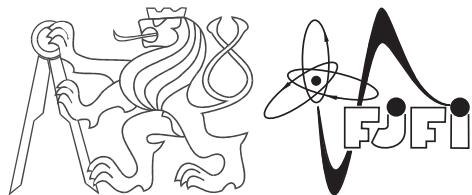


CZECH TECHNICAL UNIVERSITY IN PRAGUE  
FACULTY OF NUCLEAR SCIENCES AND PHYSICAL  
ENGINEERING  
DEPARTMENT OF PHYSICS

Programme: Mathematical Engineering  
Branch of Study: Mathematical Physics



# High $p_T$ jets in RunII of the ATLAS Experiment

MASTER'S DEGREE PROJECT

Author: Jan Lochman  
Supervisor: Ing. Zdeněk Hubáček, Ph.D.  
Submitted in: May 2015



Zadani prace



### **Statement**

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

In Prague on .....  
.....

Jan Lochman



## **Acknowledgment**

There are many people to whom I am indebted. First and the most important of these is my supervisor, Ing. Zdeněk Hubáček, Ph.D.. Two years ago, he threw me into the unknown world of particle physics. With his help, I learned not to be frightened from the unknown. He taught me, how to use knowledge, I have obtained at the university, to shed some light into the darkness. With his help, the beautiful world of particle physics has revealed to me.

I deeply appreciate a support from my Faculty, when I was sent after my supervisor at CERN. The leading support, which allowed this to happen, came from Prof. Ing. Igor Jex, DrSc., to whom I would like to thank. It was astounding feeling to be part of one of the leading experiments in particle physics and the satisfaction, when I found out, I can make some contribution to it - indescribable.

Next thanks to my family for their support during my studies.

And thanks to You, Flying Spaghetti Monster, for creating such a beautiful Universe, in which the physics reality extends far beyond the frontiers of human imagination, that it is never going to let us relax.

Jan Lochman



*Název práce:*

**Jety s vysokou příčnou hybností v RunII experimentu ATLAS**

*Autor:* Jan Lochman

*Obor:* Matematické Inženýrství

*Druh práce:* Diplomová práce

*Vedoucí práce:* Ing. Zdeněk Hubáček, Ph.D.

CERN

*Abstrakt:* Abstrakt CZ

*Klíčová slova:* Klicova slova

*Title:*

**High  $p_T$  jets in RunII of the ATLAS Experiment**

*Author:* Jan Lochman

*Abstract:* This thesis deals with the measurement of the inclusive jet double differential cross section in  $p_T$  and rapidity in proton-proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector. The measurement uses the Monte Carlo simulated data by the PYTHIA8 event generator and anti- $k_t$  jets with parameter  $R = 0.4$  reconstructed at particle and detector level. Differential cross section obtained from the detector level is unfolded to the particle level and compared with the parton level differential cross section prediction of the next-to-leading perturbative QCD. Two different approaches of unfolding denoted as the simple and 2D unfolding are implemented and results of these two approaches are compared. Both of these approaches can be applied to unfold differential cross section of inclusive jets measured by the ATLAS detector in LHC Run II, which will start in early summer 2015.

*Key words:* Quantum Chromodynamics, Jets, Unfolding



# Contents

<b>List of Tables</b>	<b>iii</b>
<b>List of Figures</b>	<b>iv</b>
<b>Introduction</b>	<b>1</b>
<b>1 QCD</b>	<b>5</b>
1.1 Theoretical Ansatz . . . . .	7
1.2 Experimental Ground . . . . .	11
1.2.1 Scattering Reactions . . . . .	11
1.2.2 Number of Colors . . . . .	14
1.3 QCD as a Gauge Theory . . . . .	15
1.4 Perturbative QCD . . . . .	17
1.5 Non-Perturbative QCD . . . . .	20
<b>2 Experimental Framework</b>	<b>23</b>
2.1 The Large Hadron Collider and The ATLAS Detector . . . . .	23
2.1.1 The Large Hadron Collider . . . . .	23
2.1.2 The ATLAS Detector . . . . .	25
2.2 Hadron Collision at LHC . . . . .	27
2.3 Jet Algorithms . . . . .	28
2.3.1 Cone algorithms . . . . .	29
2.3.2 $k_t$ algorithms . . . . .	30
2.3.3 Recombination . . . . .	31
2.3.4 Calorimeter jets . . . . .	32
2.4 Jet corrections . . . . .	33
2.4.1 Jet Energy Scale Calibration . . . . .	34
2.4.2 Unfolding . . . . .	34

<b>3 Data Analysis</b>	<b>37</b>
3.1 Data Characteristics . . . . .	37
3.2 Event Selection . . . . .	39
3.2.1 Jet Cuts . . . . .	39
3.2.2 Jet Matching . . . . .	40
3.3 Unfolding . . . . .	41
3.4 Comparison with NLO Prediction . . . . .	45
<b>Conclusion</b>	<b>51</b>
<b>Appendices</b>	<b>55</b>
<b>A Cut and Match Results</b>	<b>57</b>
A.1 Cut Results . . . . .	58
A.2 Match Results . . . . .	59
A.3 Truth and Reco $p_T$ Spectra . . . . .	60
<b>B Unfolding</b>	<b>63</b>
B.1 Matching Efficiency . . . . .	63
B.2 Slices in Unfolding Matrix . . . . .	66
B.3 Unfolding Results . . . . .	67
B.4 Simple and 2D Unfolding . . . . .	68
<b>C NLO QCD Prediction</b>	<b>69</b>
C.1 Predictions for Run I and Run II . . . . .	70
C.2 NLO Uncertainties . . . . .	73
C.3 Pythia and NLO . . . . .	75
<b>Bibliography</b>	<b>76</b>

# List of Tables

1.1	Quantum numbers of selected baryons known in 1950s. $S$ strangeness, $Y$ hypercharge, $T$ isospin, $T_3$ third component of isospin, $Q$ electrical charge. . . . .	8
1.2	Quantum numbers of three quarks which existence was predicted by Gell-Mann and Zweig in 1964. . . . .	10
3.1	The cross-sections (XS), filter efficiency (FE) and number of events for the JZXW samples which differ in the leading truth jet $p_T$ range. . . . .	38
A.1	Statistics for matching and cutting procedures described in Sections 3.2.1 and 3.2.2 displayed for all jets and for individual JZXW samples defined in Table 3.1. At the top, there is number of initial reco and truth jets respectively. For each cut, there is shown the number of jets, which were killed by it, and their relative number according to the original number of reco or truth jets respectively. Last two lines show the statistics of matching procedure including number of jets which were (un)matched. . . . .	61



# List of Figures

1	History of elementary particle physics. . . . .	2
1.1	The system of fundamental particles of the SM. . . . .	6
1.2	Eigenvalues of three-dimensional representation of $\mathfrak{su}(3)$ Lie Algebra. . . . .	9
1.3	Baryonic octuplet encapsulating baryons from Table 1.1. For baryons in this diagram, the relation $Y = S + 1$ holds. . . . .	10
1.4	Scattering reaction $e^- N \rightarrow e^- N$ with kinematics variables and vertex algebraic structures. . . . .	12
1.5	(a) $e^- e^+$ annihilation into the pair of fermion anti-fermion. Constants siting in both vertices are denoted, with $\alpha$ being the Fine Structure Constant and $Q_f$ the charge of fermion $f$ in units of positron charge. (b) $\pi^0$ meson decay into pair of photons with closed fermion loop. . . . .	14
1.6	Leading order Feynmann diagrams in the scattering reaction $q\bar{q} \rightarrow q\bar{q}$ with denoted transferred momentum $k$ . . . . .	17
1.7	Perturbative corrections to the Feynmann diagram from Figure 1.6 representing the scattering reaction $q\bar{q} \rightarrow q\bar{q}$ . Dashed line represents scalar ghost particle. . . . .	18
1.8	Experimental measurements of running coupling constant $\alpha_S(Q)$ (solid line) and its uncertainty (yellow band). $Q = \sqrt{ k^2 }$ in comparison to (1.33). . . . .	19

2.1	This diagram shows the locations of the four main experiments (ALICE, ATLAS, CMS and LHCb) that take place at the LHC. Located between 50 m and 150 m underground, huge caverns have been excavated to house the giant detectors. The Super Proton Synchotron (SPS), the final link in the pre-acceleration chain, and its connection tunnels to the LHC are also shown. . . . .	24
2.2	(a) an overview of the ATLAS detector (b) detail on the inner detector and the calorimeters - the dominant sub-detector systems used in this thesis. . . . .	26
2.3	Schematic representation of a hard scattering proton-proton collision. . . . .	27
2.4	Illustration of (a) infrared unsafety and (b) collinear unsafety of fixed cone jet algorithm. . . . .	30
2.5	Illustration of $k_t$ and anti- $k_t$ jet algorithms with $R = 1$ for calorimeter signal towers in azimuth $\Phi$ and pseudorapidity $y$ . Towers of the same color were recombined to one jet. . . . .	31
2.6	Three levels of jet reconstruction. . . . .	33
3.1	Comparison of $p_T$ spectra of reco and truth jets after event selection for selected $ y  < 0.5$ rapidity bin. Each $p_T$ bin was divided by its width so $y$ -axis has physical meaning of double differential cross section in $p_T$ and $y$ . Bottom graph contains the relative difference between reco and truth cross sections. .	41
3.2	Unfolding matrix for matched reco and truth jets with rapidity $ y  < 0.5$ corresponding to one of eight matrices used in the simple unfolding. Each cell is proportional to the number of jets with truth $p_T$ in range determined by the $x$ -axis which were reconstructed to the reco jets with $p_T$ determined by the $y$ -axis. White space signalize no input. . . . .	43

3.3	Unfolding matrix used in the 2D unfolding. Each cell is proportional to the number of jets with truth $p_T$ and rapidity $y$ determined by the $x$ -axis, which were reconstructed to the reco jets with $p_T$ and $y$ determined by the $y$ -axis. Marked square in $ y  < 0.5$ region is the matrix shown in Figure 3.2. Projection of this matrix on the $x$ and $y$ -axis corresponds to the $p_T$ spectrum of matched truth and reco jets for corresponding rapidity bin respectively. Projections on $y$ axis are shown in Appendix B.2. . . . .	44
3.4	Comparison of matching efficiency of simple and 2D unfolding for $ y  < 0.5$ rapidity bin. Matching efficiency is compared for both reco jets (left) and truth jets (right). Matching efficiency for all rapidity bins is shown in Appendix B.1. . . . .	46
3.5	Comparison of $p_T$ spectra of reco jets and the unfolded $p_T$ spectra (2D unfolding) with the $p_T$ spectra of the truth jets (left). Comparison of unfolded spectra obtained by the 2D and simple unfolding with the $p_T$ spectra of the truth jets (right). Both graphs are for $ y  < 0.5$ rapidity regions. Results for all rapidity bins are shown in Appendices B.3, B.4. . . . .	46
3.6	Comparison of NLO QCD prediction of double differential inclusive jet cross section (black) for $pp$ collisions at $\sqrt{s} = 13 \text{ TeV}$ (filled circles) corresponding to the LHC Run II and $\sqrt{s} = 8 \text{ TeV}$ (empty circles) corresponding to the LHC Run I. The cross section is multiplied by integrated luminosities and $p_T$ bin width to obtain the expected number of jets observed in each $p_T$ bin (blue). Figure shows only $ y  < 0.5$ rapidity bin, remaining rapidity bins are shown in Appendix C.1. . . . .	47
3.7	Theoretical uncertainties for NLO QCD predictions of inclusive jet differential cross section for $pp$ collisions at $\sqrt{s} = 8 \text{ TeV}$ (left) and $\sqrt{s} = 13 \text{ TeV}$ (right) for $ y  < 0.5$ rapidity bin. Uncertainties for other rapidity bins are shown in Appendix C.2. . . . .	49
3.8	Comparison of PYTHIA8 prediction with NLO QCD prediction of inclusive jet double differential cross section for $ y  < 0.5$ rapidity bin with uncertainties of NLO QCD predictions symbolized by the blue area. Comparisons for other rapidity bins are shown in Appendix C.3. . . . .	49

A.1	Impact of 4 cuts defined in Section 3.2.1 on differential cross section in $p_T$ of reco jets (top) and truth jets (bottom). Black dots represent the original uncutted spectrum, green area then these jets, which survived all four cutoffs. . . . .	58
A.2	Results of matching procedure described in Section 3.2.2 demonstrated on differential cross section in $p_T$ of reco (top) and truth (bottom) jets. Black dots represent the original uncutted spectrum. The contribution of matched and unmatched jets to green area representing all jets which survived cutoffs is shown. . . . .	59
A.3	Inclusive jet cross section of truth (top) and reco (bottom) jets as the function of jet $p_T$ and rapidity. For the convenience the cross sections for different rapidity bins are multiplied by the factor indicated in the legend. Jets were identified with anti- $k_t$ jet algorithm with $R = 0.4$ . . . . .	60
B.1	Comparison of matching efficiencies of simple and 2D unfolding for $ y  < 0.5$ (top) and $0.5 \leq  y  < 1$ (bottom) rapidity bins. Matching efficiencies are shown for both reco jets (left) and truth jets (right). . . . .	63
B.2	Comparison of matching efficiencies of simple and 2D unfolding for $1 \leq  y  < 1.5$ (top), $1.5 \leq  y  < 2$ and $2 \leq  y  < 2.5$ (bottom) rapidity bins. Matching efficiencies are shown for both reco jets (left) and truth jets (right). . . . .	64
B.3	Comparison of matching efficiencies of simple and 2D unfolding for $2.5 \leq  y  < 3$ (top), $3 \leq  y  < 3.5$ and $3.5 \leq  y  < 4$ (bottom) rapidity bins. Matching efficiencies are shown for both reco jets (left) and truth jets (right). . . . .	65
B.4	Slices in transfer matrix of 2D unfolding method from Figure on $y$ -axis represnted by the $p_T$ spectrum of reco jets. Slices are done in each rapidity bin for three different ranges of $p_T$ of truth jets. . . . .	66
B.5	Comparison of $p_T$ spectra of reco jets and the unfolded $p_T$ spectra (2D unfolding) with the $p_T$ spectra of the truth jets for 8 different rapidity bins. . . . .	67

B.6 Comparison of results obtained by two unfolding method - simple and 2D unfolding - with the $p_T$ spectra of truth jets for 8 different rapidity bins. . . . .	68
C.1 Comparison of NLO QCD prediction of double differential inclusive jet cross section (black) in $p_T$ and rapidity of $pp$ collisions at $\sqrt{s} = 13$ TeV (filled circles) corresponding to LHC Run II and $\sqrt{s} = 8$ TeV (empty circles) corresponding to LHC Run I. The cross section is multiplied by integrated luminosities and $p_T$ bin width to obtain expected number of jets observed in each $p_T$ bin (blue). Figures show the comparison for $0.5 <  y $ (top) and $0.5 \leq  y  < 1$ (bottom) rapidity bins. . . . .	70
C.2 Comparison of NLO QCD prediction of double differential inclusive jet cross section (black) in $p_T$ and rapidity of $pp$ collisions at $\sqrt{s} = 13$ TeV (filled circles) corresponding to LHC Run II and $\sqrt{s} = 8$ TeV (empty circles) corresponding to LHC Run I. The cross section is multiplied by integrated luminosities and $p_T$ bin width to obtain expected number of jets observed in each $p_T$ bin (blue). Figures show the comparison for $1 \leq  y  < 1.5$ (top) and $1.5 \leq  y  < 2$ (bottom) rapidity bins. . . . .	71
C.3 Comparison of NLO QCD prediction of double differential inclusive jet cross section (black) in $p_T$ and rapidity of $pp$ collisions at $\sqrt{s} = 13$ TeV (filled circles) corresponding to LHC Run II and $\sqrt{s} = 8$ TeV (empty circles) corresponding to LHC Run I. The cross section is multiplied by integrated luminosities and $p_T$ bin width to obtain expected number of jets observed in each $p_T$ bin (blue). Figures show the comparison for $2 \leq  y  < 2.5$ (top) and $2.5 \leq  y  < 3$ (bottom) rapidity bins. . . . .	72
C.4 NLO uncertainties for NLO QCD predictions of inclusive jet double differential cross section of $pp$ collisions at $\sqrt{s} = 8$ TeV (left) and $\sqrt{s} = 13$ TeV (right) for $ y  < 0.5$ (top), $0.5 \leq  y  < 1$ (middle) and $1 \leq  y  < 1.5$ (bottom) rapidity bins. . . . .	73
C.5 NLO uncertainties for NLO QCD predictions of inclusive jet double differential cross section of $pp$ collisions at $\sqrt{s} = 8$ TeV (left) and $\sqrt{s} = 13$ TeV (right) for $1.5 \leq  y  < 2$ (top), $2 \leq  y  < 2.5$ (middle) and $2.5 \leq  y  < 3$ (bottom) rapidity bins. . . . .	74



# Introduction

Search for the superior equation which could explain all the physical universe we observe, sometimes called the Theory of Everything, led some of the physicists to the concept of elementary particles, some of which define the building blocks of our observable universe, whereas the remaining govern their interaction.

From the beginning of the twentieth century, the term elementary particle was redefined by a new generation of physicists, as illustrated in Figure 1. The latest reform was caused by quarks and the invention of Quantum Chromodynamics, describing its strong interaction ,which, apart from the electromagnetic and weak interaction, is encapsulated by the present theory of elementary particles called the Standard Model.

Although the Standard Model contains mechanism for assigning masses to elementary particles, gravity was not included in it up to date, because the current attempts to quantize and describe gravitation as an interaction mediated by the quanta of gravity, known as gravitons, led to unrenormalizable theories. There are other questions, unresolved by the Standard Model, such as the nature of the dark part of our Universe or the origin of the matter-antimatter asymmetry.

We know that the Standard Model is not the ultimate Theory of Everything, but it successfully stands against present results from particle physics experiments. Last discoveries of elementary particles, successfully predicted by the Standard Model, occurred on  $\sim 100$  GeV energy scale with top quark discovery [2, 3] with mass  $173.34 \pm 0.98$  GeV in 1995 at Tevatron and Higgs boson discovery [4] with mass  $125.09 \pm 0.32$  GeV in 2012 at CERN. If there is a new physics beyond the Standard Model on  $\sim$  TeV scale, the LHC Run II could be the first to discover it [5].

This thesis deals with the measurement of the double differential inclu-

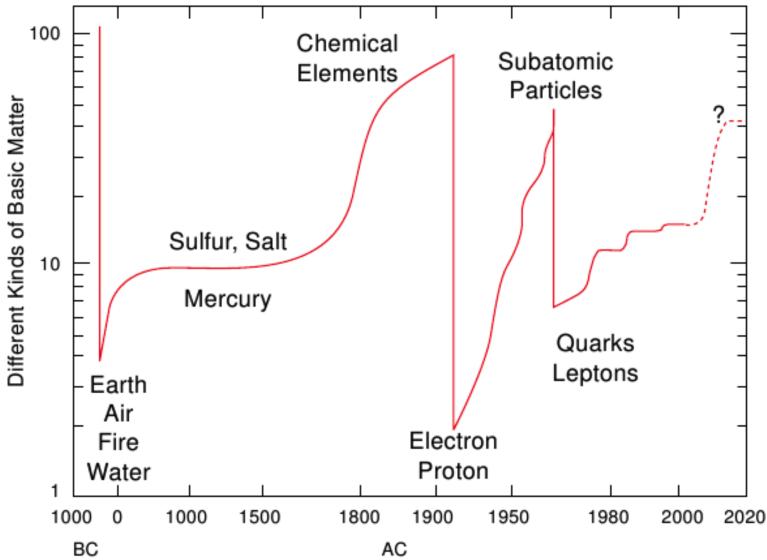


Figure 1: History of elementary particle physics. Figure taken from [1].

sive jet cross section in  $p_T$  and rapidity. The term inclusive stresses the fact, we are not interested in other particles, which can be, next to the jets, created by the collision. Jets are the dominant objects observed at inelastic collisions on hadron colliders and, covering a wide range of momentum transfers, they overshadow any other observable physics process in orders of magnitude. Inclusive jets measurements may therefore verify the ATLAS detector performance, as well as reveal some discrepancy in the Standard Model predictions.

First Chapter of this thesis discuss the Quantum Chromodynamics (QCD) and follows its historical development, including the experiments which led to the removal of the proton from the list of elementary particles, replacing it with the quarks. I will formulate the QCD as a quantum field theory and I will discuss the phenomenon known as a running coupling constant, so that the QCD can be divided into perturbative and non-perturbative regions.

The second Chapter concerns with the Large Hadron Collider along a detailed description of the ATLAS detector. I will use the basic features of the QCD, introduced in the previous Chapter, to define jets - objects predominantly observed at inelastic collisions on hadron colliders. At the end of this Chapter, I will also present the jet reconstruction on ATLAS detector, including description of jet calibration and unfolding of observed

spectra.

Third chapter describes the steps of the inclusive jet analysis, beginning with the characteristics of Monte Carlo data used, including event selection criteria. I use two approaches to unfold  $p_T$  spectra from detector level to particle level. In the end I compare results obtained from both approaches and set them against the next-to-leading order perturbative QCD prediction of parton level.



# Chapter 1

## QCD

*Is the purpose of theoretical physics to be no more than a cataloging of all the things that can happen when particles interact with each other and separate? Or is it to be an understanding at a deeper level in which there are things that are not directly observable (as the underlying quantized fields are) but in terms of which we shall have a more fundamental understanding?*

Julian Schwinger

The theoretical framework of particle physics is called the Standard Model. The Standard Model describes the way how the fundamental components of matter interact through strong, weak and electromagnetic interactions. Mathematically, the Standard Model is a gauge theory with local internal symmetries of the direct product group  $SU(3) \times SU(2) \times U(1)$ . Gauge bosons, particles with integer spin, are assigned to generators of this symmetry - there are 8 massless gluons from  $SU(3)$  and 3 massive  $W^\pm, Z$  bosons with 1 massless boson  $\gamma$  from electroweak  $SU(2) \times U(1)$  sector. In the electroweak sector, the Higgs Mechanism is introduced to assign  $W^\pm, Z$  bosons masses and as consequence, the new particle, Higgs boson, emerges in the Standard Model.

In addition to the bosons, the Standard Model introduces a spin-1/2 fermions, which are divided into three quark and three lepton families.

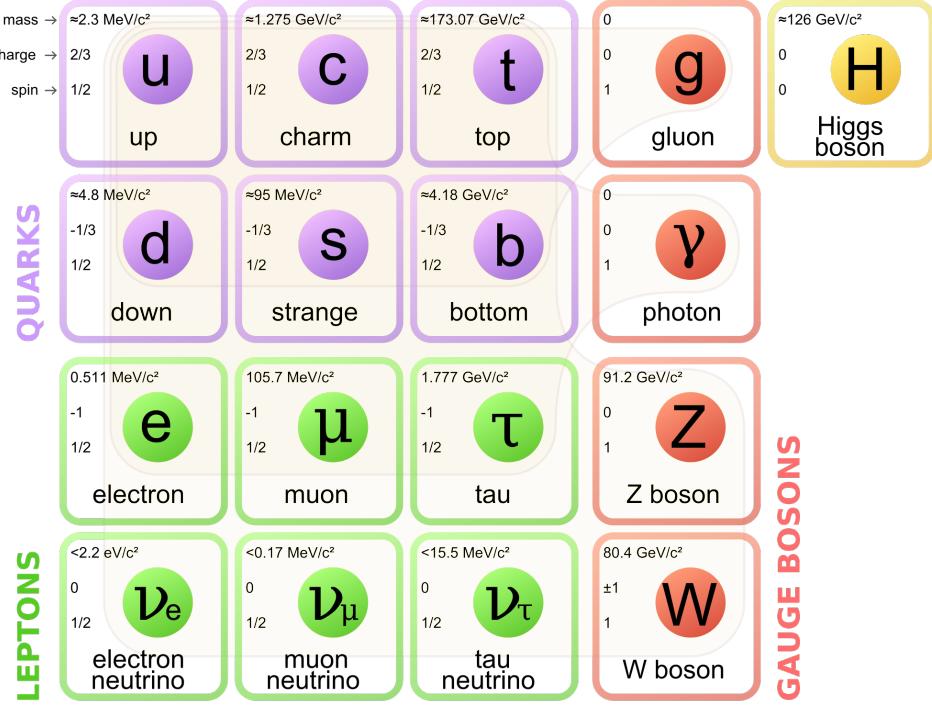


Figure 1.1: The system of fundamental particles of the SM. Figure taken from [6].

Fermions are assumed to be point-like, because there is no evidence for their internal structure to date. All fermions interact weakly, if they have electrical charge, they interact electromagnetically as well. Quarks are the only fundamental fermions which interact strongly. Figure 1.1 shows the system of fundamental particles of the Standard Model.

Quarks bind together to form hadrons and there are hundreds [7] of known hadrons up to date. Hadrons are divided into baryons (3 quarks) and mesons (quark and anti-quark pairs). A theory called Quantum Chromodynamics (QCD) describes the strong interaction between quarks. In this Chapter, I will discuss the key features of the QCD. I will give reasons for quark existence and for a description of their strong interaction as a  $SU(3)$  gauge theory. After an introduction of a QCD Lagrangian, I will derive an expression for the running coupling constant, which will be used to split the QCD into perturbative and non-perturbative regions. In these regions, the QCD has to describe the strong interaction with the use of different mathematical approaches. At the end of this Chapter, I will compare these two

approaches.

Most of ideas presented here is overtaken from the textbook [8] about the QCD. The electroweak sector of the Standard Model is described in [9]. For more concise information about the Standard Model, the following textbooks can serve [10, 11].

## 1.1 Theoretical Ansatz

In 1950s, there had already been discovered tens of hadrons, thanks to new particle accelerators, and a lot of effort was exerted to categorize them. Each hadron obtained a series of quantum numbers including isospin  $T$  with its third component  $T_3$ , hypercharge  $Y$ , electrical charge  $Q$ , strangeness  $S$ , baryon number  $B$  and others. Soon, people started to recognize some symmetries between these quantum numbers, like the famous Gell-Mann–Nishijima relation [12, 13]

$$Q = T_3 + 1/2Y \quad , \quad Y = B + S + \dots , \quad (1.1)$$

where dots denote charm, bottomness and topness and were introduced after the work of Gell-Mann and Nishijima. Some of the baryons, known in 1950s, are, together with their quantum numbers, shown in Table 1.1. In 1960s, the known hadrons were successfully categorized with the theory called Eightfold Way, which was published independently by Murray Gell-Mann [14] and George Zweig [15] in 1964. The Eightfold Way successfully predicted the existence of a new particle  $\Omega^-$  including its mass. In this Section, I present a basic ideas of the Eightfold Way.

The key feature of Eightfold Way is to understand hadron as a component of representation of infinitesimal generators of  $SU(3)$  flavor symmetry group. The infinitesimal generators of  $SU(3)$  form a real eight-dimensional Lie Algebra  $\mathfrak{su}(3)$ , which fundamental representation is usually derived from Gell-Mann matrices

	$S$	$Y$	$T$	$T_3$	$Q$
$p$	0	1	1/2	1/2 -1/2	1 0
$n$					
$\Sigma^+$				1	1
$\Sigma^0$				0	0
$\Sigma^-$	-1	0	1	-1	-1
$\Lambda$			0	0	0
$\Xi^0$				1/2	0
$\Xi^-$	-2	-1	1/2	-1/2	-1

Table 1.1: Quantum numbers of selected baryons known in 1950s.  $S$  strangeness,  $Y$  hypercharge,  $T$  isospin,  $T_3$  third component of isospin,  $Q$  electrical charge.

$$\begin{aligned} \lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \lambda_2 &= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \lambda_3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\ \lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} & \lambda_5 &= \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, & & & (1.2) \\ \lambda_6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} & \lambda_7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} & \lambda_8 &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \end{aligned}$$

The generators are usually chosen as  $g_a = \frac{1}{2}\lambda_a$  and obey the commutation relation  $[g_a, g_b] = if_{abc}g_c$  with  $f_{abc}$  being structure constants. Cartan subalgebra of fundamental representation of  $\mathfrak{su}(3)$  is generated by  $H_1 = g_3$  and  $H_2 = g_8$ . The eigenstates of three-dimensional representation of  $\mathfrak{su}(3)$  can be chosen as

$$u = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \leftrightarrow \left( \frac{1}{2}, \frac{\sqrt{3}}{6} \right), \quad d = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \leftrightarrow \left( -\frac{1}{2}, \frac{\sqrt{3}}{6} \right), \quad s = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \leftrightarrow \left( 0, -\frac{\sqrt{3}}{3} \right), \quad (1.3)$$

where I have assigned the eigenvalues to generators of the Cartan subalgebra  $H_1 u = \frac{1}{2}u$ ,  $H_2 u = \frac{\sqrt{3}}{6}u$  and similarly for  $d$  and  $s$  eigenstates. These

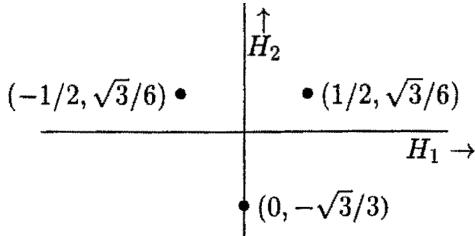


Figure 1.2: Eigenvalues of three-dimensional representation of  $\mathfrak{su}(3)$  Lie Algebra. Figure taken from [16].

eigenvalues are shown in Figure 1.2.

Next to the three-dimensional representation of  $\mathfrak{su}(3)$ , a eight-dimensional adjoint representation can be defined. This representation has the following eigenstates and eigenvalues

$$\begin{aligned} \frac{1}{\sqrt{2}} (g_1 \pm ig_2) &\leftrightarrow (\pm 1, 0), \\ \frac{1}{\sqrt{2}} (g_4 \pm ig_5) &\leftrightarrow \left( \pm \frac{1}{2}, \pm \frac{\sqrt{3}}{2} \right), \\ \frac{1}{\sqrt{2}} (g_6 \pm ig_7) &\leftrightarrow \left( \mp \frac{1}{2}, \pm \frac{\sqrt{3}}{2} \right), \end{aligned} \quad (1.4)$$

where again when denoting  $A = \frac{1}{\sqrt{2}}(g_1 + ig_2)$  then the upper sign of the first expression reads  $[H_1, A] = A$ ,  $[H_2, A] = 0$  and similarly for remaining 5 eigenstates. Defining

$$H_1 = T_3 \quad \text{and} \quad H_2 = \frac{\sqrt{3}}{2} Y \quad (1.5)$$

one can easily assign hadrons from Table 1.1 to corresponding eigenvalues of adjoint representation in (1.4) according to its third component of isospin  $T_3$  and its hypercharge  $Y$ . This is depicted in Figure 1.3.

When the same redefinition is done to the eigenstates of three-dimensional representation in (1.3), one can assign to  $u$ ,  $d$ ,  $s$  eigenstates the hypercharge  $Y$  and strangeness  $S$  as well. The concrete values for states are shown in Table 1.2.

Another representations of  $\mathfrak{su}(3)$  Lie Algebra can be found. The simplest way seems to be through highest weight defining representation. From eigenvalues of adjoint representation (1.4) one can find simple roots  $\alpha^1 = \left( \frac{1}{2}, \frac{\sqrt{3}}{2} \right)$ ,  $\alpha^2 = \left( \frac{1}{2}, -\frac{\sqrt{3}}{2} \right)$ , from which the highest weights follow  $\mu^1 = \left( \frac{1}{2}, \frac{\sqrt{3}}{6} \right)$ ,

	$S$	$Y$	$T$	$T_3$	$Q$
$u$	0	1/3	1/2	1/2	2/3
$d$			-1/2	-1/2	
$s$	-1	-2/3	0	0	-1/3

Table 1.2: Quantum numbers of three quarks which existence was predicted by Gell-Mann and Zweig in 1964.

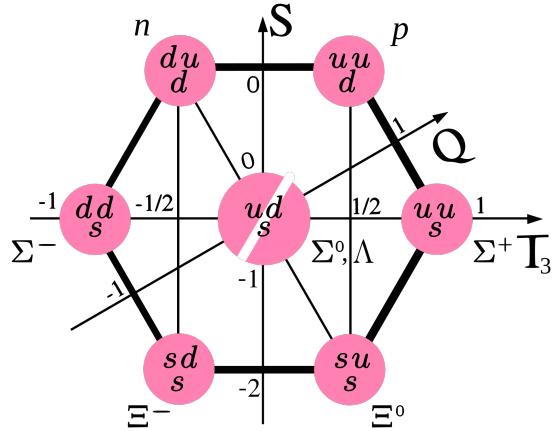


Figure 1.3: Baryonic octuplet encapsulating baryons from Table 1.1. For baryons in this diagram, the relation  $Y = S + 1$  holds. Figure taken from [17].

$\mu^2 = \left(\frac{1}{2}, -\frac{\sqrt{3}}{6}\right)$ . New representation of Lie Algebra can be constructed from highest weight. The whole procedure is described in [16] in detail.

Representations defined by the highest weight  $\mu^1$  or  $\mu^2$  respectively are called fundamental. Fundamental representation defined by  $\mu^1$  is usually denoted **3** and we have encountered it already by expressions (1.3) and weight diagram in Figure 1.2. This representation corresponds to three different quark states. The second fundamental representation, defined by the highest weight  $\mu^2$ , corresponds to three anti-quark states and is usually denoted **3̄**. The adjoint representation, depicted in Figure 1.3 is defined by the highest weight  $\mu^1 + \mu^2$ .

Special interest is in representations with dimensions 10 and 8. These are present in decompositions  $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{10} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{1}$ , which corresponds to the baryons composed of three quarks, and  $\mathbf{3} \otimes \mathbf{3̄} = \mathbf{8} \oplus \mathbf{1}$ , corresponding to mesons composed of quark and anti-quark.

Important feature of quark model just presented is its capability to predict hadron masses. This is done using Gell-Mann–Okubo mass formula [18, 19]

$$M = a_0 + a_1 S + a_2 \left( T(T+1) - \frac{1}{4} S^2 \right), \quad (1.6)$$

where  $a_0$ ,  $a_1$  and  $a_2$  are free parameters which are common for all hadrons in one multiplet.

In 1970, Sheldon Lee Glashow, John Iliopoulos and Luciano Maiani proposed [20] an extension to the Eightfold Way, which predicted existence of fourth flavor of quark - charm quark.

In 1973, Makoto Kobayashi and Toshihide Maskawa proposed [21], that the existence of 6 different quark flavors could explain the experimental observation of CP violation.

## 1.2 Experimental Ground

In the previous Section, I have shown, that the hadrons can be categorized by representations of  $\mathfrak{su}(3)$  Lie Algebra. This lead to the model, where baryons were composed of three quarks and the mesons of quark and anti-quark. In this Section, I summarize some experimental arguments to support quark model. Firstly, I will show, that the results from the lepton scattering on nucleons can be explained by assumption, that nucleons are composed of point-like spin-1/2 particles. In the second part, I will encounter the question, why the group  $SU(3)$  is connected to the theory of strong interaction.

### 1.2.1 Scattering Reactions

Inner structure of nucleon  $N$  can be investigated by one of the following scattering reactions

$$e^- (E \gg 1 \text{ GeV}) + N \rightarrow e^- + N, \quad (1.7)$$

$$\nu_e (E \gg 1 \text{ GeV}) + N \rightarrow \nu_e + N, \quad (1.8)$$

where I have explicitly written  $E \gg 1 \text{ GeV}$ , to ensure the wavelength of lepton is  $< 0.2 \text{ fm}$ . By the first scattering reaction, the information about electric charge distribution in nucleon can be extracted, whereas the second scattering reaction informs us about weak charge distribution. From now on, I work only with scattering reaction (1.7), which was experimentally

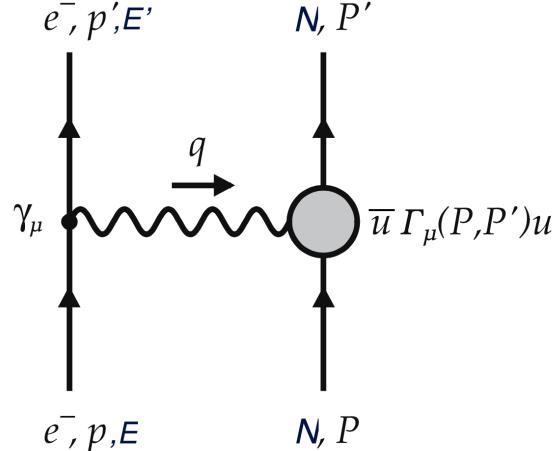


Figure 1.4: Scattering reaction  $e^- N \rightarrow e^- N$  with kinematics variables and vertex algebraic structures. Figure taken from [8].

examined as the first. Feynmann diagram of this process is, together with kinematics variables and vertex algebraic structures, depicted in Figure 1.4.

Because of Lorentz-invariance of Quantum Electrodynamics, the matrix element of the nucleon vertex  $\bar{u}(P', S')\Gamma_\mu u(P, S)$  has to be a Lorentz-vector. This restricts the possible form of  $\Gamma_\mu$  to the following algebraic structure

$$\Gamma_\mu = A\gamma_\mu + BP'_\mu + CP_\mu + iDP'^\nu\sigma_{\mu\nu} + iEP^\nu\sigma_{\mu\nu}, \quad (1.9)$$

where  $A, \dots, E$  depend only on Lorentz-invariant quantities. Next condition which has to be taken into account, is gauge invariance of matrix element, which can be written in the form

$$q^\mu\bar{u}(P', S')\Gamma_\mu u(P, S). \quad (1.10)$$

The further computation of cross section is straightforward and the result can be easily generalized to non-elastic scattering by which the nucleon in final state decays. The result is usually written using inelasticity parameter  $y = \frac{E-E'}{E}$ ,  $0 \leq y \leq 1$ ,  $y = 0$  corresponding to the elastic scattering, Bjorken variable  $x = \frac{Q^2}{2P \cdot q}$ ,  $0 < x \leq 1$ ,  $x = 1$  denoting elastic scattering and finally instead of negative value  $q^2$  the  $Q^2 = -q^2$  is used. Final result can be than written in the form

$$\left. \frac{d^2\sigma}{dx dy} \right|_{eN} = \frac{8\pi M_N E \alpha^2}{Q^4} [xy^2 F_1^{eN}(Q^2, x) + (1-y) F_2^{eN}(Q^2, x)]. \quad (1.11)$$

The  $eN$  sub(super)script stresses the fact, we are dealing with scattering (1.7).  $F_1^{eN}$  and  $F_2^{eN}$ , called Structure Functions, are not determinable by the theory just presented - they have to be measured experimentally.

Structure Functions were first measured at  $eP$  scattering at SLAC in 1968 [22] and have shown the following results

1. for  $Q^2 \geq 1 \text{ GeV}$ , there is no significant dependence of Structure Functions on  $Q^2$  and
2. for  $Q^2 \geq 1 \text{ GeV}$ ,  $F_2 \approx 2x F_1$ .

These results can be explained by assumption nucleon being composed of point-like spin-1/2 constituents, for which R. P. Feynmann used the term partons. In the following, I introduce the basic ideas of parton model.

To  $i$ th parton, we can assign momentum  $P_{i,\mu}$

$$P_{i,\mu} = \xi_i P_\mu + \Delta P_{i,\mu} \quad , \quad \max_{\mu}(\Delta P_\mu) \ll \max_{\mu} P_\mu, \quad (1.12)$$

where  $\xi_i \in \langle 0, 1 \rangle$  and  $\Delta P_{i,\mu}$  comes from the interaction between partons and we assume, the momentum coming from this interaction is much smaller than the total nucleon momentum  $P_\mu$ . In addition, probabilities  $f_i(\xi_i)$  that  $i$ th parton will carry  $\xi_i$  fraction of total momentum fulfilling

$$\int d\xi_i f_i(\xi_i) = 1 \quad (1.13)$$

must be defined. Then for scattering reaction (1.7) the formula for the cross section can be derived

$$\frac{d^2\sigma}{dxdy} \Big|_{eN} = \frac{4\pi M_N E \alpha^2}{Q^4} [y^2 + 2(1-y)] \sum_i f_i(x) q_i^2 x. \quad (1.14)$$

where for  $i$ th parton its electrical charge  $q_i$  was introduced. The last expression and expression (1.11) can be compared as polynomials in  $y$  resulting in

$$F_1^{eN}(x) = \frac{1}{2} \sum_i f_i(x) q_i^2 \quad , \quad F_2^{eN}(x) = \sum_i f_i(x) q_i^2 x. \quad (1.15)$$

It can be easily checked, that  $F_2^{eN}(x) = 2x F_1^{eN}(x)$ . Functions  $f_i(x)$  just introduced are called Parton Distribution Functions and their important role in QCD will be discussed in Section 3.4 in more details.

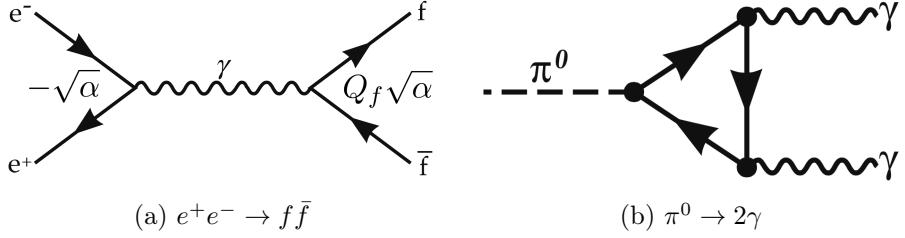


Figure 1.5: (a)  $e^-e^+$  annihilation into the pair of fermion anti-fermion. Constants sitting in both vertices are denoted, with  $\alpha$  being the Fine Structure Constant and  $Q_f$  the charge of fermion  $f$  in units of positron charge. (b)  $\pi^0$  meson decay into pair of photons with closed fermion loop.

The conclusion, which we should learn from the scattering reactions, is, that the experimental results can be explained by assumption nucleons being composed of spin-1/2 point-like partons, now called quarks.

### 1.2.2 Number of Colors

Despite the strong confidence in the parton model, a theory which would describe the interaction between partons was still missing. At the beginning of 1970s, there was no direct evidence on how the theory would look like. The theory of electroweak unification successfully suggested, that our Universe at a subatomical level, could be described by a gauge theories, but to construct a gauge theory of strong interaction the number of colors first had to be known.

Number of colors  $N_C$  is the number of different kinds of quarks of the same flavor with respect to the new interaction. In this part, I present three arguments to demonstrate, that  $N_C = 3$ .

The first argument is the analysis of the electron-positron annihilation into the pair of fermion and anti-fermion

$$e^+e^- \rightarrow f\bar{f}. \quad (1.16)$$

Feynmann diagram of this reaction is shown in Figure 1.5a, where constants sitting in two vertices are emphasized.  $\alpha$  stands for Fine Structure Constant and  $Q_f$  for the charge of the fermion  $f$  in units of positron charge. Total cross section has to be proportional to

$$\sigma(e^-e^+ \rightarrow f\bar{f}) \sim Q_f^2 \alpha^2. \quad (1.17)$$

In the case fermion  $f$  being quark, there is new degeneracy coming from different colors of quark-antiquark pair in final state - the total cross section has to be multiplied by factor  $N_C$ . Experimentally, the so called  $R$ -factor is measured

$$R = \frac{\sigma(e^+e^- \rightarrow \text{hadrons})}{\sigma(e^+e^- \rightarrow \mu^+\mu^-)} = \left( \sum_q Q_q^2 \right) N_C, \quad (1.18)$$

where the sum on the left hand side is over all possible quark flavors. When we use the quark model proposed by Gell-Mann a Zweig, and substitute the values from Table 1.2

$$R = \left[ \left( \frac{2}{3} \right)^2 + \left( \frac{-1}{3} \right)^2 + \left( \frac{-1}{3} \right)^2 \right] N_C = \frac{2}{3} N_C. \quad (1.19)$$

Experimental results for  $R$ -ratio have shown [23], that  $N_C = 3$ .

The second argument, to support  $N_C = 3$ , is the measurement of decay width of  $\pi_0$  meson, which is depicted in Figure 1.5b. For decay width  $\Gamma$  it can be derived

$$\Gamma = 7.63 \left( \frac{N_C}{3} \right)^2 \text{ eV}, \quad (1.20)$$

which, compared to the experimental value  $\Gamma = 7.57 \pm 0.32 \text{ eV}$  [23], leads again to  $N_C = 3$ .

The third argument is purely theoretical and states, that the Standard Model is internally consistent only if there are three colors [8]. This indicates, there is some linking between electroweak and strong sectors of the Standard Model, and motivates the search for Grand Unified Theories.

### 1.3 QCD as a Gauge Theory

Putting arguments of the previous Section all together, there is strong experimental evidence, that nucleons consist of point-like spin-1/2 particles called quarks and that quarks bring into the theory new degeneracy factor  $N_C = 3$ , which can be understood as three different strong charges. In this Section, I follow the Yang-Mills theory [24] and define the QCD Lagrangian.

Nowadays, the quark-quark strong interaction is understood as an  $SU(3)$  gauge theory in a degree of freedom called color. The generators of  $SU(3)$

are derived from Gell-Mann matrices (1.2) and act on quark color triplet wave functions.

$$\psi(x) = \begin{pmatrix} \psi_r(x) \\ \psi_g(x) \\ \psi_b(x) \end{pmatrix}. \quad (1.21)$$

Following the Yang-Mills theory, to each generator  $\frac{\lambda^a}{2}$  gluon field  $A_\mu^a(x)$  and gluon field strength tensor

$$F_{\mu\nu}^a = \left( \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g f^{abc} A_\mu^b A_\nu^c \right) \quad (1.22)$$

are assigned where  $g$  denotes the coupling constant of strong interaction and  $f^{abc}$  structure constant defined in Section 1.1. QCD Lagrangian

$$\mathcal{L}_{\text{QCD}} = \bar{\psi} \left( -i\partial_\mu + g \frac{\lambda}{2} A_\mu^a(x) \right) \gamma^\mu \psi - \frac{1}{4} F_{\mu\nu}^a F_a^{\mu\nu} \quad (1.23)$$

is invariant under local transformation

$$\begin{aligned} \psi(x) &\rightarrow \psi'(x) = e^{ig\Theta(x)} \psi(x), \\ A_\mu(x) &\rightarrow e^{ig\Theta(x)} \left( A_\mu(x) + \frac{i}{g} \partial_\mu \right) e^{-ig\Theta(x)}, \end{aligned} \quad (1.24)$$

where

$$\Theta(x) = \frac{1}{2} \lambda^a \Theta^a(x) \quad , \quad A_\mu(x) = \frac{1}{2} \lambda^a A_\mu^a(x). \quad (1.25)$$

There is no mass term in Lagrangian (1.23) because mass term  $m\bar{\psi}\psi$  vary under gauge transformation (1.24). To include quark mass term in QCD Lagrangian, the Higgs mechanism [25], which is explained in [9] in detail, has to be used.

QCD Lagrangian (1.23) together with gauge transformations (1.24) are sufficient for determination of Feynman rules - the key ingredient in the perturbative QCD which we will discuss in the next Section.

By derivation of a gluon propagator, the gauge-fixing term has to be added to the QCD Lagrangian.

$$\mathcal{L}_{\text{QCD}}^{\text{gauge-fixing}} = -\frac{1}{2\xi} (\partial_\mu A_a^\mu)^2. \quad (1.26)$$

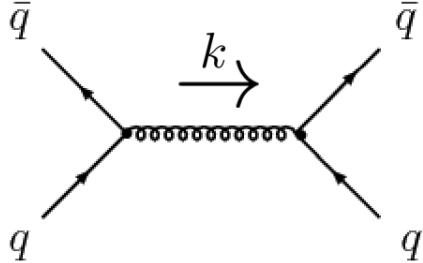


Figure 1.6: Leading order Feynmann diagrams in the scattering reaction  $q\bar{q} \rightarrow q\bar{q}$  with denoted transferred momentum  $k$ .

This term confines the possible gauges to one class parametrized by real parameter  $\xi$ . In non-Abelian gauge theories this term must be supplemented by the so called ghost term which brings into the theory new unphysical scalar particle obeying fermionic statistic. More details on the Faddev-Popov ghost field can be found in [26].

## 1.4 Perturbative QCD

The Quantum Electrodynamics and The QCD are both quantum filed gauge theories, but they differ in one killing feature - the former is Abelian whereas the latter is not. The non-Abelian character of the QCD leads to new phenomenaons of triple and quartic gluonic interactions, which have an origin in the QCD Lagrangian (1.23). In this Section, I discuss one remarkable consequence - the running coupling constant.

We start with the scattering process

$$q\bar{q} \rightarrow q\bar{q}, \quad (1.27)$$

which is depicted in the lowest order of perturbation theory by the Feynman graph in Figure 1.6. Except contribution of this graph to the scattering amplitude (which is the only contribution  $\sim g^2$ ) there are 12 other Feynman diagrams with contributions  $\sim g^4$ . These are depicted in Figure 1.7.

The contributions from all the Feynman diagrams are calculated in [8] in detail. There is shown, that all these contributions together are logarithmically divergent. This divergence can be removed, when from the scattering amplitude for arbitrary momentum transfer  $k^2$  scattering amplitude for fixed momentum transfer  $k^2 = -M^2$  is subtracted. This is how the renormalized coupling constant  $g_R$  is obtained and here is its final expression

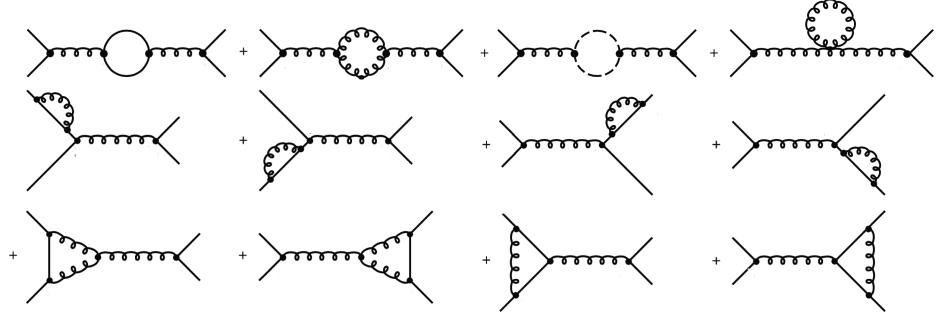


Figure 1.7: Perturbative corrections to the Feynmann diagram from Figure 1.6 representing the scattering reaction  $q\bar{q} \rightarrow q\bar{q}$ . Dashed line represents scalar ghost particle.

$$g_R = g_0 - \frac{g_0^3}{16\pi^2} \left( \frac{11}{2} - \frac{1}{3} N_F \right) \ln \left( \frac{-k^2}{M^2} \right) + \mathcal{O}(g_0^5). \quad (1.28)$$

Here  $g_0$  stands for the coupling constant measured at the renormalization scale  $k^2 = -M^2$  and  $N_F$  for the number of different quark flavors with mass  $m^2 \ll |k^2|$ . Dependence of  $g_R$  on transferred momentum  $k^2$  is evident, but there are another two intertwined dependences - on renormalization scale  $M$  and on coupling constant at renormalization scale  $g_0 = g_R|_{k^2=-M^2}$ . For next purpose, it is convenient to use the dependence schema

$$g_R = g_R(-k^2, g_0(M)) \quad (1.29)$$

which allows us to use the advantages of  $\beta$ -function. With the usage of the equation (1.28), the differential equation for  $g_0(M)$  can be obtained

$$\beta(g_0) \equiv M \left( \frac{\partial g_R}{\partial M} \right)_{-k^2=M^2} = M \left( \frac{dg_0}{dM} \right)_{-k^2=M^2} \quad (1.30)$$

$$= -b_0 g_0^3 + \mathcal{O}(g_0^5), \quad b_0 = \frac{1}{16\pi^2} \left( 11 - \frac{2N_F}{3} \right), \quad (1.31)$$

which can be solved directly to obtain coupling constant  $g_0$  for arbitrary scale  $-k^2$

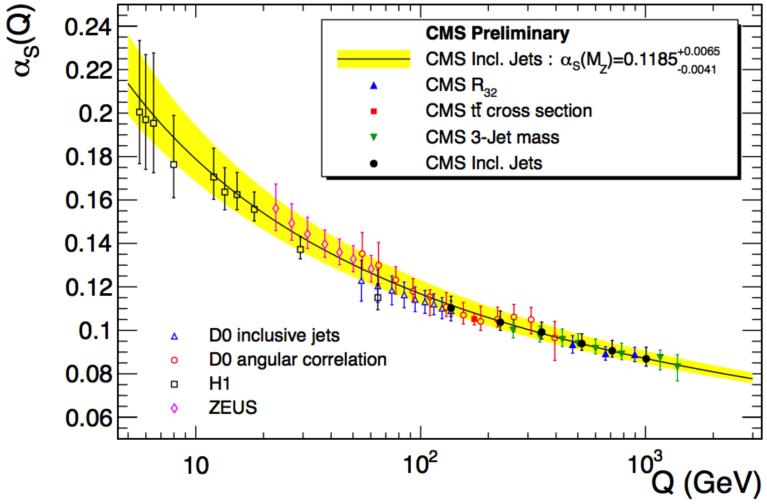


Figure 1.8: Experimental measurements of running coupling constant  $\alpha_S(Q)$  (solid line) and its uncertainty (yellow band).  $Q = \sqrt{|k^2|}$  in comparison to (1.33). Figure from [28].

$$\int_{g_0(M^2)}^{g_0(-k^2)} \frac{dg_0}{g_0^3} = -b_0 \int_{M^2}^{-k^2} \frac{dM}{M} \quad (1.32)$$

with solution

$$\alpha_S(-k^2) = \frac{\alpha_S(M^2)}{1 + \frac{\alpha_S(M^2)}{4\pi} \left( 11 - \frac{2N_F}{3} \right) \ln \left( \frac{-k^2}{M^2} \right)}, \quad g_0^2(-k^2) = 4\pi\alpha_S(-k^2), \quad (1.33)$$

which is the final expression for running coupling constant up to one-loop order. This dependence corresponds to experimental data which are depicted in Figure 1.8. Coupling constant decreases with increasing momentum transfer allowing the use of the perturbation theory. This is known as the Asymptotic Freedom [27].

On the other hand, when the momentum transfer decreases, there is a special value  $-k^2 = \Lambda^2$  for which the last expression diverges

$$-1 = \frac{\alpha_S(M^2)}{4\pi} \left( 11 - \frac{2N_F}{3} \right) \ln \left( \frac{\Lambda^2}{M^2} \right). \quad (1.34)$$

Experimental value is  $\Lambda = 213^{+38}_{-35}$  MeV [29] and demonstrates, that perturbative QCD cannot be used at low energy transfers. In fact, the run-

ning coupling constant  $\alpha_S(-k^2)$  reaches value  $\sim 1$  on momenta transfers  $\sqrt{|k^2|} \sim 500$  MeV.

The behavior of coupling constant at low energy transfers is not explainable in the language of the perturbative QCD just presented. It is non-perturbative effect known as the Principle of Color Confinement, which states, that quarks, when separate, the gluon force field between them becomes stronger instead of diminishing. The accumulated energy is consumed by the creation of quark anti-quark pair, until there is no free color charge left. This principle forbids us from observing free quarks.

To understand e.g. structure of proton with rest mass  $< 1$  GeV it is clear the non-perturbative QCD has to be used. Basic ideas of the non-perturbative QCD are introduced in the next Section.

## 1.5 Non-Perturbative QCD

The most well established non-perturbative approach to the QCD is a Lattice QCD. In this Section, I discuss the basic features of Lattice QCD. More information on this extended topic can be found in [8,30].

Lattice QCD is the QCD formulated on a hypercubic equally spaced lattice in space and time with lattice parameter  $a$  denoting the distance between neighboring sites. Quark fields are placed on sites whereas the gluon fields sit on the links between neighboring sites. From the QCD, Lattice QCD inherits the gauge invariance, which has to be formulated on the lattice structure. For  $a \rightarrow 0$  the Lattice QCD action coincides with that of QCD. The Lattice QCD contains 6 parameters - strong coupling constant and masses of 5 quarks (the top quark with lifetime  $\sim 10^{-24}$  s is not assumed by the theory).

Unlike the perturbative expansions used in perturbative QCD, the Lattice QCD uses a numerical evaluation of a path integral to perform non-perturbative calculations. Lattice QCD calculations are limited by the availability of computational resources and the efficiency of algorithms. The Lattice QCD suffers with both statistical and systematical errors, the former arising from the use of Monte-Carlo integration, the latter, e.g. from the use of non-zero values of  $a$ .

The current Lattice QCD calculations are made on supercomputers like the QCDCQ supercomputer [31] with peak speed of 500 TFlops using lattice

spacing  $a \sim 0.05 - 0.15$  fm in lattice volume  $V \sim (2 - 6 \text{ fm})^3$ .

The Importance of the Lattice QCD lies in its ability to predict masses of observed mesons and baryons, including quark masses itself, and in investigation of topological structure of a QCD vacuum. The Lattice QCD can be used to obtain Parton Distribution Functions (1.13) helping us to understand the structure of hadrons. Phenomenology of Lattice QCD also explains the Principle of Color Confinement.



## Chapter 2

# Experimental Framework

*What we observe is not nature itself, but nature exposed to our method of questioning.*

Werner Heisenberg

In the previous chapter, the key features of QCD - today's theory of strong interaction - were introduced. The predictions of QCD are tested at particle accelerators persistently with no signs for new physics so far. Large Hadron Collider which will open energy regions not observed yet can change this very soon.

Jets are the most important objects observed by inelastic collisions on hadron colliders, which allows the QCD predictions to be confronted with the experiment. In this chapter, after the introduction of the Large Hadron Collider and the ATLAS detector, jet algorithms and the way how jets are reconstructed on the ATLAS detector will be described.

### 2.1 The Large Hadron Collider and The ATLAS Detector

#### 2.1.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [32,33] is a charged particle accelerator located on the border of France and Switzerland, near Geneva at CERN (the

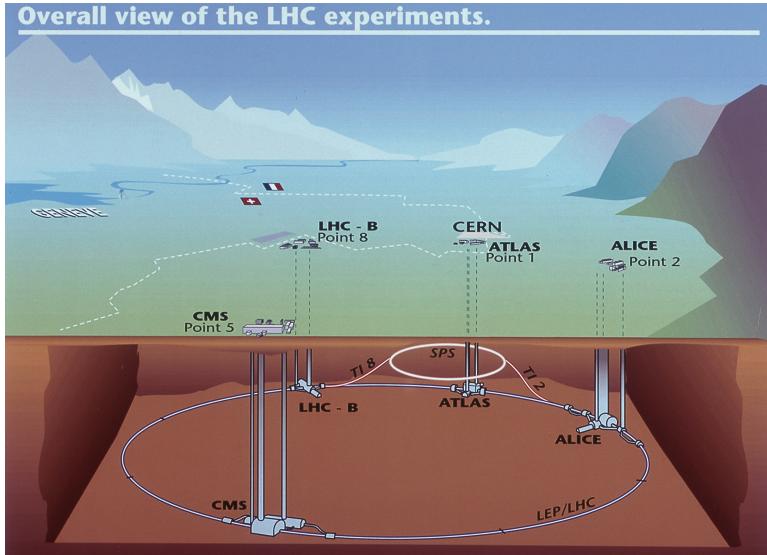


Figure 2.1: This diagram shows the locations of the four main experiments (ALICE, ATLAS, CMS and LHCb) that take place at the LHC. Located between 50 m and 150 m underground, huge caverns have been excavated to house the giant detectors. The Super Proton Synchotron (SPS), the final link in the pre-acceleration chain, and its connection tunnels to the LHC are also shown. Figure from [34].

European Organization for Nuclear Research) in Switzerland. Built using the areas of the Large Electron-Positron collider, the main accelerator ring, of a 27 km circumference, is located around 100 m below the surface. There are four main experiments located around the ring: A Large Ion Collider Experiment (ALICE), A Toroidal LHC ApparatuS (ATLAS), Compact Muon Solenoid (CMS) and Large Hadron Collider beauty (LHCb). The complete accelerator and detector system is shown in Figure 2.1.

LHC started to operate on November 23, 2009 and soon thereafter (March 30, 2010) the proton-proton collisions achieved the center-of-mass energy  $\sqrt{s} = 7 \text{ TeV}$ , which is a half of the design energy of the machine. On April 5, 2012, the machine started its successful  $\sqrt{s} = 8 \text{ TeV}$  run.

Next to the proton-proton collisions, first heavy-ion Pb-Pb collisions took place in 2010 at a center of mass energy per pair of colliding nucleons  $\sqrt{s} = 2.76 \text{ TeV}$ . Proton-Pb collisions at  $\sqrt{s} = 5.02 \text{ TeV}$  occurring on LHC during 3 weeks of 2013 successfully demonstrated LHC capability to provide asymmetric collisions.

The first running period of the LHC, Run I, was very successful and re-

sulted in the discovery of the Higgs boson on July 4, 2012 [4]. The accelerator complex including its experiments has been upgraded for two years and the Run II is expected to start in early summer 2015 [35, 36]. In Run II the center-of-mass energy of proton-proton collisions will be raised to  $\sqrt{s} = 13 \text{ TeV}$  and the beam crossing time could be reduced from the current 50 ns to 25 ns. The expected integrated luminosity is expected to be  $\sim 100 \text{ fb}^{-1}$  after three years of data collecting.

### 2.1.2 The ATLAS Detector

The ATLAS detector [37] is a general-purpose detector surrounding one of the interaction points of the LHC and with  $\sim 100$  million of individual electronic channels it is the most complicated instrument ever created. The purpose of the ATLAS detector is to record particle collisions up to the center-of-mass energy per pair of colliding nucleons  $\sqrt{s} = 14 \text{ TeV}$ . A detector overview is shown in Figure 2.2a, where the main sub-detector systems can be seen: the inner detector, used to reconstruct charged-particle tracks, the electromagnetic calorimeters, the hadronic calorimeters, and the muon spectrometer.

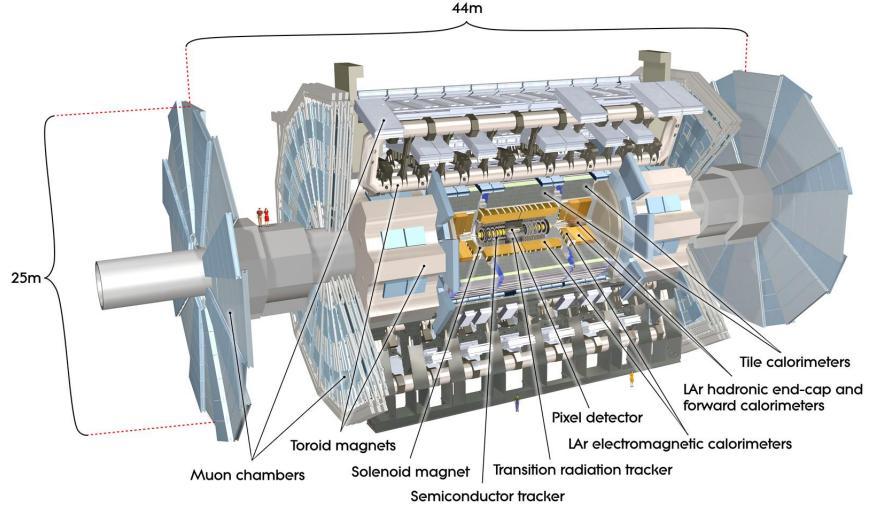
ATLAS uses a right-handed coordinate system with its origin at the interaction point in the center of the detector and the  $z$  axis along the beam pipe. The  $x$  axis points from the interaction point to the center of the LHC ring, and the  $y$  axis points upward. Cylindrical coordinates  $(r, \phi)$  are used in the transverse plane,  $\phi$  being the azimuthal angle around the beam pipe. Instead of polar angle  $\theta$ , pseudorapidity  $\eta$  is often used. In this thesis the rapidity  $y$  is used as the polar angle coordinate. In following definitions of pseudorapidity  $\eta$  and rapidity  $y$ ,  $E$  stands for the total energy and  $p$  for size of total momentum:

$$\eta = -\frac{1}{2} \ln \left( \frac{p + p_z}{p - p_z} \right) = -\ln \left[ \tan \left( \frac{\theta}{2} \right) \right], \quad (2.1)$$

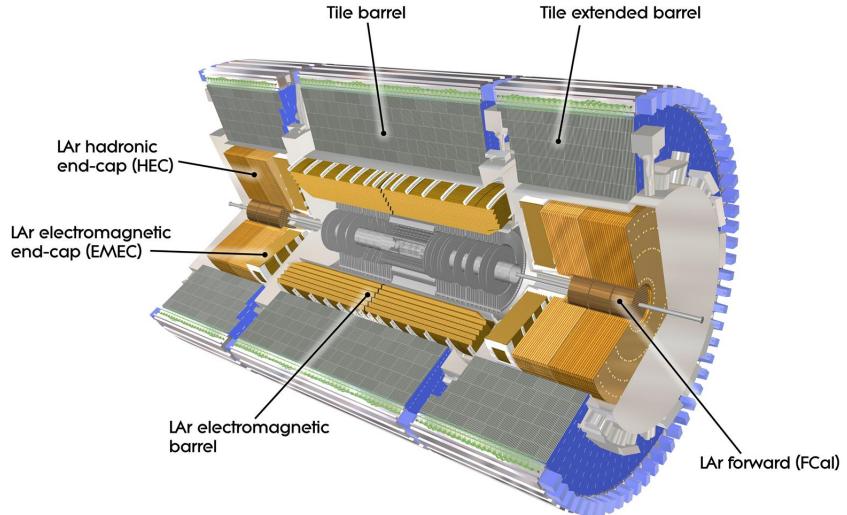
$$y = -\frac{1}{2} \ln \left( \frac{E + p_z}{E - p_z} \right). \quad (2.2)$$

The transverse momentum  $p_T = \sqrt{p_x^2 + p_y^2}$  presents the component of momentum perpendicular to the beam line.

The main detector system relevant to this thesis is the ATLAS calorimeter, which is emphasized in Figure 2.2b. The calorimeter is divided into



(a) ATLAS detector.



(b) Inner detector and calorimeter systems.

Figure 2.2: (a) an overview of the ATLAS detector (b) detail on the inner detector and the calorimeters - the dominant sub-detector systems used in this thesis. Figures taken from [38].

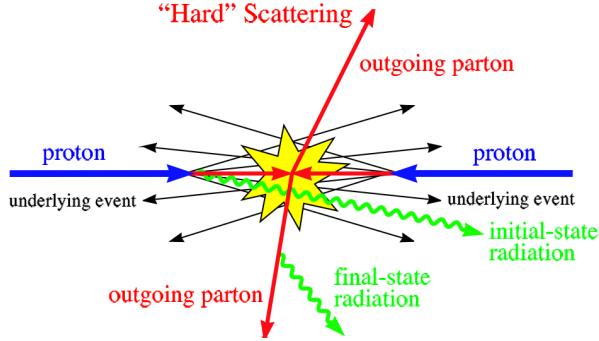


Figure 2.3: Schematic representation of a hard scattering proton-proton collision. Figure taken from [39].

sub-detectors, providing overall coverage up to  $|\eta| < 4.9$ . The electromagnetic calorimeter, covering region  $|\eta| < 3.2$ , is a high-granularity sampling detector in which the liquid argon (LAr) active medium is interspersed with layers of lead absorber. The hadronic calorimeters are divided into three sections: a tile scintillator/steel calorimeter is used in both the barrel ( $|\eta| < 1.0$ ) and extended barrel cylinders ( $0.8 < |\eta| < 1.7$ ) while the hadronic endcap ( $1.5 < |\eta| < 3.2$ ) consists of LAr/copper calorimeter modules. The forward calorimeter measures both electromagnetic and hadronic energy in the range  $3.2 < |\eta| < 4.9$  using LAr/copper and LAr/tungsten modules.

## 2.2 Hadron Collision at LHC

In this section the phenomenological description of proton-proton collisions will be presented following Figure 2.3 and Reference about Monte Carlo event generators [23].

Two incoming protons can be understood as two bags of partons. The collision is dominated by the strong interaction of two partons - one from each of the colliding hadrons. These partons are called incoming partons and the momentum transfer of their interaction is  $Q \gg \Lambda$ , so the perturbative QCD can be used to describe the process of hard scattering. The remaining energy is carried by the remaining partons, which create the so called underlying event - particles, which do not come from the hard QCD processes.

When partons are sufficiently far from each other, the non-perturbative QCD has to be used to describe the process of hadronisation, in which a

set of colored partons is transformed into a set of colorless primary hadrons which may then decay further.

During the collision, the color charges of partons interact resulting in radiation of gluons  $q \rightarrow qg$ . This process is described by perturbative QCD and leads to infrared and collinear divergences. However, infrared divergences are canceled by Kinoshita–Lee–Nauenberg theorem [40, 41], so only collinear divergences remain. There is no mechanism known up to date, which would solve the problem with collinear divergences. However, observables inclusive enough to be insensitive to processes that distinguish between different numbers of partons are not affected by infrared divergences. There is no possibility how to theoretically predict the energy of hardest outgoing particle, but it is possible to predict the energy flow in a cone from the point of scattering.

This is where the term jet comes to play. A jet can be naively seen as a group of collimated particles generated by the hadronisation of a parton in the scattering process and is the most important object used on hadron colliders for analysis of QCD processes.

### 2.3 Jet Algorithms

Jet algorithm is a generic "recipe" which takes a set of particles (or other objects with defined four-momenta) and returns jets created from them. The jet algorithm usually involves a set of parameters which together with the algorithm fully specify the jet definition. According to the remarks at the end of the previous section, jet algorithms should fulfill the following conditions

1. Infrared safety - the presence of additional soft particles should not affect the recombination of these particles into a jet.
2. Collinear safety - jet reconstruction should not depend on the fact, if the transverse momentum is carried by one particle, or if the particle is split into two collinear particles.

Two important steps must be defined in each jet algorithm

1. Clustering - description how the input objects are clustered into jets.
2. Recombination - determination of physical quantities of jets.

Additional steps may include the preclustering reducing the number of input objects for jet algorithm.

Two classes of jet algorithms are described here - cone algorithms and  $k_t$  algorithms. First of these algorithms is more illustrative, the second one is used in ATLAS. These algorithms use two different recombination schemes which description will follow. Detailed description as well as other jet algorithms can be found in [42, 43]. After the definition of jet algorithms a short description follows, how the objects (not necessary particles) with defined four-momenta are constructed from the signal observed on the ATLAS detector.

### 2.3.1 Cone algorithms

The first step of these algorithms is to order all input objects (reconstructed detector objects with four-momentum representation) in decreasing order of transverse momentum  $p_T$ . If the object with the highest  $p_T$  is above the seed threshold, all objects within a cone in rapidity  $y$  and azimuth  $\phi$  with  $\Delta R = \sqrt{\Delta y^2 + \Delta \phi^2} < R_{cone}$ , where  $R_{cone}$  is the fixed cone radius, are recombined using Snowmass recombination scheme (see Section 2.3.3). A new cone is centered around a new direction and the objects in the new cone are recombined and again the direction is updated. This process continues until the direction of the cone does not change anymore after recombination, at which point the cone is considered stable and is called a proto-jet.

At this point, the next seed is taken from the input list and a new proto-jet is formed with the same iterative procedure. This continues until no more seeds are available.

The proto-jets found by this procedure can share some constituents. Constituents shared between two proto-jets are recombined into new proto-jet and if the ratio  $E_T^{shared}/\min(E_T^{neighbor}) > f$  is over the certain threshold for example  $f = 0.5$  the neighboring proto-jets are recombined into one proto-jet (shared constituents are taken only once). If this condition is not satisfied, the shared constituents are assigned to the nearest proto-jet. When proto-jet does not share constituents it is recombined into the jet.

This algorithm is both not infrared safe (Figure 2.4a) and not collinear safe (Figure 2.4b). The infrared insensitivity can be improved by adding the midpoints between pairs of proto-jets fulfilling  $R_{cone} < \Delta R < 2R_{cone}$  and repeating the iterative procedure with midpoints being new seeds. Since the collinear unsafety arises from the use of seed towers, Seedless cone algorithm was developed, which searches the entire detector to find all stable proto-jets.

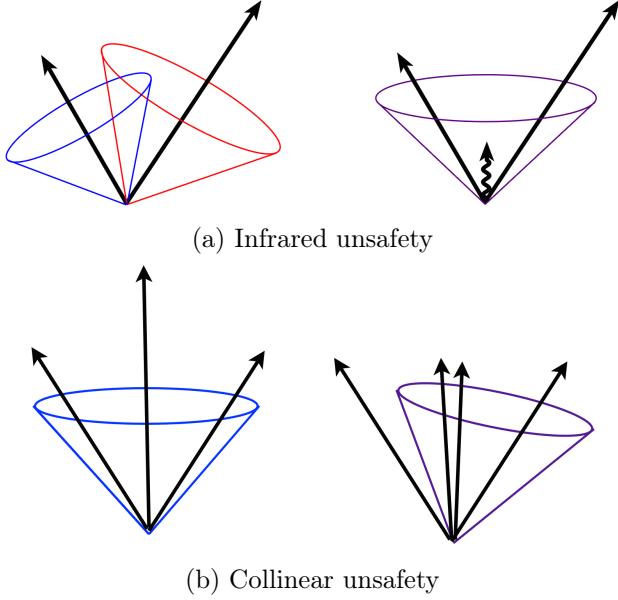


Figure 2.4: Illustration of (a) infrared unsafety and (b) collinear unsafety of fixed cone jet algorithm. Figures from [44].

Typical parameters used by fixed cone algorithm are a seed threshold of  $p_T > 1 \text{ GeV}$ , and a narrow ( $R_{cone} = 0.4$ ) or a wide cone jet ( $R_{cone} = 0.7$ ) option.

### 2.3.2 $k_t$ algorithms

In this class of algorithms all pairs  $(i, j)$  of input objects are analyzed with respect to their relative transverse momentum squared, defined by

$$d_{ij} = \min \left( p_{T,i}^{2p}, p_{T,j}^{2p} \right) \frac{\Delta R_{ij}^2}{R^2} \quad (2.3)$$

and the squared  $p_T$  of object  $i$  relative to the beam axis

$$d_i = p_{T,i}^{2p}. \quad (2.4)$$

Here  $\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$  and  $p_{T,i}$ ,  $y_i$  and  $\phi_i$  are respectively the transverse momentum, rapidity and azimuth of particle  $i$ . In addition to the radius parameter  $R$ , parameter  $p$  was added to split  $k_t$  algorithms into three categories.

- $p = 1$   $k_t$  algorithm,

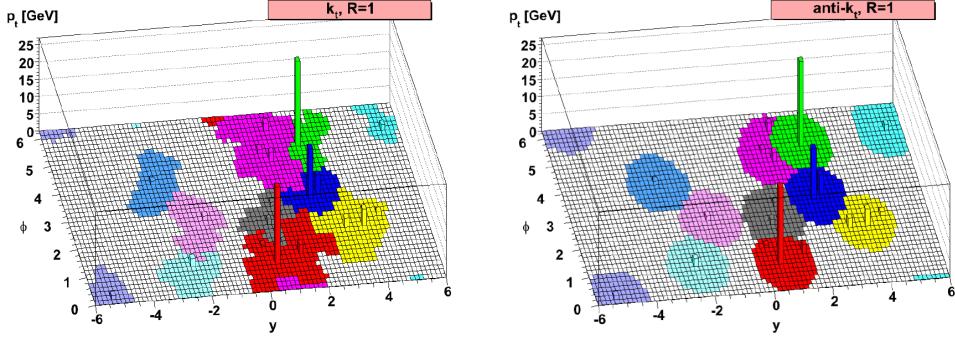


Figure 2.5: Illustration of  $k_t$  and anti- $k_t$  jet algorithms with  $R = 1$  for calorimeter signal towers in azimuth  $\Phi$  and pseudorapidity  $y$ . Towers of the same color were recombined to one jet. Figure taken from [44].

- $p = 0$  Cambridge/Aachen algorithm,
- $p = -1$  anti- $k_t$  jet-clustering algorithm.

These algorithms first find the minimum  $d_{min}$  of all  $d_{ij}$  and  $d_i$ . If  $d_{min}$  is in  $d_{ij}$ 's, the corresponding objects  $i$  and  $j$  are combined into a new object  $k$  using four-momentum recombination. Both objects  $i$  and  $j$  are removed from the list, and the new object  $k$  is added to it. If  $d_{min}$  is in  $d_i$ 's, the object  $i$  is considered to be a jet by itself and is removed from the list.

This means that all original input objects end up to be either part of a jet or to be jets by themselves. Contrary to the cone algorithm described earlier, no objects are shared between jets and the procedure is both infrared and collinear safe.

ATLAS uses anti- $k_t$  algorithm with  $R = 0.4$  for narrow and  $R = 0.6$  for wide jets. The differences between  $k_t$ -algorithms are detailed described in [45]. Recombination of calorimeter signal towers (see Section 2.3.4) in jets is for  $k_t$  and anti- $k_t$  algorithms shown at Figure 2.5.

### 2.3.3 Recombination

Let  $J$  be the index set of the input objects with the defined four-momenta  $(E^i, p_x^i, p_y^i, p_z^i)$ ,  $i \in J$  which has to be recombined into the jet with new kinematic quantities  $E^J, \mathbf{p}^J, p_T^J, y^J, \dots$  Possible recombination schemes are

- **Snowmass Scheme**

Used by fixed cone algorithm when finding proto-jets.

$$E_T^J = \sum_{i \in J} E_T^i \quad , \quad \eta^J = \frac{1}{E_T^J} \sum_{i \in J} E_T^i \eta^i \quad , \quad \phi^J = \frac{1}{E_T^J} \sum_{i \in J} E_T^i \phi^i. \quad (2.5)$$

- **Four-Momentum Recombination (E-Scheme)**

Used by  $k_t$ -algorithms and by fixed cone algorithm to final recombination of proto-jets into jets.

$$p^J = (E^J, \mathbf{p}^J) = \sum_{i \in J} (E^i, p_x^i, p_y^i, p_z^i), \quad (2.6)$$

$$p_T^J = \sqrt{(p_x^J)^2 + (p_y^J)^2} \quad , \quad y^J = \frac{1}{2} \ln \frac{E^J + p_z^J}{E^J - p_z^J} \quad , \quad \phi^J = \tan^{-1} \frac{p_y^J}{p_x^J}. \quad (2.7)$$

### 2.3.4 Calorimeter jets

The most important detectors for the jet reconstruction are the ATLAS calorimeters. The ATLAS calorimeter system has about 200,000 individual cells of various sizes and with different readout technologies and cell geometries. For jet finding it is necessary to first combine these cell signals into larger signal object with physically meaningful four-momenta. The two concepts available are the calorimeter signal towers and the topological cell clusters.

In the case of calorimeter signal towers, the cells are projected onto a fixed grid in pseudorapidity  $\eta$  and azimuth  $\phi$ . The tower bin size is  $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$  in the whole acceptance region of the calorimeters, i.e. in  $|\eta| < 5$  and  $-\pi < \phi < \pi$  with approximately  $100 \times 64 = 6,400$  towers in total.

The alternative representation of the calorimeter signals for jet reconstruction are topological cell clusters, which are basically an attempt to reconstruct three-dimensional "energy blobs" representing the showers developing for each particle entering the calorimeter. The clustering starts with seed cells with a signal-to-noise ratio, or signal significance  $\Gamma = E_{cell}/\sigma_{noise,cell}$ , above a certain threshold  $S$ , i.e.  $|\Gamma| > S = 4$ . All directly neighboring cells of these seed cells, in all three dimensions, are collected into the cluster. Neighbors of neighbors are considered for those added cells which have  $\Gamma$  above a certain secondary threshold  $N$  ( $|\Gamma| > N = 2$ ). Finally, a ring of guard cells with signal significances above a basic threshold  $|\Gamma| > P = 0$  is

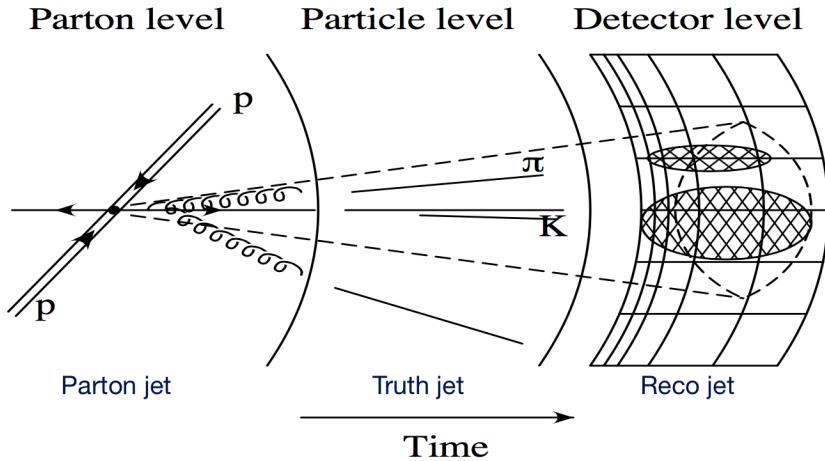


Figure 2.6: Three levels of jet reconstruction. Figure from [49]

added to the cluster. After the initial clusters are formed, they are analyzed for local signal maxima by a splitting algorithm, and split between those maxima.

## 2.4 Jet corrections

Before jets can proceed to the data analysis, corrections have to be applied to minimize detector effects including calorimeter non-compensation, noise, losses in dead material and cracks, longitudinal leakage and particle deflection in the magnetic field. Indispensable tool for jet corrections are Monte Carlo event generators - PYTHIA8 [46] generating high-energy-physics events and GEANT4 [47] or ATLASFASTII [48] detector simulations for simulating the ATLAS detector response of PYTHIA8 generated events.

Using these tools it is possible to reconstruct jets from Monte Carlo events on three different stages of collision indicated in Figure 2.6. First there are parton jets which are reconstructed from the quarks, gluons and other elementary particles created just after the collision. Stable particles (with lifetime  $c\tau \sim 10^{-15}$  m) created by hadronization are recombined into the truth jets. When collision reaches the detector, the detector simulation is used and the recorded signal is reconstructed into reco jets.

First, the reco jets are corrected to the truth jets leading to modification of kinematic properties of individual reco jet in the process called jet energy scale calibration.

### 2.4.1 Jet Energy Scale Calibration

Energy  $E_{reco}$  of the jet measured by the detector may differ from the energy  $E_{truth}$  of the corresponding particle jet. The goal of the jet energy scale calibration is to remove some detector effects and correct  $E_{reco}$  to  $E_{truth}$ . The detector effects can be summarized by the formula

$$E_{truth} = \frac{E_{reco} - O}{R \cdot S}, \quad (2.8)$$

where  $O$  is the offset representing subtraction of additional energy, which is represented by the detector noise and pile-up with contributions from other  $pp$  collisions occurring during beam crossing. Hadronic character of jets observed at LHC is the reason why the response  $R$  is the largest correction. Showering factor  $S$  describes the particle flow out/from jet recombination cells. More concise information about the parameters just introduced can be found in [49].

Because the calibration is persistently evolving, each jet analysis uses as the input the uncalibrated reco jets which are then easily calibrated using standard library `APPLYJETCALIBRATION` [50].

However, there are some detector effects which can't be fixed by the calibration. These effects include the limited detector resolution (detector cells have finite dimensions) and the limited acceptance (not all events are recorded). The former leads to the smearing of jet kinematic properties whereas the latter to decrease of observed cross section against the value theoretically predicted. Both of these effects are negatively affecting the observables and can be partially removed by the unfolding procedure, which unlike the jet calibration, is analysis dependent.

### 2.4.2 Unfolding

In this analysis, the distribution  $f(p_T)$  of inclusive jet  $p_T$  is measured for  $p_T \in \langle a, b \rangle$ . Due to the detector imperfections, instead of physical variable  $p_T$  new variable  $x$  and its distribution  $g(x)$  are measured. New distribution can be expressed as

$$g(x) = \int_a^b A(x, p_T) f(p_T) dp_T, \quad (2.9)$$

with the function  $A(x, p_T)$  describing the detector response as can be seen when the detector is exposed to a particle beam with well known  $p_T = p'_T$  meaning  $f(p_T) = \delta(p_T - p'_T)$ , leading to  $g(x) = A(x, p'_T)$ . The reconstruction of  $f(p_T)$  from measured  $g(x)$  is called unfolding.

For practical purposes the equation (2.9) should be discretized so instead of continuous distribution  $g(x)$  the discretized values  $g_i = \int_{N(i)} g(x)dx$  of discretized observable  $f_i = \int_{N(i)} f(p_T)dp_T$  are measured, where the integration is done over measurable  $N(i) \subset \langle a, b \rangle$ . For simplicity assume  $x \in \langle a, b \rangle$  is discretized in the same way as the physical  $p_T$ . Equation (2.9) then reads

$$g = Af, \quad (2.10)$$

with  $g$  and  $f$  being vectors of  $g_i$ 's and  $f_i$ 's respectively and  $A$  matrix derived from  $A(x, p_T)$ , later in Section 3.3 called the transfer matrix. If the limited acceptance would be the only detector problem, then  $A$  would be diagonal matrix with some elements  $< 1$ . When the limited resolution comes to play, the diagonal entries start to smear out of the diagonal and the matrix  $A$  starts to complicate.

The unfolding result which offers the solution of (2.10) by the inversion of matrix  $A$  is mostly disappointing as is illustrated e.g. in [51]. For result improvement, different unfolding methods were developed including Iterative Bayesian Unfolding [52], Singular Value Decomposition [53], or Iterative, Dynamically Stabilized (IDS) method [54], which is the method used in this thesis.



## Chapter 3

# Data Analysis

*In physics, you don't have to go around making trouble for yourself - nature does it for you.*

Frank Wilczek

QCD jets are the most common hard objects observed at inelastic collisions at hadron colliders, with their cross section exceeding any other physics process by orders of magnitude. Measurement of inclusive jet cross section provides the test for both QCD predictions and the detector performance up to the momentum transfers not reachable by any other physics processes.

In this Chapter, I will describe the details of the double differential inclusive jet cross section analysis, which I have done in this thesis. I will begin with the characteristics of the Monte Carlo data, I have used. These data together with the event selection criteria and matching procedure, which description will follow, fully specify the input for the unfolding procedure. Two approaches to the unfolding, which I have implemented in this thesis, will be described and compared with each other.

At the end of this Chapter, I will compare the results of my data analysis, with the next-to-leading order perturbative QCD prediction of my supervisor.

JZXW	$p_T$ range (GeV)	Cross-section (fb)	Filter Efficiency	# events
JZ0W	0 - 20	7.8420e+13	9.7193e-01	3498000
JZ1W	20 - 80	7.8420e+13	2.7903e-04	2998000
JZ2W	80 - 200	5.7312e+10	5.2261e-03	500000
JZ3W	200 - 500	1.4478e+09	1.8068e-03	499500
JZ4W	500 - 1000	2.3093e+07	1.3276e-03	477000
JZ5W	1000 - 1500	2.3793e+05	5.0449e-03	499000
JZ6W	1500 - 2000	5.4279e+03	1.3886e-02	493500
JZ7W	2000 +	9.4172e+02	6.7141e-02	497000

Table 3.1: The cross-sections (XS), filter efficiency (FE) and number of events for the JZXW samples which differ in the leading truth jet  $p_T$  range.

### 3.1 Data Characteristics

Data used in this thesis are Monte Carlo generated events of  $pp$  collisions at the center-of-mass energy  $\sqrt{s} = 13$  TeV with PYTHIA8 [46] event generator using CT10 PDFs [55] and ATLAS underlying event tune AU2 [56]. QCD calculations are done only to the leading order in PYTHIA8. The response of the ATLAS detector on these events was simulated with GEANT4 [47] software toolkit.

Particles were recombined into jets using anti- $k_t$  jet algorithm with parameter  $R = 0.4$ . There are particle jets reconstructed from the PYTHIA8 output, which further in this thesis are denoted truth jets, and next to them, there are the reco jets reconstructed from the output of GEANT4 detector simulation from the ATLAS detector topological cell clusters.

Reco jets were calibrated using the APPLYJETCALIBRATION [50] library version 3.28 with configuration parameters loaded from the JES\_Prerecommendation2015\_Feb2015 with calibration sequence JetArea\_Residual\_EtaJES. In next, reco jets denote the reco calibrated jets.

Events are generated in slices according to the leading truth jet  $p_T$ . These samples differ in event weight which is for the whole event calculated as

$$\text{weight} = \frac{(\text{XS}) \cdot (\text{FE}) \cdot w_0}{(\# \text{ events})}, \quad (3.1)$$

with XS being cross-section, FE filter efficiency and  $w_0$  additional weight factor stored in EventInfoAux container. Concrete values for datasets used in this theses are given in Table 3.1.

The analysis uses jets with transverse momentum  $p_T > 15 \text{ GeV}$  and rapidity  $|y| < 4$  and is done in double binning in  $p_T$  and  $|y|$  with the following edges which are the same as in the analysis from 2011/2012 [57].

$$\begin{aligned}
p_T = & 15 - 20 - 25 - 35 - 45 - 55 - 70 - 85 - 100 - 116 - 134 - 152 - \\
& 172 - 194 - 216 - 240 - 264 - 290 - 318 - 346 - 376 - 408 - \\
& 442 - 478 - 516 - 556 - 598 - 642 - 688 - 736 - 786 - 838 - \\
& 894 - 952 - 1012 - 1076 - 1162 - 1310 - 1530 - 1992 - 2300 - \\
& 2800 - 3400 - 4100 - 5000 - 6000 - 7200 \text{ GeV} \\
|y| = & 0.0 - 0.5 - 1.0 - 1.5 - 2.0 - 2.5 - 3.0 - 3.5 - 4.0
\end{aligned} \tag{3.2}$$

The edges in  $p_T$  are chosen to uncertainty in each bin was smaller than the differences in values in neighboring bins.

## 3.2 Event Selection

In this section the jet selection criteria and matching of truth with reco jets are described. The former is needed to cut those jets (or those events) off, which were misinterpreted by the detector, by the later the inputs for the unfolding procedure are obtained. Description of the unfolding procedure will follow in the next section. More details including graphical display and numerical results for procedures described in this section are given in Appendix A.

### 3.2.1 Jet Cuts

- **$p_T$  Cut**

Reco and truth jets with  $p_T > 15 \text{ GeV}$  were kept.

- **$y$  Cut**

Reco and truth jets with  $|y| < 4$  were kept.

- **Zero jet (0-jet) Cut**

Only those events which has at least one reco or one truth jet are considered.

- **Leading Ration (LR) Cut**

In this cut the reco and truth jets with the highest  $p_T$  were used. If there was only one reco jet left, the ratio  $LR = p_T^{reco,leading}/p_T^{truth,leading}$  was calculated. If there were two reco jets, instead of  $p_T^{reco,leading}$  the average  $p_T$  of two leading reco jets was calculated. If  $0.6 < LR < 1.4$  the event are considered.

Numbers of reco and truth jets removed in each step are shown in Table A.1, where also the cut efficiencies for individual JZXW samples are shown. The impact of each cut on jet  $p_T$  spectrum of reco and truth jets is displayed in Figure A.1.

It can be seen that the most important cut is the 0-jet cut which removes approximately 80 % of reco jets in JZ0W sample whereas the truth jets remain intact. According to Table 3.1 for event from the JZ0W sample the leading truth jet  $p_T < 20 \text{ GeV}$  which has no longer to hold for reco jets which were in some cases reconstructed with  $p_T \sim 100 \text{ GeV}$ . Because of Monte Carlo event weight of events from JZ0W sample is dominant over event weights of other JZXW samples by several orders, the misreconstructed reco jets from JZ0W sample were negatively influencing the observed  $p_T$  spectrum of reco jets as can be seen from top of the Figure A.1.

### 3.2.2 Jet Matching

To find, how the truth jets are reconstructed by the detector, the jet matching has to be done, i.e. for each truth jet it is needed to find corresponding reco jet. Jet matching used in this thesis is based on minimal angular distance between matched and its description follows in this section.

For each pair  $(i, j)$  of reco and truth jet, the quantity  $dR_{ij} = \sqrt{d\phi_{ij}^2 + dy_{ij}^2}$  was calculated with  $d\phi_{ij}$  being angle between  $\phi_i^{reco}$  and  $\phi_j^{truth}$  and  $dy_{ij} = y_i^{reco} - y_j^{truth}$ . The minimum was found between all of  $dR_{ij}$ 's. If this was smaller than the defined cutoff  $\min(dR_{ij}) = dR_{pq} < dR^{cutoff} = 0.2$ , the jets  $(p, q)$  were matched and further not assumed in matching procedure. This continued until condition  $\min(dR_{ij}) < dR^{cutoff}$  was not satisfied or all of the reco or truth jets were matched.

Numbers of reco and truth jets, both matched and unmatched are shown in Table A.1 where also the matching efficiencies for individual JZXW samples are shown. Figure A.2 shows the  $p_T$  spectra of truth and reco jets after

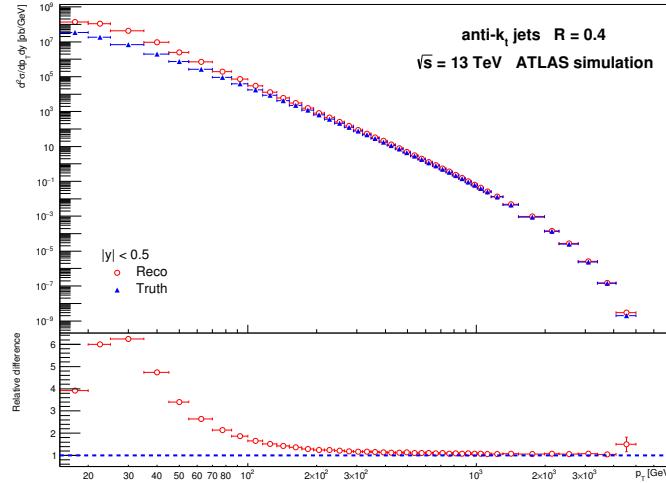


Figure 3.1: Comparison of  $p_T$  spectra of reco and truth jets after event selection for selected  $|y| < 0.5$  rapidity bin. Each  $p_T$  bin was divided by its width so  $y$ -axis has physical meaning of double differential cross section in  $p_T$  and  $y$ . Bottom graph contains the relative difference between reco and truth cross sections.

event selection, which are composed from  $p_T$  spectra of matched and unmatched jets, which are shown also.  $p_T$  spectra of truth and reco jets are for all rapidity bins assumed in this thesis shown in Figure A.3.

### 3.3 Unfolding

After the four cuts from section 3.2.1, the sets of jets denoted reco and truth jets were obtained. Matching procedure, described in section 3.2.2, divided both reco and truth jets into two categories depending on successful matching - there is correspondence 1 : 1 between matched reco and matched truth jets. Reco jets, which were not matched, formed the unmatched reco jets, and similarly a set of unmatched truth jets was created. All these 6 sets of jets are needed by the unfolding procedure which description follows in this section.

Figure 3.1 shows the  $p_T$  spectra of reco and truth jets. It can be seen, that observed  $p_T$  spectrum, represented by the reco jets, differs from the  $p_T$  spectrum theoretically expected which is represented by the  $p_T$  spectrum of truth jets. Unfolding should transform the observed  $p_T$  spectrum to the spectrum theoretically expected. If this transformation would be done on

real data, it should ideally preserve additional structures, which are presented in data, but not included by the theory.

The main ingredient for the unfolding procedure is the transfer matrix  $A_{ij}$  which contains the number of reco jets in bin  $i$  with a matched truth jet that was generated in bin  $j$  and describes thus the smearing effects of the detector. In this thesis, the double binning (3.2) is used which complicates the situation because the matched reco jet can simply migrate of the transfer matrix from Figure 3.2, when, for example, its rapidity  $|y| > 0.5$  and when it was matched with truth jet with  $|y| < 0.5$  or vice versa. Two ways of dealing with double binning are tested in this thesis.

### 1. Simple unfolding

In this case, only those reco and truth jets were used in the transfer matrix, which were matched within the same rapidity bin. Remaining matched jets were added to the unmatched jets. 8 transfer matrices  $46 \times 46$  are filled (one for each rapidity bin,  $46 = \text{number of } p_T \text{ bins}$ ) and unfolding is done for each of these matrices separately. One of these matrices for  $|y| < 0.5$  rapidity bin is shown in Figure 3.2.

### 2. 2D unfolding

In this case, the unfolding matrix was redefined to encapsulate the matching of jets between different rapidity bins. In this case only one transfer matrix  $368 \times 368$  is created ( $368 = 46 \times 8$ ) with unfolding being done only for this matrix shown at Figure B.4, from which the way how the transfer matrix was redefined from the simple unfolding should follow.

Transfer matrix from Figure B.4 used by the 2D unfolding contains at the diagonal 8 submatrices which are the unfolding matrices used by the simple unfolding. Next to these diagonal submatrices transfer matrix of 2D unfolding contains 14 additional submatrices beside diagonal. These corresponds to the matched jets with migration in rapidity bins and in case of simple unfolding, these jets are assumed to be unmatched. Appendix B.2 shows some of the slices in the transfer matrix of 2D unfolding.

Dominant elements of each of the submatrices are on the main diagonal, which corresponds to the fact, there is no significant bias in  $p_T$  reconstruction. The finite  $p_T$  resolution causes the smearing of the diagonal and finite rapidity resolution is the cause of the 14 minor submatrices. Appendix B.2

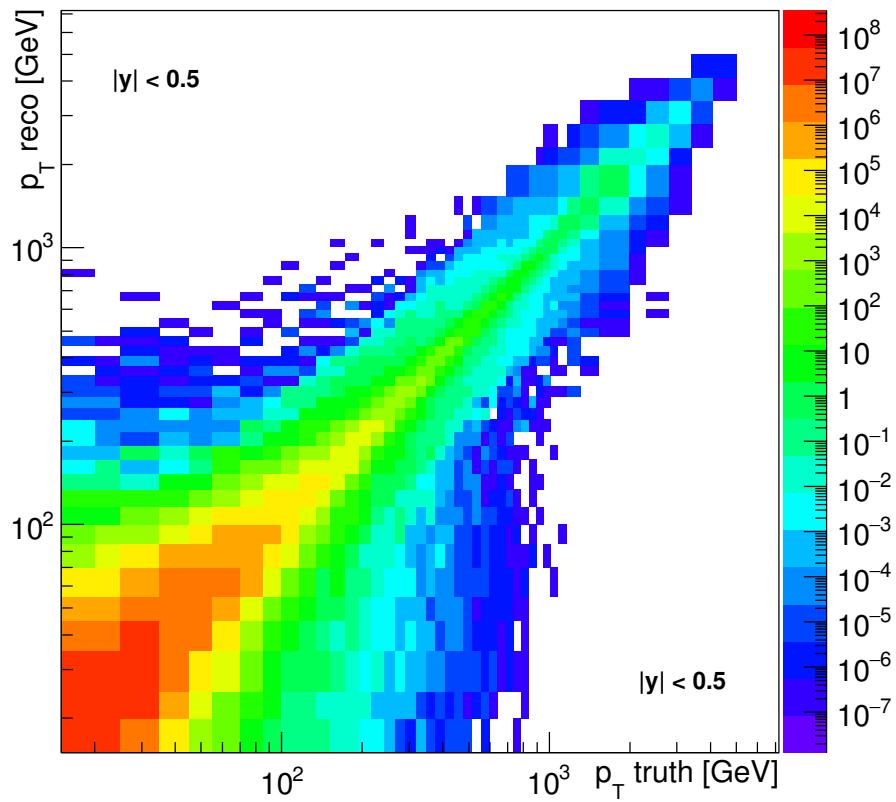


Figure 3.2: Unfolding matrix for matched reco and truth jets with rapidity  $|y| < 0.5$  corresponding to one of eight matrices used in the simple unfolding. Each cell is proportional to the number of jets with truth  $p_T$  in range determined by the  $x$ -axis which were reconstructed to the reco jets with  $p_T$  determined by the  $y$ -axis. White space signalize no input.

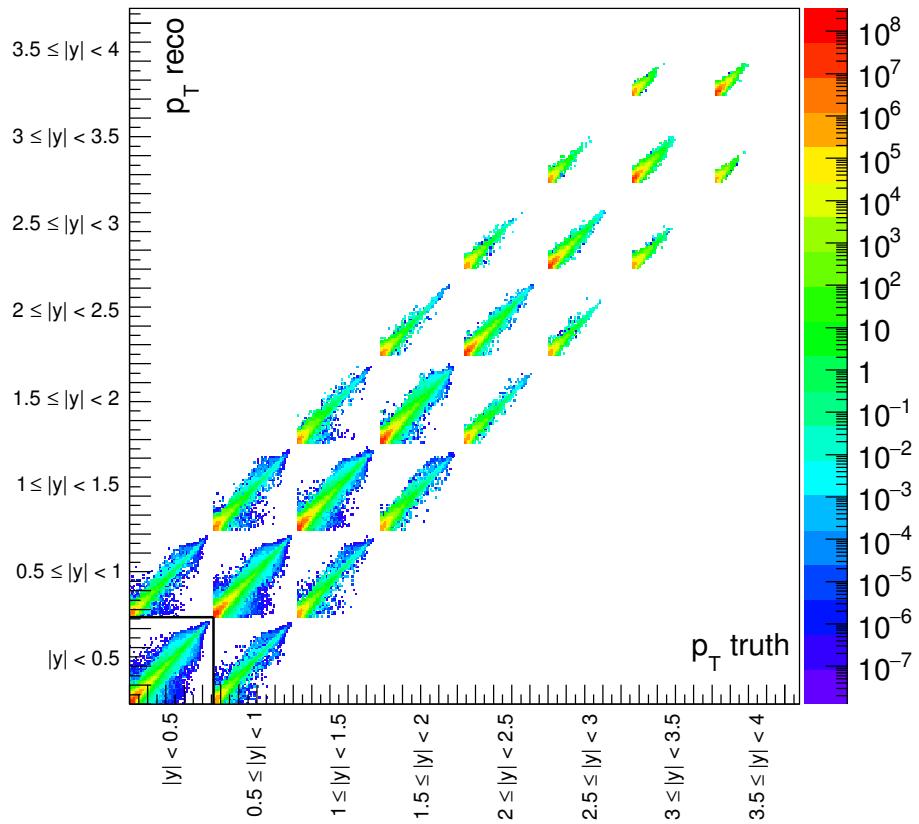


Figure 3.3: Unfolding matrix used in the 2D unfolding. Each cell is proportional to the number of jets with truth  $p_T$  and rapidity  $y$  determined by the  $x$ -axis, which were reconstructed to the reco jets with  $p_T$  and  $y$  determined by the  $y$ -axis. Marked square in  $|y| < 0.5$  region is the matrix shown in Figure 3.2. Projection of this matrix on the  $x$  and  $y$ -axis corresponds to the  $p_T$  spectrum of matched truth and reco jets for corresponding rapidity bin respectively. Projections on  $y$  axis are shown in Appendix B.2.

shows the slices in transfer matrix of 2D unfolding which demonstrates the facts just stated.

Next to the transfer matrix, numbers of matched and unmatched reco and truth jets are needed for each  $(y, p_T)$  bin by unfolding procedure. These serve for calculation of matching efficiency which is the key ingredient in the first and the last step of unfolding procedure. Matching efficiencies for  $|y| < 0.5$  rapidity bin are for both simple and 2D unfolding shown in Figure 3.4 and for other rapidity bins, the results are shown in Appendix B.1

Unfolding procedure can be divided into three main steps

1. Input data are multiplied by the matching efficiency of reco jets.
2. Transfer matrix is used to correct data spectrum for detector effects. For this purpose, the Iterative Dynamical Stabilized (IDS) [54] unfolding method was used which uses the series of iterations to improve unfolding results. In this thesis the iteration was done once.
3. Spectrum obtained by the step 2 is divided by the matching efficiency of truth jets in order to correct resulting spectrum for the unmatched truth jets.

Figure 3.5 shows the comparison of  $p_T$  spectra of reco jets and unfolded spectrum (by 2D unfolding method) with the  $p_T$  spectrum of truth jets (left) and the comparison of simple and 2D unfolded spectra with the spectrum of truth jets (right) for  $|y| < 0.5$  rapidity bin. Results for all rapidity bins are shown in Appendix B.3.

From figures it follows, the unfolding procedure corrects the  $p_T$  spectrum of reco jets to  $p_T$  spectrum of truth jets up to the systematic error  $< 10^{-3}\%$  and that the differences between results of simple and 2D unfolding are even smaller.

### 3.4 Comparison with NLO Prediction

In the previous sections, I have described the jet calibration and the unfolding procedures. These serve to remove the detector imperfections and correct the reco jet  $p_T$  spectrum to the truth jet  $p_T$  spectrum. For this purpose I have used the PYTHIA8 generated events, which uses the leading-order QCD calculations to simulate initial  $pp$  collision. Nowadays the QCD predictions

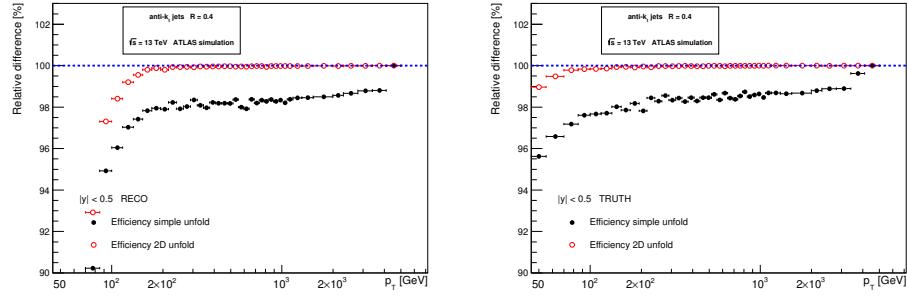


Figure 3.4: Comparison of matching efficiency of simple and 2D unfolding for  $|y| < 0.5$  rapidity bin. Matching efficiency is compared for both reco jets (left) and truth jets (right). Matching efficiency for all rapidity bins is shown in Appendix B.1.

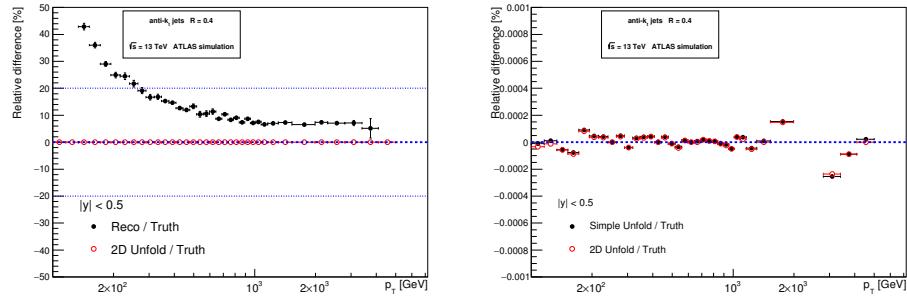


Figure 3.5: Comparison of  $p_T$  spectra of reco jets and the unfolded  $p_T$  spectra (2D unfolding) with the  $p_T$  spectra of the truth jets (left). Comparison of unfolded spectra obtained by the 2D and simple unfolding with the  $p_T$  spectra of the truth jets (right). Both graphs are for  $|y| < 0.5$  rapidity regions. Results for all rapidity bins are shown in Appendices B.3, B.4.

are tested up to next-to-leading order and for LHC Run II, new calculations assuming next-to-next-to-leading order QCD processes are preparing [?, ?].

My supervisor has determined the theoretical prediction of  $p_T$  spectra of parton jets using NLOJET++ program v4.1.2 [58]. This program computes the QCD processes up to next-to-leading order with CT10 parton distribution functions [?, 55]. In this thesis I have used his computations for center-of-mass energies  $\sqrt{s} = 8 \text{ TeV}$  and  $\sqrt{s} = 13 \text{ TeV}$ , first corresponding to the LHC Run I and second to LHC Run II. In this section I compare the truth jet  $p_T$  spectrum from PYTHIA8 with the theoretical prediction of parton jet  $p_T$  spectrum from next-to-leading order QCD predictions of my supervisor.

NLO predictions were compared for two different center-of-mass energies of  $pp$  collisions - first corresponding to the LHC Run I ( $\sqrt{s} = 8 \text{ TeV}$ ) and second to Run II ( $\sqrt{s} = 13 \text{ TeV}$ ). This comparison is shown for  $|y| < 0.5$  rapidity bin in Figure 3.6, where also the differential cross section is multiplied by the  $p_T$  bin width and by the integrated luminosity of Run I ( $20 \text{ fb}^{-1}$ ) and expected integrated luminosity of Run II ( $100 \text{ fb}^{-1}$ ) to obtain expected number of jets observed in each  $p_T$  bin. Comparisons for other rapidity bins are shown in Appendix C.1.

It can be seen, that the increase in the center-of-mass energy is the most significant for jets with high  $p_T$ . In the NLO QCD theoretical computations, several uncertainties are taken into account. These include [57]

- **Scale uncertainty**

Coming from the choice of renormalization and factorization scales, including neglecting the higher order terms beyond NLO

- **$\alpha_S$  uncertainty**

Because  $\alpha_S$  is experimentally determined with uncertainty.

- **PDF uncertainty**

Prediction depends on the concrete choice of PDF.

- **Nonperturbative corrections uncertainty**

Hadronization can split outgoing hard parton into a several jets.

- **Electroweak corrections uncertainty**

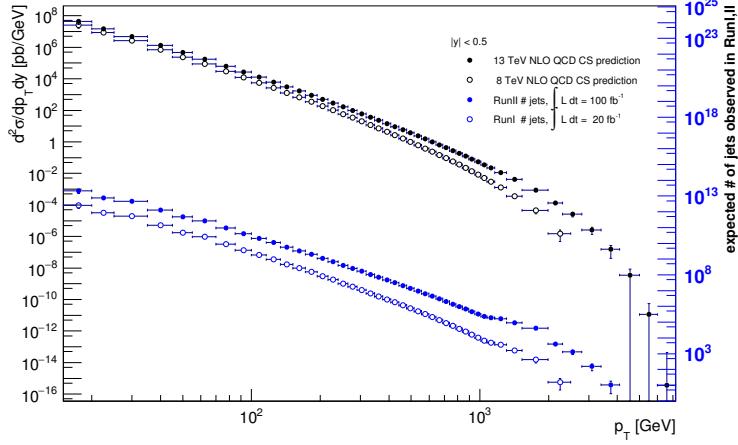


Figure 3.6: Comparison of NLO QCD prediction of double differential inclusive jet cross section (black) for  $pp$  collisions at  $\sqrt{s} = 13$  TeV (filled circles) corresponding to the LHC Run II and  $\sqrt{s} = 8$  TeV (empty circles) corresponding to the LHC Run I. The cross section is multiplied by integrated luminosities and  $p_T$  bin width to obtain the expected number of jets observed in each  $p_T$  bin (blue). Figure shows only  $|y| < 0.5$  rapidity bin, remaining rapidity bins are shown in Appendix C.1.

Next to the QCD processes, the electroweak processes has to be taken into account. These processes becomes more important, as the momentum transfer increases.

In this analysis, only first three of these corrections are assumed. The corrections were extracted from the files with NLO QCD predictions, where each correction is represented by the set of equally likely histograms expressing the deviation from the default prediction. Corrections are for  $|y| < 0.5$  rapidity bin shown in Figure 3.7, other rapidity bins are shown in Appendix C.2.

Comparison of  $p_T$  spectra of truth jets with NLO QCD prediction is for  $|y| < 0.5$  rapidity bin shown at Figure 3.8 for other rapidity bins see Appendix C.3. It can be seen that the truth  $p_T$  spectrum is for low  $p_T$  jet greater then the next-to-leading order QCD prediction and that for a few  $p_T$  bins with the highest  $p_T$  the situation is reversed. Physically, this is caused by the splitting of parton jet in the process of hadronization into several truth jets.

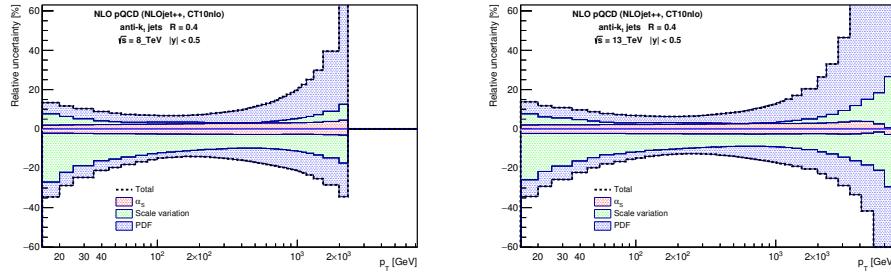


Figure 3.7: Theoretical uncertainties for NLO QCD predictions of inclusive jet differential cross section for  $pp$  collisions at  $\sqrt{s} = 8$  TeV (left) and  $\sqrt{s} = 13$  TeV (right) for  $|y| < 0.5$  rapidity bin. Uncertainties for other rapidity bins are shown in Appendix C.2.

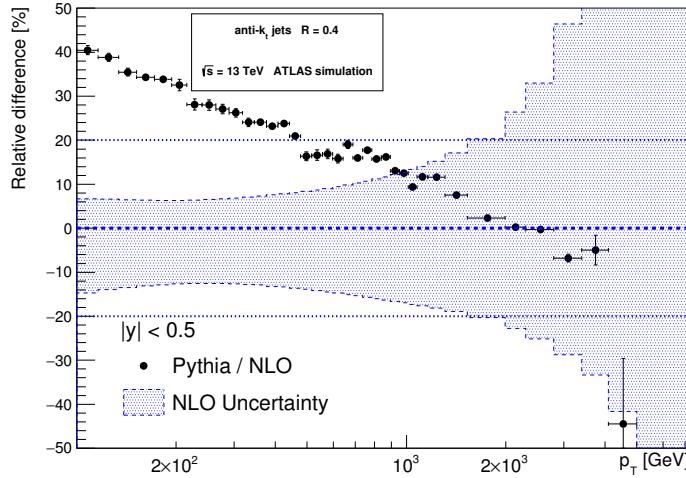


Figure 3.8: Comparison of PYTHIA8 prediction with NLO QCD prediction of inclusive jet double differential cross section for  $|y| < 0.5$  rapidity bin with uncertainties of NLO QCD predictions symbolized by the blue area. Comparisons for other rapidity bins are shown in Appendix C.3.



# Conclusion

This thesis deals with the measurement of the inclusive jet double differential cross section in  $p_T$  and rapidity in the ATLAS experiment. Inclusive jets are the prevailing objects created in the inelastic collisions at the LHC with  $p_T$  covering range from a few GeV to a few TeV. Nowhere is the increase in the center-of-mass energy appreciated as is in the case of inclusive jets as it can be seen from Figure 3.6. According to the preliminary analysis, the  $pp$  collision in LHC Run II with the center-of-mass energy  $\sqrt{s} = 13$  TeV could create thousands of jets with  $p_T$  in the interval  $1 \text{ TeV} < p_T < 4 \text{ TeV}$ .

Inclusive jets are theoretically straightforward and hence powerful test of perturbative QCD and with wide range of momentum transfers, the inclusive jet cross section is sensitive to the properties of the running coupling constant  $\alpha_S$ . Momentum transfers in orders of  $\sim 1 \text{ TeV}$  will probe the structure of proton at small distance scales  $\lambda \sim 1/p_T \sim \text{TeV}^{-1} \sim 10^{-19} \text{ m}$  and will contribute to our understanding of the proton structure (parton distribution functions). If there is new physics at these scales (such as the structure of quark), the inclusive jets may reveal it.

This thesis begins with a brief description of QCD as one of the components of the Standard Model. My intention was to connect historical development of QCD with definitions of its key concepts, which I am using later in this thesis. These concepts include the parton distribution functions, running coupling constant  $\alpha_S$  and asymptotic freedom phenomena. The first chapter ends with the splitting of QCD into perturbative and non-perturbative regions.

The second chapter starts with the description of the LHC with the ATLAS detector, and introduces the most important concept of this thesis - a jet. Using the QCD, I am trying to explain the necessity of the jet on hadron colliders, and to define the jet itself. This is done by the definition

of jet algorithms, which I am formulating with the emphasis on possible jet definition at three different stages of collision: at parton, particle and detector level. The second chapter ends with the description of jet energy scale calibration and unfolding procedures, which objective is to reduce negative detector effects and to lower differences between jets at detector and particle levels.

The last chapter of my thesis describes the analysis of the double differential inclusive jet cross section in  $p_T$  and rapidity  $y$ . As the Run II data are not yet available, I have used as the input the PYTHIA8 generated events of  $pp$  collisions with the center-of-mass energy  $\sqrt{s} = 13$  TeV using CT10 parton distribution functions and ATLAS underlying event tune AU2. These data contains the jets reconstructed on particle and detector levels and as a jet algorithm, the anti- $k_t$  jet algorithm with parameter  $R = 0.4$  is used.

The first step of my analysis is energy scale calibration of jets on detector level, which is followed by the description of event selection criteria and jet matching procedure. To correct the reconstructed  $p_T$  spectrum on detector level, I implement two different approaches of data unfolding. A simple unfolding, the first of unfolding approaches, allows jet matching only within the same rapidity bins. In addition to the simple unfolding I introduce a 2D unfolding, which allows matching between different rapidity bins. These approaches differ in definitions of transfer matrices and in matching efficiencies which are shown in Appendix B.1. It can be seen the matching efficiency in case of simple unfolding is  $\sim 2 - 5\%$  worse than the matching efficiency in 2D unfolding. This should cause the unfolded spectrum from 2D unfolding will be more precise than that of simple unfolding.

The results from both of the unfolding approaches are compared with the  $p_T$  spectrum of particle level jets in Appendix B.4. It can be seen, that in both cases, the unfolded  $p_T$  spectrum is in agreement with the  $p_T$  spectrum of particle level jets up to systematic error  $< 10^{-3}\%$  and that the relative difference between two unfolding approaches is even smaller.

The second input for this analysis is the next-to-leading (NLO) order perturbative QCD prediction of double differential inclusive jet cross section in  $p_T$  and rapidity for center-of-mass energies  $\sqrt{s} = 8$  TeV (corresponding to LHC Run I) and  $\sqrt{s} = 13$  TeV (Run II) calculated using NLOJET++ program using the same parton distribution functions as PYTHIA8. The predictions include uncertainties in coupling constant  $\alpha_S$ , PDFs and factor-

ization and normalization scales. These uncertainties are shown in Appendix C.2.

Cross section predictions for  $\sqrt{s} = 8 \text{ TeV}$  and  $\sqrt{s} = 13 \text{ TeV}$  are compared in Appendix C.1 where the integrated luminosities  $L = 20 \text{ fb}^{-1}$  (Run I) and  $L = 100 \text{ fb}^{-1}$  (Run II expected) are used. The numbers of jets expected to be observed in each  $p_T$  bin were calculated. It can be seen, that according to this prediction, in Run II  $\sim 1000$  times more jets with  $p_T > 1 \text{ TeV}$  will be observed than it was in Run I. In addition LHC run II could create a few jets with  $p_T \sim 4 \text{ TeV}$ .

The  $p_T$  spectrum of particle jets obtained from PYTHIA8 is compared with  $p_T$  spectrum of parton jets from NLO prediction in Appendix C.3. Only for few highest  $p_T$  bins the cross section predicted by the NLO is bigger than that of PYTHIA8 whereas for low  $p_T$  the situation is opposite. The explanation is, that PYTHIA8 generates events of  $pp$  collisions only using LO QCD.

There are several ways to extend this analysis. Although this analysis is only preparation for the analysis of real data which will be collected in LHC Run II, the implementations of the IDS unfolding method developed could serve for unfolding of the real  $p_T$  spectra measured by the ATLAS detector in Run II. The unfolding results could be further improved by the running on new datasets generated by the PYTHIA8 with the usage of newer parton density functions.

The unfolding approaches introduced in this thesis could be probed further. One of the possibility is the event reweighting and checking, if the unfolding is not dependent on the shape of the  $p_T$  spectra. Next possibility is to compare, how the simple and 2D unfolding will deal with new structure in data.



# Appendices



## Appendix A

### Cut and Match Results

## A.1 Cut Results

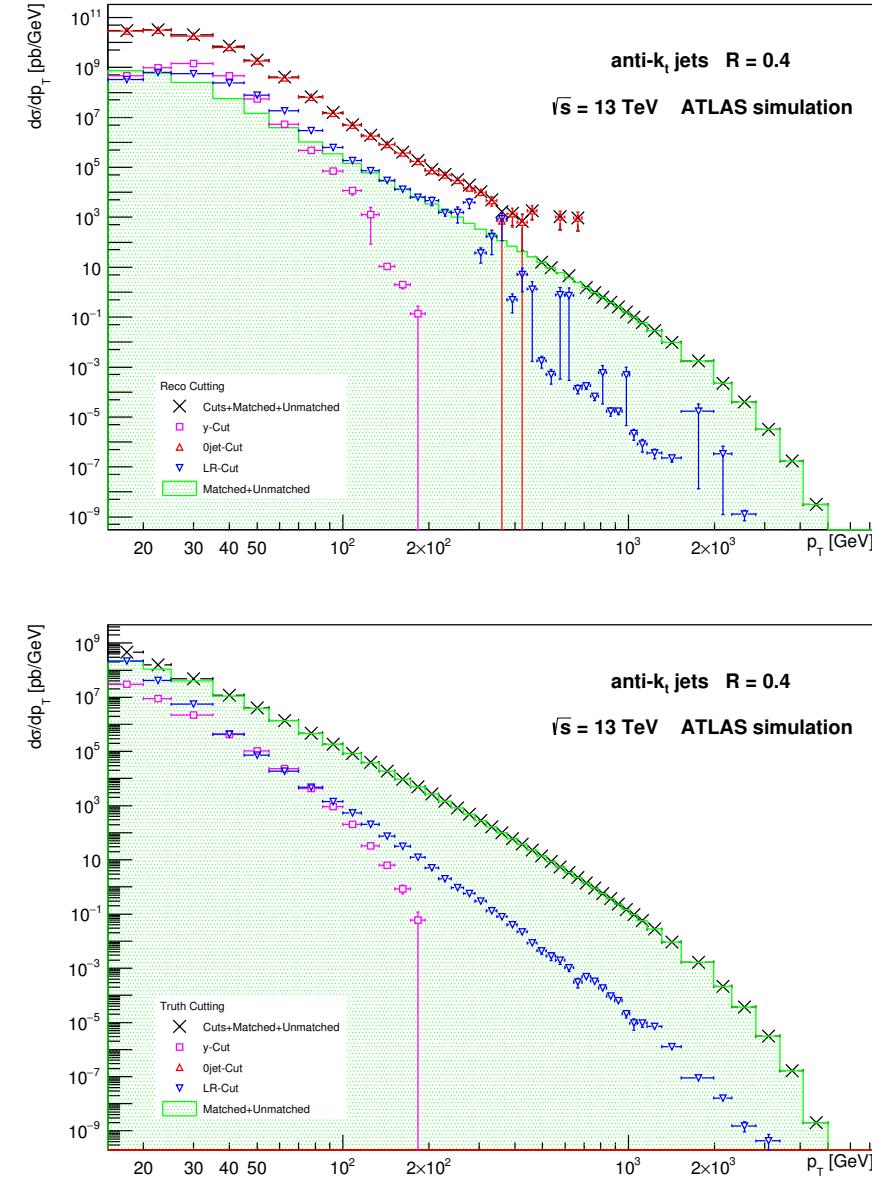


Figure A.1: Impact of 4 cuts defined in Section 3.2.1 on differential cross section in  $p_T$  of reco jets (top) and truth jets (bottom). Black dots represent the original uncutted spectrum, green area then these jets, which survived all four cutoffs.

## A.2 Match Results

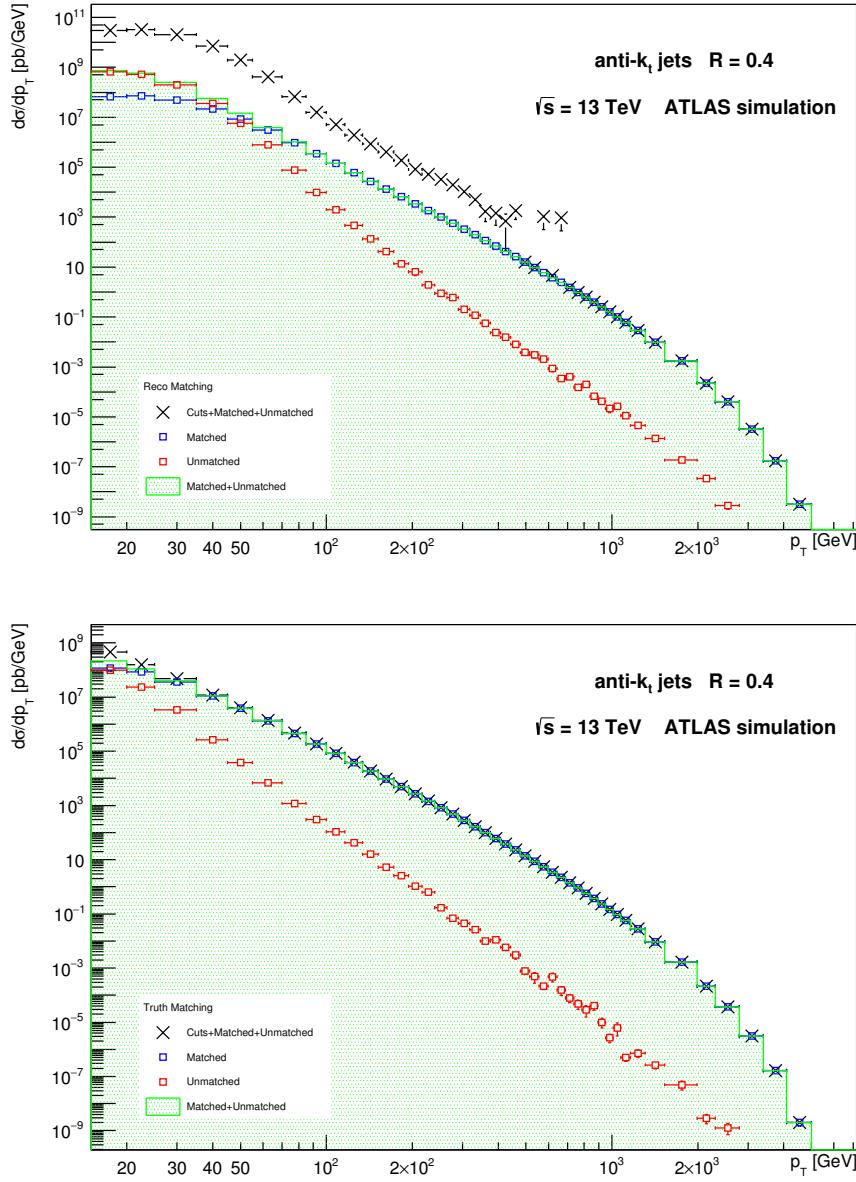


Figure A.2: Results of matching procedure described in Section 3.2.2 demonstrated on differential cross section in  $p_T$  of reco (top) and truth (bottom) jets. Black dots represent the original uncutted spectrum. The contribution of matched and unmatched jets to green area representing all jets which survived cutoffs is shown.

### A.3 Truth and Reco $p_T$ Spectra

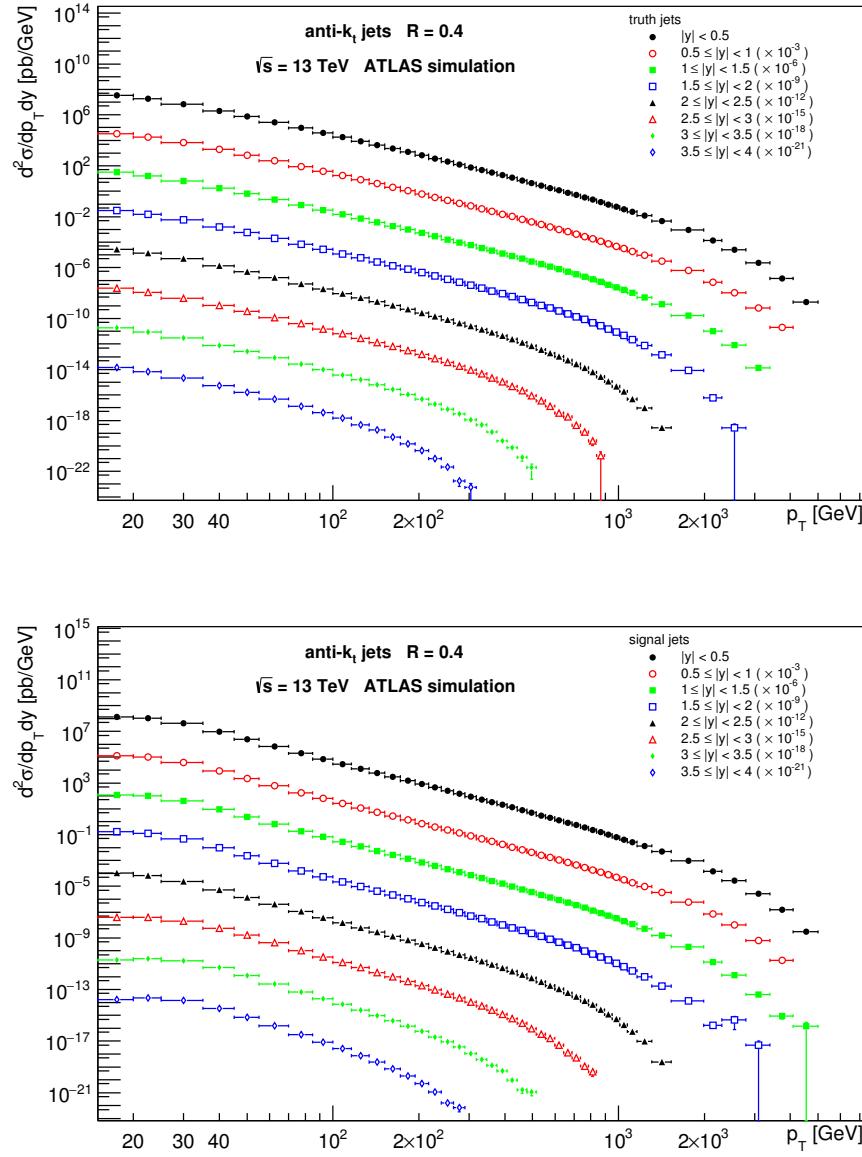


Figure A.3: Inclusive jet cross section of truth (top) and reco (bottom) jets as the function of jet  $p_T$  and rapidity. For the convenience the cross sections for different rapidity bins are multiplied by the factor indicated in the legend. Jets were identified with anti- $k_t$  jet algorithm with  $R = 0.4$ .

	# jets	ALL	JZ0W	JZ1W	JZ2W	JZ3W	JZ4W	JZ5W	JZ6W	JZ7W
Reco	<b>1.09e+08</b>	3.11e+07	3.59e+07	6.67e+06	7.07e+06	6.28e+06	7.29e+06	7.13e+06	7.11e+06	7.11e+06
Truth	<b>7.28e+07</b>	3.04e+06	3.00e+07	6.17e+06	6.91e+06	6.20e+06	6.98e+06	6.53e+06	6.25e+06	6.25e+06
CutPt	Reco	<b>9.36e+06</b> 8.6 %	3.74e+06 12.0 %	3.13e+06 8.7 %	6.50e+05 9.7 %	5.87e+05 8.3 %	4.76e+05 7.6 %	5.48e+05 7.5 %	5.52e+05 7.7 %	5.63e+05 7.9 %
	Truth	<b>4.70e+07</b> 64.6 %	3.00e+06 98.7 %	2.20e+07 73.1 %	3.86e+06 62.6 %	4.00e+06 57.8 %	3.42e+06 55.1 %	3.74e+06 53.6 %	3.43e+06 52.5 %	3.23e+06 51.6 %
Cuty	Reco	<b>3.43e+06</b> 3.1 %	1.19e+06 3.8 %	1.29e+06 3.6 %	1.42e+05 2.1 %	1.28e+05 1.8 %	1.03e+05 1.6 %	1.16e+05 1.6 %	1.10e+05 1.5 %	1.08e+05 1.5 %
	Truth	<b>5.06e+05</b> 0.7 %	3.04e+03 0.1 %	3.19e+05 1.1 %	4.54e+04 0.7 %	3.79e+04 0.5 %	2.88e+04 0.5 %	2.78e+04 0.4 %	2.22e+04 0.3 %	1.83e+04 0.3 %
Cut0jet	Reco	<b>2.64e+07</b> 24.2 %	2.59e+07 83.2 %	5.70e+05 1.6 %	5.00e+00 0.0 %	0.00e+00 0.0 %				
	Truth	<b>0.00e+00</b> 0.0 %	0.00e+00 0.0 %							
Cut1LR	Reco	<b>4.09e+06</b> 3.7 %	2.38e+05 0.8 %	3.82e+06 10.6 %	2.99e+04 0.4 %	7.07e+03 0.1 %	2.33e+03 0.0 %	1.63e+03 0.0 %	7.14e+02 0.0 %	6.31e+02 0.0 %
	Truth	<b>5.40e+05</b> 0.7 %	2.19e+04 0.7 %	4.96e+05 1.7 %	1.82e+04 0.3 %	4.45e+03 0.1 %	1.33e+03 0.0 %	9.03e+02 0.0 %	4.37e+02 0.0 %	2.78e+02 0.0 %
Matched	Reco	<b>2.17e+07</b> 19.8 %	7.62e+03 0.0 %	6.03e+06 16.8 %	1.95e+06 20.1 %	2.54e+06 31.7 %	2.46e+06 36.8 %	2.88e+06 39.6 %	2.78e+06 41.3 %	2.72e+06 42.5 %
	Truth	<b>2.17e+07</b> 29.8 %	7.62e+03 0.3 %	6.03e+06 20.1 %	1.95e+06 31.7 %	2.54e+06 36.8 %	2.46e+06 39.6 %	2.88e+06 41.3 %	2.78e+06 42.5 %	2.72e+06 43.5 %
Unmatched	Reco	<b>4.42e+07</b> 40.5 %	5.36e+04 0.2 %	2.10e+07 58.6 %	3.89e+06 58.4 %	3.81e+06 53.8 %	3.24e+06 51.6 %	3.75e+06 51.4 %	3.69e+06 51.8 %	3.72e+06 52.3 %
	Truth	<b>3.07e+06</b> 4.2 %	6.18e+03 0.2 %	1.25e+06 4.2 %	2.89e+05 4.7 %	3.29e+05 4.8 %	2.95e+05 4.8 %	3.29e+05 4.7 %	3.03e+05 4.6 %	2.88e+05 4.6 %

Table A.1: Statistics for matching and cutting procedures described in Sections 3.2.1 and 3.2.2 displayed for all jets and for individual JZXW samples defined in Table 3.1. At the top, there is number of initial reco and truth jets respectively. For each cut, there is shown the number of jets, which were killed by it, and their relative number according to the original number of reco or truth jets respectively. Last two lines show the statistics of matching procedure including number of jets which were (un)matched.



## Appendix B

# Unfolding

### B.1 Matching Efficiency

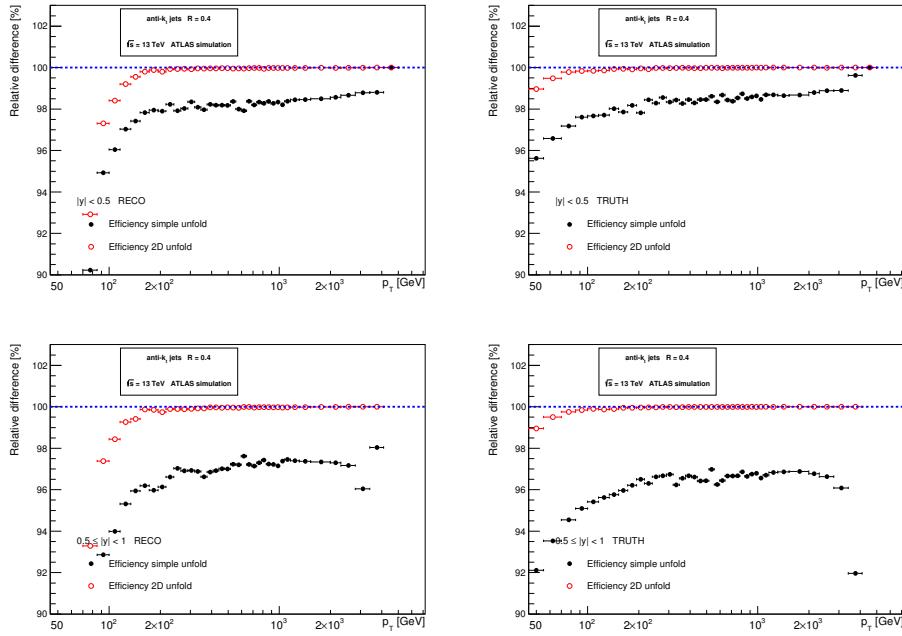


Figure B.1: Comparison of matching efficiencies of simple and 2D unfolding for  $|y| < 0.5$  (top) and  $0.5 \leq |y| < 1$  (bottom) rapidity bins. Matching efficiencies are shown for both reco jets (left) and truth jets (right).

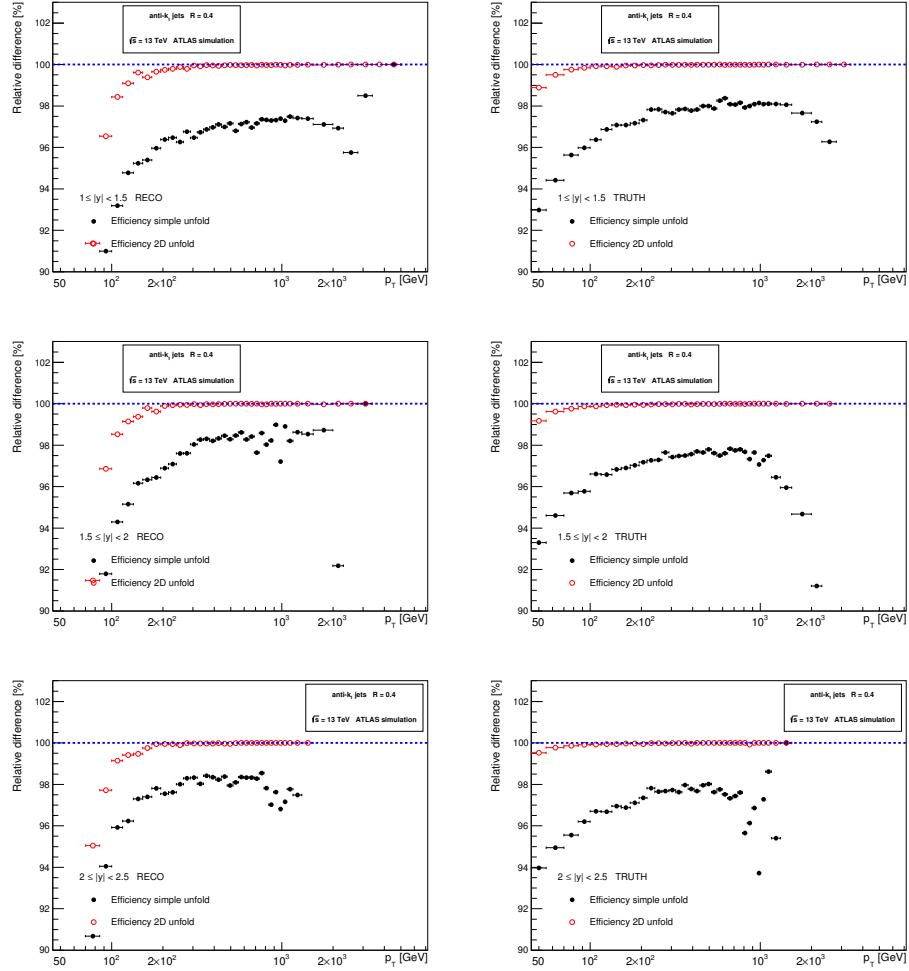


Figure B.2: Comparison of matching efficiencies of simple and 2D unfolding for  $1 \leq |y| < 1.5$  (top),  $1.5 \leq |y| < 2$  and  $2 \leq |y| < 2.5$  (bottom) rapidity bins. Matching efficiencies are shown for both reco jets (left) and truth jets (right).

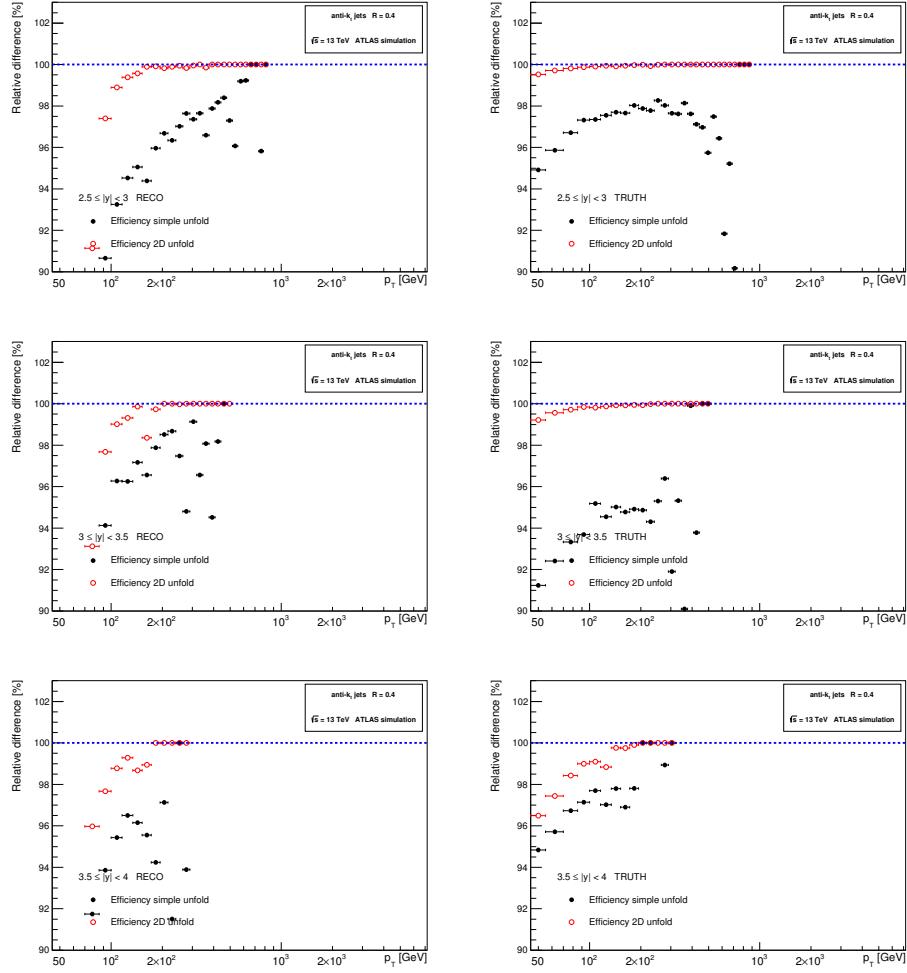


Figure B.3: Comparison of matching efficiencies of simple and 2D unfolding for  $2.5 \leq |y| < 3$  (top),  $3 \leq |y| < 3.5$  and  $3.5 \leq |y| < 4$  (bottom) rapidity bins. Matching efficiencies are shown for both reco jets (left) and truth jets (right).

## B.2 Slices in Unfolding Matrix

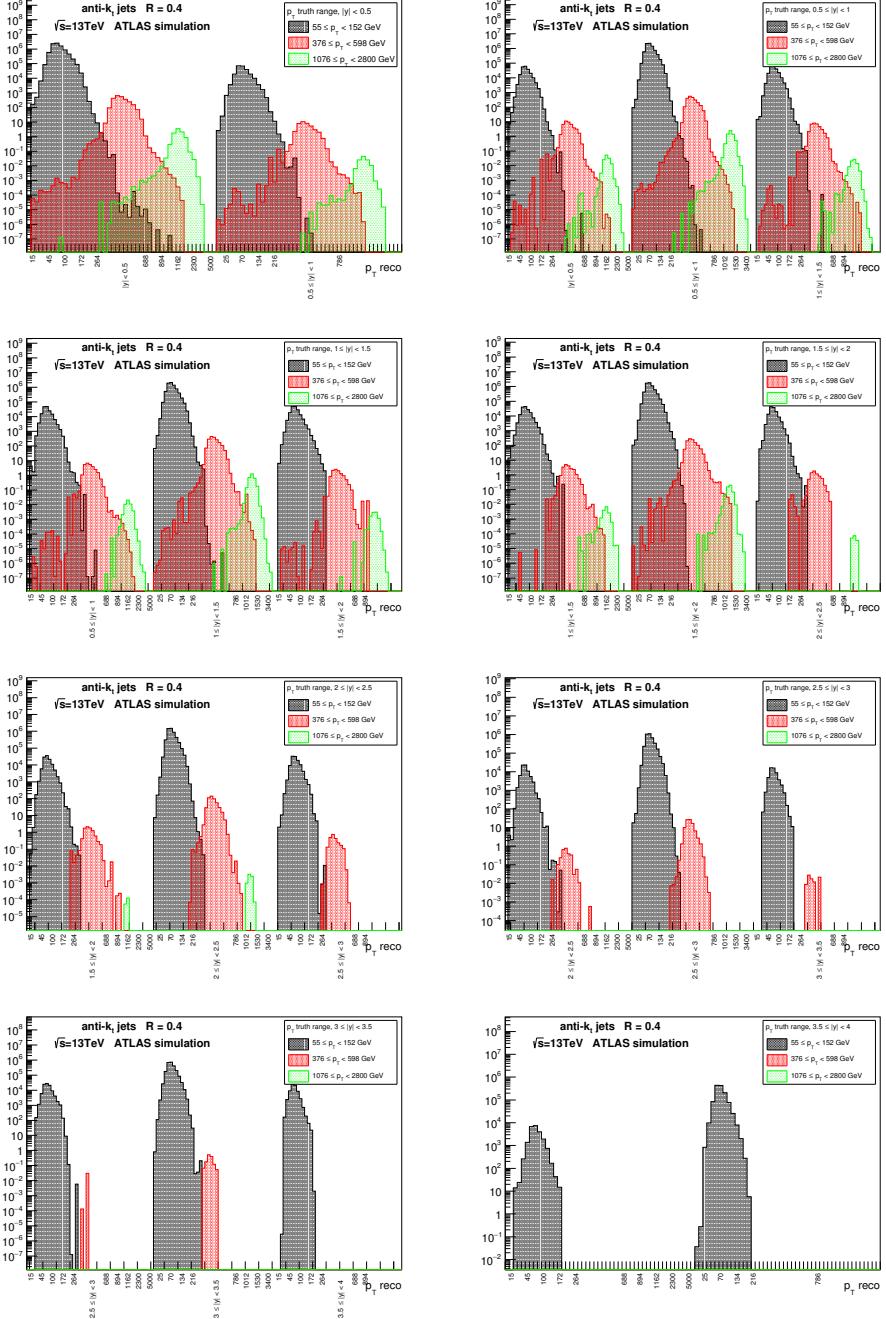


Figure B.4: Slices in transfer matrix of 2D unfolding method from Figure on  $y$ -axis represented by the  $p_T$  spectrum of reco jets. Slices are done in each rapidity bin for three different ranges of  $p_T$  of truth jets.

### B.3 Unfolding Results

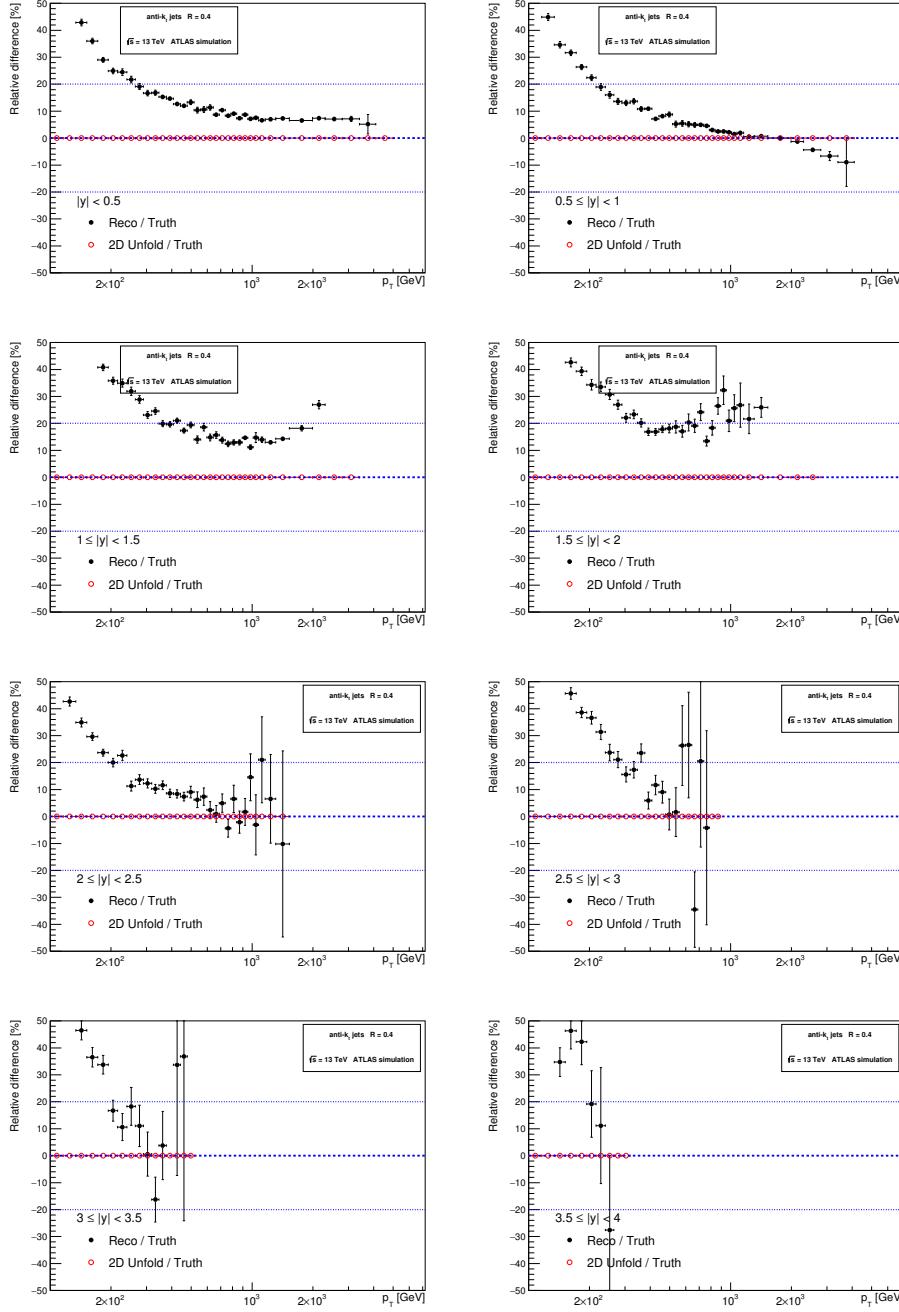


Figure B.5: Comparison of  $p_T$  spectra of reco jets and the unfolded  $p_T$  spectra (2D unfolding) with the  $p_T$  spectra of the truth jets for 8 different rapidity bins.

## B.4 Simple and 2D Unfolding

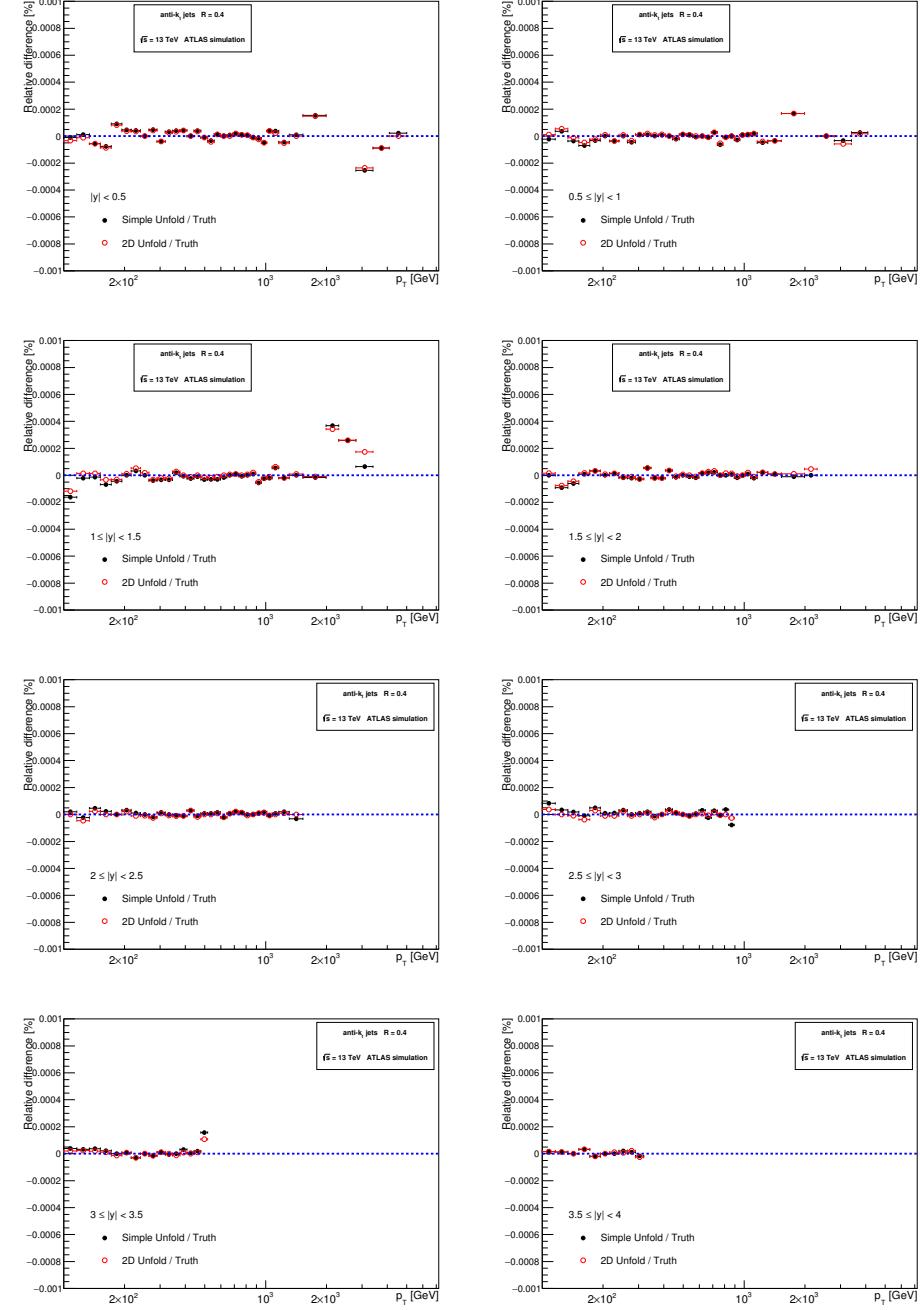


Figure B.6: Comparison of results obtained by two unfolding method - simple and 2D unfolding - with the  $p_T$  spectra of truth jets for 8 different rapidity bins.

## Appendix C

## NLO QCD Prediction

### C.1 Predictions for Run I and Run II

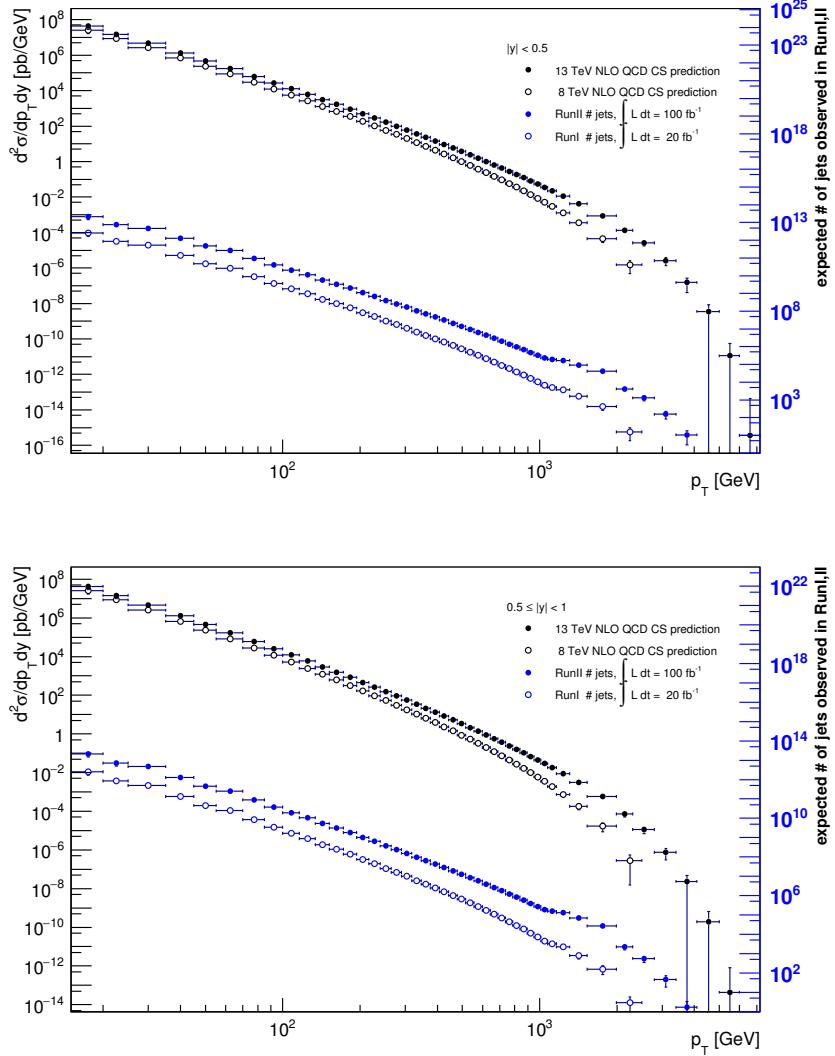


Figure C.1: Comparison of NLO QCD prediction of double differential inclusive jet cross section (black) in  $p_T$  and rapidity of  $pp$  collisions at  $\sqrt{s} = 13 \text{ TeV}$  (filled circles) corresponding to LHC Run II and  $\sqrt{s} = 8 \text{ TeV}$  (empty circles) corresponding to LHC Run I. The cross section is multiplied by integrated luminosities and  $p_T$  bin width to obtain expected number of jets observed in each  $p_T$  bin (blue). Figures show the comparison for  $0.5 < |y|$  (top) and  $0.5 \leq |y| < 1$  (bottom) rapidity bins.

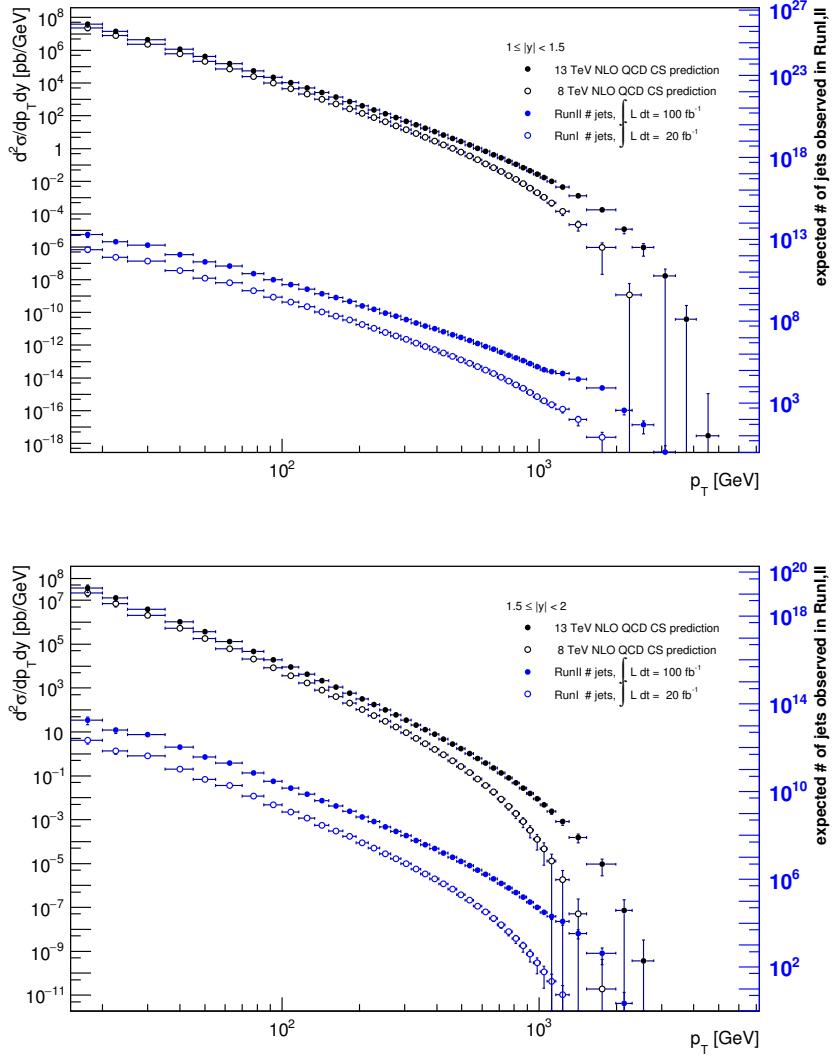


Figure C.2: Comparison of NLO QCD prediction of double differential inclusive jet cross section (black) in  $p_T$  and rapidity of  $pp$  collisions at  $\sqrt{s} = 13 \text{ TeV}$  (filled circles) corresponding to LHC Run II and  $\sqrt{s} = 8 \text{ TeV}$  (empty circles) corresponding to LHC Run I. The cross section is multiplied by integrated luminosities and  $p_T$  bin width to obtain expected number of jets observed in each  $p_T$  bin (blue). Figures show the comparison for  $1 \leq |y| < 1.5$  (top) and  $1.5 \leq |y| < 2$  (bottom) rapidity bins.

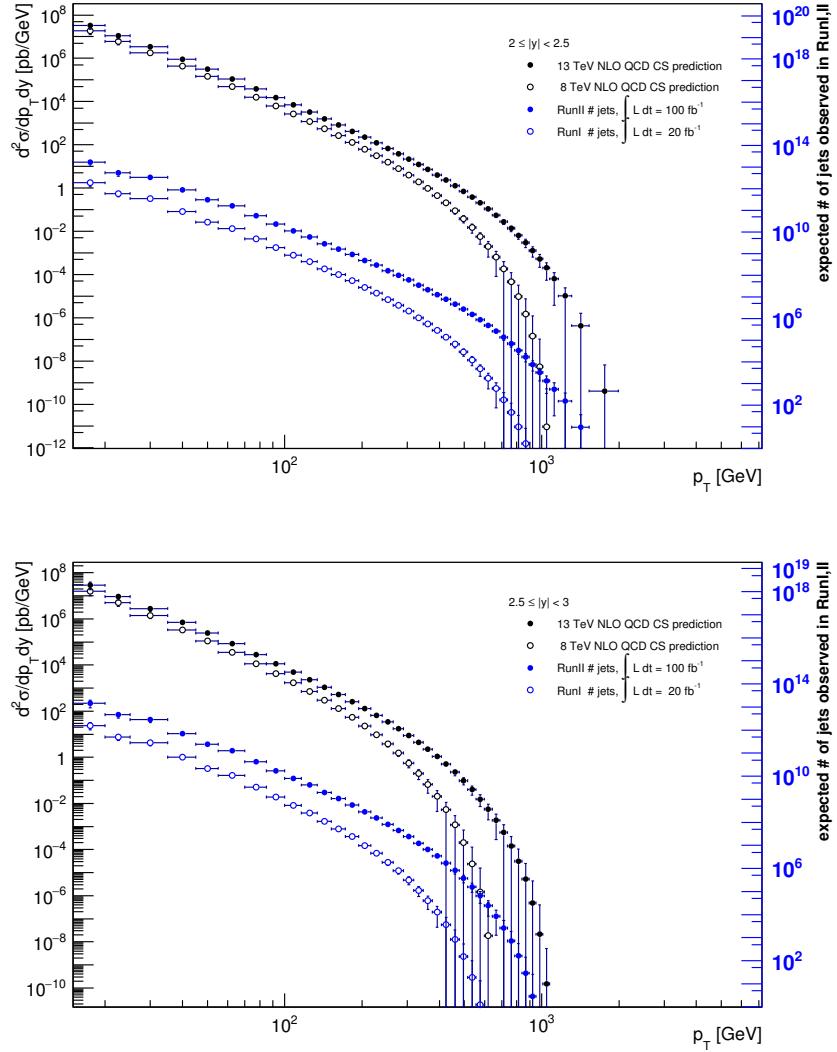


Figure C.3: Comparison of NLO QCD prediction of double differential inclusive jet cross section (black) in  $p_T$  and rapidity of  $pp$  collisions at  $\sqrt{s} = 13$  TeV (filled circles) corresponding to LHC Run II and  $\sqrt{s} = 8$  TeV (empty circles) corresponding to LHC Run I. The cross section is multiplied by integrated luminosities and  $p_T$  bin width to obtain expected number of jets observed in each  $p_T$  bin (blue). Figures show the comparison for  $2 \leq |y| < 2.5$  (top) and  $2.5 \leq |y| < 3$  (bottom) rapidity bins.

## C.2 NLO Uncertainties

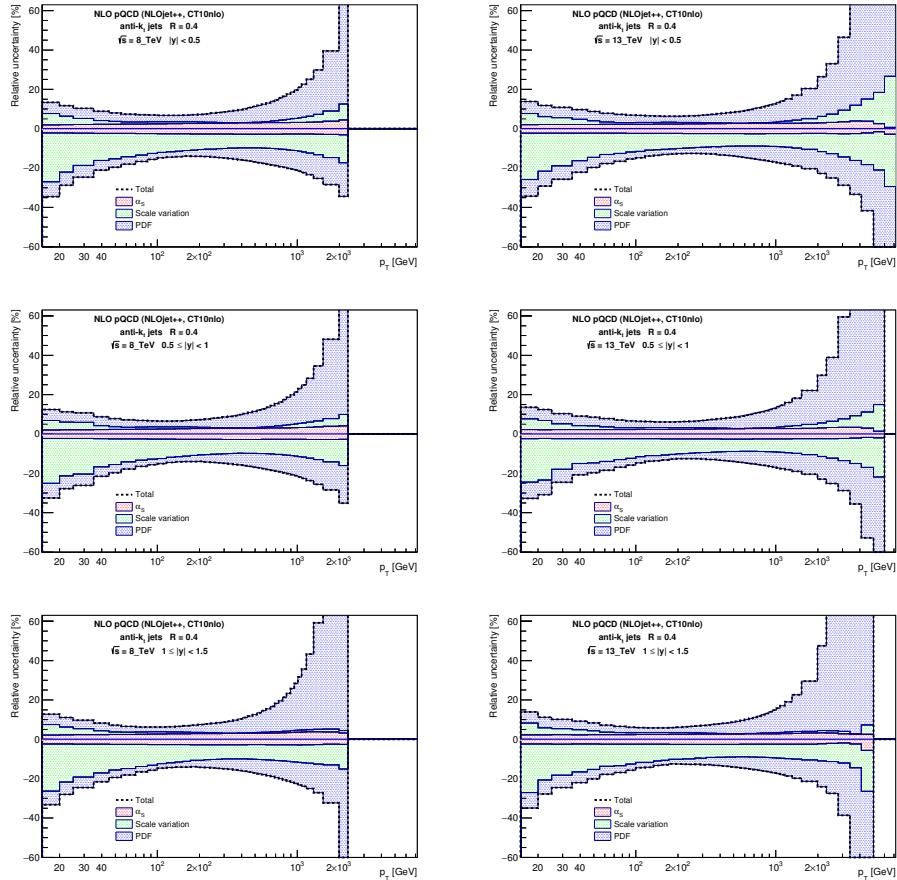


Figure C.4: NLO uncertainties for NLO QCD predictions of inclusive jet double differential cross section of  $pp$  collisions at  $\sqrt{s} = 8 \text{ TeV}$  (left) and  $\sqrt{s} = 13 \text{ TeV}$  (right) for  $|y| < 0.5$  (top),  $0.5 \leq |y| < 1$  (middle) and  $1 \leq |y| < 1.5$  (bottom) rapidity bins.

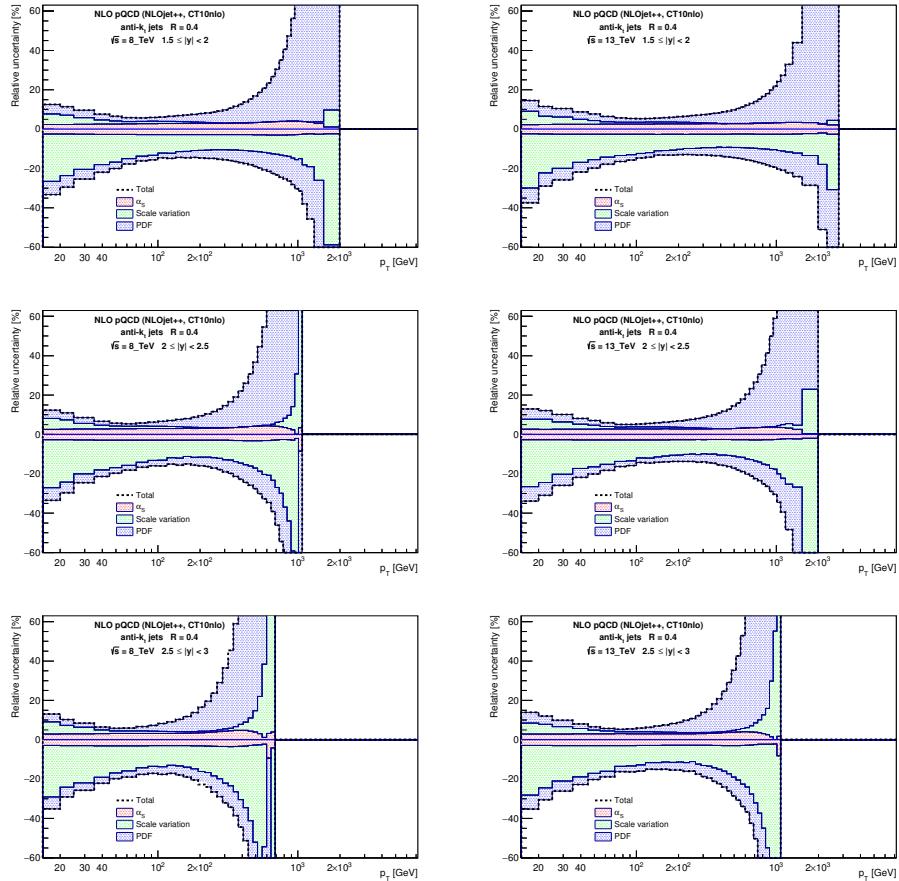


Figure C.5: NLO uncertainties for NLO QCD predictions of inclusive jet double differential cross section of  $pp$  collisions at  $\sqrt{s} = 8$  TeV (left) and  $\sqrt{s} = 13$  TeV (right) for  $1.5 \leq |y| < 2$  (top),  $2 \leq |y| < 2.5$  (middle) and  $2.5 \leq |y| < 3$  (bottom) rapidity bins.

### C.3 Pythia and NLO

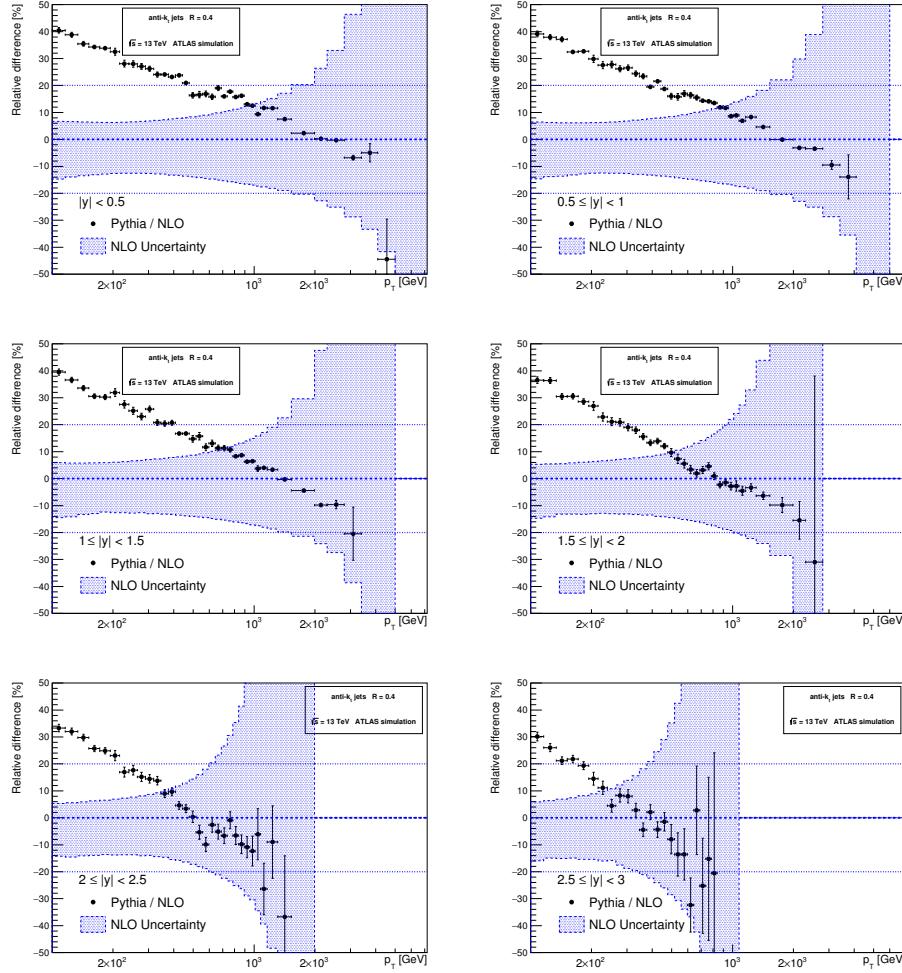


Figure C.6: Comparison of PYTHIA8 prediction with NLO QCD prediction of inclusive jet double differential cross section in  $p_T$  and rapidity for six different rapidity bins. Blue area represents the uncertainties of NLO QCD predictions.



# Bibliography

- [1] S. Durr, “Lattice qcd for pedestrians.” [http://www.physi.uni-heidelberg.de/Forschung/he/LHCb/documents/WorkshopNeckarzFeb13/durr\\_neckarzimmern.pdf](http://www.physi.uni-heidelberg.de/Forschung/he/LHCb/documents/WorkshopNeckarzFeb13/durr_neckarzimmern.pdf), 2015. [Online; accessed 16-April-2015].
- [2] C. Colaboration, “Observation of top quark production in  $\bar{p}p$  collisions with the collider detector at fermilab,” *Phys. Rev. Lett.*, vol. 74, pp. 2626–2631, Apr 1995.
- [3] D. Colaboration, “Search for high mass top quark production in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.8$  tev,” *Phys. Rev. Lett.*, vol. 74, pp. 2422–2426, Mar 1995.
- [4] G. Aad *et al.*, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC,” *Phys.Lett.*, vol. B716, pp. 1–29, 2012.
- [5] “Physics at run 2 the ultimate search engine: Lhc.” <http://indico.cern.ch/event/251191/session/0/material/slides/0?contribId=1>, 2015. [Online; accessed 14-April-2015].
- [6] “Standard model — Wikipedia, the free encyclopedia.” [http://en.wikipedia.org/wiki/Standard\\_Model](http://en.wikipedia.org/wiki/Standard_Model), 2015.
- [7] K. Olive *et al.*, “Review of Particle Physics,” *Chin.Phys.*, vol. C38, p. 090001, 2014.
- [8] W. Greiner, D. Bromley, S. Schramm, and E. Stein, *Quantum Chromodynamics*. Springer, 2007.
- [9] J. Horejsi, *Fundamentals of Electroweak Theory*. Karolinum Press, 2002.

- [10] D. Griffiths, *Introduction to Elementary Particles*. Physics textbook, Wiley, 2008.
- [11] W. Cottingham and D. Greenwood, *An Introduction to the Standard Model of Particle Physics*. Cambridge University Press, 2007.
- [12] T. Nakano and K. Nishijima, “Charge independence for v-particles,” *Progress of Theoretical Physics*, vol. 10, no. 5, pp. 581–582, 1953.
- [13] M. Gell-Mann, “The interpretation of the new particles as displaced charge multiplets,” *Il Nuovo Cimento*, vol. 4, no. 2, pp. 848–866, 1956.
- [14] M. Gell-Mann and Y. Ne’eman, *The eightfold way*. Frontiers in Physics, New York, NY: Benjamin, 1964.
- [15] G. Zweig, “An SU<sub>3</sub> model for strong interaction symmetry and its breaking; Version 2,” p. 80 p, Feb 1964.
- [16] H. Georgi, *Lie Algebras in Particle Physics: From Isospin to Unified Theories*. Frontiers in Physics Series, Westview Press, 1999.
- [17] “Eightfold way — Wikipedia, the free encyclopedia.” [http://de.wikipedia.org/wiki/Eightfold\\_Way](http://de.wikipedia.org/wiki/Eightfold_Way), 2015.
- [18] M. Gell-Mann and H. Fritzsch, *Murray Gell-Mann: selected papers*. World Scientific series in 20th century physics, Singapore: World Scientific, 2010.
- [19] S. Okubo, “Note on unitary symmetry in strong interactions,” *Progress of Theoretical Physics*, vol. 27, no. 5, pp. 949–966, 1962.
- [20] S. L. Glashow, J. Iliopoulos, and L. Maiani, “Weak interactions with lepton-hadron symmetry,” *Phys. Rev. D*, vol. 2, pp. 1285–1292, Oct 1970.
- [21] M. Kobayashi and T. Maskawa, “Cp-violation in the renormalizable theory of weak interaction,” *Progress of Theoretical Physics*, vol. 49, no. 2, pp. 652–657, 1973.
- [22] E. D. Bloom, D. H. Coward, H. DeStaebler, J. Drees, G. Miller, L. W. Mo, R. E. Taylor, M. Breidenbach, J. I. Friedman, G. C. Hartmann, and H. W. Kendall, “High-energy inelastic  $e - p$  scattering at 6° and 10°,” *Phys. Rev. Lett.*, vol. 23, pp. 930–934, Oct 1969.

- [23] K. Olive *et al.*, “Review of Particle Physics,” *Chin.Phys.*, vol. C38, p. 090001, 2014.
- [24] C. N. Yang and R. L. Mills, “Conservation of isotopic spin and isotopic gauge invariance,” *Phys. Rev.*, vol. 96, pp. 191–195, 1954.
- [25] P. W. Higgs, “Broken symmetries and the masses of gauge bosons,” *Phys. Rev. Lett.*, vol. 13, pp. 508–509, Oct 1964.
- [26] L. Faddeev and V. Popov, “Feynman diagrams for the yang-mills field,” *Physics Letters B*, vol. 25, no. 1, pp. 29 – 30, 1967.
- [27] D. J. Gross and F. Wilczek, “Ultraviolet behavior of non-abelian gauge theories,” *Phys. Rev. Lett.*, vol. 30, pp. 1343–1346, Jun 1973.
- [28] M. Dordevic, “Standard Model physics results from ATLAS and CMS,” Tech. Rep. CMS-CR-2014-251, CERN, Geneva, Oct 2014.
- [29] “The history of qcd — cern courier.” <http://cerncourier.com/cws/article/cern/50796>, 2015.
- [30] R. Gupta, “Introduction to lattice QCD: Course,” pp. 83–219, 1997.
- [31] A. Ukawa, “Kenneth Wilson and lattice QCD,” 2015.
- [32] L. Evans and P. Bryant, “Lhc machine,” *Journal of Instrumentation*, vol. 3, no. 08, p. S08001, 2008.
- [33] G. Landsberg, “LHC: Past, Present, and Future,” 2013.
- [34] “Atlas experiment@2015 cern.” <http://www.atlas.ch/photos/index.html>, 2015.
- [35] R. Bruce, G. Arduini, S. Fartoukh, M. Giovannozzi, M. Lamont, *et al.*, “Baseline LHC machine parameters and configuration of the 2015 proton run,” 2014.
- [36] “Future prospects for higgs measurements at lhc run 2+3.” [http://www.nikhef.nl/pub/BND2014/projects/ILC-vs-LHC-HL/LHC\\_Aspen\\_2014\\_Kroha.pdf](http://www.nikhef.nl/pub/BND2014/projects/ILC-vs-LHC-HL/LHC_Aspen_2014_Kroha.pdf), 2014.
- [37] T. A. Collaboration, “The atlas experiment at the cern large hadron collider,” *Journal of Instrumentation*, vol. 3, no. 08, p. S08003, 2008.

- [38] *Expected performance of the ATLAS experiment: detector, trigger and physics.* Geneva: CERN, 2009.
- [39] E. Halkiadakis, “Proceedings for TASI 2009 Summer School on ‘Physics of the Large and the Small’: Introduction to the LHC experiments,” 2010.
- [40] T. D. Lee and M. Nauenberg, “Degenerate systems and mass singularities,” *Phys. Rev.*, vol. 133, pp. B1549–B1562, Mar 1964.
- [41] T. Kinoshita, “Mass singularities of feynman amplitudes,” *Journal of Mathematical Physics*, vol. 3, no. 4, 1962.
- [42] G. Aad *et al.*, “Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics,” 2009.
- [43] G. C. Blazey, J. R. Dittmann, S. D. Ellis, V. D. Elvira, K. Frame, *et al.*, “Run II jet physics,” pp. 47–77, 2000.
- [44] B. Isildak, “Measurement of the differential dijet production cross section in proton-proton collisions at  $\sqrt{s} = 7$  tev,” 2013.
- [45] M. Cacciari, G. P. Salam, and G. Soyez, “The Anti-k(t) jet clustering algorithm,” *JHEP*, vol. 0804, p. 063, 2008.
- [46] T. Sjostrand, S. Mrenna, and P. Z. Skands, “A Brief Introduction to PYTHIA 8.1,” *Comput.Phys.Commun.*, vol. 178, pp. 852–867, 2008.
- [47] S. Agostinelli *et al.*, “GEANT4: A Simulation toolkit,” *Nucl.Instrum.Meth.*, vol. A506, pp. 250–303, 2003.
- [48] W. Lukas, “Fast Simulation for ATLAS: Atlfast-II and ISF,” 2012.
- [49] Z. Hubacek, “Measurement of the Three-jet Mass Cross Section in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.96$  TeV,” 2010.
- [50] J. Bossio, “Applyjetcalibration2014.” <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ApplyJetCalibration2014>, 2015. [Online; accessed 16-April-2015].
- [51] V. Blobel, “Unfolding methods in high-energy physics experiments,” p. 40 p, Dec 1984.

- [52] G. D'Agostini, "Improved iterative Bayesian unfolding," *Phys.Rev.*, p. 31, 2010.
- [53] A. Hocker and V. Kartvelishvili, "SVD approach to data unfolding," *Nucl.Instrum.Meth.*, vol. A372, pp. 469–481, 1996.
- [54] B. Malaescu, "An Iterative, Dynamically Stabilized(IDS) Method of Data Unfolding," 2011.
- [55] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, *et al.*, "New parton distributions for collider physics," *Phys.Rev.*, vol. D82, p. 074024, 2010.
- [56] S. Wahrmund, "Atlas tunes for mpi." <http://indico.cern.ch/event/184925/session/4/contribution/19/material/slides/0.pdf>, 2015. [Online; accessed 16-April-2015].
- [57] T. Carli, E. Ertel, E. Fullana Torregrosa, A. Hrynevich, Z. Hubacek, B. Malaescu, N. Meric, C. Meyer, V. Pleskot, S. Shimizu, P. Starovoitov, G. Vardanyan, and S. Yanush, "Measurement of inclusive jet production in  $pp$  collisions at  $\sqrt{s} = 7$  TeV using the ATLAS detector (Supporting Documentation)," Tech. Rep. ATL-COM-PHYS-2013-1390, CERN, Geneva, Oct 2013.
- [58] Z. Nagy, "Next-to-leading order calculation of three jet observables in hadron hadron collision," *Phys.Rev.*, vol. D68, p. 094002, 2003.

