Leptonic decays of light states and rare B decays

Aidan Grummer
University of New Mexico

on behalf of the ATLAS and CMS Collaborations

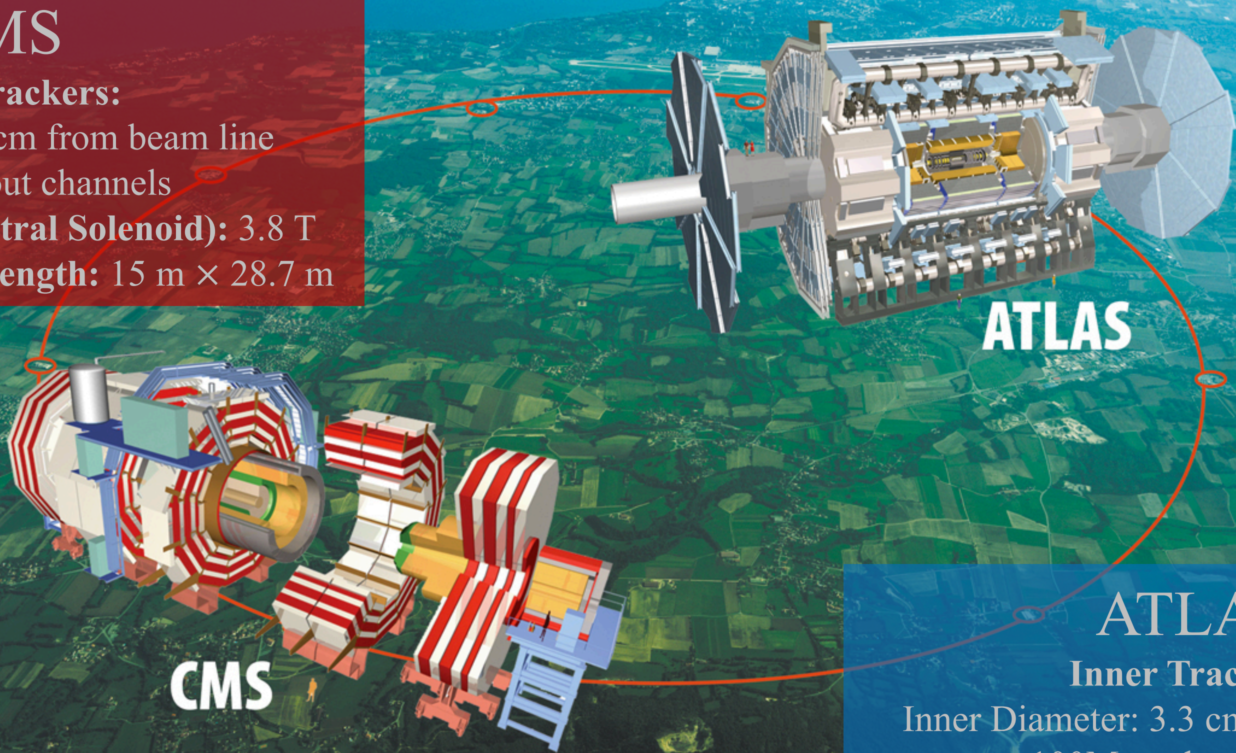
LHCP
May 26, 2020

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

$\tau \rightarrow 3\mu$

Introduction

Analysis of leptonic decays from ATLAS and CMS



CMS

Inner Trackers:

Inner Diameter: 4 cm from beam line

~80M readout channels

Magnetic Field (Central Solenoid): 3.8 T

Outer Diameter and length: 15 m × 28.7 m

ATLAS

Inner Trackers:

Inner Diameter: 3.3 cm from beam line

~100M readout channels

Magnetic Field (Central Solenoid): 2 T

Outer Diameter and length: 25 m × 46 m

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

- ATLAS: JHEP 04 (2019) 098

https://doi.org/10.1007/JHEP04(2019)098

26.3 fb^{-1} of $\sqrt{s} = 13$ TeV (2015 and 2016)

25 fb^{-1} of $\sqrt{s} = 7$ and 8 TeV (2011 and 2012)

- CMS: JHEP 04 (2020) 188

https://doi.org/10.1007/JHEP04(2020)188

36 fb^{-1} of $\sqrt{s} = 13$ TeV* (2016A and 2016B)

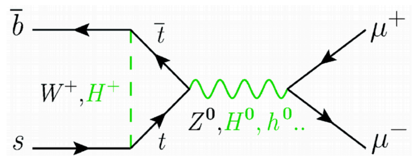
20 fb^{-1} of $\sqrt{s} = 8$ TeV (2012)

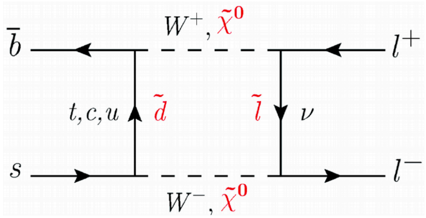
5 fb^{-1} 7 $\sqrt{s} = 7$ TeV (2011)

*Due to operational instabilities experienced with the CMS microstrip detector, CMS Run 2 data are divided into two separate running periods, denoted 2016A and 2016B. Data are further separated into the forward and central regions of the detector.

Motivation for Measurement of $B^0_{(s)} \to \mu^+\mu^-$

- The smallness and precision of the predicted branching fractions* provides a favorable environment for observing contributions from new physics
  - $\mathcal{B}(B^0_s \to \mu^+\mu^-) = (3.66 \pm 0.14) \times 10^{-9}$
  - $\mathcal{B}(B^0 \to \mu^+\mu^-) = (1.03 \pm 0.05) \times 10^{-10}$
- Probe the Standard Model, which predicts that only the heavy mass eigenstate contributes to the $B^0_s \to \mu^+\mu^-$ effective lifetime, $\tau_{\mu^+\mu^-}$
  - Experimental World Average from PDG**: $\tau_{\mathrm{B^0_{sH}}} = 1.615 \pm 0.009$ ps
- Significant deviations could arise in models involving non-SM heavy particles such as those predicted in
  - Minimal Supersymmetric Standard Model***
  - Minimal Flavor Violation †
  - Two Higgs-Doublet Models ‡





"New Physics"

* M. Beneke, C. Bobeth and R. Szafron, "Power-enhanced leading-logarithmic QED corrections to $B_q \to \mu^+\mu^-$," JHEP 10 (2019) 232 [arXiv:1908.07011].

** Particle Data Group collaboration, "Review of particle physics," Phys. Rev. D 98 (2018) 030001.

*** Huang, Chao-Shang and Liao, Wei and Yan, Qi-Shu, "Promising process to distinguish supersymmetric models with large tan β from the standard model: $B \to X_s$ μ+ μ−," Phys. Rev. D 59 (1998) 011701, arXiv: hep-ph/9803460 [hep-ph].

† G. D'Ambrosio, G. F. Giudice, G. Isidori and A. Strumia, "Minimal flavor violation: an effective field theory approach," Nucl. Phys. B 645 (2002) 155, arXiv: hep-ph/0207036 [hep-ph].

‡ K. S. Babu and C. F. Kolda, "Higgs mediated $B^0 \to$ μ+ μ− in minimal supersymmetry," Phys. Rev. Lett. 84 (2000) 228, arXiv: hep-ph/9909476 [hep-ph].

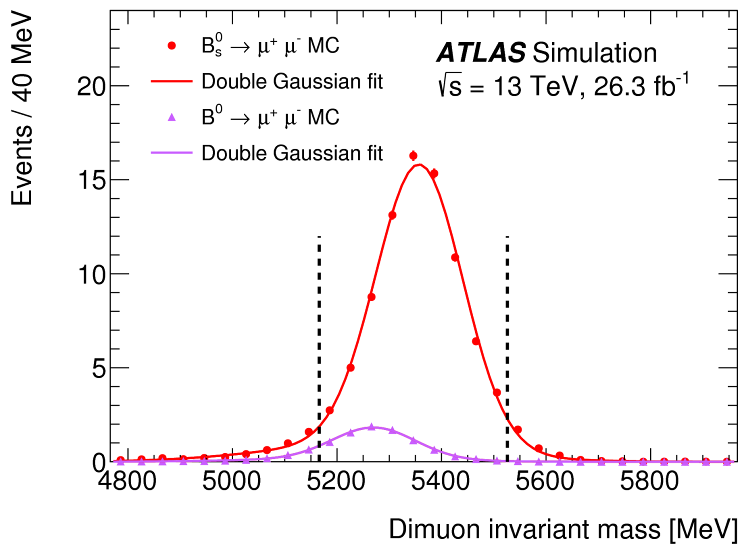
Branching Fraction Measurement

- The aim is to obtain the branching fraction of the $B^0_{(s)} \to \mu^+\mu^-$ channels
  - Utilize a reference channel: $B^+ \to J/\psi K^+$ which is abundant and has a well measured branching fraction

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

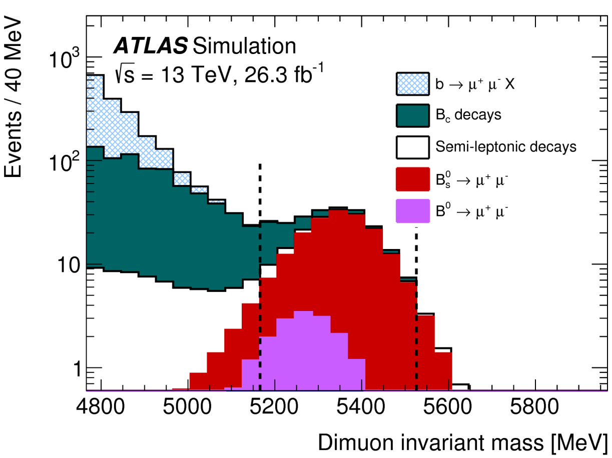
  - Here $N_{d(s)}$ is the signal yield, $N_{J/\psi K^+}$ is the reference yield, $\varepsilon_{\mu^+\mu^-}$ and $\varepsilon_{J/\psi K^+}$ are the acceptance times the efficiencies and $f_u/f_{d(s)}$ is the ratio of the hadronization probabilities of a b-quark into $B^+$and $B^0_{(s)}$.

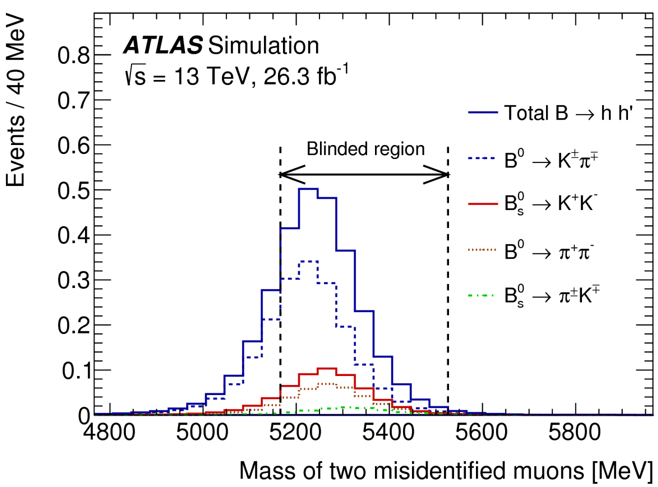
- Perform a **blind analysis**
  - Conceal the signal region of the dimuon invariant mass while procedures of the event selection and signal extraction are defined
  - ATLAS: $m_{\mu\mu}$ in [5166, 5526] MeV
  - CMS: $m_{\mu\mu}$ in [5200, 5450] MeV
- MC simulated samples
  - Dimuon events – for signal and background regions
  - $B^+ \to J/\psi K^+$candidates (reference channel)

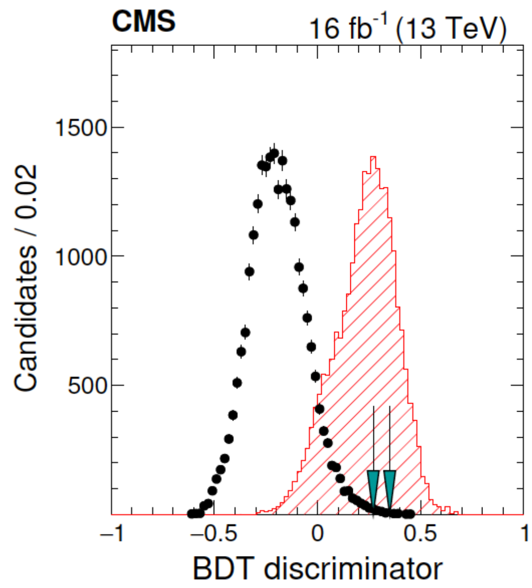


Background Composition

- Continuum background: the dominant combinatorial component
  - Consists of muons from uncorrelated hadron decays
  - The background distribution is characterized by a weak dependence on the dimuon invariant mass
  - A BDT is used to suppress the continuum background*
    - The BDT discriminator boundaries are indicated with arrows in the figure on the right
    - Signal yield extraction and systematic uncertainty determinations are performed on the highest BDT intervals
- Partially reconstructed decays: one or more of the final-state particles (X) in a b hadron decay are not reconstructed
  - These candidates accumulate in the low dimuon invariant mass sideband
- Peaking background: $B^0_{(s)} \rightarrow hh'$ decays with both hadrons misidentified as muons



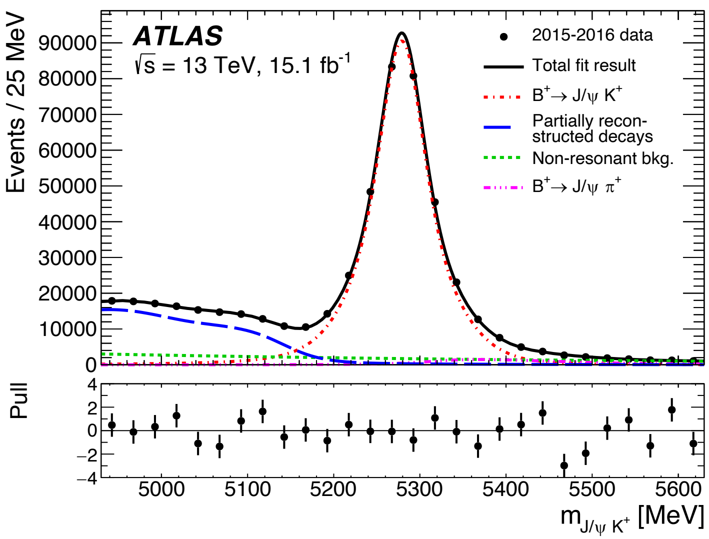




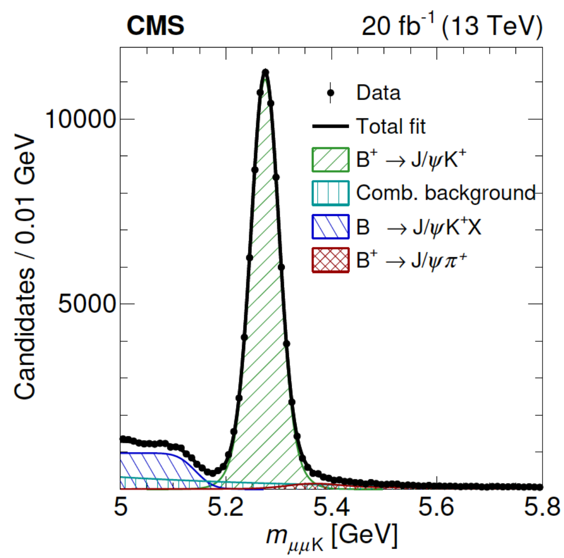
*See backup slides for more details

Reference Channel $B^+ \to J/\psi K^+$

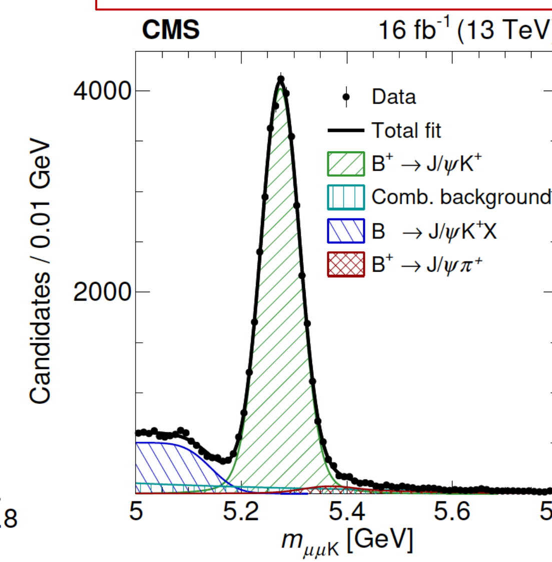
- The $B^\pm$ yield for the reference channel is extracted with an unbinned extended maximum-likelihood fit to the $J/\psi K^+$ invariant mass distribution
  - The two CMS figures are for the two subsamples of the 2016 dataset in different regions of pseudorapidity based on the most forward muon.
- The fit includes 4 components
  - $B^+ \to J/\psi K^+$ decays
  - Cabibbo-suppressed $B^+ \to J/\psi \pi^+$ decays
    - The $J/\psi \pi^+$ events are reconstructed using the K mass
  - Partially reconstructed B decays ($B^+ \to J/\psi K^+ X$)
  - Continuum background (composed mostly of $b\bar{b} \to J/\psi X$ decays)
- ATLAS: $B^+ \to J/\psi K^+$ yield for 2015-2016 data: 334,351 with a statistical uncertainty of 0.3%
- CMS: $B^+ \to J/\psi K^+$ yield for all data subsets is $1.43 \pm 0.06 \times 10^6$



2016A with $|\eta_\mu^f| < 0.7$



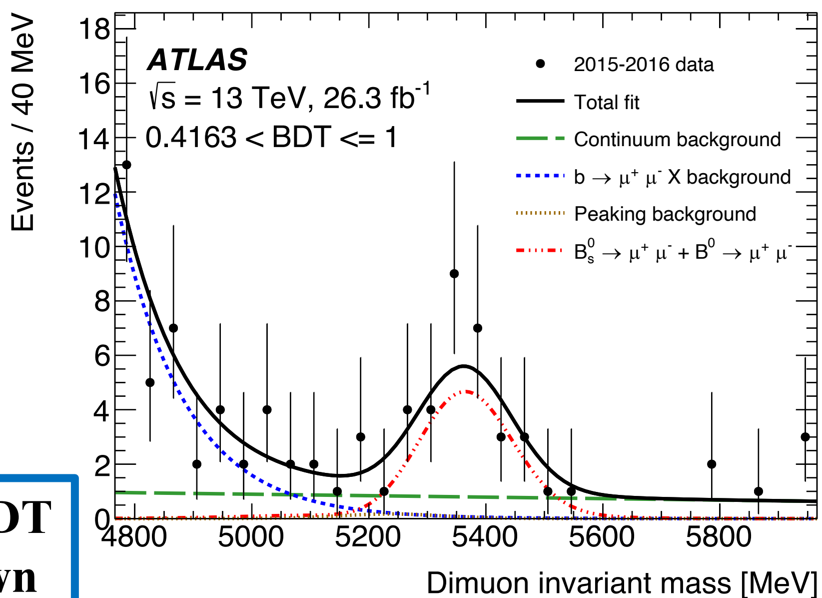
2016B with $0.7 < |\eta_\mu^f| < 1.4|$

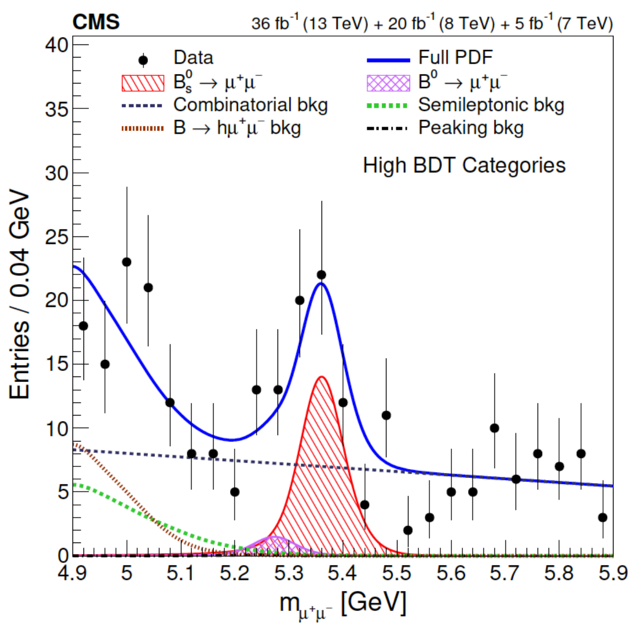


Signal Extraction and Yield Results

- The dimuon candidates are classified according to the BDT output
- ATLAS yield, determined from the unbinned maximum likelihood fit of highest three BDT bins
  - SM Expected: $N_s = 91$ and $N_d = 10$
  - $N_s = 80 \pm 22$ and $N_d = -12 \pm 20$
- CMS yield is determined from each BDT bin and data subset category (separated by year and detector region)
  - $N_s = 61^{+15}_{-13}$, results* are consistent the SM expectations

Highest BDT bin is shown





*Yield results for each data subset category for N_s and N_d are in the backup slides

Branching Fractions

- The branching fraction measurements for $B_s^0 \to \mu^+\mu^-$ and the upper limits on the $B^0 \to \mu^+\mu^-$ at 95% CL are:

ATLAS

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = \left(2.8^{+0.8}_{-0.7}\right) \times 10^{-9}$$

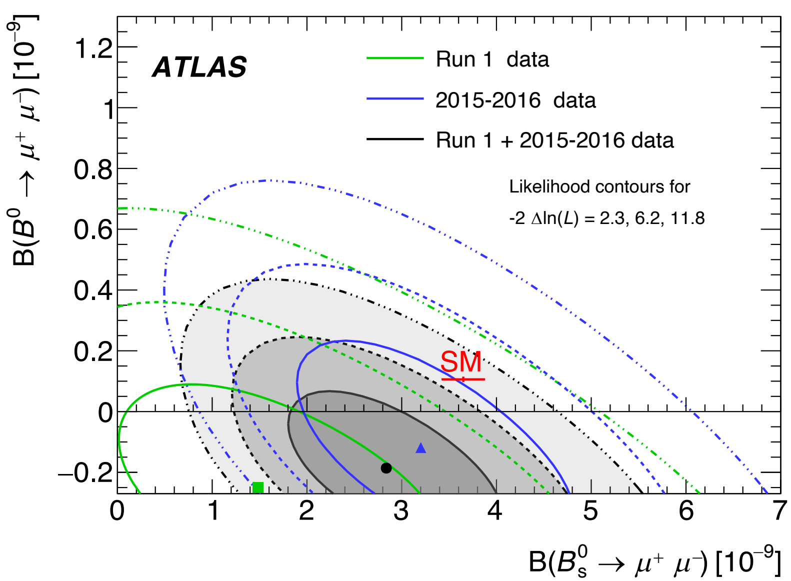
$$\mathcal{B}(B^0 \to \mu^+\mu^-) < 2.1 \times 10^{-10}$$

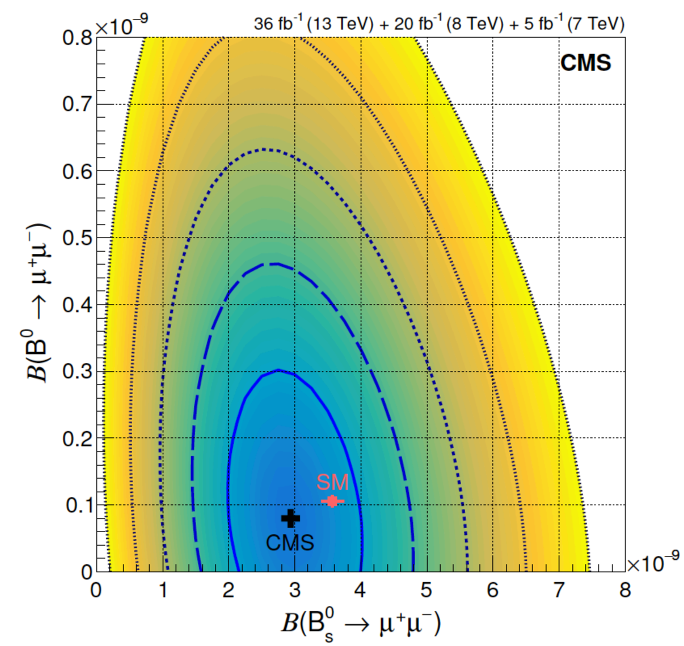
CMS

$$\mathcal{B}(\mathrm{B_s^0} \to \mu^+\mu^-) = [2.9 \pm 0.7\,(\mathrm{exp}) \pm 0.2\,(\mathrm{frag})] \times 10^{-9}$$

$$\mathcal{B}(\mathrm{B^0} \to \mu^+\mu^-) < 3.6 \times 10^{-10}$$

- The likelihood contours for the branching fractions are shown in the figures (the Neyman construction is used for ATLAS results)





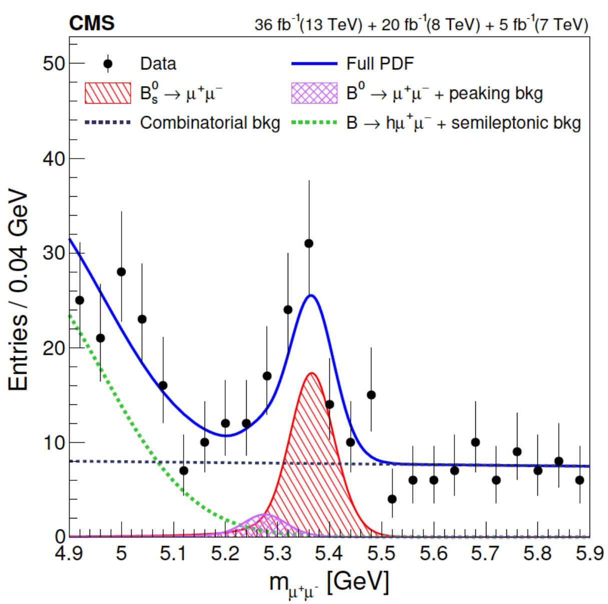
Lifetime Measurement

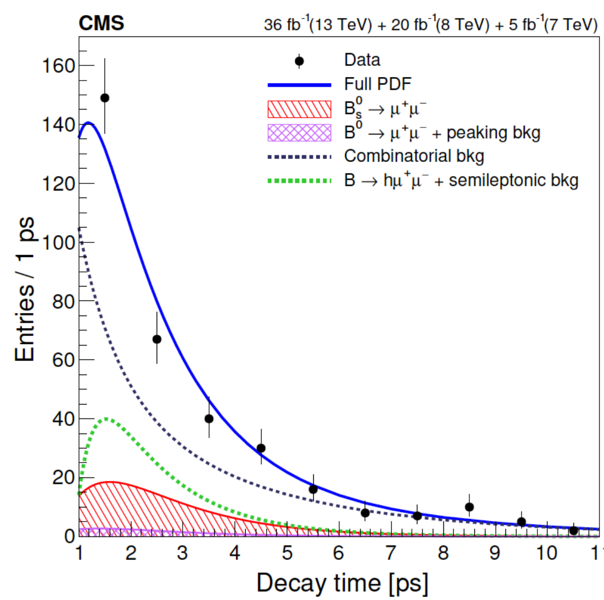
CMS

- A two dimensional unbinned maximum likelihood fit to the dimuon invariant mass and the proper decay time is implemented for extracting the $B_s^0 \to \mu^+\mu^-$ effective lifetime
  - The fit includes the signal and each background component

$$\tau_{\mu^+\mu^-} = [1.70\,^{+0.60}_{-0.43}\,(\mathrm{stat}) \pm 0.09\,(\mathrm{syst})]\ \mathrm{ps}$$

- Experimental World Average from PDG*: $\tau_{B^0_{sH}} = 1.615 \pm 0.009$ ps





* Particle Data Group collaboration, "Review of particle physics," Phys. Rev. D 98 (2018) 030001.

$\tau \to 3\mu$

- ATLAS: *Eur. Phys. J. C* 76 (2016) 232

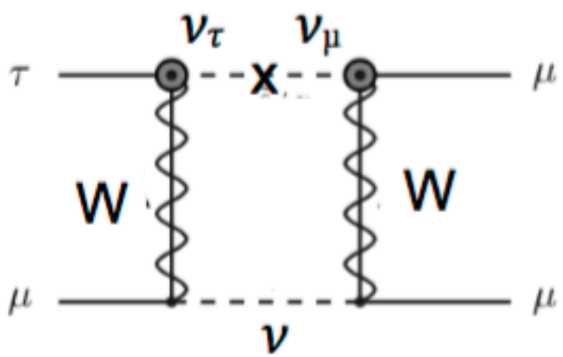
  20.3 fb^{-1} of $\sqrt{s} = 8$ TeV

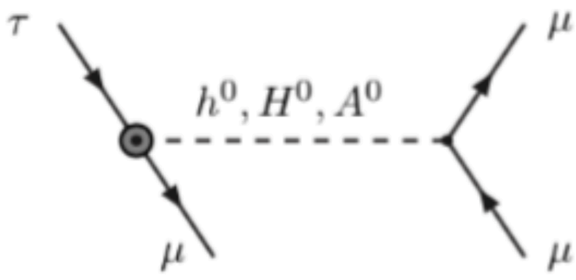
- CMS: CMS-PAS-BPH-17-004 (2019)

  33 fb^{-1} of $\sqrt{s} = 13$ TeV

Motivation

- Charged lepton flavor violation would be a major breakthrough in understanding the matter content of the universe
- Branching fraction is expected to very small in the SM*:
  - $\mathcal{B}(\tau \to 3\mu) < 10^{-14}$
- Some extensions to the Standard Model**, † lead to a branching fraction orders of magnitude greater (10^{-10}– 10^{-8}), within reach of experimental confirmation





* X.-Y. Pham, "Lepton flavor changing in neutrinoless tau decays," Eur. Phys. J. C 8 (1999) 513–516
**M. Raidal et al., "Flavour physics of leptons and dipole moments," Eur. Phys. J. C57 (2008) 13–182, arXiv:0801.1826 [hep-ph].
† W. J. Marciano, T. Mori, and J. M. Roney, "Charged Lepton Flavor Violation Experiments," Ann. Rev. Nucl. Part. Sci. 58 (2008) 315, doi:10.1146/annurev.nucl.58.110707.171126.

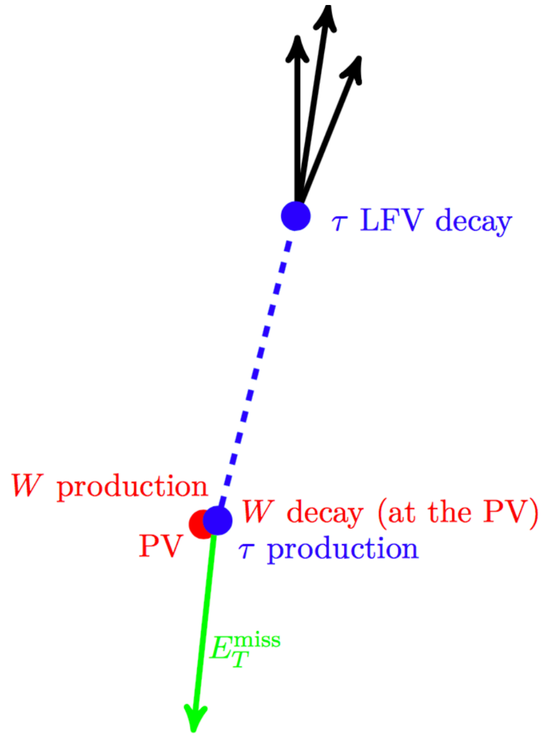
Analysis Overview

ATLAS

- The search for $\tau \to 3\mu$ uses the source channel: $W \to \tau\nu$
- A tau-neutrino from the W boson appears as missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$)
- The three muons from the τ lepton flavor violating decay are expected to have close geometric proximity
- The branching fraction is calculated as:

$$\mathrm{Br}(\tau \to 3\mu) = \frac{N_{\mathrm{s}}}{(\mathcal{A}_{\mathrm{s}} \times \epsilon_{\mathrm{s}})\, N_{W\to\tau\nu}}$$

  - where, N_{s} is the number of signal events, $\mathcal{A}_{\mathrm{s}}\times\epsilon_{\mathrm{s}}$ is the acceptance times efficiency of the signal and $N_{W\to\tau\nu}$ is the number of τ leptons produced via the decay $W \to \tau\nu$

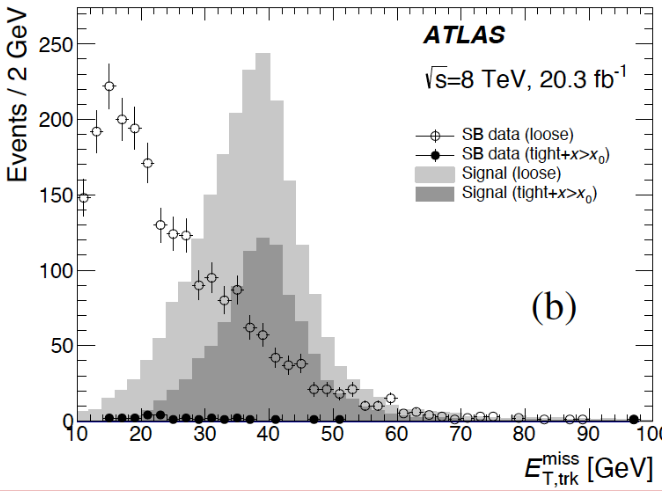


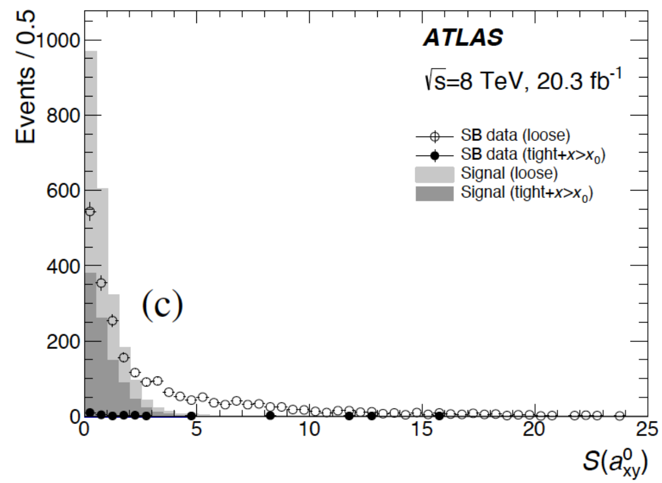
Analysis Procedures

ATLAS

1. Three muon vertex events are preselected and required to meet *loose* selection criteria* (including selections on the three-muon vertex fit quality and missing energy)
2. A BDT discriminator is trained with signal MC and background events in the BDT training region
   - BDT training region: $m_{3\mu}$ in [750, 1450] MeV and [2110, 2500] MeV
   - A loose cut on the BDT score ($x > x_0$, where $x_0 = -0.9$) is applied to remove background-like events
3. *Tight* selection criteria* are applied (including tightening of the loose requirements and mass restrictions for two muons with the same-charge and opposite-charge)

- Two analysis variables are shown, the track based missing transverse energy, $E_{\mathrm{T,trk}}^{\mathrm{miss}}$, and significance of the three-muon vertex fit a_{xy}^{0} significance, $S(a_{\mathrm{xy}}^{0})$



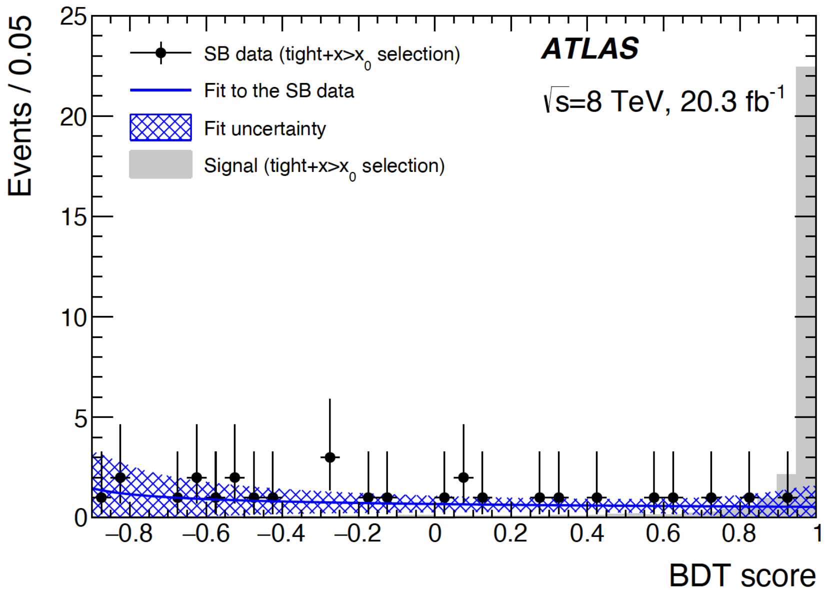


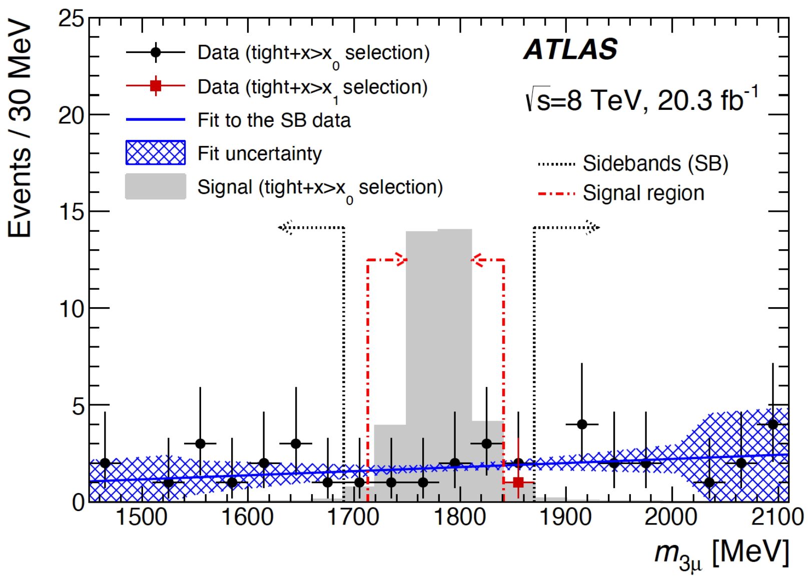
*Tight and loose selection criteria are described in the backup slides

Results

ATLAS

- The BDT distribution for the *tight* $+ x > x_0$ selection upon the sideband data and signal MC is shown in the figure on the left
  - A fit to the BDT score distribution of the sideband data, excluding the blinded data, is shown with the corresponding fit uncertainty
    - The fit is used to estimate the background in the signal region, and to scale the quantities measured in $x > x_0$ to the corresponding quantities in $x > x_1$
  - The optimal BDT cut is found to be $x_1 = 0.933$, optimizing the expected upper limit on the branching fraction
- The three-muon mass distribution for *tight* $+ x > x_0$, *tight* $+ x > x_1$, the fit to the sidebands, and the signal MC are shown in the figure on the right
- The observed upper limit on the branching fraction is:
  - $\mathcal{B}(\tau \to 3\mu) < 3.76 \times 10^{-7}$ at a 90% CL





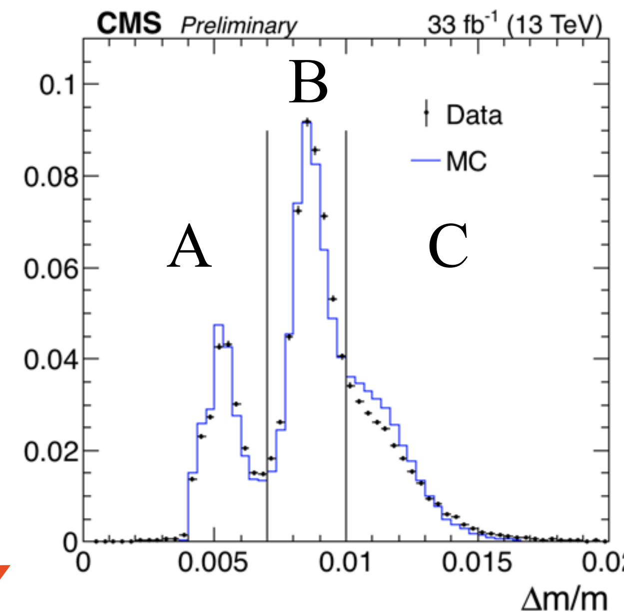
Analysis Overview

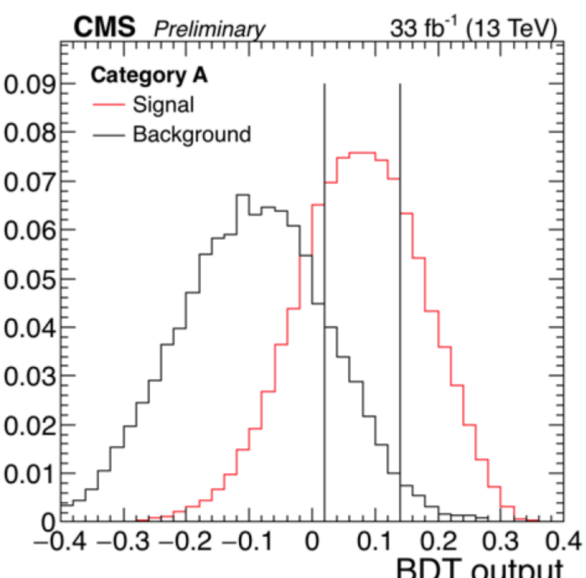
CMS

- Use the source of τ-leptons from D and B mesons
- With the 33 fb^{-1} of integrated luminosity used in the analysis:
  - $D \rightarrow \tau\nu$: 4×10^{12} expected number of τ leptons produced
  - $B \rightarrow \tau\nu$: 1.5×10^{12} expected number of τ leptons produced
  - $B \rightarrow D \rightarrow \tau\nu$: 6.3×10^{11} expected number of τ leptons produced
- Event selection trigger is for two muons and one track
- The branching fraction is measured using the normalization channel*

$$D_s \rightarrow \phi\pi \rightarrow \mu\mu\pi$$

- Data are separated into three categories (A, B, and C) based on the three-muon mass resolution
- A BDT is trained on signal simulation data and the three-muon mass sidebands for each mass resolution category
  - The BDT regions are optimized by maximizing the expected search sensitivity and used for signal extraction and uncertainty estimations.

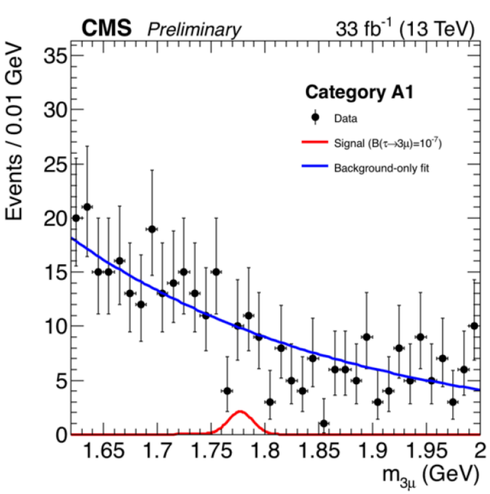


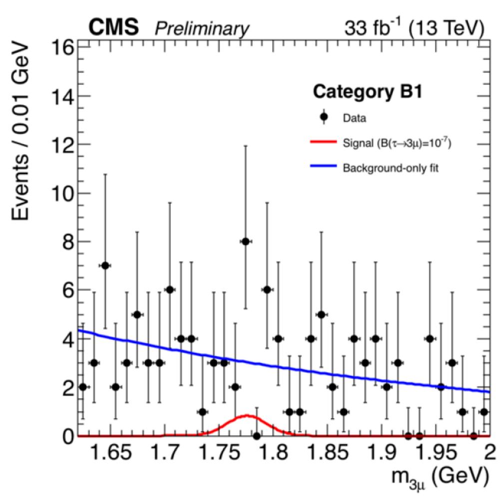


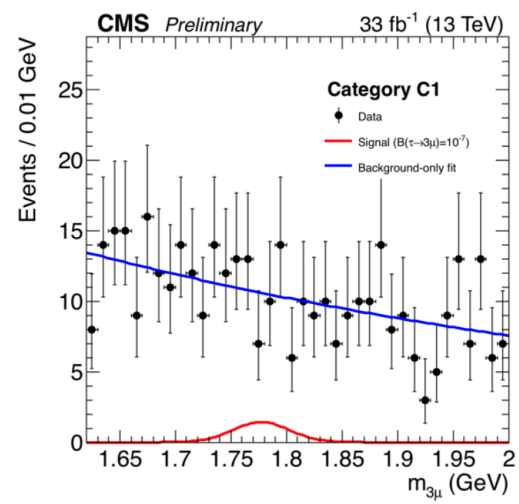
*See backup slides for more details

Mass Distributions and Results

- The three-muon mass distributions for the three mass resolution categories are shown in the highest BDT bins
- The signal is normalized in each category, assuming a branching fraction $\mathcal{B}(\tau \rightarrow 3\mu) = 10^{-7}$
- The yield results for the signal and data are summarized in the table for the mass range $1.62 - 2.00$ GeV (and for the signal mass region: $1.78 \pm 2\sigma$ in parentheses)
- The upper limit on the branching fractions is:
  - $\mathcal{B}(\tau \rightarrow 3\mu) < 8.8\times10^{-8}$ at a 90% CL







Mass resolution categories

	Signal		Data	
	sub-category 1	sub-category 2	sub-category 1	sub-category 2
Category A	6.3	10.3	360(44)	2502(319)
Category B	3.9	18.5	110(27)	2229(449)
Category C	9.4	9.6	389(107)	1549(400)

BDT Categories

Summary

- Results from ATLAS and CMS have been presented for

ATLAS | CMS

1. **The branching fraction measurements for $B_s^0 \to \mu^+\mu^-$**

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = \left(2.8^{+0.8}_{-0.7}\right) \times 10^{-9}$$

$$\mathcal{B}(\mathrm{B_s^0} \to \mu^+\mu^-) = [2.9 \pm 0.7\,(\mathrm{exp}) \pm 0.2\,(\mathrm{frag})] \times 10^{-9}$$

2. **Upper limits on $B^0 \to \mu^+\mu^-$**

$$\mathcal{B}(B^0 \to \mu^+\mu^-) < 2.1 \times 10^{-10}$$

$$\mathcal{B}(\mathrm{B^0} \to \mu^+\mu^-) < 3.6 \times 10^{-10}$$

3. **Effective lifetime measurement for $B_s^0 \to \mu^+\mu^-$ by CMS**

$$\tau_{\mu^+\mu^-} = [1.70\,^{+0.60}_{-0.43}\,(\mathrm{stat}) \pm 0.09\,(\mathrm{syst})]\ \mathrm{ps}$$

4. **Upper limits on $\tau \to 3\mu$ branching fraction**

$\mathcal{B}(\tau \to 3\mu) < 3.76\times10^{-7}$ at a 90% CL

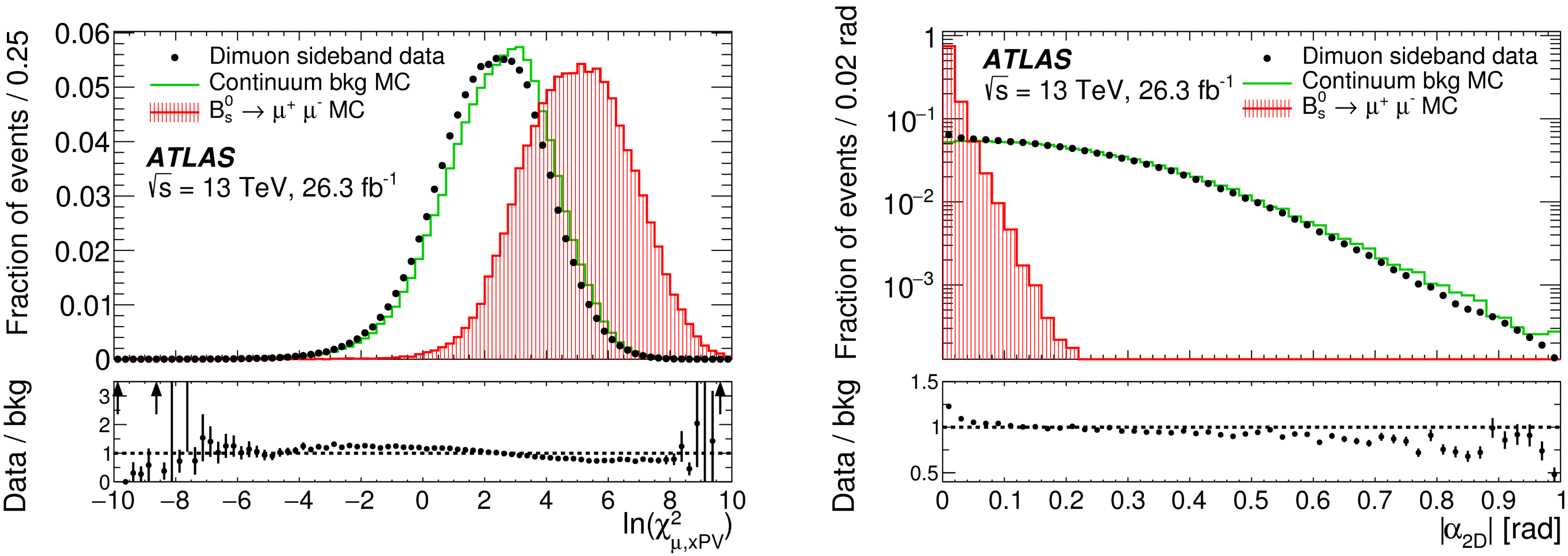
$\mathcal{B}(\tau \to 3\mu) < 8.8\times10^{-8}$ at a 90% CL

Additional Slides

$B^0_{(s)} \to \mu^+\mu^-$ Data-Simulation Comparisons

ATLAS

- The BDT is optimized when trained with 15 selected input variables - used to characterize a B meson event and the produced muons
- A grid search is performed to optimize the other BDT parameters

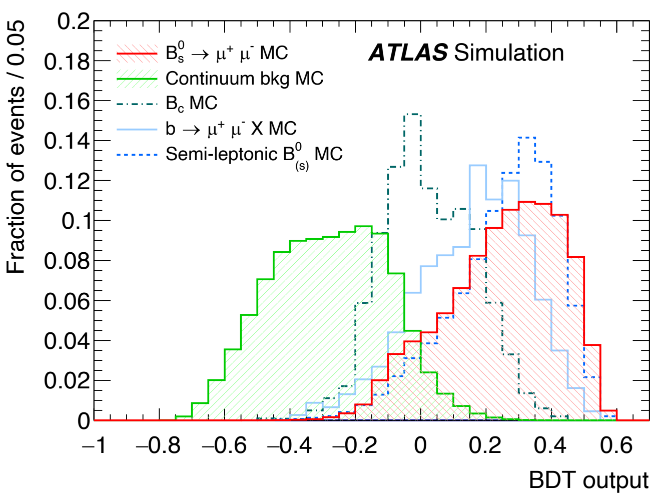


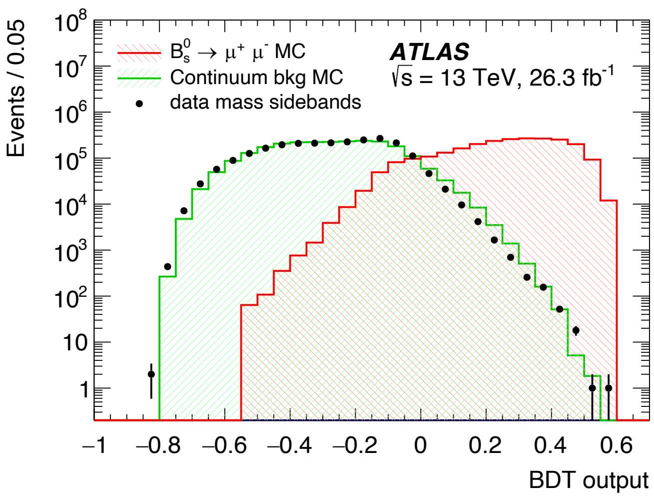
- Shown here are two of the input variables used in the training
- Care is taken to ensure that BDT output is not correlated with the invariant mass of the muons

BDT Continuum Background Suppression

ATLAS

- A multivariate approach, implemented as a Boosted Decision Tree (BDT), is used to enhance the signal relative to the continuum background
- Here is the BDT output for various datasets used in the analysis.
- A larger BDT output corresponds to more suppression of the continuum background
- Four BDT intervals are defined to give an equal efficiency of 18% for signal MC events, ordered according to increasing signal-to-background ratio
  - The lowest two BDT intervals contribute to background modelling.
  - Signal yield extraction and systematic uncertainty determinations are performed on the highest three BDT intervals.

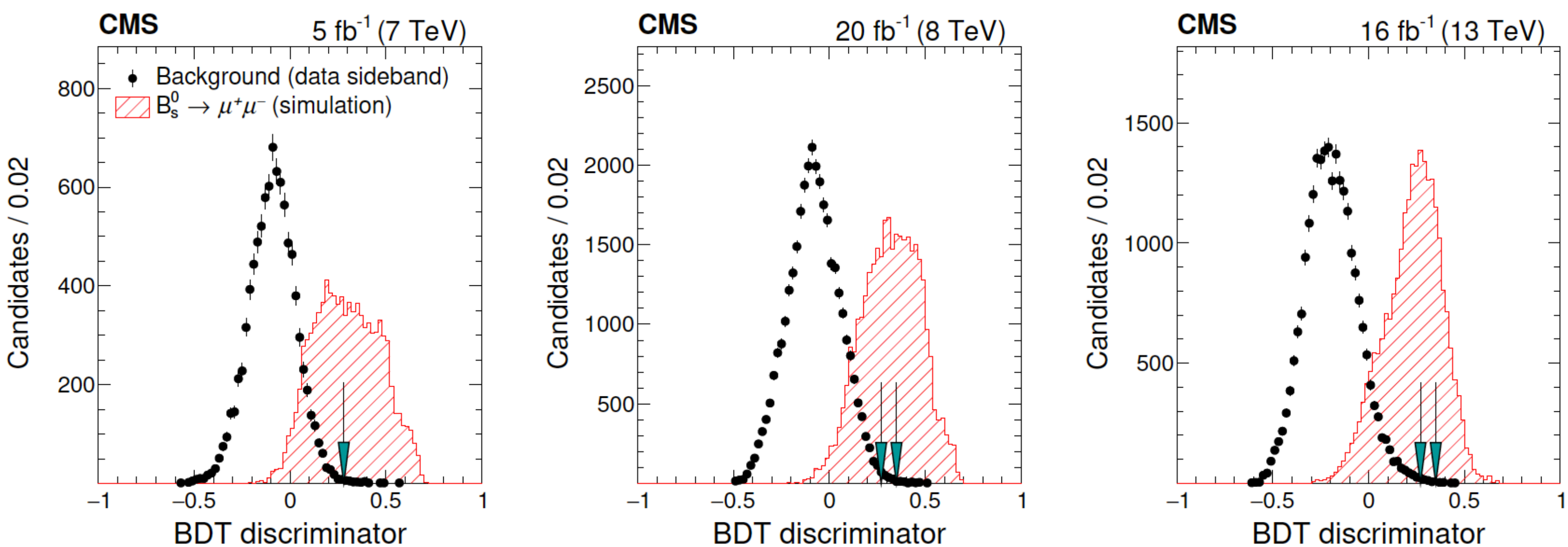




BDT Background Suppression

CMS

- One BDT is used to improve muon identification and suppress the peaking background
- A second (analysis) BDT is used to suppress the continuum background
  - The analysis BDT output is shown in the plots below for the signal MC and the sideband data for 7 TeV, 8 TeV, and 13 TeV datasets
  - The BDT boundaries are indicated with arrows in the figures
  - The binning of the analysis BDT discriminator distributions are used for the result extraction



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

ATLAS

- The efficiency ratio is required for the calculation of the signal branching fraction:

$$R_{\varepsilon} = \frac{\varepsilon(B^{+} \to J/\psi\, K^{+})}{\varepsilon(B^{0}_{(s)} \to \mu^{+}\mu^{-})}$$

- Both channels are measured in the fiducial acceptance for the B meson:
  - $p_T^B > 8$ GeV and $|\eta_B| < 2.5$
- The total efficiencies include acceptance and trigger, reconstruction and selection efficiencies.
  - Muon acceptance: $p_T^{\mu_1} > 6.0$ GeV, $p_T^{\mu_2} > 4.0$ GeV and $|\eta_{\mu_{1,2}}| < 2.5$
  - Kaon acceptance: $p_T^K > 1.0$ GeV and $|\eta_K| < 2.5$
  - The signal reference BDT selection: BDT > 0.2455

$B^0_{(s)} \to \mu^+\mu^-$ Extracted Yields by Category

CMS

Category	$N(\mathrm{B^0_s})$	$N(\mathrm{B^0})$	N_{comb}	$N^{\mathrm{B^+}}_{\mathrm{obs}}/100$	$\langle p_{\mathrm{T}}(\mathrm{B^0_s})\rangle$ [GeV]	$\varepsilon_{\mathrm{tot}}/\varepsilon^{\mathrm{B^+}}_{\mathrm{tot}}$
2011/central	$3.6\,^{+0.9}_{-0.8}$	$0.4\,^{+0.7}_{-0.6}$	2.3 ± 1.0	750 ± 30	16.4	3.9 ± 0.5
2011/forward	$2.0\,^{+0.5}_{-0.4}$	$0.2\,^{+0.4}_{-0.3}$	0.7 ± 0.5	220 ± 10	14.9	7.5 ± 0.8
2012/central/low	$3.7\,^{+0.9}_{-0.8}$	$0.4\,^{+0.6}_{-0.6}$	29.9 ± 2.9	790 ± 30	16.1	3.8 ± 0.5
2012/central/high	$9.3\,^{+2.3}_{-2.1}$	$1.0\,^{+1.7}_{-1.6}$	7.6 ± 1.8	2360 ± 100	17.3	3.2 ± 0.4
2012/forward/low	$1.7\,^{+0.4}_{-0.4}$	$0.2\,^{+0.3}_{-0.3}$	29.9 ± 2.9	190 ± 10	14.3	7.3 ± 1.0
2012/forward/high	$4.7\,^{+1.2}_{-1.1}$	$0.5\,^{+0.9}_{-0.8}$	8.3 ± 1.7	660 ± 30	15.5	5.9 ± 0.8
2016A/central/low	$2.2\,^{+0.5}_{-0.5}$	$0.2\,^{+0.4}_{-0.4}$	10.3 ± 1.7	580 ± 20	17.5	3.1 ± 0.4
2016A/central/high	$4.0\,^{+1.0}_{-0.9}$	$0.4\,^{+0.8}_{-0.7}$	3.4 ± 1.2	1290 ± 60	19.3	2.5 ± 0.3
2016A/forward/low	$3.7\,^{+0.9}_{-0.8}$	$0.4\,^{+0.7}_{-0.7}$	43.5 ± 3.5	780 ± 30	15.8	3.9 ± 0.5
2016A/forward/high	$8.1\,^{+2.0}_{-1.8}$	$0.8\,^{+1.5}_{-1.4}$	15.9 ± 2.4	1920 ± 80	17.5	3.4 ± 0.4
2016B/central/low	$4.1\,^{+1.0}_{-0.9}$	$0.4\,^{+0.8}_{-0.7}$	34.4 ± 3.2	1020 ± 40	17.2	3.3 ± 0.4
2016B/central/high	$3.6\,^{+0.9}_{-0.8}$	$0.4\,^{+0.7}_{-0.6}$	2.2 ± 1.0	1320 ± 50	20.8	2.2 ± 0.2
2016B/forward/low	$6.1\,^{+1.5}_{-1.4}$	$0.6\,^{+1.1}_{-1.0}$	33.4 ± 3.1	1260 ± 50	16.2	3.9 ± 0.4
2016B/forward/high	$3.9\,^{+1.0}_{-0.9}$	$0.4\,^{+0.8}_{-0.7}$	4.0 ± 1.3	1180 ± 50	19.5	2.7 ± 0.3

$B^0_{(s)} \to \mu^+\mu^-$ HL-LHC Prospects

ATLAS: ATL-PHYS-PUB-2018-005

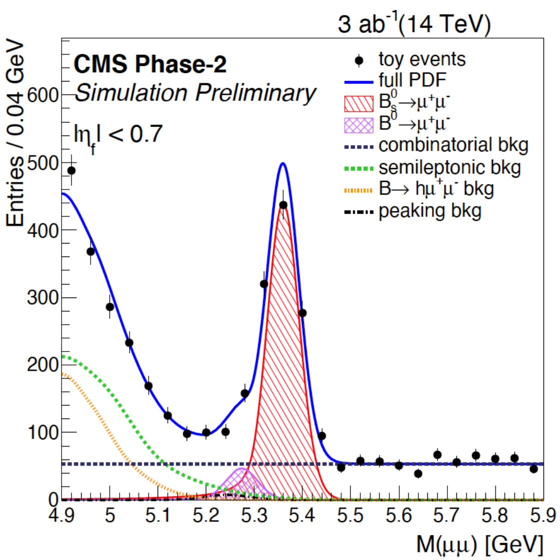
CMS: CMS-PAS-FTR-18-013

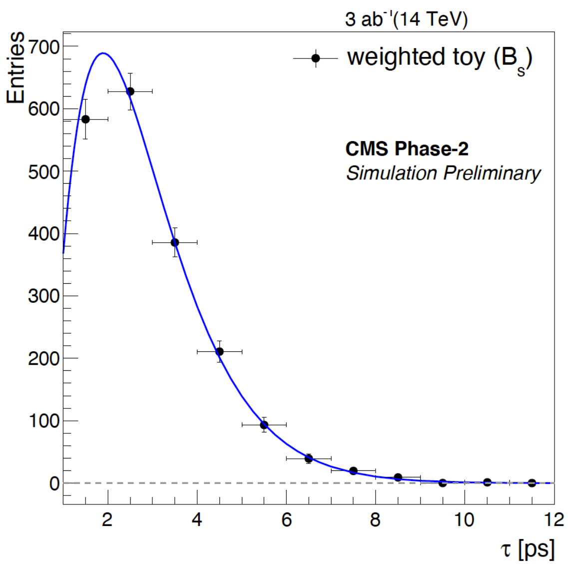
$B^0_{(s)} \to \mu^+\mu^-$ HL-LHC Projections

CMS

- The CMS Phase-2 inner tracker provides an order of 40-50% improvement on the mass resolutions over the Run-2 case, as determined from detailed MC studies.
- For the full Run 2 and HL-LHC statistics
  - The anticipated signal event yields, uncertainties and range of significance (**over 5$\boldsymbol{\sigma}$ for $\boldsymbol{\mathcal{B}(B^0 \to \mu^+\mu^-)}$**) are shown
  - The last column shows the anticipated statistical uncertainty on the effective lifetime measurement (**0.05 ps with 3 ab^{-1}** integrated luminosity)

$\mathcal{L}$ (fb^{-1})	$N(B_s)$	$N(B^0)$	$\delta\mathcal{B}(B_s \to \mu\mu)$	$\delta\mathcal{B}(B^0 \to \mu\mu)$	$\sigma(B^0 \to \mu\mu)$	$\delta[\tau(B_s)]$(stat-only)
300	205	21	12%	46%	$1.4 - 3.5\sigma$	0.15 ps
3000	2048	215	7%	16%	$6.3 - 8.3\sigma$	0.05 ps



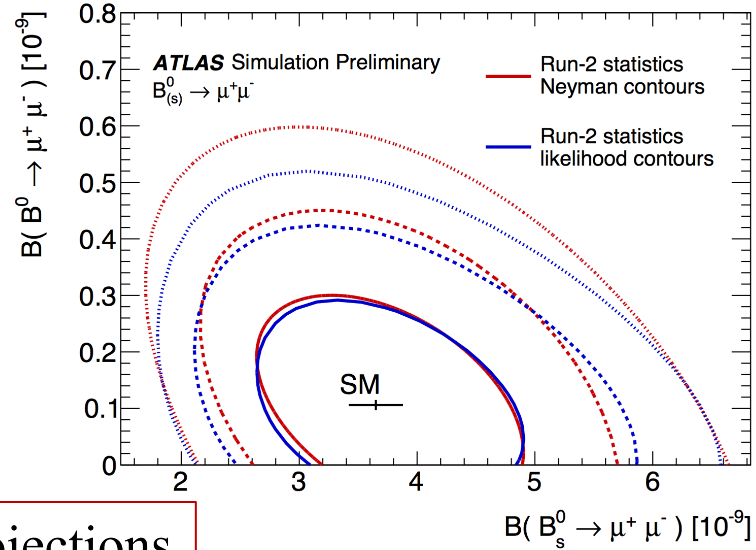


$B^0_{(s)} \to \mu^+\mu^-$ Projections

ATLAS

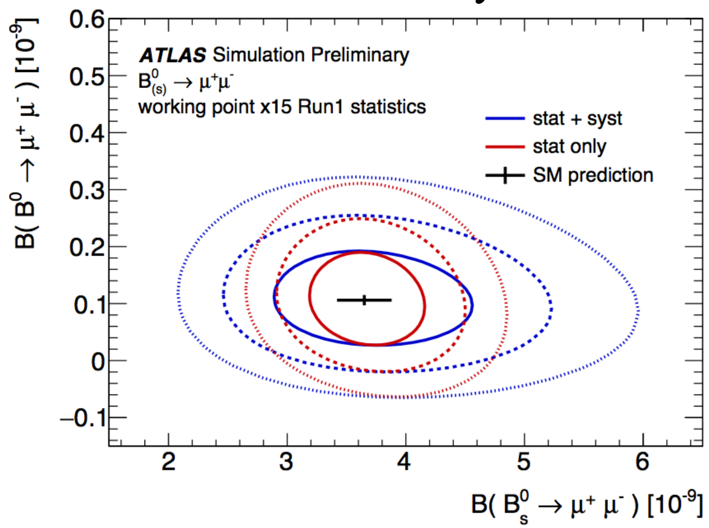
- Analysis sensitivities are expressed with the branching ratio contour plots.
- The full Run 2 and HL-LHC sensitivities are extrapolated from the Run 1 yields.
- Two scenarios for different di-muon trigger scenarios in the HL-LHC era are shown here.

Full Run 2 projections

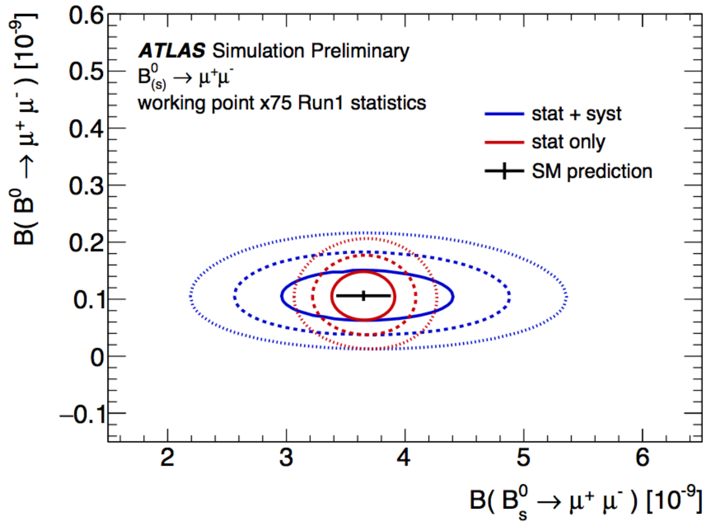


HL-LHC projections

Conservative yield



High-yield

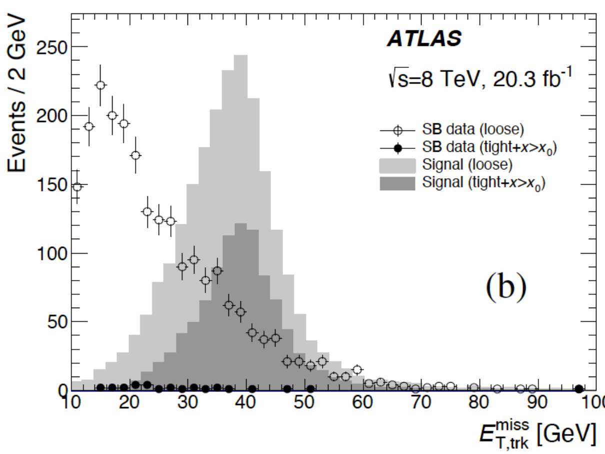


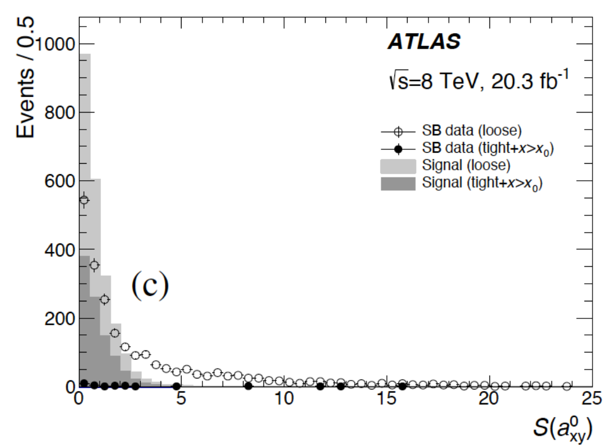
$\tau \rightarrow 3\mu$ Analysis Procedures

ATLAS

1. Events with three muons originating from a common vertex with combined mass less than 2.5 GeV are selected
2. *Loose* selection criteria* (including selections on the three-muon vertex fit quality and missing energy) are applied to obtain a background sample and a BDT is trained to discriminate the background
   - The data events are separated into three mutually exclusive regions: a blinded region (including a signal region), a sideband region and a training region
3. *Tight* selection criteria* are applied (including tightening of the loose requirements and mass restrictions for two muons with the same-charge and opposite-charge) and a loose cut on the BDT score ($x > x_0$)
4. Optimization of the final BDT cut, x_1, and statistical analysis of the *tight* $+ x > x_1$ sample

- Two analysis variables are shown, the track based missing transverse energy, $E_{\mathrm{T,trk}}^{\mathrm{miss}}$, and significance of the three-muon vertex fit a_{xy}^{0} significance, $S(a_{\mathrm{xy}}^{0})$



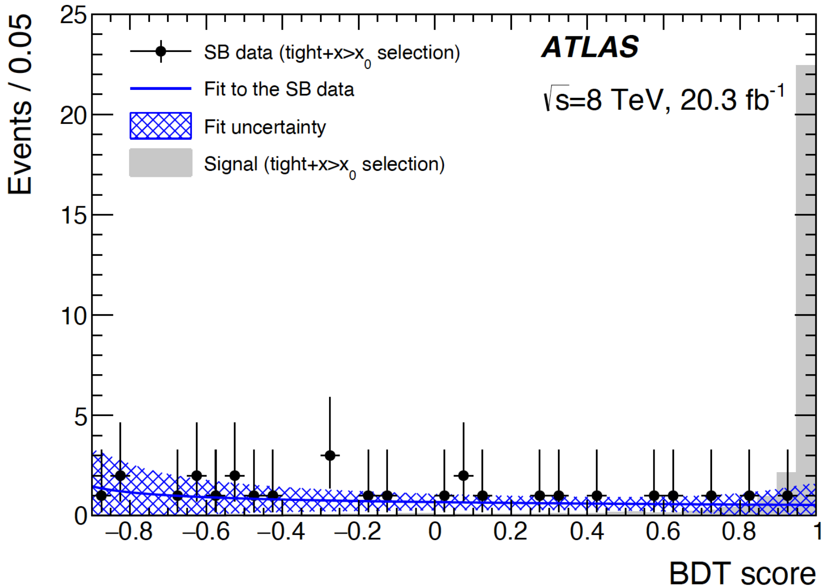


*Tight and loose selection criteria are described in the backup slides

BDT Optimization

ATLAS

- The BDT score for the *tight* and $x > x_0$ selection upon the sideband data and signal MC is shown in the figure
  - The loose x_0 BDT score cut is set to -0.9
  - A fit to the BDT score distribution of the sideband data, excluding the blinded data, is shown with the corresponding fit uncertainty
    - The fit is used to estimate the background in the signal region, and to scale the quantities measured in $x > x_0$ to the corresponding quantities in $x > x_1$
  - The optimal BDT cut is found to be $x_1 = 0.933$, optimizing the expected upper limit on the branching fraction



Loose Requirements for $\tau \to 3\mu$ Analysis

ATLAS

- The L_{xy} significance, $S(L_{xy}) = L_{xy}/\sigma_{L_{xy}}$ must satisfy $-10 < S(L_{xy}) < 50$, where $\sigma_{L_{xy}}$ is the uncertainty in the L_{xy}.
- The a_{xy}^0 significance, $S(a_{xy}^0) = a_{xy}^0/\sigma_{a_{xy}^0}$ must satisfy $S(a_{xy}^0) < 25$, where $\sigma_{a_{xy}^0}$ is the uncertainty in the a_{xy}^0.
- The three-muon track-fit probability product, $\mathcal{P}_{\text{trks}} = p_1 \times p_2 \times p_3$ (where p_i is the track fit p-value of track i), must satisfy $\mathcal{P}_{\text{trks}} > 10^{-9}$.
- The three muon transverse momentum must satisfy $p_{\text{T}}^{3\mu} > 10$ GeV.
- The calorimeter-based and track-based missing transverse energies, $E_{\text{T,cal}}^{\text{miss}}$ and $E_{\text{T,trk}}^{\text{miss}}$, respectively, must both satisfy $10 < E_{\text{T}}^{\text{miss}} < 250$ GeV.
- The calorimeter-based and track-based transverse masses, $m_{\text{T}}^{\text{cal}}$ and $m_{\text{T}}^{\text{trk}}$, respectively, must both satisfy $m_{\text{T}} > 20$ GeV.
- The three-muon track isolation is obtained from the sum of the p_{T} of all tracks $p_{\text{T}}^{trk} > 500$ MeV in a cone of $\Delta R_{\text{max}}^{3\mu} + 0.20$ (and $\Delta R_{\text{max}}^{3\mu} + 0.30$) around the three muon momentum while excluding its constituent tracks; it must satisfy $\sum p_{\text{T}}^{trk}(\Delta R_{\text{max}}^{3\mu} + 0.20)/p_{\text{T}}^{3\mu} < 0.30$ (and $\sum p_{\text{T}}^{trk}(\Delta R_{\text{max}}^{3\mu} + 0.30)/p_{\text{T}}^{3\mu} < 1$). The largest separation, $\Delta R_{\text{max}}^{3\mu}$, between any pair of the three-muon tracks is on average 0.07 for the signal.

Tight Requirements for $\tau \to 3\mu$ Analysis

ATLAS

- A number of the *loose* requirements are tightened, namely $\mathcal{P}_{\text{trks}} > 8\times10^{-9}$, $m_{\text{T}}^{\text{cal}} > 45$ GeV, $m_{\text{T}}^{\text{trk}} > 45$ GeV, and $1 < S(L_{\text{xy}}) < 50$.
- Three-muon vertex fit probability must have p-value > 0.2.
- The angle between $\sum_{\text{T}}$ and $E_{\text{T,cal}}^{\text{miss}}$ ($E_{\text{T,trk}}^{\text{miss}}$) directions is required to be $\Delta\phi_{\sum_T}^{\text{cal}} > 2$ ($\Delta\phi_{\sum_T}^{\text{trk}} > 2$).
  - $\sum_{\text{T}}$ is the transverse component of the vector sum of the three-muon and leading jet momenta.
- The same-charge two muon mass, m_{SS}, and opposite-charge two muon mass, m_{OS1} or m_{OS2}, satisfy $m_{\text{SS}} > 300$ MeV, $m_{\text{OS1}} > 300$ MeV , and $m_{\text{OS2}} > 300$ MeV, where $m_{\text{OS1}}(m_{\text{OS2}})$ is the mass of the two opposite-charge muon pairs with highest (second highest) summed scalar p_{T} among the three muons
- The event is rejected if $|m_{\text{OS}} - m_{\omega}| < 50$ MeV or $|m_{\text{OS}} - m_{\phi}| < 50$ MeV if either the $p_{\text{T}}^{3\mu}$, the $E_{\text{T,cal}}^{\text{miss}}$, or the $E_{\text{T,trk}}^{\text{miss}}$ is lower than 35 GeV.
- The event is rejected if $|m_{\text{OS}} - m_{\phi}| < 50$ MeV if $|m_{3\mu} - m_{D_s}| < 100$ MeV

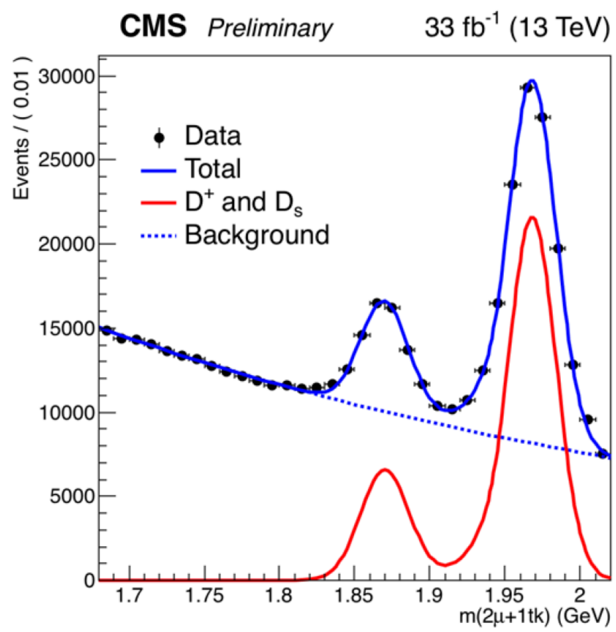
In the above notation, m_{OS} is m_{OS1} or m_{OS2}, and m_{ω}, m_{ϕ}, m_{D_s} are the masses of the ω, ϕ, and D_s mesons*, respectively

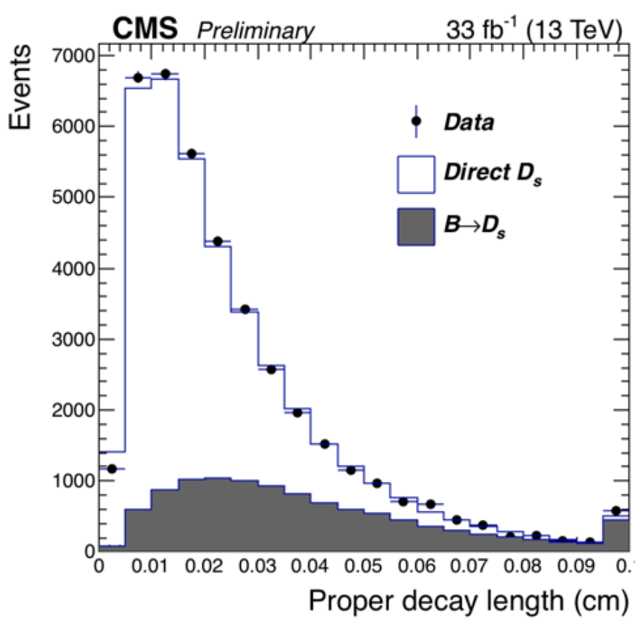
*K. Olive et al., Review of Particle Physics, Chin. Phys. C38 (2014) 090001.

Normalization Channel

CMS

- (left plot) The invariant mass distribution for two muons and a pion is shown
  - Kinematic cuts are applied to the two muons and the pion, the muons are required to have opposite signs and an invariant mass consistent with the ϕ meson mass.
  - The two peaks are associated with D_s(1.97 GeV) and D^+ (1.87 GeV) decays, and modelled with Crystal Ball functions, while the background is fitted with an exponential function.
- (right plot) Simulated prompt and non-prompt contributions of D_s are compared to data in the proper decay length distribution

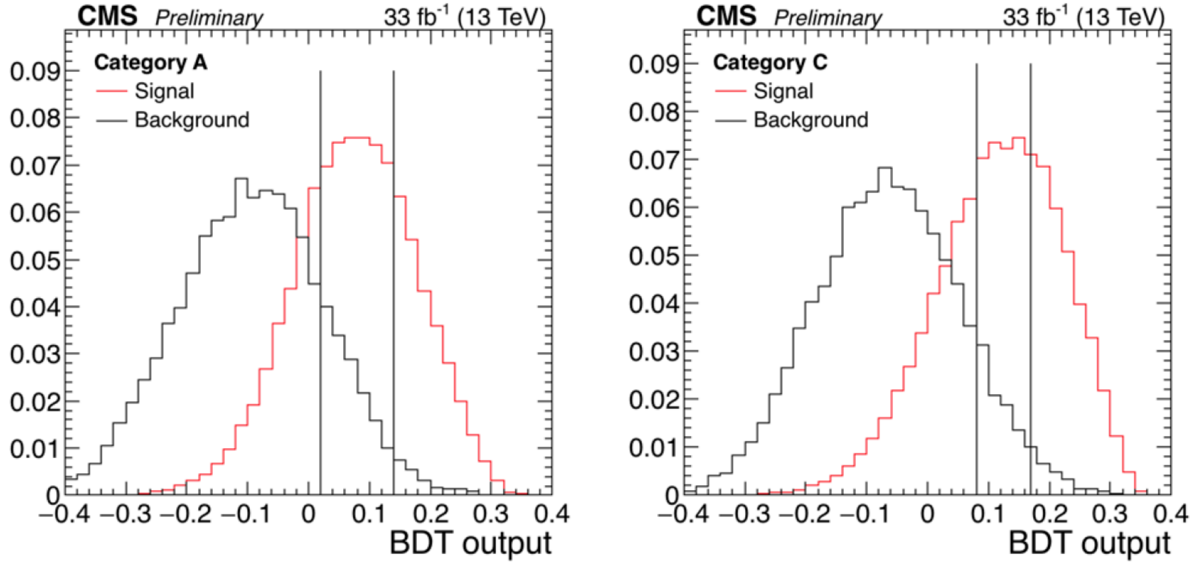




BDT Training

CMS

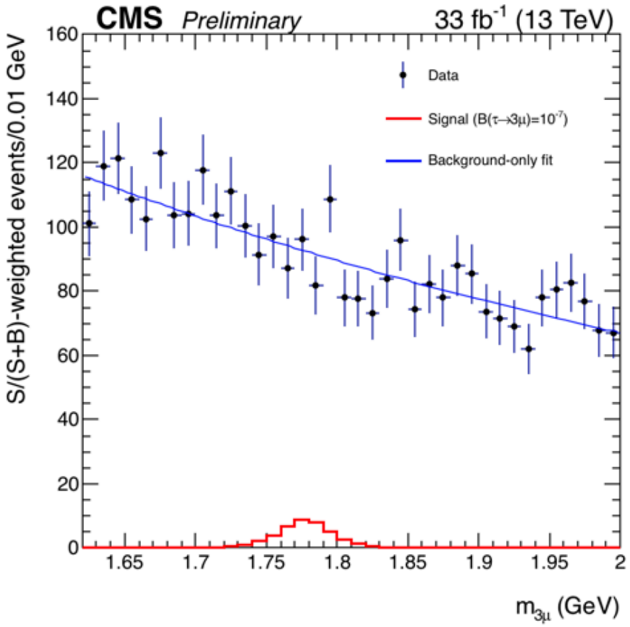
- A BDT is trained for each of the three mass resolution categories
- The BDTs are trained on signal simulation data and the three-muon mass sidebands
  - Ten analysis observables are used in the training and are separated into two categories: (1) variables associated with the three-muon vertex properties and (2) variables associated with reducing hadrons misidentified as muons and muons originating from a pion or kaon source
- The boundaries defining three BDT regions are optimized by maximizing the expected search sensitivity.
  - The two BDT regions with the best signal-to-background purity are retained for signal extraction and uncertainty estimations.



$\tau \rightarrow 3\mu$ Results

CMS

- S/(S+B) weighted three-muon mass distribution including events from all mass resolution categories used in the analysis is shown
- The upper limit on the branching fractions is:
  - $\mathcal{B}(\tau \rightarrow 3\mu) < 8.8\times10^{-8}$ at a 90% CL



$\tau \rightarrow 3\mu$ HL-LHC Prospects

ATLAS: ATL-PHYS-PUB-2018-032

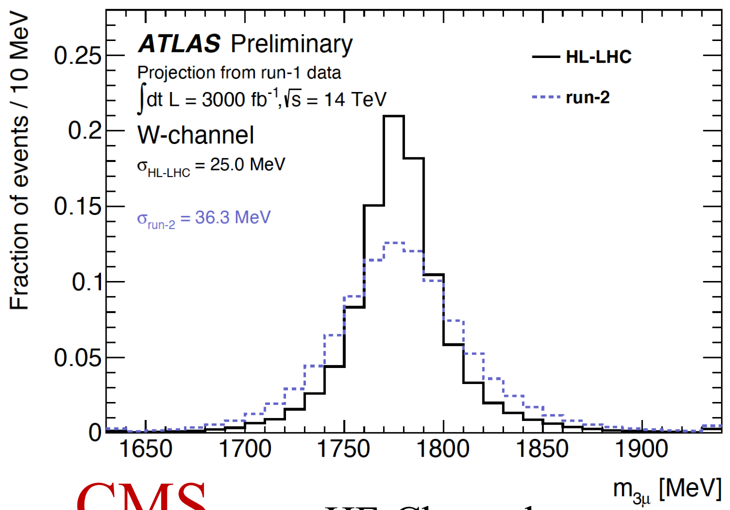
- HL-LHC prospects are also summarized in:

"Opportunities in Flavour Physics at the HL-LHC and HE-LHC,"
CERN-LPCC-2018-06

HL-LHC Projections $\tau \to 3\mu$

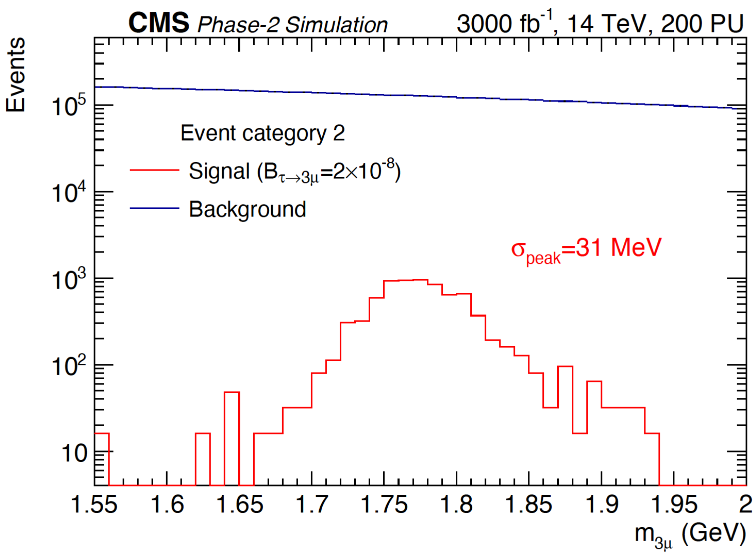
ATLAS

W-Channel



CMS

HF-Channel



- Simulated three-muon mass distributions under HL-LHC detector conditions are shown
  - Widths are estimated from double gaussian fits.
  - For ATLAS, several conditions are investigated including improvements to vertex and momentum resolution with a new tracking system and improved low muon trigger thresholds
  - For CMS, two categories are considered: (1) all three muons reconstructed only with the Phase-1 detectors, and (2) at least one muon reconstructed by the new triple Gas Electron Multiplier (GEM) detectors in the upgraded muon system
- 90% confidence level limits are summarized in the tables below for (top table) ATLAS and (bottom table) CMS

ATLAS

Scenario	W-channel 90% CL UL [10^{-9}]	HF-channel 90% CL UL [10^{-9}]
ATLAS High	5.4	1
ATLAS Medium	6.2	2.3
ATLAS Low	13.5	6.4

CMS

	Category 1	Category 2
Number of background events	2.4×10^6	2.6×10^6
Number of signal events	4580	3640
Trimuon mass resolution	18 MeV	31 MeV
$\mathcal{B}(\tau \to 3\mu)$ limit per event category	4.3×10^{-9}	7.0×10^{-9}
$\mathcal{B}(\tau \to 3\mu)$ 90% C.L. limit	3.7×10^{-9}	
$\mathcal{B}(\tau \to 3\mu)$ for 3-σ evidence	6.7×10^{-9}	
$\mathcal{B}(\tau \to 3\mu)$ for 5-σ observation	1.1×10^{-8}	