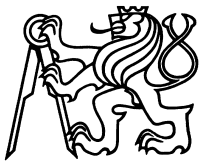
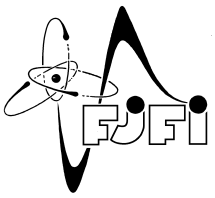


CZECH TECHNICAL UNIVERSITY IN PRAGUE
Faculty of Nuclear Sciences and Physical Engineering



Measurement of di-jet events in the ATLAS experiment

Měření dvou-jetových událostí v experimentu ATLAS

Bachelor's Thesis

Author: **Vladimír Žitka**

Supervisor: **Ing. Miroslav Myška, Ph.D.**

Academic year: 2016/2017

- Zadání práce -

- Zadání práce (zadní strana) -

Prohlášení:

Prohlašuji, že jsem svou bakalářskou práci vypracoval samostatně a použil jsem pouze podklady uvedené v přiloženém seznamu.

Nemám závažný důvod proti použití tohoto školního díla ve smyslu § 60 Zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon).

Prague, July 10, 2017 Vladimír Žitka

Acknowledgment:
I would like to thank Miroslav Myška for his expert guidance, patience and his equanimity in face of my questions. Furthermore, I would like to express my gratitude to my family for their loving support and the opportunity to freely pursue my education.

Název práce:

Měření dvou-jetových událostí v experimentu ATLAS

Autor: Vladimír Žitka

Obor: Experimentální jaderná a částicová fyzika

Druh práce: Bakalářská práce

Vedoucí práce: Ing. Miroslav Myška Ph.D., FJFI ČVUT v Praze

Abstrakt: Hlavním tématem této práce je analýza dat z proton-protonových srážek měřených detektorem ATLAS na urychlovači LHC v CERN. Cílem této analýzy je testování a ověřování platnosti Standardního modelu částicové fyziky pomocí porovnávání skutečných naměřených dat s daty získanými pomocí MC generátoru. Pro lepší pochopení fungování analýzy jako takové je připraven lehký úvod do problematiky sprškových (jetových) algoritmů. Předposlední kapitola se zabývá tématem substruktury těchto částicových spršek a vlivu různých nastavení generátoru Herwig na tvar distribucí popisující tuto substrukturu. V poslední kapitole je pak uveden příklad analýzy dat, která vyhledává události kde byly vyprodukovány dvě částicové spršky a vyhodnocuje jejich topologii. Výsledky této analýzy dat jsou pak porovnány s výsledky vyhodnocování oficiálních simulací poskytnutými kolaborací ATLAS, které byly generovány programy Pohweg + Pythia.

Klíčová slova: ATLAS, dvou jetové události, jetová substruktura

Title:

Measurement of di-jet events in the ATLAS experiment

Author: Vladimír Žitka

Abstract: The main theme of this thesis is the analysis of proton proton collision measured by the ATLAS detector at the LHC in CERN. The main goal of this analysis is to test the prediction of the Standard model of particle physics by comparing the predictions of MC generator with the actual measured data from ATLAS. For a better understanding of the inner working of this analysis, a succinct overview of jet algorithms is included. The penultimate chapter deals with jets substructure and the influence of settings of Herwig on the shape of distributions describing said jets substructure. And finally the last chapter includes an example of data analysis, that seeks out events, where only two jets (dijets) are produced and then studies the topology of such events. The results are then compared to those acquired from the official Powheg + Pythia simulations provided by the ATLAS collaboration.

Key words: ATLAS, dijets, jet substructure

Contents

1 Standard Model Of Particle Physics **15**
- 1.1 Elementary particles . . . 15
- 1.2 Fundamental interactions . . . 16
  - 1.2.1 Weak interaction . . . 16
  - 1.2.2 Electromagnetic interaction . . . 17
  - 1.2.3 Strong interaction . . . 17
  - 1.2.4 Standard model . . . 17

2 Hard Scattering Processes **19**
- 2.1 Feynman diagrams . . . 19
- 2.2 Hard scattering formalism . . . 19

3 LHC And The ATLAS Experiment **23**
- 3.1 ATLAS . . . 24
  - 3.1.1 Overview of the detector . . . 24
  - 3.1.2 Tracking . . . 25
  - 3.1.3 Calorimetry . . . 27
  - 3.1.4 Muon system . . . 29
  - 3.1.5 Forward detectors . . . 31
  - 3.1.6 Trigger, readout, data acquisition and control system . . . 31

4 Jet Algorithms **33**
- 4.1 Cone algorithms . . . 33
- 4.2 Sequential recombination algorithms . . . 34
  - 4.2.1 Generalised k_t algorithm . . . 35
- 4.3 Example of anti-k_t vs k_t results . . . 35

5 Jet Substructure **39**
- 5.1 Review of the reference analysis . . . 39
  - 5.1.1 Data description . . . 40
- 5.2 Reconstruction of chosen properties . . . 40
  - 5.2.1 Cross section as a function of p_T . . . 41
  - 5.2.2 Charged particle multiplicity per jet . . . 41
  - 5.2.3 Number density of charged particle tracks . . . 43

6 Data Analysis **45**
6.1 Event selection and weight 45
6.2 Systematic uncertainties 47
6.3 Measured quantities 48
6.3.1 Inclusive jets p_T spectrum and multiplicity 48
6.3.2 Properties of dijets 50

Conclusion **57**

Appendix A Data Analysis pictures **59**
A.1 Boost of the dijet system and dijet rapidity separation. 59
A.2 Systematic uncertainties of measured quantities 60

Introduction

The main theme of this thesis is the analysis of data from the detector ATLAS at the Large Hadron Collider (LHC). Such data contain information about collisions of two protons which can result in many different things. The main focus is on the events that result in production of jets. What is a jet and how is it defined in particle physics is explained in latter chapters. There are only a few basic from particle and accelerator physics presented in this short introduction, with other endeavoured to explain as they are necessary.

In physics and science in general there is an idea that from observation of a process for a sufficient amount of time the general laws governing the process can be derived. This has been the case i.e. for Johannes Kepler, who was through analysis of data from observed movements of celestial bodies able to formulate his famous laws. In particle physics the aim is to find and study the properties of particles that comprise or otherwise influence matter.

The idea of deriving laws from observed data is utilized in the scattering experiments, where a particle (also called probe) of certain energy is sent to collide with a target and through this collision shed some light on the structure of the target. There is a simple relation between the energy of the probe and the resolution it provides that is derived from the Heisenberg uncertainty relations:

$$\Delta p \Delta x \leq \frac{\hbar}{2} \tag{1}$$

If the target in question is the size of x then the momentum of the probe must be greater than $\frac{\hbar}{2x}$ in order to reveal the structure. A fine example was the discovery of the nucleus of atom by Rutherford, Geiger and Marsden in 1911 during one of the first experiments that set out to probe the structure of atom. Their idea was to send α-particles onto a thin sample of gold. At the time, the popular model of atom was Thompsons' "pudding" model: the atom has a positively charged medium throughout which negatively charged electrons are spattered in such a way that the resulting charge is zero. If this model were correct then the α particles sent in Rutherford's experiment that would collide with the gold atom should shoot through nearly without a deviation of their path caused by coulomb interaction with electrons. Their experimental setup was quite simple: they surrounded the golden sample with scintillating material with only a narrow gap for the alpha particles to come in and recorded the number of scintillations in certain segments of space angle. After a few days of thorough observations, they discovered that some alpha particles are deflected in large space angles which was in direct contradiction with the pudding model and paved the way for modern versions of atomic models with compact nucleus at the centre surrounded by electrons.

The cornerstone of particle physics is the measurement of various cross-sections. That is a concept constructed from the idea taken from a simple thought experiment: If two spheres are launched at each other the probability that they scatter off each other is proportional to the areas $S_i = \pi r_i^2$ where r_i is the radius of the spheres or to be more precise to the part, where these two areas overlap in the plane perpendicular to the axis that is defined by their movement. This area is then called cross-section. This concept of describing the probability of a process through area is now adapted in particle physics when

describing the scattering of a particle on a potential or a force centre. In order to better understand some concepts surrounding cross sections, Fig. 1 shows the general layout of scattering experiments. Where b denotes the so called impact parameter - the transverse distance of the initial asymptotic trajectory of the particle from the force centre , ϑ denotes the scattering angle - the angle between the initial and final asymptotic trajectories of the particle and J is the particle flux - the number of particles per a unit surface, the concept of particle flux is no longer widely used and is replaced by luminosity L ,which is the number of recorded events per unit of time divided by the cross section of such an event:

$$L = \frac{1}{\sigma}\frac{dN}{dt}. \tag{2}$$

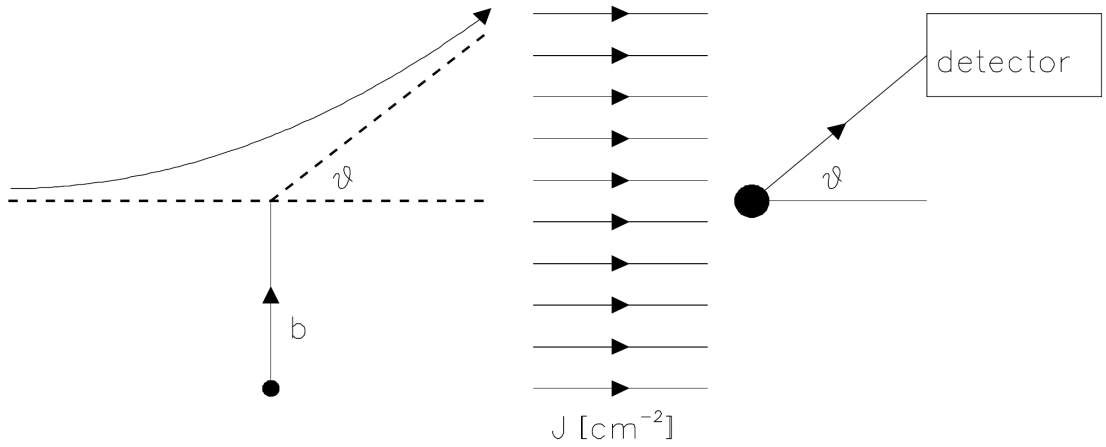


Fig. 1. General layout of a scattering experiment.Taken from [1].

It is customary to express cross sections of most scattering experiments (such as Rutherford's) in their differential forms with respect to the space angle. The general form of the differential cross section is usually dependent on the potential, which defines the full set of observables that can be measured in the process in question. These observables can be for example the energy, (four) momentum, impact parameter, scattering angle or the transferred momentum q :

$$\vec{q} \equiv \vec{p'} - \vec{p} \quad \Rightarrow \quad q^2 = 4p^2 \sin^2 \frac{\vartheta}{2}. \tag{3}$$

This definition of transferred momentum is however not universal and the implication is only valid for the Rutherford scattering cross section:

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2 m^2}{4p^4 \sin^4(\vartheta/2)} = \frac{4\alpha^2 m^2}{q^4}. \tag{4}$$

For example in the hard scattering experiments (experiments that use highly energetic particles as probes of structure of the subatomic particles) the momentum transfer is defined as the sum of four-momenta of incoming particles and is especially useful in the calculation of cross sections of processes involving the deep inelastic scattering (DIS). The question of what is the meaning of DIS is one exceeding the theoretical knowledge this thesis is intending to encompass, an introduction into the topic can be found for example in the article [2].

This thesis is comprised of a basic look at the Standard Model of the particle physics, followed by a foray in to concepts from the hard scattering theory that could prove useful during the struggle to understand the framework which is used in chapters about event analysis. After this follows a little detour into the design of the ATLAS (A Toroidal LHC ApparatuS) experiment. Thereafter, the first chapter about software comes in which the basic use of a MC generator Herwig and the overview of some general jet algorithms is put forth. The content of the next to last chapter is jet substructure, where two of the basic processes generated by Herwig are introduced and their ability to reconstruct basic quantities is tested. And in the last chapter there is a dijet analysis of the data from proton-proton collisions at the centre of mass energy of $\sqrt{s} = 8$ TeV measured by the ATLAS experiment at the Large Hadron Collider (LHC).

Chapter 1

Standard Model Of Particle Physics

Now follows a summary of particles and interactions along with their relation to the Standard Model(SM) of particle physics. As this is not a theoretical work and these frameworks and theories are rather advanced, the rigorous derivation of all concepts and formulas is omitted and the focus is instead on more practical explanations.

One of the most fundamental questions of physics is: what are the elementary particles in the universe and what are the interactions. The known interactions are: Electromagnetic, Weak, Strong and Gravitational. They are discussed below and now an overview of elementary particles is in order.

1.1 Elementary particles

Every elementary particle is a manifestation of its quantum field and can be viewed as its material fluctuation. Every particle can exist only within the boundaries of Heisenberg uncertainty principle. This means that even if the energy available in system is not sufficient to reach the mass of the particle, the particle can be said to exist withing a sufficiently small time window. Each particle is characterized by its mass, lifetime and a set of various quantum numbers for example intrinsic angular momentum (spin), electric charge and colour charge (charge of the strong interaction), parity, flavour and lepton number. The common classification is into quarks, leptons and intermediate bosons. A brief overview of their properties can be found in Fig. 1.1. It should be mentioned, that bosons are generally particles with integral spin and so the particles in Fig. 1.1 are not the only bosons in existence. Quarks and leptons are included in the category called fermions, those are particles with half-integral spin that abide the Pauli exclusion principle.

Quarks: As can be seen in Fig. 1.1 there are three generations of quarks with each generation more massive than the other. In each generation there are two quarks with different mass and charge. Each quark has its flavour, three colours and the charge of either $-\frac{1}{3}$ or $+\frac{2}{3}$ of elementary charge (e). This means that quarks are susceptible to all four interactions. Another interesting property of quarks is that they can never be found isolated and are ever in a bound into composite particles called hadrons. The reason for this property is mentioned below when the strong interaction is discussed. Hadrons can be divided into mesons (bound states of a quark and an anti-quark, another example of bosons) and baryons (fermions that have three valence quarks).

Leptons: Those are the particles that do not participate in the strong interaction. They have the electric charge of either -1 e for electron, muon and tauon or 0 for their neutrinos.

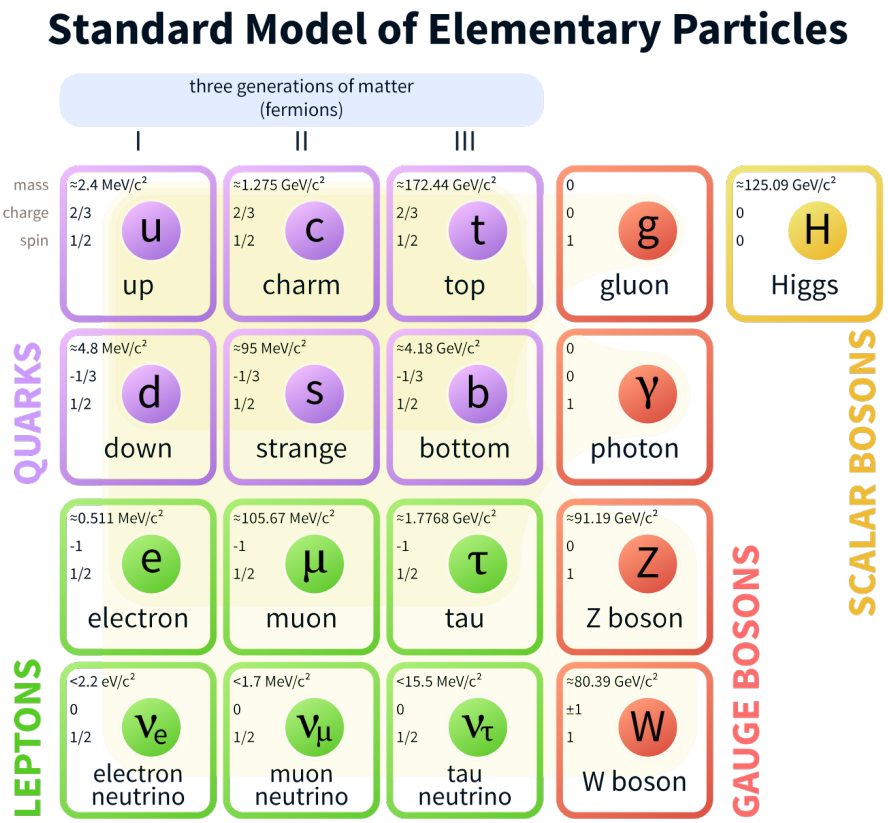


Fig. 1.1. A table of quarks, leptons and intermediate bosons. [3]

Intermediate Bosons: The last category of particles shown in Fig. 1.1 is intermediate bosons. The gauge bosons, are the quanta of strong (8 gluons), weak (W^+,W^- and Z^0) and electromagnetic (photon) field. The one scalar boson is the Higgs boson which is the boson that belongs to higgs field - the field that causes through interaction with all the other fields that their quanta gain their mass.

1.2 Fundamental interactions

The four forces above are all except for the gravitational described in terms of quantum field theory. And the theory that describes the behaviour of Electromagnetic, Weak and Strong interaction together is called the Standard Model (SM) of particle physics. The gravitational interaction is the weakest interaction and plays almost no role in particle physic (except for experiments like AeGiS or GBar that explore the effects of gravity on antimatter). It is described using general relativity that is separate from SM and will not be discussed in this thesis.

1.2.1 Weak interaction

It is the weakest of the three interactions in SM with about $1.166 \cdot 10^{-5}$ the strength of the strong interaction at the m_Z scale [4]. Its quantum field formulation is called Quantum Flavour Dynamics (QFD), although this name is not usually used. It can affect every known particle but it has a finite

range of about 10^{-18} m because of the high mass of its gauge bosons. The high mass is a consequence of spontaneous symmetry breaking caused by the Higgs mechanism. It is interesting that existence of the boson Z was discovered and explained first after unification of the electromagnetic and weak interaction into one Electoweak (EW) theory and the subsequent the discovery of neutral currents in data measurement. Up until that point it was thought that there were only the charged bosons $W^{\pm}$ that are responsible for the β-decay of nuclei.

1.2.2 Electromagnetic interaction

It has the middle strength in the SM with about $\frac{1}{137.035}$ the strength of the strong interaction [4]. Its quantum field formulation is called Quantum Electrodynamics (QED). It can only influence particles that carry electric charge and it has infinite range due to the zero mass of photon. It is the best known interaction and is described by the Maxwell equations. It is the interaction that governs our day-to-day life because it binds electrons and nuclei into atoms and atoms into molecules and even light itself is an electromagnetic field.

1.2.3 Strong interaction

It is the strongest known interaction. Its quantum field formulation is called Quantum Chromodynamics (QCD). It influences only particles that carry colour charge(there are three colours and three anti-colours)but it has a finite range of about 10^{-15} m which is roughly the diameter of a proton. This of course does not mean that the interaction does not reach beyond this threshold and the remnant force that remains is the nuclear force that binds nuclei together in a loose analogy to the Van der Waals force, a remnant of EM interaction, that can bind atoms into molecules. The short range of the strong force is caused by the peculiar property of gluons that they themselves carry colour charge and thus can interact with other gluons. Another peculiarity of QCD is that there can be no free colour charge and thus all particles that carry colour charge must be grouped in such a manner that their collective colour charge is zero (sometimes called white).

More consequences and applications of QCD are discussed in the next chapter about hard scattering processes.

1.2.4 Standard model

The Standard Model of particle physics is the theory that binds together the three interactions above along with the Higgs mechanism into one compact theory. The Lagrangian of this theory is rather extensive and so it is not mentioned here because it plays only a background role in the context of this thesis as an underlying principle. The main theoretical idea behind its construction was that it should be invariant under the groups of symmetries SU(3)×SU(2)×SU(1) from which properties and/or existence of elementary particles arise. It describes how quarks and leptons come to be and how they acquire their mass through the Higgs mechanism. Another part is describing the electroweak interaction and the breaking of the SU(2) symmetry that causes gauge bosons of weak interaction to gain mass. Yet another part describes the existence of colour and why there are 8 gluons (this is caused by the fact that the Lagrangian is invariant under the group of SU(3)). The last part of the Lagrangian deals with the virtual particles. It is clear that the standard model is not perfect as evidenced by observations of neutrino oscillations [5] that hint at the non-zero mass of neutrinos which is in contradiction with the SM condition that the neutrinos have similar to the photon zero mass. However this does not mean that SM is completely wrong because its predictions are in many experiments highly precise.

Chapter 2

Hard Scattering Processes

Topic of this chapter is the hard scattering formalism and its use in calculating cross sections of chosen processes in perturbative QCD (pQCD). At first the Feynman diagram formalism is put forth along with the factorization theorem for calculation of cross section in hard scattering.

2.1 Feynman diagrams

The Feynman diagrams are one of the core concepts in pQCD. Each diagram is the graphical representation of certain order of perturbative expansion of a scattering process. The used orders of pQCD in this thesis are either Leading Order (LO) or Next to Leading Order (NLO). This use is in the generators Herwig 7 (LO) and POWHEG (NLO). A few examples of Feynman diagrams for 2 to 2 process (two partons scattering to two partons) in LO and NLO is shown in Fig. 2.1 and Fig. 2.2.

Most of the LO Feynman diagrams for parton parton (p) scattering resulting in two partons (j) are drawn in Fig. 2.1 drawing all would be redundant because the shape of the diagram would be essentially unchanged. The drawing is done using the MadGraph web application [6] with the configuration $p\ p > j\ j$ for quarks $u\ d$ and their antiparticles. It can be seen that the processes shown are usually quark interacting with another quark or anti-quark through a gluon. Another type of interaction is with where gluons directly interact with quark or another gluon or even completely without the presence of any quarks. About half of the diagrams shown in Fig. 2.1 could be omitted due to them being the same interactions between gluon and quark only instead of quark there is an anti quark. It was however deemed educational to show all of the diagrams as they are as to remind oneself of how many diagrams are needed to compute even a seemingly simple QCD process. It is plain that leading order diagrams are straightforward, simple structures that consist of only two interaction vertices (or extraordinarily of one vertex for gluon-gluon scattering). A few examples of NLO Feynman diagrams are shown in Fig. 2.2. Such processes are identified by the presence of one more interaction vertex than in LO diagram. Such vertex can be a gluon loop, the additional radiation of a gluon from initial/final state quark or gluon, the radiation of a gluon by one quark and its absorption by another quark or gluon.

2.2 Hard scattering formalism

The formalism of hard scattering is based on the results of the deep inelastic scattering (DIS) development from which we obtain some basic functions and concepts. DIS describes the scattering of energetic lepton on a proton resulting in dissociation of proton into system of particles X with total of

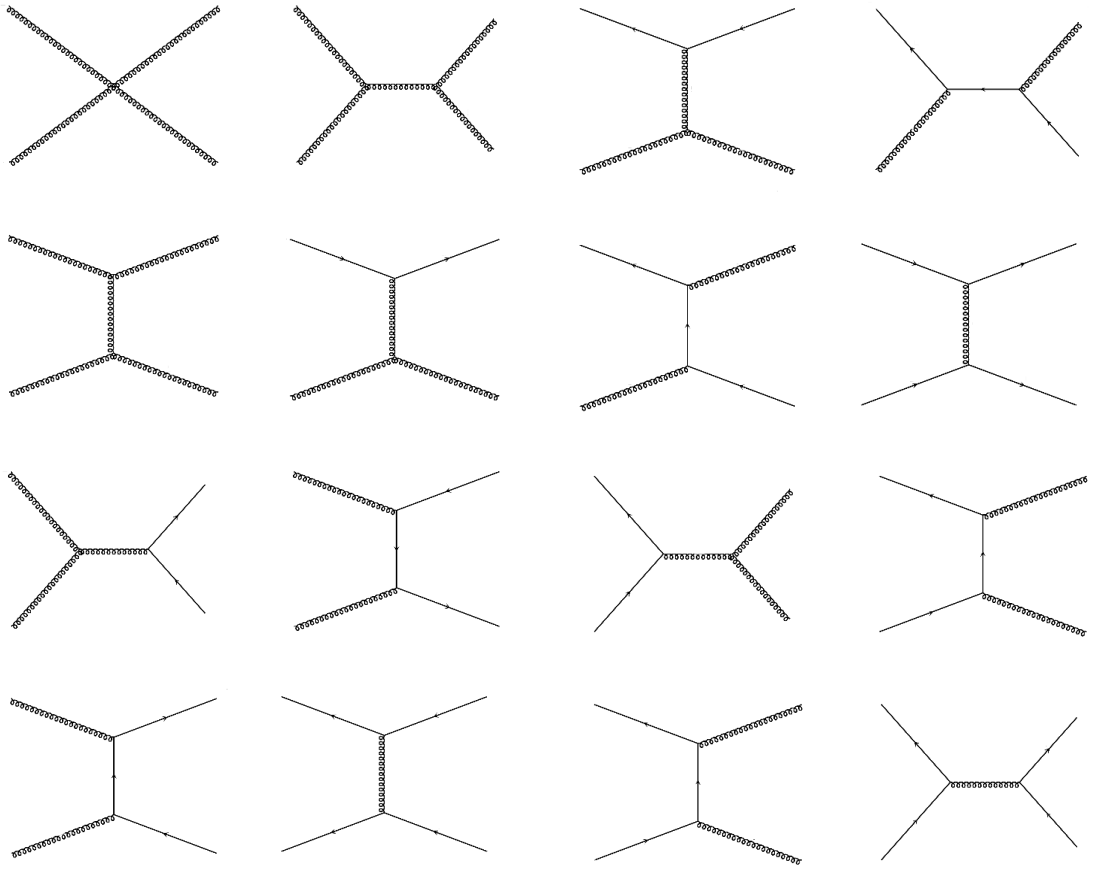


Fig. 2.1. A few examples of LO Feynman diagrams for proton proton scattering drawn using MadGraph [6]

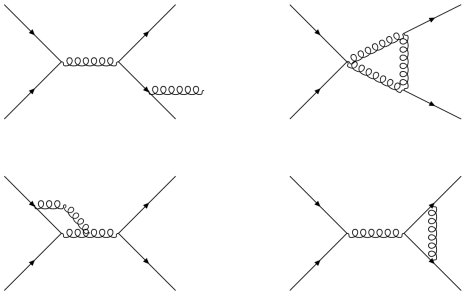


Fig. 2.2. A few examples of NLO Feynman diagrams.

mass M_X. In DIS process the lepton scatters off of a quark via a virtual photon exchange. The momentum of this photon is denoted q^2.

The first concept is virtuality, quantity defined as negative of squared transfer of momenta $Q^2 = -q^2$. If Q^2 is less than the mass of the particle then the particle is called off mass-shell. Second concept is the parton distribution function (PDF) $f_{a/A}(x)$ for each type of parton in a hadron A. PDFs are the experimentally obtained functions that can be interpreted as weighting functions for calculating of cross

sections at the hadron level. The simplest interpretation is at LO: a PDF determines the probability of finding a parton a, of the given flavour, in the hadron A and carrying the fraction x of the total momentum of the hadron. The resulting cross section $\sigma_{(AB\to X)}$ for the hard scattering of two hadrons resulting in a set of general final state particles ($AB \to X$) is derived using the factorization theorem:

$$\sigma_{AB\to X} = \int dx_a dx_b f_{a/A}(x_a) f_{b/B}(x_b) \hat{\sigma}_{ab\to X}, \tag{2.1}$$

where $\hat{\sigma}_{ab\to X}$ are the cross sections of all subprocesses that contribute to the whole process and these are weighted by the appropriate PDFs. This cross section is valid in the analogue of Bjorken scaling limit in DIS: M_X, $s \to \infty$, $\tau = M_X^2/s$ fixed. A graphical representation of such a process is shown in Fig. 2.3

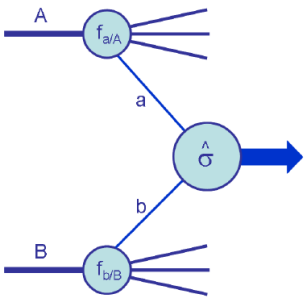


Fig. 2.3. A graphic representing the hard scattering process. [7]

This predictions are successful for many hard scattering processes at first. However as more data are taken in and the perturbative correction of both virtual and real gluon emissions are calculated there seems to arise some problems. The convergence of the perturbative expansion seems to be spoiled by logarithms from partons emitted collinear with the incoming partons. This can be fixed using the parton evolution equation DGLAP (Dokshitzer–Gribov–Lipatov–Altarelli–Parisi) (2.2) and (2.3) for redefining the quark an gluon PDFs into their renormalized forms where the factorization scale μ_F is introduced to ensure the logarithmic scaling.

$$\frac{\partial q_i(x,\mu_F^2)}{\partial \log \mu_F^2} = \frac{\alpha_S}{2\pi} \int_x^1 \frac{dz}{z} \left[P_{q_i q_j}(z,\alpha_S) q_j(\frac{x}{z},\mu_F^2) + P_{q_i g}(z,\alpha_S) g(\frac{x}{z},\mu_F^2) \right], \tag{2.2}$$

$$\frac{\partial g(x,\mu_F^2)}{\partial \log \mu_F^2} = \frac{\alpha_S}{2\pi} \int_x^1 \frac{dz}{z} \left[P_{g q_j}(z,\alpha_S) q_j(\frac{x}{z},\mu_F^2) + P_{gg}(z,\alpha_S) g(\frac{x}{z},\mu_F^2) \right], \tag{2.3}$$

where $q_i(z,\mu^2)$ and $g(z,\mu^2)$ are the PDFs of the quark of flavour i and gluon and μ_F is the appropriate scale. The distinction between the gluon and quark PDF is crucial, because they behave differently. The $P_{ab}(z,\alpha_S)$ are splitting functions where the first index denotes the final particle that carries the momentum fraction x and the second index is the initial particle that undergoes splitting. The α_S is the strong coupling constant along which the splitting function is expanded into perturbation series:

$$P_{ab}(z,\alpha_S) = P_{ab}^0(z) + \frac{\alpha_S}{2\pi} P_{ab}^1(z) + \dots \tag{2.4}$$

In the following it is assumed that all perturbations are taken only in the Leading Order. If the PDFs acquired from DGLAP equations (2.2) and (2.3) are used in (2.1) the cross section becomes:

$$\sigma_{AB} = \int dx_a dx_b f_{a/A}(x_a, Q^2) f_{b/B}(x_b, Q^2) \hat{\sigma}_{ab\to X}, \tag{2.5}$$

where Q^2 is the large momentum scale that characterizes hard scattering. The last step in the though process is to realize that final corrections that remain after the factorization of the logarithms into PDFs are not universal and need to be calculated independently for each subprocess and so the perturbative expansion of its cross section is used:

$$\sigma_{AB} = \int dx_a dx_b f_{a/A}(x_a, \mu_F^2) f_{b/B}(x_b, \mu_F^2) \times [\hat{\sigma}_0 + \alpha_S(\mu_R)\hat{\sigma}_1 \,]_{ab\to X}\,, \tag{2.6}$$

for μ_F now denoting the factorization scale, which separates short- and long-distance physics of the process and μ_R is the renormalization scale for the QCD running coupling. The usual choice is to take $\mu_F = \mu_R$ and then simply change the notation to $\mu_F = \mu_R = Q$ which is some typical scale for the process. Now follows a certain "cookbook" for the calculation of LO cross section for a inclusive final state X + anything:

1. Find every LO subprocesses that contribute into X.
2. Choose the right factorization and renormalization scale μ_F and μ_R.
3. Compute appropriate $\hat{\sigma}_0$ for all such subprocesses including their interferences.
4. Combine this $\hat{\sigma}_0$ with the right PDFs of incoming partons a and b.
5. Numerically integrate over x_a and x_b and every other phase space variable that can be associated with the final state X.

It is important to note that every time the perturbative expansion is used it should be calculated into the same order as to abide good manners.

Chapter 3

LHC And The ATLAS Experiment

The LHC is currently the largest particle accelerator ever build by the mankind. It is located at CERN in a tunnel near Geneva in Switzerland approximately 100 m, depending on various factors, under the surface. The tunnel has a 27 km circumference in order to allow the LHC to reach its immense power of 6.5 TeV per beam in proton-proton collisions. Surely enough the LHC is not capable of accelerating particles from scratch on its own, and so there is a number of pre-accelerators needed to boost the particles before they are injected into the LHC itself. In the following paragraph there is na elaboration as to how this accelerating process takes place, however as this is not the main point of this theses, it is not extremely thorough.

Accelerator cascade

- **Proton source:** The process starts with a simple bottle of hydrogen gas, which is stripped of its electrons using an electric field in order to yield protons.
- **LINAC 2:** The first step of the cascade is the Linear accelerator 2, where the protons reach the energy of 50 MeV before entering the next step.
- **PSB:** The proton synchrotron booster consists of 4 superimposed synchrotron rings that accelerate the 50 MeV protons from LINAC 2 up to 1.4 GeV. This step is important because it allows 100 times more protons to be injected in the next step of the cascade than it would be possible going directly from LINAC 2.
- **PS:** The Proton Synchrotron is an important part of the accelerator cascade, due to its role as a supplier of a whole range of particles from protons and electrons, light nuclei, to heavy ions and antiparticles. The supplied protons have the energy of 25 GeV.
- **SPS:** The penultimate step in acceleration of the particles is the Super Proton Synchrotron. Having almost 7 km in circumference it is the second largest machine in CERN's accelerator complex and has served a great many things such as the Nobel-prize-winning W and Z boson discovery in 1983, which utilized the SPS's versatility using it as a proton-antiproton collider. In the SPS the protons are accelerated to 450 GeV
- **LHC:** The Large Hadron Collider is the final step of the cascade and is fed from the SPS into two beam pipes in which the protons circulate in opposite directions. The process of filling the LHC takes 4 minutes and 20 seconds per pipe and then it takes about 20 additional minutes to accelerate the protons to their maximum energy of 6.5 TeV.

The LHC is distinguished from other accelerators at CERN not only by its impressive size but the main advantage is in the technology used in constructing it. Where for the rest of the machines conventional electromagnets had been used, the LHC is equipped with superconducting electromagnets. Because of this LHC is able to drive and contain truly astonishing energies. To be precise the LHC has 1232 bending dipole magnets, each 15 m in length, a peak magnetic field of 8.33 T and 392 quadrupole magnets, 3-7 meters long, that focus the beams inside the ultrahigh vacuum pipes (the pressure in these pipes is $1.01325 \cdot 10^{-13}$ bar). The protons in the LHC are not distributed uniformly, but are being grouped to 2808 bunches (approximately $1.15 \cdot 10^{11}$ protons per bunch). The time distance between the bunches is 25 ns.The luminosity of LHC is 10^{34} cm^{-2}s^{-1}. The beam pipes intersect at four points along the path and at each intersection an experiment takes place, these are: ALICE, CMS, LHCb and ATLAS.

Owning to the theme of this thesis an introduction of the basic technical design of the ATLAS detector is presented in the next section.

3.1 ATLAS

First a brief overview of the detector is put forth followed by some information on tracking, calorimetry, muon system and forward detectors. Lastly in this section some designs concerning trigger and data acquisition systems are introduced.

However before it comes to these, the coordinate system and a few other concepts used throughout this thesis are described here. The origin of the coordinate system is defining the nominal interaction point, the beam direction defines the z-axis and the x-y plane is transverse to the beam direction. The positive x-axis is defined so that it points from interaction point to the centre of LHC ring and the y-axis points upwards. As per usual the azimuthal angle ϕ is measured around the beam axis and the polar angle Θ is the angle from the beam axis. The pseudorapidity is defined as $\eta = -\ln\tan(\frac{\Theta}{2})$ and the rapidity is defined as $y = \frac{1}{2}\ln[\frac{E+p_z}{E-p_z}]$. Furthermore all the transverse quantities (momentum, energy and missing energy) p_t, E_t and E_t^{miss} are defined in the $x-y$ plane unless stated otherwise. The distance ΔR in the pseudorapidity-azimuthal angle space is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

3.1.1 Overview of the detector

The general layout of the detector can be found in Figure 3.1. A Toroidal LHC ApparatuS (ATLAS) detector is set up to be forward-backward and azimuthally symmetric with respect to the interaction point. ATLAS is 25 m in height and 44 meters in length. In the text terms barrel region (or just barrel) and end-cap region (or just end-caps) are used. Those refer only to the appropriate region in the respective sub-detector and should they refer to the whole ATLAS it would be explicitly stated. Furthermore these regions are defined by the pseudorapidity range that is covered by the machinery. The magnet configuration is as follows: the inner-detector cavity is surrounded by a thin superconducting solenoid and around the calorimeters are eight times three large superconducting toroids (that is always one in the barrel and two in the end-caps) arranged so that an eight-fold azimuthal symmetry around the calorimeters is achieved.

In the Inner Detector (ID), a combination of discrete, high-resolution pixel and strip detectors in the inner part of the tracking volume, and straw-tube tracking detectors capable of generating and detecting transition radiation in the outer part, allows to achieve pattern recognition, momentum and vertex measurement and electron identification.

Energy and position resolution is provided by high granularity liquid-argon electromagnetic sampling calorimeters. The hadronic calorimetry is provided by a scintillator-tile calorimeter, segmented into a large barrel and two smaller extended barrel cylinders on both sides of the central barrel. For the

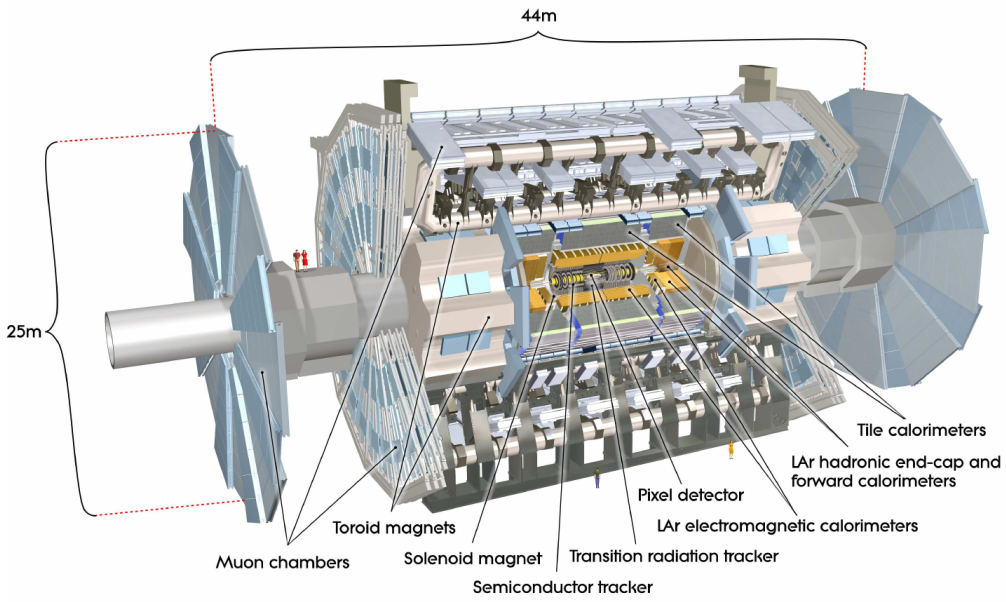


Fig. 3.1. The general layout of the ATLAS detector. Taken from [8]

hadronic calorimetry is also used the liquid-argon technology matching the outer limits in pseudorapidity for the end-cap EM calorimeters. The liquid-argon forward calorimeters provide electromagnetic as well as hadronic energy measurements and extend the coverage in the high $|\eta|$ range.

The last part that defines the overall dimensions of the ATLAS detector is the muon spectrometer. It surrounds the calorimeter and its air-core toroid system, with long barrel and two inserted end caps generates a strong bending force in a large volume. Excellent muon momentum resolution is achieved with three layers of high precision tracking chambers. Some of the key components in the muon spectrometer are the trigger chambers with time resolution approximately 1.5-4 ns. With this being the brief overall introduction of the detector some further elaboration on its functions is presented in the following sections.

3.1.2 Tracking

The detail of the ATLAS inner detector is to be found in Figure 3.2. Because of the high track density in the detector, caused by approximately 1000 emerging particle every 25 ns from the collision point within $|\eta| < 2.5$, the detector needs to be able to do high-precision measurements and have sufficient granularity. Therefore pixel and silicon microstrip (SCT) trackers, along with the straw tubes of the Transition Radiation Tracker (TRT), are used.

The inner detector is immersed in a 2 T magnetic field generated by the central solenoid whose dimensions are 5.3 m in length and 2.5 m in diameter. The aforementioned critical region of $|\eta| < 2.5$ is covered by the precision tracking detectors. Within the barrel region, these trackers are arranged on concentric cylinders aroun the beam axis, whereas in the end-cap region they are located on discs perpendicular to the beam axis. In order to achieve the highest granularity around the vertex region silicon pixel detectors are implemented. All pixel sensors are identical and their minimum pixel size in $R-\phi \times z$ is 50×400 μm^2. Pixel layers are segmented in $R-\phi$ and z with three pixel layers crossed by one track in average. The intrinsic accuracies in the barrel are 10 μm in $R-\phi$ and 115 μm in z, in the discs 10 μm in $R-\phi$ and 115 μm in R. The pixel detector by itself has approximately 80.4 million readout channels.

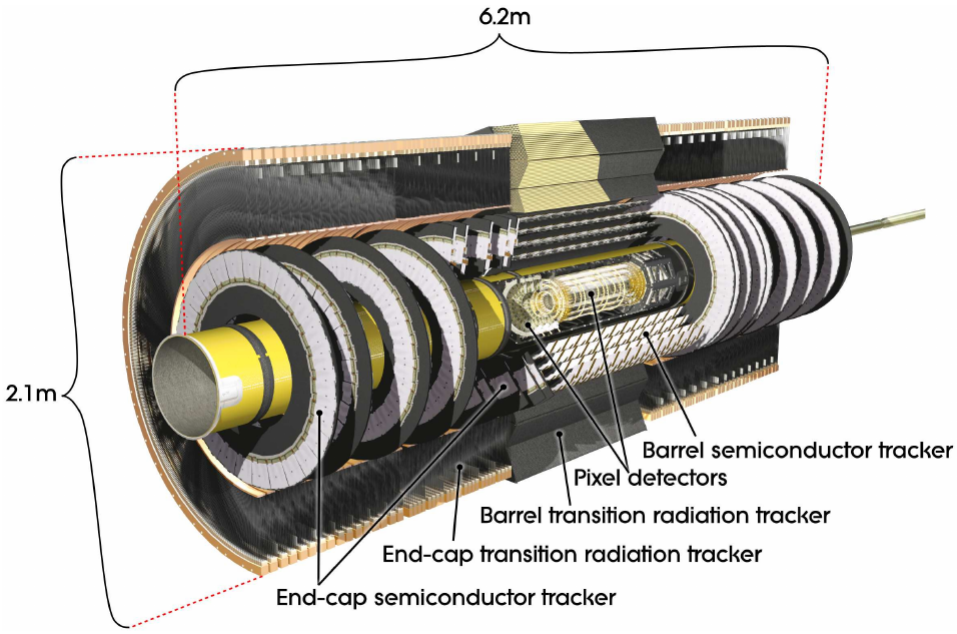


Fig. 3.2. The view of the ATLAS inner detector. Taken from [8]

Where the SCTs are concerned, eight strip layers are crossed by each track. 40 mrad - small angle stereo strips are used for measuring both coordinates in the barrel region with one set of strip in each layer parallel to the beam direction measuring $R - \phi$. They consist of two 6.4 cm long daisy-chained sensors with a strip pitch of 80 μm. In the end-caps, a set of strips running radially and a set of stereo strips at 40 mrad angle are implemented in the detectors. The mean strip pitch of those is approximately 80 μm as well. The intrinsic accuracies per module in the barrel region are 17 μm in $R - \phi$ and 580 μm in z, in the discs are 17 μm in $R - \phi$ and 580 μm in R. The total number of readout channels in the SCT is roughly 6.3 million.

Another part of the tracking system is the transition radiation tracker. It provides a large number of hits (about 36 per track) in its 4 mm diameter straw tubes and enables track-following up to $|\eta| = 2.0$. Alas, TRT is only able to provide $R - \phi$ information with accuracy of 130 μm per straw. In the barrel region the straws are installed parallel to the beam axis. Their length is 144 cm and their wires are divided into two halves at approximately $\eta = 0$. In the end caps, the straws are 37 cm long and radially arranged in wheels. The Total number of TRT readout channels is about 351 000.
The overall combination of precision trackers at small radii and the TRT at larger radius allows for a very robust pattern recognition and high precision in both $R - \phi$ and z coordinates. The straw hits at the outer radius are especially important for momentum measurement because the lower precision per point compared to the silicon is compensated by the large number of measurements and longer measured track length.

To conclude, the inner detector provides tracking measurements in a range matched by the precision measurements of the EM calorimeter. The e^- indentification capabilities are enhanced via the detection of photons occuring due to transition radiation in the xenon-based gas mixture of the straw tubes. The semiconductor trackers allow impact parameter measurements and vertexing for heavy flavour and τ lepton tagging. The enhancement of the secondary vertex measurement is possible thanks to the innermost layer of pixels at a radius of about 5 cm.

3.1.3 Calorimetry

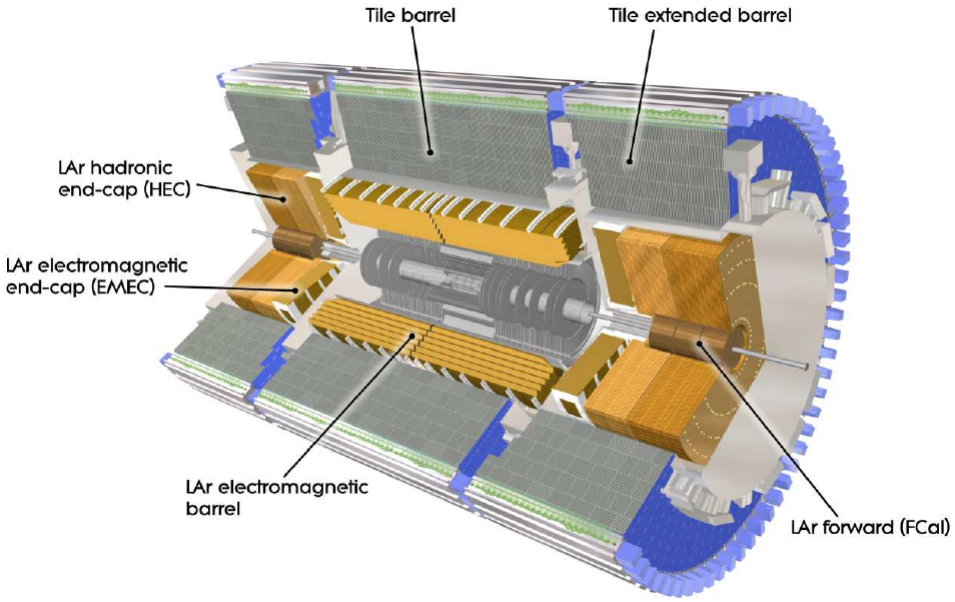


Fig. 3.3. The overall view of ATLAS calorimeter system. Taken from [8]

In order to have a better sense of the layout of the calorimeters, the system is depicted in Figure 3.3. The range covered by the calorimeters is $|\eta| < 4.9$, using different techniques that are tuned to fit the widely varying requirements of the physics processes of interest and the radiation environment over this large pseudorapidity range. In the region matching the inner detector, the fine granularity of the electromagnetic calorimeter is suited for precision measurements of electrons and photons. The rest of the calorimeters is designed with sufficient granularity to satisfy the physics requirements for jet reconstruction and E_T^{miss} measurements. It is crucial that the calorimeters are able to provide good containment for EM and hadronic showers and also limit punch-through into the muon system. Therefore, the depth of the calorimeter is an important parameter to be taken into account. The toal thickness of the EM calorimeter is >22 radiation lengths in the barrel region and >24 radiation lengths in the end-caps. The approximate 9.7 interaction lengths (λ) of active calorimeter in the barrel and 10λ in the end-caps are deemed adequate to provide good resolution for high-energy jets. The overall thickness, including 1.3 interaction lengths from the outer support, is 11λ at zero pseudorapidity and has been shown both by measurements as well as simulations to be sufficient to reduce punch-through bellow the irreducible level of prompt or decay muons. In the following paragraphs the main types of calorimeters are introduced

3.1.3.1 LAr electromagnetic calorimeter

The EM calorimeter is divided into two parts: the barrel part covering $|\eta| < 1.475$ and the two end-caps components covering $1.375 < |\eta| < 3.2$. Each of these parts is housed in their own cryostat. The central solenoid and the LAr calorimeter share a common vacuum vessel as a consequence of optimisation of the material in order to achieve the desired calorimeter performance and thereby eliminating two vacuum walls. The barrel calorimeter is comprised of two identical half-barrels that are separated by a small gap of 4 mm at $z = 0$. Each end-cap calorimeter is mechanically divided into two coaxial wheels: the inner wheel covering the range of $2.5 < |\eta| < 3.2$ and the outer wheel covering the range of

$1.375 < |\eta| < 2.5$. The electromagnetic calorimeter is a lead-LAr detector with accordion-shaped kapton electrodes and lead absorber plates over its full coverage. The accordion geometry provides complete ϕ symmetry without cracks. The lead thickness in the absorber plates has been decided to be best suited for the according pseudorapidity in terms of EM calorimeter performance and energy resolution. Over the precision physics region ($|\eta| < 2.5$) the EM calorimeter is segmented into three sections in depth. For the end-cap inner wheel, the calorimeter is segmented only in two sections in depth and has a coarser lateral granularity than the rest of the acceptance. For the $|\eta| < 1.8$ region a presampler detector is used to correct for the energy lost by electrons and photons upstream of the calorimeter. The presampler consists of an active LAr layer of 1.1 cm and 0.5 cm thickness for the barrel and end-cap region respectively.

3.1.3.2 Hadronic calorimeters

Tile Calorimeter The tile calorimeter can be found directly outside the electromagnetic calorimeter envelope. The region $|\eta| < 1.0$ is covered by its main barrel and its two extended barrels cover the range $0.8 < |\eta| < 1.7$. In this sampling calorimeter steel is used as the absorber and scintillating tiles as the active material. The tile calorimeter is azimuthally divided into 64 modules both in the barrel and the extended barrel regions. Radially, the tile calorimeter extends from an inner radius of 2.28 m to an outer radius of 4.25 m. The tile calorimeter is segmented in depth in three layers, approximately 1.5, 4.1 and 1.8 interaction lengths thick for the barrel and 1.5, 2.6 and 3.3 interaction lengths in the extended barrels. The overall thickness of the detector at the outer edge of the tile detector region is 9.7 interaction lengths at zero pseudorapidity. The readout of two sides of the scintillating tiles is secured by wavelength shifting fibres into two separate photomultiplier tubes. In η, the readout cells built by grouping fibres into the photomultipliers are pseudo-projective towards the interaction region.

LAr HEC The Hadronic End-cap Calorimeter (HEC) is comprised of two independent wheels per end-cap, located right behind the end-cap EM calorimeter and sharing the same LAr cryostats. In order to soften the drop in material density at the transition between the end-cap and the forward calorimeter at about $|\eta| = 3.1$, the HEC is extended out to $|\eta| = 3.2$ and is thereby overlapping with the forward calorimeter. In this fashion, the HEC η range is also overlapping that of the tile calorimeter by being extended to $|\eta| = 1.5$. Each wheel is build from 32 identical wedge-shaped modules assembled with fixtures at the periphery and at the central bore. Every wheel is divided into two segments in depth for a total of four layers per end-cap. For all wheels the plate closer to the interaction point is half the thickness of the further, namely it is 25 mm of copper for the closer and 50 mm of copper for the further one. The outer radius of the copper plates is 2.03 m, while the inner radius is 0.475, except in the region overlapping with the forward calorimeter where the radius becomes 0.372 m. There are 8.5 mm thick gaps between the copper plates. These gaps are filled with LAr thus providing the active medium for this sampling calorimeter.

LAr FCal The Forward Calorimeter(FCal) is integrated into the end-cap cryostats because this provides clear benefits to uniformity of the calorimetric coverage as well reduces radiation background levels in the muon spectrometer. As it was aimed to reduce the amount of neutron albedo in the inner detector cavity, the front face of the FCal is recessed by approximately 1.2 m with respect to the EM calorimeter front face. Because this limits the depth of the calorimeter, it was therefore decided to implement a

high-density design. The depth of FCal is about 10 interaction lengths, its composition is divided into three modules in each end-cap. The first is made of copper and suited for EM measurements. The other two are made of tungsten and are used to measure predominantly the energy of hadronic interactions. Each of these modules consists of a metal matrix with regularly spaced longitudinal channels filled with the electrode structure that is made of concentric rods and tubes parallel to the beam axis. The sensitive medium is the LAr between the rod and the tube. This manner of geometry was chosen for allowing excellent control of the gaps, which are as small as 0.25 mm in the first section, in order to avoid problems due to ion build-up.

3.1.4 Muon system

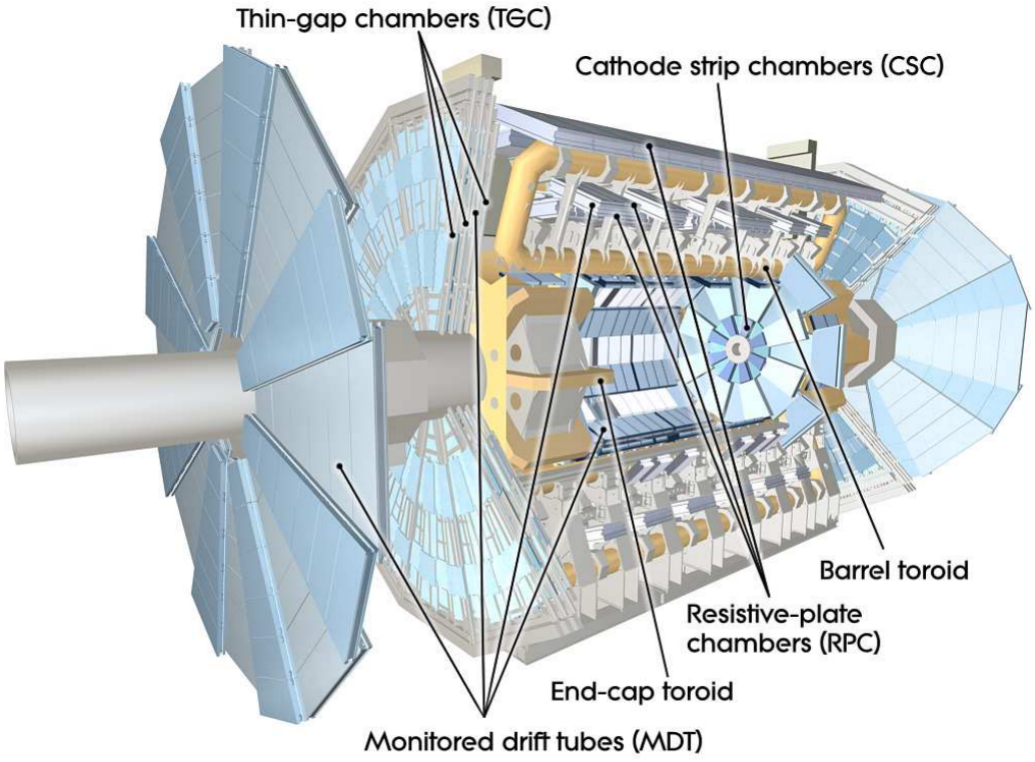


Fig. 3.4. The cut-away view of the ATLAS muon system. Taken from [8]

The layout of the muon spectrometer is shown in Figure 3.4. The main concept of the spectrometer is the deflection of muon tracks in the large superconducting air-core toroid magnets, instrumented with separate trigger and high-precision tracking chambers. For $|\eta| < 1.4$ magnetic bending is provided by the large barrel toroid. Over the range $1.6 < |\eta| < 2.7$, two smaller end-cap magnets inserted into both ends of the barrel toroid are used for track bending. Over the transition region $1.4 < |\eta| < 1.6$ magnetic deflection is provided by a combination of barrel and end-cap fields. This configuration is of course not random but provides a field that is mostly orthogonal to the muon trajectory while minimising the degradation of resolution caused by multiple scatterings. In the barrel region, muon tracks are measured in chambers arranged in three cylindrical layers around the beam axis. In the transition and end-cap region, the chambers are installed in planes perpendicular to the beam, also in three layers. Now the specification of the muon system is discussed more thoroughly. [8]

3.1.4.1 The toroid magnets

The magnetic field for the muon spectrometer is generated by a system of three large air-core toroids. The two end-cap toroids are inserted at both ends of the barrel toroid and line up with the central solenoid. Each of the three toroids consists of eight coils assembled radially and symmetrically around the beam axis. In order to provide radial overlap and optimised the bending power in the transition region between the barrel and end-cap regions, the end-cap toroid coil system is rotated by 22.5°.

The coils in the barrel are housed in eight individual cryostats with linking elements between them ensuring the overall mechanical stability. Each end-cap toroid consists of eight coils with racetrack-like winding housed in an aluminium alloy. Each coil has two double-pancake type windings. They are cold-linked and assembled as a single cold mass, that is they are housed in one large cryostat. Therefore the internal forces in the end-cap toroids are balanced by the cold supporting structure between the coils, which is a different approach than in the barrel toroid.

The performance in terms of bending power is characterised by the field integral $\int Bdl$, where B is the component of the magnetic field normal to the muon direction and the integral is computed along an infinite-momentum muon trajectory, between the innermost and outermost muon-chamber planes. In the range of $|\eta| \in (0, 1.4)$ provides the barrel toroid 1.5 to 5.5 Tm of bending power. In the region of $|\eta| \in (1.6, 2.7)$ the bending power 1 to 7.5 T·m is supplied by the end-cap toroids. In the transition region $|\eta| \in (1.4, 1.6)$ where the two magnets overlap the bending power is lower.

3.1.4.2 Muon chamber types

In the bulk of the η-range, a precision measurement of the track coordinates in the principal bending direction of the magnetic field is provided by Monitored Drift Tubes (MDT). The mechanical insulation in the drift tubes of each sense wire from its neighbours guarantees a robust and reliable operation. For $|\eta| \in (2.0, 2.7)$ in the innermost plane the Cathode Strip Chambers (CSC) are used because of their higher granularity to withstand the demanding rate and background conditions. In order to satisfy the stringent requirements on the relative alignment of the muon chamber layers, a combination of precision mechanical assembly techniques and optical alignment system both within and between muon chambers are used.

The pseudorapidity range $|\eta| < 2.4$ is covered by the trigger system. Resistive Plate Chambers (RPC) are used in the barrel and Thin Gap Chamber (TGC) in the end-caps. The three purposes of the trigger chambers for the muon spectrometer are the following: bunch-crossing identification, well-defined p_T thresholds and the measurement of the muon coordinate in the direction orthogonal to that determined by the precision-tracking chambers. [8]

3.1.4.3 Muon chamber alignment and B-field reconstruction

Due to the large areas involved, particularly at the highest momenta, the overall performance of the detector is dependent on the alignment of the muon chambers with respect to each other and with respect to the overall detector.

The accuracy of the stand-alone muon momentum measurement necessitates a precision of 30 μm on the relative alignment of chambers, both within each projective tower and between consecutive layers in immediately adjacent towers. The internal deformations and relative positions of the MDT chambers are monitored by roughly 12000 precision-mounted alignment sensors, all based on the optical monitoring of deviations from straight lines. The reconstruction and/or monitoring of the chamber positions is relying on slightly different strategies and sensors in the end-cap and barrel regions due to geometrical constraints.

The accuracy required for the relative positioning of non-adjacent towers, needed to obtain adequate mass resolution for multi-muon final states, lies in the few millimetre range. The relative alignment of the barrel and forwards region of the muon spectrometer, of the calorimeters and of the inner detector relies on high-momentum muon trajectories.

For magnetic field reconstruction, the goal is to determine the bending power along the muon trajectory to a few parts in a thousand. The field is continuously monitored by a total of about 1800 Hall sensors distributed throughout the spectrometer volume. Their outputs are compared with simulations and used for reconstructing the position of the toroid coils in space, as well as to account for magnetic perturbations induced by the tile calorimeter and other nearby metallic structures. [8]

3.1.5 Forward detectors

The forward region of ATLAS is covered by three smaller detectors. The main function of the first two systems is to determine the luminosity delivered to ATLAS. At 17 m from the interaction point on both sides lies the LUminosity measurement using Cerenkov Integrating Detector(LUCID). It detects inelastic $p - p$ scattering in the forward direction and is the main online relative-luminosity monitor for ATLAS. The second detector is the Abolute Luminosity For ATLAS (ALFA) located at both sides 240 m from the interaction point. It consist of scintillating fibre trackers located inside Roman pots which are designed to approach as close as 1 mm to the beam. The third system is the Zero-Degree Calorimeter(ZDC), which plays a key role in determining the centrality of heavy-ion collisions. It is located 140 m from interaction point again on both sides, just beyond the point where the common straight-section vacuum-pipe divides back into two independent beam-pipes. The ZDC modules consist of layers of alternating quartz rods and tungsten plates which measure neutral particles at $|\eta| \geq 8.2$. [8]

3.1.6 Trigger, readout, data acquisition and control system

The Trigger and Data Acquisition, collectively TDAQ, systems, the timing and trigger control logic and the Detector Control System are partitioned into subsystems, typically associated with sub-detectors, which have the same logical components and building blocks. However those are not discussed here and can be found thoroughly explained for example in the 8. chapter of the ATLAS Technical Design Report [TDR].

The trigger system is divided into three main levels. L1, L2 and the event filter. On Each trigger level the decisions made at the previous level are refined and if deemed necessary additional selection criteria are applied. The data acquisition system receives and buffers the event data from the detector-specific readout electronics, that is at the L1 trigger rate over 1600 point-to-point readout links. Only a limited amount of the total detector data is used on the first level to make a decision in less than 2.5 μs, reducing the rate to about 75 kHz. On the other two levels more detector information are accessed for a final rate of up to 200 Hz with an event size of approximately 1.3 MB. [8]

Chapter 4

Jet Algorithms

The study of jets is an essential part of collider experiments for its use throughout the field. The main goal of this chapter is to bring some basic understanding of jet algorithms that are commonly used. To start, an intuitive albeit a little laymen definition of a particle jet is that it is a collimated spray of hadrons that were created during the hadronisation of a high-energy quark or gluon. Jets tend to be easily visually identifiable in an event display and when correctly measured their energy and direction should be the same as those of the original parton from which the jet came. It is however not good enough to visually identify the jets. Especially in high energy experiments where the number of particles and jets is high and so the need for automatisation arises. That is where jets algorithms step into the fray.

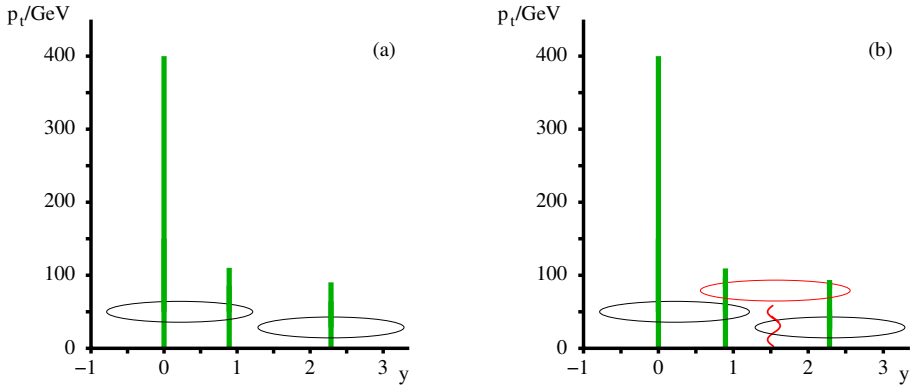
Jet algorithm is a set of rules that allows projecting a group of particles onto a group of jets. Each jet algorithm is typically dependent on one or more parameters. The combination of jet algorithm and its parameters is called jet definition. These are applicable onto a many different definitions of particles for example data from calorimeter towers or from the partonic events of perturbative QCD calculations. The main strength of this is that using these definitions we can meaningfully compare resulting jets from different kinds of inputs and thus compare experimental data with MC simulations.

Usually, most of the jet algorithms are classified into one of two broad classes: sequential recombination algorithms and cone algorithms.

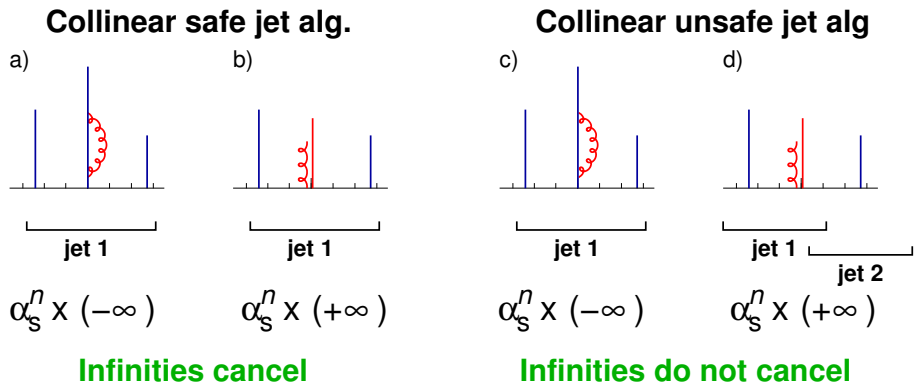
4.1 Cone algorithms

Cone algorithms utilize the premiss that jets can be defined by drawing a cone around some direction of dominant energy flow. Nowadays they serve only as a historical relict and their influence is mainly pedagogical and their description is accordingly brief. Usually, cone algorithms take some or all of the particles as seeds which define these trial directions. For each seed, the sum of 4-momenta of all particles in the trial cone then defines a new trial direction. Thus an iterative process is begun and its end is spelled when the direction of the resulting 4-momentum no longer changes and the result is called a stable cone. Because the direction and energy of a parton is within reasonable constraints unchanged by QCD radiation and hadronisation, stable cones should be close in direction and energy to the original parton. Cone algorithms in general are bound with the problems of collinear and infrared safety. Cone algorithm is considered infrared safe if addition of an infinitely soft particle does not change the number or directions of stable cones present without this particle. Infrared safety is sometimes solved by imposing a requirement of minimal p_T to be accepted as jet. Which brings us to the other problem that is collinear safety. It means that if one particle is divided into two of lower energy, the number of stable cones does not change. If there is some p_T threshold imposed on particles in a jet algorithm then the algorithm can be collineary unsafe. The issues of infrared and collinear safety are illustrated in

Fig. 4.1. There has been attempts at "fixes" for the iterative algorithms such as taking midpoint between stable cones as new seeds and reiterating the process. Those efforts are only postponing the inevitable divergence into higher-order of perturbative calculations.



(a) Infrared safety vs. unsafety



(b) Collineary safety vs. unsafety

Fig. 4.1. Illustrations of collinear and infrared safety in an iterative cone algorithm. The implications for perturbative calculations of the collinear (un)safety are hinted at as well. Horizontal axis is indicative of particle rapidity and the height of the bars is proportional to particle p_T. [9]

The problem of infrared and collinear safety may be bypassed by using exact seedless algorithms. Those are algorithms that instead of iterative approach find some exact way of solving the equation [10]:

$$D(p_{cone}, a) = 0, \text{ with } p_{cone} = \Sigma_i p_i \Theta(R - D(p_i, a)), \tag{4.1}$$

where $D(p, a)$ is angular distance measure of four-momentum p and the cone axis a, R is the given opening angle of the cone sometimes called cone radius. This manner of search for stable cones is usually computationally expensive, however the algorithm SISCone from [10] has successfully solved this issue.

4.2 Sequential recombination algorithms

All sequential recombination algorithms have some definition of distance between particles. Then using this definition, a pair of particles that are the closest to each other is identified and recombined. The recombined particles is then added to the list of so called pseudojets while the original particles are removed from the list of particles. After that the identification and recombination (the list of pseudojets

is included into the particle list) is repeated until some sort of stopping condition is met and the resulting pseudojet is added to the list of final jets.

4.2.1 Generalised k_t algorithm

As an example of a sequential jet algorithm the generalised k_t algorithm is now introduced. The core of the algorithm is the distance measure d_{ij} between the particles i and j that is defined as [11]:

$$d_{ij} = d_{ji} = min(p_{Ti}^{2p}, p_{Tj}^{2p}) \frac{\Delta R_{ij}^2}{R^2}, \tag{4.2}$$

$$\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2, \tag{4.3}$$

$$y_i = \frac{1}{2} \ln \left(\frac{E_i + p_{zi}}{E_i - p_{zi}} \right), \tag{4.4}$$

where p_{Ti} is the transverse momentum of particle i, p_{zi} is the component of its momentum parallel to the beam axis z, y_i is its rapidity and ϕ_i its azimuth. The parameter R is usually called the jet radius and it defines the angular reach of the algorithm. The notation is intentionally the same as is in the cone algorithm because for example jets produced by the anti-k_t algorithm mentioned below are cone-like and the R is indeed similar to its radius. The continuous parameter p then specifies which type of the k_t algorithm is currently used. The algorithm also involves the distance measure between particle and the beam defined as

$$d_{iB} = p_{Ti}^{2p}. \tag{4.5}$$

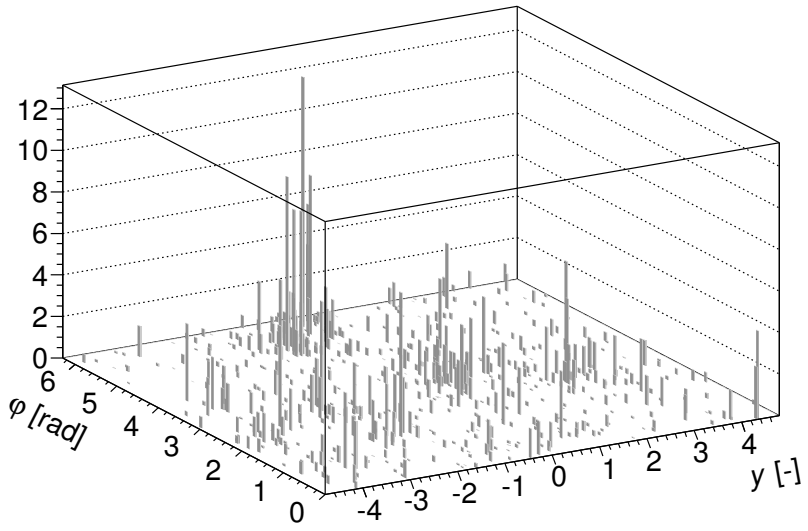
There are two formulations of the algorithm and those are the exclusive and inclusive. In the exclusive (generally used with $R = 1$) one, the smallest of d_{iB} and d_{ij} is marked. If it is the d_{ij}, the procedure is to replace both particles i and j with a single new object whose momentum is $p_i + p_j$ and is often called a "pseudojet", because it is neither a particle, nor yet a full jet. If on the other hand is d_{iB} the smaller number then the particle/pseudojet i is removed from the list and is declared a part of the beam jet. The whole process is then repeated until the smallest of d_{iB} or d_{ij} is above some chosen threshold d_{cut}. All the particles/pseudojets that are left are then the events non-beam jets.

In the inclusive formulation, the difference is following: when a d_{iB} is the smaller number, then i is removed from the list of particles/pseudojets and instead of the beam jet it is added to the list of final inclusive jets. There is also no threshold d_{cut}, because the clustering continues until no particles/pseudojets remain. Another difference is that, there is usually some p_T cut applied to the final jets before their further usage.

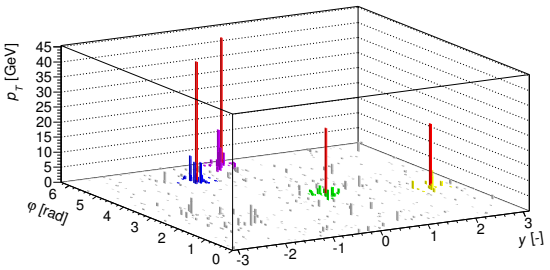
If $p = 1$ then the generalised k_t algorithm becomes the standard longitudinally invariant k_t algorithm, if $p = 0$ then it becomes the Cambridge/Aachen jet algorithm and if $p = -1$ then it becomes the anti-k_t algorithm that is at the time the most used one.

4.3 Example of anti-k_t vs k_t results

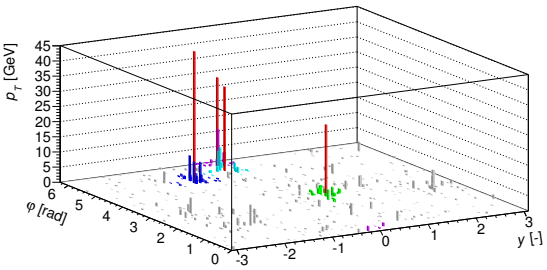
This section demonstrates the properties of jets obtained using Herwig 7.0.2 MC generator [12, 13] and reconstructed using the k_t and anti-k_t algorithms, as implemented in [11], and the influence of the jet radius R. One special event with sufficient hadron activity was selected for the visual comparison visible in Fig. 4.2. The generation was set as follows: proton-proton collisions at $\sqrt{s} = 13$ GeV with the matrix element calculating QCD scattering 2 to 2 at LO(leading order) together with LO parton



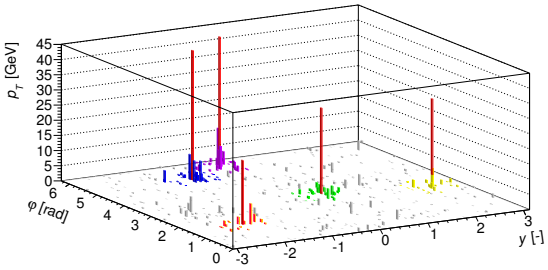
(a) Raw event



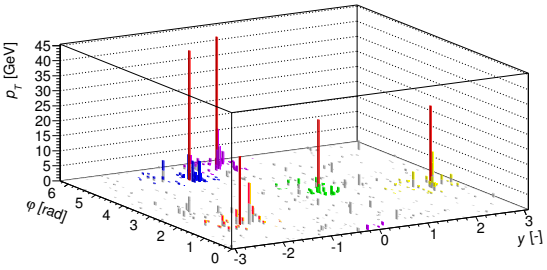
(b) anti-k_t with $R = 0.4$



(c) k_t with $R = 0.4$



(d) anti-k_t with $R = 0.6$



(e) k_t with $R = 0.6$

Fig. 4.2. Comparison of jet structure and number using k_t and anti-k_t algorithm with radii 0.4 and 0.6.

distribution function MMHT2014lo68cl loaded from LHAPDF library [13]. The event selection criteria were: minimum jet $p_T = 20$ GeV, at least four jets obtained using each algorithm and that the number of jets are different for varied jet radius R. In the horizontal plane of each histogram in Fig. 4.2 is drawn the azimuthal angle φ and their rapidity y. The height of each bar is proportional to p_T of particle, or to the sum of p_T of all the particles that fall into that relevant $y - \varphi$ bin. The raw event is shown in Fig. 4.2a,

where all final state particles are visible in grey. The other histograms are the very same ones, only with the colour differentiation of particles that belong to any reconstructed jet. The grey bars therefore represent background particles ignored by any subsequent jet analysis. The particles in different jets are highlighted in different colours. The tall, red bar in each jet is the representation of its axis and the height of the bar represents the p_T of the sum of 4-momenta of all particles in the jet. It is clear that the jet radius R is a key parameter in both algorithms. This importance is highlighted in the difference between Fig. 4.2b vs 4.2d for the anti-k_t and 4.2c vs 4.2e for the k_t algorithm, where the (anti-)k_t algorithm with $R = 0.6$ identifies one more jet, due to the wider radius that allows it to gather more particles into one jet and in doing so reach the p_T threshold of 20 GeV. One noteworthy occurrence is reconstruction of the aquamarine jet in Fig. 4.2c that is not there for the anti-k_t04 or to be more precise most of the particles are included in the violet jet and the rest is left out. Yet another difference in the k_t04 vs anti-k_t04 is that the k_t04 fails to reconstruct the yellow jet in Fig. 4.2b. The differences between the two algorithms fades in the $R = 0.6$ case because the thicker jets are less sensitive to soft particles in the event. In Fig. 4.2c and Fig. 4.2e can be seen that the k_t algorithm includes a few particles into the violet jet that are seemingly disconnected that is naturally only a misconception because the horizontal plane is the unfolded surface of a cylinder.

More extensive information on both cone algorithms as well as sequential recombination algorithms can be found for example in [9–11].

Chapter 5

Jet Substructure

Now that the concept of jet algorithms is introduced a brief foray into jet substructure and properties of jet tracks is in order. All the data throughout this chapter are generated by LO generator Herwig 7 for either minimum bias (MB) or two to two LO pQCD (QCD). From computational point of view minimum bias interaction means that the main process modeled is only colourless scattering with large rapidity. This of course does not mean that there can be no hard particles and jets found in a MB simulation because the influences from underlying event (UE) processes such as multiparton interaction (MPI), hadronisation etc need to be taken into account. Such model is entirely sufficient because from the experimental point of view (what concerns the settings of the detector) MB means that there is only the slightest cut-off on the acceptance of the detector, which means that even though the hard processes can happen the possibility of such an event actually happening is minimal and the dominant processes are all soft. The difference for QCD simulation is that the main process is no longer only a colourless scattering but fully perturbative colour scattering is under way. The minimum parton transverse momentum $\hat{p_T}$ is one of parameters characterising the QCD simulations. In case of simulations used in this work it was set on 15 GeV instead of the default 20 GeV in order to explore its influence on the shape of the p_T spectrum and find the smallest value possible that would still correctly describe the data. If the value of $\hat{p_T}$ was below 15 GeV the QCD started to significantly overvalue the measurement to a point where the cross section showed signs of divergence (which was expected). Although QCD does not produce sufficient number of jets in the p_T range below 15 GeV (due to the settings of the simulation) it does actually describes the jets just as well as MB. The reason for this is that in both cases jets/tracks are generated from the same (UE) modeling and the only difference is in the cross section scaling as is demonstrated below.

The standard rivet analysis ATLAS 2011 I919017 [14] is chosen as a reference and guiding for educational reasons. The aim of this chapter is to study jets into a greater depth, acquire some insight into their inner structure and observe the dynamic consequences of higher energy on the shape of some distributions and to establish a forecast for the $\sqrt{s} = 13\,\text{TeV}$ measurements.

5.1 Review of the reference analysis

The analysis [14] includes measurements of jets and their properties at center of mass energy $\sqrt{s} = 7\,\text{TeV}$ in proton-proton collisions at the LHC using charged particles measured by the ATLAS detector. In order to include jets at very low transverse momenta and to study their characteristics in the transition to high-momentum fully perturbative jets, the events are selected using the MB trigger. There are five quantities related to jet track properties that are measured and those are: the inclusive charged jet transverse momentum cross section in the range from 4 to 100 GeV, the transverse and longitudinal

momentum fractions of charged particles in jets, charged particle multiplicity per jet and charged particle density as a function of radial distance from the jet axis. The measurements are divided into four rapidity ranges: $(0, 0.5)$, $(0.5, 1.0)$, $(1.0, 1.5)$, $(1.5, 1.9)$ and (apart from measurement of the p_T spectrum itself) five p_T ranges : (4, 6) GeV, (6, 10) GeV, (10, 15) GeV, (15, 24) GeV, (24, 40) GeV. All results of the analysis are compared for jets reconstructed using the anti-k_t algorithm with the radius parameter R taken as 0.4 and 0.6. The article [14] includes an extensive discussion of systematic uncertainties and comparison of MC simulations from various tunes of MC generators Herwig and Pythia. It is important to point out that Herwig++ 2.4.2 and 2.5.1 is used in the original article [14] but Herwig 7.0.2 is used in this analysis. However the standard Rivet analysis remains the same for 7 TeV and with a few necessary updates in the analysis for 13 GeV.

5.1.1 Data description

The data, with which the 7 TeV simulations are compared, are a sample of early ATLAS data at centre-of-mass energy $\sqrt{s} = 7\,\text{TeV}$ for which the MB trigger was used. The requirements for track reconstructions are following [14]:

- transverse momentum $p_T > 300$ MeV
- pseudorapidity $|\eta| < 2.5$
- transverse impact parameter with respect to the primary vertex $|d_0| < 1.5$ mm for tracks with $p_T < 10\,\text{GeV}$.
- transverse impact parameter with respect to the primary vertex $|d_0| < 0.2$ mm for tracks with $p_T > 10\,\text{GeV}$.
- longitudinal impact parameter with respect to the primary vertex z_0 satisfying $|z_0 \sin\theta| < 1.5$ mm
- if a hit is expected in the innermost pixel detector layer then such a hit is required wit one pixel hit in any layer required otherwise
- at least 6 hits in silicon microstrip detector.

5.2 Reconstruction of chosen properties

The main idea is to take some chosen properties in specific rapidity and p_T ranges from [14] analysis, try to reconstruct them using ROOT histograms and apply the reconstructed analysis to simulations from Herwig 7 with $\sqrt{s} = 13\,\text{TeV}$. Due to the kinematic range being on the borders between hard and soft QCD, both cases are included below (as demonstrated on the p_T spectrum of inclusive jets in Fig. 5.1). The reference analysis is quite extensive, because its goal was to tune parameters of MC generators and so it includes a number of ranges that is unnecessarily big for this thesis that has only the intention to show a few examples of jet properties that can be measured. Because of this, in the rest of this chapter, jets are reconstructed for simplicity using only the anti-k_t algorithm with radius parameter $R = 0.6$. The study is divided into two rapidity ranges: $|y| < 0.5$ and $1.5 < |y| < 1.9$ for the double differential cross section $\frac{d^2\sigma_{jet}}{dp_T dy_{jet}}$ with respect to p_T and y of the jet, as shown in the inclusive jet p_T spectrum Fig. 5.1. And only one rapidity range $|y| < 0.5$ along with transverse momentum range of $p_T \in [4, 6]\,\text{GeV}$ for the plots of charged particle multiplicity per jet $\frac{1}{N_{jet}}\frac{dN_{jet}}{dN^{ch}_{jet}}$ (Fig. 5.2 and 5.3) and charged particle density $\rho_{ch}(r)$ (Fig. 5.4), where r is the distance of charged particle track from the jet axis in the y-ϕ plane defined as:

$$r = \sqrt{(\phi_{ch} - \phi_{jet})^2 + (y_{ch} - y_{jet})^2}, \tag{5.1}$$

for index t denoting the track in question.

5.2.1 Cross section as a function of p_T

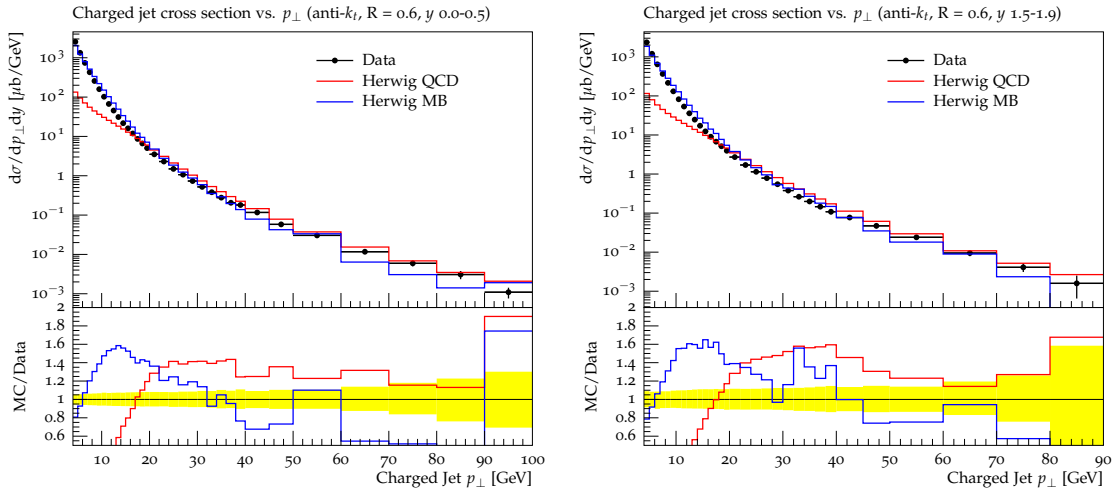


Fig. 5.1. The comparison of p_T spectra from minimum bias and two to two QCD generated by Herwig 7 for $\sqrt{s} = 7$ TeV in two rapidity ranges: $|y| < 0.5$ (left) and $1.5 < |y| < 1.9$ (right). Plots are obtained using standard Rivet analysis from [14].

The double differential cross section for inclusive charged jets in respect to transverse momentum and rapidity is one of the basic properties studied in any analysis. The results are shown in Fig. 5.1.

The red lines represent QCD simulation. The cross section in the range up to 15 GeV is the result of the UE effects such as multiparton interaction during the generation. In the region of the validity ($p_T > 15$ GeV) of the pQCD approach the simulated data generally exceeds measured datapoints and the ratio of data/MC fluctuates ±0.2 about the value of 1.4, although the right plot ($1.5 < |y| < 1.9$) is showing strange excess over data in MB in the region of 30-40 GeV. From this can be assumed that if the next to leading order correction were to be calculated its sign should be negative.

The blue lines represent the MB simulation. The plots in Fig. 5.1 show that the MB simulations are in good agreement with the data for jets with p_T up until about 40 GeV and then they start to more significantly undervalue the measured data, although until 60 GeV this can still be included within thrice the uncertainty of measured data. What is interesting however is that in both rapidity ranges the MB has a peak in the ratio plot for $p_T \approx 12$ GeV. It would be interesting to obtain a combined spectrum from both MB and QCD in order to see if it would still undervalue the measurement or if it would be in better agreement.

5.2.2 Charged particle multiplicity per jet

The acquisition of the charged particle multiplicity per jet goes: First the jets from charged final state particles are reconstructed using the anti-k_t06 algorithm mentioned above. Once the jets are reconstructed and the chosen cuts, mentioned in the beginning of this chapter, are imposed then a loop over all jets that clear the cuts fills the histogram with the number of particles in each jet. In the end a normalization of

the histogram is needed and so it is scaled to the total number of jets in the rapidity and p_T range.

The general shape of charged jet multiplicity in Fig. 5.2 tells us that the highest multiplicity is for the jets with four charged particles and from the parabolic shape of the data in logarithmic scale can be approximated that the distribution is gaussian. It can be seen in Fig. 5.2 that both QCD and MB tend to be shifted towards greater multiplicities. They undervalue the multiplicity for number of charged particles fewer than two with significant disproportions for one-particle jets. On the other hand both simulations overshoots the data from six for QCD and seven for MB on. It can be said that the MB simulations are more accurate because they agree with the data for $N = 3 \ldots 6$, whereas QCD is good enough only for $N = 4$, 5 and then quickly looses its accuracy. This can be at least partially explained as an effect of the showering of low-p_T partons, which might contribute to formation of UE jets. The reason can be that the perturbative calculations in QCD are far more precise in higher p_T regions and so the range $p_T \in [4, 6]$ GeV is far in the region of inaccuracy of these simulations[1]. This disparity in accuracy between QCD and MB simulations are only logical, for MB has been developed to better suit the low-energy regions where softer jets comprised of fewer particles are found.

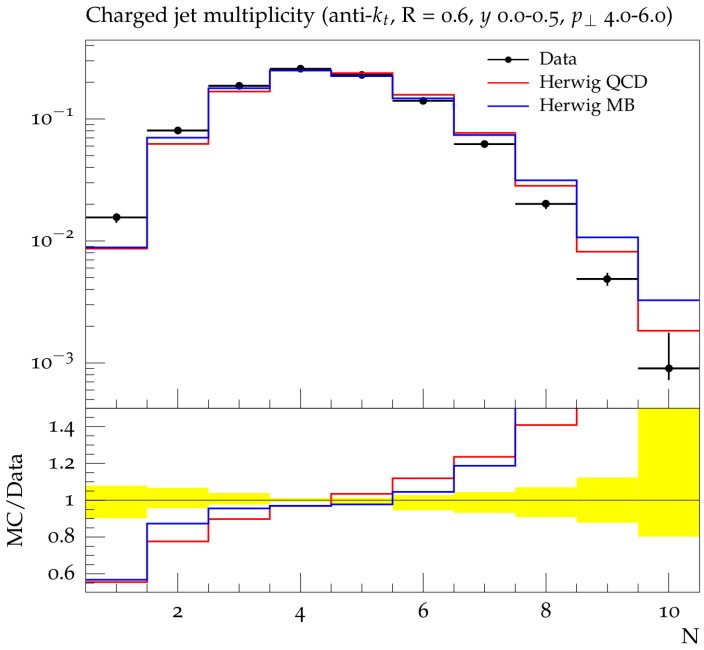


Fig. 5.2. Comparison of charged particle multiplicity per jet as generated using Herwig 7.0.2 MB, LO QCD at $\sqrt{s} = 7$ TeV and data. Obtained using standard Rivet analysis [14].

The same process is used in creation of Fig. 5.3 as was in Fig. 5.2 only the centre of mass energy $\sqrt{s}$ has been changed to 13 TeV in both simulations and the notation $N \to N^{ch}_{jet}$. Because at the time of writing of this thesis no data for charged particle multiplicity per jet at the energy of $\sqrt{s} = 13$ TeV have been available to the author only comparison between MB and QCD simulation is showed in the ratio plot instead of the ratio plot between monte carlo simulation and data in Fig. 5.2. The difference in the UE tracks from QCD and MB is growing with energy, which is an interesting fact that has no clear explanation as of yet. It seems that a particle in UE where at least one hard process has taken place (QCD) has tendency to produce fewer additional particles due to the conservation laws. It is interesting

[1] See Fig. 5.1 left, where the p_T spectrum is showing inaccuracies for $p_T \lesssim 15$ GeV.

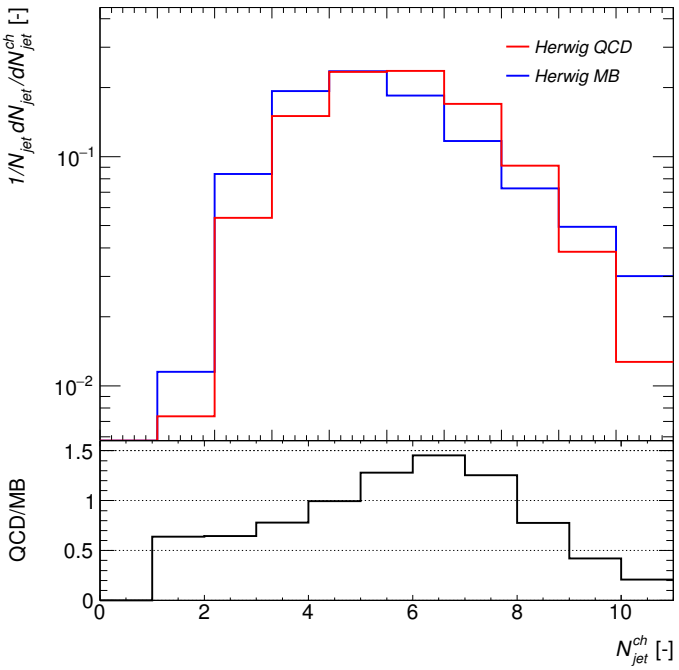


Fig. 5.3. Comparison of charged jet multiplicity as generated using Herwig 7.0.2 MB and LO QCD at $\sqrt{s} = 13$ TeV.

that both MB and QCD are showing no one particle jets even in this low p_T range. Similarly to the case in Fig. 5.2 QCD and MB agree with each other only in one bin and that is for $N^{ch}_{jet} = 5$.

5.2.3 Number density of charged particle tracks

The number density of charged particle tracks ρ_{ch} is the property that describes how the charged particles/tracks are generally distributed in an average jet. The cuts imposed on data during computation of ρ_{ch} are the same as those for charged particle multiplicity per jet. The drawing of the distribution is however not as straightforward as before. First step is to take all the jets and select those that satisfy all the cuts and for each such jet take all its comprising particles. Then the distance r of each particle from the jet axis[2]is computed from (5.1) and added into the distribution with the weight:

$$w_\rho(r) = \frac{u - d}{\pi(u^2 - d^2)}, \tag{5.2}$$

where u and d are the values of upper and lower edges of bin containing r. The important weighting function w_ρ for ρ_{ch} corresponds to standard Rivet analysis [14]. The histogram is in the end normalized to the total number of jets in taken rapidity and p_T range.

The general shape of the charged particle density in Fig. 5.4 is telling us that the greatest concentration of charged particles is immediately around the axis of the jet and then, after a prominent step, it exponentially diminishes until just before the edge of the jet another prominent step occurs signalizing that there are very few particles found around the outer perimeter of jet. Both QCD and MB in Fig. 5.4 is in far better agreement with data than was the case for the charged particle multiplicity per

[2]The axis of the jet is in this case taken as the sum of four-momenta of all particles comprising the jet.

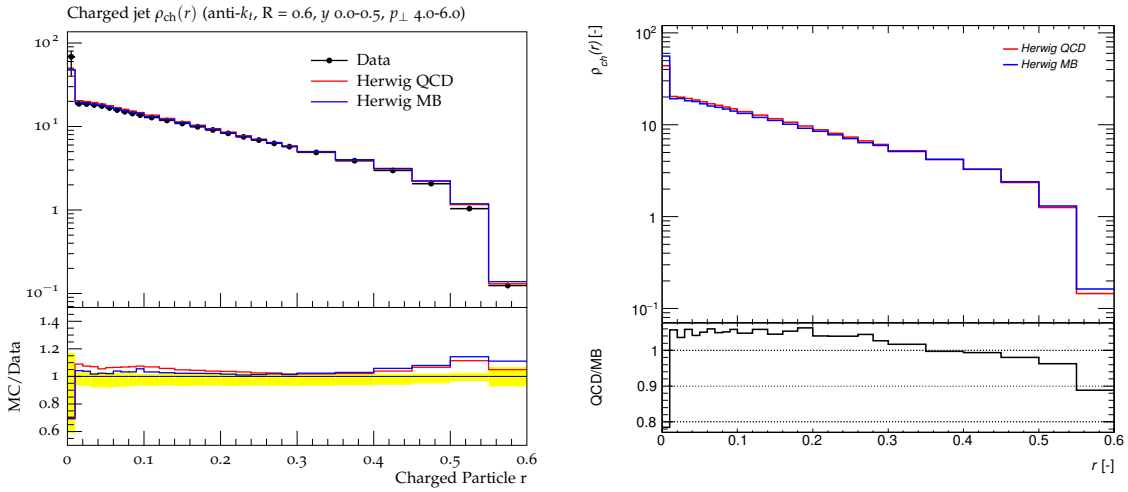


Fig. 5.4. Comparison of charged particle density as generated using Herwig 7.0.2 MB and LO QCD at $\sqrt{s} = 13$ TeV.

jet. Nevertheless the MB simulations (blue line) are still generally better approximation especially for $r \in [0, 0.4]$. In the last few bins ($r \in [0.4, 0.6]$) the disagreement of MB with data rises and the density is better approximated by the QCD simulations. This disagreement is however only marginal because both simulations are in the entire range within thrice the value of uncertainty of the measured data.

Similarly to the previous section, the course of action is to reconstruct the standard analysis [14] for the energy of $\sqrt{s} = 13$ TeV, which yields Fig. 5.4. The general shape of ρ_{ch} remains this time unchanged but still there is no comparison to data from RunII and so this remains unconfirmed. The difference between the simulations is again mostly marginal and bigger disparity is only in the values of the steps in the first and the last bin.

Chapter 6

Data Analysis

The content of this last chapter is the discussion of jet analysis created as the main outcome of this thesis. The motivation for this analysis is to study how well the Powheg + Pythia generated simulations describe the data from p-p collision at $\sqrt{s} = 8$ TeV measured by the ATLAS detector at LHC. The simulations which are used come from the official ATLAS dataset, where Powheg computes the parton level events and Pythia completes them into the hadron level (showers and hadronizes partons). The versions of Powheg and Pythia 8 used are taken for atlas release 17.2.7. with the PDF CT10 and UE tuned to AUCT10 and the detector response is simulated using Geant4 [15]. The goal is to at first study the events where any jets in general are measured in order to gain knowledge of the shape of their inclusive p_T spectrum and multiplicity and in doing so check the soundness of this analysis. The latter part of the analysis is focused only on those special types of events where exactly two jets are recorded. This type of events is called dijet event. The whole list of quantities measured in the dijet part of this analysis is as follows:

- leading and subleading jet p_T spectra
- dijet mass,
- dijet transverse momentum imbalance Δ_{ij},
- normalized dijet momentum imbalance Δ_{ij}^{n},
- half of the rapidity separation $y^* = \frac{1}{2}|y_1 - y_2|$ of dijets,
- boost of the dijet system $y_b = \frac{1}{2}|y_1 + y_2|$,
- 2D dijet event topology using y^* and y_b.

The dijet events and their topology are studied in order to decipher the kinematical differences between events where only two jets are spawned and events where a vector boson with the associated production of a dijet pair is created. One such study can be found in [16].

6.1 Event selection and weight

The jets are reconstructed using the Anti-k_t, locally calibrated jet algorithm with the radius parameter $R = 0.4$. However not every jet can be accepted into the final event that will be analysed, which leads to a need for event selection and cleaning criteria. One of the problems that plagues the jet measurements in modern physics is pileup. With the high luminosities and numbers of protons per bunch that LHC

has there are more than one proton collisions per bunch crossing and so the need arises to filter out the uninteresting collisions and focus only on the most interesting one. The place of the most energetic collision is called the primary vertex. It is however not the only vertex (place of a collision) in an event. Jets that stem from other vertices (from the "uninteresting" collisions) are called pileup jets and need to be dealt with accordingly. This leads to event selection criteria.

The first kind of event selection is based on the property of pileup, that virtually no pileup jets are found with p_T above 50 GeV and so through imposition of the general cuts:

$$p_T > 50\,\text{GeV} \qquad |\eta| < 4.8, \tag{6.1}$$

onto every jet, the pileup is effectively removed. The rapidity range is chosen to encompass the whole of ATLAS detector. The second kind of event selection is chosen to allow the study of jets with as low p_T as possible. Such jets are then subjected to general cuts:

$$p_T > 25\,\text{GeV} \qquad |\eta| < 2.1. \tag{6.2}$$

The rapidity range in this case is chosen so that every particle in each jet (the radius of the jet is 0.4) is inside the acceptance of the ATLAS inner detector ($|\eta| < 2.5$). For this selection type the mechanism that deals with the pileup jets is the so called jet vertex fraction cleaning. The jet vertex fraction (JVF) is the scalar sum of the p_T of the tracks in the jet area coming from the primary vertex, divided by the scalar p_T sum of all tracks in the jet area. Jet is rejected if its JVF is below 0.5 for jets with p_T below 50 GeV and it effectively reduces jets from pile-up. Both types of general selection are then followed up by additional cleaning of bad events that can be caused by detector defects and inefficiencies in certain runs. For this cleaning the standard ATLAS procedure [**?**] has been followed (in this case the recommendations for processed data has been used).

For data, the trigger and the p_T of the hardest jet is determining the weight of the event in dependence on the pre-scaling of the trigger. Pre-scaling of a trigger can be though of as the luminosity taken by the trigger. If no such weight can be accorded, the event is rejected. Additional event cleaning comes from information about the state of operationability of the liquid argon and tile calorimeters from the Standard Good Run List for 2012 data. This operationability can be affected by excessive noise-bursts and readout errors in certain lumiblocks and runs. The summary of used triggers and relative luminosities taken by each trigger can be found in Tab. 6.1.

For MC the task of cleansing transgressing events and assigning event weights is similar only there is additional condition on consistency with truth jets (the jets that are generated directly and have no applied detector distortion): the average p_T of two leading jets divided by the p_T of leading truth jet must be between 0.8 and 1.2. This consistency checking is present in order to hinder the detector simulation from generating non-existing jets. MC simulations are divided into eight samples along the outgoing parton p_T in order to produce sufficiently large data samples even for high p_T jet production. The event weight w for MC is determined according to the channel number and the assigned cross section σ_{JZX} and filter efficiency F off the appropriate JZX (X = 0, 1...7) sample:

$$\sigma = \sigma_{JZX} \times F \Rightarrow w = \frac{1}{\sigma}, \tag{6.3}$$

where σ denotes the actual cross section of an event in the sample. MC has the corrections of weight on the effect of pileup, but these do not affect the event selection. The summary of JZX channel numbers, number of events in sample, cross section and filter efficiency can be found in Tab. 6.2

Jet p_T [GeV]	Fired trigger	Luminosity weight
25 GeV < p_T < 50 GeV	EF_j15_a4tchad	8/0.014734
50 GeV < p_T < 100 GeV	EF_j25_a4tchad	8/0.078440
100 GeV < p_T < 136 GeV	EF_j55_a4tchad	1/0.441786
136 GeV < p_T < 190 GeV	EF_j80_a4tchad	1/2.3164
190 GeV < p_T < 204 GeV	EF_j110_a4tchad	1/9.81141
204 GeV < p_T < 252 GeV	EF_j145_a4tchad	1/36.2647
252 GeV < p_T < 304 GeV	EF_j180_a4tchad	1/78.7753
304 GeV < p_T < 404 GeV	EF_j220_a4tchad	1/261.379
404 GeV < p_T < 522 GeV	EF_j280_a4tchad	1/1164.69
522 GeV < p_T	EF_j360_a4tchad	1/20277.1

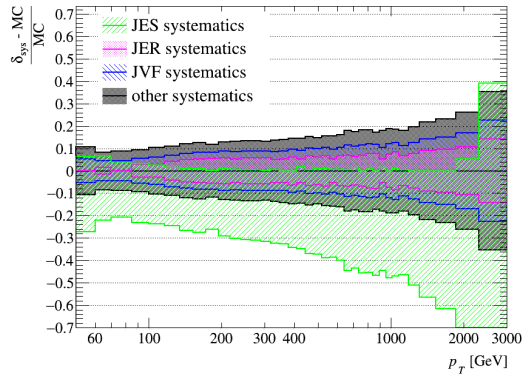
Tab. 6.1. The summary of triggers and accepted p_T of the leading jet in order to reach 99.5% efficiency threshold of each trigger.

Sample name	Channel #	# of good	$\sigma_{\mathbf{JZX}}$ [pb]	$\mathbb{F}$
PJZ0	184150	1499998	3.4366×10^{10}	9.7917×10^{-2}
PJZ1	184151	7483277	3.4366×10^{10}	2.0776×10^{-2}
PJZ2	184152	1499992	8.7759×10^{7}	2.2481×10^{-2}
PJZ3	184153	1499392	1.6440×10^{6}	1.5487×10^{-2}
PJZ4	184154	1499480	7.5748×10^{3}	2.0292×10^{-2}
PJZ5	184156	1494466	4.8925×10^{1}	2.2823×10^{-2}
PJZ6	184155	1487136	2.8028	8.1361×10^{-3}
PJZ7	184157	1486932	2.6161×10^{-2}	1.9762×10^{-2}

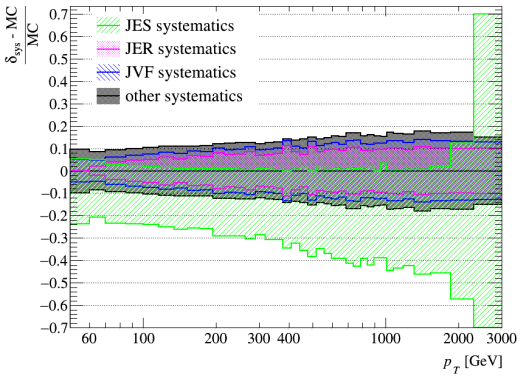
Tab. 6.2. The summary of MC sample names, ID numbers, number of good events in sample, cross section $\sigma_{\mathbf{JZX}}$ of the sample and its filter efficiency $\mathbb{F}$.

6.2 Systematic uncertainties

During the computation of systematic uncertainties, the standard ATLAS procedures have been followed. Systematic uncertainties are computed from the MC generations and then are applied to the data, because there is abundance of MC simulations over data, which will provide better results. The listed systematic uncertainties (systematics) are taken as uncorrelated so they can be summed as squares. These systematics are called the nuisance parameters of the analysis. Six of these are the in-situ calibration nuisance parameters. Two are the intercalibration parameters for pseudorapidity (one for modelling one for statistics). One parameter deals with the behaviour of high p_T jets caused by propagation of single hadron uncertainties. Four uncertainty sources are from the nuisance parameters from pile-up. Another source of uncertainty is from the close by jets nuisance parameter. Two other uncertainties are from flavour response and composition which are discussed in detail in [17]. All uncertainties above are grouped (summed as squares) into the jet energy scale (JES) systematic. Another systematic stems from the JVF of jets and yet another from the jet energy resolution (JER). The rest of systematics, listed as other are systematics caused by detector defects. More information about the computation and definition of in-situ calibration parameters can be found in [18] The graphical representations of relative uncertainties of respective systematics for all quantities can be found in the appendix A, Fig. A.4a to A.8

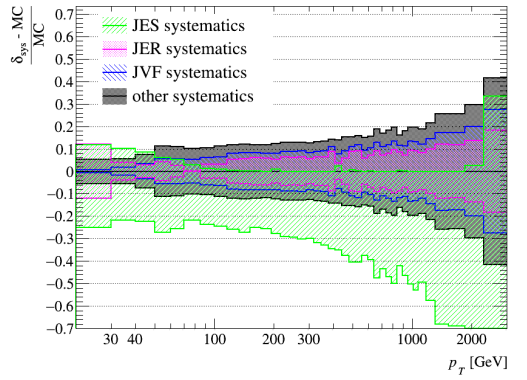


(a) Leading jet

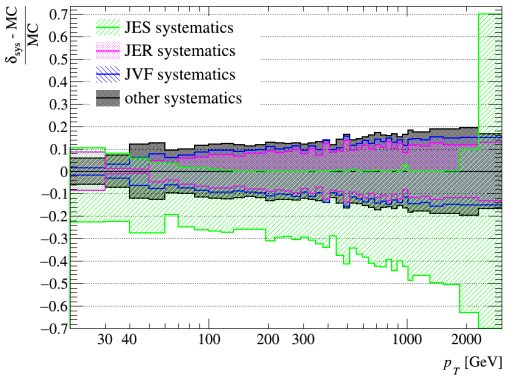


(b) Subleading jet

Fig. 6.1. Systematic uncertainties of the leading and subleading jet transverse momentum spectrum for the first selection type.



(a) Leading jet



(b) Subleading jet

Fig. 6.2. Systematic uncertainties of the leading and subleading jet transverse momentum spectrum for the second selection type.

except for the leading and subleading jet systematics that are shown here as an example. These results are however only preliminary due some unexpected behaviour of the JES systematics which has mostly negative tendency instead of being symmetric.

6.3 Measured quantities

6.3.1 Inclusive jets p_T spectrum and multiplicity

Two of the most basic properties that are used to test the applicability of MC generators are the inclusive jet p_T spectrum and multiplicity. Inclusive, in this case, means that the number of jets in an event does not matter and all jets that came unscathed from the event selections and additional cleaning are accounted for. In all following figures the notation is this: Black markers and black error-bars denote data with total uncertainty and the surrounding yellow area denotes the statistical uncertainty of measured

data. MC simulation from Powheg + Pythia are drawn as red lines. If a ratio plot is included, the red line is the MC simulation value divided by measured value and the grey area is representing the total uncertainty of measurement. By the p_T spectrum is meant the double differential cross section $\frac{d\sigma}{dp_T dy}$ with respect to p_T and y of inclusive jets. Due to the abundance of measured data, the spectrum is divided into eight rapidity ranges for the first type of event selection type (6.1) and this division is drawn in Fig. 6.3. The specra in Fig. 6.3 have been scaled in order to better differentiate among respective rapidity ranges. The values drawn are consistent with the expectation, that jets with the highest p_T are found with the lowest rapidity and the highest rapidity have only the least energetic jets. Therefore we can find no particles that have rapidity higher than 3.5 and at the same time p_T above 200 GeV and conversely jets harder than 2 TeV have exclusively the rapidity below 1.5.

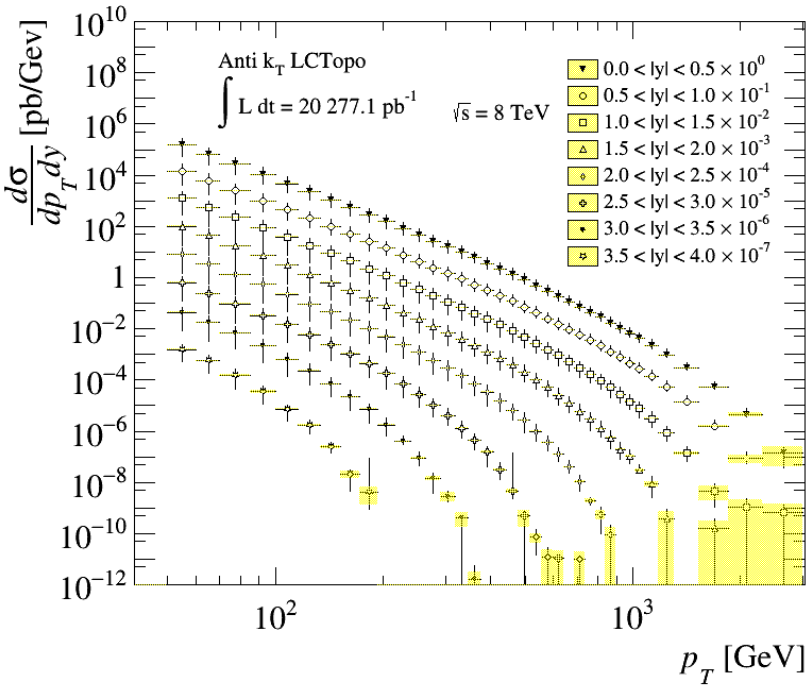


Fig. 6.3. Transverse momentum spectrum of inclusive jets for the first selection type in multiple rapidity ranges.

The inclusive p_T spectrum for the second selection type (6.2) is drawn in Fig 6.4 without any subdivision of rapidity acceptance. For this selection the spectrum is smooth without any noteworthy peaks. The highest probability of production is for jets that are just above the p_T threshold. Powheg data (red line) generally overvalue the data with exception of a few high pT bins. The general agreement is mostly within 10% except the peak at 40 GeV where the difference is 40%. Nevertheless, the agreement still seems to be within the bounds of systematic uncertainty.

The jet multiplicity expresses the total number of jets produced during an event. This quantity is included in order to bring a better understanding of what portion of events is studied during a dijet analysis. The inclusive jet multiplicity distribution for event selection type one and two can be found in Fig. 6.5 and 6.6 respectively. In both cases the highest multiplicity is for events with no jets whatsoever and then the value decreases roughly with the powers of 10. For the event selection type the step between 0 and 1 multiplicity is more pronounced due to the higher p_T threshold that a group of particles need to have in order to be taken as a valid jet.

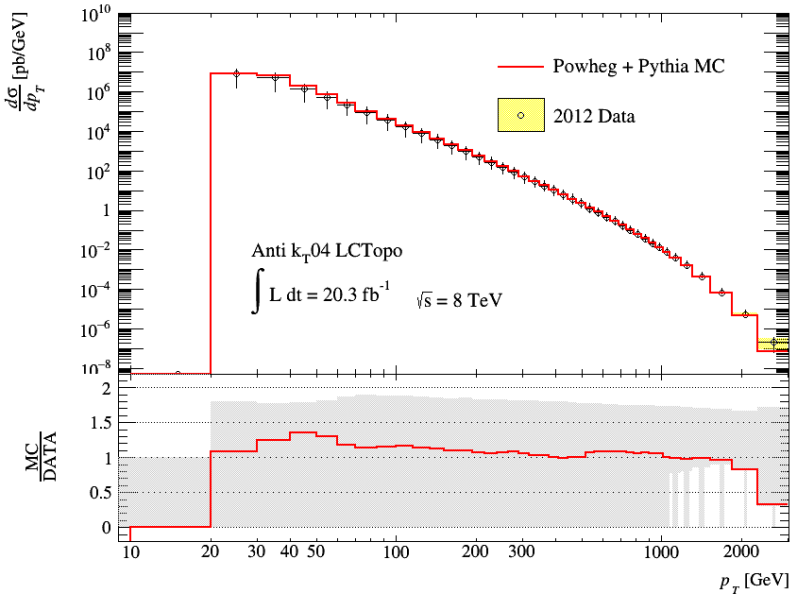


Fig. 6.4. Transverse momentum spectrum of inclusive jets for the second selection type in the rapidity range (0.0, 2.1).

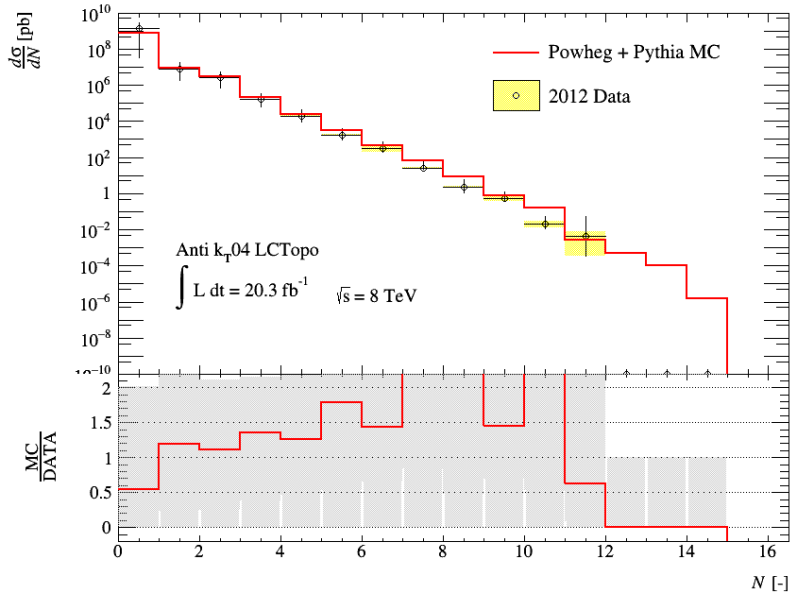


Fig. 6.5. Jet multiplicity for the first selection type, in the rapidity range (0.0, 4.8).

6.3.2 Properties of dijets

Leading and subleading jet p_T spectra

By leading and subleading jet is generally meant the jet in with the first and second highest transverse momentum. Which of course in dijet event effectively means the harder and the softer jet. The leading

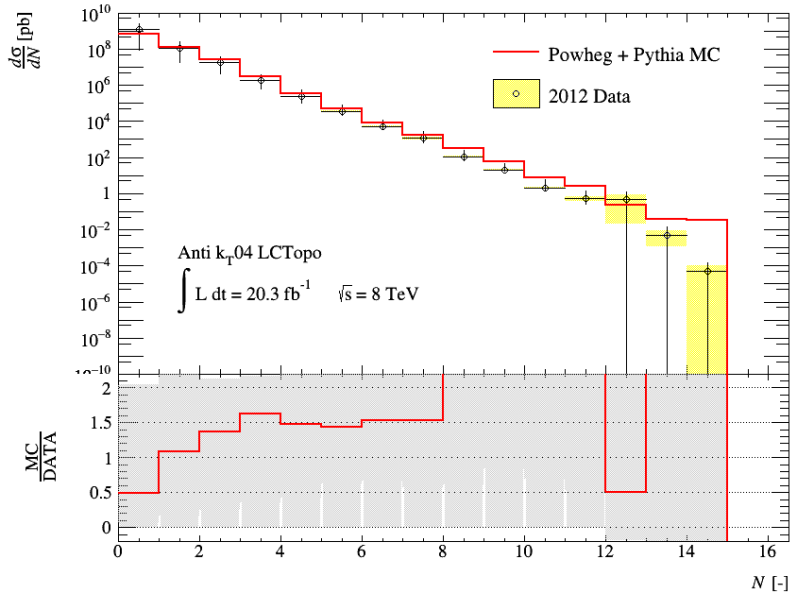
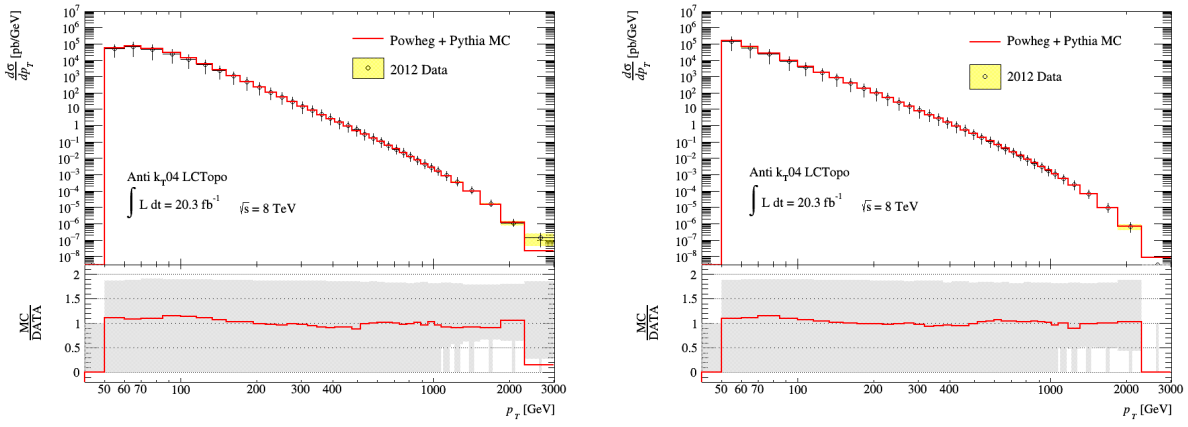


Fig. 6.6. Jet multiplicity for the second selection type, in the rapidity range (0.0, 2.1).

and subleading jet p_T spectra can be found in Fig. 6.7 for the first selection type. The leading jet spectrum is from $p_T \approx 70$ GeV always above the subleading jet spectrum, which is to be expected because the probability of producing less energetic jets is higher. Below 70 GeV it is then expected that the number of leading jets falls, because around the p_T threshold of 50 GeV there is the real danger that the would be subleading jet would not be able to pass the minimum p_T requirement. In both leading and subleading jet spectrum the data seem to be in good agreement with the MC simulations.



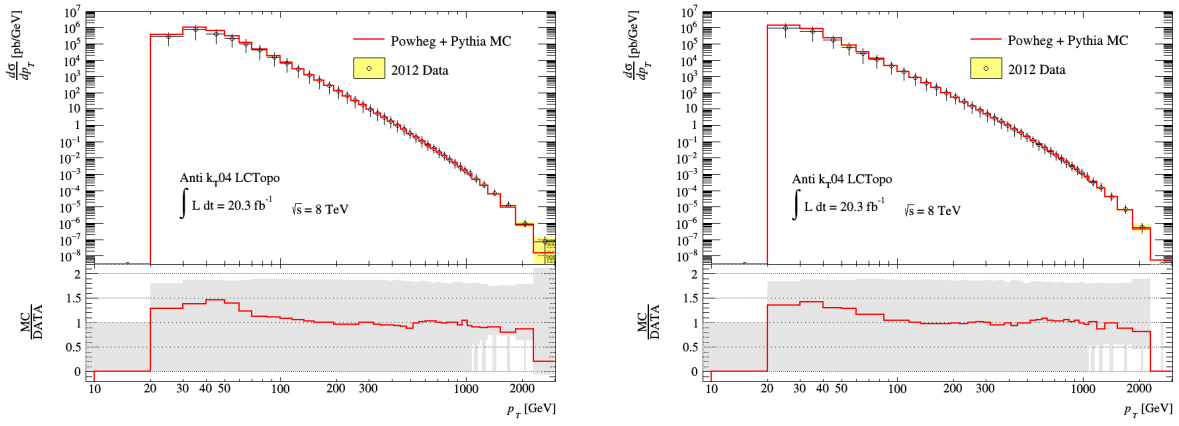
(a) Leading jet.

(b) Subleading jet.

Fig. 6.7. The transverse momentum spectrum of dijet events for the first selection type in the rapidity range (0.0, 4.8).

The behaviour of the leading and subleading jet spectra remains essentially unchanged in the second event selection type as can be observed in Fig. 6.8. The only notable difference is that the peak of the

leading jet spectrum is now at 30 to 40 GeV instead of 70 to 80 GeV and the whole spectrum naturally begins earlier due to a lower p_T threshold.



(a) Leading jet. **(b)** Subleading jet.

Fig. 6.8. The transverse momentum spectrum of dijet events for the second selection type in the rapidity range (0.0, 2.1).

Dijet mass

Mass is one of the basic properties of any particle, it is only logical then to study such a concept for systems of jet as well. For the dijets specifically, the two jets are combined to acquire the invariant dijet mass M_{JJ}. The dijet mass spectra are generally used during the search for new phenomena such as new quark resonances or quantum black holes. For the dataset used here one such a thorough analysis can be found in [19]. The mass spectrum for the first selection type can be found in Fig. 6.9. In this case, the maximum of the dijet mass spectrum is between 110 and 160 GeV. As can be observed the discrepancy between MC and data rises from about 2 TeV, although the statistical uncertainty of the data rises as well.

For the second event selection type is the mass spectrum similar, only the peak of the spectrum is shifted into the range between 70 to 110 GeV. which can be attributed to the dijet production associated with the creation of a vector as discussed in the beginning of this chapter.

Dijet transverse momentum imbalance

The law of conservation of four-momenta leads on parton level to partons moving from each other, that however does not mean that there can be no boost to the system due to many effects like UE, which means effects like MPI and soft QCD radiation, or pileup activity, which is on some level filtered out, but first on the jet level and only for jets with p_T above 25 or 50 GeV. Therefore this cuts only most of the pileup activity but not all of it. The conservation law of four momentum further dictates, that the momenta of each jet needs to cancel out in the transverse plain which lead to so called jets back to back in p_T. In the past some analyses utilized the variable $\Delta\phi$, in this analysis the chosen variable is the dijet transverse momentum imbalance Δ_{JJ} that is defined as the transverse part of the vector sum of the two jet four-momenta:

$$\Delta_{JJ} = (\vec{pJ1} + \vec{pJ2})_T. \tag{6.4}$$

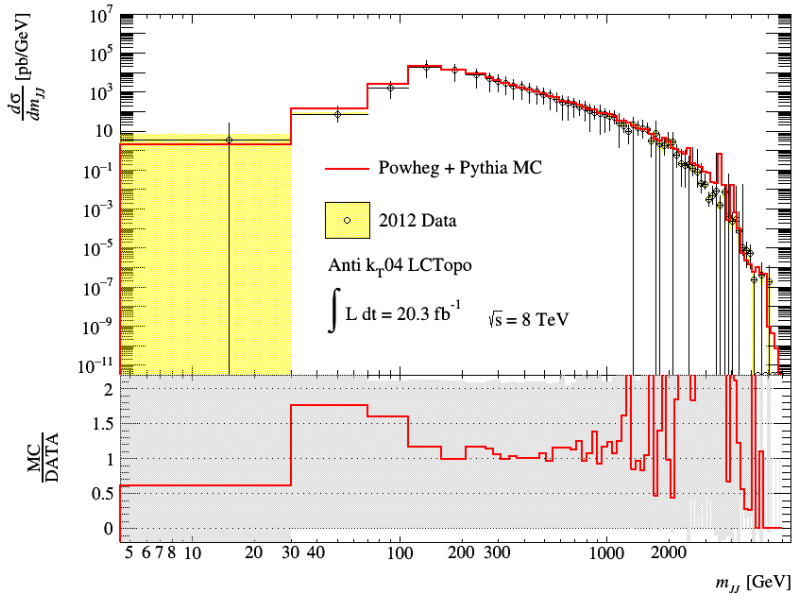


Fig. 6.9. Dijet mass spectrum for the first selection type, in the rapidity range (0.0, 4.8).

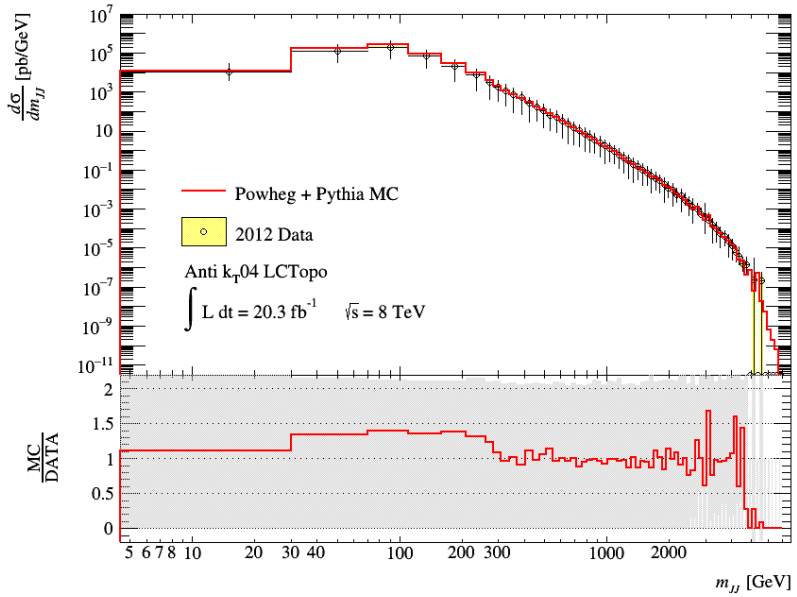
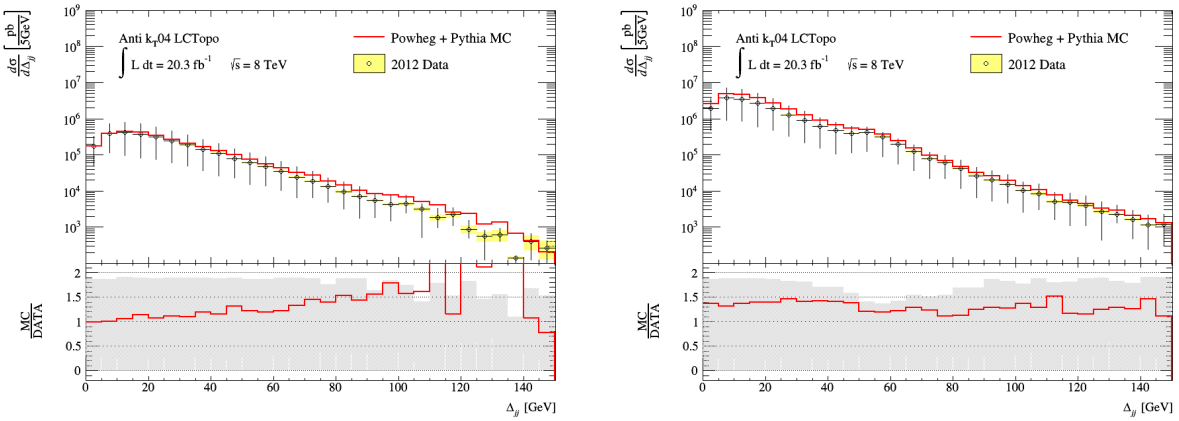


Fig. 6.10. Dijet mass spectrum for the second selection type, in the rapidity range (0.0, 2.1).

Along with its normalised variant Δ^n_{JJ}, which does not depend on the energy scale of the process, defined in equation (6.5) are first vairables that can shed some light onto the shape of the dijet event. Both quantities describe the measure of asymmetry of the dijet sytem in the p_T phase-space. If the jet momenta are distributed in such a manner that their transverse momenta cancel each other out or in other words if the jets are fully back to back in *phi*, then both observables go to zero.

$$\Delta^n_{JJ} = \frac{|\vec{p_{J1}} + \vec{p_{J2}}|_T}{|p_{T,J1}| + |p_{T,J1}|} \tag{6.5}$$

The transverse momentum imbalance distribution for the first event selection type can be found in Fig. 6.11a. The distribution rises up to ~ 10 GeV where is its maximum and then it begins to steadily decrease, with a hint of peak between 90 and 110 GeV, which is probably an effect of event selection cut on p_T as is explained below on the better example in the second event selection type. For which the Δ_{JJ} distribution in Fig. 6.11b is much higher than in the first case, especially in the region up to 90 GeV. This is caused by the lower p_T threshold. The shape of the distribution is however quite interesting. There is the starting peak again at about 10 GeV but then there is a significant peak about 50 GeV. This can be explained as a result of the chosen cut in that as Δ_{JJ} goes from zero the jets are closer to each other in the p_T space and at 50 GeV they are facing the same direction and in order to be accepted they have to have at least 25 GeV each. The fact, that in both event selection type the maximum is approximately at 10 GeV means that most of dijets are not fully back to back in p_T but are somehow boosted.



(a) First selection type.

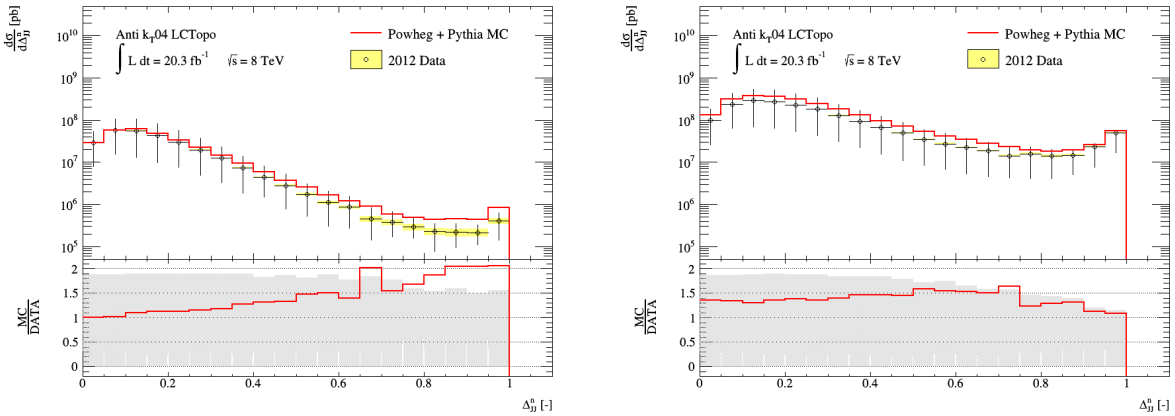
(b) Second selection type.

Fig. 6.11. The comparison between dijet momentum imbalance Δ_{JJ} for both selection types.

The normalized transverse momentum imbalance Δ^n_{JJ} provides another measure that describes the approximate shape of the event. In Fig. 6.12a for the first selection type the most of the dijets are towards 0, which means their momenta are placed in such a way that their transverse parts cancel each other out and so the jets are almost back to back, which is consistent with the expectation due to the conservation laws. In Fig. 6.12b The shape of the Δ^n_{JJ} distribution is again more interesting. The back to back peak at around 0.1 remains, however another peak can be seen towards 0.9-1. This can be caused by the lowered acceptance of this selection type which may result in some jets "escaping" either from our acceptance or even the acceptance of the detector.

Dijet event topology

In order to acquire even better understanding of the dijet event topology the two variables dijet system boost and the half of dijet rapidity separation are used. These can be interpreted as the average rapidity of the dijet system and the distance that separates the two jets in the $\phi - y$ plane. The idea is to use them in order to create a 2D plot that has the y_b on the horizontal and the y^* on the vertical axis. The number of events with appropriate settings is colour coded. Such a 2D plot for the first configuration can be



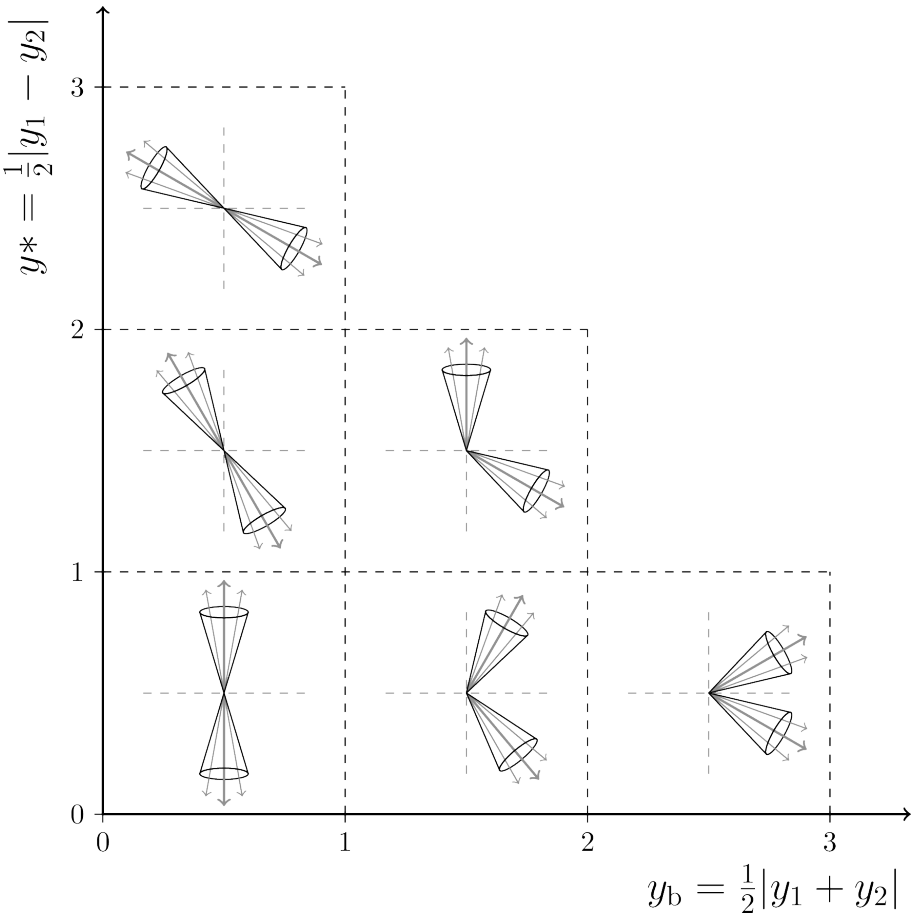
(a) First selection type.

(b) Second selection type.

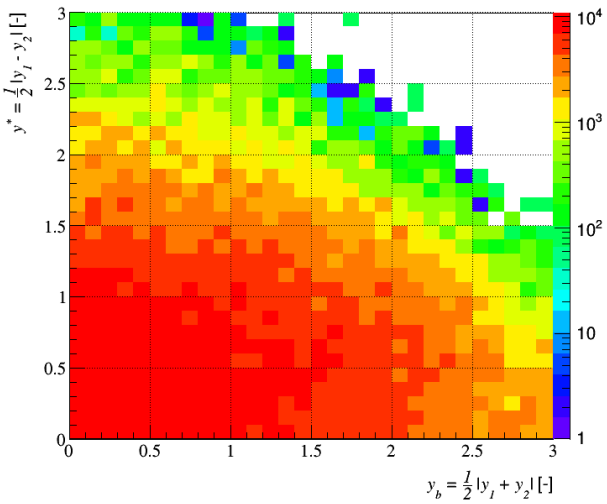
Fig. 6.12. The comparison between dijet momentum imbalance Δ_{JJ} for both selection types.

found in Fig. 6.13b along with the legend that deciphers the meaning of this plot in Fig. 6.13a. It is plain that the bulk of the data is concentrated in the region $(0,\ 1) \times (0,\ 1)$ which means that the dijets are back to back and their direction is roughly perpendicular to the beam axis. Another significant class of jets are those in the region of $(1,\ 3) \times (0,\ 1)$ that are roughly symmetrical with respect to beam axis but their momentum is not conserved and so there has to be another process that carries away momentum. As can be seen the amount of back to back jets with high rapidities in oposite directions (region $(0,\ 1) \times (2,\ 3)$) is dwindling in comparison with the rest of significant regions. This tells us that most back to back jets are created in the events where a hard scattering with low rapidity takes place.

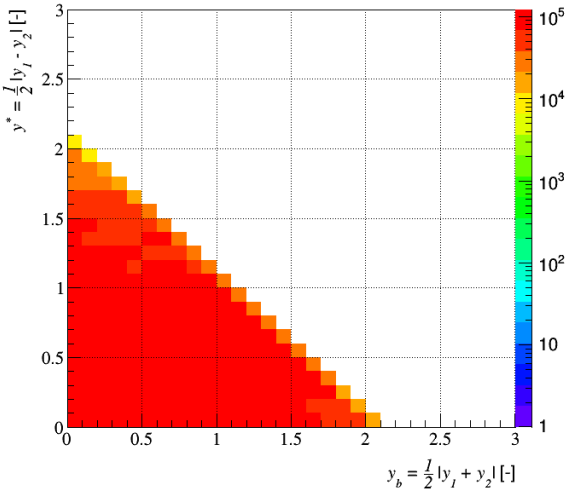
For the second selection type, the evidence in Fig. 6.13c is clear, that the most of the jets inside the inner detector are back to back or slightly boosted which hints at a process that carries away part of the total momentum in the event. The distributions of y_b and y^* can be found in the appendix A (Fig. A.1a to A.1d).



(a) Event topology legend taken from [20]



(b) First selection type.



(c) Second selection type.

Fig. 6.13. The topology of dijet events for both selection types.

Conclusion

The main theme of this thesis has been the study of particle jets in general and the analysis of dijet event in the ATLAS experiment at the Large Hadron Collider in CERN. In order to acquire some understanding of the context of this thesis, a brief introduction to the Standard model of particle physics has been put forth, along with an overview of the ATLAS detector. A brief foray into the hard scattering processes has been made in order to shed some light on the framework and tools that are being used during the analysis.

A powerful tool that almost every modern particle physicist uses and needs is MC simulation. Throughout this thesis has been used the LO generator Herwig 7.0.2 with the NLO generator of matrix elements Powheg in order to acquire simulated events to test our analysis and demonstrate certain discussed topics.

The cornerstone of any jet analysis is the use of jet algorithms in order to properly define and reconstruct particle jets from either measured data or MC simluations. To this end, a succinct elaboration on the workings of k_t algorithm, generalised Anti-k_t algorithm and SISCone has been included. A display of application of the k_t and Anti-k_t algorithm on a multi-jet event simulated using Herwig 7.0.2 has been done to emphasize the difference in the shape and size of jets acquired using different algorithms and the superiority of the Anti-k_t algorithm.

Then, choosing to further use only Anti-k_t, a simple study of jet substructure has been put forth with the intention to show the distinctions between Minimum Bias event generation and Leadig order of perturbative QCD and their effects on the shape of p_T spectrum of inclusive charged jets, the charged particle multiplicity per jet and charged track density in jets. The standard analysis [14] has been used as reference point for tuning of the MC generator input and as a source of data from proton proton collision at the centre of mass energy of $\sqrt{s} = 7$ TeV as measured by the ATLAS detector at the LHC. Furthermore a prediction about the shape of these distributions for the centre of mass energy of $\sqrt{s} = 13$ TeV has been made.

And as the title of this thesis suggests the last part has been about dijet events. The whole set of data from measurements of proton proton collisions at the centre of mass energy $\sqrt{s} = 8$ TeV recorded by the ATLAS detector at the LHC in 2012 has been analysed. The results of the measurement has been compared with the Powheg + Pythia generated simulation, where Powheg computed the parton level events and Pythia completed them into into the hadron level. The analysis has been aimed mainly on the measurement of properties that are useful for description of dijet event topology, however inclusive jet p_T spectrum and multiplicity for the two event selection types has been included as well. The main measured quantities has been: The p_T spectrum of the leading and subleading jet, the mass spectrum of the dijet pair, which have been used to determine the level of functionality of the analysis. The dijet transverse momentum balance, along with its normalised variant, which yield the information of event topology in the p_T phase space. Lastly the boost of the dijet system and the rapidity separation of dijets which has then been combined to create a 2D plot of event topologies in the rapidity phase space. All MC generated predictions are in good agreement with the measurement withing the constraints of total un-

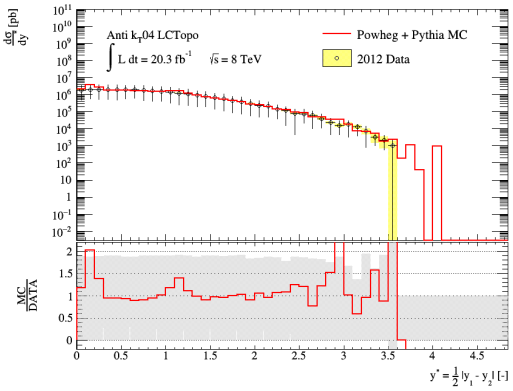
certainty, which was taken as the square sum of systematic and statistical uncertainties. This agreement notwithstanding there is a room for improvement in the determination of the systematic uncertainties and even cleaning process.

In further work the plan is to learn more precise ways of determination of systematic uncertainties and apply these in the analysis of new data. Another step forward would be to implement the study of unfolding and truth jets and apply the results of dijet analysis into the study of vector boson creation associated with dijet production.

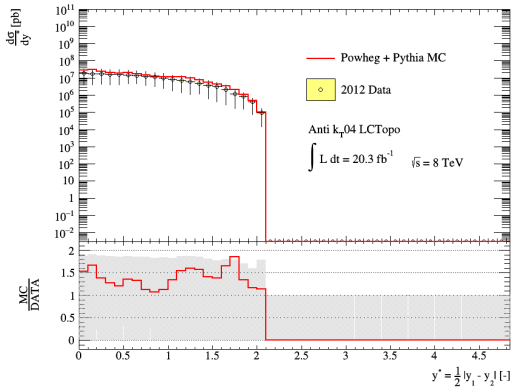
Appendix A

Data Analysis pictures

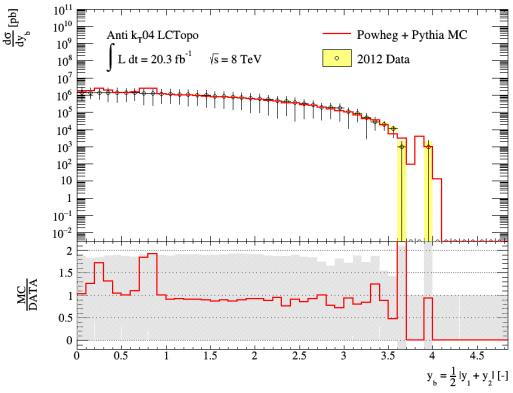
A.1 Boost of the dijet system and dijet rapidity separation.



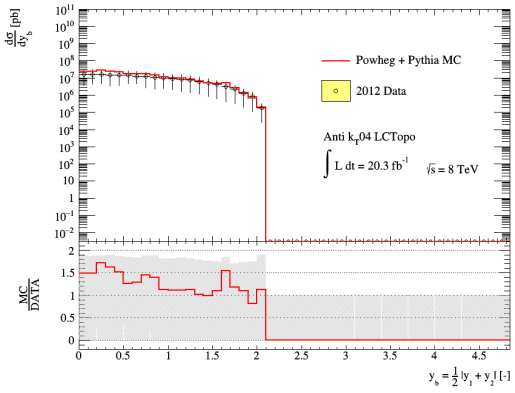
(a) Dijet rapidity separation, first selection type



(b) Dijet rapidity separation, second selection type



(c) Boost of the dijet system, first selection type



(d) Boost of the dijet system, second selection type

Fig. A.1. The dijet rapidity separation distribution and the boost of the dijet system distribution for both selection types

A.2 Systematic uncertainties of measured quantities

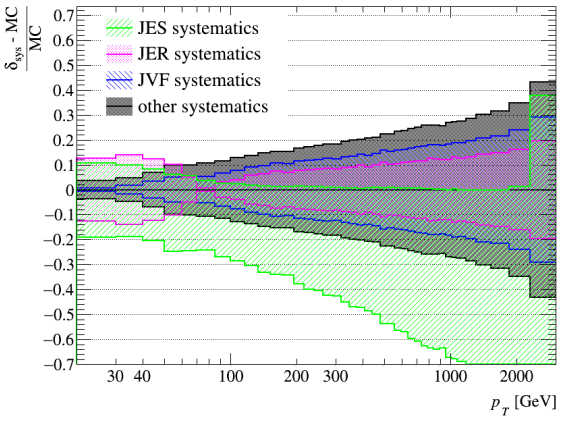
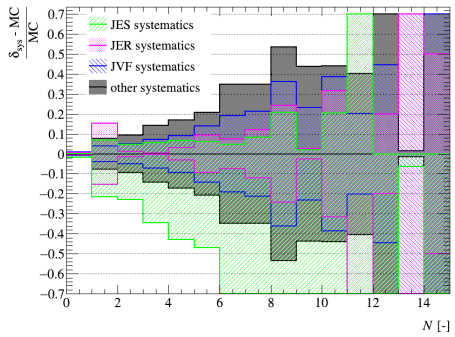
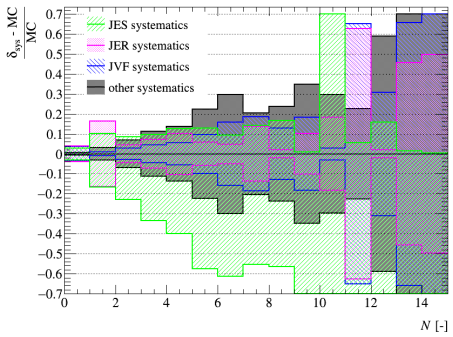


Fig. A.2. Systematic uncertainties of the transverse momentum spectrum of inclusive jets for the second selection type in the rapidity range (0.0, 2.1).

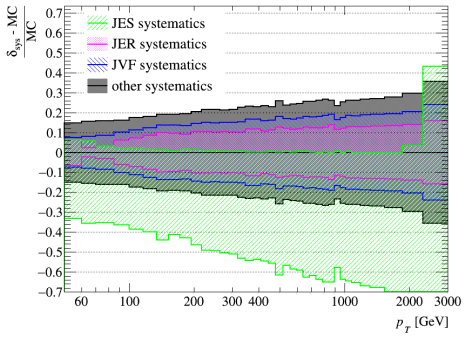


(a) First selection type

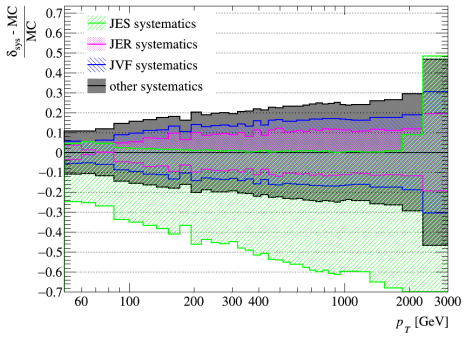


(b) Second selection type

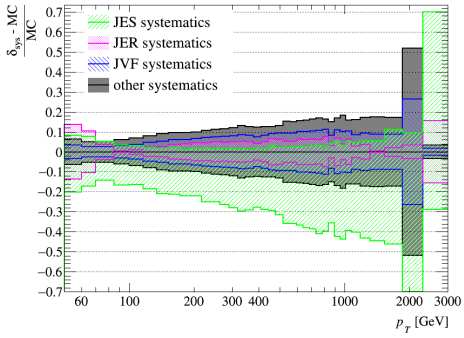
Fig. A.3. Systematic uncertainties of the charged jet multiplicity for both selection types.



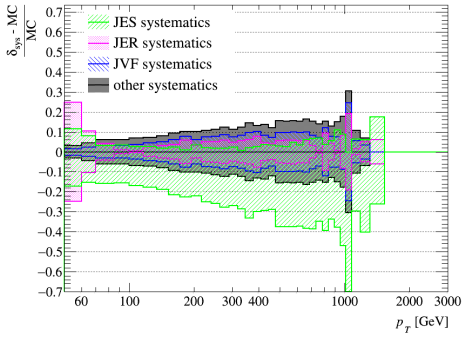
(a) Rapidity range (0.0, 0.5)



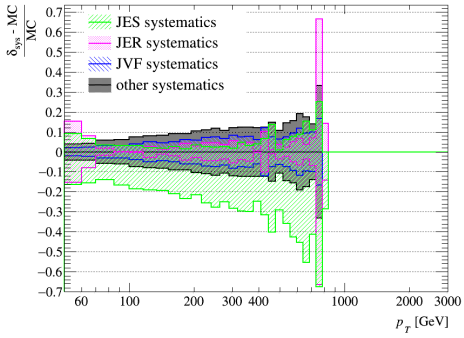
(b) Rapidity range (0.5, 1.0)



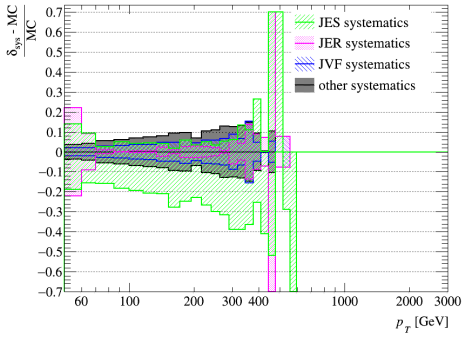
(c) Rapidity range (1.0, 1.5)



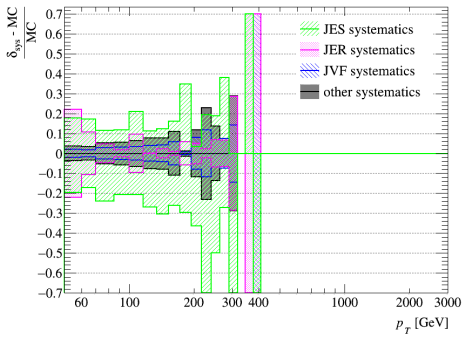
(d) Rapidity range (1.5, 2.0)



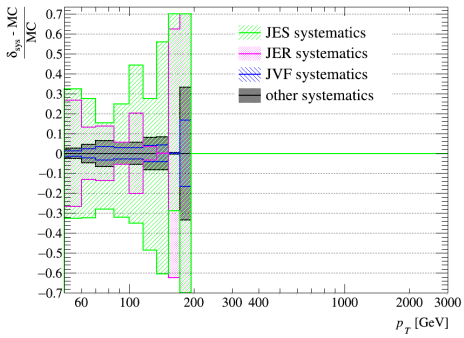
(e) Rapidity range (2.0, 2.5)



(f) Rapidity range (2.5, 3.0)

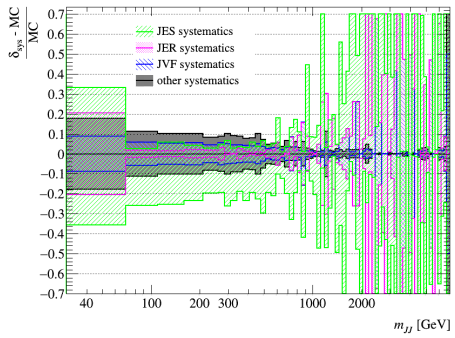


(g) Rapidity range (3.0, 3.5)

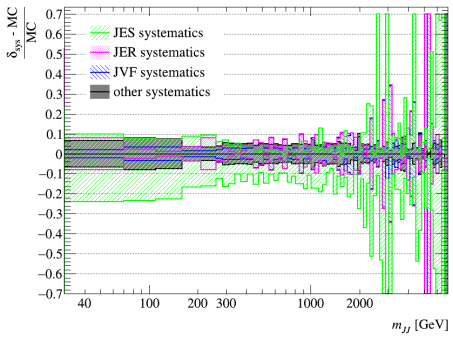


(h) Rapidity range (3.5, 4.0)

Fig. A.4. Systematic uncertainties of the transverse momentum spectrum of inclusive jets for the first selection type.

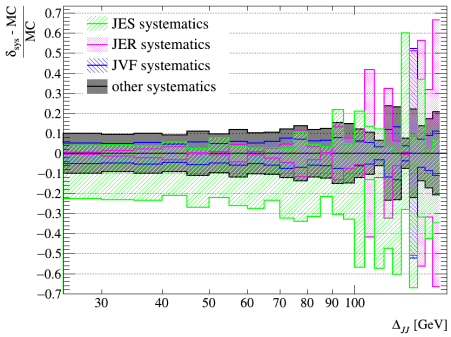


(a) First selection type

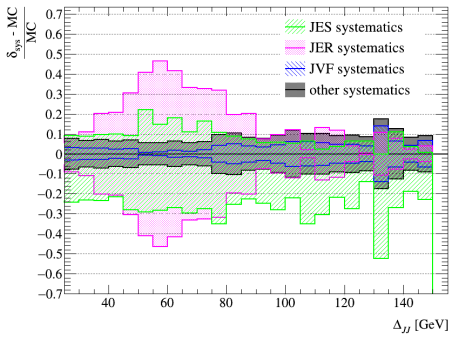


(b) Second selection type

Fig. A.5. Systematic uncertainties of the dijet mass spectra for both selection types.

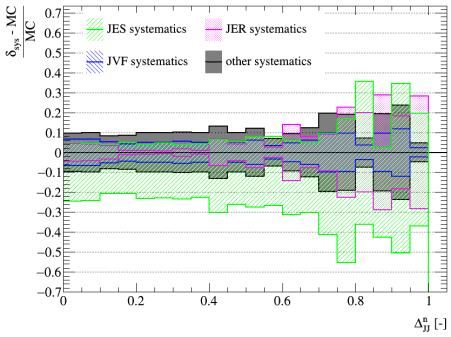


(a) First selection type.

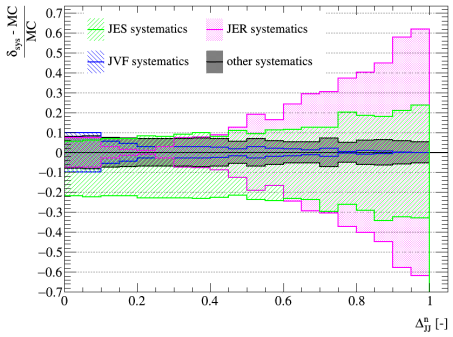


(b) Second selection type.

Fig. A.6. The comparison between the systematic uncertainties of the dijet momentum balance Δ_{JJ} for both selection types.

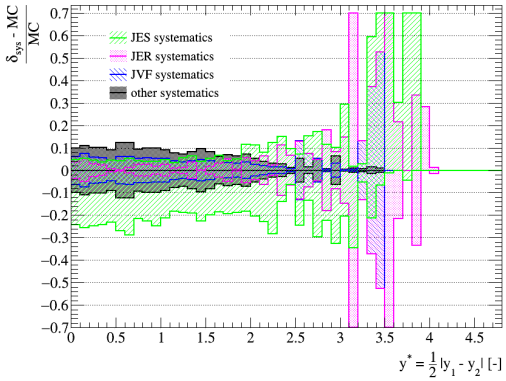


(a) First selection type.

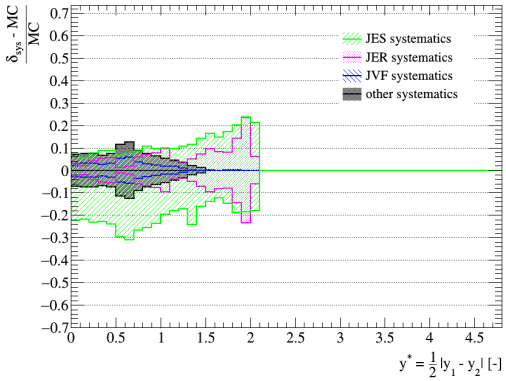


(b) Second selection type.

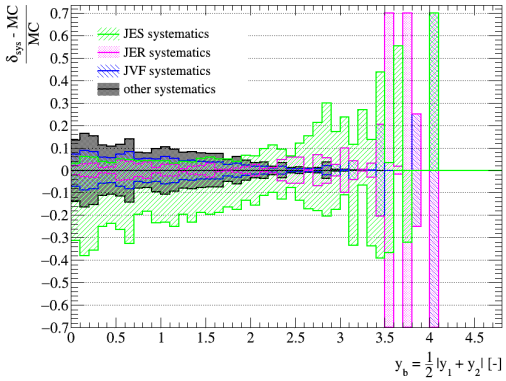
Fig. A.7. The comparison between the systematic uncertainties of the dijet momentum balance Δ_{JJ} for both selection types.



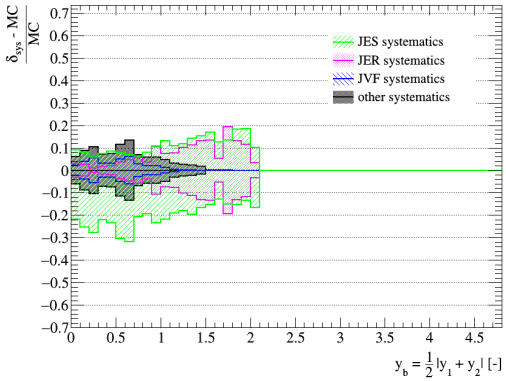
(a) Dijet rapidity separation, first selection type



(b) Dijet rapidity separation, second selection type



(c) Boost of the dijet system, first selection type



(d) Boost of the dijet system, second selection type

Fig. A.8. The systematic uncertainties of the dijet rapidity separation distribution and the boost of the dijet system distribution for both selection types

Bibliography

[1] Jiří Chýla. *Quarks, partons and Quantum Chromodynamics*. Accessed: 14.02.2016.

[2] Alan D. Martin. Proton structure, Partons, QCD, DGLAP and beyond. *Acta Phys. Polon.*, B39:2025–2062, 2008.

[3] Wikipedia.org. Standard model of elementary particles. 2008.

[4] K. A. Olive et al. Review of Particle Physics. *Chin. Phys.*, C38:090001, 2014.

[5] Takaaki Kajita. Nobel Lecture: Discovery of atmospheric neutrino oscillations. *Rev. Mod. Phys.*, 88(3):030501, 2016.

[6] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014.

[7] John M. Campbell, J. W. Huston, and W. J. Stirling. Hard Interactions of Quarks and Gluons: A Primer for LHC Physics. *Rept. Prog. Phys.*, 70:89, 2007.

[8] G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.

[9] Gavin P. Salam. Towards Jetography. *Eur. Phys. J.*, C67:637–686, 2010.

[10] Gavin P. Salam and Gregory Soyez. A Practical Seedless Infrared-Safe Cone jet algorithm. *JHEP*, 05:086, 2007.

[11] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet User Manual. *Eur. Phys. J.*, C72:1896, 2012.

[12] M. Bahr et al. Herwig++ Physics and Manual. *Eur. Phys. J.*, C58:639–707, 2008.

[13] Johannes Bellm et al. Herwig 7.0/Herwig++ 3.0 release note. *Eur. Phys. J.*, C76(4):196, 2016.

[14] Georges Aad et al. Properties of jets measured from tracks in proton-proton collisions at center-of-mass energy $\sqrt{s} = 7$ TeV with the ATLAS detector. *Phys. Rev.*, D84:054001, 2011.

[15] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.

[16] Pieter van der Deijl. *Double Parton Interactions with a Z+2 Jets Signature in Proton-Proton Collisions at the LHC*. PhD thesis, Twente U., Enschede, 2016.

[17] Light-quark and Gluon Jets in ATLAS. 2011.

[18] The ATLAS collaboration. Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS. 2015.

[19] Georges Aad et al. Search for new phenomena in the dijet mass distribution using $p - p$ collision data at $\sqrt{s}$ = 8 TeV with the ATLAS detector. *Phys. Rev.*, D91(5):052007, 2015.

[20] Albert M Sirunyan et al. Measurement of the triple-differential dijet cross section in proton-proton collisions at sqrt(s) = 8 TeV and constraints on parton distribution functions. 2017.