diately determine the connection to a particular SUSY model, region of parameter space, SUSY masses, or cosmology. To establish such a connection, a search for an excess of opposite-sign minus like-sign $(OS-LS)$ $\tau_h\tau_h$ pairs is also performed following the prescription outlined in [6]. The $OS-LS$ method [6] makes use of the idea that most jet based background processes that remain after requiring the presence of high p_T non-tau jets, two hadronically decaying taus, and large $\not\!\!E_{\mathrm{T}}$ can be removed on a statistical basis because the fragmentation of the quarks and gluons produce charge blind ditau pairs.

If an excess remains in $OS-LS$ $\tau_h\tau_h$ pairs in the high $\not\!\!H_{\mathrm{T}}$ region, then we proceed to quantify the significance of the excess for potential discovery. If no excess exists, then we proceed to set a 95% C.L. upper limit on the cross section using the CMS recommended CL_s method [12].

It is important to note that the analysis has been designed with two goals in mind: (1) provide a general search for BSM processes such as SUSY; (2) if a possible excess or discovery is observed, make precision measurements of the kinematic endpoints and attempt to extract SUSY masses and model parameters. This note focuses on goal (1) ~~and~~ Figure 1 shows the overall analysis strategy ranging from the definition of the signal region to the determination of control regions.

3 Backgrounds and Samples

3.1 Backgrounds

W + jets and $t\bar{t}$ are the dominant backgrounds for this analysis. W + jets background becomes a dominant process because the decay of the W boson creates a very clean prompt tau that passes all the tau related requirements at the same rate as signal events. Moreover, jets from quarks and gluons can be identified as tau candidates. W + jets also becomes a background when the W decays to $q\bar{q}$ giving rise to tau fakes or high p_T jets if the W boson is significantly boosted. However, because W + jets events mostly contain only one tau lepton from the W boson decay, $\tau\tau$ pairs are charge blind due to the presence of at least one jet that fakes a hadronically decaying tau. Therefore, any W + jets contamination that remains after requiring the presence of 2 high p_T jets, 2 τ_h's, and large $\not\!\!E_{\mathrm{T}}$ can be removed on a statistical basis using the OS-LS subtraction technique that will be discussed in the sections that follow.

Events containing $t\bar{t}$ contribute to the expected background due to the production of real taus from one of the W boson decays ($t \to bW \to b\tau\nu$) and/or jet fakes from the second W boson decay. Additionally, events can have high p_T b-jets from the top quark decays and large $\not\!\!E_{\mathrm{T}}$ from the leptonic decays of the W bosons which give rise to energetic neutrinos. In many $\tau\tau$ related searches, $t\bar{t}$ is minimized by requiring events with zero jets tagged as b-jets. However, this is not possible for this analysis because SUSY production often occurs via stop ($\tilde{t}$) and sbottom ($\tilde{b}$) pairs that can give rise to b-jets via e.g. $\tilde{t} \to \tilde{\chi}^{\pm}_{1,2}b$ or $\tilde{t} \to \tilde{\chi}^{0}_{1,2}t \to \tilde{\chi}^{0}_{1,2}bW$. ~~However, we do make~~ use of b-tagging to provide a data-driven background extraction method for $t\bar{t}$ (section 8.1).

QCD ~~dijet~~ events are expected to be a very small contribution to the signal region due to the requirement of 4 jet-like objects (2 non-tau jets and 2 tau jets) and large $\not\!\!E_{\mathrm{T}}$. This process can only become a background if one of the jets is badly mismeasured, giving rise to large $\not\!\!E_{\mathrm{T}}$, and additional jets from initial state or final state radiation provide fake $\tau\tau$ pairs. ~~However, these cases can also be removed~~ on a statistical basis using the OS-LS subtraction technique. As the QCD contribution is expected to be small, we measure a scale factor from a QCD enriched sample to correct the MC prediction in the signal region and validate that QCD is indeed a small contribution.

$Z \to \nu\nu$ + jets is the only other important background of interest. This background does have large $\not{E}_T$ due to the neutrinos from the Z boson decay. ~~However,~~ the very low multiplicity of jets ~~means that~~ the probability to obtain 2 tau-like jets and 2 additional high p_T jets is extremely low. However, $Z(\rightarrow \nu\nu)$ + jets can become a non-negligible background in control regions where the τ requirements are dropped to enhance the background of interest.

3.2 Datasets and Monte Carlo Samples

The entire dataset collected by the CMS detector in 2011 is used in this analysis. Table 1 shows the collision datasets used. Two sets of samples are used: (1) sample of events selected with a $\not{H}_T$ trigger (HLT_PFMHT150); (2) sample of events selected with $l + \tau_h$ cross-triggers, where $l = e/\mu$. The sample of events selected with the $\not{H}_T$ trigger is used for the final analysis path, while the sample of events selected with the $l + \tau_h$ cross-triggers is used for validation studies of τ_h identification, trigger, and PU performance. The official JSON file was used to select "good" run ranges and lumi sections. The total integrated luminosity of the collision data samples used is 4.6 ~~fb~~$^{-1}$.

The official Summer 2011 MC samples are used for all SUSY and SM processes. Next to leading order processes (NLO) were included using the POWHEG or MADGRAPH framework and the TAUOLA package was used to properly decay the tau leptons. The MADGRAPH event generator was used for $t\bar{t}$, $W + Jets$, $Z \rightarrow \nu\nu + Jets$, $DY/\gamma^{*}(\rightarrow \mu\mu) + Jets$ and QCD. Table 2 shows the entire list of MC samples used. In all cases, the NNLO cross sections are used for EWK processes. The cross sections used are summarized in Table 3.

Table 1: Collision Data Samples

Physics Sample	Official CMS Datasets
	Datasets for main analysis
Run 2011A METBTag May 10 ReReco	*/METBTag/Run2011A-May10ReReco-v1/RECO*
Run 2011A MET PromptReco-v4	*/MET/Run2011A-PromptReco-v4/RECO*
Run 2011A MET PromptReco-v5	*/MET/Run2011A-PromptReco-v5/RECO*
Run 2011A MET PromptReco-v6	*/MET/Run2011A-PromptReco-v6/RECO*
Run 2011B MET PromptReco-v1	*/MET/Run2011B-PromptReco-v1/RECO*
	Datasets for Validation Studies in the $\mu\tau_h$ and $e\tau_h$ channels
Run 2011A Tau+X May 10 ReReco	*/TauPlusX/Run2011A-May10ReReco-v1/RECO*
Run 2011A Tau+X PromptReco-v4	*/TauPlusX/Run2011A-PromptReco-v4/RECO*
Run 2011A Tau+X PromptReco-v5	*/TauPlusX/Run2011A-PromptReco-v5/RECO*
Run 2011A Tau+X PromptReco-v6	*/TauPlusX/Run2011A-PromptReco-v6/RECO*
Run 2011B Tau+X PromptReco-v1	*/TauPlusX/Run2011B-PromptReco-v1/RECO*

Because the MC simulated samples contain a PU distribution that does not match that of data, the MC needs to be properly weighted to fit the PU distribution observed in data. The reweighting of MC events is performed by determining the probabilities to obtain n vertices in data ($P_{data}(n)$) and MC ($P_{MC}(n)$) and using the event weights

$$w_{PU}(n) = \frac{P_{data}(n)}{P_{MC}(n)} \qquad (1)$$

to reweigh MC events based on the number of vertices.

Table 2: MC Samples

Physics Sample	Official CMS Datasets
$Z \to \tau\tau$	/DYToTauTau_M-20_TuneZ2_7TeV-pythia6-tauola/Summer11-PU_S3_START42_V11-v2/
$Z \to \mu\mu$	/DYToMuMu_M-20_TuneZ2_7TeV-pythia6/Summer11-PU_S3_START42_V11-v2/
$Z \to ee$	/DYToEE_M-20_TuneZ2_7TeV-pythia6/Summer11-PU_S3_START42_V11-v2/
$W + jets$	/WJetsToLNu_TuneZ2_7TeV-madgraph-tauola/Summer11-PU_S4_START42_V11-v1/
$t\bar{t}$	/TTJets_TuneZ2_7TeV-madgraph-tauola/Summer11-PU_S4_START42_V11-v1/
$Z \to \nu\nu + Jets$	/ZinvisibleJets_7TeV-madgraph/Spring11-PU_S1_START311_V1G1-v1/
$Z \to \mu\mu + Jets$	/DYJetsToLL_TuneZ2_M-50_7TeV-madgraph-tauola/Summer11-PU_S4_START42_V11-v1/
qcd 15 → 30	/QCD-Pt_15to30_TuneZ2_7TeV_pythia6/Summer11-PU_S4_START42_V11-v1/
qcd 30 → 50	/QCD-Pt_30to50_TuneZ2_7TeV_pythia6/Summer11-PU_S4_START42_V11-v1/
qcd 50 → 80	/QCD-Pt_50to80_TuneZ2_7TeV_pythia6/Summer11-PU_S4_START42_V11-v1/
qcd 80 → 120	/QCD-Pt_80to120_TuneZ2_7TeV_pythia6/Summer11-PU_S4_START42_V11-v1/
qcd 120 → 170	/QCD-Pt_120to170_TuneZ2_7TeV_pythia6/Summer11-PU_S4_START42_V11-v1/
qcd 170 → 300	/QCD-Pt_170to300_TuneZ2_7TeV_pythia6/Summer11-PU_S4_START42_V11-v1/
qcd 300 → 470	/QCD-Pt_300to470_TuneZ2_7TeV_pythia6/Summer11-PU_S4_START42_V11-v1/
qcd 470 → 600	/QCD-Pt_470to600_TuneZ2_7TeV_pythia6/Summer11-PU_S4_START42_V11-v1/
qcd 600 → 800	/QCD-Pt_600to800_TuneZ2_7TeV_pythia6/Summer11-PU_S4_START42_V11-v1/
qcd 800 → 1000	/QCD-Pt_800to1000_TuneZ2_7TeV_pythia6/Summer11-PU_S4_START42_V11-v1/
qcd 1000 → 1400	/QCD-Pt_1000to1400_TuneZ2_7TeV_pythia6/Summer11-PU_S4_START42_V11-v1/
qcd 1400 → 1800	/QCD-Pt_1400to1800_TuneZ2_7TeV_pythia6/Summer11-PU_S4_START42_V11-v1/
qcd 1800 → ∞	/QCD-Pt_1800_TuneZ2_7TeV_pythia6/Summer11-PU_S4_START42_V11-v1/
qcd 20 → 30 EM	/QCD_Pt-20to30_EMEnriched_TuneZ2_7TeV-pythia6/Summer11-PU_S4_START42_V11-v1/
qcd 30 → 80 EM	/QCD_Pt-30to80_EMEnriched_TuneZ2_7TeV-pythia6/Summer11-PU_S4_START42_V11-v1/
qcd 80 → 120 EM	/QCD_Pt-80to120_EMEnriched_TuneZ2_7TeV-pythia6/Summer11-PU_S4_START42_V11-v1/

4 Physics Object Reconstruction

4.1 Jet Reconstruction

This analysis uses corrected PFJets that are reconstructed using the Anti-kT clustering algorithm with a reconstruction cone of R = 0.5. More details on the jet objects can be found in CMS AN-11-164 [11]

4.2 Electron Reconstruction and Identification

In this analysis, electrons are used to validate tau identification by using the $e\tau_h$ final state to obtain a clean sample of $Z \to \tau\tau$ events. Electrons are also used to validate the performance of the HLT_PFMHT150 trigger. Electrons are energy superclusters in the ECAL detector matched to tracks in the silicon strip tracker and to seed hits in the pixel detector. The energy clustering algorithms, “Hybrid” for the barrel and “Island” for the endcaps, are used to measure the energy of electrons and photons while the the track that best matches the reconstructed energy of the object in the Ecal is chosen to be the reconstructed track. Table 4 shows the list of electron identification (eID) cuts applied. In addition, electrons are also required to be isolated in the Calorimeter and the tracker. Calorimeter isolation requires Ecal RecHits to have an energy threshold of > 0.1 GeV inside a cone of $\Delta R = 0.4$ in the endcaps and > 0.08 GeV for the barrel. Track isolation is defined as the sum of the p_T of tracks inside a cone of $\Delta R = 0.4$, each track with a p_T threshold of 0.7 GeV for both, barrel and endcaps. More details on electron objects used in this analysis can be found in CMS AN-11-164 [11].