CMS Draft Analysis Note

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

2011/02/10
Head Id: 38642
Archive Id: 26698:38642M
Archive Date: 2011/02/10
Archive Tag: trunk

Further interpretation of the RA1 SUSY search

R. Bainbridge[1], B. Betchart[2], O. Buchmüller[1], H. Flächer[2], E. Laird[3], B. Mathias[1], T. Rommerskirchen[4], P. Sphicas[4], M. Stoye[4], and T. Whyntie[1]

[1] Imperial College, London, UK
[2] University of Rochester, Rochester, NY, USA
[3] Princeton University, Princeton, NJ, USA
[4] CERN, Geneva, Switzerland

Abstract

A first search for supersymmetry in events with jets and missing energy was carried out in Ref. [1], using 35 $\mathrm{pb}^{-}1$ of integrated luminosity at $\sqrt{s} = 7$ TeV. In this search, the variable α_{T} was used as the main discriminator between events with real and fake missing transverse energy and no excess of events over the Standard Model expectation was found. In this note an extended interpretation of the above result is presented. An increased sensitivity to supersymmetric signals is obtained by splitting the signal region, defined by values of the scalar sum of the jet transverse momenta greater 350 GeV, into two bins. Based on these results, the dependence of the exclusion limit in the constrained minimal supersymmetric model is studied dependent on the value of $\tan\beta$. Furthermore, upper limits on the cross-section for different Simplified Model Spectra are presented.

Contents

1 Introduction

In this note an extended interpretation of the results of Ref. [1] is presented. These additional interpretations are based on the full selection, definition of signal and background control regions, as well as the data driven background methods, of the RA1 search. No re-analysis of the data or changes in the cutflow are carried out. The results are interpreted in two new physics scenarios: the Constraint Minimal SuperSymmetric Model (CMSSM) [2] and the more generic Simplified Model Spectra (SMS) [3]. An increased sensitivity of the analysis to physics processes with higher-mass states is achieved by splitting the signal region $H_\mathrm{T} > 350\,\mathrm{GeV}$ into two bins, namely $350 < H_\mathrm{T} < 450\,\mathrm{GeV}$ and $H_\mathrm{T} > 450\,\mathrm{GeV}$. Furthermore, the 1-Sigma band for the limits in the CMSSM are added to the analysis result.

2 Summary of analysis and results

In the following we give a brief summary of the analysis in Ref. [1], which selects events with two or more high-p_T jets. Specifically, jets are reconstructed using the anti-$\mathrm{k_T}$ algorithm [4] with a size parameter of 0.5 and are required to have $E_\mathrm{T} > 50\,\mathrm{GeV}$, $|\eta| < 3$ and to pass jet identification criteria [5] designed to reject spurious signals in the calorimeters. The pseudorapidity of the jet with the highest E_T (leading jet) is required to be within $|\eta| < 2.5$ and the transverse energy of each of the two leading jets must exceed $100\,\mathrm{GeV}$.

Events with jets passing the E_T threshold but not satisfying the jet identification criteria or the η acceptance requirement are vetoed, as this deposited energy is not accounted for in the event kinematics. Similarly, events in which an isolated lepton (electron [6] or muon [7]) with $p_\mathrm{T} > 10\,\mathrm{GeV}$ is identified are rejected to suppress events with genuine missing energy from neutrinos. Furthermore, to select a pure multi-jet topology, events are vetoed in which an isolated photon [8] with $p_\mathrm{T} > 25\,\mathrm{GeV}$ is found.

Events are required to fulfill $H_\mathrm{T} = \sum_{i=1}^{\mathrm{N_{jet}}} E_\mathrm{T}^{\mathrm{j}_i} > 350\,\mathrm{GeV}$. As the main discriminator against QCD multijet production the variable α_T, defined as:

$$\alpha_\mathrm{T} = E_\mathrm{T}^{\mathrm{j}_2} / M_\mathrm{T} = \frac{E_\mathrm{T}^{\mathrm{j}_2}}{\sqrt{\left(\sum_{i=1}^{2} E_\mathrm{T}^{\mathrm{j}_i}\right)^2 - \left(\sum_{i=1}^{2} p_x^{\mathrm{j}_i}\right)^2 - \left(\sum_{i=1}^{2} p_y^{\mathrm{j}_i}\right)^2}},$$

is used and events are required to have $\alpha_\mathrm{T} > 0.55$.

To protect against multiple jets failing the $E_\mathrm{T} > 50\,\mathrm{GeV}$ selection requirement, the jet-based estimate of the missing energy, $\not{H}_\mathrm{T}$, is compared to the calorimeter tower-based estimate, $\not{E}_\mathrm{T}^{\mathrm{calo}}$, and events with $R_\mathrm{miss} = \not{H}_\mathrm{T} / \not{E}_\mathrm{T}^{\mathrm{calo}} > 1.25$ are rejected.

Finally, to protect against severe energy losses, events with significant jet mismeasurements caused by masked regions in the ECAL, which amount to about 1% of the ECAL channel count, are removed with the following procedure. The jet-based estimate of the missing transverse energy, $\not{H}_\mathrm{T} = |\vec{\not{H}}_\mathrm{T}| = |-\sum_\mathrm{jets} \vec{p}_{\mathrm{T}_\mathrm{jet}}|$, is used to identify the jet most likely to have given rise to the $\not{H}_\mathrm{T}$ as the jet whose momentum is closest in ϕ to the total $\vec{\not{H}}_\mathrm{T}$ which results after removing the jet from the event. The azimuthal distance between this jet and the recomputed $\not{H}_\mathrm{T}$ is referred to as $\Delta\phi^*$ in what follows. Events with $\Delta\phi^* < 0.5$ are rejected if the distance in the (η, ϕ) plane between the selected jet and the closest masked ECAL region, ΔR_ECAL, is smaller than 0.3.

This selection results in 13 observed events in a data sample of 35 pb^{-1}.

3 Further interpretations of the RA1 result

In the following we describe extended interpretations of the RA1 result.

3.1 1-Sigma band for 95% CL limit in the CMSSM

For a more comprehensive statistical interpretation and as a measure of quality of the 95% CL in CMSSM, we have added the 1-Sigma band to the expected limit, as illustrated in Fig. 1, which shows the 95% CL excluded region in the CMSSM for $\tan\beta = 3$. The 1-Sigma band is obtained by generating toy MC for the event yield in the signal-like region and in each of the control regions. These toys are generated for the background only scenario. From these pseudo-measurements, upper limits are calculated and the 68% central confidence interval is taken as 1-Sigma band. The expected limit is taken as the median of the upper limits from the pseudo-measurements.

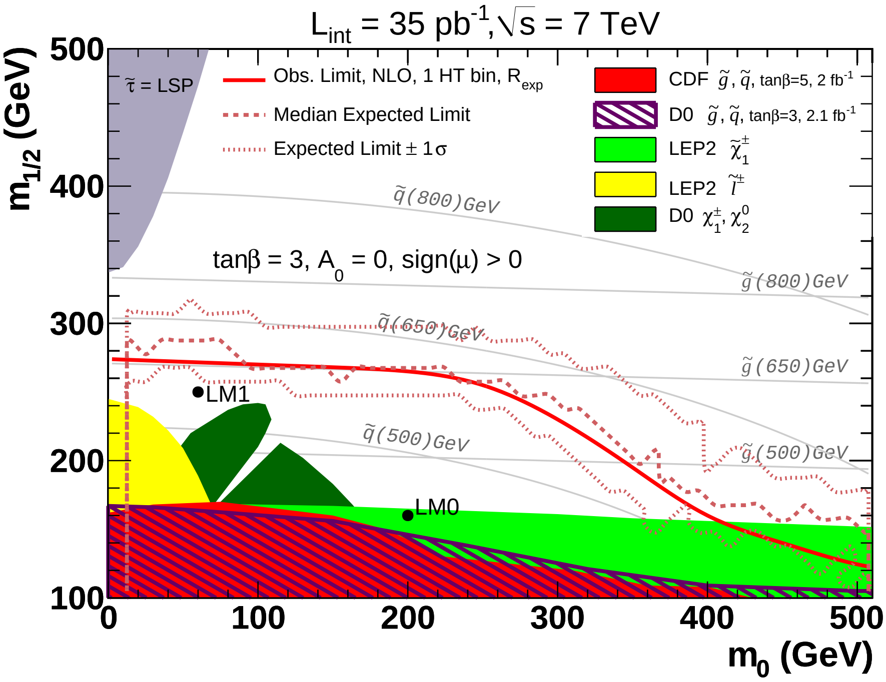


Figure 1: Observed and expected 95% CL exclusion contours (the latter shown with a 1-Sigma band) in the CMSSM $m_0 - m_{1/2}$ plane ($\tan\beta = 3$, $A_0 = 0$, $\mathrm{sign}(\mu) > 0$). *Note: The plot will be replaced with a soothed version after the pre-approval.*

3.2 Improved sensitivity to higher-mass states

The signal region in the RA1 analysis is defined by $H_T > 350\,\mathrm{GeV}$. For reasons of simplicity and robustness it was decided to carry out a simple cut-and-count interpretation of the final results for the RA1 publication. In order to gain additional sensitivity to higher-mass states it is possible to make a re-interpretation of the observed and expected number of events by splitting the original signal region into multiple H_T bins. Therefore, the simple cut-and-count interpretation of one signal bin becomes a H_T shape analysis based on several bins. This requires that the data driven background methods used to determine the expected number of SM background events in the signal region provide an estimate for each of the H_T bins in the signal region of $H_T > 350\,\mathrm{GeV}$. The data driven background methods used for the RA1 analysis were explicitly designed with this use-case in mind. In the following, the results for the different background prediction methods are presented when splitting the signal region into two bins: $350 < H_T < 450\,\mathrm{GeV}$ and $H_T > 450\,\mathrm{GeV}$, which is the optimal number of bins that can be exploited for the available data set. It should be noted that this approach does not involve any changes of cuts or even a re-analysis of the data. It simply represents a re-interpretation of the final results in terms of multiple H_T bins instead of only one.

3.2.1 Total background estimation using H_T dependence of R_{α_T}

The total Standard Model background is estimated based on an extrapolation of the ratio $R_{\alpha_T} = N^{\alpha_T>0.55}/N^{\alpha_T<0.55}$ from a low H_T control region to the signal region $H_T > 350\,\mathrm{GeV}$.

As in Ref. [1], the following dependencies of R_{α_T} on H_T are considered:

- a constant behaviour
- an exponential dependence

For the RA1 publication the more conservative approach of assuming an exponential dependence was chosen as the default method, while the more precise constant assumption was used as a cross check. Table 1 shows how the predicted and observed number of events are divided between the two H_T bins in the signal region. Figure 2 displays the observed H_T evolution of R_{α_T} for data, SM and SM+LM1. Also superimposed on each plot are $\pm 1\sigma$ bands representing the expected evolution of R_{α_T} as a function of H_T, based on the measured R_{α_T} values in the control H_T region and the assumption of exponential or constant dependence on H_T.

Table 1: Observed and predicted event yields in the different H_T bins of both the control and signal regions. The quoted uncertainties are statistical only.

H_T (GeV)	250-300	300-350	350-450	>450
$N^{\alpha_T>0.55}$	33	11	8	5
$N^{\alpha_T<0.55}$	844459	331948	225649	110036
R_{α_T} (10^{-5}) (Data)	$3.91^{+0.72}_{-0.64}$	$3.31^{+1.11}_{-0.91}$	$3.55^{+1.42}_{-1.12}$	$4.54^{+2.39}_{-1.77}$
R_{α_T} (10^{-5}) (const)	$3.72^{+0.61}_{-0.52}$		$3.72^{+0.61}_{-0.52}$	$3.72^{+0.61}_{-0.52}$
$N^{\alpha_T>0.55}$ (const)	–	–	$8.40^{+1.37}_{-1.18}$	$4.10^{+0.67}_{-0.58}$
R_{α_T} (10^{-5}) (exp)	$3.91^{+0.72}_{-0.64}$	$3.31^{+1.11}_{-0.91}$	$2.66^{+2.17}_{-1.79}$	$1.84^{+3.06}_{-2.54}$
$N^{\alpha_T>0.55}$ (exp)	–	–	$6.00^{+4.89}_{-4.03}$	$2.02^{+3.36}_{-2.79}$

3.2.2 Estimation of background from $t\bar{t}$ and W + jets events using a muon control sample

Table 2 shows the split of the muon control sample numbers and corresponding background prediction in the two H_T bins.

3.2.3 Estimation of background from Z $\rightarrow \nu\bar{\nu}$+ jets from photon + jets events

Table 3 shows the split of the photon control sample numbers and corresponding background prediction in the two H_T bins.

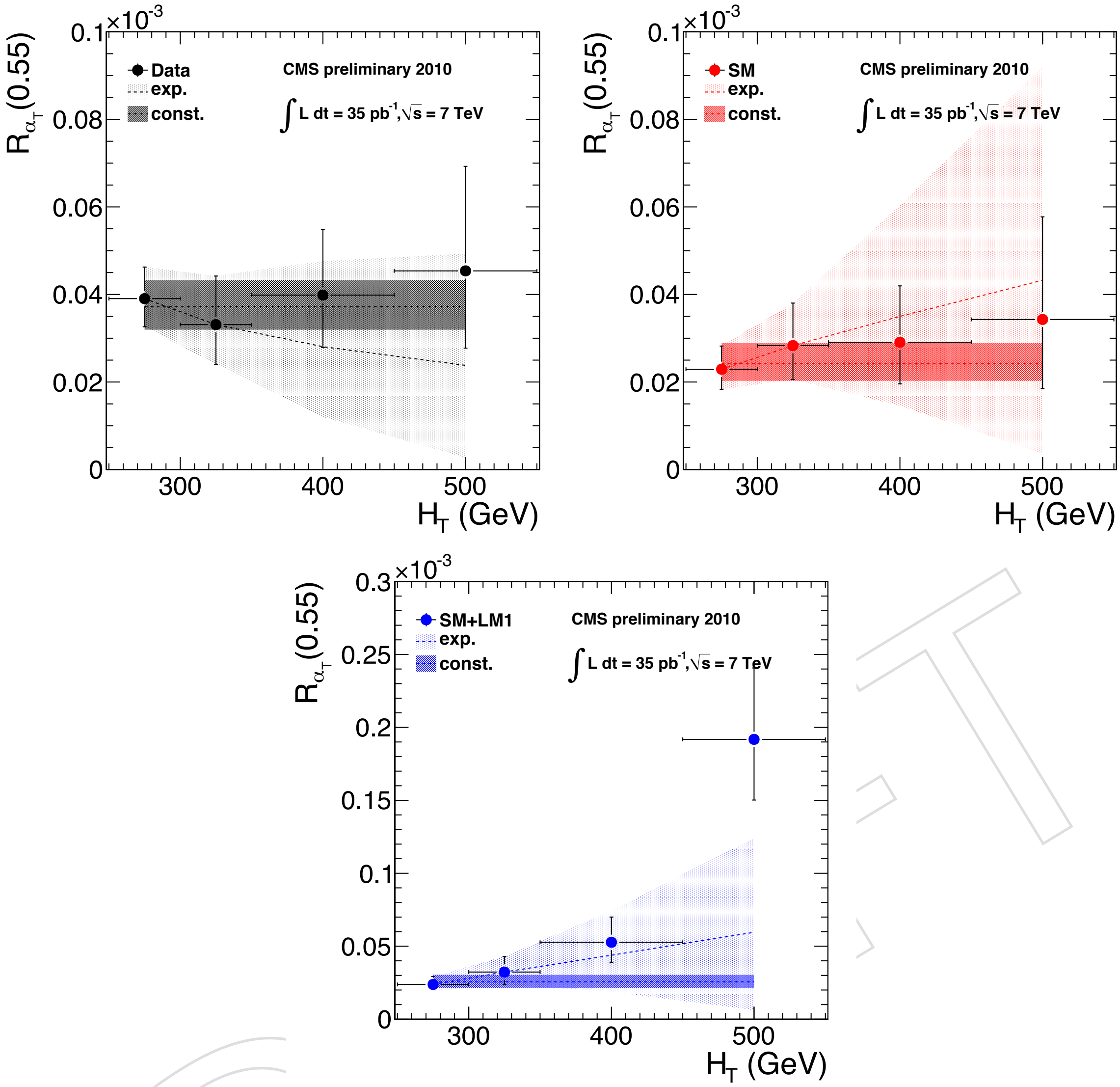


Figure 2: R_{α_T} as a function of H_T for data (left), SM (right) and SM+LM1 (bottom). The light and dark bands represent the expected R_{α_T} values ($\pm 1\sigma$) for each of the H_T bins in the signal region ($H_T > 350$ GeV), for the exponential and constant behaviours, respectively.

3.2.4 Impact of the H_T shape interpretation on the 95% CL exclusion limit in the CMSSM

Figure 3 shows a comparison of the 95% CL exclusion limit for the published cut-and-count (i.e. one signal bin) interpretation of the data and the limit obtained from the two-bin shape analysis interpretation. As expected, there is a significant gain in the observed and expected exclusion limits, of approximately 20 GeV in $m_{1/2}$ for fixed m_0 when using the shape interpretation of the final result.

3.3 Improved Background Estimations

For the RA1 publication, very conservative assumptions were used in the data driven background methods. For the inclusive background prediction, the double-ratio method, which makes the least assumptions about the H_T extrapolation into signal region but suffers from very large statistical uncertainties, was chosen as the default inclusive prediction over the more

Table 2: Observed number of events in data and MC simulation for the μ + jets control sample and the MC expectation $t\bar{t}$/W + jets events in the hadronic signal sample.

Sample	$350 < H_T < 450$ GeV	$H_T > 450$ GeV
μ + jets (Data; μ sample)	5	2
μ + jets (MC; μ sample)	4.1	1.9
$t\bar{t}$/W + jets (MC; hadr. sample)	3.4	1.7
$\tau(N_{MC}^{t\bar{t}/W;had}/N_{MC}^{t\bar{t}/W;mu})$	$0.83 \pm 30\%$	$0.89 \pm 30\%$
Predicted $t\bar{t}$/W + jets BG	$4.2 \pm {}^{+1.8}_{-2.1}\,{}_{stat} \pm 1.3\,{}_{syst}$	$1.8 \pm {}^{+1.4}_{-1.8}\,{}_{stat} \pm 0.5\,{}_{syst}$

Table 3: Observed number of events in data and MC simulation for the photon + jets control sample and the MC expectation for $Z \to \nu\bar{\nu}$+ jets events in the hadronic signal sample.

Sample	$350 < H_T < 450$ GeV	$H_T > 450$ GeV
γ + jets (Data.; μ sample)	6	1
γ + jets (MC; μ sample)	4.4	2.1
$Z \to \nu\bar{\nu}$ +jets (MC; hadr. sample)	2.6	1.5
$\tau(N_{MC}^{Z\to\nu\bar{\nu};had}/N_{MC}^{\gamma;\gamma})$	$0.59 \pm 40\%$	$0.71 \pm 40\%$
Predicted $Z \to \nu\bar{\nu}$ BG	$3.5 \pm {}^{+1.4}_{-1.6}\,{}_{stat} \pm 1.4_{syst}$	$0.7 \pm {}^{1.0}_{-0.7}\,{}_{stat} \pm 0.3_{syst}$

aggressive but also more precise "flat" scaling assumption (see Section 3.2.1). For the estimate of the EWK background based on muon (see Section 3.2.2 and photon (see Section 3.2.3 control samples, very conservative assumptions for event selection and systematic error determination were used for the final result. In the following, we discuss the possible improvement of the result when more aggressive assumptions for the data driven background methods are exploited.

3.3.1 Using the constant R_{α_T} assumption for the inclusive background prediction

As discussed as a cross check in the RA1 paper, another variant of the H_T scaling analysis, based on the independence of R_{α_T} on H_T when the data sample is dominated by EWK processes, i.e. for $\alpha_T > 0.55$, uses the weighted average of the R_{α_T} values measured in the two H_T control regions. This value is then also used in the signal region to obtain an estimate of the total background. As it can be seen from Table 1, the constant assumption of R_{α_T} with H_T yields a significantly more precise prediction of the total background in the two H_T bins than the double ratio scaling, which only assumes that the double ratio is constant.

Figure 4 shows a comparison of the 95% CL exclusion limit of the default analysis with the one obtained when using the more precise background prediction based on the constant R_{α_T} assumption. The gain in the observed and expected exclusion limit is approximately 20 GeV in $m_{1/2}$ for fixed m_0 and therefore of comparable size to the gain achieved with the two-bin interpretation (see Fig. 3 for comparison).

4 Final Plots and Results

In this section we provide a selection of plots useful for characterisation of the results of the RA1 search. The interpretation is carried out in the CMSSM as well as the simplified models

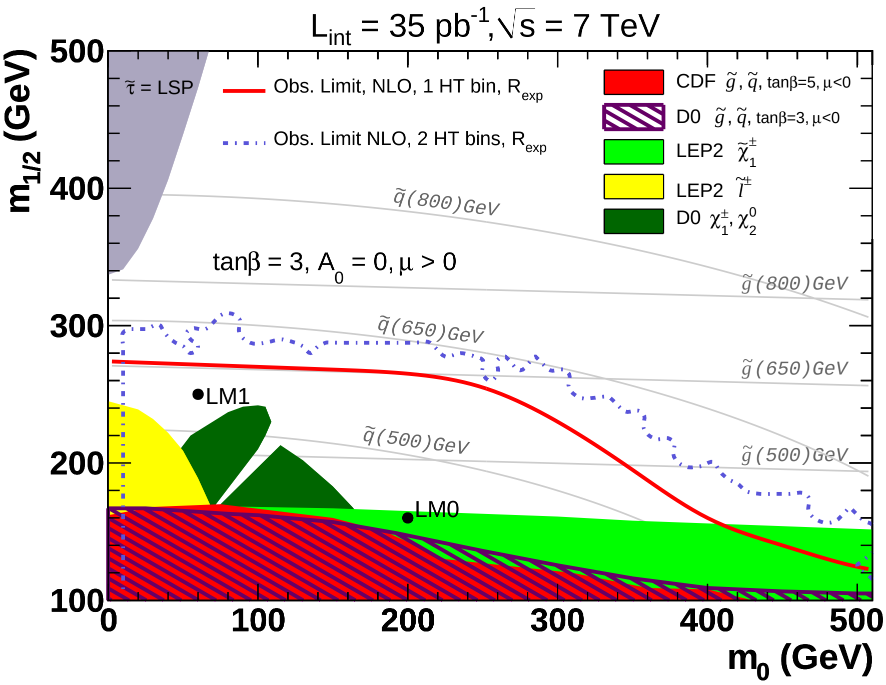


Figure 3: Comparison of the observed 95% CL exclusion contour in the CMSSM $m_0 - m_{1/2}$ plane ($\tan\beta = 3$, $A_0 = 0$, sign(μ) > 0) for the default analysis and re-interpretation using two H_T bins in the signal region. *Note: Since producing the expected limit with toy experiments is CPU time consuming, we present here the comparison for the observed limit which leads to the same conclusion. The plot will be replaced with a soothed version showing the expected limit after the pre-approval.*

T1 and T2. The presented results are based on the two-bin H_T shape interpretation of the signal region (Section 3.2.1) using the assumption of constant R_{α_T} for the inclusive background estimate (Section 3.3.1) and the default prediction method of the EWK background estimate [1] (Sections 3.2.2 and 3.2.3). Furthermore, signal contamination in the H_T control region and the EWK control data samples is properly taken into account for all the results.

4.1 Interpretation in the CMSSM

4.1.1 Analysis efficiency as a function of m_0 and $m_{1/2}$

Figure 5 shows the variation of the analysis efficiency (i.e. selection efficiency times acceptance) over the $m_0 - m_{1/2}$ plane in the CMSSM for values of $\tan\beta = 3$, 10, and 50. The analysis efficiency is normalised to the total number of signal events expected for a given model point in the CMSSM.

4.1.2 Exclusion Limits in the CMSSM

Figures 6, 7 and 8 show the 95% CL excluded regions in the CMSSM for $\tan\beta = 3, 10$ and 50.

[1] It should be noted that we can provide plots for all scenarios (e.g. the two-bin approach with exponential H_T scaling instead of the constant R_{α_T} assumption, or the one-bin cut-and-count interpretation instead of the two-bin shape analysis interpretation, etc) and the decision of what scenario(s) to provide for the approval should be taken in consultation with the SUSY community and ARC during the pre-approval.

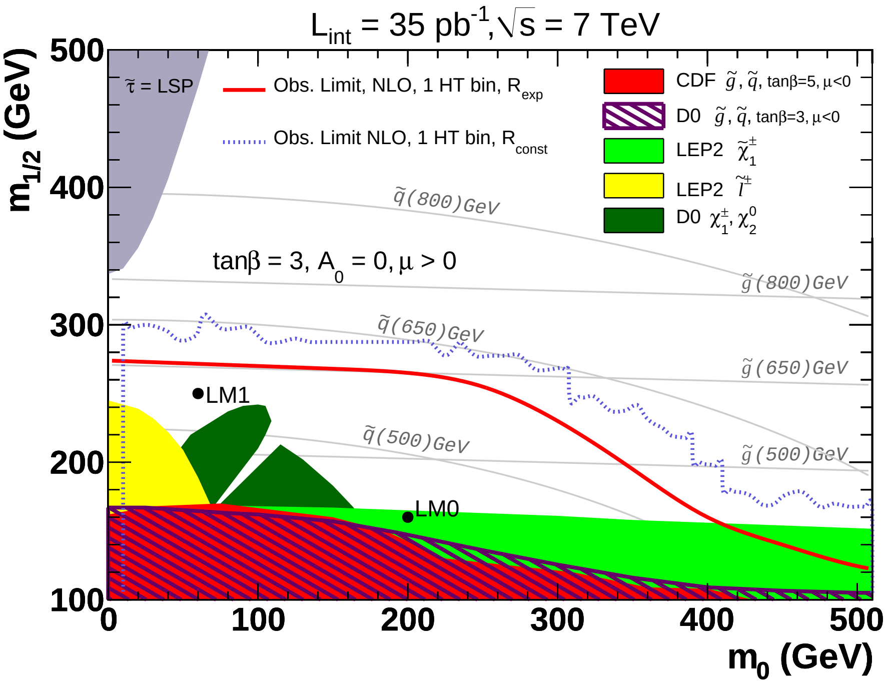


Figure 4: Comparison of the observed 95% CL exclusion contours in the CMSSM $m_0 - m_{1/2}$ plane ($\tan\beta = 3$, $A_0 = 0$, sign$(\mu) > 0$) for the default analysis and the more precise inclusive background estimate based on constant R_{α_T}. *Note: Since producing the expected limit with toy experiments is CPU time consuming, we present here the comparison for the observed limit which leads to the same conclusion. The plot will be replaced with a soothed version showing the expected limit after the pre-approval.*

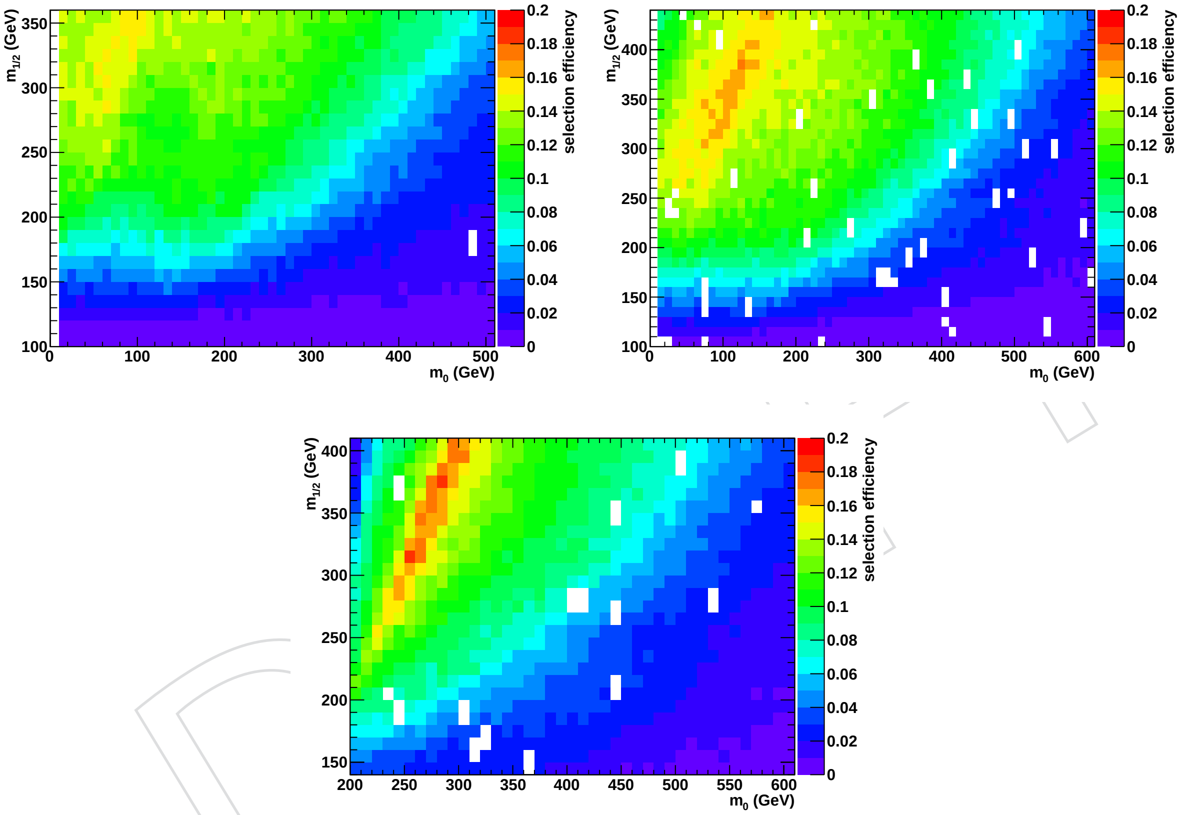


Figure 5: Analysis efficiency in the CMSSM $m_0 - m_{1/2}$ plane ($A_0 = 0$, sign$(\mu) > 0$) for $\tan\beta = 3$ (left), $\tan\beta = 10$ (right), and $\tan\beta = 50$ (bottom).

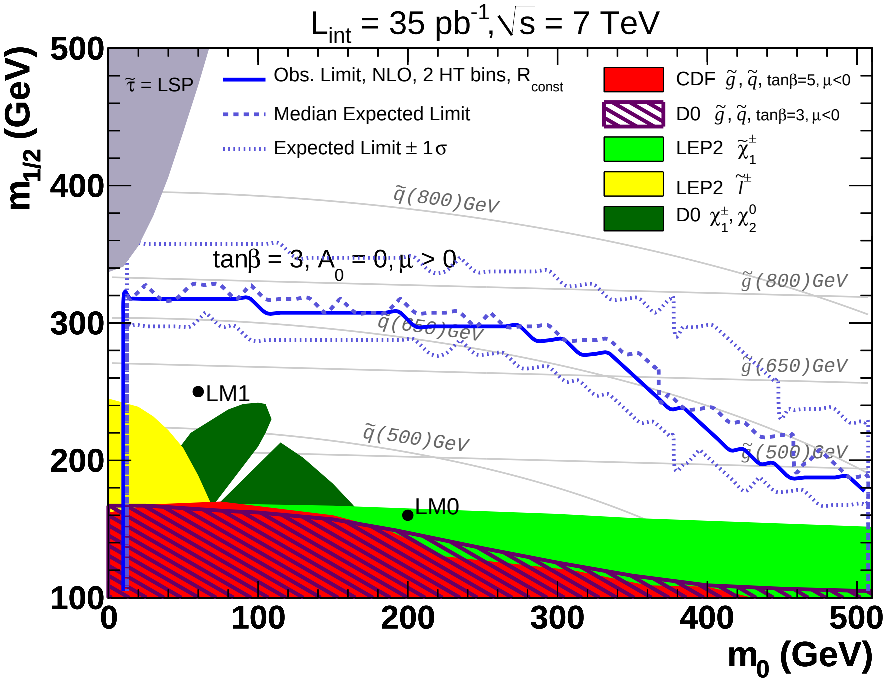


Figure 6: Expected and observed 95% CL exclusion contours in the CMSSM $m_0 - m_{1/2}$ plane ($\tan\beta = 3$, $A_0 = 0$, $\text{sign}(\mu) > 0$). *Note: The plot will be replaced with a soothed version after the pre-approval.*

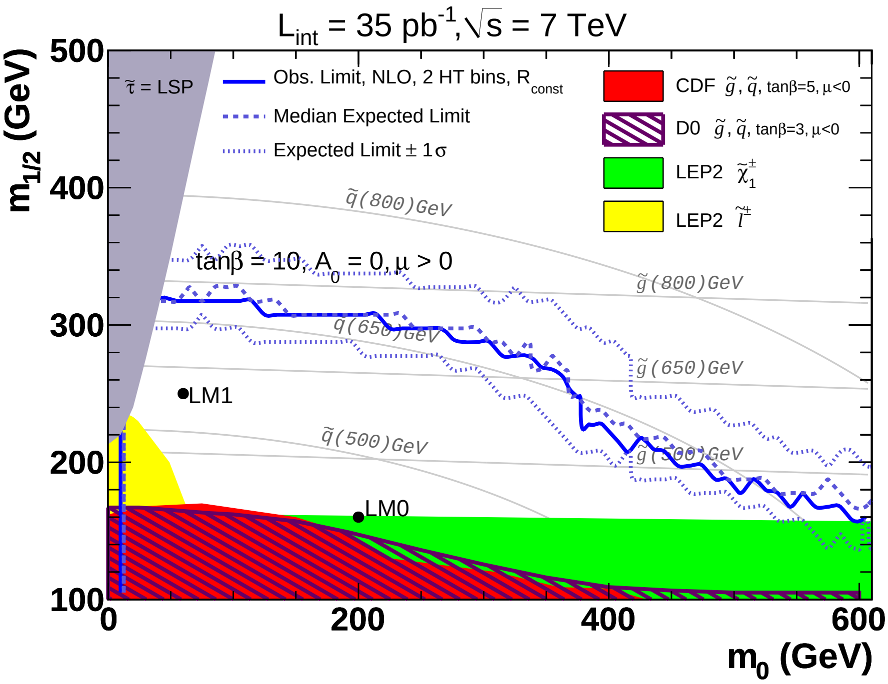


Figure 7: Expected and observed 95% CL exclusion contours in the CMSSM $m_0 - m_{1/2}$ plane ($\tan\beta = 10$, $A_0 = 0$, $\text{sign}(\mu) > 0$). *Note: The plot will be replaced with a soothed version after the pre-approval.*

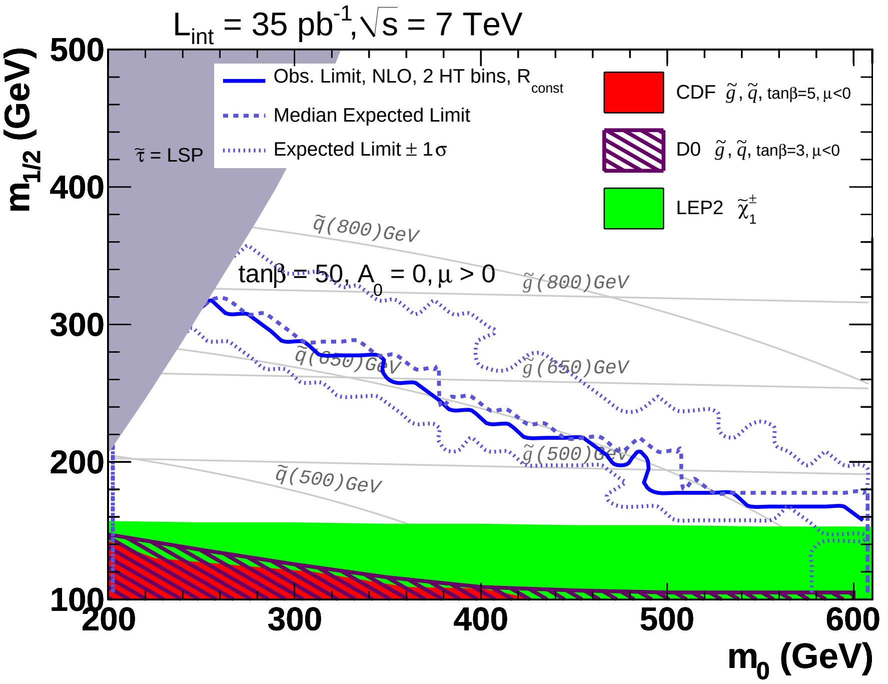


Figure 8: Expected and observed 95% CL exclusion contours in the CMSSM $m_0 - m_{1/2}$ plane ($\tan\beta = 50$, $A_0 = 0$, $\text{sign}(\mu) > 0$). *Note: The plot will be replaced with a soothed version after the pre-approval.*

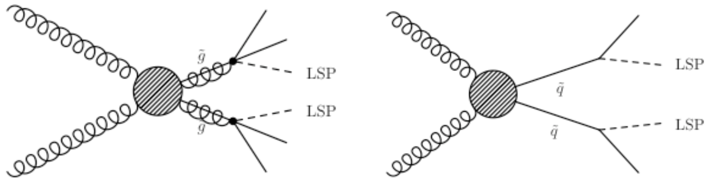


Figure 9: Diagram of simplified models. Top left: gluino pair production; top right: squark pair production.

4.2 Cross-section limits for Simplified Model Spectra

The following description of the Simplified Model Spectra is taken from Refs. [3, 9] where they are described in detail.

In short, they are [2]:

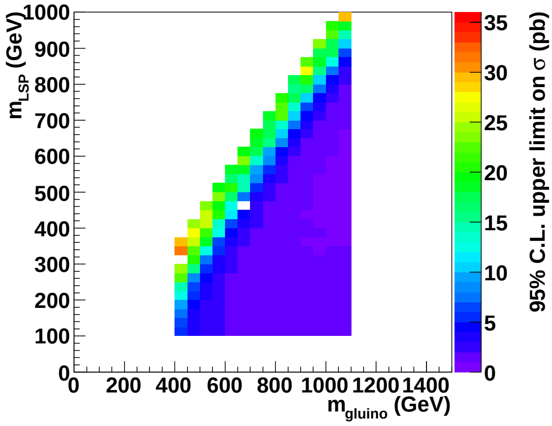
- pair-produced gluinos where each gluino directly decays to two light quarks and the LSP;
- pair-produced squarks where each squark decays to one jet and the LSP;

Figure 9 shows the respective diagrams for the simplified topologies. Fast simulation Monte Carlo samples are generated for different combinations of squark (gluino) and LSP masses.

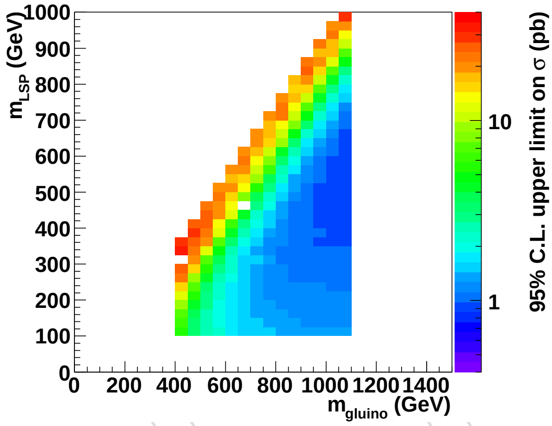
Figs. 10 and 11 show the efficiency and upper limit on the cross-section for the T1 and T2 topologies, respectively. There is not a one-to-one relation between efficiency and cross section limit because of signal contamination in the background control samples.

It can be seen that the efficiency of the analysis is much reduced in regions of parameter space where the squark (gluino) and LSP masses are similar as in this case the production of hard jets is suppressed. It is highest for heavy squarks (gluinos) and LSP mass roughly half the squark (gluino) mass which leads to hard jets and sizeable missing energy.

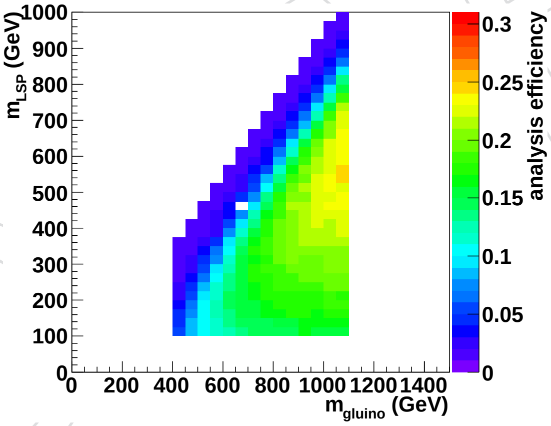
[2]So far only the two models listed below have been considered. Additionally, models with pair-produced gluinos or squarks decaying through a one stage cascade resulting in jets, LSP and a W are available and could be investigated as well if desired.



(a)



(b)



(c)

Figure 10: Top: Cross-section for the T1 topology excluded at the 95% CL. Bottom: efficiency times acceptance of the analysis for the T1 topology.

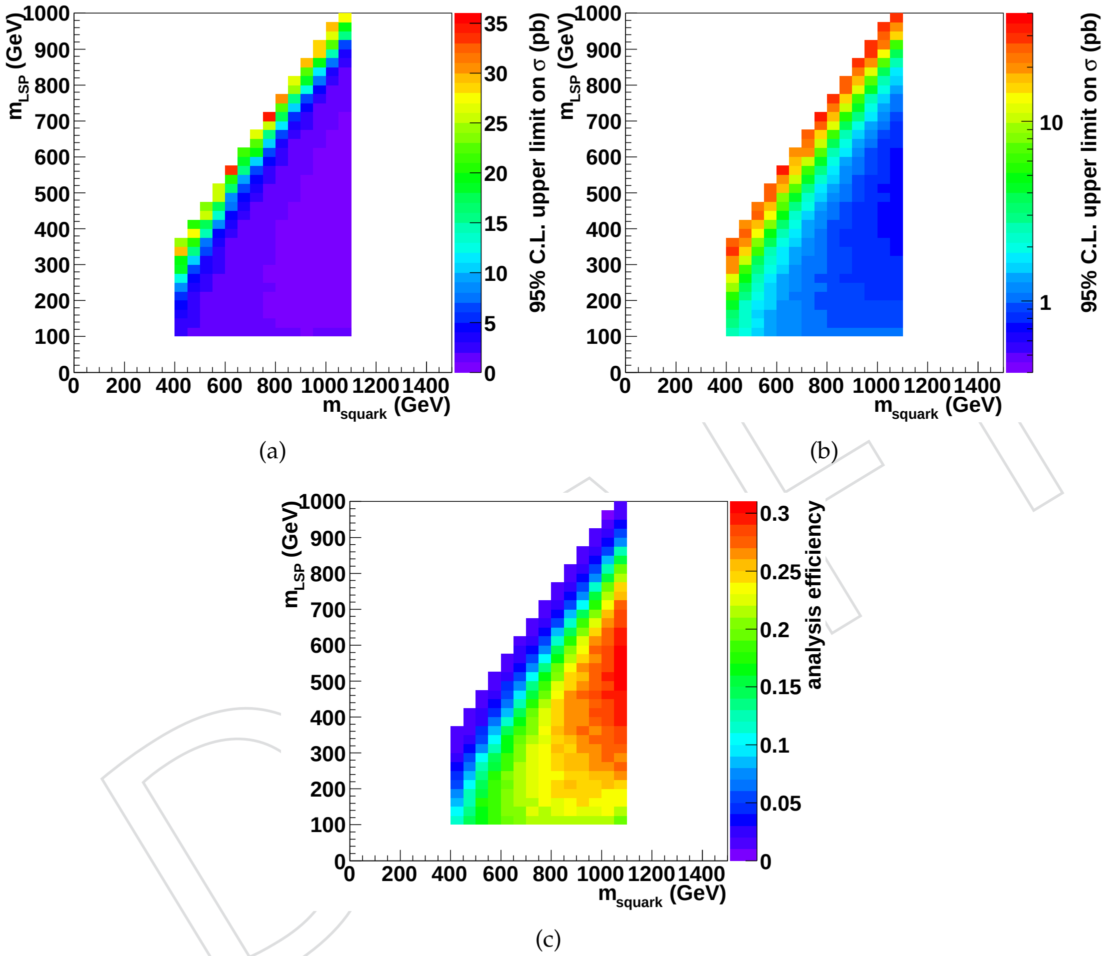


Figure 11: Top: Cross-section for the T2 topology excluded at the 95% CL. Bottom: efficiency times acceptance of the analysis for the T2 topology.

5 Conclusions

References

[1] CMS Collaboration, "Search for Supersymmetry in pp Collisions at 7 TeV in Events with Jets and Missing Transverse Energy", arXiv:1101.1628.

[2] G. L. Kane, C. F. Kolda, L. Roszkowski et al., "Study of constrained minimal supersymmetry", *Phys. Rev.* **D49** (1994) 6173. doi:10.1103/PhysRevD.49.6173.

[3] https://twiki.cern.ch/twiki/pub/CMS/SUSY/referenceTopologiesDec13.pdf.

[4] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-kt jet clustering algorithm", *JHEP* **0804:063** (2008). doi:10.1088/1126-6708/2008/04/063.

[5] CMS Collaboration, "Calorimeter Jet Quality Criteria for the First CMS Collision Data", *CMS Physics Analysis Summary* **JME-09-008** (2009).

[6] CMS Collaboration, "Electron reconstruction and identification at $\sqrt{s} = 7$ TeV", *CMS Physics Analysis Summary* **EGM-10-004** (2010).

[7] CMS Collaboration, "Performance of muon identification in pp collisions at $\sqrt{s} = 7$ TeV", *CMS Physics Analysis Summary* **MUO-10-002** (2010).

[8] CMS Collaboration, "Photon reconstruction and identification at $\sqrt{s} = 7$ TeV", *CMS Physics Analysis Summary* **EGM-10-005** (2010).

[9] Search Optimization and Limit Setting Using Simplified Model Spectra for the RA2 Inclusive Hadronic SUSY Search, CMS AN-2010/404.