Summer Student report

Sensitivity study on Dark Matter inspired pMSSM SUSY scenarios with the ATLAS experiment using soft-leptons, b-jets and missing transverse momentum in the final state

Gernot Knippen

Supervisors: Moritz Backes, Max Baak

August 31, 2013

ATLAS

CERN

1 Introduction

The hierarchy problem is an important topic in particle physics theory. One elegant possibility to solve this problem is the introduction of supersymmetry (SUSY) to the existing standard model. Additionally supersymmetry has the potential to explain dark matter. At the LHC experiments ATLAS and CMS experimental physicists are trying to find hints on the existence of SUSY with great effort. The most important part of this research is the existence of potent SUSY analyzes.

The aim of my work was to have a look at recent SUSY analyzes and check their sensitivity on different pMSSM scenarios, which include the latest known dark matter constraints. Unfortunately the existing analyzes I had a look at are not powerful enough to detect possible SUSY stop and sbottom decays in the given pMSSM phase space configurations. Thus I reoptimized one for a specific class of scenarios, for which one phase space point could be excluded.

2 pMSSM grids

I had a look at 232 pMSSM Dark Matter inspired phase space points divided into four grids varying M_1 and m_{qL3}. [1] For each of the phase space points of the grids Monte Carlo generated samples were available, which were approximately 20 million events in total. Only truth information coming from the event generator Herwig++ was available.

3 Cut setup

The signal regions "SR1L1Ba′" and "SR1L1Bc′" are defined by the phase space cuts given in table 1 while the cuts of signal regions "SR1L2Ba′" and "SR1L2Bc′" are given by table 2.

The auxiliary variables $H_{T,x}$ and m_{CT} are defined as follows:

$$H_{T,x} = \sum_{i=x+1} |\mathbf{p}_{T,jet_i}| \tag{1}$$

is the scalar sum over all jet p_Ts excluding the first x leading jets.

$$m_{CT}^2 = (E_T(b\text{-}jet_1) + E_T(b\text{-}jet_2))^2 \tag{2}$$

$$- (\mathbf{p}_T(b\text{-}jet_1) + \mathbf{p}_T(b\text{-}jet_2))^2 \tag{3}$$

with $E_T = \sqrt{\mathbf{p}_T^2 + m^2}$ is the squared cotransverse mass, which is an invariant quantity for back to back boosted particles in the transverse plane. I did not use the normal cotransverse mass but the boost corrected one as described in [6].

Table 1: Cut setup for "SR1L1Ba′" and "SR1L1Bc′"

cut name	SR1L1Ba′	SR1L1Bc′
$lep1$	leading lepton: $e: 7\,\text{GeV} < p_T < 25\,\text{GeV} \wedge !(1.37 < \lvert\eta\rvert < 1.52)$; $\mu: 6\,\text{GeV} < p_T < 25\,\text{GeV}$	
lep^{add}	additional leptons: $e: p_T < 7\,\text{GeV}$; $\mu: p_T < 6\,\text{GeV}$	
$jet\text{-}p_T$	leading jet: $p_T > 180\,\text{GeV}$	
$jet1\ not\ b$	leading jet is no b-jet	
$soft\ b\text{-}jet$	at least one of the following jets is a b-jet	
$\#jets$	$\#jets \geq 3$	$\#jets \geq 2$
$\frac{MET}{m_{\text{eff}}^{\text{inc}}}$	$\frac{MET}{m_{\text{eff}}^{\text{inc}}} > 0.3$	$\frac{MET}{m_{\text{eff}}^{\text{inc}}} > 0.35$
m_T	$m_T > 100\,\text{GeV}$	$m_T > 90\,\text{GeV}$
MET	$E_{T,\text{miss}} > 300\,\text{GeV}$	$E_{T,\text{miss}} > 370\,\text{GeV}$

Table 2: Cut setup for “SR1L2Ba′” and “SR1L2Bc′”

cut name	SR1L2Ba′	SR1L2Bc′
$lep1$	leading lepton: $e: 7\,\text{GeV} < p_T < 25\,\text{GeV} \wedge !(1.37 < \vert\eta\vert < 1.52)$ $\mu: 6\,\text{GeV} < p_T < 25\,\text{GeV}$	
lep^{add}	additional leptons: $e: p_T < 7\,\text{GeV}$ $\mu: p_T < 6\,\text{GeV}$	
$jet_2\text{-}p_T$	second leading jet: $p_T > 60\,\text{GeV}$	
$nB2Jet60$	2 b-jets in the leading 2 jets with $p_T > 60\,\text{GeV}$	
$\Delta\phi_{min}$	$\Delta\phi(E_{T,miss}, every\,jet) > 0.4$	
M_{bb}	invariant mass of two leading b-jets (M_{bb}) $> 150\,\text{GeV}$	
$jet_3\text{-}p_T$	$jet_3\text{-}p_T < 50\,\text{GeV}$	-
$H_{T,2}$	$H_{T,2} < 50\,\text{GeV}$	-
m_{CT}	$m_{CT,corr} > 150\,\text{GeV}$	$m_{CT,corr} > 200\,\text{GeV}$
MET	$E_{T,\text{miss}} > 150\,\text{GeV}$	$E_{T,\text{miss}} > 250\,\text{GeV}$

4 Efficiencies / Correction factors

Because only truth information was available correction factors had to be calculated to model the detector response. Since the b-tagging efficiency is very important I established two methods to include the b-tagging efficiency into the truth n-tuples, one for the 1 b-jet signal region SR1L1Ba′ & SR1L1Bc′ and one for the 2 b-jet signal regions SR1L2Ba′ & SR1L2Ba′. All other corrections where combined to the overall fudge factor ξ, which was calculated by different but similar samples.

4.1 B-tagging efficiency $\epsilon_{\textbf{b-tag}}$ for 1B signal regions

I estimated the b-tagging effiency $\epsilon_{\text{b-tag}}(p_{T,jet})$ by the the date given by the ATLAS CONF note [3] and the usage of the following formula:

$$\begin{aligned}\epsilon_{\text{b-tag}}(p_{T,jet}) &= 1 - P(\text{no soft b-jet detected}) \\ &= 1 - \prod^{\#\text{soft b jets}} 1 - \epsilon_b(p_{T,jet}) \end{aligned} \tag{4}$$

where P(no soft b-jet detected) is the given probability, that no non-leading b-jet is b tagged.

4.2 B-tagging efficiency for 2B signal regions

Because some cut variables of the 2 b-jet signal region depend on which b-jet is successfully detected, a simple introduction of a weight as for the 1B signal regions is not possible. Henceforth I modelled the b-tagging efficiency in the preparation of the n-tuples. A given truth b-jet was used as b-jet if a diced uniform number was smaller than the 60% b-tagging efficiency as given in the ATLAS CONF note [4].

4.3 fudge factor ξ

At the time of this work there were no reconstructed n-tuples for the pMSSM grid points available. Hence, I calculated an overall correction factor to compensate between reconstructed samples, which shouldn't be so far away from real data, and truth samples. I used Monte Carlo data from an 8 TeV simplified model grid. The fudge factor ξ is defined as follows

$$\xi = \frac{N_{expect,r}}{N_{expect,t}} \tag{5}$$

where $N_{expect,r}$ and $N_{expect,t}$ are the number of expected events from reconstructed respectively truth n-tuples. As ξ is nearly uniform over the given phase space area its arithmetic mean over each calculated point is used for the following calculations.

Because the simplified model Monte Carlo data had a preselection of $E_{T,miss} > 250\text{GeV}$, ξ could not be calculated within the given signal regions SR1L2Ba′ and SR1L2Bc′. Therefore I estimated ξ by using the signal regions with an artificial high missing transverse energy cut of 300 GeV. The values of ξ for the different signal regions used can be found in table 3.

Table 3: Arithmetic mean of the fudge factor ξ for the different signal regions

signal region	ξ	σ_ξ
SR1L1Ba′	0.534543	0.100387
SR1L1Bc′	0.495381	0.10325
SR1L2Ba′	0.20	0.10
SR1L2Bc′	0.20	0.10

5 Acceptance A

The acceptance A describes the sensitivity of a given cut setup. It is calculated by

$$A = \frac{N_{SR}}{N_{total}} \tag{6}$$

where N_{SR} is the number of selected events in the signal region and N_{total} is the total number of event on which the selection is applied.

6 Number of expected events

Given the acceptance A, the correction factors, the production cross sections and the integrated luminosity $\mathcal{L}$ one can calculate the number of events one would have expected to be observed by

$$N_{signal} = A \times \xi \times \sum \sigma. \tag{7}$$

I used NLO+NLL cross sections of sbottom and stop production if available. For some phase space point I had to use NLO only cross sections.

6.1 Error on number of expected events

The error on the number of expected events δN_{signal} was calculated by assuming that the error on the cross sections is negligible and that the error of the acceptance and the fudge factor are independent by usage of gaussian propagation of uncertainty. The error on the acceptances were calculated by a binomial error approach as presented by [5].

7 Model independent upper limits

If the number of SUSY events is higher than a certain number of background events the null hypothesis, meaning that statistical fluctuations of the background could have resulted in the same number of events as background and SUSY events combined, can be rejected. This implies that the analyzes with the given selections are sensitive.

The number of events from which onward one can reject the null hypothesis, the model independent upper limits, were calculated via HistFitter. The results can be seen in table 4.

Table 4: Model independent upper limits as calculated by HistFitter

Signal Region	SR1L1Ba′	SR1L1Bc′	SR1L2Ba′	SR1L2Bc′
s^{95}_{exp}	$14.0^{+5.9}_{-4.1}$	$7.9^{+3.8}_{-2.4}$	$11.8^{+5.1}_{-3.4}$	$8.3^{+3.9}_{-2.5}$

8 Reoptimization

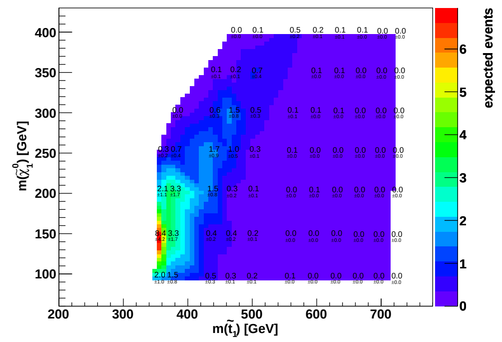
I reoptimized the signal region SR1L1Bc′ by varying its cut parameters. First I let the first lepton p_T upper cut constant to be orthogonal to the one hard lepton analyzes. Later I varied it, too. I calculate the statistical significance at each phase space point via the approximation

$$S \simeq \frac{N_{expect,SR}}{\sqrt{N_{expect,Background} + (0.3 \times N_{expect,Background})^2}}. \tag{8}$$

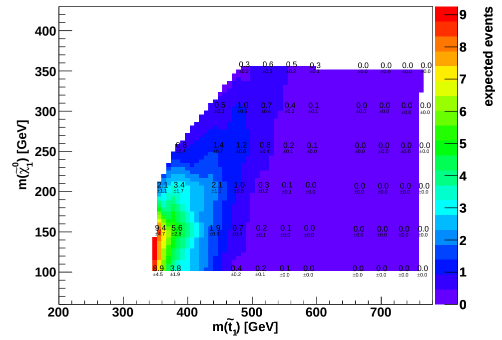
If S is greater than 1.64 for a given phase space point (corresponding to 95% confidence level) the null hypothesis can be rejected, which means that the analysis is sensitive there.

9 Recent signal regions

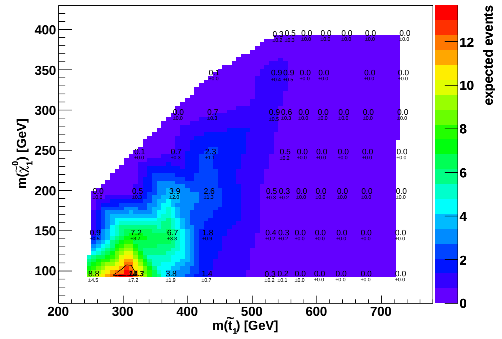
The results of the calculation of the number of expected events show, that only one signal region, SR1L2Ba′ is slightly sensitive and that one phase space point (cp. figure 1c) can be excluded with the given signal region.



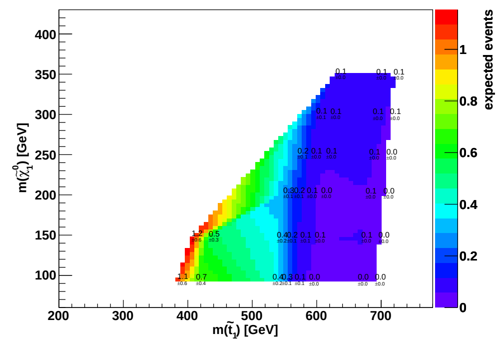
(a) $\tan\beta = 20,\ M_2 = 2\,\mathrm{TeV},\ \tilde{t}_1 \sim \tilde{t}_L$



(b) $\tan\beta = 20,\ M_2 = \mu,\ \tilde{t}_1 \sim \tilde{t}_L$



(c) $\tan\beta = 20,\ M_2 = 2\,\mathrm{TeV},\ \tilde{t}_1 \sim \tilde{t}_R$



(d) $\tan\beta = 50,\ M_2 = 2\,\mathrm{TeV},\ \tilde{t}_1 \sim \tilde{t}_L$

Figure 1: Number of expected events for the signal region SR1L2Ba′ on the DM inspired pMSSM grids. Excluded phase space points are written in bold italic. The exclusion contour in plotted in black.

10 Reoptimized signal regions

10.1 Cut variation with constant first lepton p_T upper cut of 25 GeV

After reoptimizing with a constant first lepton p_T upper cut of 25 GeV and dropping the requirement that the first jet should not be a b-jet the result seemed to be acceptable. The cut setup of this reoptimized SR which I call SR1L1Bd can be found in table 5.

Table 5: Cut setup of the recent signal region SR1L1Bc′ and the reoptimized signal region SR1L1Bd.

cut name	SR1L1Bc′ *(old)*	SR1L1Bd *(new)*
$lep1$	leading lepton:	e : 7 GeV $< p_T <$ 25 GeV $\wedge$!(1.37 $< \|\eta\| <$ 1.52) μ : 6 GeV $< p_T <$ 25 GeV
lep^{add}	additional leptons:	e : $p_T <$ 7 GeV μ : $p_T <$ 6 GeV
$\#b\text{-}jets$	at least one	
$p_T(\text{jet}_1)$	$>$ 180 GeV	$>$ 145 GeV
$jet1\ not\ b$	leading jet is no b-jet	-
$\#jets$	$\#jets \geq 2$	$\#jets \geq 1$
$\frac{MET}{m_{\text{eff}}^{\text{inc}}}$	$\frac{MET}{m_{\text{eff}}^{\text{inc}}} > 0.35$	$\frac{MET}{m_{\text{eff}}^{\text{inc}}} > 0.5$
m_T	$m_T >$ 90 GeV	$m_T >$ 90 GeV
MET	$E_{T,\text{miss}} >$ 370 GeV	$E_{T,\text{miss}} >$ 250 GeV

Because the assumption that the fudge factor ξ is constant throughout the reoptimization is an approximation the fudge factor has to be recalculated. The given simplified model reconstructed and truth n-tuples are preselected to fulfill $E_{T,miss} > 250\,\text{GeV}$. For this reason the same problem as for signal regions SR1L2Ba′ and SR1L2Bc′ occurs, the optimized cut ($E_{T,miss} > 250\,\text{GeV}$) is to close to the preselection. To be able to calculate ξ nevertheless an artificial high mising transverse energy cut of 300 GeV was used to estimate the fudge factor to ~ 0.20. The final significance with adapted fudge factor can be seen in figure 2.

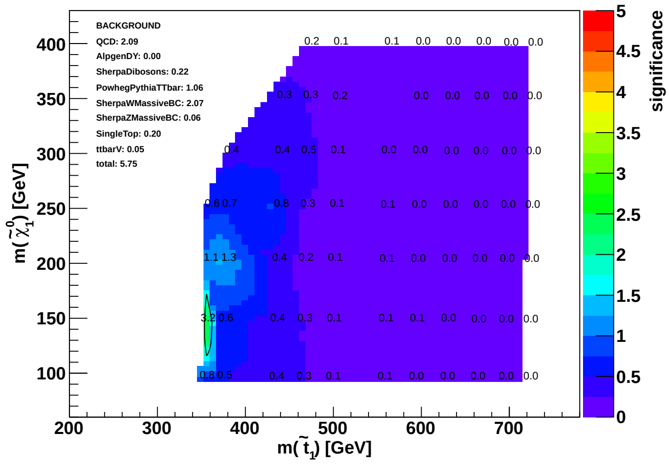


Figure 2: Significance of reoptimized signal region SR1L1Bd with adapted fudge factor. The given region is only sensitive at one phase space point. The area in which the SR can exclude SUSY is marked by a black contour.

11 Conclusion and Summary

In this study I tested the sensitivity of the existing stop soft lepton signal regions [2] on Dark Matter inspired pMSSM grids. I didn't find any sensitivity with the signal regions SR1L1Ba′, SR1L1Bc′ or SR1L2Bc′ but one point could be excluded with SR1L2Ba′.

To reach higher sensitivity I performed a signal region reoptimization but could not find any significant improvement in sensitivity with the reoptimized soft lepton signal regions. I could show that the sensitivity is considerably increased when loosening the upper lepton p_T threshold to 50 GeV.

Possible reasons for the lack of sensitivity with soft lepton selections are the suboptimal modeling of ISR-Jets with the generator Herwig++ and the softer p_T spectra of pMSSM models compared to simplified ones due to richer particle spectrum and decay phenomenology.

The result can be possibly improved by more precise fudge factors (calculated from more suitable reconstructed and truth n-tuples) or usage of reconstructed samples. Perhaps one more inclusive selection would improve the results, too.

References

[1] Y. Abulaiti, B. Petersen, T. Eifert, T. Lari, and D. Cote. DM-inspired (well-tempered) 3rd generation pmssm grids. https://indico.cern.ch/getFile.py/access?contribId=2&resId=0&materialId=slides&confId=260215.

[2] M. Baak, M. Backes, A. Chitan, M.-H. Genest, N. Omachi, A. Koutsman, E. Khramov, M. Kuze, L. Morvaj, T. Nobe, Y. Sasaki, M. Tomoto, V. Tudorache, D. Xu, T. Yamanaka, and X. Zhuang. Search for Supersymmetry with compressed mass spectra in events involving soft-leptons, jets plus missing transverse momentum at $\sqrt{s} = 8\,TeV$. Technical Report ATL-COM-PHYS-2013-061, CERN, Geneva, Jan 2013.

[3] ATLAS Collaboration. Measuring the b-tag efficiency in a top-pair sample with 4.7 fb^{-1} of data from the ATLAS detector. 2012.

[4] ATLAS Collaboration. Measurement of the b-tag Efficiency in a Sample of Jets Containing Muons with 5 fb-1 of Data from the ATLAS Detector. 2012.

[5] M. Paterno. Calculating efficiencies and their uncertainties. May 2003.

[6] G. Polesello and . R. Tovey. Supersymmetric particle mass measurement with the boost-corrected contransverse mass. *JHEP*, 1003:030, 2010.