

# CMS Draft Analysis Note

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

2015/05/23

Head Id: 219929

Archive Id: 231409M

Archive Date: 2013/12/08

Archive Tag: trunk

## A xxx Hadronic Search for Direct Stop with MT2 Variable Using full 2012 Data

Hamed Bakhshian<sup>1</sup>, Shirin Chenarani<sup>3</sup>, Esmaeel Eskandari<sup>1</sup>, Ali Fahim<sup>2</sup>, Abideh Jafari<sup>1</sup>, Saeid Paktinat Mehdiabadi<sup>1</sup>, and Maryam Zeinali<sup>1</sup>

<sup>1</sup> School of Particles and Accelerators, IPM, Tehran, Iran

<sup>2</sup> University of Tehran, Tehran, Iran, also at IPM

<sup>3</sup> Isfahan University of Technology, Isfahan, Iran, also at IPM

### Abstract

A hadronic search for direct production of stops is performed on  $19.6 \text{ fb}^{-1}$  of data from proton-proton collision in the center of mass energy of 8 TeV at CMS. The most important backgrounds, are estimated using data driven methods. It is shown that this analysis can access some parts of the SMS phase space which are not accessible by the  $E_T^{\text{miss}}$  analysis.

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

PDFAuthor: IPM MT2Top group

PDFTitle: A Hadronic Search for Direct Stop with MT2 Variable Using full 2012 Data

PDFSubject: CMS

PDFKeywords: CMS, physics, software, computing

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



# 1 Introduction

Supersymmetry [1] (SUSY) is one of the most promising extensions of the Standard Model of the elementary particles (SM) which solves the quadratic divergencies and therefore hierarchy problem simultaneously. It introduces a new symmetry between bosons and fermions and for every particle a sparticle is defined which is exactly the same, but differ in spin by 1/2. Since the super particles are not discovered yet, the supersymmetry should be a broken symmetry. Various mechanisms are introduced to break the symmetry softly without changing the other interesting features of the theory.

A search for new physics using  $19.6 \text{ fb}^{-1}$  of data from CMS taken in 2012 is documented in this note. Although the search is sensitive to any high scale new physics with a missing transverse momentum, R-parity conserving SUSY model is used to illustrate the performance of the method.

The search variable is the stransverse mass ( $M_{T2}$ ) which is the natural extension of the known transverse mass ( $m_T$ ) to a case when two massive particles with equal mass are created in pairs and decay via a chain of jets and leptons to two invisible particles. In the case of R-Parity conserving SUSY, the Lightest Supersymmetric Particle (LSP) escapes the detection and appears as a missing transverse momentum. The distribution of  $M_{T2}$  reflects the scale of the produced particles and is much higher for sparticles compared to the SM particles. Hence, SUSY should appear as an excess in the tail of the  $M_{T2}$  distribution. It was shown previously [2] that  $M_{T2}$  is a powerful variable to search for SUSY. Due to consistency of the data with background only hypothesis, low mass gluino and squarks have been ruled out. A main direction suggested by the theoreticians and phenomenologists is to search for the third generation of the sparticles. Since the third generation of the SM particles are heavier than the first two generations, in the SUSY sector, this generation can be much lighter. The current analysis is optimized to search for the direct production of the supersymmetric partner of the top quark (stop) in the hadronic final states. It is assumed that the pair produced stops undertake the following decay chain:

$$\tilde{t} \rightarrow t + \tilde{\chi}_1^0 \quad (1)$$

when top decays hadronically:

$$t \rightarrow b + W \rightarrow b + q + q' \quad (2)$$

and  $\tilde{\chi}_1^0$  can not be detected and appears as missing transverse momentum ( $E_T^{\text{miss}}$ ).

The previous version of the analysis which used only  $5.1 \text{ fb}^{-1}$  of 2012 data was documented in another analysis note [3]. In this version,  $19.6 \text{ fb}^{-1}$  of 2012 data is used and some parts of the analysis have been modified.

After the introduction in the next section the  $M_{T2}$  variable is introduced. A sepcial method for top reconstruction is described in section 3. The data and MC samples are defined in section 4. Different physical objects used in this analysis are introduced in section 5. Sections 6-7 review the procedure to select the trigger and cuts to have a better reach in this search. Our strategy to search for stop is explained in section 8. Data driven methods are used to estimate the contribution of the main SM backgrounds. Section 9 shows the methods and their performance. The statistical methods are used to interpret the results in section 10 and finally section 11 concludes the note.

## 2 The definition of $M_{T2}$

The variable transverse mass is used to measure the mass of the W-boson [4–7] in its decay to a lepton and a neutrino, where only the transverse missing energy due to undetected neutrino could be measured. It is defined as

$$m_T^2 = 2(E_T^l E_T^\nu - \vec{p}_T^l \cdot \vec{p}_T^\nu), \quad (3)$$

where for neutrino,  $E_T^\nu = p_T^\nu$ . The kinematic endpoint of  $m_T$  is an estimator for the W-mass, i.e.  $m_T^2 \leq m_W^2$ .

The  $M_{T2}$  variable [8, 9] is introduced and used in this analysis to discriminate between SUSY signal and the SM backgrounds while it is originally intended to estimate the mass of unseen particles. The kinematic endpoint of  $M_{T2}$  carries model independent information about the mass difference between the primary and the secondary supersymmetric particles. It is in particular useful to study events containing two simultaneous decays of a supersymmetric particle into a visible and an undetectable particle (e.g. neutralino). It is defined as

$$M_{T2}(m_\chi) = \min_{p_T^{\chi 1} + p_T^{\chi 2} = E_T^{miss}} [\max \{ m_T(p_T^{\chi 1}; m_\chi), m_T(p_T^{\chi 2}; m_\chi) \}], \quad (4)$$

where  $\chi$  stands for the neutralino whose mass is a free parameter in the evaluation of  $M_{T2}$ . The choice of maximum  $m_T$  is reasonable since none of the two transverse mass exceeds the mass of parents. The chosen transverse mass is minimized over the range of  $m_\chi$  which again ensures that  $m_T$  is less than the parents mass.

While for boosted systems in the transverse plane  $M_{T2}$  can be computed only numerically, there are analytic solutions [10] for unboosted scenarios. There, one can write the  $M_{T2}$  endpoint as a function of the masses.

To reconstruct the visible system as the input for  $M_{T2}$  calculation, the visible part of the event (jets in this analysis) is decomposed into two *pseudojets*. The procedure is known as *hemisphere reconstruction* and is already used in [11]. The two massless jets with the highest invariant mass define the primary two directions of hemispheres. Other jets are added to one of the hemispheres based on the minimal Lund distance (see e.g. [12]). The resulting  $M_{T2}$  variable is proven to well reject the multi-jet processes with non-genuine  $E_T^{miss}$  [11].

## 3 Top Reconstruction

To reconstruct top quarks a special method is used. The main features of the method are using a  $\chi^2$  and mass of the jets. A  $\chi^2$  is constructed based on the known masses of W and top, e.g.

$$\chi^2 = \frac{(M_{2j} - M_W)^2}{\sigma_W^2} + \frac{(M_{3j} - M_t)^2}{\sigma_t^2}; \quad (5)$$

$M_{2j}$  is usually the invariant mass of 2 jets making a W boson, but it can also be a heavy single jet, and  $M_{3j}$  is the invariant mass of 3 jets making a top quark. The uncertainty on invariant masses is computed as

$$\sigma_x = \sqrt{\frac{1}{4} \sum_i \left( \frac{\Delta p_i}{p_i} \right)^2 \left( \frac{\sum_{j \neq i} M_{ij}^2}{M} \right)^2 + \Gamma_x^2}; \quad (6)$$

where  $p_i$  uncertainty is taken as  $\frac{\Delta p_i}{p_i} = \frac{100\%}{\sqrt{p_i}}$  and  $\Gamma_x$  is the width of W (2.1 GeV) and top (10 GeV). Top reconstruction is started by reconstructing all possible W's from either 2 jets (W2j) or

1 heavy jet (W1j). Solutions are kept only if they have a  $\chi^2$  which is less than a fixed maximum value (2 in this analysis) for  $W$  in W2j and in W1j giving the  $W$  mass in a small window ( $80.4 \pm 10.0 \text{ GeV}/c^2$ ). To calculate the  $\chi^2$  for W1j, the width ( $\Gamma_x$ ) in Equation 6 is set equal to  $10 \text{ GeV}/c^2$ . If a heavy jet is in a given mass range and is  $b$ -tagged, it is considered as a coalescence of a  $b$  with a jet from  $W$  (W1b). The mass window for W1b is set to  $[40,130] \text{ GeV}/c^2$ .  $\chi^2 = 1$  is assigned to W1b candidates to avoid any systematic decrease of  $\chi^2$  for the top combinations containing such objects. All reconstructed  $W$ 's are ordered in their  $\chi^2$  value. In case of overlapping  $W$ 's, only the best  $\chi^2$  solution is kept. In the next step, the reconstructed  $W$ 's are used to reconstruct the top candidates by adding a free jet. To reduce the correlations, before using the  $W$ 's, their 4-vector is rescaled to give the correct  $W$  mass ( $80.4 \text{ GeV}/c^2$ ). If there is any overlap between the tops, the combination with a correct  $b$ -tagged jet ( $b+W$ ) or a W1b is preferred over the best  $\chi^2$  solution.

### 3.1 Performance of the Algorithm

The number of generated top quarks which are found at the reconstruction level demonstrates the performance of the top search algorithm. Only hadronically decaying top quarks are considered. The efficiency of the algorithm is defined as:

$$\epsilon_{topSearch} = N_{reco-top}^{matched} / N_{gen-top}^{hadronic}$$

where reconstructed top quarks are matched with generated ones if  $\Delta R(top_{rec}, top_{gen}) < 0.1$ .

The overall results are shown in Table 1.

#events	#gen-top-hadronic	#rec-top-all	#rec-top-matched	Overall efficiency
1.16 ME	2601	1331	936	36%

Table 1: The total efficiency of the top reconstruction algorithm. The efficiency is defined as the fraction of the generated top quarks which are reconstructed by the top search algorithm.

The efficiency versus different event kinematic variables is studied and the results are shown in Figure 1. The probability to find a hadronically decaying top is higher in higher jet multiplicities (Figure 1, top-left).

Although there is no constraint for the combination to contain a  $b$ -tagged jet, in case of a tagged jet in the top combination, it is preferred over minimum  $\chi^2$ . The efficiency is stable in  $N_b$ -jets bins, as expected. The last bin in Figure 1 (top-right) suffers from low statistics.

As the efficiency versus the top  $p_T$  is shown in Figure 1, if top quarks are generated with higher  $p_T$ , there is a higher probability for them to be reconstructed by our algorithm.

This study shows that efficiency is stable in  $M_{T2}$  bins, apart from the last two bins which have few entries.

The fake rate of the top reconstruction algorithm is also investigated. It is defined as the probability of reconstructing top from each 3 jets in a  $W + \text{jets}$  event. So the ratio is normalized to the number of jets. The results after applying all MT2b cuts on the WJetsToLNu-HT-400ToInf-8TeV sample are shown in Figure 2. It can be seen that the fake rate value is around 20%.

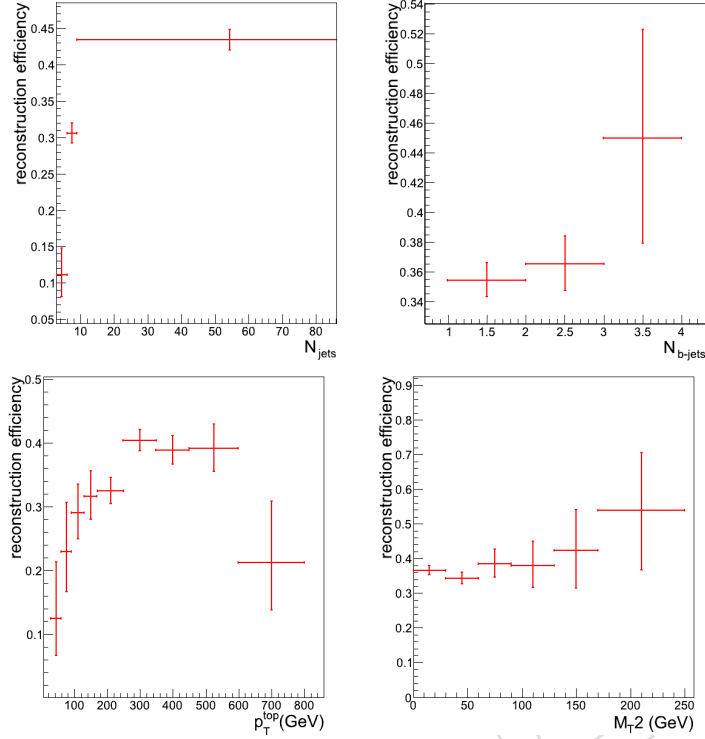


Figure 1: The efficiency of top quark reconstruction algorithm is shown vs. number of jets (top-left) and number of b-tagged jets (top-right), top  $p_T$  (bottom-left) and  $M_{T2}$  and (bottom-right).

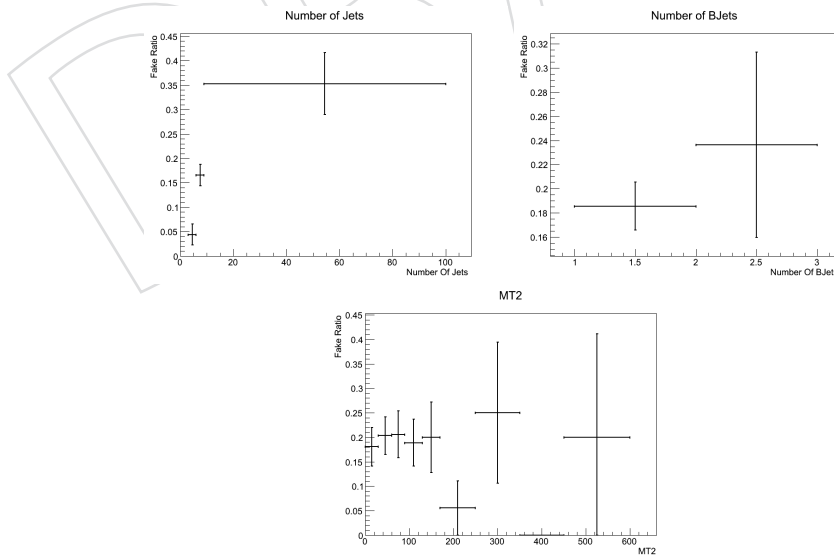


Figure 2: The fake rate of the top reconstruction algorithm is shown vs. number of jets (top-left), number of b-tagged jets (top-right) and  $M_{T2}$  (bottom).

## 4 Datasets and MC samples

To reconstruct the objects, the CMSSW\_5\_3\_7\_patch5 is used for both data and MC. The data used in the analysis corresponds to  $19.6 \text{ fb}^{-1}$  of proton-proton collisions in the center of mass energy of  $\sqrt{s} = 8 \text{ TeV}$  which was taken in 2012. The list of the datasets, the run range and the corresponding integrated luminosities are as follow:

- /MultiJet/Run2012A-13Jul2012-v1/AOD (190456-193621,  $952.6 \text{ pb}^{-1}$ )
- /MultiJet/Run2012A-recover-06Aug2012-v1/AOD (190782-190949,  $95.4 \text{ pb}^{-1}$ )
- /MultiJet/Run2012B-13Jul2012-v1/AOD (193834-196531,  $4.94 \text{ fb}^{-1}$ )
- /MultiJet/Run2012C-24Aug2012-v1/AOD (198022-198523,  $520.4 \text{ pb}^{-1}$ )
- /MultiJet/Run2012C-PromptReco-v2/AOD (198941-203742,  $6.9 \text{ fb}^{-1}$ )
- /MultiJet/Run2012D-PromptReco-v1/AOD (203777-208686,  $7.7 \text{ fb}^{-1}$ )

Only the lumisections with fully operative CMS subdetectors are used in this analysis (golden JSON files). To optimize the search method, MC samples are used for different Standard Model backgrounds and signals. These samples are officially generated and reconstructed by the CMS collaboration. The full list of the samples and their cross sections are given in Table 2. For most of the samples the most accurate calculation of the cross sections available in the literature (usually NLO and NNLO) are used. The contribution from the diboson samples was found to be very small.

## 5 Physics Object Definition and Preselections

This section, describes the physics objects used in this analysis.

### 5.1 PF Jets

- PF-CHS jets with  $p_T > 20 \text{ GeV}$  and  $|\eta| < 2.4$  are kept for the analysis.
- Jets are required to pass loose pf-jet id cuts listed below:
  - Number of constituents  $> 1$ ,
  - Neutral hadronic fraction  $< 0.99$ ,
  - Neutral electromagnetic fraction  $< 0.99$ ,
  - Charged hadronic fraction  $> 0$ ,
  - Charged electromagnetic fraction  $< 0.99$ ,
  - Charged multiplicity  $> 0$ .

### 5.2 PF Electrons

- PF electrons with  $p_T > 10 \text{ GeV}$  and  $|\eta| < 2.4$  are selected with ECAL gap veto.
- Electrons are required to pass cut-based electron id cuts corresponding to VBTF 95 working point, which is used for veto [13]. These set of cuts contain the requirements on  $|d0| < 0.04 \text{ cm}$  and  $|dz| < 0.2 \text{ cm}$ , for which both of them are calculated with respect to the primary vertex.
- Combined relative PF isolation below 0.15.

Table 2: List of the MC samples used in this analysis.

Sample name	$\sigma$ (pb)
QCD	
QCD-HT-100To250-TuneZ2star-8TeV-madgraph-pythia-Summer12-DR53X-PU-S10-START53-V7A-v1	10360000
QCD-HT-250To500-TuneZ2star-8TeV-madgraph-pythia6-...-2	276000
QCD-HT-500To1000-TuneZ2star-8TeV-madgraph-pythia6-...-2	8426
QCD-HT-1000ToInf-TuneZ2star-8TeV-madgraph-pythia6-...-2	204
Top	
T-t-channel-TuneZ2star-8TeV-powheg-tauola-...-v3	56.4
Tbar-t-channel-TuneZ2star-8TeV-powheg-tauola-...	30.7
T-tW-channel-DR-TuneZ2star-8TeV-powheg-tauola-...	22.4
Tbar-tW-channel-DR-TuneZ2star-8TeV-powheg-tauola-...	11.1773
T-s-channel-TuneZ2star-8TeV-powheg-tauola-...	3.79
Tbar-s-channel-TuneZ2star-8TeV-powheg-tauola-...	1.76
TTJets-MassiveBinDECAY-TuneZ2star-8TeV-madgraph-tauola-Summer12-DR53X-PU-S10-START53-V7A-v1	245.8
TTGJets-8TeV-madgraph-Summer12-DR53X-PU-S10-START53-V19-v1	2.166
TTH-Inclusive-M-125-8TeV-pythia6-Summer12-DR53X-PU-S10-START53-V7A-v1	0.130
TTWJets-8TeV-madgraph-Summer12-DR53X-PU-S10-START53-V7A-v1	0.232
TTWWJets-8TeV-madgraph-Summer12-DR53X-PU-S10-START53-V7A-v1	0.002037
TTZJets-8TeV-madgraph-v2-Summer12-DR53X-PU-S10-START53-V7A-v1	0.2057
WJets	
WJetsToLNu-HT-250To300-8TeV-madgraph-Summer12-DR53X-PU-S10-START53-V7A-v1	57.26
WJetsToLNu-HT-300To400-8TeV-madgraph-...	45.68
WJetsToLNu-HT-400ToInf-8TeV-madgraph-...	30.08
WJetsToLNu-HT-200To250-8TeV-madgraph-...-V7C-v1	90.27
ZJets	
DYJetsToLL-M-10To50filter-8TeV-madgraph-...-Summer12-DR53X-PU-S10-START53-V7A-v1	876.8
DYJetsToLL-M-50-TuneZ2Star-8TeV-madgraph-tarball- ...	3503.71
ZJetsToNuNu-50-HT-100-TuneZ2Star-8TeV-madgraph-ext-...	452.75
ZJetsToNuNu-200-HT-400-TuneZ2Star-8TeV-madgraph-ext-...	49.28
ZJetsToNuNu-400-HT-inf-TuneZ2Star-8TeV-madgraph-ext-...	6.26
ZJetsToNuNu-100-HT-200-TuneZ2Star-8TeV-madgraph-ext-...V7C-v1	190.39
SMS	
SMS-T2tt-mStop-150to350-mLSP-0to250-8TeV-Pythia6Z-Summer12-START52-V9-FSIM-v1-2	

### 5.3 PF Muons

- PF muons with  $p_T > 10$  GeV and  $|\eta| < 2.4$  are selected, which are asked to be global muons.
- Normalized  $\chi^2$  is required to be below 10.
- At least one valid track hit and at lease one valid pixel hit is required.
- Number of chambers with matched segments is required to be greater than one and



number of silicon layers should be above 5.

- Cuts on the  $|d0| < 0.2$  cm and the  $|dz| < 0.5$  cm, both with respect to the primary vertex, are applied.
- Combined relative PF isolation below 0.2.

## 5.4 PF Taus

In this note taus always mean hadronically decaying taus, unless stated otherwise.

- Hadron Plus Strip (HPS) algorithm identified PF-taus
- $p_T > 20$  GeV and  $|\eta| < 2.3$
- A decay into one or three prongs, plus eventually a  $\pi^0$ , is required
- Loose Electron Rejection: electron pion MVA discriminator  $< 0.6$
- Tight Muon Rejection: Tau Lead Track not matched to chamber hits, and no DT, CSC or RPC Hits in last 2 stations, and large enough energy deposit in ECAL + HCAL in 1 prong + 0 strip decay mode ( $\sum(\text{ECAL}+\text{HCAL}) > 0.2 \cdot p_T$ ).
- Loose Isolation ( $\Delta\beta$ -corrected):  $\Delta\beta$ -corrected  $\sum p_T$  of PF charged and PF gamma isolation candidates ( $p_T > 0.5$  GeV) less than 2 GeV (in a cone of  $\Delta R = 0.3$  around the tau axis), requiring 3 hits on tracks of charged isolation candidates.

## 5.5 PF $E_T^{\text{miss}}$

- Type1 corrected PF  $E_T^{\text{miss}}$  is used.

## 5.6 Jet/MET Smearing

Simulated events show discrepancies with data specially in  $E_T^{\text{miss}}$  and  $M_{T2}$  distributions. As shown in Figure 3, the data over MC shows a trend rather than fluctuating differences.

The energy of jets and  $E_T^{\text{miss}}$  in simulation are calibrated based on data. There are however

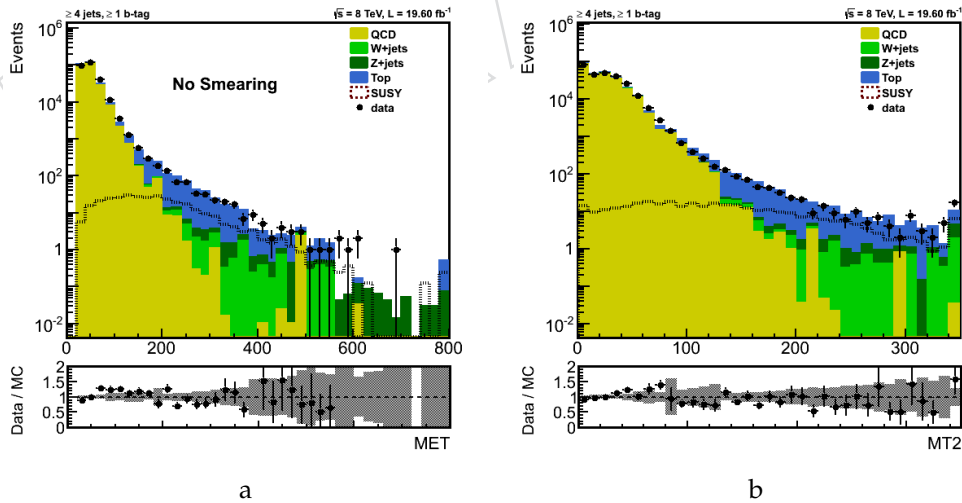


Figure 3: The distribution of  $E_T^{\text{miss}}$  (a) and  $M_{T2}$  (b) before smearing.

residual differences between data and simulation which are not covered by those corrections. Such differences could be improved by altering the jet energy resolution to match the data

and by correcting the  $E_T^{miss}$  accordingly. The CMS official recipe is followed to for the jet-  
 $E_T^{miss}$  smearing. Figure 4 illustrates the improvement achieved after smearing in  $E_T^{miss}$  and  $M_{T2}$   
distributions. Smeared jets and  $E_T^{miss}$  are used in event selection and in the rest of the analysis.

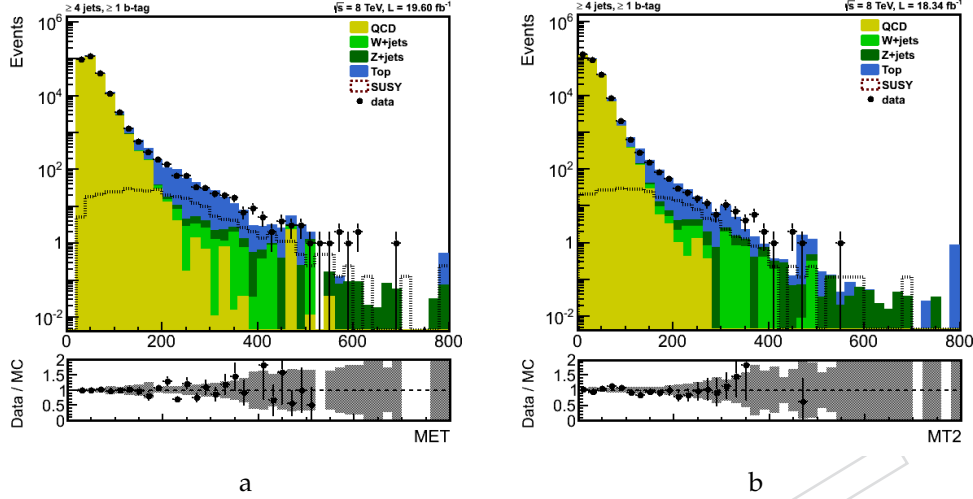


Figure 4: The distribution of  $E_T^{miss}$  (a) and  $M_{T2}$  (b) after smearing.

## 5.7 Preselections: The filters against the noises and the instrumental effects

- At least one good primary vertex, with  $\rho < 2$  cm and  $|z| < 24$  cm and  $N_{dof} > 4$  is requested.
- There are some cleaning cuts which are applied against instrumental effects, including those listed below:
  - HCAL noise We apply an isolation based HBHE noise filter [14]. Moreover, we veto HCAL laser events (laser fired accidentally) which produce very huge events. The events were rejected if they have  $> 25$  jets and  $HCALsumET > 10$  TeV.
  - Beam Halo Deposits in the CMS detector due to beam halo particles can be a source of fake  $E_T^{miss}$ . For this reason, we apply the recommended CSCtTightHaloId to filter events identified as beam halo. More information can be found here [15].
  - Jets failing Jet-ID. We reject events containing jets with  $p_T > 50$  GeV and  $|\eta| < 2.4$  that fail the Jet-ID. Such jets are mostly noise and mis-measurements, but also photons as the PF-JetID imposes a cut on the maximal EMF. small.
  - Other sources of noise. To check the events in the signal regions against instrumental noise, we also applied the following noise rejection filters:
    - Bad ECAL rechit recovery filter [16]
    - RA2 tracking failure filter [17]
    - Particle-based noise rejection [18]
    - ECAL TP filter [19]

## 5.8 MT2b Cuts

This section provides a review on the cuts which are started with to study the triggers. This set of cuts are those mainly used in the MT2b analysis [2]. Once the trigger is fixed, the optimized set of selection cuts which are used in the main stream of the current analysis will be described in detail in Section 7.

- The preselection cuts which was outlined in Section 5.7.
- At least 4 jets with  $p_T > 40$  GeV and  $|\eta| < 2.4$  are required which are asked to pass loose pf-jet id cuts.
- The leading jet- $p_T$  should be greater than 150 GeV.
- It is also required that all jets with  $p_T > 50$  GeV pass loose pf-jet id cuts. Events with non-identified high  $p_T$  jets are discarded.
- At least one b-quark jet is requested with  $p_T > 20$  GeV within the tracker acceptance, which is tagged by the Simple Secondary Vertex algorithm with a tight working point.
- The difference between  $E_T^{miss}$  and the vectorial  $p_T$  sum of the selected jets, hereafter referred to as VectorSumPt (VSPT), should be below 70 GeV.
- $E_T^{miss}$  is required to be greater than 30 GeV.
- The minimum  $\Delta\phi$  between  $E_T^{miss}$  and the four leading jets, hereafter referred to as  $\Delta\phi_4^{min}$ , should be greater than 0.3. There is no requirement on the id or  $p_T$  of the jets when looking for the minimum azimuthal angle between  $E_T^{miss}$  and jets.
- A cut on  $M_{T2} > 125$  GeV is applied.
- Leptons, being either electrons or muons, are vetoed.

## 6 Trigger

### 6.1 Trigger Study

To have the best reach, two sets of triggers are compared. Their names and run ranges are shown in Table 3. The corresponding prescaled triggers which are used to find the trigger plateau are also shown.

To take into account the statistics (the peak of the selected events by the un-prescaled trigger figure 5 (middle)), we look at the efficiencies bin-by-bin and distribution of efficiencies (un-prescaled divided by the prescaled, figure 5 (right)) with different HT cuts are weighted according to the statistics of the un-prescaled histogram. The cut that gives the mean value greater than 95% is chosen as the offline cut on the trigger parameter. An example of this method for different cuts on the HT for HLT\_PFHT650\_vX is shown in figure 6.

The result of this method is  $HT > 700$  GeV, but we use 725 GeV conservatively. For the multijet triggers, same method is used and depending on the number of jets different cuts on the  $P_T$  of the jets are found. The result is summarized here:

- HLT\_SixJet45\_vX, 6 jets with  $p_T > 65$  GeV/ $c$  or 7 jets with  $p_T > 55$  GeV/ $c$
- HLT\_QuadJet80\_vX, 4 jets with  $p_T > 100$  GeV/ $c$  or 5 jets with  $p_T > 85$  GeV/ $c$

As another possibility, one can think of decreasing the number of jets and increasing the  $p_T$  threshold, but it does not reach the plateau and is excluded from the list. Asking for 7 jets means that we rely on ISR/FSR and it is not safe from the systematic point of view, the increase

Table 3: On line triggers, their references and run ranges. A logical OR between SixJet and QuadJet is used.

HT		
Trigger Path	Prescaled Trigger	Run Range
HLT_PFHT650_v5	HLT_PFHT350_v3	190650-190750
HLT_PFHT650_v6	HLT_PFHT350_v4	191000-191400
HLT_PFHT650_v7	HLT_PFHT350_v5	191500-193750
HLT_PFHT650_v8	HLT_PFHT350_v6	193750-196030
HLT_PFHT650_v9	HLT_PFHT350_v7	196046-196531
MultiJet		
HLT_SixJet45_v1	HLT_SixJet35_v1	190456 - 190738
HLT_SixJet45_v2	HLT_SixJet35_v2	190782 - 196027
HLT_SixJet45_v3	HLT_SixJet35_v3	196046 - 196531
HLT_QuadJet80_v1	HLT_QuadJet70_v1	190456 - 190738
HLT_QuadJet80_v2	HLT_QuadJet70_v2	190782 - 196027
HLT_QuadJet80_v3	HLT_QuadJet70_v3	196046 - 196531

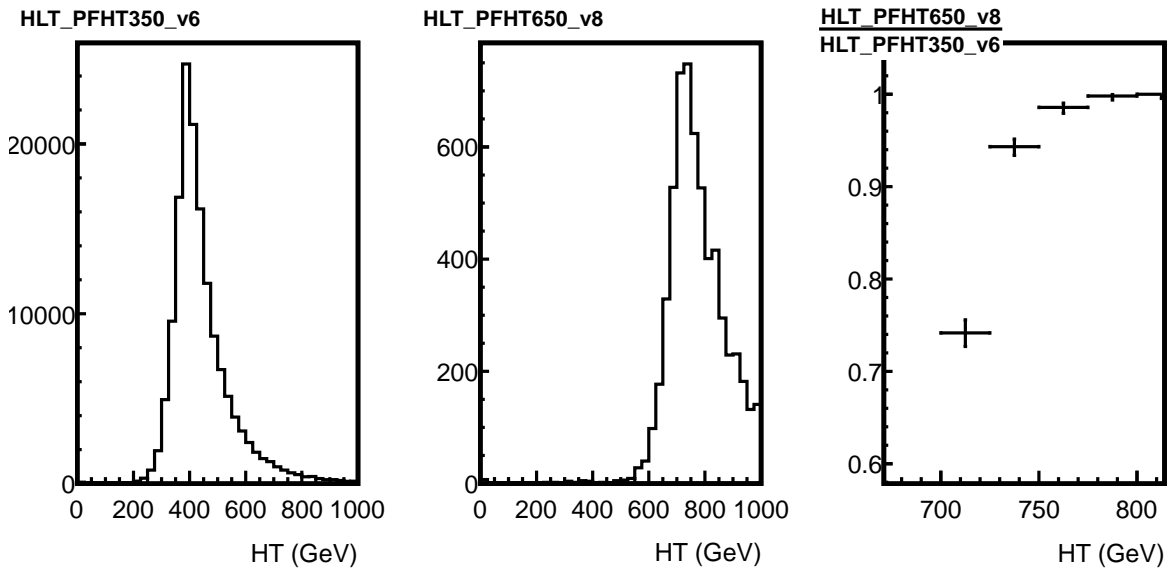


Figure 5: The prescaled (left) and un-prescaled (middle) HT triggers. The right plot shows the ratio of the two previous histograms zoomed in the interesting part.

239 of the yield due to adding this cut is negligible (606 events in data increases to 615), so this part  
 240 is dropped from the offline cuts.

241 After increasing the statistics to  $19.6 \text{ fb}^{-1}$ , the following online triggers with the same offline  
 triggers are added to the analysis.

Table 4: On line triggers and run ranges. A logical OR between SixJet and QuadJet is used.

HLT_SixJet45_v4	198022 - 199608
HLT_SixJet45_v6	199698 - 209151
HLT_QuadJet80_v4	198022 - 199608
HLT_QuadJet80_v6	199698 - 209151

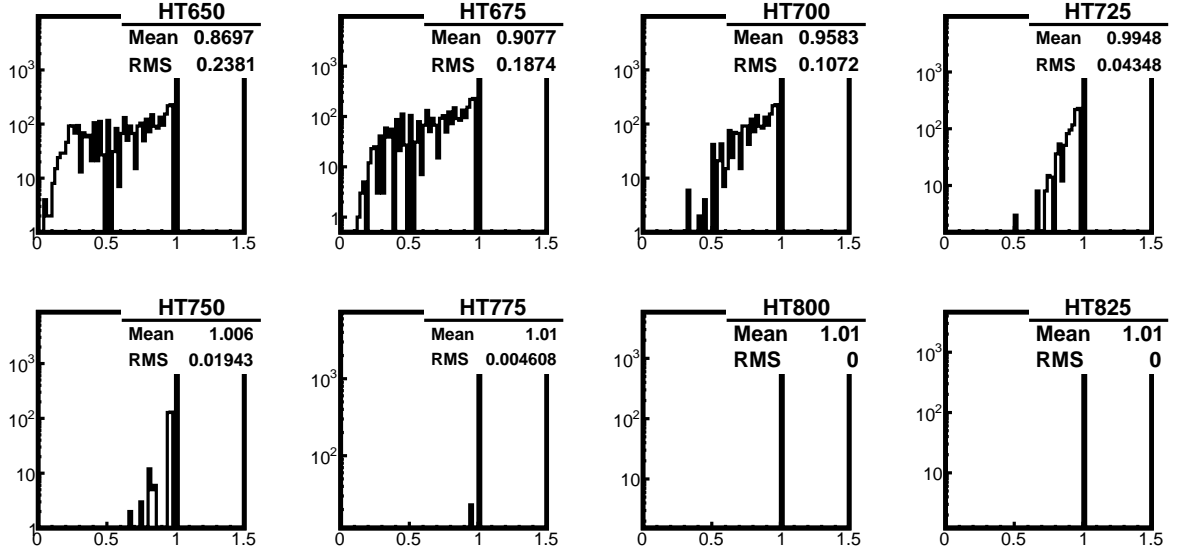


Figure 6: The weighted mean of the efficiencies in figure 5(right) for different cuts on HT. HT > 700 GeV gives 95% efficiency.

## 6.2 Trigger Selection

To investigate the efficiency of different trigger sets the SMST2tt sample is used. The selection cuts described in section 5.8 are applied on top of the trigger selection. The ratio of the signal events passing all the cuts is shown for two different sets of triggers as a function of  $\tilde{t}$  mass and  $\tilde{\chi}_1^0$  mass in figure 7. Although the signal efficiency is  $\sim 4$  times larger when the HT trigger is used, MC studies show that the number of remaining backgrounds are so larger that the multi-jet trigger is more powerful to exclude. The estimated exclusion power of both triggers are compared in figure 8.

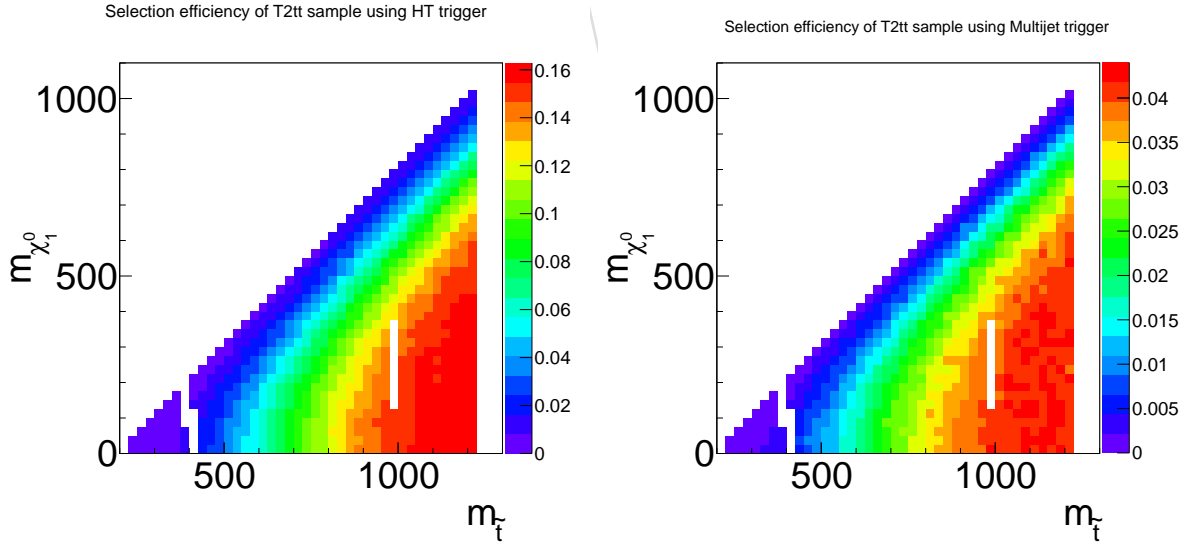


Figure 7: The efficiency of different trigger sets (Left : HT trigger, Right : Multijet trigger) for the SMST2tt sample. The results are shown as a function of the  $\tilde{t}$  mass and  $\tilde{\chi}_1^0$  mass.

The multi-jet trigger is used in the rest of the analysis.

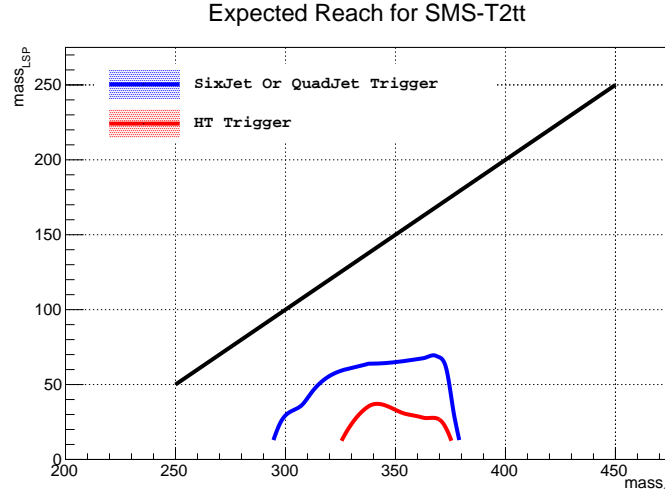


Figure 8: The estimated exclusion power for two different sets of triggers. The multijet trigger is used in this analysis.

## 7 Selection Cuts

In order to select signal events and suppress SM backgrounds, a set of cuts which are listed below, is applied.

- The preselection cuts which was outlined in Section 5.7.
- Offline trigger cuts mentioned in Section 6.1. They ask for at least 4 jets with  $|\eta| < 2.4$
- Among these set of jets, first and second leading jets are needed to have a  $p_T$  greater than 100 GeV and 60 GeV, respectively.
- It is also required that each jet with  $p_T > 50$  GeV pass loose pf-jet id cuts.
- At least one b-quark jet is requested with  $p_T > 20$  GeV within the tracker acceptance, which is tagged by the Combined Secondary Vertex (CSV) algorithm with a tight working point.
- The difference between  $E_T^{miss}$  and the vectorial  $p_T$  sum of the selected jets, electrons and muons should be below 70 GeV.
- The  $\Delta\phi_4^{min}$  of the four leading jets should be greater than 0.3. There is no requirement on the id or  $p_T$  of the jets when looking for the minimum azimuthal angle between  $E_T^{miss}$  and jets.
- $E_T^{miss}$  is required to be greater than 30 GeV.
- Leptons, being either electrons or muons, are vetoed.
- A cut on  $M_{T2} > 125$  GeV is applied.

The effect of the selection cuts on different backgrounds is shown in Table 5

## 8 Search Strategy

The  $M_{T2}$  distribution in Figure 9 is used as a variable to search for SUSY. With comparing the data and MC in the QCD dominated region ( $M_{T2} < 60$  GeV), a flat scale factor of 1.474 is found for the QCD samples to have a good agreement between data and MC. To increase the power of the analysis, a multibinning approach is used. We select 4 bins in  $M_{T2}$  with the edges

Cut	QCD	W+jets	Z+jets	Top	Total
Trigger	5404833	8155.97	2490.53	62465.45	$5477944.96 \pm 24563.93$
Jet Id	5398500.97	8155.97	2490.53	62458.38	$5471605.85 \pm 24530.03$
Lepton Veto	5395465.53	5328.26	1950.00	53172.66	$5455916.45 \pm 24529.43$
BJet	548011.88	361.80	163.32	35523.88	$584060.88 \pm 10321.31$
$\Delta\phi_4^{min} > 0.3$	297576.19	221.37	109.44	19851.70	$317758.70 \pm 8400.29$
$ E_T^{miss} - MHT  < 70$	253056.18	192.73	98.83	16860.00	$270207.73 \pm 8276.51$
$M_{T2} > 125$	113.19	30.87	22.74	500.33	$667.14 \pm 28.28$

Table 5: Event yields after applying different cuts. "Trigger" contains all of the preselections and offline trigger cuts.

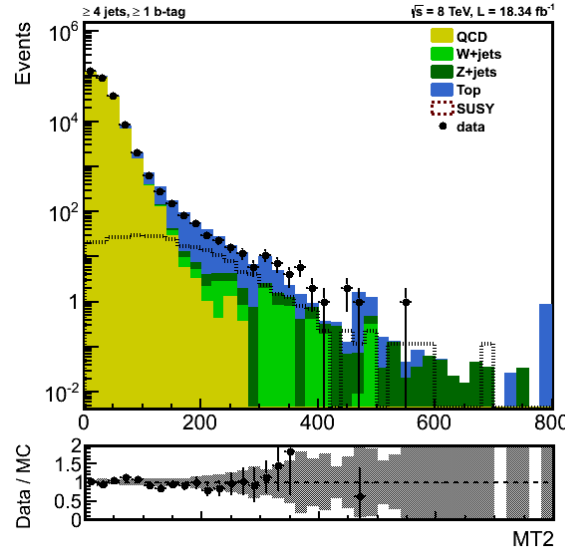


Figure 9:  $M_{T2}$  distribution after applying the full selection cuts.

of 125, 150, 200, 250 and infinity. Every  $M_{T2}$  bin is divided to two bins with number of the reconstructed top quarks equal to or greater than 0. This leads to 8 bins in total. In this round of analysis, we try to emphasize the complementary role of this analysis for the common cut and count hadronic search for the direct stop production. Since this analysis does not use the  $E_T^{miss}$  explicitly, it is more sensitive to the small mass differences between stop and LSP.

## 9 Backgrounds

In this section, data driven methods are proposed and applied to estimate the contribution of the main background processes. Most of the methods are similar to what were used in the  $M_{T2}$  analysis [2] with some minor changes which are explained here.

### 9.1 Data-driven background estimation of QCD

Due to inadequate statistics of QCD Monte-Carlo samples and complicated nature of this background, we use a data driven method to estimate its rate in the tail of the  $M_{T2}$  distribution, while the simulation shows that it is small.

We follow the method, fully discussed and applied by the  $M_{T2}$  and  $M_{T2b}$  groups [2], but the parameters are finely tuned to the conditions of our analysis. The method indeed relies on dif-



ferent distributions of QCD and SUSY-like events in the plane of  $M_{T2}$  and  $\Delta\phi_4^{min}$ , the azimuthal difference between the  $E_T^{miss}$  vector and the closest selected jet.

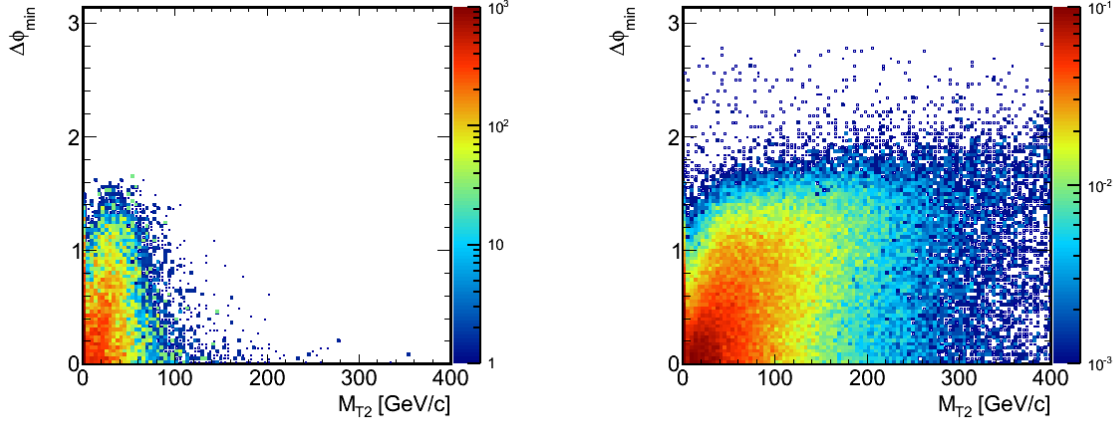


Figure 10: Distribution of  $\Delta\phi_4^{min}$  versus  $M_{T2}$  for (left) QCD and (right) SUSY-like (SMS) simulated events. QCD events are populated in the low  $\Delta\phi_4^{min}$  and  $M_{T2}$  region, while SUSY events spread over the plane.

Figure 10 shows such distributions for QCD (left) and SMS samples (right). Unlike the broad spread of SMS events in this plane, QCD events are densely populated in the low  $\Delta\phi_4^{min}$  and  $M_{T2}$  region. Due to the strong correlation between the two variables of  $\Delta\phi_4^{min}$  and  $M_{T2}$ , the usual ABCD method is inefficient, whereas a factorization method [2] is still applicable. The method works based on the ratio of  $r(M_{T2}) = N(\Delta\phi_4^{min} \geq 0.3) / N(\Delta\phi_4^{min} \leq 0.2)$  as a function of  $M_{T2}$  for QCD events. Figure 11 shows the ratio  $r(M_{T2})$  in the QCD simulation. It indicates an exponentially descending behavior in the region of  $M_{T2} > 50$  GeV (the lower bins of  $M_{T2}$  could be biased by the minimal cut on  $E_T^{miss}$ ). Hence, we characterize such specification of the QCD events by the model of

$$r(M_{T2}) = \frac{N(\Delta\phi_4^{min} \geq 0.3)}{N(\Delta\phi_4^{min} \leq 0.2)} = e^{a-b \cdot M_{T2}} + c \quad (7)$$

where  $a$  and  $b$  parameters indicate respectively the slope and the intercept of the straight line in the logarithmic scale. Ratio  $r(M_{T2})$  tends towards constant value,  $c$ , at large values of  $M_{T2}$ . The red curve in Figure 11 shows the fit of model (Equation 7) to the QCD simulation and Table 6 presents the value of parameters as a result of the fit in the range of  $M_{T2} > 60$  GeV (the first column).

Parameter	$M_{T2} > 60$ GeV	$60 < M_{T2} < 80$ GeV
$a$	$2.07 \pm 0.14$	$2.22 \pm 0.46$
$b$ (GeV $^{-1}$ )	$0.0226 \pm 0.0019$	$0.0245 \pm 0.0062$
$c$	$0.0504 \pm 0.0039$	-

Table 6: The result of the two different parametrizations for ratio  $r(M_{T2})$  in QCD simulated events.

In real data, to have a pure QCD sample with the minimal contamination from non-QCD backgrounds, we have to concentrate on the region of low  $M_{T2}$  ( $60 < M_{T2} < 80$  GeV). The fit of ratio  $r(M_{T2})$  on this short range of  $M_{T2}$  can be reasonably described as a straight line in



the logarithmic scale. Thus, it is not able to give parameter  $c$ . The green curve in Figure 11 shows the linear fit and the second column of Table 6 presents the relevant parameters,  $a$  and  $b$ . As seen from Figure 11, both fits (green and red) are in a very good agreement at low  $M_{T2}$ , while the second fit (the green straight line), called optimistic parameterization, gives the lower values for ratio  $r(M_{T2})$  at high  $M_{T2}$ . Hence, a realistic model needs also the parameter  $c$  to parameterize the ratio  $r(M_{T2})$  in the entire range of  $M_{T2}$ . we conservatively take the parameter  $c$  from the straight line at  $M_{T2} = 200$  GeV. The blue curve of Figure 11 represents such a fit, namely pessimistic parameterization.

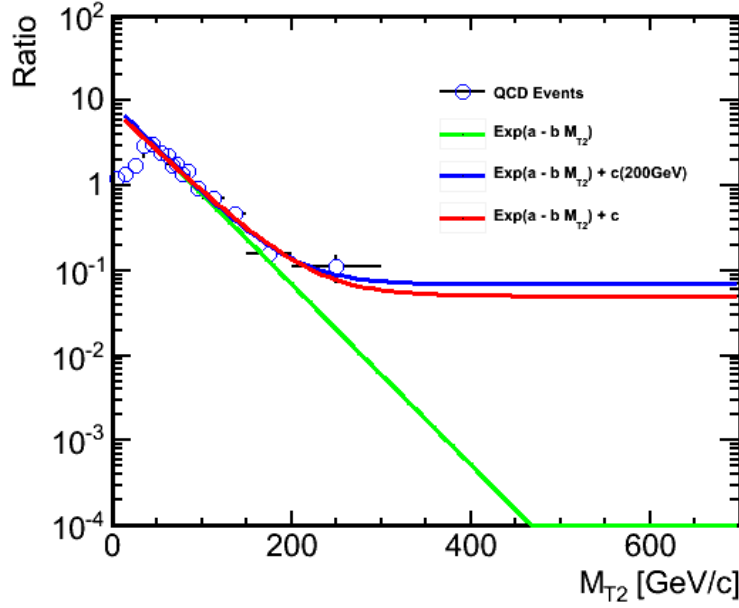


Figure 11: Three different fits of ratio  $r(M_{T2})$  in QCD simulated events. The red curve is an exponential function plus a constant. It uses the entire range of  $M_{T2} > 60$  GeV for parametrization (fully-MC) of ratio  $r(M_{T2})$ . The green curve is just an exponential function and uses the range of  $60 < M_{T2} < 80$  GeV for parameterization (optimistic). The blue curve is also an exponential function plus a constant, however it uses the range of  $60 < M_{T2} < 80$  GeV for parameterization (pessimistic).

Figure 12 depicts both parameterizations (optimistic and pessimistic by green and blue curves respectively) as a consequence of employing the method in the cleaned data. The non-QCD contaminations, taken from the Monte-Carlo simulation, are subtracted from data before calculating the parameters. Table 7 presents the parameters  $a$  and  $b$  extracted from the fit. These data-driven parameters eventually fulfill the functional form of ratio  $r(M_{T2})$ .

Parameter	$60 < M_{T2} < 80$ GeV
$a$	$2.39 \pm 0.20$
$b$ (GeV <sup>-1</sup> )	$0.0247 \pm 0.003$

Table 7: The parametrization results for ratio  $r(M_{T2})$  in real data (non-QCD events are subtracted, using simulation).

In the last step of procedure, we apply the ratio  $r(M_{T2})$  to the observed cleaned data in the QCD control region (high  $M_{T2}$ , low  $\Delta\phi_4^{min}$ ) to estimate the number of QCD events in the signal region (high  $M_{T2}$ , high  $\Delta\phi_4^{min}$ ). Figure 13 shows the  $M_{T2}$  distribution of QCD truth observed

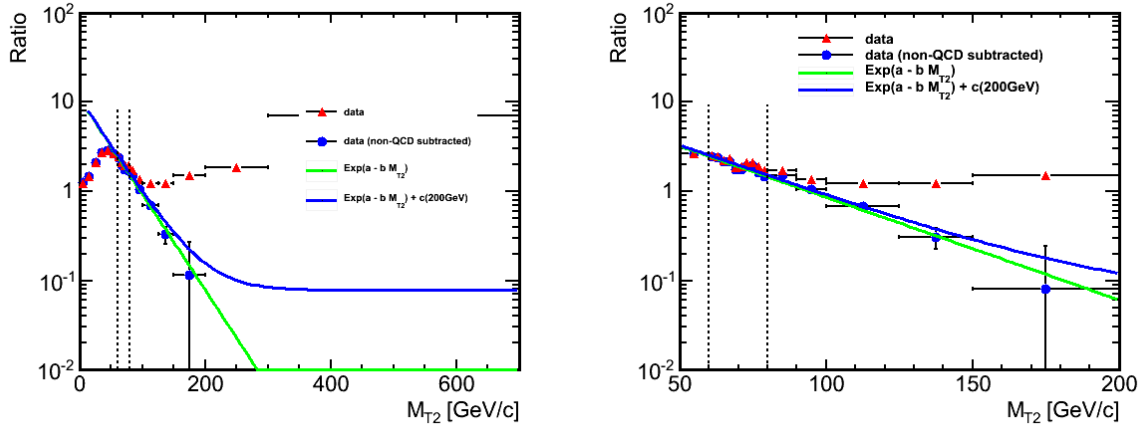


Figure 12: Fits of ratio  $r(M_{T2})$  in the non-QCD subtracted data. The green and blue curves are related to optimistic and pessimistic parameterization respectively. The right plot is a focus on the desired range of  $M_{T2}$  for the parametrizations.

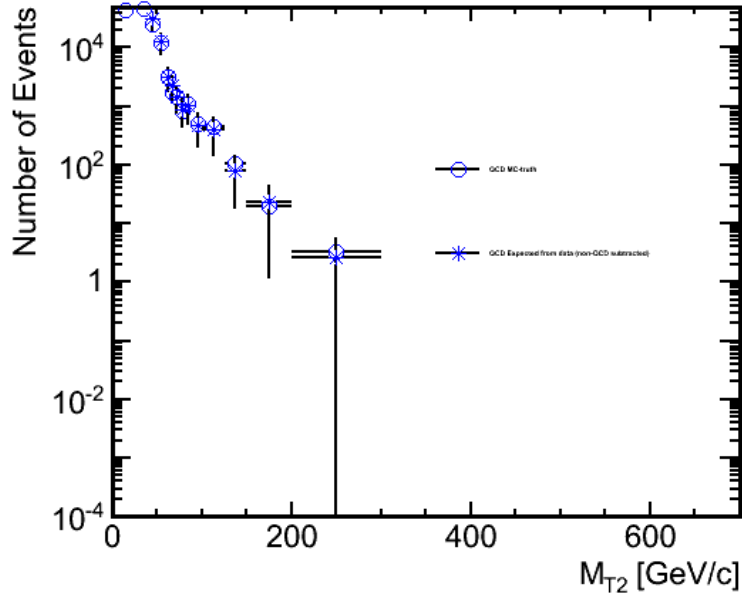


Figure 13: QCD MC-truth and data-driven prediction for the distribution of  $M_{T2}$ .

events and the expected distribution from data (non-QCD subtracted). Furthermore, Table 8 compares the estimated with observed QCD events for several bins of  $M_{T2}$ . In addition to the statistical uncertainties, the predicted numbers incorporate the systematic ones, coming from the fit range. Indeed, the standard deviation of a 10% fluctuation at the boundaries of the fit range, ( $60 < M_{T2} < 80$ ) GeV, induces the systematic uncertainties reported in Table 8. Considering the uncertainties, the method prediction is in good agreement with the QCD truth.

Data-driven estimation is consistent with MC truth within the uncertainties.

$M_{T2}$ bins	MC-truth	Data-prediction
[125, 150)	$102.8 \pm 23.22$	$80.67 \pm 63.56$
[150, 200)	$19.12 \pm 2.97$	$23.57 \pm 22.43$
[200, 300)	$3.34 \pm 1.24$	$2.59 \pm 3.22$
[300, $\infty$ )	$0.00 \pm 0.00$	$0.00 \pm 0.21$

Table 8: QCD MC-truth and data-driven prediction for the several bins of  $M_{T2}$ .

## 9.2 Data-Driven Estimation of Lost Lepton from W+jets and Top

After applying the selection cuts, described in detail in Section 7, the background events are dominated by  $t\bar{t}$  events. Among all decay channels of top pair system, it is mainly the semi-leptonic decay which contributes to the background. This can be understood because genuine neutrino is produced in the semi-leptonic decay of top pair system,  $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow b\bar{b}l\nu_{lj}$ , which can pass the  $M_{T2}$  cut while the full-hadronic decay products,  $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow b\bar{b}jjjj$ , do not contain any neutrino. This section describes a method to estimate the backgrounds from the leptonic decay of W bosons, either from prompt production in W+jets events or from W bosons produced in single top and top pair events, shown as  $t(\bar{t})$  for simplicity. The lepton is considered to be electron or muon.

Although the leptons are vetoed in the main analysis, there are still some background from  $W \rightarrow l\nu_l$ , referred to as lost lepton background events, contributing to the full-hadronic analysis. This is due to the acceptance cuts or inefficiencies in the lepton isolation and identification criteria.

In order to estimate the backgrounds due to the lost lepton events, all selection cuts are applied except lepton veto which is inverted. The distribution of the  $p_T$  of the leptons in the events with exactly one lepton, being either electron or muon, are shown in Figure 14, where it can be seen that the number of MC events are greater than the observed number of data events.

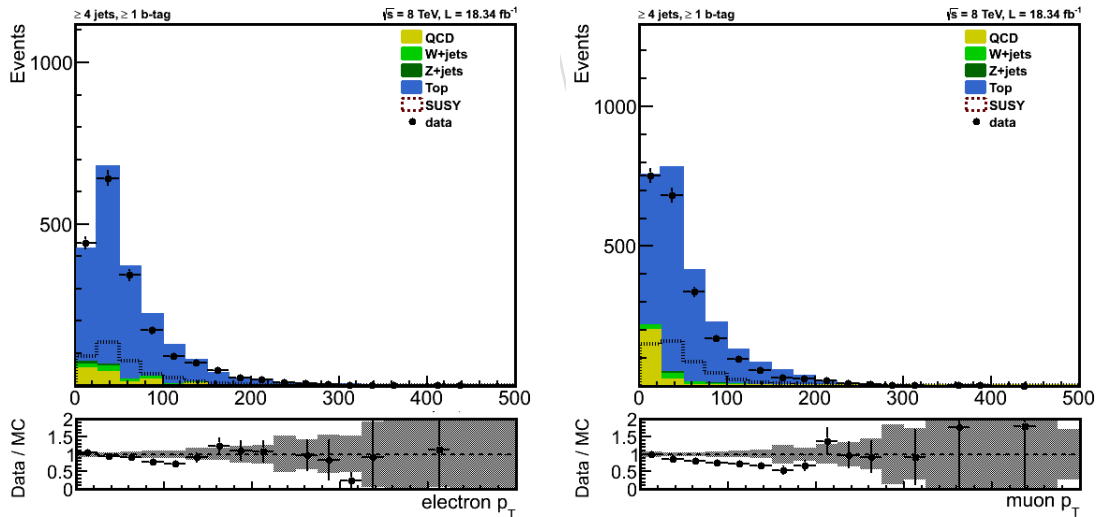


Figure 14: Left: The  $p_T$  distribution of the electrons in the events with one electron passing all selection cuts but the  $\Delta\phi_4^{min}$  cut. The reason for this is stated in the text. No  $M_{T2}$  cut is applied in order to have more statistics. Right: The same plot for muons.

In order to increase the data statistics, the cut on the  $\Delta\phi_4^{min}$  is relaxed. This cut was introduced in the main analysis to suppress the QCD background events. Now that the lepton veto is

reversed and exactly one lepton is required, the QCD events are still under control. Hence relaxing the  $\Delta\phi_4^{min}$  cut would not be harmful. The only thing which should be taken into account is the efficiency of this cut, called as  $f$ , which is explained in the following.

The contribution of the lost lepton background events passing the lepton veto, shown as  $N_l^{pass}$ , is estimated with the following formula

$$\begin{aligned} N_l^{pass} &= (N_l^{reco} - N_l^{bg}) \frac{1}{\varepsilon_l} - (N_l^{reco} - N_l^{bg}) \\ &= (N_l^{reco} - N_l^{bg}) \frac{1 - \varepsilon_l}{\varepsilon_l}, \end{aligned} \quad (8)$$

where  $N_l^{reco}$  refers to the number of data events with all selection cuts except the lepton veto, which is replaced by asking for one lepton. For this set of cuts, the number of background events from processes other than  $W \rightarrow l\nu_l$  is represented by  $N_l^{bg}$  and is taken from MC. The  $\varepsilon_l$  contains the efficiency for a generated  $W \rightarrow l\nu_l$  passing all selection cuts but the inverted lepton veto to have a lepton reconstructed. Here, the electron and muon efficiencies are obtained from both  $t\bar{t}$  and  $W + jets$  events and a relative contribution is used in the above formula. It should also be noted that, at the generator level, those  $t\bar{t}$  events containing a tau lepton decaying hadronically are vetoed since these kind of events are considered when backgrounds from tau are estimated.

In order to reduce the signal contamination in the leptonic signal region, a cut on the transverse mass of the lepton,  $m_T$ , is applied which is defined as

$$m_T = \sqrt{2p_T(e, \mu)E_T^{miss}(1 - \cos(\Delta\Phi))} < 100 \text{ GeV},$$

where  $\Delta\Phi$  is the angle between lepton- $p_T$  and  $E_T^{miss}$  in the transverse plane. In the  $W \rightarrow l\nu_l$  events, the  $m_T$  cut represents the transverse mass of the  $W$  bosons whose distribution drops at 80 GeV. Hence the leptonic signal events are not affected by this cut, while the contamination from SUSY events are strongly suppressed. The distribution of the  $m_T$  of either electrons or muons in the events with exactly one electron and one muon respectively, are shown in Figure 15. In this analysis, it is found that, e.g. for electron  $m_T$  distribution, the  $S/B$  decreases from 1.03% to 0.60% when  $m_T < 100$  GeV cut is introduced. In the rest of this section, in addition to all selection cuts, events are required to pass  $m_T < 100$  GeV cut.

The fraction of events with all selection cuts with respect to the events with all selection cuts but the  $\Delta\phi_4^{min}$  are shown in Figure 16 for data and MC. Since in the signal region, defined as region with  $M_{T2} > 125$  GeV, the ratios become flat; one can fit the ratios with a straight line. For both electrons and muons, the MC ratio is fitted and the fitted parameters  $f$  are quoted in Table 9.

	electrons	muons
Fit value for the ratio $f$	$0.811 \pm 0.028$	$0.804 \pm 0.024$

Table 9: Fit values  $f$  obtained from the MC ratios for electrons and muons.

The results of estimation of the lost lepton background events from data are summarized in Table 10. It should be mentioned that, the number of data events with one lepton selection and its corresponding background events are obtained from the relaxed cut selection, where  $\Delta\phi_4^{min}$  is dropped. Therefore the prediction is corrected back by multiplying the event yield with the

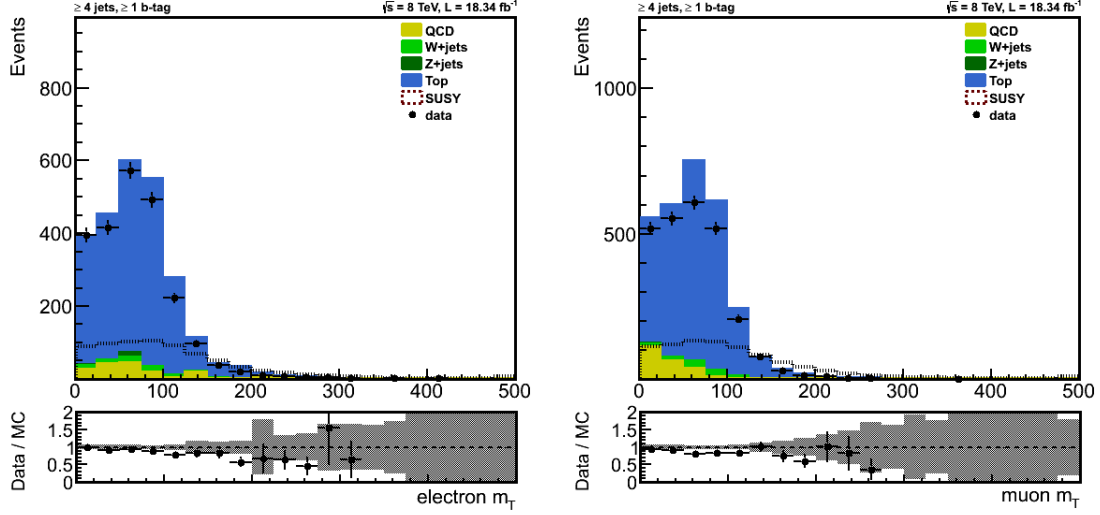


Figure 15: Left: The  $m_T$  distribution of the electrons for the events with one electron passing all selection cuts but the  $\Delta\phi_4^{min}$  cut. The reason for this is stated in the text. No  $M_{T2}$  cut is applied in order to have more statistics. Right: The same plot for muons.

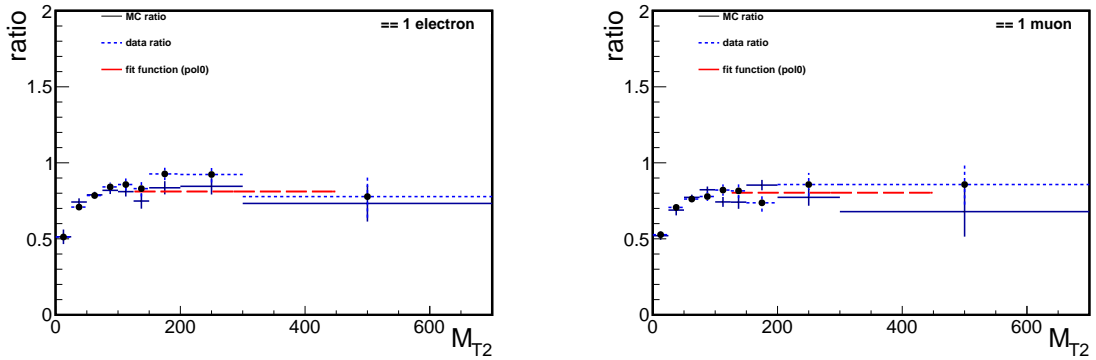


Figure 16: Left: Ratio between events with one electron passing all selection cuts versus events with one electron passing all selection cuts but the  $\Delta\phi_4^{min}$  for data (blue) and MC (black). The fit line for the MC ratio over all  $M_{T2}$  signal bins is drawn in red. Right: The same plot for events with one muon.

372 fitted ratio value  $f$ .

	$N^{reco}$	$N^{bg}$	$R_{LL}$	$N^{pass}$ MC-Truth	$N^{pass}$ data-prediction
electrons	129	20.78	$1.30 \pm 0.17$	$189.32 \pm 10.86$	$139.63 \pm 14.72(stat) \pm 33.48(sys)$
muons	150	25.42	$0.73 \pm 0.13$	$133.08 \pm 9.16$	$91.29 \pm 8.97(stat) \pm 24.97(sys)$

Table 10: Data-Driven Estimation of Lost Lepton from  $W$ +jets and  $t(\bar{t})$  for electrons and muons. The lost lepton ratio  $R_{LL}$  is given by  $f^{\frac{1-\varepsilon_l}{\varepsilon_l}}$ .

373

374 It should be noted that for the systematic uncertainty, two possible sources are taken into ac-  
 375 count. The first one is a systematic uncertainty of 100% on the number of backgrounds. The  
 376 second one is a systematic uncertainty of 5%, considered when calculating the efficiencies  $\varepsilon_l$   
 377 from MC, to account for possible difference between data and simulation. Considering the un-  
 378 certainties, the data-driven estimation for the lost lepton channel is slightly less than MC truth.  
 379 We use MC estimation in limit calculation to be on the safe side.

### 380 9.3 Estimation of the Tau Leptons

381 Tau leptons can decay hadronically and appear as a thin jet and enter the hadronic searches.  
 382 To estimate the contamination from such events a method similar to what is used for the lost  
 383 lepton background is used here. The number of events with exactly one real tau is corrected by  
 384 accounting for the reconstruction and acceptance efficiencies. In the other words:

$$N_{W \rightarrow \tau \nu} = \frac{N_{\tau}^{reco} - N_{\tau}^{bg}}{\varepsilon_{\tau}}, \quad (9)$$

385 where  $N_{\tau}^{reco}$  is the number of events with one reconstructed tau,  $N_{\tau}^{bg}$  is the number of events  
 386 with a fake tau and  $\varepsilon_{\tau}$  denotes the probability for a generated  $W \rightarrow \tau \nu$ ,  $\tau \rightarrow had$  event passing  
 387 the selection cuts to have a reconstructed and identified tau. The efficiency  $\varepsilon_{\tau}$  is extracted from  
 388 simulation. In average,  $\varepsilon_{\tau}$  is found to be  $\sim 24\%$ . 5% systematic uncertainty is assigned to this  
 389 value to take into account the differences between data and MC. The transverse mass ( $m_T$ ) of  
 390 the system of the reconstructed tau and  $E_T^{miss}$  is forced to be less than  $100 \text{ GeV}/c^2$  to decrease  
 391 the signal contamination. The number of events with a fake tau,  $N_{\tau}^{bg}$  is found from the MC  
 392 simulation and a 50% systematic uncertainty is assigned to this value.

393 The  $M_{T2}$  distribution of events with all selection cuts which have an identified tau in the final  
 394 state is shown in Figure 17. Scale factors of the tau selection are not applied and it can be the  
 395 source of the discrepancies between data and MC. In Table 11 contribution of different samples  
 396 in the plot of Figure 17 is shown. It can be seen that the statistics in the signal region is poor.  
 397 To decrease the uncertainties of the predictions, the cut on  $\Delta\phi_4^{min}$  is relaxed. The last row of the  
 398 table shows the statistics after this relaxation. The scale factor to compensate this relaxation is  
 399 read from MC.

400 Table 12 shows the performance of the method on MC and data. The quoted uncertainties of the  
 401 predictions are statistical and systematical, respectively. Data-driven estimation is consistent  
 402 with MC truth within the uncertainties.

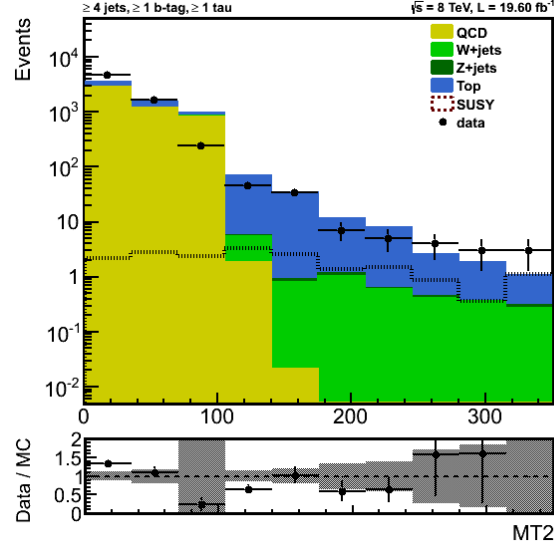


Figure 17:  $M_{T2}$  distribution for events with at least one tau in data and MC with all selection cuts.

$M_{T2}$ (GeV)	QCD	Z+Jets	W+Jets	Top	MC(sum)	data
full range	4544.26	8.68	27.99	1074.53	$5655.47 \pm 794.11$	6226.00
125 – $\infty$	0.00	0.18	5.37	57.90	$63.45 \pm 6.27$	61.00
cut on $\Delta\phi_4^{min}$ is relaxed						
125 – $\infty$	5.83	0.19	10.06	81.38	$97.46 \pm 8.57$	86.00

Table 11: MC and data event yields in full range and signal region. The last row shows the yields after relaxing the cut. The error on the total background is purely statistical.

$M_{T2}$ bin	MC Truth	Prediction in MC	Prediction in Data
125 – $\infty$	$224.51 \pm 11.86$	$227.89 \pm 27.72 \pm 32.45$	$217.07 \pm 34.53 \pm 31.13$

Table 12: Prediction of the tau contamination in the signal region in both data and MC.

## 9.4 Estimation of Invisible Z Background from Data Using W +jets Events

To estimate  $Z \rightarrow \nu\bar{\nu}$  background we use  $W \rightarrow \mu\nu$ +jets events. The kinematics of leptons as well as the jets are very similar in both Z+jets and W+jets processes. Besides, the larger cross-section of W+jets allows for a more precise estimation of  $Z \rightarrow \nu\bar{\nu}$ . This is a well studied method in various analyses within the CMS Collaboration (see e.g. [20–22]). To make the event kinematics compatible from the  $E_T^{miss}$  point of view, the  $p_T$  of muon is added to the one of neutrino in W+jets events. The  $M_{T2}$  variable and other quantities related to  $E_T^{miss}$  are recalculated accordingly. This estimation can be described as:

$$N_{Z \rightarrow \nu\bar{\nu}}(est) = N_{W(\mu\nu)} R^{MC} \frac{1}{\epsilon_{acc} \epsilon_{reco/iso}}. \quad (10)$$

where,

- $\epsilon_{acc}$  is the muon acceptance derived from MC.

- $\epsilon_{reco/iso}$  is the muon reconstruction and isolation efficiency, taken from data using the Tag&Probe method.

- $R^{MC}$  corrects kinematic, selection and cross-section differences between  $Z \rightarrow \nu\bar{\nu}$  and  $W \rightarrow \mu\nu$ +jets processes.
- $N_{W(\mu\nu)}$  is the number of selected  $W \rightarrow \mu\nu$ +jets events.

The selection is similar to the one of signal where the lepton veto is reduced to an electron veto. In addition we request for the presence of exactly one reconstructed muon passing all the quality and isolation cuts, with  $p_T > 10$  GeV and  $|\eta| < 2.4$ . The W-boson transverse mass (using default  $E_T^{miss}$ ) is required to be  $m_T < 100$  GeV in order to reduce other backgrounds and signal contaminations. To enrich the sample with W+jets and to reject  $t\bar{t}$  events, we veto events with at least one b-tagged jet where the medium working point of CSV b-tagging algorithm is applied on jets with  $p_T > 20$  and  $|\eta| < 2.4$ . The results of this selection for MC samples and data are summarized in Table 13. The distributions of muon  $p_T$ ,  $M_{T2}$  and  $m_T$  for this region are shown in Figures 18a, 18b and 18c and as it is seen there is a good agreement between data and MC in W enriched region.

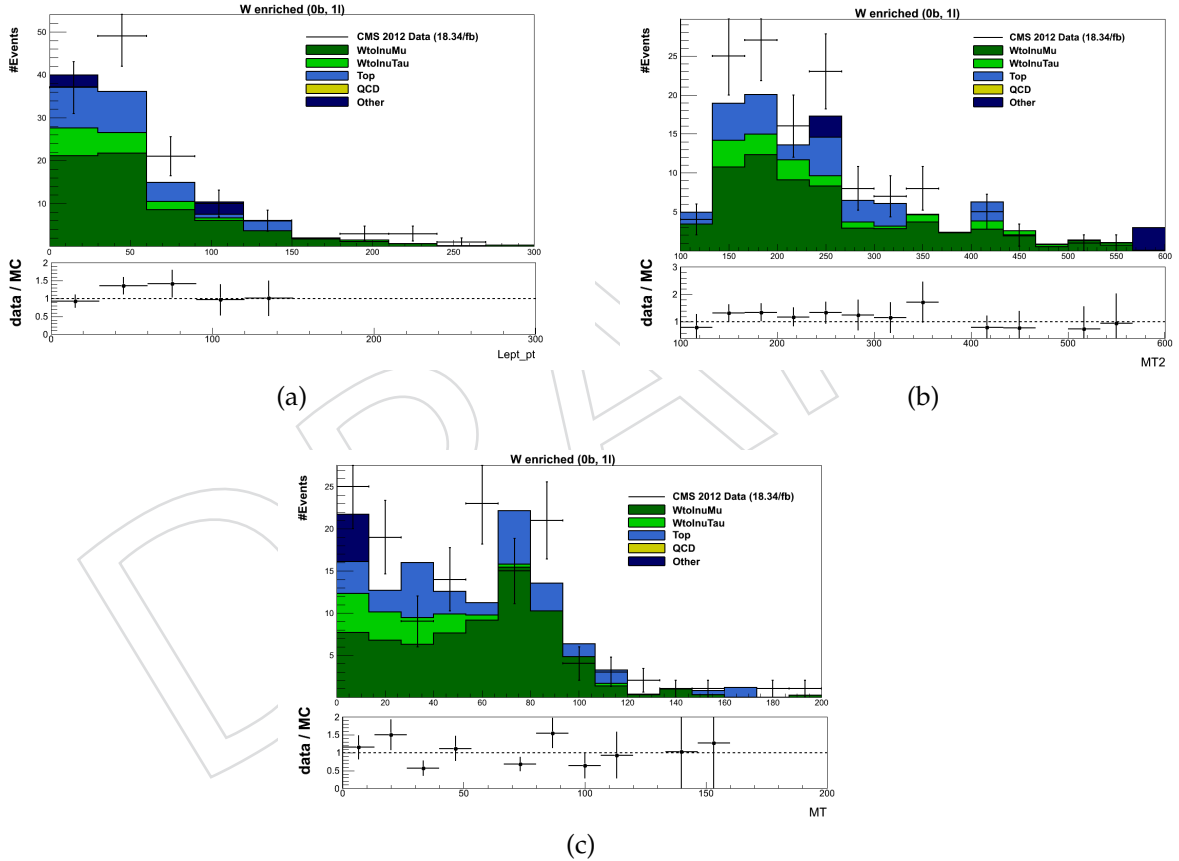


Figure 18: Muon  $p_T$ ,  $M_{T2}$  and  $m_T$  distributions for W-enriched region

#### 9.4.1 $t\bar{t}$ background estimation in W-enriched sample

Despite of b-tag veto some top-pair events remain in W enriched region. The contribution of  $t\bar{t}$  is estimated from data while for the rest of backgrounds we trust on simulation. The b-tag veto is relaxed and at least one b-jet is requested to obtain a sample enriched in  $t\bar{t}$  events. The selection results are shown in Table 14. To find out the number top events in b-tag veto region



	WtolnuMu	WtolnuTau	QCD	Zinv	Top	Other	MC	data
All events (jets $\geq$ X)	136.44	215.93	569.94	220.29	1300.06	172.98	2615.64 +- 61.07	2510.00
Analysis selection cuts	136.44	215.93	569.94	220.29	1300.06	172.98	2615.64 +- 61.07	2510.00
Electron Veto	136.44	200.65	569.33	220.18	1056.97	90.44	2274.01 +- 59.24	2192.00
Muon Selection	85.86	16.48	0.00	0.11	273.96	5.62	382.03 +- 15.87	329.00
$m_T < 100$ GeV	77.91	16.19	0.00	0.06	238.63	5.62	338.40 +- 14.96	293.00
b-jets Selection	65.19	14.31	0.00	0.03	27.55	5.62	112.69 +- 7.72	130.00
$M_{T2}$ 125 - 150 GeV	8.40	1.46	0.00	0.00	4.67	0.00	14.53 +- 2.52	12.00
$M_{T2}$ 150 - 200 GeV	18.04	4.62	0.00	0.03	6.73	0.00	29.43 +- 3.43	44.00
$M_{T2}$ 200 - 275 GeV	18.20	3.83	0.00	0.00	8.45	2.70	33.18 +- 4.45	41.00
$M_{T2}$ 275 - 375 GeV	9.17	2.10	0.00	0.00	4.14	0.00	15.41 +- 2.54	21.00
$M_{T2}$ 375 - 500 GeV	6.92	1.95	0.00	0.00	2.49	0.00	11.36 +- 2.18	7.00
$M_{T2} > 500$ GeV	3.82	0.34	0.00	0.00	1.06	0.00	8.15 +- 3.21	5.00

Table 13: Yields for the W-enriched selection

(W enriched), b-tagging (in)efficiency has to be considered. This process can be described as:

$$N_{top}(b - veto) = N_{top}(\geq 1b - tag) \frac{\epsilon(b - veto)}{\epsilon(\geq 1b - tag)}, \quad (11)$$

where the  $\epsilon(b - veto)$  and  $\epsilon(\geq 1b - tag)$  are the efficiencies of vetoing or selecting b-tagged events and are taken from simulation, thus corrected by the data-simulation scale factors given by the b-tag POG for the CSV (0.963  $\pm$  0.020) and CSV (0.947  $\pm$  0.025) working points, respectively [23]. As it is apparent from Table 14 there is a good agreement between data and MC in top-enriched region and also the muon  $p_T$ ,  $M_{T2}$  and  $m_T$  distributions are shown in Figures 19a, 19b and 19c, respectively.

	WtolnuMu	WtolnuTau	QCD	Zinv	Top	Other	MC	data
All events (jets $\geq$ X)	159.38	249.89	593.54	253.19	1109.83	198.00	2563.82 +- 61.63	2510.00
Analysis selection cuts	159.38	249.89	593.54	253.19	1109.83	198.00	2563.82 +- 61.63	2510.00
Electron Veto	159.38	232.49	593.04	253.06	903.14	102.42	2243.53 +- 60.09	2192.00
Muon Selection	100.27	19.43	0.00	0.12	233.72	5.48	359.01 +- 14.39	329.00
$m_T < 100$ GeV	91.11	19.05	0.00	0.06	203.23	5.48	318.94 +- 13.59	293.00
b-jets Selection	5.35	0.99	0.00	0.00	139.18	0.00	145.51 +- 9.57	119.00
$M_{T2}$ 125 - 150 GeV	0.59	0.39	0.00	0.00	15.93	0.00	16.91 +- 3.25	28.00
$M_{T2}$ 150 - 200 GeV	0.29	0.32	0.00	0.00	52.16	0.00	52.77 +- 5.94	43.00
$M_{T2}$ 200 - 275 GeV	1.37	0.00	0.00	0.00	43.02	0.00	44.39 +- 5.28	26.00
$M_{T2}$ 275 - 375 GeV	1.81	0.00	0.00	0.00	23.92	0.00	25.73 +- 3.88	16.00
$M_{T2}$ 375 - 500 GeV	0.63	0.00	0.00	0.00	2.05	0.00	2.68 +- 1.10	3.00
$M_{T2} > 500$ GeV	0.65	0.28	0.00	0.00	1.30	0.00	2.23 +- 1.03	2.00

Table 14: Yields for the top-enriched selection

#### 9.4.2 Z Estimation Results

After finding the number of  $t\bar{t}$  events in the b-tag veto (W-enriched) region, it is subtracted from the number of W's of this region, derived from data, to obtain the correct number of  $W \rightarrow \mu\nu$  events. Due to requesting one b-jet in the final state we need to have the number of W's in 1 b-tag region to be able to estimate the number of Z in this region. Therefore we must multiply the number of W's in b-tag veto region to  $\frac{\epsilon(1bW)}{\epsilon(0bW)}$  to reach the number of W's in 1 b-tag region. This ratio is coming from MC and it is considered b-tag scale factor.

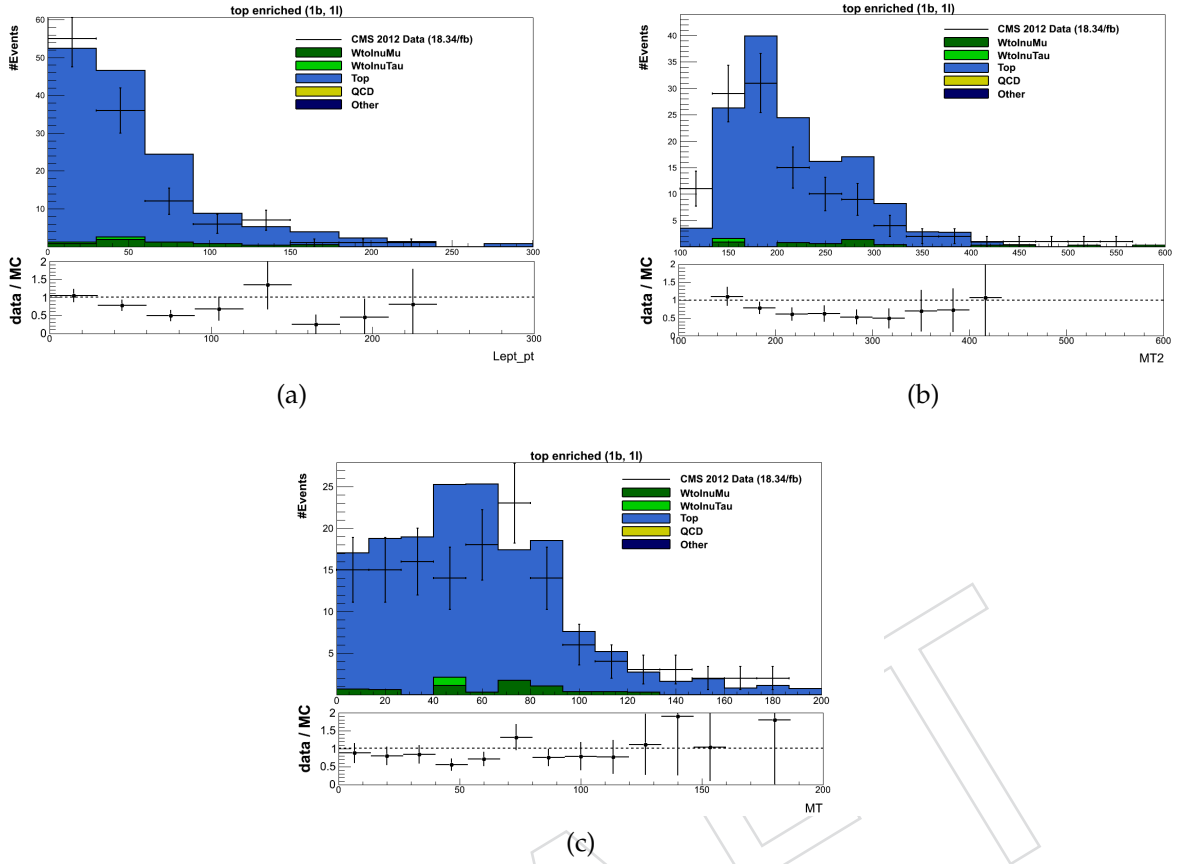


Figure 19: Muon  $p_T$ ,  $M_{T2}$  and  $m_T$  distributions for top-enriched region

### 9.4.3 Systematic uncertainties

The systematic uncertainty on  $Z \rightarrow \nu\nu$  estimation has contributions from different sources, as can be seen in Equation 10. There, the uncertainty on  $R^{MC}$  is taken from simulation where it includes the uncertainties due to the PDF set and the k-factor in Z and W bosons production rates. The uncertainty on the muon acceptance efficiency is derived from simulation, too. The muon selection efficiency ( $\epsilon_{reco/iso}$ ) as well as its uncertainty are data-driven, obtained from the Tag&Probe method. Another uncertainty in this estimation arises from the requirement of  $m_T < 100$  GeV which is estimated from simulation.

For the  $N_{W(\mu\nu)}$  in the analysis region with at least one b-tagged jet, the  $N_{W(\mu\nu)}$  estimation in b-tag veto region is corrected with the data-driven b-tagging and b-tag veto efficiencies. The uncertainties on these efficiencies are taken from data, accordingly.

Other than  $t\bar{t}$ , all backgrounds and their uncertainties are estimated from simulation in  $N_{W(\mu\nu)}$  calculation. The  $t\bar{t}$  contribution in W-enriched (b-tag veto) region is obtained using Equation 11. In this estimation, the uncertainties on b-tagging efficiencies are taken from data while the background uncertainties are derived from simulation.

The final estimations together with their uncertainties are summarized in Table 15.

Data-driven estimation is consistent with MC truth within the uncertainties.

	MC	Data Estimation
top (0b, 1l)	$27.55 \pm 4.38$	$18.99 \pm 4.44$ ( 1.74 (stat) $\pm 4.08$ (syst) )
W (0b, 1l)	$65.19 \pm 4.49$	$91.08 \pm 13.69$ ( 8.18 (stat) $\pm 10.99$ (syst))
Zinv (1b, 0l)	$21.28 \pm 0.79$	$19.92 \pm 7.96$ ( 1.79 (stat) $\pm 7.75$ (syst))

Table 15: Z-invisible Estimation

## 10 Statistical Interpretation of the results

Since no excess of data over the background prediction has been observed, we close our study with setting upper limits on the testing signals. This is conducted using a modified frequentist approach, namely CLs method [24]. In this method, the test statistic  $q_\mu$  [25] is a function of the profile likelihood-ratio,

$$q_\mu = -2 \ln \frac{\mathcal{L}(\text{data}; b + \mu s)}{\mathcal{L}(\text{data}; b + \hat{\mu} s)}, \quad (12)$$

where  $\hat{\mu}$  is the *signal strength modifier*  $\mu$  at the maximum point of the likelihood  $\mathcal{L}$ . Then CLs is given by the following probability-ratio,

$$CL_s = \frac{p(q_\mu \geq q_\mu^{obs} | b + \mu s)}{p(q_\mu \geq q_\mu^{obs} | b)}. \quad (13)$$

We compute CLs using a software package provided by the CMS Higgs PAG [26]. After incorporating systematic uncertainties, an observed CLs smaller than 0.05 for a signal strength of  $\mu = 1$ , excludes the given signal at 95% CL. Indeed, the package determines which signal strength  $\mu$  excludes the testing signal at 95% CL. Therefore all resulting  $\mu \leq 1$  define the excluded region in the parameter space of the given signal.

In this study, we analyze data in 8 different bins (multi-bin analysis) to utilize more information from the observed and the predicted distributions. The bins are defined in reconstructed top quark multiplicity, zero or more. In addition, events are categorized based on the  $M_{T2}$  values:  $125\text{GeV} \leq M_{T2} < 150\text{GeV}$ ,  $150\text{GeV} \leq M_{T2} < 200\text{GeV}$ ,  $200\text{GeV} \leq M_{T2} < 250\text{GeV}$ ,  $250\text{GeV} \leq M_{T2} < \infty$ .

To investigate the exclusion power of our search, we study the topology of direct stop pair production in Simplified Models [27], with  $\tilde{t} \rightarrow \tilde{\chi}_1^0 t$  (T2tt). Calculation of the expected exclusion limit shows that the research has potential to exclude a sizable region of the phase space, surrounded by the lines of  $m_{\tilde{t}} = 600\text{GeV}$  and  $m_{LSP} = 175\text{GeV}$  with an integrated luminosity of  $19.6\text{fb}^{-1}$ .

Figure 20 shows the expected upper limit on the cross section of the stop pair production in terms of Simplified Models. Furthermore, the figure shows the expected exclusion power considering 40% systematic uncertainties on signal and background rates which are predicted using Monte-Carlo simulations. The black dashed curve represents the expected reach by the common Cut&Count [28] search using  $E_T^{miss}$  trigger. As the figure shows our analysis (the blue solid curve) can be comparable with other analyses and it has the potential to be complementary to other analyses in some regions of the phase space.

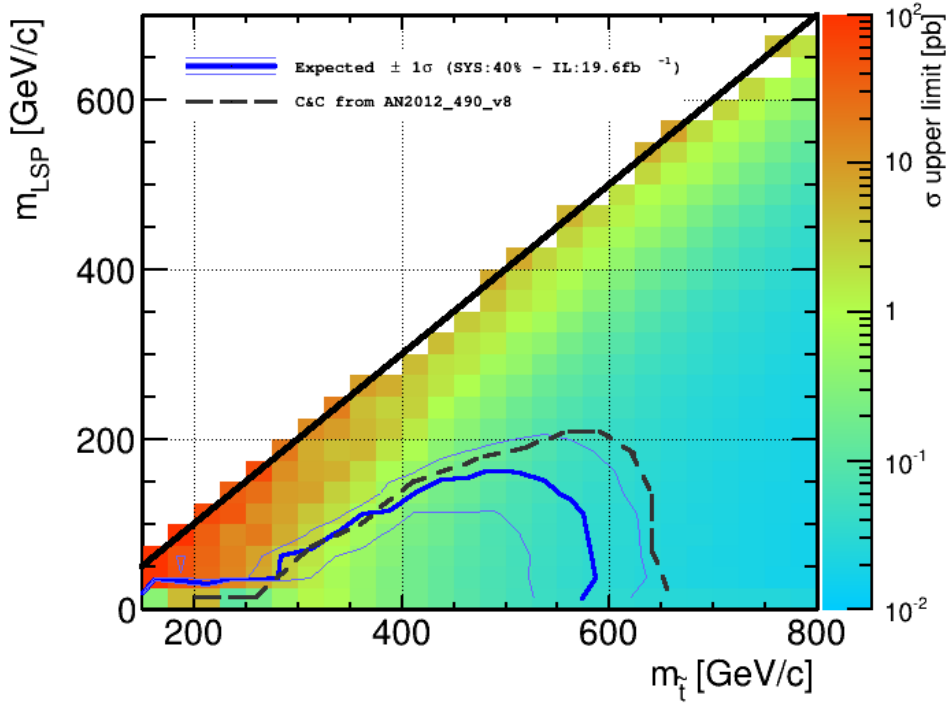


Figure 20: Expected exclusion power in terms of Simplified Models (T2tt-topology) with an integrated luminosity of  $19.6 \text{ fb}^{-1}$ . Backgrounds are predicted using Monte-Carlo simulations and a rough estimate of systematic uncertainties equal 40% is taken into account.

## 11 Conclusion

A hadronic search for direct stop production is presented using the  $M_{T2}$  variable. Data driven methods are used to estimate the main backgrounds. It is shown that the methods close properly on MC. Since the analysis uses a multijet trigger and  $M_{T2}$  does not depend explicitly on  $E_T^{\text{miss}}$ , it can be complementary to the common cut and count search for stop which uses  $E_T^{\text{miss}}$  trigger. It is shown that in the regions with low mass difference between  $m_{\tilde{t}}$  and  $m_{LSP}$  this analysis can have a comparable reach.

## 12 Acknowledgements

This analysis benefits highly from the computing resources and codes developed in ETHZurich. We appreciate their help and generosity. The method used for the top reconstruction was firstly introduced and implemented by L. Pape. We thank him for providing us the code and giving useful comments in some parts of the analysis. The authors would like to thank the conveners of the SUSY-TBT working group, Rick Cavanaugh and Juan Alcaraz Maestre, for their help and support. Batool Safarzadeh has read the analysis and provided some useful comments, we appreciate her help. The authors would like to thank the management and staff of the school of particles and accelerators of IPM, especially Prof. Arfaei for their help and support. Thanks to all of the members of the CMS collaboration for their outstanding results discussed partly here.

## References

- [1] S. P. Martin, “A Supersymmetry primer”, [arXiv:hep-ph/9709356](#).
- [2] Casal et al., “Search for supersymmetry in hadronic final states using MT2 based on 4.4/fb of CMS data at  $\sqrt{s} = 7$  TeV”, CMS Analysis Note 2012/040, CERN, (2012).
- [3] H. Bakhshian et al., “A Hadronic Search for Direct Stop with MT2 Variable”, CMS Analysis Note 2013/140, CERN, (2013).
- [4] UA1 Collaboration Collaboration, “Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at  $\sqrt{s} = 540$ -GeV”, *Phys.Lett.* **B122** (1983) 103–116, doi:10.1016/0370-2693(83)91177-2.
- [5] UA2 Collaboration Collaboration, “Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider”, *Phys.Lett.* **B122** (1983) 476–485, doi:10.1016/0370-2693(83)91605-2.
- [6] CDF Collaboration Collaboration, “Measurement of the W boson mass with the Collider Detector at Fermilab”, *Phys.Rev.* **D64** (2001) 052001, doi:10.1103/PhysRevD.64.052001, [arXiv:hep-ex/0007044](#).
- [7] D0 Collaboration Collaboration, “Improved W boson mass measurement with the DØ detector”, *Phys.Rev.* **D66** (2002) 012001, doi:10.1103/PhysRevD.66.012001, [arXiv:hep-ex/0204014](#).
- [8] C. Lester and D. Summers, “Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders”, *Phys.Lett.* **B463** (1999) 99–103, doi:10.1016/S0370-2693(99)00945-4, [arXiv:hep-ph/9906349](#).
- [9] A. Barr, C. Lester, and P. Stephens, “m(T2): The Truth behind the glamour”, *J.Phys.* **G29** (2003) 2343–2363, doi:10.1088/0954-3899/29/10/304, [arXiv:hep-ph/0304226](#).
- [10] W. S. Cho, K. Choi, Y. G. Kim, and C. B. Park, “Measuring superparticle masses at hadron collider using the transverse mass kink”, *JHEP* **0802** (2008) 035, doi:10.1088/1126-6708/2008/02/035, [arXiv:0711.4526](#).
- [11] CMS Collaboration, “Search for Supersymmetry in hadronic final states using  $M_{T2}$  with the CMS detector at  $\sqrt{s} = 7$  TeV”, CMS Physics Analysis Summary CMS-PAS-SUS-12-002, CERN, (2012).
- [12] T. Sjöstrand, S. Mrenna, and P. Skands, “PYTHIA 6.4 physics and manual”, *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, [arXiv:hep-ph/0603175](#).
- [13] “Egamma Cut-Based Identification”, [https://twiki.cern.ch/twiki/bin/view/CMS/EgammaCutBasedIdentification#Electron\\_ID\\_Working\\_Points](https://twiki.cern.ch/twiki/bin/view/CMS/EgammaCutBasedIdentification#Electron_ID_Working_Points).
- [14] J. Chou et al. <https://twiki.cern.ch/twiki/bin/view/CMS/HcalNoiseInfoLibrary>.
- [15] D. A. et al.
- [16] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFilters#RecovRecHitFilter>.

- [17] [https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFilters#Tracking\\_failure\\_filter](https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFilters#Tracking_failure_filter).
- [18] C. Bernet, L. Gauthier, , and L. Gouskos.
- [19] <https://twiki.cern.ch/twiki/bin/view/CMS/SusyEcalMaskedCellSummary>.
- [20] CMS Collaboration, “Data-Driven Estimation of the Invisible Z Background to the SUSY MET Plus Jets Search”, CMS Physics Analysis Summary CMS-PAS-SUS-08-002, CERN, (2008).
- [21] CMS Collaboration, “Performance of Methods for Data-Driven Background Estimation in SUSY Searches”, CMS Physics Analysis Summary CMS-PAS-SUS-10-001, CERN, (2010).
- [22] CMS Collaboration, “Search for supersymmetry in all-hadronic events with MT2”, CMS Physics Analysis Summary CMS-PAS-SUS-11-005, CERN, (2011).
- [23] CMS Collaboration, “Combination of b-tagging efficiency measurements in 2012 data at 8 TeV pp collision”, CMS Physics Analysis Summary CMS-BTV-AN-12-470, CERN, (2012).
- [24] A. L. Read., “Presentation of search results: the CLs technique.”, *J.Phys.G: Nucl. Part. Phys.* **28** (2002).
- [25] E. G. Glen Cowan, Kyle Cranmer and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics.”, *Eur.Phys.J.* **C71:1554** (2011).
- [26] “The RooStats-based statistics tools for Higgs PAG”, <https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideHiggsAnalysisCombinedLimit>.
- [27] D.Alves et al., “Simplified Models for LHC New Physics Searches.”, [arXiv:1105.2838](https://arxiv.org/abs/1105.2838).
- [28] T. S. W. G. for Hadronic Stops, “Search for Direct Scalar Top-quark Pair Production in the All-hadronic Channel at 8 TeV using a Top-quark tagger”, CMS Analysis Note 2012/490, CERN, (2012).