CMS Physics Analysis Summary

Contact: cms-pag-conveners-susy@cern.ch 2015/12/15

Search for new physics in the all-hadronic final state with the MT2 variable

The CMS Collaboration

Abstract

A search for new physics is performed using events with jets and the M_{T2} variable, which is a measure of the transverse momentum imbalance in an event. Results are based on a sample of proton-proton collisions collected at a center-of-mass energy of 13 TeV with the CMS detector and corresponding to an integrated luminosity of 2.2 fb^{-1}. No excess above the standard model background is observed. The results are interpreted as limits on the masses of potential new colored particles in a variety of simplified models of supersymmetry.

1 Introduction

Searches for new physics based on final states with jets and large transverse momentum imbalance are sensitive to broad classes of new physics models, including supersymmetry. Such searches were previously conducted by both the ATLAS [1, 2] and CMS [3–7] collaborations using datasets of 8 TeV proton-proton (pp) collisions. No excess above Standard Model (SM) expectations was observed and the results of these earlier searches were used to place lower limits on the masses of pair produced colored particles near the TeV scale for a broad range of production and decay scenarios. These searches provide some of the most stringent constraints on the production of supersymmetric particles and are particularly interesting at this time as they will be some of the earliest and largest beneficiaries of the increase in LHC beam energy.

In this note we present the result of a search for new physics in events with jets and significant transverse momentum imbalance. The search was performed on a dataset corresponding to an integrated luminosity of 2.2 fb^{-1} of pp collisions collected at a center-of-mass energy of 13 TeV with the Compact Muon Solenoid (CMS) detctor at the LHC. This search builds upon previous searches at 7 and 8 TeV with a similar final state selection [3, 8].

In this analysis, we select events with at least one jet and veto events with an identified lepton. The number of observed events in signal regions defined by the number of jets, the number of b-tagged jets, the scalar sum of jet transverse momenta (H_T), and the stransverse mass (M_{T2}) [9] is compared with the SM expectation. We observe no evidence for an excess above the expected background from SM processes and interpret the results as limits on the production of gluino pairs using simplified models of supersymmetry.

2 Event selection and Monte Carlo simulation

Event reconstruction is based on the particle-flow (PF) algorithm [10, 11], which combines information from the tracker, calorimeter, and muon systems to reconstruct and identify charged and neutral hadrons, photons, muons and electrons. To select collision events we require at least one reconstructed vertex. If more than one vertex is reconstructed, we designate as the primary vertex the one for which the sum p_{T}^2 of the associated particle-flow candidates is largest. Neutral particles and charged particle-flow candidates associated to the primary vertex are clustered into jets using the anti-k$_T$ algorithm [12] with a size parameter 0.4.

The jet energy is calibrated using the standard set of corrections, similar to the procedure used during the 8 TeV run [13]: L1, an offset correction accounting for neutral energy arising from both in-time and out-of-time pile-up interactions and the area of the reconstructed jet; L2, a relative correction that makes the jet response uniform in p_T and η; L3, an absolute correction that restores the average jet energy response to unity; and a residual L2L3 correction, applied to account for remaining data/MC disagreement.

Jets originating from b-quarks are identified by the combined secondary vertex (CSV) algorithm. We use a working point with a tagging efficiency of approximately 65% for jets originating from b-quarks typical in top pair events and a corresponding mis-tag rate for jets originating from light flavor quarks of about 1.5% [14]. The transverse hadronic energy, H_T, is defined as the scalar sum of the jet transverse momenta while the missing transverse hadronic momentum, $\vec{H}_T^{\text{miss}}$, is obtained from the negative vector sum of the momenta of the same jets. The transverse momentum imbalance ($\vec{E}_T^{\text{miss}}$), whose magnitude $E_{\mathrm{T}}^{\text{miss}}$ is commonly referred to as MET, is defined as the negative vector sum of all reconstructed charged and neutral particle-flow candidates.

Electron candidates are reconstructed as a cluster of energy deposits in the electromagnetic calorimeter matched to signals in the silicon tracker. We identify electrons by loose requirements on the shape of these energy deposits, on the ratio of energy in associated hadron and electromagnetic calorimeter cells (H/E), on the geometric matching between the energy deposits and the associated track, and on consistency between the energy reconstructed from calorimeter deposits and the momentum measured in the tracker [15]. We additionally require that the associated track is consistent with originating from the primary vertex.

Muon candidates are reconstructed by identifying signals from the CSC, DT and RPC subsystems together with those from the silicon tracker. Candidates are required to be identified as either global or tracker muons [16] as well as muons by the particle-flow algorithm. The associated silicon track is required to be consistent with originating from the primary vertex.

Lepton isolation is determined using a cone size dependent on the lepton p_T according to Eq. 1. The isolation sum is calculated using neutral and charged particle-flow candidates and is corrected for the effects of pile-up.

$$\Delta R = \begin{cases} 0.2, & p_{\mathrm{T}}^{\mathrm{lep}} \leq 50 \text{ GeV} \\ \frac{10 \text{ GeV}}{p_{\mathrm{T}}^{\mathrm{lep}}}, & 50 < p_{\mathrm{T}}^{\mathrm{lep}} \leq 200 \text{ GeV} \\ 0.05, & p_{\mathrm{T}}^{\mathrm{lep}} \geq 200 \text{ GeV}, \end{cases} \tag{1}$$

Photon candidates, used in the estimation of the $Z \to \nu\bar{\nu}$ background, are reconstructed from deposits in the electromagnetic calorimeter and selected by making requirements on the shower shape ($\sigma_{i\eta i\eta}$), the ratio H/E, and by requiring that the sum p_T of charged hadrons is less than 2.5 GeV within an isolation cone of $\Delta R < 0.3$.

2.1 Baseline event selection

We select events with $E_T^{\mathrm{miss}} > 30$ GeV and count the number of jets with $p_T > 30$ GeV, $|\eta| < 2.5$, passing loose requirements on the jet composition designed to reject hadronic objects arising from noise and failures of the event reconstruction [17]. When counting the number of jets identified as originating from b-quarks, we relax the p_T threshold to 20 GeV to increase sensitivity to potential signal scenarios with a compressed mass spectrum. For events with at least two reconstructed jets, we start with the pair having the largest di-jet invariant mass and iteratively cluster all selected jets using a hemisphere algorithm that minimizes the Lund distance measure [18, 19] until two stable pseudo-jets are obtained. The resulting psuedo-jets together with the E_T^{miss} are used to determine the stransverse mass. A detailed discussion of the M_{T2} variable and its properties as used in this and previous searches can be found in [8].

To suppress the background from SM multi-jet production, we require an M_{T2} value of at least 200 GeV in events with at least two reconstructed jets. Even after this requirement, a residual component with large M_{T2} value, arising primarily from jet under-measurement, remains. To protect against jet mis-measurement, we thus require that the minimum difference $\Delta\phi_{\mathrm{min}}$ in azimuthal angle between the $\vec{E}_T^{\mathrm{miss}}$ vector and each of the leading four jets is greater than 0.3 radians. For the purpose of the $\Delta\phi_{\mathrm{min}}$ calculation only, we consider jets up to $|\eta| < 4.7$. Additionally, we require that the magnitude of the vector difference in the transverse momentum imbalance determined using jets and all particle flow candidates is not too large: $\frac{|\vec{E}_T^{\mathrm{miss}} - \vec{H}_T^{\mathrm{miss}}|}{E_T^{\mathrm{miss}}} < 0.5$. The latter protects against large imbalances arising from reconstructed objects with $p_T <$

30 GeV or $|\eta| > 2.5$. Finally, events with possible contributions from beam halo processes or anomalous noise in the calorimeter are rejected using dedicated filters [20].

To reduce the background from SM processes with genuine E_T^{miss} arising from the decay of a W boson, we reject events with an identified electron or muon with $p_T > 10$ GeV, $|\eta| < 2.4$. Only electrons (muons) with a relative isolation, as defined by Eq. 1, less than 0.1 (0.2) are considered in the veto. An event is also vetoed if it contains an isolated, charged particle-flow candidate. If the charged candidate is identified as a PF electron or muon, we veto the event if the candidate has $p_T > 5$ GeV, $|dz| < 0.1$ with respect to the PV, and a relative track isolation of less than 0.2. If the charged candidate is identified as a PF hadron, we veto the event only if the candidate has $p_T > 10$ GeV, $|dz| < 0.1$, and a relative track isolation of less than 0.1. In all cases, the track isolation is calculated using all charged particle-flow candidates within a cone $\Delta R < 0.3$ and $|dz| < 0.1$ with respect to the PV. To avoid loss of efficiency in potential signals with large jet multiplicities, we only veto events if the transverse mass formed by the isolated, charged PF candidate and the E_T^{miss} is less than 100 GeV.

Collision events are selected using triggers with varied requirements on the H_T, E_T^{miss} and H_T^{miss}. Table 1 summarizes the triggers and corresponding offline selections used in this analysis, after which the triggers are found to be $> 98\%$ efficient.

Table 1: The three signal triggers and the corresponding offline selections.

online trigger selection	offline selection
H_T >800 GeV	H_T >1000 GeV & E_T^{miss} >30 GeV
H_T >350 GeV & E_T^{miss} >100 GeV	H_T >450 GeV & E_T^{miss} >200 GeV
E_T^{miss} >90 GeV & H_T^{miss} >90 GeV & 1 jet passing Tight ID	H_T >200 GeV & E_T^{miss} >200 GeV

2.2 Signal regions

Signal regions are defined separately for events with exactly one jet passing the counting criteria above and those with two or more. Events with at least two jets are categorized based on the H_T, N_j, N_b as follows:

- 5 bins in H_T: [200,450], [450, 575], [575, 1000], [1000, 1500], [1500,∞]
  These bins are also referred to as very low H_T, low H_T, medium H_T, high H_T, and extreme H_T regions.
- 11 bins in N_j and N_b: 2–3j 0b, 2–3j 1b, 2–3j 2b, 4–6j 0b, 4–6j 1b, 4–6j 2b, $\geq$7j 0b, $\geq$7j 1b, $\geq$7j 2b, 2–6j $\geq$3b, $\geq$7j $\geq$3b

Each bin defined by the H_T, N_j, N_b requirements above is referred to as a *topological region* and we further divide each topological region in bins of M_{T2}.

- 3 bins in M_{T2} at Very Low H_T: [200,300], [300,400], [400,∞]
- 4 bins in M_{T2} at Low H_T: [200,300], [300,400], [400,500], [500,∞]
- 5 bins in M_{T2} at Medium H_T: [200,300], [300,400], [400,600], [600,800], [800,∞]
- 5 bins in M_{T2} at High H_T: [200,400], [400,600], [600,800], [800, 1000], [1000,∞]
- 5 bins in M_{T2} at Extreme H_T: [200,400], [400,600], [600,800], [800,1000], [1000,∞]

For events with $\text{N}_\text{j} == 1$, the M_{T2} variable is not defined. We instead opt for a simpler strategy with signal regions defined by the p_T of the jet and N_b:

- N_b: 0b, $\geq$1b
- 7 bins in jet p_T: [200,250], [250,350], [350,450], [450,575], [575,700], [700,1000], [1000,∞]

Within each topological or N_b categorization, we merge M_{T2} or jet p_T bins, respectively, which are expected to contain less than one background event. All the edges of the previous H_T, M_{T2}, and p_T signal regions are expressed in GeV.

2.3 Monte Carlo simulation

Monte Carlo (MC) simulations including pile-up are used in the estimate of some of the SM backgrounds (see Section 3), as well as to calculate the efficiency for various new physics scenarios. In this analysis, SM background samples and signal samples of gluino and squark pair production are generated with Madgraph V5 [21] interfaced with Pythia V8.1 [22]. Standard model processes are simulated using a Geant4-based model [23] of the CMS detector. The simulation of new physics signals is performed using the CMS fast simulation package [24]. All simulated events are processed with the same chain of reconstruction programs used for collision data.

3 Backgrounds

There are three sources of SM background to potential new physics signals in a jets plus E_T^{miss} final state:

- "lost lepton", i.e. events with a lepton from a W decay where the lepton is either out of acceptance, not reconstructed, not identified, or not isolated. This background comes from both W+jets and $t\bar{t}$+jets events and thus is one of the dominant backgrounds in nearly all search regions. It is estimated using a one lepton control sample obtained by inverting the lepton veto in each topological region.
- "irreducible", i.e. Z+jets events where the Z boson decays to neutrinos. This background is most similar to potential signals. It is a major background in nearly all search regions, its importance decreasing with the N_b requirement. This background is estimated using a γ+jets control sample for each topological region.
- "instrumental background", i.e. mostly QCD multi-jet events with no real E_T^{miss}. These events enter a search region due to either significant jet momentum over-measurement, under-measurement, or sources of anomalous noise. To suppress this background we make the selections described in Section 2.1 after which it is subdominant in almost all search regions. The background is estimated from a control sample obtained by inverting the $\Delta\phi_{\text{min}}$ requirement in each topological region.

3.1 Estimation of the background from leptonic W boson decays

Single lepton control regions are used to estimate this background, arising predominantly from W+jets and $t\bar{t}$+jets processes. Control region events are selected using the same triggers as for signal regions and the baseline selections of Section 2.1 are applied with the exception of the lepton veto. Instead we require exactly one candidate passing the veto lepton or PF lepton selections (e or μ only). We additionally require $M_T(\text{cand}, E_{\text{T}}^{\text{miss}}) < 100$ GeV to reduce potential signal contamination.

Selected events are further subdivided into the categories described in Section 2.2, binning the single lepton control regions in the H_T, N_j, and N_b dimensions but not in M_{T2} to preserve statistics. The binning in N_j and N_b is the same as the signal regions except for signal bins with $\geq$7j,$\geq$1b. These are all predicted using control region bins with the same H_T selection as the signal bin and $\geq$7j,1–2b. This is motivated by the low control region statistics in bins with

$\geq$7j,$\geq$2b as well as potential signal contamination in bins with $\geq$7j,$\geq$3b. For events with $N_j = 1$, one control region is defined for each bin of jet p_T.

The background in each signal bin is obtained from the corresponding single lepton control region by the application of transfer factors according to the following equation:

$$N_{1\ell}^{\mathrm{SR}}(M_{T2}) = N_{1\ell}^{\mathrm{CR}}(H_T, N_j, N_b) \times R_{\mathrm{MC}}^{O\ell/1\ell}(H_T, N_j, N_b) \times k_{\mathrm{MC}}(M_{T2}) \quad (2)$$

The transfer factors are obtained from MC and corrected for measured differences between data and simulation. The population of events for which we fail to reconstruct or identify an isolated lepton candidate is obtained via the factor $R_{\mathrm{MC}}^{O\ell/1\ell}(H_T, N_j, N_b)$, which accounts for lepton acceptance and efficiency and the expected contribution from the decay of W bosons to hadrons through an intermediate τ lepton. The fraction of events in each topological region expected to populate a particular M_{T2} bin, $k_{\mathrm{MC}}(M_{T2})$, is used to obtain the estimate in each search bin.

The transfer factors used in the background estimate are obtained from simulation. Normalization to data control regions reduces reliance on the MC modeling of most kinematic quantities except notably M_{T2}. The uncertainty on the fraction of events in a particular M_{T2} bin is evaluated in simulation by variations of the most important experimental and theoretical quantities. Reconstruction uncertainties, assessed by varying the tagging efficiency for b-quarks and by evaluating the impact of variations in jet response on the counting of jets and b-tagged jets, E_T^{miss} and M_{T2}, are typically found to be less than 10%, but can reach as much as 40% in some bins. Variations in the renormalization and factorization scales, PDF sets and the relative composition of W+jets and $t\bar{t}$+jets are performed to assess the dominant theoretical uncertainties which are found to be as large as 30%.

The MC modeling of M_{T2} is checked in data using high statistics control regions enriched in events originating from either W+jets or $t\bar{t}$+jets as shown in the left and right panels of Figure 1, respectively. An additional check is performed by comparing the standard estimate with that obtained by replacing the factor $k_{\mathrm{MC}}(M_{T2})$ with an extra dimension in the binning of the control region: $N_{1\ell}^{\mathrm{CR}}(H_T, N_j, N_b, M_{T2})$. The two estimates agree to the extent permitted by the available control region statistics.

The single lepton control region typically has 1–2 times as many events expected as compared to the corresponding signal region. The statistical uncertainty on this yield ranges from 1–100%, depending on the region, and is propagated to the final estimate. The transfer factor $R_{\mathrm{MC}}^{O\ell/1\ell}$ also depends on the MC modeling of the lepton veto and M_T cut efficiencies. Tag and probe with $Z \to \ell^+\ell^-$ events is used to evaluate the MC modeling of lepton efficiencies and the resulting uncertainty propagated to the background estimate is found to be as large as 10%, representing the dominant uncertainty on the factor $R_{\mathrm{MC}}^{O\ell/1\ell}$. The M_T cut efficiency is cross-checked using a similar sample and removing one of the leptons to mimic events where the W decays to a lepton. Finally, the uncertainty on the MC modeling of the lepton acceptance, assessed by varying the renormalization and factorization scales and PDF sets as well as the b-tagging efficiency and the jet energy scale, is typically less than 10% although it can be as large as 40% in some bins.

3.2 Estimation of the background from $Z(\nu\bar{\nu})$+jets

The $Z \to \nu\bar{\nu}$ background is obtained from a γ+jets control sample selected using a single photon trigger. We select events where the photon has $p_T > 180$ GeV, $|\eta| < 2.5$ to mimic

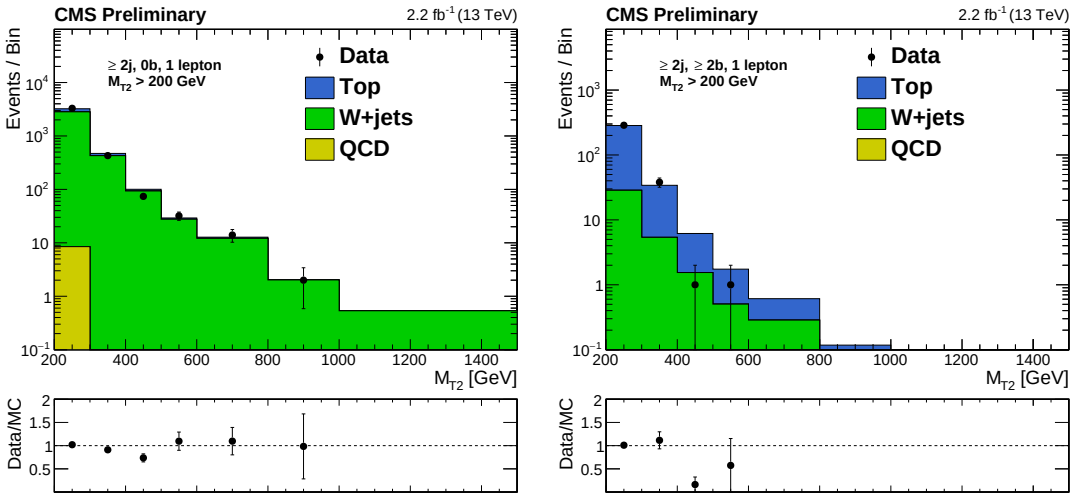

Figure 1: Shape comparison between simulation and data for the M_{T2} observable. The left and right panels correspond to W+jets and $t\bar{t}$+jets enriched control samples, respectively.

the implicit requirement on the p_T of the Z boson arising from the baseline selection $M_{T2} > 200$ GeV. The full baseline selection requirements are made based on kinematic variables re-calculated after removing the photon from the event to replicate the $Z \to \nu\bar{\nu}$ kinematics.

Adopting a similar strategy as that used for the estimate of the lost lepton background, selected events are further sub-divided into the categories described in Section 2.2, binning the photon control regions in the H_T, $\mathrm{N_j}$, and $\mathrm{N_b}$ dimensions but not in M_{T2} to preserve statistics. For events with $\mathrm{N_j} = 1$, one control region is defined for each bin of jet p_T. The background in each signal bin is obtained by the application of transfer factors according to Eq. 3:

$$N^{\mathrm{SR}}_{Z\to\nu\bar{\nu}}\left(M_{T2}\right) = N^{\mathrm{CR}}_{\gamma}\left(H_T, \mathrm{N_j}, \mathrm{N_b}\right) \times P_{\gamma} \times f \times R^{Z/\gamma}_{\mathrm{MC}}\left(H_T, \mathrm{N_j}, \mathrm{N_b}\right) \times k_{\mathrm{MC}}\left(M_{T2}\right) \tag{3}$$

The prompt photon purity, P_γ, is measured in data by performing a template fit of the charged hadron isolation distribution for each H_T, $\mathrm{N_j}$, and $\mathrm{N_b}$ region. The signal shape is obtained from data by measuring the charged hadron isolation in random cones well-separated from the photon and any jets. The background shape also comes from data using a sideband in $\sigma_{i\eta i\eta}$. A prompt photon purity of 90–100% as measured in data is well-reproduced by simulation as seen in the left panel of Figure 2. A separate determination of the prompt photon purity using a tight-to-loose ratio method obtained from the charged isolation sideband is found to yield consistent results.

The $Z \to \nu\bar{\nu}$ background in each bin of H_T, $\mathrm{N_j}$, and $\mathrm{N_b}$ is obtained from the corresponding photon control region yield via the factor $R^{Z/\gamma}_{\mathrm{MC}}$, which accounts for the photon acceptance and efficiency and the relative cross sections for the production of Z+jets and γ+jets events. The ratio $R^{Z/\gamma}_{\mathrm{MC}}$ is obtained from simulation and validated in data using $Z \to \ell^+\ell^-$ events. The right panel of Figure 2 shows the double ratio, $R^{Z\to\ell\ell/\gamma}_{\mathrm{data}} / R^{Z\to\ell\ell/\gamma}_{\mathrm{MC}}$, in bins of H_{T}. The double ratio shows no significant trend and a correction factor of 0.95 is applied to $R^{Z/\gamma}_{\mathrm{MC}}$ to account for the observed deviation from unity. The fraction of events in each topological region expected to populate a particular M_{T2} bin, $k_{\mathrm{MC}}\left(M_{T2}\right)$, is used to obtain the estimate in each search bin.

The ratio $R^{Z/\gamma}_{\mathrm{MC}}$ is obtained from γ+jet events simulated with Madgraph with an implicit re-

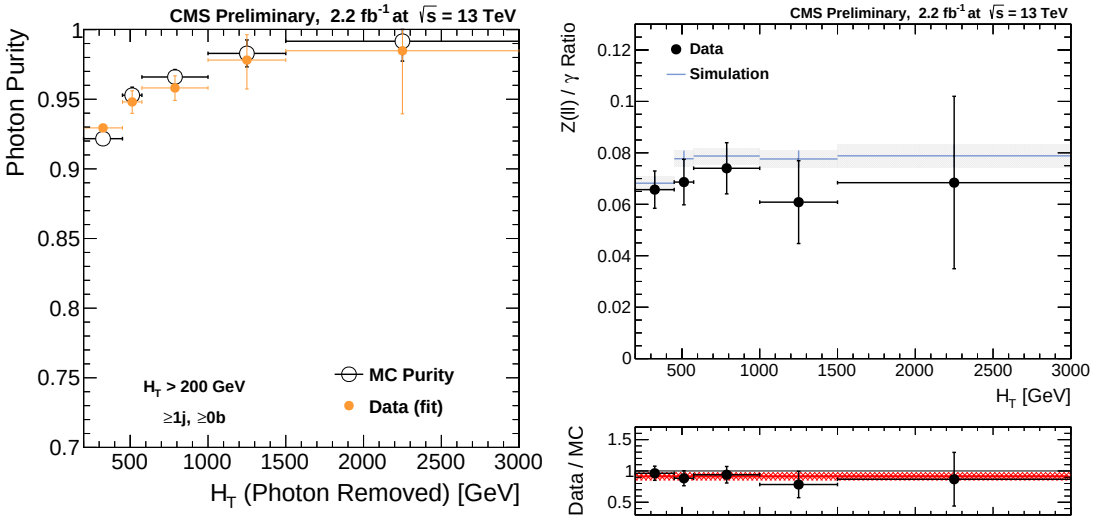

Figure 2: The left plot shows the photon purity measured in data for the single photon control sample compared with the values extracted from simulation. The right plot shows $R(Z/\gamma)$ the Z/γ ratio in simulation and data as a function of H_{T}, and the corresponding double ratio (bottom panel).

quirement $\Delta R > 0.4$ between the prompt photon and the nearest parton. As no such requirement can be made in data, a correction factor $f = 0.92$ as determined from studies with Madgraph+Pythia and Pythia QCD samples is applied to account for the fraction of selected photons passing the ΔR requirement.

As for the estimate of the single lepton background, normalization to data control regions reduces reliance on the MC modeling to a single dimension, M_{T2}. The uncertainty on the fraction of events in a particular M_{T2} bin is evaluated in simulation by variations of the most important experimental and theoretical variables. Theoretical uncertainties represent the largest contribution and are assessed by variations of the renormalization and factorization scales and PDF sets. Smaller contributions from reconstruction uncertainties are determined by varying the tagging efficiency for b-quarks and by evaluating the impact of variations in jet response on the counting of jets and b-tagged jets, E_T^{miss} and M_{T2}. Experimental and theoretical uncertainties on $k_{\mathrm{MC}}\,(M_{T2})$ total as much as 30% at large values of M_{T2}.

The MC modeling of M_{T2} is checked in data using high statistics control samples of γ+jets and $W \to \ell\nu$ events. Figure 3 shows good agreement between the M_{T2} distribution obtained from γ and W data control samples with that from $Z \to \nu\bar{\nu}$ MC in the medium H_{T} and high H_{T} regions. In this comparison, the γ sample is corrected based on P_γ, f and $R_{\mathrm{MC}}^{Z/\gamma}$, while the W sample is corrected for top contamination and $R_{\mathrm{MC}}^{Z/W}$. An additional check is performed by comparing the standard estimate with that obtained by replacing the factor $k_{\mathrm{MC}}\,(M_{T2})$ with an extra dimension in the binning of the control region: $N_\gamma^{\mathrm{CR}}\left(H_T, \mathrm{N_j}, \mathrm{N_b}, M_{T2}\right)$. The two estimates agree to the extent permitted by the available control region statistics.

The single photon control region typically has 2–3 times as many events expected as compared to the corresponding signal region. The statistical uncertainty on this yield ranges from 1–100%, depending on the region, and is propagated to the final estimate. The dominant uncertainty on the MC modeling of $R_{\mathrm{MC}}^{Z/\gamma}$ comes from the validation of the ratio using $Z \to \ell^+\ell^-$ events. One-dimensional projections of the double ratio are constructed – separately in bins of the number of jets (Figure 2, right), number of b-tagged jets and H_T – and an uncertainty on $R_{\mathrm{MC}}^{Z/\gamma}$ in each

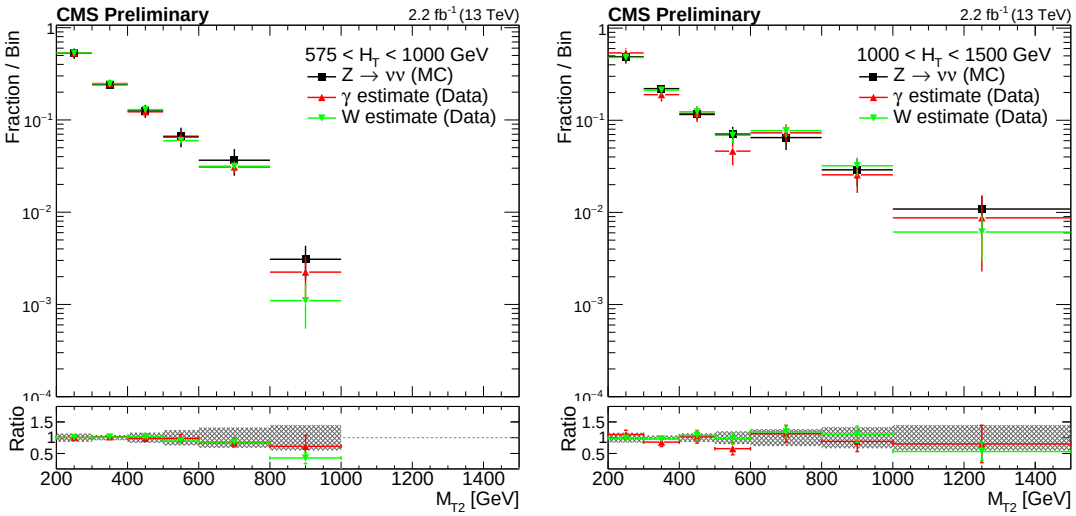

Figure 3: The shape of the M_{T2} distribution from $Z \to \nu\bar{\nu}$ simulation compared to shapes extracted from γ and W data control samples in the medium H_T and high H_T regions.

bin of H_T, N_j, and N_b is determined by adding in quadrature the corresponding uncertainty on the ratio $R^{Z\to\ell\ell/\gamma}$ in background subtracted data. As insufficient statistics are available to evaluate the double ratio for regions with $N_b \geq 3$, we conservatively assign twice the uncertainty obtained for the corresponding $N_b = 2$ bin. This uncertainty ranges from 10–100% depending on the search region. An additional 11% uncertainty on the transfer factor based on the observed offset of the double ratio from unity is added in quadrature with the above.

The uncertainty on the measurement of the prompt photon purity includes a statistical contribution from yields in the isolation sideband that is typically 5–10% but can reach as much as 100% for extreme search regions. Additional uncertainties due to variations in purity from modifications of the signal and background templates and non-closure of the method observed in tests with simulation amount to another 5%. Finally, an uncertainty of 8% is assigned to cover differences in the correction fraction f observed between Madgraph+Pythia and Pythia-only simulation.

3.3 Estimation of the multi-jet background

The multi-jet background consists predominantly of QCD events, but also includes contributions from the production of top-quark pairs or vector bosons decaying to hadrons. Though this background is expected to be small after requiring $M_{T2} > 200$ GeV, we have developed a method based on data control samples to estimate any residual background. For events with at least two jets, a mutli-jet enriched control region is obtained in each H_T bin by inverting the $\Delta\phi_{\text{min}}$ requirement described in Section 2.1. For the high and extreme H_T bins, control region events are selected using the same trigger as for signal events. For lower H_T regions, the online E_T^{miss} requirement precludes the use of the signal trigger and the control sample is instead selected using pre-scaled H_T triggers.

The extrapolation from low to high $\Delta\phi_{\text{min}}$ is based on the following ratio:

$$r_\phi(M_{T2}) = N(\Delta\phi_{\text{min}} > 0.3)/N(\Delta\phi_{\text{min}} < 0.3). \quad (4)$$

Studies in simulation show the ratio to be well described by a power law function, $a \cdot M_{T2}^b$. The parameters a, b are determined separately in each H_T bin by fitting the ratio $r_\phi(M_{T2})$ in

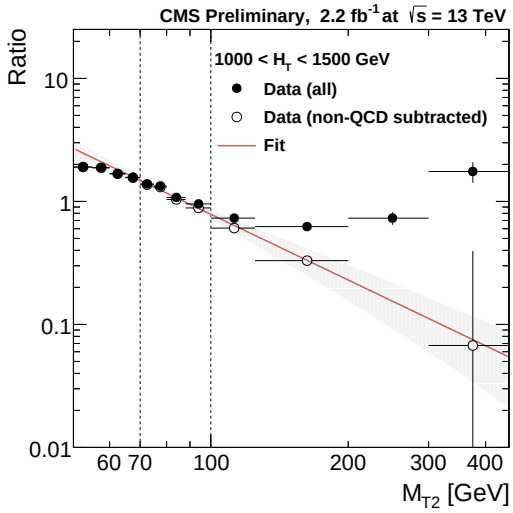

Figure 4: Distribution of the ratio r_ϕ as a function of M_{T2} for the high H_T region. The fit is performed to the hollow, background-subtracted data points. The full points represent the data before subtracting non-QCD backgrounds using simulation. Data point uncertainties are statistical only. The red line and the band around it show the fit to a power-law function perfomed in the window $70 < M_{T2} < 100$ GeV and the associated fit uncertainty.

a sideband in data, i.e. $60 < M_{T2} < 100\,\text{GeV}$, after subtracting non-QCD contributions using simulation. For the high and extreme H_T regions, the fit is performed in a slightly narrower M_{T2} window with the lower edge increased to 70 GeV. An example in the high H_T region is shown in Figure 4. The inclusive multi-jet contribution in each H_T region, $N_{\text{inc}}^{SR}(M_{T2})$, is estimated using the ratio $r_\phi(M_{T2})$ measured in the M_{T2} sideband and the number of events in the low $\Delta\phi_{\text{min}}$ control region, $N_{\text{inc}}^{CR}(M_{T2})$:

$$N_{\text{inc}}^{SR}(M_{T2}) = N_{\text{inc}}^{CR} \cdot r_\phi(M_{T2}). \tag{5}$$

From the inclusive multi-jet estimate in each H_T region, the predicted background in bins of N_j, N_b is obtained from the following equation

$$N_{j,b}^{SR}(M_{T2}) = N_{\text{inc}}^{SR}(M_{T2}) \cdot f_j(H_T) \cdot r_b\left(\text{N}_\text{j}\right), \tag{6}$$

where f_j is the fraction of QCD events falling in bin N_j, and r_b is the fraction of all events in bin N_j that fall in bin N_b. Simulation indicates that f_j and r_b attain similar values in low and high $\Delta\phi_{\text{min}}$ regions and that the values are independent of M_{T2}. We take advantage of this to measure the values of f_j, r_b using events with M_{T2} between 100–200 GeV in the low $\Delta\phi_{\text{min}}$ sideband, where f_j is measured separately in each H_T bin while r_b is measured in bins of N_j, integrated over H_T, as r_b is found to be independent of the latter. Values of f_j, r_b measured in data are shown in Figure 5 compared to simulation.

An estimate based on $r_\phi(M_{T2})$ is not viable in the mono-jet search region so a different strategy must be employed. A control region is obtained by asking for a second jet with $30 < p_T < 60$ GeV and inverting the $\Delta\phi_{\text{min}}$ requirement. After subtracting non-QCD contributions using simulation, the data yield in the control region is taken as an estimate of the background in the mono-jet search regions. As the likeliness of jet under-measurement is a steeply falling function of the degree of mis-measurement, the estimate obtained in this way is expected to be conservative. Closure tests in MC indicate a small over-estimation but the multi-jet background is not expected to exceed 8% in any mono-jet search region.

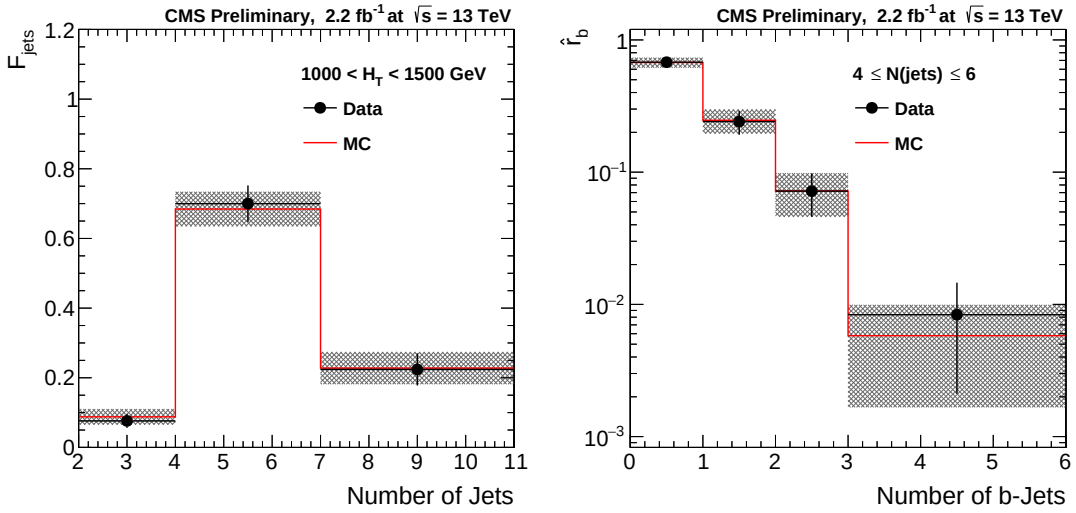

Figure 5: Values of f_j (left) and r_b (right) measured in data after requiring $\Delta\phi_{\text{min}} < 0.3$ radians and $100 < M_{T2} < 200$ GeV. The bands represent both statistical and systematic uncertainties.

Statistical uncertainties due to the event yields in the control regions, where the $r_\phi(M_{T2})$ fit is performed and the f_j and r_b values are measured, are propagated to the final estimate. A variation of 50% on the non-QCD background that is subtracted from the control region is propagated to the final uncertainty on $r_\phi(M_{T2})$. The invariance of f_j with M_{T2} and r_b with M_{T2} and H_T is evaluted in simulation and residual differences are taken as additional uncertainties. An additional uncertainty is assigned to cover the sensitivity of the r_ϕ value to variations in the fit window. These variations result in an uncertainty that increases with M_{T2} and ranges from 15-200%. The total uncertainty on the estimate covers the differences observed in closure tests based on simulation and in a validation based on data. The latter is performed in the $100 < M_{T2} < 200$ GeV sideband. For the mono-jet regions, the statistical uncertainty from the data yield in the trailing jet p_T sideband is combined with a 50% systematic uncertainty in all bins.

3.4 Alternative multi-jet background estimation

As a cross-check, the multi-jet background is also estimated using the "rebalance and smear" (R&S) method described in Ref. [25]. The method is not yet well understood for kinematical regions where the dominant multi-jet background contribution is from lost jets, and therefore it is not used for the mono-jet selection. It is also not applied to the very-low-HT region because of lack of statistics in the relevant QCD data sample due to the high trigger prescales at very low HT.

The performance of the method has been tested in QCD Monte Carlo, as well as data control regions defined by inverting the $\Delta\phi_{\text{min}}$ requirement or by selecting a sideband of M_{T2} (i.e. 100 GeV $< M_{T2} <$200 GeV). Based on these studies, we assign total systematic uncertainties of 50% (low-HT and medium-HT) and 40% (high-HT and extreme-HT) on the R&S background predictions for the M_{T2} >200 GeV selections. These uncertainties also include a small (<7%) uncertainty due to the contamination from EWK events to the QCD data sample used in the R&S procedure.

In Figure 6 we compare the QCD predictions from the R&S method and the method described in Section 3.3, which is the default multi-jet estimation method used in this analysis. The level of agreement between the two methods serves as a further validation of the QCD estimation.

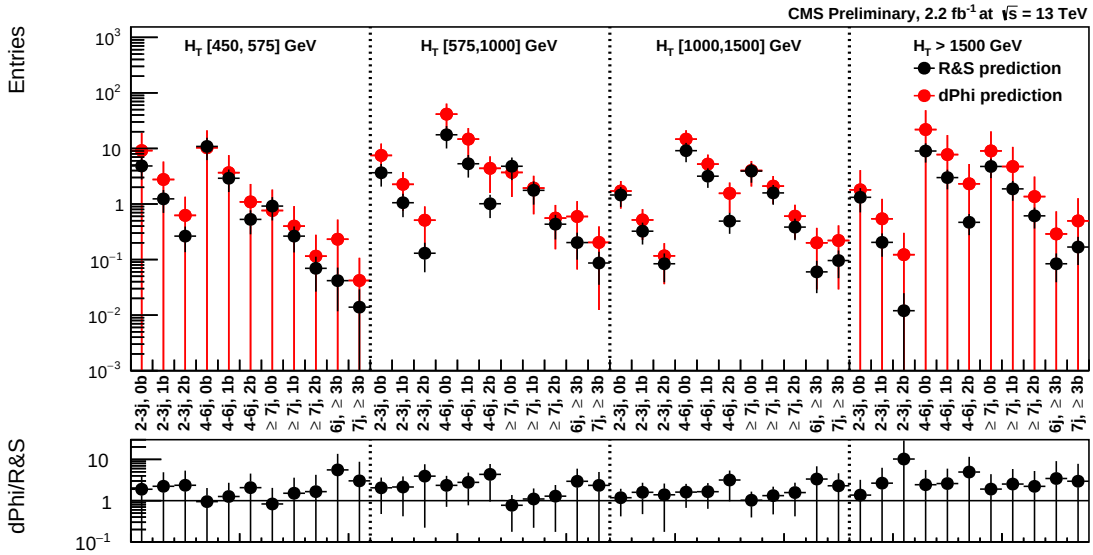

Figure 6: Comparison of the data-driven predictions of the multi-jet background in the topological regions (M_{T2} >200 GeV) from the R&S method and the $\Delta\phi_{\min}$-ratio method. The uncertainties are statistical and systematic. Within each of the four H_T categories, all the estimates from the $\Delta\phi_{\min}$-ratio method are correlated because they are derived from the same fit to the $\Delta\phi_{\min}$-ratio data.

4 Results

A summary of the results of this search is shown in Figure 7. No significant deviations are observed from standard model expectations. Each bin in the upper panel corresponds to a single H_T, N_j, N_b search region integrated over M_{T2}. The lower panel further breaks down the background estimates and observed data yields into M_{T2} bins for the medium H_T region. Summary distributions for the remaining H_T regions can be found in Appendix A.

4.1 Interpretation

The results of the search may be used to constrain specific models of new physics such as those shown in Figure 8.

For a given signal scenario, limits are derived by combining search regions using a modified frequentist approach [26–29]. Typical values of the uncertainties considered on the signal yield are listed in Table 2. The largest uncertainties come from MC statistics, b-tagging efficiency, and for model points with small mass splittings, the recoil "ISR" uncertainty [30]. The uncertainty on acceptance due to the renormalization and factorization scales was found to be relatively small and a flat value of 5% is taken. The uncertainties due to variations of the PDF set and jet energy scale were found to be compatible with statistical fluctuations for bins with lower MC statistics, so flat values of 10% and 5% are taken respectively, motivated by high statistics search bins. Uncertainties due to the luminosity, "ISR" recoil, and b-tagging and lepton efficiencies are treated as correlated across search bins. Figures 9–11 shows exclusion limits at 95% confidence level for gluino mediated bottom-squark, top-squark production, and light-flavor squark production, respectively. The analysis probes gluino masses up to 1725 GeV and the mass of the lightest neutralino (LSP) up to 1100 GeV.

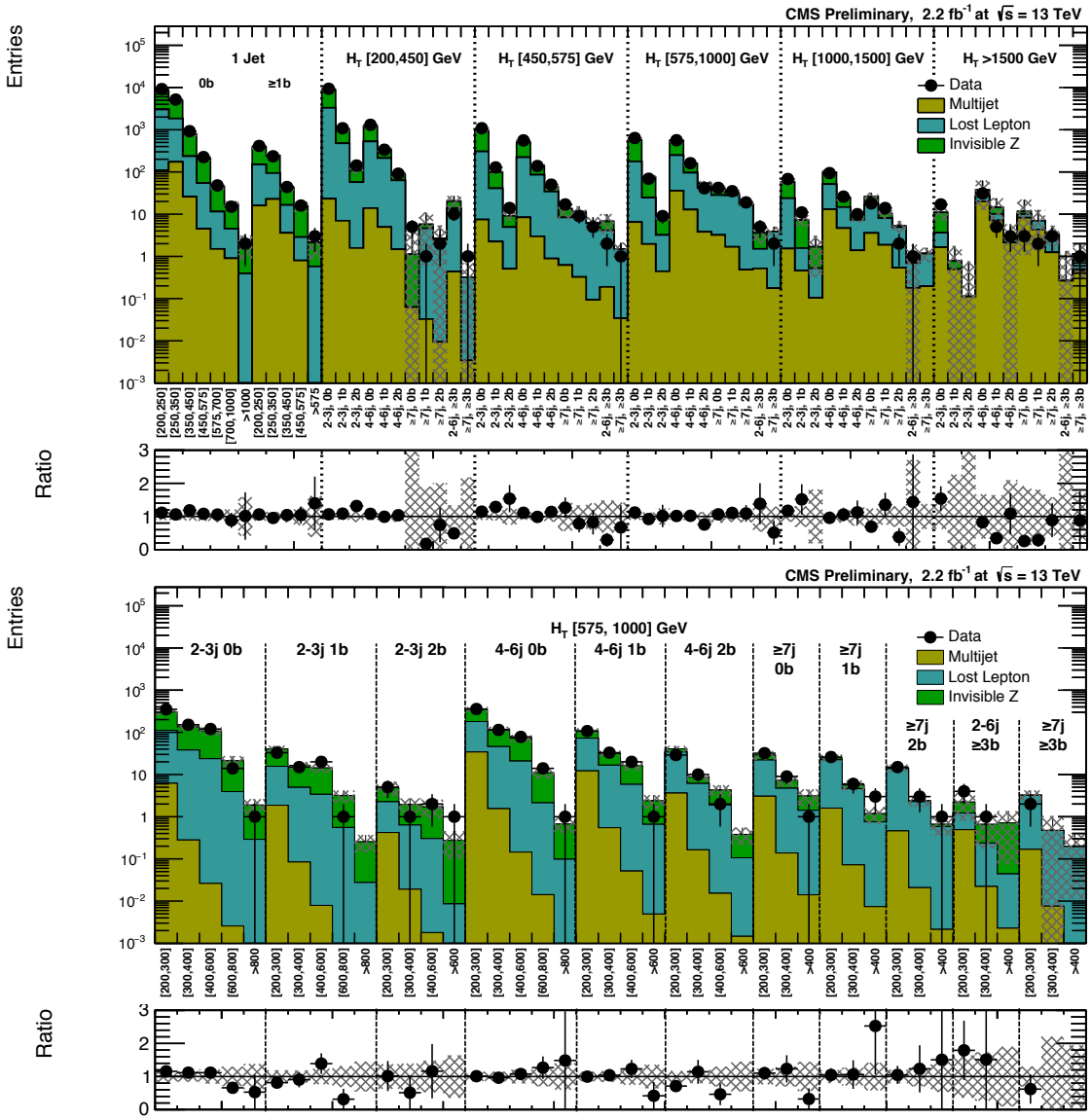

Figure 7: (Above) Comparison of estimated background and observed data events in each topological region. Hatched bands represent the full uncertainty on the background estimate. (Below) Same for individual M_{T2} signal bins in the medium H_T region. On the x-axis, the M_{T2} binning is shown (in GeV). Bins with no entry for data have an observed count of 0.

5 Summary

This note presents the result of a search for new physics using events with jets and the M_{T2} variable. Results are based on a 2.2 fb^{-1} data sample of proton-proton collisions at $\sqrt{s} =$ 13 TeV collected with the CMS detector. No significant deviations from the standard model expectations are observed. Results with this first year of data already extend sensitivity to the production of new, massive colored particles beyond that achieved in Run 1. We probe gluino masses up about 1725 GeV and LSP masses up to 1100 GeV, extending the reach of Run 1 searches by more than 300 GeV.

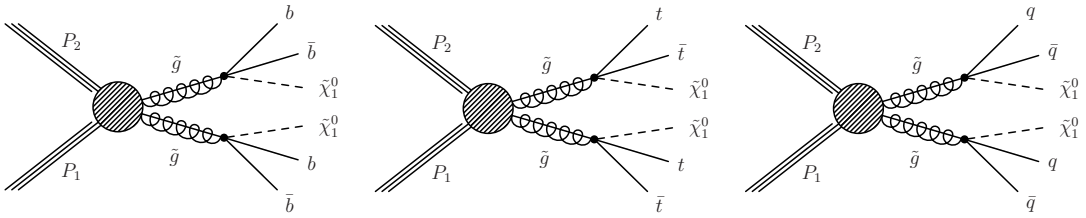

Figure 8: Diagrams for the three scenarios of gluino mediated bottom squark, top squark and light flavor squark production considered.

Table 2: Typical values of the signal systematic uncertainties as evaluated for the T1bbbb signal model. Uncertainties evaluated on other signal models are consistent with these ranges of values.

Source	Typical Values
Luminosity	4.6%
MC statistics	1–100%
Renormalization and factorization scales	5%
Parton distribution functions	10%
"ISR" recoil	0–30%
B-tagging efficiency, heavy flavor	0–40%
B-tagging efficiency, light flavor	0–20%
Lepton efficiency	0–20%
Jet energy scale	5%

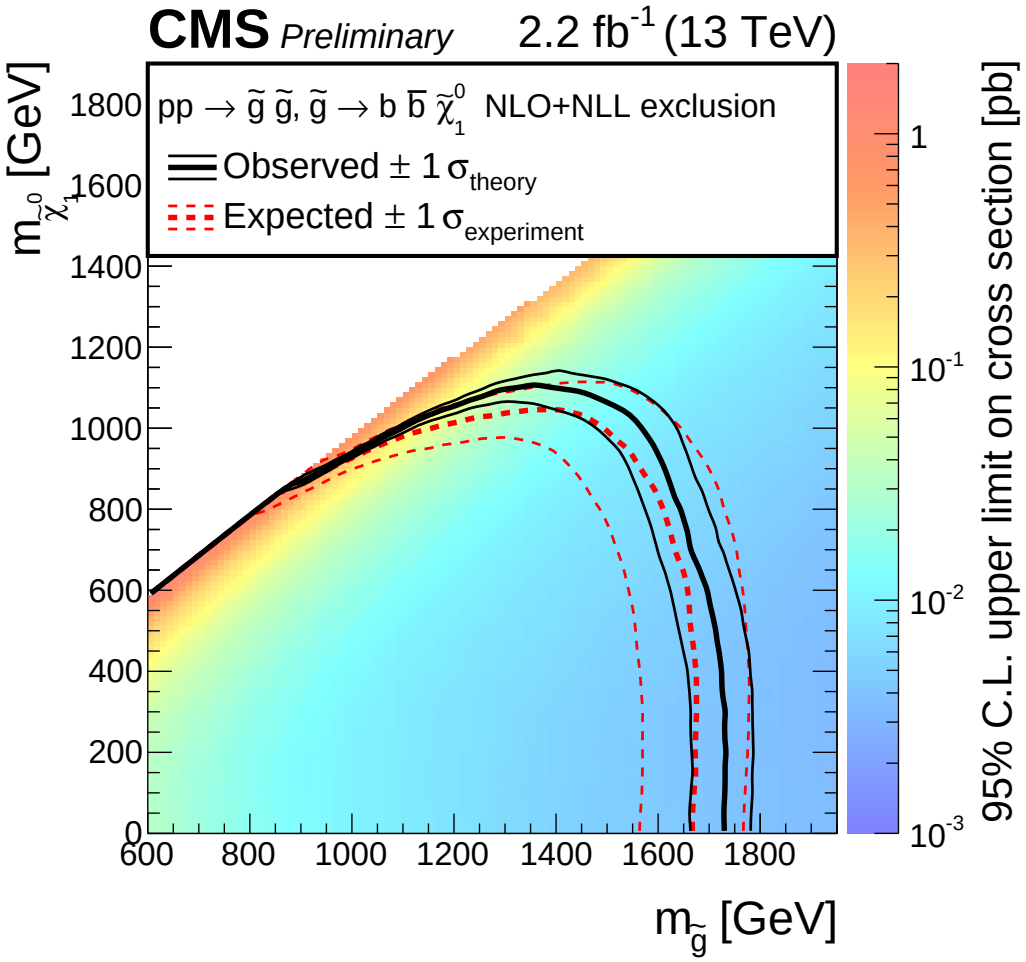

Figure 9: Exclusion limit at 95% CL for gluino mediated bottom-squark production. The area to the left of and below the thick black curve represents the observed exclusion region, while the dashed red lines indicate the expected limit and ± 1 standard-deviation. The thin black lines show the effect of the theoretical cross section uncertainties.

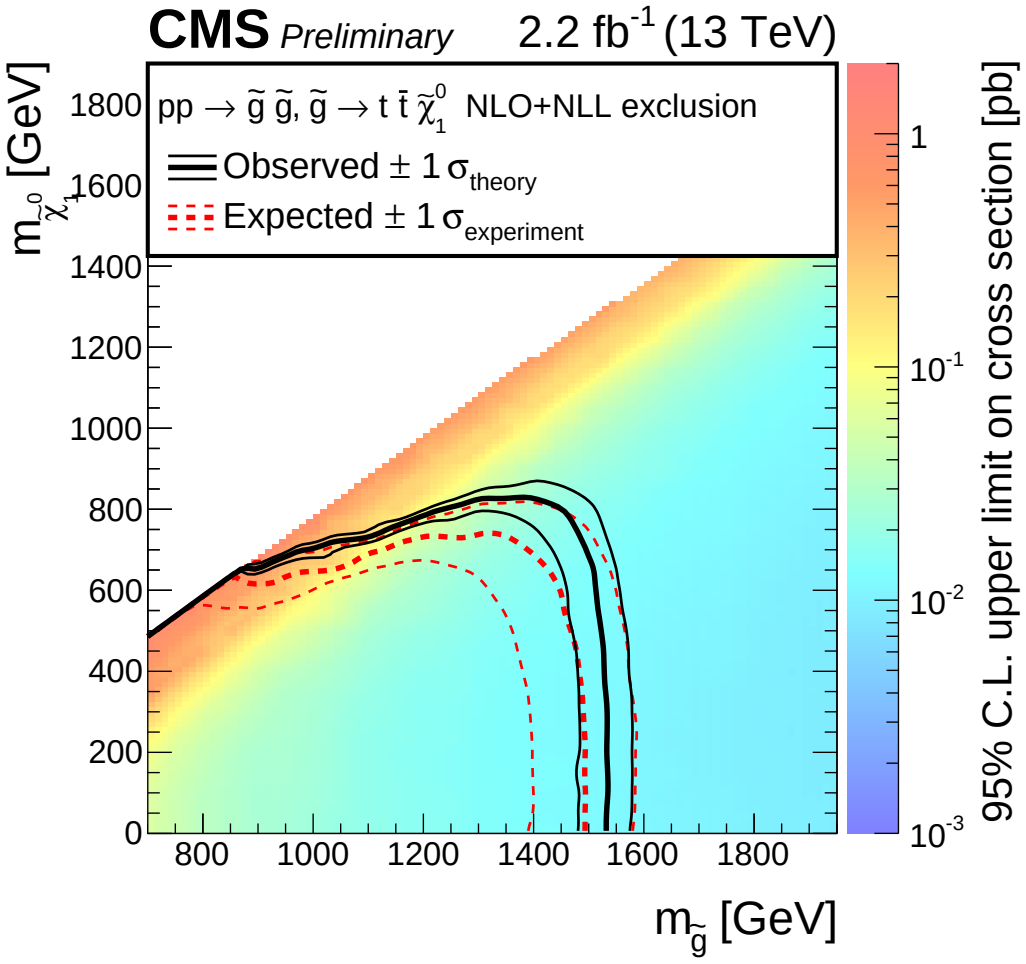

Figure 10: Exclusion limit at 95% CL for gluino mediated top-squark production. The area to the left of and below the thick black curve represents the observed exclusion region, while the dashed red lines indicate the expected limit and ± 1 standard-deviation. The thin black lines show the effect of the theoretical cross section uncertainties.

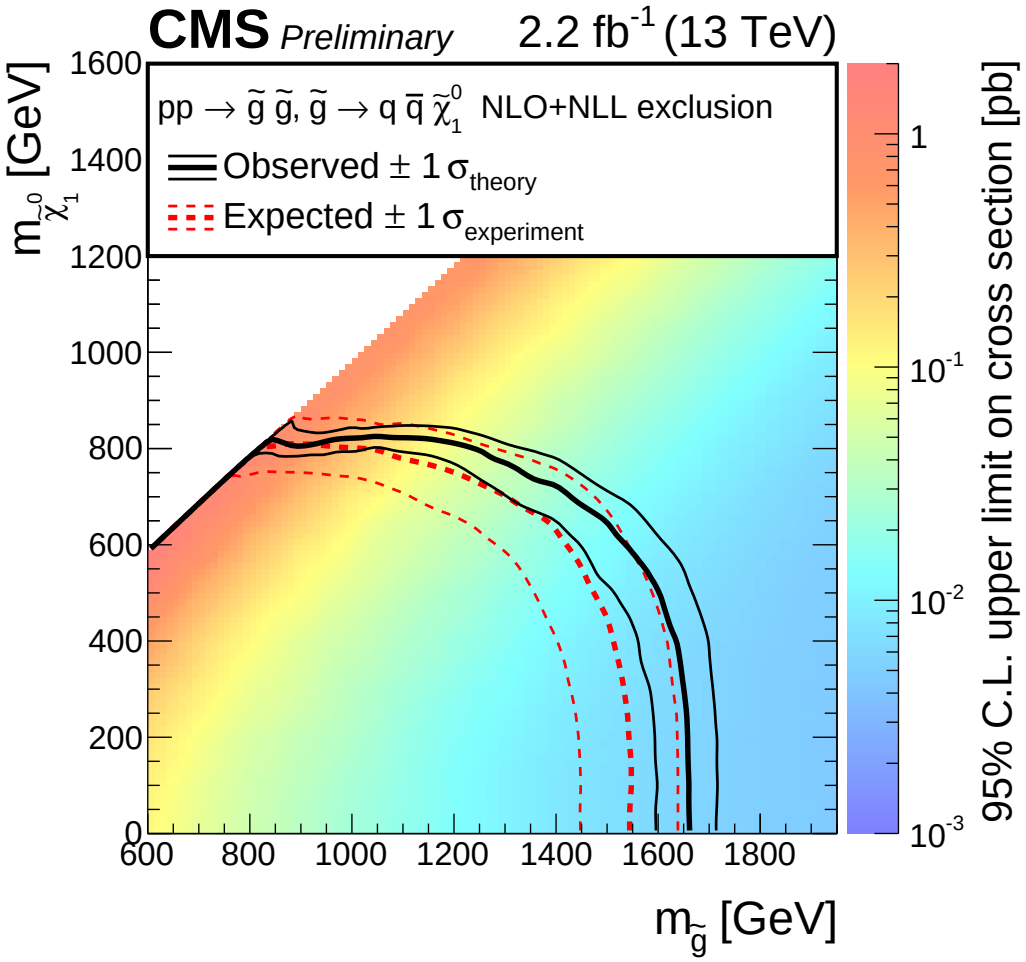

Figure 11: Exclusion limit at 95% CL for gluino mediated squark production, where the squark can be any of the first two generations. The area to the left of and below the thick black curve represents the observed exclusion region, while the dashed red lines indicate the expected limit and ± 1 standard-deviation. The thin black lines show the effect of the theoretical cross section uncertainties.

References

[1] ATLAS Collaboration, "Summary of the searches for squarks and gluinos using $\sqrt{s} = 8$ TeV pp collisions with the ATLAS experiment at the LHC", *JHEP* **10** (2015) 054, doi:10.1007/JHEP10(2015)054, arXiv:1507.05525.

[2] ATLAS Collaboration, "ATLAS Run 1 searches for direct pair production of third-generation squarks at the Large Hadron Collider", *Eur. Phys. J.* **C75** (2015), no. 10, 510, doi:10.1140/epjc/s10052-015-3726-9, arXiv:1506.08616.

[3] CMS Collaboration, "Searches for Supersymmetry using the M_{T2} Variable in Hadronic Events Produced in pp Collisions at 8 TeV", *JHEP* **05** (2015) 078, doi:10.1007/JHEP05(2015)078, arXiv:1502.04358.

[4] CMS Collaboration, "Search for Supersymmetry Using Razor Variables in Events with *b*-Tagged Jets in *pp* Collisions at $\sqrt{s} = 8$ TeV", *Phys. Rev.* **D91** (2015) 052018, doi:10.1103/PhysRevD.91.052018, arXiv:1502.00300.

[5] CMS Collaboration, "Search for new physics in the multijet and missing transverse momentum final state in proton-proton collisions at $\sqrt{s}$= 8 TeV", *JHEP* **06** (2014) 055, doi:10.1007/JHEP06(2014)055, arXiv:1402.4770.

[6] CMS Collaboration, "Search for gluino mediated bottom- and top-squark production in multijet final states in pp collisions at 8 TeV", *Phys. Lett.* **B725** (2013) 243–270, doi:10.1016/j.physletb.2013.06.058, arXiv:1305.2390.

[7] CMS Collaboration, "Search for supersymmetry in hadronic final states with missing transverse energy using the variables α_T and b-quark multiplicity in pp collisions at $\sqrt{s} = 8$ TeV", *Eur. Phys. J.* **C73** (2013), no. 9, 2568, doi:10.1140/epjc/s10052-013-2568-6, arXiv:1303.2985.

[8] CMS Collaboration, "Search for supersymmetry in hadronic final states using M_{T2} in pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **10** (2012) 018, doi:10.1007/JHEP10(2012)018, arXiv:1207.1798.

[9] C. G. Lester and D. J. Summers, "Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders", *Phys. Lett. B* **463** (1999) 99, doi:10.1016/S0370-2693(99)00945-4, arXiv:hep-ph/9906349.

[10] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET", Technical Report CMS-PAS-PFT-09-001, CERN, 2009.

[11] CMS Collaboration, "Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector", Technical Report CMS-PAS-PFT-10-001, CERN, 2010.

[12] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[13] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.

[14] CMS Collaboration, "Performance of b-Tagging Algorithms in 50ns Data at 13 TeV", Technical Report CMS-DP-2015-045, Sep, 2015.

[15] CMS Collaboration, "Electron reconstruction and identification at sqrt(s) = 7 TeV", Technical Report CMS-PAS-EGM-10-004, CERN, Geneva, 2010.

[16] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV", *JINST* **7** (2012) P10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.

[17] CMS Collaboration, "Jet Performance in pp Collisions at 7 TeV", Technical Report CMS-PAS-JME-10-003, CERN, Geneva, 2010.

[18] T. Sjöstrand, "The Lund Monte Carlo for e^+e^- Jet Physics", *Comput. Phys. Commun.* **28** (1983) 229, doi:10.1016/0010-4655(83)90041-3.

[19] 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.

[20] CMS Collaboration, "Missing transverse energy performance of the CMS detector", *JINST* **6** (2011) P09001, doi:10.1088/1748-0221/6/09/P09001, arXiv:1106.5048.

[21] J. Alwall et al., "MadGraph 5: going beyond", *JHEP* **06** (2011) 128, doi:10.1007/JHEP06(2011)128, arXiv:1106.0522.

[22] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1", *Comput. Phys. Commun.* **178** (2008) 852–867, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.

[23] GEANT4 Collaboration, "GEANT4—a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

[24] S. Abdullin et al., "The fast simulation of the CMS detector at LHC", *J. Phys. Conf. Ser.* **331** (2011) 032049, doi:10.1088/1742-6596/331/3/032049.

[25] CMS Collaboration, "Search for New Physics with Jets and Missing Transverse Momentum in pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **08** (2011) 155, doi:10.1007/JHEP08(2011)155, arXiv:1106.4503.

[26] A. L. Read, "Presentation of search results: The CL_s technique", *J. Phys. G* **28** (2002) 2693, doi:10.1088/0954-3899/28/10/313.

[27] A. L. Read, "Modified frequentist analysis of search results (The CL_s method)", *CERN-OPEN* **205** (2000).

[28] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727.

[29] ATLAS and CMS Collaborations, "Procedure for the LHC Higgs boson search combination in summer 2011", CMS NOTE/ATL-PHYS-PUB ATL-PHYS-PUB-2011-011, CMS-NOTE-2011-005, 2011.

[30] CMS Collaboration, "Search for top-squark pair production in the single-lepton final state in pp collisions at $\sqrt{s}$ = 8 TeV", *Eur. Phys. J.* **C73** (2013), no. 12, 2677, doi:10.1140/epjc/s10052-013-2677-2, arXiv:1308.1586.

A Detailed Results

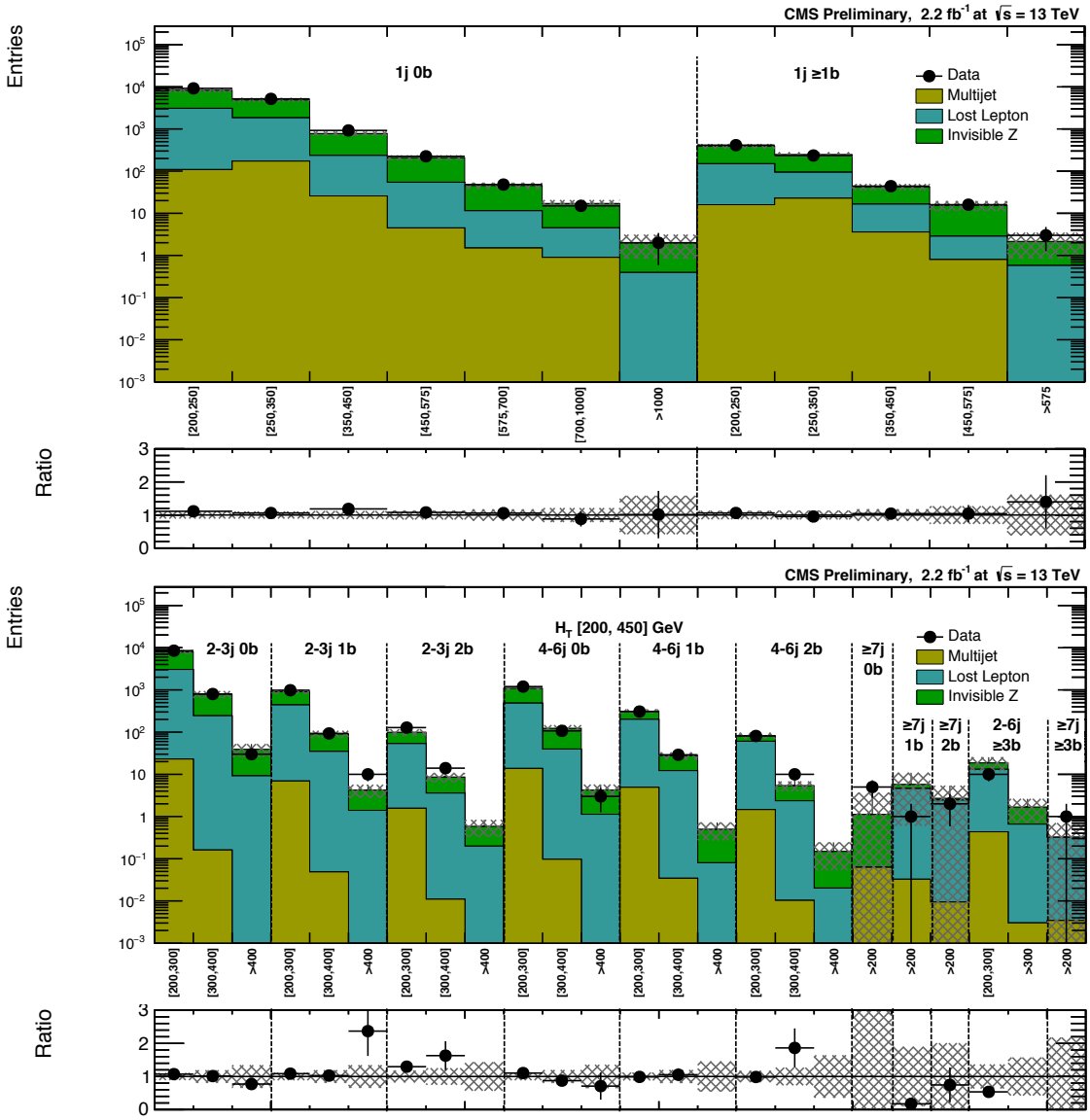

Figure 12: (Above) Comparison of the estimated background and observed data events in each signal bin in the mono-jet region. On the x-axis, the H_T binning is shown (in GeV). Hatched bands represent the full uncertainty on the background estimate. (Below) Same for the very low H_T region. On the x-axis, the M_{T2} binning is shown (in GeV). Bins with no entry for data have an observed count of 0.

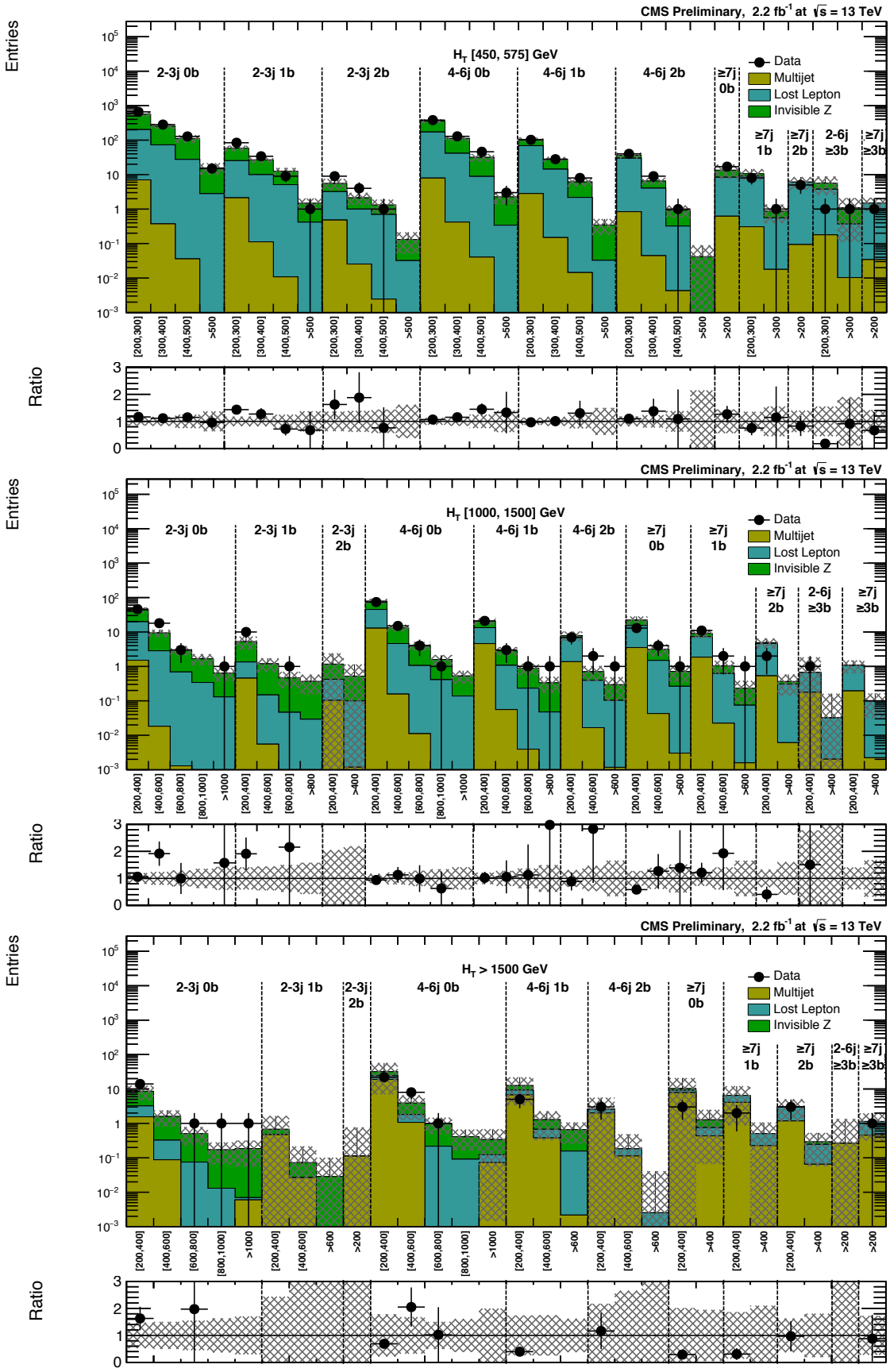

Figure 13: (Top) Comparison of the estimated background and observed data events in each signal bin in the low H_T region. Hatched bands represent the full uncertainty on the background estimate. Same for the high (middle) and extreme (bottom) H_T regions. On the x-axis, the M_{T2} binning is shown (in GeV). Bins with no entry for data have an observed count of 0.

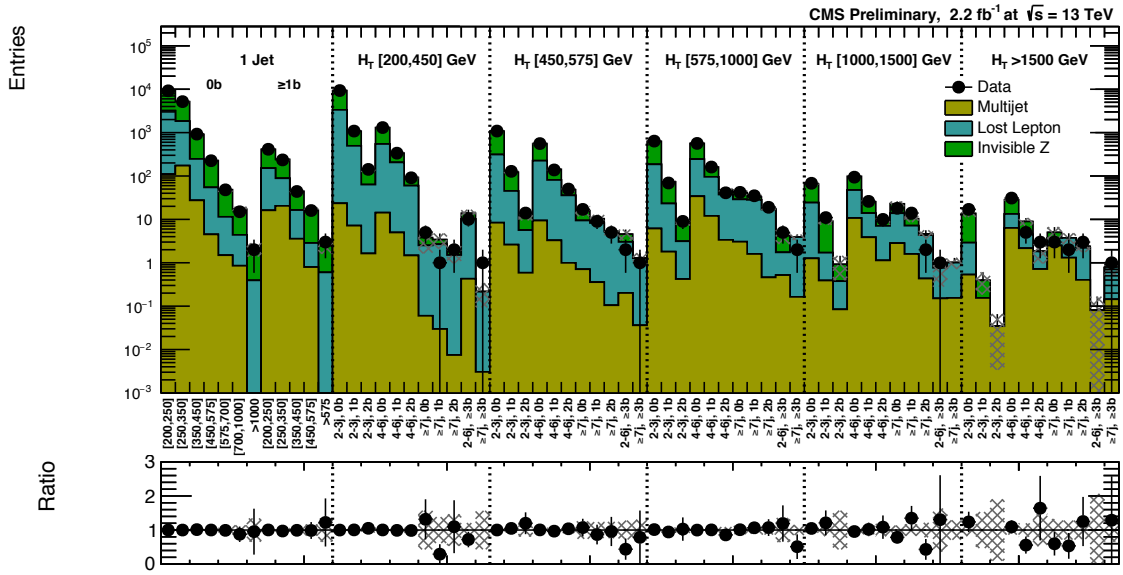

Figure 14: Comparison of post-fit background prediction and observed data events in each topological region. Hatched bands represent the post-fit uncertainty on the background prediction. For the monojet, on the x-axis the H_T binning is shown (in GeV).

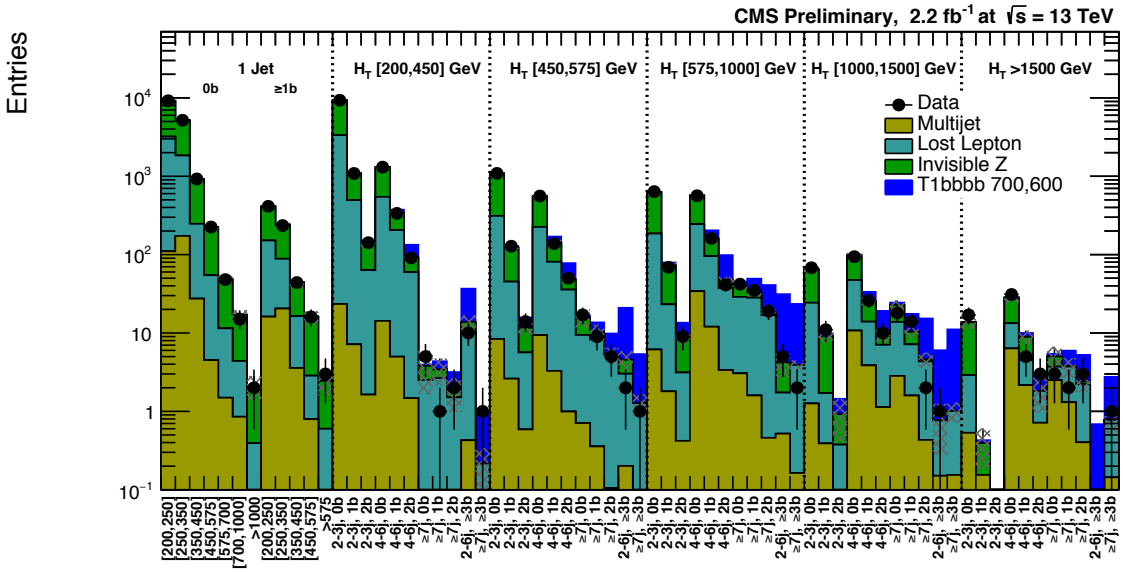

Figure 15: Post-fit background prediction, expected signal yields and observed data events in each topological region. Hatched bands represent the post-fit uncertainty on the background prediction. For the monojet, on the x-axis the H_T binning is shown (in GeV). The compressed-spectra signal model considered here is gluino-mediated bottom-squark production with mass of the gluino and LSP equal to 700 and 600 GeV, respectively.

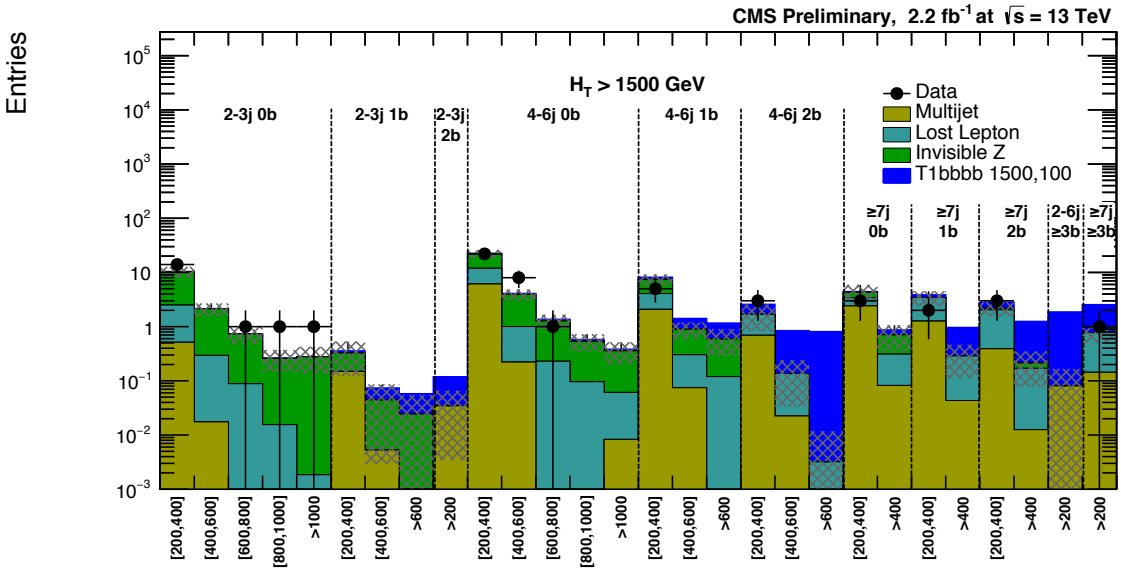

Figure 16: Post-fit background prediction, expected signal yields and observed data events in each signal bin in the extreme H_T region. Hatched bands represent the post-fit uncertainty on the background prediction. On the x-axis the M_{T2} binning is shown (in GeV). The open-spectra signal model considered here is gluino-mediated bottom-squark production with mass of the gluino and LSP equal to 1500 and 100 GeV, respectively.

Table 3: Binning in M_{T2} for each topological region of the multi-jet search regions with very-low, low, and medium H_T.

H_T Range [GeV]	Jet Multiplicities	Binning [GeV]
[200, 450]	$2-3$j, 0b	[200, 300, 400, ∞]
	$2-3$j, 1b	[200, 300, 400, ∞]
	$2-3$j, 2b	[200, 300, 400, ∞]
	$4-6$j, 0b	[200, 300, 400, ∞]
	$4-6$j, 1b	[200, 300, 400, ∞]
	$4-6$j, 2b	[200, 300, 400, ∞]
	≥ 7j, 0b	[200, ∞]
	≥ 7j, 1b	[200, ∞]
	≥ 7j, 2b	[200, ∞]
	$2-6$j, ≥ 3b	[200, 300, ∞]
	≥ 7j, ≥ 3b	[200, ∞]
[450, 575]	$2-3$j, 0b	[200, 300, 400, 500, ∞]
	$2-3$j, 1b	[200, 300, 400, 500, ∞]
	$2-3$j, 2b	[200, 300, 400, 500, ∞]
	$4-6$j, 0b	[200, 300, 400, 500, ∞]
	$4-6$j, 1b	[200, 300, 400, 500, ∞]
	$4-6$j, 2b	[200, 300, 400, 500, ∞]
	≥ 7j, 0b	[200, ∞]
	≥ 7j, 1b	[200, 300, ∞]
	≥ 7j, 2b	[200, ∞]
	$2-6$j, ≥ 3b	[200, 300, ∞]
	≥ 7j, ≥ 3b	[200, ∞]
[575, 1000]	$2-3$j, 0b	[200, 300, 400, 600, 800, ∞]
	$2-3$j, 1b	[200, 300, 400, 600, 800, ∞]
	$2-3$j, 2b	[200, 300, 400, 600, ∞]
	$4-6$j, 0b	[200, 300, 400, 600, 800, ∞]
	$4-6$j, 1b	[200, 300, 400, 600, ∞]
	$4-6$j, 2b	[200, 300, 400, 600, ∞]
	≥ 7j, 0b	[200, 300, 400, ∞]
	≥ 7j, 1b	[200, 300, 400, ∞]
	≥ 7j, 2b	[200, 300, 400, ∞]
	$2-6$j, ≥ 3b	[200, 300, 400, ∞]
	≥ 7j, ≥ 3b	[200, 300, 400, ∞]

Table 4: Binning in M_{T2} for each topological region of the multi-jet search regions with high and extreme H_T.

H_T Range [GeV]	Jet Multiplicities	Binning [GeV]
[1000, 1500]	$2-3$j, 0b	[200, 400, 600, 800, 1000, ∞]
	$2-3$j, 1b	[200, 400, 600, 800, ∞]
	$2-3$j, 2b	[200, 400, ∞]
	$4-6$j, 0b	[200, 400, 600, 800, 1000, ∞]
	$4-6$j, 1b	[200, 400, 600, 800, ∞]
	$4-6$j, 2b	[200, 400, 600, ∞]
	≥ 7j, 0b	[200, 400, 600, ∞]
	≥ 7j, 1b	[200, 400, 600, ∞]
	≥ 7j, 2b	[200, 400, ∞]
	$2-6$j, ≥ 3b	[200, 400, ∞]
	≥ 7j, ≥ 3b	[200, 400, ∞]
[1500, ∞]	$2-3$j, 0b	[200, 400, 600, 800, 1000, ∞]
	$2-3$j, 1b	[200, 400, 600, ∞]
	$2-3$j, 2b	[200, ∞]
	$4-6$j, 0b	[200, 400, 600, 800, 1000, ∞]
	$4-6$j, 1b	[200, 400, 600, ∞]
	$4-6$j, 2b	[200, 400, 600, ∞]
	≥ 7j, 0b	[200, 400, ∞]
	≥ 7j, 1b	[200, 400, ∞]
	≥ 7j, 2b	[200, 400, ∞]
	$2-6$j, ≥ 3b	[200, ∞]
	≥ 7j, ≥ 3b	[200, ∞]