CMS Draft Analysis Note

The content of this note is intended for CMS internal use and distribution only

2022/09/12
Archive Hash: bcaab30-D
Archive Date: 2022/09/11

Search for a high mass dimuon resonance associated with b quark jets at 13 TeV

Claudio Campagnari[1], Philip Chang[2], Keegan Downham[1], Yanxi Gu[2], Slava Krutelyov[2], Mario Masciovecchio[2], Ulaşcan Sarıca[1], Manos Vourliotis[2], Avi Yagil[2], Frank Golf[3], and Giacomo Zecchinelli[3]

[1] Univ. of California, Santa Barbara
[2] Univ. of California, San Diego
[3] Univ. of Nebraska-Lincoln

Abstract

A search for high-mass dimuon resonance production in association with one or more b quark jets is presented, using data collected with the CMS experiment at the LHC that correspond to an integrated luminosity of 138 fb^{-1} at a center-of-mass energy of 13 TeV. Model-independent constraints are derived on the numbers of events with one and more than one b quark jet. Results are also interpreted in terms of models that involve possible $\mathrm{Z'sb}$, $\mathrm{Z'bb}$, and $\mathrm{Z'}\mu\mu$ couplings.

This box is only visible in draft mode. Please make sure the values below make sense.

PDFAuthor: The Z′ + b-tags Analysis Team
PDFTitle: Search for a high mass dimuon resonance associated with b quark jets at 13 TeV
PDFSubject: CMS
PDFKeywords: dimuon, b-tag, anomaly, anomalies, Z′

Please also verify that the abstract does not use any user defined symbols

Contents

1 Main changes in AN versions

- Changes and additions from V2 to V3:
    - Added information on the primary datasets used and the corresponding luminosities (Sec. 4.1).
    - Added references to the trigger efficiencies and the relevant scale factors (Sec. 4.3).
    - Extended signal MC description (Sec. 4.2.1): slightly expanded description of the "Allanach et al." models (Sec. 4.2.1.1) and added description of the EXO-22-006 "bottom-fermion fusion" (BFF) models (Sec. 4.2.1.2).
    - Added Appendix G with a description of the signal MC reweighting procedure.
    - Added Appendix H with a comparison of the sensitivity of this anaysis with that of EXO-22-006.
    - Added Section 7.3 with a discussion of the expected limits before unblinding (Section 7.3.1 for the model independent limits and Section 7.3.2 for the limits in the models by Allanach et al.).

2 Introduction

This is a search for $Z' \to \mu\mu$ in the presence of one or more b-tagged jets. The search is motivated by a series of b-anomalies in decays of the type $B \to K^{(*)}e^+e^-$ and $B \to K^{(*)}\mu^+\mu^-$. Briefly, current data is pointing to a 3.1σ discrepancy between $b \to se^+e^-$ and $b \to s\mu^+\mu^-$ transitions [1]. This discrepancy could be driven by the exchange of either a leptoquark (LQ) or a new vector bosons (Z') with different couplings to electrons and muons, as shown in Fig. 1.

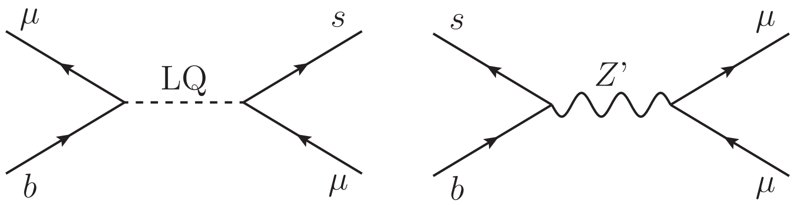


Figure 1: Possible leptoquark (left) and Z' (right) contributions to $b \to s\mu^+\mu^-$

The theoretical prejudice is that BSM contributions are in the second and third generation. This is then a motivation for a $Z' \to \mu\mu$ search. Inclusive searches for Z' have been performed several times at the LHC, see for example Refs. [2] and [3]. The power of these "standard" searches, however, are limited by the Drell–Yan background. Possible BSM $pp \to Z' \to \mu^+\mu^-$ signals motivated by the b-anomalies, on the other hand, would result in b quarks in the final state, see Fig. 2. This then motivates a dedicated $Z' \to \mu\mu$ search with an explicit requirement on the presence of b-tagged jets.

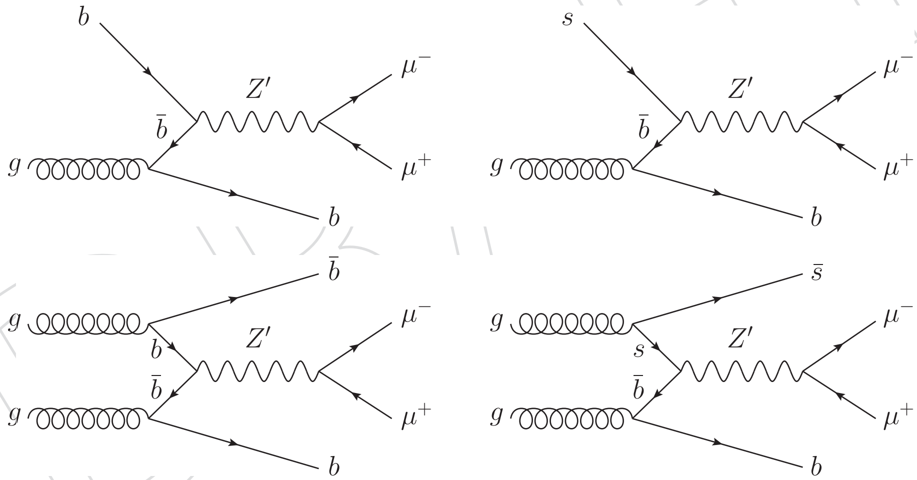


Figure 2: Feynman diagrams for $Z' \to \mu^+\mu^-$ with a Z' produced via $s\overline{b} \to Z'$ or $b\overline{b} \to Z'$ and at least on b-quark in the final state. Note that while $b\overline{b}Z'$ couplings are not needed to explain the flavor anomalies, these couplings arise naturally in more complete Z' models motivated by the anomalies.

Studies of dilepton invariant mass distributions in the presence of b-tagged jets have already been performed at the LHC [4–6]. These studies suffered from large dilepton $t\bar{t}$ backgrounds. In this analysis we gain sensitivity using requirements that substantially reduce this background.

As will be discussed in Section 4.2, there exist BSM Z' parameter space favored by the flavor anomalies, consistent with other electroweak data, and not excluded by current searches. We will show in this note that a dedicated analysis such as the one presented here can probe this parameter space with existing LHC data.

Briefly our analysis can be summarized as follows

- We search for narrow $\mu^+\mu^-$ high mass resonances in the presence of at least one b-tagged jet. Narrow because as shown in Section 4.2 phenomenological Z′ models developed to "explain" the anomalies result in Z′ widths that are small compared to the $\mu^+\mu^-$ invariant mass resolution.
- We apply requirements to reduce the $t\bar{t}$ dilepton background.
- We categorize the events by the multiplicity of b-tagged jets: $N_b = 1$ and $N_b \geq 2$.
- We extract possible signal yields by fitting the $\mu^+\mu^-$ invariant mass distribution to analytic functions (power laws or exponentials or polynomials for the continuum background and double-crystal ball for the signal). For each mass hypothesis we fit the invariant mass distribution in a relatively narrow region around the mass of interest ($\pm 10\sigma$, where σ is the mass resolution). This is the same procedure used in the search for displaced dimuon resonances [7]. The background functional form is extracted from the data. **Monte Carlo simulations of backgrounds are only used to guide the design of the analysis requirements.**
- In the absence of signal, we then:
  - Set constraints on the parameter space of specific Z′ models.
  - Provide upper limits on the number of detected Z′ events as a function of mass with $N_b = 1$, $N_b \geq 2$, and $N_b \geq 1$. Since our requirements are straightforward, this would allow reinterpetation of our limits in BSM models that we are not explicitly considering.
- Finally, we use a data sample corresponding to an integrated luminosity of 138 fb^{-1} collected with the CMS experiment during the LHC Run 2 (2016–2018) at a center-of-mass-energy of 13 TeV (2016: 36.3 fb^{-1}; 2017: 41.5 fb^{-1}; 2018: 59.8 fb^{-1}). The analysis is performed using the UltraLegacy reconstruction in the NanoAOD (v9) format.

3 The CMS detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter, and a brass and scintillator hadron calorimeter, each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity (η) coverage provided by the barrel and endcap detectors. Muons are measured in gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid. A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [8] The pixel tracker was upgraded before the start of the data taking period in 2017, providing one additional layer of measurements compared to the older tracker [9].

A two-level trigger system is used to select events of potential physics interest. The first level of the CMS trigger system (L1T), composed of custom hardware processors, uses information from the calorimeters and muon detectors to select the most interesting events in a fixed time interval of less than 4 μs. The high-level trigger (HLT) processor farm further decreases the event rate from around 100 kHz to about 1 kHz, before data storage. A more detailed description of the CMS trigger system can be found in Ref. [10].

4 Data sets and triggers

4.1 Data

The analysis is performed using the NanoAOD v9 data format of the ‘SingleMuon’ primary data sets. Table 1 outlines the data sets for each data period, and also the reconstruction conditions used in analyzing the simulation. The golden JSON files used in filtering the data are listed in Table 2. The corresponding luminosities for each data set is summmarized in Table 3. More details can be found in Refs. [11–14] for the data sets, and Refs. [15–18] for data certification.

Table 1: The data sets and the global tags used to produce them. These data sets correspond to the latest NanoAOD v9 data format, and the global tags are for the Run 2 Ultra Legacy reconstruction [11].

2016 pre-VFP data sets
Run2016B-ver{1,2}_HIPM_UL2016_MiniAODv2_NanoAODv9-v2: 106X_dataRun2_v35
Run2016[C-F]-HIPM_UL2016_MiniAODv2_NanoAODv9-v2: 106X_dataRun2_v35
Simulation: 106X_mcRun2_asymptotic_preVFP_v11
2016 post-VFP data sets
Run2016[F-H]_UL2016_MiniAODv2_NanoAODv9-v1: 106X_dataRun2_v35
Simulation: 106X_mcRun2_asymptotic_v17
2017 data sets
Run2017[B-H]-UL2017_MiniAODv2_NanoAODv9-v1: 106X_dataRun2_v35
Simulation: 106X_mc2017_realistic_v9
2018 data sets
Run2018[A-D]-UL2018_MiniAODv2_NanoAODv9_GT36-v1: 106X_dataRun2_v35
Simulation: 106X_upgrade2018_realistic_v16_L1v1

Table 2: The golden JSON files used in filtering the data runs and luminosity blocks. For each data-taking period, the normalization tags are listed in the set of second rows, and all data certification "golden" JSON file locations are listed relative to /afs/cern.ch/cms/CAF/CMSCOMM/COMM_DQM/certification/ [15–18].

2016 data sets
Collisions16/13TeV/Legacy_2016/Cert_271036-284044_13TeV_Legacy2016_Collisions16_JSON.txt
/cvmfs/cms-bril.cern.ch/cms-lumi-pog/Normtags/normtag_PHYSICS.json (norm. tag)
2017 data sets
Collisions17/13TeV/Legacy_2017/Cert_294927-306462_13TeV_UL2017_Collisions17_GoldenJSON.txt
/cvmfs/cms-bril.cern.ch/cms-lumi-pog/Normtags/normtag_PHYSICS.json (norm. tag)
2018 data sets
Collisions18/13TeV/Legacy_2018/Cert_314472-325175_13TeV_Legacy2018_Collisions18_JSON.txt
/cvmfs/cms-bril.cern.ch/cms-lumi-pog/Normtags/normtag_PHYSICS.json (norm. tag)

Table 3: The corresponding luminosities for each one of the data sets used in the analysis [19].

Data set	Luminosity (fb^{-1})
2016 pre-VFP data sets	19.52
2016 post-VFP data sets	16.81
2017 data sets	41.48
2018 data sets	59.83

4.2 Monte Carlo simulation

In this analysis, the SM background is estimated directly from data as a continuum background in the $m_{\mu\mu}$ spectrum parametrized by analytical functions (see Section 6.2). Thus, the search does not rely on Monte Carlo (MC) simulation in order to estimate the background. However, SM background MC samples were used for the analysis strategy optimization, as well as to visually compare observed data to the expected background yields. We also use MC of various BSM Z′ models to motivate the event selection and, eventually, to set model-dependent limits.

4.2.1 Signal Monte Carlo

Fits over several b-physics observables suggest that the effective BSM operator responsible for the anomalies is $\mathcal{L}_{BSM} \sim (g_X/\Lambda)^2(\bar{s}\gamma^\alpha P_L b)(\bar{\mu}\gamma_\alpha\mu) + \text{h.c.}$, with $\Lambda \approx g_X$ (30 TeV), see for example Ref. [20]. Naively this would suggest that any Z′ responsible for the anomalies would have a mass M_X of order 30 TeV, too heavy for production at the LHC. However, not only this argument does not take into account the effects of a coupling constant $g_X \lesssim 1$, but it is also natural to expect some CKM-like suppression from the non-flavor-diagonal $s\bar{b}$ coupling. As a result, Z′ models with Z′ mass M_X around a TeV or even lower can be viable.

4.2.1.1 MC models from the literature For our analysis, we choose four distinct MC models that were specifically inspired by the b-anomalies. These are the so-called "Third Family Hypercharge" variant models (Y_3, DY_3, and DY_3') [21–23] from B. Allanach and collaborators, and the "Third Family Baryon Number Minus Second Family Lepton Number" models $(B_3 - L_2)$ [24, 25].

For our purpose, these models are characterized by three parameters: M_X, g_X, and a mixing angle between the second and third quark generation θ_{23}. In Fig. 3 we show the allowed values of these parameters for the "Third Family Hypercharge" models, given the size of the b-anomalies, and other b-physics constraints, such as those from $B_s - \overline{B}_s$ mixing. The dashed lines in Fig. 3 are parametrized as:

$$\theta_{23} = \frac{1}{2}\sin^{-1}\left(\frac{a}{x^2 + bx}\right), \tag{1}$$

where

$$x := g_X(1\,\text{TeV}/M_X) \tag{2}$$

The a and b parameters, in the "good fit" regions are given in Table 4 for each model, including the $(B_3 - L_2)$ model that was not displayed in Fig. 3.

Model	a	b	$x = g_X(1\,\text{TeV}/M_X)$
Y_3	-0.01	0.12	0.08–0.2
DY_3	0.0045	0	0.1–0.2
DY_3'	-0.0045	0.067	0.04–0.13
$(B_3 - L_2)$	-0.0005	0	0.05–0.62

Table 4: Parametrization values and domain of x for the 95% CL region for each model from Ref. [26].

The Monte Carlo events at a given M_X are generated near the 'best-fit' of each scenario, with the parameters displayed in Table 5, using the UFO implementation of these models in MADGRAPH5_aMC@NLO, versions 2.9.9–2.9.11. Since we are interested in final states with b jets, we

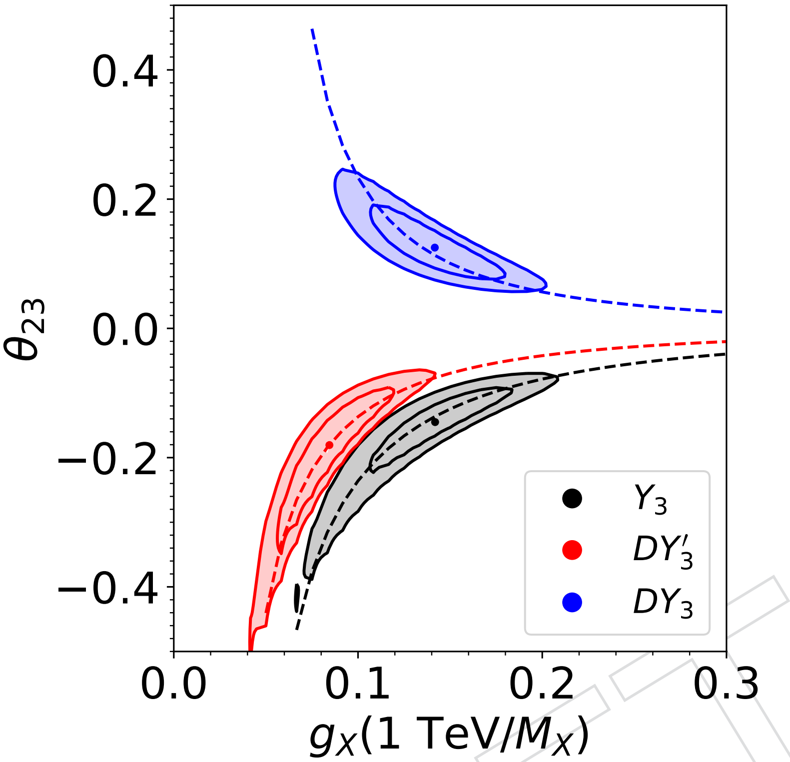


Figure 3: The colored regions show the 95% allowed regions in the $\theta_{23} - g_X(1\,\text{TeV}/M_X)$ plane for "Third Family Hypercharge models" based on an consistency with various electrweak and B-physics data (including the anomalies, of course). The inner contours show the 68% contours. The dashed lines are ad-hoc parametrizations through the center of the regions. From Ref. [26].

only generate at leading-order (LO) in QCD and BSM the following processes $gg \to Z'q\bar{q}'$, $gq \to Z'gq'$, and $q\bar{q} \to Z'g$, where q and q' are b and/or s. The possible values of the Z' total widths that correspond to the parameters in Table 5 are displayed in Fig. 4. The generated masses are 100, 200, 250, 400, 550, 700, 850, 1000, 1250, 1500, 2000.

When constraints for parameter values other than those in Table 4 are presented, the generated events are reweighted properly for the product of their couplings at the $Z'bb$, $Z'sb$, $Z'ss$, and $Z'\mu\mu$ vertices, as well as for the change in the Z' total decay width, see Appendix G.1.

Model	$x = g_X(1\,\text{TeV}/M_X)$	θ_{23}
Y_3	0.14	-0.15
DY_3	0.14	0.13
DY_3'	0.08	-0.18
$(B_3 - L_2)$	0.05	0.01

Table 5: Parameters of the signal Monte Carlo generations. Note that according to the parametrization in Table 4 the θ_{23} parameter in the B3-L2 model should be negative. However, this parameter in Figure 2 in Ref. [20] is positive. In any case, the MC predictions for $Z' \to \mu^+\mu^-$ are insensitive to the sign of θ_{23}.

4.2.1.2 MC model from EXO-22-006 (BFF) The EXO-22-006 [27] group developed a MC model (denoted as "bottom-fermion fusion", or BFF) based on the Lagrangian function $\mathcal{L}_{BSM}$:

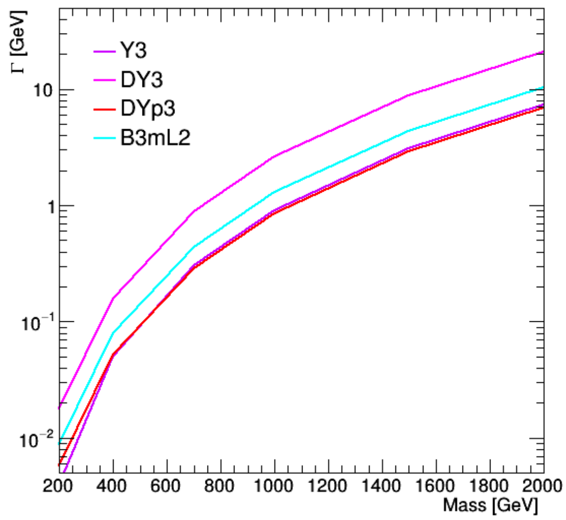


Figure 4: Widths of the Z′ for the four model of Section 4.2.1.1. The model parameters are those from Table 5.

$$\mathcal{L}_{BSM} \sim Z'_{\eta}\,[g_{\mu}\bar{\mu}\gamma^{\eta}\mu \;+\; g_{\mu}\bar{\nu}_{\mu}\gamma^{\eta}\nu_{\mu} \;+ g_b \sum \bar{q}\gamma^{\eta}P_L q \;+\; g_b\delta_{bs}\bar{s}\gamma^{\eta}P_L b] \;+\text{h.c., with } q = b \text{ or } t. \quad (3)$$

This is very similar to the models from Section 4.2.1.1, except that instead of three parameters there are now four, i.e., M_X, g_b, g_μ, and δ_{bs}. In other words, in the BFF model the $Z'\mu\mu$ coupling g_μ is a fourth free parameter, while in the models of Section 4.2.1.1 it is fixed by the structure of the model. The interesting parameter region is $g_b\delta_{bs}g_\mu(100\ \text{GeV}/M_X)^2 \approx 6.5\times10^{-6}$.

The BFF MC samples were generated under EXO-22-006 control at (mostly) LO in QCD and NLO in BSM. They include more parton-level diagrams than in the models of Section 4.2.1.1, e.g., $b\bar{s} \to Z'$ which has no final-state b quarks in the matrix element, and $b\bar{s} \to Z'b\bar{b}$ with the final $b\bar{b}$ pair from gluon splitting (this is NLO in QCD, actually). In the models of Section 4.2.1.1 the latter process is generated in the Pythia shower starting from $b\bar{s} \to Z'g$ at matrix element.

The EXO-22-006 samples were originally reconstructed in pre-Ultra Legacy (pre-UL). We separately generated UL samples using the exact EXO-22-006 MadGraph setup and the parameters specified in the EXO-22-006 analysis note (AN-18-258). In addition, since the EXO-22-006 authors concentrate their attention on masses below ≈ 400 GeV, we have also generated samples at higher mass. In particular, we have a set of MC samples at different M_X (100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 700, 750, 850, 1000, 1250, 1500, and 2000 GeV) for $g_\mu = 0.17$, $g_b = 0.02$, $\delta_{bs} = 0.04$ (these are the default parameters from AN-18-258).

Reweighting of the BFF MC samples is described in SectionG.2.

4.2.2 Background simulation

In this analysis, the SM background is estimated directly from data as a continuum background in the $m_{\mu\mu}$ spectrum parametrized by analytical functions (see Section 6.2). Thus, the search does not rely on MC simulation in order to estimate the background. However, SM background MC samples (listed Table 6) were used for the analysis strategy optimization, as well as to visually compare observed data to the expected background yields.

Sample name	Cross section (fb)
/ZToMuMu_M-50To120_TuneCP5_13TeV-powheg-pythia8/*/*	2.11×10^{6}
/ZToMuMu_M-120To200_TuneCP5_13TeV-powheg-pythia8/*/*	2.06×10^{4}
/ZToMuMu_M-200To400_TuneCP5_13TeV-powheg-pythia8/*/*	2.89×10^{3}
/ZToMuMu_M-400To800_TuneCP5_13TeV-powheg-pythia8/*/*	2.52×10^{2}
/ZToMuMu_M-800To1400_TuneCP5_13TeV-powheg-pythia8/*/*	1.71×10^{1}
/ZToMuMu_M-1400To2300_TuneCP5_13TeV-powheg-pythia8/*/*	1.37
/ZToMuMu_M-2300To3500_TuneCP5_13TeV-powheg-pythia8/*/*	8.18×10^{-2}
/ZToMuMu_M-3500To4500_TuneCP5_13TeV-powheg-pythia8/*/*	3.19×10^{-3}
/ZToMuMu_M-4500To6000_TuneCP5_13TeV-powheg-pythia8/*/*	2.79×10^{-4}
/ZToMuMu_M-6000ToInf_TuneCP5_13TeV-powheg-pythia8/*/*	9.57×10^{-6}
/TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8/*/*	8.73×10^{4}
/ST_tW_top_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8/*/*	1.96×10^{4}
/ST_tW_antitop_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8/*/*	1.96×10^{4}
/tZq_ll_4f_ckm_NLO_TuneCP5_13TeV-amcatnlo-pythia8/*/*	7.58×10^{1}
/TTWJetsToLNu_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8/*/*	2.04×10^{2}
/TTZToLLNuNu_M-10_TuneCP5_13TeV-amcatnlo-pythia8/*/*	2.53×10^{2}
/ttHJetTobb_M125_TuneCP5_13TeV_amcatnloFXFX_madspin_pythia8/*/*	2.92×10^{2}
/ttHJetToNonbb_M125_TuneCP5_13TeV_amcatnloFXFX_madspin_pythia8/*/*	2.16×10^{2}
/WW_TuneCP5_13TeV-pythia8/*/*	1.19×10^{5}
/WZ_TuneCP5_13TeV-pythia8/*/*	4.71×10^{4}
/ZZ_TuneCP5_13TeV-pythia8/*/*	1.65×10^{4}
/DYBBJetsToLL_M-50_TuneCP5_13TeV-madgraphMLM-pythia8/*'/*	9.78×10^{3}

Table 6: The background MC samples used for the development of the analysis selection. Only the first part of the sample name is given, while the * for the second and third part of the name corresponds to the latest UL, NanoAOD v9 version of the samples available for each of the 2016 APV and non-APV, 2017, and 2018 data periods. The cross section used for each sample (including any potential branching fractions that need to be applied) is also quoted. The DY+$b\bar{b}$ process is included for completeness using the only available version from NanoAOD v7. Its contribution is negligible, and in order to avoid overlaps with the DY inclusive MC samples, only events with two incoming gluons are counted. This corresponds to an effective reduction in cross section by roughly 1/3 with respect to the original LO calculations.

4.3 Triggers

The HLT paths used for the data sets from the three years of data taking used in this analyses are listed in Table 7, and are used to select both data and simulated events. The HLT_Mu50 is the main trigger path, selecting events that include at least one muon with $p_{\mathrm{T}} > 50\,\mathrm{GeV}$ and $|\eta| < 2.4$. The HLT_TkMu50 path is used for contigency in 2016, applying the same requirements but on tracker muons. The p_{T} threshold of this backup HLT path was raised to 100 GeV for 2017 and 2018 (HLT_TkMu100). In the meanwhile, a new L3 muon trigger algorithm, documented in [28], was used for the HLT_Mu50 path after 2016. To recover some inefficiencies of the new algorithm at high p_{T} that were related to its tuning at the first stages of its implementation, an HLT path using the old algorithm but with higher p_{T} threshold (HLT_OldMu100) is also used. The matching between reconstructed muons and the correspnding HLT objects are made with a $\Delta R < 0.02$ requirement. The efficiency measurement for the combination of trigger paths mentioned above, as well as the calculation of the relevant data to MC scale factors have been performed centrally by the CMS Muon POG. Details on the efficiency and the scale factors can be found in [29–31] for 2016, 2017 and 2018, respectively.

Table 7: The HLT paths used for the 2016, 2017, and 2018 data sets are listed. The version of the trigger is suppressed.

2016
HLT_(Tk)Mu50_v*
2017, 2018
HLT_Mu50_v*
HLT_(Tk,Old)Mu100_v*

4.4 $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ filters

A fraction of events in data is affected by the presence of cosmic rays, beam-gas interactions, and beam halo or calorimetric noise. Such features bias the missing transverse momentum ($\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$) calculation (Section 5.1) and are removed thanks to the dedicated filters developed by the JetMET POG [32]. The following filters are applied to both data and simulated events, according to the Run 2 recommendations [33]:

- goodVertices;
- globalSuperTightHalo2016Filter;
- HBHENoiseFilter;
- HBHENoiseIsoFilter;
- EcalDeadCellTriggerPrimitiveFilter;
- BadPFMuonFilter;
- BadPFMuonDzFilter;
- eeBadScFilter;
- ecalBadCalibFilter (only 2017+2018);
- hfNoisyHitsFilter (only 2017+2018).

5 Event reconstruction and selection

The CMS event reconstruction is based on a particle-flow (PF) algorithm [34]. The algorithm combines information from all CMS subdetectors to identify charged and neutral hadrons, photons, electrons, and muons, collectively referred to as PF candidates. Below, we describe the different identication and selection criteria of the physics objects used in this analysis.

5.1 Jet reconstruction, b-tagging, and $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$

Each event must contain at least one reconstructed pp interaction vertex. The reconstructed vertex with the largest value of the summed p_{T}^2 of physics objects is taken to be the primary vertex. The physics objects are the objects reconstructed by the anti-k_{T} jet finding algorithm [35–37] with a distance parameter of 0.4 and the associated tracks assigned to the vertex as inputs. The primary vertex is supposed to satisfy $|z| < 24\,\mathrm{cm}$ and $|\rho| < 2\,\mathrm{cm}$ in cylindrical geometry, and needs to have a number of degrees of freedom greater than 4, corresponding roughly to having two associated tracks. It is also required to be marked as valid and not fake vertex fits by the vertex reconstruction algorithm.

The jets selected for the analysis are required to satisfy $p_{\mathrm{T}} \geq 20\,\mathrm{GeV}$, unless otherwise specified, and $|\eta| < 2.5$. They must also pass at least the tight jet identification criterion [38] and be separated from all selected leptons with a requirement on the distance parameter $\Delta R > 0.4$, where $(\Delta R)^2 = (\Delta\phi)^2 + (\Delta\eta)^2$ is the distance between the two objects in the $\eta - \phi$ plane. Jets are tagged as b-tagged jets using the DeepJet algorithm [39, 40], which provides performance improvements over the DeepCSV algorithm [41, 42] by using approximately 650 input variables related to PF candidates, vertexing and jet constitutents, and improved neural network training. The b-tagging algorithm applies to all jets with $p_{\mathrm{T}} \geq 20\,\mathrm{GeV}$ and $|\eta| < 2.5$. We consider medium and tight b-tagging in order to accept events in the analysis. These tagging requirements correspond to the 'medium' ($\sim 76\%$ b-tagging efficiency for 1% misidentification probability) and 'tight' ($\sim 58\%$ b-tagging efficiency for 0.1% misidentification probability) working points (WPs) of the BTV POG, respectively. The energy of the jets is corrected for the calorimetric energy scale (jet energy scale, JES, corrections, using tags listed in Table 8). The p_{T} of the simulated jets are smeared further for the differences in resolution between the real and simulated data (jet energy resolution, JER, corrections, using tags listed in Table 9). However, the effect of such smearing is found to be negligible (see Appendix E.4). Potential discreepancies in the b-tagging efficiencies between data and simlation are corrected by using the latest BTV POG scale factor (SF) recommendations (method A) [43]. To this purpose, MC b-tagging efficiencies are computed using $\mathrm{t\bar{t}}$ MC events (see Section 4.2.2).

The missing transverse momentum, $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$, is estimated from the negative of the vector sum of all transverse momenta of PF candidates, and its magnitude is denoted as $p_{\mathrm{T}}^{\mathrm{miss}}$ [46]. The JES corrections applied on jets are propagated to $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ ('Type-1' corrections) if the ratio of sum of charged and neutral electromagnetic energy enclosed within the jet to its uncorrected jet energy is less than or equal to 0.9, and the corrected jet momentum after the subtraction of PF muons is greater than 15 GeV. The JER corrections, as well as the xy-shift correction, aiming to reduce the ϕ_{miss}modulation caused by detector effects, such as anisotropic detector responses, inactive calorimeter cells or tracking regions, the detector misalignment, the displacement of the beam spot, are also applied. The pileup-per-particle-identification (PUPPI) algorithm [47] is applied to reduce the pileup dependence of the $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ observable. The $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ is computed from the PF candidates weighted by their probability to originate from the primary interaction vertex [46].

For runs after and including 319077 in 2018, the HEM 15/16 detectors are not operational. These subdetectors correspond to $-3.2 < \eta < -1.3$, $-1.57 < \phi < -0.87$ in the $\eta - \phi$ plane.

Table 8: The jet energy correction tags for each year, separately for data and simulation [44].

2016 non-APV data sets	
Data:	Summer19UL16_RunBCDEFGH_Combined_V7_DATA
MC:	Summer19UL16_V7_MC
2016 APV data sets	
Data:	Summer19UL16_RunBCDEFGH_Combined_V7_DATA
MC:	Summer19UL16APV_V7_MC
2017 data sets	
Data:	Summer19UL17_RunB_V5_DATA
	Summer19UL17_RunC_V5_DATA
	Summer19UL17_RunD_V5_DATA
	Summer19UL17_RunE_V5_DATA
	Summer19UL17_RunF_V5_DATA
MC:	Summer19UL17_V5_MC
2018 data sets	
Data:	Summer19UL18_RunA_V5_DATA
	Summer19UL18_RunB_V5_DATA
	Summer19UL18_RunC_V5_DATA
	Summer19UL18_RunD_V5_DATA
MC:	Summer19UL18_V5_MC

Table 9: The jet energy resolution tags for simulation of each year [45].

2016 non-APV data sets	
MC:	Summer20UL16_JRV3_MC
2016 APV data sets	
MC:	Summer20UL16APV_JRV3_MC
2017 data sets	
MC:	Summer19UL17_JRV2_MC
2018 data sets	
MC:	Summer19UL18_JRV2_MC

It is observed that there are less jets in this region after all event selection criteria are applied, and that $p_{\mathrm{T}}^{\mathrm{miss}}$ has larger tails for events with at least one jet in this region. Therefore, events are vetoed if jets and/or electrons are found in this region (enlarged by 0.2 in η and ϕ for jets to account for the size of the jet cone) in the affected data taking runs. The same veto, referred to as 'HEM veto', is applied to simulated events, accounting for the fraction of affected data, and it leads to a decrease in the signal yield by at most 5% for the whole analyzed data set (2016–2018), independently of the dimuon mass (see Appendix D for more details on the performed checks).

The pileup profile in the simulation is reweighted based on the instantaneous luminosity per bunch crossing per luminosity section as a function of the number of true pileup vertices. The inelastic pp collision cross section is taken to be 69.2 mb with a 4.6% uncertainty [48–50] (see Appendix E.1 for more details on the performed checks).

5.2 Muons

The methods and algorithms to reconstruct muons are described in detail in Ref. [51]. The basic muon identification criteria used in this analysis correspond to the cut-based 'high-p_{T}' identification requirements [52], as well as longitudinal and transverse distance of closest approach requirements to increase selection purity of prompt muons. The identification criteria are as follows:

- The muon must be reconstructed as a 'global' and 'tracker' muon.
- The global muon track or the TuneP track should contain at least one valid muon hit in the muon system.
- The tracker muon must be matched to segments in at least one muon station if:
    - the muon is passing through the barrel crack and less than two segments are expected, or
    - the muon station is not on the first layer of the muon system, or
    - the tracker muon is matched to one muon station on the first layer and more than two additional RPC layers.

  If none of these conditions is satisfied, the tracker muon must be matched to segments in at least two muon stations.
- The relative p_{T} error $\delta p_{\mathrm{T}}/p_{\mathrm{T}}$ is required to be smaller than 0.3 to ensure the quality of the p_{T} measurement.
- The best track of the muon should satisfy the longitudinal closest approach requirement $|d_{\mathrm{z}}| < 0.1$ cm, and the transverse closest approach requirement $|d_{\mathrm{xy}}| < 0.02$ cm. The distance values are computed with respect to the primary vertex (PV) of the event.
- The global muon track fit must include at least one hit from each of the pixel detector and the muon system.
- The global muon track must have at least 6 tracker layers with hits in the fit.

The muons used in this analysis are required to satisfy $p_{\mathrm{T}}^{\mu} \geq 10$ GeV and $|\eta| < 2.4$, with the muons selcted to form a dimuon resonance further required to satisfy $p_{\mathrm{T}}^{\mu} \geq 53$ GeV.

All muons must pass a relative tracker-only isolation ($\mathcal{I}^{\mathrm{trk}}$) requirement: the scalar sum of the p_{T} of all other tracks in a cone of $\Delta R > 0.3$, where $(\Delta R)^2 = (\Delta\phi)^2 + (\Delta\eta)^2$, excluding the tracker track of the muon, must be less than 5 GeV and less than 5% of the tracker track p_{T} of the muon. To be used in the calculation of the tracker isolation, tracks have to be within

$|d_z| = 0.2\,\mathrm{cm}$ of the primary vertex with which the muon candidate is associated.

A subset of the muon selection requirements used in this analysis are tighter than those recommended by the Muon POG in order to minimize backgrounds from non-prompt muons and cosmic rays:

- we require the absolute muon tracker-ony isolation $\mathcal{I}^{\mathrm{trk}} < 5\,\mathrm{GeV}$;
- we require the muon transverse impact parameter $|d_{\mathrm{xy}}| < 0.02\,\mathrm{cm}$ (instead of 0.2 cm);
- we require the muon longitudinal impact parameter $|d_{\mathrm{z}}| < 0.1\,\mathrm{cm}$ (instead of 0.5 cm).

Such tighter requirements result in a reduction of the signal acceptance by less than 5%, independently of the signal mass hypothesis and signal model (see Appendix F).

Since a number of the recommended requirements for the high-p_{T} indentification do not make use of the PF algorithm, extra care needs to be taken when applying requirements on global-level variables. In addition, the muon p_{T} assignment for this analysis is performed with the 'TuneP' algorithm, which has better performance for the high-p_{T} muons that this analysis targets. A special procedure is utilized to account for possible differences in the only global-level variable used in the analysis, $p_{\mathrm{T}}^{\mathrm{miss}}$, due to the special muon reconstruction and p_{T} assignmentment algorithms implemented. The contribution of the selected muons is first excluded from the $p_{\mathrm{T}}^{\mathrm{miss}}$ computation and it is then included back after the new p_{T} assignmeent has been applied. The 'corrected' $p_{\mathrm{T}}^{\mathrm{miss}}$ is used for the event selection.

5.3 Electrons

The methods and algorithms to reconstruct electrons are described in detail in Refs. [53, 54]. The identification requirements applied for electrons used in the analysis correspond to the 'veto' cut-based criteria provided by the Egamma POG [55].

Electron isolation is computed from the flux of particle flow candidates found within a cone of $\Delta R < (10\,\mathrm{GeV}/\min(\max(p_{\mathrm{T}}^{\mathrm{e}}, 50\,\mathrm{GeV}), 200\,\mathrm{GeV}))$ built around the lepton direction. The flux of particles is computed independently for the charged hadrons ($\mathcal{I}_{\mathrm{ch}}$), neutral hadrons ($\mathcal{I}_{\mathrm{nh}}$), and photon candidates ($\mathcal{I}_{\gamma}$). The neutral hadron flux $\mathcal{I}_{\mathrm{nh}}$ is corrected for pileup by using the average energy density (ρ) due to pileup and underlying event in the central region of the detector, and an effective area correction ($A_{\mathrm{eff}}^{\mathrm{e}}$) to normalize this estimator in such a way that the isolation is independent of the number of pileup interactions. The values of $A_{\mathrm{eff}}^{\mathrm{e}}$ vary between the $|\eta_{\mathrm{SC}}|$ range and are listed in Table 10. With these quantities, the electron isolation is therefore defined as:

$$\mathcal{I}_{\mathrm{rel}}^{\mathrm{e}} = \frac{\mathcal{I}_{\mathrm{ch}} + \max(\mathcal{I}_{\mathrm{nh}} + \mathcal{I}_{\gamma} - A_{\mathrm{eff}}^{\mathrm{e}} \times \rho, 0)}{p_{\mathrm{T}}^{\mathrm{e}}}, \tag{4}$$

with $p_{\mathrm{T}}^{\mathrm{e}}$ in the denominator after electron energy corrections. The electrons used in this analysis are required to satisfy $\mathcal{I}_{\mathrm{rel}}^{\mathrm{e}} < 0.1$, $p_{\mathrm{T}}^{\mathrm{e}} \geq 10\,\mathrm{GeV}$, $|\eta| < 2.5$, $|d_{xy}| < 0.2$, and $|d_z| < 0.5$.

5.4 Isolated tracks

Roughly 85% of all tau decays result in only one charged track, being an electron 17.8% or a muon roughly 17.4% of the time, and a charged pion or kaon for the remaining 50% [56]. A single charged hadron plus possibly multiple neutral pions is thus the single largest fraction of all visible tau decay products. As a result, vetoing isolated tracks would be the most powerful veto, with tracker isolation used in order to avoid having the neutral pion decay products count

Table 10: The effective area $A^{\mathrm{e}}_{\mathrm{eff}}$ values used in each $|\eta_{\mathrm{SC}}|$ range to mitigate the dependence of the isolation requirement on pileup. The same values are used in all three years.

$\|\eta_{\mathrm{SC}}\|$ range	$A^{\mathrm{e}}_{\mathrm{eff}}$
< 1	0.1440
$[1, 1.479)$	0.1562
$[1.479, 2)$	0.1032
$[2, 2.2)$	0.0859
$[2.2, 2.3)$	0.1116
$[2.3, 2.4)$	0.1321
≥ 2.4	0.1654

towards the isolation sum. For the purpose of such a veto, isolated track are required to satisfy the following criteria:

- The isolated track must be reconstructed as a PF candidate with $|\eta| < 2.5$.
- $|d_{xy}| < 0.2$ and $|d_{\mathrm{z}}| < 0.1\,\mathrm{cm}$ with respect to the primary vertex (PV).
- $p_{\mathrm{T}} \geq 5\,(10)$ GeV for electrons or muons (charged hadrons).
- Tracker-only isolation must be less than 20% (10%) of the isolated track p_{T} for electrons or muons (charged hadrons).

5.5 Event selection

The events used in the analysis are progressively selected based on the characteristics of the final state objects and global variables. The selection starts from muons, which are the central final state objects of the analysis. First, requirements are applied to individual muons to determine the muon candidates that will form the final dimuon pair. Muons are required to have $p_T > 53\,\text{GeV}$ and be within acceptance, i.e., $|\eta| < 2.4$. They are also required to be tracker muons and pass the high-p_T identification criteria as well as the impact parameter and isolation criteria described in detail in Section 5.2. Unless at least two muons fulfill the above requirements, the event is discarded.

Then, out of these candidates muons, the two highest p_T ones are selected to form the final dimuon pair, provided they have opposite sign charge and one of them is matched to the trigger object that fired the HLT path within a cone of $\Delta R = 0.02$. The three-dimensional angle between the two muons must be smaller than $\pi - 0.02$ to minimize the contribution of potential cosmic muons. In the case where any of the previously stated dimuon pair requirements fails, the next, higher in p_T dimuon combination is tested. If no compatible pairs are found in an event, the event is discarded. The event is also discarded if the mass of the selected dimuon pair, $m_{\mu\mu}$, is less than 175 GeV.

The p_T spectra of the two selected muons, after the full event selection is applied, are shown in Fig. 5.

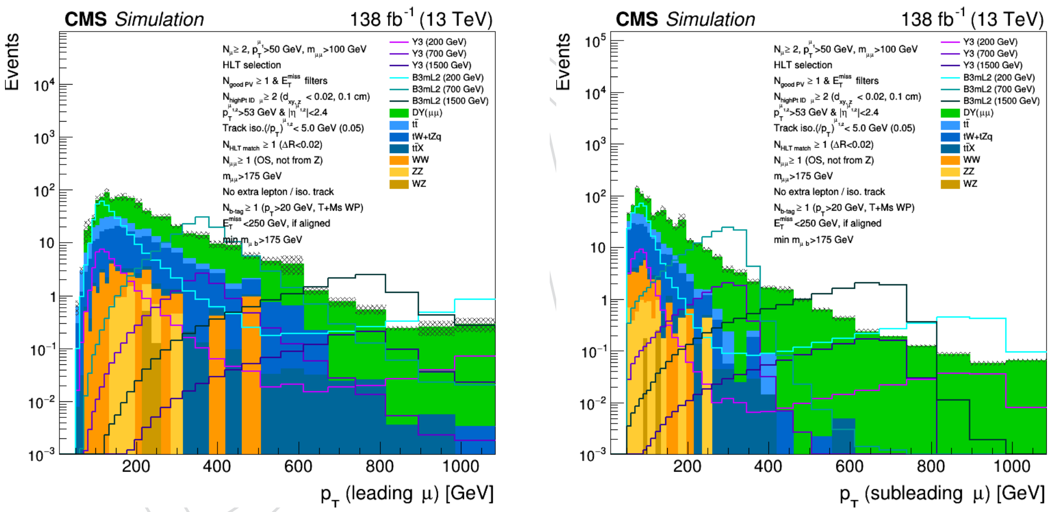


Figure 5: Distribution of the (left) leading and (right) subleading muon p_T, after the full event selection is applied.

The signal is expected to have exactly two muons in the final state, coming from the decay of the Z′ boson. As a result, any events with a third muon with $p_T > 10\,\text{GeV}$ and $|\eta| < 2.4$, passing the same single muon selection as the one mentioned above, are rejected. By the same logic, events with an electron with $p_T > 10\,\text{GeV}$ and $|\eta| < 2.5$ and/or an isolated track, fulfilling the the requirements of Section 5.3 and Section 5.4, respectively, are also rejected, in order to reduce the contribution of the WZ process. Any of these extra leptons are considered, only if it is well separated by each of the selected muons, i.e. if they satisfy $\Delta R > 0.3$ (size of isolation cone) from each muon.

Events are required to have at least one tight b-tagged jet, fulfiling the conditions described in Section 5.1. All medium b-tagged jets are included in the multiplicity count N_b. More details of event categorization are discussed in Section 6.1.

No source of significant genuine $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ is expected, except for neutrinos in the cases of semileptonic decays of the final state heavy flavor quarks. Thus, events with large missing transverse energy, $p_{\mathrm{T}}^{\mathrm{miss}} > 250\,\mathrm{GeV}$, when this is aligned ($|\Delta\phi| < 0.3$) or anti-aligned ($|\Delta\phi| > \pi - 0.3$) with any of the selected muons or b-tagged jets. This way events with largely misreconstructed muons or b-tagged jets are rejected. These requirements were chosen based on studies of $\mathrm{t\bar{t}}$ and Drell Yan MC. They reject a few instances of pathological events, eg, mismeasured muons that result in spurious high mass pairs. They have negligible effect on the efficiency of the signal under consideration, while at the same time preserving efficiency for possible other exotic signals with Z' and genuine $p_{\mathrm{T}}^{\mathrm{miss}}$ that we have no explicitly considered. The distribution of $p_{\mathrm{T}}^{\mathrm{miss}}$ for simulated signal and background events is shown in Fig. 6 together with the dimuon invariant mass distributions, before and after this selection is applied.

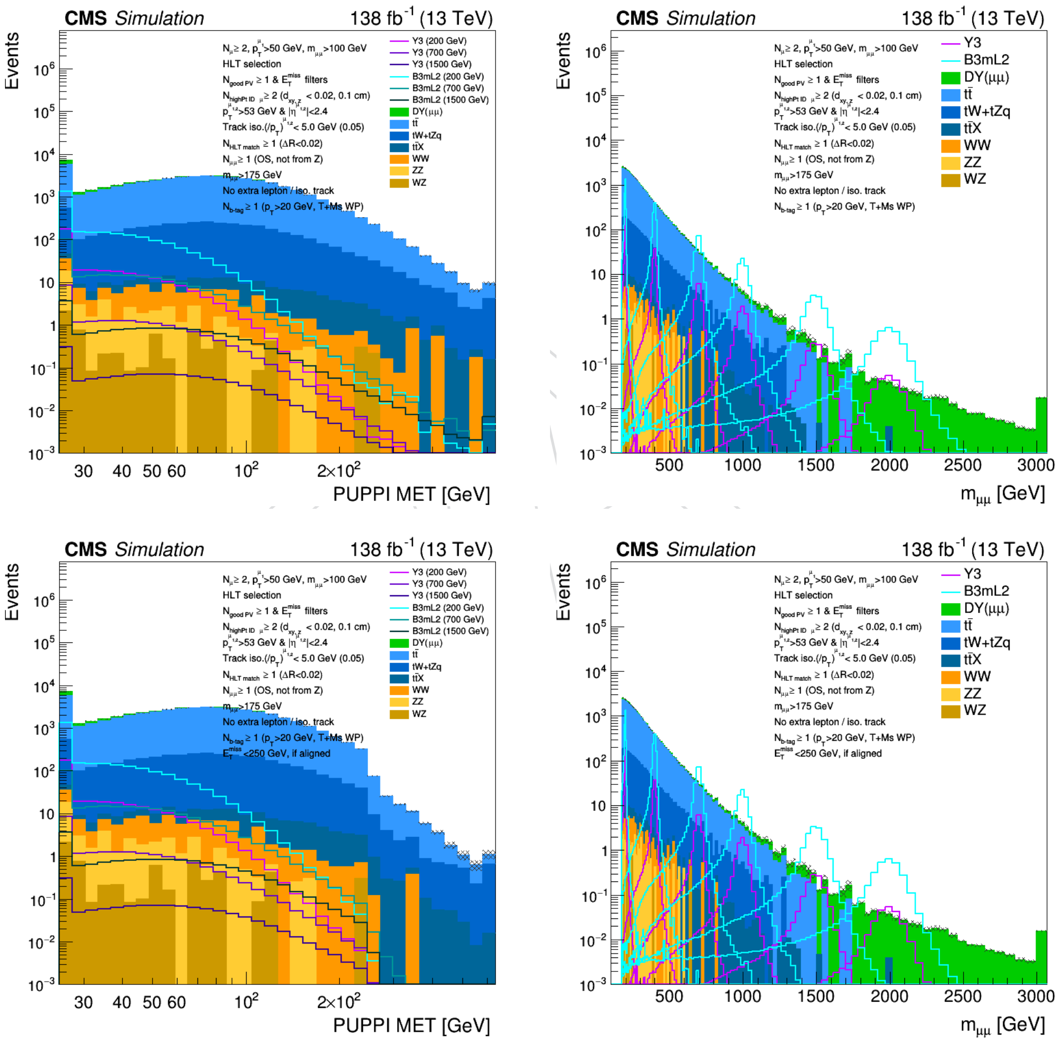


Figure 6: Distribution of (left) $p_{\mathrm{T}}^{\mathrm{miss}}$ and (right) $m_{\mu\mu}$ in the signal region, with $N_{\mathrm{b}} \geq 1$, (top) before and (bottom) after rejecting events with $p_{\mathrm{T}}^{\mathrm{miss}} > 250\,\mathrm{GeV}$ when $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ is aligned ($|\Delta\phi| < 0.3$) or anti-aligned ($|\Delta\phi| > \pi - 0.3$) with any of the selected muons or b-tagged jets.

Finally, the invariant mass of the pair of each one of them and each of the selected muons ($m_{\mu\mathrm{b}}$) is calculated. The minimum mass of all the resulting pair combination, $m_{\mu\mathrm{b}}^{\mathrm{min}}$, is used as a handle to reduce the $\mathrm{t\bar{t}}$ background, by requiring it to be greater than 175 GeV (mass of the

408 top quark). The distribution of $m_{\mu b}^{\min}$ for simulated signal and background events is shown in
409 Fig. 7, together with the dimuon invariant mass distributions after this selection is applied.

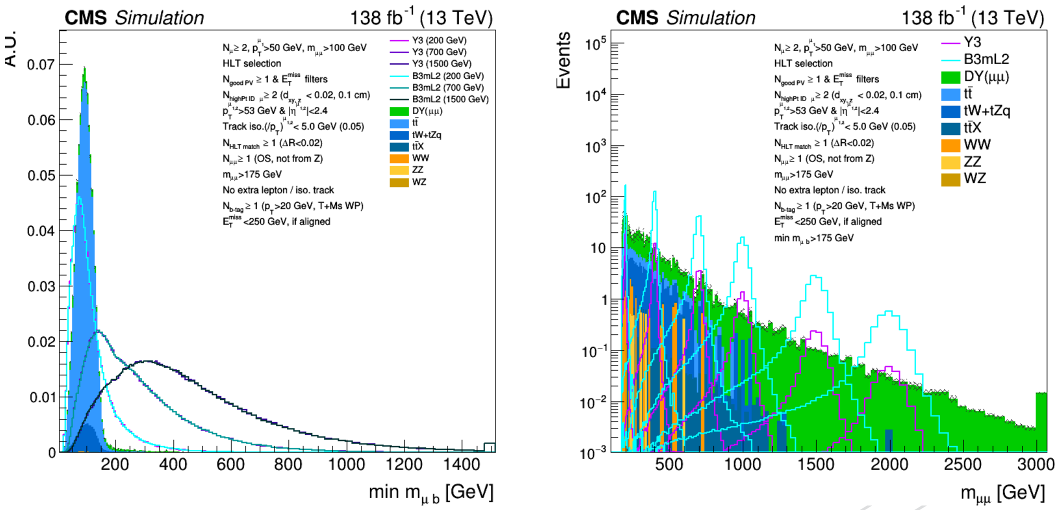


Figure 7: (Left) Distribution of $m_{\mu b}^{\min}$ in the signal region, with $N_b \geq 1$, normalized to unity for visualization. (Right) Disttributions of $m_{\mu\mu}$ in the signal region, with $N_b \geq 1$, after requiring $m_{\mu b}^{\min} > 175\,\text{GeV}$.

410 The event selection requirements are summarized in Table 11.

Table 11: The event selection requirements in the signal region are summarized. The different rows list the requirements for each quantity as well as any other explanation for how they are used.

Quantity	Requirement
p_{T}^{μ}	$> 53\,\mathrm{GeV}$ for both muons
$\lvert\eta_{\mu}\rvert$	< 2.4 for both muons
$m_{\mu\mu}$	$> 175\,\mathrm{GeV}$
N_{ℓ}	Exactly two muons passing identification criteria, no extra lepton with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$
N_{trk}	No μ or e (charged hadron) isolated track with $p_{\mathrm{T}} > 5\ (10)\,\mathrm{GeV}$
$p_{\mathrm{T}}^{\mathrm{j}}$	$> 20\,\mathrm{GeV}$
$\lvert\eta_{\mathrm{j}}\rvert$	< 2.5
N_{b}	At least one tight b-tagged jet, with possibly more medium b-tagged jets in counting
$\lvert\Delta\theta_{\mu\mu}\rvert$	$< \pi - 0.02$
$p_{\mathrm{T}}^{\mathrm{miss}}$	Event rejected if $> 250\,\mathrm{GeV}$, and $\lvert\Delta\phi_{\mathrm{miss}}^{\mathrm{b\ or\ }\mu}\rvert < 0.3$ or $\lvert\Delta\phi_{\mathrm{miss}}^{\mathrm{b\ or\ }\mu}\rvert > \pi - 0.3$ with $\Delta\phi_{\mathrm{miss}}^{\mathrm{b\ or\ }\mu}$ between $\vec{p}_{\mathrm{T}}$ of the muon or b-tagged jets, and $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$
$m_{\mu\mathrm{b}}^{\mathrm{min}}$	$> 175\,\mathrm{GeV}$ among all selected muons and b-tagged jets

6 Analysis strategy

Events selected as described in Section 5.5 are then categorized as described in Section 6.1. In each event category, the dimuon invariant mass ($m_{\mu\mu}$) distributions of signal and SM background are parametrized as described in Section 6.2, and the search is performed by fitting mass distributions to the sum of signal plus SM background models.

6.1 Event categorization

Events are categorized according to the number of b-tagged jets present in the final state, N_b. We define two categories: $N_\mathrm{b} = 1$ and $N_\mathrm{b} \geq 2$. We also define an inclusive category with $N_\mathrm{b} \geq 1$ in model-independent constraints to become fully agnostic on the signal model and in the specifics of b quark production. A dedicated study (see details in Appendix C) is performed to optimize the b-tagging working points and p_T requirements for jets in these categories. In both categories, there has to be at least one tight b-tagged jet. In the $N_\mathrm{b} \geq 2$ category, the additional b-tagged jets can be tagged according to the medium WP. The distribution of N_b for both the SM backgrounds and a few representative signal models, as obtained from simuation, is shown in Fig. 8.

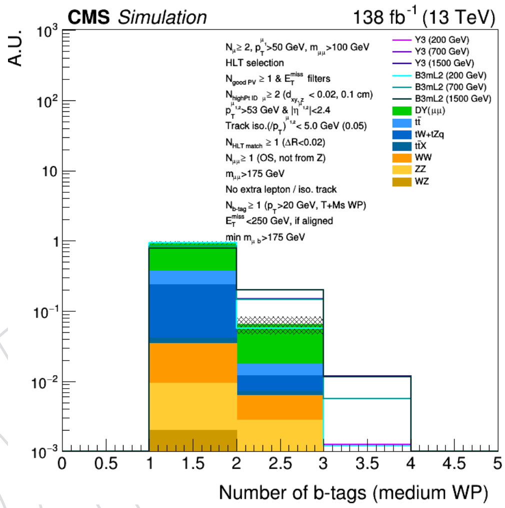


Figure 8: Distribution of N_b in the signal region, with $N_\mathrm{b} \geq 1$, normalized to unity for visualization.

In addition to the signal region event categorization described above, two additional regions enriched in SM background and depleted in signal are defined:

- DY-enriched region, defined with $N_\mathrm{b} = 0$ and with no selection requirement on $m_{\mu\mathrm{b}}^{\mathrm{min}}$, that is enriched in DY SM background;
- $\mathrm{t\bar{t}}$-enriched region, defined with $N_\mathrm{b} \geq 1$ and with an inverted requirement on $m_{\mu\mathrm{b}}^{\mathrm{min}}$ with respect to the signal region selection (i.e., with $m_{\mu\mathrm{b}}^{\mathrm{min}} < 175\,\mathrm{GeV}$), that is enriched in $\mathrm{t\bar{t}}$ SM background.

These additional regions are used to evaluate the goodness of the MC modeling of the muon kinematics in data, as shown in Appendix A and in Appendix B for the DY- and $\mathrm{t\bar{t}}$-enriched regions, respectively.

6.2 Likelihood parameterization

The search is performed by fitting mass distributions in the various categories (Section 6.1) to the sum of signal plus SM background models. We target dimuon resonances whose intrinsic width Γ is assumed to be much narrower than the detector resolution, as shown in Fig. 9, where the signal dimuon mass resolution ($\sigma^{\mu\mu}_{\text{mass}}$) is parametrized from the simulation as a function of the mass. We use the sum of a Gaussian function and a double-sided Crystal Ball function [57, 58] to model signal $m_{\mu\mu}$ distributions.

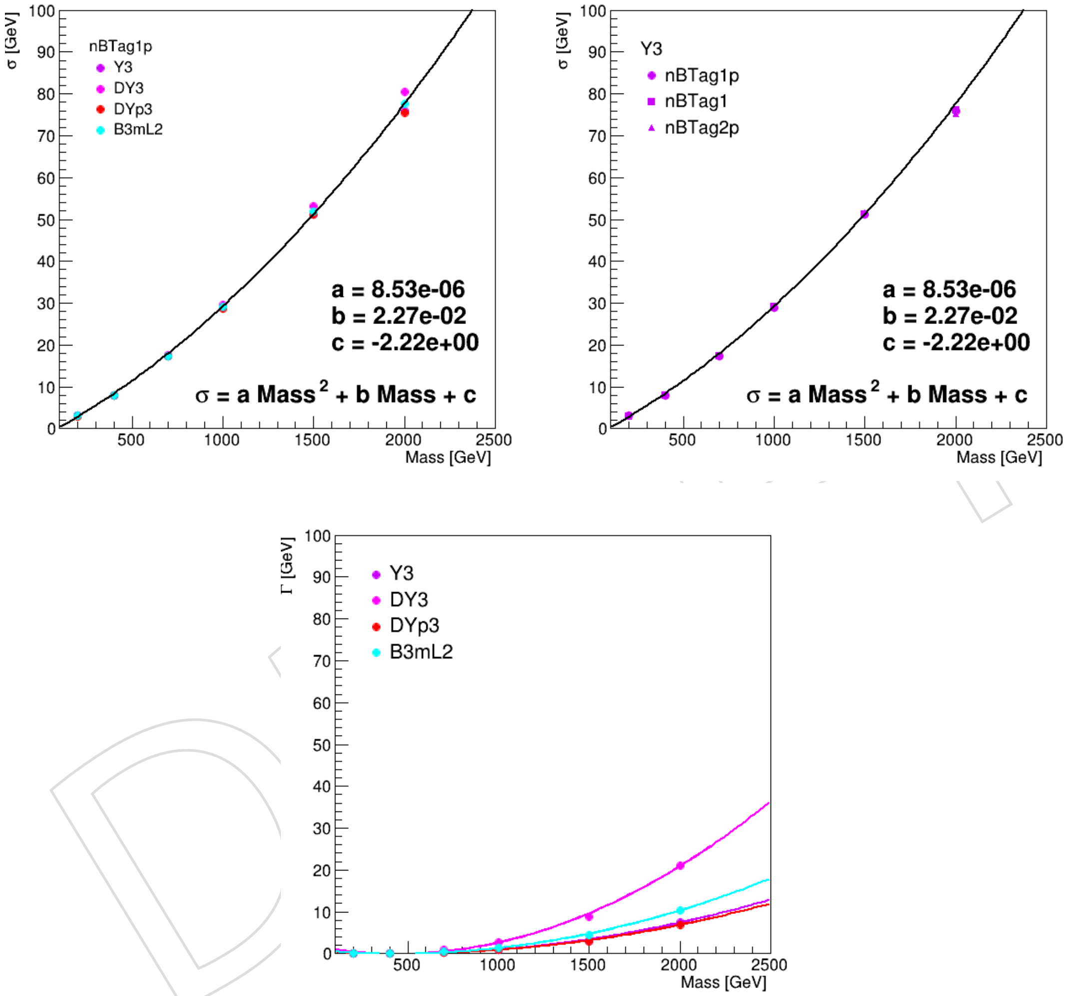


Figure 9: The expected signal dimuon invariant mass resolutioon is shown as a function of the $m_{Z'}$ mass hypothesis, together with the corresponding parametrization, in the event category with $N_{\text{b}} \geq 1$ for a number of representative signal models (top left) and for different N_{b} event categories for the Y3 signal model, as an example (top right). The signal intrinsic width Γ is also shown (bottom) as a function of the $m_{Z'}$ mass hypothesis for the same representative signal models.

In each event category, the SM background is modeled in windows of $\pm 10\sigma^{\mu\mu}_{\text{mass}}$ around the considered signal mass hypothesis. A number of functional forms are considered in order to model the background mass distribution. These include Bernstein polynomials, exponential functions, and power law functions. For Bernstein polynomials, the best order in each mass window and event category is selected by means of a Fisher test [59]. First, the lowest order

(N) function is used to fit the data. Then, the next-order ($N+1$) is used, and the difference $2\Delta\mathrm{NLL}_{N+1} = 2\,(\mathrm{NLL}_{N+1} - \mathrm{NLL}_N)$ (with NLL denoting the negative logarithm of the likelihood of the fit) is evaluated to determine whether the data support the need for a higher order function. This decision is based on the fact that $2\Delta\mathrm{NLL}_{N+1}$ is asymptotically described by a χ^2 distribution with M degrees of freedom, where M is the difference in the number of free parameters in the $(N+1)$th and Nth order functions. A p-value is calculated as $P_M\left(\chi^2 > 2\Delta\mathrm{NLL}_{N+1}\right)$ where $P_M\left(\chi^2_{\mathrm{min}}\right)$ is the χ^2 tail probability for M degrees of freedom. If the p-value is less than 0.05, the higher order function is retained, since it is determined to significantly improve the description of the data. Once the best order N for each family of functions has been determined, the corresponding functional forms are entered in an envelope then used in the *discrete profiling* method [60], where the choice of the background function is treated as a discrete nuisance parameter in the fit to account for the uncertainty associated with the arbitrary choice of the function. For Bernstein polynomials, we also include orders $N-1$ (if the selected best order $N > 0$) and $N+1$ in the list of suitable functions. Typically, the selected best order N is ≤ 2 and is smaller in event categories with a small number of observed events. In addition, we evaluate the goodness of the fit for each model based on a χ^2 test statistic, which is converted into a p-value: models with $p < 0.01$ are not considered.

The potential presence of a bias in the measurement of a signal due to the choice of the functional forms used to model the SM background will also be assessed, by means of pseudo-experiments. First, a varying amount of signal will be injected on top of the background generated according to the selected functional form. Then, a background+signal fit will be performed allowing the signal yield to float freely. The bias will be quantified as the difference between the measured and injected signal yields relative to the statistical uncertainty in the measured signal yield.

The results of the fits of the dimuon invariant mass distributions expected for representative signals in search bins with $N_{\mathrm{b}} \geq 1$ are shown in Fig. 10. The background-only fit results for the selected functional forms in the corresponding search bins with $N_{\mathrm{b}} \geq 1$ are shown in Figs. 11–16, using a toy MC dataset in place of the actual distribution in data.

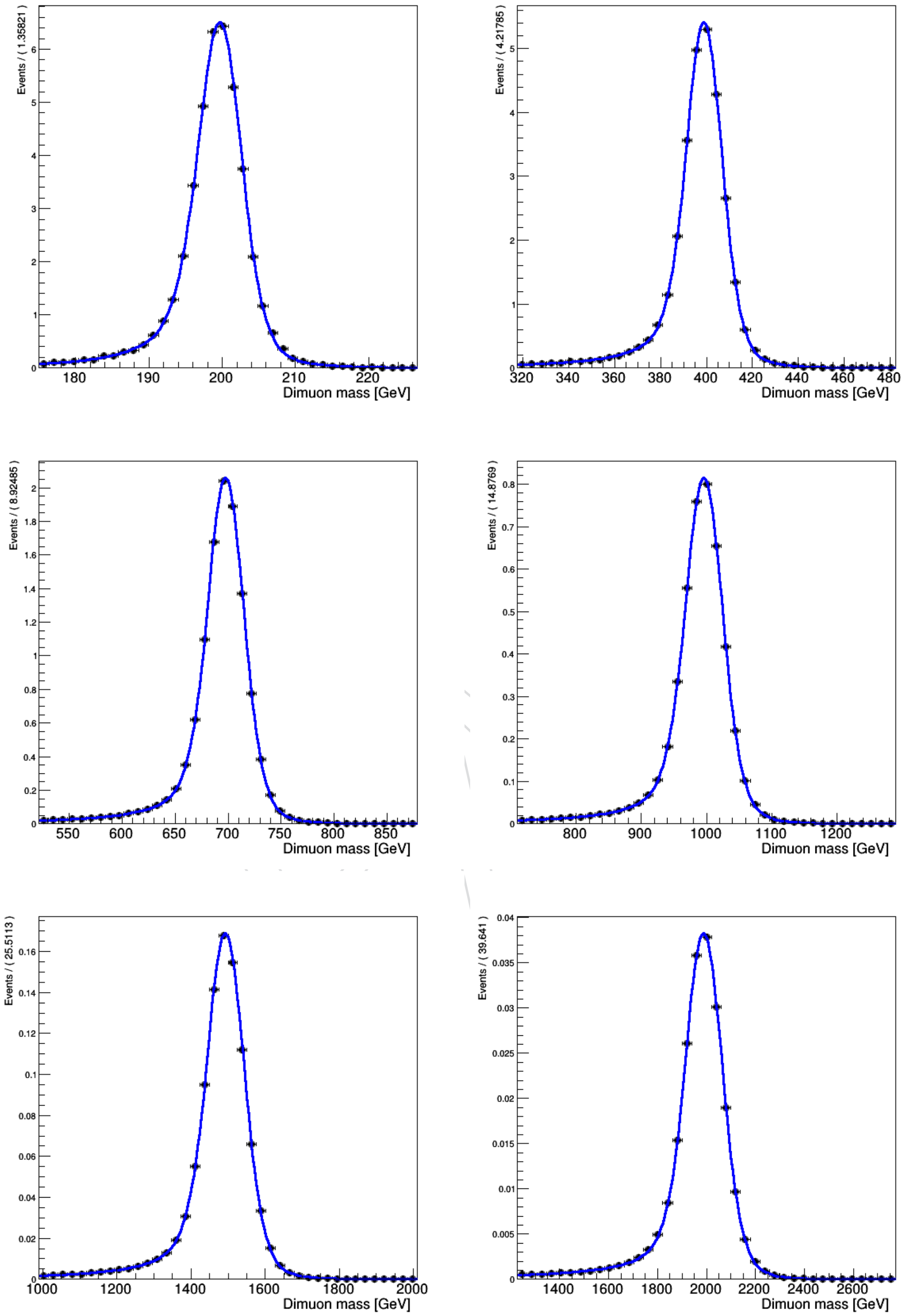


Figure 10: The invariant mass distributions expected for the Y3 signal model in the event category with $N_b \geq 1$ are shown with $m_{Z'} = 200$ (top left), 400 (top right), 700 (middle left), 1000 (middle right), 1500 (bottom left), and 2000 GeV (bottom right), together with the corresponding fit results (blue).

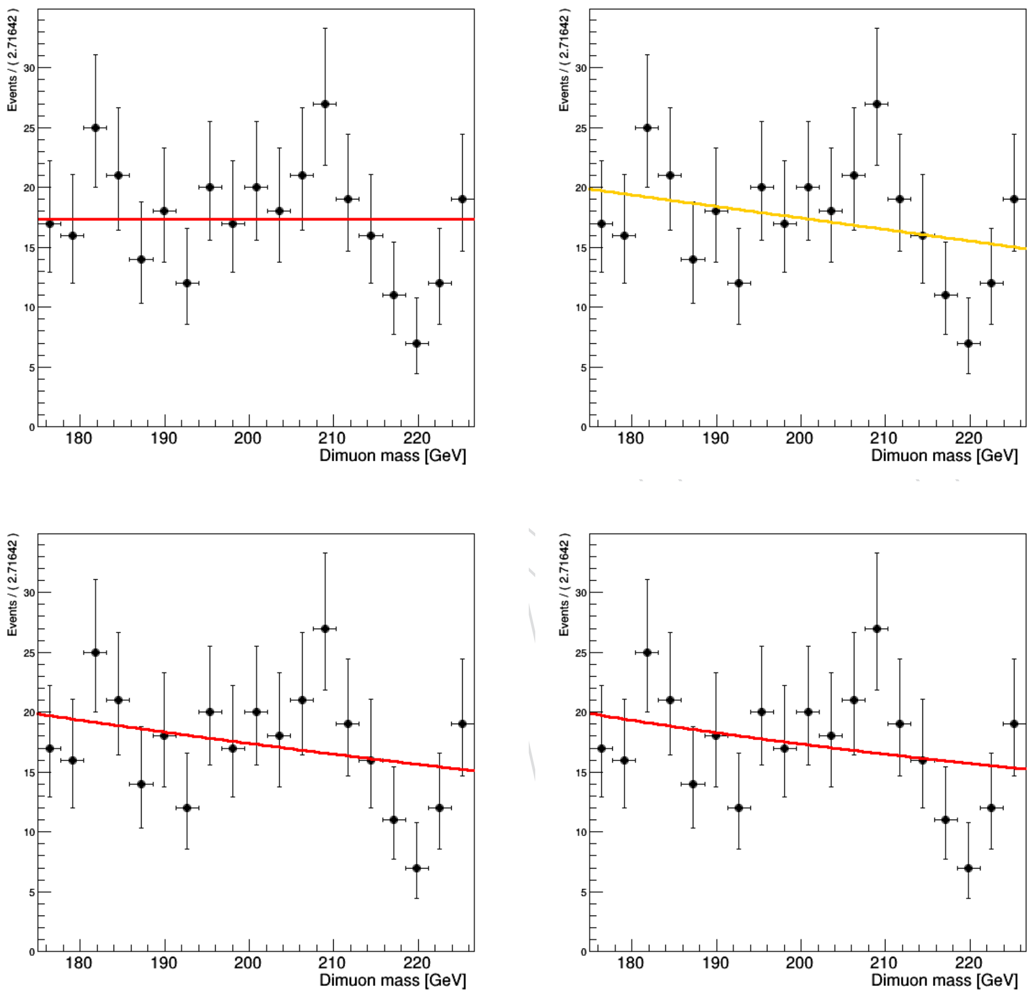


Figure 11: The invariant mass distribution expected for background (using a toy MC dataset) in the event category with $N_b \geq 1$ is shown in the search bin corresponding to $m_{Z'} = 200\,\text{GeV}$, together with the fit resuts using a Bernstein polynmial model of the best selected order $N = 0$ (top left) and of order $N + 1$ (top right), an exponential model (bottom left) and a power-law model (bottom right).

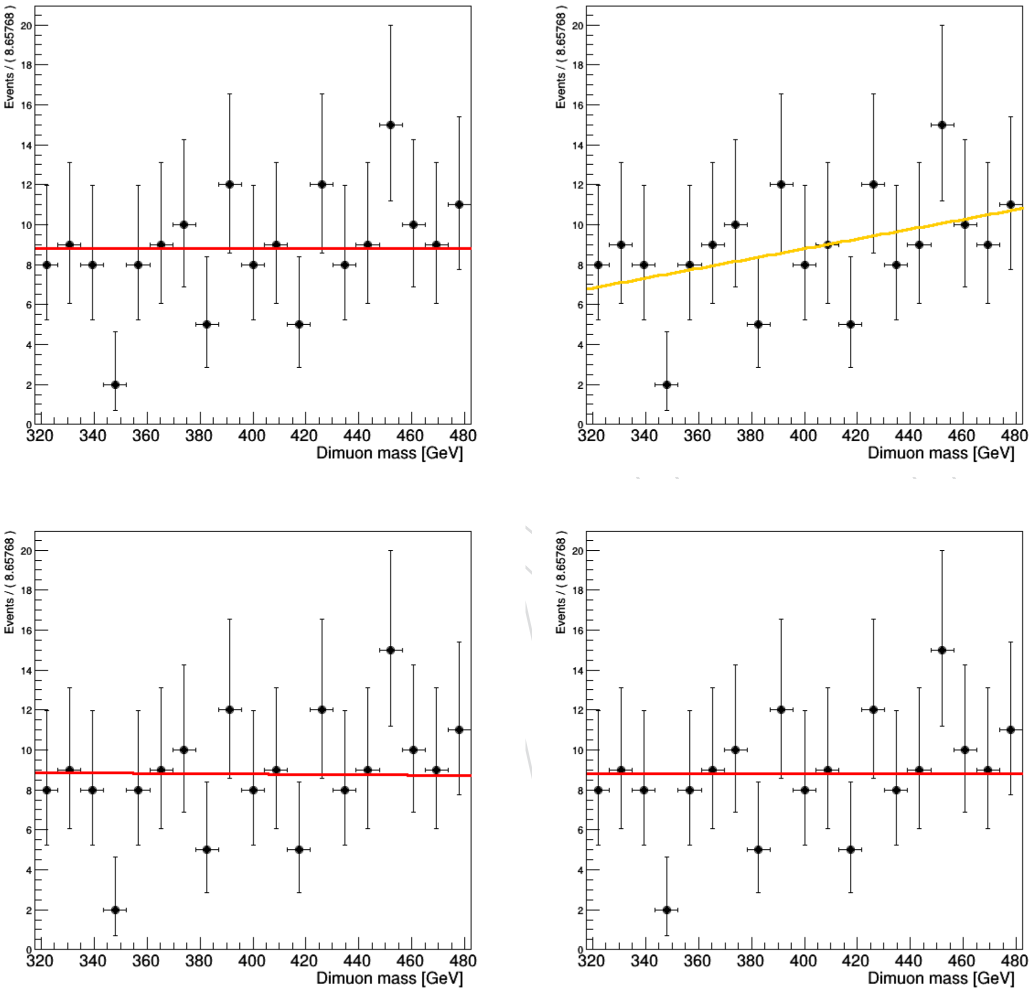


Figure 12: The invariant mass distribution expected for background (using a toy MC dataset) in the event category with $N_\mathrm{b} \geq 1$ is shown in the search bin corresponding to $m_{Z'} = 400\,\mathrm{GeV}$, together with the fit resuts using a Bernstein polynmial model of the best selected order $N = 0$ (top left) and of order $N + 1$ (top right), an exponential model (bottom left) and a power-law model (bottom right).

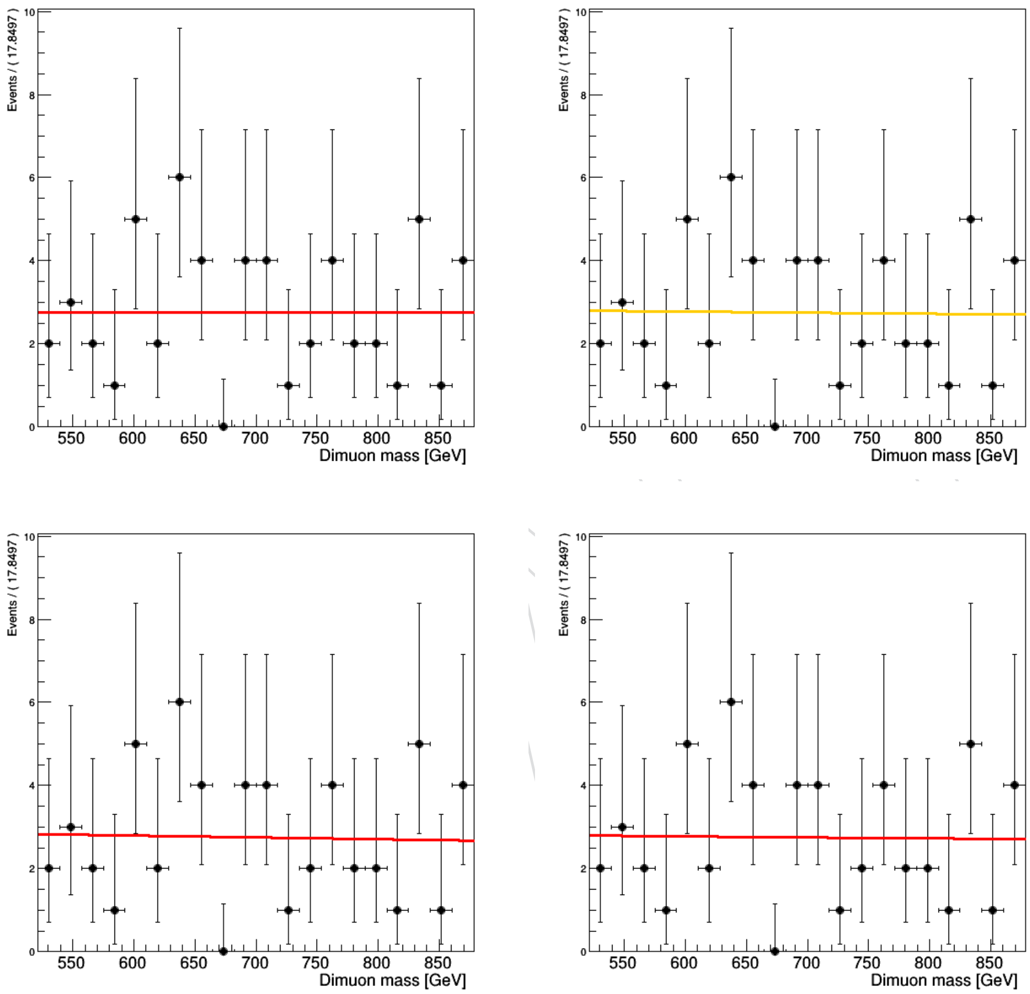


Figure 13: The invariant mass distribution expected for background (using a toy MC dataset) in the event category with $N_b \geq 1$ is shown in the search bin corresponding to $m_{Z'} = 700\,\text{GeV}$, together with the fit resuts using a Bernstein polynmial model of the best selected order $N = 0$ (top left) and of order $N + 1$ (top right), an exponential model (bottom left) and a power-law model (bottom right).

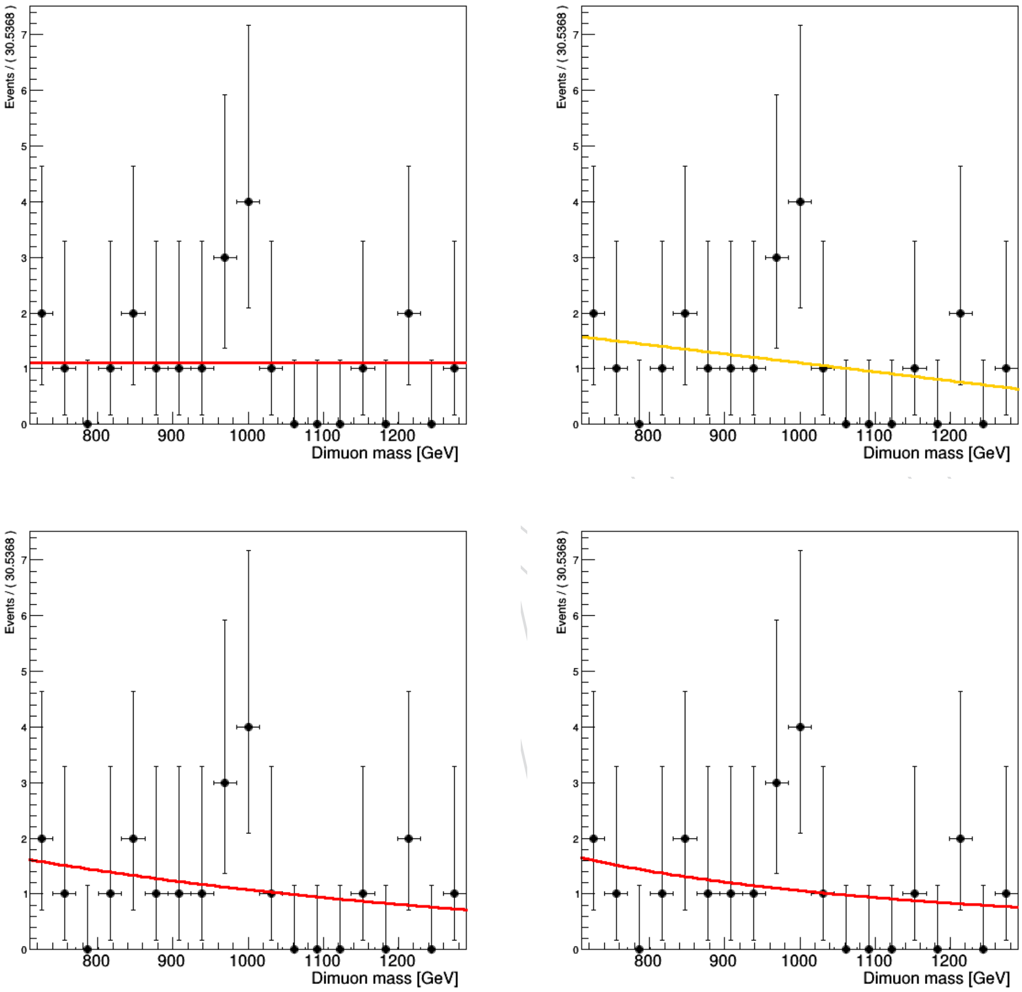


Figure 14: The invariant mass distribution expected for background (using a toy MC dataset) in the event category with $N_b \geq 1$ is shown in the search bin corresponding to $m_{Z'} = 1000\,\text{GeV}$, together with the fit resuts using a Bernstein polynmial model of the best selected order $N = 0$ (top left) and of order $N + 1$ (top right), an exponential model (bottom left) and a power-law model (bottom right).

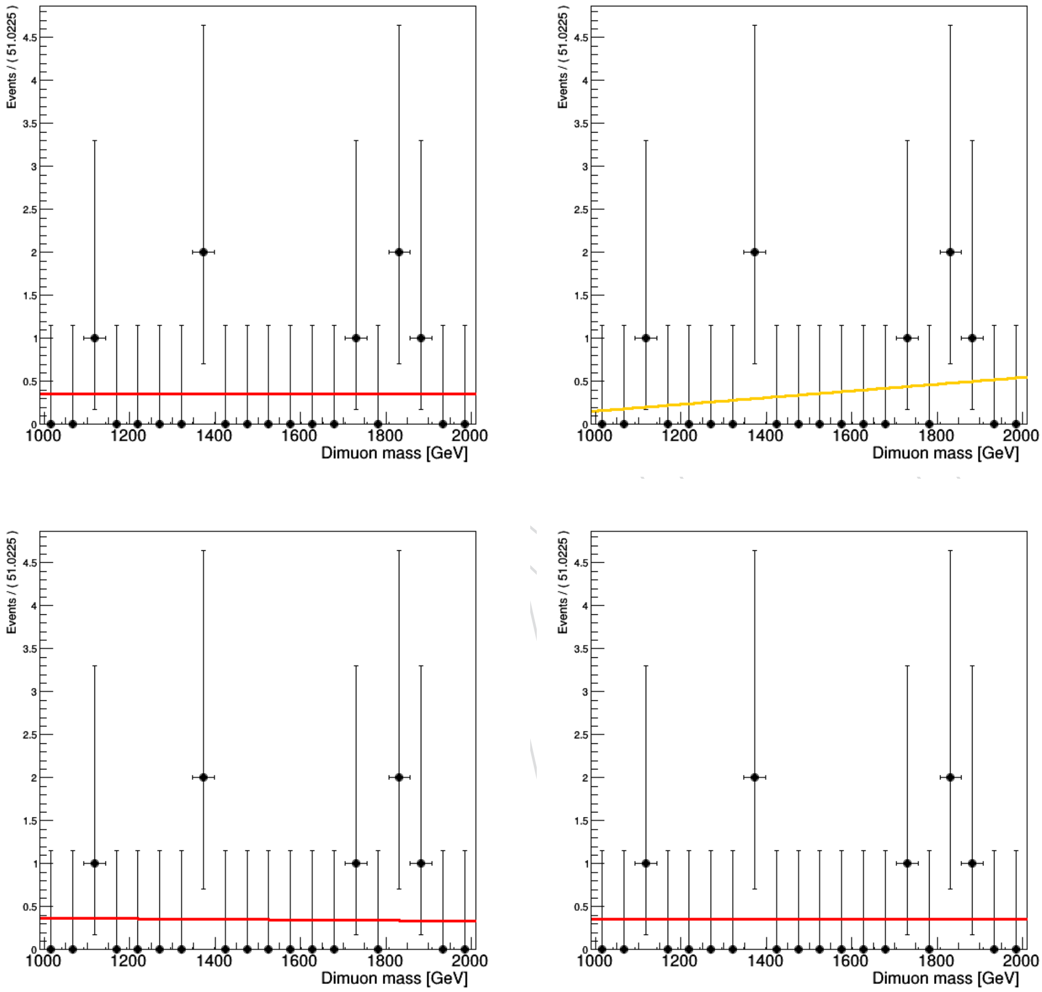


Figure 15: The invariant mass distribution expected for background (using a toy MC dataset) in the event category with $N_b \geq 1$ is shown in the search bin corresponding to $m_{Z'} = 1500\,\text{GeV}$, together with the fit resuts using a Bernstein polynmial model of the best selected order $N = 0$ (top left) and of order $N + 1$ (top right), an exponential model (bottom left) and a power-law model (bottom right).

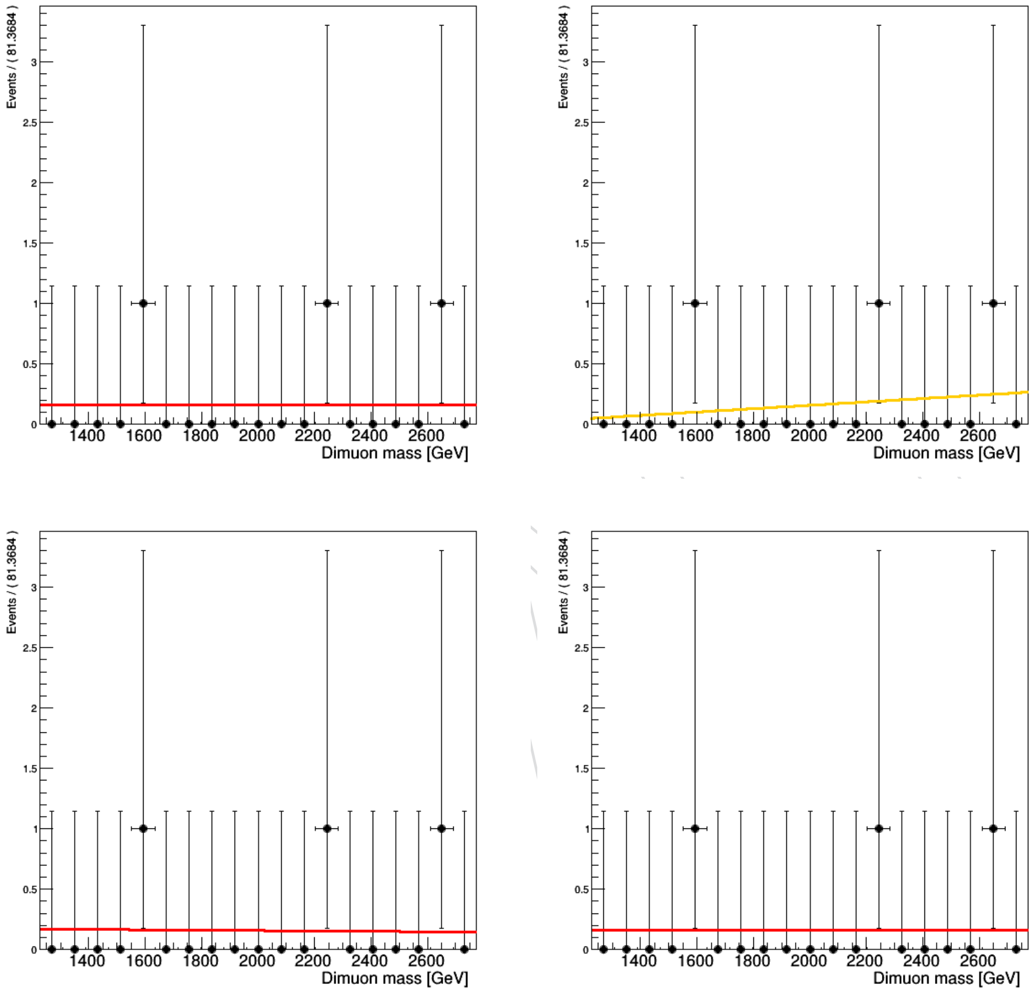


Figure 16: The invariant mass distribution expected for background (using a toy MC dataset) in the event category with $N_{\mathrm{b}} \geq 1$ is shown in the search bin corresponding to $m_{Z'} = 2000\,\mathrm{GeV}$, together with the fit resuts using a Bernstein polynmial model of the best selected order N (top left) and of order $N+1$ (top right), an exponential model (bottom left) and a power-law model (bottom right).

6.3 Signal systematic uncertainties

In this Section we discuss the systematic uncertainties in the analysis. For context, it is useful to keep in mind that, when operating in the Poisson regime, efficiency uncertainties in limit setting have very little effect, see for example Fig. 1 in Ref. [61]. These effects are often smaller than those introduced by the (arbitrary) choice of method used to calculate the limit itself (Bayesian, Classic Frequentist, Feldman-Cousins, various types of CL_s, etc.), see for example [62].

A systematic uncertainty equal to 1.6% in the expected signal yields arising from the uncertainty in the luminosity measurement is assessed [63–65]. Additionally, we evaluate the effect of all sources of uncertainty related to pileup modeling, trigger efficiency measurement, as well as physics object recontruction:

- pileup modeling;
- trigger efficiency measurement;
- jet energy resolution (JER);
- jet energy scale (JES);
- b-tagging efficiency;
- muon reconstruction, identification, and isolation.

The effect of each source of uncertainty is described in detail in Appendix E. As described in Section 5.2, we select muons using tight requirements, that are tighter than the recommended ones. Such tighter requirements result in a reduction of the signal acceptance by less than 5%, independently of the signal mass hypothesis and signal model (see Appendix F). This reduction in signal acceptance is conservatively accounted for as an additional systematic uncertainty in the expected signal yield. Finally, we account for uncertainties resulting from the limited sizes of the simulated signal samples. Uncertainties arising from the choice of the PDF and of the renormalization and factorization scales used in the event generator are negligible compared to others.

The resulting uncertainties are summarized in Table 12, together with their typical (range of) values. Uncertainties whose effect is measured to be negligible are omitted.

Uncertainties arising from integrated luminosity, trigger efficiency, b-tagging efficiency, and muon reconstruction, identification and isolation are treated as correlated across event categories. Other uncertainties are taken as uncorrelated. Since we find no significant difference between data and simulations in 2016, 2017 and 2018, data and simulations from different data taking periods are used as a whole, i.e., with full correlation across different data taking periods.

Table 12: Summary of uncertainties, with their typical (range of) values. Uncertainties whose effect is measured to be negligible are omitted.

Source	Normalization		Shape
	$N_b = 1$	$N_b \geq 2$	
Integrated luminosity	1.6%		—
Trigger efficiency	1–5%		—
JES	1–1.5%	2–5%	—
b-tagging	1%	5%	—
μ reconstruction, ID and isolation	2.5%		—
Additional μ selection criteria	5%		—
Limited MC size	$< 1\%$	$< 5\%$	—

7 Results

7.1 Event yields and distributions

The dimuon invariant mass distributions in each event category are shown in Fig. 17, as obtained from simulation for both the SM backgrounds and a few representative signal models.

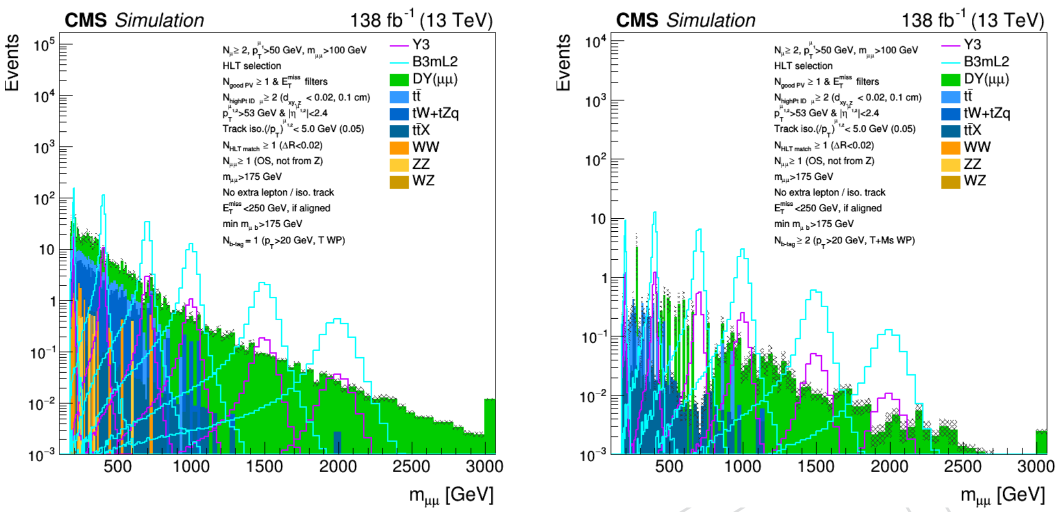


Figure 17: Distributions of $m_{\mu\mu}$ in the signal regions with (left) $N_{\mathrm{b}} = 1$ and (right) $N_{\mathrm{b}} \geq 2$, as epected from simulation, for both the SM backgrounds and a few representative signal models.

7.2 Interpretations

Unbinned maximum likelihood fits to the data are performed simultaneously in all event categories, under either background-only or background+signal hypotheses, using background and signal models and uncertainties described in Section 6.2. Additional log-normal constraint terms are used to account for the uncertainties in the signal yields, when considered (Section 6.3).

The combined fits for the signal and background are used to set 95% confidence level (CL) upper limits on the production cross section for the signal models under consideration. Limits are set using a modified frequentist approach, employing the CL_s criterion [66–69]. These limits are then used, in conjunction with the theoretical cross section calculations, to exclude ranges of masses for the BSM particles of the signal models.

For specific signal models, simultaneous fits in the event categories with $N_b = 1$ and $N_b \geq 2$ are performed. In order to set model-indepdendent constraints, fits are performed either inclusively in N_b, i.e., in the event category with $N_b \geq 1$, or simultaneously in the event categories with $N_b = 1$ and $N_b \geq 2$ varying the relative signal acceptance in each category to probe a range of hypotheses of signal production in association with b quarks.

7.3 MC expected limits before unblinding

In this Section we document the expected limits based on studies of MC data. We emphasize that the background in the analysis is taken directly from the mass distribution after unblinding. The overall background level used to extract the expected limits shown in this Section is taken from SM Monte Carlo. Thus the expected limits shown here may be a bit different from the final ones.

Results are based on the procedure discussed in more detail in Section XX.

7.3.1 Model independent limits

Limits on the number of possible BSM events passing the requirements (SR1 + SR2) as a function of mass and of the fraction (f_2) of BSM events in SR2 is shown in Fig. 18. The limit on the number of events is more stringent as f_2 increases, since the background level in SR2 ($N_b \geq 2$) is smaller.

These limits are provided for "reinterpretation". A phenomenologist interested in some Z′ BSM model not considered here should be able to use a MC event generator and a fast Delphes-like detector simulation to approximately predict the number of expected events in 138 fb^{-1} in SR1 and SR2 for that BSM model. Then, the number of expected events and the expected f_2 can be compared with the limits in Fig. 18 to decide whether the model is excluded by our analysis.

7.3.2 Limits on models from the literature

In this Section we present expected limits for the four models (Y_3, DY_3, DY_3', and $B_3 - L_2$) discussed in Section 4.2.1.1. The expected yields in SR1 and SR2 for each model as a function of mass, θ_{23}, and $x = g_X(1\,\mathrm{TeV}/M_X)$, are obtained starting from the MC samples described in Section 4.2.1.1 and using the reweighting procedure described in Appendix G.1. These yields are then compared with the curves of Fig. 18 to obtain exclusion contours in the relevant parameter space.

In Fig. 19 we show limits in the θ_{23} vs. x plane for a few choices of Z′ mass. Note that at a given mass the limit curve is almost independent of θ_{23}. This can be understood by the fact

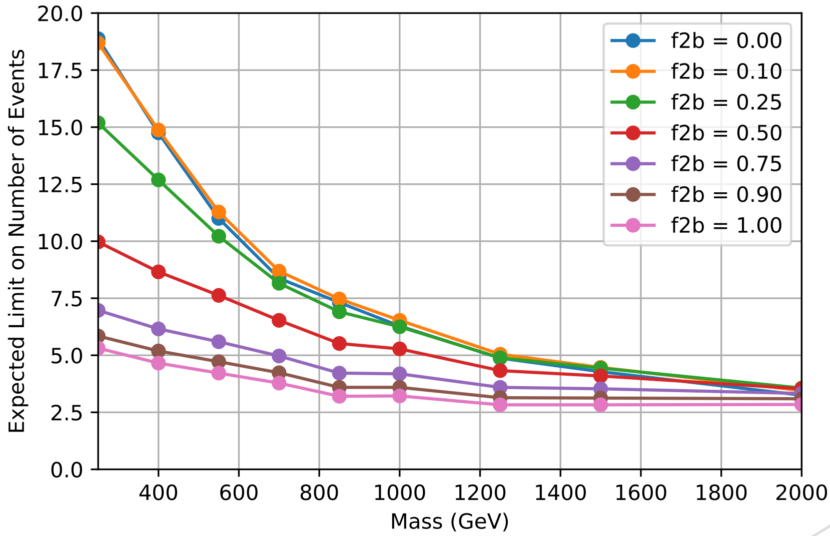


Figure 18: Expected limits on the number of detected BSM events as a function of mass and f_2. The quantity f_2 is the fraction of BSM events passing the analsis that have at least two b tags.

.

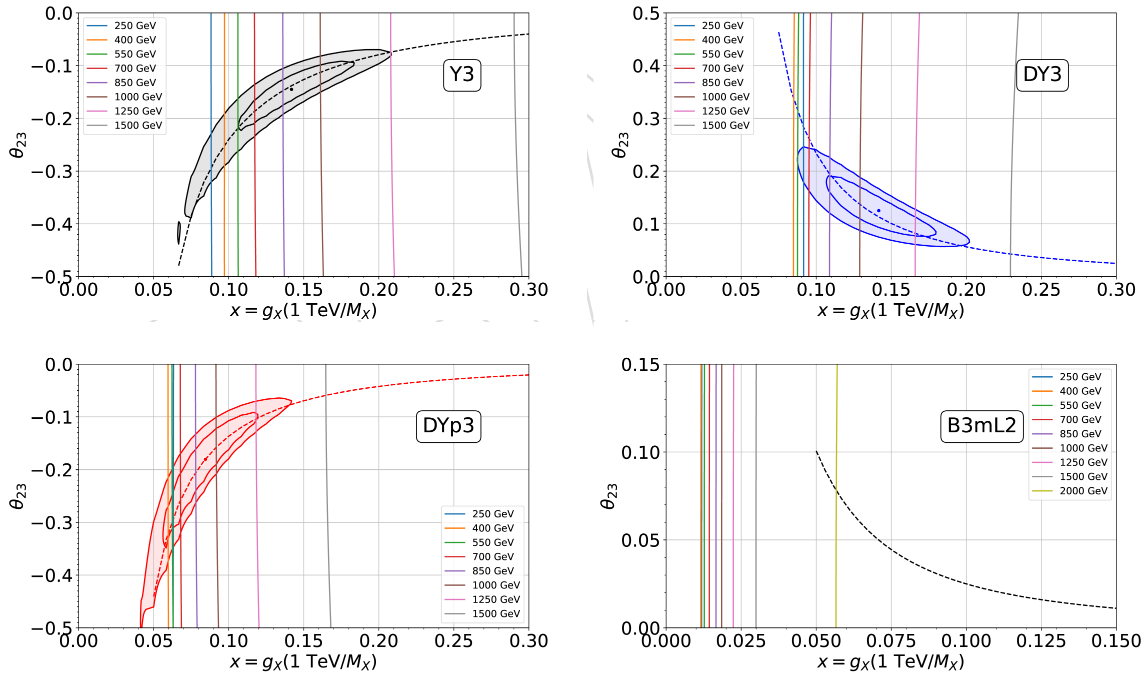


Figure 19: Expected limits in the θ_{23} vs. x plane for the four models from the literature for a few values of M_X. For a given mass, the region to the right of the corresponding line is excluded. In the case of the "Third Family Hypercharge" models (Y_3, DY_3, and DY_3') we superimpose the allowed contours (at 68 and 95% confidence) from Fig. 3. Such a contour does not exist for the $B_3 - L_2$ model. In this case the dashed line shows the preferred (and allowed!) value for θ_{23} as a function of x.

.

that the bulk of the cross section in these models originates from the $b\bar{b}Z'$ coupling which is

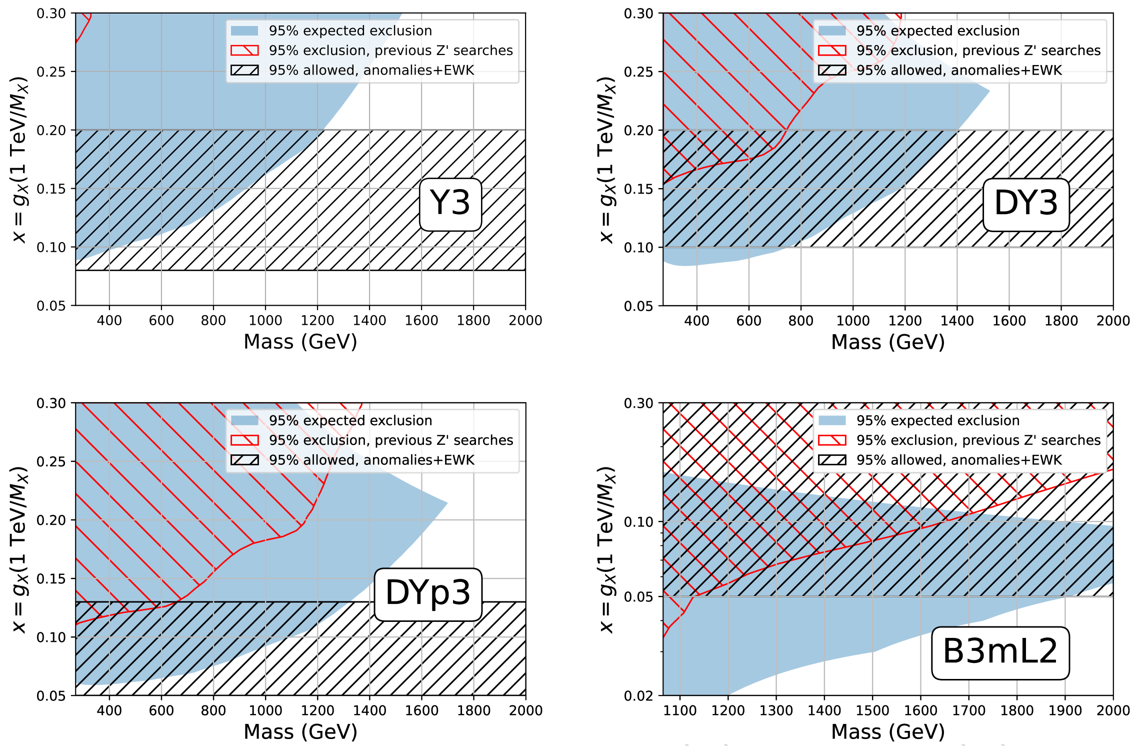


Figure 20: Expected limits in the x vs. mass plane for the four models from the literature. The constraints from inclusive Atlas and CMS searches, EWK data, and LHCb anomalies are from Ref. [26]. The expected limits are computed at model-dependent values of θ_{23} listed in Table 4. Limits are only show for the regions of parameter space where the Z' width is smaller than one-half the $\mu\mu$ invariant mass resolution.

independent of θ_{23}.

Neglecting the small θ_{23} dependence, it is then interesting to display the limits in the x vs. M_X plane. These limits, shown in Fig. 20, can then also be compared with limits obtained by inclusive Z' searches at the LHC. Note that as discussed in Appendix G.1 and shown in Fig.128, as x increases the expected width Γ of the Z' can approach and exceeds the $\mu\mu$ invariant mass resolution σ. We are calculating limits in the "narrow width" approximation, ie, assuming that the width of the $\mu\mu$ invariant mass peak is dominated by resolution effects. When this assumption fails, the calculation of the limits would need a different treatment. For now we exclude from the plots in Fig. 20 points where $\Gamma > 2\sigma$.

8 Summary

A search for high-mass dimuon resonance production in association with one or more b quark jets is presented, using data collected with the CMS experiment at the LHC that correspond to an integrated luminosity of 138 fb^{-1} at a center-of-mass energy of 13 TeV. Model-independent constraints are derived on the numbers of events with $N_{\mathrm{b}} = 1$ and ≥ 2, and also their total with ≥ 1. The constraints are presented as a function of the analyzed Z′ mass values $m_{\mathrm{Z}'}$. Results are also interpreted in terms of models that involve possible Z′sb, Z′bb, and Z′$\mu\mu$ couplings, with the constraints presented in terms of the coupling strength g_X, the b-s mixing angle θ_{sb}. and $m_{\mathrm{Z}'}$.

References

[1] LHCb Collaboration, "Test of lepton universality in beauty-quark decays", *Nature Phys.* **18** (2022), no. 3, 277–282, doi:10.1038/s41567-021-01478-8, arXiv:2103.11769.

[2] CMS Collaboration, "Search for resonant and nonresonant new phenomena in high-mass dilepton final states at $\sqrt{s} = 13$ TeV", *JHEP* **07** (2021) 208, doi:10.1007/JHEP07(2021)208, arXiv:2103.02708.

[3] ATLAS Collaboration, "Search for high-mass dilepton resonances using 139 fb^{-1} of pp collision data collected at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Phys. Lett. B* **796** (2019) 68–87, doi:10.1016/j.physletb.2019.07.016, arXiv:1903.06248.

[4] ATLAS Collaboration, "Search for New Phenomena in Final States with Two Leptons and One or No *b*-Tagged Jets at $\sqrt{s} = 13$ TeV Using the ATLAS Detector", *Phys. Rev. Lett.* **127** (2021), no. 14, 141801, doi:10.1103/PhysRevLett.127.141801, arXiv:2105.13847.

[5] ATLAS Collaboration, "Search for scalar resonances decaying into $\mu^+\mu^-$ in events with and without *b*-tagged jets produced in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *JHEP* **07** (2019) 117, doi:10.1007/JHEP07(2019)117, arXiv:1901.08144.

[6] CMS Collaboration, "Search for MSSM Higgs bosons decaying to $\mu + \mu -$ in proton-proton collisions at s=13TeV", *Phys. Lett. B* **798** (2019) 134992, doi:10.1016/j.physletb.2019.134992, arXiv:1907.03152.

[7] CMS Collaboration, "Search for long-lived particles decaying into muon pairs in proton-proton collisions at $\sqrt{s}$ = 13 TeV collected with a dedicated high-rate data stream", *JHEP* **04** (2022) 062, doi:10.1007/JHEP04(2022)062, arXiv:2112.13769.

[8] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[9] CMS Collaboration, "CMS technical design report for the pixel detector upgrade", Technical Report CERN-LHCC-2012-016, CMS-TDR-011, 2012. doi:10.2172/1151650.

[10] CMS Collaboration, "The CMS trigger system", *JINST* **12** (2017) P01020, doi:10.1088/1748-0221/12/01/P01020, arXiv:1609.02366.

[11] CMS Collaboration, "Run-2 UltraLegacy Datasets for Analysis", 2022. `https://twiki.cern.ch/twiki/bin/view/CMS/PdmVRun2LegacyAnalysis`.

[12] CMS Collaboration, "UL2016 Datasets for Analysis", 2021. `https://twiki.cern.ch/twiki/bin/view/CMS/PdmVDatasetsUL2016`.

[13] CMS Collaboration, "UL2017 Datasets for Analysis", 2021. `https://twiki.cern.ch/twiki/bin/view/CMS/PdmVDatasetsUL2017`.

[14] CMS Collaboration, "UL2018 Datasets for Analysis", 2022. `https://twiki.cern.ch/twiki/bin/view/CMS/PdmVDatasetsUL2018`.

[15] CMS Collaboration, "Legacy-2016 (pre-VFP) Analysis", 2021. `https://twiki.cern.ch/twiki/bin/view/CMS/PdmVLegacy2016preVFPAnalysis`.

[16] CMS Collaboration, "Legacy-2016 (post-VFP) Analysis", 2021. `https://twiki.cern.ch/twiki/bin/view/CMS/PdmVLegacy2016postVFPAnalysis`.

[17] CMS Collaboration, "Legacy-2017 Analysis", 2021. `https://twiki.cern.ch/twiki/bin/view/CMS/PdmVLegacy2017Analysis`.

[18] CMS Collaboration, "Legacy-2018 Analysis", 2021. `https://twiki.cern.ch/twiki/bin/view/CMS/PdmVLegacy2018Analysis`.

[19] CMS Collaboration, "Luminosity Physics Object Group (Lumi POG)", 2020. `https://twiki.cern.ch/twiki/bin/viewauth/CMS/TWikiLUM`.

[20] B. C. Allanach, "$U(1)_{B_3-L_2}$ explanation of the neutral current B-anomalies", *Eur. Phys. J. C* **81** (2021), no. 1, 56, `doi:10.1140/epjc/s10052-021-08855-w`, `arXiv:2009.02197`. [Erratum: Eur.Phys.J.C 81, 321 (2021)].

[21] B. C. Allanach and J. Davighi, "Third family hypercharge model for $R_{K^{(*)}}$ and aspects of the fermion mass problem", *JHEP* **12** (2018) 075, `doi:10.1007/JHEP12(2018)075`, `arXiv:1809.01158`.

[22] B. C. Allanach and J. Davighi, "Naturalising the third family hypercharge model for neutral current *B*-anomalies", *Eur. Phys. J. C* **79** (2019), no. 11, 908, `doi:10.1140/epjc/s10052-019-7414-z`, `arXiv:1905.10327`.

[23] B. C. Allanach, J. E. Camargo-Molina, and J. Davighi, "Global fits of third family hypercharge models to neutral current B-anomalies and electroweak precision observables", *Eur. Phys. J. C* **81** (2021), no. 8, 721, `doi:10.1140/epjc/s10052-021-09377-1`, `arXiv:2103.12056`.

[24] R. Alonso, P. Cox, C. Han, and T. T. Yanagida, "Flavoured $B - L$ local symmetry and anomalous rare B decays", *Phys. Lett. B* **774** (2017) 643–648, `doi:10.1016/j.physletb.2017.10.027`, `arXiv:1705.03858`.

[25] C. Bonilla, T. Modak, R. Srivastava, and J. W. F. Valle, "$U(1)_{B_3-3L_\mu}$ gauge symmetry as a simple description of $\mathrm{b} \to \mathrm{s}$ anomalies", *Phys. Rev. D* **98** (2018), no. 9, 095002, `doi:10.1103/PhysRevD.98.095002`, `arXiv:1705.00915`.

[26] B. C. Allanach, J. M. Butterworth, and T. Corbett, "Large Hadron Collider constraints on some simple Z' models for $\mathrm{b} \to \mathrm{s}\mu^+\mu^-$ anomalies", *Eur. Phys. J. C* **81** (2021), no. 12, 1126, `doi:10.1140/epjc/s10052-021-09919-7`, `arXiv:2110.13518`.

[27] “EXO-22-006 CADI”. https://cms.cern.ch/iCMS/analysisadmin/cadilines?line=EXO-22-006.

[28] CMS Collaboration, “Performance of the CMS Muon Trigger during Run 2 of the LHC”, CMS Note 2019/010, 2019.

[29] CMS Collaboration, “Muon recommendations for 2016 ultra legacy data and monte carlo”, 2022. https://twiki.cern.ch/twiki/bin/view/CMS/MuonUL2016#Trigger_efficiency_AN2.

[30] CMS Collaboration, “Muon recommendations for 2017 ultra legacy data and monte carlo”, 2022. https://twiki.cern.ch/twiki/bin/view/CMS/MuonUL2017#Trigger_efficiency_AN2.

[31] CMS Collaboration, “Muon recommendations for 2018 ultra legacy data and monte carlo”, 2022. https://twiki.cern.ch/twiki/bin/view/CMS/MuonUL2018#Trigger_efficiency_AN2.

[32] CMS Collaboration, “Missing transverse energy performance of the CMS detector”, *JINST* **6** (2011) P09001, doi:10.1088/1748-0221/6/09/P09001, arXiv:1106.5048.

[33] CMS Collaboration, “MET Filter Recommendations for Run II”, 2020. https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFiltersRun2.

[34] CMS Collaboration, “Particle-flow reconstruction and global event description with the cms detector”, *JINST* **12** (2017) P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.

[35] M. Cacciari and G. P. Salam, “Dispelling the N^3 myth for the k_T jet-finder”, *Phys. Lett. B* **641** (2006) 57, doi:10.1016/j.physletb.2006.08.037, arXiv:hep-ph/0512210.

[36] M. Cacciari, G. P. Salam, and G. Soyez, “The anti-k_T jet clustering algorithm”, *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[37] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet user manual”, *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[38] CMS Collaboration, “Jet Identification”, 2022. https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetID.

[39] E. Bols et al., “Jet Flavour Classification Using DeepJet”, arXiv:2008.10519.

[40] CMS Collaboration, “Performance of the DeepJet b tagging algorithm using 41.9/fb of data from proton-proton collisions at 13 TeV with Phase 1 CMS detector”, CMS Detector Performance Note CMS-DP-2018-058, 2018.

[41] CMS Collaboration, “Identification of b-quark jets with the CMS experiment”, *JINST* **8** (2013) P04013, doi:10.1088/1748-0221/8/04/P04013, arXiv:1211.4462.

[42] CMS Collaboration, “Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV”, *JINST* **13** (2018) P05011, doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.

[43] CMS Collaboration, "Recommendation for Using b-tag Objects in Physics Analyses", 2022. `https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation`.

[44] CMS Collaboration, "Recommended Jet Energy Corrections and Uncertainties For Data and MC", 2022. `https://twiki.cern.ch/twiki/bin/viewauth/CMS/JECDataMC`.

[45] CMS Collaboration, "Jet Energy Resolution", 2022. `https://twiki.cern.ch/twiki/bin/view/CMS/JetResolution`.

[46] CMS Collaboration, "Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector", *JINST* **14** (2019) P07004, `doi:10.1088/1748-0221/14/07/P07004`, `arXiv:1903.06078`.

[47] D. Bertolini, P. Harris, M. Low, and N. Tran, "Pileup per particle identification", *JHEP* **10** (2014) 059, `doi:10.1007/JHEP10(2014)059`, `arXiv:1407.6013`.

[48] CMS Collaboration, "Pileup mitigation at CMS in 13 TeV data", *JINST* **15** (2020), no. 09, P09018, `doi:10.1088/1748-0221/15/09/P09018`, `arXiv:2003.00503`.

[49] CMS Collaboration, "Measurement of the inclusive production cross sections for forward jets and for dijet events with one forward and one central jet in pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **06** (2012) 036, `doi:10.1007/JHEP06(2012)036`, `arXiv:1202.0704`.

[50] CMS Collaboration Collaboration, "Measurement of the inelastic pp cross section at $\sqrt{s} = 7$ TeV", technical report, CERN, Geneva, 2011.

[51] CMS Collaboration, "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV", *JINST* **13** (2018), no. 06, P06015, `doi:10.1088/1748-0221/13/06/P06015`, `arXiv:1804.04528`.

[52] CMS Collaboration, "Performance of high-pt muons collected with CMS at $\sqrt{s} = 13$ TeV during proton-proton collisions", CMS Note 2018/008, 2018.

[53] CMS Collaboration, "Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC", *JINST* **16** (2021) P05014, `doi:10.1088/1748-0221/16/05/P05014`, `arXiv:2012.06888`.

[54] CMS Collaboration, "ECAL 2016 refined calibration and Run2 summary plots", CMS Detector Performance Summary CMS-DP-2020-021, 2020.

[55] CMS Collaboration, "Cut based electron id for run 2", 2020. `https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedElectronIdentificationRun2`.

[56] Particle Data Group Collaboration, "Review of Particle Physics", *PTEP* **2020** (2020), no. 8, 083C01, `doi:10.1093/ptep/ptaa104`.

[57] M. J. Oreglia, "A study of the reactions $\psi' \to \gamma\gamma\psi$". PhD thesis, Stanford University, 1980. SLAC Report SLAC-R-236, see Appendix D.

[58] J. E. Gaiser, "Charmonium spectroscopy from radiative decays of the J/ψ and ψ'". PhD thesis, Stanford University, 1982. SLAC Report SLAC-R-255.

[59] R. A. Fisher, "On the interpretation of χ^2 from contingency tables, and the calculation of P", *J. R. Stat. Soc* **85** (1922) 87, doi:10.1098/rsta.1922.0009.

[60] P. D. Dauncey, M. Kenzie, N. Wardle, and G. J. Davies, "Handling uncertainties in background shapes: the discrete profiling method", *JINST* **10** (2015) P04015, doi:10.1088/1748-0221/10/04/P04015, arXiv:1408.6865.

[61] C. M., "Inclusion of systematic uncertainties in upper limits and hypothesis tests",.

[62] http://hep.ucsb.edu/people/claudio/Phys250/Poisson-notebooks/ComparisonOfLimits.pdf.

[63] CMS Collaboration, "Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS", *Submitted to EPJ C* (2021) arXiv:2104.01927.

[64] CMS Collaboration, "CMS luminosity measurement for the 2017 data taking period at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-LUM-17-004, 2018.

[65] CMS Collaboration Collaboration, "CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV", Technical Report CMS-PAS-LUM-18-002, CERN, Geneva, 2019.

[66] A. L. Read, "Presentation of search results: The $\mathrm{CL_s}$ technique", *J. Phys. G* **28** (2002) 2693, doi:10.1088/0954-3899/28/10/313.

[67] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth. A* **434** (1999) 435, doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006.

[68] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727. [Erratum: doi:10.1140/epjc/s10052-013-2501-z].

[69] ATLAS and CMS Collaborations, "Procedure for the LHC Higgs boson search combination in summer 2011", ATL-PHYS-PUB-2011-011, CMS NOTE-2011/005, 2011.

[70] CMS Collaboration, "Reference muon id, isolation and trigger efficiencies for 2016 legacy re-reco data", 2020. https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffs2016LegacyRereco.

[71] CMS Collaboration, "Reference muon id, isolation and trigger efficiencies for 2017 data", 2019. https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffs2017.

[72] CMS Collaboration, "Reference muon id, isolation and trigger efficiencies for 2018 data", 2020. https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffs2018.

[73] CMS Collaboration, "Reweighting recipe to emulate Level 1 ECAL and Muon prefiring", 2022. https://twiki.cern.ch/twiki/bin/viewauth/CMS/L1PrefiringWeightRecipe.

[74] M. Masciovecchio, "Zprime to mumu + btag(s): update". https://tinyurl.com/muvcefn5.

A DY-enriched region

To validate the muon kinematic distributions in the simulation, a region enriched in the DY process is constructed. This region, referred to as 'DY-enriched region', is orthogonal to the search regions of this analysis by requiring that the number of medium b-tagged jets is equal to zero. The p_T, η and absolute isolation variable of the leading and subleading muon of the analysis are shown in Figs. 21 and 22 respectively. The agreement between data and simulation is remarkable over the full range for all the variables.

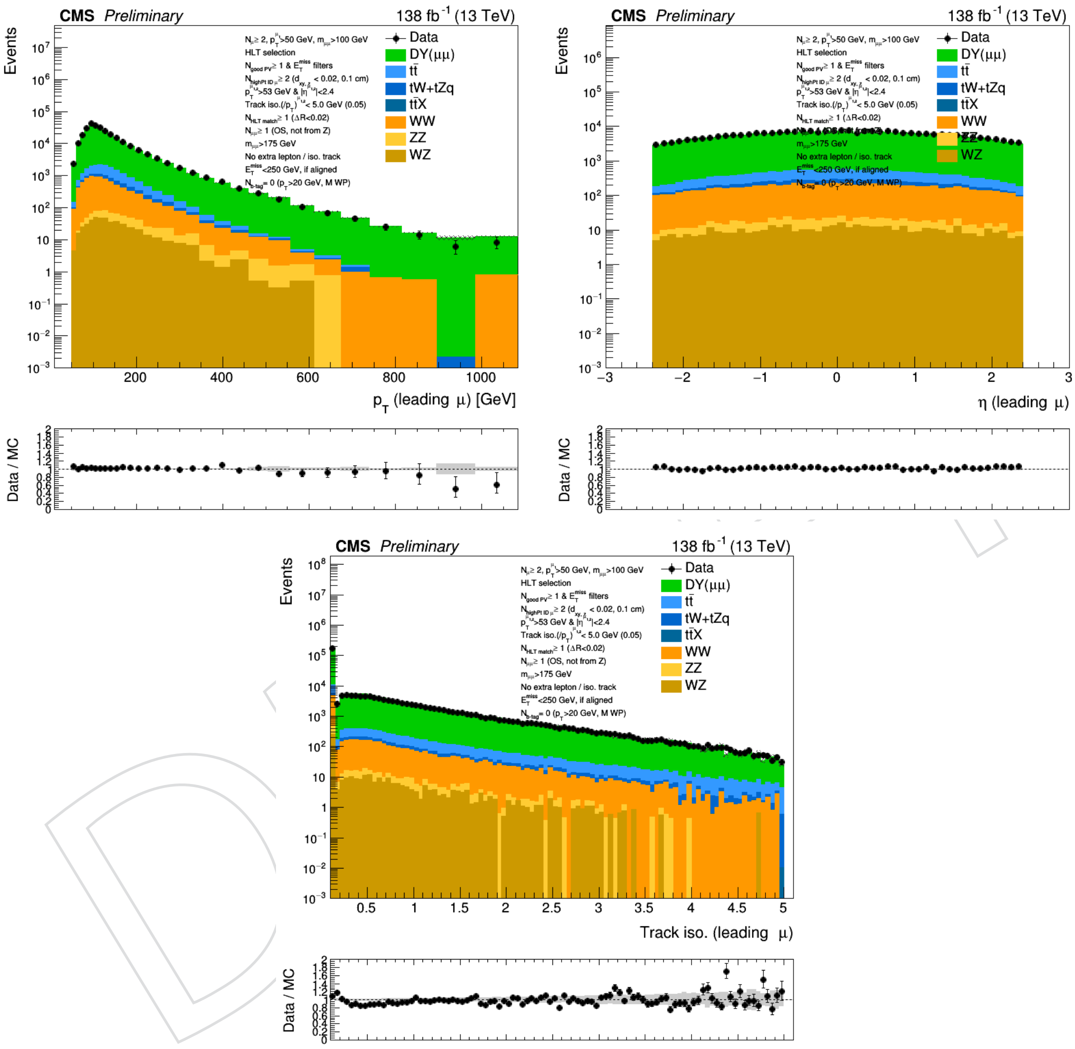


Figure 21: Distributions of p_T (upper left), η (upper right) and absolute isolation (lower) of the leading muon in the DY-enriched region.

Figure 23 shows the dimuon mass spectrum from 175 to 3000 GeV, for which the simulation is also found in excellent agreement with the data.

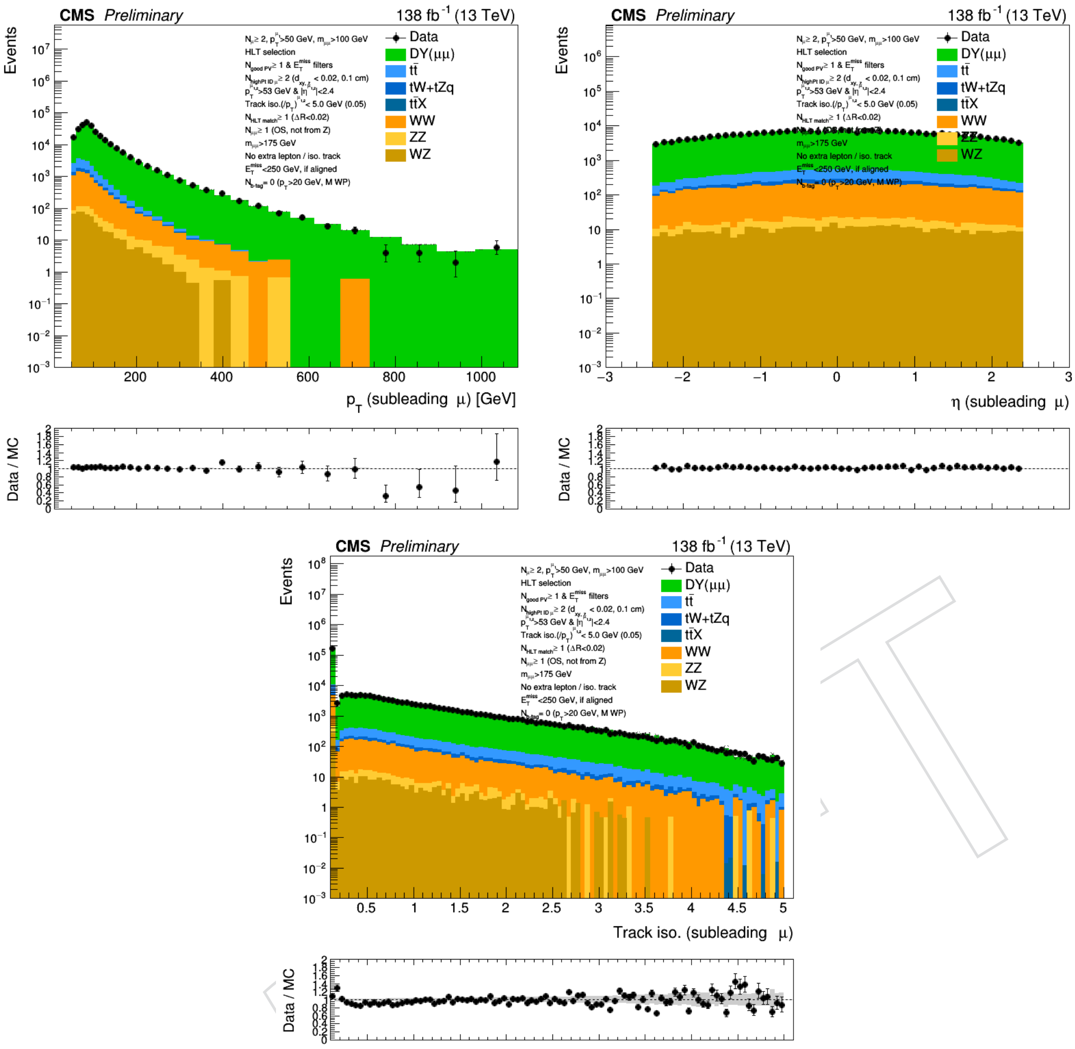


Figure 22: Distributions of p_{T} (upper left), η (upper right) and absolute isolation (lower) of the subleading muon in the DY-enriched region.

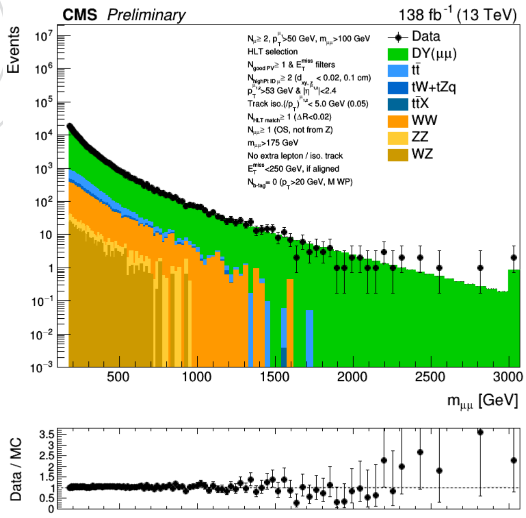


Figure 23: Distribution of dimuon mass spectrum, $m_{\mu\mu}$, in the DY-enriched region.

Finally, Fig. 24 shows the PUPPI $p_{\mathrm{T}}^{\mathrm{miss}}$ magnitude (left) and ϕ (right) in the DY-enriched region. The data and simulation are found in agreement, overall, for both distrbutions. Five data events are observed over the simulation at very large PUPPI $p_{\mathrm{T}}^{\mathrm{miss}}$, larger than 550 GeV. These events were explictly inspected to understand their origin. They have similarly high values of PF $p_{\mathrm{T}}^{\mathrm{miss}}$, which, in four out of the five cases, originates from a jet anti-aligned to the $p_{\mathrm{T}}^{\mathrm{miss}}$ in the transverse plane. In the search region of the analysis, a requirement is applied to reject such events with $p_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV when it is (anti-)aligned with one of the b-tagged jets. This requirement is (expected to be) ineffective in th DY-enriched region, due to the absenc of b-tagged jets. The dimuon invariant mass in these events was also calculated and, in all of the cases, was found to be below 265 GeV, i.e., in a region of the dimuon mass spectrum where five events do not affect the level of agreement that we are trying to establish. Finally, it was verified that the high values of $p_{\mathrm{T}}^{\mathrm{miss}}$ are not connected to any muon misreconstruction related to the 'tuneP' algorithm, as the ratio of p_{T}^{μ} values from this algorithm and from default muon reconstruction is compatible with unity within 1%.

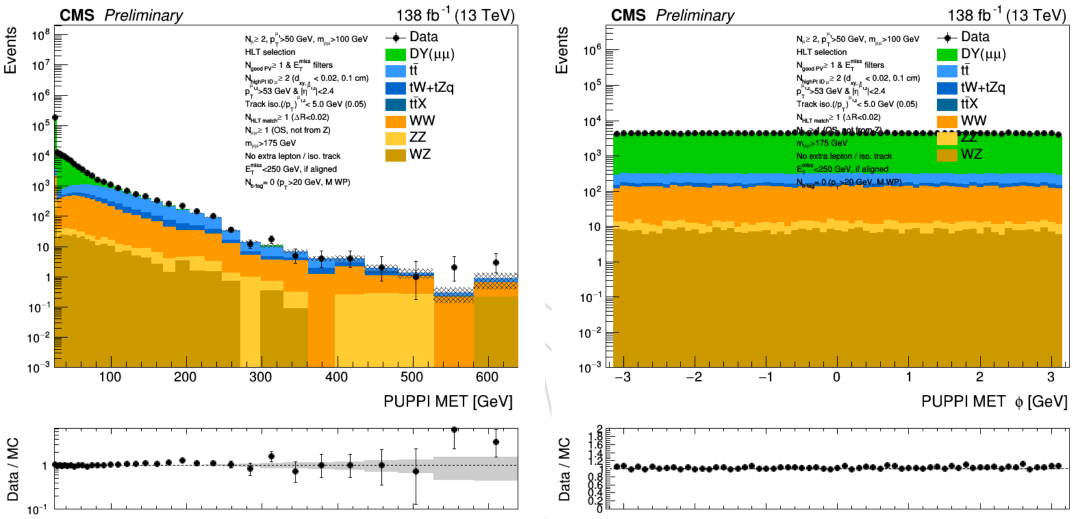


Figure 24: Distributions of $p_{\mathrm{T}}^{\mathrm{miss}}$ magnitude (left) and ϕ (right) in the DY-enriched region.

In the following, the same distributions are shown separately for each year of data taking.

A.1 2016 data taking

The p_T, η and absolute isolation variable of the leading and subleading muon of the analysis are shown in Figs. 25 and 26 respectively, for 2016 data and simulation. The agreement between data and simulation is remarkable over the full range for all the variables.

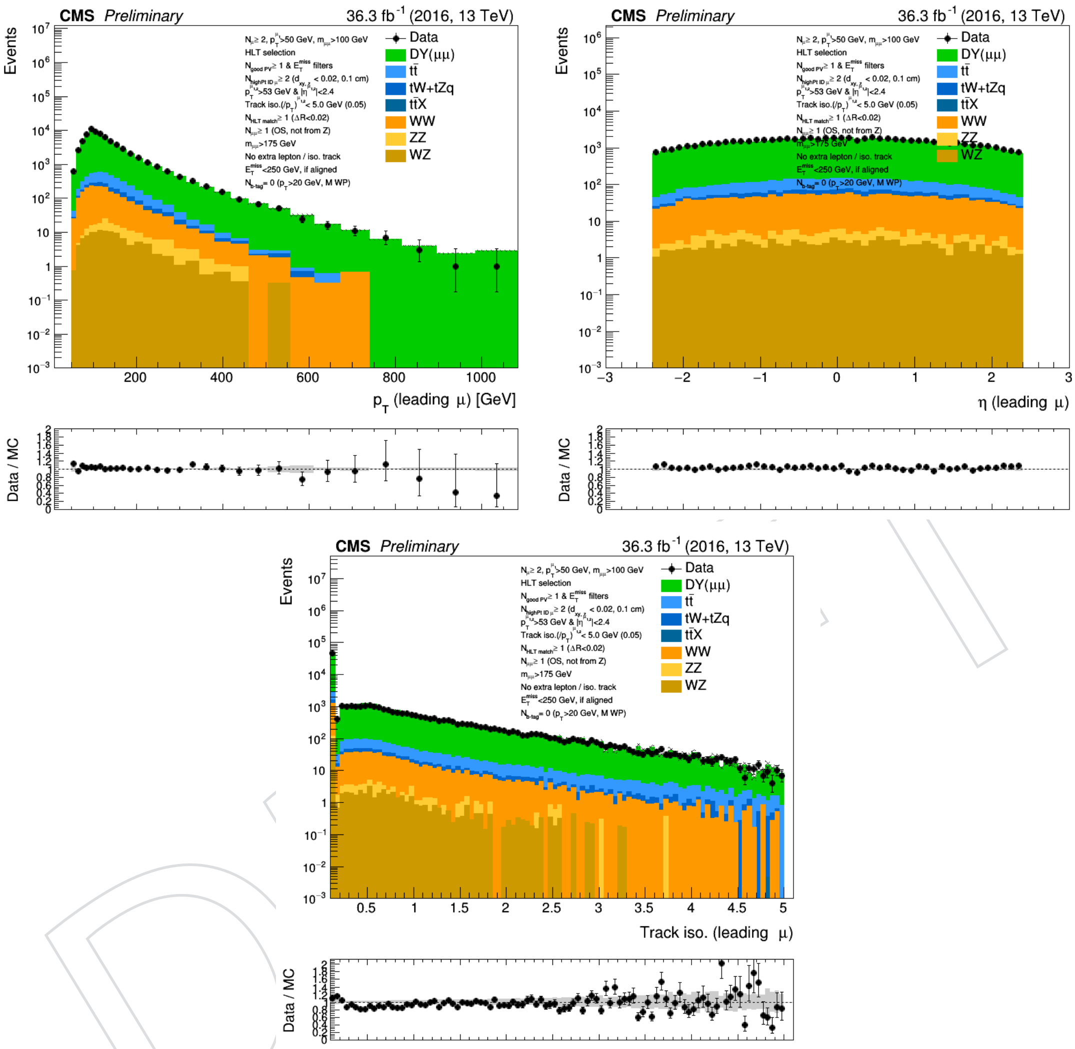


Figure 25: Distributions of p_T (upper left), η (upper right) and absolute isolation (lower) of the leading muon in the DY-enriched region, for 2016 data and simulation.

Figure 27 shows the dimuon mass spectrum as obtained from 2016 data and simulation, from 175 to 3000 GeV, for which the simulation is also found in excellent agreement with the data.

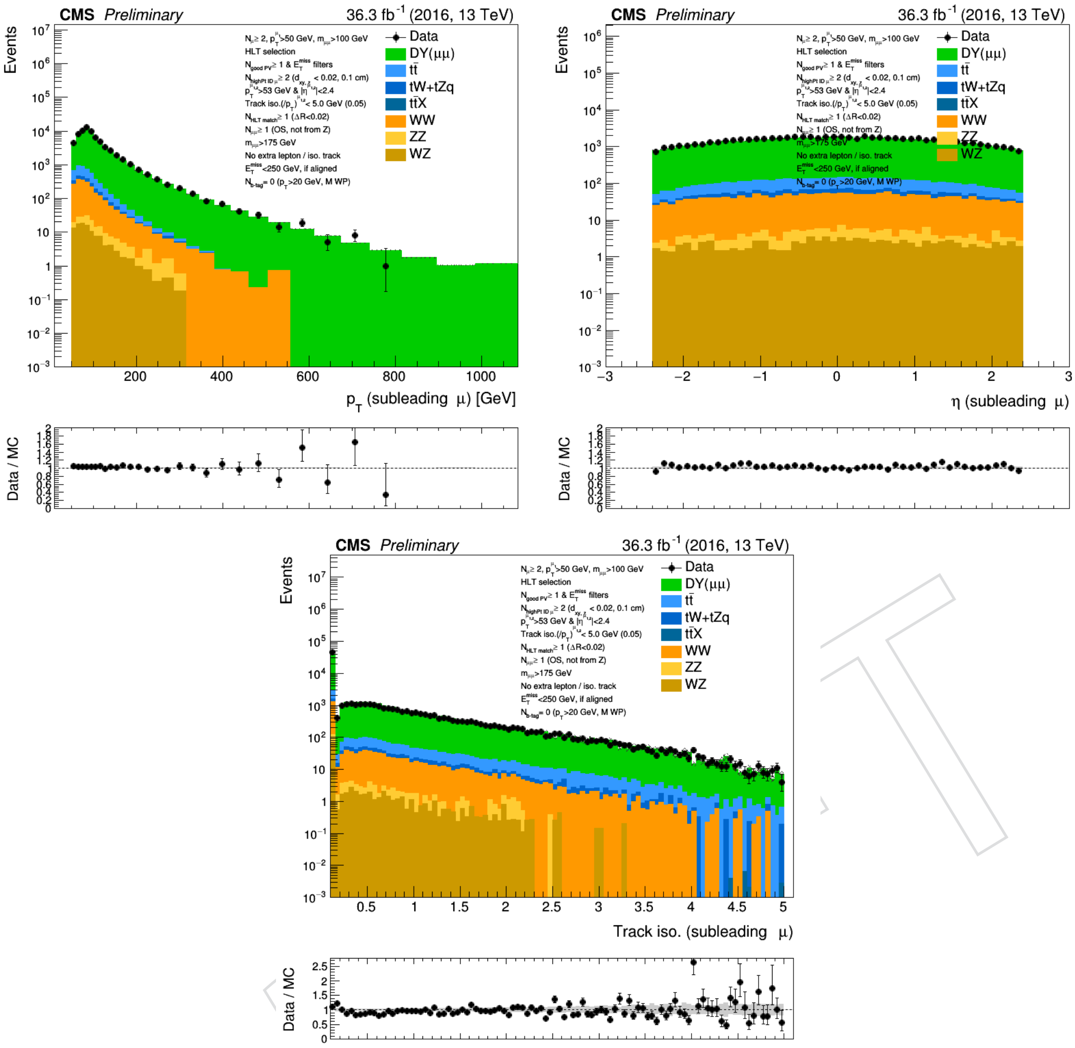


Figure 26: Distributions of p_{T} (upper left), η (upper right) and absolute isolation (lower) of the subleading muon in the DY-enriched region, for 2016 data and simulation.

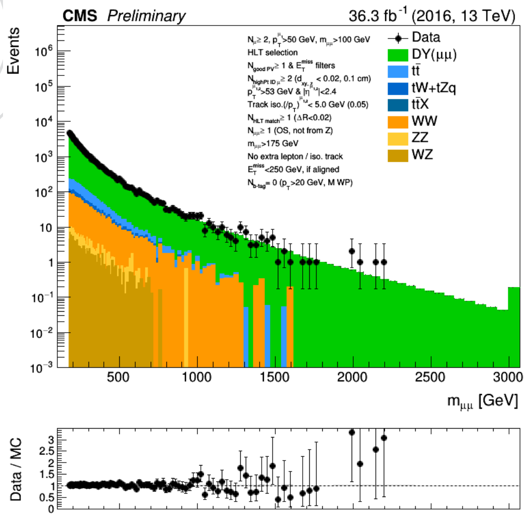


Figure 27: Distribution of dimuon mass spectrum, $m_{\mu\mu}$, in the DY-enriched region, for 2016 data and simulation.

Finally, Fig. 28 shows the PUPPI $p_{\mathrm{T}}^{\mathrm{miss}}$ magnitude (left) and ϕ (right) in the DY-enriched region, for 2016 data and simulation. The data and simulation are found in agreement, overall, for both distrbutions.

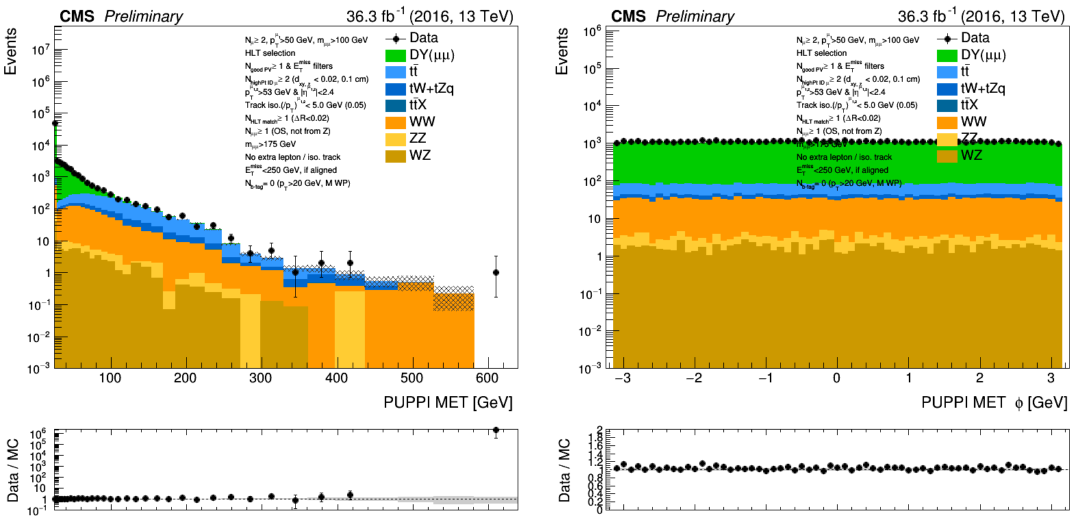


Figure 28: Distributions of $p_{\mathrm{T}}^{\mathrm{miss}}$ magnitude (left) and ϕ (right) in the DY-enriched region, for 2016 data and simulation.

A.2 2017 data taking

The p_{T}, η and absolute isolation variable of the leading and subleading muon of the analysis are shown in Figs. 29 and 30 respectively, for 2017 data and simulation. The agreement between data and simulation is remarkable over the full range for all the variables.

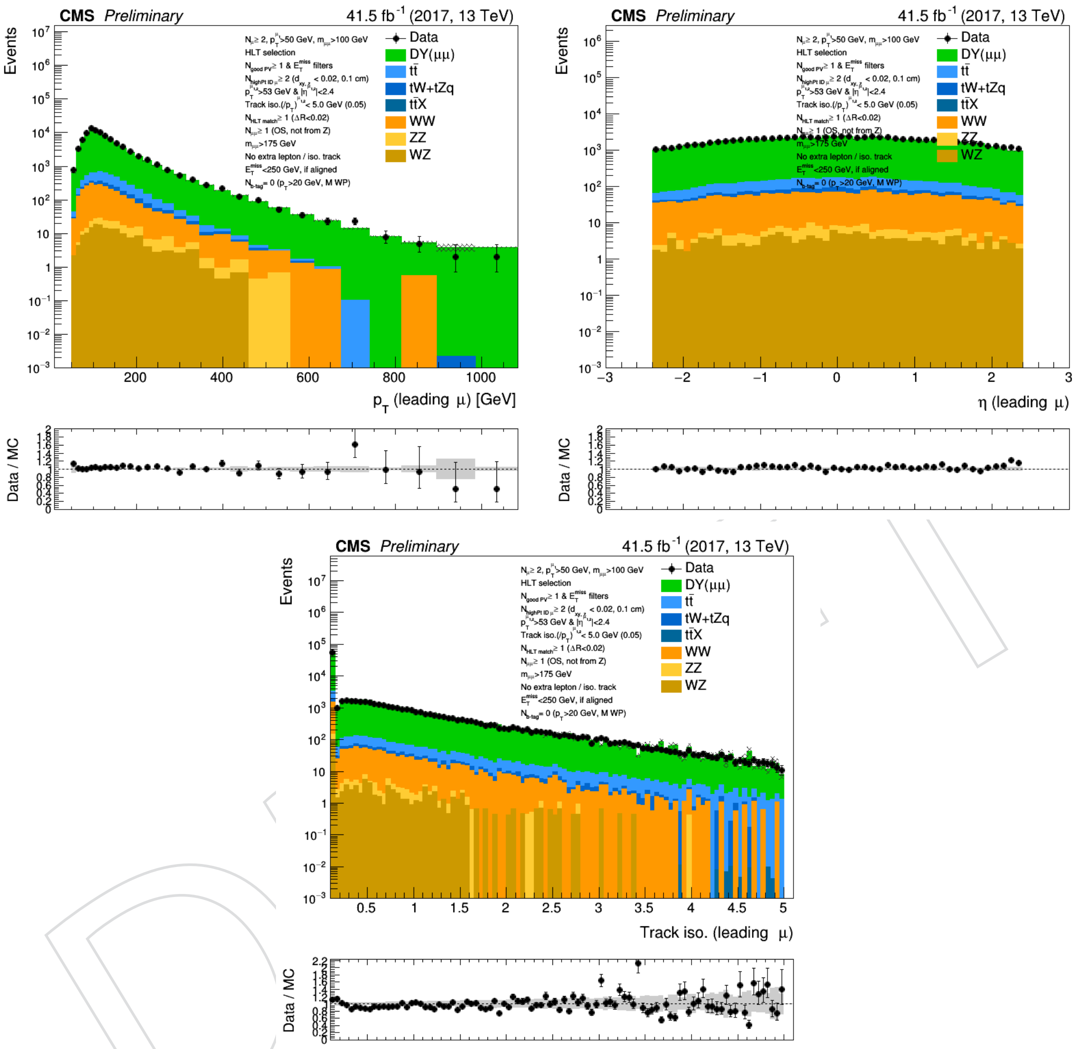


Figure 29: Distributions of p_{T} (upper left), η (upper right) and absolute isolation (lower) of the leading muon in the DY-enriched region, for 2017 data and simulation.

Figure 31 shows the dimuon mass spectrum as obtained from 2017 data and simulation, from 175 to 3000 GeV, for which the simulation is also found in excellent agreement with the data.

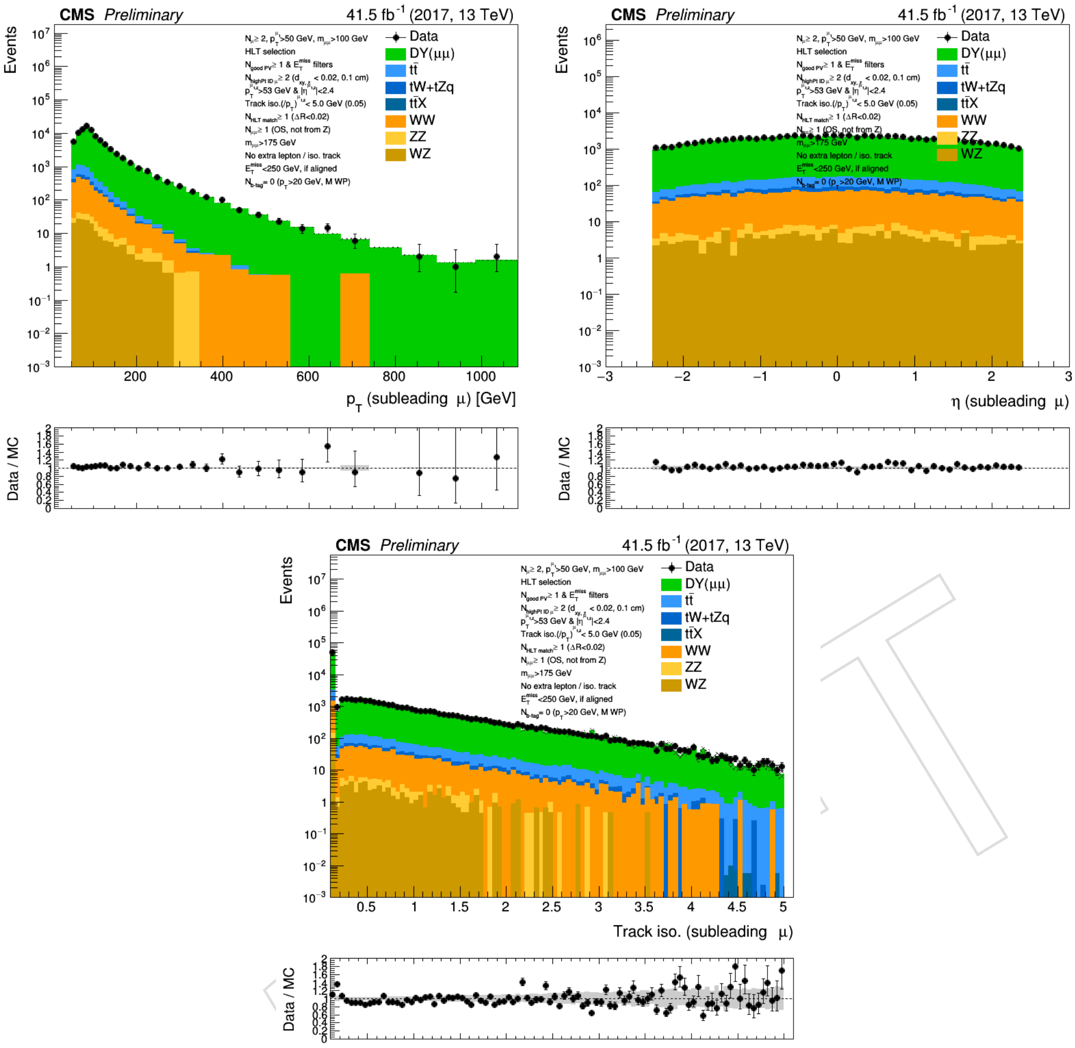


Figure 30: Distributions of p_{T} (upper left), η (upper right) and absolute isolation (lower) of the subleading muon in the DY-enriched region, for 2017 data and simulation.

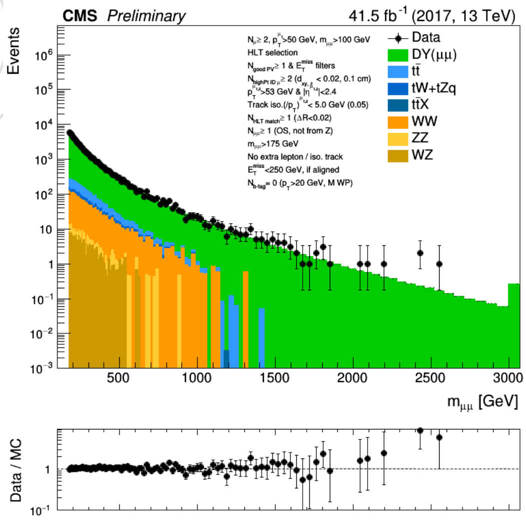


Figure 31: Distribution of dimuon mass spectrum, $m_{\mu\mu}$, in the DY-enriched region, for 2017 data and simulation.

Finally, Fig. 32 shows the PUPPI p_T^{miss} magnitude (left) and ϕ (right) in the DY-enriched region, for 2017 data and simulation. The data and simulation are found in agreement, overall, for both distrbutions.

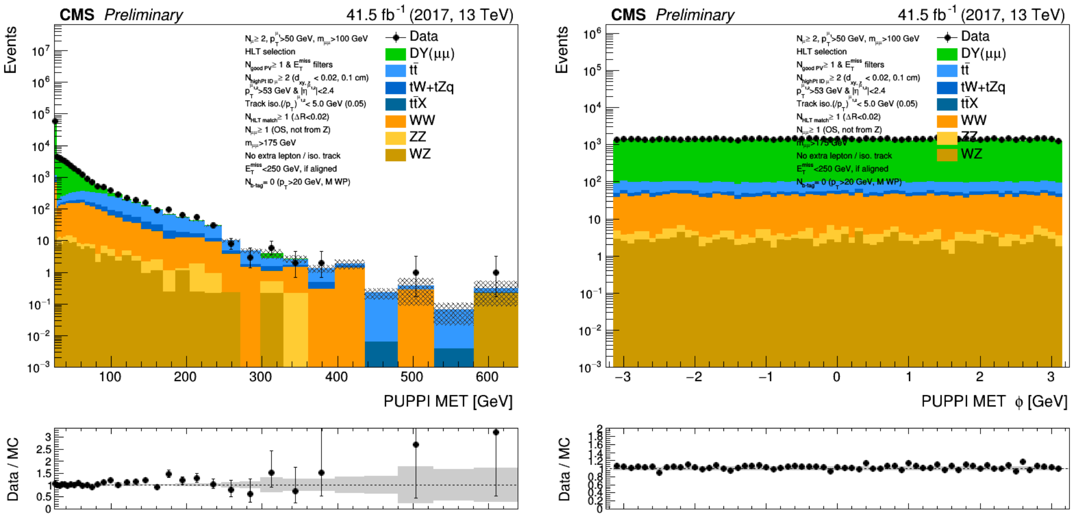


Figure 32: Distributions of p_T^{miss} magnitude (left) and ϕ (right) in the DY-enriched region, for 2017 data and simulation.

A.3 2018 data taking

The p_T, η and absolute isolation variable of the leading and subleading muon of the analysis are shown in Figs. 33 and 34 respectively, for 2018 data and simulation. The agreement between data and simulation is remarkable over the full range for all the variables.

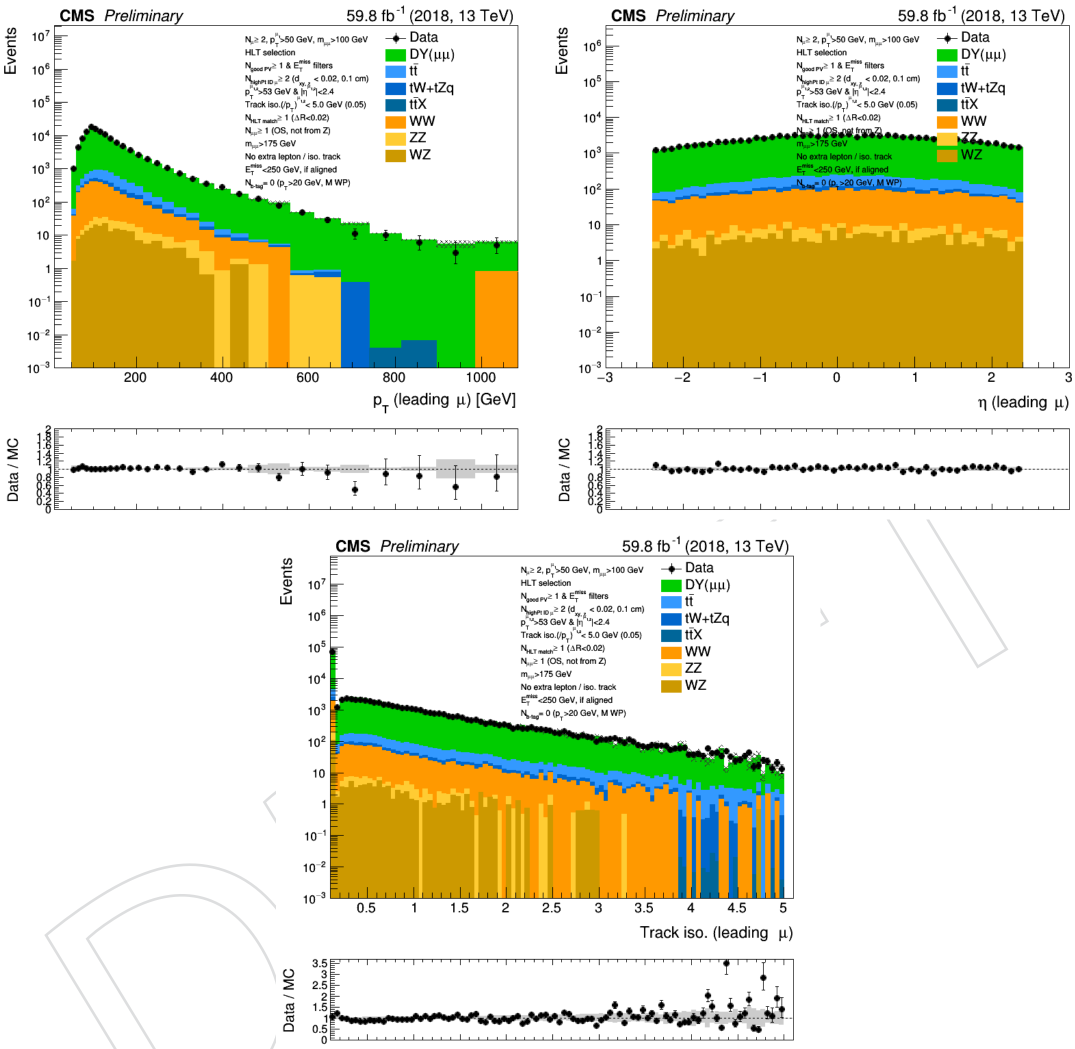


Figure 33: Distributions of p_T (upper left), η (upper right) and absolute isolation (lower) of the leading muon in the DY-enriched region, for 2018 data and simulation.

Figure 35 shows the dimuon mass spectrum as obtained from 2018 data and simulation, from 185 to 3000 GeV, for which the simulation is also found in excellent agreement with the data.

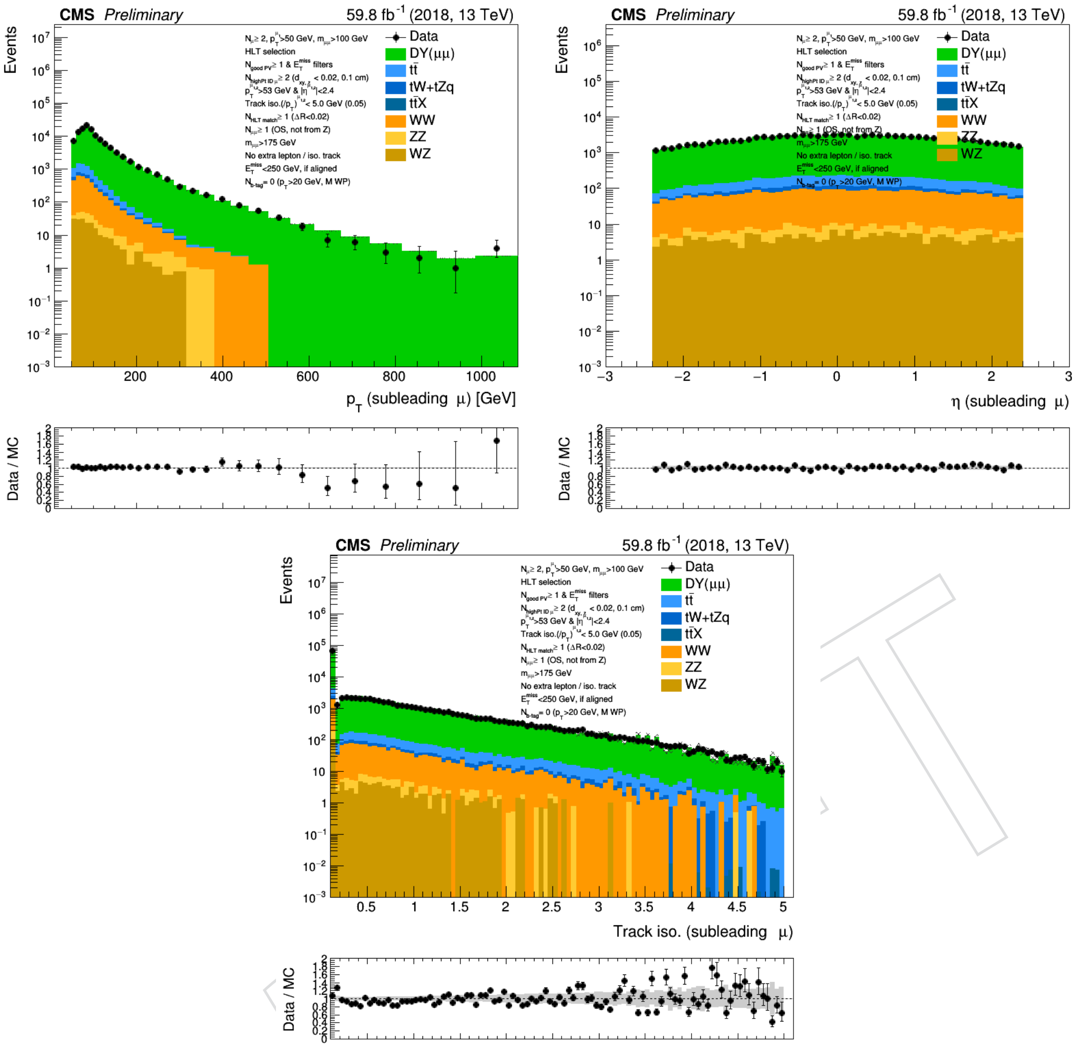


Figure 34: Distributions of p_T (upper left), η (upper right) and absolute isolation (lower) of the subleading muon in the DY-enriched region, for 2018 data and simulation.

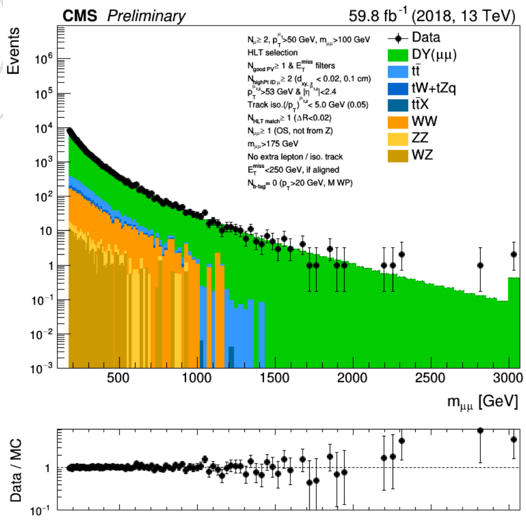


Figure 35: Distribution of dimuon mass spectrum, $m_{\mu\mu}$, in the DY-enriched region, for 2018 data and simulation.

Finally, Fig. 36 shows the PUPPI $p_{\mathrm{T}}^{\mathrm{miss}}$ magnitude (left) and ϕ (right) in the DY-enriched region, for 2018 data and simulation. The data and simulation are found in agreement, overall, for both distrbutions.

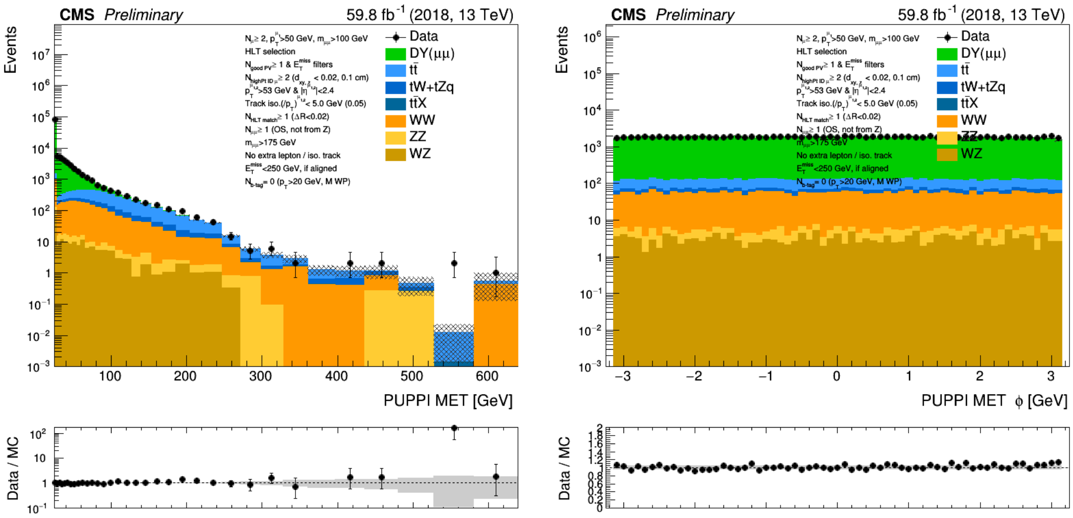


Figure 36: Distributions of $p_{\mathrm{T}}^{\mathrm{miss}}$ magnitude (left) and ϕ (right) in the DY-enriched region, for 2018 data and simulation.

B $t\bar{t}$-enriched region

To validate the muon kinematic distributions in the simulation, a region enriched in the $t\bar{t}$ process is constructed. This region, referred to as '$t\bar{t}$-enriched region', is orthogonal to the search regions of this analysis by inverting the selection requirement on $m_{\mu b}^{min}$, i.e., by requiring $m_{\mu b}^{min} < 175\,GeV$. The $t\bar{t}$-enriched region is fully unblinded for dimuon masses greater than 350 GeV to minimize any potential overlap with the EXO-22-006 analysis.

Since this validation region also includes b-tagged jets, Fig. 37 shows the values of N_b, in which excellent agreement between data and simulation is observed. The distributions of some basic variables for the selected muons, b-tagged jets and p_T^{miss} in this region are first presented inclusively in terms of N_b (Figs. 38–43), and then in the analysis categories based on N_b (Sections B.1 and B.2).

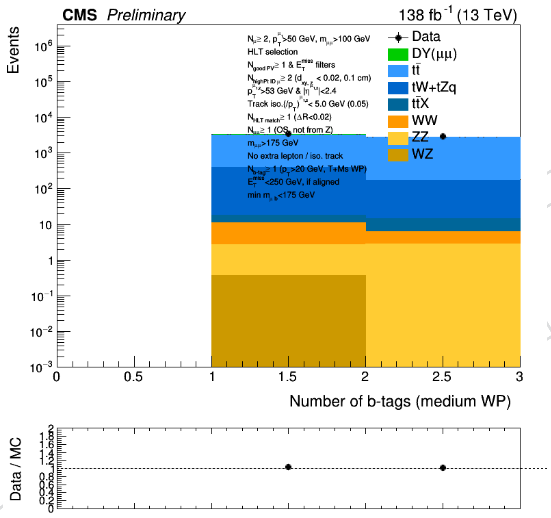


Figure 37: Distribution of the number of selected b-tagged jets, N_b, in the $t\bar{t}$-enriched region.

The p_T, η and absolute isolation variable of the leading and subleading muon of the analysis are shown in Figs. 38 and 39 respectively. Figs. 40 and 40 show the p_T and η variables of the leading and subleading b-tagged jet respectively. The agreement between data and simulation is remarkable over the full range for all the variables.

Figure 42 shows the dimuon mass spectrum, for which the simulation is also found in excellent agreement with the data.

Finally, Fig. 43 shows the PUPPI p_T^{miss} magnitude (left) and ϕ (right) in the $t\bar{t}$-enriched region. The data and simulation are found in agreement, overall, for both distrbutions.

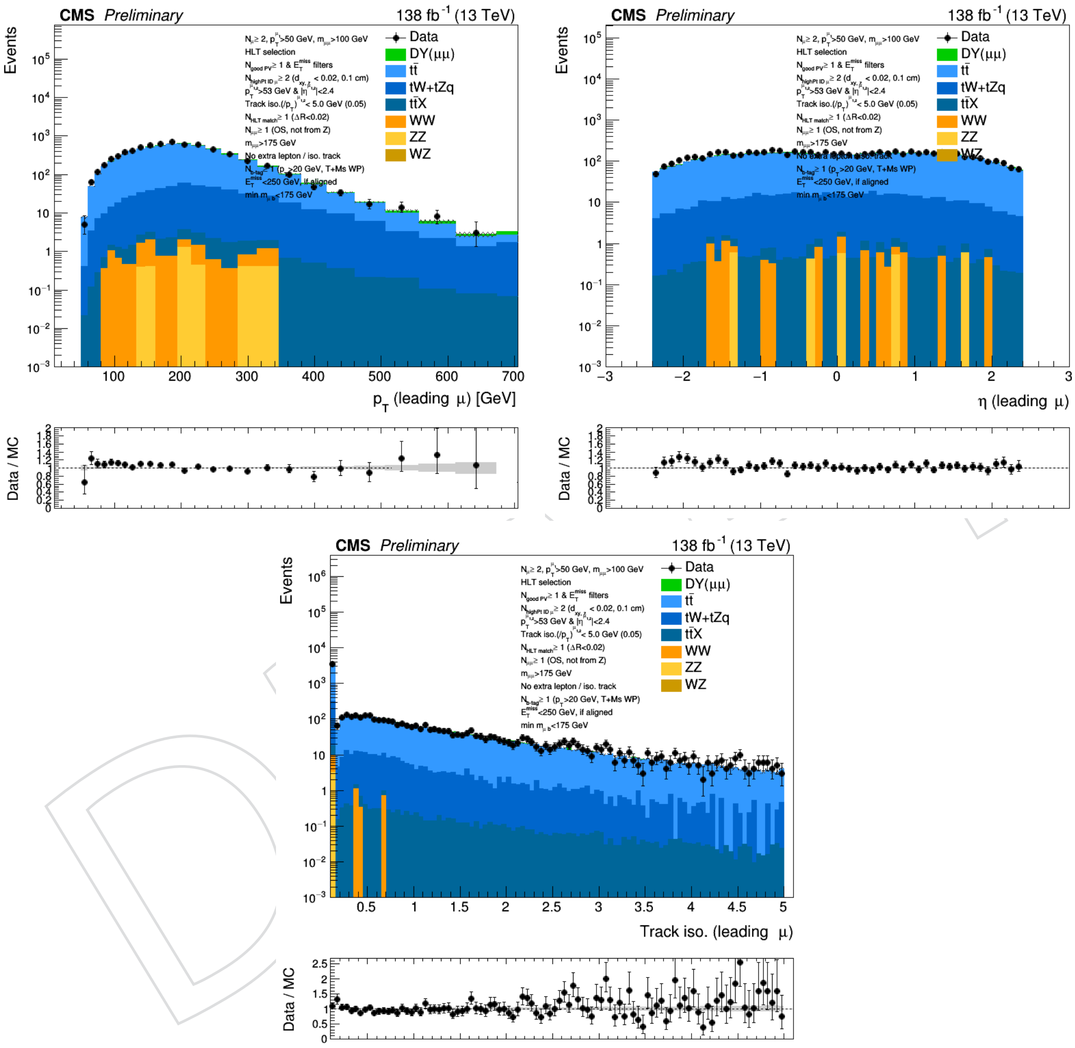


Figure 38: Distributions of p_{T} (upper left), η (upper right) and absolute isolation (lower) of the leading muon in the $\mathrm{t\bar{t}}$-enriched region.

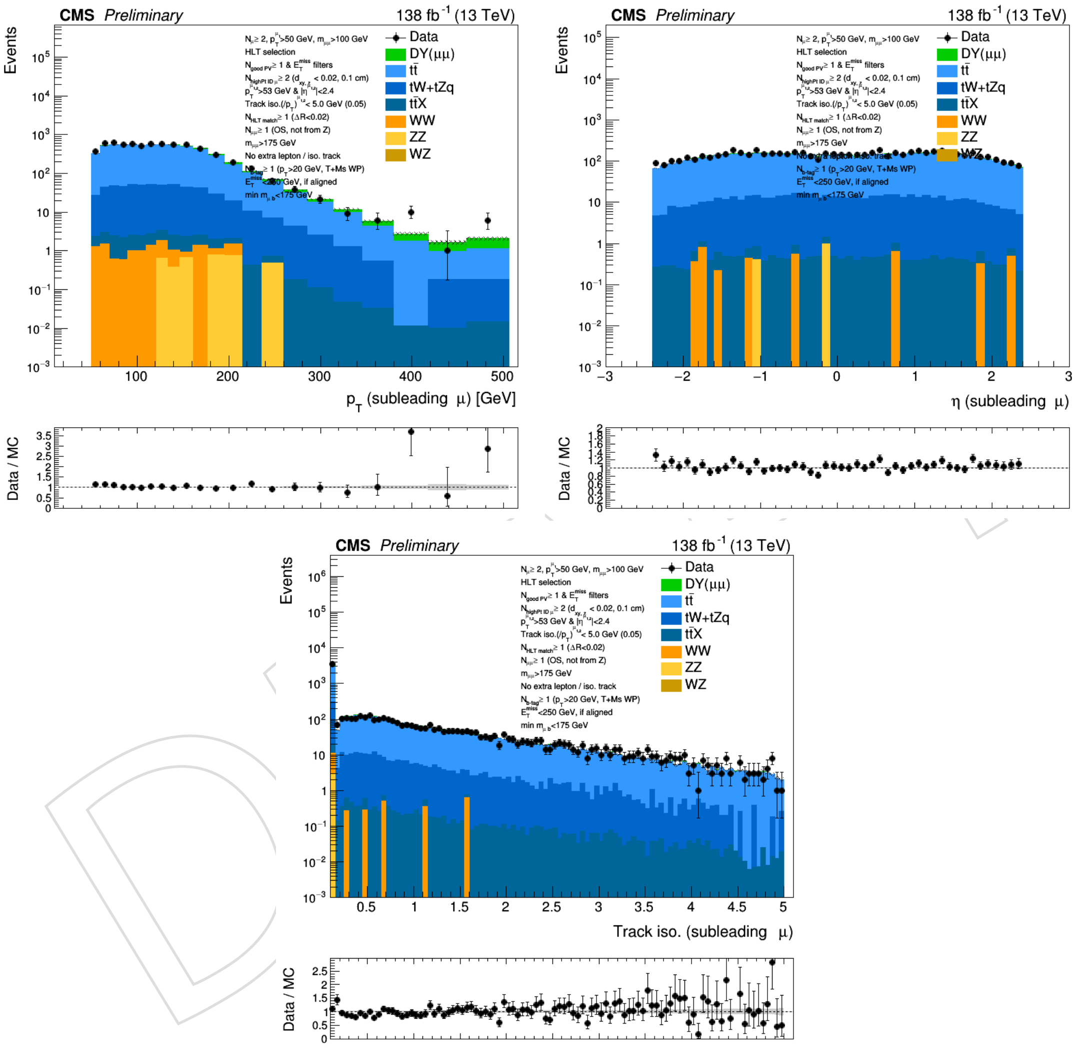


Figure 39: Distributions of p_T (upper left), η (upper right) and absolute isolation (lower) of the subleading muon in the $t\bar{t}$-enriched region.

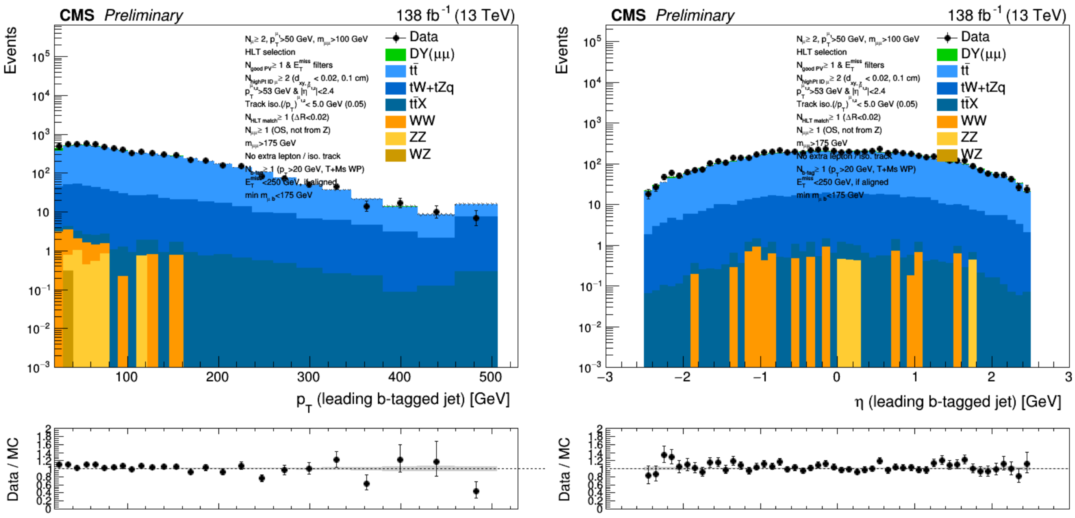


Figure 40: Distributions of p_T (upper left), η (upper right) and absolute isolation (lower) of the leading b-tagged jet in the $t\bar{t}$-enriched region.

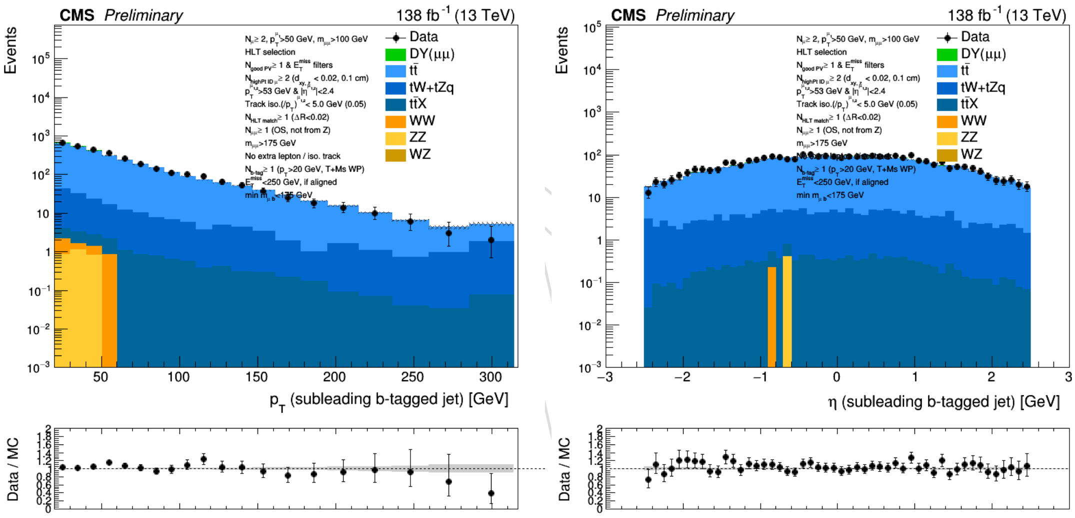


Figure 41: Distributions of p_T (upper left), η (upper right) and absolute isolation (lower) of the subleading b-tagged jet in the $t\bar{t}$-enriched region.

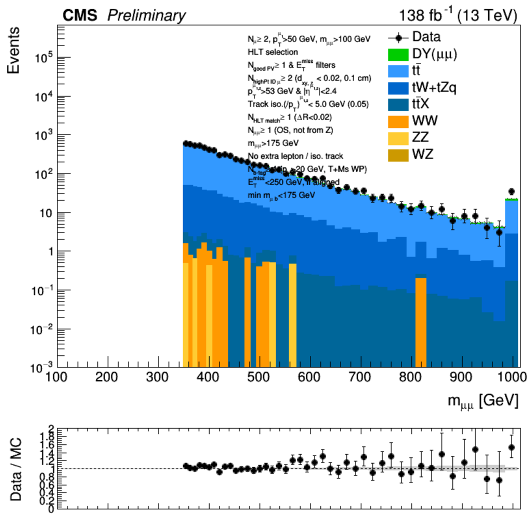


Figure 42: Distribution of dimuon mass spectrum, $m_{\mu\mu}$, in the $t\bar{t}$-enriched region.

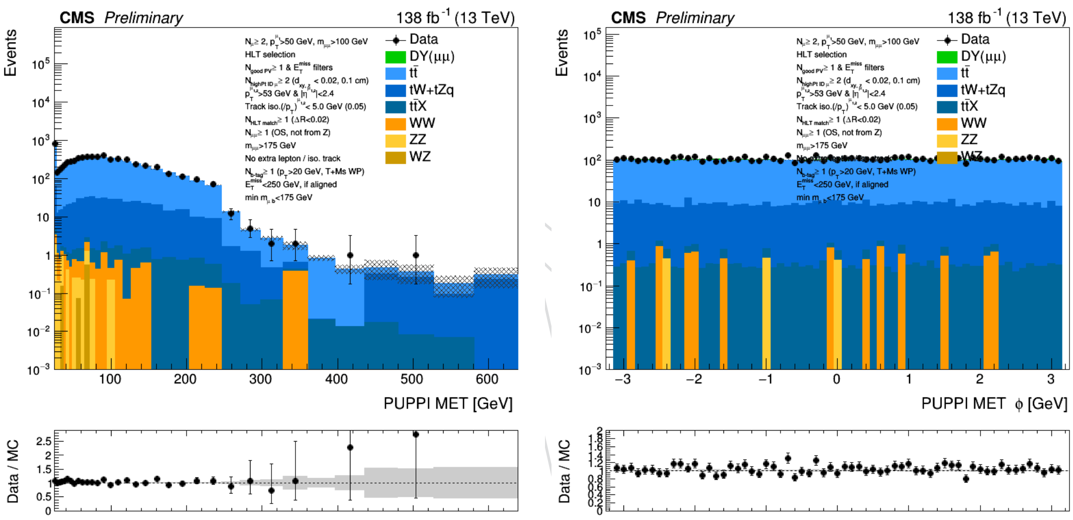


Figure 43: Distributions of $p_{\mathrm{T}}^{\mathrm{miss}}$ magnitude (left) and ϕ (right) in the $t\bar{t}$-enriched region.

B.1 $N_b = 1$ category

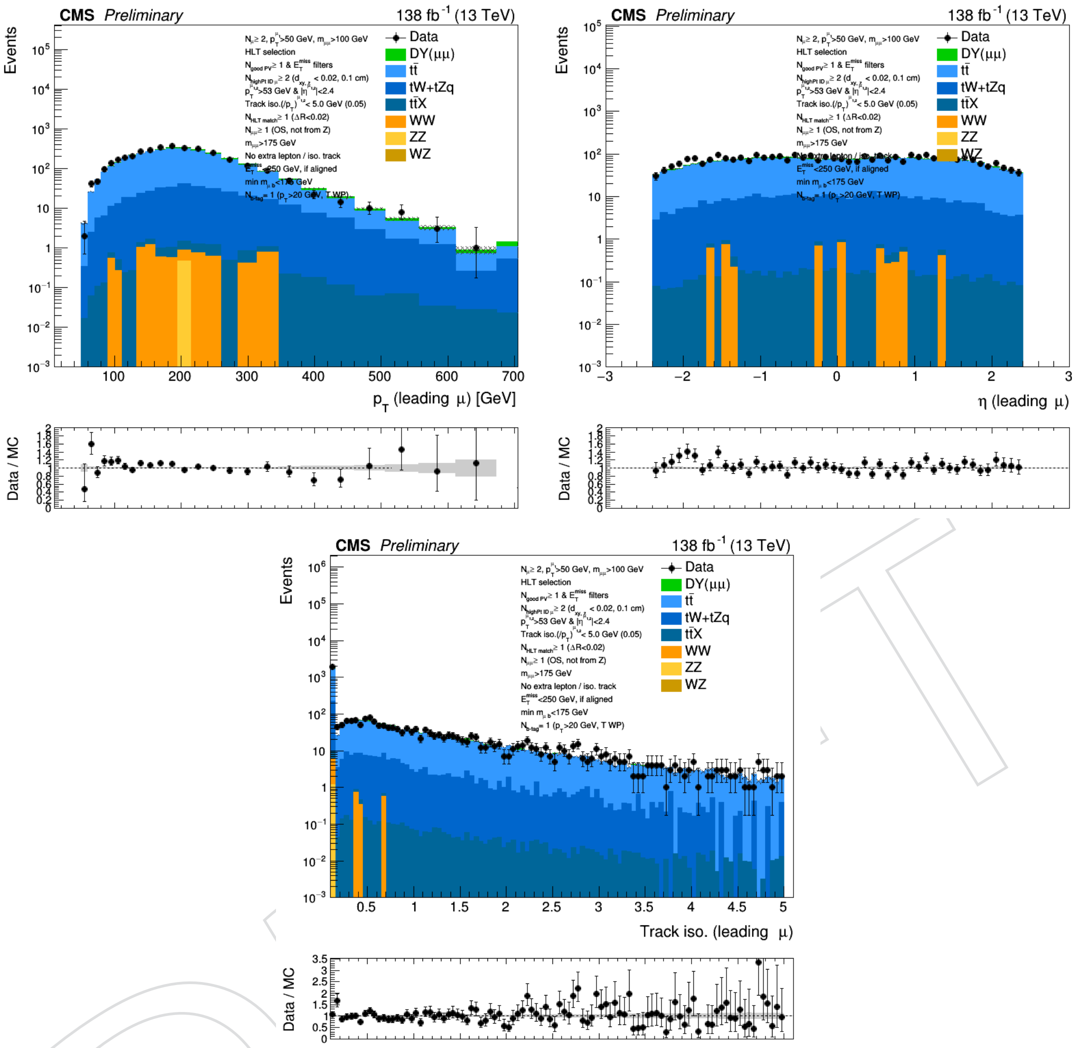


Figure 44: Distributions of p_T (upper left), η (upper right) and absolute isolation (lower) of the leading muon in the $t\bar{t}$-enriched region for $N_b = 1$.

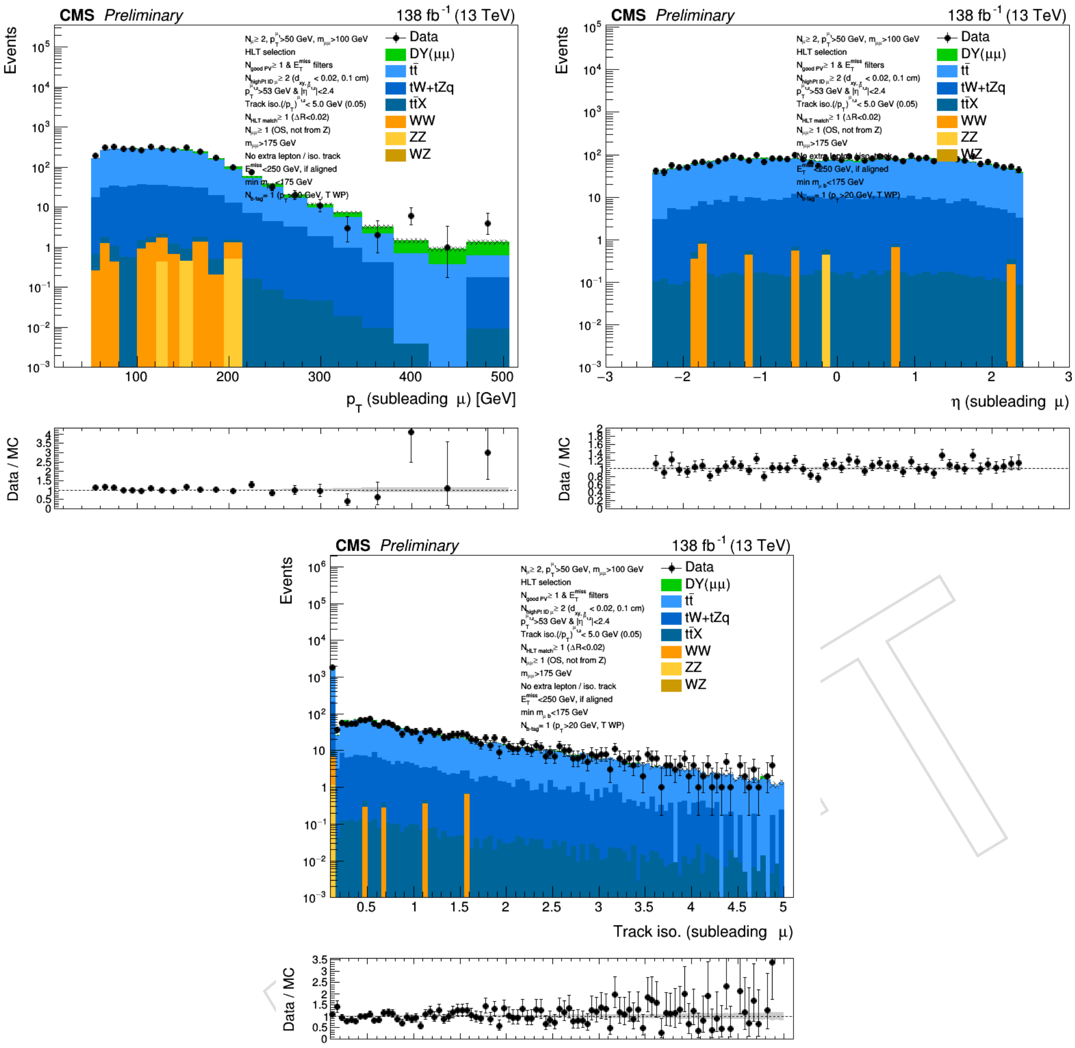


Figure 45: Distributions of p_T (upper left), η (upper right) and absolute isolation (lower) of the subleading muon in the $t\bar{t}$-enriched region for $N_b = 1$.

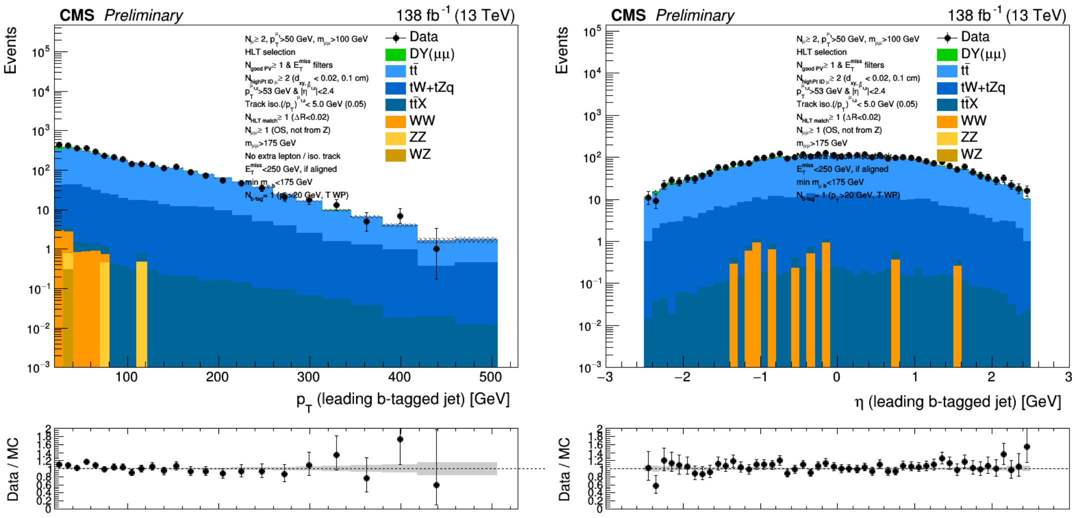


Figure 46: Distributions of p_T (upper left), η (upper right) and absolute isolation (lower) of the leading b-tagged jet in the $t\bar{t}$-enriched region for $N_b = 1$.

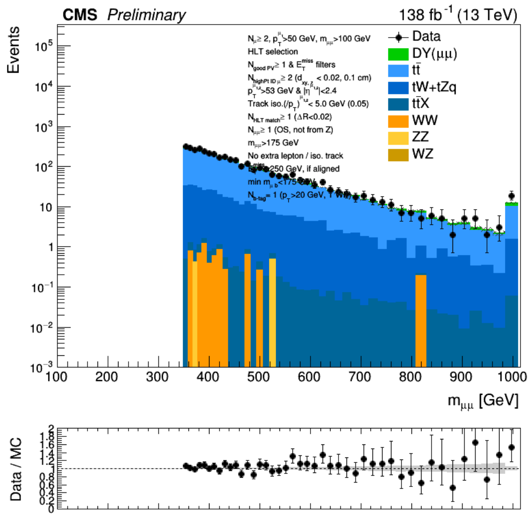


Figure 47: Distribution of dimuon mass spectrum, $m_{\mu\mu}$, in the $\mathrm{t\bar{t}}$-enriched region for $N_\mathrm{b} = 1$.

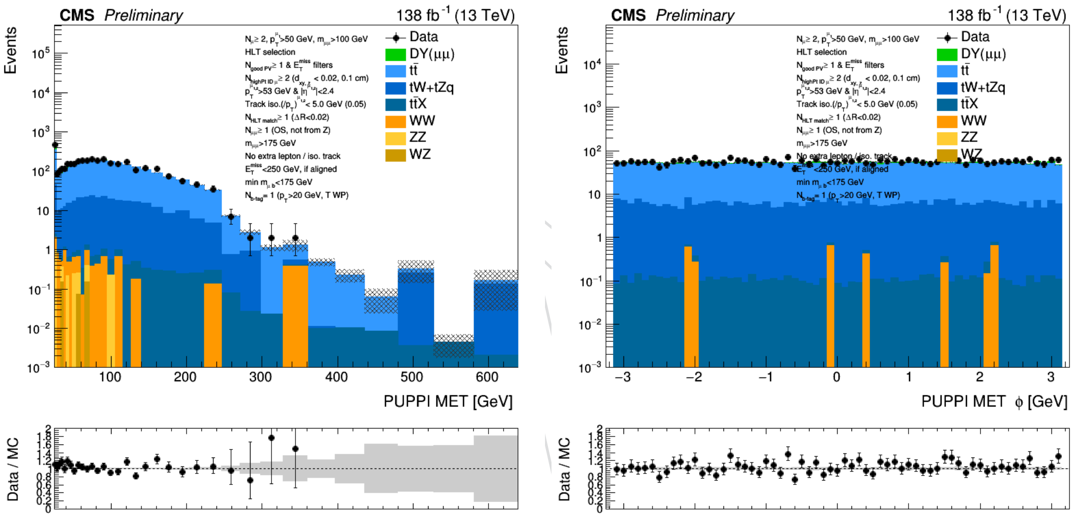


Figure 48: Distributions of $p_\mathrm{T}^\mathrm{miss}$ magnitude (left) and ϕ (right) in the $\mathrm{t\bar{t}}$-enriched region for $N_\mathrm{b} = 1$.

B.2 $N_b > 1$ category

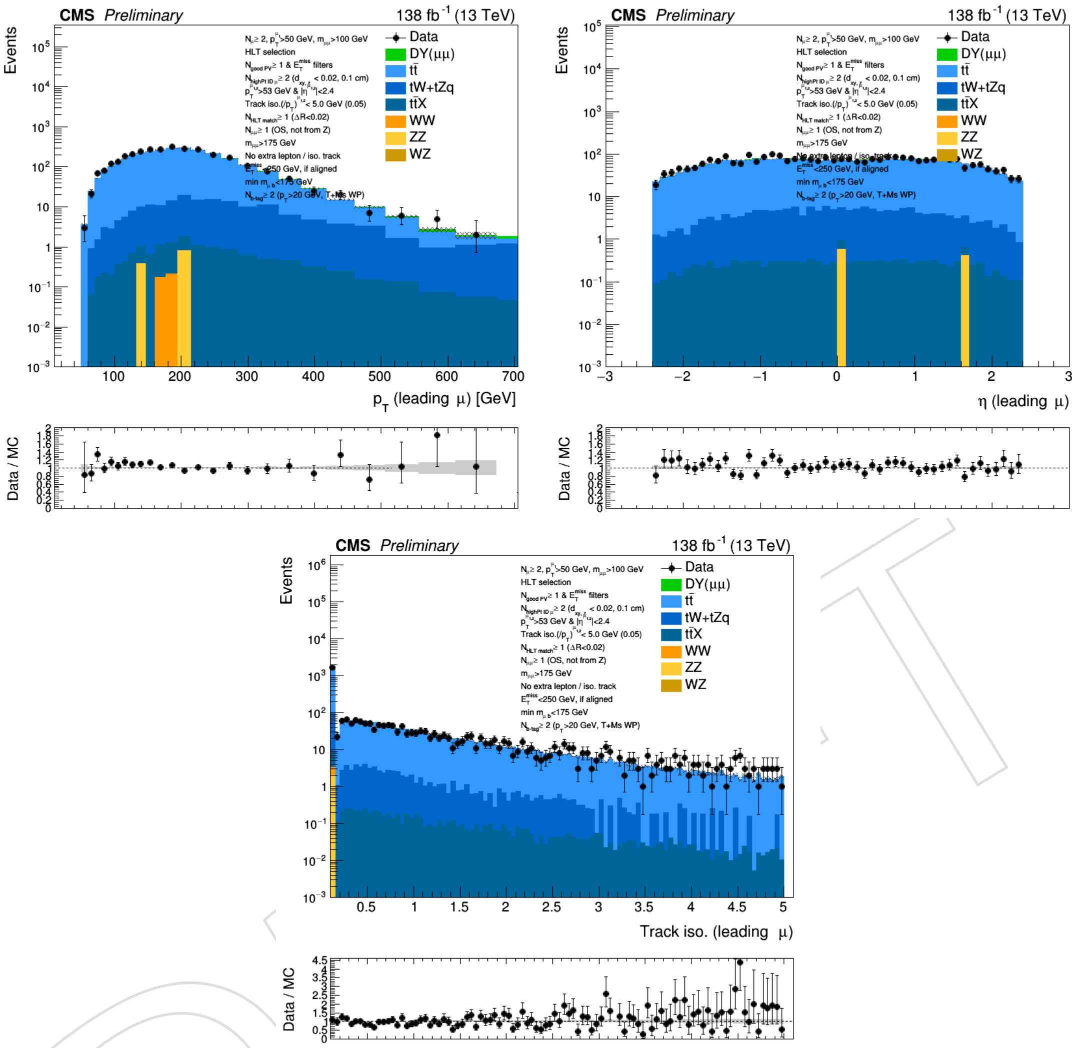


Figure 49: Distributions of p_T (upper left), η (upper right) and absolute isolation (lower) of the leading muon in the $t\bar{t}$-enriched region for $N_b > 1$.

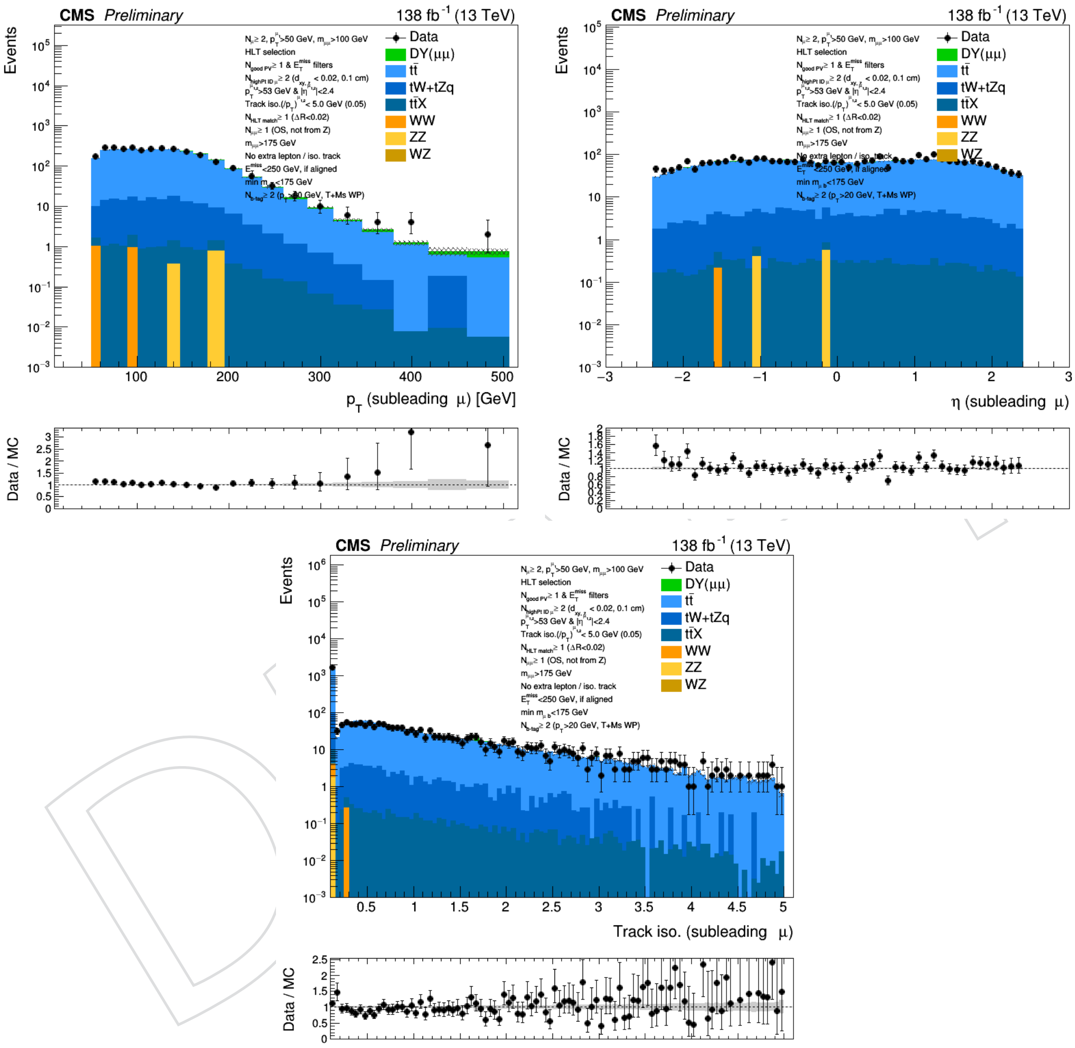


Figure 50: Distributions of p_T (upper left), η (upper right) and absolute isolation (lower) of the subleading muon in the $t\bar{t}$-enriched region for $N_b > 1$.

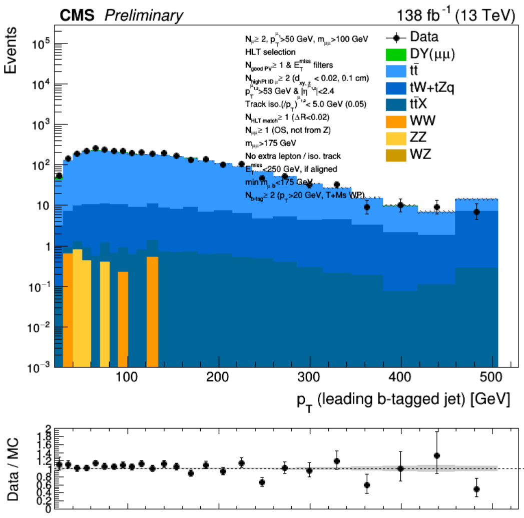


Figure 51: Distributions of p_T (upper left), η (upper right) and absolute isolation (lower) of the leading b-tagged jet in the $t\bar{t}$-enriched region for $N_b > 1$.

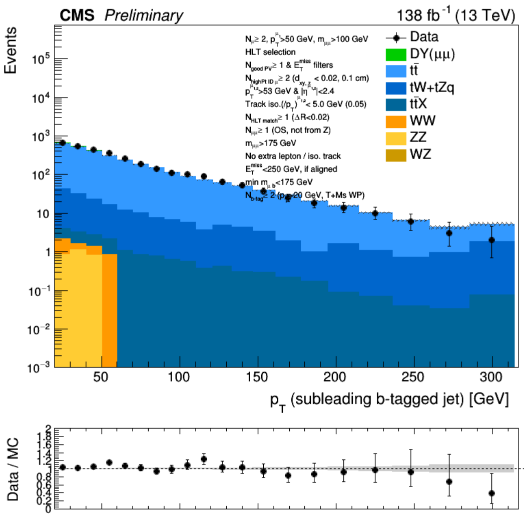


Figure 52: Distributions of p_T (upper left), η (upper right) and absolute isolation (lower) of the subleading b-tagged jet in the $t\bar{t}$-enriched region for $N_b > 1$.

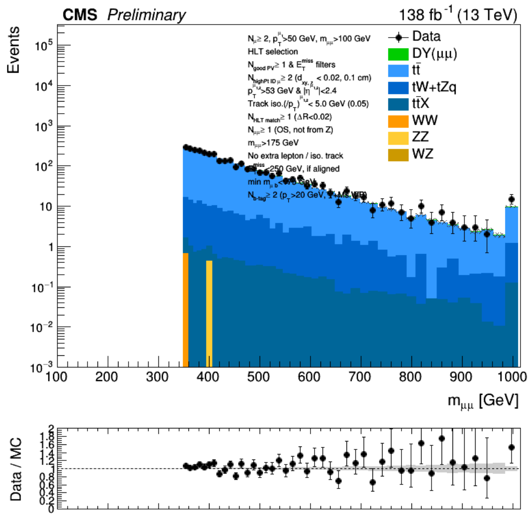


Figure 53: Distribution of dimuon mass spectrum, $m_{\mu\mu}$, in the $t\bar{t}$-enriched region for $N_b > 1$.

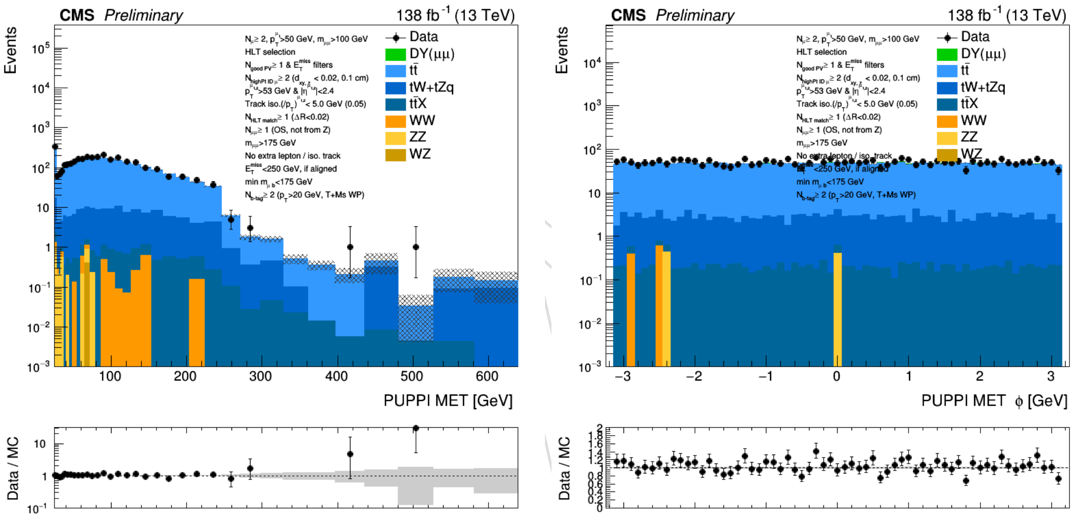


Figure 54: Distributions of p_T^{miss} magnitude (left) and ϕ (right) in the $t\bar{t}$-enriched region for $N_b > 1$.

C On the optimization of the b-tagging working points

As described in Section 6.1, we define two mutually exclusive event categories based on the number of b-tagged jets ($N_b = 1$ and $N_b \geq 2$). The b-tagging working points as well as the minimum p_T threshold for b-tagged jet counting are the result of an optimization study described here.

The optimization is based on the the 95% CL expected exclusion limits [66–69] on the number of signal events resulting from a two-bin counting experiment ($N_b = 1$ and $N_b \geq 2$). Limits are calculated across signal models and Z′ mass ($m_{Z'}$) hypotheses, with event counts based on MC simulation for both signal and background being drawn from a window in $m_{\mu\mu}$ around the $m_{Z'}$ hypothesis under test.

First, we optimized the $N_b = 1$ category with respect to the b-tagging working point and the minimum p_T of the b-tagged jets. For this optimization, we considered the 'medium' and 'tight' b-tagging working points, and a minimum b-tagged jet p_T of either 20 or 30 GeV, finding that the most stringent exclusion limits are achieved by requiring the b-tagged jet with $p_T > 20$ GeV be tagged according to the 'tight' working point.

Using this intermediate result, in the combined optimization of the $N_b = 1$ and $N_b \geq 2$ categories we imposed a minimum p_T threshold of 20 GeV for all b-tagged jets and considered three different combinations of b-tagging working points, where one b-tagged jet must always be tagged according to the 'tight' working point. The three combinations are listed in Table 13.

Table 13: Definitions of the three combinations of b-tagging working points (WP) considered for the combined optimization of the $N_b = 1$ and $N_b \geq 2$ event catgories. In all combinations, one b-tagged jet is required to be tagged accordding to the 'tight' working point, with $p_T >$ 20 GeV.

Option	$N_b = 1$	$N_b \geq 2$
1	1 'tight'-WP b jet	$+ \geq 1$ 'tight'-WP b jets
2	1 'tight'-WP b jet	$+ \geq 1$ 'medium'-WP b jets
3	1 'tight'-WP b jet	$+ \geq 1$ 'loose'-WP b jets

We then computed the expected 95% CL limits on the number of signal events for a number of representative signal models and Z′ mass hypotheses for the combination of the $N_b = 1$ and $N_b \geq 2$ categories under test. The results are summarized in Table 14.

Based on the expected limits across signal models and Z′ mass hypotheses, we found that the optimal b-tagging working point combination is achieved for option 2 of Table 13:

- in the $N_b = 1$ category, we require the presence of (exactly) one tight b jet with $p_T > 20$ GeV tagged according to the 'tight' working point;
- in the $N_b \geq 2$ category, we require the presence of additional medium b jets with $p_T > 20$ GeV tagged according to the 'medium' working point.

Table 14: Expected 95% CL limits on the number of signal events for representative signal models and $m_{Z'}$ hypotheses, as obtained from the combination of the $N_b = 1$ and $N_b \geq 2$ categories under test (see Table 13).

$m_{Z'} = 200$ GeV				
Option	Y_3	DY_3	DY'_3	$B_3 - L_2$
1	56.9	56.9	56.4	56.6
2	48.8	48.8	49.1	49.1
3	45.8	45.8	45.7	45.7
$m_{Z'} = 400$ GeV				
Option	Y_3	DY_3	DY'_3	$B_3 - L_2$
1	26.0	26.0	26.0	26.0
2	25.0	25.0	25.0	24.8
3	24.4	24.4	24.4	24.3
$m_{Z'} = 700$ GeV				
Option	Y_3	DY_3	DY'_3	$B_3 - L_2$
1	18.9	19.0	18.9	18.9
2	18.3	18.4	18.3	18.3
3	20.4	20.4	20.3	20.3
$m_{Z'} = 1000$ GeV				
Option	Y_3	DY_3	DY'_3	$B_3 - L_2$
1	12.8	12.8	12.8	12.8
2	12.9	12.9	12.9	12.9
3	14.3	14.3	14.3	14.3
$m_{Z'} = 1500$ GeV				
Option	Y_3	DY_3	DY'_3	$B_3 - L_2$
1	10.8	10.7	10.7	10.7
2	12.1	12.1	12.1	12.1
3	11.9	11.9	11.9	11.9
$m_{Z'} = 2000$ GeV				
Option	Y_3	DY_3	DY'_3	$B_3 - L_2$
1	9.3	9.3	9.3	9.3
2	10.2	10.2	10.2	10.2
3	10.1	10.1	10.1	10.1

D On the effect of the HEM 15/16 veto on signal acceptance

As described in Section 5.1, for runs after and including 319077 in 2018, the HEM 15/16 detectors are not operational. These subdetectors correspond to $-3.2 < \eta < -1.3$, $-1.57 < \phi < -0.87$ in the $\eta - \phi$ plane. It is observed that there are less jets in this region after all event selection criteria are applied, and that $p_{\mathrm{T}}^{\mathrm{miss}}$ has larger tails for events with at least one jet in this region. Therefore, events are vetoed if jets and/or electrons are found in this region (enlarged by 0.2 in η and ϕ for jets to account for the size of the jet cone) in the affected data taking runs.

The same veto, referred to as 'HEM veto', is applied to simulated events, accounting for the fraction of affected data, and it leads to a decrease in the signal yield by at most 5% for the whole analyzed data set (2016–2018), independently of the dimuon mass.

Figures 55–57 show the effect of the HEM veto for a few representative signal models, for 2018 only. The reduction in signal acceptance by roughly 12–13% in 2018 is independent of the signal mass hypothesis and of the N_{b} event category. Over the full data set in use (2016–2018), the reduction in signal accptance from the HEM veto is $\lesssim 5\%$. No statistically significant effect is found on the shape of the signal expected dimuon invariant mass distribution.

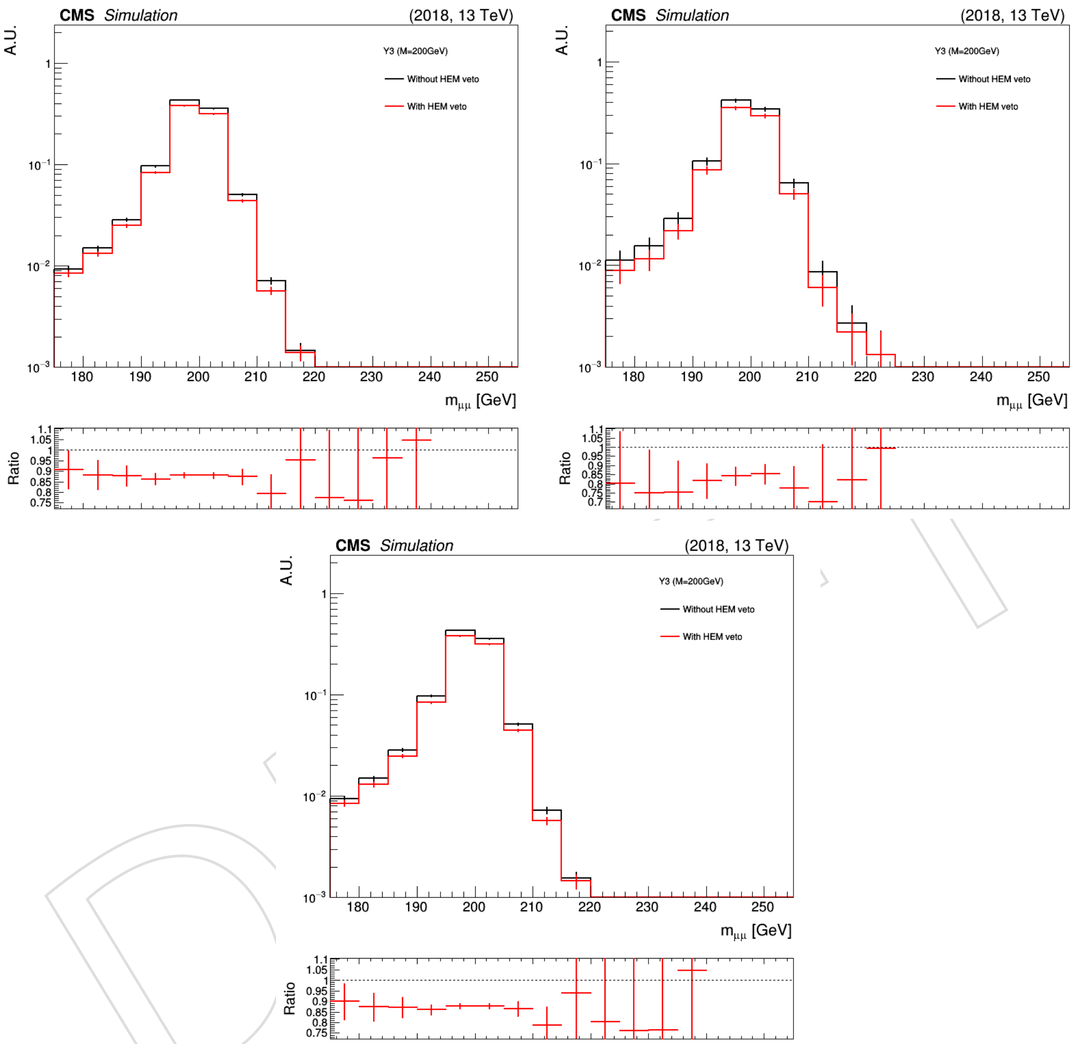


Figure 55: The effect on the expected signal yield and signal $m_{\mu\mu}$ shape from the application of the HEM veto is shown, as measured in a representative 2018 signal MC sample (Y3, with $m_{Z'} = 200\,\text{GeV}$), after the full event selection, in event categories with $N_b = 1$ (top left), $N_b \geq 2$ (top right), and $N_b \geq 1$ (bottom; equal to the sum of $N_b = 1$ and $N_b \geq 2$). The black histogram represents the signal dimuon invariant mass distribution without HEM veto. The red histogram represents the same distribution once the veto is applied. Both distributions are normalized to the area of the histogram with no veto applied.

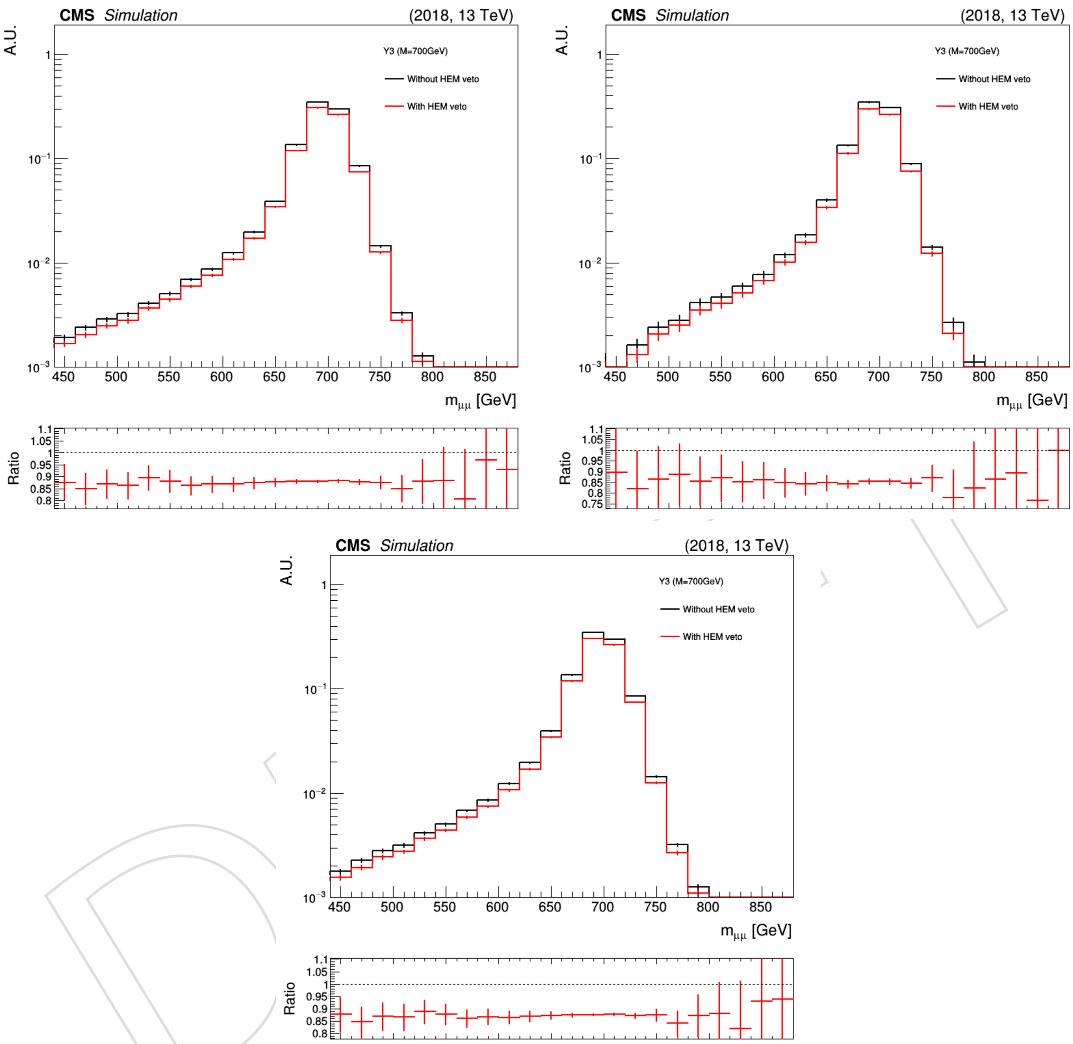


Figure 56: The effect on the expected signal yield and signal $m_{\mu\mu}$ shape from the application of the HEM veto is shown, as measured in a representative 2018 signal MC sample (Y3, with $m_{Z'} = 700\,\text{GeV}$), after the full event selection, in event categories with $N_b = 1$ (top left), $N_b \geq 2$ (top right), and $N_b \geq 1$ (bottom; equal to the sum of $N_b = 1$ and $N_b \geq 2$). The black histogram represents the signal dimuon invariant mass distribution without HEM veto. The red histogram represents the same distribution once the veto is applied. Both distributions are normalized to the area of the histogram with no veto applied.

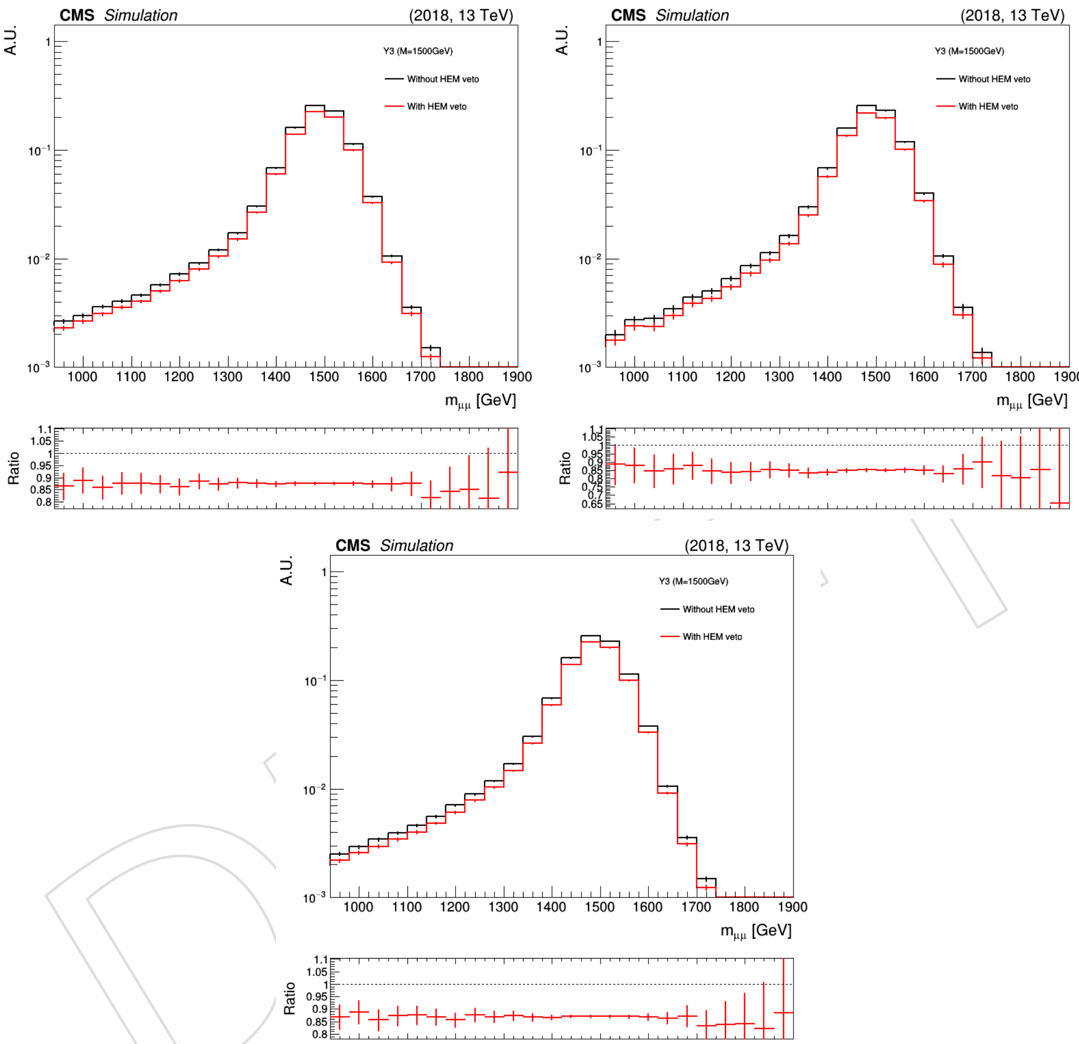


Figure 57: The effect on the expected signal yield and signal $m_{\mu\mu}$ shape from the application of the HEM veto is shown, as measured in a representative 2018 signal MC sample (Y3, with $m_{Z'} = 1500\,\text{GeV}$), after the full event selection, in event categories with $N_{\text{b}} = 1$ (top left), $N_{\text{b}} \geq 2$ (top right), and $N_{\text{b}} \geq 1$ (bottom; equal to the sum of $N_{\text{b}} = 1$ and $N_{\text{b}} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution without HEM veto. The red histogram represents the same distribution once the veto is applied. Both distributions are normalized to the area of the histogram with no veto applied.

E On the effect of systematic uncertainties

In the following, we evaluate the effect of all sources of uncertainty related to pileup modeling, trigger efficiency measurement, as well as physics object recontruction on the signal MC:

- pileup modeling;
- trigger efficiency measurement;
- jet energy resolution;
- jet energy scale;
- b-tagging efficiency;
- muon reconstruction, identification, and isolation.

Based on the measured effect, corresponding systematic uncertainties are assessed in the expected signal yield and/or in the expected signal $m_{\mu\mu}$ shape.

E.1 Uncertainty in pileup modeling

Figures 58–60 show the effect of the propagation of the upward uncertainty in the minimum bias cross section used for the MC pileup profile [19], for a few representative signal models. Similarly, Figs. 61–63 show the effect of the propagation of the downward uncertainty in the minimum bias cross section used for the MC pileup profile, for the same signal models.

The effect in the signal yield is negligible compared to the statistical uncertainties, and subdominat with respect to other systematic uncertainties. No statistically significant effect is found on the shape of the signal expected dimuon invariant mass distribution. Thus, no dedicated systematic uncertainty in the pileup modeling is assessed.

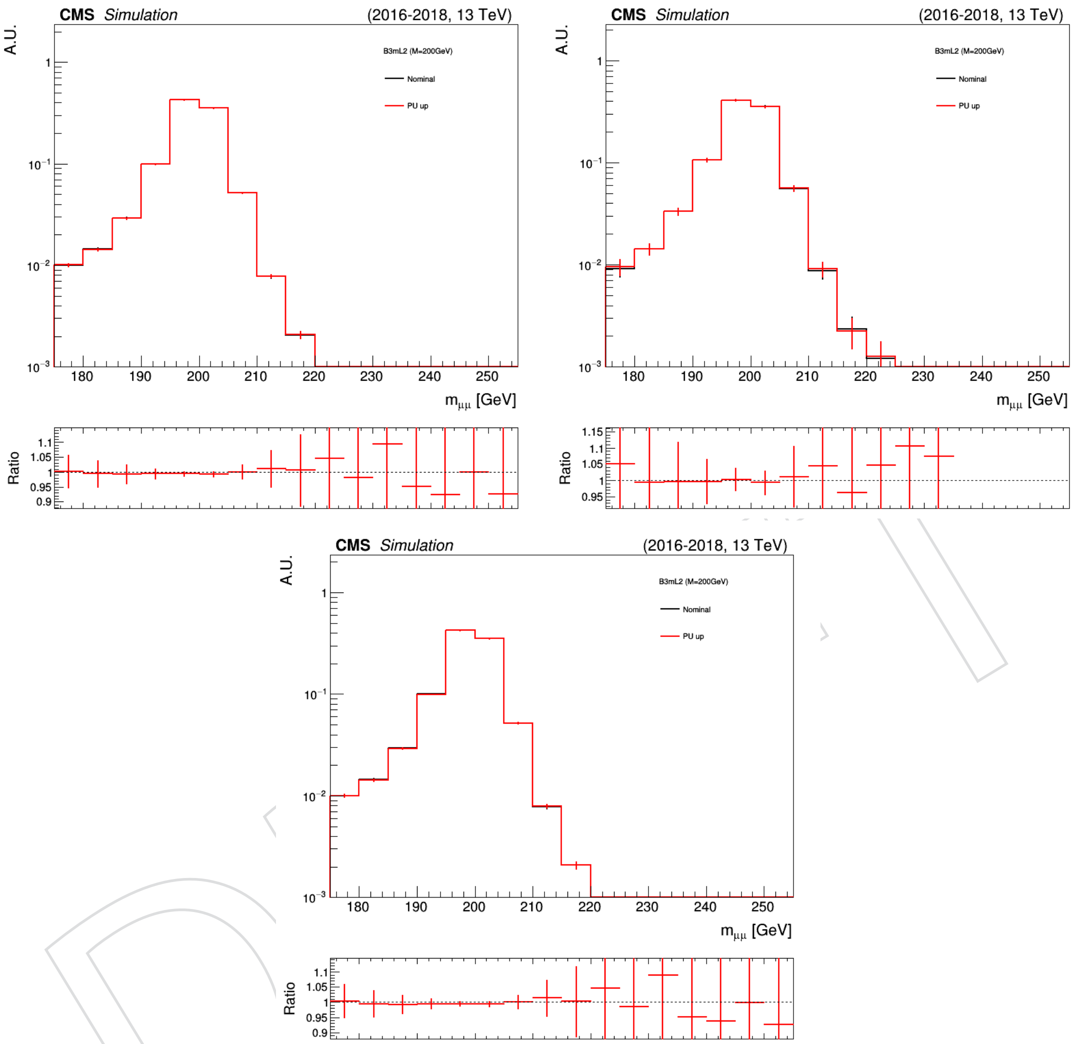


Figure 58: The effect of the propagatation of the upward uncertainty in the minimum bias cross section used for the pileup profile on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 200\,\text{GeV}$), after the full event selection, in event categories with $N_{\text{b}} = 1$ (top left), $N_{\text{b}} \geq 2$ (top right), and $N_{\text{b}} \geq 1$ (bottom; equal to the sum of $N_{\text{b}} = 1$ and $N_{\text{b}} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the nominal value of the minimum bias cross section (equal to 69.2 mb) for the MC pileup profile. The red histogram represents the same distribution using a minimum bias cross section equal to the nominal value plus the uncertainty (4.6%) in its measurement (equal to 72.4 mb). Both distributions are normalized to the area of the nominal histogram.

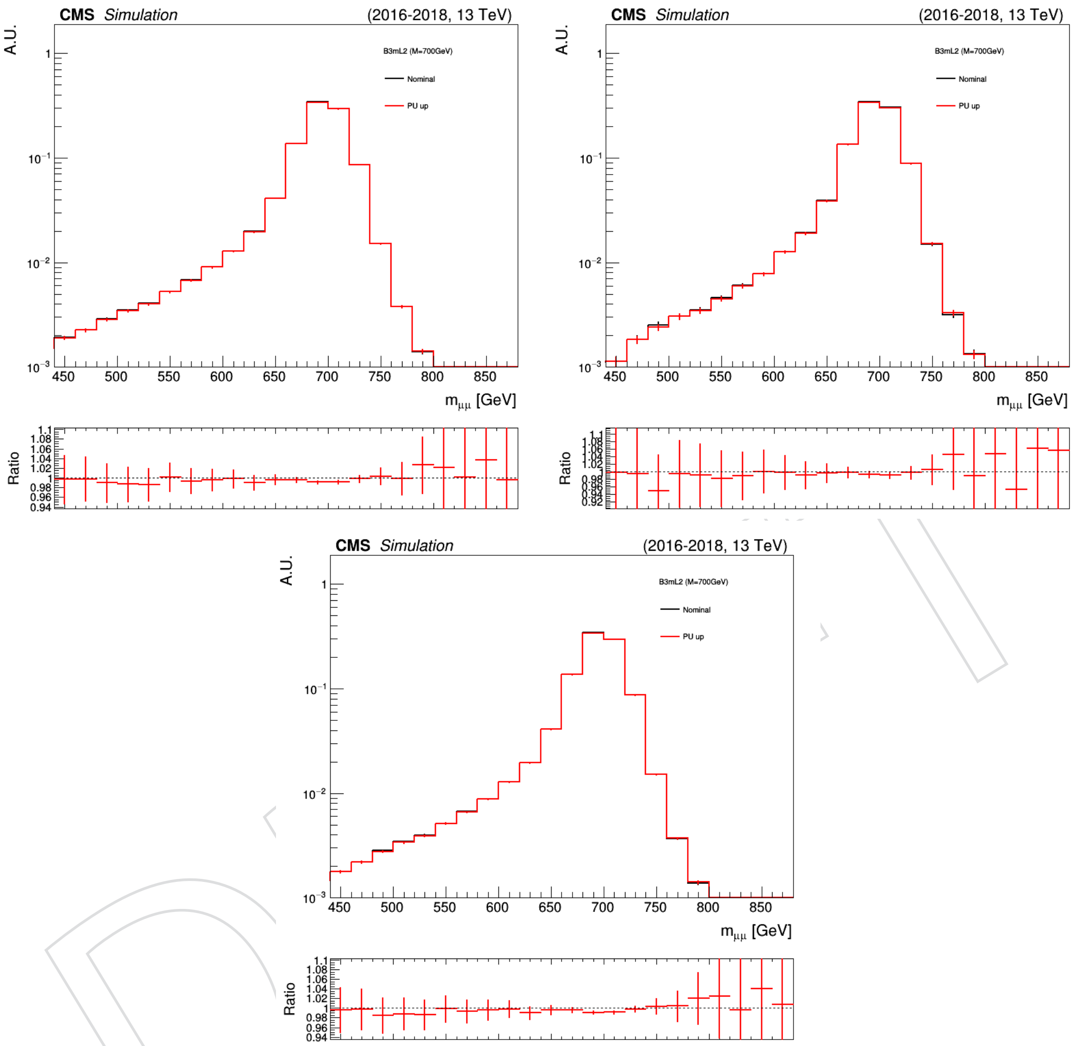


Figure 59: The effect of the propagation of the upward uncertainty in the minimum bias cross section used for the pileup profile on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 700\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the nominal value of the minimum bias cross section (equal to 69.2 mb) for the MC pileup profile. The red histogram represents the same distribution using a minimum bias cross section equal to the nominal value plus the uncertainty (4.6%) in its measurement (equal to 72.4 mb). Both distributions are normalized to the area of the nominal histogram.

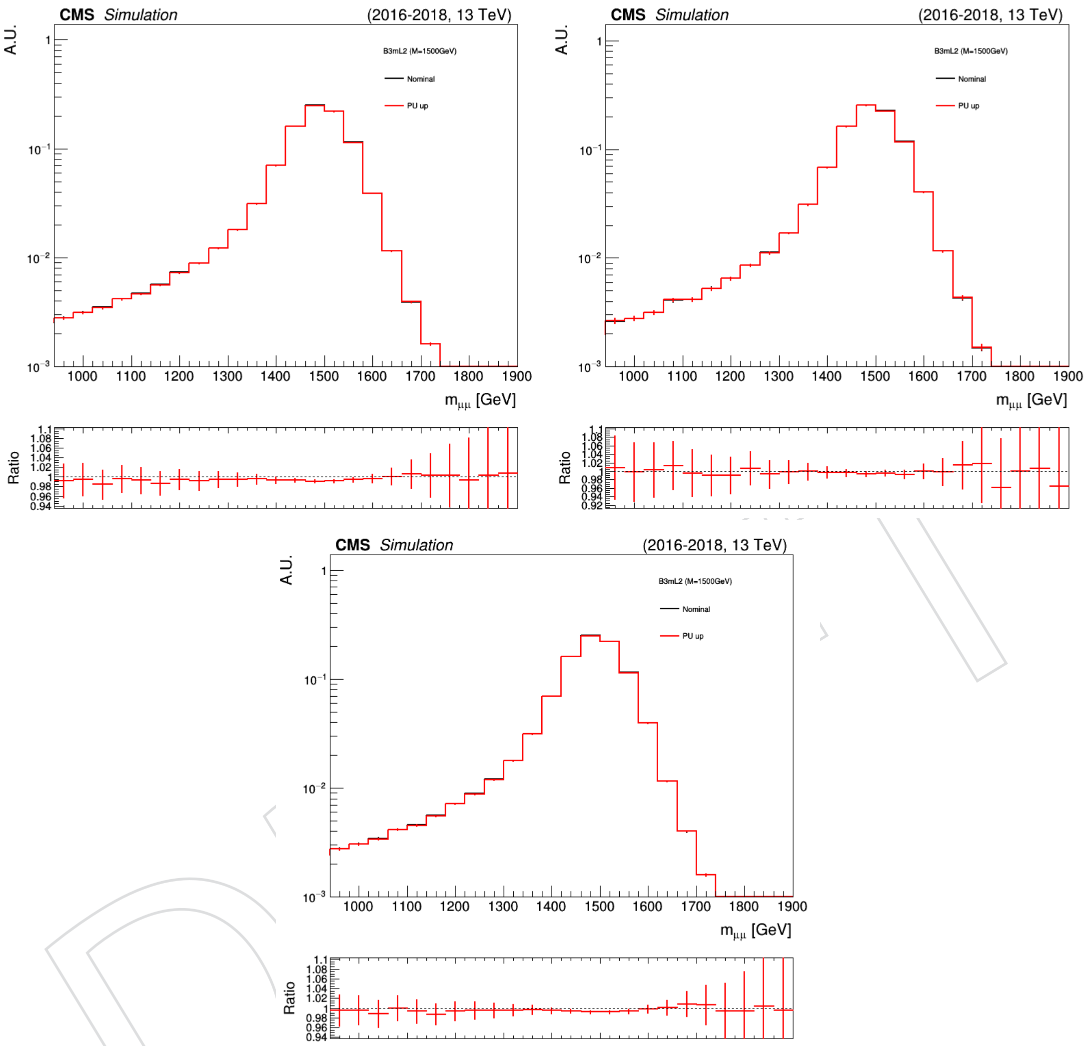


Figure 60: The effect of the propagatation of the upward uncertainty in the minimum bias cross section used for the pileup profile on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 1500\,\mathrm{GeV}$), after the full event selection, in event categories with $N_\mathrm{b} = 1$ (top left), $N_\mathrm{b} \geq 2$ (top right), and $N_\mathrm{b} \geq 1$ (bottom; equal to the sum of $N_\mathrm{b} = 1$ and $N_\mathrm{b} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the nominal value of the minimum bias cross section (equal to 69.2 mb) for the MC pileup profile. The red histogram represents the same distribution using a minimum bias cross section equal to the nominal value plus the uncertainty (4.6%) in its measurement (equal to 72.4 mb). Both distributions are normalized to the area of the nominal histogram.

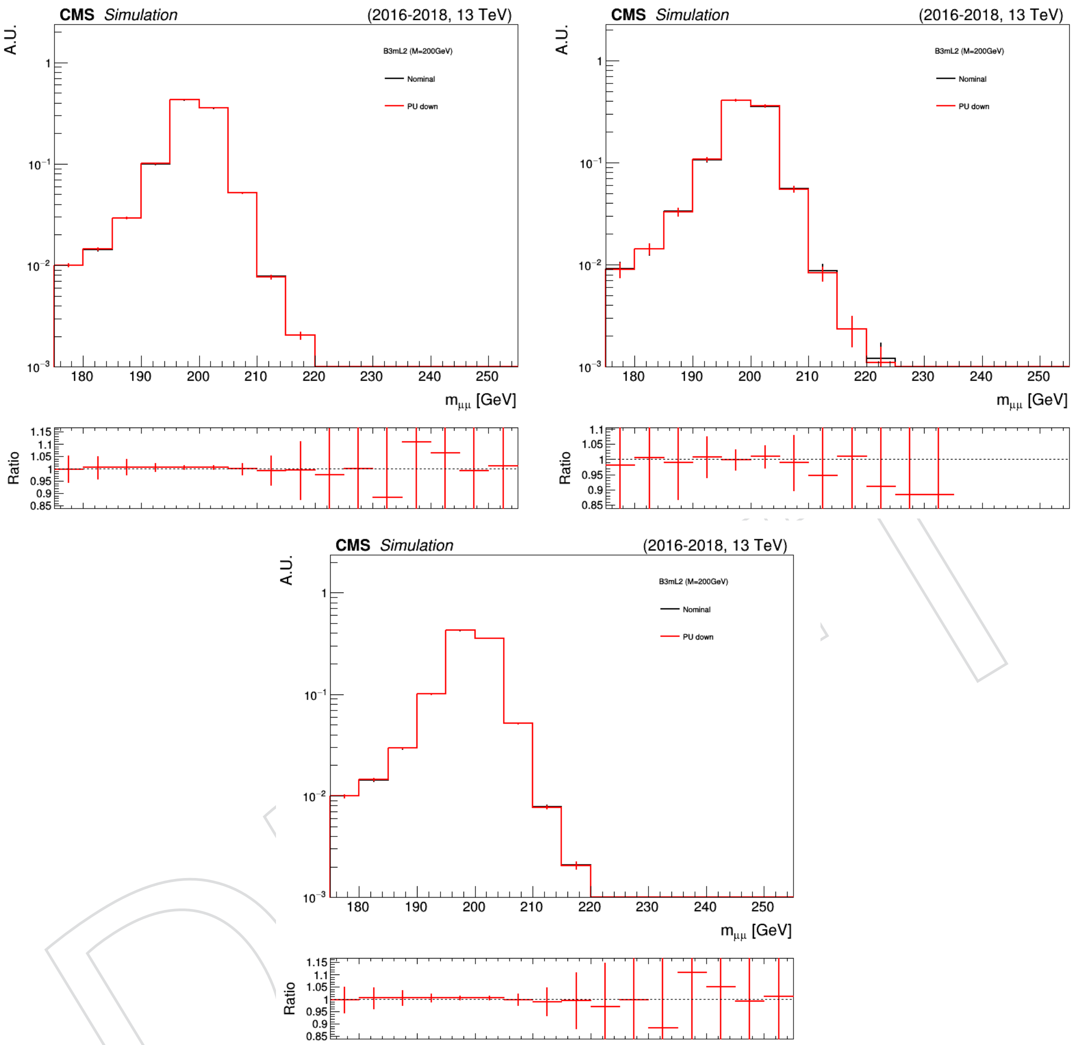


Figure 61: The effect of the propagatation of the downward uncertainty in the minimum bias cross section used for the pileup profile on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 200\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the nominal value of the minimum bias cross section (equal to 69.2 mb) for the MC pileup profile. The red histogram represents the same distribution using a minimum bias cross section equal to the nominal value minus the uncertainty (4.6%) in its measurement (equal to 66.0 mb). Both distributions are normalized to the area of the nominal histogram.

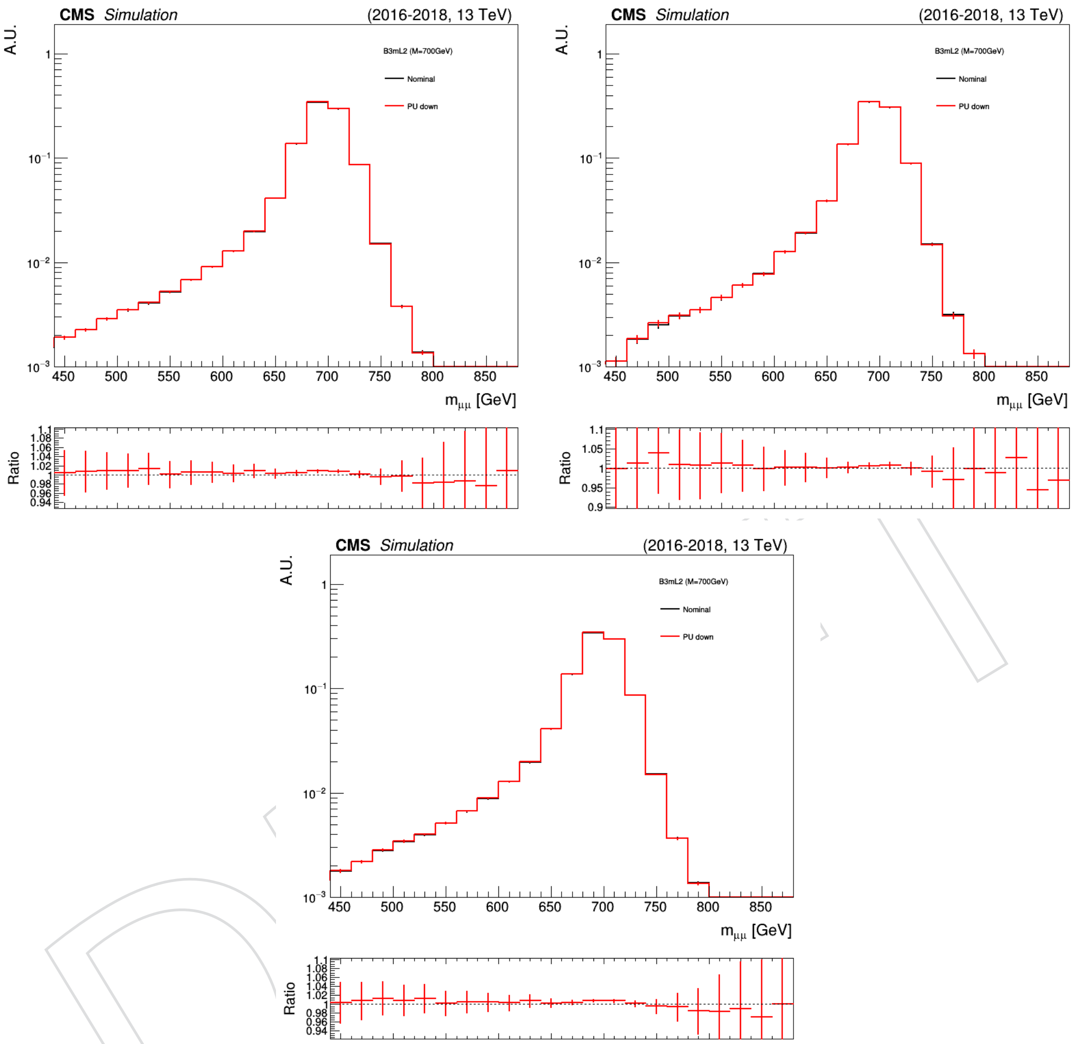


Figure 62: The effect of the propagatation of the downward uncertainty in the minimum bias cross section used for the pileup profile on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 700\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the nominal value of the minimum bias cross section (equal to 69.2 mb) for the MC pileup profile. The red histogram represents the same distribution using a minimum bias cross section equal to the nominal value minus the uncertainty (4.6%) in its measurement (equal to 66.0 mb). Both distributions are normalized to the area of the nominal histogram.

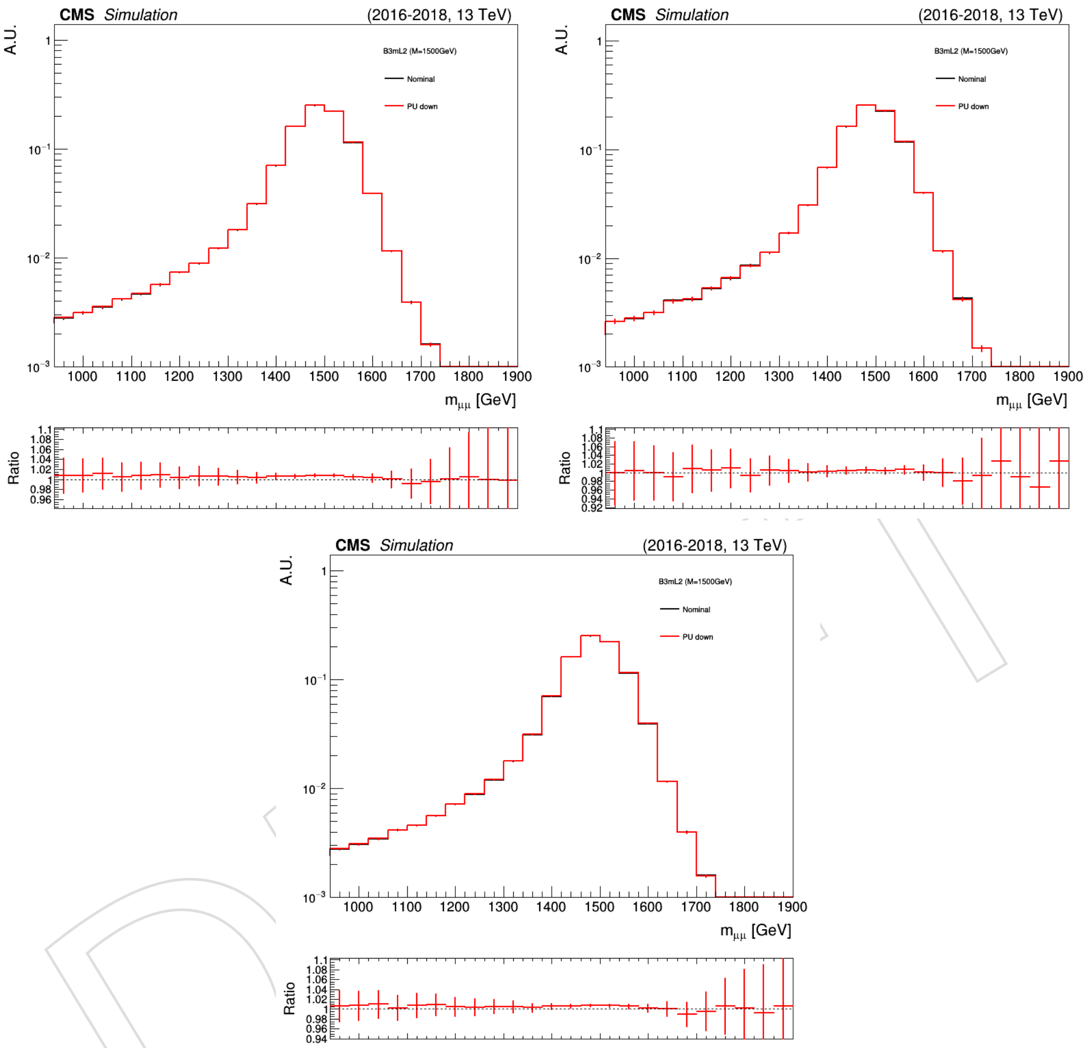


Figure 63: The effect of the propagatation of the downward uncertainty in the minimum bias cross section used for the pileup profile on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 1500\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the nominal value of the minimum bias cross section (equal to 69.2 mb) for the MC pileup profile. The red histogram represents the same distribution using a minimum bias cross section equal to the nominal value minus the uncertainty (4.6%) in its measurement (equal to 66.0 mb). Both distributions are normalized to the area of the nominal histogram.

E.2 Uncertainty in trigger efficiency

Figures 64–66 show the effect of the propagation of the upward uncertainty in the trigger efficiency measurement [70–72] for a few representative signal models. Similarly, Figs. 67–69 show the effect of the propagation of the downward uncertainty in the trigger efficiency measurement for the same signal models.

The effect in the signal yield is increasing at increasing $m_{Z'}$, ranging roughly from 1% at $m_{Z'} = 200\,\text{GeV}$ to 5% at $m_{Z'} = 2000\,\text{GeV}$, independently of N_b. No statistically significant effect is found on the shape of the signal expected dimuon invariant mass distribution. Thus, a dedicated systematic uncertainty in the expected signal yield is assessed.

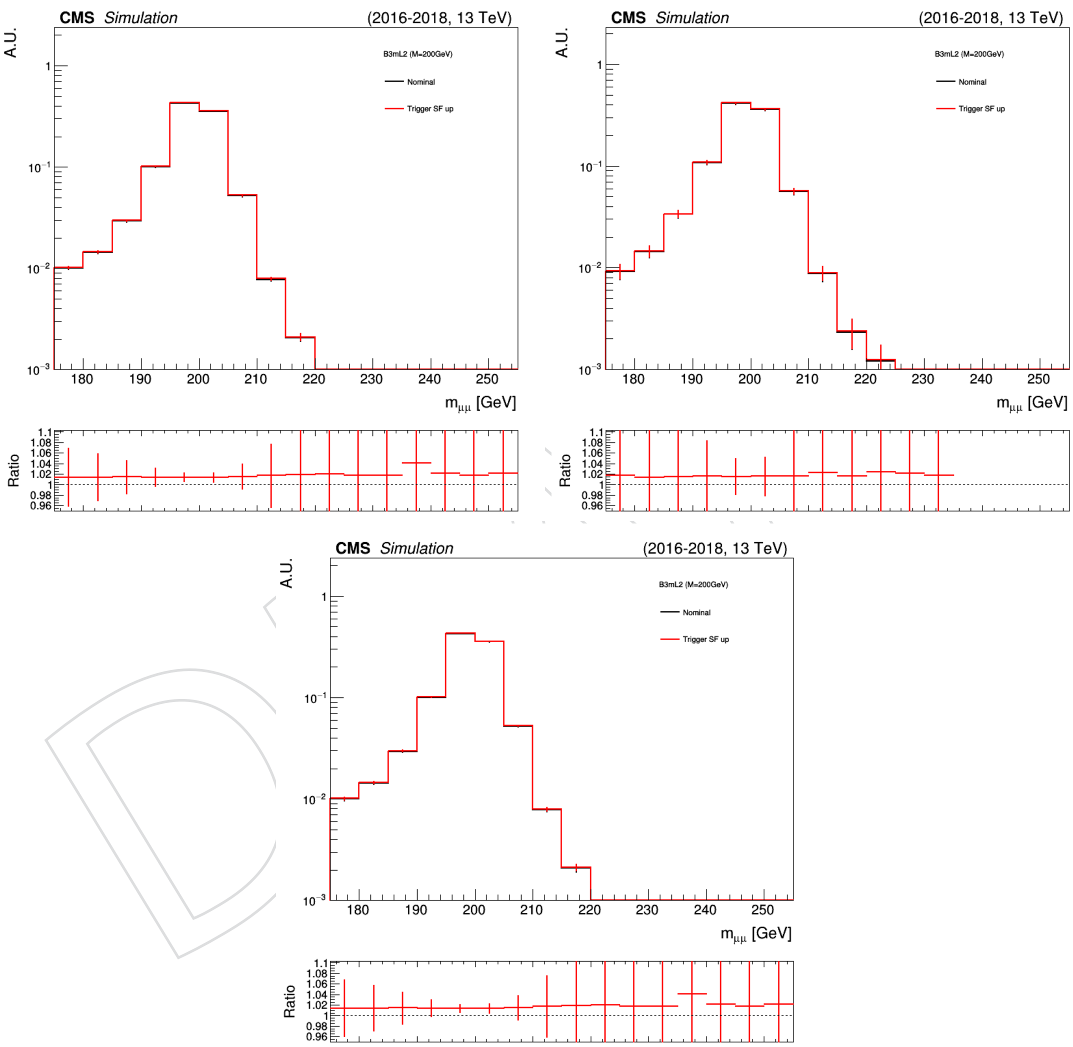


Figure 64: The effect of the propagatation of the upward uncertainty in the trigger efficiency measurement on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 200\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the central value of the trigger data/MC scale factors. The red histogram represents the same distribution using the central value plus the corresponding uncertainty. Both distributions are normalized to the area of the nominal histogram.

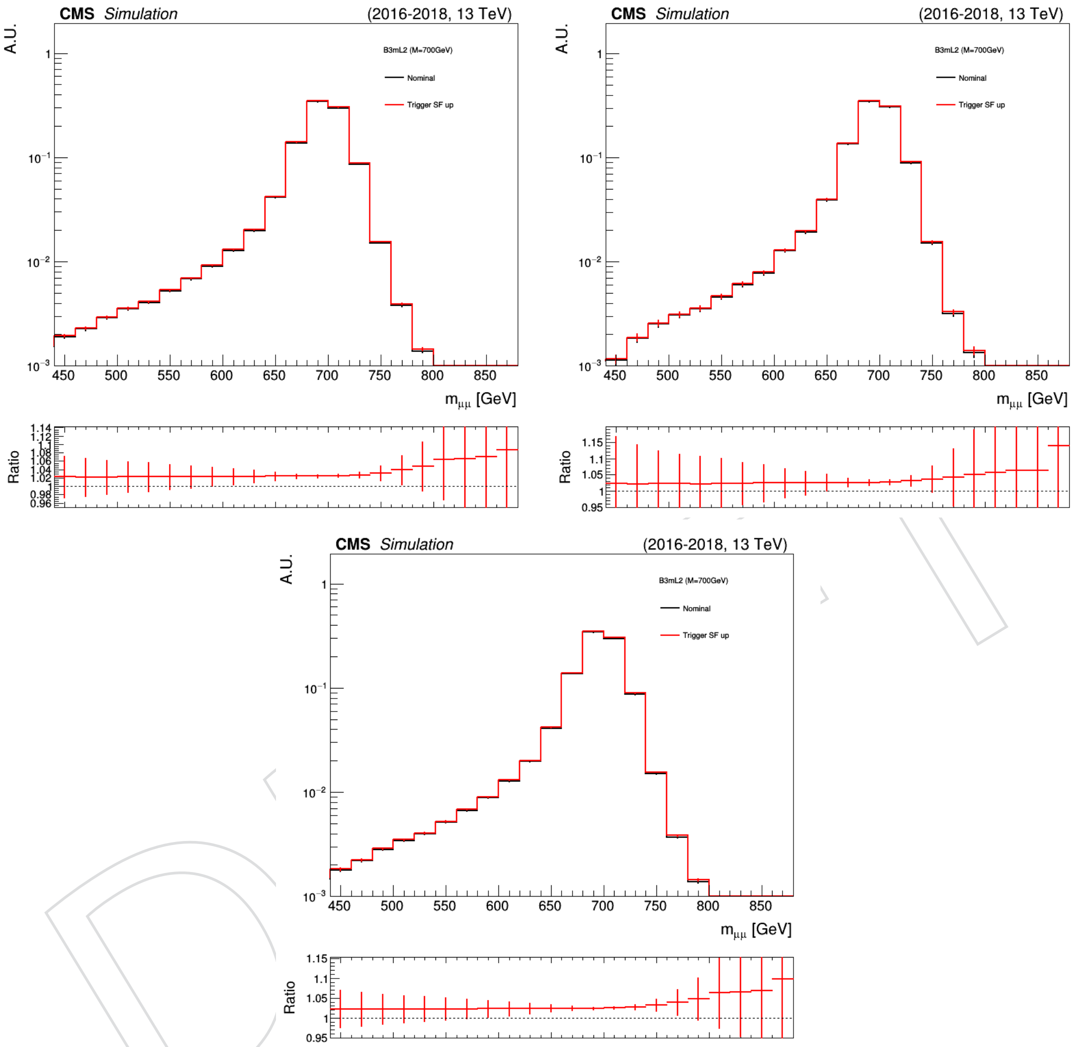


Figure 65: The effect of the propagatation of the upward uncertainty in the trigger efficiency measurement on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 700\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the central value of the trigger data/MC scale factors. The red histogram represents the same distribution using the central value plus the corresponding uncertainty. Both distributions are normalized to the area of the nominal histogram.

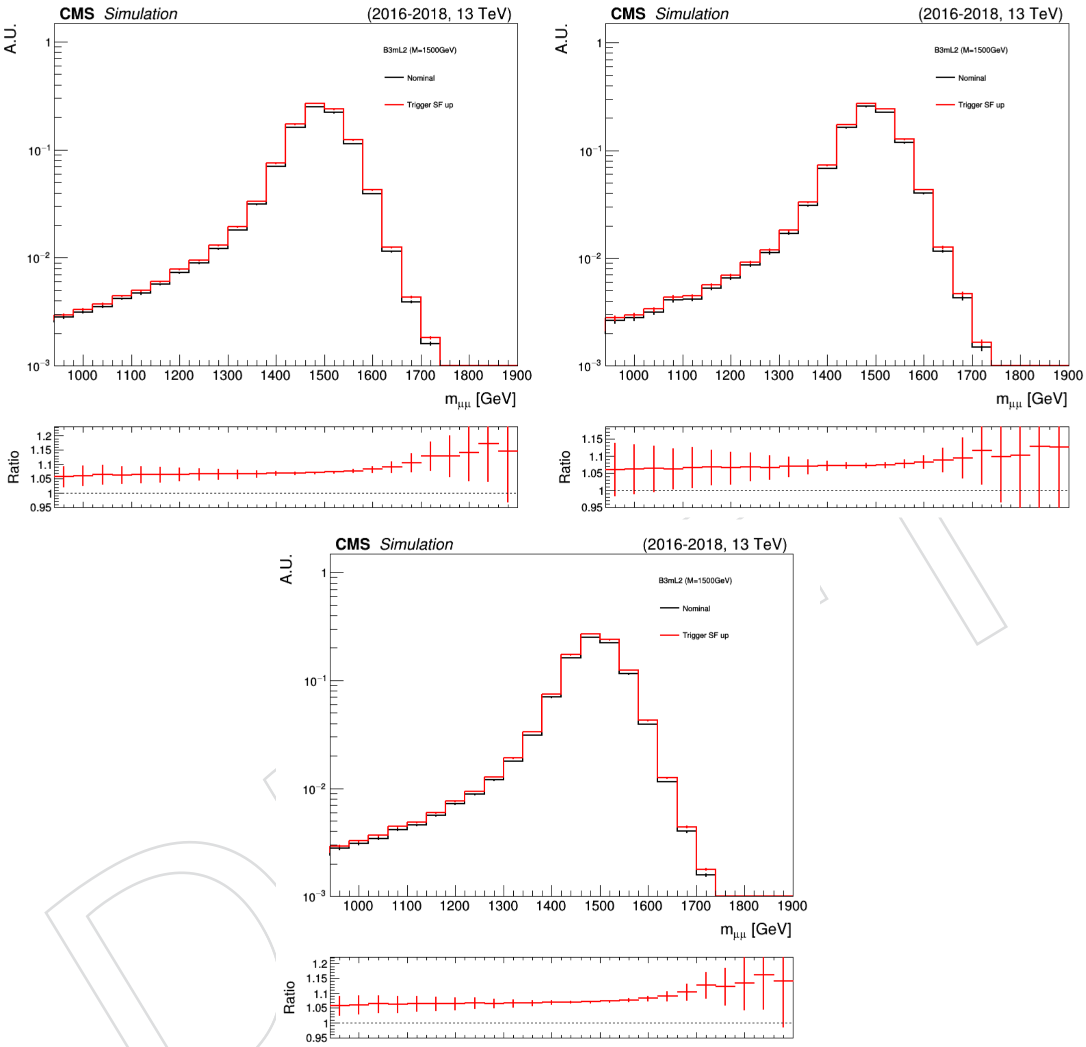


Figure 66: The effect of the propagatation of the upward uncertainty in the trigger efficiency measurement on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 1500\,\text{GeV}$), after the full event selection, in event categories with $N_{\text{b}} = 1$ (top left), $N_{\text{b}} \geq 2$ (top right), and $N_{\text{b}} \geq 1$ (bottom; equal to the sum of $N_{\text{b}} = 1$ and $N_{\text{b}} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the central value of the trigger data/MC scale factors. The red histogram represents the same distribution using the central value plus the corresponding uncertainty. Both distributions are normalized to the area of the nominal histogram.

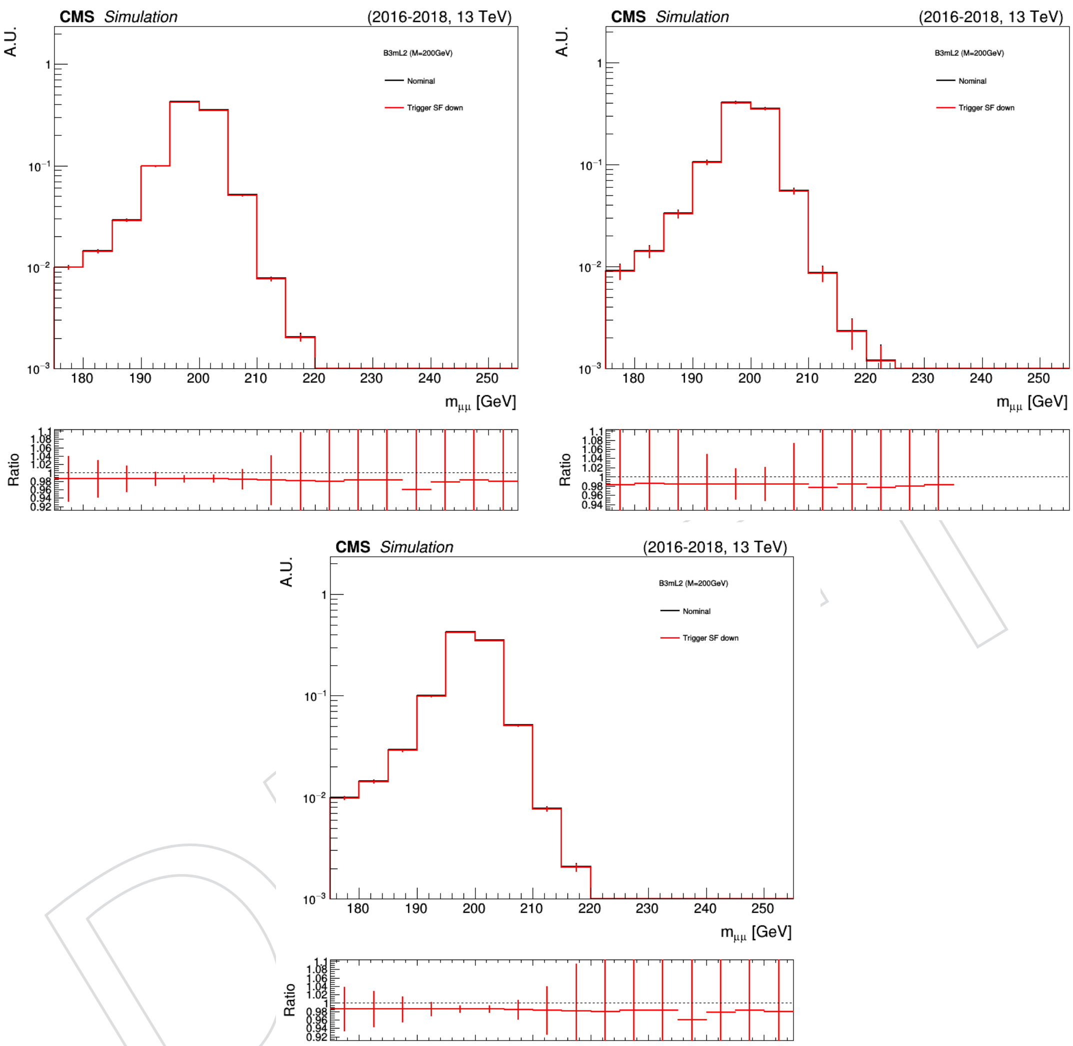


Figure 67: The effect of the propagatation of the downward uncertainty in the trigger efficiency measurement on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 200\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the central value of the trigger data/MC scale factors. The red histogram represents the same distribution using the central value minus the corresponding uncertainty. Both distributions are normalized to the area of the nominal histogram.

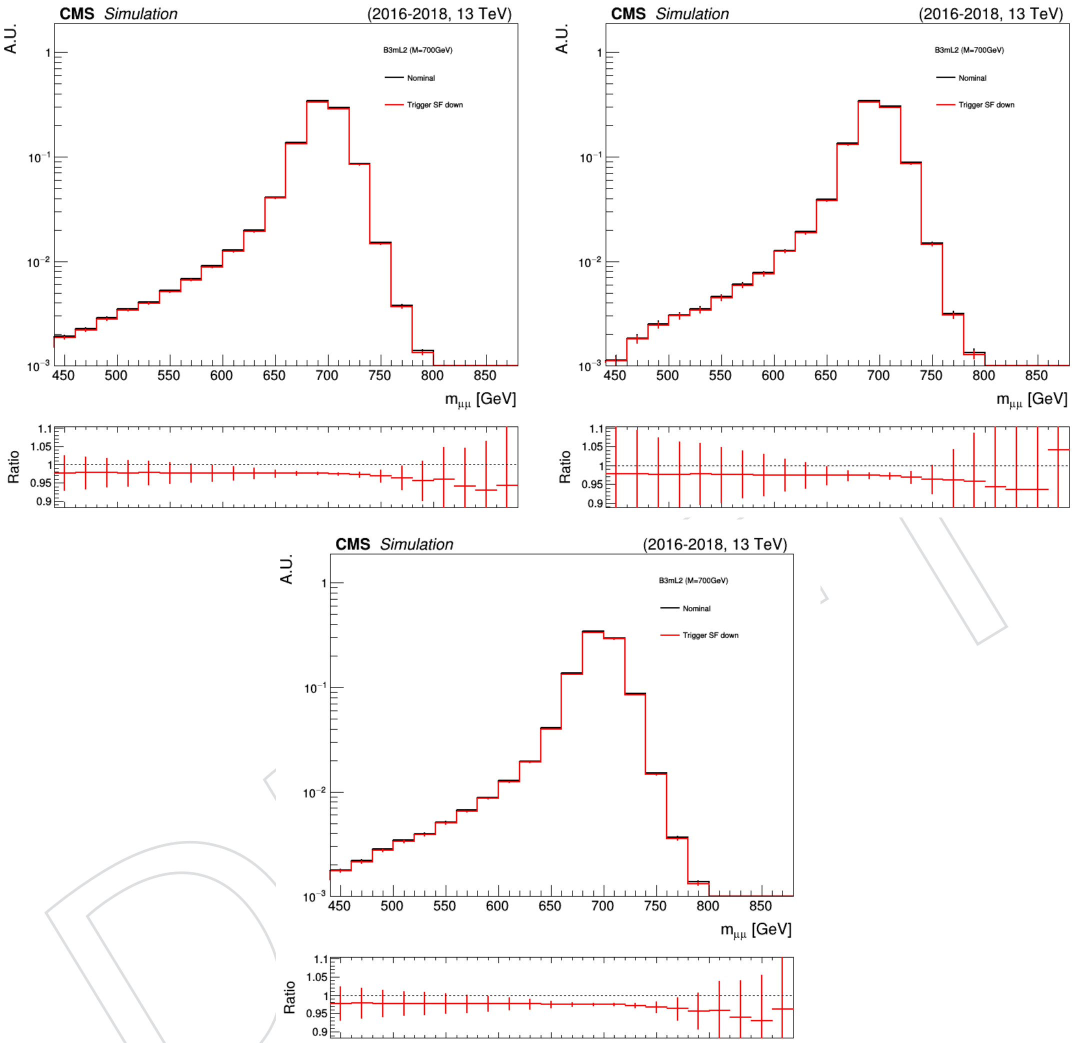


Figure 68: The effect of the propagatation of the downward uncertainty in the trigger efficiency measurement on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 700\,\mathrm{GeV}$), after the full event selection, in event categories with $N_{\mathrm{b}} = 1$ (top left), $N_{\mathrm{b}} \geq 2$ (top right), and $N_{\mathrm{b}} \geq 1$ (bottom; equal to the sum of $N_{\mathrm{b}} = 1$ and $N_{\mathrm{b}} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the central value of the trigger data/MC scale factors. The red histogram represents the same distribution using the central value minus the corresponding uncertainty. Both distributions are normalized to the area of the nominal histogram.

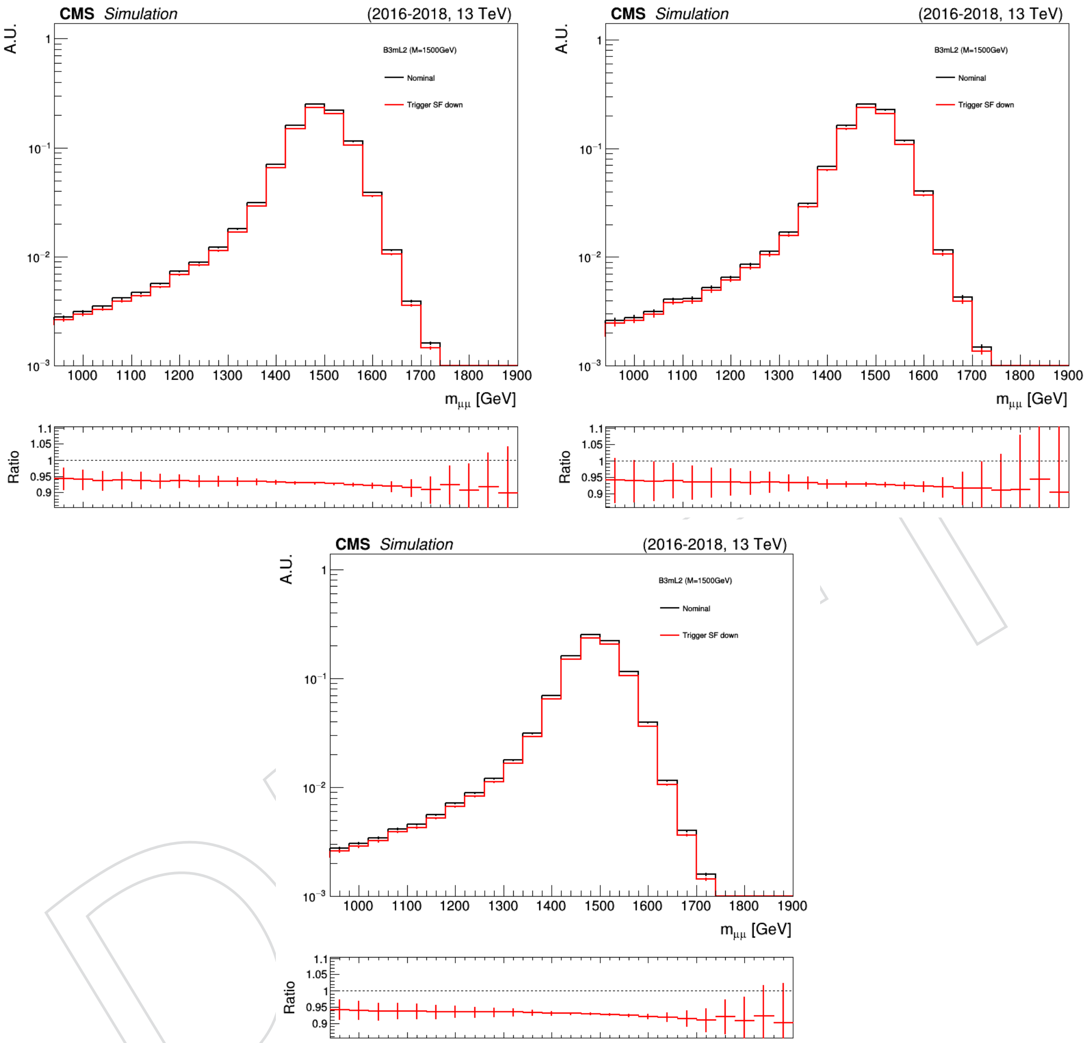


Figure 69: The effect of the propagatation of the downward uncertainty in the trigger efficiency measurement on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 1500\,\mathrm{GeV}$), after the full event selection, in event categories with $N_\mathrm{b} = 1$ (top left), $N_\mathrm{b} \geq 2$ (top right), and $N_\mathrm{b} \geq 1$ (bottom; equal to the sum of $N_\mathrm{b} = 1$ and $N_\mathrm{b} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the central value of the trigger data/MC scale factors. The red histogram represents the same distribution using the central value minus the corresponding uncertainty. Both distributions are normalized to the area of the nominal histogram.

E.3 Uncertainty in L1 prefire weight

Figures 70–72 show the effect of the propagation of the upward statistical uncertainty in the L1 prefire weight [73] for a few representative signal models. Similarly, Figs. 73–75 show the effect of the propagation of the downward statistical uncertainty in the L1 prefire weight for the same signal models.

The effect in the signal yield is negligible compared to the statistical uncertainties, and subdominat with respect to other systematic uncertainties. No statistically significant effect is found on the shape of the signal expected dimuon invariant mass distribution. Thus, no dedicated systematic uncertainty arising from the statistical uncertainty in the L1 prefire weight is assessed.

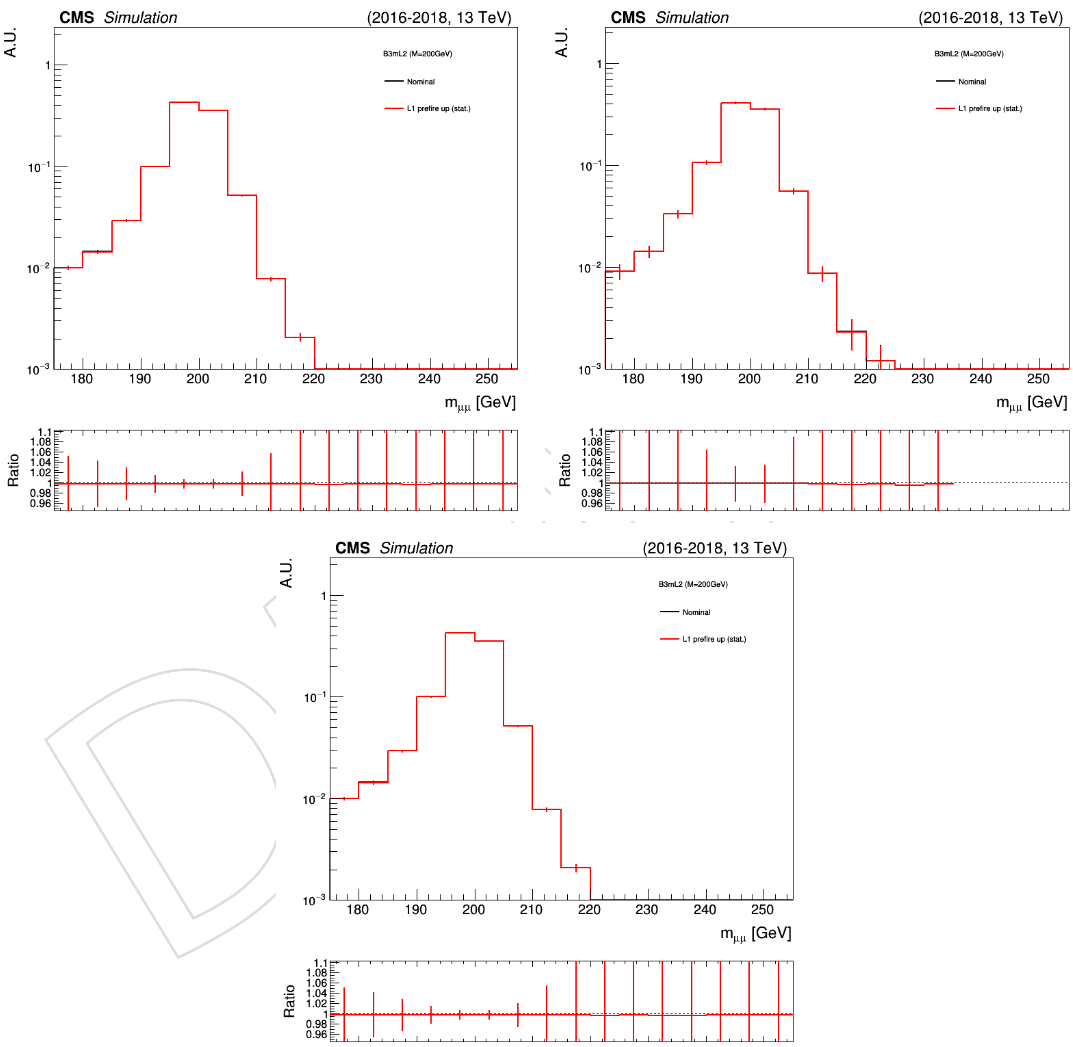


Figure 70: The effect of the propagatation of the upward statistical uncertainty in the L1 prefire weight on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 200\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the nominal L1 prefire weights. The red histogram represents the same distribution using the nominal value plus the corresponding statistical uncertainty. Both distributions are normalized to the area of the nominal histogram.

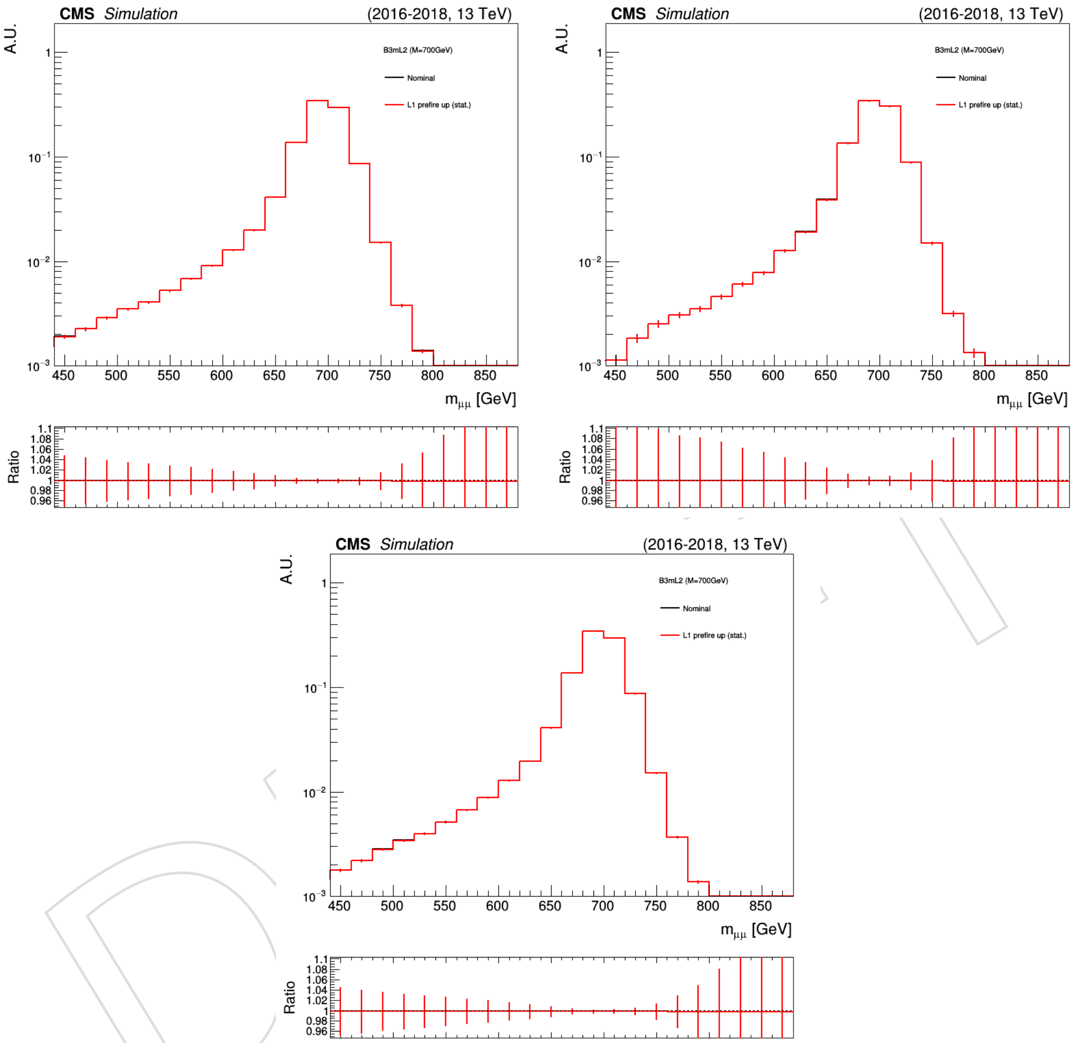


Figure 71: The effect of the propagatation of the upward statistical uncertainty in the L1 prefire weight on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 700\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the nominal L1 prefire weights. The red histogram represents the same distribution using the nominal value plus the corresponding statistical uncertainty. Both distributions are normalized to the area of the nominal histogram.

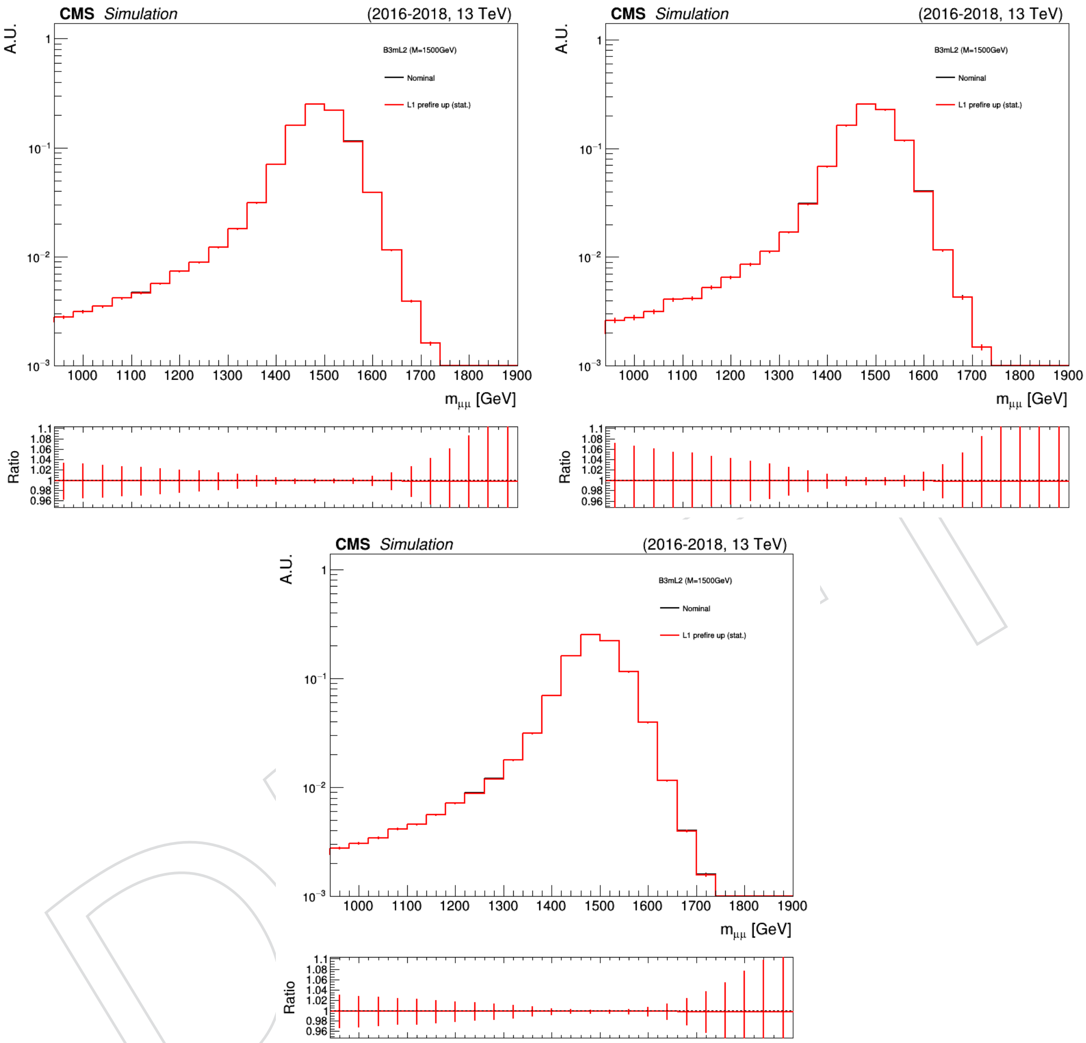


Figure 72: The effect of the propagatation of the upward statistical uncertainty in the L1 prefire weight on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 1500\,\text{GeV}$), after the full event selection, in event categories with $N_{\text{b}} = 1$ (top left), $N_{\text{b}} \geq 2$ (top right), and $N_{\text{b}} \geq 1$ (bottom; equal to the sum of $N_{\text{b}} = 1$ and $N_{\text{b}} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the nominal L1 prefire weights. The red histogram represents the same distribution using the nominal value plus the corresponding statistical uncertainty. Both distributions are normalized to the area of the nominal histogram.

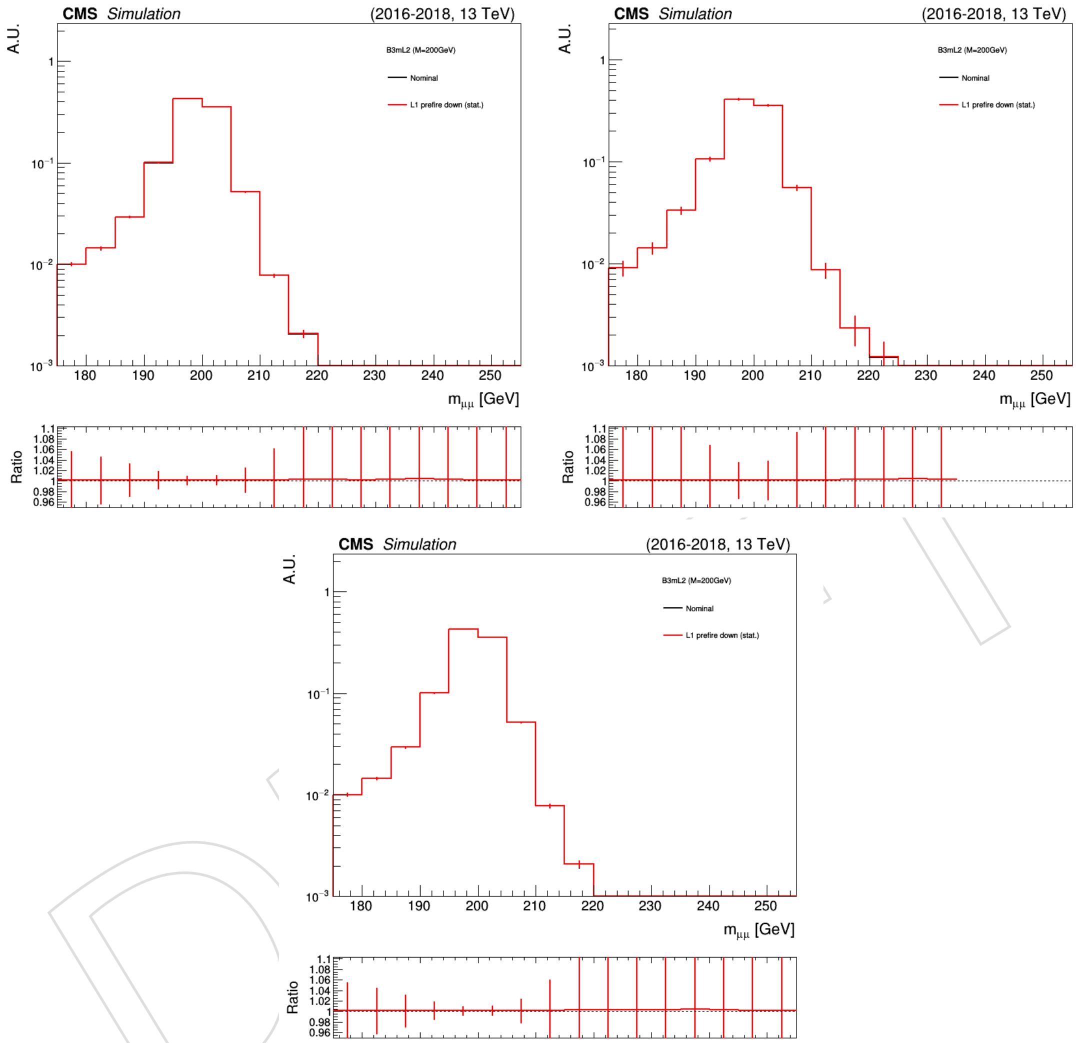


Figure 73: The effect of the propagatation of the downward statistical uncertainty in the L1 prefire weight on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 200\,\text{GeV}$), after the full event selection, in event categories with $N_b = 1$ (top left), $N_b \geq 2$ (top right), and $N_b \geq 1$ (bottom; equal to the sum of $N_b = 1$ and $N_b \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the nominal L1 prefire weights. The red histogram represents the same distribution using the nominal value minus the corresponding statistical uncertainty. Both distributions are normalized to the area of the nominal histogram.

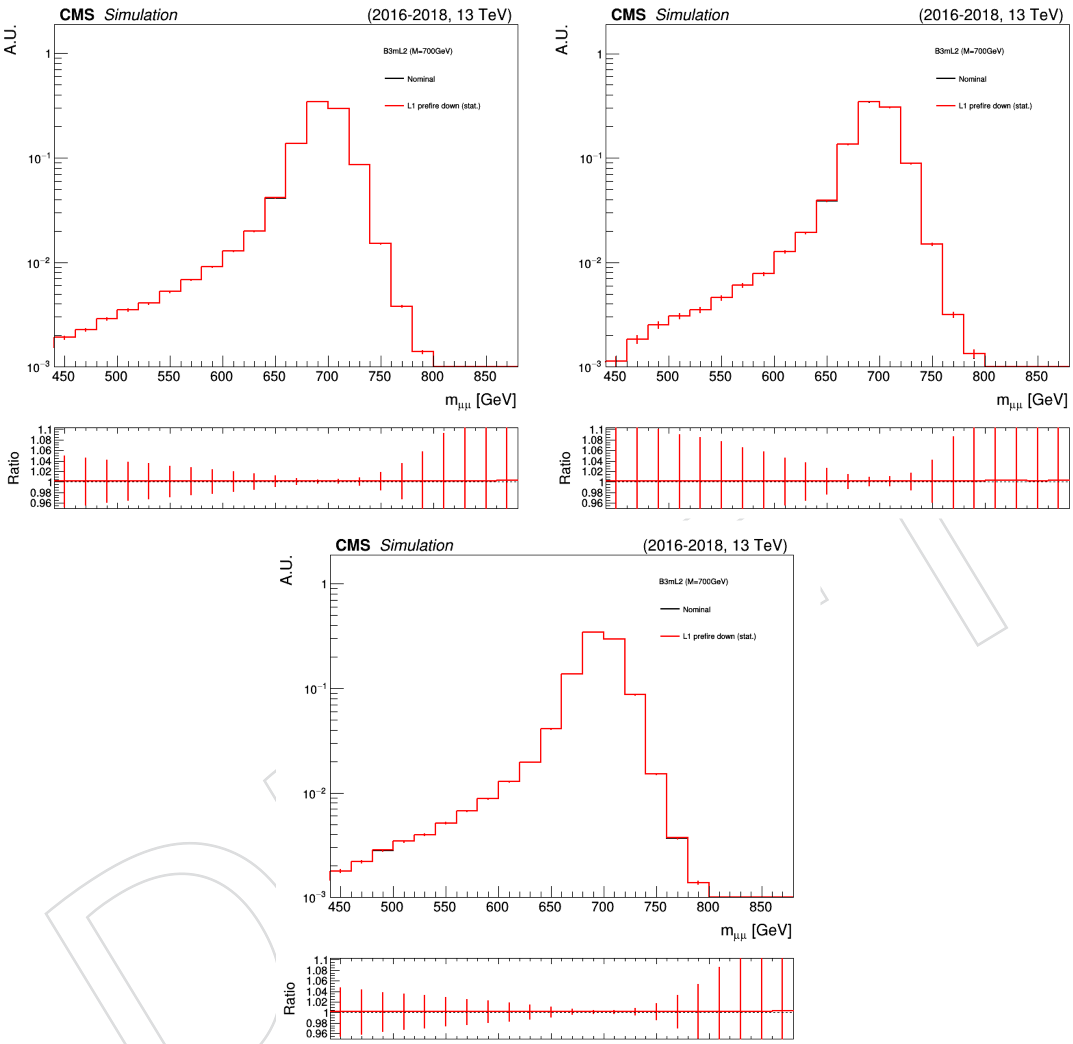


Figure 74: The effect of the propagatation of the downward statistical uncertainty in the L1 prefire weight on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 700\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the nominal L1 prefire weights. The red histogram represents the same distribution using the nominal value minus the corresponding statistical uncertainty. Both distributions are normalized to the area of the nominal histogram.

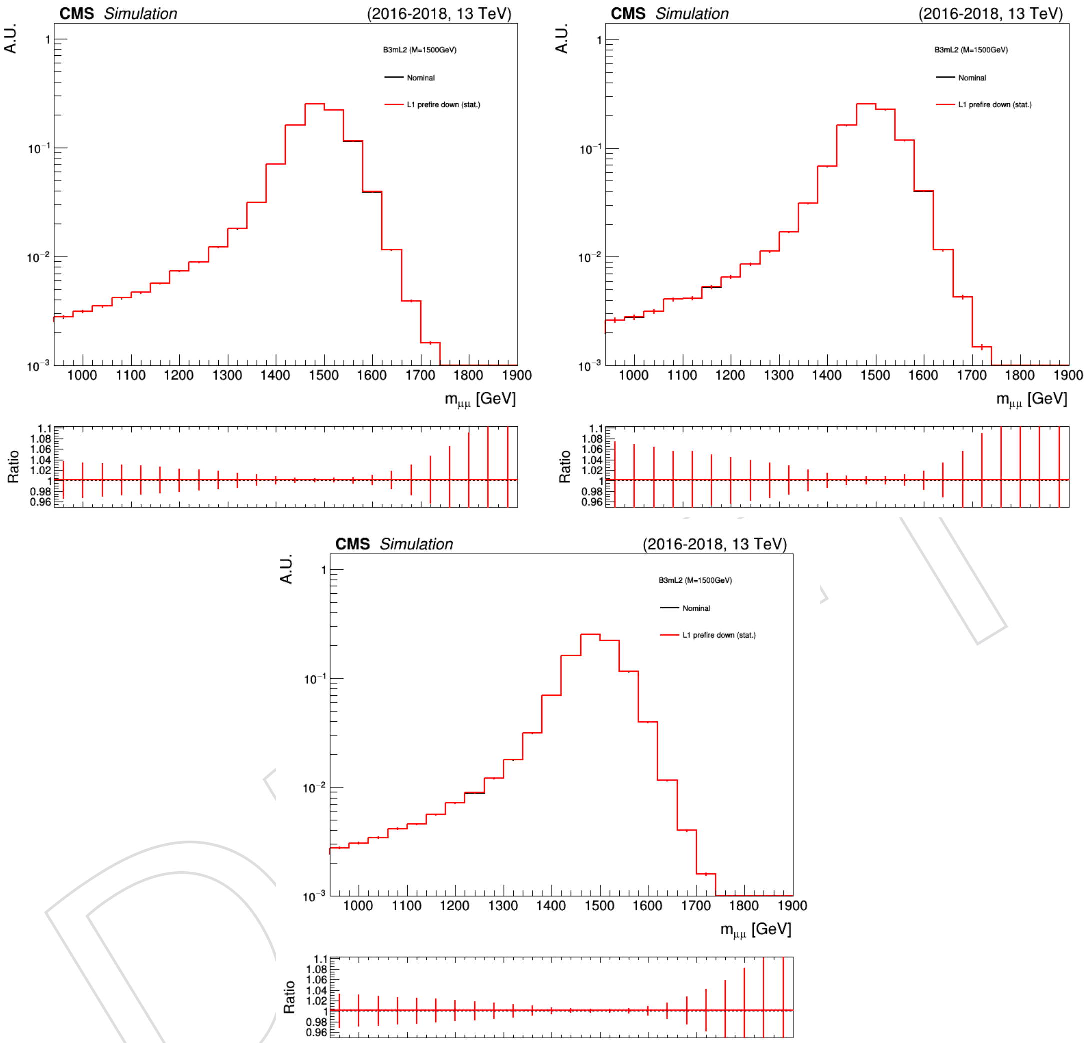


Figure 75: The effect of the propagatation of the downward statistical uncertainty in the L1 prefire weight on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 1500\,\text{GeV}$), after the full event selection, in event categories with $N_b = 1$ (top left), $N_b \geq 2$ (top right), and $N_b \geq 1$ (bottom; equal to the sum of $N_b = 1$ and $N_b \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the nominal L1 prefire weights. The red histogram represents the same distribution using the nominal value minus the corresponding statistical uncertainty. Both distributions are normalized to the area of the nominal histogram.

Figures 76–78 show the effect of the propagation of the upward systematic uncertainty in the L1 prefire weight for a few representative signal models. Similarly, Figs. 79–81 show the effect of the propagation of the downward systematic uncertainty in the L1 prefire weight for the same signal models.

The effect in the signal yield is negligible compared to the statistical uncertainties, and subdominat with respect to other systematic uncertainties. No statistically significant effect is found on the shape of the signal expected dimuon invariant mass distribution. Thus, no dedicated systematic uncertainty arising from the systematic uncertainty in the L1 prefire weight is assessed.

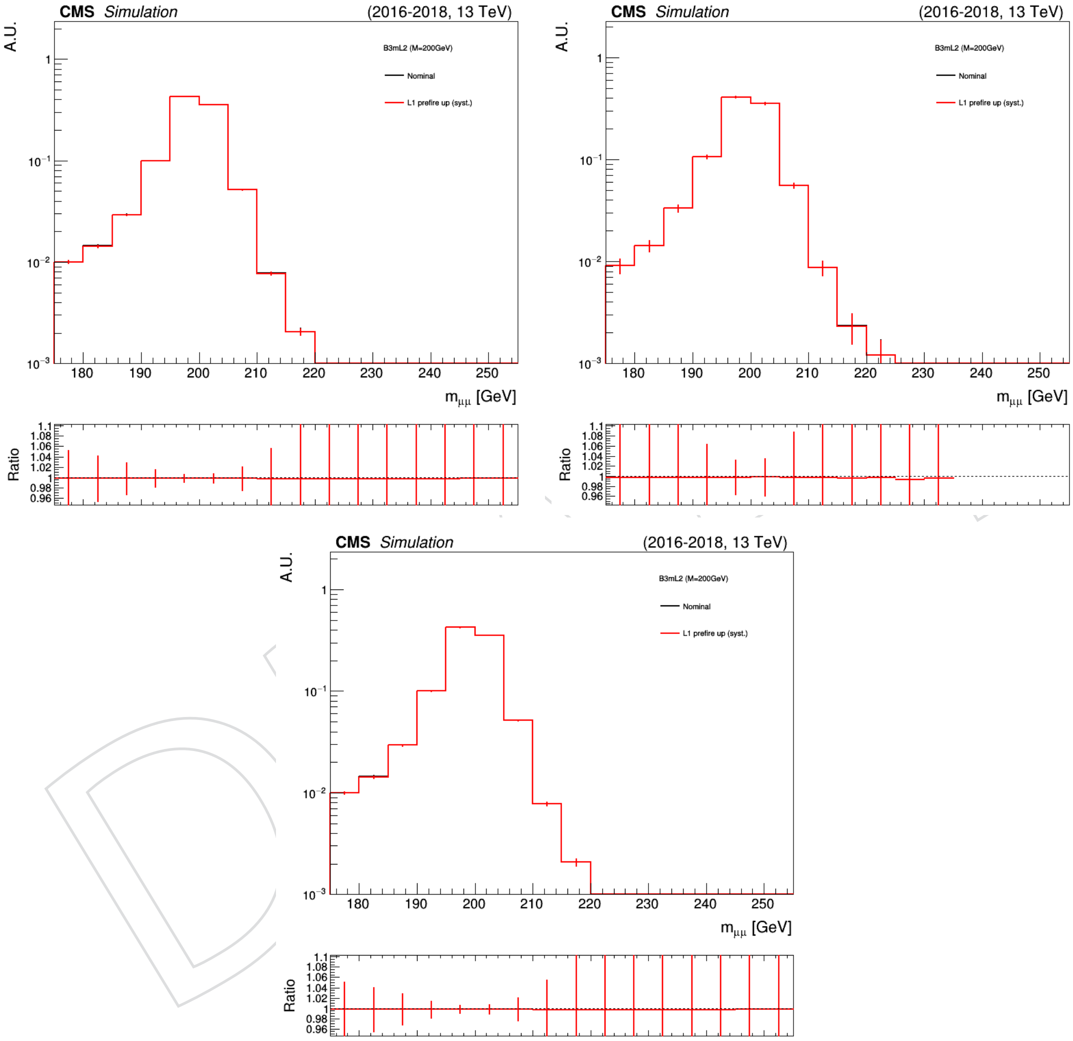


Figure 76: The effect of the propagatation of the upward systematic uncertainty in the L1 prefire weight on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 200\,\text{GeV}$), after the full event selection, in event categories with $N_{\text{b}} = 1$ (top left), $N_{\text{b}} \geq 2$ (top right), and $N_{\text{b}} \geq 1$ (bottom; equal to the sum of $N_{\text{b}} = 1$ and $N_{\text{b}} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the nominal L1 prefire weights. The red histogram represents the same distribution using the nominal value plus the corresponding systematic uncertainty. Both distributions are normalized to the area of the nominal histogram.

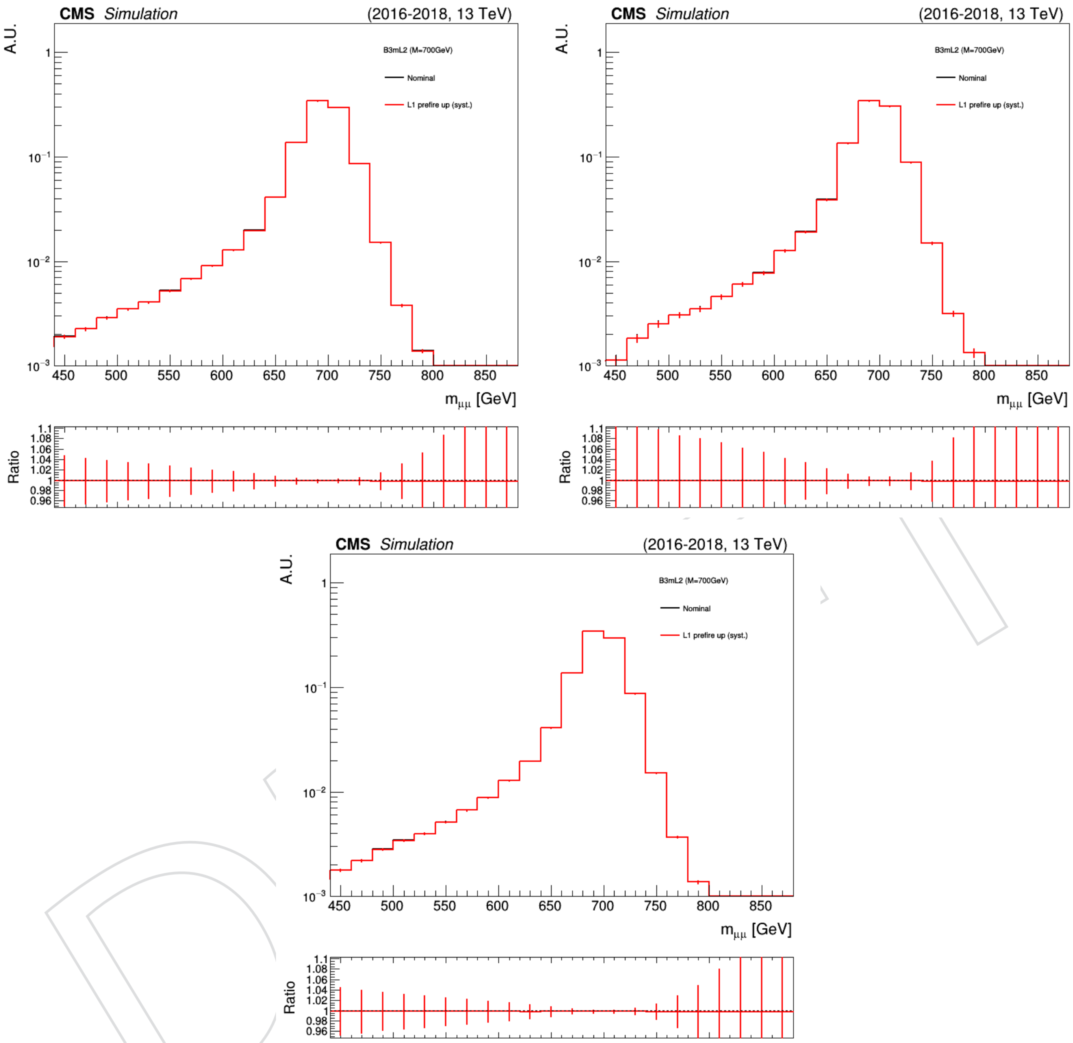


Figure 77: The effect of the propagatation of the upward systematic uncertainty in the L1 prefire weight on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 700\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the nominal L1 prefire weights. The red histogram represents the same distribution using the nominal value plus the corresponding systematic uncertainty. Both distributions are normalized to the area of the nominal histogram.

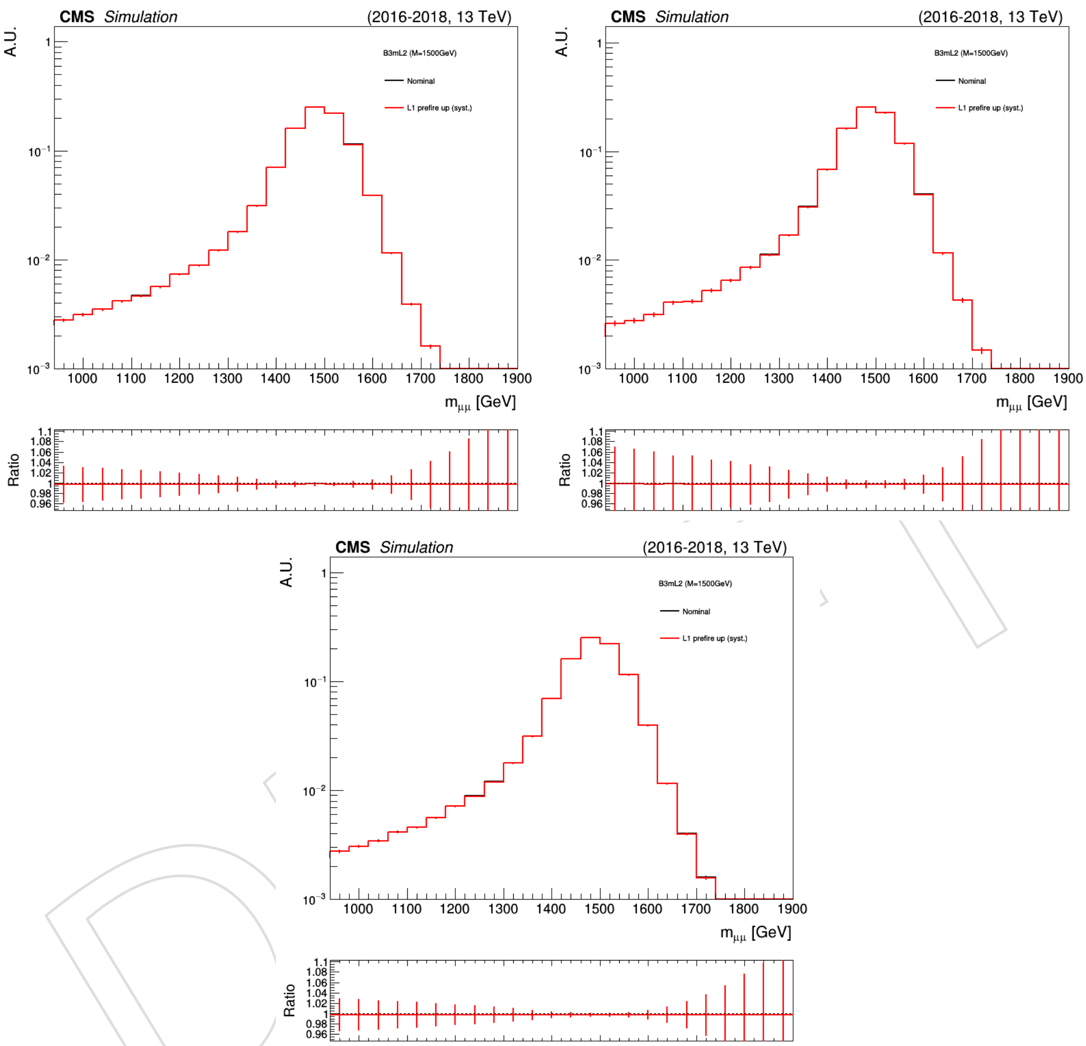


Figure 78: The effect of the propagatation of the upward systematic uncertainty in the L1 prefire weight on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 1500\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the nominal L1 prefire weights. The red histogram represents the same distribution using the nominal value plus the corresponding systematic uncertainty. Both distributions are normalized to the area of the nominal histogram.

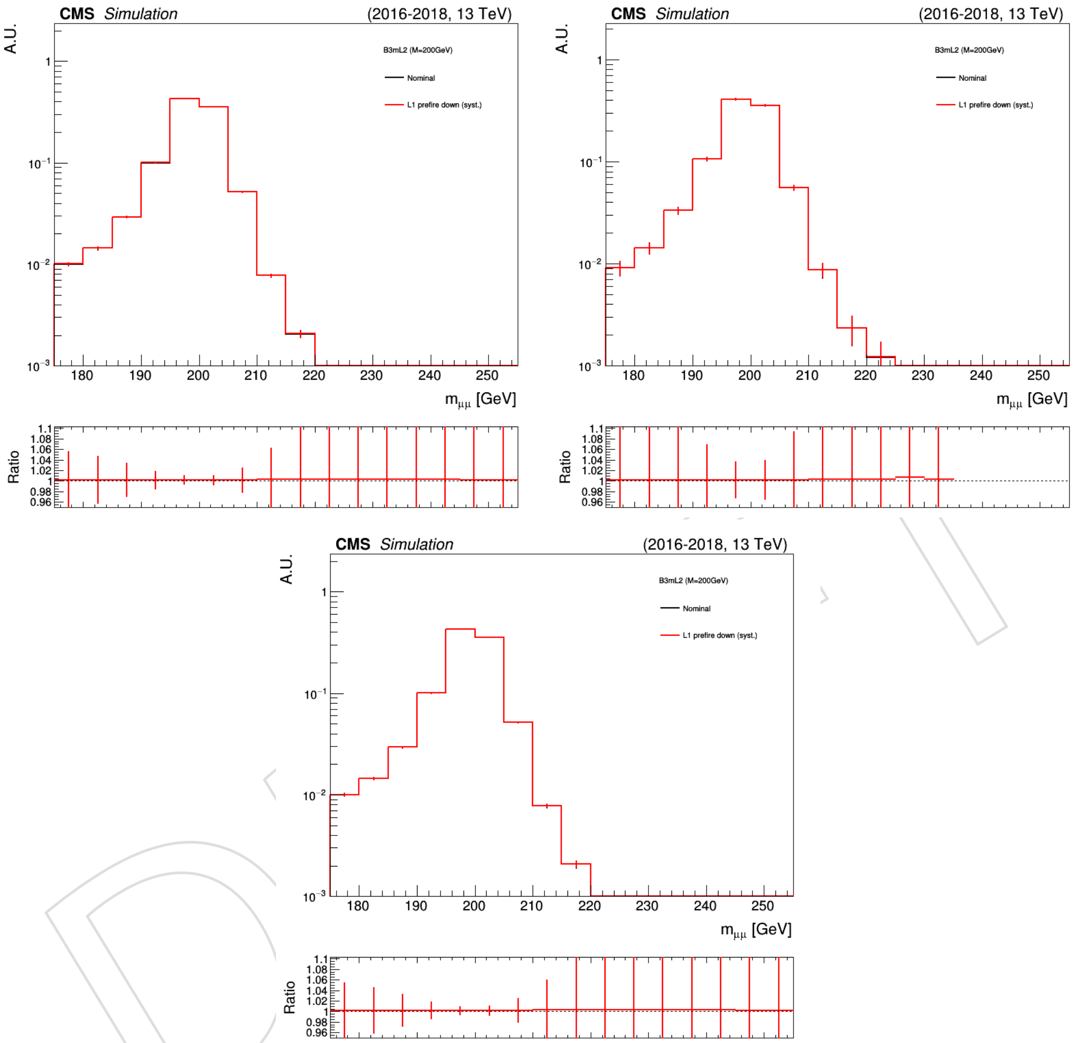


Figure 79: The effect of the propagatation of the downward systematic uncertainty in the L1 prefire weight on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 200\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the nominal L1 prefire weights. The red histogram represents the same distribution using the nominal value minus the corresponding systematic uncertainty. Both distributions are normalized to the area of the nominal histogram.

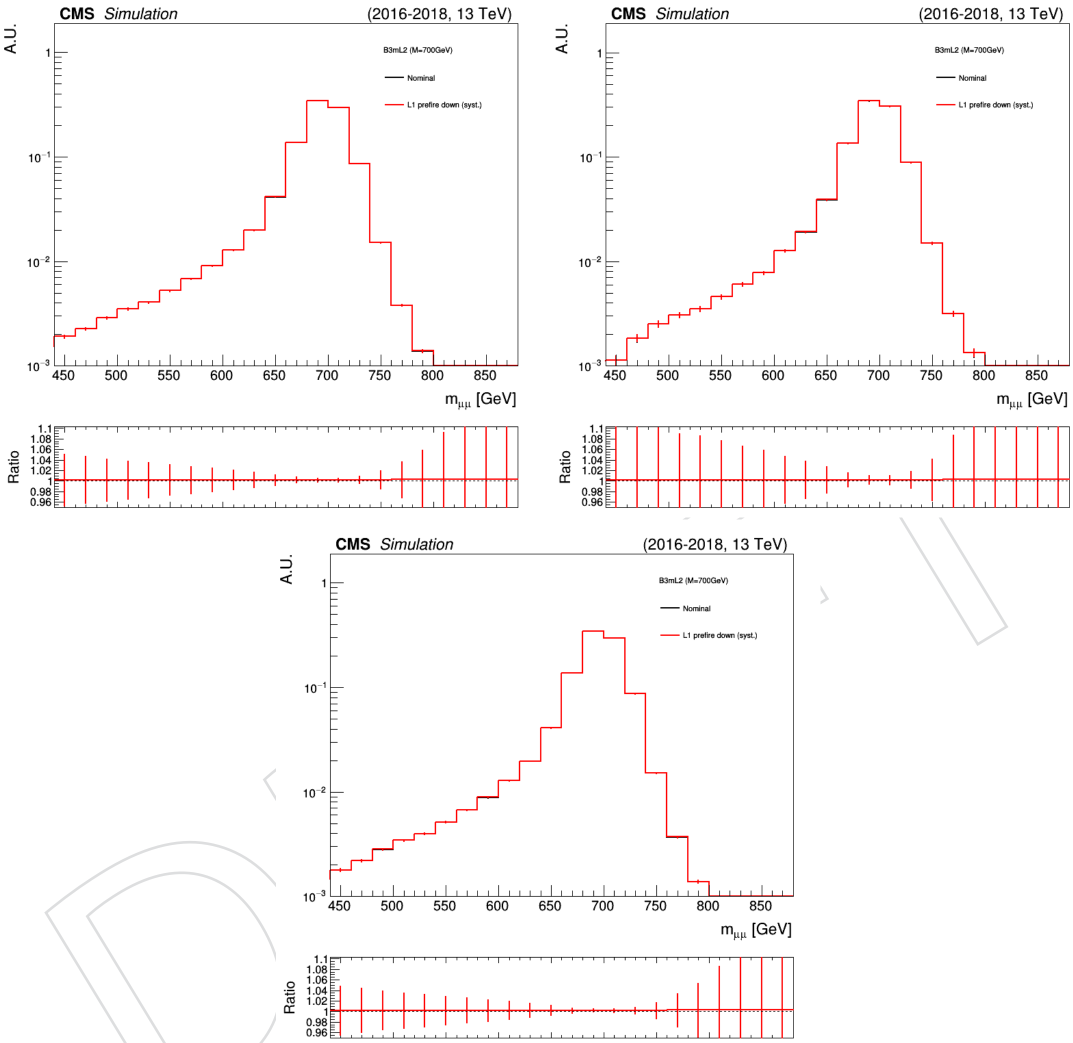


Figure 80: The effect of the propagatation of the downward systematic uncertainty in the L1 prefire weight on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 700\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the nominal L1 prefire weights. The red histogram represents the same distribution using the nominal value minus the corresponding systematic uncertainty. Both distributions are normalized to the area of the nominal histogram.

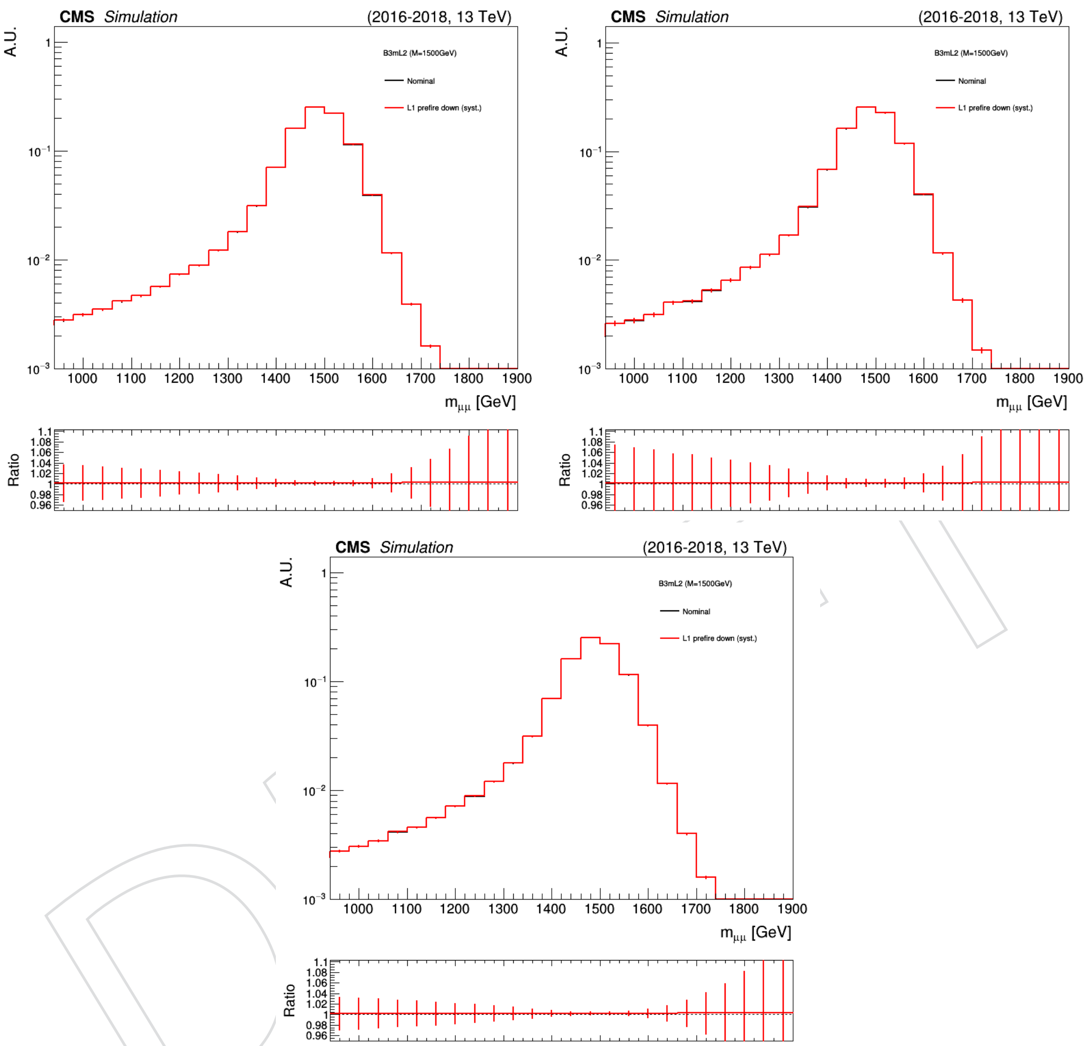


Figure 81: The effect of the propagatation of the downward systematic uncertainty in the L1 prefire weight on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 1500\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution using the nominal L1 prefire weights. The red histogram represents the same distribution using the nominal value minus the corresponding systematic uncertainty. Both distributions are normalized to the area of the nominal histogram.

E.4 Uncertainty in jet energy resolution

The impact of jet energy resolution (JER) [45] on the expected signal $m_{\mu\mu}$ distribution is shown in Figs. 82–84 for a few representative signal models. The impact is negligible compared to statistical uncertainties, showing that this analysis has no dependency on JER.

Figures 85–87 show the effect of the propagation of the upward uncertainty in the JER for a few representative signal models. Similarly, Figs. 88–90 show the effect of the propagation of the downward uncertainty in the JER for the same signal models. The effect in the signal yield is negligible compared to the statistical uncertainties, and subdominat with respect to other systematic uncertainties. No statistically significant effect is found on the shape of the signal expected dimuon invariant mass distribution. Thus, no dedicated systematic uncertainty arising from the uncertainty in the JER is assessed.

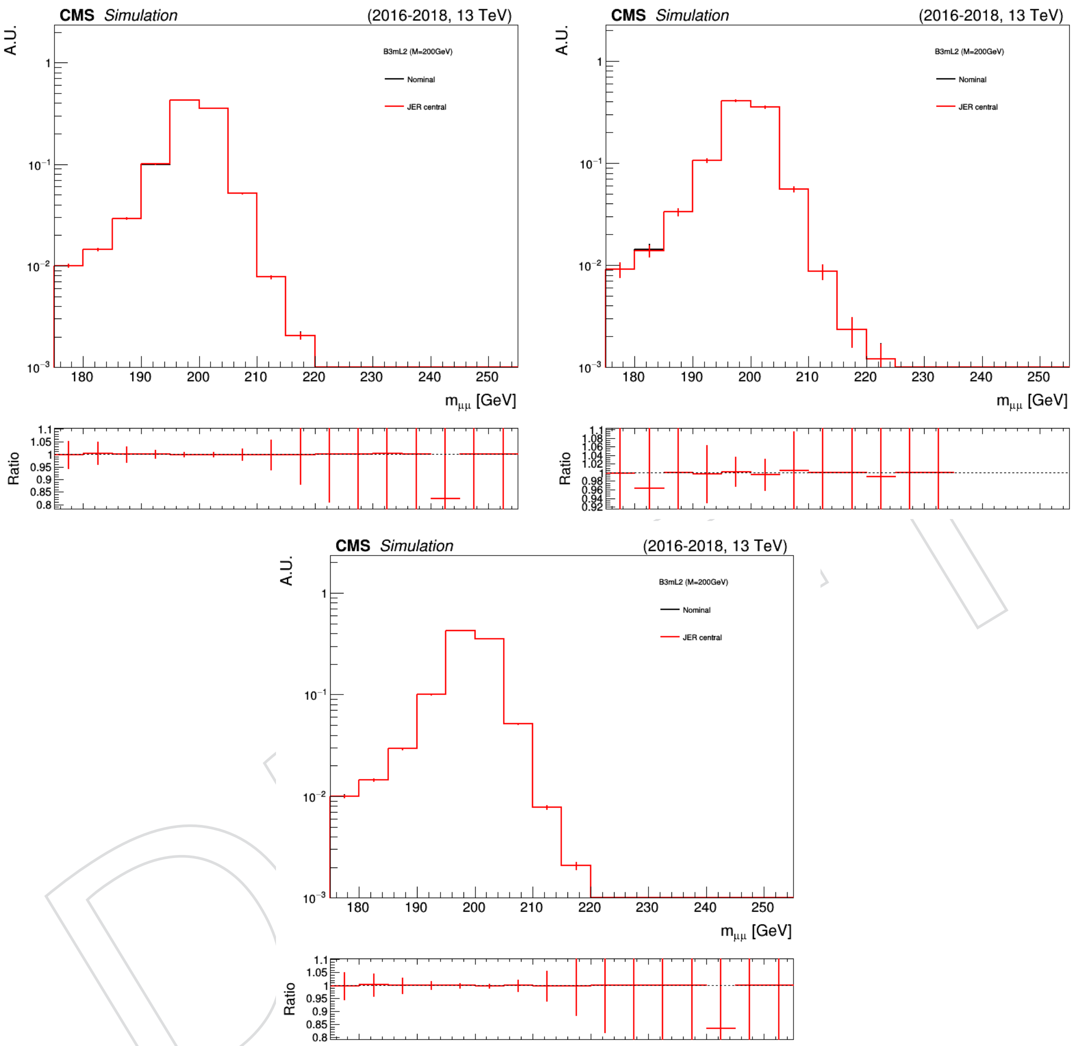


Figure 82: The effect of the propagatation of the JER on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 200\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the JER is propagated. Both distributions are normalized to the area of the nominal histogram.

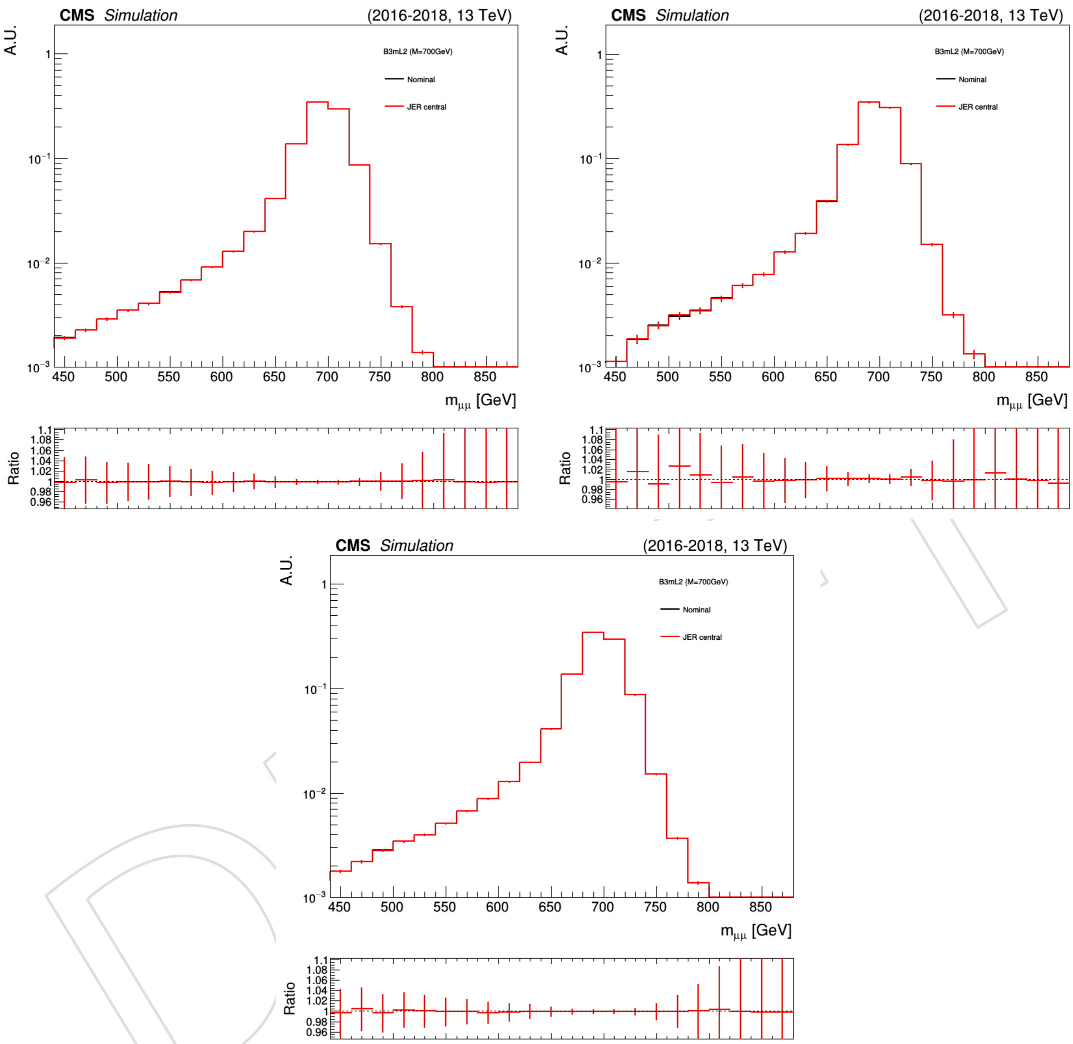


Figure 83: The effect of the propagatation of the JER on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 700\,\text{GeV}$), after the full event selection, in event categories with $N_b = 1$ (top left), $N_b \geq 2$ (top right), and $N_b \geq 1$ (bottom; equal to the sum of $N_b = 1$ and $N_b \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the JER is propagated. Both distributions are normalized to the area of the nominal histogram.

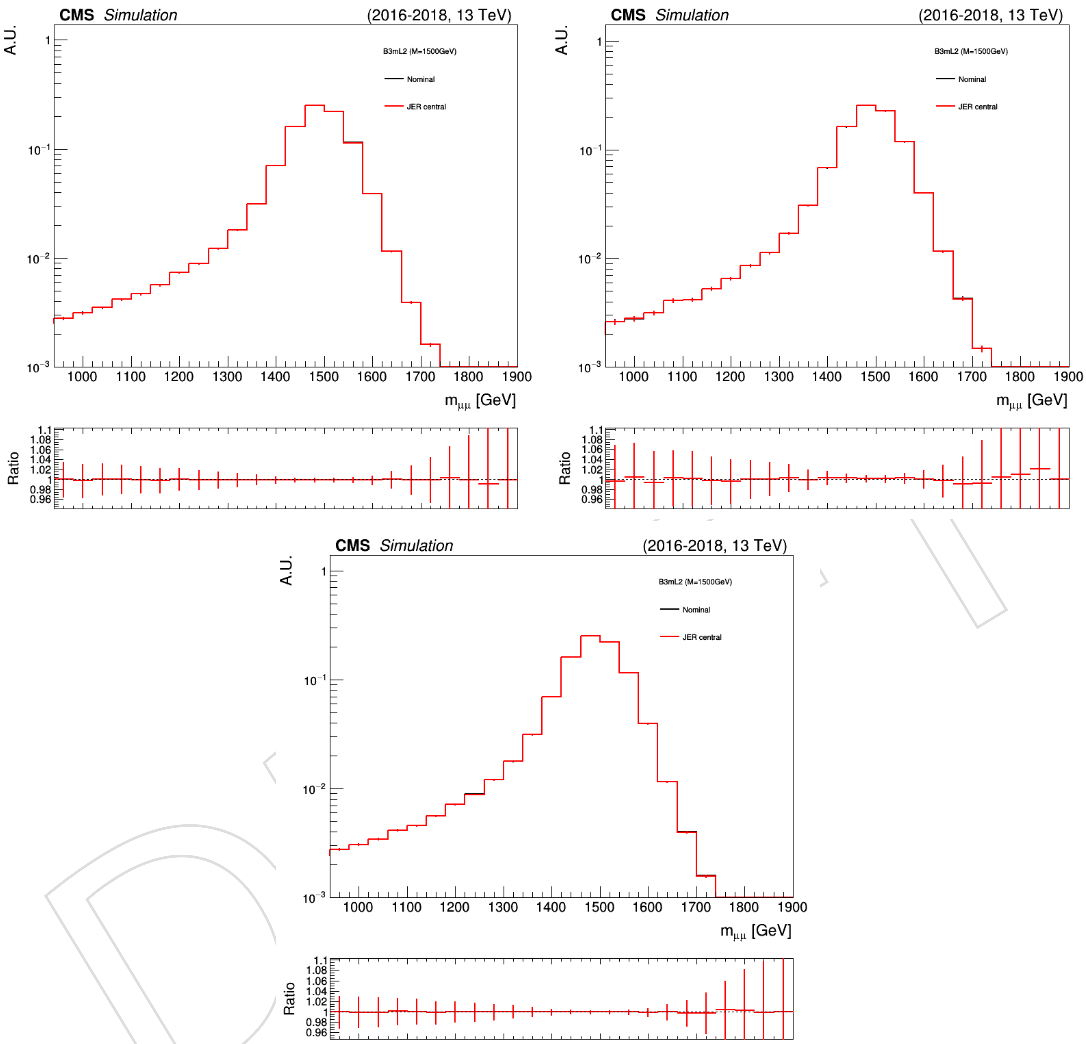


Figure 84: The effect of the propagatation of the JER on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 1500\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the JER is propagated. Both distributions are normalized to the area of the nominal histogram.

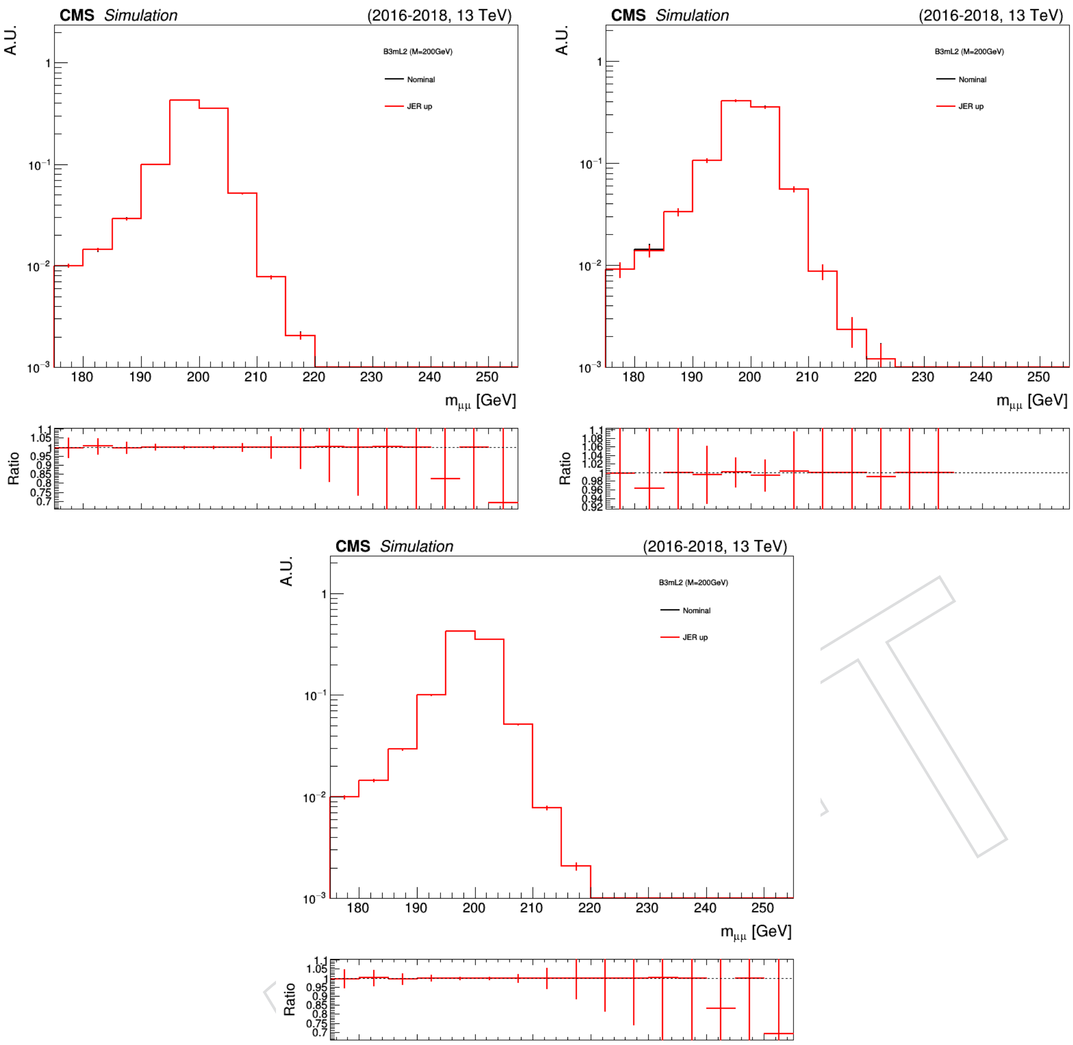


Figure 85: The effect of the propagatation of the upward uncertainty in the JER on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 200\,\text{GeV}$), after the full event selection, in event categories with $N_b = 1$ (top left), $N_b \geq 2$ (top right), and $N_b \geq 1$ (bottom; equal to the sum of $N_b = 1$ and $N_b \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the upward uncertainty in the JER is propagated. Both distributions are normalized to the area of the nominal histogram.

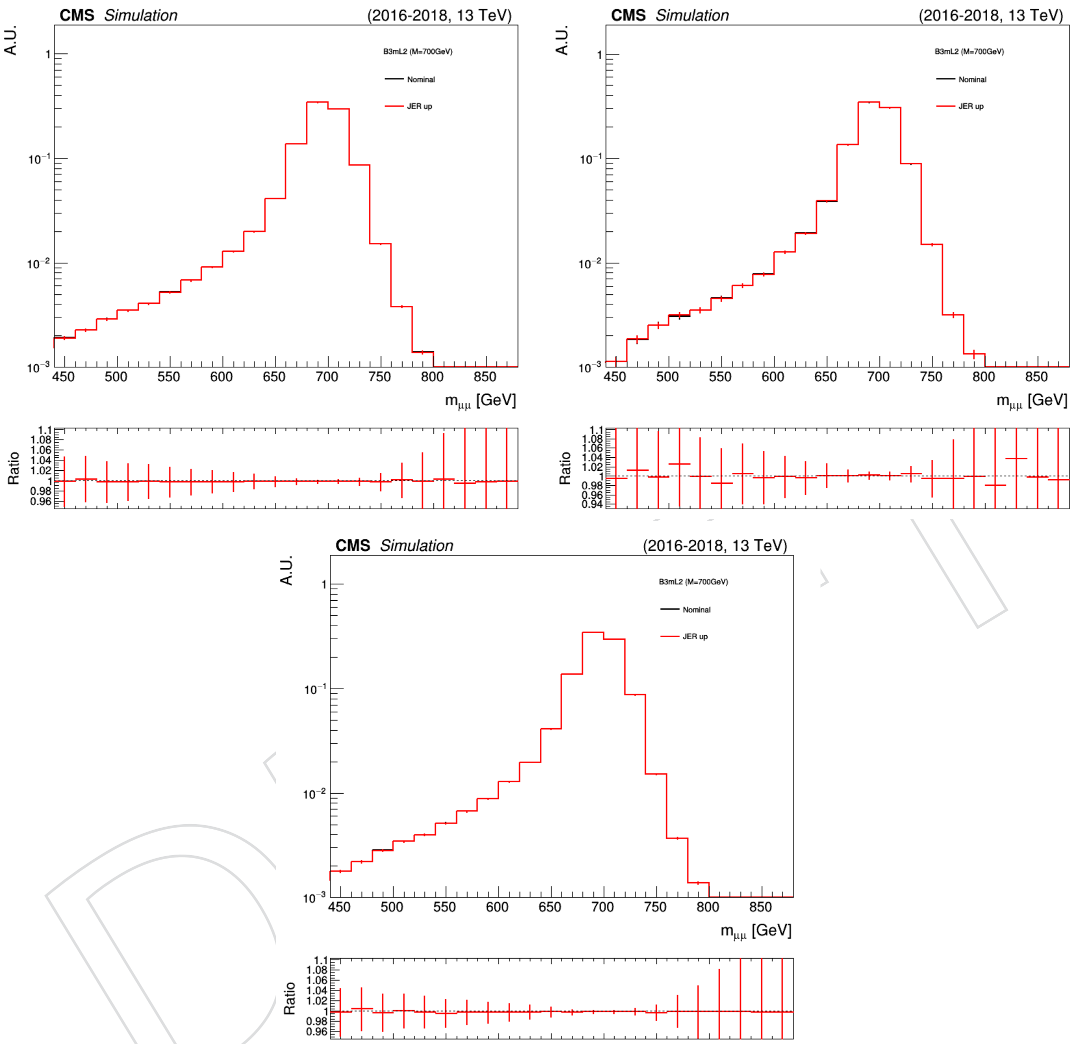


Figure 86: The effect of the propagatation of the upward uncertainty in the JER on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 700\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the upward uncertainty in the JER is propagated. Both distributions are normalized to the area of the nominal histogram.

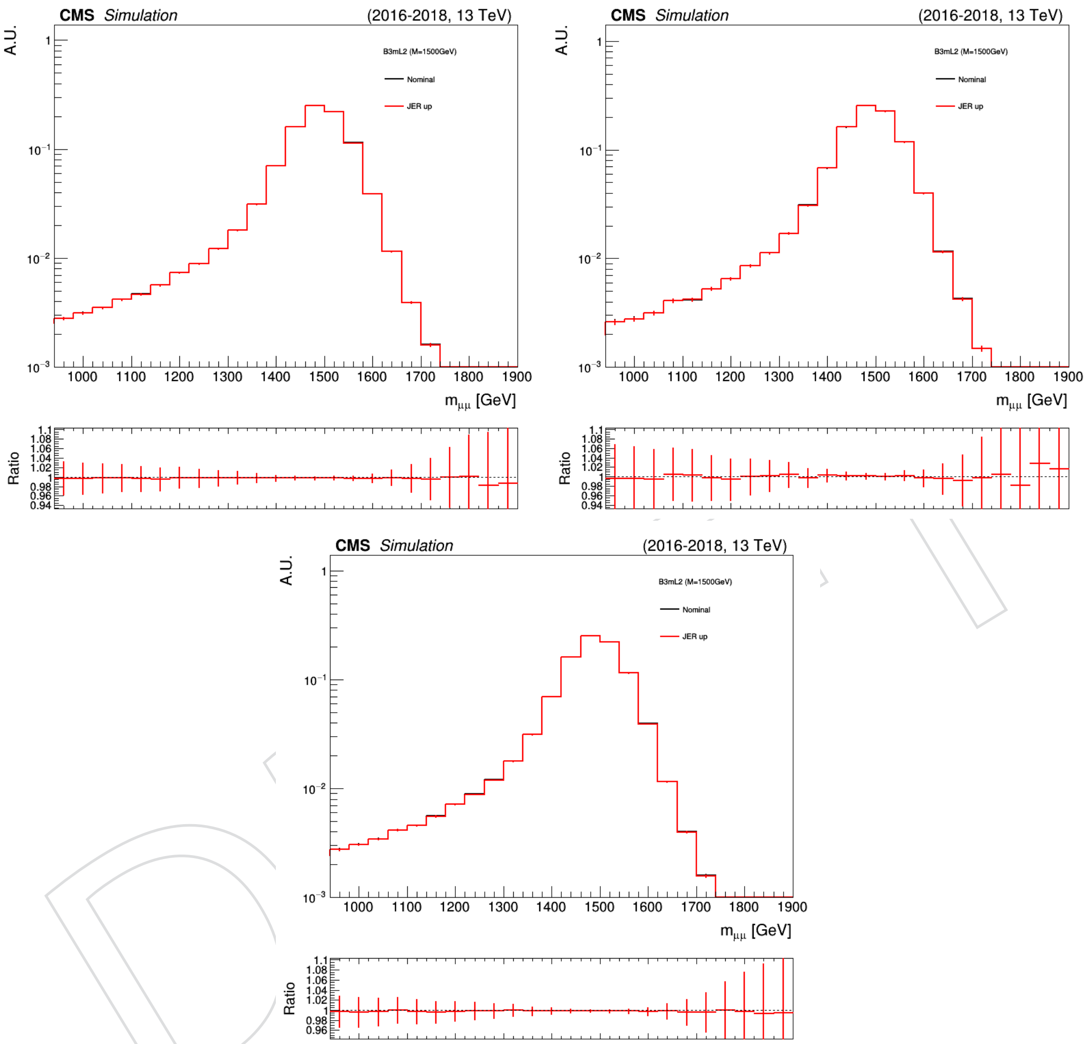


Figure 87: The effect of the propagatation of the upward uncertainty in the JER on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 1500\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the upward uncertainty in the JER is propagated. Both distributions are normalized to the area of the nominal histogram.

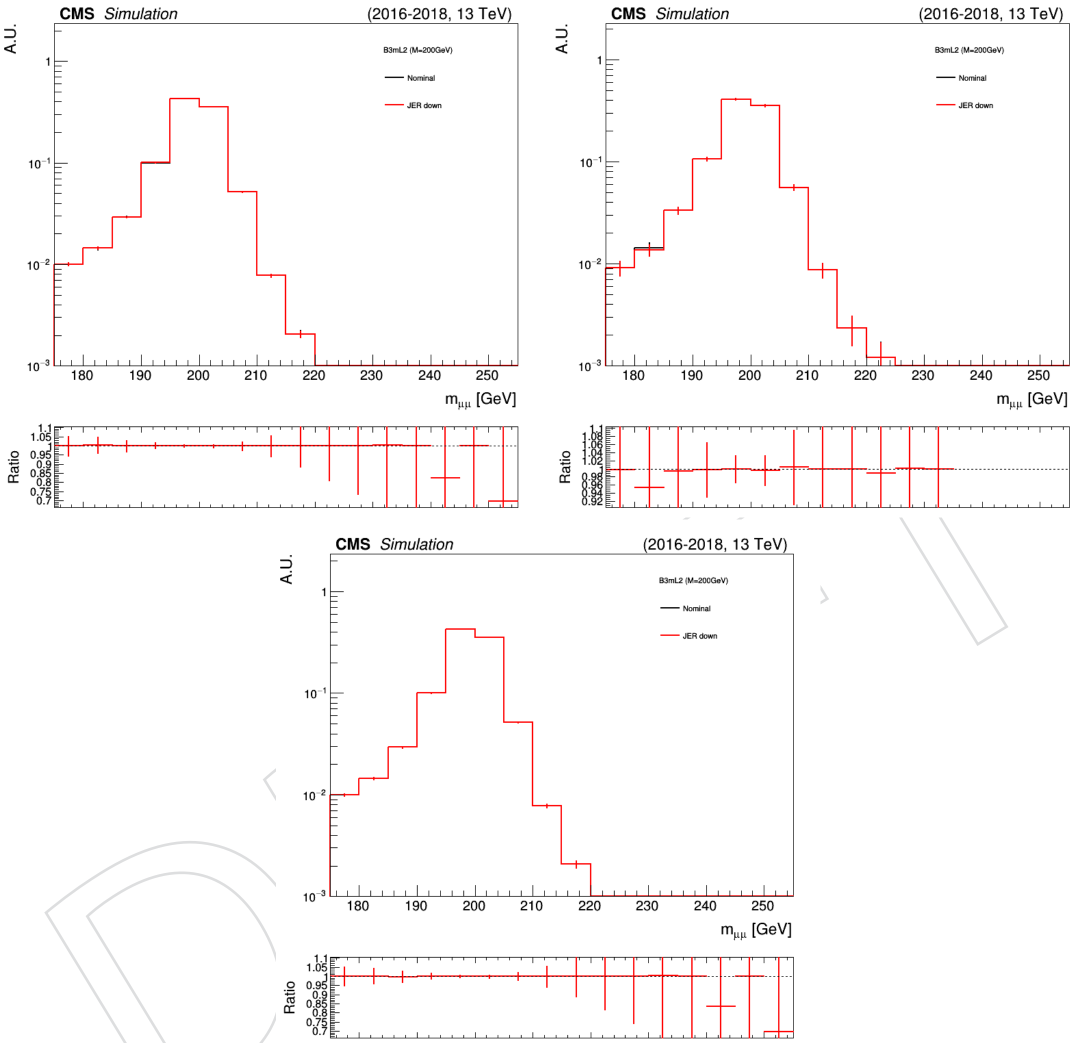


Figure 88: The effect of the propagatation of the downward uncertainty in the JER on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 200\,\mathrm{GeV}$), after the full event selection, in event categories with $N_\mathrm{b} = 1$ (top left), $N_\mathrm{b} \geq 2$ (top right), and $N_\mathrm{b} \geq 1$ (bottom; equal to the sum of $N_\mathrm{b} = 1$ and $N_\mathrm{b} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the downward uncertainty in the JER is propagated. Both distributions are normalized to the area of the nominal histogram.

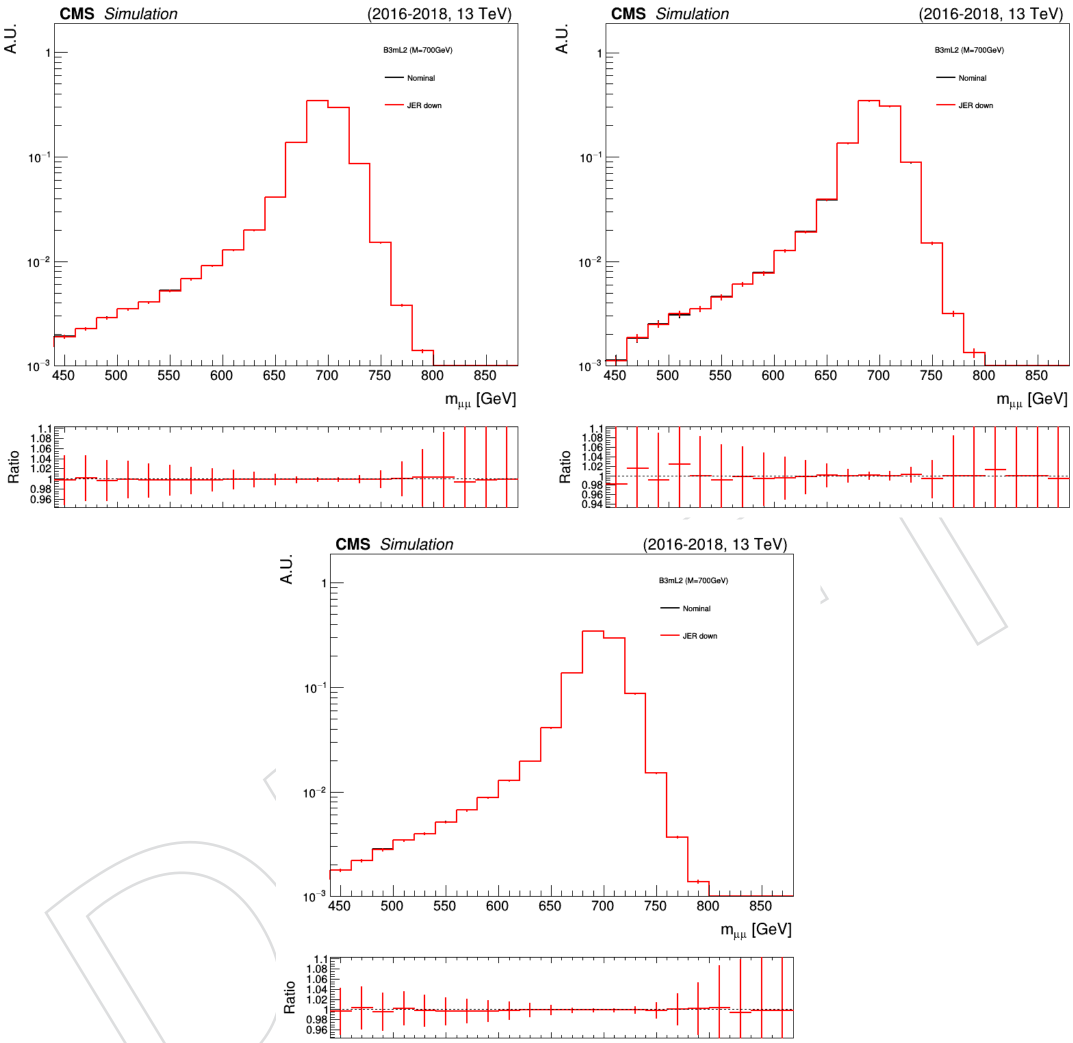


Figure 89: The effect of the propagatation of the downward uncertainty in the JER on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 700\,\mathrm{GeV}$), after the full event selection, in event categories with $N_{\mathrm{b}} = 1$ (top left), $N_{\mathrm{b}} \geq 2$ (top right), and $N_{\mathrm{b}} \geq 1$ (bottom; equal to the sum of $N_{\mathrm{b}} = 1$ and $N_{\mathrm{b}} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the downward uncertainty in the JER is propagated. Both distributions are normalized to the area of the nominal histogram.

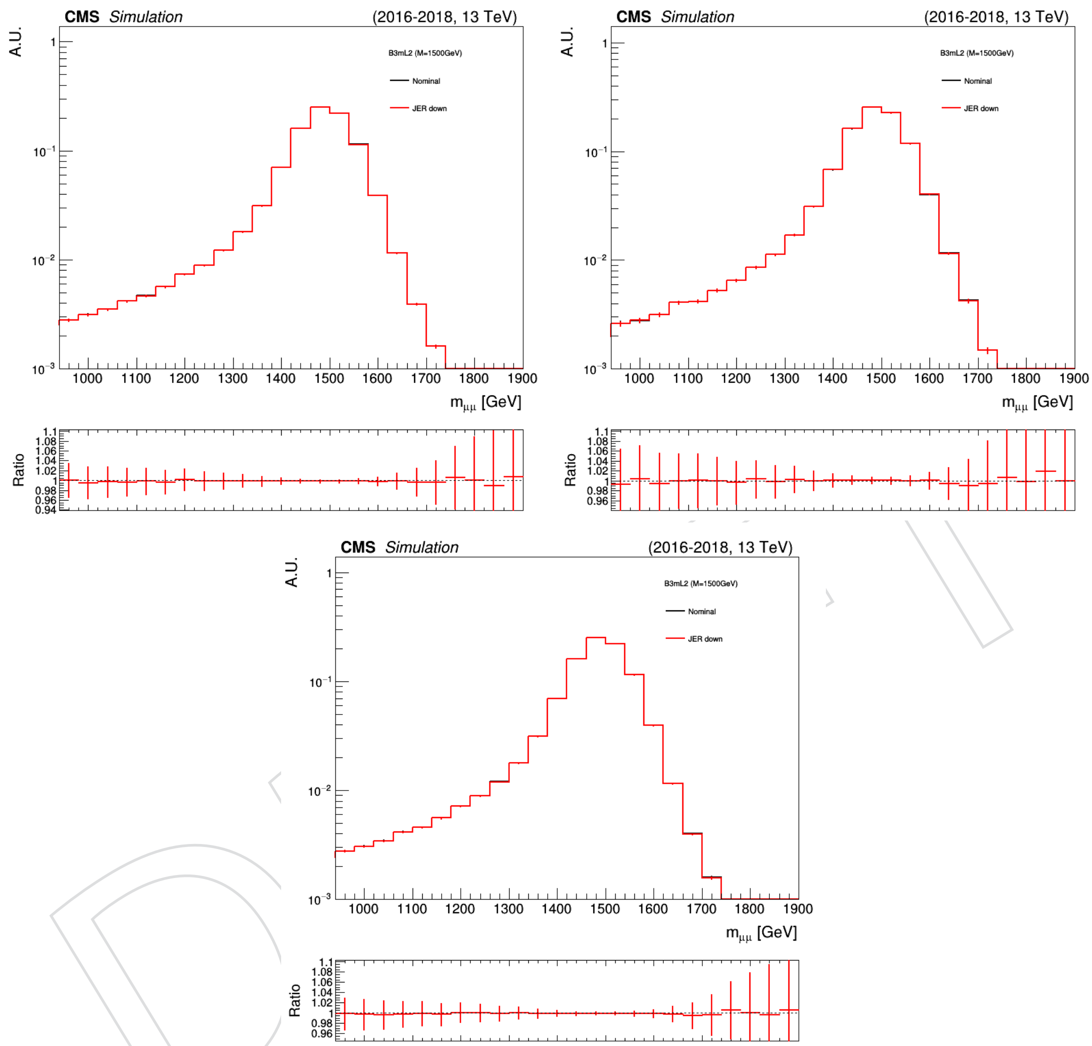


Figure 90: The effect of the propagatation of the downward uncertainty in the JER on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 1500\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the downward uncertainty in the JER is propagated. Both distributions are normalized to the area of the nominal histogram.

E.5 Uncertainty in jet energy scale

Figures 91–93 show the effect of the propagation of the upward uncertainty in the JES [44] for a few representative signal models. Similarly, Figs. 94–96 show the effect of the propagation of the downward uncertainty in the JES for the same signal models.

The effect in the signal yield is decreasing at increasing $m_{Z'}$ and is larger in the event category with $N_\mathrm{b} \geq 2$:

- in the category with $N_\mathrm{b} = 1$, the effect in the expected signal yield is ranging from 1.5% at $m_{Z'} = 200\,\mathrm{GeV}$ to less than 1% at $m_{Z'} = 2000\,\mathrm{GeV}$;
- in the category with $N_\mathrm{b} \geq 2$, the effect in the expected signal yield is ranging from 5% at $m_{Z'} = 200\,\mathrm{GeV}$ to less than 2% at $m_{Z'} = 2000\,\mathrm{GeV}$.

No statistically significant effect is found on the shape of the signal expected dimuon invariant mass distribution. Thus, a dedicated systematic uncertainty in the expected signal yield is assessed.

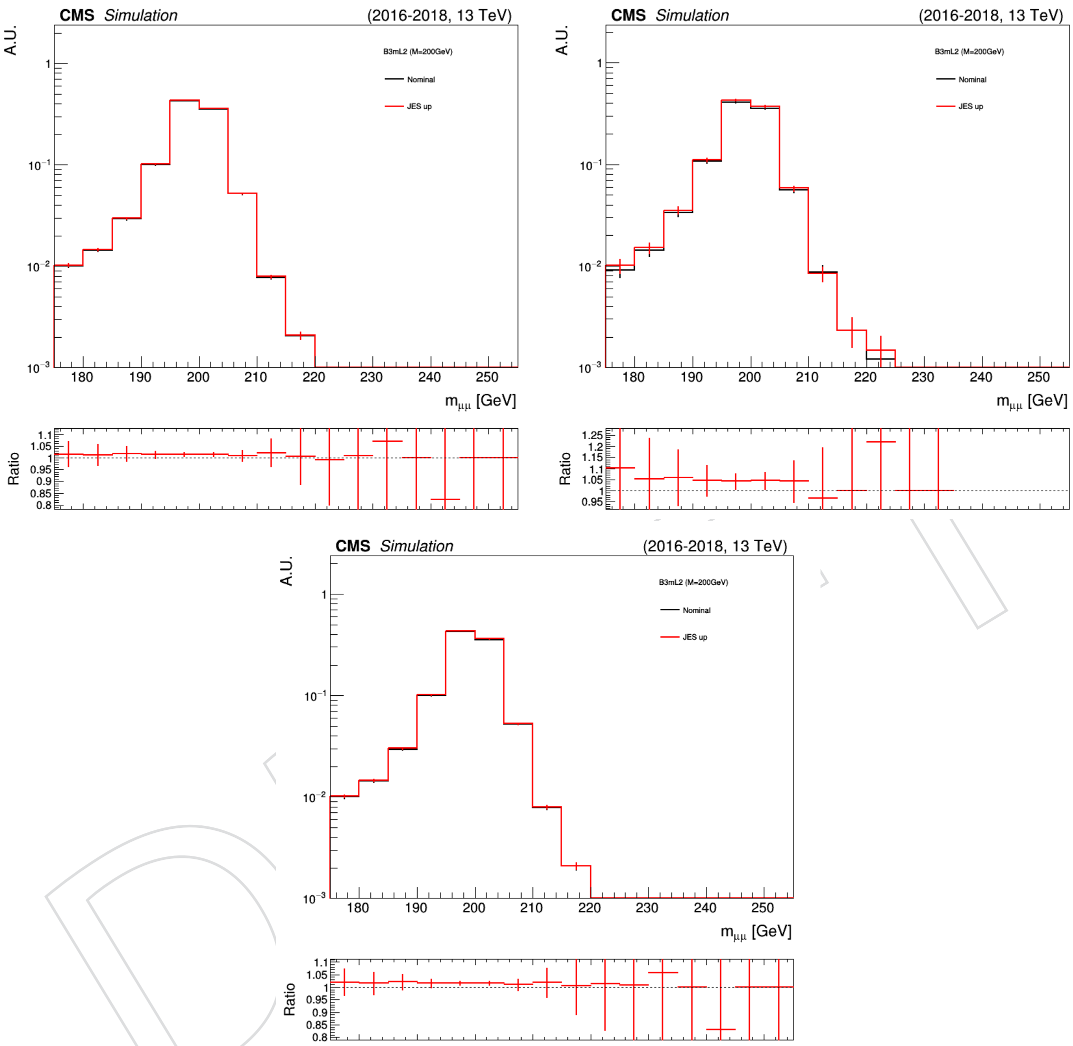


Figure 91: The effect of the propagatation of the upward uncertainty in the JES on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 200\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the upward uncertainty in the JES is propagated. Both distributions are normalized to the area of the nominal histogram.

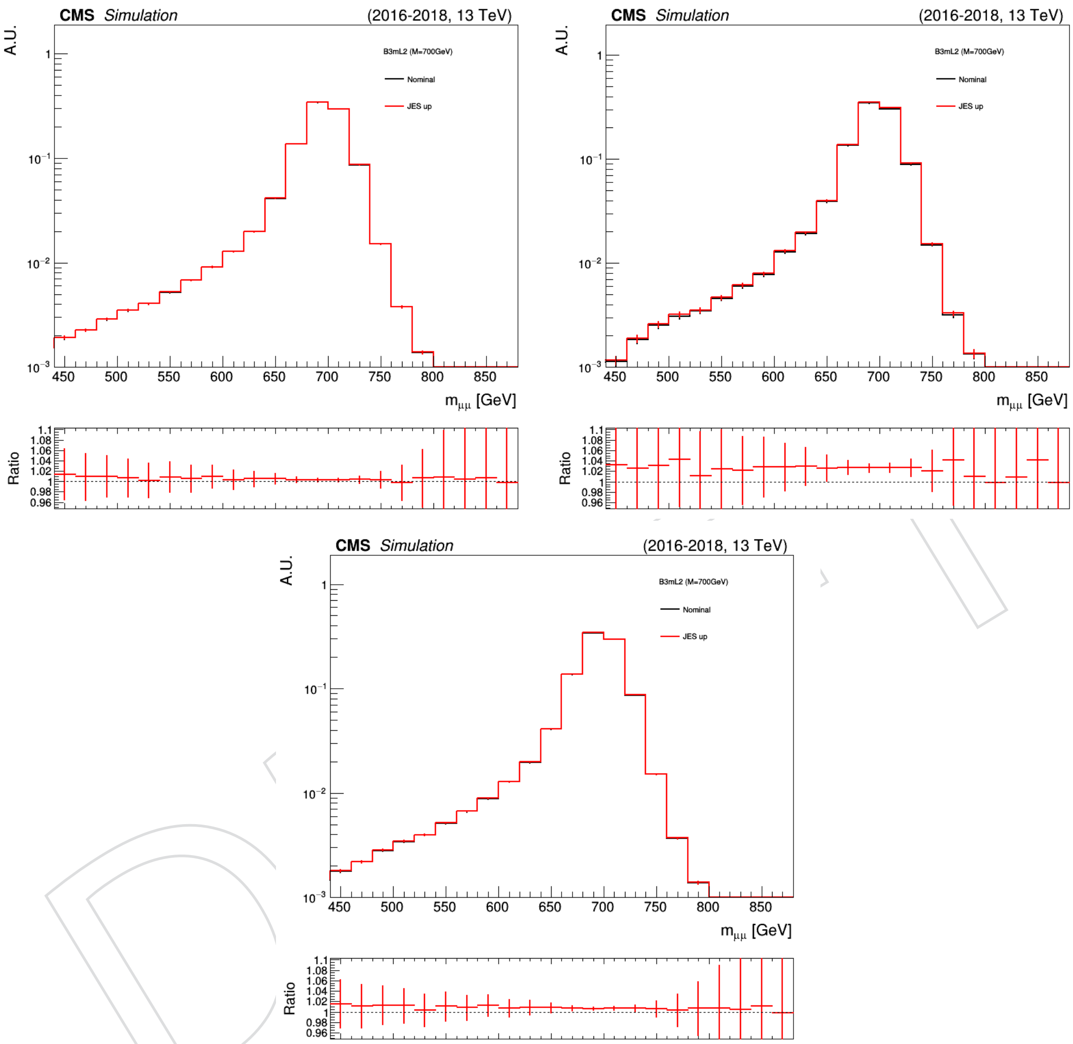


Figure 92: The effect of the propagatation of the upward uncertainty in the JES on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 700\,\text{GeV}$), after the full event selection, in event categories with $N_{\text{b}} = 1$ (top left), $N_{\text{b}} \geq 2$ (top right), and $N_{\text{b}} \geq 1$ (bottom; equal to the sum of $N_{\text{b}} = 1$ and $N_{\text{b}} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the upward uncertainty in the JES is propagated. Both distributions are normalized to the area of the nominal histogram.

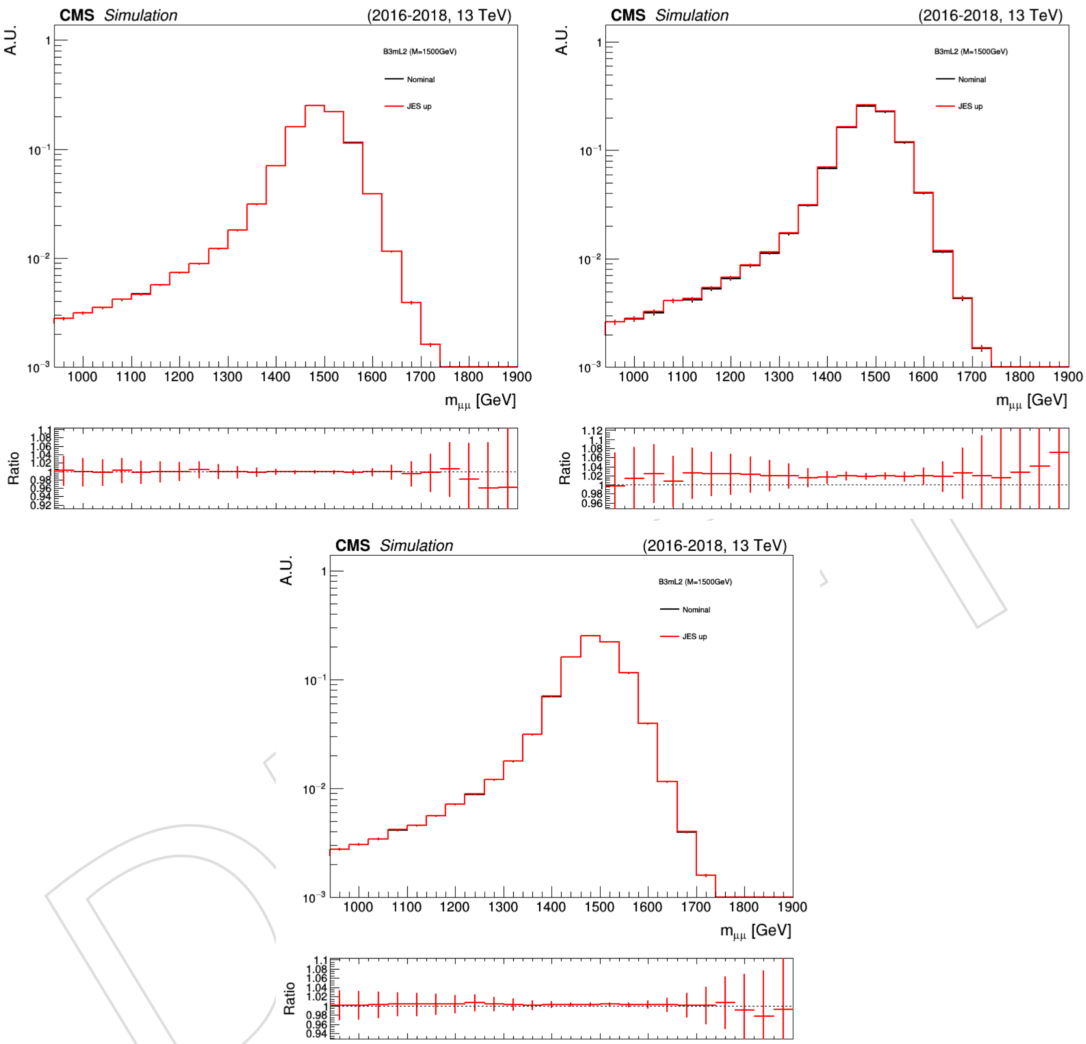


Figure 93: The effect of the propagatation of the upward uncertainty in the JES on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 1500\,\mathrm{GeV}$), after the full event selection, in event categories with $N_{\mathrm{b}} = 1$ (top left), $N_{\mathrm{b}} \geq 2$ (top right), and $N_{\mathrm{b}} \geq 1$ (bottom; equal to the sum of $N_{\mathrm{b}} = 1$ and $N_{\mathrm{b}} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the upward uncertainty in the JES is propagated. Both distributions are normalized to the area of the nominal histogram.

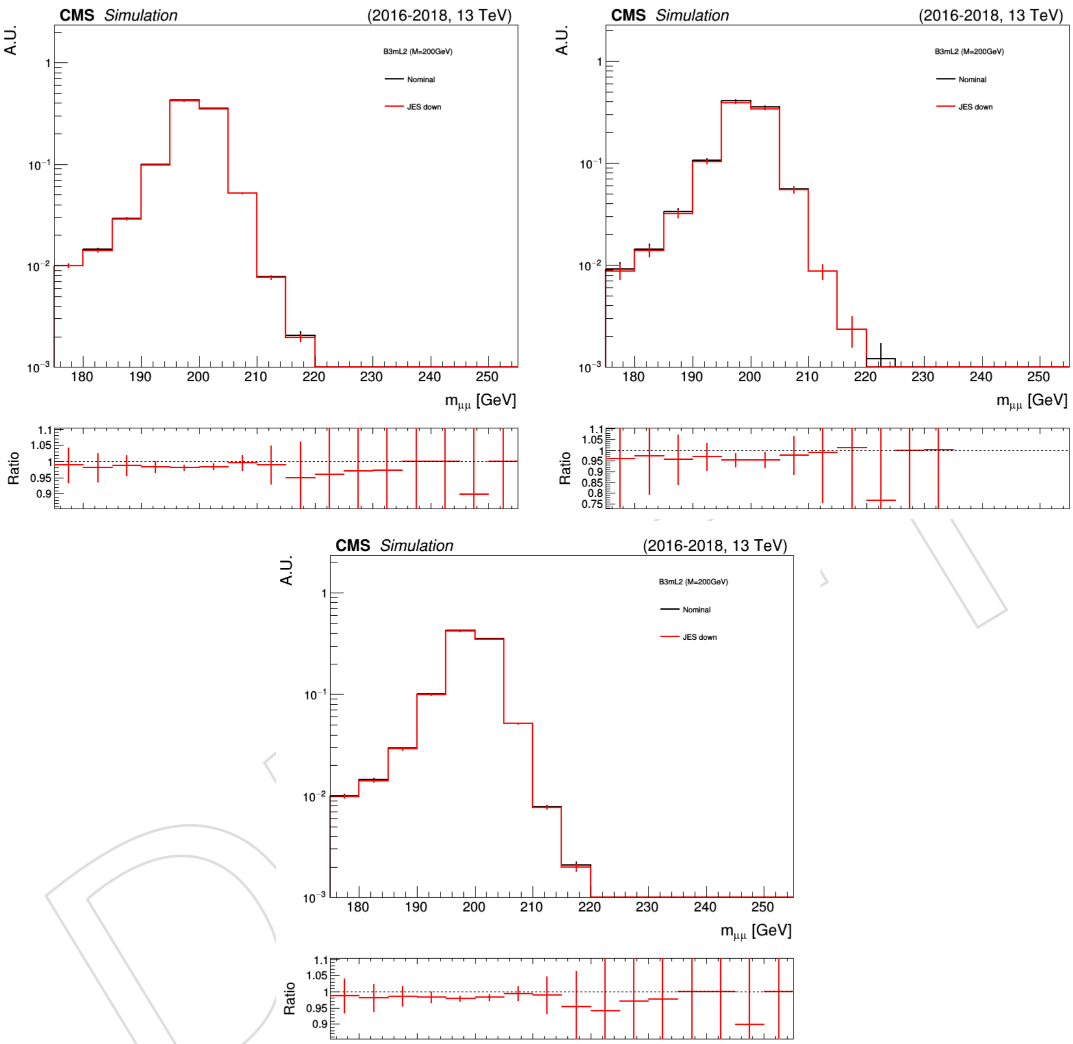


Figure 94: The effect of the propagatation of the downward uncertainty in the JES on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 200\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the downward uncertainty in the JES is propagated. Both distributions are normalized to the area of the nominal histogram.

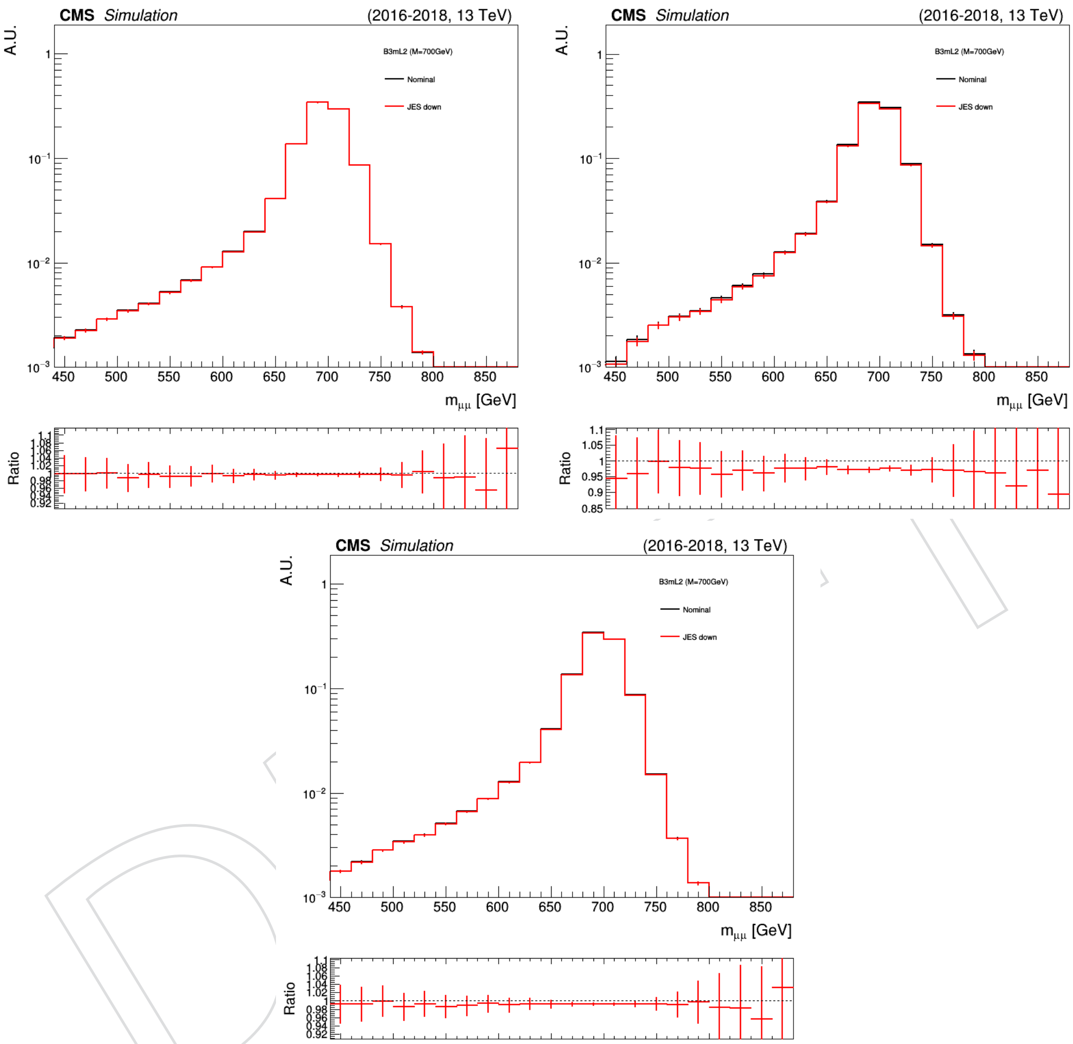


Figure 95: The effect of the propagatation of the downward uncertainty in the JES on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 700\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the downward uncertainty in the JES is propagated. Both distributions are normalized to the area of the nominal histogram.

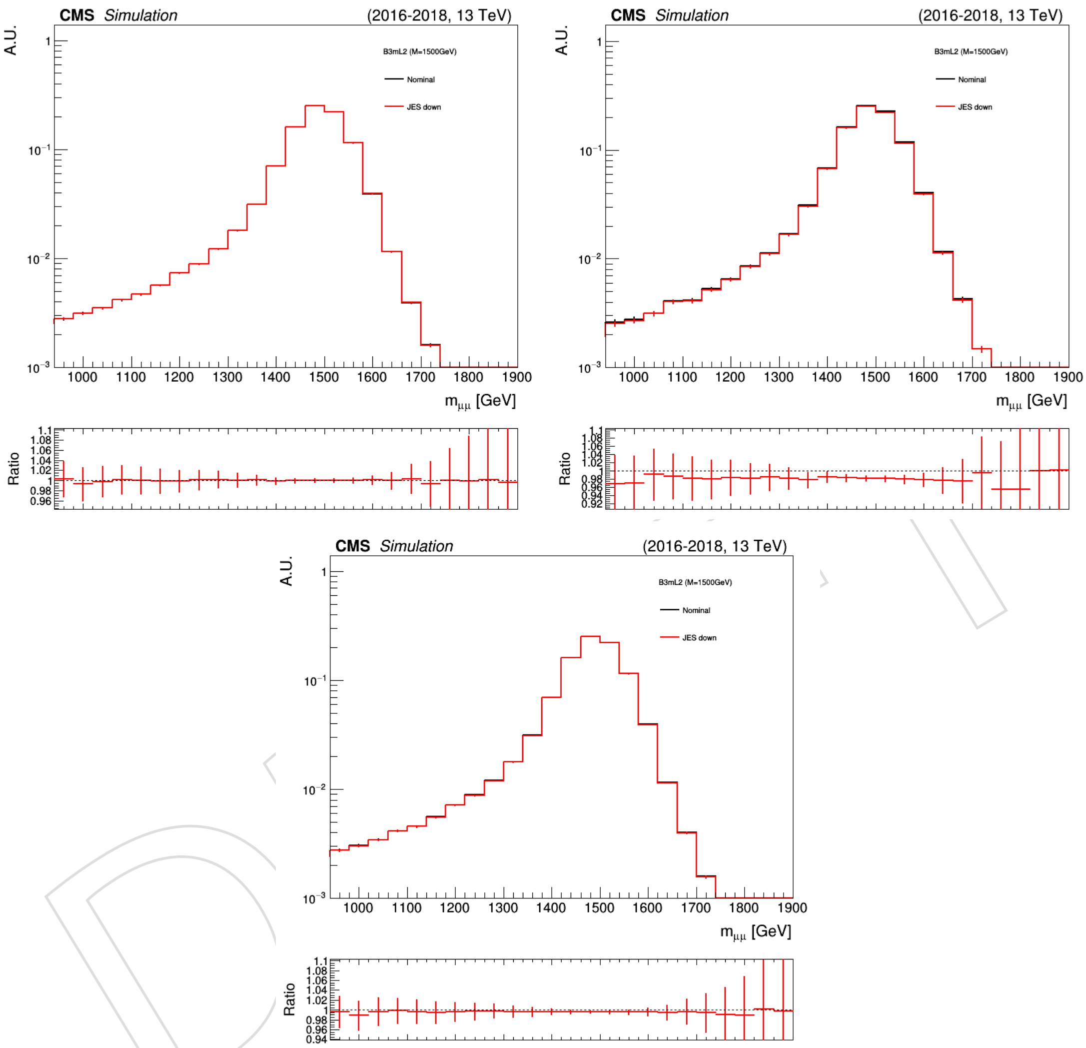


Figure 96: The effect of the propagatation of the downward uncertainty in the JES on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 1500\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the downward uncertainty in the JES is propagated. Both distributions are normalized to the area of the nominal histogram.

E.6 Uncertainty in b-tagging efficiency

Figures 97–99 show the effect of the propagation of the upward uncertainty in the b-tagging data/MC scale factors [43] for a few representative signal models. Similarly, Figs. 100–102 show the effect of the propagation of the downward uncertainty in the b-tagging data/MC scale factors for the same signal models.

The effect in the signal yield is independent of $m_{Z'}$ and is larger in the event category with $N_{\mathrm{b}} \geq 2$:

- in the category with $N_{\mathrm{b}} = 1$, the effect in the expected signal yield is $\lesssim 1\%$;
- in the category with $N_{\mathrm{b}} \geq 2$, the effect in the expected signal yield is $\lesssim 5\%$.

No statistically significant effect is found on the shape of the signal expected dimuon invariant mass distribution. Thus, a dedicated systematic uncertainty in the expected signal yield is assessed.

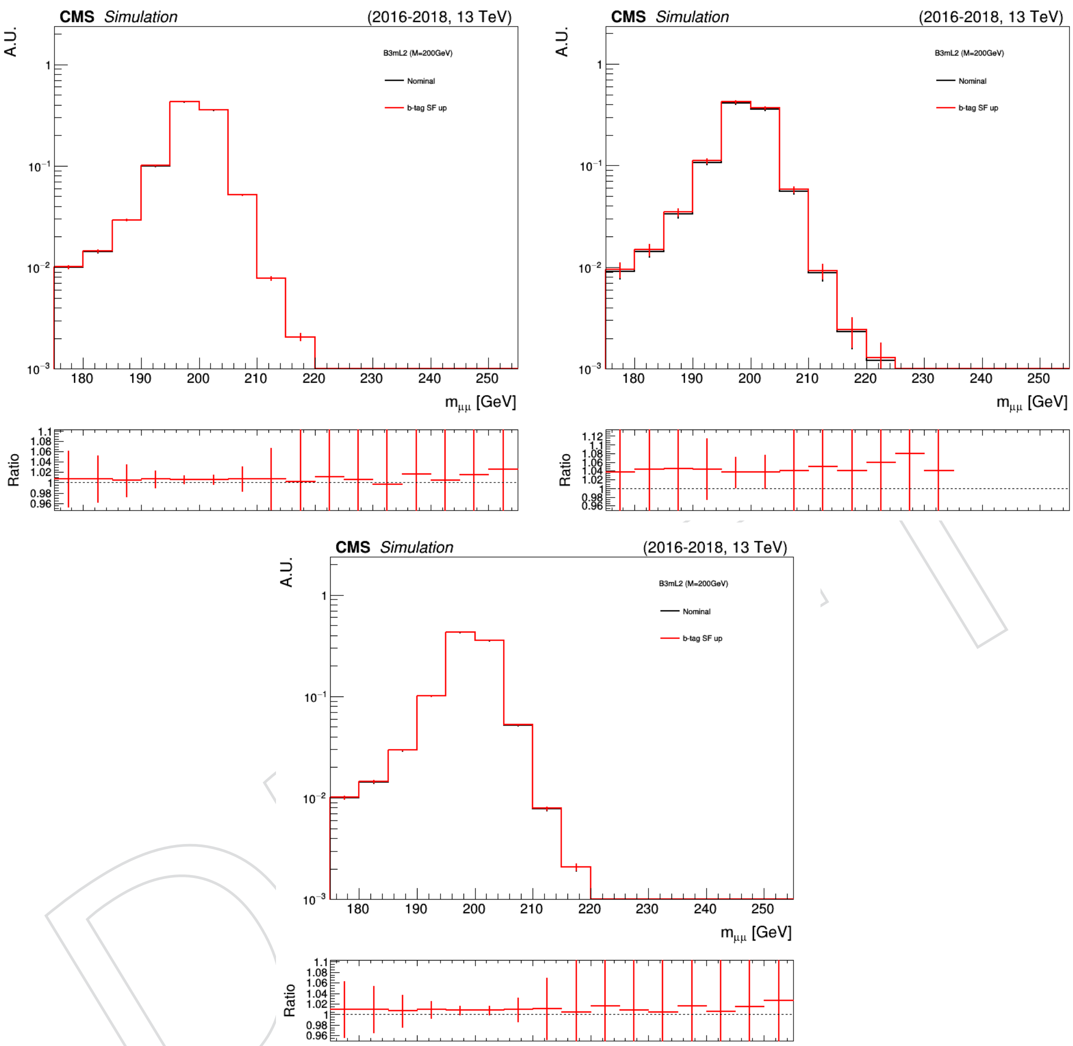


Figure 97: The effect of the propagatation of the upward uncertainty in the b-tagging data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 200\,\text{GeV}$), after the full event selection, in event categories with $N_{\text{b}} = 1$ (top left), $N_{\text{b}} \geq 2$ (top right), and $N_{\text{b}} \geq 1$ (bottom; equal to the sum of $N_{\text{b}} = 1$ and $N_{\text{b}} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the upward uncertainty in the b-tagging data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

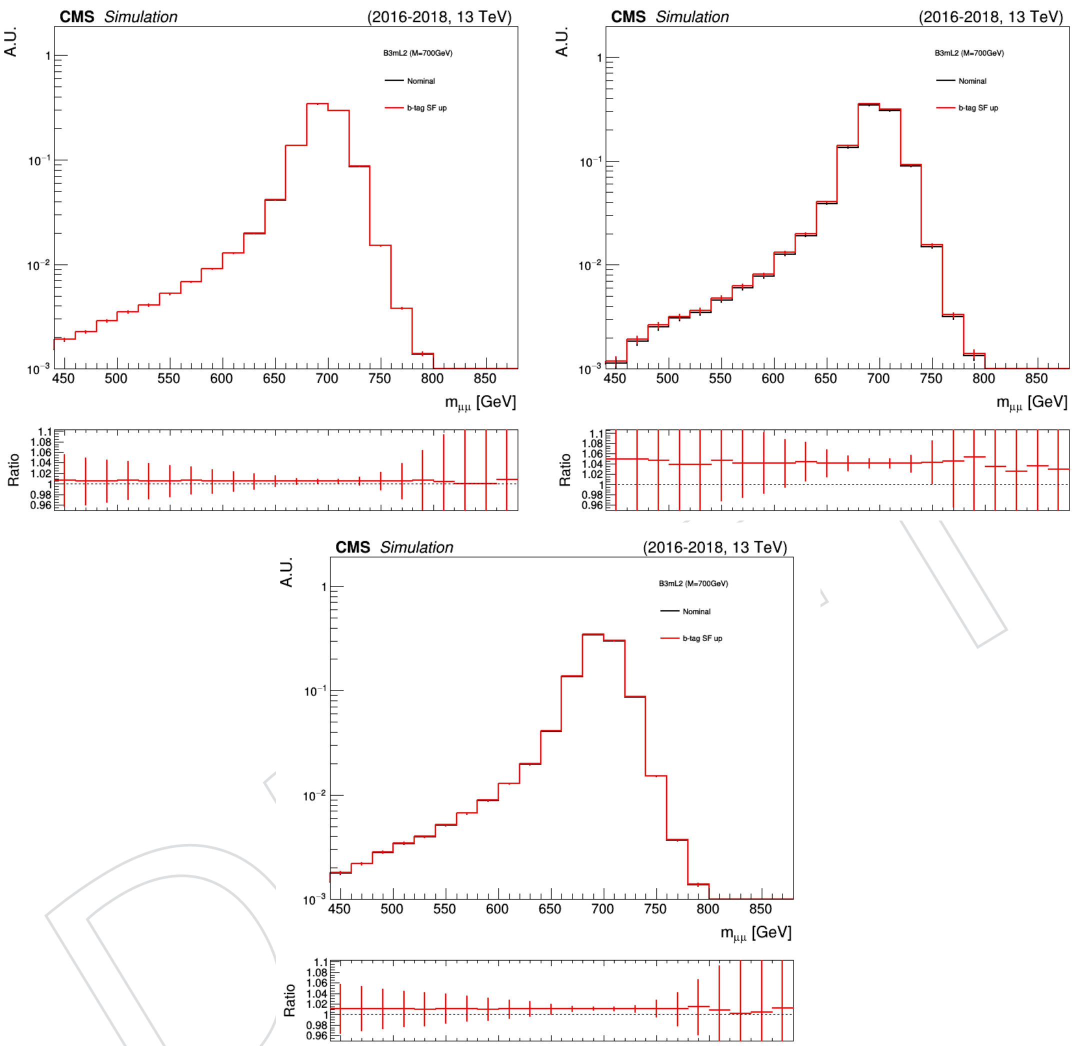


Figure 98: The effect of the propagatation of the upward uncertainty in the b-tagging data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 700\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the upward uncertainty in the b-tagging data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

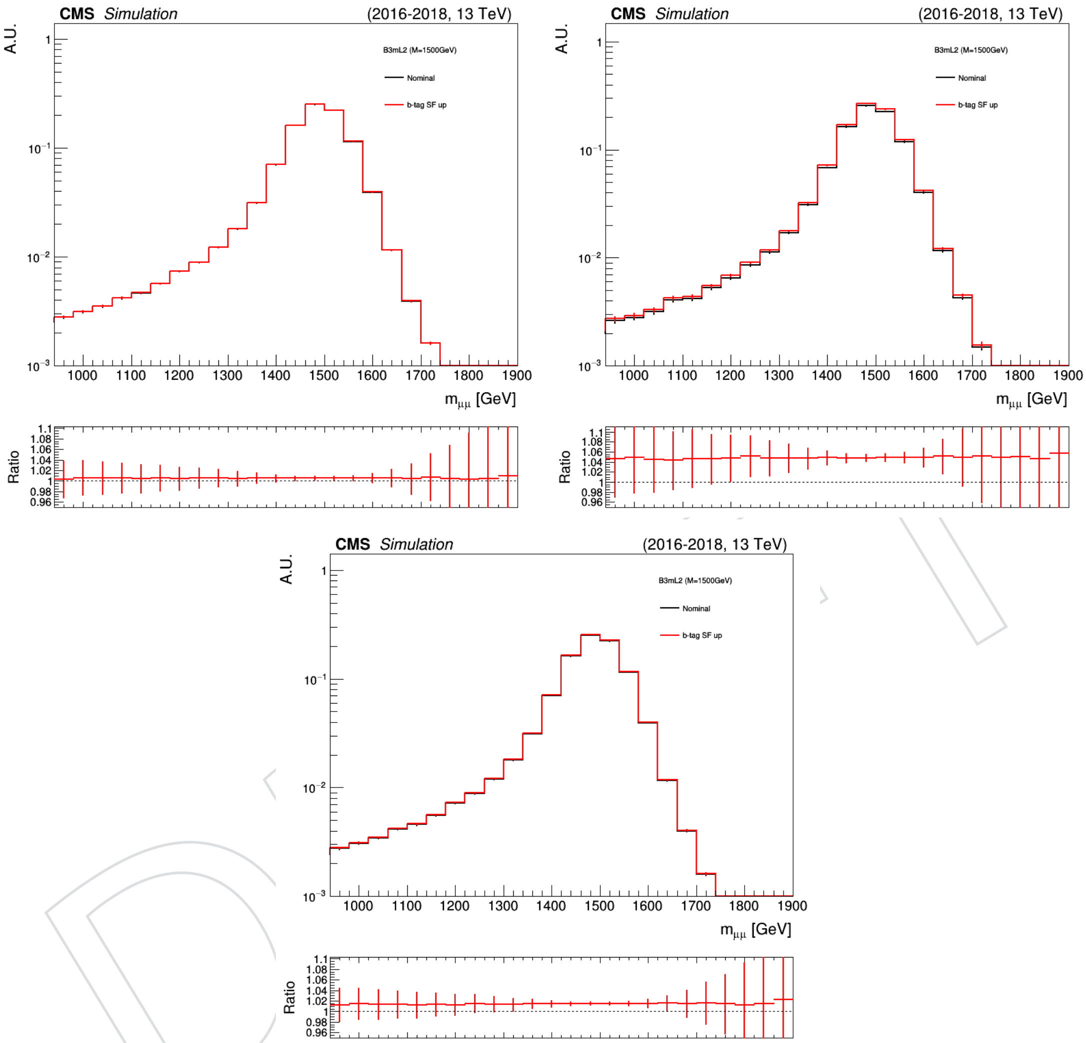


Figure 99: The effect of the propagatation of the upward uncertainty in the b-tagging data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 1500\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the upward uncertainty in the b-tagging data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

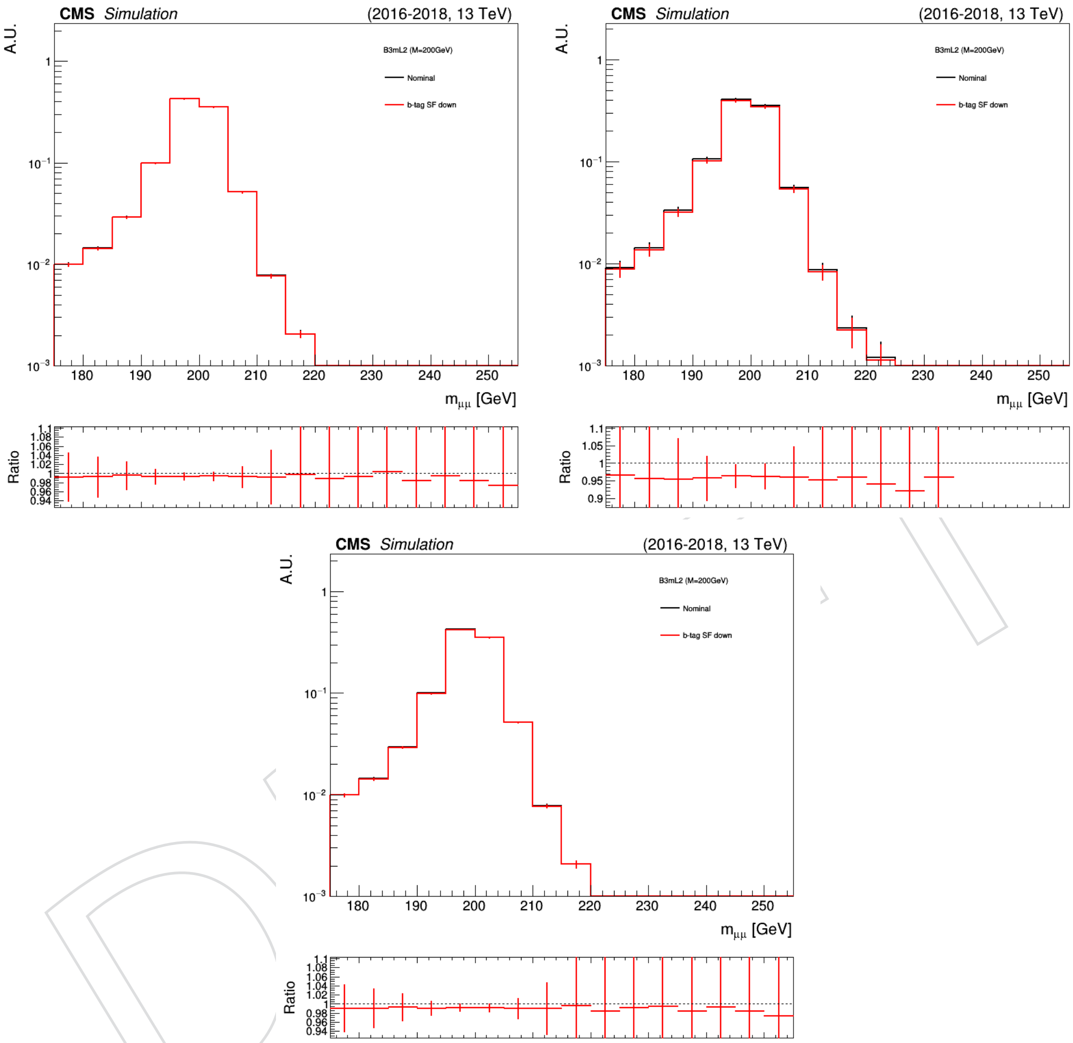


Figure 100: The effect of the propagatation of the downward uncertainty in the b-tagging data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 200\,\text{GeV}$), after the full event selection, in event categories with $N_b = 1$ (top left), $N_b \geq 2$ (top right), and $N_b \geq 1$ (bottom; equal to the sum of $N_b = 1$ and $N_b \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the downward uncertainty in the b-tagging data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

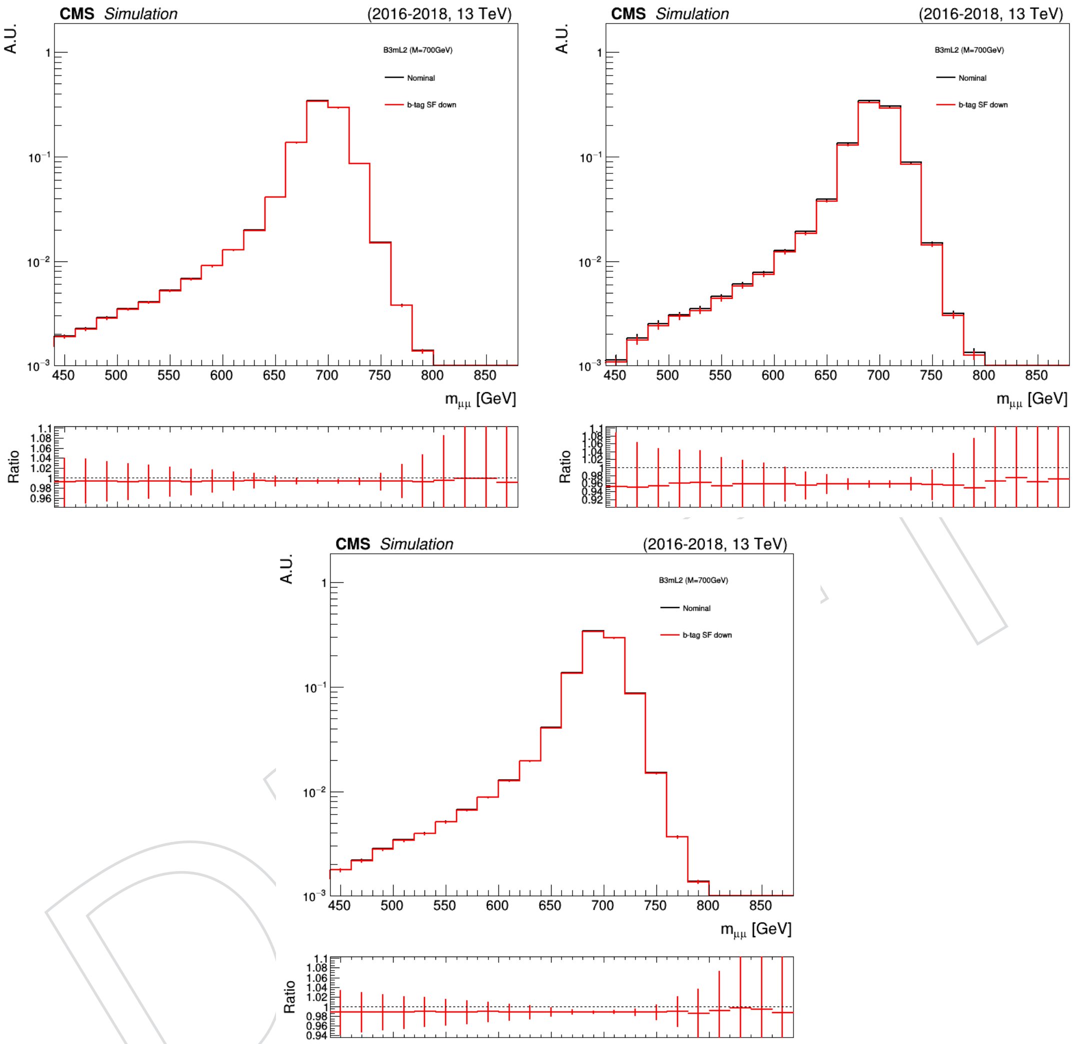


Figure 101: The effect of the propagatation of the downward uncertainty in the b-tagging data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 700\,\text{GeV}$), after the full event selection, in event categories with $N_{\text{b}} = 1$ (top left), $N_{\text{b}} \geq 2$ (top right), and $N_{\text{b}} \geq 1$ (bottom; equal to the sum of $N_{\text{b}} = 1$ and $N_{\text{b}} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the downward uncertainty in the b-tagging data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

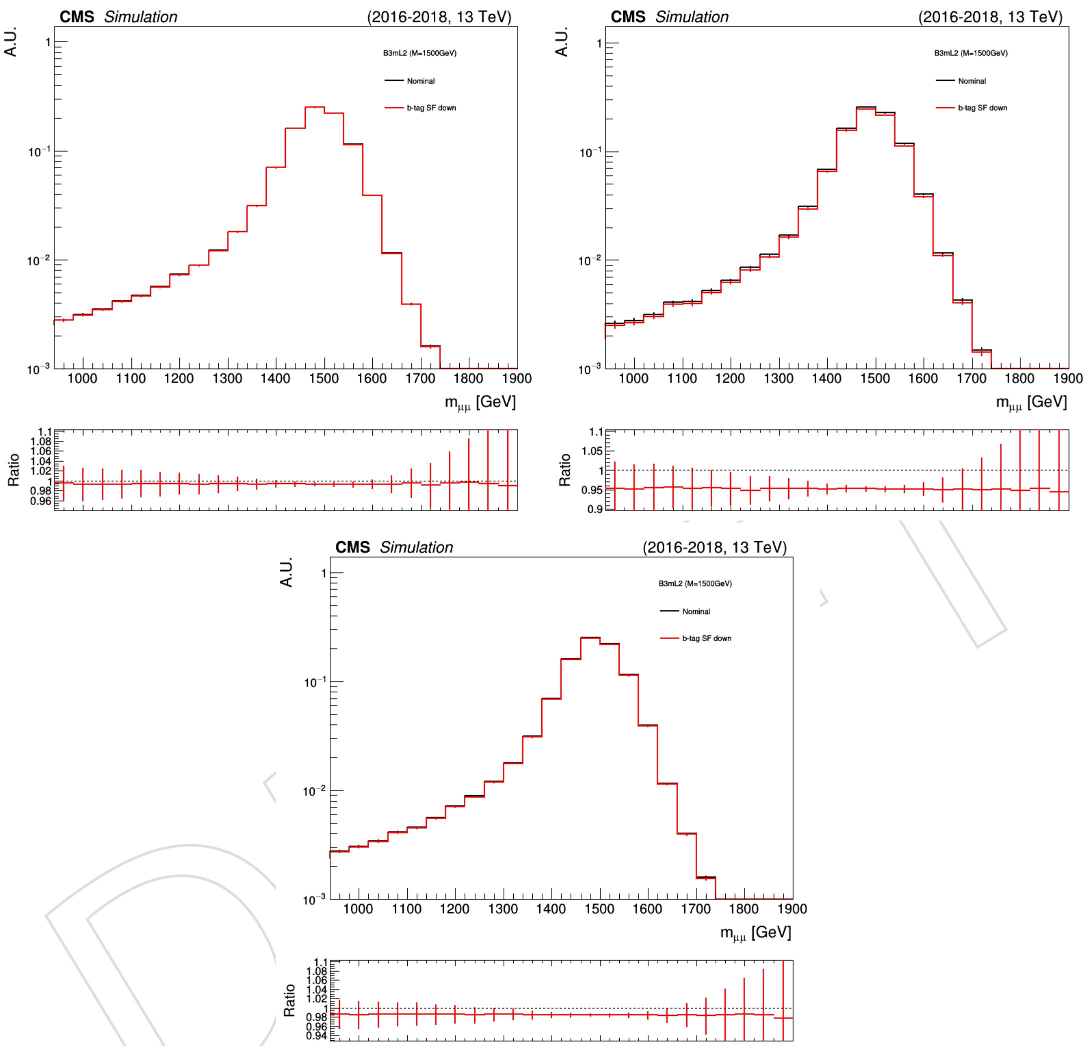


Figure 102: The effect of the propagatation of the downward uncertainty in the b-tagging data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 1500\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the downward uncertainty in the b-tagging data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

E.7 Uncertainty in muon reconstruction, identification and isolation efficiency

E.7.1 Muon reconstruction

Figures 103–105 show the effect of the propagation of the upward uncertainty in the muon reconstruction data/MC scale factors [70–72] for a few representative signal models. Similarly, Figs. 106–108 show the effect of the propagation of the downward uncertainty in the muon reconstruction data/MC scale factors for the same signal models.

The effect in the signal yield is slightly increasing at increasing $m_{Z'}$, independently of N_b, roughly ranging from 1.5% at $m_{Z'} = 200\,\text{GeV}$ to 3% at $m_{Z'} = 2000\,\text{GeV}$. No statistically significant effect is found on the shape of the signal expected dimuon invariant mass distribution. Thus, a dedicated systematic uncertainty in the expected signal yield is assessed.

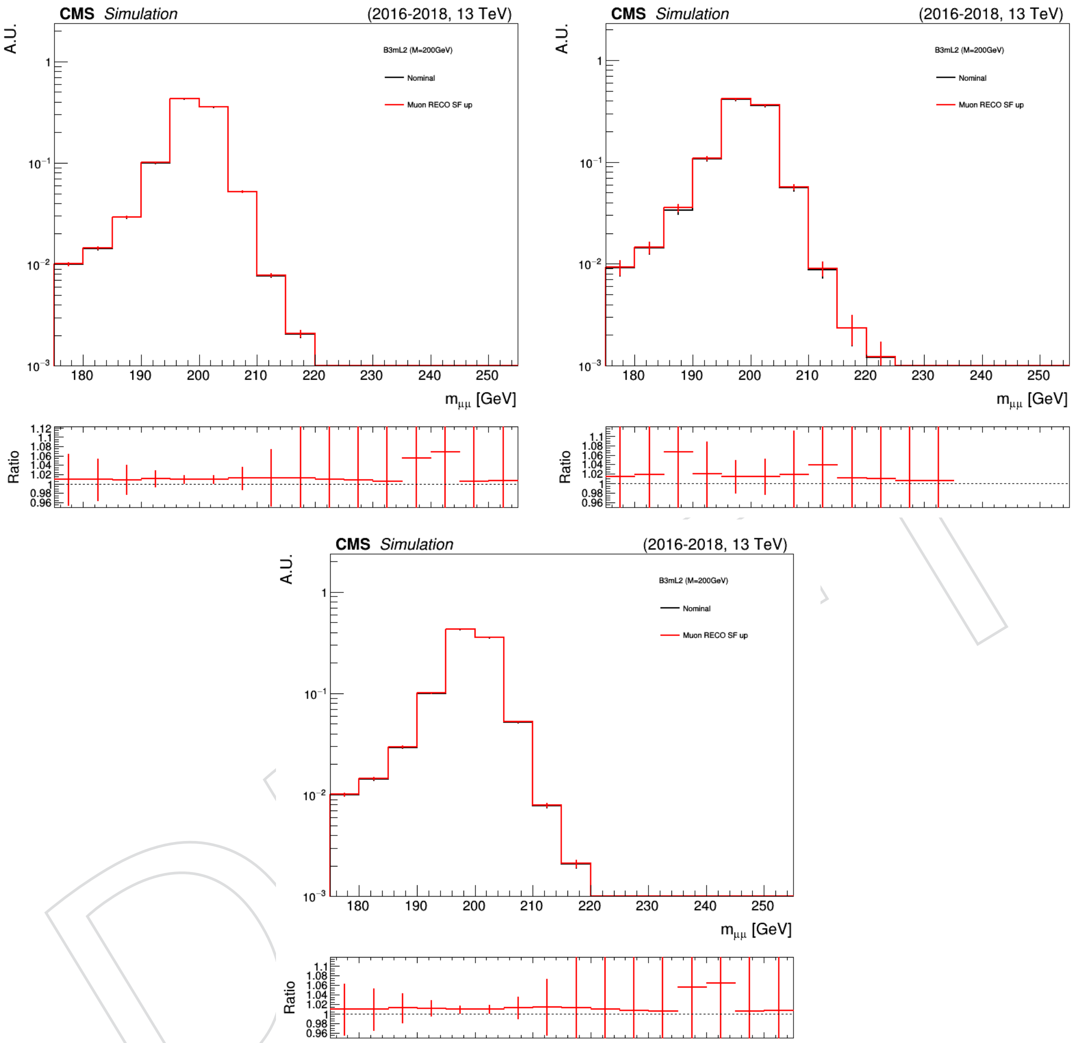


Figure 103: The effect of the propagatation of the upward uncertainty in the muon reconstruction data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 200\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the upward uncertainty in the muon reconstruction data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

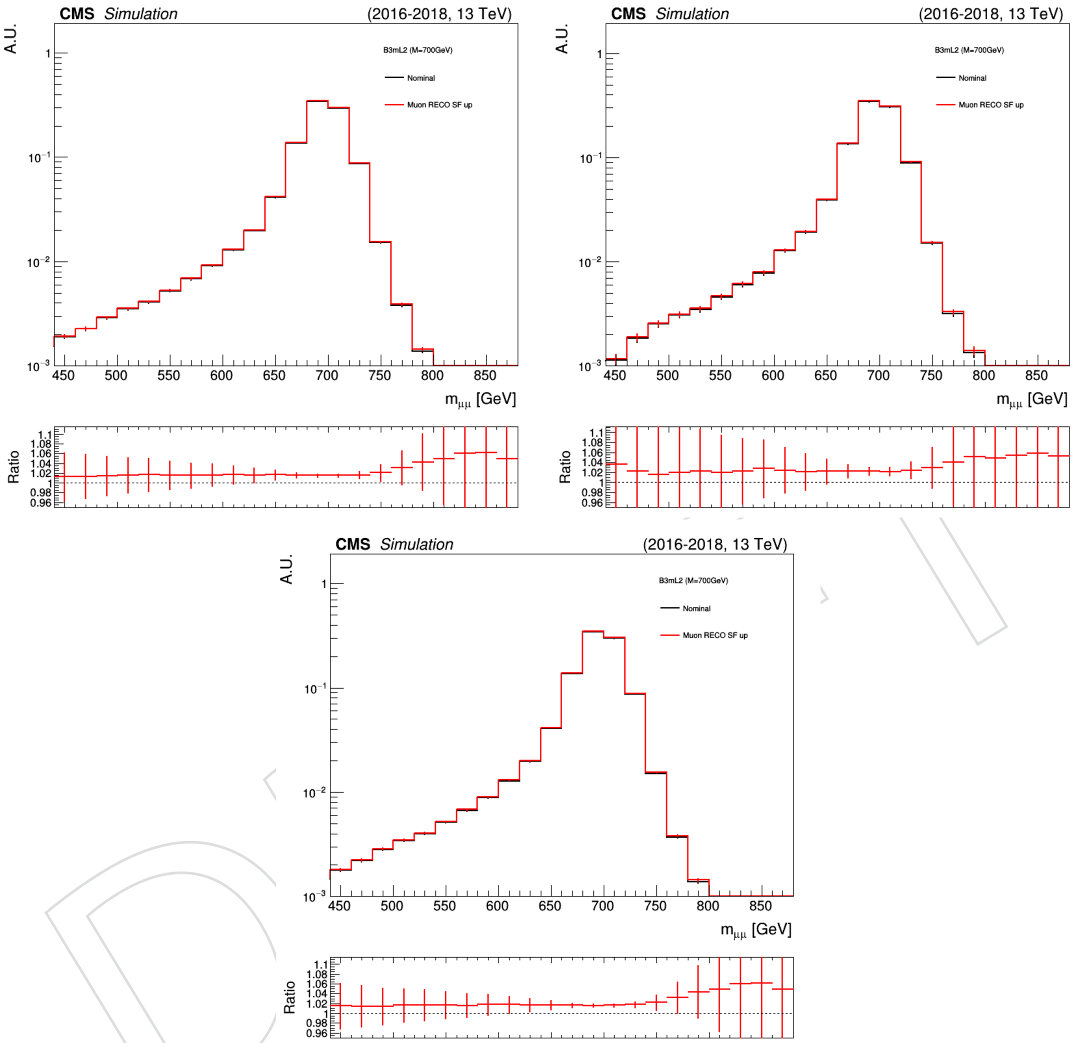


Figure 104: The effect of the propagatation of the upward uncertainty in the muon reconstruction data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 700\,\text{GeV}$), after the full event selection, in event categories with $N_b = 1$ (top left), $N_b \geq 2$ (top right), and $N_b \geq 1$ (bottom; equal to the sum of $N_b = 1$ and $N_b \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the upward uncertainty in the muon reconstruction data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

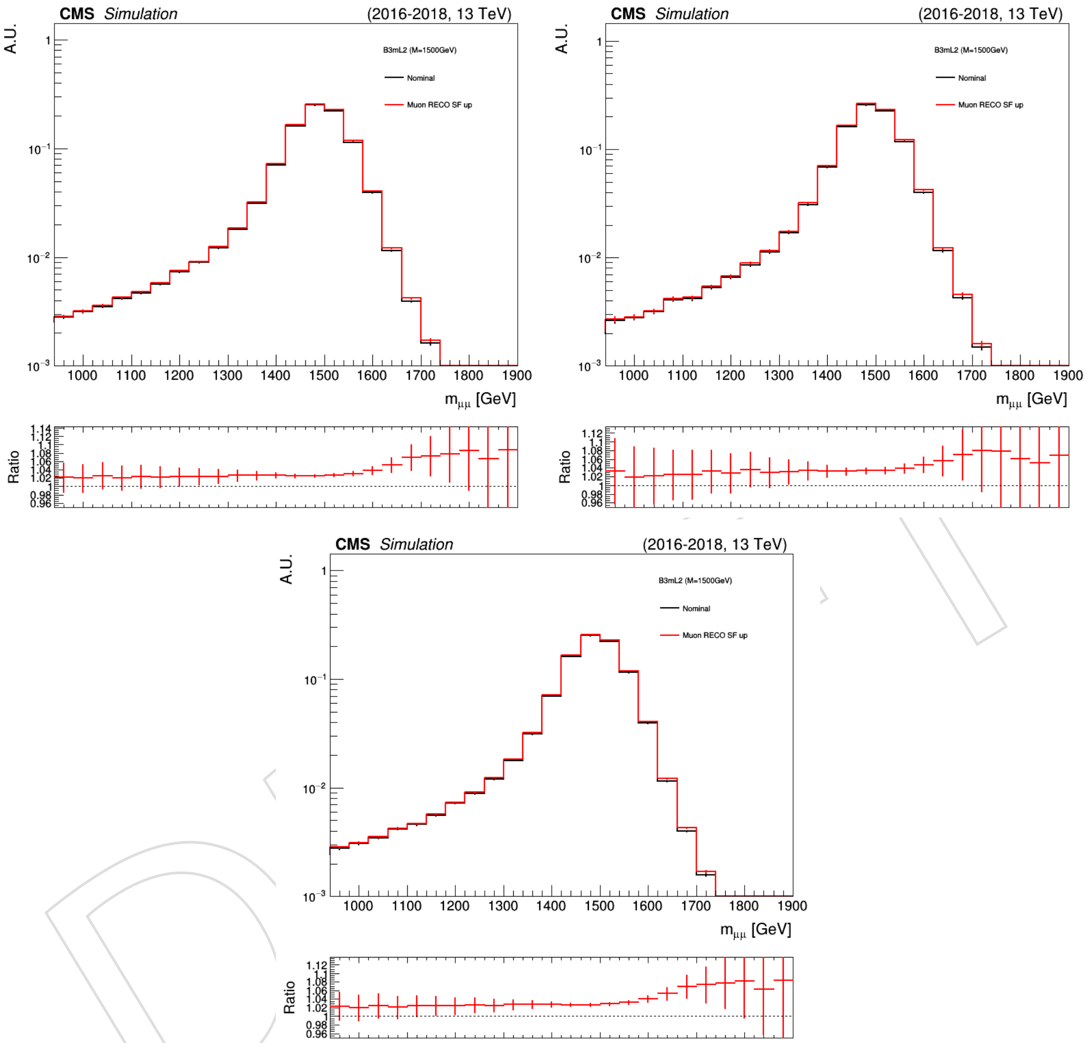


Figure 105: The effect of the propagatation of the upward uncertainty in the muon reconstruction data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 1500\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the upward uncertainty in the muon reconstruction data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

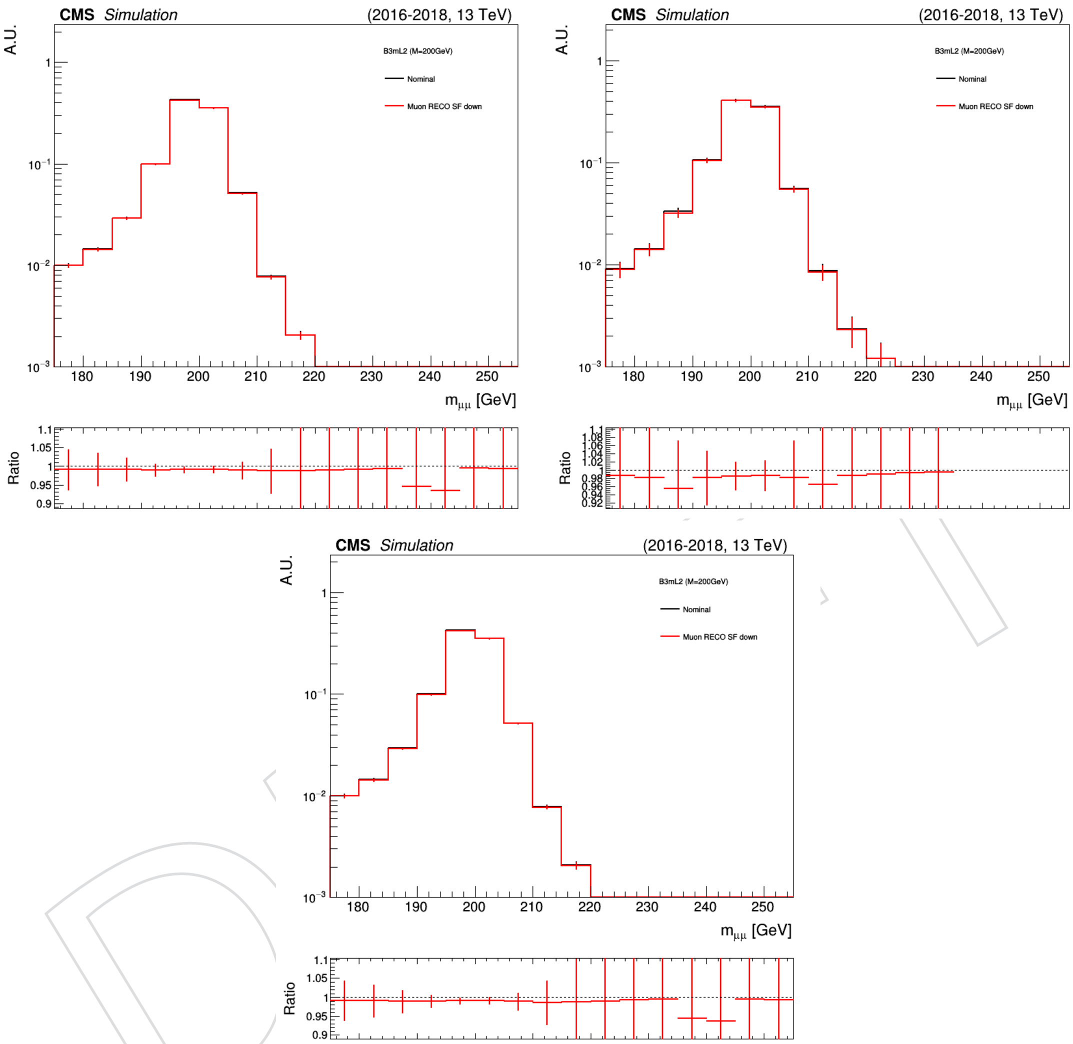


Figure 106: The effect of the propagatation of the downward uncertainty in the muon reconstruction data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 200\,\mathrm{GeV}$), after the full event selection, in event categories with $N_\mathrm{b} = 1$ (top left), $N_\mathrm{b} \geq 2$ (top right), and $N_\mathrm{b} \geq 1$ (bottom; equal to the sum of $N_\mathrm{b} = 1$ and $N_\mathrm{b} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the downward uncertainty in the muon reconstruction data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

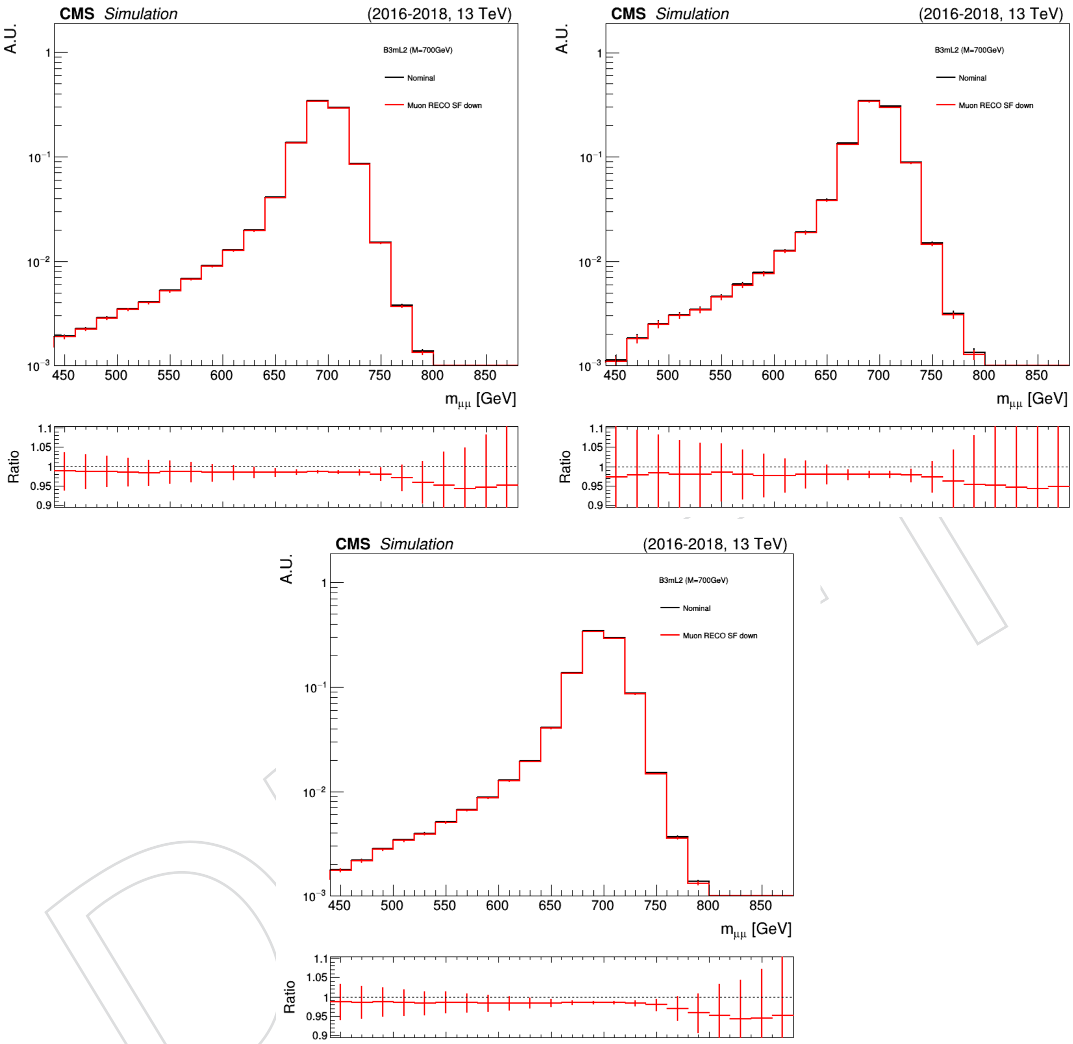


Figure 107: The effect of the propagatation of the downward uncertainty in the muon reconstruction data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 700\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the downward uncertainty in the muon reconstruction data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

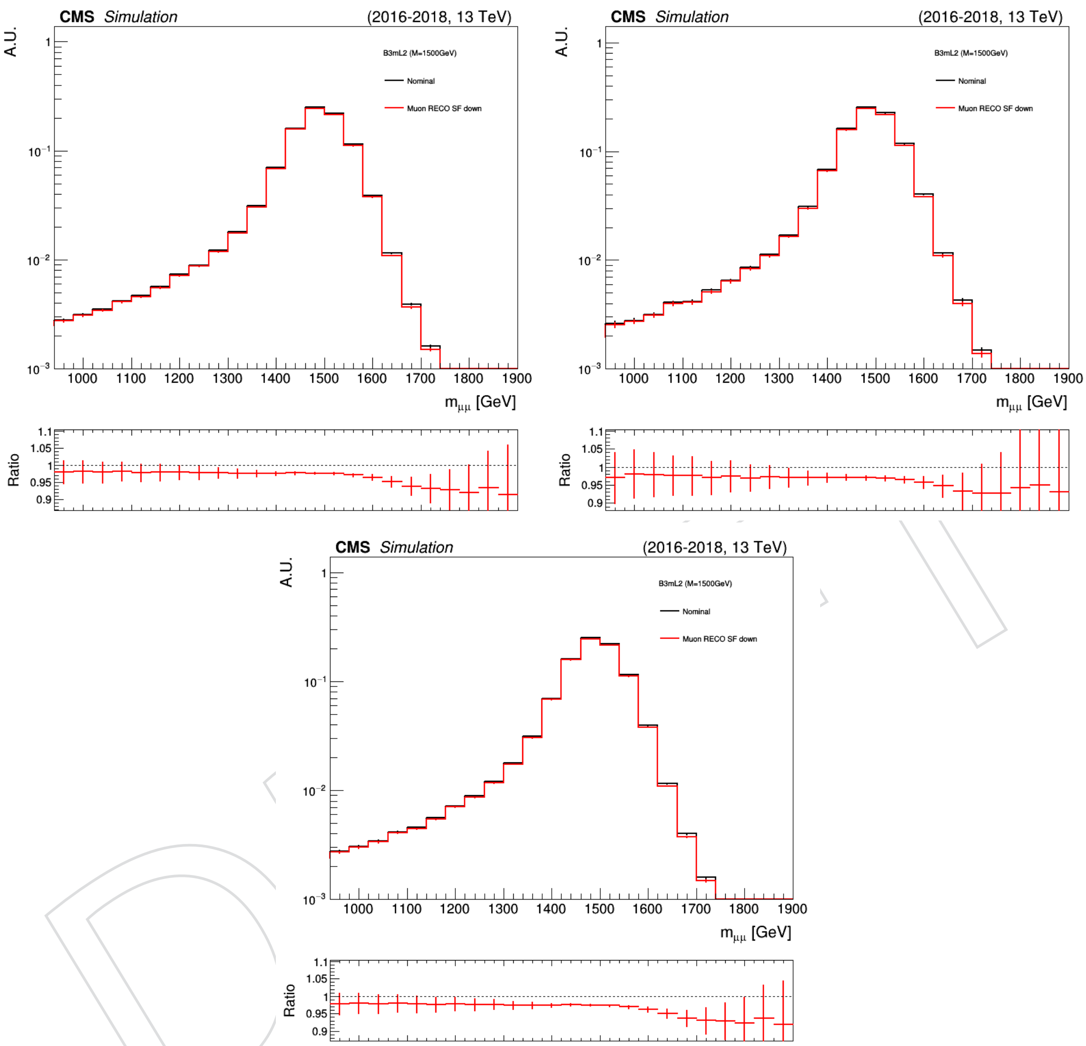


Figure 108: The effect of the propagatation of the downward uncertainty in the muon reconstruction data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 1500\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the downward uncertainty in the muon reconstruction data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

E.7.2 Muon identification

Figures 109–111 show the effect of the propagation of the upward uncertainty in the muon identification data/MC scale factors [70–72] for a few representative signal models. Similarly, Figs. 112–114 show the effect of the propagation of the downward uncertainty in the muon identification data/MC scale factors for the same signal models.

The effect in the signal yield is negligible compared to the statistical uncertainties, and subdominat with respect to other systematic uncertainties. No statistically significant effect is found on the shape of the signal expected dimuon invariant mass distribution. Thus, no dedicated systematic uncertainty arising from the uncertainty in the muon isolation data/MC scale factors is assessed.

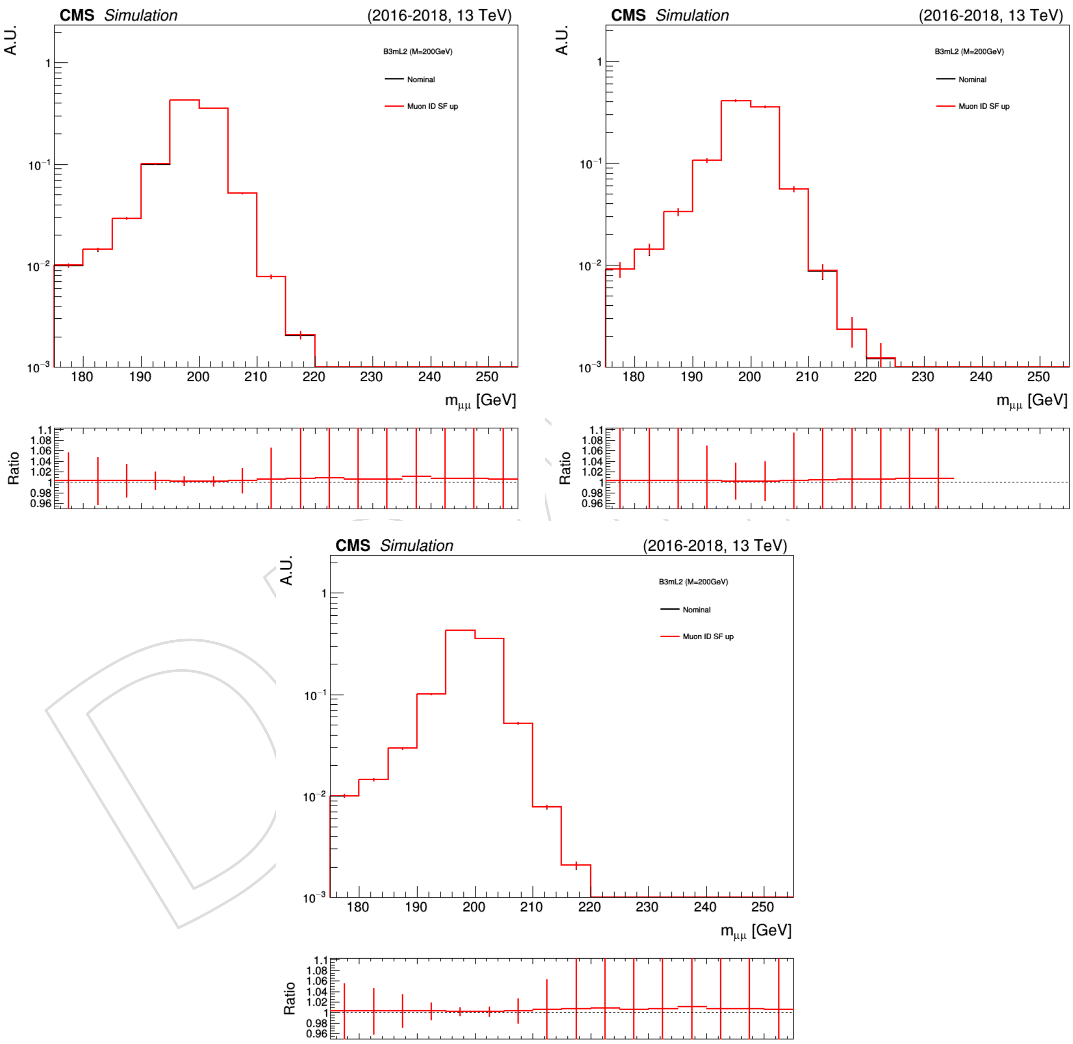


Figure 109: The effect of the propagatation of the upward uncertainty in the muon identification data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 200\,\text{GeV}$), after the full event selection, in event categories with $N_b = 1$ (top left), $N_b \geq 2$ (top right), and $N_b \geq 1$ (bottom; equal to the sum of $N_b = 1$ and $N_b \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the upward uncertainty in the muon identification data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

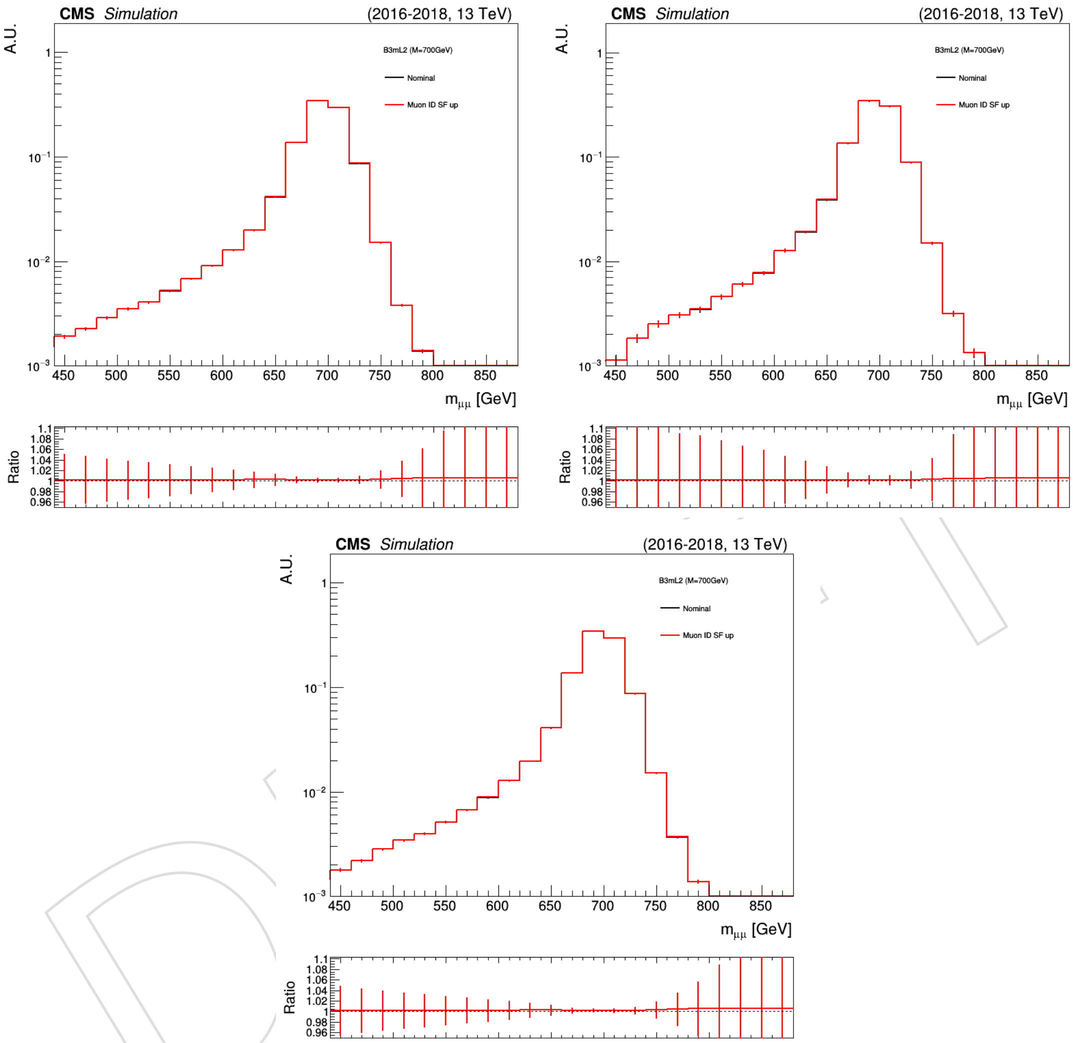


Figure 110: The effect of the propagatation of the upward uncertainty in the muon identification data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 700\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the upward uncertainty in the muon identification data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

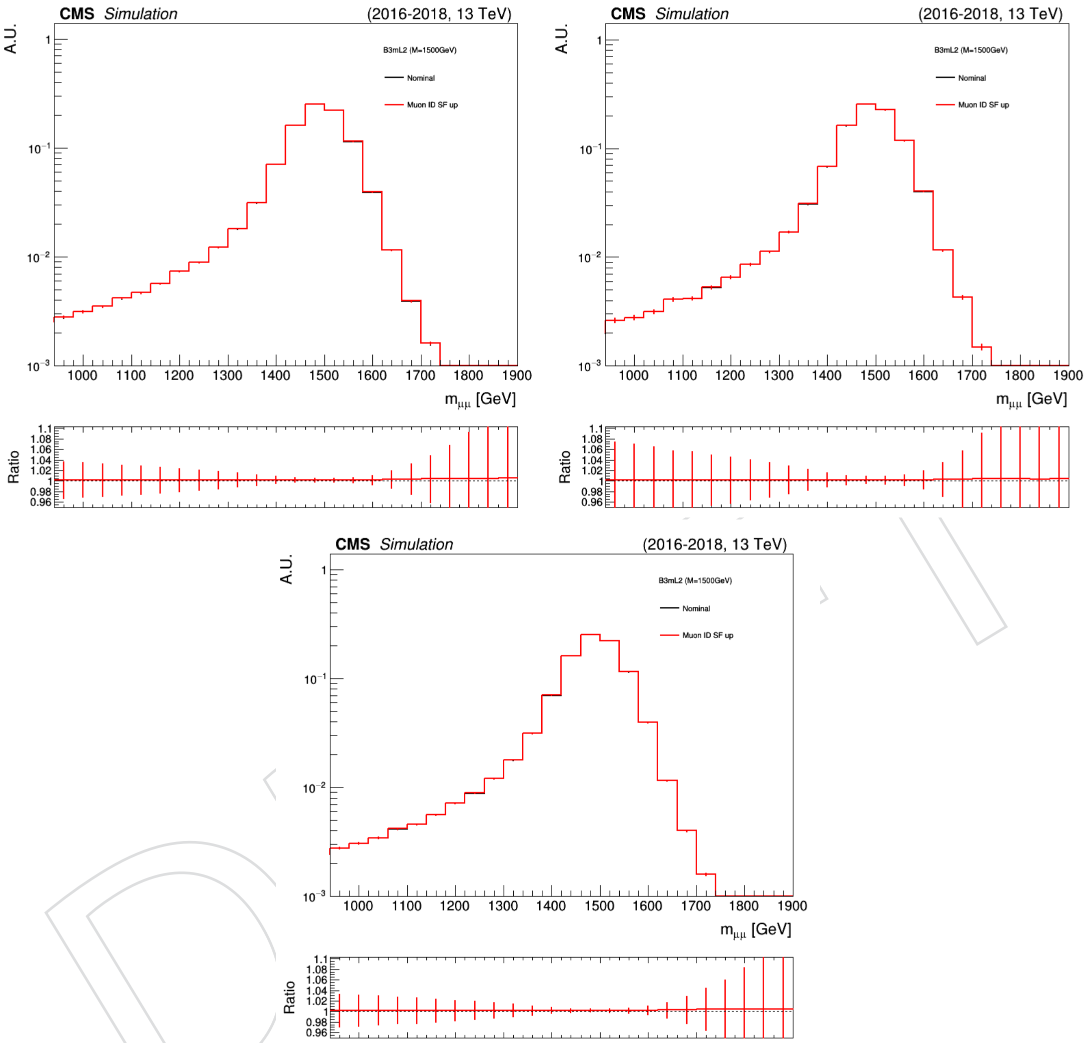


Figure 111: The effect of the propagatation of the upward uncertainty in the muon identification data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 1500\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the upward uncertainty in the muon identification data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

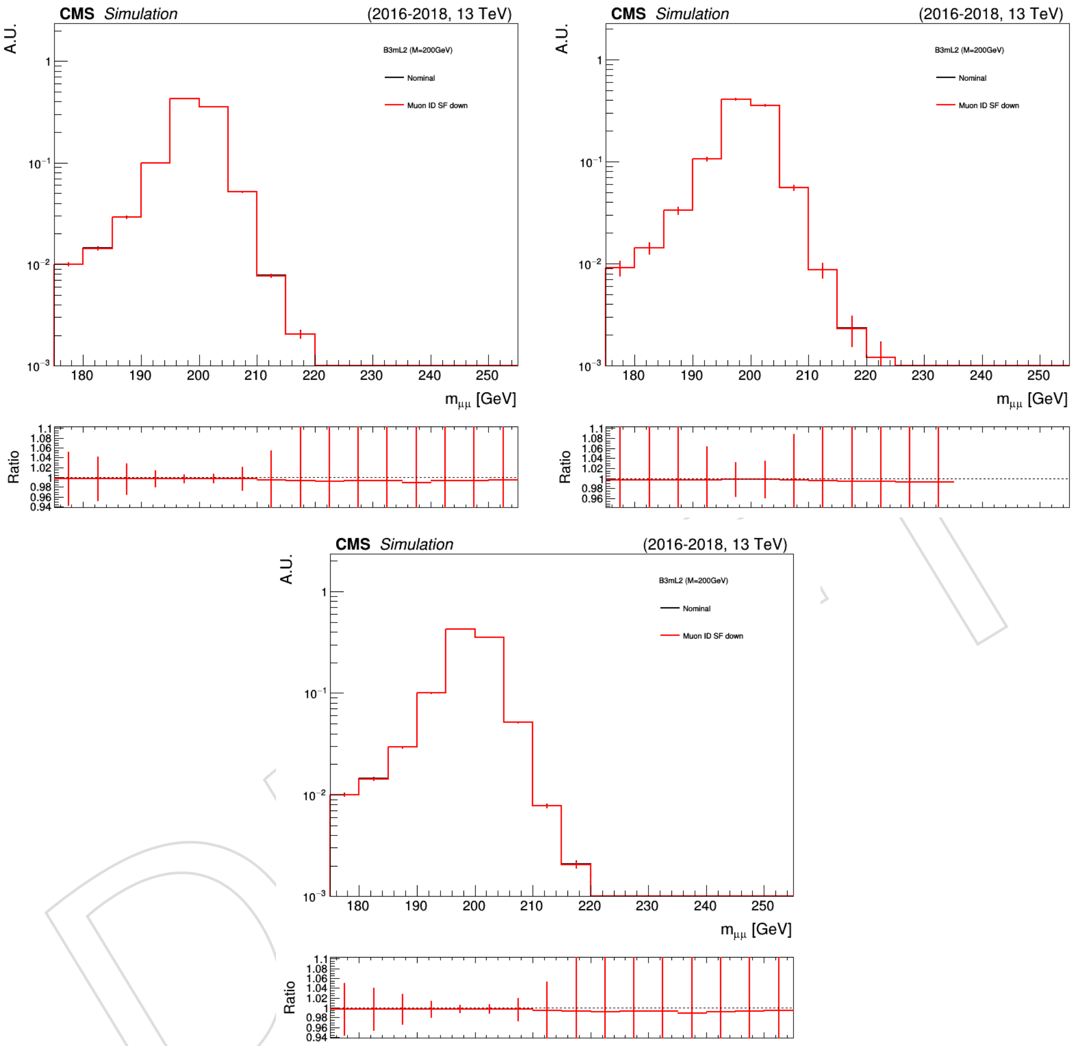


Figure 112: The effect of the propagatation of the downward uncertainty in the muon identification data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 200\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the downward uncertainty in the muon identification data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

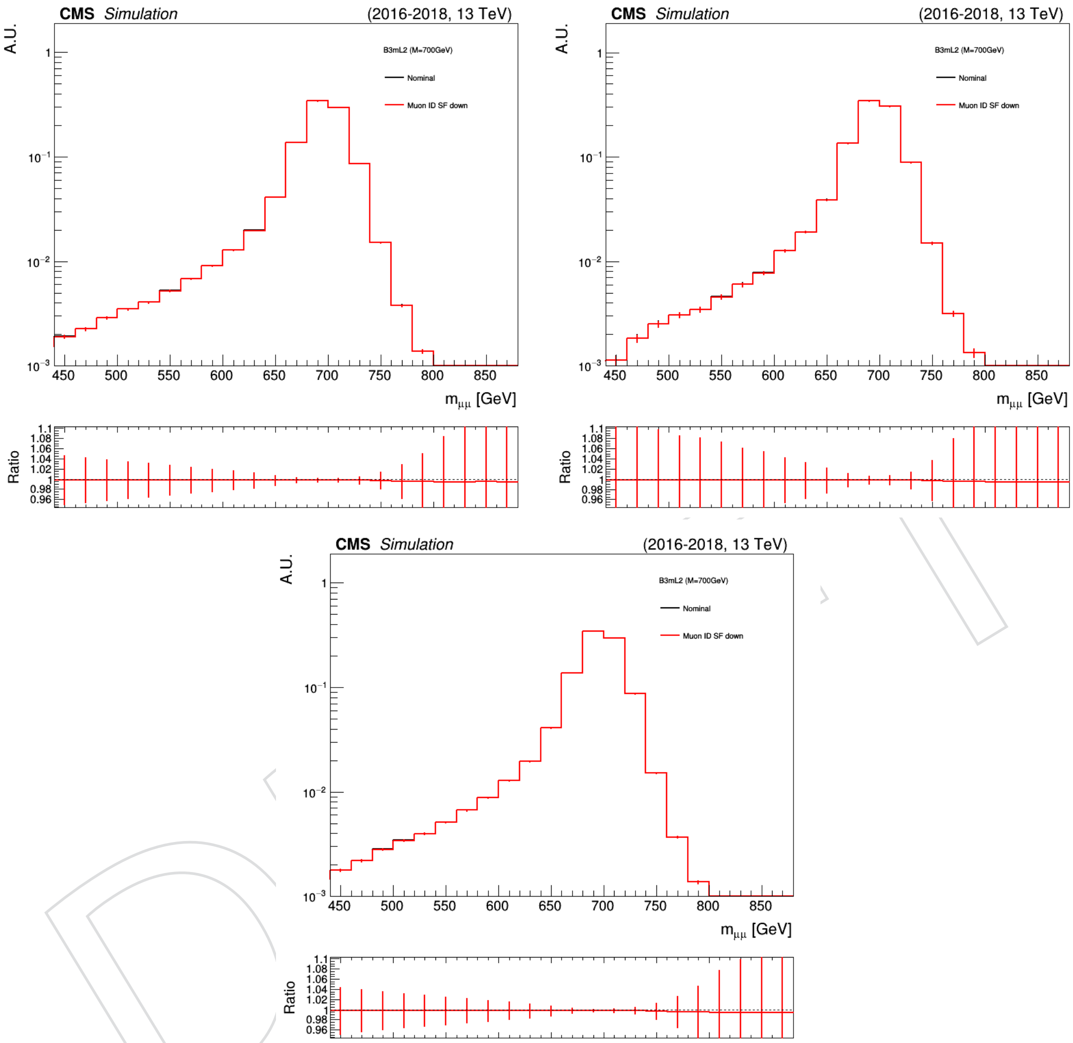


Figure 113: The effect of the propagatation of the downward uncertainty in the muon identification data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 700\,\text{GeV}$), after the full event selection, in event categories with $N_b = 1$ (top left), $N_b \geq 2$ (top right), and $N_b \geq 1$ (bottom; equal to the sum of $N_b = 1$ and $N_b \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the downward uncertainty in the muon identification data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

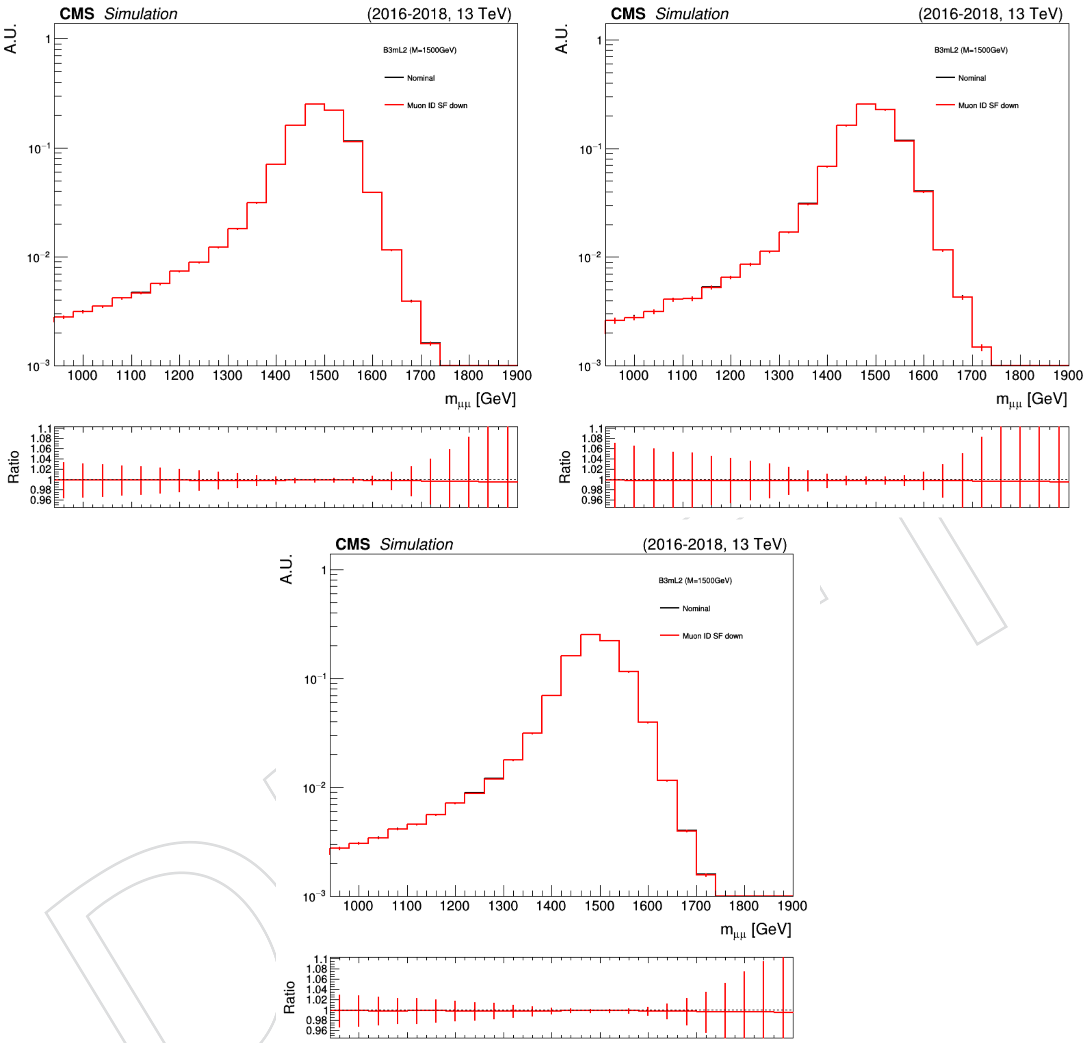


Figure 114: The effect of the propagatation of the downward uncertainty in the muon identification data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 1500\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the downward uncertainty in the muon identification data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

E.7.3 Muon isolation

Figures 115–117 show the effect of the propagation of the upward uncertainty in the muon isolation data/MC scale factors [70–72] for a few representative signal models. Similarly, Figs. 118–120 show the effect of the propagation of the downward uncertainty in the muon isolation data/MC scale factors for the same signal models.

The effect in the signal yield is negligible compared to the statistical uncertainties, and subdominat with respect to other systematic uncertainties. No statistically significant effect is found on the shape of the signal expected dimuon invariant mass distribution. Thus, no dedicated systematic uncertainty arising from the uncertainty in the muon isolation data/MC scale factors is assessed.

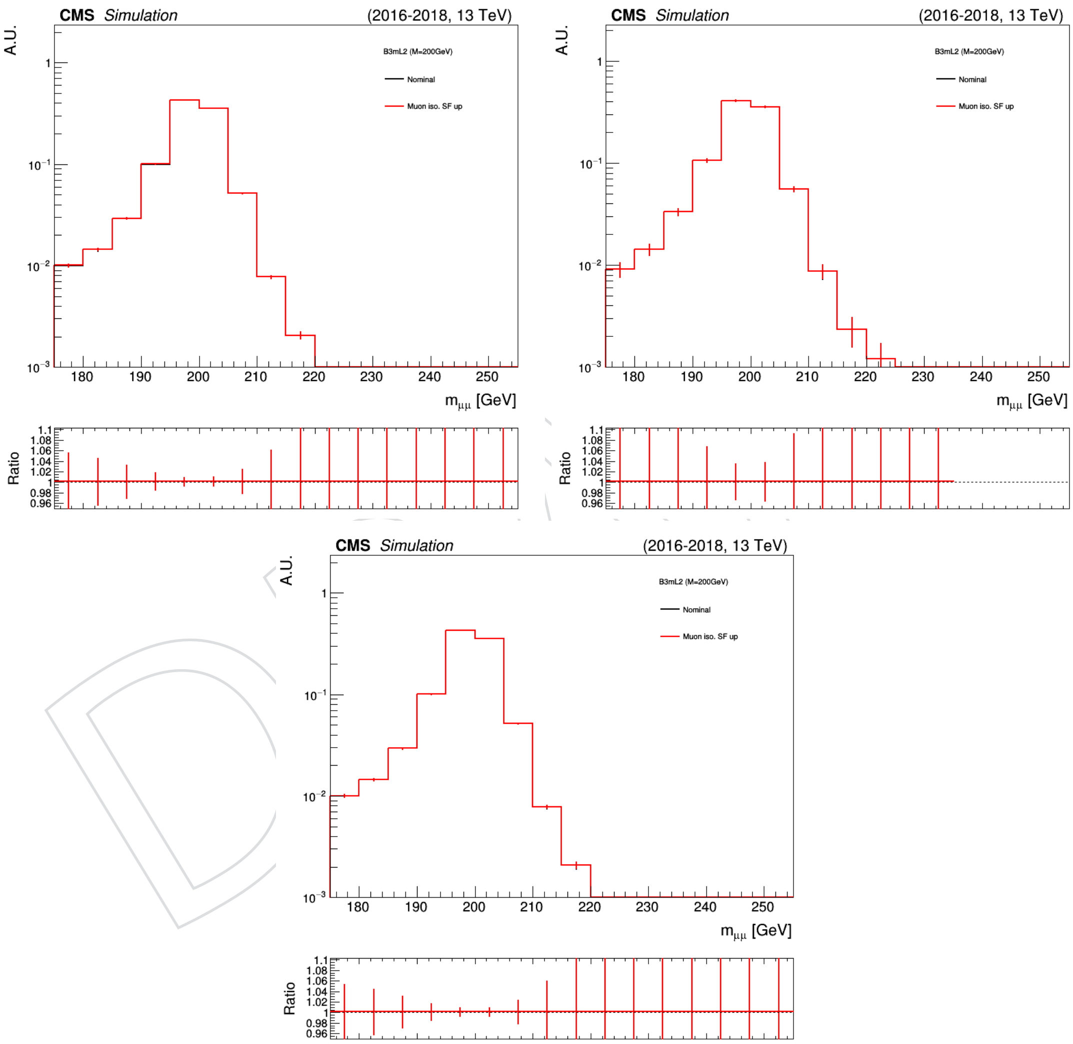


Figure 115: The effect of the propagatation of the upward uncertainty in the muon isolation data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 200\,\text{GeV}$), after the full event selection, in event categories with $N_b = 1$ (top left), $N_b \geq 2$ (top right), and $N_b \geq 1$ (bottom; equal to the sum of $N_b = 1$ and $N_b \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the upward uncertainty in the muon isolation data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

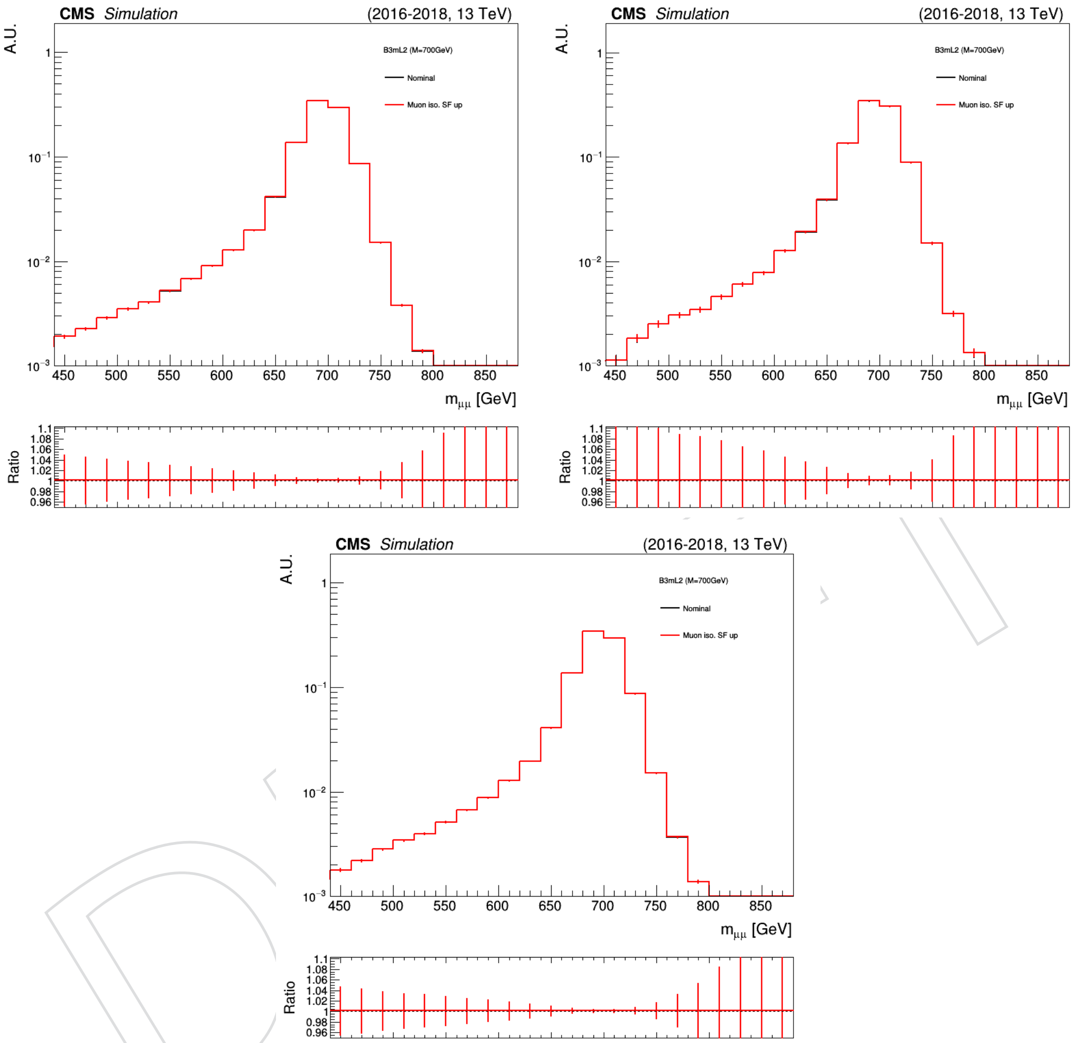


Figure 116: The effect of the propagatation of the upward uncertainty in the muon isolation data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 700\,\text{GeV}$), after the full event selection, in event categories with $N_b = 1$ (top left), $N_b \geq 2$ (top right), and $N_b \geq 1$ (bottom; equal to the sum of $N_b = 1$ and $N_b \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the upward uncertainty in the muon isolation data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

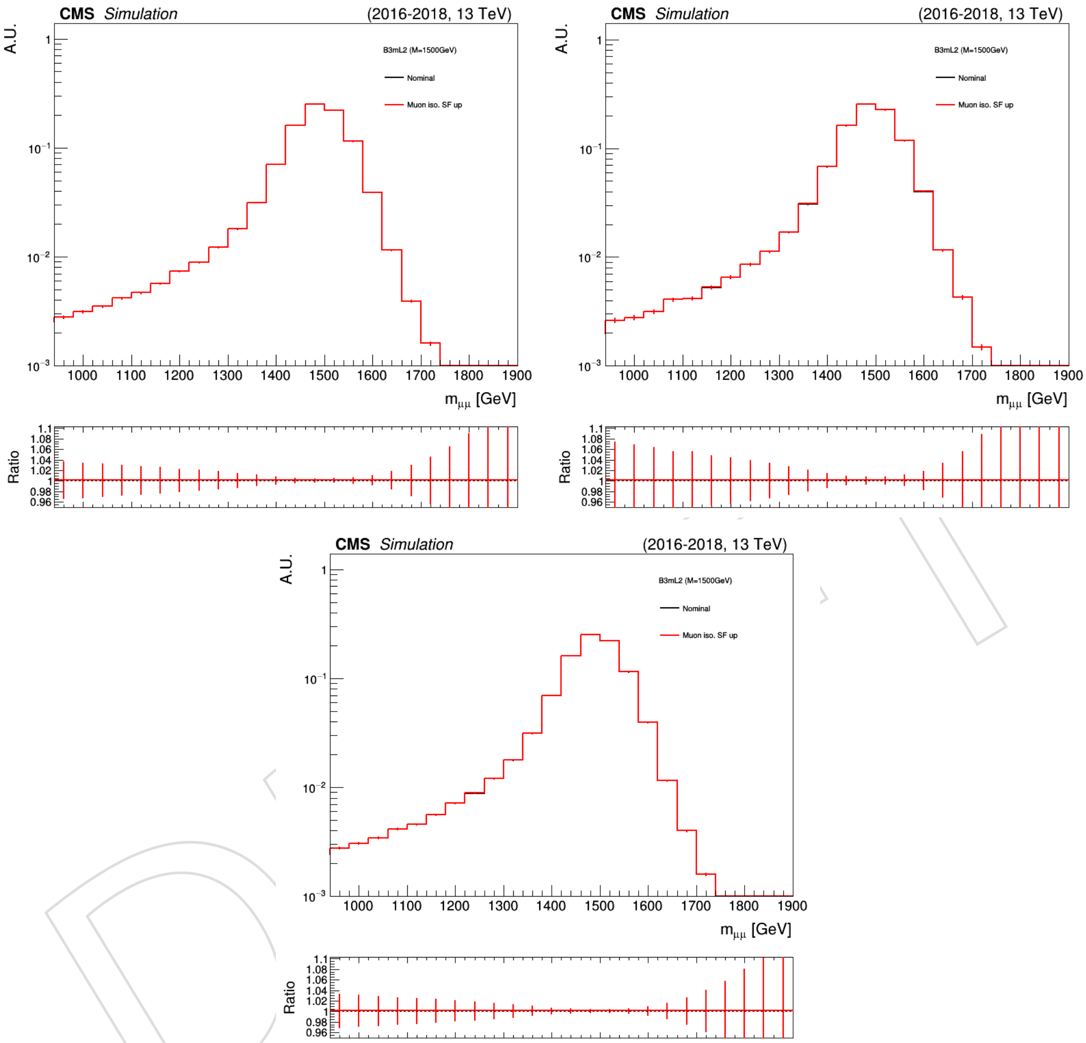


Figure 117: The effect of the propagatation of the upward uncertainty in the muon isolation data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 1500\,\text{GeV}$), after the full event selection, in event categories with $N_b = 1$ (top left), $N_b \geq 2$ (top right), and $N_b \geq 1$ (bottom; equal to the sum of $N_b = 1$ and $N_b \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the upward uncertainty in the muon isolation data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

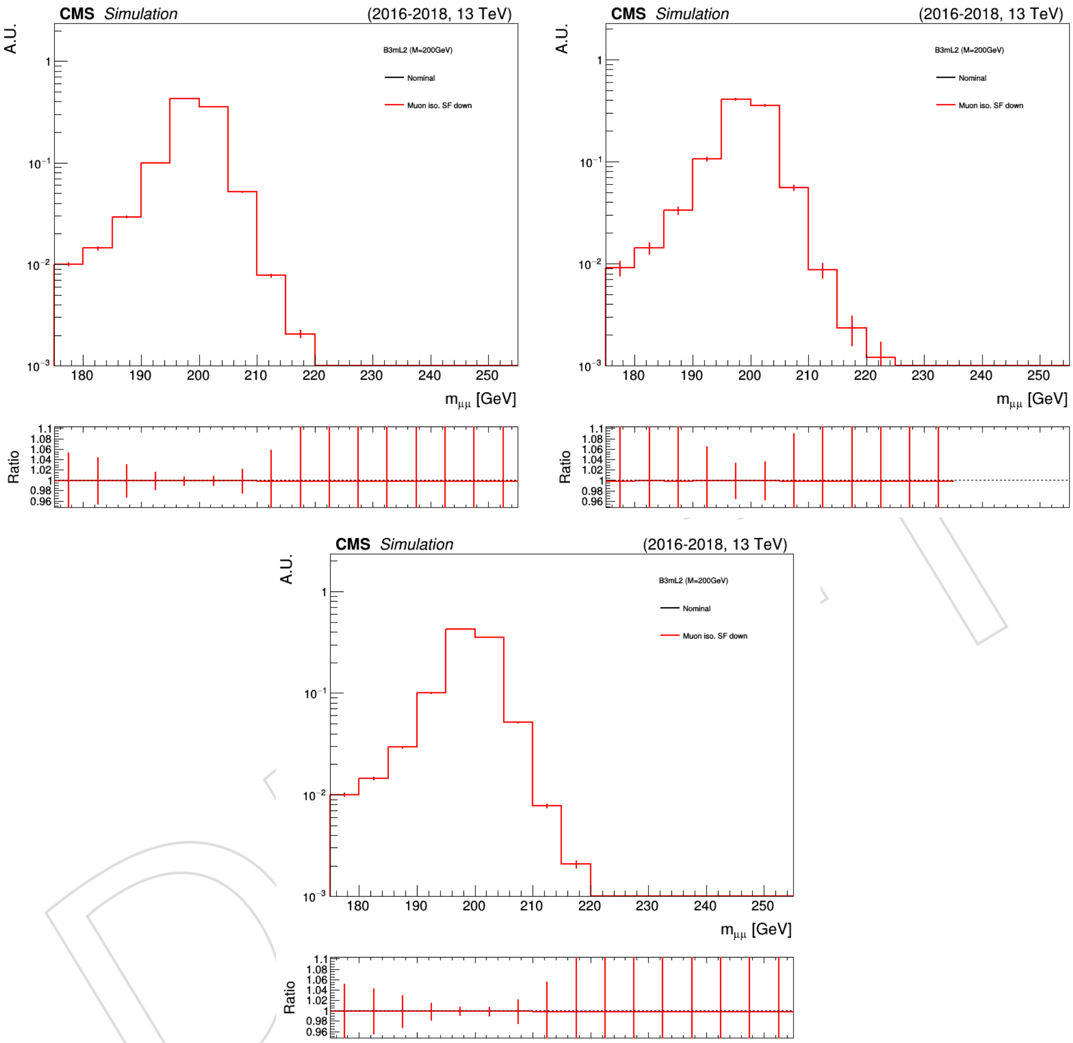


Figure 118: The effect of the propagatation of the downward uncertainty in the muon isolation data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 200\,\text{GeV}$), after the full event selection, in event categories with $N_b = 1$ (top left), $N_b \geq 2$ (top right), and $N_b \geq 1$ (bottom; equal to the sum of $N_b = 1$ and $N_b \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the downward uncertainty in the muon isolation data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

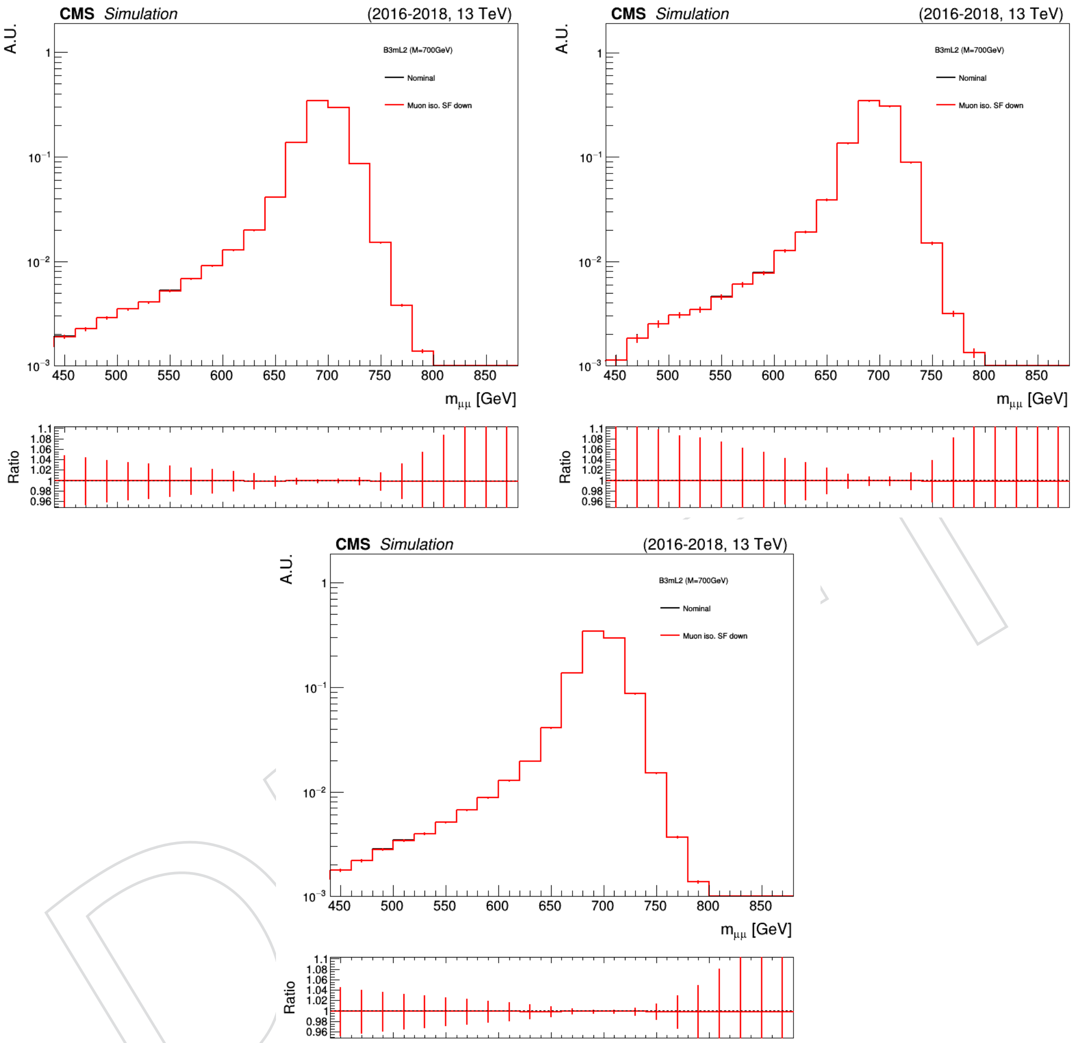


Figure 119: The effect of the propagatation of the downward uncertainty in the muon isolation data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 700\,\mathrm{GeV}$), after the full event selection, in event categories with $N_\mathrm{b} = 1$ (top left), $N_\mathrm{b} \geq 2$ (top right), and $N_\mathrm{b} \geq 1$ (bottom; equal to the sum of $N_\mathrm{b} = 1$ and $N_\mathrm{b} \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the downward uncertainty in the muon isolation data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

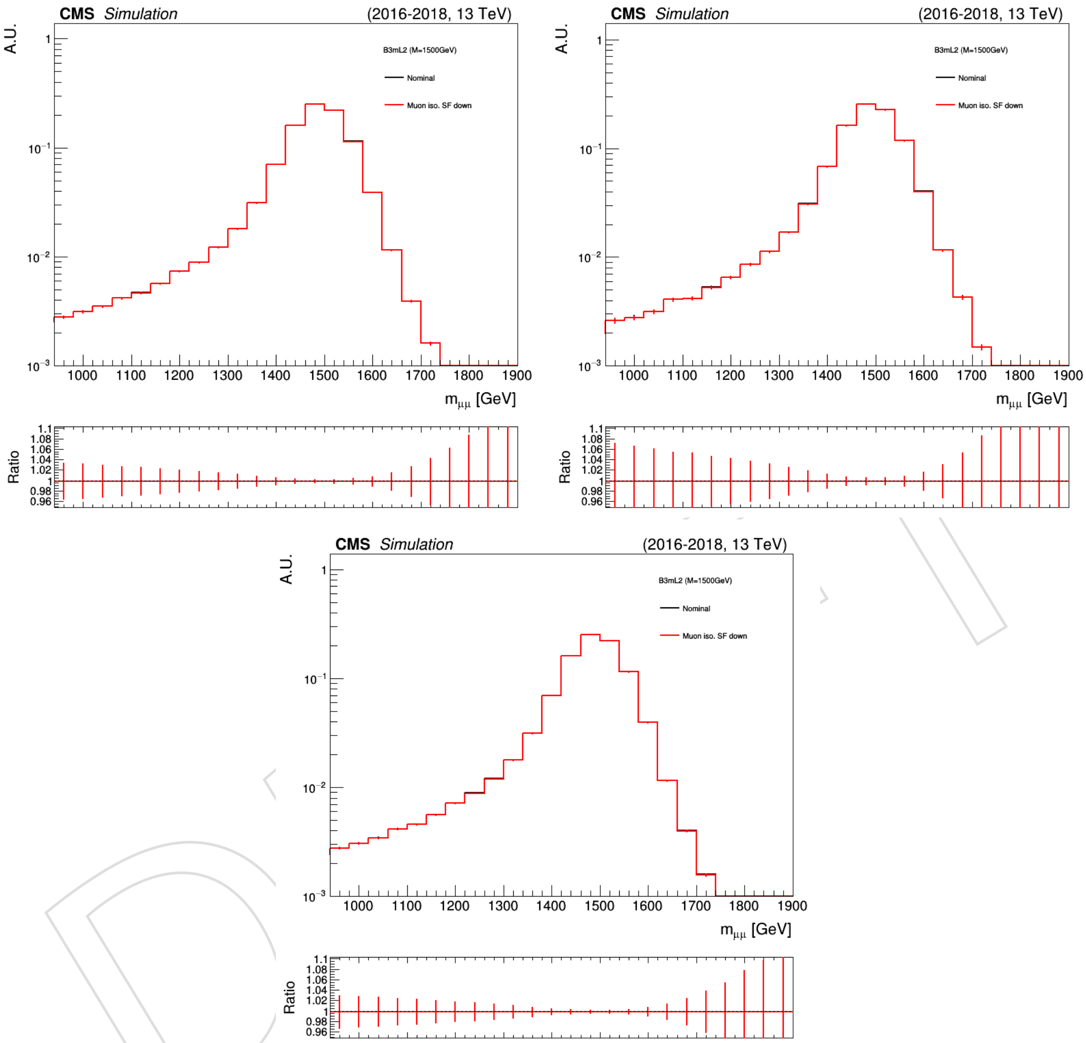


Figure 120: The effect of the propagatation of the downward uncertainty in the muon isolation data/MC scale factors on the expected signal yield and signal $m_{\mu\mu}$ shape is shown, as measured in a representative signal MC sample (B3-L2, with $m_{Z'} = 1500\,\text{GeV}$), after the full event selection, in event categories with $N_b = 1$ (top left), $N_b \geq 2$ (top right), and $N_b \geq 1$ (bottom; equal to the sum of $N_b = 1$ and $N_b \geq 2$). The black histogram represents the nominal signal dimuon invariant mass distribution. The red histogram represents the same distribution once the downward uncertainty in the muon isolation data/MC scale factors is propagated. Both distributions are normalized to the area of the nominal histogram.

F On the effect of tighter muon identification and isolation criteria

As described in Section 5.2, a subset of the muon selection requirements used in this analysis are tighter than those recommended by the Muon POG. We require

- the absolute muon tracker-only isolation to satisfy $\mathcal{I}^{\mathrm{trk}} < 5\,\mathrm{GeV}$,
- the muon transverse impact parameter to satisfy $|d_{\mathrm{xy}}| < 0.02\,\mathrm{cm}$ (instead of $0.2\,\mathrm{cm}$), and
- the muon longitudinal impact parameter to satisfy $|d_{\mathrm{z}}| < 0.1\,\mathrm{cm}$ (instead of $0.5\,\mathrm{cm}$).

Such tighter requirements result in a reduction of the signal acceptance by less than 5%, independent of the signal model or mass.

Figures 121–123 show the effect of the tighter muon selection requirements for a few representative signal models. The reduction in signal acceptance by $\lesssim 5\%$ is independent of the signal mass hypothesis and of the N_{b} event category, and includes the effect of a tighter requirement on the muon relative tracker-only isolation with respect to other similar searches (i.e., $< 5\%$ of the muon p_{T} instead of 10%) [2]. No statistically significant effect is found on the shape of the signal expected dimuon invariant mass distribution.

As an additional check, we validate the agreement between data and simulation for our choice of tighter impact parameter and isolation requirements. Figures 124–127 show the distributions of the relevant variables in the DY-enriched region, split in two bins: one for the range of values we select and one for the range of values we reject after tightening the selection. As shown in the figures, there is excellent agreement, down to the percent level, for the range of values that we accept. Some disagreement is observed for the range of values that we reject with our tighter selection. This makes sense, since the tightening of these requirements is applied specifically to reject processes that are not well-modelled in simulation, such as cosmic muons or decays-in-flight. The inclusive data-to-simulation disagreement amounts to less than approximately 1%. Given that the effect is so small, we don't apply any scale factors for our tighter selection on these variables.

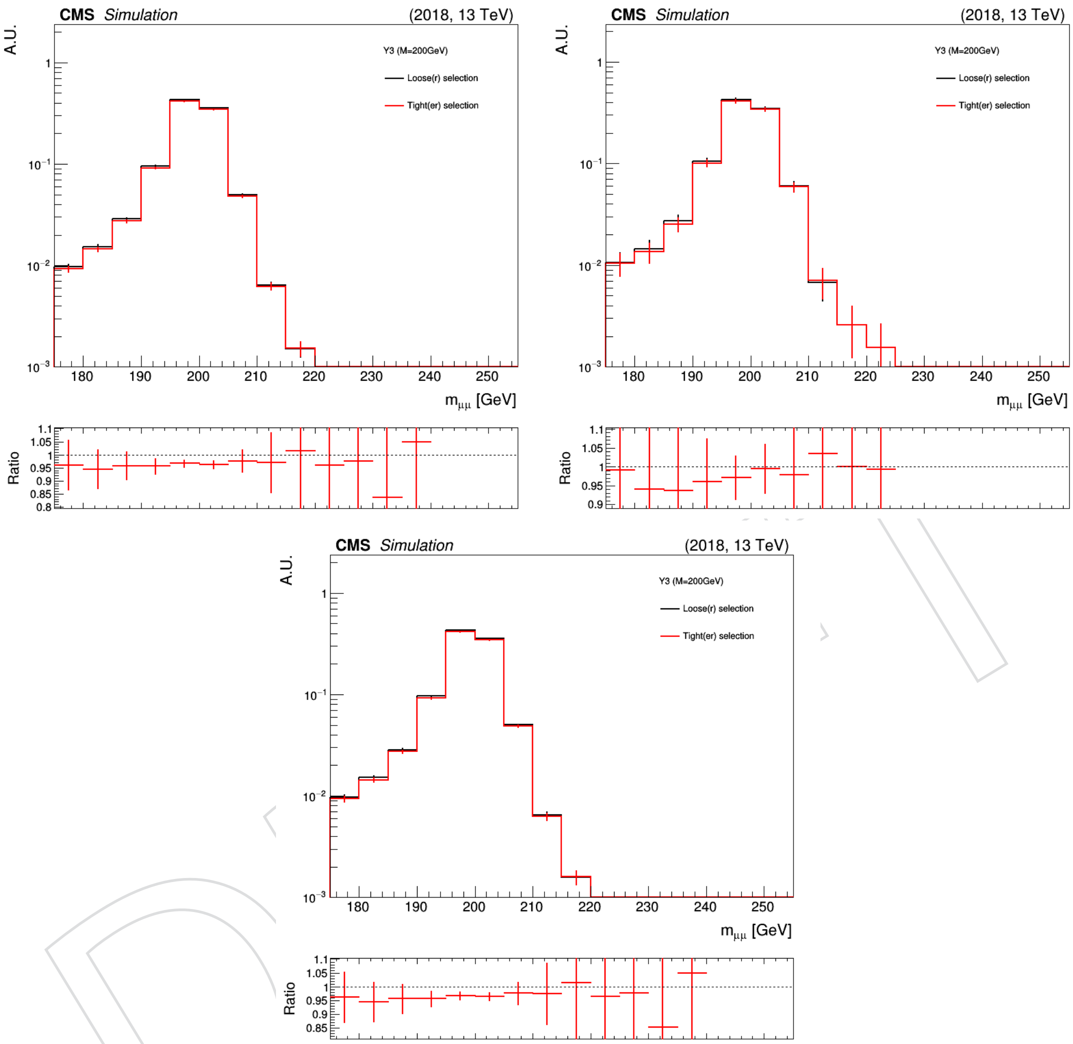


Figure 121: The effect on the expected signal yield and signal $m_{\mu\mu}$ shape from the application of the tight(er) muon selection requirements is shown, as measured in a representative 2018 signal MC sample (Y_3, with $m_{Z'} = 200\,\text{GeV}$), after the full event selection, in event categories with $N_b = 1$ (top left), $N_b \geq 2$ (top right), and $N_b \geq 1$ (bottom; equal to the sum of $N_b = 1$ and $N_b \geq 2$). The black histogram represents the signal dimuon invariant mass distribution with muon identification and isolation criteria as recommended by the Muon POG and used by other similar searches [2], including a relaxed requirement on the muon relative tracker-only isolation (i.e., $< 10\%$ of the muon p_T instead of 5%). The red histogram represents the same distribution once the tighter muon selection requirements are applied. Both distributions are normalized to the area of the histogram with looser muon selection requirements.

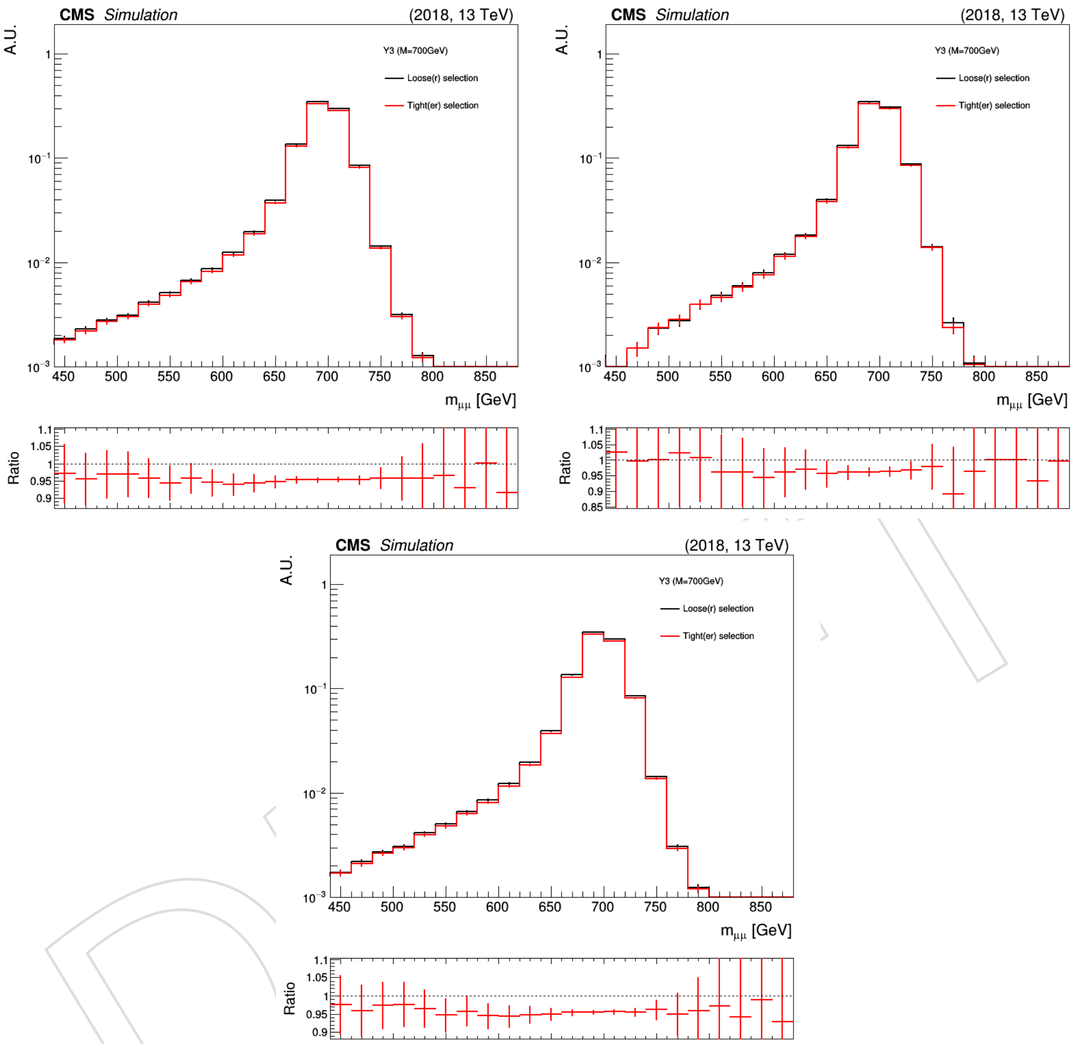


Figure 122: The effect on the expected signal yield and signal $m_{\mu\mu}$ shape from the application of the tight(er) muon selection requirements is shown, as measured in a representative 2018 signal MC sample (Y_3, with $m_{Z'} = 700\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution with muon identification and isolation criteria as recommended by the Muon POG and used by other similar searches [2], including a relaxed requirement on the muon relative tracker-only isolation (i.e., $< 10\%$ of the muon p_T instead of 5%). The red histogram represents the same distribution once the tighter muon selection requirements are applied. Both distributions are normalized to the area of the histogram with looser muon selection requirements.

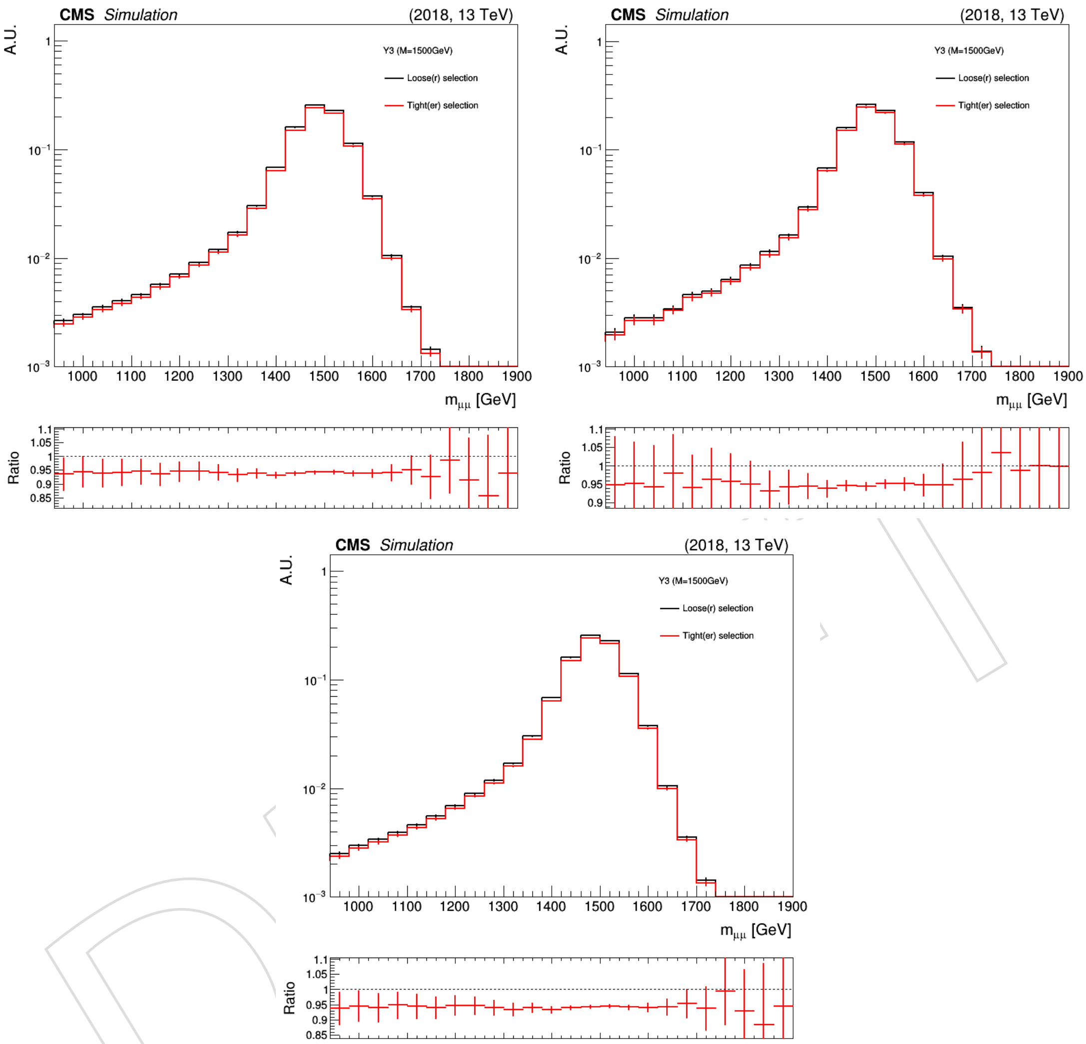


Figure 123: The effect on the expected signal yield and signal $m_{\mu\mu}$ shape from the application of the tight(er) muon selection requirements is shown, as measured in a representative 2018 signal MC sample (Y_3, with $m_{Z'} = 1500\,\text{GeV}$), after the full event selection, in event categories with $N_\text{b} = 1$ (top left), $N_\text{b} \geq 2$ (top right), and $N_\text{b} \geq 1$ (bottom; equal to the sum of $N_\text{b} = 1$ and $N_\text{b} \geq 2$). The black histogram represents the signal dimuon invariant mass distribution with muon identification and isolation criteria as recommended by the Muon POG and used by other similar searches [2], including a relaxed requirement on the muon relative tracker-only isolation (i.e., $< 10\%$ of the muon p_T instead of 5%). The red histogram represents the same distribution once the tighter muon selection requirements are applied. Both distributions are normalized to the area of the histogram with looser muon selection requirements.

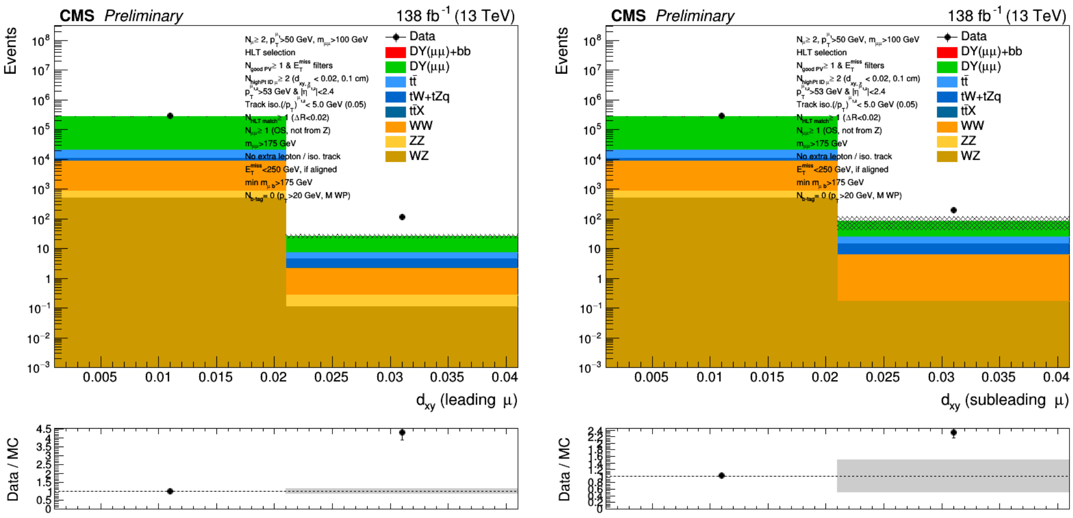


Figure 124: The distributions of the $|d_{xy}|$ variable of the leading (left) and subleading muon (right) in the DY-enriched region. The data and the simulation are compared in two bins, one for the range of values that we select (first bin) and one for the range of values that we reject with the tighter selection (second bin, including overflow yield) of $|d_{xy}| < 0.02$ cm. The data to simulation agreeement is excellent in the range of values we accept in the analysis, while any disagreement in the range of values we exclude with the tighter selection amounts to less than 1% of the total yields.

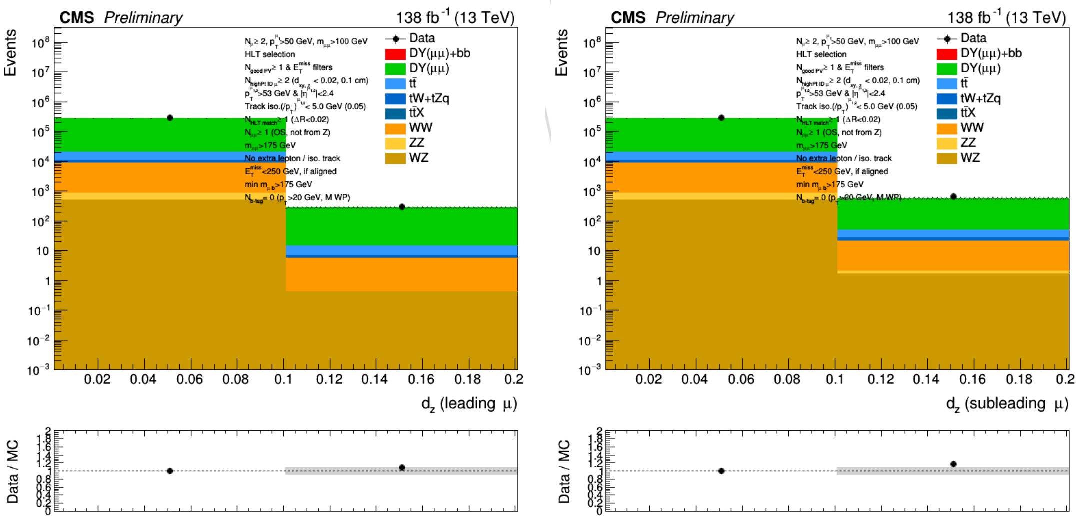


Figure 125: The distributions of the $|d_z|$ variable of the leading (left) and subleading muon (right) in the DY-enriched region. The data and the simulation are compared in two bins, one for the range of values that we select (first bin) and one for the range of values that we reject with the tighter selection (second bin, including overflow yield) of $|d_z| < 0.1$ cm. The data to simulation agreeement is excellent in the range of values we accept in the analysis, while any disagreement in the range of values we exclude with the tighter selection amounts to less than 1% of the total yields.

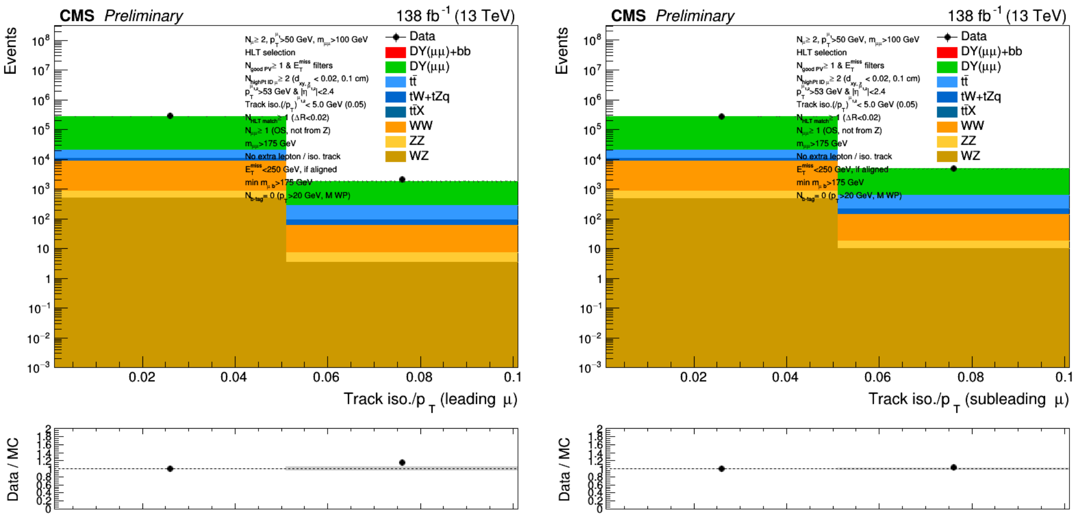


Figure 126: The distributions of the relative tracker-only isolation variable of the leading (left) and subleading muon (right) in the DY-enriched region. The data and the simulation are compared in two bins, one for the range of values that we select (first bin) and one for the range of values that we reject with the tighter selection (second bin, including overflow yield) of $\mathcal{I}^{\mathrm{trk}}_{\mathrm{rel}} < 0.05$. The data to simulation agreement is excellent in the range of values we accept in the analysis, while any disagreement in the range of values we exclude with the tighter selection amounts to less than 1% of the total yields.

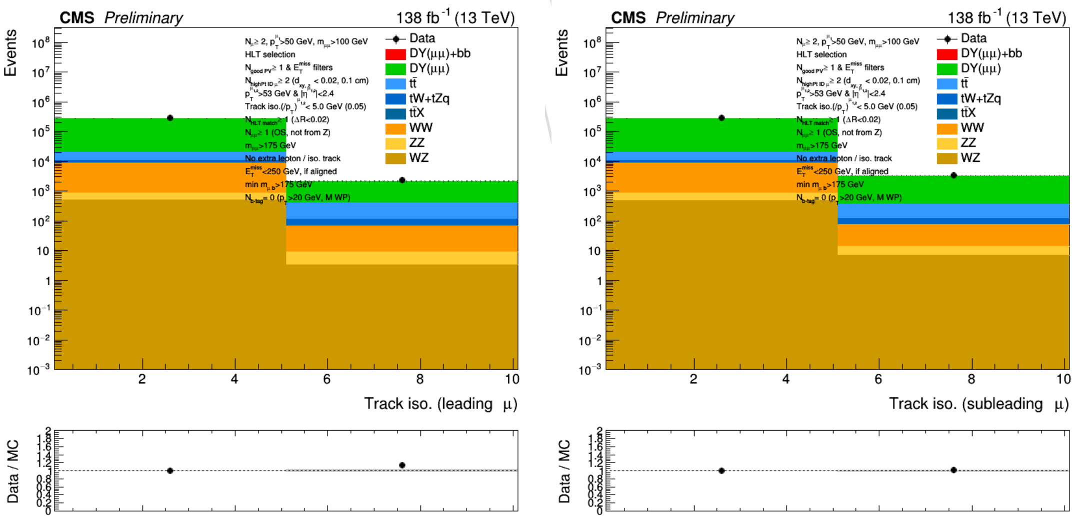


Figure 127: The distributions of the absolute tracker-only isolation variable of the leading (left) and subleading muon (right) in the DY-enriched region. The data and the simulation are compared in two bins, one for the range of values that we select (first bin) and one for the range of values that we reject with the tighter selection (second bin, including overflow yield) of $\mathcal{I}^{\mathrm{trk}} < 5\,\mathrm{GeV}$. The data to simulation agreement is excellent in the range of values we accept in the analysis, while any disagreement in the range of values we exclude with the tighter selection amounts to less than 1% of the total yields.

G Signal Monte Carlo reweighting

The MC samples described in Section 4.2.1 are generated at several Z' masses M_X and at specific (default) values of the coupling constants used in the models. In this section we describe the procedure by which we predict the yields in our two SRs for different sets of masses and coupling constants.

G.1 Reweighting the MC models from the literature

The MC samples desribed in Section 4.2.1.1 were generated with the $x = g_X(1\,\text{TeV}/M_X)$ and θ_{23} parameters listed in Table 4. The generated mass values are 100, 200, 250, 400, 550, 700, 850, 1000, 1250, 1500, 2000 GeV.

The reweighting procedure is based on the fact that the parton-level processes that we have generated and are listed below can be divided into three distinct categories (charge conjugate processes are implied):

1. $gg \to Z'b\bar{b}$, $gb \to Z'b$, $b\bar{b} \to Z'g$
2. $gg \to Z's\bar{b}$, $gs \to Z'b$, $s\bar{b} \to Z'g$
3. $s\bar{s} \to Z'g$

The cross sections for processes 1 (2 and 3) are proportional to the square of the $Z'bb$ ($Z'sb$, $Z'ss$) couplings, and these couplings are fully specified in a given model (Y_3, DY_3, DY_3', $B_3 - L_2$) by the g_X and θ_{23} parameters. The key idea of our reweighting procedure is based on the fact that from Monte Carlo truth it is possible to categorize each events as belonging to one of the three categories (bb, sb, ss).

Events from the third category do not include final state bottom quarks at the matrix element level. As a consequence their contributions to this analysis are negligible.

G.1.1 Reweighting MC yields to different couplings at one of the generated mass values

The on-shell cross sections for a process involving the $Z'q_iq_j$ coupling g_{ij} is proportional to $g^2_{q_iq_j}\Gamma(Z' \to \mu^+\mu^-)/\Gamma(Z')$. This is in turn proportional to $\alpha_{q_iq_j} := \Gamma(Z' \to q_i\bar{q}_j)\Gamma(Z' \to \mu^+\mu^-)/\Gamma(Z')$, where we introduced the parameter $\alpha_{q_iq_j}$ which is a function of M_X, g_X, and θ_{23}, as well as the model (Y_3, DY_3, DY_3', or $B_3 - L_2$).

Thus, to reweight a MC sample in a given model generated with a default M_X, g_X, and θ_{23} to the same model and same mass but different (new) g_X and θ_{23} it is sufficient to rescale each event by the ratio of the α_{ij} parameters calculated at the new and default values of g_X and θ_{23} for the given model.

This procedure does not account for the broadening or sharpening of the Breit-Wigner line shape associated with the change in couplings. This effect, which is neglected due to the narrowness of the Z' width, could in principle be included. However, it can result in a few events with very large weights, with an associated loss of statistical precision even if the $m_{\mu\mu}$ distribution after reconstruction is not supposed to change significantly.

It should be noted however that as the coupling g_X becomes very large, the width of the Z' can become quite large. At some point the assumption that the Z' width Γ is small compared with the $m_{\mu\mu}$ resolution σ breaks down. Eventually perturbation theory itself breaks down, see for example the black dashed lines in Fig. 9 and 10 of Ref. [26]. Contour plots for σ/Γ as a function of $x = g_X(1\,\text{TeV}/M_X)$ and mass are shown in Fig. 128. The "problem" can become significant at high values of x that are perhaps not so interesting, and at masses a bit beyond the reach of this analysis, see for example Table 4 and Figs. 3, 19, and 20. (Admittedly the $B_3 - L_2$ model is a bit of an exception).

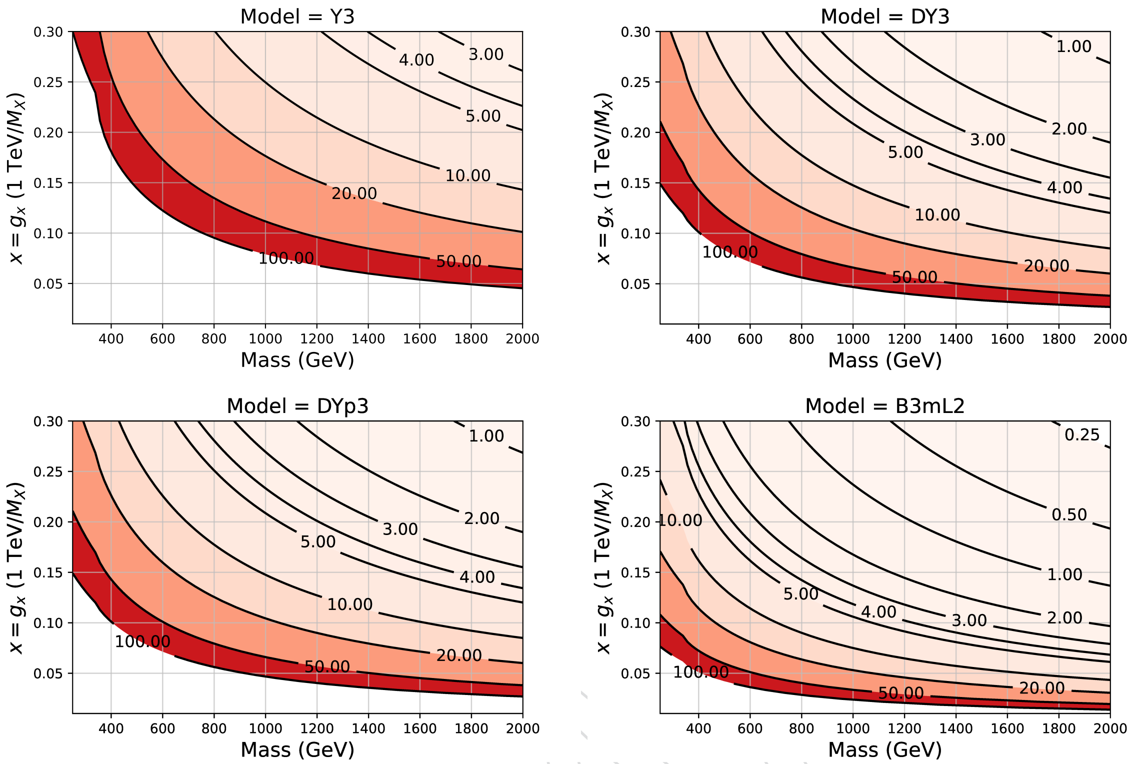


Figure 128: Contours of σ/Γ as a function of x and Z′ mass for different models. Here σ is the $m_{\mu\mu}$ resolution from Fig. 9, and Γ is the Z′ width computed at the model-dependent values of θ_{23} listed in Table 4.

In practice, the reweighting is done by first calculating the expected yields for each SR at the default values of g_X and θ_{23} separately for the three categories bb, sb, ss. These expected yieds are then rescaled by the three appropriate $\alpha_{q_i q_j}$ factors. The total yield in a given SR is then the sum of the yields from the three categories.

G.1.1.1 Check of the reweighting procedure To check the reweighting procedure, we compare the yields obtained directly from Monte Carlo at a particular point A in parameter space (model = DY_3, M_X = 1000 GeV, $x = 0.06$, and $\theta_{23} = 0.20$) with the prediction from the reweighted yields from a different point B (model = DY_3, M_X = 1000 GeV, $x = 0.14$, and $\theta_{23} = 0.13$) Note: the x and θ_{23} parameters from point B are the standard choices from Table 4. For simplicity, the test is performed at generator level, before showering. Kinematical requirements are applied to mimic the events selection (cuts on η, p_T and $m_{\mu b}^{min}$ for muons and b jets, and muon-b overlap removal). Lepton and b tagging efficiencies are taken to be perfect. Out of 60K generated events for point A, 15314 land in SR1 or SR2. Of those, 989 are in SR2. The reweighted yields from point B to point A are $15220 \pm 122 \pm 67$ and $924 \pm 30 \pm 4$, respectively. Here, the first uncertainty is statistical and the second uncertainty is from the Madgraph normalization (basically the spread in the cross-sections reported in different Madgraph runs). The agreement is very good: $0.6 \pm 1.2\%$ and $6.6 \pm 4.4\%$ for SR1 + SR2 and SR2-only, respectively.

G.1.2 Reweighting MC yields to different couplings and different masses

The following ingredients are needed to determine the expected yields for each SR and for each model (Y_3, DY_3, DY_3', $B_3 - L_2$) with arbitrary M_X, g_X, and θ_{23} parameters.

- Parametrizations of the cross sections for each model and for each category (bb, sb, ss) as function of mass with the default $g_X(1\,\text{TeV}/M_X)$ and θ_{23} parameters. These are shown in Fig. 129.
- Parametrizations as a function of mass of the acceptances in each SR for each category. We expect these acceptances to be independent of model, and indeed this is what we find, see Fig. 130.
- The ability to reweight yields at a fixed mass from one set of couplings g_X and θ_{23} to a different one.

For an arbitrary mass M_X, the products of cross sections, acceptances, and integrated luminosity give the expected yields for each SR and for each category for the deafult $g_X(1\,\text{TeV}/M_X)$ and θ_{23} values. The category-dependent SR yields can then be reweighted to arbitrary values of g_X and θ_{23} using the method described in Section G.1.1. In a given SR the yields from the different categories can then be added up to obtain the expected total yield in that SR.

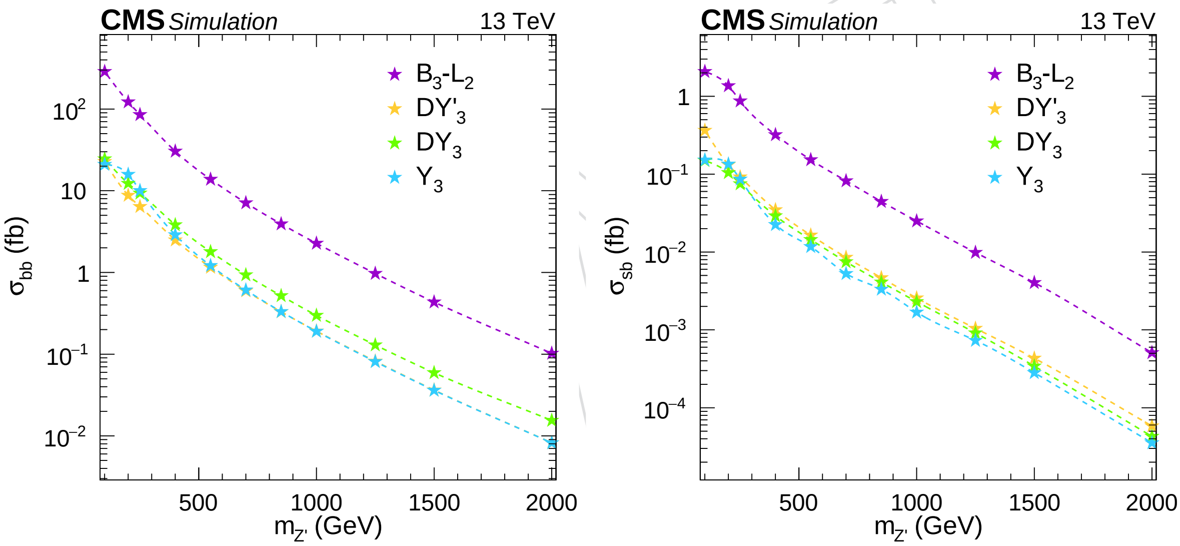


Figure 129: MadGraph cross sections LO in QCD for the bb (left) and sb (right) categories as a function of mass for the four models. The dashed lines are spline fits through the points obtained from the MC generation. The ss cross section is not shown, since this category is essentially irrelevant.

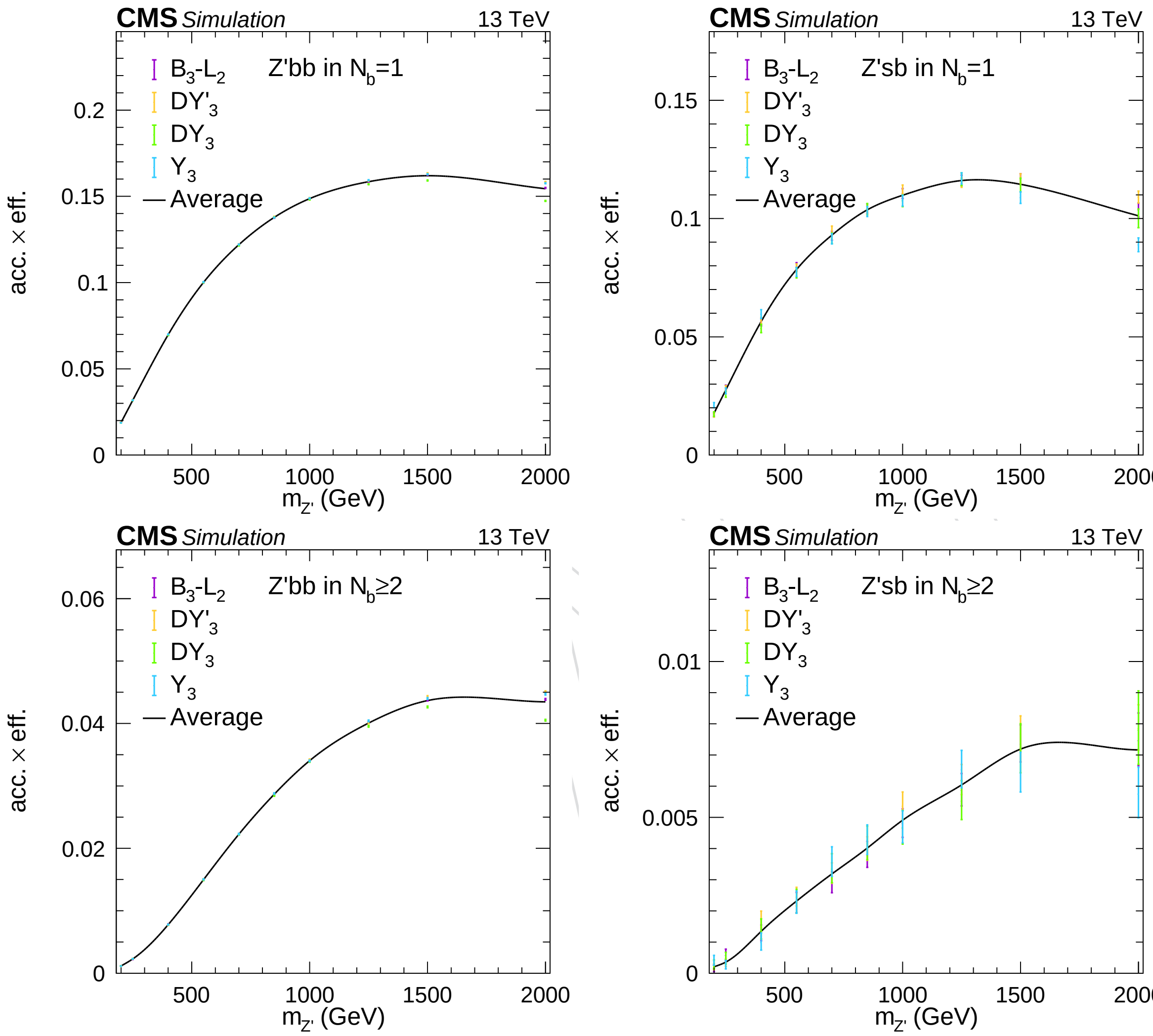


Figure 130: Products of acceptance times efficiency as a function of mass for the four models. Top row, left: SR1 ($N_\mathrm{b} = 1$) bb category. Top row, right: SR1 ($N_\mathrm{b} = 1$) sb category. Bottom row, left: SR2 ($N_\mathrm{b} \geq 2$) bb category. Bottom row, right: SR2 ($N_\mathrm{b} \geq 2$) sb category. The solid lines are spline fit through the model averages. The ss category is not shown, since it is essentially irrelevant.

G.2 Reweighting the BFF models

The reweighting of these models is based on the same ideas described in Section G.1.

There is one small complication due to the fact that the BFF models are generated at NLO in BSM couplings. For example, the parton level process $b\overline{b} \rightarrow Z'b\overline{b}$ at LO in BSM includes only one BSM $Z'bb$ vertex, whereas at NLO there are diagrams with three BSM vertices (either three $Z'bb$ or one $Z'bb$ and two $Z'bs$, see Fig. 131). As a result, the squared amplitude for this process would go as $|\mathcal{A}|^2 \approx g_b^2|\alpha_S + ag_b^2 + b\delta_{sb}^2 g_b^2|^2$, where a and b take some complicated numerical values that we do not attempt to calculate (!). An exact reweighting at NLO in BSM is then not possible. There are also processes that exist at NLO in BSM but not at LO, e.g., $b\overline{s} \rightarrow Z's\overline{b}$. However, to the extent that the BSM couplings are small, one suspects that the NLO BSM effects in this model are negligible.

To verify this conjecture, we generated a single BFF MC sample with $M_X = 1$ TeV at LO in BSM, and compared our expected yields with those obtained from the equivalent sample generated at NLO in BSM. We find that the cross sections and the expected yields are the same to better than 1%. As a result, we can reweight the BFF models generated at NLO in BSM simply based on the uniquely determined LO BSM coupling for any given parton level process. For example, for the $b\overline{b} \rightarrow Z'b\overline{b}$ process shown in Fig. 131, the reweighting is just based on one power of the $Z'bb$ coupling in the amplitude.

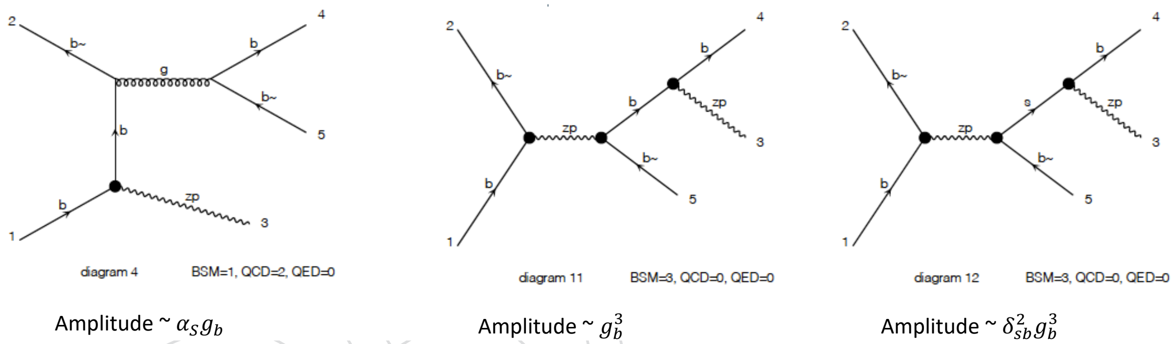


Figure 131: Some of the diagrams for $b\overline{b} \rightarrow Z'b\overline{b}$. The diagram at the left is LO in BSM, while the other two are NLO in BSM. Also shown are the coupling factors that enter in each amplitude.

FIXME: add plots of cross section and splined acceptances. Note that the splined acceptances will not be the same as those in the Allanach model section because the set of diagram being generated is diferent. For example in BFF we generate $b\bar{b} \rightarrow Z'$, which will have very low acceptance.

H Sensitivity comparison with EXO-22-006

In this Section we summarize the studies that were performed to compare the sensitivity of this analysis to the similar analysis in EXO-22-006. The differences in the two analysis strategies can be summarized as follows:

- In this analysis the background is predicted by fitting the dimuon invariant mass to analytic functions, while in EXO-22-006 the background is estimated from an ABCD method that uses dielectron events as well as events with no b tags.
- Both analyses use two signal regions. In this analysis the SRs are distinguished by the b jet multiplicity ($N_b = 1$ and $N_b > 1$). In EXO-22-006 the signal regions are disinguished by the inclusive jet multiplicity ($N_b = N_j = 1$ and $N_b \geq 1$, $N_j > 1$).
- In this analysis the strategy is to essentially eliminate the $t\bar{t}$ background, see for example Fig. 17. The price that one has to pay is a low acceptance for signal events at low Z' mass, see for example Fig. 7 (left) and Fig. 129. In contrast, the requirements against $t\bar{t}$ in EXO-22-006 are less stringent. This approach leads to a larger $t\bar{t}$ background but a better acceptance at low Z' mass.

A comparison in signal sensitivity was performed by the two groups under the following common conditions:

- The background was predicted using common $t\bar{t}$ and Drell-Yan non-UL samples, 2016 MC only, no corrections or scale factors.
- The BFF signal samples were non-UL with $\delta_{bs} = 0.04$, mass = 250, 350, 500 GeV.
- The Y3 sampless were UL with mass=250, 400, 700, 1000 TeV.
- The expected signal and background yields were computed by adding up the events within $\pm 10\%$ of the mass hypothesis.

The number of expected signal and background events as a function of mass, as defined above, are shown in Fig. 132. Approximate expected signal strength limits are also extracted from this study as a function of mass in the "counting experiment" approximation. Results are shown in Fig. 133. These studies demonstrate that the EXO-22-006 analysis is more sensitive at low mass, while this analysis is more sensitive at high mass. The cross-over point is somewhere in the neighborhood of 300-350 GeV.

More details on these studies have been prsented at the EXO non-hadronic meeting of August 10, 2022 [74].

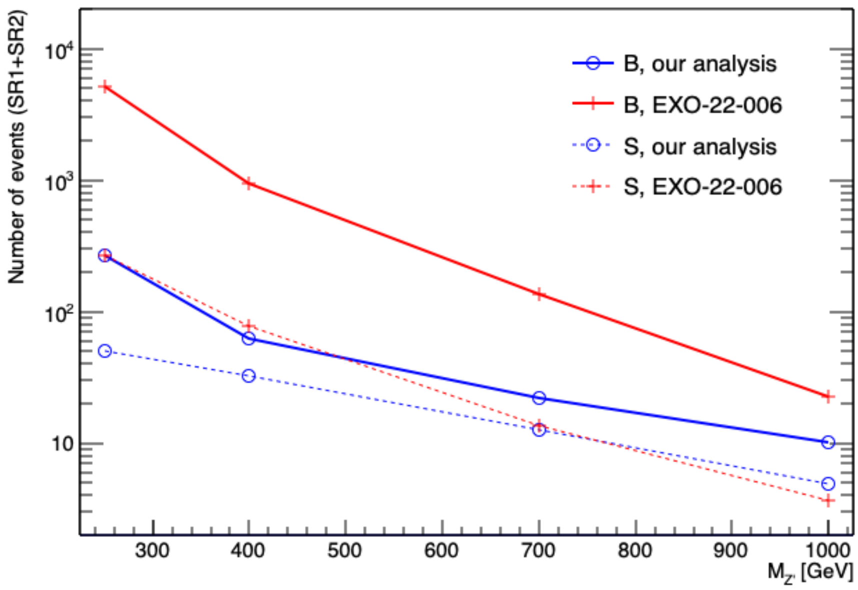
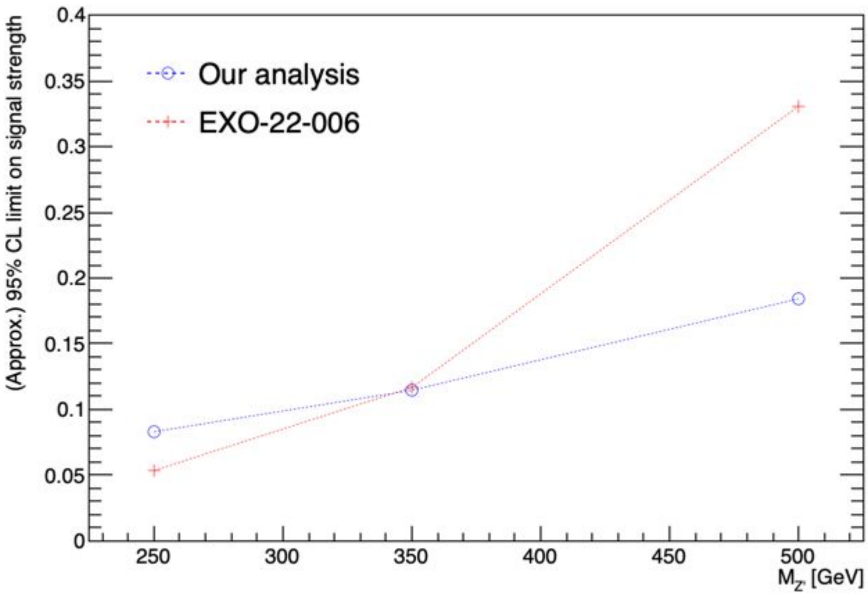


Figure 132: Signal and background yields as a function of mass for this ananysis and the EXO-22-006 analysis. The yields are defined as the sum of the yields in the two signal regions (SR1 and SR2) in the respective analyses. The signal yields are from the Y3 Monte Carlo.



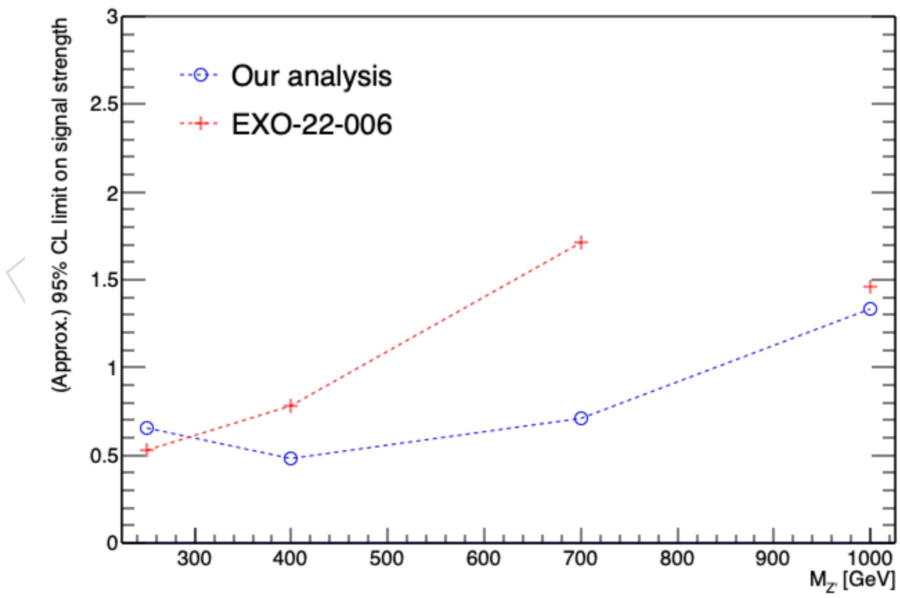


Figure 133: Approximate 95% CL signal strength limits for this ananysis and the EXO-22-006 analysis. Top: BFF model. Bottom: Y3 model. The last point (mass=1 TeV) for the EXO-22-006 analysis in the bottom plot cannot be trusted due to statistical fluctiations in the MC that was processed.