Annual Progress Monitoring (First Year, June 2017) **First Year Report**

June 01, 2017

Dark Matter searches at CMS at $\sqrt{s} = 13$ TeV

Eshwen Bhal

University of Bristol, United Kingdom

Contact information: **eshwen.bhal@bristol.ac.uk**; **eshwen.bhal@cern.ch**

Supervisors: Henning Flaecher (**Henning.Flacher@cern.ch**); Bjoern Penning (**penning@cern.ch**)

Abstract

This document outlines the progress I have made during the first year of my PhD. The over-arching goal is to search for dark matter in Supersymmetry frameworks, for the time being, and possibly expand into other theories at a later date. In order to build up the necessary skills, I have learned to use many software packages that generate Monte Carlo events and analyse data. I have also completed tasks for the RA1 group at CMS as part of my primary research. These include cut flow tables for the signal models in their 2015 paper, and I am currently analysing Supersymmetry and dark matter models for their 2016 paper. In addition to research, I have performed Jet Energy Corrections in the Level-1 Trigger as part of my service work to CMS. In the future, I aim to become a more integral part of RA1 and CMS. This includes the continuation of work on Monte Carlo signal and analysing the full Run-2 dataset at $\sqrt{s} = 13$ TeV generated by the LHC.

1 Introduction

Dark matter is a non-baryonic form of matter that permeates much of the observable universe. Thought to be created in the hot, early universe when spontaneous pair production of particles was plentiful, dark matter was produced in abundance along with other heavy exotic particles. When the universe cooled, a thermal freeze out occurred: the average temperature became too low to allow significant pair production [1]. As the cosmos expanded and matter grew further apart, the dark matter annihilation rate decreased and what was left is known as a ”thermal relic”. These remaining particles were attracted via gravity, the only known force by which dark matter interacts. They formed filaments throughout the universe, and the potential wells they generated allowed the progenitors of galaxies to form within. Recent results from the Planck mission estimate dark matter to account for 25.8% of the universe's energy density [2].

Although dark matter cannot be directly studied with conventional astronomy, there are several independent astrophysical observations that suggest its existence. The rotation curves of most galaxies are roughly flat [3]. This is in contradiction to the Keplerian curve ($v \propto 1/\sqrt{r}$) that would be exhibited if only baryonic matter were present. On the galactic scale, dark matter is sprinkled in a roughly spherical halo that spans beyond the disc. Whilst its surface density decreases with radius, the inclusive dark matter mass increases linearly [4] to compensate for the Keplerian decrease from baryonic matter [5][6]. Weak gravitational lensing is another phenomenon that supports the existence of dark matter. Images of galaxies can appear distorted due to dark matter warping the spacetime along the line of sight from the galaxy to the observer [7].

From these observations, several properties of dark matter can be inferred. It is electrically neutral, because if it could couple to the electromagnetic field it would be detectable by conventional astronomy. It is ”cold” (non-relativistic), which would imply that it is heavy. Current estimates suggest its mass is at the GeV or TeV scale. If dark matter were light, like neutrinos, and relativistic, it would be too diffuse to condense and allow galaxy formation. This supports the idea of ”bottom-up” structure formation in the universe – smaller galaxies form around dark matter clumps, then merge to form larger structures [8]. This also asserts that dark matter is stable, at least on the timescale of the universe's current age, to allow structure formation around it.

1.1 Dark matter searches

Whilst all evidence has been astrophysical in nature, discovering the properties of dark matter falls into the realm of particle physics. There are three types of dark matter searches: direct detection (dark matter recoiling from baryonic matter); indirect detection (dark matter annihilating into baryonic matter); and production (from high energy collisions). The latter of these methods is being probed using CERN's Large Hadron Collider (LHC). Protons are collided at energies sufficient to produce the heavy particles that existed in the high-temperature early universe. The Compact Muon Solenoid (CMS) experiment utilises its general purpose detector to allow physicists to search for dark matter in different theoretical frameworks.

Despite the Standard Model of Particle Physics (SM) providing precise predictions of three of the four fundamental forces and the particles that they interact with, no dark matter models exist within it. Several theories that are beyond, or extend, the Standard Model can accommodate dark matter candidates such as sterile neutrinos [9], axions [10], and Kaluza-Klein states [11].

For the moment, I am searching for dark matter in the context of Supersymmetry (SUSY) [12]. This is currently the leading candidate for physics beyond the Standard Model for its numerous solutions to SM shortcomings. The theory introduces a spin symmetry that predicts a fermionic superpartner for each boson, and vice versa. Colloquially, these are known as ”sparticles”. If the lightest sparticle (LSP) is stable and electrically neutral, it would provide a promising dark matter candidate. The conservation of R-parity [12] prevents it from decaying into SM particles even if they are lighter, and forces unstable sparticles to decay into lighter sparticles and regular matter. This means the decay cascades of SUSY particles produced in a collider would end with the LSP and many Standard Model particles. These are normally energetic hadrons, which then create showers of new hadrons from pair production that are detected as ”jets”. As LSPs are undetectable, a reconstructed event from a detector would show a momentum imbalance: SM particles would appear to recoil from nothing, as they would likely travel in one direction perpendicular to the beam, and nothing would be detected travelling in the opposite direction. These events contain ”missing” transverse energy (MET, $E_{\mathrm{T}}^{\mathrm{miss}}$) which are required to satisfy energy and momentum conservation. The predominant sources of MET in accelerator events are from neutrinos, misreconstructed particles, and possibly new particles. So the characteristics of SUSY in a collider would be high MET from the LSPs, and several hadronic jets.

There is significant motivation to study dark matter from a wider, as well as a more personal, viewpoint. It is important to understand how the universe operates, and dark matter opens up the potential for new physics that improves our understanding of nature. My personal interests include the blend of particle physics and astrophysics, and the opportunity to discover and add to humanity’s collective wisdom. With a projected 150 fb^{-1} from the LHC’s Run-2 at a centre-of-mass energy $\sqrt{s} = 13$ TeV, there is great potential to constrain some of the properties of dark matter.

1.2 Postgraduate courses

In order to obtain a comprehensive understanding of particle physics, I have undertaken several postgraduate courses to supplement the analysis-based work that forms the bulk of my PhD. I have attended modules such as Quantum Field Theory, Introduction to Particle Physics, Statistics for Particle Physicists, and others from the MPAGS programme. I have also taken part in the CMS Induction Course at CERN that focused on the experiment itself. More recently, I attended the Warwick Week school that covered particle physics from a more experimental standpoint to supplement the upcoming theory-heavy STFC summer school.

2 Analysis and research

The bulk of my research as a PhD student is focused on data generation and analysis. The first weeks of the course involved learning various pieces of software that are integral to this. Then, as I became more experienced with these tools and integrated within the RA1 (Reference Analysis 1) group at CMS, I earned more responsibility. My first task with this group was to create cut flow tables for their 2015 paper. The second, and current, assignment is now to analyse the signal models that are the focus of their 2016 paper.

2.1 Initial work

There are many tools and software that are required in high energy physics research, and I spent the first months learning these as well as getting to grips with other aspects of the PhD. ROOT (see `https://root.cern.ch/`) is the primary tool used to analyse LHC data and generate plots. ROOT uses a data container referred to as a ″tree″ or ″ntuple″ that stores the data from collider events. These are then categorised into ″branches″ and ″leaves″ that store event objects and variables, respectively. ROOT often visualises data as histograms, as seen in most high energy physics papers. I have also spent time learning programs that can generate Monte Carlo events, such as MadGraph [13]. An accompanying program MadAnalysis [14] can analyse other data formats, such as LHE files. As most of these programs operate using the command line interface (CLI) or scripting, I have gained skills in both areas. ROOT analyses generally use C++ macros or CLI interactivity. Programs like these usually generate large amounts of data, which necessitates the use of remote servers to store it all. This also allows other group members to access it and perform different analyses in parallel.

I began learning many of these skills and programs in the context of generating backgrounds for dark matter searches. The most common processes are $pp \to t\bar{t}$, $pp \to W$ + jets, and $pp \to Z$ + jets. MadGraph was used to generate the Monte Carlo for these processes. Then I could write C++ macros to analyse them in ROOT, or a shell script to analyse them in MadAnalysis, all done on the Bristol Particle Physics group's remote server, Soolin. An example histogram is displayed below. This shows the number of weighted events against $E_{\mathrm{T}}^{\mathrm{miss}}$ for different H_{T} (the scalar sum of momentum in GeV) regimes for $pp \to W$ + jets and Z + jets.

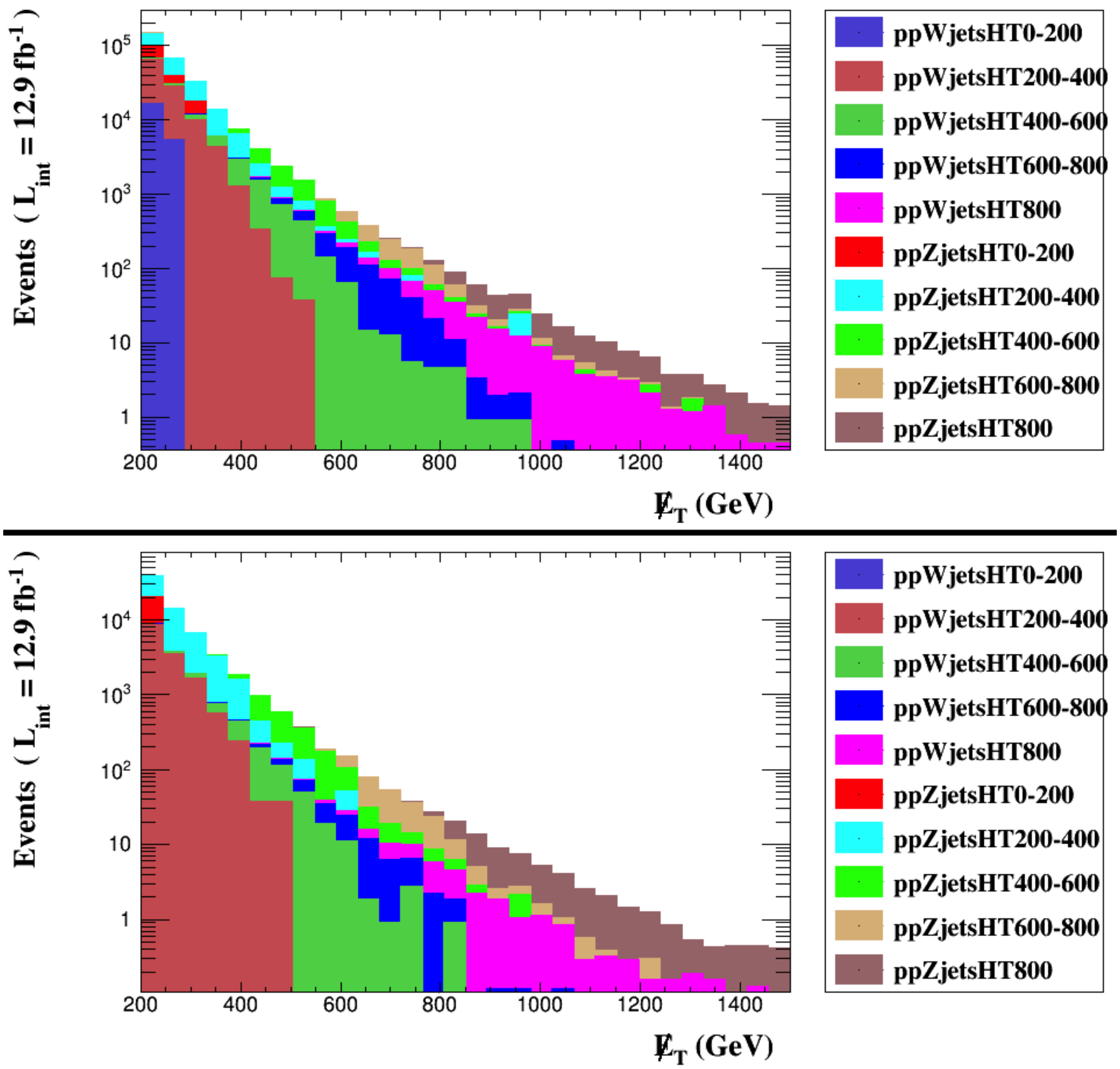

Figure 1: The number of weighted events vs. $E_{\mathrm{T}}^{\mathrm{miss}}$ for the Standard Model background processes $pp \to W$ + jets and Z + jets. These were generated using Monte Carlo random sampling with different ranges of H_{T}. The upper panel shows the events before any cuts, with the lower panel showing the effects of several cuts: jet $p_{\mathrm{T}} > 100$ GeV, $E_{\mathrm{T}}^{\mathrm{miss}}$ > 200 GeV, $|\eta| < 2.5$, no b-quarks, and no leptons.

2.2 Cut flow tables for SUS-15-005

My first significant task for the RA1 group was making cut flow tables for the signal models analysed in their 2015 paper (SUS-15-005, [15]). The motivation for these is the fact that cut flow tables are an important part of analyses based on searches. They are useful for understanding how the signal behaves when certain cuts are applied. These can then be updated in the future to implement different selections (or invent new variables like α_{T} [16]) that maximise the signal and minimise the background after the full cut flow has been applied. One can also see whether there are bugs/anomalies in the signal if the efficiencies seem out of place. In terms of personal motivation, this task helped me get accustomed to the software the RA1 group uses for tree production and analysis.

The signal datasets I analysed are known as SMS – or Simplified Model Spectra – models. These contain Monte Carlo events whose decay products from pp collisions are a heavy SUSY particle that decays with a 100% branching fraction into, typically, light squarks, the LSP and SM particles. The first port of call was to produce trees of these SMS models where no cuts are applied. Loose cuts are usually applied during production, but can bias the values in the cut flow tables so were removed. I used the CMGTOOLS software (the tree production code) that also required HEPPY (a Python framework for high energy physics), all within a CMSSW (CMS SoftWare) environment. Once these trees were made, I used a variant of ALPHATWIRL – Python code which converts ROOT trees into Pandas dataframes, making them easier to analyse – to find the cumulative efficiencies of the models after the cuts were applied. I wrote a Python dictionary that contained the cuts in the form of strings corresponding to leaves in the trees. By skimming over the SMS trees I found the efficiencies of each model. Then I passed the data to the group coordinator to include them in the auxiliary material for the paper, which can be found at `http://cms-results.web.cern.ch/cms-results/public-results/publications/SUS-15-005/index.html`. The group thought my contribution was significant enough to add my name to the authors list of the internal paper.

Event selection	Benchmark model (m_{SUSY}, m_{LSP})					
	T1qqqq (1300, 100)	T1qqqq (900, 700)	T2qq_8fold (1050, 100)	T2qq_8fold (650, 550)	T2qq_1fold (600, 50)	T2qq_1fold (400, 250)
Before selection	100	100	100	100	100	100
Event veto for muons and electrons	99	100	100	100	100	100
Event veto for single isolated tracks	94	91	96	95	96	95
Event veto for photons	92	90	95	94	95	95
Event veto for forward jets ($\|\eta\| > 3.0$)	81	78	82	81	80	80
$n_{\mathrm{jet}} \geq 2$	81	78	81	72	80	75
$p_{\mathrm{T}}^{\mathrm{j}_1} > 100\,\mathrm{GeV}$	81	71	81	57	79	66
$\|\eta^{\mathrm{j}_1}\| < 2.5$	81	70	81	55	79	65
$H_{\mathrm{T}} > 200\,\mathrm{GeV}$	81	69	81	50	79	60
$H_{\mathrm{T}}^{\mathrm{miss}} > 130\,\mathrm{GeV}$	77	50	78	33	71	40
$H_{\mathrm{T}}^{\mathrm{miss}}/E_{\mathrm{T}}^{\mathrm{miss}} < 1.25$	74	44	75	28	65	33
H_{T}-dependent α_{T} requirements ($H_{\mathrm{T}} < 800\,\mathrm{GeV}$)	74	30	71	15	50	17
$\Delta\phi^{*}_{\mathrm{min}} > 0.5$	22	18	44	10	33	13
Four most sensitive n_{jet} event categories	22	13	43	5.5	31	6.1

Table 1: The cumulative signal acceptance times efficiency for various benchmark models from the RA1 2015 analysis. The variable $\Delta\phi^{*}_{\mathrm{min}}$ is is based on the minimum azimuthal separation between a jet and the negative vector $\vec{p}_{\mathrm{T}}$ sum of all of the other jets in the event. The names of the benchmark models refer to the parent SUSY particle and its decay products.

2.3 Current work: Signal model analysis for SUS-16-016

My second, and current task for the RA1 group is to analyse the signal models for their 2016 paper (SUS-16-016). This involves finding the systematic uncertainties and efficiencies for all of the SUSY SMS and heavy flavour dark matter models. I first produced the trees from miniAODs using CMGTOOLS. All of the SUSY models are from the Spring16 dataset, and the scalar and pseudoscalar DM models are from Summer16. The next step is to calculate the systematic uncertainties using our analysis code, ALPHATOOLS. First, I will produce histograms of the yields under all systematic variations (b-tag scale factors, jet energy corrections, etc.). It is these variations that are taken as the systematic uncertainties. These are then fed to our statistical framework ALPHASTATS to calculate the signal acceptance times efficiency – similar to the overall cut flow efficiencies in the previous subsection – for different mass combinations of the parent sparticle and LSP for a given model. These are usually presented as surface plots in the auxiliary material for a paper.

3 Service work: Jet Energy Corrections in the Level-1 Trigger

In addition to analysis, researchers that are a part of CMS must undertake service work, or Experimental Physics Responsibility. These are normally maintenance-based tasks related to the trigger or infrastructure that keep CMS running smoothly. My service work is performing Jet Energy Corrections (JECs) in Layer-2 of the calorimeter in the Level-1 (L1) Trigger.

JECs are necessary to compensate for various losses when recording jet properties in the trigger. These losses depend on the transverse momentum p_{T} and pseudorapidity η – an angle relative to the beam axis – and so JECs ensure that the performance of the trigger is uniform across the detector. Firstly, some ideal (or reference) jets are needed to compare against given L1 jets. For Monte Carlo these are called "GenJets" and for data are "PF jets". Then L1 jets need to be matched against reference jets. From there, various studies can be performed such measuring the response of the detector, and its position and energy resolutions.

In terms of the workflow, all software is used within the CMSSW framework. There are specific L1 Trigger packages and integration tags that must be included and compiled. Once set up, the first task is to create Monte Carlo ntuples of L1-specific jets *without* any corrections applied. These are prepared using various Python scripts and then sent to CERN's computing grid via CRAB (CMS Remote Analysis Builder). Then the L1 jets are matched with the reference jets using the variable

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \tag{1}$$

where ϕ is the azimuthal angle of the jet around the detector. These are expressed as differences because the angular components are discretised in the detector, and resolution can be poor within each trigger crystal/tower. The algorithm used to match the jets does so by inspecting each L1 jet in descending p_{T} and searching for reference jets with $\Delta R < 0.25$. If there is more than one match, the reference jet with the smallest ΔR is taken. Then the next L1 jet follows the same procedure, with the previous reference jet being removed from the collection of possible matches to avoid double counting. This is then repeated for all L1 jets. After matching, calibrations are derived. Correction graphs as a function of p_T are plotted for each $|\eta|$ bin. Then, Gaussian correction curves are fitted to them. If they look good, closure tests are then conducted as the final step. The L1 ntuples are remade with JECs and then matched with the reference jets to check that the calibrations have been properly applied. Plots are then passed to the Trigger Studies Group to check over and then continue with the long chain of trigger corrections and calibrations. Currently, I have completed two rounds of JECs, each taking up to a few weeks to complete due to setbacks or complications.

4 Future research

Short term goals for research include finishing the 2016 signal model analysis and continuing with JECs when needed. In the intermediate future, I will be attending the STFC summer school. This is a theory-based course consisting of postgraduate-level particle physics lectures and assignments. Shortly after, I expect to move to France/Switzerland for my long term posting at CERN of between 12 and 24 months. During this time, I will largely be undertaking analysis and research work similar to this year. It is assumed that I will also carry on with JECs and possibly begin maintenance shifts on the Trigger to contribute to my service work quota. In the longer term, both during and after my time spent at CERN, I will be analysing the Run-2 data from the LHC as it becomes available. Dark matter searches will be performed, perhaps in a SUSY framework, and may also be extended to include other theories, as mentioned in the introduction. The completion of these aforementioned goals will hopefully lead to a comprehensive experimental PhD and a contribution to the field of particle physics in pursuit of understanding dark matter.

References

[1] Iason Baldes and Kalliopi Petraki. "Asymmetric thermal-relic dark matter: Sommerfeld-enhanced freeze-out, annihilation signals and unitarity bounds". In: (2017). arXiv: 1703.00478 [hep-ph].

[2] Planck Collaboration et al. "Planck 2015 results. XIII. Cosmological parameters". In: *Astronomy & Astrophysics* 594, A13 (Sept. 2016), A13. DOI: 10.1051/0004-6361/201525830. arXiv: 1502.01589.

[3] M. Persic, P. Salucci, and F. Stel. "The universal rotation curve of spiral galaxies - I. The dark matter connection". In: *Monthly Notices of the Royal Astronomical Society* 281 (July 1996), pp. 27–47. DOI: 10.1093/mnras/281.1.27. eprint: astro-ph/9506004.

[4] J. Einasto. "Dark Matter". In: *ArXiv e-prints* (Jan. 2009). arXiv: 0901.0632 [astro-ph.CO].

[5] K. C. Freeman. "On the Disks of Spiral and S0 Galaxies". In: *The Astrophysical Journal* 160 (June 1970), p. 811. DOI: 10.1086/150474.

[6] A. H. Broeils. "The mass distribution of the dwarf spiral NGC 1560". In: *Astronomy & Astrophysics* 256 (Mar. 1992), pp. 19–32.

[7] D. Huterer. "Weak lensing, dark matter and dark energy". In: *General Relativity and Gravitation* 42 (Sept. 2010), pp. 2177–2195. DOI: 10.1007/s10714-010-1051-z. arXiv: 1001.1758 [astro-ph.CO].

[8] S. D. M. White and M. J. Rees. "Core condensation in heavy halos: a two-stage theory for galaxy formation and clustering". In: *Monthly Notices of the Royal Astronomical Society* 183.3 (1978), p. 341. DOI: 10.1093/mnras/183.3.341.

[9] Marco Drewes. "The Phenomenology of Right Handed Neutrinos". In: *International Journal of Modern Physics E* 22.08 (2013), p. 1330019. DOI: 10.1142/S0218301313300191.

[10] M. Dine, W. Fischler, and M. Srednicki. "A simple solution to the strong CP problem with a harmless axion". In: *Physics Letters B* 104 (Aug. 1981), pp. 199–202. DOI: 10.1016/0370-2693(81)90590-6.

[11] Tao Han, Joseph D. Lykken, and Ren-Jie Zhang. "On Kaluza-Klein states from large extra dimensions". In: *Phys. Rev.* D59 (1999), p. 105006. DOI: 10.1103/PhysRevD.59.105006. arXiv: hep-ph/9811350 [hep-ph].

[12] Stephen P. Martin. "A Supersymmetry primer". In: (1997). [Adv. Ser. Direct. High Energy Phys.18,1(1998)]. DOI: 10.1142/9789812839657_0001, 10.1142/9789814307505_0001. arXiv: hep-ph/9709356 [hep-ph].

[13] Johan Alwall et al. "MadGraph 5 : Going Beyond". In: *JHEP* 06 (2011), p. 128. DOI: 10.1007/JHEP06(2011)128. arXiv: 1106.0522 [hep-ph].

[14] Eric Conte, Benjamin Fuks, and Guillaume Serret. "MadAnalysis 5, A User-Friendly Framework for Collider Phenomenology". In: *Comput. Phys. Commun.* 184 (2013), pp. 222–256. DOI: 10.1016/j.cpc.2012.09.009. arXiv: 1206.1599 [hep-ph].

[15] CMS Collaboration. "A search for new phenomena in pp collisions at $\sqrt{s} = 13$ TeV in final states with missing transverse momentum and at least one jet using the α_{T} variable". In: *Submitted to: Eur. Phys. J. C* (2016). CMS-SUS-15-005, CERN-EP-2016-246. arXiv: 1611.00338 [hep-ex].

[16] Lisa Randall and David Tucker-Smith. "Dijet Searches for Supersymmetry at the LHC". In: *Phys. Rev. Lett.* 101 (2008), p. 221803. DOI: 10.1103/PhysRevLett.101.221803. arXiv: 0806.1049 [hep-ph].