Stony Brook University

Searches for Beyond the Standard Model τ jet at the proposed EIC detector

A Thesis Presented

by

Raghav Kunnawalkam Elayavalli

to

The Graduate School

in Partial Fulfillment of the Requirements

for the Degree of

Master of Arts

in

Physics

Stony Brook University

December 2012

Stony Brook University

The Graduate School

Raghav Kunnawalkam Elayavalli

We, the thesis committee for the above candidate for the Master of Arts degree, hereby recommend acceptance of this thesis.

Abhay Deshpande – Thesis Advisor
Professor, Department of Physics and Astronomy

Klaus Dehmelt – Chairperson of Defense
Research Scientist, Department of Physics and Astronomy

Dmitri Kharzeev
Professor, Department of Physics and Astronomy

Matthew Dawber
Professor, Department of Physics and Astronomy

This thesis is accepted by the Graduate School.

Charles Tabler
Dean of the Graduate School

Abstract of the Thesis

Searches for Beyond the Standard Model τ jet at the proposed EIC detector

by

Raghav Kunnawalkam Elayavalli

Master of Arts

in

Physics

Stony Brook University

2012

One of the measurements proposed at the Electron Ion Collider (EIC) detector is the search for lepton flavor violation. This is a Beyond the Standard Model (BSM) process in which the electron and the quark from the proton form an intermediate boson called the Leptoquark (LQ) which in turn decays into another lepton and quark. In the case of the final lepton being a τ, we are faced with a situation in which we have an electron transforming into a τ, violating lepton flavor conservation in the process.The EIC Detector, if designed appropriately, could have the ability to check if such a LQ state exists. A recent study showed that the ratio of the width of the τ jet to the DIS jet at the generator level is about half when taking into account the characteristic 3 pion decay of the τ. In this report we find that the relation observed above does not persist in the detector when only taking the response of the calorimeters into account. The detector simulation and analysis have been performed using a framework called Fairroot.

To my dear brother K E Srivatsav who inspires me to succeed everyday.

Contents

List of Figures

List of Tables

Acknowledgements

There are several people who have made this masters thesis possible. Since they are numerous in number and it is not possible to mention them all here, I would group them into two groups which are family and work.

I shall always be indebted to my family for giving me the freedom and opportunity to choose my path in life. A triangle is often symbolized as the most stable configuration in the universe and the main reasons for my stability are the love and care showered by my mother, logic and rationality rendered by my father, and a drive to excel that my inspiring brother arms me with.

This masters thesis work has been performed under the guidance of Prof. Abhay Deshpande and Dr. Klaus Dehmelt. I came across Prof. Deshpande during my seminar presentation on the spin of the proton. I was intrigued by the topic and I started working with him after talking to him a couple of times. I still remember him telling me during one of our first meetings 'It does not matter what you have to do to get a better understanding of physics, you should be ready to even do ballet' when I said that I did not like to program in C++. That sentence alone changed my entire outlook to my research and helped me to become the person I am now.

I would like to thank the Fairroot developers team at GSI, Damstadt in Germany for their help during the initial stages of software development.

Several Professors/Teachers from my student life are responsible for shaping my knowledge and passion in physics. Prominent among them are Mr. Sadagoppan Rajesh, Dr. T.R.Subramaniam (India), Dr. Raghavan (India), Dr. Derin Sherman (Cornell College), Dr. Jim Freeman (Cornell College), Dr. Abhay Deshpande (Stony Brook), Dr. Dmitri Kharzeev (Stony Brook) and Dr. Ismael Zahed (Stony Brook).

Amongst my peers, Yakov Kulinich and Will Foreman stand out for helping me with programming and for understanding the initial process of producing Monte Carlo data. Raphael Cervantes, for helping me better understand my software.

Sensei Nick Panebianco at VMA Setauket deserves special mention for teaching me that an active and healthy mind always goes with an active and

healthy body. After a long hard day at the office, his classes in Kempo Jiu Jitsu have always cleared my mind of the stress I had and helped me to achieve peace at heart.

Finally I would like to thank God Almighty for her belief and support in my activities.

Chapter 1

EIC - what and why?

The Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Lab (BNL), NY is one of only two operating heavy-ion colliders, and the only spin-polarized proton collider ever built. By using RHIC to collide gold ions traveling at relativistic speeds, physicists study the primordial form of matter that existed in the universe shortly after the Big Bang, the quark gluon plasma. By colliding spin-polarized protons, the spin structure of the proton is explored.

1.1 Introduction

The Electron Ion Collider (EIC)[2] is primarily a QCD machine with the ability to provide some answers in the Electro-Weak (EW) and Beyond the Standard Model (BSM) realms of physics. It could collide electrons having energy from $5-30$ GeV with protons having energy between $50-325$ GeV. The main tool which we use to study these collisions is Deep Inelastic Scattering (DIS) which is a process used to probe the inside of hadrons (protons) using the electrons.

The EIC working group has proposed two possible realizations in the future[13]. One is called eRHIC and the proposal is to add a $5-30$ GeV electron beam facility to the existing RHIC facility at BNL to collide with one of its hadron (polarized nucleons and nuclei) beams. The other option is to build a polarized hadron beam facility called mEIC (medium energy EIC) at Jefferson Laboratory (JLab) [14] adding to its already existing electron beam. In this report we will simulate the eRHIC detector.

We will look at the existence of LFV processes in the detector due to the production of a Leptoquark boson, which decays into a lepton (of a different flavor) and a quark. Our main question here is to differentiate between a τ jet arising from a leptoquark boson and a DIS quark jet at the EIC detector.

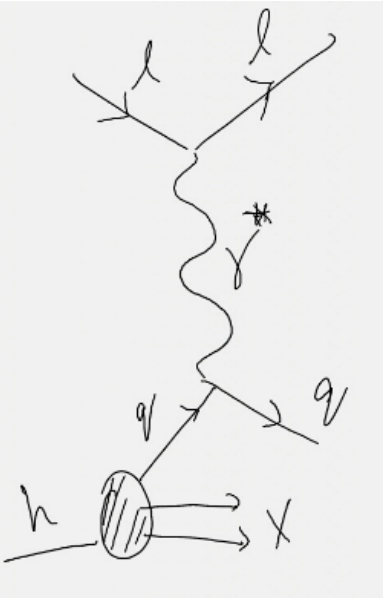


Figure 1.1: Deep inelastic scattering of a lepton on a hadron at leading order in perturbative expansion

The following aspects about the detector like its design, response and which particular process we are looking at i.e our signal, what is our background etc... need to be clarified and understood before we start plotting the widths of the respective jets.

1.2 Leptoquark

Leptoquarks (LQ) are particles coupling to leptons and quarks which arise in models such as the Pati-Salam color-$SU(4)$ and SU(5) Grand Unified Theories (GUT) [5]. Because of their coupling, they posses both lepton and baryon numbers and they are also color-triplet bosons. LQs are very useful for the initial analysis of an $e \to \tau$ conversion because they allow the process to happen at tree level and with particularly large cross-sections relative to other models which allow LFV at loop levels.

LQs have various properties such as: spin 0 or 1; fermion number $F = 3B + L$ (B and L are Baryon and Lepton number respectively)$= 0$ or ± 2; $SU(2)_L$ singlet, doublet, or triplet representations; and chiral couplings to $L-$ or $R-$handed leptons. We will use the Buchmüller-Rückl-Wyler (BRW) parameterization of the LQ [8]. In this model there are 14 different LQ encompassing all allowed combinations of the listed properties.

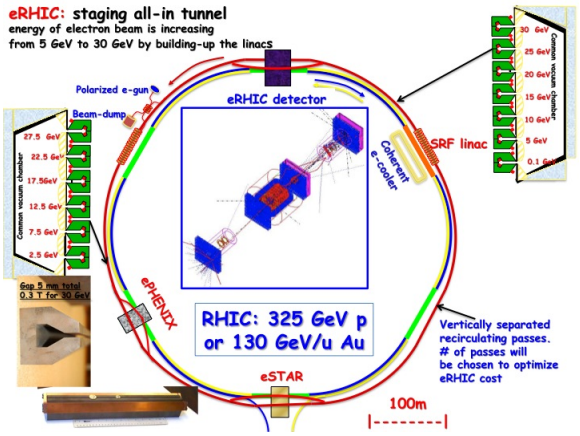
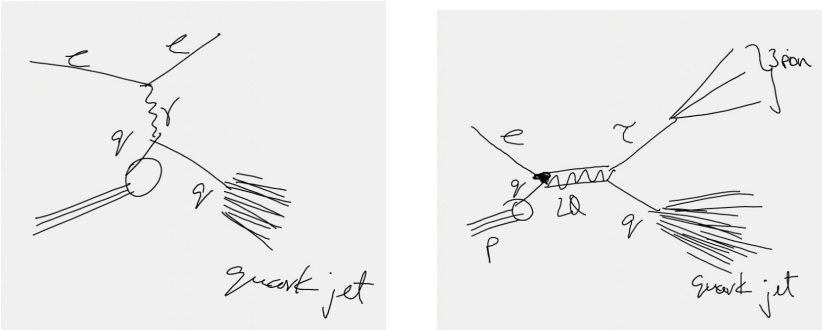


Figure 1.2: Proposed design of eRHIC, with the new electron ring shown in red



(a) DIS (b) LQ

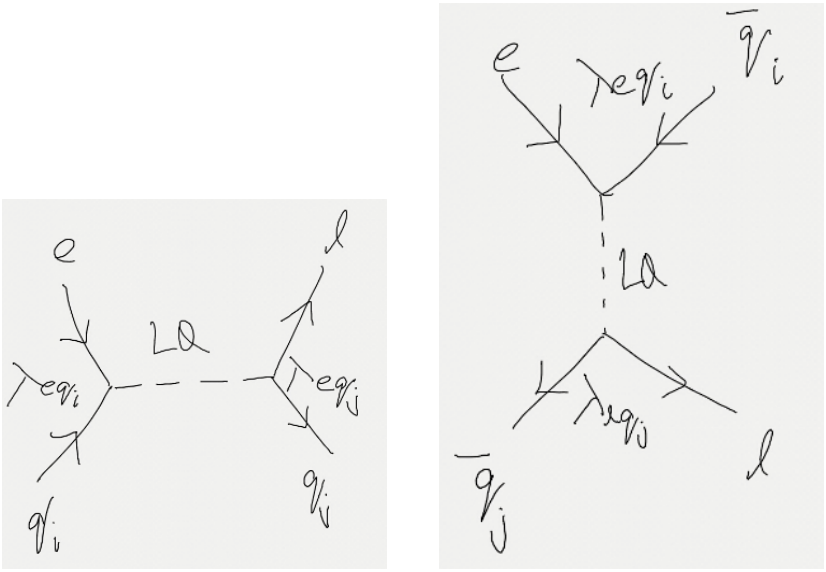
Figure 1.3: Feynman Diagrams of the DIS quark jet and the LQ tau jet arising from its 3 pion decay

The fermion number is assumed to be conserved, taking values of $F = 2$ for $e^- q$ processes and $F = 0$ for $e^+ q$ processes. Here we will only look at $F = 2$ because the EIC is built to have an electron ring. LQ proceesses proceed via s channel resonant LQ production or u channel virtual LQ exchange as shown in Fig:1.4

The cross section of the LQ mediated $e \to \tau$ conversion is calculated in [9] and is found to be proportional to a factor z

$$z = \frac{\lambda_{1\alpha}\lambda_{3\beta}}{M_{LQ}^2} \tag{1.1}$$

where $\lambda_{1\alpha}$ is the coupling of the electron to a quark, $\lambda_{3\beta}$ is the coupling of



(a) s-channel (b) u-channel

Figure 1.4: a): s-channel resonant LQ production and decay to a lepton quark pair. b): u-channel exchange of a LQ. The indices i and j represent quark generation indices, such that λ_{eq_i} denotes the coupling of an electron to a quark of generation i and λ_{lq_j} is the coupling of the outgoing lepton l to a quark of generation j. For $l = \mu, \tau$, the LQ introduces LFV.

the τ to a quark and M^2_{LQ} is the mass squared of the LQ.

We will now look at the τ decay modes and then come back to this z factor which is important for the EIC.

1.3 Tau Decay Modes

LQ with couplings to the first and third lepton generations (e and τ) can be produced in ep collisions and may decay to a τ and a quark. These τ leptons have three kinds of decays namely electronic, muonic and hadronic decays. We will look specifically for the hadronic decays of the tau which is our **sample** and more specifically at the $\tau \to 3\pi^{\pm}$ channel which is our **signal**. We have to keep in mind that these π's are charged pions. These charged pion decays are special in that their signature is often called "pencil like" because they are very narrow when compared to other hadronic jets. The branching ratio of the τ decay is given in table:1.1.

Decay Mode	BR
$\tau \to \nu_\tau + \bar{\nu}_e + e$	17.8%
$\tau \to \nu_\tau + \bar{\nu}_\mu + \mu$	17.4%
$\tau \to \nu_\tau + \pi^\pm$	11.0%
$\tau \to \nu_\tau + \pi^\pm + \pi^0$	25.4%
$\tau \to \nu_\tau + \pi^\pm + 2\pi^0$	10.8%
$\tau \to \nu_\tau + \pi^\pm + 3\pi^0$	1.4%
$\tau \to \nu_\tau + K^\pm + n\pi^0$	1.6%
$\tau \to \nu_\tau + 3\pi^\pm + n\pi^0$	15.2%

Table 1.1: Tau decay channels and their branching ratios

1.4 Previous Experimental Searches

The search for LQ involved in LFV has been carried out at HERA and very stringent limits were set by both the H1[7] and ZEUS [10] collaborations. In both detectors, no $e \to \tau$ events were observed during their runs. Since the physicists did not find any evidence for LFV they put limits on the mass and couplings of the LQ in the BRW model. They have excluded the LQ which couple to the τ up to masses of 379 GeV.

The best existing limits on the $e \leftrightarrow \tau$ conversion come from the BaBar collaboration [11] with the process $\tau \to e\gamma$ [1] and the BELLE collaboration [12] $\tau \to lhh'$ $(h, h' = \pi^\pm$ or $K^\pm)$ [2].

1.5 Advantages at EIC

There are several reasons to suggest that the EIC will improve on the existing measurements. The EIC would in principle be sensitive to the $e \to \tau$ conversion cross section to a level of 0.001 fb with an integrated luminosity of 1000 fb^{-1}[2]. This would give us on the order of one $e \to \tau$ event per sample, without taking into account the background and reconstruction efficiency. Using this number for the cross-section and assuming a center-of-mass energy of $\sqrt{s} = 90$GeV, we can calculate the factor z from eq:1.1. We understand that for all the combinations of lepton and quark generations α and β respectively, the EIC could probe values of z which are smaller than the HERA limits by a factor between 10 and 200 [13]. Since z is inversely proportional to M^2_{LQ}, we can probe for massive LQs beyond our energy threshold.

[1] $B < 3.3 \times 10^{-8}$

[2] $B \sim (2.0 - 8.4) \times 10^{-8}$

Chapter 2

Simulating the EIC detector and Generator Level Analysis

In order to study the τ jet at the EIC detector we must know how the particles produced in the collision interact with the materials in the detector. The exact structural details for the EIC are still at the R&D stage so we have used the current geometrical representation given in the EIC letter of intent [2]. For simulating the interaction of particles with matter there are lots of software packages available but the one that the particle physics community often uses is Geant [19]. Geant is a toolkit for producing computer models of complex systems comprising of geometrical shapes and to calculate the effect of particles traveling through them. We have used a new framework called Fairroot[20] in this analysis breaking away from the tradition of using Geant alone.

2.1 Fairroot

Fairroot was developed at GSI Helmholtzzentrun für Schweironenforschung GmbH at Darmstadt, Germany by Mohammad Al-Turany and Florian Uhlig. GSI is a Large scale accelerator for heavy ions. The "Fair" in Fairroot comes from their Facility for Antiproton and Ion Research (FAIR) and "root" is a data analysis framework developed at CERN in 1995 by Rene Brun and Fons Rademakers.

The Fairroot framework is fully based on the ROOT system. Fairroot lets the user simulate data and/or perform analysis with the same framework. Moreover, Geant3 and Geant4 (which is the latest version of Geant) transport engines are supported making it convenient for the user. It is very portable in that the user code that creates simulated data does not depend on a particular Montecarlo engine. The framework delivers base classes which enable the user

to construct their detectors and/or analyze tasks in a simple way. It also delivers some general functionality like track visualization and event display.

Fairroot is ideal for physicists as it allows the user to completely avoid going deep into software components in order to change it to their needs. It is a unified package with generic mechanisms to deal with the most commonly used tasks in the high energy physics community. Using Fairroot is analogous to driving a car. One does not need to know how the gear system works in order to drive properly. Similarly, the user can concentrate on detector performance details and physics deliverables while avoiding purely software engineering issues like storage, retrieval, code organization etc.

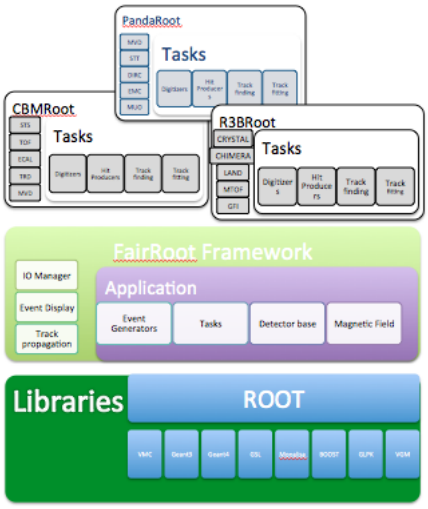


Figure 2.1: Structure of the Fairroot class hierarchy

2.1.1 Why Fairroot?

Almost all of the detector simulation presently working in the HEP community is created using Geant4 and then the output is later converted into root files for analysis. Fairroot seeks to avoid that step in between and unify the process. Of course, Fairroot also uses Geant4 but it is integrated in such a way that the person does not need to know how to code in Geant.

In order to create a detector geometry in Fairroot one just has to specify the geometry details (according to Geant) in a .geo file and every thing else is accomplished by Fairroot. On the other hand, in Geant, one has to effectively create new individual classes for every module which can quickly become tedious for a complicated geometry structure.

2.2 Simulating EIC from scratch using Fairroot

The EIC detector is a 4π detector with emphasis on particle identification and similarity of angular coverage for tracking and calorimetry.

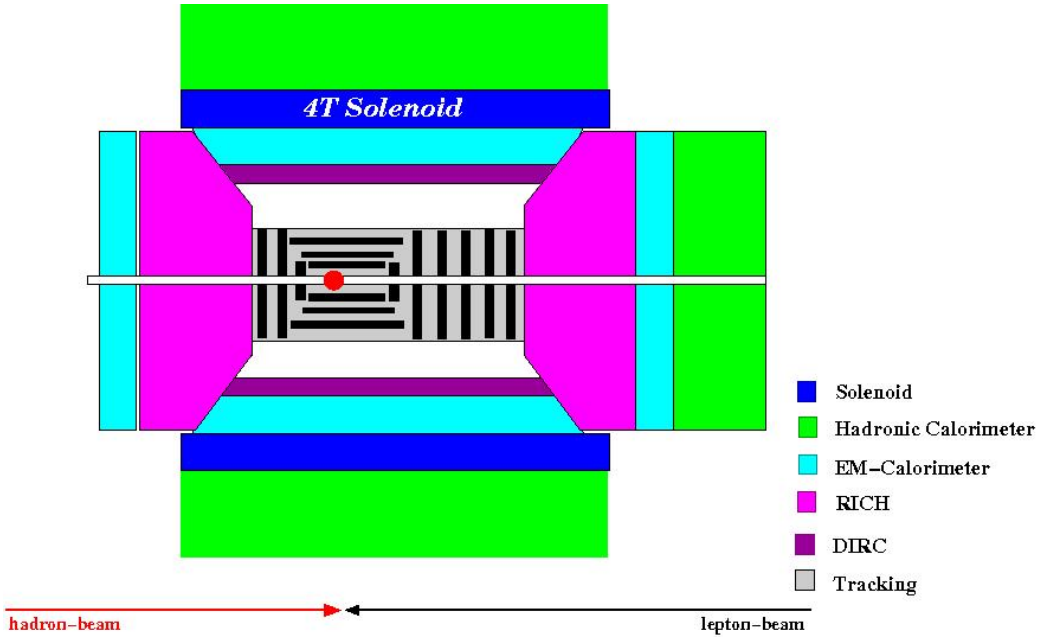


Figure 2.2: Cross-section of the EIC central detector

The EIC detector which we have simulated for this study is by no means the final structure and hence not ideal. The tracker system for the EIC is divided into three segments - the vertex, the barrel and the forward trackers. Each is modeled separately after the ZEUS micro-vertex detector, BaBar barrel drift chambers and the HERMES forward drift chambers respectively. For the forward trackers, there are multiple possibilities. For example, we might use Gas Electron Multipliers (GEM)[2] and line them parallel to each other.

For Particle Identification (PID), we are modeling the Detection of Internally Reflected Cherenkov (DIRC) on the BaBar DIRC and the Rich Imaging Cherenkov detector (RICH) is based on the HERMES dual radiator RICH[22] which has a layer of aerogel in front of the volume.

Now that we have our EIC detector simulated in fairroot, we have to simulate *ep* collisions and pass it through the detector to look at the output. The two data sets which we will use are DIS events and LQ events.

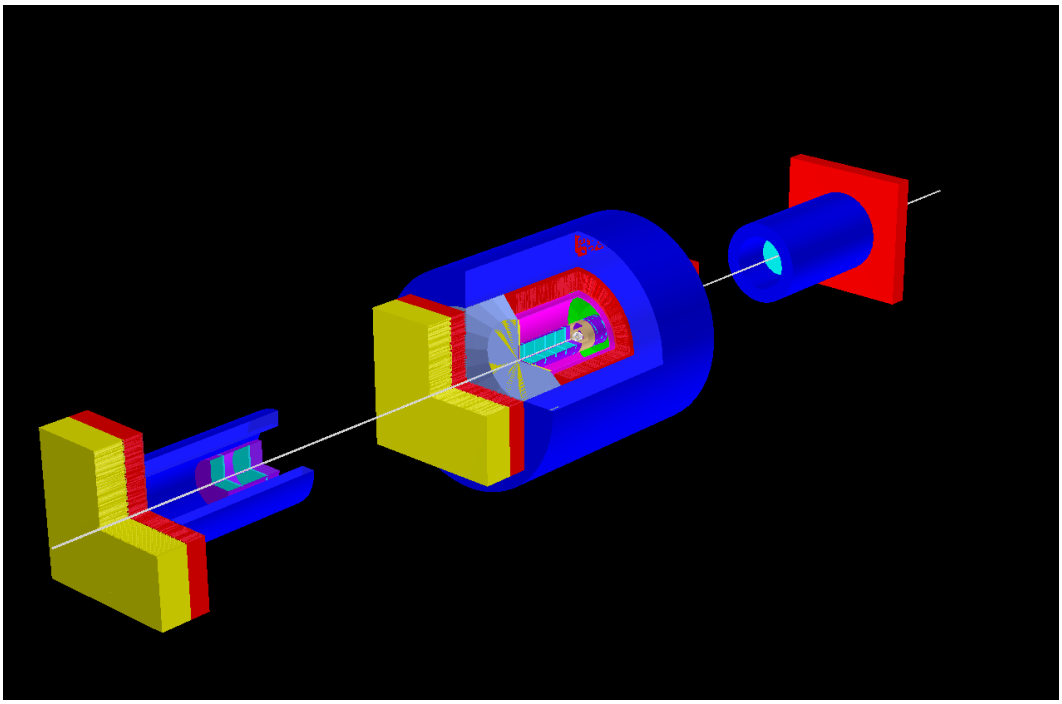


Figure 2.3: EIC detector simulated in Fairroot

2.3 Creation of Montecarlo data sets

These data sets were created using PYTHIA, a MonteCarlo (MC) simulation program for particle collisions at very high energies in particle accelerators. We used PYTHIA6.4 [23] for this study. It simulates hard processes, parton showers, secondary level interactions, hadronization, decays, short lived resonances and many other processes involved in full reconstruction of collisions between particles.

2.3.1 DIS generator

DIS events are of the form $e + q \rightarrow e + q + X$, where the final state q and X both hadronize and we get 2 or more hadron jets in the final state. The event generator is maintained at BNL by the EIC group and all these events follow SM branching ratios and processes. These events are a mixture of leading order DIS, QCD Compton scattering ($\gamma^*_{T/L} q \rightarrow qg$), photon-gluon fusion ($\gamma^*_{T/L} \rightarrow q\bar{q}$) as well as resolved and soft vector meson dominated processes (VMD). The cuts imposed on this data set is that $Q^2 > 0.1$ GeV^2 and x (Bjorken variable - momentum fraction) range is from 10^{-9} to 1.

Detector	Material
Barrel EM Calorimeter	Lead Glass (O, Si, Ti, As, Pb)
DIRC	Fused Silica (SiO_2)
Silicon Detectors	Si
Trackers	Mylar, Ar/CH_4/CO_2 90 : 5 : 5
High Threshold Cherenkov (HTCK)	CO_2
RICH	C_4F_{10}, Aerogel (O_2SiH_2)
Barrel Tracker inner surface	Be
Barrel Tracker end plate	Al
Barrel Tracker Gas Volume	He, C_4H_{10} 80:20
Solenoid	Fe
Dipole Magnets	Fe
Beam Pipe	Be
Hadronic Calorimeters	Fe
Electromagnetic Calorimeters	Lead Glass

Table 2.1: List of Detectors and their material make up for the EIC detector

Detector	η range
Barrel EM Calorimeter	-0.65 to 1.35
ECal (forward, inner)	1.516 to 5.545
ECal (forward, outer)	2.706 to 6.78
ECal (Backward, inner)	-0.94 to -4.852
ECal (Backward, outer)	-2.508 to -6.579
HCal (inner)	1.625 to 5.663
HCal (outer)	2.742 to 6.815

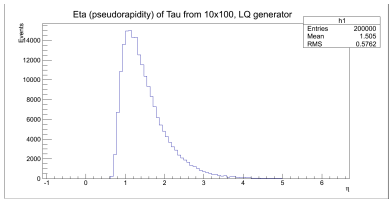
Table 2.2: η coverage of Calorimeters (Hardonic and Electromagnetic)

2.3.2 LQ generator

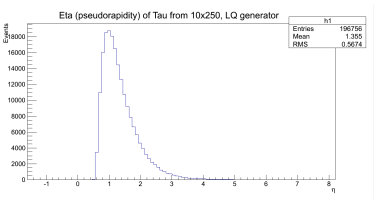
In order to produce LQ mediated τ events, we have used another PYTHIA-like generator called LQGENEP [24]. It is a modification to the DIS event producer where the principle process of creating events involve the presence of the LQ according to the BRW model described in chapter 1. The user input parameters for this MC generation are the number of events, mass of the LQ, energy of the lepton beam, energy of the proton beam, initial quark flavor and final quark flavor. we have chosen the following values for the parameters: mass of the scalar LQ was set to $M_{LQ} = 1,936.5$ GeV (calculated from Yukawa coupling $\lambda_i = \lambda_j = 0.3$), and the initial and final state quarks to be d quarks.

2.4 Generator Level Analysis

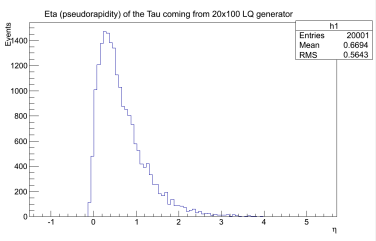
Before we can put the MC generated tracks in the detector, we analyzed them to get an idea about the distribution of the particles. The τ jets in the LQ data are all present in the forward direction with $\eta > 0$ (η : Pseudorapidity) and so we should be able to get all the required information from them using only the forward calorimeters and the barrel calorimeter.



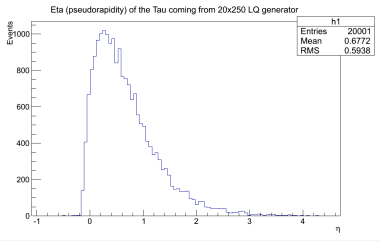
(a) 10x100



(b) 10x250



(c) 20x100



(d) 20x250

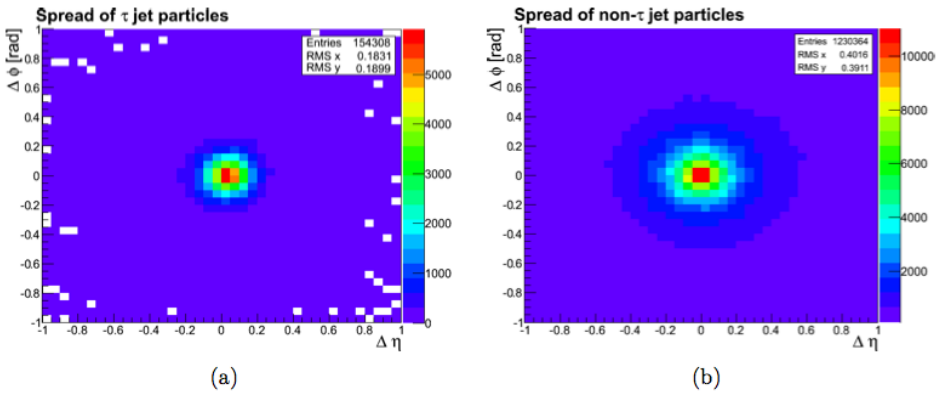
Figure 2.4: Pseudorapidities η for τ at different energies of collision

Here we have to concede that there is a gap [1] in our detector where in some tracks can escape and that is between the barrel ECAL and forward inner ECAL. This gap is from $\eta = 1.35$ to $\eta = 1.516$. This is a problem because according to Fig:2.4 (b) we will miss quite a number of τ jets.

Recent analysis of these τ jets at the generator level by our group[25] have shown that the width of the τ jet when we look only at the 3 pion decay is about half of the standard width of the quark jet that comes from DIS. For fig:2.5 the particle with the highest momentum was chosen as the lead and every particle around it was plotted in $\Delta\phi$ and $\Delta\eta$(for example the $\Delta\phi$ refers to the difference between the lead particle's ϕ and the other particle's ϕ) on an event basis. So if we compare the color scales in fig:2.5 we see that the

[1] by the gap I refer to a volume in the detector wherein we do not have calorimeters to collect the information

number of particles in the $\Delta\phi \sim \Delta\eta \sim 0$ bin in (b) is about four times that of (a). Even though the number of events for the nonτ jet particles is larger than the τ jet particles, we see that the jet structure is pretty well defined.



(a) (b)

Figure 2.5: Plot (a) Angular distribution of hadronic τ decay products in LQ data. Plot (b) shows the corresponding distribution using leading charged pions that were not associated with the known τ jet

2.5 Running the events through the Detector

The MC data sets were given as input to the Fairroot macro which runs them in the EIC detector. It is very important to look at the time taken to complete the simulation. There are 2 factors that we have to look at here. First the number of events (obviously) and the number of tracks per event. For a sample event in the DIS or LQ generator we get on average 20 to 30 tracks. Each of these tracks has to interact with the material in the detector and that in turn produces more particles/tracks which also have a chance to interact. So finally we end up anywhere from 1000 to 100000 tracks per event. Therefore our macro/simulation output file has to store information about all the hundreds of thousands of tracks as they hit/exit any volume/detector that is sensitive.

This volume of memory and hard drive usage puts a constraint on the number of events that can be processed in one run. The time scale for different events is given in table:2.3.

A linear fit gives us an equation for the time as $t = 2.22678N + 7.72731$ in seconds where N is the number of events with $R^2 = 0.999992$. The EIC computing system is still in its beginning stages and so we have only performed the

No of Events	Time (seconds)
1	8.28
10	27.78
100	234.7
1000	2234.1

Table 2.3: CPU usage timings in seconds for Fairroot to simulate particle tracks through the EIC detector

analyses required for this report with 10,000 events at the maximum. The EIC machines are normal desktops with 16 core 2.88 GHz speed processor having 4GB of ram. For events around 2000, the ram usage exceeds the available space and due to the absence of parallel processing systems, it is not efficient to keep doing these for more and more number of events.

Now we have created the detector, the MC data sets and have information from the detector on the tracks. The next step is to group them into jets and look at their area in the calorimeter cells. Since we are dealing with analysis of jets, it is very important that we give a fair bit of thought and effort into looking at the different possibilities of clustering jets in the detector. This subject will be dealt with in detail in the coming chapter on Jet Selection.

Chapter 3

Jet Selection and Detector Level Analysis

A jet is a narrow cone of hadrons and other particles produced by the hadronization of a quark or gluon in the detector. As I have mentioned before we will look at two different types of jets in this analysis, the quark jet from DIS event and the τ jet from the LQ event. Since we are only interested in the three pion decay of the τ, our τ jet will consist of the 3 charged pions and the particles it produces on interactions with the detector. Let us now look at how we can convert hits in the detector to jets using various well known algorithms.

3.1 Jet Clustering Algorithms

Fairroot will tell us the momentum and energy of all particles which hit out calorimeters, which we can use to trace them back using Jet clustering algorithms. There are essentially two classes of jet algorithms in use: cone-type algorithms 'Sternman-Weinberg et al' [26] and clustering type algorithms [27] which were first introduced by the Jade collaboration. In cone-type algorithms jets are defined by maximizing the amount of energy which can be covered by cones of defined size, whilst in clustering algorithms particles are assigned to jets iteratively according to whether a given energy-angle resolution variable $d_{i,j}$ exceeds a fixed resolution parameter. We will now look at three varieties of the iterative clustering algorithms and how they are defined.

3.1.1 Longitudinally invariant k_t

The Longitudinally invariant k_t jet algorithm [31] comes in inclusive and exclusive variants. the inclusive variant is formulated as follows (k_t represents the

parton's transverse momentum/energy. If we choose energy then it is called E-scheme or Energy-scheme recombination):

For each pair of particles i,j calculate the k_t distance.

$$d_{i,j} = min(p_{t,i}^2, p_{t,j}^2)\Delta R_{i,j}^2/R^2$$

with $\Delta R_{i,j}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, where $p_{t,}$, y_i, ϕ_i are the transverse momentum, rapidity and azimuth of the particle i. R is a jet-radius parameter usually set to 1. For each parton i also calculate the beam distance $d_{i,B} = p_{t,i}^2$.

Then we proceed to find the minimum d_{min} of all the $d_{i,j}$, $d_{i,B}$. If d_{min} is a $d_{i,j}$ merge particles i and j into a single particle, summing their four momenta; if it is a $d_{i,B}$ then declare particle i to be a final jet and remove it from the list. This procedure continues till there are no initial state particles remaining and all have been converted to beam particles.

The exclusive variant of the longitudinally invariant k_t algorithm is similar except that when $d_{i,B}$ is the smallest value, that particle is considered to become part of the beam jet (i.e. discarded) and clustering is stopped when all $d_{i,j}$ and $d_{i,B}$ are above some d_{cut}. In the exclusive mode R is commonly set to 1.

3.1.2 Cambridge/Aachen

The *pp* Cambridge/Aachen (C/A) jet algorithm [28] is provided only in an inclusive version whose formulation is identical to that of the k_t jet algorithm, except as regards the distance measure which are:

$$d_{i,j} = \Delta R_{i,j}^2/R^2$$

with $d_{i,B} = 1$.

This clusters jets purely based on their distance towards each other with disregard to the transverse momentum of the respective particles.

3.1.3 anti k_t

This algorithm introduced and studied in [33] is defined exactly like the standard k_t algorithm, except for the distance measures which are now given by

$$d_{i,j} = min(1/p_{t,i}^2, 1/p_{t,j}^2)\Delta R_{i,j}^2/R^2$$

with $d_{i,B} = 1/p_{t,i}^2$. While it is a sequential recombination algorithm like k_t and Cambridge/Aachen, the anti-k_t algorithm behaves in some sense like a perfect cone algorithm, in that its hard jets are exactly circular on the $y - \phi$ cylinder.

3.2 Issue with Sequential algorithms

The above mentioned sequential algorithms have one main problem in that it is heavily dependent on the number of tracks/particles per event. Since we are creating 2 arrays $d_{i,j}$ and $d_{i,B}$ and we are looping over them to find the least value, we have an algorithm which is $O(N^3)$ where N is the number of elements in each array. Now when we do *ep* collision, we create tens of thousands of particles per event and we have million of events which puts a constraint on the timing of algorithms. An algorithm of $O(N^3)$ is not ideal on that scale [35].

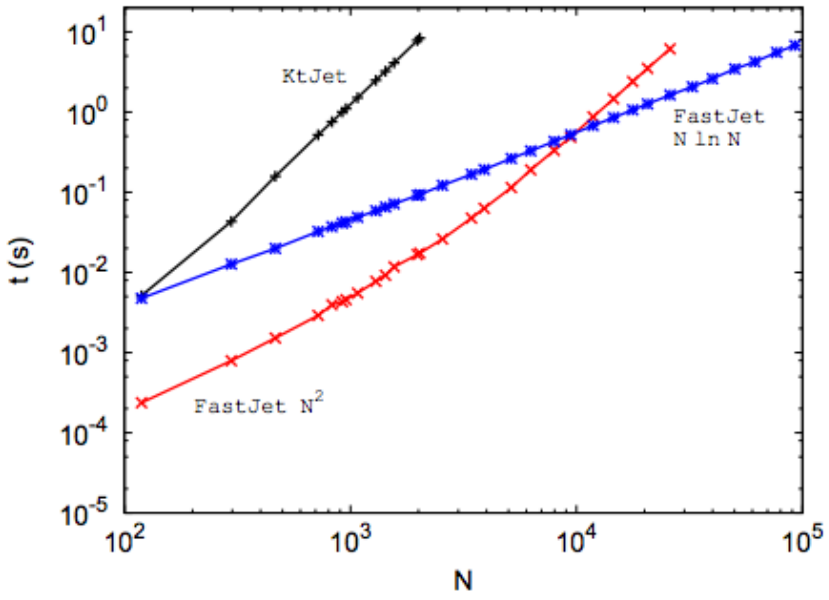


Figure 3.1: The running times of the KtJet and FastJet implementations of the k_t clustering jet-finder versus the number of initial particles. [35]

So that is why we have decided to go to a package which is freely available called FastJet.

3.3 FastJet Package

FastJet is a C++ package that provides a broad range of jet finding and analysis tools. It includes efficient native implementations of all widely used 2 to 1 sequential recombination/clustering jet algorithms [34]

To alleviate the timing problem, FastJet makes use of the observation that the smallest pairwise distance remains the same if one uses the following alternative (non-symmetric) $d_{i,j}$ distance measure:

$$d_{i,j} = p_{t,i}^2 \Delta R_{i,j}^2 / R^2; dj, i = p_{t,j}^2 \Delta R_{i,j}^2 / R^2$$

For a given i, the smallest of the $d_{i,j}$ is simply found by choosing the j that minimizes the $\Delta R_{i,j}$, i.e by identifying i's geometrical nearest neighbor on the $y - \phi$ cylinder. Geometry adds many constraints to closest pair and nearest neighbor type problems, e.g. if i is geometrically close to k and j is geometrically close to k, then i and j are also geometrically close; such a property is not true for the $d_{i,j}$. The factorization of the problem into momentum and geometrical parts makes it possible to calculate and search for minima among a much smaller set of distances. This approach is sufficiently powerful that with the help of the external Computational Geometry Algorithms Library (CGAL) [29] and using the Delaunay Triangulation modules (For a set P of points in a plane, the Delaunay triangulation is a triangulation such that no point in P is inside the circumcircle of any triangle formed by the triangulation), FastJet achieves a timing of $O(N \ln(N))$.

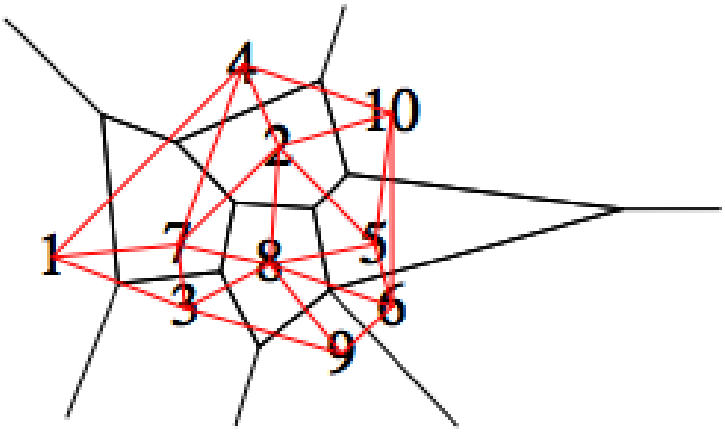


Figure 3.2: Shown are the voronoi diagram and the Delaunay triangulation (in red) for randomly generated points - wikipedia.org/voronoi-diagram

Given an ensemble of vertices in a plane (specified by the y_i and ϕ_i of the particles), one can find the nearest neighbor of each vertex by using a structure

known as a Voronoi diagram [30] or its dual, a Delaunay Triangulation. The Voronoi diagram divides the plane into cells (one per vertex), such that every point in the cell surrounding a vertex i has i as it nearest vertex. The structure is useful for nearest neighbor location because the vertex nearest to vertex i is always in one of the (few, i.e $O(1)$) cells that share an edge with the cells of vertex i. This gives us a final $O(N \ln(N))$

Almost all of these jet algorithms use neural networks as an efficient way of accessing/processing large amounts of data.

3.3.1 Neural Networks

The recent wave in high energy physics computing involves the use of artificial neural networks in off-line data analysis, particular for event reconstruction in particle detectors [37]. An artificial neural network consists of a set of interconnected units (neurons). The state or activation of a given i-neuron, γ_i is a real number, determined as a function of the activation of the neurons connected to it. Each pair of neurons (i, j) is connected by a synapsis characterized by a real number $w_{i,j}$ (weight), where the weights need not be symmetric. The activation of each neuron is a function g of the difference between a weighted average of input from other neurons and a threshold θ_i. This is sometimes known as multi-layer perception in Feed-Forward architecture [36].

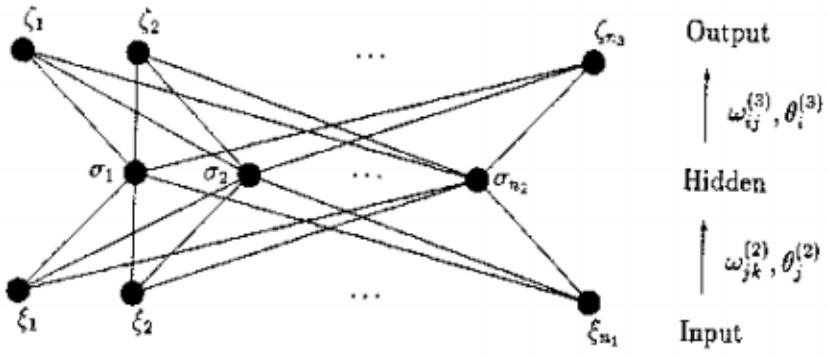


Figure 3.3: A three lair feed forward neural network consisting of input, hidden and output layers

For a given activation function, the parameters can be tuned in such a way that the neural networks can reproduce any continuous function. The behavior of a neural network is determined by the joint behavior of all its connections

and thresholds, and it can thus be build to be redundant, in the sense that modifying, adding or removing a neuron has little impact on the final output.

Learning Process

The usefulness of neural networks is due to the availability of a training algorithm. This algorithm allows one to select the values of weights and thresholds such that the neural network reproduces a given set of input-output data (or patterns). This procedure is called learning since unlike a standard fitting procedure, there is no need to know in advance the underlying rule which describes the data. Rather, the neural network generalizes the examples used to train it. This is very analogous to how humans develop intuition for the solution to a problem based on past experiences.

The FastJet algorithms use some of these neural networks which have been trained on present ee, ep, pp data in their package.

3.4 Jet Areas

Since the information we have about the respective jets at the generator level is their width/area, we will look at the area of the jets when they hit the calorimeters.

FastJet package includes an option which calculates areas of jets based on the information provided by the calorimeters. The area provides a measure of the surface in the $y - \phi$ plane over which a jet extends, or equivalently a measure of a jet's susceptibility to soft contamination. We want to look at the area that a jet makes in the calorimeters since the square root of the ratio of the areas is proportional to the ratio of the jet widths. There are three predefined jet areas in the package

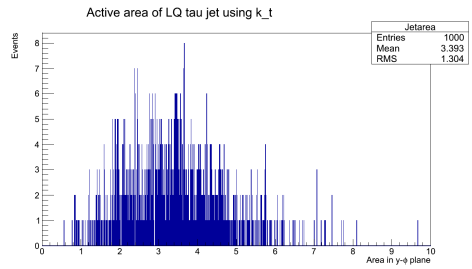
- Active area adds a uniform background of extremely soft massless 'ghost' particles to the event and allow them to participate in the clustering. The area of a given jet is proportional to the number of ghosts it contains. Because the ghosts are extremely soft (and sensible jet algorithms are infrared safe), the presence of the ghosts does not affect of the set of user particles that end up in a given jet. Active areas give a measure of a jet's sensitivity to diffuse background noise.

- Passive areas are defined as follows: one adds a single randomly placed ghost at a time to the event. One examines which jet (if any) the ghost ends up in. Then we repeat the procedure many times and the passive area of a jet is then proportional to the probability of it containing the

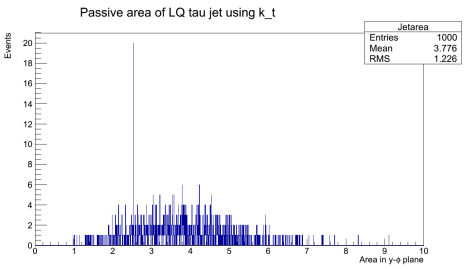
ghost. Passive areas give a measure of a jet's sensitivity to point-like background noise.

- The Voronoi area of a jet is the sum of the Voronoi areas of its constituent particles. The voronoi area of a particle is obtained by determining the Voronoi diagram for the event as a whole, and intersecting the Voronoi cell of the particle with a circle of radius R centered on the particle. Note that for the k_t algorithm the Voronoi area of a jet coincides with its passive area.

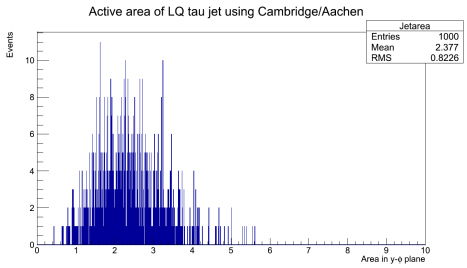
For this analysis we will be using the anti-k_t algorithm and active area as the area definition based on the plots Fig:3.4 which are made from a sample of 1000 LQ generated τ events.



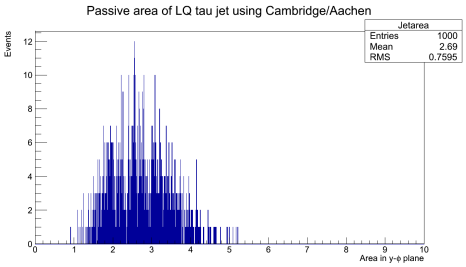
(a) k_t; Active area



(b) k_t; Passive area



(c) C/A; Active area

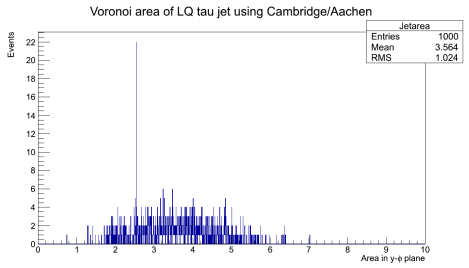


(d) C/A; Passive area

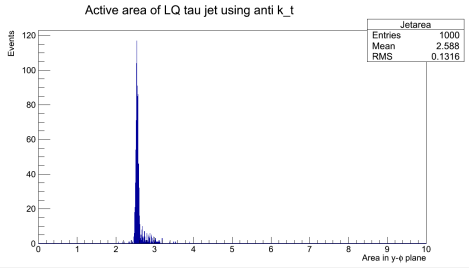
Figure 3.4: Sample events through the FastJet area package with different options of clustering algorithms and area definitions

3.5 Jets in the Detector

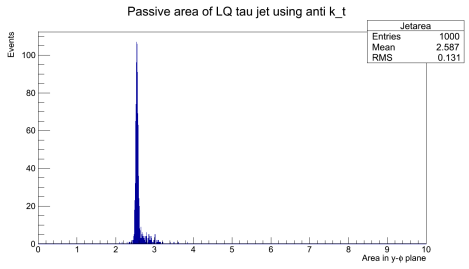
The LQ generator produces events which all have a τ in them and that particular τ can decay into the $3\pi^{\pm}$ channel about 15% of the time table:1.1. For



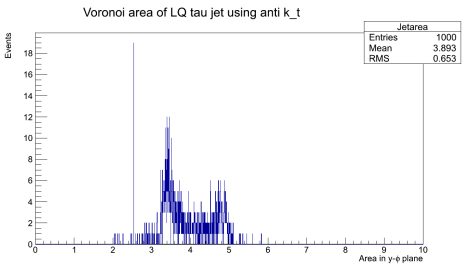
(a) C/A; Voronoi Area



(b) Anti k_t; Active area



(c) Anti k_t; Passive area



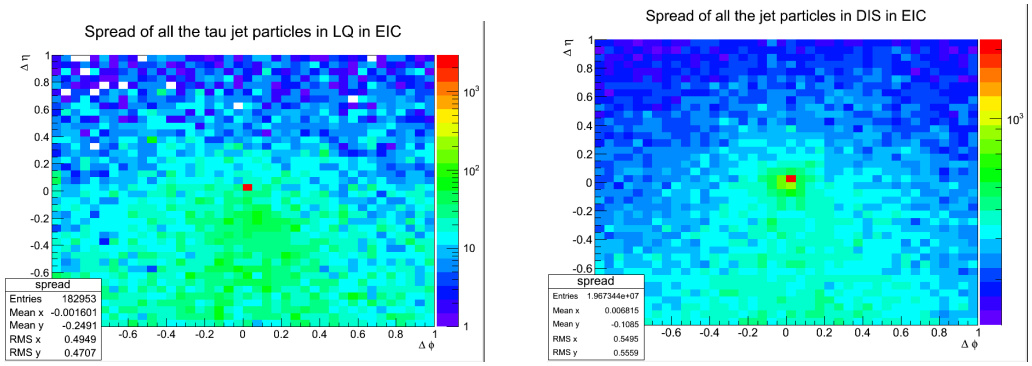
(d) Anti k_t; Voronoi area

Figure 3.5: Sample events through the FastJet area package with different options of clustering algorithms and area definitions

this analysis we will take only the tracks which have come from the τ and have them as input to the detector simulation. Now we have eliminated the need to tag the τ jet from other jets.

For the DIS event, we have put all the tracks into the detector and we are looking at the jet with high pt, which will correspond to the quark jet.

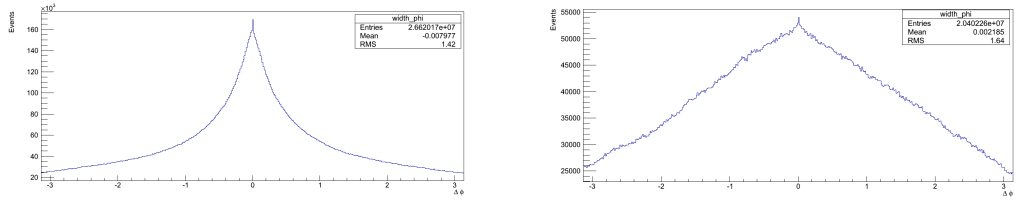
Plotting the particles hits in the detector according to the prescription outlaid in plots 2.5 we get fig:3.6. This plot, when compared to fig:2.5 shows little difference between the jets. We can see that sizes of the two jets are different just by looking at the values in the color scales, but there are certainly a lot more extraneous particles which have bigger η than the lead particles. This is due to particles traveling with high η which are produced from interaction with the material near the interaction region. The following plots indeed confirm that while there is a difference between the two jets in the detector, it is very premature to say that its a signature. The problem with this kind of analysis done in the detector is that it assumes that there is only one jet and all the particles are part of that jet. So when we do this analysis at an event basis, some events can have very closely moving particles of different jets grouped as one and on the other hand completely different jets moving in the opposite



(a) LQ tau jet spread (b) DIS quark jet spread

Figure 3.6: Angular distribution of the different jets in the detector : Hits collected by the EM calorimeter

directions can also be counted as one jets.



(a) LQ, $\Delta\phi$ projection in the detector (b) DIS, $\Delta\phi$ projection in the detector

Figure 3.7: $\Delta\phi$ projection of the widths of the different jets in the detector for similar number of events : Hits collected by the EM calorimeter

Therefore this method is not very helpful to us and so we plot the areas of the jets calculated using the FastJet package (which is now integrated with Fairroot).

We began this analysis asking what is the ratio of the jet width of the DIS quark jet to the LQ τ jet in the EIC detector and if we can use that factor to indicate a special signature of the LFV violating third generation LQ production. Our results in the detector indicate that the τ jet (once it has interacted with matter in the detector) cannot be differentiated from the quark jet based on its area alone. If it had followed the ratio at the generator level, we should have seen a ratio of $1/4^{th}$ for the areas. Just by looking at the areas of the jets in the calorimeter cells we cannot increase the sensitivity of the EIC towards the τ jet.

When we study Figs:3.9 and 3.10 we see that:

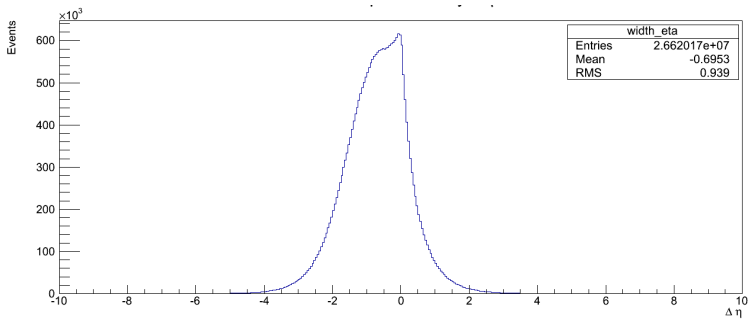


Figure 3.8: Plot showing the hits from the proton debris in the high η direction. This heavily impacts our lead particle algorithm to find the jet width

- they both rise sharply towards their respective peak from the left side which means that there is a threshold for the jet area.
- the peaks have a very small width (which is a consequence of our area calculating algorithm).
- there exists a tail towards the high areas which is enhanced for the DIS quark jets.

The direction that the τ's are produced coincide with the gap mentioned in chapter 2 between the barrel ECAL and the forward inner block ECAL. At that η, the tracks which are coming from the interaction region would hit the calorimeters at a much bigger angle causing it to cover a wider spread of the calorimeters. This would suggest to us that jets which have really big areas would be found near the gap. On the other hand, the DIS quark jets do not have this preference in the forward direction and thus have a pretty good area description. The following plots 3.11 to 3.13 show the η of LQ τ jets which are in different regions in the area plot (peak, end of peak and the tail regions).

We have looked at a specific BSM process involving the production of a third generation LQ at the EIC detector, which we have simulated uses a preliminary geometry model. The whole of the analysis was performed in a relatively new framework called Fairroot which integrates the geometry construction, transport engines and an analysis framework (ROOT) all into one package. This is the first time that such an analysis has been performed for the EIC using Fairroot and this was a test to see if this recent developmental framework will benefit the experimental nuclear/particle physics communities in the future.

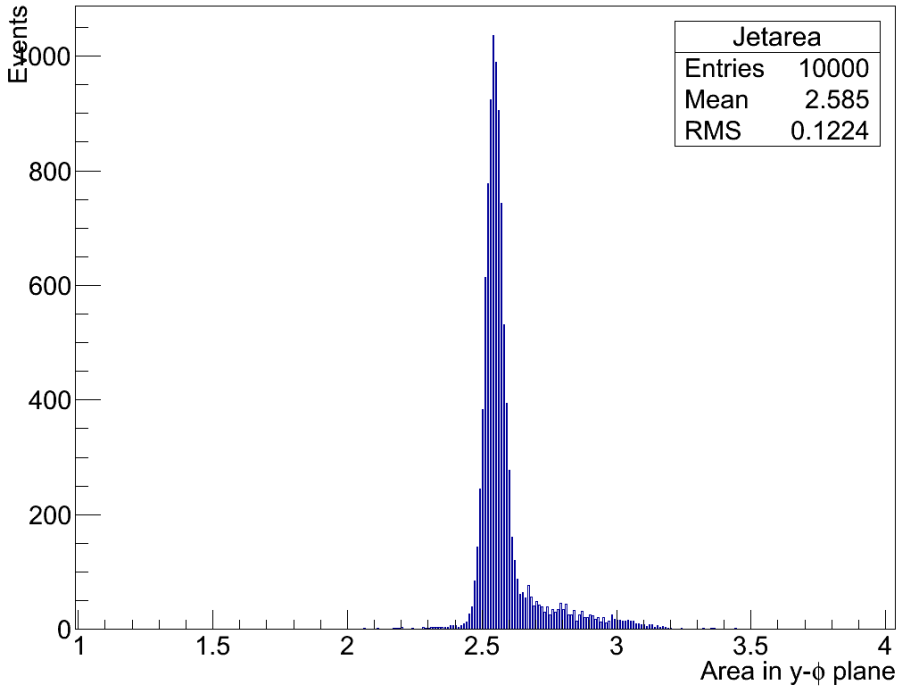


Figure 3.9: Area of the LQ tau jet in the EIC detector

3.6 Conclusion from Analysis

From our study, we see that the areas of the respective jets (DIS quark and LQ τ) inside the EIC detector does not follow the same ratio which it exhibited at the generator level. This ratio does not give us a specific signature when we look at that processes taking only the calorimeter into account. All the other modules n the EIC like the RICH (Ring Imaging Cherenkov), Barrel trackers etc.. were all just volumes to simulate particle interaction with their materials.

At the generator level we saw that the width of the tau jet was about half when compared to a standard DIS jet. This means that that area of the tau jet is about $1/4^{th}$ of the quark jet. Even though we are seeing a slight difference in the jet areas in the detector for the two jets, it is not enough to classify it as a signature when we look for it based on its size alone.

3.7 Future work - Updates

In this section we will outline the plan for the future of this particular analysis.

- Creating EICRoot: The current simulations were all based on the ex-

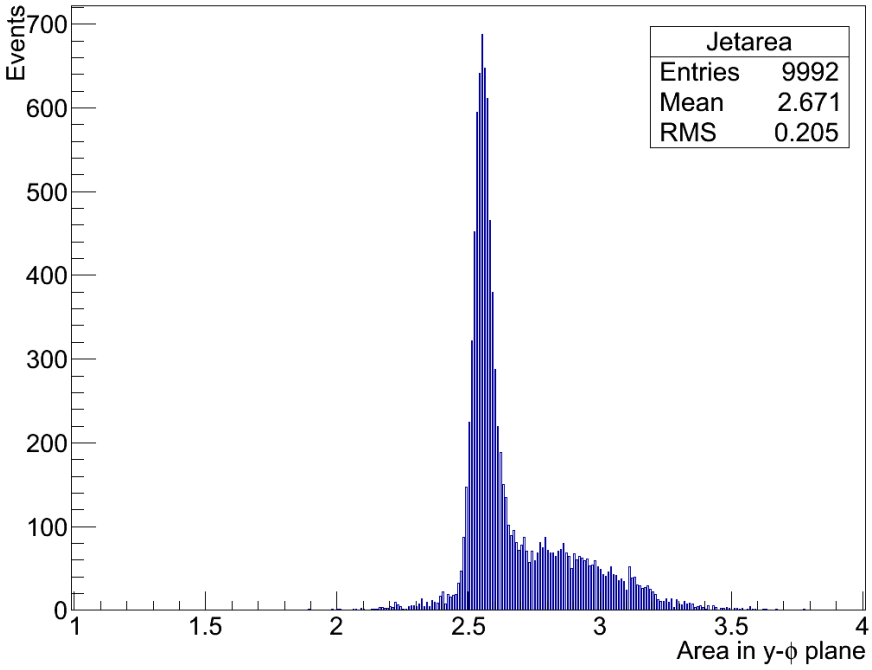


Figure 3.10: Area of the DIS quark jet in the EIC detector

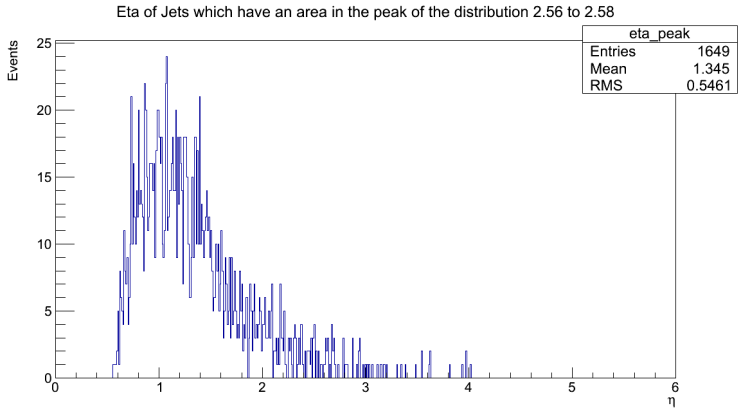


Figure 3.11: η of τ jets in the peak area of plot 3.9

ample distribution of Fairroot called fairbase. It would be very useful to create a stand-alone EICRoot with its own modified base classes and individual detector classes for future analysis.

- Better Readout Geometry: Due to the current development stage of the EIC detector, all the detector modules and calorimeters have very limited

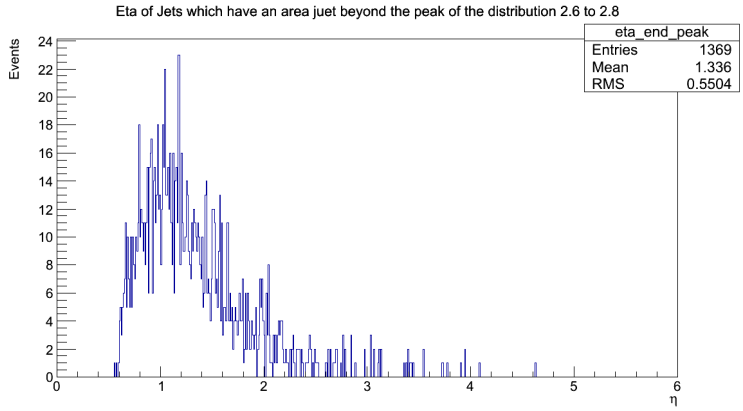


Figure 3.12: η of τ jets in the end of the peak of plot 3.9

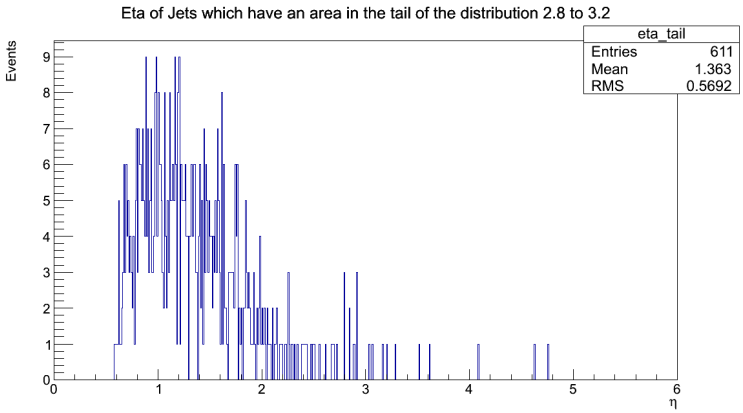


Figure 3.13: η of τ jets in the tail of plot 3.9

geometry structure. Once individual groups which work on their specific modules come up with a good design for it, we have to implement that in the geometry files and also have the classes for that module appended appropriately.

- Jet selection and Tagging at RECO level: Using the information from the hits in the calorimeter and the track information from the RICH and the trackers, we should be able to select the τ jet at the reconstruction level. This involves detailed detector specific class structure in EICRRoot and a comprehensive description of the readout mechanism and parameters.

Finally I would like to thank Prof. Abhay Deshpande, Dr. Klaus Dehmelt, Dr. Dmitri Kharzeev, and Dr. Matthew Dawber for being part of my thesis committee, Stony Brook University Physics department for giving me the opportunity to take part in this research.

Bibliography

[1] J. Ashman *et al.*. [European Muon Collaboration], *Phys. Lett.* B 206 (1988) 364

[2] C.Aidala *et al.*, 'Letter of intent for Detector R&D Towards an EIC detector', (2011)

[3] Q.R. Ahmad *et al.* [SNO Collaboration], *Phys. Rev. Lett.* 89, 011301 (2002) [arXiv:nucl-ex/0204008].

[4] Y. Fukuda *et al.* [Super-Kamiokamde Collaboration], *Phys. Rev. Lett.* 81, 1562 (1998) [arXiv:hep-ex/9807003].

[5] C.H.Albright and M.C.Chen, *Phys. Rev.* D77, 113010 (2008).

[6] H1 Collaboration, F.D.Aaron *et al.*, *Phys. Lett.* B701, 20 (2011).

[7] H1 Collaboration, A.Aktas *et al. Eur. Phys.* J. C52, 833 (2007).

[8] W.Buchmuller, R.Ruckl, and D.Wyler, *Phys.Lett.* B191, 442 (1987).

[9] ZEUS Collaboration, S.Chekanov *et al. Phys. Rev.* D65, 092004 (2002).

[10] ZEUS Collaboration, S.Chekanov *et al. Eur. Phys.* J. C44, 464 (2005).

[11] BABAR Collaboration, B.Aubert *et al.*, *PHYS. Rev. Lett.* 104, 021802 (2010).

[12] BELLE Collaboration, Y. Miyazaki *et al.*, *Phys. Lett.* B660, 154 (2008).

[13] D.Boer *et al.*, Gluons and the quark sea at high energies: distributions, polarizations, tomography; arXiv:1108.1713

[14] Jefferson Lab website

[15] Matthew Gonderinger and Michael J. Ramsey-Musolf. 'Electron-to-tau Lepton Flavor Violation at the Electron-Ion-Collider'; arXiv:hep-ph/1106.5063 (2010).

[16] C Kolda, 'Theory of Leoton Flavor Violation: A Mini-Review' *AIP Conf. Prof.* 1182, 652 (2009)

[17] M.A. Doncheski and R.W.Robinett, 'Third-generation leptoquark decays and collider searches', arXiv:hep-ph/9707486 (1997).

[18] EIC Wiki (webpage), The EIC group at Brookhaven National Laboratory (2012). [https://wiki.bnl.gov/eic/index.php]

[19] S.Agostinelli *et al.*, 'GEANT4 - a simulation toolkit'. Nuclear Instruments and Methods in Physics Research A 506 250-303 (2003).

[20] https://fairroot.gsi.de

[21] C.Woody *et al.* The Calorimeter R&D Consortium, 'Joint Proposal to Develop Calorimeters for EIC'.

[22] D.Ryckbosch *et al.* The HERMES RICH detector HERMES Collaboration. Nuclear Instruments and Methods in Physics Research Section A: Volume 433, Issues 1-2, 21 August (1999).

[23] Torbjorn Sjostrand *et al.*, 'PYTHIA 6.4 physics and manual' (2006).

[24] Lorenzo Bellagamba, 'LQGENEP: a Leptoquark generator for ep scattering" (2010).

[25] Foreman, W.M; Sensitivity to Lepton Flavor Violation via leptoquark Exchange at an Electron-Ion Collider. stony Brook University (2012)

[26] G.Sterman & S.Weinberg, *Phys.Rev.Lett.* 39 1436(1977).

[27] Jade Collaboration, W.Bartel *et al.*, *Zeit. Phys.* C33 23 (1986); S.Bethke *et al.*, *Phys. Lett.* 213B 235 (1988).

[28] Dokshitzer Y.L, Leder G.D, Moretti S, Webber B.R, *JHEP* 9708, 001 (1997) [hep-ph/970723].

[29] Fabri A. *et al.*, *Softw. Pract. Exper.* 30 1167 (2000)

[30] Voronoi, Georgy. "Nouvelles applications des paramtres continus la thorie des formes quadratiques". Journal fr die Reine und Angewandte Mathematik 133: 97178 (1908)

[31] Catani, S.; Dokshitzer, Y. L.; Seymour, M. H. & Webber, B. R., 'Longitudinally-invariant k$_p$*erp*-clustering algorithms for hadron-hadron collisions', *Nucl. Phys.* B 406 (CERN-TH-6775-93. LU-TP-93-2) , 187-224 .(1993)

[32] Ellis, S. D. & Soper, D. E., 'Successive Combination Jet Algorithm For Hadron Collisions', *Phys. Rev.* D48 No.7 (1993)

[33] Cacciari, M.; Salam, G. P. & Soyez, G., 'The anti-k_t jet clustering algorithm', *JHEP* 04 , 063 (2008).

[34] Cacciari, M.; Salam, G. P. & Soyez, G., 'FastJet user manual', *Eur. Phys.* J. C72 , 1896 (2012).

[35] Cacciari, M. & Salam, G. P., 'Dispelling the N^3 myth for the k_t jet-finder', *Phys. Lett.* B641 , 57-61 (2006).

[36] Peterson.C; 'Neural Networks in High Energy Physics' LU TP 92-93 (1992)

[37] Forte. S;, Garrido. L;, Latorre.J.I; & Piccione. A . 'Neural Network Parameterization of Deep-Inelastic Structure Functions'. arXiv:hep-ph/0204232

[38] Private Communication: Mohammad Al-Turaney, Alexander Kisslev