

# CMS Draft Analysis Note

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

2014/09/23

Head Id: 254219

Archive Id: 261156M

Archive Date: 2014/08/02

Archive Tag: trunk

## Search for long-lived particles decaying to final states that include dimuons reconstructed using only the muon chambers

Z. Hu, N. Leonardo, M. de Mattia, I. Shipsey, M. Solmaz  
Purdue University

### Abstract

A search for long-lived neutral particles decaying to pairs of muons by using only the muon chambers is presented. Events were collected by the CMS detector during pp collisions at  $\sqrt{s} = 8$  TeV, and selected from data samples corresponding to  $20.5 \text{ fb}^{-1}$  of integrated luminosity. No excess of events is observed above the Standard Model expectations and upper limits are derived for two specific models. The first model predicts a heavy scalar decaying to two long-lived particles, each of which decays to muon pairs. The second model predicts stop pair production. Each of the stops decays to a long-lived neutralino which subsequently decays to two muons and a neutrino. Combined upper limits with an analysis utilizing the CMS silicon tracker are also presented.

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

PDFAuthor: Z. Hu, N. Leonardo, M. de Mattia, I. Shipsey, M. Solmaz  
PDFTitle: Search for long-lived particles decaying to final states that include dimuons  
reconstructed using only the muon chambers  
PDFSubject: CMS  
PDFKeywords: CMS, physics

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



---

1 NOTICES & MILESTONES

2     ● August, 2014 (V2)

- 3       - Charge misassignment in signal MC.  
4       - Appendix explaining low trigger efficiency due to transverse momentum  
5       bias as a function of transverse decay length.  
6       - Additional background symmetry plot is added to "Background" section.  
7       - Updated table 5 (cutflow). The previous version was using an outdated  
8       selection.  
9       - Switched to higher statistics signal MC samples.  
10      - Upper limits for the squark-neutralino model and corresponding com-  
11      bined limits.  
12      - Added variation of the normalization for the background systematic.

13     ● July, 2014 (v1)

14     This analysis note – AN-2014/176 – serves as supporting documentation for EXO-  
15     **14-012**.

DRAFT

**16    Contents**

17	1	Introduction . . . . .	3
18	2	The CMS detector . . . . .	4
19	3	Data and Monte Carlo simulation samples . . . . .	6
20	4	Event reconstruction and selection . . . . .	8
21	4.1	Displaced track reconstruction . . . . .	8
22	4.2	Muon selection . . . . .	10
23	4.3	Selection of long-lived exotica . . . . .	11
24	4.4	Results of blind analysis . . . . .	13
25	4.5	Selection efficiency and acceptance . . . . .	15
26	5	Estimated background and associated systematic uncertainties . . . . .	21
27	5.1	Background validation . . . . .	21
28	5.2	Background systematic uncertainties . . . . .	24
29	6	Signal systematic uncertainties . . . . .	26
30	6.1	Luminosity . . . . .	27
31	6.2	Effect of pileup . . . . .	27
32	6.3	Track finding and selection efficiency . . . . .	27
33	6.4	Trigger efficiency measurement . . . . .	36
34	6.5	Effect of the $p_T$ resolution . . . . .	40
35	7	Analysis sensitivity and expected upper limits . . . . .	41
36	7.1	Comparison with the tracker-based analysis . . . . .	42
37	7.2	Combined Limits . . . . .	43
38	8	Summary . . . . .	47
39	A	Muon timing information in the events removed by cosmic rejection . . . . .	47
40	B	Muon trigger efficiency . . . . .	49
41	C	Agreement between data and the background simulation . . . . .	52
42	D	Effect of the RSA muon $p_T$ resolution on the analysis . . . . .	60
43	E	$L_{xy}$ resolution of RSA muons for prompt decays . . . . .	66
44	F	Secondary vertex reconstruction efficiency . . . . .	69
45	F.1	Secondary vertex reconstruction efficiency in signal MC . . . . .	69
46	F.2	Secondary vertex reconstruction efficiency in cosmics data and simulation . . . . .	70
47	G	Alternative background systematics estimate . . . . .	71
48	H	Bias in $p_T$ and low trigger efficiency . . . . .	73

---

## 49 1 Introduction

- 50 Massive long-lived particles conjectured by several new physics models, such as “split SUSY”  
 51 [1] or SUSY with very weak R-parity violation [2], “hidden valley” models [3] and Z' mod-  
 52 els that contain long-lived neutrinos [4], might be produced at the LHC. In the models where  
 53 the long-lived massive particles decay to lepton pairs, they can be differentiated from Stan-  
 54 dard Model (SM) particles by virtue of the significant distance they travel in the volume of the  
 55 detector.
- 56 As a benchmark for this physics signature, two models are considered to quantify the sensitiv-  
 57 ity of the analysis. The first model postulates pair production of long-lived X particles by the  
 58 decay of a non-SM Higgs boson,  $H^0 \rightarrow XX$ , where  $H^0$  is produced by gluon-gluon fusion and  
 59 X is a spinless boson decaying to lepton pairs,  $X \rightarrow \ell^+ \ell^-$  [5]. In the second model, a squark  
 60 decays as  $\tilde{q} \rightarrow q\tilde{\chi}^0$ ,  $\tilde{\chi}^0 \rightarrow \ell^+ \ell^- \nu$ , where the long-lived neutralino,  $\chi^0$ , produces a lepton pair  
 61 and a neutrino as a result of R-parity violation. Even though the results are given for these spe-  
 62 cific models, they are presented in a way that is approximately model independent, and can be  
 63 used to derive concrete limits on a wide array of models producing two-body and three-body  
 64 decays of long lived particles to muons.
- 65 This study presents the results of a search for long-lived neutral particles decaying to muon  
 66 pairs reconstructed using only the muon chambers of the Compact Muon Solenoid (CMS) de-  
 67 tector. The analysis uses data taken during 2012 in pp collision at  $\sqrt{s} = 8$  TeV, corresponding to  
 68 a total integrated luminosity of  $20.5 \text{ fb}^{-1}$ . It shares some similarities with another CMS analy-  
 69 sis [6], searching for particles of the same nature yielding displaced electron and muon signals  
 70 by utilizing both the silicon tracker and the muon system for particle reconstruction. The two  
 71 analyses utilize similar techniques for event selection, background estimation, evaluation of  
 72 systematic uncertainties, and upper limit calculations. Nevertheless, they are orthogonal by  
 73 construction, as explained later in this document.
- 74 Although the tracker-based analysis benefits from the precision with which tracks are mea-  
 75 sured in the silicon tracker, the major constraint comes from the fact that the reconstruction  
 76 efficiency for a track in the silicon tracker is essentially zero for tracks with transverse impact  
 77 parameter ( $d_0$ ), the closest distance between the track and the interaction point in the trans-  
 78 verse plane of the detector, greater than 40 cm. The tracker-based analysis has little sensitivity  
 79 to particles with longer lifetimes. On the other hand, the muon chambers give non-vanishing  
 80 reconstruction efficiency even a few meters away from the interaction point. To illustrate this,  
 81 the reconstruction efficiencies of the tracker and the muon chambers as a function of  $d_0$  are  
 82 given in Figure 1. Additional selection requirements are applied to derive the reconstruction  
 83 efficiency of the muon system, including quality selection efficiency. More details can be found  
 84 in Section 6.3. Therefore, the muon chambers can be used to extend the lifetime sensitivity of  
 85 this analysis. Note, however, that the muon chambers have a much lower muon  $p_T$  resolution  
 86 and higher level of cosmic muon background compared to the tracker-based analysis. Impor-  
 87 tantly, the double muon trigger that is used to collect the events has a vanishing efficiency  
 88 beyond 2.5 meters of the collision point in the transverse plane. Hence, the effective range of  
 89 the analysis is 2.5 meters, which is only halfway through the muon chambers.
- 90 The analysis is fully complementary to the tracker-based analysis in that the displaced muons  
 91 reconstructed by the muon chambers that are matched to the tracker muons are rejected. That  
 92 is, the set of events passing the full selection of the analysis does not overlap with the one  
 93 satisfying the selection criteria of the tracker-based analysis.
- 94 The D0 Collaboration has performed similar searches for leptons from displaced decays within

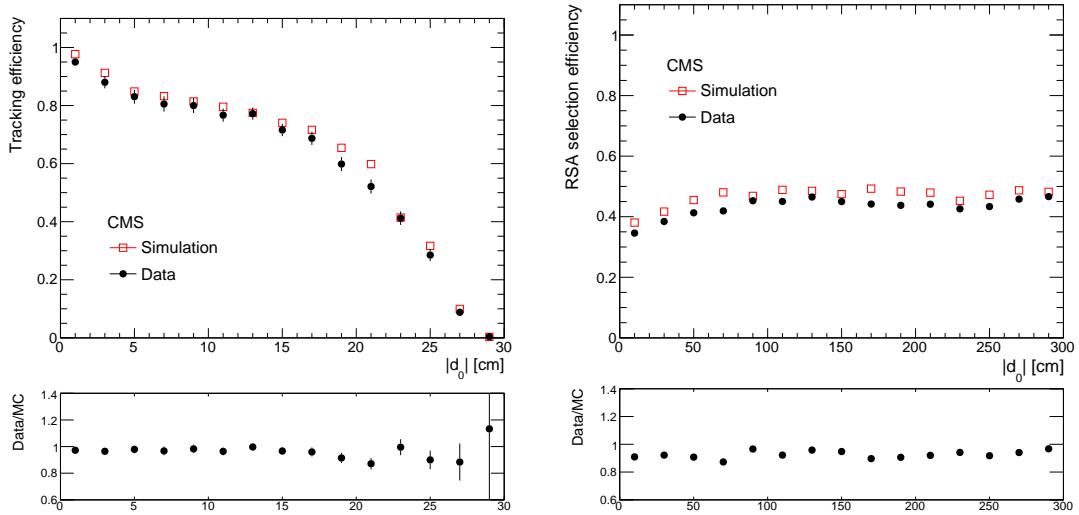


Figure 1: (Upper left) Efficiency of the tracker to find a track given a cosmic ray muon as a function of the transverse impact parameter of the muon. Only the tracker muons with  $|z_0| < 10$  cm are used. (Upper right) Efficiency of the muon chambers for muons with  $|z_0| < 50$  cm. The lower left and lower right plots show the ratio of the efficiency in data to the simulation for the tracker and the muon chambers, respectively.

95 its tracker volume [7, 8], yet the scope of these searches covers a much smaller kinematic phase  
 96 space region than CMS. The ATLAS Collaboration has performed searches that are sensitive  
 97 to decay lengths up to about 20 m by exploiting the ATLAS muon spectrometer [9, 10], using  
 98 different decay channels from those considered in this analysis.

## 99 2 The CMS detector

100 The central feature of the CMS apparatus [11] is a superconducting solenoid of 6 m internal  
 101 diameter providing an axial field of 3.8 T. Within the field volume are the silicon pixel, strip  
 102 tracker, the lead-tungstate crystal electromagnetic calorimeter (ECAL) and the brass/scintillator  
 103 hadron calorimeter (HCAL). Muons are identified in gas-ionisation detectors embedded in the  
 104 steel magnetic-flux return yoke of the solenoid. The transverse view of the detector is illus-  
 105 trated in Figure 2.

106 The silicon tracker can reconstruct the trajectories of charged particles such as muons, electrons  
 107 and hadrons as well as their momentum with high precision. It is composed of pixel detectors  
 108 (three barrel layers and two forward disks on either end of the detector) surrounded by strip  
 109 detectors (ten barrel layers plus three inner disks and nine forward disks at each end of the  
 110 detector). The tracker covers the pseudorapidity range  $|\eta| < 2.5$ , where  $\eta = -\ln[\tan(\theta/2)]$   
 111 and  $\theta$  is the polar angle with respect to the anticlockwise-beam direction.

112 In order to stop electrons and photons, the electromagnetic calorimeter is placed on the periph-  
 113 ery of the silicon tracker. The ECAL consists of nearly 76 000 lead tungstate crystals in a barrel  
 114 and two endcap sections, which provide coverage in pseudorapidity  $|\eta| < 3$ . The hadron cal-  
 115 orimeter is between the muon chambers and the electromagnetic calorimeter. It measures the  
 116 energy of hadrons and it is made up of barrel, endcap and forward sections.

117 Muons are measured in the pseudorapidity range  $|\eta| < 2.4$  with detection planes based on one  
 118 of three technologies: drift tubes in the barrel region, cathode strip chambers in the endcaps,

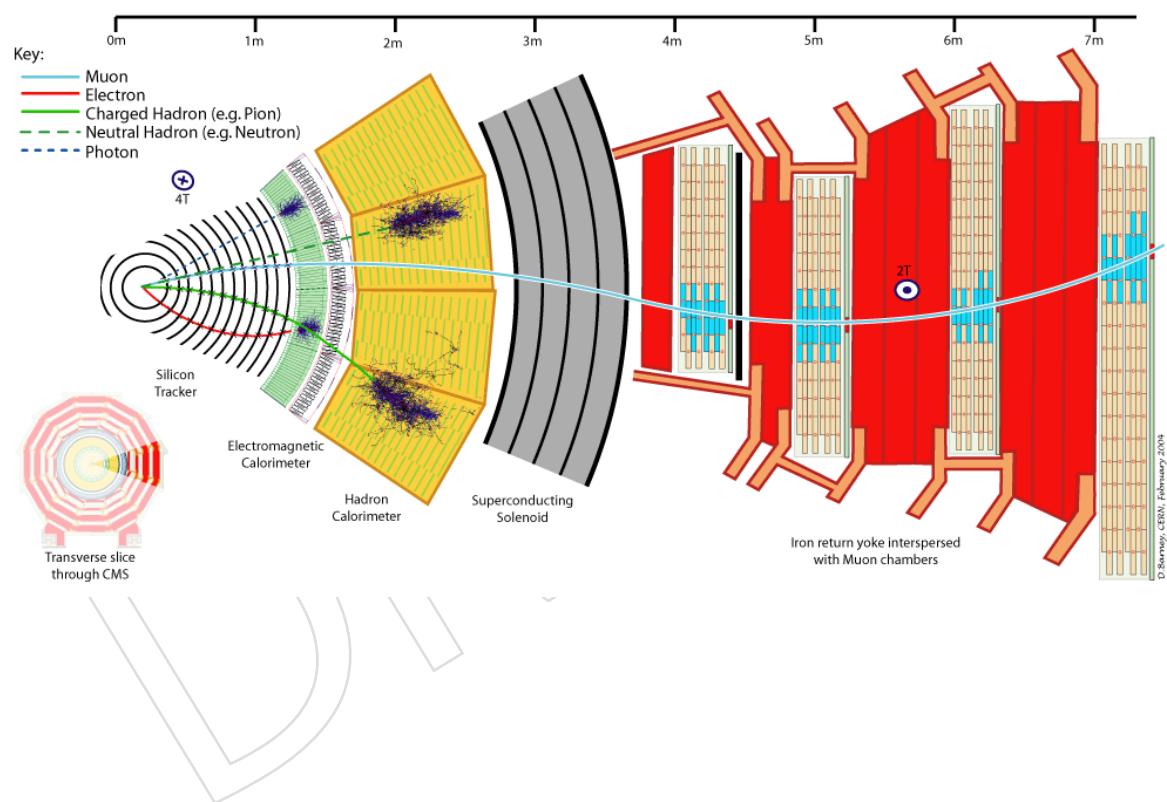


Figure 2: The CMS detector in the plane transverse to the beam.

and resistive plate chambers in the barrel and endcaps. The muon system, shown in Figure 3 has three main functions: triggering on muons, muon identification, and the improvement of muon momentum measurement. The drift tube subsystem, which covers  $|\eta| < 1.2$  region, is responsible for determining the muon position through the process of ionization in the gas tubes. Cathode strip chambers consist of anode wires and cathode strips positioned perpendicular to each other. They provide two position coordinates in the region,  $0.9 < |\eta| < 2.4$ . The resistive plate chambers located in the pseudorapidity range of  $|\eta| < 1.6$  provide additional fast muon trigger capability. Track reconstruction can be achieved in the tracker and the muon system independently and it can be improved by combining the two. Muon reconstruction performance has been studied in great detail with data [12].

The first level of the CMS trigger system, composed of custom hardware processors, selects events of interest using information from the calorimeters and the muon detectors. A high-level trigger processor farm then employs the full event information to further decrease the event rate.

### 3 Data and Monte Carlo simulation samples

The analysis uses data taken from pp collisions at a centre-of-mass energy of 8 TeV corresponding to an integrated luminosity of  $20.5 \pm 0.5 \text{ fb}^{-1}$ . The CMS datasets utilized are reprocessed under CMSSW 5\_3\_7 in January 2013. They are known as the “rereco” data. Table 1 lists them along with the associated good run range.

Table 1: The CMS datasets used by the analysis.

Dataset	Run range
Run2012A-22Jan2013-v1/AOD	190456–193621
Run2012B-22Jan2013-v1/AOD	193833–196531
Run2012C-22Jan2013-v1/AOD	198022–203742
Run2012D-22Jan2013-v1/AOD	203777–208686

HLT\_L2DoubleMu23\_NoVertex\_2Cha\_Angle2p5 is the Level 2 (L2) double muon trigger that collects the events used in this analysis. It requires two muons in an event, each reconstructed in the muon detectors without imposing any beam spot constraint and having  $p_T > 23 \text{ GeV}/c$ . Both muons are also required to have at least two reconstructed hits in at least two cathode strip chambers (CSC) or drift tubes (DT). To prevent cosmic ray muons from passing these criteria, the opening angle between the two muons must be less than 2.5 radians. The trigger is independent of the silicon tracker activity.

For the  $H^0 \rightarrow XX$  model, simulated signal samples are generated using PYTHIA V6.426 [13] to simulate  $H^0$  production through gluon fusion ( $gg \rightarrow H^0$ ). Subsequently the  $H^0$  is forced to decay to  $XX$ , with the  $X$  bosons each decaying to lepton pairs ( $X \rightarrow \ell^+ \ell^-$ ). The analysis focuses on the final states with at least one muon pair. The generated samples have  $M_{H^0} = 125, 200, 400, 1000 \text{ GeV}/c^2$  and  $M_X = 20, 50, 150, 350 \text{ GeV}/c^2$ . Each sample is produced with three  $X$  boson lifetimes. After the boost, mean transverse decay lengths are of approximately 20 cm, 200 cm and 2000 cm with respect to the laboratory frame. The sensitivity of the analysis is determined only for the decays with the longest lifetime in each sample. Figure 4 displays a simulated event with  $M_{H^0} = 1000 \text{ GeV}/c^2$  and  $M_X = 350 \text{ GeV}/c^2$ .

For the  $\tilde{\chi}^0 \rightarrow \ell^+ \ell^- \nu$  model, squark pair production is simulated by PYTHIA as well as the decay to  $\tilde{\chi}^0$ . The following four pairs of squark and neutralino masses are used :  $(M_{\tilde{q}}, M_{\tilde{\chi}^0}) =$

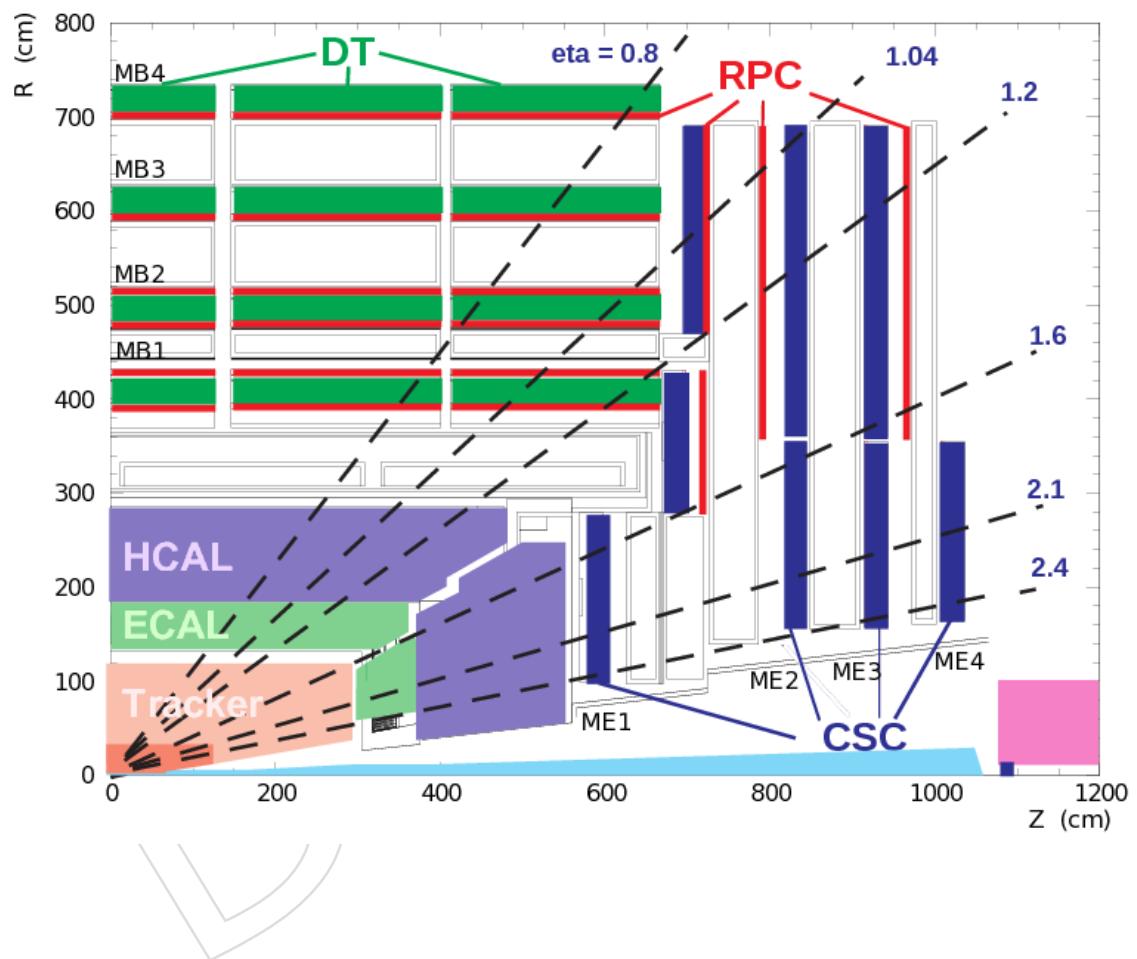


Figure 3: One quadrant of the CMS detector in the longitudinal plane. The drift tube (DT) stations are represented by dark green rectangles. The blue and red rectangles denote the four cathode strip chamber (CSC) stations and resistive plate chamber (RPC) stations, respectively.

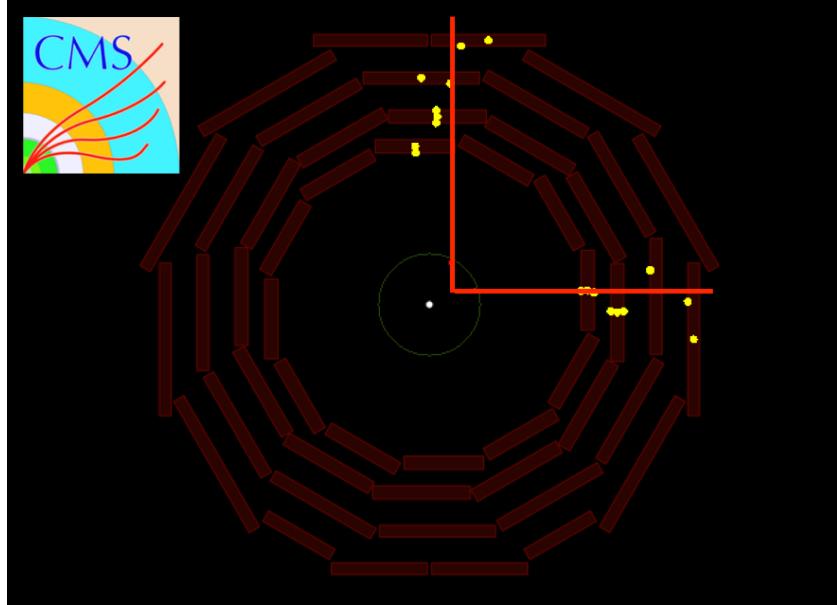


Figure 4: Transverse view of a simulated event with  $M_{H^0} = 1000 \text{ GeV}/c^2$  and  $M_X = 350 \text{ GeV}/c^2$ . In this event, one X boson decays to a pair of muons, identified by the hits in the muon system. The other X boson decays to an electron pair which is not shown in the figure.

(1500, 494), (1000, 148), (350, 148) and (120, 48)  $\text{GeV}/c^2$ . The R-parity violating parameter  $\lambda'_{211}$  is set to a non-zero value so that the decay of the  $\tilde{\chi}^0$  into two charged leptons and a neutrino is permitted. The value of  $\lambda'_{211}$  is chosen to give a mean transverse decay length of approximately 200 cm. Table 2 summarizes the properties of the signal Monte Carlo (MC) samples used in the analysis.

All MC background samples, reconstructed under CMSSW 5\_3\_2, used in the analysis are listed in Table 3. They are generated with PYTHIA and correspond to ‘Summer12 DR53X’ production. The major background for this analysis comes from the Drell-Yan process yielding dileptons,  $\mu^+ \mu^-$  and  $\tau^+ \tau^-$  at significant rates. Even though the branching ratio is low, tau decays might also lead to displaced muons that can fake the signals we are looking for. The Drell-Yan background is simulated at Next-Leading-Order (NLO) with POWHEG [14]. Other simulated backgrounds are  $t\bar{t}$ , W/Z boson pair production with leptonic decays, and QCD multijet events. All these backgrounds produce negligible contributions. However, the random cosmic background is not simulated and the background MC cannot provide a good description of the expected background. The expected background is estimated from the data. In all the samples, the response of the detector is simulated in detail using GEANT4 [15]. The samples are then processed through the trigger emulation and event reconstruction chain of the CMS experiment.

Table 2: The list of simulated signal samples used in the analysis. In the upper part of the table the  $H^0$  and  $X$  mass values are presented along with three different lifetimes for  $X$ . After the boost, the mean transverse decay lengths are of approximately 20 cm, 200 cm and 2000 cm with respect to the laboratory frame. For the RPV SUSY samples, the masses of the  $\tilde{q}$  and  $\tilde{\chi}_0^1$  are shown in the lower part of the table. The mean proper decay lengths of these samples correspond to a mean transverse decay length in the laboratory frame of approximately 200 cm.

$M_{H^0}$ (GeV/c $^2$ )	$M_X$ (GeV/c $^2$ )	$c\tau$ (cm)
1000	350	(35.0, 350.0, 3500.0)
1000	150	(10.0, 100.0, 1000.0)
1000	50	(4.0, 40.0, 400.0)
1000	20	(15.0, 150.0, 1500.0)
400	150	(40.0, 400.0, 4000.0)
400	50	(8.0, 80.0, 800.0)
400	20	(4.0, 40.0, 400.0)
200	50	(20.0, 200.0, 2000.0)
200	20	(7.0, 70.0, 700.0)
125	50	(50.0, 500.0, 5000.0)
125	20	(13.0, 130.0, 1300.0)

$M_{\tilde{q}}$ (GeV/c $^2$ )	$M_{\tilde{\chi}_0^1}$ (GeV/c $^2$ )	$c\tau$ (cm)
1500	494	160
1000	148	60
350	148	173
120	48	165

Table 3: The simulated background samples used in the analysis. The `DYJetsToLL` samples are the leading background and include Drell-Yan production of all three lepton flavours. The QCD background is smaller. It is modelled with the `Mu-Enriched QCD` samples. The `Mu-Enriched QCD` samples contain QCD events where there is at least one generator level muon with  $p_T > 15$  GeV/c (or 5 GeV/c at low  $\hat{p}_T$ ). All samples are from the Summer 2012 `DR53X` production. The event weighting factor is shown for an integrated luminosity of 20 fb $^{-1}$ . The cross section includes the efficiency of the generator-level filter, if applicable.

Dataset name	Cross section (pb)	Number of events	Weight Factor
<code>DYJetsToLL_M-10To50</code>	1.25e+04	3.78e+07	6
<code>DYJetsToLL_M-50</code>	3.5e+03	3.05e+07	2.3
<code>WW</code>	54.8	1e+07	0.11
<code>WZ</code>	33.2	1e+07	0.0664
<code>ZZ</code>	17.6	9.8e+06	0.0359
<code>TTJets_FullLeptMGDecay</code>	24.8	1.21e+07	0.041
<code>WJetsToLNu</code>	3.63e+04	1.84e+07	39.4
<code>QCD_Pt_15to20_MuEnrichedPt5</code>	2.74e+06	1.72e+06	3.18e+04
<code>QCD_Pt_20_MuEnrichedPt_15</code>	1.35e+05	2.15e+07	125

## 174 4 Event reconstruction and selection

### 175 4.1 Displaced track reconstruction

176 By design, the analysis does not use the silicon tracker information in muon track reconstruction.  
 177 At CMS, there are a handful of algorithms which utilize only the hits in the muon chambers  
 178 to perform track reconstruction of muons. The two muon collections which are based  
 179 only on the muon chambers are refittedStandAlone (RSA) and standAlone (SA) muons. Whilst  
 180 there are structural similarities between the two, the differentiation arises due to the fact that  
 181 the RSA muon algorithm computes an additional final fit of the tracks by excluding the beam  
 182 spot, which provides more accuracy for displaced muon measurements. On the other hand,  
 183 SA muons preserve the inherent bias towards the collision point, which is designed to analyze  
 184 muons coming directly from the beam spot, so called prompt muons.

185 The RSA muon collection improves transverse impact parameter,  $d_0$ , and transverse momentum,  
 $p_T$ , resolutions for displaced muons compared to those of the SA muons, as expected.  
 186 This has been confirmed in a study reported in [16] where the performances of RSA and SA  
 187 muons are compared. Therefore, the RSA muon collection is chosen for this analysis.

### 189 4.2 Muon selection

190 We require the RSA muons to satisfy  $p_T > 26 \text{ GeV}/c$  and pseudorapidity,  $|\eta| < 2$ . The mo-  
 191 mentum threshold is slightly higher than the corresponding trigger requirement, which is  
 192  $p_T > 23 \text{ GeV}/c$ , to ensure that the trigger has a good efficiency and its systematic uncertainty is  
 193 minimal.

194 A distinctive track rejection step is applied to make the analysis fully complementary to the  
 195 tracker-based analysis described in Ref. [6] and to exclude prompt muons in the most effective  
 196 way. All muons reconstructed in the muon chambers are rejected if they can be matched to  
 197 a track reconstructed in the silicon tracker with  $p_T > 10 \text{ GeV}/c$ . The matching is done by  
 198 extrapolating the track from the silicon tracker to the muon's innermost hit in the muon system.  
 199 The track and the muon are considered matched if  $\Delta R$  (where  $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$  between  
 200 the innermost hit of the muon and the extrapolated position of the tracker track in the muon  
 201 chambers) is less than 0.1. The  $p_T$  requirement on the tracker tracks is relaxed compared to  
 202 Ref. [6] to account for the low  $p_T$  resolution of RSA muons. Loosening this requirement further  
 203 does not lead to the removal of extra prompt events in data. No overlap is found in the events  
 204 selected by the two analyses when applying their full selection to all signal MC samples.

205 To select muons of good quality, the fit of the hits in the muon chambers to build each muon  
 206 track should meet the condition  $\chi^2/\text{dof} < 2$ . Each muon must have at least 3 muon stations  
 207 with at least a valid hit. Given non-negligible cosmic muon contamination, each muon is also  
 208 required to have at least 17 valid hits (DT+CSC+RPC) as a sanity check since in most cases out-  
 209 of-time muons with cosmic origin tend to have lower number of valid hits compared to in-time  
 210 muons, that is, muons arising from pp collisions. A detailed study of in-time and out-of-time  
 211 muons is presented in Appendix A. Finally, muons should have a transverse impact parameter  
 212 significance,  $|d_0|/\sigma_d > 4$ , where  $|d_0|/\sigma_d$  is the ratio of the transverse impact parameter to its  
 213 error. For muons passing this selection, except the track rejection step, the typical  $d_0$  resolution  
 214 is of  $\sim 2\text{cm}$  and the typical relative  $p_T$  resolution is  $\sim 40\%$  [16].

215 This analysis does not impose any explicit requirement on the isolation of the muons. Non-  
 216 isolated muons, such as those produced from a semileptonic b decay, can have a significant  
 217 impact parameter and represent a possible source of background for this analysis. To suppress  
 218 this source of background the muon is required to be isolated in [6]. In this analysis, however,

219 the track rejection step suppresses this source of background and no additional isolation re-  
 220 quirement is necessary. The absence of an isolation requirement makes this analysis sensitive  
 221 to models producing highly displaced b quarks, such as long-lived particles decaying to  $b\bar{b}$   
 222 where the two b produce a muon each.

### 223 4.3 Selection of long-lived exotica

224 The long-lived (LL) particle candidates are formed by pairing all muons in the event in all  
 225 possible combinations. There is no opposite charge requirement enforced when building the  
 226 dimuon candidates to eliminate the unfavourable effect of charge mis-measurement by the  
 227 muon system. As shown in Figure 5, the efficiency of correctly signed muons drops as a func-  
 228 tion of the generated muon  $p_T$  and as a function of the generated muon  $d_0$ . The efficiency goes  
 229 down to around 80% when the generated  $d_0 > 50$  cm, which suggests the opposite charge re-  
 230 quirement should not be applied in this analysis. Among all LL candidates that share the same  
 231 lepton, the one with the smallest  $\chi^2/\text{dof}$  of the secondary vertex, which two muon tracks are  
 232 fitted to, is kept. The procedure avoids the double counting of muons.

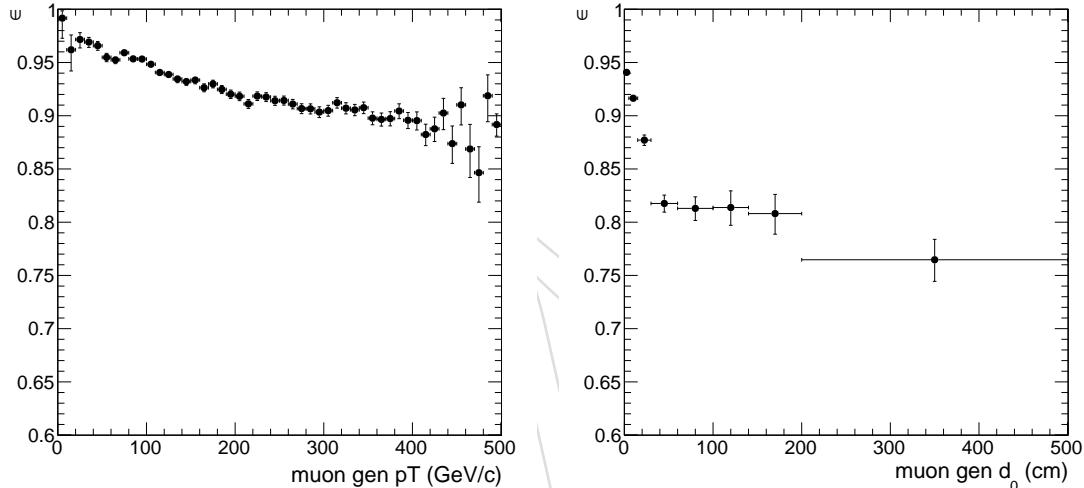


Figure 5: The efficiency of correctly signed muons as a function of the generated muon  $p_T$  (left) and as a function of the generated muon  $d_0$  (right). The signal sample shown on the plots has the following mass points:  $M_{H^0} = 1000 \text{ GeV}/c^2$  and  $M_X = 350 \text{ GeV}/c^2$  with  $c\tau = 350 \text{ cm}$ .

233 We discard dimuons consistent with coming from  $J/\psi$  and  $\Upsilon$  decays and  $\gamma$  conversions by  
 234 requiring an invariant mass greater than  $15 \text{ GeV}/c^2$ . Although the tracker track rejection step  
 235 should already remove this background, the minimum mass cut is kept as a sanity check. The  
 236 two muon tracks are required to form a secondary vertex with  $\chi^2/\text{dof} < 4$ . The detailed study  
 237 on the secondary vertex reconstruction is presented in Appendix F. The angular difference  
 238 in the azimuthal plane,  $\Delta\Phi$ , between the dimuon momentum vector and the vector from the  
 239 primary vertex to the dilepton vertex should satisfy  $|\Delta\Phi| < \pi/2$ , where  $\Delta\Phi$  is measured in  
 240 the range  $-\pi < \Delta\Phi < \pi$ . The diagram showing the simple geometry of a dimuon decay in  
 241 Figure 6 describes this collinearity angle pictorially. The region,  $|\Delta\Phi| < \pi/2$ , is called signal  
 242 region and the one with  $|\Delta\Phi| > \pi/2$  is defined as control region. The control region should be  
 243 signal-free, whereas the background should be symmetrically distributed in both regions.

244 A significant amount of background arises from cosmic rays, which may be reconstructed as  
 245 back-to-back muons that are often displaced from the primary vertex. Such events should,  
 246 in principle, be removed at trigger level. However, the trigger requirement  $\cos(\alpha) > -0.8$

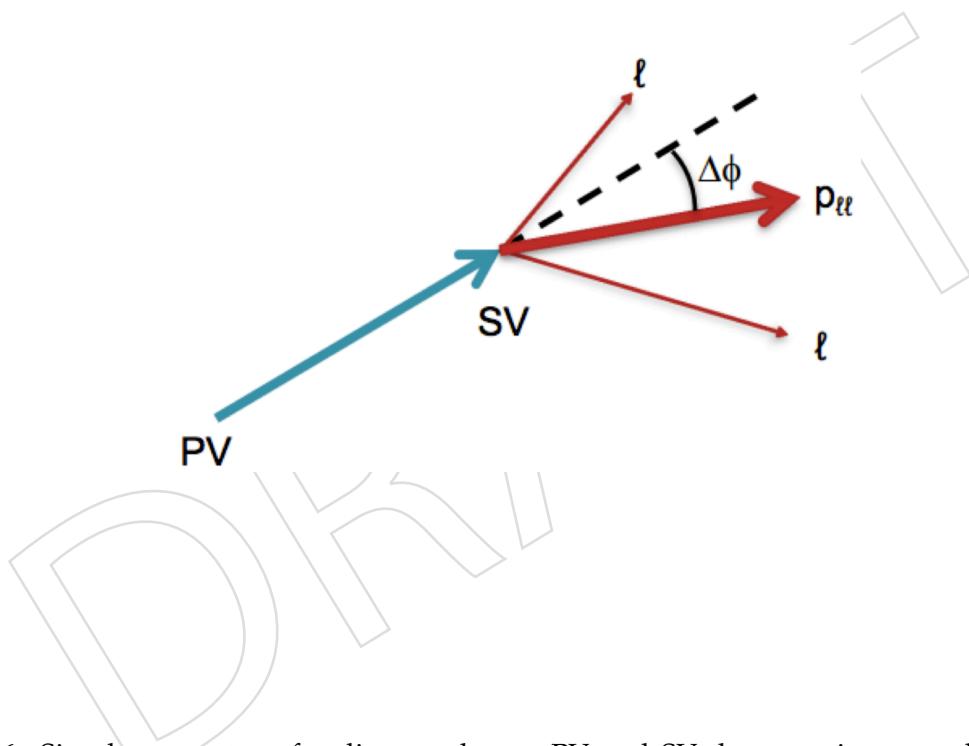


Figure 6: Simple geometry of a dimuon decay. PV and SV denote primary and secondary vertices, respectively. The dimuon momentum vector is represented by the thick red arrow.

is tightened to  $\cos(\alpha) > -0.75$ , where  $\alpha$  is the 3D opening angle between the two muons. Furthermore, a dimuon candidate can also be reconstructed from half a cosmic and another (fake or real) muon in the event. To remove these combinations, candidates are rejected when one of the two muons is back-to-back ( $\cos(\alpha) \leq -0.75$ ) to another muon in the same event that is not included in another dimuon candidate. An example event removed by this cut is shown in Figure 7.

The double muon trigger efficiency becomes difficult to model when the two muons are very close to each other. Hence, it is required that the two muons are separated by  $\Delta R > 0.2$ . Finally, LL candidates should have a transverse decay length significance of  $L_{xy}/\sigma_{L_{xy}} > 12$ , where  $L_{xy}$  is defined as the distance between the primary and the secondary vertices in the transverse plane and its resolution is studied comprehensively in Appendix E. The full selection is summarized in Table 4.

Table 4: Summary of the analysis selection.

Selection cut	Cut value
Min. muon $p_T$ (GeV/ $c$ )	26
Max. muon $ \eta $	2
Max. normalized muon $\chi^2$	2
Max. normalized vertex $\chi^2$	4
Min. $\Delta R$ between the two muons	0.2
Min. $\cos(\alpha)$	-0.75
Min. dimuon mass (GeV/ $c^2$ )	15
Max. $ \Delta\Phi $	$\pi/2$
Min. number of muon DT + CSC stations	3
Min. number of valid muon hits	17
Min. muon $ d_0 /\sigma_d$	4
Min. dimuon $L_{xy}/\sigma_{L_{xy}}$	12

We generate the signal to be within the CMS detector acceptance region given by:

- The generated transverse decay length  $L_{xy}$  of the LL particle must be  $< 500$  cm.
- The generated muon pseudorapidity must be  $|\eta| < 2$ .
- The generated muon momentum must satisfy  $p_T > 26$  GeV/ $c$ .

This defines a region where it is possible to reconstructed long-lived particle decays in CMS. We call acceptance (A) the fraction of long-lived particle decays that fall in the acceptance region. By considering the efficiency for decays within the acceptance  $\epsilon_1/A$  we obtain a quantity that is less model dependent.

Figure 8 demonstrates that our analysis is mostly sensitive to LL particles with long lifetimes while being completely insensitive to prompt events. That plot also suggests that although RSA muons have non-null reconstruction efficiency up to 5 meters away from the beam spot in the transverse plane [16] the effective range of the analysis is restricted to 2.5 meters since the dimuon trigger efficiency vanishes around that distance. To illustrate this limitation more clearly, the trigger efficiency given that the event is within the acceptance vs. generated transverse decay length,  $L_{xy}$ , graph is drawn for three different signal MC samples in Figure 9. The inefficiency of the trigger is due to a bias in the reconstructed muon  $p_T$  as a function of the transverse impact parameter. This bias causes the  $p_T$  requirement of the trigger to be inefficient for longer decay lengths. The effect is discussed in detail in Appendix H.

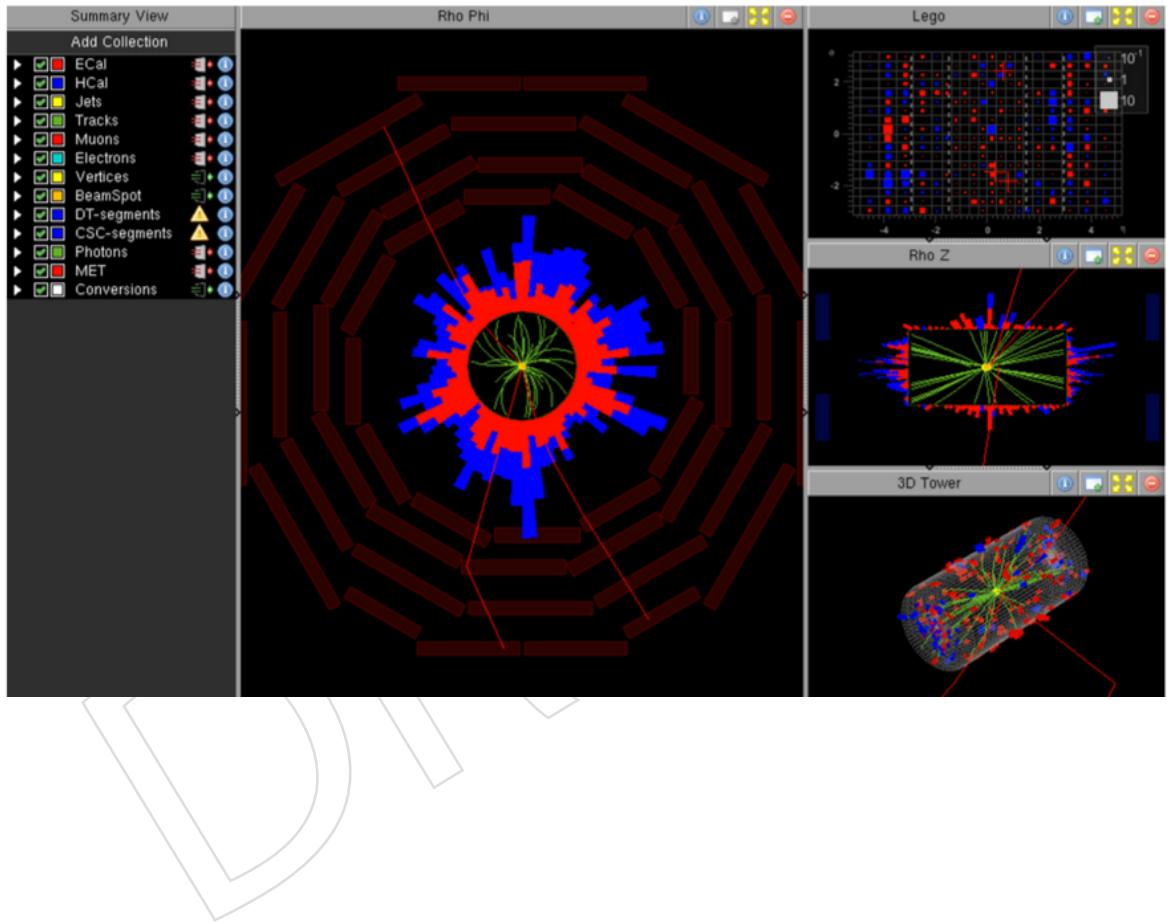


Figure 7: An example event removed by the cosmic rejection cut. Three muons are shown in red, two of which emerge from a cosmic ray muon. A LL candidate is reconstructed from half a cosmic and the other muon in the event that is independent of the cosmic. Note that only the track segments in the muon chambers are representative of the track direction. The segments at smaller radius are instead interpolated to the beamspot position by the visualization software, as it is designed to display prompt tracks.

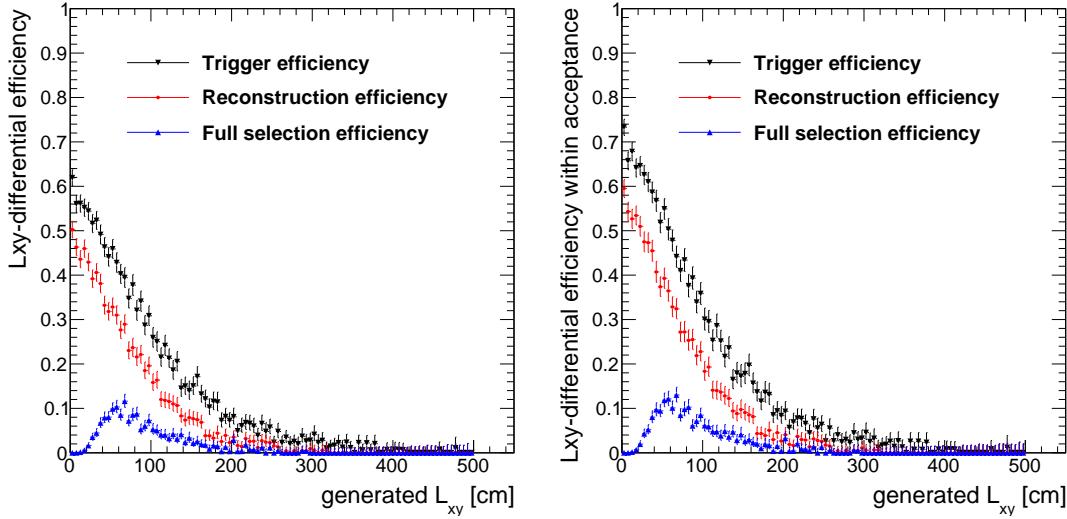


Figure 8: (Left) The efficiency to find muons from long-lived particles decays as a function of the generated  $L_{xy}$ . (Right) The efficiency computed only for particles decaying within the detector acceptance. Although the RSA muon reconstruction efficiency is non-null up to  $\approx 5$  meters in the transverse plane, the effective range of the analysis is up to 2.5 meters in the transverse plane due to the trigger efficiency. The signal sample shown on the diagram has the following mass points:  $M_{H^0} = 1000 \text{ GeV}/c^2$  and  $M_X = 350 \text{ GeV}/c^2$  with  $c\tau = 350 \text{ cm}$ .

#### 277 4.4 Results of blind analysis

278 The blind analysis has been completed with the full selection described above. The plots from  
 279 Figure 10 to Figure 17 show the distributions of the cut parameters in both control,  $|\Delta\Phi| >$   
 280  $\pi/2$ , and signal,  $|\Delta\Phi| < \pi/2$ , regions with all the selection applied except the one plotted. No  
 281 event in data passes the full selection in the control region of the analysis. This implies that  
 282 the number of expected background events in the signal region of data is zero as well, given  
 283 the established symmetry between the two regions, as validated in Section 5.1. The systematic  
 284 uncertainty from the expected background is estimated in Section 5.2.

285 In addition, the efficiency of each individual cut in the analysis selection is explicitly shown  
 286 in Table 5. The table contains the cut efficiencies of the signal and background MC samples in  
 287 the signal region and those of data in the control region. As expected, the cosmic rejection has  
 288 no effect on the background and signal MC samples, whilst it reduces the background level in  
 289 data by about a half.

290 The agreement between background MC and data samples used is shown in Appendix C.  
 291 The agreement is quite reasonable in the phase space of the analysis, though several minor  
 292 discrepancies are observed. We use data-driven methods for background estimation in this  
 293 analysis.

#### 294 4.5 Selection efficiency and acceptance

295 The selection efficiency and the limits are determined in terms of the number of events passing  
 296 our selection, rather than the number of the dimuon candidates. The full signal efficiency is  
 297 simply the ratio of the total number of generated events to the number of events in which at  
 298 least a LL candidate passes the full selection. It is computed separately for two different cases.  
 299 In the first case, the events that have only one generated LL particle ( $X$  or  $\tilde{\chi}^0$ ) decaying to  
 300 muons give the efficiency  $\epsilon_1$ ; whereas the efficiency for the events in which two LL particles

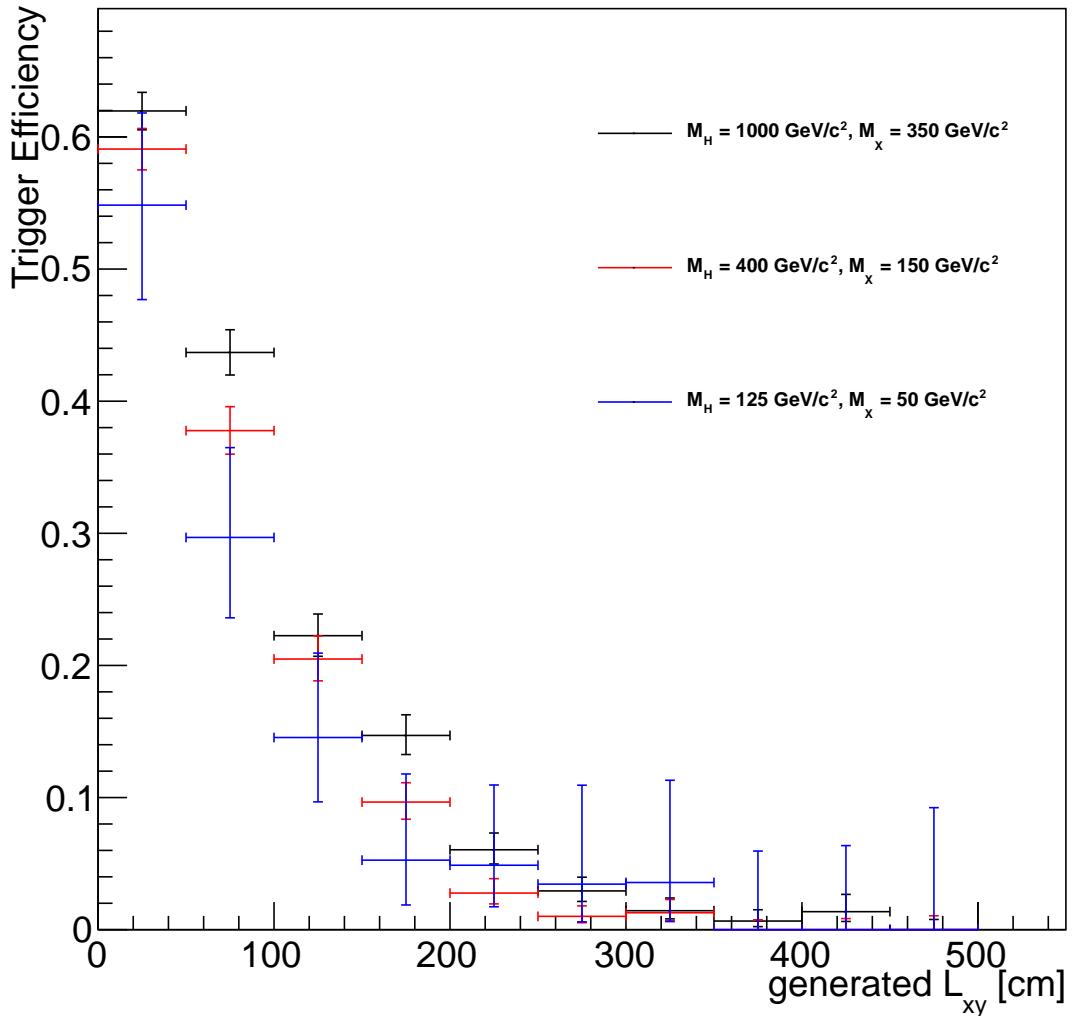


Figure 9: Trigger efficiency vs. generated transverse decay length for the  $H^0 \rightarrow XX$  signal model with three different mass combinations. The trigger efficiency is computed as the fraction of events within the acceptance that satisfy the trigger requirement.

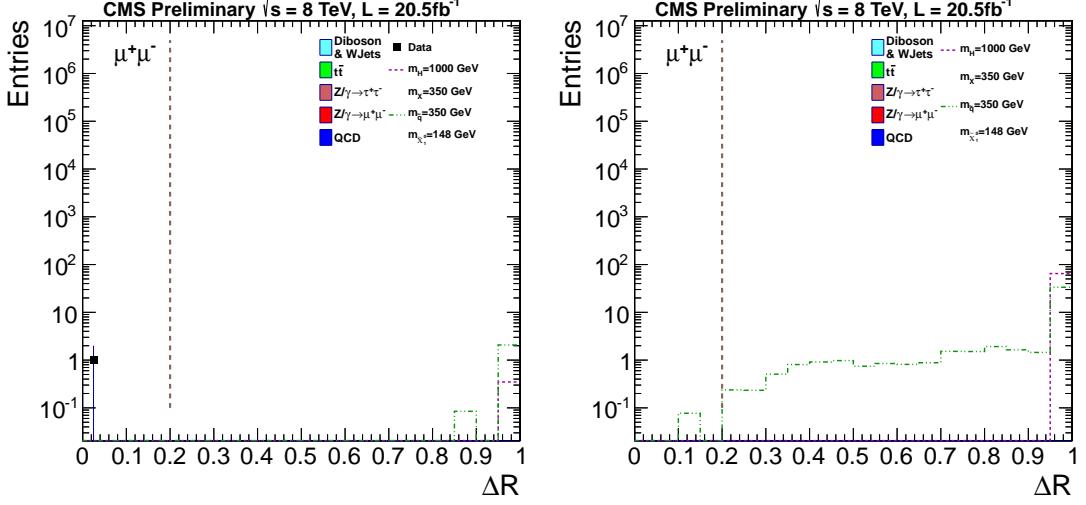


Figure 10: Distribution of  $\Delta R$  separation between the two muons for the dimuon candidates passing the full selection except the one plotted in the control region,  $|\Delta\Phi| > \pi/2$  (left) and in the blinded signal region,  $|\Delta\Phi| < \pi/2$  (right). An arbitrary cross section of 1 pb is assigned to the signal MC sample for good visualization of the signal events. The dashed line indicates the cut value of the parameter shown.

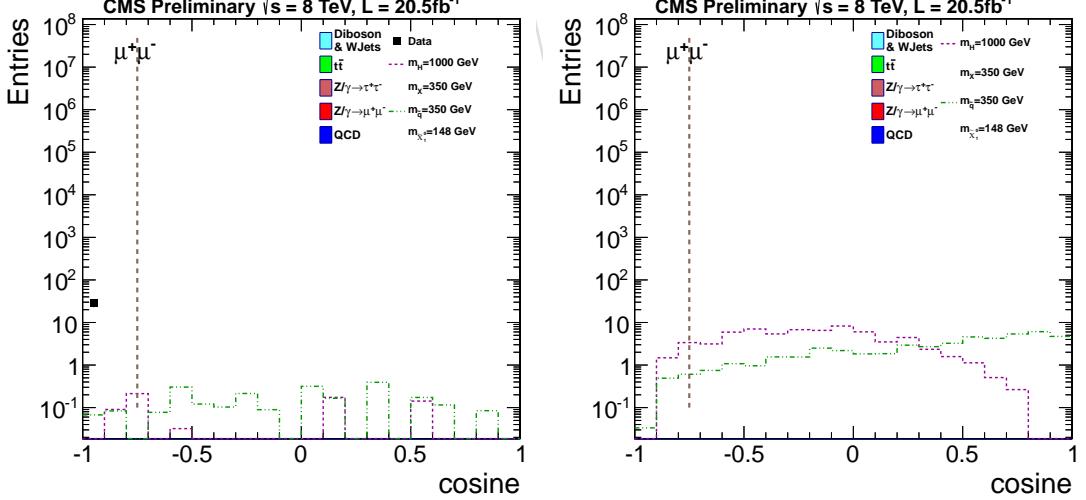


Figure 11: Distribution of the cosine of the 3D opening angle,  $\cos(\alpha)$ , between the two muons for the dimuon candidates passing the full selection except the one plotted in the control region,  $|\Delta\Phi| > \pi/2$  (left) and blinded signal region,  $|\Delta\Phi| < \pi/2$  (right).

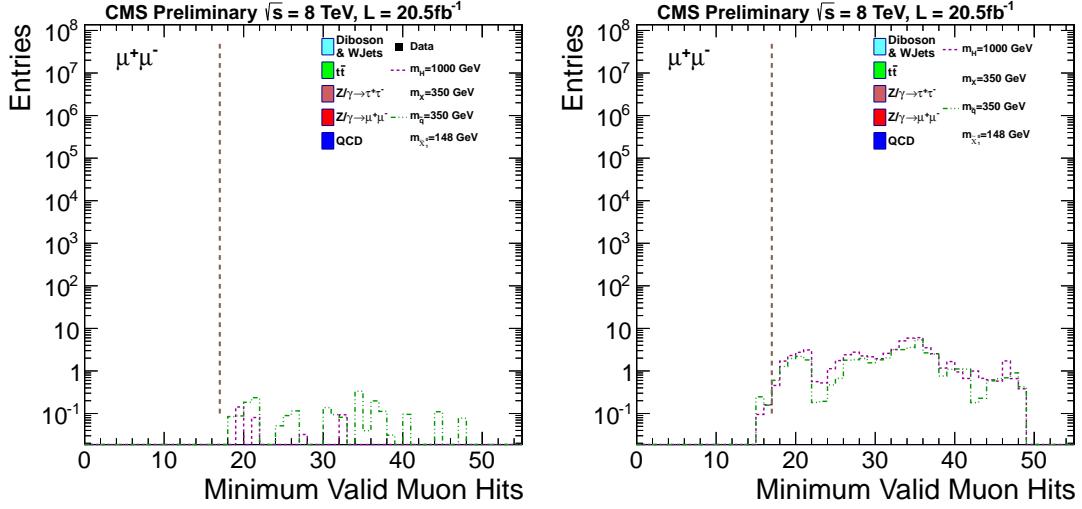


Figure 12: Distribution of the minimum number of valid muon hits of the two muons for the dimuon candidates passing the full selection except the one plotted in the control region,  $|\Delta\Phi| > \pi/2$  (left) and in the blinded signal region,  $|\Delta\Phi| < \pi/2$  (right).

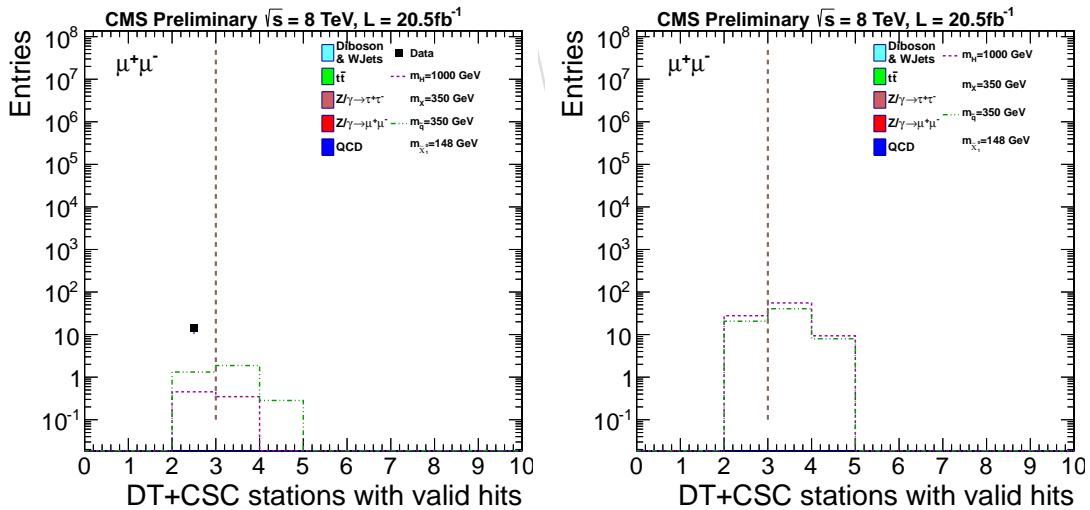


Figure 13: Distribution of the minimum number of valid muon stations of the two muons for the dimuon candidates passing the full selection except the one plotted in the control region,  $|\Delta\Phi| > \pi/2$  (left) and in the blinded signal region,  $|\Delta\Phi| < \pi/2$  (right).

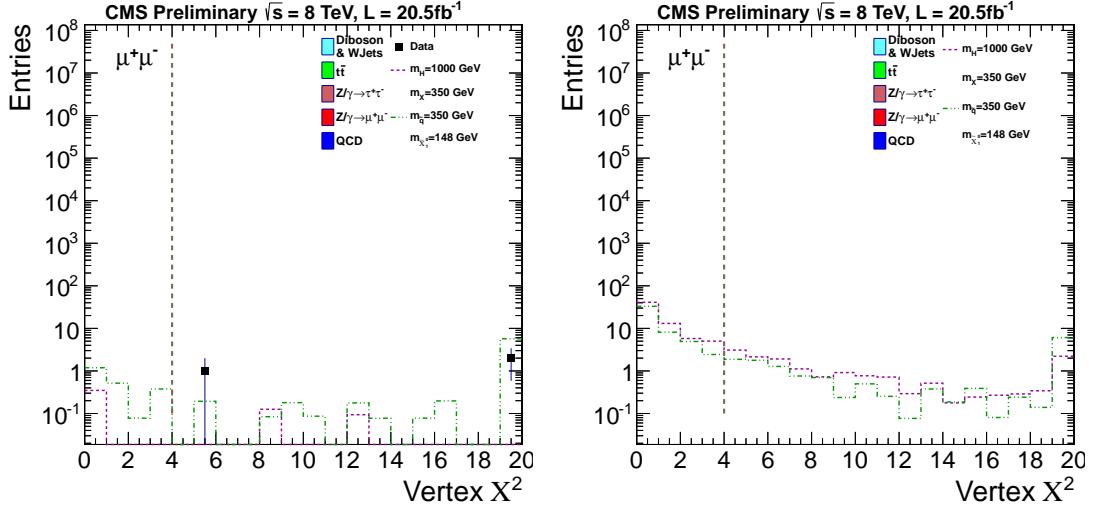


Figure 14: Distribution of the normalized vertex  $\chi^2$  of the dimuon candidates passing the full selection except the one plotted in the control region,  $|\Delta\Phi| > \pi/2$  (left) and in the blinded signal region,  $|\Delta\Phi| < \pi/2$  (right).

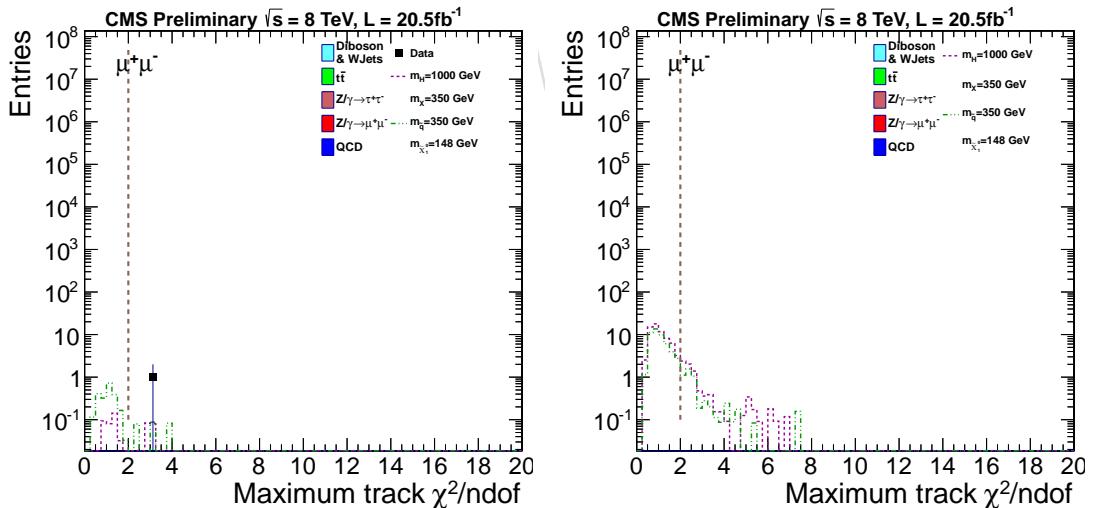


Figure 15: Distribution of the maximum normalized track  $\chi^2$  of the two muons for the dimuon candidates passing the full selection except the one plotted in the control region,  $|\Delta\Phi| > \pi/2$  (left) and in the blinded signal region,  $|\Delta\Phi| < \pi/2$  (right).

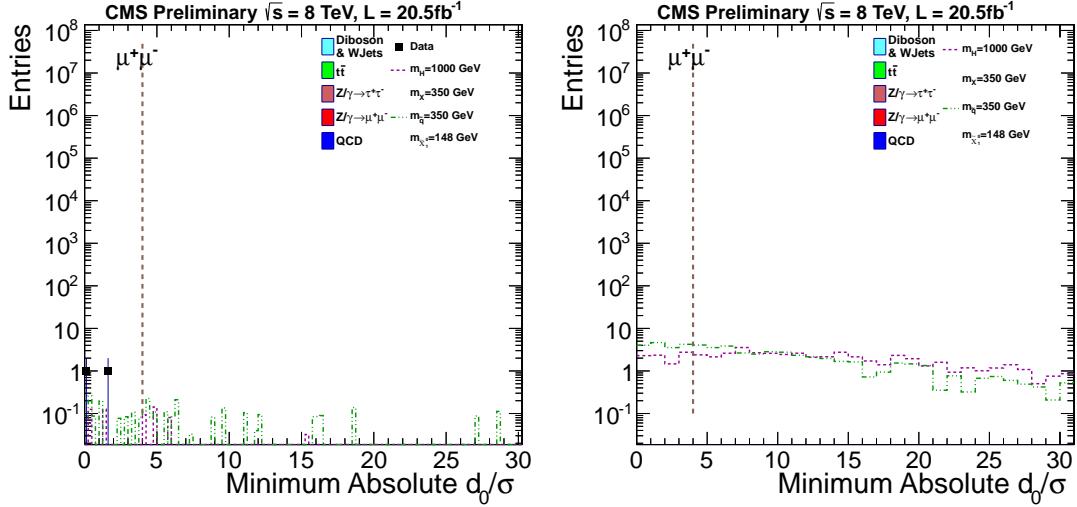


Figure 16: Distribution of the minimum absolute transverse impact parameter significance of the two muons for the dimuon candidates passing the full selection except the one plotted in the control region,  $|\Delta\Phi| > \pi/2$  (left) and in the blinded signal region,  $|\Delta\Phi| < \pi/2$  (right).

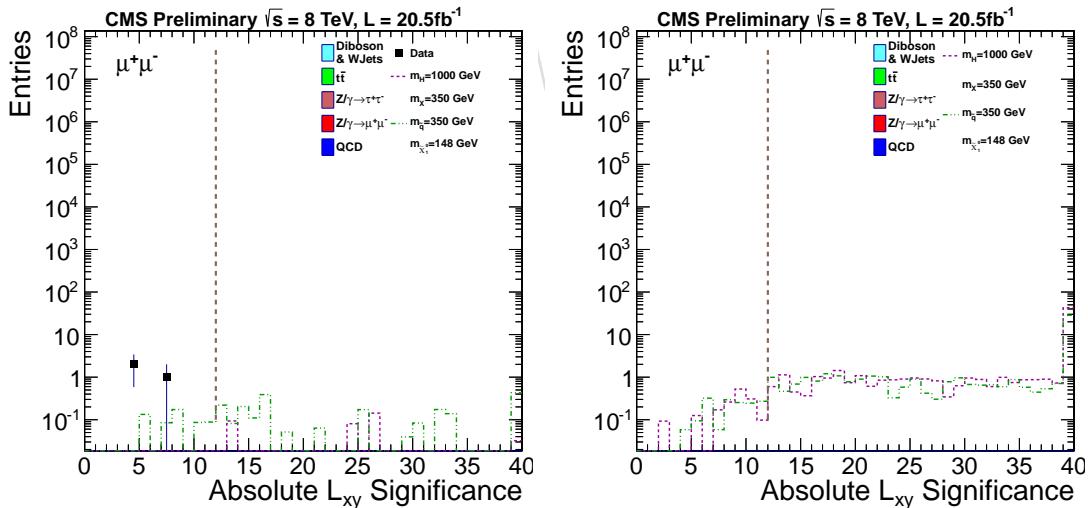


Figure 17: Distribution of the transverse decay length significance of the dimuon candidates passing the full selection except the one plotted in the control region,  $|\Delta\Phi| > \pi/2$  (left) and in the blinded signal region,  $|\Delta\Phi| < \pi/2$  (right).

Table 5: Dimuon candidate selection efficiencies. Each cut efficiency is the fraction of the candidates passing the cut given that the previous ones are already passed. Preselection is a cut that requires the presence of at least two muons with a transverse momentum,  $p_T > 26 \text{ GeV}/c$  in a triggered event. The cut efficiencies of the signal samples are shown for the events in which there is only one LL particle generated decaying to muons.

Selection Cut	Cut Efficiency			
	Data (Control Region)	Background MC	$H^0 \rightarrow XX(1000/350)$	$H^0 \rightarrow XX(125/20)$
Trigger	-	-	27.1%	33.4%
Preselection	-	-	68.6%	51.0%
Muon Matched to Tracker Trackers	2.8%	0.7%	56.8%	66.7%
Vertex $\chi^2$	67.7%	12.4%	85.3%	90.5%
Track $\chi^2$	58.3%	69.2%	90.7%	93.0%
Lepton $p_T$	99.0%	99.7%	100.0%	100.0%
Lepton $ \eta $	72.6%	78.6%	100.0%	100.0%
Dilepton Mass	93.6%	78.2%	100.0%	95.4%
$\cos(\text{angle between leptons})$	31.7%	97.4%	95.3%	100.0%
$\Delta R$	82.6%	100.0%	100.0%	99.3%
$ \Delta\Phi $	52.1%	64.6%	97.8%	99.5%
Candidate Contaminated with Cosmics	54.9%	99.0%	95.7%	96.1%
Minimum Number of Valid DT + CSC Stations	16.6%	11.6%	53.5%	45.9%
Minimum Number of Valid Muon Hits	96.7%	100.0%	99.2%	98.6%
Lepton $ d_0 /\sigma_d$	0.7%	0.0%	83.3%	56.4%
Dilepton	0.0%	0.0%	97.6%	97.4%

decay to muons is denoted by  $\epsilon_2$ . The efficiencies are estimated by reweighting the generated events with respect to the generated lifetimes. The efficiencies of the signal samples with the longest lifetime to pass the full selection of the analysis are tabulated in Table 6. In addition, the efficiencies of the simulated events that are within acceptance, as described in Section 4.2, to satisfy the selection criteria is given in Table 7.

Finally, the full selection is also implemented on the cosmics enriched sample to test how effective the analysis is to reject the events with cosmic origin. The analysis is run over the two cosmics datasets taken in 2012 between the run periods, Run B-Run C and Run C-Run D, at the LHC. The datasets contain 13 million events triggered by the dedicated cosmic muon trigger. We reconstruct 160,000 events consisting of dimuons. Only one of these events passes the analysis selection, corresponding to an efficiency of 0.000625% which is about 200 times smaller than the lowest signal efficiency of the analysis. Therefore, the events originating from cosmics are suppressed by the analysis.

## 5 Estimated background and associated systematic uncertainties

The signal and the background have a different distribution in  $|\Delta\Phi|$ . The signal is expected to have small  $|\Delta\Phi|$  values, while the background distribution is expected to be uniform in  $|\Delta\Phi|$  due to the absence of a genuine secondary vertex, as shown in Figure 18.

No data events are observed after the full selection is applied in the control region. Signal-control region symmetry sets the nominal value of the background expectation in the signal region to zero as well. The systematic uncertainty on this estimate is computed in Section 5.2.

Table 6: Efficiencies of the signal MC samples to pass the full selection of the analysis for the events where only one simulated LL particle decays to muons ( $\epsilon_1$ ), and for the events where two generated LL particles decay to muon pairs ( $\epsilon_2$ ).

$M_{H^0}$ (GeV/c $^2$ )	$M_X$ (GeV/c $^2$ )	$c\tau$ (cm)	Efficiency	
			$\epsilon_1$	$\epsilon_2$
1000	350	35.0	0.033	0.057
		350.0	0.023	0.049
		3500.0	0.0033	0.0069
1000	150	10.0	0.018	0.025
		100.0	0.044	0.073
		1000.0	0.0069	0.014
1000	50	4.0	0.0037	0.005
		40.0	0.018	0.03
		400.0	0.0027	0.0057
1000	20	1.5	0	4.7e-05
		15.0	0.00043	0.00096
		150.0	0.00029	0.00018
400	150	40.0	0.028	0.053
		400.0	0.017	0.038
		4000.0	0.0021	0.0042
400	50	8.0	0.0075	0.012
		80.0	0.029	0.053
		800.0	0.0058	0.01
400	20	4.0	0.0016	0.0013
		40.0	0.0079	0.014
		400.0	0.0015	0.0023
200	50	20.0	0.0089	0.013
		200.0	0.0081	0.02
		2000.0	0.0012	0.0019
200	20	7.0	0.0018	0.0024
		70.0	0.0085	0.015
		700.0	0.0014	0.0034
125	50	50.0	0.003	0.0081
		500.0	0.0012	0.0033
		5000.0	0.00017	0.00021
125	20	13.0	0.0016	0.0029
		130.0	0.0018	0.0047
		1300.0	0.00039	0.00081
$M_{\tilde{q}}$ (GeV/c $^2$ )	$M_{\tilde{\chi}_0^1}$ (GeV/c $^2$ )	$c\tau$ (cm)		
			$\epsilon_1$	$\epsilon_2$
1500	494	160.0	0.025	0.056
1000	148	60.0	0.027	0.058
350	148	173.0	0.017	0.036
120	48	165.0	0.0023	0.005

Table 7: Efficiencies of the signal MC samples to satisfy the full selection for events within the acceptance. The efficiencies are computed separately for the events where only one simulated LL particle decays to muons ( $\epsilon_1$ ), and for the events where two generated LL particles decay to muon pairs ( $\epsilon_2$ ).

$M_{H^0}$ (GeV/c $^2$ )	$M_X$ (GeV/c $^2$ )	$c\tau$ (cm)	Efficiency	
			$\epsilon_1$	$\epsilon_2$
1000	350	35.0	0.038	0.069
		350.0	0.032	0.071
		3500.0	0.022	0.034
1000	150	10.0	0.022	0.034
		100.0	0.065	0.12
		1000.0	0.048	0.082
1000	50	4.0	0.0045	0.0071
		40.0	0.029	0.045
		400.0	0.023	0.036
1000	20	1.5	0	0
		15.0	0.00068	0.00095
		150.0	0.0019	0
400	150	40.0	0.038	0.078
		400.0	0.029	0.069
		4000.0	0.016	0.034
400	50	8.0	0.013	0.023
		80.0	0.06	0.11
		800.0	0.054	0.093
400	20	4.0	0.0027	0.0027
		40.0	0.017	0.03
		400.0	0.017	0.021
200	50	20.0	0.028	0.049
		200.0	0.034	0.1
		2000.0	0.027	0
200	20	7.0	0.0058	0.0073
		70.0	0.033	0.067
		700.0	0.033	0.12
125	50	50.0	0.027	0.11
		500.0	0.018	0.075
		5000.0	0.015	0
125	20	13.0	0.014	0.0066
		130.0	0.022	0.041
		1300.0	0.032	0
$M_{\tilde{q}}$ (GeV/c $^2$ )	$M_{\tilde{\chi}_0^1}$ (GeV/c $^2$ )	$c\tau$ (cm)	$\epsilon_1$	$\epsilon_2$
1500	494	160.0	0.033	0.073
1000	148	60.0	0.04	0.088
350	148	173.0	0.034	0.07
120	48	165.0	0.031	0.059

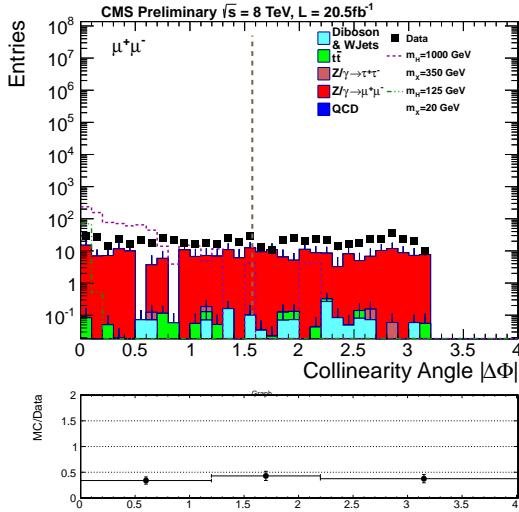


Figure 18: Collinearity angle,  $|\Delta\Phi|$ , distribution for the dimuon candidates passing the full selection except  $|d_0|/\sigma_d$ ,  $L_{xy}/\sigma_{L_{xy}}$  and  $|\Delta\Phi|$  cuts. The signal events plotted have only one LL particle generated decaying to muons. An arbitrary cross section value of 1 pb is assigned to the signal MC sample for good visualization of the signal events. The dashed line indicates the cut value of the parameter shown.

### 321 5.1 Background validation

322 The background symmetry is confirmed for data by comparing the  $L_{xy}/\sigma_{L_{xy}}$  tail-cumulative  
 323 distribution in the signal region with that in the control region at modest  $L_{xy}/\sigma_{L_{xy}}$  and  $|d_0|/\sigma_d$   
 324 values where the data is background-dominated. Similarly, the study is also repeated using  
 325 simulated background events, though the distribution is not expected to have similar normal-  
 326 ization as in data since background MC samples do not fully describe the data. For both studies,  
 327 the full selection except the  $L_{xy}/\sigma_{L_{xy}}$  cut is implemented and the cut on the minimum  $|d_0|/\sigma_d$  is  
 328 reversed, minimum  $|d_0|/\sigma_d < 4$ . Note that this requires only the smallest of the two  $|d_0|/\sigma_d$  to  
 329 be less than four, while no explicit requirement is made on the biggest of the two. The  $L_{xy}/\sigma_{L_{xy}}$   
 330 tail-cumulative plot for data excludes the region  $L_{xy}/\sigma_{L_{xy}} > 6$  where potential signal events  
 331 might appear.

332 Figure 19 shows the tail-cumulative distributions (i.e., the integral from the cut value on the  
 333 horizontal axis to infinity) of  $L_{xy}/\sigma_{L_{xy}}$  in the signal and control regions for both data and simu-  
 334 lated background events. Additionally, the discrepancy between control and signal regions is  
 335 expressed in terms of the combined uncertainty for each bin. The difference is found to be in  
 336 agreement with the symmetric background hypothesis. In addition, the background symmetry  
 337 is confirmed for data also by loosening the selection further as given in Figure 20.

### 338 5.2 Background systematic uncertainties

339 There are three main classes of systematic uncertainty in this analysis. These are the uncertainty  
 340 on the integrated luminosity, the uncertainty in the signal selection, which will be discussed in  
 341 Section 6, and the uncertainty that arises when deriving the background estimate. To derive a  
 342 systematic uncertainty on the estimated background from data, a fit to the  $L_{xy}/\sigma_{L_{xy}}$  distribution  
 343 is performed in a background dominated region. The fit function is extrapolated to the signal  
 344 region and is used to obtain an estimated background. The difference of this estimate, plus  
 345 its uncertainty, from the nominal estimated background of zero events is taken as systematic

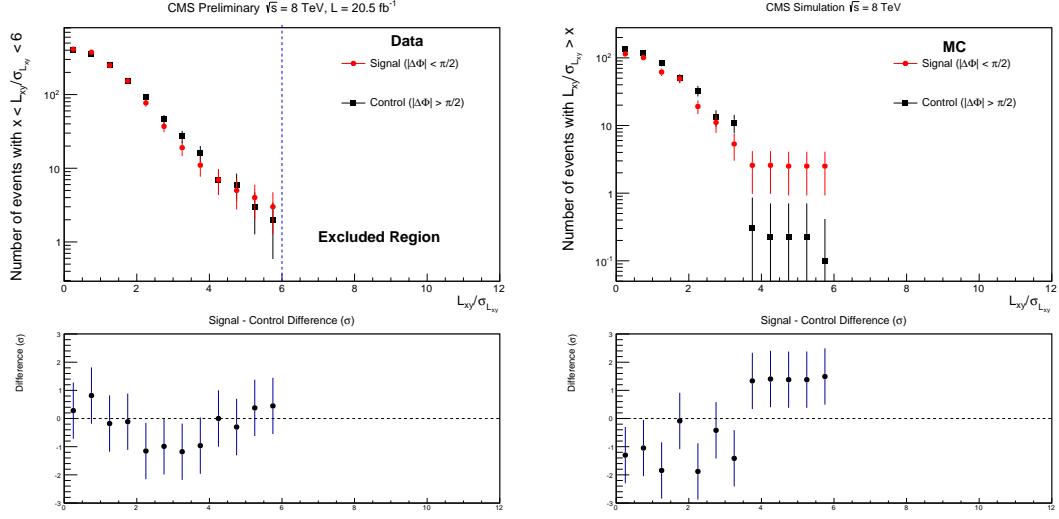


Figure 19: Comparisons of  $L_{xy}/\sigma_{L_{xy}}$  tail-cumulative distributions between signal,  $|\Delta\Phi| < \pi/2$ , and control,  $|\Delta\Phi| > \pi/2$ , regions for both data (upper left) and background MC samples (upper right). The full selection is applied with the exception of the  $L_{xy}/\sigma_{L_{xy}}$  cut. The  $|d_0|/\sigma_d$  cut is reversed to  $|d_0|/\sigma_d < 4$ . The plots on the bottom left and the bottom right show the statistical significance of the difference between the two regions for data and MC, respectively. The  $L_{xy}/\sigma_{L_{xy}} > 6$  region is excluded in data to avoid the possible signal contamination.

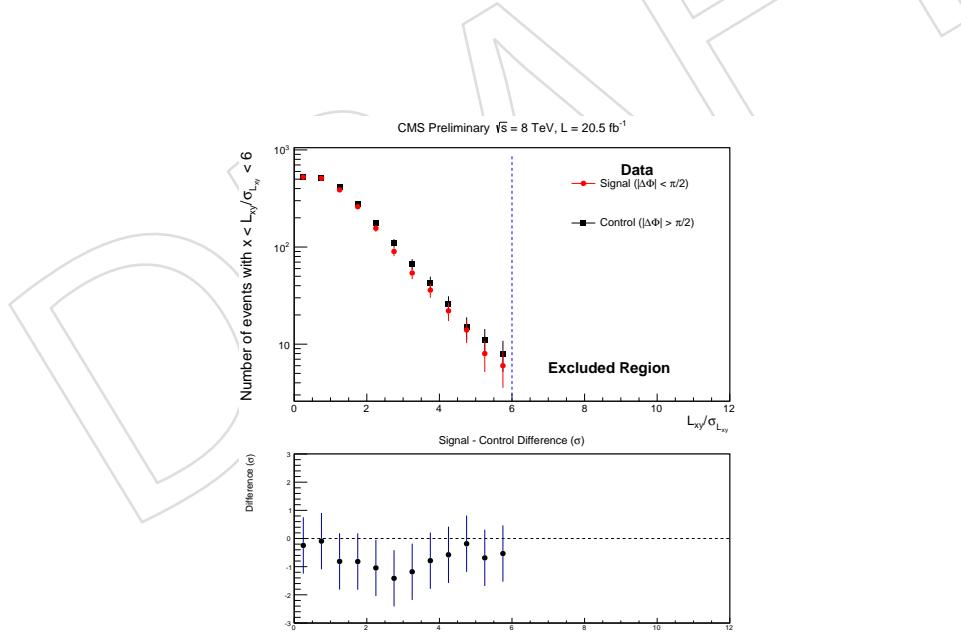


Figure 20: Comparison of  $L_{xy}/\sigma_{L_{xy}}$  tail-cumulative distributions between signal,  $|\Delta\Phi| < \pi/2$ , and control,  $|\Delta\Phi| > \pi/2$ , regions for data (top). The full selection is applied with the exception of the  $L_{xy}/\sigma_{L_{xy}}$ , vertex  $\chi^2/\text{dof}$ ,  $\cos(\alpha)$  and cosmic rejection cuts. The  $|d_0|/\sigma_d$  cut is reversed to  $|d_0|/\sigma_d < 4$ . The plot on the bottom shows the statistical significance of the difference between the two regions for data. The  $L_{xy}/\sigma_{L_{xy}} > 6$  region is excluded to avoid the possible signal contamination.

346 uncertainty. However, as there is not enough statistic to perform a meaningful fit of the back-  
 347 ground shape after the full selection, shown in the right plot of Figure 21, we remove the track  
 348 rejection cut and fit the resulting  $L_{xy}/\sigma_{L_{xy}}$  distribution, shown in the left plot of Figure 21. The  
 349 shape of the distribution is due to the  $|d_0|/\sigma_d$  cut. Since we are only interested in modelling  
 350 the shape of the right tail of the distribution, we perform a fit using a simple exponential of  
 351 the form  $Ae^{-\alpha L_{xy}/\sigma_{L_{xy}}}$  and we only fit the region  $L_{xy}/\sigma_{L_{xy}} > 7$ . The resulting fit is represented  
 352 by the red curve in the figure. Furthermore, it can be also inferred from Figure 21 that there  
 353 is only one event with  $L_{xy}/\sigma_{L_{xy}} > 7$  in the control region of the data. Under the assumption  
 354 above, the parameter  $A$  can be rescaled such that the overall function is normalized to unity  
 355 since there is one event in the region of the fit. Finally, the number of estimated background  
 356 events from this method can be calculated by integrating the rescaled function between 12 and  
 infinity. Back-of-the-envelope calculation yields  $0.18 \pm 0.03$ . We assign a systematic uncertainty

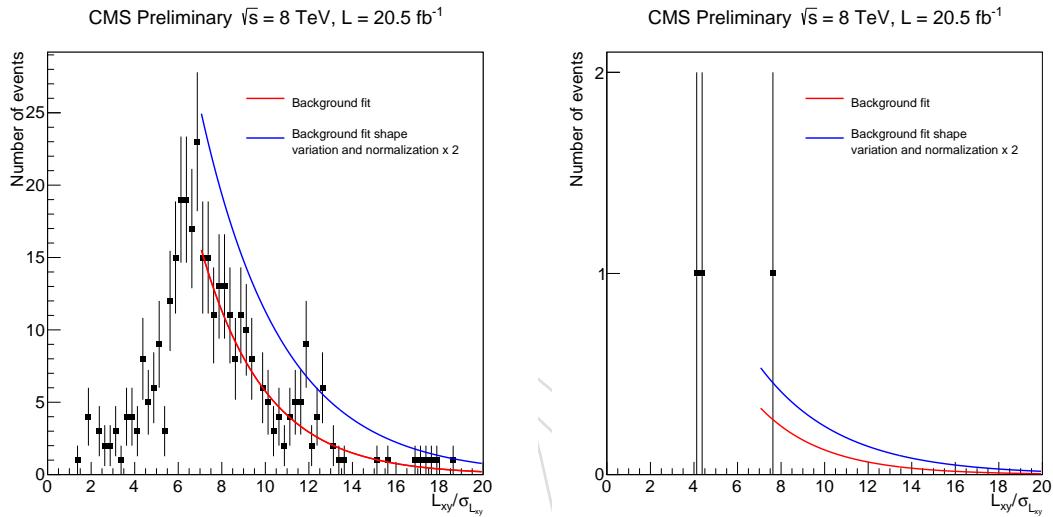


Figure 21: (Left)  $L_{xy}/\sigma_{L_{xy}}$  distribution after applying the full selection except the  $L_{xy}/\sigma_{L_{xy}}$  and the tracker track rejection cuts on the data in the control region. The shape is largely determined by the  $|d_0|/\sigma_d$  cut. (Right)  $L_{xy}/\sigma_{L_{xy}}$  distribution after applying the full selection except the  $L_{xy}/\sigma_{L_{xy}}$  cut. Only one event is found with  $L_{xy}/\sigma_{L_{xy}}$  larger than 7. In both figures the red curve shows the background shape normalized in the  $L_{xy}/\sigma_{L_{xy}} > 7$  region and the blue curve shows the variation of the shape with double normalization as explained in the text.

357 to the shape determination by varying the slope parameter,  $\alpha$ , by  $\pm 20\%$ . Furthermore, we take  
 358 into account the uncertainty on the background normalization due to observing only one event  
 359 after the full selection in the region  $L_{xy}/\sigma_{L_{xy}} > 7$  by doubling the normalization to two events.  
 360 This corresponds to a  $1\sigma$  coverage for the Poisson distribution given the observation of one  
 361 event. The red curve in Figure 21 shows the exponential curve obtained with both shape and  
 362 normalization variations. Combining the background estimate above and the variational effect  
 363 in the background shape and normalization gives an upper bound on the background estimate  
 364 of  $0.52 + 0.06$  (*stat*) events. We take this value (0.58 events) as systematic uncertainty on the  
 365 background estimated from data in the control region.

## 367 6 Signal systematic uncertainties

368 The dominant systematic uncertainties in the analysis are associated with the signal efficiency  
 369 and are caused by uncertainties in the trigger and reconstruction efficiencies of the displaced

370 RSA muons and by the pileup modelling in the simulation. A summary of the sources of  
 371 systematic uncertainty affecting the signal efficiency is presented in Table 8. In addition, we  
 372 consider the effect of pileup on the cosmic muon and tracker track rejection cuts and the effect  
 373 of RSA muon  $p_T$  resolution. They appear to be negligible and we do not assign any additional  
 374 systematic uncertainty, as detailed later.

Table 8: Systematic uncertainties related to the signal selection. The uncertainty specified is a relative uncertainty. The relative uncertainty in the luminosity is 2.6%.

Source	Uncertainty
Pileup modelling	2%
Tracking efficiency from cosmics	18%
Trigger efficiency	17%

## 375 6.1 Luminosity

376 For the running period corresponding to this analysis, CMS estimates the relative uncertainty  
 377 on the luminosity to be 2.6% [17]. This uncertainty is used when calculating the final cross  
 378 section estimates.

## 379 6.2 Effect of pileup

380 In order for the simulation to describe the pileup events in data realistically, the background  
 381 simulation events are reweighted to match the pileup in data by following the procedure given  
 382 in Ref. [18]. Data and reweighted background MC events are compared in Figure 22 in terms  
 383 of the number of reconstructed primary vertices, which is an estimate of the pileup.

384 The systematic uncertainty on the pileup modelling is estimated varying the average number  
 385 of reconstructed primary vertices in the background MC events by  $\pm 5\%$  as recommended in  
 386 Ref.[19]. The variation is realized through the re-reweighting of the simulation. The systematic  
 387 uncertainty on the signal efficiencies due to the pileup modelling is found to be less than 2%  
 388 for all signal MC samples used.

389 The more collisions that occur during the bunch crossing, the more tracks would be recon-  
 390 structed by the silicon tracker. This would increase the probability that we reject extra sig-  
 391 nal events from a mismatch. Therefore, the possible dependence of the track rejection and  
 392 cosmic rejection cuts on pileup is investigated on signal simulated  $H^0 \rightarrow XX$  events with  
 393  $M_{H^0} = 1000 \text{ GeV}/c^2$ ,  $M_X = 350 \text{ GeV}/c^2$  and  $c\tau = 350 \text{ cm}$ . The two cut efficiencies vs. the number  
 394 of reconstructed primary vertices are plotted in Figure 23 by applying the rest of the analysis  
 395 selection. Due to the limited statistics, a solid conclusion could not be drawn. Figure 24 shows  
 396 the pileup independence of the two cuts when the  $|d_0|/\sigma_d$  and  $L_{xy}/\sigma_{L_{xy}}$  cuts are removed from  
 397 the full selection and the minimum valid muon station requirement is loosened from 3 to 2.

## 398 6.3 Track finding and selection efficiency

399 To assess if the efficiency to reconstruct displaced muons in the muon chambers is correctly  
 400 modeled by the simulation, a direct measurement is performed utilizing cosmic ray muons.  
 401 Events are selected from dedicated runs with no beam activity (CRAFT) and the cosmic ray  
 402 muons are reconstructed as two separate RSA muons in opposite halves of the CMS detector.  
 403 The trigger used to collect the events during cosmic runs is a dedicated RPC technical trigger  
 404 that requires a signal in the RPC in the current and the previous two bunch crossings. This  
 405 requirement ensures that if a muon is coming from above the detector, it will reach the bottom

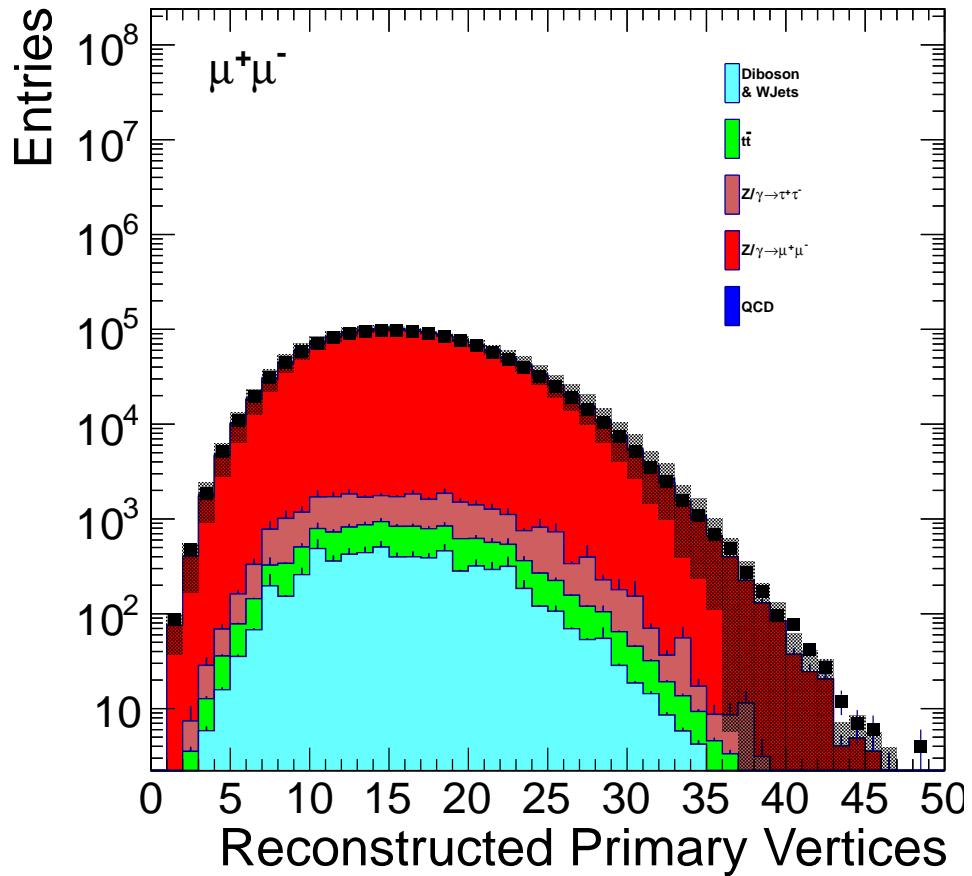


Figure 22: Distribution of the number of reconstructed primary vertices for data and the background simulation. The simulation is reweighted according to the procedure explained in the text. The full selection except the tracker track rejection and the cuts on  $|d_0|/\sigma_d$  and  $L_{xy}/\sigma_{L_{xy}}$  is applied. The grey vertical band represents the systematic uncertainty from varying the pileup weights by  $\pm 5\%$ .

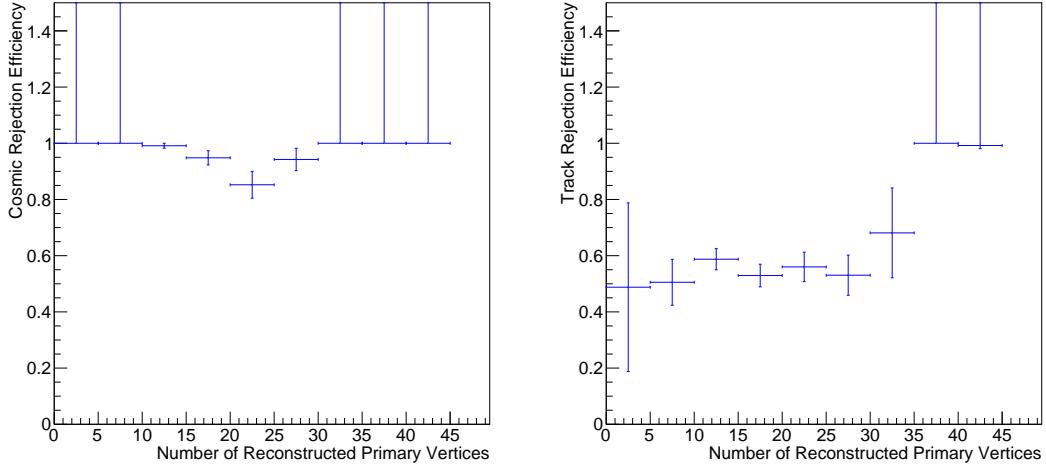


Figure 23: Efficiency of the cosmic muon rejection (left) and of the tracker track rejection (right) vs. the number of reconstructed primary vertices for a signal MC sample with  $M_{H^0} = 1000 \text{ GeV}/c^2$ ,  $M_X = 350 \text{ GeV}/c^2$  and  $c\tau = 350 \text{ cm}$ . The efficiency of the cut is computed after all other selection cuts are applied. In both cases no significant dependence on the number of reconstructed primary vertices is observed.

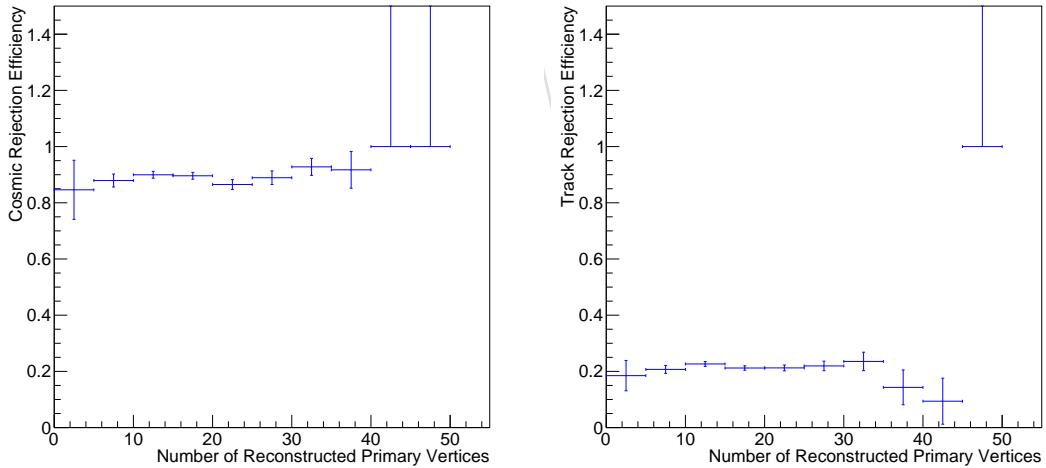


Figure 24: Efficiency of the cosmic muon rejection (left) and of the tracker track rejection (right) vs. the number of reconstructed primary vertices for a signal MC sample with  $M_{H^0} = 1000 \text{ GeV}/c^2$ ,  $M_X = 350 \text{ GeV}/c^2$  and  $c\tau = 350 \text{ cm}$ . The efficiency of the cut is computed after all other selection cuts are applied except the  $|d_0|/\sigma_d$  and  $L_{xy}/\sigma_{L_{xy}}$  cuts. The minimum muon valid station requirement is also loosened for both plots. In both cases no significant dependence on the number of reconstructed primary vertices is observed.

406 muon chambers in time with the readout of the detector. In essence, the bottom half of a cosmic  
 407 muon has the same timing, from the point of view of the detector readout, as a muon coming  
 408 from a collision. The same trigger is not available for the simulation. However, by generating  
 409 cosmic muons with a production time delayed by 25ns we achieve a similar distribution of  
 410 cosmic muons. The distribution of timing related variables for cosmic data and simulation for  
 411 all muon candidates is shown in Figure 25 and Figure 26. The simulation only reproduces  
 412 the main peak of the timeAtIpInOut for data. This is expected as muons arriving with 25 ns  
 413 or 50 ns before or after the ideal timing in data might still have a chance to pass the trigger  
 414 while the simulation does not generate such events. To select a consistent sample between data  
 415 and simulation we require that the muon at the top has a timeAtIpInOut in [-40, -20]ns. This  
 416 requirement also limits any bias due to timing on the measured efficiency since the bottom half  
 417 of the muon is in time with the detector readout.

418 The timing related variables after this selection for muons in the top half of CMS are shown  
 419 in Figure 27 and Figure 28. The comparisons between the kinematic distributions for cosmic  
 420 data and simulation are shown in Figure 29 and Figure 30 for all candidate muons and for the  
 selected ones, respectively.

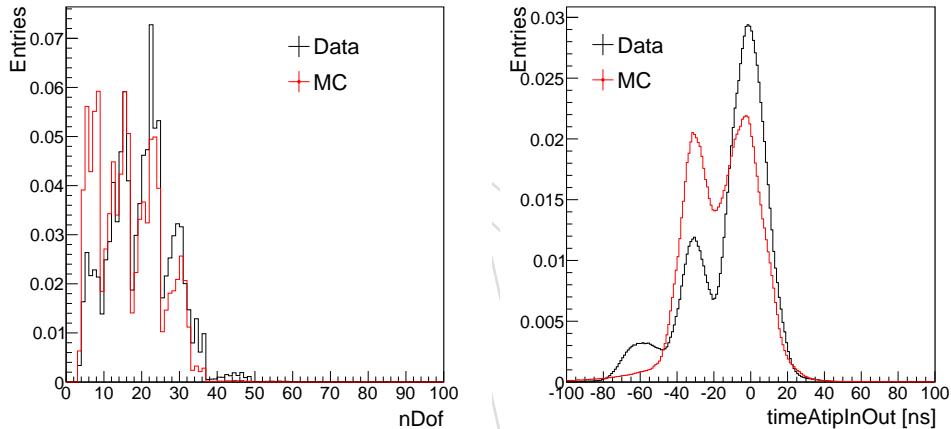


Figure 25: Number of degrees of freedom (left) and timeAtIpInOut (right) in cosmic data and simulation for all muon candidates.

421  
 422 We perform two measurements of the track finding efficiency. The first one is relative to the  
 423 silicon tracker and the second one uses only the muon chambers. In the first measurement we  
 424 require at least one track reconstructed in the silicon tracker passing the following selection:

- 425 •  $p_T > 26 \text{ GeV}/c$ ,  
 426 •  $|\eta| < 2$ ,  
 427 • at least 6 valid hits.

428 The efficiency is defined as the ratio of the number of reconstructed RSA muons in the bottom  
 429 half of CMS to the total number of tracks found in the silicon tracker. The RSA muons must  
 430 additionally satisfy the following criteria:

- 431 •  $p_T > 26 \text{ GeV}/c$ ,  
 432 •  $|\eta| < 2$ ,  
 433 • at least 17 valid hits in the DT or CSC muon chambers,

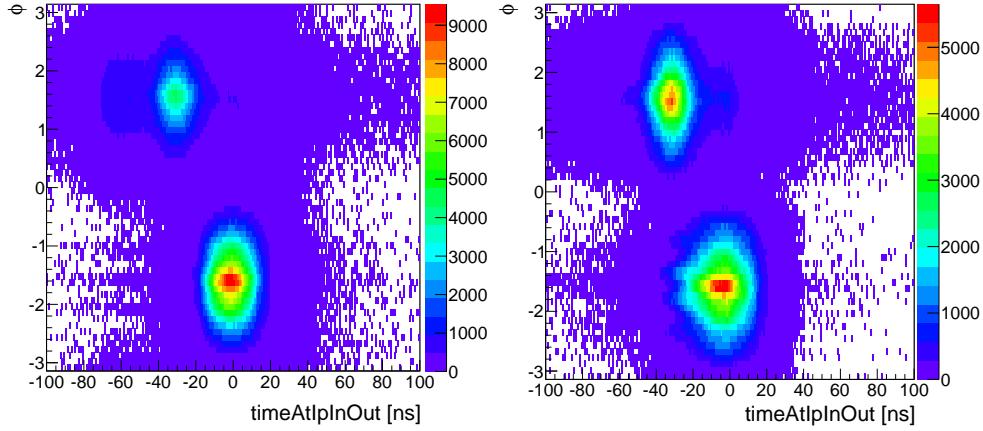


Figure 26:  $\text{timeAtIpInOut}$  vs  $\phi$  for data (left) and MC (right) for all muon candidates.

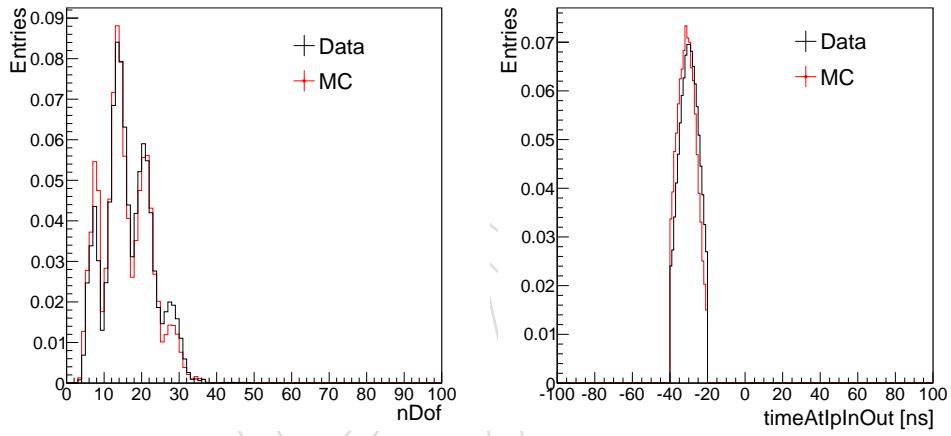


Figure 27: Number of degrees of freedom (left) and  $\text{timeAtIpInOut}$  (right) in cosmic data and simulation for muon candidates in the top half of CMS passing the timing selection.

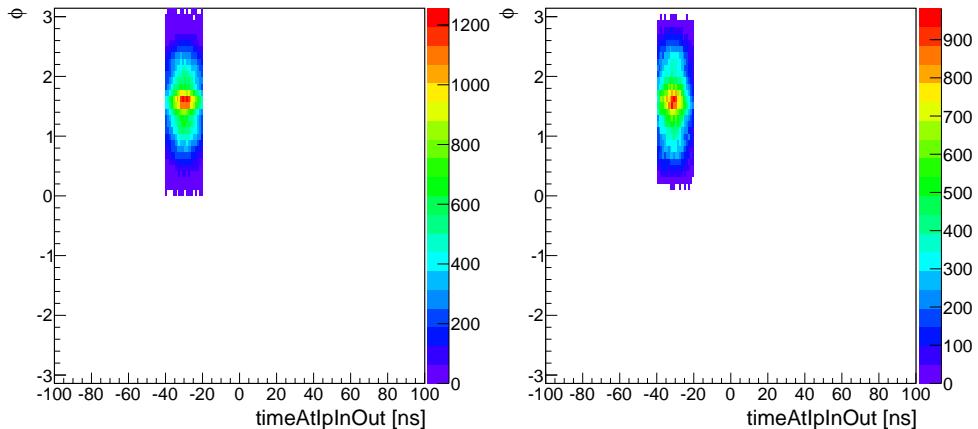


Figure 28:  $\text{timeAtIpInOut}$  vs  $\phi$  for data (left) and MC (right) for muon candidates in the top half of CMS passing the timing selection.

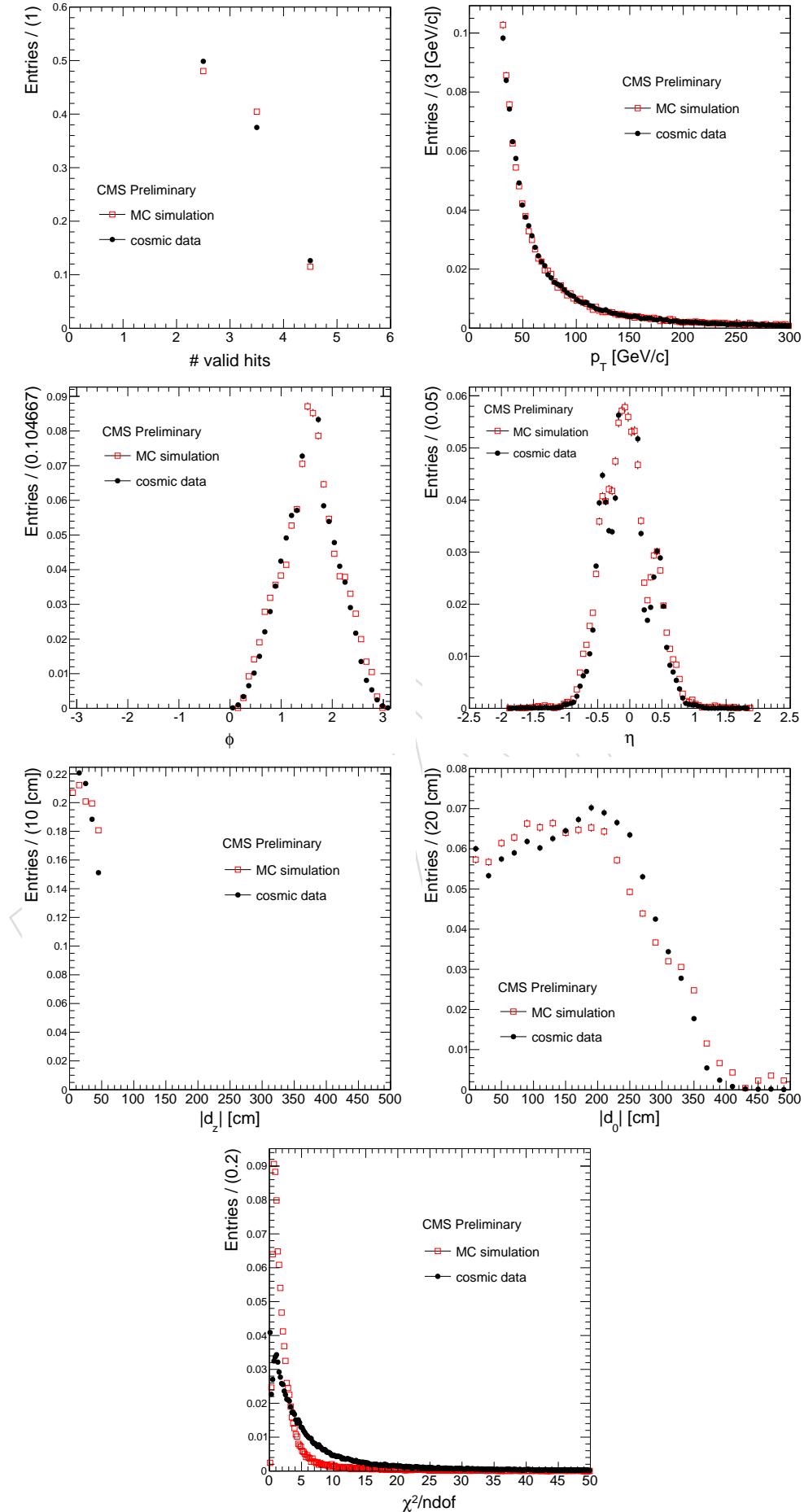


Figure 29: Comparison of kinematic distributions in cosmic data and simulation for all muon candidates.

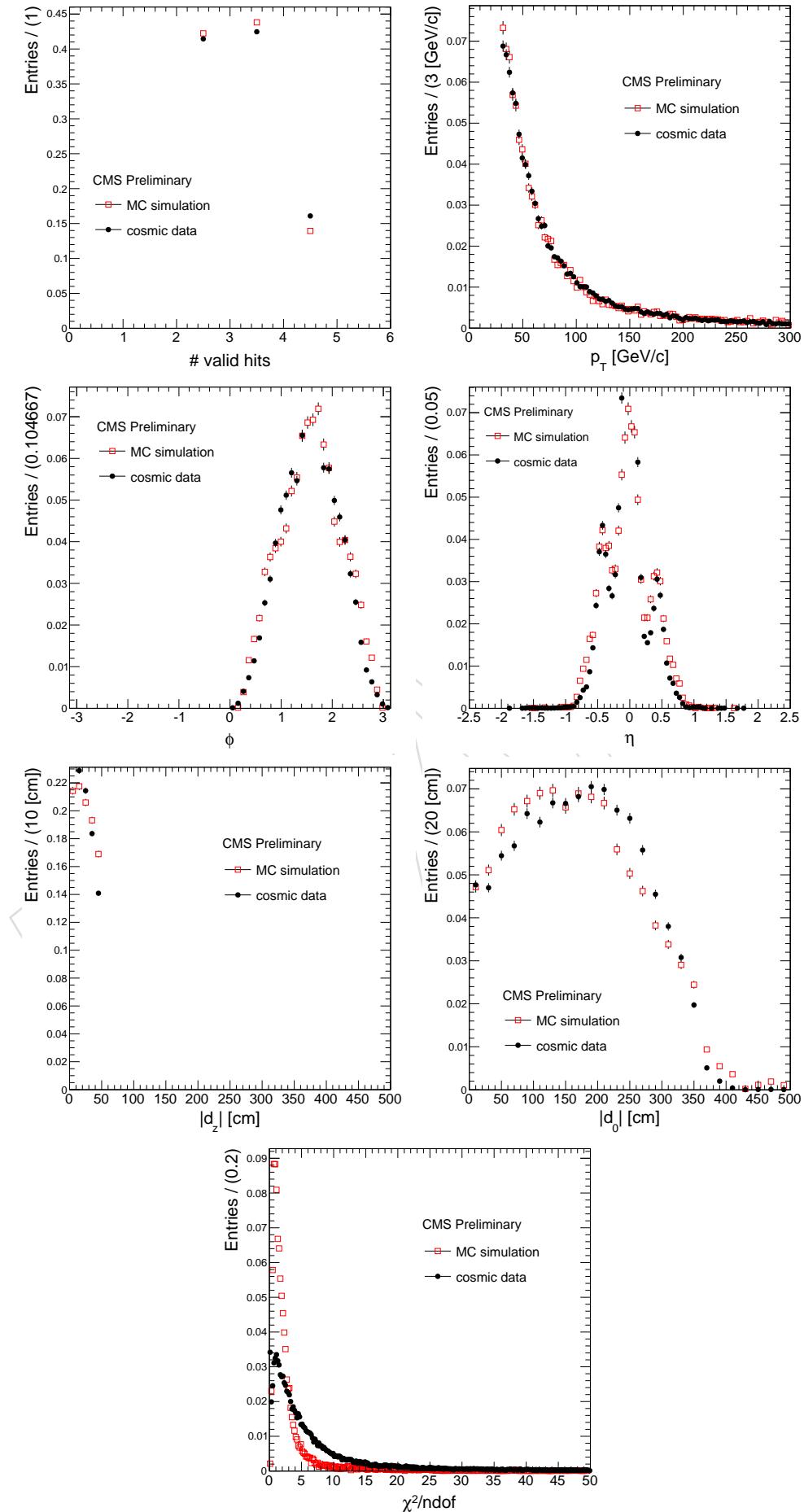


Figure 30: Comparison of kinematic distributions in cosmic data and simulation for muon candidates passing the selection detailed in the text.

- 434 • at least one valid hit in three DT or CSC stations.

435 The result is shown in Figure 31 as a function of the transverse impact parameter of the track.  
 436 By construction, this method is only sensitive to the impact parameters up to a few tens of  
 437 centimeters since the track finding efficiency for the silicon tracker is zero for higher values of  
 438 the impact parameters.

439 The second measurement allows to explore efficiencies for much higher impact parameter val-  
 440 ues. It requires that a RSA muon is reconstructed in the top half of CMS and that it passes the  
 441 following selection:

- 442 •  $p_T > 30 \text{ GeV}/c$ ,  
 443 •  $|\eta| < 2$ ,  
 444 •  $-40\text{ns} < \text{timeAtIpInOut} < -20\text{ns}$ ,  
 445 • at least two valid hits in two DT or CSC stations,  
 446 • maximum transverse and longitudinal impact parameter errors of 10cm,

447 while the bottom muon is required to satisfy:

- 448 •  $p_T > 26 \text{ GeV}/c$ ,  
 449 •  $|\eta| < 2$ ,  
 450 • at least 17 valid hits in the DT or CSC muon chambers,  
 451 • at least three valid hits in two DT or CSC stations.

452 The efficiency is computed as the fraction of bottom muons found when a top muon is also  
 453 found and is shown in Figure 32.

454 The results of these two measurements do not need to yield the same absolute value as they  
 455 are integrated over different timing distributions. We can select the timing for the top muons,  
 456 while for the silicon tracker tracks the timing is constrained by the charge integration time of the  
 457 detector and no direct measurement of the track arrival time is available. Additionally, because  
 458 the resolution on the impact parameters from silicon tracker tracks and from muon chamber  
 459 tracks is significantly different, the effective ranges analyzed are affected in different ways by  
 460 bin migration effects. The aim of these measurements is to provide a comparison between data  
 461 and simulation, so the conditions need only to be consistent within a given method.

To estimate the systematic uncertainty associated to the simulation of the track reconstruction and selection efficiency for the dimuon candidates, we take into account the  $|d_0|$  distributions of each muon in signal MC samples. If  $\langle \epsilon_{Data} \rangle$  and  $\langle \epsilon_{MC} \rangle$  are the weighted mean efficiencies to reconstruct both muon tracks in  $X \rightarrow \ell^+ \ell^-$  or  $\tilde{\chi}^0 \rightarrow \ell^+ \ell^- \nu$  decays, then the ratio of the two can be written as:

$$\frac{\langle \epsilon_{Data} \rangle}{\langle \epsilon_{MC} \rangle} = \frac{\sum_i \sum_j g_{MC}(i, j) \epsilon_{Data}(i) \epsilon_{Data}(j)}{\sum_i \sum_j g_{MC}(i, j) \epsilon_{MC}(i) \epsilon_{MC}(j)}$$

462 where:

- 463 •  $i$  and  $j$  are bins in  $|d_0|$  distributions of the two muons respectively as shown in Fig-  
 464 ure 33.  
 465 •  $g_{MC}(i, j)$  is the number of generated signal decays in which the two muons have  $|d_0|$   
 466 in bin  $(i, j)$  in a given MC sample.  
 467 •  $\epsilon_{Data}(i)$  and  $\epsilon_{MC}(i)$  are the efficiencies to reconstruct a single muon with  $|d_0|$  in bin  $i$   
 468 in data and MC, as given in Figure 32.

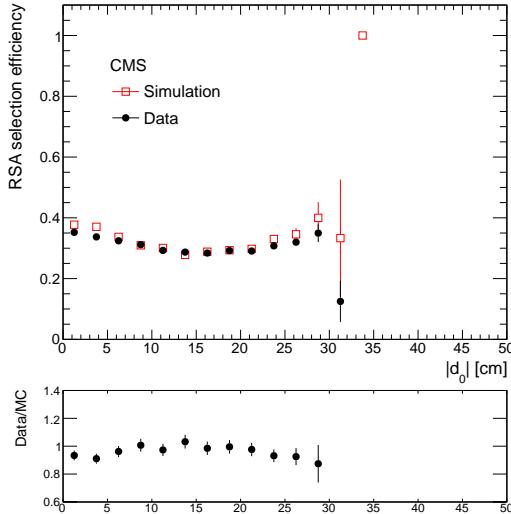


Figure 31: RSA muon reconstruction and selection efficiency measured by requiring the presence of a reconstructed track in the silicon tracker as a function of  $|d_0|$ .

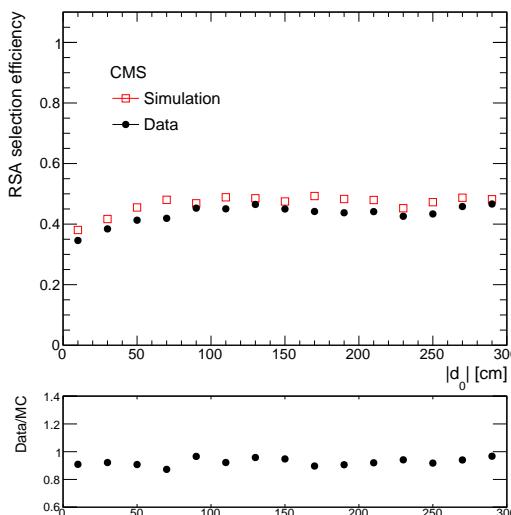


Figure 32: RSA muon reconstruction and selection efficiency measured using only the muon chambers as a function of  $|d_0|$  for  $|z_0| < 50\text{cm}$ .

469 The ratios for all signal MC samples are shown in Figure 34. We conclude from this figure that  
 470 a 18% systematic uncertainty per candidate would cover all considered signals.

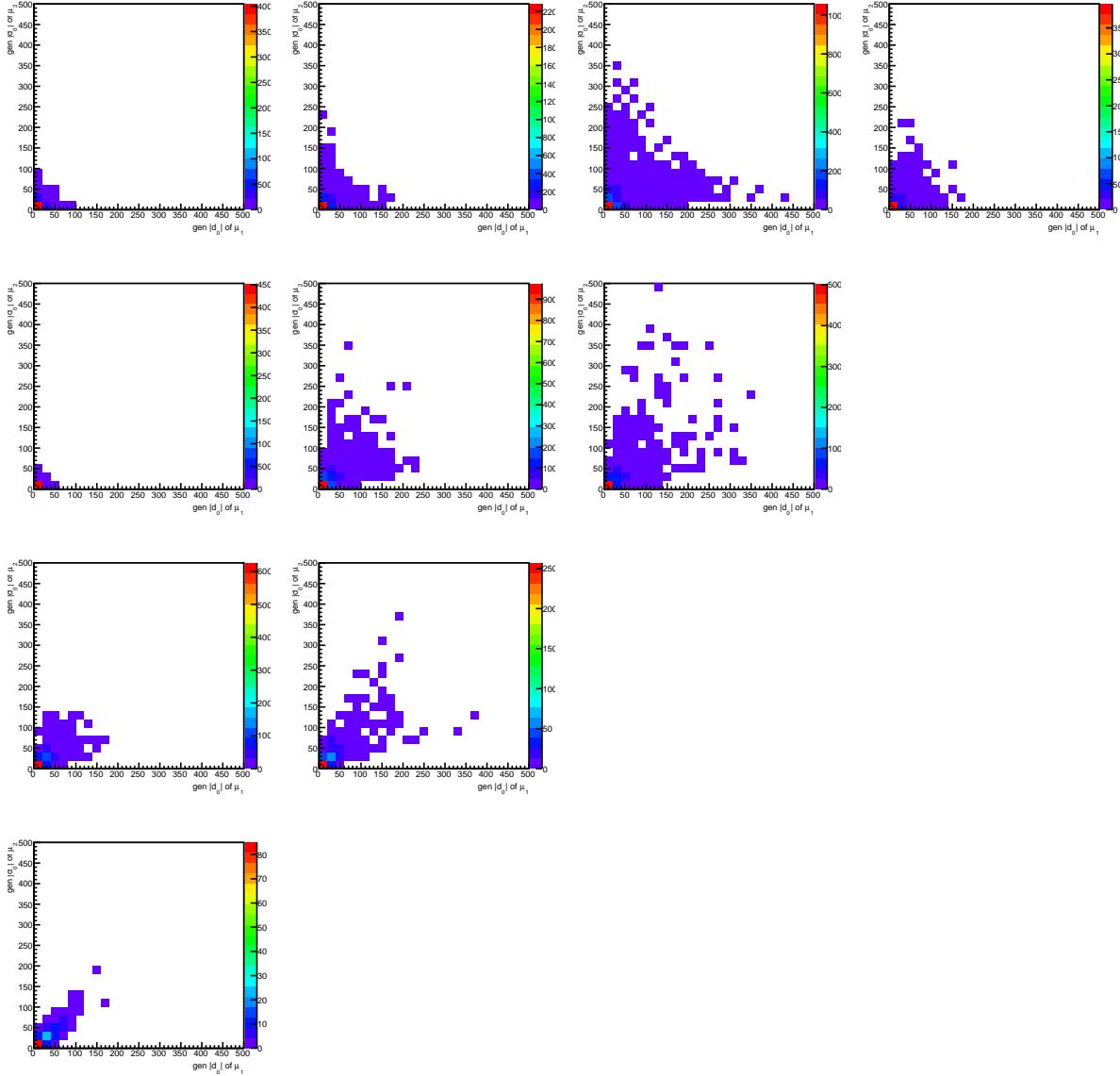


Figure 33:  $|d_0|$  distributions of both muons from the decay of the same LL particle in Monte Carlo signal samples. From left to right,  $M_X = 20, 50, 150,$  and  $350 \text{ GeV}/c^2$  and from top to bottom,  $M_H = 1000, 400, 200,$  and  $125 \text{ GeV}/c^2$ . The generated  $c\tau$  of each sample can be found from Figure 34.

#### 471 6.4 Trigger efficiency measurement

472 The systematic uncertainty assigned to the trigger efficiency to select the events analyzed is  
 473 simply the discrepancy associated with that measurement between data and the simulation.  
 474 The trigger efficiency is measured via the *Tag and Probe* method which can be exploited pro-  
 475 vided that there is a mass resonance, such as Z boson, decaying to muon pairs.  
 476 One of the muons, which comes from Z boson's decay, is labeled as *tag* that survives the tight  
 477 selection criteria that ensure that it is very unlikely to be fake. The other muon, which is as-

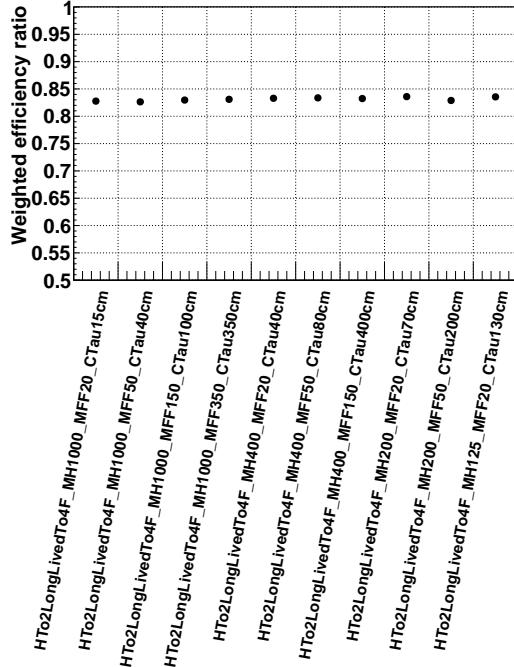


Figure 34: Ratio of the weighted mean efficiencies convoluted with the signal MC distributions.

478 sumed to have no correlation with the tag muon, is called probe. A passing probe, satisfying  
 479 selection cuts, is required to be matched to the trigger that selects the events for this analysis.  
 480 Finally, one fits the tag-probe mass distribution to extract the number of Z candidates for failing  
 481 and passing probes.

482 The analysis selects events collected by the trigger `HLT_L2DoubleMu23_NoVertex_2Cha_`–  
 483 `Angle2p5`, which will be denoted by `Trigger A` for convenience. This dimuon trigger re-  
 484 quires:

- 485 • Two L2 muons with  $p_T > 23 \text{ GeV}/c$  reconstructed with no vertex constraint.
- 486 • Each muon must have at least two DT or CSC muon stations with hits.
- 487 • The three dimensional angle between the two muons must be larger than 2.5 radians  
 488 (cosine  $> \sim -0.8$ ).

489 Because of the angle requirement in the trigger, the collected dataset is unsuitable to apply the  
 490 *Tag and Probe* method directly. Instead, we factor the trigger efficiency measurement into two  
 491 parts. First, we use the events that are collected by another double muon trigger, `HLT_L2Dou-`  
 492 `bleMu23_NoVertex`, henceforth called `Trigger B`, for the measurement. This trigger is iden-  
 493 tical to `Trigger A` except for the lack of the angle cut and of the requirement of a minimum  
 494 number of muon stations with hits. Secondly, since the actual dimuon trigger is equivalent to  
 495 `Trigger B` plus two additional cuts, we measure the efficiency of these additional cuts sepa-  
 496 rately to get the overall efficiency of `Trigger A` used in this search. Note that `Trigger B` is  
 497 prescaled in data by a factor of 20, which is accounted for in the measurement.

498 To measure the efficiency of `Trigger B`, the tag is chosen from the global muon collection and  
 499 is matched to an *IsoMu24* single muon trigger object within  $\Delta R < 0.1$ . Then, the following  
 500 selection cuts are applied for the tag muon:

- 501 •  $p_T > 26 \text{ GeV}/c$  and  $|\eta| < 2$
- 502 • Relative isolation ( $\text{isolationR}03.\text{sumPt})/\text{pt} < 0.1$
- 503 • Number of tracker layers with measurement  $\geq 6$
- 504 •  $|d_{xy}| < 30 \text{ cm}$  and  $|d_z| < 30 \text{ cm}$

505 The probe muons are required to be RSA muons used in the analysis. The conditions to be met  
506 for the probe muons are:

- 507 •  $p_T > 17 \text{ GeV}/c$  and  $|\eta| < 2.4$
- 508 • Number of DT + CSC muon stations with valid hits  $\geq 2$
- 509 • The probe is a passing probe if it is matched to the Trigger B object within  $\Delta R <$   
510 0.5

511 In addition, the following two criteria are applied for the pair selection between the tag and the  
512 probe:

- 513 •  $\Delta R > 0.2$
- 514 • Cosine of the angle  $> -0.79$

515 The efficiency for this measurement is input as an unknown parameter to the fitting. The trigger  
516 efficiency vs.  $p_T$  of the probe muon for Trigger B is shown in Figure 35.

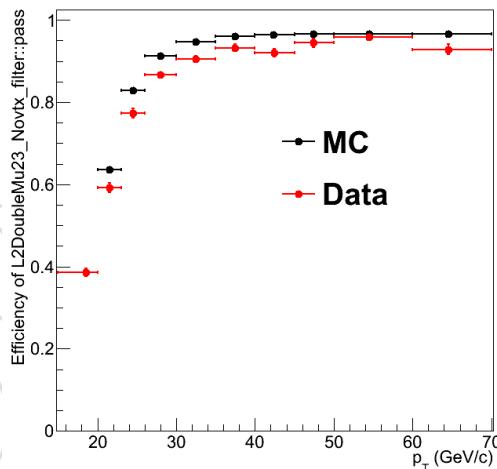


Figure 35: Trigger efficiency of `HLT_L2DoubleMu23_NoVertex` vs.  $p_T$  of the probe muon for both data and  $Z \rightarrow \mu\mu$  simulation, obtained with the *Tag and Probe* method.

517 The efficiency of the additional cuts included in Trigger A relative to Trigger B is mea-  
518 sured for the RSA muons. This efficiency is defined as the ratio of the number of events trig-  
519 gered by Trigger A, given that Trigger B is fired, to the number of events triggered by  
520 Trigger B only.

521 While keeping the selection requirements on RSA muons the same as described previously, the  
522 simultaneous fit is implemented to extract the ratio of the two trigger efficiencies as a function  
523 of the  $p_T$  of one of the two RSA muons, which is chosen randomly, in the event triggered by  
524 both triggers. In Figure 36, the ratio of the Trigger A efficiency to the Trigger B efficiency  
525 vs.  $p_T$  is presented. The efficiency to select two muons equals the square of the Trigger B  
526 efficiency multiplied by the ratio of Trigger A efficiency to Trigger B efficiency. The dis-

crepancy between data and simulation is no larger than 10%, which is taken as systematic uncertainty on the trigger efficiency measured for Z decays. The study is detailed in Appendix B.

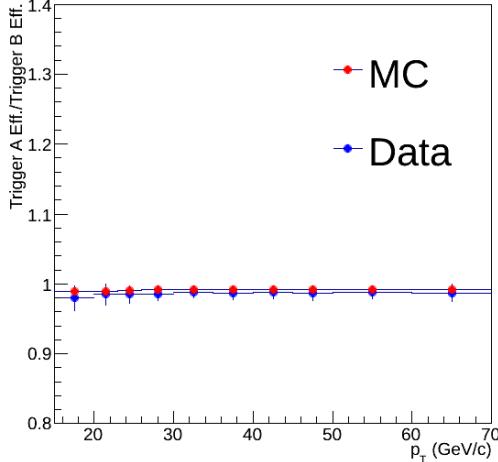


Figure 36: Ratio of HLT\_L2DoubleMu23\_NoVertex\_2Cha\_Angle2p5 trigger efficiency to HLT\_L2DoubleMu23\_NoVertex trigger efficiency vs.  $p_T$  for both data and  $Z \rightarrow \mu\mu$  simulation.

Figure 37 shows that Trigger A efficiency is highly dependent on the lifetime,  $c\tau$ , of the decays in the effective range of the analysis. Note that the plots are obtained with the lifetime reweighting for the two signal MC samples with  $M_{H^0} = 1000 \text{ GeV}/c^2$  and  $M_X = 350 \text{ GeV}/c^2$ ;  $M_{H^0} = 400 \text{ GeV}/c^2$  and  $M_X = 20 \text{ GeV}/c^2$ .

The results discussed in Section 6.3 show that the agreement between RSA muon reconstruction and selection efficiencies in data and simulation is approximately independent of  $|d_0|$ . Since the algorithms used in the trigger muon reconstruction are similar to those used in the offline muon reconstruction, it is also reasonable to expect that the agreement between the trigger efficiencies does not strongly depend on  $|d_0|$ . Nevertheless, because we do not directly measure the trigger efficiency as a function of the decay length, we assign an additional systematic uncertainty by assuming that the difference in the trigger efficiency between data and simulation increases linearly as a function of the transverse decay length. The dimuon trigger efficiency to select the signal events is parameterized by a simple function of the transverse decay length,  $L_{xy}$ :

$$F(L_{xy}) = 0.9 \pm \frac{\beta L_{xy}}{250} \quad (1)$$

indicating that the simulated trigger efficiency should be reduced by 10% for prompt events to overlap that of the data and that the additional variation by a factor of  $\beta$  is added.  $\beta$  is set to be 10% such that the trigger efficiency is varied by another 10% at 2.5 meters away from the interaction point, corresponding to the assumption that the discrepancy between data and simulation at that value is twice as big as for prompt decays.

To conclude, the largest deviation in the signal selection efficiency is found to be not greater than 17%, which we take as the relative systematic uncertainty on the trigger efficiency measurement in this analysis.

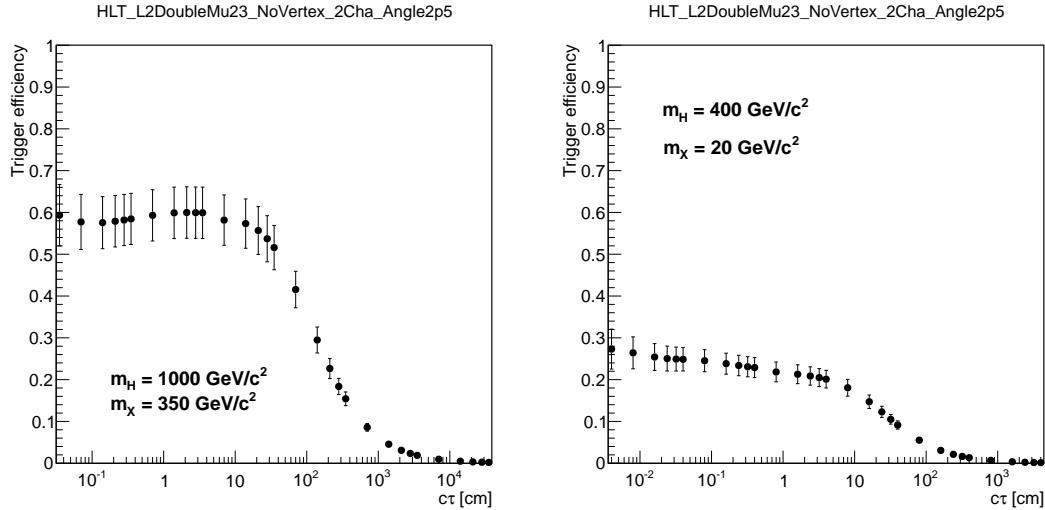


Figure 37: Trigger efficiency of `HLT_L2DoubleMu23_NoVertex_2Cha_Angle2p5` vs. generated decay length,  $c\tau$ , for the signal simulated events with  $M_{H^0} = 1000 \text{ GeV}/c^2$  and  $M_X = 350 \text{ GeV}/c^2$ ;  $M_{H^0} = 400 \text{ GeV}/c^2$  and  $M_X = 20 \text{ GeV}/c^2$ . Lifetime reweighting is applied to cover the full  $c\tau$  range. The trigger efficiency starts to drop when  $c\tau > 10 \text{ cm}$ .

## 550 6.5 Effect of the $p_T$ resolution

551 We have studied the effect of the modest RSA muon  $p_T$  resolution on this analysis and conclude  
 552 that no systematic uncertainty needs to be assigned.

## 553 7 Analysis sensitivity and expected upper limits

554 The upper limits on the signal production mechanism based on a particular model with various  
 555 mass points and lifetime values are set with 95% confidence level (CL). The computation is  
 556 performed via the statistics software package developed by the CMS Higgs Group [20], which  
 557 applies the Bayesian method established in Ref. [21]. The limits are derived by comparing the  
 558 number of events  $N_S$  expected in the signal region with the number of events that the signal  
 559 plus background hypothesis predicts.

560 The systematic uncertainties on the signal selection efficiency given in Section 6 are introduced  
 561 in the limit calculation as nuisance parameters with log-normal prior distributions. The ex-  
 562 pected number of background events in the signal region,  $\mu_B$ , is taken as an additional nu-  
 563 iance parameter, which depends on the number of observed events in the control region,  $N_C$ ,  
 564 and, therefore, in the signal region. The probability distribution of  $\mu_B$ ,  $p(\mu_B)$ , is given by  
 565  $p(\mu_B) = \frac{\mu_B^{N_C}}{N_C!} \exp(-\mu_B)$ , as can be shown using Bayesian method assuming a flat prior in  $\mu_B$   
 566 [21].

567 The expected number of signal events,  $\mu_S$ , takes the following form:

$$\begin{aligned}\mu_S &= \mathcal{L}\sigma [2B(1-B)\epsilon_1 + \epsilon_2 B^2] (1-f) \\ &= 2\mathcal{L}\sigma B\epsilon_1 \left[ 1 - B \left( 1 - \frac{\epsilon_2}{2\epsilon_1} \right) \right] (1-f)\end{aligned}\quad (2)$$

568 where  $\mathcal{L}$  is the integrated luminosity,  $\epsilon_{(1,2)}$  are the signal efficiencies defined in Section 4.5,  $\sigma$   
 569 is the production cross section of  $H^0 \rightarrow XX$  and  $B$  is the branching fraction for the decay  $X \rightarrow$

$\ell^+\ell^-$  where each lepton refers to a muon. The parameter  $f$  represents the ratio of the number of signal events falling into the control region as fake background to the number of signal events in the signal region. Although the effect is negligible for all signal samples used, the conservative value, 5%, is set for this parameter. If the efficiencies,  $\epsilon_1$  and  $\epsilon_2$ , are independent of each other, it can be shown that  $\epsilon_2 = 1 - (1 - \epsilon_1)^2$ . However, since the triggering of one LL decay can give rise to the triggering of two LL decays, the two efficiencies are correlated, that is,  $\epsilon_2 \geq 1 - (1 - \epsilon_1)^2$ . To calculate the upper limits conservatively, the value of the expected number of signal events,  $\mu_S$ , can be minimized with  $\epsilon_2 = 1 - (1 - \epsilon_1)^2$  in the equation (2),

$$\mu_S = 2\mathcal{L}\sigma B\epsilon_1 [1 - B\epsilon_1/2] (1 - f) \quad (3)$$

In equation (3), the upper bounds on  $\sigma B$  depend on the branching ratio. Hence, the limits are derived for two extreme cases,  $B\epsilon_1 \ll 1$  and  $B = 1$ .

The 95% CL upper limits are calculated for all mass points of  $H^0 \rightarrow XX$  signal samples, listed in Table 2, as a function of X boson lifetime. The expected limits are illustrated in the plots in Figure 38. The analysis is least sensitive to the  $M_{H^0} = 125 \text{ GeV}/c^2$  case due to the low signal selection efficiencies in particular when  $M_X = 20 \text{ GeV}/c^2$ . The corresponding limits on  $\sigma(\tilde{q}\tilde{q} + \tilde{q}\tilde{q})B(\tilde{q} \rightarrow q\tilde{\chi}^0, \tilde{\chi}^0 \rightarrow \ell^+\ell^-\nu)$  are shown in Fig. 39. The limits improve as the decay lifetime increases, as expected, since the analysis has a negligible sensitivity for transverse decay lengths less than 40 cm. The green bands in these limit plots represent the  $\pm 1\sigma$  range of variation of the expected 95% CL limits.

The same lifetime reweighting procedure is applied as in Ref. [6] to obtain the upper limits on  $\sigma B$ . The signal efficiency for a given lifetime is estimated by reweighting the lifetime distribution of the sample that has the closest generated lifetime value to the one for which the estimation is performed. If the uncertainty in the signal efficiency being recomputed is greater than 40%, the upper bound for that lifetime is discarded.

## 7.1 Comparison with the tracker-based analysis

In this section we compare the results of this analysis with those of the tracker-based analysis [6].

The signal efficiencies that the two analyses yield for the largest lifetime samples of  $H^0 \rightarrow XX$  simulated events are tabulated in Table 9. In addition, the signal efficiency ratio of the two analyses for various  $H^0$  and X masses as a function of  $c\tau$  is shown in Figure 40. For signal events with smaller  $c\tau$ , the ratio is near zero because this analysis is not sensitive in that range by design. For signal events with larger  $c\tau$ , the efficiency of this analysis is similar to that of the tracker-based analysis.

## 7.2 Combined Limits

The analysis described in this document and the one based on the silicon tracker are orthogonal in the sense that there is no overlap in the events selected by the two analyses. The results are combined to yield the best possible exclusion limits. To produce the combination, we use the same framework that was used to compute the exclusion limits. The two analyses are treated as independent measurement channels of the same physics signal.

All the systematic uncertainties of the two channels are uncorrelated except for the integrated luminosity (fully correlated) and the trigger efficiency (partially correlated). For the trigger efficiency correlation we take the most conservative assumption when computing the combined

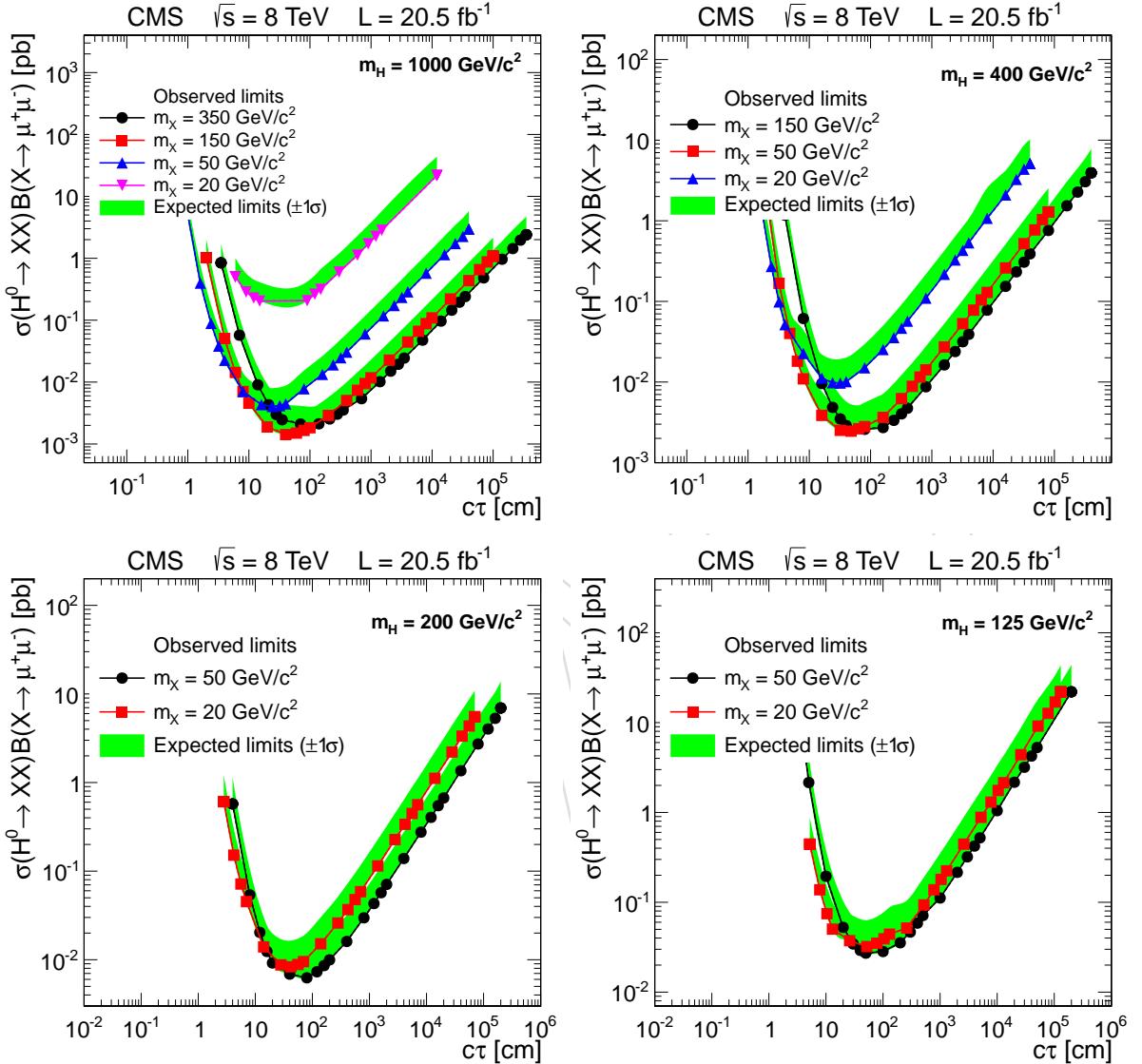


Figure 38: 95% CL upper limits on  $\sigma(H^0 \rightarrow XX)B(X \rightarrow \ell^+\ell^-)$  for  $M_{H^0} = 1000 \text{ GeV}/c^2$ ,  $400 \text{ GeV}/c^2$ ,  $200 \text{ GeV}/c^2$  and  $125 \text{ GeV}/c^2$  with various  $X$  mass points. The limits derived for  $B\epsilon_1 \ll 1$  are illustrated by the solid curves, whereas the dotted curves represent those for  $B = 1$  (the dotted curves are difficult to discern due to the overlap with the solid curves). Green shaded bands show the  $\pm 1\sigma$  range of variation of the expected 95% CL limits.

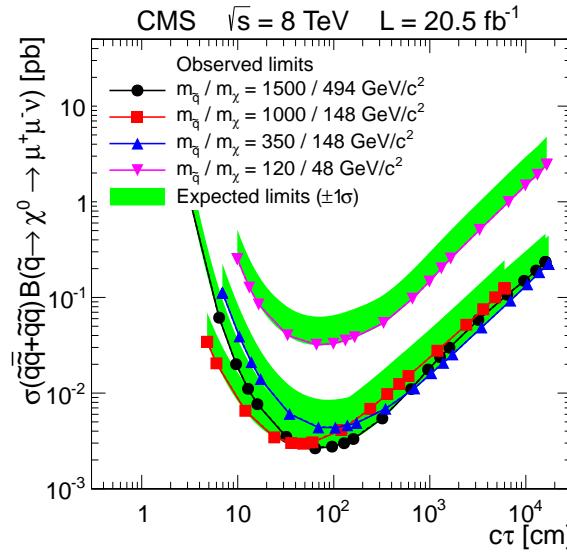


Figure 39: The 95% CL upper limits on  $\sigma(\tilde{q}\bar{q} + \tilde{q}\bar{q})B(\tilde{q} \rightarrow q\tilde{\chi}^0, \tilde{\chi}^0 \rightarrow \ell^+\ell^-\nu)$  for the muon channel as a function of the neutralino lifetime. The limits derived for  $B\epsilon_1 \ll 1$  are illustrated by the solid curves, whereas the dotted curves represent those for  $B = 1$  (the dotted curves are difficult to discern due to the overlap with the solid curves). Green shaded bands show the  $\pm 1\sigma$  range of variation of the expected 95% CL limits.

Table 9: Efficiencies of the signal MC samples to satisfy the full selection for the two analyses. The efficiencies for events in which only one simulated LL particle decays to the muons,  $\epsilon_1$ , and for the ones where there are two generated LL particles decaying to the muon pairs,  $\epsilon_2$ , are shown separately.

$M_{H^0}$ ( $\text{GeV}/c^2$ )	$M_X$ ( $\text{GeV}/c^2$ )	$c\tau$ (cm)	Muon Chambers		Silicon Tracker	
			$\epsilon_1$	$\epsilon_2$	$\epsilon_1$	$\epsilon_2$
1000	350	350	0.023	0.049	0.03	0.058
1000	150	100	0.044	0.073	0.05	0.1
1000	50	40	0.018	0.03	0.043	0.093
1000	20	15	0.00043	0.00096	0.0035	0.009
400	150	400	0.017	0.038	0.018	0.039
400	50	80	0.029	0.053	0.036	0.082
400	20	40	0.0079	0.014	0.023	0.056
200	50	200	0.0081	0.02	0.0094	0.026
200	20	70	0.0085	0.015	0.016	0.039
125	50	500	0.0012	0.0033	0.0013	0.0026
125	20	130	0.0018	0.0047	0.0021	0.0054

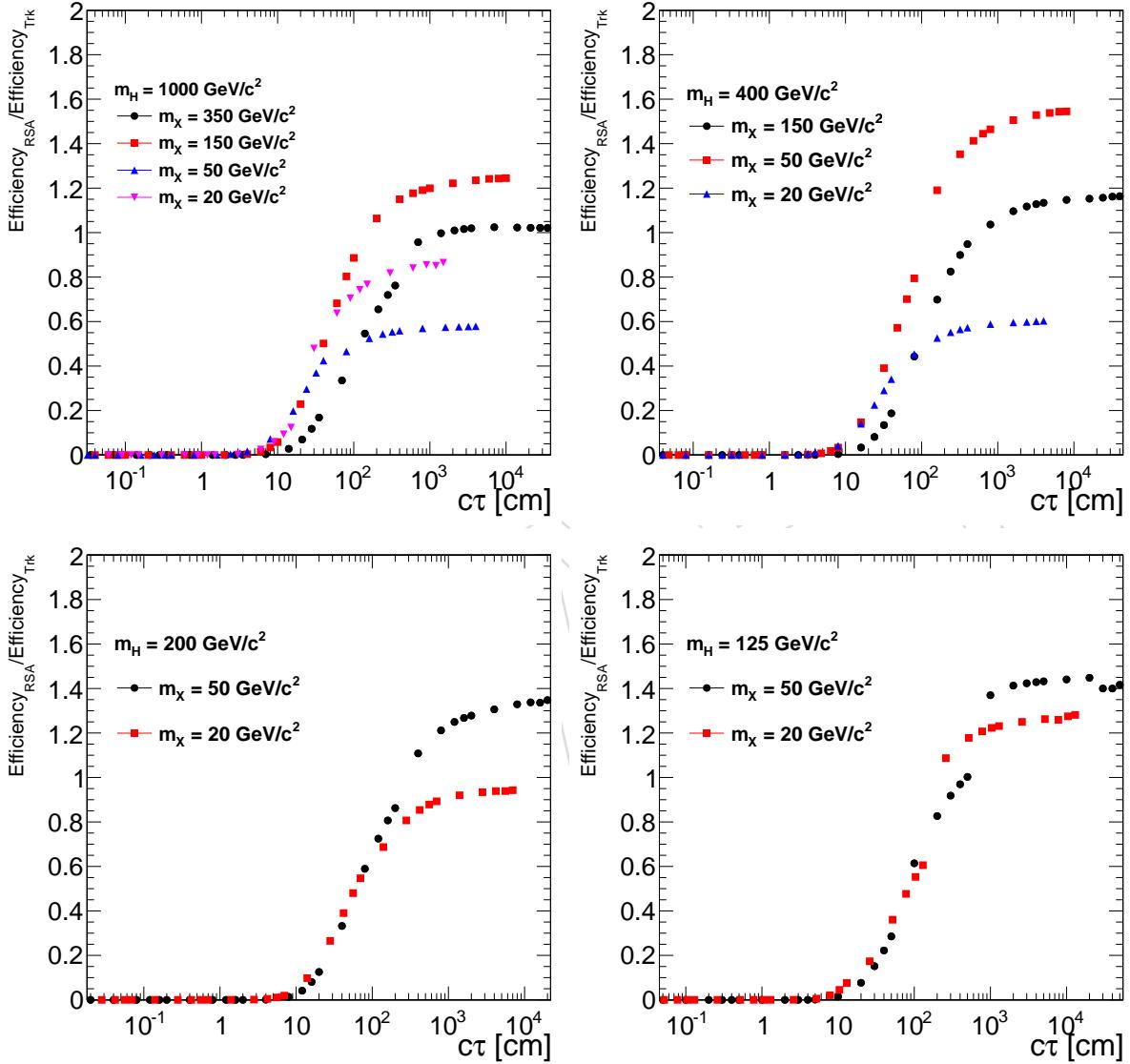


Figure 40: Signal efficiency ratio of the two analyses  $\epsilon_{RSA}/\epsilon_{trk}$  for  $H^0$  mass values of  $1000 \text{ GeV}/c^2$ ,  $400 \text{ GeV}/c^2$ ,  $200 \text{ GeV}/c^2$  and  $125 \text{ GeV}/c^2$  with various  $X$  mass points.

expected limits. The tracker tracking efficiency systematic uncertainty is also partially correlated because of the tracker track rejection cut applied in the muon chambers-based analysis. However, if this cut is removed, there is at most 2% overlap in the set of events satisfying the selection criteria of both analyses, meaning that only about 2% of the systematics is correlated. We consider this effect to be negligible and we assume no correlation for the systematic uncertainty on the tracker tracking efficiency. We consider this effect to be negligible and we assume no correlation for the tracker tracking efficiency systematic uncertainty.

The results of the combination are shown in Figure 41 for the  $H^0 \rightarrow XX$  and in Figure 42 for the  $\tilde{q} \rightarrow q\tilde{\chi}^0, \tilde{\chi}^0 \rightarrow \ell^+\ell^-\nu$ . Figure 43 shows the comparison of the upper limits for the tracker-based analysis, the RSA-based analysis, and the combination of the two for the  $H^0 \rightarrow XX$  and the  $\tilde{q} \rightarrow q\tilde{\chi}^0, \tilde{\chi}^0 \rightarrow \ell^+\ell^-\nu$  for two example mass points. The comparisons are similar for the other mass points. The limits coincide with the ones from the tracker-based analysis for lower lifetime values, where the tracker-based analysis dominates the efficiency. For higher lifetime values the exclusions are significantly improved by the combination, up to a factor of two.

## 8 Summary

A search for long-lived particles decaying to dimuon final states using only the muon chambers has been performed on pp collision data taken by the CMS detector at  $\sqrt{s} = 8$  TeV in 2012. It extends the scope of a similar search for displaced dimuon signatures based on the silicon tracker. No background events are expected to pass the selection criteria of the analysis. Expected upper limits are computed for a model predicting a heavy scalar, with mass in the range  $125 - 1000 \text{ GeV}/c^2$ , decaying to pairs of long-lived neutral particles, with masses in the range  $20 - 350 \text{ GeV}/c^2$ , which decay to dimuon pairs. The limits are typically in the range  $1 - 50 \text{ fb}$ , and can weaken to a few pb for the lowest masses and longest lifetimes, and are given for lifetimes in the range  $1 < c\tau < 1000 \text{ cm}$ . The expected upper limits are comparable to, and in some cases improve on, those set by the tracker-based analysis, for longer lifetimes. Given that the two analyses are fully orthogonal, combined upper limits are presented which provide the most stringent limits for this kind of search in the dimuon channel so far.

## A Muon timing information in the events removed by cosmic rejection

As an additional source of background in our analysis, we might have events in which at least one cosmic muon may enter and make fake dimuon combinations, which might be so highly displaced that they pass the lifetime selection, with other muons emanating from the pp collisions. Hence, the cosmic rejection cut, described in Section 4.3, is placed to avoid such cosmic contamination.

If a muon originating from cosmics hits the detector components randomly, it would most likely be unsynchronized with the time of the collisions. Even though the muon can still be reconstructed, the fit used to extract the timing information could fail. To examine the events vetoed by the cosmic rejection in detail, we check the timing information of the muons in those events. If a muon has the timing information reconstructed, it is labeled as in-time muon and out-of-time muon, if not.

It is checked that the efficiency of reconstructing muon timing is slightly above 99% for the prompt data and the background simulation events. The same efficiency is around 93% for the signal simulation events. In other words, if a muon emerges from the collision, then, it

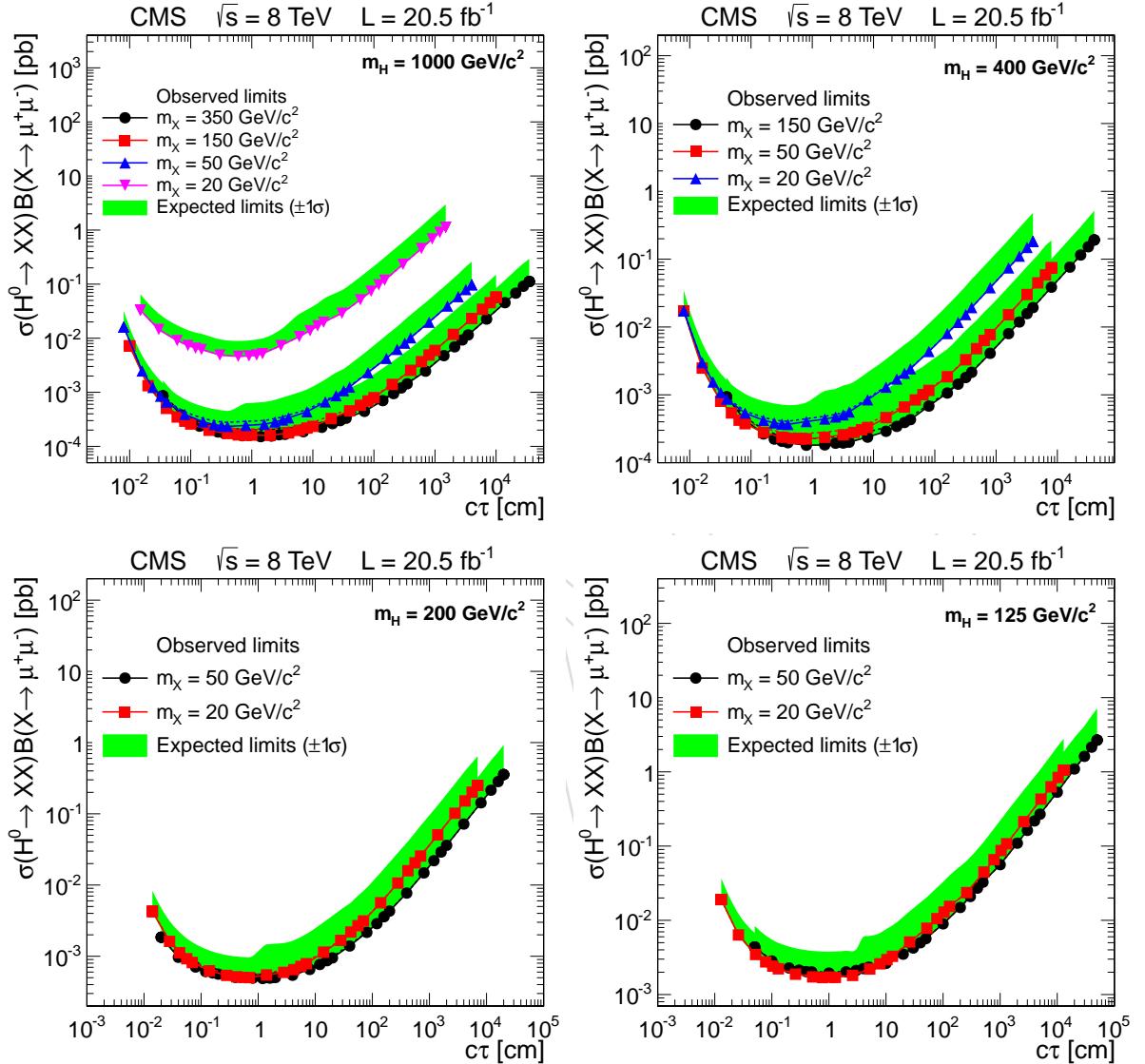


Figure 41: Combined 95% CL upper limits on  $\sigma(H^0 \rightarrow XX)B(X \rightarrow \ell^+\ell^-)$  for all  $H^0$  mass values of  $1000 \text{ GeV}/c^2$ ,  $400 \text{ GeV}/c^2$ ,  $200 \text{ GeV}/c^2$  and  $125 \text{ GeV}/c^2$  with various  $X$  mass points. The limits derived for  $B\epsilon_1 \ll 1$  are illustrated by the solid curves, whereas the dotted curves represent those for  $B = 1$  (the dotted curves are difficult to discern due to the overlap with the solid curves). Green shaded bands show the  $\pm 1\sigma$  range of variation of the expected 95% CL limits.

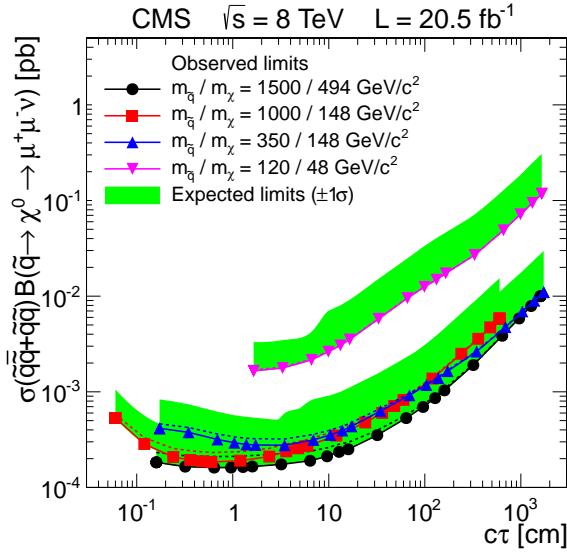


Figure 42: Combined 95% CL upper limits on  $\sigma(\tilde{q}\bar{q} + \tilde{q}\bar{q})B(\tilde{q} \rightarrow q\tilde{\chi}^0, \tilde{\chi}^0 \rightarrow \ell^+\ell^-\nu)$  for the muon channel as a function of the neutralino lifetime. The limits derived for  $B\epsilon_1 \ll 1$  are illustrated by the solid curves, whereas the dotted curves represent those for  $B = 1$  (the dotted curves are difficult to discern due to the overlap with the solid curves). Green shaded bands show the  $\pm 1\sigma$  range of variation of the expected 95% CL limits.

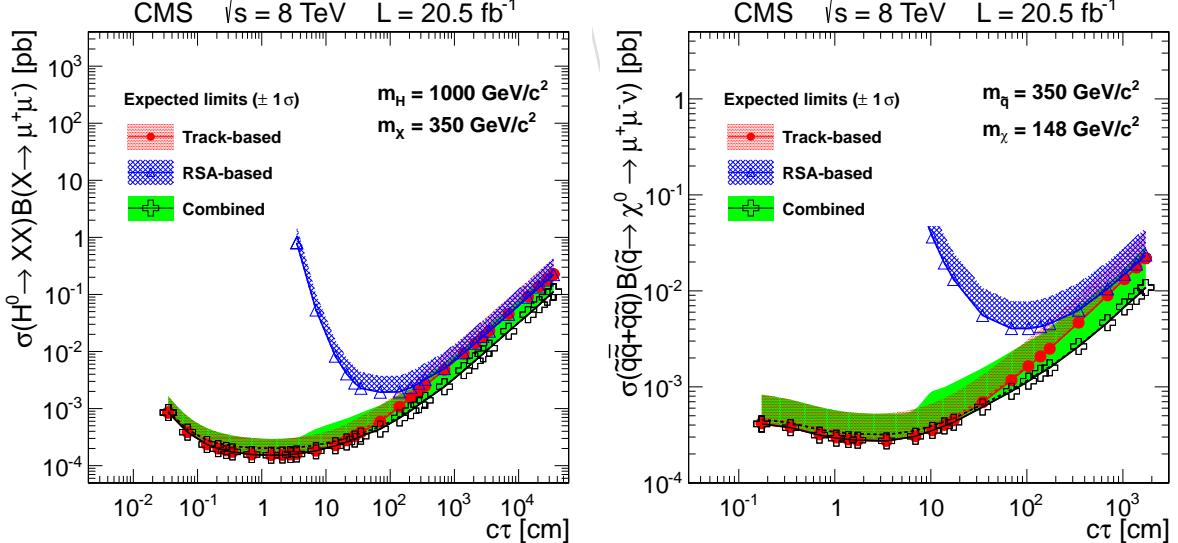


Figure 43: Comparison of the upper limit for the tracker-based analysis, the RSA-based analysis and the combination of the two for the  $H^0 \rightarrow XX$  for  $M_{H^0} = 1000 \text{ GeV}/c^2$  and  $M_X = 350 \text{ GeV}/c^2$  (left) and for the  $\tilde{q} \rightarrow q\tilde{\chi}^0, \tilde{\chi}^0 \rightarrow \ell^+\ell^-\nu$  for  $M_{\tilde{q}} = 350 \text{ GeV}/c^2$  and  $M_\chi = 148 \text{ GeV}/c^2$ .

654 is an in-time muon with great efficiency. On the other hand, if the events removed by the  
 655 cosmic rejection are carefully investigated, more than  $\approx 85\%$  of the cases, we find at least one  
 656 out-of-time muon in these events. Therefore, no decisive disagreement with the cosmic muon  
 657 interference hypothesis is shown.

658 Another important finding is that there is a correlation between the muon timing reconstruction  
 659 and the number of muon hits. To confirm it, we examine the valid muon hit distributions of  
 660 in-time and out-of-time muons in the events removed by the cosmic rejection cut. Figure 44  
 661 demonstrates the clear separation in the number of valid muon hits between the two. Thus,  
 662 one could reasonably argue that out-of-time muons tend to have lower number of hits in the  
 663 muon chambers and the timing reconstruction fails for the same reason.

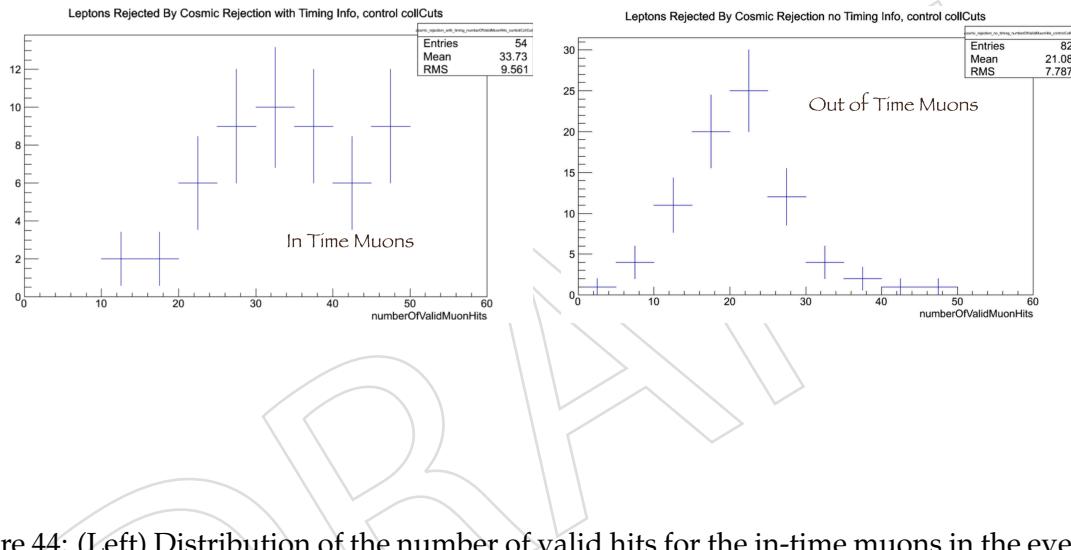


Figure 44: (Left) Distribution of the number of valid hits for the in-time muons in the events rejected by the cosmic rejection cut. (Right) Same distribution for the out-of-time muons. There is an obvious shift between the two distributions, suggesting that out-of-time muons, in general, have lower number of hits in the muon chambers.

664 In Figure 45, the distribution of the minimum number of valid hits of the muons passing the  
 665 full selection is shown for one of the signal samples. After checking the same distribution in all  
 666 signal simulated samples, it is concluded that if one requires at least 17 valid hits for each muon,  
 667 the signal efficiencies would not be significantly affected, whereas the level of background in  
 668 data would be slightly reduced. Therefore, this cut is included in the analysis selection as a  
 669 further protection against the cosmic contamination.

## 670 **B Muon trigger efficiency**

671 The systematic uncertainty assigned to the trigger efficiency to select the events analyzed is  
 672 simply the discrepancy associated with that measurement between data and the simulation.  
 673 The trigger efficiency is measured via the *Tag and Probe* method which can be exploited pro-  
 674 vided that there is a mass resonance, such as Z boson, decaying to muon pairs.

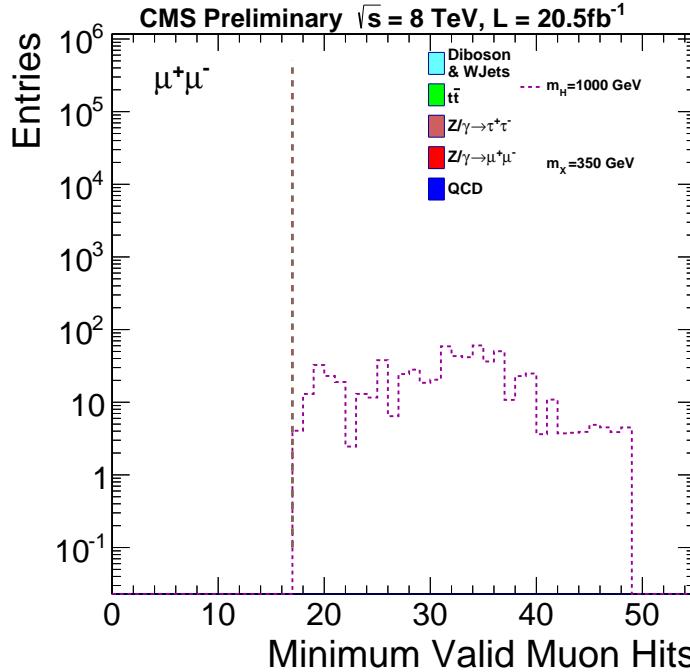


Figure 45: Distribution of the minimum valid muon hits of the two muons for the dimuon candidates passing the full selection except the cut on the minimum valid muon hits. This plot is obtained for the signal sample:  $M_{H^0} = 1000 \text{ GeV}/c^2$  and  $M_X = 350 \text{ GeV}/c^2$  with  $c\tau = 350 \text{ cm}$

675 One of the muons, which comes from Z boson's decay, is labeled as *tag* that survives the tight  
 676 selection criteria that ensure that it is very unlikely to be fake. The other muon, which is as-  
 677 sumed to have no correlation with the tag muon, is called *probe*. A passing probe, satisfying  
 678 selection cuts, is required to be matched to the trigger that selects the events for this analysis.  
 679 Finally, one fits the tag-probe mass distribution to extract the number of Z candidates for failing  
 680 and passing probes.

681 The analysis selects events the collected by the trigger, `HLT_L2DoubleMu23_NoVertex_2Cha_-`  
 682 `Angle2p5`. This dimuon trigger requires:

- 683 • Two L2 muons with  $p_T > 23 \text{ GeV}/c$  reconstructed with no vertex constraint.
- 684 • Each muon must have at least two DT or CSC muon stations with any hits.
- 685 • The three dimensional angle between the two muons must be larger than 2.5 radians  
 686 (cosine  $> \sim -0.8$ ).

687 Because of the angle requirement in the trigger, the collected dataset is unsuitable to apply the  
 688 *Tag and Probe* method directly. Instead, we factor the trigger efficiency measurement into two  
 689 parts. First, we use the events that are collected by another double muon trigger, `HLT_L2Dou-`  
 690 `bleMu23_NoVertex` for the measurement. This trigger is identical to `HLT_L2DoubleMu23_-`  
 691 `NoVertex_2Cha_Angle2p5` except for the lack of the angle cut and of the requirement of a  
 692 minimum number of muon stations with hits. Secondly, since the actual dimuon trigger is  
 693 equivalent to `HLT_L2DoubleMu23_NoVertex` plus two additional cuts, we measure the effi-  
 694 ciency of these additional cuts separately to get the overall efficiency of `HLT_L2DoubleMu23_-`  
 695 `NoVertex_2Cha_Angle2p5` used in this search. Note that `HLT_L2DoubleMu23_NoVertex` is  
 696 prescaled by a factor of 20, which is accounted for the measurement.

697 The MC dataset used for this measurement is /DYJetsToLL\_M-50\_TuneZ2Star\_8TeV-madg-  
 698 raph-tarball/Summer12\_DR53XPU\_S10\_START53\_V7Av1/AODSIM and all CMS datasets  
 699 in this study are processed as 22Jan ReReco. In the part in which HLT\_L2DoubleMu23\_-  
 700 NoVertex trigger efficiency is measured, SingleMu CMS dataset is used. DoubleMu CMS  
 701 dataset is used when the ratio of the two trigger efficiencies is computed. The official *Tag and*  
 702 *Probe* package is exploited to measure the trigger efficiency of HLT\_L2DoubleMu23\_NoVertex.  
 703 The package is run under CMSSW\_5\_3\_7.

704 To measure the HLT\_L2DoubleMu23\_NoVertex trigger efficiency, the tag is chosen from the  
 705 global muon collection matched to *IsoMu24* single muon trigger object within  $\Delta R < 0.1$ . Then,  
 706 the following selection cuts are applied for the tag muon:

- 707 •  $p_T > 26 \text{ GeV}/c$  and  $|\eta| < 2$
- 708 • Relative isolation ( $\text{isolationR}03.\text{sumPt})/pt < 0.1$
- 709 • Number of tracker layers with measurement  $\geq 6$
- 710 •  $|d_{xy}| < 30 \text{ cm}$  and  $|d_z| < 30 \text{ cm}$

711 The probe muons are required to be RSA muons used in the analysis. The conditions to be met  
 712 for the probe muons are:

- 713 •  $p_T > 17 \text{ GeV}/c$  and  $|\eta| < 2.4$
- 714 • Number of DT + CSC muon stations with valid hits  $\geq 2$
- 715 • The probe is a passing probe if it is matched to the dimuon trigger, HLT\_L2Dou-  
 716 bleMu23\_NoVertex, object within  $\Delta R < 0.5$

717 In addition, two generic criteria are applied for the pair selection between the tag and the probe:

- 718 •  $\Delta R > 0.2$
- 719 • Cosine of the angle  $> -0.79$

720 The efficiency for this measurement is input as an unknown parameter to the simultaneous  
 721 fitting. Once the fitting is done, the parameter is simply extracted and plotted. The trigger  
 722 efficiency vs.  $p_T$  graph for the dimuon trigger, HLT\_L2DoubleMu23\_NoVertex, is shown in  
 723 Figure 46. It can be deduced from the same graph that data-MC discrepancy is less than 5%.  
 724 Additionally, in order to check the validity of the efficiency result from the simultaneous fitting,  
 725 a comparison with the efficiency results from counting method is performed for the simulation.  
 726 Figure 47 shows that the fit results agree with those from counting.

727 To demonstrate the accuracy of the fitting, one may look at Figure 48 and Figure 49 in which  
 728 the tag-probe dimuon mass distributions are shown for each  $p_T$  bin of the probe muon. Note  
 729 that green line represents the dimuon mass distribution of the tags and the passing probes and  
 730 blue line represents that of tags and all probes. The red curve is for the same distribution with  
 731 the failing probes. It can be said that even though the fitting is not perfectly accurate, it is  
 732 reasonable. Given that the offline muon  $p_T$  threshold is  $26 \text{ GeV}/c$  for the analysis, one should  
 733 not consider the bins below that value for the accuracy check. Another important feature of  
 734 the RSA muons seen from these figures is that Z mass peak location is different for each bin. It  
 735 is mainly due to the low  $p_T$  resolution of RSA muons. Besides, the fact that we have binning  
 736 in  $p_T$  makes the shift more obvious to discern. One may also suspect that some of the passing  
 737 probes are fake. However, an additional study is performed with the MC sample from which  
 738 MC *Truth* information is collected determining if the muons actually come from Z boson decay.  
 739 In Figure 50, red, blue and violet dots are the results from fitting, counting and MC *Truth* fitting.  
 740 The plot illustrates that MC *Truth* fitting results are almost identical to the results in matching

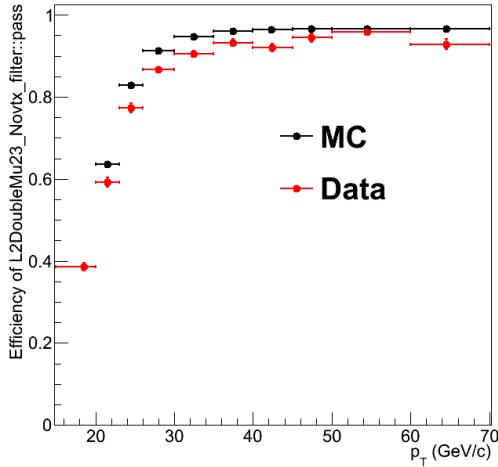


Figure 46: Trigger efficiency of `HLT_L2DoubleMu23_NoVertex` vs.  $p_T$  of the probe muon for both data and  $Z \rightarrow \mu\mu$  simulation, obtained with the *Tag and Probe* method.

- 741 not-required case. The fitting and counting results are in agreement with each other as well.
- 742 `HLT_L2DoubleMu23_NoVertex` trigger efficiency as a function of pseudorapidity,  $|\eta|$ , is also  
743 derived. Figure 51 shows that the trigger efficiency is highest in the barrel region, which is  
744 expected. In addition, in order to have an idea of how the efficiency varies for data and the  
745 simulation in two dimensional plane of  $p_T$  and  $|\eta|$ , the plots in Figure 52 and Figure 53 are  
746 prepared for a further inspection.
- 747 The efficiency of the additional cuts included in `HLT_L2DoubleMu23_NoVertex_2Cha_Angle2p5`  
748 relative to `HLT_L2DoubleMu23_NoVertex` is measured for the RSA muons. This efficiency is  
749 defined as the ratio of the number of events triggered by `HLT_L2DoubleMu23_NoVertex_-`  
750 `2Cha_Angle2p5`, given that `HLT_L2DoubleMu23_NoVertex` is fired, to the number of events  
751 triggered by `HLT_L2DoubleMu23_NoVertex` only.
- 752 While keeping the selection requirements on RSA muons same as described previously, the  
753 simultaneous fit is implemented to extract the ratio of the two trigger efficiencies as a function  
754 of the  $p_T$  of one of the RSA, probe, muons in the event. In Figure 54, the ratio of the *new* trigger,  
755 `HLT_L2DoubleMu23_NoVertex_2Cha_Angle2p5`, efficiency to the *old* trigger, `HLT_L2Dou-`  
756 `bleMu23_NoVertex`, efficiency vs.  $p_T$  plot is presented. Note that the terms, *new* and *old*, are  
757 used for simplicity. As it can be inferred from the plot, the efficiency of the additional cuts is  
758 around 98% and the simulation-data discrepancy is less than 1%. It is also advised to look at  
759 both Figure 55 and Figure 56 that justify that the fits are accurate enough for which Crystal Ball  
760 function is employed.
- 761 The efficiency to trigger two muons equals the square of `HLT_L2DoubleMu23_NoVertex` effi-  
762 ciency multiplied by the ratio of `HLT_L2DoubleMu23_NoVertex_2Cha_Angle2p5` efficiency  
763 to `HLT_L2DoubleMu23_NoVertex` efficiency. The discrepancy between data and simulation is  
764 no larger than 10%, which is taken as systematic uncertainty on the trigger efficiency measured  
765 for  $Z$  decays.
- 766 An accompanying cross check performed is to test if the cosine cut,  $\cos \theta > -0.79$ , applied to  
767 the muon pairs is reasonable. If it is correct, then one may expect to have a jump in `HLT_-`  
768 `L2DoubleMu23_NoVertex_2Cha_Angle2p5` trigger efficiency around this value. Figure 57

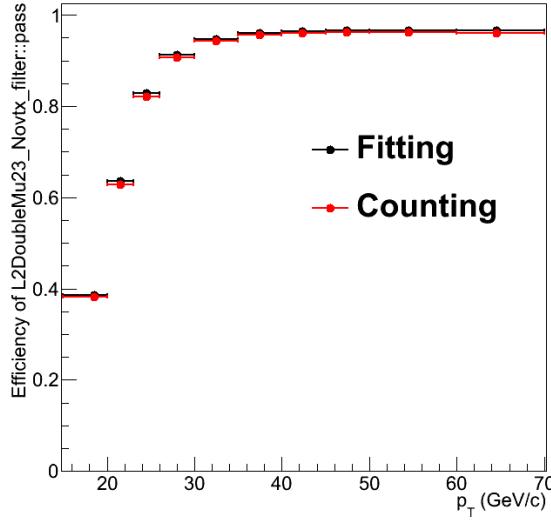
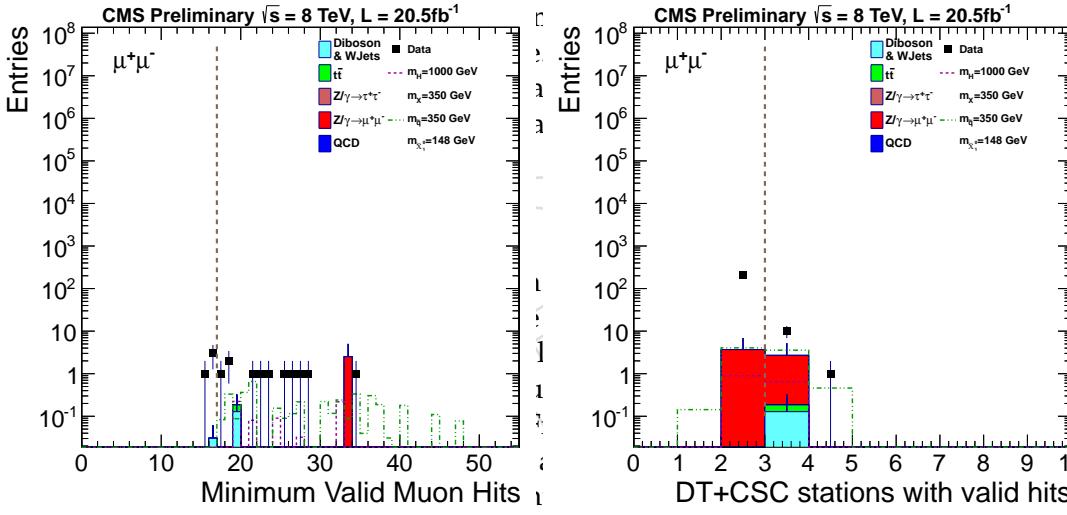


Figure 47: The plot compares HLT\_L2DoubleMu23\_NoVertex trigger efficiency results from the fitting, already presented in Figure 35 and with those from the counting method for the MC sample.



tought to check if in the muon pairs peated due to the

## ulation

ccount for the ex-  
e background MC  
d by applying the  
< rejection and the  
selection criteria (

hey show that the

Some discrepan-

Figure 66: (Left) Distribution of the minimum number of valid muon hits of the two muons  
<sup>781</sup> for the dimuon candidates passing the selection. (Right) Distribution of the minimum number  
<sup>782</sup> of valid muon stations of the two muons for the dimuon candidates passing the selection but  
<sup>783</sup> also different than that of data. However, the detailed study presented in Appendix D shows  
<sup>784</sup> that the analysis is significantly insensitive to the  $p_T$  resolution of the RSA muons.  
<sup>785</sup> with loose cuts at  $|u_0|/v_d \geq 1$  and  $\Sigma_{xy}/\sigma_{L_{xy}} \geq 3$  in the control region.

Therefore, even though there are minor discrepancies between data and the simulation, they do not play an effective role within the scope of the analysis. We utilize the simulation only to compute the efficiencies of the signal samples. The comparison to the background MC is only used to assess whether the simulation is modelling correctly the variable distributions. We estimate the expected background directly from the data. Note also that this check is employed without unblinding the analysis. That is, only the agreement in the control region is examined.

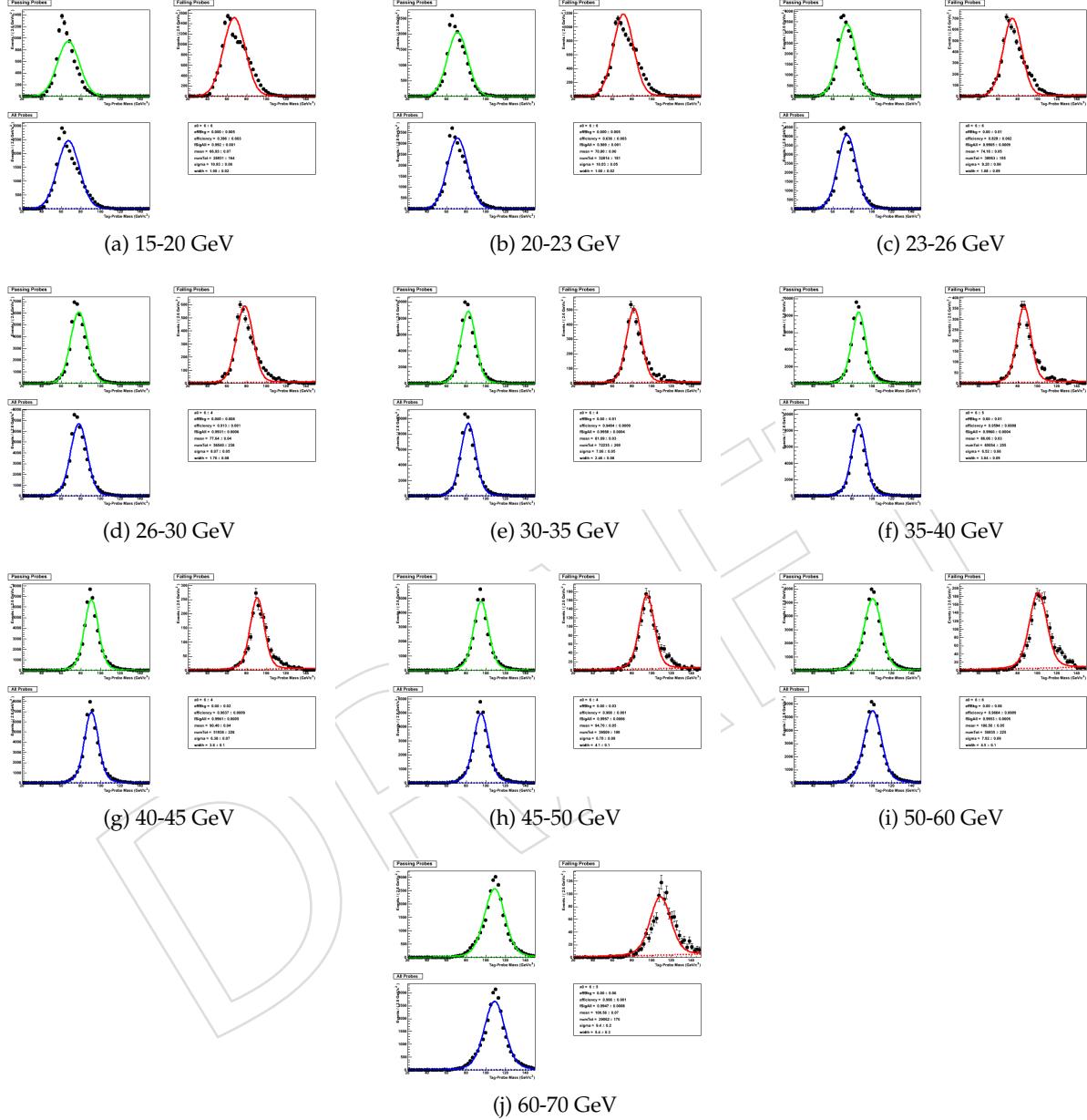


Figure 48: MC fitting results of tag-probe dimuon mass distributions for each  $p_T$  bin with *Tag and Probe* applied for the trigger, `HLT_L2DoubleMu23_NoVertex`. Note that the range of each  $p_T$  bin is indicated under each plot.

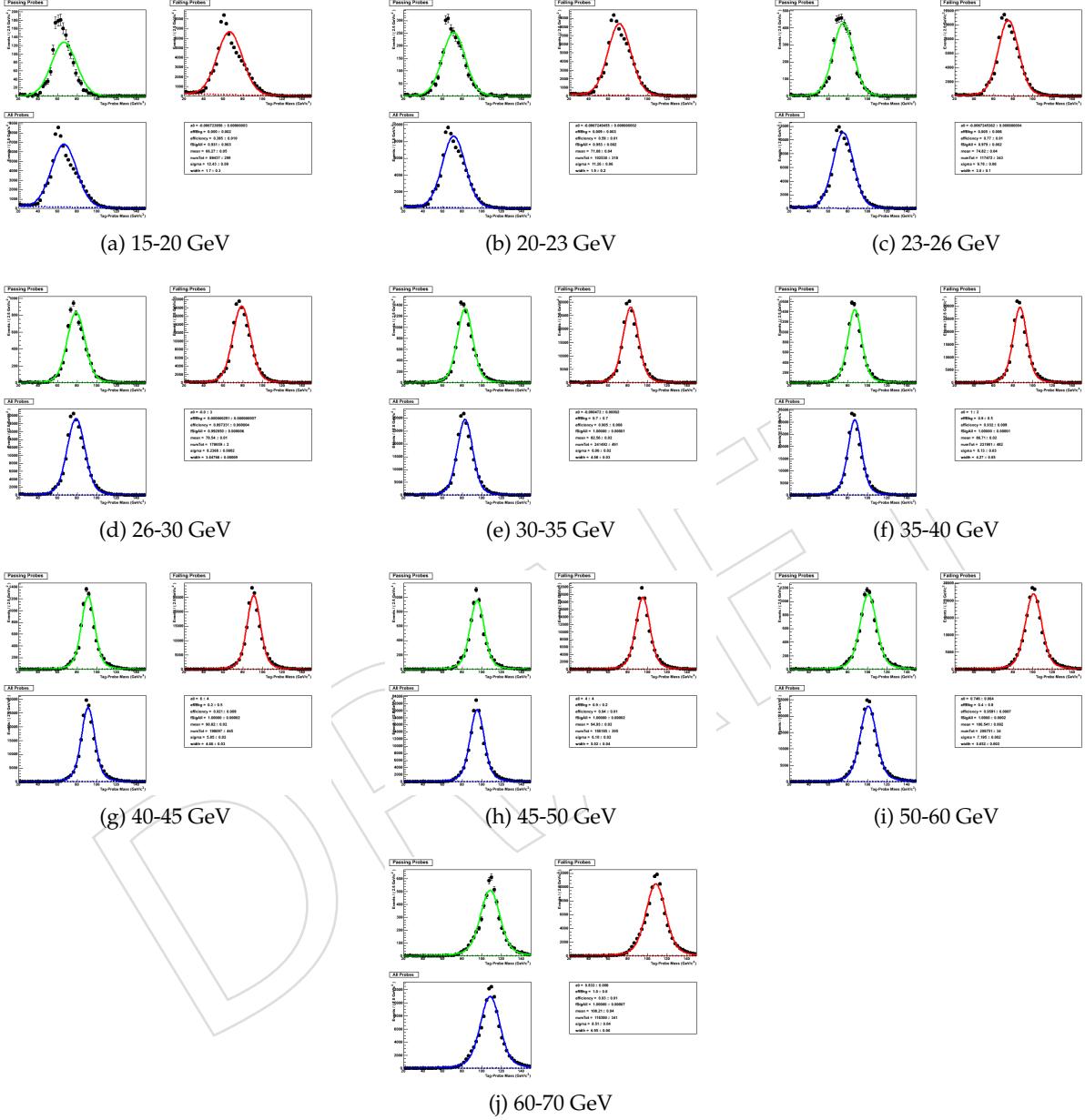


Figure 49: Data fitting results of tag-probe dimuon mass distributions for each  $p_T$  bin with *Tag* and *Probe* applied for the trigger, `HLT_L2DoubleMu23_NoVertex`. Note that the range of each  $p_T$  bin is indicated under each plot.

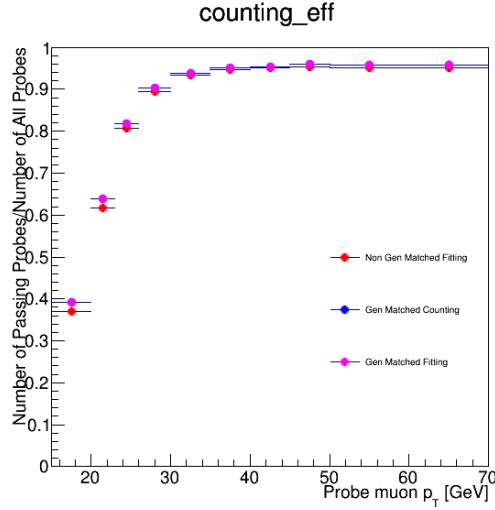


Figure 50: HLT\_L2DoubleMu23\_NoVertex trigger efficiency vs.  $p_T$  graph. The results from fitting, counting and MC Truth fitting methods are compared. Note that Tag and Probe method is applied.

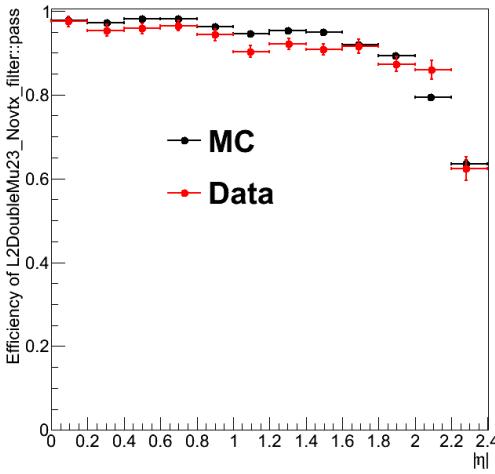


Figure 51: HLT\_L2DoubleMu23\_NoVertex trigger efficiency vs.  $|\eta|$  graph for both the simulation and data with Tag and Probe applied.

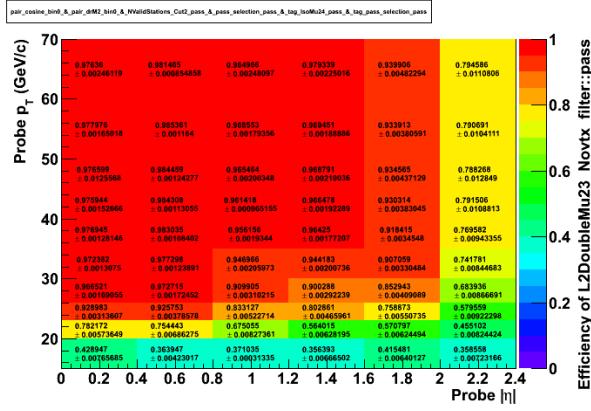


Figure 52: HLT\_L2DoubleMu23\_NoVertex trigger efficiency distribution on the  $p_T$  and  $|\eta|$  plane for MC with *Tag and Probe* applied.

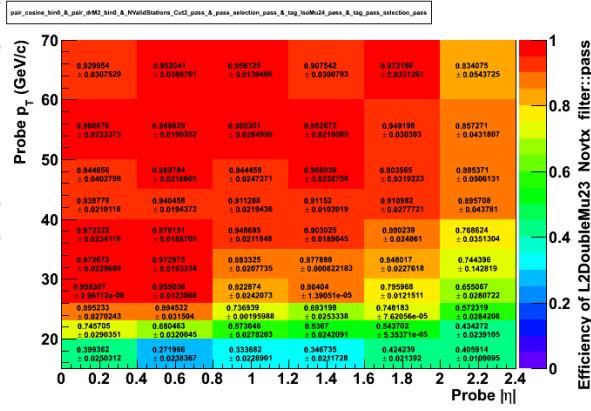


Figure 53: HLT\_L2DoubleMu23\_NoVertex trigger efficiency distribution on the  $p_T$  and  $|\eta|$  plane for data with *Tag and Probe* applied.

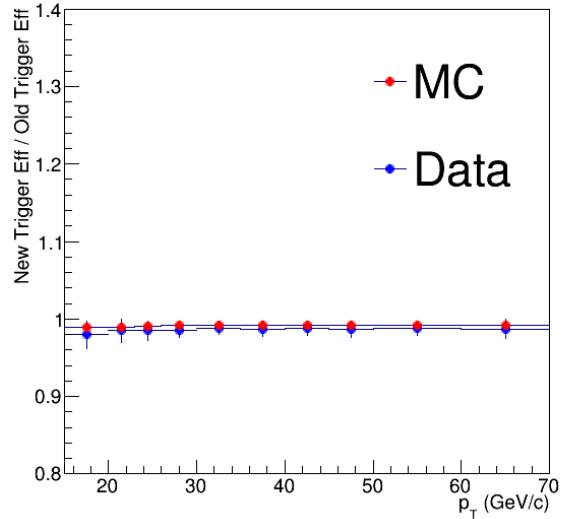


Figure 54: The ratio of HLT\_L2DoubleMu23\_NoVertex\_2Cha\_Angle2p5 trigger efficiency to HLT\_L2DoubleMu23\_NoVertex trigger efficiency vs.  $p_T$  graph for both data and the simulation.

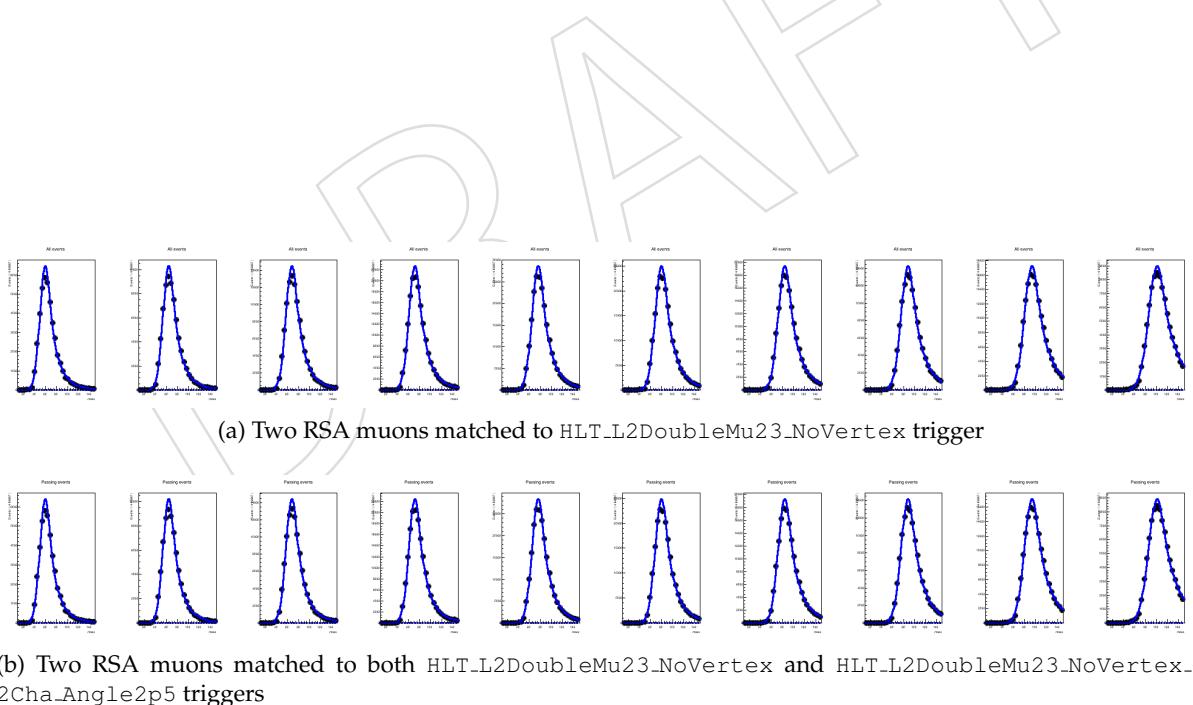


Figure 55: MC fitting results of dimuon mass distributions. The ratio of the integrals of the two curves in each  $p_T$  bin gives the fraction of dimuons matched to HLT\_L2DoubleMu23\_NoVertex trigger given that HLT\_L2DoubleMu23\_NoVertex is fired.

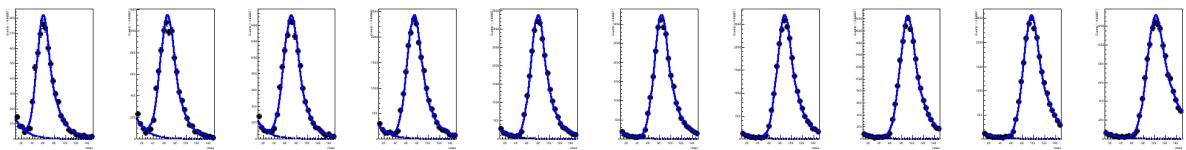
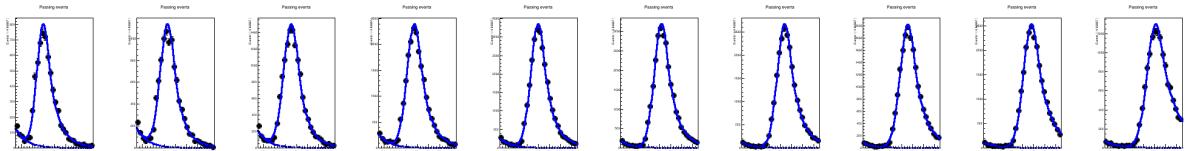
(a) Two RSA muons matched to `HLT_L2DoubleMu23_NoVertex` trigger(b) Two RSA muons matched to both `HLT_L2DoubleMu23_NoVertex` and `HLT_L2DoubleMu23_NoVertex_2Cha_Angle2p5` triggers

Figure 56: Data fitting results of dimuon mass distributions. The ratio of the integrals of the two curves in each  $p_T$  bin gives the fraction of dimuons matched to `HLT_L2DoubleMu23_NoVertex_2Cha_Angle2p5` trigger given that `HLT_L2DoubleMu23_NoVertex` is fired.

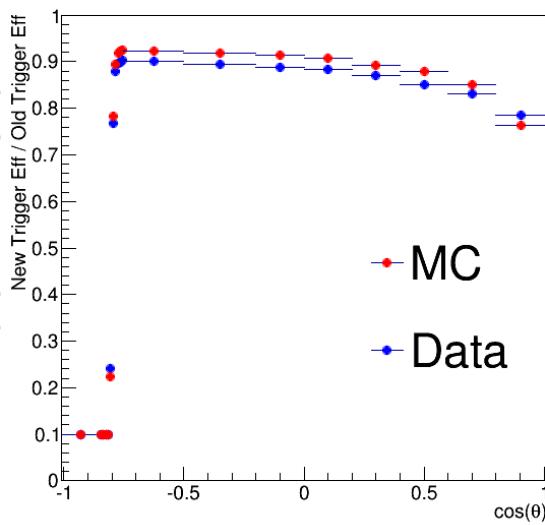


Figure 57: The ratio of `HLT_L2DoubleMu23_NoVertex_2Cha_Angle2p5` trigger efficiency to `HLT_L2DoubleMu23_NoVertex` trigger efficiency vs.  $\cos \theta$ . Note that  $\theta$  is the 3D angle between the two muons.

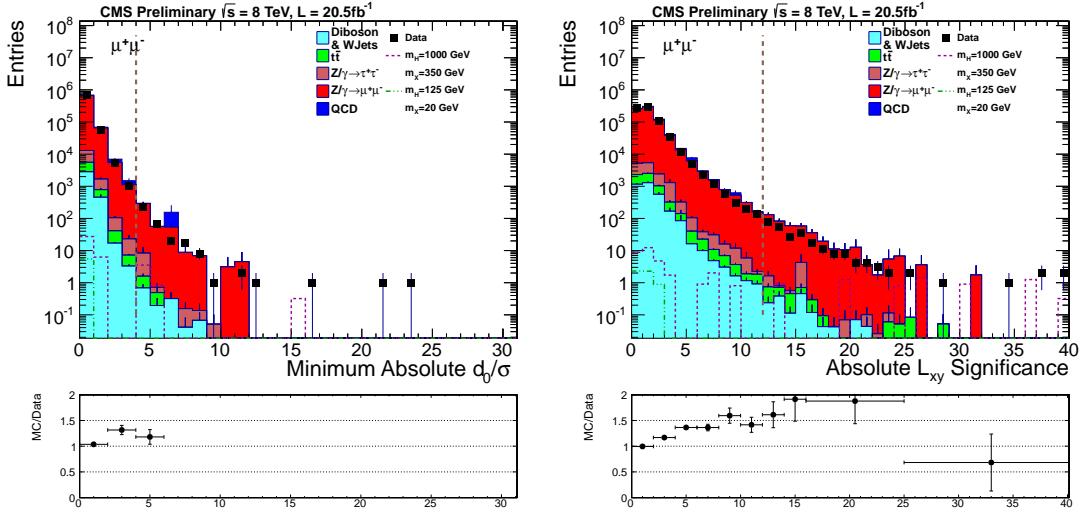


Figure 58: (Left) Distribution of the minimum absolute transverse impact parameter significance of the two muons for the dimuon candidates passing the selection. (Right) Distribution of the transverse decay length significance of the dimuon candidates passing the selection. The selection here is defined as the full analysis selection except  $|d_0|/\sigma_d$ ,  $L_{xy}/\sigma_{L_{xy}}$  and track rejection cuts in the control region. In addition to the selection defined, no cut is applied to the parameter shown in the plots.

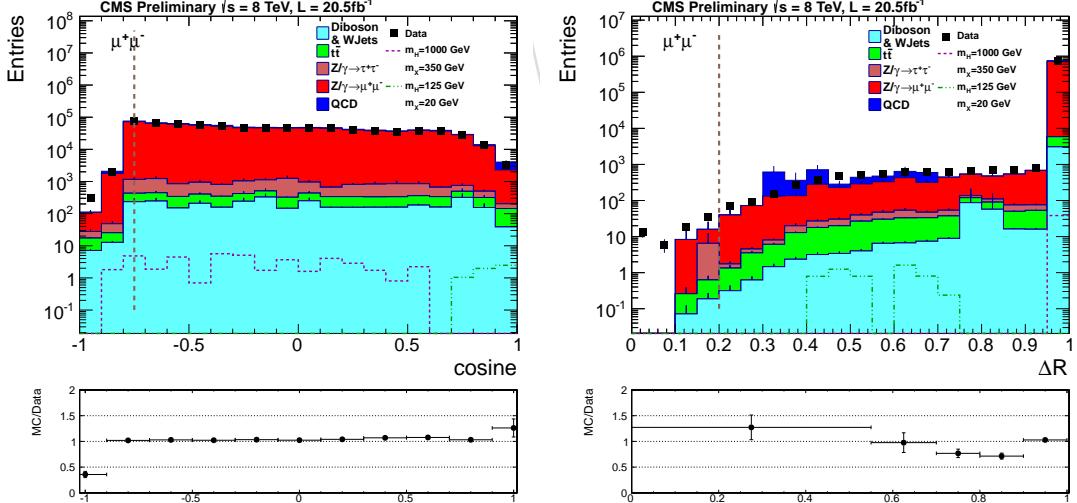


Figure 59: (Left) Distribution of the cosine of the 3D opening angle,  $\cos(\alpha)$ , between the two muons for the dimuon candidates passing the selection. (Right) Distribution of  $\Delta R$  separation between the two muons for the dimuon candidates passing the selection except  $|d_0|/\sigma_d$ ,  $L_{xy}/\sigma_{L_{xy}}$  and track rejection cuts in the control region.

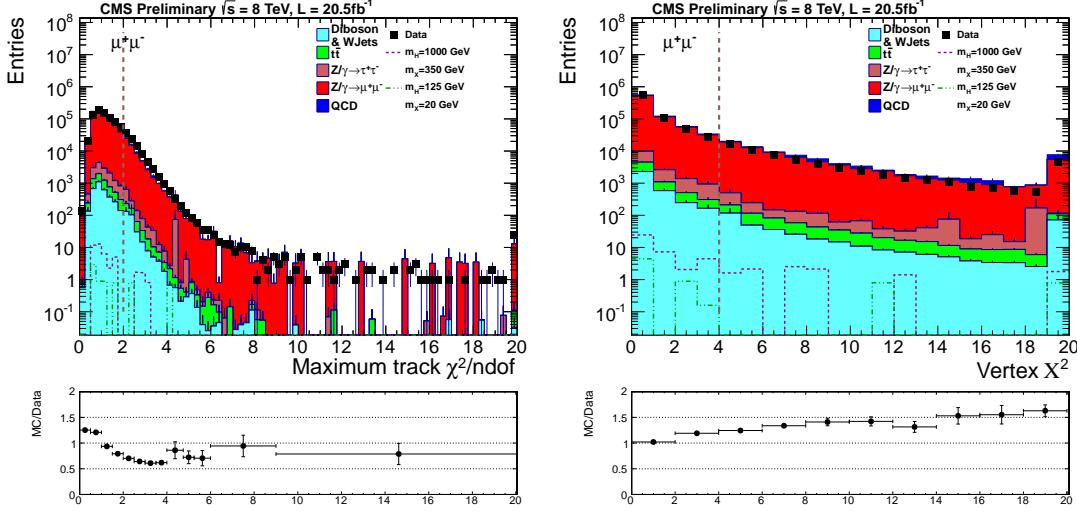


Figure 60: (Left) Distribution of the maximum normalized track  $\chi^2$  of the two muons for the dimuon candidates passing the selection. (Right) Distribution of the normalized vertex  $\chi^2$  of the dimuon candidates passing the selection except  $|d_0|/\sigma_d$ ,  $L_{xy}/\sigma_{L_{xy}}$  and track rejection cuts in the control region.

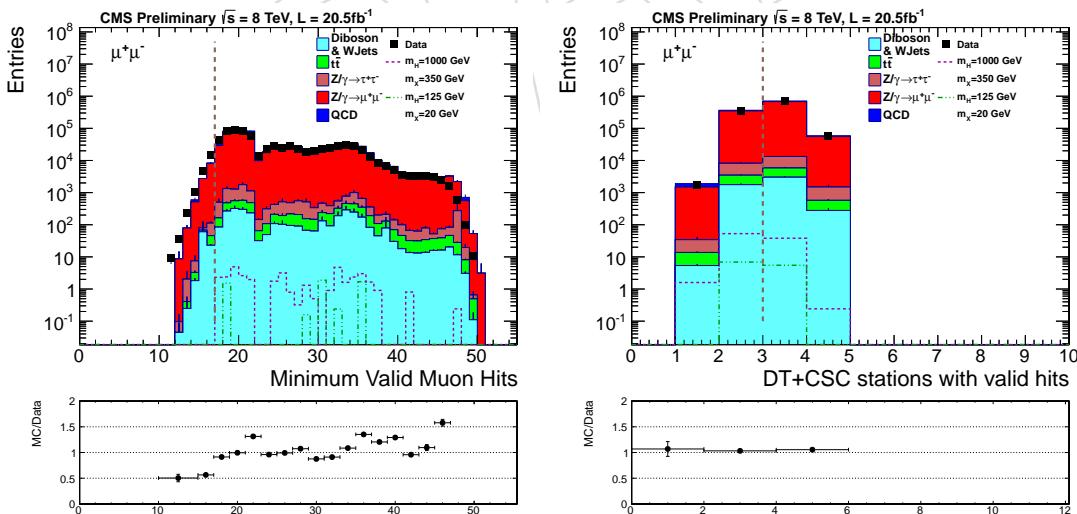


Figure 61: (Left) Distribution of the minimum number of valid muon hits of the two muons for the dimuon candidates passing the selection. (Right) Distribution of the minimum number of valid muon stations of the two muons for the dimuon candidates passing the selection except  $|d_0|/\sigma_d$ ,  $L_{xy}/\sigma_{L_{xy}}$  and track rejection cuts in the control region.

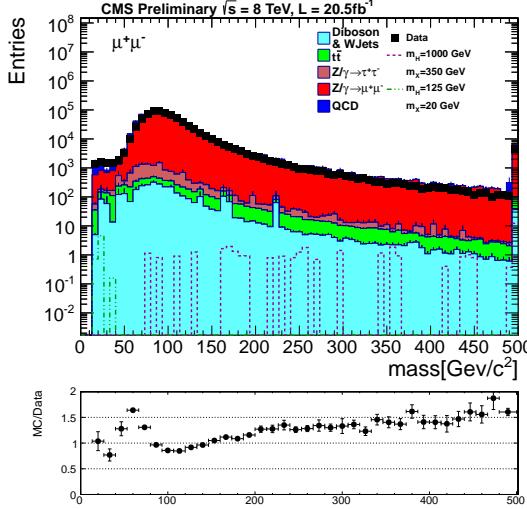


Figure 62: Distribution of the mass for the dimuon candidates passing the selection except  $|d_0|/\sigma_d$ ,  $L_{xy}/\sigma_{L_{xy}}$  and track rejection cuts in the control region.

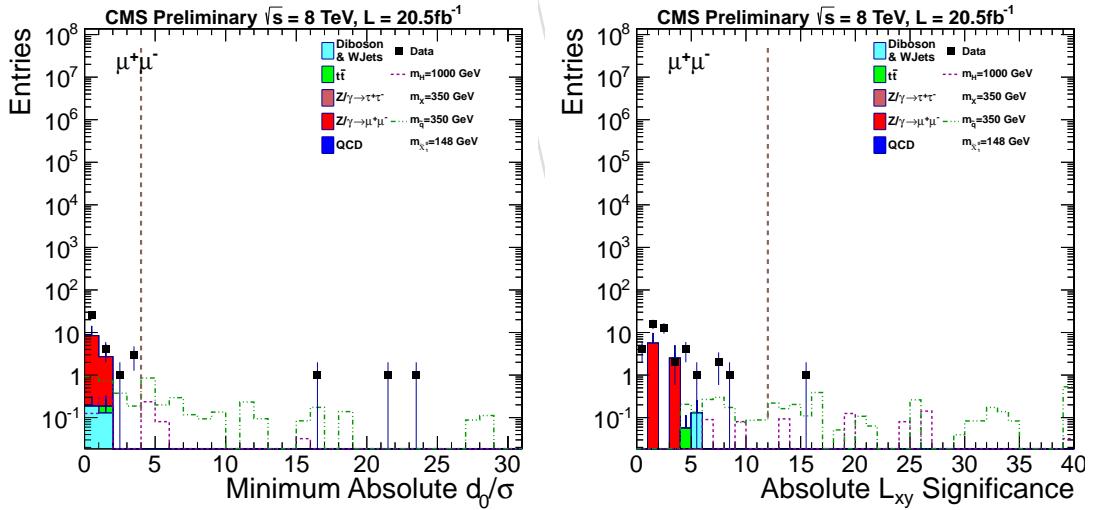


Figure 63: (Left) Distribution of the minimum absolute transverse impact parameter significance of the two muons for the dimuon candidates passing the selection. (Right) Distribution of the transverse decay length significance of the dimuon candidates passing the selection. The selection here is defined as the full analysis selection but with loose cuts at  $|d_0|/\sigma_d > 1$  and  $L_{xy}/\sigma_{L_{xy}} > 3$  in the control region. In addition to the selection defined, no cut is applied to the parameter shown in the plots.

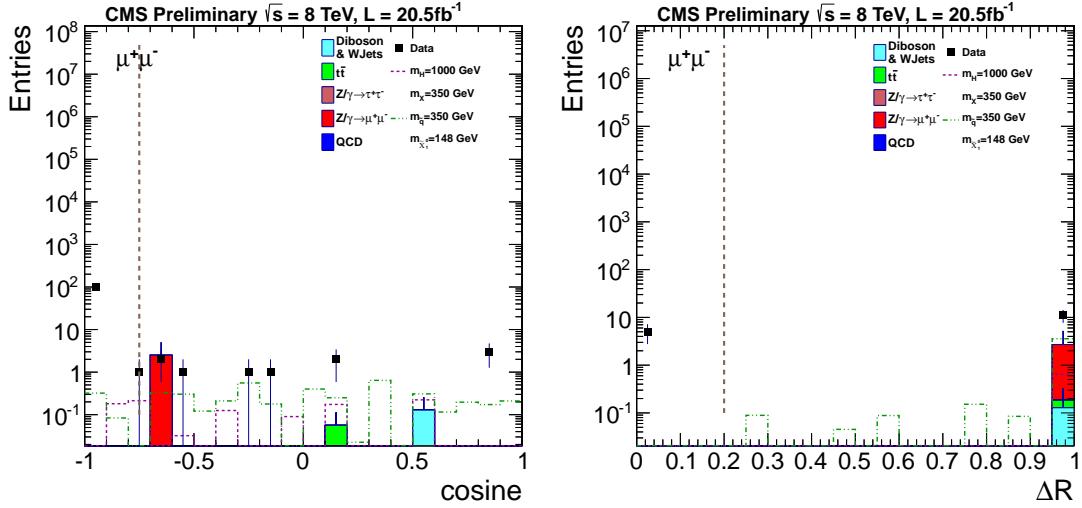
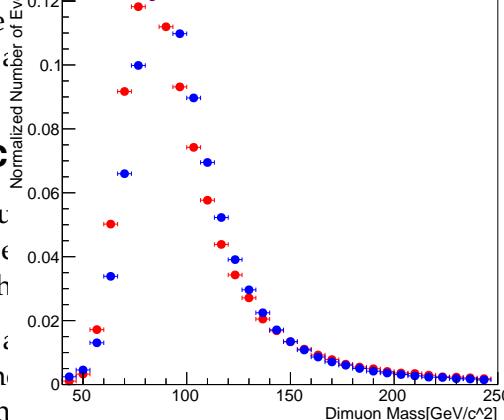


Figure 64: (I)

muons for the selection, we give that the muons are selected such that



## D Effect

The RSA muon selection, we given that the muons are selected such that the efficiency. In

consistent with coming from  $J/\psi$  and  $\Upsilon$  decays and  $\gamma$  conversions. The effect is expected to be negligible because loose selection criteria are applied to the muon transverse momentum and the dimuon mass. However, we performed a check to determine how much the analysis is affected by the  $p_T$  resolution of the RSA muons quantitatively.

Dimuon mass distribution is obtained for data and the background simulation by applying the full selection except the tracker track rejection and lifetime cuts, as given in Figure 68. The figure shows that the two distributions are peaked in different positions and that they have slightly different resonance widths.  $p_T$  resolution wise, data is not perfectly described by the simulation.

To quantify this discrepancy between the simulation and data, the distributions are fitted by Crystal Ball function due to the asymmetric long tail as presented in Figure 69. Mean values of the two fit functions are found to be within the uncertainties of each other and the difference in these mean values between data and MC can be expressed by a relative difference of 1.13%. Then,  $p_T$  of each muon in the simulation is increased by this factor in order to test how much the background MC dimuon mass distribution is shifted. As it can be seen in Figure 70, this operation does not shift MC dimuon mass distribution enough to overlap with that of data. Furthermore, the signal selection efficiencies remain unchanged after this rescaling.

In order to assess to what extent the analysis is sensitive to  $p_T$  resolution, the  $p_T$  of each muon in the background and signal simulations is increased by an unreasonably big factor of 10%.

opening angle,  $\cos(\alpha)$ , between the two muons for the selection but with loose cuts at

## on on the analysis

pared to the tracker muons [16]. In this uncertainty on this should be assigned to depend on the  $p_T$  measurement.

at trigger level. In the offline analysis, to ensure that the trigger has a good signal to background ratio. It should be 15  $\text{GeV}/c^2$  to discard the dimuons

consistent with coming from  $J/\psi$  and  $\Upsilon$  decays and  $\gamma$  conversions. The effect is expected to be negligible because loose selection criteria are applied to the muon transverse momentum and the dimuon mass. However, we performed a check to determine how much the analysis is affected by the  $p_T$  resolution of the RSA muons quantitatively.

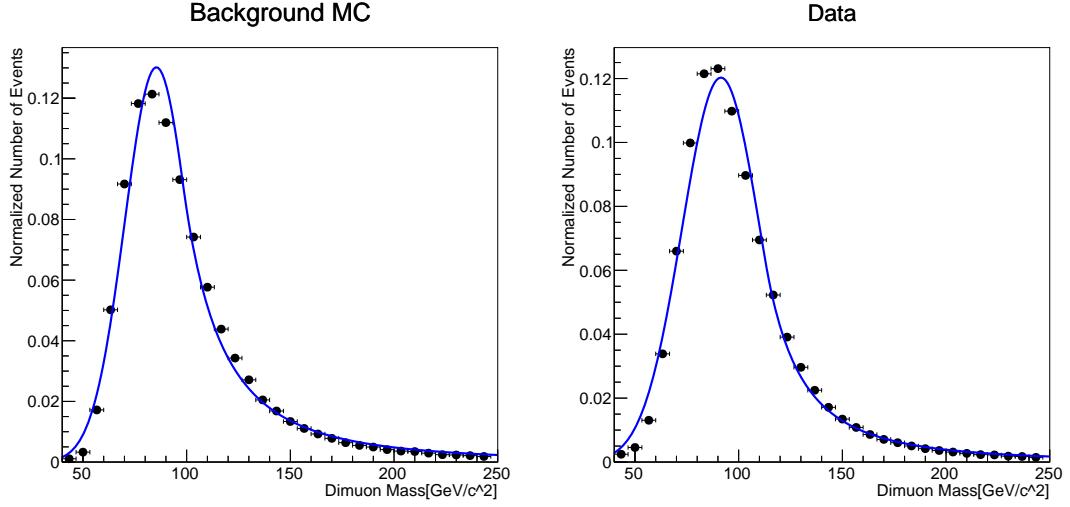


Figure 69: Dimuon mass distribution for background MC (left) and data (right) fitted by Crystal Ball function. Note that the number of events is normalized.

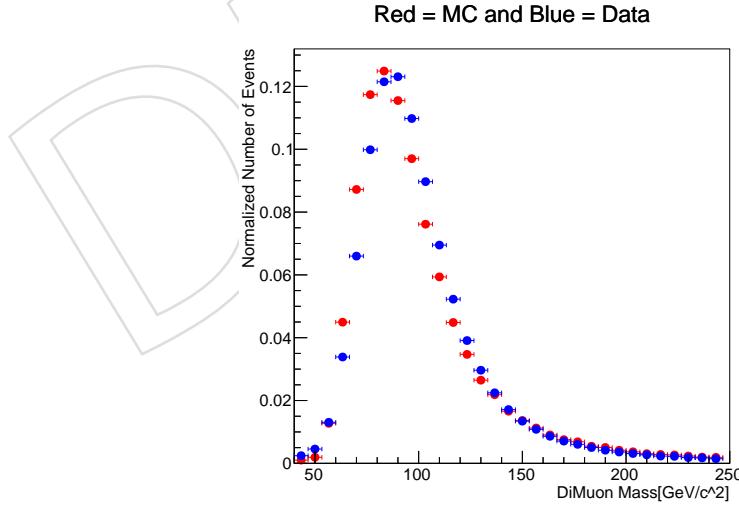


Figure 70: Dimuon mass distribution for both rescaled background MC, shown in red, and data, in blue.  $p_T$  is increased by a factor of 1.13% for each simulated muon. The dimuon masses are recomputed. Note that the number of events is normalized.

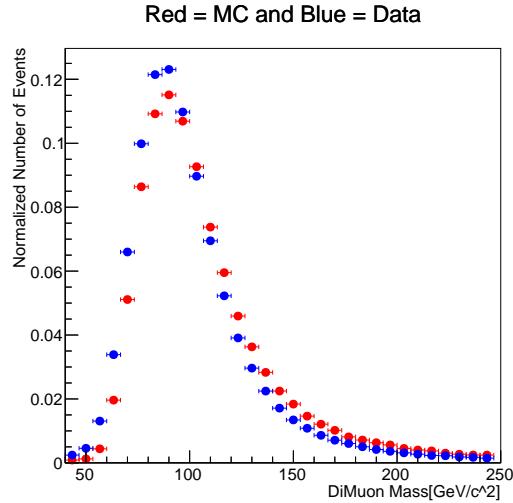


Figure 71: Dimuon mass distribution for both rescaled background MC, shown in red, and data, in blue.  $p_T$  is increased by a factor of 10% for each simulated muon. The dimuon masses are recomputed. Note that the number of events is normalized.

817 Then, the procedure is repeated as previously. Figure 71 indicates that the background MC  
 818 dimuon mass distribution is shifted even beyond that of data under such scaling. However, no  
 819 significant variation in the signal efficiencies of all mass points is observed. Therefore, assign-  
 820 ing an additional systematic uncertainty on  $p_T$  resolution of the RSA muons is not found to be  
 821 necessary since our analysis is significantly insensitive to it.

## 822 E $L_{xy}$ resolution of RSA muons for prompt decays

823 To evaluate the resolution of the transverse decay length, we used simulated  $Z \rightarrow \mu^+ \mu^-$  events  
 824 that have a reasonable agreement with data as given in Appendix C. Since  $Z$  bosons decay  
 825 promptly, there is no genuine secondary vertex. One might be led to think that this could yield  
 826 a half gaussian distribution for  $L_{xy}$  peaking at zero owing to the resolution effect. Nonetheless,  
 827 it will be shown why it is not the case.

828 Note that  $L_{xy}$  is the distance between the primary and secondary vertices, hence, always posi-  
 829 tive:

$$L_{xy} = \sqrt{(x_{PV} - x_{SV})^2 + (y_{PV} - y_{SV})^2}$$

830 where  $x_{PV} - x_{SV}$  and  $y_{PV} - y_{SV}$  are the distance between primary and secondary vertices in x  
 831 and y directions.

832 If the secondary vertex is assigned symmetrically around the primary vertex, then, the vertex  
 833 distances in x and y directions are expected to have a gaussian distribution as justified in Fig-  
 834 ure 72. That is, the secondary vertex is positioned around the primary vertex symmetrically  
 835 for the prompt decays. However, same symmetry can not be met by  $L_{xy}$  since it is square root  
 836 of the square sum of the two gaussian distributions. Even though Figure 73 might, at first  
 837 glance, infer that the secondary vertex positions are distributed unevenly around the primary  
 838 vertex, there is actually two fold symmetry coming from both x and y directions. So, it is not  
 839 straightforward to derive the resolution on  $L_{xy}$  in the first place.

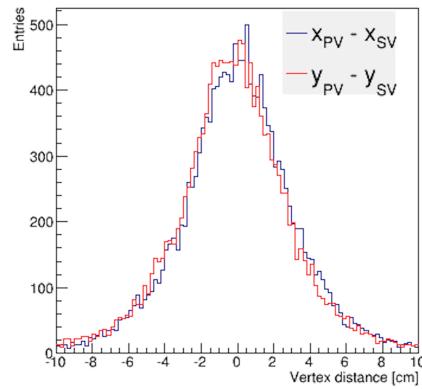


Figure 72: Distribution of the distance between primary and secondary vertices in x and y planes for the simulated  $Z \rightarrow \mu^+ \mu^-$  events. The shapes look gaussian indicating the symmetry of the decay.

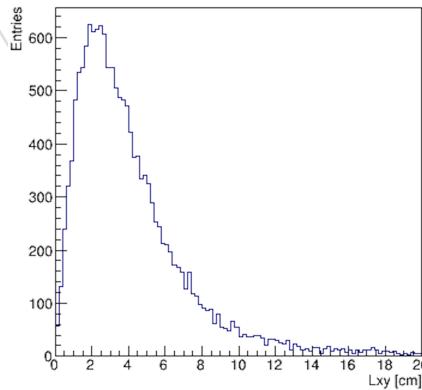


Figure 73: Distribution of the transverse decay length for the simulated  $Z \rightarrow \mu^+ \mu^-$  events. Since it is square root of the square sum of the two gaussian distributions, the resolution of  $L_{xy}$  can not be trivially extracted from this shape.

840  $L_{xy}$  resolution is computed by a rough estimate on the resolution of  $x_{PV} - x_{SV}$ . Figure 74 shows  
 841 that the gaussian fit applied is reasonable. The sigma value of the gaussian fit is about 3 cm.  
 842 Then, if two resolution effects are combined, one could roughly estimate the resolution of  $L_{xy}$   
 843 as  $3\sqrt{2} \approx 4.2$ cm. This result is valid only for prompt dimuon decays.

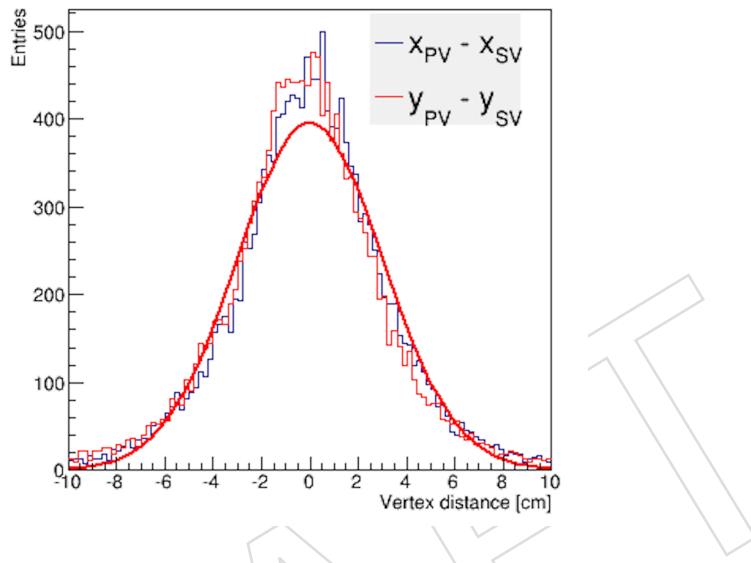


Figure 74: Distribution of the distance between primary and secondary vertices in x and y planes for the simulated  $Z \rightarrow \mu^+ \mu^-$  events. Bold red curve is the gaussian fit applied to one of the two. The sigma value of the fit function is 3.00918.

## 844 F Secondary vertex reconstruction efficiency

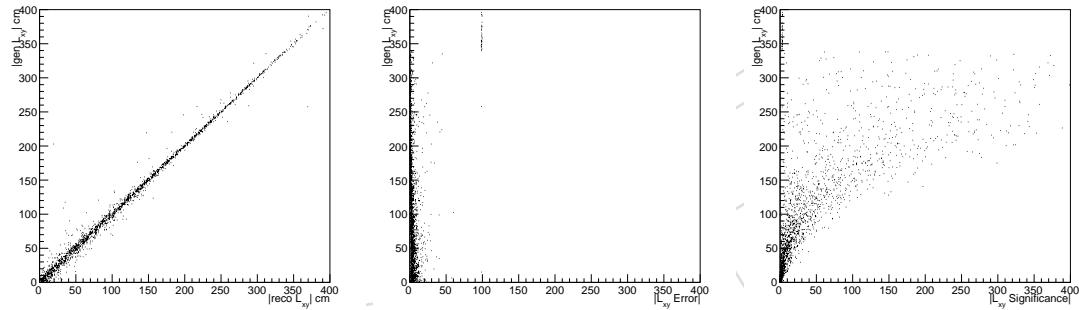
845 The algorithm used to fit a common secondary vertex from two tracks is the KalmanVertexFitter  
 846 [22]. The default implementation in the CMSSW release used in this analysis contains a cutoff  
 847 of the size of the silicon tracker. We extended the cutoff to the approximate beginning of the  
 848 muon chambers in order to recover efficiency for particles decaying outside the tracker. The  
 849 following changes were made to RecoVertex/VertexTools/src/SequentialVertexFitter.cc:

- 850 • TrackerBoundsRadius was changed from 112 to 500,
- 851 • TrackerBoundsHalfLength was changed from 273.5 to 1000.

852 To verify the performance of the algorithm for decays outside its previously intended range,  
 853 we study the secondary vertex reconstruction efficiency in signal MC. We also perform an ad-  
 854 dditional cross check by using cosmic data and comparing them with results obtained on cosmic  
 855 simulation.

### 856 F.1 Secondary vertex reconstruction efficiency in signal MC

857 We compute the secondary vertex reconstruction efficiency on signal MC. The reference sample  
 858 used is  $H \rightarrow XX$  with  $M_{H^0} = 400$  GeV/c $^2$ ,  $M_X = 50$  GeV/c $^2$  and  $c\tau = 800$  cm. We consider  
 859 the generated  $L_{xy}$  as the position of the secondary vertex. The primary vertex is reconstructed  
 860 using tracker tracks and it is measured with an accuracy of the order of 100 microns or less.  
 861 The resolution on  $L_{xy}$  is dominated by the resolution on the secondary vertex position. A  
 862 good agreement between measured and generated  $L_{xy}$  indicates a good agreement between  
 863 the generated and reconstructed position of the secondary vertex in the transverse plane. In all  
 864 the results shown in this section we do not apply any explicit trigger requirement. Figure 75  
 865 shows the correlation of  $L_{xy}$ , its error and its significance with the generator level  $L_{xy}$ . A good  
 866 correlation is found for the  $L_{xy}$  value up to 4 meters. The  $L_{xy}$  error shows a cutoff at about 3.4  
 867 meters where it assumes a default value of 100 cm. As a consequence, the  $L_{xy}$  significance is  
 868 generally below the cut of 12 used in the analysis for this region. However, the effective range  
 869 of the analysis is limited by the trigger efficiency to approximately 2.5 meters and this behavior  
 870 does not affect the results. It should be certainly improved in case the analysis sensitivity is  
 extended to even longer decay length values in the future.



871  
 872 Figure 75: Correlation of the reconstructed  $L_{xy}$  (left),  $L_{xy}$  error (center), and  $L_{xy}$  significance  
 873 (right) with the generated  $L_{xy}$ . The  $L_{xy}$  error defaults to a value around 100cm after approxi-  
 874 mately 3.4 meters. This is outside the effective region of the analysis that is limited to approxi-  
 875 mately 2.5 meters by the trigger efficiency.

876 To evaluate the secondary vertex reconstruction efficiency we consider all generated long-lived  
 877 particles decaying to dimuons where the two muons are matched to RSA muons and among  
 878 these events (denominator), we consider the cases where a secondary vertex has been suc-  
 879 cessfully reconstructed (numerator). The efficiency is computed as the ratio of the numerator  
 880 and the denominator terms. The distributions of the number of all generated long-lived par-  
 881 ticles decaying to dimuons and of those passing the numerator and denominator selections  
 882 are shown in Figure 76. Figure 77 shows the secondary vertex reconstruction efficiency. No  
 883 significant dependence on  $L_{xy}$  is observed and we do not assign any additional systematic  
 884 uncertainty.

### 885 F.2 Secondary vertex reconstruction efficiency in cosmics data and simulation

886 To evaluate the efficiency for data as a function of the displacement of the tracks with respect  
 887 to the expected beamspot, we use cosmic data and cosmic simulation. The datasets used are  
 888 described in Section 6.3 with the difference that for the data we utilize cosmics collected in the  
 889 "Run C" period instead of the CRAFT period.

890 In cosmic events there is no well defined  $L_{xy}$  so it is not possible to express the efficiency as a

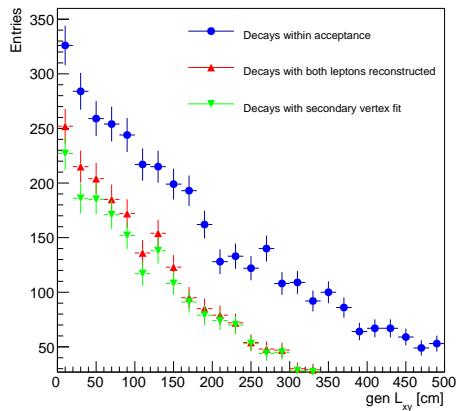


Figure 76: Distribution of the generated  $L_{xy}$  for all generated long-lived particles decaying to muons (blue), for those that have both muons reconstructed (red), and for the ones where a secondary vertex is also successfully reconstructed (green).

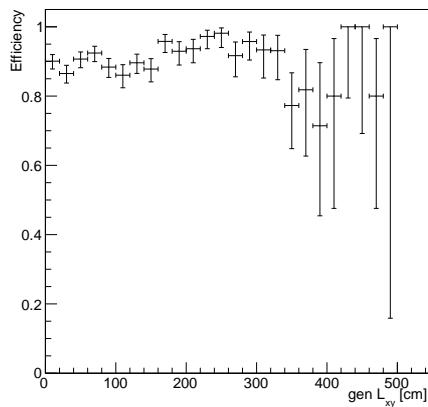


Figure 77: Efficiency to reconstruct a secondary vertex in signal MC events for  $M_{H^0} = 400$   $\text{GeV}/c^2$ ,  $M_X = 50$   $\text{GeV}/c^2$  and an average lifetime of 800 cm.

function of this variable. The impact parameter of the tracks, however, is a good measurement of the displacement of the reconstructed secondary vertex. We therefore show the efficiency vs  $d_0$  and we compare the results obtained in data and simulation. The result cannot be directly compared to the efficiency obtained in the previous section but it provides a cross check of the agreement between the secondary vertex reconstruction efficiencies in data and simulation. Note, however, that if the track is completely straight, the position of the secondary vertex is arbitrary since the two RSA tracks, when extrapolated, will coincide (ideally). To reduce this effect we require that the two RSA muons are not completely back-to-back. We select events with only two reconstructed RSA muons with an angle between them,  $\alpha$ , such that  $-0.95 < \cos(\alpha) < 0.90$ . The upper bound is to avoid possible cases of duplicates where the two RSA tracks are actually reconstructed from a subset of the hits of the same half of the cosmic. The two muons are also required to be within the acceptance of the analysis ( $p_T > 26$  GeV and  $|\eta| < 2$ ). Among those events we select the ones where a secondary vertex was successfully reconstructed. The ratio of these two categories of events is a measurement of the (fake) secondary vertex reconstruction efficiency in the cosmic events. The distributions of number of events with two reconstructed RSA muons within the acceptance as a function  $d_0$  of the highest  $p_T$  track in the event and of those where a secondary vertex is reconstructed are shown in Figure 78 for data and simulation. The resulting efficiencies are presented in Figure 79. Good agreement is found between data and simulation and we do not assign any additional systematic uncertainty from secondary vertex reconstruction efficiency. The choice of the  $d_0$  from the highest  $p_T$  RSA muon in the event is arbitrary. However, the highest  $p_T$  RSA muon is expected to be the better measured half of the cosmic and thus have a better resolution on  $d_0$ .

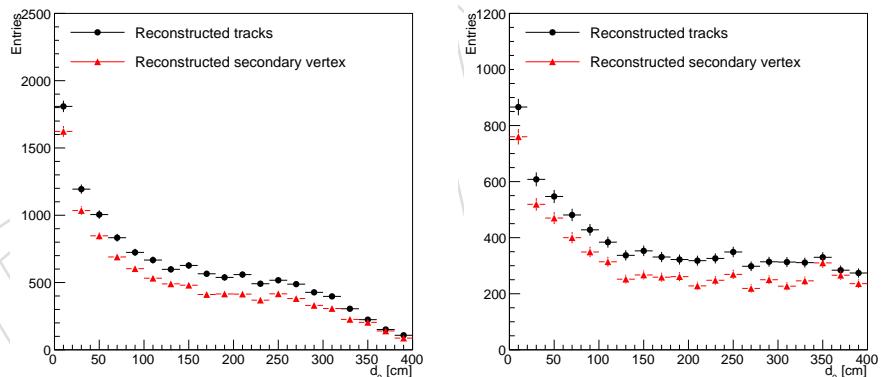


Figure 78: Distribution of the number of cosmic events in data (left) and simulation (right) with two reconstructed RSA muons within the acceptance (black) and with the additional requirement of a reconstructed secondary vertex (red) as a function of the  $d_0$  of the muon with the highest  $p_T$ .

909

## 910 G Alternative background systematics estimate

911 As an additional method to the principal one to derive a systematic uncertainty on the back-  
 912 ground estimate presented in Section 5.2, the  $L_{xy}/\sigma_{L_{xy}}$  tail-cumulative distribution for the data  
 913 is obtained after applying all the remaining cuts in the control region. Only 3 events are left  
 914 in the control region after the full selection except the  $L_{xy}/\sigma_{L_{xy}}$  cut is applied. Given the low  
 915 statistics, an unbinned fit is performed on the  $L_{xy}/\sigma_{L_{xy}}$  cumulative distribution with an expo-

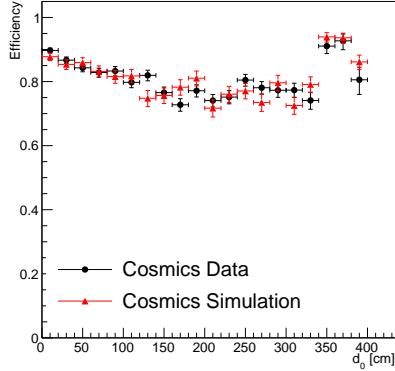


Figure 79: Efficiency to reconstruct a (fake) secondary vertex in cosmic data (black) and simulation (red) as a function of the  $d_0$  of the highest  $p_T$  RSA muon in the event.

916 nential function. The evaluation of the fit function at  $L_{xy}/\sigma_{L_{xy}} = 12$ , the cut value, gives another  
 917 estimate on the number of expected background events in the signal region, that is  $0.23 \pm 0.24$ .  
 918 Secondly, the procedure can also be iterated for the  $|d_0|/\sigma_d$  cumulative distribution. In this  
 919 case, there are only 2 events left if the complete selection is applied except the  $|d_0|/\sigma_d$  cut.  
 920 Figure 80 demonstrates both tail-cumulative distributions. Fitting the  $|d_0|/\sigma_d$  cumulative dis-  
 921 tribution yields an estimation of zero background events in the signal region. Note that the  
 922 errors in the fit results obtained by this method do not take into account that the data points  
 923 in these cumulative distributions are correlated to each other. Even though the two results are  
 924 consistent with each other, having very few statistics necessitates a more reliable method.

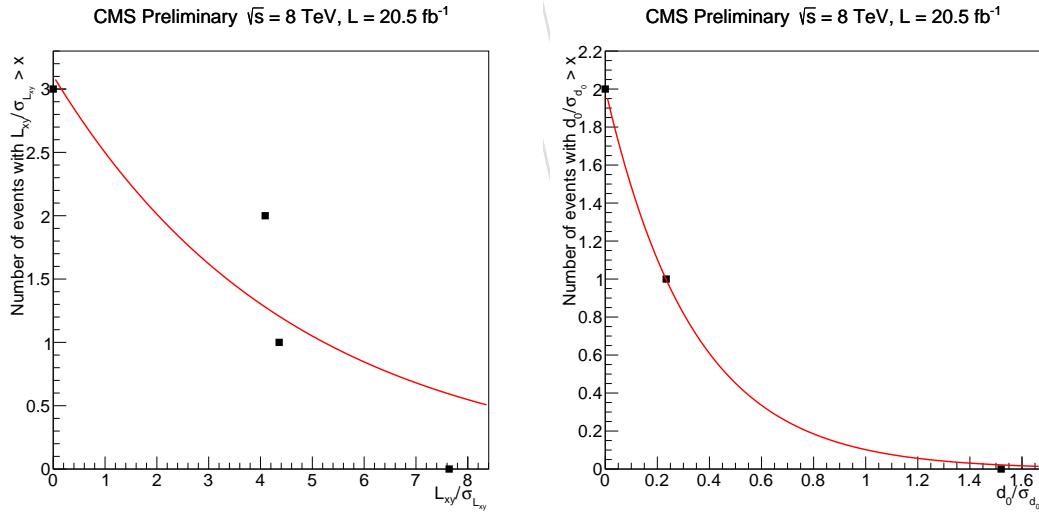


Figure 80: (Left)  $L_{xy}/\sigma_{L_{xy}}$  cumulative distribution as all the remaining cuts are applied. (Right)  $|d_0|/\sigma_d$  cumulative distribution. The most conservative estimate of background comes from the left one,  $0.23 \pm 0.24$ , suggesting that the systematic uncertainty on the number of estimated background can be set as high as 0.47. Note that the red curves represent the exponential unbinned fit functions.

## 925 H Bias in $p_T$ and low trigger efficiency

926 To understand the origin of the low trigger efficiency for displaced muons we study the  $p_T$   
 927 distribution as a function of the generated transverse decay length in a signal MC sample with  
 928  $M_H = 1000 \text{ GeV}/c$ ,  $M_X = 350 \text{ GeV}/c$  and  $c\tau = 350 \text{ cm}$ . We only consider cases where only one  
 929 of the X bosons decays to muons. Figure 81 shows the generator-level  $L_{xy}$  distribution used  
 930 in this study and Figure 83 shows how the  $d_0$  of the lowest  $p_T$  muon in the pair is correlated  
 931 with  $L_{xy}$ . The distribution of the generated  $p_T$  vs. the generated  $L_{xy}$  for all the events and for  
 932 those that pass the di-muon trigger is shown in Figure 82. For all the plots in this section the  $p_T$   
 933 shown is the one for the lowest  $p_T$  muon in the pair. As expected, there is no strong correlation  
 934 between  $p_T$  and  $L_{xy}$  when considering all the events. However, despite the fact that the trigger  
 935 only requires a  $p_T$  cut at 23  $\text{GeV}/c$ , most of the events removed by the trigger are not those  
 936 with generated  $p_T$  below the cut, but are those with high  $L_{xy}$ .

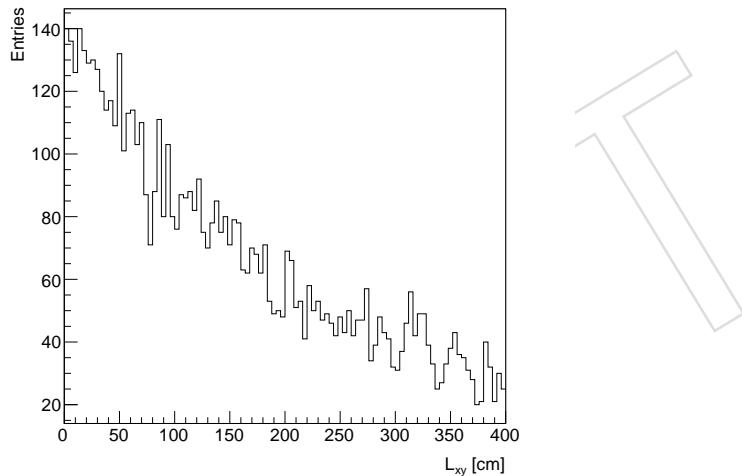


Figure 81: Generator-level  $L_{xy}$  distribution for a signal sample with  $M_H = 1000 \text{ GeV}/c$ ,  $M_X = 350 \text{ GeV}/c$  and  $c\tau = 350$ .

937 The reason for this becomes clear when looking at the distribution of  $p_T$  vs. the generator-level  
 938  $L_{xy}$  for the level 2 muons (L2Muons) used in the trigger, as shown in Figure 84. The  $p_T$  of  
 939 L2Muons shows a dependence on the  $L_{xy}$  and it is lower for muons produced in decays with  
 940 longer  $L_{xy}$ . Because of this bias a  $p_T$  cut in the trigger is partially also an  $L_{xy}$  cut and it tends to  
 941 reject more events with higher displacement.

942 Figures 85 and 86 show the same distribution for RSA muons and regular stand-alone (SA)  
 943 muons, respectively. A similar  $p_T$  bias is present also in those two cases and it appears to be  
 944 less pronounced for RSA muons and more for SA muons, as expected.

945 To better appreciate the effect, Figure 87 shows the generator-level  $p_T$  distribution in differ-  
 946 ent generator-level  $L_{xy}$  ranges and Figure 88 shows the same distribution for L2Muons, RSA  
 947 muons and SA muons before and after quality selections. A peak for a  $p_T$  below approximately  
 948 25  $\text{GeV}/c$  grows as the  $L_{xy}$  range increases. There is only reasonable correlation with the gen-  
 949 erated  $p_T$  distribution for the smallest  $L_{xy}$  range.

950 The  $p_T$  measured for a muon depends on how displaced it is from the beamspot. This is true,  
 951 at different levels, for L2Muons, RSA muons, and SA muons. The  $d_0$  for RSA is essentially  
 952 unbiased [16]. Despite this the  $p_T$  seems to be biased towards the beamspot. Given these

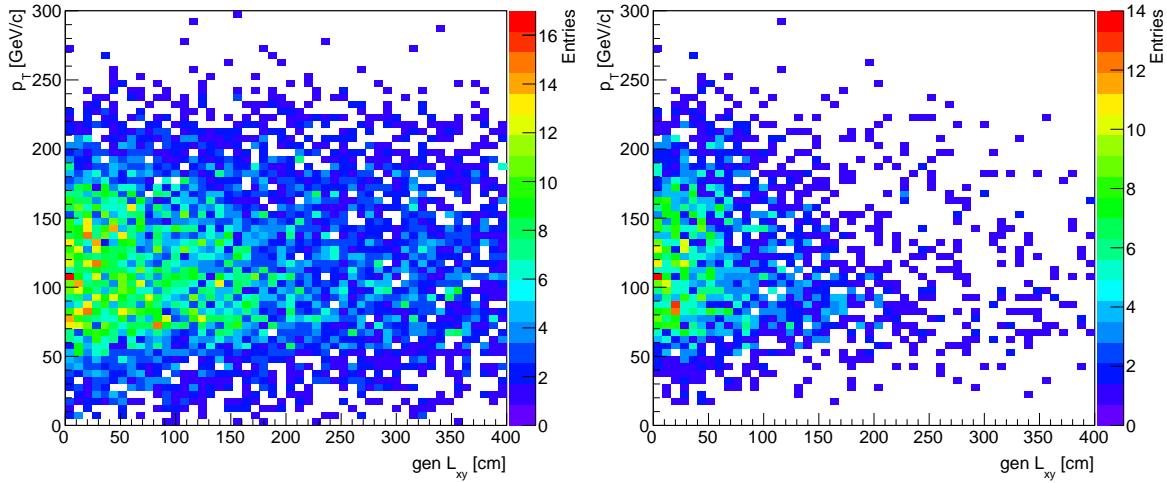


Figure 82: Distribution of generated  $p_T$  of the lowest  $p_T$  muon in the pair vs. generated  $L_{xy}$  for all events (left) and for events passing the trigger (right).

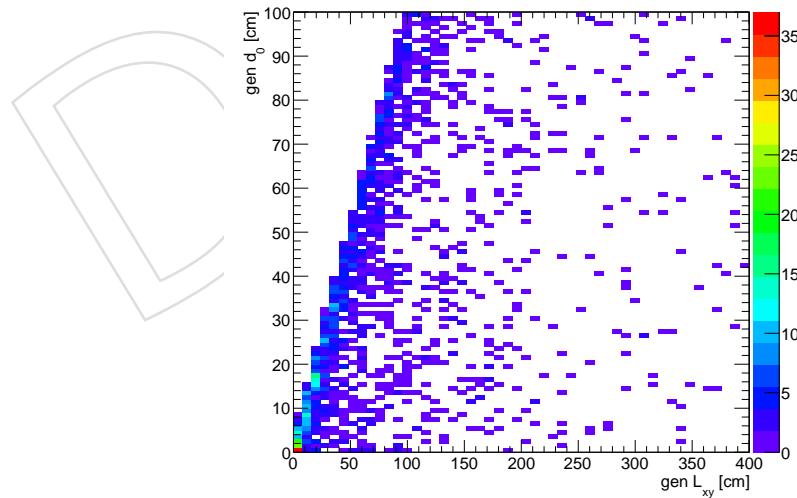


Figure 83: Generator-level  $d_0$  vs.  $L_{xy}$  for a signal sample with  $M_H = 1000$  GeV/c,  $M_X = 350$  GeV/c and  $c\tau = 350$ .

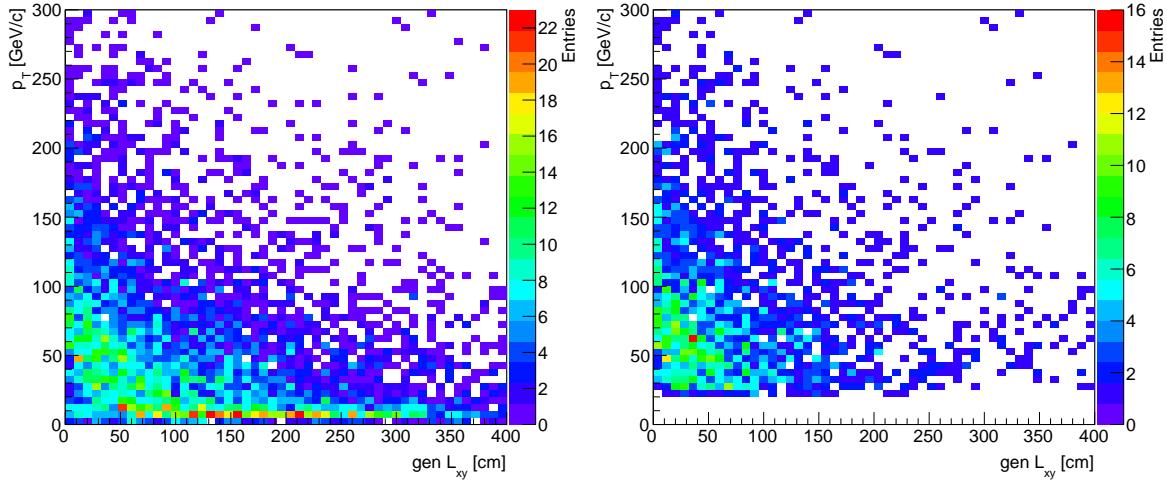


Figure 84: L2Muons  $p_T$  of the lowest  $p_T$  muon in the pair vs. generated  $L_{xy}$  for all events (left) and for events passing the trigger (right).

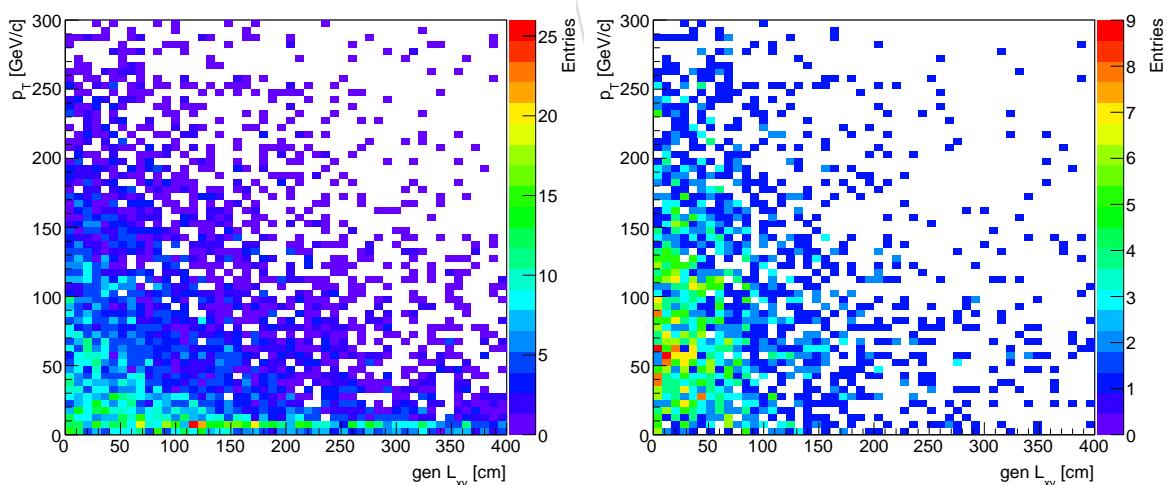


Figure 85: RSA muons  $p_T$  of the lowest  $p_T$  muon in the pair vs. generated  $L_{xy}$  for all events (left) and for events passing the trigger (right).

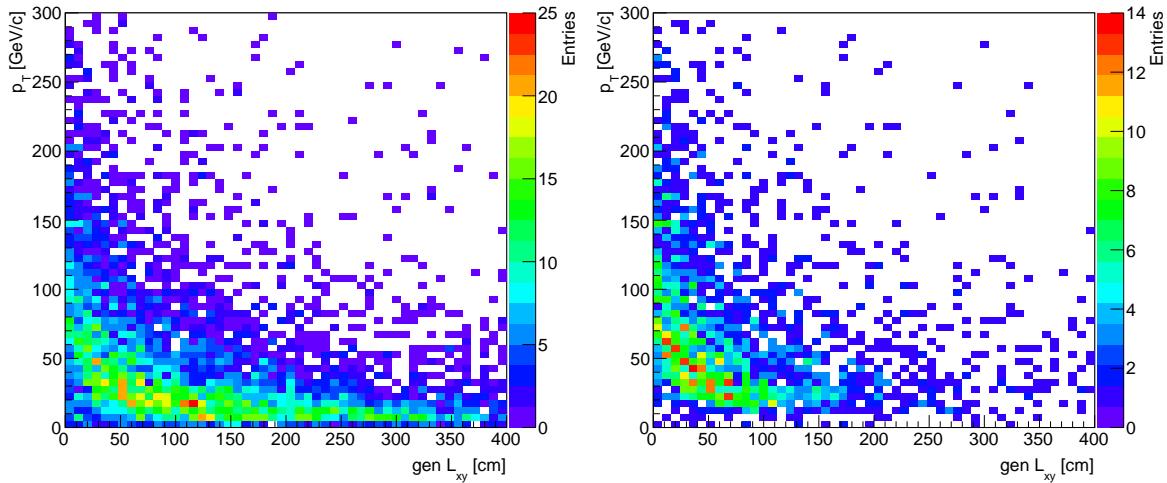


Figure 86: SA muons  $p_T$  of the lowest  $p_T$  muon in the pair vs. generated  $L_{xy}$  for all events (left) and for events passing the trigger (right).

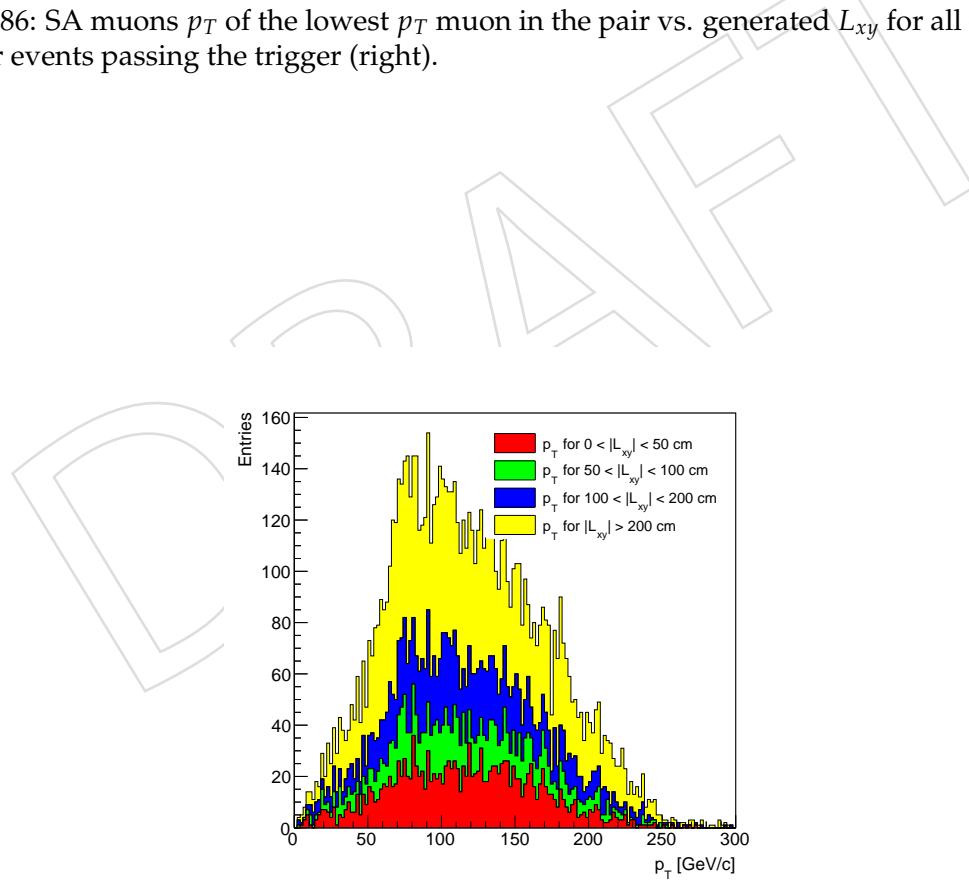


Figure 87: Generator-level  $p_T$  of the lowest  $p_T$  muon in the pair for different generated  $L_{xy}$  ranges for all events.

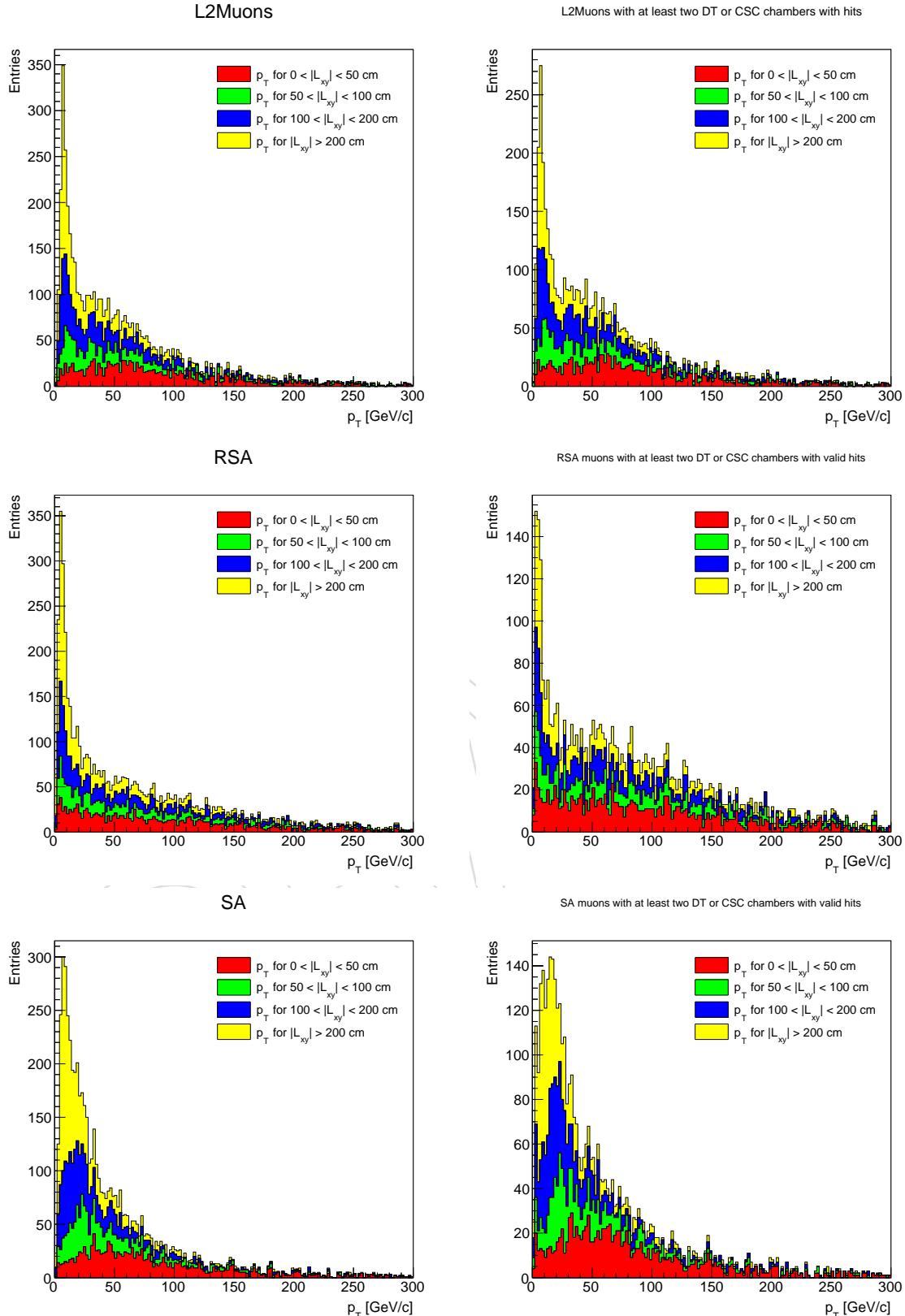


Figure 88: Distribution of the  $p_T$  of the lowest  $p_T$  muon in the pair for different generated  $L_{xy}$  ranges for all events for L2Muons (top row), RSA muons (middle row) and SA muons (bottom row). The left column shows the distributions without any selection cut and the right column after quality requirements. The quality requirement for L2Muons is at least two DT or CSC chambers with hits, and for SA and RSA muons is at least two DT or CSC chambers with valid hits.

953 observations the  $p_T$  cut in the trigger is effectively a cut on  $L_{xy}$  and it causes large inefficiencies  
 954 for displaced decays.

## 955 References

- [1] J. L. Hewett, B. Lillie, M. Masip et al., "Signatures of long-lived gluinos in split supersymmetry", *JHEP* **09** (2004) 070, doi:10.1088/1126-6708/2004/09/070, arXiv:hep-ph/0408248.
- [2] R. Barbier, C. Berat, M. Besancon et al., "R-parity violating supersymmetry", *Phys. Rept.* **420** (2005) 1, doi:10.1016/j.physrep.2005.08.006, arXiv:hep-ph/0406039.
- [3] T. Han, Z. Si, K. M. Zurek et al., "Phenomenology of hidden valleys at hadron colliders", *JHEP* **07** (2008) 008, doi:10.1088/1126-6708/2008/07/008, arXiv:0712.2041.
- [4] L. Basso, A. Belyaev, S. Moretti et al., "Phenomenology of the minimal B-L extension of the Standard model:  $Z'$  and neutrinos", *Phys. Rev. D* **80** (2009) 055030, doi:10.1103/PhysRevD.80.055030, arXiv:0812.4313.
- [5] M. J. Strassler and K. M. Zurek, "Discovering the Higgs through highly-displaced vertices", *Phys. Lett. B* **661** (2008) 263, doi:10.1016/j.physletb.2008.02.008, arXiv:hep-ph/0605193.
- [6] CMS Collaboration Collaboration, "Search for long-lived particles decaying to final states that include dileptons".
- [7] D0 Collaboration, "Search for neutral, long-lived particles decaying into two muons in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$ -TeV", *Phys. Rev. Lett.* **97** (2006) 161802, doi:10.1103/PhysRevLett.97.161802, arXiv:hep-ex/0607028.
- [8] D0 Collaboration, "Search for long-lived particles decaying into electron or photon pairs with the D0 detector", *Phys. Rev. Lett.* **101** (2008) 111802, doi:10.1103/PhysRevLett.101.111802, arXiv:0806.2223.
- [9] ATLAS Collaboration, "Search for a light Higgs boson decaying to long-lived weakly-interacting particles in proton-proton collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector", *Phys. Rev. Lett.* **108** (2012) 251801, arXiv:1203.1303.
- [10] ATLAS Collaboration, "Search for displaced vertices arising from decays of new heavy particles in 7 TeV pp collisions at ATLAS", *Phys. Lett. B* **707** (2012) 478, doi:10.1016/j.physletb.2011.12.057, arXiv:1109.2242.
- [11] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **03** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.
- [12] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at  $\sqrt{s} = 7$  TeV", *Journal of Instrumentation* **7** (2012), no. 10, P10002.
- [13] S. M. T. Sjöstrand and P. Z. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* **05** (2006) 576, arXiv:0603175.
- [14] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method", *JHEP* **0711** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

- 992 [15] GEANT4 Collaboration, “GEANT4: A simulation toolkit”, *Nucl. Instrum. Meth. A* **506**  
993 (2003) 250, doi:10.1016/S0168-9002(03)01368-8.
- 994 [16] D. Mattia et al., “Studies of Stand-Alone Muons Reconstruction for Displaced Muons”,  
995 CMS Analysis Note 2011/487, (2011).
- 996 [17] CMS Collaboration, “CMS luminosity based on pixel cluster counting - Summer 2013  
997 update”, CMS Physics Analysis Summary CMS-PAS-LUM-13-001, (2013).
- 998 [18] “CMS reweighting procedure for pile-up”.  
999 <https://twiki.cern.ch/twiki/bin/view/CMS/PileupReweighting>.
- 1000 [19] “CMS estimating systematic errors due to pile-up modeling”. <https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupSystematicErrors>.
- 1002 [20] “RooStats-based statistics tools for Higgs PAG”. <https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideHiggsAnalysisCombinedLimit>.
- 1004 [21] ATLAS Collaboration, CMS Collaboration, LHC Higgs Combination Group, “Procedure  
1005 for the LHC Higgs boson search combination in Summer 2011”, Public Note  
1006 ATL-PHYS-PUB-2011-011, ATL-COM-PHYS-2011-818, CMS-NOTE-2011-005, (2011).
- 1007 [22] T. Speer, K. Prokofiev, R. Fruehwirth et al., “Vertex fitting in the CMS Tracker”, CMS  
1008 Internal Note CMS-NOTE-2006/032, (2006).

DRAFT