Search for dark sector showering in ATLAS using semi-visible jets

Sukanya Sinha

School of Physics, University of Witwatersrand, Johannesburg, South Africa

E-mail: sukanya.sinha@cern.ch

Abstract. Recent studies in particle physics have shown that there are myriad possibilities for strong dark sector studies at the LHC. One signature is the case of semi-visible jets, where parton evolution includes dark sector emissions, resulting in jets overlapping with missing transverse energy. Owing to the unusual MET-along-the-jet event topology, this is mostly an unexplored domain within ATLAS. In this contribution, the public results of the first t-channel ATLAS search for semi-visible jets are discussed, that focussed on overcoming the performance and optimisation challenges associated with such a unique final state, specifically looking at the angle difference between the hardest jet and the missing transverse energy.

1. Introduction

Collider searches for Dark Matter (DM) until this date have mostly focussed on scenarios where DM particles are produced in association with either heavy Standard Model (SM) particles, photons or jets. However, no confirmed evidence of DM has been observed so far. Several models [1, 2, 3, 4] have been proposed that include a strongly-coupled dark sector, giving rise to unusual and unexplored collider topologies. Semi-visible jets (SVJ) [5] is one such example. A result in the s-channel production mode has been presented by the CMS collaboration [6]. Here, a preliminary result from the ATLAS search for the t-channel production mode [7] is presented. Searches for t-channel production modes allow to probe a broad class of non-resonant signals and can potentially have higher mass reach, as they are not limited only to finding resonance peaks as in the s-channel searches.

2. Signal and background modelling

In the t-channel production mode, the scalar bi-fundamental mediator (Φ) acts as a portal between SM and dark sectors. It splits into two dark quarks, which decay partially to dark hadrons, and partially to SM hadrons as shown in Fig. 1. The ratio of the rate of stable dark hadrons over the total number of hadrons in the event is termed R_{inv}, which can be set in the model. This results in reconstructed jets geometrically encompassing the dark hadrons, termed semi-visible jets (SVJ). The resultant missing transverse momentum ($E_{\text{T}}^{\text{miss}}$) direction is aligned along one of the jets. As this is also a signature of mis-measured jets in a detector, this class of jets as new physics signal has been mostly unexplored at the Large Hadron Collider (LHC).

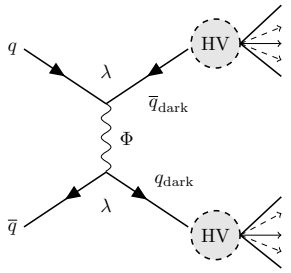

Figure 1. The Feynman diagram and subsequent production mechanism for semi-visible jets. The mediator Φ decays to two dark quarks, which decay partially to dark hadrons, and partially to SM hadrons, governed by R_{inv} fraction. The coupling is denoted by λ. HV denotes the Pythia8 Hidden Valley module interactions connecting the dark sector with the SM sector.

The Madgraph [8] event generator at leading order, with up to two extra jets was used to generate matrix element (ME) level signal events with the DMsimp_tchannel model [9] and NNPDF30LO [10] PDF. The pair production of mediators with their decays governed by the same R_{inv} fraction would result in roughly back-to-back semi-visible jets. The addition of extra jets boosts the hadronic activity, and consequently the $E_{\text{T}}^{\text{miss}}$ in the event, thereby making the search viable over the Standard Model background. The dark hadron mass is set to be 10 GeV based on [5], while the mediator mass is varied within the range of 1000 - 5000 GeV in 500 GeV intervals. Another free parameter in the model is the coupling connecting the SM and DM sectors, λ (as shown in Fig. 1). The nominal samples are generated with $\lambda = 1$ but λ can be varied between $10^{-4} - 4\pi$ [11, 12]. The cross-section is scaled as λ^4 without having any impact on the kinematic distributions or on the validity of the model if the mediator mass is 2.5 TeV or higher, as at lower mediator masses, there are contributions from s-channel processes as well. The Hidden Valley (HV) module [13] of Pythia8 [14] is used to shower the ME level event, using the NNPDF2.3LO [10] parton distribution function (PDF) set and the A14 set of tuned parameters [15]. The MLM [16] jet matching scheme, with xqcut of 100, is employed. The samples were generated with R_{inv} values of 0.2, 0.4, 0.6 and 0.8 for each mediator mass.

The contributing background processes are multijet, $t\bar{t}$, single top, Z/W+jets, and diboson, generated using Pythia8.230, PowhenBoxV2 [17, 18, 19] plus Pythia8.230 for showering, and Sherpa2.2.11 [20, 21] respectively. The Monte Carlo (MC) simulated samples are processed through the full ATLAS detector simulation [22] based on Geant4 [23], and then reconstructed and analysed using the same procedure and software that are used for the data.

3. Analysis Strategy

This analysis uses 139 fb^{-1} of proton-proton collision data with 25 ns bunch spacing collected by ATLAS from 2015 to 2018. The description of the ATLAS detector can be found elsewhere [24]. The $E_{\text{T}}^{\text{miss}}$ is defined as the negative vectorial sum of the transverse momenta of all selected objects. Events in the analysis are selected with the un-prescaled $E_{\text{T}}^{\text{miss}}$ trigger having the lowest threshold [26]. Events with $E_{\text{T}}^{\text{miss}} > 200$ GeV are selected in order to be on trigger efficiency plateau. Particle-flow (PFlow) jets are constructed using the anti-k_t algorithm [27, 28] with a radius parameter of $R = 0.4$, using charged constituents associated with the primary vertex and neutral PFlow constituents as inputs [29]. Events are required to have at least two jets within $|\eta| < 2.8$, the leading jet is required to have $p_{\text{T}} > 250$ GeV, while other jets are required to have $p_{\text{T}} > 30$ GeV. Events are also required to have at least one jet within $\Delta\phi < 2.0$ of $E_{\text{T}}^{\text{miss}}$ direction, as the distance of the closest jet to $E_{\text{T}}^{\text{miss}}$ direction depends somewhat on the R_{inv} fraction. Jets are considered tagged as b-jets if they pass the 77% efficiency working

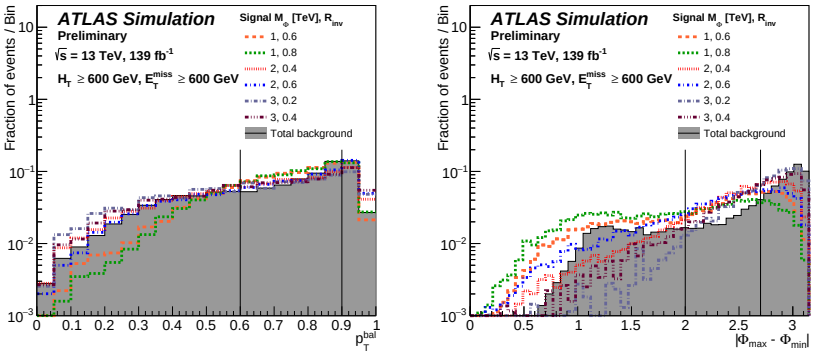

Figure 2. Comparisons of shape of $p_{\mathrm{T}}^{\mathrm{bal}}$ (a) and $|\phi_{\mathrm{max}} - \phi_{\mathrm{min}}|$ (b) distributions between the total background before the fit and six signal predictions covering a representative mediator mass and invisible fraction range. The solid vertical lines represent how these distributions are divided to form the nine-bin grid subsequently.

point of the *DL1r* algorithm [30]. Events with two or more b-tagged jets are vetoed to reduce $t\bar{t}$background contributions.

Any events with a τ-lepton candidate with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$ are rejected. Events with any electrons or muons satisfying $p_{\mathrm{T}} > 7$ GeV and within the tracking volume $|\eta| < 2.5$ are discarded for the nominal analysis. However, for estimation of leptonic backgrounds, additional leptonic selections are defined. In all cases, $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger is employed, and $E_{\mathrm{T}}^{\mathrm{miss}}$ is recalculated considering muons to be invisible, to mimic the nominal analysis trigger level $E_{\mathrm{T}}^{\mathrm{miss}}$ definition. Events with any electron with $p_{\mathrm{T}} \geq 7$ GeV are discarded. The muons must have $p_{\mathrm{T}} \geq 7$ GeV. The 1L selection requires exactly one muon and no b-tagged jet. The 1L1B selection requires exactly one muon as before, but exactly one b-tagged jet in addition. Finally, the 2L selection requires two opposite charged muons with the pair's invariant mass between 66 GeV and 116 GeV, and no b-tagged jets. The 1L-region is dominated by W+jets events, the 1L1B-region is dominated by semi-leptonic $t\bar{t}$and single top quark induced processes, and the 2L-region almost exclusively contains Z+jets events.

The analysis uses $E_{\mathrm{T}}^{\mathrm{miss}}$ and H_{T}, the latter defined as the scalar sum of p_{T} of jets in the event. The the region with $E_{\mathrm{T}}^{\mathrm{miss}} > 600$ GeV and $H_{\mathrm{T}} > 600$ GeV after the pre-selection is defined as the signal region (SR). The corresponding 1L, 1L1B and 2L control regions (CR) are defined using the muon and b-tagged jet requirements with the same $E_{\mathrm{T}}^{\mathrm{miss}}$ and H_{T} requirements as in the SR. Low and intermediate $E_{\mathrm{T}}^{\mathrm{miss}}$ validation regions (VR) for multijet process are defined by requiring $E_{\mathrm{T}}^{\mathrm{miss}}$ to be between 250 GeV to 300 GeV and between 300 GeV to 600 GeV respectively, with the same $H_{\mathrm{T}} > 600$ GeV requirement after the pre-selection. The CR and VRs have negligible signal contamination.

The search then makes use of other two key observables, which are found to be largely uncorrelated:

(i) the p_{T} balance between the closest jet (j_1) and farthest jet (j_2) from $E_{\mathrm{T}}^{\mathrm{miss}}$ direction, termed as $p_{\mathrm{T}}^{\mathrm{bal}}$, defined using two-dimensional p_{T} vectors:

$$p_{\mathrm{T}}^{\mathrm{bal}} = \frac{|\vec{p_{\mathrm{T}}}(j_1)+\vec{p_{\mathrm{T}}}(j_2)|}{|\vec{p_{\mathrm{T}}}(j_1)|+|\vec{p_{\mathrm{T}}}(j_2)|}.$$

(ii) the difference in the azimuthal angle between j_1 and j_2 as defined above, termed $|\phi_{\mathrm{max}} - \phi_{\mathrm{min}}|$:

Fig. 2 shows the signal against total background shape comparison for $p_{\mathrm{T}}^{\mathrm{bal}}$ and $|\phi_{\mathrm{max}} - \phi_{\mathrm{min}}|$ distributions for the nominal selection. There is a distinct shape difference between the different signal benchmark points, and the total background, which is utilised in designing the fit strategy. The $|\phi_{\mathrm{max}} - \phi_{\mathrm{min}}|$ and the $p_{\mathrm{T}}^{\mathrm{bal}}$ distributions are divided into nine bins as seen in Fig. 3. These bins are defined identically in SR and in each CR. Yields in these nine bins in each case are treated as the observables.

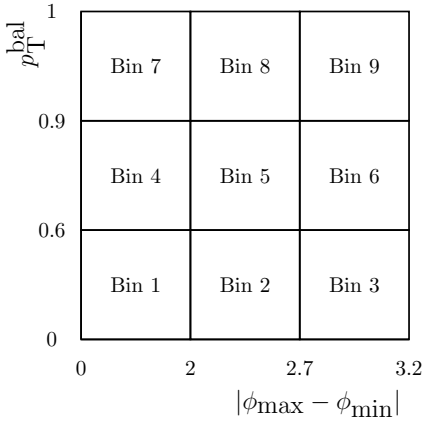

Figure 3. The definition of the 9-bins in $|\phi_{\mathrm{max}} - \phi_{\mathrm{min}}|$ and $p_{\mathrm{T}}^{\mathrm{bal}}$, defined identically in SR, VR and in each CR.

The background contributions are from multijet processes with mis-measurement of jet momenta and angles, and from W/Z+jets, diboson and semi-leptonic top processes with real $E_{\mathrm{T}}^{\mathrm{miss}}$. In order to address the known mismodelling in multijet MC samples, reweighting factors using an additional low $E_{\mathrm{T}}^{\mathrm{miss}}$ VR of 250–300 GeV were derived in multijet-rich bins 3, 6 and 9 and applied to the rest of the bins in the same $p_{\mathrm{T}}^{\mathrm{bal}}$ range. The background from $t\bar{t}$and W+jets arises either because an electron or a muon is not detected or a hadronically decaying τ lepton is misidentified as a jet.

The systematic uncertainties on signal and background yields and shapes result from experimental uncertainties and theoretical modelling effects. The former are due to the jet energy scale (JES) and resolution (JER), computation of $E_{\mathrm{T}}^{\mathrm{miss}}$ soft term, flavour-tagging performance, rescaling of simulation to match the pile-up profile in data, and an absolute uncertainty on the luminosity estimation. Uncertainties in the reconstruction, identification, isolation and trigger efficiencies of muons, electrons and τ-leptons and on their energy scale and resolution are also considered. Theoretical uncertainties common on the MC samples are due to renormalisation and factorisation scales and due to PDF choices. Additionally, initial and final state radiation scale uncertainties were assessed for $t\bar{t}$and single-top processes as well, treated as uncorrelated between them.

4. Results

In order to estimate the background, a simultaneous binned maximum likelihood function fit is performed using all the nine bin yields, using the MC templates, by employing the SR and the corresponding CRs (1L, 1L1B, and 2L). This is done to simultaneously search for the signal while improving the background prediction in the SR. The scale factors for the individual backgrounds are determined from the fit: The post-fit distributions of H_{T}, $E_{\mathrm{T}}^{\mathrm{miss}}$, $|\phi_{\mathrm{max}} - \phi_{\mathrm{min}}|$ and $p_{\mathrm{T}}^{\mathrm{bal}}$ observables are shown in Fig. 4 for the SR. Excellent agreement of data with SM background predictions are seen for all the observables.

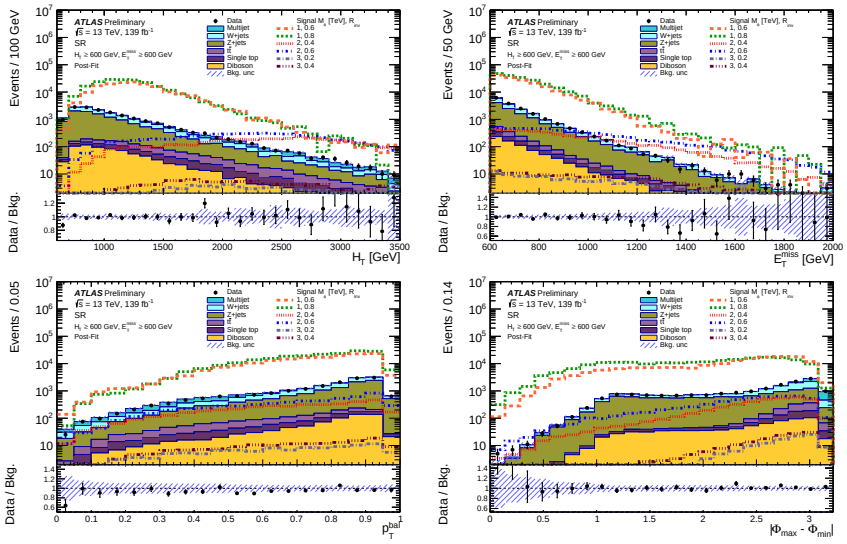

Figure 4. The post-fit distributions for H_{T} (a), $E_{\mathrm{T}}^{\mathrm{miss}}$ (b), $p_{\mathrm{T}}^{\mathrm{bal}}$ (c), and $|\phi_{\mathrm{max}} - \phi_{\mathrm{min}}|$ (d) are shown for the SR. Data is compared against background predictions, and six signal predictions covering a representative mediator mass and invisible fraction range are overlaid. The uncertainties include all systematic and statistical components. The last bin in (a) and (b) contains the overflow.

Upper limits on the contribution of events from new physics are computed by using the modified frequentist approach CL_s based on asymptotic formulas at 95% confidence level [31]. The nominal signal cross-sections for each signal mass point can be scaled by λ^4 for mediator masses larger than 2.5 TeV. For each mediator mass point, the limit on the cross-section is obtained, and the corresponding λ is calculated. This λ value corresponding to the cross-section upper limit is presented for the SR in Fig. 5. It can be seen that for lower mass points, the nominal cross-sections are excluded, whereas for higher mass points only higher values of cross-sections can be excluded. The advantage of this representation is that it sets stringent limits on the signature in general for a wide range of λ values, and can help in recasting this analysis for future model predictions.

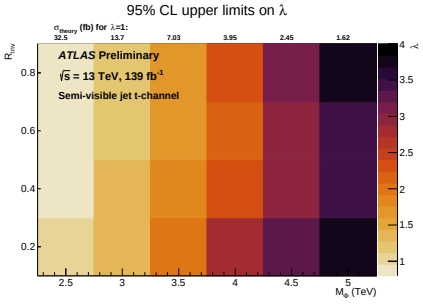

Figure 5. The grid shows the observed 95% CL upper limit on λ with M_ϕ on the x-axis, R_{inv} on the y axis. It also includes over each M_ϕ column the predicted cross-section for that specific mass value as a reference.

5. Summary

This article presents the first limits on the SVJ t-channel production for mediator masses ranging from 1000–5000 GeV, and for R_{inv} of 0.2–0.8. The observed yields are in agreement with the SM background expectations. The upper limits at the 95% confidence level on the mediator mass range from 2.4 TeV to 2.7 TeV, depending on the values of the invisible energy fraction. They are translated into upper limits on the coupling strength between the mediator, a Standard Model quark and a dark quark.

6. Acknowledgments

SS thanks University of Witwatersrand Research Council for Sellschop grant and NRF for Extension Support Doctoral Scholarship.

7. References

[1] Schwaller P, Stolarski D and Weiler A 2015 *JHEP* **05** 059 (*Preprint* 1502.05409)
[2] Cohen T, Lisanti M and Lou H K 2015 *Phys. Rev. Lett.* **115** 171804 (*Preprint* 1503.00009)
[3] Park M and Zhang M 2019 *Phys. Rev. D* **100** 115009 (*Preprint* 1712.09279)
[4] Bernreuther E, Kahlhoefer F, Krämer M and Tunney P 2020 *JHEP* **01** 162 (*Preprint* 1907.04346)
[5] Cohen T, Lisanti M, Lou H K and Mishra-Sharma S 2017 *JHEP* **11** 196 (*Preprint* 1707.05326)
[6] CMS Collaboration 2021 (*Preprint* 2112.11125)
[7] ATLAS Collaboration 2022 Search for non-resonant production of semi-visible jets using Run 2 data in ATLAS ATLAS-CONF-2022-038 URL http://cds.cern.ch/record/2815284
[8] Alwall J, Frederix R, Frixione S, Hirschi V, Maltoni F, Mattelaer O, Shao H S, Stelzer T, Torrielli P and Zaro M 2014 *JHEP* **07** 079 (*Preprint* 1405.0301)
[9] Dmsimp model https://github.com/smsharma/SemivisibleJets accessed: 2020-11-10
[10] Ball R D *et al.* (NNPDF) 2015 *JHEP* **04** 040 (*Preprint* 1410.8849)
[11] Cohen A G, Kaplan D B and Nelson A E 1997 *Phys. Lett. B* **412** 301–308 (*Preprint* hep-ph/9706275)
[12] Colucci S, Fuks B, Giacchino F, Lopez Honorez L, Tytgat M H G and Vandecasteele J 2018 *Phys. Rev. D* **98** 035002 (*Preprint* 1804.05068)
[13] Seth M S 2011 *A first study of Hidden Valley models at the LHC* Bachelor thesis Lund Observ. (*Preprint* 1106.2064)
[14] Sjöstrand T, Ask S, Christiansen J R, Corke R, Desai N, Ilten P, Mrenna S, Prestel S, Rasmussen C O and Skands P Z 2015 *Comput. Phys. Commun.* **191** 159 (*Preprint* 1410.3012)
[15] ATLAS Collaboration 2014 ATLAS Pythia 8 tunes to 7 TeV data ATL-PHYS-PUB-2014-021 URL https://cds.cern.ch/record/1966419
[16] Mangano M L, Moretti M, Piccinini F and Treccani M 2007 *JHEP* **01** 013 (*Preprint* hep-ph/0611129)
[17] Nason P 2004 *JHEP* **11** 040 (*Preprint* hep-ph/0409146)
[18] Frixione S, Nason P and Oleari C 2007 *JHEP* **11** 070 (*Preprint* 0709.2092)
[19] Alioli S, Nason P, Oleari C and Re E 2010 *JHEP* **06** 043 (*Preprint* 1002.2581)
[20] Bothmann E *et al.* 2019 *SciPost Phys.* **7** 034 (*Preprint* 1905.09127)
[21] Schumann S and Krauss F 2008 *JHEP* **03** 038 (*Preprint* 0709.1027)
[22] ATLAS Collaboration 2010 *Eur. Phys. J. C* **70** 823 (*Preprint* 1005.4568)
[23] GEANT4 Collaboration, Agostinelli S *et al.* 2003 *Nucl. Instrum. Meth. A* **506** 250
[24] ATLAS Collaboration 2008 *JINST* **3** S08003
[25] ATLAS Collaboration 2018 *Eur. Phys. J. C* **78** 903 (*Preprint* 1802.08168)
[26] ATLAS Collaboration 2020 *JHEP* **08** 080 (*Preprint* 2005.09554)
[27] Cacciari M, Salam G P and Soyez G 2008 *JHEP* **04** 063 (*Preprint* 0802.1189)
[28] Cacciari M, Salam G P and Soyez G 2012 *Eur. Phys. J. C* **72** 1896 (*Preprint* 1111.6097)
[29] ATLAS Collaboration 2017 *Eur. Phys. J. C* **77** 466 (*Preprint* 1703.10485)
[30] ATLAS Collaboration 2020 *JHEP* **03** 145 (*Preprint* 1910.08447)
[31] Cowan G, Cranmer K, Gross E and Vitells O 2011 *Eur. Phys. J. C* **71** 1554 (*Preprint* 1007.1727)