

# CMS Draft Analysis Note

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

2021/08/27

Head Id: 388561

Archive Id: 447036

Archive Date: 2017/02/21

Archive Tag: trunk

## Search for displaced leptons with 2016–2018 data

J. Alimena<sup>1</sup>, L. Antonelli<sup>2</sup>, B. Cardwell<sup>2</sup>, C. Hill<sup>2</sup>, E. Nunez Ornelas<sup>2</sup>, and I. Tomalin<sup>3</sup>

<sup>1</sup> CERN

<sup>2</sup> Ohio State University

<sup>3</sup> Rutherford-Appleton Laboratory

### Abstract

A search for new long-lived particles decaying to leptons is presented using proton-proton collisions produced by the CERN LHC at a center-of-mass energy of 13 TeV. Events are selected with two leptons (an electron  $e$  and a muon  $\mu$ , two electrons, or two muons) that both have transverse impact parameter values between 0.01 cm and 10 cm. Data used for the analysis were collected by the CMS detector in 2016, 2017, and 2018 and correspond to an integrated luminosity of  $113 \text{ fb}^{-1}$  for the  $e\mu$  and  $\mu\mu$  channels and  $118 \text{ fb}^{-1}$  for the  $ee$  channel. The search is designed to be sensitive to a wide range of models with nonprompt  $e\mu$ ,  $ee$ , and  $\mu\mu$  final states. The results are interpreted with a model involving top squarks that decay into displaced leptons via R-parity-violating interactions, a gauge-mediated supersymmetry breaking model with lepton superpartners that decay to gravitinos and displaced leptons, and a model involving exotic Higgs bosons that decay to long-lived scalars, which in turn decay to displaced leptons. This is the first search at CMS for displaced leptons at a center-of-mass energy of 13 TeV, and the first search at CMS for displaced leptons in the  $ee$  and  $\mu\mu$  channels that does not require the leptons to come from a common displaced vertex.

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

PDFAuthor: Displaced SUSY Team

PDFTitle: Search for displaced leptons with 2016, 2017, and 2018 data

PDFSubject: CMS

PDFKeywords: CMS, physics, supersymmetry

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



## 1 Contents

2	1	Introduction and theoretical motivation . . . . .	5
3	2	Data and Monte Carlo simulation samples . . . . .	7
4	2.1	Data samples . . . . .	7
5	2.2	Standard model background process simulation . . . . .	7
6	2.3	Signal process simulation . . . . .	7
7	3	Event selection . . . . .	8
8	3.1	Triggers . . . . .	8
9	3.2	Preselection . . . . .	9
10	3.3	Prompt control region . . . . .	18
11	3.4	Inclusive signal search region . . . . .	18
12	4	Corrections to the simulation . . . . .	25
13	4.1	Event pileup reweighting . . . . .	25
14	4.2	Lepton ID corrections . . . . .	25
15	4.3	Lepton $d_0$ corrections . . . . .	26
16	4.4	Trigger efficiency correction . . . . .	30
17	5	Background sources . . . . .	33
18	6	Background estimation . . . . .	34
19	6.1	Data-driven ABCD method . . . . .	34
20	6.2	Signal search regions (SRs) . . . . .	36
21	6.3	Closure tests in one-prompt/one-displaced sidebands . . . . .	36
22	6.4	Correction to estimate and systematic uncertainty . . . . .	38
23	6.5	Closure tests in signal regions . . . . .	42
24	7	Additional background checks . . . . .	44
25	7.1	Material interactions . . . . .	44
26	7.2	Cosmic-ray muons . . . . .	44
27	7.3	Heavy-flavor mesons . . . . .	44
28	7.4	Low-mass SM hadrons . . . . .	46
29	8	Systematic uncertainties in the signal efficiency . . . . .	47
30	8.1	Integrated luminosity . . . . .	47
31	8.2	Pileup . . . . .	47
32	8.3	Tracking efficiency for displaced particles . . . . .	47
33	8.4	Trigger efficiency . . . . .	54
34	8.5	Lepton ID and isolation . . . . .	54
35	8.6	Muon pixel hit efficiency . . . . .	55
36	8.7	Lepton $ d_0 $ smearing . . . . .	55
37	8.8	Summary of systematic uncertainties in the signal efficiency . . . . .	55
38	9	Results . . . . .	56
39	9.1	Observed events . . . . .	60
40	9.2	Limits . . . . .	60
41	10	Summary . . . . .	64
42	A	Study with $K_S^0$ . . . . .	69
43	B	Privately generated signal samples . . . . .	71

44	C	Mismeasured lepton $ d_0 $ at large $ \eta $	73
45	D	Additional plots	74
46	D.1	GMSB sleptons: kinematic distributions and cutflows	74
47	D.2	Exotic Higgs bosons: kinematic distributions and cutflows	78
48	D.3	Modified isolation	82
49	D.4	Prompt control regions	84
50	D.5	Lepton $d_0$ corrections	100
51	D.6	Background estimation	102
52	D.7	Data and background simulation in the inclusive signal region	103
53	E	Likelihood tests	115
54	F	Study of R-hadron treatment in simulation	118

DRAFT

## 55 Change log

### 56 Changes wrt v12

- 57 • Implemented first round of ARC comments

### 58 Changes wrt v11

- 59 • Changed to take symmetric systematic uncertainties in background estimation for  
60 channels/years where no correction is needed (see Sections 6 and 9)
- 61 • Unblinded results included, including a brief description of the observed events in  
62 the signal regions and the observed limits (see Section 9)
- 63 • Limits updated to include the signal systematic uncertainty updates already in-  
64 cluded in v11 (see Section 9)
- 65 • Added appendix with likelihood tests (see Appendix E)

### 66 Changes wrt v10

- 67 • Added additional signal systematic uncertainty for muon pixel hit efficiency in sig-  
68 nals (see Section 8, not yet applied to limits)

### 69 Changes wrt v9

- 70 • Added an appendix with performance plots for the modified isolation (see Sec-  
71 tion D.3)

### 72 Changes wrt v8

- 73 • Added additional signal systematic uncertainty to cover change in muon trigger  
74 efficiency in signal at large  $|d_0|$  (see Section 8)
- 75 • Updated results including the limits

### 76 Changes wrt v7

- 77 • Improved trigger efficiency scale factors by making the lepton ID closer to what is  
78 used elsewhere in the analysis (see Section 4.4)
- 79 • Added 500um–10cm closure tests using background MC with the DY $\rightarrow \tau\tau$  events  
80 removed (see Section 6.3.2)
- 81 • Added section to describe additional studies of backgrounds (see Section 7).

### 82 Changes wrt v6

- 83 • Added description of additional models we will use to interpret the results, and  
84 some placeholders for limits (see Sections 1, 2, 9).
- 85 • Added additional systematic uncertainty to background estimate, from varying ex-  
86 trapolation point by  $\pm 50 \mu\text{m}$  (see Section 6.4.1). Not yet applied to results (Section 9).

### 87 Changes wrt v4

- 88 • Replaced lifetime reweighted 0.1 mm signal samples with privately generated sam-  
89 ples
- 90 • Updated preselection criteria (see Section 3.2), including:

- 91     • Tighter muon  $\eta$  cut (also see the new Appendix C)
- 92     • A cut to remove background from material interactions
- 93     • Updated cut flows (see Section 3.2) and remade prompt control region data/MC
- 94       plots (see Section 3.3)
- 95     • Updated background estimation method (see Section 6)
- 96     • Finished signal systematic uncertainties for pileup, lepton ID and isolation, and lep-
- 97       ton d0 smearing (see Section 8)
- 98     • Updated results and limits. The limits were produced with Hybrid New, the  $\tilde{t}\tilde{t} \rightarrow$
- 99        $d\bar{l} d\bar{l}$  limits are added, and the 3 channels have been combined. The limits include the
- 100      interpolated lifetime points (from reweighting the generated signal). All systematic
- 101      uncertainties are included in the limits (see Section 9)

## 102   Changes wrt v2

- 103     • Updated preselection criteria (see Section 3.2), including:
  - 104       • A custom PV-agnostic PF isolation
  - 105       • Tighter  $p_T$  cuts in the  $e\mu$  and  $\mu\mu$  channels, after measuring where the trig-
  - 106       ger turn-on plateau starts
  - 107       • Tighter electron  $\eta$  cut, to better reject electrons from heavy flavor and in
  - 108       order to have better measured  $|d_0|$
  - 109       •  $\Delta R$  cuts, to reject leptons from heavy flavor
  - 110       • Veto leptons in eta-phi regions affected by pixel power supply issues
  - 111       • Remove the overlap among channels' signal regions
- 112     • Updated cut flows (see Section 3.2) and remade prompt control region data/MC
- 113       plots (see Section 3.3)
- 114     • Updated trigger efficiency scale factors (see Section 4.4)
- 115     • Added a section on background sources (see Section 5)
- 116     • Changed background estimation method (see Section 6)
- 117     • Updated signal systematic uncertainties in the trigger efficiency and the displaced
- 118       tracking efficiency (see Section 8)
- 119     • Updated results, including limits (see Section 9)

## 120 1 Introduction and theoretical motivation

121 To date, no direct evidence of a particle beyond the standard model (BSM) has been found  
 122 by the CERN LHC experiments or any other facilities. However, the vast majority of LHC  
 123 searches assume the lifetimes of the new particles are short enough so that their decay prod-  
 124 ucts are prompt. The search strategies are often not optimized for particles with measurable  
 125 lifetimes whose decays produce displaced signatures in the detectors. Therefore, new physics  
 126 with displaced signatures can escape such searches and be less constrained by the current re-  
 127 sults.

128 While the majority of searches are only sensitive to prompt signatures, both the CMS and AT-  
 129 LAS Collaborations have performed dedicated searches for decays of BSM particles with long  
 130 lifetimes. Direct search strategies include finding BSM particles via anomalous energy loss  
 131 and/or low velocity [1–3] or via a disappearing track signature [4–6]. There are also numer-  
 132 ous indirect searches targeting the decay products of long-lived particles, such as nonprompt  
 133 final-state jets [7–11], photons [12–14], leptons [15–20], or combinations thereof [21, 22]. For  
 134 simplicity, we have just cited the most recent updates of these searches here.

135 We performed two previous CMS searches in the displaced electron-muon ( $e\mu$ ) final state [16,  
 136 23]. These were optimized to the phase space just beyond the sensitivity of prompt searches  
 137 but with smaller displacements than other searches for long-lived BSM signatures. These were  
 138 unique in that they dispensed with the usual requirement that the displaced final state par-  
 139 ticles originate from a common vertex. The search described here is an update of the search  
 140 described in Ref. [16] using the data taken during 2016, 2017, and 2018. However, we have also  
 141 implemented a number of changes and additions; most notably, we have changed the back-  
 142 ground estimation technique, added two same-flavor channels, and added an interpretation of  
 143 the results.

144 This search is designed to be maximally model independent and to be sensitive to as many  
 145 event topologies as possible. Consequently, the event selection focuses exclusively on a dis-  
 146 placed, isolated dilepton signature and does not try to identify signal events using hadronic  
 147 activity or missing transverse momentum ( $p_T^{\text{miss}}$ ). In this way, we retain sensitivity to any  
 148 model that can produce leptons with displacements on the order of 0.01 cm to 10 cm, regardless  
 149 of whether these leptons are accompanied by jets,  $p_T^{\text{miss}}$ , or other interesting kinematic features.

150 We interpret the search results in the context of a “displaced supersymmetry” [24] model. This  
 151 model introduces R-parity violating terms in the superpotential of the minimally supersym-  
 152 metric standard model. R-parity violation (RPV) allows the lightest supersymmetric particle  
 153 (LSP) to decay into standard model (SM) particles. This can frustrate standard SUSY searches  
 154 by lessening or removing the missing energy signature generally present in supersymmetry  
 155 (SUSY) topologies. Only lepton number-violating operators are considered in order to avoid  
 156 constraints from proton decay. Due to sufficiently small couplings for these operators, the LSP  
 157 will have a long enough lifetime that its decay products are measurably displaced from the re-  
 158 gion of beam-beam overlap in which the proton-proton (pp) interactions occur. For a range of  
 159 these couplings, displaced supersymmetry can generate a natural SUSY model that could have  
 160 easily escaped detection thus far at the LHC, because the displacement of the decay products  
 161 will be in the gap of sensitivity between the previous searches based on prompt decays and  
 162 those based on very long-lived BSM signatures.

163 We focus on the case in which the LSP is the top squark, the superpartner of the top quark.  
 164 At the LHC, top squarks would be dominantly produced in pairs. We generate the top squark  
 165 decay through an RPV vertex,  $\tilde{t} \rightarrow q l$ , where  $\tilde{t}$  is a top squark,  $q$  is a b or d quark, and  $l$  is an

electron, muon, or tau lepton ( $\tau$ ) (see Fig. 1). The decay to a d quark and a lepton has been added since the 2015 analysis. For simplicity and to reduce dependency on any specific model, we assume lepton universality in the top squark decay vertex, so that the branching fraction to any lepton flavor is equal to one third. We conduct a search for top squark decays in which there are one electron and one muon, two electrons, or two muons in the final state, where both of the leptons are displaced from the nominal proton-proton (pp) collision interaction point (IP)

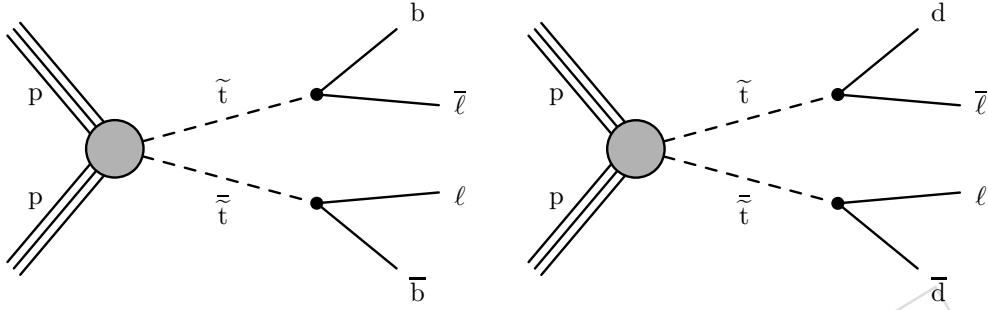


Figure 1: Feynman diagrams for  $\tilde{t}$  pair production and  $\tilde{t}$  decay to a lepton and a b (left) or a d quark (right) via an R-parity violating vertex.

We also interpret the results with two other models (see Fig. 2). The first is a gauge-mediated SUSY breaking model in which the next-to-lightest SUSY particle (NLSP) is long-lived due to its small gravitational coupling to the LSP gravitino  $\tilde{G}$ , which is nearly massless [25]. In this model, the NLSP is a slepton  $\tilde{l}$ ; we consider selectrons  $\tilde{e}$ , smuons  $\tilde{\mu}$ , and staus  $\tilde{\tau}$  separately as well as together as co-NLSPs. The sleptons would be pair-produced at the LHC and which would decay to a lepton of the same flavor and a gravitino. In addition, we consider a model that produces BSM Higgs bosons (H) through gluon-gluon fusion [26]. The exotic H decays to two long-lived scalars S, each of which decays to two oppositely-charged leptons, which have equal probability of being an electron or a muon.

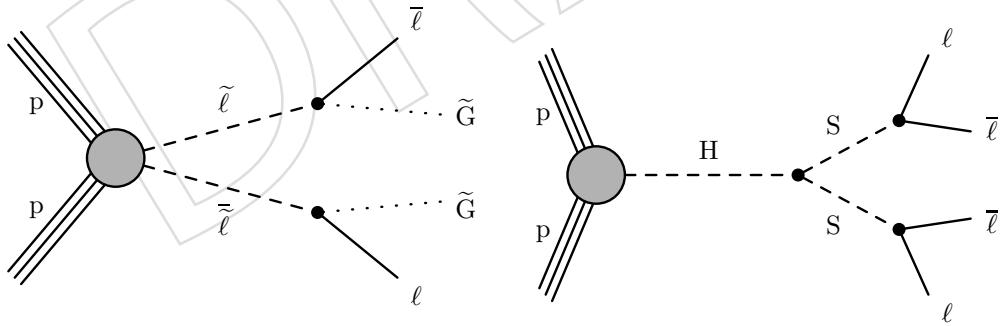


Figure 2: Feynman diagrams for  $pp \rightarrow \tilde{l}\tilde{l} \rightarrow \tilde{l}\tilde{G}\tilde{l}\tilde{G}$  (left), and  $pp \rightarrow H \rightarrow SS \rightarrow \tilde{l}\tilde{l}\tilde{l}\tilde{l}$  (right).

This note is organized as follows. Section 2 catalogs all of the LHC datasets and simulated datasets used in this search and gives relevant details about their production. Section 3 details the event selections that were used to perform the search. Section 4 explains the corrections made to the simulations. Section 5 explains the major background processes for this search, Section 6 describes the methods employed to estimate their contributions, and Section 7 describes additional checks that were performed related to the backgrounds. Section 8 covers the systematic uncertainties considered in the final results of the search. Section 9 gives the results

189 of the search and shows the limits placed on different theoretical models. Finally, Section 10  
 190 contains a summary of the search.

## 191 **2 Data and Monte Carlo simulation samples**

### 192 **2.1 Data samples**

193 This analysis uses pp collision data taken in 2016, 2017, and 2018 at a center-of-mass energy  
 194  $\sqrt{s} = 13$  TeV, applying the Golden JSON files. Only the last run periods in 2016 (eras G and  
 195 H) are considered in order to avoid the periods with lower displaced tracking efficiency due to  
 196 the analog pipeline voltage (APV) saturation problem in the silicon strip detector. Appendix  
 197 A provides a detailed study of this effect using  $K_S^0$ . Thus, in all three channels, an integrated  
 198 luminosity of  $16.1 \pm 0.4 \text{ fb}^{-1}$  from 2016 is used. Furthermore, in 2017, some of the triggers used  
 199 in this analysis were not available for the entire run period, and so an integrated luminosity of  
 200  $41.5 \pm 1.0$  ( $36.7 \pm 0.8$ )  $\text{fb}^{-1}$  from 2017 is used in the ee ( $e\mu$  and  $\mu\mu$ ) channel. In all three channels,  
 201 an integrated luminosity of  $59.7 \pm 1.5 \text{ fb}^{-1}$  from 2018 is used.

202 The MuonEG primary dataset is used as our search sample in the  $e\mu$  channel, as described in  
 203 Section 3. The DoubleEG (EGamma) primary dataset is used as our search sample in the ee  
 204 channel in 2016–2017 (2018). The DoubleMu primary dataset is used as our search sample  
 205 in the  $\mu\mu$  channel. The MET dataset is used for the trigger efficiency studies in Section 4.4,  
 206 and the Cosmics and NoBPTX datasets is used in the displaced tracking efficiency studies  
 207 as described in Section 8.3. All data collected in 2016, 2017, and 2018 were reconstructed  
 208 in the 07Aug17, 31Mar2018, 17Sep2018 reprocessing campaigns with software versions  
 209 CMSSW\_8\_0\_31, CMSSW\_9\_4\_8, and CMSSW\_10\_2\_0, respectively, and the MiniAOD event for-  
 210 mat. However, for the EGamma 2018D dataset, the 22Jan2019 campaign is used, following the  
 211 PPD recommendation.

### 212 **2.2 Standard model background process simulation**

213 This analysis employs a fully data-driven background estimation technique. However, centrally-  
 214 produced SM simulation is used to verify the compositions of various control regions. The  
 215 samples corresponding to 2016, 2017, and 2018 data conditions are from the  
 216 PdmVMCcampaignRunIIISummer16, PdmVMCcampaignRunIIFall117, and  
 217 PdmVMCcampaignRunIIIAutumn18 production campaigns and were reconstructed in  
 218 CMSSW\_8\_0\_31, CMSSW\_9\_4\_8, CMSSW\_10\_2\_0 with the MiniAODSIM event format. The sam-  
 219 ples simulating Z+jets, W+jets, and tt production are generated using MADGRAPH5\_aMC@NLO [27–  
 220 29], while those simulating diboson (WW, WZ, and ZZ with leptonic and semi-leptonic decays)  
 221 and single-top-quark production are simulated with POWHEG v2 [30–34]. PYTHIA 8.2 [35] is  
 222 used to simulate the parton showering and hadronization for all processes. The modeling of  
 223 the underlying event is generated using the CUETP8M1 [36] and CP5 tunes [37] for simulated  
 224 samples corresponding to the 2016 and 2017–18 data sets, respectively.

### 225 **2.3 Signal process simulation**

226 Monte Carlo simulation samples of the process  $\text{pp} \rightarrow \tilde{t}\tilde{t}$ , with the top squarks decaying via  $\tilde{t} \rightarrow b l$   
 227 or  $\tilde{t} \rightarrow d l$ , where b is a b quark and d is a d quark, were generated using PYTHIA 8.2 [35] at  
 228 leading order. (The decay to a d quark and a lepton has been added since the 2015 analy-  
 229 sis [16].) For simplicity, all lepton flavors are generated with equal branching fractions. The top  
 230 squarks can form strongly produced hadronic states called “R-hadrons,” which are generated  
 231 with PYTHIA. In the samples used in this analysis, the interactions of the R-hadrons with matter

are not simulated in GEANT4. However, such interactions are not expected to have a significant impact because these particles do not traverse a significant number of interaction lengths before decaying. As described in Appendix F, we nevertheless study the impact of the R-hadron interactions using the “cloud model,” which assumes that the top squark is surrounded by a cloud of colored, light constituents that interact during scattering [38, 39], and find the effect on the signal efficiency to be negligible. Starting with a SUSY Les Houches Accord (SLHA) file [40] corresponding to Snowmass Points and Slopes (SPS) point 1a [41], the mass and width of the top squark were modified according to the sample being produced. Samples were generated for  $\tilde{t}$  masses from 100 GeV to 1800 GeV at 100 GeV intervals, and for  $\tilde{t}$  proper decay lengths at each decade from 0.1 mm to 1 m. Most of the signal samples were generated centrally, although a few needed to be made privately (see Appendix B). The production cross sections for each  $\tilde{t}$  mass hypothesis are taken from the website of the LHC SUSY Cross Section Working Group.

The signal process  $pp \rightarrow \tilde{l}\tilde{l} \rightarrow \tilde{l}\tilde{G}\tilde{l}\tilde{G}$  is generated at leading order using MADGRAPH5\_aMC@NLO. Samples were generated for  $\tilde{l}$  masses from 50 GeV to 900 GeV, and for  $\tilde{l}$  proper decay lengths at each decade from 0.1 mm to 1 m. These samples were produced privately (see Appendix B).

The signal process  $pp \rightarrow H \rightarrow SS \rightarrow l\bar{l} l\bar{l}$  is generated using POWHEG v2 and PYTHIA 8.2 at next-to-leading order. Samples were generated for  $H$  masses from 125 GeV to 1000 GeV, for various  $S$  masses, and for  $S$  proper decay lengths at each decade from 1 mm to 10 m. These samples were produced privately (see Appendix B).

A proper decay length reweighting technique was used to alter the signal samples to correspond to an additional nine points between the original proper decay length in a sample and the next-shortest-proper decay length sample available, and in the case of the  $\tilde{t}$  1 m samples, also to correspond to proper decay lengths up to 10 m. All signal samples were reconstructed in the same campaigns and with the same conditions as described above for the SM background samples.

## 3 Event selection

### 3.1 Triggers

The events are required to pass different triggers in the three channels. Because standard CMS electron triggers are not designed to recognize displaced tracks, we use photon triggers instead to ensure efficiency for finding displaced electrons.

In the  $e\mu$  channel, in 2016 data and corresponding MC simulation, the events are required to pass the logical OR of two HLT paths, `HLT_Mu38NoFiltersNoVtx_Photon38_CaloIdL_v*` OR `HLT_Mu28NoFiltersNoVtxDisplaced_Photon28_CaloIdL_v*`, which were designed for this analysis. These triggers require at least one L3 muon with  $p_T > 38(28)$  GeV, without any of the usual vertex constraints and without any maximum requirement on the impact parameter. The second trigger also requires that the absolute value of the L3 muon transverse impact parameter ( $|d_0|$ ) is greater than 0.01 cm. (The transverse impact parameter will be described in more detail in Section 3.2.) These two triggers also require at least one photon with  $E_T > 38(28)$  GeV and with loose calorimeter ID. We use a photon in these triggers instead of an electron so that the efficiency of selecting a displaced electron will be high, since there are no tracking requirements. The HLT paths are seeded by `L1_Mu5_EG20` OR `L1_Mu20_EG15`. Compared to the standard muon-photon HLT paths, the signal efficiency with these dedicated triggers is significantly higher.

In 2017 and 2018 data and corresponding MC simulation, the events are required to pass

276 HLT\_Mu43NoFiltersNoVtx\_Photon43\_CaloIdL\_v\*. The muon  $p_T$  and photon  $E_T$  thresholds were  
 277 raised with respect to 2016 due to increased pileup. Furthermore, no version of this trigger with  
 278 displaced muons was available in 2017 or 2018.

279 In the ee channel, in 2016 data and corresponding MC simulation, the events are required to  
 280 pass the logical OR of two HLT paths, HLT\_Diphoton30\_18\_R9Id\_OR\_IsoCaloId\_AND\_HE\_R9Id\_Mass90\_v\*  
 281 OR HLT\_DoublePhoton60\_v\*. The first trigger requires a leading photon with  $E_T > 30\text{ GeV}$  and  
 282 a subleading photon with  $E_T > 18\text{ GeV}$ . Calorimeter identification, isolation,  $H/E$ , and  $R_9$   
 283 requirements are made on both photons, and the diphoton invariant mass must be  $> 90\text{ GeV}$ .  
 284 This HLT path is seeded by a suite of single-photon and double-photon L1 seeds. This path  
 285 was chosen because it is highly efficient at low  $\tilde{t}$  mass. The second trigger simply requires at  
 286 least two photons with  $E_T > 60\text{ GeV}$ . This HLT path is seeded by a suite of nonisolated single-  
 287 photon, double-photon, single-jet, and single-tau-jet L1 seeds. This path is highly efficient at  
 288 large  $\tilde{t}$  mass and long lifetime.

289 In 2017 and 2018 data and corresponding MC simulation, the events are required to pass  
 290 HLT\_Diphoton30\_22\_R9Id\_OR\_IsoCaloId\_AND\_HE\_R9Id\_Mass90\_v\* OR HLT\_DoublePhoton70\_v\*.  
 291 The photon  $E_T$  thresholds were raised with respect to 2016 due to increased pileup.

292 In the  $\mu\mu$  channel, in 2016 data and corresponding MC simulation, the events are required to  
 293 pass the logical OR of two HLT paths, HLT\_DoubleMu33NoFiltersNoVtx\_v\* OR  
 294 HLT\_DoubleMu23NoFiltersNoVtxDisplaced\_v\*, which were designed for this analysis. These  
 295 triggers require at least two L3 muons with  $p_T > 33(23)\text{ GeV}$ , without any of the usual vertex  
 296 constraints and without any maximum requirement on the impact parameter. The second trig-  
 297 ger also requires that the L3 muon  $|d_0|$  is greater than 0.01 cm. The HLT paths are seeded by the  
 298 lowest  $p_T$  threshold double-muon L1 seeds that were unprescaled. Compared to the standard  
 299 dimuon HLT paths, the signal efficiency with these dedicated triggers is significantly higher.

300 In 2017 and 2018 data and corresponding MC simulation, the events are required to pass  
 301 HLT\_DoubleMu43NoFiltersNoVtx\_v\*. The muon  $p_T$  threshold was raised with respect to 2016  
 302 due to increased pileup. Furthermore, no version of this trigger with displaced muons was  
 303 available in 2017 or 2018.

## 304 3.2 Preselection

305 After requiring that the events pass the triggers described above, we preselect good offline  
 306 reconstructed electrons and muons in each channel.

307 In the  $e\mu$  channel, we preselect events with at least one particle flow (PF) electron and at least  
 308 one global PF muon, while in the same-flavor channels, we preselect events with at least two  
 309 PF electrons or muons [42]. We make requirements on these electrons and muons as shown in  
 310 Tables 1, 2 and 3. The offline electron and muon  $p_T$  threshold requirements are chosen in order  
 311 to be in the plateau of the trigger turn-on curve. We require electrons and muons to have  $|\eta| <$   
 312 1.5 in order to remove leptons with poorly measured  $|d_0|$ , which are predominantly at large  $|\eta|$   
 313 (see Appendix C for further discussion). In addition, electrons that traverse the gap between  
 314 the endcap and barrel detectors are rejected as it is known that the electron reconstruction  
 315 performance is not optimal in this region. Thus, in effect, electrons are required to have  $|\eta| <$   
 316 1.44.

317 We use tight cut-based identification (ID) on the electrons and muons to select well-reconstructed  
 318 leptons, but we remove any selections placed on  $d_0$ , the longitudinal impact parameter, or the  
 319 charge product, as these three selections would limit the signal efficiency and make the analy-  
 320 sis more model dependent. For electrons, the ID corresponds to the

Table 1: The  $e\mu$  preselection criteria. The electron and muon  $p_T$  thresholds increased in 2017 because the HLT electron and muon  $p_T$  thresholds increased.

Selection variable	Object-level selections	
	Electron	Muon
Number	$\geq 1$	$\geq 1$
$p_T$	$> 42 \text{ GeV}$ (2016) $> 45 \text{ GeV}$ (2017 and 2018)	$> 40 \text{ GeV}$ (2016) $> 45 \text{ GeV}$ (2017 and 2018)
$ \eta $	$< 1.44$	$< 1.5$
$\eta - \phi$ (pixel power supply issue)	veto ( $1.0 < \eta < 1.5$ and $\phi > 2.7$ ) (2017) veto ( $0.3 < \eta < 1.2$ and $0.4 < \phi < 0.8$ ) (2018)	veto ( $1.0 < \eta < 1.5$ and $\phi > 2.7$ ) (2017) veto ( $0.3 < \eta < 1.2$ and $0.4 < \phi < 0.8$ ) (2018)
ID	Tight (cut-based)	Tight (cut-based)
Custom isolation	Tight	Tight
Event-level selections		
	Zero $\mu\mu$ pairs with $\cos\alpha < -0.99$ Reject $\Delta t < -20 \text{ ns}$ , if both timing ndof $> 7$ At least one $e\mu$ pair with $\Delta R(e, \mu) > 0.2$	
	Reject events where the candidate electron and muon form a good displaced vertex that overlaps with the tracker material	

Table 2: The ee preselection criteria. The electron  $p_T$  threshold increased in 2017 and 2018 because the HLT electron  $p_T$  threshold increased at the start of 2017.

Selection variable	Object-level selections
Number	$\geq 2$
$p_T$	$> 65 \text{ GeV (2016)}$ $> 75 \text{ GeV (2017 and 2018)}$
$ \eta $	$< 1.44$
$\eta - \phi$ (pixel power supply issue)	veto ( $1.0 \leq \eta < 1.5$ and $\phi > 2.7$ ) (2017) veto ( $0.3 \leq \eta < 1.2$ and $0.4 < \phi < 0.8$ ) (2018)
ID	Tight (cut-based)
Custom isolation	Tight
	Event-level selections
	At least one ee pair with $\Delta R(e, e) > 0.2$
	Reject events where the candidate electrons form a good displaced vertex that overlaps with the tracker material Reject events with displaced muons in the ee channel inclusive signal region

Table 3: The  $\mu\mu$  preselection criteria. The muon  $p_T$  threshold increased in 2017 and 2018 because the HLT muon  $p_T$  threshold increased at the start of 2017.

Selection variable	Object-level selections	
	Muon	$\geq 2$
Number		
$p_T$		
$ \eta $		
$\eta - \phi$ (pixel power supply issue)		veto ( $1.0 < \eta < 1.5$ and $\phi > 2.7$ ) (2017) veto ( $0.3 < \eta < 1.2$ and $0.4 < \phi < 0.8$ ) (2018)
ID	Tight (cut-based)	
Custom isolation	Tight	
Event-level selections		
Zero $\mu\mu$ Pairs with $\cos\alpha < -0.99$		
Reject $\Delta t < -20\text{ ns}$ , if both timing ndof $> 7$		
At least one $\mu\mu$ pair with $\Delta R(\mu, \mu) > 0.2$		
Reject events where the candidate muons form a good displaced vertex that overlaps with the tracker material		
Reject events with displaced electrons in the $e\mu$ channel inclusive signal region		

Table 4: The electron tight ID requirements. They are identical to the tight cut-based ID from the egamma POG, except we remove any requirements on  $d_0$  or the longitudinal impact parameter [43]. Only the electron ID in the barrel is relevant, given our electron  $|\eta|$  cut and ECAL gap veto.

Electron ID requirements	
full5x5 $\sigma I\eta I\eta <$	0.0104 (2018, 2017) 0.00998 (2016)
$ \delta\eta_{\text{Seed}}  <$	0.00255 (2018) 0.00353 (2017) 0.00308 (2016)
$ \delta\phi_{\text{In}}  <$	0.022 (2018) 0.0499 (2017) 0.0816 (2016)
H/E <	0.026 + 1.15/E + 0.0324 $\rho$ /E (2018) 0.026 + 1.12/E + 0.0368 $\rho$ /E (2017) 0.0414 (2016)
Rel. comb. PF iso with EA corr <	0.0287 + 0.506/ $p_T$ (2018) 0.0361 (2017) 0.0588 (2016)
$ 1/E - 1/p  <$	0.159 (2018) 0.0278 (2017) 0.0129 (2016)
expected missing inner hits $\leq$	1
pass conversion veto	yes

<sup>321</sup> egmGsfElectronIDs:cutBasedElectronID-Summer16-80X-V1-tight,

<sup>322</sup> egmGsfElectronIDs:cutBasedElectronID-Fall17-94X-V1-tight, and

<sup>323</sup> egmGsfElectronIDs:cutBasedElectronID-Fall17-94X-V2-tight

<sup>324</sup> versioned ID for 2016, 2017, and 2018 samples, respectively [43]. The electron and muon tight

<sup>325</sup> ID requirements we use are summarized in Tables 4 and 5. The  $p_T$  and  $\eta$  distributions are  
shown for selected top squark signal points in Fig. 3, for events that pass the  $e\mu$  preselection.

<sup>326</sup> Kinematic distributions of the the GMSB slepton and the exotic Higgs boson signals are shown  
in Appendices D.1 and D.2, respectively.

<sup>327</sup>

<sup>328</sup>

Table 5: The muon tight ID requirements. They are identical to the tight cut-based ID from the muon POG, except we remove any requirements on  $d_0$  or the longitudinal impact parameter [44].

Muon ID requirements
Is a global muon
Is a PF muon
$\chi^2/n_{\text{dof}}$ of the global-muon track fit is $< 10$
At least one muon-chamber hit included in the global-muon track fit
Muon segments in at least two muon stations
At least 1 valid pixel hit
At least 6 tracker layers with hits

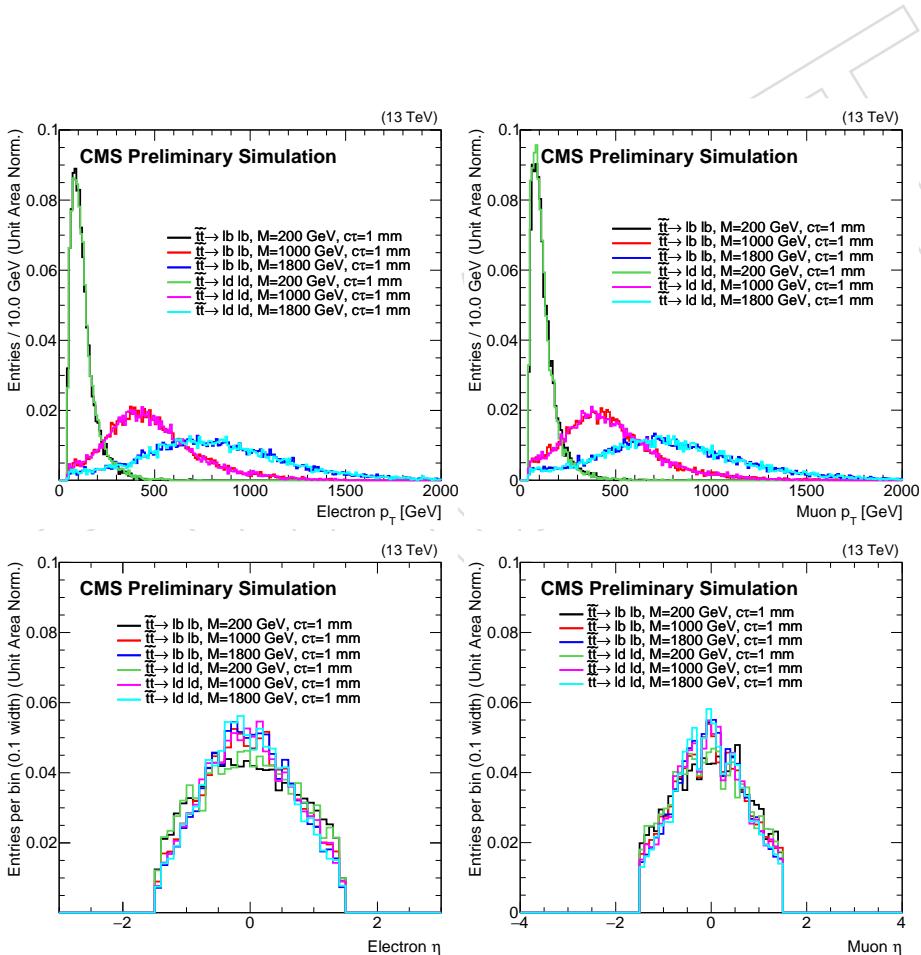


Figure 3: The electron (left) and muon (right)  $p_T$  (upper) and  $\eta$  (lower) distributions in 2018 top squark signal for events that pass the  $e\mu$  preselection. The distributions are shown for both decay processes, one choice of  $\tilde{t}$  proper decay length, and 3 sample  $\tilde{t}$  masses. Each histogram is normalized to unity.

We also require tight PF isolation on the electrons and muons. However, we use a modified isolation definition that accounts for the fact that displaced leptons may be associated with the wrong primary vertex. The standard PF isolation assumes all energy from primary vertices other than the leading primary vertex is due to pileup, which is not true when the primary vertex ordering is altered by an incorrectly associated lepton. We have therefore modified the pileup correction to be agnostic to the primary vertex ordering by allowing PF candidates from any primary vertex to contribute to the isolation sum and by using a simple  $\rho$ -based pileup correction, where  $\rho$  is the total transverse energy of all the PF candidates divided by the total detector area. We use the modified isolation definition for both electrons and muons while keeping the original tight working point for electrons and shifting the muon tight working point from less than 0.15 to less than 0.10. Thus, we require that the relative isolation is  $< 0.10$  for muons,  $< 0.0588$  (0.0571) for electrons in the barrel (endcap) in 2016, and  $< 0.0287 + 0.506 / p_{T,e}$  ( $0.0445 + 0.963 / p_{T,e}$ ) for electrons in the barrel (endcap) in 2017 and 2018, where  $p_{T,e}$  is the electron  $p_T$ . This modified PF isolation rejects more background when the leptons are displaced without significantly altering the signal yield. We note, however, that there may still be some minor dependence on the primary vertex selection in the PF muon requirement because the PF muon selection includes some loose isolation requirements where the charged hadron component is constrained to the selected primary vertex. See Appendix D.3 for performance plots with the modified isolation.

We also reject electrons and muons in certain regions of the  $\eta - \phi$  plane that contain relatively more leptons with large  $|d_0|$  due to mismeasurements. We identify these regions as highly populated bins in the electron  $\eta - \phi$  distribution in a prompt-muon, displaced-electron control region in 2017 and 2018 data (see Fig. 4). No such bins are present in 2016 data. The identified regions coincide with regions found by a previous CMS analysis [4] to have pixel power supply issues in 2017 and 2018. The  $\eta - \phi$  variation is more apparent for displaced electrons than displaced muons, so we use data in a prompt muon ( $|d_0| < 40 \mu\text{m}$ ), displaced electron ( $100 < |d_0| < 500 \mu\text{m}$ ) control region to define the rejected regions used for both electrons and muons. In 2017, the rejected region is  $1.0 < \eta < 1.5$ ,  $\phi > 2.7$ , and in 2018 the rejected region is  $0.3 < \eta < 1.2$ ,  $0.4 < \phi < 0.8$ .

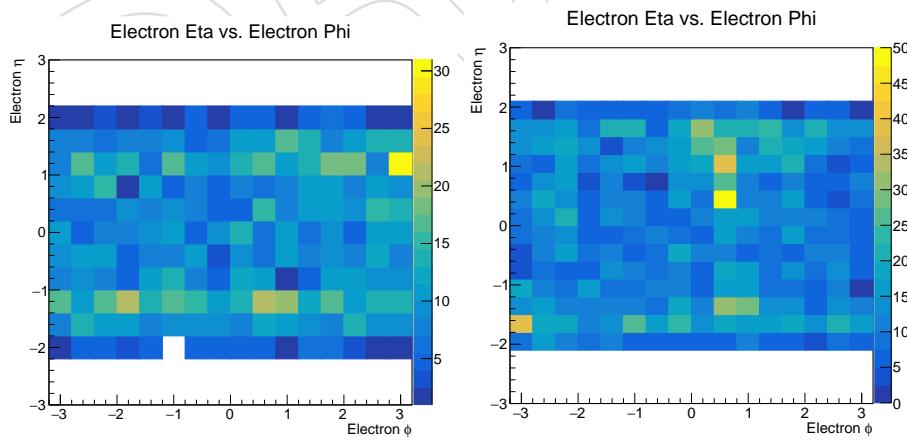


Figure 4: The electron  $\eta - \phi$  distribution in a prompt-muon, displaced-electron control region in 2017 (left) and 2018 (right) data.

Besides these object-level selections, we also impose just a few event-level selections. To remove cosmic-ray muons in the  $\mu\mu$  and  $e\mu$  channels, we require that there are zero pairs of muons with  $\cos \alpha < -0.99$ , where  $\alpha$  is the 3D angle between the muons. In addition, we require that the relative time between the leading and subleading muons is not consistent with

the timing of cosmic-ray muons, in the  $\mu\mu$  and  $e\mu$  channels. To do this, we look at the muon time as measured at the IP from the DTs and CSCs, assuming the muons are outgoing. We then determine which muon is above the other based on their  $\phi$  measurements, and find  $\Delta t$ , the time of the lower muon subtracted from the time of the upper muon. We reject events with  $\Delta t < -20$  ns, if the number of degrees of freedom of the timing measurements for both muons is greater than seven. See Section 7.2 for a study where we invert the cosmic-ray muon rejection criteria. We also require the relevant leptons in each channel are separated: a loose selection of  $\Delta R > 0.2$  is sufficient to significantly reduce the contribution from heavy flavor backgrounds. In addition, we reject events where the candidate leptons form a good displaced vertex that overlaps with the tracker material. The vertices are reconstructed with the Kalman Vertex Fitter, and a “good” vertex is one with a  $\chi^2/n_{\text{dof}} < 20$  and a resolution of less than  $10 \mu\text{m}$ . The tracker material was obtained from the tracker material budget measurements [45, 46]. See Section 7.1 for a study of inverting the material interactions criterion.

To remove overlaps between the three channels, we reject events that would fall into the  $e\mu$  inclusive signal region in the same-flavor channels. That is, in the  $ee$  ( $\mu\mu$ ) channel, we reject events with at least one muon (electron) that passes the  $e\mu$  channel preselection and has  $|d_0| > 100 \mu\text{m}$ .

If there are more than the minimum number of electrons or muons in an event that pass the preselection, the highest  $p_T$  leptons are used. In contrast, in the 2015 analysis [16], exactly one electron and exactly one muon were required. Also in the 2015 analysis, the leptons were required to be oppositely charged, while we have removed the opposite-sign requirement for this iteration of the analysis. We made these changes at the request of several theorists, including the authors of Ref. [25].

The events passing each channel preselection criteria are further categorized by using the 2D track impact parameter of the selected leptons in the plane transverse to the proton beams. This transverse impact parameter ( $|d_0|$ ) is defined as the distance of closest approach in the transverse plane of the helical trajectory of the lepton track to the center of the luminous region<sup>1</sup>. Figure 5 shows a diagram of the definition of  $d_0$  used in this analysis. The lepton  $d_0$  is strongly correlated with the lifetime of the particle from which it decayed. The  $d_0$  spectra of leptons from promptly decaying particles, such as top quarks and W bosons, will peak at zero and fall sharply. For leptons from particles that travel farther than the resolution of lepton  $|d_0|$ , such as tau leptons and mesons containing bottom or charm quarks, the  $d_0$  spectrum will be noticeably wider. For leptons from much longer-lived particles such as the top squarks in the benchmark signal process, the  $d_0$  spectrum will be very wide, approaching a flat distribution for the longest top squark lifetimes we consider. The electron and muon  $|d_0|$  distributions are shown for top squark signals at different lifetimes and backgrounds in Fig. 6, for events that pass the  $e\mu$  preselection criteria. For a given lifetime, the  $|d_0|$  distributions are the same for the two  $\tilde{t}$  decay modes considered and are independent of  $\tilde{t}$  mass.

Figures 7 and 8 show the cumulative number of events that pass the  $e\mu$ ,  $ee$ , and  $\mu\mu$  preselection criteria for signal events with different  $\tilde{t}$  proper decay lengths, masses, and both decay modes considered. In general, the smallest proper decay length samples are the most efficient. In the  $e\mu$  and  $\mu\mu$  channels, slightly more  $\tilde{t} \rightarrow l b$  events than  $\tilde{t} \rightarrow l d$  events pass the trigger, due to the presence of soft muons from the b quark decay. However, once we require that the muons are

<sup>1</sup>Another option of a discriminating variable could be the  $d_0$  significance  $|d_0|/\sigma$ , where  $\sigma$  is the uncertainty in  $|d_0|$ . However, we choose to simply use  $|d_0|$  in this analysis because of its straightforward correspondence to the particle lifetime, which makes it an easier variable for theorists to use in their efforts to reinterpret this search. Furthermore, we found that the uncertainty in  $|d_0|$  is often underestimated, which reduces any potential benefit of using  $|d_0|/\sigma$ .

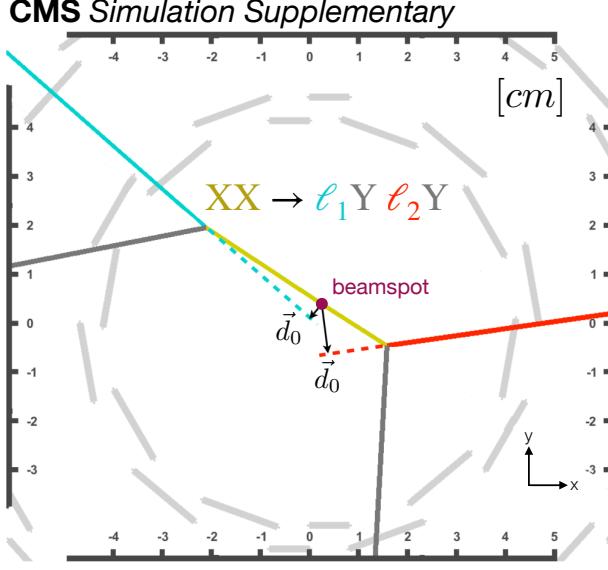


Figure 5: Diagram of a generic signal event, from a transverse view of the interaction point. The black arrows show the lepton  $d_0$ s.

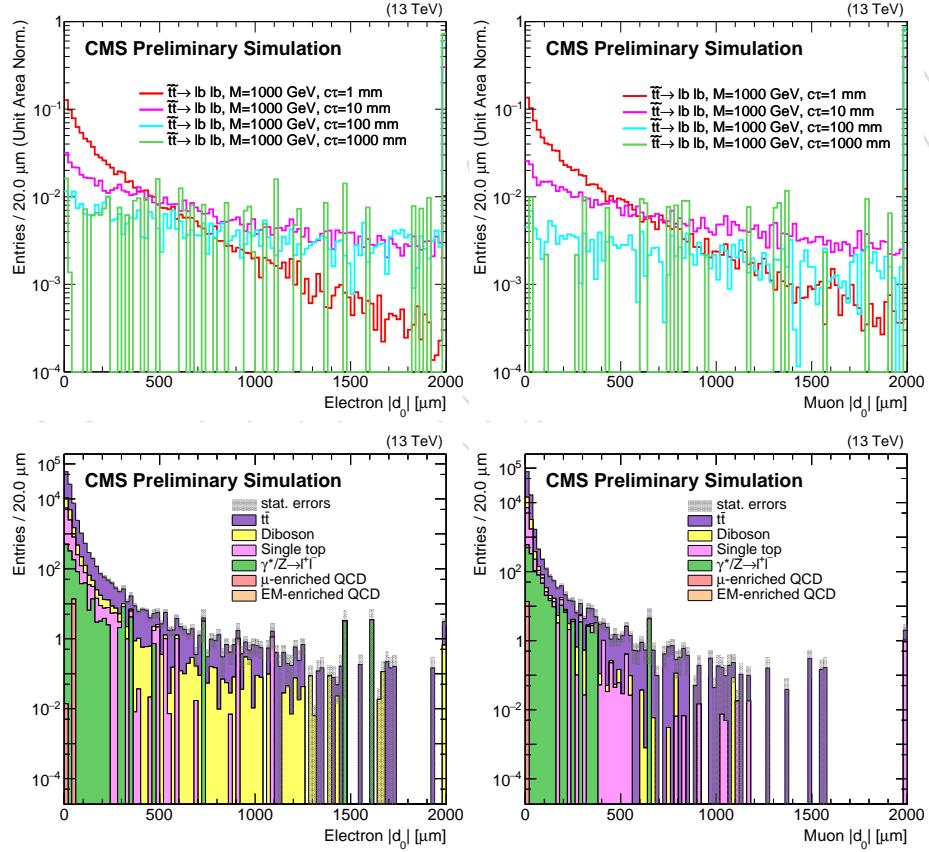


Figure 6: The electron (left) and muon (right)  $|d_0|$  distributions for 2018 simulation events that pass the  $e\mu$  preselection criteria. The upper two plots show  $\tilde{t} \rightarrow b \bar{b}$  simulation for a single  $\tilde{t}$  mass and 4 different proper decay lengths, where each histogram is normalized to unity. The lower two plots show the background MC simulation, normalized to the integrated luminosity. In all of the histograms, the last bin includes the overflow. All of the corrections from Section 4 are applied.

405 isolated, the efficiency of the two decay modes are the same. The signal efficiency per channel  
 406 is consistent among the 3 years. Cutflows of the the GMSB slepton and the exotic Higgs boson  
 407 signals are shown in Appendices D.1 and D.2, respectively.

### 408 3.3 Prompt control region

409 In order to check the data to simulation agreement, we define a prompt control region. Events  
 410 in each channel prompt control region are selected by requiring that they pass all of the pres-  
 411 election criteria described above, and then additionally requiring the selected leptons to have  
 412  $|d_0| < 50 \mu\text{m}$ . The prompt control region is dominated by promptly decaying SM events, and  
 413 it is depleted of events with signal or leptons from heavy flavor QCD decays. We define this  
 414 region in each channel in order to verify the agreement between MC simulation and data, after  
 415 the corrections described in Section 4 are applied. Figures 9, 10, and 11 show the  $p_T$ ,  $\eta$ , and  $|d_0|$   
 416 distributions of the leptons in the  $e\mu$ ,  $ee$ , and  $\mu\mu$  prompt control regions, respectively, for 2016  
 417 data and MC simulation. Similar figures for 2017 and 2018 are in Appendix D.4. Distributions  
 418 of the dilepton pairs in each channel and year can also be found in Appendix D.4. All plots  
 419 shown in this section and in Appendix D.4 are after the corrections described in Section 4 are  
 420 applied.

### 421 3.4 Inclusive signal search region

422 For events to enter the inclusive signal region, they must pass the preselection criteria and  
 423 have lepton  $100 \mu\text{m} < |d_0| < 10 \text{ cm}$ . The signal regions will be defined in Section 6.2. We do  
 424 not select leptons more displaced than 10 cm because the tracking efficiency drops sharply after  
 425 this point, as shown in Section 8.3. Furthermore, this requirement ensures that the leptons are  
 426 well within the pixel volume, which is necessary since the tight lepton IDs that we use require  
 427 pixel hits. Tables 6, 7, and 8 show the cumulative efficiency for  $\tilde{t}\tilde{t} \rightarrow b\bar{b}$ ,  $pp \rightarrow \tilde{l}\bar{\tilde{l}} \rightarrow \tilde{l}\tilde{G} l\tilde{G}$ , and  
 428  $pp \rightarrow H \rightarrow SS \rightarrow l\bar{l} l\bar{l}$  events, respectively, to pass the full 2018 inclusive signal region selection  
 429 for several signal points.

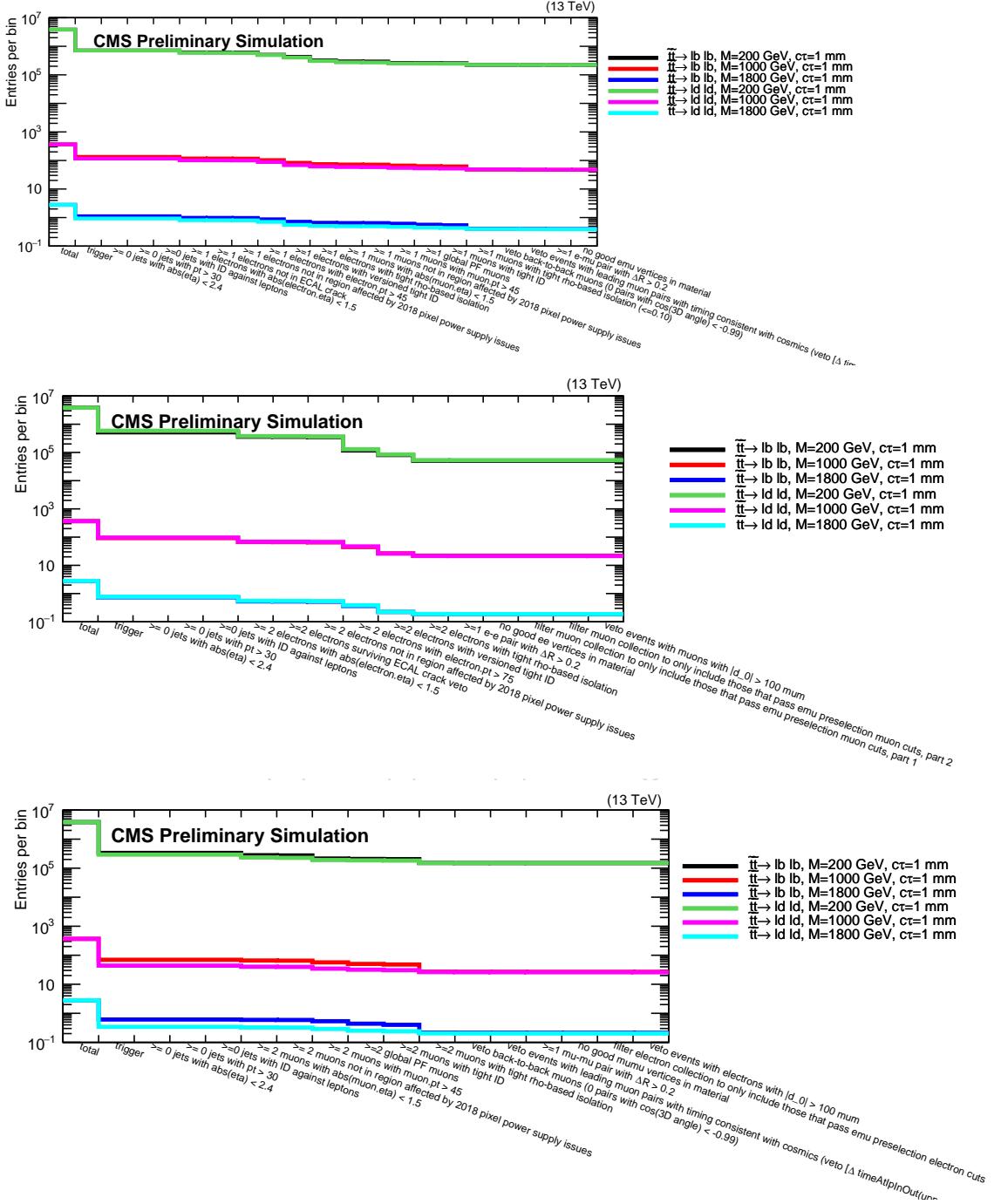


Figure 7: The cumulative number of events that pass each criterion in the  $e\mu$  (upper),  $ee$  (middle), and  $\mu\mu$  (lower) preselection, using 2018 top squark signal simulation. Several  $\tilde{t}$  masses and both decay modes are shown. The jet criteria do not exclude any events and are simply an artifact of the analysis framework.

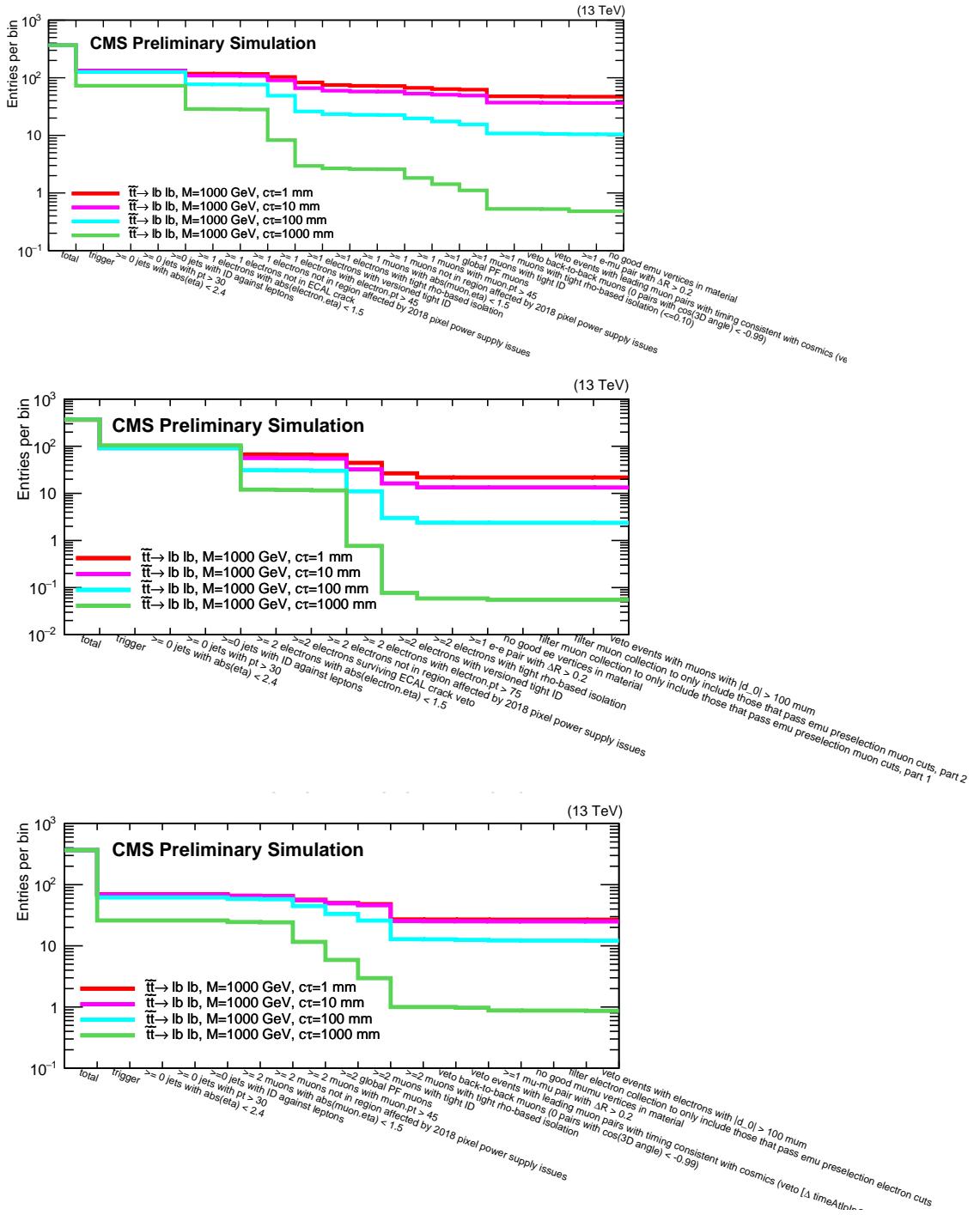


Figure 8: The cumulative number of events that pass each criterion in the  $e\mu$  (upper),  $ee$  (middle), and  $\mu\mu$  (lower) preselection, using 2018 top squark signal simulation. Several  $\tilde{t}$  proper decay lengths are shown. The jet criteria do not exclude any events and are simply an artifact of the analysis framework.

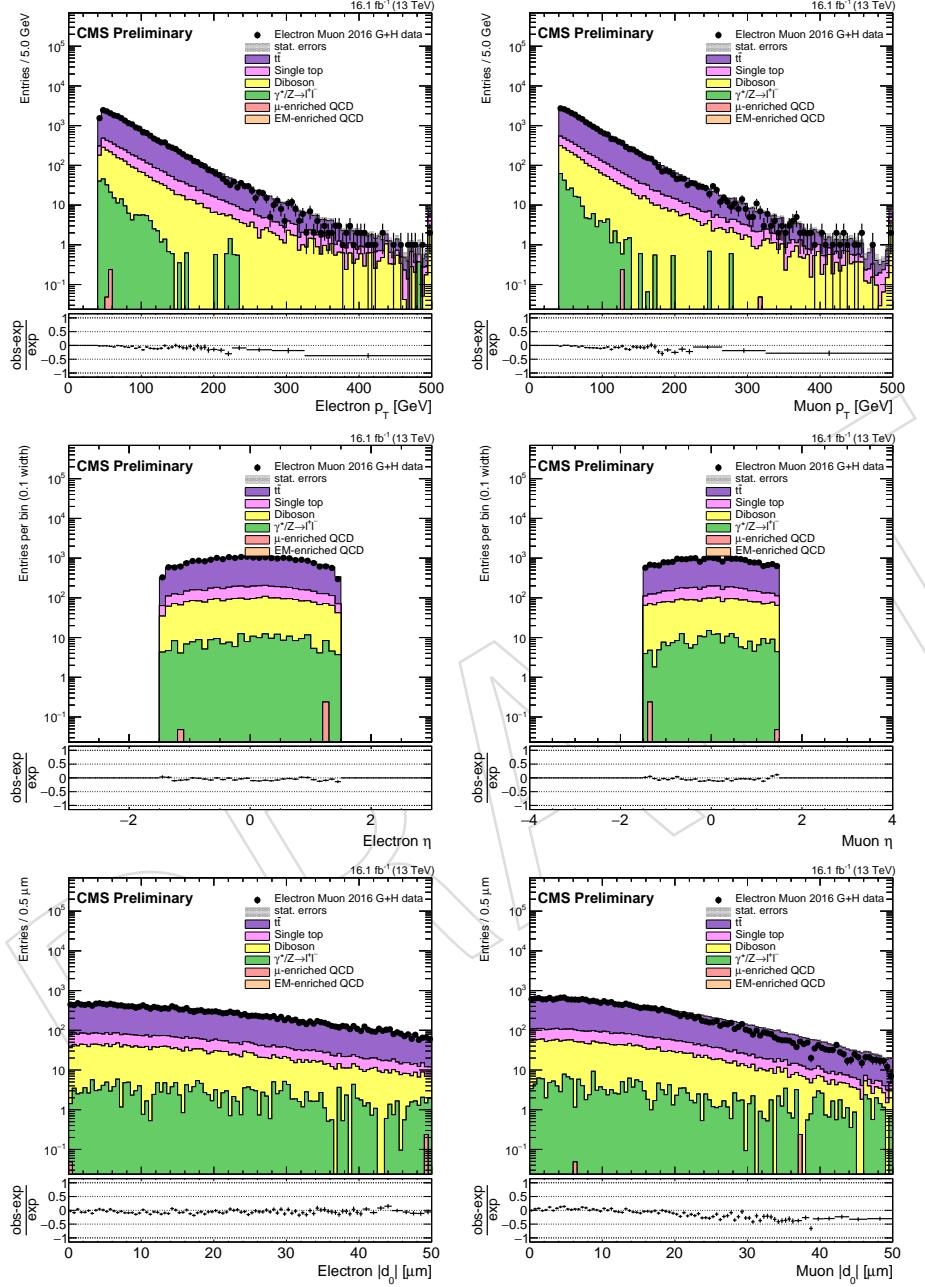


Figure 9: The lepton  $p_T$  (upper),  $\eta$  (middle), and  $|d_0|$  (lower) distributions in the  $e\mu$  prompt control region, for electrons (left) and muons (right), for 2016 data and MC simulation. The rightmost bin in each plot contains the overflow entries.

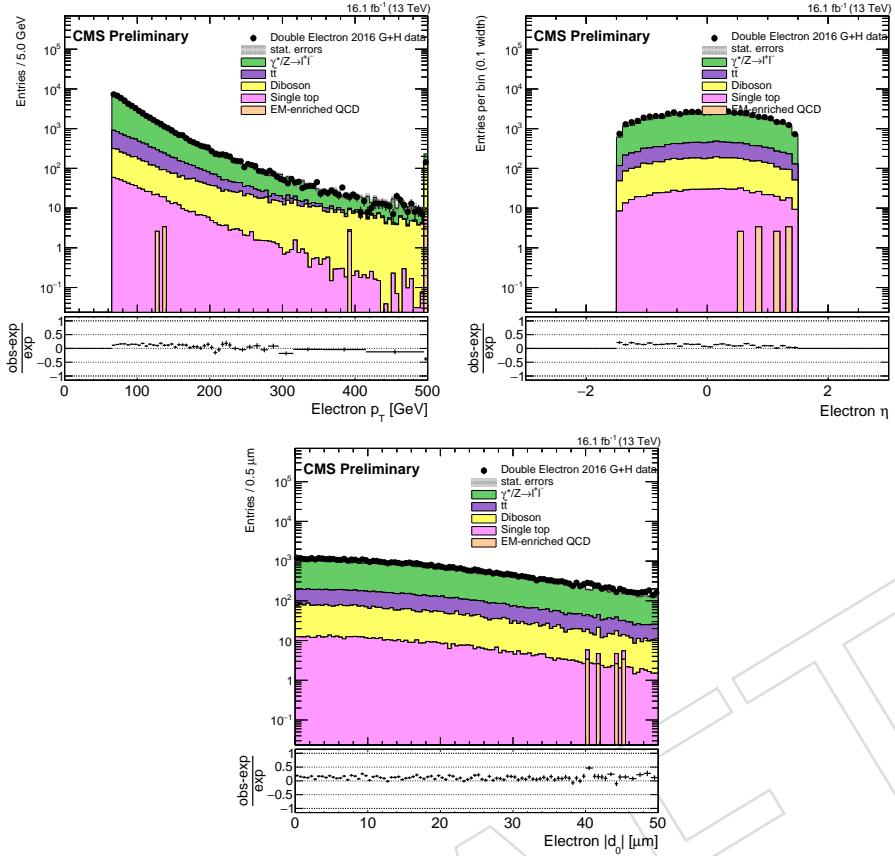


Figure 10: The electron  $p_T$  (upper),  $\eta$  (middle), and  $|d_0|$  (lower) distributions in the ee prompt control region for 2016 data and MC simulation. The rightmost bin in each plot contains the overflow entries.

Table 6: The cumulative efficiency for  $\tilde{t} \rightarrow b\ell$  signal events to pass the 2018 inclusive signal region selection, for several choices of  $\tilde{t}$  mass and  $c\tau$ . The corrections described in Section 4 are applied.

e $\mu$ inclusive signal region			
	200 GeV	1000 GeV	1800 GeV
0.1 cm	2.1%	4.6%	4.6%
1 cm	3.5%	7.9%	8.8%
10 cm	1.0%	2.7%	3.4%
100 cm	0.047%	0.13%	0.16%
ee inclusive signal region			
	200 GeV	1000 GeV	1800 GeV
0.1 cm	0.47%	2.1%	2.1%
1 cm	0.60%	2.8%	3.3%
10 cm	0.11%	0.59%	0.76%
100 cm	0.003%	0.014%	0.022%
$\mu\mu$ inclusive signal region			
	200 GeV	1000 GeV	1800 GeV
0.1 cm	1.4%	2.6%	2.5%
1 cm	3.1%	5.6%	5.9%
10 cm	1.5%	3.2%	3.7%
100 cm	0.11%	0.23%	0.34%

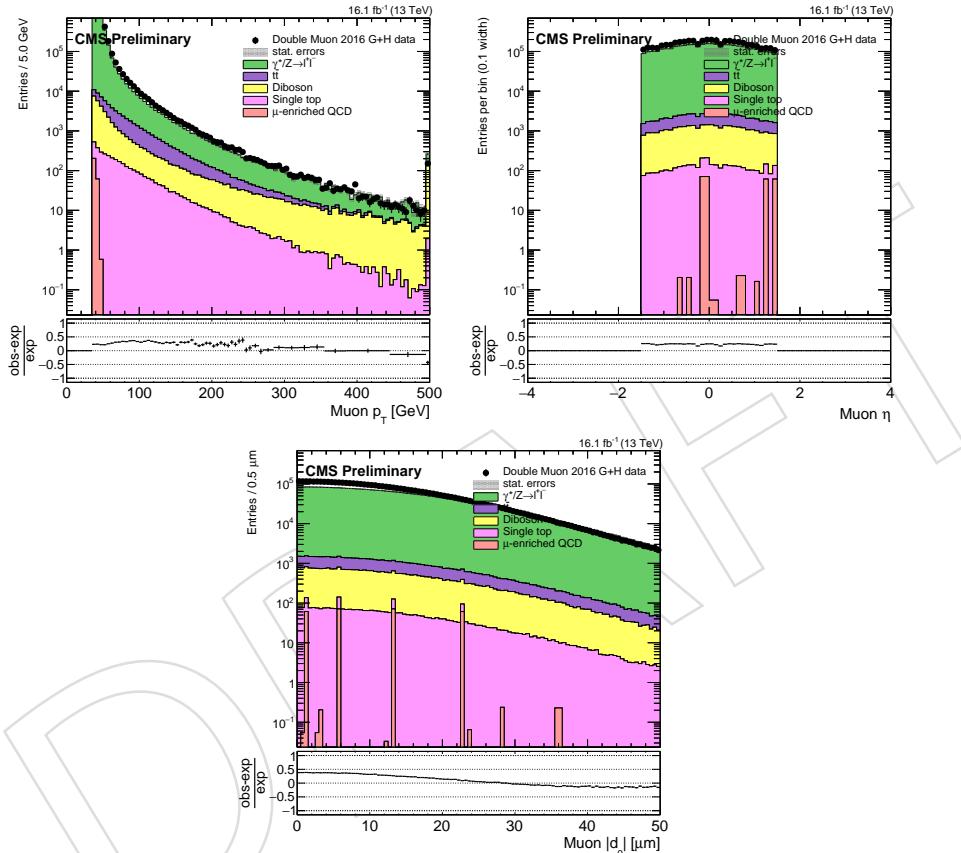


Figure 11: The muon  $p_T$  (upper),  $\eta$  (middle), and  $|d_0|$  (lower) distributions in the  $\mu\mu$  prompt control region for 2016 data and MC simulation. The rightmost bin in each plot contains the overflow entries.

Table 7: The cumulative efficiency for  $\text{pp} \rightarrow \tilde{l}\tilde{l} \rightarrow \tilde{l}\tilde{G} l\tilde{G}$  signal events to pass the 2018 inclusive signal region selection, for several choices of  $\tilde{l}$  mass and  $c\tau$ . The corrections described in Section 4 are applied.

ee inclusive signal region, $\tilde{e}$ NLSP			
	100 GeV	500 GeV	900 GeV
0.1 cm	1.8%	16%	20%
1 cm	1.8%	17%	21%
10 cm	0.13%	1.8%	2.8%
100 cm	0.002%	0.02%	0.08%
$\mu\mu$ inclusive signal region, $\tilde{\mu}$ NLSP			
	100 GeV	500 GeV	900 GeV
0.1 cm	6.4%	19%	21%
1 cm	11%	33%	35%
10 cm	4.1%	13%	16%
100 cm	0.12%	0.48%	0.66%
e $\mu$ inclusive signal region, $\tilde{\tau}$ NLSP			
	100 GeV	200 GeV	400 GeV
0.1 cm	0.96%	1.8%	4.3%
1 cm	0.67%	2.3%	5.7%
10 cm	0.14%	0.36%	1.1%
100 cm	0.006%	0%	0.02%
ee inclusive signal region, $\tilde{\tau}$ NLSP			
	100 GeV	200 GeV	400 GeV
0.1 cm	0.07%	0.20%	0.90%
1 cm	0.10%	0.30%	1.2%
10 cm	0.02%	0.04%	0.11%
100 cm	0%	0%	0%
$\mu\mu$ inclusive signal region, $\tilde{\tau}$ NLSP			
	100 GeV	200 GeV	400 GeV
0.1 cm	0.55%	1.2%	2.0%
1 cm	0.99%	1.9%	3.6%
10 cm	0.20%	0.75%	1.7%
100 cm	0%	0%	0.07%

Table 8: The cumulative efficiency for  $\text{pp} \rightarrow H \rightarrow SS \rightarrow l\bar{l}l\bar{l}$  signal events to pass the 2018 inclusive signal region selection, for a Higgs mass of 125 GeV and several choices of S mass and  $c\tau$ . The corrections described in Section 4 are applied.

$e\mu$ inclusive signal region		
	30 GeV	50 GeV
0.1 cm	0.33%	0.27%
1 cm	0.58%	0.42%
10 cm	0.07%	0.07%
100 cm	0.004%	0.005%
ee inclusive signal region		
	30 GeV	50 GeV
0.1 cm	0.04%	0.03%
1 cm	0.03%	0.02%
10 cm	0.003%	0%
100 cm	0%	0%
$\mu\mu$ inclusive signal region		
	30 GeV	50 GeV
0.1 cm	0.50%	0.35%
1 cm	0.83%	0.55%
10 cm	0.26%	0.22%
100 cm	0.01%	0.01%

## 4 Corrections to the simulation

Several corrections are applied to the MC simulations, in order to account for known differences between simulation and data.

### 4.1 Event pileup reweighting

The MC simulation is corrected so that its distribution of pileup interactions matches that of 2016, 2017, and 2018 data. The pileup distribution in 2016, 2017, and 2018 data is obtained using the standard `pileupCalc.py` script, the JSON files mentioned above, and the pileup JSON files

[https://cms-service-dqm.web.cern.ch/cms-service-dqm/CAF/certification/Collisions16/13TeV/PileUp/pileup\\_latest.txt](https://cms-service-dqm.web.cern.ch/cms-service-dqm/CAF/certification/Collisions16/13TeV/PileUp/pileup_latest.txt),  
[https://cms-service-dqm.web.cern.ch/cms-service-dqm/CAF/certification/Collisions17/13TeV/PileUp/pileup\\_latest.txt](https://cms-service-dqm.web.cern.ch/cms-service-dqm/CAF/certification/Collisions17/13TeV/PileUp/pileup_latest.txt), and  
[https://cms-service-dqm.web.cern.ch/cms-service-dqm/CAF/certification/Collisions18/13TeV/PileUp/pileup\\_latest.txt](https://cms-service-dqm.web.cern.ch/cms-service-dqm/CAF/certification/Collisions18/13TeV/PileUp/pileup_latest.txt).

We use the recommended minimum bias cross section of 69.2 mb with an uncertainty of 5%. Each simulated sample is reweighted, event-by-event, by the pileup distribution in data divided by the pileup distribution in that simulated sample.

### 4.2 Lepton ID corrections

The standard Physics Object Group (POG) tight lepton scale factors are applied to the MC simulation in order to account for differences in the reconstruction and ID of the leptons between data and MC simulation. We also use the electron reco scale factors, but we do not use the electron scale and smearing scale factors, as they are not relevant for the analysis. Although we do not apply the tight ID selections on  $|d_0|$  and  $|d_z|$ , the standard scale factors are still appropriate for our use for a few reasons. First of all, very few of the leptons from Z decays are

Table 9: The average  $\sigma_{align}$  for electrons and muons, for the 2017 and 2018 analyses.

	2017	2018
Electrons	$14.75 \pm 0.36 \mu\text{m}$	$9.18 \pm 0.41 \mu\text{m}$
Muons	$7.57 \pm 0.12 \mu\text{m}$	$8.11 \pm 0.08 \mu\text{m}$

454 rejected with the standard impact parameter selections. Secondly, most of the displaced leptons in the search region come from QCD multijet events, whose contribution is estimated with  
 455 a data-driven method. Therefore, different lepton scale factors do not affect the background  
 456 estimation.  
 457

458 We have a dedicated study for the displaced tracking efficiency, which is described in Sec-  
 459 tion 8.3.

### 460 4.3 Lepton $d_0$ corrections

461 As can be seen in Fig. 12, there was initially poor data to MC simulation agreement in the 2017  
 462 and 2018 muon and electron  $|d_0|$  distributions. We found that the average muon and electron  
 463  $|d_0|$  fluctuated periodically with respect to  $\phi$  in 2017 and 2018 data but not in background  
 464 simulation, as can be seen in Fig. 13. This periodic fluctuation in data but not in simulation  
 465 indicates that the tracker alignment is overly optimistic in simulation, and it drives the data to  
 466 simulation disagreement in the  $|d_0|$  distributions.

467 As a result, we smeared the electron and muon  $d_0$  in 2017 and 2018 background MC simu-  
 468 lation in each channel’s prompt control region to match that of data. To do this, we first fit  
 469 the central regions of the background and the data  $d_0$  distributions with Gaussian functions in  
 470 each channel’s prompt control region, and then we compared the widths of the Gaussian fits.  
 471 The fitted distributions are shown in Figs. 14 and 15 for the  $e\mu$  channel. We assume that the  
 472 width of each Gaussian fit is mostly determined by the  $d_0$  resolution. Furthermore, we assume  
 473 that the discrepancy between the data and background MC simulation  $d_0$  distributions comes  
 474 from an overly optimistic alignment in the simulation, which creates an unrealistically ideal  
 475  $d_0$  resolution. That is, we define  $\sigma_{data}^2 = \sigma_{bgd}^2 + \sigma_{align}^2$ , where  $\sigma_{data}$  is the data Gaussian width,  
 476  $\sigma_{bgd}$  is the current background simulation Gaussian width, and  $\sigma_{align}$  is the additional piece  
 477 that is needed to make the background simulation and data agree in each channel’s prompt  
 478 control region. We find  $\sigma_{data}$  and  $\sigma_{bgd}$  from the fits, and compute  $\sigma_{align}$  (see Table 9). The fit  
 479 results are similar in the  $e\mu$  channel (shown in this section) and the same-flavor channels (see  
 480 Appendix D.5). We average the  $\sigma_{align}$  derived in the ee channel and the  $e\mu$  channel for elec-  
 481 trons, and in the  $\mu\mu$  channel and the  $e\mu$  channel for muons. The average  $\sigma_{align}$  is shown in  
 482 Table 9. We then smear the simulation  $d_0$  values by a Gaussian distribution centered at 0 and  
 483 with a width of the average  $\sigma_{align}$ . The smearing was applied to both background and signal  
 484 MC simulation. The corrected  $|d_0|$  distributions (after this smearing) are shown in Section 3.3  
 485 and Appendix D.4.

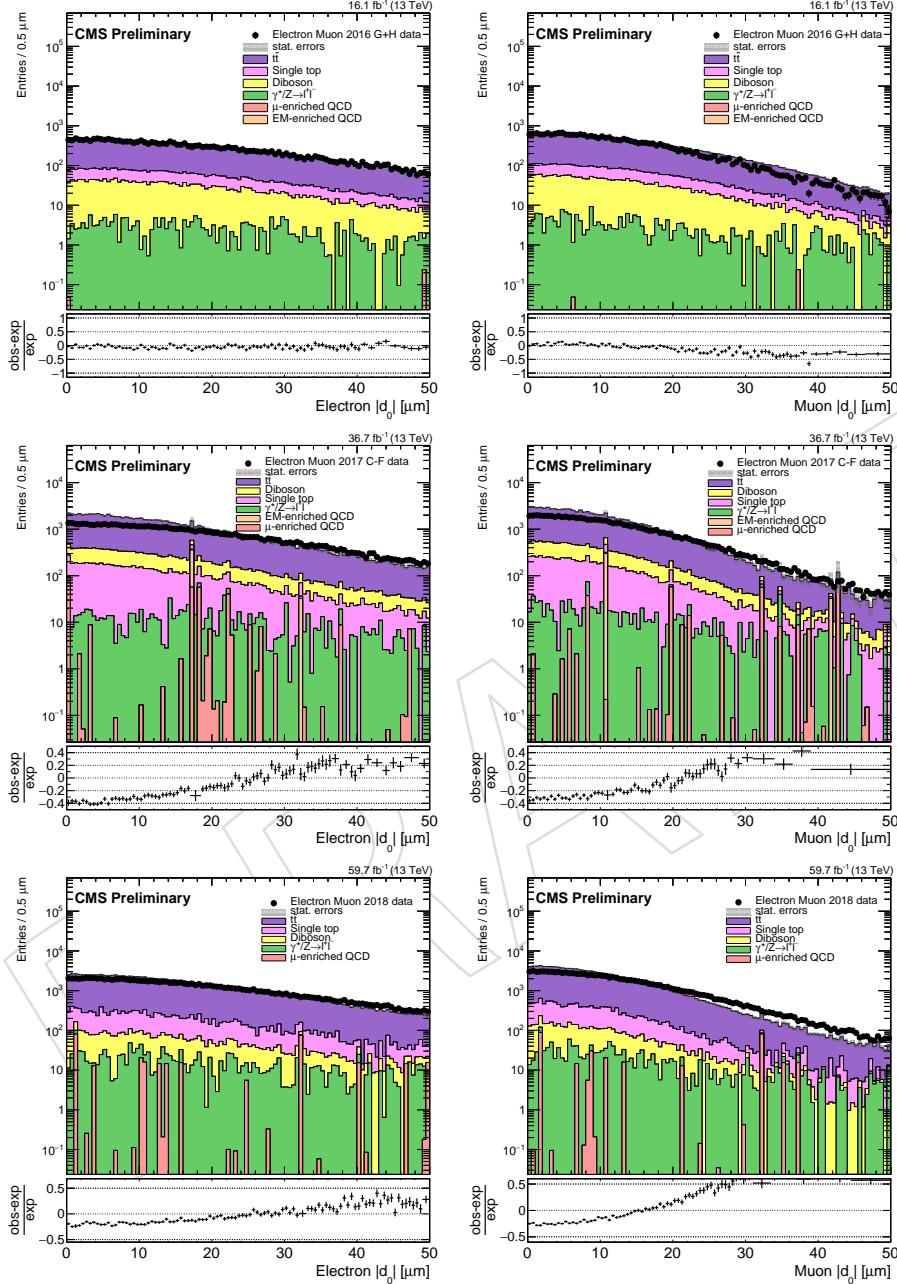


Figure 12: The uncorrected lepton  $|d_0|$  distributions in the  $e\mu$  prompt control region, for electrons (left) and muons (right), for 2016 data and MC simulation (upper), 2017 data and MC simulation (middle), and 2018 data and MC simulation (lower). The rightmost bin in each plot contains the overflow entries.

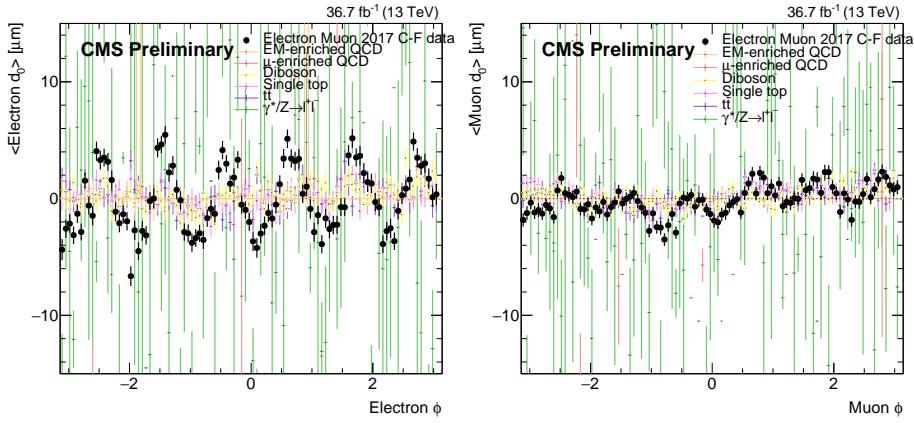


Figure 13: The average lepton  $|d_0|$  as a function of  $\phi$  in the  $e\mu$  prompt control region, for electrons (left) and muons (right), for 2017 data and MC simulation.

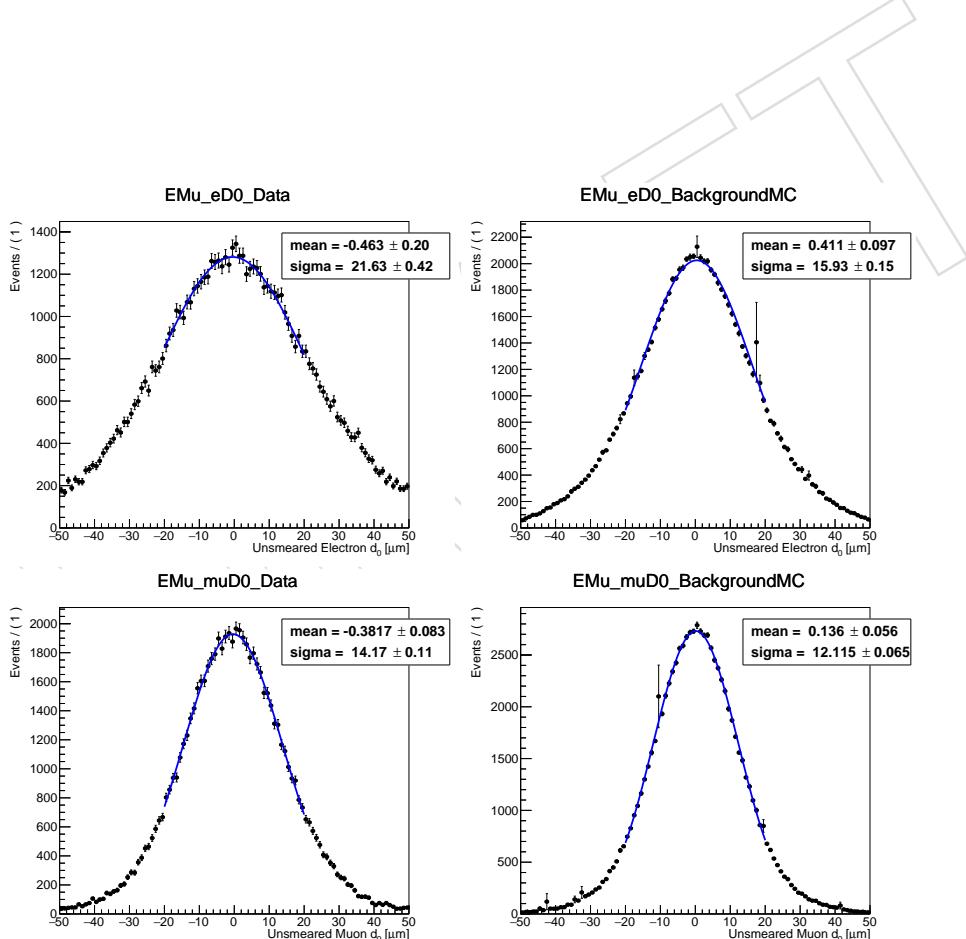


Figure 14: The lepton  $d_0$  distributions with Gaussian fits in data (left) and background MC simulation (right) for electrons (upper) and muons (lower) in the 2017  $e\mu$  prompt control region. The widths of the Gaussian fits are used to determine the width of the Gaussian distribution used to smear the  $d_0$ .

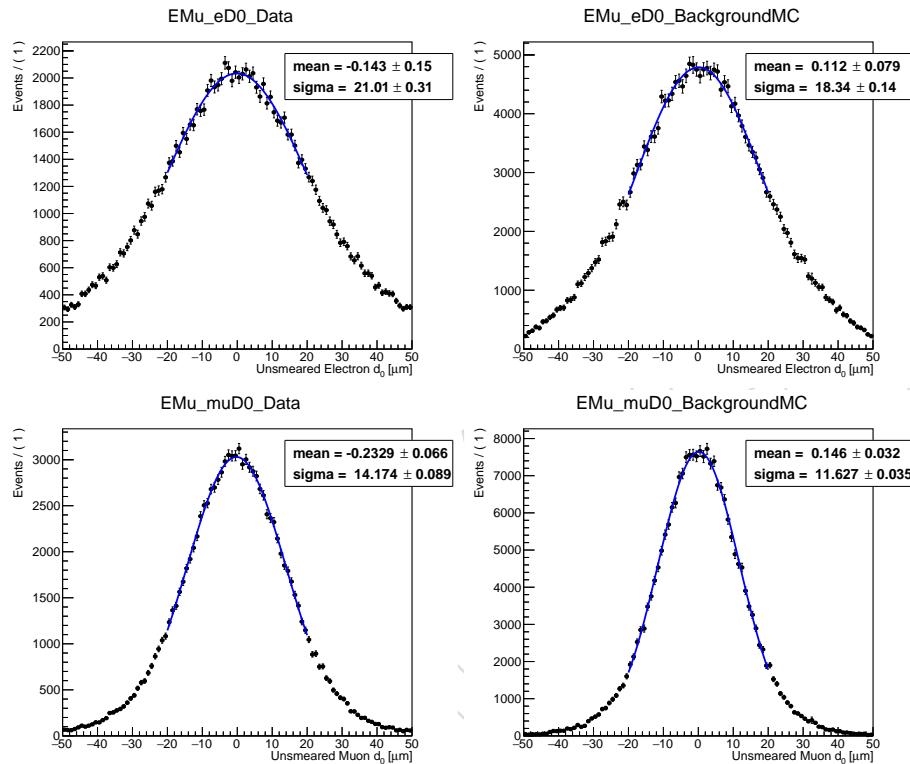


Figure 15: The lepton  $d_0$  distributions with Gaussian fits in data (left) and background MC (right) for electrons (upper) and muons (lower) in the 2018  $e\mu$  prompt control region. The widths of the Gaussian fits are used to determine the width of the Gaussian distribution used to smear the  $d_0$ .

Table 10: The unprescaled MET triggers used to create an orthogonal data sample for the trigger efficiency calculation.

<b>2016</b>
HLT_MET200
HLT_MonoCentralPFJet80_PFMETNoMu110_PFMHTNoMu110_IDTight
HLT_PFMET120_PFMHT120_IDTight
HLT_PFMET170_HBHECleaned
HLT_PFMET300
HLT_PFMETNoMu120_PFMHTNoMu120_IDTight
<b>2017</b>
HLT_CaloMET350_HBHECleaned
HLT_MonoCentralPFJet80_PFMETNoMu120_PFMHTNoMu120_IDTight
HLT_PFMET120_PFMHT120_IDTight
HLT_PFMET250_HBHECleaned
HLT_PFMETNoMu120_PFMHTNoMu120_IDTight
<b>2018</b>
HLT_CaloMET350_HBHECleaned
HLT_MonoCentralPFJet80_PFMETNoMu120_PFMHTNoMu120_IDTight
HLT_PFMET120_PFMHT120_IDTight
HLT_PFMET200_HBHE_BeamHaloCleaned
HLT_PFMET250_HBHECleaned
HLT_PFMETNoMu120_PFMHTNoMu120_IDTight

#### 486 4.4 Trigger efficiency correction

487 We also apply scale factors to the MC simulation to correct for differences in trigger efficiency  
 488 between data and simulation. To measure the trigger efficiency, we first require that events  
 489 pass an OR of several unprescaled MET triggers (see Table 10) and the offline preselection  
 490 criteria, excluding the lepton  $p_T$  criteria, in order to produce a sample of dilepton events that is  
 491 unbiased with respect to the main triggers used in the analysis. In the  $e\mu$  channel, the electron  
 492 (muon)  $p_T$  is required to be greater than 50 GeV when plotting against the muon (electron)  $p_T$   
 493 to disentangle the effect from the other leg of the muon-photon trigger. Events in data from the  
 494 MET primary dataset are used, and the background MC simulation is the  $t\bar{t}$  simulation for the  
 495  $e\mu$  channel and DY simulation for the same-flavor channels. The efficiency is then defined as  
 496 the ratio of events that pass the main triggers in addition to the OR of MET triggers and the  
 497 offline preselection, to events that pass the OR of MET triggers and the offline preselection. The  
 498 scale factor is the ratio of the efficiency in data to the efficiency in simulation.

499 Figures 16 and 17 show the trigger efficiency turn-on curves for data and background simula-  
 500 tion as a function of electron and muon  $p_T$ . The lepton  $p_T$  selections in the preselection criteria  
 501 are chosen to be in the plateaus of these turn-on curves, which allows us to use the plateau  
 502 efficiency values to determine the scale factors.

503 The trigger efficiency scale factors are shown in Table 11. The uncertainty on these scale factors  
 504 is used as a systematic uncertainty (see Section 8). An additional check of the trigger efficiency  
 505 at large  $|d_0|$ , and subsequent additional systematic uncertainty, is shown in Section 8.4.

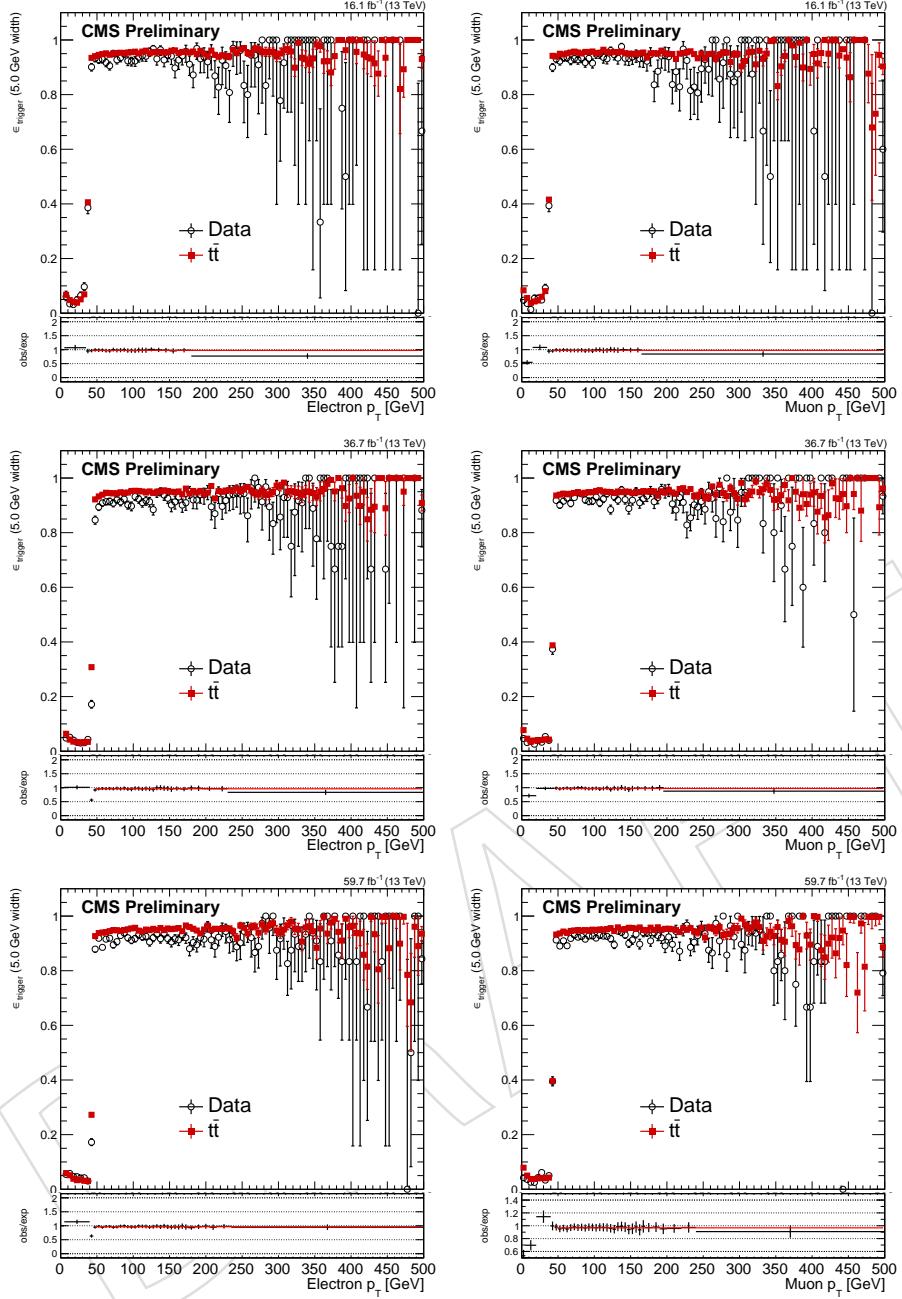


Figure 16: Trigger efficiency as a function of leading electron (left) or leading muon (right)  $p_T$  for the  $e\mu$  channel. The efficiency for the 2016, 2017, and 2018 analyses is shown in the top, middle, and bottom rows, respectively.

Table 11: The trigger efficiency scale factors in each channel and year. These scale factors are applied to each lepton.

	2016	2017	2018
$e\mu$ channel, electrons	$0.974 \pm 0.016$	$0.961 \pm 0.013$	$0.965 \pm 0.012$
$e\mu$ channel, muons	$0.975 \pm 0.016$	$0.972 \pm 0.014$	$0.967 \pm 0.012$
ee channel	$1.000 \pm 0.099$	$1.000 \pm 0.131$	$1.000 \pm 0.185$
$\mu\mu$ channel	$0.956 \pm 0.012$	$0.930 \pm 0.010$	$0.935 \pm 0.011$

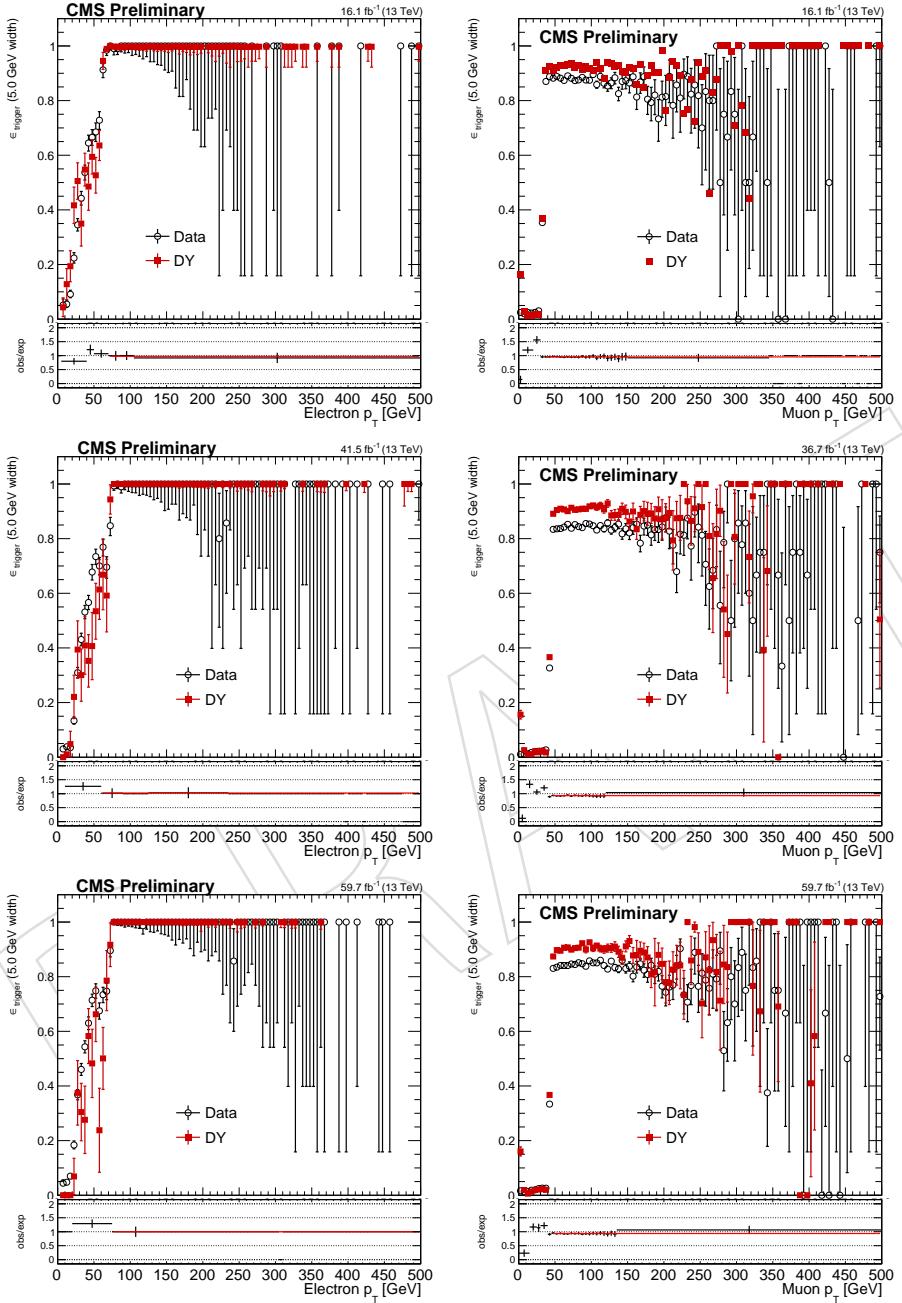


Figure 17: Trigger efficiency as a function of subleading lepton  $p_T$ , for the ee channel (left) and the  $\mu\mu$  channel (right). The upper, middle, and lower rows show the curves for 2016, 2017, and 2018, respectively.

## 5 Background sources

Most leptons resulting from SM processes are from particles that decay promptly. However, displaced leptons can arise from mismeasurements of prompt tracks or from decays of tau leptons and B and D mesons, which have proper decay lengths of  $87 \mu\text{m}$ ,  $500 \mu\text{m}$ , and less than  $100 \mu\text{m}$ , respectively. The tight isolation criteria used in this analysis reject the vast majority of heavy flavor background.

The data-driven background estimation method will account for all three types of background sources, namely, mismeasurements, tau lepton decays, and heavy flavor decays, as will be shown in Section 6. However, some processes in which both leptons come from long-lived SM parents, such as  $\text{DY} \rightarrow \tau\tau \rightarrow ll$ , will cause the leptons to have correlated  $|d_0|$  values. The muon  $|d_0|$  resolution is better than that of electrons, so we expect the correlation to be more pronounced in the  $\mu\mu$  channel. The difference in the electron and muon  $|d_0|$  resolution is clear when one observes the fraction of leptons from different parents as a function of lepton  $|d_0|$ , as shown in Fig. 18 for  $t\bar{t}$  events that pass the  $e\mu$  channel preselection. Mismeasurements dominate for electrons at all  $|d_0|$ . In contrast, muons from long-lived SM parents, namely, tau leptons and heavy-flavor processes, are most important for  $|d_0| > 100 \mu\text{m}$ .

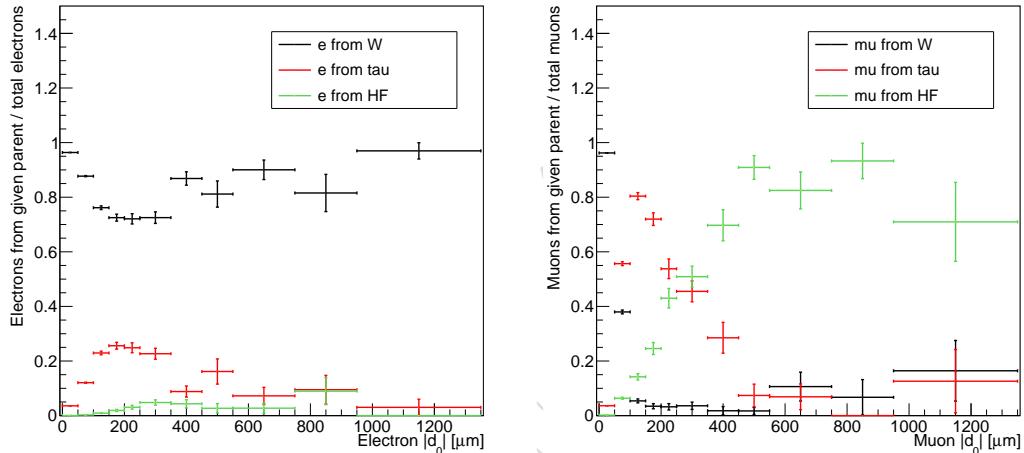


Figure 18: The fraction of electrons (left) and muons (right) from different parents as a function of lepton  $|d_0|$ , for  $t\bar{t}$  events in 2018 conditions that pass the  $e\mu$  channel preselection.

In the  $\mu\mu$  channel, it is worth examining which long-lived SM parents will contribute to  $|d_0|-|d_0|$  correlation. The correlation specifically comes from DY-type processes in which the parentage is correlated between muons. Figure 19, which shows the fraction of muons from different background sources in DY simulation, indicates that tau lepton decays are the main source of muons that may be correlated in this way, and that the heavy-flavor contribution is negligible. This is reasonable because while tau leptons and heavy-flavor mesons both produce displaced muons, the isolation criteria rejects the vast majority of muons from heavy-flavor mesons. Muons from tau leptons contribute significantly from about  $100$  to  $500 \mu\text{m}$ , so we expect the most significant  $|d_0|-|d_0|$  correlation to appear in this range and peak around  $200 \mu\text{m}$ . Furthermore, the correlation will be most pronounced in the regions where the  $|d_0|$  measurements are the best.

Also see Appendix C for more discussion on other sources of  $|d_0|$  correlation.

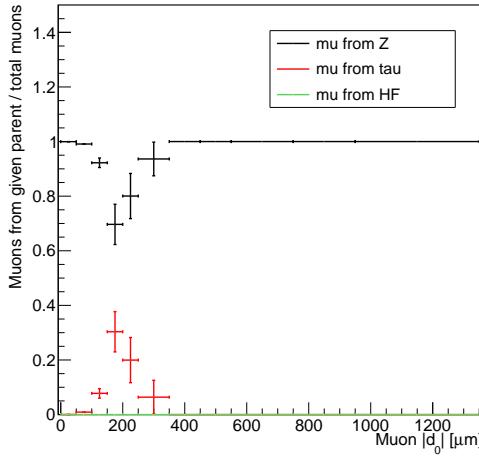


Figure 19: The fraction of muons from different parents as a function of muon  $|d_0|$ , for DY events in 2018 conditions that pass the  $\mu\mu$  channel preselection.

## 534 6 Background estimation

535 We use a data-driven ABCD method to estimate the background in the signal region. Section  
 536 6.1 gives an overview of the method, and Section 6.2 describes the signal regions. Section 6.3  
 537 shows closure tests in sideband regions where one lepton is prompt and the other is displaced.  
 538 Section 6.4 describes a correction to the estimate and the associated systematic uncertainty.  
 539 Then, Section 6.5 shows closure tests in background simulation in the signal regions. The back-  
 540 ground estimates themselves are listed with the results in Section 9.

### 541 6.1 Data-driven ABCD method

542 The data-driven method does not distinguish between background processes, but instead uses  
 543 the  $|d_0|$  distributions as composite models of all background processes together. To estimate  
 544 the amount of background in the SRs, we employ an ABCD method that uses each lepton  $|d_0|$ .  
 545 First, we categorize the events that pass the preselection criteria into four regions (A, B, C, and  
 546 D) based on each lepton  $|d_0|$  as shown in Fig. 20.

547 To easily diagram the ABCD method in all channels, we define  $|d_0^a|$  and  $|d_0^b|$ . For the  $e\mu$  channel,  
 548  $|d_0^a|$  is defined as the leading electron  $|d_0|$  and  $|d_0^b|$  is defined as the leading muon  $|d_0|$ . For the  
 549 ee ( $\mu\mu$ ) channel,  $|d_0^a|$  is defined as the leading electron (muon)  $|d_0|$  and  $|d_0^b|$  is defined as the  
 550 subleading electron (muon)  $|d_0|$ . Two-dimensional plots of these two variables are shown for  
 551 each channel in Fig. 21. Additional plots of these variables are shown in Appendix D.6.

552 We use the number of background events in regions A, B, and C to estimate the expected back-  
 553 ground in region D, which is the inclusive signal region. We assume that  $N_B/N_A = N_D/N_C$   
 554 or that the number of background events in D is  $N_B N_C / N_A$ , where  $N_X$  is the number of back-  
 555 ground events in the given region. This assumption is only true if  $|d_0^a|$  and  $|d_0^b|$  are uncorrelated,  
 556 and so the baseline ABCD method will not account for the correlation discussed in Section 5.  
 557 As will be shown in Section 6.3, we indeed observe the expected correlation in the 100–500  $\mu\text{m}$   
 558 region of the one-prompt/one-displaced sidebands. Using sideband closure test results, we  
 559 assign a correction and systematic uncertainty to cover this correlation in all three channels, as  
 560 will be described in Section 6.4.

561 When performing the background estimate and closure tests, we separate 2016 data and sim-  
 562 ulation from 2017–2018 data and simulation because we observed a difference in the lepton  $d_0$

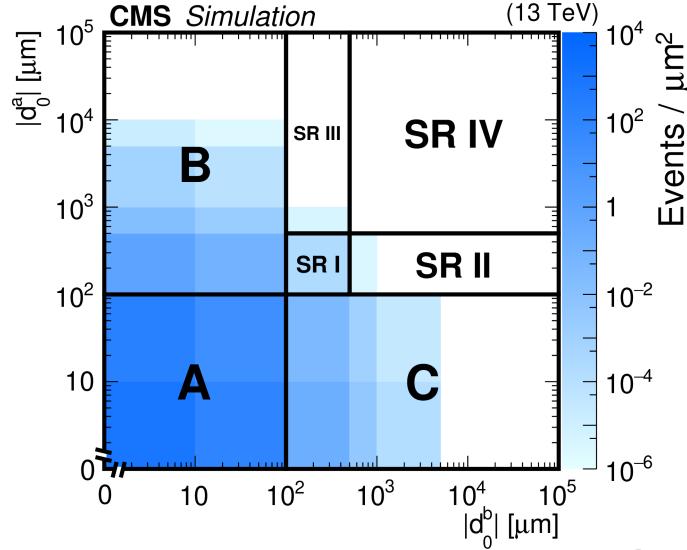


Figure 20: A diagram of the ABCD method, shown for illustration on background simulation passing the  $e\mu$  preselection with 2018 conditions. In each  $|d_0|$ - $|d_0|$  bin, the number of events divided by the bin area is plotted. A, B, and C are control regions, and D is the inclusive SR. SRs I–IV will be described below.

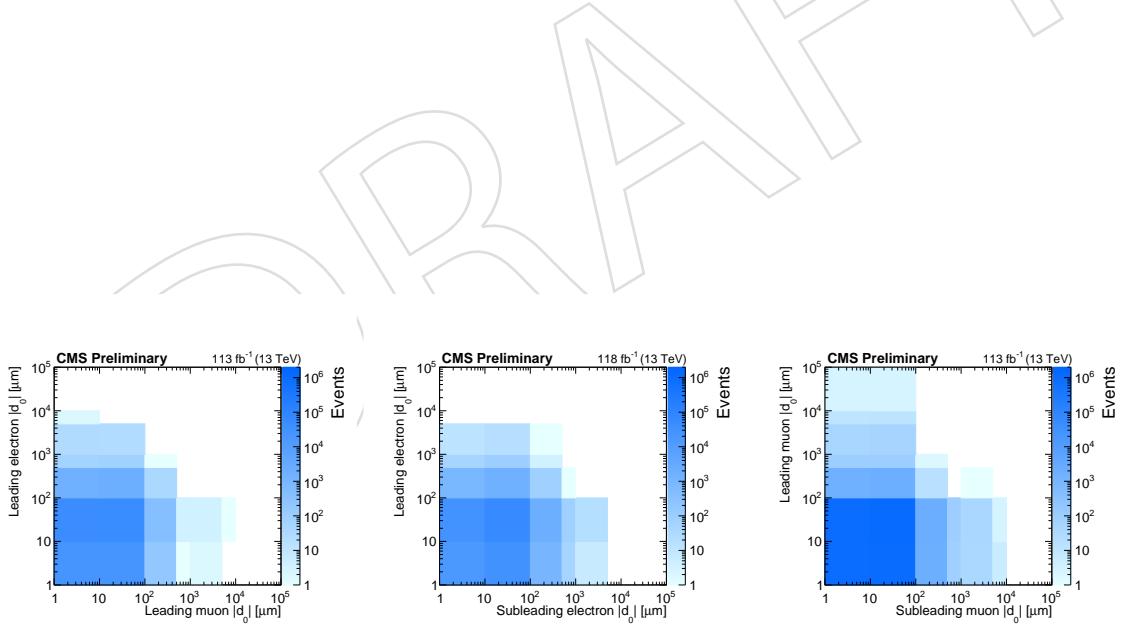


Figure 21: The  $|d_0^a|$  as a function of the  $|d_0^b|$ , for the background simulated events that pass the  $e\mu$  (left),  $ee$  (middle), and  $\mu\mu$  (right) preselection with 2018 conditions. If a  $|d_0|$  value is less than unity, it is set to unity for plotting. The inclusive signal region covers the region between  $100 \mu\text{m}$  and  $10 \text{ cm}$  in each  $|d_0|$  variable shown.

resolution in 2016 due to the Phase 0 tracker and in 2017–2018 due to the Phase 1 tracker (see Appendix C).

## 6.2 Signal search regions (SRs)

In each channel, we estimate the background in several SRs binned in  $|d_0|$  and  $p_T$ .

For events to enter the inclusive SR, they must pass the preselection criteria and both  $|d_0^a|$  and  $|d_0^b|$  must satisfy  $100\text{ }\mu\text{m} < |d_0| < 10\text{ cm}$ . We define nonoverlapping SRs in  $|d_0|$ , as shown in Fig. 20 and as given below:

- SR I:  $100 < \text{both } |d_0| < 500\text{ }\mu\text{m}$
- SR II:  $100 < |d_0^a| < 500\text{ }\mu\text{m}, 500\text{ }\mu\text{m} < |d_0^b| < 10\text{ cm}$
- SR III:  $500\text{ }\mu\text{m} < |d_0^a| < 10\text{ cm}, 100 < |d_0^b| < 500\text{ }\mu\text{m}$
- SR IV:  $500\text{ }\mu\text{m} < \text{both } |d_0| < 10\text{ cm}$

These SRs are defined primarily because of the expected contribution of the different background sources, as was described in Section 5. In addition, we obtain loose SRs with some amount of background contamination but high signal efficiency and tight SRs with little background contamination but also smaller signal efficiency. Because they are nonoverlapping, we can use these SRs simultaneously in the limit setting procedure.

We also bin SR I, which has the largest amount of background events, in one of the lepton  $p_T$  in order to maximize sensitivity.

We perform a separate ABCD estimate for each SR. When performing the estimates, we subdivide regions B and C into  $100\text{--}500\text{ }\mu\text{m}$  and  $500\text{ }\mu\text{m}\text{--}10\text{ cm}$  regions to match the SR definitions and bin region A and the  $100\text{--}500\text{ }\mu\text{m}$  subregions of B and C in  $p_T$  to match the binning of SR I.

## 6.3 Closure tests in one-prompt/one-displaced sidebands

### 6.3.1 100–500 $\mu\text{m}$ region

We perform closure tests in control regions where one lepton is more prompt ( $20\text{--}100\text{ }\mu\text{m}$ ) and the other is more displaced ( $100\text{--}500\text{ }\mu\text{m}$ ). (The prompt lepton region A starts at  $20\text{ }\mu\text{m}$ , but region D is defined from  $30\text{ }\mu\text{m}$  and greater.) Then, from these closure tests, we use the ratio of the actual number of events to the estimated number of events as the measure of nonclosure.

Using the procedure outlined in 6.4, we extrapolate these ratios from the closure tests to the inclusive signal region.

Table 12 shows the extrapolated ratios for background simulation, with and without  $DY \rightarrow \tau\tau$  events, and for data, in the one-prompt/one-displaced sidebands. Without the  $DY \rightarrow \tau\tau$  events, the ratios are compatible with 1, as shown in the second column, indicating there is no correlation. When the  $DY \rightarrow \tau\tau$  events are added back in, as shown in the third column, the ratios are generally greater than 1, which indicates some positive correlation. Furthermore, the level of correlation observed with the full background simulation (column 3) is compatible with the level of correlation observed in data (column 4). Thus, we indeed observe correlation in the one-prompt/one-displaced ( $100\text{--}500\text{ }\mu\text{m}$ ) sidebands due to  $DY \rightarrow \tau\tau$  events.

### 6.3.2 500 $\mu\text{m}$ –10 cm region

We then perform closure tests in control regions where one lepton is more prompt ( $20\text{--}100\text{ }\mu\text{m}$ ) and the other is more displaced ( $500\text{ }\mu\text{m}\text{--}10\text{ cm}$ ). (The prompt lepton region A starts at  $20\text{ }\mu\text{m}$ , but region D is defined from  $30\text{ }\mu\text{m}$  and greater.) From these closure tests, we again use the ratio

Table 12: Closure test results in background simulation (with and without  $DY \rightarrow \tau\tau$  events) and in data, in the  $100\text{--}500\,\mu\text{m}$  region. The extrapolated ratios (averaged over the two one-prompt/one-displaced sidebands) and their statistical uncertainties are given. The A, B, C, and D regions are defined as follows: A is  $20\text{--}30\,\mu\text{m}$  in prompt lepton  $|d_0|$  and  $20\text{--}100\,\mu\text{m}$  in displaced lepton  $|d_0|$ , B is  $20\text{--}30\,\mu\text{m}$  in prompt-lepton  $|d_0|$  and  $100\text{--}500\,\mu\text{m}$  in displaced lepton  $|d_0|$ , C is always  $20\text{--}100\,\mu\text{m}$  in displaced lepton  $|d_0|$ , D (the test region) is always  $100\text{--}500\,\mu\text{m}$  in displaced lepton  $|d_0|$ , and we perform repeated tests while simultaneously varying the C and D prompt lepton  $|d_0|$ s within the  $30\text{--}100\,\mu\text{m}$  range.

	Ratio for bkg. simulation without $DY \rightarrow \tau\tau$	Ratio for all bkg. simulation	Ratio for data
2016 e $\mu$	$0.9 \pm 0.3$	$1.6 \pm 0.6$	$0.9 \pm 1.3$
2017+2018 e $\mu$	$1.1 \pm 0.4$	$1.6 \pm 0.7$	$3.1 \pm 0.8$
2016 ee	$0.8 \pm 0.5$	$0.8 \pm 0.5$	$0.6 \pm 0.6$
2017+2018 ee	$0.8 \pm 1.0$	$1.6 \pm 0.9$	$1.5 \pm 0.4$
2016 $\mu\mu$	$1.1 \pm 0.8$	$2.0 \pm 0.8$	$2.5 \pm 0.9$
2017+2018 $\mu\mu$	$2.6 \pm 2.8$	$7.8 \pm 3.7$	$4.2 \pm 1.5$

of the actual number of events to the estimated number of events as the measure of nonclosure.  
We extrapolate these ratios from the closure tests to the inclusive signal region.

Table 13 shows the ratios for background simulation and for data in the one-prompt/one-displaced sidebands. In both data and background simulation, there is no evidence of positive correlation beyond  $500\,\mu\text{m}$ .

Table 13: Closure test results in data and background simulation (with and without  $DY \rightarrow \tau\tau$  events), in the  $500\,\mu\text{m}\text{--}10\,\text{cm}$  region, for the two one-prompt/one-displaced sidebands. The ratios of the actual yield to the estimated yield and their statistical uncertainties are given. “Sideband 1 (2)” refers to the case when “lepton a” is displaced (prompt) and “lepton b” is prompt (displaced). The A, B, C, and D regions are defined as follows: A is  $20\text{--}30\,\mu\text{m}$  in prompt lepton  $|d_0|$  and  $20\text{--}100\,\mu\text{m}$  in displaced lepton  $|d_0|$ , B is  $20\text{--}30\,\mu\text{m}$  in prompt lepton  $|d_0|$  and  $500\,\mu\text{m}\text{--}10\,\text{cm}$  in displaced lepton  $|d_0|$ , C is  $30\text{--}100\,\mu\text{m}$  in prompt lepton  $|d_0|$  and  $20\text{--}100\,\mu\text{m}$  in displaced lepton  $|d_0|$ , and D (the test region) is  $30\text{--}100\,\mu\text{m}$  in prompt lepton  $|d_0|$  and  $500\,\mu\text{m}\text{--}10\,\text{cm}$  in displaced lepton  $|d_0|$ .

	Sideband 1			Sideband 2		
	Data	Bkg. simulation	Bkg. simulation w/o $DY \rightarrow \tau\tau$	Data	Bkg. simulation	Bkg. simulation w/o $DY \rightarrow \tau\tau$
2016 e $\mu$	$0.4^{+1.0}_{-0.4}$	$1.1 \pm 0.3$	$1.1 \pm 0.3$	$1.0\,(0\text{ vs }0)$	$0.8^{+0.4}_{-0.3}$	$0.8^{+0.4}_{-0.3}$
2017+2018 e $\mu$	$0.7 \pm 0.3$	$0.9^{+0.3}_{-0.2}$	$0.9^{+0.3}_{-0.2}$	$0.7^{+1.3}_{-0.7}$	$0.8^{+0.3}_{-0.2}$	$0.8^{+0.3}_{-0.2}$
2016 ee	$1.4^{+1.6}_{-0.9}$	$0.4^{+0.6}_{-0.3}$	$0.4^{+0.6}_{-0.3}$	$0.7^{+1.0}_{-0.6}$	$4.0^{+5.8}_{-3.1}$	$4.0^{+5.8}_{-3.1}$
2017+2018 ee	$1.0 \pm 0.3$	$0.3^{+0.4}_{-0.2}$	$0.5^{+0.8}_{-0.4}$	$1.0 \pm 0.3$	$2.1^{+2.6}_{-1.5}$	$3.5^{+2.6}_{-1.8}$
2016 $\mu\mu$	$0.8 \pm 0.3$	$0.7 \pm 0.3$	$0.7 \pm 0.3$	$0.6^{+0.4}_{-0.3}$	$1.3^{+0.6}_{-0.4}$	$1.2^{+0.5}_{-0.4}$
2017+2018 $\mu\mu$	$1.8^{+0.6}_{-0.7}$	$0.4^{+1.0}_{-0.4}$	$0.8^{+1.8}_{-0.7}$	$0.5^{+0.3}_{-0.2}$	$0.5^{+0.5}_{-0.3}$	$0.4^{+0.4}_{-0.3}$

## 609 6.4 Correction to estimate and systematic uncertainty

610 We have shown that there is positive correlation due to tau leptons in the 100–500  $\mu\text{m}$  region,  
 611 and we find no evidence for such correlation in the 500  $\mu\text{m}$ –10 cm region. As a result, we now  
 612 define a procedure to correct for this correlation and to assign a systematic uncertainty for it.

613 First, we divide each one-prompt/one-displaced sideband into two subregions in the displaced  
 614 lepton  $|d_0|$ : (1) the 100–500  $\mu\text{m}$  region, where we expect correlation from tau leptons to be  
 615 significant, and (2) the 500  $\mu\text{m}$ –10 cm region, where we expect correlation from tau leptons to  
 616 be insignificant. The 100–500  $\mu\text{m}$  (500  $\mu\text{m}$ –10 cm) sideband region is used as a control region  
 617 for SR I (SRs II–IV). We perform closure tests in data in each subregion, and use the ratio of the  
 618 actual to the estimated number of events as measure of nonclosure. From the 500  $\mu\text{m}$ –10 cm  
 619 region tests, we take the largest deviation of the ratio from 1.0 as a systematic uncertainty in  
 620 SRs II–IV, and apply no correction. In the 100–500  $\mu\text{m}$  region, we fit the ratio as function of  
 621 the prompt lepton  $|d_0|$  and extrapolate to 200  $\mu\text{m}$  (within SR I). We take the average of the  
 622 two extrapolated ratios (one from each one-prompt/one-displaced sideband), and derive a  
 623 correction and systematic uncertainty from this average ratio. The details of this correction and  
 624 systematic uncertainty are shown below, for each of the two subregions.

### 625 6.4.1 100–500 $\mu\text{m}$ region

626 As shown above, a correction to the background estimate, due to lepton  $|d_0|$  correlation, is  
 627 needed in the 100–500  $\mu\text{m}$  region. Figures 22, 23, and 24 show the results of the closure tests in  
 628 the  $e\mu$ ,  $ee$ , and  $\mu\mu$  channels, respectively, in the one-prompt (20–100  $\mu\text{m}$ )/one-displaced (100–  
 629 500  $\mu\text{m}$ ) sidebands in the 100–500  $\mu\text{m}$  region, in data. These plots show the ratio of the actual  
 630 to the estimated number of events as a function of the prompt lepton  $|d_0|$ . In all of these plots,  
 631 the binning of the prompt lepton axis is initially 10  $\mu\text{m}$  wide. Starting from most-displaced  
 632 bin, we test to see if any bin has fewer than 5 events, and if so, we combine it with whichever  
 633 neighboring bin has fewer events, repeating until all bins have at least 5 events.

634 The data are fitted with a straight line, where the slope and y-intercept are allowed to vary<sup>2</sup>.  
 635 This fit is used to find the ratio at 200  $\mu\text{m}$  (the extrapolation into SR I), which is where we  
 636 expect the largest contribution from tau lepton decays, as was shown in Section 5. A second  
 637 reason why we extrapolate to 200  $\mu\text{m}$  is that this is approximately the center-of-mass of the  
 638 100–500  $\mu\text{m}$  bin in background simulation. We average the two extrapolated ratios (one from  
 639 each one-prompt/one-displaced sideband), and derive a correction and systematic uncertainty  
 640 from this average ratio.

641 If the average is  $> 1.0$ , we use the average as a correction to the background estimate and we  
 642 use the uncertainty in the average as a systematic uncertainty in the background estimate. In  
 643 this case, we also vary the 200  $\mu\text{m}$  extrapolation point by  $\pm 50 \mu\text{m}$ , as we find that the width of  
 644 the tau lepton contribution as a function of  $|d_0|$  is about 50  $\mu\text{m}$ . We apply the difference from  
 645 this variation in extrapolation point as an additional systematic uncertainty in the background  
 646 estimate. If the average is  $\leq 1.0$ , we set the correction equal to 1.0 and use the uncertainty in

<sup>2</sup>We have also tried fitting the data with an exponential, and it generally results in a similar p-value as the linear fit. In a few cases, the fit parameter values are such that the exponential function effectively mimics a linear function over the relevant range and results in a very similar extrapolated ratio. In all other cases, the sensitivity of the function to the exponential parameter produces very large uncertainties on the extrapolated ratio (generally  $> 100\%$  relative uncertainty). In a couple cases, the exponential fit results in an unreasonably large extrapolated ratio and uncertainty that we do not believe is physical. Taking systematic uncertainties from the differences between the exponential and linear extrapolated ratio central values would result in unreasonably large uncertainties in some cases and negligible uncertainties in others, and the linear fit results are generally more robust, so we choose to simply use the existing uncertainties on the linear fit.

the average as a symmetric systematic uncertainty about 1.0. The correction factors and the uncorrected and corrected background estimates for the 100–500  $\mu\text{m}$  region are summarized in Table 14.

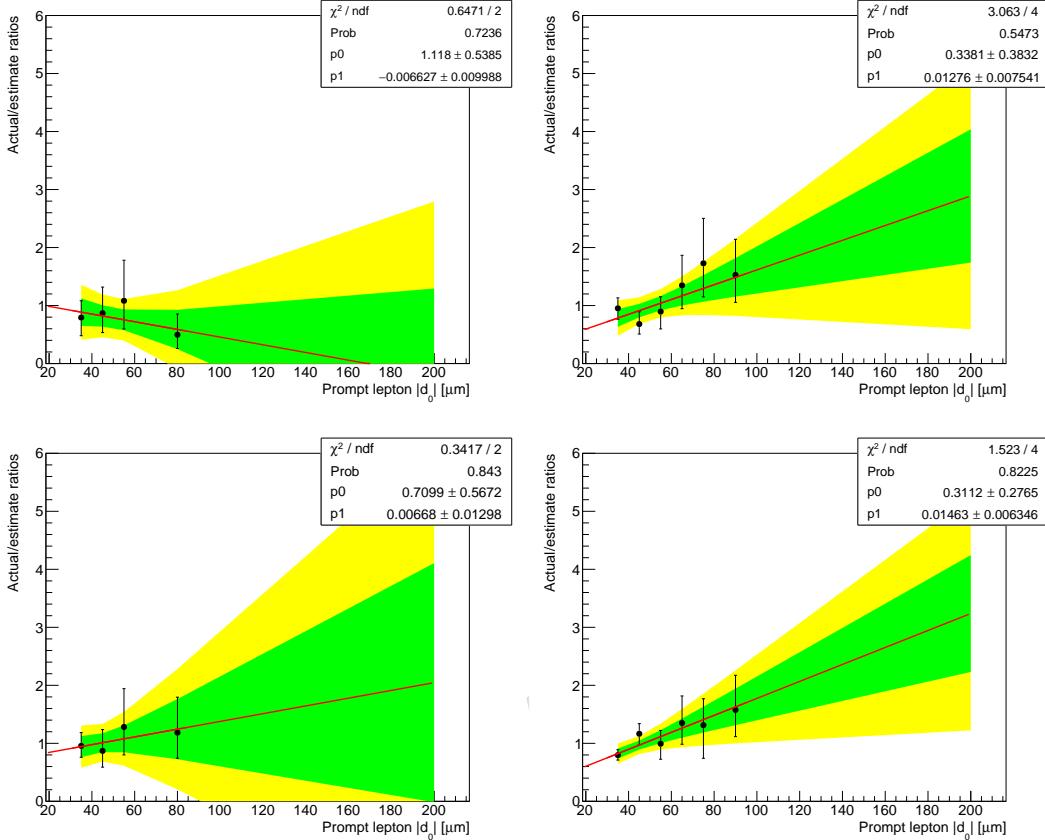


Figure 22: Background estimation closure tests in data, in the one-prompt (20–100  $\mu\text{m}$ )/one-displaced (100–500  $\mu\text{m}$ ) sidebands, in the  $e\mu$  channel. The prompt leading electron/ displaced leading muon sideband is shown in the upper row, and the prompt leading muon/ displaced leading electron sideband is shown in the lower row. The plots on the left show the results for 2016 data, and the plots on the right are for combined 2017 and 2018 data. The plots show the ratio of the actual to the estimated number of events as a function of the prompt lepton  $|d_0|$ . The data are fitted with a straight line, where the slope and y-intercept are allowed to vary. The  $1\sigma$  and  $2\sigma$  confidence intervals are shown in the green and yellow bands, respectively.

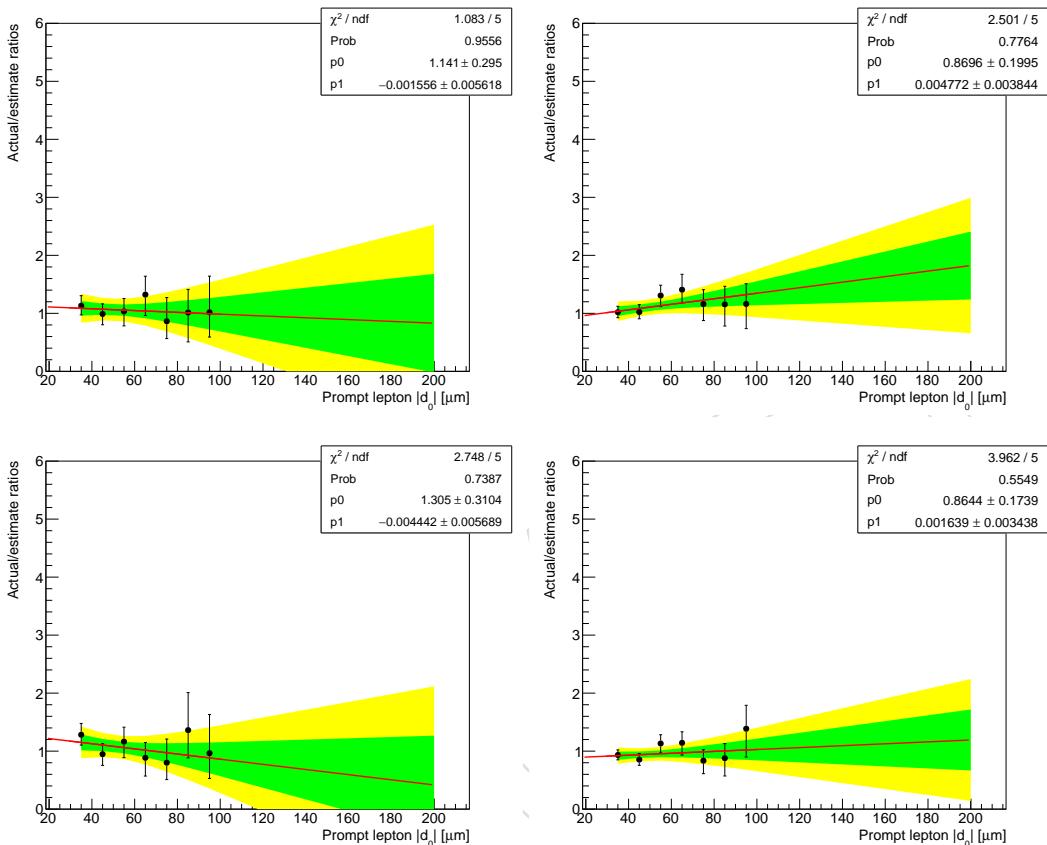


Figure 23: Background estimation closure tests in data, in the one-prompt ( $20\text{--}100\,\mu\text{m}$ )/one-displaced ( $100\text{--}500\,\mu\text{m}$ ) sidebands, in the ee channel. The prompt leading electron/ displaced subleading electron sideband is shown in the upper row, and the prompt subleading electron/ displaced leading electron sideband is shown in the lower row. The plots on the left show the results for 2016 data, and the plots on the right are for combined 2017 and 2018 data. The plots show the ratio of the actual to the estimated number of events as a function of the prompt lepton  $|d_0|$ . The data are fitted with a straight line, where the slope and y-intercept are allowed to vary. The  $1\sigma$  and  $2\sigma$  confidence intervals are shown in the green and yellow bands, respectively.

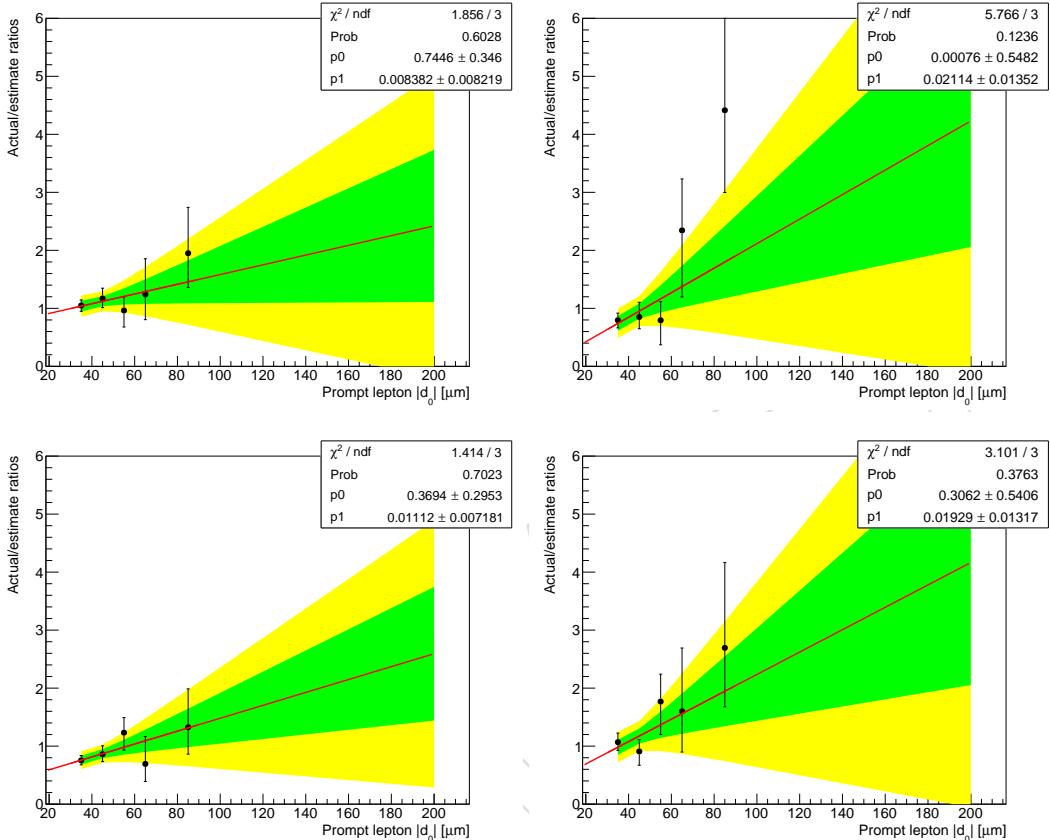


Figure 24: Background estimation closure tests in data, in the one-prompt ( $20\text{--}100\,\mu\text{m}$ )/one-displaced ( $100\text{--}500\,\mu\text{m}$ ) sidebands, in the  $\mu\mu$  channel. The prompt leading muon/ displaced subleading muon sideband is shown in the upper row, and the prompt subleading muon/ displaced leading muon sideband is shown in the lower row. The plots on the left show the results for 2016 data, and the plots on the right are for combined 2017 and 2018 data. The plots show the ratio of the actual to the estimated number of events as a function of the prompt lepton  $|d_0|$ . The data are fitted with a straight line, where the slope and y-intercept are allowed to vary. The  $1\sigma$  and  $2\sigma$  confidence intervals are shown in the green and yellow bands, respectively.

Table 14: The correction factors and the uncorrected and corrected background estimates for the  $100\text{--}500\,\mu\text{m}$  region, as derived from data. The correction factor uncertainty includes both the uncertainty in the average and the additional uncertainty obtained from varying the fit extrapolation point. The total uncertainty (statistical plus systematic) is given for the corrected background estimate.

	Correction factor	Uncorrected estimate	Corrected estimate
2016 $e\mu$	$1.0^{+1.3}_{-1.0}$	$4.21^{+0.38}_{-0.40}$	$4.2^{+5.4}_{-4.2}$
2017+2018 $e\mu$	$3.1 \pm 1.0$	$12.53^{+0.64}_{-0.61}$	$38 \pm 13$
2016 ee	$1.00 \pm 0.60$	$18.30^{+0.94}_{-0.91}$	$18 \pm 11$
2017+2018 ee	$1.51^{+0.43}_{-0.42}$	$41.6 \pm 1.3$	$63^{+18}_{-17}$
2016 $\mu\mu$	$2.5 \pm 1.0$	$3.07 \pm 0.08$	$7.7 \pm 3.1$
2017+2018 $\mu\mu$	$4.2 \pm 1.8$	$1.00 \pm 0.04$	$4.2 \pm 1.8$

#### 650 6.4.2 $500\,\mu\text{m}\text{--}10\,\text{cm}$ region

651 As shown above, no correction to the background estimate is needed in the  $500\,\mu\text{m}\text{--}10\,\text{cm}$  re-  
 652 gion, since the tau lepton contribution is negligible here. In this region, we derive a system-  
 653 atic uncertainty in the background estimate from the one-prompt ( $20\text{--}100\,\mu\text{m}$ )/one-displaced  
 654 ( $500\,\mu\text{m}\text{--}10\,\text{cm}$ ) closure tests in data. We take the largest deviation from 1 that occurs in the  
 655 ratio of the actual to the estimated number of events plus its uncertainty, in either of the two  
 656 one-prompt/one-displaced closure tests, as a systematic uncertainty. This is a conservative  
 657 approach that produces a large systematic uncertainty in the small background estimates that  
 658 we find in this region. Table 15 shows the systematic uncertainty and the predicted number of  
 659 events in SRs II–IV.

Table 15: The systematic uncertainty and the background estimates in SRs II–IV (the  $500\,\mu\text{m}\text{--}10\,\text{cm}$  region). The total uncertainty (statistical plus systematic) is given for the background estimates in SRs II–IV.

	Systematic uncertainty for SRs II–IV	SR II	SR III	SR IV
2016 $e\mu$	98%	$0.15 \pm 0.15$	$0.09^{+0.12}_{-0.09}$	$0.003^{+0.004}_{-0.003}$
2017+2018 $e\mu$	106%	$0.71^{+0.76}_{-0.71}$	$0.23^{+0.27}_{-0.23}$	$0.01^{+0.02}_{-0.01}$
2016 ee	199%	$0.51^{+1.02}_{-0.51}$	$0.43^{+0.85}_{-0.43}$	$0.01^{+0.02}_{-0.01}$
2017+2018 ee	37%	$3.6 \pm 1.4$	$2.8 \pm 1.1$	$0.24^{+0.10}_{-0.09}$
2016 $\mu\mu$	64%	$0.17 \pm 0.11$	$0.19 \pm 0.12$	$0.01 \pm 0.01$
2017+2018 $\mu\mu$	140%	$0.14^{+0.19}_{-0.14}$	$0.08^{+0.12}_{-0.08}$	$0.01^{+0.02}_{-0.01}$

#### 660 6.5 Closure tests in signal regions

661 With the aforementioned correction applied, we can perform closure tests in background sim-  
 662 ulation in the four SRs. In these tests, both leptons are displaced. The results of these closure  
 663 tests in the SRs, namely, the ratio of the actual to the estimated yield, is shown in Table 16.  
 664 Since the ratios are compatible with 1, these closure tests show that the correction performs as  
 665 expected, and the systematic uncertainties are sufficient to cover any unforeseen correlation.

Table 16: Closure test results in background simulation in the SRs, with the correction applied. The estimated number of events, the actual number of events, and their total uncertainties (statistical plus systematic) are given. In cases where the actual number of events is zero, the uncertainty is given by the product of the average background simulation event weight and the upper bound of the 68% Poisson interval given by a single observation of zero events.

	SR I	SR II	SR III	SR IV
2016 $e\mu$ estimated	$7.4^{+4.8}_{-4.2}$	$0.07 \pm 0.07$	$0.096^{+0.105}_{-0.096}$	$0.001 \pm 0.001$
2016 $e\mu$ actual	$5.0^{+1.5}_{-1.2}$	$0.07^{+0.09}_{-0.05}$	$0.005^{+0.011}_{-0.004}$	$0.000^{+0.037}_{-0.000}$
2017+2018 $e\mu$ estimated	$13.5 \pm 6.4$	$0.37^{+0.40}_{-0.37}$	$0.34^{+0.36}_{-0.34}$	$0.02 \pm 0.02$
2017+2018 $e\mu$ actual	$19.1^{+11.4}_{-7.6}$	$0.52^{+0.41}_{-0.25}$	$0.00^{+0.24}_{-0.00}$	$0.00^{+0.24}_{-0.00}$
2016 $ee$ estimated	$9.3 \pm 5.0$	$0.12^{+0.23}_{-0.12}$	$0.14^{+0.28}_{-0.14}$	$0.002^{+0.004}_{-0.002}$
2016 $ee$ actual	$13.4^{+3.4}_{-2.8}$	$0.15^{+0.19}_{-0.09}$	$1.03^{+1.36}_{-0.67}$	$0.000^{+0.550}_{-0.000}$
2017+2018 $ee$ estimated	$18 \pm 11$	$0.59^{+0.27}_{-0.26}$	$0.45^{+0.21}_{-0.20}$	$0.02 \pm 0.01$
2017+2018 $ee$ actual	$8.2^{+6.5}_{-3.9}$	$0.17^{+0.23}_{-0.11}$	$0.00^{+0.17}_{-0.00}$	$0.00^{+0.17}_{-0.00}$
2016 $\mu\mu$ estimated	$1.3 \pm 0.6$	$0.04 \pm 0.04$	$0.03 \pm 0.03$	$0.002 \pm 0.002$
2016 $\mu\mu$ actual	$3.3^{+1.8}_{-1.2}$	$0.11^{+0.14}_{-0.07}$	$0.06^{+0.14}_{-0.05}$	$0.000^{+0.110}_{-0.000}$
2017+2018 $\mu\mu$ estimated	$2.7 \pm 1.4$	$0.04 \pm 0.04$	$0.02 \pm 0.02$	$0.002 \pm 0.002$
2017+2018 $\mu\mu$ actual	$7.1^{+6.9}_{-3.8}$	$0.00^{+0.15}_{-0.00}$	$0.00^{+0.15}_{-0.00}$	$0.078^{+0.179}_{-0.064}$

## 666 7 Additional background checks

667 We performed a few additional studies to check other potential sources of background in the  
 668 SRs. We find that they are either negligible in the SRs or already covered by the background  
 669 estimation method as described above.

### 670 7.1 Material interactions

671 In order to further study the material interactions, we invert the criterion in the preselection  
 672 that rejects good vertices in the material. In data, we find seven events, across all channels  
 673 and years, that pass the preselection with this inverted selection, as shown in Table 17. Of  
 674 these seven events, three are in the prompt control region, and four are in the sidebands (either  
 675 region B or region C of the ABCD method). We checked the vertex for each of these events, and  
 676 they are indeed truly displaced (two vertices are in the beampipe, one in the BPIX inner shield,  
 677 and four in BPIX layer 1). Thus, none of the events found with the material veto inverted are  
 678 in the signal regions, and so we are confident that material interactions are not a significant  
 679 background after the full selection is applied.

Table 17: Some details of the seven events found in data with the material interactions selection inverted.

Channel, year	$d_0$ lept0, $d_0$ lept1 [ $\mu\text{m}$ ]	vertex position (x, y, z) [ cm ]
e $\mu$ 2017C	-14 (electron), -9.5 (muon) (PCR)	(-2.5, 1.4, 6.8) (BPIX L1)
e $\mu$ 2018D	46 (electron), -14 (muon) (PCR)	(0.89, 2.1, 0.071) (beampipe)
ee 2018D	198 (electron0), -34 (electron1) (sideband)	(-1.9, 0.47, 2.7) (beampipe)
$\mu\mu$ 2016G	407 (muon0), -8.3 (muon1) (sideband)	(-1.4, 4.0, 6.3) (BPIX L1)
$\mu\mu$ 2016G	-17 (muon0), -2215 (muon1) (sideband)	(-2.6, 3.1, 6.6) (BPIX L1)
$\mu\mu$ 2016H	1.6 (muon0), 0.099 (muon1) (PCR)	(-1.6, -3.5, 12) (BPIX inner shield)
$\mu\mu$ 2017F	522 (muon0), -13 (muon1) (sideband)	(-1.1, -3.0, -7.5) (BPIX L1)

### 680 7.2 Cosmic-ray muons

681 We performed a study in which we invert the  $\Delta t$  and  $\cos \alpha$  criteria in the  $\mu\mu$  preselection, and  
 682 then ask how many of these events in data are in the SRs. We find three events in data with these  
 683 criteria (one event per year, all in SR IV). This is a very maximum upper bound on the number  
 684 of cosmic-ray muons we expect in the signal region. Then, we find the efficiency to pass the  $\Delta t$   
 685 and  $\cos \alpha$  criteria in NoBPTX data, which is predominantly cosmic-ray muon events, on top of  
 686 the rest of the  $\mu\mu$  preselection criteria. Zero events in NoBPTX data pass all of the preselection  
 687 criteria, and 3736 events in NoBPTX data pass all of the preselection criteria except these two  
 688 cosmic rejection selections. To be conservative, we fluctuate the 0 events up to 1 event, and thus  
 689 find an efficiency of 1/3736. If we multiply this efficiency by the three cosmic-ray muon events  
 690 observed in  $\mu\mu$  data in the SRs, we find 0.0008 events as a reasonable upper bound on the  
 691 number of expected cosmic-ray muon events, which is negligible compared to the background  
 692 estimation in each SR.

### 693 7.3 Heavy-flavor mesons

694 To estimate an upper limit on the heavy-flavor background in the signal region, two studies  
 695 were performed.

696 First, we performed the ABCD method in 2018  $\mu\mu$  data, but we additionally required at least  
 697 one medium CSVv2 b-tagged jet. The  $\mu\mu$  channel contains the smallest relative SR contribution  
 698 from mismeasurements, so it should therefore be the most sensitive to heavy flavor. As shown

699 in Table 18, the background estimates with this check are about an order of magnitude smaller  
700 than when no b jet requirement is applied in our usual preselection.

Table 18: Background estimates in the 2018  $\mu\mu$  channel, for the nominal prediction and with the additional requirement of at least one medium b-tagged jet. The estimates with at least one b jet are about an order of magnitude below the nominal prediction, indicating that any additional contribution from heavy flavor events in the SRs is negligible.

	SR I	SR II	SR III	SR IV
2018 $\mu\mu$ preselection (corrected)	$2.64 \pm 0.96$	$0.09^{+0.12}_{-0.09}$	$0.050^{+0.070}_{-0.050}$	$0.0068^{+0.0096}_{-0.0069}$
2018 $\mu\mu$ preselection + 1 b jet	$0.190^{+0.029}_{-0.026}$	$0.0075^{+0.0073}_{-0.0041}$	$0.0046^{+0.0037}_{-0.0023}$	$0.00018^{+0.00023}_{-0.00013}$

701 In addition, we look at samples in 2018 data and QCD simulation in which we invert the isolat-  
702 ion criterion for events that pass the  $\mu\mu$  preselection. These samples are dominated by muons  
703 from the decays of B mesons, and the QCD simulation describes the data well in the region  
704 outside of the Z boson peak, as can be seen in Fig. 25. We use this QCD sample to test the  
705 heavy-flavor background in two ways. First, we perform a naive ABCD estimate in the QCD  
706 simulation to check for  $|d_0|$  correlation. As shown in Table 19, there is no sign of correlation,  
707 indicating that the nominal background estimation already accounts for the heavy-flavor back-  
708 ground. Secondly, we estimate the approximate heavy-flavor background in the SRs. We take  
709 the ratio of the SR to prompt control region events from the QCD simulation in the anti-isolated  
710 region and the normalization from the number of QCD simulated events that pass the nominal  
711  $\mu\mu$  preselection. Using this approach, we estimate that the heavy-flavor background is  $0.06^{+0.13}_{-0.05}$   
712 events in SR I and  $0.0015^{+0.0034}_{-0.0012}$  events in SR IV, which is small compared with the nominal pre-  
713 diction shown in the first row of Table 18.

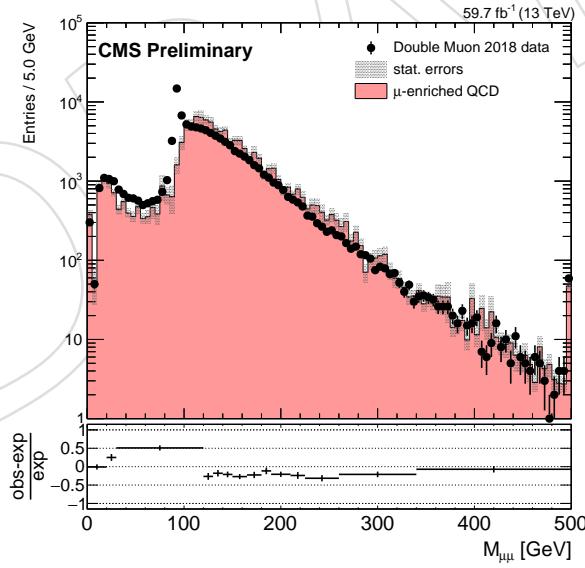


Figure 25: The dimuon invariant mass distribution in the  $\mu\mu$  channel with the muon isolation criterion inverted, for 2018 data and QCD simulation.

714 Thus, from these tests, we show that the heavy-flavor contribution in the SRs is small and  
715 already accounted for with the nominal background prediction.

Table 19: A closure test of the ABCD method in 2018 QCD simulation in the  $\mu\mu$  channel with the muon isolation criterion inverted. The estimates from the ABCD method, the actual yields in simulation, and the ratios of the actual to the estimated yields are shown.

Region	Estimated yield	Actual yield	Ratio of actual to estimate
SR I	$9500 \pm 1100$	$11000 \pm 1000$	$1.19^{+0.18}_{-0.17}$
SR II	$1740^{+310}_{-280}$	$2200^{+330}_{-290}$	$1.26^{+0.29}_{-0.26}$
SR III	$1450^{+280}_{-240}$	$1500^{+180}_{-160}$	$1.04^{+0.23}_{-0.21}$
SR IV	$265^{+62}_{-54}$	$268^{+61}_{-50}$	$1.01^{+0.33}_{-0.28}$

## 7.4 Low-mass SM hadrons

To estimate an upper limit on decays of low-mass SM hadrons in the signal region, we examine 2018 data and QCD simulation in the  $\mu\mu$  channel with both the muon isolation and the  $\Delta R$  requirements inverted. As can be seen in Fig. 26, this region is dominated by low-invariant mass  $\mu\mu$  events, with clear  $J/\psi$ ,  $\psi'$ , and  $\Upsilon$  mass peaks. This region also has many displaced events, especially in the  $J/\psi$  mass range. We take the ratio of SR to prompt control region events of SM hadrons that decay to leptons from data in this region.

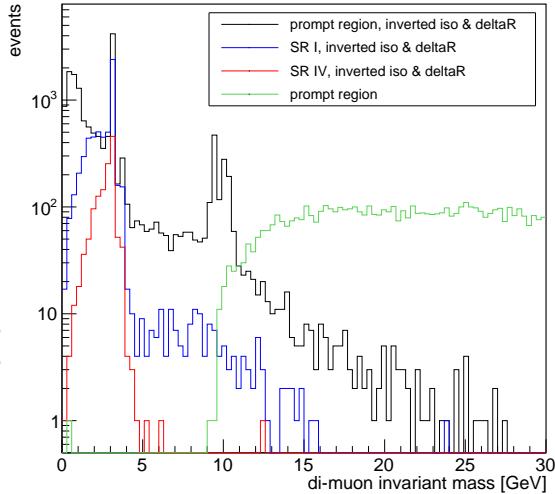


Figure 26: The dimuon invariant mass distribution in 2018 data in the  $\mu\mu$  channel, in the prompt control region (black), SR I (blue), SR IV (red), with the muon isolation and  $\Delta R$  criteria inverted. The nominal prompt control region is also shown in green.

However, the only QCD simulation event that survives the 2018  $\mu\mu$  preselection has a large invariant mass (300 GeV) and  $\Delta R$  (3), which is inconsistent with the low-mass SM hadron events that dominate the region with the inverted isolation and  $\Delta R$  criteria. Thus, we use the inverted isolation sample from the heavy-flavor background check to help with the normalization. In this sample, the ratio of events with  $\Delta R < 0.5$  to events with  $2.8 < \Delta R < 3.2$  is about 0.1. We find 0.2 QCD simulated events that pass the nominal preselection, and so we estimate there are 0.02 events from SM hadrons that decay to leptons. We estimate the SR contributions from this preselection normalization and from the ratio of SR to prompt control region events from the sample of SM hadrons that decay to leptons in data. We find this contribution is less than  $0.006^{+0.013}_{-0.005}$  events in SR I and less than  $0.001^{+0.002}_{-0.001}$  events in SR IV, which, if compared with the nominal prediction shown in the first row of Table 18, are respectively negligible and covered

734 by the 140% systematic uncertainty applied in this region.

## 735 8 Systematic uncertainties in the signal efficiency

### 736 8.1 Integrated luminosity

737 The integrated luminosities of the 2016, 2017, and 2018 data-taking periods are individually  
 738 known with uncertainties in the 1.2–2.5% range [47–49], which combined for the data set used  
 739 in this analysis has a total uncertainty of 1.8%, the improvement in precision reflecting the  
 740 uncorrelated time evolution of some systematic effects. The full table of 2016–2018 integrated  
 741 luminosity uncertainties are taken from [50], with the correlations specified therein.

### 742 8.2 Pileup

743 The simulation of pileup events assumes a total inelastic pp cross section of 69.2 mb, with an  
 744 associated uncertainty of 5% [51]. The systematic uncertainty arising as a result of the modeling  
 745 of pileup events is estimated by varying the cross section of the minimum bias events by 5%  
 746 when generating the target pileup distributions. The pileup weights are recomputed with these  
 747 new distributions and applied to the simulated events to obtain the variation in the yields in  
 748 the inclusive signal region. The average uncertainty is <1%. We treat these uncertainties as  
 749 100% correlated across the three years of data taking.

### 750 8.3 Tracking efficiency for displaced particles

751 The displaced tracking efficiency is measured in both data and simulation using cosmic rays,  
 752 as described below. In addition,  $K_S^0$  are used to check for any dependence on instantaneous  
 753 luminosity or run period, as described in Appendix A.

#### 754 8.3.1 Study with cosmic ray muons

755 The efficiency to reconstruct displaced, isolated, high  $p_T$  muons can be measured using cosmic  
 756 rays. The basic idea is to reconstruct cosmic rays in the muon chambers and then measure for  
 757 what fraction of these, a corresponding track is reconstructed in the tracker.

758 Two different kinds of cosmic ray dataset are available within CMS.

759 1. '/Cosmics/Run2016(7)(8)\*/\*' : Dedicated cosmic runs, reconstructed with dedi-  
 760 cated cosmic ray reconstruction algorithms only.

761 2. '/NoBPTX/Run2016(7)(8)\*/\*' : This contains all the cosmic rays from '/Cosmics/Run2016(7)  
 762 plus additional cosmic rays from *Parasitic cosmic runs*. The latter are taken during nor-  
 763 mal pp collision running, using a trigger which vetoes events in which colliding proton  
 764 bunches are present. They are reconstructed with standard pp-reconstruction algorithms  
 765 plus some dedicated cosmic ray reconstruction algorithms.

766 In both datasets, the strip tracker readout electronics is in 'deconvolution' mode, which is also  
 767 the standard mode used for pp data taking. (Historically, CMS used to take the dedicated  
 768 cosmic ray runs in 'peak' mode, but this is no longer the case). Deconvolution mode provides  
 769 good time resolution, but has the disadvantage that the tracker can be inefficient for cosmic  
 770 rays that arrive out of phase with the 40 MHz LHC clock.

771 For the measurement of the tracking efficiency, we choose the "NoBPTX" dataset, since they  
 772 include the pp-reconstruction sequences. As will be shown below, one can avoid the loss of  
 773 tracker efficiency expected in this data for out-of-time cosmic rays, by vetoing the latter using  
 774 the muon arrival time measured in the muon chambers.

775 Events were taken using the trigger '`HLT_L2Mu10_NoVertex_NoBPTX3BX_v*`', which be-  
 776 sides vetoing events with colliding pp bunches, requires the presence of a muon reconstructed  
 777 in the muon chambers with momentum component transverse to the beam-axis exceeding  
 778 10 GeV. This muon is reconstructed by the trigger without using any beam-spot constraint.  
 779 The 2016, 2017 and 2018 data was taken from the "07Aug17", "17Nov2017" and "17Sep2018"  
 780 processings of the AOD, respectively. For the 2016 data, only eras 'G' and 'H' were used, since  
 781 the earlier eras that year were affected by problems with the tracker readout chips.

782 No reliable 'JSON' file was available to select good data taking periods in the cosmic runs,  
 783 when the detector was working reliably. We therefore produced the JSON file to do this our-  
 784 selves, following the prescription in [52]. The following key software flags were used by the  
 785 script producing this JSON file:

- 786 • `runClassName = 'Cosmics16(17)(18)'`: Selects events with cosmic trigger  
 787 timing configured suitably for taking cosmic rays. (This is important, since the alter-  
 788 native choice of `'Collisions16(17)(18)'` gives very poor results).
- 789 • `Dataset = /PromptReco/%/%`: Use the offline run registry to access the data  
 790 quality information, as opposed to the inferior online one.
- 791 • `QFLAGS = L1t:good,HLT:good,Pix:good,Strip:good,Dt:good,Rpc:good,`  
 792 `Csc:good,Track::good`: Selects data in which the trigger, tracker and muon  
 793 chambers are known to be functioning well, and track reconstruction is behaving  
 794 correctly.
- 795 • `DSSTATE = COMPLETED`: Requires run registry data quality information to have  
 796 been checked by the DQM shift crew and by the DPG expert.
- 797 • `BField_thr = 3.8`: Requires the magnetic field to be on and have a value close  
 798 to its nominal one. This was necessary to avoid runs where the magnet had been  
 799 malfunctioning.

800 Simulated cosmic ray events were produced using the CMSCGEN generator, simulating cos-  
 801 mic rays of  $p_T > 20$  GeV. Only those passing within 40 cm of the beam-axis and doing so no  
 802 more than 80 cm from the center of CMS along the z-axis were retained. The time window in  
 803 which the cosmics were generated was 30 ns wide, and centered on a time chosen (with some  
 804 trial and error) to correspond to that in which the tracker readout efficiency was maximal. Sim-  
 805 ulated of the detector response for 2016, 2017 and 2018 samples was performed with GEANT  
 806 within CMSSW 7\_1\_25, 9\_3\_13, and 10\_2\_23, respectively. The corresponding event reconstruc-  
 807 tion algorithm was run within CMSSW 8\_0\_25, 9\_4\_10 and 10\_2\_5, respectively. (Note for experts:  
 808 Both detector simulation and event reconstruction used configuration files produced with the  
 809 cmsDriver script using `--scenario pp` instead of `--scenario cosmics`, since this corre-  
 810 sponds to the data taken during parasitic cosmic runs.)

811 Within the data and simulated events, cosmic rays are reconstructed in the tracker using the  
 812 same track reconstruction algorithm used in pp data taking (yielding the '`generalTrack`'  
 813 collection). As this algorithm is designed to reconstruct tracks produced near the beam-line,  
 814 it typically reconstructs each cosmic ray as two separate tracks, one moving upwards and one  
 815 moving downwards with respect to the cosmic rays point of closest approach to the beam-  
 816 line. The cosmic rays are also reconstructed in the muon chambers alone, using dedicated

algorithms designed to give good efficiency for highly displaced muons. Two such algorithms are run: a ‘2-leg’ algorithm reconstructs each cosmic ray as two separate muons, one found in the top-half of CMS and the other in the bottom-half; and a ‘1-leg’ algorithm reconstructs the cosmic rays as a single muon transversing the whole detector. The 1-leg muons have superior helix parameter resolution to the 2-leg ones, owing to the longer lever arm over which they measure the muon trajectory.

The event selection requires a 1-leg muon with  $p_T > 20 \text{ GeV}$  and at least 50 hits in the muon chambers. This is required to match two 2-leg muons within a tolerance in azimuthal angle of 0.3 radians. This requirement rejects a few percent of cosmic rays candidates that either failed to traverse CMS or were not genuine. The RMS angular difference in azimuthal angle between the 1-leg and associated 2-leg muons is  $\approx 0.04$  radians, so well below the allowed tolerance. This event selection uses a lower  $p_T$  cut than is employed in the physics analysis, in order to increase the usable number of cosmic rays. The tracking efficiency does not depend significantly on  $p_T$  in the range considered here.

The tracking efficiency is determined by measuring the fraction of these cosmic rays which are associated within an azimuthal angle of 0.2 radians with a track reconstructed in the tracker, where only tracks of  $p_T > 15 \text{ GeV}$  are considered. In addition, tracks are required to have at least one pixel hit to mirror the track selection used in the physics analysis. Two such efficiencies can be defined: a ‘downward’ and an ‘upward’ going efficiency, where the muon is matched to downward-going or upward-going tracks, respectively. (In the case of upward-going tracks, the track momentum vector must be flipped before matching it to the muon). As will be seen later, these efficiencies are almost, but not exactly identical. The RMS angular difference in azimuthal angle between the 1-leg muons and associated tracks is  $\approx 0.005$  radians, so well below the allowed tolerance. The resolution in  $q/p_T$  is  $\approx 0.002/\text{GeV}$ .

One additional tracking efficiency: the ‘perpendicular’ efficiency can also be defined, which measures the fraction of cosmic muons that have a reconstructed track moving perpendicular to them, within a tolerance of 0.2 radians. This efficiency provides a measurement of the probability that residual tracks from pp collisions by chance are matched with a cosmic muon, and so bias the tracking efficiency measurement. No such bias is seen.

The arrival time of the cosmic ray at its point of closest approach to the beam-line can be measured using the the muon chambers. Both the 1-leg and 2-leg muons provide an estimate of this time. It was observed that taking the average time measurement of the two 2-leg muons gives a much more precise estimate of the time (resolution  $\approx 5 \text{ ns}$ ) than does taking the time measurement of the 1-leg muon (resolution  $\approx 50 \text{ ns}$ ), so the former was used. This may be because the timing estimates are produced assuming that the muon moves either away from the center of CMS or towards it. For the 1-leg muons, this assumption is false, since the muon initially moves towards the center of CMS and then away from it.

The sample of 1-leg muons with  $|d_0| < 8 \text{ cm}$  and  $|z_0| < 20 \text{ cm}$  is used to study the dependence of the tracking efficiency upon the cosmic ray arrival time. Figure 27 shows the distribution of the measured arrival time. In the case of the simulated events, the distribution is determined by the time window specified when generating the events (described above), so it is not surprising that this differs from the data. Figure 28 shows the measured downward tracking efficiency versus the arrival time. The efficiency shows a clear peak, with width of approximately 15 ns, which is roughly what one would expect based on the known time resolution of the strip tracker readout electronics in deconvolution mode, although the time resolution of the pixel tracker and muon chambers may also contribute. Based on these plots, the following selection requirements are imposed on the measured time  $t_{\mu\text{on}}$  when measuring the downward

864 tracking efficiency:

- 865 • ‘Downward’ tracking efficiency in 2016-2017 data:  $-13 \text{ ns} < t_{\mu\text{on}} < -7 \text{ ns}$ .
  - 866 • ‘Downward’ tracking efficiency in 2018 data:  $-8 \text{ ns} < t_{\mu\text{on}} < -2 \text{ ns}$ .
  - 867 • ‘Downward’ tracking efficiency in 2016-2017 simulation:  $-38 \text{ ns} < t_{\mu\text{on}} < -32 \text{ ns}$ .
  - 868 • ‘Downward’ tracking efficiency in 2018 simulation:  $-40 \text{ ns} < t_{\mu\text{on}} < -34 \text{ ns}$ .
- 869 Similar results are obtained for the ‘upward’ efficiency, except that the optimum time windows  
870 are shifted 5 ns later. This was expected, since cosmic rays will reach the top half of the tracker  
871 approximately 5 ns earlier than the bottom half.

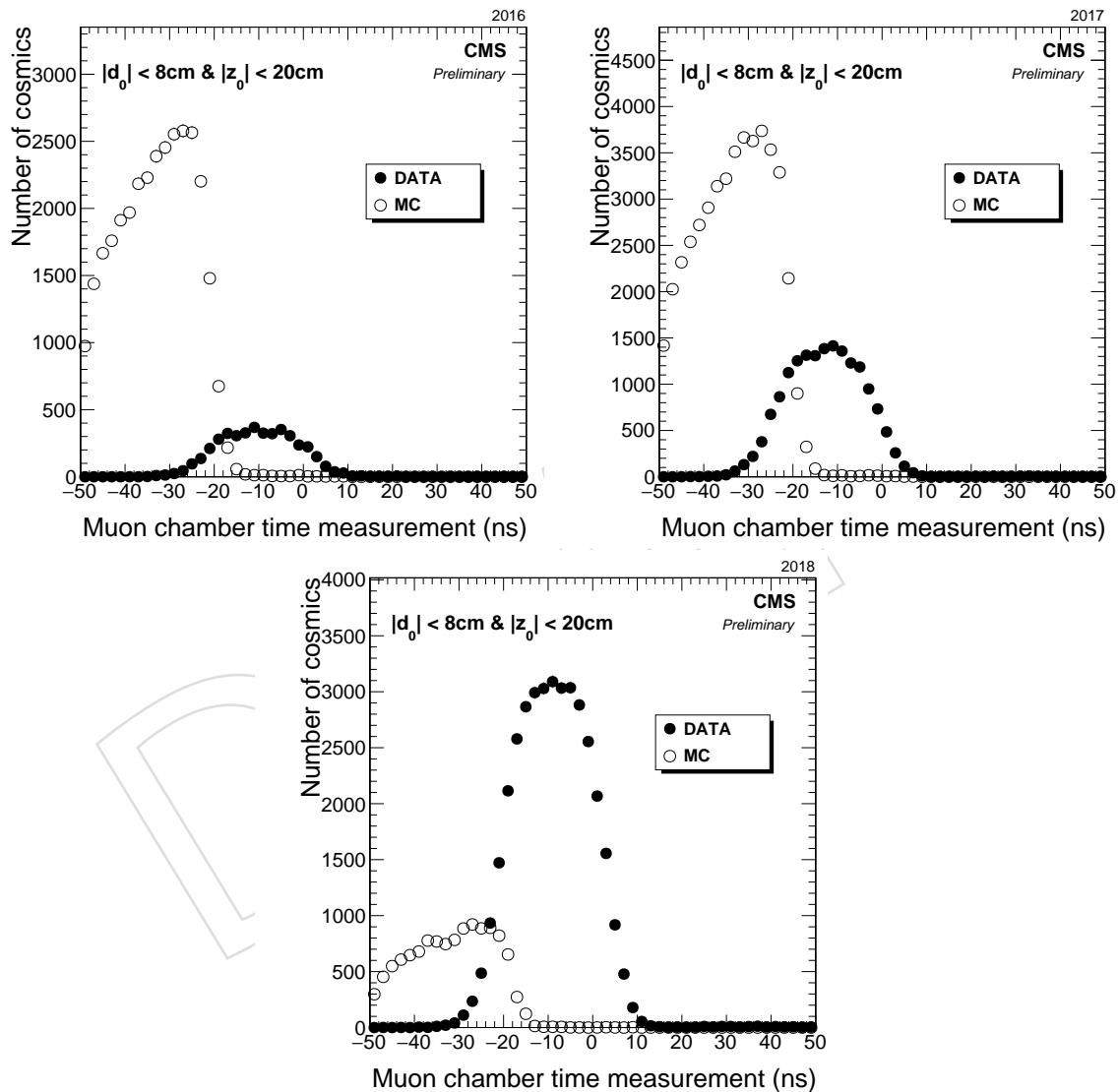


Figure 27: Distribution of the arrival time of cosmic rays at their point of closest approach to the beam-line, measured with the muon chambers, in 2016 (top-left), 2017 (top-right) and 2018 (below), for both data and simulated events. Only cosmic rays passing close to the center of CMS are used.

- 872 The time window needed in the simulated events is 30 ns earlier than in the data. This suggests  
873 that CMS has not adequately tuned the timing of the tracker or muon chamber digitization  
874 in the simulation. However, this difference does not prevent one selecting the above timing

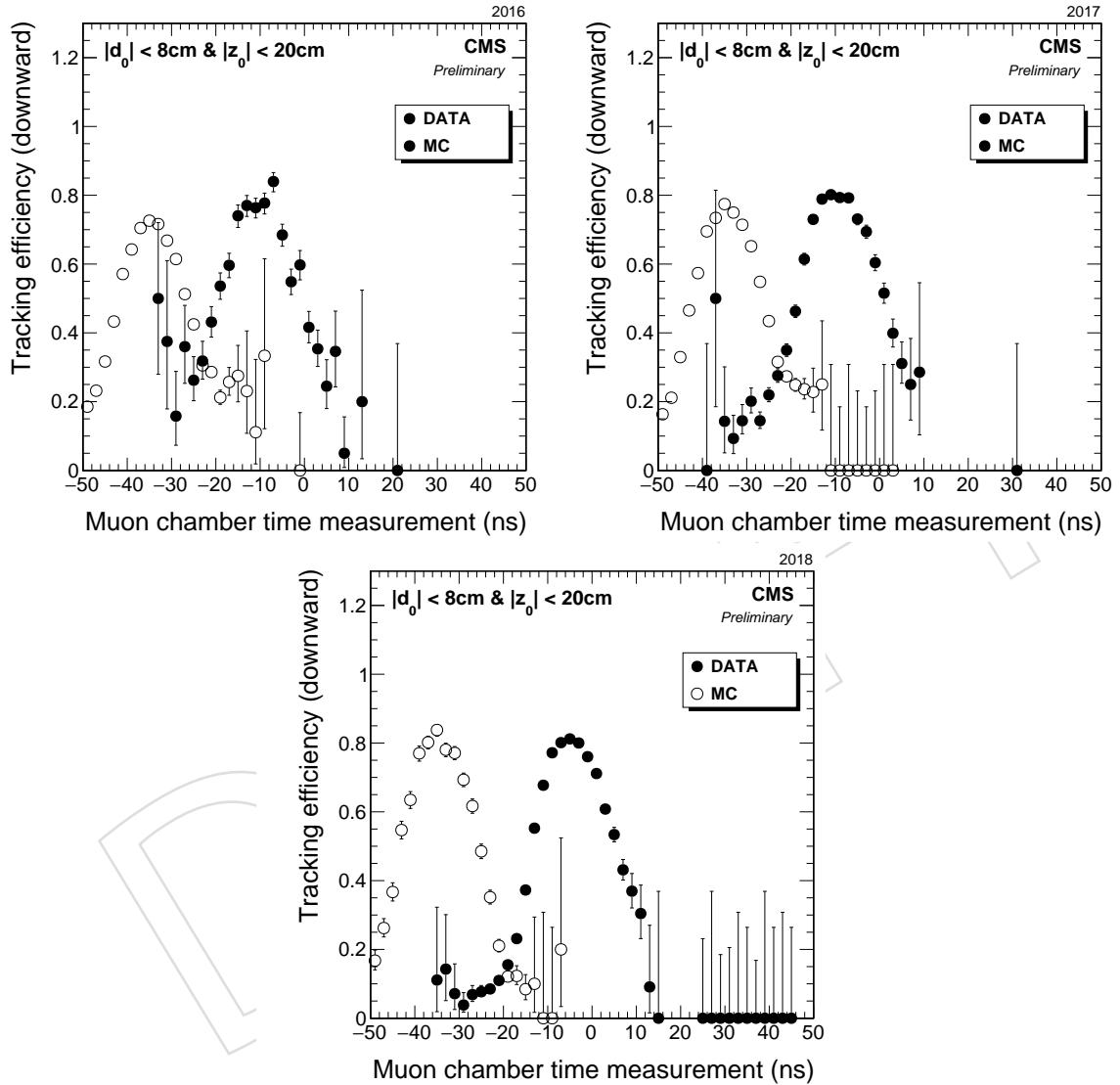


Figure 28: Measured downward tracking efficiency versus arrival time of the cosmic ray in 2016 (top-left), 2017 (top-right) and 2018 (below), where only cosmic rays with  $|d_0| < 8\text{ cm}$  and  $|z_0| < 20\text{ cm}$  are considered.

875 windows where the tracking efficiency is high, so it was not considered necessary to fix it.

876 The tracker readout electronics are programmed to read the hit signals a few nanoseconds later  
 877 from the outer layers of the tracker than from the inner layers, to take into account that particles  
 878 moving away from the beam-line reach the outer layers later. When cosmic rays traverse the  
 879 bottom half of the tracker, they also pass through the inner tracker layers before the outer ones,  
 880 so this should be approximately correct. In contrast, when cosmic rays traverse the top half of  
 881 the tracker, they traverse the outer tracker layers before the inner ones. The tracker timing can  
 882 therefore never be perfect for the tracks used in the upward tracking efficiency measurement,  
 883 so one might predict that this would be somewhat lower than the downward one. However, the  
 884 effect is only a few percent. Nonetheless, we will use the measurements of downward tracking  
 885 efficiency as our definitive measurements of the displaced tracking efficiency, and only use the  
 886 upward efficiency as a check.

887 1-leg muons with  $|d_0| < 8\text{ cm}$  and  $|z_0| < 20\text{ cm}$  can also be used to additional checks. In  
 888 particular, one can use them to measure the tracking efficiency as a function of run number.  
 889 This allows one to check if any bad runs have escaped the JSON good run selection mentioned  
 890 above. No such bad runs were discovered.

891 Relaxing the 1-leg muon impact parameter requirements, and imposing the above timing win-  
 892 dows, the tracking efficiency is measured as a function of  $|d_0|$  (for  $|z_0| < 20\text{ cm}$ ) and  $|z_0|$  (for  
 893  $|d_0| < 1.2\text{ cm}$ ). The results are shown for the downward tracking efficiency in Fig. 29, compar-  
 894 ing data and simulation. Results for the upward efficiency, not shown, are just a few percent  
 895 lower.

896 It should be noted that if the requirement that tracks have an associated pixel hit were removed,  
 897 then the tracking efficiency would remain non-zero out to  $|d_0| \approx 30\text{ cm}$  and  $|z_0| \approx 50\text{ cm}$ .

898 To estimate the systematic uncertainty arising from the displaced tracking efficiency, a simu-  
 899 lated signal sample with a top squark mass of  $1800\text{ GeV}$  and lifetime  $c\tau = 100\text{ cm}$  was con-  
 900 sidered. This sample produces leptons with the largest impact parameters of any of those  
 901 considered in this paper, so represents the most challenging scenario for the displaced track  
 902 reconstruction. Only those events in which both top squarks decayed within the volume of the  
 903 pixel detector were retained, since in any events failing this requirement, both leptons would  
 904 not have pixel hits and so would not be selected by the physics analysis. The  $|d_0|$  and  $|z_0|$   
 905 impact parameters of both leptons in this subset of signal events were then noted. Next, a  
 906 2-dimensional plot of tracking efficiency  $\epsilon(|d_0|, |z_0|)$  as a function of the transverse and lon-  
 907 gitudinal impact parameters was produced from the cosmic rays in both data and simulated  
 908 cosmic ray events. Using this plot, the mean efficiency to reconstruct both lepton tracks in  
 909 the simulated signal events was evaluated as  $\frac{1}{N} \sum_i \epsilon(|d_0|_i^{(1)}, |z_0|_i^{(1)}) \epsilon(|d_0|_i^{(2)}, |z_0|_i^{(2)})$ , where the  
 910 sum extends over the  $N$  events in the signal sample, and the superscripts “(1)” and “(2)” denote  
 911 the two leptons in each event. For the 2018 [2017] (2016) run, the efficiency to reconstruct the  
 912 dilepton tracks evaluates to  $56.1 \pm 0.7\%$  [ $55.3 \pm 1.1\%$ ] ( $57.5 \pm 2.1\%$ ) when taking the efficiency  
 913 from data cosmic rays, and to  $57.5 \pm 1.1\%$  [ $52.3 \pm 0.7\%$ ] ( $50.3 \pm 1.0\%$ ) when taking it from sim-  
 914 ulated cosmic rays. From the ratio of these two efficiencies, the relative systematic uncertainty  
 915 in the efficiency to reconstruct both dilepton tracks is estimated to be  $2.4 \pm 2.2\%$  [ $5.8 \pm 2.3\%$ ]  
 916 ( $14.1 \pm 4.3\%$ ). (In the ratio, the contribution to the statistical uncertainty arising from the finite  
 917 number of simulated signal events partially cancels). The 2017 and 2018 systematic uncertain-  
 918 ties are treated as fully correlated, while the 2016 uncertainty is treated as uncorrelated with  
 919 the 2017 and 2018 uncertainties, since the pixel detector was upgraded after 2016 data taking.

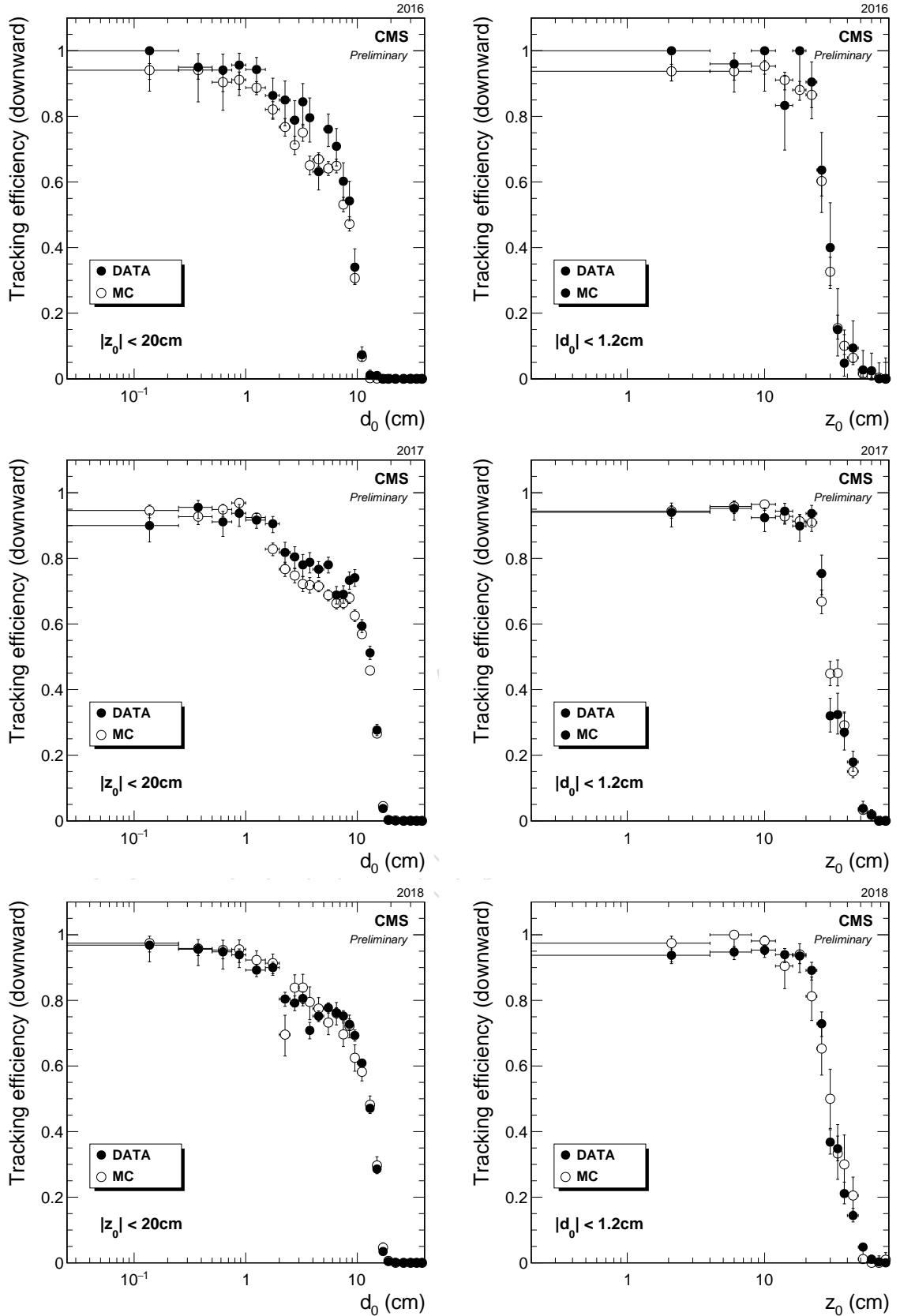


Figure 29: Measured ‘downward’ tracking efficiency versus impact parameter in 2016 (top), 2017 (middle) and 2018 (bottom): specifically  $|d_0|$  (for  $|z_0| < 20$  cm) in the left-hand plots, and  $|z_0|$  (for  $|d_0| < 1.2$  cm) in the right-hand plots. Both data and simulation are shown.

## 920 8.4 Trigger efficiency

- 921 The trigger efficiency systematic uncertainty is given by the uncertainty in the measured trigger  
 922 efficiency scale factors (see Section 4.4). These uncertainties are about 1% for the  $e\mu$  and  $\mu\mu$   
 923 channels and 10–19% for the ee channel.
- 924 In addition, we have also studied the trigger efficiency in signal as a function of  $|d_0|$ , as shown  
 925 in Fig. 30, for events in the trigger  $p_T$  plateau. To cover the change observed in the muon trigger  
 926 efficiency over the full  $|d_0|$  range, we assign an additional 20% uncertainty.

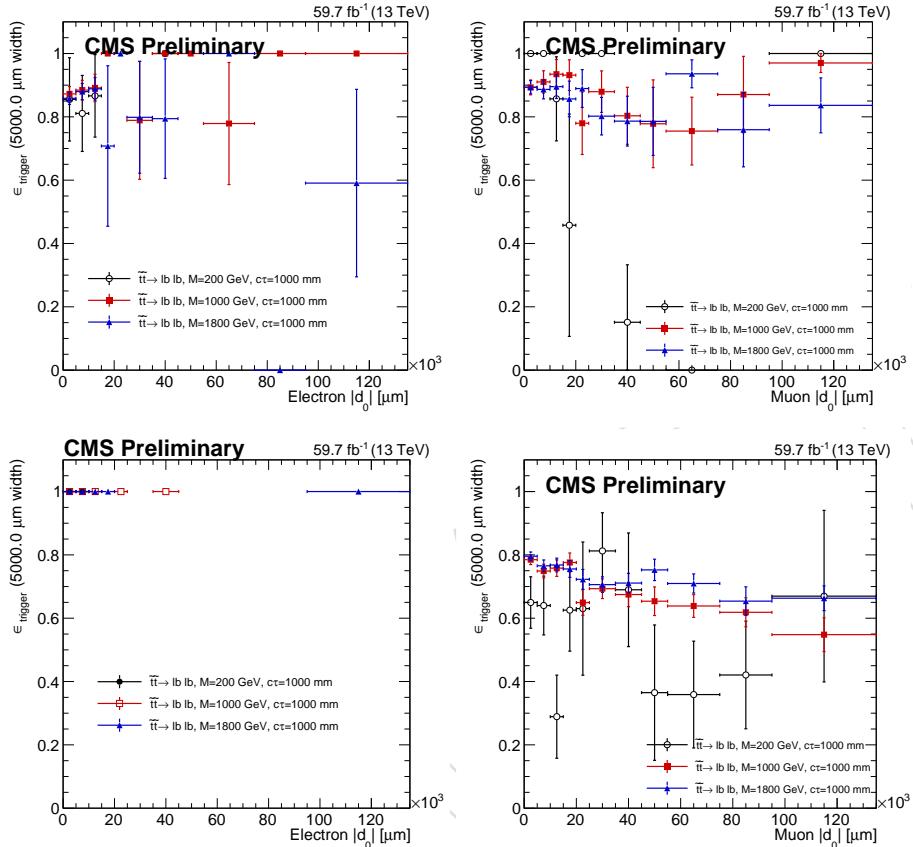


Figure 30: Trigger efficiency as a function of lepton  $|d_0|$ , for the  $e\mu$  channel (upper row), the ee channel (lower left), and the  $\mu\mu$  channel (lower right) in 2018 signal, for events in the trigger  $p_T$  plateau.

- 927 We treat the trigger efficiency uncertainties as 100% correlated across the three years of data  
 928 taking.

## 929 8.5 Lepton ID and isolation

- 930 To find the systematic uncertainty associated with the corrections to the lepton ID and isolation,  
 931 we fluctuate the lepton scale factors up and down by their uncertainty and observe the change  
 932 in the event yields in the inclusive signal region. The average uncertainty for electrons is about  
 933 3% in the  $e\mu$  channel and about 7% in the ee channel, while the average uncertainty for muons  
 934 is < 1%. We treat these uncertainties as 100% correlated across the three years of data taking.

### 8.6 Muon pixel hit efficiency

One selection within the muon ID could have some dependence with respect to  $|d_0|$ , namely, the requirement that the muons have at least one pixel hit. Figure 31 shows the efficiency of this requirement in cosmic simulation and NoBPTX data as a function of muon  $|d_0|$ . For events in the denominator of these plots, we require that at least 2 global, PF muons have  $|\eta| < 1.0$ ,  $p_T > 25 \text{ GeV}$ , no displaced vertices in the tracker material,  $|z_0| < 15 \text{ cm}$ , and that they pass all the tight ID criteria except the pixel hit requirement. We also require the muons to be separated by  $\Delta R > 0.2$ . The same requirements are made on the numerator, except in addition, we require that the muons have at least one pixel hit. Using this plot, the mean efficiency to identify the muons in the simulated signal events is evaluated in the same way as done for the displaced tracking efficiency systematic uncertainty. That is, we find the efficiency to identify muons that pass the pixel hit requirement in cosmic simulation and in NoBPTX data, and using the ratio of these two efficiencies, we derive the relative systematic uncertainty in the signal. The average uncertainty is about 16% (32%) in the  $e\mu$  ( $\mu\mu$ ) channel. The 2017 and 2018 systematic uncertainties are treated as fully correlated, while the 2016 uncertainty is treated as uncorrelated with the 2017 and 2018 uncertainties, since the pixel detector was upgraded after 2016 data taking.

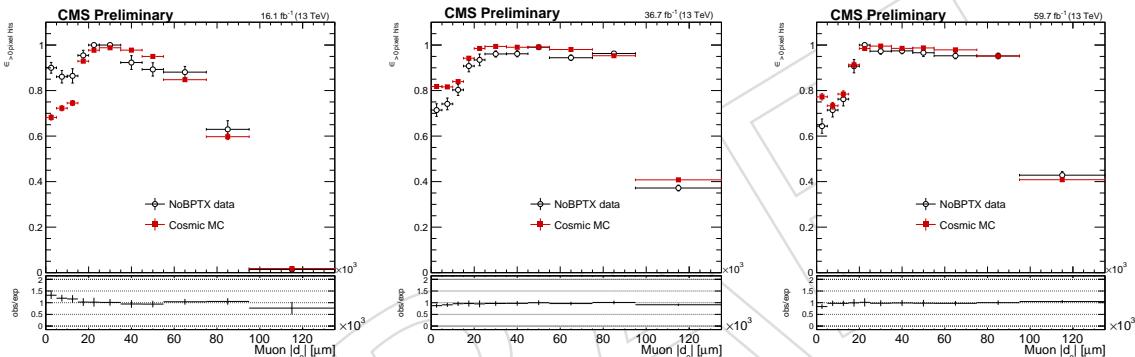


Figure 31: The pixel hit efficiency as a function of muon  $|d_0|$ , for cosmic simulation and NoBPTX data in 2016 (left), 2017 (middle), and 2018 (right) conditions.

### 8.7 Lepton $|d_0|$ smearing

To find the systematic uncertainty associated with the corrections to the lepton  $|d_0|$ , we fluctuate the lepton  $|d_0|$  corrections up and down by their uncertainty and observe the change in the event yields in the inclusive signal region. The average uncertainty is  $< 1\%$ . We treat these uncertainties as 100% correlated in 2017 and 2018. There was no  $|d_0|$  correction needed for 2016 simulation.

### 8.8 Summary of systematic uncertainties in the signal efficiency

The systematic uncertainties in the signal efficiency are summarized in Table 20.

Table 20: Systematic uncertainties in the signal efficiency, for all three years and the three channels. For many sources of uncertainty, a range indicating the 68% CL of the spread is given, followed by the mean. Uncertainties in the same row are treated as correlated among the years of data taking, except for the displaced tracking and muon pixel hit efficiencies, where the 2016 uncertainty is treated as uncorrelated with the 2017 and 2018 uncertainties.

Systematic uncertainty	2016	2017	2018
<i>Integrated luminosity</i>	1.8% (2016–2018)		
<i>Pileup</i>			
- $e\mu$ channel	0.2–0.7% / 0.5%	0.4–0.9% / 0.6%	0.3–0.7% / 0.5%
- ee channel	0.1–0.9% / 0.5%	0.5–1.2% / 0.9%	0.4–1.2% / 0.8%
- $\mu\mu$ channel	0–0.4% / 0.2%	0–0.3% / 0.1%	0–0.4% / 0.2%
<i>Trigger efficiency</i>			
- $e\mu$ channel, electrons	1.6%	1.3%	1.2%
- $e\mu$ channel, muons	1.6%	1.4%	1.2%
- ee channel	10%	13%	19%
- $\mu\mu$ channel	1.2%	1.0%	1.1%
<i>Muon trigger efficiency at large <math> d_0 </math></i>			
- $e\mu$ channel, muons	20%	20%	20%
- $\mu\mu$ channel	20%	20%	20%
<i>Displaced tracking efficiency</i>	14%	5.8%	2.4%
<i>Muon pixel hit efficiency</i>			
- $e\mu$ channel, muons	30–32% / 32%	10–13% / 12%	12–17% / 16%
- $\mu\mu$ channel	70–74% / 73%	20–24% / 23%	27–31% / 30%
<i>Lepton identification and isolation</i>			
- $e\mu$ channel, electrons	1.2–1.3% / 1.2%	3.1–4.0% / 3.6%	3.1–3.9% / 3.5%
- $e\mu$ channel, muons	0.04–0.05% / 0.05%	0.06–0.07% / 0.07%	0.05–0.06% / 0.06%
- ee channel	2.3–2.5% / 2.4%	6.4–7.9% / 7.2%	6.3–7.7% / 7.0%
- $\mu\mu$ channel	0.09–0.10% / 0.10%	0.14–0.15% / 0.14%	0.11–0.13% / 0.12%
<i>Lepton <math> d_0 </math> correction</i>			
- $e\mu$ channel, electrons	—	0–0.003% / 0.001%	0–0.003% / 0.001%
- $e\mu$ channel, muons	—	0–0.005% / 0.003%	0–0.002% / 0.001%
- ee channel	—	0–0.27% / 0.11%	0–0.27% / 0.11%
- $\mu\mu$ channel	—	0–0.27% / 0.11%	0–0.27% / 0.11%

## 9 Results

Table 21 shows the expected number of background events and the observed data in each  $p_T$  bin and each SR, for each channel, and Fig. 32 shows the same with a representative signal yield overlaid. We split SR I into two bins. In the  $e\mu$  and  $\mu\mu$  channels, these bins are in the leading muon  $p_T$ , and in the ee channel, these bins are in the leading electron  $p_T$ . The  $p_T$  bins are chosen such that the bin with higher  $p_T$  contains  $<1$  background event, which maximizes the sensitivity to small lifetimes. The observed number of events are consistent with the predicted amount of background.

Figure 33 shows two-dimensional plots of the  $|d_0|$  variables in data in each channel, for events that pass the preselection, and Fig. 34 shows the same but for data events in the inclusive SR. Figure 35 shows the  $|d_0|$  variables for data and a representative signal point.

Table 21: The number of estimated background and observed events in each channel and SR. For each estimate, the total uncertainty is given.

	SR I $p_T$ boundary [GeV]	SR I, low $p_T$ bin	SR I, high $p_T$ bin	SR II	SR III	SR IV
<i>2016 e<math>\mu</math></i>						
- estimated	90	$3.8^{+4.8}_{-3.8}$	$0.41^{+0.53}_{-0.41}$	$0.09^{+0.12}_{-0.09}$	$0.15 \pm 0.15$	$0.003^{+0.004}_{-0.003}$
- observed	90	8	1	0	0	0
<i>2017+2018 e<math>\mu</math></i>						
- estimated	140	$38 \pm 13$	$0.75^{+0.41}_{-0.34}$	$0.23^{+0.27}_{-0.23}$	$0.71^{+0.76}_{-0.71}$	$0.01^{+0.02}_{-0.01}$
- observed	140	28	3	0	1	0
<i>2016 ee</i>						
- estimated	300	$18 \pm 11$	$0.22^{+0.17}_{-0.16}$	$0.51^{+1.02}_{-0.51}$	$0.43^{+0.85}_{-0.43}$	$0.01^{+0.02}_{-0.01}$
- observed	300	40	0	0	1	0
<i>2017+2018 ee</i>						
- estimated	400	$62^{+18}_{-17}$	$0.85^{+0.33}_{-0.35}$	$2.8 \pm 1.1$	$3.6 \pm 1.4$	$0.24^{+0.10}_{-0.09}$
- observed	400	48	0	1	4	0
<i>2016 <math>\mu\mu</math></i>						
- estimated	100	$7.4 \pm 3.0$	$0.25 \pm 0.11$	$0.17 \pm 0.11$	$0.19 \pm 0.12$	$0.01 \pm 0.01$
- observed	100	15	0	0	1	0
<i>2017+2018 <math>\mu\mu</math></i>						
- estimated	100	$3.5 \pm 1.5$	$0.69 \pm 0.31$	$0.08^{+0.12}_{-0.08}$	$0.14^{+0.19}_{-0.14}$	$0.01^{+0.02}_{-0.01}$
- observed	100	1	1	1	1	0

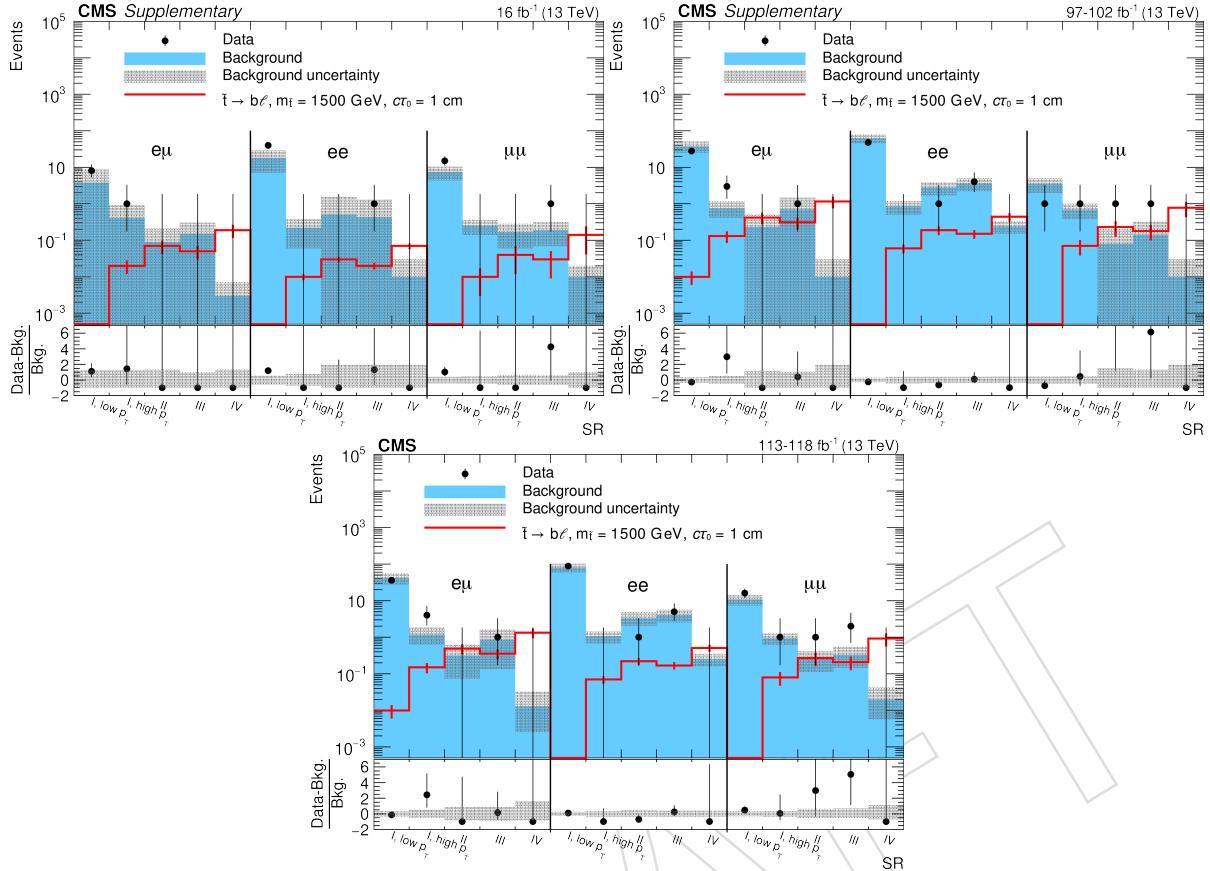


Figure 32: The number of estimated background and observed events in each channel and SR, with a representative signal overlaid, for 2016 (upper left), 2017–2018 (upper right), and the full Run 2 luminosity (lower). For each background estimate and signal yield, the total uncertainty is given.

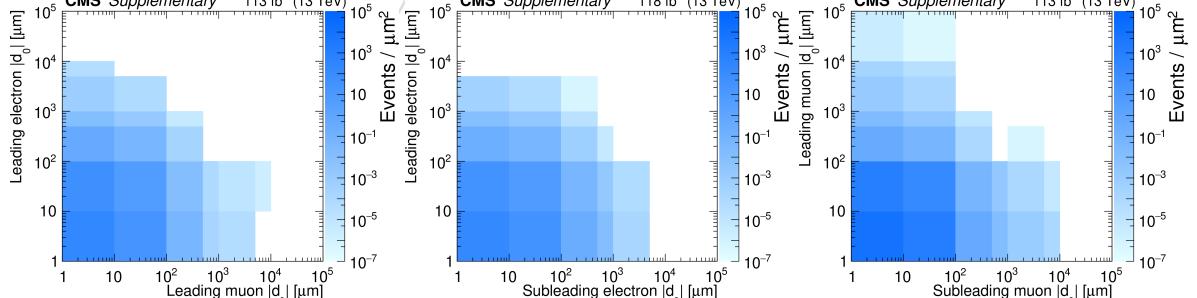


Figure 33: Two-dimensional distributions of  $|d_0^a|$  and  $|d_0^b|$ , for the events in data that pass the  $e\mu$  (left),  $ee$  (middle), and  $\mu\mu$  (right) preselection. In each  $|d_0|$ - $|d_0|$  bin, the number of events divided by the bin area is plotted. The inclusive signal region covers the region between 100  $\mu\text{m}$  and 10 cm in each  $|d_0|$  variable shown.

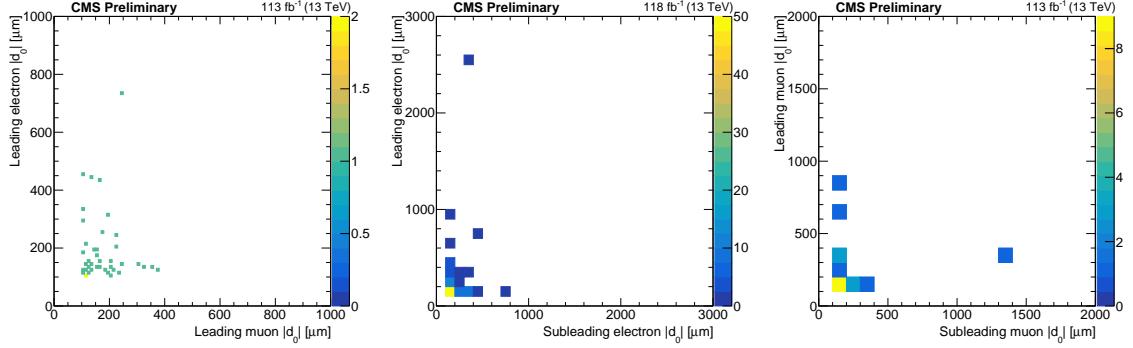


Figure 34: Two-dimensional distributions of  $|d_0^a|$  and  $|d_0^b|$ , for data events in the inclusive SR in the  $e\mu$  (left),  $ee$  (middle), and  $\mu\mu$  (right) channels.

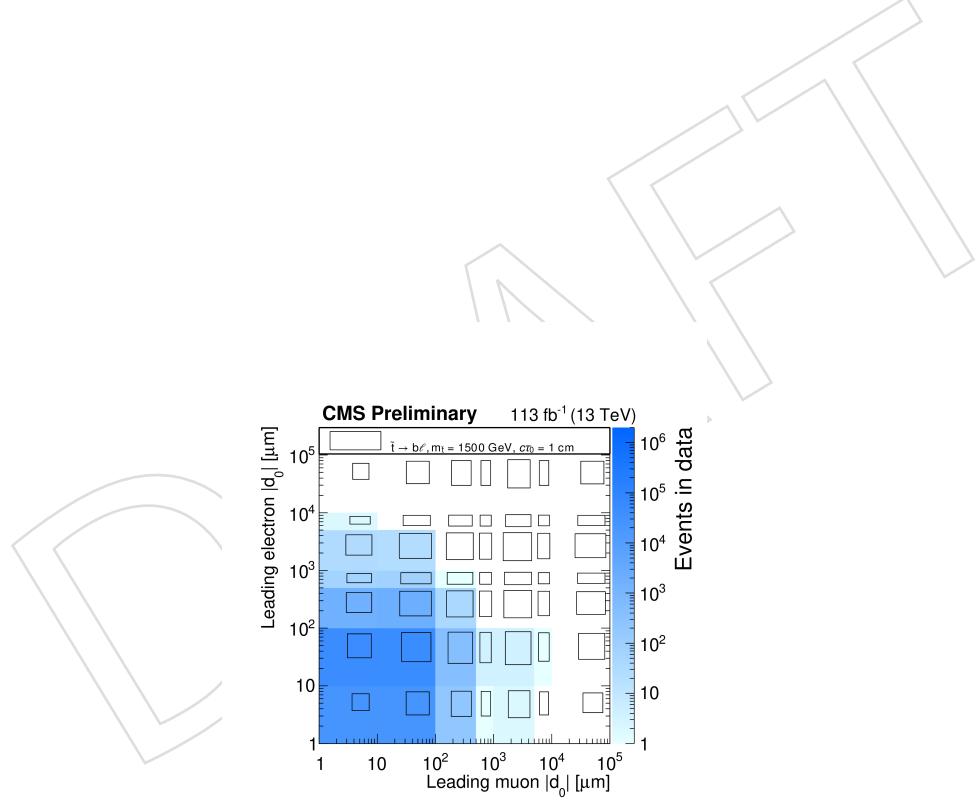


Figure 35: The two-dimensional distribution of the leading electron and leading muon  $|d_0|$ , for the events in data (colors) and signal (black boxes) that pass the  $e\mu$  preselection. The size of the black boxes are proportional to the size of the bin content. If a  $|d_0|$  value is less than unity, it is set to unity in order to plot in log scale. The inclusive signal region covers the region between  $100 \mu\text{m}$  and  $10 \text{ cm}$  in each  $|d_0|$  variable shown.

971 **9.1 Observed events**

972 We have looked in some detail at the observed events in the inclusive SR for each channel and  
 973 year. Plots showing background simulation and the data in the inclusive SR for different kinematic  
 974 variables are shown in Appendix D.7. We record here the summary of our observations.

975 In general, the unblinded data appear to be SM events from the pp collision. In the  $e\mu$  channel,  
 976 the SR events tend to have several jets and often have significant MET. Many events have muon  
 977  $\phi$  values such that the muon system hits are all near the edges of detector sections, or muon  
 978  $\eta$  values such that the muon is near the barrel/endcap transition in the muon system. There  
 979 are also a few events in which the electron and/or muon are associated with different PV than  
 980 their associated track.

981 In the ee channel, the majority of SR events contain at least one electron with  $|eta| > 1.1$ ,  
 982 indicating a high probability that their  $|d_0|$  values are mismeasured. Most events from all three  
 983 years fall into one of three categories:

- 984 1. Events with two electrons that appear to be from a boosted Z boson, with an invariant  
 985 mass between 80 and 100 GeV, opposite one or two jets
- 986 2. Events with two electrons back-to-back in  $\phi$  with an invariant mass greater than 100 GeV  
 987 and MET usually between 10 and 40 GeV
- 988 3. Events that are similar to type 2 but with at least one jet and frequently, but not always,  
 989 MET between 70 and 110 GeV

990 In the  $\mu\mu$  channel, many events have an invariant mass consistent with the mass of the Z boson  
 991 and MET less than about 60 GeV. Most of the events found in 2017 and 2018 have an invariant  
 992 mass higher than the Z boson mass and could be  $t\bar{t}$  events. Out of the events found in  
 993 2016, eight of them have a muon  $\phi$  value of about  $\pm\pi/2$ . Most of these eight events appear  
 994 in the 2016 low  $p_T$  bin of SR I, and they all have an invariant mass consistent with a Z boson.  
 995 Because of this fact and because of the  $\cos(\alpha)$  and timing distribution of these muons, there is  
 996 no indication that they are cosmic-ray muons. 13 of the 16 muons in these eight events have  
 997 only 1 valid pixel hit. For these 13 muons, 5 muons have this pixel hit at layer 1, 3 muons have  
 998 it at layer 2, and 5 muons have it at layer 3. From event displays of these eight events, we have  
 999 found that the inner track of the muons often passes between or at the edge of pixel modules,  
 1000 and so they may be poorly reconstructed and have few pixel hits because of this.

1001 **9.2 Limits**

1002 The data show no excess over background, and we set upper limits on the product of the signal  
 1003 production cross section ( $\sigma$ ) and branching fraction ( $\mathcal{B}$ ) using the HybridNew statistical  
 1004 method of the “Combine” tool, which was developed by the CMS Higgs working group [53–  
 1005 56]. The ABCD method is implemented in Combine, which has the advantage that all signal  
 1006 contamination in the control regions is automatically accounted for. We perform a simultaneous  
 1007 counting experiment in each signal region bin for most of the interpretations we consider.  
 1008 However, the ee and  $\mu\mu$  channels are fit individually to calculate limits on GMSB models with  
 1009 a  $\tilde{e}$  or  $\tilde{\mu}$  NLSP. Figure 36 shows the 95% confidence level (C.L.) upper limits on the top squark  
 1010 proper decay length as a function of its mass, for each of the three channels.

1011 In the high  $p_T$  SR I bin, which is the most sensitive bin for large top squark masses and small  
 1012 lifetimes, the  $e\mu$  channel has the largest signal yield, relative to the other two channels. This is  
 1013 because there are twice as many chances to have one electron and one muon, since the stops

decay to each lepton flavor with equal probability. In this bin, the ee and  $\mu\mu$  channel signal yields are similar. In SR IV, which is the most sensitive bin for large top squark masses and large lifetimes, the  $\mu\mu$  channel has the largest signal yield for  $c\tau \gtrsim 10$  cm, relative to the other two channels. This is because the muon reconstruction is better than that of electrons, so the  $|d_0|$  values of muons have less uncertainty than those of electrons. In this bin, the  $e\mu$  and  $\mu\mu$  channels have similar amounts of background, and the ee channel has the smallest signal yield out of the three channels, for  $c\tau \gtrsim 10$  cm. Therefore, for large top squark masses, the  $e\mu$  channel is the most sensitive for  $c\tau \lesssim 10$  cm, the  $\mu\mu$  channel is the most sensitive for  $c\tau \gtrsim 10$  cm, and the ee channel is the least sensitive of the three for  $c\tau \gtrsim 10$  cm.

Figure 37 shows the 95% C.L. upper limits on the top squarks for the three channels combined. The top squark limits assume  $\mathcal{B}(\tilde{t} \rightarrow bl) = \mathcal{B}(\tilde{t} \rightarrow dl) = 100\%$ , and each l has an equal probability of being an electron, a muon, or a tau lepton.

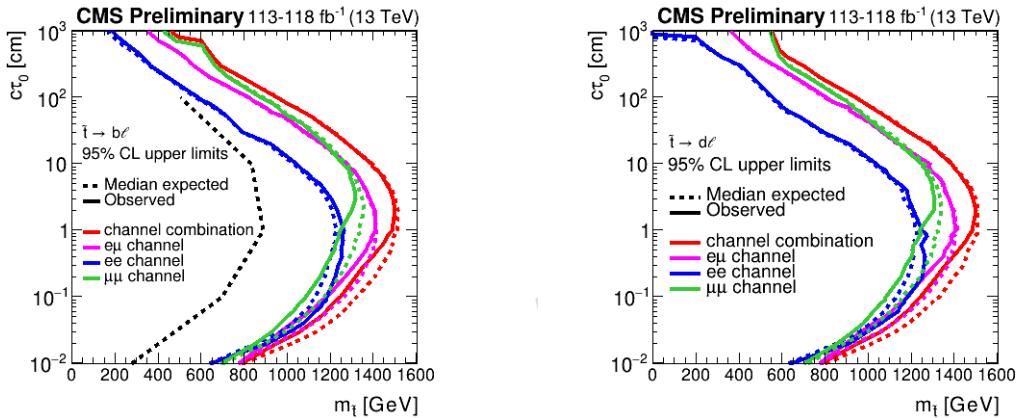


Figure 36: The 95% C.L. upper limits on the long-lived top squark proper decay length ( $c\tau$ ) as a function of its mass, for the  $e\mu$ , ee, and  $\mu\mu$  channels. The  $\tilde{t}\tilde{t} \rightarrow \bar{b}\bar{b}$  (left) and  $\tilde{t}\tilde{t} \rightarrow \bar{d}\bar{d}$  (right) processes are shown.

Figures 38 and 39 show the 95% confidence level (C.L.) upper limits for the sleptons and exotic Higgs bosons, respectively. The slepton limits assume the superpartners of the left- and right-handed leptons are mass degenerate. The Higgs boson limits assume  $\mathcal{B}(H \rightarrow SS = 100\%$ , and each l has an equal probability of being an electron or a muon.

See Appendix E for more information on the likelihood.

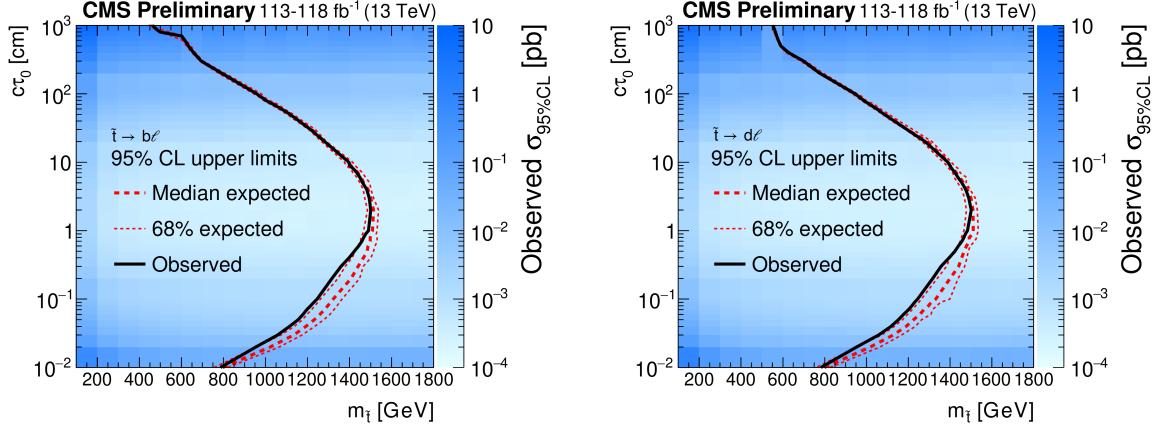


Figure 37: The 95% C.L. upper limits on the top squark  $c\tau$  as a function of its mass for the three channels combined. The colors indicate the observed 95% C.L. upper limit on the cross section. The  $\tilde{t}\tilde{t} \rightarrow \bar{l}b\bar{l}b$  (left) and  $\tilde{t}\tilde{t} \rightarrow \bar{l}d\bar{l}d$  (right) processes are shown.

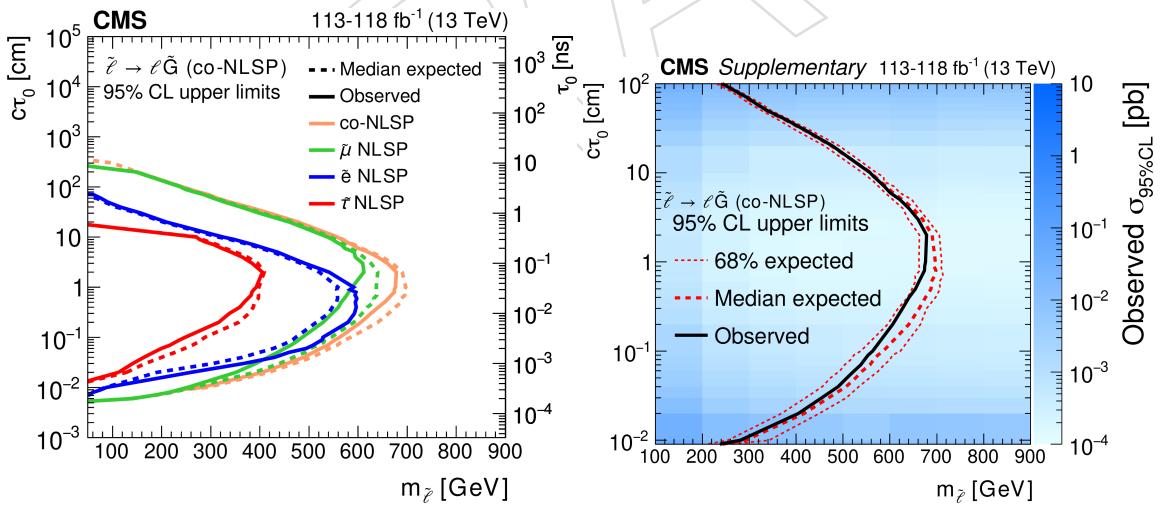


Figure 38: The 95% C.L. upper limits on the long-lived slepton production cross section, in the  $\tau$  – mass plane. The  $\tilde{\ell}$  and co-NLSP limits are shown for the three channels combined, while the  $\tilde{e}$  and  $\tilde{\mu}$  NLSP limits are shown for the ee and  $\mu\mu$  channels, respectively. The area to the left of the solid curves represents the observed exclusion region, and the dashed lines indicate the expected limits.

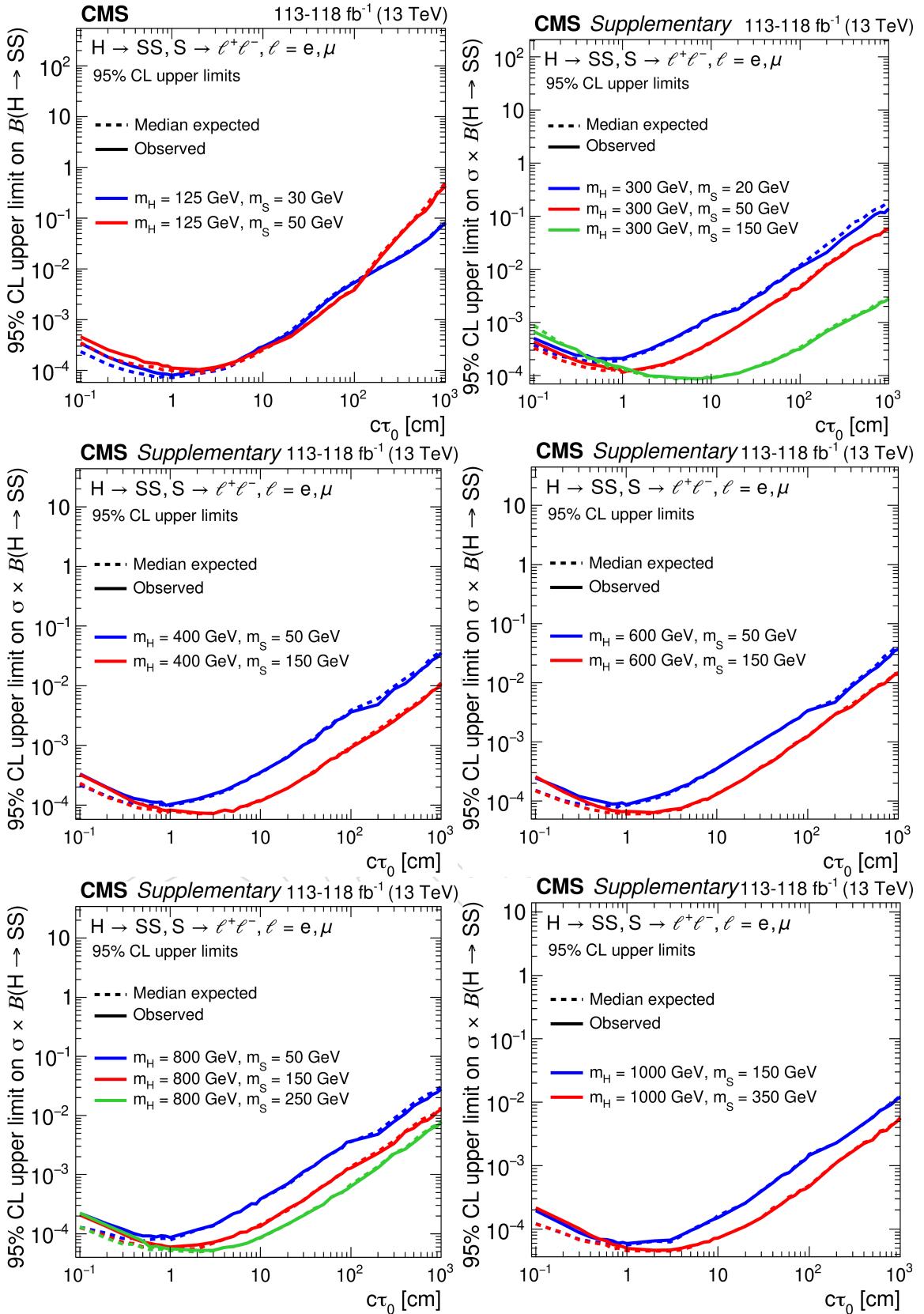


Figure 39: The 95% C.L. upper limits on the branching fraction or the product of the cross section and branching fraction  $\text{pp} \rightarrow H \rightarrow SS \rightarrow l\bar{l}l\bar{l}$  as a function of  $c\tau$ , for different exotic Higgs boson masses, for the three channels combined. The area above the solid (dashed) curve represents the observed (expected) exclusion region.

## 1031 10 Summary

1032 A search has been presented for long-lived particles decaying to displaced leptons in proton-  
 1033 proton collisions at a center-of-mass energy of 13 TeV at the CERN LHC. With collision data  
 1034 recorded in 2016, 2017, and 2018, and corresponding to an integrated luminosity of 113–118 fb<sup>−1</sup>,  
 1035 no excess above the estimated background has been observed. Upper limits have been set at  
 1036 95% confidence level. Top squarks with masses between 100 and at least 460 GeV have been  
 1037 excluded for proper decay lengths between 0.01 and 1000 cm, with a maximum exclusion of  
 1038 1500 GeV occurring at a proper decay length of 2 cm, assuming 100% of the top squarks de-  
 1039 cay to a lepton and a b or d quark, where the lepton has an equal probability of being an  
 1040 electron, muon, or tau lepton. The following exclusions assume that the superpartners of the  
 1041 left- and right-handed leptons are mass degenerate. Electron superpartners with masses of  
 1042 at least 50 GeV have been excluded for proper decay lengths between 0.007 and 70 cm, with  
 1043 a maximum exclusion of 610 GeV occurring at a proper decay length of 0.7 cm. Muon super-  
 1044 partners with masses of at least 50 GeV have been excluded for proper decay lengths between  
 1045 0.005 and 265 cm, with a maximum exclusion of 610 GeV occurring at a proper decay length of  
 1046 3 cm. Tau lepton superpartners with masses of at least 50 GeV have been excluded for proper  
 1047 decay lengths between 0.015 and 20 cm, with a maximum exclusion of 405 GeV occurring at  
 1048 a proper decay length of 2 cm. In the case that electron, muon, and tau lepton superpartners  
 1049 are mass degenerate, lepton superpartners with masses between 50 and at least 270 GeV have  
 1050 been excluded for proper decay length between 0.005 and 265 cm, with a maximum exclusion  
 1051 of 680 GeV occurring at a proper decay length of 2 cm. For proper decay lengths between 0.10  
 1052 and 12 cm, Higgs bosons with a mass of 125 GeV and with branching ratios to two long-lived  
 1053 scalars greater than 0.03% have been excluded, assuming each scalar has a mass of 30 GeV and  
 1054 decays with equal probability to electrons or muons. This is the first search at CMS for dis-  
 1055 placed leptons at a center-of-mass energy of 13 TeV, and the first search at CMS for displaced  
 1056 leptons in the electron-electron and muon-muon channels that does not require the leptons to  
 1057 come from a common displaced vertex. As a result of the larger center-of-mass energy and  
 1058 integrated luminosity, as well as the addition of the same-flavor channels, the mass exclusion  
 1059 limits for this search improve upon previous CMS results by approximately a factor of 2.

## 1060 Acknowledgments

1061 Frank Jensen and Kevin Stenson performed an earlier study of K<sub>S</sub><sup>0</sup>, similar to that presented in  
 1062 Section A, but based on the prompt, rather than the reprocessed data. Their advice facilitated  
 1063 the newer K<sub>S</sub><sup>0</sup> analysis presented here.

## References

- [1] CMS Collaboration, "Search for heavy stable charged particles with  $12.9 \text{ fb}^{-1}$  of 2016 data", *CMS-PAS-EXO-16-036* (2016).
- [2] ATLAS Collaboration, "Search for magnetic monopoles and stable high-electric-charge objects in 13 TeV proton-proton collisions with the ATLAS detector", *PRL* **124** (2020) 031802, doi:10.1103/PhysRevLett.124.031802, arXiv:1905.10130.
- [3] ATLAS Collaboration, "Search for heavy charged long-lived particles in the ATLAS detector in  $36.1 \text{ fb}^{-1}$  of proton-proton collision data at  $\sqrt{s} = 13 \text{ TeV}$ ", *PRD* **99** (2019) 092007, doi:10.1103/PhysRevD.99.092007, arXiv:1902.01636.
- [4] CMS Collaboration, "Search for disappearing tracks in proton-proton collisions at  $\sqrt{s} = 13 \text{ TeV}$ ", *PLB* **806** (2020) 135502, doi:10.1016/j.physletb.2020.135502, arXiv:2004.05153.
- [5] CMS Collaboration, "Searches for physics beyond the standard model with the  $M_{T2}$  variable in hadronic final states with and without disappearing tracks in proton-proton collisions at  $\sqrt{s} = 13 \text{ TeV}$ ", *EPJC* **80** (2020) 3, doi:10.1140/epjc/s10052-019-7493-x, arXiv:1909.03460.
- [6] ATLAS Collaboration, "Search for long-lived charginos based on a disappearing-track signature in pp collisions at  $\sqrt{s} = 13 \text{ TeV}$  with the ATLAS detector", *JHEP* **06** (2018) 022, doi:10.1007/JHEP06(2018)022, arXiv:1712.02118.
- [7] CMS Collaboration, "Search for long-lived particles using displaced jets in proton-proton collisions at  $\sqrt{s} = 13 \text{ TeV}$ ", arXiv:2012.01581.
- [8] CMS Collaboration, "Search for long-lived particles using nonprompt jets and missing transverse momentum with proton-proton collisions at  $\sqrt{s} = 13 \text{ TeV}$ ", *PLB* **797** (2019) 134876, doi:10.1016/j.physletb.2019.134876, arXiv:1906.06441.
- [9] ATLAS Collaboration, "Search for long-lived neutral particles in pp collisions at  $\sqrt{s} = 13 \text{ TeV}$  that decay into displaced hadronic jets in the ATLAS calorimeter", *EPJC* **79** (2019) 481, doi:10.1140/epjc/s10052-019-6962-6, arXiv:1902.03094.
- [10] ATLAS Collaboration, "Search for long-lived particles produced in pp collisions at  $\sqrt{s} = 13 \text{ TeV}$  that decay into displaced hadronic jets in the ATLAS muon spectrometer", *PRD* **99** (2019) 052005, doi:10.1103/PhysRevD.99.052005, arXiv:1811.07370.
- [11] ATLAS Collaboration, "Search for long-lived neutral particles produced in pp collisions at  $\sqrt{s} = 13 \text{ TeV}$  decaying into displaced hadronic jets in the ATLAS inner detector and muon spectrometer", *PRD* **101** (2020) 052013, doi:10.1103/PhysRevD.101.052013, arXiv:1911.12575.
- [12] CMS Collaboration, "Search for long-lived particles using delayed photons in proton-proton collisions at  $\sqrt{s} = 13 \text{ TeV}$ ", *PRD* **100** (2019) 112003, doi:10.1103/PhysRevD.100.112003, arXiv:1909.06166.
- [13] CMS Collaboration, "Search for displaced photons using conversions at 8 TeV", CMS Physics Analysis Summary CMS-PAS-EXO-14-017, 2015.

- [14] ATLAS Collaboration, "Search for nonpointing and delayed photons in the diphoton and missing transverse momentum final state in 8 TeV  $pp$  collisions at the LHC using the ATLAS detector", *PRD* **90** (2014) 112005, doi:10.1103/PhysRevD.90.112005, arXiv:1409.5542.
- [15] CMS Collaboration, "Search for long-lived particles that decay into final states containing two electrons or two muons in proton-proton collisions at  $\sqrt{s} = 8$  TeV", *PRD* **91** (2015) 052012, doi:10.1103/PhysRevD.91.052012, arXiv:1411.6977.
- [16] CMS Collaboration, "Search for displaced leptons in the e-mu channel", *CMS-PAS-EXO-16-022* (2016).
- [17] ATLAS Collaboration, "Search for displaced leptons in  $\sqrt{s} = 13$  TeV  $pp$  collisions with the ATLAS detector", arXiv:2011.07812.
- [18] ATLAS Collaboration, "Search for heavy neutral leptons in decays of  $W$  bosons produced in 13 TeV  $pp$  collisions using prompt and displaced signatures with the ATLAS detector", *JHEP* **10** (2019) 265, doi:10.1007/JHEP10(2019)265, arXiv:1905.09787.
- [19] ATLAS Collaboration, "Search for long-lived, massive particles in events with a displaced vertex and a muon with large impact parameter in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector", *PRD* **102** (2020) 032006, doi:10.1103/PhysRevD.102.032006, arXiv:2003.11956.
- [20] ATLAS Collaboration, "Search for displaced vertices of oppositely charged leptons from decays of long-lived particles in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector", *PLB* **801** (2020) 135114, doi:10.1016/j.physletb.2019.135114, arXiv:1907.10037.
- [21] CMS Collaboration, "Search for decays of stopped exotic long-lived particles produced in proton-proton collisions at  $\sqrt{s} = 13$  TeV", *JHEP* **05** (2018) 127, doi:10.1007/JHEP05(2018)127, arXiv:1801.00359.
- [22] ATLAS Collaboration, "Search for light long-lived neutral particles produced in  $pp$  collisions at  $\sqrt{s} = 13$  TeV and decaying into collimated leptons or light hadrons with the ATLAS detector", *EPJC* **80** (2020) 450, doi:10.1140/epjc/s10052-020-7997-4, arXiv:1909.01246.
- [23] CMS Collaboration, "Search for Displaced Supersymmetry in events with an electron and a muon with large impact parameters", *Phys. Rev. Lett.* **114** (2015), no. 6, 061801, doi:10.1103/PhysRevLett.114.061801, arXiv:1409.4789.
- [24] P. Graham, D. Kaplan, S. Rajendran, and P. Sarawat, "Displaced supersymmetry", *JHEP* **07** (2012) 149, doi:10.1007/JHEP07(2012)149, arXiv:1204.6038.
- [25] J. A. Evans and J. Shelton, "Long-Lived Staus and Displaced Leptons at the LHC", *JHEP* **04** (2016) 056, doi:10.1007/JHEP04(2016)056, arXiv:1601.01326.
- [26] M. J. Strassler and K. M. Zurek, "Discovering the Higgs through highly-displaced vertices", *PLB* **661** (2008) 263, doi:10.1016/j.physletb.2008.02.008, arXiv:hep-ph/0605193.
- [27] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

- 1145 [28] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", *JHEP* **12** (2012)  
1146 061, doi:[10.1007/JHEP12\(2012\)061](https://doi.org/10.1007/JHEP12(2012)061), arXiv:[1209.6215](https://arxiv.org/abs/1209.6215).
- 1147 [29] J. Alwall et al., "Comparative study of various algorithms for the merging of parton  
1148 showers and matrix elements in hadronic collisions", *Eur. Phys. J. C* **53** (2008) 473,  
1149 doi:[10.1140/epjc/s10052-007-0490-5](https://doi.org/10.1140/epjc/s10052-007-0490-5), arXiv:[0706.2569](https://arxiv.org/abs/0706.2569).
- 1150 [30] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower  
1151 simulations", *JHEP* **06** (2002) 029, doi:[10.1088/1126-6708/2002/06/029](https://doi.org/10.1088/1126-6708/2002/06/029),  
1152 arXiv:[hep-ph/0204244](https://arxiv.org/abs/hep-ph/0204244).
- 1153 [31] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo  
1154 algorithms", *JHEP* **11** (2004) 040, doi:[10.1088/1126-6708/2004/11/040](https://doi.org/10.1088/1126-6708/2004/11/040),  
1155 arXiv:[hep-ph/0409146](https://arxiv.org/abs/hep-ph/0409146).
- 1156 [32] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton  
1157 Shower simulations: the POWHEG method", *JHEP* **11** (2007) 070,  
1158 doi:[10.1088/1126-6708/2007/11/070](https://doi.org/10.1088/1126-6708/2007/11/070), arXiv:[0709.2092](https://arxiv.org/abs/0709.2092).
- 1159 [33] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO vector-boson production matched with  
1160 shower in POWHEG", *JHEP* **07** (2008) 060,  
1161 doi:[10.1088/1126-6708/2008/07/060](https://doi.org/10.1088/1126-6708/2008/07/060), arXiv:[0805.4802](https://arxiv.org/abs/0805.4802).
- 1162 [34] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO  
1163 calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043,  
1164 doi:[10.1007/JHEP06\(2010\)043](https://doi.org/10.1007/JHEP06(2010)043), arXiv:[1002.2581](https://arxiv.org/abs/1002.2581).
- 1165 [35] T. Sjstrand et al., "An introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015)  
1166 159, doi:[10.1016/j.cpc.2015.01.024](https://doi.org/10.1016/j.cpc.2015.01.024), arXiv:[1410.3012](https://arxiv.org/abs/1410.3012).
- 1167 [36] CMS Collaboration, "Event generator tunes obtained from underlying event and  
1168 multiparton scattering measurements", *Eur. Phys. J. C* **76** (2016), no. 3, 155,  
1169 doi:[10.1140/epjc/s10052-016-3988-x](https://doi.org/10.1140/epjc/s10052-016-3988-x), arXiv:[1512.00815](https://arxiv.org/abs/1512.00815).
- 1170 [37] CMS Collaboration, "Extraction and validation of a new set of CMS PYTHIA8 tunes from  
1171 underlying-event measurements", *EPJC* **80** (2020) 4,  
1172 doi:[10.1140/epjc/s10052-019-7499-4](https://doi.org/10.1140/epjc/s10052-019-7499-4), arXiv:[1903.12179](https://arxiv.org/abs/1903.12179).
- 1173 [38] R. Mackeprang and A. Rizzi, "Interactions of coloured heavy stable particles in matter",  
1174 *Eur. Phys. J. C* **50** (2007) 353, doi:[10.1140/epjc/s10052-007-0252-4](https://doi.org/10.1140/epjc/s10052-007-0252-4),  
1175 arXiv:[hep-ph/0612161](https://arxiv.org/abs/hep-ph/0612161).
- 1176 [39] R. Mackeprang and D. Milstead, "An updated description of heavy-hadron interactions  
1177 in GEANT-4", *Eur. Phys. J. C* **66** (2010) 493,  
1178 doi:[10.1140/epjc/s10052-010-1262-1](https://doi.org/10.1140/epjc/s10052-010-1262-1), arXiv:[0908.1868](https://arxiv.org/abs/0908.1868).
- 1179 [40] J. Alwall et al., "A standard format for Les Houches event files", *Comput. Phys. Commun.*  
1180 **176** (2007) 300, doi:[10.1016/j.cpc.2006.11.010](https://doi.org/10.1016/j.cpc.2006.11.010), arXiv:[hep-ph/0609017](https://arxiv.org/abs/hep-ph/0609017).
- 1181 [41] B. Allanach et al., "The Snowmass points and slopes: Benchmarks for SUSY searches",  
1182 *EPJC* **25** (2002) 113, doi:[10.1007/s10052-002-0949-3](https://doi.org/10.1007/s10052-002-0949-3), arXiv:[hep-ph/0202233](https://arxiv.org/abs/hep-ph/0202233).
- 1183 [42] CMS Collaboration, "Particle-flow reconstruction and global event description with the  
1184 CMS detector", *JINST* **12** (2017) P10003, doi:[10.1088/1748-0221/12/10/P10003](https://doi.org/10.1088/1748-0221/12/10/P10003),  
1185 arXiv:[1706.04965](https://arxiv.org/abs/1706.04965).

- 1186 [43] EGamma POG, “Cut Based Electron ID for Run 2, Tight Electron”, <https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedElectronIdentificationRun2>.
- 1187
- 1188 [44] Muon POG, “Baseline muon selections for Run-II, Tight Muon”, [https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2#Tight\\_Muon](https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2#Tight_Muon).
- 1189
- 1190 [45] CMS Collaboration, “Precision measurement of the structure of the CMS inner tracking system using nuclear interactions”, *JINST* **13** (2018) P10034,  
1191 doi:10.1088/1748-0221/13/10/P10034, arXiv:1807.03289.
- 1192
- 1193 [46] CMS Collaboration, “Precision measurement of the structure of the CMS inner tracking system using nuclear interactions with data collected in 2018”, *CMS-DP-2019-001* (2019).
- 1194
- 1195 [47] CMS Collaboration, “Precision luminosity measurement in proton-proton collisions at  $\sqrt{s} = 13$  TeV in 2015 and 2016 at CMS”, (2021). arXiv:2104.01927. Submitted to *Eur. Phys. J. C*.
- 1196
- 1197
- 1198 [48] CMS Collaboration, “CMS luminosity measurements for the 2017 data-taking period at  $\sqrt{s} = 13$  TeV”, CMS Physics Analysis Summary CMS-PAS-LUM-17-004, 2018.
- 1199
- 1200 [49] CMS Collaboration, “CMS luminosity measurements for the 2018 data-taking period at  $\sqrt{s} = 13$  TeV”, CMS Physics Analysis Summary CMS-PAS-LUM-18-002, 2018.
- 1201
- 1202 [50] Lumi POG, “Luminosity POG Twiki”,  
<https://twiki.cern.ch/twiki/bin/view/CMS/TWikiLUM#LumiComb>.
- 1203
- 1204 [51] CMS Collaboration, “Measurement of the inelastic proton-proton cross section at  $\sqrt{s} = 13$  TeV”, *JHEP* **07** (2018) 161, doi:10.1007/JHEP07(2018)161, arXiv:1802.02613.
- 1205
- 1206 [52] Data quality group, “Documentation on how to produce JSON files”, <https://twiki.cern.ch/twiki/bin/view/CMS/InstructionForJSONfileCreation>.
- 1207
- 1208 [53] 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.
- 1209
- 1210
- 1211 [54] A. L. Read, “Presentation of search results: the  $CL_s$  technique”, in *Durham IPPP Workshop: Advanced Statistical Techniques in Particle Physics*, p. 2693. Durham, UK, March, 2002. [*J. Phys. G* **28** (2002) 2693]. doi:10.1088/0954-3899/28/10/313.
- 1212
- 1213
- 1214 [55] 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].
- 1215
- 1216
- 1217
- 1218 [56] The ATLAS Collaboration, The CMS Collaboration, The LHC Higgs Combination Group, “Procedure for the LHC Higgs boson search combination in Summer 2011”, Technical Report CMS-NOTE-2011-005, ATL-PHYS-PUB-2011-11, 2011.
- 1219
- 1220
- 1221 [57] CMS Tracker Collaboration, “The effect of highly ionising particles on the CMS silicon strip tracker”, *Nucl. Instrum. Meth. A* **543** (2005) 463, doi:10.1016/j.nima.2004.11.049.
- 1222
- 1223

---

1224 **A Study with  $K_S^0$**

1225 Data taken in 2016 prior to run 278802 were affected by the APV saturation effect [57], in which  
 1226 inelastic nuclear interactions in the tracker silicon sensors result in the deposit of a huge amount  
 1227 of charge there (equivalent to that of hundreds of minimum ionizing particles). This in turn  
 1228 saturates the front-end readout chip, and causes it to temporarily lose efficiency. The effect  
 1229 increases with instantaneous luminosity, so had not been significant in the lower luminosity  
 1230 runs taken in previous years. From run 278802 onwards, the tracker front-end electronics was  
 1231 reconfigured, which substantially reduced its sensitivity to the APV saturation effect.

1232 The APV saturation effect causes some tracker hits to be lost, which reduces the tracking effi-  
 1233 ciency. One would expect the efficiency loss to be larger for displaced particles from long-lived  
 1234 particle decay, since such displaced particles traverse fewer silicon layers, so can less afford to  
 1235 lose some of their hits. The same effect should reduce the efficiency to reconstruct pions from  
 1236  $K_S^0 \rightarrow \pi^+ \pi^-$  decay, so the latter can be used to understand if data affected by the APV satura-  
 1237 tion effect can be safely used for displaced particle searches. However, as pions from  $K_S^0$  decay  
 1238 tend to be softer and have lower impact parameter than displaced particles from long-lived  
 1239 exotica decay, this will only give a rough estimate.

1240 Data collected in 2016, 2017 and 2018 with the `HLT-ZeroBias` trigger were taken from the  
 1241 "07Aug17", "17Nov2017" and "17Sep2018" processings of the AOD, respectively. Only 9, 6 and  
 1242 8 different runs were studied from 2016, 2017 and 2018, respectively, selected as they together  
 1243 explored a wide range of instantaneous luminosity and data taking eras and covered periods  
 1244 before and after the APV saturation effect was largely eliminated by reconfiguring the front-  
 1245 end electronics.

1246 All  $K_S^0$  candidates in the `generalV0Candidates:Kshort` collection were plotted. These con-  
 1247 sist of pairs of oppositely charged tracks consistent with the  $K_S^0$  mass and coming from a com-  
 1248 mon vertex that lies at least 2 cm from the beam-line. The tracks were required to have at least  
 1249 one pixel hit. For 2016 data, it was required that both tracks have  $p_T > 0.7 \text{ GeV}$  and  $|\eta| < 2$ ,  
 1250 since the displaced tracking efficiency is poor for particles failing these requirements. The  
 1251 higher statistics of the 2017-2018 data allowed the  $p_T$  requirement to be increased to 1.5 GeV.

1252 Figure 40 shows the reconstructed decay length of the  $K_S^0$  candidates, measured in the plane  
 1253 perpendicular to the beamline, for data from the three years. The distributions shown for each  
 1254 run are normalized to the integrated luminosity of that run. In the 2016 plot, it can be seen that  
 1255 in runs taken before the APV saturation effect was fixed by reconfiguring the electronics (shown  
 1256 by solid lines), the transverse decay length distribution falls much more rapidly in runs with  
 1257 higher instantaneous luminosity. On the other hand, after the APV saturation effect was fixed  
 1258 (shown by dashed lines in the 2016 plot or all lines in the 2017 and 2018 plots), the distributions  
 1259 show a much smaller dependence on instantaneous luminosity. The natural interpretation of  
 1260 this is that the change shown by the solid lines in Fig. 40 reflects the loss of displaced tracking  
 1261 efficiency when the APV saturation effect was present.

1262 The size of this effect leads to the conclusion that 2016 data taken before run 278802 should  
 1263 not be used for this search. The displaced tracking efficiency would depend heavily on instan-  
 1264 taneous luminosity and be difficult to quantify. The systematic uncertainties in the displaced  
 1265 tracking efficiency in this period would be sizeable, and because of the poor efficiency, any sig-  
 1266 nal would be suppressed. Furthermore, studies of displaced tracking with cosmic ray data are  
 1267 insensitive to the APV saturation effect, since the instantaneous luminosity during dedicated  
 1268 cosmic runs is zero.

1269 Some small dependence of the efficiency on the instantaneous luminosity remains even in runs

taken after the front-end electronics was reconfigured. This is particularly noticeable in 2017-2018, when the luminosity was especially high. This may be due to a residual APV saturation effect or due to luminosity related inefficiency of the pixel tracker.

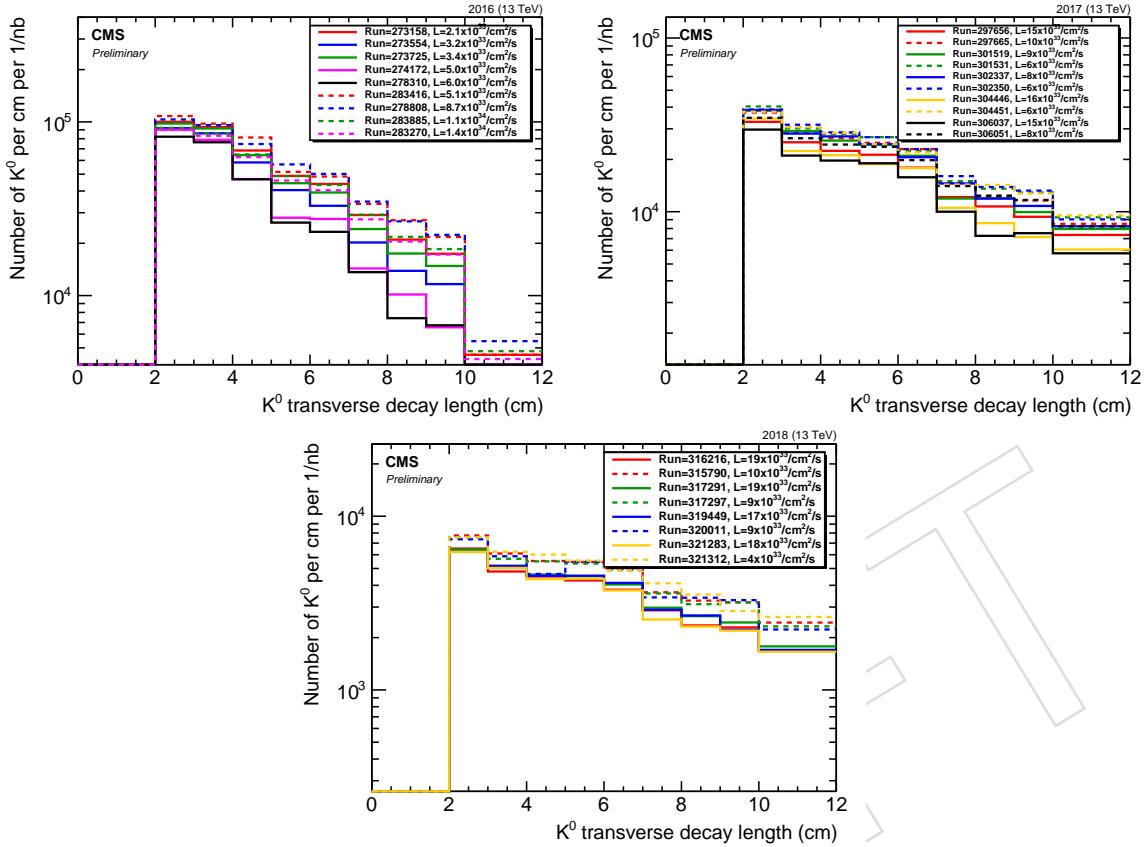


Figure 40: Transverse decay length distribution of reconstructed  $K_S^0 \rightarrow \pi^+\pi^-$  candidates for various runs in 2016 (top-left), 2017 (top-right) and 2018 (below). The peak instantaneous luminosity of each run is indicated in the legend. Each distribution is normalized to the integrated luminosity of the run. In the 2016 plot, runs taken after the APV saturation effect was eliminated by reconfiguring the tracker front-end electronics are shown by dashed lines, whereas those taken before this are shown by solid lines. In the 2017 plot, the run with lowest (highest) instantaneous luminosity examined in each data taking era is plotted with a dashed (solid) line. Towards the end of 2017 (orange and black lines) the LHC ran with approximately 20% fewer proton bunches, meaning that the instantaneous luminosity per bunch crossing was higher than one might naively expect from the instantaneous luminosities shown in the legend.

Table 22: Privately generated signal MC simulation samples.

Private signal samples
/StopToL*_M_100_1mm_13TeV_2016MC/jalimena-MiniAod-bd3e7bcff6c9bcad356ea4ed7e4f08b4/USER
/StopToL*_M_100_10mm_13TeV_2016MC/jalimena-MiniAod-bd3e7bcff6c9bcad356ea4ed7e4f08b4/USER
/StopToL*_M_100_100mm_13TeV_2016MC/jalimena-MiniAod-bd3e7bcff6c9bcad356ea4ed7e4f08b4/USER
/StopToL*_M_100_1000mm_13TeV_2016MC/jalimena-MiniAod-bd3e7bcff6c9bcad356ea4ed7e4f08b4/USER
/StopToLB_M_1300_1mm_13TeV_2016MC/jalimena-MiniAod-bd3e7bcff6c9bcad356ea4ed7e4f08b4/USER
/StopToLB_M_1400_100mm_13TeV_2016MC/jalimena-MiniAod-bd3e7bcff6c9bcad356ea4ed7e4f08b4/USER
/StopToLB_M_1500_1mm_13TeV_2016MC/jalimena-MiniAod-bd3e7bcff6c9bcad356ea4ed7e4f08b4/USER
/StopToLB_M_1700_1000mm_13TeV_2016MC/jalimena-MiniAod-bd3e7bcff6c9bcad356ea4ed7e4f08b4/USER
/StopToLB_M_1800_1000mm_13TeV_2016MC/jalimena-MiniAod-bd3e7bcff6c9bcad356ea4ed7e4f08b4/USER
/StopToL*_M_*_0p1mm_13TeV_2016MC/jalimena-MiniAod-53f8667ba4b240d5eaf36e71bf34742/USER
/StopToL*_M_100_1mm_13TeV_2017MC/jalimena-MiniAOD-18783c0a07109245951450a1a4f55409/USER
/StopToL*_M_100_10mm_13TeV_2017MC/jalimena-MiniAod-18783c0a07109245951450a1a4f55409/USER
/StopToL*_M_100_100mm_13TeV_2017MC/jalimena-MiniAod-18783c0a07109245951450a1a4f55409/USER
/StopToL*_M_100_1000mm_13TeV_2017MC/jalimena-MiniAod-18783c0a07109245951450a1a4f55409/USER
/StopToLB_M_200_1mm_13TeV_2017MC/jalimena-MiniAOD-18783c0a07109245951450a1a4f55409/USER
/StopToLB_M_200_10mm_13TeV_2017MC/jalimena-MiniAod-18783c0a07109245951450a1a4f55409/USER
/StopToLB_M_200_100mm_13TeV_2017MC/jalimena-MiniAod-18783c0a07109245951450a1a4f55409/USER
/StopToLB_M_200_1000mm_13TeV_2017MC/jalimena-MiniAod-18783c0a07109245951450a1a4f55409/USER
/StopToLB_M_300_1mm_13TeV_2017MC/jalimena-MiniAod-18783c0a07109245951450a1a4f55409/USER
/StopToLB_M_300_10mm_13TeV_2017MC/jalimena-MiniAod-18783c0a07109245951450a1a4f55409/USER
/StopToLB_M_300_100mm_13TeV_2017MC/jalimena-MiniAod-18783c0a07109245951450a1a4f55409/USER
/StopToLB_M_300_1000mm_13TeV_2017MC/jalimena-MiniAod-18783c0a07109245951450a1a4f55409/USER
/StopToLB_M_400_1mm_13TeV_2017MC/jalimena-MiniAod-18783c0a07109245951450a1a4f55409/USER
/StopToL*_M_*_0p1mm_13TeV_2017MC/bcardwel-MiniAod-4fe98f39b775e67c69bc92a03424ad6b/USER
/StopToL*_M_100_1000mm_13TeV_2018MC/bcardwel-MiniAod-3ee3afd6b5a1410aea6d0b4d52723d06/USER
/StopToL*_M_100_100mm_13TeV_2018MC/bcardwel-MiniAod-3ee3afd6b5a1410aea6d0b4d52723d06/USER
/StopToL*_M_100_10mm_13TeV_2018MC/bcardwel-MiniAod-3ee3afd6b5a1410aea6d0b4d52723d06/USER
/StopToL*_M_100_1mm_13TeV_2018MC/bcardwel-MiniAod-3ee3afd6b5a1410aea6d0b4d52723d06/USER
/StopToL*_M_*_0p1mm_13TeV_2018MC/jalimena-MiniAod-c21dec93027231dc6f615dfe5c662834/USER

1273 **B Privately generated signal samples**

1274 Some signal samples needed to be generated privately, as they failed during the central production.  
 1275 The privately generated signal samples can be found in "phys03" DBS instance and are  
 1276 listed in Table 22. They were generated with the same conditions, and for the same number of  
 1277 events, as the centrally produced samples. Figure 41 shows a few key distributions to validate  
 1278 the privately generated samples against the centrally produced ones.

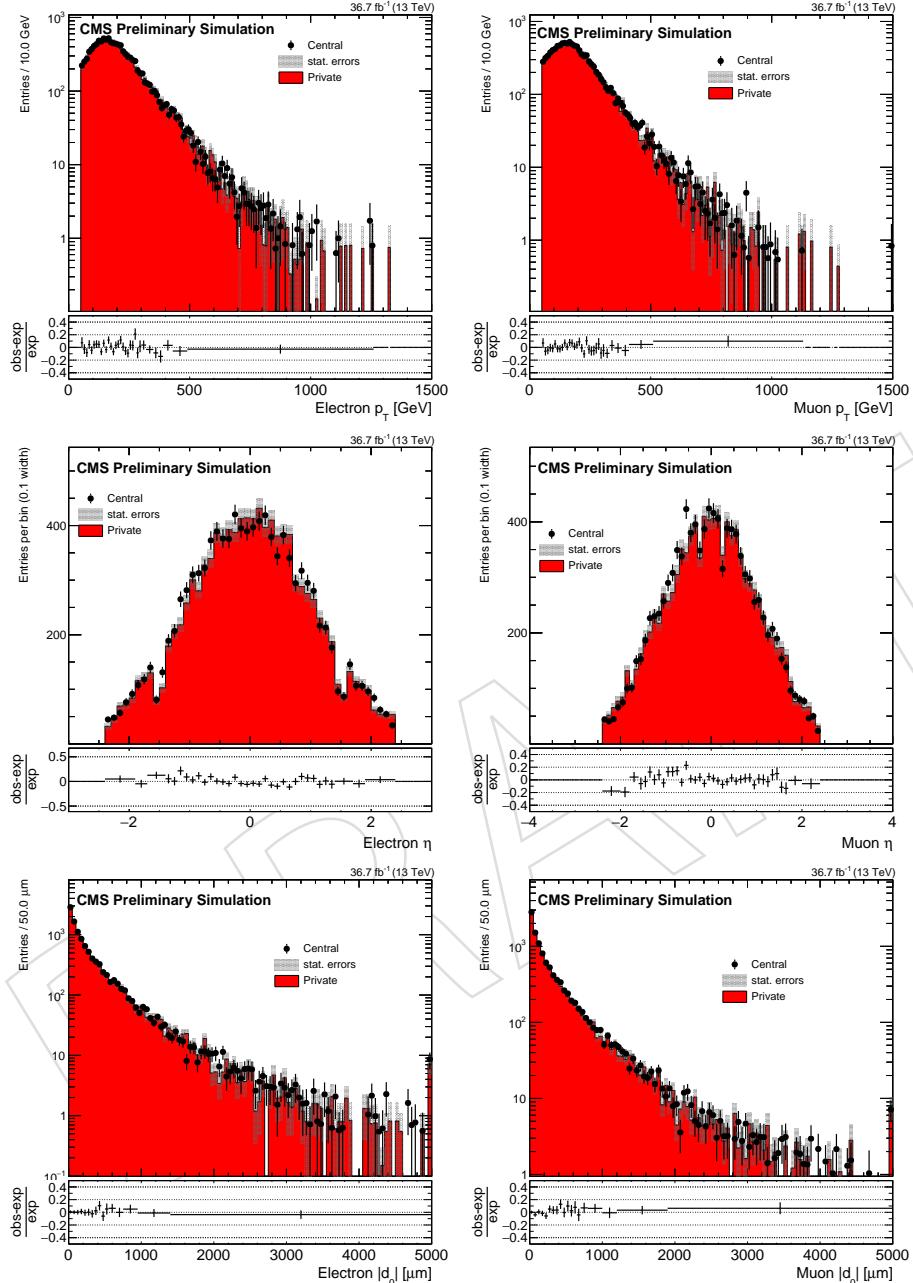


Figure 41: The lepton  $p_T$  (upper),  $\eta$  (middle), and  $|d_0|$  (lower) distributions for electrons (left) and muons (right), for events that pass the  $e\mu$  preselection, for privately and centrally-produced  $\tilde{t}\tilde{t} \rightarrow b1b1$  simulation with 2017 conditions. The  $\tilde{t}$  mass is 400 GeV and the lifetime is 1 mm. The rightmost bin in each plot contains the overflow entries.

## 1279 C Mismeasured lepton $|d_0|$ at large $|\eta|$

1280 We require muons to have  $|\eta| < 1.5$  due to the observed increase in width of the muon  $d_0$   
 1281 distribution at large  $|\eta|$  in DY simulation with  $DY \rightarrow \tau\tau$  events removed (see Fig. 42 left). This  
 1282 width increases at large  $|\eta|$  dramatically in 2016 conditions, but also to a certain extent in 2017  
 1283 and 2018 conditions. Requiring muon  $|\eta| < 1.5$  has two effects: (1) it dramatically reduces  
 1284 the background in 2016 data in the  $\mu\mu$  channel due to mismeasurements, and (2) it removes a  
 1285 (non- $\tau$  lepton) source of correlation between the muon  $d_0$ s, since the muons in DY events are  
 1286 likely to have similar  $\eta$ s. The signal is fairly central (see Fig. 42 right), and so we can afford to  
 1287 tighten the muon  $\eta$  selection to  $< 1.5$ .

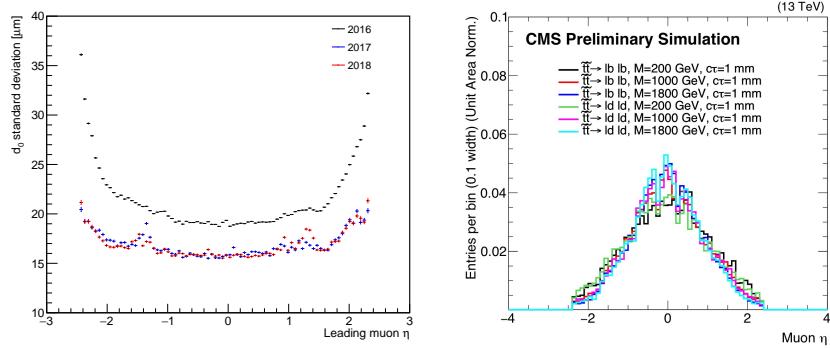


Figure 42: left: The standard deviation of the leading muon  $|d_0|$ , as a function of the leading muon  $\eta$ , for muons that pass the  $\mu\mu$  preselection (except for tight  $|\eta|$  cut). The sample used is DY simulation with  $DY \rightarrow \tau\tau$  events removed, in 2016, 2017, and 2018 conditions. right: Muon  $\eta$  distribution for signal simulated events that pass the the  $e\mu$  preselection in 2018 conditions (before the tighter  $\eta$  selection).

1288 We also require electrons to have  $|\eta| < 1.5$ , since the electron  $d_0$  resolution worsens at large  $|\eta|$ .  
 1289 Furthermore, we plotted background simulated events that pass the preselection, where the  
 1290 leptons come from light or heavy flavor mesons, and compared that to signal. We noticed that  
 1291 electrons from light mesons are particularly concentrated at large  $|\eta|$ , while the signal is fairly  
 1292 central (see Fig. 43). Based on these plots, we tightened the electron  $|\eta|$  selection to require  
 1293 electron  $|\eta| < 1.5$ .

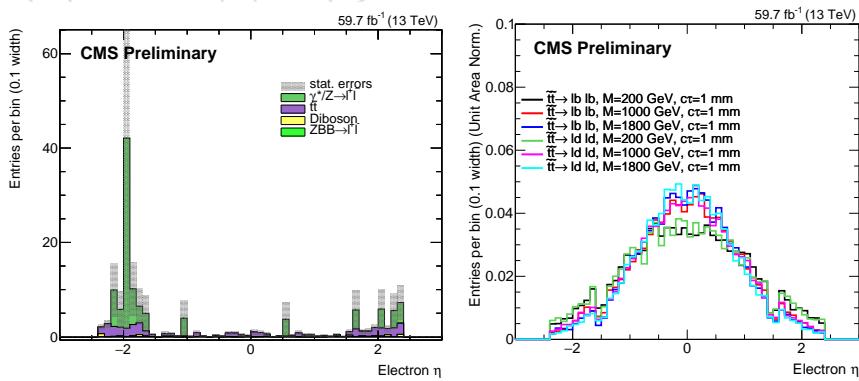


Figure 43: Electron  $\eta$  distributions for background simulated events, where the leptons come from light mesons (left) and signal simulation (right), for events that pass the  $e\mu$  preselection in 2018 conditions (before the tighter  $\eta$  selection).

## D Additional plots

### D.1 GMSB sleptons: kinematic distributions and cutflows

The  $p_T$  and  $\eta$  distributions are shown for selected selectron and smuon signal points in Fig. 44, for events that pass the ee and  $\mu\mu$  preselection. Likewise, the  $p_T$  and  $\eta$  distributions are shown for selected stau signal points in Fig. 45, for events that pass the  $e\mu$  preselection.

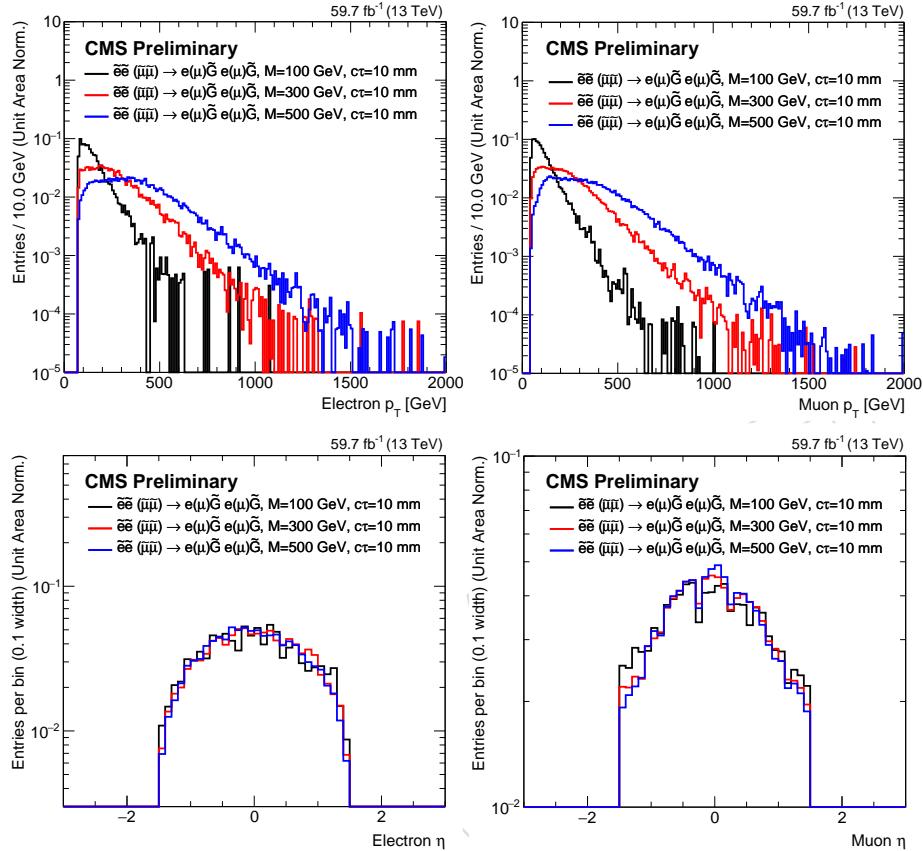


Figure 44: The electron (left)  $p_T$  (upper) and  $\eta$  (lower) distributions in 2018 selectron signal for events that pass the ee preselection, and the muon (right)  $p_T$  (upper) and  $\eta$  (lower) distributions in 2018 smuon signal for events that pass the  $\mu\mu$  preselection. The distributions are shown for one choice of slepton proper decay length and 3 sample masses. Each histogram is normalized to unity.

- The electron and muon  $|d_0|$  distributions are shown for selectron and smuon signals at different lifetimes in Fig. 46, for events that pass the ee and  $\mu\mu$  preselection criteria.
- Figures 47 and 48 show the cumulative number of events that pass the  $e\mu$ , ee, and  $\mu\mu$  preselection criteria for slepton signal events with different proper decay lengths and masses.

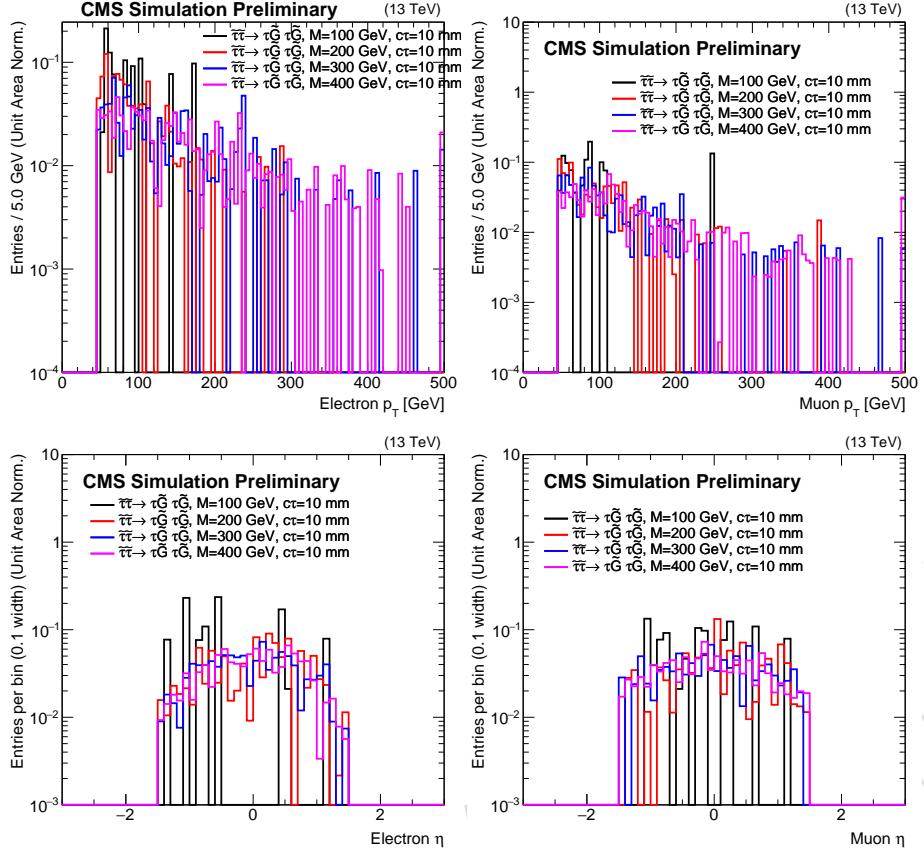


Figure 45: The electron (left) and the muon (right)  $p_T$  (upper) and  $\eta$  (lower) distributions in 2018 stau signal for events that pass the  $e\mu$  preselection. The distributions are shown for one choice of stau proper decay length and 4 sample masses. Each histogram is normalized to unity.

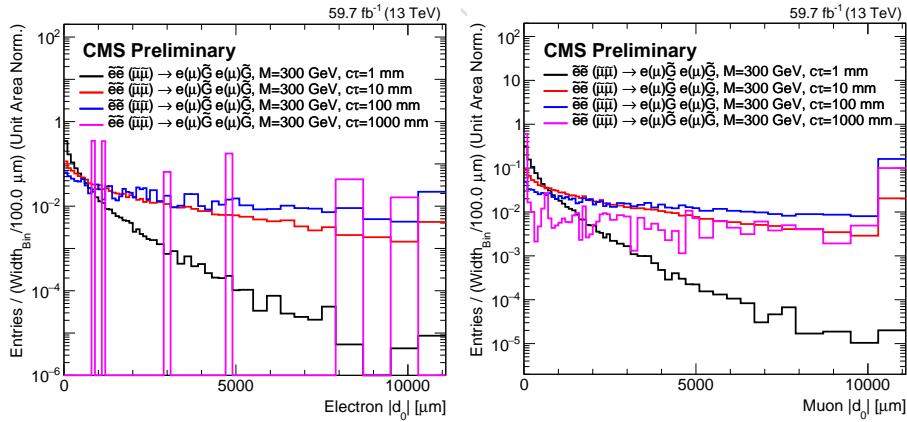


Figure 46: left: The electron  $|d_0|$  distribution for 2018 selectron events that pass the  $ee$  pre-selection criteria. right: The muon  $|d_0|$  distribution for 2018 smuon events that pass the  $\mu\mu$  preselection criteria. The plots are shown for a single selectron or smuon mass and 4 different proper decay lengths, where each histogram is normalized to unity. In all of the histograms, the last bin includes the overflow. All of the corrections from Section 4 are applied.

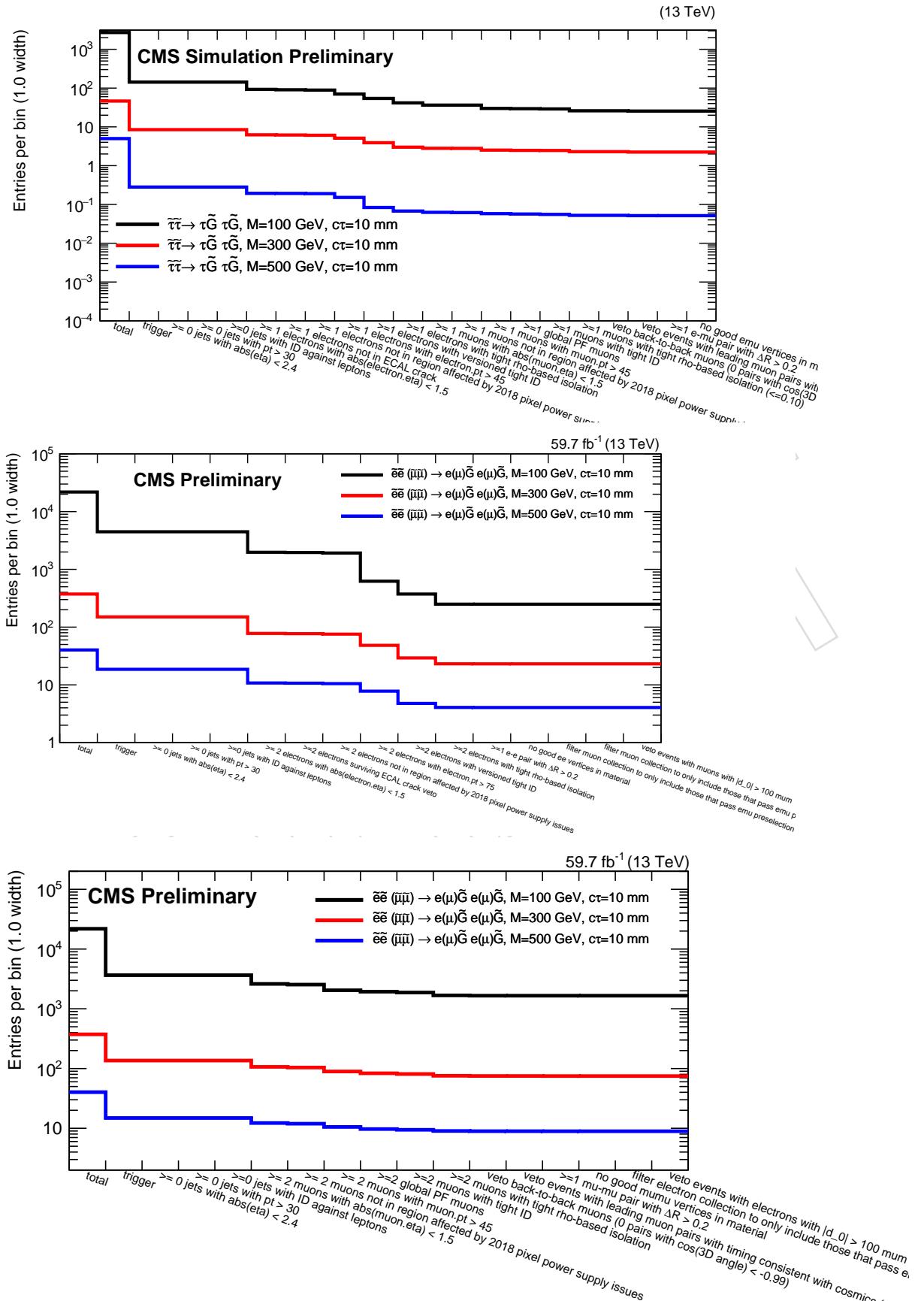


Figure 47: The cumulative number of events that pass each criterion in the  $e\mu$  (upper),  $ee$  (middle), and  $\mu\mu$  (lower) preselection, using 2018 signal simulation. Status are shown for the  $e\mu$  channel, selectrons for the  $ee$  channel, and smuons for the  $\mu\mu$  channel. Several slepton masses are shown. The jet criteria do not exclude any events and are simply an artifact of the analysis framework.

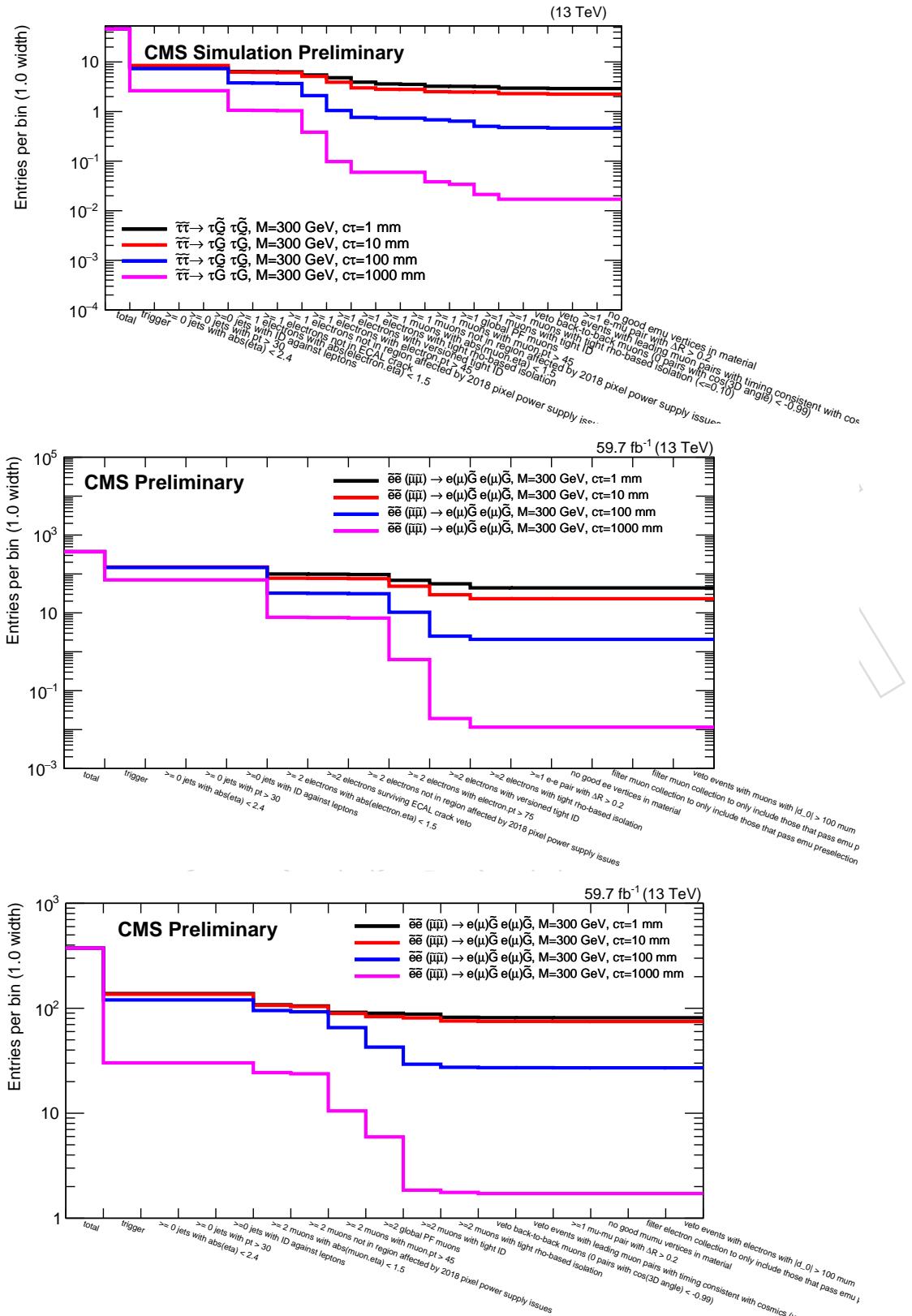


Figure 48: The cumulative number of events that pass each criterion in the  $e\mu$  (upper),  $ee$  (middle), and  $\mu\mu$  (lower) preselection, using 2018 signal simulation. Status are shown for the  $e\mu$  channel, selectrons for the  $ee$  channel, and smuons for the  $\mu\mu$  channel. Several slepton proper decay lengths are shown. The jet criteria do not exclude any events and are simply an artifact of the analysis framework.

## D.2 Exotic Higgs bosons: kinematic distributions and cutflows

The  $p_T$  and  $\eta$  distributions are shown for selected Higgs signal points in Fig. 49, for events that pass the  $e\mu$  preselection.

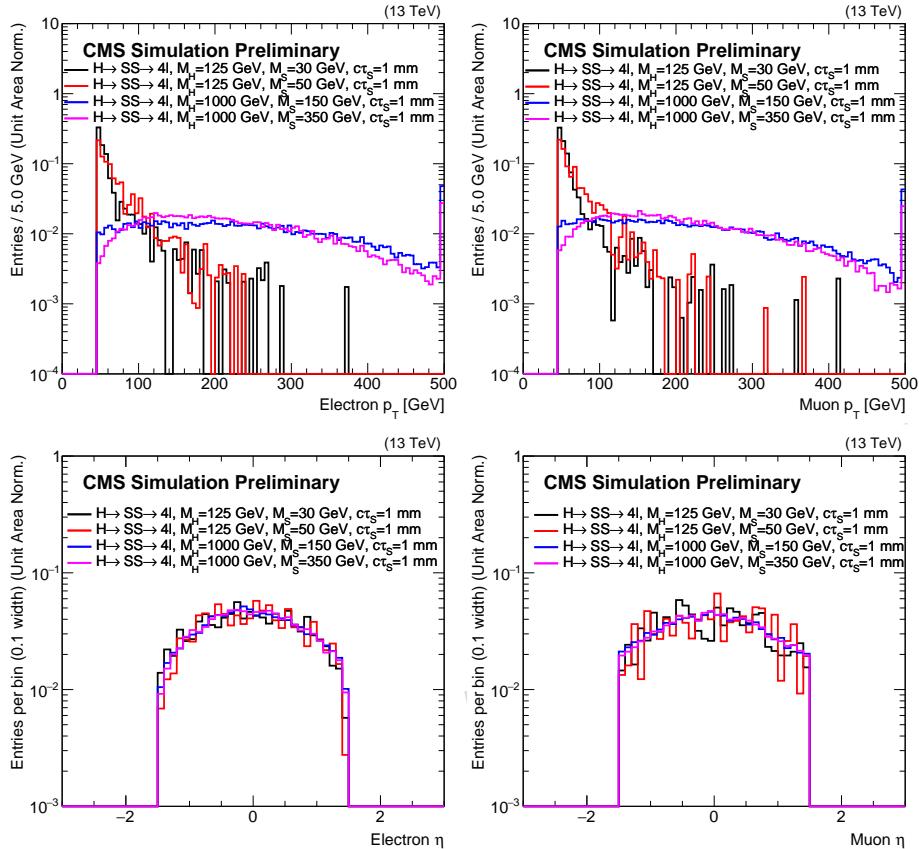


Figure 49: The electron (left) and the muon (right)  $p_T$  (upper) and  $\eta$  (lower) distributions in 2018 Higgs signal for events that pass the  $e\mu$  preselection. The distributions are shown for one choice of scalar proper decay length and 4 sample Higgs and scalar mass combinations. Each histogram is normalized to unity.

The electron and muon  $|d_0|$  distributions are shown for Higgs signals at different lifetimes in Fig. 50, for events that pass the  $e\mu$  and preselection criteria.

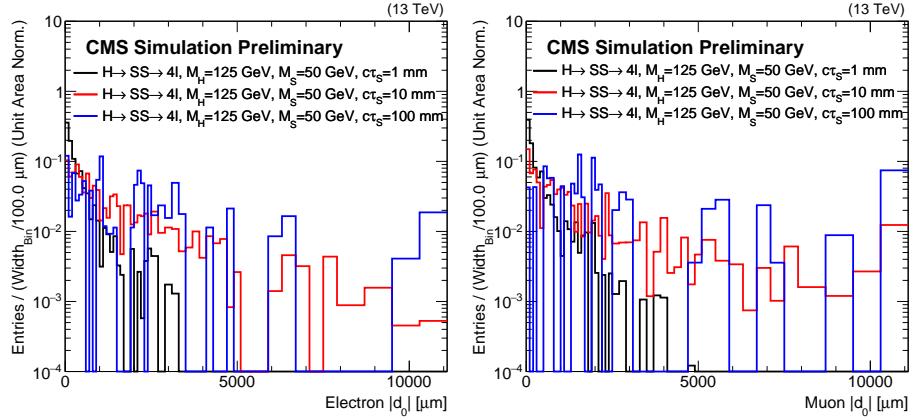


Figure 50: left: The electron  $|d_0|$  distribution for 2018 Higgs events that pass the  $e\mu$  preselection criteria. right: The muon  $|d_0|$  distribution for 2018 Higgs events that pass the  $e\mu$  preselection criteria. The plots are shown for a single Higgs mass and 3 different proper decay lengths, where each histogram is normalized to unity. In all of the histograms, the last bin includes the overflow. All of the corrections from Section 4 are applied.

1308 Figures 51 and 52 show the cumulative number of events that pass the  $e\mu$ ,  $ee$ , and  $\mu\mu$  preselec-  
 1309 tion criteria for Higgs signal events with different proper decay lengths and masses.

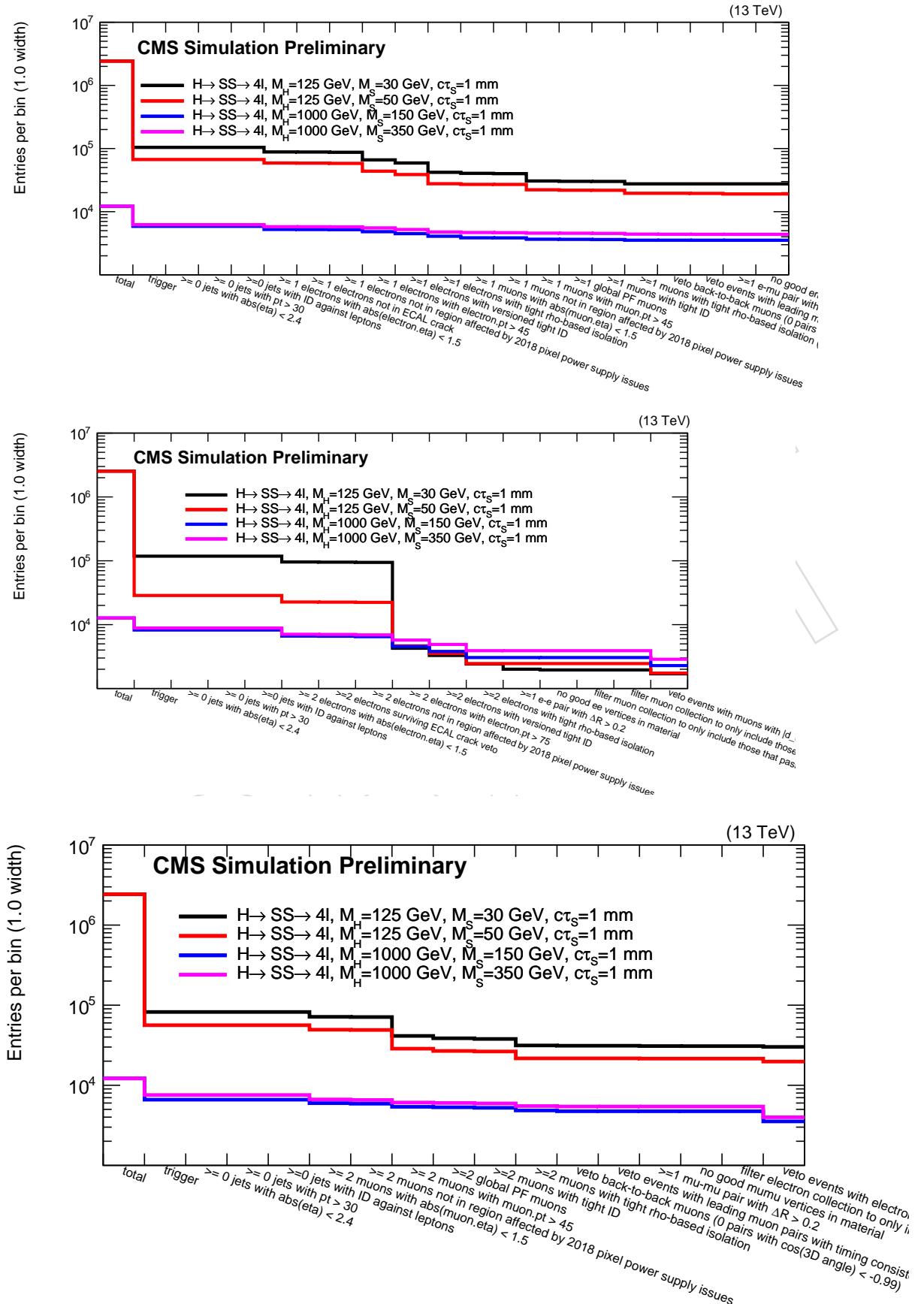


Figure 51: The cumulative number of events that pass each criterion in the  $e\mu$  (upper),  $ee$  (middle), and  $\mu\mu$  (lower) preselection, using 2018 signal simulation. Several Higgs and scalar masses are shown. The jet criteria do not exclude any events and are simply an artifact of the analysis framework.

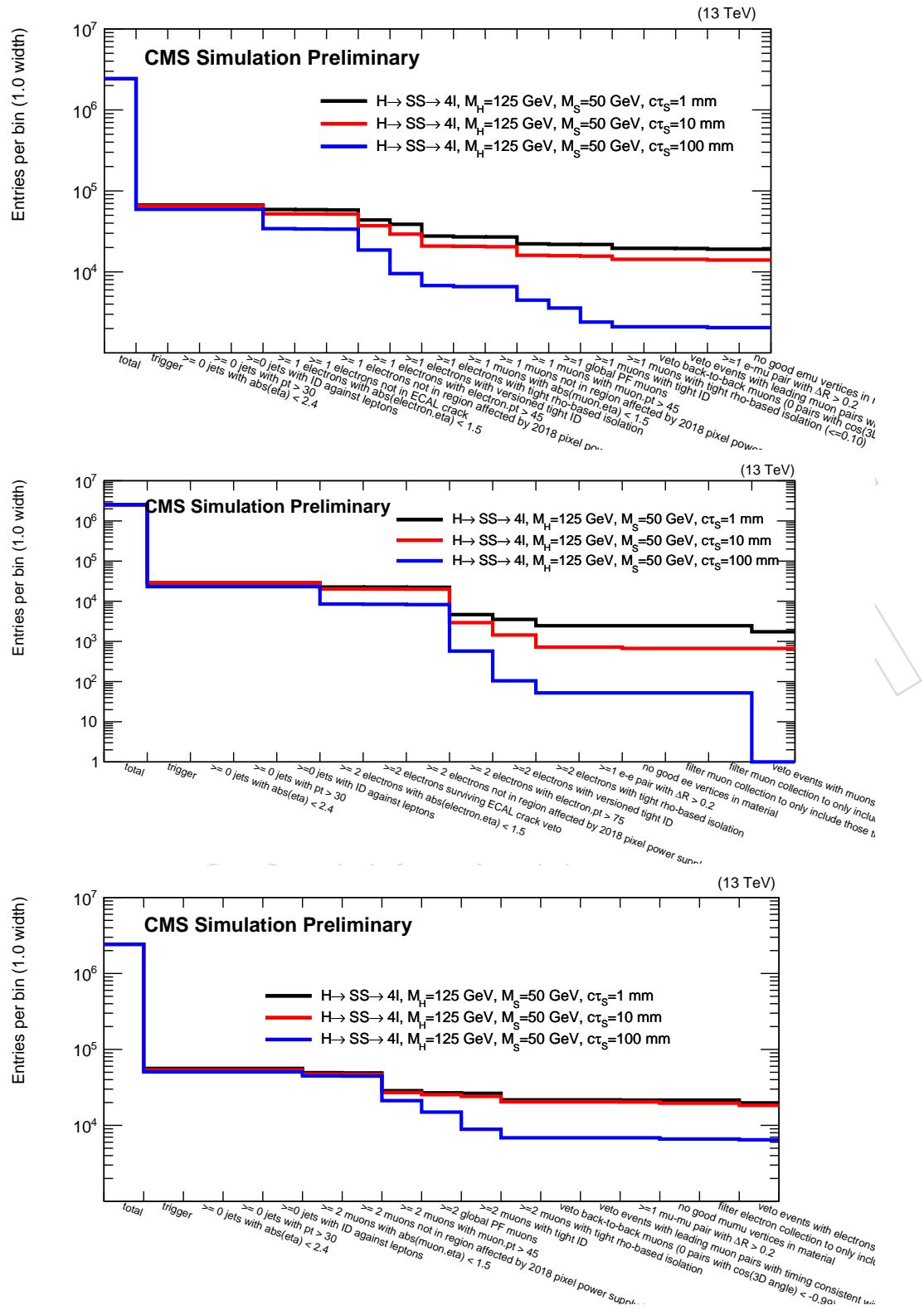


Figure 52: The cumulative number of events that pass each criterion in the  $e\mu$  (upper),  $ee$  (middle), and  $\mu\mu$  (lower) preselection, using 2018 signal simulation. Several scalar proper decay lengths are shown. The jet criteria do not exclude any events and are simply an artifact of the analysis framework.

### D.3 Modified isolation

Figure 53 shows the muon isolation pileup correction term, for the standard muon isolation and for the modified muon isolation, for  $t\bar{t}$  simulation that passes the  $e\mu$  preselection. The standard isolation pileup correction term develops a long tail as  $|d_0|$  increases, while the custom isolation pileup correction is independent of  $|d_0|$ .

Figure 54 shows the electron  $|d_0|$  versus the muon  $|d_0|$ , for  $t\bar{t}$  simulated events that pass the  $e\mu$  preselection and where at least one lepton comes from a heavy-flavor meson. The plot on the left uses the standard isolation, and the plot on the right uses the modified isolation. From these plots, we can see that the custom isolation rejects more leptons from heavy-flavor mesons, especially in the inclusive signal region, which contains 28  $t\bar{t}$  events when the standard isolation is used and only 2 events when the modified isolation is used.

Figure 55 shows the muon custom isolation distribution for  $t\bar{t}$  background simulation and signal simulation. From this plot, we chose a muon working point of isolation  $< 0.10$ .

The modified isolation reduces the heavy-flavor background by a factor of 10 without significantly altering the signal acceptance.

Figure 56 shows the modified isolation distributions for electrons and muons in data and simulated background events in the 2018  $e\mu$  prompt control region with the isolation cuts removed. The agreement between data and simulation is generally within 20%.

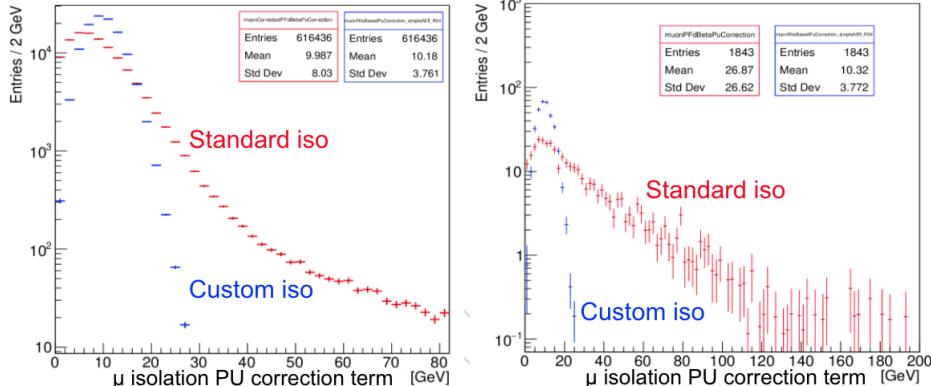


Figure 53: The muon isolation pileup correction term, for the standard muon isolation and for the modified muon isolation, for  $t\bar{t}$  simulation that passes the  $e\mu$  preselection in 2018 conditions. The plot on the left is for muon  $|d_0| < 100 \mu\text{m}$ , and the plot on the right is for  $500 < |d_0| < 1000 \mu\text{m}$ .

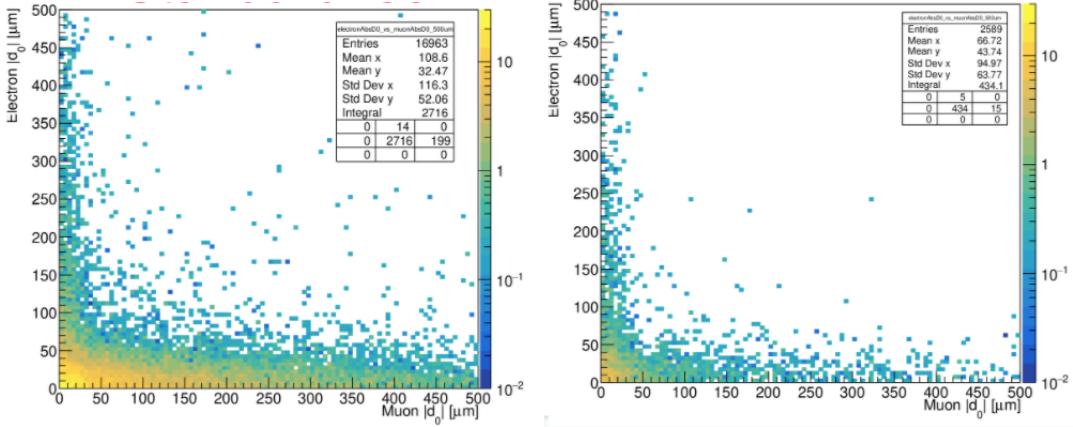


Figure 54: The electron  $|d_0|$  versus the muon  $|d_0|$ , for  $t\bar{t}$  simulated events that pass the  $e\mu$  preselection and where at least one lepton comes from a heavy-flavor meson. The plot on the left uses the standard isolation, and the plot on the right uses the modified isolation.

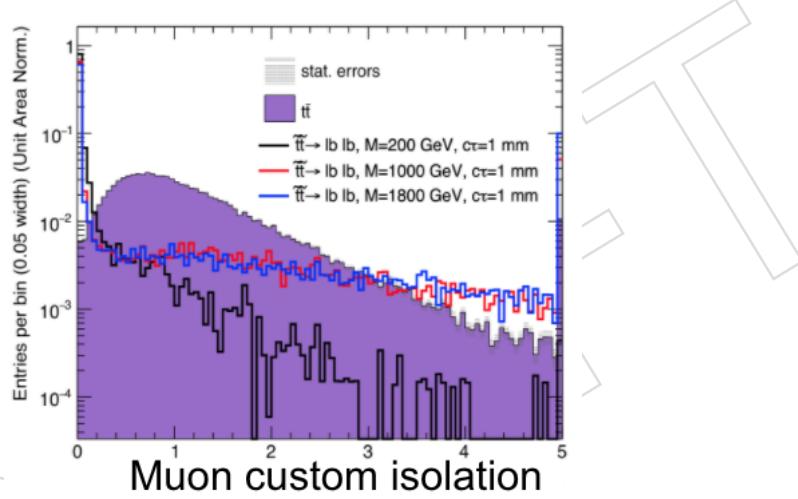


Figure 55: The muon custom isolation distribution for  $t\bar{t}$  background simulation and signal simulation in 2018 conditions.

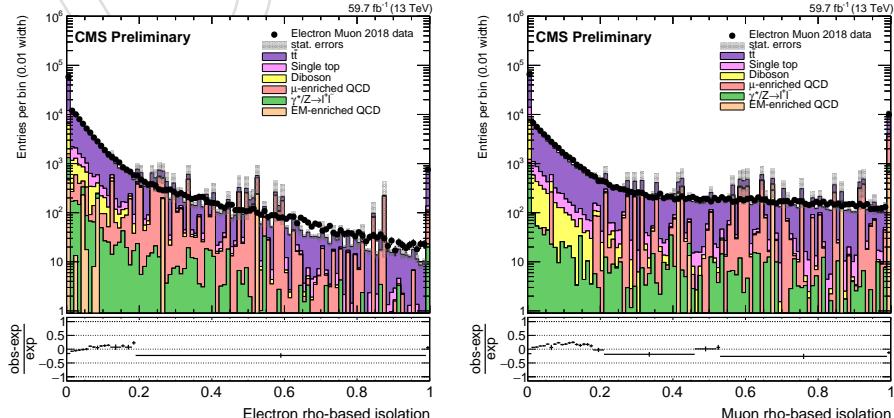


Figure 56: The electron (left) and muon (right) modified isolation distributions in data and simulated background events that pass the 2018  $e\mu$  prompt control region with the isolation cuts removed

**D.4 Prompt control regions**

- Figures 57 and 58 show the  $p_T$ ,  $\eta$ , and  $|d_0|$  distributions of the leptons in the  $e\mu$  prompt control region for 2017 and 2018 data and MC simulation, respectively. Figures 59 and 60 shows the  $p_T$ ,  $\eta$ , and  $|d_0|$  distributions of the electrons in the  $ee$  prompt control region for 2017 and 2018 data and MC simulation, respectively. Likewise, Figs. 61 and 62 show the  $p_T$ ,  $\eta$ , and  $|d_0|$  distributions of the muons in the  $\mu\mu$  prompt control region for 2017 and 2018 data and MC simulation, respectively.
- Figures 63, 64, and 65 show some distributions for the dilepton pairs in the  $e\mu$  channel in 2016, 2017, and 2018 data and MC simulation, respectively. Figures 66, 67, and 68 show the some distributions for the dielectron pairs in the  $ee$  channel for 2016, 2017, and 2018 data and MC simulation, respectively. Likewise, Figs. 69, 70, and 71 show some distributions for the dimuon pairs in the  $\mu\mu$  channel for 2016, 2017, and 2018 data and MC simulation, respectively.

DRAFT

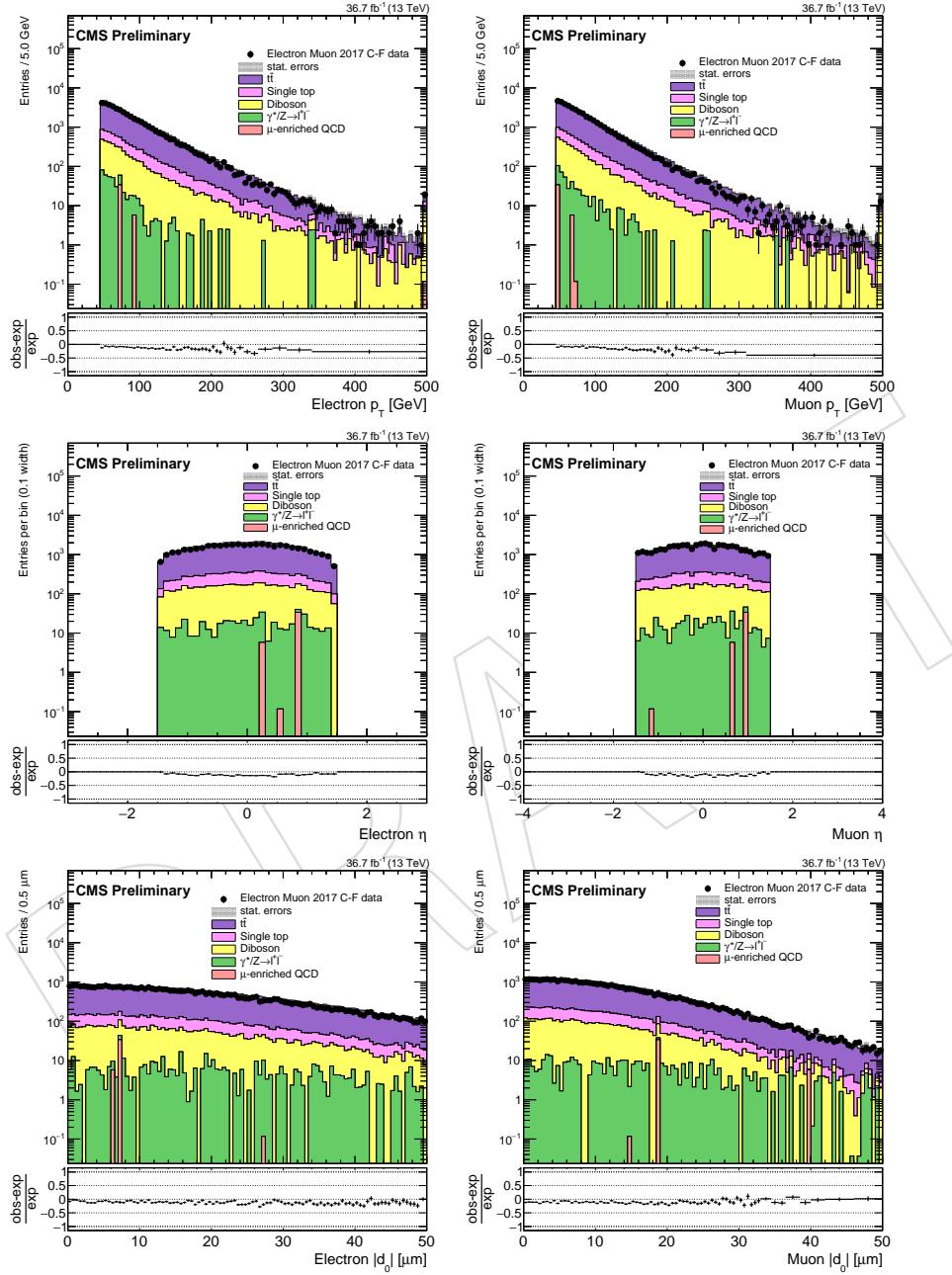


Figure 57: The lepton  $p_T$  (upper),  $\eta$  (middle), and  $|d_0|$  (lower) distributions in the  $e\mu$  prompt control region, for electrons (left) and muons (right), for 2017 data and MC simulation. The rightmost bin in each plot contains the overflow entries.

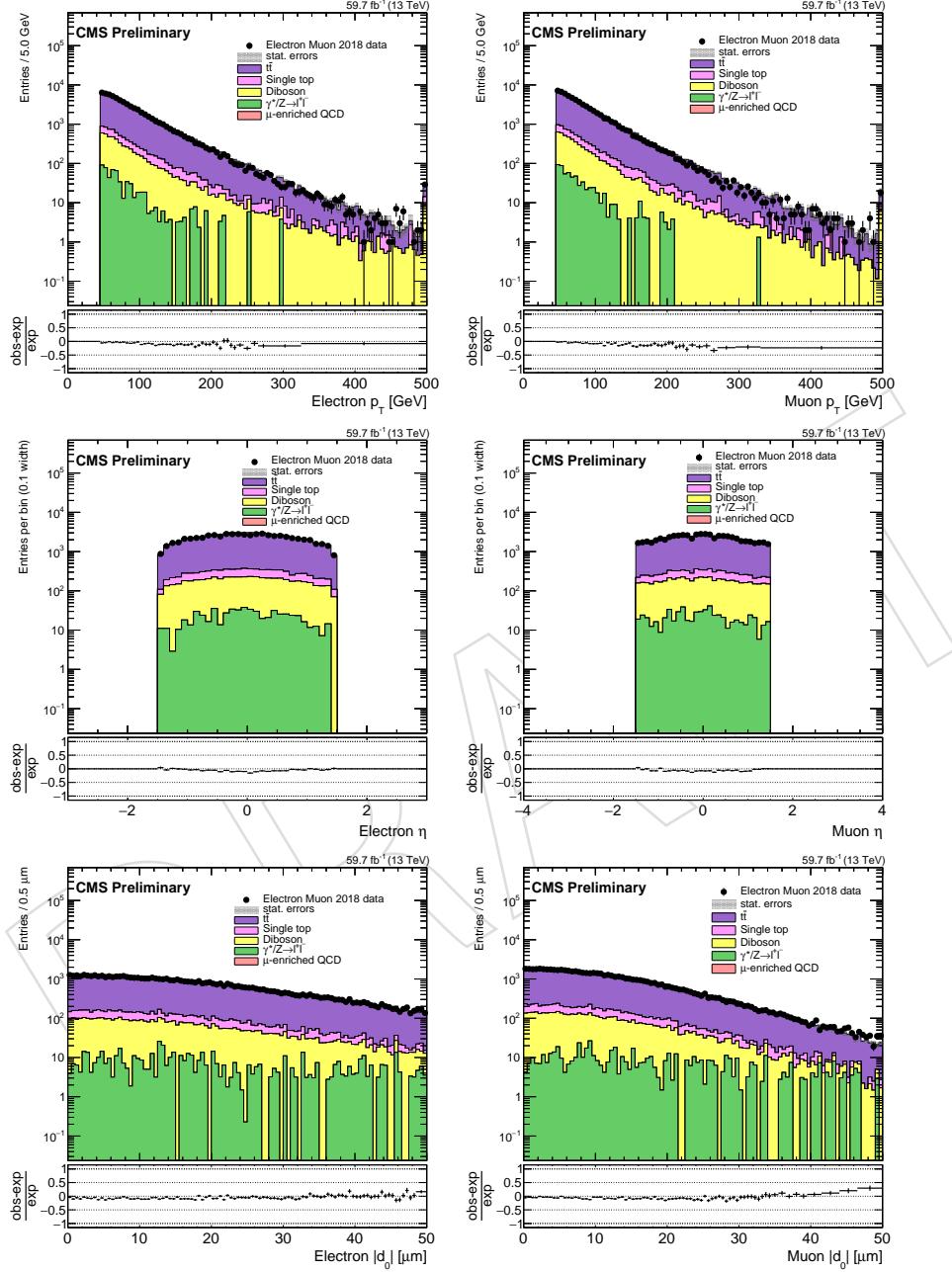


Figure 58: The lepton  $p_T$  (upper),  $\eta$  (middle), and  $|d_0|$  (lower) distributions in the  $e\mu$  prompt control region, for electrons (left) and muons (right), for 2018 data and MC simulation. The rightmost bin in each plot contains the overflow entries.

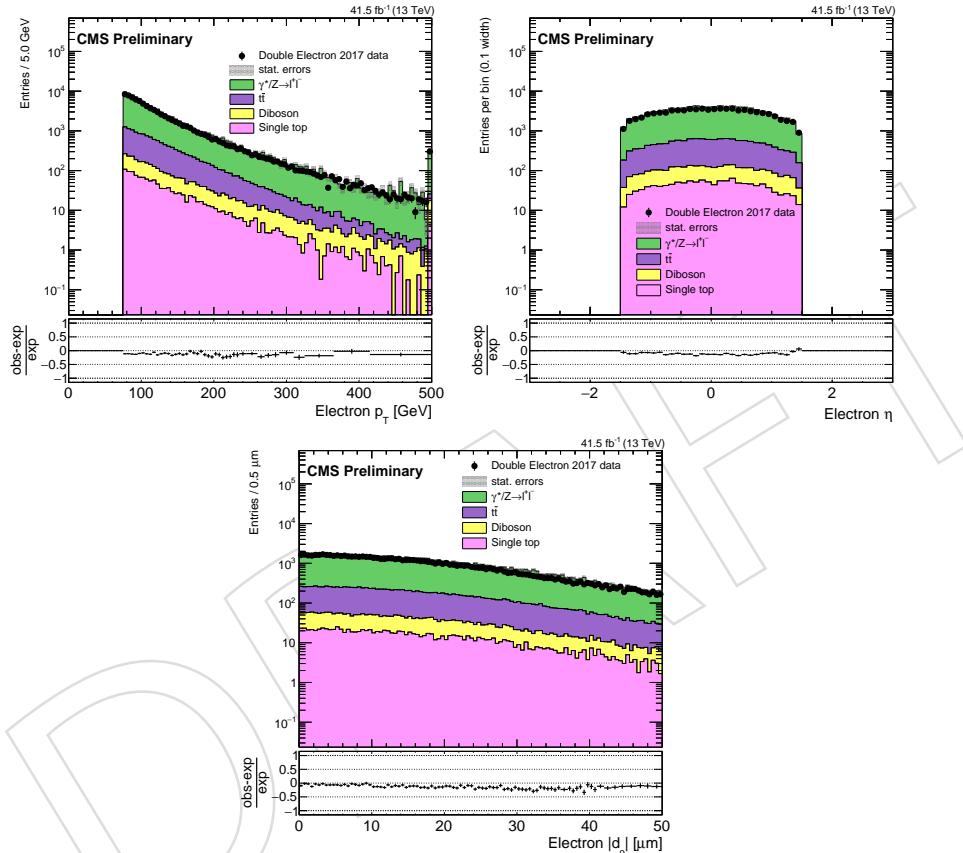


Figure 59: The electron  $p_T$  (upper),  $\eta$  (middle), and  $|d_0|$  (lower) distributions in the ee prompt control region for 2017 data and MC simulation. The rightmost bin in each plot contains the overflow entries.

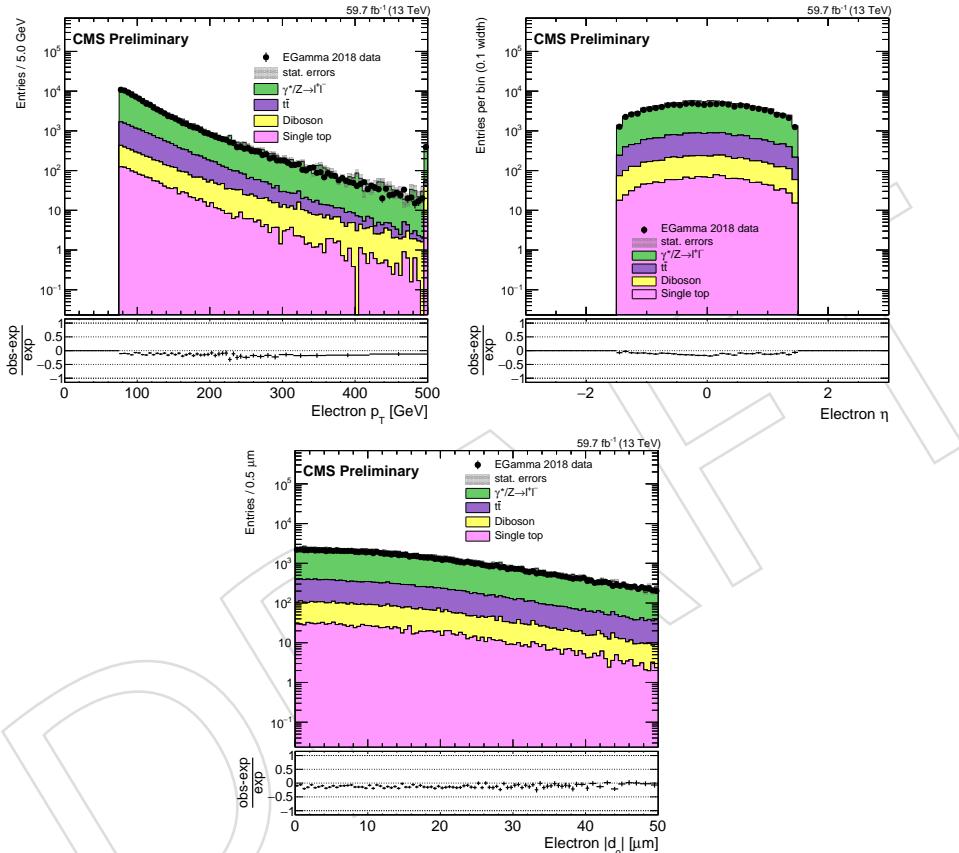


Figure 60: The electron  $p_{\text{T}}$  (upper),  $\eta$  (middle), and  $|d_0|$  (lower) distributions in the ee prompt control region for 2018 data and MC simulation. The rightmost bin in each plot contains the overflow entries.

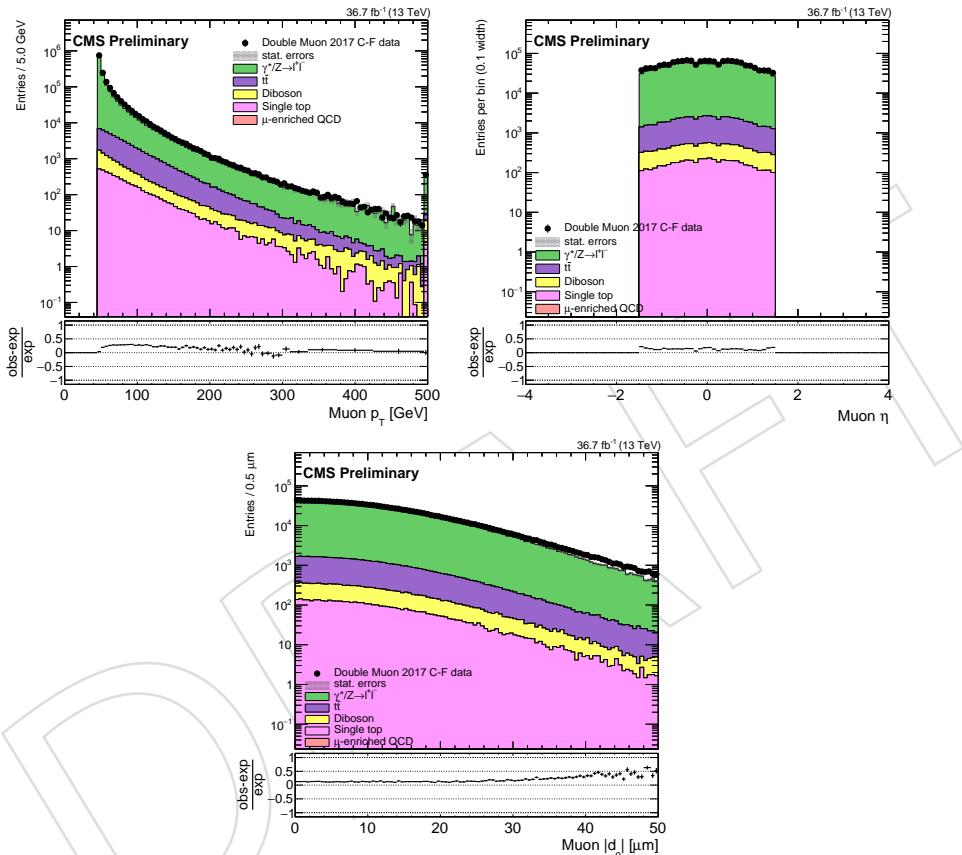


Figure 61: The muon  $p_T$  (upper),  $\eta$  (middle), and  $|d_0|$  (lower) distributions in the  $\mu\mu$  prompt control region for 2017 data and MC simulation. The rightmost bin in each plot contains the overflow entries.

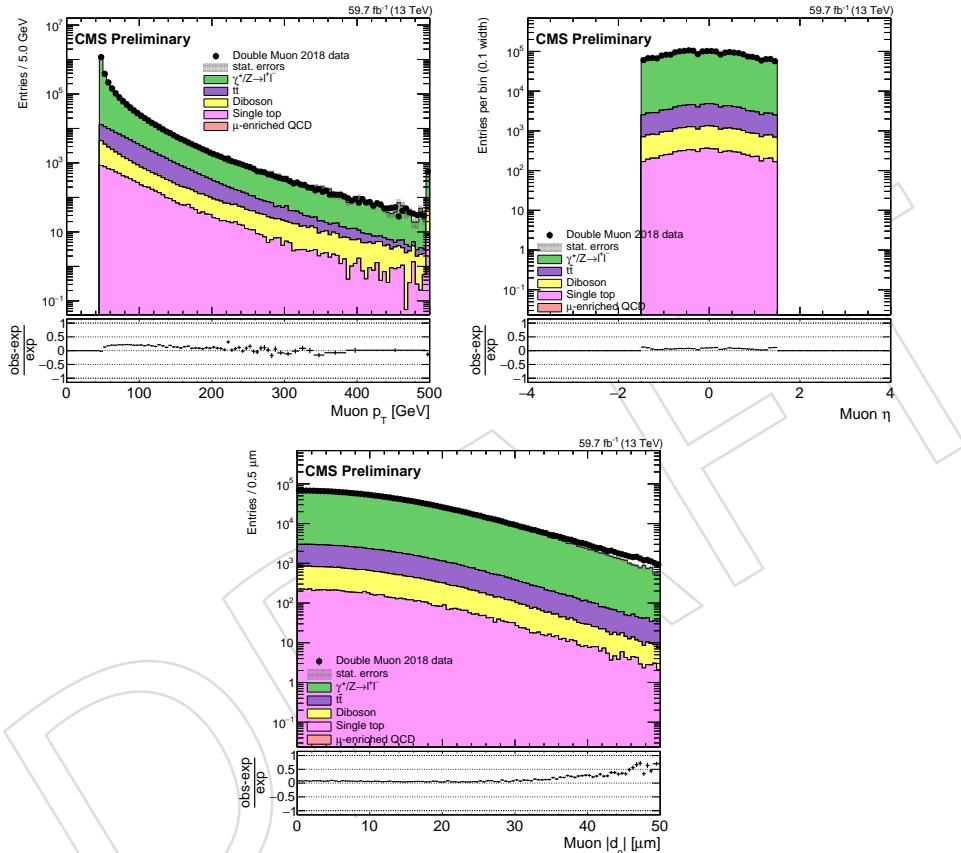


Figure 62: The muon  $p_T$  (upper),  $\eta$  (middle), and  $|d_0|$  (lower) distributions in the  $\mu\mu$  prompt control region for 2018 data and MC simulation. The rightmost bin in each plot contains the overflow entries.

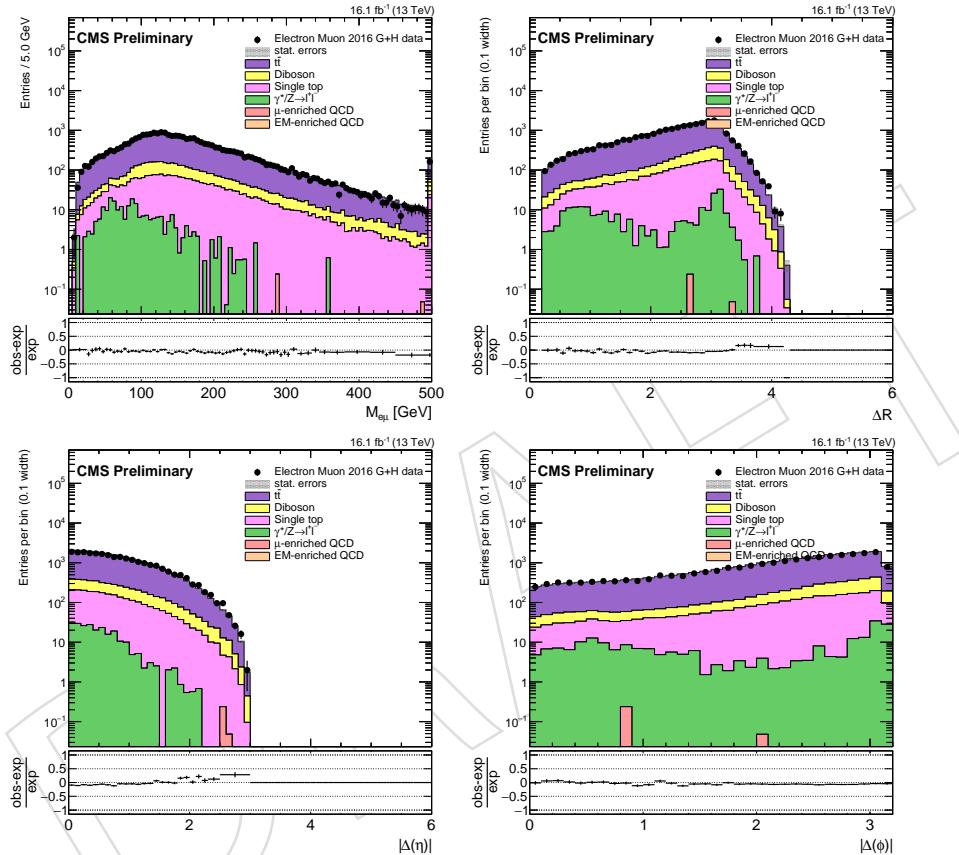


Figure 63: The invariant mass (upper left),  $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$  (upper right),  $\Delta\eta$  (lower left), and  $\Delta\phi$  (lower right) distributions of the electron-muon pairs in the  $e\mu$  prompt control region, for 2016 data and MC simulation. The rightmost bin contains the overflow entries.

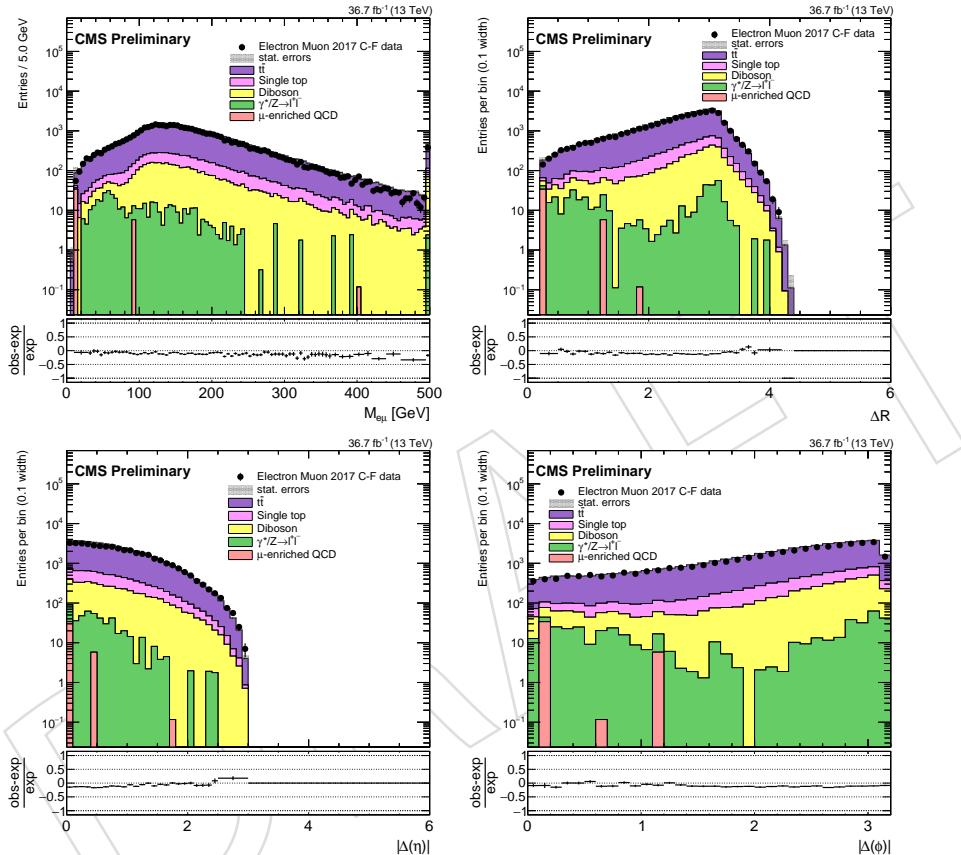


Figure 64: The invariant mass (upper left),  $\Delta R$  (upper right),  $\Delta\eta$  (lower left), and  $\Delta\phi$  (lower right) distributions of the electron-muon pairs in the  $e\mu$  prompt control region, for 2017 data and MC simulation. The rightmost bin contains the overflow entries.

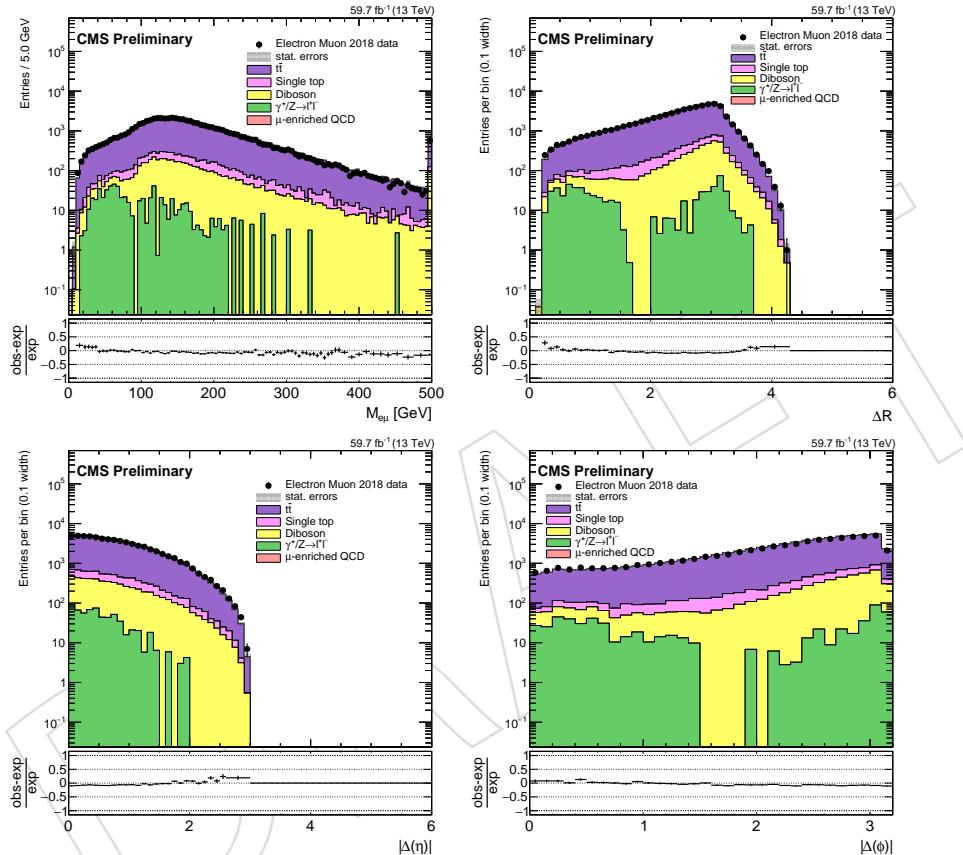


Figure 65: The invariant mass (upper left),  $\Delta R$  (upper right),  $\Delta\eta$  (lower left), and  $\Delta\phi$  (lower right) distributions of the electron-muon pairs in the  $e\mu$  prompt control region, for 2018 data and MC simulation. The rightmost bin contains the overflow entries.

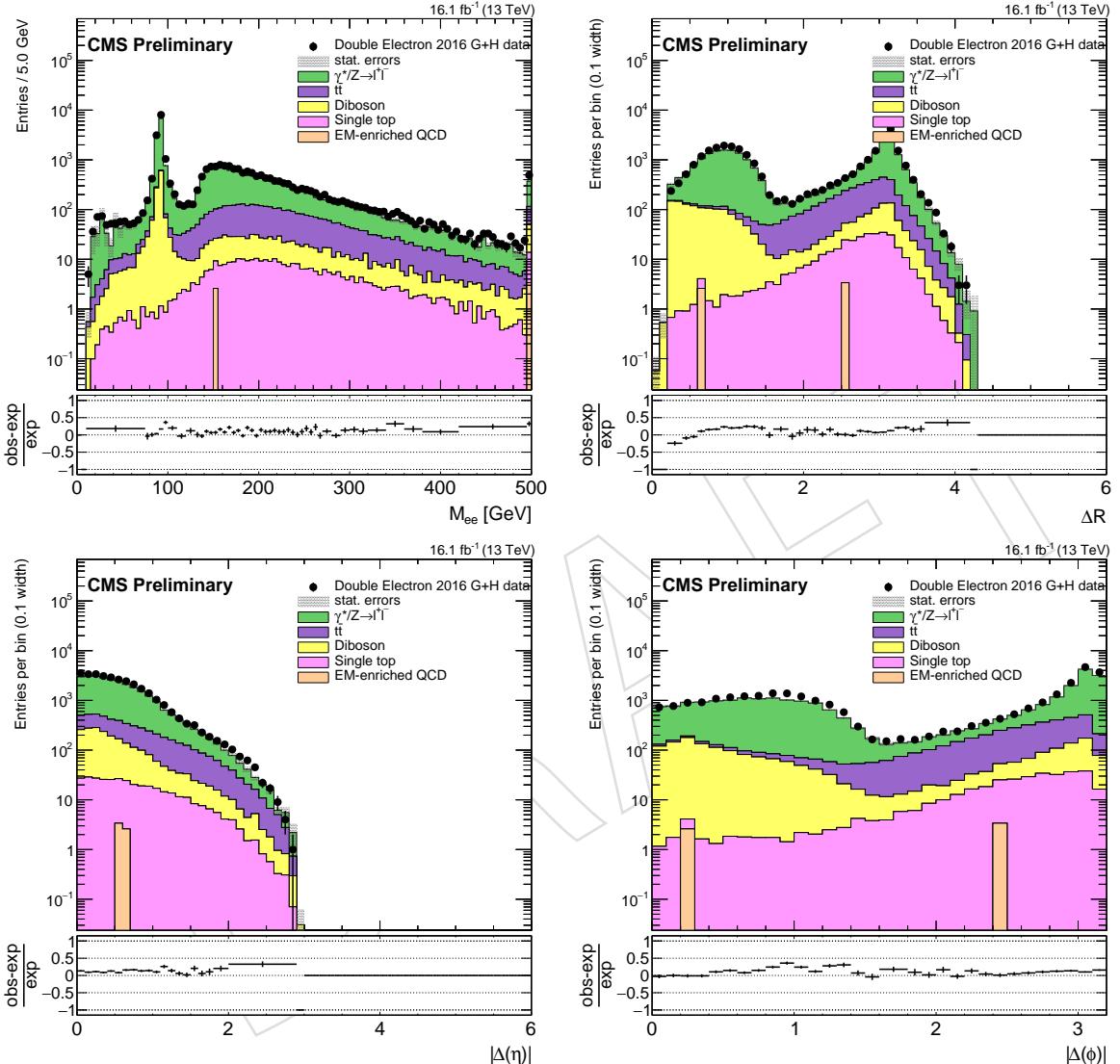


Figure 66: The invariant mass (upper left),  $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$  (upper right),  $\Delta\eta$  (lower left), and  $\Delta\phi$  (lower right) distributions of the electron-electron pairs in the ee prompt control region for 2016 data and MC simulation. The rightmost bin contains the overflow entries.

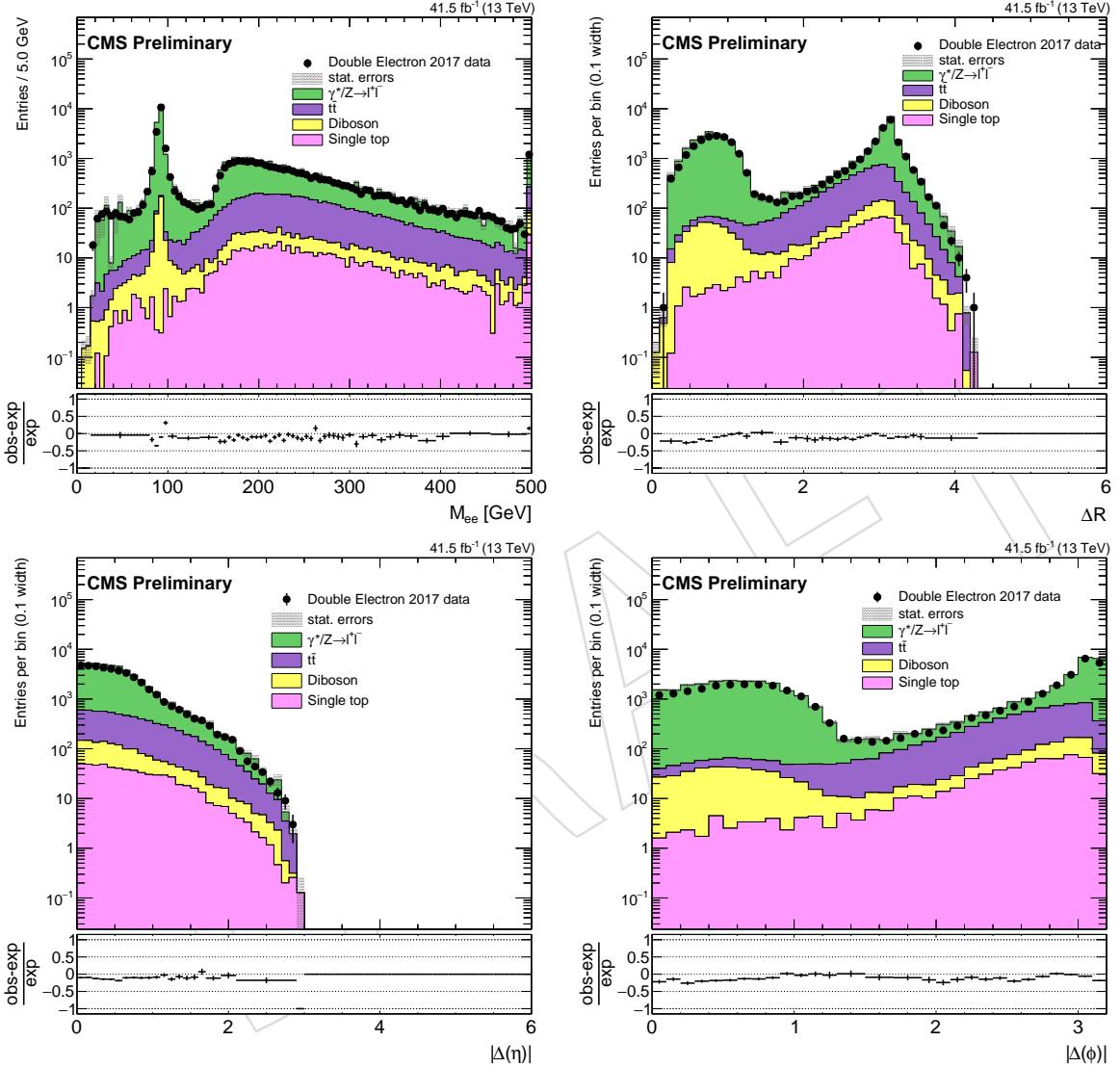


Figure 67: The invariant mass (upper left),  $\Delta R$  (upper right),  $\Delta\eta$  (lower left), and  $\Delta\phi$  (lower right) distributions of the electron-electron pairs in the ee prompt control region for 2017 data and MC simulation. The rightmost bin contains the overflow entries.

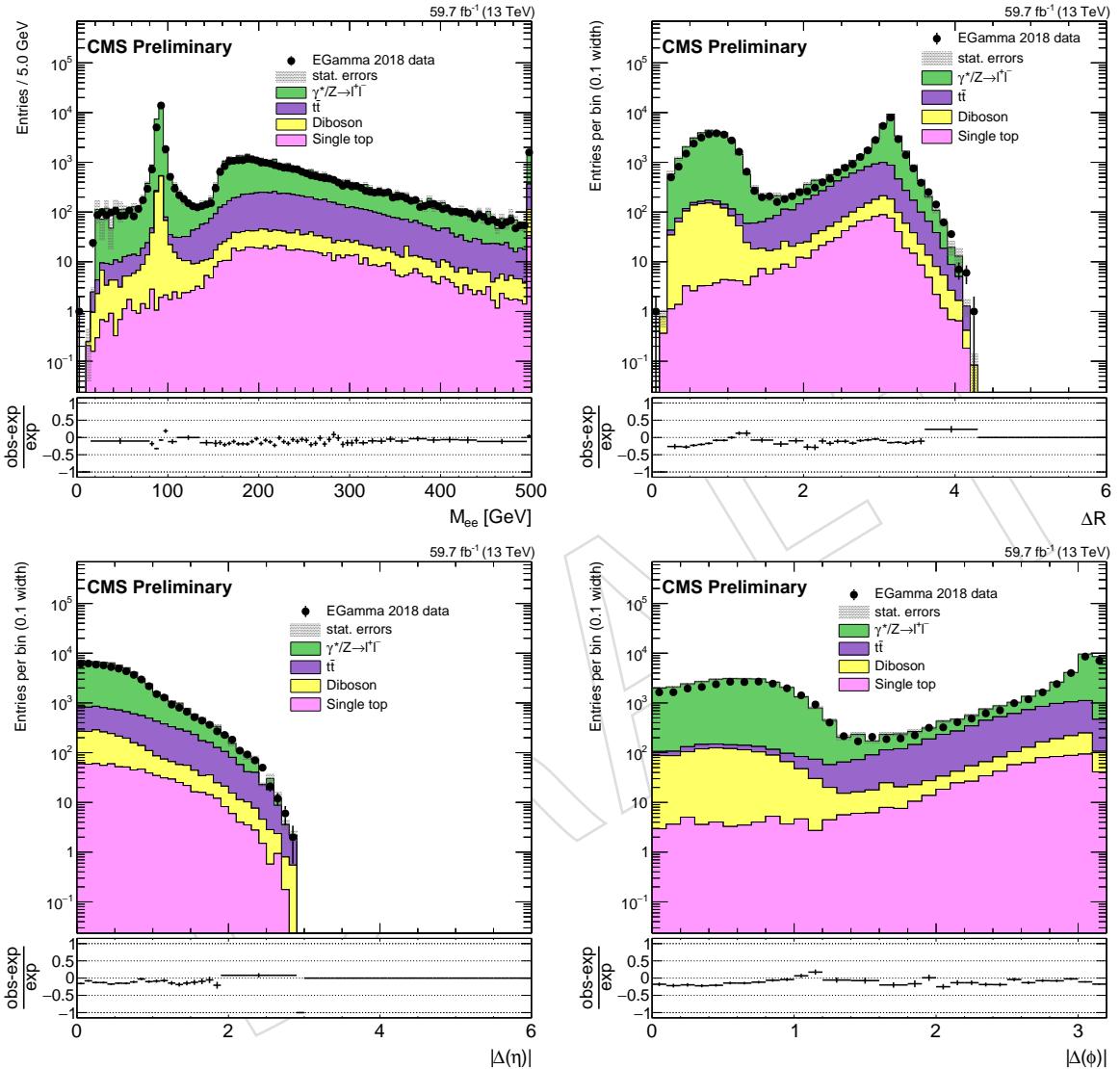


Figure 68: The invariant mass (upper left),  $\Delta R$  (upper right),  $\Delta\eta$  (lower left), and  $\Delta\phi$  (lower right) distributions of the electron-electron pairs in the ee prompt control region for 2018 data and MC simulation. The rightmost bin contains the overflow entries.

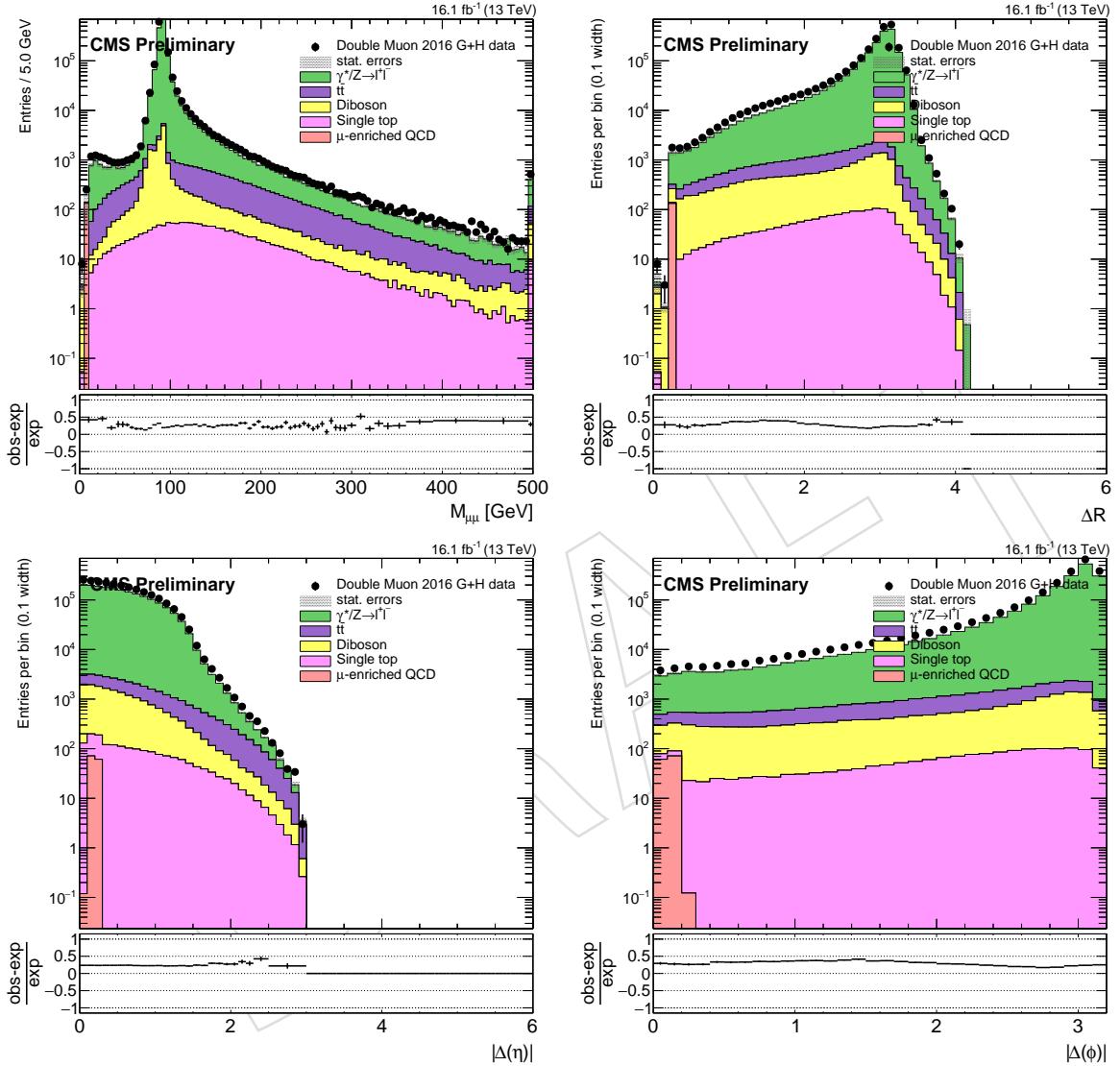


Figure 69: The invariant mass (upper left),  $\Delta R$  (upper right),  $\Delta\eta$  (lower left), and  $\Delta\phi$  (lower right) distributions of the muon-muon pairs in the  $\mu\mu$  prompt control region for 2016 data and MC simulation. The rightmost bin contains the overflow entries.

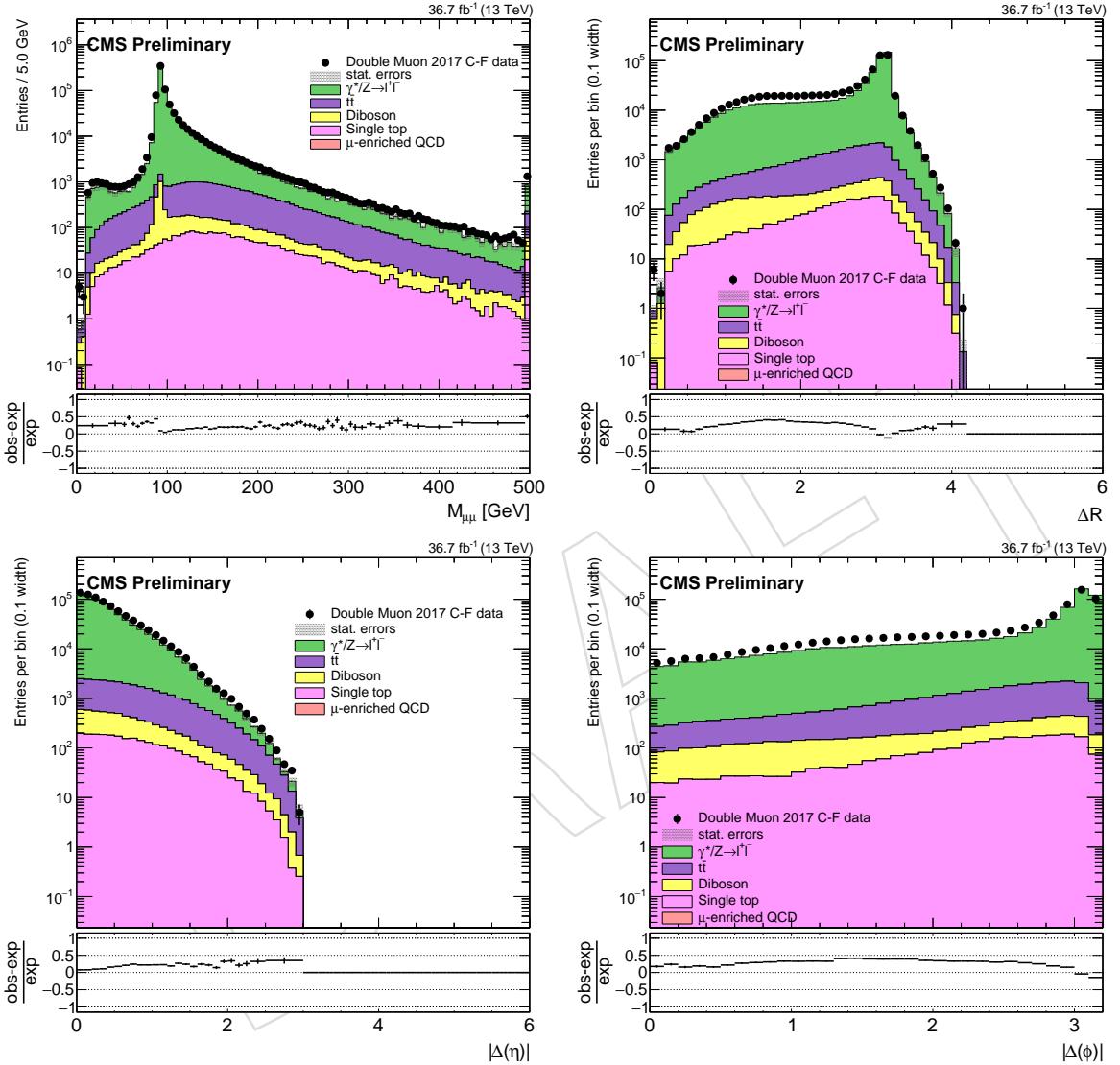


Figure 70: The invariant mass (upper left),  $\Delta R$  (upper right),  $\Delta\eta$  (lower left), and  $\Delta\phi$  (lower right) distributions of the muon-muon pairs in the  $\mu\mu$  prompt control region for 2017 data and MC simulation. The rightmost bin contains the overflow entries.

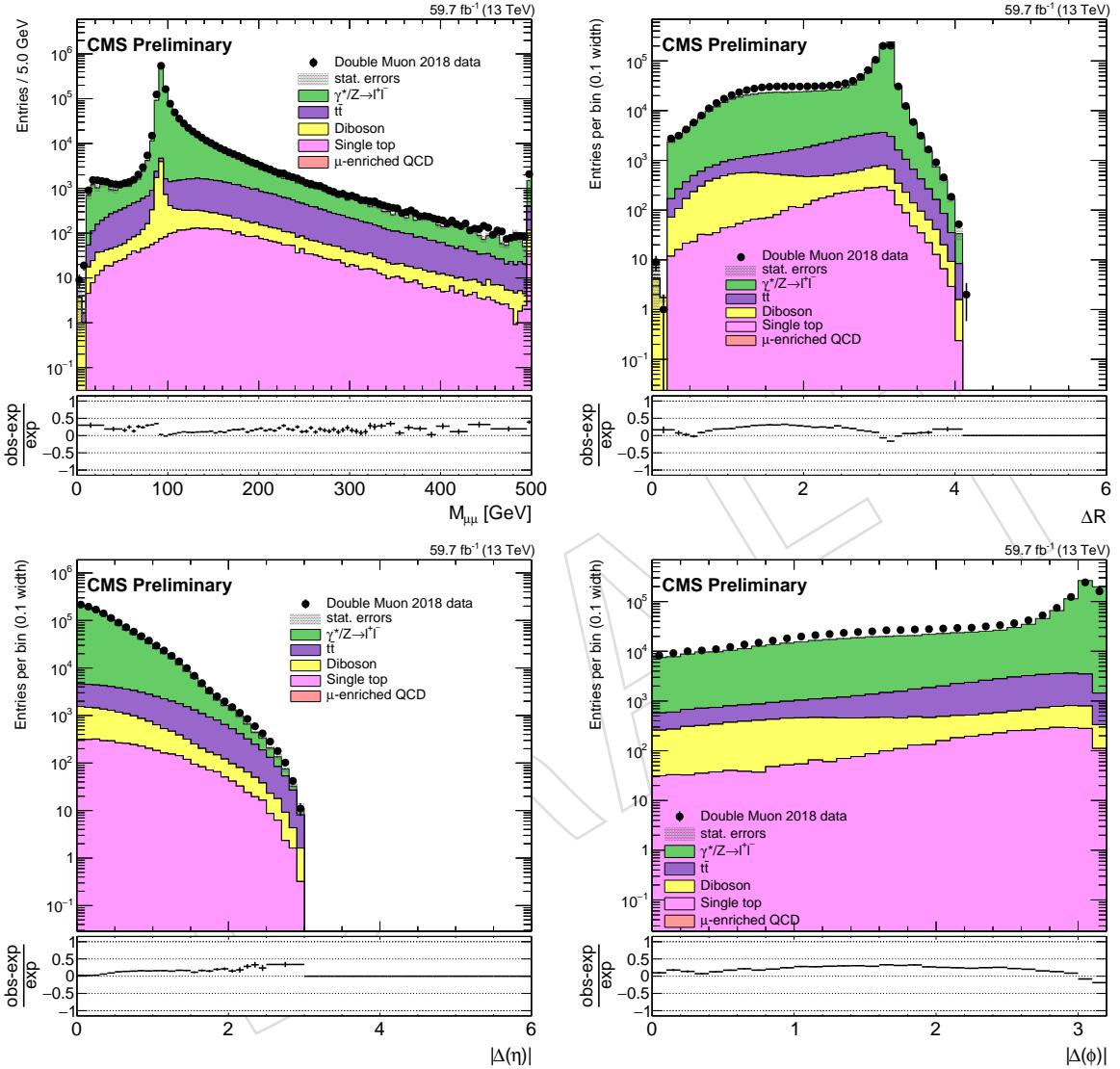


Figure 71: The invariant mass (upper left),  $\Delta R$  (upper right),  $\Delta\eta$  (lower left), and  $\Delta\phi$  (lower right) distributions of the muon-muon pairs in the  $\mu\mu$  prompt control region for 2018 data and MC simulation. The rightmost bin contains the overflow entries.

1340 **D.5 Lepton  $d_0$  corrections**

1341 The fitted  $d_0$  distributions are shown in Figs. 72 and 73. These are shown for the same-flavor  
 1342 channels, as opposed to the plots in Section 4.3, which are shown for the  $e\mu$  channel.

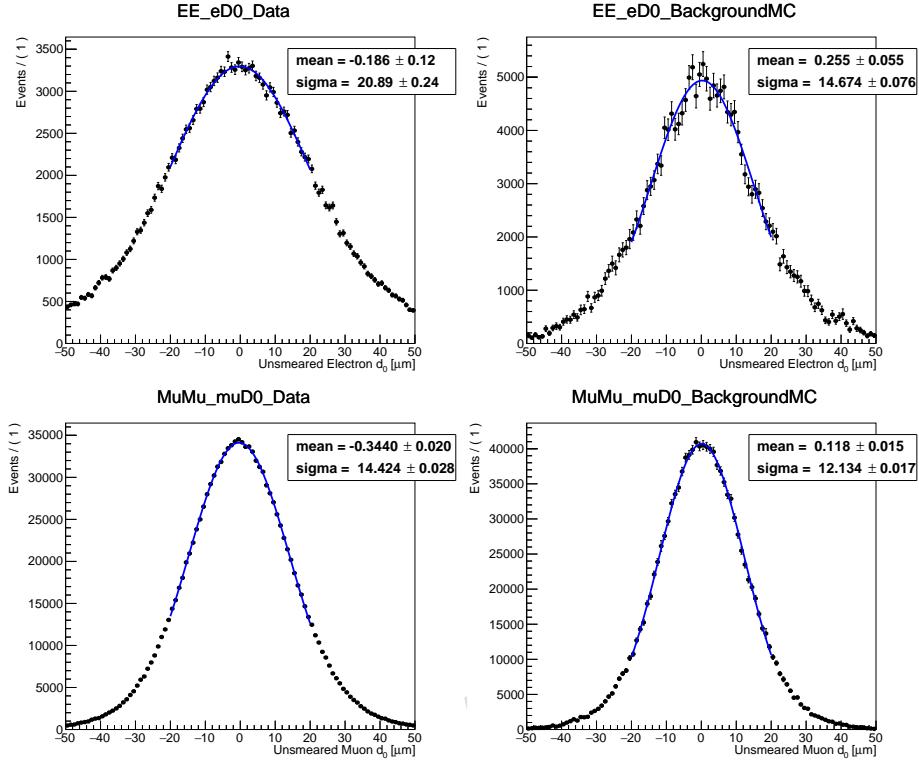


Figure 72: The lepton  $d_0$  distributions with Gaussian fits in data (left) and background MC (right) for electrons (upper) and muons (lower) in the 2017 ee and  $\mu\mu$  prompt control regions. The widths of the Gaussian fits are used to determine the width of the Gaussian distribution used to smear the  $d_0$ .

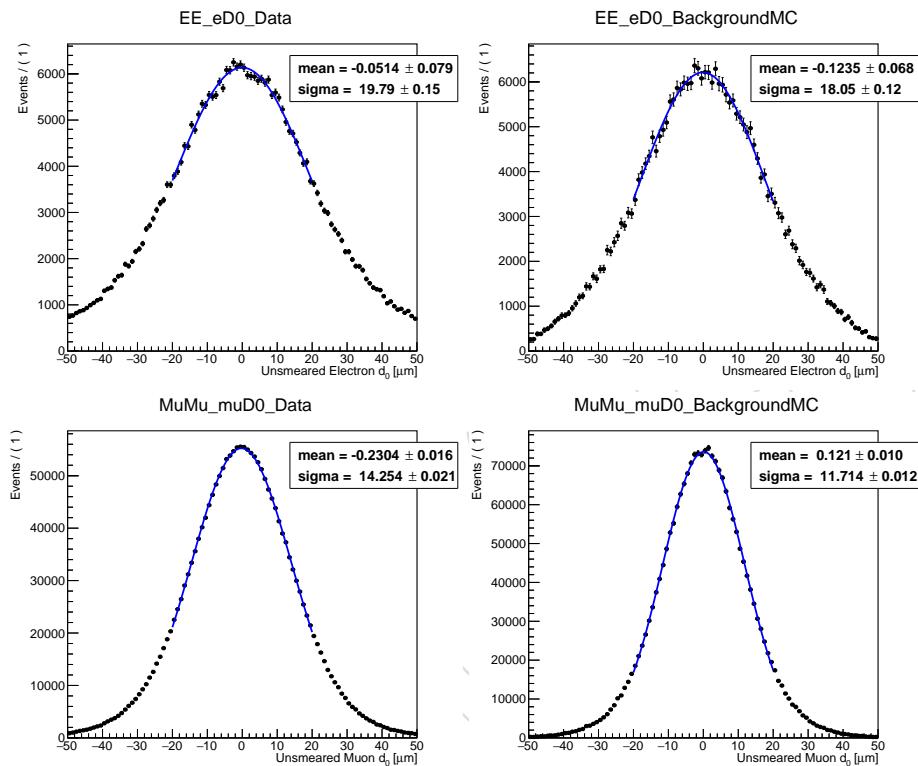


Figure 73: The lepton  $d_0$  distributions with Gaussian fits in data (left) and background MC (right) for electrons (upper) and muons (lower) in the 2018 ee and  $\mu\mu$  prompt control regions. The widths of the Gaussian fits are used to determine the width of the Gaussian distribution used to smear the  $d_0$ .

1343 D.6 Background estimation

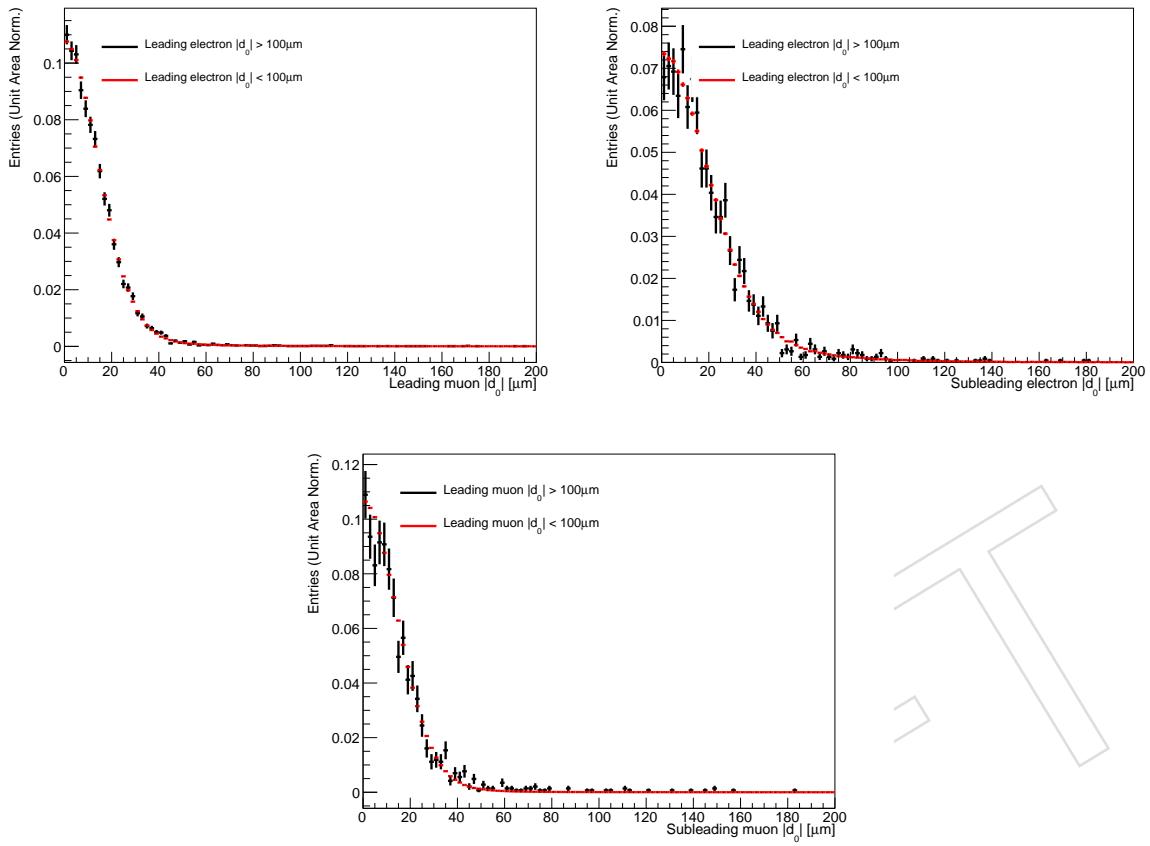


Figure 74: The  $|d_0^b|$  distributions for  $|d_0^a| > 100 \mu\text{m}$  and  $|d_0^a| < 100 \mu\text{m}$ , for background simulation that passes the  $e\mu$  (upper left),  $ee$  (upper right), and  $\mu\mu$  (lower) preselection with 2018 conditions.

<sup>1344</sup> **D.7 Data and background simulation in the inclusive signal region**

<sup>1345</sup> Figures 75, 76, 77, 78, 79, and 80 show data and background simulation in the  $e\mu$  channel in the  
<sup>1346</sup> inclusive SR.

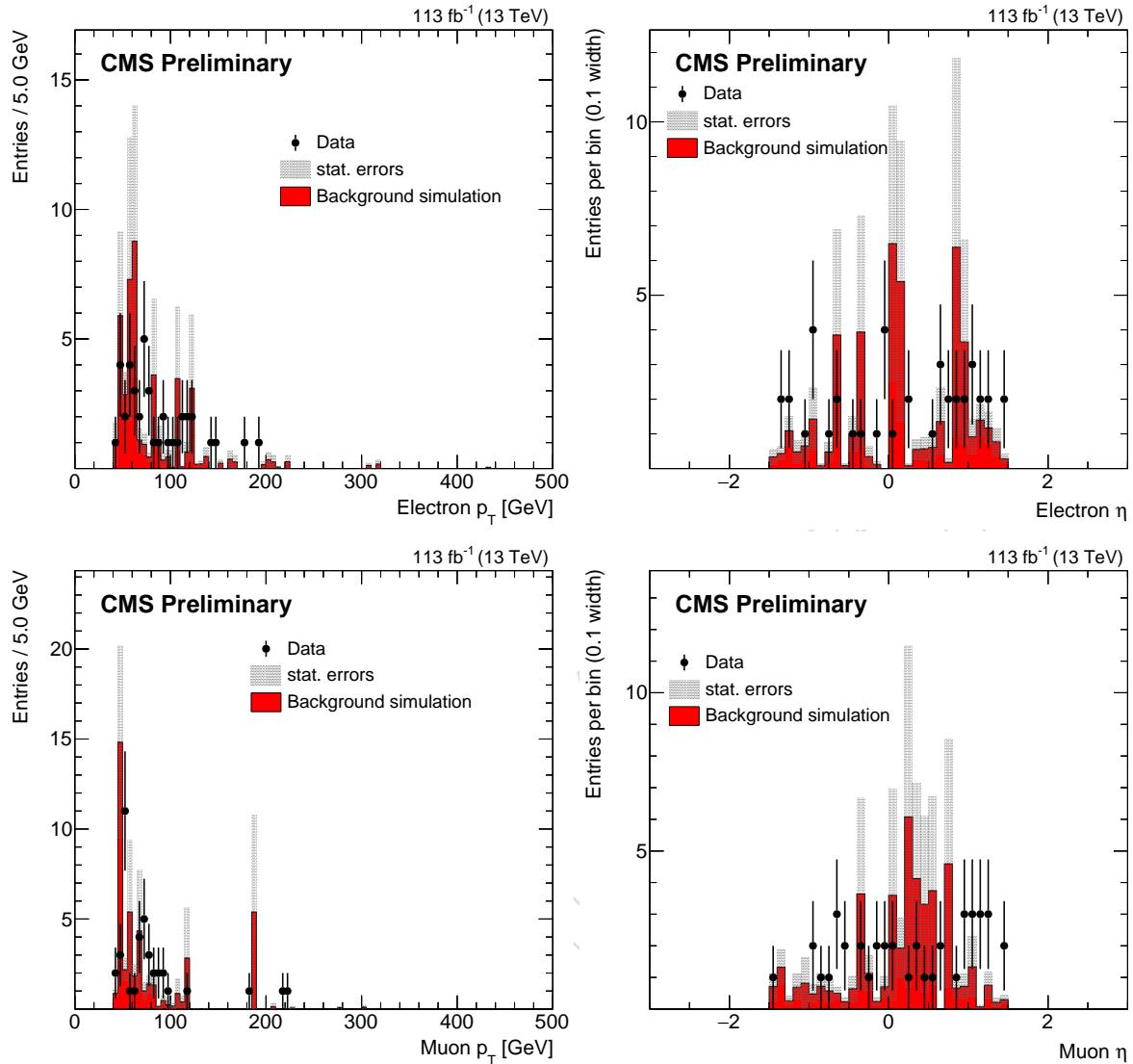


Figure 75: Electron and muon  $p_T$  and  $\eta$  distributions for data and composite background simulation in the  $e\mu$  channel in the inclusive SR. The background simulation is normalized to the ABCD method background prediction.

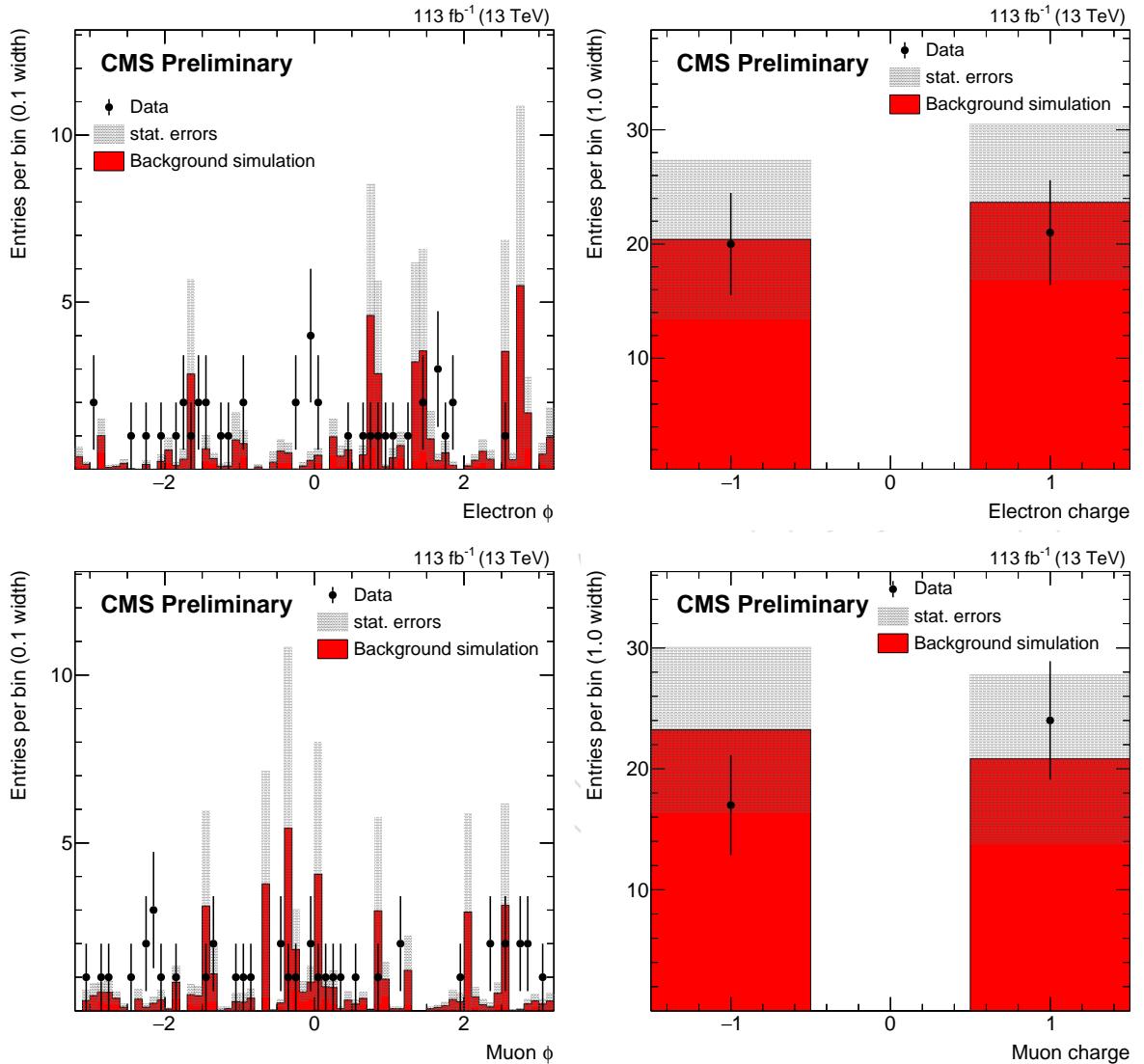


Figure 76: Electron and muon  $\phi$  and charge distributions for data and composite background simulation in the  $e\mu$  channel in the inclusive SR. The background simulation is normalized to the ABCD method background prediction.

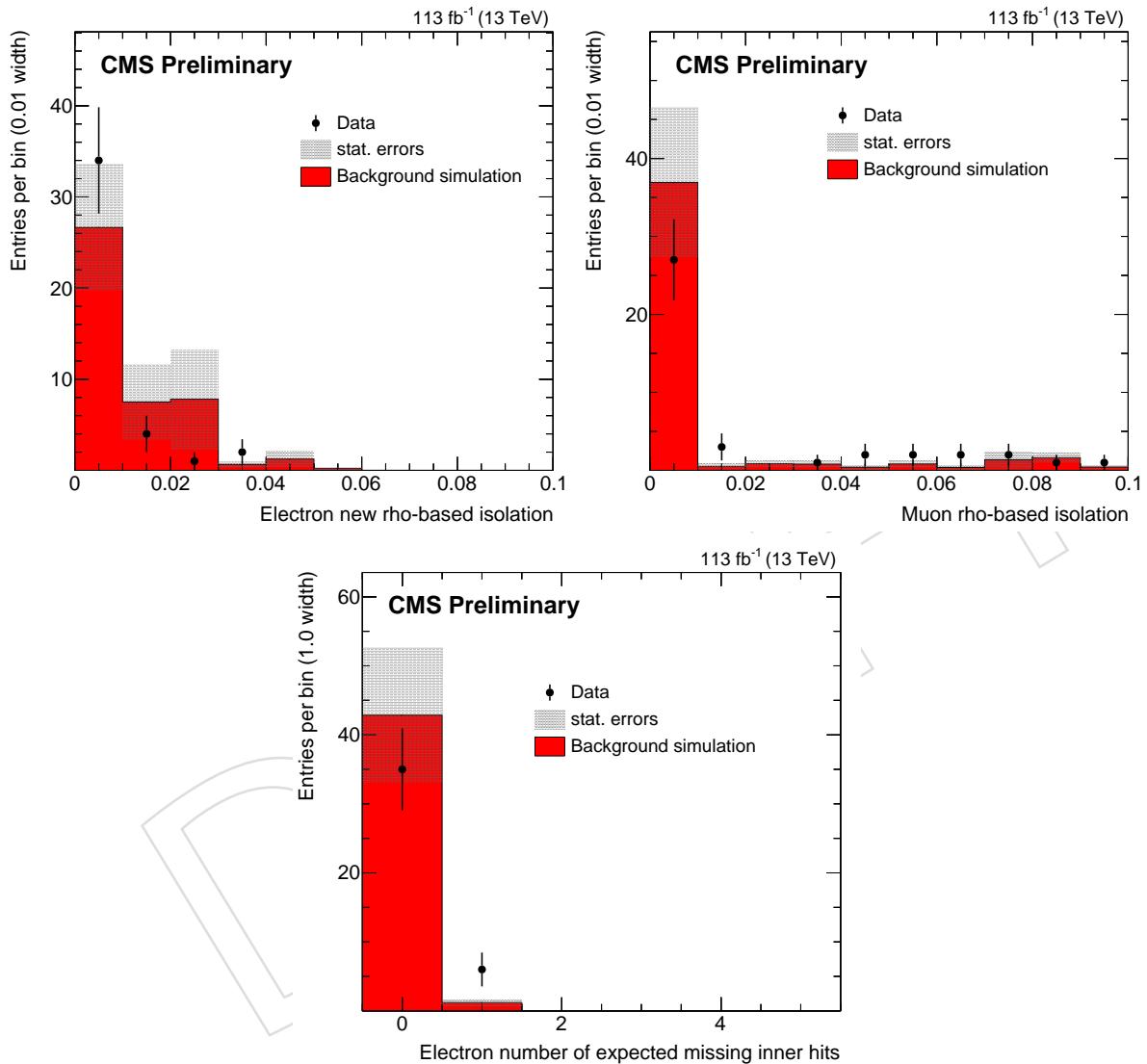


Figure 77: Electron and muon isolation and missing inner hits distributions for data and composite background simulation in the  $e\mu$  channel in the inclusive SR. The background simulation is normalized to the ABCD method background prediction.

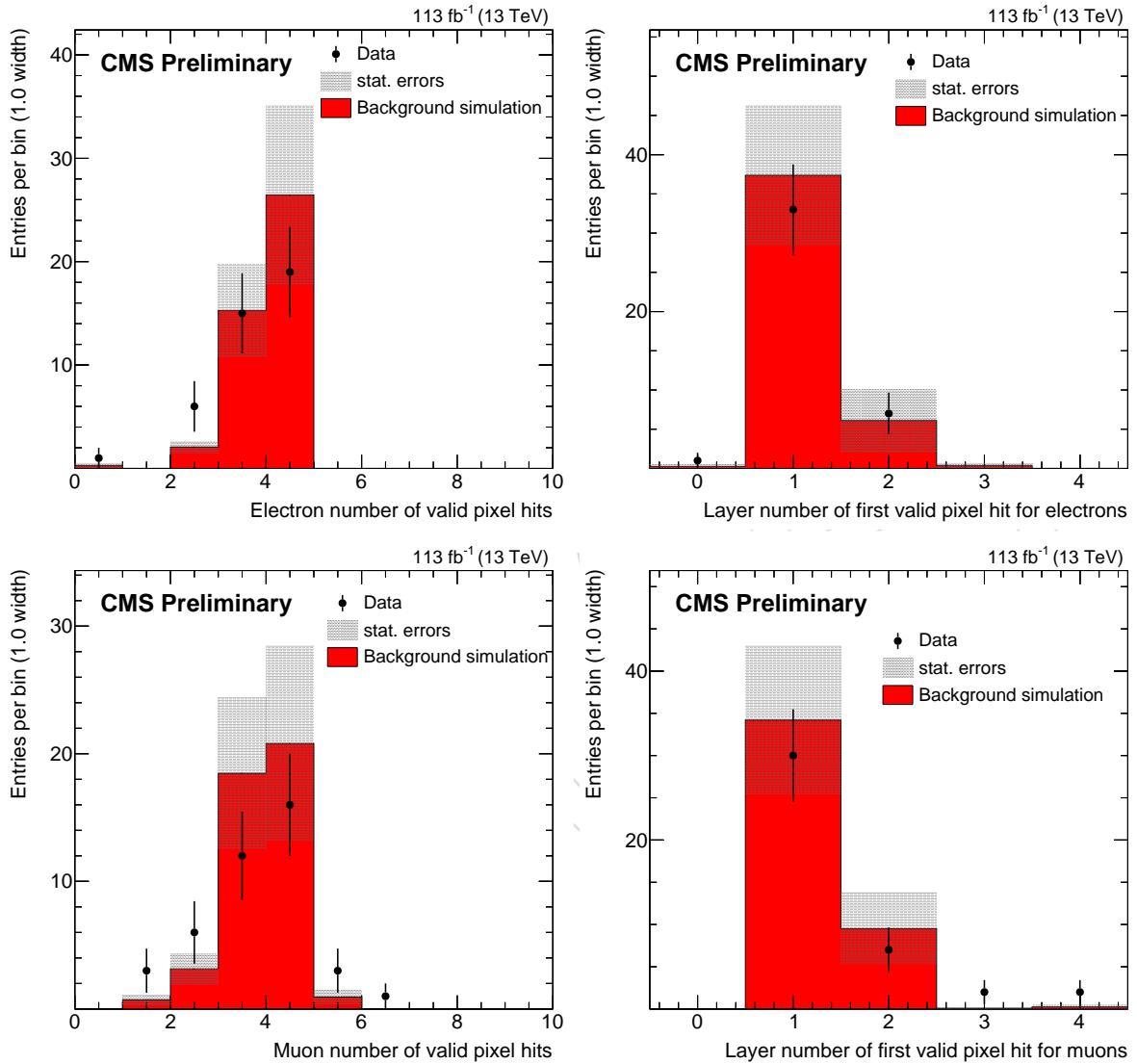


Figure 78: Electron and muon pixel hit distributions for data and composite background simulation in the  $e\mu$  channel in the inclusive SR. The background simulation is normalized to the ABCD method background prediction.

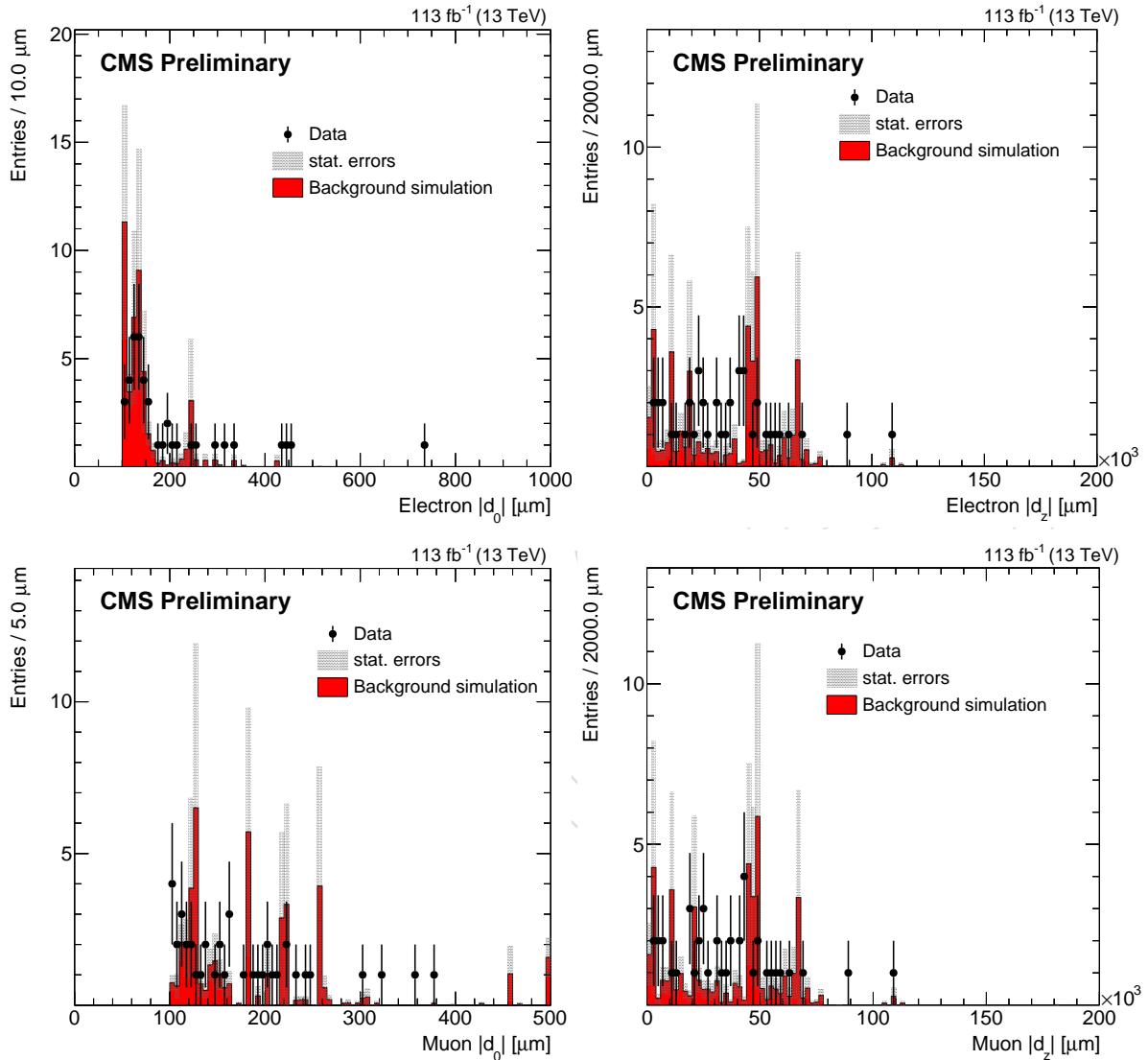


Figure 79: Electron and muon  $|d_0|$  and  $|z_0|$  distributions for data and composite background simulation in the  $e\mu$  channel in the inclusive SR. The background simulation is normalized to the ABCD method background prediction.

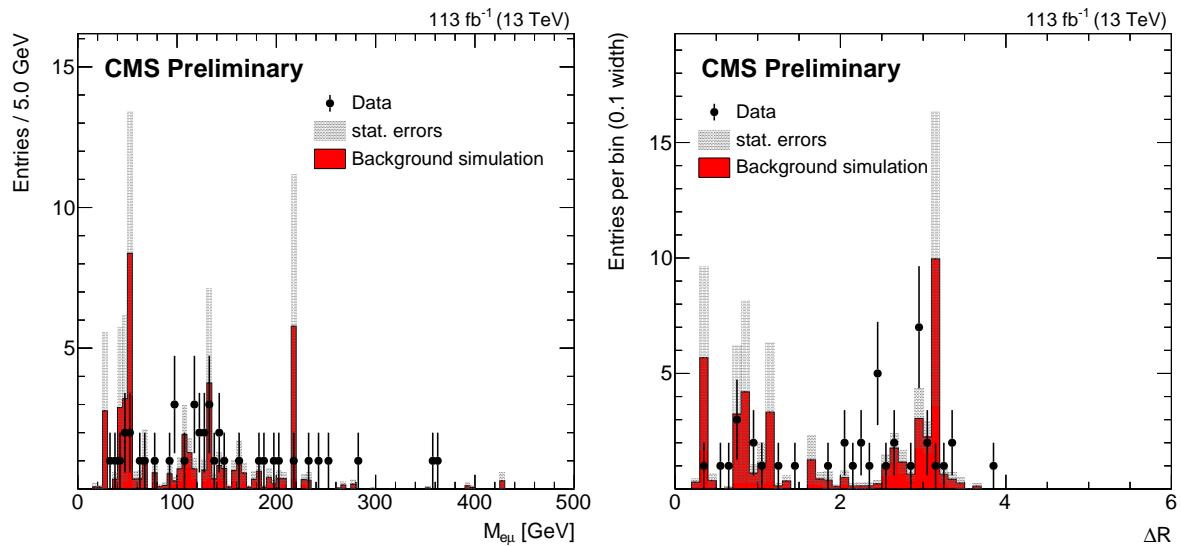


Figure 80: Invariant mass and  $\Delta R$  distributions for data and composite background simulation in the  $e\mu$  channel in the inclusive SR. The background simulation is normalized to the ABCD method background prediction.

<sup>1347</sup> Figures 81, 82, and 83 show data and background simulation in the ee channel in the inclusive  
<sup>1348</sup> SR.

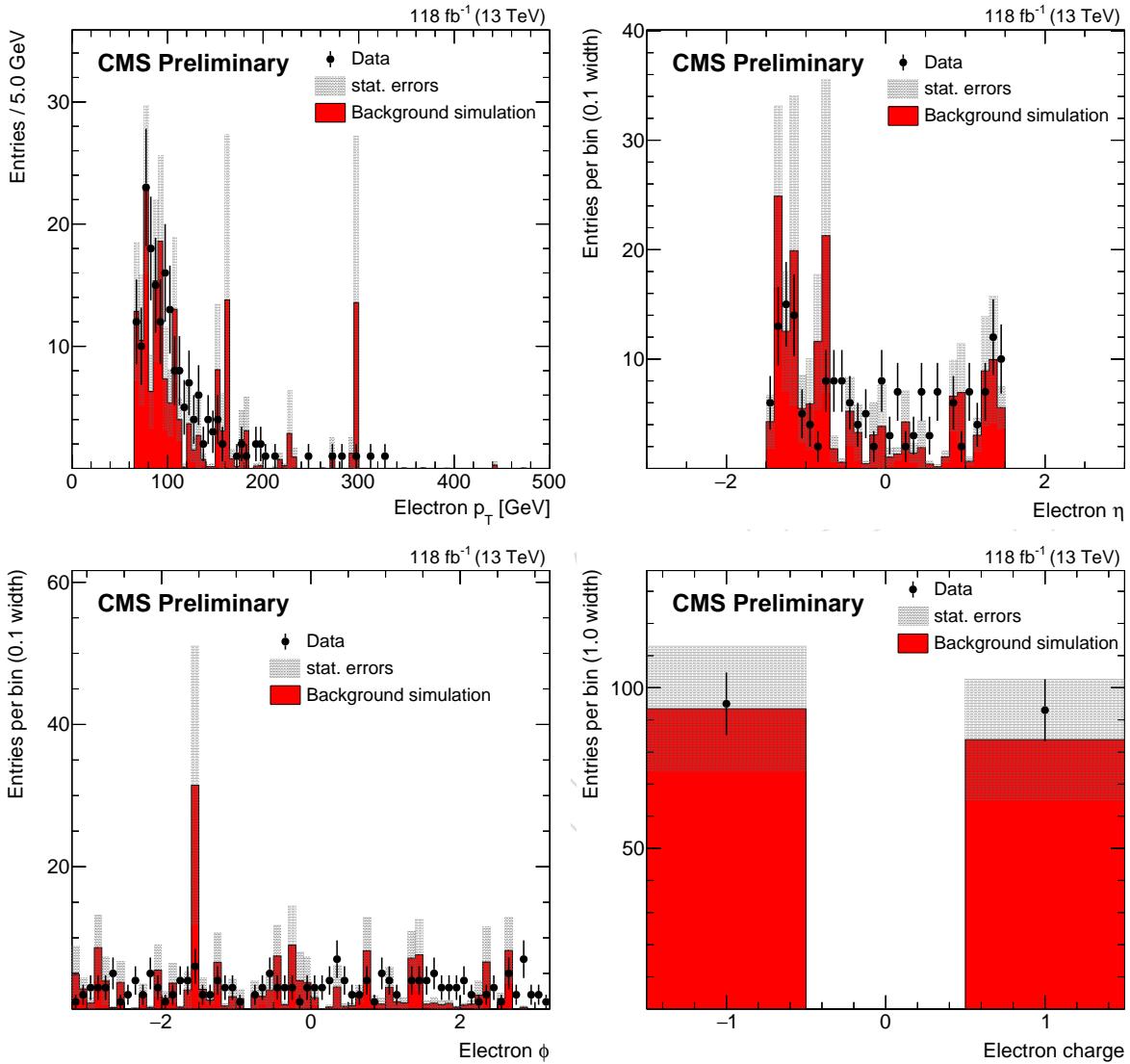


Figure 81: Electron  $p_T$ ,  $\eta$ ,  $\phi$ , and charge distributions for data and composite background simulation in the ee channel in the inclusive SR. The background simulation is normalized to the ABCD method background prediction.

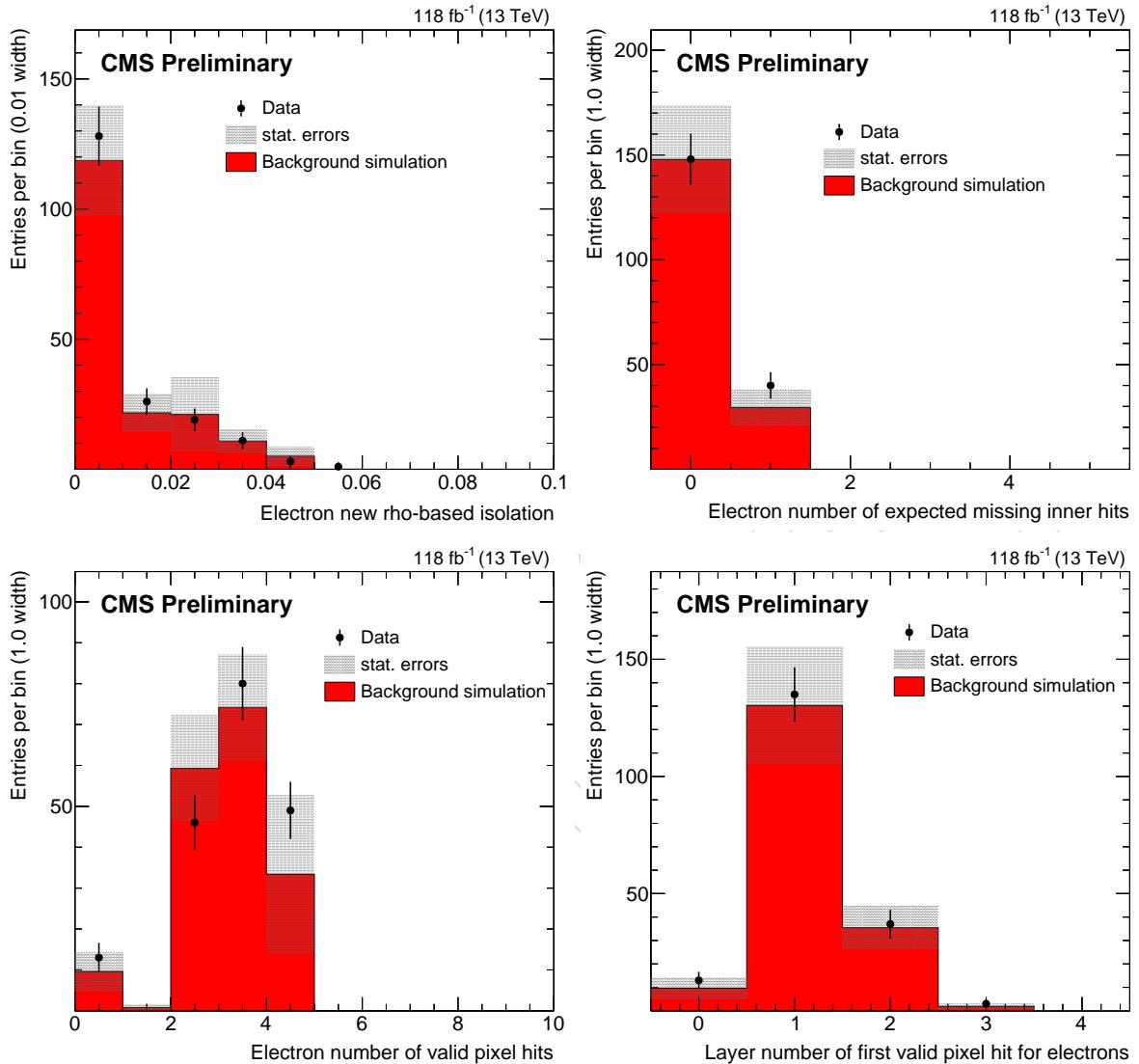


Figure 82: Electron isolation, missing inner hits, and pixel hit distributions for data and composite background simulation in the ee channel in the inclusive SR. The background simulation is normalized to the ABCD method background prediction.

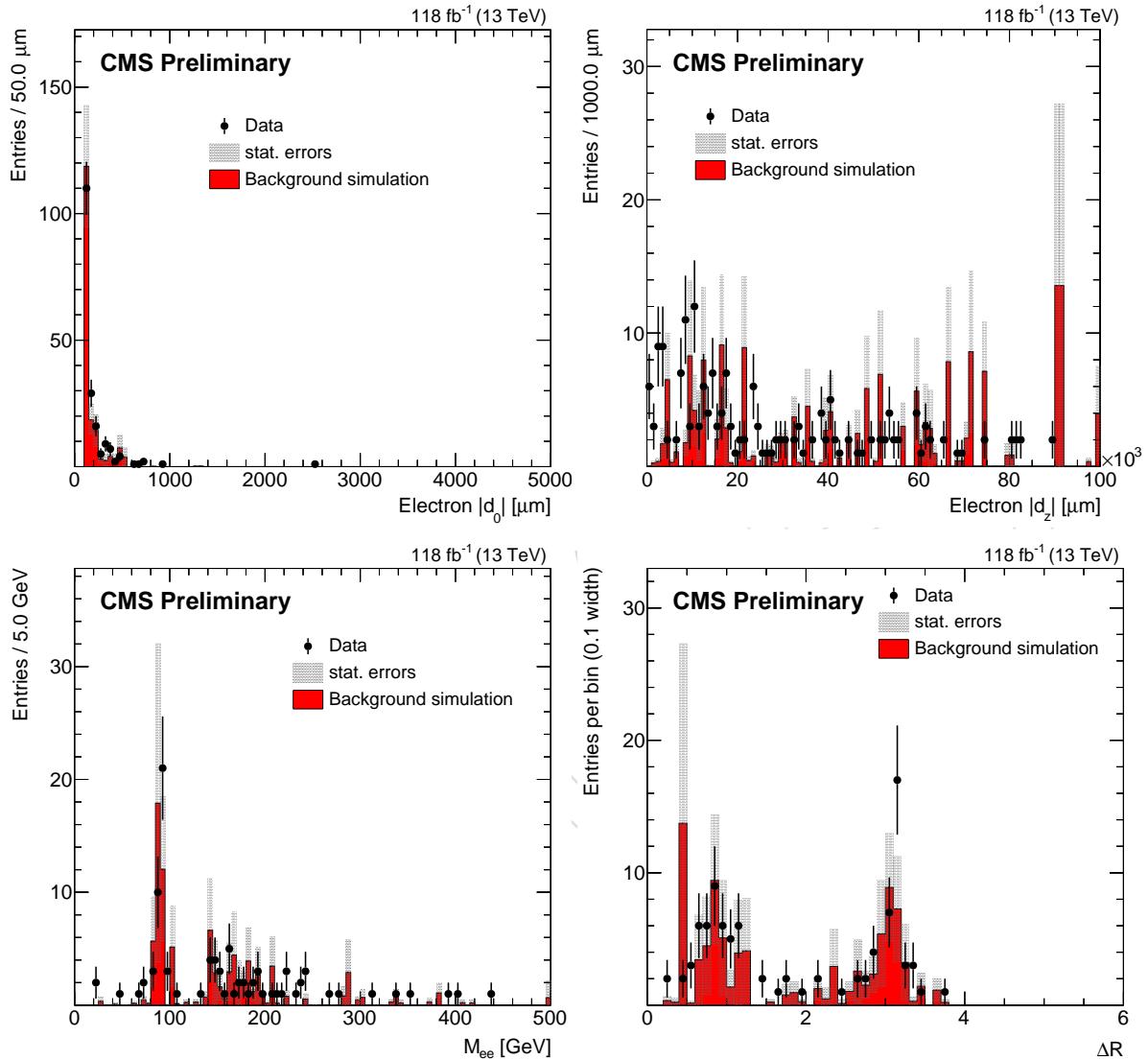


Figure 83: Electron  $|d_0|$ ,  $|z_0|$ , invariant mass, and  $\Delta R$  distributions for data and composite background simulation in the ee channel in the inclusive SR. The background simulation is normalized to the ABCD method background prediction.

<sup>1349</sup> Figures 84, 85, and 86 show data and background simulation in the ee channel in the inclusive  
<sup>1350</sup> SR.

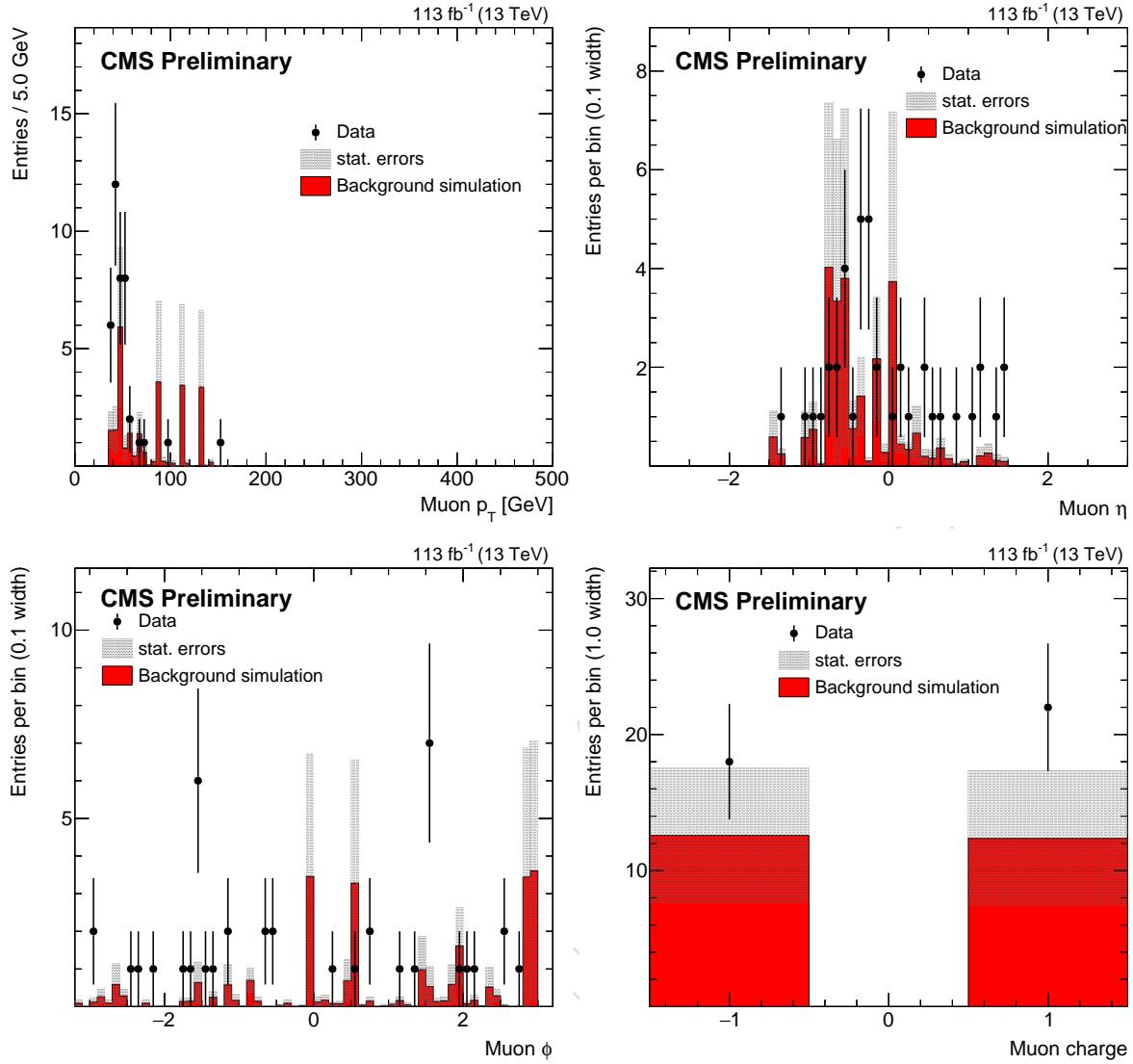


Figure 84: Muon  $p_T$ ,  $\eta$ ,  $\phi$ , and charge distributions for data and composite background simulation in the  $\mu\mu$  channel in the inclusive SR. The background simulation is normalized to the ABCD method background prediction.

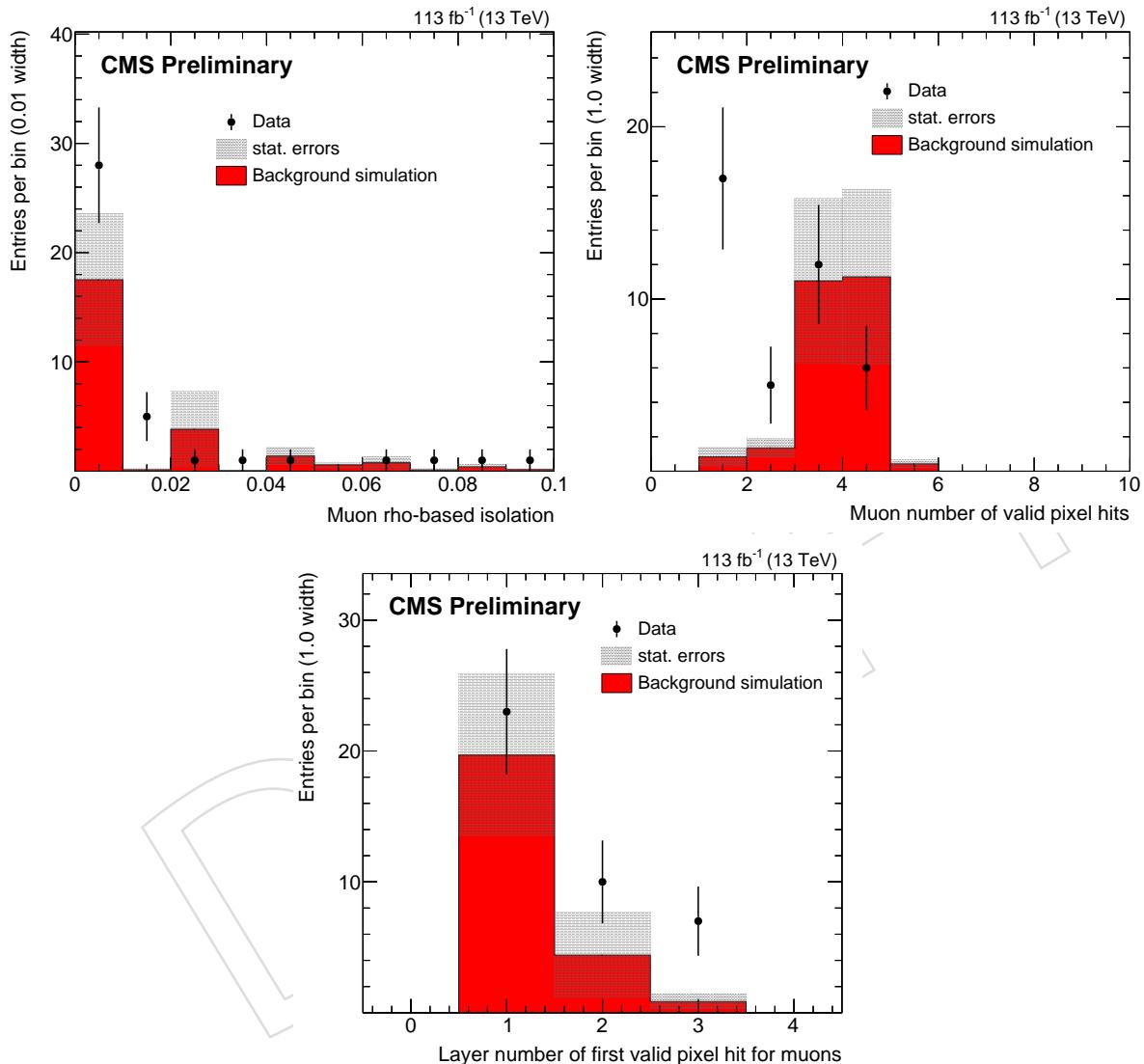


Figure 85: Muon isolation and pixel hit distributions for data and composite background simulation in the  $\mu\mu$  channel in the inclusive SR. The background simulation is normalized to the ABCD method background prediction.

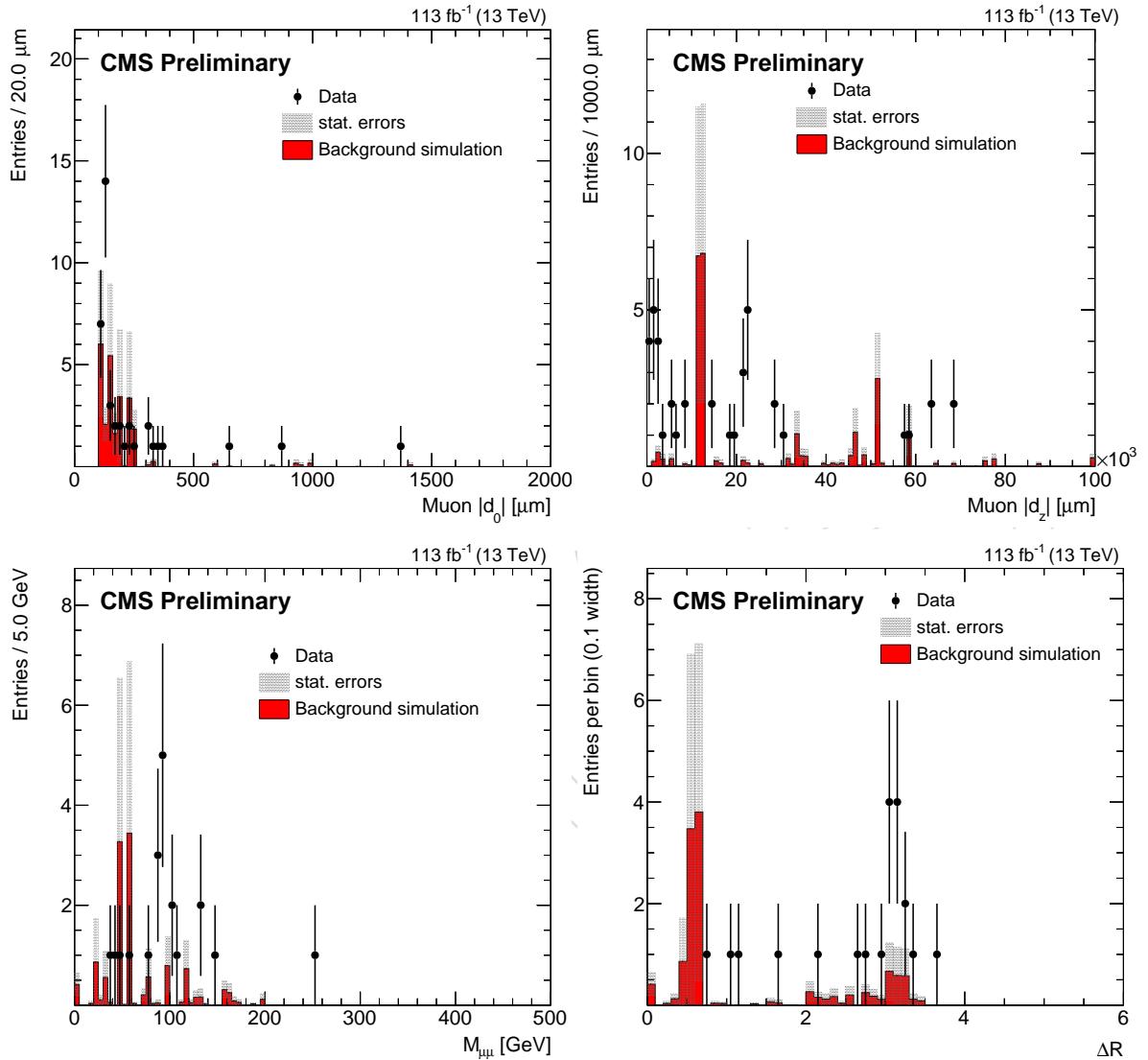


Figure 86: Muon  $|d_0|$ ,  $|z_0|$ , invariant mass, and  $\Delta R$  distributions for data and composite background simulation in the  $\mu\mu$  channel in the inclusive SR. The background simulation is normalized to the ABCD method background prediction.

1351 **E Likelihood tests**

1352 After unblinding, we performed several tests to understand how the likelihood handles the  
 1353 observed yields. Figure 87 shows the distribution of pulls for each signal region bin, Figure 88  
 1354 shows the pulls for each background estimate nuisance parameter, Table 23 shows the post-fit  
 1355 background yields in each signal region, and Figure 89 shows the observed significances for  
 1356 the  $\tilde{t}\tilde{t} \rightarrow \bar{l}b\bar{l}b$  process.

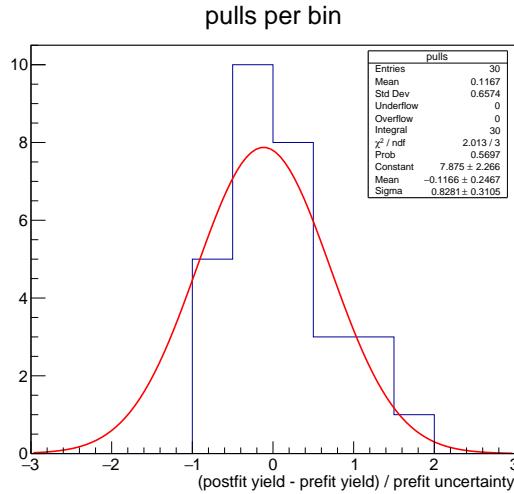


Figure 87: The distribution of pulls for each signal region bin. Pulls are calculated as the difference between the post-fit background yield and the pre-fit background yield divided by the pre-fit background uncertainty. The post-fit values are obtained from the signal-plus-background fit with no regions masked and the pre-fit values are obtained from the background-only fit with all signal regions masked.

Post-fit nuisance pull distribution

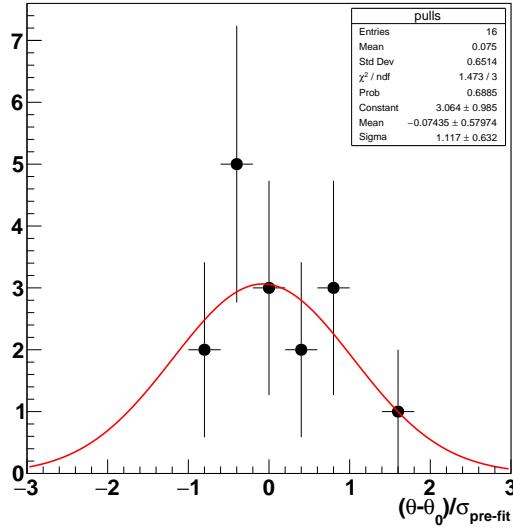


Figure 88: The distribution of pulls for each background nuisance parameter. Pulls are calculated as the difference between the post-fit value and the pre-fit value divided by the pre-fit uncertainty.

Table 23: The post-fit predictions and pulls for each signal region bin.

	SR I, low $p_T$ bin	SR I, high $p_T$ bin	SR II	SR III	SR IV
<i>2016 eμ</i>					
- post-fit prediction	$7.1 \pm 2.0$	$0.76 \pm 0.31$	$0.08 \pm 0.08$	$0.14 \pm 0.14$	$0.003 \pm 0.003$
- pull	0.86	0.79	-0.11	-0.11	0.00
<i>2017+2018 eμ</i>					
- post-fit prediction	$31 \pm 5$	$0.68 \pm 0.25$	$0.20 \pm 0.17$	$0.65 \pm 0.48$	$0.01 \pm 0.01$
- pull	-0.51	-0.18	-0.08	-0.07	-0.05
<i>2016 ee</i>					
- post-fit prediction	$35 \pm 5$	$0.40 \pm 0.14$	$0.50 \pm 0.75$	$0.44 \pm 0.53$	$0.01 \pm 0.02$
- pull	1.58	1.03	0.00	0.01	0.00
<i>2017+2018 ee</i>					
- post-fit prediction	$50 \pm 6$	$0.65 \pm 0.19$	$2.5 \pm 0.7$	$3.2 \pm 0.9$	$0.22 \pm 0.06$
- pull	-0.72	-0.67	-0.38	-0.32	-0.34
<i>2016 μμ</i>					
- post-fit prediction	$11 \pm 2$	$0.37 \pm 0.10$	$0.19 \pm 0.10$	$0.21 \pm 0.12$	$0.01 \pm 0.01$
- pull	1.12	1.06	0.20	0.21	0.33
<i>2017+2018 μμ</i>					
- post-fit prediction	$2.5 \pm 1.1$	$0.51 \pm 0.22$	$0.14 \pm 0.36$	$0.23 \pm 0.63$	$0.02 \pm 0.05$
- pull	-0.58	-0.57	0.48	0.49	0.53

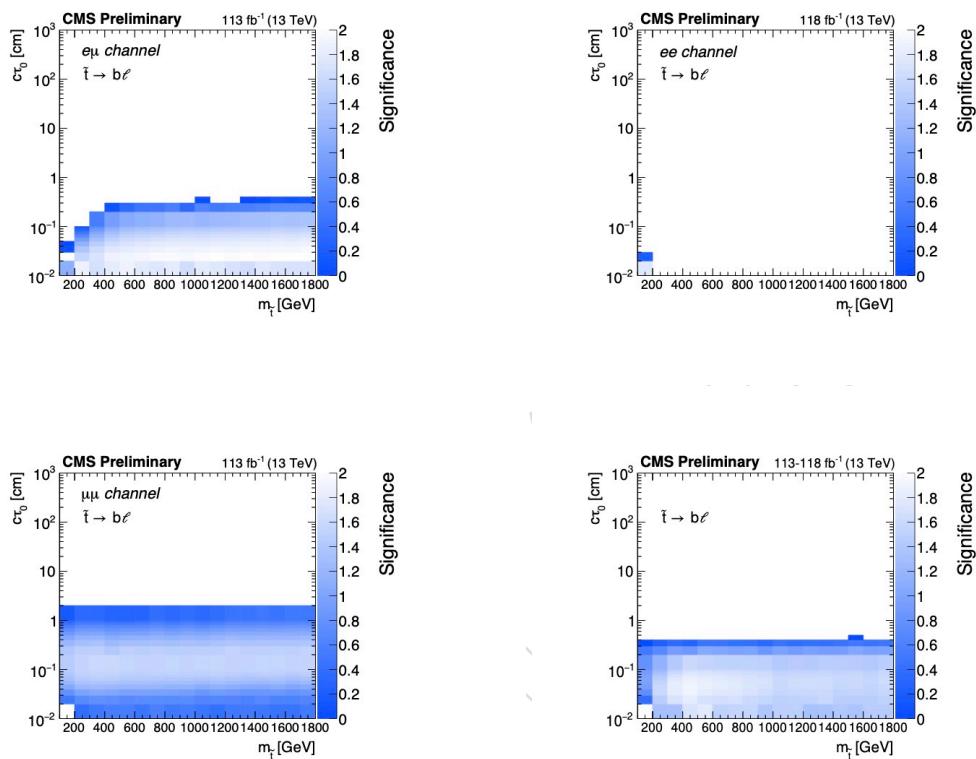


Figure 89: The observed significances for the  $\tilde{t}\tilde{t} \rightarrow \bar{l}b\bar{l}b$  process as a function of  $\tilde{t}$  mass and lifetime, for the  $e\mu$  channel (upper left), the  $ee$  channel (upper right), the  $\mu\mu$  channel (lower left), and the three channels combined (lower right).

## 1357 F Study of R-hadron treatment in simulation

1358 In the  $pp \rightarrow \tilde{t}\tilde{t}$  signal processes, the top squarks can form strongly produced hadronic states  
 1359 called “R-hadrons.” In the  $\tilde{t} \rightarrow bl$  and  $\tilde{t} \rightarrow dl$  signal samples used in this analysis, R-hadron for-  
 1360 mation and decay are modeled with PYTHIA while the R-hadron interactions with material are  
 1361 not modeled. Given that the the lepton identification criteria effectively limit the number of  
 1362 interaction lengths an R-hadron can traverse before decaying, we expect the impact of such  
 1363 interactions to be negligible. To test this hypothesis, we generate alternative signal samples  
 1364 in which the R-hadron material interactions and decay are modeled in GEANT4 and compare  
 1365 the resulting signal efficiency to that of the nominal samples. We find differences in signal effi-  
 1366 ciency on the order of 20% that are completely attributable to differences in the treatment of jet  
 1367 formation between the PYTHIA and GEANT4 implementations and therefore conclude that the  
 1368 R-hadron material interactions do not meaningfully impact the analysis.

1369 In the alternative samples, the R-hadron interactions are modeled following the “cloud model”,  
 1370 which assumes that the top squark is surrounded by a cloud of colored, light constituents  
 1371 that interact during scattering [38, 39]. Such interactions can potentially alter the rate of R-  
 1372 hadron energy loss or cause the R-hadron charge value to change mid-flight. In the context  
 1373 of CMSSW, simultaneously enabling R-hadron interactions with material and R-hadron decay  
 1374 within the detector volume constrains us to perform both the R-hadron propagation and decay  
 1375 with GEANT4. We generate two additional sets of signal samples for this study. In the first, the  
 1376 R-hadron decays to a lepton and bottom quark. Because GEANT4 does not perform the parton  
 1377 shower and hadronization that lead to jet formation, the bottom quark is in this case unrealis-  
 1378 tically treated as a final-state particle. To help study the effect of this missing jet formation, we  
 1379 generate a second sample in which the R-hadron decays to a lepton and a neutral pion, which  
 1380 then decays to two photons and initiates an electromagnetic shower.

1381 Figure 90 shows how modeling the R-hadron interactions and decay in GEANT4 affects the  
 1382 inclusive signal region cutflows in each channel for a top squark mass of 1000 GeV and proper  
 1383 decay length of 1 cm. The alternative samples generally pass the full selection at a slightly  
 1384 higher rate than the existing signal samples. The overall magnitude of the effect is similar  
 1385 across channels and signal points, and a few clear trends are apparent in the cutflows. First,  
 1386 events from the neutral pion sample pass the ee and  $e\mu$  triggers at a higher rate than other  
 1387 samples because of the increased number of high-momentum photons and electrons from the  
 1388 neutral pion decay, an effect that carries through the full selection. Second,  $\tilde{t} \rightarrow bl$  events pass  
 1389 the  $\mu\mu$  and  $e\mu$  triggers at a higher rate due to the increased number of muons from b jets, but  
 1390 this effect is neutralized by the muon isolation criteria. Third, events in which the R-hadron  
 1391 decays to a lepton and a final-state b quark in GEANT4 pass the  $e\mu$  and ee triggers at lower  
 1392 rates due to the reduced number of electrons from jets but pass the lepton quality and isolation  
 1393 criteria at higher rates because the absence of jets increases the probability that leptons from  
 1394 the top squark decay are well isolated. A similar effect can be seen in the neutral pion sample  
 1395 in the  $\mu\mu$  cutflow.

1396 The stated causes of the effects described above are supported by the kinematic distributions of  
 1397 electrons, muons, and jets in each of these samples before and after the inclusive signal region  
 1398 selections are applied. For example, Fig. 91 shows that before any selection is applied, events  
 1399 in which the R-hadron decays to a lepton and a neutral pion contain a clear relative excess of  
 1400 high- $p_T$ , isolated electrons;  $\tilde{t} \rightarrow bl$  events contain a clear relative excess of high- $p_T$ , non-isolated  
 1401 muons; and events in which the R-hadron decays to a lepton and a non-physical final-state b  
 1402 quark (and to a lesser degree, events in which the R-hadron decays to a lepton and a neutral  
 1403 pion) show a distinct lack of central, high- $p_T$  jets.

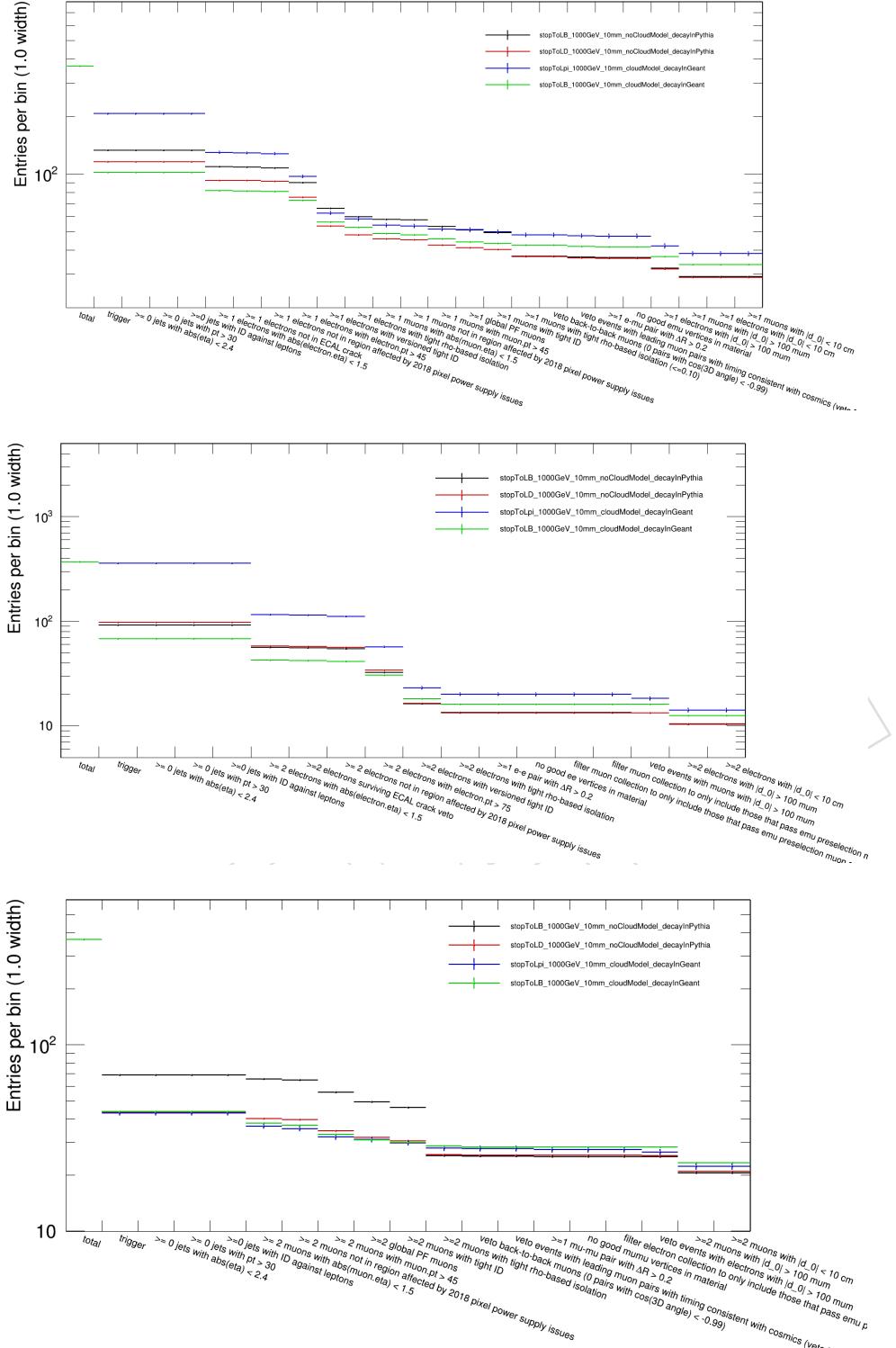


Figure 90: Inclusive signal region cutflows for four signal samples in the  $e\mu$  (top),  $ee$  (middle), and  $\mu\mu$  (bottom) channels. The top squark mass and proper decay length are 1000 GeV and 1 cm. In the sample corresponding to the black (red) curves, R-hadron material interactions are not modeled, but the top squark decay is performed in PYTHIA and the resulting b (d) quark produces a jet. In the samples corresponding to the blue and green curves, the R-hadron material interactions and decay are modeled with GEANT4. In the sample corresponding to the blue curves, the R-hadron decays to a lepton and a neutral pion, and in the sample corresponding to the green curves, the R-hadron decays to a lepton and a non-physical final-state quark.

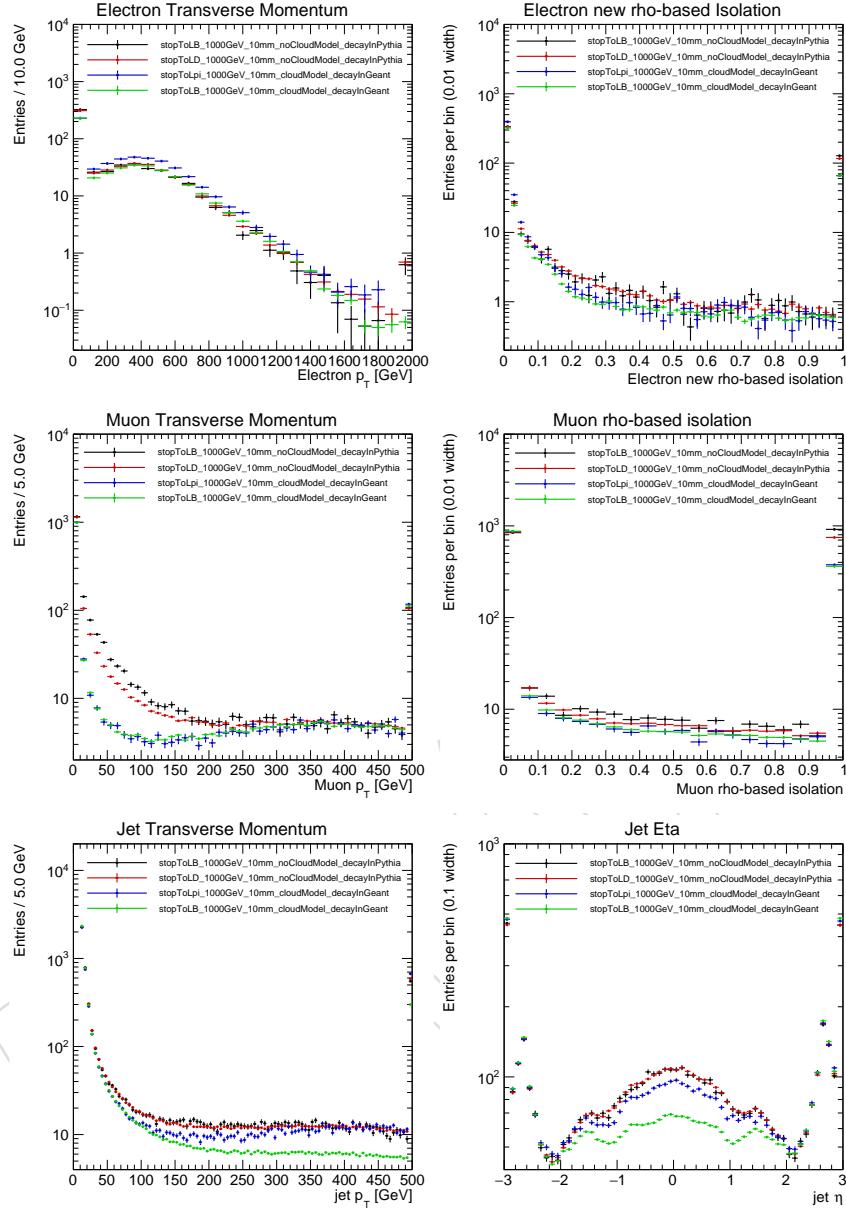


Figure 91: Kinematic distributions of electrons (top), muons (middle), and jets (bottom) before any selection is applied for four signal samples in which the top squark mass and proper decay length are 1000 GeV and 1 cm. In the sample corresponding to the black (red) curves, R-hadron material interactions are not modeled, but the top squark decay is performed in PYTHIA and the resulting b (d) quark produces a jet. In the samples corresponding to the blue and green curves, the R-hadron material interactions and decay are modeled with GEANT4. In the sample corresponding to the blue curves, the R-hadron decays to a lepton and a neutral pion, and in the sample corresponding to the green curves, the R-hadron decays to a lepton and a non-physical final-state quark. In each plot, the rightmost bin contains overflow entries.

1404 Figure 92 shows how the effects described above carry through to the events that pass the  
1405 inclusive signal region selections. As expected, the events in which the R-hadron decays to a  
1406 lepton and a neutral pion have a relative excess of high- $p_T$ , well-isolated electrons when the  
1407 full ee signal region selection is applied. The muon isolation cut eliminates the differences  
1408 between the two samples in which the top squark is decayed in PYTHIA when the  $\mu\mu$  inclusive  
1409 signal region selection is applied, but the relative excess of well-isolated muons from the two  
1410 other samples becomes more pronounced. Finally, the stark differences in jet activity remain  
1411 apparent.

1412 In the end, modeling the R-hadron interactions and decay with GEANT4 has an approximately  
1413 20% effect on the signal efficiency that does not vary meaningfully with analysis channel, top  
1414 squark mass, or top squark lifetime. We find that all observable differences are due to known,  
1415 non-physical effects of the modeling of the R-hadron decay and therefore conclude that our  
1416 nominal signal samples provide the best estimate of the signal efficiency and any real effects  
1417 from R-hadron interactions must be comparatively small and therefore insignificant to the anal-  
1418 ysis as a whole.

DRAFT

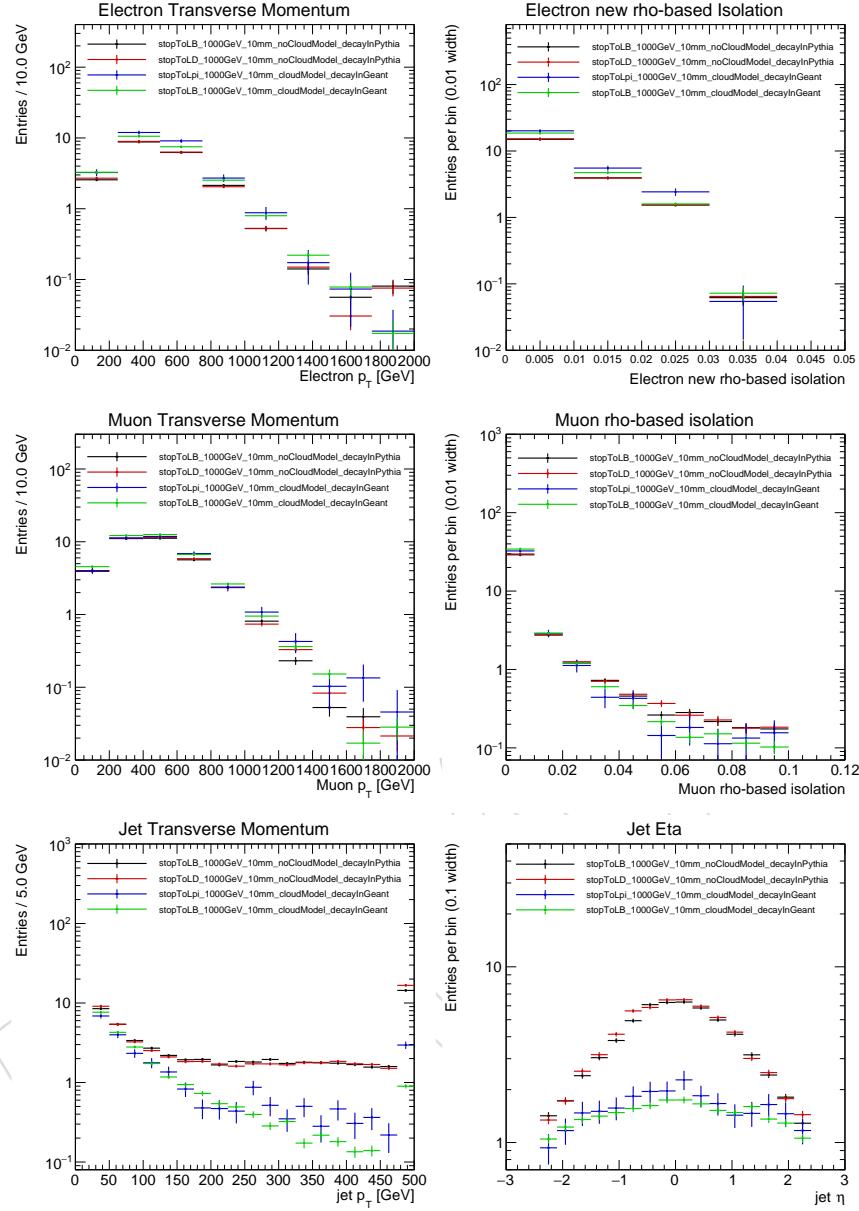


Figure 92: Kinematic distributions of electrons in the ee signal region (top), muons in the  $\mu\mu$  signal region (middle), and jets in the  $\mu\mu$  signal region (bottom) for four signal samples in which the top squark mass and proper decay length are 1000 GeV and 1 cm. In the sample corresponding to the black (red) curves, R-hadron material interactions are not modeled, but the top squark decay is performed in PYTHIA and the resulting b (d) quark produces a jet. In the samples corresponding to the blue and green curves, the R-hadron material interactions and decay are modeled with GEANT4. In the sample corresponding to the blue curves, the R-hadron decays to a lepton and a neutral pion, and in the sample corresponding to the green curves, the R-hadron decays to a lepton and a non-physical final-state quark. In each plot, the rightmost bin contains overflow entries.