ATLAS/CMS rare decays

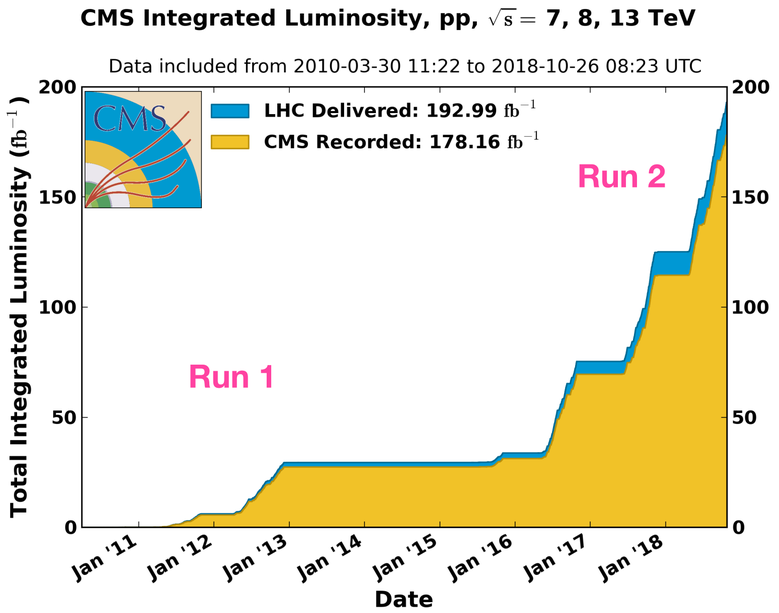
recent results and prospects

$B^0_{(s)} \to \mu^+\mu^-$ branching fraction
$B \to K^*\mu\mu$ angular analyses
$\tau \to 3\mu$

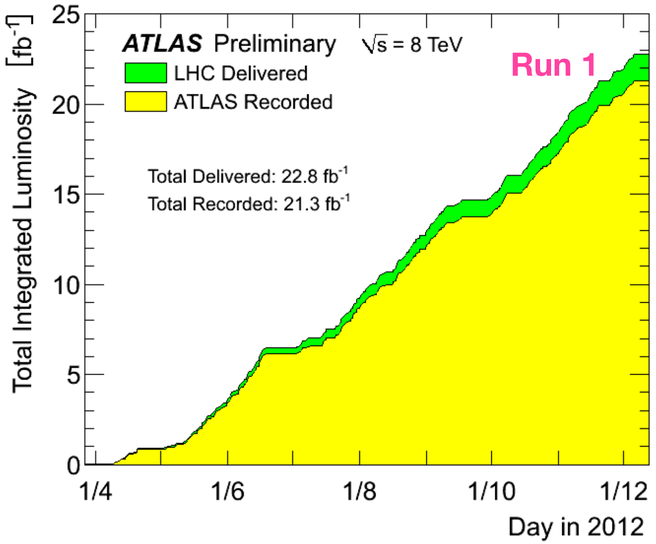
Camilla Galloni
for the ATLAS and CMS Collaborations

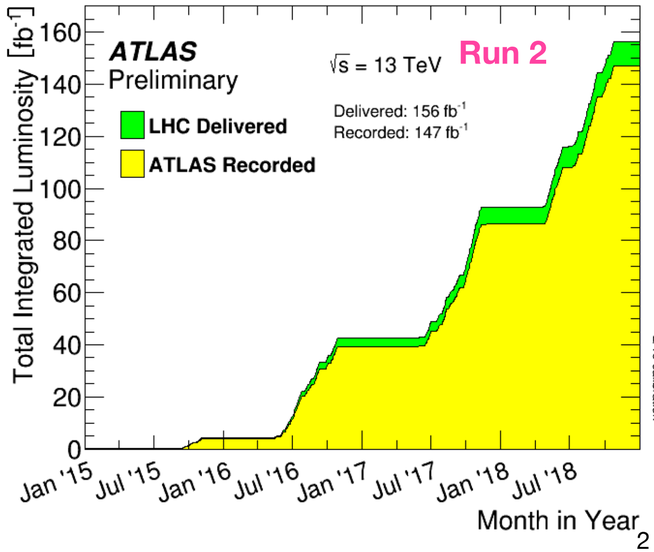
Outline

- Main results presented in this talk are:
  - $B^0_{(s)} \to \mu^+\mu^-$ branching fractions
  - $B \to K^*\mu\mu$ angular analyses
  - $\tau \to 3\mu$ decay



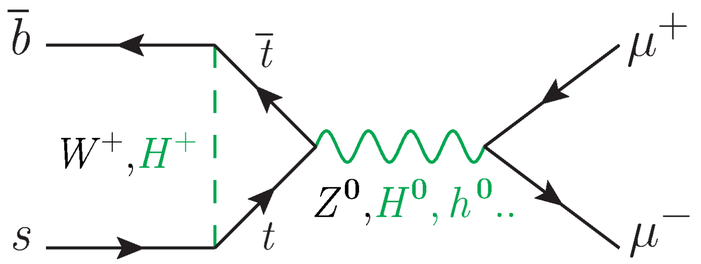
- They were produced by ATLAS and CMS using data acquired in different years or run periods:
  - Run 1 (2011-2012) 7-8 TeV
  - Run 2 (2015-2018)- 13 TeV
    - available results from 2016
- Prospects will be given for
  - HL-LHC - 3000/fb expected





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

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

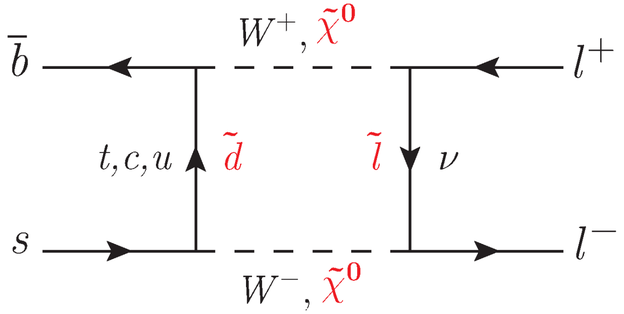


- In the standard model (SM) these decays are heavily suppressed:
  - FCNC - but also helicity-suppression $(m_\mu/m_B)^2$
  - High precision SM predictions of branching fractions for muonic $B^0_{(s)}$ decays:
    - $\mathcal{B}(B^0_s \to \mu^+\mu^-) = (3.66 \pm 0.14) \cdot 10^{-9}$
    - $\mathcal{B}(B^0 \to \mu^+\mu^-) = (1.03 \pm 0.005) \cdot 10^{-10}$

~4-5% precision

Flavour Lattice Averaging Group Collaboration, "FLAG Review 2019: Flavour Lattice Averaging Group (FLAG)", Eur. Phys. J. C 80 (2020) 113, arxiv:902.08191

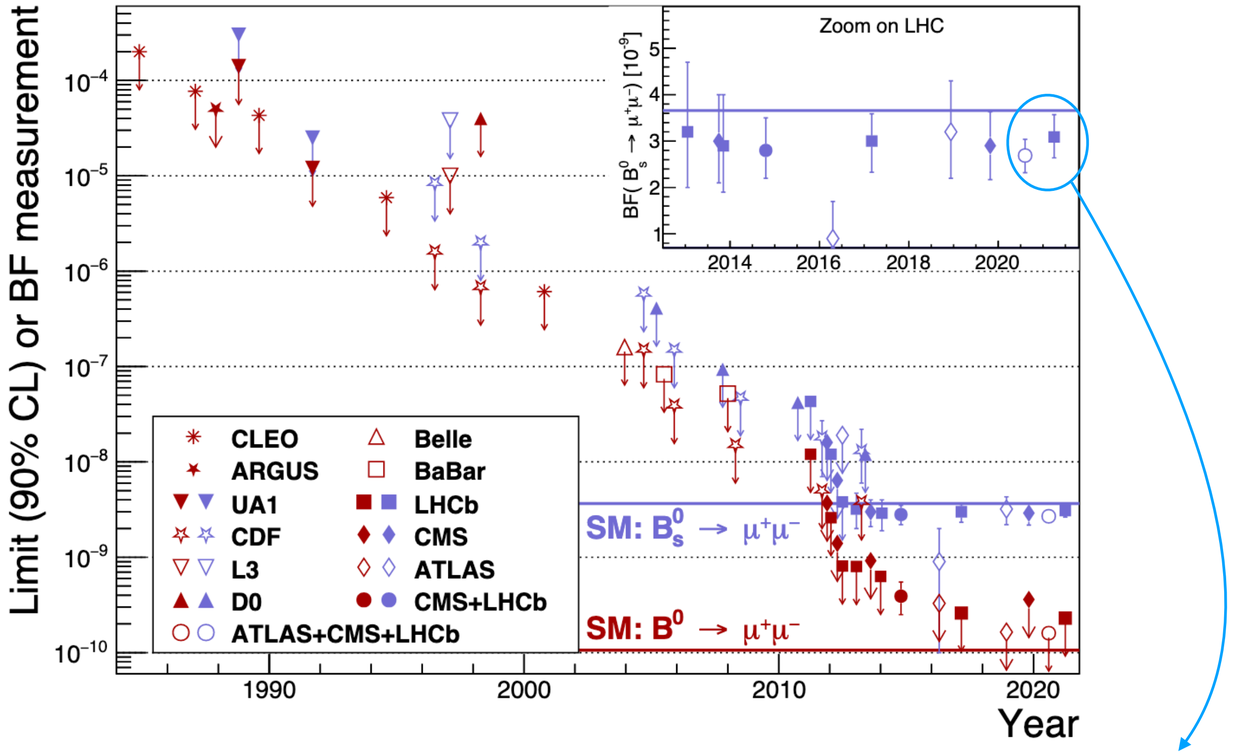
- Significant deviations could arise in new physics scenarios (Two higgs doublet Model, Minimal Flavor Violation, SUSY..)
- The experimentally clean signature provide a unique opportunity to probe for the new physics effects in the branching fractions



- And a further probe is provided by the measurement of the effective lifetime $\tau_{B^0_s \to \mu^+\mu^-}$, which receives contribution just from the heavy mass eigenstate B^0_{sH} in the SM prediction

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

- A three decades long quest
- Results presented here:
  - ATLAS: JHEP 04 (2019) 098
    - $26.3\ \text{fb}^{-1}$ at $\sqrt{s} = 13$ TeV (2015 and 2016)
    - $25\ \text{fb}^{-1}$ at $\sqrt{s} = 7$ and 8 TeV (2011 and 2012)
  - CMS: JHEP 04 (2020) 188
    - $36\ \text{fb}^{-1}$ at $\sqrt{s} = 13$ TeV (2016)
    - $5(20)\ \text{fb}^{-1}$ at $\sqrt{s} = 7(8)$ TeV (in 2011(2012))



LHCb latest result (2021)
LHCB-PAPER-2021-007, Archilli's Moriond EW and Santimaria's LHC seminar
(not presented in this talk)

Recently combined ATLAS+CMS+LHCb[*] results!!

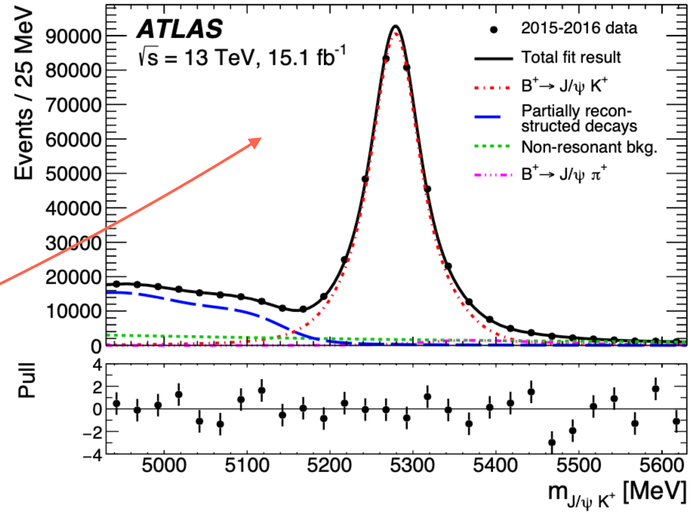
[*]LHCb: PRL 118, 191801 (2017)

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

- The branching fractions of the of the $B^0_{(s)} \to \mu^+\mu^-$ channels are measured with similar strategies:
  - $B^+ \to J/\psi K^+$ is used as reference channel since abundant and with a well measured branching fraction

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

Maximum likelihood fit in $J/\psi K^+$ data



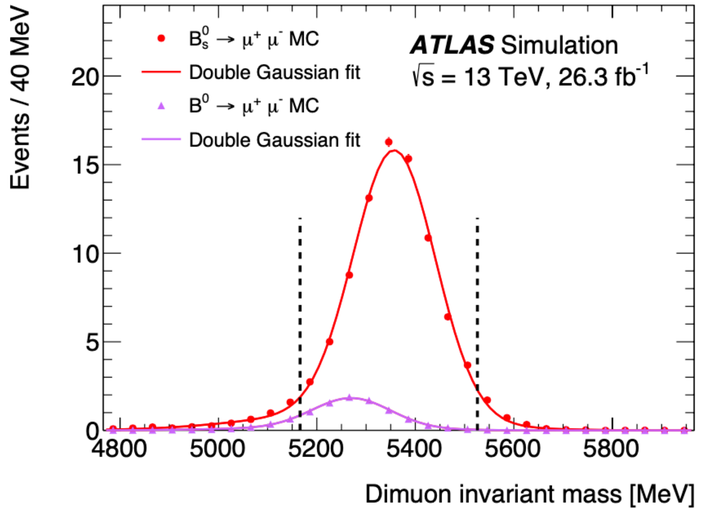
N are the yields

ϵ the acceptance times efficiency

ratio of the hadronization probabilities of b-quarks to B^0 or B^0_s

- Signal selection:
  - starting from dimuon triggered data, pairs of well identified opposite charged muons are combined to form a displaced vertex;
  - additional criteria on isolation and kinematics

- Invariant mass of the 2 muons is used to identify a signal enriched region and two sideband mass regions enriched in backgrounds (used for background prediction)

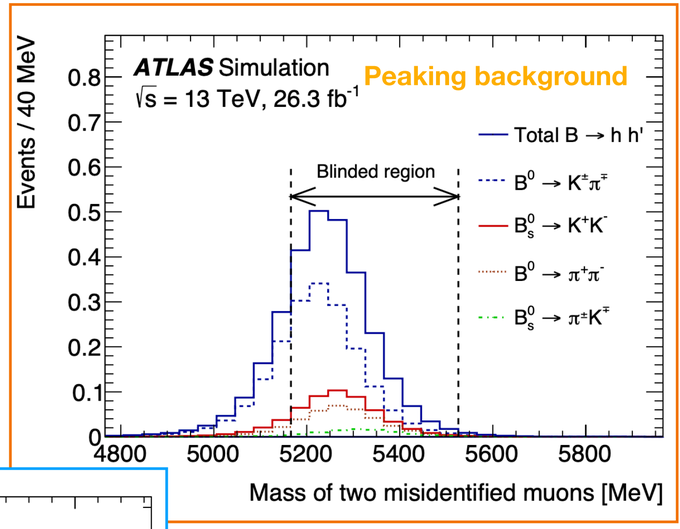


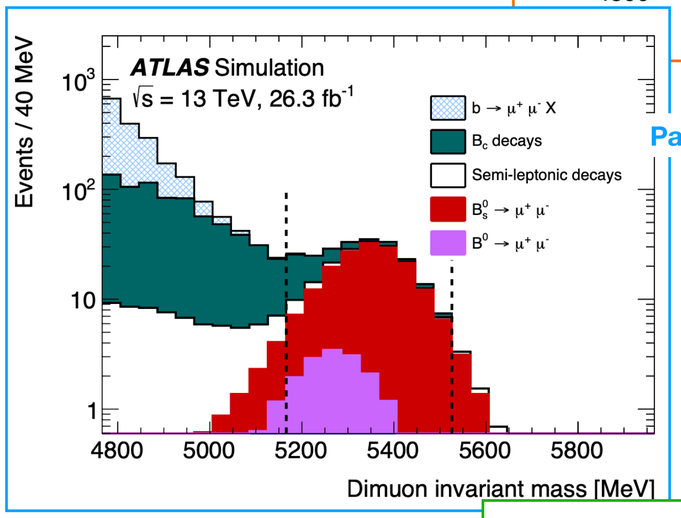
ATLAS: JHEP04(2019)098

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

ATLAS: JHEP04(2019)098

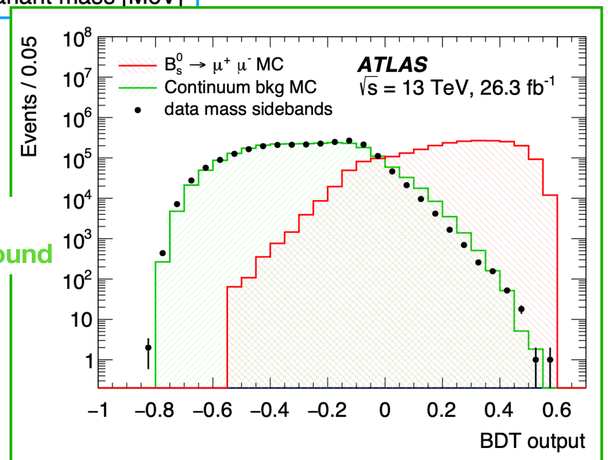
- Peaking backgrounds: $B^0_{(S)} -> hh'$ decays with both hadrons misidentified as muons
- Partially reconstructed decays: one or more of the final-state particles (X) in a b-hadron decay are not reconstructed
  - contribution in low dimuon mass region
- Continuum background - dominant
  - Combinatorial component: muons originated from uncorrelated hadron decays
    - modeled with $bb \to \mu^+\mu^- X$ MC sample, checked in the signal mass sidebands in data (ATLAS)
    - from the mass sideband in data (CMS)
  - BDT discriminator to distinguish background from signal
    - variables related to: B meson decay, muon quality and the rest of the event (isolation, number of tracks)
      - 15 ATLAS, 12 CMS





Partially reconstructed

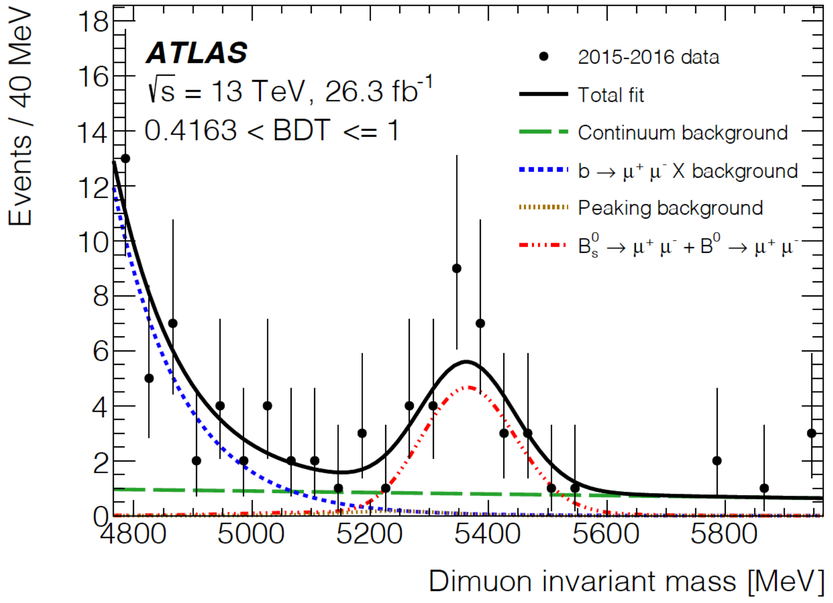
BDT output for Continuum background

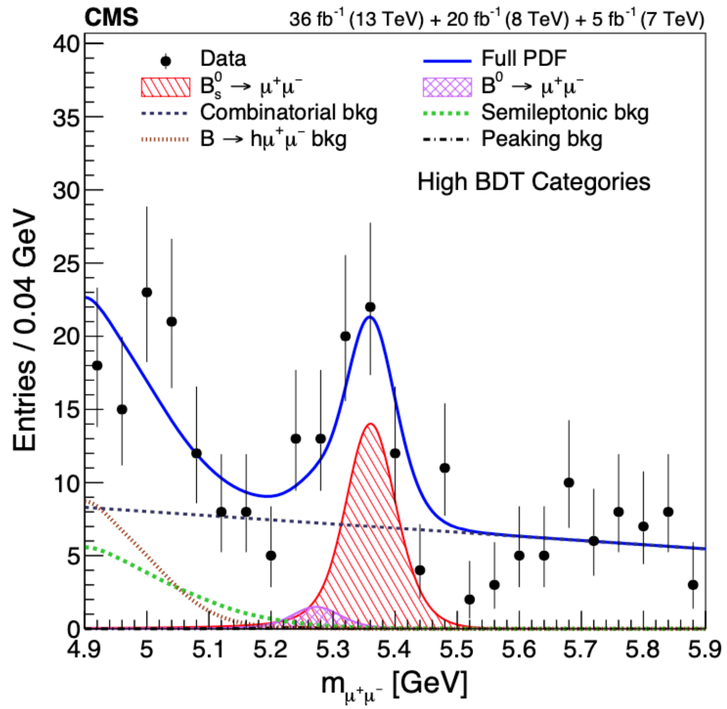


$B^0_{(s)} \to \mu^+\mu^-$: signal extraction

- The dimuon candidates are classified according to the BDT output
- ATLAS yield, determined from the unbinned maximum likelihood fit of highest (3) BDT bins:
  - N_s= 80 ±22 and N_d= -12 ±20, in agreement with SM expectations
- CMS yield is determined from each BDT bin and data subset category (14- split by year and detector region)
  - in agreement with SM expectations
- Uncertainty is statistically dominated:

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





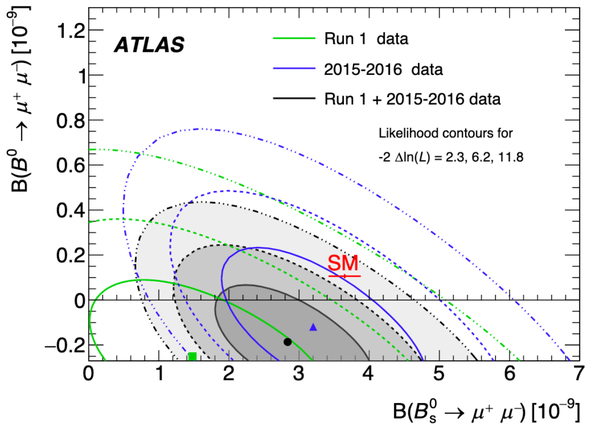
$B^0_{(s)} \to \mu^+\mu^-$: branching fraction

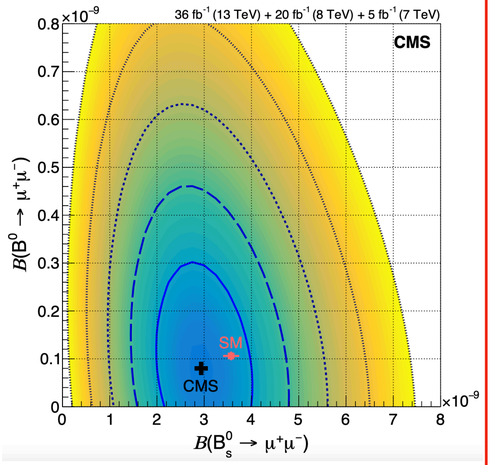
CMS-BPH-20-003, LHCb-CONF-2020-002, ATLAS-CONF-2020-049

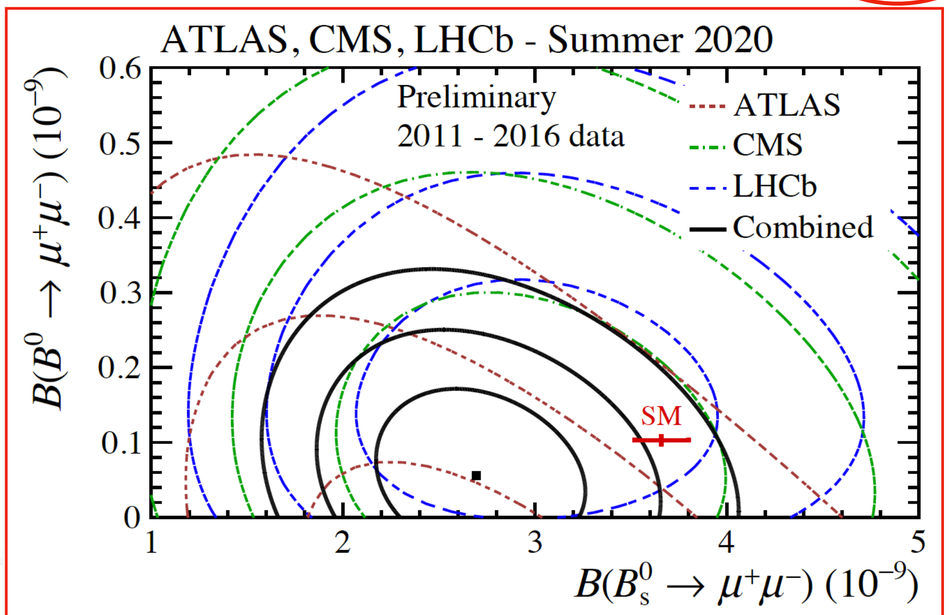
Achilli's presentation Moriond EW

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

	ATLAS	CMS	LHCb (2017)	Combination	LHCb (2021)
$\mathcal{B}(B^0_s \to \mu^+\mu^-)$	$2.8^{+0.8}_{-0.7} \cdot 10^{-9}$	$[2.9^{+0.7}_{-0.6}(\text{exp}) \pm 0.2(\text{frag})] \cdot 10^{-9}$	$[3.0 \pm 0.6^{+0.3}_{-0.2}] \cdot 10^{-9}$	$[2.69^{+0.37}_{-0.35}] \cdot 10^{-9}$	$[3.09^{+0.46+0.15}_{-0.43-0.11}] \cdot 10^{-9}$
$\mathcal{B}(B^0 \to \mu^+\mu^-)$	$(-1.9 \pm 1.6) \cdot 10^{-10}$	$(0.8^{+1,4}_{-1.3}) \cdot 10^{-10}$	$(1.5^{+1.2+0.2}_{-1.0-0.1}) \cdot 10^{-10}$	$(0.6 \pm 0.7) \cdot 10^{-10}$	
$\mathcal{B}(B^0 \to \mu^+\mu^-)$ Upper Limit at 95% CL	$< 2.1 \cdot 10^{-10}$ with Neyman procedure	$< 3.6 \cdot 10^{-10}$	$< 3.4 \cdot 10^{-10}$	$< 1.9 \cdot 10^{-10}$	$< 2.6 \cdot 10^{-10}$
$\tau_{B^0_s \to \mu^+\mu^-}$ [more in backup]		$(1.70^{+0.60}_{-0.43} \pm 0.09)$ ps	$(2.04 \pm 0.44 \pm 0.05)$ ps	$(1.91^{+0.37}_{-0.35})$ ps	$(2.07 \pm 0.29 \pm 0.03)$ ps





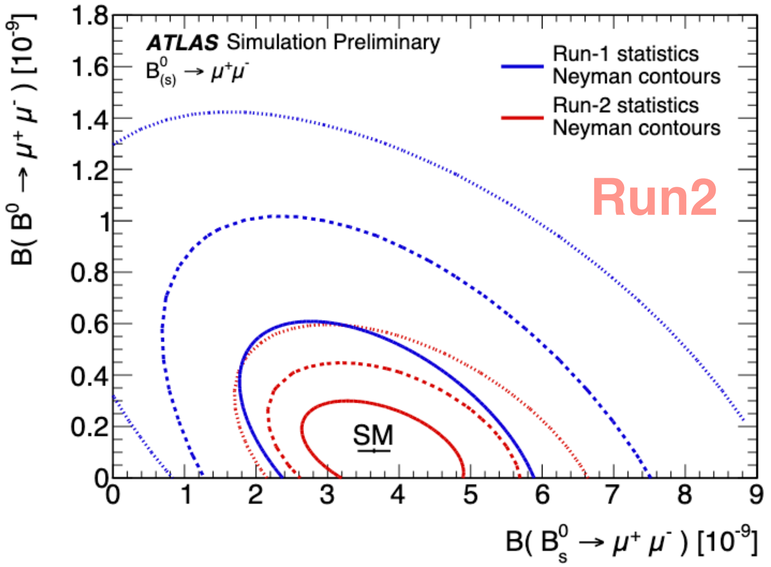


2.4σ and 0.64σ discrepancy with the SM expectations for $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ and $\mathcal{B}(B^0 \to \mu^+\mu^-)$

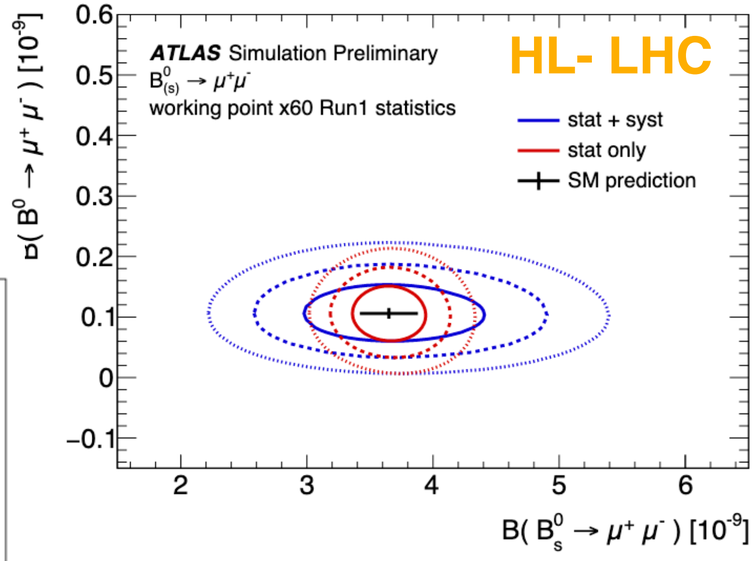
$\tau_{B^0_s \to \mu^+\mu^-}$ found compatible with B^0_{sH}-only contribution (SM)

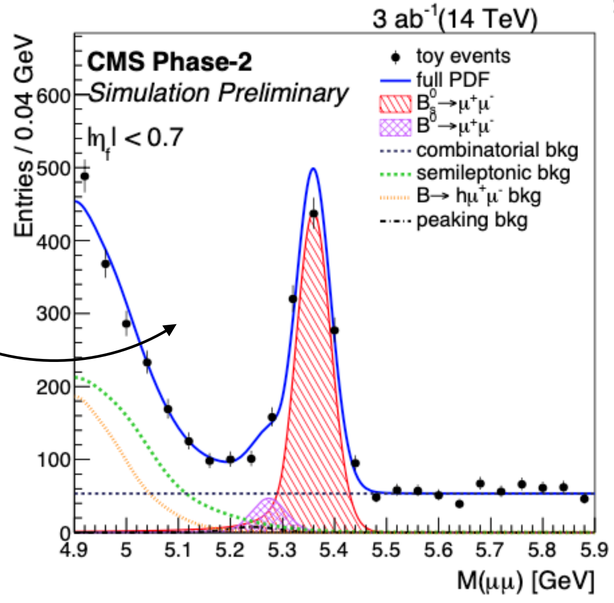
$B^0_{(s)} \to \mu^+\mu^-$: prospects

- In Phase-2 (**HL-LHC**) expected precision:
  - CMS and ATLAS upgrade of the inner tracker is expected to provide a 40-50% improvement on the mass resolution (reduce B_s-B_d mass correlation)
    - ALTAS:
      - $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ (~15%) - factor 2 improvement wrt full Run2
      - $\mathcal{B}(B^0 \to \mu^+\mu^-)$ (~30%)
    - CMS:
      - $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ (~7%)
      - $\mathcal{B}(B^0 \to \mu^+\mu^-)$ (~16%) - 5σ observation
      - 3% precision on the $\tau_{B^0_s \to \mu^+\mu^-}$



ATL-PHYS-PUB-2018-005



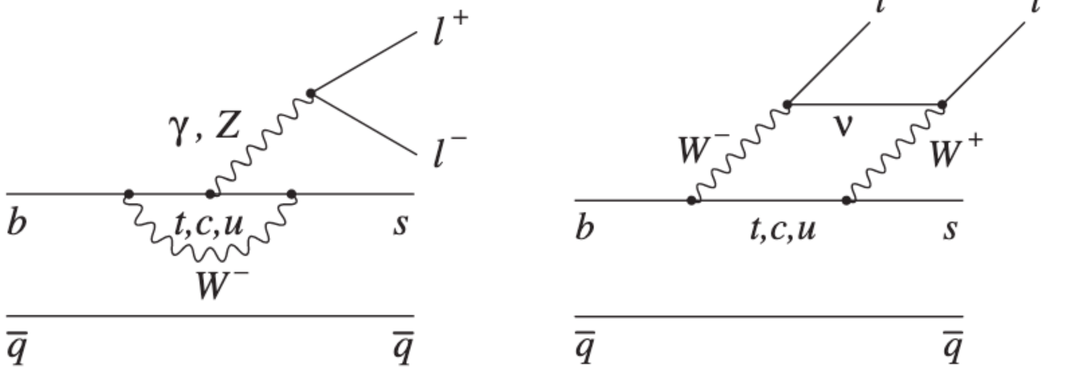


CMS-FTR-18-013

$\mathrm{B} \to \mathrm{K}^{*}\mu\mu$ angular analyses

$B \to K^* \mu^+ \mu^-$ motivation

- Flavor changing neutral current decays are also interesting for angular analysis:
  - The properties of the decay can be thoroughly investigated
  - The angular distributions of these decays can be parametrized in observables where the dependence on (and uncertainty due to) hadronic effects and form factors is minimized
- Some tensions (~2σ) with the SM predictions have been reported (LHCb and Belle)



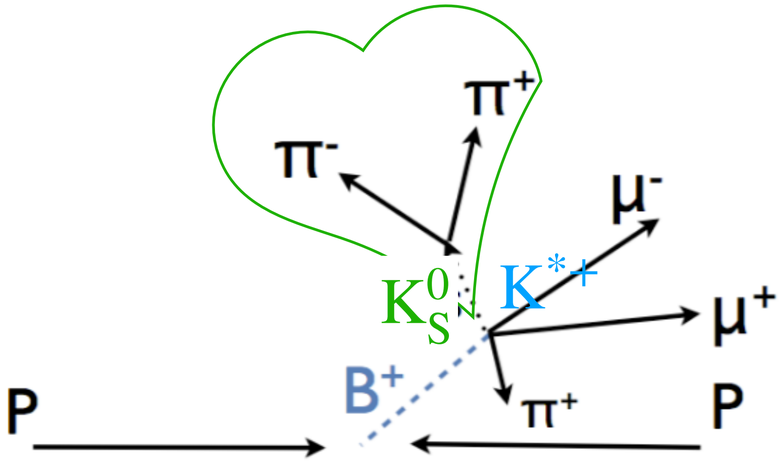
- Results
  - CMS:
    B^+:JHEP04(2021)124 *Recently published!*
    B^0:Phys. Lett. B 781 (2018) 517
    - $20\ \text{fb}^{-1}$ at $\sqrt{s} = 8$ TeV (2012)
  - ATLAS:
    B^0:JHEP10(2018)047
    - $20.3\ \text{fb}^{-1}$ at $\sqrt{s} = 8$ TeV (2012)

Phys. Rev. Lett. 111, 191801 (2013)

Phys. Rev. Lett. 103, 171801 (2009)

JHEP04(2021)124
CMS-PAS-BPH-15-009

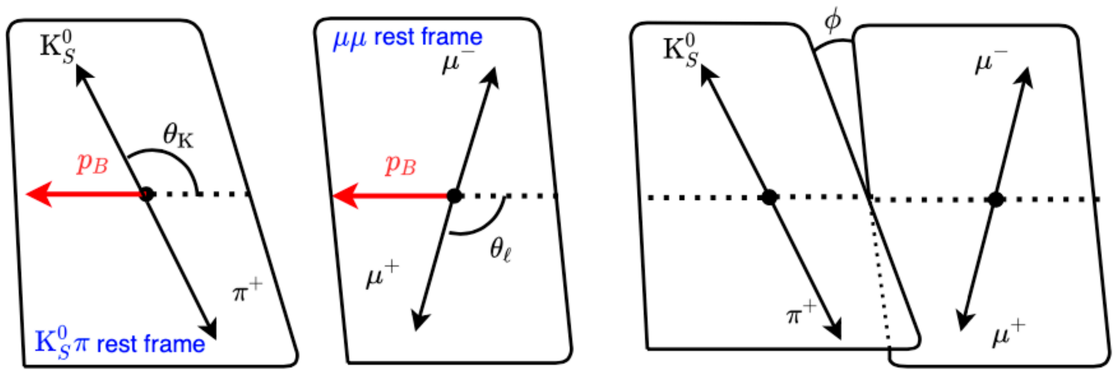
$B^+ \to K^{*+}(892)\mu^+\mu^-$



- pp collisions from 2012 corresponding to $20\ \text{fb}^{-1}$ at $\sqrt{s} = 8$ TeV:
  - Dimuon trigger with requirements on displaced vertex and transverse momentum collinear with decay flight
  - K^{*+} is reconstructed through $K^0_S(K^0_S \to \pi^+\pi^-)\pi^+$:
    - $\pi^+\pi^-$ with a common vertex and $|m_{\pi^+\pi^-} - m_{K^0_S}| < 17.3\text{MeV}\ (3\sigma_{m_{K^0_S}})$
  - $|m_{K^0_S\pi^+} - m_{K^{*+}}| < 100$ MeV , and $m_{K^0_S\pi^+\mu^+\mu^+}\epsilon$ [4.76,5.8] GeV
  - Selection on the dimuon invariant mass (q^2) to suppress $B^+ \to K^{*+}J/\psi$ and $B^+ \to K^{*+}\psi(2S)$

- Aim is to measure two decay observables: the muon forward-backward asymmetry, A_{FB}, and the K^{*+} longitudinal polarization fraction, F_L, in bins of squared dimuon invariant mass (q^2).

- Decay described by four independent kinematic variables:
  - q^2 and angles $(\theta_K, \theta_\ell, \phi)$



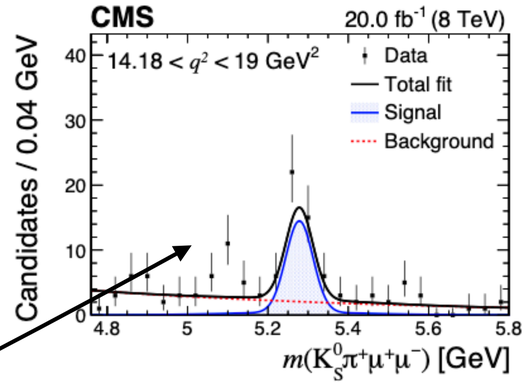
$B^+ \to K^{*+}(892)\mu^+\mu^-$: angular analysis

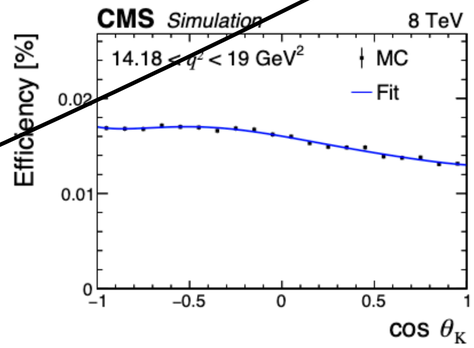
JHEP04(2021)124
CMS-PAS-BPH-15-009

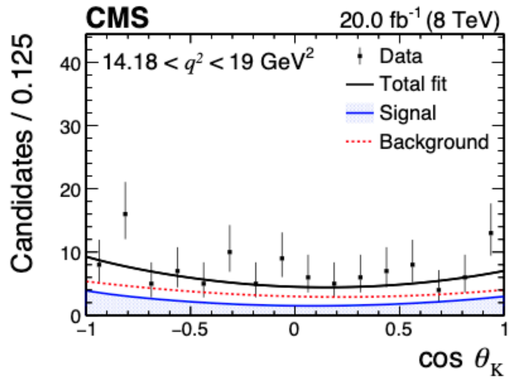
- The differential decay rate is:

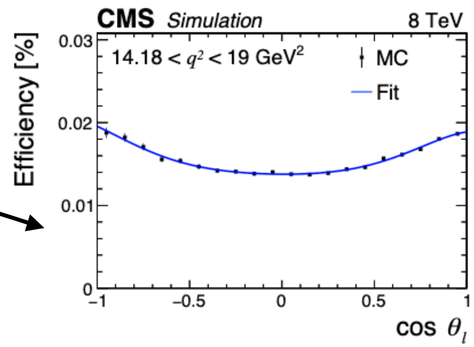
$$\frac{1}{\Gamma}\frac{d^3\Gamma}{d\cos\theta_K\, d\cos\theta_\ell\, dq^2} = \frac{9}{16}\left\{\frac{2}{3}\left[F_S + 2A_S\cos\theta_K\right](1-\cos^2\theta_\ell)\right.$$
$$+ (1-F_S)\left[2F_L\cos^2\theta_K(1-\cos^2\theta_\ell)\right.$$
$$+ \frac{1}{2}(1-F_L)\left(1-\cos^2\theta_K\right)(1+\cos^2\theta_\ell)$$
$$\left.\left. + \frac{4}{3}A_{FB}\left(1-\cos^2\theta_K\right)\cos\theta_\ell\right]\right\}.$$

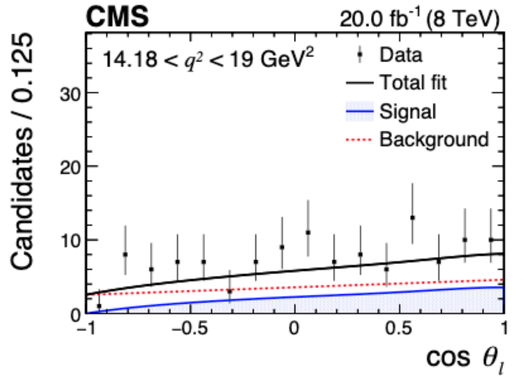
- In each bin of q^2 , A_{FB} and F_L are extracted with unbinned maximum likelihood fits, parametrized in terms of $m, \cos\theta_K, \cos\theta_\ell$:
  - signal distributions based on the prediction
  - data-driven background estimation from mass sidebands
- Efficiencies are obtained from simulations
  - Data / simulation agreement is checked in $B^+ \to K^{*+}J/\psi$ enriched control region







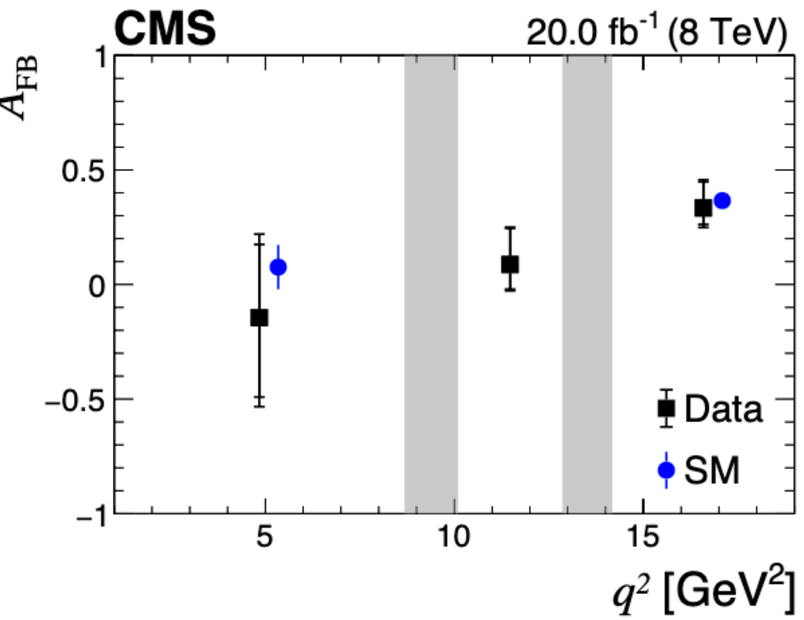


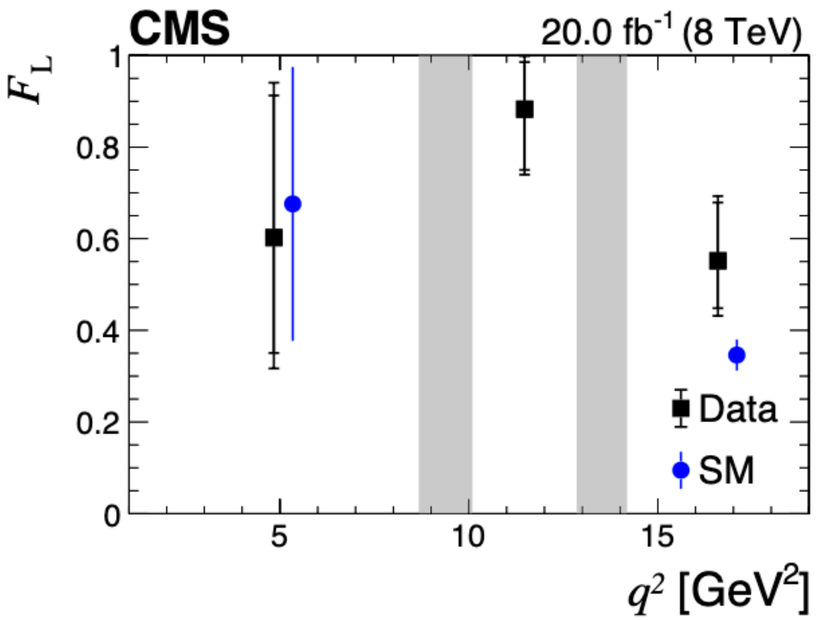


$B^+ \to K^{*+}(892)\mu^+\mu^-$: results

JHEP04(2021)124
CMS-PAS-BPH-15-009

- Results in agreement with the SM:





Source	A_{FB} (10^{-3})	F_L (10^{-3})
MC statistical uncertainty	12 – 29	18 – 38
Efficiency model	3 – 25	4 – 12
Background shape functional form	0 – 9	0 – 33
Background shape statistical uncertainty	16 – 73	20 – 87
Background shape sideband region	28 – 153	38 – 78
S-wave contamination	4 – 22	5 – 12
Total systematic uncertainty	42 – 174	55 – 127

q^2 (GeV^2)	Y_S	A_{FB}	F_L
1 – 8.68	22.1 ± 8.1	$-0.14^{+0.32}_{-0.35} \pm 0.17$	$0.60^{+0.31}_{-0.25} \pm 0.13$
10.09 – 12.86	25.9 ± 6.3	$0.09^{+0.16}_{-0.11} \pm 0.04$	$0.88^{+0.10}_{-0.13} \pm 0.05$
14.18 – 19	45.1 ± 8.0	$0.33^{+0.11}_{-0.07} \pm 0.05$	$0.55^{+0.13}_{-0.10} \pm 0.06$

uncertainty is statistically dominated

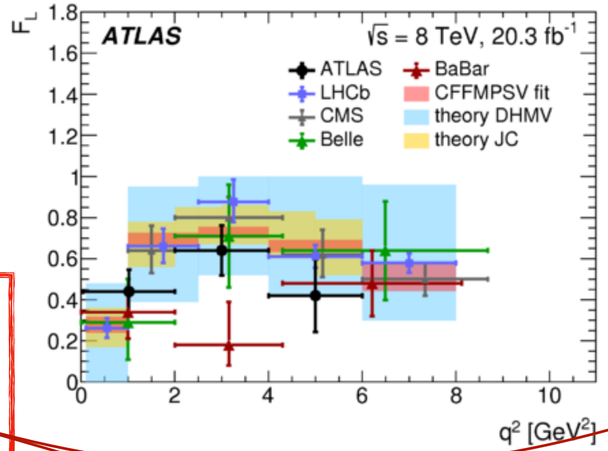
Other angular results using 8 TeV data in the following slides..

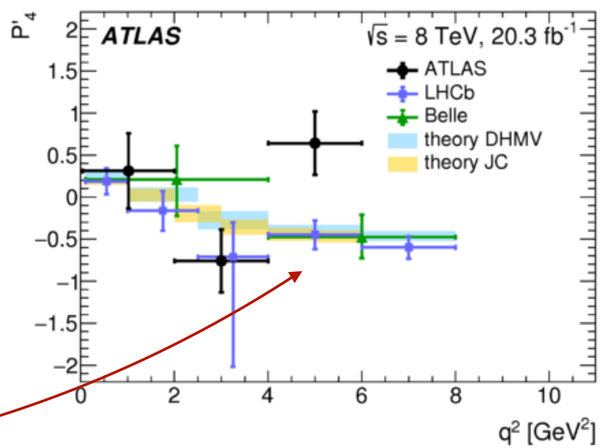
$B^0 \to K^{*0}(892)\mu^+\mu^-$

$$\frac{1}{d\Gamma/dq^2}\frac{d^4\Gamma}{d\cos\theta_L d\cos\theta_K d\phi dq^2} = \frac{9}{32\pi}\left[\frac{3(1-F_L)}{4}\sin^2\theta_K + F_L\cos^2\theta_K + \frac{1-F_L}{4}\sin^2\theta_K\cos 2\theta_L\right.$$
$$-F_L\cos^2\theta_K\cos 2\theta_L + S_3\sin^2\theta_K\sin^2\theta_L\cos 2\phi$$
$$+S_4\sin 2\theta_K\sin 2\theta_L\cos\phi + S_5\sin 2\theta_K\sin\theta_L\cos\phi$$
$$+S_6\sin^2\theta_K\cos\theta_L + S_7\sin 2\theta_K\sin\theta_L\sin\phi$$
$$\left.+S_8\sin 2\theta_K\sin 2\theta_L\sin\phi + S_9\sin^2\theta_K\sin^2\theta_L\sin 2\phi\right]$$

JHEP 10 (2018) 047
ATLAS-BPHY-2013-02

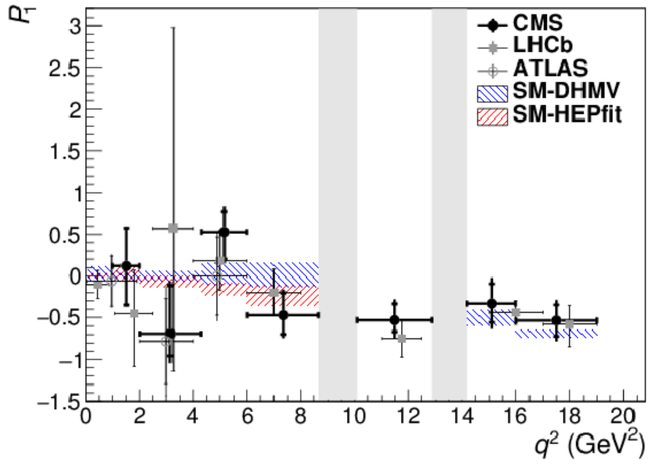
- ATLAS analysis at 8 TeV
  - Angular fit for several observables
    - Folding in ϕ and θ_ℓ for simplification
    - Extended ML fit in 3 q^2 bins

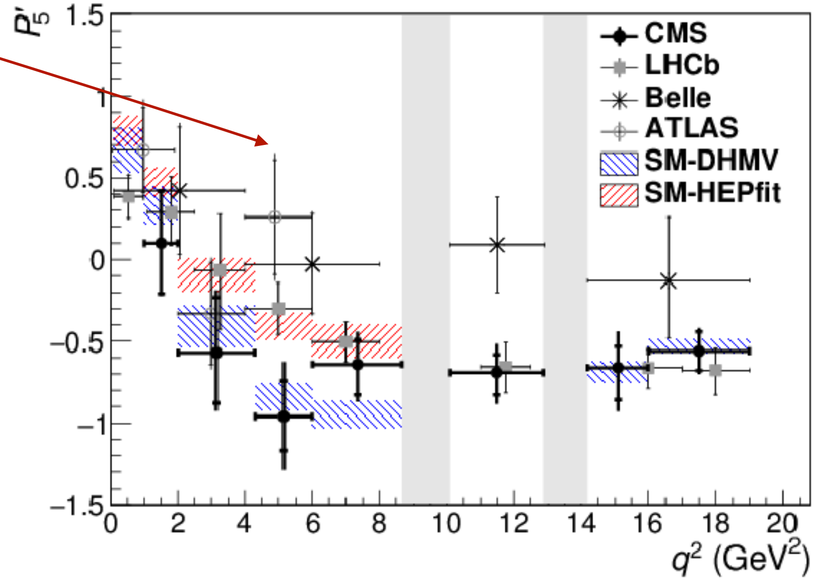




- All measurements are within 3σ range covered by predictions
- 2.7σ deviations for P'_4, P'_5 in $q^2 \in [4,6]$ GeV²

- CMS performed similar analysis at 8 TeV and fitted for P_1 and P'_5
  - precision comparable with other results
  - consistent with predictions based on the standard model.

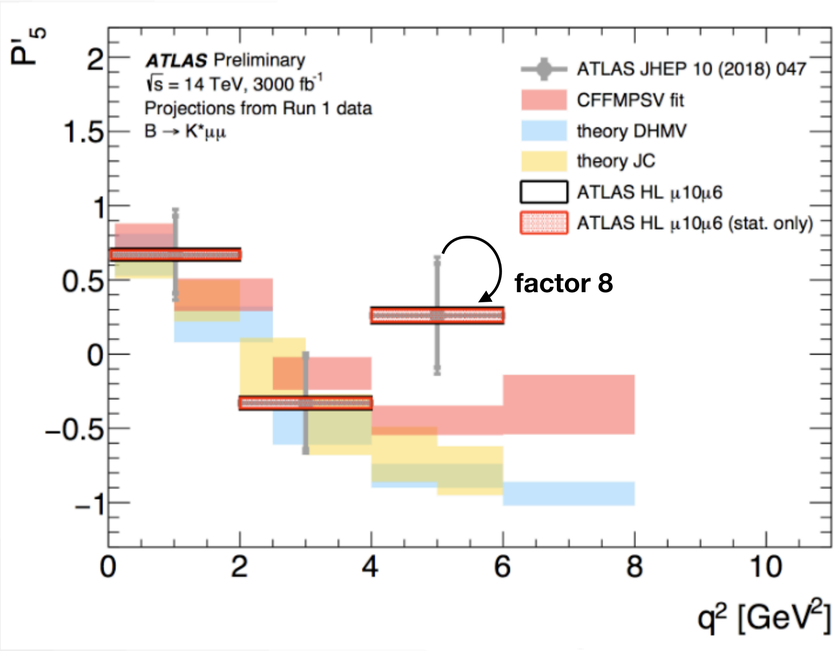




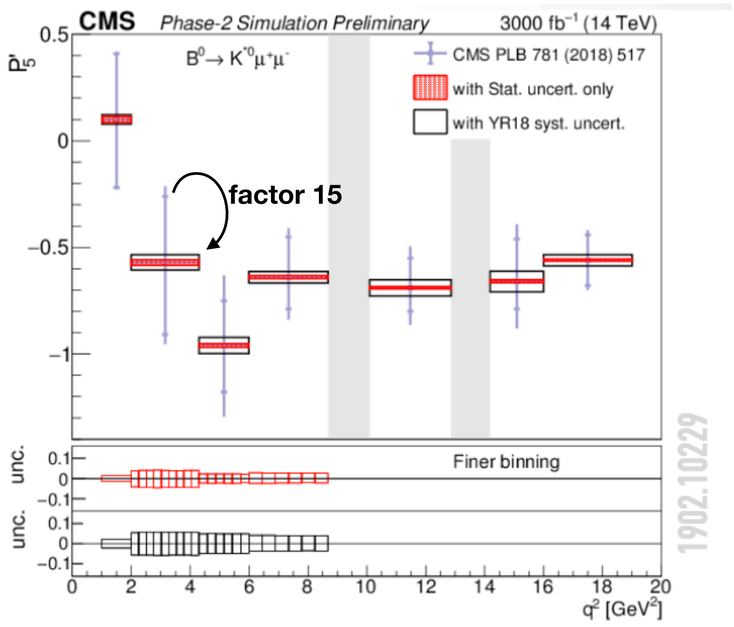
Phys. Lett. B 781 (2018) 517
CMS-BPH-15-008

$B^0 \to K^{*0}(892)\mu^+\mu^-$: prospects

CERN-LPCC-2019-01



- Estimation from Run 1 results keeping into account improved tracking performance and different trigger scenarios:
  - the precision in measuring a representative P'_5 parameter is expected to improve by factors of 8



- Estimation from Run 1 results keeping into account improved tracking performance ⇒ higher mass resolution
  - the precision in measuring a representative P'_5 parameter is expected to improve by a factor 15
  - capability to perform a full angular analysis

$$\tau \to 3\mu$$

$\tau \to 3\mu$

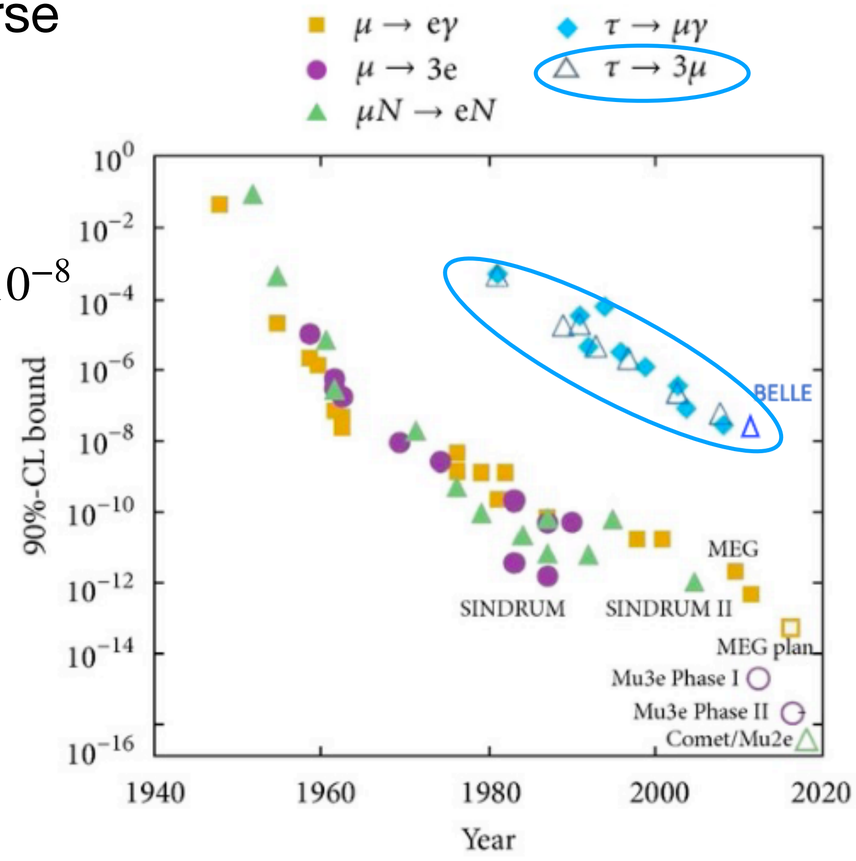
- Observing a charge-lepton flavor violating decay $\tau \to 3\mu$ would be a major breakthrough in understanding the matter content of the universe

- Allowed by neutrino oscillation, with very small branching fraction:
  - In some extension of the SM, including new physics scenarios, $\mathcal{B}(\tau \to 3\mu)$ could be enhanced to order of magnitudes of $10^{-10} - 10^{-8}$
  - Experimentally accessible and clean three-muon final state

  - Most stringent limit (Belle I): $\mathcal{B}(\tau \to 3\mu) < 2.1\ 10^{-8}$ (90% CL)

    - Results
      - ATLAS: Eur.Phys.J.C76(2016) 232
        - $20.3\ \text{fb}^{-1}$ at $\sqrt{s} = 8$ TeV (2012)
        - Results from $W \to \tau$ production
      - CMS: JHEP01(2021)163
        - $33\ \text{fb}^{-1}$ at $\sqrt{s} = 13$ TeV (2016)
        - Combination of results from $W \to \tau$ and $B, D \to \tau$ production

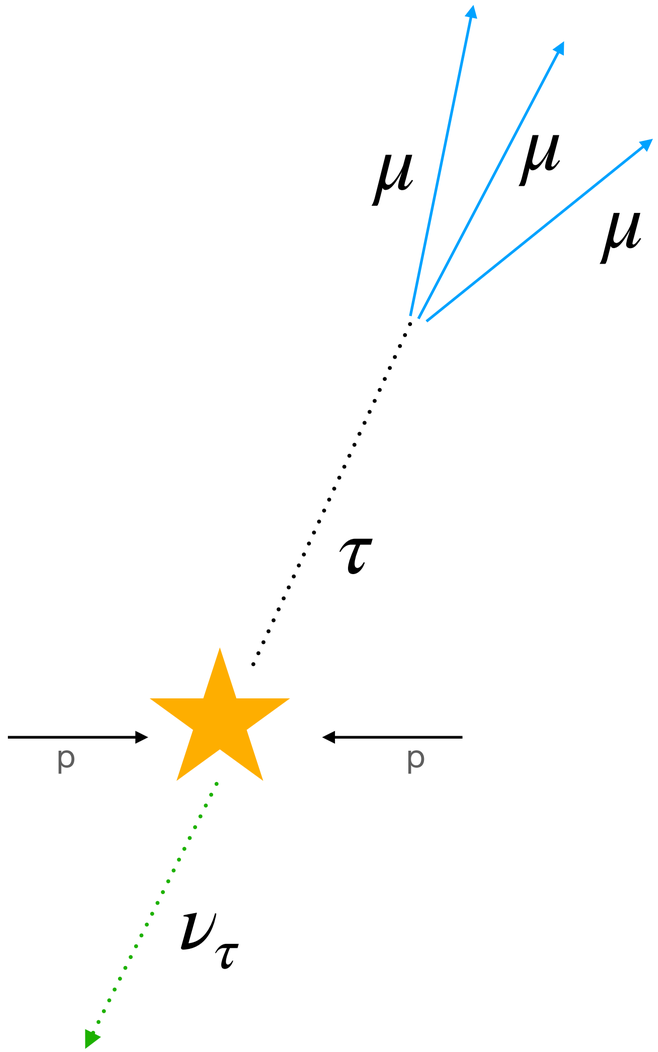
Historical view of the LFV experiments



recently published!

The European Physical Journal C, 76, 232 (2016)

$\tau \to 3\mu$: ATLAS



- Tau produced in $W \to \tau\nu_\tau$ events:
  - the neutrino manifests as missing transverse energy (E_T^{miss})
  - the 3 muons from the tau decays are expected to have geometrical proximity and form a displaced vertex(tau lifetime 10^{-12} s)
- The branching fraction is estimated as:

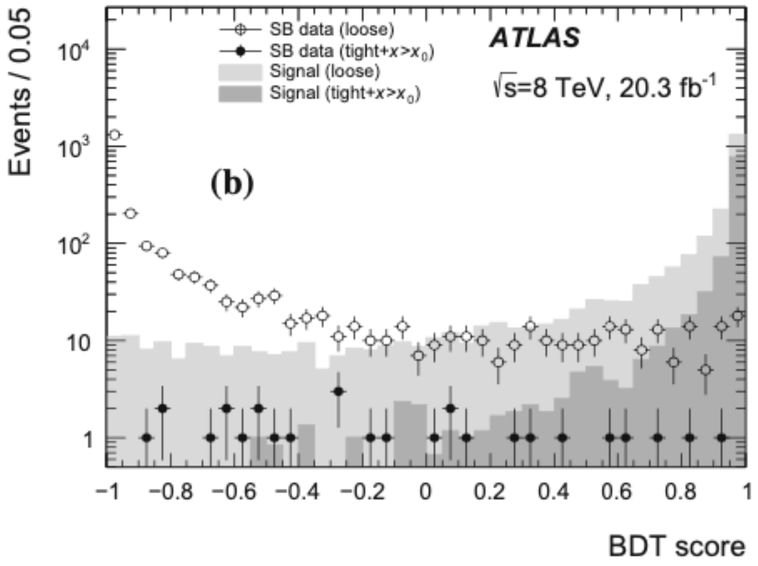
$$\mathrm{Br}(\tau \to 3\mu) = \frac{N_S}{(\mathcal{A}_S \times \epsilon_S)\, N_{W\to\tau\nu}}$$

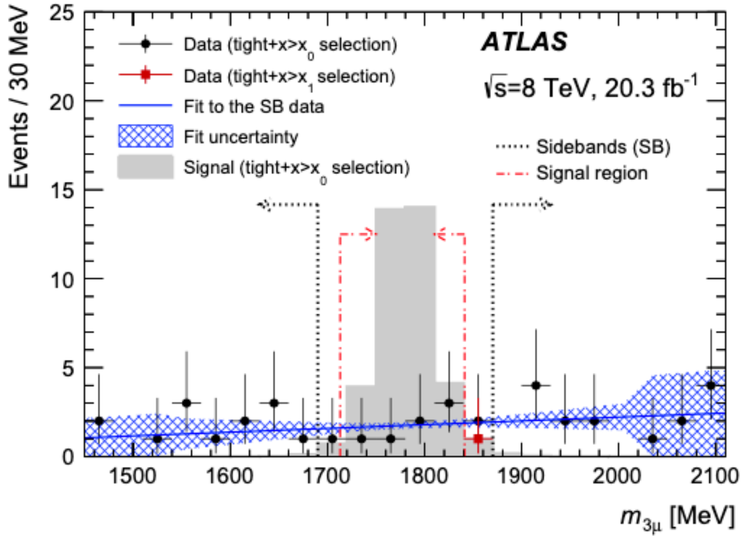
where N_S is the number of signal events, $\mathcal{A}_S \times \epsilon_S$ the acceptance times the efficiency for the signal and N is the number of tau leptons produced in the collision, extrapolated from 2011 $\sigma_{W\to\ell\nu}$ measurement

The European Physical Journal C, 76, 232 (2016)

$\tau \to 3\mu$: ATLAS

- A loose preselection on 3 muon events reject kinematically inconsistent with the signal: muon kinematics, track isolation, vertex displacement, and missing energy
- A BDT discriminator is trained with signal and background events from data with $m_{3\mu}$ sidebands ([750,1450] and [2110,2500] MeV)
- A tighter selection is applied:
  - displacement and invariant mass of two muons
    (to remove contributions from meson/light resonances)
    + with a loose BDT score cut (x_0=-0.9)
    (to further reduce the backgrounds)
  - the shape and the expected yields of backgrounds in the tau mass signal region are extrapolated from a fit in sideband regions
    (final yields are then re-scaled for the optimal BDT cut (x_1=0.933))



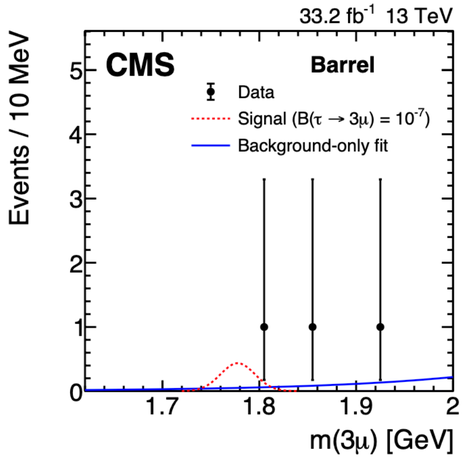


$\mathcal{B}(\tau \to 3\mu) < 38\ 10^{-8}$ at 90% CL

JHEP01(2021)163

$\tau \to 3\mu$: CMS (W, B-D production)

- Trigger: three muons or two muons and a track with mass and vertex requirements
- Candidate: three vertexing muon with $m_{3\mu}$ in [1.6,2.00] GeV + requirements on the invariant mass of any of two muons to remove backgrounds from meson/resonances



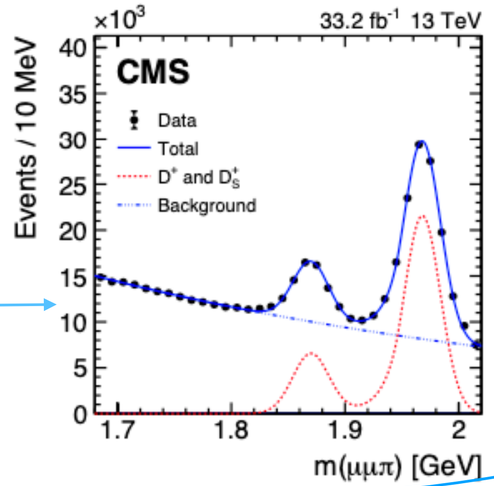
$W \to \tau$ production

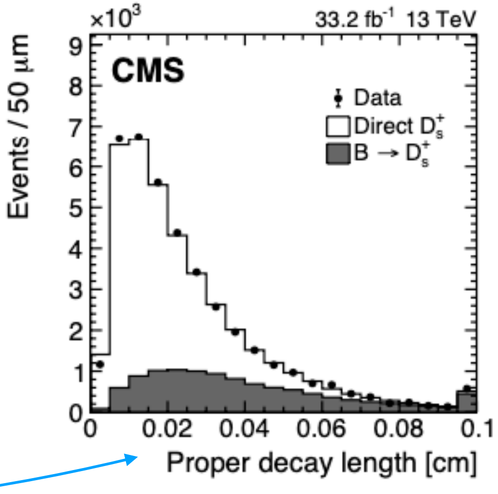
- BDT discriminator is trained with signal and background events from data mass sideband ([1.60,1.74] and [1.82,2.00] GeV)
  - 18 variables:muon quality, displacement, isolation
  - 2 categories: endcap and barrel

$$\mathcal{B}(\tau \to 3\mu) = \frac{N_{\mathrm{sig(W)}}}{\mathcal{L}\,\sigma(\mathrm{pp} \to \mathrm{W} + \mathrm{X})\,\mathcal{B}(\mathrm{W} \to \tau\nu)\,\mathcal{A}_{3\mu(\mathrm{W})}\,\epsilon_{3\mu(\mathrm{W})}}$$

B, D $\to \tau$ production

- $D \to \tau$ (65%), $B \to \tau$(25%), $B \to D \to \tau$ (10%) [~95% D_S]
- $D_S \to \phi\pi \to \mu\mu\pi$ as reference channel to reduce uncertainties of the heavy-flavor hadrons production:
  - normalization from a fit to mass spectrum in data
  - fraction of the D_S candidates from B meson decay is established by a fit to the proper decay length





$\tau \to 3\mu$: CMS

JHEP01(2021)163

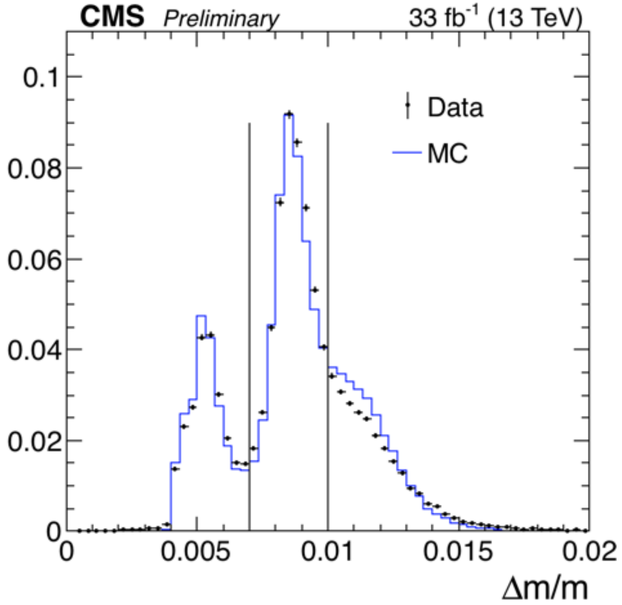
$$N_{\mathrm{sig(D)}} = N_{\mathrm{norm}} \frac{\mathcal{B}(\mathrm{D_s^+} \to \tau^+\nu)}{\mathcal{B}(\mathrm{D_s^+} \to \phi\pi^+ \to \mu^+\mu^-\pi^+)} \frac{\mathcal{A}_{3\mu(\mathrm{D})}}{\mathcal{A}_{\mu\mu\pi}} \frac{\epsilon^{\mathrm{reco}}_{3\mu(\mathrm{D})}}{\epsilon^{\mathrm{reco}}_{\mu\mu\pi}} \frac{\epsilon^{2\mu\mathrm{trig}}_{3\mu(\mathrm{D})}}{\epsilon^{2\mu\mathrm{trig}}_{\mu\mu\pi}} \mathcal{B}(\tau \to 3\mu)$$

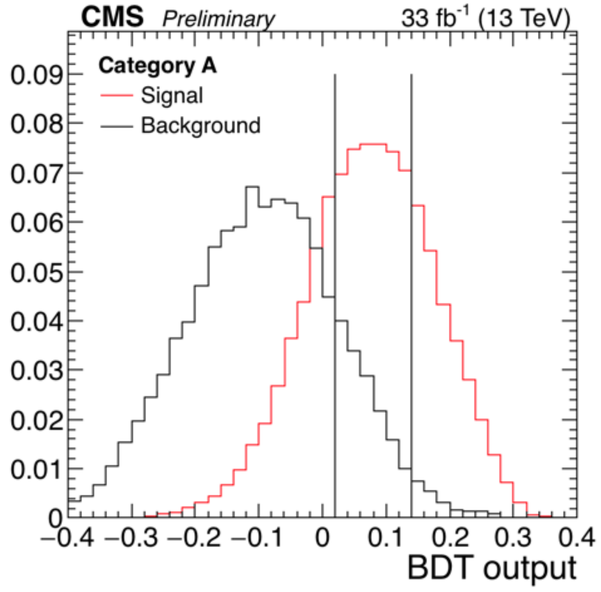
Acceptance efficiency corrections

from decay length fit

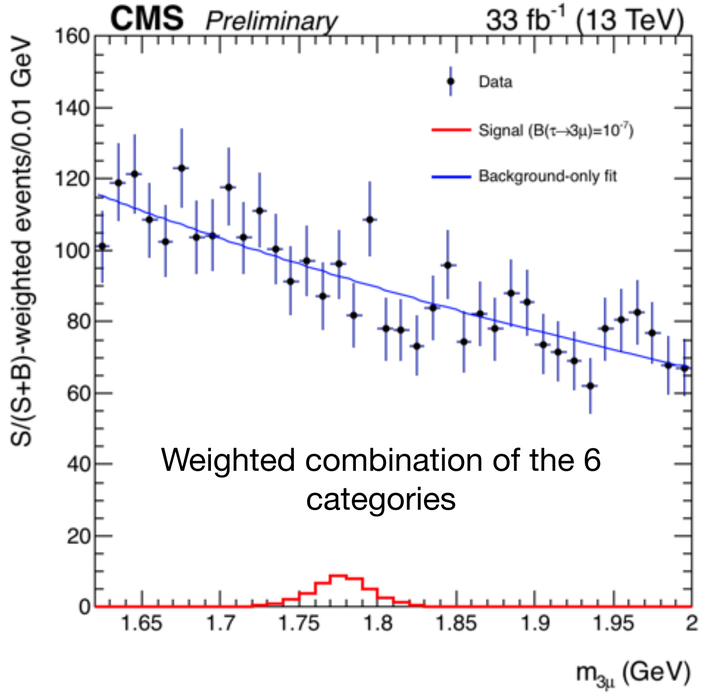
$$N_{\mathrm{sig(B)}} = N_{\mathrm{norm}} f \frac{\mathcal{B}(\mathrm{B} \to \tau + \mathrm{X})}{\mathcal{B}(\mathrm{D_s^+} \to \phi\pi^+ \to \mu^+\mu^-\pi^+)\mathcal{B}(\mathrm{B} \to \mathrm{D_s^+} + \mathrm{X})} \frac{\mathcal{A}_{3\mu(\mathrm{B})}}{\mathcal{A}_{\mu\mu\pi}} \frac{\epsilon^{\mathrm{reco}}_{3\mu(\mathrm{B})}}{\epsilon^{\mathrm{reco}}_{\mu\mu\pi}} \frac{\epsilon^{2\mu\mathrm{trig}}_{3\mu(\mathrm{B})}}{\epsilon^{2\mu\mathrm{trig}}_{\mu\mu\pi}} \mathcal{B}(\tau \to 3\mu)$$

- Maximum likelihood fit performed simultaneously in 6 categories:
  - Three signal subsamples based on mass resolution according to muon rapidity (0.4-1.5%)
  - BDT discriminator trained using vertex- and muon- quality variables
    - Signal (from simulation) discriminated from background (from data sidebands)
    - Just 2 categories (richest in signal) used in the final fit





Dominant systematic uncertainty: D_S normalisation (10%)



No excess observed in the signal region

Combination of results $W \to \tau$ & $B, D \to \tau$

$\mathcal{B}(\tau \to 3\mu) < 8.8\ 10^{-8}$ at 90% CL

arxiv:1812.07638

$\tau \to 3\mu$: prospects

$\mathcal{B}(\tau \to 3\mu)$ 90% CL limit	Exp	Comments
3.8 x 10^{-7}	ATLAS	Limit (Run 1)
8.8 x 10^{-8}	CMS	Limit (2016)
2.1 x 10^{-8}	Belle I	Limit
2.3 x 10^{-9}	ATLAS - HF channel	expected - 3000/fb
6 x 10^{-9}	ATLAS - W channel	expected - 3000/fb
3.7 x 10^{-9}	CMS - HF channel	expected - 3000/fb
3.3 x 10^{-10}	Belle II	expected - 50/ab

ATLAS: extrapolated from Run 1, including effect for tracker update:

More than 1 order of magnitude improvement

CMS: extrapolated from Run 2 (2016) for just one channel, includes expected amelioration for tracker and muon system update:

More than 1 order of magnitude improvement

Quite exciting prospective for HL-LHC!

Conclusions

- Main results presented in this talk were:
  - $\mathrm{B}^0_{(s)} \to \mu^+\mu^-$ branching fractions
  - $\mathrm{B} \to \mathrm{K}^*\mu\mu$ angular analyses
  - $\tau \to 3\mu$ decay

- Future prospects were given for the expected sensitivities of HL-LHC (3000/fb)

- However these results use just fractions of the data collected so far by ATLAS and CMS

More results are in preparation with Run 2…

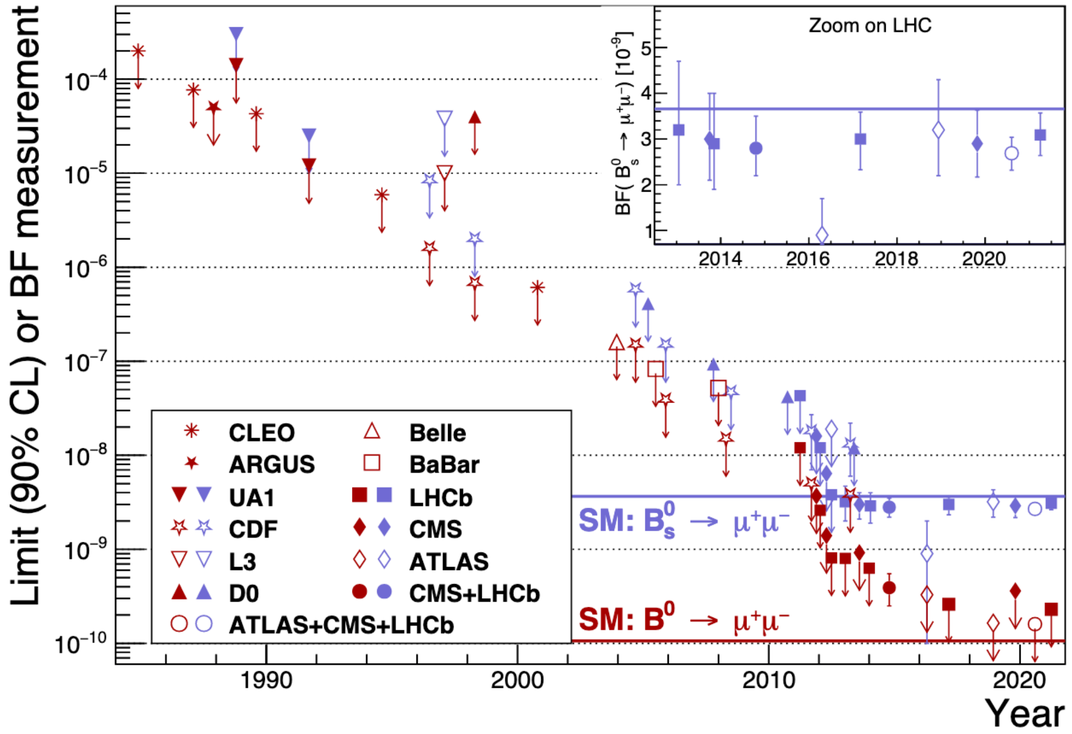
… stay tuned!

Back up

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

- A three decades long quest
- Results
  - ATLAS: JHEP 04 (2019) 098
    - $26.3\ \text{fb}^{-1}$ at $\sqrt{s} = 13$ TeV (2015 and 2016)
    - $25\ \text{fb}^{-1}$ at $\sqrt{s} = 7$ and 8 TeV (2011 and 2012)
  - CMS: JHEP 04 (2020) 188
    - $36\ \text{fb}^{-1}$ at $\sqrt{s} = 13$ TeV (2016)
    - $5(20)\ \text{fb}^{-1}$ at $\sqrt{s} = 7(8)$ TeV (in 2011(2012))

Recently combined with the LHCb !!



- LHCb: Phys. Rev. Lett. 118, 191801 (old)
  - $1.4\ \text{fb}^{-1}$ at $\sqrt{s} = 13$ TeV (2015 and 2016)
  - $1(2)\ \text{fb}^{-1}$ at $\sqrt{s} = 7(8)$ TeV (in 2011(2012))
  - $3\ \text{fb}^{-1}$ Run 1 + $6\ \text{fb}^{-1}$ Run 2 (not presented here) - LHCB-PAPER-2021-007, Archilli's Moriond EW and Santimaria's LHC seminar (new)

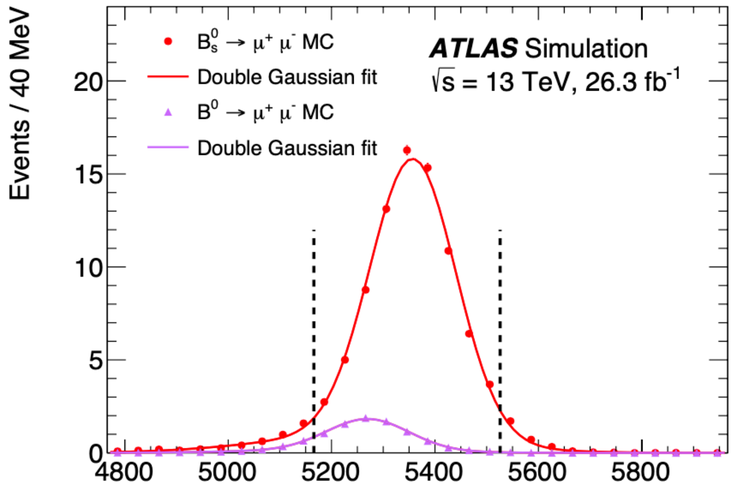
$B^0_{(s)} \to \mu^+\mu^-$

- The branching fractions of the of the $B^0_{(s)} \to \mu^+\mu^-$ channels are measured with similar strategy
  - $B^+ \to J/\psi K^+$ is used as reference channel since abundant and with a well measured branching fraction

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

N are the yields — ϵ **the acceptance times efficiency** — **ratio of the hadronization probabilities of b-quarks to** B^+ **or** $B^0_{(s)}$

- Selection:
  - starting from dimuon triggered data, pairs of well identified opposite charged muons are combined to form a displaced vertex;
  - additional criteria on isolation, kinematics and geometrical requirements
- Multivariate discriminators to distinguish signal from background events
- Analysis kept blind, i.e. signal mass window is concealed, while selection optimization and signal extraction procedure are defined
  - ATLAS: $m_{\mu^+\mu^-}$ in [5.166,5.526] GeV
  - CMS: $m_{\mu^+\mu^-}$ in [5.200,5.450] GeV

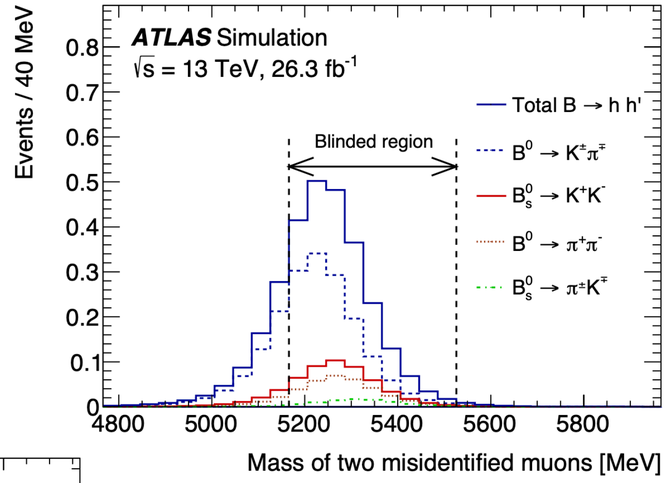


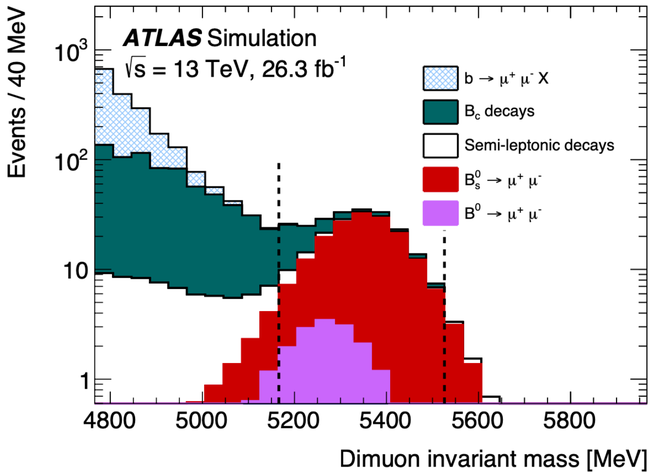
ATLAS: JHEP04(2019)098

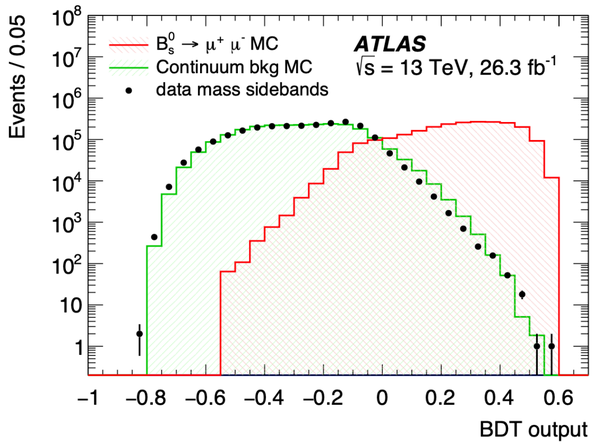
$B^0_{(s)} \to \mu^+\mu^-$: background composition

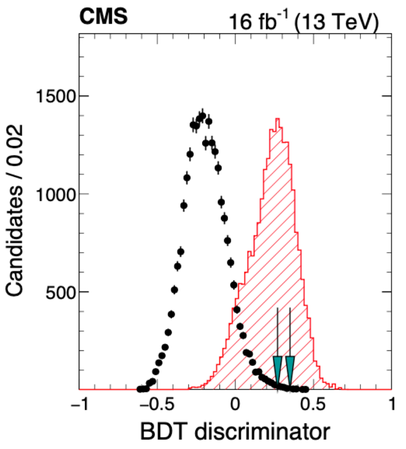
ATLAS: JHEP04(2019)098

- Peaking backgrounds: $B^0_{(S)} -> hh'$ decays with both hadrons misidentified as muons
- Partially reconstructed decays: one or more of the final-state particles (X) in a b-hadron decay are not reconstructed
  - contribution in low dimuon mass region
- Continuum background - dominant
  - Combinatorial component: muons originated from uncorrelated hadron decay
    - modeled with $bb \to \mu^+\mu^- X$ MC sample, checked in the signal mass sidebands in data (ATLAS)
    - from the mass sideband in data (CMS)
  - BDT discriminator to distinguish background form signal
    - variables related to: B meson decay, muon quality and the rest of the event (isolation, number of tracks)
      - 15 ATLAS, 12 CMS



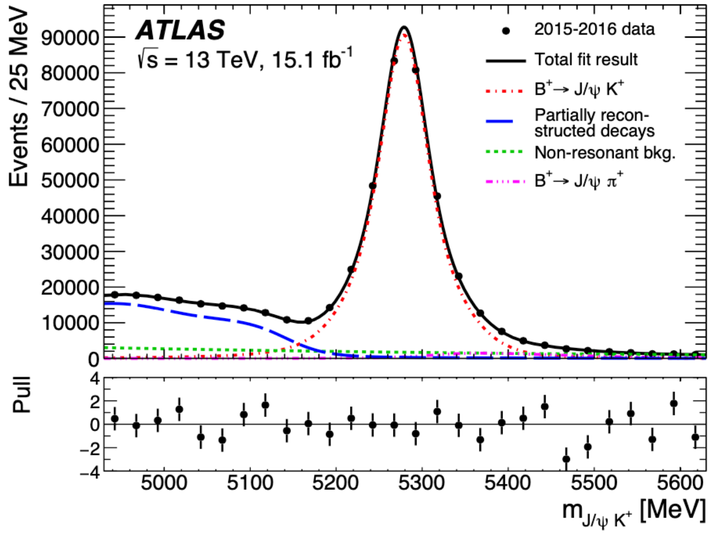




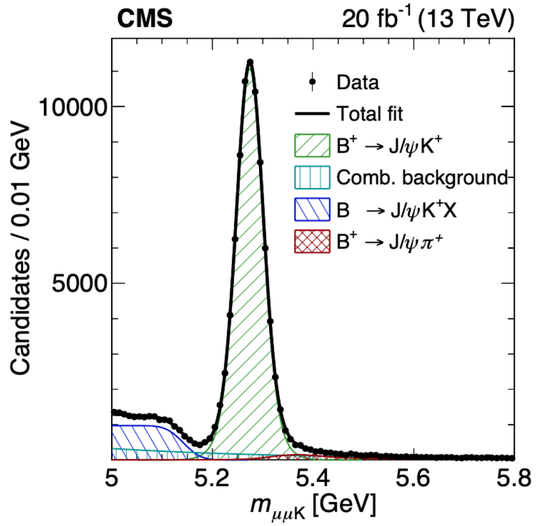


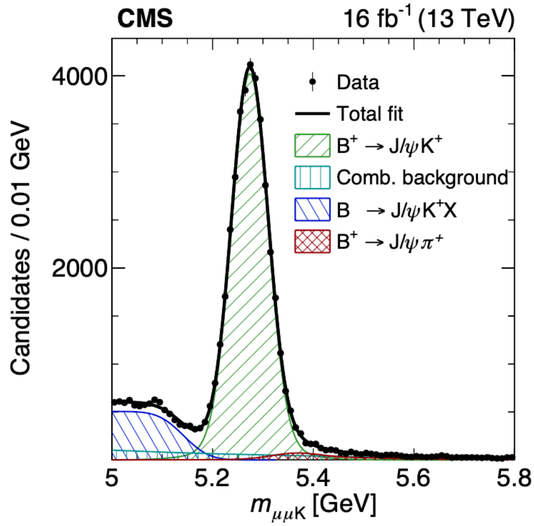
$B^0_{(s)} \to \mu^+\mu^-$: normalization channel

- The B^+ yield for the normalization channel is extracted with an unbinned extended maximum-likelihood fit to the $J/\psi K^+$ invariant mass distribution
  - In CMS split in two subsamples of the 2016 dataset in different regions of pseudorapidity regions based on the most forward muon.



- Fit includes:
  - $B^+ \to J/\psi K^+$ decays
  - Cabibbo suppressed $B \to J/\psi\pi^+$
  - Partially reconstructed decays $B \to J/\psi K^+ X$
  - Continuum combinatorial background





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

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

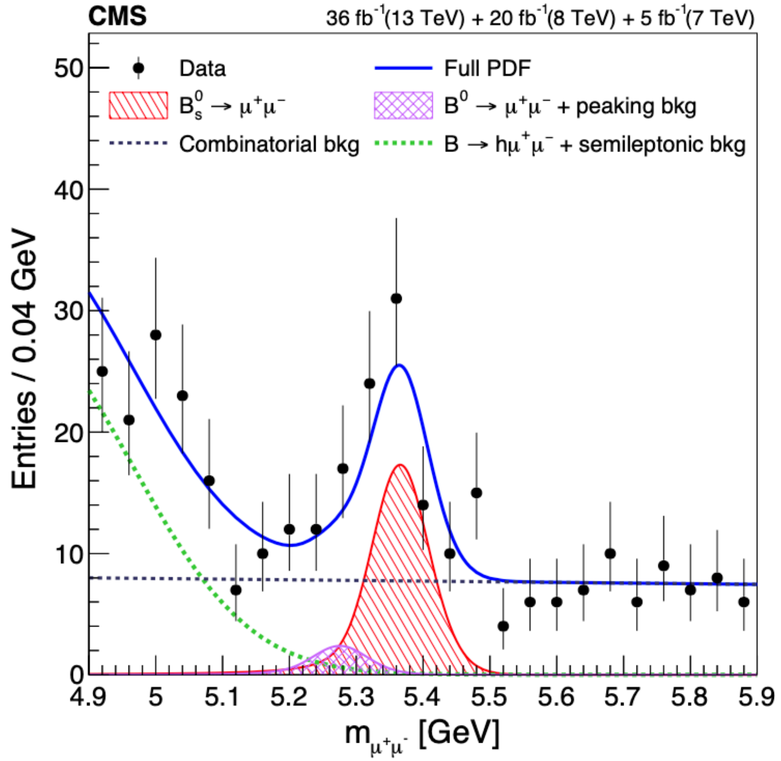
A 2D unbinned maximum likelihood fit to the dimuon invariant mass and the proper decay time is implemented for extracting the effective $B^0_s \to \mu^+\mu^-$ lifetime, including the signal and each background components, the proper decay time resolution is as a conditional parameter

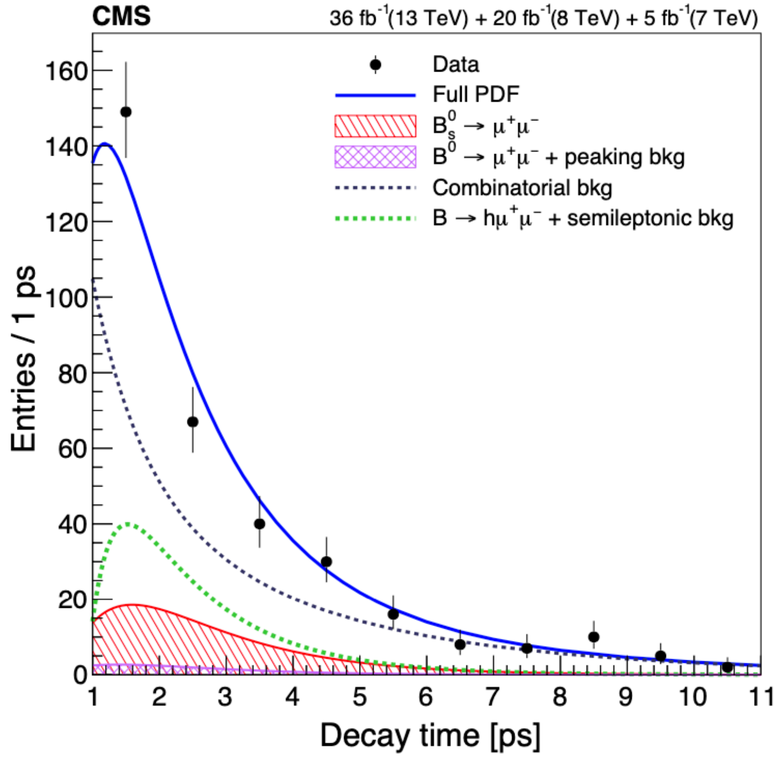
CMS

$$\tau_{B^0_s \to \mu^+\mu^-} = (1.70^{+0.60}_{-0.43} \pm 0.09)\ \text{ps}$$, which combined with

OLD LHCb

$$\tau_{B^0_s \to \mu^+\mu^-} = (2.04 \pm 0.44 \pm 0.05)\ \text{ps}$$





$$\tau_{B^0_s \to \mu^+\mu^-} = (1.91^{+0.37}_{-0.35})\ \text{ps}$$

Recent update from LHCb :

$$\tau_{B^0_s \to \mu^+\mu^-} = (2.07 \pm 0.29 \pm 0.03)\ \text{ps}$$

Archilli's presentation Moriond EW

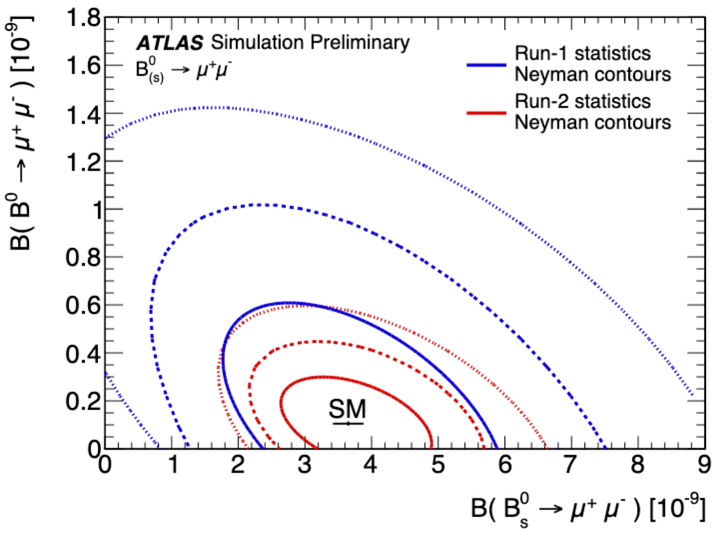
$B^0_{(s)} \to \mu^+\mu^-$: prospects

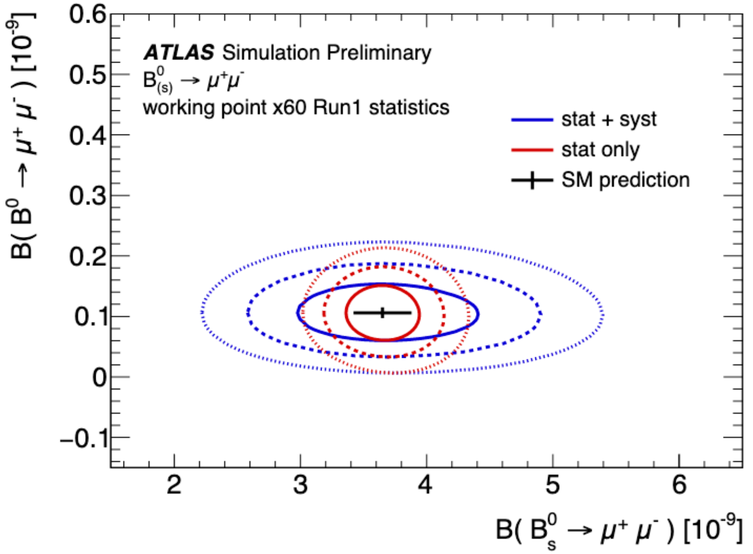
- Prediction for full Run 2 and HL-LHC

	$\mathcal{B}(B^0_s \to \mu^+\mu^-)$		$\mathcal{B}(B^0 \to \mu^+\mu^-)$	
	stat $[10^{-10}]$	stat + syst $[10^{-10}]$	stat $[10^{-10}]$	stat + syst $[10^{-10}]$
Run 2	7.0	8.3 **~30%**	1.42	1.43
HL-LHC: Conservative	3.2	5.5	0.53	0.54
HL-LHC: Intermediate	1.9	4.7 **~15%**	0.30	0.31 **~30%**
HL-LHC: High-yield	1.8	4.6	0.27	0.28

ATL-PHYS-PUB-2018-005

- In Phase-2:
  - CMS and ATLAS upgrade of the inner tracker is expected to provide a 40-50% improvement on the mass resolution (reduce B_s-B_d mass correlation)
  - Uncertainty dominated by external physics parameters on B_s (dominated by f_s/f_d)





$\mathcal{L}$ (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

precision of 3%

CMS-FTR-18-013

ATLAS BDT

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

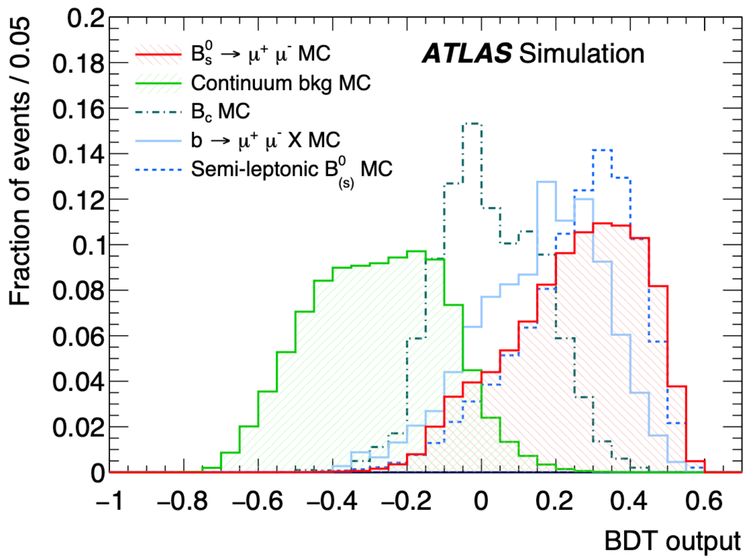


Table 1: Description of the 15 input variables used in a BDT classifier to discriminate between signal and continuum background. When the BDT classifier is applied to $B^+ \to J/\psi\, K^+$ and $B^0_s \to J/\psi\, \phi$ candidates, the variables related to the decay products of the B mesons refer only to the muons from the decay of the J/ψ. Horizontal lines separate the classifications into groups (a), (b) and (c) respectively, as described in the text. For category (c), additional tracks are required to have p_{T} >500 MeV.

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.
$\Delta R_{\mathrm{flight}}$	Three-dimensional angular distance between $\overrightarrow{p}^{B}$ and $\overrightarrow{\Delta x}$: $\sqrt{{\alpha_{\mathrm{2D}}}^2 + (\Delta\eta)^2}$
$\lvert\alpha_{\mathrm{2D}}\rvert$	Absolute value of the angle in the transverse plane between $\overrightarrow{p_{\mathrm{T}}}^{B}$ and $\overrightarrow{\Delta x}_{\mathrm{T}}$.
L_{xy}	Projection of $\overrightarrow{\Delta x}_{\mathrm{T}}$ along the direction of $\overrightarrow{p}^{B}_{\mathrm{T}}$: $(\overrightarrow{\Delta x}_{\mathrm{T}} \cdot \overrightarrow{p_{\mathrm{T}}}^{B})/\lvert\overrightarrow{p_{\mathrm{T}}}^{B}\rvert$.
$\mathrm{IP}^{\mathrm{3D}}_{B}$	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).
$\Delta\phi_{\mu\mu}$	Azimuthal angle between the momenta of the two tracks forming the B candidate.
$\lvert d_0\rvert^{\mathrm{max}}$-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.
$\lvert d_0\rvert^{\mathrm{min}}$-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.
$P_{\mathrm{L}}^{\mathrm{min}}$	The smaller of the projected values of the muon momenta 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 the transverse momenta of all additional tracks contained within a cone of size $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.7$ 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. Only tracks matched to the same PV as the B candidate are considered.
$N^{\mathrm{close}}_{\mathrm{xtrk}}$	Number of additional tracks compatible with the decay vertex (DV) of the B candidate with $\ln(\chi^2_{\mathrm{xtrk,DV}}) < 1$. Only tracks matched to the same PV as the B candidate are considered.
$\chi^2_{\mu,\mathrm{xPV}}$	Minimum χ^2 for the compatibility of a muon in the B candidate with any PV reconstructed in the event.

ATLAS efficiency systematics

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

$$\begin{aligned}\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-) &= \frac{N_{d(s)}}{\varepsilon_{\mu^+\mu^-}} \times \left[\mathcal{B}(B^+ \to J/\psi\, K^+) \times \mathcal{B}(J/\psi \to \mu^+\mu^-)\right] \frac{\varepsilon_{J/\psi K^+}}{N_{J/\psi K^+}} \times \frac{f_u}{f_{d(s)}} \\ &= N_{d(s)} \frac{\mathcal{B}(B^+ \to J/\psi\, K^+) \times \mathcal{B}(J/\psi \to \mu^+\mu^-)}{\mathcal{D}_{\text{ref}}} \times \frac{f_u}{f_{d(s)}},\end{aligned}$$

$$\mathcal{D}_{\text{ref}} = N_{J/\psi K^+} \times (\varepsilon_{\mu^+\mu^-}/\varepsilon_{J/\psi K^+}),$$

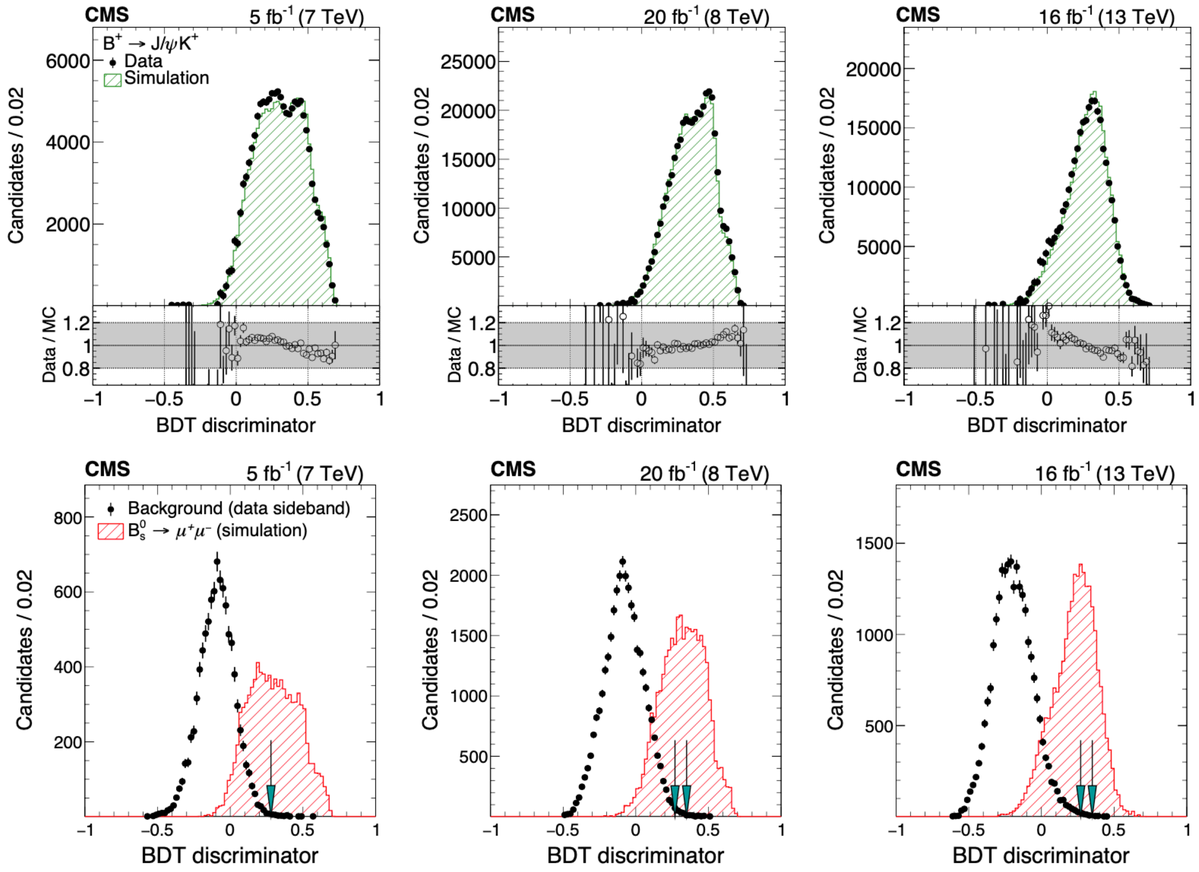
Table 2: Summary of the uncertainties in R_ε.

Source	Contribution [%]
Statistical	0.8
Kinematic reweighting (DDW)	0.8
Muon trigger and reconstruction	1.0
BDT input variables	3.2
Kaon tracking efficiency	1.5
Pile-up reweighting	0.6

Ratio between the efficiency of the reference channel and the signal

CMS BDT

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



- $l_{3D}/\sigma(l_{3D})$: The flight length significance of the B candidate (the distance between the secondary and primary vertex, divided by its uncertainty);
- $\delta_{3D}/\sigma(\delta_{3D})$: The significance of the 3D impact parameter of the B candidate with respect to the selected primary vertex;
- α_{3D}: The pointing angle of the B candidate;
- d^0_{ca}: The minimum distance of closest approach to the B candidate vertex of a track (not belonging to the B candidate) in the event;
- χ^2/ndf: The vertex fit χ^2 of the dimuon vertex;
- $N^{\text{close}}_{\text{trk}}$: The number of tracks in the vicinity of the B decay vertex;
- Isolation variables:
  - $I \equiv p_{\perp_B}/(p_{\perp_B} + \sum_{\text{trk}} p_\perp)$: The isolation of the B candidate while $p_\perp$ >0.9 GeV, $\Delta R < 0.7$ and $d^0_{\text{ca}} < 0.05$ cm.
  - $I_\mu \equiv p_{\perp_\mu}/(p_{\perp_\mu} + \sum_{\text{trk}} p_\perp)$: The isolation of the muon candidate while $p_\perp$ >0.5 GeV, $\Delta R < 0.5$ and $d^0_{\text{ca}} < 0.1$ cm.

CMS BDT

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

Category	$N(B^0_s)$	$N(B^0)$	N_{comb}	$N^{B^+}_{obs}/100$	$\langle p_T(B^0_s)\rangle$ [GeV]	$\varepsilon_{tot}/\varepsilon^{B^+}_{tot}$
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

	2011		2012		2016A		2016B	
	Central	Forward	Central	Forward	Central	Forward	Central	Forward
Low	—	—	0.27	0.23	0.19	0.19	0.18	0.23
High	0.28	0.21	0.35	0.32	0.30	0.30	0.31	0.38

Table 4: Analysis BDT discriminator minimum requirements per channel and running period for the 1D and 2D effective lifetime fits.

2011		2012		2016A		2016B	
Central	Forward	Central	Forward	Central	Forward	Central	Forward
0.22	0.19	0.32	0.32	0.22	0.30	0.22	0.29

CMS Systematics

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

Table 2: Summary of systematic uncertainty sources described in the text. The uncertainties quoted for the branching fraction $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ are relative uncertainties, while the uncertainties for the effective lifetime $\tau_{\mu^+\mu^-}$ are absolute and are given for both the 2D UML and *sPlot* analysis methods. The relative uncertainties in the upper limit on $\mathcal{B}(B^0 \to \mu^+\mu^-)$ differ for the background yields and parametrization, but have negligible impact on that result. The bottom rows provide the total systematic uncertainty and the total uncertainty in the branching fraction and the effective lifetime measurements. Contributions that are included in other items are indicated by (*). Non-applicable sources are marked with dashes.

Source	$\mathcal{B}(B^0_s \to \mu^+\mu^-)$ [%]	$\tau_{\mu^+\mu^-}$ [ps]	
		2D UML	*sPlot*
Kaon tracking	2.3–4	—	—
Normalization yield	4	—	—
Background yields	1	0.03	(*)
Production process	3	—	—
Muon identification	3	—	—
Trigger	3	—	—
Efficiency (data/MC simulation)	5–10	—	(*)
Efficiency (functional form)	—	0.01	0.04
Efficiency lifetime dependence	1–3	(*)	(*)
Running period dependence	5–6	0.07	0.07
BDT discriminator threshold	—	0.02	0.02
Silicon tracker alignment	—	0.02	—
Finite size of MC sample	—	0.03	—
Background parametrization/Fit bias	2.3	—	0.09
$\mathcal{C}$ correction	—	0.01	0.01
Absolute total systematic uncertainty	0.3×10^{-9}	0.09	0.12
Absolute total uncertainty	0.7×10^{-9}	$^{+0.61}_{-0.44}$	$^{+0.52}_{-0.33}$

CMS: Phase-2

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

Table 1: Input sources of systematic uncertainties and the propagated uncertainties on the $B \to \mu^+\mu^-$ branching fractions, $\delta\mathcal{B}(B^0_s \to \mu^+\mu^-)$ and $\delta\mathcal{B}(B^0 \to \mu^+\mu^-)$.

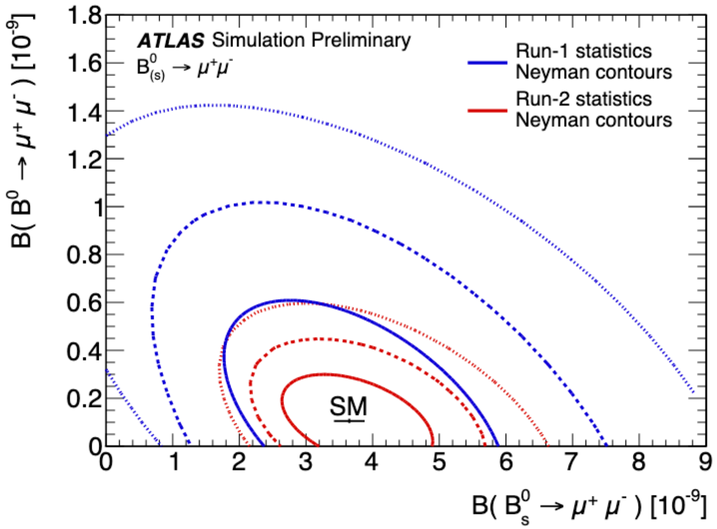
Source	Input uncertainties	$\delta\mathcal{B}(B^0_s \to \mu^+\mu^-)$	$\delta\mathcal{B}(B^0 \to \mu^+\mu^-)$
Muon ID efficiency ratio	1%	1%	1%
B^+ normalization yield	1.4%	1.4%	1.4%
f_u/f_s ratio	3.5%	3.5%	-
Effective lifetime	2%	2%	-
Trigger efficiency	1.5%	1.5%	1.5%
Other sources	3%	3%	3%
Peaking background yield	10%	0.5%	2.7%
Semileptonic background yield	7.5%		

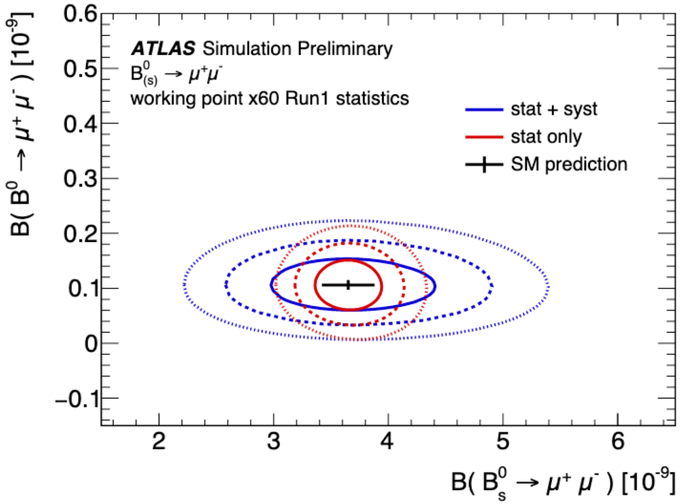
$B^0_{(s)} \to \mu^+\mu^-$: prospects

ATL-PHYS-PUB-2018-005

	$\mathcal{B}(B^0_s \to \mu^+\mu^-)$		$\mathcal{B}(B^0 \to \mu^+\mu^-)$	
	stat $[10^{-10}]$	stat + syst $[10^{-10}]$	stat $[10^{-10}]$	stat + syst $[10^{-10}]$
Run 2	7.0	8.3	1.42	1.43
HL-LHC: Conservative	3.2	5.5	0.53	0.54
HL-LHC: Intermediate	1.9	4.7	0.30	0.31
HL-LHC: High-yield	1.8	4.6	0.27	0.28

- Prediction by ALTAS:
  - Extrapolation from Run 1 results to Run-2(130fb^{-1}) and HL-LHC
  - Analysis sensitivities are shown with BR contour plots
  - Different scenarios for HL-LHC, depending on trigger [15x, 60x,75xRun 1]





Phase-2: CMS and ATLAS Phase-2 inner tracker is expected to provide a 40-50% improvement on the mass resolution
Uncertainty dominated by external physics on B_s dominated by f_s/f_d

- Prediction by CMS:

$\mathcal{L}$ (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

CMS-FTR-18-013

Run 2 results are in preparation from both ATLAS and CMS - *stay tuned!*

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

JHEP04(2021)124
CMS-PAS-BPH-15-009
$20\ \text{fb}^{-1}$ at $\sqrt{s} = 8$ TeV (2012)

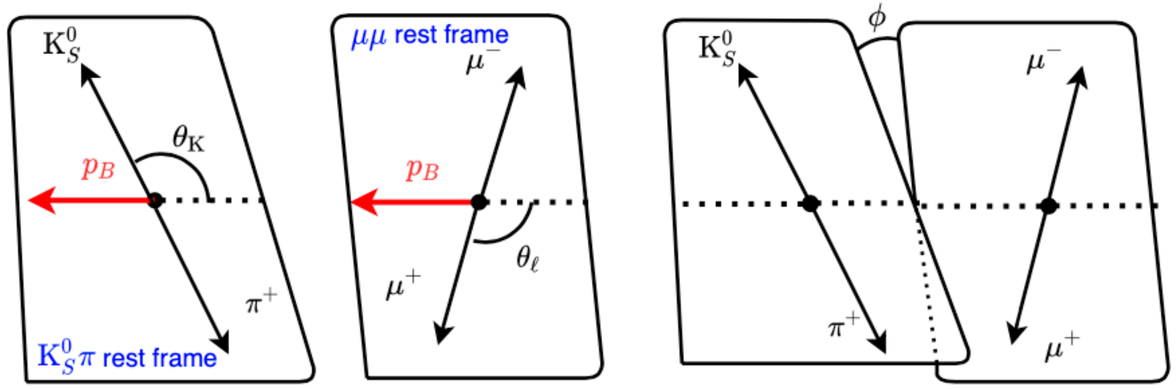
- pp collisions from 2012 corresponding to $20\ \text{fb}^{-1}$ at $\sqrt{s} = 8$ TeV:
  - dimuon trigger with requirements on displaced vertex and transverse momentum collinear with decay flight
  - K^{*+} is reconstructed through $K^0_S\pi^+$, where K^0_S is reconstructed as $\pi^+\pi^-$ with a common vertex and $|m_{\pi^+\pi^-} - m_{K^0_S}| < 17.3\text{MeV}\ (3\sigma_{m_{K^0_S}})$ [vetoed mass window compatible with $\Lambda \to p\pi^-$ decay]
  - $|m_{K^0_S\pi^+} - m_{K^{*+}}| < 100\ \text{MeV}$, and $m_{K^0_S\pi^+\mu^+\mu^+}\epsilon$ [4.76,5.8] GeV
  - selection on the dimuon invariant mass (q^2) to suppress $B^+ \to K^{*+}J/\psi$ and $B^+ \to K^{*+}\psi(2S)$ (<7-10%)
- Four independent kinematic variables: square of the dimuon invariant mass (q^2) and angles ($\theta_K, \theta_\ell, \phi$)
- Two angular distributions are used to measure two decay observables, the muon forward-backward asymmetry, A_{FB}, and the K^{*+} longitudinal polarization fraction, F_L, in bins of q^2.
- dimuon trigger with requirements on displaced vertex and transverse momentum collinear with decay length $p_{T,\mu} > 3.5\text{GeV},\ |\eta_\mu| < 2.2,\ \chi^2 > 0.1, p_{T,\mu\mu} > 6.9\text{GeV}, m_{\mu\mu}\epsilon[1,4.8]$ GeV

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

JHEP04(2021)124
CMS-PAS-BPH-15-009

- pp collisions from 2012 corresponding to $20\ fb^{-1}$ at $\sqrt{s} = 8$ TeV:
  - dimuon trigger with requirements on displaced vertex and transverse momentum collinear with decay flight
  - K^{*+} is reconstructed through $K^0_S\pi^+$, where K^0_S is reconstructed as $\pi^+\pi^-$ with a common vertex and $|m_{\pi^+\pi^-} - m_{K^0_S}| < 17.3 MeV\ (3\sigma_{m_{K^0_S}})$ [vetoed event with mass compatible with $\Lambda \to p\pi^-$ decay]
  - $|m_{K^0_S\pi^+} - m_{K^{*+}}| < 100\ MeV$, and $m_{K^0_S\pi^+\mu^+\mu^+\epsilon}$ [4.76,5.8] GeV
  - selection on the dimuon invariant mass (q^2) to suppress $B^+ \to K^{*+}J/\psi$ and $B^+ \to K^{*+}\psi(2S)$
  - Main background $K^+\mu^+\mu^-$: data-driven in mass sidebands

- Decay described by four independent kinematic variables:
  - square of the dimuon invariant mass (q^2) and angles ($\theta_K, \theta_\ell, \phi$)
- Aim is to measure two decay observables: the muon forward-backward asymmetry, A_{FB}, and the K^{*+} longitudinal polarization fraction, F_L, in bins of q^2.



$B^+ \to K^{*+}(892)\mu^+\mu^-$: angular analysis

JHEP04(2021)124
CMS-PAS-BPH-15-009

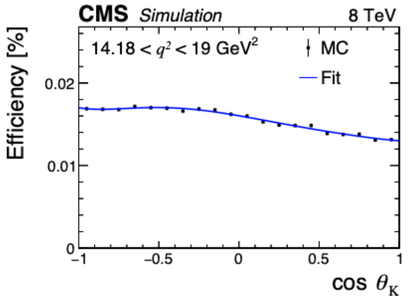
- The differential decay rate is:

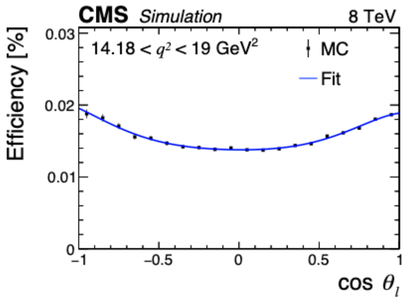
$$\frac{1}{\Gamma}\frac{d^3\Gamma}{d\cos\theta_K\, d\cos\theta_\ell\, dq^2} = \frac{9}{16}\left\{\frac{2}{3}\left[F_S + 2A_S\cos\theta_K\right](1-\cos^2\theta_\ell)\right.$$
$$+ (1-F_S)\left[2F_L\cos^2\theta_K(1-\cos^2\theta_\ell)\right.$$
$$+ \frac{1}{2}(1-F_L)\left(1-\cos^2\theta_K\right)(1+\cos^2\theta_\ell)$$
$$\left.\left. + \frac{4}{3}A_{FB}\left(1-\cos^2\theta_K\right)\cos\theta_\ell\right]\right\}.$$

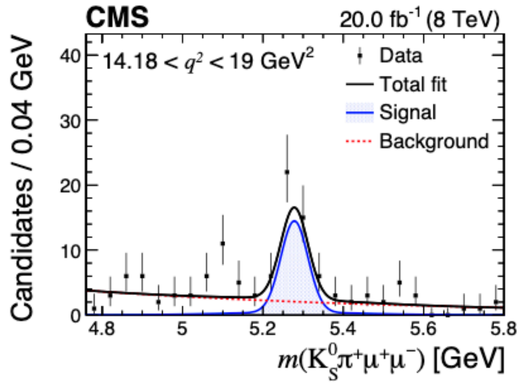
S-wave fraction

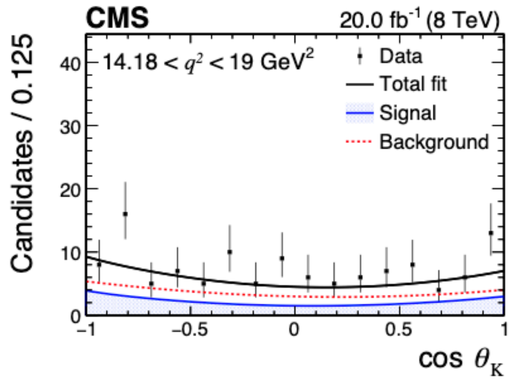
interference amplitude btw S- and P- wave

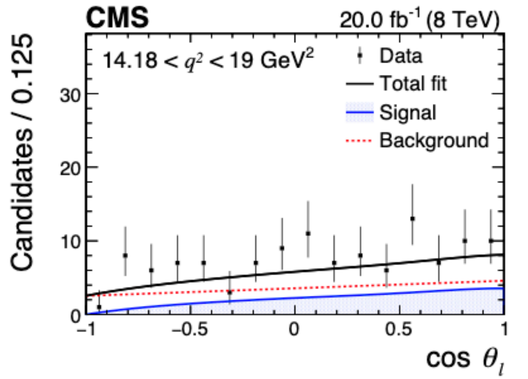
- In each bin of q^2 , A_{FB} and F_L are extracted with unbinned maximum likelihood fits, parametrized in terms of $m, \cos\theta_K, \cos\theta_\ell$:
  - F_s and A_s fixe at zero in the nominal fit, since negligible
  - Yields are obtained from fit in data
    - data-driven background from mass sidebands
  - Efficiency are obtained fitting simulations first separately on each angular variable and then considering the correlation terms
  - Data / simulation agreement are checked in $B^+ \to K^{*+}J/\psi$ enriched control region











PLB 753 (2016) 424
CMS-BPH-15-008

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

Phys. Rev. D 98, 112011
CMS-BPH-15-001

- Results in agreement with the SM:
- pp collisions from 2012 corresponding to $20\ \text{fb}^{-1}$ at $\sqrt{s} = 8$ TeV:
  - dimuon trigger with requirements on displaced vertex and transverse momentum collinear with decay flight
  - K^{*+} is reconstructed as $K^+\pi^-$ ($|m_{K^+\pi^-} - m_{K^{*0}}| < 90\text{MeV}$) with a common vertex [vetoed event with mass compatible with $\phi(1020) \to K^+K^+$ decay]
  - selection on the dimuon invariant mass (q^2) to suppress $B^0 \to K^{*0}J/\psi$ and $B^0 \to K^{*0}\psi(2S)$ (10% contamination)
  - $|m_{K^+\pi^-} - m_{K^{*0}}|$ is minimized to avoid B^0 or $\bar{B}^0$ candiate mis-assingment (12-14%)
- Background $K^*\mu^+\mu^-$:
  - Negligible contributions from misID muons from hadronic and semileptonic b-dacays
  - Few events from $\Lambda_B \to J/\psi X$ (not peaking-same as combinatorial) and $B_S \to K^{*0}\mu^+\mu^-$ <1 ev.
  - data-driven in mass sidebands

PLB 753 (2016) 424
CMS-BPH-15-008

Phys. Rev. D 98, 112011
CMS-BPH-15-001

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

- $K^+\mu^+\mu^-$:
  - NP contribution could lie in forward backward asimmetry
  - Pseudo-scalar/scalar/tensor contribution F_H
  - 23000 events across different bins

$$\frac{1}{\Gamma_\ell}\frac{d\Gamma_\ell}{d\cos\theta_\ell} = \frac{3}{4}(1-F_H)(1-\cos^2\theta_\ell) + \frac{1}{2}F_H + A_{FB}\cos\theta_\ell.$$

- $K^*\mu^+\mu^-$:
  - Folding in ϕ and θ_ℓ to simplify the fit
  - 1400 events across different bins

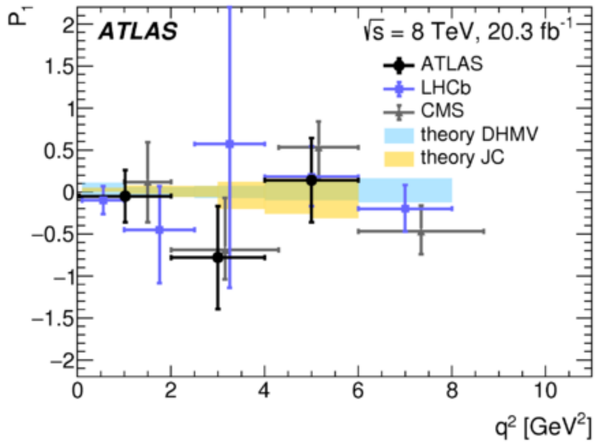
$$\begin{aligned}\frac{1}{d\Gamma/dq^2}\frac{d^4\Gamma}{dq^2\, d\cos\theta_\ell\, d\cos\theta_K\, d\varphi} = \frac{9}{8\pi}\Bigg\{&\frac{2}{3}\Big[(F_S + A_S\cos\theta_K)(1-\cos^2\theta_\ell) \\ &+ A_S^5\sqrt{1-\cos^2\theta_K}\sqrt{1-\cos^2\theta_\ell}\cos\varphi\Big] \\ &+ (1-F_S)\Big[2F_L\cos^2\theta_K(1-\cos^2\theta_\ell) \\ &+ \frac{1}{2}(1-F_L)(1-\cos^2\theta_K)(1+\cos^2\theta_\ell) \\ &+ \frac{1}{2}P_1(1-F_L)(1-\cos^2\theta_K)(1-\cos^2\theta_\ell)\cos 2\varphi \\ &+ 2P_5'\cos\theta_K\sqrt{F_L(1-F_L)}\sqrt{1-\cos^2\theta_K}\sqrt{1-\cos^2\theta_\ell}\cos\varphi\Big]\Bigg\}\end{aligned}$$

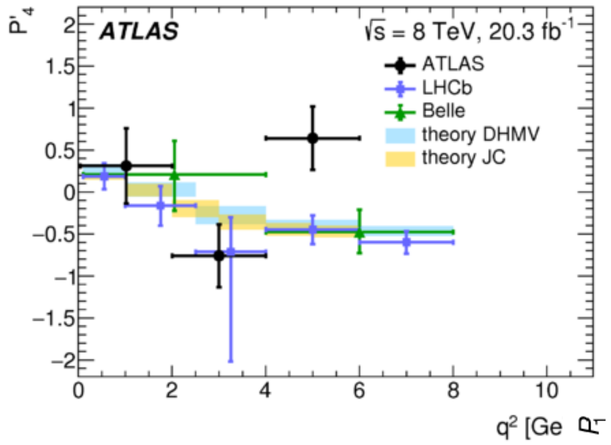
$B^0 \to K^{*0}(892)\mu^+\mu^-$

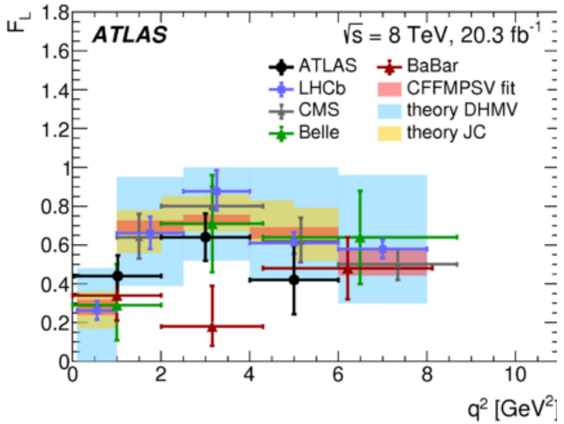
$$\frac{1}{d\Gamma/dq^2}\frac{d^4\Gamma}{d\cos\theta_L d\cos\theta_K d\phi dq^2} = \frac{9}{32\pi}\left[\frac{3(1-F_L)}{4}\sin^2\theta_K + F_L\cos^2\theta_K + \frac{1-F_L}{4}\sin^2\theta_K\cos 2\theta_L\right.$$
$$-F_L\cos^2\theta_K\cos 2\theta_L + S_3\sin^2\theta_K\sin^2\theta_L\cos 2\phi$$
$$+S_4\sin 2\theta_K\sin 2\theta_L\cos\phi + S_5\sin 2\theta_K\sin\theta_L\cos\phi$$
$$+S_6\sin^2\theta_K\cos\theta_L + S_7\sin 2\theta_K\sin\theta_L\sin\phi$$
$$\left.+S_8\sin 2\theta_K\sin 2\theta_L\sin\phi + S_9\sin^2\theta_K\sin^2\theta_L\sin 2\phi\right]$$

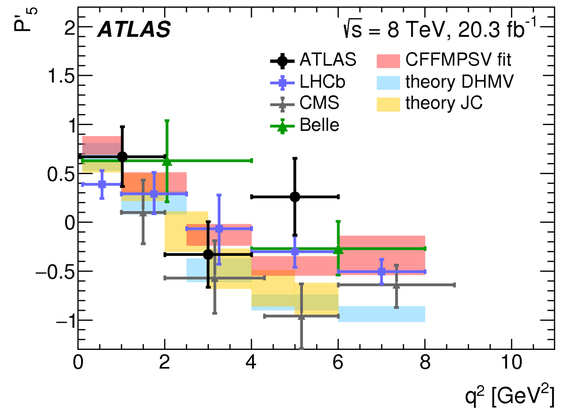
JHEP 10 (2018) 047
ATLAS-BPHY-2013-02

- ATLAS analysis at 8 TeV
  - Angular fit for several observables
    - Folding in ϕ and θ_ℓ for simplification
    - Extended ML fit in 3 q^2 bins





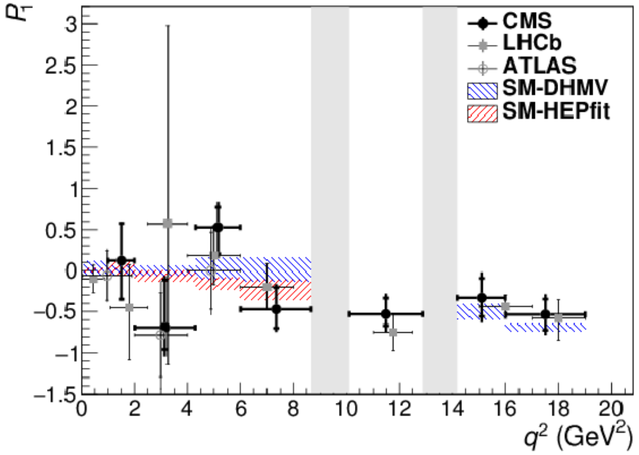


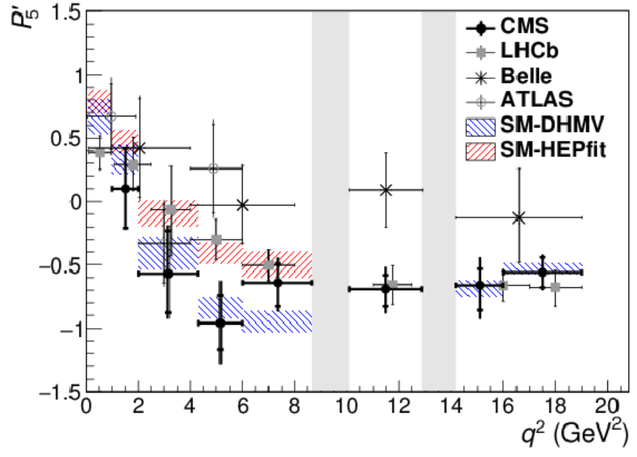


- All measurements are within 3σ range covered by predictions
- 2.7 σdeviations for P'_4, P'_5 in q^2∈[4,6] GeV²
- Consistent with what reported by LHCb

- CMS performed similar analysis at 8 TeV and fitted for P_1 and P'_5
  - precision comparable with other results
  - consistent with predictions based on the standard model.

Phys. Lett. B 781 (2018) 517
CMS-BPH-15-008





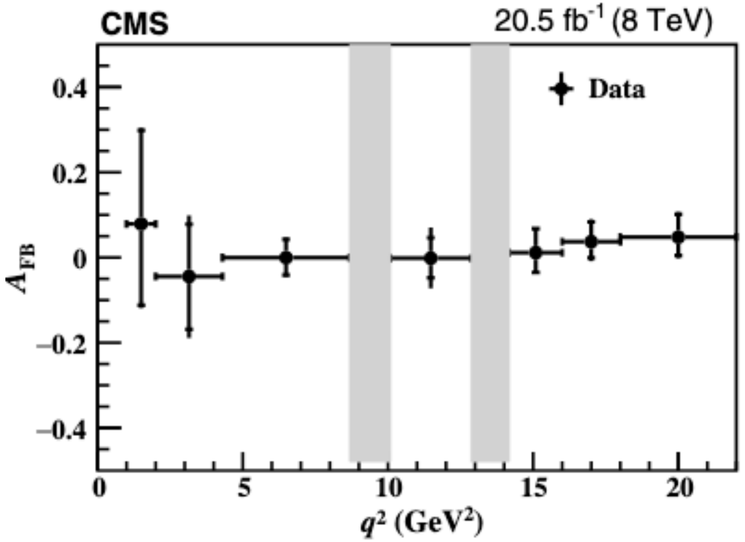
PLB 753 (2016) 424
CMS-BPH-15-008

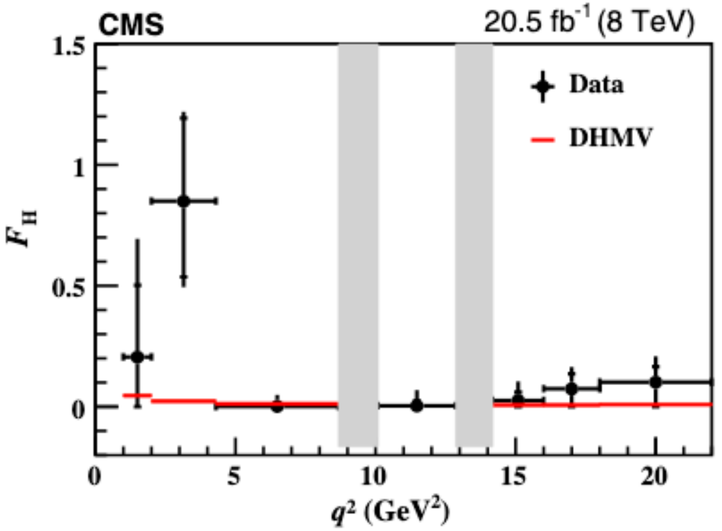
Phys. Rev. D 98, 112011
CMS-BPH-15-001

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

- $K^+\mu^+\mu^-$:

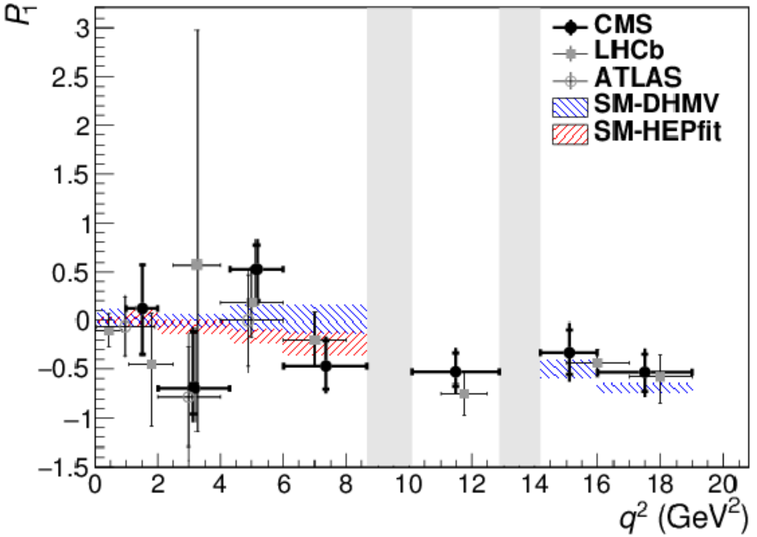
Systematic uncertainty	$A_{FB}(\times 10^{-2})$	$F_H(\times 10^{-2})$
Finite size of MC samples	0.4–1.8	0.9–5.0
Efficiency description	0.1–1.5	0.1–7.8
Simulation mismodeling	0.1–2.8	0.1–1.4
Background parametrization model	0.1–1.0	0.1–5.1
Angular resolution	0.1–1.7	0.1–3.3
Dimuon mass resolution	0.1–1.0	0.1–1.5
Fitting procedure	0.1–3.2	0.4–25
Background distribution	0.1–7.2	0.1–29
Total systematic uncertainty	1.6–7.5	4.4–39

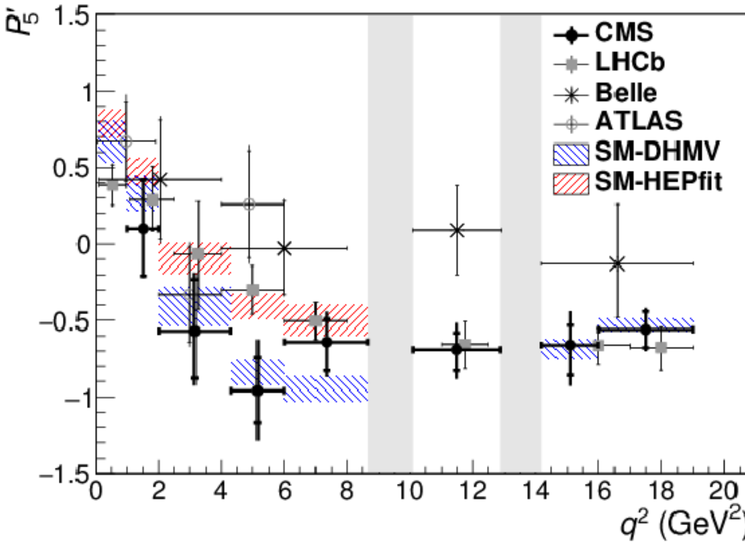




- $K^*\mu^+\mu^-$:
  - Folding in ϕ and θ_ℓ to simplify the fit
  - 1400 events across different bins

Source	$P_1(\times 10^{-3})$	$P_5'(\times 10^{-3})$
Simulation mismodeling	1–33	10–23
Fit bias	5–78	10–120
Finite size of simulated samples	29–73	31–110
Efficiency	17–100	5–65
Kπ mistagging	8–110	6–66
Background distribution	12–70	10–51
Mass distribution	12	19
Feed-through background	4–12	3–24
F_L, F_S, A_S uncertainty propagation	0–210	0–210
Angular resolution	2–68	0.1–12
Total	100–230	70–250





Consistent with SM prediction

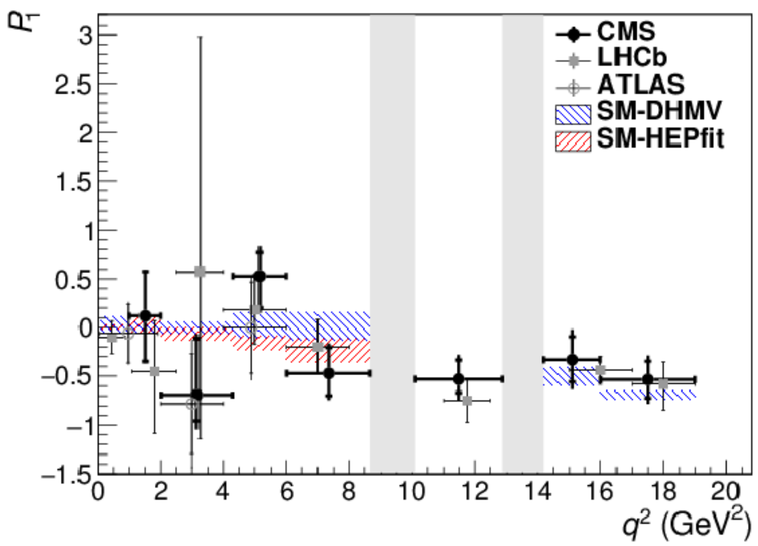
PLB 753 (2016) 424
CMS-BPH-15-008

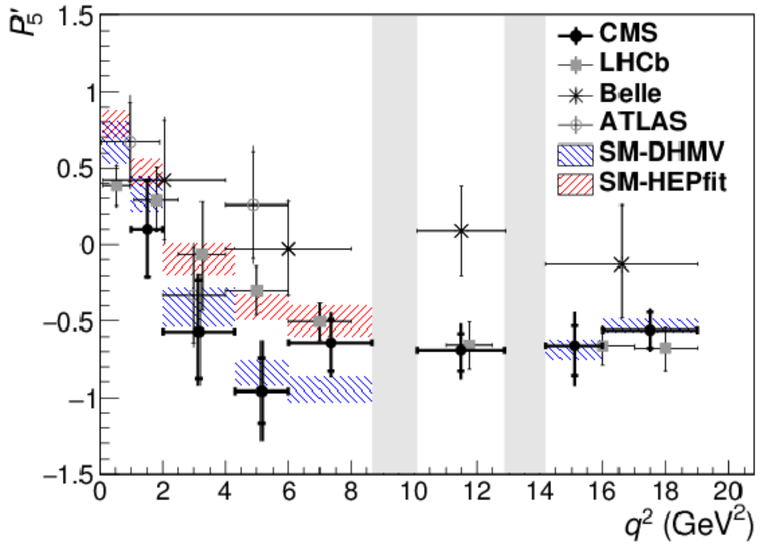
Phys. Rev. D 98, 112011
CMS-BPH-15-001

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

- $K^*\mu^+\mu^-$:
  - Folding in ϕ and θ_ℓ to simplify the fit
  - 1400 events across different bins

Source	$P_1(\times 10^{-3})$	$P_5'(\times 10^{-3})$
Simulation mismodeling	1–33	10–23
Fit bias	5–78	10–120
Finite size of simulated samples	29–73	31–110
Efficiency	17–100	5–65
$K\pi$ mistagging	8–110	6–66
Background distribution	12–70	10–51
Mass distribution	12	19
Feed-through background	4–12	3–24
F_L, F_S, A_S uncertainty propagation	0–210	0–210
Angular resolution	2–68	0.1–12
Total	100–230	70–250



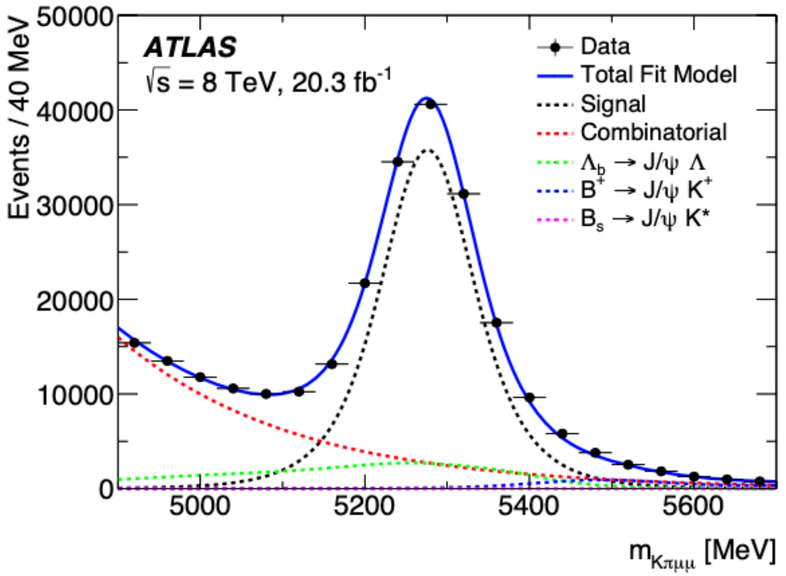


q^2 (GeV^2)	Signal yield	P_1	P_5'	Correlations
1.00–2.00	80 ± 12	$+0.12\ ^{+0.46}_{-0.47} \pm 0.10$	$+0.10\ ^{+0.32}_{-0.31} \pm 0.07$	-0.0526
2.00–4.30	145 ± 16	$-0.69\ ^{+0.58}_{-0.27} \pm 0.23$	$-0.57\ ^{+0.34}_{-0.31} \pm 0.18$	-0.0452
4.30–6.00	119 ± 14	$+0.53\ ^{+0.24}_{-0.33} \pm 0.19$	$-0.96\ ^{+0.22}_{-0.21} \pm 0.25$	$+0.4715$
6.00–8.68	247 ± 21	$-0.47\ ^{+0.27}_{-0.23} \pm 0.15$	$-0.64\ ^{+0.15}_{-0.19} \pm 0.13$	$+0.0761$
10.09–12.86	354 ± 23	$-0.53\ ^{+0.20}_{-0.14} \pm 0.15$	$-0.69\ ^{+0.11}_{-0.14} \pm 0.13$	$+0.6077$
14.18–16.00	213 ± 17	$-0.33\ ^{+0.24}_{-0.23} \pm 0.20$	$-0.66\ ^{+0.13}_{-0.20} \pm 0.18$	$+0.4188$
16.00–19.00	239 ± 19	$-0.53 \pm 0.19 \pm 0.16$	$-0.56 \pm 0.12 \pm 0.07$	$+0.4621$

Consistent with SM prediction

JHEP 10 (2018) 047
ATLAS-BPHY-2013-02

$B^0 \to K^{*0}(892)\mu^+\mu^-$: ATLAS



- pp collisions from 2012 corresponding to $20.3\ \mathrm{fb}^{-1}$ at $\sqrt{s} = 8$ TeV:
  - single, di- and tri-muon trigger
  - Displacement, PV-pointing and vertex χ^2
  - K^{*+} is reconstructed as $K^+\pi^-$ ($|m_{K^+\pi^-} - m_{K^{*0}}| < 50\mathrm{MeV}$)
    [vetoed event with mass compatible with $\phi(1020) \to K^+K^+$ decay]
  - selection on q^2 to suppress $B^0 \to K^{*0}J/\psi$ and $B^0 \to K^{*0}\psi(2S)$:
    - used as control sample to extract values for nuisance parameters describing the signal probability density function from data (M_B,resolution)
  - $|m_{K^+\pi^-} - m_{K^{*0}}|$ is minimized to avoid B^0 or $\bar{B}^0$ candiate mis-assingment (10%)
- Background $K^*\mu^+\mu^-$:
  - Combinatorial background is suppressed by requiring a B^0 candidate lifetime significance $\tau/\sigma_\tau > 12.5$
  - Radiative $J/\psi K^* \to J/\psi\gamma K^*$ contribution at high q^2 (>6 GeV)
  - Partially reconstructed $B \to \mu^+\mu^- X$ suppressed excluding low tail:
    - residual contribution accounted as systematic
  - Moreover, accounted in the systematic evaluation:
    - 2 background peaking at $cos(\theta_K) = 1$: mis-reconstructed $B^+ \to K^+\mu^+\mu^-$ and $B^+ \to \pi^+\mu^+\mu^-$, Fake K* from 2 charged tracks
    - peaking at $cos(\theta_\ell) = 0.7$ partially reconstructed $B \to D$ decays

JHEP 10 (2018) 047
ATLAS-BPHY-2013-02

$B^0 \to K^{*0}(892)\mu^+\mu^-$: ATLAS systematics

Table 3: The values of P_1, P_4', P_5', P_6' and P_8' parameters obtained for different bins in q^2. The uncertainties indicated are statistical and systematic, respectively.

q^2 [GeV^2]	P_1	P_4'	P_5'	P_6'	P_8'
[0.04, 2.0]	$-0.05 \pm 0.30 \pm 0.08$	$0.31 \pm 0.40 \pm 0.20$	$0.67 \pm 0.26 \pm 0.16$	$-0.18 \pm 0.21 \pm 0.04$	$-0.29 \pm 0.48 \pm 0.18$
[2.0, 4.0]	$-0.78 \pm 0.51 \pm 0.34$	$-0.76 \pm 0.31 \pm 0.21$	$-0.33 \pm 0.31 \pm 0.13$	$0.31 \pm 0.28 \pm 0.19$	$1.07 \pm 0.41 \pm 0.39$
[4.0, 6.0]	$0.14 \pm 0.43 \pm 0.26$	$0.64 \pm 0.33 \pm 0.18$	$0.26 \pm 0.35 \pm 0.18$	$0.06 \pm 0.27 \pm 0.13$	$-0.24 \pm 0.42 \pm 0.09$
[0.04, 4.0]	$-0.22 \pm 0.26 \pm 0.16$	$-0.30 \pm 0.24 \pm 0.17$	$0.32 \pm 0.21 \pm 0.11$	$0.01 \pm 0.17 \pm 0.10$	$0.38 \pm 0.33 \pm 0.24$
[1.1, 6.0]	$-0.17 \pm 0.31 \pm 0.13$	$0.05 \pm 0.22 \pm 0.14$	$0.01 \pm 0.21 \pm 0.08$	$0.03 \pm 0.17 \pm 0.12$	$0.23 \pm 0.28 \pm 0.20$
[0.04, 6.0]	$-0.15 \pm 0.23 \pm 0.10$	$0.05 \pm 0.20 \pm 0.14$	$0.27 \pm 0.19 \pm 0.06$	$0.03 \pm 0.15 \pm 0.10$	$0.14 \pm 0.27 \pm 0.17$

Table 2: The values of F_L, and S_3, S_4, S_5, S_7 and S_8 parameters obtained for different bins in q^2. The uncertainties indicated are statistical and systematic, respectively.

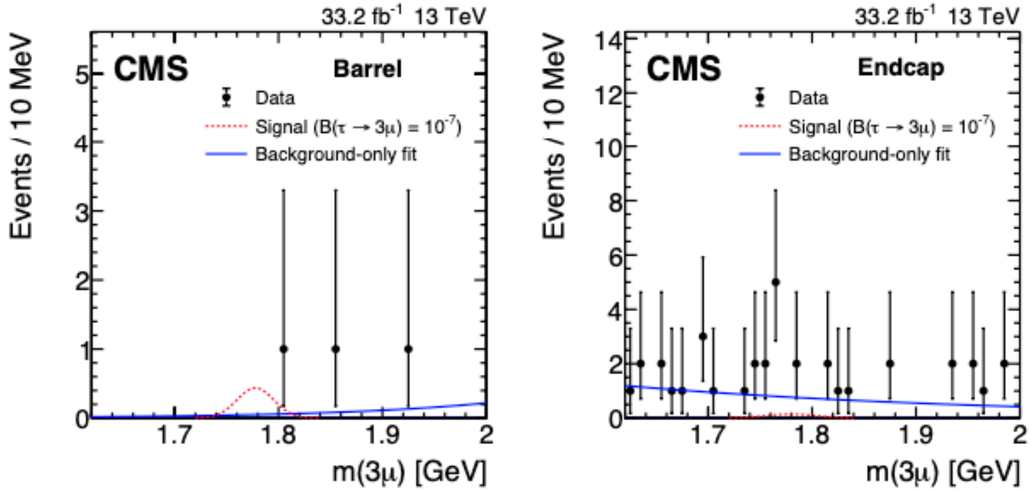
q^2 [GeV^2]	F_L	S_3	S_4	S_5	S_7	S_8
[0.04, 2.0]	$0.44 \pm 0.08 \pm 0.07$	$-0.02 \pm 0.09 \pm 0.02$	$0.15 \pm 0.20 \pm 0.10$	$0.33 \pm 0.13 \pm 0.08$	$-0.09 \pm 0.10 \pm 0.02$	$-0.14 \pm 0.24 \pm 0.09$
[2.0, 4.0]	$0.64 \pm 0.11 \pm 0.05$	$-0.15 \pm 0.10 \pm 0.07$	$-0.37 \pm 0.15 \pm 0.10$	$-0.16 \pm 0.15 \pm 0.06$	$0.15 \pm 0.14 \pm 0.09$	$0.52 \pm 0.20 \pm 0.19$
[4.0, 6.0]	$0.42 \pm 0.13 \pm 0.12$	$0.00 \pm 0.12 \pm 0.07$	$0.32 \pm 0.16 \pm 0.09$	$0.13 \pm 0.18 \pm 0.09$	$0.03 \pm 0.13 \pm 0.07$	$-0.12 \pm 0.21 \pm 0.05$
[0.04, 4.0]	$0.52 \pm 0.07 \pm 0.06$	$-0.05 \pm 0.06 \pm 0.04$	$-0.15 \pm 0.12 \pm 0.09$	$0.16 \pm 0.10 \pm 0.05$	$0.01 \pm 0.08 \pm 0.05$	$0.19 \pm 0.16 \pm 0.12$
[1.1, 6.0]	$0.56 \pm 0.07 \pm 0.06$	$-0.04 \pm 0.07 \pm 0.03$	$0.03 \pm 0.11 \pm 0.07$	$0.00 \pm 0.10 \pm 0.04$	$0.02 \pm 0.08 \pm 0.06$	$0.11 \pm 0.14 \pm 0.10$
[0.04, 6.0]	$0.50 \pm 0.06 \pm 0.04$	$-0.04 \pm 0.06 \pm 0.03$	$0.03 \pm 0.10 \pm 0.07$	$0.14 \pm 0.09 \pm 0.03$	$0.02 \pm 0.07 \pm 0.05$	$0.07 \pm 0.13 \pm 0.09$

Source	F_L	S_3	S_4	S_5	S_7	S_8
Combinatoric $K\pi$ (fake K^+) background	0.03	0.03	0.05	0.04	0.06	0.16
D and B^+ veto	0.11	0.04	0.05	0.04	0.01	0.06
Background pdf shape	0.04	0.04	0.03	0.03	0.03	0.01
Acceptance function	0.01	0.01	0.07	0.01	0.01	0.01
Partially reconstructed decay background	0.03	0.05	0.02	0.08	0.05	0.06
Alignment and B field calibration	0.02	0.04	0.05	0.04	0.04	0.04
Fit bias	0.01	0.01	0.02	0.03	0.01	0.05
Data/MC differences for p_T	0.02	0.02	0.01	0.01	0.01	0.01
S-wave	0.01	0.01	0.01	0.01	0.01	0.03
Nuisance parameters	0.01	0.01	0.01	0.01	0.01	0.01
Λ_b, B^+ and B_s background	0.01	0.01	0.01	0.01	0.01	0.01
Misreconstructed signal	0.01	0.01	0.01	0.01	0.01	0.01
Dilution	–	–	-	< 0.01	–	< 0.01

$\tau \to 3\mu$: CMS (W production)

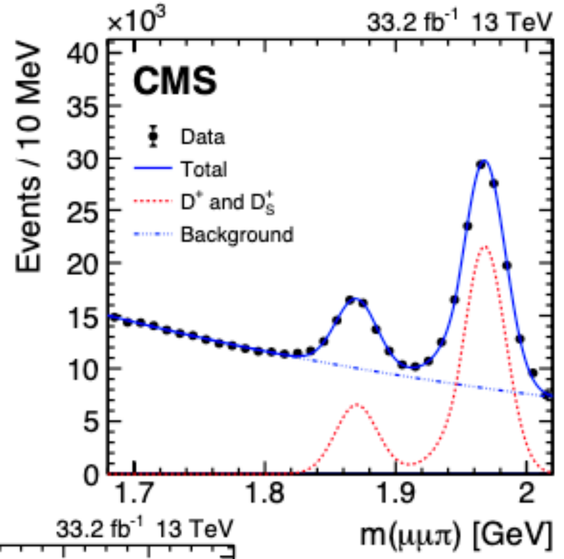
- Trigger: three muons or two muons and a track with mass and vertex requirements
- Candidate: three vertexing muon with $m_{3\mu}$ in [1.6,2.00] GeV
  - requirements on the invariant mass of any of two muons to remove background from meson/resonances
- Combination of results from $W \to \tau$ and $B, D \to \tau$ production
  - BDT discriminator is trained with signal and background events from data ([1.60,1.74] and [1.82,2.00]) with 18 variables regarding muon quality, displacement, isolation
    - 2 categories: endcap and barrel

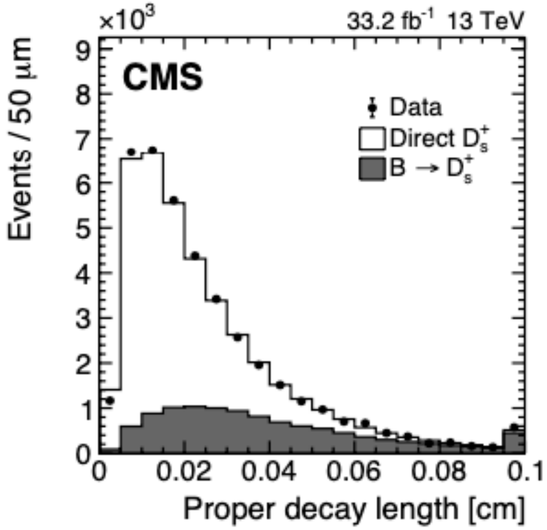
$$\mathcal{B}(\tau \to 3\mu) = \frac{N_{\mathrm{sig(W)}}}{\mathcal{L}\,\sigma(\mathrm{pp} \to \mathrm{W} + \mathrm{X})\,\mathcal{B}(\mathrm{W} \to \tau\nu)\,\mathcal{A}_{3\mu(\mathrm{W})}\,\epsilon_{3\mu(\mathrm{W})}}$$



$\tau \to 3\mu$: CMS (B-D production)

- Trigger: three muons or two muons and a track with mass and vertex requirements
- Candidate: three vertexing muon with $m_{3\mu}$ in [1.6,2.00] GeV
  - requirements on the invariant mass of any of two muons to remove background from meson/resonances
  - from $B, D \to \tau$ production
  - $D \to \tau$ (65%), $B \to \tau$(25%), $B \to D \to \tau$ (10%) [~95% from D_S]
  - $D_S \to \phi\pi \to \mu\mu\pi$ used as normalization channel to reduce uncertainties in the production of heavy-flavor hadrons
    - selected similarly to signal for trigger and kinematic
    - fraction of the D_S candidate from B meson decay is established by a fit to the proper decay length





$\tau \to 3\mu$:CMS syst

JHEP01(2021)163

Source	Uncertainty (%) Barrel	Endcap
Signal efficiency	7.9	32
Limited size of simulated samples	4.3	6.2
Integrated luminosity	2.5	2.5
pp → W cross section	2.9	2.9
$\mathcal{B}(W \to \mu\nu)$	0.2	0.2
$\mathcal{B}(W \to \tau\nu)$	0.2	0.2
Total	9.8	33

Source of uncertainty	Uncertainty (%)	Yield (%)
D_s^+ normalization	10	10
$\mathcal{B}(D_s^+ \to \tau^+\nu)$	4	3
$\mathcal{B}(D_s^+ \to \phi\pi^+ \to \mu^+\mu^-\pi^+)$	8	8
$\mathcal{B}(B \to D_s^+ + X)$	16	5
$\mathcal{B}(B \to \tau + X)$	11	3
B/D ratio f	11	3
Number of events from L1 trimuon trigger	12	3
Acceptance ratio $\mathcal{A}_{3\mu}/\mathcal{A}_{\mu\mu\pi}$	1	1
Muon reconstruction efficiency	1	1
BDT requirement efficiency	5	5
Total		16

$\tau \to 3\mu$:ATLAS

Table 2 The event count for the different steps of the analysis in the sideband and signal regions. The signal sample used to evaluate the $\mathcal{A}_s \times \epsilon_s$ has 2×10^5 events

Phase	Data SB	Data SR	Signal MC SR [out of 2×10^5]
loose	2248	580	12672
loose+$x>x_0$	736	203	12557
tight	42	9	5503
tight+$x>x_0$	28	7	5501
tight+$x>x_1$	0	0	4616

The European Physical Journal C, 76, 232 (2016)

- 1)Loose selection: displacement of the vertex,
  3 muon kinematics, track isolation and missing energy

- 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^0_{xy} significance, $S(a^0_{xy})=a^0_{xy}/\sigma_{a^0_{xy}}$, must satisfy $S(a^0_{xy})<25$, where $\sigma_{a^0_{xy}}$ is the uncertainty in a^0_{xy}.
- 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 missing transverse energies, $E_{\text{T,cal}}^{\text{miss}}$ and $E_{\text{T,trk}}^{\text{miss}}$, must both satisfy $10 < E_{\text{T}}^{\text{miss}} < 250$ GeV.
- The transverse masses, $m_{\text{T}}^{\text{cal}}$ and $m_{\text{T}}^{\text{trk}}$, 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 with $p_{\text{T}}^{\text{trk}} > 500$ 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 $\Sigma p_{\text{T}}^{\text{trk}}(\Delta R_{\text{max}}^{3\mu} + 0.20)/p_{\text{T}}^{3\mu} < 0.3$ (and $\Sigma p_{\text{T}}^{\text{trk}}(\Delta R_{\text{max}}^{3\mu} + 0.30)/p_{\text{T}}^{3\mu} < 1$). 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.

- 2)BDT is trained

1. The calorimeter-based transverse mass, $m_{\text{T}}^{\text{cal}}$.
2. The track-based missing transverse momentum, $E_{\text{T,trk}}^{\text{miss}}$.
3. The isolation variable, $\Sigma p_{\text{T}}^{\text{trk}}(\Delta R_{\text{max}}^{3\mu} + 0.20)/p_{\text{T}}^{3\mu}$.
4. The transverse component of the vector sum of the three-muon and leading jet momenta, Σ_{T}.
5. The track-based transverse mass, $m_{\text{T}}^{\text{trk}}$.
6. The difference between the $E_{\text{T,cal}}^{\text{miss}}$ and $E_{\text{T,trk}}^{\text{miss}}$ directions, $\Delta\phi_{\text{trk}}^{\text{cal}}$.
7. The calorimeter-based missing transverse momentum, $E_{\text{T,cal}}^{\text{miss}}$.
8. The track-based missing transverse momentum balance $p_{\text{T}}^{3\mu}/E_{\text{T,trk}}^{\text{miss}} - 1$.
9. The difference between the three-muon and $E_{\text{T,cal}}^{\text{miss}}$ directions, $\Delta\phi_{3\mu}^{\text{cal}}$.
10. Three-muon vertex fit probability, p-value.
11. The three-muon vertex fit a^0_{xy} significance, $S(a^0_{xy})$.
12. The track fit probability product, $\mathcal{P}_{\text{trks}}$.
13. The three-muon transverse momentum, $p_{\text{T}}^{3\mu}$.
14. The number of tracks associated with the PV (after refitting the PV while excluding the three-muon tracks), $N_{\text{trk}}^{\text{PV}}$.
15. The three-muon vertex fit L_{xy} significance, $S(L_{xy})$.
16. The calorimeter-based missing transverse momentum balance, $p_{\text{T}}^{3\mu}/E_{\text{T,cal}}^{\text{miss}} - 1$.

- 3)Tight selection: tighter requirements and minimal BDT score

Additional *tight* cuts are applied after the BDT training and the application of the $x>x_0$ cut on the BDT score. The following requirements are tightened or added:

- A number of the *loose* requirements are tightened, namely $\mathcal{P}_{\text{trks}} > 8 \times 10^{-9}$, $m_{\text{T}}^{\text{cal}} > 45$ GeV, $m_{\text{T}}^{\text{trk}} > 45$ GeV and $1<S(L_{xy})<50$.
- Three-muon vertex fit probability must have $p-\text{value} > 0.2$.
- The angle between the Σ_{T} and $E_{\text{T,cal}}^{\text{miss}}$ ($E_{\text{T,trk}}^{\text{miss}}$) directions is required to be $\Delta\phi_{\Sigma_{\text{T}}}^{\text{cal}} > 2$ ($\Delta\phi_{\Sigma_{\text{T}}}^{\text{trk}} > 2$).
- The same-charge two-muon mass, m_{SS}, and opposite-charge two-muon mass, m_{OS1} or m_{OS2}, satisfy $m_{\text{SS}} > 300\,\text{MeV}, m_{\text{OS1}} > 300\,\text{MeV}$ and $m_{\text{OS2}} > 300\,\text{MeV}$, where m_{OS1} (m_{OS2}) is the mass of the two opposite-charge muon pairs with the highest (second highest) summed scalar p_{T} among the three muons.
- The event is rejected if $|m_{\text{OS}} - m_\omega| < 50\,\text{MeV}$ or $|m_{\text{OS}} - m_\phi| < 50\,\text{MeV}$ if either of 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 $|m_{\text{OS}} - m_\phi| < 50\,\text{MeV}$ if $|m_{3\mu} - m_{D_s}| < 100\,\text{MeV}$.

$\tau \to 3\mu$: prospects

arxiv:1812.07638

ATLAS: extrapolated from Run 1, including effect for tracker update

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: extrapolated from Run 2 (2016), includes expected amelioration for tracker and muon system update

	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}	

BR($\tau \to 3\mu$) (90% CL limit)	Ref.	Comments
3.8×10^{-7}	ATLAS [429]	Actual limit (Run 1)
4.6×10^{-8}	LHCb [428]	Actual limit (Run 1)
3.3×10^{-8}	BaBar [417]	Actual limit
2.1×10^{-8}	Belle [423]	Actual limit
3.7×10^{-9}	CMS HF-channel at HL-LHC	Expected limit (3000 fb^{-1})
6×10^{-9}	ATLAS W-channel at HL-LHC	Expected limit (3000 fb^{-1})
2.3×10^{-9}	ATLAS HF-channel at HL-LHC	Expected limit (3000 fb^{-1})
$\mathcal{O}(10^{-9})$	LHCb at HL-LHC	Expected limit (300 fb^{-1})
3.3×10^{-10}	Belle-II [196]	Expected limit (50 ab^{-1})

Quite exciting prospective for HL-LHC!