$B^0_{(s)} \rightarrow \mu^+\mu^-$ at ATLAS: Measurement and Prospects

Iskander Ibragimov
UNIVERSITÄT SIEGEN

The 12th Annual Meeting of the Helmholtz Alliance "Physics at the Terascale"

DESY, Hamburg
November 26-28 2018

SPONSORED BY THE
Federal Ministry of Education and Research

Outline

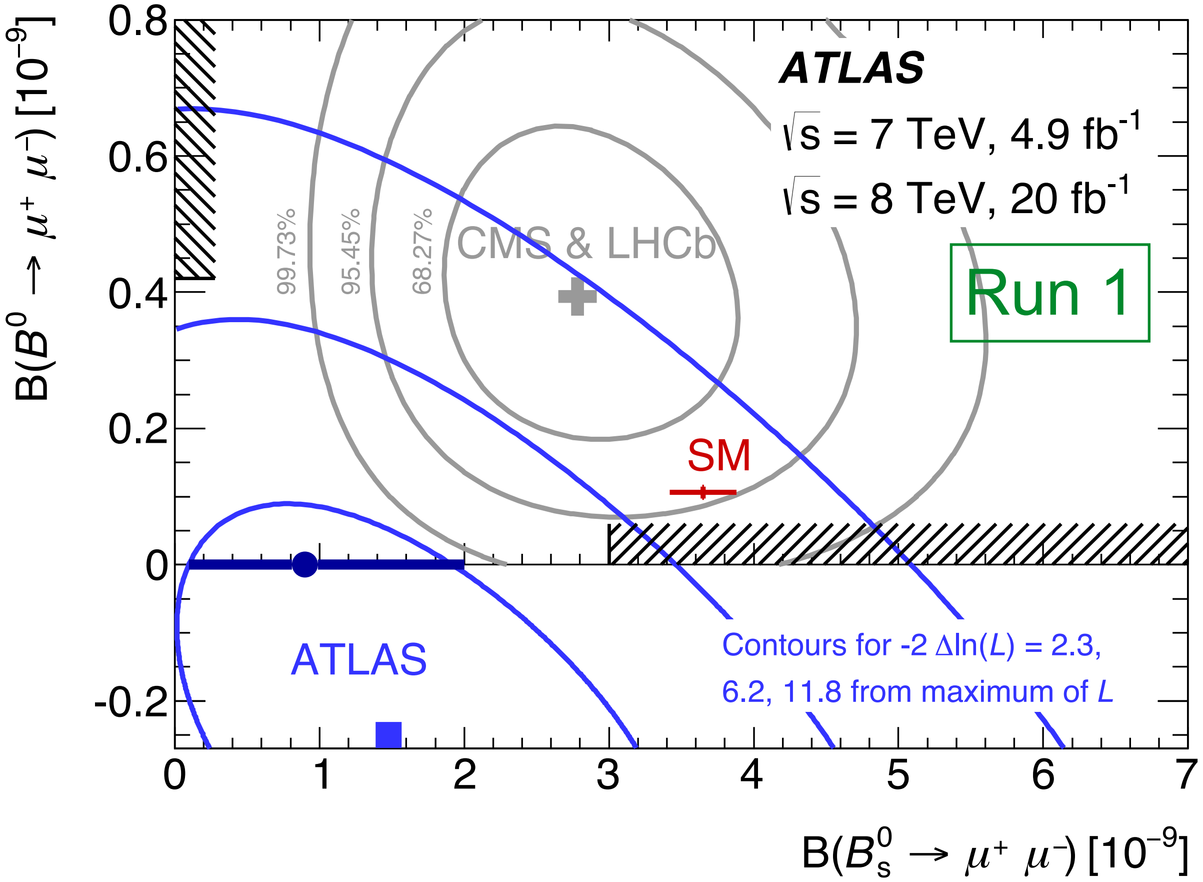
- Introduction
- $B^0_{(s)} \to \mu^+\mu^-$ analysis with 2015 and 2016 data [ATLAS-CONF-2018-046]
- $B^0_{(s)} \to \mu^+\mu^-$ prospects for Run 2 and HL-LHC [ATL-PHYS-PUB-2018-005]
- Summary

Introduction

- strongly suppressed and precisely calculated in the SM:
  - $BR(B^0_s \to \mu^+\mu^-) = (3.65 \pm 0.23)\times 10^{-9}$
  - $BR(B^0 \to \mu^+\mu^-) = (1.06 \pm 0.09)\times 10^{-10}$

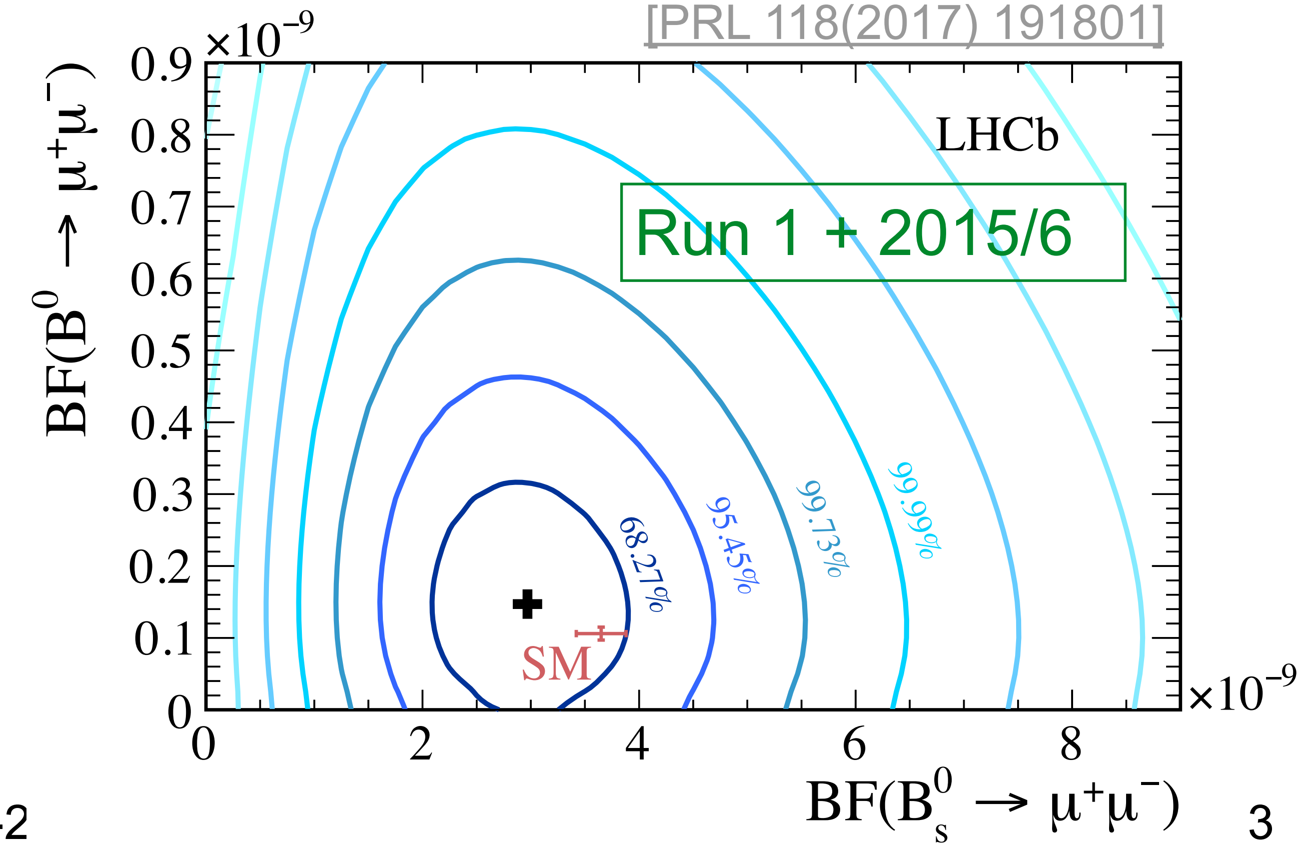
➡ powerful indirect search for New Physics

[Nature 522 (2015) 68-72, EPJC 76 (2016) 513]



- Run 1 CMS & LHCb combination
  - $BR(B^0_s \to \mu^+\mu^-) = (2.8^{+0.7}_{-0.6})\times 10^{-9}$
  - $BR(B^0 \to \mu^+\mu^-) = (3.9^{+1.6}_{-1.4})\times 10^{-10}$
- ATLAS Run 1 result
  - $BR(B^0_s \to \mu^+\mu^-) = (0.9^{+1.1}_{-0.8})\times 10^{-9}$
  - $BR(B^0 \to \mu^+\mu^-) < 4.2 \times 10^{-10}$ (at 95% CL)
    - compatible with SM at ~2σ

[PRL 118(2017) 191801]



- Run 1 + 2015/6 LHCb result
  - $BR(B^0_s \to \mu^+\mu^-) = (3.0\pm 0.6^{+0.3}_{-0.2})\times 10^{-9}$
  - $BR(B^0 \to \mu^+\mu^-) < 3.4 \times 10^{-10}$ (at 95% CL)

all LHC results consistent with SM (so far)

$B^0_{(s)} \to \mu^+\mu^-$ analysis with 2015 and 2016 data

Analysis Strategy

- relative BR measurement (to $B^\pm \to J/\psi K^\pm$ reference channel)
  - partial cancelation of uncertainties (luminosity, cross-section, ...)

[HFLAV average]

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

$B^0_{(s)}$ yield from UML fit to $m_{\mu\mu}$ data

[world averages, PDG]

$$\mathcal{D}_{\mathrm{norm}} = N_{J/\psi K^+} \alpha \frac{\varepsilon_{\mu^+\mu^-}}{\varepsilon_{J/\psi K^+}}$$

$B^\pm$ yield from UML fit to $m_{\mu\mu K}$ data

prescaling factor

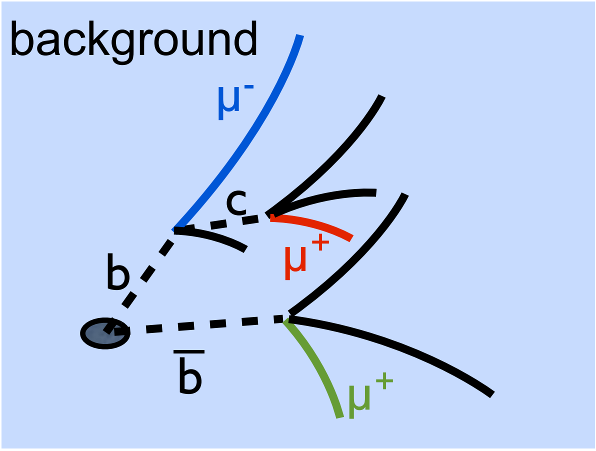
ratio of efficiencies evaluated on MC tuned to data

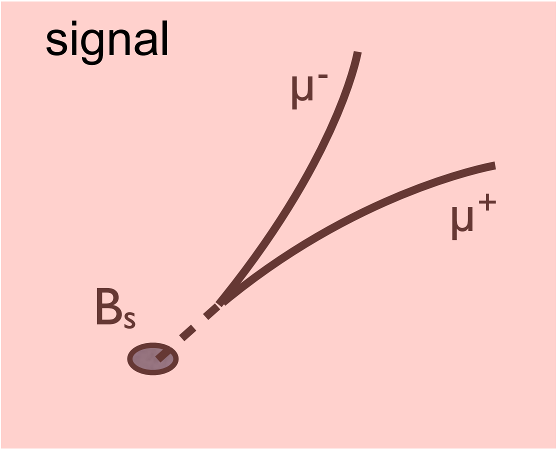
- use 'Tight' muon working point to suppress misidentified hadrons
- keep S/B discrimination unbiased - region [5166, 5526] MeV blinded
- use BDT against combinatorial background
- check Data/MC agreement on $B^\pm \to J/\psi K^\pm$ and $B^0_s \to J/\psi \phi$, $\phi \to K^+K^-$
- signal extraction with ML fit over $m_{\mu\mu}$ in four intervals of BDT
- use Neyman construction to improve statistical treatment of the result

Non-resonant Background

- combinatorial from real muons
  - dominant source
  - mostly from $b\bar{b} \rightarrow \mu^+\mu^- X$
  - modelled with dedicated 0.7G events MC

small mass dependence over entire search region

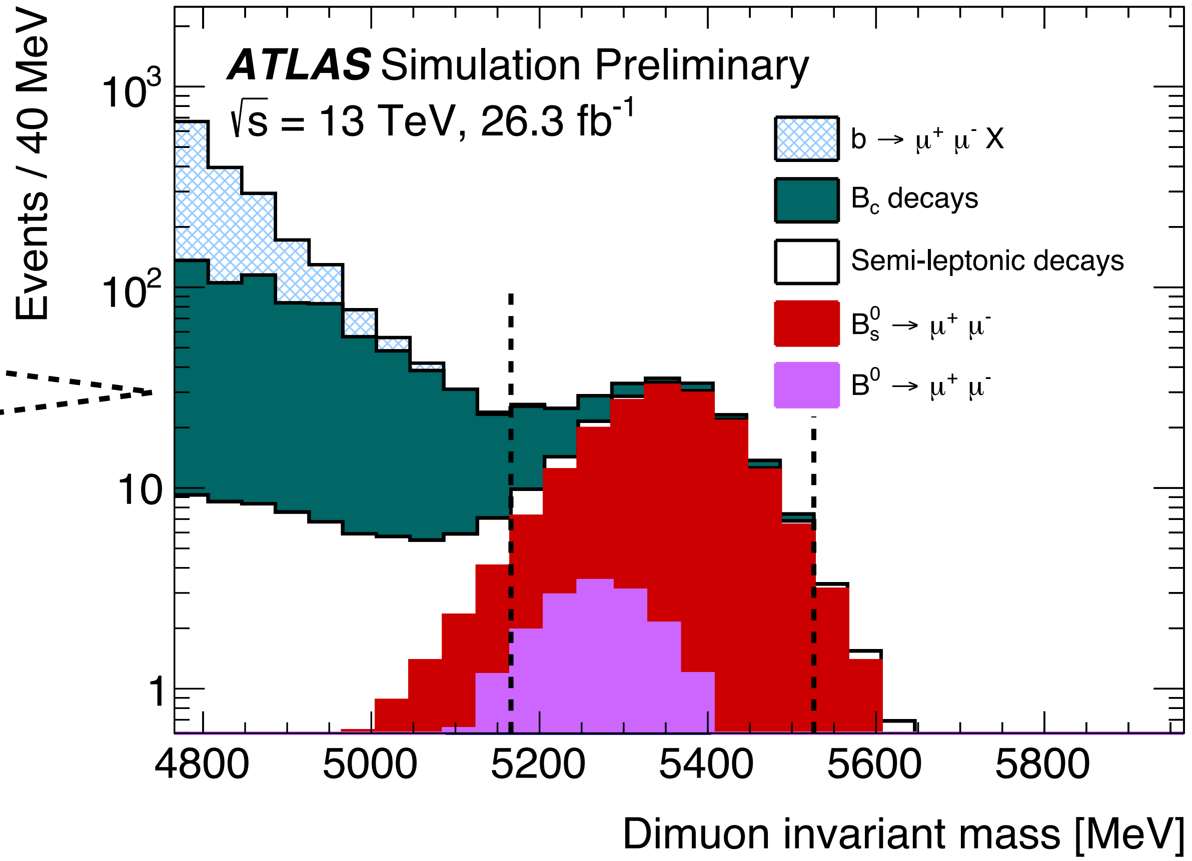




- partially reconstructed $B \rightarrow \mu\mu X$ decays
  - same vertex, e.g. $B^0 \rightarrow K^*\mu\mu$
  - semi-leptonic decay cascades $b \rightarrow c\mu\nu \rightarrow s(d)\mu\mu\nu\nu$
  - $B^{\pm}_c \rightarrow J/\psi\mu^{\pm}\nu \rightarrow \mu^{\pm}\mu^+\mu^-\nu$ decays
- semi-leptonic $B_{(s)} \rightarrow \mu h\nu$
  - hadron misidentified as muon
  - $B^0 \rightarrow \pi^-\mu^+\nu$, $B^0_s \rightarrow K^-\mu^+\nu$, $\Lambda_b \rightarrow p\mu\bar{\nu}$

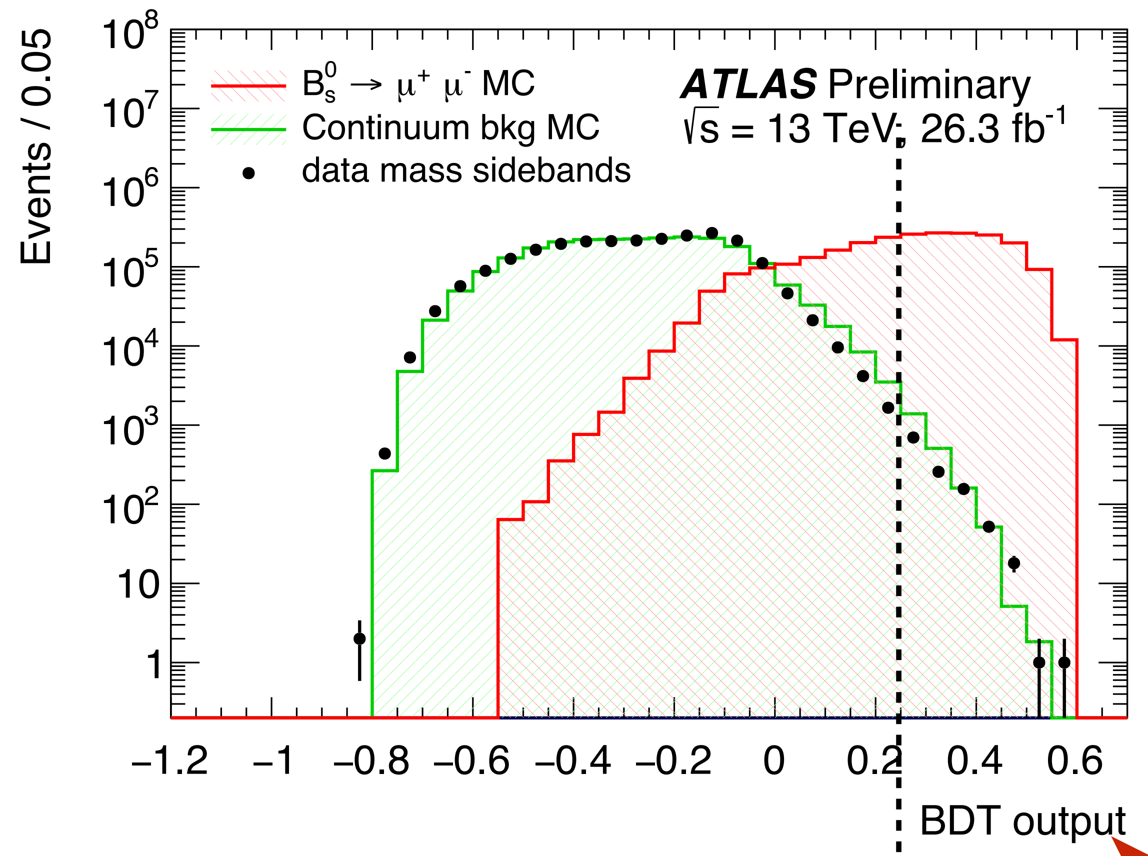
populate mainly left sideband

expectation based on PDG' BRs and integrated luminosity:

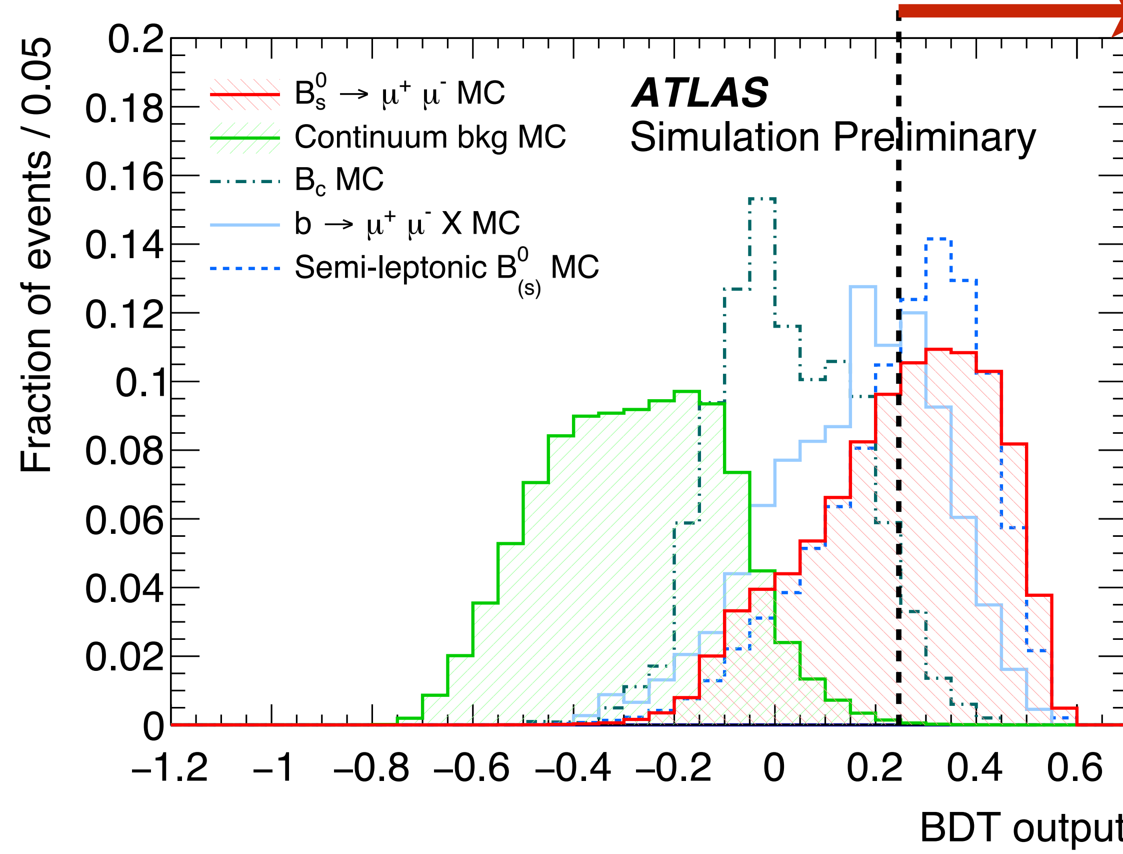


BDT Discriminant

- against combinatorial background
- 15 variables used for training
  - topology of reconstructed B-decay
  - properties of muons
  - underlying event
- split mass sidebands into 3 subsets
  - unbiased training/evaluation of 3 independent BDT's
  - similar performance of BDT's
- signal sensitive region BDT>0.25
  - ε_{sig} = 54%, ε_{bkg} = 0.03%
- suppresses other background components as well

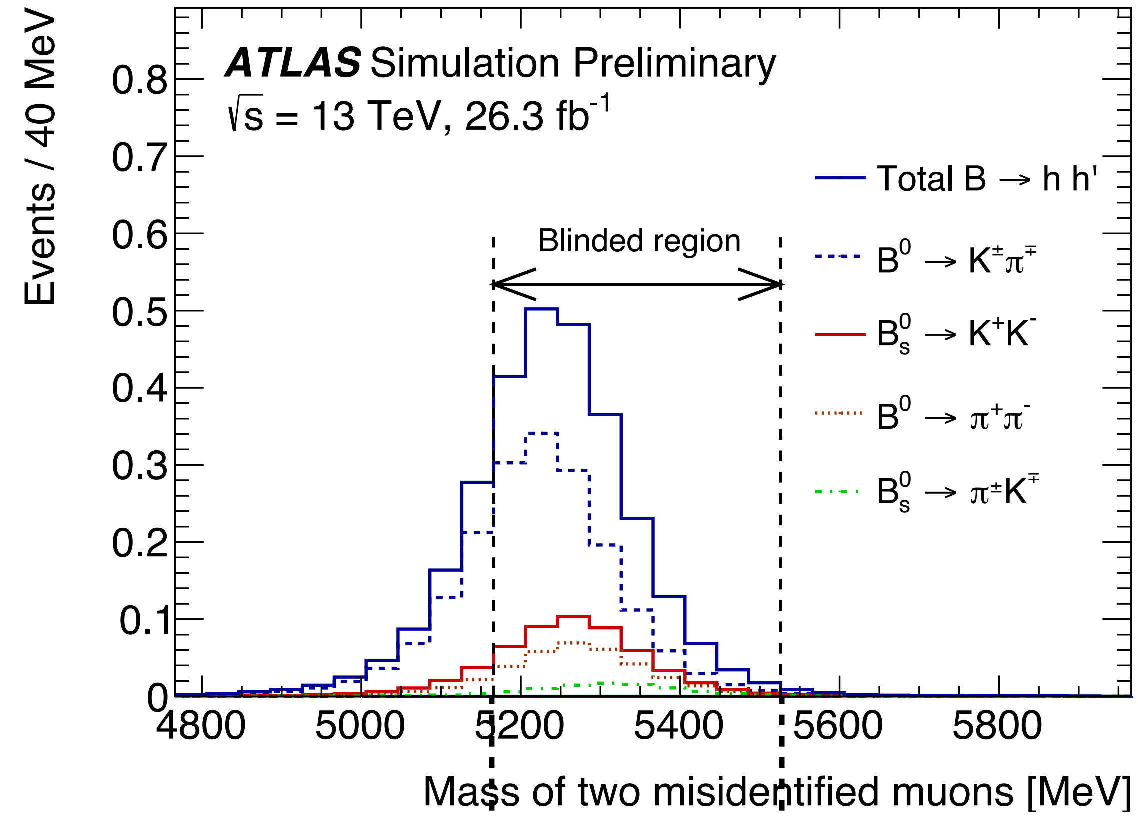


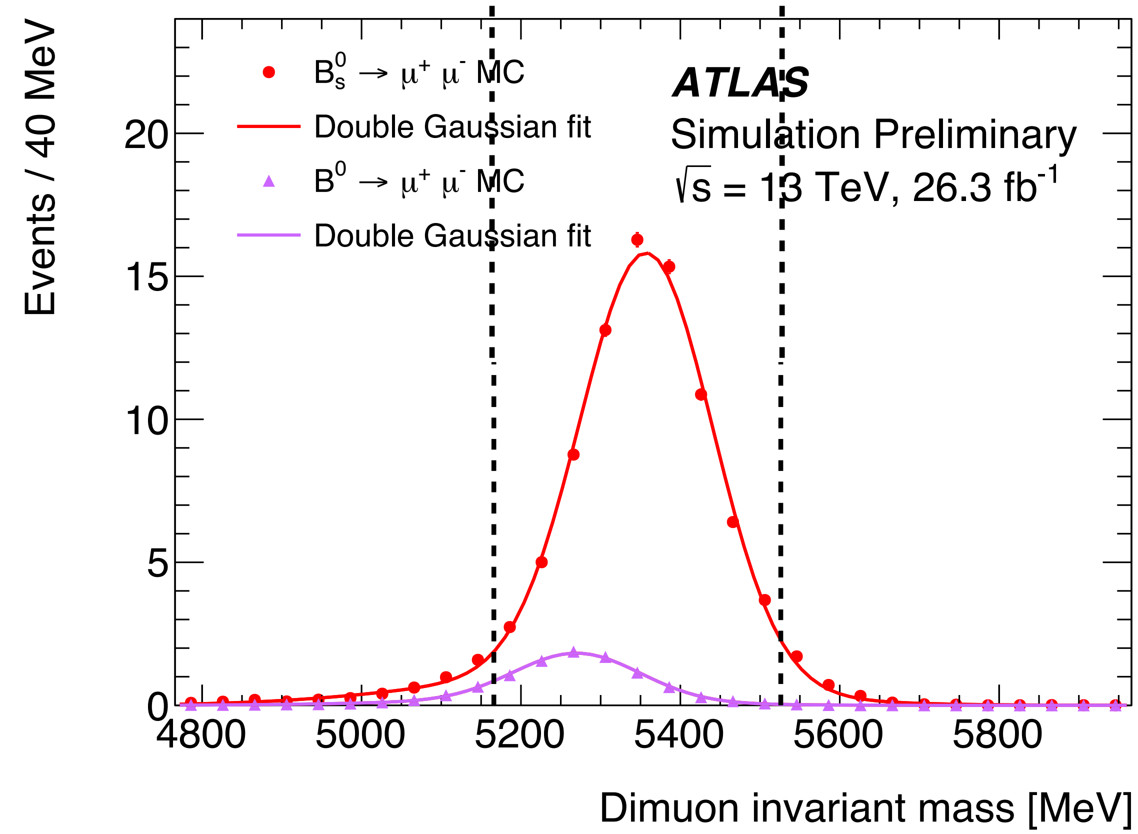
range of signal extraction



Peaking Background $B \to hh'$

- both *h* reconstructed as *μ* mainly due to decays in flight
- from simulation studies: $B^0_s \to K^+K^-$, $B^0 \to K^\pm\pi^\pm$, $B^0_s \to \pi^+\pi^-$ decays
- low rate but $B^0 \to \mu^+\mu^-$-like topology
- use 'Tight' μ working point: misID reduction by x 0.39^2 with ε_μ = 90%
- final P(misID) = 0.08% ($K^\pm$), 0.1% ($\pi^\pm$), <0.01% (p)
- signal region yield obtained inverting 'Tight' μ selection: $N_{peak\ bkg}$ = 2.9 ± 2.0 events
  - agrees with simulation: (2.7 ± 1.3) events
  - split equally among 3 BDT bins in the signal fit

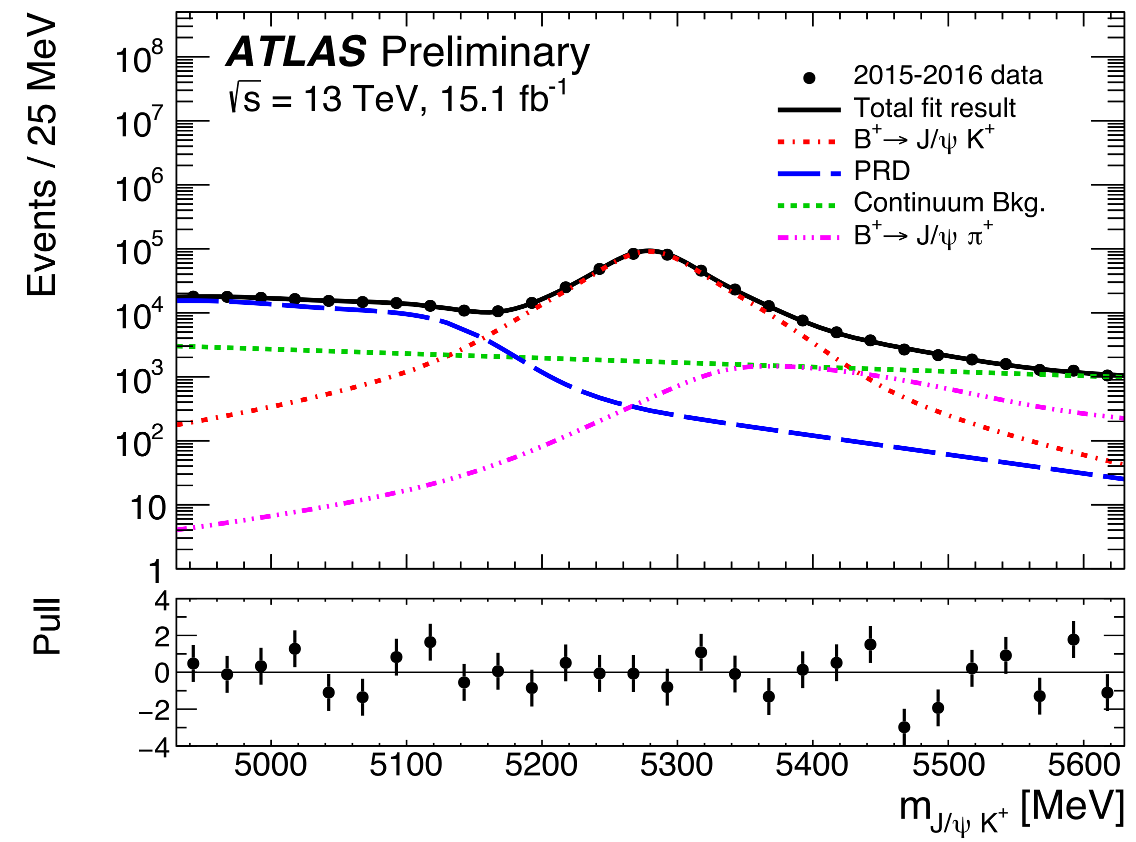




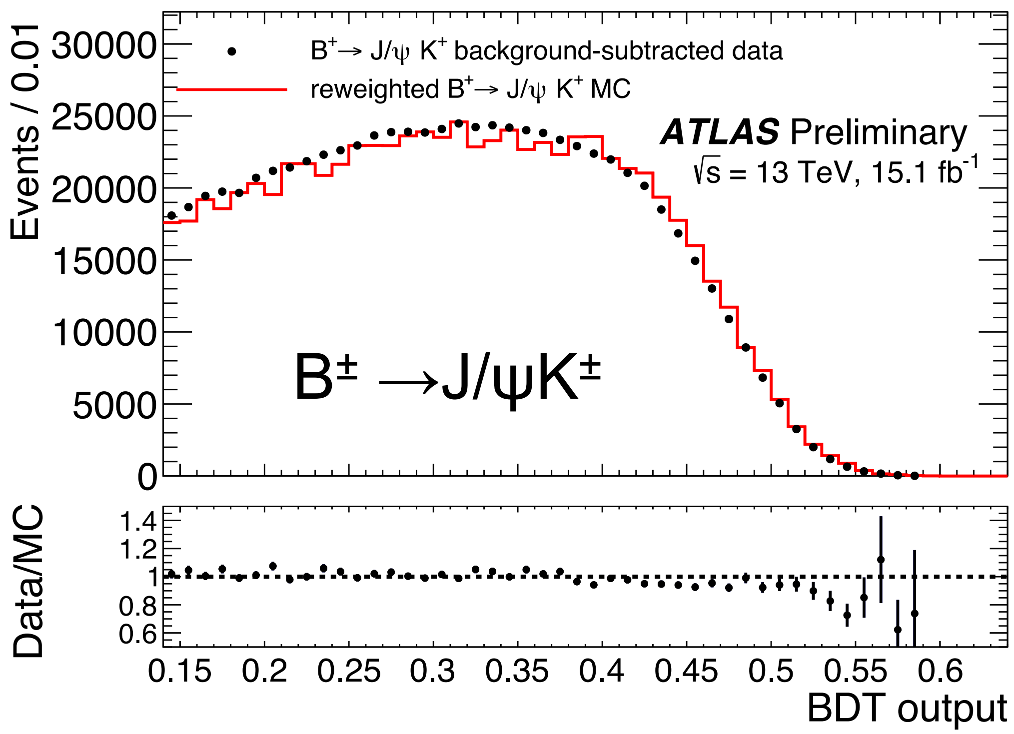
Reference and Control Channels

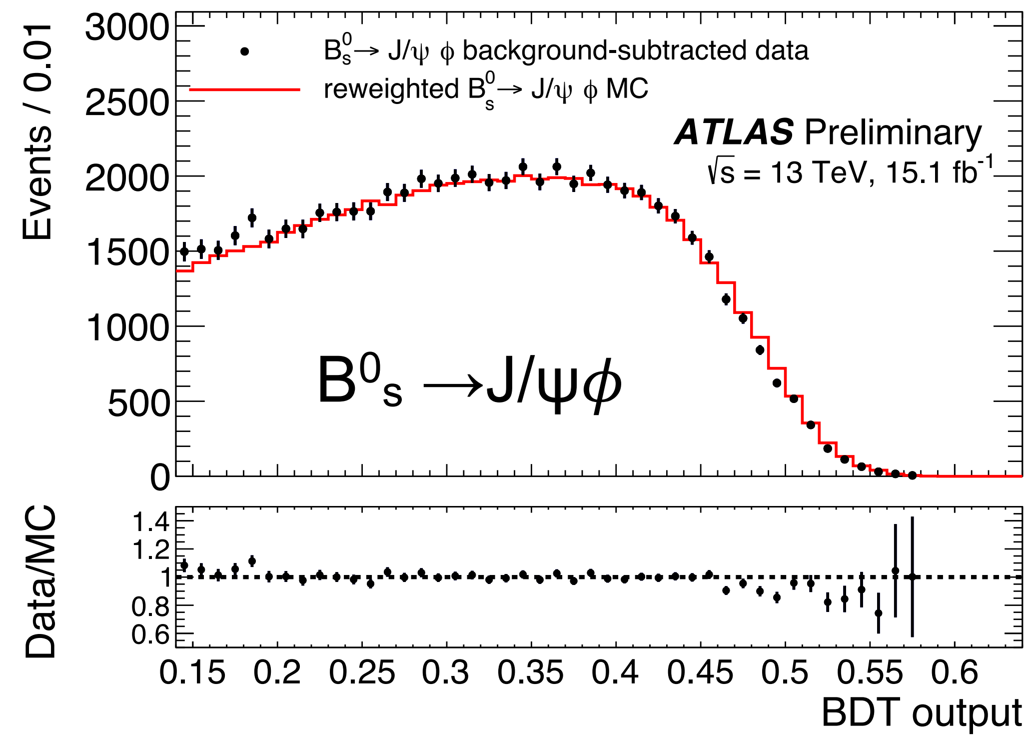
- $B^{\pm}$ yield from UML fit of $m_{\mu\mu K}$ distributions
  - shape parameters obtained from simultaneous fit to data and MC samples of sig. and bkg.
  - crosscheck:

    $N(B^{\pm} \to J/\psi\pi^{\pm})/N(B^{\pm} \to J/\psi K^{\pm}) = (3.71 \pm 0.09)\%$ agrees with world average: $(3.84 \pm 0.16)\%$
- data/MC discrepancies dominate systematic uncertainty on $\varepsilon_{\mu\mu}/\varepsilon_{\mu\mu K}$ ratio (3.2% out of the total 4.1%)

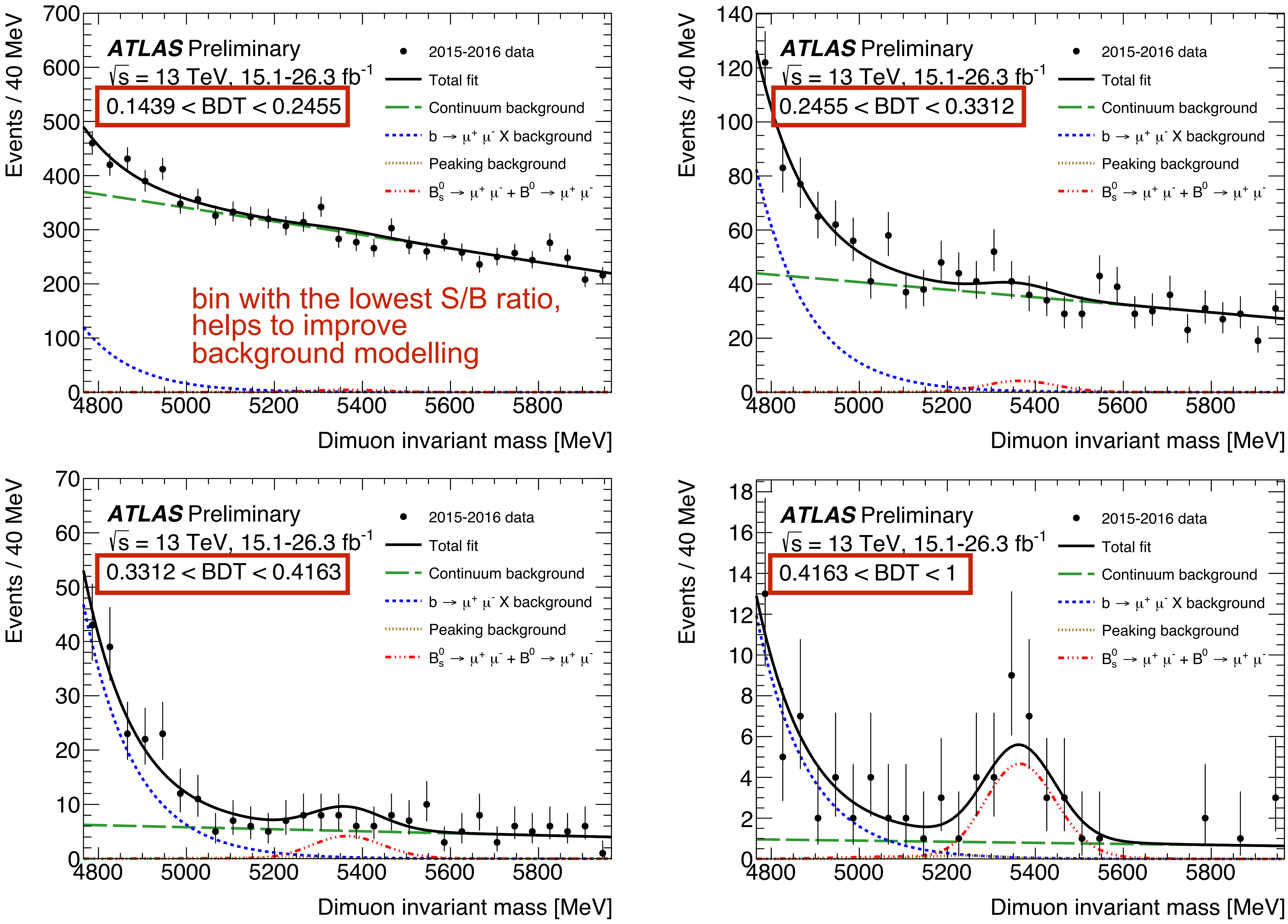


Data/MC agreement in the most sensitive BDT range:



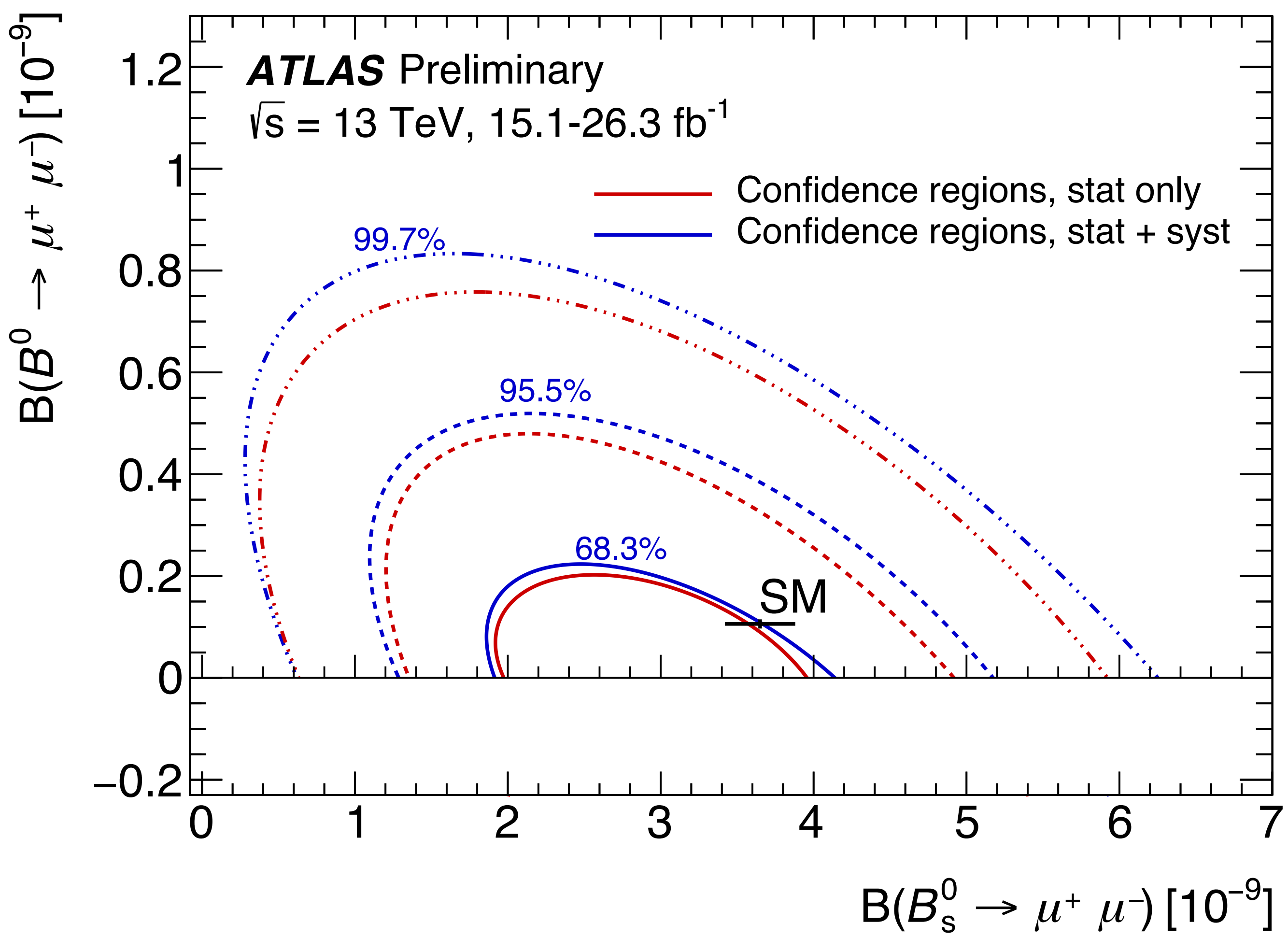


Signal Yield Extraction



Results with 2015/16 Data

- extracted (expected in SM) signal yields:
  - N_s = 80 ± 22 (91) and N_d = -12 ± 20 (10) events → consistent with expectations



- likelihood maximum:

$BR(B^0_s \to \mu^+\mu^-) = (3.2 \pm 0.9)\times10^{-9}$
$BR(B^0 \to \mu^+\mu^-) = (-1.3 \pm 2.1)\times10^{-10}$

- result from Neyman 2D contours:

$BR(B^0_s \to \mu^+\mu^-) = (3.21^{+0.96}_{-0.91}{}^{+0.49}_{-0.30})\times10^{-9}$
$BR(B^0 \to \mu^+\mu^-) < 4.3\times10^{-10}$ at 95% CL

- expected result in SM hypothesis:

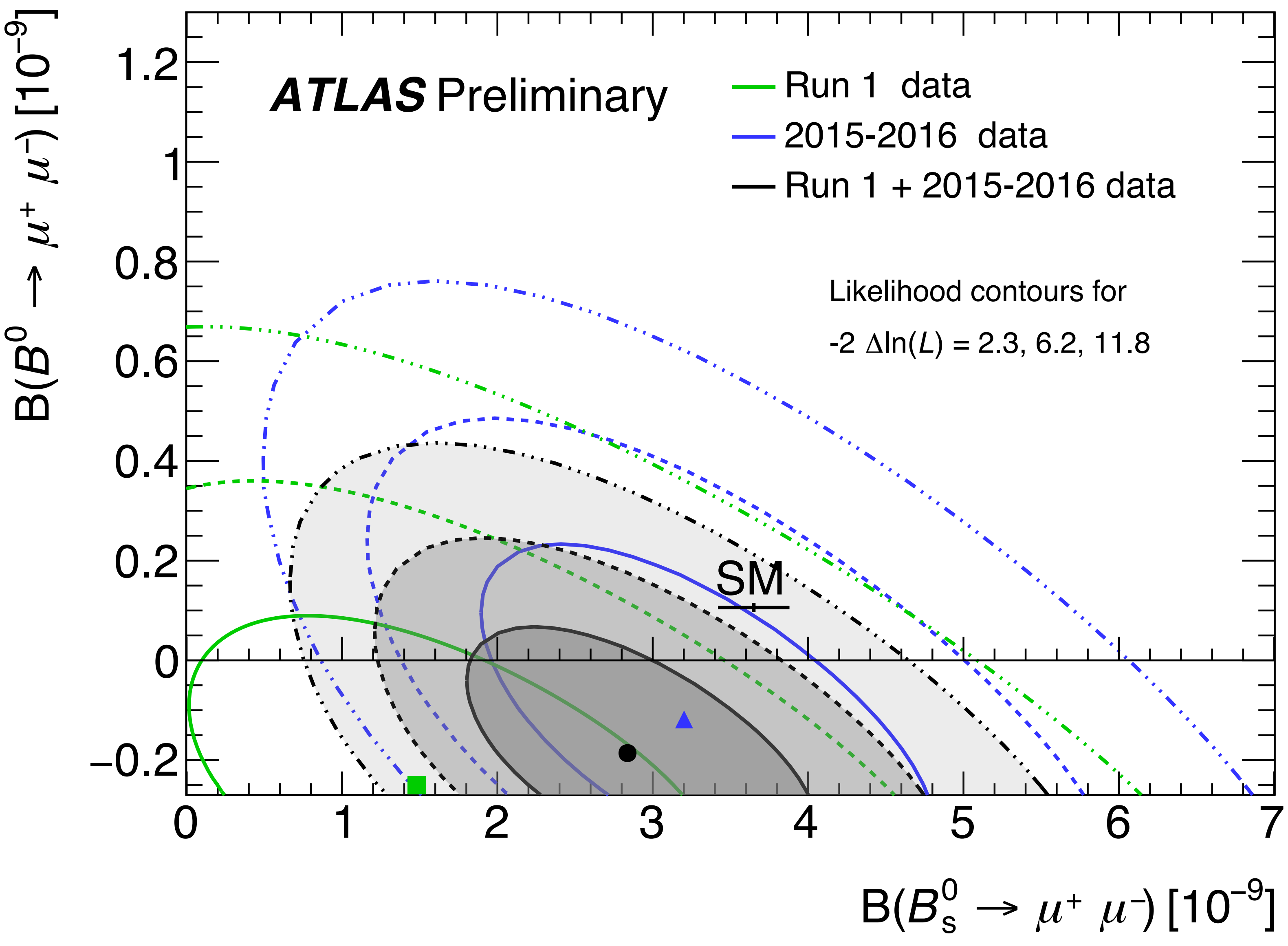
$BR(B^0_s \to \mu^+\mu^-) = (3.6^{+1.1}_{-1.0})\times10^{-9}$
$BR(B^0 \to \mu^+\mu^-) < 7.1\times10^{-10}$ at 95% CL

- statistical uncertainties dominate

Source	B^0_s (%)	B^0 (%)
f_s/f_d	5.1	-
B^+ Yield	4.8	4.8
R_ε	4.1	4.1
$\mathcal{B}(B^+ \to J/\psi K^+) \times \mathcal{B}(J/\psi \to \mu^+\mu^-)$	2.9	2.9
Fit Systematic Uncertainties	8.7	65
Stat. Uncertainty (from Likelihood est.)	27	150

[ATLAS-CONF-2018-046]

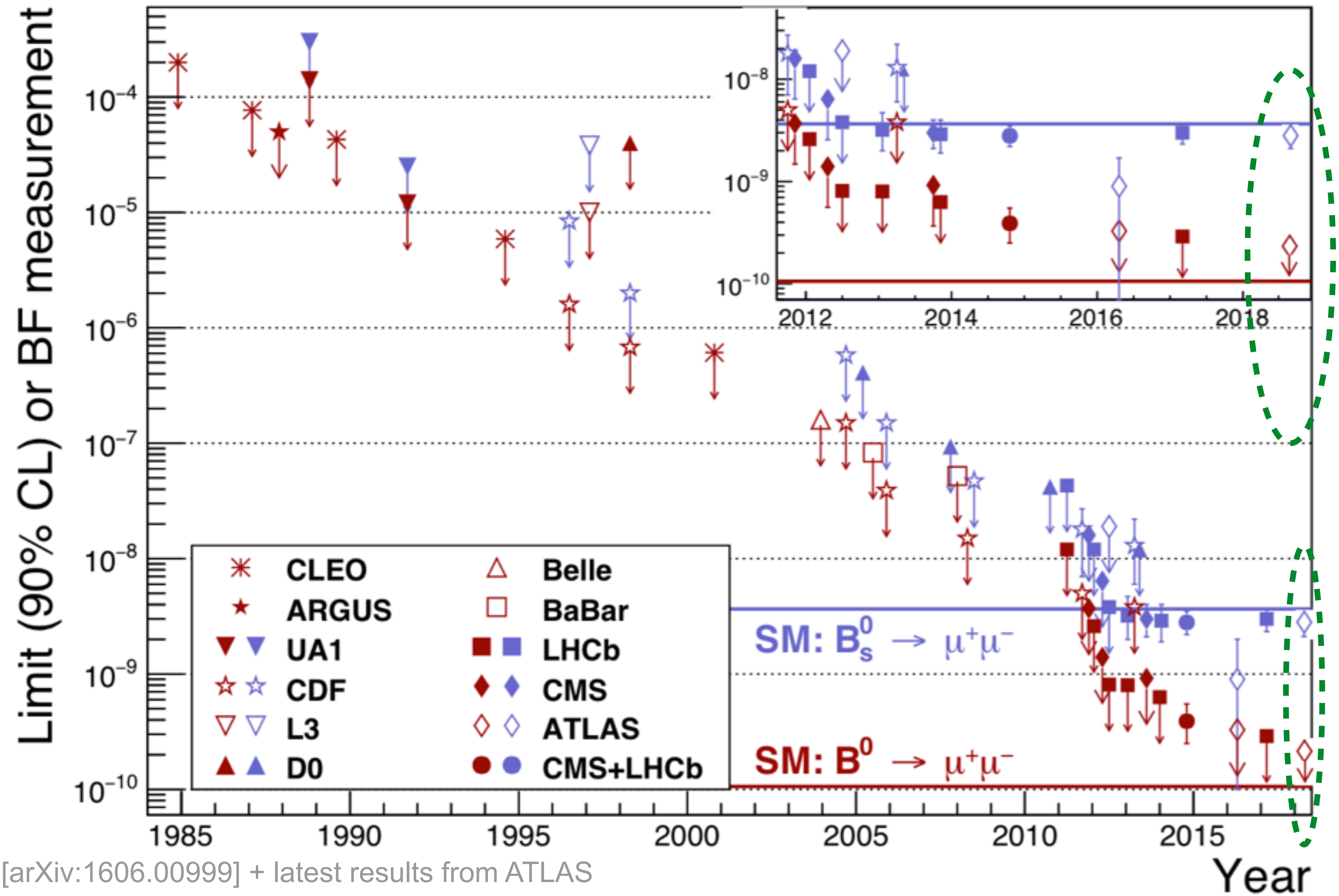
Combination with Run 1 Result



- likelihood maximum:
  $BR(B^0_s \to \mu^+\mu^-) = (2.8 \pm 0.7) \times 10^{-9}$
  $BR(B^0 \to \mu^+\mu^-) = (-1.9 \pm 1.6) \times 10^{-10}$
- result from likelihood contours:
  $BR(B^0_s \to \mu^+\mu^-) = (2.8^{+0.8}_{-0.7}) \times 10^{-9}$
  $BR(B^0 \to \mu^+\mu^-) < 2.1 \times 10^{-10}$ at 95% CL
- expected result in SM hypothesis:
  $BR(B^0_s \to \mu^+\mu^-) = (3.6^{+0.9}_{-0.8}) \times 10^{-9}$
  $BR(B^0 \to \mu^+\mu^-) < 5.6 \times 10^{-10}$ at 95% CL

- compatible with SM at ~2.4σ
- combined significance for $B^0_s \to \mu^+\mu^-$ ~4.6σ

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



[arXiv:1606.00999] + latest results from ATLAS

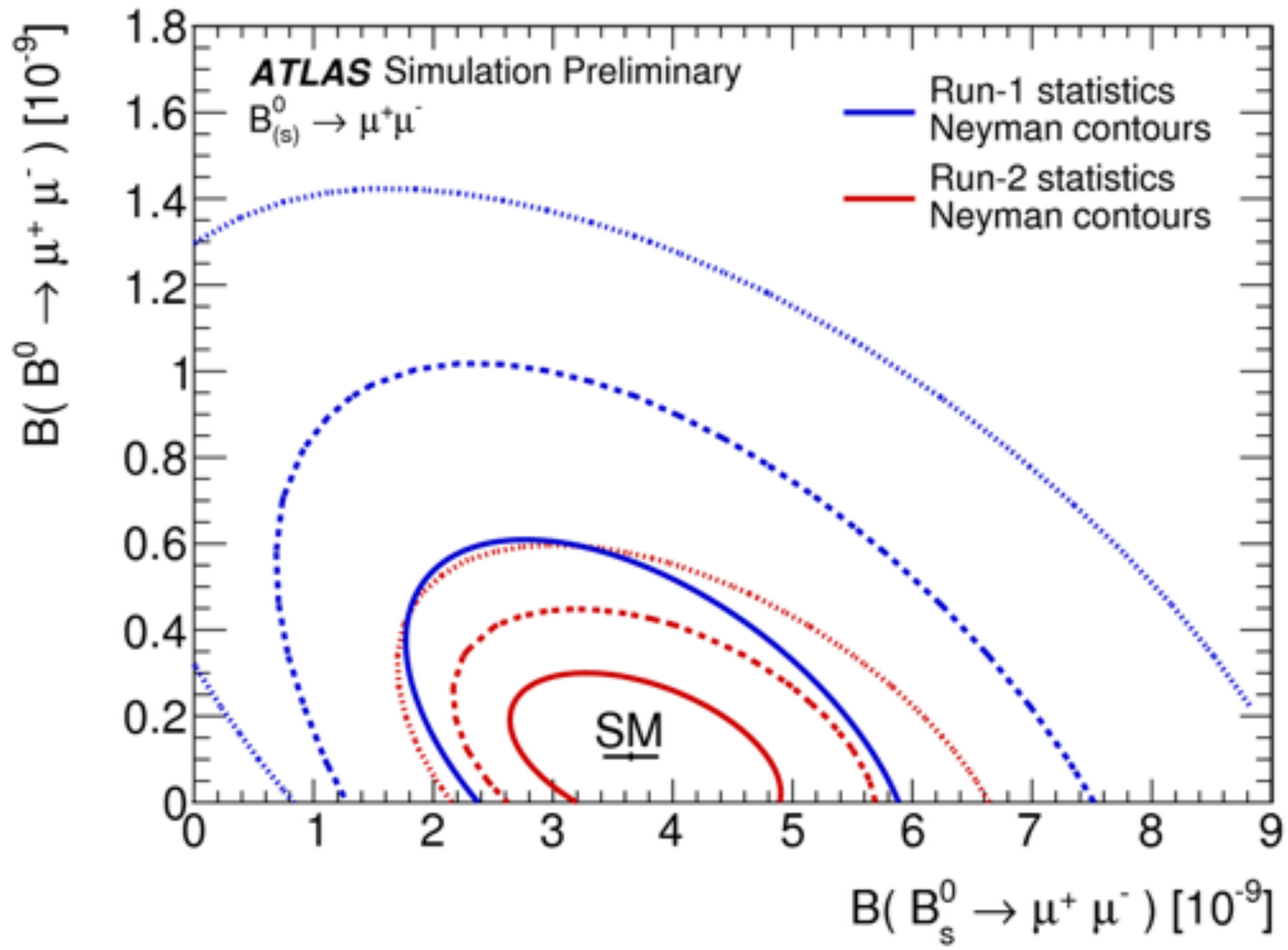
Prospects of $B^0_{(s)} \rightarrow \mu^+\mu^-$ for LHC Run 2 and HL-LHC

Analysis Prospects with Full Run 2 Data (~130 fb^{-1})

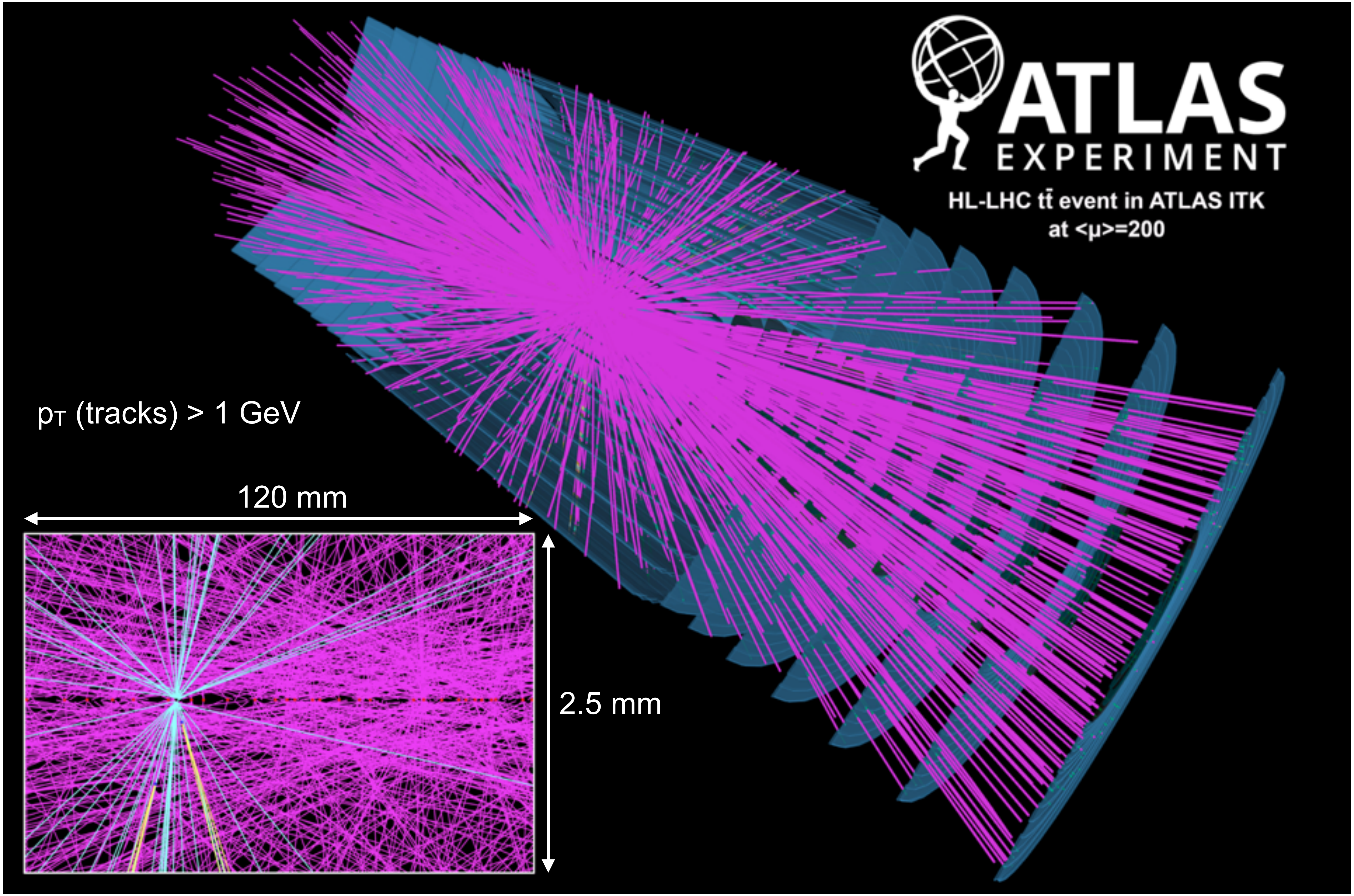
- pseudo-MC experiments **based on Run 1** measurement
- 2D Neyman belt construction to estimate CL contours
  - total statistics is scaled in the likelihood
    - σ_{bb} ~1.7x Run 1 (8 TeV ⟶ 13 TeV)
    - 2MU6 || MU6_MU4 topological triggers
    - ⇒ estimated to ~7x Run 1
  - same S/B ratio as in Run 1 - "conservative" as background suppression expected to improve due to IBL

Systematic uncertainties:
(30% of Run 1 total error)

- external: f_s/f_d, BR($B^\pm \to J/\psi K^\pm$)
  - kept as in Run 1
- internal: efficiencies, fit shapes, background extrapolation, trigger modelling,...
  - scaled with statistics

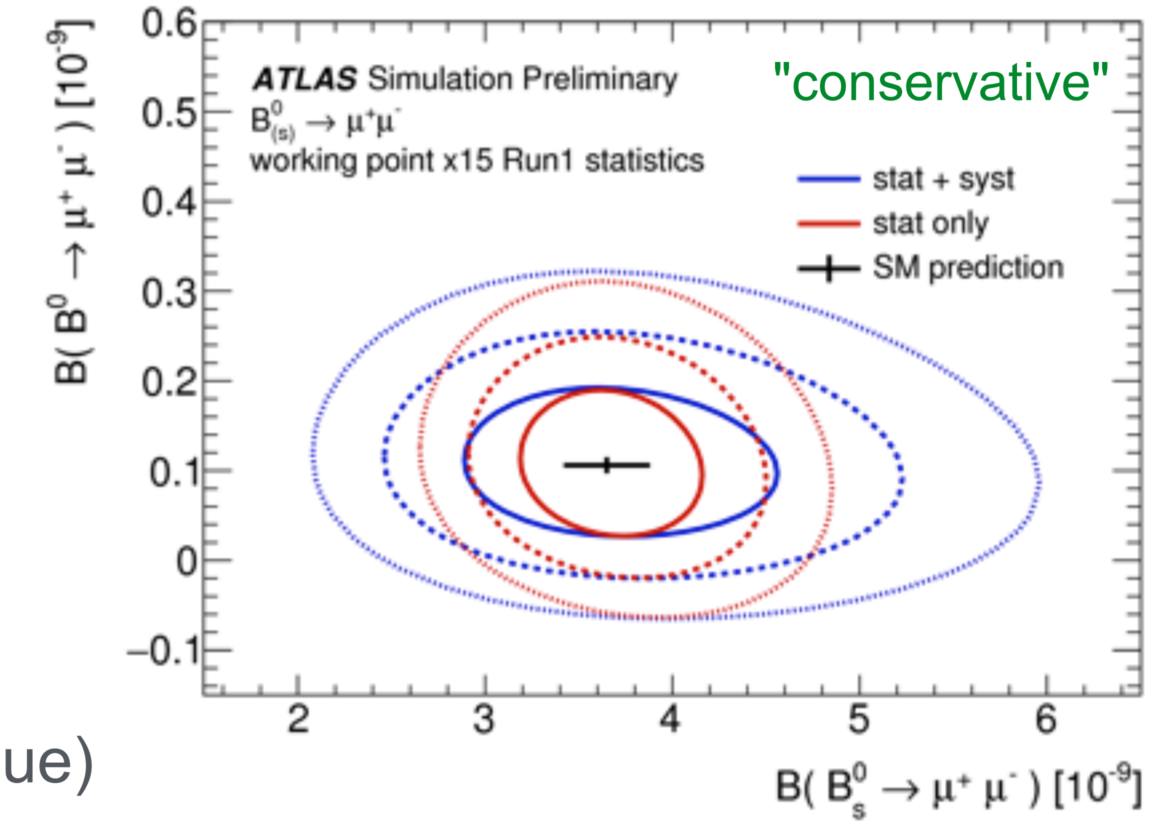


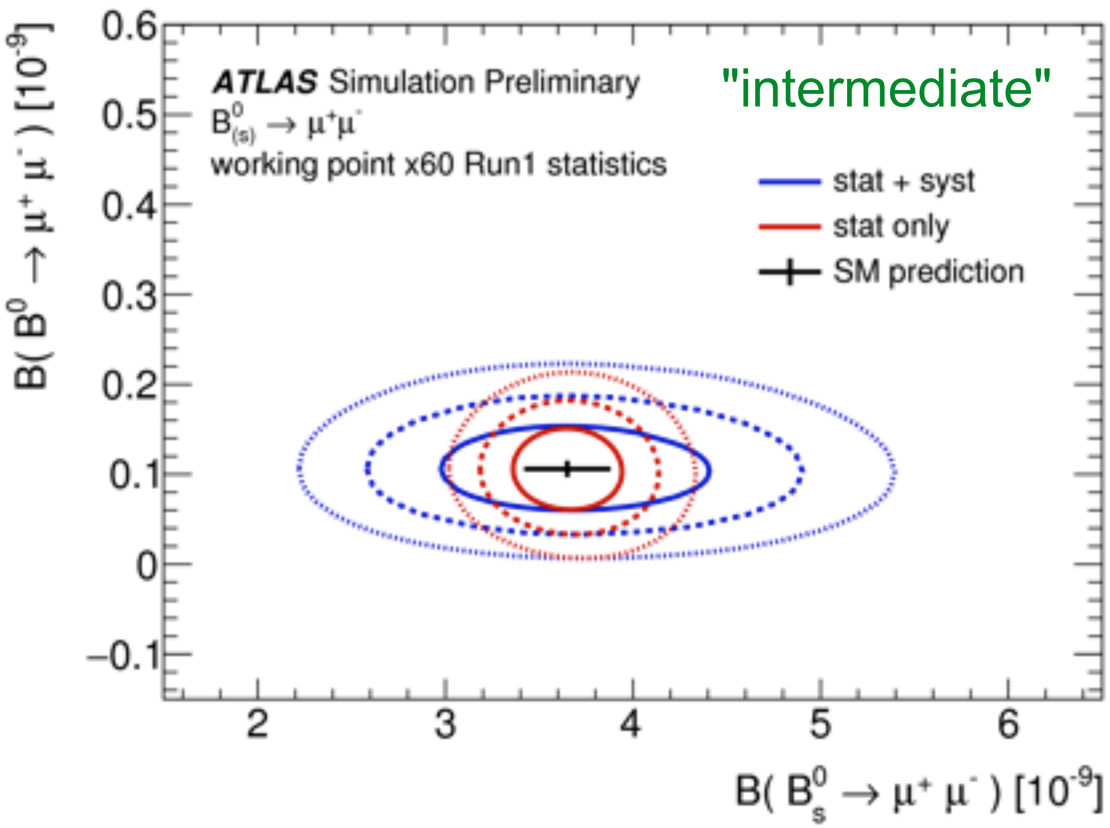
HL-LHC

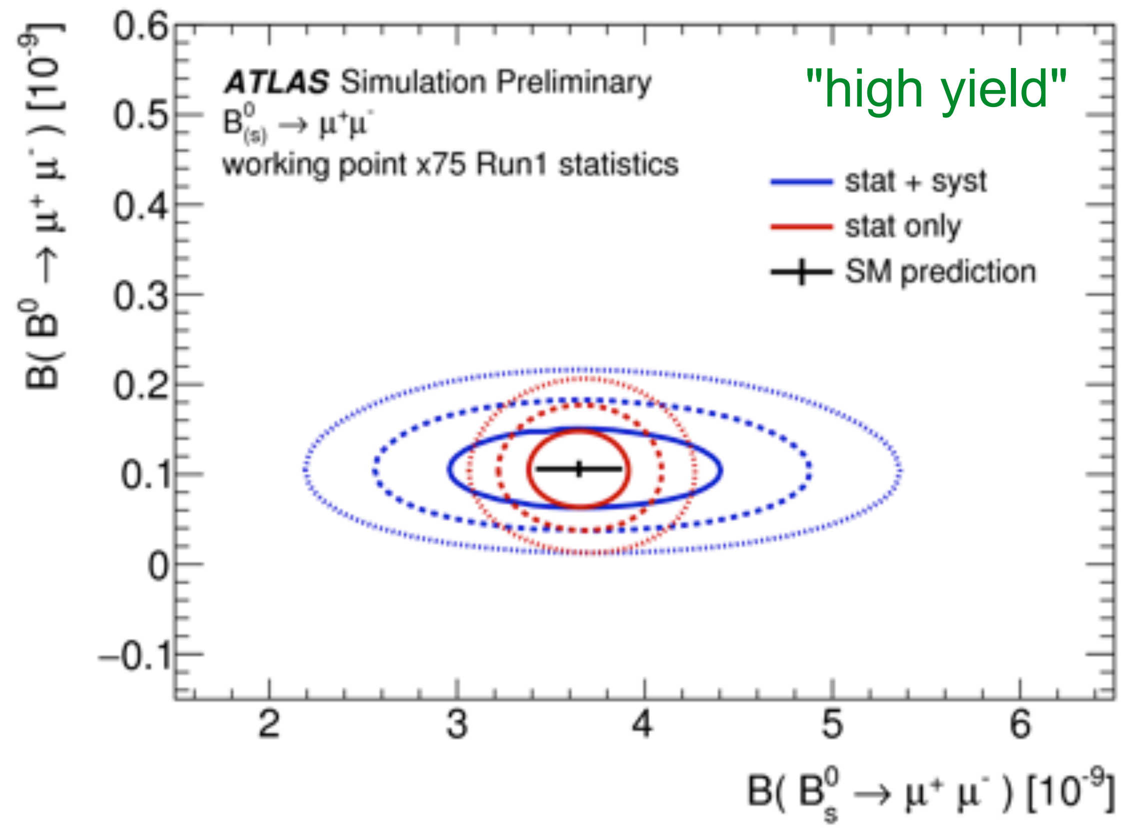

ATLAS
EXPERIMENT
HL-LHC $t\bar{t}$ event in ATLAS ITK
at <μ>=200
p_T (tracks) > 1 GeV
120 mm
2.5 mm

HL-LHC Prospects (3 ab^{-1})

- better $\sigma(m_{\mu\mu})$ expected → improvement of $m(B^0_s)$ and $m(B^0)$ separation
- 3 trigger scenarios considered:
  - 2MU10 ⟶ 15 x N(Run1)
  - MU6_MU10 ⟶ 60 x N(Run1)
  - 2MU6 ⟶ 75 x N(Run1)
- pseudo-MC experiments based on Run 1 likelihood
- "conservative" systematics
  - dominant $\sigma(f_s/f_d)$ ~ 8.3% (yet Run 1 value)







Summary & Outlook

- $B^0_{(s)} \to \mu^+\mu^-$ analysis with 2015/16 data presented
  - $BR(B^0_s \to \mu^+\mu^-) = (3.2^{+1.1}_{-1.0})\times10^{-9}$
  - $BR(B^0 \to \mu^+\mu^-) < 4.3 \times10^{-10}$ at 95% CL
- combination with Run 1 result yields
  - $BR(B^0_s \to \mu^+\mu^-) = (2.8^{+0.8}_{-0.7})\times10^{-9}$
  - $BR(B^0 \to \mu^+\mu^-) < 2.1 \times10^{-10}$ at 95% CL
    - results compatible with SM at 2.4σ
- reach for expected full Run 2 and HL-LHC statistics presented
- Run 2 *pp* data-taking just finished - stay tuned for new results!

BACKUP

BDT Input Variables

Variable	Description
p_{T}^{B}	Magnitude of the B candidate transverse momentum $\overrightarrow{p_{\mathrm{T}}}^{B}$.
$\chi^2_{\mathrm{PV,DV}\ xy}$	Compatibility of the separation $\overrightarrow{\Delta x}$ between production (*i.e.* associated PV) and decay (DV) vertices in the transverse projection: $\overrightarrow{\Delta x}_{\mathrm{T}} \cdot \Sigma^{-1}_{\overrightarrow{\Delta x}_{\mathrm{T}}} \cdot \overrightarrow{\Delta x}_{\mathrm{T}}$, where $\Sigma_{\overrightarrow{\Delta x}_{\mathrm{T}}}$ is the covariance matrix.
ΔR	three-dimensional opening between $\overrightarrow{p}^{B}$ and $\overrightarrow{\Delta x}$: $\sqrt{\alpha_{2\mathrm{D}}{}^2 + \Delta\eta^2}$
$\lvert\alpha_{2\mathrm{D}}\rvert$	Absolute value of the angle between $\overrightarrow{p_{\mathrm{T}}}^{B}$ and $\overrightarrow{\Delta x}_{\mathrm{T}}$ (transverse projection).
L_{xy}	Projection of $\overrightarrow{\Delta x}_{\mathrm{T}}$ along the direction of $\overrightarrow{p}^{B}_{\mathrm{T}}$: $(\overrightarrow{\Delta x}_{\mathrm{T}} \cdot \overrightarrow{p_{\mathrm{T}}}^{B})/\lvert\overrightarrow{p_{\mathrm{T}}}^{B}\rvert$.
$\mathrm{IP}_{B}^{3\mathrm{D}}$	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.
$\lvert d_0\rvert^{\max}$-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.
$P_{\mathrm{L}}^{\min}$	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 $\lvert\overrightarrow{p_{\mathrm{T}}}^{B}\rvert$ to the sum of $\lvert\overrightarrow{p_{\mathrm{T}}}^{B}\rvert$ and of the transverse momenta of all additional tracks contained within a cone of size $\Delta R < 0.7$ around the B direction. Only tracks matched to the same PV as the B 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.
$N_{\mathrm{xtrk}}^{\mathrm{close}}$	Number of additional tracks compatible with the decay vertex (DV) of the B candidate with $\ln(\chi^2_{\mathrm{xtrk,DV}}) < 1$. The tracks matched to a PV different from the B candidate are excluded.
$\chi^2_{\mu,\mathrm{xPV}}$	Minimum χ^2 for the compatibility of a muon in the B candidate with any PV reconstructed in the event.

[ATLAS-CONF-2018-046]

Systematic uncertainties on $\varepsilon_{\mu\mu}/\varepsilon_{\mu\mu K}$ ratio

Source	Contribution (%)
Statistical	0.8
BDT Input Variables	3.2
Kaon Tracking Efficiency	1.5
Muon trigger and reconstruction	1.0
Kinematic Reweighting (DDW)	0.8
Pile-up Reweighting	0.6

[ATLAS-CONF-2018-046]

Signal Fit Model

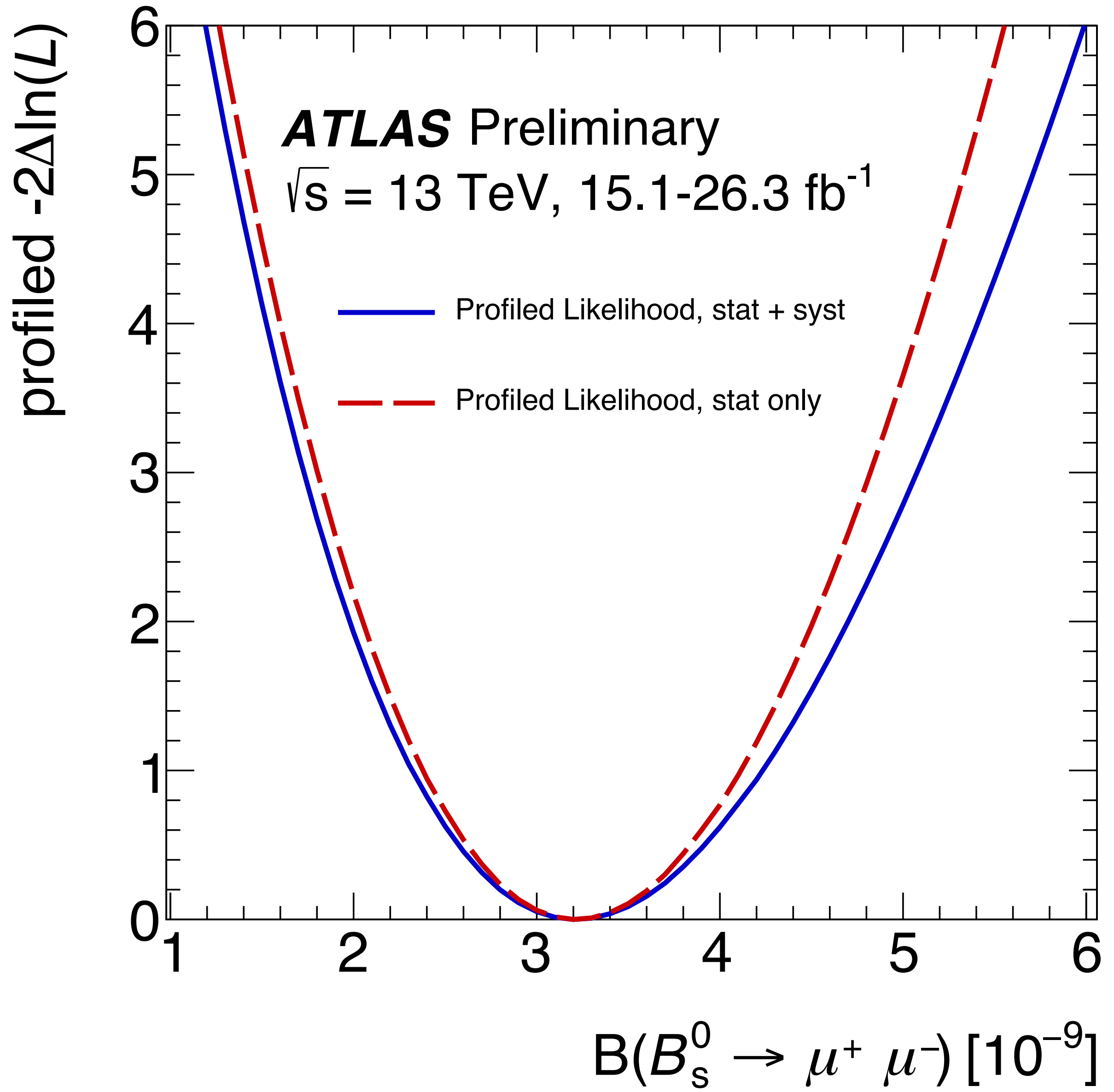
- $m_{\mu\mu}$ distributions split into 4 BDT intervals (bins), each with $\varepsilon_{sig} = 18\%$
  - N_s and N_d extracted simultaneously in 3 bins with unbinned extended ML fit
  - 4th bin with the lowest S/B ratio mainly for background control

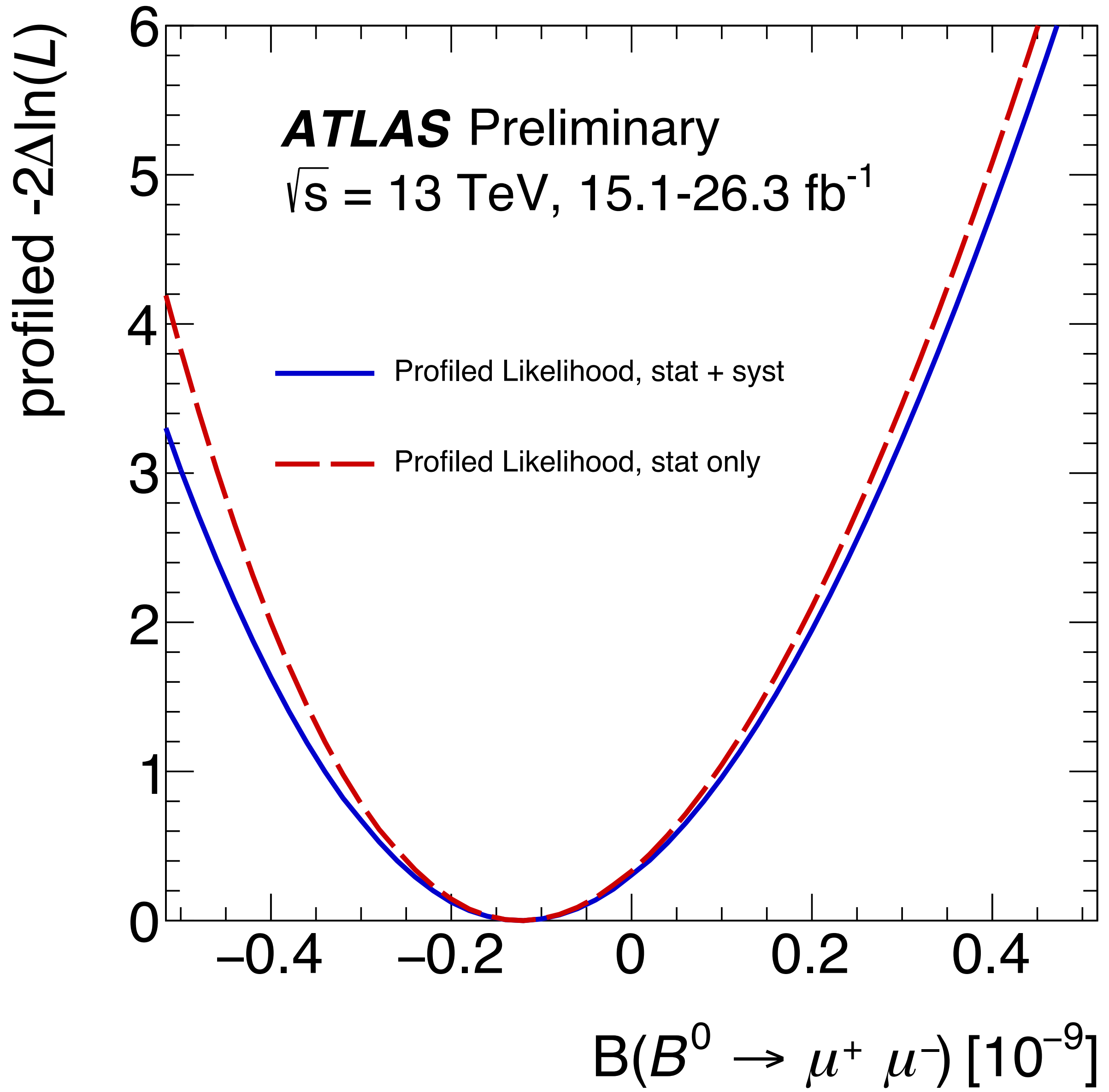
- fit model
  - signals: double-Gaussians with common mean each
  - continuum background: linear in $m_{\mu\mu}$, minimal correlation between $m_{\mu\mu}$ and BDT, shape and normalisation taken from sideband data
  - $b \rightarrow \mu\mu X$ background: exponential in $m_{\mu\mu}$, determined from sideband data
  - peaking background: double-Gaussian, same mean, equal shape and amplitude in BDT bins

- fit systematics $\sigma^{N_s}_{\text{syst}} = 3 + 0.05 N_s$ and $\sigma^{N_d}_{\text{syst}} = 2.9 + 0.05 N_s + 0.05 N_d$

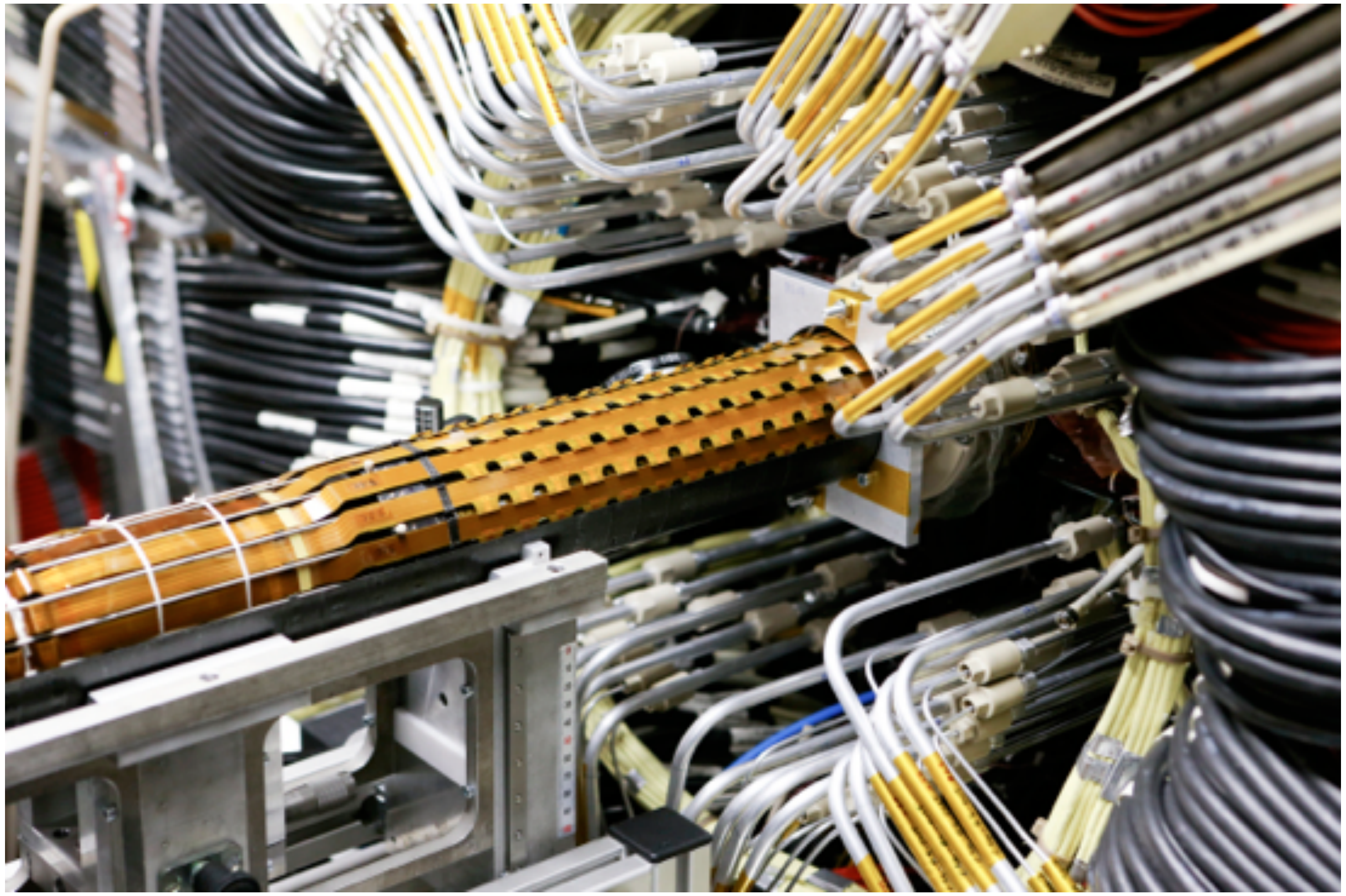
  most systematic shifts for N_s and N_d are correlated: $\rho_{\text{syst}} = -0.83$

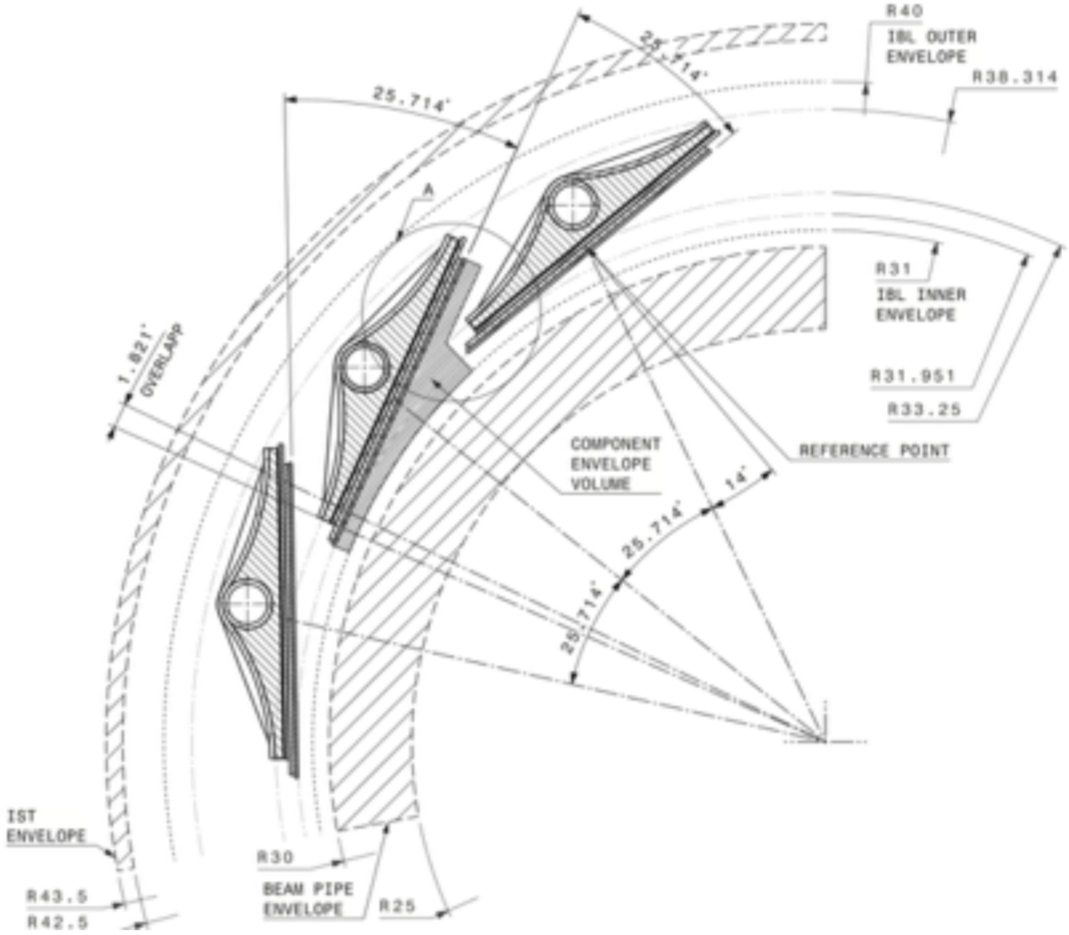
Profiled Likelihood Scan





Tracker Upgrade for Run 2 (IBL)

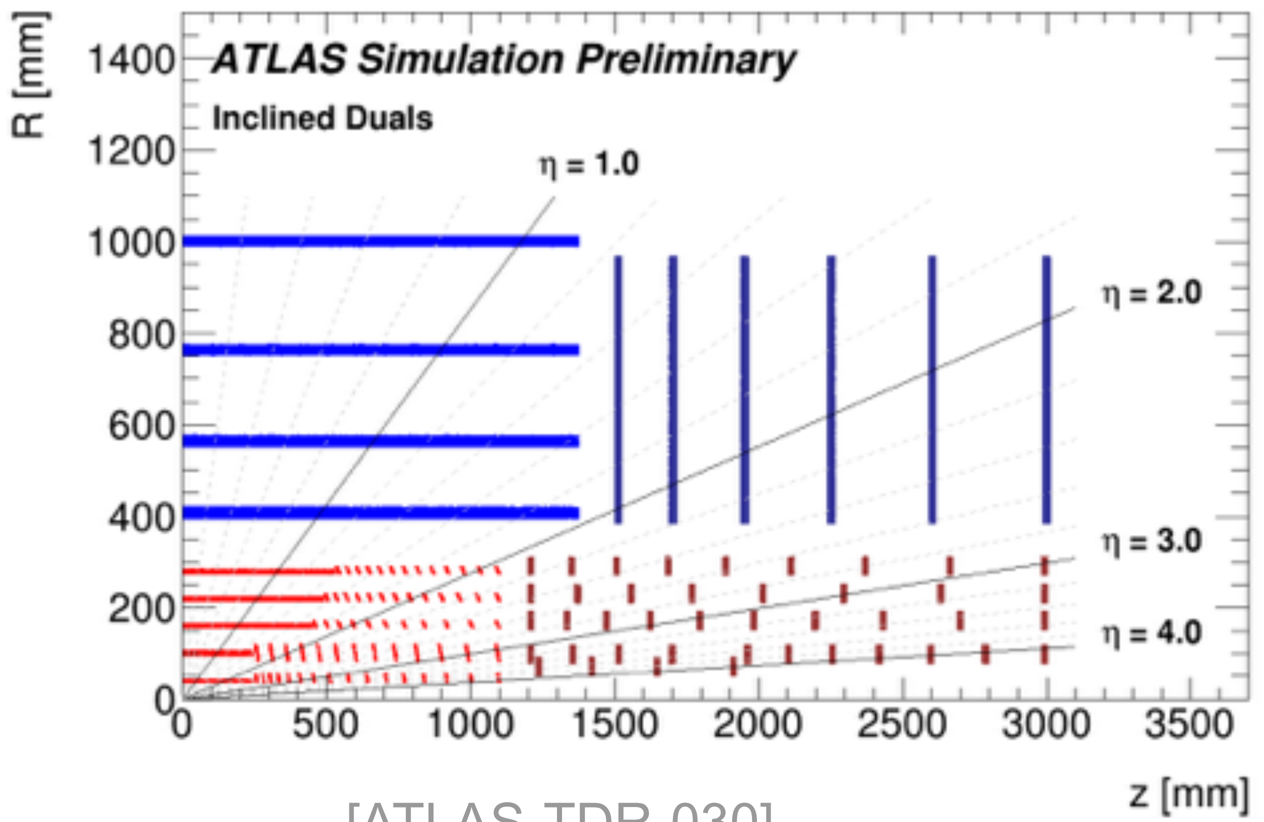




- additional innermost pixel layer at ~33 mm
- improvement in impact parameter resolution for low momentum tracks

  $\Longrightarrow$ beneficial for B-Physics

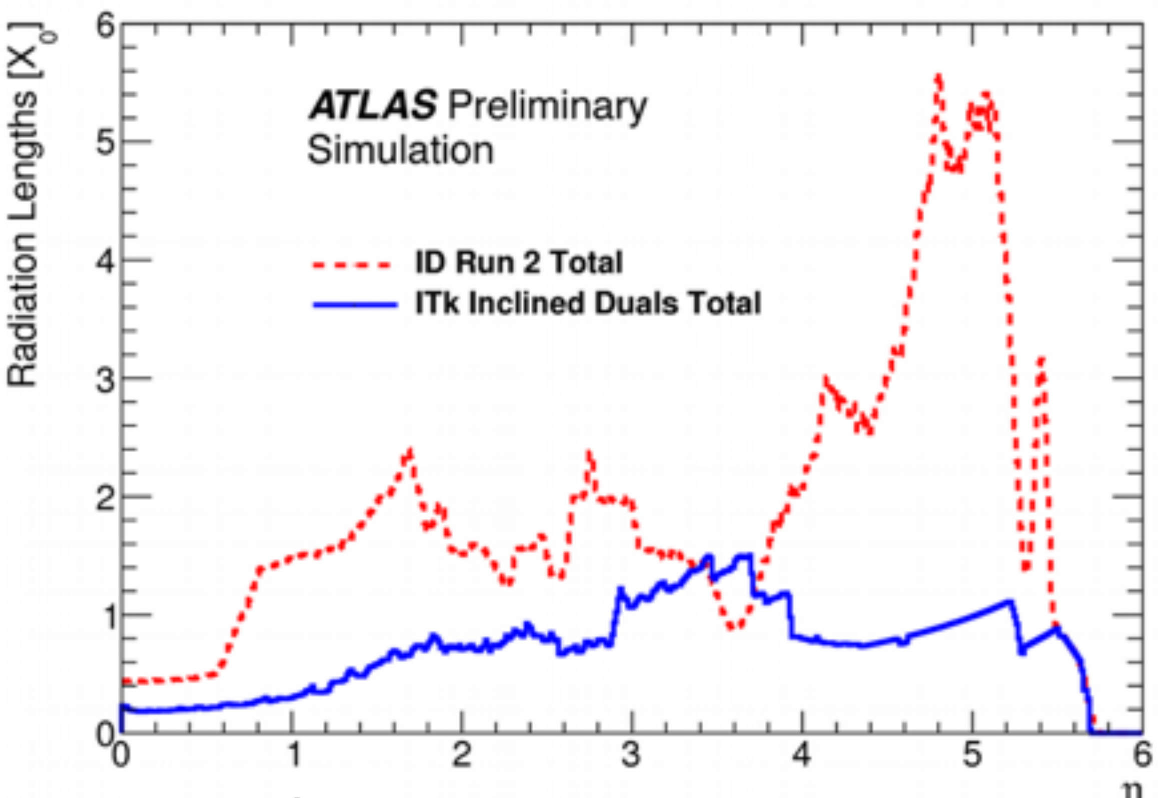
ATLAS Inner Tracker (ITk) Upgrade



[ATLAS-TDR-030]

- ITk Pixel (13 m^2):
  - 5 barrel, 5 EC layers (with rings)
  - inclined design, $|\eta_{max}| < 4$ (2.5 now)
  - innermost layer at 36 mm
  - ~580 M channels (~92 M now)
- ITk Strips (160 m^2):
  - 4 barrel layers, 6 EC rings
  - ~50M channels (6M now)

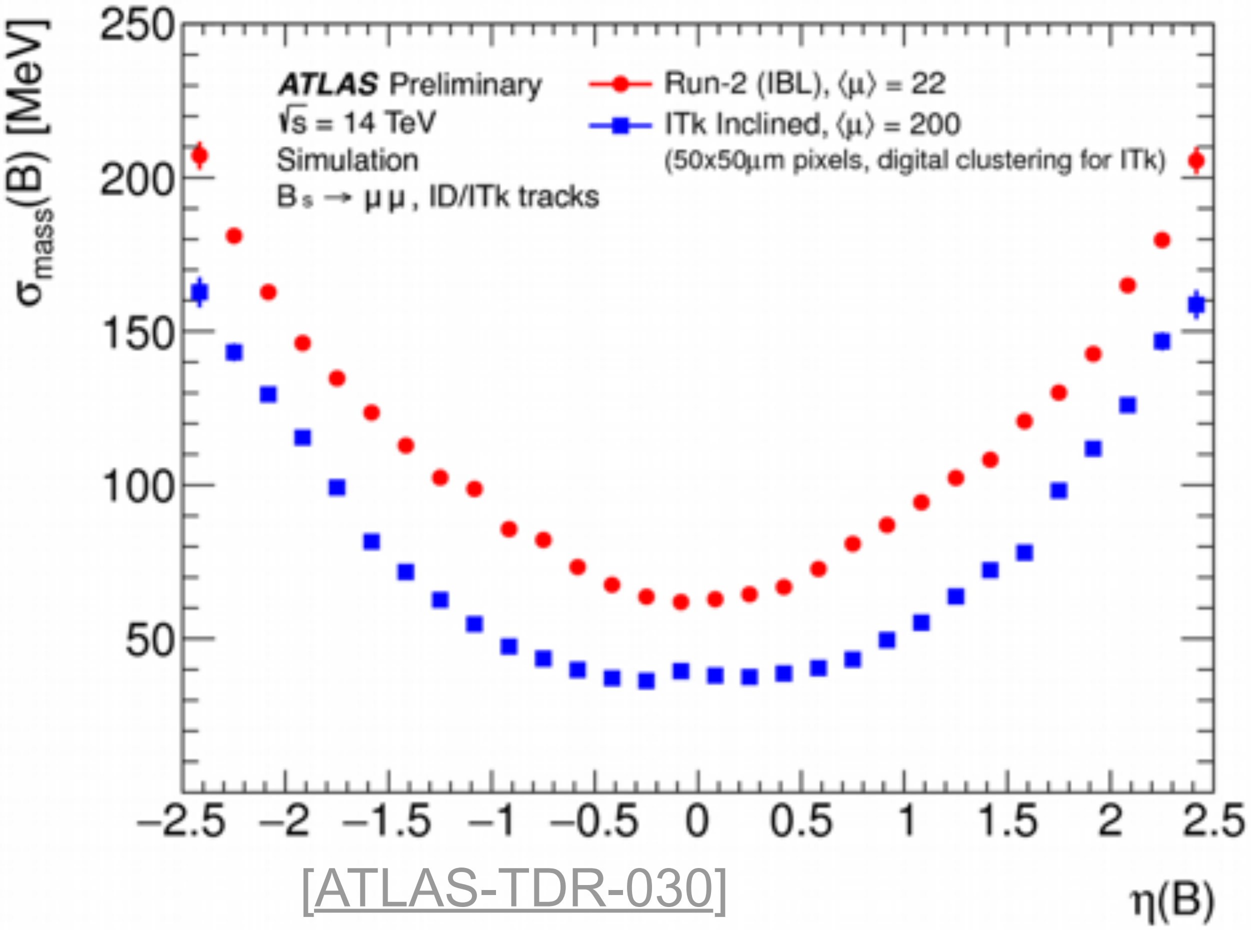
- ITk material considerably less than current ID:
  - better tracking efficiency
  - better mass resolution



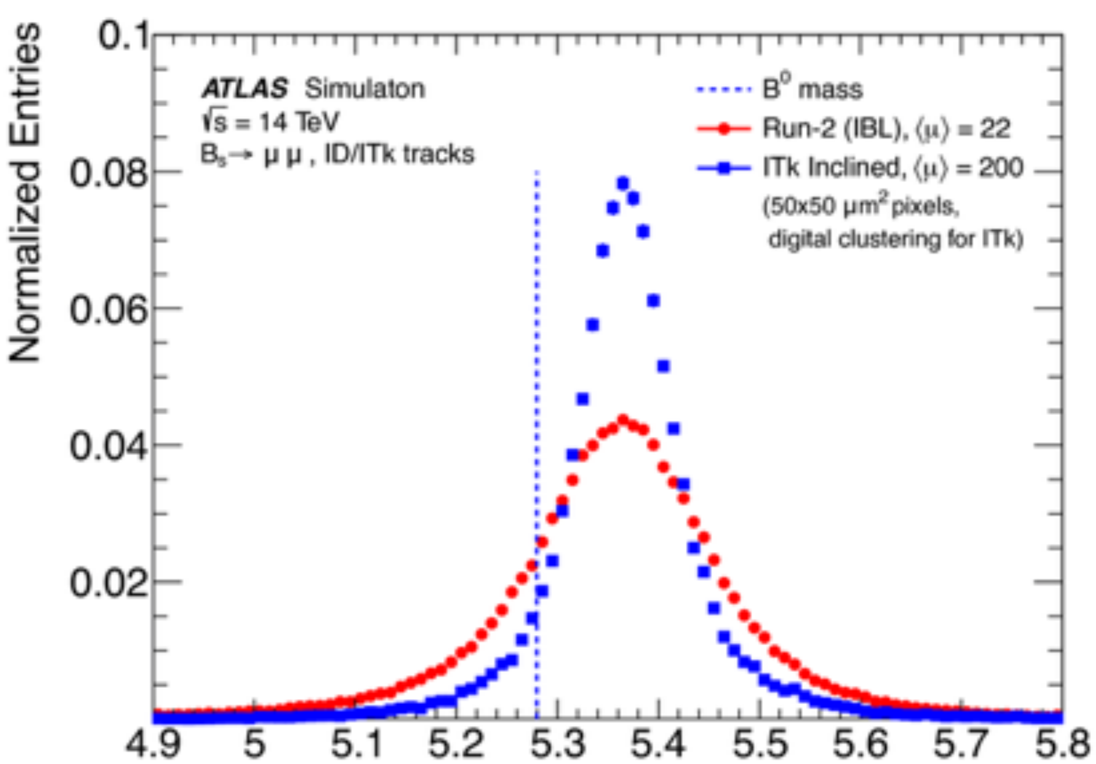
[ATLAS-TDR-030]

Mass Resolution with HL-ATLAS

- dedicated $B^0_s \to \mu^+\mu^-$ MC:
  - Run 2 (2015) conditions
  - new tracker: ITk, inclined design, $|\eta| < 4$, 50x50 μm^2 pixels
- candidate selection ~Run 1:
  - B^0_s from $\mu^+\mu^-$ pairs with $p_T(\mu^\pm) >$ 5.5 GeV
  - two-track vertex fit
  - $m(B^0_s)$ from ITk-only tracks



[ATLAS-TDR-030]



[ATLAS-TDR-030]

improvement of $m(B^0_s)$ and $m(B^0)$ separation:

- barrel by x 1.65:
  - 1.4σ (Run 1) ⟶ 2.3σ
- end-caps by x ~1.5:
  - 0.85σ (Run 1) ⟶ 1.3σ

[ATL-PHYS-PUB-2016-026]