

CZECH TECHNICAL UNIVERSITY IN  
PRAGUE

Faculty of Nuclear Sciences and Physical  
Engineering  
Department of Physics



## Master's thesis

Measurement of the underlying event and its  
impact on top mass determination at the LHC  
accelerator

Matouš Vozák

Supervisor: Mgr. Oldřich Kepka, Ph.D.

Prague, 2017



ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ  
V PRAZE

Fakulta Jaderná a Fyzikálně Inženýrská  
Katedra Fyziky



## Diplomová práce

Měření pozorovatelných citlivých na underlying  
event a studium jeho vlivu na určení hmoty top  
kvarku na urychlovači LHC

Matouš Vozák

Supervisor: Mgr. Oldřich Kepka, Ph.D.

Praha, 2017



**Prohlášení:**

Prohlašuji, že jsem svoji diplomovou práci vypracoval samostatně a použil jsem pouze podklady ( literaturu, projekty, SW atd ... ) uvedené v příloženém seznamu.

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

V Praze dne



*Title:*

**Measurement of the underlying event and its impact on top mass determination at the LHC accelerator**

*Author:* Matouš Vozák

*Specialization:* Experimental nuclear and particle physics

*Sort of project:* Master's thesis

*Supervisor:* Mgr. Oldřich Kepka, Ph.D

*Abstract:* This thesis introduces two separate analyses of parton interactions in both soft and hard processes. At first, ATLAS measurement of additional activity termed underlying event which includes effect of parton radiation and multi-parton interactions is provided using early minimum bias data from proton-proton collisions at 13 TeV. This activity was studied using several observables such as charged particle multiplicity or their transverse momentum in specific regions with different sensitivity to these phenomena. The measured distributions were corrected for detector effects and relevant systematic uncertainties were estimated. The second analysis concerns influence of the colour reconnection modeling on the top mass determination. This QCD effect represents colour exchanges between partons, before the process where colourless hadrons and mezoons are formed and may lead to different final states. Various colour reconnection models embedded in Monte Carlo generator PYTHIA 8 were used to simulate mass of the top quark. By studying the spread between the models it is possible to estimate systematic uncertainty from modeling of colour reconnection. To reduce this uncertainty, potential sensitive observables were studied in order to constrain these models.

*Key words:* underlying event, multi-parton interactions, parton showers, the Large Hadron Collider, detector ATLAS, colour reconnection.



*Název:*

**Měření pozorovatelných citlivých na underlying event a studium jeho vlivu na určení hmoty top kvarku na urychlovači LHC**

*Autor:* Matouš Vozák

*Specializace:* Experimentální jaderná a částicová fyzika

*Projekt:* Diplomová práce

*Supervisor:* Mgr. Oldřich Kepka, Ph.D

*Abstract:* Tato práce obsahuje dvě specializované analýzy, které se zabývají partonovými interakcemi jak v tvrdých tak v měkkých procesech. Nejprve je představeno měření dodatečné aktivity nazývané underlying event s využitím detektoru ATLAS a minimum bias dat z proton-protonových kolizí při těžšíškové energii 13 TeV. Tato aktivita byla studována pomocí několika pozorovatelných jako například počet nabitých částic nebo jejich transverzální hybnost ve specifických regionech s jinou citlivostí na tento jev. Výsledky naměřených pozorovatelných byly opraveny na detektorové efekty a zahrnují systematické nejistoty z různých zdrojů s dominantním příspěvkem simulace materiálu v detektoru. Druhá analýza je věnována vlivu barevné rekonexe na hmotu top kvarku. Tento QCD efekt představuje barevnou výměnu mezi partony před tím, než se zformují do bezbarvých hadronů a mezonů a může vést na změnu finálních stavů. Několik modelů barevné výměny, implementovaných v Monte Carlo generátoru PYTHIA 8, bylo použito k určení vlivu barevné rekonexe na výslednou hmotu top kvarku. Studii rozptylu mezi těmito modely je možné určit teoretickou systematickou chybu způsobenou modelováním barevné výměny. K snížení této chyby bylo dodatečně studováno několik pozorovatelných, jejichž měření bylo možné dané modely zpřesnit.

*Klíčová slova:* underlying event, multi-partonové interakce, partonové spršky, Velký Hadronový Srážeč, detektor ATLAS, barevná rekonexe.



## **Acknowledgement**

I would like to thank my supervisor Mgr. Oldřich Kepka, Ph.D. for his patience, friendly approach, great opportunities and shared wisdom. He has been always an inspiration for me and I value him not only as a great physicist but also as a friend.

My thanks go to my family and friends who have supported me the whole time and kept me sane in the time of need. I sincerely apologize to them if I have not been always nice and talkative.

# Contents

<b>1</b>	<b>Introduction</b>	<b>18</b>
<b>2</b>	<b>Simulation of pp collisions</b>	<b>22</b>
2.1	Cross section calculation . . . . .	24
2.2	Parton showers . . . . .	25
2.3	Multi-parton interactions . . . . .	25
2.4	Hadronization . . . . .	26
2.5	MC modelling . . . . .	26
<b>3</b>	<b>ATLAS detector</b>	<b>28</b>
3.1	ATLAS Inner detector . . . . .	29
3.2	ATLAS calorimeters . . . . .	31
3.3	ATLAS muon spectrometer . . . . .	31
3.4	ATLAS trigger system . . . . .	32
<b>4</b>	<b>Underlying event analysis</b>	<b>33</b>
4.0.1	Plane regions . . . . .	33
4.1	Observables . . . . .	34
4.2	Selection . . . . .	35
4.3	Reconstruction efficiencies . . . . .	36
4.4	Correction . . . . .	38
4.4.1	Event and track weights . . . . .	39
4.4.2	Hit backspace once more method . . . . .	39
4.4.3	Effect of HBOM correction . . . . .	40
4.5	Systematic uncertainties . . . . .	41
4.6	Results . . . . .	45
4.6.1	Leading transverse momentum distribution . . . . .	45
4.6.2	Angular distributions . . . . .	46
4.6.3	$N_{\text{ch}}$ and $\sum p_{\text{T}}$ distributions with respect to $p_{\text{T}}^{\text{lead}}$ . . . . .	46
4.6.4	Mean $p_{\text{T}}$ distribution . . . . .	47
4.7	Summary of UE analysis . . . . .	49
<b>5</b>	<b>Effects of colour reconnection on top mass measurement</b>	<b>50</b>
5.1	Historic motivation for CR . . . . .	51
5.2	Colour reconnection and top mass reconstruction . . . . .	52
5.3	Colour reconnection schemes . . . . .	52

5.3.1	MPI-based models . . . . .	53
5.3.2	Extreme models . . . . .	54
5.3.3	Full scale models . . . . .	55
5.4	Top reconstruction analysis . . . . .	55
5.5	Constrain of colour reconnection . . . . .	58
5.5.1	Activity in jets . . . . .	58
5.5.2	Jet shapes . . . . .	61
5.5.3	CR and UE at 13 TeV . . . . .	64
5.5.4	Fragmentation function . . . . .	64
5.5.5	Pull angle . . . . .	65
5.6	Summary of CR analysis . . . . .	68

# List of Figures

1.1	A sketch of all currently known building blocks. Leptons, quarks and gauge bosons are depicted on the green, red and blue/violet background respectively [1]. . . . .	19
2.1	Three main stages of particle propagation and simulation from a proton-proton collision. After parton-parton interactions, partons are bound to form colourless hadrons and then they are interacting in various layers of detector. Each step is simulated differently. . . . .	22
2.2	Sketch of simulation of proton-proton interaction via MC generators [2]. . . . .	23
3.1	A sketch of CERN accelerators and experiments [3]. . . . .	28
3.2	Computer generated image of A Toroidal Large Apparatus (ATLAS) [4]. . . . .	30
4.1	The sketch of azimuthal plane perpendicular to the beam pipe. The plane is divided into toward, transverse and away regions by using the track with the highest transverse momentum and its difference in the azimuthal angle compared to other tracks $\Delta\phi$ . . . . .	34
4.2	Trigger efficiency $\epsilon_{\text{trig}}$ plotted as a function of the number of selected tracks $n_{\text{sel}}^{\text{BL}}$ . A reconstructed vertex is not required. The statistical uncertainty is shown as a black line, the total uncertainty as a green shaded area [5]. . . . .	37
4.3	Data driven vertex efficiency plotted as a function of selected tracks $n_{\text{sel}}^{\text{BL}}$ (left). The statistical uncertainty is shown as a black line, the total uncertainty as a green shaded area [5]. The multiplicity distribution of selected tracks $n_{\text{sel}}^{\text{BL}}$ (right), with the requirement on leading track transverse momentum $p_{\text{T}}^{\text{lead}} > 1$ GeV (red) and without (blue). . . . .	37
4.4	Tracking efficiency as a function of $\eta$ which is averaged over $p_{\text{T}}$ (left) and $p_{\text{T}}$ which is on the other side averaged over $\eta$ (right). The statistical uncertainty is shown as a black line, the total uncertainty as a green shaded area [5]. . . . .	38
4.5	Result of extrapolation to -1 using 6 iteration and parameterization by a polynomial of second degree in one particular bin of sum of transverse momentum distribution in transverse region with respect to the highest transverse momentum in an event $p_{\text{T}}^{\text{lead}}$ (left). Red line represents the main fit, gray lines are 1000 toy experiments fits. For comparison, the particle level value (tru) is added as a dotted horizontal line with a green band as the statistical uncertainty. Distribution of toy experiments values in -1. Red dotted vertical line represents main fit from initial values which is close to the mean of toys(right). . . . .	40

4.6	Comparison of particle level spectrum (black), track-event corrected spectrum (green) and HBOM corrected spectrum (red) for charged particle density in transverse diff region on the left and sum of transverse momentum in the transverse region on the right both with respect to the highest transverse momentum in an event $p_T^{\text{lead}}$ . The uncertainties are statistical only. . . . .	41
4.7	Comparison of particle level spectrum (black), track-event corrected spectrum (green) and HBOM corrected spectrum (red) for mean transverse momentum $\langle p_T \rangle$ with respect to highest transverse momentum in the event $p_T^{\text{lead}}$ in transverse max region (left) and $\langle p_T \rangle$ with respect to $N_{\text{ch}}$ in the away region (right). . . . .	42
4.8	Comparison of truth (black), track-event corrected (green) and HBOM (red) spectrums for charged particle density on the left and sum of $p_T$ on the right both with respect to $ \Delta\phi $ . . . . .	43
4.9	4 HBOM iteration with 2 <sup>nd</sup> polynomial parametrization and 6 iteration with 3 <sup>rd</sup> polynomial parametrization compared to the default 6 iteration with 2 <sup>nd</sup> polynomial parametrization. $\sum p_T$ with respect to $p_T^{\text{lead}}$ in transverse max on the left and with respect to $ \Delta\phi $ on the right. . . . .	43
4.10	Combination of all used systematic uncertainties for mean $p_T$ plotted as a function of $N_{\text{ch}}$ in transverse min region on the left, where the iteration systematic is included in the non-closure band, and $\sum p_T$ as a function of $ \Delta\phi $ on the right. . . . .	44
4.11	Combination of all used systematic uncertainties for $\sum p_T$ in the toward region on the left and $N_{\text{ch}}$ in the away region on the right both plotted as functions of $p_T^{\text{lead}}$ . . . . .	44
4.12	Normalized distribution of leading particle transverse momentum $p_T^{\text{lead}} > 1$ GeV (left) and $> 5$ GeV. Blue shaded bands represent combination of statistical and systematic uncertainties. . . . .	45
4.13	Distribution of normalized charged particle multiplicities $N_{\text{ch}}$ (left) and scalar sum of transverse momentum $\sum p_T$ (right) plotted as a difference in the azimuthal angle $ \Delta\phi $ with respect to leading particle. Both observables contain data and MC for two different constraints on the leading particle transverse momentum $p_T^{\text{lead}} > 1, 10$ GeV. Blue shaded bands represent combination of statistical and systematic uncertainties. . . . .	46
4.14	Normalized charged particle multiplicity $N_{\text{ch}}$ (left) and scalar sum of transverse momentum $\sum p_T$ (right) plotted as a function of $p_T^{\text{lead}}$ in all studied regions atop each other. . . . .	47
4.15	Normalized charged particle multiplicity $N_{\text{ch}}$ (left) and scalar sum of transverse momentum $\sum p_T$ (right) plotted as a function of $p_T^{\text{lead}}$ in the transverse region. Blue shaded bands represent combination of statistical and systematic uncertainties. . . . .	48
4.16	Normalized mean transverse momentum $p_T$ plotted as a function of $p_T^{\text{lead}}$ (left) and as a function of multiplicities $N_{\text{ch}}$ (right) in the transverse region. Blue shaded bands represent combination of statistical and systematic uncertainties. . . . .	48
5.1	Static representation of colour reconnection (left), dynamic representation of colour reconnection (right) [6]. . . . .	50

5.2	Average number of charged particle transverse momentum with respect to their multiplicity in transverse region. 13 TeV minimum bias data are compared to PYTHIA 8 models where colour reconnection is turned off (red) and where the default model of colour reconnection is turned on (blue). . . . .	51
5.3	Normalized mass spectrum of top quark reconstructed from $b$ jet and two jets coming from the decay of W boson (left), normalized mass spectrum of W reconstructed from the jets originating from the decay of W (right). . . . .	57
5.4	Charged particle multiplicity $n_{\text{ch}}$ in the $b$ jet from top decay (left), charged particle transverse momentum $p_{\text{T}}$ in the higher $p_{\text{T}}$ jet originating from W decay termed lightjet1 (right). . . . .	59
5.5	Distribution of average charged particle transverse momentum in the higher $p_{\text{T}}$ jet originating from W decay termed lightjet1 (left) and an average charged particle transverse momentum with respect to the multiplicity in the lightjet1 (right). . . . .	60
5.6	Multiplicity of charged particles $n_{\text{ch}}$ in the underlying event which are not clustered in jets originating from $t\bar{t}$ decayed quarks nor MPI (left) and their transverse momentum $p_{\text{T}}$ with(right). . . . .	61
5.7	Sum of charged particles in the underlying event transverse momentum $\sum p_{\text{T}}$ . These particles are either not associated with any selected jets (left) or only not associated with jets originating from decay of $t\bar{t}$ system(right). . . . .	62
5.8	Average transverse momentum of charged particles in the UE which are not associated with any selected jets $\langle p_{\text{T}} \rangle$ (left), and also $\langle p_{\text{T}} \rangle$ plotted against the multiplicity of particles in question (right). . . . .	62
5.9	Differential jet shape $\rho(r)$ representing charged particle transverse momentum $p_{\text{T}}$ in the annulus with a distance $r$ from the jet axis, normalized by jet transverse momentum $p_{\text{Tjet}}$ and width of the annulus $\Delta r$ . Left figure is for $r = 0.06$ and right for $r = 0.34$ . . . . .	63
5.10	Differential jet shape representing charged particle multiplicity $n_{\text{ch}}$ in the annulus with a distance $r$ from the jet axis, normalized by the total multiplicity inside jet $n_{\text{ch}}(\text{tot})$ and width of the annulus $\Delta r$ . Left figure is for $r = 0.06$ and right for $r = 0.38$ . . . . .	63
5.11	Number of charged particle in the transverse region plotted with respect to the highest transverse momentum in the event (left), average charge particle transverse momentum in the transverse region shown with respect to their multiplicity. . . . .	64
5.12	Distribution of $F$ related to fragmentation function (left) and transverse profile $f$ (right) for jets with transverse momentum $25 < p_{\text{T}} < 40$ . . . . .	65
5.13	Decay of top quark into $b$ and $W^+$ which further decays into colour connected quarks (left) and the same example with distortion of colours resulting into colour connection of one of the quarks from $W$ with $b$ [7]. . . . .	66
5.14	A diagram showing two jets $J_1, J_2$ in a $\eta - \phi$ difference space and constructed jet pull angle between the pull vector and vector connecting $J_1$ with $J_2$ [7] (left). Pull angle between pull vector constructed from lightjet1 and vector connecting both lightjets for various CR scheme at 13 TeV (right). Only charged particles inside the jet are used to create the pull vector. . . . .	67

5.15	Mean charged-particle multiplicity $N_{\text{ch}}$ (left), and mean sum of transverse momentum $\sum p_{\text{T}}$ (right) plotted as a function of $ \Delta\phi $ with respect to the leading charged particle $p_{\text{T}}^{\text{lead}} > 1$ GeV (top), 5 GeV (middle), and 10 GeV (bottom). The error bars represent statistical uncertainty and the blue band the total combined statistical and systematic uncertainty. . . . .	70
5.16	Mean charged-particle multiplicity $N_{\text{ch}}$ (left), and mean sum of transverse momentum $\sum p_{\text{T}}$ (right) plotted as a function of leading particle transverse momentum $p_{\text{T}}^{\text{lead}}$ for $p_{\text{T}}^{\text{lead}} > 1$ GeV for the transverse (top), towards (middle), and away (bottom) azimuthal regions. The error bars on data points represent statistical uncertainty and the blue band the total combined statistical and systematic uncertainty. . . . .	71
5.17	Mean charged-particle average transverse momentum as a function of azimuthal region charged-particle multiplicity $N_{\text{ch}}$ (left) and as a function of $p_{\text{T}}^{\text{lead}}$ (right), for each of the towards (top) and away (bottom) regions. The error bars on data points represent statistical uncertainty and the blue band the total combined statistical and systematic uncertainty. . . . .	72
5.18	Mean charged-particle average transverse momentum in trans-min (left) and trans-max (right) regions depicted as a function of azimuthal region charged particle multiplicity $N_{\text{ch}}$ (trans-min) and $N_{\text{ch}}$ (trans-max), respectively. The error bars represent statistical uncertainty and the blue band the total combined statistical and systematic uncertainty. . . . .	73

# Chapter 1

## Introduction

It is remarkable how the view on the building blocks of the world has changed in the last century. The increasing knowledge has challenged the mankind to built technologically advanced machines in order to discover new particles and probe their structure. To describe observed experimental results new theories arose which needed to be constrained thus craved for additional measurement eventually leading to massive advance in the field of particle physics. Nowadays most of what is known in particle physics was combined into the unified "package" called **Standard model** (SM). This model comprises all known building blocks of the world and describes how these blocks behave if one of the fundamental interactions is in the play.

### Standard model

In Fig 1.1 there is a complete list of elementary particles which seem to have no further inner structure. The list can be divided into two main blocks, particles with intrinsic angular momentum (spin)  $\pm 1/2$  called **fermions** and **intermediate bosons** for which the value of their spin is an integer. The former are further divided into **quarks** and **leptons**. Bosons are referred to as intermediate because they mediate the interactions between quarks and leptons. Therefore, these interactions can be viewed as an exchange of respective bosons.

Leptons include the electron and its heavier "brothers" muon and tauon together with their respective neutrinos. Except neutrinos they all carry an elementary electric charge. As such, they can participate in the electromagnetic interaction which is mediated by a massless photon. The theory describing this processes is called **quantum electrodynamics** (QED). Leptons also take part in a weak interaction in which massive W and Z bosons are mediated. W carries a charge thus is associated with charge currents whereas Z mediates neutral currents without a charge transfer. Large mass of W and Z leads to their quick decay back to either leptons or quarks. The existence of neutral currents and the building structure of the theory of weak interaction led to unification of weak and electromagnetic interaction under one electro-weak theory.

Quarks are building blocks of all the currently known surrounding matter. Quarks were actually first predicted before their discovery. Except the electric charge they carry also additional charge which was termed colour. Strong interaction is responsible for interactions between coloured particles and is mediated via gluon which carries two different colours at once hence can mediate various colour exchanges. Theory describing the behavior in question is called

**quantum chromodynamics** (QCD). The character of the electromagnetic interaction is quite different from the strong force. Whereas the former is weakening with a greater distance, completely opposite applies for the latter. For this reason it is impossible to study separated colour charged object. Instead, physicists are trying to examine quarks by looking at composite structures containing two quarks **mezons** (pion, kaon,...) and three quarks **hadrons** (neutron, proton, ...). These are detected in a form of "clumps" traditionally referred to as **jets** due to a busy collision environment described later on. It was found out that the inside of hadrons and mezons is actually much more complex and when the resolution of a probe is high enough, instead of having two or three quarks inside, particles look more like an active "soup" of many quarks and gluons.

The last piece of the SM is the Higgs boson whose existence was proven quite recently in 2012. Higgs field is responsible for mass generation of particles through the Higgs mechanism and its detection was a challenge for particle physicists and engineers.

Standard model is surely not the complete theory as it fails in several aspects. For example unification of gravity with the rest of interactions, the fact that neutrinos are within SM massless but their measurement and evidence of their oscillation [8] proved that they must acquire a mass etc. Despite that, the SM is the best description of what is the world made of currently at hand.

A task for particle physicists is to either do more precise measurement within the SM or search for physics beyond it. This thesis is focused on the former and it contains two separated analyses.

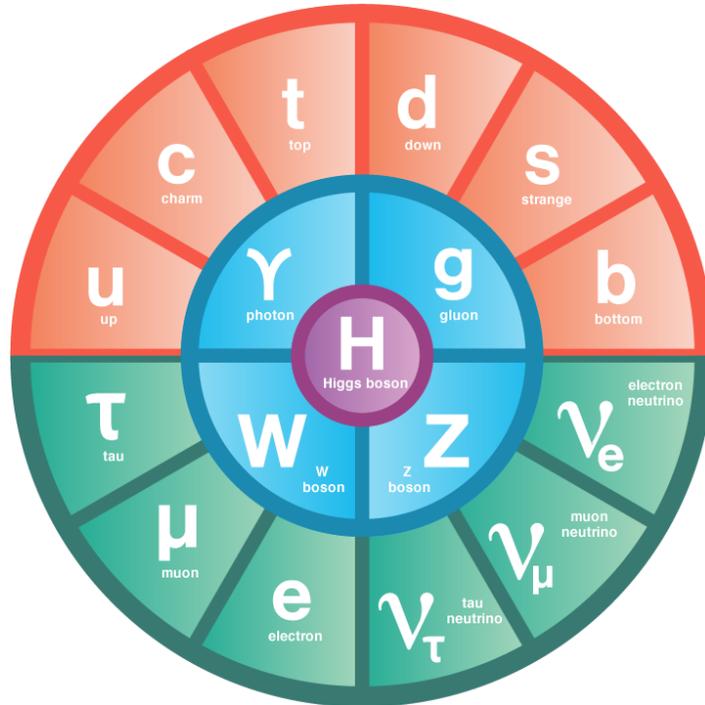


Fig. 1.1: A sketch of all currently known building blocks. Leptons, quarks and gauge bosons are depicted on the green, red and blue/violet background respectively [1].

## Underlying event

The first analysis concerns the ATLAS detector measurement of particle activity in proton-proton collisions at the **Large Hadron Collider** (LHC) at 13 TeV. As mentioned above, protons have an inner structure and consist of quarks and gluons collectively termed partons<sup>1</sup> thus their interactions can be viewed rather as an interplay between their constituents. When protons collide, signals in the form of collimated particle jets are detected. These jets are manifestations of quarks coming from a hard parton-parton interaction in which the collided partons may carry a significant amount of proton's momentum. This process which can be calculated from the first principle is underlined with an additional activity originating from the dynamic structure of proton, properties of particles and other effects. The phenomena contributing to the underlying activity of the hard process are collectively termed **underlying event** (UE).

One of the example of UE are **multi-parton interactions**. As there are more partons in the proton it is possible for additional interactions between more parton pairs to occur. Another example which contributes to the additional activity are parton showers from multi-parton interactions. These QCD showers are created due to the presence of moving partons with colour charge similarly to QED in which electrons may radiate photons. However, in this case gluons are capable of radiating additional gluons which can further radiate as well, eventually creating a shower of partons. The radiation can take place either before the proton collide or after and is referred to as the **initial** or the **final state radiation** respectively. Once partons are taken out of proton during hard interaction or MPI, the proton dissociates and loses its structure. Nevertheless, some of the remnant partons are kept "almost intact" and keep traversing in the initial way of the proton. They are hard to detect as they typically move in the direction of accelerator beam pipe where a detection system can not be placed. These particles are termed as **beam-beam remnants**, and they are commonly included in UE.

Measurement in this thesis is the extension of ATLAS 900 GeV and 7 TeV UE measurement [9] to new LHC center of mass energy 13 TeV in order to see how the character of UE changes with the collision energy. As the analysis used early Run II data it was important to provide first results as quickly as possible. The early results were summarized in the ATLAS public note [10]. The work presented here is a continuation of the preliminary result. Correction procedure to remove detector effects has been carried out and the corresponding systematic uncertainties estimated. The results were published in the beginning of 2017 [11] and can be found in the HepData format [12] with the corresponding Rivet analysis at [13].

## Top mass reconstruction

The second analysis is focused on studying one particular element in the modeling of UE that is the **colour reconnection** effect. This phenomenon allows possible colour exchanges between partons which occur before the hadronization process and as such can influence the final state particles. There are different interpretation of colour reconnection among Monte Carlo generators and even different models within the individual generators. For example

---

<sup>1</sup>Partons as they carry a part of proton's momentum

13 different colour reconnection models are currently embedded in PYTHIA 8. In general all models share the common approach and include two main steps. In the first step systems which should participate in colour exchanges are determined and in the second the actual colour exchange takes place based on various limits.

This exchanges can occur between all colour parton in an event thus may lead to interaction of both particles within jets and particles from underlying event eventually resulting into potential leakage of hadrons out of the jet cone. Therefore, jet properties such as number of particles inside jet or jet transverse momentum are influenced and the precision of analyses utilizing jets might be decreased depending on the observable of interest.

Measurement of top quark mass is one example of such analysis. Top decays dominantly to bottom quark and  $W$  boson which can further on decay into a quark pair. Mass of the top quark in this particular process is dependent on three jets which constituent partons may have undergone colour reconnection thus the top mass is obscured.

Currently the systematic uncertainty from the colour reconnection modeling often represents a small contribution to the total systematic uncertainty for experiments which are measuring the top quark mass. However, this uncertainty is commonly taken from comparison of models which were proved to be unrealistic and is most likely underestimated. As the reconnection phenomenon is not very well understood, better approach would be to study all possible models which exhibit not unrealistic character and then estimate the uncertainty from the spread of all these models. This thesis studies values of top quark mass reconstructed from  $t\bar{t}$  system for range of PYTHIA 8 colour reconnection models. In addition, plenty of observables were examined in order to look for possible constrains for these models.

## **Thesis structure**

The thesis goes as follows. Firstly a short introduction to simulation of proton-proton collision utilizing Monte Carlo generators in the chapter 2 is given which continues with the brief mention about the Large Hadron Collider and description of one of its main detector ATLAS in the chapter 3. The underlying event analysis at the 13 TeV with ATLAS detector including selection, correction procedure, list of relevant systematics and finally description of results is in the chapter 4. The last chapter is dedicated to studies of colour reconnection effects on the top quark mass measurement where at first colour reconnection is explained in more details, followed by the manifestation of its existence. Then the motivation to study colour reconnection connected with its theoretical uncertainty is provided continuing with the description of PYTHIA 8 colour reconnection models. The last part summarizes the result of top mass by various colour reconnection schemes and explored observables to sensitivity of colour reconnection with either positive or negative result.

## Chapter 2

# Simulation of pp collisions

Simulations play an essential part in particle physics due to the lack of information about events at early stages before particles reach detection systems. There is no proof that a theory is wrong without an experiment and no understanding of an experiment without a theory. Thus simulations provide the complementary theoretical side to particle experiments.

It is impossible to determine which event will occur as the next one and which particles are going to be measured in a detector. Anything that is allowed to happen might happen next. Therefore, to simulate a particle collision is a complex procedure including solving n-dimensional integrals. Monte Carlo (MC) generators are very efficient in solving these integrals thus are put into use to describe for example proton-proton collisions at particle accelerators. This chapter will be focused on these collisions as this analysis was done in the collaboration with the LHC experiment which accelerates and collides protons and heavy nuclei.

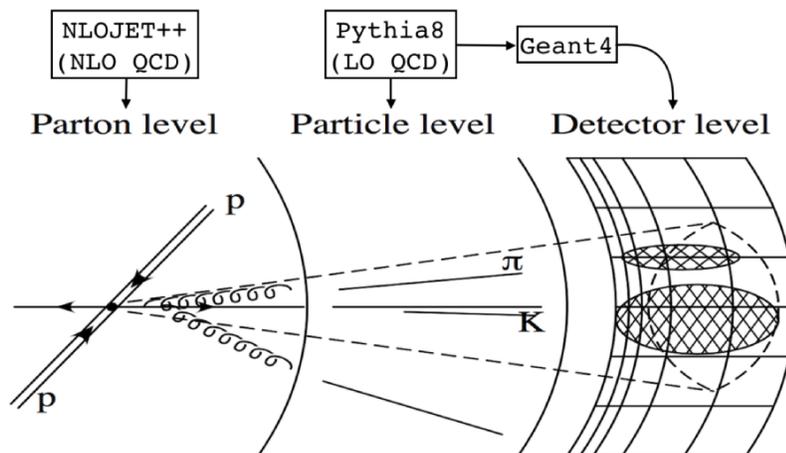


Fig. 2.1: Three main stages of particle propagation and simulation from a proton-proton collision. After parton-parton interactions, partons are bound to form colourless hadrons and then they are interacting in various layers of detector. Each step is simulated differently.

In Fig. 2.1 there are three main stages of event evolution. At first, there is a so called **parton**

**level** where partons from protons are interacting together producing new particles and flying to all directions from the collision point. As some of them carry a colour charge they bundle together to form colourless particles. This stage is denoted as a **particle level**. Particles are observed at a **detector level**, once they reach a detector they leave signals in different layers from which their trajectories can be reconstructed (more about detection in chapter 3). From the definition it can be seen that parton and particle levels are purely simulated. Compensation factors must be derived which are used to correct data to represent the event before it was obscured by the detector so various experiment can compare the results. Simulation of proton-proton interaction is based on several steps Fig. 2.2. At the core of the event there is a process with a significant amount of energy transfer between interacting partons termed hard process from which heavy particles might emerge. This hard component is accompanied with the initial state or the final state radiation (ISR/FSR) of additional partons which can take place either before the interaction or after. Radiated partons can participate in further radiation eventually creating parton showers (PS). Not only one pair of partons may interact and create a hard process but collision environment may be filled with additional interactions between parton pairs<sup>1</sup> leading to effect called multi-parton interaction (MPI). All together PS, MPI and beam-beam remnants (BBR) are collectively termed as underlying event because they "lie under" the hard component<sup>2</sup>. During the hadronization partons form colourless objects which may be stable or may decay into stable particles. Short introduction of individual steps in event simulation is provided in the following sections.

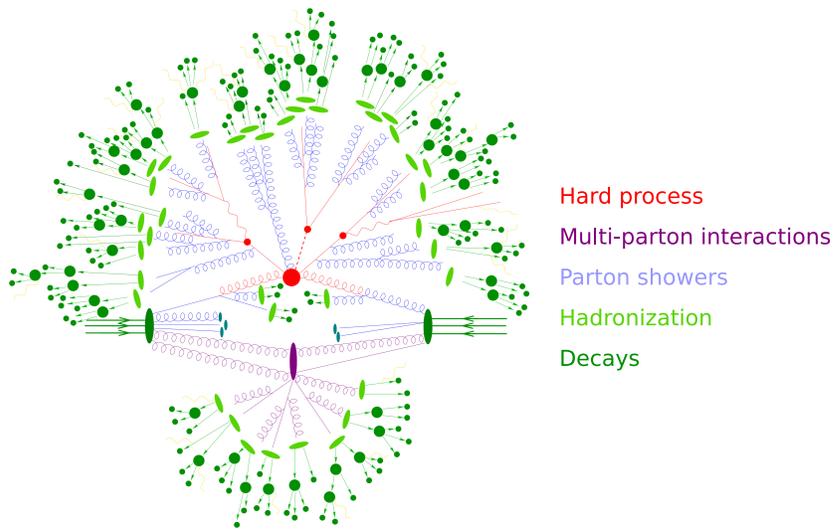


Fig. 2.2: Sketch of simulation of proton-proton interaction via MC generators [2].

<sup>1</sup>Not to be confused with a pile-up which is due to interactions of more proton pairs not parton pairs from the same pair of protons

<sup>2</sup>The definition of underlying event is ambiguous as PS may also represent an additional hard component thus the structure of underlying event would be a mix of both hard and soft component

## 2.1 Cross section calculation

As described above any interactions that can happen may happen, however, each one of them is characterized by their probability. This probability is "no ordinary" probability but has units in [area<sup>2</sup>] and can be interpreted in a following example. If there are two particles with a finite volume going towards each other, the probability of them colliding is proportional to the cross section of the particles. If they are interacting for example electromagnetically the area over which the force takes place is much larger than their cross section. Nevertheless, this unorthodox probability of particles to interact was termed **cross section**  $\sigma$  and is one of the most important observable in many analyses.

Cross section can be calculated from a particle flux, phase spaces and **matrix element** (ME). In general for QCD process  $ab \rightarrow n$  where partons  $a$  and  $b$  interact creating  $n$  partons,  $\sigma$  can be written as

$$d\sigma_{ab \rightarrow n} = \int \int \int \frac{1}{F} \prod_{i=1}^n \frac{d^3 p_i}{2E_i (2\pi)^3} \delta^4(p_a + p_b - \sum_{i=1}^n p_i) |\mathcal{M}_{ab \rightarrow n}|^2, \quad (2.1)$$

where  $F$  represents the initial flux<sup>3</sup>, phase spaces for individual particles are represented by the momentum differentials  $dp_i$  with a chosen normalization  $(2E_i(2\pi)^3)$  and delta functions  $\delta$  indicates energy conservation in the interaction. The last part is a square of matrix element which contains the physics of the process, meaning what process is being calculated (for example interaction  $qq \rightarrow qq$  through gluon). ME can be calculated using Feynman rules which are derived from a Lagrangian of the theory. It can then be written as a sum over all possible interactions which are for simplicity interpreted through Feynman diagrams  $\mathcal{F}$ .

$$\mathcal{M}_{ab \rightarrow n} = \sum_i \mathcal{F}_{ab \rightarrow n}^i \quad (2.2)$$

To calculate precisely the value of  $\sigma$  is analytically impossible, however, it can be decomposed into expansion in parameter of strong interaction  $\alpha_s$ <sup>4</sup>. Most of the generator employ only **leading order** (LO) term (the first term in the expansion) e.g. PYTHIA but with a rising demand on the precision, there are generators such as POWHEG which for generating ME uses **next to leading order** (NLO) term. The reason for not using higher terms such as NLO or NNLO right away is the increasing number of diagrams thus the complexity of the calculation which becomes cumbersome. ME are used to calculate hard processes in which is the perturbation theory still applicable.

To obtain total proton-proton cross section at the LHC, the result from (2.1) needs to be convoluted with parton distribution functions (PDF)  $f$  and integrated over  $x_a, x_b$  which are fractions of proton's momentum carried by parton  $a$  in A proton and  $b$  in B proton.

$$\sigma = \sum_{a,b} \int_0^1 dx_a dx_b f_a^A(x_a, \mu_F) f_b^B(x_b, \mu_F) d\sigma_{ab \rightarrow n} \quad (2.3)$$

PDFs are not provided in advance and need to be extracted from the measurement. They dependent on the normalization scale  $\mu_F$  which is in general arbitrary.

---

<sup>3</sup>Initial flux can be written as the number of beam particles traversing unit area per unit time times the number of target particles per unit volume

<sup>4</sup>For large transfer momenta  $Q^2$  between partons  $\alpha_s \rightarrow 0$

## 2.2 Parton showers

Similar to radiation of photons by e.g. electron in QED, quarks are also capable of radiating gluons. The difference is that gluons can interact with each other thus a radiated gluon can further radiate more gluons and so on. This results into parton showers which brings additional particles in the final states. As mentioned in the previous section this might cause complications. Luckily  $\sigma$  can be decomposed. In the simple example of three partons  $ggg$  in the final state, the process can be viewed as a production of  $gg$  pair (for example from  $gg \rightarrow gg$ ) followed by a radiation of additional gluon. PS model uses following approach to introduce the showering into  $\sigma$ .

The probability that something happens at the time  $T$  can be written as probability that something happens at time  $T$  for the first time times the probability that it did not happen yet.  $\sigma$  can be written in the same way using Sudakov factor [14] but for example in a transverse momentum  $p_T$ <sup>5</sup> instead of time. The evolution of a shower thus goes downward in  $p_T$  as more particles are radiated, the energy is redistributed. The PS stops once the evolution reaches  $\sim 1$  GeV which is a cut where hadronization model takes place as  $\sigma$  diverges in soft regions where  $p_T$  is low.

As there are charged partons even before the actual interaction the shower can happen also at the early stage. The description of these initial state radiation is more complicated then the case with the final state radiation due to the fact that parton distribution functions need to be included. As the hard scale is given by ME and radiating incoming partons have to fit the scale of the hard process, it is more effective to start ISR from the ME and go backwards.

## 2.3 Multi-parton interactions

Possible interaction of more than one pair of partons between protons should not come as such a surprise due to the dynamic structure of hadrons. Another hint came from the measurement of total hadronic cross section which consists of an inelastic and an elastic part with contribution about 1:2 respectively. QCD  $2 \rightarrow 2$  interaction producing two jets in first approximation represents the inelastic process but when the theoretical value of this interaction cross section  $\sigma_{int}$  is compared to the experimental  $\sigma_{tot}$  it actually exceeds it. This excess was interpreted by MPI mechanism, where each event has on average more than one parton-parton interactions. For example PYTHIA uses Sjöstrand-Zijl model [15] in which the average number of interactions is proportional to the overlap of the hadron matter reaching saturation for central collisions.

MPI model deals with soft divergence of  $\sigma$ , which would mean infinite MPI interactions, by introducing phenomenological dampening factor to reduce the contribution of low  $p_T$  exchanges to  $\sigma$ . These low  $p_T$  gluons<sup>6</sup> have large wavelength and their phase space is across the whole hadron. As hadrons are colourless objects, contributions from all colour low  $p_T$  contributions must add to zero.

---

<sup>5</sup>Different MC generators use different scales such as transverse momentum, virtuality, mass, angle, where each has its highs and lows.

<sup>6</sup>QCD  $2 \rightarrow 2$  process is dominated by gluon exchange Feynman diagram.

## 2.4 Hadronization

When particles reach a detector, they are already in the form of colourless hadrons. Therefore, there is a missing link between so far described parton's evolution and the detection, describing creation of hadrons. Unfortunately, this part can not be calculated using the perturbation as it happens typically at scales of 1 GeV and lower. Various models are used instead from which the **Lund string model** and **Cluster model** are the most famous ones.

In the string model hadrons are produced by breaking the string stretched between partons. In the simplest example having quark antiquark  $q\bar{q}$  pair moving away from each other, strong interaction "span a string between them". The tension of the string represents the potential energy of the system. Eventually the string breaks up by a production of another  $q\bar{q}$  pair from the vacuum using quantum tunneling mechanism. In general string may break up on several places leading to production of multiple hadrons.

In the cluster model, all gluons are split to  $q\bar{q}$  pairs at the end of the parton showers. By keeping track of colour flow through the event,  $qq\bar{q}$  pairs form clusters representing the basic unit from which hadrons are made.

## 2.5 MC modelling

In the first analysis of this thesis, the measurement are compared with MC generators typically used in ATLAS in particular PYTHIA 8 [16], HERWIG 7 [17], EPOS [18]. PYTHIA 8 was also used in the second analysis to study effect of colour exchange on the top mass measurement. For this reason a brief description of the mentioned generators is in place.

### PYTHIA

PYTHIA with its last release PYTHIA 8 is a commonly used generator for LHC physics. It is capable of either generating a relevant matrix element or using externally generated elements with further merging with parton showers. PYTHIA separates total inelastic cross section into diffractive and non-diffractive (ND) components. Diffractive can be further on divided into single and double diffractive (SD,DD) where either one or both protons are dissociated. This type of events exhibit no colour exchange between proton remnants and forms characteristic regions in the detector called gaps in which no signals are detected. In non-diffractive components both protons dissociate as well but allowed colour exchanges lead to particle production in the whole rapidity range.

As already mentioned PYTHIA uses Sjöstrand-Zijl model for simulation of MPI and Lund string hadronization model to form colourless objects from outgoing partons. Generator posses 13 models to simulate colour reconnection mechanism which is going to be described in details in the chapter 5. PYTHIA includes also decays of unstable particles.

There are several free parameters from the models which need to be set by hand, but ATLAS provides several tunes which are in a common use. They were made using preferably the data from the LHC at 7 and 8 TeV. These are minimum bias tunes A2 and underlying event tunes A14 and MONASH with the description in the Tab. 2.1.

Generator	Version	Tune	PDF	Focus	From
PYTHIA 8	8.185	A2	MSTW2008 LO	MB	ATLAS
PYTHIA 8	8.185	A14	NNPDF2.3 LO	UE	ATLAS
PYTHIA 8	8.186	MONASH	NNPDF2.3 LO	MB/UE	Authors
HERWIG 7	7.0.1	UE-MMHT	MMHT2014 LO	UE/DPS	Authors
EPOS	3.4	LHC	—	MB	Authors

Tab. 2.1: MC models used for underlying event analysis at the LHC at 13 TeV. Used PYTHIA tunes are focused on describing both underlying event and minimum bias whereas HERWIG 7 is dedicated to underlying event and double parton scattering distribution. EPOS provides a very good description of 13 TeV minimum bias data [19].

## HERWIG

HERWIG is another example of frequently used generator for collider physics with its last release HERWIG 7. Instead of the string model, EPOS uses cluster model in which a track of every colour is kept to form colourless clusters. These may later on decay into stable hadrons. HERWIG MPI model uses similar grounds as in PYTHIA but with different and simpler parametrization and for the description of possible colour exchanges before the hadronization stage, the generator utilizes a disrupt mechanism.

## EPOS

EPOS is a MC generator dedicated to simulation of soft-QCD/cosmic ray air showers. It utilizes Gribov-Regge theory which is an effective field theory capable of simulating simultaneously hard and soft component. The requirement of standard PDFs is absent in EPOS due to the combination of collective flow modeling taken from the heavy ion and nuclear physics and string based hadronization. EPOS is not strictly dedicated to UE as it lacks the hard component, however, it turned out to be successful in description of 13 TeV minimum bias data [19].

## Chapter 3

# ATLAS detector

The desire of particle physicist is to understand building blocks of the universe. Direct approach is to take particles and smash them either into each other or into a fixed target made out of various chemical elements. However, to be able to "see" most from the particle zoology, it is necessary to accelerate them as Nature does not provide particles with sufficient energies<sup>1</sup>. To produce massive particles such as Z boson or to look deeper into the structure of particles, complex facilities are built in order to accelerate particles to required energies.

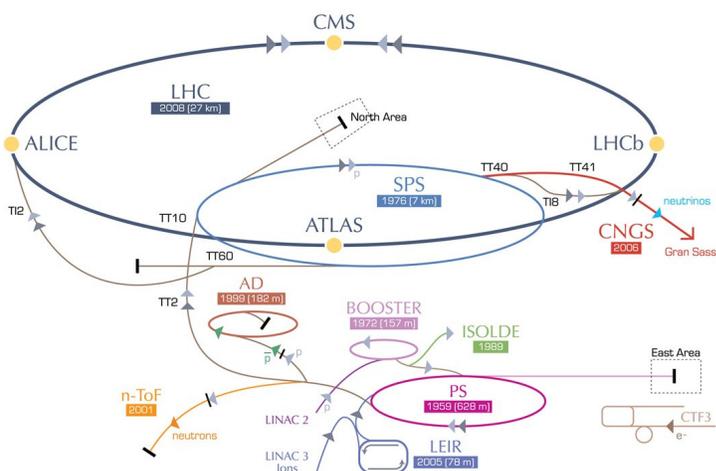


Fig. 3.1: A sketch of CERN accelerators and experiments [3].

### Large Hadron Collider

One of the mentioned facilities is the LHC situated in Geneva which is currently the largest circular accelerator in the world with its 27 km circumference. It is capable of colliding both protons and heavier nuclei. The LHC acceleration cycle of protons shown in Fig. 3.1 is quite

<sup>1</sup>Actually it does, there are high energetic particles coming from space, however, humanity have limited resources and time to detect them.

complex but starts with a simple bottle of hydrogen. Firstly, atoms of hydrogens are stripped of electrons using stripping foils. The remaining protons go through a linear accelerator Linac 2 where they gain energy of 50 GeV. From linac protons go to the Booster where they are further accelerated to 1.4 GeV. Next acceleration up to 25 GeV takes place in the Proton Synchrotron (PS) followed by the Super Proton Synchrotron (SPS) which protons leave with 450 GeV. The SPS is the last link before the LHC in which protons circulate in two separate beam pipes. In the end of accelerating cycle, proton has the energy of 6.5 TeV in center of mass system.

The probability of proton interacting with each other is quite small due to their tiny size. Typical approach how to enhance the probability of "seeing" proton interacting is to increase the number of protons. Therefore, protons are "squeezed" together to form structures called bunches. To provide some numbers, one bunch contains  $1.15 \cdot 10^{11}$  protons in the beginning of the fill and there are 2808 bunches in total. The performance of the LHC is quantified by a **luminosity**  $\mathcal{L}$  which is given by a beam characteristics. It also represents a proportional factor between the number of interactions per unit of time and  $\sigma$ .  $\mathcal{L}$  is decreasing over time due to the reducing number of protons in bunches as they collide. The LHC represents the amount of data collected by integrating  $\mathcal{L}$  over period of time during which the machine was running and collision took place. The unit of the integrated  $\mathcal{L}$  over time is  $\text{m}^{-2}$  but more used are barns  $\text{b}^2$ . One of the main reasons for the LHC upgrades is to obtain higher luminosities hence more data to access less frequent processes.

As mentioned above the purpose is to smash particles together thus the beam pipes are made to intersect at a given location. Huge detectors are built around the intersection to capture or at least detect the particles coming from interactions via various detection principles. There are 4 main detectors at the LHC - ATLAS, CMS, ALICE and LHCb. First two are general purposes detectors whereas ALICE is dedicated to heavy ion collision and LHCb is focused on study of parity breaking phenomenon. The data used for this analysis were collected using the ATLAS detector, therefore, only the main parts of this detector will be specified. The whole ATLAS detection system can be seen in Fig. 3.2. Different behavior and interactions of particles requires such a monstrous size of the detector to determine their presence.

### 3.1 ATLAS Inner detector

The innermost part of ATLAS is the inner detector (ID) in which most of the detection materials are made of silicon. ID consists of Insertable B-layer (IBL), Pixel, semiconductor tracker (SCT) and transition radiation tracker (TRT). These parts come in two geometries - discs (called end caps) and barrels around the beam pipe with attached detection systems on the surfaces to obtain almost the full coverage in azimuthal angle  $\phi$  and pseudorapidity<sup>3</sup>  $|\eta| < 2.5$ .

Increase in  $\mathcal{L}$  demanded better radiation resistance of the material in order to prevent fast

---

<sup>2</sup> $10^{-28} \text{ m} = 1 \text{ b}$ , which is a large unit, much common are for example  $\mu\text{b}$

<sup>3</sup>Instead of using second angle  $\theta$  to give full spatial angle  $\Omega$ , pseudorapidity  $\eta$  which is directly related to  $\theta$  through relation  $\eta = -\ln \tan \frac{\theta}{2}$  is preferred observable due to its convenient behavior under lorentz transformation

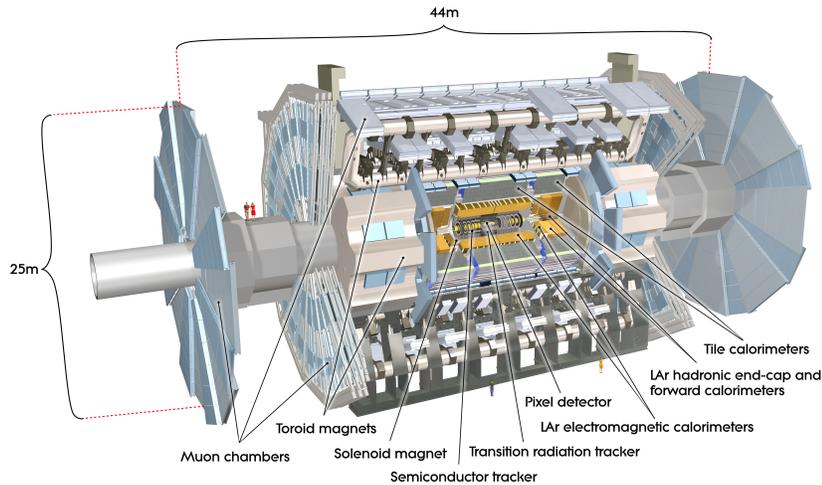


Fig. 3.2: Computer generated image of A Toroidal Large Apparatus (ATLAS) [4].

degradation. Newly developed IBL [20] was inserted as the closest part of the detector to the beam pipe in the last upgrade before Run II. For a signal detection IBL uses hybrid pixel detectors. Thanks to its low material density and closer position to the collisions, IBL increases quality of particle detection<sup>4</sup>.

Next part of the ID is the Pixel which consists of two barrels and two end caps each with three discs. In this part of ATLAS particle detection is done through silicon pixel detectors. Pixel is followed by SCT which has four barrels and nine discs at each side, where silicon micro strip sensors are used for the detection. The farthest part of the ID from the beam pipe is TRT which comes in end cap with 18 discs per side and one barrel. In the opposition to the parts described so far, TRT uses many gas filled tubes to produce transition radiation and provides additional information on particle position thus enhance the precision of particle trajectory reconstruction. Furthermore, TRT is also capable of discrimination between electrons and pions due to their different energy deposition in the tubes.

The whole ID is embedded in the 2 Tesla magnetic field produced by a solenoid. This magnetic field forces charged particles to change their trajectories via Lorentz force. The purpose of the ID is then to reconstruct propagation of charged particles through the detector from signals which they leave behind in the layers of the ID as they move away from the collision. Particle's trajectory is often referred to as **track** thus the method of track reconstruction is termed **tracking**. Tracking is a multi-step procedure where various algorithms are used to reconstruct a track from fit of detected hit signals [21]. Each track is characterized by several parameters.

- $q/p$  charge of the track over its momentum.
- $\theta, \phi$  polar and azimuthal angles of the track point of the closest approach to the origin respectively.

<sup>4</sup>In particular impact parameter reconstruction

- $d_0, z_0$  transversal and longitudinal distances of the track point of the closest approach to the ATLAS geometric origin respectively, also termed as impact parameters.

Unfortunately, some particles might get lost in the supporting material or might be badly reconstructed. Therefore, it is necessary to estimate this inefficiency of detector reconstruction. As there is no measured information about particles between collision and the point where they reach the first layers of ATLAS, full detector simulations are used to provide the efficiencies.

Another task for the reconstruction is to determine the position of interaction point referred to as vertex. Large number of protons in the bunches may lead to several proton-proton interactions in a small time and space window thus occurrence of several vertices. It is important for the analysis to know the assignment of the tracks to vertices as some of the signal seen in the detector may actually represent a background originating from other proton-proton interaction. Similar to the tracks, finding and reconstructing vertex is associated with various algorithms. This brings additional systematic uncertainty to the measurement. For this reason, the efficiency of vertex reconstruction is also another important topic for studies.

## 3.2 ATLAS calorimeters

To measure energies of particles, ATLAS has the Liquid argon calorimeter (Lar) and Tile calorimeter. Lar calorimeter follows right after the ID and its purpose is to capture electrons, gamma radiation, jets and missing energy  $E_T^{\text{miss}}$  carried away by hardly detectable neutrinos. It has a sandwich structure of absorbers and active materials where the former serves to capture the particles in a bulk and the latter to transform the energy of the captured particle into a signal. Absorbers are in general made out of dense materials to enhance  $e^-e^+$  pair production to create electromagnetic cascade whereas opposite applies for active materials which utilizes scintillation, ionization and bremsstrahlung. For example Lar in particular uses lead as the absorber and liquid argon as the active material. Regarding the calorimeter overlay, Lar also contains barrel and end-caps geometry which are accompanied by forward calorimeter to cover higher  $\eta$ .

Tile calorimeter is dedicated mainly for detection of hadrons through strong interaction, however, the electromagnetic interaction occurs as well due to the existence of processes such as  $\pi \rightarrow \gamma\gamma$ . The name "Tile" comes from calorimeter's active material which is a scintillator in a shape of tile. As the absorber, ATLAS hadronic calorimeter uses iron. Tile comes in a barrel and end-cap geometry as well.

## 3.3 ATLAS muon spectrometer

Muons are high penetrating particles and are able to pass the calorimeters without being captured. They occur in many decay processes such as  $H \rightarrow 4l$  thus the ability to detect them was a matter of great importance for ATLAS. For this purpose the outermost detection system of ATLAS is the muon spectrometer which consists of two end-cap toroidal magnets and one barrel toroidal magnet to bend muon trajectories. Monitored drift tubes (MDT) and Cathode strip chambers (CSC) are used to reconstruct muon tracks. Spectrometer also possesses additional Resistive plate chambers (RPC) and Thin gap chambers for triggering.

### 3.4 ATLAS trigger system

It is impossible for the current technologies to collect information from all collisions as the event rate is about 40 MHz. Trigger system is an essential part of the ATLAS detector which reduces this rate to 1kHz by selecting only the "interesting" events. The ATLAS trigger system contains Level-1 (L1) trigger and High level trigger (HLT). The L1 represents the hardware part of the trigger and the selection informations are provided by individual detector systems in particular by the calorimeter and muon detectors. The HLT on the other hand consists of a computer farm and represents the software part of the trigger. The most essential trigger part used in the first analysis is **minimum bias trigger scintillator** (MBTS) dedicated to the low luminosity LHC runs. These contain production of minimum bias events which are events with as least constrain as possible. The acceptance of the MBTS is for the outer ring  $2.08 \leq |\eta| < 2.76$  and for the inner ring  $2.76 \leq |\eta| \leq 3.86$ .

# Chapter 4

## Underlying event analysis

As mentioned in the chapter 2 underlying event consists of initial and final state radiation (ISR/FSR), multi-parton interactions (MPI) and beam-beam remnants (BBR). The presence of this additional activity hinders to obtain better precision as it can contaminate sensitive measurements such as measurement of low energetic jets, reconstruction of top mass etc. UE phenomenological models introduce free parameters to generators, therefore, new data can be used to provide a better tuning of these parameters. To summarize, motivation of the analysis is to enhance precision measurements by providing tuning informations and of course trying to understand character and behavior of underlying event at 13 TeV.

Several measurements regarding the additional activity in question were performed. In particular measurement of hadron interactions with CDF detector at 1.8 TeV [22] and also 1.9 TeV [23] where UE was studied in Drell-Yan process. CMS collaboration provided the measurement of UE at 7 TeV [24] and later on also at 13 TeV [25].

### 4.0.1 Plane regions

The analysis uses the same approach to study UE as in ATLAS studies at center of mass energies 900 GeV and 7 TeV [9] and introduce few regions sensitive to the UE. In each event particle with the highest transverse momentum termed **leading particle** divide the plane perpendicular to the beam pipe by difference in azimuthal angle  $|\Delta\phi|$  from leading particle into 3 main regions.

- Toward region  $|\Delta\phi| < \pi/3$
- Transverse region  $\pi/3 < |\Delta\phi| < 2\pi/3$
- Away region  $2\pi/3 < |\Delta\phi|$

As can be seen from Fig 4.1, by definition the hardest scattering event contributes mostly to the toward and the away regions leaving the transverse region occupied solely by UE. Both transverse regions can be further subdivided and studied separately. As suggested in [26] one of the transverse region could be more active (transverse max) in the case of three jet situation and contain MPI with FSR whereas the other (transverse min) mainly MPI. Subtraction of these two transverse region (transverse diff) could provide information about activity solely from typically harder radiation.

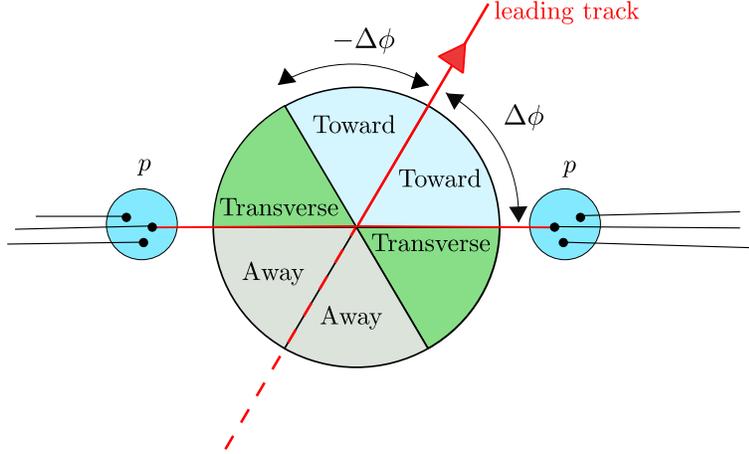


Fig. 4.1: The sketch of azimuthal plane perpendicular to the beam pipe. The plane is divided into toward, transverse and away regions by using the track with the highest transverse momentum and its difference in the azimuthal angle compared to other tracks  $\Delta\phi$ .

## 4.1 Observables

All measured observables defined at particle and detector level are in the Tab. 4.1. At the particle level these observables are constructed from primary charged particles. These are coming either from collision and have the mean life time  $\tau_{prim} \geq 0.3 \cdot 10^{-10}$  s or come from decays of particles with  $\tau \leq \tau_{prim}$ . Strange baryons (representing mostly by  $K^0$  and  $\Lambda^0$ ) are excluded from the measurement due to their high dependence on generator modeling. At the detector level the observables are constructed from selected tracks (more about selection in section 4.2). Charged particle density is defined as the number of charged particles  $N_{ch}$  in a given region divided by the region's  $\eta - \phi$  area. The same applies for sum of transverse momentum  $p_T$  of all particles in the region. These two observables are plotted with respect to the transverse momentum of the leading particle  $p_T^{lead}$  and also with the respect to the difference in azimuthal angles  $|\Delta\phi|$ . Last observable is the mean particle transverse momentum which is constructed on event-by-event basis and plotted against the number of particles  $N_{ch}$  and also  $p_T^{lead}$ .

	$\langle d^2 N_{ch} / d\eta d\phi \rangle$	$\langle d^2 \sum p_T / d\eta d\phi \rangle$	$\langle \text{mean } p_T \rangle$
Particle level	Mean number of stable charged particles per unit $\eta\phi$	Mean scalar $p_T$ sum of stable charged particles per unit $\eta\phi$	average $p_T$ of stable charges particles
Detector level	Mean number of selected tracks per unit $\eta\phi$	Mean scalar $p_T$ sum of selected tracks per unit $\eta\phi$	average $p_T$ of selected tracks

Tab. 4.1: Observables at the detector and particle level.

## 4.2 Selection

The experimental data used for the analysis include early proton-proton Run II LHC data at center of mass energy  $\sqrt{s} = 13$  TeV. The total integrated luminosity of all samples is  $1618.4 \mu\text{b}^{-1}$ . The average number of interactions per bunch crossing often referred to as pile-up ranges between the samples from 0.3 % to 3%. An event and track selection was applied to reduce contributions from background effects (cosmic radiation, detector noises, ...).

### 1. Event selection

- Good run list applied (GRL) to account only events with good luminosity.
- Hit on either side of Minimum bias trigger scintillator (MBTS) to fire L1 trigger, this trigger is passed through the HLT and the event is always recorded.
- Presence of a primary vertex which is the vertex with the highest  $\sum p_T$  of associated tracks.
- No secondary vertex with four or more tracks to reduce the small pile-up contribution.
- At least one good track in the event with a  $p_T > 1$  GeV and for  $|\Delta\phi|$  observable also  $p_T > 5, 10$  GeV. This requirement basically changes the environment and event's topologies from isotropic to more anisotropic to allow separation of UE from the hard component.
- Number of events after above requirement are listed in the Tab. 4.2.

Run	All	GRL	Trigger	Vtx	$p_T^{\text{lead}} > 1\text{GeV}$	L [ $\mu\text{b}^{-1}$ ]	$\mu$
267358	8665704	7272839	4545977	3775181	2725118	62.0	0.3%
267359	12292491	10504126	6309244	5387438	3880743	89.1	0.7%
267360	12558421	8309283	5178427	4487985	3211685	74.1	3%
267367	17111228	16469031	8438677	7237217	5172711	120.8	3%
267385	71217242	71095463	48021030	41001542	29358370	690.4	3%
267599	105611891	105610048	41221102	34963098	25063706	582.3	1%-3%
Total	227456977	219260790	113714457	96852461	69412333	1618.4	-

Tab. 4.2: Number of events in used runs after selection. Each column has also applied selections from all previous left columns. "All" represents number of events before any selection. Numbers in third column are after requirement on good run list. Trigger selection uses HLT\_noalg\_L1\_MBTS\_1. Number of events after requirement on primary vertex and cut on transverse momentum of leading track  $p_T^{\text{lead}}$  are presented as well. Last two columns represent integrated luminosity and pile-up fraction in individual runs.

As can be seen from the Tab. 4.2 around 230 million events were used in total. This number was reduced almost by half with the trigger requirement. Another 17 millions were lost due to the vertex condition and the last selection on  $p_T^{\text{lead}}$  left around 70 million events. Therefore, around 30% events survived all the event selection criteria.

### 2. Track selection

- $p_T > 0.5$  GeV and  $|\eta| < 2.5$ .

- At least 1 pixel hit.
- If a hit is expected in IBL, then one is required. The active area coverage of the IBL layer is more than 99%. If a track passes through an inactive IBL module, then a hit is required in the next layer if one is expected.
- At least 6 SCT hits. If a track passes through an inactive layer, it is counted as a hit which makes the selection less sensitive to differences in the number of dead module in data and simulation.
- The transverse impact parameter  $|d_0|$  calculated with the respect to the mean beam line<sup>1</sup> is required to be less than 1.5 mm to reduce the amount of badly reconstructed tracks and non-primary tracks.
- The longitudinal impact parameter  $z_0$  is calculated with respect to the primary vertex. It is required that the distance between the primary vertex and the track at the point where  $d_0$  is measured multiplied by  $\sin(\theta)$  is  $|z_0| \sin(\theta) < 1.5$  mm.
- If the track  $p_T$  exceeds 10 GeV, the probability of track  $\chi^2$  must be  $> 0.01$  in order to suppress mismeasured tracks.

### 4.3 Reconstruction efficiencies

The fact that the detector and other used tools are not 100% efficient has to be accounted for. Knowledge of these efficiencies acquired from the MC simulations is then used to estimate correction of observables to particle level in order to be able to compare results from various experiments. These observables are then corrected by applying weights which are constructed from the respective efficiencies. Detailed studies of trigger, vertex and tracking efficiencies were performed by minimum bias group [5].

#### Trigger efficiency

From the trigger section, MBTS was essential part used to select the events and as mentioned previously hit on either side was required. The trigger efficiency  $\epsilon_{\text{trig}}$  is taken as the ratio of MBTS triggered events to events in a control sample.

$\epsilon_{\text{trig}}$  as a function of selected tracks  $n_{\text{sel}}^{\text{BL}}$  is depicted in Fig. 4.2. These selected tracks are different from selected tracks in the section 4.2 due to loosen requirement on  $|z_0^{\text{BL}}| < 1000$  mm which is applied in order to avoid biasing the vertex position along the beam line. The trigger efficiency is above 98 % and even reaches 100 % for  $n_{\text{sel}}^{\text{BL}} > 6$ .

#### Vertex efficiency

Vertex efficiency  $\epsilon_{\text{vtx}}$  is taken as a ratio of MBTS triggered events with reconstructed vertex to all MBTS triggered events after subtraction of non-collision background. Dependence of  $\epsilon_{\text{vtx}}$  on  $\eta$  and  $n_{\text{sel}}^{\text{BL}}$  is depicted in the left Fig. 4.3 where for  $\eta$  plot it is only for events with  $n_{\text{sel}}^{\text{BL}} = 1$ . If  $n_{\text{sel}}^{\text{BL}} > 2$  then  $\epsilon_{\text{vtx}}$  is 100 %. The  $p_T^{\text{lead}} > 1$  GeV requirement in the analysis in question shifts the  $n_{\text{sel}}^{\text{BL}}$  into higher values in the right Fig. 4.3. Therefore, for most of the events,  $\epsilon_{\text{trig}}$  times  $\epsilon_{\text{vtx}}$  is equal to 1 and only for small portion of events ( $\sim 2$  %) value differs.

---

<sup>1</sup>Beam line also referred as beam spot is a mean value of primary vertex (vertex with the highest sum of tracks  $p_T$  in an event ) coordinates x,y,z in run.

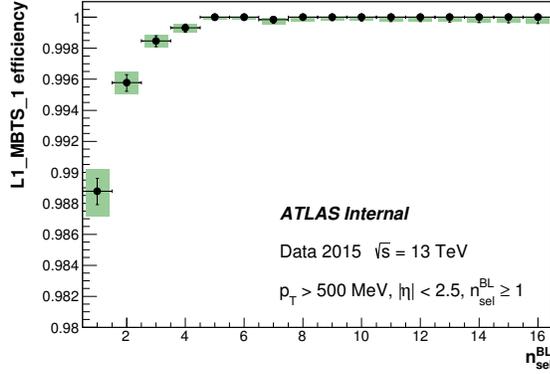


Fig. 4.2: Trigger efficiency  $\epsilon_{\text{trig}}$  plotted as a function of the number of selected tracks  $n_{\text{sel}}^{\text{BL}}$ . A reconstructed vertex is not required. The statistical uncertainty is shown as a black line, the total uncertainty as a green shaded area [5].

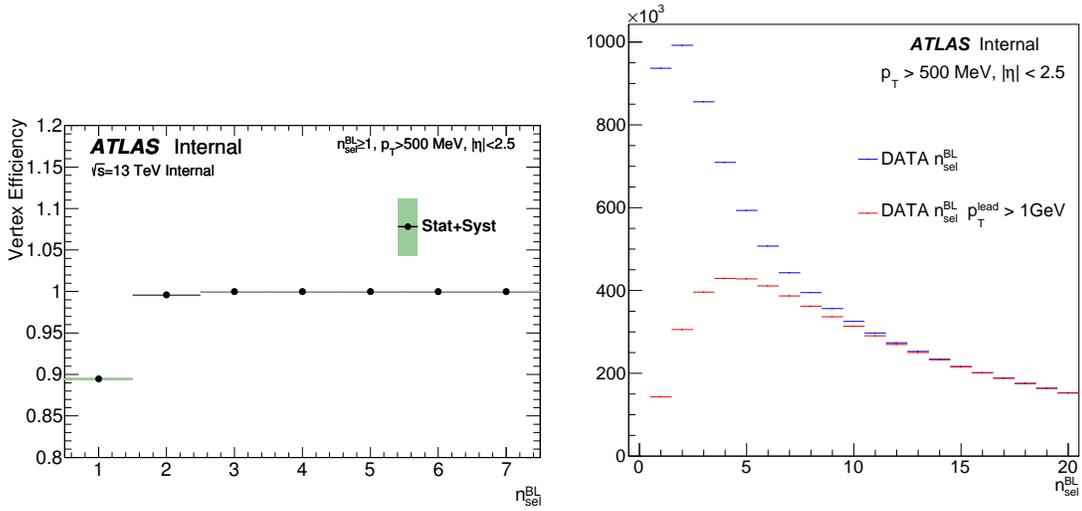


Fig. 4.3: Data driven vertex efficiency plotted as a function of selected tracks  $n_{\text{sel}}^{\text{BL}}$  (left). The statistical uncertainty is shown as a black line, the total uncertainty as a green shaded area [5]. The multiplicity distribution of selected tracks  $n_{\text{sel}}^{\text{BL}}$  (right), with the requirement on leading track transverse momentum  $p_{\text{T}}^{\text{lead}} > 1 \text{ GeV}$  (red) and without (blue).

## Tracking efficiency

The efficiency to reconstruct a track  $\epsilon_{\text{trk}}$  is determined from a simulation where tracks are matched to truth particles in the generator.  $\epsilon_{\text{trk}}$  is parameterized by  $\eta$  and  $p_T$  shown in the left and right Fig. 4.4 respectively. As can be seen in  $\eta$  figure, tracks are reconstructed with 85 % probability in the central rapidity regions. This efficiency decreases towards forward region reaching  $\sim 67\%$  for  $|\eta| = 2.5$ . Dependence of  $\epsilon_{\text{trk}}$  is rising from  $\sim 87.5\%$  at 500 MeV to  $\sim 92\%$  at 100 GeV.

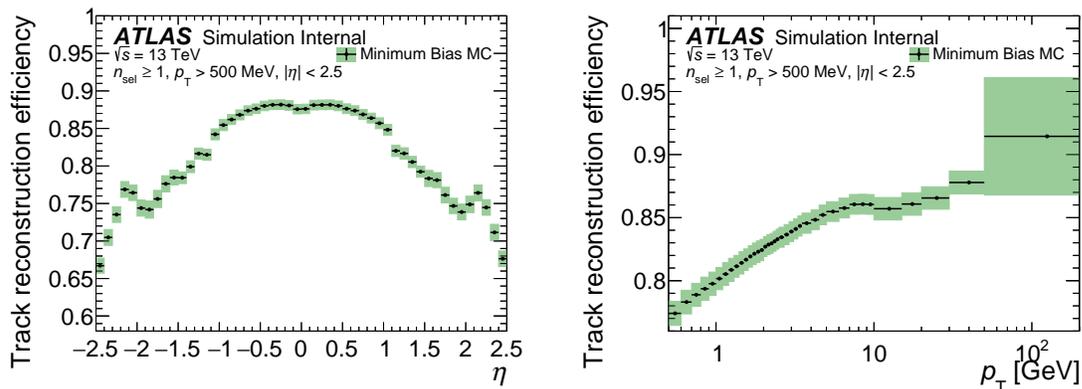


Fig. 4.4: Tracking efficiency as a function of  $\eta$  which is averaged over  $p_T$  (left) and  $p_T$  which is on the other side averaged over  $\eta$  (right). The statistical uncertainty is shown as a black line, the total uncertainty as a green shaded area [5].

## Fraction of non-primary tracks

The rate of particles which are non-primary is used in determining the final corrections. These particles come from hadronic interaction of particles with material and also from photon conversion. The rate of non-primary particles was obtained using side-band fits of the MC  $d_0$  and  $z_0$  distributions to data. The resulting fraction of non-primaries in data is  $0.0228 \pm 0.005$ . More information about the procedure in [5].

## Fraction of strange baryons

Contribution of strange baryons is taken as a background due to their bad reconstruction efficiency and significant model dependence. EPOS was chosen to estimate this background for its best description of strange baryons in the ALICE measurements. The fraction of strange baryons turned out to be negligible within the studied kinematic range.

## 4.4 Correction

To obtain observables at the particle level, they have to be corrected for detector effects. Two step correction is applied. At first observables are reweighted using trigger, vertex, track efficiencies introduced in the previous section 4.3. The requirement of second correction is due

to introduction of regions more or less sensitive to UE. If the leading track is not reconstructed the next track with the highest momentum (subleading) may takes its place. However, as the subleading track might have different  $\phi$ , the topology of the event is changed and some activity for example in transverse region will contribute to other regions or vice versa. The hit backspace once more (HBOM) method was used to account for this reorientation effect. Application of weights will be discussed at first, followed by the description of HBOM method.

#### 4.4.1 Event and track weights

Event weight is constructed as one over efficiencies of trigger and vertex reconstruction.

$$w_{\text{ev}} = \frac{1}{\epsilon_{\text{trig}}(n_{\text{sel}}^{\text{BL}})} \frac{1}{\epsilon_{\text{vtx}}(n_{\text{sel}}^{\text{BL}}, \eta)} \quad (4.1)$$

The multiplicity of selected tracks with no impact parameter cuts  $n_{\text{sel}}^{\text{BL}}$  is for  $p_{\text{T}}^{\text{lead}} > 1$  GeV high enough making the correction close to 1 for most of the events.

The correction for tracking effectivity is embedded in a tracking weight and contributions of non-primary tracks, strange baryons and tracks outside kinematic regions are subtracted from this correction.

$$w_{\text{trk}}(p_{\text{T}}, \eta) = \frac{1}{\epsilon_{\text{trk}}(p_{\text{T}}, \eta)} \cdot (1 - f_{\text{okr}}(p_{\text{T}}, \eta) - f_{\text{nonpr}}(p_{\text{T}}, \eta) - f_{\text{sb}}(p_{\text{T}})) \quad (4.2)$$

Due to a finite resolution of detector tracks outside the kinematic region ( $p_{\text{T}}, \eta$ ) might migrate inside the fiducial region of the measurement. Fraction of these tracks  $f_{\text{okr}}$  was estimated using MC and represents effect of a few percent at the edges of the fiducial region, but is overall negligible.  $f_{\text{nonpr}}$  and  $f_{\text{sb}}$  correspond to fraction of non-primary particles and fraction of strange baryons respectively which were discussed in 4.3 and 4.3. The track and event weights are applied to the measured observables as follows:

- $\sum_i p_{\text{T}}^i \rightarrow w_{\text{ev}} \sum_i p_{\text{T}}^i w_{\text{trk}}^i$
- $N_{\text{ch}} \rightarrow w_{\text{ev}} \sum_i w_{\text{trk}}^i$
- $\text{mean } p_{\text{T}} \rightarrow (\sum_i p_{\text{T}}^i w_{\text{trk}}^i) / (\sum_i w_{\text{trk}}^i)$

#### 4.4.2 Hit backspace once more method

Unfortunately, applying only track-event weight corrections is not sufficient to bring the measured spectra to the particle level. To account for other effects such as mentioned reorientation the HBOM method [27] was performed. This method is based on randomly loosing tracks from events according to their track reconstruction efficiency in several iteration. For  $i^{\text{th}}$  iteration the obtained spectrum corresponds to distribution with detector effects folded  $i$  times where  $i=0$  case corresponds to the raw data or simulation. The particle level result cleared out of detector effects is then obtained by extrapolating to  $i=-1$  utilizing fit procedure. The optimal number of HBOM iteration was studied and set to 6. The method is performed in each bin for all observables. As an example, one particular bin of  $\sum p_{\text{T}}$  distribution is depicted in the left Fig. 4.5<sup>2</sup>. Each iteration starts with a different seed to avoid correlation between the different

<sup>2</sup>Note that this is just an illustrative example in less statistically populated bin where the HBOM correction is not that effective

iterations. In each iteration, surviving tracks are again reweighted by additional  $w_{\text{trk}}(\eta, p_T)$ . In this way, most of the correction is carried out by simple weighting that reflects particular track kinematics leaving HBOM correction to deal with other effects such as the reorientation. This additional correction is typically within 2.5% for all the distributions.

The statistical uncertainty of the method is taken as a 68% confidence interval around the mean value of 1000 generated toys in  $i = -1$  shown in the right Fig. 4.5. For each toy, distribution in every iteration is smeared according to a gaussian distribution with the mean set equal to the initial value of the statistical uncertainty of the observable in that iteration. The advantage of using HBOM instead of orthodox application of bayesian unfolding is its simplicity. The method is not generator dependent and was applied on both data and simulation.

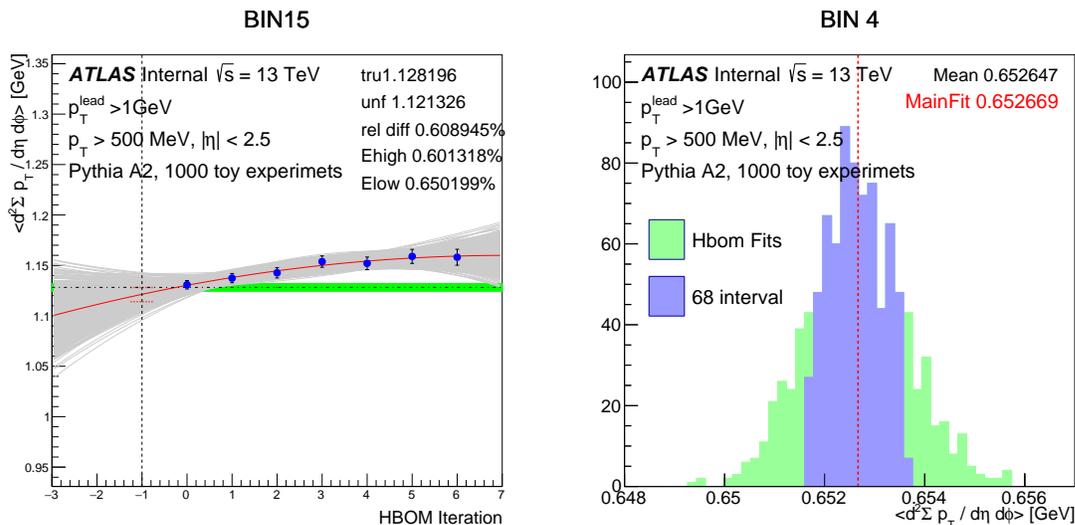


Fig. 4.5: Result of extrapolation to -1 using 6 iteration and parameterization by a polynomial of second degree in one particular bin of sum of transverse momentum distribution in transverse region with respect to the highest transverse momentum in an event  $p_T^{\text{lead}}$  (left). Red line represents the main fit, gray lines are 1000 toy experiments fits. For comparison, the particle level value (tru) is added as a dotted horizontal line with a green band as the statistical uncertainty. Distribution of toy experiments values in -1. Red dotted vertical line represents main fit from initial values which is close to the mean of toys(right).

#### 4.4.3 Effect of HBOM correction

Effect of HBOM correction were studied with MC generator PYTHIA 8 A2 using only dominant non-diffractive sample and EPOS to ensure that the correction method is not dependent on a generator. HBOM was studied for all the observables but only a few examples are provided below. In all the following figures in this section, there are comparisons between particle (referred also as truth) level represented by a black line, track-event corrected level in green

line and HBOM atop track-event correction depicted with a red line. Each figure is also appended by the ratio plot below to demonstrate the effect of HBOM correction.

All corrections on  $\sum p_T$  and  $N_{ch}$  distributions with respect to  $p_T^{lead}$  show similar trend depicted in the Fig. 4.6. The HBOM method is not sufficient in the first 2-3 bins which is the region with the highest probability to loose the leading tracks. Correction bias between HBOM and particle level (tru) spectrum in these bins is within 1 % showing improvement compared to the spectrum obtained by using only track-event weights. For the rest of the bins the correction to particle level works within the statistical uncertainty except for the tail of the distribution for some regions where low statistics resulted in low quality fits.

The difference between the track-event weighted spectrum and the particle level spectrum

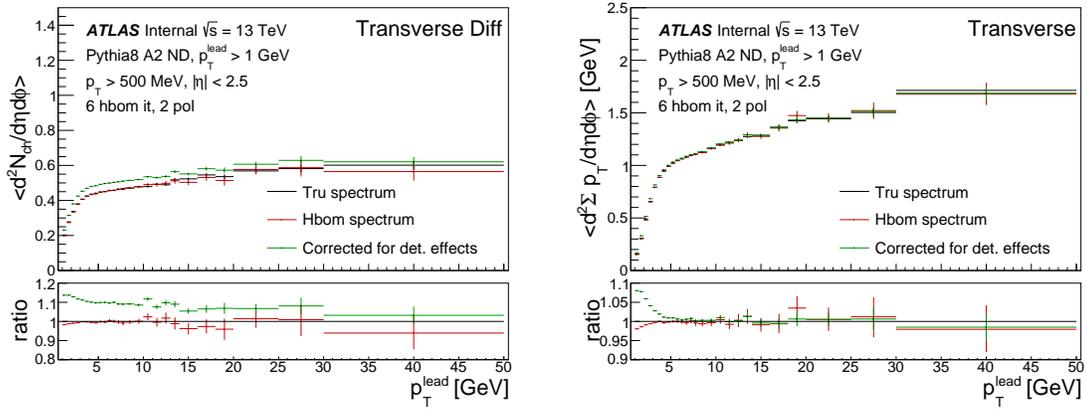


Fig. 4.6: Comparison of particle level spectrum (black), track-event corrected spectrum (green) and HBOM corrected spectrum (red) for charged particle density in transverse diff region on the left and sum of transverse momentum in the transverse region on the right both with respect to the highest transverse momentum in an event  $p_T^{lead}$ . The uncertainties are statistical only.

for mean  $p_T$  distributions plotted as a function of  $p_T^{lead}$  depicted in the left Fig. 4.7 is small and within 2%, however the HBOM method helps to reduce this correction bias. The HBOM also appears to help significantly the mean  $p_T$  distribution with respect to  $N_{ch}$  where its application brings very good agreement with the particle level except for the last bins.

For distributions  $\sum p_T$  and  $N_{ch}$  plotted as a function of  $|\Delta\phi|$  the correction bias is within 1%. As these two distributions are normalized to number of events taken from distribution of  $p_T^{lead}$ , HBOM was performed on unnormalized spectrum and also on  $p_T^{lead}$  distribution separately and only then were both normalized.

In general, HBOM correction was applied on distributions only in the areas where the correction bias turned out to be larger than statistical uncertainty.

## 4.5 Systematic uncertainties

The total systematic uncertainty consists of material modelling, non-primary contribution, HBOM non-closure and difference in number of HBOM iteration shown in Tab. 4.3.

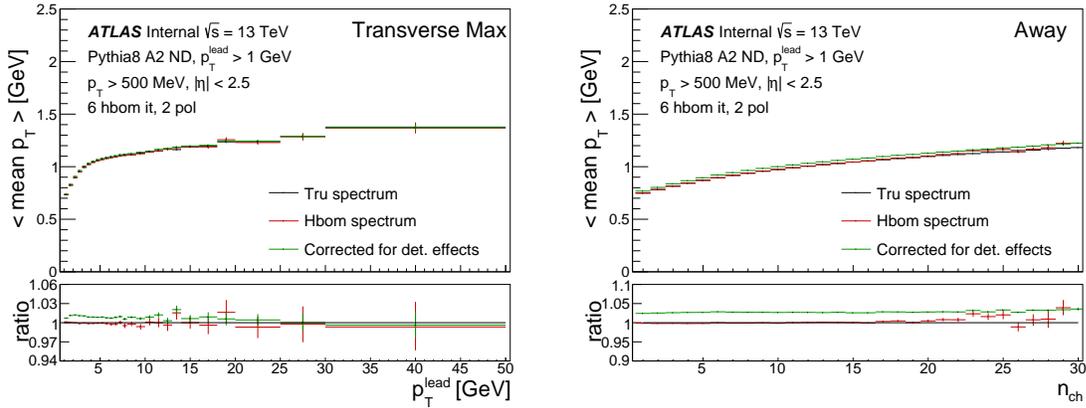


Fig. 4.7: Comparison of particle level spectrum (black), track-event corrected spectrum (green) and HBOM corrected spectrum (red) for mean transverse momentum  $\langle p_T \rangle$  with respect to highest transverse momentum in the event  $p_T^{\text{lead}}$  in transverse max region (left) and  $\langle p_T \rangle$  with respect to  $N_{\text{ch}}$  in the away region (right).

Uncertainties from bad material modeling and contribution of non-primary particles were propagated by modification of the track weights. These uncertainties includes both up and down variation which for example for the material modeling simulates less or more material in the detector thus larger possibility to loose track due to the hadronic interaction with the material.

As non-closure still persists for some distributions, this difference between truth and HBOM level from PYTHIA A2 ND is taken as a correction and also as a systematic uncertainty. As already said, the choice to use one certain MC is justified by the fact that the HBOM method is MC independent. The correction biased and uncertainty is applied for distributions plotted as a function of  $p_T^{\text{lead}}$  only in the first bins shown in the Fig. 4.6 where it reaches maximum of 2% for most of the distributions. For distributions plotted as a function of  $|\Delta\phi|$ , the difference is plotted in Fig. 4.8 below 0.5 % for all bins. Taken uncertainties for these azimuthal distributions were smoothed.

The number of HBOM iterations is by default 6 with parameterization by 2<sup>nd</sup> order polynomial. Behavior of all observables with different HBOM setups was studied by repeating the unfolding with 4 iterations and 2<sup>nd</sup> order polynomial and with 6 iterations and 3<sup>rd</sup> order polynomial. Some examples are provided in Fig. 4.9. The difference is significant in the first bins for all  $p_T^{\text{lead}}$  distributions. The same method as in previous case is applied for distributions plotted as a function of  $|\Delta\phi|$ . Contribution of iteration systematic uncertainty to  $N_{\text{ch}}$  distribution is small thus is combined with the correction bias to provide 0.5 % band for all bins.

Examples of combinations of all used systematic uncertainties are presented in Fig. 4.10,4.11.

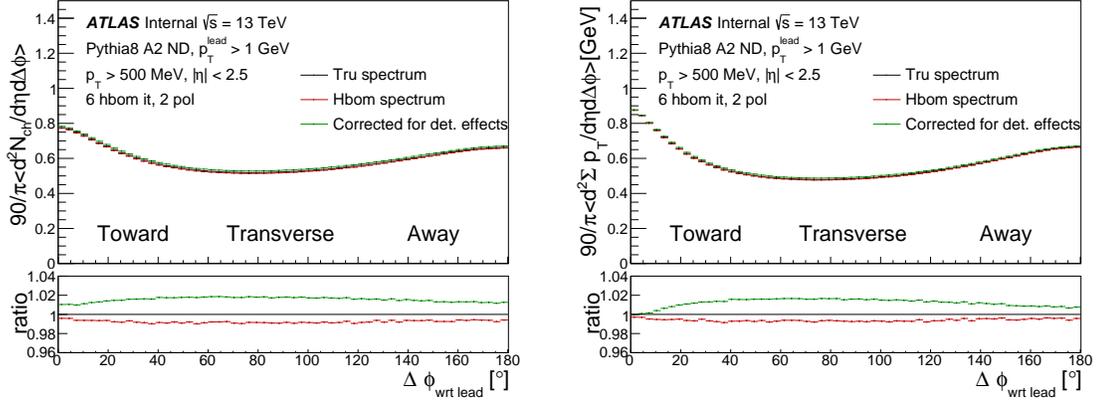


Fig. 4.8: Comparison of truth (black), track-event corrected (green) and HBOM (red) spectra for charged particle density on the left and sum of  $p_T$  on the right both with respect to  $|\Delta\phi|$ .

Observable	Range of values			
	Material	Non primaries	Non-closure	Parameterisation
$N_{ch}$ or $\sum p_T$ vs. $ \Delta\phi $	0.9%	0.6%	0–0.6%	0–0.4%
$N_{ch}$ or $\sum p_T$ vs. $p_T^{lead}$	0.5–1.0%	0.3–0.6%	0–2.5%	0–0.4%
$\langle p_T \rangle$ vs. $N_{ch}$	0–0.5%	0–0.5%	— 0.5% (combined)	—
$\langle p_T \rangle$ vs. $p_T^{lead}$	0–0.4%	0–0.3%	— 0.5% (combined)	—

Tab. 4.3: Summary of systematic uncertainties for each class of UE observable, broken down by origin.

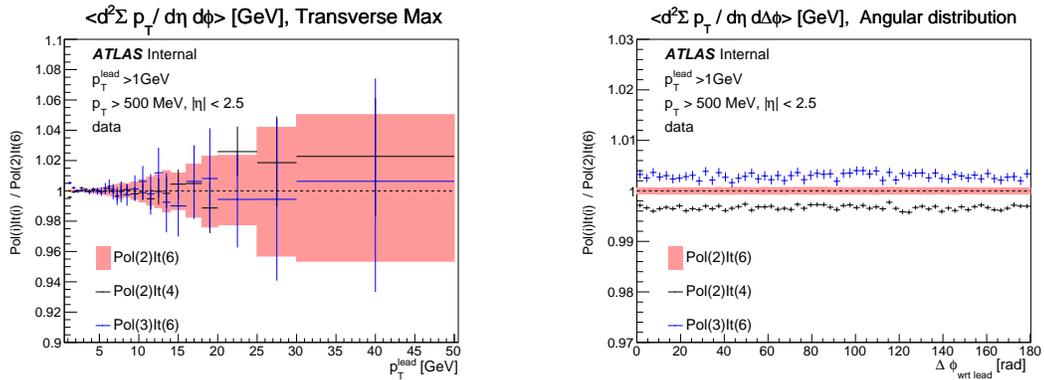


Fig. 4.9: 4 HBOM iteration with 2<sup>nd</sup> polynomial parametrization and 6 iteration with 3<sup>rd</sup> polynomial parametrization compared to the default 6 iteration with 2<sup>nd</sup> polynomial parametrization.  $\sum p_T$  with respect to  $p_T^{lead}$  in transverse max on the left and with respect to  $|\Delta\phi|$  on the right.

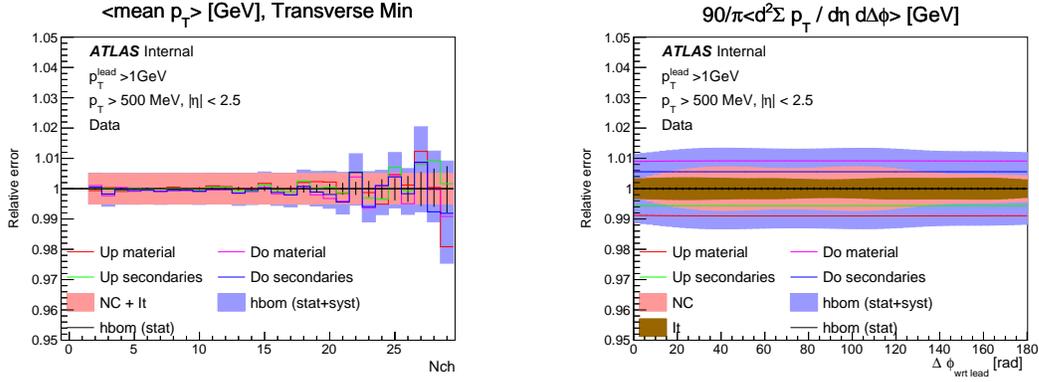


Fig. 4.10: Combination of all used systematic uncertainties for mean  $p_T$  plotted as a function of  $N_{\text{ch}}$  in transverse min region on the left, where the iteration systematic is included in the non-closure band, and  $\sum p_T$  as a function of  $|\Delta\phi|$  on the right.

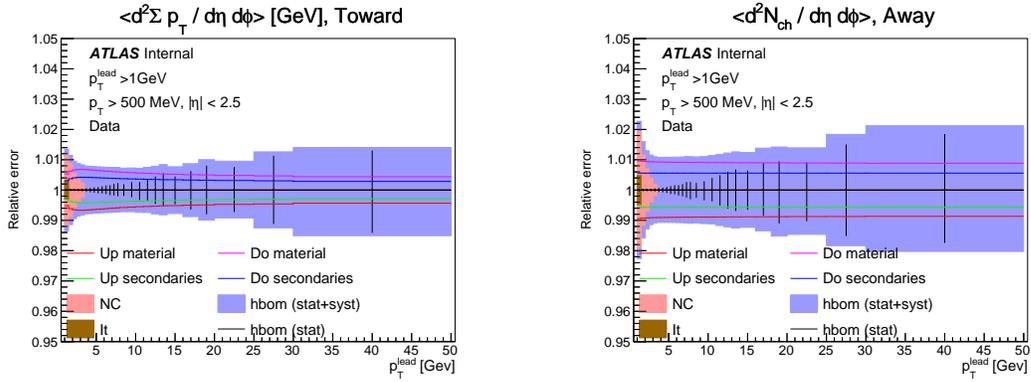


Fig. 4.11: Combination of all used systematic uncertainties for  $\sum p_T$  in the toward region on the left and  $N_{\text{ch}}$  in the away region on the right both plotted as functions of  $p_T^{\text{lead}}$ .

## 4.6 Results

Unfolded data are compared to several MC generators, in particular HERWIG 7, EPOS and from PYTHIA 8 family A2, MONASH and A14 tunes. MC generators were processed by Rivet software [28] invented as a tool for easy comparison of MC predictions with data. All distributions are normalized to number of events and to  $\eta - \phi$  space region. Distributions of  $|\Delta\phi|$  were divided by a degree conversion factor  $180/2\pi$ . All figures are accompanied with a ratio plot to provide easier comparison. Only several examples, in particular transverse regions, are shown in this section for better readability of the thesis. The whole set of results is enclosed in the Appendix.

### 4.6.1 Leading transverse momentum distribution

Normalized distribution of leading particle transverse momentum  $p_T^{\text{lead}} > 1$  GeV is depicted in the left Fig. 4.12. The distribution exhibits steep fall over the whole range with increasing hardness of the process. The best description is provided by EPOS and A14. The former is within the uncertainties approximately from  $\sim 7$  GeV whereas the latter shows offset around 10 %. The rest of the PYTHIA family seems to have an offset as well reaching almost 40% for A2 in the high  $p_T^{\text{lead}}$  region. HERWIG 7 deviates from the data in soft regions and even change the behavior in the first bins, the continuing range between 7-14 GeV is described reasonable well, whereas in the rest HERWIG 7 undershoots the data.

Better description of data is obtained when reducing the contribution of badly modeled low  $p_T$  events by applying condition on  $p_T^{\text{lead}} > 5$  GeV which is shown in the right Fig. 4.12.

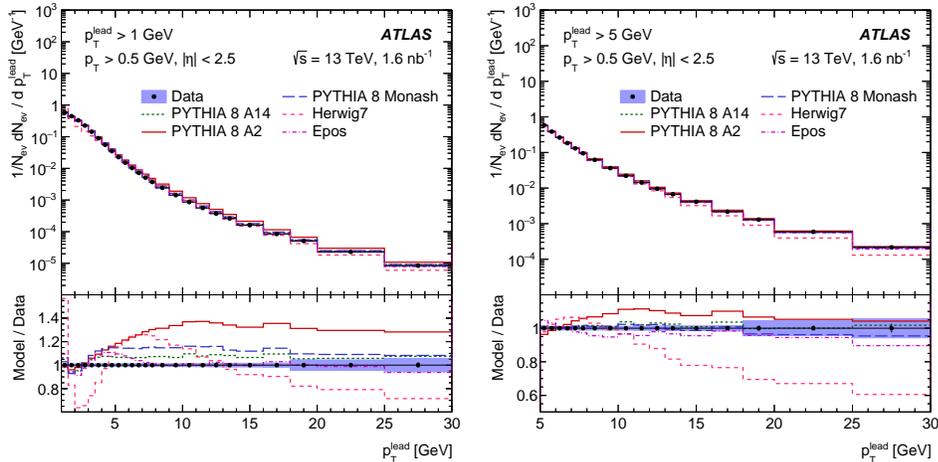


Fig. 4.12: Normalized distribution of leading particle transverse momentum  $p_T^{\text{lead}} > 1$  GeV (left) and  $> 5$  GeV. Blue shaded bands represent combination of statistical and systematic uncertainties.

## 4.6.2 Angular distributions

In Fig. 4.13 there are  $\sum p_T$  and  $N_{ch}$  distributions plotted as a function of  $|\Delta\phi|$  for  $p_T^{\text{lead}} > 1$  and 10 GeV. As can be seen from both distributions the topology of events is changed from more isotropic shapes with  $p_T^{\text{lead}} > 1$  GeV to shapes where energy is focused in certain directions. The toward region, which is situated between  $0 < |\Delta\phi| < 60^\circ$  has the highest activity for both observables with steeper drop in  $\sum p_T$  towards the transverse region which resembles the plateau. The slope is again gradually rising in the away region. None of the MCs are describing well both  $\sum p_T$  and  $N_{ch}$  in the whole toward region neither in away for  $p_T^{\text{lead}} > 1$  GeV. On the contrary, the transverse region seems to behave similar for data and most of the MCs except MONASH with  $\sim 20\%$  offset and A2 with  $\sim 40\%$  offset. With rising hard scale  $p_T^{\text{lead}} > 10$  GeV the description is changed for A2 and MONASH which now provides better agreement with the data, however, EPOS and A14 description deteriorates.

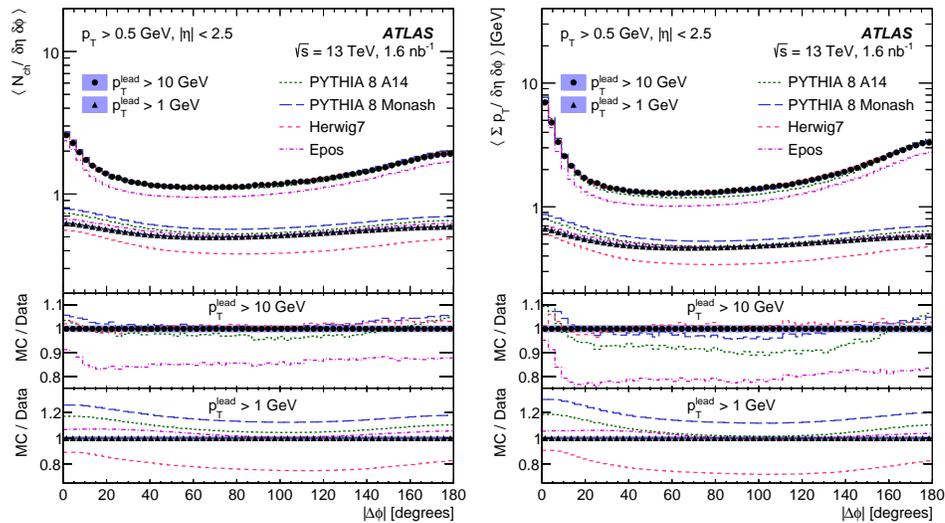


Fig. 4.13: Distribution of normalized charged particle multiplicities  $N_{ch}$  (left) and scalar sum of transverse momentum  $\sum p_T$  (right) plotted as a difference in the azimuthal angle  $|\Delta\phi|$  with respect to leading particle. Both observables contain data and MC for two different constraints on the leading particle transverse momentum  $p_T^{\text{lead}} > 1, 10$  GeV. Blue shaded bands represent combination of statistical and systematic uncertainties.

## 4.6.3 $N_{ch}$ and $\sum p_T$ distributions with respect to $p_T^{\text{lead}}$

In the region of low transverse momentum  $p_T^{\text{lead}} < 5$  GeV, charged particle density shows similar rising behavior between all three regions depicted in the left Fig. 4.14. The rest of the range is however slightly different. Contributions in both toward and away regions are slowly rising with  $p_T^{\text{lead}}$ . A visible plateau termed **underlying event pedestal** in the transverse region from  $p_T^{\text{lead}} > 5$  GeV means that the density of charged particles remains constant even with the harder scattering events. This phenomenon is associated with the fact that the spectrum is rising as impact parameter of protons decreases with interactions becoming

more central. As already mentioned in the section 2.3, number of MPI increases with the higher overlap, however, once the overlap is maximal the rate of MPI saturates. Therefore, the density of underlying event from MPIs is constant and can be altered only by initial and final state radiations.

Interesting point is that charged particle density is in the away region even higher than in the toward. There is less energy to redistribute among other particles in the toward regions as most of it is taken by the leading track. Dominance of the toward region in  $\sum p_T$  contribution shown in the right Fig. 4.14 is obvious and comes from the definition. The away region exhibits also steady rise through the whole  $p_T^{\text{lead}}$  range whereas rise in the transverse is mild compared to previous two regions.

Majority of MCs are peaking in the first bins of soft region  $N_{\text{ch}}$  distribution in transverse region depicted in the left Fig. 4.15. But as  $p_T^{\text{lead}}$  rises, the discrepancy between data and MCs start to decrease and after  $p_T^{\text{lead}} > 15$  GeV the deviance is within 5% for all MCs except EPOS. Similar trend is visible for  $\sum p_T$  distribution shown in the right Fig. 4.15.

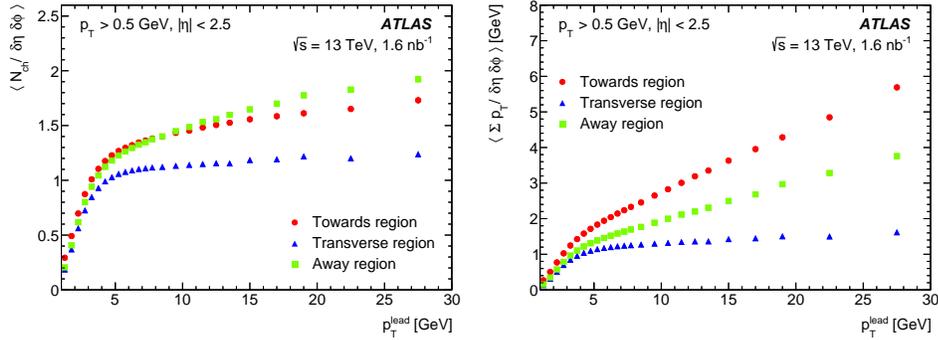


Fig. 4.14: Normalized charged particle multiplicity  $N_{\text{ch}}$  (left) and scalar sum of transverse momentum  $\sum p_T$  (right) plotted as a function of  $p_T^{\text{lead}}$  in all studied regions atop each other.

#### 4.6.4 Mean $p_T$ distribution

The mean charged particle transverse momenta  $\langle p_T \rangle$  observable probes the connection between  $\sum p_T$  and  $N_{\text{ch}}$  in a given region in the event. If the MPI systems are uncorrelated, equal redistribution of the energy would lead to a flat  $\langle p_T \rangle$ . However, a rise of  $\langle p_T \rangle$  can be observed in the Fig. 4.16. Therefore, MPI systems exhibit some correlation modeled through a mechanism of colour exchanges between systems called colour reconnection. This phenomenon effectively reduce the number of observed charged particles in the final states and as the contribution of  $\sum p_T$  should be approximately the same, because energy is conserved,  $\langle p_T \rangle$  rises. The more detailed description is postponed to the next chapter which is dedicated to understanding of colour exchanges. Left Fig. 4.16 shows the distribution of mean  $p_T$  with respect to  $p_T^{\text{lead}}$  in the transverse region. The spectrum has characteristic steep rise up to  $\sim 4$  GeV which continues with mild increase due to the shape of  $\sum p_T$  in the region in question.

The comparison between all MC and the data shows a bump in the soft regions. Description is then improved only for HERWIG 7 in higher  $p_T^{\text{lead}}$  regions whereas the rest of the MCs have offset ranging from 5% up to 10%.

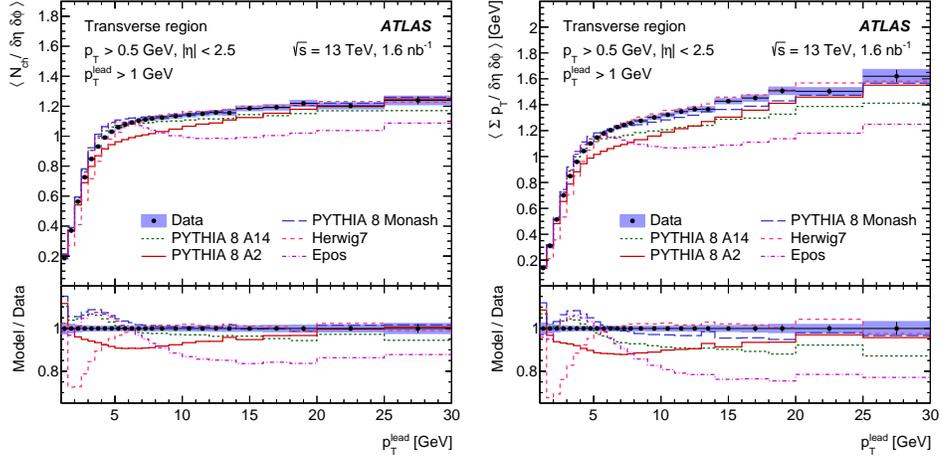


Fig. 4.15: Normalized charged particle multiplicity  $N_{ch}$  (left) and scalar sum of transverse momentum  $\sum p_T$  (right) plotted as a function of  $p_T^{\text{lead}}$  in the transverse region. Blue shaded bands represent combination of statistical and systematic uncertainties.

Mean  $p_T$  plotted as a function of  $N_{ch}$  is given in the right Fig. 4.16. The distribution shows that particles have higher  $p_T$  in more populated events. All MC generators are undershooting data in lower multiplicities. For  $N_{ch} \sim 12 - 15$  character changes and A2 HERWIG 7 and MONASH exceed the data. Nevertheless the deviation is within 5%.

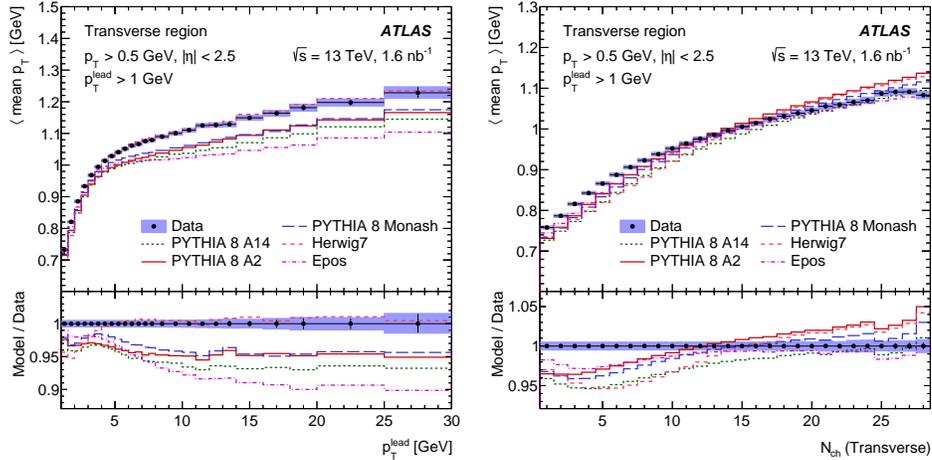


Fig. 4.16: Normalized mean transverse momentum  $p_T$  plotted as a function of  $p_T^{\text{lead}}$  (left) and as a function of multiplicities  $N_{ch}$  (right) in the transverse region. Blue shaded bands represent combination of statistical and systematic uncertainties.

## 4.7 Summary of UE analysis

The behavior of underlying event was studied in approximately 30 distributions sensitive to these effects and constructed from primary charged particles. The approach of the analysis was to separate an azimuthal plane perpendicular to the beam pipe into regions with different contribution of underlying event, using a track with the highest transverse momentum in an event with  $p_T^{\text{lead}}$  above 1 GeV.

The observables were measured by the ATLAS detector using data with typically single proton-proton interaction produced at the LHC at 13 TeV.

To obtain particle level information, measured observables were corrected by weights to account for trigger, vertex and track reconstruction efficiency. Additional correction using Hit backspace once more method was performed to correct for reorientation of events. This method helped to improve correction to particle level and reduce correction bias. Systematic uncertainties include material modeling, contribution of non-primary particles, correction bias and difference between the number of HBOM iteration.

The constant behavior so called underlying plateau with rising  $p_T^{\text{lead}}$  in the transverse region is observed for the distribution of sum of charged particle transverse momentum  $\sum p_T$  and also for charged particle multiplicity  $N_{\text{ch}}$ . It is even more visible in distributions calculated in transverse min region inhibited mostly by MPI than in transverse max with combination of MPI, ISR and FSR.

Data were compared with several MC generators namely HERWIG 7, EPOS, and from PYTHIA 8 family A2, A14, MONASH. The MC describe data reasonably well except for low  $p_T^{\text{lead}}$ .

The UE ATLAS measurement at 13 TeV shows rise in activity about 20% in comparison to 7 TeV ATLAS study [9].

The UE results were published in the beginning of 2017 [11] and can be found in the HepData format [12] with the corresponding Rivet analysis at [13].

## Chapter 5

# Effects of colour reconnection on top mass measurement

The fact that there is a large number of coloured partons in proton-proton interactions due to the MPI, PS and BBR was already discussed in the chapter 2. In this colour field, partons are looking for their partner with opposite colour charge to form colourless particles via hadronization. Lund hadronization model, mentioned in chapter 2, treats production of hadrons as the breaking of gluon string between colour dipoles. These strings may be stretched and overlap in space and time with possible interaction between them leading to the colour reconnection (CR).

### Static and dynamic representation

The CR can be viewed either in a static or a dynamic representation depicted in Fig. 5.1. Decays of top quark to bottom quark and W boson,  $t \rightarrow bW$ , with further decay of W to a quark-antiquark pair,  $W \rightarrow q\bar{q}$ , provides a nice example of the static representation. It may happen that the colour of  $b$  quark is opposite to the colour of the quark(antiquark) from W and the string is formed rather between  $bq$  ( $b\bar{q}$ ) than  $q\bar{q}$ . The dynamic representation on the other side views CR as a mentioned string-string interaction.

This effect obviously reshuffles with hadron composition and multiplicity in the final state thus may eventually lead to decrease in a precision of measurements. First requirement to introduce such phenomenon was made already in the previous century.

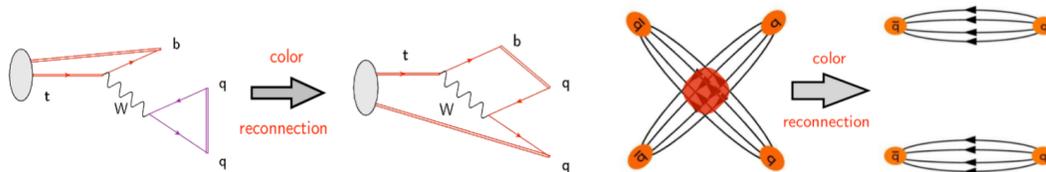


Fig. 5.1: Static representation of colour reconnection (left), dynamic representation of colour reconnection (right) [6].

## 5.1 Historic motivation for CR

### Minimum bias

CR was introduced to provide explanation of minimum bias data collected by UA1 experiment [29], in particular the measurement of the mean transverse momentum distribution  $\langle p_T \rangle$  with respect to multiplicity of charged particles  $N_{ch}$ . In Fig. 5.2 there is the mentioned distribution plotted using ATLAS data at 13 TeV and two MC predictions. One is PYTHIA 8 where the model of CR is turned off (red) and the other is default modeling of CR by PYTHIA 8 (blue). Simulations where CR is off are undershooting data by 20% in higher multiplicity region whereas the blue CR model provides reasonably good description.

Obviously there is a mechanism affecting the distribution leading to rise of  $\langle p_T \rangle$ . This can

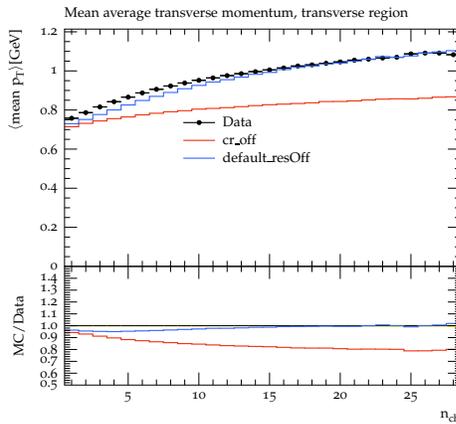


Fig. 5.2: Average number of charged particle transverse momentum with respect to their multiplicity in transverse region. 13 TeV minimum bias data are compared to PYTHIA 8 models where colour reconnection is turned off (red) and where the default model of colour reconnection is turned on (blue).

be achieved in two ways, either the total sum of charged particle transverse momentum  $\sum p_T$  is increasing or their multiplicity  $N_{ch}$  is decreasing. The first one will happen unlikely as the energy must be conserved (and no leakage to other regions is observed) leaving only the second option.

### $J/\psi$ production in B decays

Evidence of CR can be found also in the measurement of B mezon decay [30].  $b$  quark from a mezon can decay through the weak interaction to W boson and  $c$  quark  $b \rightarrow cW$ , where W can further decay hadronically into  $\bar{c}s$  quarks which are surely colour connected as they are produced from a colour singlet. The enhanced production of  $J/\psi$  particles were found which can be explained through the CR especially static representation.

Leaving aside for a while the details how various models implement the CR with resulting increase in  $\langle p_T \rangle$  in higher multiplicities, it is important to state the motivation for study of CR in this thesis.

## 5.2 Colour reconnection and top mass reconstruction

The exchange of colours may lead to scenarios with different final state hadrons. In the case of jets where partons are moving in a collimated cone, surrounding additional underlying event activity may take part in the CR interactions with partons inside jet resulting to leakage of hadrons from jet. This of course effects the jet transverse momentum  $p_T$  and hence precision of analyses which utilize jet  $p_T$  in their calculation.

### Obscuring of top quark mass

One example is the production of  $t\bar{t}$  system either from  $qq$  or  $gg$  interaction. Both  $t$  and  $\bar{t}$  essentially decay into  $b$  and  $\bar{b}$  quark. Mediating W bosons may decay hadronically hence  $t\bar{t}$  system can produce from 2 up to 6 jets in the final state. In the most direct approach, mass of the top is reconstructed from momenta of jets originated from  $b$  and  $q\bar{q}$  quarks.

$$m_{\text{top}}^2 = (p_b + p_q + p_{\bar{q}})^2 \quad (5.1)$$

Therefore, different CR modeling may obscure the reconstructed mass of top quark in a various way thus the CR represents additional systematic uncertainty which needs to be estimated.

### Underestimation of CR uncertainty

In the article about the CR and its effect on  $t\bar{t}$  system [31] S. Argyropoulos and T. Sjöstrand pointed out that systematic of top mass reconstruction coming from CR is underestimated. Although there has not been theoretical development the CR systematic uncertainty has been decreasing over the years. Currently, the uncertainty is estimated from a comparison of unrealistic CR models which are contrary to provided data. The values of top mass measured by various experiments together with CR systematic uncertainties are listed in Tab. 5.2. Suggestion from the article is to establish the CR effect on top mass by studying spread of various CR models embedded in PYTHIA 8 and provide constrains of these models using new measurement. New two sets of CR models with one being universal and the other specially dedicated for events with the top quark were introduced for the first time in the article in question.

### Aim of the analysis

The aim of this analysis is to study the spread in top mass using various CR schemes and to find observables which could be sensitive to CR thus might serve to supplement the constraints. First, an overview of various CR schemes embedded in PYTHIA 8 is provided.

## 5.3 Colour reconnection schemes

Currently, PYTHIA 8.219 posses 13 modes to study the effect of CR mechanism. All of them contain two main steps. In the first step, all pairs of partons/systems which are going to undergo colour reconnection are determined based on various criteria. The second step concerns

Experiment	$m_{\text{top}}$ [GeV]	CR $\Delta m_{\text{top}}$ [MeV]	Reference
ATLAS	$173.71 \pm 1.50(\text{stat}) \pm 1.44(\text{syst})$	140	[32]
CMS	$172.22 \pm 0.19(\text{stat}) \pm 0.75(\text{syst})$	150	[33]
D0	$174.98 \pm 0.58(\text{stat}) \pm 0.49(\text{syst})$	100	[34]
World combined	$173.34 \pm 0.27(\text{stat}) \pm 0.71(\text{syst})$	310	[35]

Tab. 5.1: Examples of top mass measurement from CDF and LHC experiments. The third column represents theoretical systematic uncertainty on top mass due to the CR modeling.

how the actual colour exchange happens.

### Top quark colour reconnection

A question arises which objects are actually affected by the CR in the case with a hadronically decaying top quark  $t \rightarrow bW \rightarrow bq\bar{q}$ . Top quark travels a distance  $\sim 0.2$  fm before it decays but hadronization scale is around 1 fm. There are two possibilities, either the CR happens only between the top quark and underlying event or top decays before the CR is invoked and its decay products are dumped to the environment in which they can further reconnect.

### List of all Pythia 8 CR models

The basic CR model in PYTHIA 8 is MPI-based model with two options. The first one is a default late resonance model in which only top quark participates in CR with the rest of the event. The second model is termed default early resonance that accounts also for a possibility where top decay products take part in CR. The rest of the 13 models can be separated into two main groups - extreme and full scale where the first was introduced to maximally affect top decay products whereas the second treats all colour partons equally.

Default models	Extreme models	Full scale models
<ul style="list-style-type: none"> <li>• late resonance</li> <li>• early resonance</li> </ul>	<ul style="list-style-type: none"> <li>• forced random</li> <li>• forced nearest</li> <li>• forced farthest</li> <li>• forced smallest <math>\Delta\lambda</math></li> <li>• smallest <math>\Delta\lambda</math></li> </ul>	<ul style="list-style-type: none"> <li>• move</li> <li>• move flip (excl. junctions)</li> <li>• move flip (incl. junctions)</li> <li>• swap</li> <li>• swap flip (excl. junctions)</li> <li>• swap flip (incl. junctions)</li> </ul>

#### 5.3.1 MPI-based models

In this model coloured partons are treated with respect to their MPI system. At first all MPI systems are found and sorted in increasing  $p_T$  of the system.

- Starting with the MPI system with the lowest  $p_T$ , probability to reconnect is given according to the inverse of  $p_T$ , phenomenological factor  $R_{\text{rec}}$  and  $p_{T0}^{\text{ref}}$  which is related to

the MPI model  $p_{T0}$  to cutoff soft divergence. This character of  $p_T$  dependence is driven by the fact that the  $P_{\text{rec}}$  is higher for softer gluons which are more expanded over the phase space thus have higher chance to overlap with other strings.

$$P_{\text{rec}}(p_T) = \frac{(R_{\text{rec}} p_{T0})^2}{(R_{\text{rec}} p_{T0})^2 + p_T^2} \quad (5.2)$$

All MPI systems are then iterated over and the reconnection happens stochastically based on  $P_{\text{rec}}$ . If reconnection occurs, new probability  $P_{\text{rec}}$  based on  $p_T$  from the next MPI system is calculated and then reconnection with the next in line MPI system is checked. However, if reconnection does not happen, reconnection with the next MPI in a line is tried such that the total probability of reconnection for the lowest MPI system is  $1 - (1 - P_{\text{rec}})^n$  in an event with  $n$  MPI. Note that more than 2 systems might be connected together.

- MPI systems which should be taking part in the colour reconnection are now determined from the previous step. Now to the actual exchange of colours. Contrary to the first step, starting with the MPI with the highest  $p_T$  and working in decreasing  $p_T$  order, all colour dipoles ( $ab$ ) are found. Then the MPI systems are iterated over in decreasing  $p_T$  and all gluons in the systems are located. To establish which gluons from one MPI system should be connected to which dipoles from the other MPI system, the model utilizes  $\lambda$  measure [36]. This property is proportional to the string length thus effect hadron production. When looked upon in the terms of energy,  $\lambda$  may represent available energy for particle creation and Nature seems to prefer states with lesser energy. Therefore, gluons  $g$  are exchanged to dipoles  $ab$  based on the minimalization of the increase of  $\lambda$  measure  $\Delta\lambda$ .

$$\Delta\lambda = \lambda_{ag} + \lambda_{bg} - \lambda_{ab} \quad (5.3)$$

In the end there are two colour dipoles  $ag, bg$  instead of original  $ab$ . In addition also  $q\bar{q}$  pairs which originate from the gluon splitting are allowed to move similar to the case with gluons.

### 5.3.2 Extreme models

These models are specially dedicated to top events. They all treat the CR of top quark utilizing default late resonance but varies in the modeling CR with top decay products. In this set of models gluons from top decay products  $g_{\text{tdp}}$  and from underlying event  $g_{\text{ue}}$  are put into two separate groups.

- Unlike MPI-based models in which probability to reconnect have a dependence on  $p_T$ ,  $g_{\text{tdp}}$  are forced to exchange gluons with  $g_{\text{ue}}$  in the extreme models. However, strength parameter  $\alpha \in (0, 1)$  was introduced to reduce the CR due to unrealistic large effects. Another feature is that  $g_{\text{ue}}$  are able to participate in CR multiple times. Exchange of gluons is determined based on various criteria depending on the particular model.
- **Forced random:** Gluon from  $g_{\text{tdp}}$  is forced to exchange colours with a gluon  $g_{\text{ue}}$  which is randomly selected.

- **Forced farthest:** Gluon from  $g_{\text{tdp}}$  is forced to exchange colours with a gluon  $g_{\text{ue}}$  that minimizes  $m(g_{\text{tdp}}, g_{\text{ue}})^2 = (p(g_{\text{tdp}}) + p(g_{\text{ue}}))^2$ .
- **Forced nearest:** The same as forced farthest but maximizes  $m(g_{\text{tdp}}, g_{\text{ue}})^2$ .
- **Forced smallest  $\Delta\lambda$ :** Force exchange of gluon  $i$  colour connected to partons  $ab$  with gluon  $j$  colour connected to partons  $cd$  in a way that  $\Delta\lambda$  is minimal.

$$\Delta\lambda(i, j) = \lambda_{j;ab} + \lambda_{i;cd} - \lambda_{i;ab} - \lambda_{j;cd}. \quad (5.4)$$

- **smallest  $\lambda$ :** Similar to forced smallest  $\Delta\lambda$  but only if  $\Delta\lambda < 0$ .

### 5.3.3 Full scale models

Contrary to extreme models, full scale models treat all gluons regardless of their origin and no default is used at the early stage.

- All gluons take part in CR, however, this can be controlled with a parameter  $\alpha \in (0, 1)$ .
- **Swap:** Utilizes smallest  $\Delta\lambda$  model with tunable parameter  $\Delta\lambda_{\text{cut}} < 0$ . CR occurs if  $\min \Delta\lambda < \Delta\lambda_{\text{cut}}$  and continues until the condition is broken.
- **Move:** Instead of swapping gluons between two dipoles, gluon  $i$  from connected to dipole  $ab$  is added to a gluon  $j$  connected to dipole  $cd$ . Again gluons are selected according to the minimal  $\Delta\lambda \leq \Delta\lambda_{\text{cut}}$ .

$$\Delta\lambda(i; cd) = \lambda_{i;cd} - \lambda_{i;ab} = \lambda_{ci} + \lambda_{di} + \lambda_{ab} - \lambda_{ia} - \lambda_{ib} - \lambda_{cd} \quad (5.5)$$

- **Flip:** This model lifts the restriction of fixed dipole quarks and allows exactly the possibility where  $b$  quark connects to quark from  $W$  in a decay of top. Again the flip is based on minimalization of  $\Delta\lambda \leq \Delta\lambda_{\text{cut}}$ .

$$\Delta\lambda(ab; cd) = \lambda_{ad} + \lambda_{bc} - \lambda_{ab} - \lambda_{cd} \quad (5.6)$$

As an improvement of the flip, model also incorporates junction topologies [37]. These represent structure with three quarks in a Y shaped configuration. Currently there are three variations of flip atop move or swap. It can be completely turned down so that the end of dipoles are fixed or turned on creating possibility for string to be flipped between the dipoles. Flip in junctions in addition to string flip represents the third option.

## 5.4 Top reconstruction analysis

PYTHIA 8.219 was used to generate approximately 10 million events with center of mass energy 13 TeV for different CR schemes. Semi-leptonic events where  $W^+ \rightarrow q\bar{q}$  and  $W^- \rightarrow l\nu_l$  were chosen instead of full leptonic or hadronic events as a compromise between too small and too large contribution of CR. Also due to the presence of two neutrinos, leptonic channel is complicated for the reconstruction whereas hadronic channel suffers with combinatorial complexity (which jet originates from which quark).

The particles of interest were produced from QCD interactions  $qq \rightarrow t\bar{t}$  and  $gg \rightarrow t\bar{t}$ . Top mass was set to  $m_{\text{top}} = 172.5$  GeV and W mass to  $m_W = 80.385$  GeV. W bosons were allowed to decay in the case of  $W^+ \rightarrow q\bar{q}$  where  $q \in (u, d, s, c, b)$  and regarding the negative boson  $W^- \rightarrow l\nu_l$ ,  $l \in (e, \mu)$  were enabled. Preselection was made on the presence of at least one electron or muon and on the number of final jets to be at least 4 within the acceptance of ATLAS to enhance statistic of the selected events. For extreme and full scale models  $\alpha$  parameter was set in correspondence with [31] using data from jet shape measurement in  $t\bar{t}$  events [38] and minimum bias [39] at  $\sqrt{s} = 7$  TeV to  $\alpha = 0.075$  and  $\Delta\lambda_{\text{cut}} = 0^1$ . The whole list of setting is in Tab. 5.3 in the appendix.

The analysis of simulation was done using Rivet [28] and the starting point was Rivet routine of  $t\bar{t}$  cross section used in ATLAS measurement [40] with following criteria:

- One lepton with transverse momentum  $p_T > 25$  GeV and  $|\eta| < 2.5$  and at most one lepton with  $p_T > 15$  GeV with the same pseudorapidity range. No more than one neutrino.
- Missing transverse energy  $E_T^{\text{miss}} > 30$  GeV and transverse mass  $m_T > 35$  GeV.  
 $m_T = \sqrt{2p_T^l E_T^{\text{miss}}(1 - \cos(\theta^l - \theta^\nu))}$ , with  $l$  indicating lepton and  $\nu$  neutrino.
- Jet reconstructed using anti-kt algorithm [41] with transverse momentum  $p_T > 25$  GeV and  $|\eta| < 2.5$ .
- At least four jets, two of them must be b-tagged<sup>2</sup>.
- Events with overlap<sup>3</sup> between jets lesser than 0.5 and overlap between jet and lepton lesser than 0.4 were not considered.
- Events where mass of  $W^+$  reconstructed from jets that were no b-tagged were outside 20 GeV range around experimental W mass were rejected.

Top mass spectra depicted in Fig. 5.3 were fitted with a gaussian distribution  $f_g$  (5.7) utilizing ROOT framework [42] and results for CR schemes are presented in Tab. 5.2.

$$f_g(m) = A \exp\left(\frac{-(m - m_{\text{top}})^2}{2\sigma^2}\right) \quad (5.7)$$

All models are compared to default late resonance (default resOff) model which is the nominal CR scheme for PYTHIA 8. As can be seen effect of CR with different treatment of gluon exchange produce spread in both direction ranging from -837 MeV for forced farthest up to 1012 MeV for model with CR turned off. In general it seems that forced gluon exchanges predict smaller value of the top mass whereas full scale models exhibit more massive top quark.

There is a slight difference between Tab. 5.2 and the fitted mass obtained in the mentioned older CR  $t\bar{t}$  analysis [31]. This discrepancy is mostly caused due to the different selection

<sup>1</sup>Values of  $p_{T0}^{\text{ref}}$  parameters which is directly connected to low  $p_T$  cutoff were used according to the mentioned article.

<sup>2</sup>Selected as jets coming from b quark.

<sup>3</sup>An overlap in a phase-space of azimuthal angle and pseudorapidity, where a distance is measured as  $\Delta R = \sqrt{(\phi(\text{jet1}) - \phi(\text{jet2}))^2 + (\eta(\text{jet1}) - \eta(\text{jet2}))^2}$

used in the reconstruction of  $t\bar{t}$ . Another reason is concerning the choice of baseline tune. All models in the thesis use the standard MONASH 2013 tune whereas 4C tune is used in the article. Lastly, only rough fit of the top masses was performed which can be seen from systematic uncertainties ( $\sim 130$  MeV). This can be further improved by more sophisticated fitting procedure.

CR model	$m_{\text{top}}$ (fit) [GeV]	$\Delta m_{\text{top}}$ (fit) [MeV]
default resOff	$168.99 \pm 0.12$	0.00
cr off	$170.00 \pm 0.12$	1012.52
default resOn	$169.50 \pm 0.13$	510.62
forced random	$168.18 \pm 0.14$	-812.60
forced nearest	$168.66 \pm 0.12$	-330.53
forced farthest	$168.15 \pm 0.14$	-837.96
swap	$169.47 \pm 0.12$	483.90
swap flip1	$169.51 \pm 0.12$	516.55
swap flip2	$169.40 \pm 0.13$	414.20
move	$169.67 \pm 0.13$	682.12
move flip1	$169.63 \pm 0.13$	638.39
move flip2	$169.76 \pm 0.12$	772.17

Tab. 5.2: Comparison of fit values of top mass reconstructed in various CR schemes. The columns go as follows: in the first are names of CR scheme, second includes fit values of top mass with fit uncertainty, third represents comparison to default resOff model.

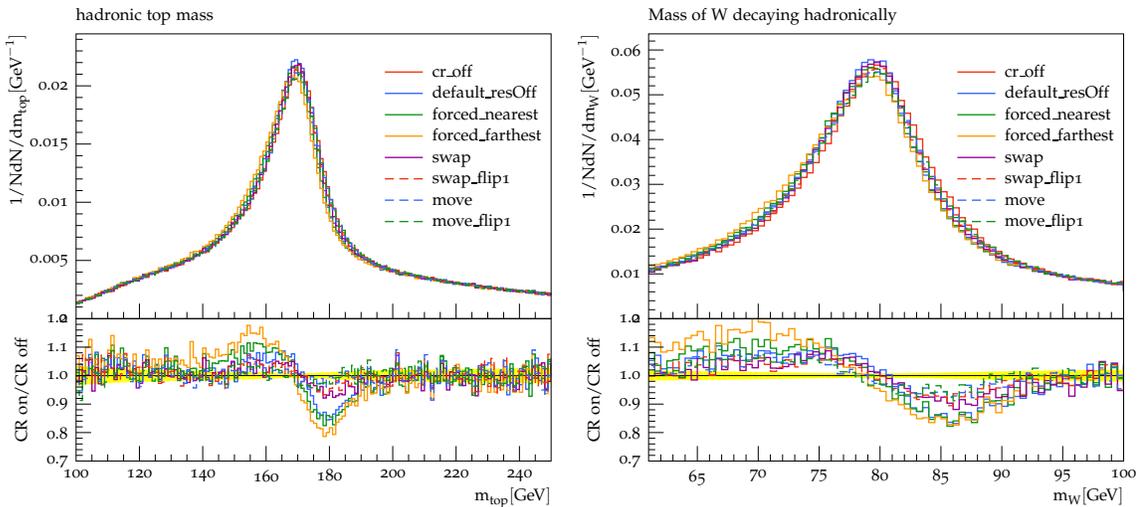


Fig. 5.3: Normalized mass spectrum of top quark reconstructed from  $b$  jet and two jets coming from the decay of W boson (left), normalized mass spectrum of W reconstructed from the jets originating from the decay of W (right).

## 5.5 Constrain of colour reconnection

Further investigation of CR effect is required in order to constrain models in question using measured data or at least suggest possible measurement. Several observables which might have sensitivity to CR effect were studied.

At first, basic observables such as transverse momentum  $p_T$  of charged particle, their multiplicity  $N_{ch}$ , sum of their  $p_T$  or mean  $p_T$  were explored in jets and underlying event in  $t\bar{t}$  system. Four main jets coming from  $b$ ,  $\bar{b}$  and both jets originating from quarks from  $W^+$  termed as lightjet1 and lightjet2 in decreasing  $p_T$  were studied separately to probe if CR have different effects on them due to the topology of events. Plane azimuthal to the beam pipe were divided into toward, transverse, away regions but in this case based on  $t\bar{t}$  system in order to look for sensitivity in different areas. However, no additional information were obtained compared to the case where all regions are combined.

One of the candidate for CR sensitivity are jet shapes distributions of  $p_T$  and  $N_{ch}$  in the jets studied separately in different annuli from the jet axis. CR between jet constituents and underlying event may lead to a leakage of particles from the jet cone which various reconnection schemes treat differently.

Looking for constrain elsewhere than in  $t\bar{t}$  system, underlying event measurement in minimum bias collisions with ATLAS at 13 TeV provided some information although only for the full scale and default models as extremes are specially dedicated to production of  $t$  quark.

Another possibility where to look for difference between CR models is the colour flow in the event. A pull vector and with it associated a pull angle observable provides information about coloured history of events. Unfortunately, this distribution turned out to be insensitive. Nevertheless, the results of this measurement are provided as well.

As already mentioned, CR strongly influence hadronization stage. One of the ATLAS analysis at 7 TeV was focused on the measurement of the fragmentation function which describes fraction of partons momenta carried by final state hadrons. Investigation of CR fragmentation functions is one of the recent studies and is included in the end of this chapter. More observables and plots potentially sensitive to CR were under study. One example may be the number of particles in an area between jets given by the distance of their axis and a parameter  $\Delta R_0$ , however, not much proved to be useful and are not discussed.

### 5.5.1 Activity in jets

In the Fig. 5.4 there is a distribution of charged particle multiplicity  $n_{ch}$  inside the jet originating from  $b$  quark (left) and charged particle transverse momentum  $p_T$  in the lightjet1 (right). As can be seen from this plot there are typically bundles of models with the similar behavior across other properties. For example both default models have almost identical behavior and require further understanding. For the purpose of figure's clarity, in what follows models which exhibit the same behavior are left out except only one representative highlighted below.

All the models typically exhibit higher contribution in lower multiplicity regions compared

- **default late resonance** (resOff)
- **forced farthest**
- **move flip1**
- **swap flip1**
- default early resonance
- forced random
- move flip2
- swap flip2

to the case without CR. The difference between the models is erased towards the mean value of  $n_{\text{ch}}$ . The contrary applies toward the distribution tail where most of the models are less populated except the extreme schemes which again change the shape rise. These shapes are common also in all the other jets. Full scale models are the most dominant within the whole  $n_{\text{ch}}$  range in particular move flip model allowing only string flips.

The difference between the models in jet  $p_{\text{T}}$  distribution is not that significant in comparison to  $n_{\text{ch}}$  and reaches only up to several percent. Right Fig. 5.4 shows the dependence for lightjet1, however, the same trend within a difference of percent is visible for all jets. All full scale models predicts larger contribution of higher  $p_{\text{T}}$  particles than CR off except the beginning of the spectrum. Default is insensitive to any changes in the given distribution as lightjet1 is a decay remnant of  $t$  thus no CR is applied. Extreme models gives slightly lesser contribution than CR off.

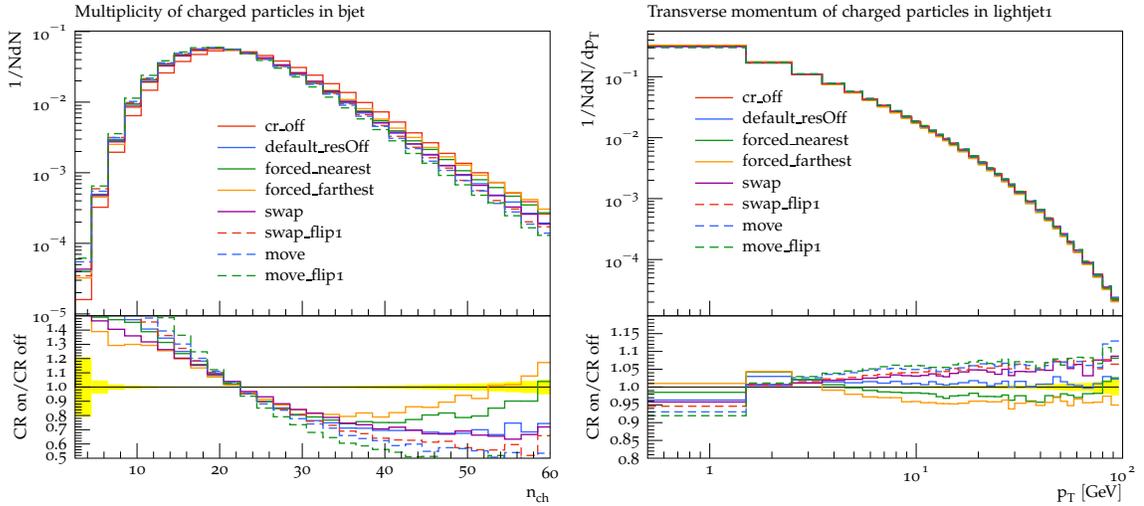


Fig. 5.4: Charged particle multiplicity  $n_{\text{ch}}$  in the  $b$  jet from top decay (left), charged particle transverse momentum  $p_{\text{T}}$  in the higher  $p_{\text{T}}$  jet originating from  $W$  decay termed lightjet1 (right).

Average charged particle transverse momentum  $\langle p_{\text{T}} \rangle$  distribution in lightjet1 and the same  $\langle p_{\text{T}} \rangle$  plotted against charged particle multiplicity  $n_{\text{ch}}$  are depicted on the left and right side Fig. 5.5 respectively. The left plot shows also the transition behavior of CR models with respect to CR off. All the different variability of CR eventually results into lesser  $\langle p_{\text{T}} \rangle$  at the beginning of the spectra. At  $\sim 4$  GeV CR character changes and  $\langle p_{\text{T}} \rangle$  is actually higher. The difference between the models shares the the same ordering as in  $n_{\text{ch}}$  with dominant move flip1.

There is no significant sensitivity to CR in  $\langle p_{\text{T}} \rangle$  distribution for low  $n_{\text{ch}}$ . Some can be seen in the tail where the steep spectrum falls and the energy is more redistributed between the particles. This is however only within a few percent. Interesting feature is that forced farthest resembles almost the case with no CR at all. Again, this behavior is typical for the rest of the jets from which  $t\bar{t}$  system is constructed.

The sum of charged particle transverse momentum in jets were investigated as well. However,

no significant discrepancy was observed.

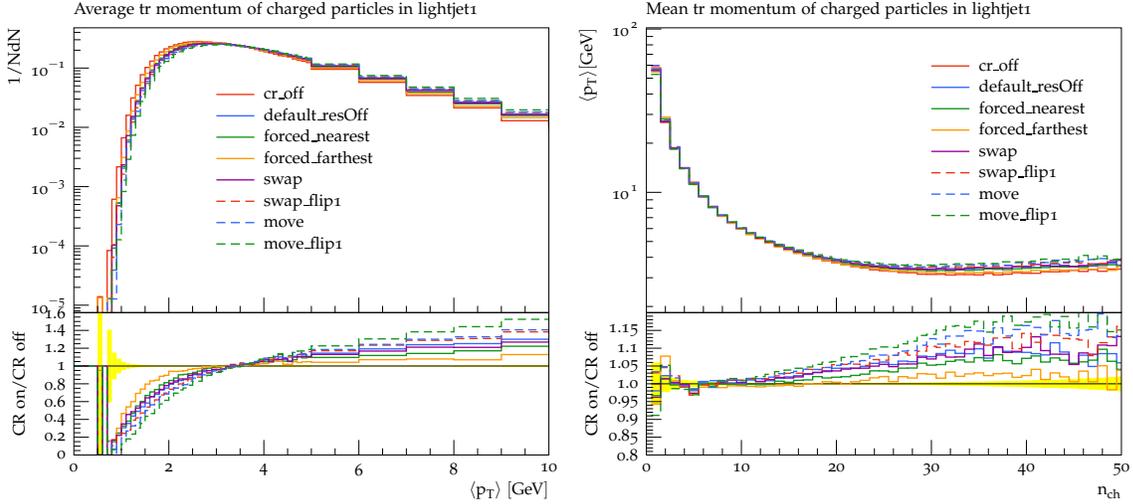


Fig. 5.5: Distribution of average charged particle transverse momentum in the higher  $p_T$  jet originating from W decay termed lightjet1 (left) and an average charged particle transverse momentum with respect to the multiplicity in the lightjet1 (right).

### Activity in underlying event

Slightly different effect of colour exchanges was found in underlying event spectra of  $t\bar{t}$  system. Left Fig. 5.6 shows the multiplicity of charged particles which are not associated with the jets coming from top quark nor any other jet created for example from MPI interactions. As can be seen  $n_{ch}$  in UE distribution is much broader with a mean value of  $\sim 50$  charged particles. Contrary to multiplicity in jets, forced farthest model shows a lesser contribution already in the beginning of the spectrum. Then it exceed the CR off model between  $25 < n_{ch} < 65$  and drops again. Interesting feature occurs approximately in the same region where the forced farthest is in the excess. Flip models which typically exhibit the largest difference compared to CR off are overcome by default, nearest and farthest. CR models start to display three similar trends in the higher multiplicities.

Spectrum of transverse momentum  $p_T$  from charged particles in UE is depicted in the right Fig. 5.6. All the models exceed the CR off within the whole range.  $p_T$  is a very sensitive observable allowing to discriminate between models (difference 10-40%) and can be easily accessed experimentally. Also  $\sum p_T$  distributions Fig. 5.7 in UE contain some sensitivity to CR which was not the case for jets. This sensitivity is even more enhanced if  $\sum p_T$  is taken only from charged particles which are not clustered in the additional event jets (left Fig. 5.7). There is a similar trend of CR schemes at the beginning of distribution for particles which were not associated with any jets and UE particles which are not in the jets from the  $t\bar{t}$  system decay but are allowed to be in the ambient additional jets (right Fig. 5.7). The figure drastically changes in the more energetic environment. The left distribution is much steeper in the tail and CR models are well distinguished in comparison to the right plot with a mild sensitivity.

Probably the largest deviation of CR models can be seen in the average UE charged particle transverse momentum  $\langle p_T \rangle$  in the left Fig. 5.8. Models indicate lesser contribution in the lowest  $\langle p_T \rangle$  region but rise from around  $\sim 1.2$  GeV. Once they reach the maximum they slightly drop and then steadily decrease their slope. The difference for the highest move flip1 model is almost up to 300%. On the other hand force nearest is in the tail identical to CR off. The rest of the models are spread among this large range.  $\langle p_T \rangle$  plotted with respect to UE charged particle multiplicity  $n_{ch}$  is plotted in the right Fig. 5.8. All models are steadily deviating from the CR off with rising  $n_{ch}$  leaving the possibility for more reconnection options.

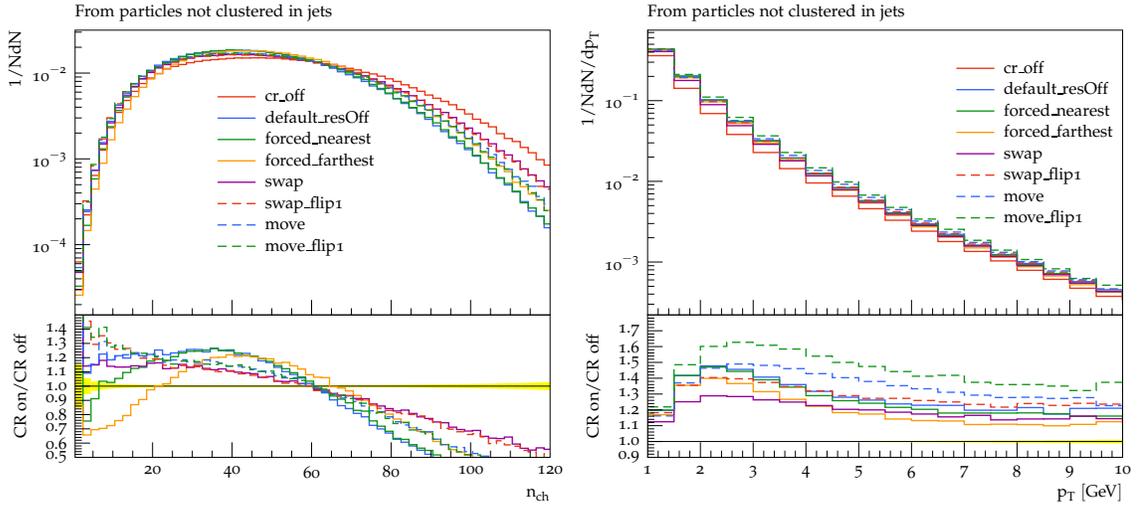


Fig. 5.6: Multiplicity of charged particles  $n_{ch}$  in the underlying event which are not clustered in jets originating from  $t\bar{t}$  decayed quarks nor MPI (left) and their transverse momentum  $p_T$  with(right).

### 5.5.2 Jet shapes

Activity in an annuli around jet axis is potential observable where to look for CR sensitivity. Particles inside jets may interact with each other but also with underlying event surrounding the jets. Data from  $t\bar{t}$  jet shapes measurement at 7 TeV [38] were actually used to tune the parameters in the mentioned analysis concerning the CR effect at 7 TeV. The main observable is a differential jet shape  $\rho(r)$  in an annulus with the distance  $r$  from the jet axis and inner and outer radius  $r - \Delta r, r + \Delta r$  respectively. Therefore, sum of transverse momentum in annuli normalized to jet transverse momentum. Radius of the jet  $R$  is 0.4 and radius width  $\Delta r$  of annuli was set to 0.02.

$$\rho(r) = \frac{1}{\Delta r} \frac{p_T(r - \Delta r, r + \Delta r)}{p_T(0, R)} \quad (5.8)$$

Provided distributions are reproduced only for charged particles while both charged and neutral particles were used in the article. In Fig. 5.9 there are differential jet shapes in  $t\bar{t}$  system at 13 TeV for two different annuli of b tagged jet, one very close to the jet axis (left figure)

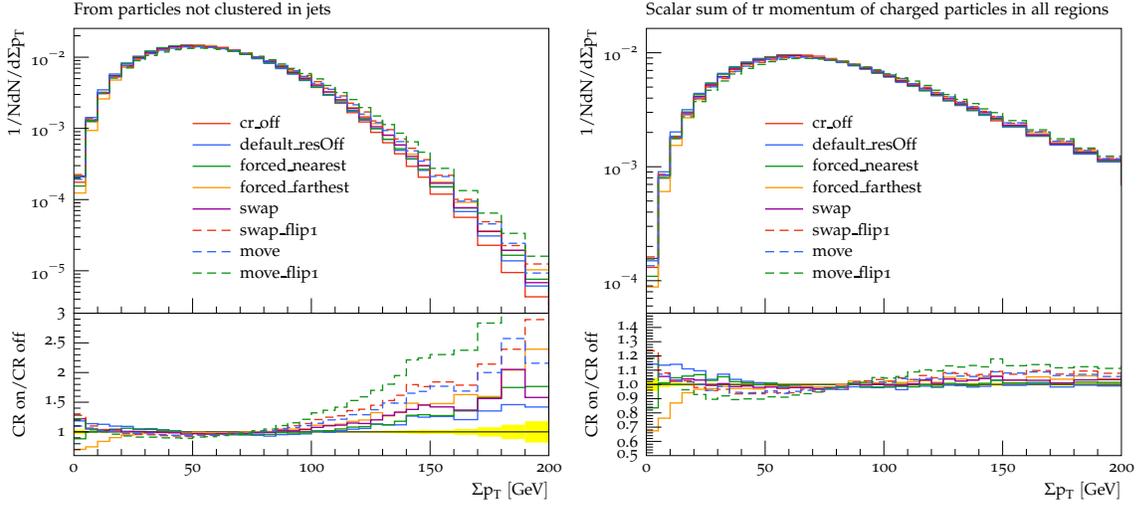


Fig. 5.7: Sum of charged particles in the underlying event transverse momentum  $\Sigma p_T$ . These particles are either not associated with any selected jets (left) or only not associated with jets originating from decay of  $t\bar{t}$  system(right).

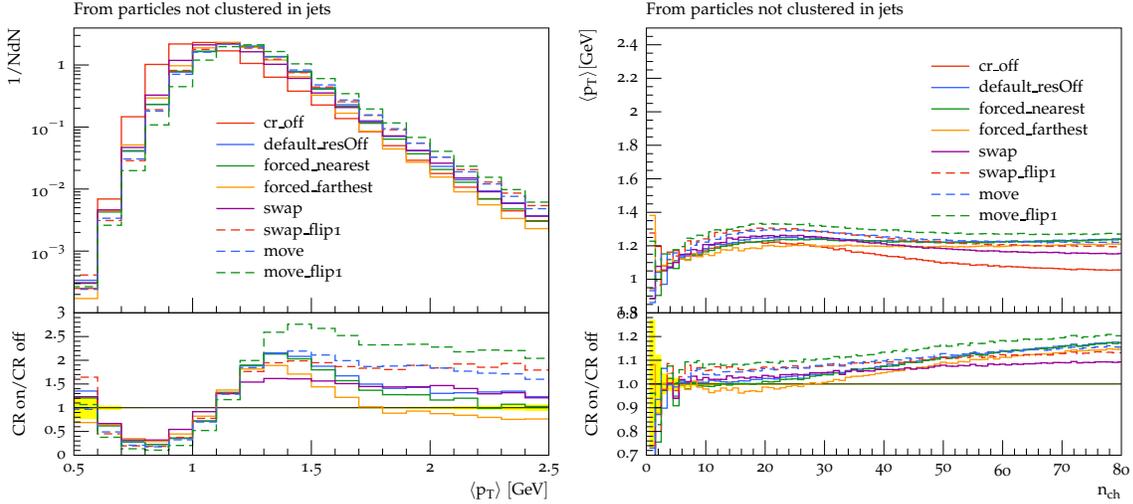


Fig. 5.8: Average transverse momentum of charged particles in the UE which are not associated with any selected jets  $\langle p_T \rangle$  (left), and also  $\langle p_T \rangle$  plotted against the multiplicity of particles in question (right).

and one near the edge of the jet (right figure). As can be seen, tuned CR models exhibit no discrepancy in  $\rho(r)$  near the jet axis nor edge of the jet at 13 TeV.

As most of the sensitivity in the previous chapter was observed in the charged particle multiplicity, jet shape constructed similarly to  $\rho(r)$  but based on the  $n_{ch}$  were studied Fig. 5.10 in different annuli as well. This shape was normalized on the total number of charged particles

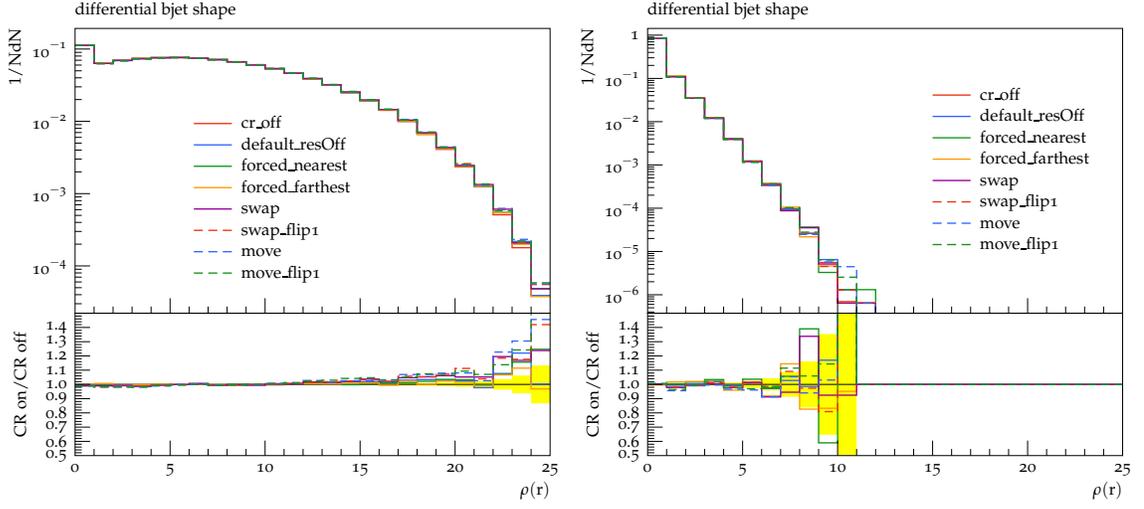


Fig. 5.9: Differential jet shape  $\rho(r)$  representing charged particle transverse momentum  $p_{Tj}$  in the annulus with a distance  $r$  from the jet axis, normalized by jet transverse momentum  $p_{Tjet}$  and width of the annulus  $\Delta r$ . Left figure is for  $r = 0.06$  and right for  $r = 0.34$ .

inside a jet. Both annuli shows higher effect of CR at higher multiplicity fractions.  $n_{ch}$  jet shapes exhibit typical ordering of the CR schemes.

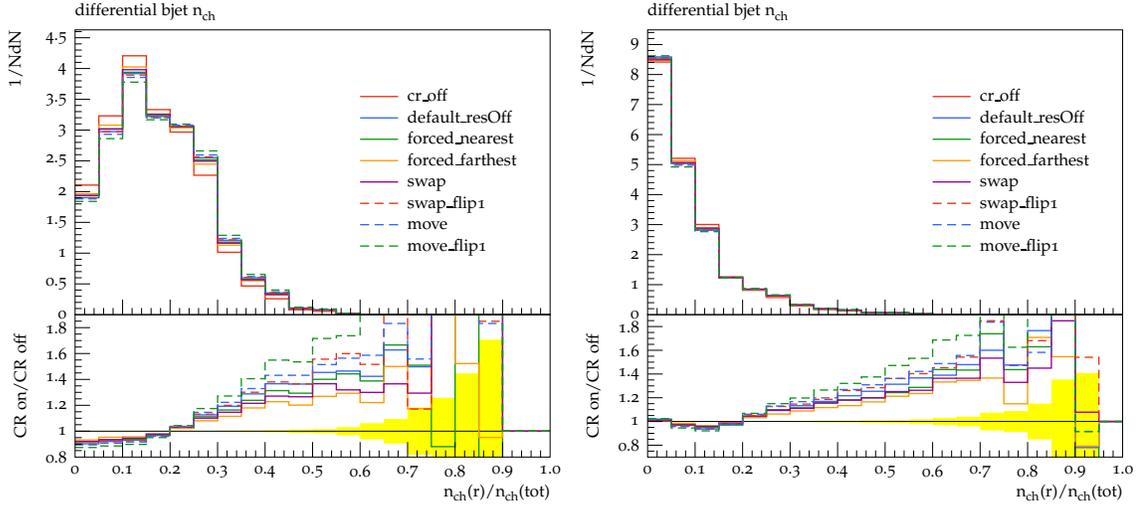


Fig. 5.10: Differential jet shape representing charged particle multiplicity  $n_{ch}$  in the annulus with a distance  $r$  from the jet axis, normalized by the total multiplicity inside jet  $n_{ch}(tot)$  and width of the annulus  $\Delta r$ . Left figure is for  $r = 0.06$  and right for  $r = 0.38$ .

### 5.5.3 CR and UE at 13 TeV

Full scale models were tuned to minimum bias data but some constraints of the models can be also made using underlying event distributions from the first part of the thesis. Only default and full scale models can be used as the data describe minimum bias collisions where the environment to create top particle is too soft. Therefore, forced models are excluded from the comparison. Also the default early resonance model which allows early decay of top quark is excluded as well. Small ( $\sim 1$  M) samples with different CR settings were generated in order to compare with provided data.

In the left Fig. 5.11 there is a mean charged particle multiplicity  $N_{\text{ch}}$  distribution with respect to the leading particle transverse momentum  $p_{\text{T}}^{\text{lead}}$  in the transverse region. All CR models predict lesser contribution compared to CR off model up to 12 GeV where the simulations are less populated due to the limited size of the sample. The best description of data is provided by default late resonance MONASH tune which was already shown in Fig. 4.15.

Average transverse momentum of charged particles  $\langle p_{\text{T}} \rangle$  with respect to their multiplicity  $N_{\text{ch}}$  in transverse region was depicted in Fig. 5.2 but in this case it includes more CR schemes for comparison. All CR schemes show increase in  $\langle p_{\text{T}} \rangle$  in higher multiplicities compared to CR off model.

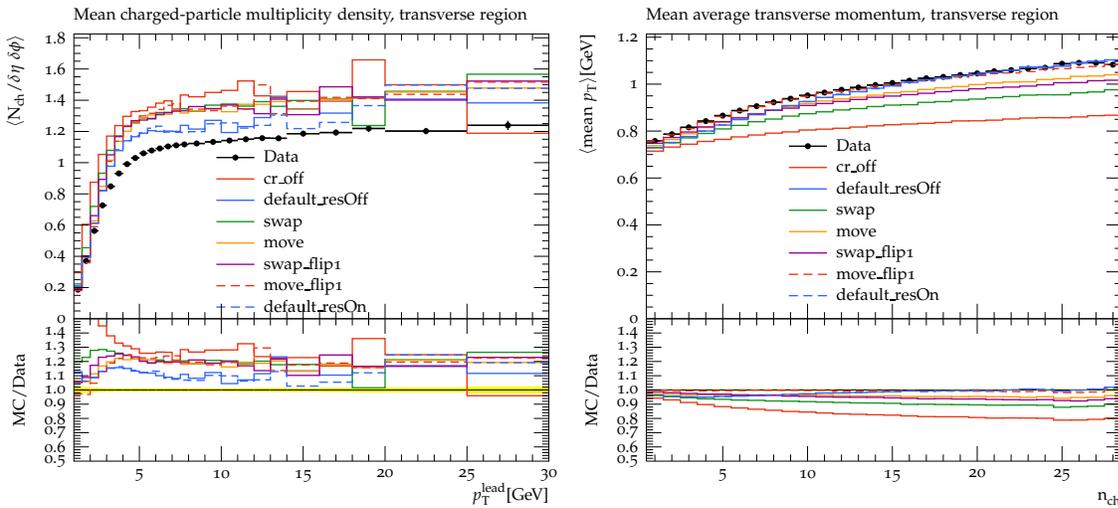


Fig. 5.11: Number of charged particle in the transverse region plotted with respect to the highest transverse momentum in the event (left), average charge particle transverse momentum in the transverse region shown with respect to their multiplicity.

### 5.5.4 Fragmentation function

Colour exchanges eventually lead to different final state particles formed during hadronization process also termed as fragmentation<sup>4</sup>. Probability that hadron  $h$  carries longitudinal momentum fraction  $z$  of parton  $i$  with momentum  $p_i$  is described by fragmentation function

<sup>4</sup>In the Lund string model, colour string **fragments** into hadrons and mezos.

$D_i^h(z, Q)$ . Other observable with some relation to fragmentation function must be used instead due to the lack of information about parton's momentum. Jet fragmentation function and transverse profile were measured by ATLAS at 7 TeV [43] using the following  $F$  function

$$F(z, p_{T\text{jet}}) = \frac{1}{N_{\text{jet}}} \frac{dN_{\text{ch}}}{dz}; \quad z = \frac{p_{\text{jet}} \cdot p_{\text{ch}}}{|p_{\text{jet}}|^2}, \quad (5.9)$$

where  $p_{\text{jet}}$  is the momentum of jet and  $p_{\text{ch}}$  momentum of charged particle inside the jet in question.  $F$  function represents sum over  $D_i^h(z, Q)$  weighted by production of  $i$  parton in hard scattering. To compare with the provided data 7 TeV CR samples ( $\sim 1$  M events) were generated with PYTHIA 8 4C tune as the baseline. The example and comparison of CR schemes for the  $F$  calculated from jets with transverse momentum  $25 < p_T < 40$  is in the left Fig. 5.12. All the models are compared to 7 TeV data. Although more enhanced statistic simulations can lead to better description, no significant difference between the CR models can be obtained from this distribution. It seems that all the CR models describe data much better than CR off at the beginning of the distribution and also the tail is described reasonably well within the statistical uncertainty. In addition to  $F$  distribution, transverse shape of the jets  $f$  containing the information about particle production in the plane transverse to the jet axis was measured. Nevertheless even  $f$  appear to be not very sensitive.

$$f(p_{T\text{rel}}, p_{T\text{jet}}) = \frac{1}{N_{\text{jet}}} \frac{dN_{\text{ch}}}{dp_{T\text{rel}}}; \quad p_{T\text{rel}} = \frac{|p_{\text{jet}} \times p_{\text{ch}}|}{|p_{\text{jet}}|} \quad (5.10)$$

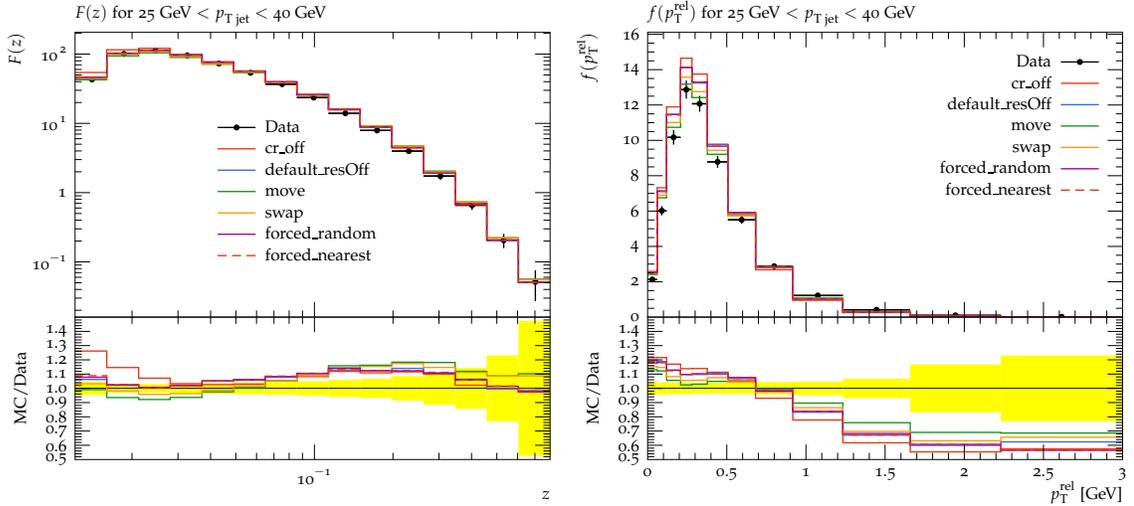


Fig. 5.12: Distribution of  $F$  related to fragmentation function (left) and transverse profile  $f$  (right) for jets with transverse momentum  $25 < p_T < 40$ .

### 5.5.5 Pull angle

Another interesting analysis which was considered as an opportunity to probe CR effects is ATLAS analysis concerning colour flow in the  $t\bar{t}$  events at 8 TeV [7]. Having the possibility

to determine flow of the colour in jets would provide additional information to jet  $p_T$ . Fig. 5.13 shows  $t \rightarrow bW$  process where colour flows between the colour dipole from  $W$  on the left and distortion of colour by connecting  $b$  with one of the quarks from  $W$  on the right. Coloured partons are allowed to radiate and this radiation is predominantly in the direction of the other end of the colour dipole [44] as being *pulled* toward it. Production of showers is a manifestation of the colour flow, therefore the processes can be distinguished on the basis of this colour tracking.

Jet pull vector  $\vec{v}_p^J$  is an observable which provides information about the flow of the colour. It is constructed as

$$\vec{v}_p^J = \sum_{i \in J} \frac{p_T^i |\vec{r}^i| \vec{r}^i}{P_T^J}, \quad (5.11)$$

where the sum runs over all particles inside the jet.  $p_T^i$  is the transverse momentum of particle and  $\vec{r}^i$  is the relative distance of particles to the jet axis in the space of rapidity  $y$  and azimuthal angle  $\phi$ . Angle between the pull vector  $\vec{v}_p^{J_1}$  made of jet  $J_1$  and vector connecting  $J_1$  and some other jet  $J_2$  is called **pull angle**  $\theta_p$  left Fig. 5.14 and tends to be sensitive to the colour connection. When is the  $\theta_p$  calculated between the jets that are not colour connected its value is typically around  $\pi$  and 0 if the jets are colour connected. This effect can be seen in the right Fig. 5.14. In this plot  $\theta_p$  is constructed from pull vector constructed only from charged particles in the lightjet1 and vector connecting both lightjets. Provided figure shows effect of CR schemes only within a few percent, dropping the distribution from the list of potential observables useful for constrain of CR.

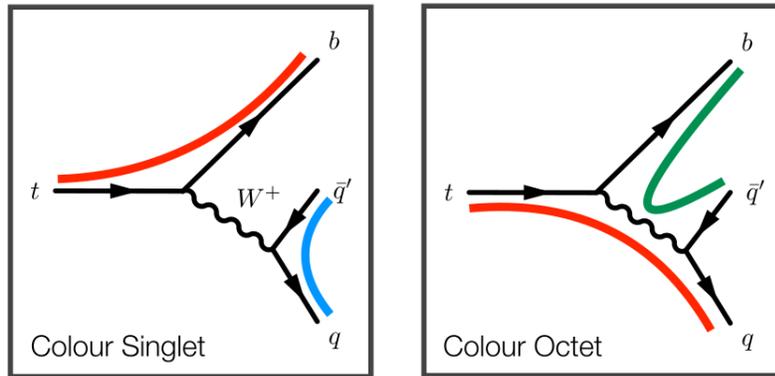


Fig. 5.13: Decay of top quark into  $b$  and  $W^+$  which further decays into colour connected quarks (left) and the same example with distortion of colours resulting into colour connection of one of the quarks from  $W$  with  $b$  [7].

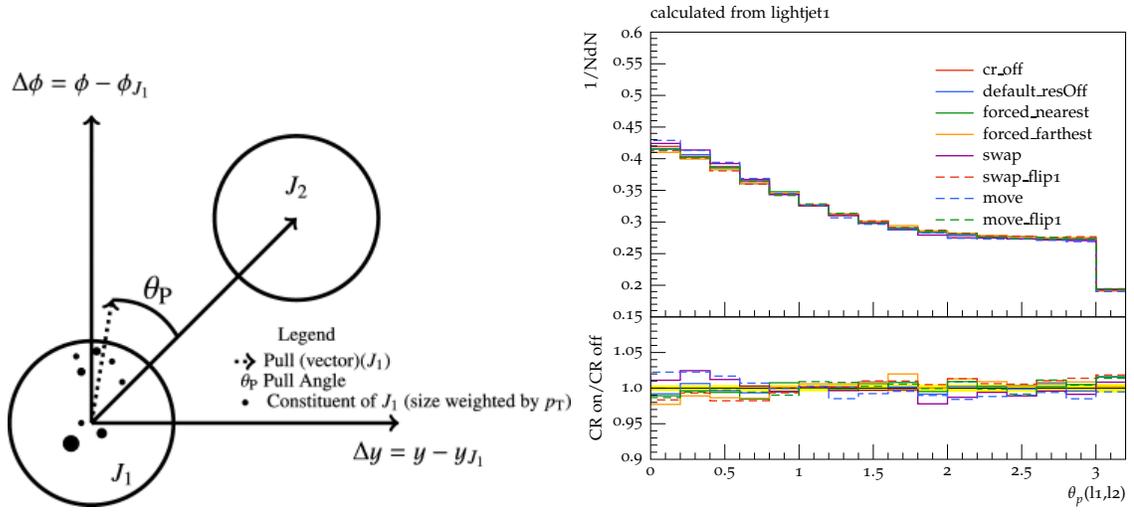


Fig. 5.14: A diagram showing two jets  $J_1, J_2$  in a  $\eta - \phi$  difference space and constructed jet pull angle between the pull vector and vector connecting  $J_1$  with  $J_2$  [7] (left). Pull angle between pull vector constructed from lightjet1 and vector connecting both lightjets for various CR scheme at 13 TeV (right). Only charged particles inside the jet are used to create the pull vector.

## 5.6 Summary of CR analysis

This analysis focused on studying effects of colour reconnection on the reconstruction of top quark mass. Semi-leptonic decays were considered in which  $t$  quark decays into  $b$  quark and  $W^+$  which further decays to quarks whereas  $\bar{t}$  decays to  $\bar{b}$  and  $W^+$  decaying leptonically. Simulation of various PYTHIA 8 colour reconnection models were performed to study their difference in reconstructed mass of top quark. The spread of top mass between the models was found to be around 1.8 GeV. To better estimate systematic uncertainty from the modeling of colour reconnection (for the ATLAS detector currently underestimated and represents less than 10% of the total systematic) it is important to find observables which are sensitive to colour reconnection effects. By measuring the observables models can be constrained to avoid unrealistic behaviors.

Basic properties such as charged particle transverse momentum, their multiplicities and others were studied in jets and underlying event separately. Some of these observables in particular multiplicity provides good discrimination between the various PYTHIA 8 schemes. Further studies of the basic observable in annuli around jets were performed and proved to provide some sensitivity. Discrepancy between the models were found also in the minimum bias simulation which were compared to data discussed in the first part of this thesis. Substantial effort has been made to use existing measurement for the constraints such as study of fragmentation function or colour flow. However both turned out to be not very sensitive.

Several observables were found to be sensitive to the colour exchanges and it is inevitable to perform their measurement which could constrain and better estimate top mass systematic uncertainty originating from this phenomenon.

## Appendix

CR model	$p_{T0}^{\text{ref}}$ [GeV]	$R_{\text{rec}}$	$\alpha$	$\Delta\lambda_{\text{cut}}$
cr off	2.28	-	-	-
default resOff	2.28	10	-	-
default resOn	2.28	10	-	-
move	2.25	-	1	0
move flip1	2.15	-	1	0
move flip2	2.15	-	1	0
swap	2.30	-	1	0
swap flip1	2.20	-	1	0
swap flip2	2.20	-	1	0
forced random	2.28	-	0.075	-
forced nearest	2.28	-	0.075	-
forced farthest	2.28	-	0.075	-
forced smallest $\Delta\lambda$	2.28	-	0.075	-
smallest $\Delta\lambda$	2.28	-	1	-

Tab. 5.3: Selected tuning parameters for CR models in  $t\bar{t}$  semileptonic system. The values of parameters are used with correspondence to article concerning colour reconnection in  $t\bar{t}$  at 7 TeV [31] where parameters were tuned employing ATLAS data from minimum bias and jet shapes measurement.

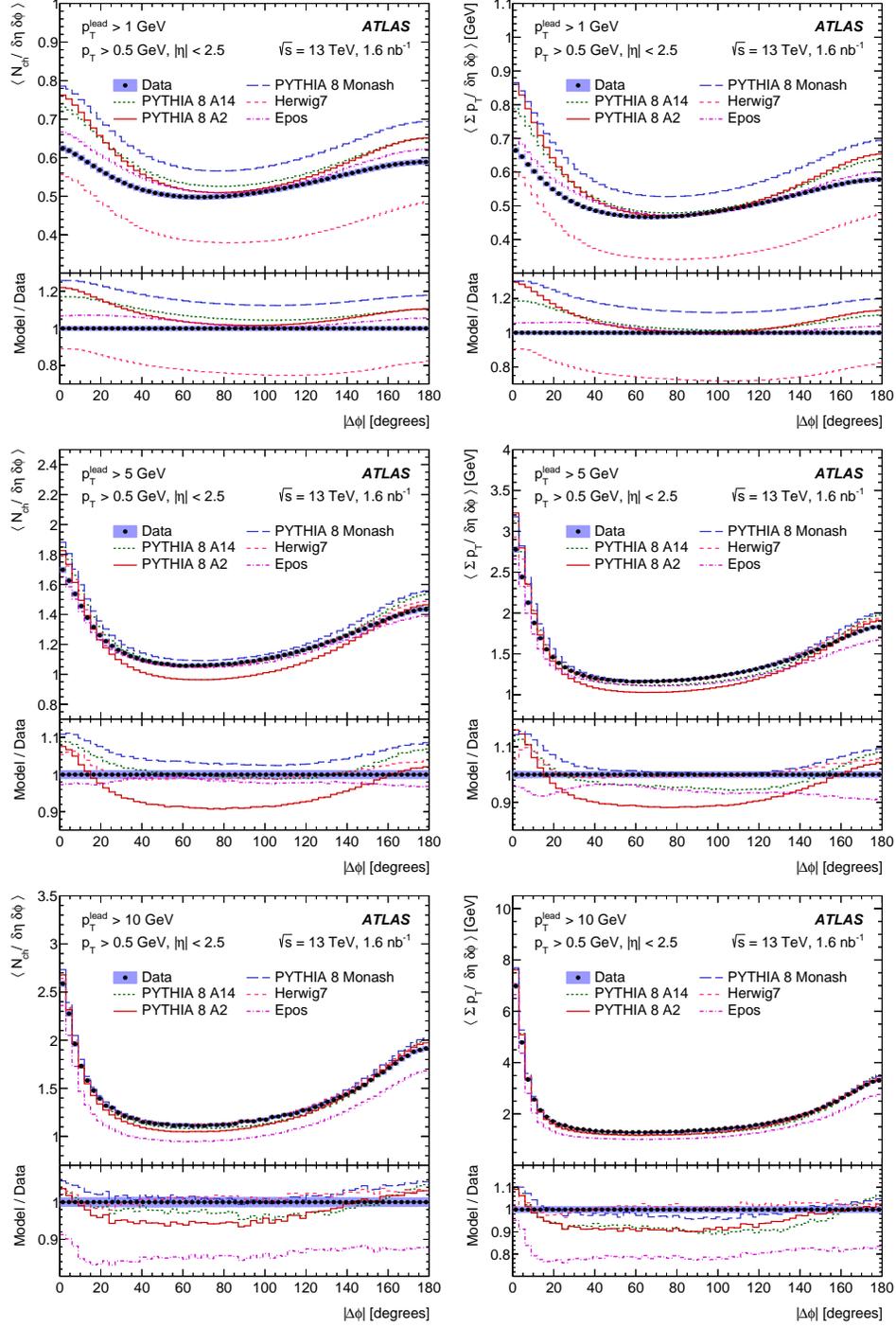


Fig. 5.15: Mean charged-particle multiplicity  $N_{\text{ch}}$  (left), and mean sum of transverse momentum  $\sum p_T$  (right) plotted as a function of  $|\Delta\phi|$  with respect to the leading charged particle  $p_T^{\text{lead}} > 1 \text{ GeV}$  (top),  $5 \text{ GeV}$  (middle), and  $10 \text{ GeV}$  (bottom). The error bars represent statistical uncertainty and the blue band the total combined statistical and systematic uncertainty.

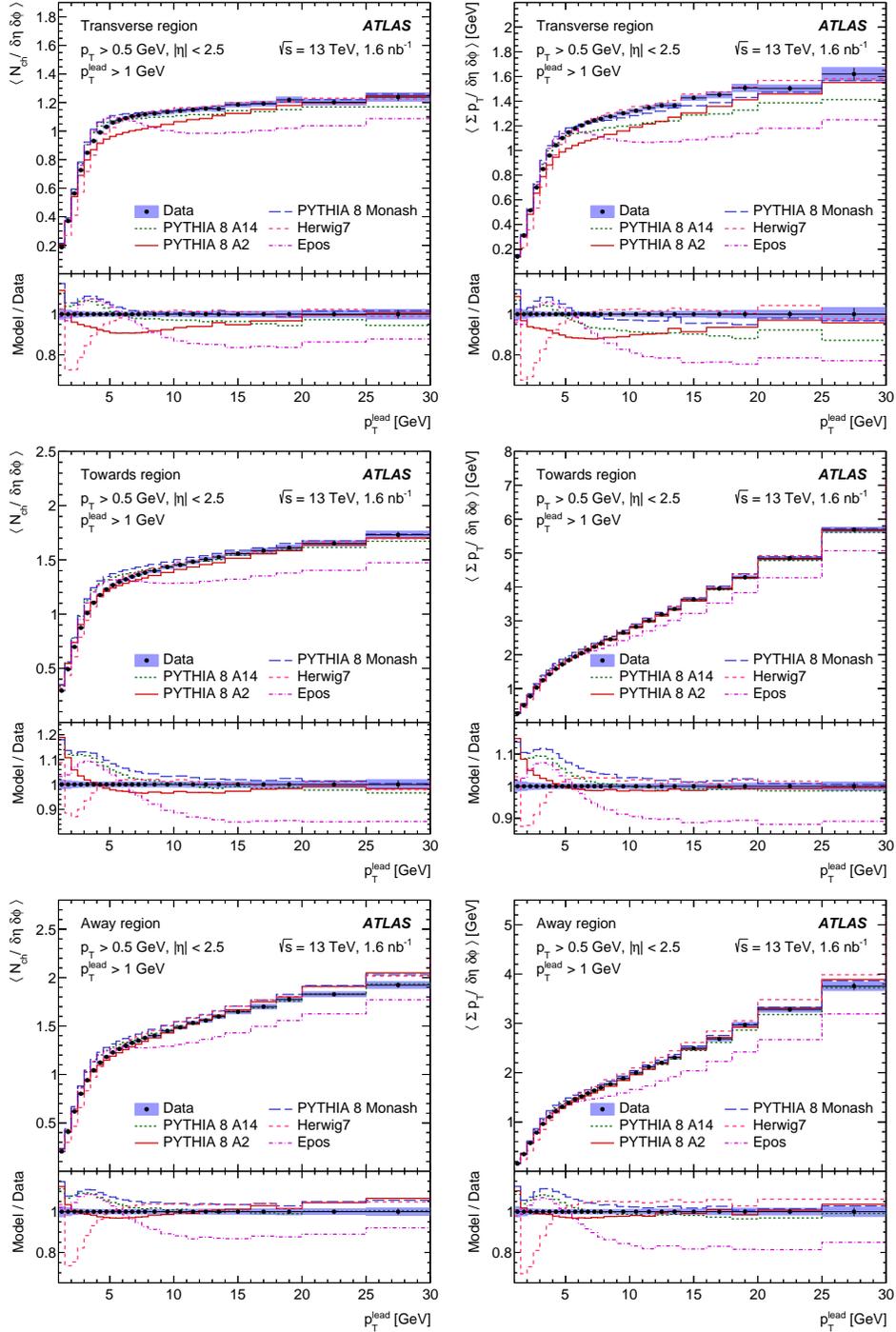


Fig. 5.16: Mean charged-particle multiplicity  $N_{ch}$  (left), and mean sum of transverse momentum  $\sum p_T$  (right) plotted as a function of leading particle transverse momentum  $p_T^{\text{lead}}$  for  $p_T^{\text{lead}} > 1$  GeV for the transverse (top), towards (middle), and away (bottom) azimuthal regions. The error bars on data points represent statistical uncertainty and the blue band the total combined statistical and systematic uncertainty.

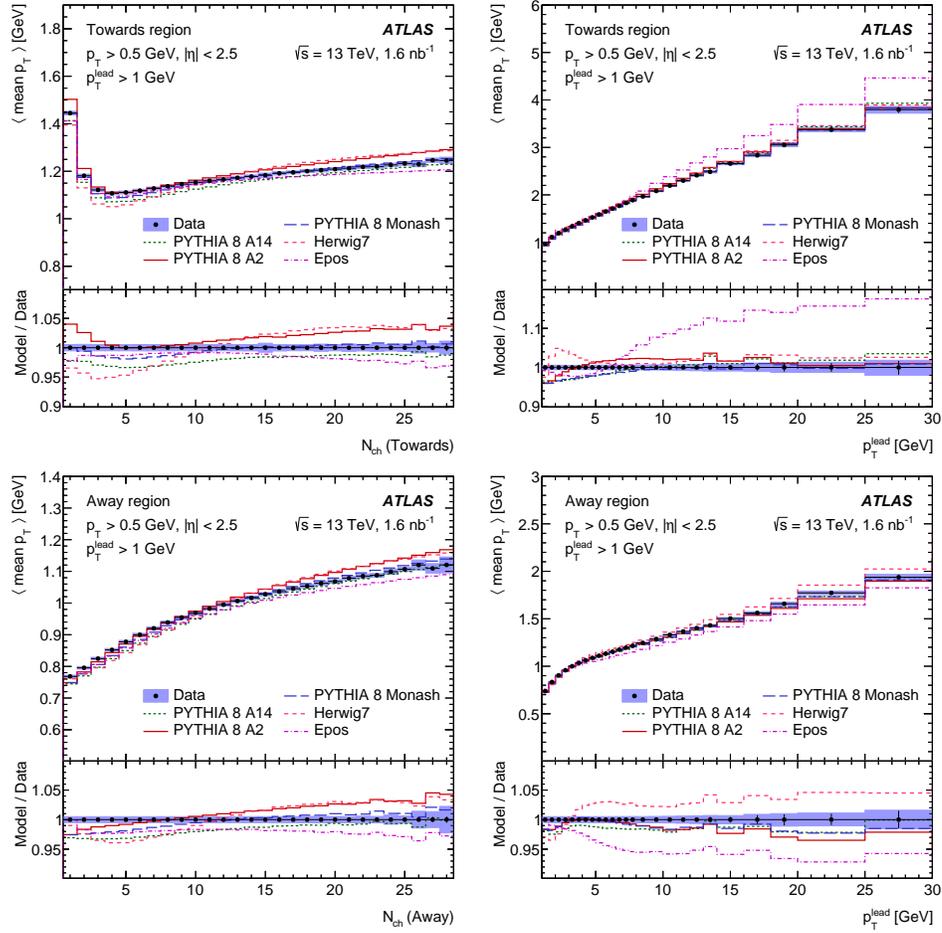


Fig. 5.17: Mean charged-particle average transverse momentum as a function of azimuthal region charged-particle multiplicity  $N_{\text{ch}}$  (left) and as a function of  $p_T^{\text{lead}}$  (right), for each of the towards (top) and away (bottom) regions. The error bars on data points represent statistical uncertainty and the blue band the total combined statistical and systematic uncertainty.

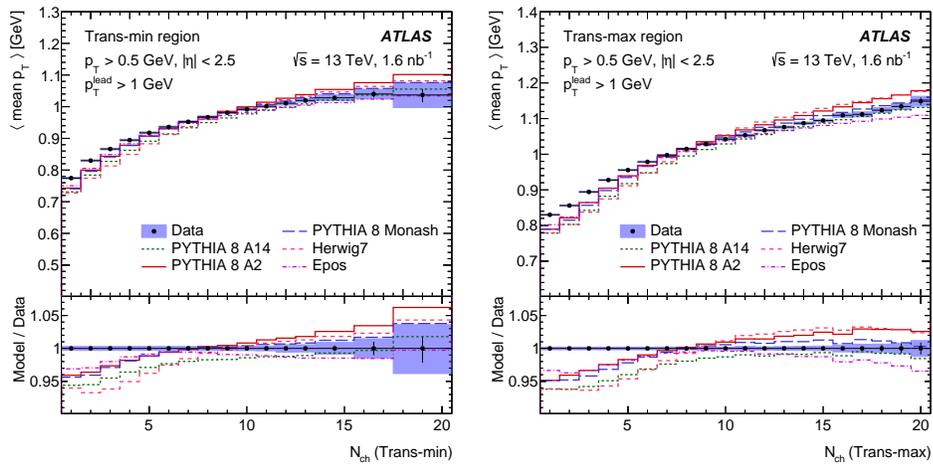


Fig. 5.18: Mean charged-particle average transverse momentum in trans-min (left) and trans-max (right) regions depicted as a function of azimuthal region charged particle multiplicity  $N_{\text{ch}}$  (trans-min) and  $N_{\text{ch}}$  (trans-max), respectively. The error bars represent statistical uncertainty and the blue band the total combined statistical and systematic uncertainty.

# Bibliography

- [1] Symmetrymagazine, *The standard model of particle physics*, <http://www.symmetrymagazine.org/standard-model/>, cited May 5, 2017.
- [2] S. Höche, *Introduction to parton-shower event generators*, 2014. <https://inspirehep.net/record/1328513/files/arXiv:1411.4085.pdf>.
- [3] CERN, *CERN accelerator complex*, <https://bigscience.web.cern.ch/bigscience/en/lhc/lhc2.html>, cited May 5, 2017.
- [4] J. Pequeno, *Computer generated image of the whole ATLAS detector*, <http://cds.cern.ch/record/1095924>, cited May 5, 2017.
- [5] ATLAS Collaboration, A. Morley et al., *Supporting note for the Minimum Bias publication: Charge particle multiplicity in pp interactions at 13 TeV measured with the ATLAS detector*, Tech. Rep. ATL-COM-PHYS-2015-313, CERN, Geneva, Apr, 2015. <https://cds.cern.ch/record/2011604>.
- [6] S. Argyropoulos, *Overview of CR models in Pythia8*, <https://indico.cern.ch/event/602289/>, cited May 5, 2017.
- [7] ATLAS Collaboration, G. Aad et al., *Measurement of colour flow with the jet pull angle in  $t\bar{t}$  events using the ATLAS detector at  $\sqrt{s} = 8$  TeV*, Phys. Lett. **B750** (2015) 475–493, arXiv:1506.05629 [hep-ex].
- [8] Super-Kamiokande Collaboration, Y. Fukuda et al., *Evidence for Oscillation of Atmospheric Neutrinos*, Phys. Rev. Lett. **81** (1998) 1562–1567, <https://link.aps.org/doi/10.1103/PhysRevLett.81.1562>.
- [9] ATLAS Collaboration, G. Aad et al., *Measurement of underlying event characteristics using charged particles in pp collisions at  $\sqrt{s} = 900\text{GeV}$  and 7 TeV with the ATLAS detector*, Phys. Rev. **D83** (2011) 112001, arXiv:1012.0791 [hep-ex].
- [10] ATLAS Collaboration, *Leading Track Underlying Event at 13 TeV*, Tech. Rep. ATL-PHYS-PUB-2015-019, CERN, Geneva, Jul, 2015.
- [11] ATLAS Collaboration, M. Aaboud et al., *Measurement of charged-particle distributions sensitive to the underlying event in  $\sqrt{s} = 13$  TeV proton-proton collisions with the ATLAS detector at the LHC*, arXiv:1701.05390 [hep-ex].

- [12] ATLAS Collaboration Collaboration,, *HEPData to ATLAS underlying event measurement at 13 TeV*, <https://hepdata.net/record/ins1509919>.
- [13] ATLAS Collaboration Collaboration,, *Rivet routine to ATLAS underlying event measurement at 13 TeV*, [https://www.hepforge.org/archive/rivet/contrib/ATLAS\\_2017\\_I1509919.tar.gz](https://www.hepforge.org/archive/rivet/contrib/ATLAS_2017_I1509919.tar.gz).
- [14] J. C. Collins, *Sudakov form-factors*, Adv. Ser. Direct. High Energy Phys. **5** (1989) 573–614, [arXiv:hep-ph/0312336](https://arxiv.org/abs/hep-ph/0312336) [hep-ph].
- [15] T. Sjostrand and M. van Zijl, *A Multiple Interaction Model for the Event Structure in Hadron Collisions*, Phys. Rev. **D36** (1987) 2019.
- [16] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852–867, [arXiv:0710.3820](https://arxiv.org/abs/0710.3820) [hep-ph].
- [17] J. Bellm, S. Gieseke, D. Grellscheid, S. Plätzer, M. Rauch, C. Reuschle, P. Richardson, P. Schichtel, M. H. Seymour, A. Siódmok, A. Wilcock, N. Fischer, M. A. Harrendorf, G. Nail, A. Papaefstathiou, and D. Rauch, *Herwig 7.0/Herwig++ 3.0 release note*, The European Physical Journal C **76** (2016) 196, [http://dx.doi.org/10.1140/epjc/s10052-016-4018-8](https://dx.doi.org/10.1140/epjc/s10052-016-4018-8).
- [18] T. Pierog and K. Werner, *EPOS Model and Ultra High Energy Cosmic Rays*, Nuclear Physics B - Proceedings Supplements **196** (2009) 102 – 105, [http://www.sciencedirect.com/science/article/pii/S0920563209006537](https://www.sciencedirect.com/science/article/pii/S0920563209006537).
- [19] ATLAS Collaboration, G. Aad et al., *Charged-particle distributions in  $\sqrt{s} = 13$  TeV pp interactions measured with the ATLAS detector at the LHC*, Phys. Lett. **B758** (2016) 67–88, [arXiv:1602.01633](https://arxiv.org/abs/1602.01633) [hep-ex].
- [20] ATLAS Collaboration, F. A. C. Hügging, *The ATLAS Pixel Insertable B-layer (IBL)*, [arXiv:1701.05390](https://arxiv.org/abs/1701.05390) [hep-ex].
- [21] T. Kanai and O. Jinnouchi, *Measurement of the material in the ATLAS Inner Detector using hadronic interactions for an improvement in the track reconstruction*. PhD thesis, Tokyo Inst. Tech., Feb, 2013. <https://cds.cern.ch/record/1528163>. Presented 18 Feb 2013.
- [22] CDF Collaboration Collaboration, D. e. a. Acosta, *Underlying event in hard interactions at the Fermilab Tevatron  $\bar{p}p$  collider*, Phys. Rev. D **70** (2004) 072002, <https://link.aps.org/doi/10.1103/PhysRevD.70.072002>.
- [23] CDF Collaboration, T. Aaltonen et al., *Studying the Underlying Event in Drell-Yan and High Transverse Momentum Jet Production at the Tevatron*, Phys. Rev. **D82** (2010) 034001, [arXiv:1003.3146](https://arxiv.org/abs/1003.3146) [hep-ex].
- [24] S. e. Chatrchyan, *Measurement of the underlying event in the Drell-Yan process in proton-proton collisions at  $\sqrt{s} = 7$  TeV*, The European Physical C **72** (2012) 2080.
- [25] CMS Collaboration,, *Underlying Event Measurements with Leading Particles and Jets in pp collisions at  $\sqrt{s} = 13$  TeV*, <http://inspirehep.net/record/1406033>, 2015.

- [26] G. Marchesini and B. R. Webber, *Associated transverse energy in hadronic jet production*, Phys. Rev. D **38** (1988) 3419–3426, <https://link.aps.org/doi/10.1103/PhysRevD.38.3419>.
- [27] J. W. Monk and C. Oropeza-Barrera, *The HBOM Method for Unfolding Detector Effects*, Nucl. Instrum. Meth. **A701** (2013) 17–24, [arXiv:1111.4896](https://arxiv.org/abs/1111.4896) [hep-ex].
- [28] A. Buckley, J. Butterworth, L. Lonnblad, D. Grellscheid, H. Hoeth, J. Monk, H. Schulz, and F. Siegert, *Rivet user manual*, Comput. Phys. Commun. **184** (2013) 2803–2819, [arXiv:1003.0694](https://arxiv.org/abs/1003.0694) [hep-ph].
- [29] UA1 Collaboration, *A study of the general characteristics of proton-antiproton collisions at  $\sqrt{s} = 0.2$  to  $0.9$  TeV*, Nucl. Phys. **B335** (1990).
- [30] D. Eriksson, G. Ingelman, and J. Rathsmann, *Color rearrangements in B-meson decays*, Phys. Rev. D **79** (2009) 014011, <https://link.aps.org/doi/10.1103/PhysRevD.79.014011>.
- [31] S. Argyropoulos and T. Sjöstrand, *Effects of color reconnection on  $t\bar{t}$  final states at the LHC*, JHEP **11** (2014) 043, [arXiv:1407.6653](https://arxiv.org/abs/1407.6653) [hep-ph].
- [32] ATLAS Collaboration, G. Aad et al., *Determination of the top-quark pole mass using  $t\bar{t} + 1$ -jet events collected with the ATLAS experiment in 7 TeV pp collisions*, JHEP **10** (2015) 121, [arXiv:1507.01769](https://arxiv.org/abs/1507.01769) [hep-ex].
- [33] CMS Collaboration, *Measurement of the top-quark mass in  $t$   $t$ -bar events with lepton+jets final states in pp collisions at  $\sqrt{s}=8$  TeV*, Tech. Rep. CMS-PAS-TOP-14-001, CERN, Geneva, 2014. <https://cds.cern.ch/record/1690093>.
- [34] D0 Collaboration, V. M. Abazov et al., *Precision measurement of the top-quark mass in lepton+jets final states*, Phys. Rev. Lett. **113** (2014) 032002, [arXiv:1405.1756](https://arxiv.org/abs/1405.1756) [hep-ex].
- [35] ATLAS, CDF, CMS, D0 Collaboration, *First combination of Tevatron and LHC measurements of the top-quark mass*, [arXiv:1403.4427](https://arxiv.org/abs/1403.4427) [hep-ex].
- [36] B. Andersson, G. Gustafson, and B. Soderberg, *A Probability Measure on Parton and String States*, Nucl. Phys. **B264** (1986) 29.
- [37] J. R. Christiansen, *Colour Reconnection - Models and Tests*, PoS **EPS-HEP2015** (2015) 443, [arXiv:1510.04859](https://arxiv.org/abs/1510.04859) [hep-ph].
- [38] ATLAS Collaboration, G. Aad et al., *Measurement of jet shapes in top-quark pair events at  $\sqrt{s} = 7$  TeV using the ATLAS detector*, Eur. Phys. J. **C73** (2013) 2676, [arXiv:1307.5749](https://arxiv.org/abs/1307.5749) [hep-ex].
- [39] ATLAS Collaboration, G. Aad et al., *Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC*, New J. Phys. **13** (2011) 053033, [arXiv:1012.5104](https://arxiv.org/abs/1012.5104) [hep-ex].

- [40] ATLAS Collaboration, G. Aad et al., *Differential top-antitop cross-section measurements as a function of observables constructed from final-state particles using pp collisions at  $\sqrt{s} = 7$  TeV in the ATLAS detector*, JHEP **06** (2015) 100, arXiv:1502.05923 [hep-ex].
- [41] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph].
- [42] R. Brun and F. Rademakers, *ROOT - An Object Oriented Data Analysis Framework*, <https://root.cern.ch/>.
- [43] ATLAS Collaboration, G. Aad et al., *Measurement of the jet fragmentation function and transverse profile in proton-proton collisions at a center-of-mass energy of 7 TeV with the ATLAS detector*, Eur. Phys. J. **C71** (2011) 1795, arXiv:1109.5816 [hep-ex].
- [44] J. Gallicchio and M. D. Schwartz, *Seeing in Color: Jet Superstructure*, Phys. Rev. Lett. **105** (2010) 022001, arXiv:1001.5027 [hep-ph].