



Signatures of Scalar Dark Energy in $t\bar{t} + E_T^{miss}$ events in ATLAS

Áron Csaba Bodor, Eötvös Loránd University, Budapest, Hungary

September 9, 2017

Abstract

Here put a short abstract of what one can find in this document

Contents

1	Introduction	3
1.1	Dark energy	3
1.2	Cosmological constant problem	3
2	Theory	3
2.1	Characteristics, signals	4
3	Analysis	5
3.1	Aims and strategy	5
3.2	Processing signals	6
3.2.1	Applying the cuts to truth level signals	8
3.2.2	Preselection	8
3.2.3	Optimised selection for each signal region	9
3.2.4	Efficiencies	10
3.3	Significance and results	13
4	Conclusion and future plans	14
5	Acknowledgements	16

1 Introduction

1.1 Dark energy

The Standard Model of cosmology or the so-called Λ -CDM model had a lot of success in describing the cosmic microwave background, the abundance of light nuclei, the large scale distribution of galaxies and the accelerated expansion of the universe. The latter one requires the introduction of the cosmological constant, Λ . The energy density associated to this is called Dark Energy (later denoted by DE), which is the most dominant contribution to the total energy density of the universe.

1.2 Cosmological constant problem

When one tries to reconcile the standard model of particle physics (referred later to as SM) and of cosmology faces the cosmological constant problem. On a cosmological scale the SM could be thought of an effective field theory of matter (\mathcal{L}_M with the matter field denoted by Ψ), and gravity is classical with the action:

$$S = \int d^4x \sqrt{-g} \left(\frac{M_{Pl}^2}{2} R - \mathcal{L}_M(\Psi, g^{\mu\nu}) - \Lambda \right). \quad (1)$$

If one were to match the renormalized value of the cosmological constant, Λ_{ren} to the measured experimental value, one would find that the difference between the bare Λ and the renormalized one is 60 orders of magnitude on the tree level. So we can say less technically that the cosmological constant arising from cosmology is far less with its $(\text{meV})^4$ value than the one coming from SM with its $\gtrsim (\text{TeV})^4$. This is referred to as the cosmological constant problem. See a more detailed portrayal of this phenomena in [4].

To solve the problem one needs the UV-completion of the above theory. One way of doing this is to include new particles associated to DE into the SM. This way the existence of these particles could be put to test by colliders - at least on terrestrial scales.

2 Theory

In developing a theory to extend the SM two things should come to one's mind: simplicity and generality. We will turn our heads to the so-called Horndeski-theories (see [1]), the effective field theories which implement these features as they are the most general description of scalar DE particles (produced by the field $\hat{\Phi}$) that are coupled to gravity and to SM matter universally through the following metric:

$$g_{\mu\nu} = A(\Phi, X) \tilde{g}_{\mu\nu} + B(\Phi, X) \partial_\mu \Phi \partial_\nu \Phi, \quad (2)$$

where in our case \tilde{g} will become the flat metric, because in colliders one can neglect gravitational effects. In (2) the symbol X stands for $\frac{1}{2} \partial^\mu \Phi \partial_\mu \Phi$ and functions A and B

could be expanded in the powers of (X/M^4) , where M is the cutoff scale of the effective field theory. With this expansion and the requirement of shift-symmetry - the model should be symmetric regarding the transformation $\Phi \rightarrow \Phi + c$ - the lowest dimension interaction operators we can have are denoted with \mathcal{L}_1 and \mathcal{L}_2 :

$$\mathcal{L}_1 = \frac{\partial_\mu \Phi \partial^\mu \Phi}{M^4} T_\nu^\nu, \quad (3)$$

$$\mathcal{L}_2 = \frac{\partial_\mu \Phi \partial^\nu \Phi}{M^4} T_\nu^\mu, \quad (4)$$

where T denotes the energy-momentum tensor of the SM. There are further operators suppressed by the M cutoff-scale but we are for now only interested by these two. The collider phenomenological model implementing these interactions reads:

$$\mathcal{L}_{BSM} = \mathcal{L}_{SM} + \sum_i C_i \mathcal{L}_i + \frac{1}{2} m^2 \Phi^2, \quad (5)$$

where the factors C_i called the Wilson-factors, we set these to be one for \mathcal{L}_1 and \mathcal{L}_2 and zero for anything with higher dimensions. The mass of the DE particle, m is set to be really small compared to particle physics scales.

2.1 Characteristics, signals

Working out the interaction with the SM, we can say that \mathcal{L}_1 will scale with SM masses, because of the trace of the energy-momentum is determined by these parameters. That's why we try to look for this interaction in $t\bar{t} + E_T^{miss}$ final states. The other operator, \mathcal{L}_2 is determined by the mass transfer, so searching for this interaction could be more successful with monojet + E_T^{miss} final states, as momentum transfer could be the highest this way. As the theory has universal coupling to matter, we will try to observe through pair production induced by a top quark and by a gluon, which decays to $t\bar{t}$ after. See the according graphs on Figure 1.

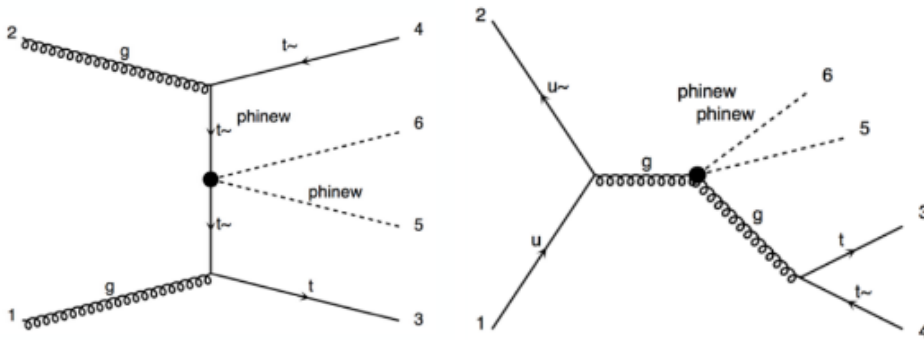


Figure 1: Feynman graph of the possible interaction of SM and DE through pair production. The label "phnew" denotes the DE particle.

3 Analysis

3.1 Aims and strategy

The aim of my project was twofold: investigate whether signal regions from previous analyses are usable for the search of the specified DE signal and if so, to put stricter constraints on the model's M parameter by exclusion. To do this I obtained signal samples of the model. These samples were on truth (or particle) level as reconstructed ones were not yet generated. The signal regions were from previous dark matter and supersymmetry searches, namely [2, 3]. The corresponding truth level signal samples used as benchmark for these regions were made available for me and the expected number of background events (on reconstructed level, both before and after fitting) was accessible in the internal notes. These resources proved enough for a truth level study on the DE model.

The mentioned dark matter signal regions were optimised to the processes of dark matter pair production via a scalar mediator, Φ . The process is parametrised by the mass of the mediator, m_Φ . The graph of this process could be seen on Figure 2.

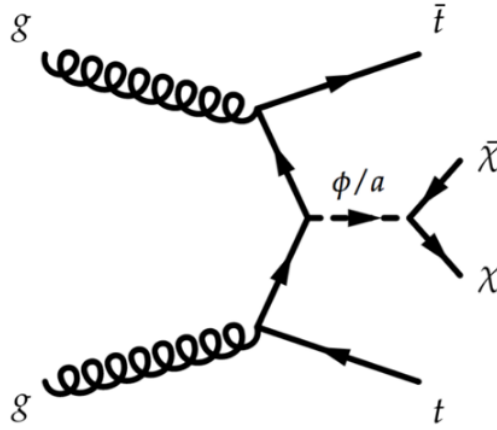
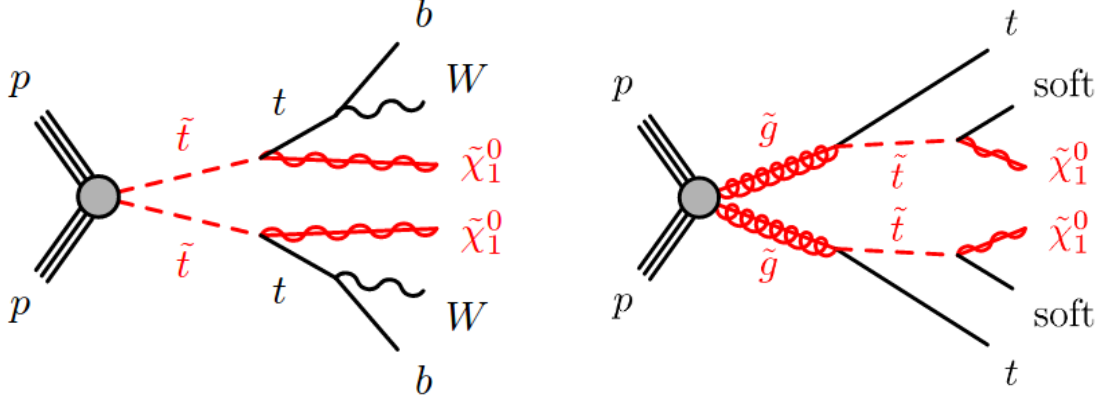


Figure 2: Feynman diagram for the dark matter pair production process via a scalar mediator.

In the case of supersymmetry two schemes were considered: a process in which stop-pairs produced directly and one in which they are from the decay of gluinos (gluino mediated stop production). These processes are governed by the supersymmetric particle masses, namely $m_{\tilde{g}}$, $m_{\tilde{t}}$ $m_{\tilde{\chi}}$. Figure 3 show the graphs of these processes.



(a) Direct stop production

(b) Gluino mediated stop production

Figure 3: Feynman graph of supersymmetry processes used to develop the signal regions.

3.2 Processing signals

The signal regions used were 'HIGH' and 'LOW' from the dark matter analysis [2], and 'A', 'B', 'E' from the supersymmetry investigation [3]. In table 1 the relevant physical parameters of the signals used to optimise the region are listed.

Before starting the detailed analysis it is advised to have a look at the E_T^{miss} distributions of the processes. Comparing the generated DE signals to DM on Figure 4 we can see that the DE signal has a harder E_T^{miss} spectrum. We can also see, that due to cuts in this variable (typically $E_T^{miss} > 300$ GeV) the resulting DMLOW yield from the according benchmark signal will be low. The kink in the DMLOW reference spectrum at around 60 GeV is due to the skimming filter applied to the sample.

We can also see that the SUSY signals have harder spectrum averaging in all cases above the DM ones. The SRA and SRE reference signal averages above the DE signal, but not the SRB. We will see that this is correlates to the yields after cuts.

Signal region name	Benchmark sample parameters
DMLOW	$m_\Phi = 20$ GeV
DMHIGH	$m_\Phi = 300$ GeV
SRA	$m_{\tilde{t}} = 1$ TeV, $m_{\tilde{\chi}} = 1$ GeV
SRB	$m_{\tilde{t}} = 600$ GeV, $m_{\tilde{\chi}} = 300$ GeV
SRE	$m_{\tilde{g}} = 1.7$ TeV, $m_{\tilde{t}} = 400$ GeV

Table 1: Physical parametrisation of the benchmark samples.

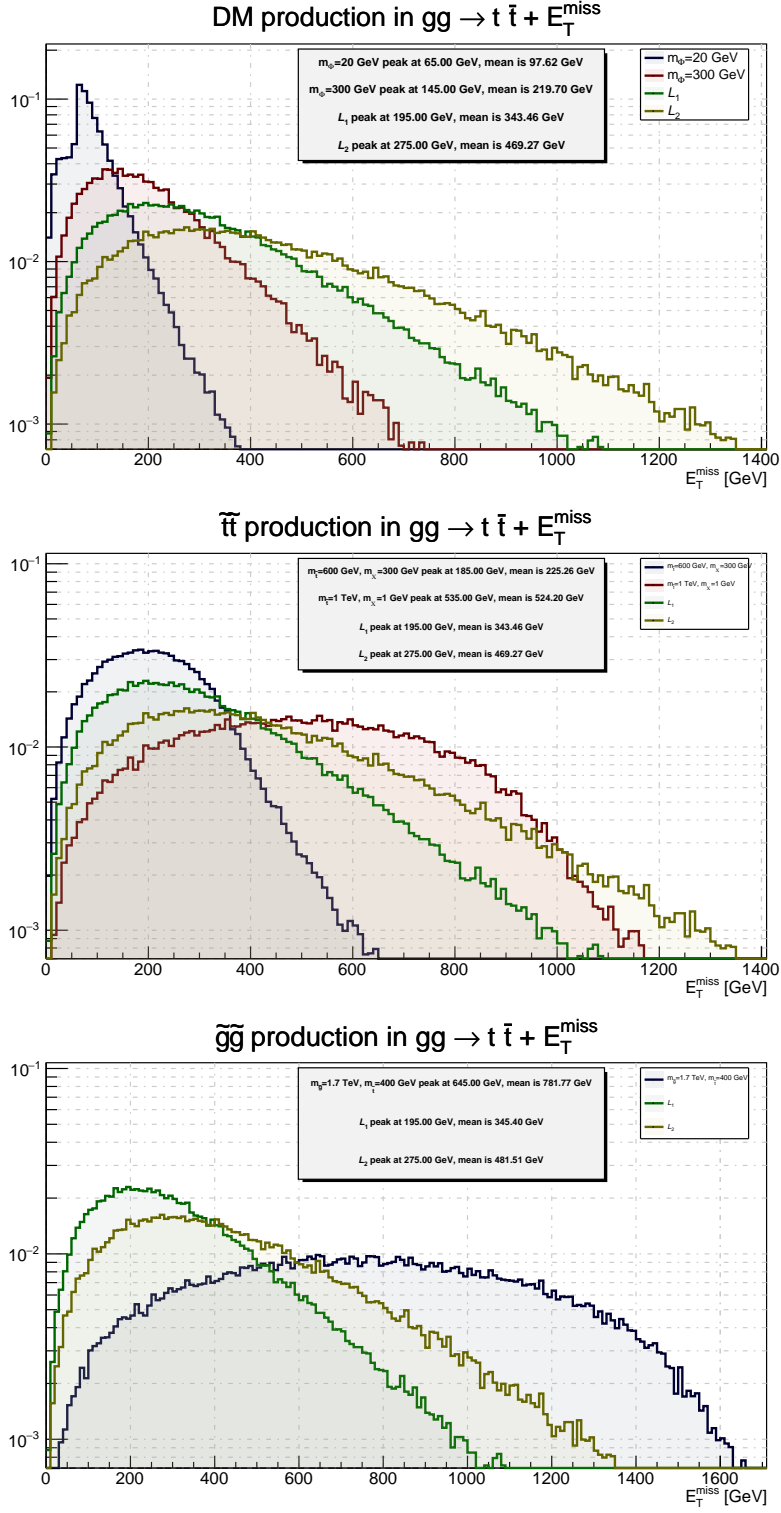


Figure 4: Missing transverse energy spectra of the reference signals (blue, red) compared to the DE signals (green, yellow).

3.2.1 Applying the cuts to truth level signals

The cuts prescribed in [2, 3] are developed using reconstructed samples. Issues arise if one tries to apply them directly to truth level samples, because in some cases the required variable in which the cut is made is not defined at truth level. During the analysis two of these issues were addressed, one of which concerns the τ -veto, see that later.

The other problem considered is connected to the b -jettagging efficiency. As on truth level we don't have the MV2c10 b -tagging discriminant, all of them considered tagged with 100% efficiency. However the cuts require a 77% tagging efficiency, so we addressed the issue with random generating numbers for all b -jets from a uniform distribution on $[0, 1]$. Those with scores below 0.77 remain a b -jet, the others are not considered as b -jets any more. This implies that all b -jetconnected variables (e.g. $m_T^{b, min}$) has to be recalculated during this process.

3.2.2 Preselection

The DM and SUSY analyses both define similar preselection cuts to insure that we are defining signal regions in the correct decay channels and to improve purity. Aiming to capture the correct decay channel we apply:

- lepton-veto (electrons, muons),
- number of jets at least four with $p_T > [80, 80, 40, 40]$ GeV,
- number of b -jets at least two,
- τ (-jet) veto (except for SRE).

These cuts comes trivially seeing the decay scheme. As it was pointed out in the previous subsection some issues arise with the last cut, as it requires the reconstructed variable $E_T^{miss, track}$, the missing transverse energy measured by the tracker. To circumvent these a truth level τ - jet overlap removal was applied the following way:

1. select the non- b -jet closest to E_T^{miss} with $|\Delta\Phi(E_T^{miss}, \text{jet})| < \pi/5$,
2. scan truth level τ particles, reject event if $\Delta R(\text{truth } \tau, \text{selected jet}) < 0.2$.

To reduce the background and reject fake events from mismeasured jet p_T :

- $E_T^{miss} > 300$ GeV (DM), 250 GeV (SUSY),
- $|\Delta\Phi(p_T, E_T^{miss})| > 0.4$ for the leading 4 (DM) or 3 (SUSY) jets.

3.2.3 Optimised selection for each signal region

To maximise the expected sensitivity more detailed cuts are applied, this time these are different for all signal regions. The discriminating variables used to make these cuts:

- m_{AkT8} , m_{AkT12} : reclustered jet masses with anti- k_T parameters $R = 0.8$ and $R = 1.2$; cuts in these variables ensure that we select events with W -boson and top-quark involved,
- $m_T^{b, min}$, $m_T^{b, max}$: transverse mass of the b -jet lying closest and furthest to E_T^{miss} ; these variables provide good discrimination to the semileptonic $t\bar{t}$ -background.
- H_T : scalar sum of the jet transverse momenta,
- $E_T^{miss} / \sqrt{H_T}$: missing transverse energy significance,
- $\Delta R(b, b)$: the distance of the two b -jets with highest MV2c10 scores; discriminates against the $Z + b$ -jets background, where the b -jets arise from gluon-splitting,
- m_{T2} : also known as stransverse momentum.

The concrete cuts for each region are listed in Table 2. As mentioned in 3.2.1 every b -jetrelated variable had to be recalculated during the analysis. This could be carried unambiguously for all variables of this kind, except for $\Delta R(b, b)$. As we have no proper discrimination on which two b -jets to use, all possible combinations were calculated and the smallest ΔR was chosen.

Cut	DMLOW	DMHIGH	SRA	SRB	SRE
m_{AkT8}^0	>80	-	>60	-	>120
m_{AkT8}^1	>80	-	-	-	>80
m_{AkT12}^0	-	>140	>120	>120	-
m_{AkT12}^1	-	>80	>120	>120	-
$m_T^{b, min}$	>150	>200	>200	>200	>200
$m_T^{b, max}$	>250	-	-	>200	-
m_{T2}	-	-	>400	-	-
H_T	-	-	-	-	>800
$E_T^{miss} \sqrt{H_T} [\sqrt{\text{GeV}}]$	-	>12	-	-	>18
E_T^{miss}	-	-	>400	-	>550
$\Delta R(b, b)$	>1.5	>1.5	-	>1.2	-

Table 2: Cuts specified for each investigated signal region. All variables are understood in GeV, except for the missing transverse energy significance, where units are shown explicitly.

3.2.4 Efficiencies

After applying the cuts, the efficiencies could be calculated for all samples and regions:

$$\text{Efficiency} = \frac{\# \text{ events after cuts}}{\# \text{ events before cuts}}. \quad (6)$$

Evaluating (6) after each cut results in the cutflows. Plotting these can show which cut made the most difference to a signal sample. Cutflow plots are also a suitable tool to compare different signals in the same region. Figures 5 to 9 show these plots. Referring back to the introduction of section 3.2 it is apparent that whichever signal has the harder E_T^{miss} spectrum will dominate in terms of efficiency after the cuts. This correlates with the fact that the cuts in E_T^{miss} made the most change in the efficiencies (not considering the necessary cuts of the decay channel). The numerical values in form of percentages of the efficiencies can be found in Table 3.

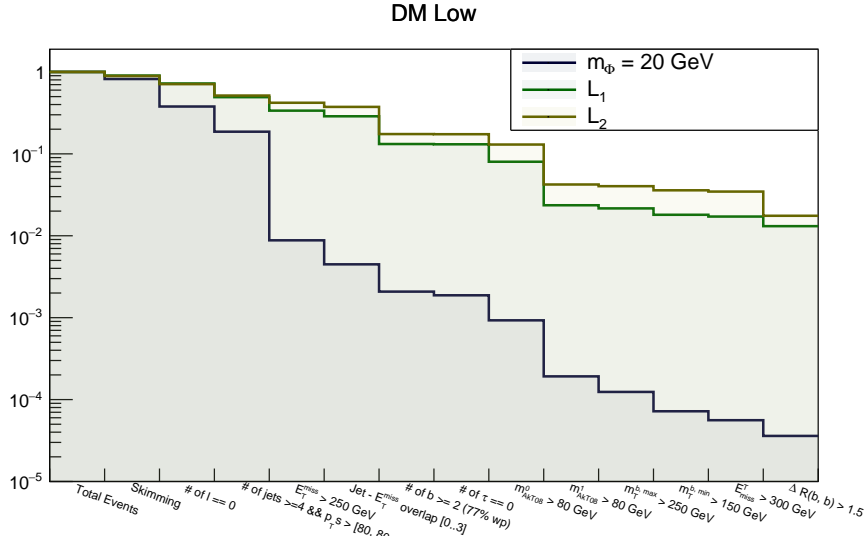
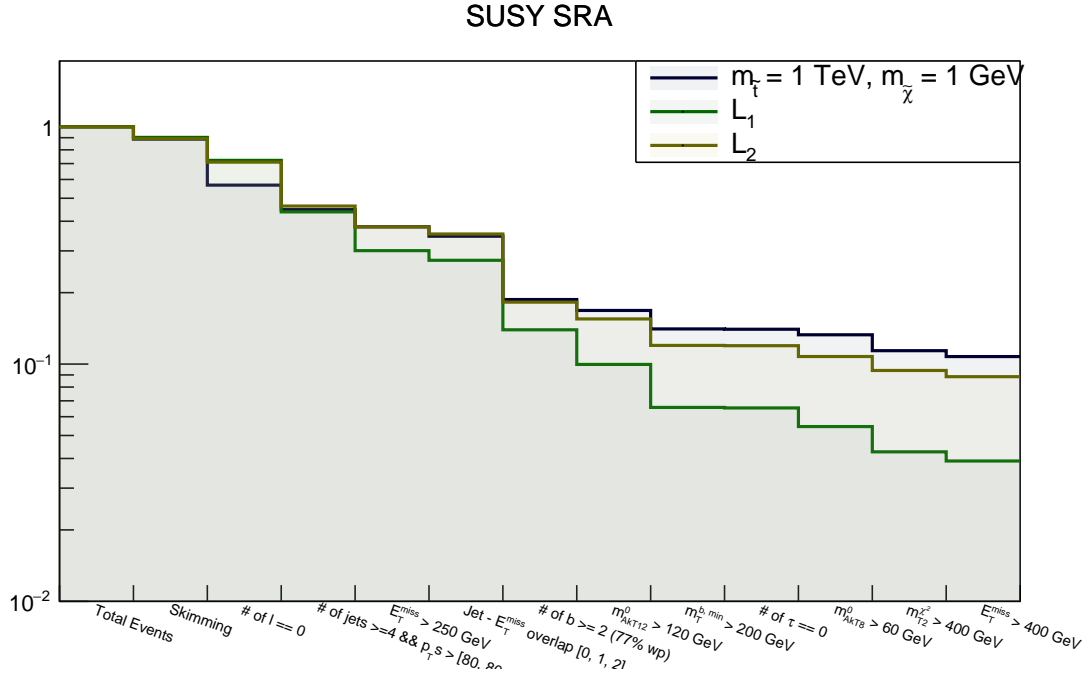
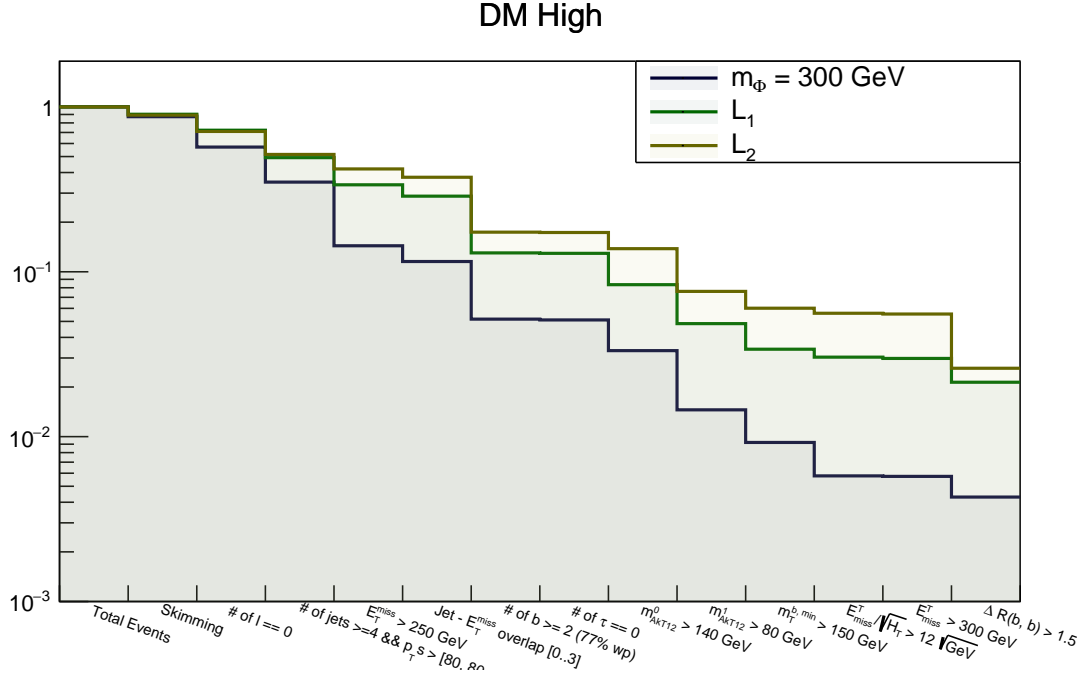


Figure 5: Cutflow plot of the signal region DMLOW.

Signal region	Reference	\mathcal{L}_1	\mathcal{L}_2
DMLOW	0.0036 %	1.31 %	1.75 %
DMHIGH	0.43 %	2.14 %	2.60 %
SRA	10.76 %	3.90 %	8.85 %
SRB	1.48 %	5.22 %	6.90 %
SRE	7.00 %	0.56 %	1.63 %

Table 3: Efficiencies of DE samples compared to the benchmark samples.



SUSY SRB

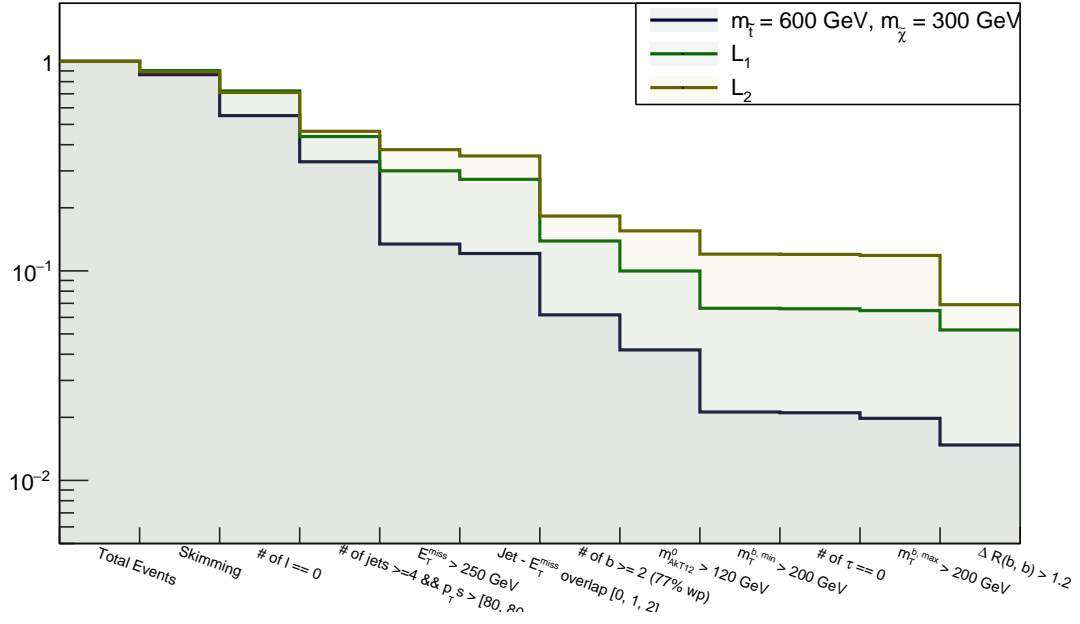


Figure 8: Cutflow plot of the signal region SRB.

SUSY SRE

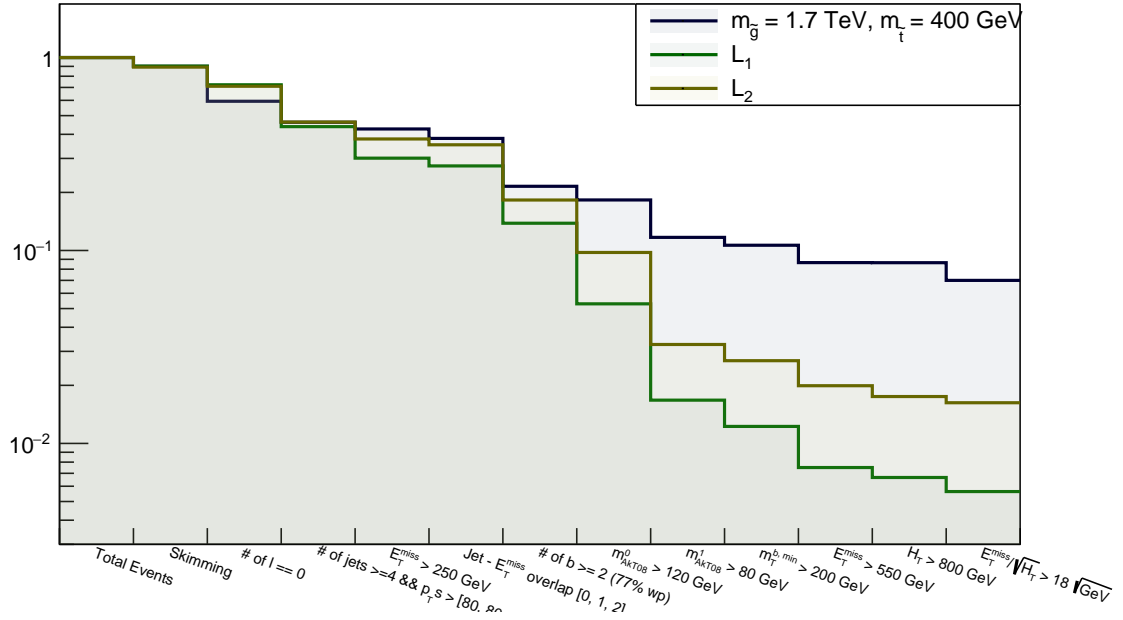


Figure 9: Cutflow plot of the signal region SRE.

3.3 Significance and results

Using the results from Subsection 3.2 we can calculate the truth level signal yields (denoted by S) for arbitrary luminosity by the formula

$$S(\text{SR}, \text{sample}, \text{truth}) = \sigma_{LO}(\text{sample}) \times k(\text{sample}) \times \int \mathcal{L} dt \times \text{Efficiency}(\text{SR}, \text{sample}, \text{truth}), \quad (7)$$

where the k factor is ratio of the NLO and LO cross sections. We will denote the expected number of background events before fit by B . The values of these for all regions could be found in [] for the DM analysis and in [] for the SUSY one. It is important to note, that these background yields are given after reconstruction, so comparing signal to it is not a one-by-one comparison. To quantify the significance we used the formula:

$$\text{Significance}(\text{SR}, \text{sample}, \text{truth}) = \frac{S(\text{SR}, \text{sample}, \text{truth})}{\sqrt{B(\text{SR}, \text{reco}) + (0.3B(\text{SR}, \text{reco}))^2}}. \quad (8)$$

In (8) we are comparing the signal yield to the statistical and systematic uncertainty of the background. The 30% systematic is an assumption adopted from the SUSY analysis. On Figure 10 the significance is shown over the M -scale parameter of the model. This plot shows that there is a possibility of being sensitive to the model (the constraints put on the model by [?])s $M >$ for \mathcal{L}_1 and $M >$ for \mathcal{L}_2 in case of $t\bar{t} + E_T^{\text{miss}}$ final states).

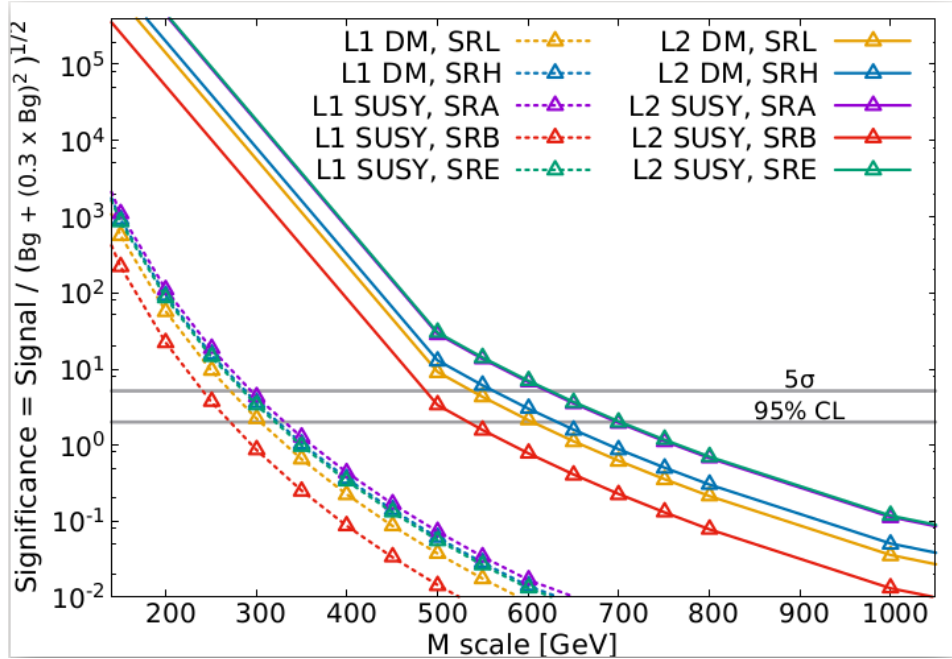


Figure 10: Significances on truth level. It apparent that we can be sensitive to \mathcal{L}_2 to higher scale than to \mathcal{L}_1 , but monojet analyses expected to be more significant to \mathcal{L}_2 than $t\bar{t}$ analyses.

To have a rough estimation of what could we expect with reconstructed signal samples we introduced a transfer factor (T) which describes a rescaling of the signal yield

calculated by

$$T(\text{SR}) = \frac{S(\text{SR}, \text{benchmark sample}, \text{reco})}{S(\text{SR}, \text{benchmark sample}, \text{truth})}. \quad (9)$$

This formula ought to be calculated for all signal regions with the according benchmark sample. The reconstructed level signal yields are available in [2] and [3]. Using (8) and (9) we can derive a naive extrapolation from the truth level significance to the reconstructed level significance:

$$\text{Significance}(\text{SR}, \text{sample}, \text{reco}) = T(\text{SR}) \times \text{Significance}(\text{SR}, \text{sample}, \text{truth}). \quad (10)$$

On Figures 11 and 12 we can see this extrapolated significance and the according transfer factors and background events for \mathcal{L}_2 and \mathcal{L}_1 . It is apparent that according to this analysis we may in some parameter regions - $M < 600$ GeV for \mathcal{L}_2 and $M < 300$ GeV for \mathcal{L}_1 using the best available signal regions - very well be sensitive to the dark energy signals suggested by [1]. The fact that no significant signal was observed during the Run 2 in 2015 and 2016 inspire the idea of making exclusions - or rather enforcing the ones in place.

As mentioned in the theoretical introduction the monojet studies supposed to be more sensitive for the \mathcal{L}_2 operator, so an analysis on those final states probably will end up in much better exclusions. The result of this analysis is:

$$\boxed{\mathcal{L}_2 : M \gtrsim 700 \text{ GeV}.} \quad (11)$$

Turning now to \mathcal{L}_1 to which we are the more sensitive ones, we can say based on Figure 12 that

$$\boxed{\mathcal{L}_1 : M \gtrsim 300 \text{ GeV}.} \quad (12)$$

This is the main result of this analysis as it improved the exclusion in [1], which is $M \gtrsim 237.4$ GeV.

Although it seems as a good result, it is important to note two issues. Calculating the transfer factor in signal region DM-HIGH according to (9) resulted in a number greater than one, namely 2.6731. This result is hard to interpret as it means that applying more cuts would result in higher yield. The source of the issue is unknown to date.

The other problem is that samples were not fitted to data. The source of this is that I met issues running the program `HistFitter` which is usually used to create exclusion fits by the SUSY analysis teams. So the limits in (11) and (12) are to be treated with reservation. On the other hand, as this study will be reproduced in the near future with reconstructed samples the absence of proper exclusion fits is not significant.

4 Conclusion and future plans

As written in the previous subsection I achieved the aims set in 3.1 to some extent. I have managed to show that previous signal regions are perceptive to the dark energy

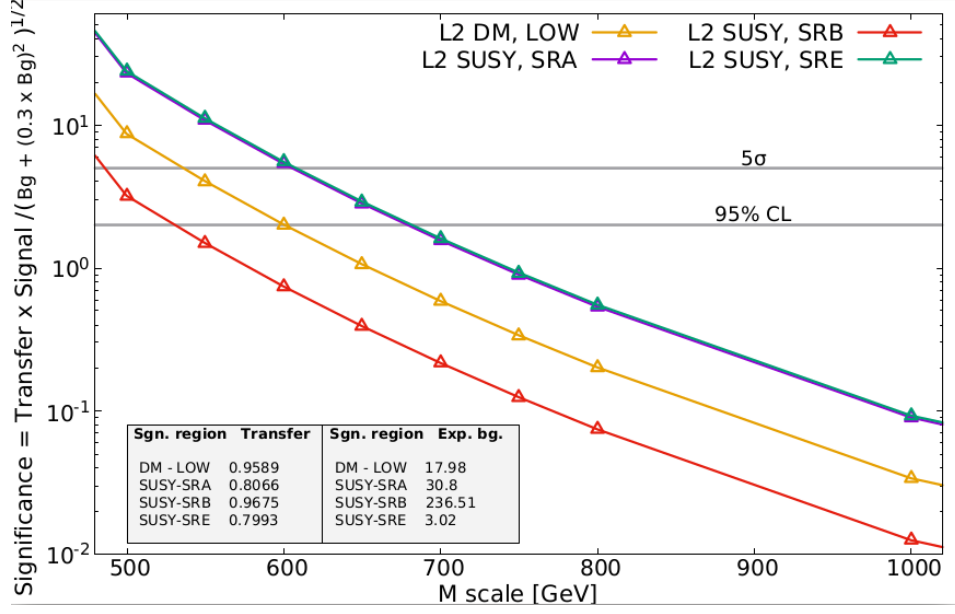


Figure 11: Significances involving the \mathcal{L}_2 operator extrapolated to reconstructed level, see (10).

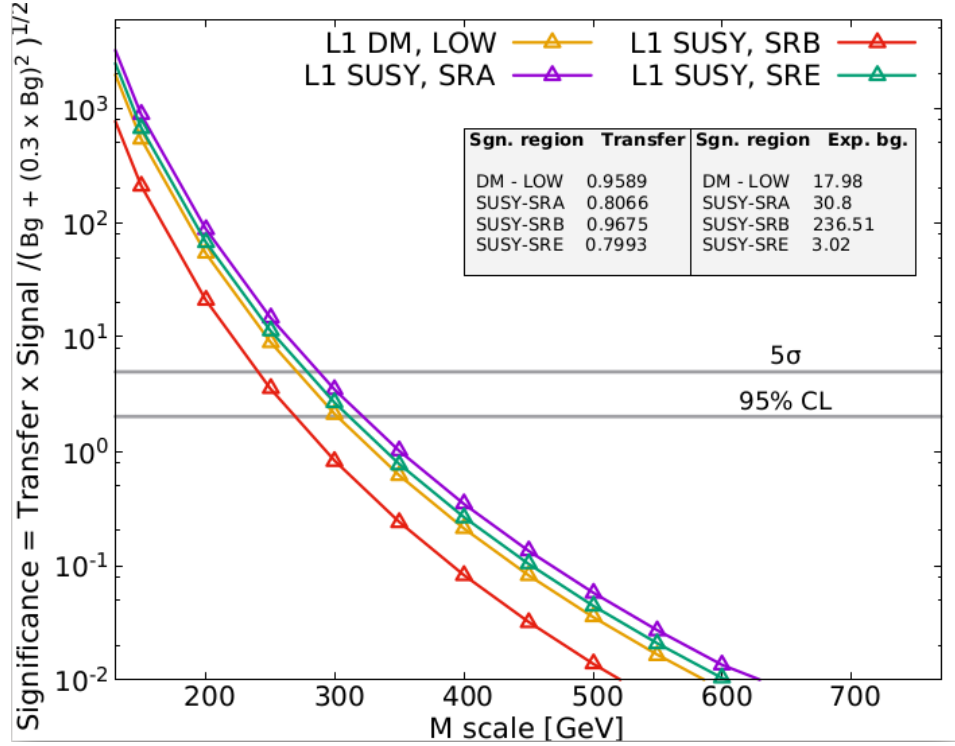


Figure 12: Significances involving the \mathcal{L}_1 operator extrapolated to reconstructed level. It is apparent that we can expect better exclusion limits with a detailed analysis than [1] according to this naive extrapolation.

model. The cuts could be improved to maximise sensitivity, possibly as a part of a dedicated analysis. I made a rough estimate on the expected significance which proved as a good tool to exclude the model up to higher values in it's parameter.

By the production of reconstructed level dark energy samples these studies will continue on by reproducing this analysis. Hopefully it will result in similar numbers, excluding the dark energy model the same way. This would be the first study published on dark energy exclusion made by hadronic colliders.

5 Acknowledgements

I am deeply indebted to my supervisors, Krisztian Peters and Matthias Saimpert for selecting me to complete this project and for their guidance and care which was accompanied by the outmost patience. I hereby say thank you for all of my friends whom i met here for making this stay an absolutely incredible time. I am very grateful to the organisers as well for making this program. I am beholden to my girlfriend and family for letting me miss out on summer and for their encouragements.

References

- [1] LHC Signatures Of Scalar Dark Energy. arXiv:1604.04299v1 (2016). *P. Brax, C. Burrage, C. Englert, M. Spannowsky*
- [2] Search for Dark Matter associated production with top quarks in pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector at the LHC. Atlas Internal Note, 21 June 2017. *Y. Afik et al.*
- [3] Search for the Supersymmetric Partner of the Top Quark in the Jets + E_T^{miss} Final State at $\sqrt{s} = 13$ TeV. Atlas Internal Note, 18 May 2017. *M. Albergo et al.*
- [4] Lectures on the Cosmological Constant Problem. arXiv:1502.05926v1 (2015). *Antonio Padilla*