Christina Nelson
University of Hawaii Manoa
canelson@hawaii.edu

May 14, 2015

Abstract

This study investigates the search for a new heavy resonance decaying into two Higgs bosons with a pair of $b\bar{b}$ quarks in the final state, with focus on highly boosted Higgs. This analysis is done with Monte Carlo (MC) simulations from Data Challenge (DC) 14 at $\sqrt{s} = 13$ TeV with a benchmark model of the Kaluza-Klein graviton decaying to hh. Combining techniques of jet substructure and flavour tagging will be the method applied to aid in maximizing sensitivity while decreasing background. In order to study dominance of background particle decay a $\sqrt{s} = 13$ TeV DC14 MC sample of $t\bar{t}$ events is used to provide pre-recommendations for run 2 of the Large Hadron Collider. It is found that pre-recommendations for jet tracking systematics do not exceed $5\%, 10\%, 20\%$ respectively for b-jets, c-jets, and light-jets.

1 Introduction

The Large Hadron Collider (LHC) at CERN is at the forefront in pushing our understanding of the fundamental nature of matter. The discovery of the Higgs boson has opened up opportunities for searches of new physics at the TeV and electro-weak scale. Out of the four experiments at the LHC, one of the two largest is the ATLAS (A Toroidal LHC Apparatus) experiment. Many new physics models predict TeV-scale resonances decaying to pairs of electroweak-scale bosons.[2]. The decay resonances produced at the TeV scale yield highly boosted final states such that jets from b hadrons become collimated resulting in jet algorithms failing to define the separte b hadron jets.

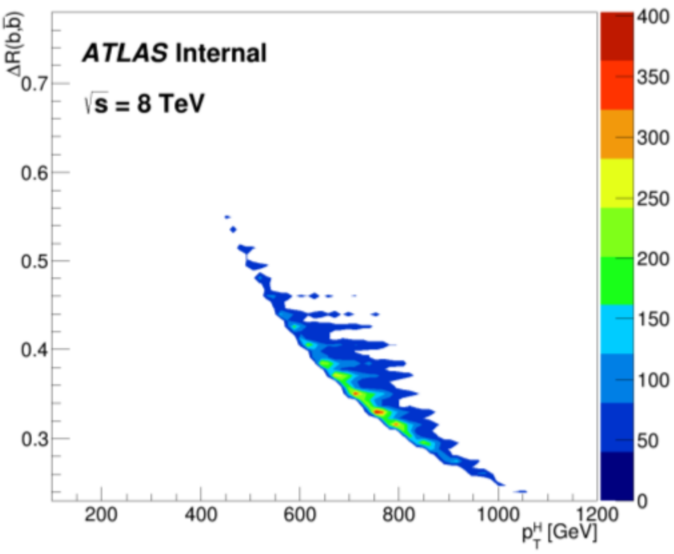

Figure 1: delta R between two b hadrons as a funciton of the transverse momentum, p_T, of the Higgs boson

2 Motivation

A possible solution to the Plank-weak hierarchy problem of the Standard Model (SM) is the proposed solution based on the Randall-Sundrum framework with a warped extra dimension. A distinct description of this model is the existence of a spin-2 Kaluza-Klein (KK) gravitons whose masses and couplings to the SM are set by the TeV scale. The KK graviton would be the first excitation in the fifth dimension that we are able to experimentally detect.[2]. In general it is taken that the angular separation of the 2-body decay products of a heavy particle is approximately

$$\Delta R \approx \frac{2m}{p_T} \tag{1}$$

where $\Delta R = \sqrt{(\Delta y)^2 + (\Delta\phi)^2}$ between the two decay products, p_T is the transverse momentum and m is the mass. [3] It should be noted that directly observable quantities form the ATLAS detector are y which is known as rapidity, but often times taken as equivalent to η which is knwon as pseudorapidity, and the measure of the angle ϕ [11]. Figure 1 shows a kinematic plot of delta R between two b hadrons as a function of the transverse momentum of the Higgs. One can easily see that for a Higgs of mass 125 GeV, the point where the standard 0.4 jet finding algorithm will begin to fail to find separate jets starts at a transverse momentum of approximately 625 GeV; as momenta increases so the jets become more collimated.

A grip that can be found in these dense jet and highly boosted environments is in $b-$tagging methods. If a b quark is produced, the hadronization process will create a jet of hadrons, one of which will contain the b quark, known as b hadrons. The b quarks are relatively long lived with life times of the order 1.5×10^{-12}s. Due to their lifetime, the b hadrons travel a few millimeters before decaying; $b-$tagging depends on resolving this decaying point, or secondary vertex, with the primary vertex produced by the hard scattering interaciton of the particles. Also, the branching ratio of a 125 GeV Higgs to two b quarks is approximately 58%, which means that it happens that 58% of the time the Higgs decays to two b quarks. Compare this to the branching ratio of a 125 GeV Higgs to two Ws which is approximately 21.6%.[4].

3 The Detector and Data Aquisition

The ATLAS detector consists of four major components. The inner detector (ID) measures the momentum of charged particles and consists of a silicon pixel detector, a silicon microstrip detector (SCT), and a straw-tube transition radiation tracker (TRT). The electromagnetic calorimeter (EM) measures the energies carried by the particles and is divided into a central barrel and end-cap regions on either end of the detector. The muon spectrometer identifies and measures the momenta of muons, and the magnet system bends charged particles for momentum measurement. The interactions in the ATLAS detectors create more data than is possible for current technology to handle. To handle this enormous data, a trigger system where events are selected by 100 interesting events per second out of 1000 million other events. Then data is distributed across the world stored on tape or disk for analysers to process, this is known as the grid. [11].

For run 2 there has been hardware upgrades to the detector in order to cope with bringing the Luminosity up to $5 \times 10^{34}\ cm^{-2}s^{-1}$. The ID has been completely rebuilt for run 2, and consists of an all-new all-silicon tracker. Also upgrades have been made to the EM, pixel system, barrel and end-cap strips. The muon detector, muon and electron triggers have also been modified to include tracking regions of interest and improve muon resolution. [7]. In addition to hardware, there has been a migration of software framework for analysers from DPDs (Derived Physics Dataset) to xAODs (x-Analysis Object Data). When data comes directly out of the detector it is called RAW data and from there is reconstruced into an AOD (analysis object data) or ESD (Event Data Model). The framework that reconstructs the raw data is known as Athena and is an ATLAS online and offline software. From the reconstructed AOD or ESD file, a merge between AODs and DPDs into a common analysis format known as xAOD is almost completely implemented across ATLAS software now. The advantage and reason for migration to

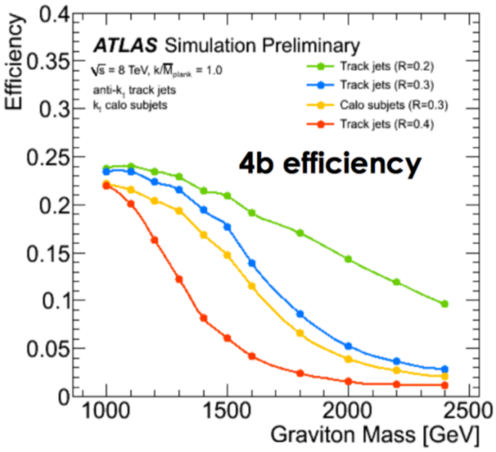

Figure 2: Signal efficiency after selection for individual Randal-Sundrum graviton (RSG) mass points using different track jet radius (R) parameters that are associated to a large radius (large-R jet of 1.0)

xAOD format is that it is readable in Athena and ROOT (an object oriented framework for large scale data analysis). Also xAODs are fast and compact, uniform across all reconstructed object types (ie. election, muon, jets, etc.), and is more object oriented. [8]. All of the work presented here done with 13 TeV monte carlo samples are analyzed using xAOD in either Athena or the framework known as RootCore.

4 Analysis Strategy

A novel strategy is followed that the boosted diHiggs to 4b analysis employed in run1 which combines jetsubstructure algorithms and b-tagging methods; in fact this group was the first in ATLAS to use such an approach in an analysis. The method is to use large-R trimmed jets in order to capture the fully boosted Higgs boson decay products, and require two b-tagged small-R track jets to be associated to the large-R jet. The benefits of using track jets are many. The b-tagging algorithms use pattern recognition from the inner detector information, with track jets the b-tagging training and calibration can be decoupled form the calorimeter, and optimized separately to interpret the hadronic final state. Track jets may be chosen directly from the primary vertex, thus allowing b-tagging performance to not be deteriorated due to pileup. This allows for smaller-R jets to be used which greatly enhances signal efficiency in the boosted regions. Track jets also have a good angular resolution with respect to b-hadrons, even in dense environments, which allows for greater accuracy in determination of the primary vertex leading to better resolution and efficiency at high boosted values.[3] [5]. Figure 2 shows the gain in signal efficiency of a resonant mass of the Randall-Sundrum graviton (RSG) for different radius parameters. Around 1 TeV there is observed a decrease in efficiency with track jets of radius 0.3

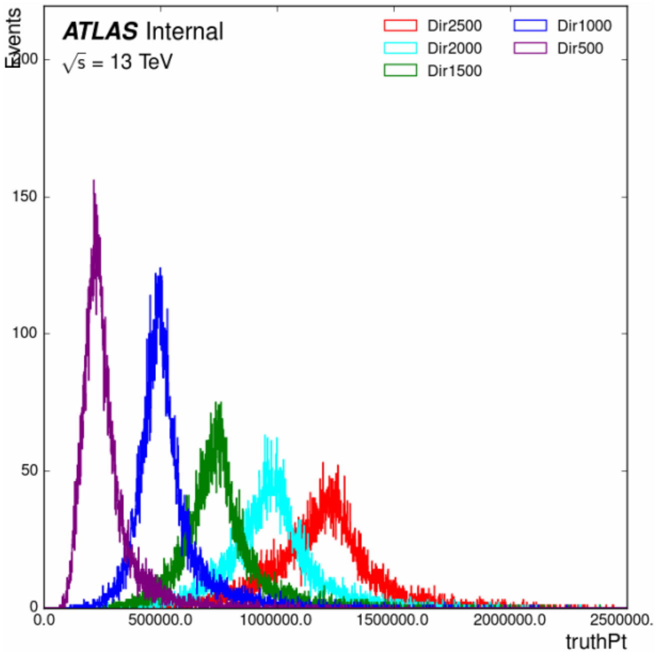

Figure 3: spectrum of truth jets from monte carlo sample of RSG at different mass resonances

to 0.2. This analysis currently looks at a small-R of 0.3 track jets because it was what was done in the run 1 analysis, but is in the process of looking into R =0.2 track jets as well.

5 Software Framework and Object selection

The Monte Carlo samples used are generated wtih Pythia and Powheg, and consist of five samples of the RSG decaying to a pair of Higgs bosons with a $b\bar{b}b\bar{b}$ final state, with mass resonances of 500GeV, 1000GeV, 1500GeV, 2000GeV, and 2500GeV. Figure 3 shows a plot of the resonat mass distrubution of the RSG signals with truth particle information, which is direct information from the event generator. To investigate backgrounds in flavour tagging, a $t\bar{t}$ 13 TeV sample is used specifically to observe b-tagging with Monte Carlo samples.

I wrote a software package that uses RootCore framework and implements helper tools contributed by the Chicago group who also work on the diHiggs to 4b analysis. The package algorithm follows the analysis strategy outlined above. Calorimeter jets are defined by using the anti-k_t algorithm with a radius parameter of 1.0. The antik_t algorithm is an iterative algorithm that reclusters energy deposits read from the calorimetr into what we call jets based on a distance and angular parameter.[6].Then track jets are reconstructed from the ID track with R =0.3 and required to have a $p_T > 0.5$GeV. Only track jets with at least two track jets that are b-tagged with a $p_T > 7$GeV and $\eta < 2.5$ are considered for this analysis. Track jets are then so-called ghost associated to the large-R calorimeter jet. What ghost association does is to instead of being confined to the cone jet-algorithms of anti-k_t findings, a ghost associated jet with infinitely small p_T is linked to the calorimeter jet and track jet. Then a test is done and the jet ultimately kept

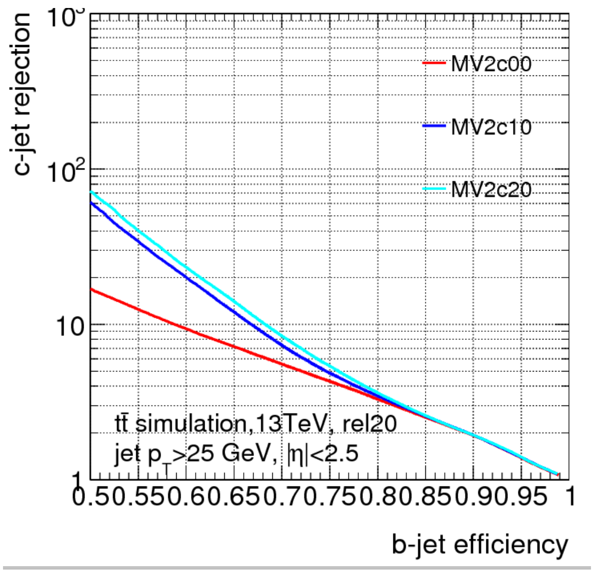

Figure 4: c-jet efficiency after b-tagging

for analysis is the linked jet that either falls within the defined jet-algorithm cone, or has the highest momentum. This allows for greater functionality in keeping highly boosted particles that otherwise may have been cut off due to a fixed radius parater. Finally, the jet flavour tagging is spatially matched to the jet. First if a weakly decaying b-hadron is found within $\Delta R < 0.3$ of the jet direction and jet axis, the closest jet is labeled a b jet. If no b hadron is found, the algorithm is repeated and looks for a c-hadron. If no c-hadrons are found then the process is repeated and labeled a light-flavour jet.[3][5].

6 Observations of simulations and Results

In order to provide pre-recommendations for early run 2 analyses, a background sample of ttbar monte carlo 13 TeV events is used to run b-tagging algorithms over and study their efficiency. This is done for b-jets, c-jets, and *light*-jets to study b-tagging efficiency. The framework used here has been provided by the combined performance (CP) and flavour tagging groups and implements a frame work using the inner detector geometry to study tracking systematics and efficiency. Figure 4 shows a reciever operating characteristic (ROC) curves produced after re-running b-tagging algorithms on this signal for c-jets. Results of the ROC curves produced for b-jets. c-jets, and *light*-jets are found to be within pre-recommendation systematics of 5%, 20%, and 10% respectively, as provided by the CP group.

7 Conclusion

While building the Boosted analysis package it was found that in the MC sample the b-tagger was not switched on, so re-derivations of the sample needed to be made. Next steps are to implement b tagging in the boosted package and provide an efficiency study for various signal regions. Since the framework is already built, this should be straightforward to implement.

Overall, results, although be them with monte carlo and very preliminary yield no great deviations from expectation. One cannot base an entire analysis and study on simulations, yet the simulations will be essential for comaparison with data when 13 TeV collisions begin in a couple of weeks. Indeed no one knows exactly what will happen at higher luminosity, and how the jet finding algorithms and b-tagging algorithms will fare in higer p_T and dense hadronic environment, all the more exciting it is to see what happens. To be able to work for five months at the forefront of international physics in search for a deeper understanding of reality and the world we are part of has been a truly wonderful experience that I will value always. I've learned what it is to write software with different groups and huge collaborations, which has even further developed my appreciation for software development. As my joy in physics grows, and full luminosity collisions are right around the corner it is truly a special time in the world of physics!

Acknowledgement

Special and warm acknowledgements to my mentors Stephane Willocq and Massimmiliano Bellomo. And thank you to the Loundsbery Foundation, University of Michigan, Professor Jean Kirsch, Professor Homer Neal, Professor Tom Schwartz, and Steven Goldfarb.

References

[1] Flavor Tagging with Track Jets in Boosted Topologies with the ATLAS Detector `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/FlavourTagging`

[2] A Large Mass Hierarchy from a Small Extra Dimension. Lisa Randall, Raman Sundrum. `http://arxiv.org/pdf/hep-ph/9905221v1.pdf`

[3] A search for resonant Higgs-pair production in the $b\bar{b}b\bar{b}$ final state in pp collisions at $\sqrt{s} = 8$ TeV `https://cds.cern.ch/record/1666518/files/ATLAS-CONF-2014-005.pdf`

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

[5] Flavor Tagging with Track Jets in Boosted Topologies with the ATLAS Detector `http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2014-013/ATL-PHYS-PUB-2014-013.pdf`

[6] AntiK-t jet clustering algorithm https://inspirehep.net/record/779080/plots?ln=en

[7] Plans for the Phase II upgrade to the ATLAS detector http://iopscience.iop.org/1748-0221/9/02/C02033/pdf/1748-0221_9_02_C02033.pdf

[8] ATLAS Analysis Model: How analysis is done https://indico.cern.ch/event/345626/session/7/contribution/24/material/slides/0.pdf

[9] Data Preparation and Physics Analysis https://indico.cern.ch/event/345626/session/8/contribution/26/material/slides/0.pdf

[10] Boosted W-Boson Identification at $\sqrt{s} = 8$TeV with the ATLAS Detector. Chris Malena Delitzch. http://cds.cern.ch/record/2015208/files/ATL-PHYS-PROC-2015-023.pdf

[11] Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC http://arxiv.org/pdf/1207.7214v2.pdf

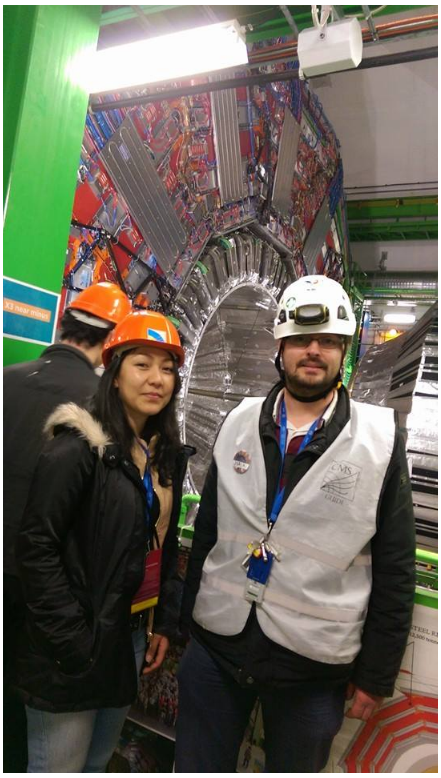

Figure 5: our guide and me in front of the CMS detector