Blois 2023:
34th Rencontres de Blois on "Particle Physics and Cosmology"

Blois - May 17th, 2023

Rare decays and flavour anomalies at CMS

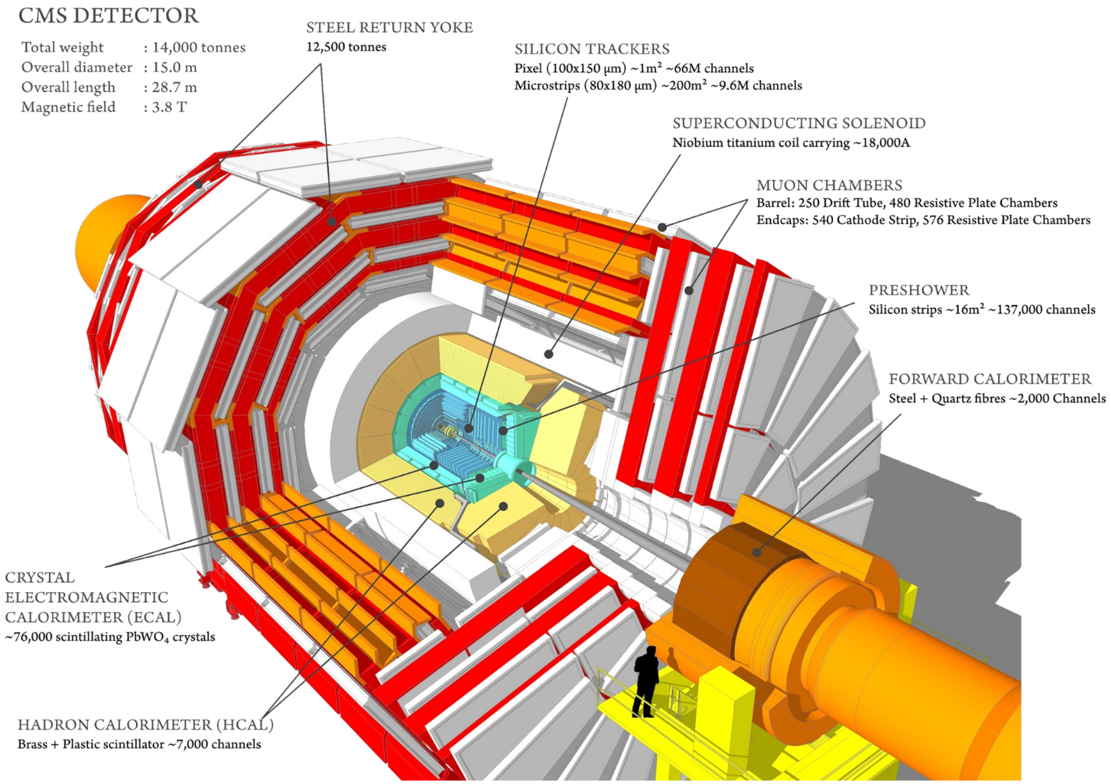
Luca Guzzi – INFN Milano-Bicocca

Overview

results from rare decays and flavour anomalies searched in B-physics at CMS

- observation of the rare $\eta \to 4\mu$ decay at CMS in Run2 scouting data
- search for $B^0_{(s)} \to \mu^+\mu^-$ events at CMS and B^0_s life-time measurement in Run2 data
- search for the LFV $\tau \to 3\mu$ decays in CMS in 2016 data
- angular analysis $B0 \to K^{0*}\mu^+\mu^-$ at CMS in Run1 data

The CMS detector



collected luminosity:

- Run1: X 25 /fb pp @ 7 and 8 TeV
- Run2: X 140 /fb pp @ 13 TeV
- Run3 ongoing, 37 /fb collected in 2022

- cylindric compact (15m x 21m) detector
- high granularity pixel + strip silicon tracker for excellent track, PV and SV measurements
- $PbWO_4$ crystal ECAL and brass+plastic HCAL to achieve hermeticity and for jet+EG shower measurement
- 3.8T solenoid for pT measurement
- external muon chambers outside steel return yoke for a clean muon detection and pT measurement
- two level trigger system (hardware + software)

$\eta \to 4\mu$

doi.org/10.48550/arXiv.2305.04904

η→4μ
introduction

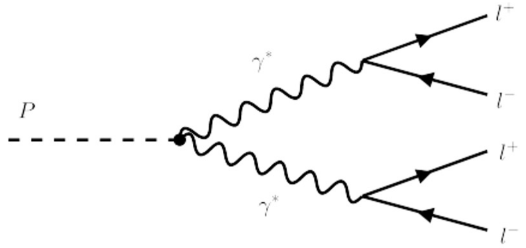
motivation

- η→4μ decay predicted with a very low branching fraction ($3.9x10^{-9}$)
    - never observed so far: precision test of the Standard Model (SM)
    - sensible to new physics scenarios

doi.org/10.1016/j.physrep.2021.11.001

result

- first observation of the rare η→4μ decay



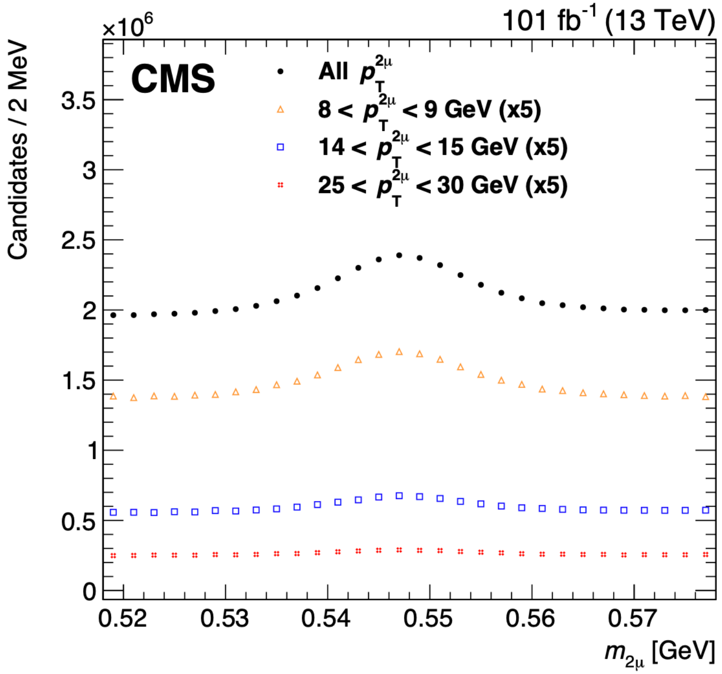
data scouting

- trigger thresholds limited by the computing power and bandwidth of the experiment
- reduce event size and fasten data acquisition
    - limit the amount of information to muon tracks
    - save HLT reconstruction and skip *prompt* event processing
    - event size reduced to ~kB (from ~MB)

→ can use looser muon thresholds → allow for low transverse momentum (pT) rare decays searches

η→4μ
event selection

- pp collisions @ 13 TeV 101 /fb collected in 2017 and 2018
- CMS trigger system
  - L1 trigger: di-muon patterns select low-pT collimated muons (pT>~4 GeV)
  - HLT trigger: di-muon pattern with mild pT selection (pT>3 GeV)
  - di-muon triggers select both 4μ (signal) and 2μ (control channel) η decays
- trigger scouting for low pT analysis
  - higher trigger rate possible (2 kHz vs. 30 Hz of standard di-muon triggers)
  - size reduction: 4 (8) kB per event in 2017 (2018)
  - 4.5 M of η→2μ events recorded → several billions η mesons produced in the CMS acceptance
- further signal skimming: charge-zero 4μ events with common vertex

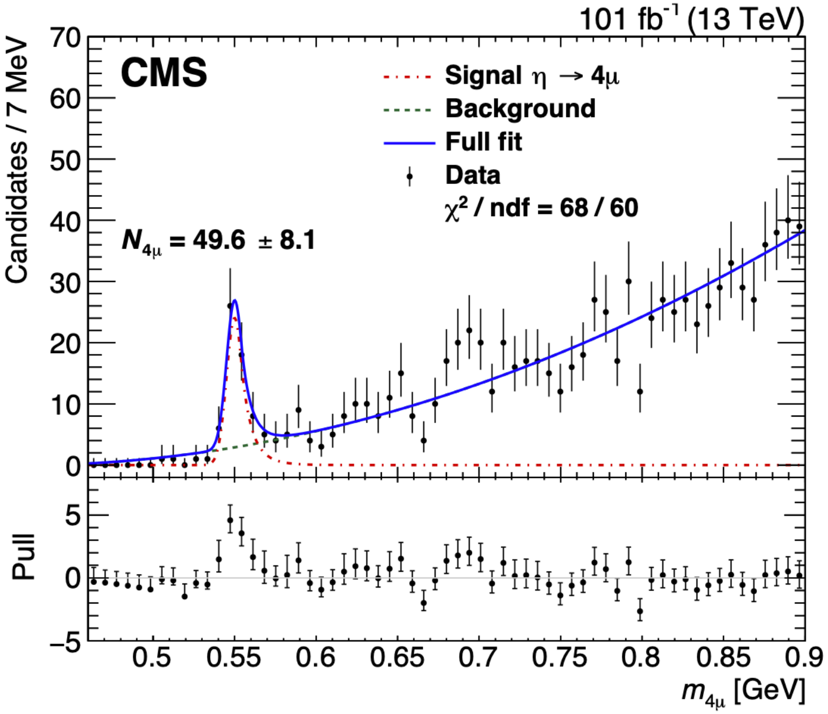


invariant mass of di-muon events in the eta range, collected by 2017 and 2018 CMS parking triggers

η→4μ

results

- η→4μ yield is normalized to the η→2μ yield
  - relatively precise normalization strategy (13.8% uncertainty)
- efficiency and acceptance corrections from MC samples
  - MC correction for 2μ-4μ differences
- η→4μ yield fit with CB function + polynomial
  - ~50 η→4μ events observed: 5 sigma excess from background (estimated with LLR)
  - resonant backgrounds faking 4μ in the signal region excluded by MC studies (see backup)



$$\mathcal{B}(\eta \to 4\mu) = 5.0 \pm 0.8(stat) \pm 0.7(syst) \pm 0.7(\mathcal{B}) \times 10^{-9}$$

- in agreement with SM prediction 3.98 ± 0.15 x 10^{-9}

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

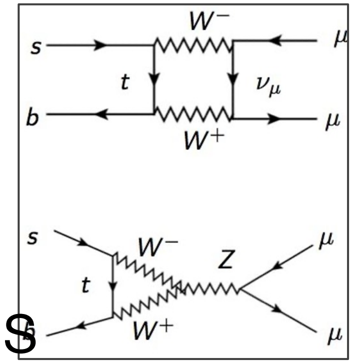
cds.cern.ch/record/2815334

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

the physics case

motivations

- $B^0_{(s)} \to \mu^+\mu^-$ strongly suppressed in the SM (FCNC and helicity)
- connected to $b \to s l^+ l^-$ transitions via the EFT operators can help understand $b \to s$ anomalies doi.org/10.1140/epjc/s10052-021-09725-1
- probe SM though lifetime measurements
- clear final state and experimental signature at CMS



result

- pp @ 13 TeV Run2 data (2016-2018) 140 /fb
  - updates the published result on 2016 data (30 /fb)
- 12.5 sigma observation of the $B^0_{(s)} \to \mu^+\mu^-$ decay, upper limit on the $\mathcal{B}(B^0 \to \mu^+\mu^-)$ and life time measurement of $B^0_{(s)} \to \mu^+\mu^-$

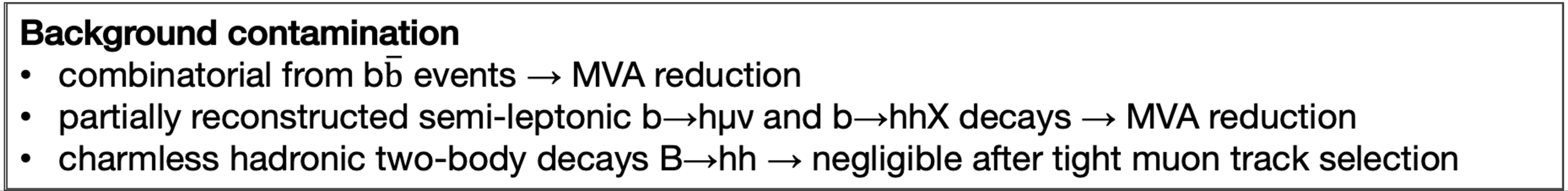
$B^0_{(s)} \rightarrow \mu^+\mu^-$
event selection

Data collection

- trigger selection: di-muon triggers with tight quality tracks and a valid secondary vertex (SV)
- similar selection for the control channels B→J/Ψ K^+ and B→ J/Ψϕ

signal selection

- two opposite-sign muons with pT > 4 GeV and |η| < 1.4
- decay vertex of B meson→ kinematic re-fit of the muon tracks with additional SV constraint
- 16 categories: 4 years x 2 BDT bins x 2 detector |η| regions



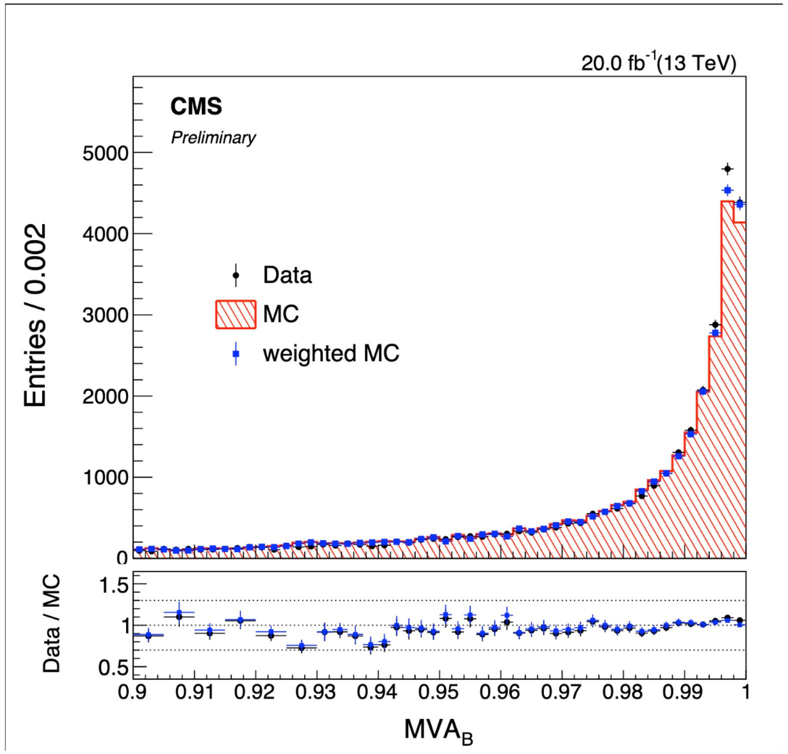
Background contamination

- combinatorial from $b\bar{b}$ events → MVA reduction
- partially reconstructed semi-leptonic b→hμν and b→hhX decays → MVA reduction
- charmless hadronic two-body decays B→hh → negligible after tight muon track selection

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

MVA analysis

- exploit several weak discrimination variables with a BDT (XGBoost)
  - features: pointing angles (2D and 3D) → effective vs. all non-two-body backgrounds
  - features: SV (quality and displacement) → effective vs. combinatorial
  - features: isolation (sum of pT surrounding the signal) → effective vs. semi-leptonic decays
- trained on data from the signal mass sidebands and MC signal samples
  - validate on $B^+ \to J/\Psi\ K^+$ events



MVA score distribution for data (black dots), MC (bars) and re-weighted MC (blue dots) for 2016a $B^+ \to J/\Psi K^+$ events

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

signal extraction

- 2D UML fit to the μμ mass x mass-resolution to extract the B→μμ signal yields. Two strategies for B_s^0 normalization:
  - $B^+\to J/\Psi(\to\mu^+\mu^-)\,K^+$ normalization → rely on the knowledge of fs / fu
  - $B_s^0\to J/\Psi(\to\mu^+\mu^-)\,\phi(\to K^+K^-)$ normalization → higher systematic (additional kaon)

$$\mathscr{B}(B_s^0 \to \mu\mu) = \mathscr{B}(B^+ \to J/\Psi K^+) \cdot \frac{N_{B_s^0\to\mu\mu}}{N_{B^+\to J/\Psi K^+}} \cdot \frac{\epsilon_{B^+\to J/\Psi K^+}}{\epsilon_{B_s^0\to\mu\mu}} \cdot \boxed{\frac{f_u}{f_s}}$$

derived from LHCb measurement: 0.231 ± 0.008

doi.org/10.1103/PhysRevD.104.032005

$$\mathscr{B}(B_s^0 \to \mu\mu) = \mathscr{B}(B_s^0 \to J/\Psi\Phi) \cdot \frac{N_{B_s^0\to\mu\mu}}{N_{B_s^0\to J/\Psi\Phi}} \cdot \frac{\epsilon_{B_s^0\to J/\Psi\Phi}}{\epsilon_{B_s^0\to\mu\mu}}$$

$$\mathscr{B}(B^0 \to \mu\mu) = \mathscr{B}(B^+ \to J/\Psi K^+) \cdot \frac{N_{B^0\to\mu\mu}}{N_{B^+\to J/\Psi K^+}} \cdot \frac{\epsilon_{B^+\to J/\Psi K^+}}{\epsilon_{B_s^0\to\mu\mu}} \cdot \frac{f_u}{f_d}$$

- UML fit to the decay time to extract τ (3D fit: decay time, its uncertainty and μμ mass)

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

$\mathcal{B}(\mathbf{B^0_s}\rightarrow\mu^+\mu^-) = 3.83^{+0.38}_{-0.36}(stat)^{+0.14}_{-0.13}(syst)\ ^{+0.14}_{-0.13}(fs/fu)$

$\times 10^{-9}$ (from J/Ψ K$^+$)

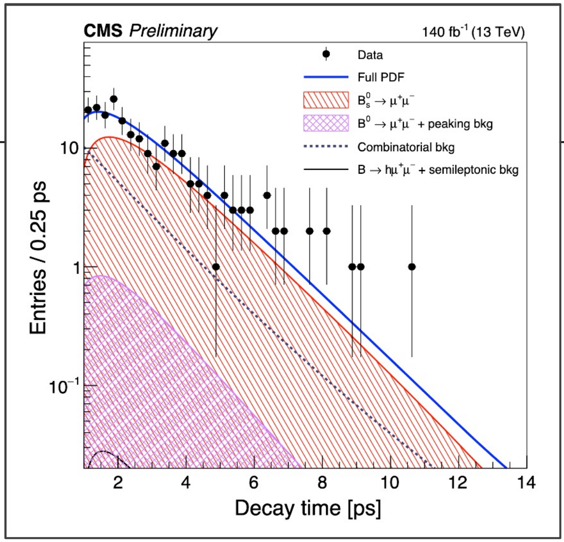
$\mathcal{B}(\mathbf{B^0_s}\rightarrow\mu^+\mu^-) = 3.95^{+0.39}_{-0.37}(stat)^{+0.27}_{-0.22}(syst)\ ^{+0.21}_{-0.19}(BF)$

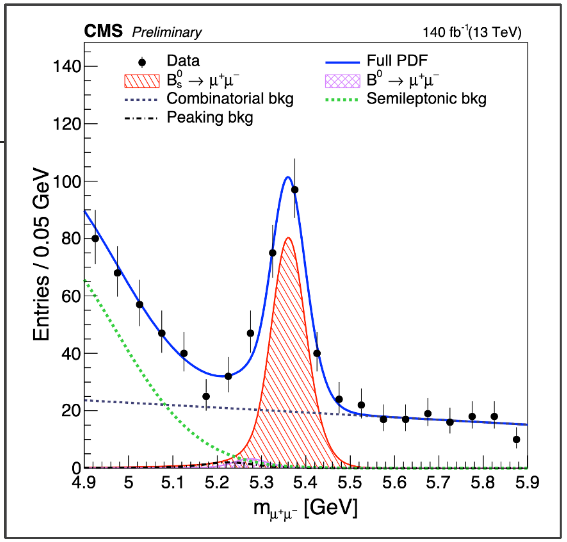
$\times 10^{-9}$ (from J/Ψϕ)

$\mathcal{B}(B^0\rightarrow\mu^+\mu^-) < 1.5 \times 10^{-10}$ @ 90% CL

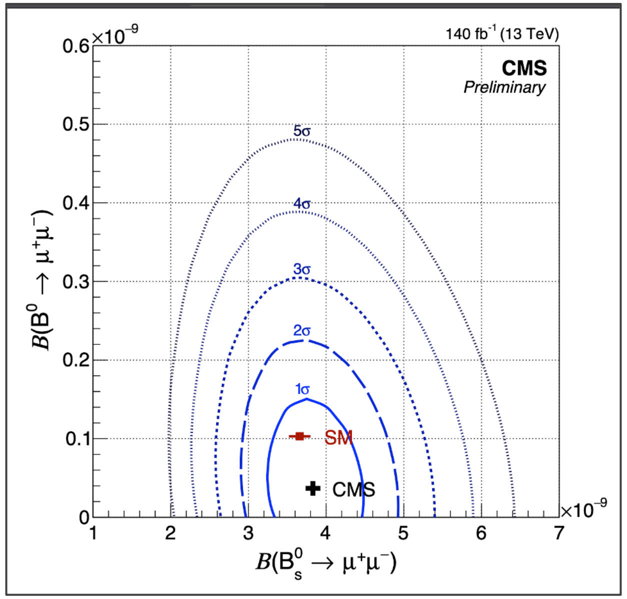
$\mathcal{B}(B^0\rightarrow\mu^+\mu^-) < 1.9 \times 10^{-10}$ @ 95% CL

$\tau(\mathbf{B^0_s}) = 1.83^{+0.23}_{-0.20}(stat)^{+0.04}_{-0.04}\ (syst)$ ps





- All UML fit results are compatible with the SM prediction within 1 sigma
- <u>most precise measurement of $B^0_s\rightarrow\mu^+\mu^-$ branching fraction and lifetime to date</u>

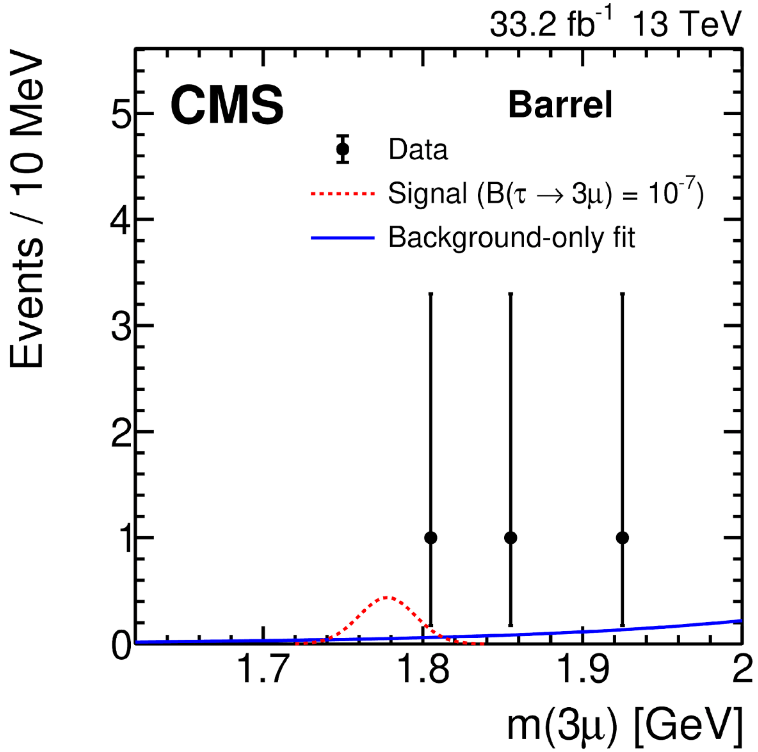


$\tau^+ \longrightarrow \mu^+ \mu^+ \mu^-$

doi.org/10.1007/JHEP01%282021%29163

$\tau^+\rightarrow\mu^+\mu^+\mu^-$

- $\tau\rightarrow3\mu$ excellent candidate for new physics searches
  - LFV process, strongly suppressed in the SM (~10^{-55}), but predicted at the level of $10^{-8} - 10^{-10}$ by some BSM models Bordone et al. 10.1007/JHEP10(2018)148
  - clear final state signature
  - fairly abundant in pp collisions (per /fb)
- CMS targets τ leptons produced via D/B mesons and via W bosons
- analysis on 2016 pp data @ 13 TeV (30 /fb)
- select three-muon events and reduce the background contamination via BDT
- observed (expected) UL from three-muon invariant mass distribution
  - $B^{HF}(\tau\rightarrow3\mu) < 9.2\ (10.0) \times 10^{-8}$ @ 90% CL
  - $B^{W}(\tau\rightarrow3\mu) < 20.0\ (13.0) \times 10^{-8}$ @ 90% CL
  - **$B(\tau\rightarrow3\mu) < 8.0\ (6.9) \times 10^{-8}$ @ 90% CL**



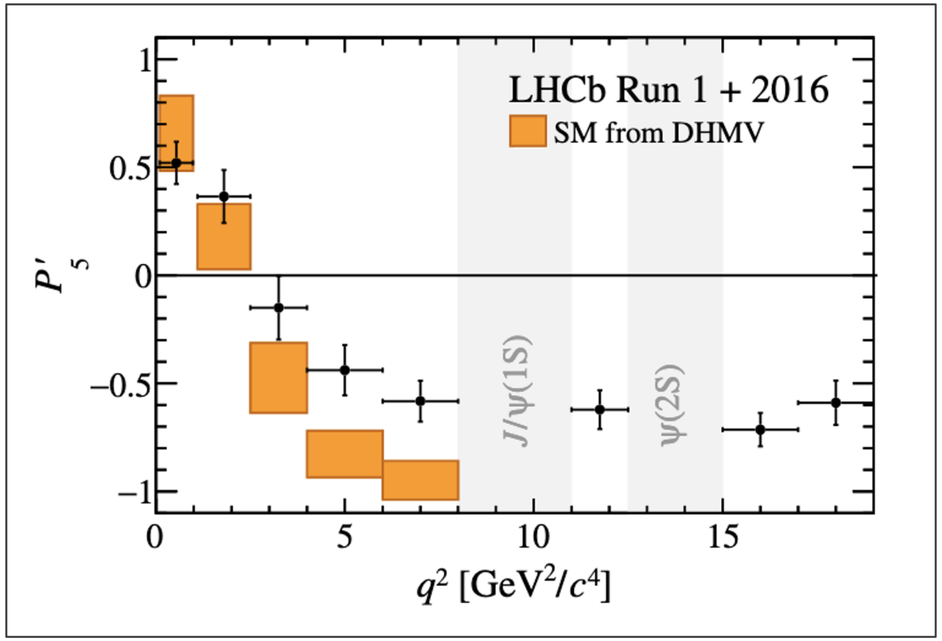
$B^0 \rightarrow K^{0*} \mu^+ \mu^-$

doi.org/10.1016/j.physletb.2018.04.030

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

the physics case

- FCNC strongly suppressed in the SM
- angular analysis
- study the deviation from the SM of the angular parameters describing the decay
    - ○ P5' most interesting following 2–3 sigmas deviation from SM observed by LHCb in the $4<q^2<6$ and $6<q^2<8$ GeV^2 bins in Run1



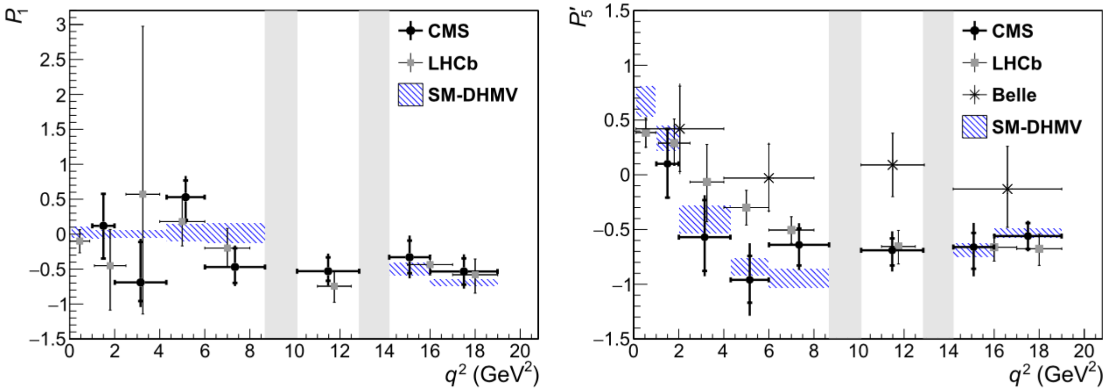
result

- pp @ 8 TeV Run1 data 20.5 /fb
- P5' distributions compatible with the SM

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

analysis results

- q^2 binned
- trigger di-muon events with displaced vertex
- build signal candidates by requiring two muons and two tracks (K* candidate) from the same vertex
  - resolve PID ambiguity by selecting the Kπ hypothesis closest to the K* mass (misid: 12-14%)
- background mainly from combinatorial
  - negligible contamination from peaking B and Λ_b decays and b→cX decays
- UML fit to the B0 mass x angular distributions to extract the POIs (P_1 and P_5')
  - B→JPsi/Ψ' K* for fit validation and systematic assesment



- both the angular parameters are compatible with the SM

Summary

Summary of the talk

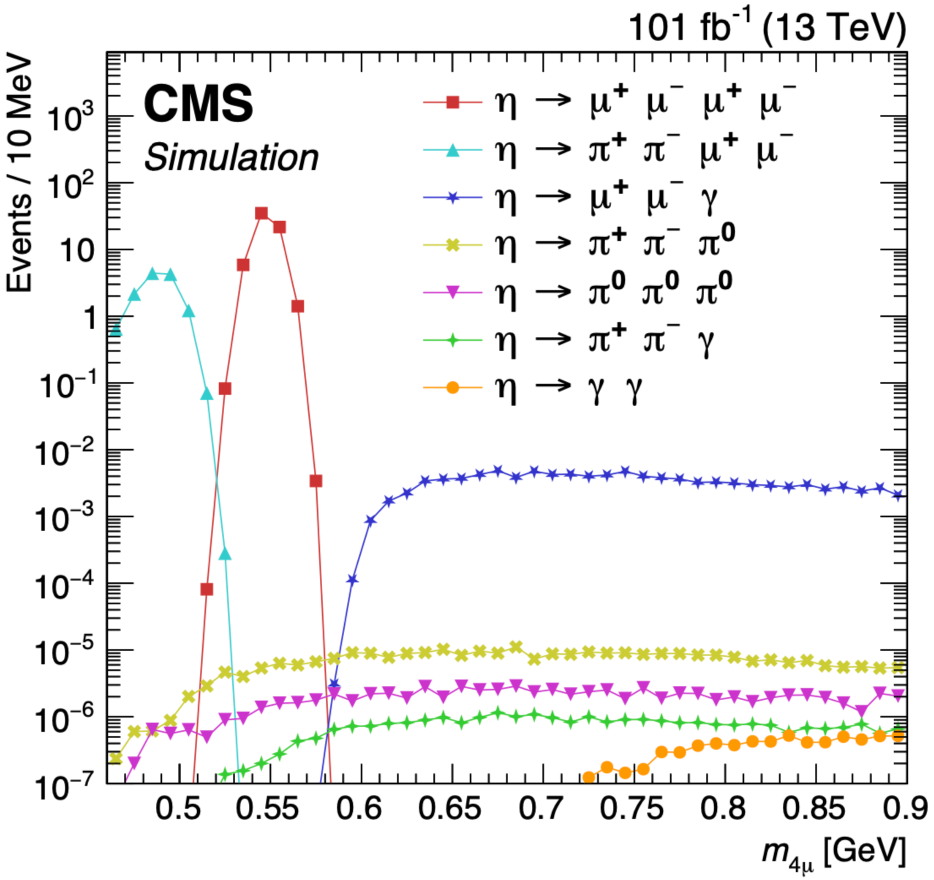
- **First η→4μ observation in CMS Run2 scouting data @ 13 TeV (101 /fb)**
  - B(η→4μ) = 5.0 ± 0.8 (stat) ± 0.7 (syst) ± 0.7 (B) · 10^{-9}
- **$B^0_{(s)}\rightarrow\mu^+\mu^-$ at CMS on pp collisions @ 13 TeV (140 /fb)**
  - $B(B^0_s\rightarrow K^{0*}\mu^+\mu^-) = 3.83^{+0.38}_{-0.36}$ (stat) $^{+0.14}_{-0.13}$ (syst) $^{+0.14}_{-0.13}$ (fs/fu) · 10^{-9} (*)
  - $B(B^0\rightarrow\mu^+\mu^-) < 1.5\ (1.9) \cdot 10^{-10}$ @ 90% (95%) CL
  - $\tau(B^0_s) = 1.83^{+0.23}_{-0.20}$(stat)$^{+0.04}_{-0.04}$ ps (*)
- **τ→3μ (W and D/B channels) at CMS in pp collisions @ 13 TeV (30 /fb)**
  - B(τ→3μ) < 8.0 x 10^{-8} @ 90% CL
- **$B^0\rightarrow K^{0*}\mu^+\mu^-$ at CMS in pp collisions @ 8 TeV (20.5 /fb)**
  - angular analysis: all angular parameters compatible with the SM, no deviation observed in the P5’ parameter

(*) most precise up to date

Backup

$\eta \rightarrow 4\mu$

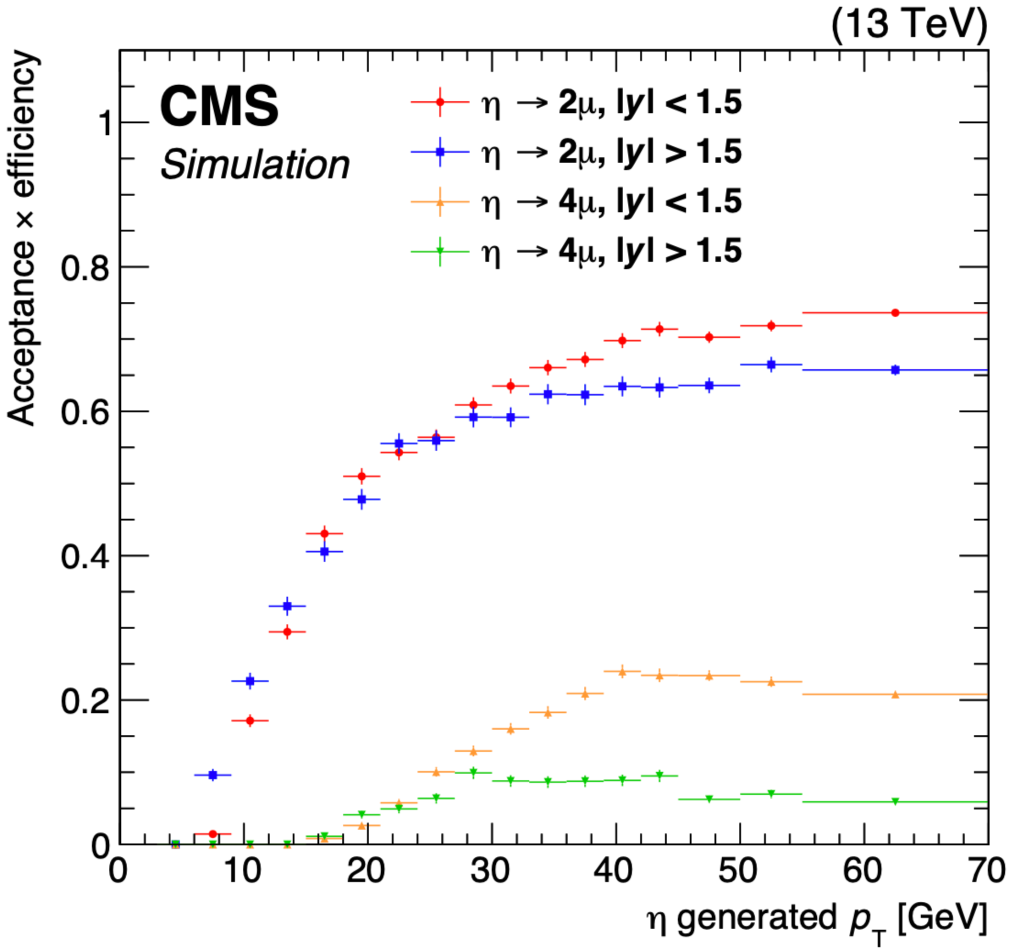
resonsnat background contamination



- no peaking decay under the η peak
- note: unobserved decays are normalized to their upper limit

η→4μ

acceptance correction



- 4μ and 2μ efficiencies in bins of pT and rapidity

η→4μ
systematic uncertainties

- **track pT threshold uncertainty [9%]:** imperfect modeling of turn-on behaviour of single-muon reconstruction efficiency in simulated data
- **trigger pT threshold uncertainty [8.4%]**: imperfect modeling of turn-on behaviour of single-muon reconstruction efficiency at HLT in simulated data
- **plateau efficiency uncertainty [3.2%]:** mismodeling of trigger efficiency plateau
- **fit bias:** subdominant
- **η→2μ branching fraction [13.8%]**

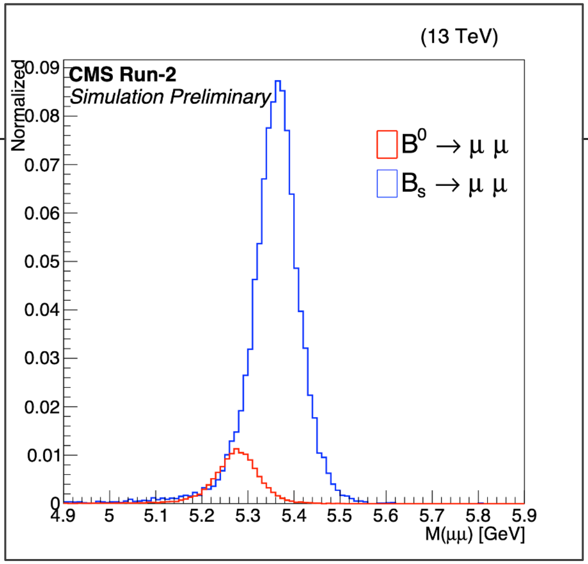
cds.cern.ch/record/2650545

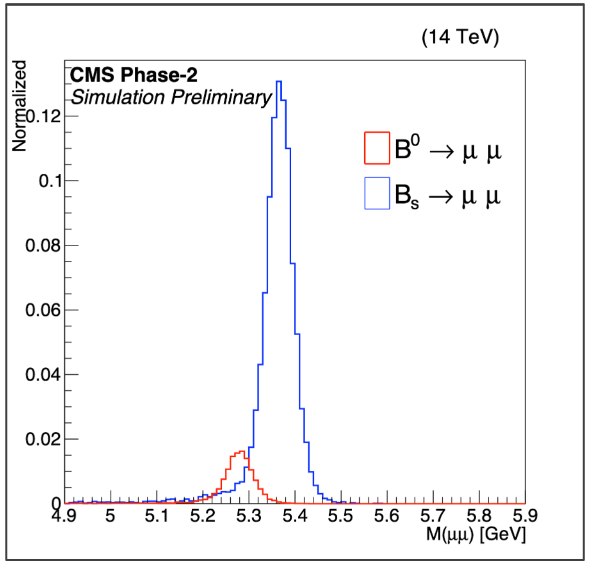
$B^0_{(s)} \rightarrow \mu^+\mu^-$ perspectives at the HL-LHC

- CMS prediction for HL-LHC (Phase 2) starting in 2029
  - 14 TeV pp collision → ~ same b production
  - x5 collision rate (200 PU) → no large impact from 200PU is expected
  - 3 /ab of luminosity → x20 Run-2
- extrapolation via MC simulation (full Phase2 detector) + toys from Run-1 results
  - reasonable projection of most of the systematic uncertainties (x0.5)
- much better mass resolution following tracker upgrade
  - less contamination from semi-leptonic fakes
  - better B^0_s - B^0 hypothesis separation

➢ Time resolution on lifetime: 0.05 ps

➢ observation of $B^0 \rightarrow \mu\mu$ at more than 5 sigmas





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

SYSTEMATIC UNCERTAINTIES

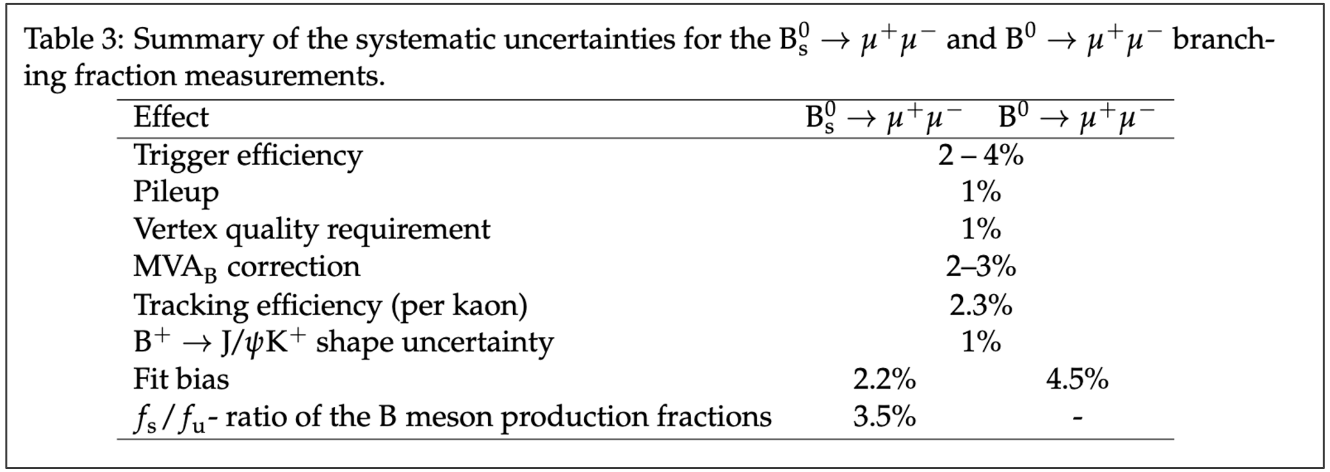


Table 3: Summary of the systematic uncertainties for the $B^0_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ branching fraction measurements.

Effect	$B^0_s \to \mu^+\mu^-$	$B^0 \to \mu^+\mu^-$
Trigger efficiency	2 – 4%	
Pileup	1%	
Vertex quality requirement	1%	
MVA_B correction	2–3%	
Tracking efficiency (per kaon)	2.3%	
$B^+ \to J/\psi K^+$ shape uncertainty	1%	
Fit bias	2.2%	4.5%
f_s/f_u- ratio of the B meson production fractions	3.5%	-

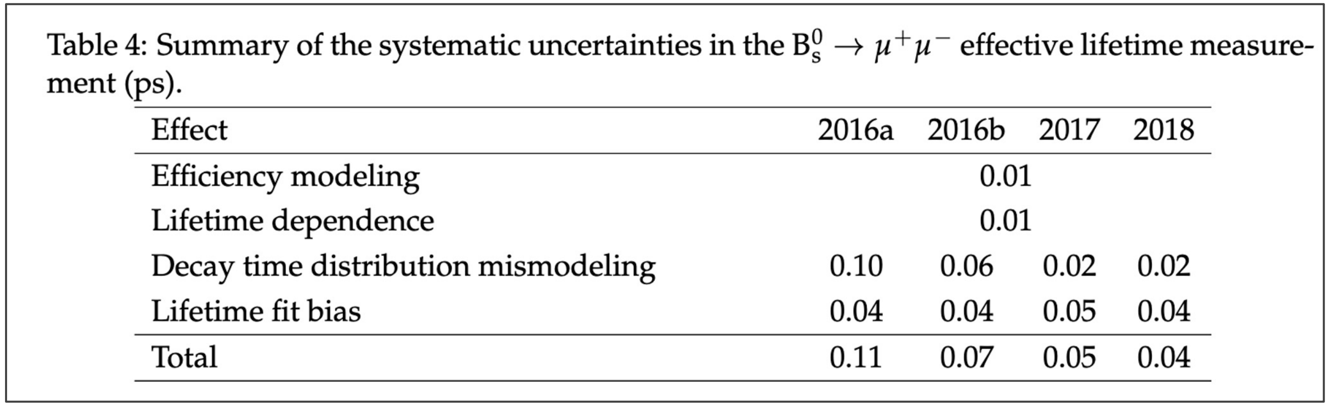
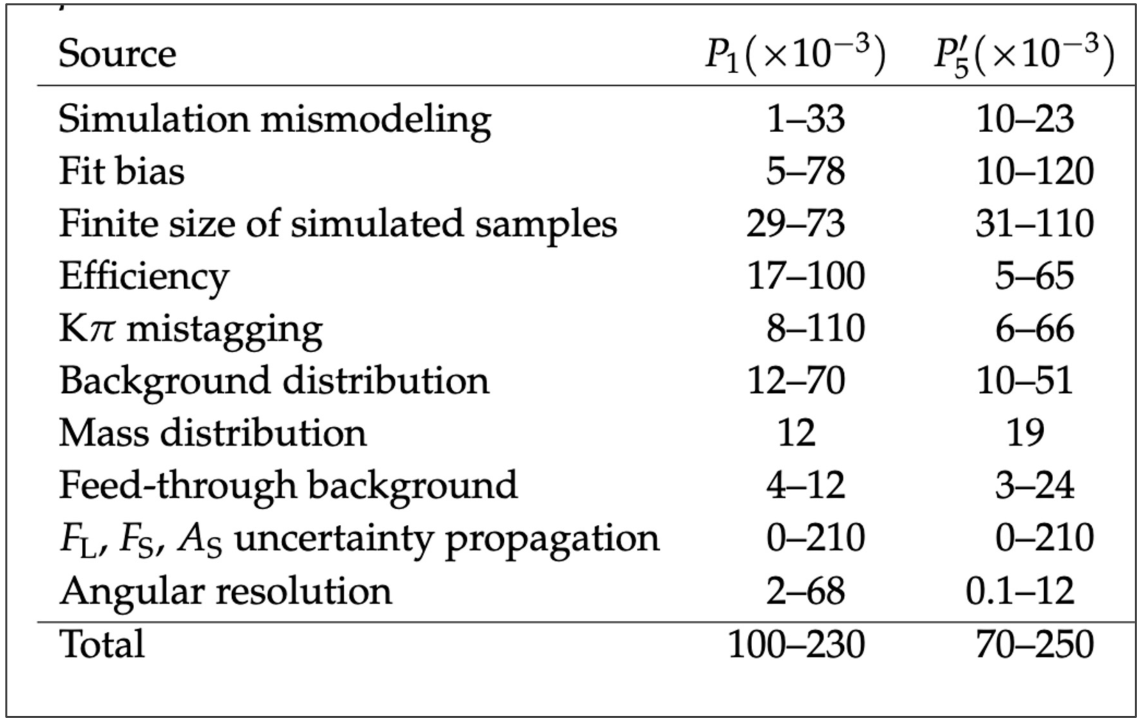


Table 4: Summary of the systematic uncertainties in the $B^0_s \to \mu^+\mu^-$ effective lifetime measurement (ps).

Effect	2016a	2016b	2017	2018
Efficiency modeling	0.01			
Lifetime dependence	0.01			
Decay time distribution mismodeling	0.10	0.06	0.02	0.02
Lifetime fit bias	0.04	0.04	0.05	0.04
Total	0.11	0.07	0.05	0.04

- **trigger:** data-MC comparison of control channels
- **pileup:** by means of reweighing
- **vertex:** the control channel triggers require a tighter selection. Evaluated the difference of the two selections.
- **MVA:** difference between data and MC efficiencies evaluated after an MVA reweight of the control channel
- **tracking:** comparing $D^0\to K\pi$ and $D^0\to K\pi\pi\pi$ ratio with world average
- **B→J/ΨK shape:** evaluating different shapes
- **fit bias:** with pseudo-experiments
- **fs/fu:** from external measurement
- **lifetime fit bias:** correlation of the BDT to the life-time. Measured by comparing the B→J/ΨK fit to the SM prediction after the BDT cut
- **decay time distribution mismodeling:** the lifetime distribution of simulated signal events is corrected using scale factors from B→J/ΨK events taken after BDT>.9 over BDT>.99. The fit difference introduced by data- or MC-derived corrections is taken as uncertainty.
- **efficiency modelling:** evaluated using different efficiency functions
- **lifetime fit bias:** measured with pseudo-experiments with different lifetimes

$B^0 \rightarrow K^{0*} \mu^+ \mu^-$
systematic uncertainties



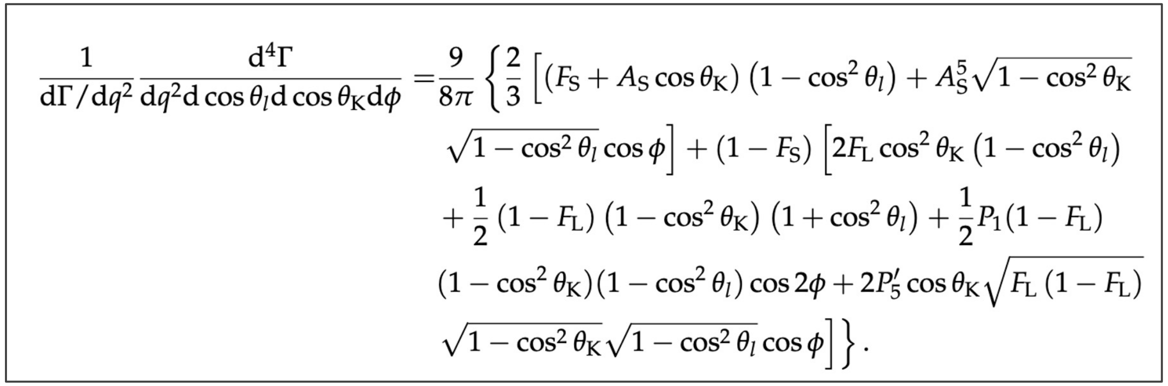
Source	$P_1 (\times 10^{-3})$	$P'_5 (\times 10^{-3})$
Simulation mismodeling	1–33	10–23
Fit bias	5–78	10–120
Finite size of simulated samples	29–73	31–110
Efficiency	17–100	5–65
$K\pi$ mistagging	8–110	6–66
Background distribution	12–70	10–51
Mass distribution	12	19
Feed-through background	4–12	3–24
F_L, F_S, A_S uncertainty propagation	0–210	0–210
Angular resolution	2–68	0.1–12
Total	100–230	70–250

- **simulation mismodeling:** fit a simulated signal sample with 400x data and see the fit difference wrt to the input
- **fit bias:** use 200 simulated signal samples + simulated data (~data size) to estimate the fit bias
- **finite size of simulated samples:** due to the finite size of the MC used to derive the efficiency shape. Generate 100 numerator and denominator shapes from the original ones and refit them to estimate the difference due to the statistical uncertainty.
- **efficiency:** fit the control channels to extrapolate fixed parameters (FL) and cross check with PDG
- **Kπ mistagging:** fit the mistag rate on the control channel B→J/ΨK* and take the diffference wrt the simulation as systematics
- **background shape:** fit the data 200 times varying the shape of the background distributions within their error (fixed in these test fits) and evaluating the POIs distribution RMS
- **signal mass shape:** fit the control channels letting their width vary alternately and see the effect on the POIs
- **background feed-through:** see the difference in the POIs after counting for an additional systematic uncertainty describing the feed-through from B→J/Ψ / Ψ' K* events
- **F_L, F_S, A_S uncertainties propagation:** fit pseudo-experiments allowing these parameters to change and compare the POIs values with the nominal procedure
- **angular resolution:** fit a siulated sample using generated or reconstructed values of the angular parameters and see the difference

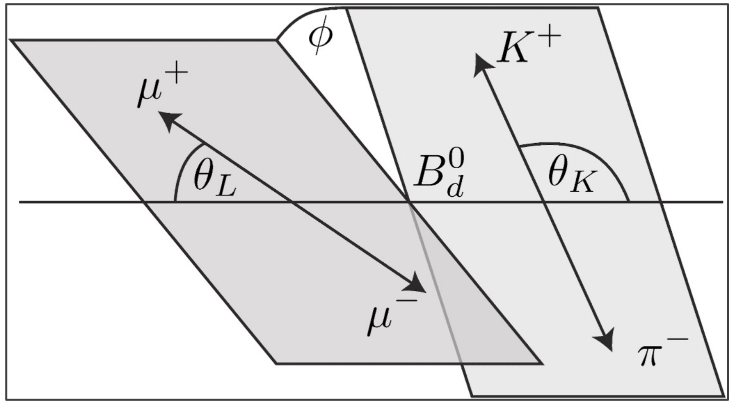
$B^0 \rightarrow K^{0*} \mu^+ \mu^-$

fit strategy

- simplify the angular pdf by considering symmetries in the ϕ=0 and θl=π/2 angles
    - reduce the POIs that can be extracted to P_1 and P_5'
- UML fit to the mass and angular distributions
    - signal mass shape: double-gaussian different for correct- and wrong-tagged events, with paramters obtained from simulated samples
    - angular pdf for background: polynomials (factorizing)
    - mass pdf for background: exponential
    - angular efficiencies: obtained from kernel-density estimators, different for correct- and wrong-tagged events
- Fit run in two steps:
    1. fit the sidebands and fix the background shapes
    2. fit the full distribution and obtain the POIs



$$\frac{1}{\mathrm{d}\Gamma/\mathrm{d}q^2}\frac{\mathrm{d}^4\Gamma}{\mathrm{d}q^2\mathrm{d}\cos\theta_l\mathrm{d}\cos\theta_\mathrm{K}\mathrm{d}\phi}=\frac{9}{8\pi}\left\{\frac{2}{3}\left[(F_\mathrm{S}+A_\mathrm{S}\cos\theta_\mathrm{K})\left(1-\cos^2\theta_l\right)+A_\mathrm{S}^5\sqrt{1-\cos^2\theta_\mathrm{K}}\sqrt{1-\cos^2\theta_l}\cos\phi\right]+(1-F_\mathrm{S})\left[2F_\mathrm{L}\cos^2\theta_\mathrm{K}\left(1-\cos^2\theta_l\right)+\frac{1}{2}(1-F_\mathrm{L})\left(1-\cos^2\theta_\mathrm{K}\right)\left(1+\cos^2\theta_l\right)+\frac{1}{2}P_1(1-F_\mathrm{L})(1-\cos^2\theta_\mathrm{K})(1-\cos^2\theta_l)\cos 2\phi+2P_5'\cos\theta_\mathrm{K}\sqrt{F_\mathrm{L}(1-F_\mathrm{L})}\sqrt{1-\cos^2\theta_\mathrm{K}}\sqrt{1-\cos^2\theta_l}\cos\phi\right]\right\}.$$



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

results

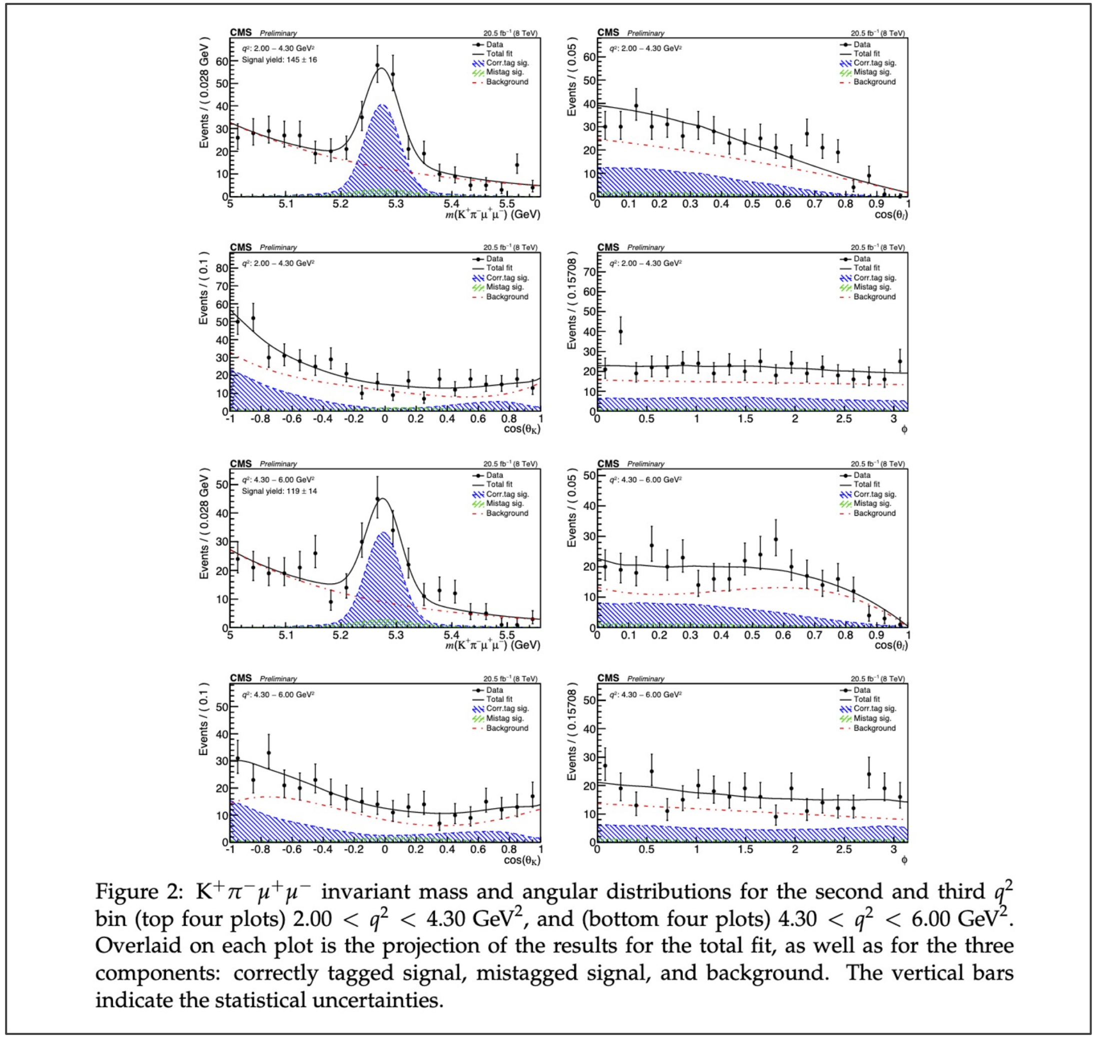
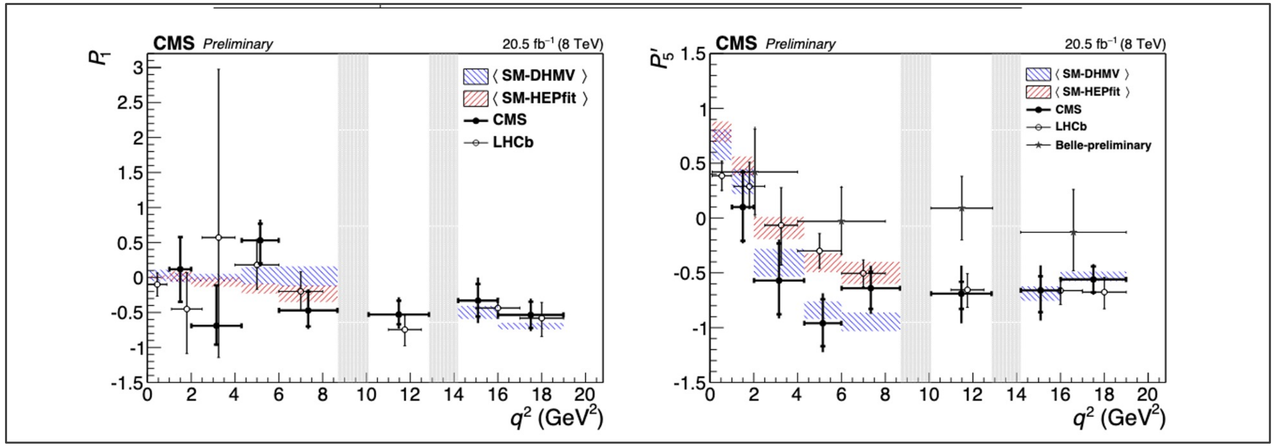


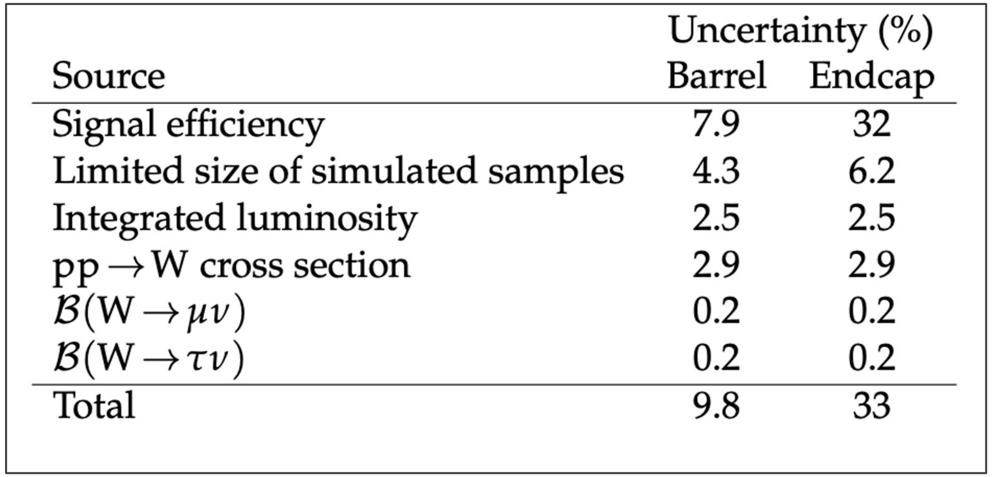
Figure 2: $K^+\pi^-\mu^+\mu^-$ invariant mass and angular distributions for the second and third q^2 bin (top four plots) $2.00 < q^2 < 4.30$ GeV^2, and (bottom four plots) $4.30 < q^2 < 6.00$ GeV^2. Overlaid on each plot is the projection of the results for the total fit, as well as for the three components: correctly tagged signal, mistagged signal, and background. The vertical bars indicate the statistical uncertainties.



$\tau^+\to\mu^+\mu^+\mu^-$

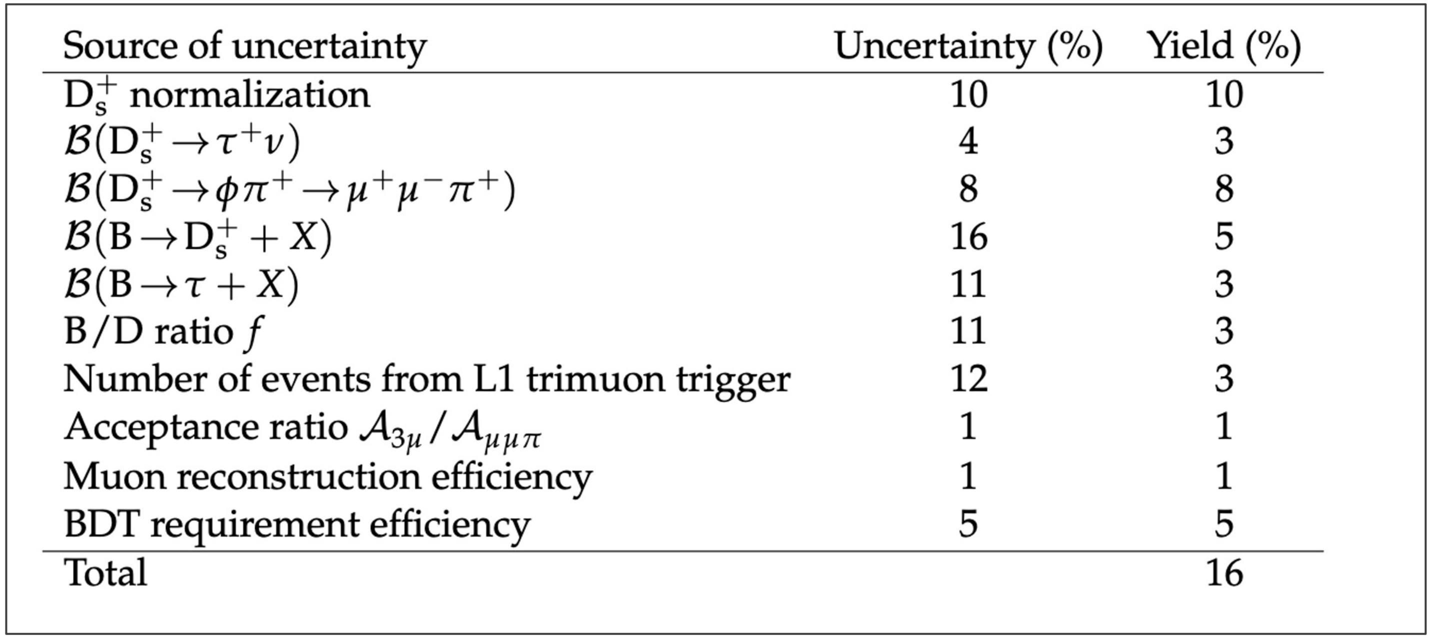
systematic uncertainties

W channel



Source	Uncertainty (%) Barrel	 Endcap
Signal efficiency	7.9	32
Limited size of simulated samples	4.3	6.2
Integrated luminosity	2.5	2.5
$pp \to W$ 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

HF channel

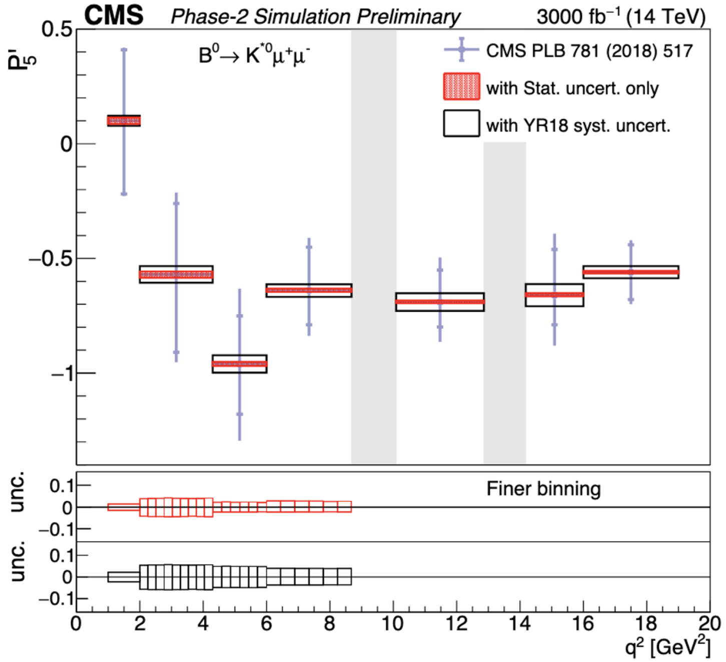


Source of uncertainty	Uncertainty (%)	Yield (%)
D_s^+ normalization	10	10
$\mathcal{B}(D_s^+ \to \tau^+\nu)$	4	3
$\mathcal{B}(D_s^+ \to \phi\pi^+ \to \mu^+\mu^-\pi^+)$	8	8
$\mathcal{B}(B \to D_s^+ + X)$	16	5
$\mathcal{B}(B \to \tau + X)$	11	3
B/D ratio f	11	3
Number of events from L1 trimuon trigger	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

$B^0 \to K^{0*}\mu^+\mu^-$ at the HL-LHC

- MC study to address the precision reached by ATLAS at the HL-LHC (3 /ab)
- analysis strategy same as Run1
  - reduce fit-related systematics by 1/sqrt(L)
  - reduce most of uncertainties (signal shape, efficiency shape, mis-tag, detector-related) by a factor 2
- precision increase up to x15 better → also explore possibilities of finer binning

cds.cern.ch/record/2651298



$\tau \to 3\mu$ at the HL-LHC

- luminosity-scaled projections based on the HF results place CMS sensitivity at 3.7×10^{-9} @ 90% CL

  arXiv:1812.07638