Rare flavor physics decays at ATLAS and CMS

Aidan Grummer
Fermilab

on behalf of the ATLAS and CMS Collaborations

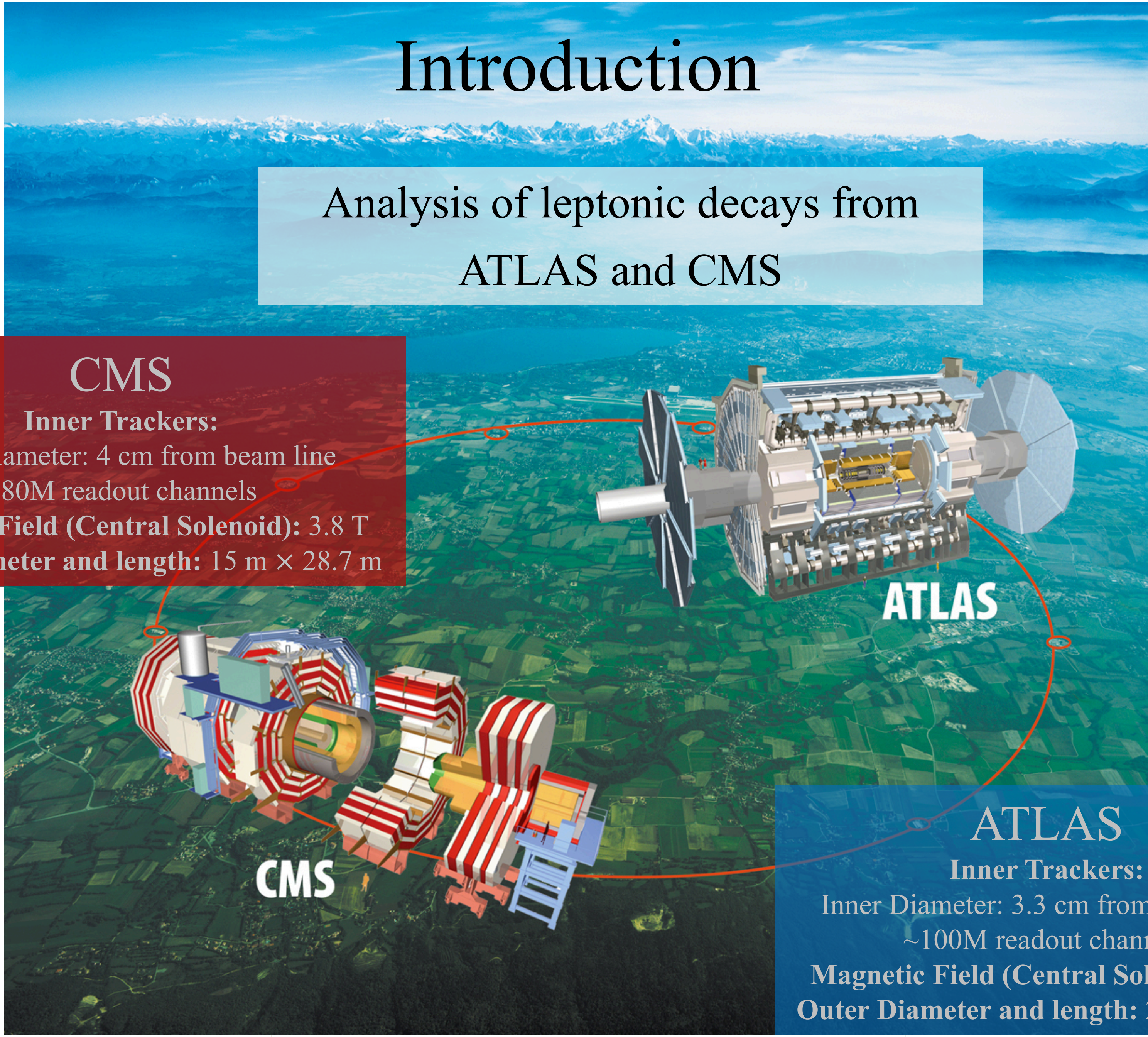
PANIC
Sept. 5, 2021

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

$\boldsymbol{B^0_d \to K^{*0}(892)\mu^+\mu^-}$
and
$\boldsymbol{B^+ \to K^{*}(892)^+\mu^+\mu^-}$

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

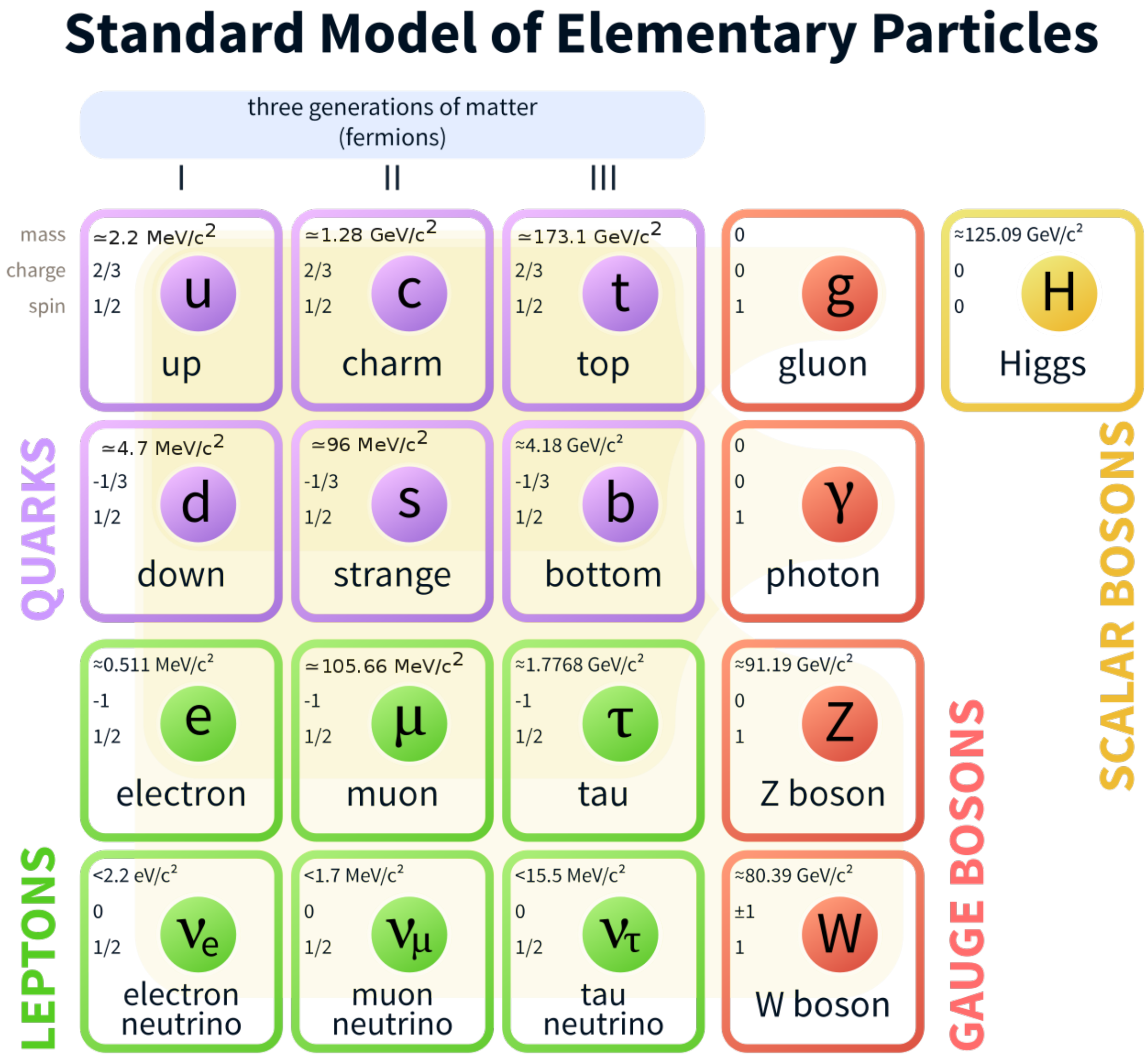
~100M readout channels

Magnetic Field (Central Solenoid): 2 T

Outer Diameter and length: 25 m × 46 m

Probing the Standard Model

- All of the known elementary particles are shown in the image on the right*
- The aim of the research presented today is to compare experiment and predictions of the several decay channels of **B mesons**
- B mesons are composite particles made up of a *bottom* quark in a bound state with another quark



* Image from: https://en.wikipedia.org/wiki/Particle_physics

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

- 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)

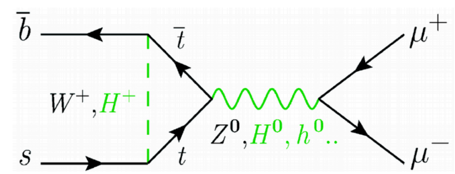
- Combination of the ATLAS, CMS and LHCb results:

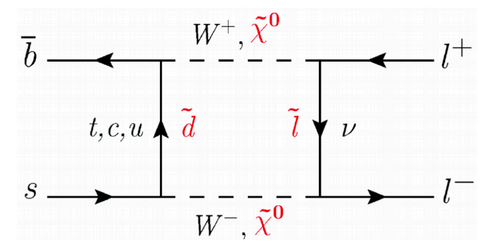
CMS-PAS-BPH-20-003, LHCb-CONF-2020-002, ATLAS-CONF-2020-049

*Due to operational instabilities experienced with the CMS microstrip detector, CMS 2016 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
  - PREDICTED $\mathcal{B}(B^0_s \to \mu^+\mu^-) = (3.66 \pm 0.14) \times 10^{-9}$
  - PREDICTED $\mathcal{B}(B^0 \to \mu^+\mu^-) = (1.03 \pm 0.05) \times 10^{-10}$
- 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**: $\tau_{B^0_{sH}} = 1.616 \pm 0.010$ 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 ‡





$\boxed{\text{“New Physics”}}$

* 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, Prog. Theor. Exp. Phys. 2020, 083C01 (2020) and 2021 update.

*** Huang, Chao-Shang and Liao, Wei and Yan, Qi-Shu, “Promising process to distinguish supersymmetric models with large tan β from the standard model: $B \to X_s \mu+ \mu-$,” 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].

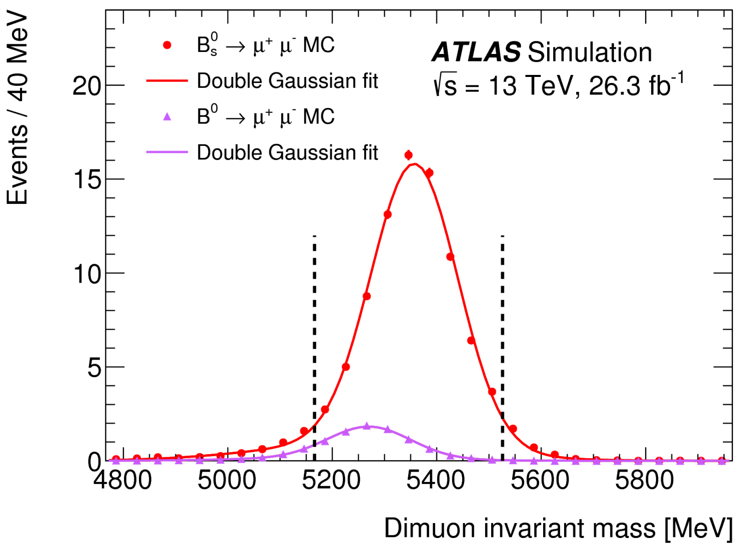
‡ K. S. Babu and C. F. Kolda, “Higgs mediated $B^0 \to \mu+ \mu-$ in minimal supersymmetry,” Phys. Rev. Lett. 84 (2000) 228, arXiv: hep-ph/9909476 [hep-ph].

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

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

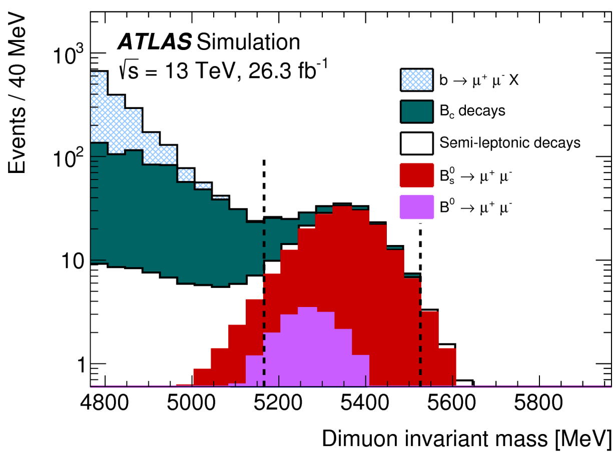
  - Here $N_{d(s)}$ is the signal yield, $N_{J/\psi K^+}$ is the reference yield, $\varepsilon_{\mu^+\mu^-}$ and $\varepsilon_{J/\psi K^+}$ are acceptances times efficiencies and $f_u/f_{d(s)}$ is the ratio of the hadronization probabilities*, **,†,‡ of a b-quark into $B^+$and $B^0_{(s)}$.
- Perform a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

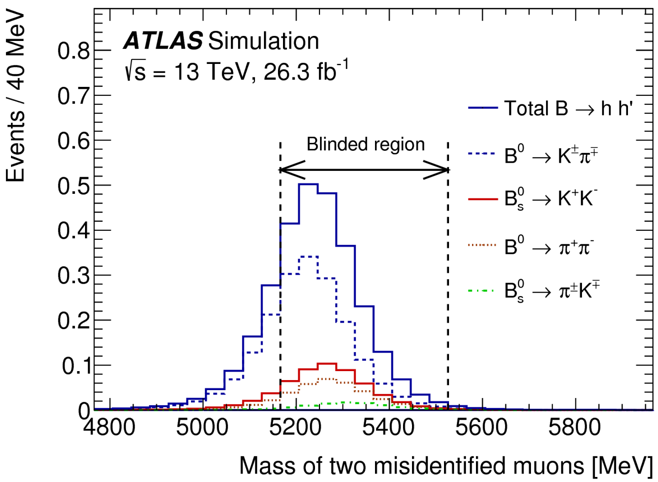


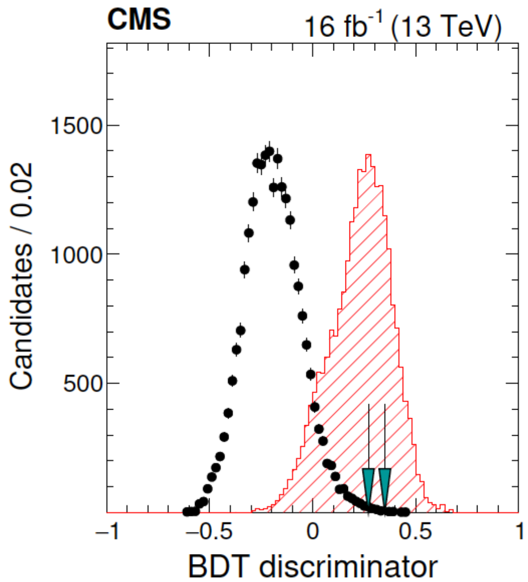
* HFLAV Group, Eur. Phys. J. C 77 (2017) 895, http://dx.doi.org/10.1140/epjc/s10052-017-5058-4 [Used in paper]
** PDG Collaboration, Phys. Rev. D 98 (2018) 030001, https://doi.org/10.1103/PhysRevD.98.030001 [Used in paper]
† HFLAV Group, Eur. Phys. J. C 81 (2021) 226, https://doi.org/10.1140/epjc/s10052-020-8156-7 [latest]
‡ Particle Data Group collaboration, Prog. Theor. Exp. Phys. 2020, 083C01 (2020) and 2021 update https://doi.org/10.1093/ptep/ptaa104 [latest]

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 boosted decision tree (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 (left plot)
  - These candidates accumulate in the low dimuon invariant mass sideband
- Peaking background: $B^0_{(s)} \to hh'$ decays with both hadrons misidentified as muons (middle plot)



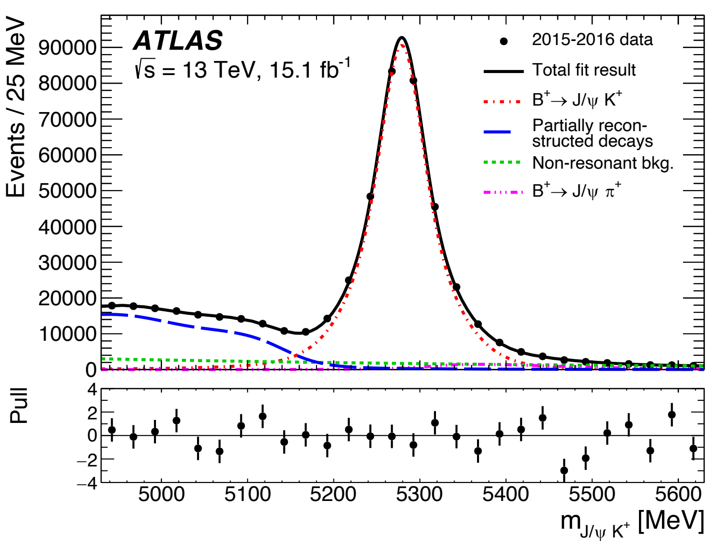




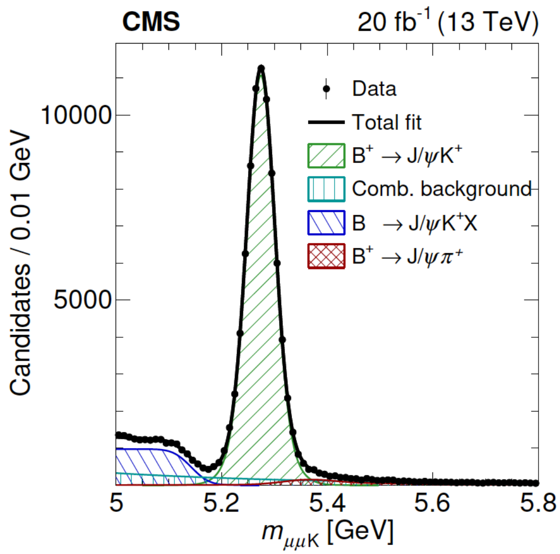
*See backup slides for more details

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

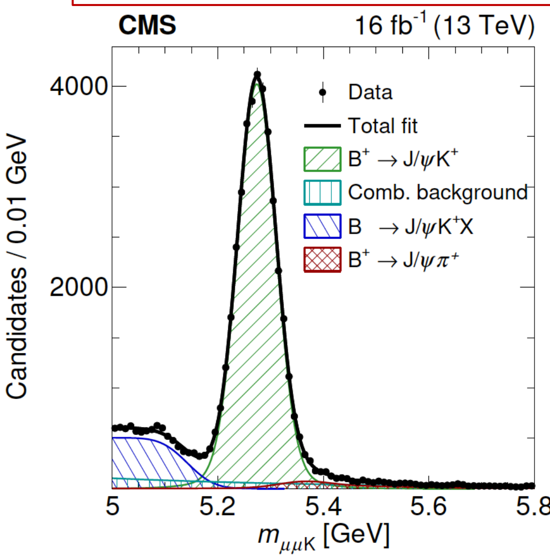
- The $B^\pm$ yield for the reference channel is extracted with an unbinned extended maximum-likelihood fit to the $J/\psi K^+$ invariant mass distribution
  - The fit is shown (left plot) for ATLAS and (middle and right plots) for two CMS subsamples of the 2016 dataset in different regions of pseudorapidity based on the most forward muon, η_μ^f.
- 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 $|\eta_\mu^f| < 0.7$



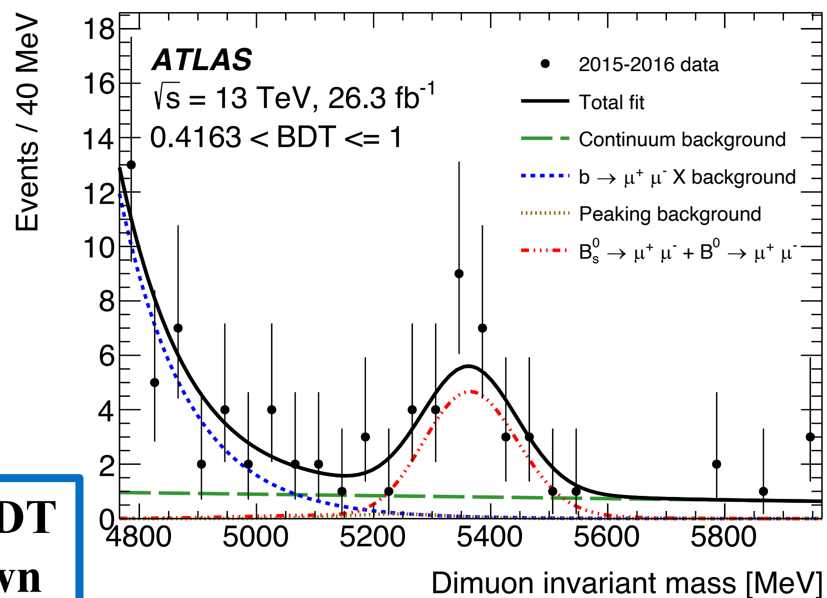
2016B with $0.7 < |\eta_\mu^f| < 1.4$

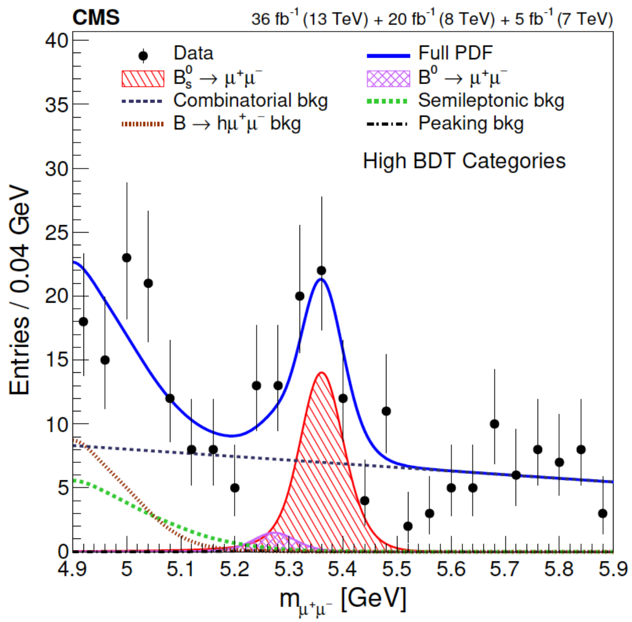


Signal Extraction and Yield Results

- The dimuon candidates are classified according to the BDT output
- ATLAS yield, determined from the unbinned maximum likelihood fit of highest three BDT bins, simultaneously
    - 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

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

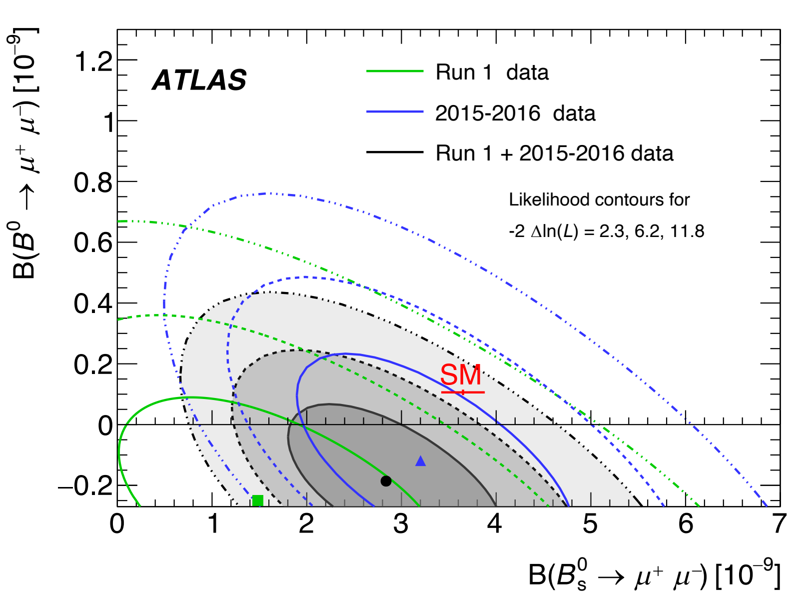
$$\mathcal{B}(B^0 \to \mu^+\mu^-) < 2.1 \times 10^{-10}$$

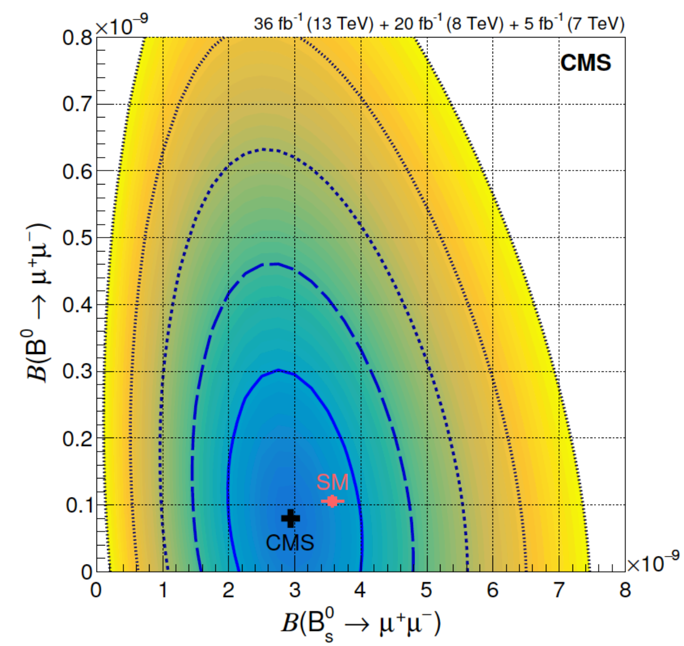
CMS

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

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

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





Fragmentation fraction uncertainty (frag): PDG Collaboration, Phys. Rev. D 98 (2018) 030001, https://doi.org/10.1103/PhysRevD.98.030001

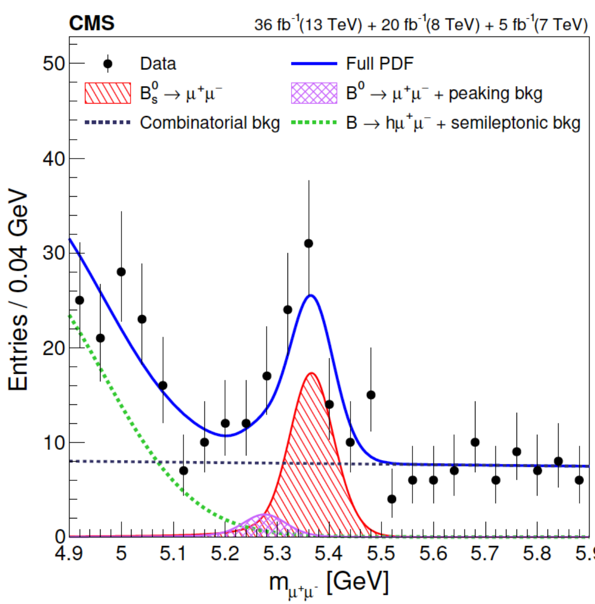
Lifetime Measurement

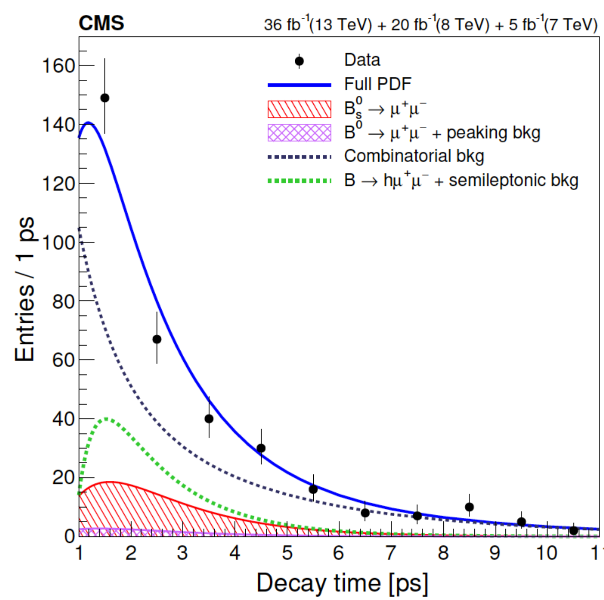
CMS

- 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

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

- Experimental World Average from PDG*: $\tau_{B^0_{sH}} = 1.616 \pm 0.010$ ps





* Particle Data Group collaboration, Prog. Theor. Exp. Phys. 2020, 083C01 (2020) and 2021 update.

LHC Combination

- The combination of the ATLAS, CMS, and LHCb* branching ratios is performed using binned two-dimensional profile likelihoods obtained by each experiment from their fit to the dimuon invariant mass distributions.
  - This combination does not include the latest LHCb results (for 9 fb^{-1})**, †
  - The uncertainties on the current measurements are dominated by the statistical component.
  - The systematic uncertainties affecting the three measurements are considered to be independent with the exception of the fragmentation fractions f_d/f_s
    - In order to take this into account, this factor was profiled separately in each likelihood, retaining its uncertainty in only one (LHCb) of the three experiments.
- The CMS and LHCb experiments also measured the effective lifetime of the observed $B_s^0 \to \mu^+\mu^-$ candidates.
  - For both experiments, the measurement is fully dominated by its statistical uncertainty, hence the two results are uncorrelated
  - Two variable-width Gaussian likelihoods are used to describe the CMS and LHCb original likelihoods and the value of $-2\Delta\ln L$ obtained from these functions

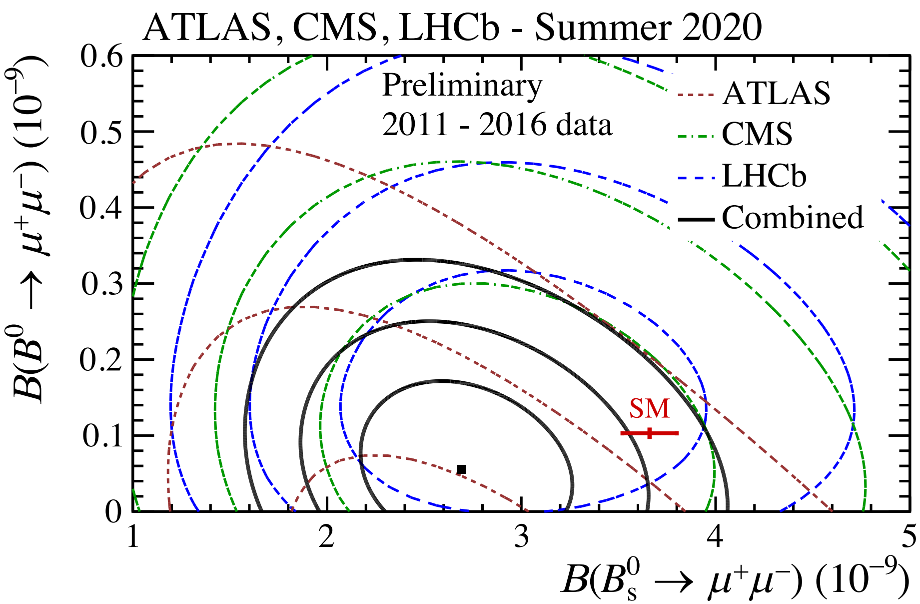
* LHCb Collaboration, Phys. Rev. Lett. 118, 191801, http://dx.doi.org/10.1103/PhysRevLett.118.191801

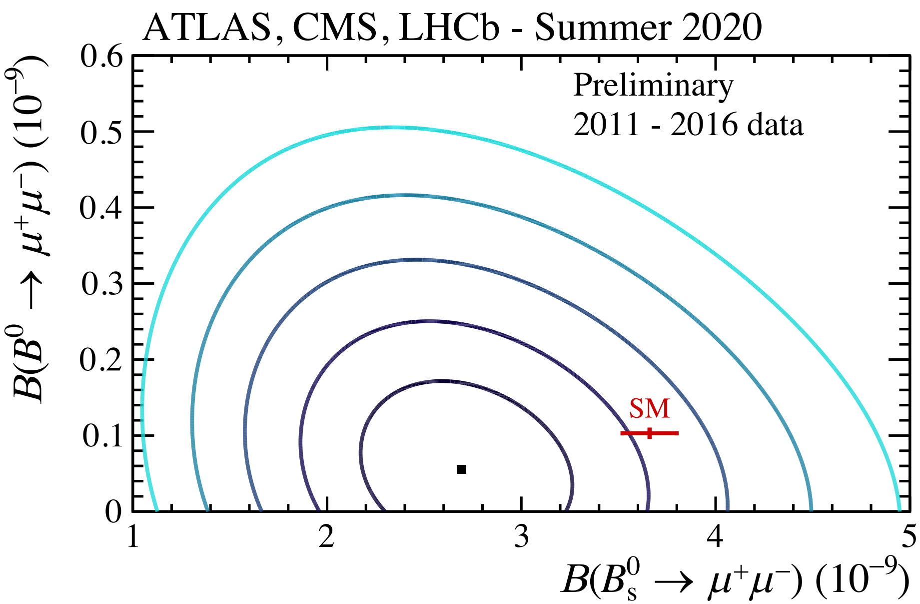
** LHCb Collaboration, https://arxiv.org/abs/2108.09284

† LHCb Collaboration, https://arxiv.org/abs/2108.09283

LHC Combination – Branching Ratios

- The two-dimensional likelihood contours of the branching ratio results for the $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ decays for the three experiments and their combination are shown on the left
- The two-dimensional likelihood contours for the branching ratios of the combined results are shown in the plot on the right
- Both plots are compared to the current SM prediction, shown with the red cross.

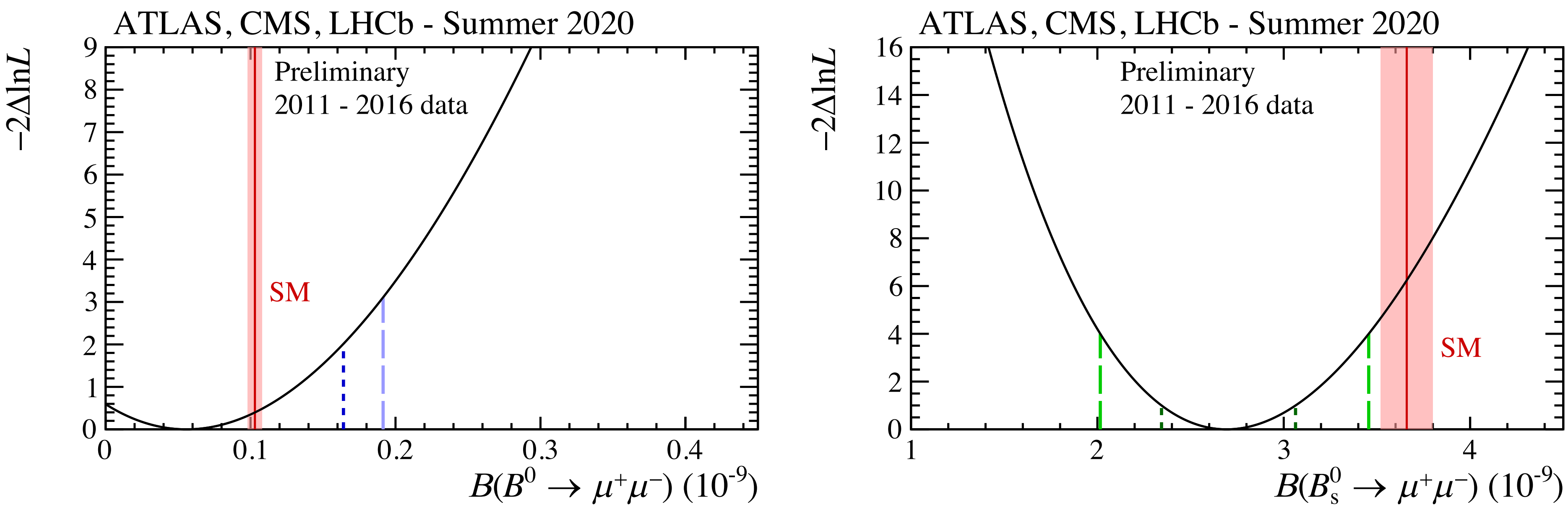




- The negative log-likelihood fit values of the LHC experiments combination results for (left) $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ and (right) $\mathcal{B}(B^0 \to \mu^+\mu^-)$ compared to the SM and its uncertainty (red line and band, respectively).
- The branching fraction measurements for $B_s^0 \to \mu^+\mu^-$ and the upper limits on the $B^0 \to \mu^+\mu^-$ at 90% CL and 95% CL are:

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = \left(2.69\,^{+\,0.37}_{-\,0.35}\right) \times 10^{-9}$$

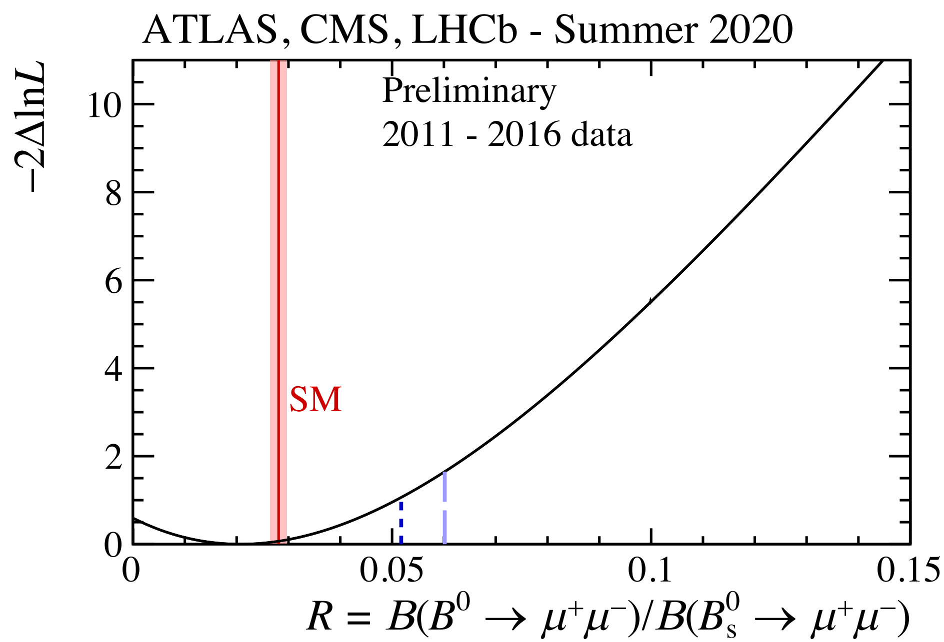
$$\begin{aligned} \mathcal{B}(B^0 \to \mu^+\mu^-) &< 1.6 \times 10^{-10} \text{ at 90\% CL, and} \\ \mathcal{B}(B^0 \to \mu^+\mu^-) &< 1.9 \times 10^{-10} \text{ at 95\% CL.} \end{aligned}$$

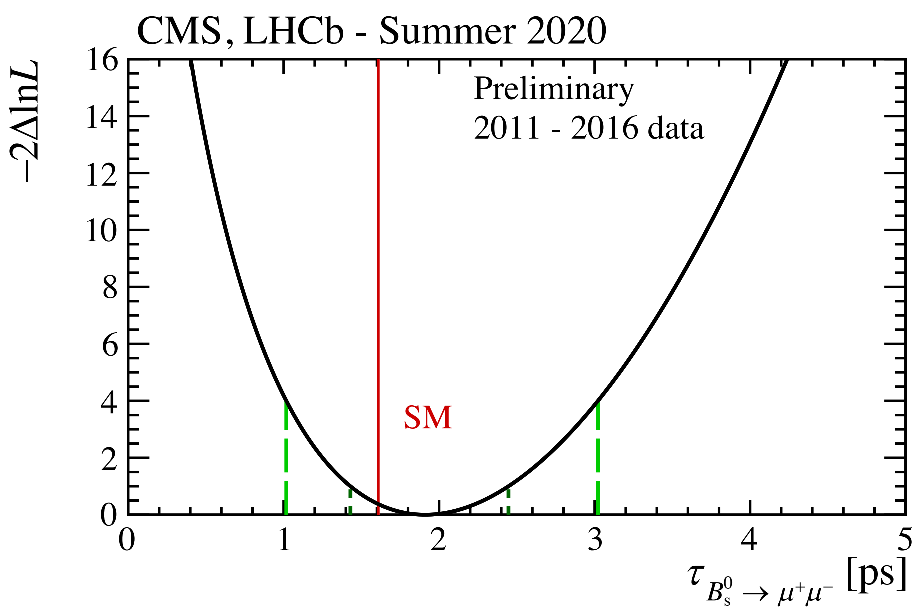


- The negative log-likelihood for the **ratio** of the $B^0 \to \mu^+\mu^-$ and $B_s^0 \to \mu^+\mu^-$ branching fractions of the LHC experiments combination results. The results are compared to the SM and its uncertainty (red line and band, respectively)

- The value of the ratio is determined to be:

$$\mathcal{R} = 0.021^{+0.030}_{-0.025}$$





- The negative log-likelihood, for the combination of the CMS and LHCb measurements of the $B_s^0 \to \mu^+\mu^-$ effective lifetime, shown as a solid black line on the left

- The minimized $-2\Delta\ln L$ combined value is:

$$\tau_{B_s^0 \to \mu^+\mu^-} = 1.91^{+0.37}_{-0.35}\,\text{ps}$$

$B_d^0 \to K^{*0}\mu^+\mu^-$

- ATLAS: JHEP 10 (2018) 047

https://doi.org/10.1007/JHEP10(2018)047

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

- CMS: PLB 781 (2018) 517

https://doi.org/10.1016/j.physletb.2018.04.030

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

and

$B^+ \to K^{*+}\mu^+\mu^-$

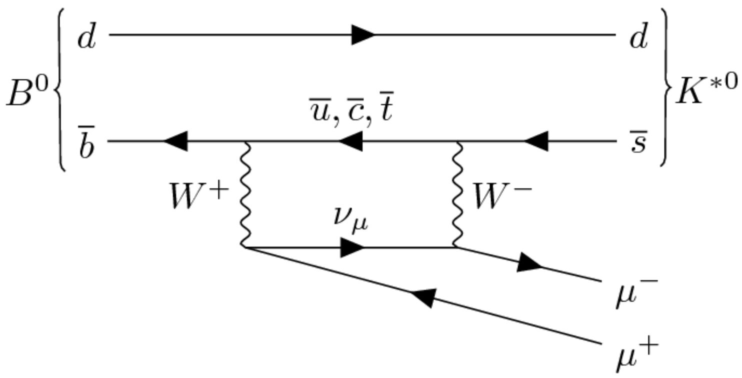
- CMS: JHEP 04 (2021) 124

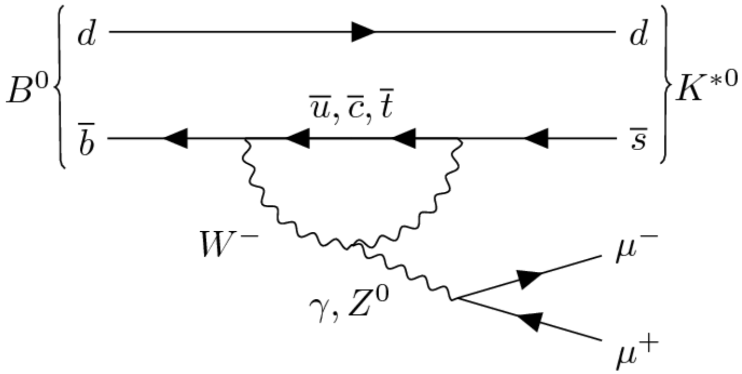
https://doi.org/10.1007/JHEP04(2021)124

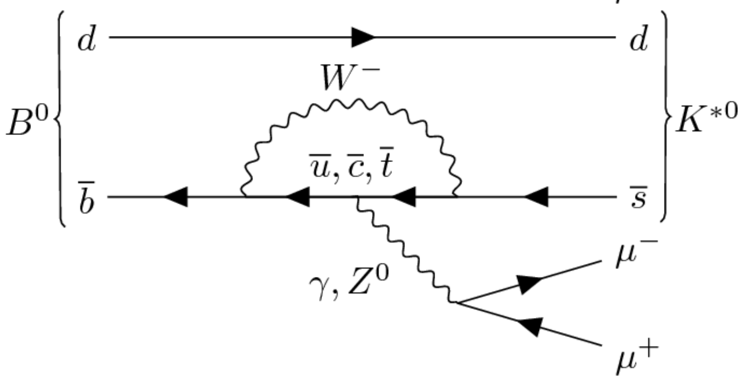
20.0 fb^{-1} of $\sqrt{s}$ = 8 TeV (2012)

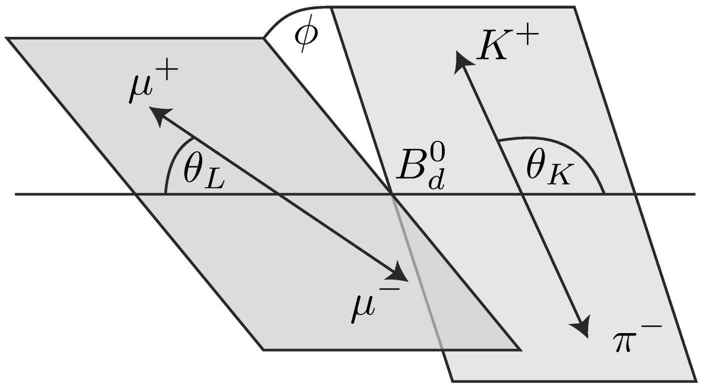
There is a related CMS analysis of $B^+ \to K^+\mu^+\mu^-$: PRD 98 (2018) 112011 https://doi.org/10.1103/PhysRevD.98.112011

Angular Analysis $B_d^0 \to K^{*0}\mu^+\mu^-$









$B_d^0 \to K^{*0}(\to K^+\pi^-)\,\mu^+\mu^-$

- Search for heavy new particles that may contribute to flavor changing neutral current (FCNC) decay amplitudes
- The lowest order Feynman diagrams for $B_d^0 \to K^*\mu^+\mu^-$ are the box diagram shown on the upper left and the two penguin diagrams
- The angular parameters of the measurement are shown in the upper right diagram. The angles θ_K, θ_L, and ϕ are measured in the rest frame of the K^{*0}, dimuon system, and B_d^0 meson, respectively.

Analysis Scheme

- The differential decay amplitude of $B_d^0 \to K^{*0}\mu^+\mu^-$ can be written in terms of:
  - $q^2 = 4m_\mu^2$, $\cos(\theta_K)$, $\cos(\theta_L)$ and ϕ
  - The fraction of longitudinally polarized K^* mesons (F_{L})
  - And 7 angular coefficients, S_i where i = 3,4,5,6,7,8,9
  - Theoretical uncertainties can be reduced in the decay amplitude using ratios of F_{L} and S_i to form P_1, P_2, P_3 and P_j' where j = 4,5,6,8
- The equation can be simplified using trigonometric transformations to "fold" certain angular distributions so that only 3 coefficients remain in the decay amplitude

Folding Schemes:

$$F_L, S_3, S_4, P_4' : \begin{cases} \phi \to -\phi & \text{for } \phi < 0 \\ \phi \to \pi - \phi & \text{for } \theta_L > \frac{\pi}{2} \\ \theta_L \to \pi - \theta_L & \text{for } \theta_L > \frac{\pi}{2}, \end{cases}$$

$$F_L, S_3, S_5, P_5' : \begin{cases} \phi \to -\phi & \text{for } \phi < 0 \\ \theta_L \to \pi - \theta_L & \text{for } \theta_L > \frac{\pi}{2}, \end{cases}$$

$$F_L, S_3, S_7, P_6' : \begin{cases} \phi \to \pi - \phi & \text{for } \phi > \frac{\pi}{2} \\ \phi \to -\pi - \phi & \text{for } \phi < -\frac{\pi}{2} \\ \theta_L \to \pi - \theta_L & \text{for } \theta_L > \frac{\pi}{2}, \end{cases}$$

$$F_L, S_3, S_8, P_8' : \begin{cases} \phi \to \pi - \phi & \text{for } \phi > \frac{\pi}{2} \\ \phi \to -\pi - \phi & \text{for } \phi < -\frac{\pi}{2} \\ \theta_L \to \pi - \theta_L & \text{for } \theta_L > \frac{\pi}{2} \\ \theta_K \to \pi - \theta_K & \text{for } \theta_L > \frac{\pi}{2}. \end{cases}$$

Resulting angular variable ranges:

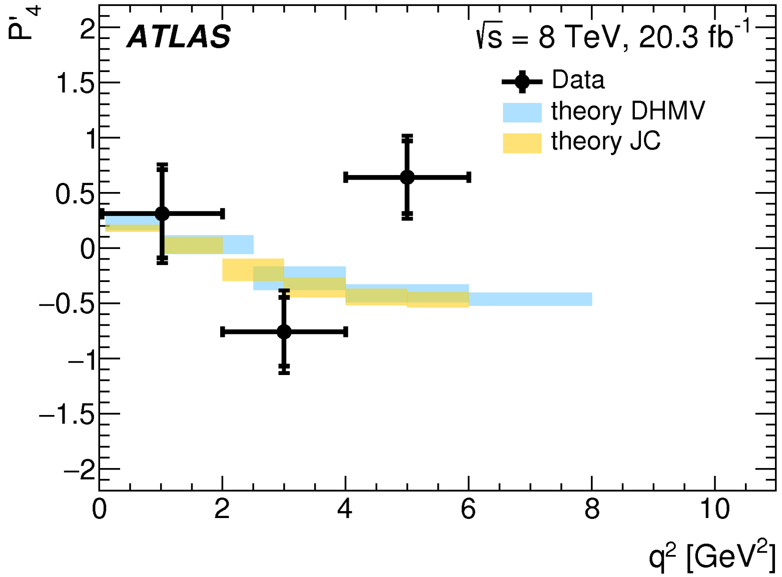
$$\cos\theta_L \in [0,1], \quad \cos\theta_K \in [-1,1] \quad \text{and} \quad \phi \in [0,\pi],$$
$$\cos\theta_L \in [0,1], \quad \cos\theta_K \in [-1,1] \quad \text{and} \quad \phi \in [0,\pi],$$
$$\cos\theta_L \in [0,1], \quad \cos\theta_K \in [-1,1] \quad \text{and} \quad \phi \in [-\pi/2,\pi/2],$$
$$\cos\theta_L \in [0,1], \quad \cos\theta_K \in [-1,1] \quad \text{and} \quad \phi \in [-\pi/2,\pi/2],$$

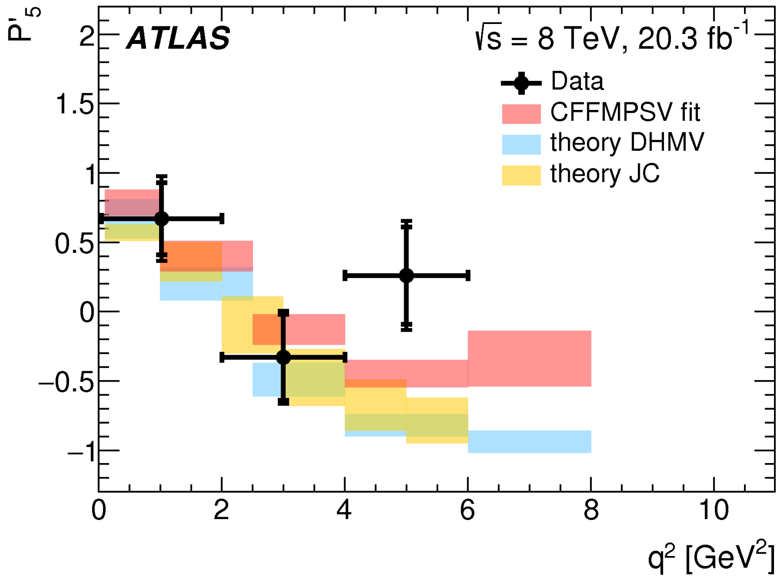
Extended Maximum Likelihood Fit

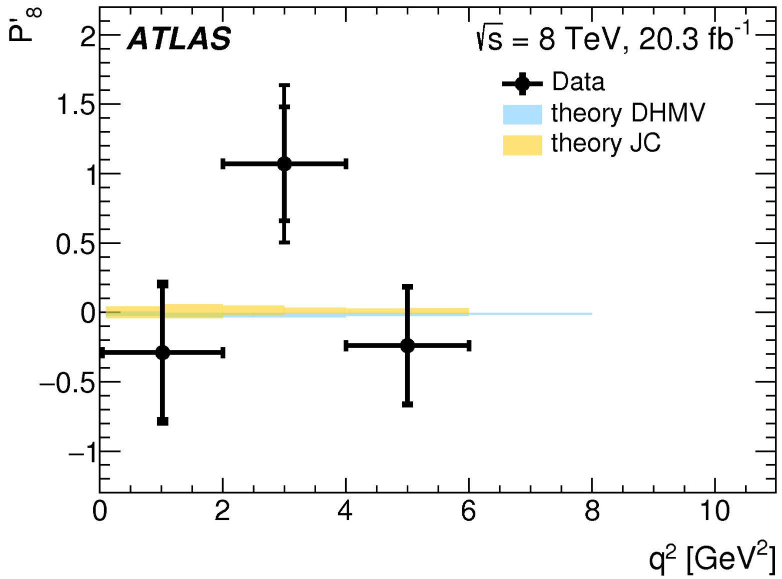
- An extended unbinned maximum-likelihood fit of the angular distribution of the signal decay is performed in order to extract the 3 coefficients of interest in a particular folding scheme:
  - F_{L}, S_3 and S_j where j = 4,5,7,8 (or F_{L}, S_3 and P_j' where j = 4,5,6,8)
  - Nuisance parameters (mass of $K\pi\mu\mu$ and mass width coefficient) are also extracted
- For ATLAS, the fitting procedure is performed in 6 bins of q^2 between 0.04 and 6 GeV^2 for each folding scheme in order to probe the dependence on q^2
- For CMS, an unbinned maximum-likelihood fit is performed in 7 bins between 1 and 19 GeV^2

ATLAS Results

- There is good agreement between theory and measurement for all regions except:
  - the P_4' and P_5' parameters in $q^2 \in [4.0, 6.0]$ GeV^2, each 2.7σ deviation from the SM (DHMV)
  - and P_8' parameter in $q^2 \in [2.0, 4.0]$ GeV^2
- The results of theoretical approaches CFFMPSV*, DHMV** and JC†, ‡ are shown







* M. Ciuchini et al., JHEP 06 (2016) 116 https://doi.org/10.1007/JHEP06(2016)116

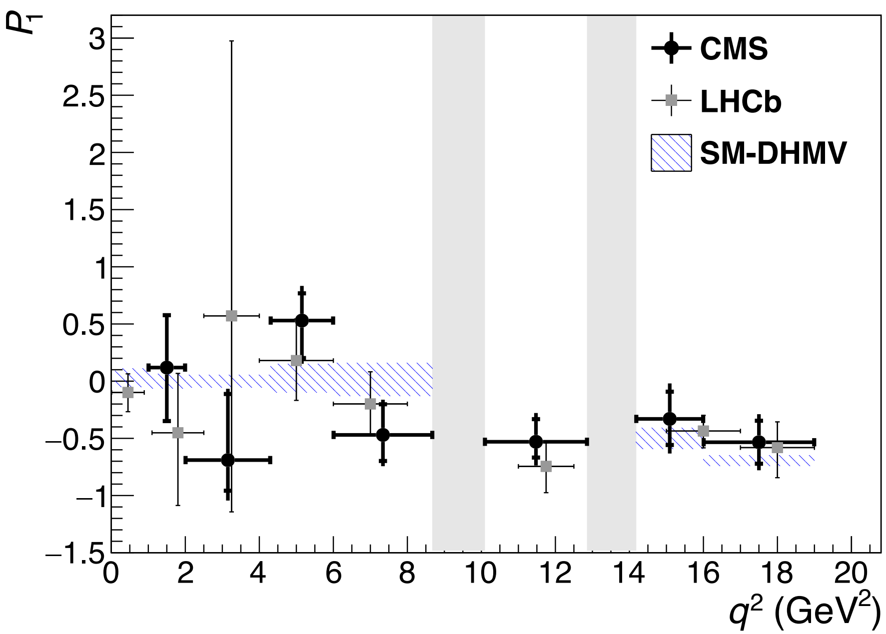
** S. Descotes-Genon, L. Hofer, J. Matias and J. Virto, JHEP 12 (2014) 125, https://doi.org/10.1007/JHEP12(2014)125

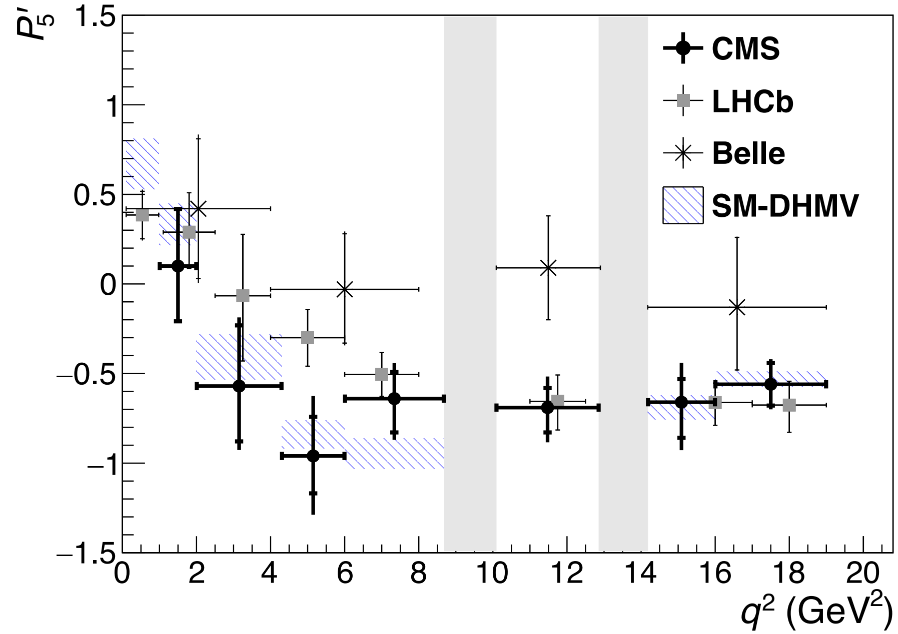
† S. Jäger and J. Martin Camalich, JHEP 05 (2013) 043, https://doi.org/10.1007/JHEP05(2013)043

‡ S. Jäger and J. Martin Camalich, Phys. Rev. D 93 (2016) 014028, https://doi.org/10.1103/PhysRevD.93.014028

CMS Results

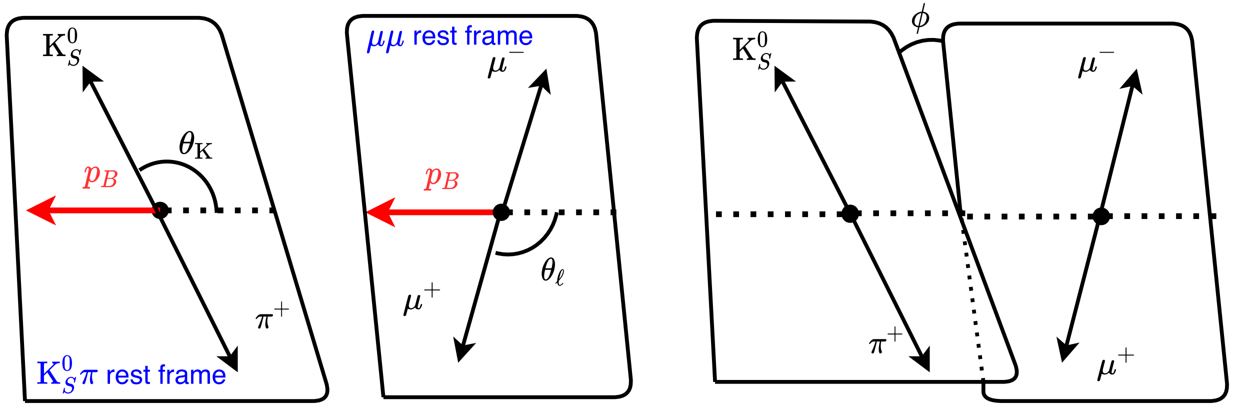
- CMS measurements of the (left plot) P_1 and (right plot) P_5' angular parameters versus q^2 for $B_d^0 \to K^{*0}\mu^+\mu^-$ decays, in comparison to results from the LHCb* and Belle Collaborations.
- The vertical shaded regions correspond to the J/ψ and ψ' resonances
- The results are compared to the Standard Model calculations (SM-DHMV[†, ‡]), averaged over each q^2 bin.





* LHCb Collaboration, J. High Energy Phys. 02 (2016) 104, https://doi.org/10. 1007/JHEP02(2016)104
** Belle Collaboration, Phys. Rev. Lett. 118 (2017) 111801, https://doi.org/10.1103/ PhysRevLett.118.111801
† S. Descotes-Genon, J. Matias, M. Ramon, J. Virto, J. High Energy Phys. 01 (2013) 048, https://doi.org/10.1007/JHEP01(2013)048
‡ S. Descotes-Genon, T. Hurth, J. Matias, J. Virto, J. High Energy Phys. 05 (2013) 137, https://doi.org/10.1007/JHEP05(2013)137

$B^+ \to K^{*+}\mu^+\mu^-$ Analysis



$B^+ \to K^{*+}\left(\to K_S^0\pi^+\right)\mu^+\mu^-$

- Using heavy-flavor hadrons to search for effects of unknown heavy particles that might modify the standard model (SM) prediction.
  - Probing FCNC involving $b \to s\mu^+\mu^-$
  - Differences can appear in the overall decay rate or as modifications to the angular distributions of the decay products
- Definition of the angular observables is provided in the image:
  - θ_K (left), θ_l (middle), and ϕ (right)

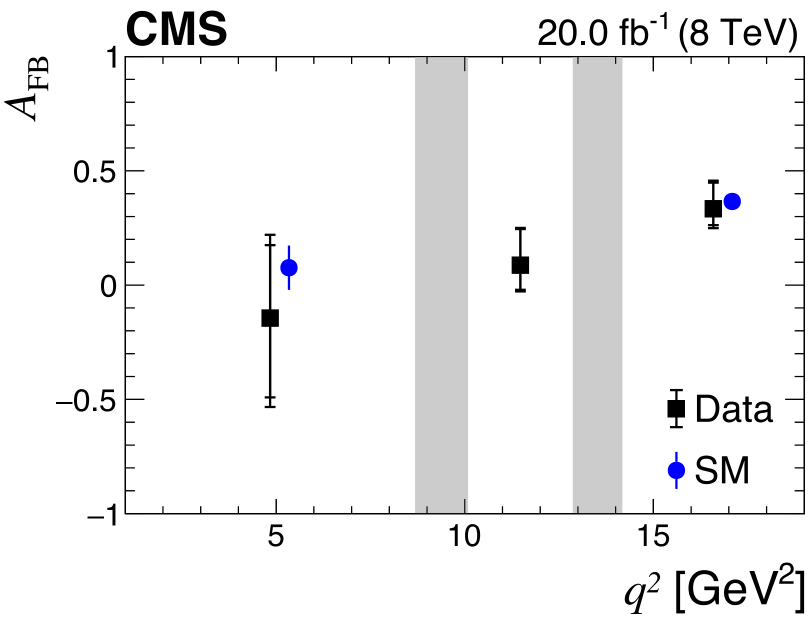
Measurement

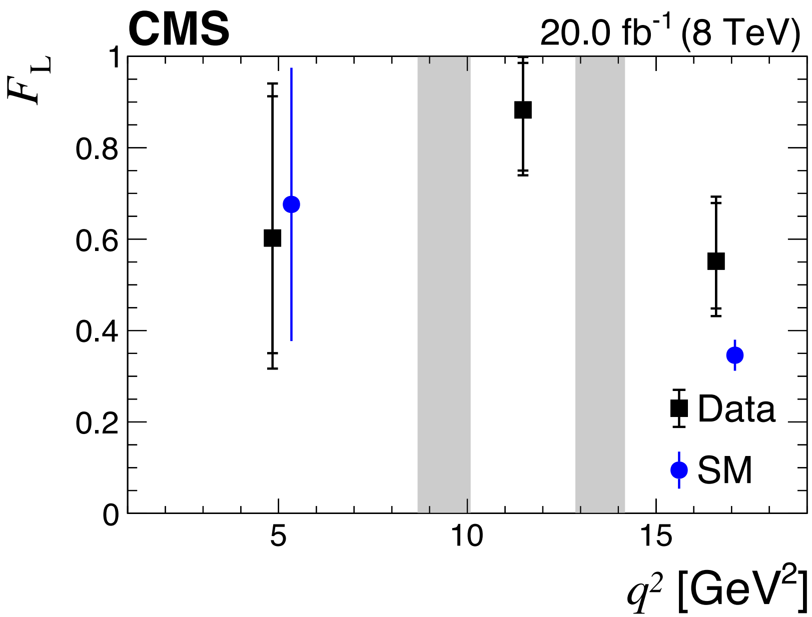
- This CMS paper* reports the first measurement of A_{FB} and F_{L} in the exclusive decay $B^{+} \to K^{*+}\mu^{+}\mu^{-}$
    - A_{FB} is the muon forward-backward asymmetry (given by the normalised difference between the number of positive muons going in the forward and backward directions with respect to the direction opposite to B^{+} momentum in the dimuon rest frame.)
    - F_{L} is the fraction of longitudinally polarised K^{*+} mesons
- For each q^2 bin, the observables A_{FB} and F_{L} are extracted by performing an unbinned extended maximum likelihood fit with three independent variables: m, $\cos(\theta_K)$, $\cos(\theta_l)$
- Fits to the data are performed in three independent q^2 bins between 1 and 19 GeV^2

* CMS Collaboration, JHEP 04 (2021) 124 , https://doi.org/10.1007/JHEP04(2021)124

$B^+ \to K^{*+}\mu^+\mu^-$ Results

- The measured values of A_{FB} (left) and F_{L} (right) versus q^2 for $B^+ \to K^{*+}\mu^+\mu^-$ decays are shown with filled squares, centered on the q^2 bin.
- The statistical (total) uncertainty is shown by inner (outer) vertical bars.
- The vertical shaded regions correspond to the regions dominated by $B^+ \to K^{*+}J/\psi$ and $B^+ \to K^{*+}\psi(2S)$ decays
- The SM predictions* and associated uncertainties are shown by the filled circles and vertical bars, with the points slightly offset from the center of the q^2 bin for clarity.





* S. Descotes-Genon, L. Hofer, J. Matias and J. Virto, JHEP 12 (2014) 125, https://doi.org/10.1007/JHEP12(2014)125

Summary

- Results from ATLAS and CMS have been presented for

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

ATLAS

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

CMS

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

Summer 2020 LHC combination

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = \left(2.69\,^{+\,0.37}_{-\,0.35}\right) \times 10^{-9}$$

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

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

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

$$\mathcal{B}(B^0 \to \mu^+\mu^-) \quad < \quad 1.9 \times 10^{-10}$$

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

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

$$\tau_{B_s^0 \to \mu^+\mu^-} = 1.91^{+0.37}_{-0.35}\ \mathrm{ps}$$

4. The angular analyses of the $B_d^0 \to K^{*0}(892)\mu^+\mu^-$ and $B^+ \to K^*(892)^+\mu^+\mu^-$

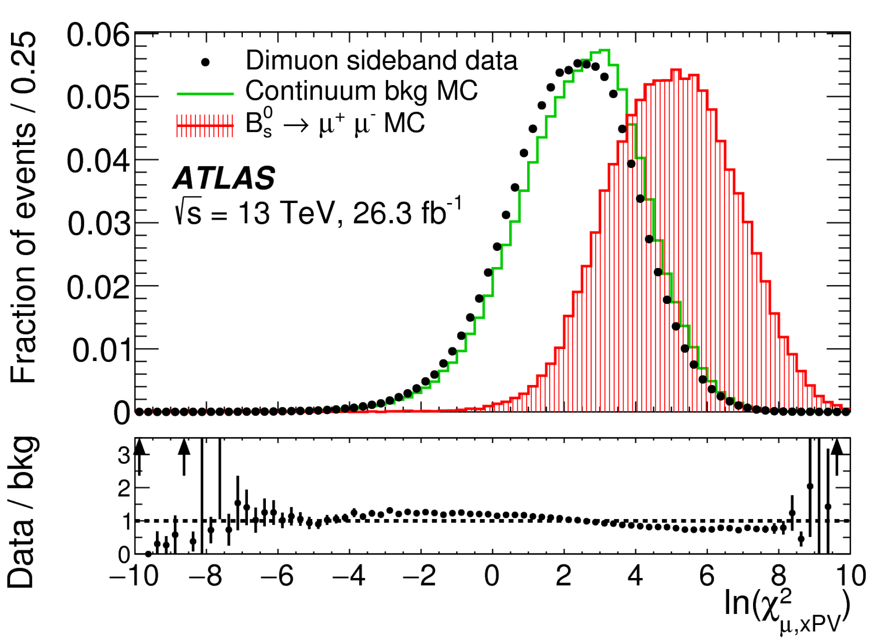
There is a related CMS analysis of $B^+ \to K^+\mu^+\mu^-$: PRD 98 (2018) 112011 https://doi.org/10.1103/PhysRevD.98.112011

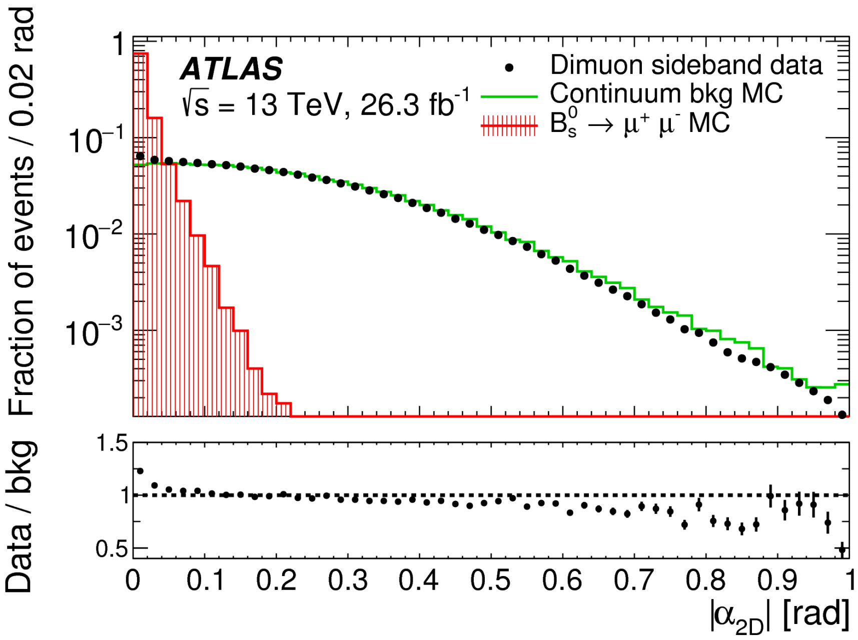
Additional Slides

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



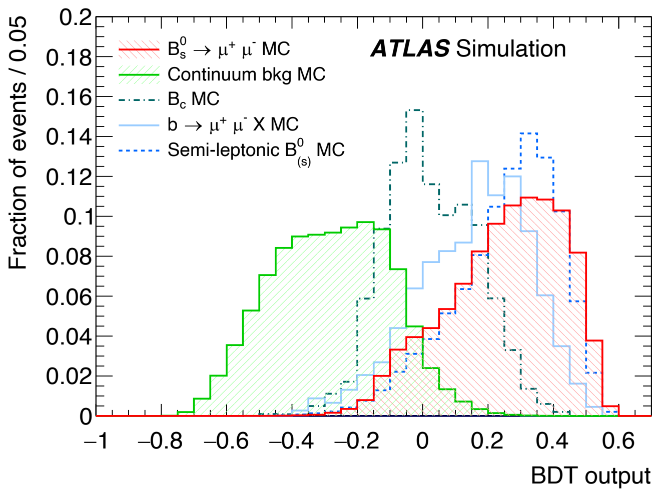


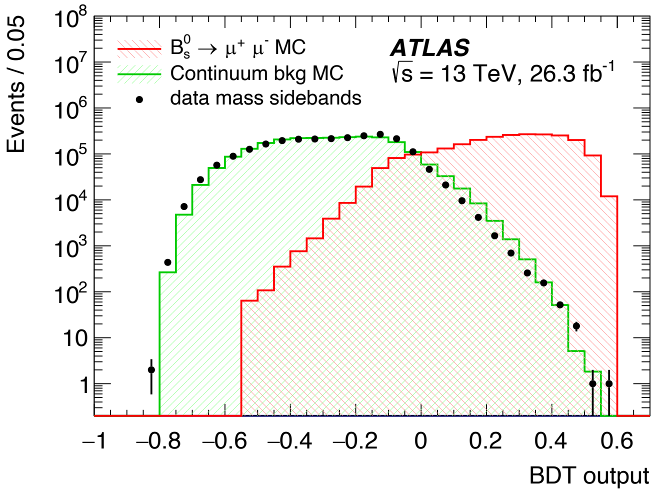
- Shown here are two of the input variables used in the training
- Care is taken to ensure that BDT output is not correlated with the invariant mass of the muons

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.

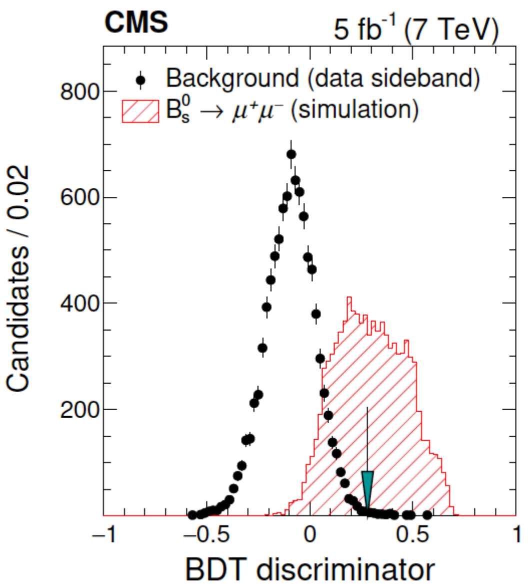


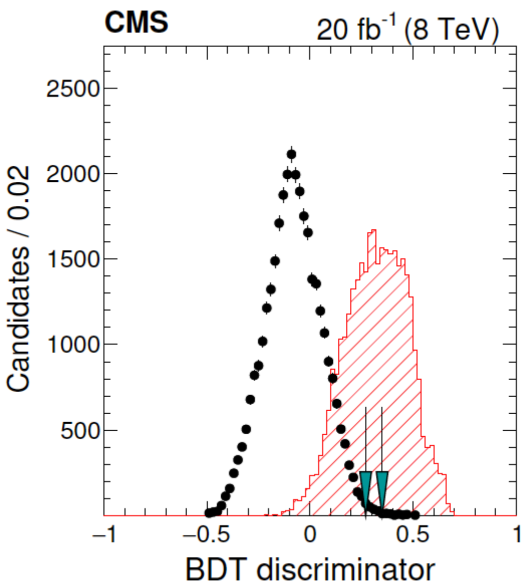


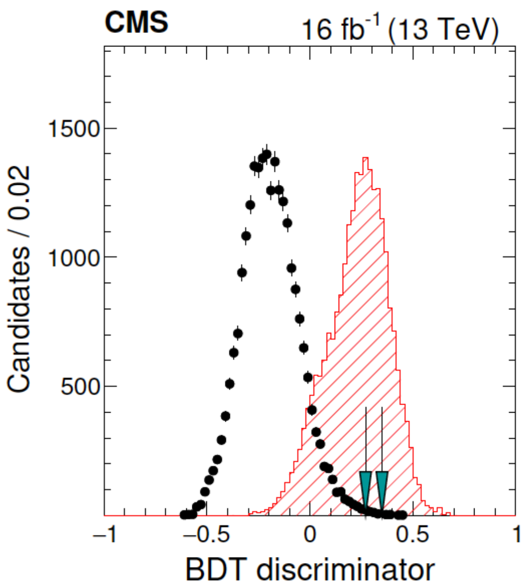
BDT Background Suppression

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:

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

- 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 $\left|\eta_{\mu_{1,2}}\right| < 2.5$
  - 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

$B_d^0 \to K^{*0}(892)\mu^+\mu^-$

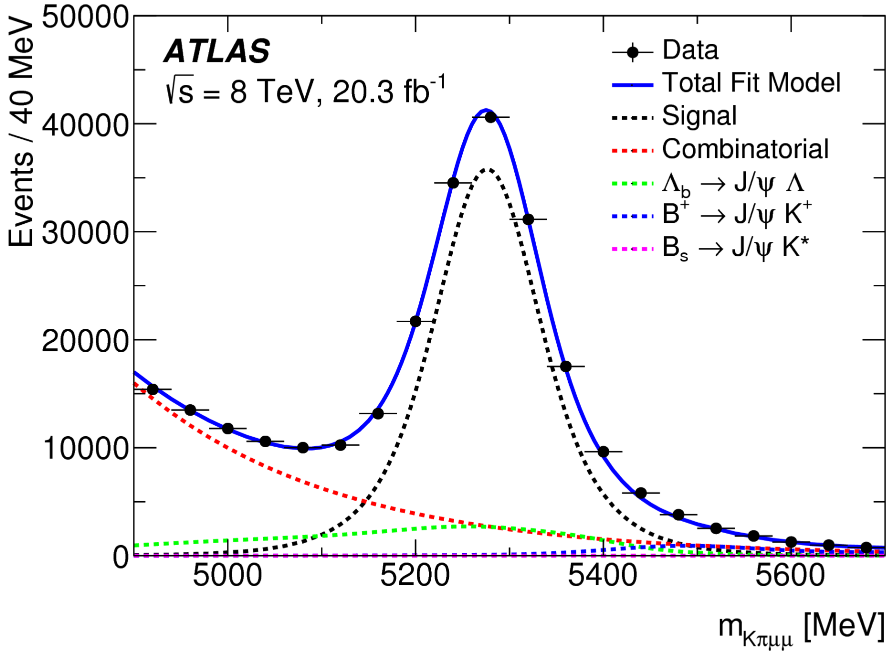
- differential decay rate as a function of the angular parameters

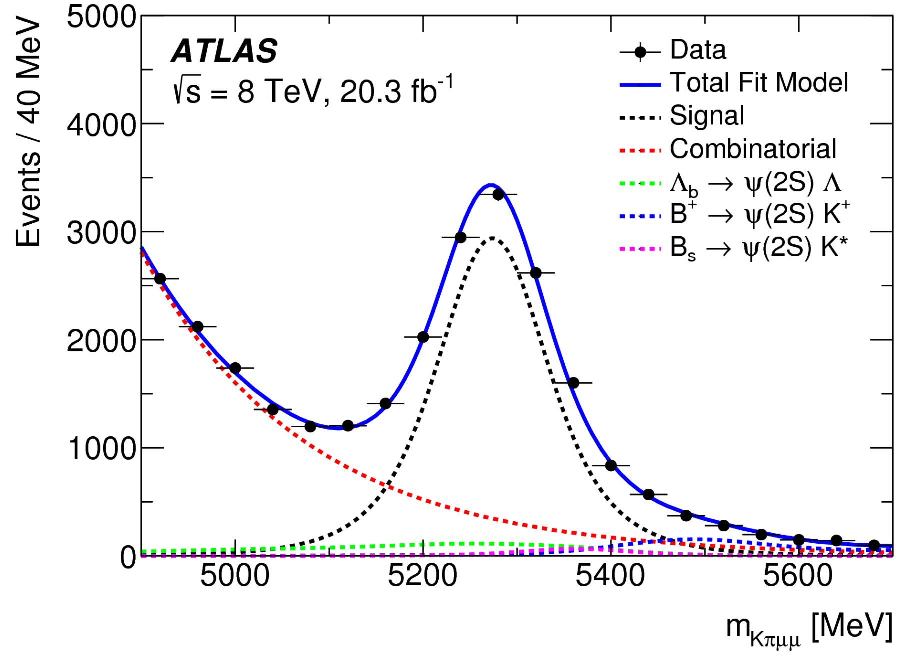
$$
\begin{aligned}
\frac{1}{\mathrm{d}\Gamma/\mathrm{d}q^2}\frac{\mathrm{d}^4\Gamma}{\mathrm{d}\cos\theta_L\,\mathrm{d}\cos\theta_K\,\mathrm{d}\phi\,\mathrm{d}q^2} = \frac{9}{32\pi}\Bigg[&\frac{3(1-F_L)}{4}\sin^2\theta_K + F_L\cos^2\theta_K + \frac{1-F_L}{4}\sin^2\theta_K\cos 2\theta_L \\
&- F_L\cos^2\theta_K\cos 2\theta_L + S_3\sin^2\theta_K\sin^2\theta_L\cos 2\phi \\
&+ S_4\sin 2\theta_K\sin 2\theta_L\cos\phi + S_5\sin 2\theta_K\sin\theta_L\cos\phi \\
&+ S_6\sin^2\theta_K\cos\theta_L + S_7\sin 2\theta_K\sin\theta_L\sin\phi \\
&+ S_8\sin 2\theta_K\sin 2\theta_L\sin\phi + S_9\sin^2\theta_K\sin^2\theta_L\sin 2\phi\Bigg]. \qquad (2.1)
\end{aligned}
$$

$$
\begin{aligned}
P_1 &= \frac{2S_3}{1-F_L} \\
P_2 &= \frac{2}{3}\frac{A_{\mathrm{FB}}}{1-F_L} \\
P_3 &= -\frac{S_9}{1-F_L} \\
P'_{j=4,5,6,8} &= \frac{S_{i=4,5,7,8}}{\sqrt{F_L(1-F_L)}}.
\end{aligned}
$$

Control Regions

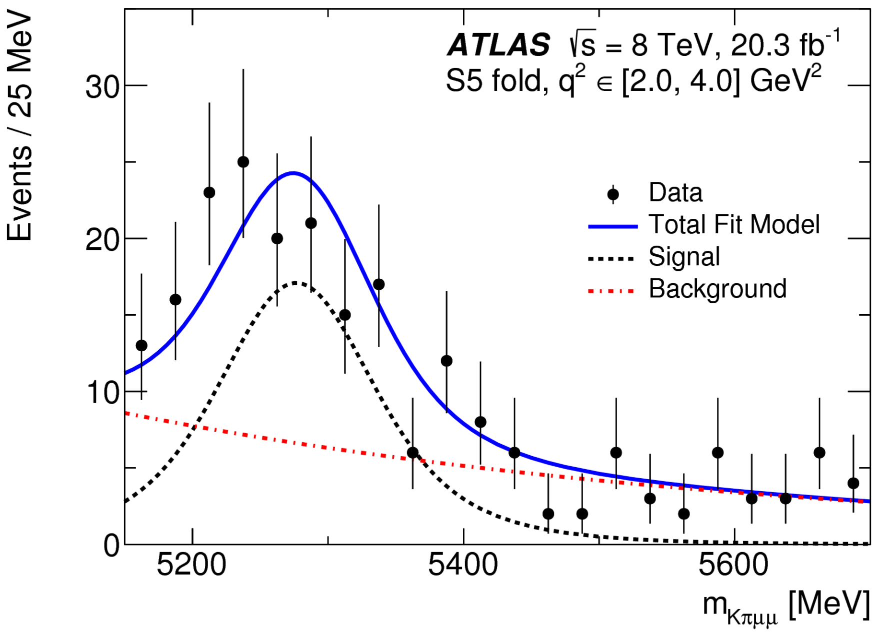
- Two $K^*c\bar{c}$ decay control sample fits, K^*J/ψ and $K^*\psi(2S)$, are shown in $q^2 \in [8, 11]$ and $[12, 15]$ GeV^2 regions, respectively
- Control samples are used to extract values for nuisance parameters describing the signal
- The fit to data includes a combinatorial background component that does not peak in the $m_{K\pi\mu\mu}$ distribution

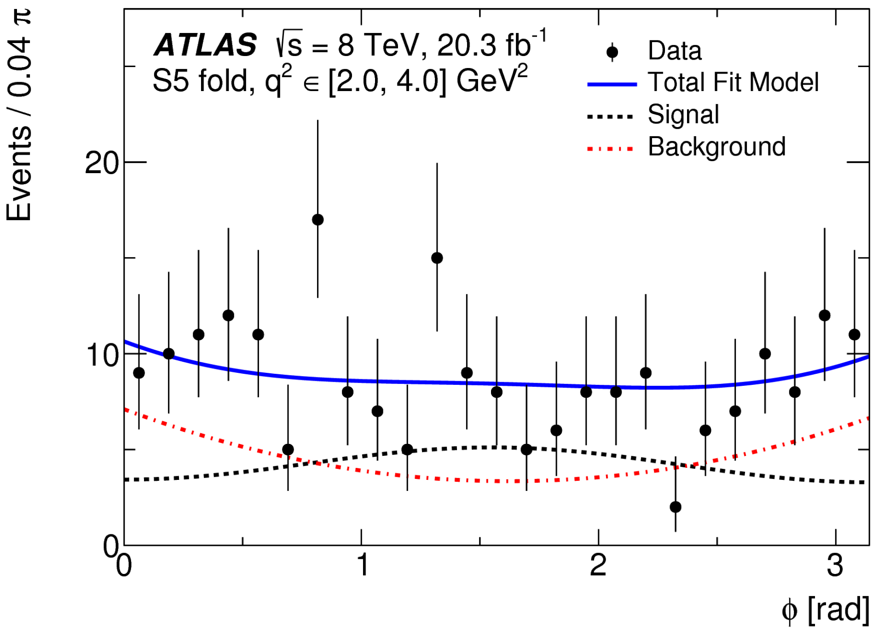




Subset of Results

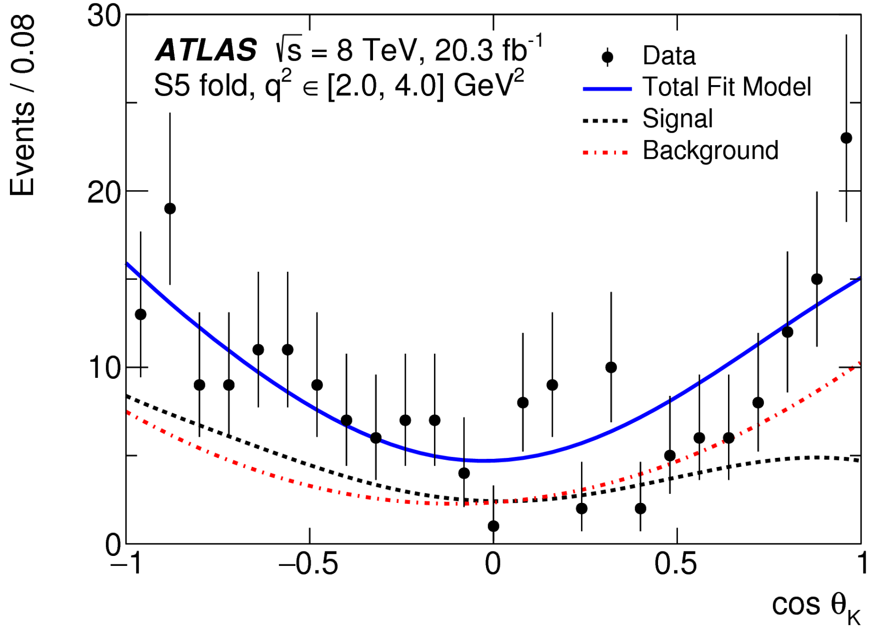
- Fit to the mass $K\pi\mu\mu$ and angle ϕ in the dilepton mass region $q^2 \in [2.0, 4.0]$ GeV^2
- The fit is performed using the F_L, S_3 and S_5 folding scheme

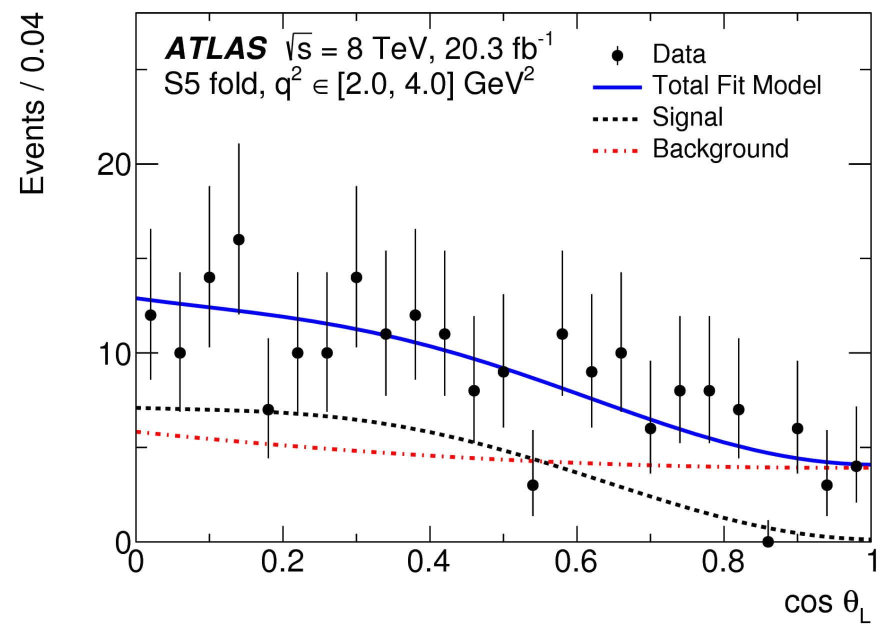




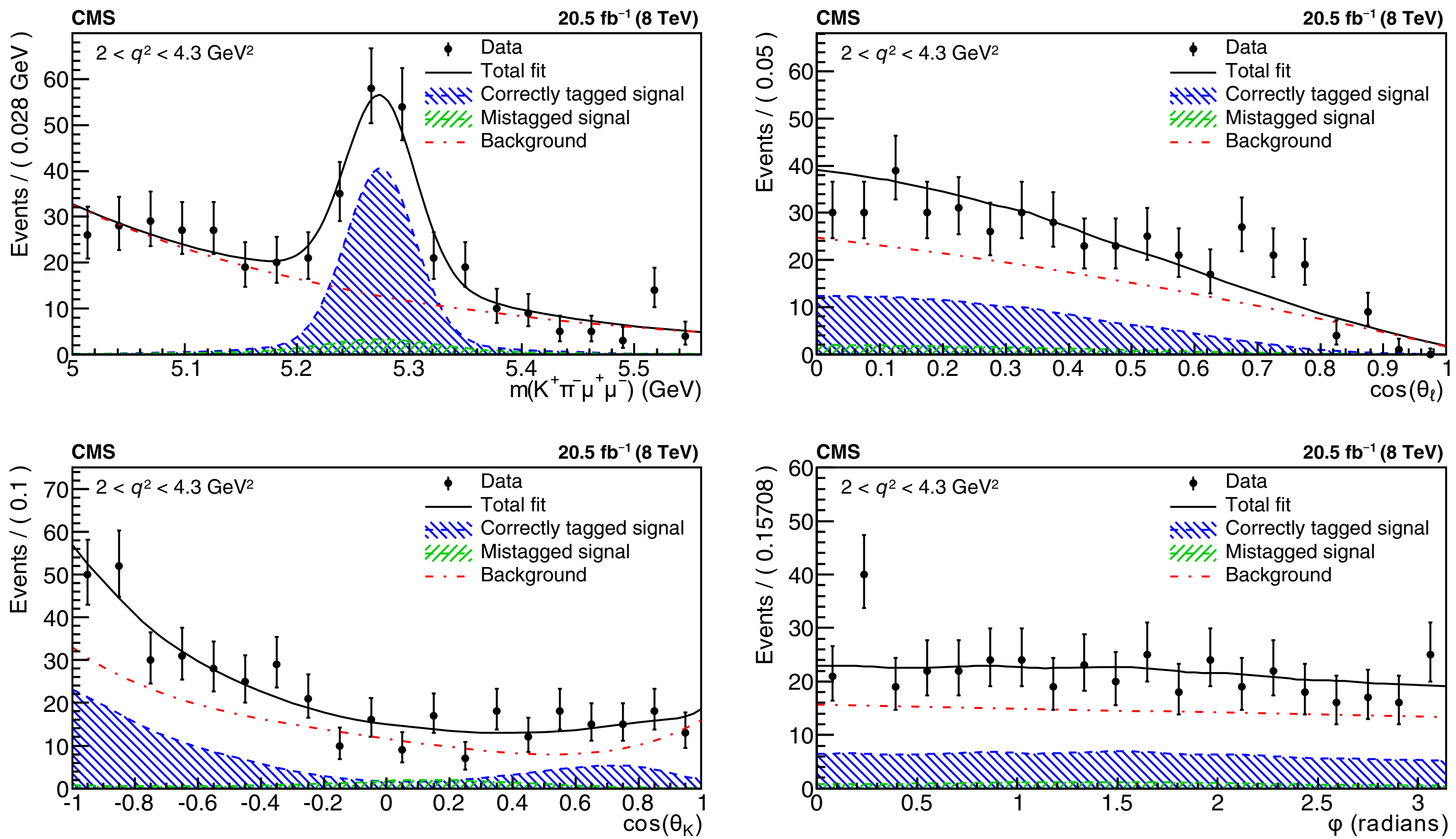
- Fitted signal and background yields are shown in the table for the various bins of q^2
- The fits to $\cos(\theta_K)$ and $\cos(\theta_L)$ in the $q^2 \in [2.0, 4.0]$ GeV^2 in the F_L, S_3 and S_5 folding scheme are shown here

q^2 [GeV^2]	n_{signal}	$n_{\text{background}}$
[0.04, 2.0]	128 ± 22	122 ± 22
[2.0, 4.0]	106 ± 23	113 ± 23
[4.0, 6.0]	114 ± 24	204 ± 26
[0.04, 4.0]	236 ± 31	233 ± 32
[1.1, 6.0]	275 ± 35	363 ± 36
[0.04, 6.0]	342 ± 39	445 ± 40

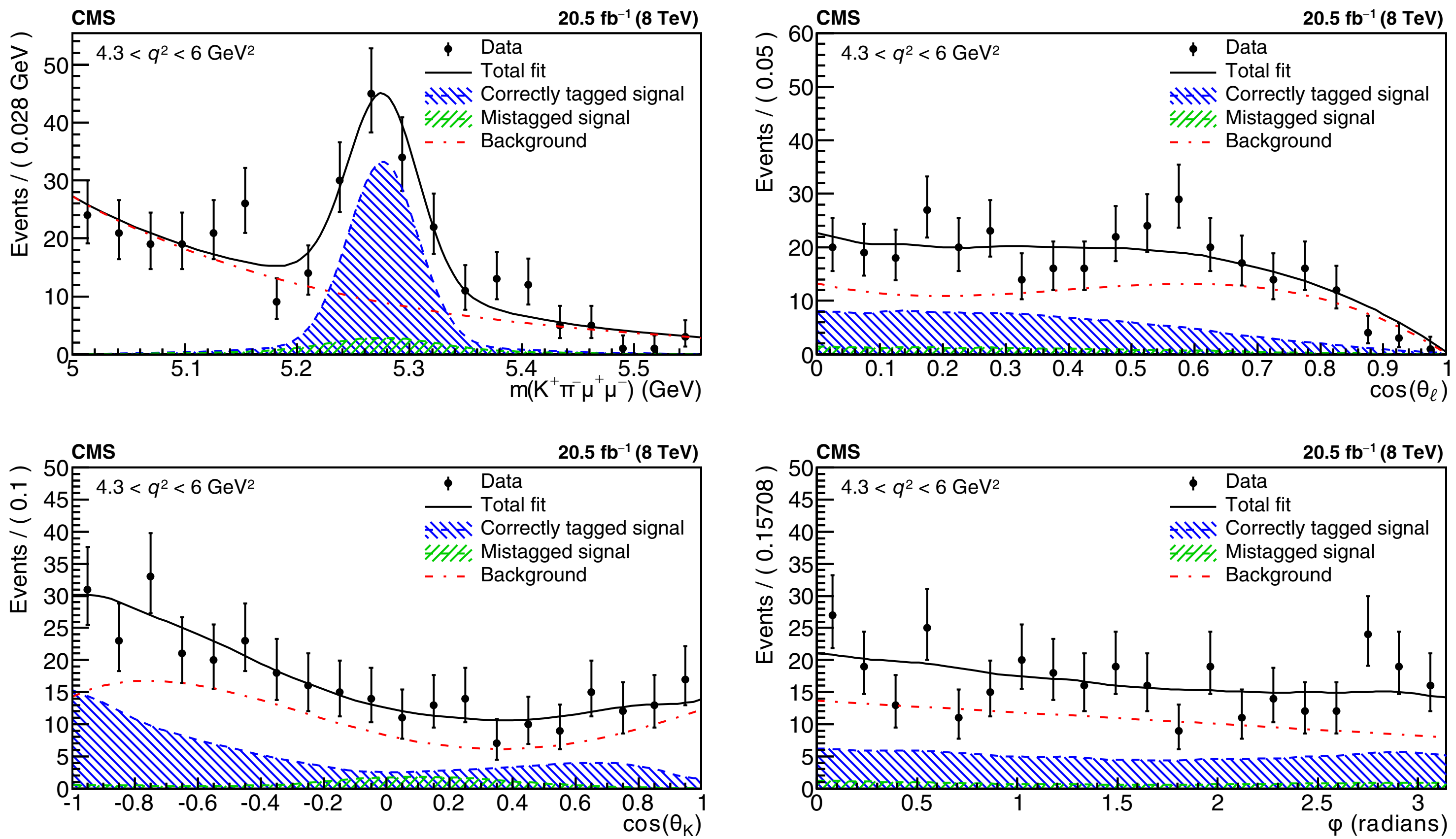




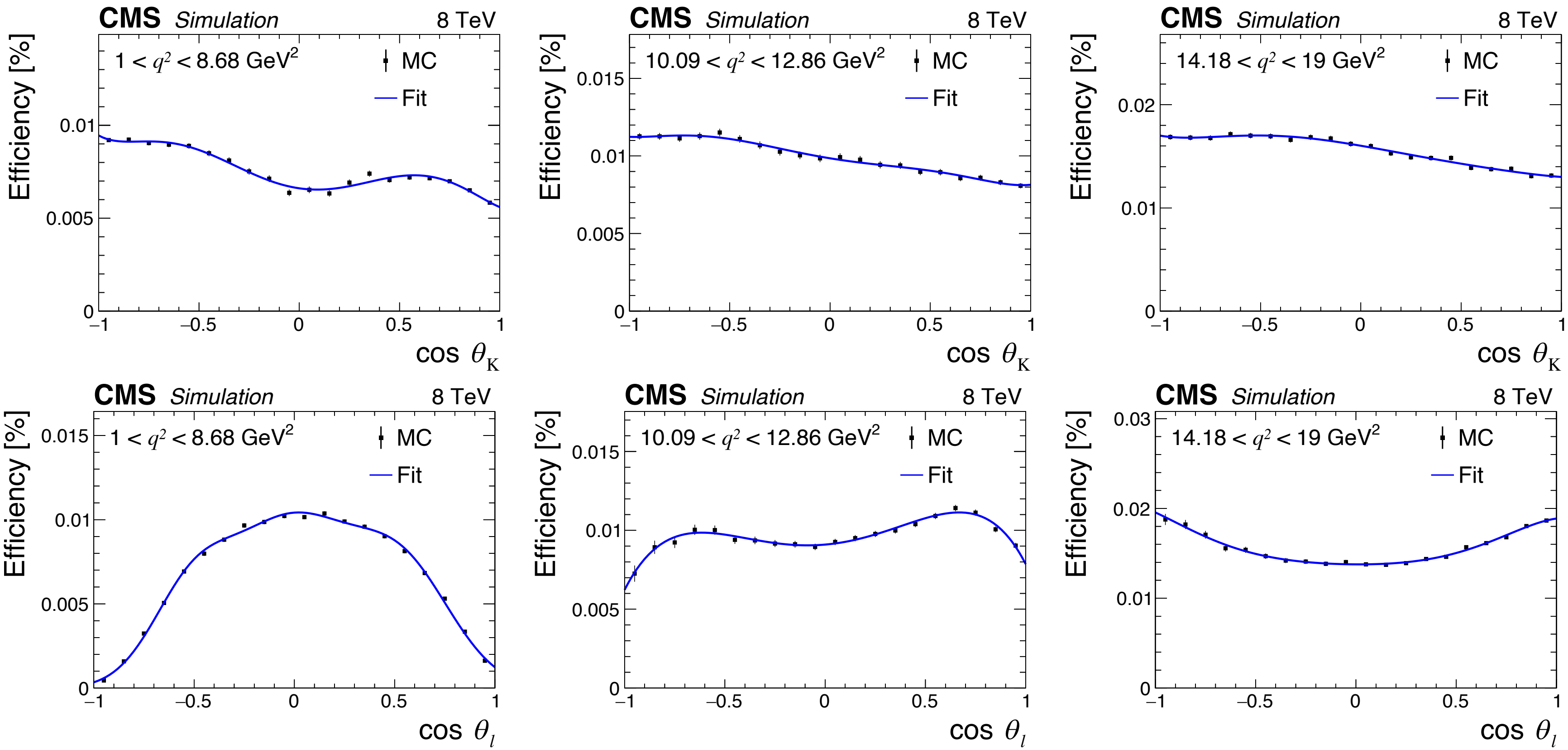
$$B_d^0 \rightarrow K^{*0}\mu^+\mu^-$$



$$B_d^0 \to K^{*0}\mu^+\mu^-$$

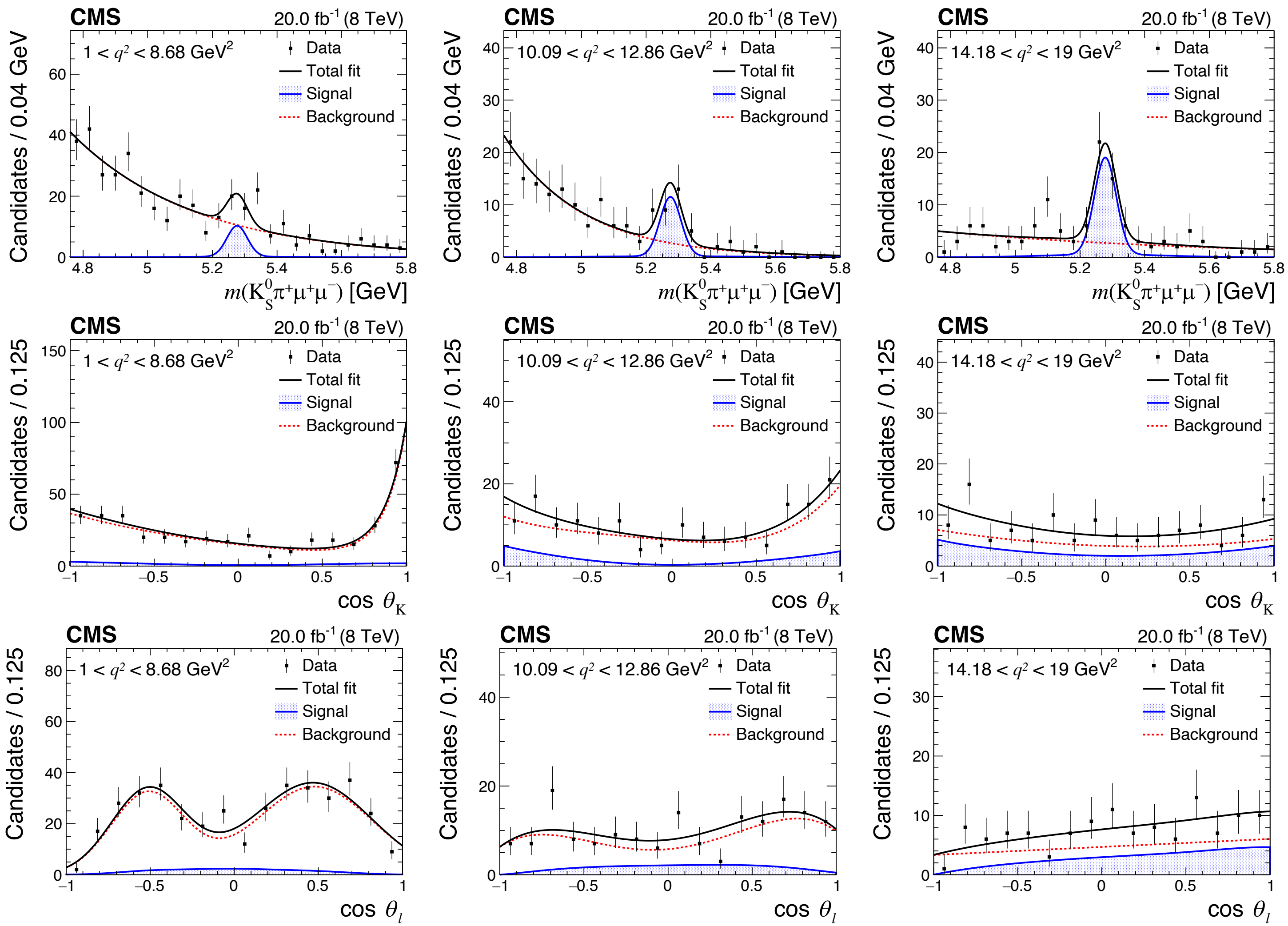


$B^+ \to K^{*+}\mu^+\mu^-$

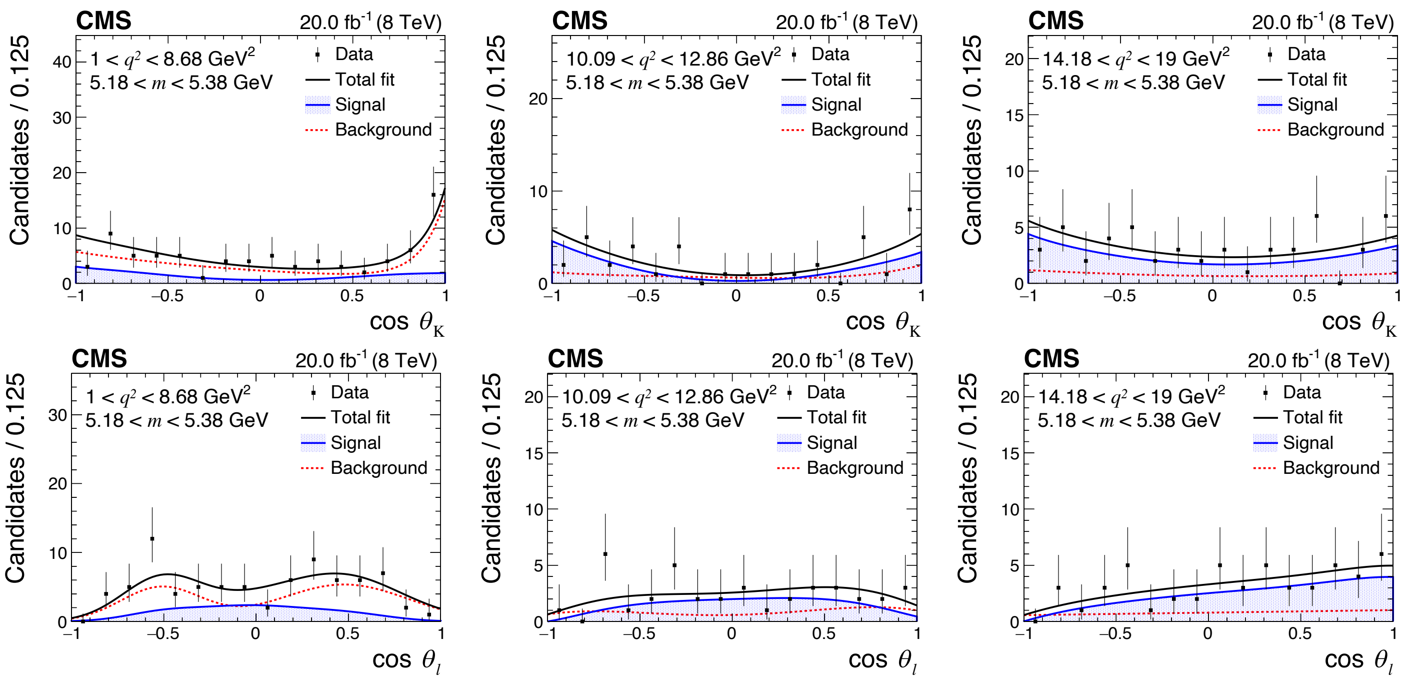


CMS Simulation 8 TeV
1 < q² < 8.68 GeV²
MC
Fit
Efficiency [%]
cos θ_K
CMS Simulation 8 TeV
10.09 < q² < 12.86 GeV²
MC
Fit
Efficiency [%]
cos θ_K
CMS Simulation 8 TeV
14.18 < q² < 19 GeV²
MC
Fit
Efficiency [%]
cos θ_K
CMS Simulation 8 TeV
1 < q² < 8.68 GeV²
MC
Fit
Efficiency [%]
cos θ_l
CMS Simulation 8 TeV
10.09 < q² < 12.86 GeV²
MC
Fit
Efficiency [%]
cos θ_l
CMS Simulation 8 TeV
14.18 < q² < 19 GeV²
MC
Fit
Efficiency [%]
cos θ_l

$$B^+ \to K^{*+}\mu^+\mu^-$$


CMS
20.0 fb^{-1} (8 TeV)
$1 < q^2 < 8.68$ GeV^2
$10.09 < q^2 < 12.86$ GeV^2
$14.18 < q^2 < 19$ GeV^2
Data
Total fit
Signal
Background
Candidates / 0.04 GeV
Candidates / 0.125
$m(K^0_S\pi^+\mu^+\mu^-)$ [GeV]
cos θ_K
cos θ_l

$B^+ \to K^{*+}\mu^+\mu^-$


CMS
20.0 fb-1 (8 TeV)
1 < q2 < 8.68 GeV2
5.18 < m < 5.38 GeV
10.09 < q2 < 12.86 GeV2
14.18 < q2 < 19 GeV2
Data
Total fit
Signal
Background
Candidates / 0.125
cos θK
cos θl