Dijet production at the LHC at next-to-leading order

Zachary Warner

B.S. Physics, University of Kansas, 2018

B.S. Astronomy, University of Kansas, 2018

Submitted to the graduate degree program in Department of Physics and Astronomy and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Master of Science.

Professor Christophe Royon, Chairperson

Committee members

Professor Michael Murray

Professor Caroline Jewers

Date defended: December 03, 2020

The Thesis Committee for Zachary Warner certifies
that this is the approved version of the following thesis :

Dijet production at the LHC at next-to-leading order

Professor Christophe Royon, Chairperson

Date approved: December 03, 2020

Abstract

This thesis presents an analysis of NLO inclusive dijets produced form proton-proton collisions at center of mass energies $\sqrt{s}$ = 13 TeV. These dijets were simulated using POWHEG+PYTHIA8 parton shower Monte Carlo event generators. The jets were reconstructed using FastJet and the anti-k_t jet clustering algorithm. The purpose of this analysis is to contribute in the calculation of the ratio of jet-gap-jet events divided by the number of inclusive dijet events, a ratio highly sensitive to the effects that have been predicted by the Balitsky–Fadin–Kuraev–Lipatov (BFKL) evolution equation. These effects have proven difficult to separate from other effects predicted by perturbative QCD. The cuts applied for the jet-gap-jet selection process and the observables chosen for study were in accordance with a recent preliminary analysis preformed by the CMS Collaboration at $\sqrt{s}$ = 13 TeV. The main observables were the difference in pseudorapidity of the two leading jets, $\Delta\eta_{jj}$, the azimuthal angle separation of the two leading jets, $\Delta\phi_{jj}$, the momentum of the subleading jet, p_{T2}, and the number of charged particles produced in the "gap" region of jet-gap-jet events. The results of our analysis found several interesting features. The first was that $\Delta\phi_{jj}$ was found to deviate from the back-to-back topology found in events produced at LO accuracies. We believe this is caused by NLO corrections that allow for the generation of three partons at the parton level instead of the maximum two partons found in LO calculations. We also found that our Monte Carlo event generators, POWHEG+PYTHIA8, were producing an overabundance of low-energy charged particles. It was not until we reached a minimum transverse momentum for the charged particles of 1.0 GeV that the distribution of the number of charged particles began to behave as expected. We conclude that further analysis needs to be done to deduce the cause of the surplus of low-energy charged particles, and more NLO inclusive dijets events should be produced to reduce statistical uncertainty and reveal subtle features in the distributions of the obsrevables.

Acknowledgements

First I would like to thank my colleagues at the University of Kansas. Dr. Christophe Royon for your willingness to work with me when I started out knowing so little. You gave me a chance to grow as a physicists, and your advice lead me to working on projects that expanded my skills and prepared me for the future. Cole Lindsey for helping me set up all of the computer programs necessary for running and analyzing particle physics simulations, your amazing chocolate chip cookies, and for teaching me to relax...everything will be alright. Tommaso Isidori, whose friendship has been invaluable to me both in development in physics and in life. I am so grateful that you convinced me to train Brazilian Jiu-Jitsu, and it is something that I will cherish for rest of my life. Federico Deganutti, for your help in understanding the theoretical aspect of my analysis, and for the discussion of the global economy. Justin Williams, for always being an enjoyable presence to have in the office, and for showing me how to disk golf. I would also like to extend my endless gratitude to Cristian Baldenegro-Berrara, without your advise throughout the entirety of my graduate career, and on writing the thesis, I would have been lost soul. I have no doubt you will go on to do great things in physics and any greasy-grad student with you as their advisor should consider themselves lucky. Thank you to everyone else in the office (Dr. Georgios Krintiras, Dr. Alexander (Sasha) Novikov, and Florian Gautier) for listening to my presentations and research updates, and for giving feedback to help me improve.

I would also like to thank my colleagues at the University of Muenster (Dr. Jens Salomon, Dr. Michael Klasen, Pablo Gonzalez,and Mats Kampshoff) for your advice on this project, and for working with a graduate student still learning the ropes. Especially Dr. Jens Salomon, your patients with my endless questions, and advice in the setup and execution of the Monte Carlo programs, saved me from weeks of heart ache and endless toiling.

Thank you to my Thesis Committee for your willingness to read this document and sit in on

my presentation. I know your time is valuable, and the energy it takes to read and sit in on the presentation for a graduate student's thesis does not go unnoticed.

Lastly I would like to thank my wife who I relied heavily on for support not just during the writing of my thesis, but during my entire graduate career. There is no way I could have accomplished all I have without you by my side. Also, thank you to my two month old son, Matias. Your very existence gives me more motivation than all the espresso in the world.

Contents

List of Figures

Preface

My interest in nuclear physics began as an undergraduate student while I was taking classes in nuclear and particle physics. Like all physics students who have several years of training under their belt, I was familiar with electric charge carried by particles like electrons and protons. Thanks to introductory quantum mechanics and electricity & magnetism, I felt confident in my understanding of how particles interacted with one another through properties like their charge and spin. Then came nuclear physics. I learned that some of these particles not only contained electric charge, but they also contained color charge. Unlike electromagnetism where the force carrier particle (the photon) is electrically neutral, the particle responsible for carrying the strong nuclear force (the gluon) is *not* color-neutral. It, holds a combination of two different color charges. This means that while the gluons were interacting with color charged particles like quarks, they were also interacting with themselves! How could this be, and what were the implications? With only introductory knowledge from my undergraduate courses, I decided that I could not leave this new concept alone and must learn more. A colorful new outlook on the subatomic world had captured my curiosity.

As it turns out, color is a subtopic of a larger field of nuclear physics called Quantum Chromodynamics (QCD), a topic we will discuss in much more detail throughout this paper. Graduate school presented the opportunity for me to explore QCD and color using a feature of QCD known as jets[1]. I began my dive into nuclear physics by studying events containing pairs of jets called Mueller-Navalet Jets. Interesting events capable of revealing nuclear physics' leading theory models of high-energy regions [8]. I was also able to help with the preliminary investigation of dijets, meaning two jets, found in collisions between protons and lead nuclei. These jets made use of a detector at the Large Hadron Collider (LHC) called CASTOR, a subdetector of the Compact Muon

[1]For now, they can be thought of as a narrow cone of particles that are generated from high-energy particle collisions. Later we will give them a more rigorous definition.

Solenoid (CMS) experiment, and a detector which holds a special place in my heart as I was fortunate enough to be part of the data taking process during its final operational run in 2018. The results of this preliminary investigation into dijets were used in a Department of Energy grant proposal. Sparing much detail, my aim was to show that data for the dijets we wished to study was behaving as expected, meaning that the angle of separation for the jets in question would peak at pi. For the curious reader, the results showing that this was the case can be seen in Figure 1.

Upon changing directions in my graduate education towards a Master's degree, my research team and I found a sub part of these projects that was feasible to investigate in my time frame. This allowed me to continue to do research that I found interesting while coinciding nicely with my added Subspecialty in Computational Physics and Astronomy now accompany my Master of Science in Physics. As part of an ongoing project with members of my research team at the University of Kansas (Dr. Christophe Royon, Cristian Baldenegro-Barrera, and Federico Deganutti) and colleagues at the University of Muenster (Dr. Jens Salomon, Dr. Michael Klasen, Pablo Gonzalez, and Mats Kampshoff) I will be running high-energy particle physics simulations of dijet events like those that are found from proton-proton collisions at the LHC. Particle physics simulations contain all of "known" physics that we believe govern these collisions. Careful comparisons of simulations to real events are capable of giving insight into the accuracy of our understanding of what goes on at particle accelerators and in the Universe.

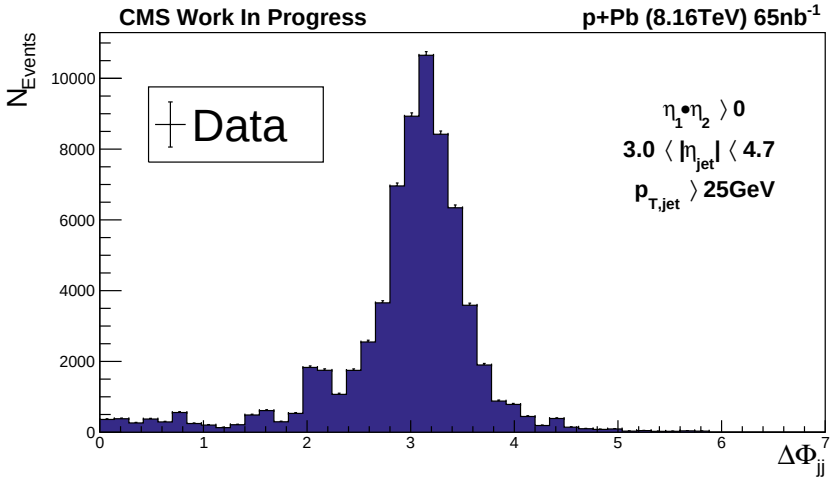

Figure 1: Azimuthal angle correlations between the leading pT jets in proton-lead collisions. This plot was created with CMS data.

Chapter 1

Introduction

1.1 Introduction

My aim in this section is to introduce the topic of study in this thesis, and familiarize the reader with the theoretical underpinnings of the topic and computational tools used to produce the simulated data for this analysis. While some of the terminology may seem technical to those uninitiated in nuclear physics, my hope is that by introducing the more intimidating terminology early and often the reader will feel more confident with the lingo as we dive into more details throughout the thesis.

In general, nuclear physics is the study of nuclei of atoms, and the forces that hold them together. It would be impossible to have a thorough discussion of nuclear physics without the strong interaction. Our best understanding of what is going on inside the nucleus of atoms comes from the strong interaction which is described by QCD. Precise and elegant, it has been extremely powerful in explaining the workings of fundamental particles called quarks at small distances where physically measurable quantities of the quarks can be computed. These calculations are done by applying perturbation theory to develop QCD as a function of the strong coupling constant,α_s, from the strong interaction. As successful as QCD has been, there remains interesting kinematic regions where predictions made by perturbative QCD have yet to be confirmed. Exploring these regions is important for understanding the initial state of collisions, like proton-proton collisions, and for studies of particle scattering at high energies.

This is where the project of this thesis comes into play, and the importance of dijets becomes apparent. The explorations of these previously unconfirmed kinematic region involve higher-order

corrections to terms in QCD equations that describe the evolution of the system. Predictions made by these higher order corrections suggest that the production of two jets separated by a large angle of measurement inside the detector, is a reasonably good process to understand the behavior of strong interactions in the high energy limit of QCD. This processes, known as jet-gap-jet or Mueller-Tang jets [8], is a great candidate for observing predicted effects we wish to investigate at the higher energy limits of QCD. To study these effects, it was proposed by Mueller and Tang to take ratio of the jet-gap-jet events to so called inclusive (to be defined later) dijets. A cancellation of various uncertainties from theory and experiment take place in the ratio of yields, which could potentially allow for further model discrimination [8]. The challenge comes with the fact that, not only does one need a good description of the events contributing to the numerator, jet-gap-jet, you also need a good handle on the events in the denominator, inclusive dijet events. Therefore, in order for jet-gap-jet to be a good QCD observable for high-energy scattering, we need reliable predictions for the inclusive dijet production.

As an added incentive, the most frequently occurring of all hard scattering processes in proton-proton collisions is dijet production. They are the benchmark measurement of any collider physics experiment. For example, jet measurements and calibrations play an important role in the determination of the jet energy scale [9]. From the standpoint of QCD, jet cross sections provide information about the strong coupling constant, leading to further insight about the quark and gluon content of the proton. Since they are a dominant final state object in proton-proton collisions, dijets act as a background to searches for physics events beyond the Standard Model of particle physics[9, 4]. Given the broad spectrum of applications for dijet productions, and the coming high luminosity upgrades to the Large Hadron Collider (LHC), it is apparent that we need well established theoretical models.

Together with colleagues at the University of Muenster and with colleagues at the University of Kansas, we have been investigating the effects of completing the calculation for the QCD evolution equations known as Balitsky–Fadin–Kuraev–Lipatov (BFKL) and developing a robust prediction for the inclusive dijet. My input in this collaboration has been mostly involved in the latter.

In this paper, I will analyze inclusive dijets produced from proton-proton collisions at center of mass energy $\sqrt{s} = 13$ TeV[1]. These dijets have been simulated using POWHEG+PYTHIA8 [10], parton shower Monte Carlo event generators. The purpose of this analysis is to contribute in the calculation of the ratio of jet-gap-jet events divided by the number of inclusive dijet events, a ratio highly sensitive to the effects that have been predicted by the BFKL evolution equation which are very difficult to separate from other effects predicted by perturbative QCD [11]. More broadly stated, we we want to know how we can better understand QCD and the strong interaction through the study of dijets.

1.2 Structural Overview

To begin this paper, we will map out the theoretical framework underlying and motivating this analysis by defining QCD, segueing into perturbative QCD models. This will allow us to better understand jets/dijets, and shed light on why we want to use jet-gap-jet to inclusive dijet ratios to search for BFKL effects. We will then briefly discuss the LHC and its Compact Muon Solenoid (CMS) detector, the detector effects which are considered and simulated during data production. The computational tools for generating the data and reconstructing the jets will be presented along with the kinematic observables used for analysis of the dijets and why. Finally, we will discuss the results of our analysis before concluding with a summary of the paper and future prospects for both this project and myself.

[1]The unit of measurement electronvolt, eV, is defined as the amount of kinetic energy gained by a single electron in a vacuum accelerating from rest through an electric potential difference of one volt.

Chapter 2

Theory and Motivation

2.1 Quantum Chromodynamics (QCD)

As a theory, the Standard Model of particle physics, Figure 2.1, has been successful in classifying all of the known elementary particles and describes three of the four fundamental forces of nature[1]. These fundamental forces describe all the observed interactions between particles, with the Standard Model explaining the particles that give rise to the electromagnetic, weak and strong interactions. Of these four forces our focus is going to be on strong interactions between quarks and gluons, QCD.

By the 1960s the idea of quarks as fundamental building blocks of nature was gaining wider acceptance, with the physical evidence being solidified when particle accelerators could provide deep-inelastic scattering experiments of electrons off nucleons [12]. Through these events, experimental physicists found that quarks acted as individual "free" particles inside the nucleon, but these quarks were never found to exist in nature on their own without the presence of at least one other quark. This intrigued experimentalists, and two lines of reasoning indicated that there was a missing piece to the puzzle. For one, people realized that the wave function which described the quantum state of quarks inside the nucleus did not seem to obey the requirements of the Pauli exclusion principle. The second came from the decay rate for a type of hadron (meaning two or more quarks bound together), called a neutral pion. The neutral pion decaying into two photons, $\pi^0 \rightarrow \gamma + \gamma$, came out to be $3^2 = 9$ times smaller than the decay rate found by observation. The missing puzzle piece

[1]The Standard Model describes the electromagnetic, weak, and strong interactions. However, it does not include a description of gravity.

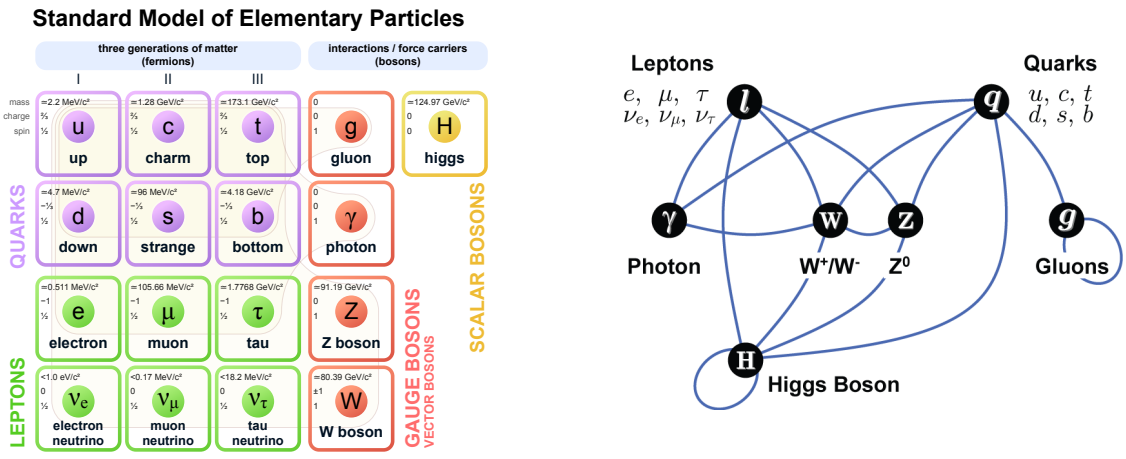

Figure 2.1: (Left) All of the elementary particles along with their basic properties. (Right) Allowed interactions between the particles. These figures were obtained from Refs. [1] and [2] respectively.

was found through the introduction of a property of quarks and gluons known as color [13]. It was found that quarks and gluons could carry one of three color charges (red, green, or blue). In addition to their color charge, quarks also carry a fraction of electric charge. Thus, we can now accurately describe the wave function of quarks in the nucleus through the introduction of color while, simultaneously, solving the 3^2 discrepancy in the neutral pion decay. Pions contain one quark and one anti-quark, with each quark and antiquark containing one of three possible colors. This means we now have $3^2 = 9$ different possible combinations. Voila! With the introduction of color charge in the late '60s and early '70s we had solved issues brought forth from deep-inelastic scattering studies and QCD was born.

What began as a means to describe the ordinary matter like protons and neutrons evolved into a foundation of our understanding of high-energy particle accelerators. QCD is a modern theory, well established at describing the interactions of the quarks which come in six flavors (up, down, top, bottom, charmed, and strange) as well as their force carrier, the gluon. It should be noted that for each quark there also exists an antiquark which carries anticolor charge. These quarks and gluons are the fundamental building blocks of hadrons. For example, at lower energy scales a familiar hadron, the proton, is made out of two up quarks and one down quark. Without the mediation of the strong interaction by gluon exchange between quarks, antiquarks, and other gluons, hadrons

could not form from the confined quarks [14]. QCD also explains why protons and neutrons stay bound to one another inside atomic nuclei by means of exchanging quark antiquark pairs, despite protons repelling one another through their shared positive charge.

To truly understand how QCD explains strong interactions we need to become more familiar with the concept of color charge. Of course the term *color* is just clever nomenclature, in actuality the three color charges have properties more analogous to electric charge which is so elegantly defined in QCD's complimentary theory, Quantum Electrodynamics (QED). For starters, both color charge and electric charge are measurably conserved quantities in physical processes. Like the photon, the massless force carrier for the electromagnetic force, the gluon is a massless particle that acts as the force carrier for the strong nuclear force. However, key differences between the gluons of QCD and the photons of QED create properties that appear truly unique to QCD. One interesting observation about gluons is that they respond to the presence of color charge more intensely than photons respond to the presence of electric charge. Also, gluons are capable of changing one color charge into another. We can conclude from these observations that gluons must themselves carry a color or anticolor charge. This is in contrast to photons that carry no electric charge. We have identified a quirky feature of QCD: gluons that carry and respond to color charge must also be capable of interacting with other gluons [15]. Quarks also carry color charge, so why do protons not carry color charge if they are made of quarks? The reason is the three quarks that make up a proton are composed of red, green, and blue color. These colors cancel out to create the color neutral proton similar to how charge neutral atoms contain the same number of negatively charged electrons and positively charged proton.

Of course it would be an injustice to discuss QCD without writing the Lagrangian density that describes the theory of the strong interaction,

$$\mathscr{L}_{QCD} = -\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} + \sum_j \bar{q}_j(i\gamma^\mu D_\mu - m_j)q_j \quad . \tag{2.1}$$

Where

$$D_\mu = \delta_\mu + igA^a_\mu T^a \tag{2.2}$$

is what is known as the gauge covariant derivative, and

$$F^{\mu\nu}_a = \delta^\mu A^\nu_a - \delta^\nu A^\mu_a - g f_{abc} A^b_\mu A^c_\nu \tag{2.3}$$

is the field strength tensor of the gluon. Here m_j is the quarks mass and q_j is the the quarks quantum field, where the flavor of the quark is indicated by the index j. The indices a, b, c keep track of the color and the spacetime indices are μ, ν. The gluon field is A and the self interaction of the gluon can be seen in the $A^a_\mu T^a$ term from equation 2.2. We can also see that all the interaction terms are proportional to the strong interaction coupling constant, $g = \sqrt{4\pi\alpha_s}$. Lastly, to ensure that the equation is symmetric under the so called SU(3) gauge group, something that must be true if QCD is to remain unchanged when quark and gluon fields transform under SU(3), the coefficients f and T are introduced [13]. That is it! While it is true that the QCD Lagrangian leads to notably difficult equations to solve, 2.1 and the subsequent equations contain all the information about the strong interaction.

Despite having a Lagrangian that describes the theory, there are regions of QCD phase space[2] that have yet to be confirmed. The reason comes from two fundamental properties of QCD that make it intriguing, but the calculations difficult:

- **Asymptotic freedom**: This means that the coupling and interactions of quarks and gluons decreases at large energy scales. At these high energies quarks inside of hadrons act as if they hardly interact with one another [12].

- **Color Confinement**: Quarks and gluons cannot be isolated in nature. When objects like partons[3] are spatially separated by ever increasing energies it eventually becomes energet-

[2]Phase space references a space which looks at multiple kinematic observables of a particle. Classically this typically refers to position and momentum space.

[3]A model of quarks and gluons introduced by Richard Feynam to explain observations made during high-energy particle collisions and QCD processes.

ically favorable to create a quark-antiquark pairs as you increase the distance, rather than elongating the "color tube" connecting the quarks. Color Confinement also means that at small energy scales the quarks and gluons stay confined inside hadrons [12].

Color confinement, while observed, has never been proven mathematically. It is believed that at very low momentum scales the theory becomes strongly coupled, and the use of non-perturbative techniques or calculations on the "lattice" (discretized spacetime) have supported the hypothesis of color confinement. This runs counter intuitive to what we learn in elementary physics where forces between interacting objects becomes smaller as we increase the distance between them. However, there is one classical analogy found in Maxwell's equation $\nabla \cdot B = 0$. Which states that there cannot exist a magnetic monopole. Any time you try to split a bar magnetic in half you do not have one north pole half and one south pole half. Instead you end up with two magnets with north and south poles. This is because it takes more energy to create a magnetic monopole than it does to recreate a magnetic dipole, similar to energetically favorable quark-antiquark pair production in color confinement [13].

Asymptotic freedom hints at a "running" coupling constant, meaning that the strong coupling constant becomes small at high energies and very small distance. This allows us to use perturbation theory to approximate this complex quantum system with a more simple system, which brings us to the subfieled of QCD know as perturbative quantum chromodynamics (pQCD)

2.2 Perturbative QCD: BFKL & DGLAP Evolution

As previously mentioned, perturbative QCD involves calculations for quarks and gluons at high energy scales and small distance by applying perturbative expansion to powers of the strong coupling constant. A simple example of perturbation technique and the order of perturbation can be shown by considering E to be the exact solution to a complex problem. Applying perturbation

with respect to small perturbed parameter, call it β, gives us

$$E = E_0 + \beta E_1 + \beta^2 E_2 + \beta^3 E_3 + ... + \beta^n E_n \quad . \tag{2.4}$$

In this equation E_0 would give us our initial state solution, E_1 gives us our first-order solution, E_2 gives us our second order solution, and so on up to some power of n. This means that an approximate second-order solution for E would give $E \approx E_0 + \beta E_1 + \beta^2 E_2$. In this paper we will use the common nomenclature of referring to first order as leading order (LO) and second order as next-to-leading-order (NLO), where the expansion is usually done in powers of α_s when $\alpha_s << 1$.

In a similar manner we can approximate the solutions to QCD using fixed-order perturbative expansion. This involves the calculation of all the processes and interactions taking place between the quarks and gluons that contribute to the cross section of an interaction to a given expansion order n of α_s. Here, the cross section is the probability a process will take place during a given interaction.

A good deal of the time fixed-order perturbative expansion applied to QCD is sufficient to explain what is going on at the level of the parton. However, sometimes we enter regions of phase space where fixed-order approximations break down due to large α_s values. The solution to this problem is to sum up all the contributions for α_s^n where n now goes from one to infinity instead of from one to some fixed value. This is what is referred to as an all-order resummation, and is what the BFKL and Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) parton evolution equations do in their respective kinematic regimes [16, 17]. These two equations form the basis of our understanding for high energy scatterings in QCD.

DGLAP is a parton evolution equation where parton emissions are strongly ordered in transverse momentum[4]. It can be thought of as the sum of all the higher order contributions of α_s that have been enhanced through the multiplication of the logarithm of Q^2. Here Q^2 is measurement of the energy of a collision involving hadrons that has been transferred to the constituent partons, i.e. how

[4]The momentum of a particle or parton that is perpendicular to the beam line of particle accelerators.

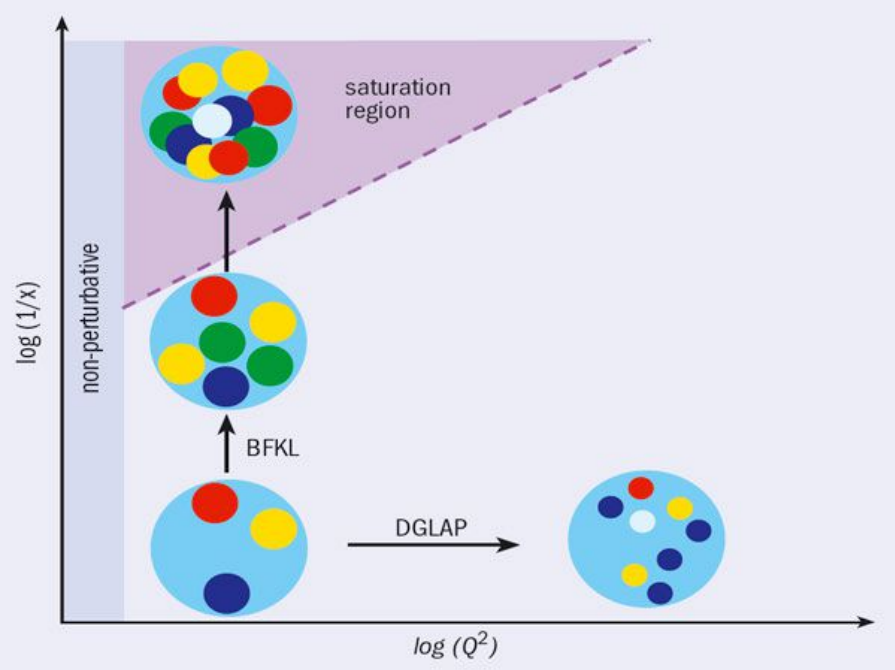

Figure 2.2: The horizontal and vertical directions correspond respectively to the DGLAP and BFKL regimes of QCD. The dash-diagonal line is the boundry to the saturation region. The thin vertical region to the left is the non-perturbative QCD region. This figure was obtained from Ref. [3].

deeply we are probing inside the hadron to observe more of its constituent partons[17]. In contrast, BFKL is a parton evolution equation where the parton emissions are strongly ordered in a spacial coordinate known as rapidity[5]. It has its higher order contributions of α_s enhanced by logarithm of square of the center of mass energy of the collision, s. For deep inelastic scattering events, it can also be thought of as being enhanced by the fraction of logarithm of x, the small momentum fraction carried by the gluons[16]. This variable, also known as the Bjorken-x scaling variable, is defined as in terms of Q^2 and s as $x = Q^2/(s+Q^2)$. The regimes where BFKL/DGLAP are relevant are portrayed nicely in Figure 2.2.

An offered interpretation of Figure 2.2 is that for the DGLAP regime as a proton is probed at increasing Q^2, the number of partons "seen" by the proton rises while the size of the partons decrease. By contract, in the BFKL regime when the fraction of momentum carried by the parton,

[5] An energy and angle dependent parameter of a particle or parton that is measured relative to the beam line of particle accelerators.

x, becomes smaller the number of partons (specifically gluons) increases until x is sufficiently small enough that gluon saturation can be achieved [18].

This power-law growth of gluons at low-x, along with it accounting for higher order terms, is the reason BFKL is considered the more complicated theory of the two. In fact, of the two parton evolution equations, studies have thus far shown no deviation from the DGLAP predictions [19, 20, 21], but they have also not not ruled out BFKL since they were operating in energy regimes where both BFKL and DGLAP effects are present. The early searches for BFKL signatures used hadron-hadron collisions, starting at HERA with the H1 and ZEUS experiments before moving on to the Tevatron D0 detector. They found a strong dependence on energies for dijet cross sections at large rapidity separations, but no BFKL effects were found [22, 23, 24]. However, this does not mean we are ready to give up on BFKL quite yet. These studies have also shown that the BFKL effects associated with multiparton splitting are very difficult to separate from other effects predicted by perturbative QCD. Therefore, we need a more restrictive final state like that offered by jet-gap-jet. To understand why this is true let us begin by defining jets and dijets before moving to jet-gap-jet and inclusive dijets.

2.3 Jets

To study high-energy collider physics, especially collisions between hadronic protons, we must consider the quarks and gluons that will be produced in the final state. However, we cannot directly observe these quarks and gluons. Instead we observe the hadrons that form out of these color confined quarks and gluons. When color charged particles are produced with high enough energies during a hard scatter process, it becomes more energetically favorable for these partons to create more color charged particles then to try to separate further. These newly formed color-charged particles are found in the vicinity of the initial partons. This process of producing radiation of color-charged particles from the initial state partons is known as a parton shower, and the process of combining the constituent quarks and gluons produced from the parton shower into hadrons is

known as hadronization[25]. Toghether they form the backbone of our definition of jets:

> Jets are collimated sprays of particles formed out of high-energy quarks and gluons that underwent parton showers followed by hadronization[26].

This allows us to define dijets as simply two jets produced in the final state that came from the same collision vertex. A good visual representation of dijets by the CMS experiment can be found in Figure 2.3. Inclusive jets are jets where we are not measuring all the particles that were produced in the final state, only those which are contributing to the measured jets. However, two jets in the final state is not the limit. Many jets can be produced in the final state, and for a proton-proton collision this can be written as $p + p \rightarrow j_1 + j_2 + ... + jn + X$ where j_1, j_2,...,j_n are the measured jets and X is everything we did not measure.

While this might seem like a simple enough definition, in practice defining and measuring the jets at the detector level is much more difficult. For example, how do we know if two hadrons that are in close proximity to one another at the detector level should be counted as part of the same jet originating from the same parton? This problem magnifies when considering that we are not colliding one proton with one proton, but "bunches" of protons with other bunches of protons in what is referred to as a bunch crossing. This requires consistent and well defined definitions and algorithms for reconstructing the jets from the hadrons that we measure in the final state. We will discuss these algorithms in more detail in the subsequent chapters when we discuss the computational tools used in this analysis.

What about jet-gap-jet, and how can we use inclusive dijets and jet-gap-jet events to probe BFKL physics? As we mentioned in our discussion of perturbative QCD, we need a more restrictive final state to probe BFKL since BFKL effects associated with multiparton splitting are difficult to separate from effects of other higher order perturbative QCD corrections. This BFKL multiparton splitting is strongly ordered in rapidity and is known to dominate at large rapidity separations between leading (most energetic) and subleading (second most energetic) jets [27]. This can be juxtaposed with DGLAP which multiparton splitting is strongly ordered in transverse momentum

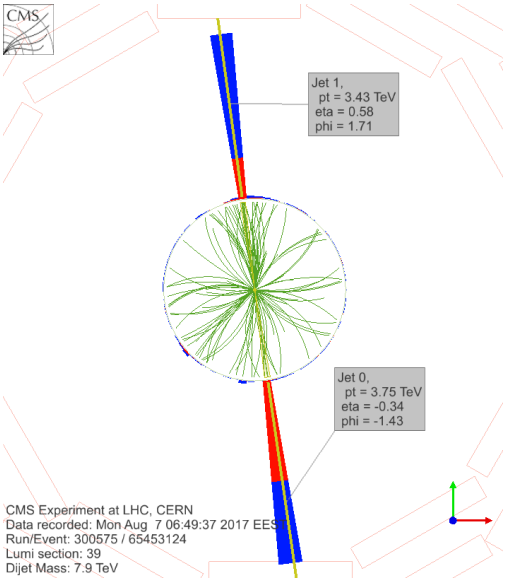

Figure 2.3: A 2D representation of dijet produciton at the LHC measured with the CMS detector. This view is as though you are looking down the beam line. This figure was obtained from Ref. [4].

and is not a function of rapidity [28]. Therefore, events that allow us to go to large rapidity separations will get us away from the regimes where DGLAP takes effect so we can observe the effects of BFKL. This occurs, for example, with the jet-gap-jet events. These events are very "clean" processes since they are defined as having a gap between two jets where this gap does not contain a particle [8].

The reason we anticipate the ratio of jet-gap-jet to inclusive dijets to be highly sensitive to dynamical effects predicted by BFKL has to do with the way color singlet exchange[6] occurs between partons. For inclusive dijets color singlet exchange between partons results in particle production over a wide range of rapidity between jets. In BFKL the color singlet exchange, which can occur between quark-quark, quark-gluon, and gluon-gluon, is thought to be predominantly dominated by t-channel[7] two gluon exchange between partons [29], as represented in Figure 2.4. Therefore, taking the ratio between them is expected to reveal effects predicted by the BFKL evolution equations [8]. In the chapter where I discuss the analysis I will present which specific dynamical variables

[6]Particles like protons and neutrons that contain color charged particles, but are themselves color neutral are said to be in a color singlet state

[7]When measuring momentum exchange between particles we are considering the t-channel. Compare this to the measurement of center of mass energy exchange known as s-channel

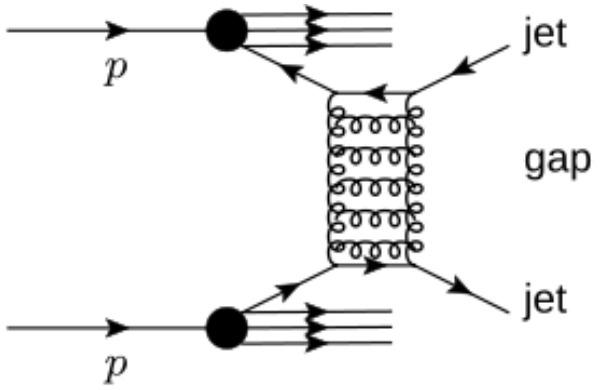

Figure 2.4: Feynman diagram of a jet-gap-jet event from t-channel two-gluon exchange in proton-proton collisions. The three lines to the right of the protons represent the breakup of the protons. This figure was obtained from Ref. [5].

we expect to reveal these BFKL effects and why.

Chapter 3

The CMS Experiment at the LHC

3.1 The Large Hadron Collider (LHC)

Located between the France-Switzerland border at the European Organization for Nuclear Research (CERN) accelerator complex, the LHC, Figure 3.2, is the largest and most powerful particle accelerator in the world. In fact, it is largest machine ever built requiring a massive collaboration of thousands of scientists and engineers, and involving over 100 countries and even more universities from around the globe. The collider consists of a 27 kilometer ring of thousands of superconducting magnets and several accelerating structures that are used to boost the energy of the particles, all of which is buried in a tunnel over 150 meters below the surface. These superconducting magnets create a strong magnetic field for the purpose of guiding two high-energy beams of charged particles traveling in opposite directions and intersecting at various collision points, where detectors can be installed to detect the debris from the collision. These beams are smashed together at high-energies in several multipurpose advanced detectors around the LHC for measurements of particles and events both common and rare[30].

The primary purpose of the LHC was to discover the Higgs boson, an excitation of the Higgs quantum field that permeates the Universe. The Higgs boson was long believed responsible for giving mass to bosons, the force carries for the weak interactions, and fermions, the fundamental building blocks of matter. This search finally paid off in 2012 when two independent papers published by both the CMS and ATLAS experiments at the LHC discovered a scalar particle with a mass of 125 Gev, and properties compatible with the Higgs boson [31, 32]. However, the story

of the LHC certainly is not over. We know that the Standard Model of particle physics is not complete and there is a need for new physics. Having reached center of mass collision energy of $\sqrt{s} = 13$TeV, higher than any collider ever created, the LHC aims to go beyond the Standard Model by reaching energy scales that mimic moments just after the Big Bang. These conditions make the data a valuable resource for probing the strong interaction of QCD and understanding how our Universe works and came to be.

To aid in these new discoveries the LHC is planned for a shutdown to impliment the High-Luminosity LHC upgrade. The objective of this upgrade is to increase the luminosity[1] of collision at the collider by a factor of 10 beyond the LHC's design value. It also aims to reach $\sqrt{s} = 14$ TeV energy [33]. This high luminosity and energy means more data to use in searches for rare events, data that creates a unique challenge requiring both sophisticated computational analysis to reconstruct events, as well as detector hardware capable of detecting inelastic events at a rate of up to a billion events per second [34].

3.2 The Compact Muon Solenoid (CMS)

Smashing particles together at energies close to that of the Big Bang and observing the results is an incredible engineering feat, one requiring sophisticated detectors capable of recording all of the different types of particles and events produced in these collisions. Not only that, but these detectors must having tracking and timing systems capable of tracing particles back to their point of collision, also known as their primary vertex. The CMS experiment at the LHC is one such detector. A multipurpose detector, Figure 3.2 shows how it is actually multiple layers of different detectors, all when put together are capable of achieving these goals for collisions at the TeV scale. Since the recent jet-gap-jet paper that serves as the resource for our analysis was done in collaboration with CMS and using CMS data, let us explore these layers of CMS in a bit more detail to get a better understanding how this detector works.

[1] Luminosity is a concept in accelerators that is proportional to the number of collisions occurring per unit time.

Figure 3.1: To give scale to the magnitude of the LHC, this image shows the above ground outline of the collider and locations of the the main detectors. This figure was obtained from Ref. [6].

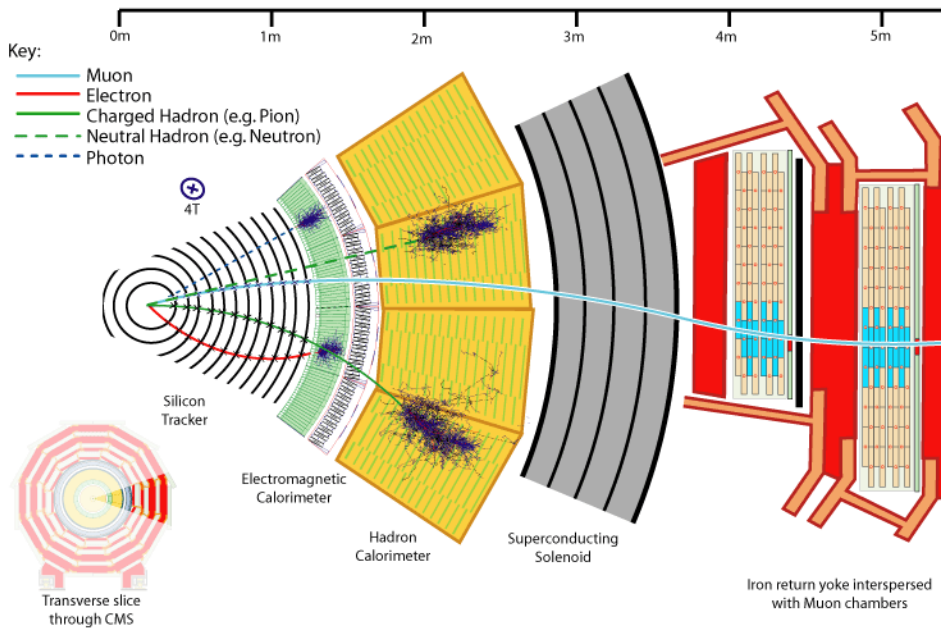

Figure 3.2: A slice of the CMS detector showing the different layers of the detector to scale. Included are examples of what types of particles would be detected in each layer. This figure was obtained from Ref. [7].

The first layer we come across is the silicon tracker. It is designed to provide precise and efficient measurements of the trajectories of charged particles immediately following the hadronic collisions that took place at the center of the detector. This is achieved by recording electrical signals triggered by the moving charges. Without this layer we would not be able to track charged particles stopped later in the detector back to their primary vertex, or be able to properly reconstruct secondary vertices[2] [34].

The next two layers are our hermetic homogeneous calorimeters. A fancy way of saying blocks of the same type of instrumented material in which particles are fully absorbed and their energy transformed into a measurable quantity, and that these blocks cover the full azimuthal range in the detector. These two layers are known as the Electromagnetic Calorimeter (ECAL) and the Hadronic Calorimeter (HCAL). The first of the two is ECAL which consists of over 60,000 lead tungstate crystals, capable of measuring energies of photons and electrons to a high precision [35]. Hadrons pass right through the lead tungstate crystals and enter the HCAL which is capable of stopping and measuring the energies of these hadrons. As these particles pass through the layers of brass material rapid pulses of light are triggered indicating information about arrival time, position, and energy [36].

So far every layer we have discussed exists inside the CMS's massive superconducting solenoid. This is because the measurements of these detectors are heavily dependent on the solenoid's extremely powerful 4.0 T homogeneous magnetic field which interacts with the charged particles in the inner region allowing us to precisely measure their momentum. More momentum means these particles will curve less in the detector, while less momentum means the particles will curve more due to the presence of the magnetic field. If the momentum of the particle is low enough it is bent so as to remove it from measurement. Of course particles with no charge will not interact with the magnetic field at all. This powerful magnetic field was achieved by a few unique features. The solenoid consists of 4 winding layers of conducting material co-constructed with pure aluminum and reinforced with aluminum alloy allowing a current of 41.7 MA-turn. It also has extremely

[2]A secondary vertex is the point at which a particle decayed into two or more new particles that are stopped elsewhere in the detector.

large dimensions with a 6.3 meter bore, 12.5 meter length, and a 220 ton mass [34]. All of these make the CMS's superconducting solenoid truly unique.

The outermost layer of detector, and arguably the most important as the Compact *Muon* Solenoid name suggests, is the Muon chambers. Muons are charged leptons like electrons, but they are 200 times heavier. Their measurements provide signatures for many important decays, including some decays of the Higgs particle. Muons and neutrinos are capable going through lots of material without interacting. Therefore, by placing the dense muon detector on the outer edge of the CMS detector, and inter-layering with iron plates, we increase the chance of measuring muon momentum, position, and charge as they pass through the muon chambers [34].

All of these layers working together make CMS one of the most precise and sophisticated detectors in the world. One that was capable of detecting the elusive Higgs boson for the first time in history, and will no doubt continue to lead the way to understanding new physics, including QCD, through tracking and detecting rare events produced at the LHC.

Chapter 4

Analysis

4.1 Event Generation

Due to advancements in mathematics and computing power, Monte Carlo (MC) event generators have become a staple in studies of particle physics. Making use of random number generators to obtain numerical solutions to complex problems, Monte Carlo simulations have allowed physicists to test their theoretical models by comparing the events they generate to the data from experiments, such as those conducted at the LHC. Their reliability allows modern collider experiments to use the simulated data produced by Monte Carlos to understand their detectors. With both current and upcoming data from the high-energy collisions at the LHC, event generators capable of NLO and even next-to-next-to-leading order (NNLO) QCD calculations play an important role for physicists wanting to probe the limits of theoretical models [37]. Whether the results of the simulated events match the actual data or not, there is much we can learn by implementing these event generators. If the simulated events do not match the data, then the Monte Carlo needs to be re-tuned or the models used to generate the events need to be reconsidered. Simulated events that match the data indicate accurate models, and could lead to the discovery of new physics. It is for this reason we will be using Monte Carlo event generators to compute the fraction of color singlet exchange dijet events, the main observable in the study of jet-gap-jet events by CMS. This requires a complete calculation at NLO for both the numerator, the jet-gap-jet events described by BFKL evlution equations, and the denominator, the inclusive dijet events described by fixed-order pQCD techniques.

In our case we will make use of Monte Carlo event generators capable of implementing NLO QCD

calculations for our inclusive dijet production. This will allow the data to be compared to NLO jet-gap-jet events generated by our colleagues at the University of Muenster who are implementing NLO BFKL calculations performed by our colleague Federico Deganutti. Due to the complexity of simulating hadronic collisions, showering the partons, hadronization of the partons, and production of final state particles, we will use two Monte Carlo event generators capable of handling the different processes. For the production of the hard scattering events from proton-proton collisions we will be making use of POWHEG, specifically POWHEG-V2, and the POWHEG-BOX computer framework. These hard processes will then be fed into PYTHIA8 [38] to supplement the parton-showering, hadronization, and final-state particle production. Originally we had planned to also investigate MG5_aMC@NLO (MadGraph5) [39], another MC event generator capable of NLO hard scattering events. However, our preliminary analysis found discrepancies between the jet cone radius, R, used by the NLO jet matching schemes in Madgraph5, and the value of R we would be using to reconstructing the jets for consistency with both the our Muenster colleagues and jet reconstructing at CMS. Their jet matching scheme was forcing us to use a value $R = 1.0$ to generate the jets. This is in contrast to our desired value of $R = 0.4$ used when reconstructing the jets for analysis. Due to the jet cone radius being defined as $R = \sqrt{\Delta\phi^2 + \Delta y^2}$ where $\Delta\phi$ is the difference in azimuthal angle for the cone, and Δy is the difference in rapidity for the cone, the discrepancy was leading to noticeable irregularities in our distributions. The main issue was caused by the double counting of jets for very small $\Delta\phi$ and Δy values. Since $0.4 + 0.4 = 0.8 < 1.0$ you are able fit two reconstructed jets inside of one of the generated jets. After further inquiry, the Madgraph5 team expressed a lack of confidence in Madgraph5's ability to perform NLO jet matching with dijets using a different jet matching scheme. They suggested that we switch to LO matching schemes if we wanted to be confidant in our events. While it should still be possible to successfully generate NLO inclusive dijets in Madgraph5, to ensure more confident in our generated events we decided to switch to POWHEG for our NLO hard scattered event production. As an added bonus of consistency, this is the same MC event generater being used by our colleagues in Muenster for the jet-gap-jet production using the BFKL calculations at NLO accuracy. Now that we have a general

layout of the process for particle collision event generation using Monte Carlo methods, let us take a deeper look at the Monte Carlos we used and our setup for inclusive dijet production.

The first step is to produce the hard scattering events from proton-proton collisions at a center of mass energy $\sqrt{s} = 13$ TeV. This is done using the POWHEG-BOX computer framework for implementing NLO QCD calculations in Monte Carlo programs using the POWHEG method. The main purpose of the POWHEG method is to generate hard events that can be fed into another shower Monte Carlo program for subsequent showering and hadronization [10]. Conveniently, the POWHEG-BOX comes with many built-in processes that can be accessed by users, including a jet pair production process that we will be using for our inclusive dijet production. These processes come with input files that are preset to LHC conditions, but allow the user to adjust key input parameters such as the center of mass energy of the collision, the minimum transverse momentum of the underlying events for jets that are produced at the level of hard scattering events, and the choice of the Parton Distribution Function (PDF)[1]. The PDF is needed since hadronic cross section for the generated events are calculated by means of the QCD factorization theorem, where the hard parton cross section is convoluted PDFs of the proton. All of the new features offered by NLO Monte Carlos expand the range of user customization, but also means one must carefully select their input parameters to obtain reliable results.

Running POWHEG successfully requires the use of two external libraries, Fastjet[40] and LHAPDF[41], as well as setting several parameters in the input file. FastJet is a software package that provides a wide range of jet finding tools, jet analysis programs, and necessary libraries for POWHEG to simulate dijet production from hadronic collisions . The Les Houches accord PDFs (LHAPDF) library is used for calling and implementing PDFs which are used to describe the partonic content of the hadrons produced in the collision. The PDF chosen from the LHAPDF library needs to be well understood to a sufficiently high enough precision to obtain theoretical predictions that will match LHC data . Our choice of PDF was NNPDF31[2], an up-to-date PDF set with NLO accuracies that

[1] the measure of the probability density for finding a particle with longitudinal momentum fraction x at a given resolution scale Q^2. This value cannot be calculated with perturbative QCD, and must come from measured data.

[2] The full name of the PDF is "NNPDF31_nlo_as_0118". We condensed this to NNPDF31 to make it easier on the

was tuned to LHC data, and the PDF of choice by our colleagues in Muenster. Fastjet is required during the setup of the POWHEG-BOX dijet process, and your choice of PDF from LHAPDF is set in the POWHEG input file along with the other input parameters.

Most of the POWHEG input parameters can be left at their default value, but there are a couple important parameters that we should mention due to their impact on the events produced. First, we chose a minimum transverse momentum of 10 GeV for the jets produced by the hard scattering events. This value is low enough that it is away from the kinematic region we wish to explore for inclusive dijets and jet-gap-jet analysis, while being large enough to produce events consistent with data. Figure 4.1 offers an example of what can happen if your choice of minimum transverse momentum for generated jets, p_{Tmin}, is too close to the minimum cut placed on the jets during analysis, p_{Tj}. In this case our choice of $p_{Tmin} = 30$ GeV caused the distributions of the leading jet to peak around 70 GeV, and the distribution of the subleading jet to peak around 55 GeV, instead of peaking closer to our of choice of 40 GeV for p_{Tj} as we would expect based on experiment. In the next chapter we will present distributions of the leading and subleading jet that were produced using an appropriate $p_{Tmin} = 10$ GeV which gives reuslts much closer to whats expected from data. We also, by recommendation of the POWHEG team, enabled the parameter "doublesfr 1". This parameter is offered as a fix to the spikes in histograms caused by events with very large weights and fairly large transverse momentum [42]. Once the hard scattering events are generated with POWHEG we are ready feed the produced events into another Monte Carlo for parton-showering and hadronization to obtain final state particles.

PYTHIA is a program for generating high-energy collisions between elementary particles. It is comprised of physics models for evolving a system from just a few-body hard processes like those produced by POWHEG to a multiparticle final-state. Inside the PYTHIA program there are libraries capable of generating hard processes for initial and final state parton showers, multiparton interactions, beam remnants, string fragmentations (hadronization), and particle decays. In our case we will be making use of the hadronization and parton showering capabilities by implementing

reader.

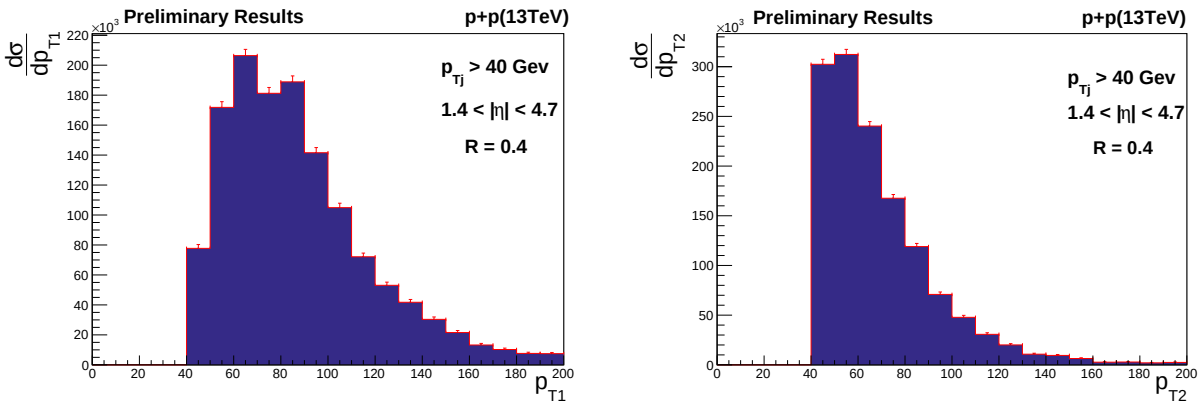

Figure 4.1: (Left) The differential cross section as a function of the p_T of the leading jet from dijet production when $p_{Tmin} = 30$ GeV. (Right) The differential cross section as a function of the p_T of the subleading jet from dijet production when $p_{Tmin} = 30$ GeV. These plots were created with data generated using POWHEG+PYTHAI8.

PYTHIA8 which is the most recent version of PYTHIA written in the C++ coding language as apposed to the older PYTHIA6 which was written in Fortran [38]. We use a special tune of PYTHIA8 for CMS underlying event measurements in accordance with a study done by the CMS collaboration [43], and for inclusive dijet events at NLO accuracies in accordance with Ref. [44]. We also used the same PDF in PYTHIA8 that we used in POWHEG, "NNPDF31". After running PYTHIA8 we are left with stable final state particles that can be used to reconstruct our jets for analysis.

4.2 Jet Reconstruction

To reconstruct jets from the stable particles that have been generated using POWHEG+PYTHIA8 we will make use of FastJet's "JetDefinition" jet clustering algorithm and recombination scheme for jet reconstruction and dijet event selection [40]. FastJet has several jet clustering algorithms to choose from, but the one we will be using is the anti-k_t algorithm. This is considered the standard inclusive jet finding algorithm for hadron-hadron collisions in high-energy physics [45]. FastJet also performs the recombination scheme for our jets which decides how to combine the four-momenta[3] of the particles merged during the clustering. Here we went with FastJet's default

[3]Four-momentum are four-vectors which are invariant under Lorentz transformations and represent a particle's relativistic energy and relativistic momentum in all three spacial dimensions (x,y,z).

E-scheme which simply adds the four-vectors. Lastly, we used the commonly accepted value of the jet cone radius $R = 0.4$ to maintain consistency with our colleagues in Muenster and previous CMS jet-gap-jet studies [5].

In essence what FastJet does is create a pseudojet vector out of the jets' four-momentum values (p_x,p_y,p_z,E) allowing for the user to easily access the different kinematic variables of the jet that are calculated using these four-momentum such as the transverse momentum, pseudorapidity, rapidity, and azimuthal angle[34]. Transverse momentum of the particles is simply defined the component of the three-momentum perpendicular to the beam line along the z-axis, i.e.

$$p_T = \sqrt{p_x^2 + p_Y^2}. \tag{4.1}$$

Rapidity is defined as

$$y = \frac{1}{2}\log\left(\frac{E + p_z}{E - p_z}\right), \tag{4.2}$$

where E and p_z are the energy and momentum beam-component of the jet. Pseudorapidity, η, is related to rapidity, but for situations were the particles mass is zero. Therefore, E is replaced with a particle's 3-momentum $|\mathbf{p}| = \sqrt{p_x^2 + p_y^2 + p_z^2}$ to give us

$$\eta = \frac{1}{2}\log\left(\frac{|\mathbf{p}| + p_z}{|\mathbf{p}| - p_z}\right), \tag{4.3}$$

Now, using our pseudojet vector, we are able to skim over all the inclusive jets created in each simulated event to select the two jets with the highest p_T as our leading and subleading jets for the inclusive dijet analysis.

4.3 Event Selection and Observables

We will now talk about the jet-gap-jet analysis, first introduced in Section 3 of Chapter 2. As anticipated in those sections, one is interested in studying events separated by a pseudorapidity

gap, which is a signature expected from hard color singlet exchange. The main observable in this study is the fraction of dijet events produced by color singlet exchange, i.e.,

$$FCSE = \frac{N_{CSE}}{N_{all\,jets}} \tag{4.4}$$

Where N_{CSE} are the events produced by color singlet exchange in BFKL, and $N_{all\,jets}$ are all the jets, dominated by standard QCD interactions between quarks and gluons. $N_{all\,jets}$ has to be calculated with the POWHEG+PYTHIA8 simulation introduced in Section 1 of this chapter.

To choose the leading and subleading jets that will constitute our dijet pair we apply several cuts to ensure our choice of inclusive dijets is congruent with the jet-gap-jet events. Such cuts follow the ones introduced by the CMS Collaboration on a recent preliminary analysis in Ref. [5].

- Our choice of leading and subleading jets are the most energetic and second most energetic jets respectively, with a minimum transverse momentum $p_{T\,j} > 40$ Gev.

- To ensure our jets are on opposite hemispheres of the detector we select inclusive dijets pseudorapidities that give $\eta_1 \times \eta_2 < 0$. We also want $1.4 < |\eta_{1,2}| < 4.7$, meaning our jets are in the region for phase space used for the production of jet-gap-jet events. Having pseudorapdity values greater than 1.4 ensures that they are at least one jet cone radius, $R = 0.4$, away from the region of no charged particles known as the gap site, $|\eta| < 1.0$, required for the jet-gap-jet events. All of these are in accordance with measurements preformed by our colleagues at the University of Kansas within the CMS Collaboration [5].

The NLO inclusive dijets that pass these cuts are now ready for analysis. We use the observables used in the measurement by CMS of jet-gap-jet events at 13 TeV. In this study, special attention was given to the following observables:

- The difference between the pseudorapdity of the two jets, $\Delta\eta_{jj} = |\eta_1 - \eta_2|$. This difference is highly sensitive to predictions made with perturbative BFKL calculations [26].

- Analysis of the momentum of the subleading jet, p_{T2} will address phenomenology studies that predict a weak dependence on the ratio of p_{T1}/p_{T2} for jet-gap-jet events at NLO accuracy divided by the number of inclusive dijet events [46].

- The azimuthal angle separation between the two jets, $\Delta\phi_{jj} = |\phi_1 - \phi_2|$, is sensitive to deviations from back-to-back (π) topology for jet-gap-jet. This is caused by higher order perturbative corrections [5].

- The number of charged particles with transverse momentum $p_T > 200$ MeV in the psudorapidity interval used for jet-gap-jet, $|\eta| < 1.0$ [5].

For each observable we want to look at the differential cross section as a function of each chosen observable. Here we define the differential cross section as a function of a given observable value, x_i, as

$$\frac{d\sigma}{dx_i} = \frac{\sigma L}{N_E} N_{x_i}. \tag{4.5}$$

Where σ is the weighted cross section of NLO inclusive dijet production in units of picobarns computed by PYTHIA8, N_E is the total number of events we produced using POWHEG, N_{xi} is the number of jets at a given value i of observable x, and L is the integrated luminosity of 0.66 pb^{-1}. This value is chosen based on data collected by the CMS experiments in proton-proton collisions during a low luminosity run in 2015 at $\sqrt{s} = 13$ TeV. The differential cross section essentially tells us the probability of an event occurring as a function of the the value of our observable of choice.

4.4 Uncertainties

There are two sources of statistical uncertainty in this analysis:

- The statistical uncertainty that comes from treating the events as a Poisson distribution when binning for histograms. More specifically, this is the uncertainty on the mean of the underlying Poisson distribution used when binning the generated data, given an observed number of generated events [47]. In order to lower this value we would need to produce more events.

- The statistical uncertainty given by POWHEG+PYTHIA8 from generateing NLO inclusive dijet events. This is provided by PYTHIA8 as the uncertainty of the weighted cross section.

The statistical uncertainty is determined and displayed by our histogram plotting package provided by CERN's object-oriented program and library, ROOT. The uncertainty in the weighted cross section is given as an output by PYTHIA8. While the uncertainty in the weighted cross section was found to be insignificant when compared to the statistical uncertainty from the Poisson distribution when binning the histograms, we will present its value along with the cross section the next chapter. Lastly, a full analysis would require an investigation into the uncertainty coming from the theory used to generate the dijet events. This would be done by varying the values of the parameters used generate the events to determine the significance of the effects and would need to be carried out in the future by colleagues at University of Kansas or Muenster.

Chapter 5

Results

5.1 Results

Making use of POWHEG+PYTHIA8 we were able to generate 4,276,893 NLO inclusive dijet events from proton-proton collisions at center of mass energy $\sqrt{s} = 13$ TeV. After reconstructing the jets with FastJet we applied the jet-gap-jet cuts described in Section 3 of Chapter 4 to the produced jets. Specifically, a minimum transverse momentum $p_{Tj} > 40$ Gev for the jets, inclusive dijets pseudorapidities that give $\eta_1 \times \eta_2 < 0$, and $1.4 < |\eta_{1,2}| < 4.7$ was used to ensure consistency. These cuts reduced our sample size to 3,752 dijet events that passed the cuts and will be used for analysis. The cross section for these events was given by PYTHIA8 as $\sigma = (5.894 \times 10^6 \pm 1.4028 \times 10^4)$ pb and will be used in equation 4.4 to calculate the differential cross sections for the observables we wish to analyze.

The first observables we analyzed were the transverse momentum of the leading and subleading jets. These distributions provide a nice check on the quality of the generated events, and a reference back to Figure 4.1 shows how we used these distributions to determine whether or not the value of p_{Tmin} for our generated jets was too large. In Figure 5.1 we can see our value of p_{T2} peaking at 40 GeV. As we would expect, the peak value for the subleading jet's transverse momentum is close to the value chosen for the minimum transverse momentum for the analyzed jets. From Figure 5.1 we can also see that the distribution for the transverse momentum of the leading jet peaks slightly higher than the 40 GeV of the subleading jet. This behavior was anticipated since, by definition, the leading jet must be more energetic then the subleading jet. Therefore, based on the analysis of

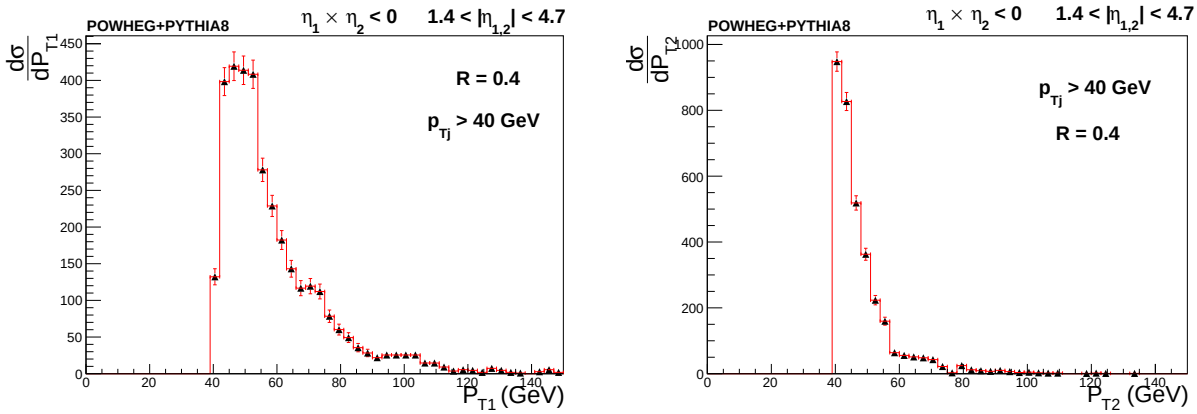

Figure 5.1: (Left) The differential cross section as a function of the p_T of the leading jet. (Right) The differential cross section as a function of the p_T of the subleading jet. In both plots the black triangles represent the data points, the red vertical bars represent the statistical uncertainty from binning the data, and the red horizontal bar represent the bin width.

these distributions we can conclude that our generated inclusive dijets are behaving as one would expect. The main reason for looking at the differential cross section for the transverse momentum of the leading and subleading jet lies in the fact that these are observables used in the measurement by CMS of jet-gap-jet events at 13 TeV, and will be used by our colleagues in Muenster producing NLO jet-gap-jet events. Specifically, p_{T2} was said to address phenomenology studies that predict a weak dependence on the ratio of p_{T1}/p_{T2} for jet-gap-jet events at NLO accuracy divided by the number of inclusive dijet event.

The next observables we looked at were the number of jets produced in the dijet analysis, and the number of charged particles that were produced in the "gap" region, $|\eta| < 1.0$, for jet-gap-jet events. The number of jets produced (also known as jet multiplicity), N_{jets}, is an important observable to look at, especially when considering different jet matching algorithms and NLO accuracies. This is because parton showering, rehadronizatoin, and jet reconstruction are affected by the logarithmic accuracy and jet matching algorithm chosen. We expect that most of the time NLO dijet production will produce two jets, but occasionally we will see the production of more than two, with higher numbers of jets being less likely. This is precisely what we observed in our generated data and can be seen in the left plot of Figure 5.2.

The number of charged particles, N_{CP} produced in the "gap" is an important value to consider if

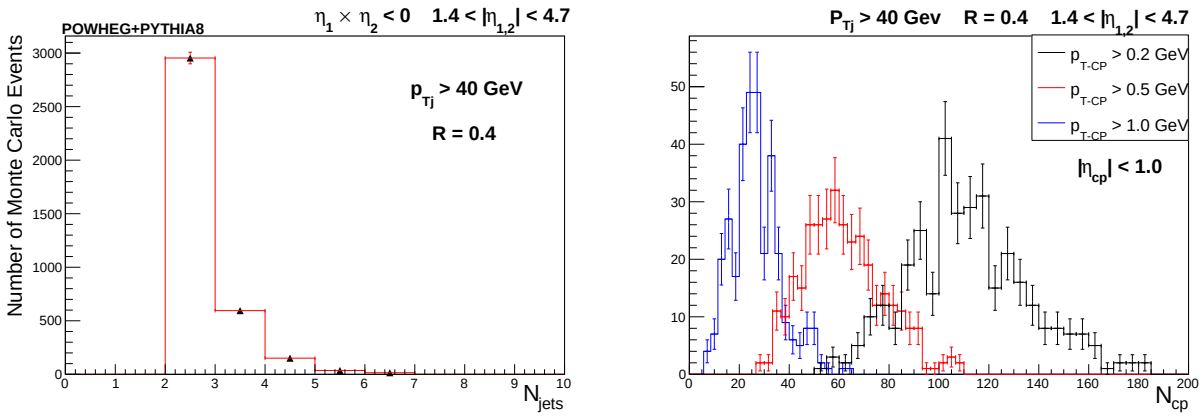

Figure 5.2: (Left) The number of jets produced in each event. The black triangles represent the data points, the red vertical bars represent the statistical uncertainty from binning the data, and the red horizontal bar represent the bin width (Right) The number of charged particles produced in the "gap" region of each event with varying minimum p_{T-CP}. The black line is for a minimum p_{T-CP} of 0.2 GeV, the red line is for a minimum p_{T-CP} of 0.5 GeV, and the blue line is for a minimum p_{T-CP} of 1.0 Gev. In each case the vertical bars represent the statistical uncertainty from binning the data, and the horizontal bar represent the bin width.

our simulated data is to be compared to NLO jet-gap-jet events. Originally we placed a cut on the allowed minimum transverse momentum of the charged particles of $p_{T-CP} > 200$ MeV = 0.2 GeV. This value ensures that the charged particles have high enough energies to be considered "real" events inside the detector, while being low enough to consider and count all the charged particles produced. However, we found our distribution was peaking around 110 charged particles per event, way more charged particles produced in each event than one would expect [48]. To investigate the cause of this we decided to raise the value chose for p_{T-CP} to see if the cause of the over production of charged particles was due to an over production of low-energy charged particles. Looking at Figure 5.2 we can see that slightly increasing p_{T-CP} had a significant impact on the number of charged particles produced in each event. By the time we had reached $p_{T-CP} > 1.0$ GeV the distribution for number of charged particles was peaking at a more reasonable vale of 20 charged particles per event. This leads us to conclude that POWHEG+PYTHIA8 overproduced the number low-energy charged particles for NLO inclusive dijet production from proton-proton collisions at $\sqrt{s} = 13$ TeV. Future investigations may want to check lower collision energies or LO accuracies to see if these have an effect. If not, the root cause of this issue may lie elsewhere.

Now we begin looking at the more pertinent observables used in the measurement by CMS of jet-gap-jet events at 13 TeV in Ref. [5], used for searching for effects that have been predicted by the BFKL evolution equation . These observables were given in Chapter 4 Section 3 as the difference in pseudorapidity of the two leading jets, $\Delta\eta_{jj} = |\eta_1 - \eta_2|$, and the azimuthal angle separation between the two leading jets, $\Delta\phi_{jj} = |\phi_1 - \phi_2|$. Let us begin by taking a look at $\Delta\eta_{jj}$.

The left plot in Figure 5.3 shows the differential cross section as a function of $\Delta\eta_{jj}$ for the two leading jets. The absolute value of the difference between η_1 and η_2 ensures our distribution is always positive, and the cuts placed on η from jet-gap-jet are responsible for lack of observed events in the region $0.0 < |\Delta\eta_{jj}| < 3.0$. While the $\Delta\eta_{jj}$ looks okay with no noticeably strange features, we decided to investigate further. For the right plot in Figure 5.3 we looked at $\Delta\eta_{jj}$ for small $\Delta\phi_{jj}$, i.e. $\Delta\phi_{jj} < 0.4$. While this investigation goes beyond the scope of our study, the reason we did this was two fold. One, we wanted to explore the properties of collinear events to see if small $\Delta\eta_{jj}$ correlated to jets with small $\Delta\phi_{jj}$. These are interesting kinematics to explore for Mueller-Navalet jets [49]. Also, we wanted to make sure nothing strange was happening at these very small angular separations where the chosen jet merging scheme can have significant effects. By comparing the left and right plots from Figure 5.3 we can see that $\Delta\phi_{jj} < 0.4$ cut caused the $\Delta\eta_{jj}$ distribution to slightly skew towards larger $\Delta\eta_{jj}$ values, but due to the low number of statistics after the $\Delta\phi_{jj}$ cut leading to larger uncertainties we can not be sure this skew is entirely physical. The main take away is that the $\Delta\phi_{jj}$ cut did not cause our simulated data to behave in "funky" ways, and served as a nice quality check on our generated events.

Now let us take a look at the distribution that contained the most striking and interesting feature, the differential cross section as a function of $\Delta\phi_{jj}$ for the two leading jets. With LO accuracies one expects a more back-to-back configuration for the leading and subleading jets rustling in a $\Delta\phi_{jj}$ distribution peaking at π. A good example of this type of distribution was presented in Figure 1 of the Preface. However, looking at Figure 5.4 we can see that our distribution for $\Delta\phi_{jj}$ deviates slightly from the back-to-back configuration and peaks slightly before π instead of at π. We believe this feature is not a mistake of the Monte Carlos used to produce the data. Instead we

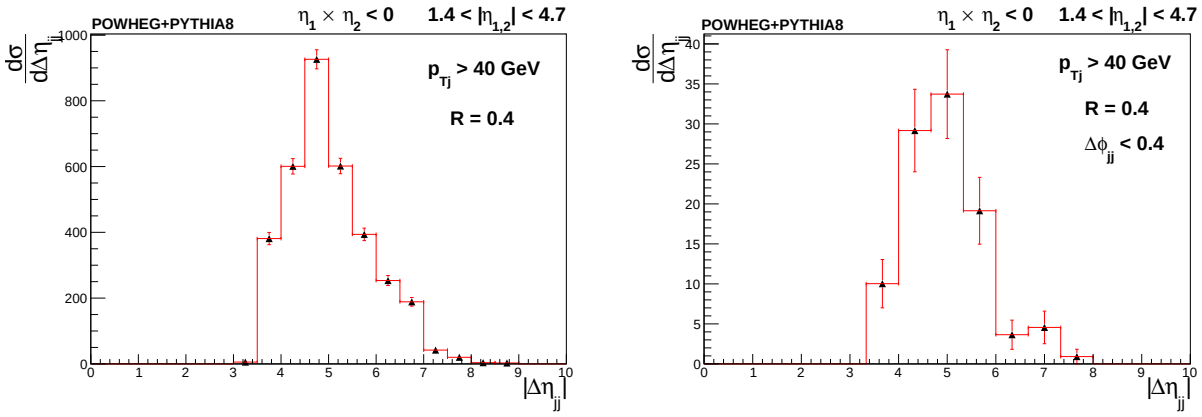

Figure 5.3: (Left) The differential cross section as a function of the $\Delta\eta_{jj}$ of the two leading jet. (Right) The differential cross section as a function of the $\Delta\eta_{jj}$ for small $\Delta\phi_{jj}$ of the two leading jets. In both plots the black triangles represent the data points, the red vertical bars represent the statistical uncertainty from binning the data, and the red horizontal bar represent the bin width.

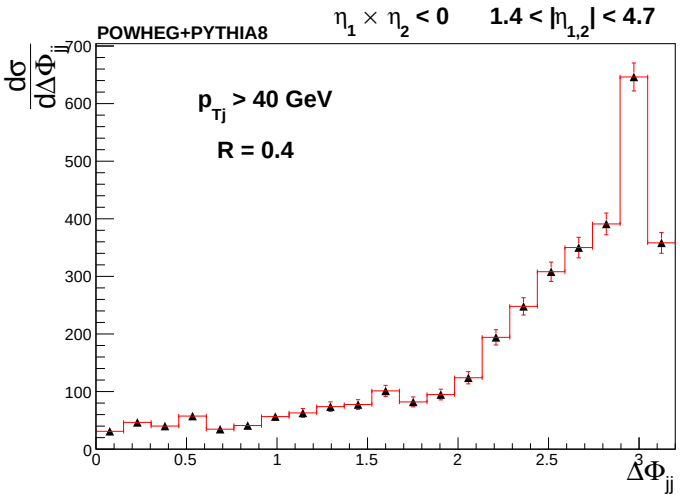

Figure 5.4: The differential cross section as a function of the $\Delta\phi_{jj}$ of the two leading jet. The black triangles represent the data points, the red vertical bars represent the statistical uncertainty from binning the data, and the red horizontal bar represent the bin width.

believe this feature is due to difficulty in creating back-to-back configurations when considering NLO corrections. This is because NLO can generate two to three partons at the parton level. Contrast this with LO accuracies which only consider two partons at the parton level. If one has three partons then it is clearly not possible to have a back-to-back configuration for our dijets that result form the partons. Therefore, we believe that this interesting feature is physical, and is a product of considering NLO accuracies.

5.2 Summary

Inclusive NLO dijet events resulting from proton-proton collisions at center of mass energy $\sqrt{s} =$ 13 TeV were generated with POWHEG and the parton showering and hadronization to produce final state particles was performed with PYTHIA8. The stable particles were then reconstructed into jets using FastJet's anti-k_t jet clustering algorithm for analysis. We applied jet-gap-jet cuts to the dijet events in accordance with previous CMS jet-gap-jet studies, and to maintain consistency with NLO jet-gap-jet events being generated by our colleagues at the University of Muenster through the implementation of our colleague Federico Deganutti's NLO BFKL calculations.

The observables we analyzed for our generated events were based on observables analyzed in the measurement of jet-gap-jet events at 13 TeV by CMS, most of which were chosen based on their believed sensitivity to NLO BFKL calculations. Specifically, we took a close look at $\Delta\eta$ between the two leading jets which is believed to be sensitive to perterbative BFKL calculations, and applied a cut to look at the effects on $\Delta\eta$ at small $\Delta\phi$. We looked at $\Delta\phi$ between the two leading jets which is sensitive to deviations from back-to-back topology caused by NLO corrections. Also analyzed was the number of charge particles in the "gap" region of jet-gap-jet, and the number of jets produced in each event. The number of charged particles was plotted as a function of different minimum transverse momentum values of the charged particles in question to investigate the overproduction of low-energy charged particles.

The results of this analysis found several interesting features. While $\Delta\eta$ and p_{T2} showed no strik-

ing features, $\Delta\phi$ was found to indeed deviate from the back-to-back topology found in events produced at the LO. We believe this is caused by NLO corrections that allow for the generation of three partons at the parton level instead of the maximum two partons found in LO calculations. We also found that our MCs, POWHEG+PYTHIA8, were producing an overabundance of low-energy charged particles. It was not until we reached a minimum transverse momentum of the charged particles of 1.0 GeV that the distribution of the number of charged particles began to behave as expected.

5.3 Future Prospects

While I plan to continue to stay in contact with colleagues at the University of Kansas and the University of Muenster to transfer both data and knowledge, the analysis for my thesis has concluded. Before ending I would like to briefly discuss some of the future plans for this project, and where I think it can go from here. I would also like to take a few sentences to mention my future plans after graduate school.

For the project I would like to start by discussing some areas that I would have investigated further if I had more time to work on this analysis. The first is concerning the number of charged particles. It is not obvious what is causing this issue, but I have several troubleshooting suggestions that might lead to an answer. One suggestion is to test a lower center of mass energy for the collision in the Monte Carlos. It might be the case that simulating events at $\sqrt{s}$ = 13 TeV leads to overproducing low-energy particles, including charged particles. Running a simulation at $\sqrt{s}$ = 7 Tev and looking at the number of charged particles in the "gap" region would answer this question. One could also try generating events with LO corrections as apposed to NLO to see if these higher-order corrections are the cause of the charged particle distribution. Lastly, it has been suggested by some of our Muenster colleagues that we considered the "weights" associated with each event produced. These are values assigned to each event and given as output by PYTHIA8. Typically these weights are the same for every event produced and, therefore, trivial. However, there are some instances

where the weighting is non-trivial and should be taken into account. While this event weight was brought to my attention to late to include it into my analysis, it is something future analysis should consider, and may get rid of the surplus of charged particles that is currently observed.

For a proper error analysis I would have liked to vary the theoretical parameters used by both POWHEG and PYTHIA8 to generate the events. By monitoring the effects changing the parameters has on the distributions of our observables, one can obtain an uncertainty on our generated events that comes from the theory used to generate the events. Currently, our analysis only considers the statistical uncertainty from the Poisson distributions during histogram binning, and the statistical uncertainty given by POWHEG+PYTHIA8 as the uncertainty in the weighted cross section.

My hopes for the future of this project is colleagues will pick up where I left on in generating a large sample of NLO inclusive dijet events for comparison with NLO jet-gap-jet events that are incorporating Federico Deganutti's NLO BFKL calculations. Currently, I have produced a little over 4 million inclusive dijet events, but a number closer to 20 million would be ideal. This would reduce the statistical uncertainty and could reveal more subtle features in the distributions. The future of this analysis should either solve or understand the source of the surplus of low-energy charged particles, and take into account the event weights associated with each event produced. It might also be nice if Madgraph5 was revisited at some point as it should still be possible to produce NLO dijet events using Madgraph5. I believe a comparison of events produced by Madgraph5 and events produced by POWHEG would be interesting for both NLO inclusive dijet studies and for studies of the difference between these two Monte Carlo event generators. Either way, a continuation of the work done in this analysis will lead to a contribution in the calculation of the ratio of jet-gap-jet events divided by the number of inclusive dijet events to search for effects that have been predicted by the BFKL evolution equation. More broadly, a continuation of this work could lead to a better understanding of QCD and the strong interaction through the study of dijets.

Lastly, I would like to mention my future plans. Now that my career as a graduate student is drawing to a close I would like to take the skills and knowledge I have obtained and apply it to a career in data science. My time as graduate student performing research in nuclear physics gave

me the opportunity to find and explore another area of interest that is closely related to my research in physics, big data analysis. It allowed me to develop an ability to analyze large sets of data and think critically about the results of the analysis. I have no doubt the experience I gained from my research will allow me to apply the new skills that I learn in industry in unique ways to find clever solutions, and I truly believe that my time spent in graduate school has given me a valuable perspective on data analysis and prepared me for a successful career.

Bibliography

[1] ABC News. How does the 'bible' of quantum physics work?. https://www.abc.net.au/news/science/2017-07-15/the-standard-model-of-particle-physics-explained/7670338, accessed: 11.22.2020.

[2] The Editors of Encyclopedia Britannica. Fundamentla force physics. https://www.britannica.com/science/fundamental-interaction, accessed: 11.22.2020.

[3] CERN Courier. Qcd scattering: from dglap to bfkl. https://cerncourier.com/a/qcd-scattering-from-dglap-to-bfkl/, accessed: 11.22.2020.

[4] CMS Collaboration. Search for Dijet Resonances in 7 TeV pp Collisions at CMS. *Phys.Rev.Lett.*, 105, 2010.

[5] CMS Collaboration. Study of hard color singlet exchange in dijet events with proton-proton collisions at $\sqrt{s}$ =13 TeV. *CMS-PAS-SMP-19-006*, 2020.

[6] Tech2. Cern's new concept for a next-gen particle smasher dwarfs the large hadron- technology news, firstpost. https://www.firstpost.com/tech/science/cerns-new-concept-for-a-next-gen-particle-smasher-dwarfs-the-large-hadron-5904071.html, accessed: 11.22.2020.

[7] Cms particle detector. https://www.mpoweruk.com/cms-detector.htm, accessed: 11.22.2020.

[8] A. Mueller and W.-K. Tang. High energy parton-parton elastic scattering in QCD. *Physics Letters B*, 284(1-2):123, 1992.

[9] Atlas Collaboration. High energy parton-parton elastic scattering in QCD. *Eur.Phys.J.*, C71:1512, 2011.

[10] S. Alioli, P. Nason, C. Oleari, and E. Re. The POWHEG BOX user manual: common features. 2011.

[11] R.D. et al. Ball. Parton distributions with small-x resummation: evidence for BFKL dynamics in HERA data. *Eur. Phys. J.*, 78, 2018.

[12] Ian J.R. Aitchison and Anthony J.G. Hey. *Gauge Theroies in Particle Physics: A practical, Introduction, Volume II*.

[13] A. Zee. *Quantum Field Theory in a Nutshell*. Princeton University Press, 2 edition, February 2010.

[14] Particle Data Group Collaboration. Review of Particle Physics. *Phys. Rev. D*, 98, 2018.

[15] Wilczek F. QCD Made Simple. *American Institute of Phys.*, Aug 2000.

[16] L. N. Lipatov. The bare pomeron in quantum chromodynamics. *Sov. Phys. JETP*, 63(904), 1986.

[17] Y. L. Dokshitzer. Calculation of the structure functions for deep inelastic scattering and e + e annihilation by perturbation theory in quantum chromodynamics. *Sov. Phys. JETP*, 46(641), 1977.

[18] J. L. Albacete and C. Marquet. Gluon saturation and initial conditions for relativistic heavy ion collisions. *Prog. Part. Nucl. Phys.*, 76, 2014.

[19] G. Aad and ATLAS Collaboration. Dijet production in $\sqrt{s}$ = 7 TeV pp collisions with large rapidity gaps at the ATLAS experiment. *Phys. Lett. B*, 754, 2016.

[20] ATLAS Collaboration. Measurement of dijet production with a veto on additional central jet activity in pp collisions with $\sqrt{s}$ = 7 TeV using the ATLAS detector. *JHEP*, 09(053), 2011.

[21] S. Chatrchyan and CMS Collaboration. Ratios of dijet production cross sections as a function of the absolute difference in rapidity between jets in proton-proton collisions at $\sqrt{s}$ =7 TeV. *Eur. Phys. J. C*, 72(2216), 2012.

[22] ZEUS Collaboration. Rapidity gaps between jets in photoproduction at HERA. *Phys. Lett. B*, 369(55), 1969.

[23] D0 Collaboration. The azimuthal decorrelation of jets widely separated in rapidity. *Phys. Rev. Lett.*, 77(595), 1969.

[24] D0 Collaboration. Probing BFKL dynamics in the dijet cross section at large rapidity intervals in p p collisions at $\sqrt{s}$ = 1800 GeV and 630-GeV. *Phys. Rev. Lett.*, 84, 1999.

[25] Steffen A. Bass. Review of parton recombination models. *J. Phys.: Conf. Ser.*, 50(033), 2006.

[26] S. Marzani, M. Spannowsky, and Soyez G. Looking inside jets: an introduction to jet substructure and boosted-object phenomenology. *Lect. Notes in Phys.*, 958, 2019.

[27] I. I. Balitsky and L. N. Lipatov. The Pomeranchuk Singularity In Quantum Chromodynamics. *Sov. J. Nucl. Phys.*, 28(822), 1978.

[28] J. D. Bjorken. Rapidity gaps and jets as a new-physics signature in very-high-energ hadron-hadron collisions. *Phys. Rev. D*, 47(101), 1993.

[29] S. Donnachie, G. Dosch, P. Landshoff, and O. Nachtmann. Pomeron physics and QCD. *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.*, 19, 2002.

[30] O. S. et al Bröning. LHC design report. *CERN Yellow Reports: Monographs*, 2004.

[31] CMS Collaboration. Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC. *Phys. Lett. B*, 716, 2012.

[32] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716, 2012.

[33] Cern accelerating science. https://home.cern/science/accelerators/high-luminosity-lhc, accessed: 11.22.2020.

[34] CMS Collaboration. The CMS Experiment at the CERN LHC. *JINST*, 3(S08004), 2008.

[35] CMS Collaboration. The CMS Electromagnetic Calorimeter at the LHC. *PoS TIPP 2014*, 029, 2014.

[36] CMS Collaboration. Performance of the CMS Hadron Calorimeter with Cosmic Ray Muons and LHC Beam data. *JINST*, 5(TO3012), 2010.

[37] A. et al. Buckley. Monte Carlo Event generators for high energy particle physics event simulations. *MCnet-19-02*, 2019.

[38] S. et al. Ask. An introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 2014.

[39] J. et al. Alwall. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 1407(079), 2014.

[40] M. Cacciari, G. P. Salam, and G. Soyez. FastJet user manual (for version 3.3.4). *CERN-PH-TH*, 297, 2011.

[41] D. Bourilkov, R.C. Group, and M.R. Whalley. The Les Houches accord PDFs (LHAPDF) and LHAGLUE. *HERA and the LHC: A Workshop on the implications of HERA for LHC phys.: Part B*, 2005.

[42] POWHEG-Box. The origin of spikes in dijet production. http://powhegbox.mib.infn.it/Spike-in-Dijet.html. accessed: 11.22.2020.

[43] CMS Collaboration. Extraction and validation of a new set of CMS PYTHIA 8 tunes from underlying-event measurements. *CMS-PAS-GEN-17-001*, 2018.

[44] A. Buckley and D. B. Gupta. Powheg–Pythia matching scheme effects in NLO simulation of dijet events. *MCNET-16-34*, 2016.

[45] M. Cacciari, G. Salam, and G. Soyez. The anti-kt jet clustering algorithm. *JHEP*, 04(063), 2008.

[46] R. Enberg, G. Ingelman, and L. Motyka. Hard color singlet exchange and gaps between jets at the Tevatron. *Phys. Lett. B*, 524(273), 2002.

[47] R. Aggarwal and A. Caldwell. Error bars for distributions of numbers of events. *Eur. Phys. J. Plus*, 127(24), 2012.

[48] CMS Collaboration. Study of dijet events with a large rapidity gap between the two leading jets in pp collisions at $\sqrt{s} = 7$ TeV. *Eur. Phys. J. C*, 78(3), 2018.

[49] F. Deganutti, D. Colferai, and A. Niccoli. Improved theoretical description of Mueller-Navelet jets at LHC. *PoS QCDEV 2016*, 2017.