MadAnalysis5 validation for the recasting of CMS-SUS-16-033

Federico Ambrogi

August 9, 2018

University of Vienna, Faculty of Physics, Bolzmanngasse 5, A-1090 Wien, Austria

email: federico.ambrogi88@gmail.com

Abstract

This document constitutes the validation note for the `MadAnalysis 5` [1–3] implementation of the analysis CMS-SUS-16-033 [4] *"Search for supersymmetry in multijet events with missing transverse momentum in proton-proton collisions at 13 TeV"* by the CMS Collaboration. The recast code can be found on `INSPIRE` [5].

Contents

1 CMS-SUS-16-033 Analysis Description

The analysis CMS-SUS-16-033 [4] is designed for the search of Supersymmetric (SUSY) particles decaying in the 'all hadronic' final state. The data analyzed, that corresponds to an integrated luminosity of 35.9 fb^{-1}, was collected with proton-proton collisions at a center of mass energy of 13 TeV.

The search is performed using four main kinematic variables: the light and b-tagged jets (i.e. jets originating from the hadronization of a bottom quark), the hadronic transverse energy (H_T) and the missing transverse energy ($\not{H}_T$).

All the information as well as the validation material used in the following is publicly available on the official wiki page of the analysis [1].

Events selection The selection of events requires no isolated leptons (electrons or muons), and the p_T of the lepton determine the isolation radius:

- $\Delta R < 0.2$ if $p_T < 50$ GeV ;
- $\Delta R = 10$ GeV $/p_T$ for $50 \leq p_T \leq 200$ GeV ;
- $\Delta R = 0.05$ if $p_T > 200$ GeV.

The isolation variable I_l is defined as:

$$I_l = \frac{\sum_{<\Delta R} p_T(h,c) + p_T(h,n) + p_T(\gamma)}{p_T(l)}$$

where the sum at the numerator includes the scalar p_T of the charged (h,c) and neutral (h,n) hadrons and the photons. The sum is performed over all the objects included in a radius $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$ around the lepton direction.

Electrons and muons are considered isolated if I_e <0.2 and I_μ <0.1 respectively. The analysis also vetoes isolated tracks, with again isolation requirements depending on the type and momentum of tracks being considered.

Coming to jets, the variable H_T is defined as the scalar sum of the p_T of the signal jets:

$$H_T = \sum_{jets(p_T > 30GeV)} |\vec{p}_T| \tag{1}$$

where only the jets with $p_T > 30\ GeV$ and $|\eta| < 2.4$ are used. The variable $\not{H}_T$ is instead

[1] http://cms-results.web.cern.ch/cms-results/public-results/publications/SUS-16-033/index.html

defined as the jets momenta vectorial sum:

$$\not{H}_T = |\vec{\not{H}}_T| = \left| \sum_{jets(p_T > 30GeV)} \vec{p}_T \right| \quad (2)$$

for jets with $|\eta| < 5.0$.

Additional event selections require an angular separation of the azimuthal angle $\Delta\phi(\vec{\not{H}}_T, \vec{p_T})$ between the jets up to the fourth highest p_T jet: the first two-highest-p_T jets must satisfy $\Delta\phi$ >0.5, while the 3rd and 4th-highest, if present in the event, must satisfy $\Delta\phi$ >0.3.

2 Samples Generation

The results of the analyses are interpreted in the context of simplified models spectra, and detailed information (cutflow tables, kinematic distributions and events yields) are provided for several benchmark points, for gluino and squark pair production. No specific information from the CMS collaboration regarding the samples production is available.

The SUSY signal samples were produced using MadGraph5_aMC@NLO [6, 7] for the hard scattering process, while the hadronization and showering were performed using PYTHIA 6.4 [8]. Samples were generated with up to one additional parton. The MLM merging scheme was chosen, with parameters $(XQcut, Qcut) = (30, 65), (30, 160)$ for the case of squark and gluino production respectively.

After the merging procedure, around 100.000 events were passed to the detector simulations in order to ensure statistical robustness.

The CMS detector simulation was performed using the Delphes 3.4.0 [9] fast simulation tool; the CMS card was tuned according to the b-tagging efficiency fitted formula [2] and b-tagging mis-identification rate as provided[3] by the CMS collaboration for 13 TeV data.

Jets were clustered using FastJet 3.2.1 [10], with the $anti - k_T$ [11] algorithm, with $\Delta R = 0.4$.

[2] https://twiki.cern.ch/twiki/bin/view/CMSPublic/SUSICHEP2016ObjectsEfficiency

[3] https://twiki.cern.ch/twiki/pub/CMSPublic/BTV13TeVICHEP2016/xSig_4_0_AErr_2_CSVv2L_0_0_2_4.pdf

3 MadAnalysis 5 Validation

The following list of simplified models were tested and validated:

- *T2qq*: $\tilde{q}\tilde{q}, \tilde{q} \to q\tilde{\chi}_1^0$ for $(m_{\tilde{q}}, m_{\tilde{\chi}_1^0}) = (1000, 800), (700, 400)$;
- *T2bb*: $\tilde{b}\tilde{b}, \tilde{b} \to b\tilde{\chi}_1^0$ for $(m_{\tilde{b}}, m_{\tilde{\chi}_1^0}) = (650, 1), (500, 300)$;
- *T1qqqq*: $\tilde{g}\tilde{g}, \tilde{g} \to qq\tilde{\chi}_1^0$ for $(m_{\tilde{g}}, m_{\tilde{\chi}_1^0}) = (1400, 100), (2000, 800)$;
- *T1bbbb*: $\tilde{g}\tilde{g}, \tilde{g} \to bb\tilde{\chi}_1^0$ for$(m_{\tilde{g}}, m_{\tilde{\chi}_1^0}) = (1500, 100), (1000, 900)$;

For each simplified model in the list above, the preselection cut flows, the kinematic histograms for the N_{jet} ,n_b , H_T and $\not{H}_T$ variables, and the signal yield for the aggregated signal regions, are validated.

The reference cross section for gluino,squark and sbottom production were taken from the LHC SUSY working group [4], and used to normalize the kinematic distributions and to calculate the yields in the aggregated signal regions.

The kinematic distributions are obtained right after the preselection cuts: all the cuts are applied but the one involving the variable shown in the plot. The histogram data obtained from `MadAnalysis 5` are weighted with the total luminosity, production cross section and divided by the number of generated events. The same normalization applies when evaluating the yields in the aggregated signal regions.

[4] https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SUSYCrossSections

3.1 T2qq Simplified Model

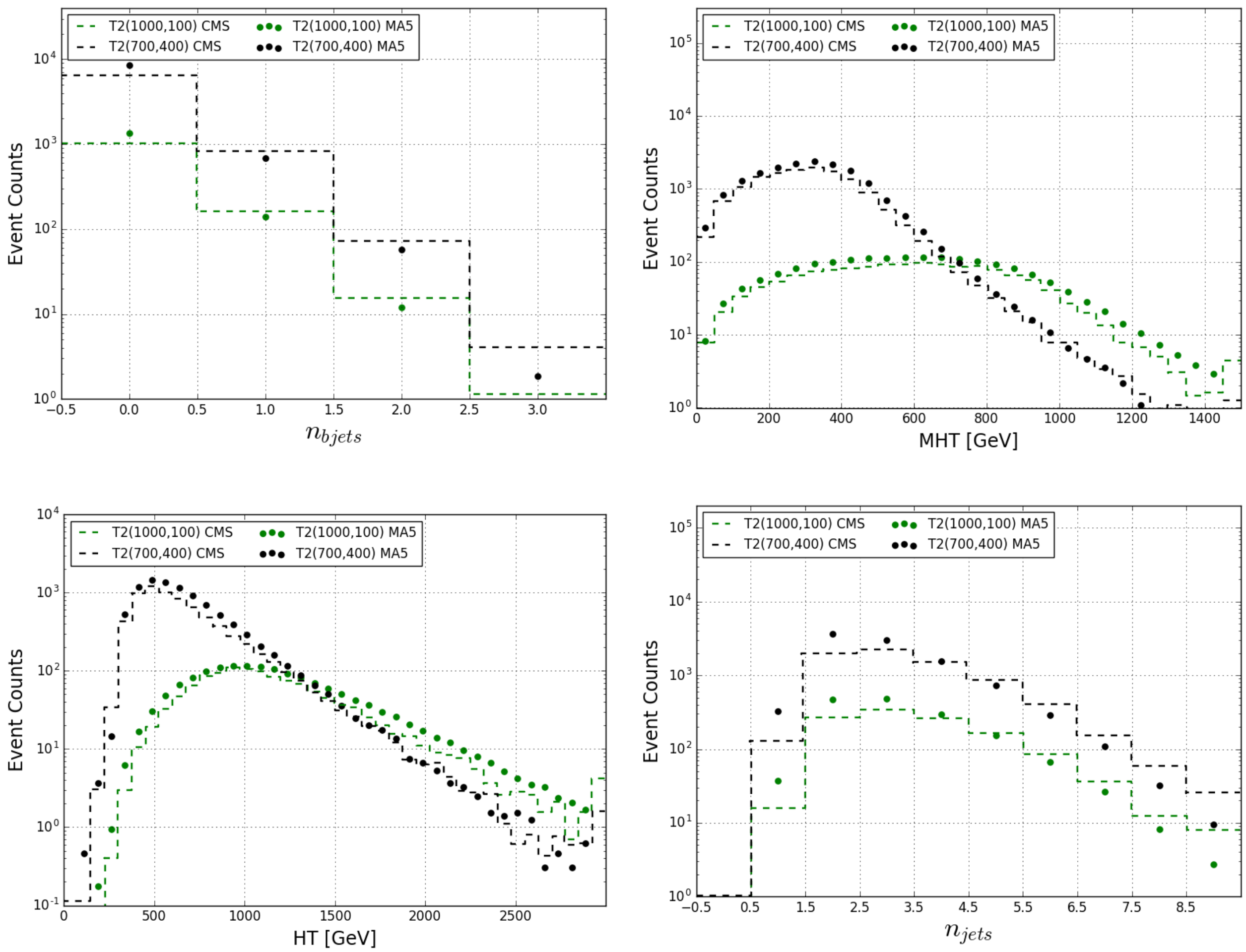

Figure 1: Kinematic distributions for the T2qq simplified models.

	T2qq(700,400)					T2qq(1000,100)				
	Absolute			Drop[%]		Absolute			Drop[%]	
Cut	**MA5**	**CMS**	**Diff(%)**	**MA5**	**CMS**	**MA5**	**CMS**	**Diff(%)**	**MA5**	**CMS**
$N_{jet} \geq 2$	96.18	98.00	1.86	-3.82	-2.00	97.82	98.90	1.09	-2.18	-1.10
H_T >300	87.85	91.30	3.78	-8.66	-6.84	97.12	98.60	1.50	-0.72	-0.30
$\not{H}_T$ >300	43.63	43.80	0.39	-50.34	-52.03	79.16	80.00	1.04	-18.49	-18.86
NoIsoMuons	43.63	43.80	0.39	0.00	0.00	79.16	79.90	0.92	-0.00	-0.12
NoMuonsTracks	43.61	43.70	0.20	-0.04	-0.23	79.14	79.80	0.82	-0.03	-0.13
NoIsoElectrons	43.61	43.50	-0.26	0.00	-0.46	79.14	79.60	0.57	0.00	-0.25
NoElectronsTracks	43.59	43.40	-0.45	-0.05	-0.23	79.12	79.30	0.23	-0.03	-0.38
NoIsoTracks	43.44	43.00	-1.03	-0.35	-0.92	78.96	78.70	-0.33	-0.20	-0.76
$\Delta\phi(\vec{\not{H}}_T, \vec{j_1})$ >0.5	43.41	42.90	-1.19	-0.07	-0.23	78.89	78.60	-0.36	-0.09	-0.13
$\Delta\phi(\vec{\not{H}}_T, \vec{j_2})$ >0.5	40.97	41.10	0.32	-5.62	-4.20	73.58	74.50	1.24	-6.73	-5.22
$\Delta\phi(\vec{\not{H}}_T, \vec{j_3})$ >0.3	39.57	39.60	0.08	-3.41	-3.65	70.24	70.60	0.50	-4.53	-5.23
$\Delta\phi(\vec{\not{H}}_T, \vec{j_4})$ >0.3	38.46	37.90	-1.46	-2.82	-4.29	67.65	67.90	0.36	-3.69	-3.82

Table 1: Pre-selection cutflow for the *T2qq* simplified model.

	T2qq(700,400)		T2qq(1000,100)	
Aggregated Signal Region	**MA5**	**CMS**	**MA5**	**CMS**
SR1-Njet2-Nb0-HT500-MHT500	1620.33	1055.38	988.50	729.30
SR2-Njet3-Nb0-HT1500-MHT750	22.23	25.55	84.27	58.40
SR3-Njet5-Nb0-HT500-MHT-500	279.11	337.33	152.72	179.20
SR4-Njet5-Nb0-HT1500-MHT750	14.51	18.24	29.50	27.98
SR5-Njet9-Nb0-HT1500-MHT750	0.15	1.03	0.68	1.16
SR6-Njet2-Nb2-HT500-MHT500	17.75	16.87	9.41	11.60
SR7-Njet3-Nb1-HT750-MHT750	17.44	21.71	46.10	50.15
SR8-Njet5-Nb3-HT500-MHT500	0.62	0.94	0.49	0.56
SR9-NJet5-Nb2-HT1500-MHT750	0.31	0.78	1.24	1.13
SR10-Njet9-Nb3-HT750-MHT750	0.00	0.05	0.01	0.04
SR11-Njet7-Nb1-HT300-MHT300	39.52	59.26	10.06	16.50
SR12-Njet5-Nb1-HT750-MHT750	10.19	13.41	17.30	21.82

Table 2: Yields in the aggregated signal region for the *T2qq* simplified model.

3.2 T2bb Simplified Model

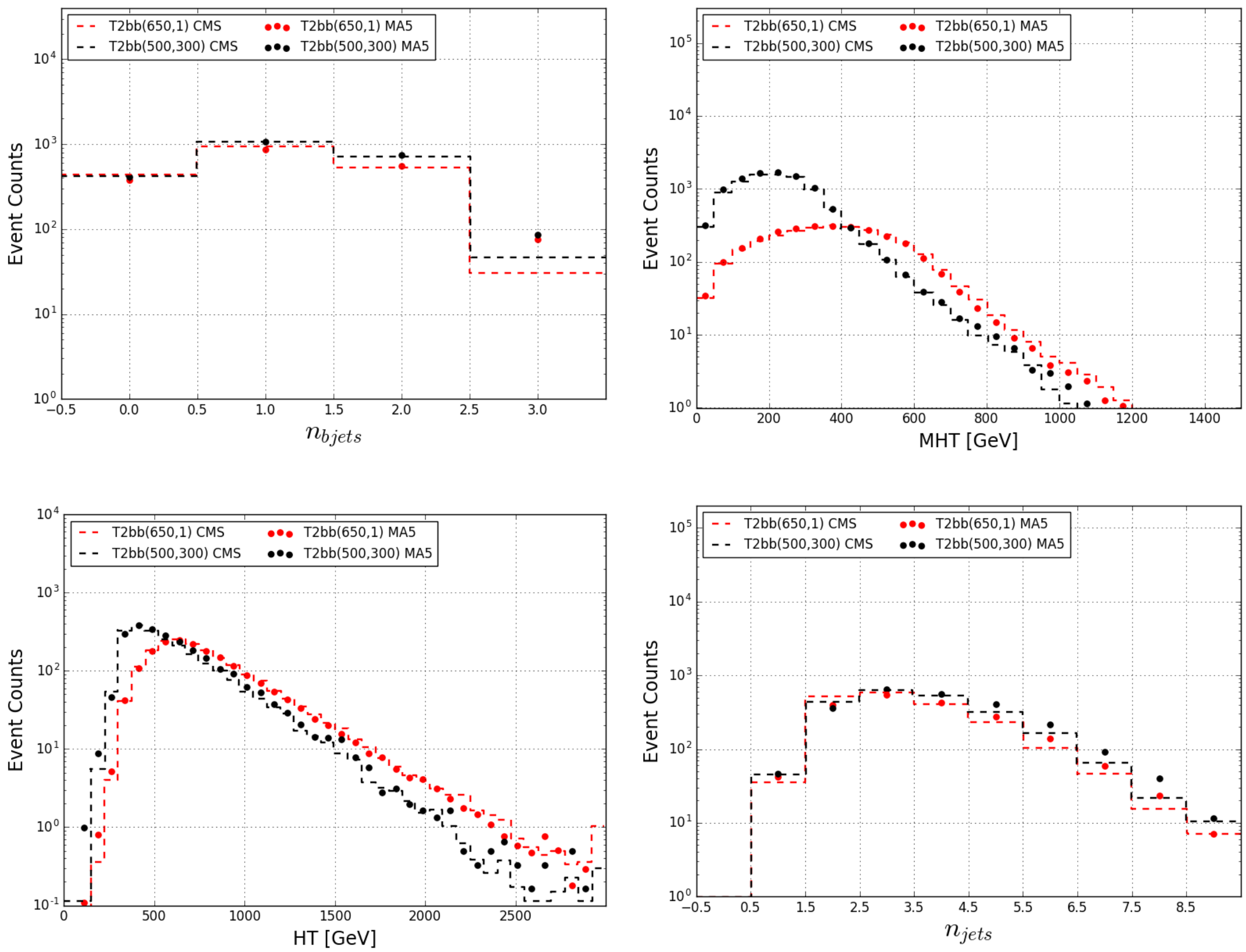

Figure 2: Kinematic distributions for the T2bb simplified models.

	T2bb(500,300)					T2bb(650,1)				
	Absolute			Drop[%]		Absolute			Drop[%]	
Cut	**MA5**	**CMS**	**Diff(%)**	**MA5**	**CMS**	**MA5**	**CMS**	**Diff(%)**	**MA5**	**CMS**
$N_{jet} \geq 2$	95.27	96.00	0.76	-4.73	-4.00	97.36	98.20	0.85	-2.64	-1.80
H_T >300	67.99	68.00	0.02	-28.64	-29.17	94.17	95.40	1.29	-3.27	-2.85
$\not{H}_T$ >300	14.89	15.60	4.53	-78.09	-77.06	56.99	59.80	4.70	-39.49	-37.32
NoIsoMuons	14.89	15.60	4.56	-0.03	0.00	56.98	59.60	4.39	-0.01	-0.33
NoMuonsTracks	14.87	15.50	4.03	-0.10	-0.64	56.94	59.50	4.31	-0.08	-0.17
NoIsoElectrons	14.87	15.40	3.42	-0.01	-0.65	56.94	59.20	3.82	0.00	-0.50
NoElectronsTracks	14.85	15.30	2.95	-0.17	-0.65	56.90	59.00	3.56	-0.06	-0.34
NoIsoTracks	14.79	15.20	2.67	-0.37	-0.65	56.70	58.50	3.07	-0.35	-0.85
$\Delta\phi(\vec{\not{H}}_T, \vec{j}_1)$ >0.5	14.77	15.10	2.21	-0.19	-0.66	56.62	58.40	3.06	-0.15	-0.17
$\Delta\phi(\vec{\not{H}}_T, \vec{j}_2)$ >0.5	13.84	14.10	1.86	-6.29	-6.62	54.15	55.70	2.78	-4.35	-4.62
$\Delta\phi(\vec{\not{H}}_T, \vec{j}_3)$ >0.3	13.34	13.50	1.22	-3.63	-4.26	51.66	53.30	3.07	-4.59	-4.31
$\Delta\phi(\vec{\not{H}}_T, \vec{j}_4)$ >0.3	12.57	13.10	4.07	-5.76	-2.96	49.09	51.60	4.87	-4.99	-3.19

Table 3: Pre-selection cutflow for the *T2bb* simplified model.

	T2bb(500,300)		T2bb(650,1)	
Aggregated Signal Region	**MA5**	**CMS**	**MA5**	**CMS**
SR1-Njet2-Nb0-HT500-MHT500	57.91	50.48	145.59	168.10
SR2-Njet3-Nb0-HT1500-MHT750	1.32	1.46	2.87	3.57
SR3-Njet5-Nb0-HT500-MHT-500	20.29	18.32	32.71	32.40
SR4-Njet5-Nb0-HT1500-MHT750	0.82	0.95	1.82	1.65
SR5-Njet9-Nb0-HT1500-MHT750	0.00	0.10	0.22	0.08
SR6-Njet2-Nb2-HT500-MHT500	98.67	86.76	224.70	202.30
SR7-Njet3-Nb1-HT750-MHT750	29.53	24.11	41.93	47.20
SR8-Njet5-Nb3-HT500-MHT500	10.56	5.54	17.28	6.90
SR9-NJet5-Nb2-HT1500-MHT750	3.63	2.02	3.19	2.30
SR10-Njet9-Nb3-HT750-MHT750	0.00	0.06	0.22	0.08
SR11-Njet7-Nb1-HT300-MHT300	128.86	80.14	79.62	57.00
SR12-Njet5-Nb1-HT750-MHT750	17.16	12.35	19.71	19.57

Table 4: Yields in the aggregated signal regions for the *T2bb* simplified model.

3.3 T1qqqq Simplified Model

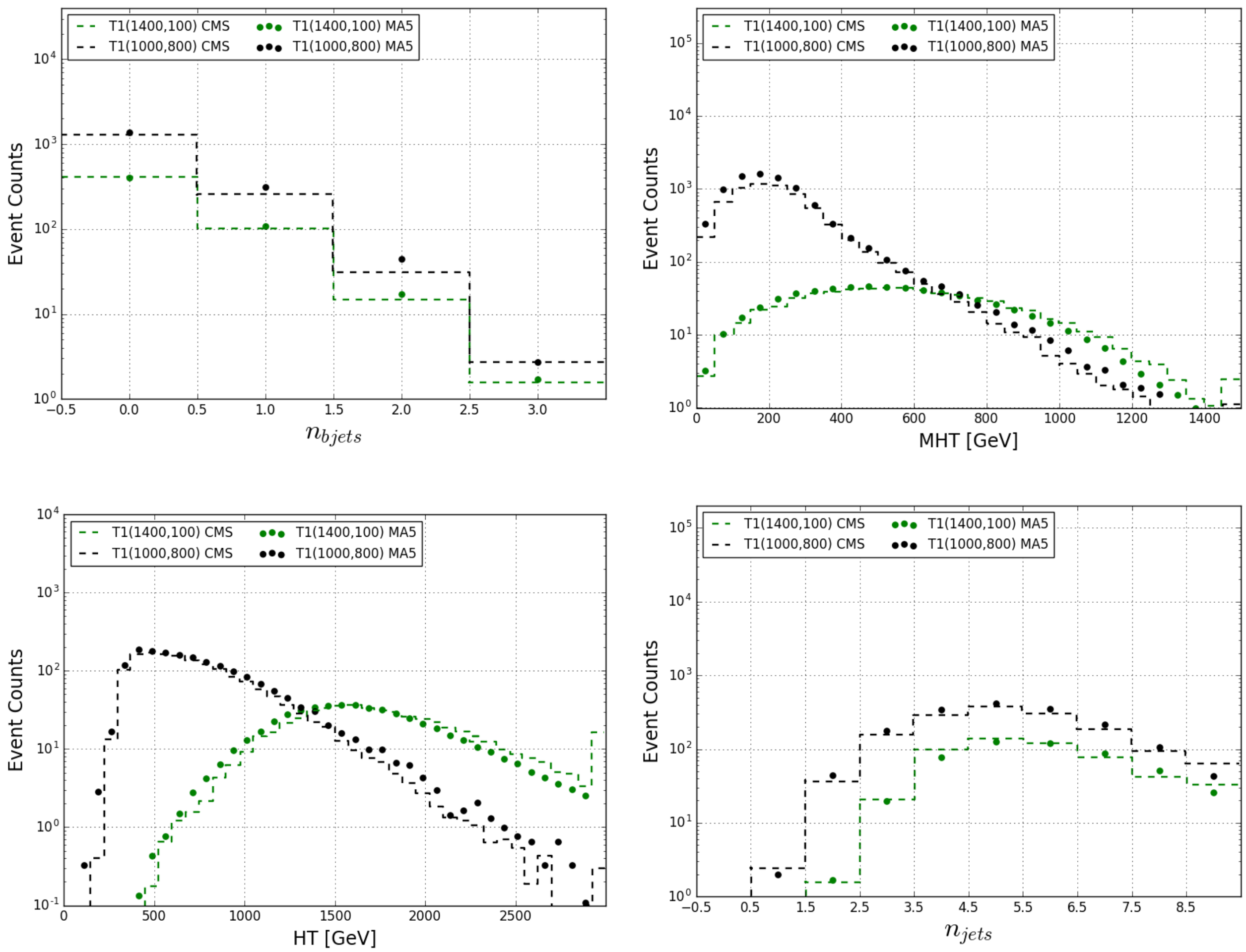

Figure 3: Kinematic distributions for the T1qqqq simplified models.

	T1qqqq(1000,800)					T1qqqq(1400,100)				
	Absolute			Drop[%]		Absolute			Drop[%]	
Cut	**MA5**	**CMS**	**Diff(%)**	**MA5**	**CMS**	**MA5**	**CMS**	**Diff(%)**	**MA5**	**CMS**
$N_{jet} \geq 2$	99.42	99.60	0.18	-0.58	-0.40	99.99	100.00	0.01	-0.01	0.00
H_T >300	74.80	81.30	8.00	-24.77	-18.37	99.98	100.00	0.02	-0.02	0.00
$\not{H}_T$ >300	14.30	19.10	25.11	-80.88	-76.51	77.79	80.00	2.77	-22.20	-20.00
NoIsoMuons	14.30	19.10	25.11	0.00	0.00	77.79	80.00	2.77	0.00	0.00
NoMuonsTracks	14.30	19.10	25.14	-0.04	0.00	77.78	79.90	2.66	-0.01	-0.12
NoIsoElectrons	14.30	19.00	24.75	0.00	-0.52	77.78	79.50	2.17	-0.00	-0.50
NoElectronsTracks	14.29	18.80	23.99	-0.06	-1.05	77.76	79.10	1.69	-0.02	-0.50
NoIsoTracks	14.13	18.40	23.23	-1.15	-2.13	77.41	78.30	1.13	-0.45	-1.01
$\Delta\phi(\vec{\not{H}}_T, \vec{j_1})$ >0.5	14.10	16.90	16.59	-0.21	-8.15	76.16	76.90	0.96	-1.62	-1.79
$\Delta\phi(\vec{\not{H}}_T, \vec{j_2})$ >0.5	12.78	15.80	19.13	-9.36	-6.51	69.28	69.80	0.75	-9.03	-9.23
$\Delta\phi(\vec{\not{H}}_T, \vec{j_3})$ >0.3	11.94	14.80	19.30	-6.53	-6.33	64.30	64.40	0.16	-7.19	-7.74
$\Delta\phi(\vec{\not{H}}_T, \vec{j_4})$ >0.3	10.94	14.60	25.04	-8.36	-1.35	58.43	59.40	1.63	-9.12	-7.76

Table 5: Pre-selection cutflow for the *T1qqqq* simplified model.

	T1qqqq(1000,800)		T1qqqq(1400,100)	
Aggregated Signal Region	**MA5**	**CMS**	**MA5**	**CMS**
Aggregated Signal Region	MA5	CMS	MA5	CMS
SR1-Njet2-Nb0-HT500-MHT500	238.50	272.50	270.35	284.40
SR2-Njet3-Nb0-HT1500-MHT750	20.33	18.79	72.06	88.90
SR3-Njet5-Nb0-HT500-MHT-500	187.56	219.10	215.10	213.30
SR4-Njet5-Nb0-HT1500-MHT750	18.87	17.51	61.38	69.90
SR5-Njet9-Nb0-HT1500-MHT750	3.64	3.06	7.17	5.83
SR6-Njet2-Nb2-HT500-MHT500	9.96	9.42	12.71	11.20
SR7-Njet3-Nb1-HT750-MHT750	18.22	15.89	35.50	37.43
SR8-Njet5-Nb3-HT500-MHT500	0.65	0.81	1.19	1.05
SR9-NJet5-Nb2-HT1500-MHT750	0.81	0.94	3.21	3.49
SR10-Njet9-Nb3-HT750-MHT750	0.00	0.07	0.10	0.12
SR11-Njet7-Nb1-HT300-MHT300	76.77	87.10	52.18	42.97
SR12-Njet5-Nb1-HT750-MHT750	15.47	14.07	29.97	30.67

Table 6: Yield in the aggregated signal region for the *T1qqqq* simplified model.

3.4 T1bbbb Simplified Model

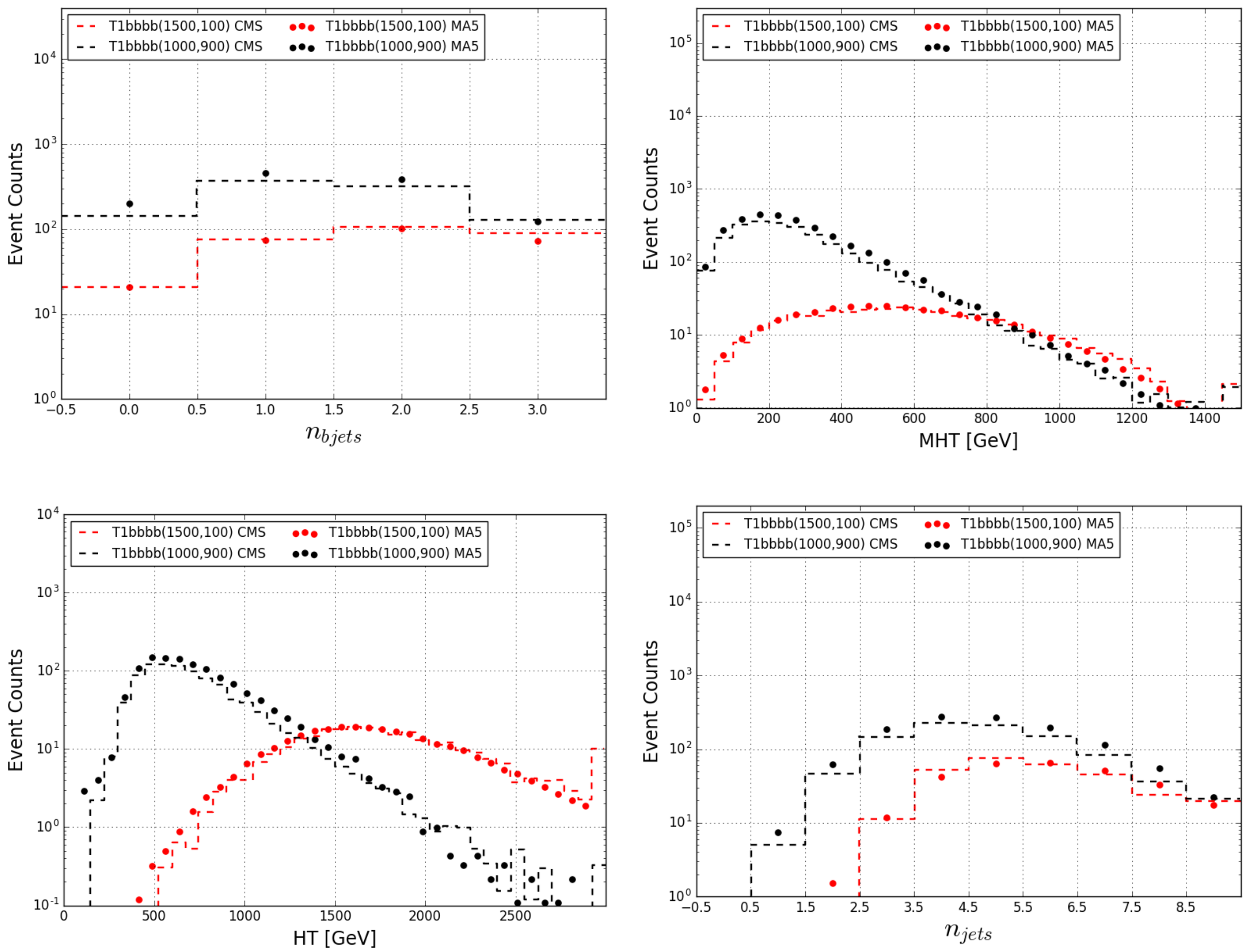

Figure 4: Kinematic distributions for the T1bbbb simplified models.

	T1bbbb(1000,900)					T1bbbb(1500,100)				
	Absolute			Drop[%]		Absolute			Drop[%]	
Cut	**MA5**	**CMS**	**Diff(%)**	**MA5**	**CMS**	**MA5**	**CMS**	**Diff(%)**	**MA5**	**CMS**
$N_{jet} \geq 2$	91.40	92.50	1.19	-8.60	-7.50	99.97	100.00	0.03	-0.03	0.00
H_T >300	39.88	38.60	-3.32	-56.36	-58.27	99.95	100.00	0.05	-0.02	0.00
$\not{H}_T$ >300	14.32	14.10	-1.57	-64.09	-63.47	79.36	80.30	1.17	-20.60	-19.70
NoIsoMuons	14.31	13.90	-2.93	-0.10	-1.42	79.36	79.80	0.55	-0.01	-0.62
NoMuonsTracks	14.06	13.60	-3.35	-1.76	-2.16	79.27	79.60	0.42	-0.12	-0.25
NoIsoElectrons	14.06	13.40	-4.90	0.00	-1.47	79.26	79.20	-0.08	-0.01	-0.50
NoElectronsTracks	13.84	13.10	-5.62	-1.56	-2.24	79.19	78.80	-0.49	-0.10	-0.51
NoIsoTracks	13.75	12.80	-7.40	-0.65	-2.29	78.90	78.00	-1.15	-0.37	-1.02
$\Delta\phi(\vec{\not{H}}_T,\vec{j_1})$ >0.5	13.74	12.80	-7.31	-0.08	0.00	77.49	76.70	-1.03	-1.78	-1.67
$\Delta\phi(\vec{\not{H}}_T,\vec{j_2})$ >0.5	12.19	11.40	-6.94	-11.24	-10.94	70.38	69.20	-1.70	-9.18	-9.78
$\Delta\phi(\vec{\not{H}}_T,\vec{j_3})$ >0.3	11.13	10.40	-7.00	-8.73	-8.77	65.22	63.90	-2.06	-7.33	-7.66
$\Delta\phi(\vec{\not{H}}_T,\vec{j_4})$ >0.3	10.28	9.60	-7.11	-7.59	-7.69	59.15	58.60	-0.94	-9.30	-8.29

Table 7: Pre-selection cutflow for the *T1bbbb* simplified model.

	T1bbbb(1000,900)		T1bbbb(1500,100)	
Aggregated Signal Region	**MA5**	**CMS**	**MA5**	**CMS**
SR1-Njet2-Nb0-HT500-MHT500	66.94	46.90	14.41	14.68
SR2-Njet3-Nb0-HT1500-MHT750	3.05	2.19	4.59	5.15
SR3-Njet5-Nb0-HT500-MHT-500	28.45	18.40	9.76	9.75
SR4-Njet5-Nb0-HT1500-MHT750	2.73	1.49	3.56	3.74
SR5-Njet9-Nb0-HT1500-MHT750	0.65	0.15	0.39	0.28
SR6-Njet2-Nb2-HT500-MHT500	173.45	138.65	141.44	137.51
SR7-Njet3-Nb1-HT750-MHT750	76.31	63.07	89.04	93.83
SR8-Njet5-Nb3-HT500-MHT500	43.83	34.77	62.94	52.78
SR9-NJet5-Nb2-HT1500-MHT750	10.57	9.68	40.87	40.20
SR10-Njet9-Nb3-HT750-MHT750	1.85	1.10	4.95	2.69
SR11-Njet7-Nb1-HT300-MHT300	190.24	133.21	110.48	82.60
SR12-Njet5-Nb1-HT750-MHT750	53.64	43.87	72.73	71.37

Table 8: Aggregated signal region yields for the *T1bbbb* simplified model.

4 Conclusion

We presented the validation note for the analysis CMS-SUS-16-033. Overall the results look in good agreement with the official public data provided by the CMS collaboration, especially for the pre-selection cutflows, which show an agreement at the level of $O(\%)$. Deviations in the signal yields in the aggregated signal regions comparison can be noted especially in the regions where the tails of the distributions are considered. Note however that the largest differences are found for SR where the simulation details, e.g. Monte Carlo parameters and detector simulation efficiencies, are important, for example in the bins of large N_{jet} and N_{bjet} distributions.

References

[1] E. Conte, B. Fuks and G. Serret, Comput. Phys. Commun. **184** (2013) 222 doi:10.1016/j.cpc.2012.09.009 [arXiv:1206.1599 [hep-ph]].

[2] E. Conte, B. Dumont, B. Fuks, and C. Wymant, "Designing and recasting LHC analyses with MadAnalysis 5," *Eur. Phys. J.* **C74** (2014), no. 10, 3103, 1405.3982.

[3] B. Dumont, B. Fuks, S. Kraml, S. Bein, G. Chalons, E. Conte, S. Kulkarni, D. Sengupta, and C. Wymant, "Toward a public analysis database for LHC new physics searches using MADANALYSIS 5," *Eur. Phys. J.* **C75** (2015), no. 2, 56, 1407.3278.

[4] A. M. Sirunyan *et al.* [CMS Collaboration], arXiv:1704.07781 [hep-ex].

[5] F. Ambrogi and J. Sonneveld, doi:10.7484/INSPIREHEP.DATA.77YH.NBR3

[6] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer and T. Stelzer, JHEP **1106** (2011) 128 doi:10.1007/JHEP06(2011)128 [arXiv:1106.0522 [hep-ph]].

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

[8] T. Sjostrand, S. Mrenna and P. Z. Skands, JHEP **0605** (2006) 026 doi:10.1088/1126-6708/2006/05/026 [hep-ph/0603175].

[9] J. de Favereau *et al.* [DELPHES 3 Collaboration], JHEP **1402** (2014) 057 doi:10.1007/JHEP02(2014)057 1307.6346

[10] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual," *Eur.Phys.J.* **C72** (2012) 1896, 1111.6097.

[11] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k(t) jet clustering algorithm," *JHEP* **0804** (2008) 063, 0802.1189.