Chapter 4

Event Selection: Baseline and Search regions

Contents

4.1 SUSY signature

In this thesis, as mentioned in section 1.2.3, a search for SUSY is performed in the context of the simplified model T5qqqqWW (see Figure 4.1)[1]. In this model, each gluino decays to two light quarks (c,s,u,d) and an intermediate chargino, with the latter decaying to a W boson and a neutralino.

In the event selection, one of the W's is chosen to be decaying leptonically. The choice of single lepton final states provides a cleaner event topology than the full hadronic final states while keeping the signal efficiency sufficiently high. Therefore, the visible final

[1] Throughout this thesis, the T5qqqqWW model is sometimes referred as the signal model.

state includes at least six jets and one lepton (can be either electron or muon). Moreover, it is required that none of the jets are tagged as b quark to increase the probablity of selecting the events with topology T5qqqqWW. Furthermore, in such events, there is a momentum in balance originated by the neutralino and the neutrino which leave no trace in the detector. Therefore, a high missing transferse energy (MET) is also required. These selections can be further enhanced with the characteristics of individual objects such as; the transverse momentum of lepton $p_{T,\ell}$ or jets $p_{T,jets}$. However, only selecting events with one lepton, many jets, and high MET or using object characteristics is not enough to reveal the existence of SUSY-like events which are rare among the many SM like events with the similar final state. The $W+$jets and $t\bar{t}+$jets events are the leading SM processes which can mimic the T5qqqqWW signature. Hence, some advanced kinematic variables need to be invented to eliminate these SM background processes. The table 4.1 shows the list of variables which are used in this analysis. This list does not include the variables that are used in the identification of particle flow objects (see section 3.1). ~~The physics variables are categorised according to their mathematical complexity.~~

Simple variables	Advanced variables
L_T , H_T	$\Delta\Phi(W,\ell)$
$p_{T,\ell}$, η_ℓ	$M_{\mathrm{T2}}(\vec{p}_{\mathrm{T}}^{\,\ell}, \vec{p}_{\mathrm{T}}^{\,t}, \vec{p}_T^{\,miss})$
$p_{T,jets(1,2)}$	
n_{jets}	
n_{b-tag}	

Table 4.1: List of kinematic variables

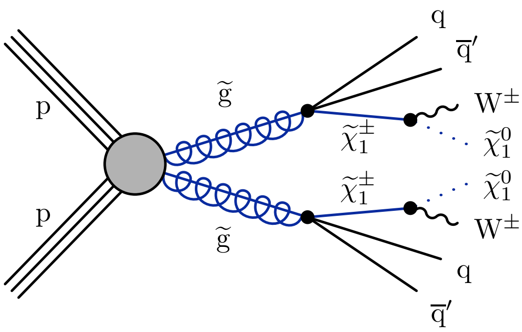


Figure 4.1: Diagram showing the simplified model T5qqqqWW.

4.1.1 Key variables

The list of simple variables consists of fundamental properties of the reconstructed physics objects, such as the transverse momentum p_T of these objects, and their linear combinations.

$\boldsymbol{L_T}$ is the scalar sum of the charged lepton p_T and the missing transverse momentum E_T^{miss}. Given no additional source of E_T^{miss}, except the SM neutrino, the variable L_T can be also written as: $\sqrt{p_{T(W)}^2 + M_{T(W)}^2}$ (see Appendix **??**). For events with a single highly boosted W boson ($p_{T(W)} \gg M_{T(W)}$), L_T $_{T(W)}$. This variable is also known as "leptonic mass scale" of the event.

$\boldsymbol{H_T}$ he scalar sum of the transverse momenta of all jets above a p_T threshold in the event. The major contribution to this sum is coming from transverse energy of the jets originated from the gluino decay, and it is related to the mass gap between the gluino and the chargino. This variable is reflecting the "hadronic ~~mass~~ scale" of the event.

$\boldsymbol{n_{jets}}$ is the number of all jets above a p_T threshold, same as in the H_T calculation, in the event.

$\boldsymbol{n_{b-tag}}$ is the number of all jets tagged as coming from a b quark above a p_T threshold, same as with other jets, in the event.

The other variables in the list such as $p_{T,\ell}$ or η_ℓ was discussed in ~~the~~ section 3.1 regarding particle flow objects.

There are two advanced variables used in this analysis.

$\boldsymbol{\Delta\Phi(W,\ell)}$ is the azimuthal angle between the W boson and the charged lepton. In the $W+$jets and $t\bar{t}+$jets events, the lepton comes from a leptonic decay of W boson, $W \to \ell\nu$. Therefore, $\Delta\Phi(W,\ell)$ strongly related to the mass of the W boson and its momentum. In SM events with high missing energy indicates that the W bosons yielding the lepton and the neutrino are boosted, hence resulting in a narrow distribution in $\Delta\Phi(W,\ell)$. On the other side, in SUSY decays, the missing transverse energy comes from two neutralinos and the neutrino, which randomizes the $\Delta\Phi(W,\ell)$, thus resulting in an even distribution in $\Delta\Phi(W,\ell)$. As a result of its features, the variable $\Delta\Phi(W,\ell)$, is the most significant variable in this analysis. The high values of $\Delta\Phi(W,\ell)$ is used as a signal region[2] while the low values of $\Delta\Phi(W,\ell)$ is the control region[3].

$\boldsymbol{M_{\mathrm{T2}}}$ is defined as:

$$M_{\mathrm{T2}}(\vec{p}_{\mathrm{T}}^{\ell}, \vec{p}_{\mathrm{T}}^{t}, \vec{p}_T^{miss}) = \min_{\vec{p}_{\mathrm{T}}^{(1)} + \vec{p}_{\mathrm{T}}^{(2)} = \vec{p}_T^{miss}} \left\{ \max\left[M_{\mathrm{T}}(\vec{p}_{\mathrm{T}}^{\ell}, \vec{p}_{\mathrm{T}}^{(1)}), M_{\mathrm{T}}(\vec{p}_{\mathrm{T}}^{t}, \vec{p}_{\mathrm{T}}^{(2)}) \right] \right\}, \tag{4.1}$$

[2] where signal to background event counts are sufficiently large.
[3] where signal event counts are negligible with respect to background event counts.

where M_{T} is the transverse mass and the indices t,ℓ represent isolated track and the selected lepton repectively. In this analysis, this variable is used only in the baseline selection to reduce the dileptonic $t\bar{t}$+jets contribution which is more important in the signal region (high $\Delta\Phi(W,\ell)$).

4.1.2 Signal samples

The simplified model T5qqqqWW is discussed in section 1.2.3. This model originally includes three free parameters: the masses of the gluino $m_{\tilde{g}}$, the intermediate chargino $m_{\tilde{\chi}_1^{\pm}}$ and the neutralino $m_{\tilde{\chi}_1^0}$. To reduce the three dimension mass space the chargino mass is fixed to midway between gluino and neutralino mass. Hence, this model is investigated in the $m_{\tilde{g}}$-$m_{\tilde{\chi}_1^0}$ mass plane. For each mass point, a separate Monte Carlo sample is produced. The mass plane is shown in the Figure 4.2 where the z-axis represents the total number of events generated. This scan of parameter space includes 657 mass points and it requires more computational power than usual SM process production therefore fastsim is used as explained in section **??**. The MADGRAPHv5 event generator is used for signal events modelling. The samples are normalized to the cross section presented by the LHC SUSY Cross Section Working Group [39]. Throughout present thesis, two mass points corresponding to different gluino and neutralino masses are used as benchmarks to study the kinematic properties of the signal.

T5qqqqWW(1.9,0.1) represents a point in the high mass gap region with $m_{\tilde{g}} = 1900$ GeV and $m_{\tilde{\chi}_1^0} = 100$ GeV. These signal events have high $H_{\mathrm{T}}$and L_{T}.

T5qqqqWW(1.5,1.0) represents a point close to the compressed region with $m_{\tilde{g}} = 1500$ GeV and $m_{\tilde{\chi}_1^0} = 1000$ GeV. The compressed region is where the $m_{\tilde{g}} - m_{\tilde{\chi}_1^0} \leq 2m_W$. These signal events have lower H_{T} and $L_{\mathrm{T}}$with respect to high mass gap ones.

4.2 Background processes

As mentioned earlier in this section the most important background processes are W+jets and $t\bar{t}$+jets events. Requiring zero b-tagged jets in the event selection, W+jets events becomes the main background. All of the background processes considered in this analysis are listed below. PYTHIA 8 is used for parton shower and hadronisation.

***$t\bar{t}$*+jets** is generated with MADGRAPH. To enhance the statistics the combination of H_{T} binned samples, dedicated semi- and dilep- tonic samples is used.

***W*+jets** is generated with MADGRAPH. H_{T} binned samples, where W decaying leptonically, are used.

QCD multijets events are produced with MADGRAPH in bins of H_{T}.

$t/\bar{t}$ samples are produced with powheg except the t decaying leptonically with s-channel is produced with AMC@NLO.
Drell-Yan evemts are produced with MADGRAPH and $m_{\ell\ell} > 50GeV$ samples are used in H_{T} bins.
Di-boson samples are produced with AMC@NLO except WW samples are produced with powheg.
$t\bar{t}V$ samples are produced with AMC@NLO.
A list of all background MC samples with cross sections can be found in Appendix **??**.

4.2.1 Scale Factors

The modeling of physics processes and the simulation of the detector responses are not perfect. Studies of the data-MC discrepancies indicate that MC needs to be corrected by some event-by-event weights, which are called scale factors. The scale factors used in this analysis are as follows:
Lepton identification and reconstruction efficiency:
MC samples have to be corrected according to the reconstruction and identification efficiencies of leptons. In this work, scale factors are applied as a function of p_T and η of the selected lepton. They are calculated for electrons and muons separately by the e-gamma and muon physics object groups respectively[44].
B-tagging efficiency:
The simulated b-tagging efficiency is corrected with respect to the one in data. For each

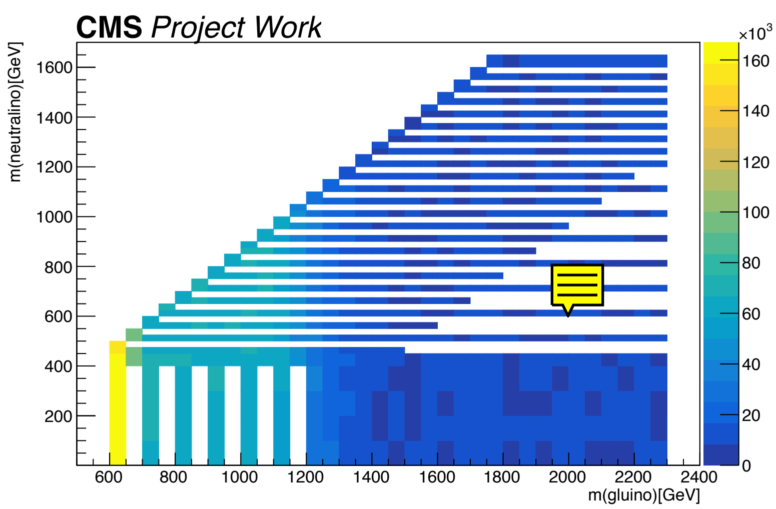


Figure 4.2: Produced signal mass points for the simplified T5qqqqWW model.

simulated event, weights corresponding to the different number of b-tagged jets values are calculated and this results in multiple weights for each eve 4]. This method allows reusing the events. The advantage of this method is that it provides to use full statistical power of the MC events.

ISR reweighting:

Since the 8 TeV Run of LHC, it is known that the p_T spectrum of $t\bar{t}$ events are not well modeled. Therefore, an event-by-event correction is obtained using the jets from initial state radiation (ISR). In the present analysis, each event is corrected according to number of ISR jets. The weights can be seen in table 4.2 where the D factor is for keeping the normalization of the overall sample invariant. This D factor is calculated using an inclusive sample i.e. without the specific selections of the analysis.

Table 4.2: Weights based on the number of ISR jets as given in Ref. [46]

nISR jet	Normalisation weight $D_{t\bar{t}} = 1.071$
0	
1	$D \times (0.920 \pm 0.005 \pm 0.040)$
2	$D \times (0.821 \pm 0.006 \pm 0.090)$
3	$D \times (0.715 \pm 0.009 \pm 0.143)$
4	$D \times (0.662 \pm 0.016 \pm 0.169)$
5	$D \times (0.561 \pm 0.027 \pm 0.219)$
≥ 6	$D \times (0.511 \pm 0.041 \pm 0.244)$

Pileup:

As discussed earlier in Section **??**, the number of pile-up interactions in simulated samples is generated with a prior distribution. Thus, the MC pile-up distribution may vary from the actual pile-up distribution in Data. The simulated distribution is rescaled with the Data/MC ratio in Figure 4.3.

4.3 Baseline selection

~~Earlier in this~~ section 4.1, the SUSY signature considered in this work is introduced. Moreover, a suitable primary event selection for removing the SM background as much as possible while keeping the signal efficiency high was discussed the following, details of this event selection are explained.

Selection of leptons:

The selected lepton, which can be an electron or a muon, is required to have a minimum p_T of 25 GeV while at the same time it satisfies the good lepton criteria introduced in section 3.1. Additionally, all the other electrons and muons with p_T greater than 10 GeV are vetoed. These veto leptons satisfy the loose lepton selection with $I_{rel} < 0.4$.

Selection of jets:

The reconstruction of jets is already introduced in section 3.1. In the present analysis, jets are required to have $p_T > 30$ GeV and $|\eta| < 2.4$. In order to avoid double counting of objects, jets that are close ($\Delta R < 0.4$) to either a veto or selected lepton are removed. As discussed earlier in this chapter the T5qqqqWW is expected to have at least five jets. Moreover, the mass difference of the gluino and the intermediate chargino affects the p_T of the jets. Therefore, it is required that two highest p_T jets satisfy the $p_T > 80$ GeV condition. Additionally, in the baseline selection of the analysis, the number of jets tagged as coming from b quarks is zero to suppress the events containing top quarks.

Energy scales thresholds:

The signal model considered in this work favors events with high hadronic and leptonic scale. The hadronic scale is chosen to be at least 500 GeV while the leptonic scale thresh-

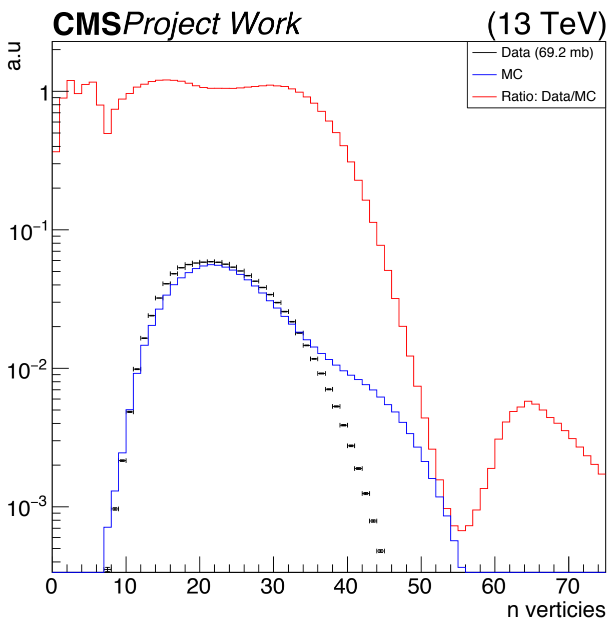
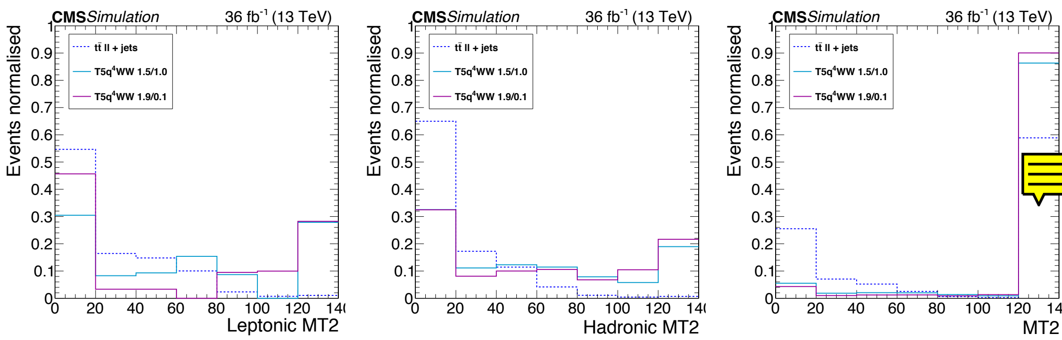


Figure 4.3: The normalised distributions of mean number of interactions per bunch crossing during the analysed data Run2016 (black) and in the MC samples (blue). And the red distribution represents the Data to MC ratio. For the data distribution the latest luminosity calibrations [47] and an inelastic pp cross-section of 69.2 mb are used.

old is 250 GeV.

Isolated track veto:

The isolated track veto is designed to suppress $t\bar{t}$ events in which both W bosons decay leptonically and one lepton does not satisfy the selection criteria for veto leptons. In the mechanism of this veto, the M_{T2} variable, which is introduced in section 4.1.1, is used with an isolated track $\vec{t}$, a lepton $\vec{\ell}$ and the missing transverse momentum $\vec{p}_T^{miss}$. In the calculation of M_{T2}, it is assumed that the missing energy source is the two neutrinos from the dileptonic $t\bar{t}$ decay. The minimization runs over all possible splitting of $\vec{p}_T^{miss}$. The figure 4.4 shows the M_{T2} distribution seperate for hadronic and leptonic tracks after the baseline selection with $H_{\mathrm{T}} > 500$ GeV, $L_{\mathrm{T}} > 250$ GeV, $n_{\mathrm{jet}} \geq 5$ and without b-tag requirement. The distribution of M_{T2} is slightly different for the two cases and a lower M_{T2} cut of 60 GeV and 80 GeV is applied for hadronic and leptonic tracks respectively. The rightmost plot in the figure 4.4 shows the M_{T2} distribution for all events i.e. including events that do not have any isolated track at all. In this case events without any isolated track are added to the overflow bin. It can also be derived from the plot that only about 20% of the signal events have an opposite charged isolated track at all compared to 40% of the dileptonic $t\bar{t}$ events.



(a) M_{T2} with leptonic tracks (b) M_{T2} with hadronic tracks (c) M_{T2} (All events)

Figure 4.4: Distributions of M_{T2} for events with electron or muon veto tracks (a) and hadronic veto tracks (b), for dileptonic $t\bar{t}$ (blue dotted) and T5qqqqWW (purple and azure) signal samples. The highest bin is always an overflow bin. The majority $t\bar{t}$ events have $M_{\mathrm{T2}} < 80$ GeV, while the signal events have longer tails in M_{T2}. The rightmost plot (c) shows the distribution for all events.

The effect of each baseline requirement is demonstrated in Figure 4.5 for the different background processes (stacked) and for two signal benchmark points. It should be noted that the background samples include an initial skim of $H_{\mathrm{T}} > 350$ GeV and $L_{\mathrm{T}} > 150$ GeV to shorten the computation time. It can be observed that the design of baseline selections works such that they reduced the background events significantly while only a small fraction of signal events are eliminated by the selections. It can be also notted that, naturally, the QCD background is almost vanished after the single muon selection. The estimation of important backgrounds will be explained in the next chapter.

Selection	Object definitions
Single lepton	Tight leptons, $p_T \geq 25$ GeV and $\lvert\eta\rvert < 2.4$ and $I_{mini} < 0.1(0.2)$ for electrons(muons)
Lepton veto	Loose leptons, $p_T \geq 10$ GeV and $\lvert\eta\rvert < 2.4$ and $I_{mini} < 0.4$
Isolated track veto	$I_{rel} < 0.3$, $\Delta R(\ell, track) < 0.1$, $charge_{track}$ = -$charge_{lepton}$
$n_{jets} \geq 5$	Good jets with $p_T \geq 30$ GeV and $\lvert\eta\rvert < 2.4$ and cleaned from close leptons
$p_{T,jets(1,2)} \geq 80$ GeV	
$H_T \geq 500$ GeV	$\sum_{jets} p_T$
$L_T \geq 250$ GeV	$\not{E}_T + p_T^{lep}$
$n_{b-tag} = 0$	b-tagged good jets with CSVv2 Medium working point (0.8484)

Table 4.3: List of event selection criteria and object requirements.

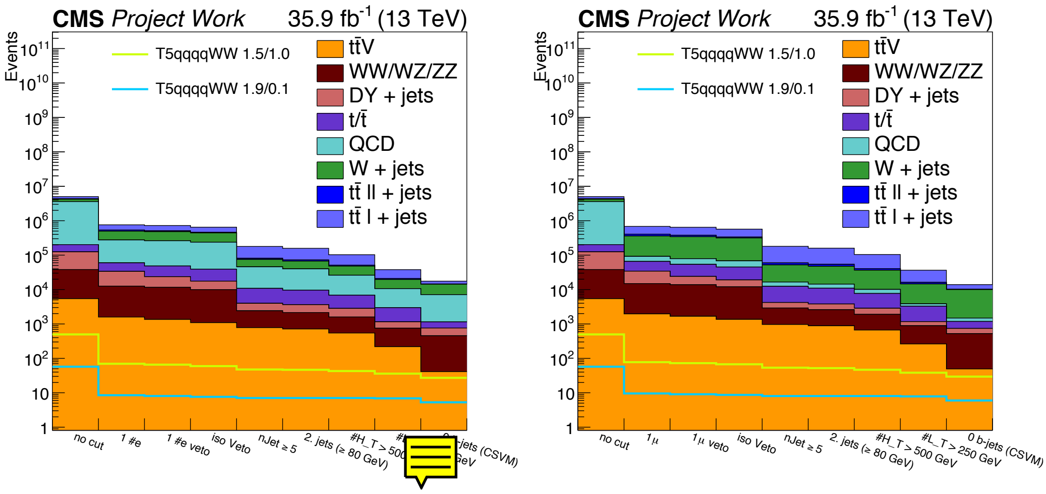


Figure 4.5: The number of events for each sample for the electron (left) and muon (right) channel. Samples are scaled to the luminosity with their cross section and scale factors discussed in section 4.2.1.

4.4 Data Samples

The data used for this analysis is recorded by the CMS detector during the 2016 LHC run at 13 TeV. As in Figure **??**, The total integrated luminosity collected by CMS is 37.76 fb^{-1}. In the present analysis, the data validated by the CMS Data Quality Monitoring (CMS-DQM) certification team, which correspond to an integrated luminosity of L=35.9 fb^{-1}, is used.

4.4.1 Triggers

In this analysis, the main trigger set is a combination of triggers containing an isolated single lepton, electron or muon, with pT of 15 GeV and HT of 350/400 GeV. To recover inefficiencies due to an update on the online electron ID and the saturated L1-jets, the trigger strategy had to be extended: The list of trigger paths can be seen in Table 4.4.

Single Electron Dataset
HLT_Ele105_CaloIdVT_GsfTrkIdT_v
HLT_Ele115_CaloIdVT_GsfTrkIdT_v
HLT_Ele50_CaloIdVT_GsfTrkIdT_PFJet165_v
HLT_Ele27_WPTight_Gsf_v
HLT_Ele15_IsoVVVL_PFHT350_v
HLT_Ele15_IsoVVVL_PFHT400_v
Single Muon dataset
HLT_Mu50_v
HLT_IsoMu24_v
HLT_IsoTkMu24_v
HLT_Mu15_IsoVVVL_PFHT350_v
HLT_Mu15_IsoVVVL_PFHT400_v
MET dataset
HLT_PFMET100_PFMHT100_IDTight_ OR HLT_PFMETNoMu100_PFMHTNoMu100_IDTight_
HLT_PFMET110_PFMHT110_IDTight_ OR HLT_PFMETNoMu110_PFMHTNoMu110_IDTight_
HLT_PFMET120_PFMHT120_IDTight_ OR HLT_PFMETNoMu120_PFMHTNoMu120_IDTight_

Table 4.4: List of HLT paths

Events recorded with these trigger paths are allocated in three primary datasets (PDs):

the **SingleElectron**, **SingleMuon** and **MET** datasets. Therefore, when successively adding PDs, triggers contained in previous PDs are vetoed to avoid double counting of events contained in more than one of the PDs. The trigger efficiency calculation is shown in the Equation 4.2. The trigger efficiencies are measured as a function of L_{T}, H_{T} and lepton p_T and can be seen in Figures 4.6, 4.7, 4.8 respectively. The efficiency distribution in the baseline region (L_{T}>250 GeV; H_{T}>500 GeV; lepton p_T >25 GeV) is flat and its value close to 100%(98%) for the muons (electrons).

$$\epsilon = \frac{N(\text{all events passing probed trigger(s)} + \text{preselection} + \text{reference trigger})}{N(\text{all events passing preselection} + \text{reference trigger})} \tag{4.2}$$

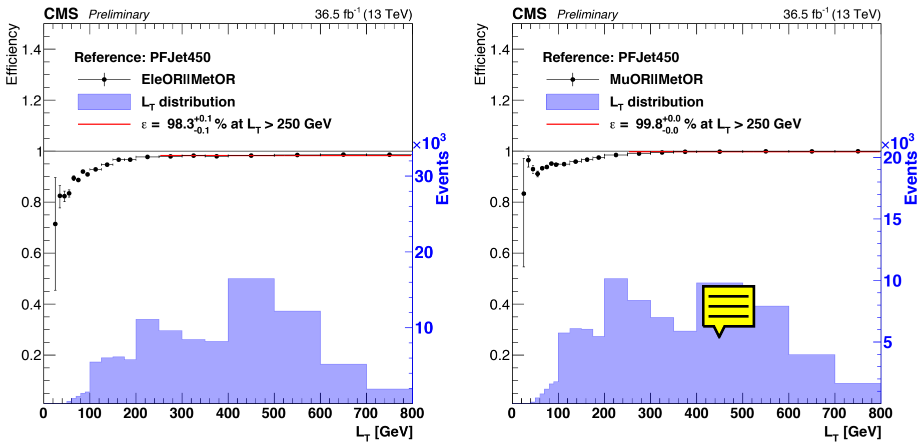


Figure 4.6: Measurement of the trigger efficiency as a function of L_T for the electron trigger selection (Ele50PFJet165 / IsoEle27 / Ele105 / Ele115 / Ele15HT400 / Ele15HT350 / METMHTTriggers) on the left and the muon trigger selection (Mu50 / IsoMu24 / Mu15HT400 / Mu15HT350 / METMHTTriggers) on the right [45].

4.5 Event cleaning: Filters

The event reconstruction can fail due to the noisy detector cells and other kinds of detector problems. These can result in incorrectly reconstructed physics objects such as: muons, jets and hence MET. ~~In CMS, the physics object groups related to the problem release event filters or cures for the falsely reconstructed objects.~~

Primary vertex filter:
A good primary vertex is required which is introduced in section 3.2.1.

Beam halo filter:
The collisions of the beam with residual gas inside LHC cause showers of secondary particles. Additionally to these scattering effects, charged particles are deflected by the magnetic field of the beam optics. These particles are called beam halo particles and one of

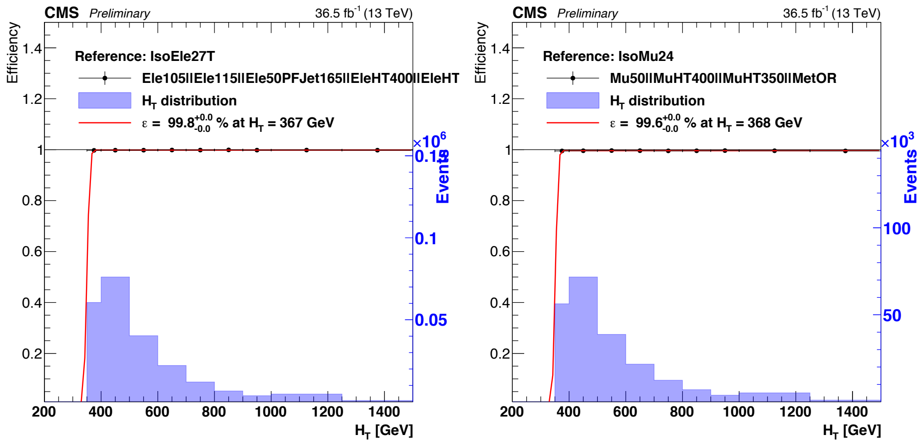


Figure 4.7: Measurement of the trigger efficiency as a function of $H_T$for the electron trigger selection (Ele50PFJet165 / IsoEle27 / Ele105 / Ele115 / Ele15HT400 / Ele15HT350 / METMHTTriggers) on the left and the muon trigger selection (Mu50 / IsoMu24 / Mu15HT400 / Mu15HT350 / METMHTTriggers) on the right [45].

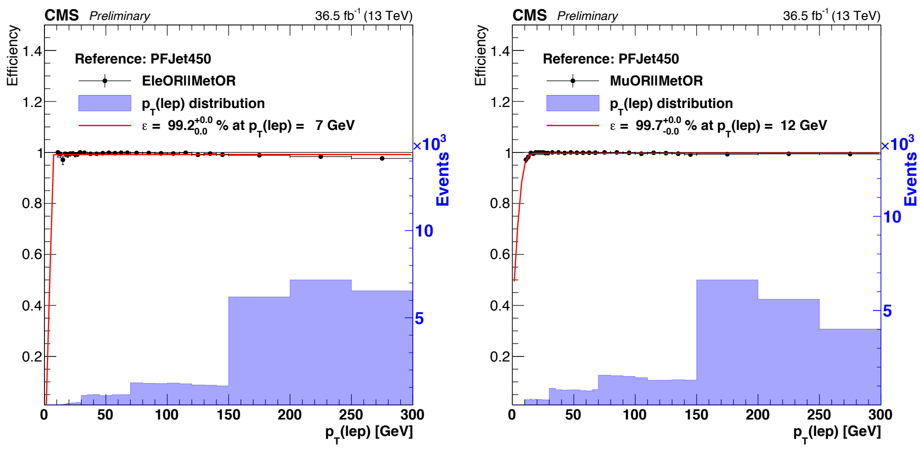


Figure 4.8: Measurement of the trigger efficiency as a function of lepton p_T for the electron trigger selection (Ele50PFJet165 / IsoEle27 / Ele105 / Ele115 / Ele15HT400 / Ele15HT350 / METMHTTriggers) is shown on the left and the muon trigger selection (Mu50 / IsoMu24 / Mu15HT400 / Mu15HT350 / METMHTTriggers) is shown on the right. The lepton p_T trigger efficiency is measured after an $L_T > 250\,\text{GeV}$ cut which is the analysis baseline selection [45].

the main sources of the beam background of the LHC. In CMS, the beam halo algorithm considers the particles produced outside the CMS cavern and to detect events with beam halo it uses the timing information and hit topology in CSC, ECAL and HCAL subsystems.

HB-HE noise filter:

This noise is originated from the Hybrid PhotoDiodes and Readout Boxes of the HCAL.

The timing, pulse shape as well as the other readout errors are used to detect the noise.

ECAL dead cell trigger primitive filter:

The existence of noisy crystals in ECAL can lead to fake MET. The events, in which the noisy cells deposit high energy, are filtered.

Bad PF muon filter:

This filter fires when there is a PF muon with too low quality and has large p_T. The quality of the muon is determined according to its tracking uncertainty, segment compatibility and other detector related features. This bad muon is required to have $p_T > 100$ GeV. Unlike the other filters explained above this filter is applied to background simulation samples as recommended.

Bad charged hadron filter:

The events, where there is a muon but it fails to be a PF muon and it still contributes to PF MET calculation as a charged hadron candidate, are filtered. Here, as in the previous case, this muon is required to have $p_T > 100$ GeV. Moreover, it is required that the muon and the charged hadron traces are almost overlapping ($\Delta R(\mu, charged\, hadron) < 0.00001$). In addition to this, their p_Ts are needed to be very close to each other. Similarly to the bad PF muon filter, this is applied to background MC samples as well as the data events.

Duplicate muon filters:

In the Re-reco data it is observed that there is an increase in events in the MET tail of Z→ $\mu\mu$ data. It is understood by the muon POG that there are dublicate muons in the events due to reconstruction failures. Additionally, two filters are recommended by the SUSY group: One is to remove events with the ratio of PF MET to calorimeter MET is more than 5. The second is to remove events containing bad jets which have $p_T > 200$ GeV, muon energy fraction > 0.5 and $\Delta\phi(me) > \pi - 4$.

Filter on fastsim:

This filter is to clean up bad jets from the fastsim events. To remember, in this work, fastsim is used for the signal MC production. The event is removed if there is a jet satisfying the following conditions: $p_T > 20$ GeV, $|\eta| < 2.5$, unmatched to a generated jet ($\Delta R(jet, gen\, jet) > 0.3$), and charged hadron fraction < 0.1.

The efficiency of all these filters after the analysis baseline selection is approximately 98%

4.6 Control plots

The distributions of main kinematic variables, also called control plots, are shown in this section. In all the control plots, the colored lines represent the signal models and the

color filled stacked histograms display the background processes. Additionally, the black dotted distribution exhibits the observed data points requiring the triggers introduced in section 4.4.1. In all the control plots, the events are cleaned by the filters discussed in section 4.5. The signal and background events are scaled by the luminosity factor and additional weights introduced in section 4.2.1. In the Figure 4.9 top left plot shows the n-btag distribution after the baseline selection (see section 4.3). The distribution of simulated signal events peaks at zero as expected. Clearly, one can see from this distribution that choosing events with zero b-tagged jets suppresses the $t\bar{t}$ background significantly. The other distributions, displaying main kinematic variables, exhibit reasonable MC-data agreement although it is not necessarily expected. The plot shows the number of jet distribution (top/right), H_{T} (bottom/left) and L_{T} (bottom/right) are shown.
The multiplicities of jets and b-tagged jets display no difference between the signal scenarios, since the mass splitting has no effect on the decay topology. The two selected signal benchmark models show differences in the distributions of H_{T} and L_{T}, this is due to the gluino-neutralino mass splitting. In the non-compressed T5qqqqWW (1900,100) model, the quarks coming from the gluino decay have a large boost, resulting in high leptonic and hadronic energy scales. For the compressed region with T5qqqqWW (1500,1000) no such effect is observed, and the shape of the distributions look similar to the SM processes.
Additional control plots are shown in Appendix **??**.

4.7 Search regions

4.7.1 Signal and control region

Signal region is where it is expected to observe a deviation from SM background in the existence of SUSY scenario. In order to locate this kinematic region a powerful discriminative varible is used. Figure 4.10 shows the $\Delta\phi$ distribution after the baseline selection. The plot exhbits that the SM background processes peak at zero while the signal models stay flat. Naturally, this shape difference leads to the high values of $\Delta\phi$ as the signal region (SR) and the low values as the control region (CR). The control region is used in the background estimation, which will be discussed in the next chapter. It can also be seen from the distribution that after the baseline selection $W+\text{jets}$ and $t\bar{t}+\text{jets}$ are the main background components. The $t\bar{t}$ composition in the CR is dominated by the single leptonic decays while in the SR the di-leptonic $t\bar{t}$ decays are dominating.

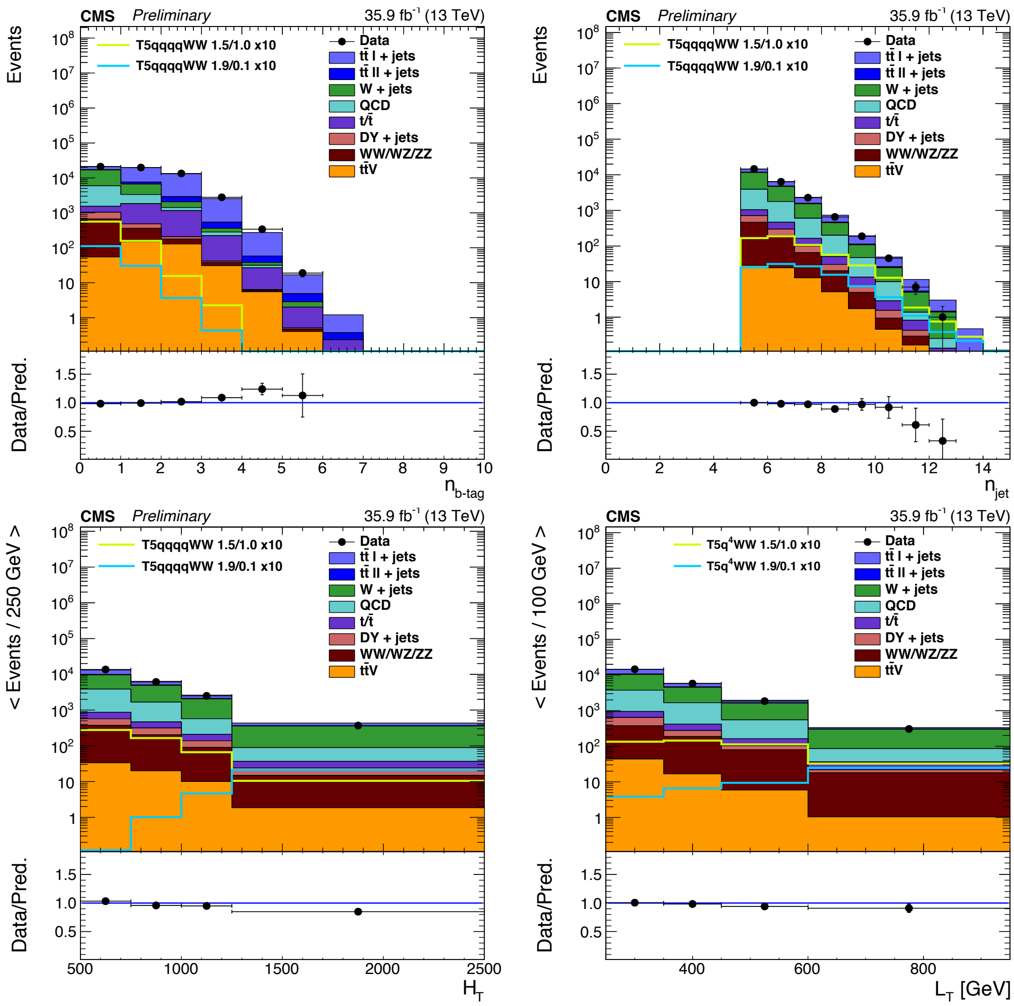


Figure 4.9: The top left distribution shows the number of b-tagged jets, after the baseline selection, requiring at least five jets, minimum H_T of 500 GeV, a minimum L_T of 250 GeV and exactly one lepton with p_T >25 GeV. The rest of the distributions are plotted after the same baseline selection additionally requiring no b-tagged jets. In the top row the number of jets (right) while in the bottom row H_T (left) and L_T (right) distributions are shown. The simulated background events are stacked on top of each other, and several signal points are overlaid for illustration without being stacked. The model T5qqqqWW (1.5,1.0) (T5qqqqWW (1.9,0.1)) corresponds to a gluino mass of 1.5 TeV (1.9 TeV) and neutralino mass of 1.1 TeV (0.1 TeV), respectively. The intermediate chargino mass is fixed at 1.25 TeV (1.0 TeV). The two benchmark signal models are scaled up by a factor of 10.

4.7.2 Main band regions

To further enhance the sensitivity, ~~the phase space is subdivided in bins of~~ n_jet, H_T and L_T. These sub bins are called main band regions (MB).

The binning is designed for an integrated luminosity of 40 fb^{-1} scenario. In the optimiza-

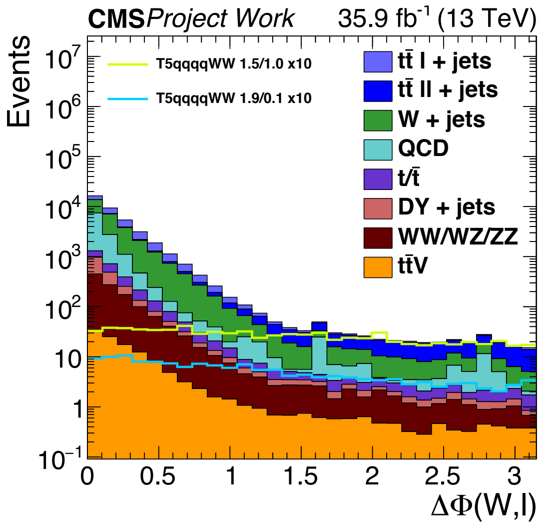


Figure 4.10: The $\Delta\phi$, after the baseline selection, requiring at least five jets and non of b-tagged, minimum H_{T} of 500 GeV, a minimum L_{T} of 250 GeV and exactly one lepton (electron of muon) with p_T >25 GeV. The simulated background events are stacked on top of each other, and several signal points are overlaid for illustration without being stacked. The model T5qqqqWW (1.5,1.0) (T5qqqqWW (1.9,0.1)) corresponds to a gluino mass of 1.5 TeV (1.9 TeV) and neutralino mass of 1.1 TeV (0.1 TeV), respectively. The intermediate chargino mass is fixed at 1.25 TeV (1.0 TeV).

tion e bins with the highest figure of merit are chosen among a spectrum of bins. Then, a fully profiled likelihood is performed for the several combinations; The bin combination with the highest sensitivity is ch. At the same time, for each bin, the yield of background MC is required to be at least one, to have robust background estimation. In addition to these, throughout the bins, boundary cuts are chosen to be symmetric if possible and sensible. For instance, the binning in L_{T} is same for different n_{jet} bins.
The $\Delta\phi$ and L_{T} variables are constructed by using the same objects; p_T^{lep} and $\not{E}_T$. Therefore, a correlation can be assume s in Figure 4.11, for increasing L_{T} the SM events are located towards the low values of $\Delta\phi$. On the other hand, the signal plots do not show the same trend. Therefore, a varying $\Delta\phi$ cut, depending on the L_{T} binning, is used to determine the signal region for each MB.
The resultant binning is given in table 4.5 and the corresponding simulated background and signal yields can be found in Figure ??.
As mentioned, the W+jets and $t\bar{t}$+jets are dominating the SM background. In this analysis, the number of the events in the MB SR is estimated with observed data. In the next chapter, the procedure for the estimation of these background components will be explained.

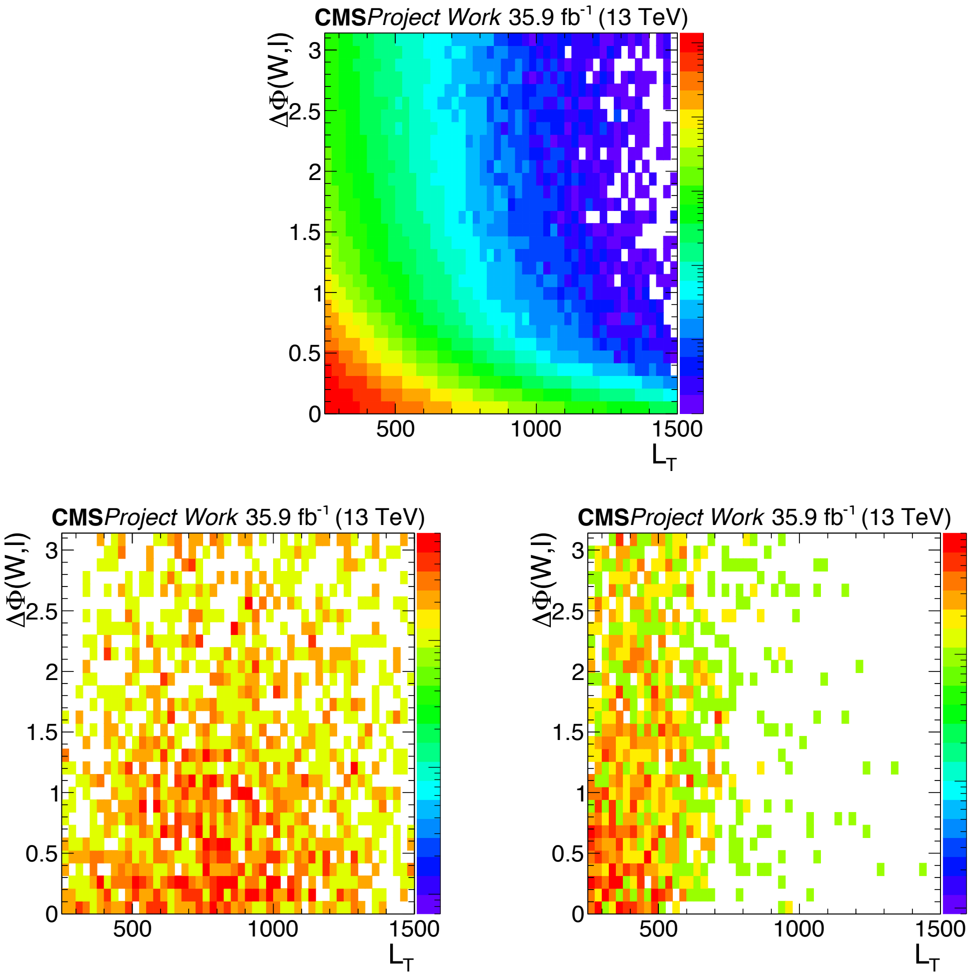


Figure 4.11: 2D distributions of event counts for the main background samples ($t\bar{t}$ + jets + W + jets) (top), and the signal T5qqqWW(1900,100) (left) and T5qqqWW(1500,1000) in the the $\Delta\phi$ vs. L_T plane after the preselection.

Signal acceptance:

The yield after baseline selection for the simulated signal events is shown in Figure 4.12 in the gluino-neutralino mass plane for the T5qqqqWW model. The acceptance of at least one event for 35.9 fb^{-1} integrated luminosity is up to 2250 GeV for gluino mass for the neutralinos masses below 1600 GeV. The equivalent selection efficiency, which is the fraction of baseline events over the total number of simulated events, is presented in the same Figure (top/right). In the bottom part of Figure, the SR efficiencies are shown. The plot on the left displays the efficiency for the constant $\Delta\phi > 1$ cut and it is changing between 50-70%. On the other hand, the left plot shows the efficiency for the variable $\Delta\phi$ cut with respect to L_T bin and this time the efficiency goes up to 80%.

Table 4.5: Signal regions.

5 jets, 0 b-tagged jets		
L_T[GeV]	H_T[GeV]	$\Delta\phi$
[250, 350]	[500, 750]	1.0
	≥ 750	1.0
[350, 450]	[500, 750]	1.0
	≥ 750	1.0
[450, 650]	[500, 750]	0.75
	[750, 1250]	0.75
	≥ 1250	0.75
≥ 650	[500, 750]	0.50
	[750, 1250]	0.50
	≥ 1250	0.50
[6,7] jets, 0 b-tagged jets		
L_T[GeV]	H_T[GeV]	$\Delta\phi$
[250, 350]	[500, 1000]	1.0
	≥ 1000	1.0
[350, 450]	[500, 1000]	1.0
	≥ 1000	1.0
[450, 650]	[500, 750]	0.75
	[750, 1250]	0.75
	≥ 1250	0.75
≥ 650	[500, 750]	0.50
	[750, 1250]	0.50
	≥ 1250	0.50
≥ 8 jets, 0 b-tagged jets		
L_T[GeV]	H_T[GeV]	$\Delta\phi$
[250, 350]	[500, 1000]	1.0
	≥ 1000	1.0
[350, 450]	[500, 1000]	1.0
	≥ 1000	1.0
[450, 650]	[500, 1250]	0.75
	≥ 1250	0.75
≥ 650	[500, 1250]	0.50
	≥ 1250	0.50

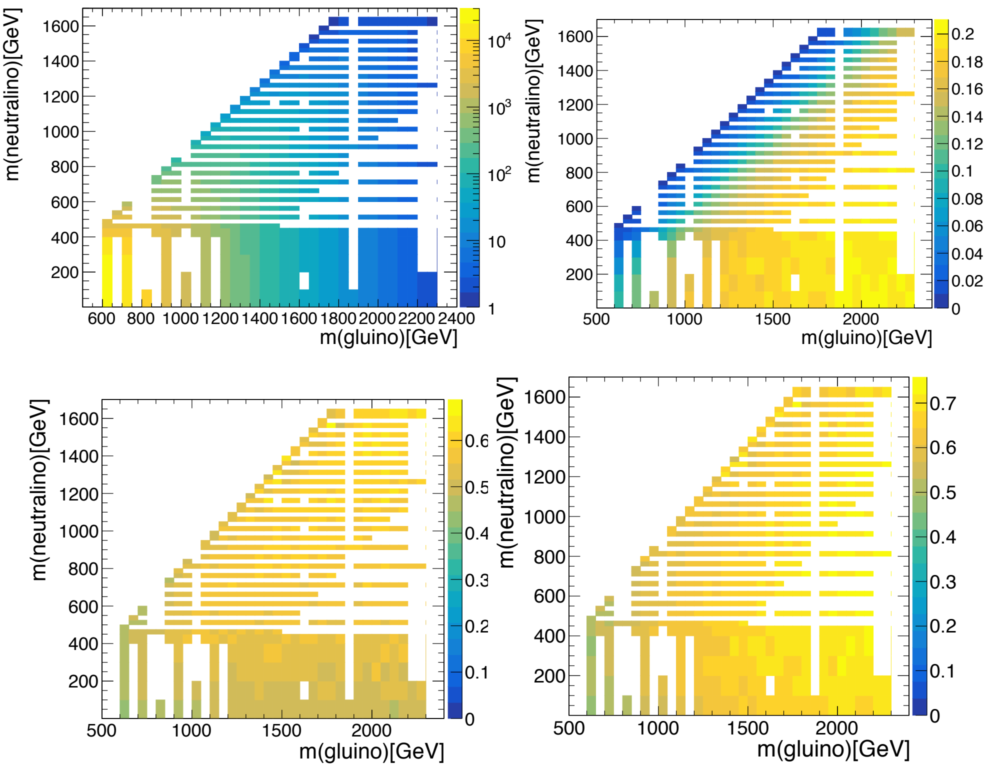


Figure 4.12: The distributions on top represent T5qqqqWW signal counts (left) and selection efficiency (right) after the baseline requirements in the gluino-lsp plane. The distributions at the bottom show the efficiency of signal region selection with respect to the baseline; $\Delta\phi > 1$ (left) and $\Delta\phi > x$ (right) where x stands for the cut value changing according to L_{T} bin.