Studies of new light resonances decaying into two hadronic jets produced in association with a radiated jet

MASTER'S THESIS

Zhiying Li

Supervisor

Caterina DOGLIONI

LUND UNIVERSITY

Department of Physics

Division of Particle Physics

LUND UNIVERSITY

May 2018

Abstract

This report presents combinatorics studies in the search for a resonance of below 800 GeV which decay product is a pair of hadronic jets from quarks. The resonance, predicted as a dark matter mediator, has couplings to dark matter particles as well as quarks and gluons. The resonance is created in association with a radiated object that can be either a photon or a jet. In the case of the radiation being a jet, the dijet from the resonance and the radiated jet cannot be distinguished. Choosing the wrong dijet to reconstruct the resonance decreases the sensitivity of the search. The study focuses on studies that help improve the search sensitivity.

The signal samples used in this thesis are generated with Monte Carlo simulation and passed through the ATLAS detector simulation. Firstly, in an attempt to identify the jet from radiation, samples where a jet is radiated and a photon is radiated are compared. However, due to the differences in the sample generations, this attempt does not lead to a successful identification of the jet from radiation. Kinematic variable distributions of radiated jet signal samples at various resonance masses (250 - 1050 GeV) are studied in detail, leading to a better understanding of the event topology.

As a next step, different variables are used to order the jets and choose two of them for the resonance reconstruction. The performances of these reconstruction methods are studied and presented. The correlations between the different ordering methods are also discussed. The most promising methods are chosen to be taken forward to a study of signal significance including the background.

The results in this thesis indicate that the second and third leading jets ordered by p_T form the best performing pair for the dijet in the mass range of interest (below 800 GeV) in this search. Jets that are close to the direction where most of the jet p_T are aligned can also be acceptable candidates for dijet choices.

Key words: dark matter mediator, ATLAS, dijet+ISR jet, light resonance, ISR jet, dijet, trijet final state, kinematics variables, combinatorics, mass combinations, significance, thrust, LHC, CERN

Declaration

This project report is the result of my own work, except where explicit reference is made to the work of others, and has not been submitted for another qualification to this or any other university.

Zhiying Li

Acknowledgements

A great many people have contributed to the completion of this thesis. I would want to thank everyone for everything, from the excellent background work to all the wonderful memories within this year.

Caterina, my main supervisor, is also the person introduced me to this subject of dijet + ISR combinatorics search. I can not thank you enough for all your effort and time into this thesis. Thank you for always being there for me, from the physics discussions, coding problems to my long hunt for Ph. D. positions and later fundings. This thesis and everything along with it in this year would not have happened without you. You are the best and thank you so much.

Kate, I would not forget all the insights into my tiny progresses on this subject. Your feedback has kept me on the right course along the way. And a special thank to Antonio for the wonderful suggestions. Thank you, Adam, Valentina, Alexandra and everyone in the resolved dijet+ISR group for the brilliant ideas and discussions we had.

My fellow Master students, Simon, Dim and Yosse. It has been a great time among great friends, thank you for hearing me whining and giving me encouragements, thank you for the tiny language class and all the jokes we shared. This time will be so different without you guys.

And all my short or long term fellow students and colleagues, thank you so much for making this corridor home, one I will always remember fondly. Our cute fika time and all the lovely cakes!

My family and my friends, thank you for always being there for me, hearing me out and offering help in life. It has been a difficult time and I do not know how I would be able to manage without your support.

Preface

This report describes my research on combinatorics studies on the search for a dark matter mediator predicted as a resonance produced in association with radiations. My thesis work is done at the Particle Physics Division of the Lund University under the supervision of Caterina Doglioni. I was also member of the resolved dijet + ISR analysis group at the ATLAS collaboration at CERN. The studies in this report are conducted to contribute to the analysis group. The plots and related code in this report (including the appendix), except for those specified, are made by myself using the ROOT package developed by experts at CERN. My other contributions to the analysis group are recorded by reports at weekly ATLAS internal meetings.

The samples used in this report are produced with the ATLAS simulation framework and are generated by the rest of the analysis group.

Contents

Chapter 1.

Introduction

For thousands of years, human beings have been wondering about what the most basic components of matter are. Particle Physics is the study of those components, in both theory and experiment. Achievements in the filed have led to the discovery of the leptons, the neutrinos, the quarks and the bosons, as fundamental components of the matter. The theory describing those particles and their interactions is called the Standard Model (SM), described briefly in subsection 1.1.1. It is a very effective model that can make precise predictions that are verified in experiments, but it also has its problems. For example it does not explain the existence of 95% of the matter-energy content in the universe, called *dark matter* (DM) and *dark energy*, described in subsection 1.1.2. To explain this and other unexplained phenomena, physicists hope to discover particles beyond those present in the Standard Model.

The Large Hadron Collider (LHC) [1], located at CERN, Geneva, is the largest particle collider in the world. Many new physics searches and SM measurements are carried on at LHC experiments. The main purpose of this report is to help improve the performance of a search for particles that mediate the interaction between SM particles and DM particles conducted with the ATLAS detector [2], one of the detectors at the LHC.

This report aims to investigate possible improvements in the sensitivity of one of the dark matter searches in ATLAS, where a dark matter mediator decays into quarks. The focus lies on understanding the kinematics of the decay products of these mediator particles and the associated processes within the event, in order to correctly reconstruct the resonance from what is recorded in the detector (chapter 3 to chapter 6).

1.1. Theory

1.1.1. Standard Model

The Standard Model (SM) is a model developed by particle physicists to describe the elementary particles and their interactions. [3] [4] This model started being built in its current formulation in the 1970s and has been developing ever since. The Standard Model includes three families of fermions, and each one has an antiparticle. The fermions have the spin of half-integer. Leptons and quarks are the components of matter. The SM describes the electromagnetic interaction, the weak interaction and the strong interaction among these particles. These interactions are mediated by gauge bosons, with integer spin. The electromagnetic interaction is mediated by photons, the weak interaction by W and Z bosons and the strong interaction by gluons.

Each fermion has an antiparticle of the same mass but opposite charges with respect to its particle. Quark and lepton families have masses that increase from the first family to the third family.

Among the fermions, the lepton families are: electron (ν_e,e), muon (ν_μ,μ) and tau (ν_τ,τ). There are also three quark families, (u,d) (c,s) and (t,b). For each lepton family, the νs are the neutrinos, which have the electric charge of zero. The other leptons, electrons, muons and taus have the electric charge of -e. Charged leptons are subjects to the electromagnetic interaction and weak interaction, but not to the strong interaction. Neutrinos can only interact weakly because they are neutral in electromagnetic charge.

In the quark families, the u, c and t quarks have electric charge of $+\frac{2}{3}$e while the d, s and b have electric charge of -$\frac{1}{3}$e. Quarks also have color charge from the strong interaction, and interact weakly as well. Quarks have color, weak and and electromagnetic charges and can interact strongly. The quarks are confined inside the baryons and mesons due to the strong interaction. Baryons are made of three quarks and mesons of two quarks. There are many baryons and mesons that have been discovered or created by colliders. The most common baryons are the protons and neutrons, which are the basic parts of the nucleus. [5]

The quark and neutrino families also present a phenomenon called mixing, that allows particles in one family to decay into particles in another family. This can be explained by the fact that the weak interaction or flavour eigenstates are not the mass eigenstates. Therefore, the family mixing can happen when they interact. Quark mixing is described with the Cabibbo-Kobayashi-Maskawa (CKM) matrix and the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix describes neutrino mixing. [6] Neutrino mixing is also the evidence for neutrino mass, since if there were no neutrino mass states there would be no mixing. [7]

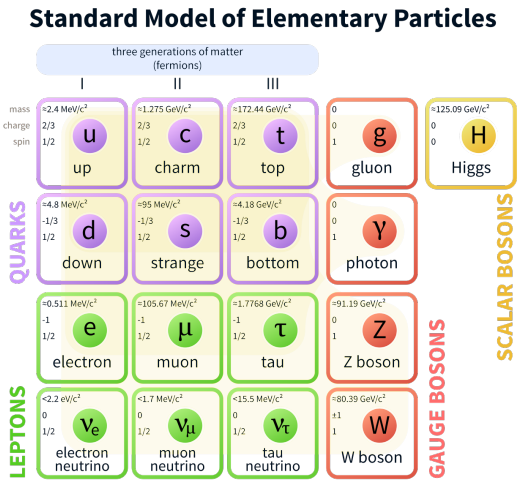


Figure 1.1.: Particles and their properties of the Standard Model. Plot source: Ref. [8] .

The particle compositions and their properties of the SM can be summarised in this plot from the Particle Data Group (the PDG) [8] as in Figure 1.1.

The γ is the mediator of the electromagnetic interaction. It is massless, not self-interacting and does not carry any charges, electric or color.

W and Z^0 are mediators of weak interactions. The charge for weak interaction is the weak isospin, which is carried by left-handed particles. W can couple with both left and right-handed particles while Z^0 does not couple with right-handed particles. As for the weakly interacting particles, it is worth noting that neutrinos can only interact weakly and right-handed neutrinos have not been discovered yet. The weak interaction and EM interaction combined have a U(1) $\times$ SU(2) symmetry, and together they are named electroweak interaction.

The strong interaction exchanges gluons with color charges, so gluons can interact with themselves. The color charge is therefore a property of the strong interaction, and there are three different color charges. Both quarks and gluons carry color charges. The color charge is named after real-world colors, red, green and blue. When quarks form hadrons, the hadrons are color neutral. The theory describing this interactions is quantum chromodynamics (QCD). According to QCD, the quarks inside hadrons are in the form of valence quarks together with a sea of quarks and gluons. When hadrons break, the quarks cannot exist on their own since they are color confined. They need to form other hadrons. The strong interaction has the SU(3) symmetry for its three color charges. Therefore in total, the SM has the symmetry of SU(3) $\times$ SU(2) $\times$ U(1).

There is also the Higgs boson, with the spin of 0. The Higgs boson has mass and it gives many fundamental particles mass through the Higgs mechanism. [9] The strength of the Higgs interaction with another particle is proportional to the particle's mass.

As a theory, the SM also has its limitations. For example, the explanation for neutrino masses is not included in the SM. The SM does not account for gravity, which is also a fundamental interaction. Moreover, astronomical and cosmological observations, (for example, the accelerating expansion of the universe), can not be explained only with particles within the SM. So physicists are now very interested in exploring the possibilities beyond it. There are many famous BSM (Beyond Standard Model) theories, e.g. supersymmetry (SUSY). We will use simplified theories postulating the existence of Dark Matter to motivate LHC searches in this report.

1.1.2. Dark Matter and Dark Energy

There are several observations from astronomy and cosmology supporting the idea that there exists some unknown form of mass-energy in our universe, which have been named as the dark matter and dark energy. These include the accelerating expansion of the universe and the gravitational lensing effect. The gravitational lensing effect means that light is bent by some clumps of mass, and this can be experimentally observed. It was noticed that the visible mass does not account for all the lensing effects caused by mass [10]. There are also arguments that the visible mass itself cannot provide gravity strong enough to form the clusters and galaxies. [11]. Another proof of the existence comes from galaxy scatterings inside a cluster [11]. If physicists are to maintain their current theory of gravity, they need to introduce the existence of new phenomena and new particles to the current understandings of the universe. Dark matter and dark energy are unaccounted for in the SM yet and take up about 95 % of the universe energy-mass composition. According to theories, Dark Matter (DM) takes up 27 % and Dark Energy (DE) takes up 68 %. Due to their large proportions in the universe, the DM and DE are attractive research subjects in Particle Physics.

There are many theories about the particle nature of DM and DE with respect to ordinary matter. The WIMP (weakly interacting massive particles) and the axions are two of the most popular DM candidates. WIMPs can interact with standard matter via gravity and other interactions that are as weak as the weak interaction, if not weaker. The WIMP is a DM candidate predicted by many theories, including SUSY.

Nowadays many experimental searches are dedicated to searching for the WIMP particles. The main search methods are the direct and indirect detection of the DM as well as collider creations of the DM. For direct detection, it is expected that the DM particles can interact with the ordinary matter we use to build detectors even if such interactions are rather weak. The DM particles should make it till underground, and if all other background radiations are removed, detectors can collect signals from DM particles that interact with nucleons in the detector and produce a recoil. In indirect detection experiments, on the

other hand, we would expect the DM particles to interact with each other and produce normal matter particles as interaction products. The detectors would then seek evidence of interactions beyond what is known in the universe. Another approach to studying DM is with colliders. If DM particles interact with SM particles, they can be created with high-energy colliders and their interaction products can be detected with the collider detectors. The collider and experiment used for DM searches in this report will be discussed in more detail in chapter 2. Even though there are many detectors and projects analysing data from all these methods, the DM particle signals remain undiscovered so far.

Chapter 2.

Experimental and Theoretical Tools

As is discussed in chapter 1, many SM and BSM searches are conducted to broaden humans' understanding towards the universe. The main focus of this work is on one of the BSM searches for DM via the LHC. In this chapter, experimental and theoretical tools for this search are introduced, including the LHC collider, and the ATLAS detector used for the search.

Located at CERN in Geneva, the Large Hadron Collider (LHC) is the largest hadron collider currently in the world. Living up to the "large" in its name, the LHC ring is 27 km in length and 100 m deep underground, and it has four major detectors distributed around the ring where collisions occur. For the searches that are discussed in this report, the main focus is on the ATLAS detector, which is a large general-purpose detector.

2.1. The Large Hadron Collider

The CERN accelerator complex [1] including the LHC, schematically drawn in Figure 2.1, is a complex system that can collide two different kinds of beams: protons and heavy ions. In this work, we will only cover proton-proton collisions. The center-of-mass energy ($\sqrt{s}$) of the colliding protons has been rising with each upgrades, to the current $\sqrt{s}$ of 13 TeV. [12]

The protons are extracted from hydrogen sources and go through the pre-accelerating phase in several different accelerators. Starting from the Linear Accelerator (LINAC) to the Proton Synchrotron (PS) then onto the Super Proton Synchrotron (SPS), the hadrons are injected in the LHC. The hadrons will only collide at certain collision points, where detectors are built to study the collision products. On the LHC ring there are four major detectors: ALICE, CMS, LHCb and ATLAS. The ALICE detector is focused on the study of heavy-ion processes generated by lead-lead collision, while LHCb mainly studies b-physics. The ATLAS and the CMS are two more general-purpose detectors, measuring

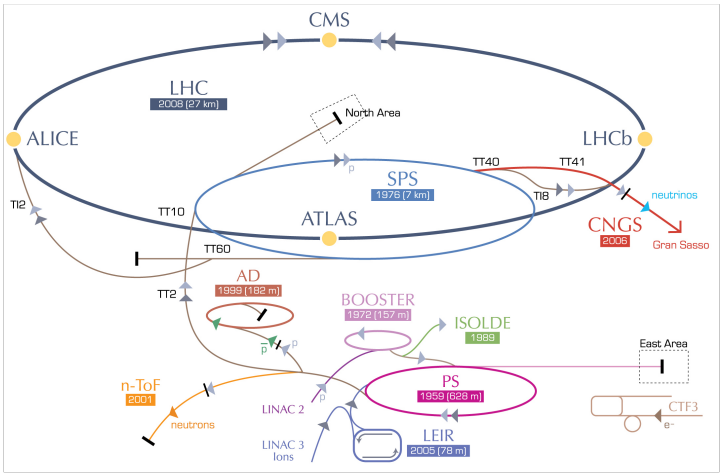


Figure 2.1.: The LHC accelerator complex. The plot is taken from Ref. [13].

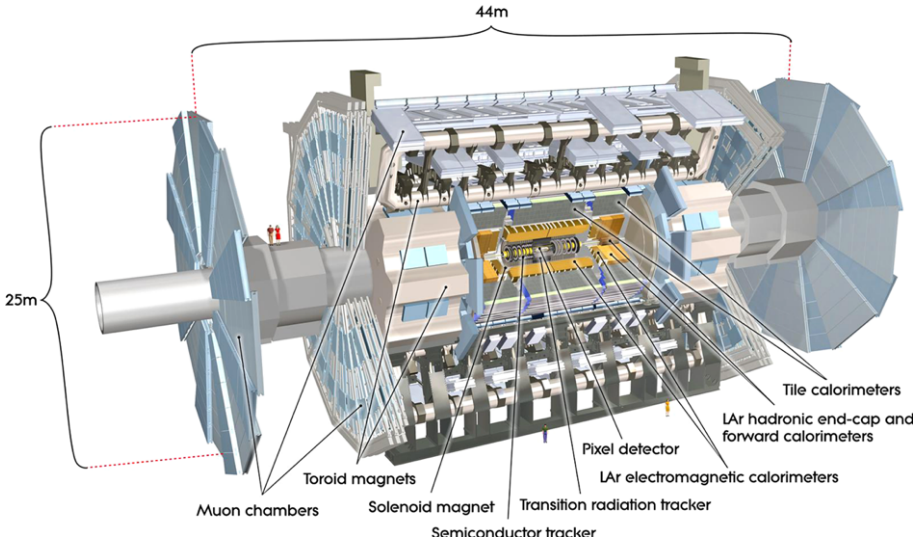
the properties of the Standard Model and the Higgs boson, as well as searching for physics Beyond the Standard Model.

2.2. The ATLAS experiment

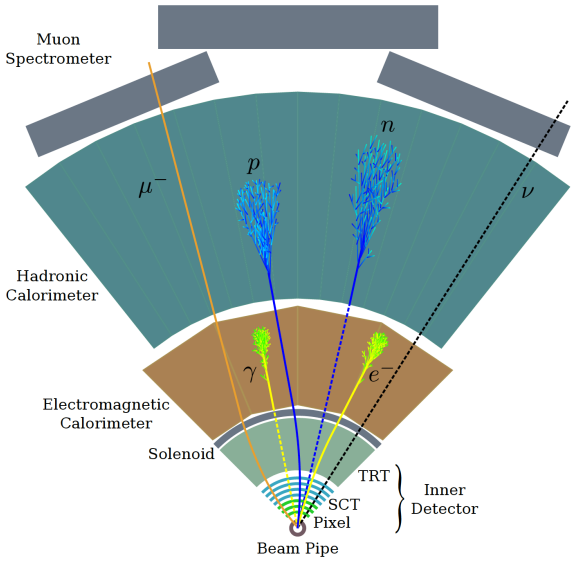
The ATLAS (A Toroidal LHC ApparatuS) detector is a general-purpose detector and has a number of different sub-detectors. A more general drawing showing the sub-detector distributions is in Figure 2.2a and sub-detector details show their distributions in the plane that is perpendicular to the beam direction. This layered distribution can ensure that the product particles can be identified and measured. [14] [12]

For the structure of the detector, the inner detector is composed of the pixel detector the semiconductor tracker (SCT), and transition radiation tracker (TRT). One layer outside is the electromagnetic calorimeter, where particles lose energy through electromagnetic showers. The hadronic calorimeter collects particles from hadronic showers with jets, which are sprays of hadron particles in the detector. The definitions of jets are discussed in subsection 2.3.2. The muon chambers are the most outside ones, mainly for muon detection.

Particles are detected inside this detector as shown in Figure 2.2b. Charged particles are detected in the tracking system as they have electromagnetic charges. Electrons and positrons deposit energy in the electromagnetic detector (ECAL) via electromagnetic showers. Photons do not have electric charge so they are invisible to the inner trackers but they produce electromagnetic showers and deposit energy in the ECAL. Hadrons leave energy depositions in both the ECAL and in the hadronic calorimeters (HCAL). Muons pass through the layers as they are minimum ionization particles, and reach the muon detectors. The energy and momentum of each of the particle in a collision event can be



(a) The ATLAS detector with sub-detectors in detail. Plot taken from Ref. [14]



(b) The ATLAS detector layers and particles detected. Plot taken from Ref. [13]

Figure 2.2.: The ATLAS detector introductions.

reconstructed. If missing transverse momentum is present, it can signal the production of invisible particles, such as neutrinos.

Bunches of protons collide at the LHC approximately every 25 ns, and for each collision, multiple protons may collide. This, combined to the large number of electronic channels in the ATLAS detector, prevents recording all of the events. As a result, the ATLAS detector uses a so-called trigger system to choose whether the event is going to be recorded. ATLAS has two levels of triggers, the first level (L1) trigger and the high-level trigger (HLT). L1 trigger is implemented in hardware and works with a lower latency. The HLT performs software selections on the events that have passed the L1 selections.

2.3. Observable Variables and Jet finding in the ATLAS detector

2.3.1. Observable Variables

The ATLAS detector has a cylindrical structure and in order to describe the spatial distributions of the detected particles, there needs to be a coordinate system for the detector. The ATLAS coordinate system uses right-handed coordinates with the z-axis along the beam direction, the x-axis pointing to the accelerator ring center and the y-axis pointing up perpendicularly to the other axes [14]. This coordinate system is described in Figure 2.3a. Apart from the coordinate system, angular variables are also used to describe spatial distributions for the particles in the event. The polar angle θ is measured from the beam direction and the azimuthal angle ϕ is measured from the x-axis direction. In this definition p_Z is the momentum along the z-axis direction. Since there is the reference system of particles and the reference system of the laboratory, using Lorentz-invariant variables serves the purpose of describing the particles better. Therefore, rapidity (y), a Lorentz-invariant variable, is introduced as Equation 2.1. The pseudo rapidity η is also a common used variable and it is related to θ, defined as Equation 2.2. It can be deduced that if the particle is traveling close to the speed of light or the mass of the particle can be neglected, pseudo rapidity and rapidity are the same. The relation between pseudo rapidity and θ is shown as Figure 2.3b, when the particles are massless. The rapidity difference is noted as y*. y* is defined as Equation 2.3.

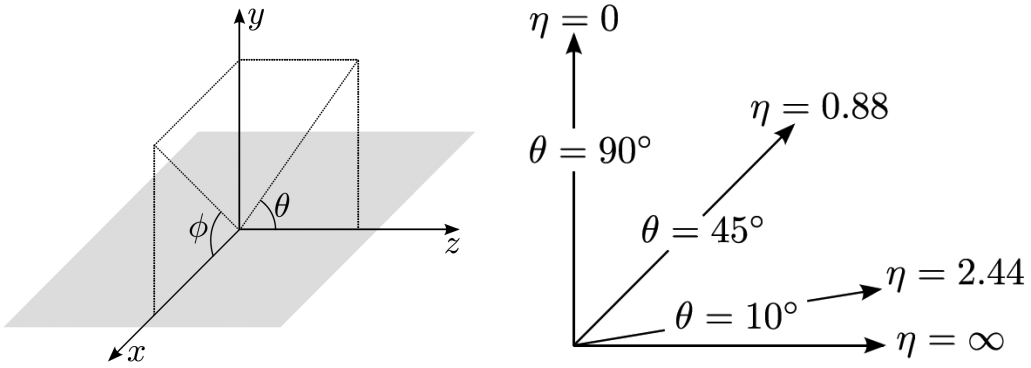
$$y = \frac{1}{2} ln(\frac{E + p_Z}{E - p_Z}) \tag{2.1}$$

$$\eta = -ln(tan(\frac{\theta}{2})) \tag{2.2}$$

With these variables, the distance inside the η - ϕ space can be described as Δ R, in Equation 2.4.

$$y* = \frac{|y_i - y_j|}{2} \tag{2.3}$$

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \tag{2.4}$$



(a) The coordinate system of the ATLAS detector, with the axises and angle measurements. Plot taken from [15] [12].

(b) The relation between η and θ in geometry. Plot taken from [15]

Figure 2.3.: The ATLAS detector coordinate system and connection between azimuthal angle and pseudorapidity. Plot taken from [16].

As a measurement of the energy in the detected particles, transverse momentum p_T is introduced. The colliding beams come from opposite direction and collide at collision points. This means there was no transverse momentum before the collision, but after the collision, the product objects would carry transverse momentum, which is the "new" momentum and the sum of all transverse momentum needs to be back to 0. These features make p_T an important variable in many searches.

2.3.2. Jet-finding Methods at the ATLAS

In detectors, jets are sprays of hadrons in a cone. The most basic idea to define the jet is to use the distance between two objects to check whether they are from the same quark or gluon, thus forming the same jet. There are now several different algorithms to do this [17].

In the family of algorithms used in this report, the distance between the objects i and j that need to be clustered is d_{ij} and the distance between an object and the beam is given as

$$d_{ij} = min(k_{t_i}^{2p}, k_{t_j}^{2p}) \frac{\Delta_{ij}^2}{R^2} \tag{2.5}$$

$$d_{iB} = k_{t_i}^{2p} \tag{2.6}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, R is the radius parameter; and $k_{t_i}^{2p}$ is the transverse momentum, ϕ_i stands for the azimuthal angle, and y_i is the rapidity of object i. The p parameter varies between different jet search algorithms. It can be -1, 0 and 1, corresponding to anti-k_t, Cambridge/Aachen, and k_t algorithm. For anti-k_t algorithm that we

use for the rest of this report, the jet finding starts from the highest-pT object. For k_t, jet finding starts from smallest-pT object while for Cambridge/Aachen, there is no energy measure [15,18,19]. The full set of steps for the anti-k_t algorithm with p = -1 is given below:

1. For each i,j (i≠j), perform d_{ij} and d_{iB} calculations according to Equation 2.5 and Equation 2.6.

2. Search for the i, j for the minimum d_{ij} and recombine i j.

3. Calculate the minimum d_{iB} and i is defined as a jet and removed from the list of objects.

4. Repeat from 1. until no object is left.

Jets obtained in this way are a set of objects combined together. In the case of jets that have been formed by grouping together detector energy deposits, a calibration procedure is needed to bring them to the correct energy of the particles that were generated by the fragmentation of the quark or gluon originating the jet, called the Jet Energy Scale [19–21]. Jet calibrations involve a set of processes that correct for the energy lost before the calorimeters and outside the jet radius, point jets back to the most energetic collision vertex rather than to the center of the detector, and remove effects of multiple proton-proton interactions. The uncertainty on the knowledge of the jet energy is called the jet energy scale uncertainty, and it is as small as 1 % in ATLAS.

2.4. Monte Carlo Simulation

The simulation of both physics processes and interaction of particles with matter can be used to obtain information of how data would appear in the detector. [22] Simulations are performed step-by-step to mimic how a proton-proton collision would appear in the detector readouts and could be reconstructed. Monte Carlo event generators simulate the physics process, while detector simulation programs simulate the interaction of the final-state particles with the detector. Typical generator examples are PYTHIA [23], HERWIG [24], and MadGraph [25]. GEANT4 [26] is a widely used detector simulator.

Monte Carlo event generators use pseudo-random numbers to mimic the probabilities in quantum mechanics [22]. In real-life detectors only the final, more long-lived or stable hadrons would be visible. In the simulation it is possible to break up the steps that happen before the particles enter the detector. Firstly, the main physics process to simulate (e.g. the creation of a dark matter mediator particle from a proton-proton collision, as well as its decay) is generated as the highest-energy interaction within the collision. This is called the "hard process". The hard process gives rise to short-lived "resonances" such as the dark matter mediators, and the decays of resonances are part of this process. For each pair of

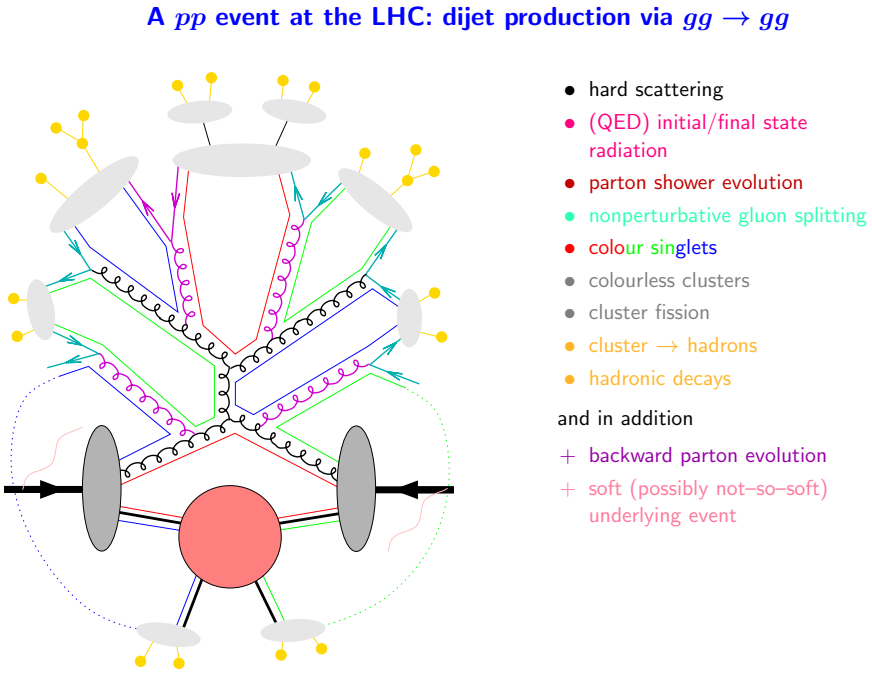


Figure 2.4.: A plot showing a simulation process at the LHC. Figure taken from Dieter Zeppenfeld's PiTP 2005 lectures.

partons interacting, further processes are simulated, for example the radiation emitted by the interacting partons, called initial state radiation (ISR) and final state radiation (FSR). The combination of all processes beyond the hard process in QCD is called the "parton" shower. Due to QCD confinement, the created partons cannot exist on their own, and "hadronize" in many new hadrons [22]. Even those hadrons may not be stable, so their decay processes also need to be simulated. In the detector, particles from hadronization can be grouped in the form of jets with same methods as introduced in 2.3.2. Jets built from final state simulated particles are called *truth jets*. After the detector simulation has taken place, it is possible to compare simulated events and real data taken from the detector. Jets at this level are "reconstructed jets". An example of a possible MC process at the LHC is shown in Figure 2.4.

Monte Carlo simulation is also important as it allows to design new searches, as in simulation it is clear what the background and what the signal is. Signals are the physics processes that are of interest, for example beyond the SM, while other processes with the same signature in the detector form the background. In the case of this search, we consider signals of dark matter mediators decaying into quarks, and the background are all the QCD processes that very frequently generate two jets in LHC detectors.

In experimental data, signals and background can only be distinguished discriminating the characteristics of their signatures in the detectors, while in MC simulations they can be generated and studied separately. By comparing signal plus background

simulation with experimental data, theories can be verified or refused. For these reasons, MC simulations are of great importance to physics analysis nowadays.

2.5. Signals and Background for the Search in This Report

The main part of the report is the study of the kinematic distributions of the light dark matter mediator decaying into two jets, and the mediator is produced in association with a jet or a photon [27]. The dark matter mediator (DMM) can decay either into DM particles or into SM particles. Many models including DMM involve a light resonance $Z^{'}$, which has same properties as Z boson but a higher mass [28] [29]. Since they can been produced at the LHC from a proton-proton (parton-parton) interaction, these resonances can have sizable couplings to quarks and gluons as well. These resonances can be produced in association with a jet or a photon, from the radiation of the initial or final state of one of the partons involved in the hard scatter.

This kind of signature has a relatively low dependence on the theoretical models since the idea of mediator resonance is proposed by many models. This search can also have a purpose beyond searching for DM mediator, since there are many models that predict resonances decaying into dijets.

DM mediators that have a mass below 1 TeV are called "light resonance". They could be produced at the energy scale reachable by the LHC. However, since signal and QCD background are nearly indistinguishable and the data collection system could not cope with recording all background, signal and background events are discarded equally. A way to overcome this is to employ a high-p_T radiation object to select the event, as this reduces the QCD background and makes the data rates sustainable for recording. [1].

The decay products of the light resonance can be either boosted (both decay products collimated, due to the Lorentz boost) [27] or resolved [27, 33]. In this study, we investigate the resolved regime.

The Feynman diagrams of the search are given in Figure 2.5.

However, when studying the final states in the resolved regime, there are experimental difficulties. In the case where the radiated object is a photon, the jets from the resonance can be deduced clearly to be the two leading jets as they are accompanied by a photon. In the radiated jet case, there are three jets involved and only two of them are decay products of the resonance. The wrong choice of the jets to reconstruct the resonance leads to the wrong reconstructed dijet mass. The resonance mass distributions with the wrong jet

[1] Another technique used to reach light resonances at the LHC is that of only saving final state information reconstructed at the trigger level, called Trigger-object Level Analysis [30–32]

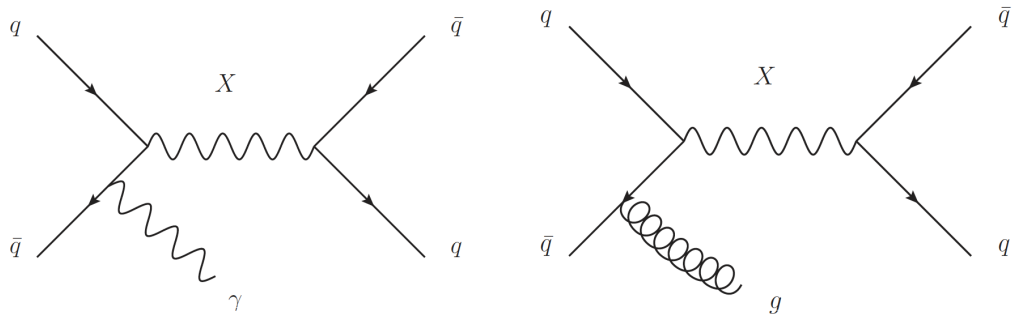


Figure 2.5.: Main Feynman diagrams of interest in this search. Plots are taken from Ref. [27].

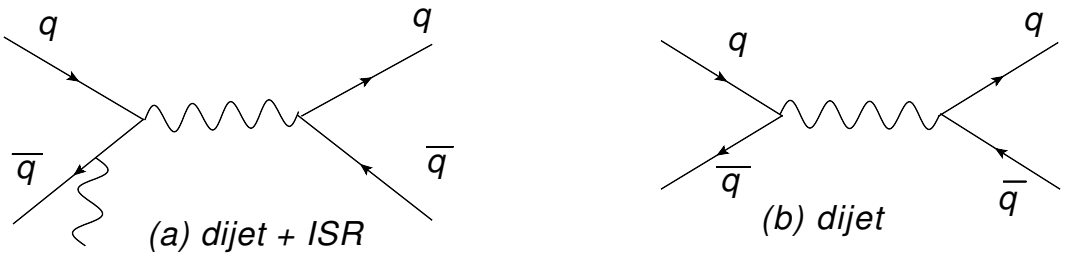


Figure 2.6.: Feynman diagrams for different sample generation. Plot made with JaxoDraw: [34].

choice is therefore wider compared to the mass peak obtained from the correct choices. Since these signals produce resonant excesses on top of the smooth QCD background, the wrong jet choices make the peak more difficult to identify and therefore reduce the sensitivity of the search. The aim of this work is to study the case where a gluon is radiate (using the naming convention of "dijet + ISR jet" in this report), in order to understand how to choose the correct dijet pair and improve the sensitivity of the search with respect to the current version where jets are simply ordered according to their transverse momentum.

Concerning the simulated samples used in this search, I use two kinds of samples that have been generated centrally in the ATLAS Monte Carlo production system. For the first kind of samples, the events are generated with specific selections on the leading jet transverse momentum and only include the dijet + ISR Feynman diagram in Figure 2.6a. These first samples use MadGraph for the process generation and Pythia 8 for the parton shower generation. They also require an extra initial state jet in the generation [27]so we name them as "dijet + ISR" samples in later discussions. Those are the signals used in the current search. Another kind of samples did not have any specifications on jet transverse momentum. They use MadGraph for the hard process but only generate the dijet process in Figure 2.6b, and let Pythia 8 generate the radiation through the parton shower. They also use Pythia 8 for parton shower. These are named as "dijet" samples. The generation of these samples is detailed in [33]. These samples are at "reconstruction level", which means that they have gone through the detector simulation after the generator generation.

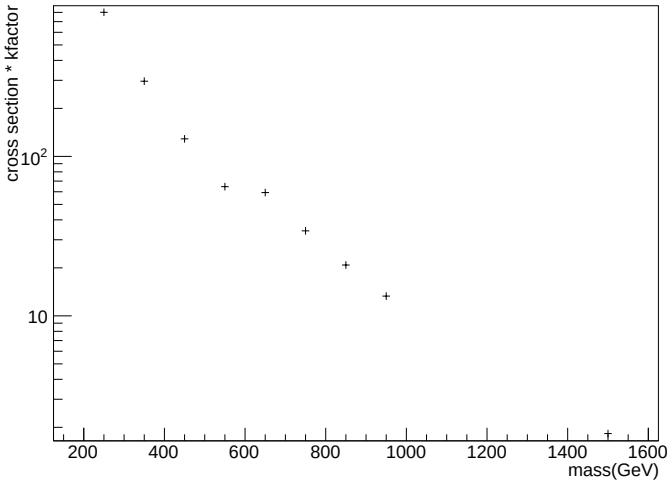


Figure 2.7.: Cross section at different mass points, below 600 GeV for dijet + ISR samples, above 600 GeV for dijet samples.

2.6. Comparisons Between the Dijet + ISR Samples and the Dijet samples

Since I am using two kinds of samples, it is necessary to check their consistency to first order. Firstly, in the generation process, the cross section for the samples should be relatively smooth: as the mass of the mediator grows, the cross section decreases. The cross sections are shown in Figure 2.7. Even though there is a transition between the cross-sections of the two kinds of samples, we consider this performance sufficient for the studies in this report. [2]

Another way to check the performance of these two samples is to check the variable distributions with different kinds of samples at the same mass. The p_T in the plots stands for the transverse momentum of the object. For most of the later discussions, the jets will be ordered by p_T and so a convention in this report is the leading p_T jet named as jet1, sub-leading p_T jet as jet2, in this order. The angular variables are related to the spatial distributions for the objects.

For 350, 450 and 550 GeV, both the dijet and dijet+ISR samples are available. Only the most basic variables, the reconstructed resonance masses, p_T distributions, $\Delta\phi$ for the 450 GeV mass point are shown here as examples.

The analysis selection that is needed to ensure that the trigger selection is fully efficient, that jets are in well-understood regions of the calorimeter, and to enhance signal over background, is below:

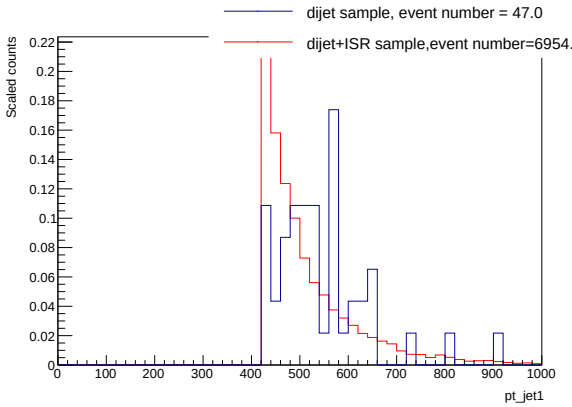
[2] Most of the analysis figures in this study are made with ROOT CERN tools, [35].

1) leading jet p_T above 420 GeV (to mimic the cut applied in the trigger system), all jet p_Ts above 25 GeV (because this is the minimum value where reliable calibrations are present);

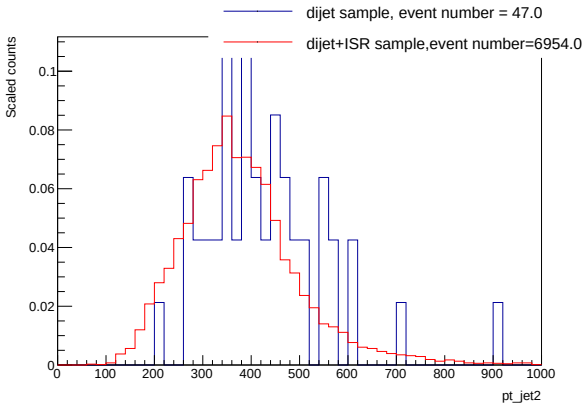
2) $y^*_{dijet} < 0.8$ (to improve discrimination between signal and background).

Figure 2.8 to Figure 2.10 are the p_T, $\Delta\phi$ and mass distributions for dijet + ISR samples overlaid with dijet samples, with this selection applied. The sharp edges in p_T distributions come from the p_T selections.

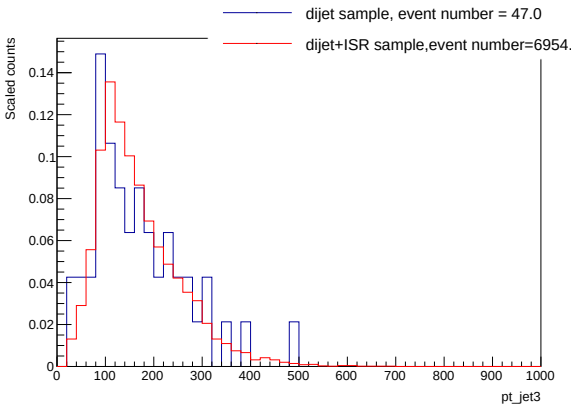
As shown in the legend of the plots, due to no generation selections, the p_T selection later applied would remove most of the events in dijet samples since the number of small-p_T events is relatively large at 450 GeV mass. Therefore, there are much fewer events in dijet samples since the leading jet p_T cut removes most of the events. This is expected. As for the distribution shapes, which indicate the physics for these samples, they are similar in two different samples with the statistical errors. This indicates that the two samples are physically comparable, so it is acceptable to use these two kinds of samples (with the appropriate selections) together in these studies. Because the number of events passing the trigger selections is rather small for dijet samples at lower masses, a decision has been made about what samples to use at what energy. The choice of samples is dijet + ISR samples for simulated resonance invariant masses below 600 GeV, and dijet samples for masses above 600 GeV.



(a) jet1 p_T distribution

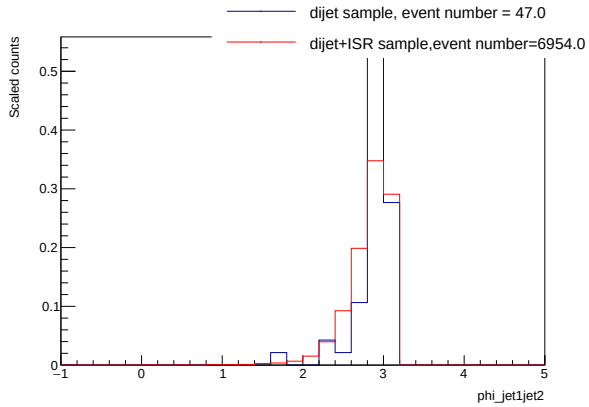


(b) jet2 p_T distribution

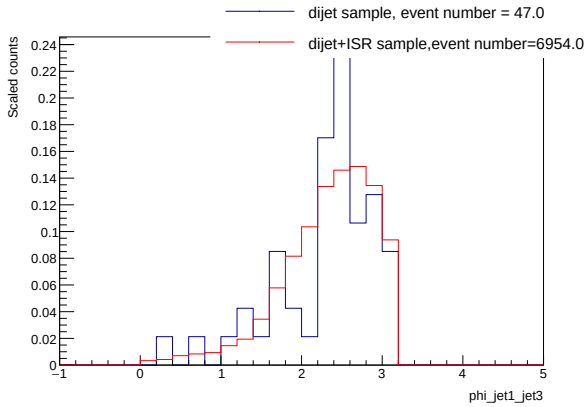


(c) jet3 p_T distribution

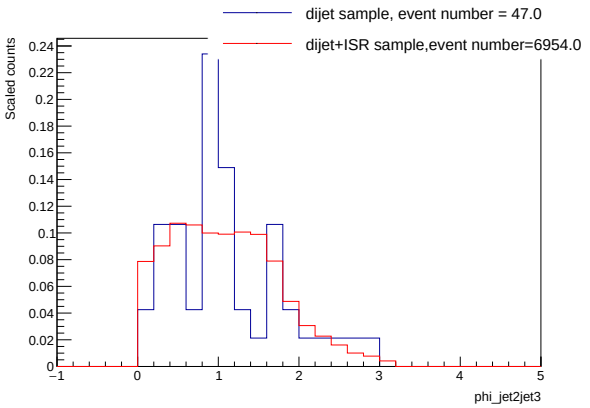
Figure 2.8.: p_T distributions of dijet + ISR and dijet sample at 450 GeV.



(a) $\Delta\phi_{jet1jet2}$

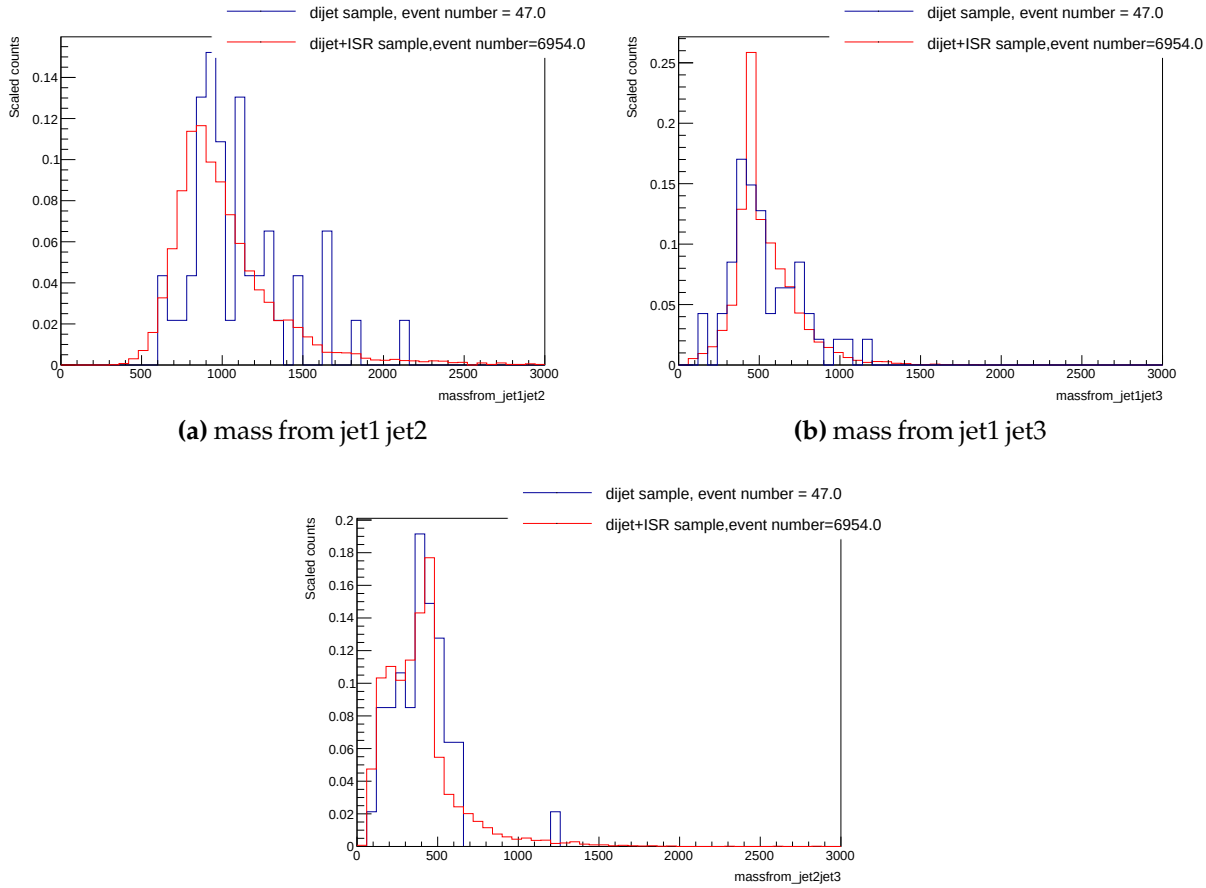


(b) $\Delta\phi_{jet1jet3}$



(c) $\Delta\phi_{jet2jet3}$

Figure 2.9.: $\Delta\phi$ distributions of dijet + ISR and dijet sample at 450 GeV.



(a) mass from jet1 jet2

(b) mass from jet1 jet3

(c) mass from jet2 jet3

Figure 2.10.: Resonance masses reconstructed from jets of dijet + ISR and dijet sample at 450 GeV.

Chapter 3.

Comparisons Between Jet and Photon Radiation

The main issue studied in this work arises from the difficulty to choose the correct jet pair in the light resonance search with dijet + ISR [1] jet events. The ISR can also be a photon. Both cases involve three objects, two jets and an ISR. In the case of ISR photon, it is easier to identify the resonance dijet from the ISR as one could identify the photon first. ISR jets instead cannot be easily distinguished by the jets from the resonance, and as discussed in the previous section this affects the sensitivity of the search to new physics models.

Firstly, to understand whether the ISR jet has characteristics that make it distinguishable from the resonance jets is to compare some kinematics variable distributions for ISR jet and ISR photon samples, checking the common characteristics of the two kinds of ISRs, regardless the ISR being a jet or a photon. Kinematics variables that are used for this comparison and for later studies are also introduced in the first part of this chapter.

3.1. Variable Introductions

In order to "select" the jets from the resonance or the ISR jet, various variables are introduced here, based on p_T, the angular observables ϕ and η. These kinematics variables can help understand the spatial distributions and energy distributions of each individual jets, and they are used for comparing the ISR jet samples in section 2.5.

[1] In the following, we adopt the convention of calling the radiation "ISR" even though it is in principle not distinguishable from FSR.

In addition to the variables related to individual jets, new variables [36] can be formed to describe the dijet system. The p_T asymmetry between jet i and jet j, p_T_asymmetry, is defined as Equation 3.1.

$$p_T\_asymmetry = \frac{|p_{T,jeti} - p_{T,jetj}|}{|p_{T,jeti} + p_{T,jetj}|} \tag{3.1}$$

The p_T_asymmetry can show the imbalance between two jets, and give an idea of how their magnitudes are compared with each other. If the two jets are similar in p_T, the p_T_asymmetry would be close to 0 while for quite different p_Ts, the p_T_asymmetry value would be larger, to the extreme of 1.0. The spatial variables are also taken into consideration. The differences in η and ϕ angles between two jets are noted as $\Delta\eta$ (DeltaEta) and $\Delta\phi$ (DeltaPhi)Similarly, Δ R (Delta R) we use is introduced as Equation 3.2 using $\Delta\phi$ and y*.

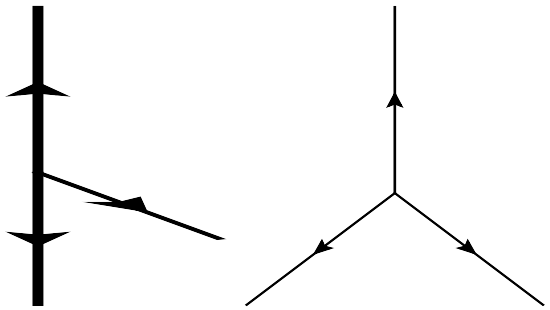
$$\Delta R(jeti,jetj) = \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2} \tag{3.2}$$

To understand the spatial distribution of all jets in each event, the thrust variables [37] are introduced to show how aligned or anti-aligned the jets in each event are. These variables may also be used to identify the ISR jet. The event shape variables [37] are calculated event by event, thus instead of individual jets, they show the characteristics of a whole event. The y_{23} is a measurement of how large the third jet p_T is compared with the first two jet p_T, definition given in Equation 3.3 where $H_{T,2} = (p_{T,1} + p_{T,2})$ is the p_T sum of the two leading p_Ts.

$$y_{23} = \frac{p_{T,3}^2}{H_{T,2}^2} \tag{3.3}$$

The thrust variable definitions are from Equation 3.4 to Equation 3.6. In the case of this discussion, only the three leading jets are taken into consideration in calculations. The key thrust variable is the $T_\perp$, given in Equation 3.4. The unit vector $\hat{n_\perp}$, is the "thrust axis", which is the direction that maximizes the $T_\perp$. For $T_\perp$, if the jets are very "back-to-back" aligned, the value would be close to 1.0 while in the more spatially distributed jet case, the $T_\perp$ value would be smaller. [2] An example is shown in Figure 3.1, it is clear to see the difference between high-$T_\perp$ jet distributions and low-$T_\perp$ jet distributions. To see the unaligned part more clearly, the variable $\tau_\perp$ (Equation 3.5) is introduced and it can show how different $T_\perp$ is from 1.0 in the samples. The $T_{m,\perp}$ as in Equation 3.6, on the other

[2] The code towards the thrust-related variable calculations are from Antonio Boveia, [38]



(a) Jet distributions in high-$T_\perp$ events **(b)** Jet distributions in low-$T_\perp$ events

Figure 3.1.: Jet distributions in events with different $T_\perp$ values.

hand, gives the sum of the p_Ts that are not aligned with the thrust axis.

$$T_\perp = \max_{\hat{n_\perp}} \frac{\sum_i |p_{Ti} \cdot \hat{n_\perp}|}{\sum_i p_{Ti}} \tag{3.4}$$

$$\tau_\perp = 1 - T_\perp \tag{3.5}$$

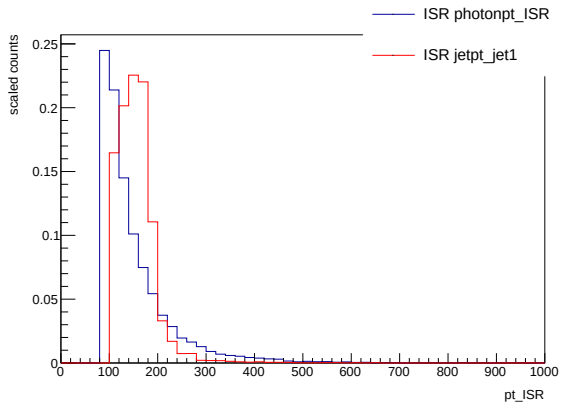
$$T_{m,\perp} = \frac{\sum_i |p_{Ti} \times \hat{n_\perp}|}{\sum_i p_{Ti}} \tag{3.6}$$

To see how the dijet combination would perform, each two of the three jets are combined into possible "resonances" by adding the four-vector of the jets and checking how close that value is to the nominal mass of the resonance. The remaining jet that is not used for this combination is assumed to be the "ISR".

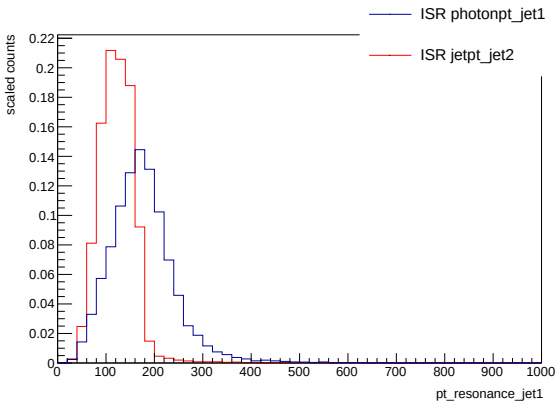
3.2. Comparisons Between the ISR Jet Samples and the ISR Photon Samples

In this section, ISR jet samples and ISR photon samples are compared in the hope of finding common characteristics of the ISR objetcts. However, several major differences exist between the centrally-generated ISR jet and ISR photon samples that prevent this comparison to be fully useful. The differences and the results are in this section.

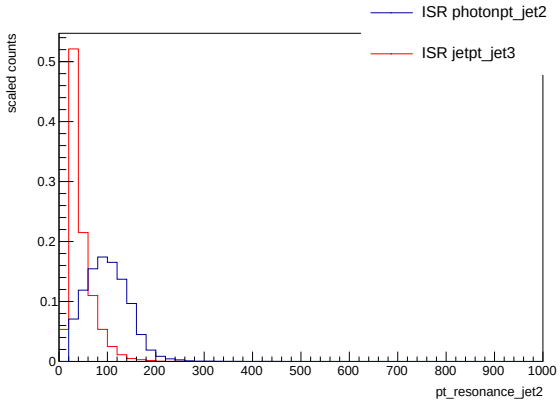
Firstly, the trigger selection requires the photon p_T to be above 85 GeV in the photon ISR samples, while the selection on the leading jet p_T is more stringent ($p_T > 420$ GeV) [27]



(a) ISR photon (ISR photon) and ISR jet (ISR jet) p_T



(b) Jet1 p_T (ISR photon) and Jet2 (ISR jet) distributions

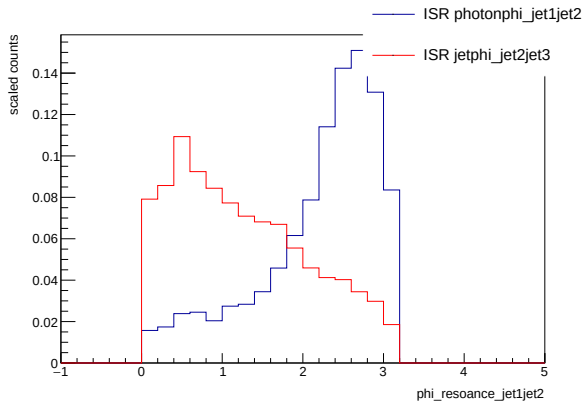


(c) Jet2 p_T (ISR photon) and Jet3 (ISR jet) distributions

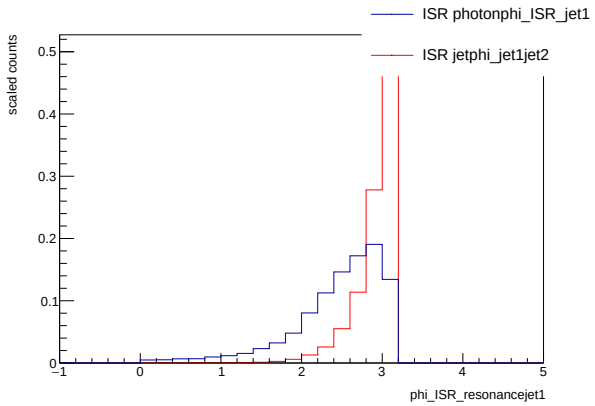
Figure 3.2.: p_T distributions of ISR photon sample compared with ISR jet sample at 250 GeV.

and it is applied to the leading jet no matter what the ISR is. For this reason the ISR jet samples have a filter where the leading truth jet p_T needs to be above 350 GeV applied already in the simulation. This means that, to be comparable, the photon samples need to have a similar cut applied as the jet samples. This reduces the statistics that is available for this comparison. Nevertheless, the distributions of p_Ts, $\Delta\phi$ and $T_\perp$ plots are made for comparisons from low mass to high mass samples with photon samples as well. From low mass to high mass, 250, 450 and 1500 GeV ISR photon samples are selected and studied. The p_T distributions and $\Delta\phi$ distributions of 250 GeV ISR photon sample are given as examples here in Figure 3.2 and Figure 3.3.

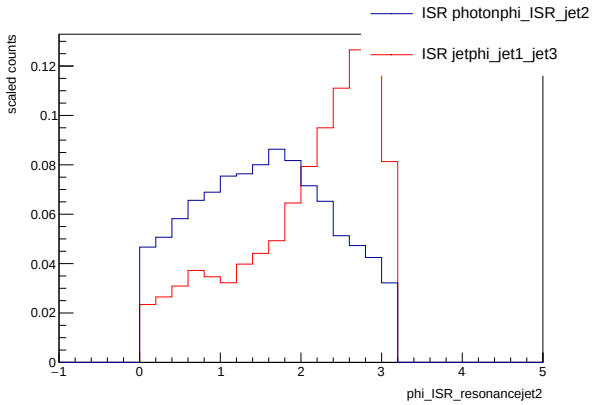
The kinematic distributions of the ISR photon samples, as understood from from these plots, is shown in Figure 3.4. The length of the jet arrows stands for the p_T magnitude and the direction of the arrows shows the spatial distribution. The jet distributions as in Figure 3.4 are named as topology in this report. At low mass, the jet p_Ts are small and comparable with photon p_T. When the mass of the resonance gets higher, the p_Ts of the jets increase, becoming larger and larger compared to the photon p_T. This behavior is described in Figure 3.4. These topology distributions are different from the ISR jet case, and do not



(a) $\Delta\phi_{jet1jet2}$ (ISR photon) and $\Delta\phi_{jet2jet3}$ (ISR jet)



(b) $\Delta\phi_{jet1ISR}$ (ISR photon) and $\Delta\phi_{jet1jet2}$ (ISR jet)



(c) $\Delta\phi_{jet2ISR}$ (ISR photon) and $\Delta\phi_{jet1jet3}$ (ISR jet)

Figure 3.3.: $\Delta\phi$ distributions of ISR photon sample compared with ISR jet sample at 250 GeV.

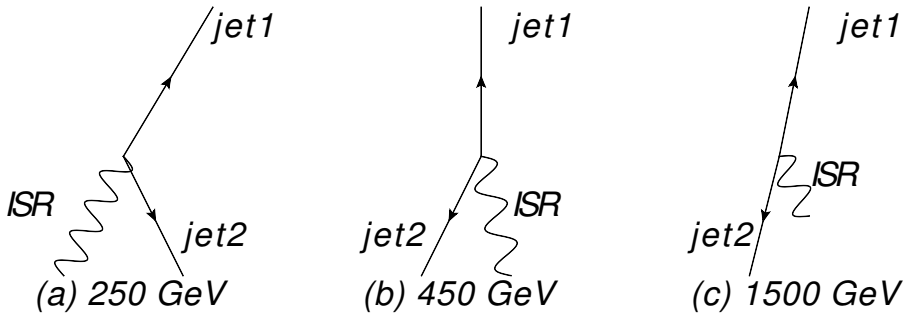


Figure 3.4.: Topology distributions for ISR photon samples from low mass (250 GeV) to medium mass (450 GeV) to high mass (1500 GeV).

facilitate further comparisons. This indicates that we need to use another approach to choosing the resonance dijet pair, focusing on ISR jet samples. chapter 4 discusses this.

Chapter 4.

Kinematics Studies

After comparing ISR jet samples with ISR photon samples, the focus is back on ISR jet samples. The kinematic distributions of the dijet + ISR samples at different masses within the range of 250 GeV to 1500 GeV are the main results shown in this chapter. All the distributions shown in this chapter are made using the simulated signal samples as introduced in section 2.5. The nominal resonance masses for these signals are known but information about which jet is the radiation and which are the resonance dijet is still unknown. For this reason, we consider the fractional difference of the dijet mass from the two chosen jets with respect to the nominal resonance mass as a discriminant on whether the right jet choice has been made.

As introduced in chapter 2, the selection cuts applied for these distributions are:

- leading jet $p_T > 420$ GeV, any other jet $p_T > 25$ GeV;
- $y^*_{dijet} < 0.8$.

For the kinematics studies, these selections are not yet applied but in the sample generation, dijet + ISR samples are generated with the cut leading jet $p_T > 350$ GeV, while the dijet samples are not. Since the dijet samples are only used at high masses, where leading jet p_T will usually be large, this should not affect the conclusion.

For these kinematic plots, we use only the three leading jets, assuming that two jets are from the resonance and one is the radiation jet. We have also checked that the sum of the energies and p_Ts of the three leading jets is above 95 % of the sum of the energies or p_Ts of all the jets in the event.

4.1. Distributions of Variables at Different Resonance Masses

For Z′ masses below 600 GeV, we choose to use the dijet + ISR samples and for Z′ masses above 600 GeV, dijet samples are used. An intuitive idea about the relation between the resonance masses and the ISR jet is that at low mass, the ISR jet tends to be the leading p_T jet and at high mass it tends to be the third leading jet. The reason for this lies in the leading jet p_T selection: if the resonance mass is 250 GeV but the leading jet is required to have a p_T of 420 GeV, then it is more likely that the leading jet generally does not come from the resonance and it is the radiated jet instead. On the other hand, when the resonance mass is higher, the two jets from the resonance will have a larger p_T, leaving the third leading jet to be identified as radiation. This idea guides the way the results are interpreted and is verified by the results in this section. From Figure 4.1, at 250 GeV, jet2 and jet3 can be reconstructed to obtain resonance masses very close to the simulated mass of the sample while at 1050 GeV, jet1 and jet2 perform well. The medium mass, on the other hand, would be a more complicated case since it is not clear how to identify the resonance jets from the radiation.

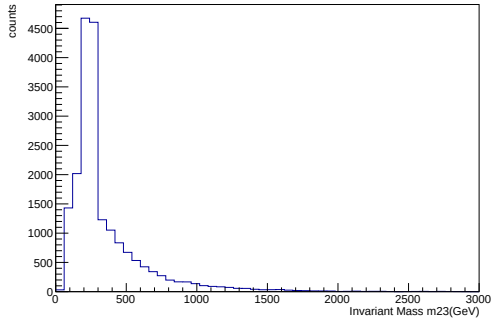
We have checked the distributions of all the variables introduced in section 3.1 but p_T, $\Delta\phi$, Δy and thrust-related variables are the most indicative ones when it comes to understanding the topology of the events, so we concentrate on those in this chapter.

To use the variable distributions as indicators for jet choices, the variables are expected to perform differently in the case that the jet pair that form the variable come from the Z′ and the case that the jet pair do not come from the Z′. The simulation samples do not contain information on which jets come from the resonance, so we need an approximate way to identify whether they are correctly reconstructing the resonance or not. The criteria is set to be that the reconstructed resonance mass from the two chosen jets falls in $\pm$ 20 % of the Z′ mass. In the case of 250 GeV, this means that the reconstructed resonance mass from jet2 and jet3 needs to fall in the mass range of 200 GeV - 250 GeV. Jet2 and jet3 here are the "preferred dijet" in this case.

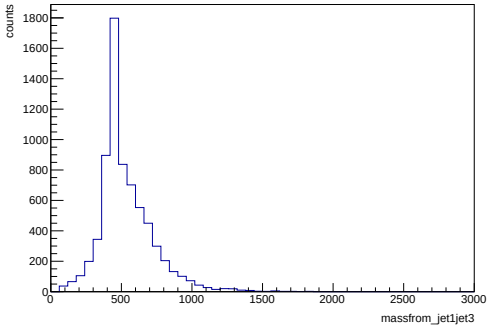
In the discussions in this and following chapters, the 250 GeV Z′ sample serves as a representative example of the low mass samples, the 450 GeV Z′ mass as a medium mass example and the 1050 GeV as a high mass example.

4.1.1. Low Mass Sample Kinematics Distributions

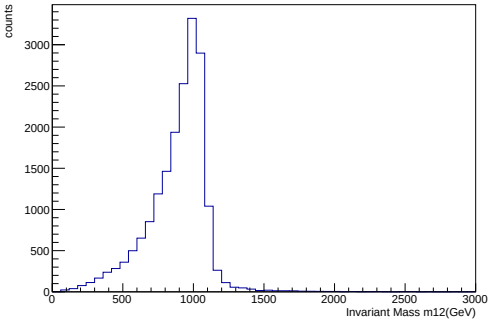
The p_T distributions are given in Figure 4.2 and $\Delta\phi$ between the jets are shown in Figure 4.3. The leading jet p_T and sub-leading jet p_T peak at around 400 GeV and 300 GeV



(a) resonance mass from jet2 jet3 for m(Z′)=250 GeV



(b) resonance mass from jet1 jet3 for m(Z′)=450 GeV



(c) resonance mass from jet1 jet2 for m(Z′)=1050 GeV

Figure 4.1.: Reconstructed resonance masses from the "preferred dijet".

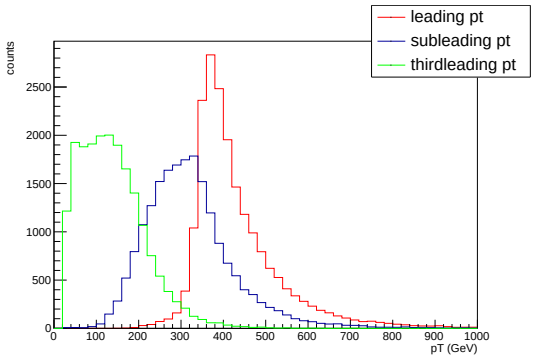


Figure 4.2.: p_T distributions at low mass of 250 GeV

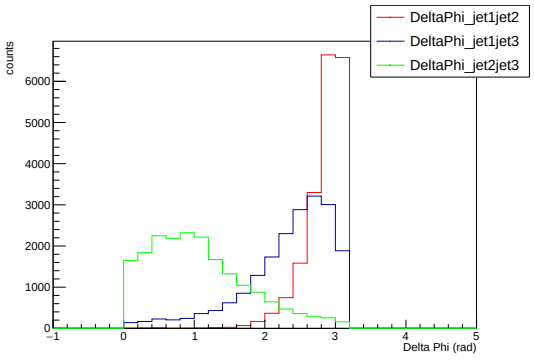
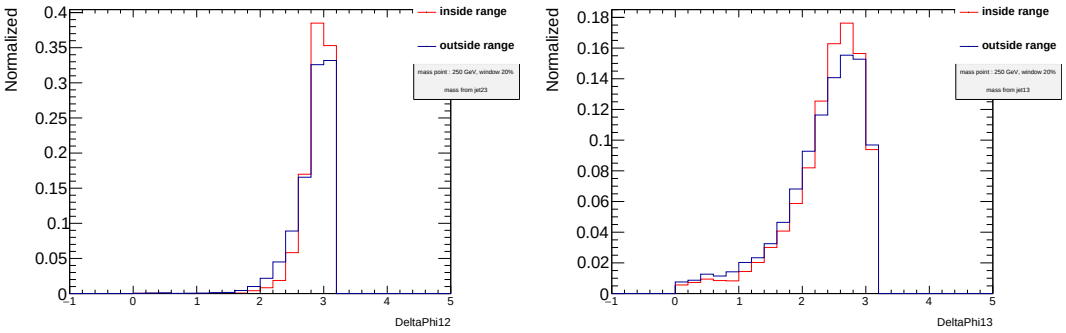


Figure 4.3.: $\Delta\phi$ relations between the jets in one event at low mass.

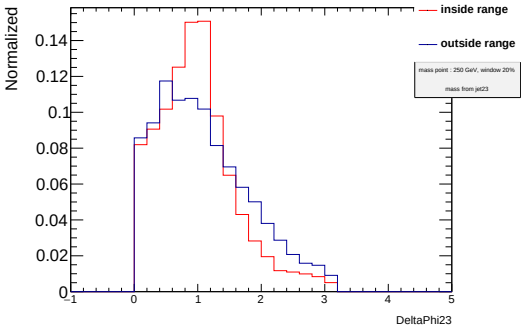
respectively. The third-leading jet p_T peaks around 150 GeV. The first and second, as well as the first and the third jets are mostly back-to-back with each other in terms of ϕ, while the second and third jets are closer together.

The distributions of $|\Delta y|$ and $|\Delta\phi|$ are shown in Figure 4.4 and Figure 4.5. From these results, it can be deduced that jet2 and jet3 are closer to each other in terms of both rapidity and ϕ for events inside the mass window than in the events falling outside the window. Jet1 tends to be more "back-to-back" with the other two jets when the event is inside the mass window. The thrust related variables can indicate how back-to-back the



(a) $|\Delta\phi_{jet1jet2}|$ inside and outside the mass window

(b) $|\Delta\phi_{jet1jet3}|$ inside and outside the mass window



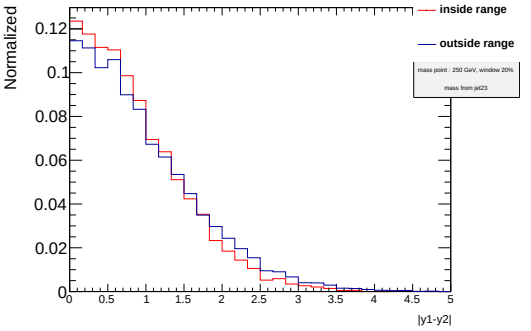
(c) $|\Delta\phi_{jet2jet3}|$ inside and outside the mass window

Figure 4.4.: $|\Delta\phi|$ of the events falling inside and outside the interested mass window range of 200 - 300 GeV.

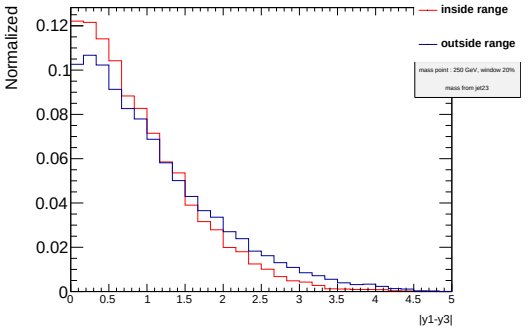
events are, as shown in Figure 4.6, the events falling inside the mass range have larger $T_{\perp}$ and smaller $T_{m,\perp}$ values. So jets in the events inside the mass window tend to be more back-to-back distributed in jet spatial distributions. This agrees with the results from rapidity and $\Delta\phi$ relations.

4.1.2. High Mass Sample Kinematics Distributions

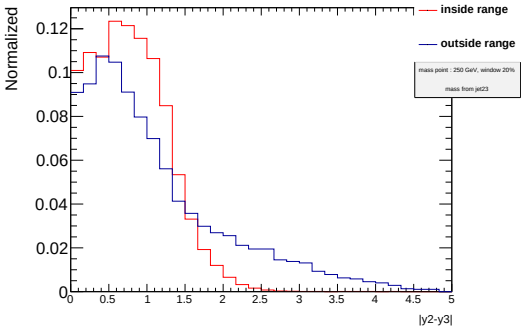
Similarly, the distributions for the high-mass sample (Z' mass of 1050 GeV) can be plotted and used to infer the topology of the jets in the events. The p_T distributions and $\Delta\phi$ distributions are shown in Figure 4.7 and Figure 4.8. Jet1 and jet2 p_Ts are significantly larger than jet3 p_T. The two leading jets are more back-to-back in topology and jet3 is closer to jet2. This agrees with the previous intuition that the p_Ts are balanced against each other.



(a) $|\Delta\ y_{jet1jet2}|$ inside and outside the mass window



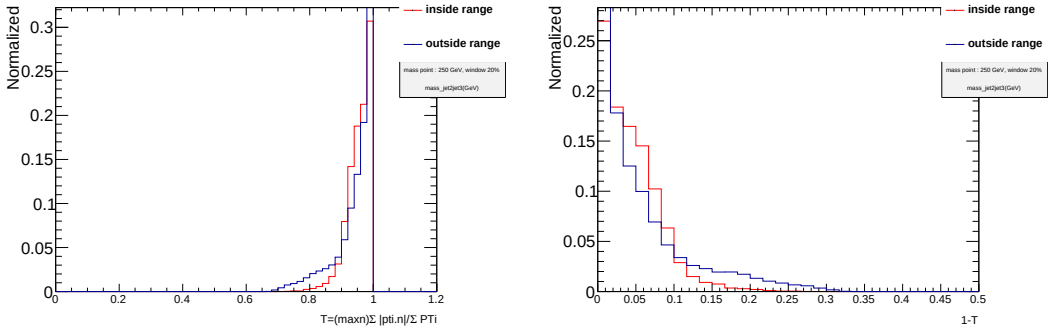
(b) $|\Delta\ y_{jet1jet3}|$ inside and outside the mass window



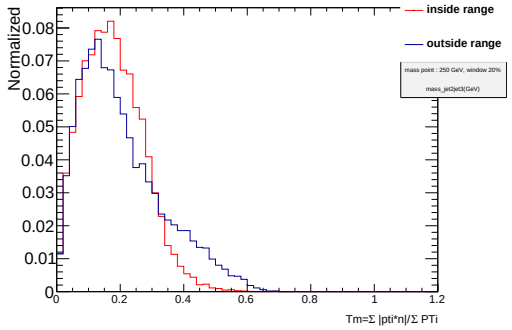
(c) $|\Delta\ y_{jet2jet3}|$ inside and outside the mass window

Figure 4.5.: $|\Delta\ y|$ of the events falling inside and outside the interested mass window range of 200 - 300 GeV.

The jets for resonance reconstruction for the high mass sample are chosen to be jet1 and jet2. Similar to the low mass case, variable distributions can be obtained when the resonance mass reconstructed from jet1 and jet2 falls inside and outside the mass window. Following the same $\pm$ 20 % mass window range criterion, the "inside" range for the 1050 GeV mass point is 840 GeV to 1260 GeV. The same variables checked in 250 GeV case can also be checked here. The $|\Delta\phi|$ and $|\Delta\ y|$ distributions are given in Figure 4.9 and Figure 4.10. Both events inside and outside the mass window are very much similar in terms of spatial distributions. But the p_T_asymmetry between jet1 and jet3 as well as jet2 and jet3, shown in Figure 4.11, indicates that when the event falls inside the mass window, it tends to have smaller third-leading jet p_T since the jet3 p_Ts in events inside the mass range tend to be much more different from the other jet p_Ts. The thrust related variables are shown in 4.12. They indicate that the jets in events inside the mass window are slightly more back-to-back distributed than jets in events outside the window.



(a) $T_{\perp}$ inside and outside the mass window **(b)** $\tau_{\perp}$ = 1 - $T_{\perp}$ inside and outside the mass window



(c) $T_{m,\perp}$ inside and outside the mass window

Figure 4.6.: Thrust related variables of events falling inside and outside the mass window of 200 - 300 GeV.

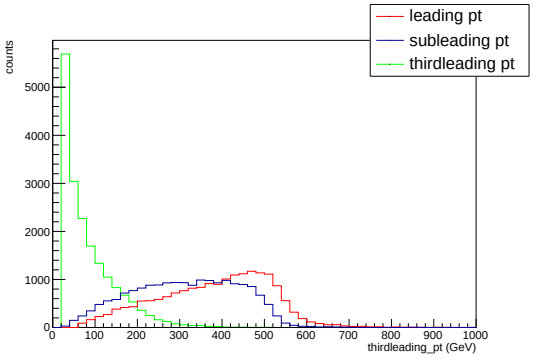


Figure 4.7.: p_T distributions at the high mass of 1050 GeV.

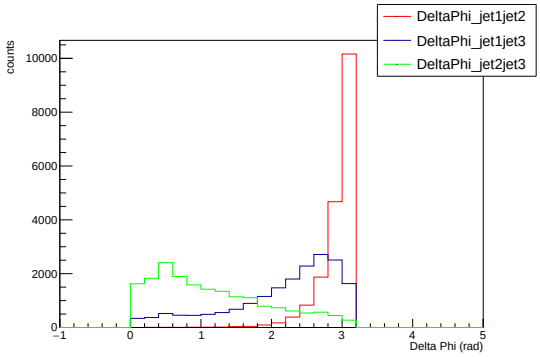
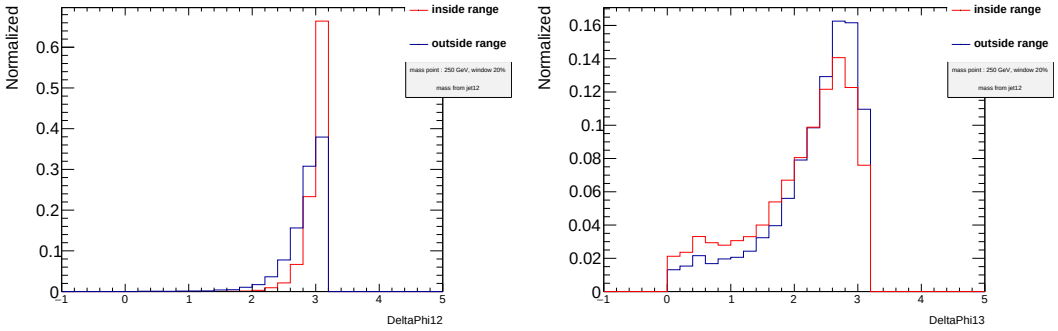


Figure 4.8.: $\Delta\phi$ distributions at the high mass of 1050 GeV.

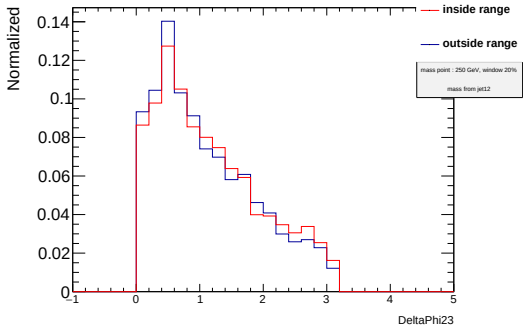
4.1.3. Medium Mass Sample Kinematics Distributions

In between the low mass and the high mass lies a more ambiguous mass range, the medium mass region. In those samples, the ISR jet identity is not clear but resonance masses reconstructed from jet1 and jet3 tend to be closer to the simulated mass, thus making jet2 the radiation jet, as given in Figure 4.13. The topology of low and high masses are very different from each other. Therefore, identifying characteristics of the process in between is not trivial as there will be a transition between the two topologies. The



(a) $|\Delta\phi_{jet1jet2}|$ inside and outside the mass window

(b) $|\Delta\phi_{jet1jet3}|$ inside and outside the mass window



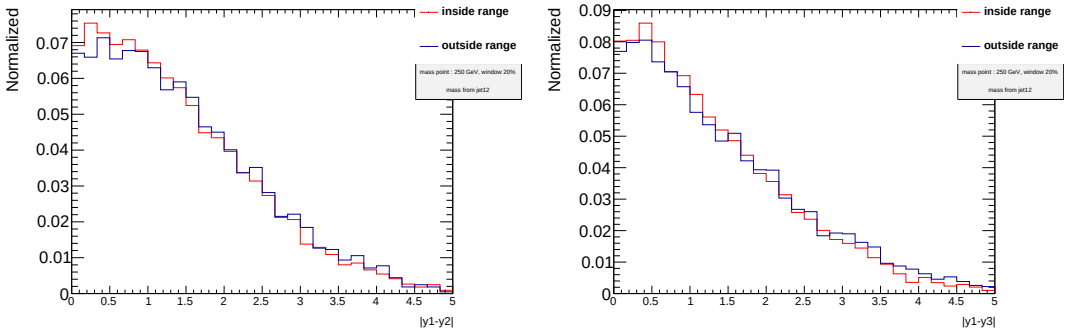
(c) $|\Delta\phi_{jet2jet3}|$ inside and outside the mass window

Figure 4.9.: $|\Delta\phi|$ of the events falling inside and outside the interested mass window range of 840 - 1260 GeV.

p_T distributions and $\Delta\phi$ distributions are shown in Figure 4.14 and Figure 4.15, they indicate that the topology of these events are in between the more spread out distribution at the low mass and the more back-to-back distribution at high mass. The thrust-related variable distributions are in Figure 4.16 and they indicate that the jet topologies are very similar inside and outside the mass range. Figure 4.14 shows that jet1 and jet3 are most likely to come from the resonance since they give the closest reconstructed mass. However, jet2 and jet3 can also give the correct mass for some events. Therefore, in medium mass region, the radiation jet is not clear.

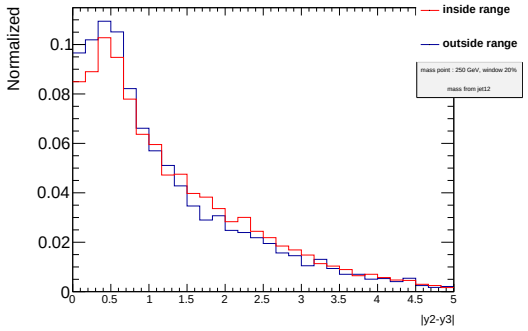
4.2. Discussion of Results on Kinematic Distributions

From the kinematic variable distributions above, it can be deduced that different events have different topologies. At the low mass and high mass, where things (and the ISR) are more clear, topology plots like Figure 4.17 can be drawn. At low mass, the jets in one event are overall more spread out. Jets of events inside the 20 % mass window are



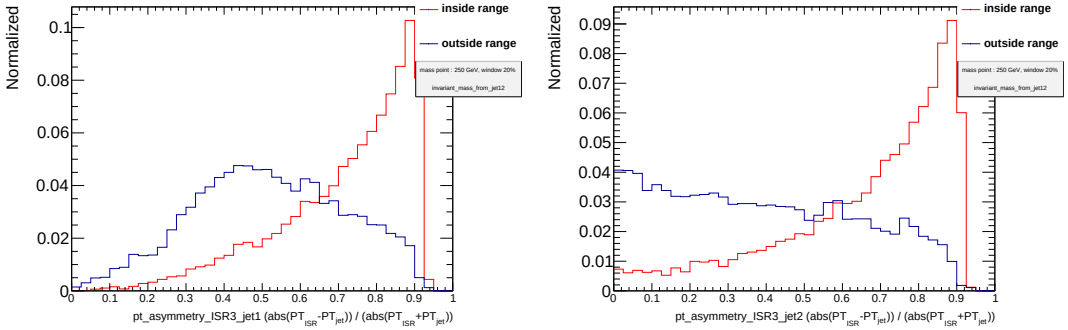
(a) $|\Delta\ y_{jet1jet2}|$ inside and outside mass window

(b) $|\Delta\ y_{jet1jet3}|$ inside and outside mass window



(c) $|\Delta\, y_{jet2jet3}|$ inside and outside mass window

Figure 4.10.: $|\Delta\ y|$ of the events falling inside and outside the interested mass window of 840 - 1260 GeV.

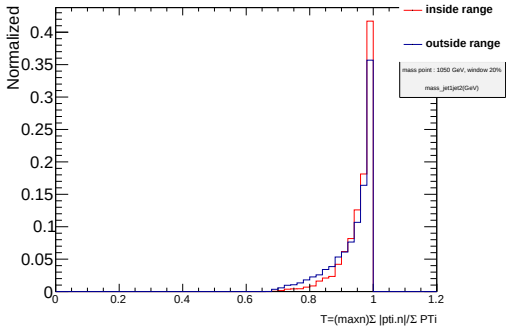


(a) pT_asymmetry$_{jet1jet3}$ inside and outside mass window

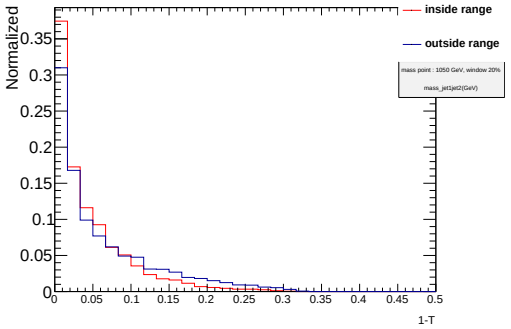
(b) pT_asymmetry$_{jet2jet3}$ inside and outside mass window

Figure 4.11.: p_T_asymmetry of the events falling inside and outside the interested mass window.

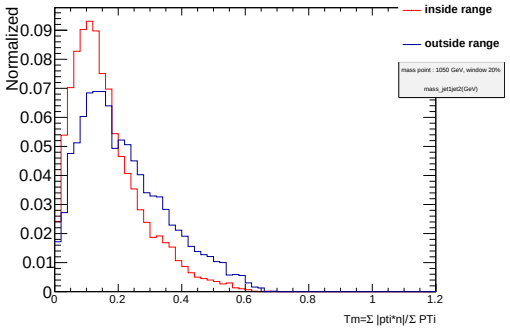
more back-to-back than jets of events outside the window. At high mass, jet3 of events inside the mass window tend to be smaller. The jets of events inside the window are more back-to-back than jets of events outside the window as well. When the mass increases from low to high, the spatial and energy distribution of the jets in an event also moves from being more spread and closer-in-p_T shape to the more back-to-back topology. In the medium mass range, we expect to have a distribution in between these two extremes.



(a) $T_{\perp}$ inside and outside the mass window



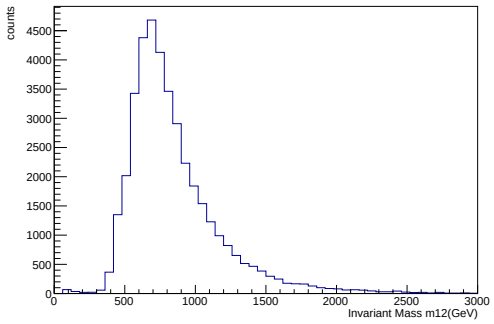
(b) τ = 1 - $T_{\perp}$ inside and outside the mass window



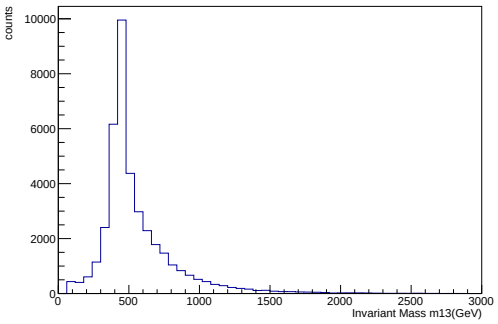
(c) $T_{m,\perp}$ inside and outside the mass window

Figure 4.12.: Thrust related variables of the events falling inside and outside the interested mass window at high mass.

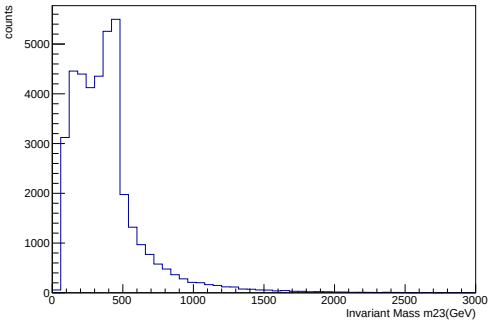
These studies of the event topology help the understanding of the kinematics of the signal samples generated, and verifies the original intuitions. The three leading jets need to be balanced in terms of momentum, and this influences their spatial distributions. The sum of p_T is 0, therefore the p_T in each direction need to balance each other, which is verified here. If the three jets are equally energetic and balance each other, it is difficult to identify the ISR jet based on the jet kinematics alone, and therefore ordering the jets by their p_T to reconstruct the resonance is not useful. For this reason, we move on to ordering the jets differently, based on other kinematic variables.



(a) resonance mass from jet1 and jet2



(b) resonance mass from jet1 and jet3



(c) resonance mass from jet2 and jet3

Figure 4.13.: Masses from jet1 jet2, jet1 jet3 and jet2 jet3 at medium mass of 450 GeV.

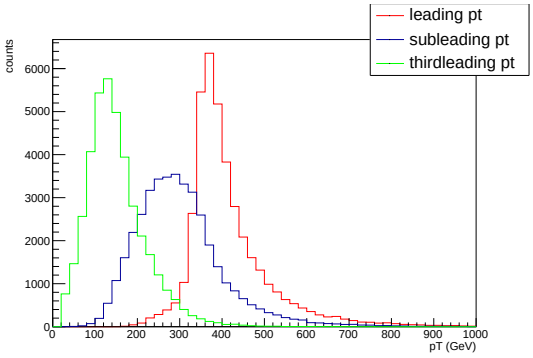


Figure 4.14.: Jet pT distributions for medium mass.

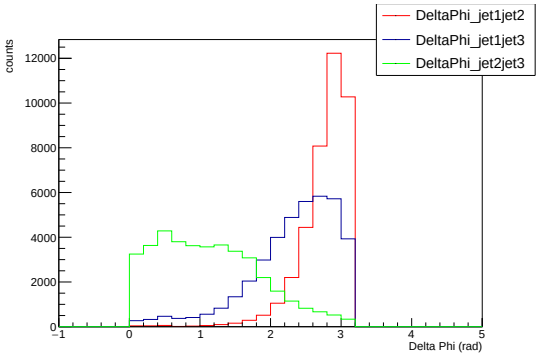
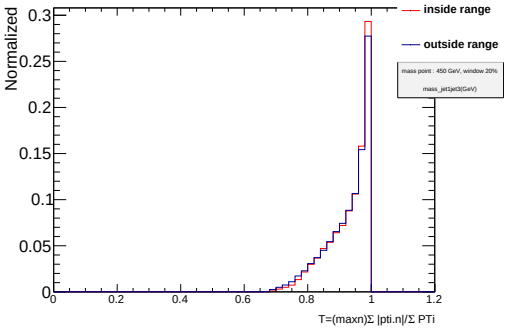
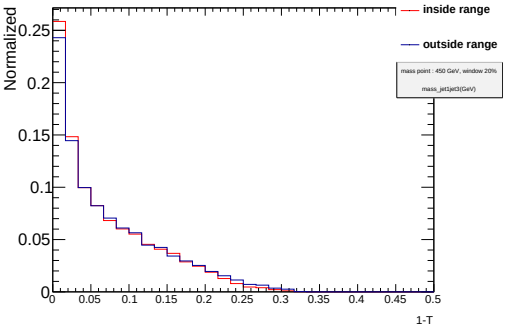


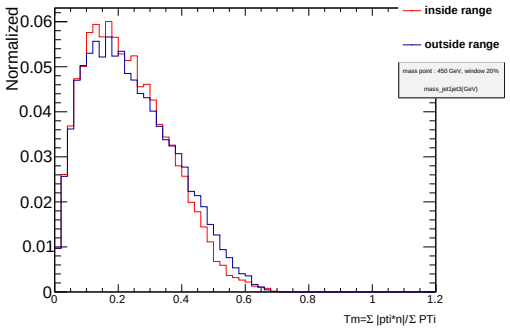
Figure 4.15.: $\Delta\phi$ between jets at medium mass of 450 GeV.



(a) $T_{\perp}$ distributions inside and outside the mass window

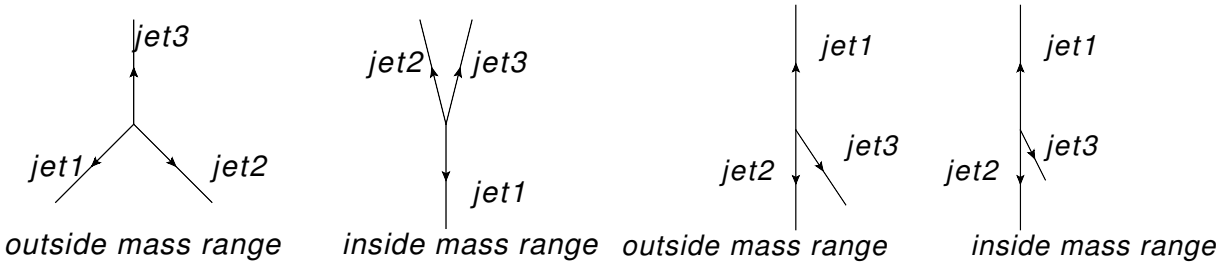


(b) τ = 1 - T_{perp} inside and outside the mass window



(c) $T_{m,\perp}$ inside and outside the mass window

Figure 4.16.: Thrust related variables of the events inside and outside the interested mass at medium mass.



(a) jet topology distribution in one event at 250 GeV. **(b)** jet topology distribution in one event at 1050 GeV.

Figure 4.17.: Topologies when the event falls inside or outside the energy window of 20 % at low and high energy. Plots made with JaxoDraw: [34].

Chapter 5.

New Approaches to Choosing Jets From Resonance

The previous chapters have introduced kinematics variables, presented their distributions and compared different samples. Those studies give a base for understanding the topologies of those samples better. This chapter comes back to the ISR jet samples and we focus on finding a better way to associate the two jets to a resonance that can perform better than simply selecting the second and third leading p_T jets. In this chapter, we show how ordering jets by different variables with respect to the jet p_T influences the sharpness of the signal peak in terms of our figure of merit, the percentage of the events falling within 20% of the signal peak. We also show supporting evidence for including more than three jets in those studies, [39], and include the fourth jet as well.

Due to the y* and p_T selections, the acceptances of different variable-sorting methods would be different. Ideal choices are the ones with better acceptance. The relations between these variables (correlated or anti-correlated or not-related) are also interesting and the results are presented here. From this chapter, most promising variables are chosen and for these background and significance studies are performed. Background and significance studies are presented in chapter 6.

5.1. Peak Sharpness Performances With Three Jets

Firstly, we try to reconstruct the invariant mass of two jets ordered using different kinematic variables with respect to the jet p_T. For example, sorting the jets in the order of $\Delta\phi_{min}$ means that the invariant mass is reconstructed starting from the two jets resulting in the minimal $\Delta\phi$. In the same way, the two jets that are closest to the thrust axis (as introduced in section 3.1) are the thrust_close jets, and the furthest two are the thrust_far jets. As discussed in section 2.5, dijet + ISR samples are used for masses below 600 GeV

and dijet samples are used for mass above 600 GeV. Following the same $\pm$ 20 % window as in section 4.1, a simple way to tell if the chosen jets come from the resonance follows:

1) Calculate the $\pm$ 20 % range of the simulated Z' mass (using the signal sample mass) for each mass points;

2) For each variable ordering chosen, select the two jets (jet_i and jet_j) and calculate their invariant mass;

3) If this mass falls inside the mass range described in 1), add a count to the number of events that falls inside the range;

4) Calculate percentage = (Events falling inside the mass range) / (total event number);

The variables that have been studied and are shown in Figure 5.1 are:

1) DeltaEtamax: the two jets that have the largest $\Delta\eta$ separation;

2) DeltaEtamin: the two jets that have the smallest $\Delta\eta$ separation;

3) DeltaEtamiddle: the two jets that have the $\Delta\eta$ separation between largest and smallest;

4) DeltaRmax: the two jets that have the largest Δ R separation;

5) DeltaRmin: the two jets that have the smallest Δ R separation;

6) Etamin: the two jets that have smallest η separation;

7) ymax: the two jets that have the largest $y^*_{jetijetj}$ separation;

8) ymin: the two jets that have the smallest $y^*_{jetijetj}$ separation;

9) thrustfar: the two jets that are further away from the thrust axis;

10) thrustclose: the two jets that are closer to the thrust axis;

With these criteria, a plot of the performance of each of the variable orderings, in terms of sharpness of the peak, is shown in Figure 5.1, where only three leading jets are considered.

However, according to the studies [39] in Figure 5.2, especially in the intermediate mass points (for example 450 GeV) there can be more than three jets involved in this process, as there is a secondary peak constituted by the first and fourth jet. This happens for example if the ISR jet is split into two secondary jets.

Based on this study, criteria for also including the fourth jet in the plots above are developed as below :

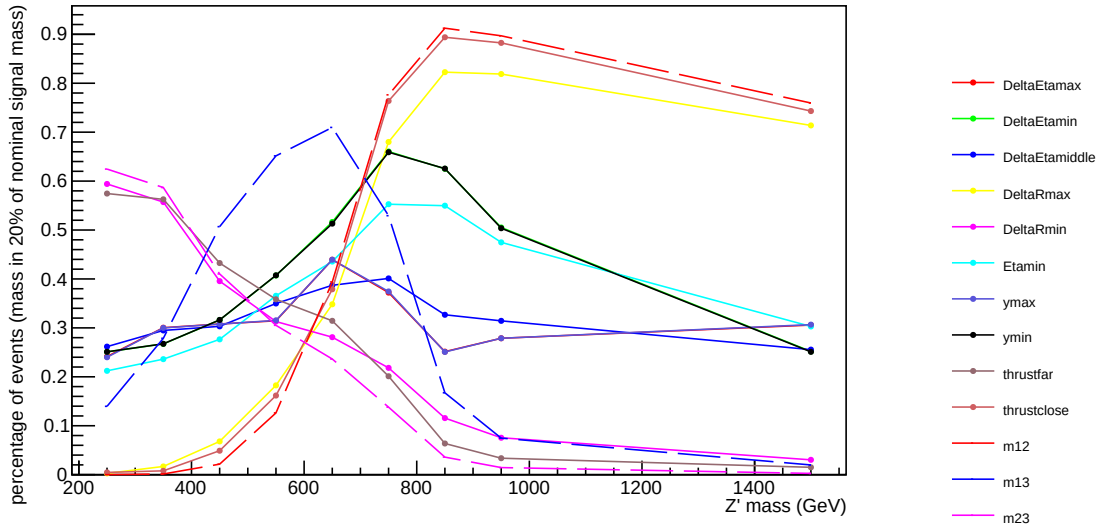


Figure 5.1.: Performance plots in terms of percentage falling within 20% of the peak for various variables. Only the three leading jets are taken into consideration. The x-axis is the simulated mass of the resonance in GeV. Below 600 GeV, dijet + ISR samples are used and above 600 GeV, dijet samples are used.

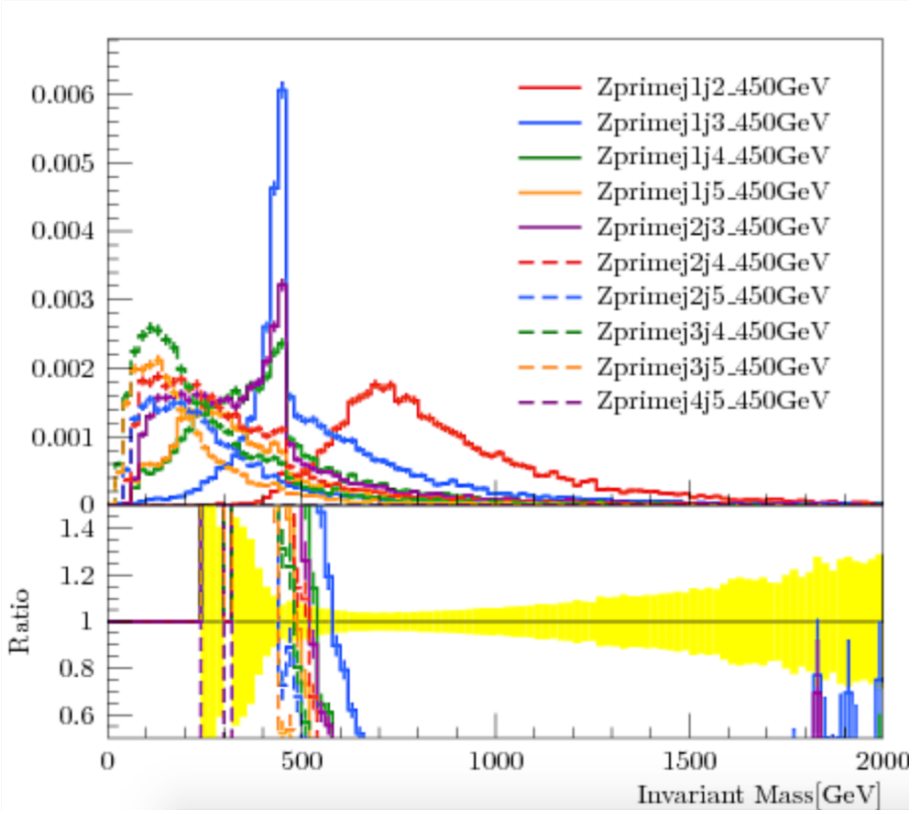


Figure 5.2.: Studies made with more than three jets at 450 GeV Z′ mass. The y-axis is the integrated number of events. Jet1 and jet3 give the best match to 450 GeV as the corresponding peak is at 450 GeV, jet2 and jet3 are the second best but jet1 and jet4 (green line) also give correct mass sometimes by peaking at 450 GeV. Studies taken from [39].

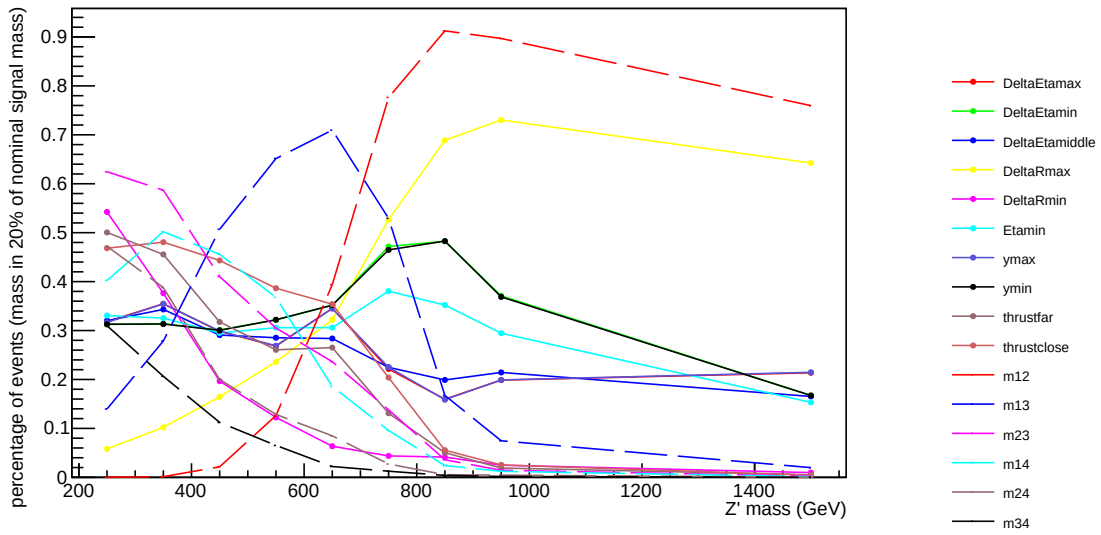


Figure 5.3.: Performance plots in terms of percentage falling within 20% of the peak. Four leading jets are taken into consideration, with the algorithm in the text. The x-axis is the simulated mass of the resonance in GeV. Below 600 GeV, dijet + ISR samples are used and above 600 GeV, dijet samples are used.

1) If there are only three jets in an event, use only the three jets. Apply selection: leading jet $p_T > 420$ GeV, all jet $p_T > 25$ GeV, y* of dijet < 0.8. (same as three jet case)

2) If there are more than three jets in the event, impose that jet4 $p_T > 25$ GeV (otherwise use only the first three jets), and if that condition is passed include the fourth jet into the variable calculation.

When considering the fourth jet, the plot shown in Figure 5.3 differs from Figure 5.1. We note that some variables are behaving differently with respect to the three-jet case and four-jet case, for example, thrust far and thrust close. This is because the fourth jet is involved in the variable calculation, and can be chosen as part of of the dijet used for the invariant mass, changing the behavior of the percentage curve.

In Figure 5.1 and Figure 5.3, the performances of all involved variables at different mass scales are presented. Using jet2 and jet3 for the invariant mass performs best at low mass, while at medium mass, jet1 and jet3 start to take over. At high mass, jet1 and jet2 are most likely to be the best choice. This agrees with previous studies. Apart from what is already known, these plots show that there are some other variables that can give a good performance. Although none of the variables is significantly better from the p_T ordering in any of the mass ranges, it is worth investigating whether there is a variable that is better than a single p_T ordering choice on a larger mass range.

To choose alternative variables for the dijet selection, acceptances and significance performances need to be studied. The acceptance can indicate how much signal is kept after selections. The significance is the value that measures how outstanding the signal

is compared to background. It is also interesting to see the correlations between the variables: as it can be seen from Figure 5.2 and Figure 5.3, some variables are related to each other because they have a similar performance. For example, Delta R Min and thrust far are related to jet2 jet3, and Delta R Max is related to jet1 jet2. A more detailed correlation study is presented in section 5.3. Since we are most interested in the behavior of the first three jets and we have shown that the performance of selecting three or four jets does not show outstanding changes, the acceptance and correlation studies are performed with the three leading jets.

5.2. Acceptances of the Variables

The acceptance of a sample is a value that can indicate how much signal is kept after selections. The acceptance is defined as the ratio of the number of events that pass all the selections over the total number of events in each sample.

In the case of this report, the selections considered are the p_T selection and the y* selection as introduced in chapter 2. For the p_T selection, all the resonance reconstruction methods have the same amount of events ruled out because the p_T magnitudes of each event in one sample do not change. But for y*, since it is a selection on the chosen "dijet". For different resonance reconstruction methods (corresponding to different variables), the "dijets" chosen are different. Therefore, the y* cut selects different numbers of events. It is also mentioned in chapter 2 that dijet + ISR samples are generated with a filter on leading jet p_T, so here the filter efficiency is also taken into consideration in the acceptance to be comparable with the dijet samples. The acceptances are shown in Figure 5.4. For masses lower than 600 GeV, dijet + ISR samples are used and above 600 GeV, dijet samples are used. It can be seen that there is a small difference in the trend between the two kinds of samples, but this is still acceptable to our studies. This is also another way to validate the consistency of the samples. Most of the variables give similar signal acceptances but Delta R max (DeltaRmax) and y* max (ymax) are not performing well as they remove much more signal than the others.

5.3. Correlation Studies

As mentioned in previous sections, some of the variables have similar performances in the plot showing the fraction of events under the peak, which means that they may be correlated.

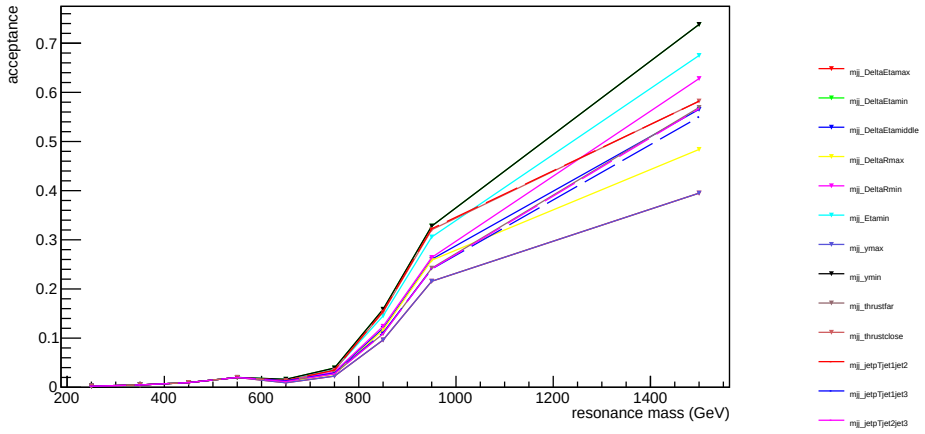
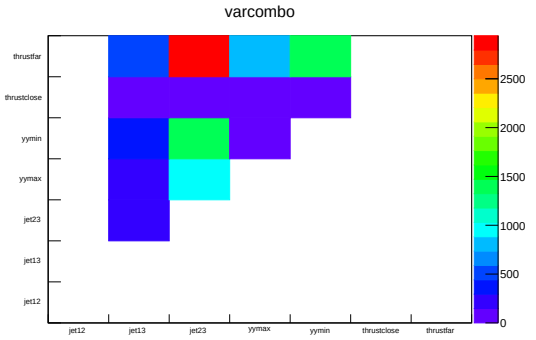


Figure 5.4.: Acceptances for different resonance reconstruction methods.

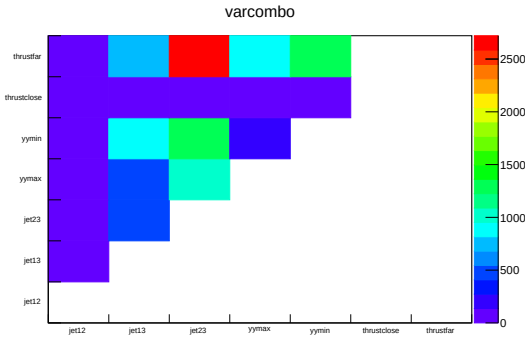
It is important that the chosen variables do not shape the background, as the search strategy looks for a bump over the smoothly falling background. If the variable chosen shapes the background to not look like a smooth falling curve, it would be impossible to locate the signal bump by relying on this. Therefore, we only show the correlation of the variables that are not shaping the background. More background related research is presented in chapter 6.

In this section, the correlation studies between jet1 jet2, jet1 jet3, jet2 jet3, y* max (yymax), y* min (yymin), thrustclose and thrustfar are presented, as they are the ones with the best performance in the lower and medium energy range, which are the ranges we are most interested in for this search. Even though we do not use this information in the following chapters, it can be used in the future if more than one variable at a time are used to select the dijet from the resonance.

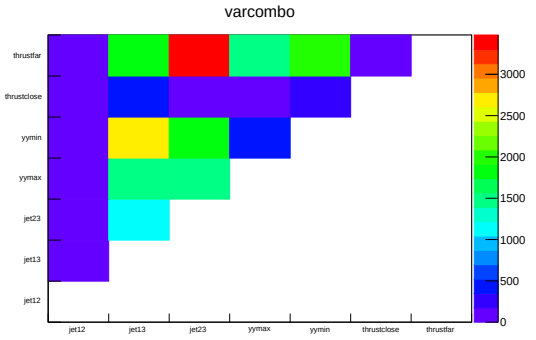
The plots in Figure 5.4 give an indication on the variable correlations. Both x-axis (variable X) and y-axis (variable Y) are the variables we are interested in. If the dijet selected based on X also gives a resonance invariant mass falling inside the $\pm$ 20 % range (meaning that they give the "correct" mass) and the dijet selected based on Y also gives the correct mass, the event would be filled in the bin of (X, Y). The more events in (X, Y) means that the two variables are more correlated. From Figure 5.4, we can see that jet2 jet3 and thrustfar are correlated variables (according to plots at lower masses), jet1 jet3 and yymin (medium masses) are correlated and jet1 jet2 and thrustclose are related (higher masses). This can be understood with the topology studies in chapter 4. At low mass, the jets are spread and jet2 and jet3 are further apart, which means they are further away from the thrust axis. At higher mass, the jets are very "back-to-back" with jet1 and jet2 balancing each other, which means they are closer to the thrust axis.



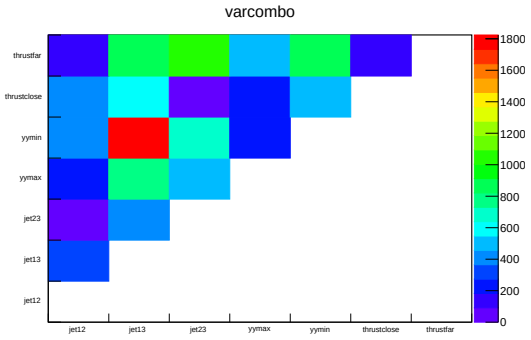
(a) mass point: 250 GeV



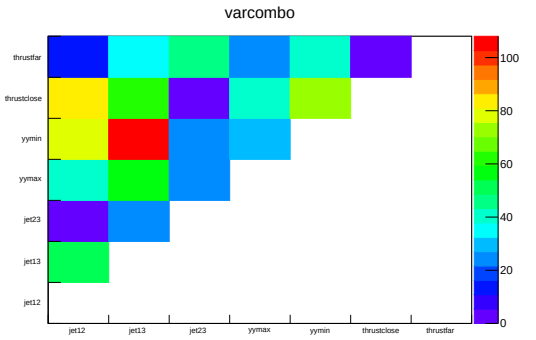
(b) mass point: 350 GeV



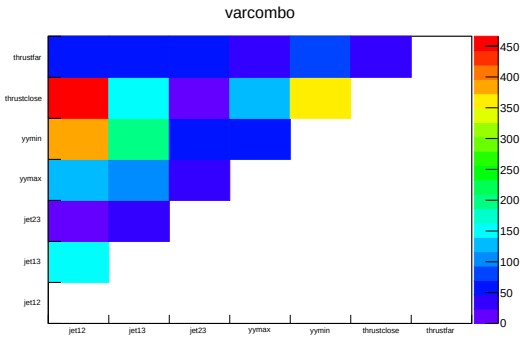
(c) mass point: 450 GeV



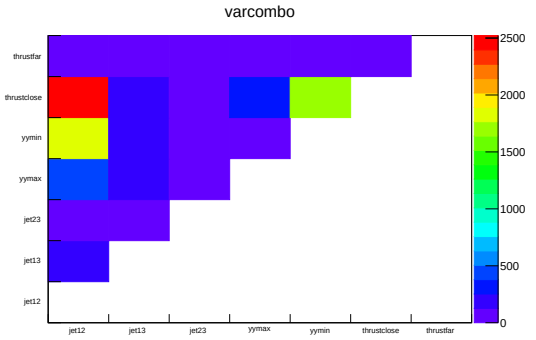
(d) mass point: 550 GeV



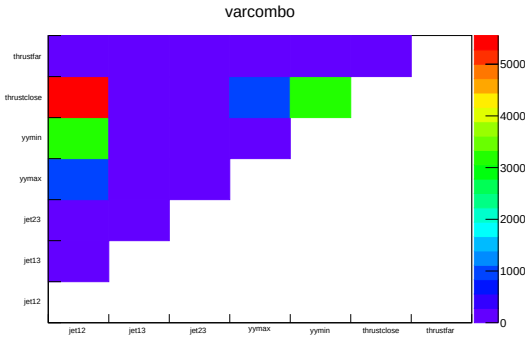
(e) mass point: 650 GeV



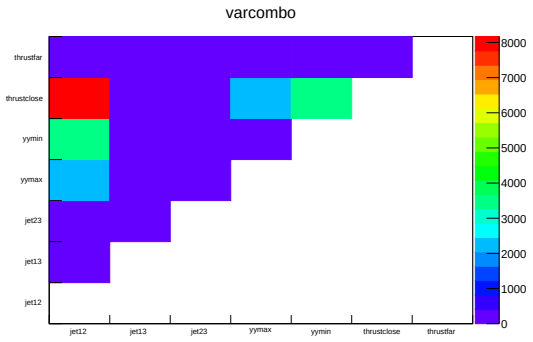
(f) mass point: 750 GeV



(g) mass point: 850 GeV



(h) mass point: 950 GeV



(i) mass point: 1500 GeV

Figure 5.4.: Correlation studies for interested variables from 250 GeV to 1500 GeV, mass under 600 GeV, dijet + ISR samples are used and above 600 GeV, dijet samples are used. More events in the (X, Y) bin means that X and Y are more related. (varcombo for variable combinatorics)

5.4. Conclusions and Discussions

When taking into account only the previous studies on kinematics and topology as in section 4.1 and section 4.2, it is not straightforward to identify which jets can be used to reconstruct the resonance. In this chapter, we have attempted to identify ordering with variables that perform better than the current choice of p_T ordering. Since the parton truth level information that directly tell if a jet comes from the resonance was unavailable, we introduced a criterion to select the "resonance dijet". According to the percentage of events that falls inside the $\pm$ 20 % range of the nominal signal mass, whether a variable is a good candidate for selecting dijets can be verified. The results are summarized in the "peak sharpnenss performance plot", Figure 5.1 and Figure 5.3. We then studied the acceptances of the methods of reconstructing resonance mass based on various variables, Figure 5.4 to see whether the new ordering would allow us to preserve sufficient signal. From the "percentage plots", we also see that some variables are correlated with each other, and Figure 5.4 shows their correlation that can be used in future studies.

The aim of our previous studies was to find a single variable ordering that helps improving the choice of the dijet from the resonance, but there is no variable that performs better than the p_T ordering in the whole mass range. The thrust-related variable orderings have a performance that is worse than the p_T ordering at any given point, but one can use them for a wider mass range.

We now move on to studies of how these variables perform on the background, in the next chapter.

Chapter 6.

Background and Significance Studies

In chapter 5, different ways to reconstruct resonances based on the ordering of the jets have been discussed. We have selected a number of variables that are more promising in terms of performance and acceptance, and now we can proceed to studies of whether these orderings leave the QCD background smooth, and to studies of the signal significance.

So far in this report, only signal samples have been studied and studies on the background are presented in this chapter. Methods applied for resonance mass reconstruction should not shape the background, otherwise it would be impossible to identify the signal in case of real data when signal and background are mixed together. Therefore, background performances for the chosen variables are presented in this chapter. The variables that shape the background curve are not studied for significance.

Significance studies are the way to measure how "significant" a signal is compared to background. If the significance is high, then the signal stands out against the background, making it easier to be noticed. The search would thus have better sensitivity to signals. On the other hand, if the significance is not large enough, the signal can not be distinguished from the background and this would decrease the search sensitivity.

6.1. Background Sample

In this search, the background is from the very frequent QCD processes, simulated with Pythia 8. These samples are at the truth level as introduced in section 2.4. They have the final-state product particles but they have not gone through the detector yet. Sample generation details are given in [33].

Figure 6.1e is a background mass curve constructed with jet2 jet3. As the mass increases, the curve starts smoothly falling since there are less events with high energy. The dijet invariant mass distribution for QCD should be a smooth falling curve. The point where the curve starts falling smoothly is the "turn-on" point. Only after this point can statistical

analysis methods be applied to the background curve. Masses below the "turn-on" point are not accessible to studies. Therefore, a lower "turn-on" point is preferred in our study.

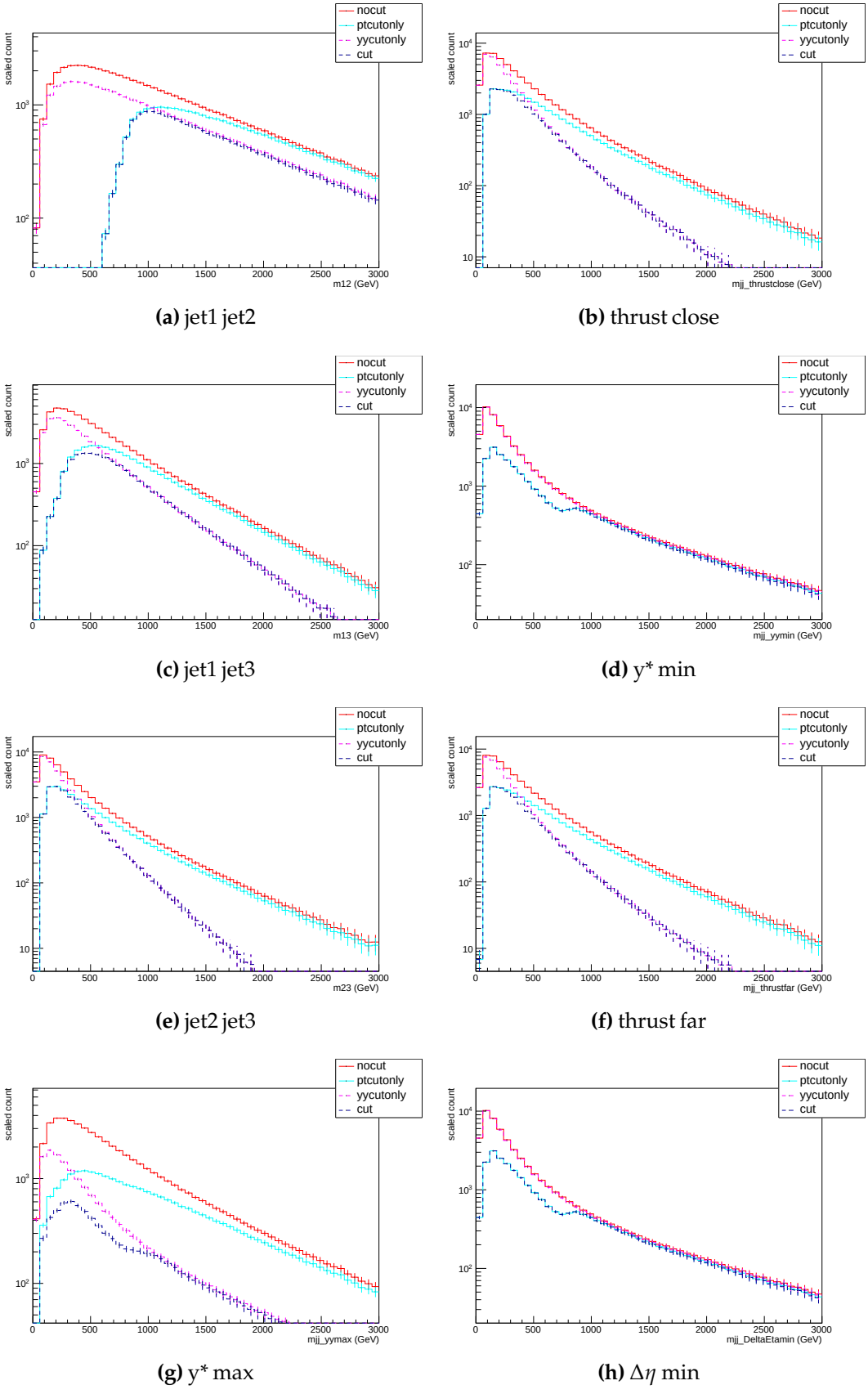
6.2. Background Studies

The background needs to go through exactly the same selections as the signals do, because in data one cannot distinguish between signal and background events and therefore the whole data spectrum has to be statistical analyzed to pick out the signal bump. The previous selections on p_T and y* are also applied here: 1) jet1 $p_T > 420$ GeV, all jet $p_T > 25$ GeV, if there is a fourth jet and its p_T above 25 GeV, we include it in the variable calculation and dijet selection. 2) y* of dijet < 0.8. This is the same selection as in chapter 5 for the four-jet selections. To see the effect of each selection step, the curves after each individual selection are overlaid with the curve before any cut (no cut) and the curve after all cut (cut). Similar studies are also conducted in the three-jet case, giving very similar results in terms of background. Therefore, only the four-jet case results are presented here in Figure 6.1 The variables shown are: jet1 jet2 (m12), jet1 jet3 (m13), jet2 jet3 (m23), thrust close (thrustclose), thrust far (thrustfar), y* min (yymin) and y* max (yymax) and $\Delta\eta$ min (DeltaEtamin).

As it can be seen from Figure 6.1, the background mass curve generated with y* max and y* min ordering are not smooth after the p_T selections. This is because requiring a high leading jet p_T also changes the jet topology distribution, and therefore influences the background curve after y* selections. This background curve shaping makes it difficult to select signals from background when signals and background are overlaid. Therefore, y* max and y* min can not be chosen as ordering variables to optimize the search sensitivity.

Since we are interested in a mass range that is as low as possible, we prefer that the background curve starts to be smooth from as low as possible in terms of invariant mass. According to this criterion, jet2 and jet3, as well as thrust close, are chosen as the best performing variables by comparing the curves in Figure 6.1.

To further compare the performance of jet2 jet3 and thrust close, Figure 6.2 is made. This plot shows that jet2 jet3 and thrust close are having similar performances. However, it also shows that the curve for jet2 jet3 starts falling smoothly two bins in the histogram (around 100 GeV) before the thrust close curve, which indicates that jet2 jet3 is still a better choice in terms of background. The background is not smooth until 350 GeV, which means that we would not be able to use the current background estimation technique until at least the invariant mass of 350 GeV. Nevertheless, we move on to comparing the signal-to-background significance for these variables.



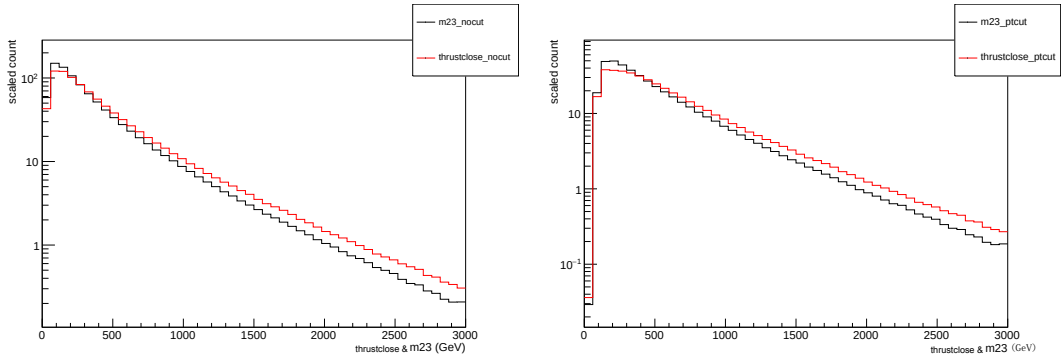
(a) jet1 jet2 **(b)** thrust close

(c) jet1 jet3 **(d)** y* min

(e) jet2 jet3 **(f)** thrust far

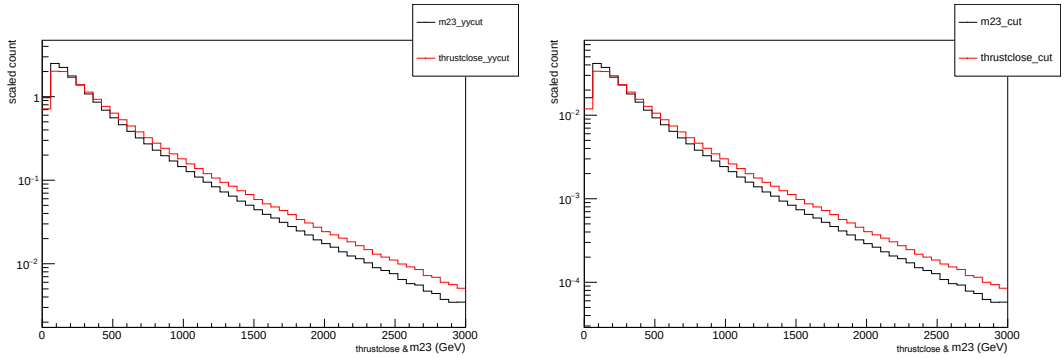
(g) y* max **(h)** $\Delta\eta$ min

Figure 6.1.: Background mass curves with different mass reconstruction variables. Different cuts specified.



(a) thrust close and jet2 jet3 comparison, no cut applied.

(b) thrust close and jet2 jet3 comparison, after p_T cut



(c) thrust close and jet2 jet3 comparison, after y* cut

(d) thrust close and jet2 jet3 comparison, after both cuts

Figure 6.2.: Comparisons on jet2 jet3 and thrust close background smoothness.

6.3. Significance Studies

The statistical significance as defined in Ref. [40] is a quantitative way of describing how signals "stand out" above the background when the events follow a Poisson distribution. This is the same significance calculation method used in dijet + ISR search before [27].

In the search discussed in this report, the background and the signals are not generated and scaled with the same technique. In order to combine them together, the weight information in the samples need to be taken into consideration. From the cross section (σ), the number of events detected (N) over a certain time (t), the luminosity can be defined as Equation 6.1. Integrated luminosity, as in Equation 6.2, is the luminosity integrated over time. Therefore the total number of events can be obtained by Equation 6.3.

$$Luminosity: L = \frac{1}{\sigma}\frac{dN}{dt} \tag{6.1}$$

$$Integrated luminosity: L_{int} = \int L dt \tag{6.2}$$

$$N = L_{int} * \sigma \tag{6.3}$$

The weight information describes how much importance each event has. In the MC processes, the weights of each simulated event for the signal and background samples are different. In order to combine the signal with background we need to scale the signal and background to the same weight, taking into account their respective cross sections.

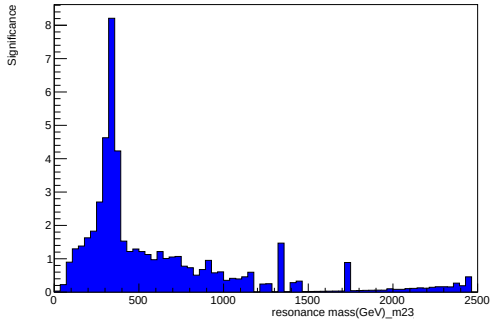
We normalize the signal and background samples to the total LHC luminosity collected in 2015, 2016 and in the first part of 2017, amounting to a total of 76.521 fb^{-1}.

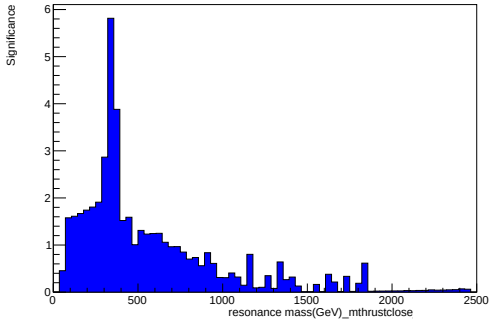
We then assume that the MC statistics is representative of the data statistics, and set errors for each bin in signal and background histograms to$\sqrt{N}$, N being the number of events in that bin [1]. The samples used in this study are simulated samples and they have weight information from the generation. However, the experimental data does not come with weight. The experimental data comes event by event from a Poisson-dsitributed process. In order to enable the simulated samples (signal and background) to have a similar statistical error as the data, similar Poisson distributions need to be added to the simulation samples. This is achieved by adding the Poisson noises bin by bin in the simulation sample histograms. Then the "data" samples are created with signals overlaid with background, by adding the histograms. Signal samples used under 600 GeV are dijet + ISR samples and above 600 GeV, dijet samples are used.

Using those samples, the significance is calculated in the manner defined in [40], with the "signed z-value" method. The p-value describes the probability of finding a deviation at least as large as observed in data, when the chosen theoretical model describes the system. Often the p-value is converted to z-value, which is the deviation at the right of the mean of a Gaussian distribution, in the units of standard deviations, and correspond to the same p-value [40]. The z-values for the chosen variables (jet2 jet3 and thrust close as mentioned in section 6.2) from resonance mass 350 GeV to 850 GeV are plotted as significances in Figure 6.2.

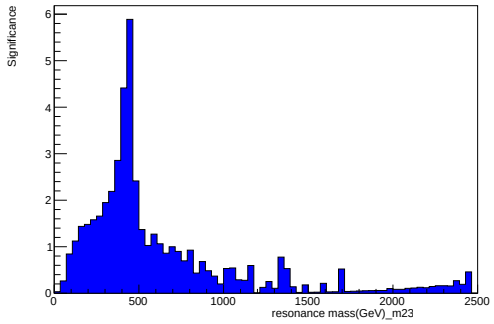
As shown in Figure 6.2, the significance performances of jet2 jet3 and thrust close is very similar. However, the significance is slightly less for thrust close than it is for jet2 jet3, as expected from the lower performance of this variable in Figure 5.3. There exist some high significance bins above 1 TeV, caused by the poor MC statistics for samples above 600 GeV. The cut on leading jet p_T to be above 420 GeV removed most of the signal for dijet samples, as discussed in section 2.5, resulting in poor statistics.

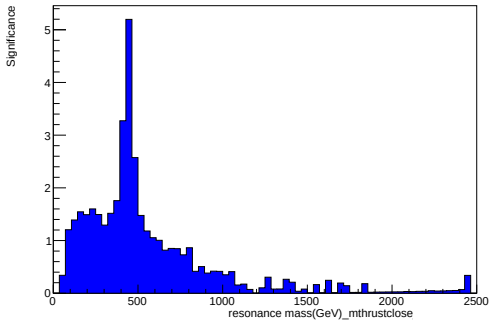
[1] This is not the case for some of the signal samples with large cross-section because not enough events were generated, but we leave a deeper investigation to further studies



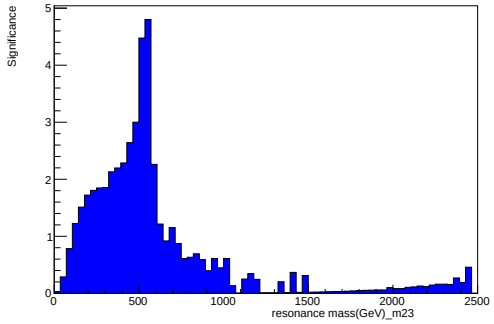


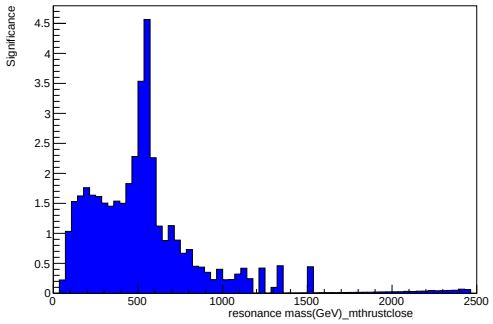
(a) Significance for jet2 jet3 at Z′ mass 350 GeV **(b)** Significance for thrust close at Z′ mass 350 GeV





(c) Significance for jet2 jet3 at Z′ mass 450 GeV **(d)** Significance for thrust close at Z′ mass 450 GeV

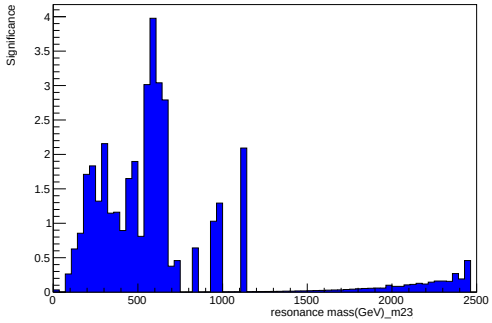




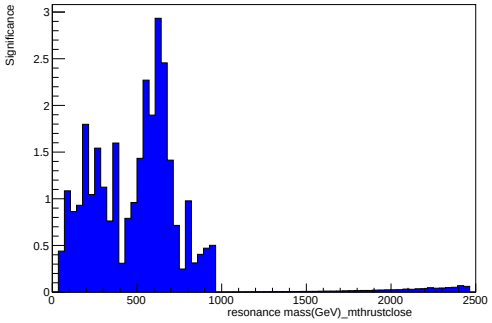
(e) Significance for jet2 jet3 at Z′ mass 550 GeV **(f)** Significance for thrust close at Z′ mass 550 GeV

6.4. Conclusions and Discussions

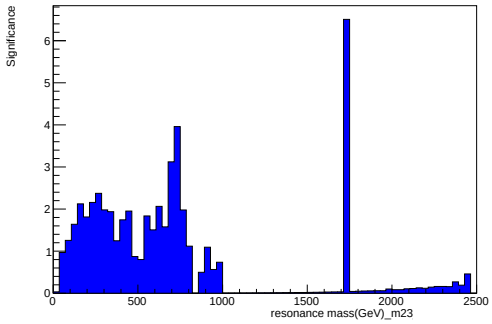
The background and significance studies in this chapter present more evidence to "which variable ordering is the best choice for the resonance dijet selection". The results confirm that jet2 jet3 is still the best pairing option at lower masses. At intermediate mass, thrust close is also a usable candidate for dijet selection, but its performance is not better than jet2 jet3. Thrust close can not be expected to be better that jet1 jet3 or jet1 jet2 from medium mass to high mass either. This is the result that can be expected from the performances of thrustclose, thrustfar and pT-ordered jet variables in the peak sharpness performance



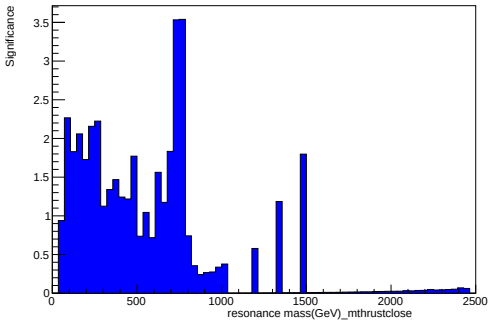
(g) Significance for jet2 jet3 at Z′ mass 650 GeV



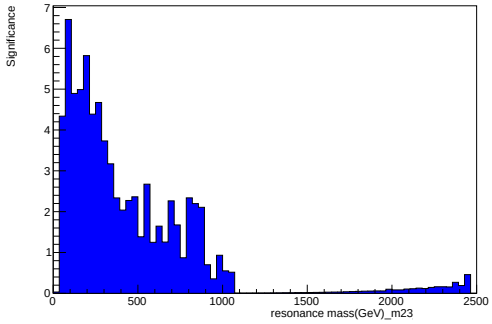
(h) Significance for thrust close at Z′ mass 650 GeV



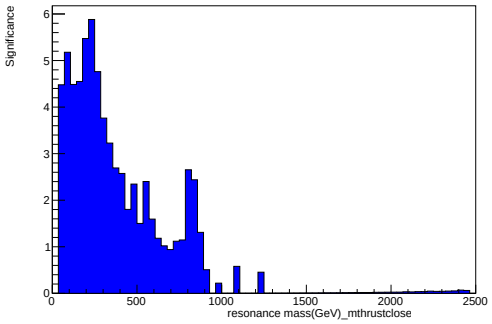
(i) Significance for jet2 jet3 at Z′ mass 750 GeV



(j) Significance for thrust close at Z′ mass 750 GeV



(k) Significance for jet2 jet3 at Z′ mass 850 GeV



(l) Significance for thrust close at Z′ mass 850 GeV

Figure 6.2.: Significance distributions for jet2 jet3 and thrust close from 350 GeV to 850 GeV.

plots (Figure 5.1 and Figure 5.3) in chapter 5. Some other promising variables, for example y* min or $\Delta\eta$ min, can not be chosen for selection because they shape the background.

Therefore, the dijet pairing method should be kept to jet2 jet3 from low mass to medium mass. We can also apply different ordering methods in different mass region, but this idea is not straightforward because there could be bumps at the connection points of mass regions in background curve. The difficulty of keeping the background smooth throughout the mass region from low to high would be largely increased by this. The same can

be said about using multiple variables at the same time to perform dijet selections. As a result, only single variable situations are studied in this work.

Chapter 7.

Summary and Outlook

The main question of this study is "How to identify the resonance dijet and the ISR jet in the dijet + ISR search"? To answer this question, different attempts are made throughout the study, and these studies further our understandings not only on this question alone but also in other aspects of this search, from sample characteristics to event topology. The overall conclusions are stated in this final chapter as well as some outlook on this project.

7.1. Summary

For the light resonance search for DM mediators, the mediator that couples with DM particles comes in form of a light resonance created in association with a radiation. The radiation (noted as "ISR" in this report but in principle the ISR is not distinguishable from the FSR) can be either a photon or a jet. The light resonance would decay into a dijet. This study concentrates on dijet + ISR jet search and the methods to choose the dijet from the three involved jets. The main background of this process comes from the QCD processes that happen at the same time in the collision. In order to identify signals from background, the signal selections are expected to give a resonance mass distribution that is sharp and outstanding from the background. This requires choosing the resonance dijet correctly. Otherwise, it would be difficult to tell signals from the background. The main purpose of this study is to find one best variable for dijet selection.

This report starts from validating the signal samples used in the dijet + ISR search, as there are two different kinds of samples involved in this process, the dijet + ISR samples and the dijet samples. The results in chapter 2 verify that the two samples are comparable and can be used together. In the following chapters, further studies are conducted on attempts to directly identify the ISR jet, firstly through comparing ISR jet and ISR photon samples as detailed in chapter 3. But this comparison is not effective because of the differences in the central sample generation, therefore, further studies only use the ISR jet samples. The kinematics and combinatorics distributions of the ISR jet samples are

detailed in chapter 4, leading to a deeper understanding in the topologies of ISR jet signals but still no indications on how to identify the ISR jet. chapter 5 is a summary of the variables and their performances when they are used to order the jets to reconstruct the resonance peak. The peak sharpness performance plots in chapter 5 guide the studies further to which variables are most promising and for which we go on to study background smoothness and significance. Background and significance studies in chapter 6 serve as finalizations of which variable serves the purpose best, especially at the resonance mass range of below 800 GeV, which is of interest to us. The result is that jet2 jet3 works best yet thrust close can also give an acceptable performance.

With the percentage plot and kinematics distributions, other results can be verified with them. The results we see so far, have been self-consistent.

7.2. Outlook

This report presented studies on single-variable identifications on the ISR jet in a dijet + ISR search. If I am to give some indications on how to proceed from here, there are several paths that can be taken. Some I have tried, others are still open questions.

1) If using different variables to select jets does not help much with the situation, will selecting which events to use from the start help?

This is a question that I have tried to answer. Two different approaches were taken. The first is only to choose samples that are more aligned. By saying "more aligned" means that selecting samples with higher $T_{\perp}$ value or lower $T_{m,\perp}$ value. This can help a tiny bit with the resonance mass distribution.

Another approach that I have tried taking is to select the mass that is "less relevant" to jet p_Ts. This approach is based on the idea that the simulated resonance has a certain mass regardless of what the jet p_Ts are. If cuts can be made to select events that have "less-likely" p_Ts but give the same resonance mass nevertheless, this can also point to successful resonance reconstructions. This method is detailed in [41].

2) If using single variables does not help much, what about multiple variables?

As discussed in the end of chapter 5, even a single variable can cause bumps and dips in the background curve, more variables would make the situation even more complicated. Multiple variables, or even different variables at different mass ranges, would add unexpected and unwanted features to the background, thus making them invalid to use. Therefore, some advice from me would be that the first thing to try with multiple variable methods is to keep the background smooth and under control. Or one can take the approach of machine learning. The problem with machine learning is that it is more

like a black-box process. The outcome of machine learning processes might be that we are actually optimizing for something else other than the radiation jet or dijet identifications, but we would not know of this. It can cause problems in the long run. The methods in this study are old-school and sometimes too obvious, and this is the way of understanding what we are asking for. But indeed, machine learning does have a chance in solving this problem in a different manner. One can try if interested.

Overall, there are different paths to take from where this study ends. The methods and variables discussed here might also be of use in other similar studies.

Appendix A.

Coding Documentations

This chapter serves as the documentations of the codes developed for this thesis.

The codes are all open-sourced and published at https://gitlab.cern.ch/atlas-phys-exotics-dijetisr/Combinatorics/tree/master/Zhiying_code under the /Zhiying_code folder. Please refer to README files in each folder for further information on the codes.

For the cross sections and filter efficiencies used in the code, for dijet + ISR samples they are recorded as : https://github.com/UCATLAS/xAODAnaHelpers/blob/54d7baeac9475cf43b952cccdbb838f81120a858/data/metadata/dijetisr_crosssections_13TeV.txt

For dijet samples, the same information is kept in : https://svnweb.cern.ch/trac/atlasphys-exo/browser/Physics/Exotic/JDM/DiJet/RunII/DijetResonanceAlgo/branches/DijetResonanceAlgo-01-00-02-branch/data/XsAcc_13TeV.txt

Colophon

This thesis was made in LaTeX 2_ε [42] using the “hepthesis” class [43] and "savetrees" package [44].

Bibliography

[1] L. Evans and P. Bryant, "Lhc machine," *Journal of instrumentation*, vol. 3, no. 08, p. S08001, 2008.

[2] G. Aad, J. Butterworth, J. Thion, U. Bratzler, P. Ratoff, R. Nickerson, J. Seixas, I. Grabowska-Bold, F. Meisel, S. Lokwitz, *et al.*, "The atlas experiment at the cern large hadron collider," *J. Instrum.*, vol. 3, p. S08003, 2008.

[3] B. R. Martin and G. Shaw, *Particle physics*. John Wiley & Sons, 2017.

[4] G. Kane, *Modern Elementary Particle Physics: Explaining and Extending the Standard Model*. Cambridge University Press, 2017.

[5] L. Jonsson, "Lecture in particle physics."

[6] S. Fukuda, Y. Fukuda, M. Ishitsuka, Y. Itow, T. Kajita, J. Kameda, K. Kaneyuki, K. Kobayashi, Y. Koshio, M. Miura, *et al.*, "Determination of solar neutrino oscillation parameters using 1496 days of super-kamiokande-i data," *Physics Letters B*, vol. 539, no. 3-4, pp. 179–187, 2002.

[7] R. N. Mohapatra and G. Senjanović, "Neutrino masses and mixings in gauge models with spontaneous parity violation," *Physical Review D*, vol. 23, no. 1, p. 165, 1981.

[8] MissMJ, "Standard model of elementary particles.svg."

[9] J. Bernstein, "Spontaneous symmetry breaking, gauge theories, the higgs mechanism and all that," *Reviews of modern physics*, vol. 46, no. 1, p. 7, 1974.

[10] D. M. Wittman, J. A. Tyson, D. Kirkman, I. Dell'Antonio, and G. Bernstein, "Detection of weak gravitational lensing distortions of distant galaxies by cosmic dark matter at large scales," *Nature*, vol. 405, no. 6783, p. 143, 2000.

[11] G. Bertone, D. Hooper, and J. Silk, "Particle dark matter: Evidence, candidates and constraints," *Physics Reports*, vol. 405, no. 5-6, pp. 279–390, 2005.

[12] A. Floderus, *Luminosity determination and searches for supersymmetric sleptons and gauginos at the ATLAS experiment*. PhD thesis, Lund University, 2014.

[13] F. Fayette, "Strategies for precision measurements of the charge asymmetry of the

w boson mass at the lhc within the atlas experiment," *arXiv preprint arXiv:0906.4260*, 2009.

[14] G. Aad, B. Abbott, J. Abdallah, A. Abdelalim, A. Abdesselam, O. Abdinov, B. Abi, M. Abolins, H. Abramowicz, H. Abreu, *et al.*, "Studies of the performance of the atlas detector using cosmic-ray muons," *The European Physical Journal C*, vol. 71, no. 3.

[15] P. Pasuwan, "Searches for low-mass dijet resonances in trijet final states using monte carlo samples," 2016.

[16] JabberWork, "Pseudorapidity2.png."

[17] G. P. Salam, "Towards jetography," *The European Physical Journal C*, vol. 67, no. 3-4, pp. 637–686, 2010.

[18] S. D. Ellis and D. E. Soper, "Successive combination jet algorithm for hadron collisions," *Physical Review D*, vol. 48, no. 7, p. 3160, 1993.

[19] L. K. Bryngemark, *Search for new phenomena in dijet angular distributions at* $\sqrt{s}$*= 8 and 13 TeV*. Springer, 2017.

[20] A. Schwartzman, "Jet energy calibration at the lhc," *International Journal of Modern Physics A*, vol. 30, no. 31, p. 1546002, 2015.

[21] A. Hrynevich, "Atlas jet and missing energy reconstruction, calibration and performance in lhc run-2," *Journal of Instrumentation*, vol. 12, no. 06, p. C06038, 2017.

[22] T. Sjöstrand, "Monte carlo tools," tech. rep., Lund University, 2009.

[23] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, "An introduction to pythia 8.2," *Computer physics communications*, vol. 191, pp. 159–177, 2015.

[24] G. Marchesini, B. R. Webber, G. Abbiendi, I. Knowles, M. H. Seymour, and L. Stanco, "Herwig 5.1-a monte carlo event generator for simulating hadron emission reactions with interfering gluons," *Computer Physics Communications*, vol. 67, no. 3, pp. 465–508, 1992.

[25] 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," *Journal of High Energy Physics*, vol. 2014, no. 7, p. 79, 2014.

[26] S. Agostinelli, J. Allison, K. a. Amako, J. Apostolakis, H. Araujo, P. Arce, M. Asai, D. Axen, S. Banerjee, G. . Barrand, *et al.*, "Geant4 simulation toolkit," *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 506, no. 3, pp. 250–303, 2003.

[27] the ATLAS collaboration, "Search for new light resonances decaying to jet pairs and produced in association with a photon or a jet in proton-proton collisions at $\sqrt{s}$= 13 tev with the atlas detector," ATLAS-CONF-2016-070, 2016.

[28] D. Abercrombie, N. Akchurin, E. Akilli, J. A. Maestre, B. Allen, B. A. Gonzalez, J. Andrea, A. Arbey, G. Azuelos, P. Azzi, *et al.*, "Dark matter benchmark models for early lhc run-2 searches: report of the atlas / cms dark matter forum," *arXiv preprint arXiv:1507.00966*, 2015.

[29] M. Chala, F. Kahlhoefer, M. McCullough, G. Nardini, and K. Schmidt-Hoberg, "Constraining dark sectors with monojets and dijets," *Journal of High Energy Physics*, vol. 2015, no. 7, p. 89, 2015.

[30] M. Nessi, M. Nordberg, P. Jenni, and K. Smith, "Atlas high-level trigger, data-acquisition and controls: Technical design report," tech. rep., ATLAS-TDR-016, 2003.

[31] the ATLAS collaboration, "Search for light dijet resonances with the atlas detector using a trigger-level analysis in lhc pp collisions at $\sqrt{s}$= 13 tev," in *ATLAS Conference Report ATLAS-CONF-2016-030*, 2016.

[32] the ATLAS Collaboration, "Search for low-mass dijet resonances using trigger-level jets with the atlas detector in pp collisions at $\sqrt{s}$= 13 tev," *arXiv preprint arXiv:1804.03496*, 2018.

[33] J.Alison, A.Boveia, *et al.*, "Search for low mass di-jet resonances using proton-proton collisions at $\sqrt{s}$ = 13 tev with the atlas detector," 2016.

[34] D. Binosi and L. Theussl, "Jaxodraw: A graphical user interface for drawing feynman diagrams," *Computer Physics Communications*, vol. 161, no. 1-2, pp. 76–86, 2004.

[35] I. Antcheva, M. Ballintijn, B. Bellenot, M. Biskup, R. Brun, N. Buncic, P. Canal, D. Casadei, O. Couet, V. Fine, *et al.*, "Root c++ framework for petabyte data storage, statistical analysis and visualization," *Computer Physics Communications*, vol. 182, no. 6, pp. 1384–1385, 2011.

[36] D. Krohn, L. Randall, and L. Wang, "On the feasibility and utility of isr tagging," *arXiv preprint arXiv:1101.0810*, 2011.

[37] the ATLAS collaboration, "Measurement of event shapes at large momentum transfer with the atlas detector in pp collisions at $\sqrt{s}$ = 7 tev," 2012.

[38] A. Boveia, "code for thrust-related variable calculations." private communication, 2017.

[39] A. Tudorache, A. Jinaru, and V. Tudorache, "Some ideas about combinatorics." private communication, 2018.

[40] G. Choudalakis and D. Casadei, "Plotting the differences between data and expectation," *The European Physical Journal Plus*, vol. 127, no. 2, p. 25, 2012.

[41] S. Chatrchyan, V. Khachatryan, A. M. Sirunyan, A. Tumasyan, W. Adam, T. Bergauer, M. Dragicevic, J. Erö, C. Fabjan, M. Friedl, *et al.*, "Search for three-jet resonances in pp collisions at $\sqrt{s}$= 7 tev," *Physical review letters*, vol. 107, no. 10, p. 101801, 2011.

[42] L. Lamport, *LaTeX: A Document Preparation System, 2/e*. Pearson Education India, 1994.

[43] A. Buckley, "hepthesis v1. 4," 2010.

[44] S. Pakin, "The savetrees package," 2016.

List of figures