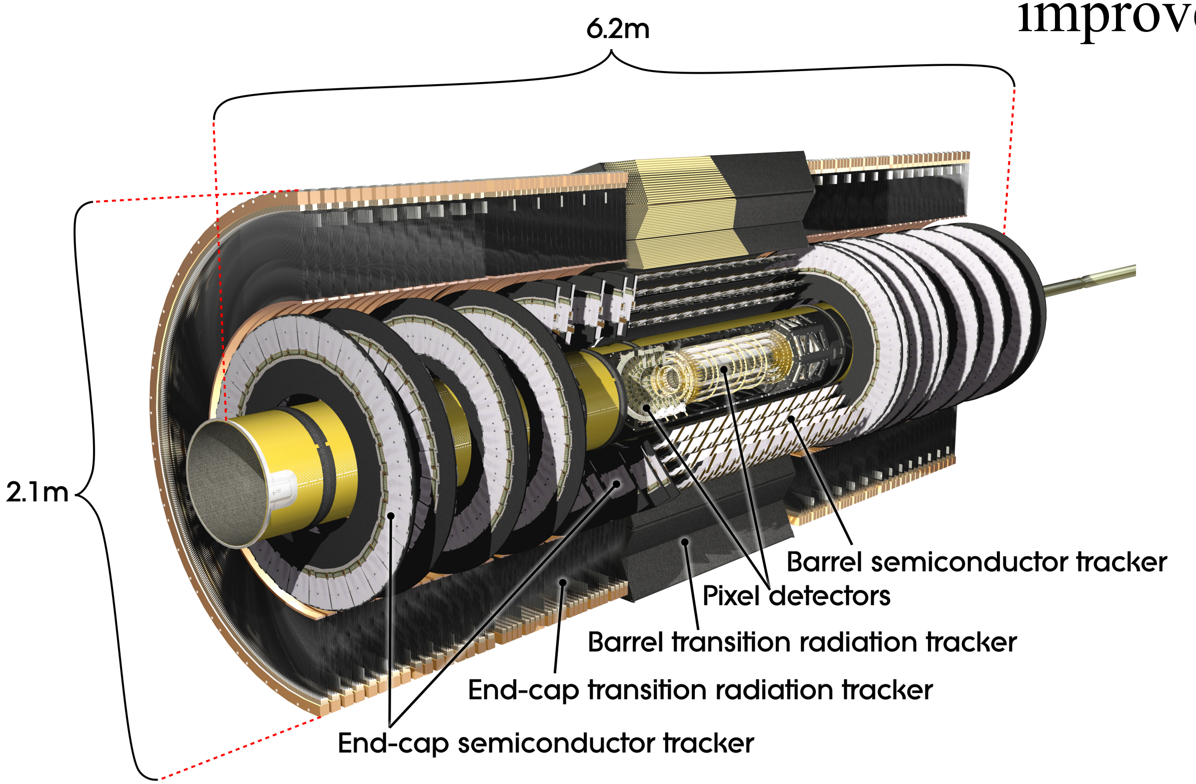
Transition radiation tracker (TRT)

Provides:

Precision tracking and vertexing

New for run 2:

Small radius pixel layer (IBL); expect improved d_0, lifetime resolution etc.



Inner detector: (ID)

- Coverage: $|\eta| < 2.5$
- Solenoid B = 2T
- Si Pixels, microstrips, and TRT straw tracker system.

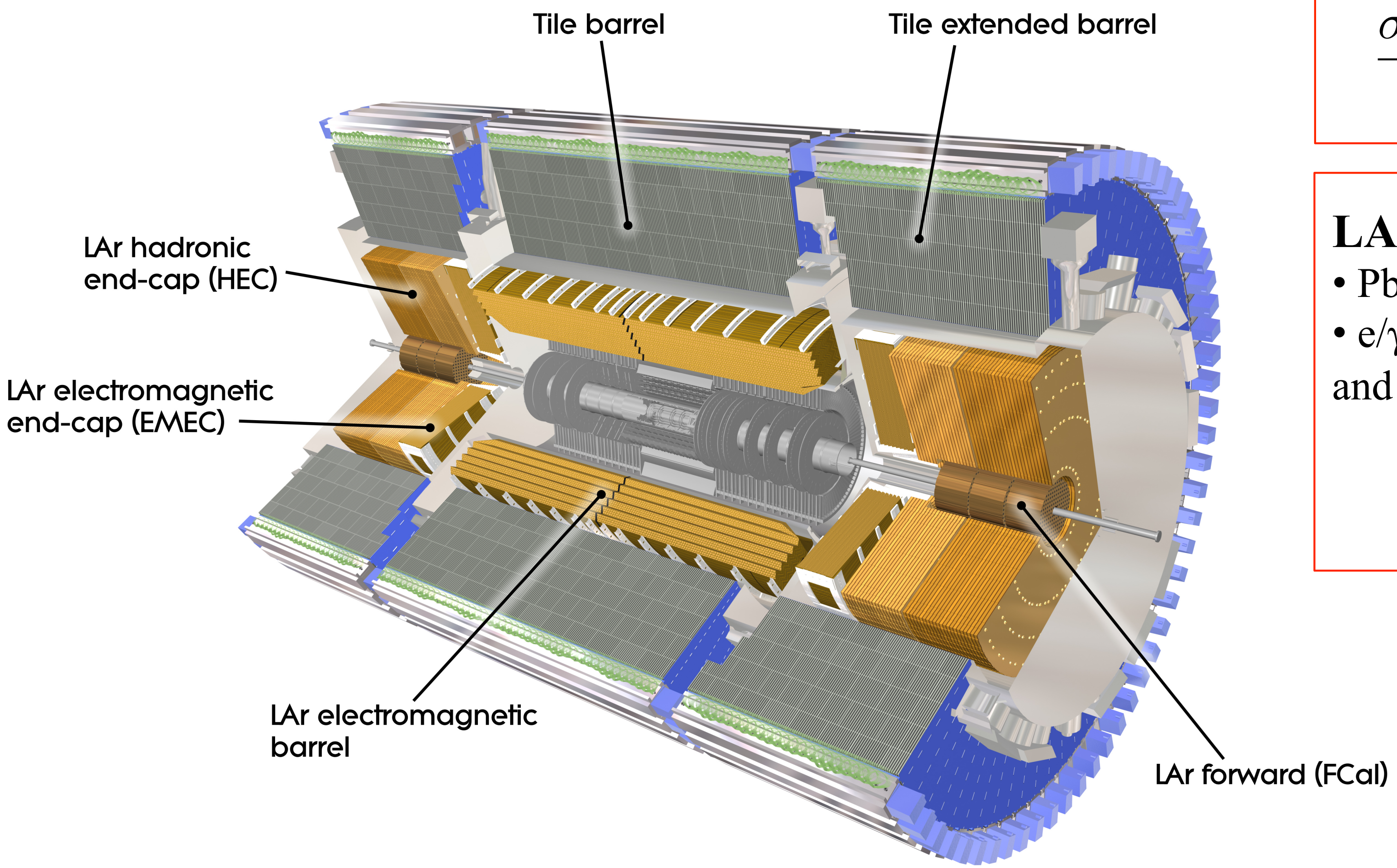
$$\frac{\sigma(p_T)}{p_T} \sim 3.8 \times 10^{-4} p_T(GeV) \oplus 0.015$$

ATLAS Detector

Calorimeter consists of:

LAr barrel and end-caps

Tile calorimeter



Hadronic Calorimeter:

- Coverage: $|\eta| < 5$
- Fe/Scintillator tiles (central)
- Cu/W-LAr (forward)
- Trigger, Jet and MET measurement.

$$\frac{\sigma(E)}{E} \sim \frac{50\%}{\sqrt{E}} \oplus 0.03$$

LAr Calorimeter:

- Pb-LAr accordion structure
- e/γ trigger, identification and measurement.

$$\frac{\sigma(E)}{E} \sim \frac{10\%}{\sqrt{E}}$$

ATLAS Detector

Magnet systems consists of:

Solenoid (surrounds the ID; provides B = 2T)

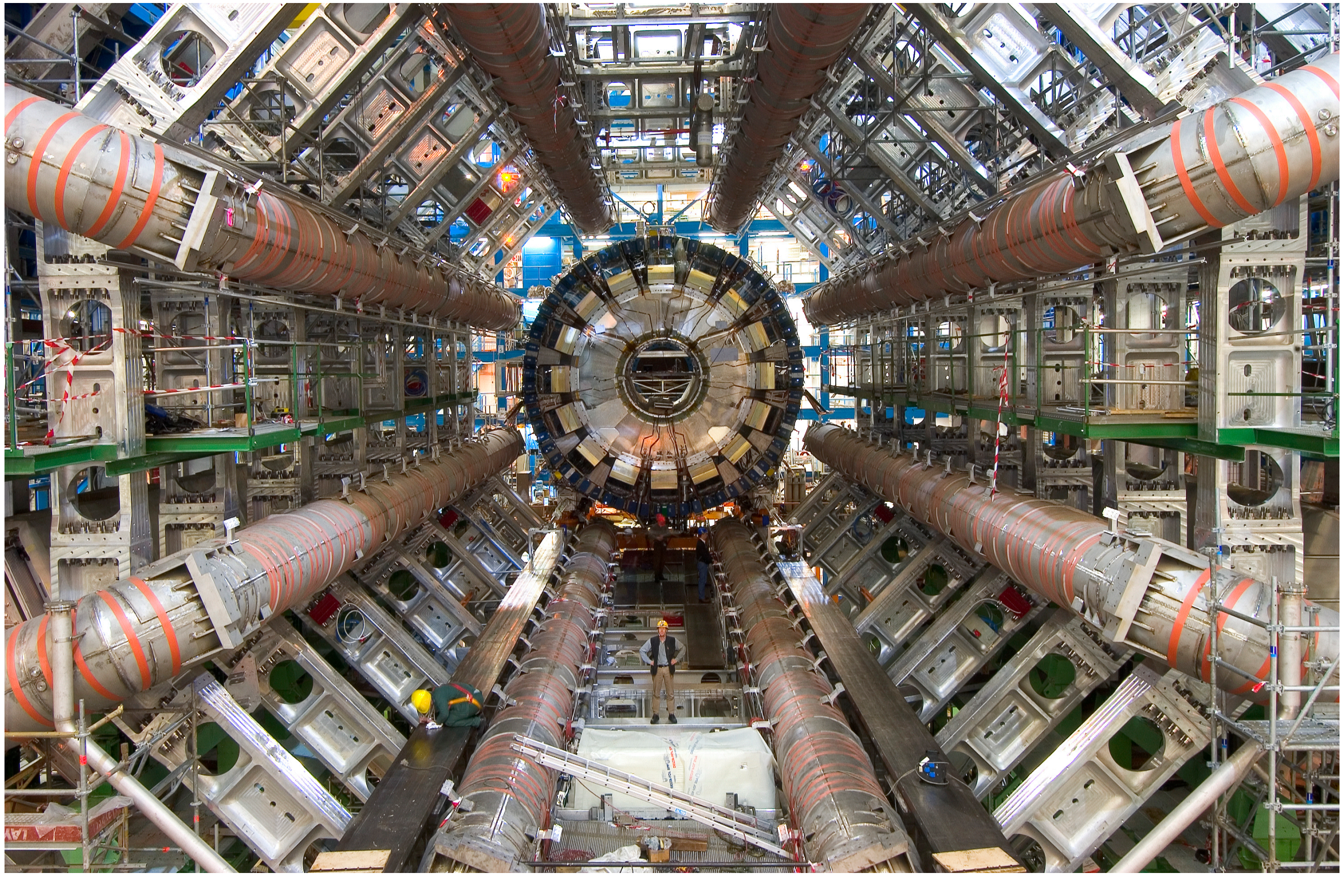
Toroid (embedded in the muon system; provides <B> ~ 0.5 T)

ATLAS Detector

Magnet systems consists of:

Solenoid (surrounds the ID; provides B = 2T)

Toroid (embedded in the muon system; provides <B> ~ 0.5 T)



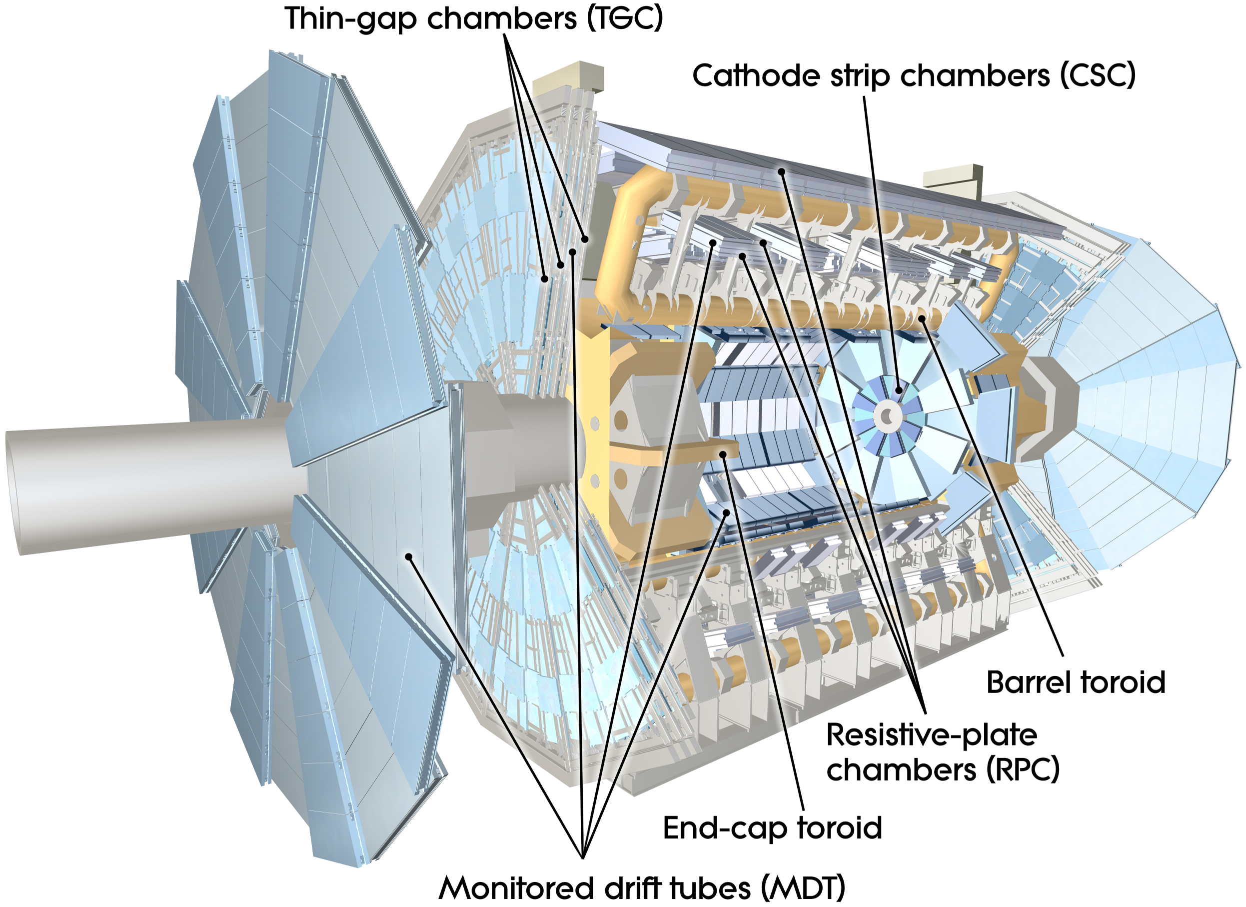
ATLAS Detector

Muon system consists of:

RPCs
Monitored Drift tubes
Thin gap chambers
Cathode strip chambers

Provides:

Muon identification
Tracking information



Muon spectrometer: (MS)

- Coverage: $|\eta| < 2.7$
- Air core toroids ($\langle B \rangle$ 0.5T)
- Gas-based muon chambers
- Provides muon trigger
- $\sigma(p)/p \sim 10\%$