Validation note : implementation of the ATLAS-EXOT-2022-37 analysis in the MadAnalysis 5 framework

Implementation : Sukanya Sinha
Validation : Thomas Wojtkowski

1 Introduction

This note describes the implementation of the analysis ATLAS-EXOT-2022-37 [1] in MadAnalysis 5 ([2], [3]), that is now available in its Public Analysis Database [4], and that uses the simplified fast detector simulation SFS [5]. This analysis is a search for semi-visible jets [6], i.e jets accounting for a significant contribution to the event's missing transverse energy and that can arise from a strongly interacting dark sector. These jets contain Standard Model (SM) particles in addition to stable dark hadrons which are invisible in a detector.

The analysis considers a production of semi-visible jets via a t-channel mediator Φ that produces a pair of dark quarks q_{dark} (Figure 1). These dark quarks then undergo parton shower and hadronization in the dark sector to form stable and unstable dark hadrons, before these latter decay to the visible sector. The parameters for the signal are the mass of the mediator m_Φ, the fraction of stable dark hadrons among all the dark hadrons in the event noted R_{inv}, and the coupling constant λ between the SM quark, the dark quark and the mediator.

The analysis uses 139 fb^{-1} of data at $\sqrt{s} = 13$ TeV, collected with the ATLAS detector of the LHC between 2015 and 2018. Events are required to have two reconstructed jets where one of these can be aligned with the direction of the missing transverse momentum.

The ATLAS collaboration made available substantial additional data via HEPData at `https://www.hepdata.net/record/ins2663256`, including cutflows and signal region yields for a few representative models, and the number of observed events and expected background with uncertainty in each signal region bin.

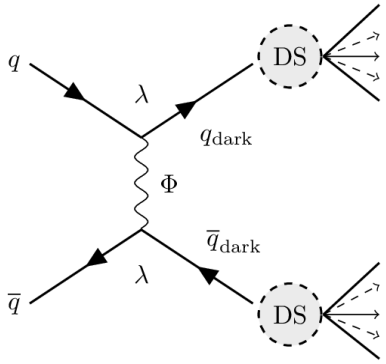

FIGURE 1 – Production of semi-visible jets jets via a t-channel mediator ϕ. DS denotes the dark shower.

2 Description of the analysis

2.1 Events selection

The events are required to pass a trigger with a threshold on the missing transverse energy E_T^{miss}. The threshold value was 70 GeV in 2015, and reached 110 GeV during 2016-2018 data-taking period. Particle-flow jets are constructed using the anti-k_t algorithm [7] with a radius of 0.4.

Events must have 2 reconstructed jets within $|\eta| < 2.8$, the leading jet must satisfy $p_T > 250$ GeV, and the other jets must have $p_T > 30$ GeV. Moreover, at least one jet has to be within $\Delta\phi = 2.0$ of the $\vec{p}_T^{miss}$ direction.

Events with 2 or more b-jets are rejected (the identification of b-jet is done using the DL1r algorithm [8] at the 77 % efficiency working point).

Hadronically decaying τ-lepton candidates are formed thanks to a Recurent Neural Network algorithm [9] with information from the calorimeters and inner tracking detectors, and all events with such a τ lepton with $p_T > 20$ GeV and $|\eta| < 2.5$ are rejected.

2.2 Signal regions

The signal region is defined with $E_T^{miss} > 600$ GeV and $H_T = \sum_{jets} p_T > 600$ GeV.

In the signal region, 9 bins are defined with two uncorrelated variables, which are :

- the p_T balance between the closest (j_1) and farthest (j_2) jets in azimuth from the $\vec{p}_T^{miss}$ direction defined as $p_T^{bal} = \frac{|\vec{p}_T(j_1)+\vec{p}_T(j_2)|}{|\vec{p}_T(j_1)|+|\vec{p}_T(j_2)|}$,
- and the azimuthal distance between j_1 and j_2 denoted by $|\phi_{max} - \phi_{min}|$,

The definition of these bins is shown on the Figure 2.

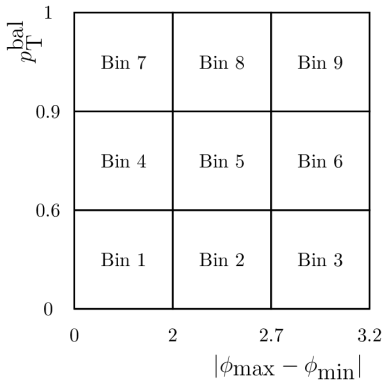

FIGURE 2 – Definition of the 9 ($|\phi_{max} - \phi_{min}|$, p_T^{bal}) bins

In each of these bins, the background contribution is estimated with uncertainty, and exclusion on the signal strength is done thanks to a simultaneous binned maximum-likelihood fit on data. The 95 % confidence level limits on the cross-section are then computed by using CLs method.

3 Validation of the implementation

3.1 Generation of signal events

ATLAS provided cutflows and signal yields in the signal region for 6 representative models, which are : (m_Φ = 1 TeV, R_{inv} = 0.6), (1 TeV, 0.8), (2 TeV, 0.4), (2 TeV, 0.6), (3 TeV, 0.2) and (3 TeV, 0.4).

For each model, around 30 000 events were generated with MG5 aMC v2.6.7 [10] for the hard scattering matrix elements and with Pythia v8.244 [11] for the simulation of the parton shower and hadronization, with the same parameters employed by ATLAS.

For the fast simplified detector simulation, the same b-tagging efficiencies as the ones implemented for the recast of the ATLAS-EXOT-2019-03 analysis [12] were used, because they both use the same DNN at the same working point for the b-jet identification.

Comparison between the ATLAS results and those obtained from the MadAnalysis code were made, on both the cutflow for the signal events and their repartition in the 9 bins of the signal region.

The cross-section upper limits computation was made thanks to the MadAnalysis output interpreter ma5_expert (`https://github.com/MadAnalysis/ma5_expert`) to compare with the limits set by ATLAS.

3.2 Cutflows

Cutflows are presented in the tables in appendix A. They show the number of events after each cut weighted by $L \times \sigma$ (L : luminosity, σ : signal cross-section given $\lambda = 1$) that are provided by ATLAS, and the corresponding ones obtained with MadAnalysis in the "ATLAS events" and "MA5 events" columns respectively.

In the "Relative difference" column, the relative difference between these numbers after each step is computed, and in the "ATLAS cut efficiency" and "MA5 cut efficiency" columns, the fraction of events surviving the considering cut are presented.

The agreement between MadAnalysis and ATLAS is excellent on both the successive and final number of events, and on the cut efficiency (especially for $E_T^{miss} > 600$ GeV).

3.3 Signal region yields

In appendix B, the number of events in the 9 bins of the signal region for ATLAS and MadAnalysis is shown in the "ATLAS events" and "MA5 events" columns respectively.

The "ATLAS proportion" and "MA5 proportion" columns account for the number of events in the bin divided by the total number of events in the signal region.

For the 6 models, the repartition of the events in these 9 bins is highly similar between ATLAS and MadAnalysis.

3.4 Cross-section exclusion limits

In the Figure 3, the expected and observed upper limits on the signal cross-section at 95 % confidence level are represented for 6 different scenarios. Based on the number of events and expected background in the 9 bins provided by ATLAS, the limit calculation was made in each bin separately, and the best value among the 9 ones was finally retained.

To be able to compare to ATLAS in a consistent way, the limits obtained are multiplied by a factor of 1.25, because it is mentioned in the article that the systematic uncertainties (coming mostly from signal modeling) weaken the limits by about 25 %, and the limits computed with MadAnalysis do not take these errors into account.

The limits are globally very well reproduced.

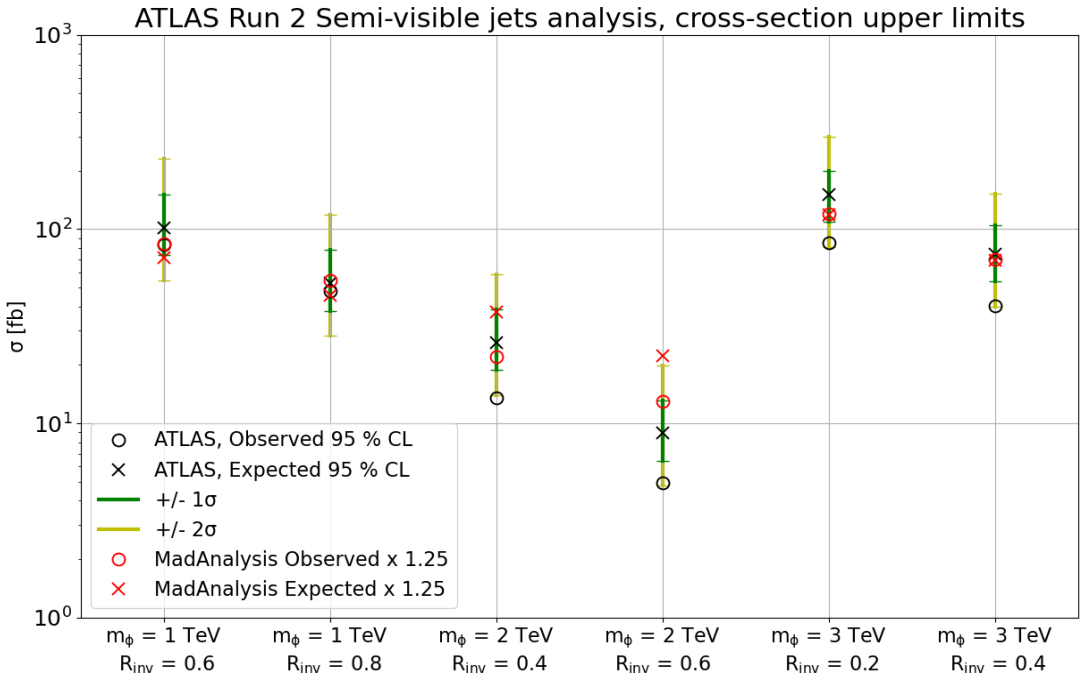

FIGURE 3 – 95% CL upper limits on σ for the 6 illustrative models

4 Conclusion

The ATLAS-EXOT-2022-37 analysis, a search for semi-visible jets via a t-channel production mode, has been implemented in the MadAnalysis 5 framework. The validation results on the signal cutflows and cross-section upper limits show very good agreement with the ATLAS values. The analysis is now available in the MadAnalysis Public Analysis Database.

Références

[1] The ATLAS Collaboration, "Search for non-resonant production of semi-visible jets using Run 2 data in ATLAS", 2024, arXiv :2305.18037, Phys. Lett. B 848 (2024) 138324

[2] Eric Conte and Benjamin Fuks and Guillaume Serret, "MadAnalysis 5, a user-friendly framework for collider phenomenology", 2012, arXiv :1206.1599, Comput. Phys. Commun. 184 (2013) 222-256

[3] Eric Conte, Benjamin Fuks, "Confronting new physics theories to LHC data with MadAnalysis 5", 2018, arXiv :1808.00480, Int.J.Mod.Phys. A33 (2018) 1830027

[4] B. Dumont and B. Fuks and S. Kraml and S. Bein and G. Chalons and E. Conte and S. Kulkarni and D. Sengupta and C. Wymant, "Towards a public analysis database for LHC new physics searches using MadAnalysis 5", 2014, arXiv :1407.3278, Eur. Phys. J. C75 (2015) 56

[5] Jack Y. Araz and Benjamin Fuks and Georgios Polykratis, "Simplified fast detector simulation in MadAnalysis 5", 2020, arXiv :2006.09387, Eur. Phys. J. C81 (2021) 329

[6] Timothy Cohen and Mariangela Lisanti and Hou Keong Lou, "Semi-visible Jets : Dark Matter Undercover at the LHC", 2015, arXiv :1503.00009, Phys. Rev. Lett. 115, 171804 (2015)

[7] Matteo Cacciari and Gavin P. Salam and Gregory Soyez, "The anti-k_t jet clustering algorithm", 2008, arXiv :0802.1189, JHEP 0804 :063,2008

[8] The ATLAS Collaboration, "ATLAS flavour-tagging algorithms for the LHC Run 2 *pp* collision dataset", 2022, arXiv :2211.16345, Eur. Phys. J. C 83 (2023) 681

[9] The ATLAS Collaboration, "Identification of hadronic tau lepton decays using neural network in the ATLAS experiment", ATL-PHYS-PUB-2019-033

[10] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, M. Zaro, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", 2014, arXiv :1405.0301 JHEP 07 (2014) 079

[11] Torbjörn Sjöstrand and Stefan Ask and Jesper R. Christiansen and Richard Corke and Nishita Desai and Philip Ilten and Stephen Mrenna and Stefan Prestel and Christine O. Rasmussen and Peter Z. Skands, "An Introduction to PYTHIA 8.2", 2014, arXiv :1410.3012 Comput.Phys.Commun. 191 (2015) 159-177

[12] Wojtkowski Thomas, 2023, "Implementation of a search for new dijet resonances (139 fb-1 ; ATLAS-EXOT-2019-03)", https ://-doi.org/10.14428/DVN/KHJ1MW, Open Data @ UCLouvain, V1

A Cutflows

1 TeV, 0.6	**ATLAS events**	**MA5 events**	**Relative difference [%]**	**ATLAS cut efficiency [%]**	**MA5 cut efficiency [%]**
Pre-selection	844520.2	802004.396	5.03		
$\Delta\phi < 2.0$	816341.4	773773.493	5.21	96.66	96.48
$p_{T,\ leading\ jet} > 250$ GeV	791042.5	747516.248	5.50	96.90	96.61
$N_{b-jet} < 2$	707151.0	747516.248	5.71	89.39	100
τ jet veto	701537.9	724219.427	3.23	99.21	96.88
$E_T^{miss} > 600$ GeV	101378.1	108504.199	7.03	14.45	14.98
$H_T > 600$ GeV	101235.2	108504.199	7.18	99.86	100

1 TeV, 0.8	**ATLAS events**	**MA5 events**	**Relative difference [%]**	**ATLAS cut efficiency [%]**	**MA5 cut efficiency [%]**
Pre-selection	958068.42	937323.509	2.17		
$\Delta\phi < 2.0$	848394.66	887855.258	4.65	88.55	94.72
$p_{T,\ leading\ jet} > 250$ GeV	812949.3	868462.621	6.83	95.82	97.82
$N_{b-jet} < 2$	760521.81	868462.621	14.19	93.55	100
τ jet veto	754783.75	769611.826	1.96	99.25	88.62
$E_T^{miss} > 600$ GeV	159627.3	164450.938	3.02	21.15	21.37
$H_T > 600$ GeV	159582.1	164236.461	2.92	99.97	99.87

2 TeV, 0.4	**ATLAS events**	**MA5 events**	**Relative difference [%]**	**ATLAS cut efficiency [%]**	**MA5 cut efficiency [%]**
Pre-selection	9743.90	8351.378	14.29		
$\Delta\phi < 2.0$	9674.17	8273.449	14.48	99.28	99.07
$p_{T,\ leading\ jet} > 250$ GeV	9625.17	7460.968	22.48	99.49	90.18
$N_{b-jet} < 2$	7640.34	7460.968	2.35	79.38	100
τ jet veto	7596.01	7413.346	2.40	99.42	99.36
$E_T^{miss} > 600$ GeV	2814.51	2728.232	3.07	37.05	36.80
$H_T > 600$ GeV	2814.51	2728.232	3.07	100	100

2 TeV, 0.6	**ATLAS events**	**MA5 events**	**Relative difference [%]**	**ATLAS cut efficiency [%]**	**MA5 cut efficiency [%]**
Pre-selection	10563.8	10170.440	3.72		
$\Delta\phi < 2.0$	10368.65	10020.355	3.36	98.15	98.52
$p_{T,\ leading\ jet} > 250$ GeV	10284.3	9500.107	7.63	99.19	94.81
$N_{b-jet} < 2$	8693.65	9500.107	9.28	84.53	100
τ jet veto	8633.8	9304.563	7.77	99.31	97.94
$E_T^{miss} > 600$ GeV	4430.74	4866.947	9.85	51.32	52.31
$H_T > 600$ GeV	4430.405	4866.947	9.85	99.99	100

3 TeV, 0.2	**ATLAS events**	**MA5 events**	**Relative difference [%]**	**ATLAS cut efficiency [%]**	**MA5 cut efficiency [%]**
Pre-selection	298.47	246.521	17.41		
$\Delta\phi < 2.0$	293.45	238.030	18.89	98.32	96.56
$p_{T,\ leading\ jet} > 250$ GeV	286.74	213.520	25.54	97.71	89.70
$N_{b-jet} < 2$	218.65	213.520	2.35	76.25	100
τ jet veto	217.98	211.868	2.80	99.69	99.23
$E_T^{miss} > 600$ GeV	59.06	55.491	6.04	27.09	26.19
$H_T > 600$ GeV	59.05	55.491	6.03	99.98	100

3 TeV, 0.4	**ATLAS events**	**MA5 events**	**Relative difference [%]**	**ATLAS cut efficiency [%]**	**MA5 cut efficiency [%]**
Pre-selection	407.31	360.258	11.55		
$\Delta\phi < 2.0$	395.97	343.597	13.23	97.22	95.38
$p_{T,\ leading\ jet} > 250$ GeV	383.73	321.336	16.26	96.91	93.52
$N_{b-jet} < 2$	315.48	321.336	1.86	82.21	100
τ jet veto	313.87	316.655	0.89	99.49	98.54
$E_T^{miss} > 600$ GeV	126.16	127.506	1.07	40.19	40.27
$H_T > 600$ GeV	126.15	127.460	1.04	99.99	99.96

B Signal region yields

1 TeV, 0.6	ATLAS events	MA5 events	ATLAS proportion [%]	MA5 proportion [%]
Bin n°1	3374.36	3389.416	3.33	3.12
Bin n°2	9515.77	10597.286	9.4	9.77
Bin n°3	12811.8	13214.429	12.66	12.18
Bin n°4	15256.2	16346.428	15.07	15.07
Bin n°5	25855.2	27458.553	25.54	25.31
Bin n°6	20496.5	23168.153	20.25	21.35
Bin n°7	4750.29	5448.807	4.69	5.02
Bin n°8	5947.91	5234.287	5.88	4.82
Bin n°9	3226.77	3646.839	3.19	3.36

1 TeV, 0.8	ATLAS events	MA5 events	ATLAS proportion [%]	MA5 proportion [%]
Bin n°1	9643.89	10125.342	6.04	6.17
Bin n°2	12578.9	12957.006	7.88	7.89
Bin n°3	11172.2	10725.998	7	6.53
Bin n°4	35597.8	37712.607	22.31	22.96
Bin n°5	39655.1	43161.416	24.85	26.28
Bin n°6	24541.3	24154.954	15.38	14.71
Bin n°7	13176.9	12442.158	8.26	7.58
Bin n°8	8608.53	8409.182	5.39	5.12
Bin n°9	4607.78	4547.823	2.89	2.77

2 TeV, 0.4	ATLAS events	MA5 events	ATLAS proportion [%]	MA5 proportion [%]
Bin n°1	48.0222	41.129	1.71	1.51
Bin n°2	527.435	449.534	18.74	16.48
Bin n°3	608.745	601.784	21.63	22.06
Bin n°4	137.336	147.199	4.88	5.4
Bin n°5	559.836	575.808	19.89	21.11
Bin n°6	540.037	534.679	19.19	19.6
Bin n°7	69.8095	72.878	2.48	2.67
Bin n°8	189.783	168.124	6.74	6.16
Bin n°9	133.502	137.097	4.74	5.03

2 TeV, 0.6	**ATLAS events**	**MA5 events**	**ATLAS proportion [%]**	**MA5 proportion [%]**
Bin n°1	135.42	155.136	3.06	3.19
Bin n°2	741.915	748.983	16.75	15.39
Bin n°3	658.42	758.363	14.86	15.58
Bin n°4	381.6975	453.142	8.62	9.31
Bin n°5	1040.295	1143.678	23.48	23.5
Bin n°6	732.785	808.873	16.54	16.62
Bin n°7	178.38	188.328	4.03	3.87
Bin n°8	332.352	366.554	7.5	7.53
Bin n°9	229.142	243.889	5.17	5.01

3 TeV, 0.2	**ATLAS events**	**MA5 events**	**ATLAS proportion [%]**	**MA5 proportion [%]**
Bin n°1	0.252929	0.184	0.43	0.33
Bin n°2	7.87065	6.977	13.33	12.57
Bin n°3	20.6843	21.205	35.02	38.21
Bin n°4	1.10346	0.505	1.87	0.91
Bin n°5	6.88926	6.747	11.67	12.16
Bin n°6	14.7363	13.035	24.95	23.49
Bin n°7	0.563986	0.413	0.95	0.74
Bin n°8	2.08462	2.295	3.53	4.14
Bin n°9	4.87367	4.131	8.25	7.44

3 TeV, 0.4	**ATLAS events**	**MA5 events**	**ATLAS proportion [%]**	**MA5 proportion [%]**
Bin n°1	1.98936	2.387	1.58	1.87
Bin n°2	17.3099	17.350	13.72	13.61
Bin n°3	35.0971	36.122	27.82	28.34
Bin n°4	3.12326	3.672	2.48	2.88
Bin n°5	18.9341	19.232	15.01	15.09
Bin n°6	30.5786	29.100	24.24	22.83
Bin n°7	2.35781	2.249	1.87	1.76
Bin n°8	6.20717	7.206	4.92	5.65
Bin n°9	10.5617	10.144	8.37	7.96