THE HENRYK NIEWODNICZAŃSKI
INSTITUTE OF NUCLEAR PHYSICS
POLISH ACADEMY OF SCIENCES

Evidence of the Exclusive Jet Production Using the ATLAS Detector

Paula Erland

Prepared under the supervision of
prof. dr hab. Janusz J. Chwastowski
dr Maciej Trzebiński (auxiliary supervisor)

Contents

Foreword

Very early in human history, people started to ask questions concerning their surroundings. This curiosity leads mankind thru ages of discoveries and innovations. Still, many questions remain unanswered. Among them are very basic ones, like how the Universe started or what are its fundamental components.

One way to find the answers to questions about the nature of our Universe is to study the basic constituents of matter. This can be done using the most powerful microscopes human-build so far – the particle accelerators – and their eyepieces – the particle detectors. Their general principle of operation is to collide accelerated bunches of particles (*e.g.* electrons, protons, ions) and detect products of such a collision. Nowadays, due to the complexity of these devices, studies are carried out by collaborations gathering physicists from all around the world. One such laboratory is CERN (Conseil Européen pour la Recherche Nucléaire) hosting the largest and the highest energy human-build accelerator – the Large Hadron Collider (LHC). This accelerator together with its detectors allows studies of various phenomena observed in interactions of colliding particles. This thesis describes the performed studies concerning the production of particle jets in diffractive processes.

The first Chapter introduces basics of the Standard Model, a commonly accepted theory describing elementary particles and interactions between them. It discusses also Quantum Chromodynamics (QCD) – a theory describing interactions of quarks and gluons. In this chapter one can also find the description of diffractive processes, one of the possible interaction types in proton-proton collisions and the structure of a proton and Pomeron. Next, the processes of Single Diffractive Di-Jet[1] (SD JJ) is described. Finally, the Exclusive Di-jet production (EXC JJ) is presented.

The second Chapter provides information about the structure of the Large Hadron Collider. Its understanding is essential since the LHC accelerator settings (the so-called machine optics) influence the scattered forward proton trajectory. The forward protons are scattered at very small angles, predominantly into the beam pipe. Their presence is one of the features of the diffractive or exclusive events. A description of "A Toroidal LHC ApparatuS" (ATLAS) detector, which is used for the jet measurement, is given in Chapter 3. Chapter 4 introduces the ATLAS Forward Proton detectors (AFP) providing information concerning the forward scattered proton. Both Chapters 3 and 4 are equally important

[1]In this thesis prefix "Di-" is often skipped.

since the data collected by the considered detectors are used in the present work.

The above introductory part is followed by the description of the analysis of the Single Diffractive Jet processes. First, the studies dedicated to the selection of the AFP jet triggers most suitable for this analysis are presented. Next, the selection of SD JJ events from the 2017 ATLAS data is discussed. Finally, Chapter 6 shows how the application of the exclusive di-jet selection impacts data. Here, the properties of EXC JJ candidates are discussed.

Technical aspects of the analyses done during my PhD are described in the appendixes. Appendix A presents the studies performed for the AFP Silicon Tracker detectors (SiT). The goal of this work was to develop algorithms for searches of low-efficiency, dead or hot pixels. The performance of these algorithms is presented in this Appendix.

In Appendix B, the performance of the AFP exclusive jet trigger algorithm is discussed. The algorithm was designed to enable efficient registration of the exclusive jet processes. The idea was to compare the proton position predicted from the di-jet system kinematics in the central detector to that of the proton registered in the AFP. During the study, various scenarios of the LHC optics were analysed as well as the possible improvements of the trigger performance (*e.g.* studies at the trigger efficiency with various radii of acceptance).

Appendix C contains studies dedicated to the performance of AFP jet trigger algorithms. The efficiencies of such triggers available during 2017 data taking were calculated. The efficiencies were analysed as a function of time, leading jet transverse momentum and proton energy loss. In addition, the dead-time effect for AFP triggers is shown and described.

Personal Contribution

I joined the IFJ PAN ATLAS Experiment group during my BSc. studies at the Cracow University of Technology. My first task was devoted to the work on the Jet-Gap-Jet (JGJ) scattering amplitude parametrisation. Based on full amplitude calculations, the parametrisation formulas for the leading logarithm (LL) and next-to-leading logarithm (NLL) approximations were obtained. These results are described in [1].

During my MSc. years, I contributed to the development of the GenEx Monte Carlo generator. I implemented the elastic scattering amplitudes for the various Pomeron models. In addition, a differential elastic cross-section as a function of the scattered proton transverse momentum for unpolarised and polarised protons was studied. This work resulted in publication [2].

At the end of my MSc. studies I obtained *Diamentowy Grant* from Polish Ministry of Science and Education [2]. This project allowed me to conduct studies presented in this work as well as to contribute to the ATLAS Roman Pots (ARP) group works concerning performance studies, technical work and on-call duties.

My collaboration with the ATLAS experiment and in particular with the ARP group started during the 2015 test beam, just before AFP installation at the LHC. I had an opportunity to take part in the tests of the Time-of-Flight system. The main goal was to study the performance of the quartz bars and the cross-talk between them. The results of these tests are described in Ref. [3].

I covered the AFP on-call expert shifts during the LHC Run 2 (2017-2018) and later during the Run 3 (2022) data-taking. This duty is to be the first contact person in case of problems related to the AFP detectors. The task is to solve them as quickly as possible to ensure an efficient operation of the AFP. In case of more serious matters, I was a bridge between the AFP detector experts and the ATLAS control room (operation) team. Taking part in these activities resulted in entrusting me with several talks about the AFP detector status during various conferences, see Refs [4–6].

During the LHC Long Shutdown 2, I participated in a few technical tasks. I performed the measurements of the temperature distribution inside the AFP pots. These measurements were done in the laboratory using the pot mock-up. The goal was to check if the temperature can be lowered by providing additional heat sinks in the pots.

To become the ATLAS author, one needs to complete the so-called qualification task. My task was to validate the efficiency of the exclusive jet trigger algorithm. This was achieved by comparing the proton position registered in the AFP to that predicted from the di-jet system properties. The results of these studies are presented in Appendix B, conference proceeding [7] and ATLAS internal note [8]. I was qualified as ATLAS author on August 15^{th}, 2019.

[2] *Badanie właściwości ekskluzywnej produkcji dżetów na podstawie danych zebranych przez detektor ATLAS*, 0138/DIA/2017/46.

During my PhD, I performed studies of the AFP SiT detectors' efficiency. Their results are described in Appendix A and the ATLAS internal note [9]. Discussion of the SD JJ analysis can be found in the ATLAS internal note [10].

[1] P. Swierska, M. Trzebinski, *BFKL Amplitude Parametrisation for the Jet-Gap-Jet Events at the LHC Energies*, Acta Phys. Pol. B 46 (2015), 2415.

[2] P. Erland, R. Staszewski, M. Trzebinski, R. Kycia, *Elastic Hadron Scattering in Various Pomeron Models*, Acta Phys. Pol. B 48 (2017) 981.

[3] J. Lange et. al (P. Swierska), *Beam tests of an integrated prototype of the ATLAS Forward Proton detector*, JINST 11 (2016) P09005.

[4] P. Erland, *ATLAS Forward Proton Detectors Status and Plans*, Proceedings of Science, 06.09.2019, Volume 348. Contribution presented at "The 27th International Workshop on Vertex Detectors" Chennai, India.

[5] P. Erland, *Performance of ATLAS Forward Proton Detector during LHC Run 2*, PoS(LHCP2022)294. Poster at "10th Edition of the Large Hadron Collider Physics Conference".

[6] P. Erland, *Overview of ATLAS Forward Proton Detectors for LHC Run 3 and Plans for the HL-LHC*, PoS (ICHEP2022) 662. Contribution at "International Conference of High Energy Physics" Bologna, Italy.

[7] P. Erland, *On the Exclusive Jet Measurements at the LHC*, Proc. SPIE Int. Soc. Opt. Eng. **11176** (2019). Contribution at "XLIV-th IEEE-SPIE Joint Symposium" Wilga, Poland.

[8] P. Erland, M. Trzebinski, *Validation of the AFP HLT Exclusive Jet Trigger*, ATL-COM-FWD-2019-021.

[9] P. Erland, M. Trzebinski, *Dead and Low Efficiency Pixels in AFP Silicon Detector*, ATL-COM-FWD-2021-006.

[10] P. Erland, *Analysis with AFP and ATLAS Detectors*, arXiv:2210.13539, doi:10.48733/978-83-63542-30-6. Contribution at "PPSS Alumni Conference", Krakow, Poland.

Chapter 1

State of the Art

Discoveries in the domain of exact sciences should be described by a theory. In the case of High Energy Physics (HEP), the widely used one is the Standard Model (SM). Despite its undoubted success in explaining many phenomena, it is not a complete theory. For example, it does not explain/include gravitation, dark matter or neutrino masses. Nevertheless, the Standard Model works well to explain diffractive jet phenomena, which are a subject of this thesis. In this chapter, the Standard Model will be briefly described. The stress will be put on Quantum Chromodynamics (QCD) which is nowadays used to describe the diffractive physics processes.

1.1 The Standard Model

The Standard Model is a theory describing elementary particles and the interactions between them. According to it, the particles are divided into two classes: fermions, *i.e.* particles with half spin and bosons – particles with integer spin. Figure 1.1 presents the list of fundamental particles.

Fermions can be further divided into leptons and quarks. Currently, six types of quarks and six types of leptons are known. Each of these particles has an anti-particle that has the same mass but the opposite electric charge – *e.g.* the electron's anti-particle is called a positron and has a positive charge +e. All fermions are subject to the Pauli exclusion principle – there cannot be two fermions in the same quantum state at a time.

Each quark carries the quantum number called colour. This concept was introduced to explain the existence of the Δ^{++} particle, which otherwise would violate Pauli's exclusion principle in the *uuu* quark system. Three types of colour charge are distinguished, commonly referred to as red, green and blue. Naturally, there exist three "anti-colours": anti-red, anti-green and anti-blue.

Only "white" particles, *i.e.* those with zero colour charge (the colour singlet states), are observed in nature. Free quarks are not observed – they are bounded into hadrons: systems

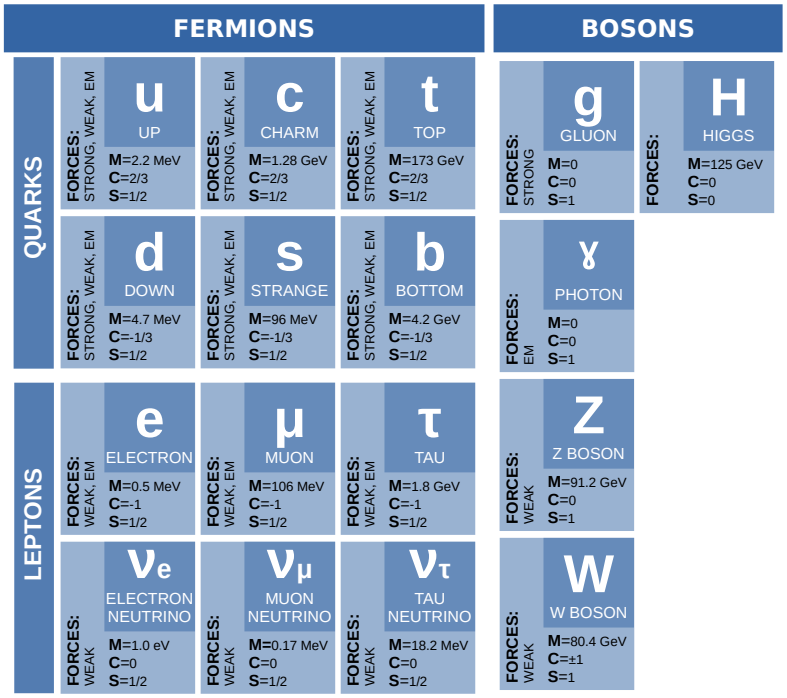


Figure 1.1: The elementary particles described in the Standard Model.

of quarks. Hadrons consisting of three quarks are called baryons (*e.g.* a proton consists of *uud* quarks), and those composed of a quark and anti-quark are mesons (*e.g.* a positively charged pion π^+ consist of quark u and anti-quark $\bar{d}$).

The Standard Model also describes interactions of fundamental particles. One can distinguish electromagnetic, weak and strong forces. They vary in terms of strength and interaction range. Each force has its own carrier being a boson. The strengths, ranges and carriers of the discussed interactions are listed in Table 1.1 which includes also gravitation for comparison. Gravitation is typically neglected in elementary particle physics due to its very weak interaction and the fact that its carrier has not yet been discovered.

Table 1.1: Properties of three types of interactions described by the Standard Model. Additionally, the gravitational force is listed.

interaction type	**carrier**	**strength**	**range [m]**
strong	gluon	1	10^{-15}
electromagnetic	photon	10^{-2}	∞
weak	W^+, W^-, Z^0	10^{-13}	10^{-18}
gravitation	graviton	10^{-38}	∞

It is also worth mentioning a particle called the Higgs boson [1, 2] whose existence was confirmed in 2012. It is responsible for giving masses to elementary particles. One should note that not all particles interact with each other and some of them (gluons, Higgs and W^+, W^- bosons) can interact with themselves. As it will be briefly described later, the fact that gluons are self-interacting has profound consequences for the theoretical description of processes, in which they are exchanged.

One should note that the number of the experimentally observed particle states is very large, *cf.* [3]. Although, each of them can be attributed to one of the above-mentioned groups relying on their quantum numbers.

1.1.1 Feynman Diagrams and Mandelstam Variables

The complicated processes of particle production happening during proton-proton collisions at the LHC are usually described using the language of the Feynman diagrams. The lines in diagrams represent different types of particles and the vertices represent interactions between them, see Table 1.2.

Table 1.2: Description of symbols used in the Feynman diagrams.

Line type	**Type of particles**
	quarks and leptons
	anti-quarks and anti-leptons
	bosons W, Z^0, photons
	gluons
	complex processes, for example a multi-gluon exchange
	interaction vertex

To each symbol, a corresponding mathematical formula is ascribed which allows the calculation of the desired processes.

The Mandelstam variables are commonly used to describe the basic scattering process of two particles into two outgoing particles. Those invariant variables, s, t and u, representing different channels in which the interaction can occur, are defined as:

$$s = (p_1 + p_2)^2 = (p_3 + p_4)^2,$$

$$t = (p_1 - p_3)^2 = (p_2 - p_4)^2,$$

$$u = (p_1 - p_4)^2 = (p_2 - p_3)^2,$$

where p_i represents the four-momentum of an i^{th} particle participating in the interaction. The corresponding Feynman diagrams are presented in Figure 1.2.

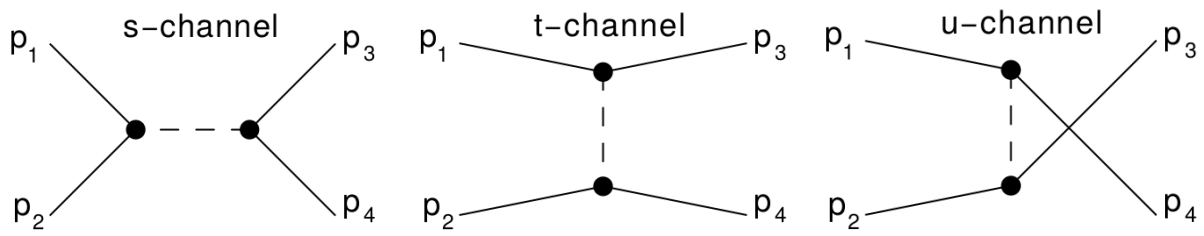


Figure 1.2: Feynman diagrams illustrating the two-to-two process.

1.2 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is a theory describing strong interactions of coloured particles: quarks and gluons. It was already stated that free quarks do not exist in nature and they are bounded in hadrons. It was also mentioned that the gluons are self-interacting. QCD describes an interaction as an exchange of a colour field. The calculations use a perturbative approach which works very well in the case of the short-range, large-scale processes due to the asymptotic freedom, which implies that the coupling constant decreases with increasing value of the scale, decreasing distance.

The perturbative calculations can be used in cases where the exact solutions are not known and approximate methods must be applied. They deliver a solution of a specific problem decomposed in a power series of a small parameter, named α_S, where the first term (A_0) is the known solution to the initial problem and the next terms (A_i) are the high-order terms. An approximation of the full solution, A, can be written as:

$$A = A_0 + \alpha_S A_1 + \alpha_S^2 A_2 + ...$$

The QCD equations are usually solved up to the leading (A_1) or next-to-leading (A_2) order terms. However, some results concerning specific problems were obtained in an even higher order.

There exist also long-range, small-scale processes to which the perturbative methods can not be applied. This is a consequence of the shape of a potential of two colour charges. With increasing separation, the field energy grows until a point where the creation of a quark/anti-quark pair is energetically preferred. This non-perturbative area is typically described using phenomenological models and also those based on analytical properties of the scattering amplitude.

1.2.1 Proton Structure

Protons are compound particles. Since this thesis exploits the data gathered for the proton-proton collisions, the proton structure is worth a brief discussion.

The parton model assumes that a compound particle consists of point-like objects called

partons. In the first approximation, a proton consists of the three so-called valence quarks: two u and one d. With an increasing value of the resolution scale, a proton reveals a very rich structure. Research on the proton structure was started at SLAC in the 1960s [4–6] and was successfully continued at the HERA accelerator [7] in (anti-)electron-proton collisions. These measurements confirmed the Bjorken scaling as well as the scaling violation. The obtained results demonstrate a very complicated picture: in addition to the valence quarks, a proton contains a "sea" of quarks, anti-quarks and gluons [8]. This structure is described by the Parton Density Functions (PDF [9–11]) representing the probability densities to find a parton, of a given kind, carrying the fraction x of the proton momentum at a given energy scale Q^2 (*e.g.* the momentum transfer).

An example of the PDF functions for $Q^2 = 1.9$ GeV^2 and $Q^2 = 10$ GeV^2 is shown in Figure 1.3. For $x > 0.1$, the valence quark contributions are the largest. With decreasing value of x, the contribution of gluons starts to dominate (it should be noted that the curves corresponding to gluons and c quarks are scaled by a factor of 20). This domination becomes stronger with increasing energy scale, Q^2 (*cf.* left and right plot). The rise of the gluon PDF with decreasing x implies also saturation at very low x values as a consequence of the unitary. Qualitatively, a similar behaviour, however not so strong, is observed for the "sea" contribution.

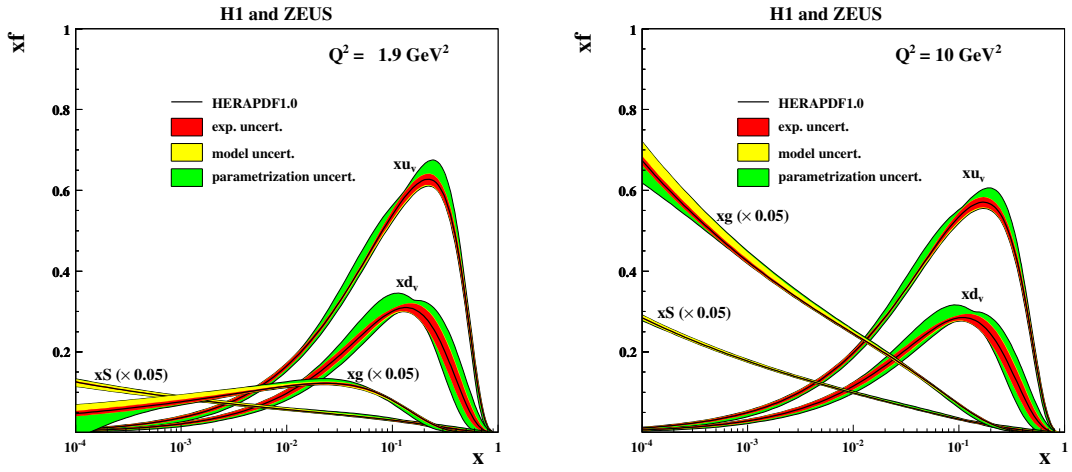


Figure 1.3: Parton Density Function of the proton from HERAPDF1.0 data plotted as a function of the momentum fraction x for $Q^2 = 1.9$ GeV^2 (**left**) and $Q^2 = 10$ GeV^2 (**right**). From [7].

1.3 Diffractive Physics

A strong interaction of two hadrons is described by an exchange of a colour field. Some of such processes are due to an exchange of a colour singlet pictured at the lowest order

as an exchange of two colour-compensated gluons. Such interactions are named diffractive ones. Historically, this comes from the observation of a diffractive-like shape of the angular distribution in the elastic scattering: $pp \to pp$, $\pi^- p \to \pi^- p$. Diffractive processes are described as those in which "no quantum numbers are exchanged between high-energy colliding particles" or processes in which "a Pomeron is exchanged", where a Pomeron is an object postulated within the Regge theory.

The Regge theory [12] is based on the analytical properties of the scattering amplitudes and the unitary nature of the scattering matrix. In this theory, the exchanged object is not a single particle, but an ensemble of particles lying on the so-called trajectory which is described as:

$$\alpha(t) = \alpha_R(0) + \alpha'_R(0)t,$$

where $\alpha_R(0)$ is the intercept, $\alpha'_R(0)$ is the slope of the trajectory (see Figure 1.4, left).

The Pomeron was introduced in the Regge theory in order to explain the observed increase of the cross-sections (see Figure 1.4, right) with energy. The phenomenological fit to the experimental data performed by Donnachie and Landshoff [13] yielded $\alpha_{\mathbb{P}}(0) = 1.0808$ and the slope $\alpha'_{\mathbb{P}}(0) = 0.25$ GeV^{-2} and is presented in Figure 1.4 (left plot, green line).

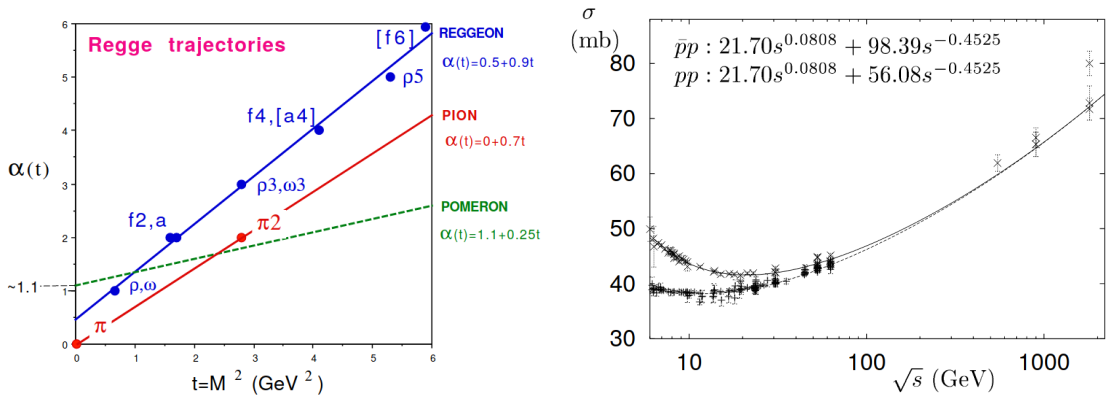


Figure 1.4: **Left:** Regge trajectories for pions and mesons ρ, ω, f and a (straight lines) and Pomeron (dotted line). From [14]. **Right:** cross-section in energy function in the centre of mass for proton-proton collisions (pp, dotted line) and proton-anti-proton ($\bar{p}p$, continues line). From [15]

.

1.3.1 Soft and Hard Diffraction

Depending on the energy scale involved, the diffractive processes can be divided into soft and hard diffraction. Soft diffractive ones are usually described as those where the energy scale is lower than the hadron scale (<1 GeV) and the perturbative QCD methods can not be used. Contrary to soft diffraction, in hard diffraction, the involved energy scale is large ($\gg$1 GeV) and the perturbative approach can be applied.

Soft processes are responsible for the majority of events in hadron-hadron collisions. One can distinguish elastic and inelastic processes. The Feynman diagram of the elastic scattering process, $pp \to pp$, is shown in Figure 1.5a. It is the simplest process one can imagine: two colliding protons remain intact, however, there is the four-momentum transfer in the t channel.

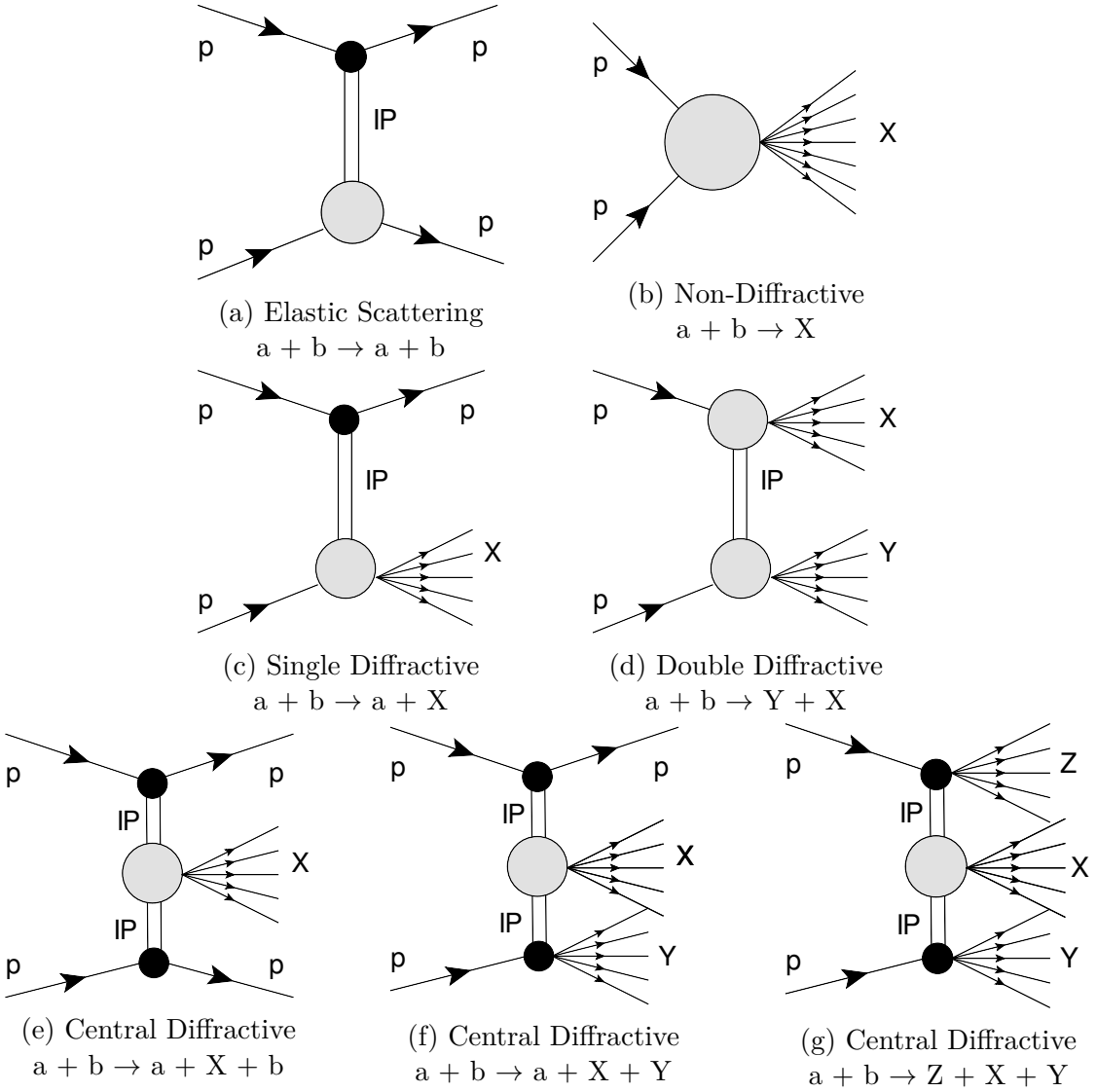
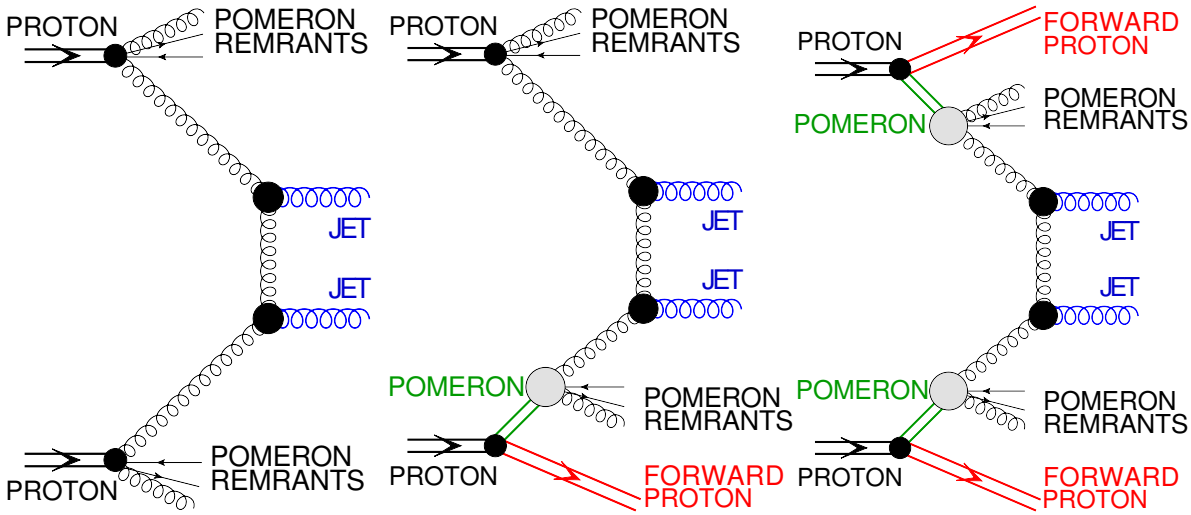


Figure 1.5: Various types of soft processes.

The inelastic processes can be separated into diffractive and non-diffractive ones. The latter class is depicted in Figure 1.5b where the final multi-particle state (marked by X) is shown. The soft diffractive processes are: Single Diffractive Dissociation (often called Single Diffraction, SD, Figure 1.5c), Double Diffractive Dissociation (Double Diffraction, DD, Figure 1.5d) and Central Diffraction (CD, Figures 1.5e,1.5f and 1.5g). In SD one, of

the two initial state protons, is dissociated and a multi-particle state with the quantum numbers of a proton (note that the spin of the state can be different) is produced. The other proton stays intact. In DD processes both protons dissociate and there is no quantum number exchange between them (similarly as in the case of one dissociated proton in the SD process). Finally, the CD processes, often called Double Pomeron Exchange, are described as the ones with both incoming particles radiating a colour singlet. In consequence, the centrally produced system, X, carries quantum numbers of the vacuum (modulo the spin).

The hard processes are usually defined as the ones in which a hard scale can be singled out. This allows for the use of perturbative calculation methods. Similarly to the soft case, the hard interactions can be divided into diffractive and non-diffractive (ND) processes. Hard diffraction $p+p \to p+X$, where system X contains two high transverse momentum jets was proposed for the first time in [16]. Such hard processes were first measured by the UA8 experiment at the SPS collider at CERN [17,18]. Hard diffraction includes Single Diffractive (SD) and Double Pomeron Exchange (DPE) processes. Examples of the Feynman diagrams related to hard SD, DPE and ND jet production are shown in Figure 1.6.



(a) Non-Diffractive (b) Single Diffractive (c) Double Pomeron Exchange

Figure 1.6: Diagrams of non-diffractive and diffractive jet production processes.

In the Non-Diffractive Jet production (Figure 1.6a), two jets are present in the final state and the remnants of protons are visible in the forward region. Unlike in the cases of SD and DPE jet production (Figure 1.6b and 1.6c), both protons are destroyed. In Single Diffraction one proton emits a Pomeron and, in consequence, one intact proton is present in the final state. In the DPE case, both protons emit Pomerons and survive an interaction. In addition to the presence of the forward protons, the diffractive processes are characterised by a large rapidity gap (a region with no particles produced) present between the Pomeron remnants and a proton.

Exclusive Jet Production (Figure 1.7) is a special case of diffractive production. Here, both interacting protons stay intact and two jets are present. There are no Pomeron nor proton remnants in the produced system.

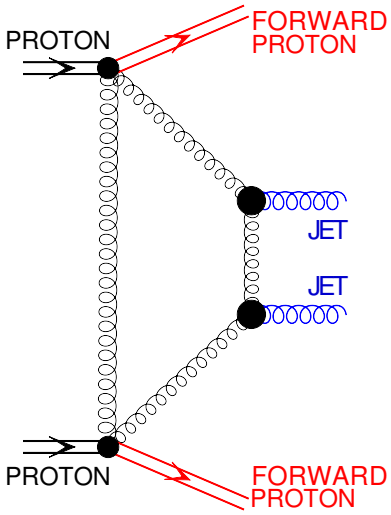


Figure 1.7: Feynman diagram of the Exclusive Jet Production.

The cross-section for a hard process can be calculated as a convolution of the Parton Density Function with the cross-section for the hard parton-parton interaction, σ_{hard}:

$$d\sigma = f_p(x_1, Q^2) f_p(x_2, Q^2) d\sigma_{hard}(x_1, x_2, Q^2),$$

where x_1 and x_2 are the proton momentum fractions carried by the interacting partons. In the above formula, Q is the scale of the process. The larger the Q^2 is the more closely one can probe the particle structure. Finally, the $f_p(x_1, Q^2)$ and $f_p(x_2, Q^2)$ are the corresponding PDFs.

To describe hard diffraction the Diffractive Parton Distribution Function (DPDF) is used. This function is similar to the PDF and takes the form:

$$f_d(x, Q^2, \xi, t) = \psi_P(\xi, t) f_P(x/\xi, Q^2).$$

In hard diffractive processes, the Pomeron is considered to be a colourless object with a partonic structure (gluons and quark) with the quantum number of the vacuum. The partons constituting a Pomeron participate in the hard scattering thus DPDF is factorised into the Pomeron flux $\psi_P(\xi, t)$ and the Pomeron partonic structure term $f_P(x/\xi, Q^2)$.

The cross-section for hard SD production can be calculated using the following formula:

$$d\sigma = S^2 \psi_P(\xi_1, t) f_P(x_1/\xi_1, Q^2) f_p(x_2, Q^2) d\sigma_{hard}(x_1, x_2, Q^2),$$

and for DPE processes:

$$d\sigma = S^2\psi_P(\xi_1, t)f_P(x_1/\xi_1, Q^2)\psi_P(\xi_2, t)f_P(x_2/\xi_2, Q^2)d\sigma_{hard}(x_1, x_2, Q^2).$$

To match the experimental data [19], it was essential to introduce an extra factor, the rapidity gap survival probability, S^2. This factor reflects the potential destruction of the rapidity gap which is possible due to additional soft initial or final state interactions. This effect was discovered by comparing diffractive deep inelastic scattering results from HERA (electron-proton) with Tevatron diffractive cross-sections (proton-antiproton). The measured diffractive cross-sections at HERA extrapolated to the Tevatron energies were larger than the measured ones by a factor close to 10. The gap survival probability depends on the centre of mass energy. For the LHC energies, it is expected to be of the order of 0.03 [20].

In this thesis, a measurement of Single Diffractive Jets based on data collected by the ATLAS experiment will be performed. In addition, an attempt to select Exclusive Jets with one proton tag will be made.

Chapter 2

Large Hadron Collider

Earth is bombarded with particles coming from outer space (cosmic rays) all the time, so it is possible to set up a detector and wait for particles to be registered. This solution is used in many experiments carried out either on the Earth's surface, underground, in the atmosphere or in orbit. Since the source of particles is not provided artificially there is no need for extra machines, to create and deliver them, which usually are costly and need manpower to operate. The drawback of using cosmic rays as a particle source is that it takes a lot of time to gather a significant amount of data at high energies and the detectors build on the ground or underground usually cover a large area or volume.

The particles can be delivered in a controlled way using particle accelerators. The simplest solution is the linear one accelerating particles in a straight line. The linear accelerators were and are used, successfully providing valuable data for many experiments. However, increasing the particle energy delivered by a linear accelerator would also eventually mean an increase in its length, which may be problematic due to available space or other factors. An alternative is to accelerate ensembles of particles in two separate circular machines till they reach desired energy and then collide them in the predefined Interaction Points (IP). During the collision in such a solution, many exchanges can take place and multiple elementary particles can be produced into the whole solid angle. Therefore, the Interaction Point is surrounded, as much as possible, by the particle detectors in order to register, ideally all, of the collision products.

This chapter describes the method of accelerating particles at the Large Hadron Collider (LHC [21]). In addition, possible LHC operation mods (optics) and their impact on proton trajectories are discussed.

2.1 CERN Accelerator Complex

The Large Hadron Collider is nowadays the most powerful human-build particle accelerator which delivers highly energetic proton or lead ion beams. The accelerator is located in a 27-

kilometre-long underground tunnel at CERN. The LHC is not the only machine hosted by CERN. The whole accelerator complex consists of a chain of accelerators (see Figure 2.1).

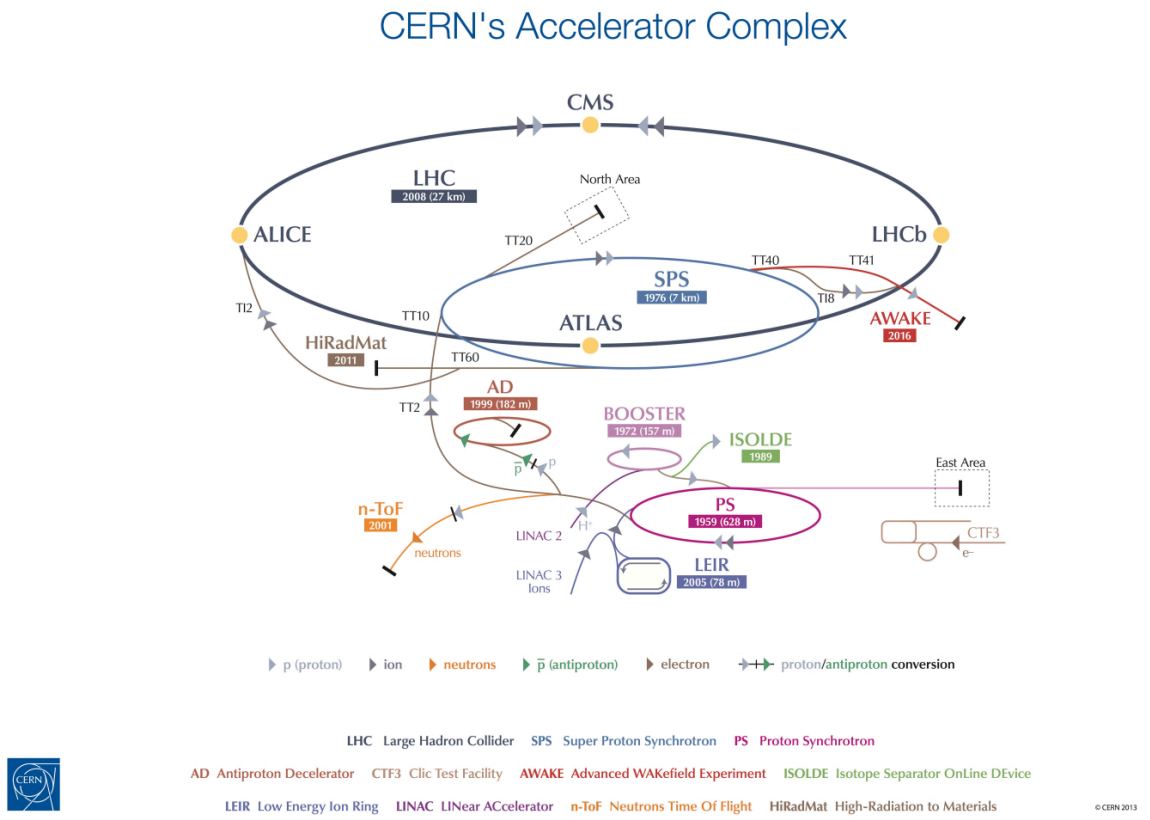


Figure 2.1: Scheme of CERN accelerators chain with experiments.

The acceleration of protons starts at their source – a hydrogen bottle. The atoms are ionised and the protons are extracted into the Linac2[1] [22] where they are accelerated to 50 MeV. From Linac2 the protons, with velocity reaching $\frac{1}{3}$ of the speed of light, are sent to the Proton Synchrotron Booster [22] (PSB). Here, the quality (brightness) of the beam is determined. The brightness corresponds directly to the beam density (related to its size and the number of protons it contains). The PSB accelerates the particles to the energy of 1.4 GeV.

In the next step, more than half of the protons is sent to Isotope Separator OnLine DEvice (ISOLDE) [23] – a nuclear physics research facility. The remaining protons travel to the Proton Synchrotron (PS) [24], which supplies three other experimental areas outside the LHC: the Antiproton Decelerator (AD) [25], the East Area and n-TOF [26]. In the PS

[1]Situation described in this thesis reflex the settings used in 2017 – a period when the analysed data were collected. Starting from Run 3 (2022) the acceleration path is a bit different, with Liniac4 used instead of Liniac2.

the protons are grouped into bunches and accelerated to the energy of 25 GeV.

PS sends batches of 72 bunches to the Super Proton Synchrotron (SPS) [27]. Usually, SPS will wait for four PS batches to form the so-called train, containing typically 288 bunches. Next, the trains are sent, clockwise or anti-clockwise, into the corresponding LHC beam pipes. At the injection into the LHC, the protons have an energy of 450 GeV. Eventually, the LHC accelerates the particles up to 6.5 TeV (years 2015–2018).

As it was mentioned, there are two beams accelerated in opposite directions in the separate beam pipes. The one moving clockwise (anti-clockwise) is commonly named beam1 (beam2). The beam pipes converge into a mutual one about 140 m away from the IP. The bending of the beam trajectory, and hence preserving an approximately circular orbit, is possible thanks to 1232 dipole magnets. To bend the trajectory of the 7 TeV protons, the LHC dipoles must be able to produce the field of 8.36 Tesla. In addition to beam bending, dipole magnets are also used for the separation of beams and their dumping.

If a beam of charged particles is left on its own, the particles belonging to a bunch will diverge from each other due to electromagnetic interactions. Therefore, the transverse dimensions of the beam must be well-controlled. This is achieved using quadrupole (and also higher-order) magnets. These magnets serve as "optical" devices and are used for focusing/defocusing the beams. Since it is impossible to construct a single quadrupole magnet focusing in both planes, the set of magnets focusing horizontally (defocusing vertically) and focusing vertically (defocusing horizontally) is used.

The beam is sensitive to any irregularities of the magnetic field and thus many elements are installed to monitor the beam quality and to correct for imperfections. For the field tuning, the Quadrupole Tuning Focusing/Defocusing, Skew Quadrupole magnets and Sextupole Magnets are used. To minimise the dispersion of the beam the quadrupole magnets called Dispersion Suppressors are employed.

During the machine's duty cycle, a beam is accompanied by a halo and stray beam particles. Also, there is debris emerging from the beam-beam collisions at the IP. Therefore, to protect the detectors, the superconducting magnets and the sensitive infrastructure of the machine various systems of collimators were devised.

2.2 LHC Operation

In order to collect a large amount of data, the LHC must operate as long as possible whilst maintaining good conditions for colliding particles. The rate of events, dN/dt, due to a specific process with the cross-section, σ, is proportional to the machine instantaneous luminosity, L:

$$\frac{dN}{dt} = \sigma \cdot L.$$

The instantaneous luminosity can be expressed in terms of the machine parameters and in a simplified form it is given by:

$$L \sim f \sum_i \frac{I_{1,i} \cdot I_{2,i}}{A_{eff}},$$

where f is the beam resolution frequency, the sum runs over the colliding bunches, $I_{1,i}$ and $I_{2,i}$ denote the currents of bunches belonging to beam1 and beam2, and A_{eff} is the overlap (effective) area of the colliding bunches in the transverse plane. The above formulas are given for the head-on collisions of the beams. The beam crossing angle, needed in the LHC case to avoid parasitic bunch-bunch collisions outside the IP region yet within the detectors, introduces effects which lead to the luminosity decrease [28].

The integrated luminosity is very often used to give the size of the gathered data sample:

$$N = \sigma \int L \cdot dt.$$

It is clear that the higher the integrated luminosity is, the rarer processes can be studied. Since at the LHC a lot of interest is in the precise studies of rare processes and the search for new physics phenomena, a lot of effort was put into increasing the instantaneous luminosity. In fact, there are a few methods to do that. These methods seriously impact the accelerator magnetic lattice as well as the infrastructure related to the beam creation. One of the methods uses the betatron function[2] at the Interaction Point (β^*) as the instantaneous luminosity is inversely proportional to the β^* value.

In order to keep the luminosity high, the procedure of a gradual change (levelling) of β^* was introduced. The levelling is based on changing the β^* value in sufficiently small steps such that the luminosity evolution will remain very smooth.

Another method hinge on the beam crossing angle (θ_C) levelling. It allows the control of the luminosity independently for each experiment, but the range of changes is limited by the long-range beam–beam separation (required to ensure good beam lifetimes) and by the LHC aperture. The luminosity levelling methods applied at the ATLAS IP during Run 2 data-taking (2015–2018) are listed in Table 2.1.

During the machine operation, the instantaneous luminosity decreases as a result of a loss of protons due to the collisions and imperfection of the machine. Therefore, the LHC is working in repeating cycles[3] with predefined phases. The standard cycle consists of:

- **Injection** – at first a low-intensity probe bunch is injected into the machine in order to adjust the beam parameters. Once good conditions are established, the nominal beam is injected in a form of trains of varying lengths.

[2]Approximately, the β function value gives the distance from a certain point to the point at which the beam transverse dimensions are doubled.

[3]Each such cycle is commonly called “a run” and has an assigned number.

Table 2.1: The luminosity levelling methods applied at the IP1 during 2017 and 2018.

Period	β^* [cm]	β^* levelling	$2 \cdot \theta_C$ [μrad]	$2\theta_C$ levelling
06.06 – 26.07.2017	40	no	280→260→230→210	in steps
29.07 – 13.09.2017	40	no	270→250→230→215	in steps
22.09 – 26.11.2017	30	no	270→230→215→195	in steps
17.04 – 23.11.2018	30→27→25	in steps	300→240	continuous

- **Ramp** – the beam energy is increased from 450 GeV to the desired energy.
- **Squeeze** – the betatron function at the interaction points is adjusted to its target value for collisions.
- **Adjust** – during this phase, the beams are brought into collision (up till now the beams were kept separated at the IPs), and the final optimisations of the beam parameters are performed.
- **Stable Beams** – the beams are maintained in the stable collision regime for the experiments. Only minor adjustments are made to the beam parameters during this phase.
- **Beam Dump** – circulating beam is dumped in a controlled way into the absorber block.
- **Ramp Down** – currents of the magnets are brought down to their injection values.

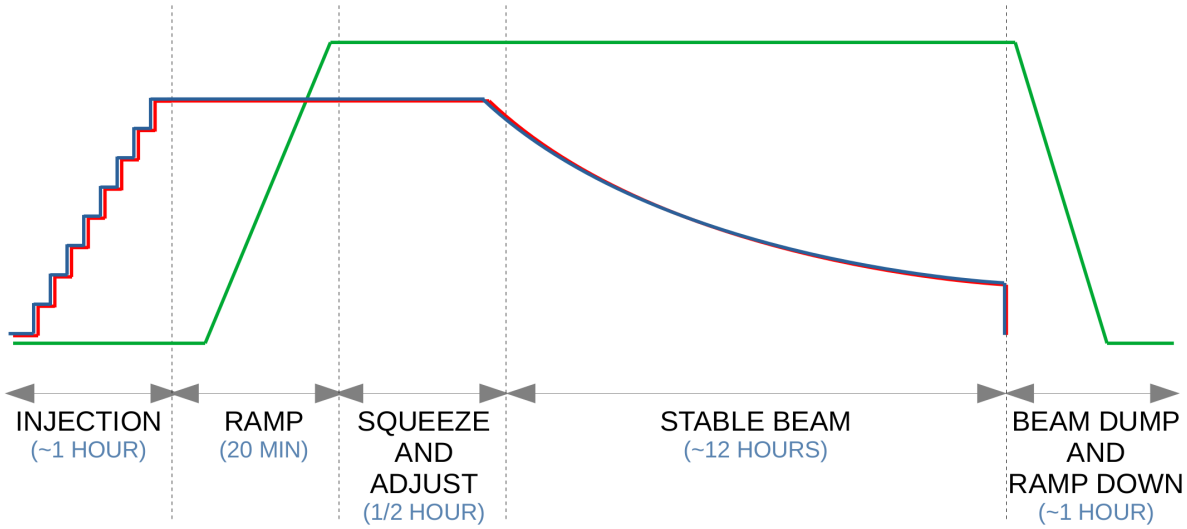


Figure 2.2: Scheme of a standard LHC cycle in Run 2. Blue and red lines represent Beam 1 (clockwise) and Beam 2 (anti-clockwise) intensity (current), whereas the green line is the magnetic circle of the LHC.

As shown in Figure 2.2 a standard cycle at the LHC lasts over a dozen hours with data collection during the stable beam phase. It is preferable to start the data collection at the highest luminosity possible as it will drop with time.

2.3 LHC Optics and Proton Trajectories

Diffractive protons are scattered at very small angles into the LHC beam pipe. This means that the forward protons detectors must be placed far away (a few hundred meters) from the Interaction Point to register particles separated from the beam. In the case of the ATLAS Forward Proton (AFP) detectors used for analysis detailed in this thesis, the distance is about 210 m.

There are several LHC elements between the IP and the detectors influencing the proton trajectory: two dipole magnets (D1, D2) for the separation of the beams (bending), five quadrupole magnets (Q1-Q5) for the beam focusing and two collimators (TCL4, TCL5) for the magnet protection. In consequence, the proton trajectory between the interaction point and the forward detectors is not a straight line.

The settings of the machine elements are called optics. Their details follow from the compromise between the LHC machine protection and the requests of the LHC experiments concerning the physics programmes in terms of luminosity. Figure 2.3 shows a typical situation for the high-luminosity ATLAS data taking. The horizontal axis gives the distance from the ATLAS IP whereas the vertical one a position in the x-direction within the horizontal, accelerator plane. The view is from the above, the trajectories are marked with black lines and the LHC elements (dipole and quadrupole magnets) are marked as black rectangles. The blue lines show the collimator positions. The red lines symbolise the AFP [29] and ALFA stations [30].

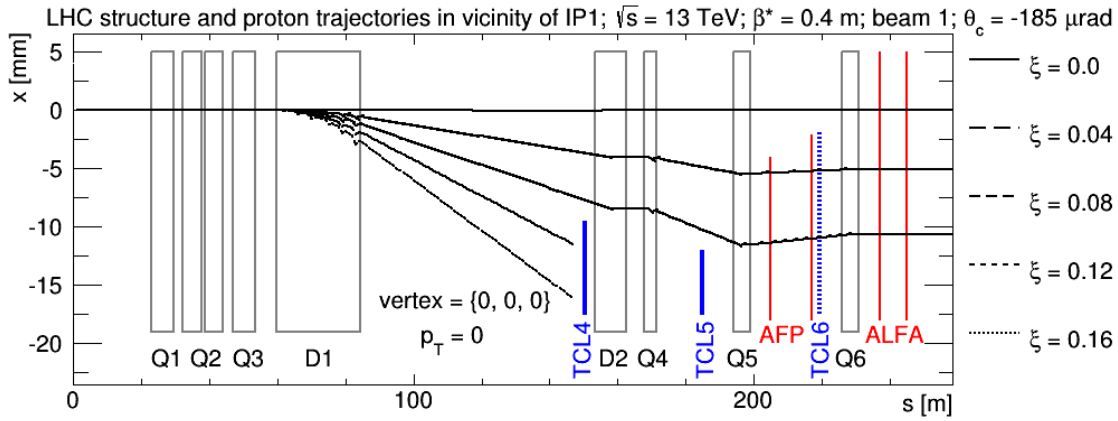


Figure 2.3: Proton trajectories between the ATLAS interacting point and AFP detectors for different values of the proton relative energy loss, ξ (see text). From [31].

The proton trajectories reflect the proton momentum. The main impact is due to the relative energy loss of a proton:

$$\xi = 1 - \frac{E_{proton}}{E_{beam}}.$$

The proton trajectories are shown in Figure 2.3 for various values of ξ assuming the

scattered proton transverse momentum $p_T = 0$ at the IP. With increasing ξ, the proton trajectories start to diverge from the beam and for $0.025 < \xi < 0.1$, corresponding to the proton energy of 160 GeV $< E_{proton} <$ 650 GeV, the protons can reach the AFP detectors. For larger ξ values, the protons will be absorbed in the machine infrastructure, typically in the collimators. One should note that there are also other parameters influencing the actual shape of the forward scattered proton trajectory. For example the proton emission angles or its transverse momentum. In the case of diffractive analysis, the settings of the LHC elements not only impact the luminosity but also define the detector acceptance. More detail about the AFP detector structure and performance will be described in Chapter 4.

Chapter 3

ATLAS Detector

There are seven experiments installed at the LHC: A Large Ion Collider Experiment (ALICE) [32], A Toroidal LHC ApparatuS (ATLAS) [33], the Compact Muon Solenoid (CMS) [34], the Large Hadron Collider beauty (LHCb) [35], the TOTal Elastic and diffractive cross section Measurement (TOTEM) [36], the Large Hadron Collider forward (LHCf) [37], and Monopole and Exotics Detector at the LHC (MoEDAL) [38]. The biggest ones: ALICE, ATLAS, CMS and LHCb are installed at the four collision points of the LHC beams. TOTEM is located close to the CMS interaction point, LHCf is installed near ATLAS and MoEDAL is close to the LHCb detector. The present chapter discusses the structure of the ATLAS components and the algorithms used to reconstruct such detector objects as clusters, tracks and jets, which later will be used in the analysis.

The ATLAS detector is currently the largest experiment at the LHC and consists of several sub-detectors of various types (see Figure 3.1, left). The sub-detectors create an "onion-like" structure that allows the reconstruction and investigation of various particle properties since each one is intended for different measurements. The Inner Detector (ID) is the closest to the IP. It is surrounded by Calorimeters. The outermost layers are the Muon Detectors with Toroid Magnets. Additionally, two modules of the ATLAS Luminosity Detector (LUCID) are placed symmetrically with respect to the interaction point (17 m from it) around the beam pipe.

The ATLAS detector uses a right-handed coordinate system with the origin defined by the nominal interaction point, with the z-axis along the beam line pointing towards the ATLAS side C (Figure 3.1, right). The (x, y) plane (the transverse plane) is perpendicular to the beam line. The x-axis points towards the LHC ring centre.

A common way to describe the detected particles is to use the (η-ϕ) space. The azimuthal angle ϕ (marked in red in the right part of Figure 3.1) is measured with respect to the x-axis. The pseudorapidity is defined as

$$\eta = -\ln \tan\left(\frac{\theta}{2}\right),$$

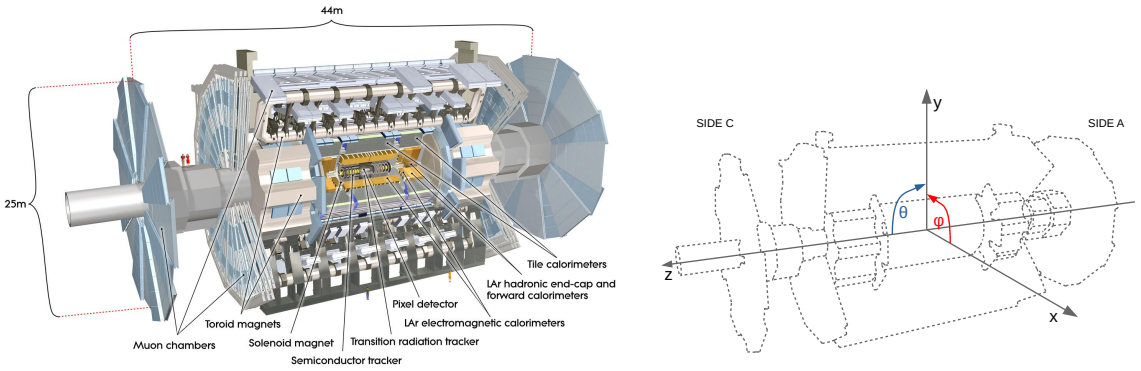


Figure 3.1: ATLAS detector with its sub-detectors (**left**) and the scheme of its coordinate system (**right**).

where θ is the particle polar angle.

3.1 Inner Detector

The Inner Detector (ID) [39] consists of several sub-detectors, see Figure 3.2. ID provides measurements of the charged particle track position within the range $|\eta| < 2.5$. It consists of a cylindrical barrel region (full coverage for $|\eta| \leqslant 1.5$) arranged around the beam pipe and two end-caps. Disks in the end-cap regions are placed perpendicularly to the beam axis and cover $1.5 < |\eta| < 2.5$.

ID consist of three sub-detectors: Pixel Detector [40] (with Insertable B-Layer (IBL) [41]), Silicon miCrostrip Tracker (SCT) [42] and Transition Radiation Tracker (TRT) [43]. These detectors are submerged into a solenoid magnetic 2 T field. The role of this field is to bend the tracks of the particles in order to measure their momentum and determine the charge.

3.1.1 Pixel Detector

The high-granularity silicon pixel detector (Pixels) [40] with the innermost layer, the Insertable B-layer (IBL) [41], covers the vertex region. It is a detector closest to the IP.

The IBL has a mean radius of 33 mm with a pixel size of 50 μm by 250 μm in the x and z directions and a sensor is 200 μm thick. The three layers of the pixel system are located at the mean radii of 50.5, 88.5, and 122.5 mm respectively. Typical pixel size is 50 μm by 400 μm with a thickness of 250 μm. The coverage in the end-cap regions is enhanced by three disks. The pixel detectors measure the charge collected in each individual pixel using the Time over Threshold (ToT)[1].

[1]ToT is the time the pulse exceeds a given threshold and is proportional to the deposited energy.

Table 3.1: Parameters of the Inner Detector. The quoted resolutions show the typical values (the actual resolution in each detector depends on the impact angle). From [44]

System	**Position**	**Area** [m^2]	**Resolution** σ [μm]	**#Channels** [10^6]	**$\lvert\eta\rvert$ coverage**
Pixels	1 removable barrel layer (B-layer)	0.2	Rϕ=12, z=66	16	2.5
	2 barrel layers	1.4	Rϕ=12, z=66	81	1.7
	5 end-cap disks on each side	0.7	Rϕ=12, R=77	43	1.7-2.5
SCT	4 barrel layers	34.4	Rϕ=16, z=580	3.2	$\pm$1.4
	9 end-cap wheels on each side	26.7	Rϕ=16, R=580	3.0	1.4-2.5
TRT	Axial barrel straws		170 (per straw)	0.1	0.7
	Radial end-cap straws 36 straws per track		170 (per straw)	0.32	0.7-2.5

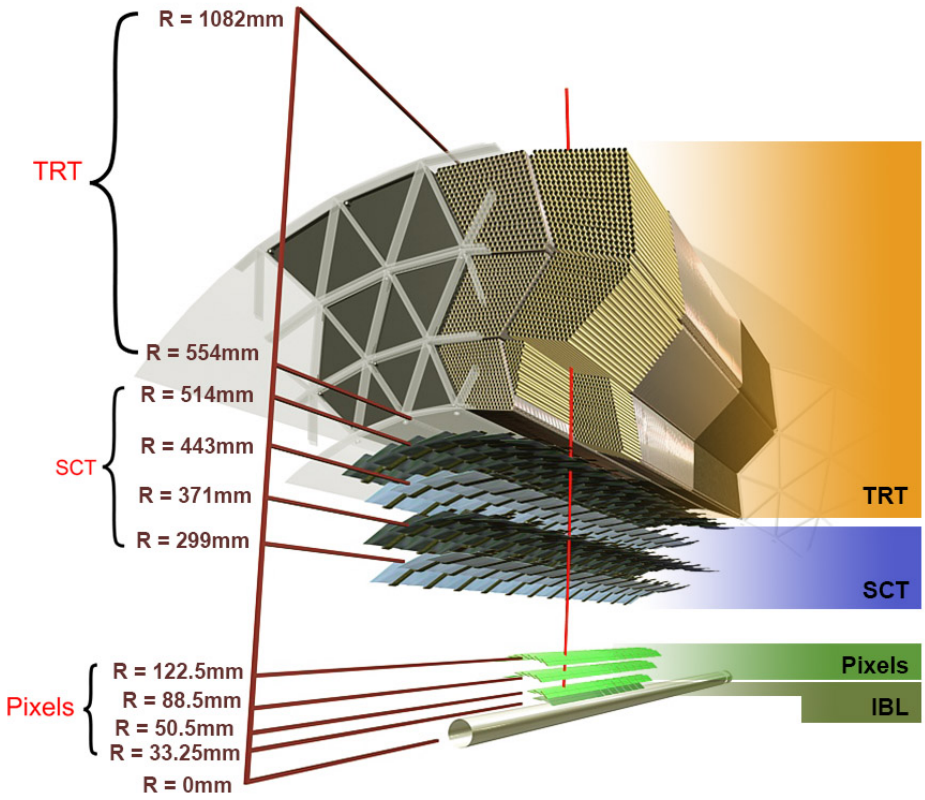


Figure 3.2: Inner ATLAS detector with its sub-detectors.

3.1.2 Silicon Microstrip Tracker

Outside the pixel detectors, a barrel of the Silicon Microstrip detectors (SCT) [42] is placed. SCT consists of four silicon strip double layers located at radii between 299 mm and 514 mm. The barrel is complemented by nine disks in each of the end-caps. A typical strip of the SCT barrel sensor has a length of 126 mm and a pitch of 80 μm. For each layer, the strips are parallel to the beam direction on one side and at the stereo angle of 40 mrad on the other. Information from both sides of a layer can be combined to provide an average of four three-dimensional measurements per track. The SCT sensors are connected to binary read-out chips, which do not provide information concerning the collected charge.

3.1.3 Transition Radiation Tracker

The silicon detectors are complemented by the Transition Radiation Tracker [43] (TRT). The TRT consists of drift tubes made of Kapton reinforced with thin carbon fibres. The tubes are 4 mm in diameter and host the gold-plated tungsten wire. The TRT is a gaseous detector and ideally should be filled with a gas mixture of 70% Xe, 27% CO2, and 3% O2.

The tracker consists of a barrel and two end-caps. The barrel region straw tubes are 1.5 m long and are placed parallel to the beam axis. They cover the radii from 0.5 m to 1.1 m and the range in pseudorapidity of $|\eta| <1$. The end-cap straws are placed perpendicularly to the beam axis (are radially arranged) and are 0.4 m long. End-cap region covers the geometrical range of 0.8 m$< |z| <$2.7 m and $1< |\eta| <2$.

Charged particle passing thru the TRT ionises the gas inside the straws. This results in free electrons drifting towards the wire, where they are amplified and the signal is read out.

In the space between the straws, the polymer fibres (barrel) and foils (end-caps) are placed to create the transition radiation. Such radiation is emitted by highly relativistic charged particles crossing the boundary of the dielectric media. This effect depends on the relativistic factor $\gamma = E/m$ and for a fixed value of E is the strongest for electrons. This means it can be exploited for particle identification.

TRT design makes it complementary to the silicon-based tracking devices: the intrinsic single-point resolution of 120 μm is larger than that of the silicon trackers, however, due to the large number of hits per track (typically more than 30) and the long lever arm this number is compensated.

3.2 Calorimeters

Calorimeters are crucial parts of the ATLAS detector. They provide precise measurements of the energies and positions of particles passing thru them. One can distinguish the Liquid-Argon (LAr) [45] and the Tile Hadronic (Tile) [46] calorimeters (all parts of the calorimeter

system are shown in Figure 3.3). The details about various parts of the calorimeter system coverage are listed in Table 3.2.

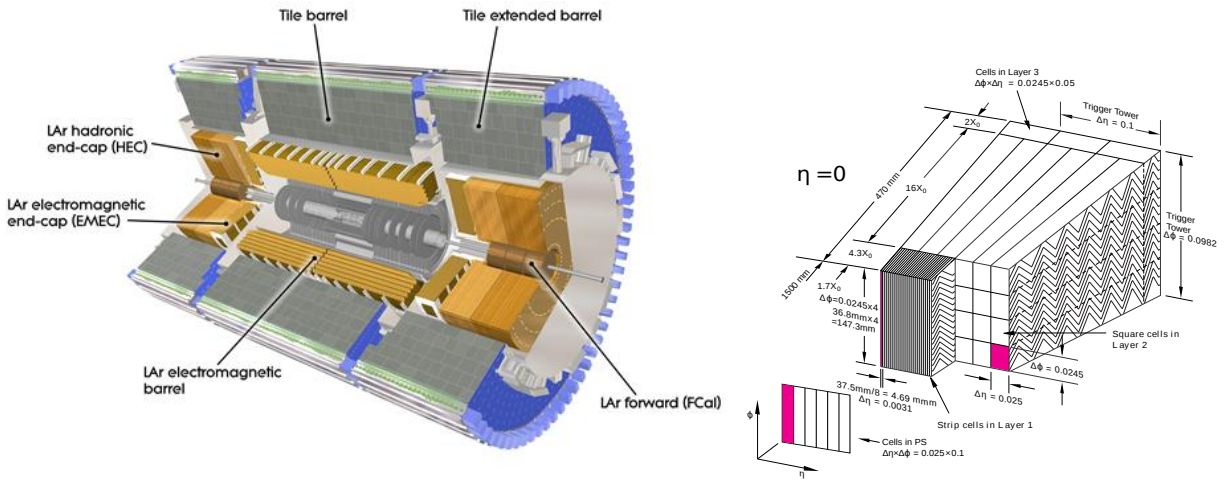


Figure 3.3: **Left:** Calorimeter system in ATLAS detector. **Right:** structure of EM ATLAS calorimeters, a cell is marked in magenta.

3.2.1 LAr Calorimeters

The Liquid Argon (LAr) [45] sampling calorimeter surrounds the ATLAS Inner Detector. It features layers of metal (either tungsten, copper or lead) that absorbs the incoming particles, "converting" them into showers of lower energy particles. The particles ionise liquid argon placed between the absorber layers, producing an electric signal that is measured. By combining these signals the energy of the incident particle can be determined.

Electromagnetic Calorimeters

The Electromagnetic calorimeter has a characteristic accordion structure, with a honeycomb pattern, to ensure that no particle escapes registration. Such an accordion geometry provides a complete ϕ symmetry without azimuthal cracks. It is divided into the barrel ($|\eta| < 1.475$) and two end-cap components ($1.375 < |\eta| < 3.2$), each housed in its own cryostat. The energy resolution of the barrel and end-caps is $\Delta E/E = 10\%/\sqrt{E(GeV)} \oplus 0.7\%$.

LAr electromagnetic barrel, designed to identify electrons and photons, consists of two identical half-barrels, separated by a small gap of 6 mm at z = 0. The end-caps are mechanically divided into two coaxial wheels: an outer covering $1.375 < |\eta| < 2.5$, and the inner one covering $2.5 < |\eta| < 3.2$.

In the region $|\eta| < 2.5$, the EM calorimeter is segmented into three longitudinal sections (samplings):

Table 3.2: Pseudorapidity coverage, granularity and longitudinal segmentation of the ATLAS calorimeters.

EM CALORIMETER	**Barrel**	**End-cap**	
Coverage	$\|\eta\| < 1.475$	$1.375 < \|\eta\| < 3.2$	
Longitudinal segmentation	3 samplings	3 samplings	$1.5 < \|\eta\| < 2.5$
		2 samplings	$1.375 < \|\eta\| < 1.5$
			$2.5 < \|\eta\| < 3.2$
Granularity ($\eta \times \phi$)			
Sampling 1	0.003×0.1	0.025×0.1	$1.375 < \|\eta\| < 1.5$
		0.003×0.1	$1.5 < \|\eta\| < 1.8$
		0.004×0.1	$1.8 < \|\eta\| < 2.0$
		0.006×0.1	$2.0 < \|\eta\| < 1.5$
		0.1×0.1	$2.5 < \|\eta\| < 3.2$
Sampling 2	0.025×0.025	0.025×0.025	$1.375 < \|\eta\| < 2.5$
		0.1×0.1	$2.5 < \|\eta\| < 3.2$
Sampling 3	0.05×0.025	0.05×0.025	$1.5 < \|\eta\| < 2.5$
PRESAMPLER	**Barrel**	**End-cap**	
Coverage	$\|\eta\| < 1.52$	$1.5 < \|\eta\| < 1.8$	
Longitudinal segmentation	1 sampling	1 sampling	
Granularity ($\eta \times \phi$)	0.025×0.1	0.025×0.1	
HADRONIC TILE	**Barrel**	**Extended barrel**	
Coverage	$\|\eta\| < 1.0$	$0.8 < \|\eta\| < 1.7$	
Longitudinal segmentation	3 samplings	3 samplings	
Granularity ($\eta \times \phi$)			
Samplings 1 and 2	0.1×0.1	0.1×0.1	
Sampling 3	0.2×0.2	0.1×0.1	
HADRONIC LAr		**End-cap**	
Coverage		$1.5 < \|\eta\| < 3.2$	
Longitudinal segmentation		4 samplings	
Granularity ($\eta \times \phi$)		0.1×0.1	$1.5 < \|\eta\| < 2.5$
		0.2×0.2	$2.5 < \|\eta\| < 3.2$
FORWARD CALORIMETER		**Forward**	
Coverage		$3.1 < \|\eta\| < 4.9$	
Longitudinal segmentation		3 samplings	
Granularity ($\eta \times \phi$)		$\sim 0.2 \times 0.2$	

- strip section (sampling 1 in Figure 3.3 right): equipped with narrow strips. This section acts as a "pre-shower" detector, enhancing the particle identification and providing a precise position measurement in η,
- middle section (sampling 2): transversely segmented into the square towers,
- back section (sampling 3): with a granularity of 0.05 in η.

For $|\eta| > 2.5$, *i.e.* for the end-cap inner wheel, the calorimeter is segmented into two longitudinal sections and has more crude granularity than for the rest of the acceptance. This is sufficient to satisfy the physics requirements (reconstruction of jets and measurement of the missing energy).

Additionally, a presampler detector, covering the range $|\eta| < 1.8$, is used to correct for the energy lost by electrons upstream of the calorimeter. The presampler consists of an active LAr layer of the barrel (end-cap) region.

Hadron Calorimeters

There are LAr Hadronic end-cap (HEC) covering $1.5 < |\eta| < 3.2$ and Forward (FCal) calorimeters ($3.1 < |\eta| < 4.9$), see Figure 3.3. Their energy resolution is respectively $\Delta E/E = 50\%/\sqrt{E(GeV)} \oplus 3\%$ and $\Delta E/E = 100\%/\sqrt{E(GeV)} \oplus 10\%$.

The Hadronic End-cap Calorimeter (HEC) consists of two independent wheels per end-cap, located directly behind the end-cap electromagnetic calorimeter. Each wheel is divided into two segments in depth, a total of four layers (samplings) per end-cap.

The Forward Calorimeter (FCal) is integrated into the end-cap cryostats, as this provides benefits in relation to uniformity of the calorimetric coverage as well as the reduced radiation background levels in the muon spectrometer. The FCal consists of three modules (samplings) in each end-cap: the first uses a copper absorber and is optimised for electromagnetic measurements, while the other two, made of tungsten, measure predominantly the energy of passing hadrons.

3.2.2 Tile Calorimeter

The Tile Calorimeter surrounds the LAr calorimeters and measures the energy of hadrons, which do not deposit all of their energy in the LAr Calorimeter. It is a sandwich-like calorimeter made of alternating layers of steel and plastic scintillating tiles. The energy deposited in the scintillators is converted into a charge whose value is proportional to the original energy of the passing particle.

A large hadronic barrel calorimeter is a sampling calorimeter using iron as the absorber and scintillator tiles as the active material. The tiles are placed radially and staggered in depth.

The tile calorimeter is composed of one barrel and two extended barrels. It is segmented in depth into three layers (samplings). The single hadron energy resolution obtained in the

beam test is described by $\Delta E/E = 52\%/\sqrt{E(GeV)} \oplus 5.7\%$ for $\eta = 0.2$ while the jet energy resolution is $\Delta E/E = 50\%/\sqrt{E(GeV)} \oplus 3\%$.

3.3 Muon Spectrometer

Muons usually pass through the Inner Detector and Calorimeter nearly unobstructed. Thus, the need for detectors dedicated to their identification: Muon Spectrometer [47]. The scheme of the detector is visible in Figure 3.4. This detector uses the air-core toroid magnets that are responsible for the bending of the muon tracks.

The four main components of the Muon Spectrometer are the Monitored Drift Tubes (MDT), the Cathode Strip Chambers (CSC), the Resistive Plate Chambers (RPC) and the Thin Gap Chambers (TGC). The measurements of the track curvature are sufficed by MDT detectors within the range $|\eta| < 2.7$. The MDT tube resolution is 80 μm. The CSC provide measurements precision coordinates at the ends of the detector ($2.0 < |\eta| < 2.7$) with the resolution of 60 μm. The TGC is used for triggering and second coordinate measurement within the range of $1.05 < |\eta| < 2.4$. A similar purpose has the RPC detector but in the central region $|\eta| < 1.05$.

The whole Muon Spectrometer detector covers the range of $|\eta|$ <2.7 and is able to detect muons within the transverse momentum within 3 GeV $< p_T <$ 1 TeV.

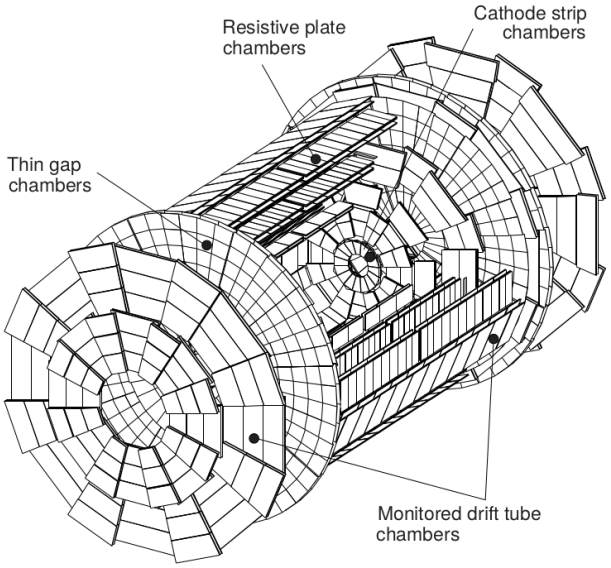


Figure 3.4: ATLAS muon spectrometer.

3.4 The ATLAS Luminosity Detector

The LUCID detector [48] is the main luminosity provider of the ATLAS experiment. It is able to deliver a reliable luminosity determination in all configurations of the beams, luminosity ranges and bunch-by-bunch measurements. The final aim of LUCID is to provide the measurement of the integrated luminosity independently of the physics data. This information is important during the data taking, where it is used to tune the trigger prescales. In offline analysis, integrated luminosity is necessary for every measurement of the cross-section.

The LUCID detector is located on both sides of the interaction point at a distance of 17 m. Each LUCID vessel contains 20 Cherenkov detectors (1.5 m long) surrounding the beam pipe. These detectors consist of aluminium tubes with 15 mm diameter pointing towards the interaction region. The gas in the detector is C_4F_{10} at the pressure of 1.1-1.5 bar. Cherenkov light is collected in 16 of the tubes by 15 mm diameter photomultipliers, which have quartz windows to increase their radiation hardness. In four of the tubes, the light is collected by a cone and then transmitted via the optical quartz fibres to the outside of the experiment Forward Shielding. These signals are used for various measurements and tests. The application of quartz fibres follows from their radiation hardness.

3.5 Reconstruction in ATLAS

The diffractive jet analysis is based on the ATLAS reconstructed objects such as clusters, tracks and jets. Unfortunately, neither the ATLAS apparatus nor the reconstruction algorithms are perfect. This results in false objects reconstructed in the events (*e.g.* due to the electronics noise) or the objects reconstructed incorrectly. It is crucial to understand each step of the reconstruction process to provide the best selection to be applied to all stages of the data analysis. In this section, the reconstruction algorithms of the clusters, tracks and jets will be briefly discussed and the methods of calculating their most important properties will be described.

3.5.1 Clusters

The reconstructed clusters represent a part of the final information one can obtain from the calorimeters. One of the most basic objects reconstructed in calorimeters are the cells (an example of a cell is marked in magenta in Figure 3.3, right). It should be mentioned that the standard ATLAS Object Data (AOD; the ATLAS data type used in this analysis) contains only information considering the clusters of the cells.

The topo-clustering algorithm [49] defines the clusters used further in an analysis. In the first step, the algorithm chooses the seed cell(s). For each cell, its signal significance (ς_{cell}^{EM}) is defined as the ratio of the cell signal (E_{cell}^{EM}) to the average (expected, estimated

for each data taking period) noise ($\sigma^{EM}_{noise,cell}$):

$$\varsigma^{EM}_{cell} = \frac{E^{EM}_{cell}}{\sigma^{EM}_{noise,cell}}.$$

Next, the cells are ordered in descending order according to their signal. In the algorithm, the seeding, growth, and boundary features of the topo-clusters are controlled by the three parameters S, N, P:

- $|\varsigma^{EM}_{cell}| > S$ – primary seed threshold, default S = 4,
- $|\varsigma^{EM}_{cell}| > N$ – threshold for growth control, default N = 2,
- $|\varsigma^{EM}_{cell}| > P$ – principal cell filter, default P = 0.

The algorithm selects all neighbouring cells with the signal significance passing the threshold defined by S. If the neighbouring cell is also a seed, the two (proto-)clusters are merged. If the neighbouring cell is attached to two different proto-clusters and passes the threshold these two proto-clusters are also merged.

This procedure is iteratively applied to further neighbours until the last set of neighbouring cells is inspected. The final proto-cluster is characterised by the core of cells with high significant signal surrounded by cells with less significant signals.

In the ATLAS calorimeters, the non-signal clusters can be present in the event due to the pile-up or, to a lesser extent, due to the electronic noise. Not only positive cell signals can seed a cluster, but also those with large negative signals.

Topo-clusters from the out-of-time pile-up seeded by $E^{EM}_{cell} > 0$ often yield $E^{EM}_{clus} > 0$, because they are typically generated by particles produced in past bunch crossings closer in time (within 100 ns). The topo-clusters with $E^{EM}_{clus} < 0$ can be used to provide an average global cancellation of contributions of clusters seeded by positive fluctuations in the out-of-time pile-up in a full event observable (an observable calculated for an event) including missing transverse energy, E^{miss}_T.

To avoid biases in the jet-finding and to support a detailed jet substructure analysis as well as a high-quality E^{miss}_T reconstruction, the proto-clusters with two or more local maxima are split.

The transition regions at $|\eta|$ ∼0.8–1.0, $|\eta|$ ∼1.4 and $|\eta|$ ∼3.2 affect the topo-cluster formation more than in the case of electromagnetic showers. Not only in terms of the maximum multiplicity of the clusters, N_{clus}, but also in terms of the range in η. In particular, the region around $|\eta|$ ∼0.8–1.0 has a larger effect on N_{clus} for hadrons than for electromagnetically interacting particles, as this is the transition from the central to the extended Tile calorimeter introducing a reduced calorimetric coverage for hadrons. The central electromagnetic calorimeter provides hermetic coverage here, without any effect on N_{clus}. A sharp drop of N_{clus} for π^- at $|\eta|$ =2.5 corresponds to the reduction of the calorimeter cell granularity by a factor of approximately four.

3.5.2 Tracks

The Inner Detector signals are used for the reconstruction of tracks within $|\eta|$ <2.4. Similarly to the clusters, the tracks can be used to reconstruct the jets and in the rapidity gap studies.

The first step in reconstructing the charged particle tracks in ATLAS is finding the track seeds. They are formed from sets of three space points. The purity of seeds that result in good-quality tracks, varies depending on which sub-detector(s) recorded the space points. The purest seed types are the SCT-only, the pixel-only and finally the mixed-detector seeds.

After choosing the seeds, the track candidates are built by including additional space points from the remaining layers of the pixel and SCT detectors. Multiple track candidates can be produced per single seed.

The selection in this algorithm starts with calculating the track candidate's score, which depends on:

- clusters assigned to a track (increasing the track score) and holes[2] (reducing the score),
- the χ^2 of the track fit (decreasing score for the track candidates with a poor fit),
- the logarithm of the track momentum is considered to promote energetic tracks and suppress the larger number of tracks with incorrectly assigned clusters (typically with low p_T).

All track candidates are processed individually, starting from the track with the highest score. After the track scores have been calculated and the track candidates have been sorted, the neural network is used to predict the cluster positions. In a dense environment, it is natural to have clusters compatible with multiple track candidates. Clusters can be shared by no more than two tracks, giving preference to a track with a higher score. Additionally, a track can maximally have two shared clusters. A cluster is removed from a track candidate if it causes either the candidate or the accepted track to not meet the shared-cluster criteria. The track candidate is then scored again and returned to the ordered list of the remaining candidates. In addition to standard selection, additional cuts were applied in order to ensure the good quality [50]. The following cuts were implemented:

- track $p_T > 200$ MeV,
- track $|\eta| < 2.5$,
- minimum 1 hit from the Pixel layer, including the B-layer,
- if track $p_T > 200$ MeV, then 4 hits in SCT layer were required,
- if track $p_T > 300$ MeV, then 6 hits in SCT layer were required,
- $|d_0| \leqslant 1.5$ mm,
- $|z_0 - z_{beam}| \sin(\theta_0) \leqslant 1.5$ mm, where z_{beam} is reconstructed z position of primary vertex.

[2]Holes are defined as the intersections of the reconstructed track trajectory with a sensitive detector element that does not contain a matching cluster. Inactive sensors or regions, such as the edge areas on the silicon sensors, are excluded from the hole definition.

The definitions of d_0, z_0 and θ are pictured in Figure 3.5. The d_0 parameter is the transverse impact parameter calculated with respect to the measured beam-line position, z_0 is the longitudinal (along the beam line) difference between the point where d_0 is measured and the primary vertex (point (0, 0) in Figure 3.5), and θ is the polar angle of a track.

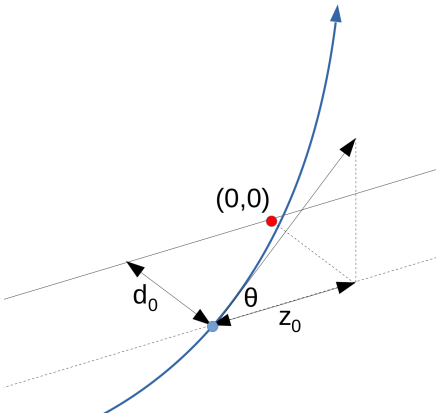


Figure 3.5: Graphical representation of track d_0, z_0 and θ variables.

3.5.3 Jets

Reconstruction of jets is not an easy task, since they are very complex objects that are impossible to be defined in a simple and faultless way. In order to reconstruct jets, sophisticated algorithms are used. The jet algorithm defines how the particles are clustered into jets. It can be applied to the calorimeter objects, tracks, combinations of the two, and the truth particles used in the Monte Carlo data analysis. In ATLAS, the jets are usually reconstructed by the anti-kt algorithm [51].

The present analysis is based on the "AntiKt4EMPFlowJets" objects. The name of these objects is composed accordingly to the ATLAS rules (see [52]) and tells the specification of a reconstructed jet. The "AntiKt4" part of the name tells that the jets are reconstructed with the anti-kt algorithm [51] with the jet radius R=0.4. These jets were reconstructed using the topo-clusters as input (see Sec. 3.5.1), in particular the clusters at the electromagnetic scale (EMTopo, "EM" in the name). The clusters can be also at the local calibration scale (LCTopo). For large-R jets, most analyses performed currently use the LCTopo jets. This provides a good performance across all p_T. For jets with small R (R=0.4) and low p_T, the particle EMTopo jets are used. The last part of the name implies which algorithm (PFlow [53]) was used for the reconstruction.

The PFlow algorithm is the cell-based energy subtraction algorithm which is designed to remove overlaps of the momentum and energy measurements from the inner detector and calorimeters. The algorithm combines information from both the tracker and the calorimeter

with the purpose of reconstruction of the hadronic jets and soft activity in the event[3].

The algorithm starts by selecting well-reconstructed (for more information about this criteria see [53]) tracks. Next, for each track, single topo-clusters are matched to them. The expected energy in the calorimeter is calculated based on the topo-cluster position and the track momentum. Usually, a single particle deposits energy in multiple topo-clusters. The probability of such an occurrence is calculated. Based on that it is decided if additional clusters must be added to the track and topo-cluster system.

The calculated expected energy in the calorimeter is subtracted cell by cell from the set of matched topo-clusters. This procedure lasts until the remaining energy in the system is consistent with the expected shower fluctuations of a single particle's signal. The topo-cluster remnants are then removed.

The described procedure is enforced starting from the tracks with the highest p_T with only a single topo-cluster matched.

Jets reconstructed in this way are saved in the data, however in order to use them in the analysis couple of corrections must be applied. First, the calibration must be applied. The calibrations are delivered by specialised ATLAS groups and are related to the named data-taking periods. In the analysis described in this work, the calibration for 2017 low-μ runs was used. In order to ensure the good quality of jets following cuts were used:

- $|\eta_{jet}| < 1.5$ – guarantee that the whole jet will be reconstructed within the acceptance of the ATLAS detector,
- $p_T^{jet} > 20$ GeV – recommendations of ATLAS for PFlow jets,
- $|timing_{jet}| < 12.5$ ns – exclusion of the so-called "out-of-time" jets,
- if $|\eta_{jet}| < 2.4$ then JVT>0.5 – Jet Vertex Tagger (JVT) is the variable telling how probable is that a jet comes from the primary vertex.

3.6 Trigger System in ATLAS

Analysing a large amount of data produced at the LHC is very challenging. During LHC Run 2 (2015–2018), the machine worked at the maximum crossing rate of 40 MHz. The size of the raw data event (around 1 MB) implies that a huge amount would need to be stored to preserve all information. However, since physicists are interested in the specific processes that occur at a relatively low rate, many events can and should be rejected. For this purpose, dedicated electronics and software algorithms, performing the triggering functions were created. The accumulation of such algorithms is commonly called the Trigger system [54].

For technical reasons during Run 2, the trigger system had to reduce the rate of events selected for permanent storage to $\sim$4 GB/s. This is done in two steps called Level 1 (L1)

[3]Additional hadronic recoil below the threshold used in the jet reconstruction. It is especially important for the calculation of the missing transverse momentum in an event.

and High-Level Trigger (HLT). Historically, the HLT consisted of two parts: Level 2 (L2) and the Event Filter (EF).

The event can be saved only if it passes the selection criteria at all trigger steps. The trigger decision is checked sequentially. This means that once an event is rejected at L1 then it is not processed in the HLT system.

The Level 1 trigger has to accept/reject an event in a very short period of time (2.5 μs) using crudely segmented data from the calorimeter and muon systems. This L1 trigger determines the so-called Regions of Interest (RoIs). They include not only information on the position and p_T of the candidate objects but also the energy (missing-E_T). The RoIs are then passed to the HLT system. At first, the event is partially reconstructed using specialised software algorithms which have an access to all detectors at full granularity. The full scans can be performed for specific triggers if needed for a small fraction of events.

The selection is narrowed depending on the specific triggering criteria. In the jet trigger case, the threshold at the L1 trigger is already sharp since the jet events are dominant at high–p_T at the LHC energies. Therefore, in the jet triggers case it is easier to increase the threshold or add more requirements in order to reduce the accepted trigger rate.

The HLT algorithms may perform the reconstruction offline and can work either on RoIs or the full event. The HLT will make the final selection of events which will be then written to the mass storage for full offline analysis.

It is possible that after careful selection of trigger conditions, the trigger rates are still too high. In such a case the so-called prescale (PS) is applied either at L1 or HLT. The PS of value "n" determines that every n^{th} event passing the trigger conditions will be saved. The prescale of 1 means that all accepted events are passed. A negative value (usually -1) means that a trigger was not storing the events. It is worth mentioning that for each stored event a full trigger decision is available. The L1 and HLT are independent, meaning that the final prescale is their product.

Chapter 4

ATLAS Forward Proton Detector

The ATLAS experiment has two sets of forward proton detectors: Absolute Luminosity For ATLAS (ALFA) [30] and ATLAS Forward Proton (AFP) [29]. ALFA is used only during special runs, mainly for the measurements of the elastic scattering and will not be discussed further as these detectors are not a subject of this thesis.

The AFP consists of two Roman Pot (RP) stations (called Near and Far) on each side of the ATLAS Interaction Point (IP1). They are located around 210 m from the ATLAS IP. The stations allow for horizontal (in the accelerator plane) insertions of the detectors into the beam pipe, up to a few millimetres from the beam. Each RP contains the Silicon Tracker (SiT) consisting of four layers of silicon pixel detectors. Additionally, the Far stations host the Time-of-Flight (ToF) detectors. The scheme of the AFP components is shown in Figure 4.1.

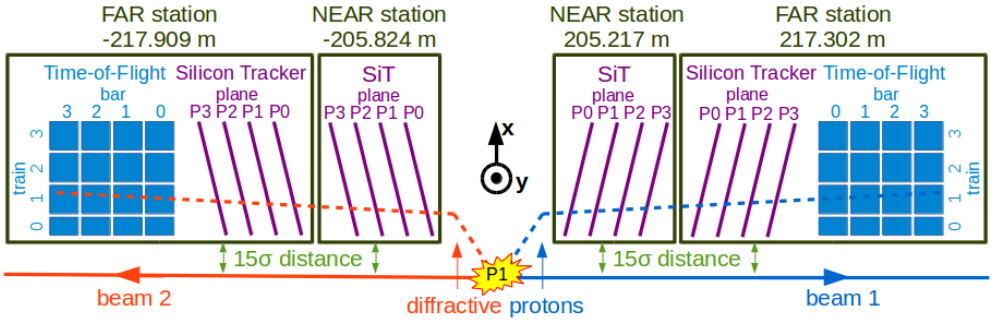


Figure 4.1: Scheme of the AFP detectors around the ATLAS interaction point.

The detector movement is possible due to the Roman Pot technology. In the case of the non-stable beams, the pots are in the "safe" position, which is about 40 mm away from the beam. When the stable beam flag is shown, the detectors can be positioned 2 to 4 mm away from the beam centre. The distance from the beam may be different for each station. This is defined by the LHC optics, accordingly to the machine safety rules. The RP motor position is cross-checked with the Linear Variable Differential Transformer (LVDT) readout

and the movement control is very precise, with 2 μm accuracy.

4.1 AFP Silicon Detectors

The AFP pixelated silicon tracking system (SiT) provides the position measurement of the scattered protons. Each station is equipped with four SiT modules tilted by 14°. A single module consists of 336×80 pixels with a pixel size of 50×250 μm^2 and thickness of 230 μm. The dead edge at the beam side is only ∼100 μm. The detectors are expected to withstand a non-uniform radiation with the fluence of $3 \cdot 10^{15} \frac{n_{eq}}{cm^2}$ per 100 fb^{-1}.

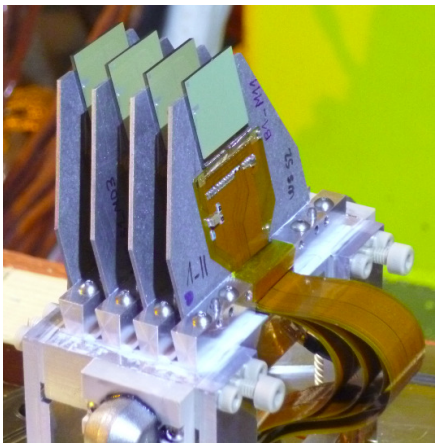


Figure 4.2: Four AFP Silicon Tracker detectors mounted on the heat exchanger.

The read-out is performed by the FE-I4B front-end chip [55], which is operating with a 40 MHz externally supplied clock. The FE-I4B have also the triggering capabilities.

Trajectories of the diffractively scattered protons are nearly parallel to the beam. The slope is about 20 μrad [31]. If the SiT planes would be perpendicular to the beam, then they would have a resolution of $50/\sqrt{12}$ μm in the horizontal direction. This is improved by titling the detector, which causes charge sharing between the pixels, thus improving the resolution. For the AFP, the optimum tilt angle was found to be 14°, resulting in the resolution of a single plane of about 6 μm [56].

4.2 AFP Time-of-Flight Detector

The main backgrounds to the hard diffractive events discussed below are due to the overlay of non-diffractive processes with forward protons originating from the pile-up interactions, *i.e.* additional pp interactions occurring during a single bunch crossing. These backgrounds can be significantly reduced by using Time-of-Flight detectors.

The idea is to measure the difference of the time of flight of the scattered protons on both sides of the IP:

$$(t_A - t_C)/2$$

and to compare it to the position of the interaction vertex reconstructed by the ATLAS. The AFP ToF detector is composed of 16 L-shaped quartz bars (see Figure 4.3, left). When the scattered proton hits a bar, the Cherenkov light created by protons is guided to a Micro-Channel Plate Photo-Multiplier (MCP-PMT). After the amplification, a readout is done

using the radiation-hard electronics.

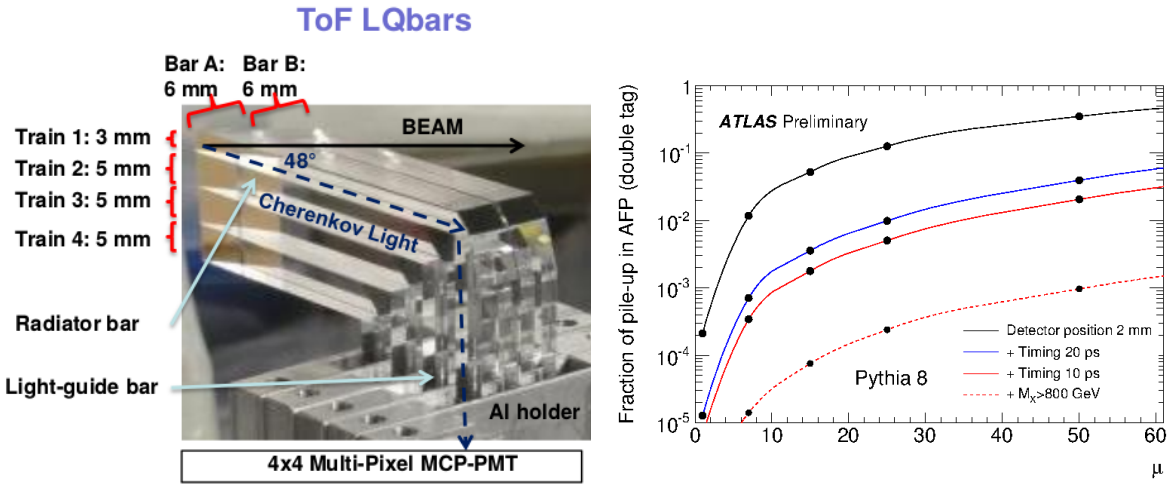


Figure 4.3: **Left:** AFP Time-of-Flight quartz bars. **Right:** Fraction of background in function of pile-up (from [56]).

Since a sizeable pile-up is expected during the regular ATLAS data-taking, the probability of the combinatorial background is high. The AFP Time-of-Flight detector was designed to reduce such a background by a factor of few, see Figure 4.3 (right). The black line represents the expected background reduction due to the forward proton tag (AFP tag) and the blue and red lines represent the background with timing measurements included assuming various values of the ToF timing resolutions.

It was demonstrated in [57] that the timing resolutions per bar can be as low as 20 ps. The per-proton resolution was found to be 20 ps and 26 ps for the A and C side, respectively.

4.3 AFP Acceptance

Before the installation of the detector packages, the simulations were performed in order to check the proton visibility in the AFP stations. A first simulation of the transport of a proton was carried out using MAD-X [58] programme. The obtained results are presented in Figure 4.4 (left). The beam starts at (0, 0, 0), the circles (squares) represent the positions of protons with $\xi = 0$ (0.07) and black and white colours represent different values of p_T. One can conclude that the protons with a higher energy loss are further away from the beam centre at the AFP detector positions. A presence of a non-zero transverse momentum results in an additional shift in x and y, accordingly to the azimuthal angle of a particle.

Another simulation (Figure 4.4, right) concerned the detector response and the forward proton track reconstruction and was done for the purpose of the project approval [29]. Diffractive processes were generated with Pythia 8 [59] and the Geant4 transport [60] was

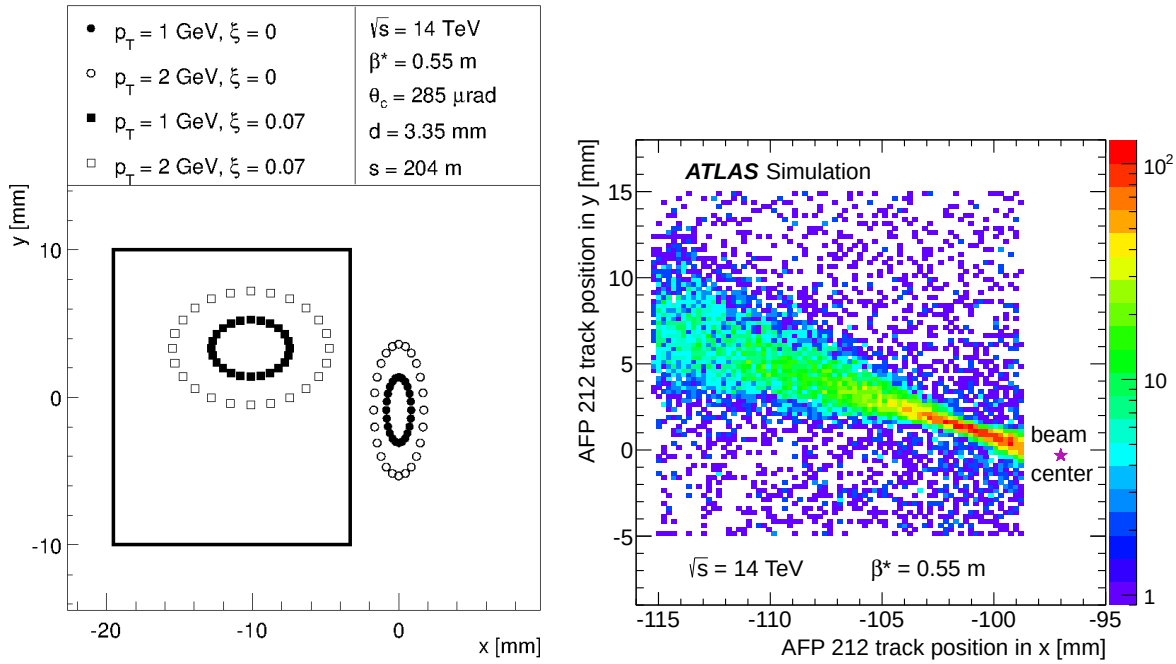


Figure 4.4: Simulated proton position in the AFP detector. **Left**: artificial protons with given ξ and p_T (from [31]). **Right**: simulation of the detector response and track reconstruction (from [29]).

used. This plot shows a characteristic distribution – the so-called diffractive pattern is visible.

Two types of acceptances are usually considered: geometric and mass. The first one is defined as the ratio of the number of protons with a given relative energy loss and transverse momentum that reached the AFP detector to the total number of the scattered protons having such ξ and p_T. An example of the geometric acceptance for one of the AFP stations is presented in Figure 4.5. The black region shows that more than 80% of the scattered protons will hit AFP. The two noticeable ξ limits are due to the beam-detector distance (lower limit) and the collimator settings (upper limit).

Mass acceptance is defined as the probability that the forward protons, emerging from a process in which the central system of a given mass is produced, will be visible in the AFP. For example: if an exclusive process $pp \to p + X + p$ with two forward protons and a hypothetical centrally produced particle of mass $M_X = 700$ GeV occurs, then there is a 70% chance to observe the scattered protons in the AFP detectors (with the ξ-values within the AFP acceptance) if the SiT modules were positioned 3.5 mm from the beam centre, (*cf.* Figure 4.5, right plot).

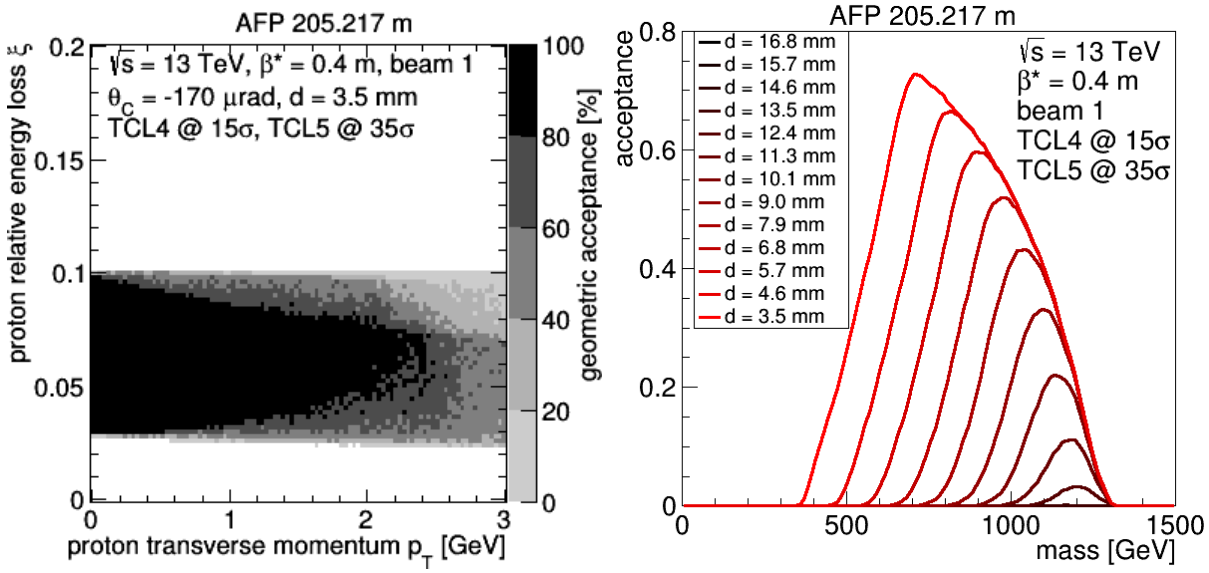


Figure 4.5: AFP Near station acceptance: geometric (**left**) and mass (**right**). From [29].

4.4 Proton Reconstruction in AFP

Before obtaining full information about the AFP-tagged proton a couple of steps must be taken. At first, the pixel clusters are formed independently within each plane and station. The algorithm starts with a pixel (a seed) not associated with any cluster, searches for neighbouring pixels and adds them to a cluster until no more neighbouring pixels are found.

The cluster position is calculated as the deposited charge-weighted average position of the centres of the associated pixels and it is calculated with respect to the edge of the sensor corrected by the inter-plane alignment[1]. Due to detector geometry, it is expected that the number of hits belonging to a cluster related to a single proton is less than four. The clusters made of a large number of pixels are expected to be formed due to the particle showering in the Roman Pot or detector material.

In the next step, tracks are formed individually for each AFP station. The procedure is similar to the clustering algorithm with the exception that it creates clusters of clusters. It is required that the (x, y)-positions of the centres of the clusters belonging to a super-cluster candidate (chosen from all SiT layers) do not deviate by more than 0.5 mm. The super-clusters are the track candidates. Then, the track parameters are reconstructed. Each candidate must have associated clusters in at least two different layers of the SiT. The x and y positions and slopes are calculated and corrected by applying the global alignment[2]. The position is calculated with respect to the edge of the first plane.

Finally, properties of the protons scattered into the AFP are found based on the reconstructed tracks. The procedure takes into account the machine optics and the detector

[1] Alignment of the SiT modules with respect to each other.
[2] Actual position of the station (SiT plane) with respect to the beam-pipe centre.

alignment [61]. Information about the LHC optics is held in the form of the proton transport parametrisation [31], which allows calculation of the proton trajectory position and slopes at a given distance from the IP.

The proton kinematics is calculated for each AFP side independently. At first, a pair of the track candidates (one in the NEAR and one in the FAR station) is found. The distance in the (x,y) plane between the reconstructed track in the NEAR station and FAR station can not exceed 2 mm. If the track found in a station is matched with more than one track in the other station, additional protons may be reconstructed.

All these algorithms are implemented into the reconstruction package called AFPToolBox. With the help of the AFPToolBox, the proton energy is unfolded from the detector measurements applying the parameterisation (see [62]) or the "minimisation" method. The latter uses the proton transport software performing the χ^2 minimisation.

AFPToolBox deals also with Monte Carlo samples. In this case, the truth-level protons are required to have sufficient energy (ξ <0.3). The proton momentum is smeared with the beams' angular spread, and its transport through the LHC lattice is simulated. Next, the Gaussian smearing of the proton position with experimental spatial resolutions is applied. The protons simulated in such a way undergo the same reconstruction procedure as the experimental data, starting from the reconstructed track level.
The AFPToolBox package was commonly used in the analysis presented in this thesis.

Chapter 5

Single Diffractive Jet Analysis

Analysis of hard diffractive events was one of the major motivations for the construction of the ATLAS Forward Proton detectors [29]. As was discussed in [63], such studies may bring insight into the Pomeron structure since in the case of Single Diffractive and Double Pomeron Exchange jet production, the leading diagram (see Fig. 1.6) results in the gluonic jets, as the $gg \to gg$ production is expected to happen in the majority of cases [64]. Moreover, the measurements of the cross-section and the gap survival probability are interesting on their own.

The Single Diffractive Jet production process can be characterised by the presence of the following features: (a) – rapidity gap (a region in rapidity devoid of produced particles), (b) – proton(s) scattered at very small angles, (c) – two jets produced in the central region.

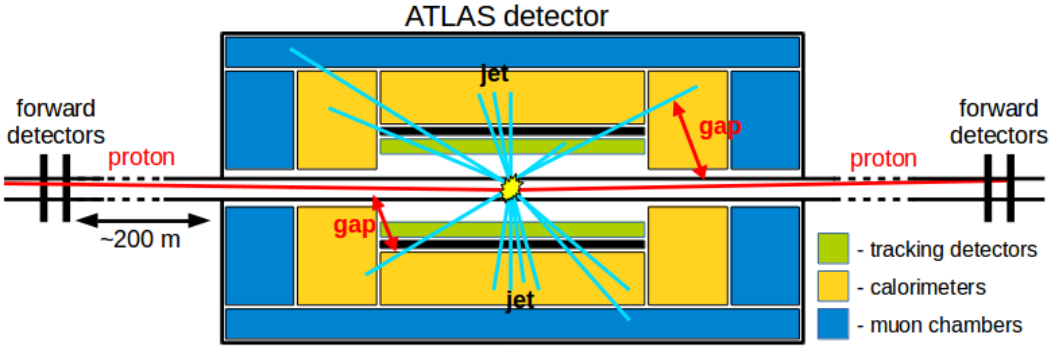


Figure 5.1: Concept of the diffractive jet measurement using the ATLAS detector: the jets are measured in the central part (tracker or calorimeter) and a proton(s) is registered in the forward detectors. The rapidity gap may also be observed.

A schematic view of the measurement using the ATLAS detector is shown in Fig. 5.1. The idea is to measure the jets in the central detectors: the tracker and calorimeters. The scattered protons should be registered by the ATLAS Forward Proton detectors. It should be noted that a rapidity gap may also be observed, however for the kinematic range defined

by the AFP acceptance, the gaps are expected to be rather small [65].

The main backgrounds to the diffractive jet production are non-diffractive jets overlaid with forward protons produced in the minimum bias (pile-up) interactions. The pile-up is quantified by the average number of pp interactions per bunch crossing, $< \mu >$. Examples of the signal and background topologies are shown in Figure 5.2.

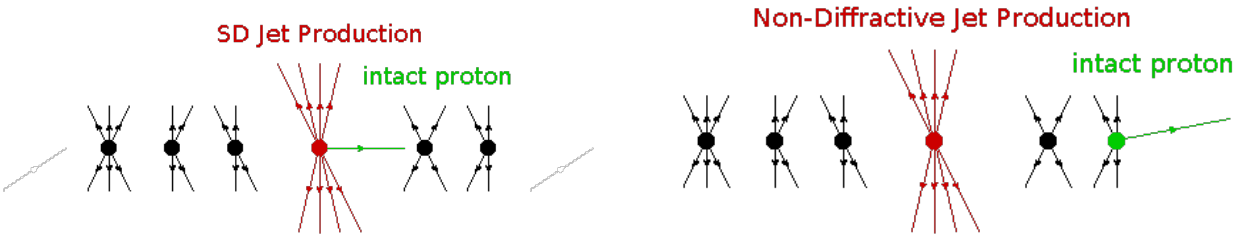


Figure 5.2: **Left:** signal – Single Diffractive jet production; diffractive proton (green line) originates from the same vertex as jets. **Right:** backgrounds to the single diffractive jet production: a proton does not originate from the jet vertex but from the pile-up interaction.

Initial feasibility studies were done within the *LHC Forward Physics Working Group* and published in [66]. Judging from Fig. 5.3, taken from Ref. [67], the preferred data-taking scenario assumes a relatively low pile-up. In the case of $\mu \gg 1$, the purity of the sample is expected to be very small, making the measurement extremely challenging, and nearly impossible, since the backgrounds would have to be perfectly controlled.

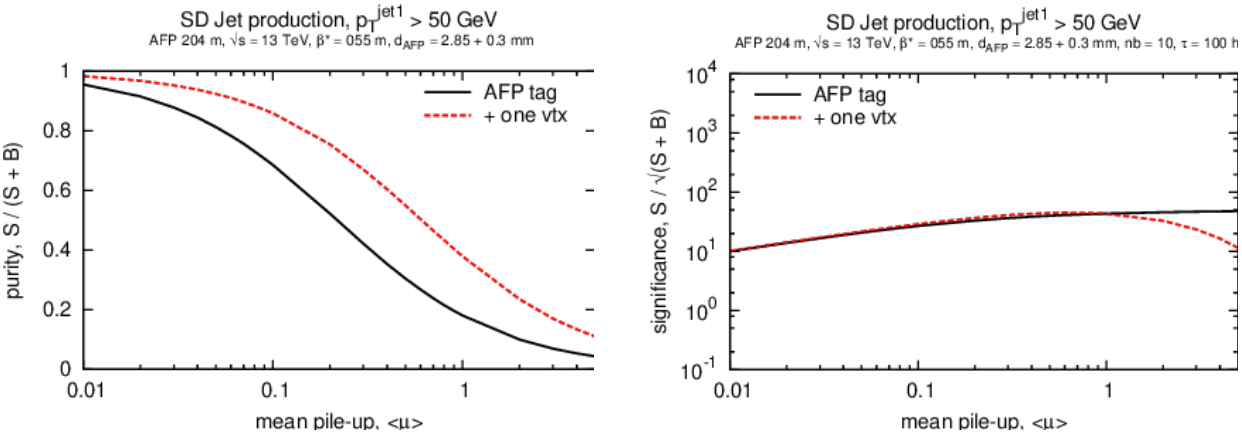


Figure 5.3: Monte Carlo predictions for Single Diffractive jet production with protons tagged in the AFP detectors for $\sqrt{s} = 13$ TeV, $\beta^* = 0.55$ m and the jet with $p_T > 50$ GeV. **Left:** purity, **right:** significance as a function of the average pile-up, $< \mu >$. The number of bunches multiplied by the data collecting time was assumed to be 1000. In the analysis, the following selection criteria were considered: the AFP proton tag (black solid line) and a tag + one reconstructed vertex (red dashed line).

As shown in Fig. 5.3, the purity of the sample can be increased if, in addition to the proton tag, a single vertex is required in ATLAS. Also, a difference in activity in the

forward region (forward calorimeters) can be expected. This can be due to the rapidity gap or differences in the properties of the proton and Pomeron remnants.

5.1 Data

The analysis presented in this Chapter is based on data collected by the AFP detectors in 2017 during a few special low-μ runs. It should be noted that the AFP recorded in total 32.0 fb^{-1} of data in 2017. However, the majority was collected at high-μ and such data would not be useful for the SD JJ analysis. The list of the data sets, together with a few basic pieces of information is shown in Table 5.1. The presented integrated luminosity is the sum over all Luminosity Blocks[1] (lumiblocks, LB) which passed the so-called "Good Run List" requirement (GRL, see [68]). Since it is interesting to study the dynamics of the system when a proton is tagged on a chosen side, the analysis was done separately for those two cases. Consequently, separate GRLs were provided for protons on sides A and C. Thus, the separate integrated luminosity values were used in this analysis for each side.

Table 5.1: An overview of 2017 low-μ runs with integrated luminosity from LBs passing the GRL requirements separate for protons on the ATLAS A and C sides.

ATLAS Run Number	**LHC Fill**	**Pile-up μ**	**Int. Luminosity [nb^{-1}] for protons on side A**	**Int. Luminosity [nb^{-1}] for protons on side C**
331020	6019	∼1.0	56.866	510.841
336505	6238	∼0.04	44.751	60.2411
341294	6404	∼2.0	709.542	709.542
341312	6405	∼2.0	18245.492	18234.639
341419	6411	∼2.0	31636.072	31593.050
341534	6413	∼2.0	47663.701	52680.387
341615	6349	∼2.0	31772.631	31772.631
341649	6417	∼2.0	6543.940	6449.680
		∼1.0	3325.167	3325.167

Data taking process during these runs was not perfectly smooth – a couple of issues were reported. First, the number of active SiT planes in station A Far was two during run 331020 and three in the other runs (instead of the default of 4). This had to be taken into consideration, especially during the trigger efficiency studies. The impact on the proton reconstruction efficiency was, fortunately, minor [69]. Another issue was the use of incorrect settings for the forward calorimeter in run 336505: a high-μ configuration was used, resulting, inter alia, in different pedestals. Other issues, which happened occasionally

[1] A luminosity block is a period of the data taking lasting about 1 minute.

and do not affect whole runs were covered by the GRLs (*cf.* the difference in luminosity collected in run 331020 with proton tagged on sides A and C).

5.1.1 Good Run List

On top of the "standard" ATLAS Good Run List prepared by the ATLAS Data Quality team, the additional quality criteria reflecting the condition of the AFP detectors had to be included. These are:

- requirement of the Roman Pots to be inserted and to take data in a stable position (no movement allowed),
- lack of problems related to the data acquisition (errors, corruption, *etc.*),
- at least two fully operational SiT planes in a given station (powered; with non-zero High Voltage).

If the above-mentioned problems were observed within a given lumiblock, this LB was removed from GRL. It should be noted that the "standard" duration of a lumiblock is 60 seconds. The list of LBs used for the analysis is listed in Tab. 5.2.

Table 5.2: List of Luminosity Blocks (LBs) passing Good Run List.

Run	**Side**	**LBs**
331020	A	441,444,446,449,456,460,468,496,507,515,521,531,541,544,546,547,561,564
	C	[438,470],[476,479],[485,489],[491,504],[507,551],[553,567]
36505	A	[216,329],[394,454]
	C	[216,454]
341294	A	[157,223]
	C	[157,223]
341312	A	[150,358],[361,545],[548,559],[563,610],[612,741]
	C	[150,358],[361,545],[549,559],[563,610],612,[614,741]
341419	A	[117,901],915,[919,1140]
	C	[117,292],[294,901],915,918,[920,1140]
341534	A	[91,549],[551,673],[675,700],[872,1106],[1108,1111],1113,1114,1116, [1118,1716],[1718,1741],[1743,1811]
	C	[91,549],[551,1105],[1120,1716],[1718,1741],[1743,1812]
341615	A	[119,243],[245,252],[257,1162]
	C	[119,243],[245,252],[257,1162]
341649	A	[152,187],[193,364],[390,603],[607,613]
	C	[152,169],[171,186],[193,225],[227,364],[390,603],[607,613]

5.2 Monte Carlo Samples

The Monte Carlo events were produced using Pythia 8 tune A3 [70]. This tune is aimed at the modelling of the low-p_T processes. It uses the NNPDF2.3LO PDF set [71]. The sample contained single diffractive events (SD MC) with the reported cross-section $\sigma_{gen} = 12.83$ mb, filtered at the generator level. The filter required the presence of a forward proton with the energy loss within the range of $0 < \xi < 0.2$ and di-jet with $p_T > 12$ GeV. The filter efficiency was 0.000853.

It is interesting to see if such a sample provides a reasonable description of the collected data. Indeed, as can be seen in Fig. 5.4, where $\mu = 1$ part of run 341649 is shown, the shape of basic distributions like the cluster and track multiplicities are in a rough agreement. It should be noted that these distributions were taken "as they were" *i.e.* neither including corrections nor the signal selection criteria. In fact, in the case of the MC, the requirements were applied only at the generator level. In the experimental data case, the trigger HLT_j20_L1AFP_A_OR_C_J12 was used. The AFP trigger delivers a high probability of a proton reconstruction in the AFP detector on side A or C. The requirement of jet in data is met by triggers on L1 and HLT. This should result in at least one jet with $p_T > 20$ GeV.

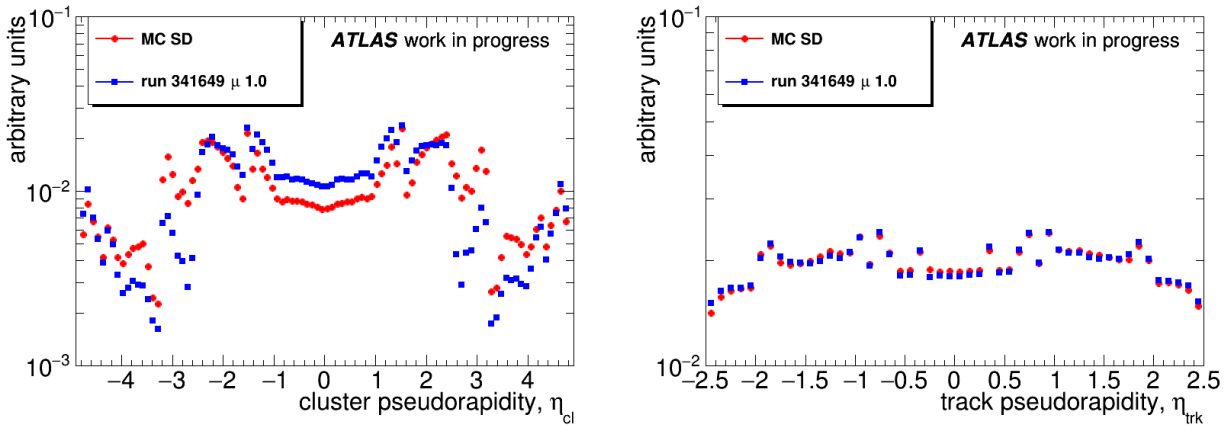


Figure 5.4: Comparison of the cluster (**left**) and track (**right**) pseudorapidity distribution for $\mu = 1$ of run 341649. Neither the selection cuts on the quality of the objects nor the background suppression were applied.

5.3 Data Properties

In the Single Diffractive jet analysis, the following objects are exploited: the jets, the forward protons, the vertex and the activity in the central parts of the detector (tracker, calorimeters). It is useful to check the basic properties of these objects and define the

quality criteria if needed. The discussion on such quality criteria will be presented later on in this Chapter.

The SD event selection is based on the requirement of two jets and a forward proton, therefore the interest in those objects is rather obvious. Information about vertices, reconstructed from the charged particle tracks, is important since it is expected to be a selection criterion. In addition, Single Diffractive processes are expected to exhibit a relation between the central system and the forward proton(s). To exploit this relation one should analyse the energy lost by a proton and compare it to that calculated from the central system. The latter can be based on the jets, clusters or tracks. For diffractive events, the energy seen in the central detector is connected to the energy lost by a proton in the interaction. Thus, in this thesis, it will be often simply referred to as the energy loss, as quantified by ξ. There exist various ways of calculating ξ and their expected limitations are listed in Table 5.3.

Table 5.3: Various ways of calculating the relative energy loss of a proton including their description. The plus and minus sign in the formulas indicates the side on which the forward proton is produced (plus/minus is for the A/C side, respectively).

Object	**Formula**	**Description**
proton	$\xi_p^{\pm} = 1 - \frac{E_{proton}}{E_{beam}}$	values limited by acceptance of the AFP detectors ($0.03 \lesssim \xi_p \lesssim 0.1$)
clusters	$\xi_{cl}^{\pm} = \frac{1}{\sqrt{s}} \sum_{cl} p_T \exp(\pm\eta)$	values expected to be similar to ξ_p
tracks	$\xi_{trk}^{\pm} = \frac{1}{\sqrt{s}} \sum_{trk} p_T \exp(\pm\eta)$	likely to be smaller than ξ_p due to the central detector acceptance ($\lvert\eta_{trk}\rvert < 2.4$)
di-jet	$\xi_{dijet}^{\pm} = \frac{M_{dijet}}{2E_{beam}} \exp(\pm\eta_{dijet})$	anticipated to be smaller than ξ_p since only a part of the central system is taken

5.3.1 Trigger Selection

Before starting the analysis, one should identify the trigger(s) providing the best statistics resulting from the selection of the SD events. Such a selection is based on the requirement of a presence of at least two jets and only one proton passing the "good quality" criteria. It includes also the requirement of one vertex and GRL. More details about the selection will be provided later, in Sec. 5.4. There is a couple of the AFP+jet triggers available in low-μ runs. In this section, the most suitable trigger for the single diffractive di-jet analysis will be identified. There is not an obvious choice as the triggers may have different efficiencies and prescales[2].

[2]In addition, in some cases the prescales changed during the data-taking.

The triggers considered below, showing a potential to be used in the analysis, were chosen among all available in the trigger menu. The general naming convention is explained in Sec. 3.6, but a few detailed comments are at the place:

- **L1_MBTS_2** means that the Level 1 trigger fired when at least 2 MBTS tiles were activated[3],
- **L1_J12** requires at least one (L1) jet; the level 1 threshold of 12 GeV should reflect the mid-point of the turn-on curve of 40 GeV for reconstructed jets (see Fig. 5.8),
- **L1_AFP_A/L1_AFP_C** require hits registered in at least 2 (out of 3 considered for triggering) AFP SiT planes in each station (NEAR and FAR) on a given side,
- in some of the trigger chains combinations of the above L1 items were used – this is reflected in the trigger name,
- **HLT_noalg** means that no additional algorithm was applied at the HLT, *i.e.* the event passed through without further selection; although, it should be noted that there might be a prescale on the HLT in addition to that potentially applied at the L1,
- **HLT_j20** means that the jet reconstruction algorithms (which are supposed to be the same as in "regular, offline" reconstruction) were applied, in this case, an event passed the selection if it contained at least one jet with p_T of at least 20 GeV,
- **HLT_mb_sptrk** requires at least one reconstructed track.

All trigger items considered to be suitable for SD JJ analysis are listed in Figure 5.5. As an example, run 331020 is shown, separately for a proton tagged on sides A (left) and C (right). Different colours represent consecutive cuts used in the analysis. The grey areas represent all triggered events, red – events after requiring one vertex, blue – events passing also AFP GRL, the magenta plot – events after requiring two jets and the green one shows events with the additional requirement of a proton either on side A or C. It should be noted that the number of all events (grey plot) saved by individual triggers is in agreement with what was roughly predicted from the trigger rates during the 2017 data-taking [72].

As was expected, the proton requirement will cut out more events initially saved by the MBTS triggers than by the AFP triggers. The requirement for an event to contain at least two jets will cut out most events accepted by triggers without the jet requirement *e.g.* HLT_mb_sptrk or HLT_noalg_L1AFP_A_OR_C. This requirement fewer influences events accepted by the J12 trigger at L1. There are also differences between events triggered by a proton on sides A and C: side C seems to be more effective. For example, in the run 331020 case, after all cuts, the HLT_j20_L1AFP_A_OR_C_J12 saved only 2654 events on side A while 52155 for side C.

At this point, it should be noted that these triggers had different prescales at L1 or HLT levels. For example, in run 331020 the L1_AFP_A_OR_C trigger prescale was 30, whereas the L1_AFP_A_OR_C_J12 trigger was not prescaled.

[3]Regardless of the side; for at least two planes but on the same side L1_MBTS_A or L1_MBTS_C could be used.

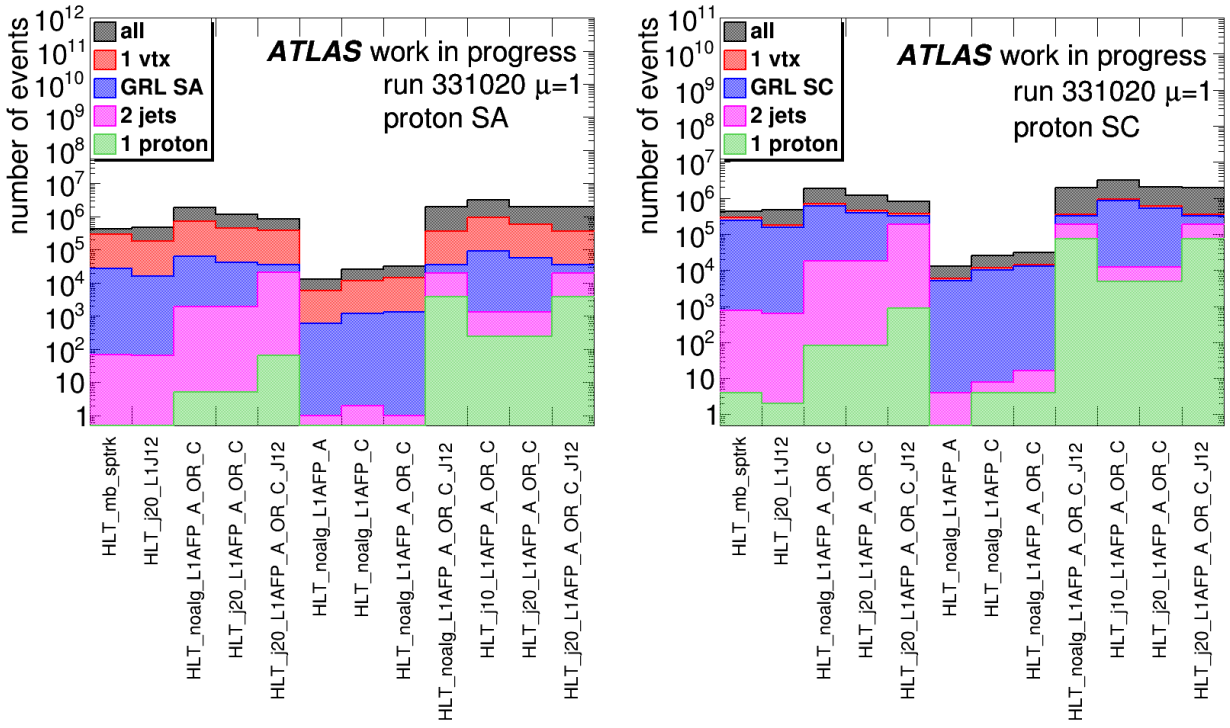


Figure 5.5: The number of events saved by a given trigger (listed on the x-axis) after consecutive cuts. The number of events without cuts and after cuts on one vertex (grey and red histograms) is the same for both plots. Further selection criteria require distinction of either side A or C (**left** and **right** plots).

It is worth observing that raw number of the accepted events triggered by a given item may not provide a full picture. This is because after the selection the events triggered by various algorithms (especially when the items had different L1 and HLT prescales) may belong to the independent samples. Therefore, it is worth checking the percentage of events shared between the triggers. This is plotted in Figure 5.6. Each bin represents the percentage of events saved simultaneously by the triggers, listed on the x and y axes, calculated using the total number of events passing the selection for side A (upper plot) or C (lower plot). In this example run 331020 with $\mu \sim 1.0$ was considered. The bin located on the diagonal (y = x) shows how many events were saved by a given trigger. For example, the triggers HLT_j20_L1AFP_A_OR_C_J12 and also HLT_noalg_L1AFP_A_OR_C_J12 saved 96% of events.

Clearly, there is a correlation between those triggers. However, reading only the values on the diagonal does not imply that all events triggered by HLT_j20_L1AFP_A_OR_C_J12 are exactly the same as those passing HLT_noalg_L1AFP_A_OR_C_J12. One must also see the percentage of common events for them (9^{th} and 12^{th} bins) which is also 96%. Only now one can safely conclude that both triggers are indeed saving exactly the same events.

The 4% of events, not saved by the previously discussed algorithms, can be retrieved from the HLT_j10_L1AFP_A_OR_C trigger. One may observe that it saves 7% of the total number of events and shares around 3% of them with the trigger containing also the L1J12 selection. This means that the use of these two triggers instead of only the first one would accept an additional 4% of events.

Eventually, since the $HLT\_j20\_L1AFP\_A\_OR\_C\_J12$ trigger provides the best statistic for all considered runs, it will be used as the main trigger in the following analysis.

Please note that the analysis described in this Section was done for all considered runs with 331020 taken as an example. The details (percentage, number of events passing a given criterion) obviously differ between the runs, but the general conclusion on trigger selection is valid for all of them. The list of the L1 and HLT prescales of the chosen main trigger for all considered runs is shown in Table 5.4.

5.3.2 Trigger Efficiency

After choosing the most suitable trigger its efficiency has to be established. In this section, the study of the HLT_j20_L1AFP_A_OR_C_J12 trigger efficiency will be presented.

It was checked (see Appendix C) that the AFP triggers efficiency is independent of the jet trigger performance (*e.g.* the L1AFP_A trigger efficiency is constant as a function of the leading jet transverse momentum). However, it depends on the energy loss of a proton as shown in Fig. 5.7. It was checked that a linear fit provides a good description ($\chi^2/NDF \sim 1$) of the efficiency dependence as a function of the proton energy loss (within the range of interest) for side A in all runs and side C in run 331020 and 336505. The efficiency of the triggers on side C in runs with $\mu \gtrsim 1$ had to be described by function:

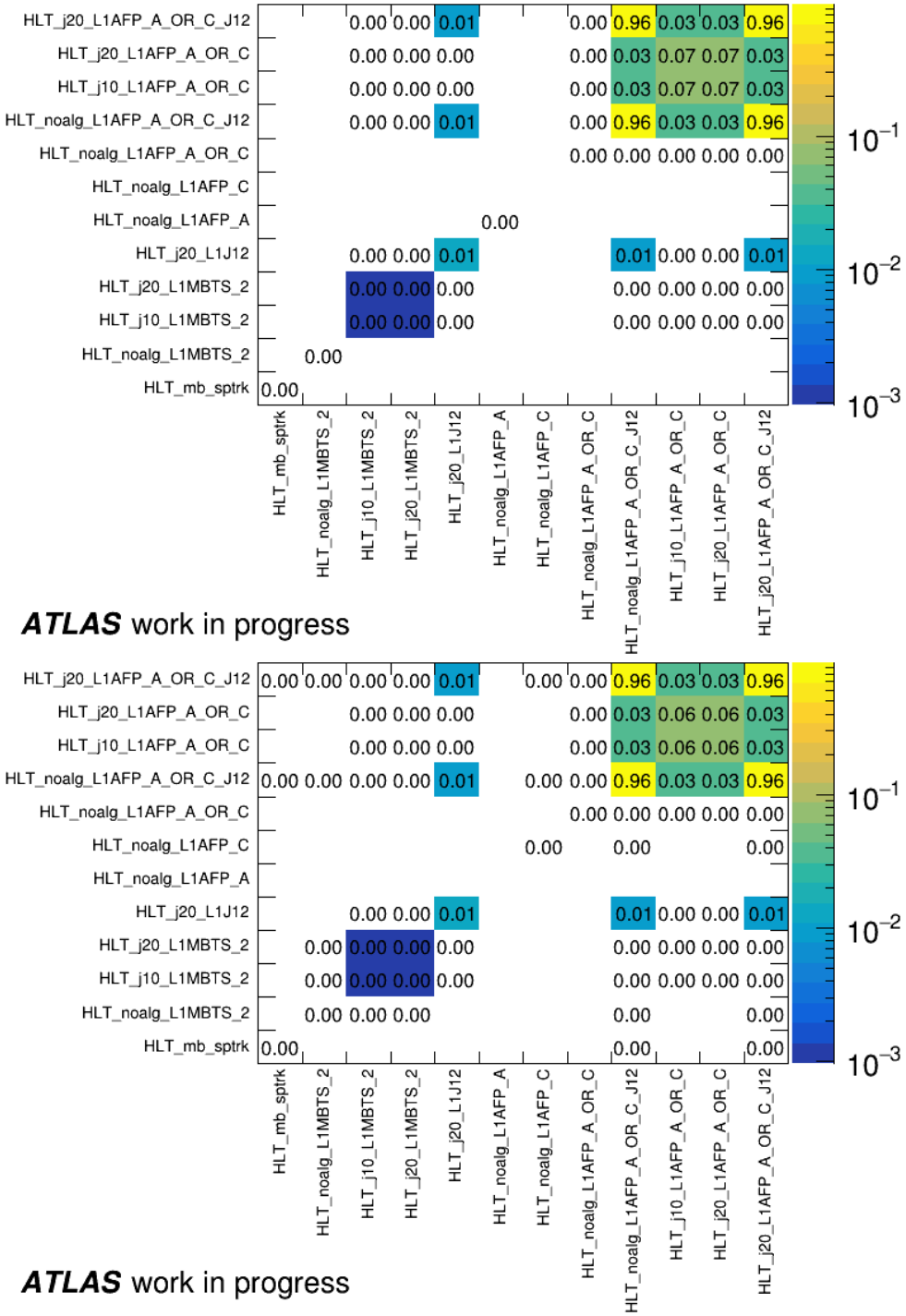


Figure 5.6: The percentage of events from run 331020 saved simultaneously by the triggers listed on the x and y axes. The percentage is calculated using the total number of events passing the event selection for side A (**top**) or C (**bottom**).

Table 5.4: List of the HLT_j20_L1AFP_A_OR_C_J12 trigger prescales for the L1 and HLT triggers in the specified LB ranges.

Run	LB range	L1 PS	HLT PS
331020	[438; 576]	1	1
336505	[216; 465]	1	1
341294	[157; 248]	1	50
341312	[150; 188]	3	50
	[189; 759]	2	50
341419	[115; 1140]	2	50
341534	[73; 1812]	2	50
341615	[106; 1163]	2	50
341649	[106; 389]	2	50
	[390; 625]	1	50

$y = p_0 + p_1/(1 + (x/p_2)^{p_3})$ in order to reproduce the observations in the lower ξ range. In Figure 5.7 the results for runs 31020, 341649 ($\mu \sim 1$) and 341649 ($\mu \sim 2$) are shown as examples, but the studies were done for all considered runs.

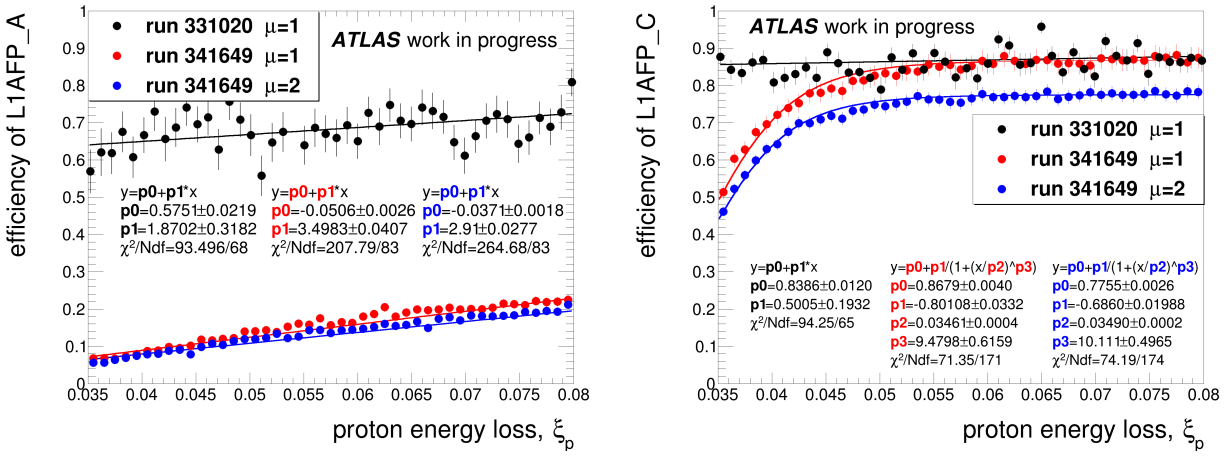


Figure 5.7: Efficiency of the Level-1 AFP trigger on side A (**left**) and C (**right**) as a function of the proton energy loss within $0.035 < \xi < 0.08$. Black, red and blue lines are for run 331020, 341649 ($\mu \sim 1$) and 341649 ($\mu \sim 2$), correspondingly.

The jet trigger efficiency depends on the transverse momentum and pseudorapidity of the leading jet. In this analysis, twelve pseudorapidity regions were considered: [-5.0;-3.0), [-3.0;-2.4), [-2.4;-1.8), [-1.8;-1.2), [-1.2;-0.6), [-0.6;0.0), [0.0;0.6), [0.6;1.2), [1.2;1.8),[1.8;2.4), [2.4;3.0), [3.0;5.0). For each such region, the efficiency of the L1J12 algorithm as a function of leading jet transverse momentum was estimated. It should be noted that for this analysis events were triggered by independent (*i.e.* unbiased for jet events) trigger on MBTS: HLT_noalg_L1MBTS_2. As an example, trigger efficiencies as a function of

leading jet transverse momentum for pseudorapidity region $-1.8 < \eta_{LJet} < -1.2$ for run 331020 and 341649 are shown in Fig. 5.8. To each such distribution an effective function $y = p0 + p1/(1+(x/p2)^{p3})$ was fitted in order to have a continuous description of a "turnover region". The fit was imposed to provide an efficiency value between 0 and 1. Such fits were done for all pseudorapidity regions in all considered runs.

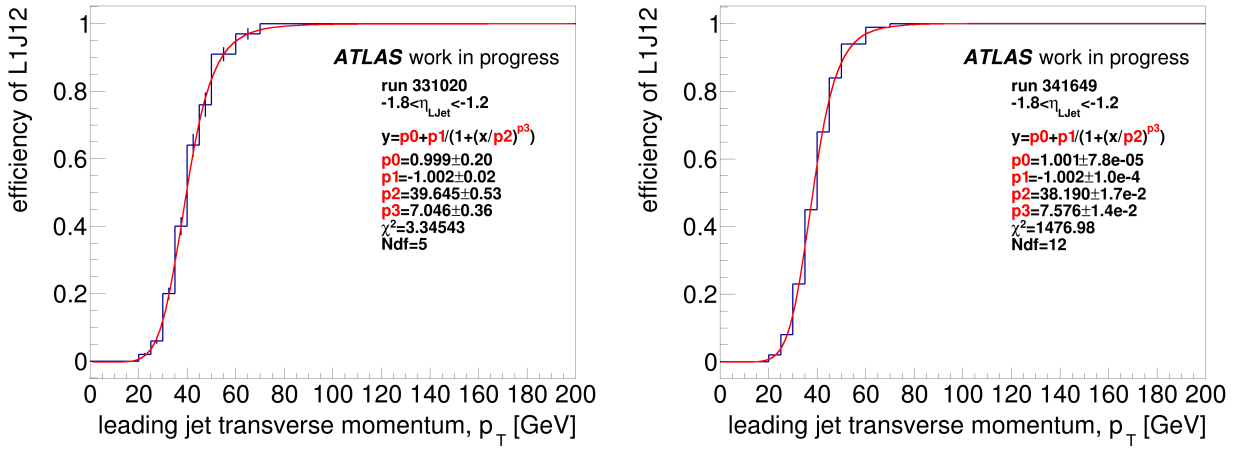


Figure 5.8: The L1J12 trigger efficiency in run 331020 (**left**) and 341649 (**right**) as a function of the leading jet transverse momentum for the pseudorapidity $\eta \in (-1.8; -1.2)$. The red line represents the fit of $y = p0 + p1/(1 + (p_T/p2)^{p3}$, where $p1, ..., p4$ are the fit parameters.

The HLT_j20_L1AFP_A_OR_C_J12 used in the analysis has also a jet selection at the HLT level. Therefore, the efficiency of the HLT algorithm should be considered. Fortunately, the efficiency of the HLT_j20 trigger for events triggered by L1J12 is 100%. It was checked and is visible in Figs. 5.5 and 5.6, where the HLT_noalg_L1AFP_A_OR_C_J12 and HLT_j20_L1AFP_A_OR_C_J12 are triggering exactly the same events.

5.3.3 Clusters

For the single diffractive jet analysis information coming from clusters reconstructed from the calorimeter signals is very valuable. It can be used for the proton energy loss reconstruction (ξ_{cl}, *cf.* Tab. 5.3), and also, as will be shown later, may serve as an additional criterion for the background suppression.

As was explained in Sec. 3.5.1, the clusters are reconstructed from the calorimeter cells. In principle, signals in cells, hence the clusters, maybe "corrupted" by *e.g.* electronic noise or malfunction of the readout system. For the majority of the ATLAS analyses, which aim for the searches or studies of objects with relatively high p_T, the above effects play a minor role. Unfortunately, for the diffractive studies, they may be crucial, thus it is worth performing an analysis to determine the dedicated quality criteria, in addition to the regular ones commonly used by ATLAS.

In the studies presented below only run 341312 is shown. However, the selection criteria (except for the list of noisy clusters which is the run specific) are valid for all analysed low-μ runs.

In order to find the characteristics of the noisy clusters the data from the colliding bunches (triggered by HLT_noalg_L1AFP_A_OR_C) were compared to the samples of events from the empty bunch[4] crossings (HLT_noalg_L1AFP_A_OR_C_EMPTY), first empty bunch (HLT_noalg_L1AFP_A_OR_C_FIRSTEMPTY) or unpaired bunches (HLT_noalg_L1AFP_A_OR_C_UNPAIRED_NONISO).

The colliding bunch events were required to have exactly one reconstructed vertex and at least 30 reconstructed tracks. Such a sample is enriched with the true "signal" clusters as the high multiplicity of tracks should correspond to a high multiplicity of particles produced in the event, thus a significant activity in the calorimeter. Events triggered on empty or unpaired bunches were required to have 0 reconstructed tracks – such samples should be enriched with noisy clusters. The trigger based on the AFP, together with the above requirement considering the tracks, should result in almost unbiased samples. Although, it should be mentioned that the AFP trigger in the majority of cases would result in having a forward proton reconstructed in an event.

In the standard ATLAS data format used in the analysis (the so-called xAOD), a number of properties describing the reconstructed clusters is available. To list a few: the depth at the cluster centroid, the isolation (the energy-weighted fraction of non-clustered perimeter cells), the energy fraction of LAr cells with quality larger than a given cut, *etc.* All of them, separately and as two-variable correlation, were studied. The vast majority of the obtained dependencies are similar for the "signal" and "noise" enriched samples mentioned above. However, five were identified to be useful for the cluster quality selection *i.e.* to reduce the noise effect in the calorimeter. It should be pointed out that the selection criteria discussed below are consequently used, *i.e.* once a given selection is identified, and are applied to obtain the distributions of other variables used for identification. In such a way the effect of correlation between variables was effectively considered.

At first, the timing of the cluster was considered (see Figure 5.9 left). Except for the empty bunch sample, the clusters' timing distribution has a characteristic peak around $|cl_t| \approx 24$ ns. Clusters with such high values of the timestamp, called the out-of-time clusters, are due to the neighbouring bunch-crossing collisions (previous or next) and should be removed. For the performed analysis, a cut $|cl_t| < 10$ ns was selected and will be consequently used from now on.

The differences between the clusters from colliding and non-colliding bunches were also noticed in the distribution of the energy fraction of the hottest cell (Figure 5.9 right). Here, all non-colliding samples have a characteristic peak at $\sim$0.96. It was decided that such clusters should be removed, thus a cut on $cl_{EngFracMax}$ <0.8 was introduced. It should be

[4]It should be noted that the EMPTY bunches are in reality not truly empty but populated with some protons. However, their number is orders of magnitude smaller wrt. the FILLED bunches.

added that such effect is observed also in other studies, *e.g* [73].

Next differentiating distribution was the cluster centroid, defined as:

$$cl_{MAG} = \sqrt{x^2 + y^2 + z^2},$$

where the x, y, z are the spatial coordinates of a cluster. From Figure 5.10 (left) one can conclude that the colliding bunches do not show a peak at around 6000. Therefore, a criterion $cl_{MAG} < 5500$ was formulated.

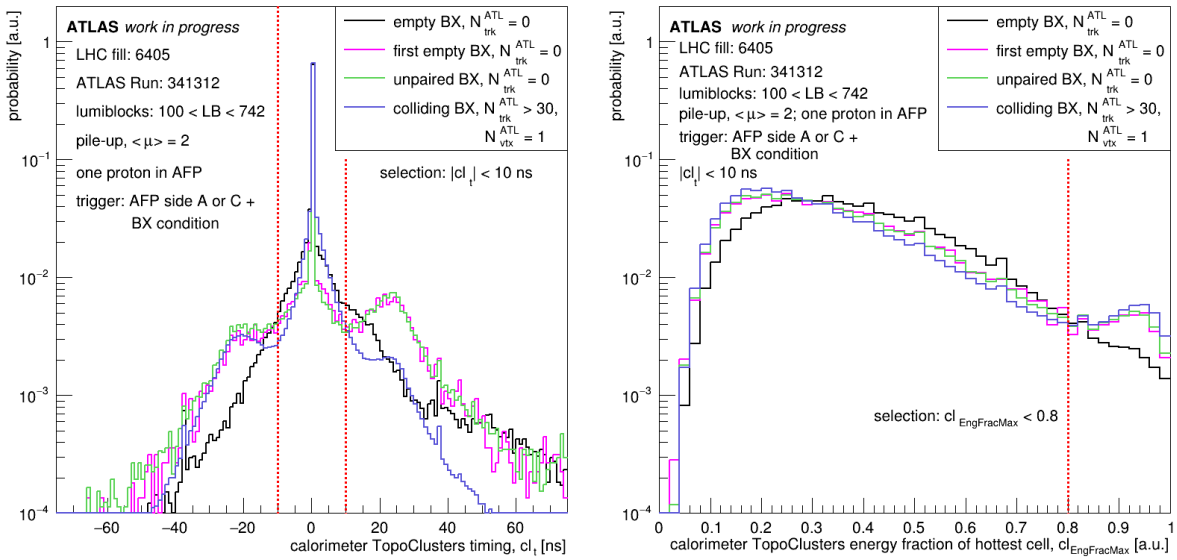


Figure 5.9: Distributions of the clusters timing (**left**) and the energy fraction of the hottest cell (**right**). The lines represent the samples with filled bunches (blue), unpaired (green), empty (black) or first empty (magenta) bunches. The timing plot was made without any additional cuts, the energy fraction was plotted after including a cut on the cluster time $|cl_t| < 10$ ns.

Finally, it is well known that the requirement of minimal cluster energy significantly decreases the noise. The distribution of the raw energy of a cluster is shown in Figure 5.10 (right). The significant amount of clusters in non-colliding and empty bunches, compared to colliding ones, have $cl_E^{raw} < 250$ MeV. Therefore such clusters are excluded from the analysis.

Next, the $\eta \times \phi$ map of clusters was considered. As this is a two-dimensional distribution, it is plotted separately for colliding (Fig. 5.11) and unpaired bunches (Fig. 5.12). It should be noted that the maps for empty bunches are similar to the unpaired ones, thus not shown. It is interesting how the map "evolves" after a given selection is applied. Starting from the top left where all clusters (no selection) are plotted, through the consecutive application of the cuts on timing (top right), energy fraction (middle left), cluster centroid (middle right), noisy pixel removal (bottom left) to minimal cell energy (bottom right).

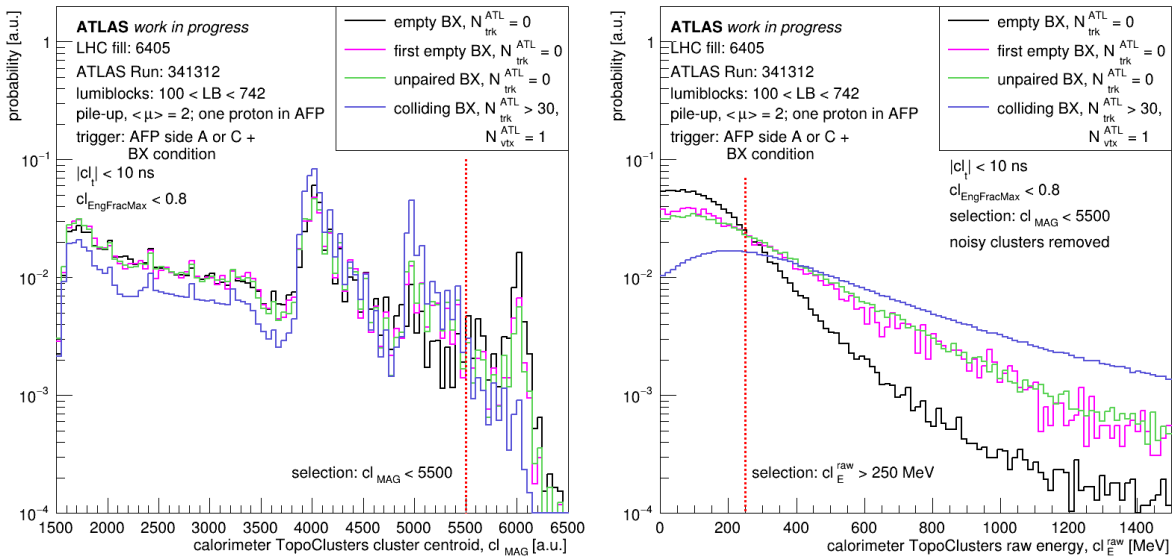


Figure 5.10: Distributions of the clusters centroid (**left**) and the raw energy (**right**). The lines represent the samples: filled (blue), unpaired (green), empty (black) or first empty (magenta) bunches. The cluster centroid distribution was made after applying the cuts on timing ($|cl_t| < 10$ ns) and energy fraction of the hottest cell (< 0.8). The cluster energy distribution include the selection criteria considering the timing, the energy fraction, the cluster centroid (< 5500) and the noisy pixel removal.

Looking at the $\eta \times \phi$ maps passing the selections 1-3 mentioned above (see Figures 5.11 and 5.12 middle right) the areas with activity much higher than that of the neighbours can be easily identified. These are the so-called hot pixels, visible in the same regions in the colliding and non-colliding samples. After applying the mentioned cuts, the hot pixels start to disappear. Unfortunately, not all. The remaining ones need to be removed "by hand". The result is visible in the plots marked as Selection 4 in the bottom left panels of Figs. 5.11 and 5.12.

The performed studies showed that there are no additional variables nor 2-variable correlations which can efficiently differentiate between the signal- and noise-enriched samples. Thus, the discussed set of five criteria is concluded to be the final one and was used in the following analysis.

Run 336505 Issues

The selection proposed in this section was applied to all low-μ runs taken in 2017 (see Tab. 5.1). Unfortunately, for the run with the smallest pile-up, the behaviour of the calorimeter (especially in the forward region) was far from expectations. It can be seen in Fig. 5.13 where the cluster multiplicity is plotted as a function of the cluster pseudorapidity, η_{cl}, and energy, E_{cl}. In run 336505 (cf. Fig. 5.13b) the number of clusters having

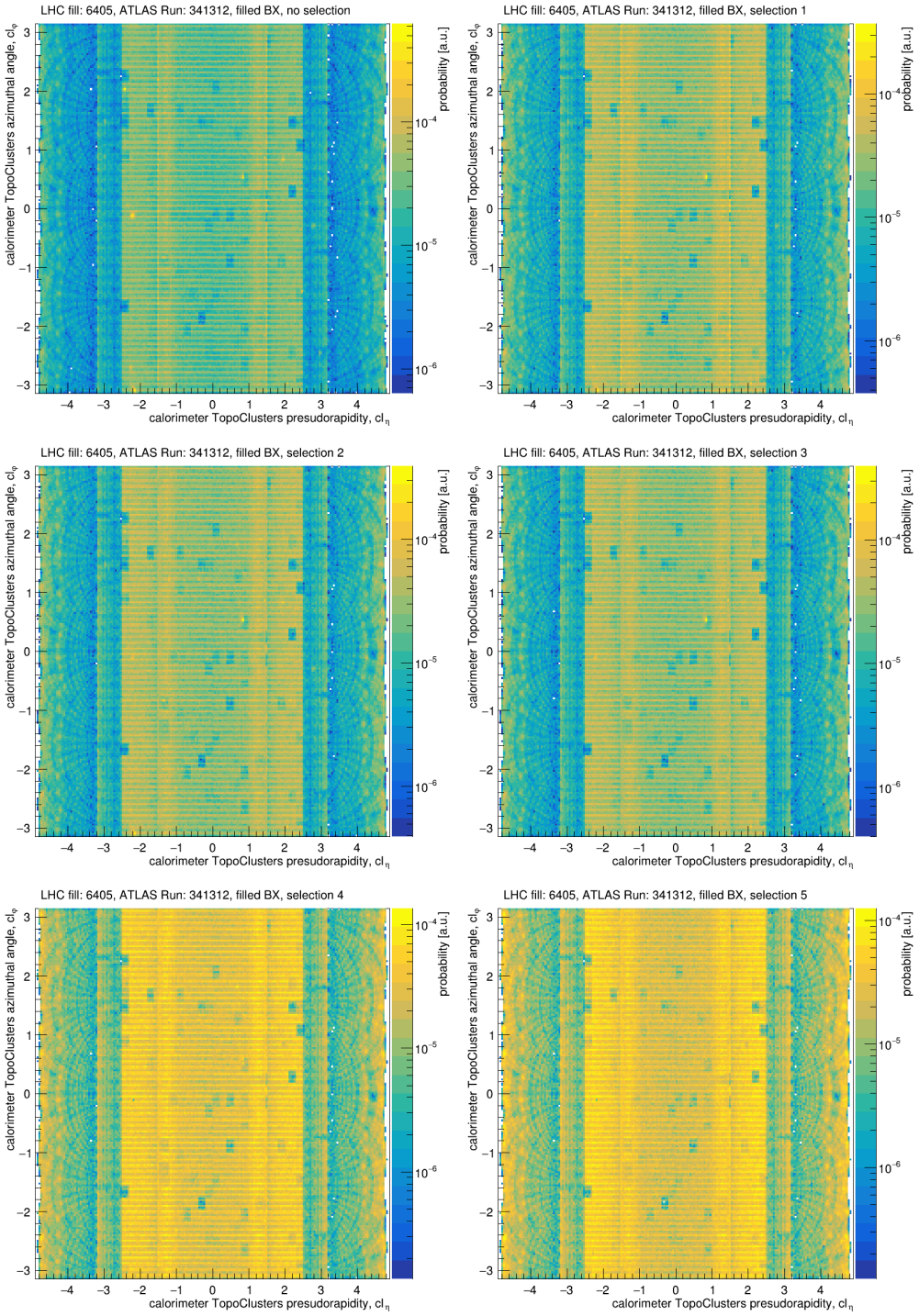


Figure 5.11: Clusters $\eta \times \phi$ map for colliding bunches. Selection 1 includes $|cl_t| < 10$ ns cut. Selection 2 includes $|cl_t| < 10$ ns and $cl_{EngFracMax}$ <0.8 cuts. Selection 3 includes $|cl_t| <$ 10 ns, $cl_{EngFracMax}$ <0.8 and $cl_{MAG} < 5500$ cuts. Selection 4 includes all previous cuts and the cut on the pixels with higher activity. Selection 5 includes all previous cuts and the requirement of $cl_E^{raw} > 250$ MeV.

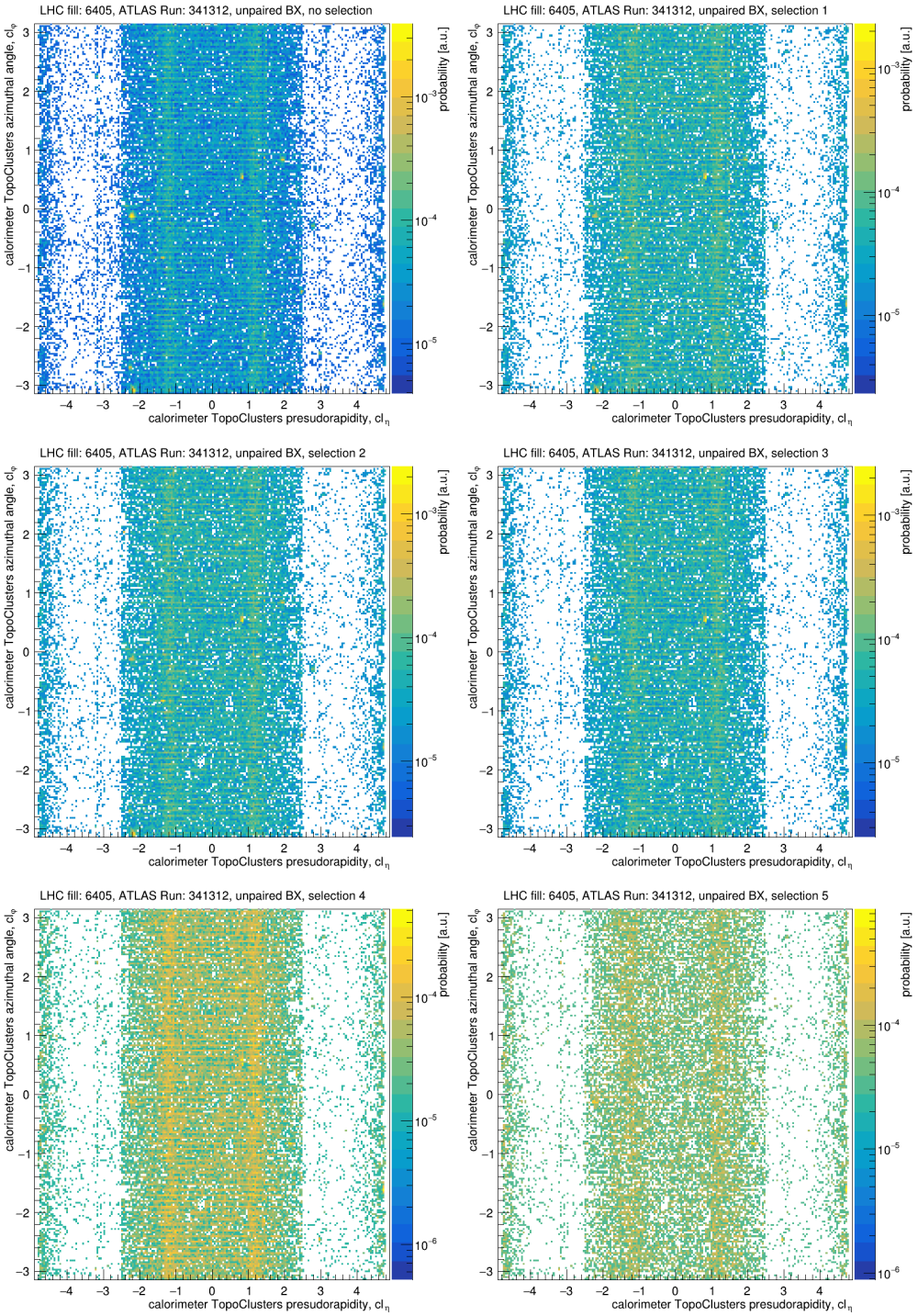
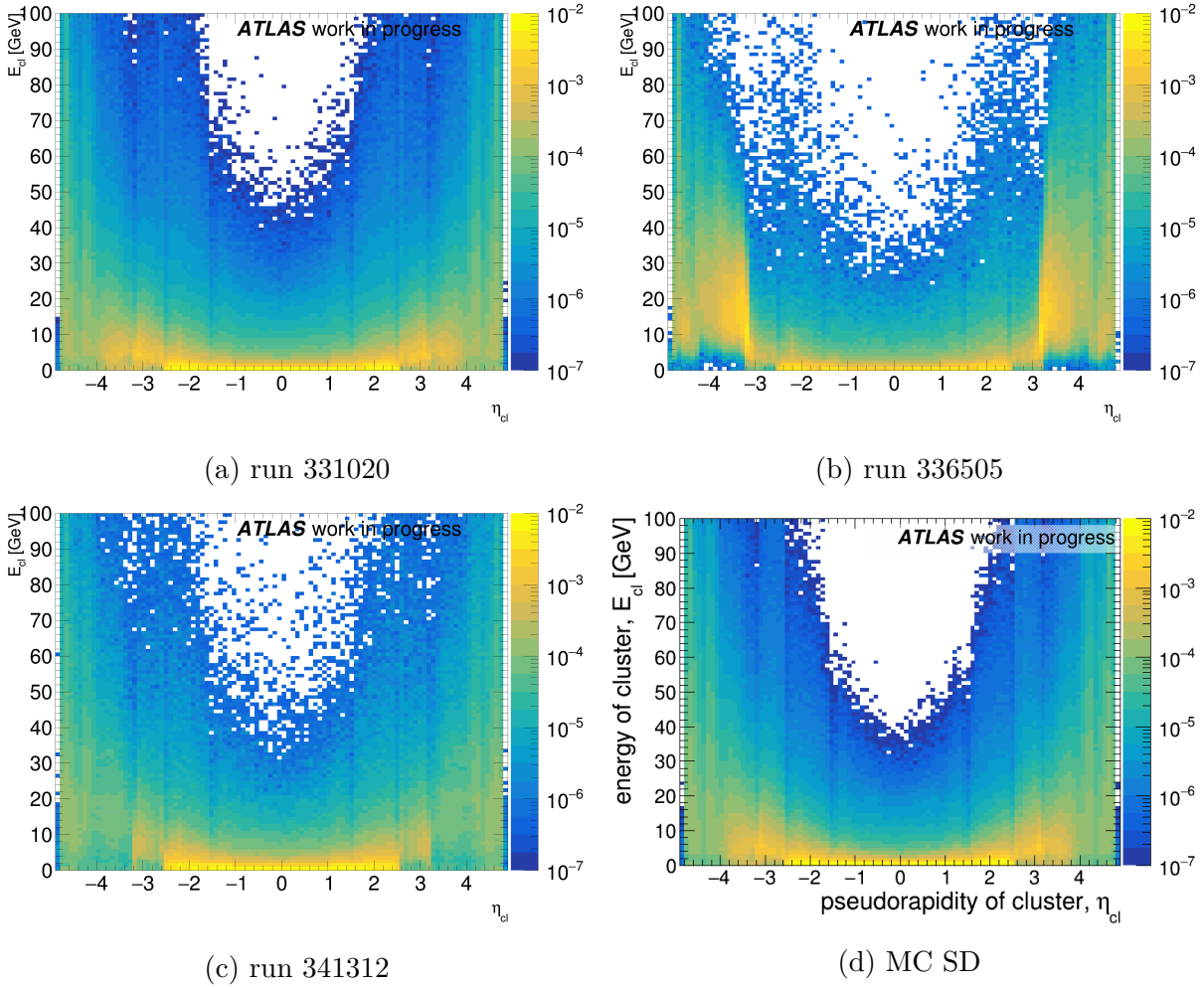


Figure 5.12: Clusters $\eta \times \phi$ map for unpaired bunches. Selection 0 – no cuts. Selection 1 includes $|cl_t| < 10$ ns cut. Selection 2 includes $|cl_t| < 10$ ns and $cl_{EngFracMax}$ <0.8 cuts. Selection 3 includes $|cl_t| < 10$ ns, $cl_{EngFracMax}$ <0.8 and $cl_{MAG} < 5500$ cuts. Selection 4 includes all previous cuts and the cut on the pixels with higher activity. Selection 5 adds the requirement of $cl_E^{raw} > 250$ MeV to all previous cuts.

high $|\eta|$ and relatively high energy is much greater than in the other runs and in the Monte Carlo sample. This effect was identified as coming from the incorrect hardware settings of the calorimeter in this run resulting in higher activity. It should be noted that there are also slight differences in the forward regions between runs 331020 and 341312. This observation is still not fully understood.



(a) run 331020

(b) run 336505

(c) run 341312

(d) MC SD

Figure 5.13: The cluster multiplicity as a function of the cluster pseudorapidity η_{cl} and energy (E_{cl}) for different runs and SD MC samples. All plots were normalised to unity.

5.3.4 Tracks

Charged particles travelling through the ATLAS tracker are expected to create signals in its detectors. Combinations of these signals should allow the reconstruction of particle trajectories. Some aspects of the track reconstruction needed for further analysis were studied using the MC simulated sample and are presented below.

The basic distributions of multiplicity, transverse momentum and pseudorapidity of tracks are shown in Figs. 5.14, 5.15 and 5.16, respectively. The analysis presented here was performed for all considered runs, but only examples of 331020, 341648 ($\mu = 1$ part) and 341648 ($\mu = 2$ part) are shown. Grey areas show the distribution of the generated (the truth level) charged particles having $p_T > 200$ MeV and $|\eta| < 2.5$. Such a selection reflects to a certain extent the response of the ATLAS tracker. The red points mark the distribution of tracks reconstructed from signals coming from the truth particles and the data distribution is shown with blue circles.

The reconstructed tracks were required to pass the selection criteria described in Section 3.5.2. The distributions are normalised to the number of events. The data events were accepted by the HLT_j20_L1AFP_A_OR_C_J12 trigger. In the MC case, the effective cross-section (including the generation filter) was multiplied by the luminosity collected during the given run (*cf.* Tab. 5.1). Only LBs passing the AFP GRL were considered. In addition, the trigger efficiency (*cf.* Section 5.3.2) was considered. In all cases, the proton was required to be reconstructed on the C side.

As can be seen from Fig. 5.14, the track multiplicity (red points) is shifted on average towards lower values w.r.t. the truth distribution (grey area). This is an expected consequence of the reconstruction. It can be observed that the data (blue points) are falling slower with increasing number of the charged particle tracks – this may be due to the presence of pile-up.

It is interesting to see if other basic MC track distributions, like the transverse momentum and pseudorapidity, reproduce the data distributions. The transverse momentum distribution is shown in Fig. 5.15. It should be noted that the ratio between the distributions of the p_T of a track reconstructed in MC and data is quite constant. The difference in normalisation may come from the pile-up (not present in MC) and a mixture of various types of events (eventually accepted by HLT_j20_L1AFP_A_OR_C_J12) compared to pure diffractive events in the Monte Carlo.

Neglecting the above-mentioned difference of normalisation, the shape of the track pseudorapidity spectrum in data is nicely described by the Monte Carlo – see Fig. 5.16. Even with a crude event selection (the trigger level), the differences in the shape are within 20%. A small left-right asymmetry (A-C; positive-negative η) is due to asymmetrical MC predictions (diffractive proton on side C).

Another interesting point is to look at the properties of tracks as a function of both: p_T and η. For the MC sample, this is shown in Fig. 5.17. The full spectrum of the generated charged particles is shown in the top left plot. To see the features for small and high transverse momenta a function – $\log_{10}(p_T)$ – is used. The effect of a simple selection (the truth level) requesting $p_T > 200$ MeV and $|\eta| < 2.5$ is shown in the top right plot. The reconstructed MC track distribution is shown in the bottom left panel. As can be seen, this distribution, in a sense originating from the top left one, has some interesting "features". For example, a large number of high-p_T tracks is reconstructed in certain regions of the

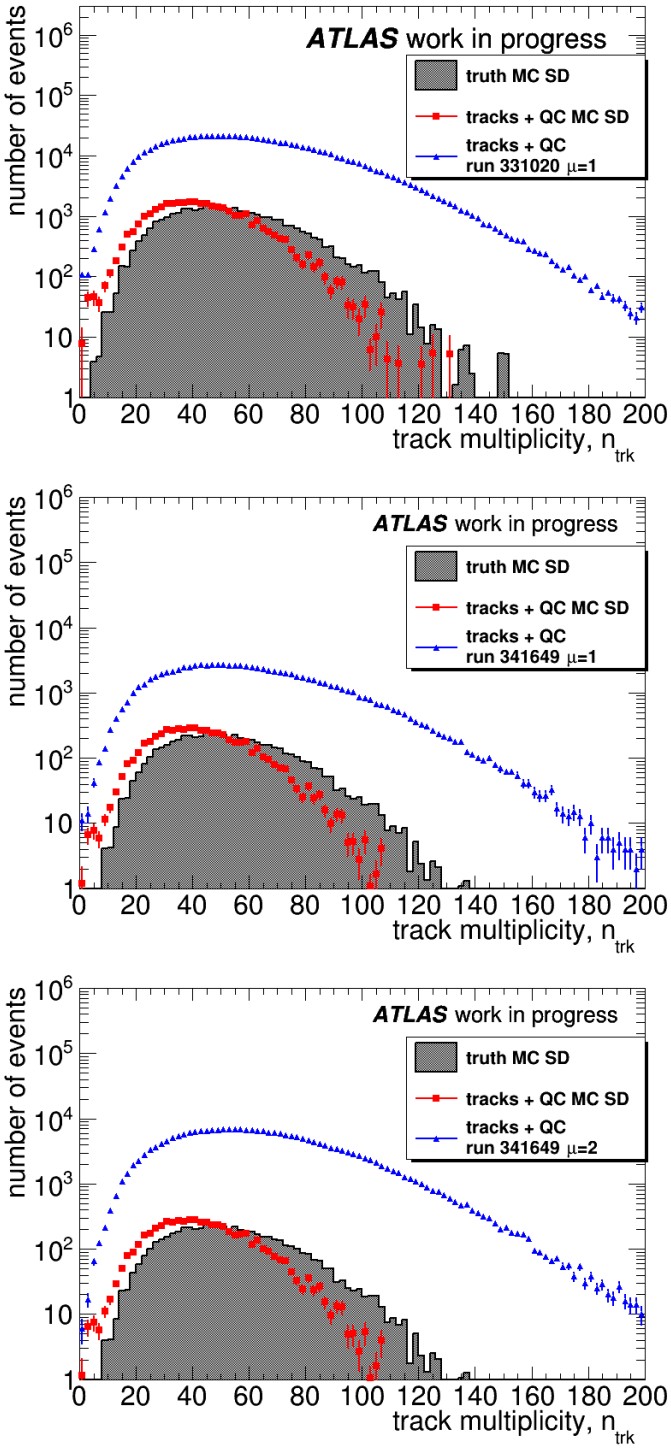


Figure 5.14: The multiplicity distribution of the truth particles (grey area), MC (red points) and data (blue points) tracks for run 331020 (**top**), 341649 ($\mu \sim 1.0$ part; **centre**) and 341649 ($\mu \sim 2.0$, **bottom**). Only events with proton reconstructed on side C were considered. See the text for the event selection criteria.

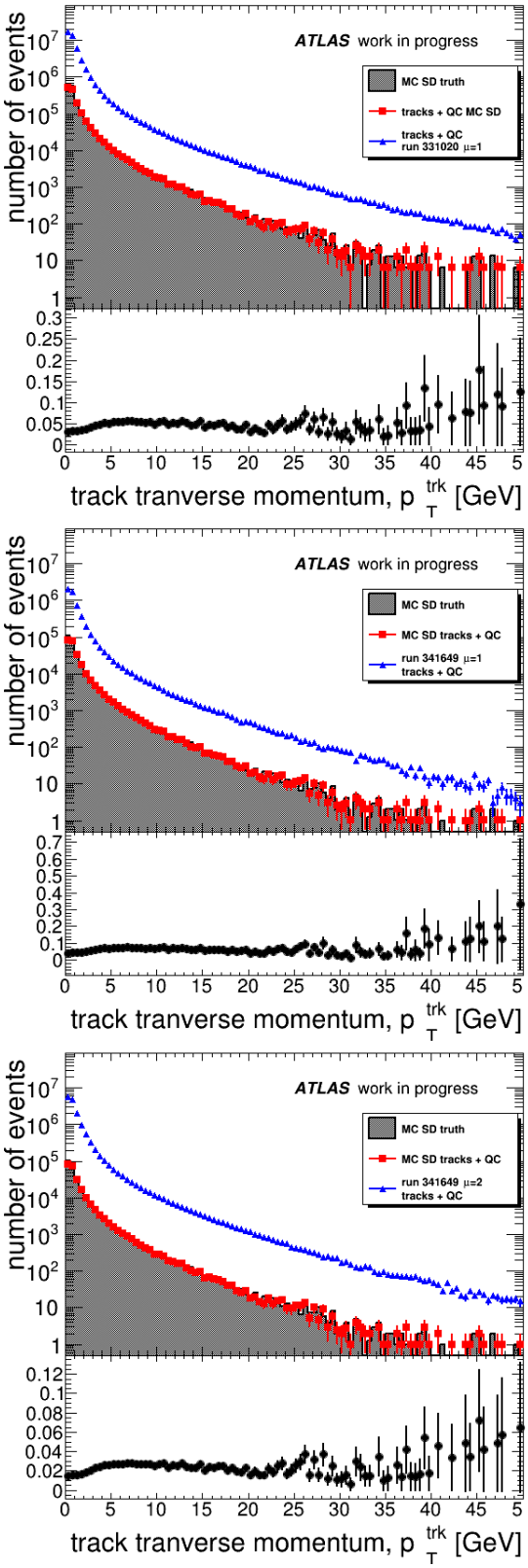


Figure 5.15: The p_T distributions of the truth particles – grey area, MC (red points) and data (blue points) tracks for run 331020 (**top**), 341649 ($\mu \sim 1.0$ part; **centre**) and 341649 ($\mu \sim 2.0$; **bottom**). For the reconstructed tracks the quality selection was applied. In all samples, only events with a proton reconstructed on side C were considered. The bottom panels of the plots show the ratio of the Monte Carlo to data distributions.

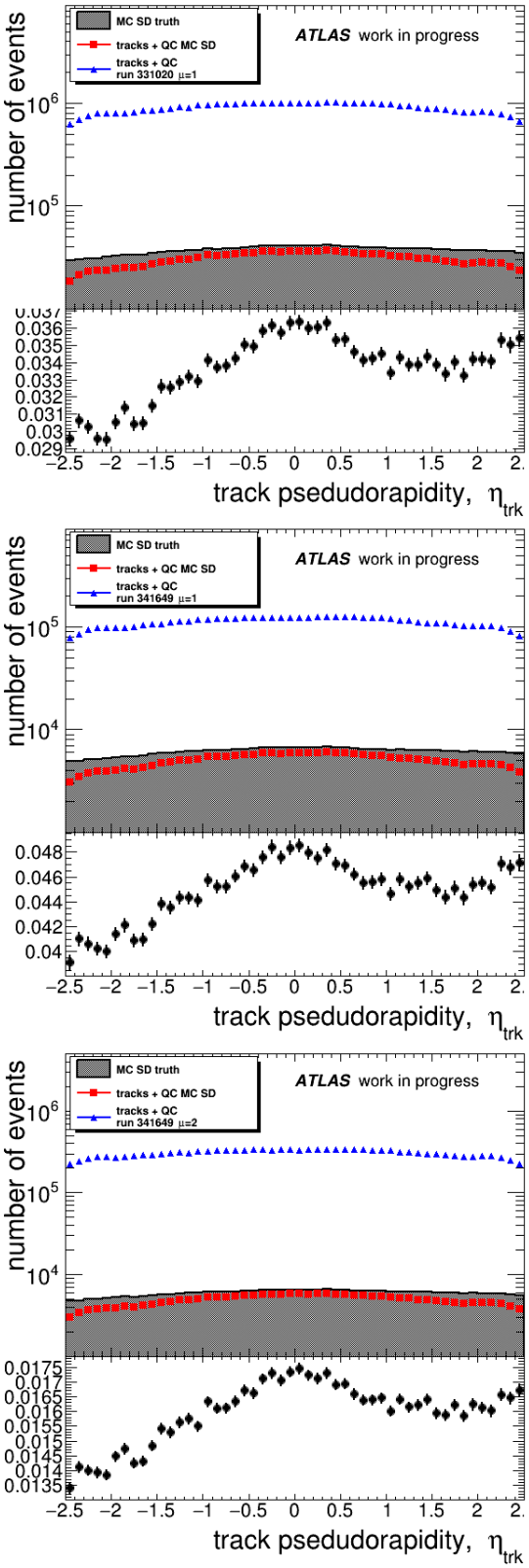


Figure 5.16: The pseudorapidity distribution of the truth particles (grey area), MC (red points) and data (blue points) tracks for run 331020 (**top**), 341649 ($\mu \sim 1.0$ part; **centre**) and 341649 ($\mu \sim 2.0$, **bottom**). For the reconstructed tracks the quality selection was applied. In all samples, only events with a proton reconstructed on side C were considered. The bottom panels of the plots show the ratio of the Monte Carlo to data distributions.

pseudorapidity ($|\eta| \sim 1.8$, $|\eta| \sim 1.0$, $|\eta| \sim 0.4$). These features, mostly the reconstruction artefacts, are removed by applying the quality selection criteria (*cf.* Sec. 3.5.2). The effect of the quality cleaning is visible in the bottom right plot. A comparison of the top right and bottom right plots confirms, to no surprise, the correctness of the reconstruction algorithms and selection criteria.

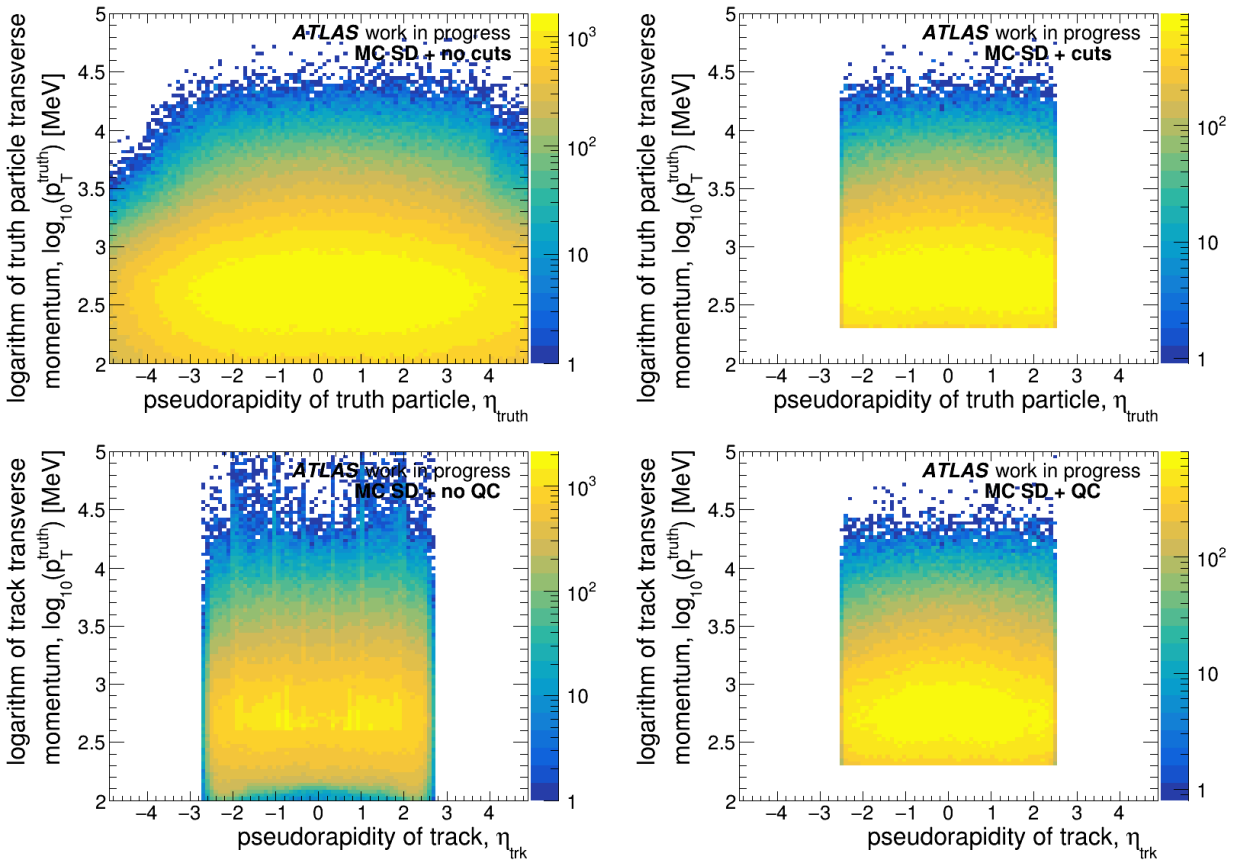


Figure 5.17: SD MC sample. The $\eta \times p_T$ map for the truth particles/tracks (**top**/**bottom**) without/with the quality cuts (**left**/**right**).

The same check was applied for all data samples, and the results for run 331020 are shown in Fig. 5.18. The reconstruction features similar to those observed in the MC are visible (see left plot). Again, they disappear after applying the quality selection criteria (right figure).

Finally, it is interesting to determine the track reconstruction efficiency. For each charged particle having $p_T > 200$ MeV and $|\eta| < 2.5$ a region in the (η, ϕ) plane centred at the expected position of the hit ($\sqrt{\eta^2 + \phi^2} < 0.1$) was checked for the presence of the reconstructed track. The reconstruction efficiency as a function of the pseudorapidity and transverse momentum of the charged truth particle is shown in Fig. 5.19 left and right, respectively.

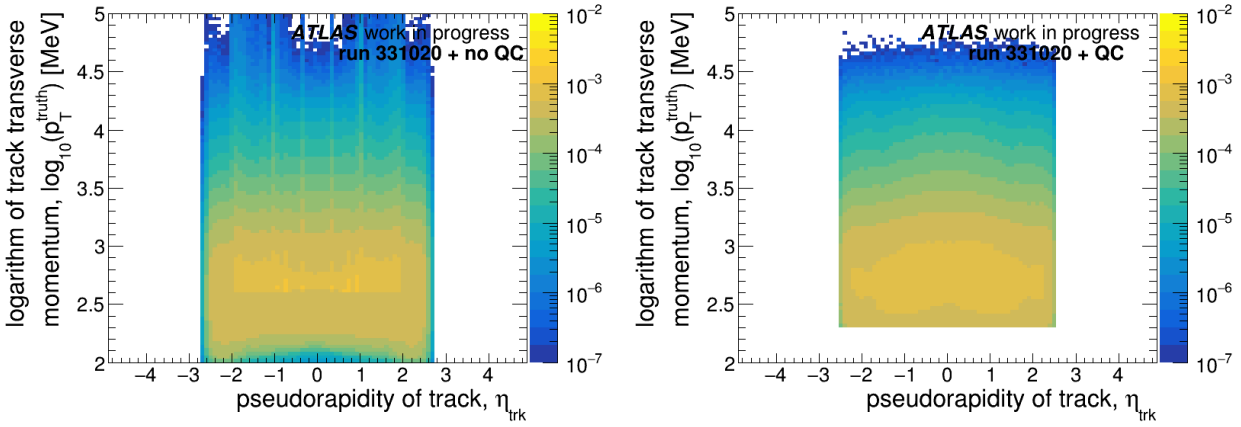


Figure 5.18: Run 331020: the $\eta \times p_T$ map of tracks without (**left**) and with (**right**) quality cuts.

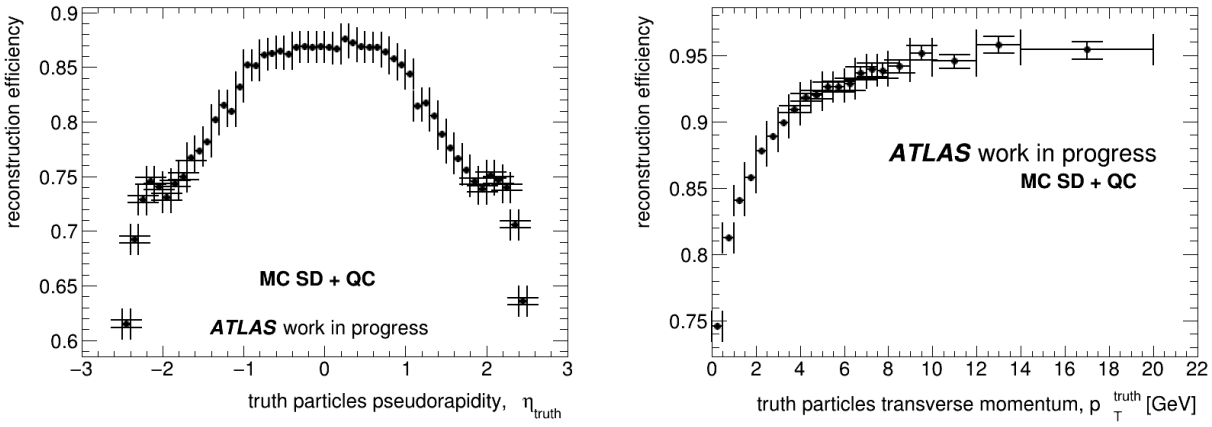


Figure 5.19: The efficiency of the track reconstruction in the SD MC sample as a function of the pseudorapidity (**left**) and transverse momentum (**right**) of the charged truth particle.

The efficiency drops with increasing $|\eta|$ of a track. The highest efficiency is obtained for the central tracks. It also increases with increasing p_T of a particle. As the efficiency depends on both the pseudorapidity and transverse momentum of a track, the two-dimensional track reconstruction efficiency is shown in Fig. 5.20. The efficiency above 90% is observed in the central region ($|\eta| \lesssim 1$) for particles with $p_T \gtrsim 1$ GeV.

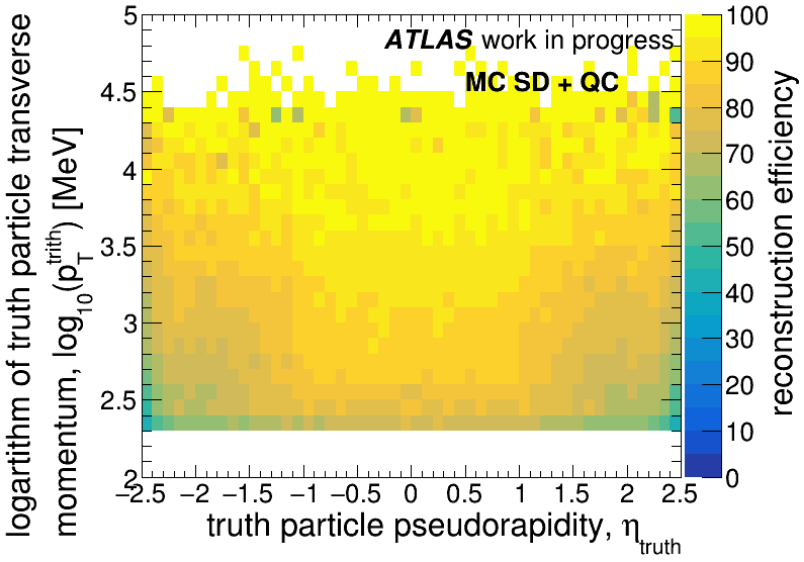


Figure 5.20: Efficiency of the track reconstruction in the SD MC sample as a function of the pseudorapidity and transverse momentum of the charged truth particle.

5.3.5 Vertex

Identification of the interaction vertex is another important factor in the analysis. In this thesis, a standard ATLAS reconstruction will be used. The distribution of the vertex multiplicity is shown in Fig. 5.21, left plots. Runs 331020 and 341649 (separately for $\mu \sim$1.0 and 2.0 part) were taken as an example.

In the ATLAS reconstruction nomenclature certain names of vertices, reflecting their properties, are used. In this analysis three of them are mentioned (see Fig. 5.21 right), namely:

- dummy – exactly one present in each event for the technical purposes; note: not considered in vertex multiplicity shown in Fig. 5.21 left,
- primary – the vertex which is pointed by the tracks having the highest p_T sum; usually corresponding to the "hard" interaction,
- pile-up – other vertices reconstructed in the vicinity of the beam spot.

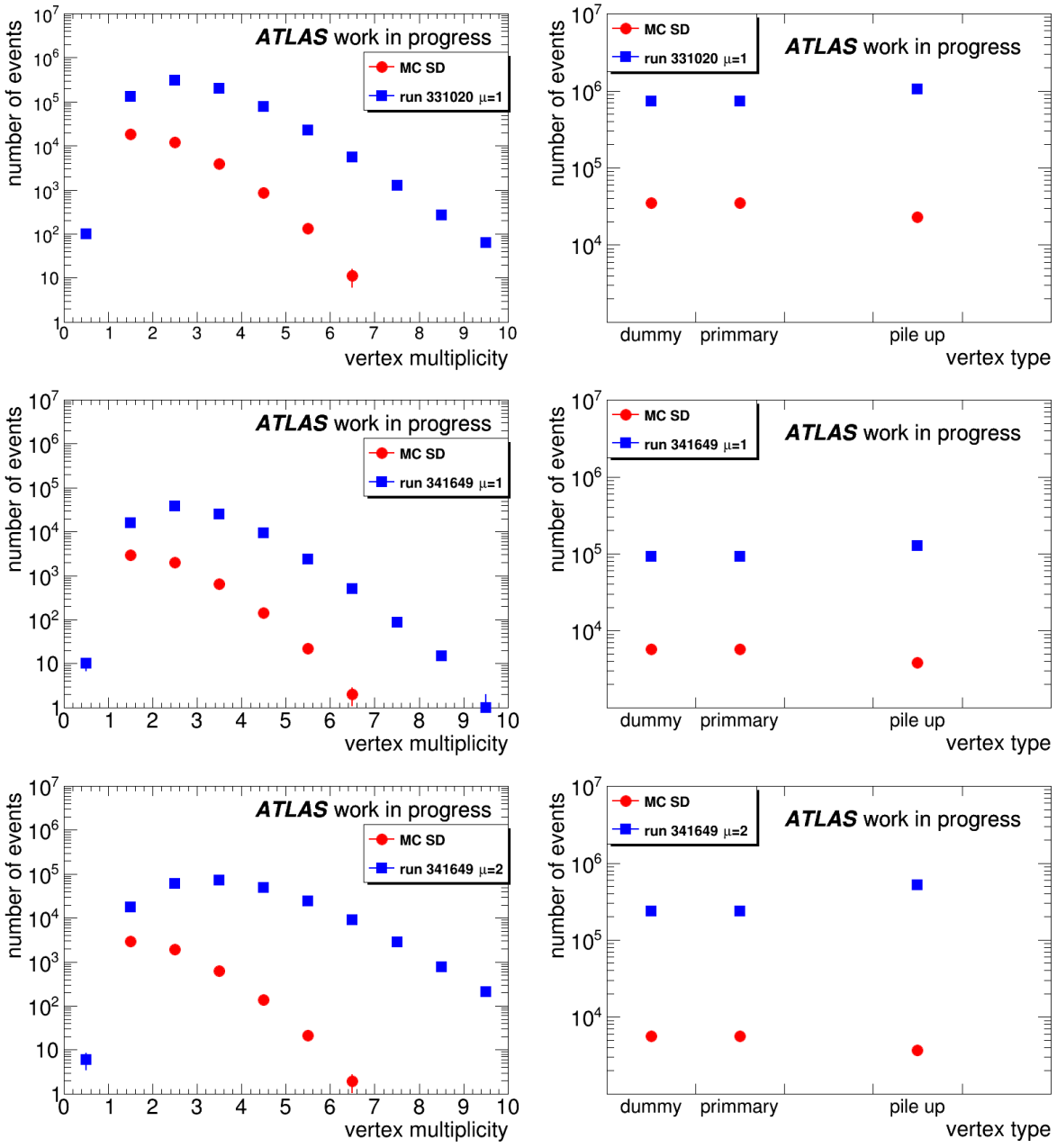


Figure 5.21: A comparison of the vertex properties for the SD MC and data samples: the vertex multiplicity (**left**) and the vertex type (**right**). From **top** to **bottom**: run 331020, run 341649 with μ ~1.0 and 2.0.

Interestingly, even for the MC sample (the blue points) the number of the reconstructed vertices is not exactly equal to 1, despite the fact that it does not contain pile-up events. These additional vertices are a feature of the ATLAS reconstruction algorithms and are visible in around 43% of the cases (the exact value can be obtained from Table 5.5).

In the data case (the red points), the vertex multiplicity profile is slightly different. It peaks at 2 for $\mu \sim 1$ data (top and middle plots) and at 3 for $\mu \sim 2$ (bottom plot). It is due to the presence of pile-up as clearly visible in the vertex type distributions where, contrary to MC, there are a lot of data events containing the pile-up vertex type. It should be noted that there are only a few events without a reconstructed vertex.

Another interesting distribution is the longitudinal (z) position of the vertex reflecting the beam profile. It is shown in Fig. 5.22 for the Monte Carlo (red points) and data (blue points; run 331020 taken as an example). In the case of all vertices (left), the distribution is double-Gaussian-like. If only the primary vertex is considered it becomes Gaussian-like. The data profile has a smaller width (σ=38.99) than the MC predicted (σ=52.95), however, it should not affect the present analysis. This observation may be considered for future MC generation campaigns.

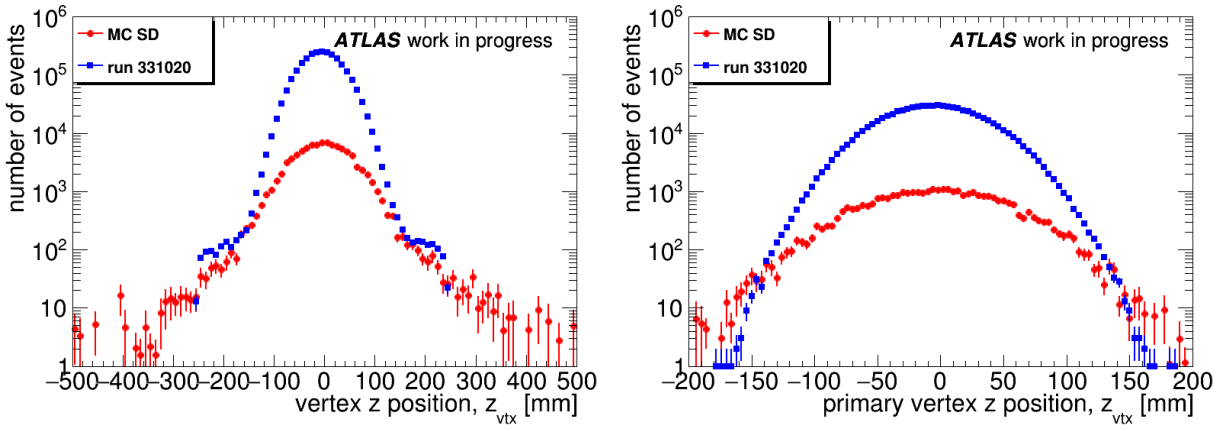


Figure 5.22: Longitudinal position of all (**left**) and primary (**right**) vertices. The red (blue) points are for the Monte Carlo (data; run 331020) samples.

5.3.6 Jets

A brief discussion of the jet reconstruction can be found in Section 3.5.3. Here, the basic jet properties visible in MC and data samples will be discussed.

Figure 5.23 shows the jet multiplicity distribution. The grey area represents the jets reconstructed from using the truth (generated, stable) particles. The blue (red) points mark the distributions obtained for the reconstructed data (MC) events. Top plots were created without selection criteria other than generation settings (MC) and trigger (data). Due to the event preselection at the generation level, the smallest multiplicity of the truth jets is

2. However, in the majority of cases, more than 2 jets are expected. In the reconstruction, events with only one reconstructed jet may appear in a few per mille of cases. However, the majority of cases will contain exactly 2 jets. Due to the applied trigger, events with one jet may also appear in the data. However, events with a multiplicity of 2-3 constitute the majority. Events with a higher multiplicity of jets are, in some cases, due to the bad-quality jets.

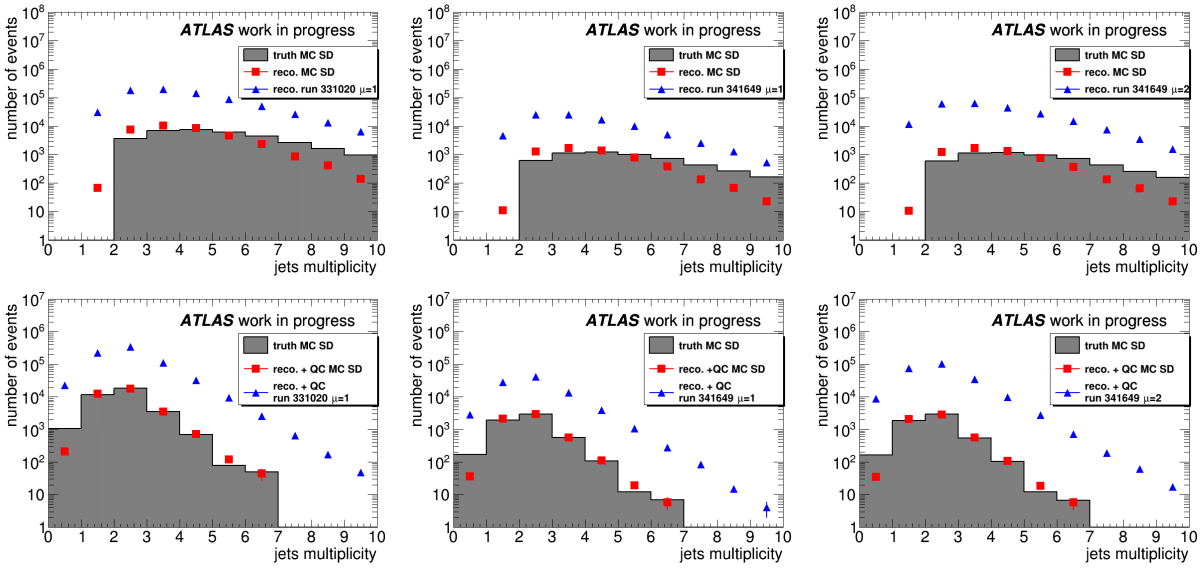


Figure 5.23: Multiplicity of jets: the SD MC truth (grey), reconstructed in the SD MC (red) and in the data (blue) samples. From **left** to **right**: run 331020, 341649 with μ ∼1.0 and 2.0. **Top** plots were constructed without the selection criteria other than the generation settings (MC) and the trigger (data). **Bottom** plots were made after applying the cuts requesting at least two good-quality jets. In the MC truth case, a good quality means $p_T > 20$ GeV for the sub-leading and $|\eta| < 4.5$ for both jets.

Indeed, by application of the jet quality criteria (see Sec. 3.5.3), the multiplicity is shifted towards smaller values – see Fig. 5.23 (bottom). Please note that in the case of the truth particles, good quality means $p_T > 20$ GeV and $|\eta| < 4.5$ for both jets.

Properties of leading (*i.e.* having the highest p_T) and next-to-leading (or the 2^{nd} leading; sub-leading) jets were checked. The distribution of the transverse momentum is presented in Fig. 5.24 and of the pseudorapidity in Fig. 5.25. Run 341649 ($\mu = 1$ part) was taken as an example. The left plots are for the leading whereas the right ones are for sub-leading jets.

The distribution of the jet transverse momentum is expected to be steeply falling. This is indeed true for the leading jet with $p_T > 35$ GeV. A smaller number of events within the region between 20 and 35 GeV is due to the requirement that the 2^{nd} leading jet $p_T > 20$ GeV. It should be noted that this effect is present, and follows the same curve, in both the MC and data. Such a "turnover" is also visible in the data for the 2^{nd} leading jets

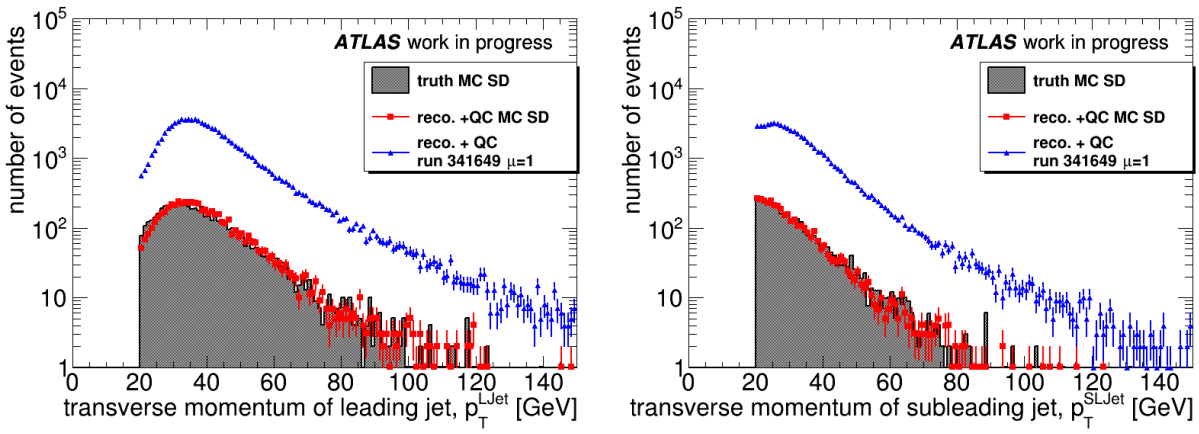


Figure 5.24: The transverse momentum distribution of the leading (**left**) and sub-leading jet (**right**) in run 341649 ($\mu = 1$ part). The distributions of jets reconstructed using the truth particles (MC) are shown in grey, the reconstructed objects (MC) in red and the data in blue. The reconstructed jets passed the quality selection. The truth jets were required to have $p_T > 20$ GeV and $|\eta| < 4.5$. A forward proton was requested on side C.

having $p_T \lesssim 25$ GeV. This effect is connected to the (jet) trigger efficiency (*cf.* Sec. 5.3.2).

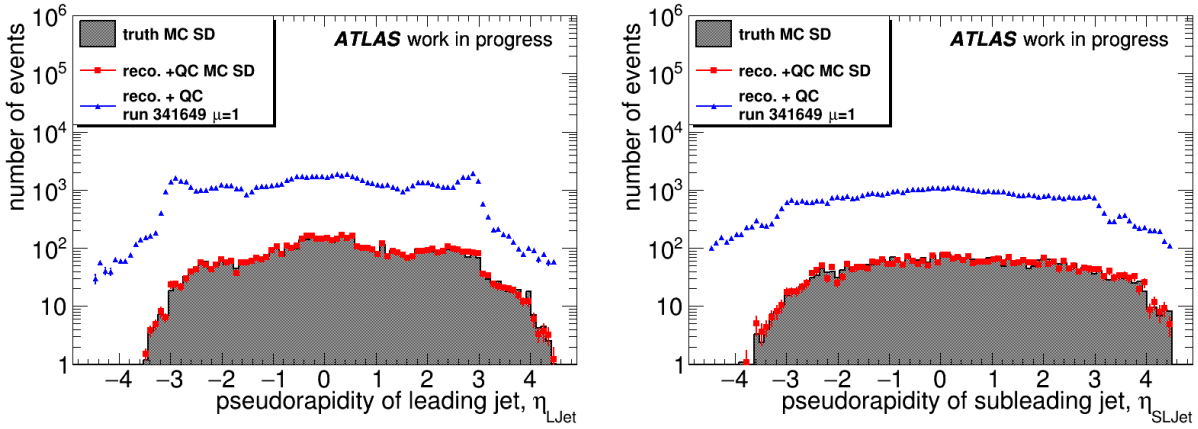


Figure 5.25: The pseudorapidity distribution of the leading (**left**) and sub-leading jet (**right**) in run 341649 ($\mu = 1$ part). The distributions of jets reconstructed using the truth particles (MC) are shown in grey, the reconstructed objects (MC) in red and the data in blue. The reconstructed jets passed the quality selection. The truth jets were required to have $p_T > 20$ GeV and $|\eta| < 4.5$. A forward proton was requested on side C.

The pseudorapidity distribution of the leading and sub-leading jet for run 341649 ($\mu = 1$ part) is presented in Fig. 5.25 left and right, respectively. Also, the Monte Carlo predictions are plotted in the figure. Despite the events pass the quality criteria, local maxima in the "calorimeter transition region" ($2.6 < |\eta| < 3.2$) are visible. This effect is not fully understood yet. MC sample also lack in events in region η_{LJet} <-3.5 and η_{SLJet} <-4. A

small left-right (positive-negative η) asymmetry in the case of the MC is due to the proton being tagged on side C.

5.3.7 Forward Protons

Finally, the basic properties of the scattered protons were investigated. In the reconstruction, the corrections due to the local and global alignment [69] were applied. The hit pattern visible in the data is shown in Fig. 5.26, run 331020 was taken as an example. In the case of the Monte Carlo, the truth protons were transported by the FPTracker tool (embedded into the AFPToolbox, see Sec. 4.4) using Run 2, $\beta^* = 0.6$ m optics with opened collimators. For both MC and data, the AFP C NEAR station was taken as an example.

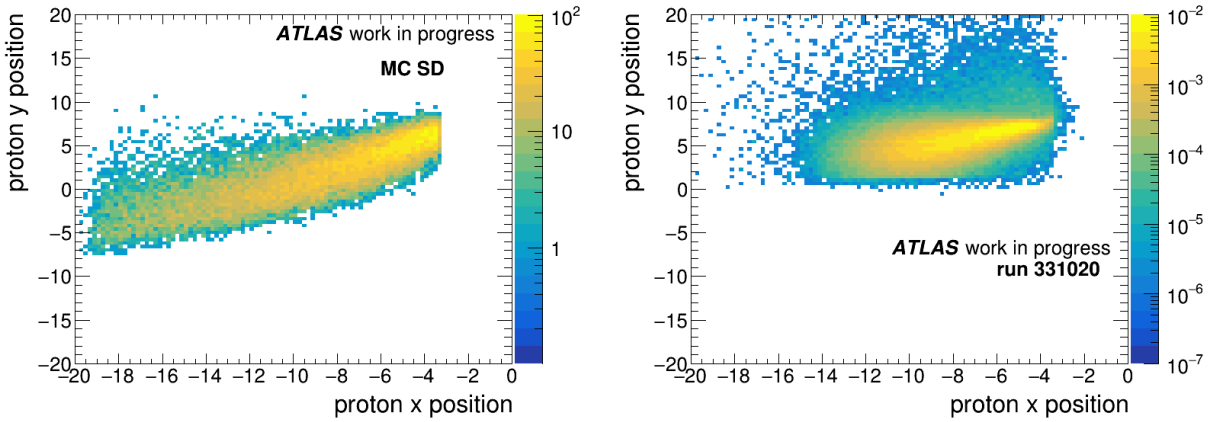


Figure 5.26: The hit pattern in the AFP C NEAR station for the SD MC (**left**) and run 331020 (**right**).

In the 2017 data-taking campaign, the crossing angle had a phase of 90^0 (pointing up at the IP) which resulted, due to the LHC optics, in a diffractive pattern going "down" as protons were further from the beam (in x). In the case of data, the effect of the collimators removing the protons with high ξ value ($x \lesssim -14$ mm) is visible. The geometry of the AFP SiT detector results in a lack of events with $y < 0$.

The multiplicity of tracks reconstructed in the AFP is shown in Fig. 5.27. Only events with a proton on side C were selected, thus in the case of the Monte Carlo, exactly one proton was expected on that side (*cf.* right figure). This is due to the lack of a pile-up and the fact the particle showers in the AFP detector were not simulated. In the data case, events with a single proton are prevailing. Two reconstructed protons appear in ∼3% of cases. A per mille fraction of events has higher proton multiplicities – an effect of showers created either downstream of the detector or in the station material.

Interestingly, in the MC sample, there is a non-zero fraction of events, $\sim 1‰$, containing a track (proton) on side A. One should recall that the sample was generated requiring a diffractively scattered proton on side C. These, side A, protons are produced in the

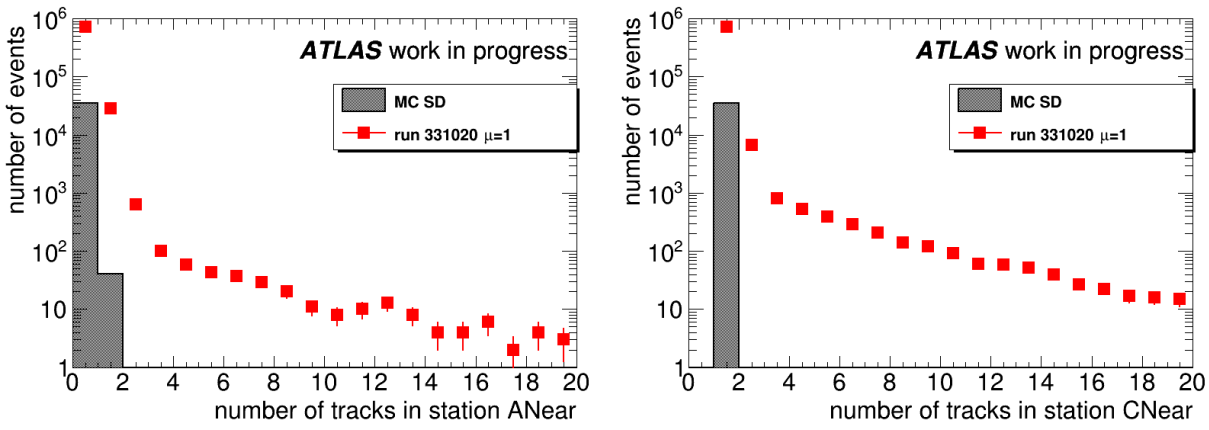


Figure 5.27: The multiplicity distribution of the tracks reconstructed by the AFPToolBox for the NEAR station of side A (**left**) and C (**right**). The grey histogram represents the SD MC predictions and the red points mark the run 331020 distribution.

hadronisation process. For the data sample such events, happening in about 5% of cases, may also come from the Double Pomeron Exchange jet production or appear as a result of a pile-up or the beam halo.

5.4 Event Selection

In the previous section properties of the data sample to be used for the Single Diffractive jet analysis were described. In addition, the quality criteria tailored for the diffractive analysis were proposed. From now on, they will be applied by default. It should be noted that the cuts on the quality of jets and protons influence the number of the accepted events whereas the cut on tracks and clusters only affects the distribution of $\xi_{trk/cl}$.

Having basic quality selection criteria established, a signal selection can be proposed. The data triggered by the AFP jet triggers will contain both the signal and background events. The main background in this analysis is expected to be due to the non-diffractive jets and an intact proton originating from the pile-up (see Fig. 5.2). In order to reduce the background-to-signal ratio, a few selection steps need to be done. It should be noted that the result of the cut-flow described below is shown for run 331020 as an example. However, the distributions were checked for all the considered samples.

5.4.1 Single Vertex Requirement

Following indications of [67], a selection based on the vertices reconstructed by ATLAS was applied. Only events having exactly one vertex, without any additional (pile-up) ones were accepted for further analysis. It should be noted that this selection impacts differently the Monte Carlo (generated w/o pile-up) and data samples. In the case of the SD JJ Monte

Carlo, the fraction of events with exactly one vertex is around 57% (*cf.* Table 5.5 columns 3 and 4). For data, it depends on the pile-up and is of about 73.6% for the run with $\mu = 0.05$, 18.5% for runs with $\mu = 1$ and around 7.5% for runs with $\mu = 2$.

The multiplicity of tracks associated with the primary vertex is shown in Fig. 5.28 (left). The data distribution shows a large tail. It is due to the presence of the pile-up allowing the merging of the neighbouring tracks to the jet vertex. From the right plot, one can conclude that not all reconstructed tracks will be assigned to the primary vertex.

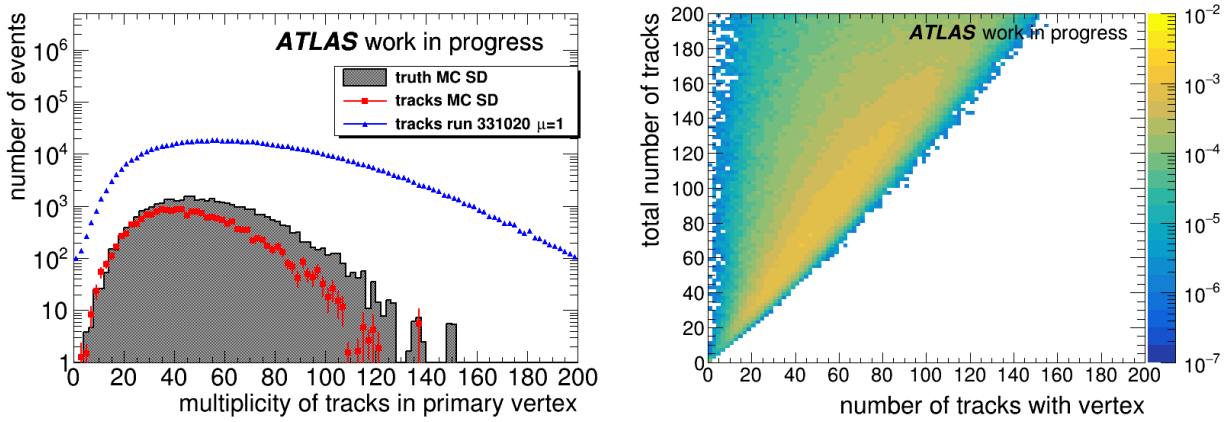


Figure 5.28: **Left**: the number of tracks associated to the primary vertex for MC (red points) and data (run 331020; blue points). The grey area represents the generated charged particles with $p_T > 200$ MeV and $|\eta| < 2.5$. **Right**: the correlation between the number of tracks associated to the primary vertex to the total number of tracks present in event (run 331020).

5.4.2 Good Run List

In the next step, the Good Run List was applied removing LumiBlocks in which some issues were spotted. Since not all 2017 low-μ runs are listed in the "standard, ATLAS GRL" a special AFP GRL was prepared [68]. In addition to the standard ATLAS criteria, the AFP-related requirements were considered. For obvious reasons, the GRL does not have to be applied to the MC samples. For the data, it is worth preparing separate lists for events with a proton on side A or C. The number of events after passing the GRL selection is listed in column 5 of Table 5.5. In all runs except 331020, the application of GRL results in a similar event reduction on both sides. For run 331020, side A was quite problematic – only about 10% of events compared to side C was accepted for further analyses.

5.4.3 Good Quality Jets

Next, the presence of at least two good-quality jets (see Sec. 5.3.6) was required. As described, the jet calibration was applied for both MC and data. Due to the efficiency of the jet reconstruction algorithms, the ATLAS recommendation of using jets with transverse momentum of at least 20 GeV was considered. In addition, the jets were required to be within $|\eta| < 4.5$ in order to avoid a part of the jets escaping the acceptance of the ATLAS detector[5]. In order to reject the out-of-time background, the value of the jet's absolute time was required to be less than 12.5 ns. Lastly, whenever possible due to the acceptance of the tracker, the jets were required to come from the same vertex. This condition was satisfied by using the Jet Vertex Tagger (JVT) tool, which provides the probability that a jet is coming from the primary vertex. Following the ATLAS recommendations, only events with jets having JVT>0.5 (at $|\eta| < 2.4$) were considered.

5.4.4 Single Reconstructed Proton

The proton object is reconstructed using the AFPToolBox. This step already implies some selection (see Sec. 4.4). In addition, the recommendation of *the AFP Combined Performance* group is to use the protons with the track x-coordinate in the NEAR station smaller than -3.5 mm and to require the proton energy loss to be within $0.035 < \xi < 0.08$. These criteria are to remove the "edge effects" of the AFP Silicon Tracker. A proton passing such a selection will be called a "good quality" one. In this analysis, exactly one such proton is required to be reconstructed without any other protons of a good or bad quality.

5.4.5 Summary

The number of events left after a given selection is listed in Table 5.5. "All" denotes the number of events: accepted by the HLT_j20_L1AFP_A_OR_C_J12 trigger (after the prescale, see Tab. 5.4) in case of data or generated in case of MC. The number of events left after the single vertex requirement is presented in column 4. The next column shows the situation after the GRL application, separately for a proton on sides A and C. Columns 6 and 7 list events accepted after the request of two good quality jets with p_T of at least 20 GeV and exactly one proton on a given side, respectively. The last column shows the number of events after an additional selection of activity in the forward region, which will be discussed later.

[5]The ATLAS detector acceptance is $|\eta| < 4.9$ and the (anti-k_T) radius of the jet is 0.4.

Table 5.5: Number of events in Monte Carlo and data samples after application of a given selection requirement. The data events were triggered by $HLT\_j20\_L1AFP\_A\_OR\_C\_J12$. The GRL was not applied to the Monte Carlo sample. Only run 331020 had a cut on 10 clusters instead of 5, and in run 336505 a cut on the forward clusters was not possible.

Sample	**Consecutive Cut**						
	Side	**All**	**1 Vertex**	**GRL**	**2 Jets**	**1 Proton**	**5/10 Clusters**
SD MC	A	239499	136390	136390	4825	2061	394
	C					2654	1130
331020	A	1952990	357983	35404	23810	4447	693
	C			321605	217344	86869	17066
336505	A	41908	30857	22712	8536	658	-
	C			30784	11602	4262	-
341294	A	116385	7124	6977	4651	1122	406
	C			6977	4651	1214	408
341312	A	940071	70859	70717	47080	3111	1083
	C			70710	47076	19056	6208
341419	A	1649173	123268	123124	81825	5436	1879
	C			123003	81745	32564	10711
341534	A	2864217	246297	212460	143152	10453	2805
	C			244662	164771	63254	17367
341615	A	1650721	130234	130114	86375	6217	2164
	C			130114	86375	34642	11398
341649	A	236404	44633	39653	26460	2011	759
$\mu \sim 1.0$	C			39653	26460	10533	3690
341649	A	354020	27503	26387	17611	1315	415
$\mu \sim 2.0$	C			26124	17439	6943	2192

5.5 Initial Results

The selection chain described in the previous Section was applied to all data samples. It is interesting to see what fraction of the SD JJ is predicted by MC and visible in the data. In order to do this, after the event selection, the Monte Carlo was weighted in the following way:

$$W_{MC} = \frac{L \cdot \sigma_{gen} \cdot \varepsilon_{gen} \cdot vtx_{frac} \cdot PS \cdot \varepsilon_{jet}(p_T, \eta) \cdot \varepsilon_p(\xi)}{N_{gen}^{MC}}, \text{where :}$$

- N_{gen}^{MC} is the number of events in the generated MC sample,
- L is the luminosity collected during a given run after consideration of GRL; the distinction is made for proton on sides A and C (see Table 5.2),
- σ_{gen} is the cross-section reported by the generator,
- ε_{gen} is generation efficiency (*cf.* Section 5.2),
- vtx_{frac} is the correction reflecting the fact that the MC was generated w/o pile-up which makes a rejection due to the single vertex requirement different for the data and MC – *cf.* the ratio of col. 3 and 4 in Tab. 5.5:

$$vtx_{frac} = \frac{N_{data}^{1vtx}/N_{data}^{all}}{N_{MC}^{1vtx}/N_{MC}^{all}}$$

  ,
- PS is the correction due to the prescale (see Tab. 5.4),
- $\varepsilon_{jet}(p_T, \eta)$ is a correction due to the trigger efficiency (dependent on p_T and η of jet, see Section 5.3.2),
- $\varepsilon_p(\xi)$ reflects the "proton part" of the trigger efficiency (ξ dependent; see Sec. 5.3.2).

The transverse momentum of the leading jet for runs 331020 and 341649 ($\mu \sim$1.0 and $\mu \sim$2.0 part) after the selection described in the previous section is shown in Fig. 5.29. The Monte Carlo is scaled as described above. Clearly, the applied selection results in much more events visible in the data than expected from the MC. It is worth noting that the shape of the data and MC distributions is similar.

It is interesting to study the kinematic properties of the selected events. In particular, the distribution of the proton x-position in the AFP should allow the distinction between diffractive and non-diffractive events when plotted as a function of the energy deposited in the calorimeter (ξ_{cl}, see Tab. 5.3). Such initial studies were done for the pilot runs from 2016 when only one side (C) of the AFP was installed – see the left plot in Fig. 5.30. The similar distribution was created for the 2017 data, with run 331020 shown in Fig. 5.30

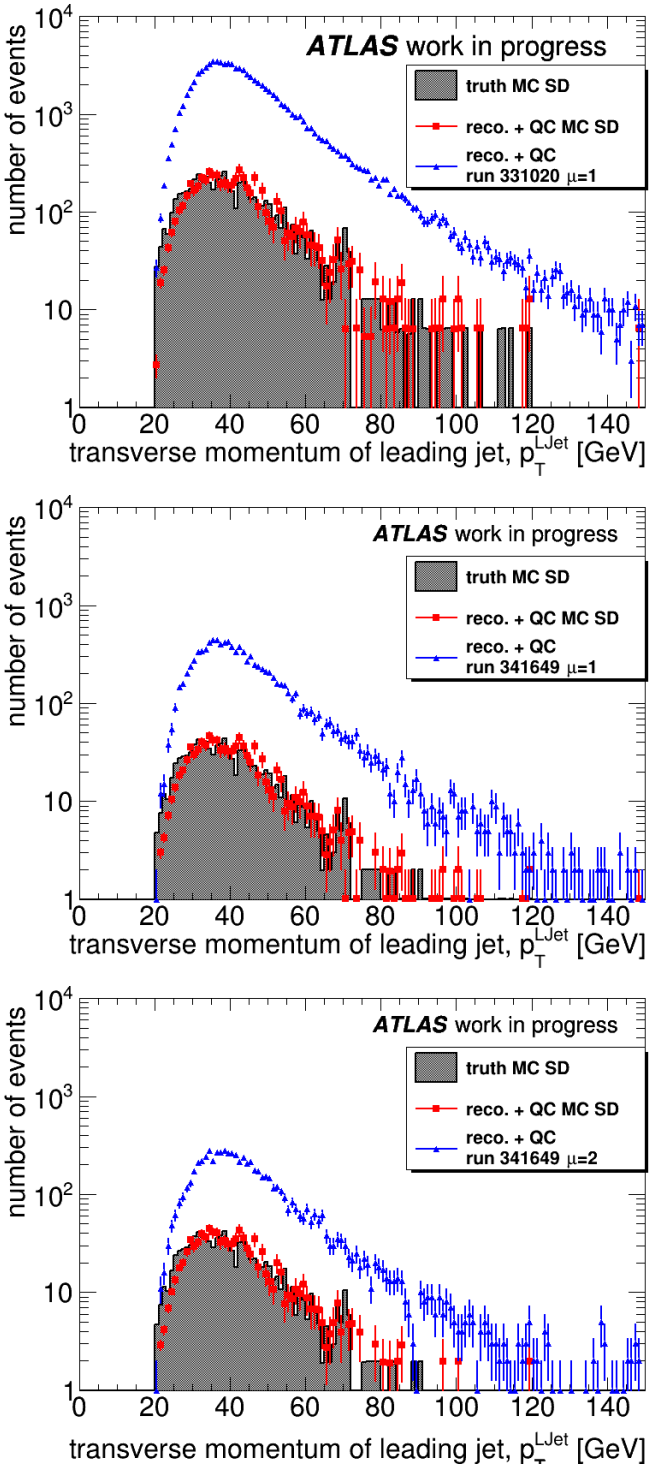


Figure 5.29: The transverse momentum of the leading jet for run 331020 (**top**), 341649 (μ ~1.0; **middle**) and 341649 (μ ~2.0; **bottom**) after the selection including the trigger, GRL, single vertex, two good quality jets and one proton on the C side.

(right) as an example. It should be noted that on y axis of the right plot the x-pixel ID is plotted, whereas the right one shows reconstructed proton x position. In order to compare the distributions one should flip vertically the right distribution.

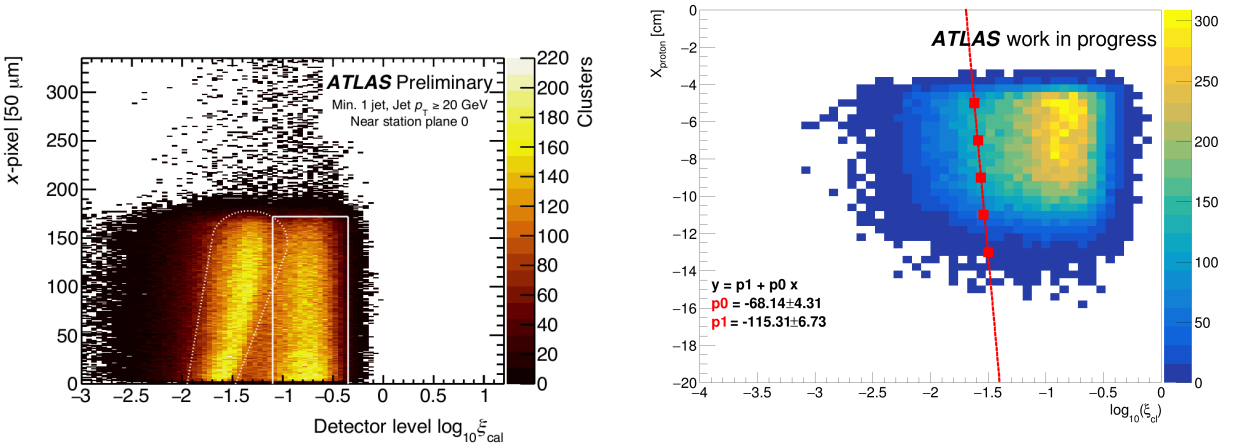


Figure 5.30: **Left:** Correlation between the x-pixel ID (in Near station plane 0) and the energy deposited in the calorimeter ($\log_{10}(\xi_{cl})$). Plot shows results of the analysis done for 2016 data. **Right:** proton x-position registered in the AFP C Near station and the energy deposited in the calorimeter ($\log_{10}(\xi_{cl})$). The results obtained from run 331020. Red points indicate places where the diffractive pattern is located.

A difference in units on the y axis does not play a role since the x-pixel position used in the left plot is strictly connected to the x-proton used in the right one (*cf.* Sec. 4). A diffractive pattern, indicated with the white dotted line (the left plot in Fig. 5.30), is not evident in the run 331020 data. This, together with the observation of the MC-data difference in the number of events visible in Fig. 5.29, may indicate an unexpectedly high background contribution.

5.6 Background Determination

In order to investigate the issue of the unexpectedly high background contribution described in the previous section, a dedicated analysis was performed. First, for each run, two data samples were prepared: the "signal" and the "background" one. The "signal" sample contained events triggered by HLT_j20_L1AFP_A_OR_C_J12 and passing the selection described in Section 5.4.

As the main background was expected to come from the Non-Diffractive jets (ND JJ) overlaid with the pile-up or beam halo protons, the idea was to use the sample dominated by the ND jets. For that purpose, the *HLT_j20_L1MBTS_2* triggered events were considered. In cases where this trigger was not active in the menu, the *HLT_mb_sptrk*

was used. For such a "background" sample, the same selection as for the "signal" one was used with the exception of the requirement of the forward proton.

For all events from the "signal" and "background" samples passing the selection criteria a distribution of ξ_{cl} (*cf.* Tab. 5.3) was made. An example of run 331020 is shown in Fig. 5.31. The "signal" distribution is plotted in red whereas the "background" one is shown with the black points.

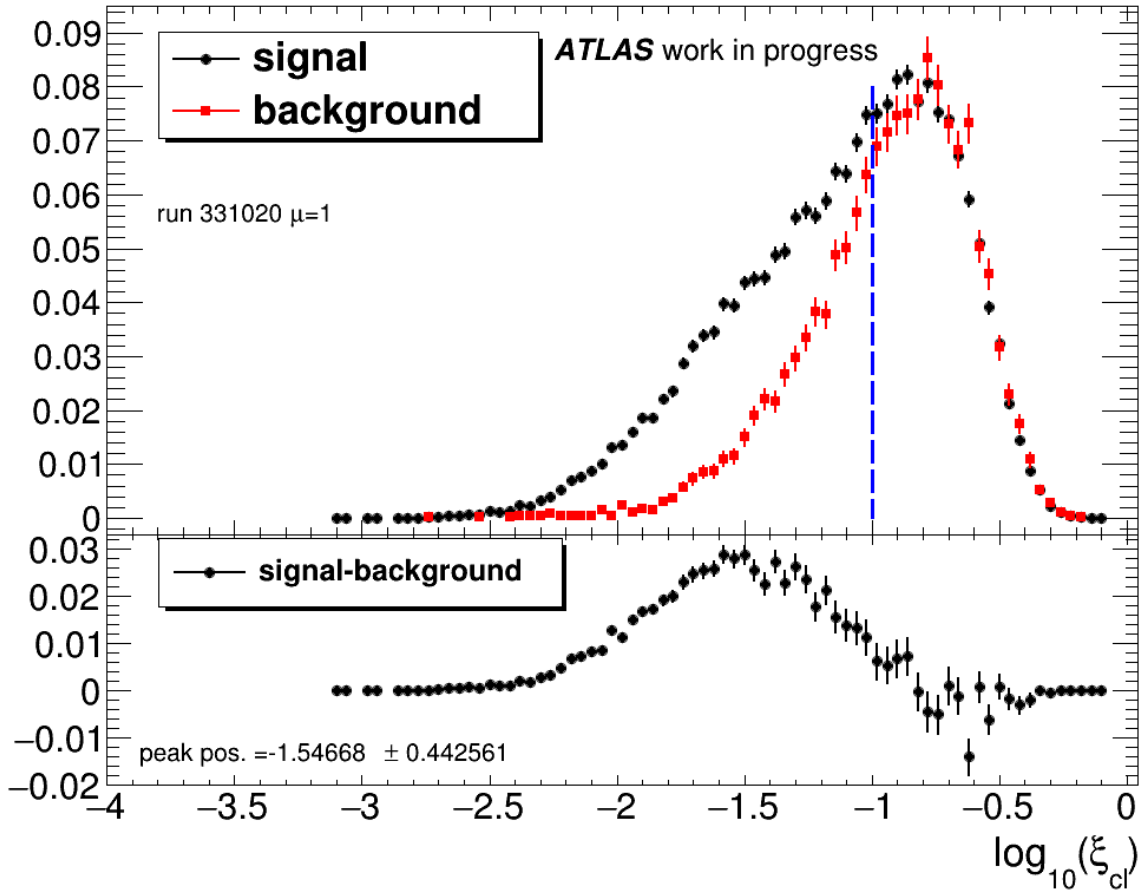


Figure 5.31: **Top:** the distribution of the energy loss computed from the ATLAS calorimeter, ξ_{cl} for the "signal" (red points) and "background" (black points) samples in run 331020. For the sample definition see text. The samples were normalised such that the integrals over the high-end 25 bins (out of 100) were equalised. **Bottom:** The difference between the signal and background distributions.

In order to determine the region dominated by the SD jets, the samples were normalised SUCH that the integrals over the high-end 25 bins (out of 100) were equalised. This is because both samples are expected to contain the background events (ND JJ + a proton from pile-up/halo), which in both cases should populate the region of high ξ_{cl}. A difference of shapes of the "signal" and "background" samples is visible in the region of $-2.5 < \log_{10}(\xi_{cl}) < -1.2$. Here the "signal" sample is expected to have an enhancement of the SD

JJ events (although the exact amount should not be directly concluded due to the specific normalisation used).

Next, a simple subtraction of the "background" from the "signal" was performed, see the bottom panel of Fig. 5.31. This procedure was repeated for several regions of x-proton, and the peak position (maximum) of the subtraction result was determined. Finally, these "peak values" were plotted on top of the x-position vs. $\log_{10}(\xi_{cl})$ distribution in order to "guide the eye" to the region where the diffractive pattern is expected to appear. See the right panel of Fig. 5.30 presenting the run 331020 data.

Finally, a line was fitted to the "peak values". It should be noted that it has a certain slope resembling the skew observed in the 2016 analysis – see Fig. 5.30 (left; focus on white dotted contour). This indicates that the "diffractive pattern" is present in 2017 data, but not visible "by an eye" probably due to the background dominating the data sample.

5.6.1 Very Low-μ Run (336505)

The best signal-to-background ratio is expected in run 336505 taken with the smallest pile-up, $\mu \approx 0.04$. However, as it was explained in Sec. 5.3.3, only information from the central calorimeter ($|\eta_{cl}| < 3.1$) can be used. Therefore, in the x-position vs. $\log_{10}(\xi_{cl})$ distribution for this run, shown in Fig. 5.32, ξ_{cl} is computed using only central clusters.

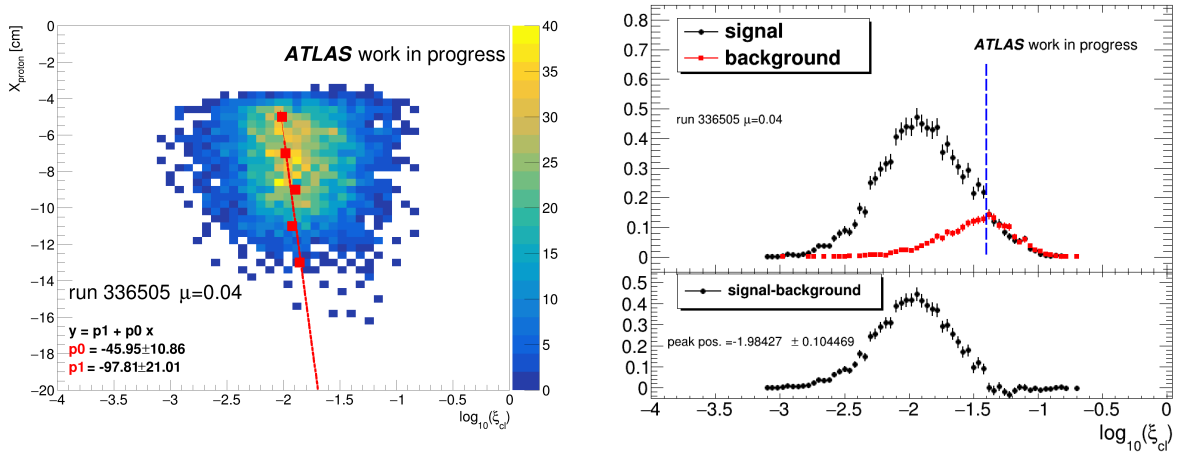


Figure 5.32: The signal-to-background ratio study for run 336505 with $\mu \approx 0.04$. **Left:** the correlation between the proton x-position and $\log_{10}(\xi_{cl})$. Only the central clusters ($|\eta_{cl}| < 3.1$) were considered. **Right top:** the distribution of the energy loss of a proton computed from the ATLAS calorimeter clusters, ξ_{cl} for the "signal" (red points) and "background" (black points) samples in run 336505. The samples were normalised such that the integrals over the high-end 35 bins (out of 100) were equalised. For the definition of samples see the text. **Right bottom:** result of the background subtraction from the signal.

A first observation is that the proton energy loss distribution is shifted towards the lower values in comparison to run 331020. This was anticipated, as neglecting the forward

part of the detector means that less energy will be used for ξ_{cl} calculations. However, a second observation is that the "signal" sample visible on the right plot in Figure 5.32 contains relatively more events within $-2.5 \lesssim \log_{10}(\xi) \lesssim -1.5$ than the "background" one if compared to run 331020. This indicates that purity of single diffractive jet signal in data taken at $\mu \sim 0.05$ is indeed higher than in runs taken with higher pile-up. The direct comparison of mentioned data-sets is, unfortunately, difficult due to the fact that the information concerning the forward region was lost. For the same reason studies of kinematic properties of diffractive jets using data from run 336505 are not straightforward.

5.7 Gap-like Signal Selection

Observations from run 336505 indicate the possibility of additional selection based on the activity in the forward region. As was mentioned at the beginning of the present Chapter, it is not exactly a rapidity gap, since the gap is expected to be very small due to kinematics imposed by the AFP acceptance, but an analogical one.

The forward cluster multiplicity distribution ($|\eta_{cl}| > 3.5$), after the quality selection described in Sec. 5.3.3, is presented in Figure 5.33. Only events which passed the signal selection (*cf.* Sec. 5.4.5) were considered. The grey area represents the SD JJ Monte Carlo predictions while the coloured points show the data: run 331020 (red), 341649 with $\mu \sim$1.0 (blue) and 341649 with $\mu \sim$2.0 (magenta).

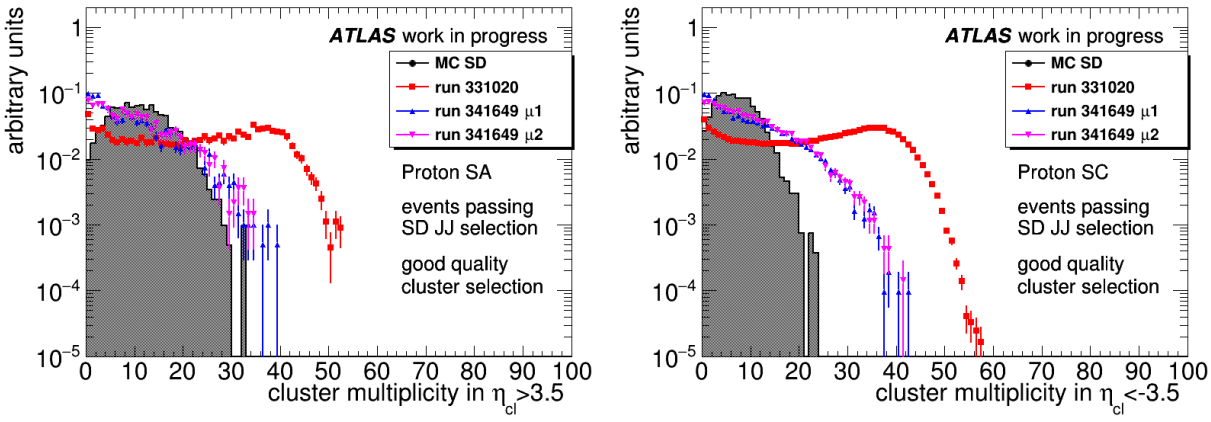


Figure 5.33: The forward cluster multiplicity distribution in events with the forward proton tagged on sides A (**left**; $\eta_{cl} > 3.5$) and C (**right**; $\eta_{cl} < -3.5$). The grey area represents the SD JJ Monte Carlo predictions, the coloured points mar the data: run 331020 (red), 341649 with $\mu \sim$1.0 (blue) and 341649 with $\mu \sim$2.0 (magenta). All plots are normalised to unity and the events had to pass the SD JJ selection. Only clusters which passed the quality cuts were considered.

The left plot shows the distribution for $\eta_{cl} > 3.5$ when the proton is tagged on side A (positive η). The right one is for clusters having $\eta_{cl} < -3.5$ with a proton on side C.

Clearly, in run 331020 the cluster multiplicities are much higher than expected from the Monte Carlo. Also, for side C in the run 341649 case, the data exhibit a longer tail than the MC. This indicates that a "gap-like" selection based on the number of clusters in the forward region may provide another was of differentiation between the signal and background.

Knowing this, it is interesting to study the multiplicities of the forward clusters in the "signal" and "background" samples (defined in Sec. 5.6). Such distribution is shown in Fig. 5.34 where the run 331020 data are presented as an example. Again, the distributions were normalised to equalise the integrals over the last bins (15 out of 50). Clearly, starting from the multiplicity of 35, the "background" falls quicker than the "signal" with decreasing multiplicity of clusters.

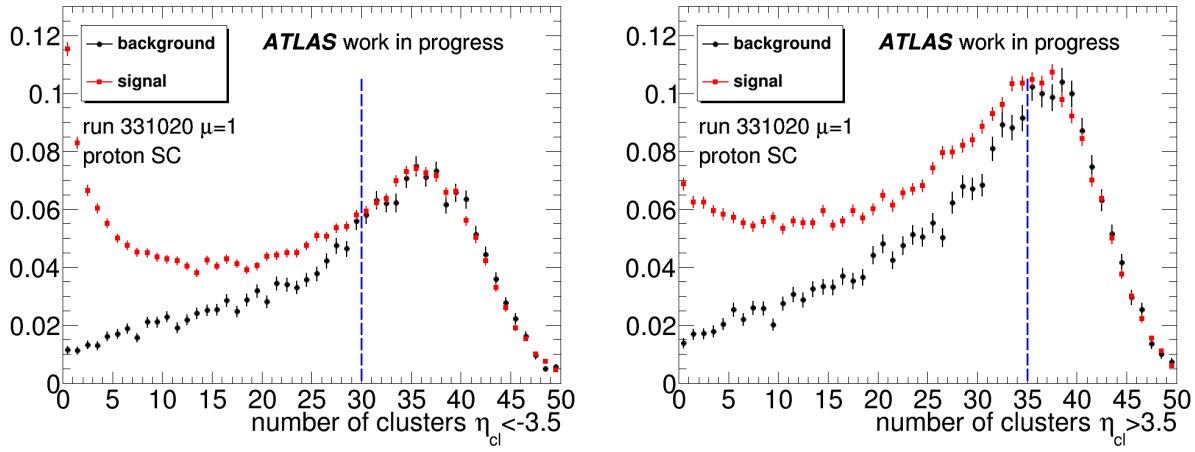


Figure 5.34: The multiplicity of clusters with $\eta < -3.5$ (**left**) and $\eta > 3.5$ (**right**) in run 331020. The symbols represent the "background" sample (black) and events containing the forward proton on side C in the "signal" sample (red). The normalisation took into account the integral from 30 to 50 for the left plot and from 35 to 50 for the right one.

The situation is similar for the other runs collected with $\mu \gtrsim 1$. Two of them are shown as an example in Fig. 5.35. This time only the clusters having $\eta < -3.5$ for events in which the forward proton was tagged on side C (negative η) were selected. Still, a different behaviour of the "signal" and "background" samples are visible, but in the region of lower multiplicities compared to run 331020.

To conclude, as there are differences in the activity in the forward calorimeters on the side of the proton between SD MC and data, an additional selection criterion can be defined. For run 331020 the event will pass the "gap-like selection" when no more than 10 good quality clusters will be present on the side of the forward proton. For the other runs the maximum number of such clusters was set to 5. Finally, it should be reminded that due to hardware problems with the calorimeter settings, the forward calorimeters will not be considered at all when processing the run 336505 data.

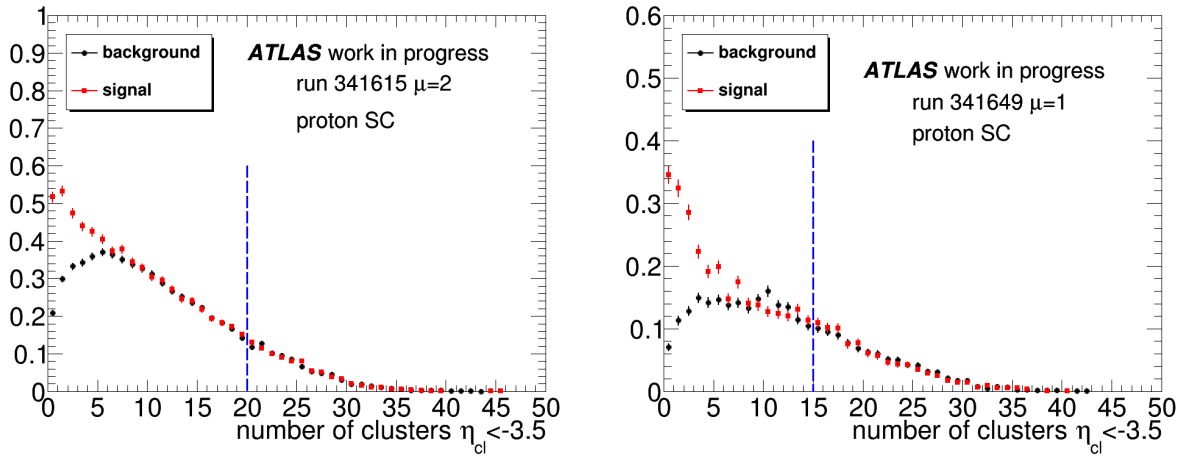


Figure 5.35: The multiplicity of clusters with $\eta < -3.5$ and the forward proton on side C for run 341615 (**left**; $\mu \sim 2.0$) and run 341649 (**right**; $\mu \sim 1.0$). The colours represent the "background" sample (black) and events containing the proton on side C in the "signal" sample (red). The samples were normalised such that the integrals over the high-end 20 bins (out of 50) were equalised. Distributions were normalised to equalise the integral of last 20 bins (out of 50).

5.7.1 Event Properties

The number of events passing this additional selection is listed in the last column of Tab. 5.5. It is also visualised in Fig. 5.36 where the transverse momentum of the leading jet in runs 331020 and 341649 is shown. This distribution should be compared to corresponding ones from Fig. 5.29. Clearly, the difference between the MC expectations and the data is getting smaller when the additional gap-like selection is applied.

For completeness, the impact of the applied selection on the properties of events was studied. Properties of the reconstructed tracks are shown in Fig. 5.37. From left to right one can see the multiplicity, transverse momentum and pseudorapidity distributions. Monte Carlo events were weighted to reflect the data (blue points) – see Sec. 5.5. The grey area is for truth particles (see Section 5.3.4 for selection definition) and red points are for reconstructed, good-quality tracks. The difference in the number of events corresponds to the one observed in Fig. 5.36. There seem to be no significant differences in the shapes of the MC and data.

The cluster multiplicity is shown in Fig. 5.38. Due to the selection related to the forward clusters only the central region (with $|\eta_{cl}| < 3.1$ taken as an example) was considered. Despite already discussed differences in the number of events between the MC and data, the data samples exhibit on average higher multiplicities than the Monte Carlo ones. This translates to even higher differences in the number of entries when the cluster pseudorapidity distribution is considered – see Fig. 5.39.

The multiplicity of the good-quality jets is shown in Fig. 5.40. The distribution of

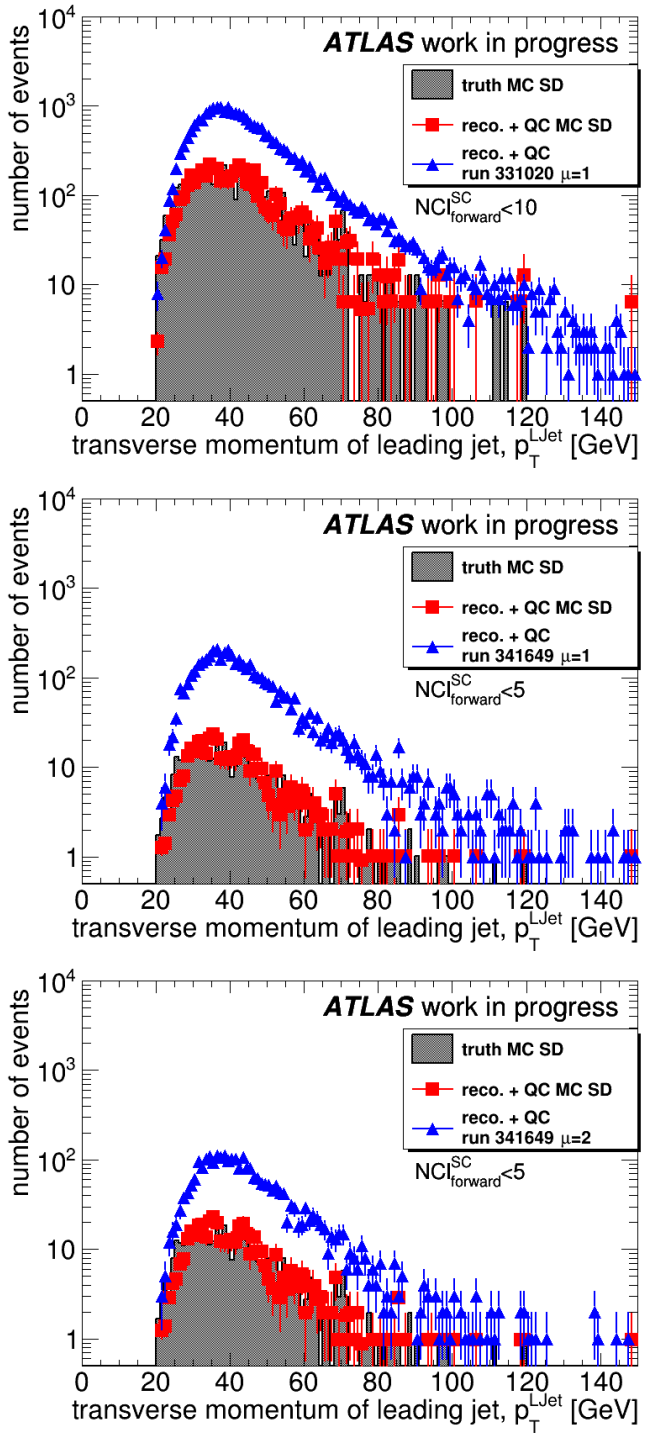


Figure 5.36: The leading jet p_T for run 331020 (**top**), 341649 (μ ~1.0; **middle**) and 341649 (μ ~2.0; **bottom**) after the selection on the trigger, GRL, single vertex, two good quality jets, one forward proton on C side and maximally 10/5 clusters on proton side for run 331020/341649. The Monte Carlo predictions (the grey area for the truth particles and red points for reconstructed tracks) were weighted to reflect the data (blue points) – see Sec. 5.5.

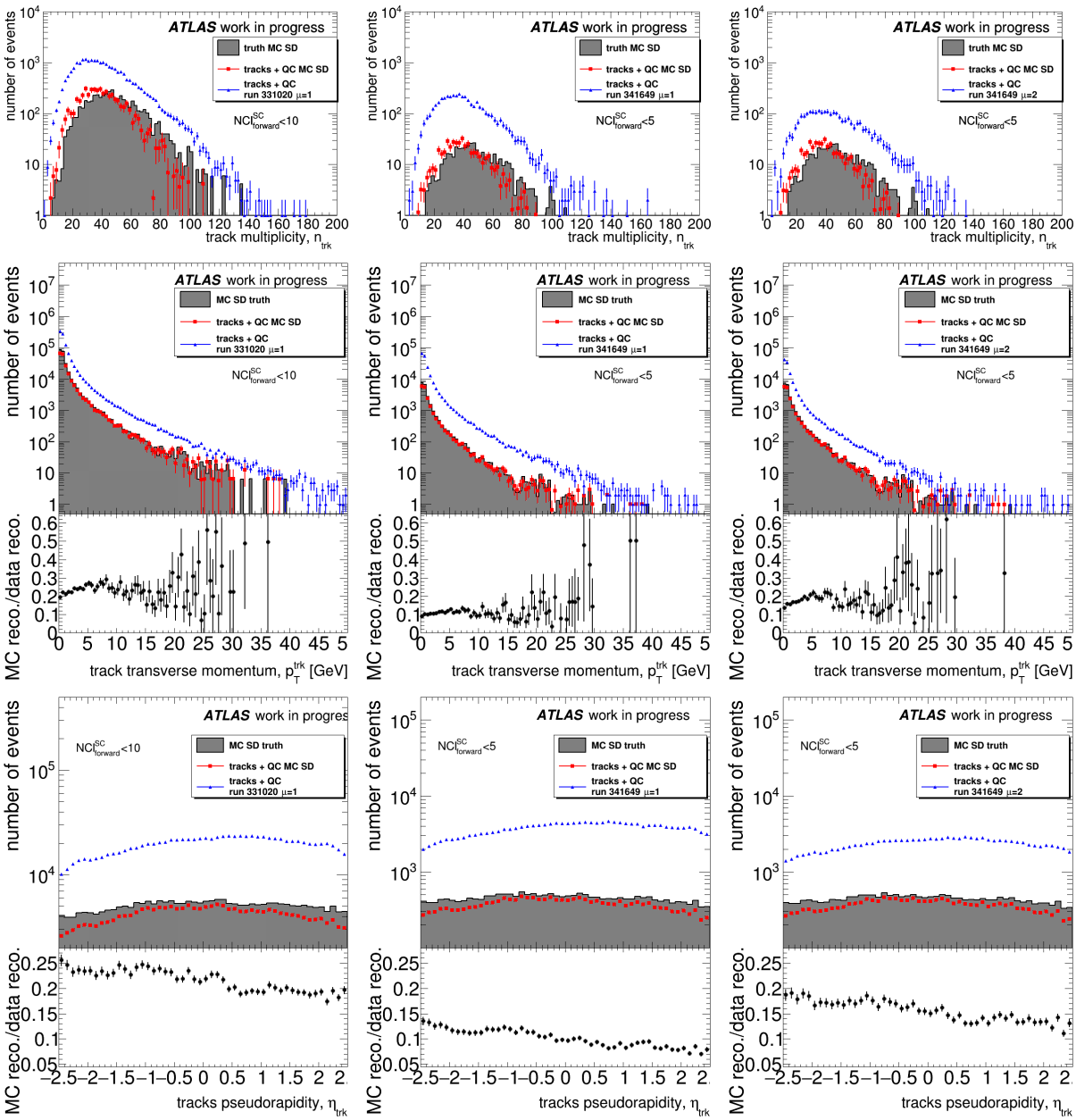


Figure 5.37: The distributions of multiplicity (**top**), transverse momentum (**middle**) and pseudorapidity (**bottom**) for good-quality tracks reconstructed in runs 331020 (**left**), 341649 ($\mu \sim 1.0$; **centre**) and 341649 ($\mu \sim 2.0$; **right**). The events passed the SD JJ signal selection with an additional requirement on the maximum number of the forward clusters ($\pm\eta_{cl} > 3.5$) on the side of the forward proton. The Monte Carlo events (grey area for truth particles and red points for reconstructed tracks) were weighted to reflect the data (blue points) – see Sec. 5.5.

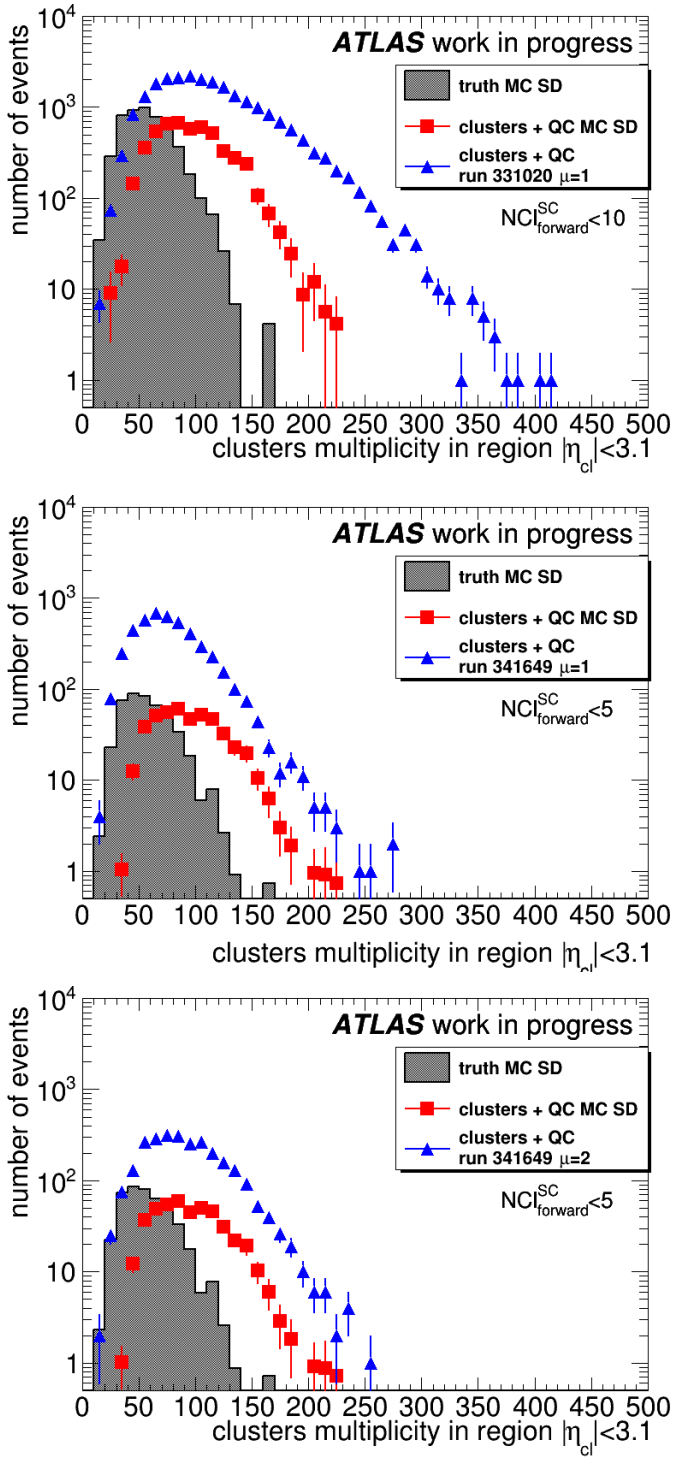


Figure 5.38: The central ($|\eta_{cl}| < 3.1$) cluster multiplicity for runs 331020 (**top**), 341649 ($\mu \sim 1.0$; **middle**) and 341649 ($\mu \sim 2.0$; **bottom**). The events pass the SD JJ signal selection with an additional requirement on the maximum amount of forward clusters ($\pm\eta_{cl} > 3.5$) on the side of the proton. Monte Carlo events (grey area for truth particles and red points for reconstructed tracks) were weighted to reflect the data (blue points) – see Sec. 5.5.

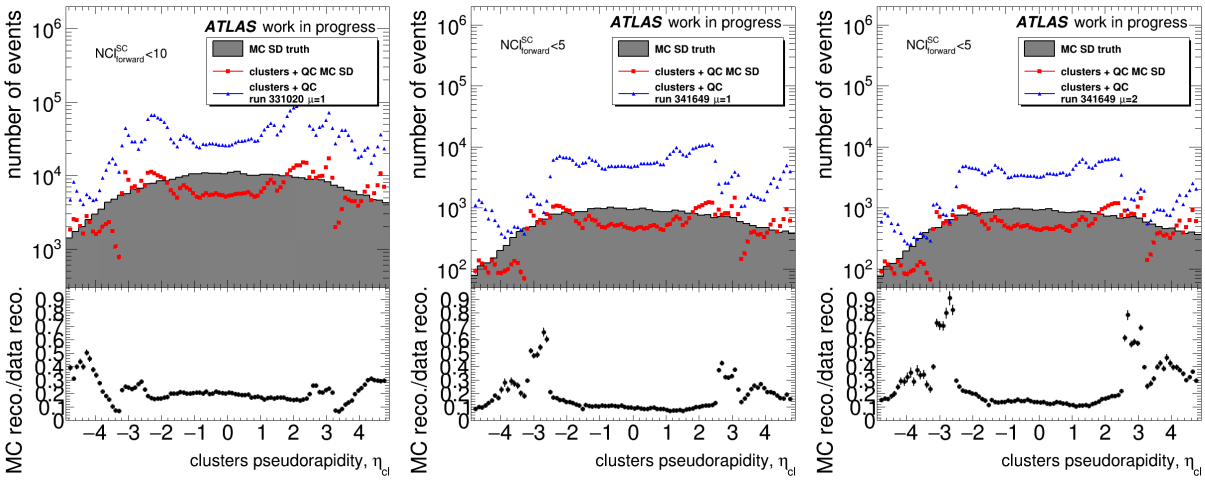


Figure 5.39: The cluster pseudorapidity for runs 331020 (**left**), 341649 ($\mu \sim 1.0$; **centre**) and 341649 ($\mu \sim 2.0$; **right**). The events passed the SD JJ signal selection with an additional requirement on the maximum number of the forward clusters ($\pm\eta_{cl} > 3.5$) on the side of the proton. The Monte Carlo events (the grey area for the truth particles and the red points for reconstructed tracks) were weighted to reflect the data (blue points) – see Sec. 5.5.

transverse momentum of the leading jet was already shown in Fig. 5.36. For completeness, the p_T distribution of the sub-leading jet is presented in Fig. 5.41. The distributions of the pseudorapidity of the leading and sub-leading jets are plotted in Fig. 5.42. For all cases, the forward proton was required on side C (negative η). For the last distribution, a left-right asymmetry is visible in data for leading jets – more events are registered on a positive pseudorapidity side.

5.7.2 Diffractive Pattern

It is interesting to observe how the additional, gap-like selection impacts the kinematic properties of events. In other words, to perform studies similar to the ones described in Sections 5.5 and 5.6.

The distributions of the proton x-position as a function of the energy deposited in the calorimeter for run 331020 considering only events with at most 10 or 5 good-quality cells on the proton side are shown in Fig. 5.43 left or right, correspondingly. The black points (and fit) are an exact copy-paste from Fig. 5.30 (right). Despite a significant contribution of the non-diffractive background (as discussed above), a diffractive pattern starts to appear. In addition, the red points, originating from the background subtraction (*cf.* Sec. 5.6) for the events with at most 10 clusters on the forward proton side, are close to the black ones (forcing the fitted functions to have the same values of the slopes and intercepts). More aggressive selection on at most 5 clusters on the forward proton side results in events shifted, on average, towards smaller values of ξ_{cl}.

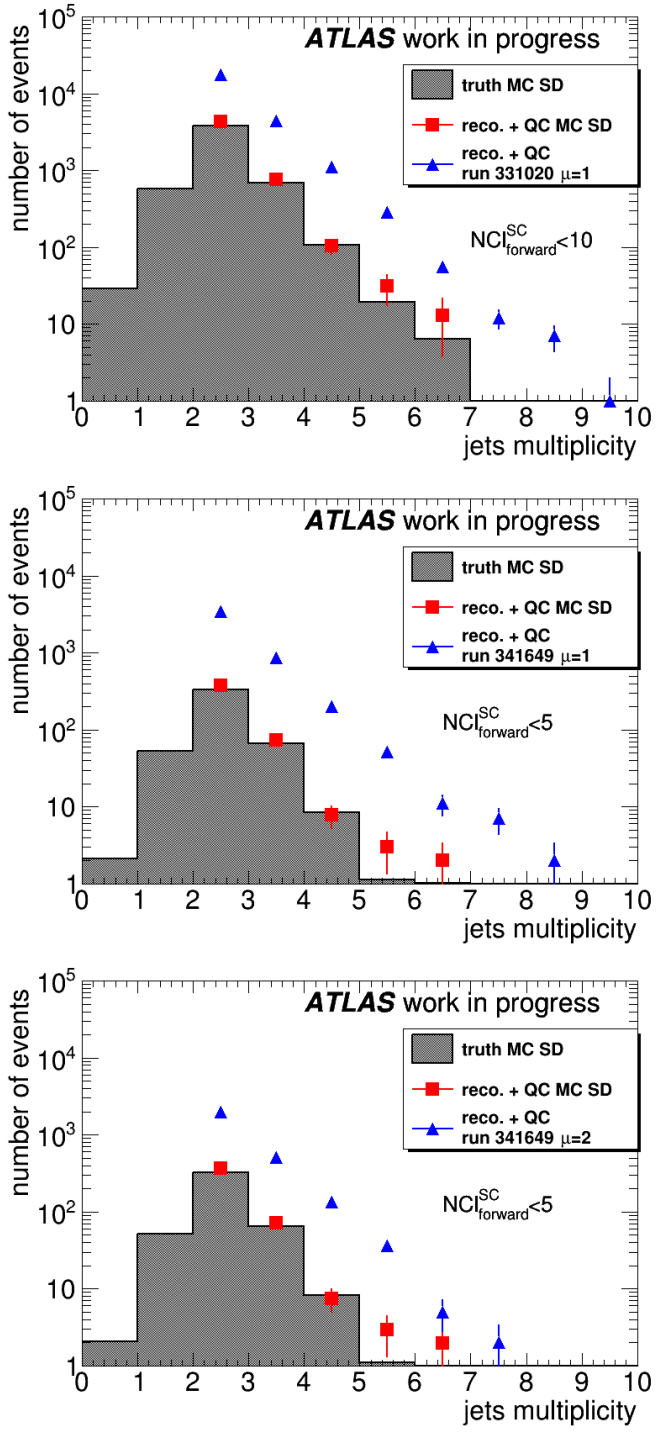


Figure 5.40: The jet multiplicity for runs 331020 (**top**), 341649 ($\mu \sim 1.0$; **middle**) and 341649 ($\mu \sim 2.0$; **bottom**). The events passed the SD JJ signal selection with an additional requirement on the maximum number of the forward clusters with the forward proton tagged on the C side. The Monte Carlo events (the grey area for the truth particles and the red points for the reconstructed tracks) were weighted to reflect the data (blue points) – see Sec. 5.5.

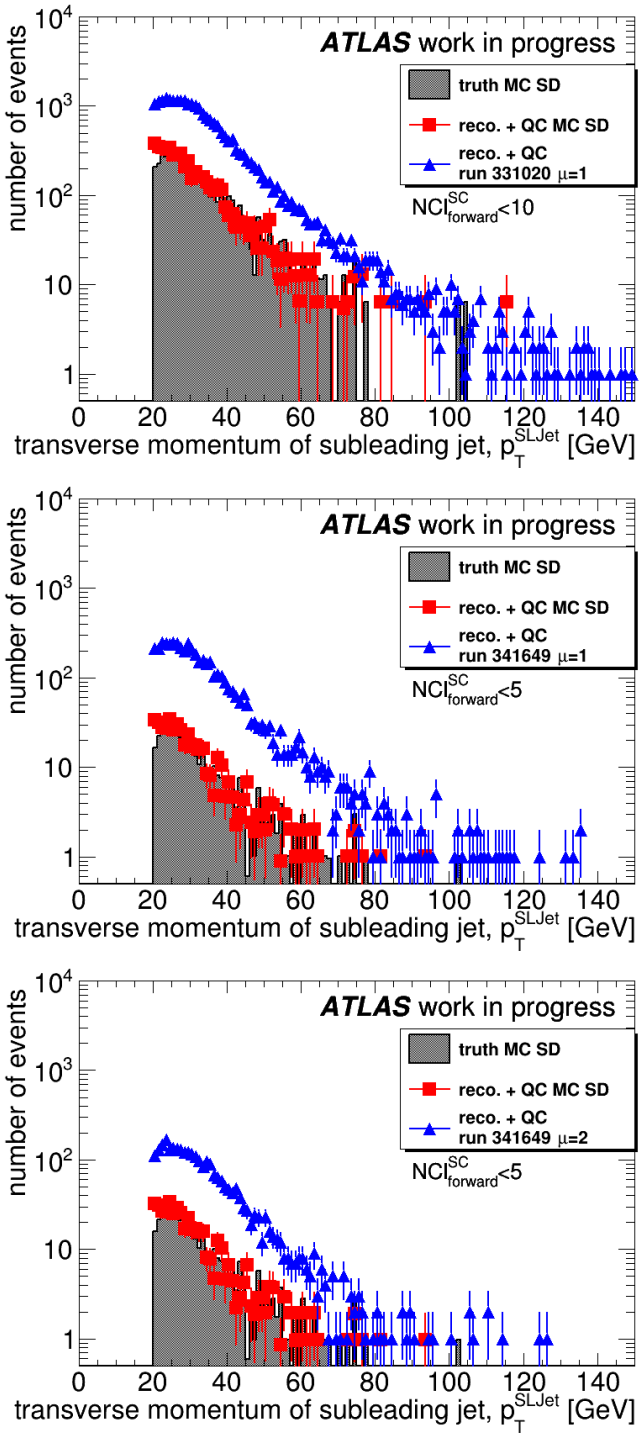


Figure 5.41: The sub-leading jet p_T for run 331020 (**top**), 341649 (μ ~1.0; **middle**) and 341649 (μ ~2.0; **bottom**) after the selection on the trigger, GRL, single vertex, two good quality jets, one proton on the C side and maximally 10/5 clusters on the forward proton side for run 331020/341649. The Monte Carlo events (the grey area for the truth particles and the red points for reconstructed tracks) were weighted to reflect the data (blue points) – see Sec. 5.5.

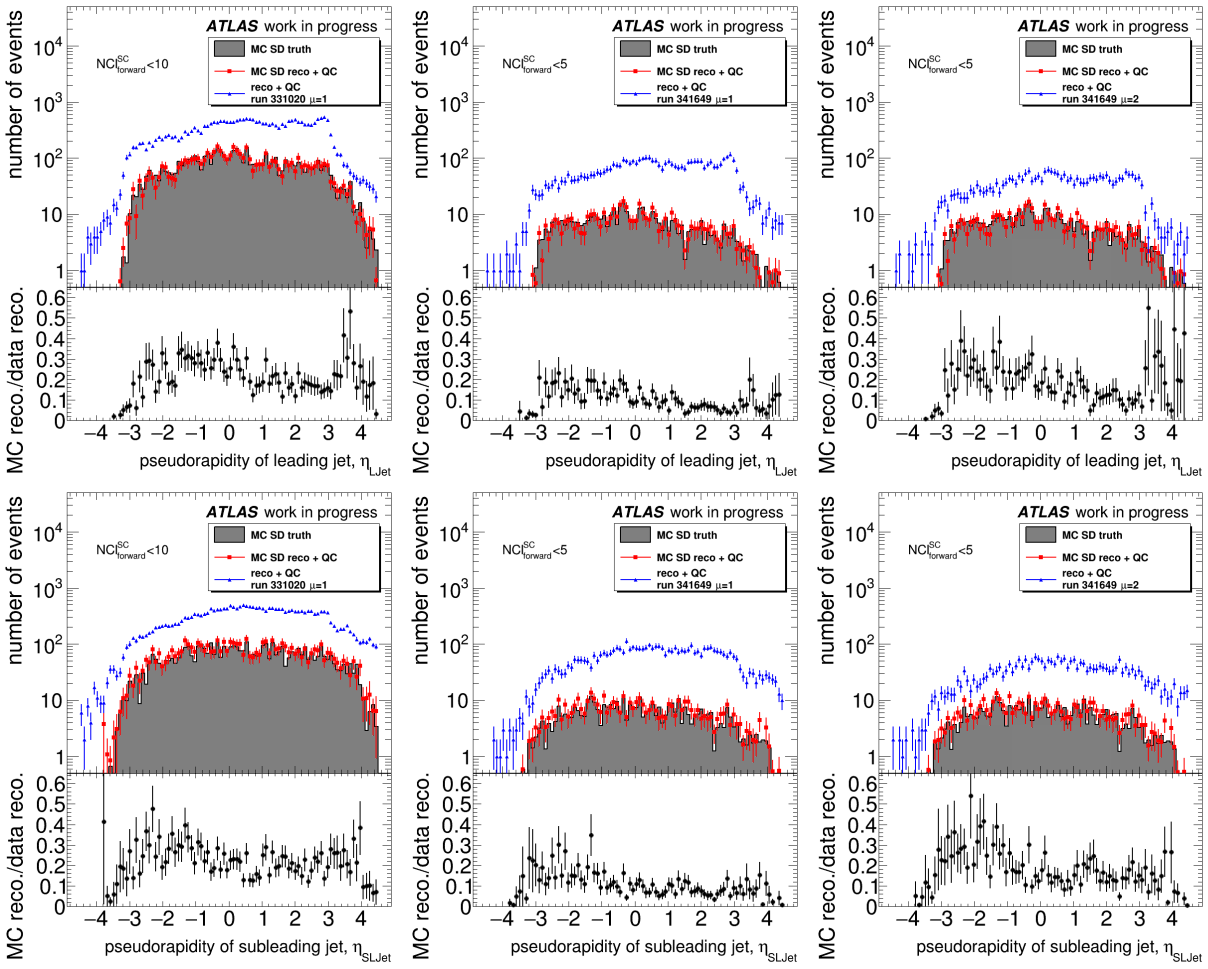


Figure 5.42: The leading (**top raw**) and sub-leading (**bottom raw**) jet pseudorapidity for runs 331020 (**left**), 341649 ($\mu \sim 1.0$; **centre**) and 341649 ($\mu \sim 2.0$; **right**). The events passed the SD JJ signal selection with an additional requirement on the maximum number of the forward clusters ($\pm\eta_{cl} > 3.5$) on the forward proton side. The Monte Carlo events (the grey area for the truth particles and the red points for the reconstructed tracks) were weighted to reflect the data (blue points) – see Sec. 5.5.

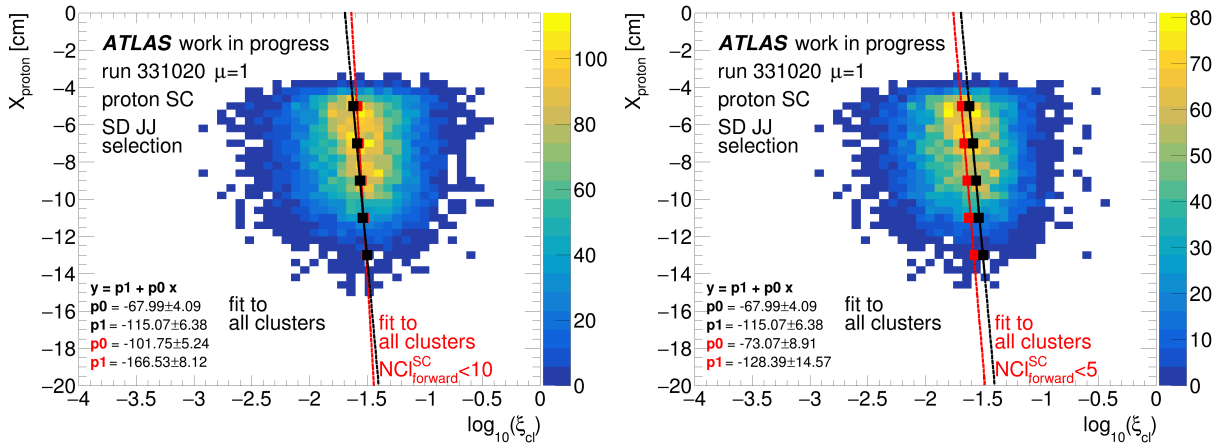


Figure 5.43: Correlation between the proton x-position registered in the AFP C NEAR station and the energy deposited in the calorimeter ($\log_{10}(\xi_{cl})$) for run 331020 ($\mu \sim 1$). The events were required to pass the SD JJ event selection (see text) and have not more than 10 (**left**) or 5 (**right**) clusters on the C side, where the forward proton was reconstructed. The red (black) points represent the mean values of the "background-subtracted" distribution for events with the forward cluster requirement (without any requirement concerning the forward clusters).

Next, a comparison with the run taken at the lowest pile-up (336505) was done. In order to make it, only clusters having $|\eta_{cl}| < 3.1$ should be considered (*cf.* Sec. 5.6.1). Distribution of x_{proton} vs. $\log_{10}(\xi_{cl})$ for run 331020 considering only events with at most 10 good quality cells on the proton side is shown in Fig. 5.44. Also here the red line represents a fit to the peak values of the subtracted ("background" – "signal") distributions. Black points are a copy-paste from Fig. 5.32.

As one can see, the peak values of the background-subtracted distributions in the case of the run taken with the smallest pile-up (black points), which also is expected to be the least impacted by the background, are shifted towards the smaller values of the energy deposited in the calorimeter in a given range of x_{proton} compared to the run 331020 taken with $\mu \sim 1$. In addition, the fitted slope is significantly different: events selected in run 336505 show, on average a stronger dependence between x_{proton} and ξ_{cl}. This difference can be due to a strong contribution of background events in the $\mu \sim 1$ sample – *cf.* Fig. 5.36 (left) and corresponding discussion.

Finally, for completeness, examples of a few other runs taken at $\mu \gtrsim 1$ are shown in Fig. 5.45. For events passing the SD JJ selection, all good-quality clusters were considered (no η_{cl} constraint). Only events with at most 5 clusters on the (forward proton) C side were taken into account. Again, the "background" was subtracted from the "signal" with the peak of the distributions plotted for a given x_{proton} range using the red points. For comparison, a copy-paste of the fit done for run 331020 (from Fig. 5.30) is plotted as black points. It should be noted that although visually the fits seem to be very similar, in the

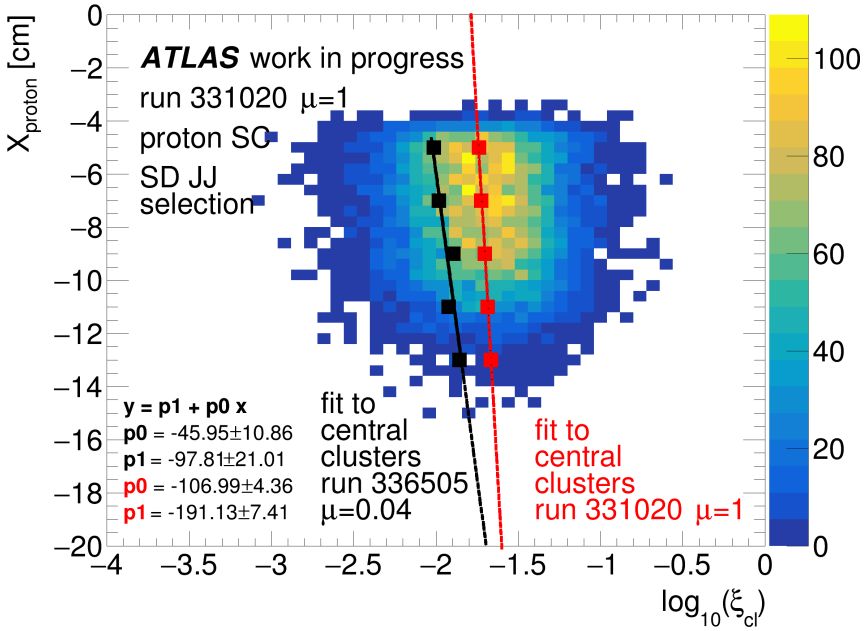


Figure 5.44: Correlation between the proton x-position registered in the AFP C NEAR station and the energy deposited in the central calorimeter (only the clusters having $|\eta_{cl}| < 3.1$ were considered in sum) for run 331020 ($\mu \sim 1$). The events were required to pass the SD JJ event selection (see text) and have not more than 10 clusters on the C side, where the forward proton was reconstructed. The red (black) points represent the mean values of the "background-subtracted" distribution for run 331020 (336505).

case of runs taken with $\mu \sim 2$ the uncertainties on the fitted intercept and slope are of the same order as the values. This is most probably a consequence of the fact that the sample is only $10 - 20\%$ pure (see Fig. 5.36 middle and right).

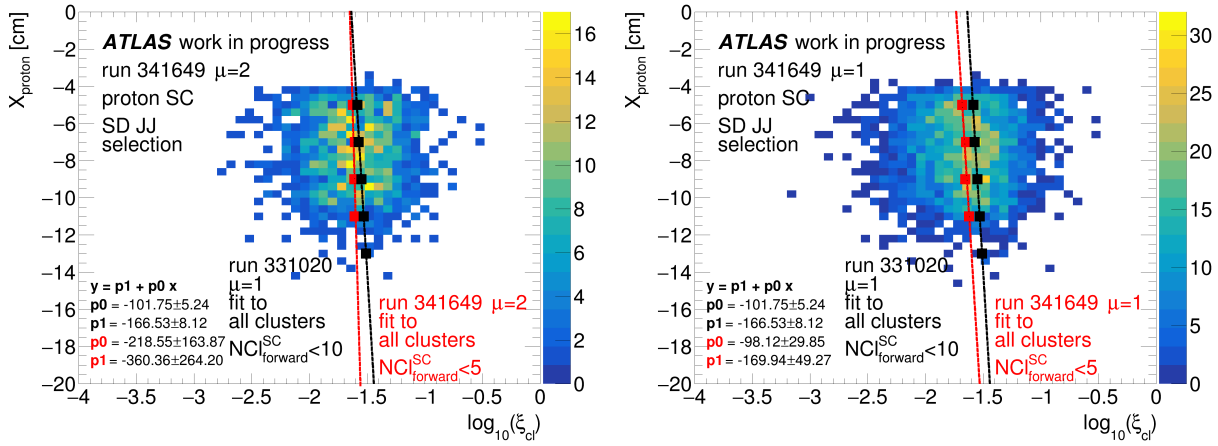


Figure 5.45: The correlations between $\log_{10}(\xi_{cl})$ calculated from all clusters and the proton x-position for the "signal" sample with the forward proton on side C after the cut on $N_{cl} < 5$ in run 341649 with μ=2.0 (**left**) and μ=1.0 (**right**). The red points represent the mean values of $\log_{10}(\xi_{cl})$ (after the cut on 5 forward clusters) in the specific x-proton range. The black points represent the mean value of distributions acquired with the background subtraction method for run 331020.

5.8 Estimation of the Pile-up Background

A sample containing non-diffractive jets was generated using Pythia 8 A14 tune [74]. The NNPDF2.3LO PDF sets were used. The sample contained events (the reported cross-section $\sigma_{gen} = 78.4$ mb) filtered at the generator level. The filter required the presence of a jet with $p_T > 20$ GeV. The generation efficiency was 0.00067152.

The sample was generated assuming various pile-up conditions, with the majority of events overlaid with 10–40 min-bias events. A fraction of events had very small pile-up values: 0.5, 1.5 or 2.5. For the purpose of this analysis, two subsets were considered: $\mu \sim 1$ containing events generated at the pile-up of 0.5 and 1.5, whereas $\mu \sim 2$ consisting of events with the pile-up of 1.5 and 2.5.

The effect of pile-up is visible in Table 5.6, where the number of events passing a given selection is listed. For the $\mu \sim 1$ sample, the rejection due to the single vertex requirement is about 53%, compared to 27% for the one with $\mu \sim 2$. The rest of the selection is exactly the same as those described in Section 5.4, except for the requirement of the proton being tagged.

After applying the normalisation as described in Sec. 5.5 it is interesting to compare

Table 5.6: The number of events in the non-diffractive Monte Carlo sample after application of a given selection.

Pile-up	Consecutive Cut			
	All	1 Vertex	2 Jets	5 Clusters
$\mu \sim$1.0	31988	17056	7094	7078
$\mu \sim$2.0	31988	8605	3526	3521

some basic MC predictions to the data. As an example, the pseudorapidity of the leading jet was taken. Its spectrum for runs 331020 and 341649 (separately for $\mu \sim 1$ and 2) is shown in Fig. 5.46. Evidently, after the event weighting, the Monte Carlo predicts much more events than are visible in the data. This is not a surprise, since for the MC a requirement of the proton tag was lifted.

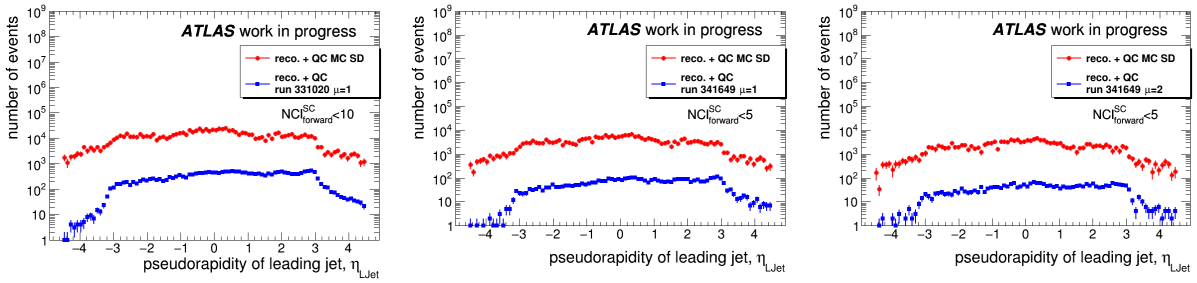


Figure 5.46: The leading jet pseudorapidity for runs 331020 (**left**), 341649 ($\mu \sim 1.0$; **cente**) and 341649 ($\mu \sim 2.0$; **right**). The events passed the SD JJ signal selection with an additional requirement on the maximum number of the forward clusters ($\pm\eta_{cl} > 3.5$) on the forward proton side. In the non-diffractive Monte Carlo case the proton tag requirement was omitted. The MC events were weighted to reflect the data (blue points) – see text.

5.9 Summary

The data collected by ATLAS during special, low-μ runs in 2017 provide an opportunity for studies of the nature of diffractive events. In the case of hard processes, the ones having the highest cross-section are those related to the (low-p_T) diffractive jet production. In addition, as the jet signature is used in other studies, like the jet-gap-jet, γ+jet or exclusive jet production analyses, it is desirable to understand the properties of such events.

In 2017, the AFP group requested two special runs to be recorded: one with $\mu \sim 0.05$ (run number 336505) and the other one wih $\mu \sim 1$ (331020). In addition, the ATLAS Standard Model group requested a few days of the data-taking at $\mu \sim 2$, which resulted in six additional medium-μ data sets (341294, 341312, 341419, 341534, 341615 and 341649).

The last one contains a part when the pile-up was lowered to 1. Unfortunately, several issues happened during these data-taking periods. The most severe one was an incorrect hardware setting of the forward calorimeter in the run with the lowest pile-up. In reality, this caused that information from the forward region of ATLAS for a part of the collected data is not useful for the diffractive data analysis.

The study presented in this thesis was the first one aimed at understanding the nature of hard diffractive events using the above-listed data sets. Therefore, a lot of dedicated mini-analyses were performed in order to investigate the detector response and data-taking conditions. These mini-studies are briefly described in this thesis, with three of them being described in more detail in Appendices.

After determining the reconstruction quality criteria, a selection of single diffractive jets was proposed. It turned out that in the runs taken with $\mu \geqslant 1$, the selection based on the proton tag and the single vertex results in samples being only $\sim 10\%$ pure – the number of collected events was 10 times higher than expected from the Monte Carlo predictions for the single diffractive jet production. Although the there are differences in the number of events between MC and data, the distributions are agreeable in shape. In addition, the kinematic relation between the forward proton registered in the AFP and the energy deposited in the ATLAS calorimeters was not obvious to reflect the diffractive nature of the events.

On the other hand, a relation between x_{proton} and ξ_{cl} was observed in the run taken with $\mu \sim 0.05$, when the central region ($|\eta_{cl}| < 3.1$) was considered. Moreover, the studies of clusters in the forward region lead to an additional, gap-like selection to be applied for $\mu \geqslant 1$. With the requirement limiting the maximum number of clusters on the forward proton side, the purity of the samples increased to about 20% with a diffractive pattern starting to be noticeable. Knowing this, one could compute a visible, effective cross-section for events passing the selection. These numbers for all considered runs are listed in Table 5.7. As an exercise, one can also compute the average visible cross-section for sides A and C getting 74.4 ± 9.0 nb and 81.1 ± 3.8 nb, correspondingly. These numbers should be, however, treated with great caution since the purity of considered samples is low an between 10-20%.

Table 5.7: The visible cross-section after a given section for all 2017 low-μ ATLAS runs. Basic selection means that the event was triggered by HLT_j20_L1AFP_A_OR_C_J12, had exactly one vertex, passed Good Run List, had two good-quality jets with $p_T > 20$ GeV and exactly one forward proton reconstructed on a given side. The number of events passing additional requirements of at most 10 or 5 clusters in the region of $\eta > 3.5$ on the forward proton side is listed in the last but one and last columns, respectively. In order to obtain the cross-section value, the number of the accepted events was scaled by the luminosity, and the reduction due to the single vertex requirement, trigger efficiency and prescale factors have to be taken into account.

Sample	**Side**	**visible cross-section [nb]**		
		basic selection	**+10>NFcl**	**+5>NFcl**
331020	A	265.4±7.0	67.4±3.6	42.6±2.7
	C	671.1±3.6	155.1±1.7	92.0±1.3
336505	A	93.2±3.3	-	-
	C	90.1±1.2	-	-
341294	A	284±14	172±12	102.0±8.4
	C	194.1±9.3	106.5±6.3	60.6±4.3
341312	A	317±11	177.8±7.0	99.1±4.5
	C	244.7±3.2	139.7±2.4	80.9±1.8
341419	A	301.6±7.3	178.5±5.8	105.3±4.2
	C	235.8±2.3	137.5±1.8	78.2±1.3
341534	A	238.7±4.6	113.3±2.8	60.5±2.0
	C	281.3±1.9	135.4±1.3	70.97±0.95
341615	A	292.2±7.1	173.5±5.7	98.1±4.2
	C	237.2±2.3	136.3±1.7	77.5±1.3
341649	A	383±16	231±11	151.0±8.5
$\mu \sim$1.0	C	353.3±6.2	209.7±4.5	134.5±3.6
341649	A	314±15	177±11	104.2±8.9
$\mu \sim$2.0	C	236.8±5.4	136.3±4.1	77.1±3.0

Chapter 6

Exclusive Jet Analysis

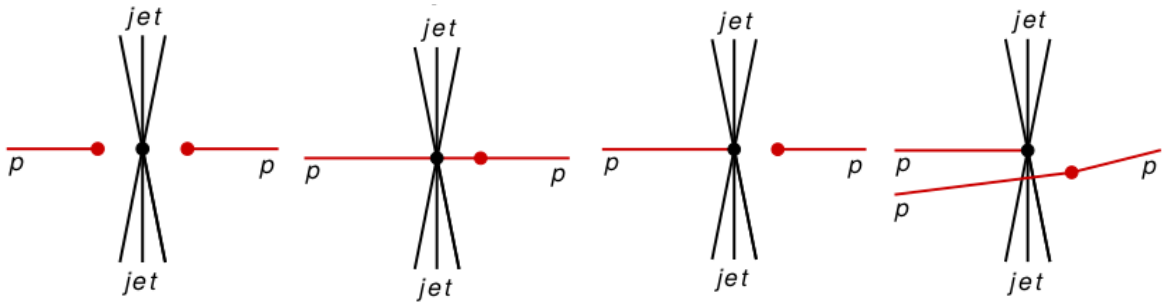
The amount of data collected during the low-μ runs in 2017 suggests attempting a search for evidence of the exclusive jet production. The signal signature is two jets and two forward scattered protons. Feasibility studies performed for the double-tagged events tagged by the ATLAS AFP [75] indicate that, due to the AFP acceptance, the jets should have the transverse momentum greater than 150 GeV. As the cross-section is steeply falling with jet p_T, such a measurement would require the data sample corresponding to the integrated luminosity of the order of inverted femtobarns. Events having jets of lower p_T can be registered if the requirement of the forward proton tag is relaxed to only a single side tag (with the other proton presumable having too small ξ to be detected in the AFP). Feasibility studies carried out for such semi-exclusive jet signature are described in [76]. They will be the base of the analysis presented in this Chapter.

The backgrounds for the exclusive jets are other processes containing jets and proton(s). The one which is expected to have the highest cross-section is due to, similarly to the single diffractive analysis case, the non-diffractive jets overlaid by a pile-up or beam halo protons. The second background contribution comes from the single diffractive jets. Finally, the Double Pomeron Exchange jet events can contribute in case only one proton is tagged. The most common configurations of those backgrounds are shown in Figure 6.1.

The initial cross-sections for the signal and backgrounds, expected for events containing a sub-leading jet with $p_T > 20$ GeV, are shown in Fig. 6.2 (first bin)

The selection proposed in [76] uses the kinematic relationship between the forward proton and the central jet systems as well as a veto concerning the activity outside the jet system. In fact, it requires:

- a proton tagged in the AFP,
- exactly one reconstructed vertex,
- the relative energy loss difference within a specific range (see text below),
- small activity in regions outside the jets.



(a) (b) (c) (d)

Figure 6.1: A schematic presentation of the backgrounds to the exclusive di-jet production: (a) non-diffractive jets + two protons from two independent pile-up vertexes, (b) non-diffractive jets + two protons from the same pile-up vertex, (c) single diffractive jets + a proton from pile-up, (d) single diffractive jets + two pile-up protons.

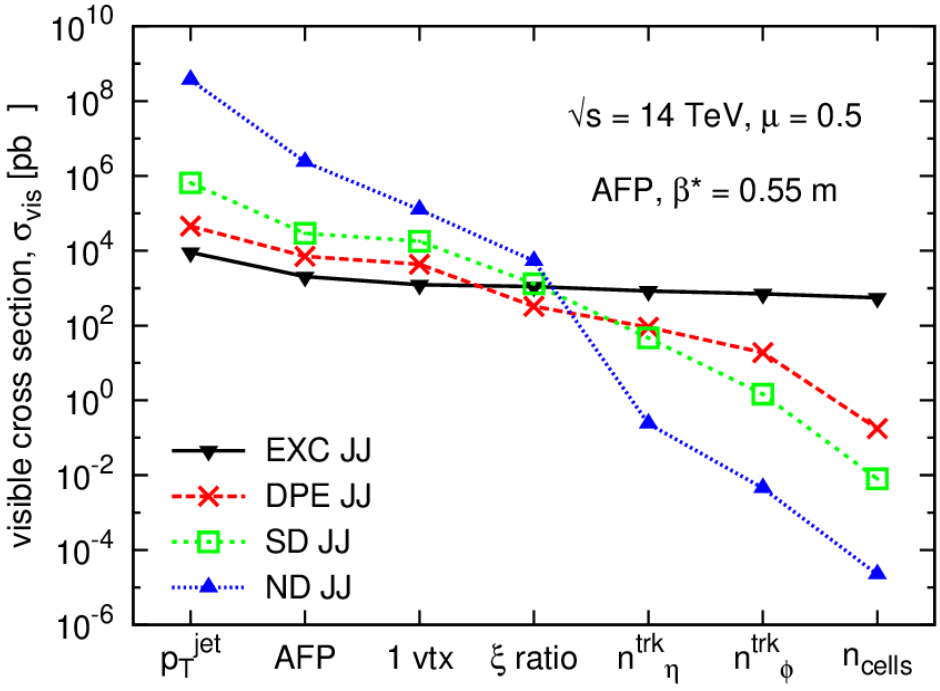


Figure 6.2: Visible cross section of the exclusive di-jet production events after a particular cut applied to the signal and background for the average number of interactions of $\mu = 0.5$ as a function of the applied consecutive cuts for AFP and for the $\beta^* = 0.55$ m LHC optics. From [76].

The first two requirements are exactly the same as those used for the Single Diffractive jet analysis. The relative energy loss difference is defined as:

$$\xi_{rel} = (\xi_p - \xi_{dijet})/(\xi_p + \xi_{dijet}),$$

where ξ_p and ξ_{dijet} are defined in Tab. 5.3. The expected distribution of ξ_{rel} is shown in Fig. 6.3. The black solid line shows the signal distribution, whereas the other colours represent the backgrounds: DPE JJ (red), SD JJ (green) and ND JJ (blue). The suggested selection criterion is $\xi_{rel} < 0.5$.

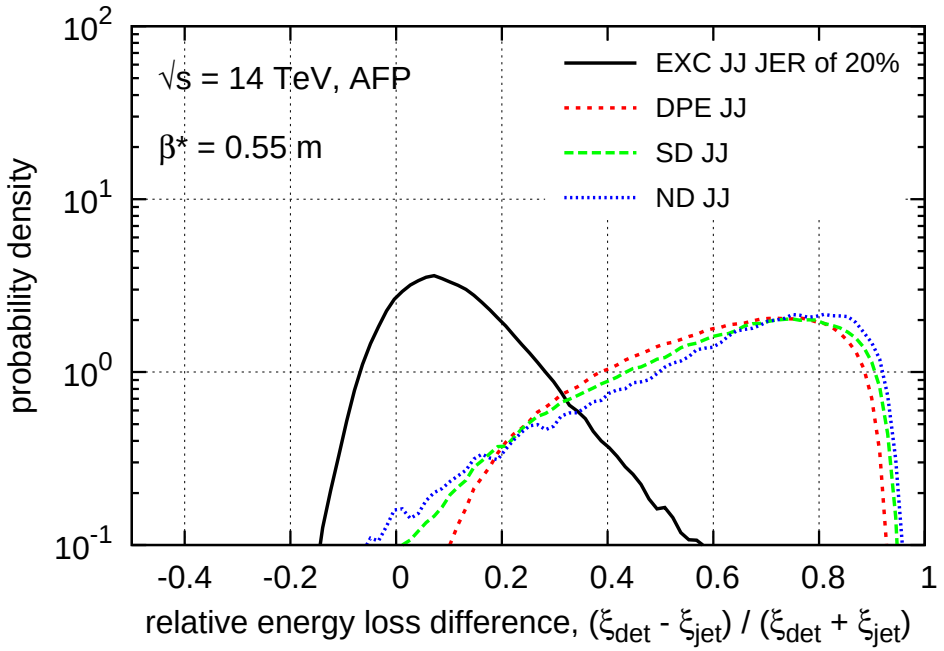


Figure 6.3: The distribution of the relative energy loss difference, $\xi_{rel} = (\xi_p - \xi_{dijet})/(\xi_p + \xi_{dijet})$ for $\sqrt{s} = 14$ TeV and $\beta^* = 0.55$ m optics. The black solid line marks the signal, whereas the other ones represent the backgrounds: double Pomeron exchange (red), single diffractive (green) and non-diffractive (blue) jets.

The last set of the selection criteria contains the cuts on the number of tracks outside the jet system in pseudorapidity (n^{η}_{trk}), the number of tracks perpendicular to the leading jet in the azimuthal angle (n^{ϕ}_{trk}) and the multiplicity of particles with the energy greater than 4 GeV produced in the pseudorapidity range of $2.5 < |\eta| < 4.9$ (n_{cells}). The suggested selection criteria are: $n^{\eta}_{trk} < 4$, $n^{\phi}_{trk} < 6$ and $n_{cells} < 2$. The predicted cross-section after a given selection is presented in Fig. 6.2.

The number of events left after the consecutive cuts for runs 331020, 336505 and 341649 taken as an example, is shown in Fig. 6.4. Events with the forward protons tagged on sides A and C are on the left and right plot, correspondingly. It should be noted that for run

336505, which suffered from the hardware settings of the calorimeter, the cut on n_{cells} was not performed.

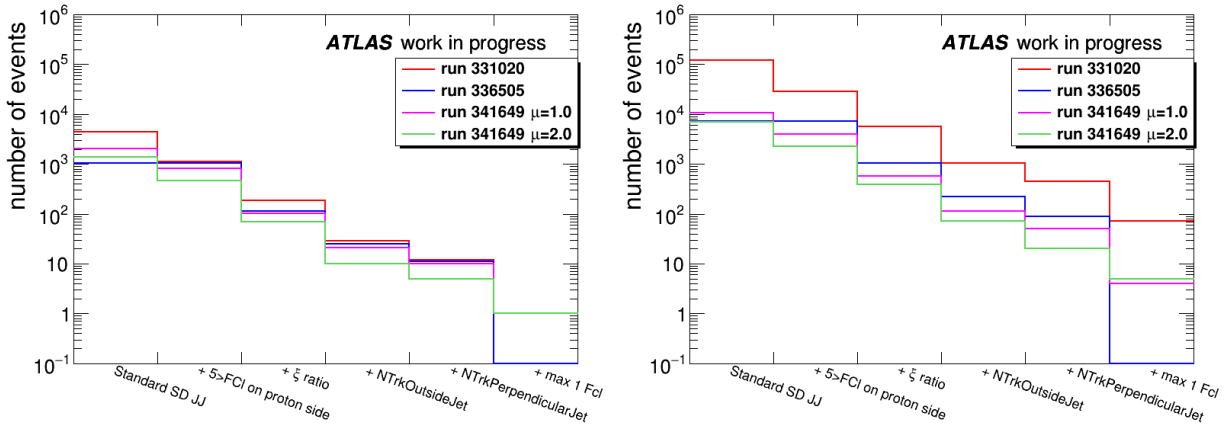


Figure 6.4: The number of events after the consecutive cuts in run 331020 (red), 336505 (blue), and 341649 with μ ∼1.0 (magenta) and 2.0 (green). Events were tagged either on side A (**left**) or C (**right**)

Before drawing any stronger conclusion, it is worth checking the behaviour of the ξ_{rel} when a selection is consecutively applied. Corresponding distributions for runs 336505 ($\mu \sim 0.05$), 331020 ($\mu \sim 1$) and 341649 ($\mu \sim 1$ and $\mu \sim 2$) are shown in Fig. 6.5. One could assume that the application of the selection criteria will make the shape of this distribution to peak around 0, as expected from Fig. 6.3. In the case of runs taken with $\mu \geqslant 1$ evidently, this is not the case. In the run taken with the lowest pile-up, the enhancement is visible already from the beginning.

From Figures 6.4 and 6.5 one could see that around 100 (in fact exactly 11/90 with proton tag on side A/C) events in run 336505 passed the selection, from which only eleven had a proton tagged on side A. From them, 6 were taken to be shown as "event displays". They are plotted in Figure 6.6 where the $\eta - \phi$ map of the ATLAS detector containing positions of the reconstructed objects is shown. The position of the reconstructed jets is shown as orange circles, with the value of the transverse momentum displayed next to it. The blue squares represent the good-quality clusters while the green triangles the good-quality tracks. The circle radius is proportional to the energy of the cluster or the transverse momentum of a jet/track. Two top plots of Fig. 6.6 show the cases when the proton was tagged on side C. The left plot shows a "very clean" event where most of the $\eta - \phi$ map is deprived of good-quality clusters and tracks. The right plot shows an example of an event containing the high-p_T jets (p_T of 150.1 and 107.4 GeV). Two middle plots show examples of events with the forward proton tagged on side A. Lastly, two bottom plots present cases where three good-quality jets are present. The left bottom plot displays a very interesting event. Here two smaller jets were produced very close to each other and their combined transverse momentum is of similar value to the third jet. It is still possible that those

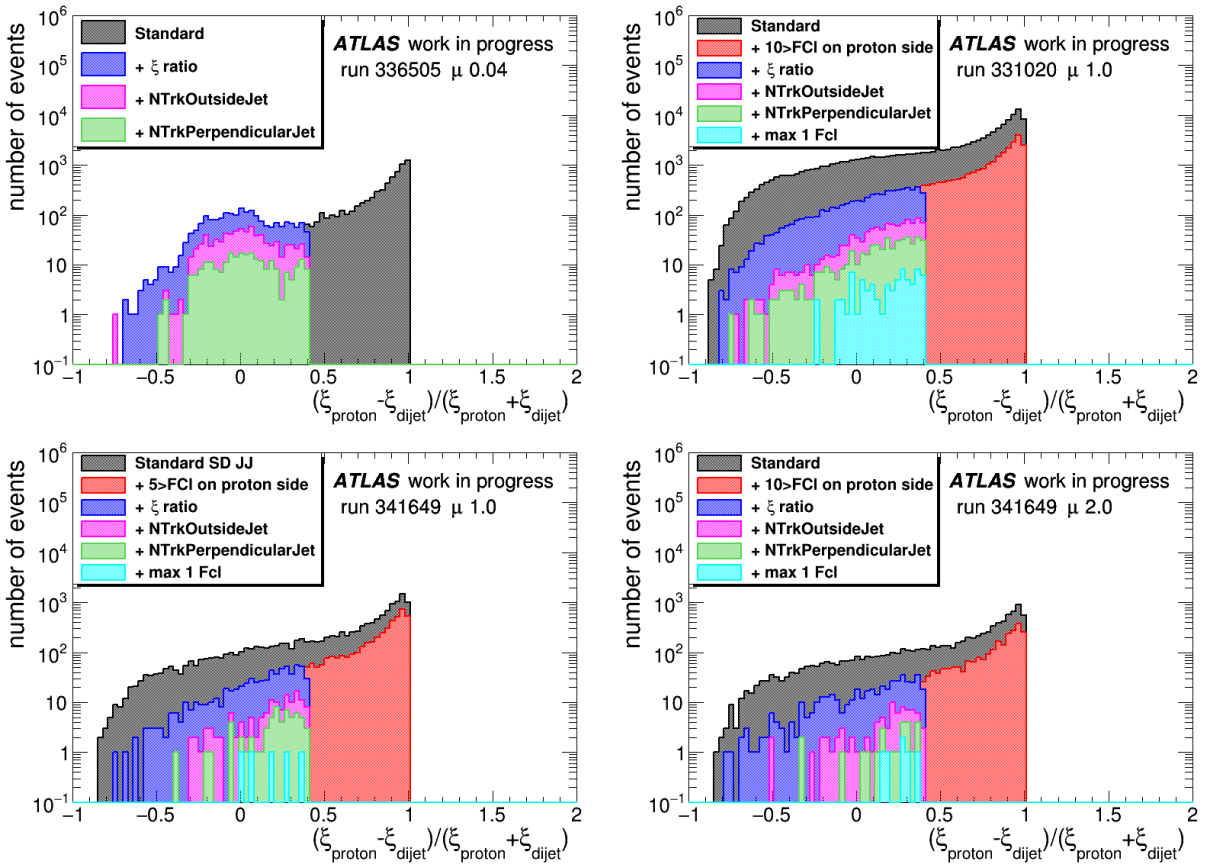


Figure 6.5: The distribution of the relative energy loss difference, $\xi_{rel} = (\xi_p - \xi_{dijet})/(\xi_p + \xi_{dijet})$ for runs 336505 ($\mu \sim 0.04$; **top left**), 331020 ($\mu \sim 1$; **top right**) and 341649 ($\mu \sim 1$ on **bottom left** and $\mu \sim 2$ on **bottom right**). Various colours represent situations after a given, consecutive cut.

smaller jets should be in fact one but the jet reconstruction algorithm split it into two.

Lastly, using the value of the visible cross-section for run 336505 (*cf.* Tab. 5.7) and considering the number of events left after the SD JJ selection (*cf.* Tab. 5.5), one can roughly estimate the visible cross-section for the production of exclusive di-jets to be $\sim$1.56$\pm$0.47 nb for the A side proton tag and $\sim$1.9$\pm$0.2 nb for the C side one. The results are in a good agreement with Monte Carlo predictions (*cf.* Fig. 6.2). A full analysis, including consideration of systematic effects, would require larger statistics. This should be considered for preparations of low-μ data-taking requests for ongoing Run 3 (2023+).

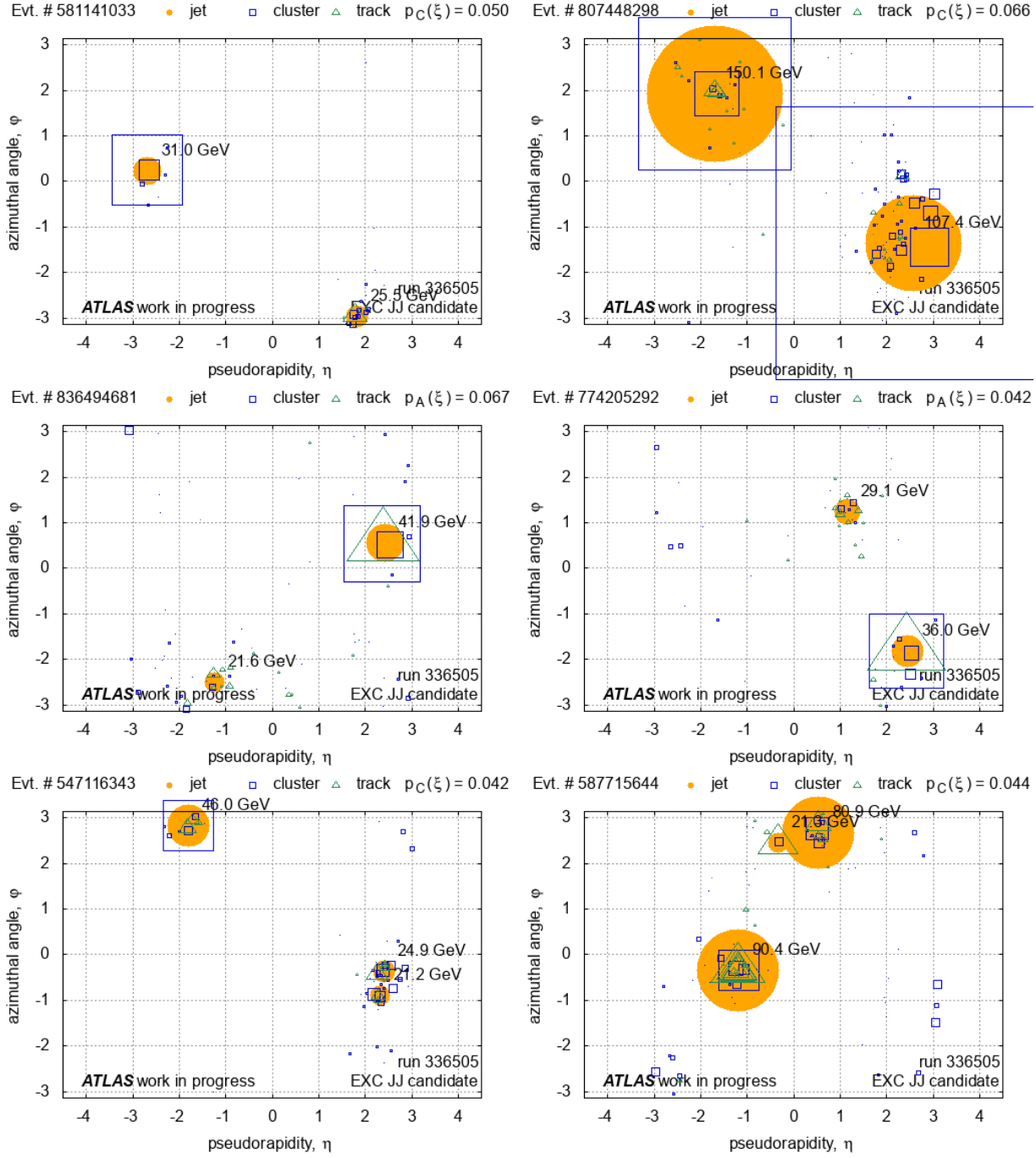


Figure 6.6: Displays of 6 events from run 336505 ($\mu \sim 0.05$) fulfilling the selection criteria for the exclusive jets.

Chapter 7

Summary and Conclusions

Diffractive processes are one of the fundamental ingredients of the Standard Model. Therefore, the study of diffraction is an important part of the physics programme at hadron colliders. They are also of wide interest to theorists. Unfortunately, despite many conducted researches, there are still many mysteries in this field. In theory, observation of jets and scattered protons enables many interesting measurements such as factorisation tests (gap survival factor), structure and nature of Pomeron and the properties of a colourless exchange within the hard system.

Since 2017 the ATLAS experiment is equipped with a full set of ATLAS Forward Proton (AFP) stations which allow measurement of scattered protons. They are horizontally moving stations installed symmetrically with respect to the ATLAS IP around 210 m. Stations located closer to the IP contain a tracker (silicon pixel detectors), whereas the outer ones are also equipped with Time-of-Flight (ToF) devices.

In 2017, the AFP detectors collected data alongside the ATLAS in both high- and low-μ LHC runs. In total AFP collected 32.0 fb^{-1} of data. Unfortunately, studies of single diffractive and (semi-)exclusive jets require using data taken at small pile-up, which was only a tiny fraction of the full data set. In fact, around 45/60 nb^{-1} for proton on side A/C was collected with $\mu \sim 0.04$ and around 140 pb^{-1} with $\mu \gtrsim 1$.

Despite the relatively small pile-up in studied runs, a careful selection of events had to be made. First, a determination of the quality of objects used for the analysis had to be done. In some cases, due to operation in "non-standard" low-μ settings, these quality criteria had to be re-defined – tailored for the diffractive analysis. For example, the energy lost by a proton can be calculated not only from the registered proton properties but also from the central system using jets or clusters. Careful studies of cluster properties revealed a need for a dedicated selection of "good-quality" clusters. The "standard way" led, in some cases to unrealistic values of energy loss.

Second, event selection criteria founded on the Monte Carlo-based feasibility studies were applied. The cuts expected to be contributing most to the background reduction were:

- requirement of exactly one vertex reconstructed by ATLAS;
- ensuring that at least two jets of a "good quality" are present in the event;
- necessity of exactly one forward proton of a "good quality" without any other forward protons present in the event.

The selection was done using events triggered with various trigger algorithms. Then, it was checked which trigger logic is the best (saves most of the events). Furthermore, the advantage of using more than one trigger was studied. In the end, it turned out that for the analysis purpose, the best and simplest choice is to use a single trigger chain HLT_j20_L1AFP_A_OR_C_J12.

Unfortunately, the three cuts discussed above were not sufficient to select pure signal events – the background was still dominating the sample. In fact, data samples were estimated to be only $\sim$ 10% pure – the number of collected events was 10 times higher than expected from the Monte Carlo predictions for the single diffractive jet production, however, the shapes of distributions are comparable. In addition, there was no evident correlation between the central system (the anticipated energy loss of a proton calculated from clusters) and the forward proton (the proton energy loss calculated from the AFP measurements). Such correlation was expected from the pilot analysis done using the 2016 data sample with proton tagged only on side C.

In order to investigate whether a "diffractive pattern" is really present in the data, dedicated studies were performed. For this purpose, special "signal" and "background" enriched samples were prepared. Studies of their properties show significant differences in the behaviour of ξ_{cl} distributions – indeed the diffractive events seem to be there, but "hidden" under overwhelming background.

Special attention was paid to the most suitable data set – the one collected at the smallest pile-up. From the feasibility studies, it was expected that this run should have the best signal-to-background ratio – a purity reaching up to 0.9. Unfortunately, during the data taking the hardware settings of the forward calorimeter were not set correctly. Therefore, the information provided by this part of the detector could not be used. This resulted in studies based on only the central part of the calorimeter. Restricting the analysis to the central region, a "diffractive pattern" was clearly visible. This observation led to further investigations of runs taken at higher μ.

To improve sample purity, further studies towards the determination of selection criteria were done. It was found that the distributions of clusters in the forward regions ($|\eta| > 3.5$) are significantly different for "signal" and "background" enriched samples. The "signal" events tended to have fewer clusters, especially on the forward proton side. This indicated the possibility to introduce a gap-like selection. With such additional requirement, the purity of the samples increased to about 20% with a diffractive pattern starting to be

noticeable in the runs with $\mu \sim 1.0$. However, such a pattern was still not visible in the runs with a pile-up of 2.0.

Finally, for the completes, an effective, visible cross-section for events passing the selection criteria for the production of single diffractive jets was computed. This was done for all considered runs considering separately for proton being tagged on sides A and C.

With the knowledge gained from a single diffractive analysis an attempt to search for the (semi-)exclusive jets was done. Based on the Monte Carlo predictions, such events with low transverse momentum jets and only one proton being tagged should be visible in the low-μ 2017 data sets.

The starting point of the analysis was selection as in the case of single diffractive jets. Next, following the suggestions from the feasibility studies, the kinematic connection between the proton and jet system was used. The distribution of the relative energy loss difference, ξ_{rel}, was expected to be peaked at 0 for the signal events. Quite surprisingly, this was the case only for data taken with the lowest pile-up, indicating that the background contribution in the case of other samples was much higher than expected. A determination of the origin of such background (beam halo? higher than expected soft diffractive contribution?) is certainly an interesting topic for separate, dedicated studies.

Finally, the most interesting cases, passing the exclusivity cuts, from the mentioned data set were selected and shown as event displays. The estimate of the visible cross-section for the production of (semi-)exclusive jets production is 1.56 ± 0.47 nb and 1.90 ± 0.20 nb for events with proton on sides A and C, correspondingly. The numbers are in reasonable agreement with the predictions of the feasibility analysis. However, a full analysis of the systematic effects requires much higher statistics of the experimental data and that of the Monte Carlo samples.

For future data taking, besides obvious remarks concerning checking that the detectors are set correctly, a few additional observations can be made. First, there is no need to have multiple trigger items in the menu as their efficiencies are now estimated. In fact, two or three trigger chains (HLT_j20_L1AFP_A_OR_C_J12 being the main one) are enough for the jet analysis. On the other hand, one should make sure that they are prescaled as low as possible (in fact, it would be worth trying to keep them un-prescaled). This should be possible, as the SD JJ data sample should be taken at $\mu << 1$ due to surprisingly high backgrounds, which are not entirely reducible in the analysis. Finally, for such data-taking one should aim at gathering the integrated luminosity of the order of several hundreds of inverse nanobarns.

Appendices

Appendix A

Dead and Low-Efficiency Pixels in AFP Silicon Detector

In an ideal situation, every pixel in AFP Silicon Detectors should provide a signal when a proton is passing thru it. Unfortunately, it is not always the case. During the data taking a pixel may become "faulty" and not deliver a signal despite the passing proton. It is also possible that a pixel provides a signal (=noise) even if there is no passing proton, resulting in false information. Before performing a physics analysis, it is important to understand which pixels are defective and how they may influence the proton trajectory reconstruction. Below, the stress is put on the identification of "faulty" pixels.

To recall, see Sec. 4.2, each AFP station contains 4 layers of silicon detectors with a spacing of about 9 mm. A single layer contains 336 pixels (0.05 mm wide; called the short side) in the x-direction and 80 pixels (0.25 mm wide; called also the long side) in the y-direction. Each pixel is 0.23 mm thick (z-direction). Here, the pixel identification number (ID) for the short side of the pixel will be referred to as a row whereas the ID of the long side will be referred to as a column.

In this work, three different types of flawed pixels are distinguished:

1. dead pixels – never give a signal; on a hit map they are visible as empty areas,

2. Low Efficiency (LEFF) pixels – not always provide a signal; on a hit map they will be visible as areas with a much lower number of hits than their neighbours,

3. noisy pixels – deliver more signal than their neighbours; visible peaks on the hit map.

The data gathered by the AFP silicon detectors shape up into a specific pattern creating the so-called hit map. In typical data-taking conditions, the edge of the pattern has a rather irregular form (see Fig. 5.26). In addition, the low statistics outside the diffractive pattern makes the identification of faulty pixels challenging.

In the following, various "faulty pixel" recognition methods will be described and analysed. The goal is to recognise as many defective pixels as possible with a minimum number of wrongly identified ones.

A.1 Dead Pixels

The first problematic type of faulty pixels is the dead ones. They can appear as a result of the automated masking procedure performed during the SiT chip tuning. This procedure was carried out by the AFP expert a couple of times during the 2017 data-taking. The masking procedure was done when there was no beam in the LHC and the AFP stations were in the garage position (see Section 4). The procedure may select (mask) different pixels from tune to tune.

Masking is not the only reason for pixels to not give any signal. A pixel may not provide any signal because of the readout electronics fault, which could be destroyed either mechanically or due to radiation.

As the dead pixels are characterised by a lack of signal, this should be easily spotted within the diffractive pattern as not filled areas. The problem starts at the edge of the pattern where there are not many events and inactive pixels could be wrongly identified as dead ones. To avoid this, one can think of a limited search range or of the use of a larger amount of the data (if available).

The data collected by the ATLAS experiment are divided into the parts called the Lumi Blocks (LB, *cf.* Section 2.2). In consequence, one can think about the amount of data needed for a creation of a hit map in terms of the number of LBs. Here, two approaches to hit map creation are discussed. The first one uses one hit map created using 100 LBs of data. This reflects the usual amount of data collected during one run. In the second one, 20 smaller hit maps (each constructed using 5 LBs) are used to check whether the identification of faulty pixels is possible for a smaller amount of data. Such an approach may be useful to spot defects which can appear during the run.

The algorithm for dead pixel identification consists of two parts. First, single ("lone") dead pixels are identified. The second one aims at finding groups of dead pixels. A single pixel is considered a dead one when the conjunction of the following conditions is fulfilled:

- there are no hits in a pixel,
- pixel has at least 5 neighbouring pixels with any hits,
- neighbouring pixels are close enough to the diffractive pattern[1].

The above algorithm will not detect all dead pixels. Especially, the "clusters" of dead pixels

[1]The pixel was considered to be within the algorithm range if the average number of hits in its neighbours was larger than 0.5% of the maximum hits of all pixels in the studied hit map. One may anticipate that the smaller hit maps will not be fully filled with hits, having thus empty pixels that could be wrongly recognised by the algorithm as the dead ones.

will not be identified. Therefore, an additional procedure is executed on pixels which were not identified as dead. Namely, groups of dead pixels (and single dead pixels not pointed out in the first step) are identified when a pixel:

- has 0 hits and,
- has less than 5 neighbouring pixels with any hits and,
- was not marked as dead in the first iteration and,
- has at least 1 dead neighbour (recognised either in the previous or the current part of the algorithm).

This part of the algorithm is executed until no new pixels are found.

In Figure A.1 an example of the results of the above method is presented. The blue squares represent the fired pixels, the white squares symbolise the pixels that never gave any signal (the dead pixel candidates), the crossed red squares mark the pixels found as dead in the first iteration and the crossed green ones in the second. Note that the area in the top left corner is the "edge of diffractive pattern", thus these pixels should not be identified as dead. With such an algorithm, identification of all dead pixels should be possible.

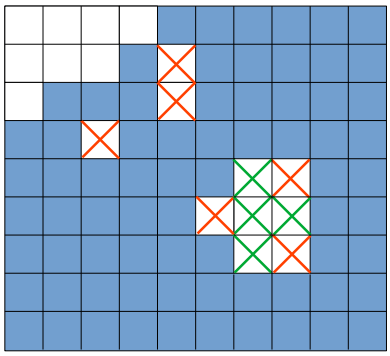


Figure A.1: Scheme of dead pixels found by the algorithm. The blue squares represent the fired pixels, the white squares symbolise the pixels that never fired (the dead pixel candidates), the crossed red squares represent the pixels found as the dead pixel candidates in the first iteration and the crossed green ones in the second one.

A.2 Tests of Dead Pixel Identification Algorithm

To check the algorithm a number of tests was performed. The main complication for a dead pixel identification is the non-uniformity of the diffractive pattern resulting in a low statistic on its edges. To recreate these conditions, the test maps should include this characteristic distribution.

To simplify the procedure of imitating the diffractive pattern, the test maps were based on data. Since data contain dead, inefficient and noisy pixels, a "smoothing procedure" was

developed. Smoothing was done by calculating the average of hits in the studied pixel and its neighbours[2]. Such an averaged number of hits in the vicinity of a given pixel was set as its value in the test map. The procedure was carried out for all pixels. A special treatment was applied for pixels located at the edge of the detector. If a pixel was on the edge its five neighbours were considered when calculating the average. In the case of pixels located in the corner, only three neighbours were taken into account. Such masking should not bias the samples for the dead pixel analysis. However, for the LEFF pixel identification, it will significantly influence the average number of hits which are used in algorithms. Fortunately, as these hit maps were used for the sole purpose of performing simple tests and drawing the first conclusions concerning the methods, it is not a major problem.

As it was mentioned before, for the tests the five LB hit maps were considered. An example of such a test hit map is shown in Figure A.2 (left). The 5 LB samples were chosen to check if a quick identification of faulty pixels is possible. This can be very useful for recognising potential problems appearing during the data taking.

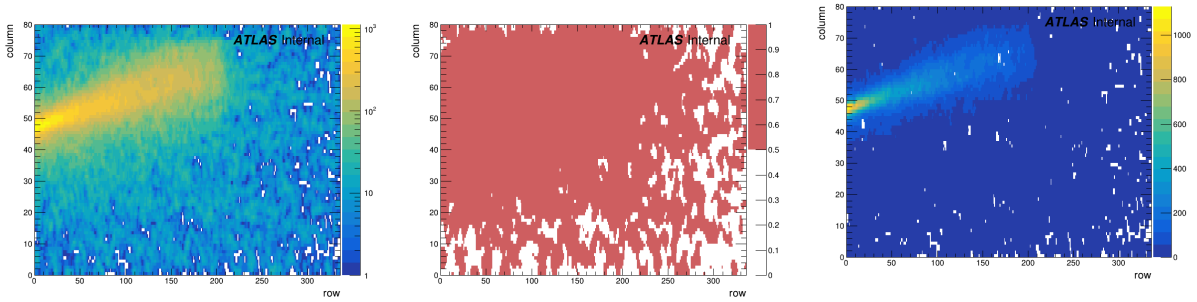
The single test starts with 20 hit maps (each being equivalent to 5 LBs of data) without any defective pixels. In the next step, 50 pixels were randomly chosen within the diffractive pattern. The area of the diffractive pattern was defined in the same way as the range of the algorithm operation, but a pixel could be chosen to be dead when the average number of hits in its neighbourhood was larger than 0.5% of maximum hits in the hit map. This was done in order to avoid picking pixels in the area without hits. The possible range of random faulty pixels is shown in Figure A.2 (centre).

Since it is expected to observe the clusters of dead pixels, in each test five groups of pixels, were randomly set. The number of pixels in such a group could be 7 or 8 and the groups could be of different shapes. This results in a maximum of 40 dead pixels being in groups. In order to reach the limit of 50 dead pixels, the rest was chosen to be the single ones, randomly distributed over the allowed area. An example of a final test hit map with added dead pixels is shown in Figure A.2 (right). A comparison of the left and right plots reveals which pixels were added within the considered range (plot at the centre). For example, the group of dead pixels can be found around (row, column) = (17,41) and a single one can be spotted at (17,37).

In the previous section, the conditions for a pixel to be recognised as a dead one within one hit map were defined. In a perfect situation, in all twenty hit maps, the dead pixels would be correctly classified. However, since the algorithm operates in different ranges for different hit maps, some pixels may be located out of the range and will not be recognised as dead ones. Therefore, an additional condition was introduced: a pixel had to be recognised as dead in at least 1 hit map and remain empty in the rest. This allowed finding all dead pixels in all 50 performed tests without any pixel being incorrectly recognised.

It should be mentioned that one way to increase the range of the algorithm is to run it

[2]Empty neighbouring pixels were not included in calculations of the average



(a) Test 5 LBs Hit Map (b) Random Range (c) The 5 LBs Test Hit Map with added dead pixels.

Figure A.2: **Left**: an example of a hit map obtained from the data after the smoothing procedure. **Centre**: the range where the faulty pixels may appear (see text). **Right**: an example of a hit map with added dead pixels.

on a hit map created from a full run, where usually all non-dead pixels were fired at least a few times. As a consequence, there is no need for setting the ranges since the algorithm will be at its peak performance. Such tests were done 50 times for a hit map created using 100 LBs of data and in every test, all dead pixels were correctly identified.

A.3 Application of Dead Pixel Identification Algorithm to 2017 Low-μ Runs

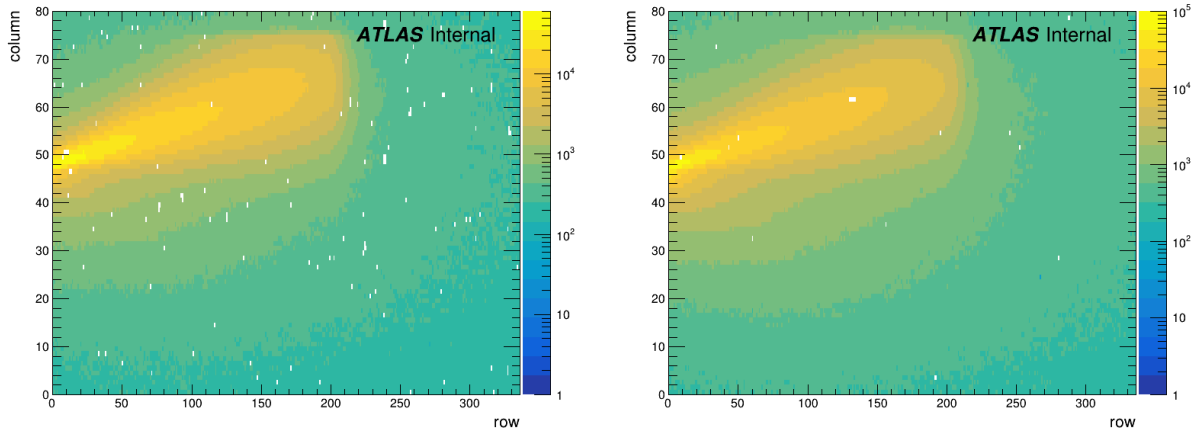
After the tests on the artificial (generated) maps, the next step was to check the performance of the algorithm using the experimental data sample. The analysis was performed for the 331020 run data from 2017. The data sample came from the calibration stream[3]. In this run, two planes, P1 and P3, in station A FAR were turned off. Also, the P1 plane of the C NEAR station had the high voltage turned off. Only events collected when the AFP was inserted were considered. For all active planes in each station, the hit maps using 100 LBs of data were created.

An example of the hit maps for the A NEAR station is shown in Figure A.3. Just by looking at them, one can immediately point out the pixels which should be identified as dead ones (white, not filled).

As was already described, some pixels were masked during the tuning procedure. The data from these pixels are not stored, thus they should appear as dead ones. Since the list of masked pixels is known, one can check if the algorithm properly detects them. The results for the A NEAR station are shown in Figure A.4. The colours on the z-axis mean:

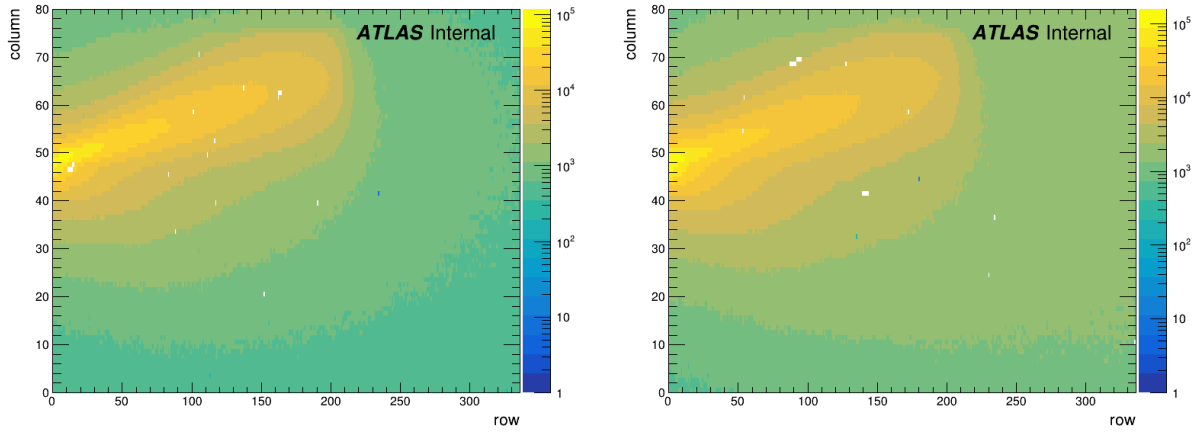
- green – pixel that was found dead and was not masked in the tuning procedure,

[3]This stream contains the data only from the AFP subsystem, resulting in a much smaller event size and, thus, a much higher rate/number of stored events compared to the physics stream.



(a) Station A NEAR Plane 0

(b) Station A NEAR Plane 1

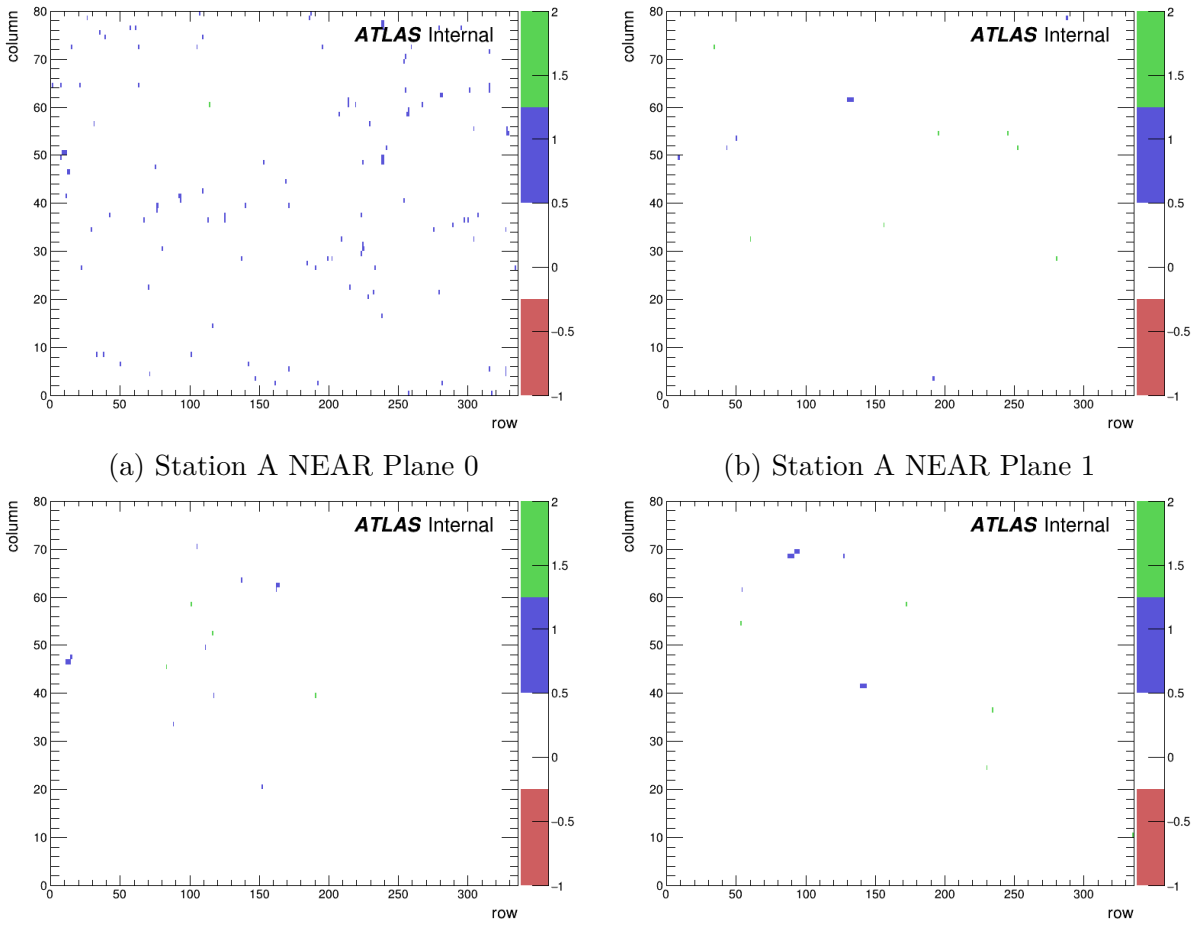


(c) Station A NEAR Plane 2

(d) Station A NEAR Plane 3

Figure A.3: Hit maps from run 331020 for station A NEAR (100 LBs of the data taken when the AFP was inserted). The dead and LEFF pixels are visible.

- blue – pixel that was found dead and was masked,
- white – pixel that was not found dead and was not masked,
- red – pixel that was not found dead but was masked.



(a) Station A NEAR Plane 0

(b) Station A NEAR Plane 1

(c) Station A NEAR Plane 2

(d) Station A NEAR Plane 3

Figure A.4: Results of the dead pixel identification method applied to the 331020 run data collected by the A NEAR station. Pixels marked green are the ones found by the algorithm but not masked, blue means that the pixel was found dead and it was masked, white represents the pixels not found as dead and not masked and, finally, red marks the pixel that was masked but not found by the algorithm. It should be noted that there are no "red" pixels in the studied sample.

In all cases, the algorithm identified all masked pixels. In addition, the algorithm also found pixels which were not masked but were not giving a signal. Usually, for studied 2017 low μ runs, less than 10 dead but not masked pixels are identified within a hit map. This does not mean that the masking method works incorrectly, but rather that during the data taking some additional pixels became faulty.

A.4 Low Efficiency Pixels

In this section, the main focus will be on the Low Efficiency (LEFF) pixels. By definition, this type of faulty element only sometimes gives a signal. Such behaviour is visible in a hit map as an area with a lower number of hits compared to its neighbourhood (*e.g.* pixel (235,42) in Figure A.3c).

The main question that immediately appears is how the pixel becomes inefficient and at which point it should be considered LEFF. In this work, 6 different algorithms for the recognition of such pixels will be used. Five of them (based on the neighbouring pixels and a Gaussian fit) will be described in this Section, whereas the last one (the track method) is discussed separately in Section A.6.

A.4.1 Neighbouring Pixels Methods

As it was mentioned before, LEFF pixels are characterised by a smaller number of hits than their neighbours. Therefore, the simplest way to determine the candidates for the LEFF pixels is to study the average number of hits in the adjacent pixels (av_{method}) and compare it to the number of hits in a studied pixel ($P_{(row,col)}$). For the LEFF pixels, the ratio defined as

$$RATIO_{method} = \frac{P_{(row,col)}}{av_{method}}$$

should be smaller than a threshold, which will not exceed one[4]. This threshold must be carefully identified since too small value can lead to a large number of incorrectly identified pixels, whereas too large threshold can lead to a smaller efficiency of the algorithm. The selection of the most effective threshold will be presented below.

Four methods based on the neighbouring pixels were defined (see also a scheme in Figure A.5):

1. 2_COL – the method using 2 neighbouring pixels in a column:

$$av_{2\_COL} = 0.5 \cdot (P_{(row,col+1)} + P_{(row,col-1)})$$

2. 2_ROW – the method using 2 pixels in a row:

$$av_{2\_ROW} = 0.5 \cdot (P_{(row+1,col)} + P_{(row-1,col)})$$

3. 4_PIX – the method using 4 pixels:

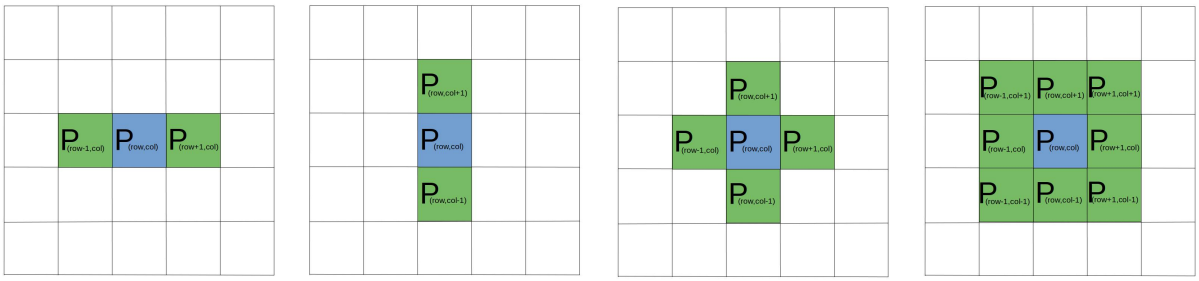
$$av_{4\_PIX} = 0.25 \cdot (P_{(row,col+1)} + P_{(row,col-1)} + P_{(row+1,col)} + P_{(row-1,col)})$$

[4]In the case of noisy pixels the value of $RATIO_{method}$ should be larger than a threshold which will be greater than 1.

4. 8_PIX – the method using all 8 pixels located around the studied one:

$$av_{4\_PIX} = 0.125 \cdot (P_{(row,col+1)} + P_{(row,col-1)} + P_{(row+1,col)} + P_{(row-1,col)} + \\ + P_{(row+1,col+1)} + P_{(row-1,col-1)} + P_{(row+1,col-1)} + P_{(row-1,col+1)})$$

Pixels on the edges of the detector (with row ID 1 or 336 or column ID 1 or 80) have to be processed more carefully and a larger threshold should be applied to recognise them as LEFFs. This problem was solved by calculating the average as if these pixels were not on the edge and treating the missing pixels as empty ones.



(a) 2_ROW (b) 2_COL (c) 4_PIX (d) 8_PIX

Figure A.5: A scheme showing which pixels are taken into consideration (marked green) when examining a surrounded one (blue). From left to right the method: 2_COL, 2_ROW, 4_PIX and 8_PIX.

A.4.2 Fit Method

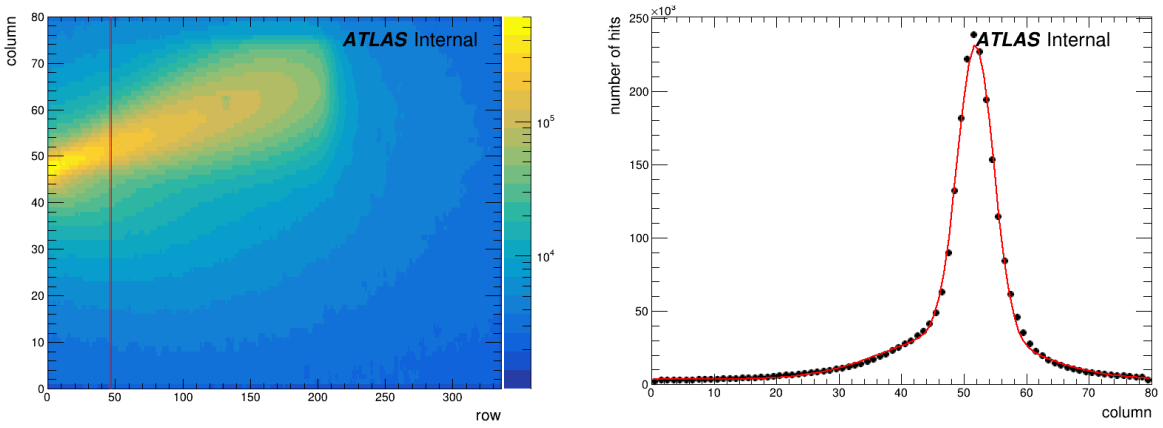
Another way of looking for the LEFF pixels is by using a double Gaussian fit to the projections of a hit map. From each row of a hit map, a one-dimensional distribution was created. An example of a hit map is shown in Fig. A.6 (left). It is a test hit map created from 100 LBs of data. It was found that the fits for smaller hit maps (*e.g.* 5 LBs) are inefficient, therefore only the tests for a full hit map will be presented. The projection of row 46 is shown on the right side of this figure. The red line marks the fitted formula.
Here, the fitted function was composed of two Gaussian functions and a constant:

$$FIT_{row}(x) = gaus_1(x) + gaus_2(x) + p_6.$$

The fit was performed for each row separately resulting in a set of functions.

The FIT(x) function has 7 parameters and the fitting procedure starts with setting initial values of the parameters for the first histogram (first row) to:

$$p_0 = 25000,\ p_1 = p_4 = Hist\_Mean(),\ p_2 = p_5 = \frac{Hist\_StdDev()}{\sqrt{2}},\ p_3 = 5000,$$



(a) Hit map (b) FIT(x) for row 48

Figure A.6: **Left**: an example of a test hit map with projected row 46 (red lines,) in station A NEAR Plane 1. **Right**: the hit distribution in row 46 (black dots) with fitted double Gaussian function (red line).

$$p_6 = max(P(row = 1, col = 1), P(row = 1, col = 2), P(row = 1, col = 3)).$$

The fit results for the current row are then used to set the initial values of the fit parameters for the next row.

The LEFF pixels are defined as the ones with a much lower number of entries than expected from the FIT(x) function. Not always a row will have "a good" fit. In some cases, some rows will not have perfectly fitted functions (especially in the maximum of the diffractive pattern, *e.g.* bins 50-53 in Figure A.6).

The following ratio was defined:

$$RATIO_{FIT_{row}} = \frac{FIT_{row}(col) - P_{(row,col)}}{\sigma},$$

where $\sigma = \sqrt{FIT_{row}(col)}$. The LEFF pixels are expected to have $RATIO_{FIT} > 0$ whereas the noisy ones will have $RATIO_{FIT} < 0$.

A.4.3 Tests of Methods

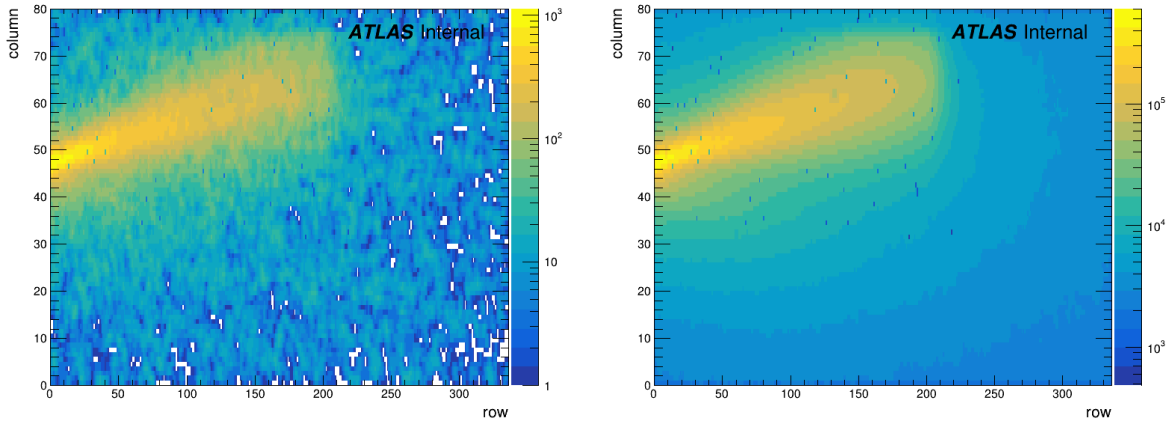
For this part of the analysis, the same 20 hit maps (5 LB each) were used as in Section A.2 for methods: 2_COL, 2_ROW, 4_PIX and 8_PIX. For the FIT method, a 100 LBs hit map was considered. For each test, 50 different single pixels were selected to be LEFF with efficiency set to 10%. An example of the produced test hit maps is shown in Figure A.7.

These tests allowed the calculation of the mean percentage of correctly found LEFF pixels and the mean number of incorrectly found ones. As was mentioned in Sec. A.4.1 it is important to choose a correct threshold. For this reason, the range of thresholds was

considered:

- $[0.1, 1.0]$ for methods using neighbouring pixels,
- $[10, 100]$ for the FIT method.

The thresholds of 2_*COL*, 2_*ROW*, 4_*PIX* and 8_*PIX* methods correspond directly to the efficiency of LEFF pixels, *e.g.* a threshold of 0.5 should detect pixels having the efficiency of 50%. A threshold for the FIT method corresponds to the σ [5] from the fitted function. Since a hit map may be not properly reproduced by the fitted functions, due to the non-Gaussian shape of the diffractive pattern (see Figure A.6 right), the thresholds were set relatively high (dozens of σ).



(a) 5 LBs Hit Map (b) 100 LBs Hit Map

Figure A.7: An example of test hit maps with added LEFF pixels, visible as pixels with much lower multiplicity (*e.g.* (73,57)). **Left:** hit map created from 5 LB of data. **Right:** hit map created from 100 LB of data.

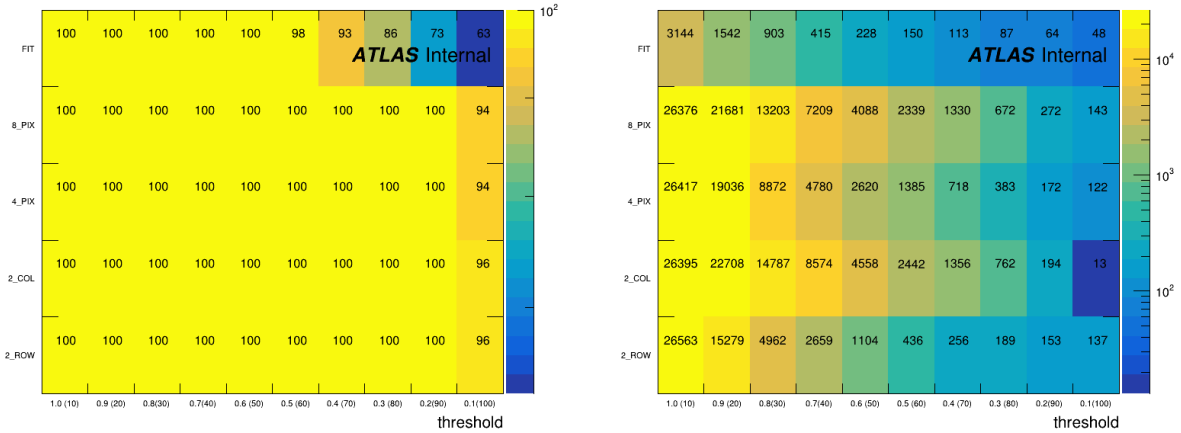
The results are shown in Figure A.8. The left plot presents the mean percentage of correctly found pixels for each method (y-axis) and given threshold (x-axis). Thresholds for the FIT algorithm are written in brackets. For methods, 2_*COL*, 2_*ROW*, 4_*PIX* and 8_*PIX* Pixel was marked as LEFF when in at least 6 hit maps (each hit map consists of 5 LB) it was identified as a LEFF candidate.

For the considered thresholds, the efficiency of 4 methods using neighbouring pixels stays at 100% up to threshold 0.2, while the FIT method efficiency starts dropping at the threshold of 60.

On the right plot of Figure A.8, the mean number of incorrectly found pixels from all tests is plotted. As previously, the x-axis represents the thresholds and the y-axis the used method. At this point, it is worth recalling that there are 26880 pixels in total. The number of incorrectly found pixels drops for all four methods (FIT method) with decreasing (increasing) value of the threshold. The FIT method finds fewer incorrect LEFF pixels than

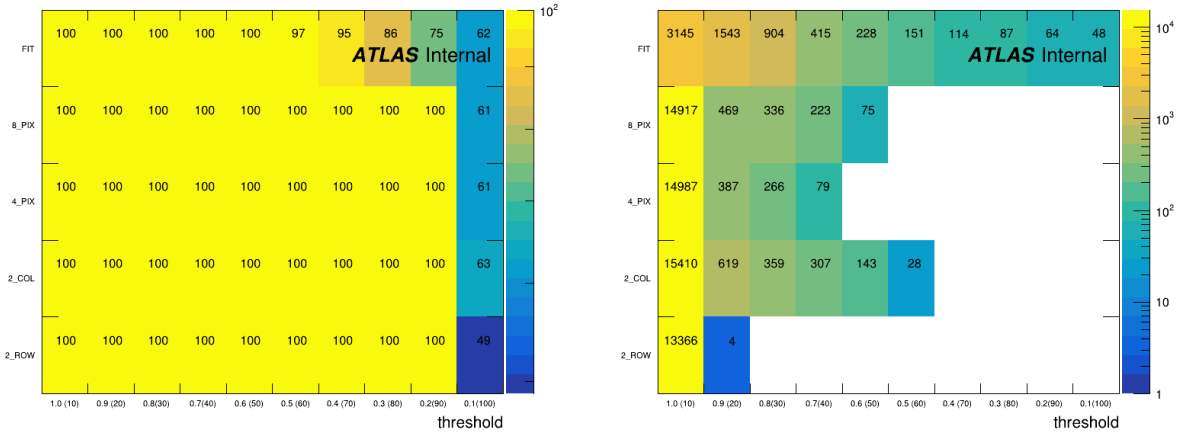
[5] Defined above as a square root of the fit value in for a specific column.

the other ones. However, this method benefits from the use of larger statistics. The best results are provided by the FIT and the 2_*ROW* methods. The high efficiency is kept while having the lowest number of incorrectly identified pixels.



(a) Percentage of correctly found LEFF. (b) Number of incorrectly found LEFF.

Figure A.8: Results for the 5 LB test hit map (FIT method used the 100 LB hit map). x-axis: the threshold used for the LEFF pixel identification, y-axis: used method. **Left:** the mean percentage of correctly found pixels for all tests. **Right:** the mean number of incorrectly found pixels for all tests.



(a) Percentage of correctly found LEFF. (b) Number of incorrectly found LEFF.

Figure A.9: Results for 100 LB test hit map. x-axis: the threshold used for the LEFF pixel identification, y-axis: used method. **Left:** the mean percentage of correctly found pixels from all tests. **Right:** the mean number of incorrectly found pixels for all tests.

The methods using neighbouring pixels (*e.g.* 2_ROW) were applied to 5 LB hit maps in order to check how much data would be sufficient for a good LEFF identification. In this paragraph, 100 LB hit maps will be used to determine how well the methods work compared

to the FIT method. The results are presented in Figure A.9. The efficiency stays at 100% but the number of the incorrectly identified pixels dropped to 0 already at the threshold of 0.3 for the 2_ROW method and 0.6 for all 4 methods. This result gives a promise of good identification when using larger hit maps.

At this point, it may seem that the 2_ROW method alone will be sufficient for finding LEFF pixels. This is not true. As was mentioned in Section A.2, the test hit maps for LEFF identification are biased. To address this issue one can think of further ways of improvement. For example, it is possible to consider the use of two methods to reduce the number of incorrectly found LEFF pixels. In Figure A.10 and A.11, the results of using such combinations are shown. The abscissa is the threshold and the Z-axis shows the mean percentage of correctly found pixels or the mean number of incorrectly found ones. There are a few outcomes when using two methods:

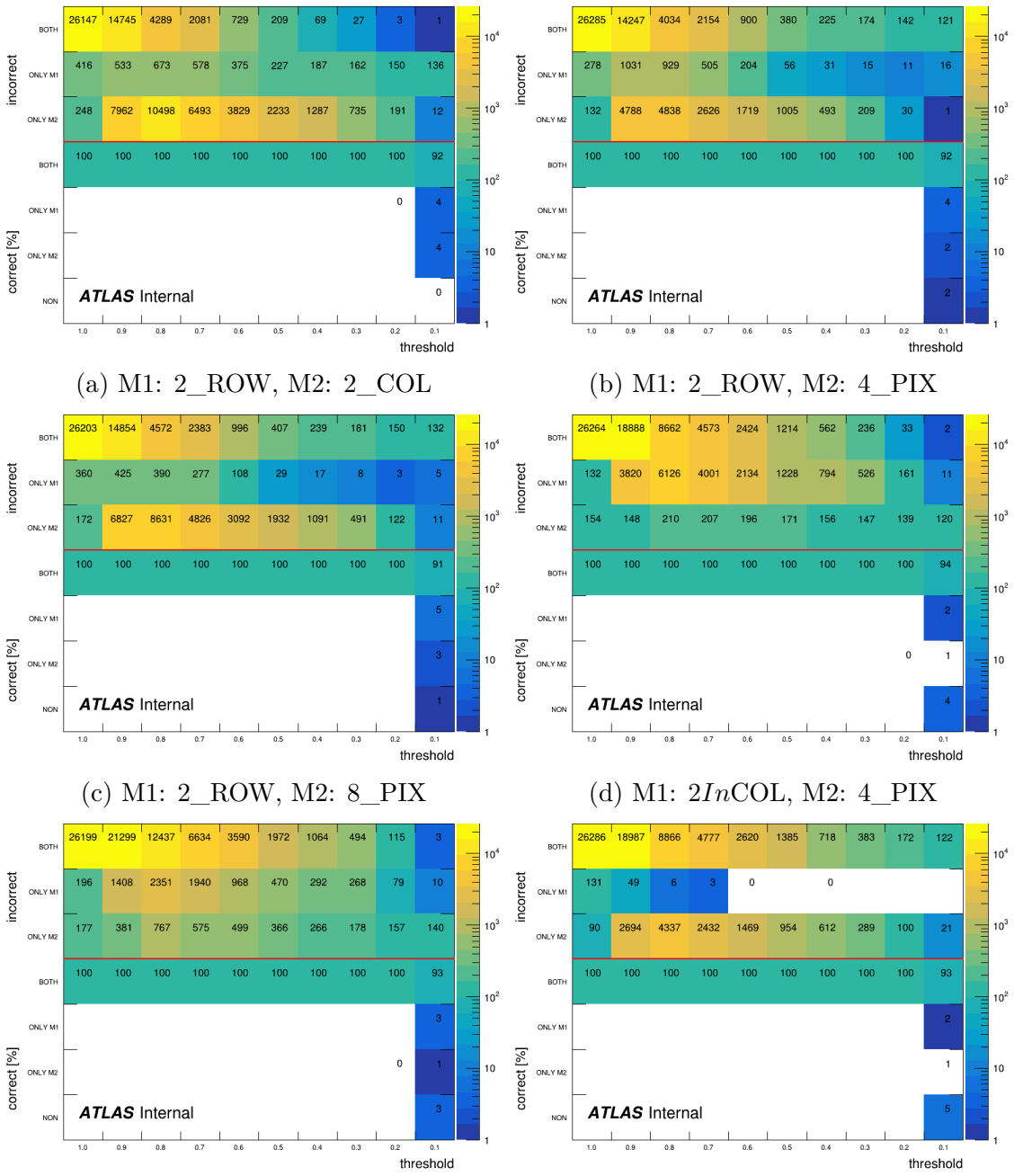
- both methods found pixel in/correctly (marked with "BOTH" on the y-axis),
- only the first/second method found pixel in/correctly (marked with "ONLY M1"/"ONLY M2" on the y-axis),
- none of the methods found pixel correctly (marked with "NON" on the y-axis).

Figure A.10 shows results for 5 LB whereas Figure A.11 those for 100 LB hit maps. On the first plot, the FIT method was skipped since it uses larger hit maps only. The most important results on these plots are those in rows marked as "BOTH" since one wants to use two methods. For the results obtained with smaller hit maps, the largest improvement was achieved by using combinations of 2_ROW and 2_COL, where for the largest threshold the number of incorrectly found pixels dropped to 28.

In Figure A.11 the 2_ROW method was omitted since for this method only 4 pixels were found incorrectly already for the threshold of 0.9 (see Figure A.9). Therefore, the 2_ROW combined with any other method would give very good results. Presented combinations with the FIT method give the best improvement. For example the combination of the 4_PIX and FIT methods where the number of the incorrectly identified pixels drops to 0 already for the threshold of 0.8.

The last item to be discussed is the efficiency of LEFF pixels and their visibility in different algorithms. The results shown previously were obtained for random LEFF pixels with the efficiency set to 10%. It is clear that this value will not be the same for all inefficient pixels in data. Therefore it is worth checking which pixel efficiencies will be visible by studied algorithms. Incorrect identification of LEFF pixels does not depend on the set efficiency and will not be discussed further.

In Figure A.12 the mean percentage of the correctly found pixels is plotted. On the x-axis, there is a threshold whereas on the y-axis the used efficiency for LEFF pixels. Within the threshold range of 1.0 to 0.9 most of the pixels are found correctly with efficiency close to 100% for 2_COL, 2_ROW, 4_PIX and 8_PIX methods. The FIT method provides satisfying results for the thresholds 10–20 for most of the efficiencies. However, it must be



(a) M1: 2_ROW, M2: 2_COL

(b) M1: 2_ROW, M2: 4_PIX

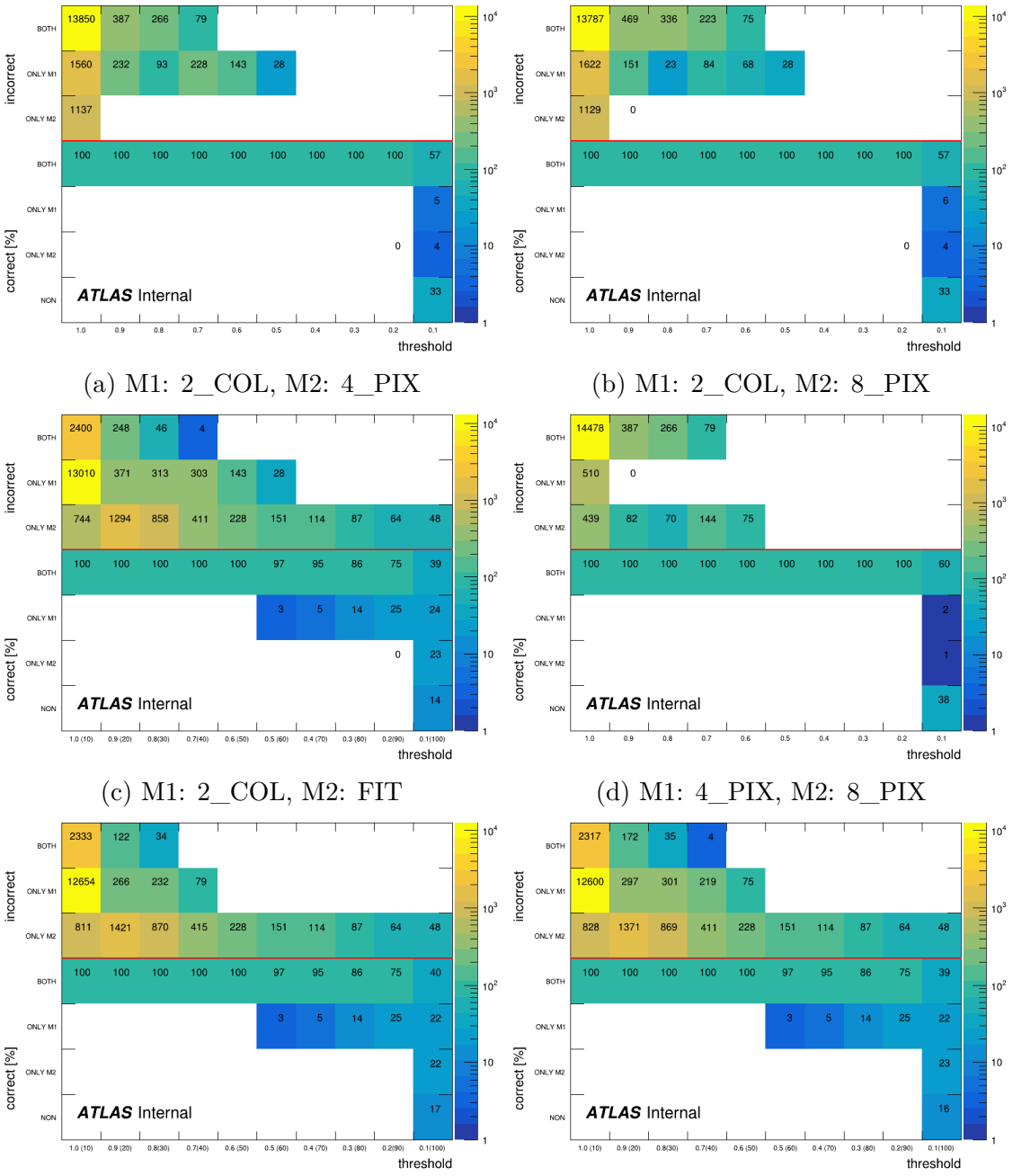
(c) M1: 2_ROW, M2: 8_PIX

(d) M1: 2InCOL, M2: 4_PIX

(e) M1: 2_COL, M2: 8_PIX

(f) M1: 4_PIX, M2: 8_PIX

Figure A.10: Numbers on the plot were obtained by calculating the mean using the results of multiple tests on 5 LB hit maps. Three top rows are filled with the number of incorrectly found pixels, whereas the four bottom show the percentage of correctly found pixels. The rows marked with the label "BOTH" mean that both analysed methods found a pixel in/correctly, labels "ONLY M1"/"ONLY M2" state that only the first/second specified method found a pixel in/correctly and the rows marked as "NON" mean that neither method found a pixel correctly. On the x-axis, there is the threshold used for each method.



(a) M1: 2_COL, M2: 4_PIX

(b) M1: 2_COL, M2: 8_PIX

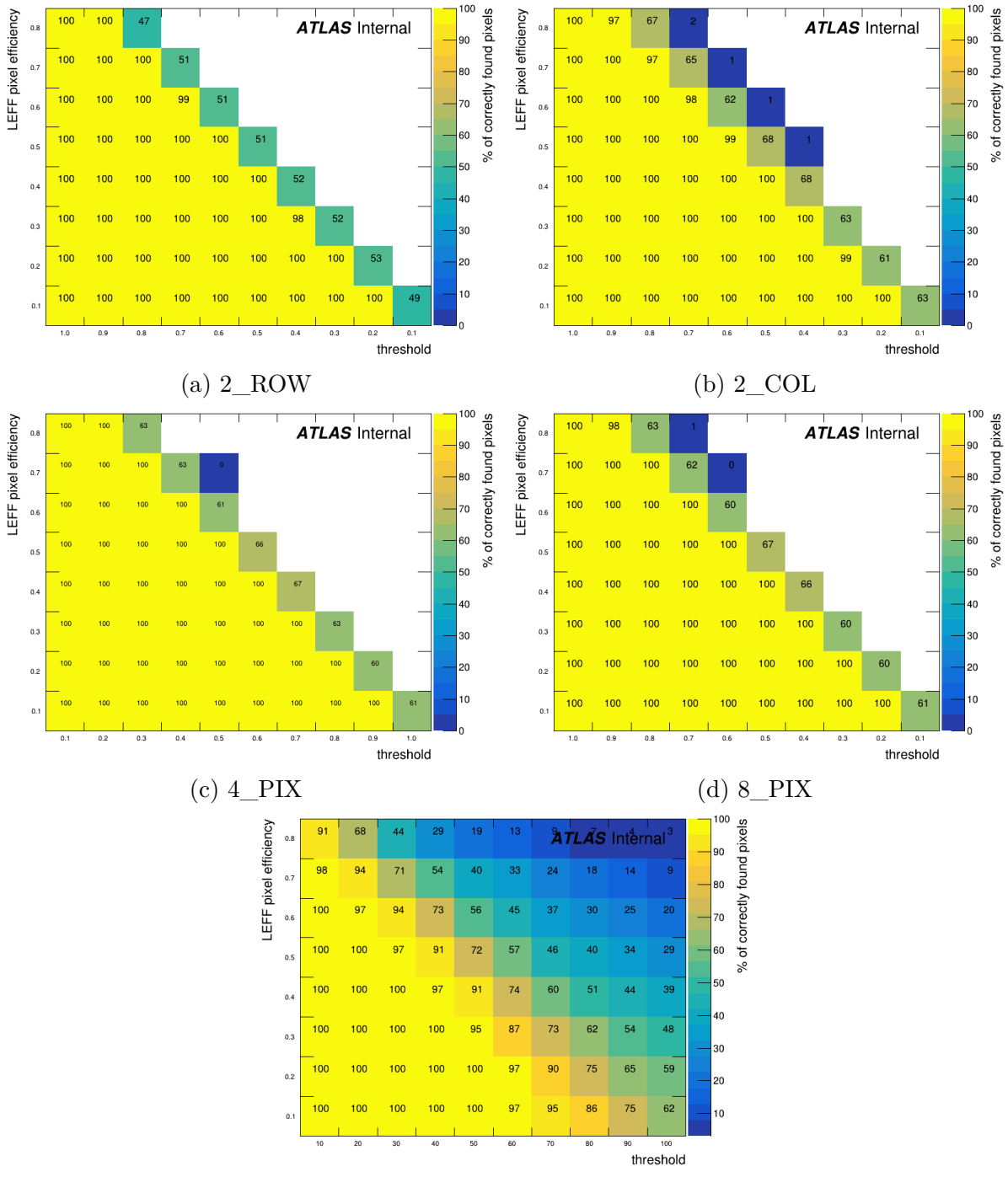
(c) M1: 2_COL, M2: FIT

(d) M1: 4_PIX, M2: 8_PIX

(e) M1: 4_PIX, M2: FIT

(f) M1: 8_PIX, M2: FIT

Figure A.11: The x-axis: the threshold used for each method. the z-axis: the mean percentage of the correctly found pixels and the mean number of incorrectly found ones. The meaning of the y-axis labels: "BOTH"– both analysed methods found a pixel in/correctly, "ONLY M1"/"ONLY M2" – only the first/second specified method found a pixel in/correctly, "NON" – neither method found a pixel correctly. Results are for analysis with 100 LB hit maps.



(a) 2_ROW

(b) 2_COL

(c) 4_PIX

(d) 8_PIX

(e) FIT

Figure A.12: Mean percentage of the correctly found pixels from all tests. On the x-axis different thresholds, and on the y-axis efficiencies of the LEFF pixels.

noticed that for such a range of the threshold, the number of incorrectly found pixels is high for all methods except for the 2_ROW method. Therefore, they should not be used alone. Additionally, the 2_COL method maintains the best efficiency in comparison to other methods (even for the threshold of 0.8).

A.4.4 LEFF Pixels in Data

It is time to see the performance of discussed algorithms on real data. As it was discussed there is a strong dependence between algorithms performance and the amount of data used for LEFF identification. Therefore, for the following studies hit maps containing data from 100 LBs run were considered and used.

The algorithms: 2_COL, 2_ROW, 4_PIX, 8_PIX were run over data gathered during 2017 data taking in run 331020 by Station A NEAR Plane 1. The threshold for those methods was set to 0.9 allowing identification of pixels with 90% efficiency or less. All of the algorithms identified one pixel as LEFF: (286,25) with an efficiency of 11%.

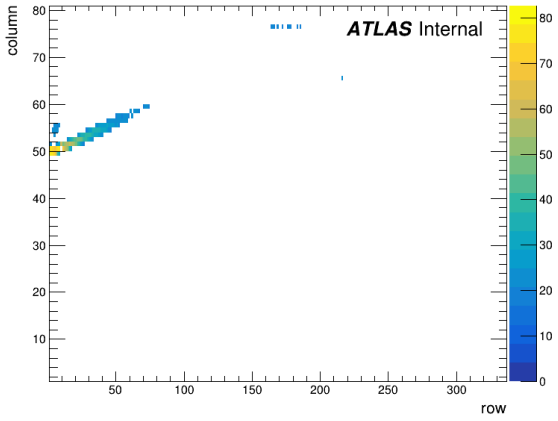


Figure A.13: The LEFF pixels found by the FIT method for the A Near Plane 1. On the z-axis, there is a value of the calculated $RATIO_{FIT}$.

The combined results of two different algorithms were discussed for tests. The next step is to check how they perform on the experimental data. The 2_ROW method was able to find all LEFF pixels in tests with only 4 incorrectly identified pixels for the threshold of 0.9. As it was mentioned before the hit maps used for tests were biased to some degree which is why it is recommended to use the 2_ROW method with another one. It was discussed before that the FIT method provides the best results combined with other methods. The disadvantage is that the fits are not always perfect and in some cases, this method can not provide the correct fit function. Because of this method's sensitivity to the provided irregular shape of the diffractive pattern, it is discouraged to use it. Instead, a combination of the 2_ROW and the 2_COL methods was used. The 2_COL method was chosen because it was providing the second best results and it's using a completely different set of

pixels than the 2_ROW method, which will be a good cross-check. For this combination of methods the list of found LEFF pixels is presented in Table A.1.

Table A.1: Pixel ID (row,column) for found LEFF pixel in run 331020 with combined methods 2_ROW and 2_COL with the threshold of 1.4.

Station	**Plane**	**Pixel ID (row,column) and Pixel Efficiency [%]**
A FAR	P0	(2,2)[3.9%] (3,2)[1.5e+02%] (3,4)[4.3%] (4,2)[3.5%] (6,2)[3.4%]
	P1	-
	P2	(2,77)[8.7%] (2,78)[2.7%] (3,77)[10%] (3,79)[4.5%] (4,79)[6.4%] (18,46)[0.036%] (70,30)[52%] (93,63)[11%] (96,68)[23%] (110,35)[0.18%] (127,1)[4.5%] (170,12)[0.33%] (237,13)[34%] (270,58)[27%] (274,58)[37%]
	P3	-
A NEAR	P0	(272,11)[44%]
	P1	(61,18)[36%] (216,65)[26%] (268,25)[11%]
	P2	(23,58)[47%] (117,78)[50%] (156,15)[51%] (235,42)[0.78%]
	P3	(9,67)[39%] (136,33)[12%] (181,45)[0.36%]
C NEAR	P0	(131,72)[39%] (151,33)[19%]
	P1	(29,65)[20%] (88,43)[17%] (292,51)[22%]
	P2	(3,53)[30%] (123,30)[32%] (226,1)[32%]
	P3	(44,10)[41%] (46,22)[24%] (95,24)[15%] (108,32)[19%] (139,1)[40%] (220,36)[48%] (221,22)[19%] (235,7)[32%] (272,19)[49%] (299,17)[36%] (307,37)[19%] (332,28)[23%] (332,68)[39%]
C FAR	P0	(50,69)[43%] (75,23)[14%] (180,68)[43%] (229,30)[12%] (275,13)[30%] (295,70)[31%]
	P1	(24,2)[7.4%] (32,38)[55%] (32,39)[3.1%] (49,28)[3.3%] (51,53)[3.3%] (88,21)[46%] (106,30)[48%] (144,20)[23%] (150,61)[5.8%] (157,44)[9.5%] (322,79)[53%]
	P2	(1,5)[52%] (19,29)[43%] (53,45)[42%] (85,36)[48%] (169,46)[16%] (189,59)[21%] (298,23)[13%] (300,4)[14%] (324,48)[17%]
	P3	(5,25)[0.69%] (293,55)[44%]

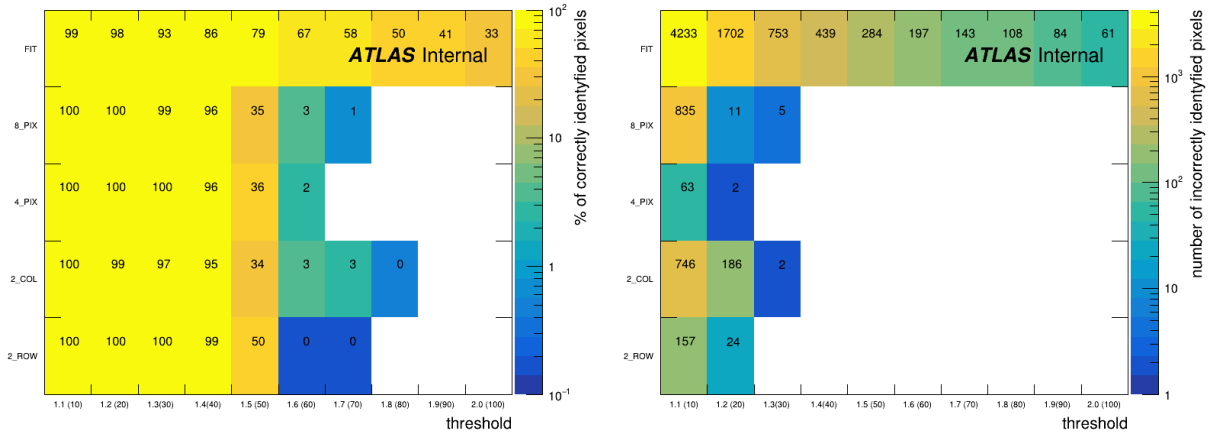
A.5 Noisy Pixels

The last group of faulty pixels which will be discussed are noisy ones. These pixels are expected to deliver more signal and should be visible as pixels with higher multiplies than their neighbours.

A.5.1 Tests

As previously, the test hit maps were created. In this case, the tests based on 100 LB hit maps were performed. The applied algorithms were the same as for the LEFF identification but with the ratio expected to be higher than one ($RATIO_{method} > 1$). The same threshold ranges were used for noisy pixels as for LEFF. The 100% efficiency of the pixel means that the pixel always fired when it should, whereas the efficiency of the pixels sometimes failing to give a signal is lower than 100%. For the technical aspects of this analysis, the efficiency above 100% means that the pixel gave the signal even when it was not supposed to – noisy pixel. For the first test, the efficiency for artificially added noisy pixels was set to 150%.

The results are shown In Figure A.14. The left plot presents the mean percentage of the correctly found pixels for each method (y-axis) and different thresholds (x-axis). The thresholds for the FIT algorithm were written in brackets. The right plot presents the number of incorrectly identified pixels for different thresholds and methods. All algorithms provide good results up to the threshold value of 1.4 (the FIT method up to the threshold of 50) and already at the threshold of 1.4, no incorrect pixels were marked for the 2_COL, 2_ROW, 4_PIX and 8_PIX methods. The FIT method provides many incorrectly assigned pixels in the whole range and it is not recommended for noisy pixel identification.

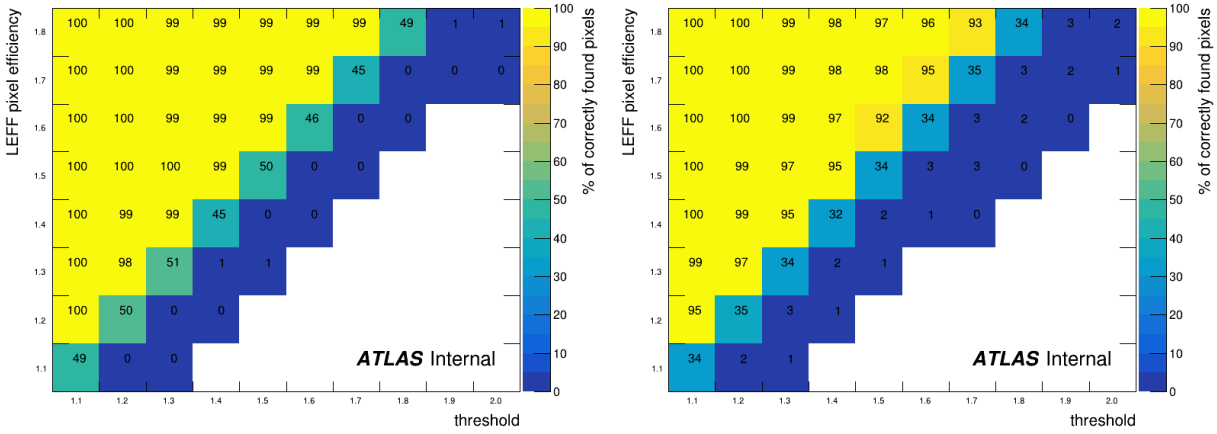


(a) Percentage of correctly found NOISY pixels. (b) Number of incorrectly found NOISY pixels.

Figure A.14: The x-axis: threshold used for NOISY pixel identification, the y-axis used method. Left: the mean percentage of correctly found pixels from all tests. Right: the mean number of incorrectly found pixels from all tests.

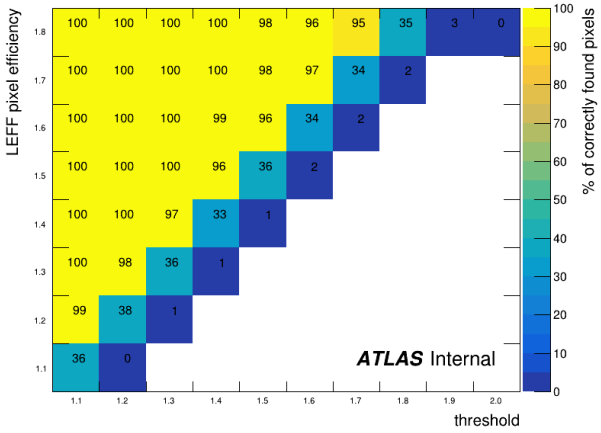
Now, same as in the LEFF pixel identification, the performance of studied methods with various thresholds and noisy pixel efficiency will be checked.

In Figure A.15 the results for different noisy pixel efficiencies for different methods are presented. The set efficiencies varied from 110% to 180% (the higher efficiency pixels are expected to be always seen by algorithms). It is understandable that with higher efficiency the performance of algorithms will be better even for higher thresholds.

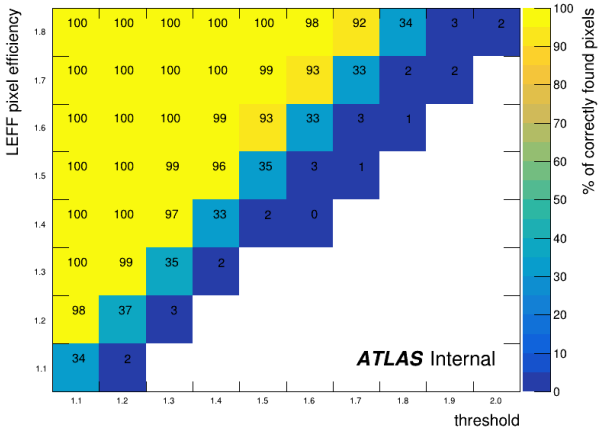


(a) 2_ROW

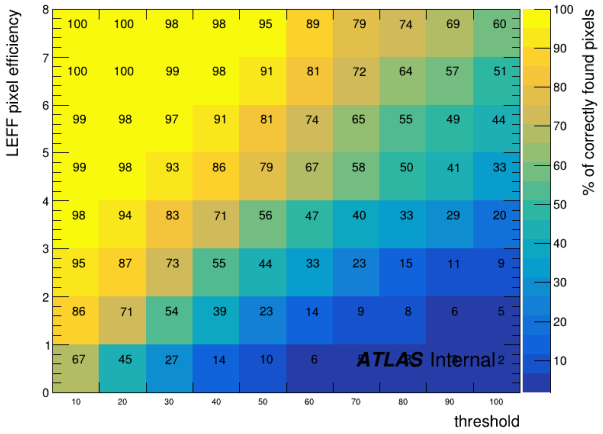
(b) 2_COL



(c) 4_PIX



(d) 8_PIX



(e) 8_PIX

Figure A.15: For each method: mean percentage of correctly found pixels from all tests. On the x-axis different thresholds, and on the y-axis efficiencies of NOISY pixels.

A.5.2 Noisy Pixels in Data

Finally, the noisy pixels were identified using the 2_ROW and 2_COL methods with the expected noisy pixel minimum efficiency of 120% (1.2 threshold). In the studied 331020 run, only a couple of noisy pixels were found.

In the studies of the experimental data, it was noticed that noisy pixels appear more frequently near the LEFF pixels. This is due to the shifts of part of the signal from the LEFF pixel to the neighbouring one. An example is shown In Figure A.16. Here, one can easily spot pixels which give more signal than their neighbours (e.g. pixel (4,1) belonging to the 0 plane). For such situations, a different approach had to be taken in order to correctly identify noisy pixels. First, the LEFF pixels are identified. Then during the calculation of the average for noisy pixels search, the LEFF pixels are excluded and in case this would lead to no available pixels for calculation average was set to be 1. For example, when estimating whether the pixel (2,77) on the right plot in Figure A.16 is noisy, the 2_ROW algorithm will take only pixel (3,77) into account.

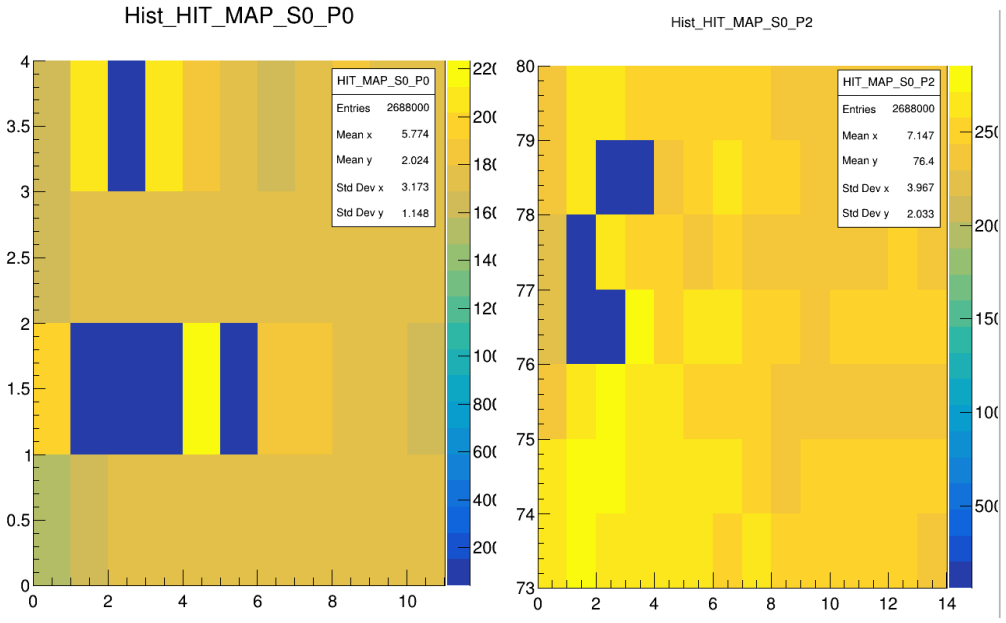
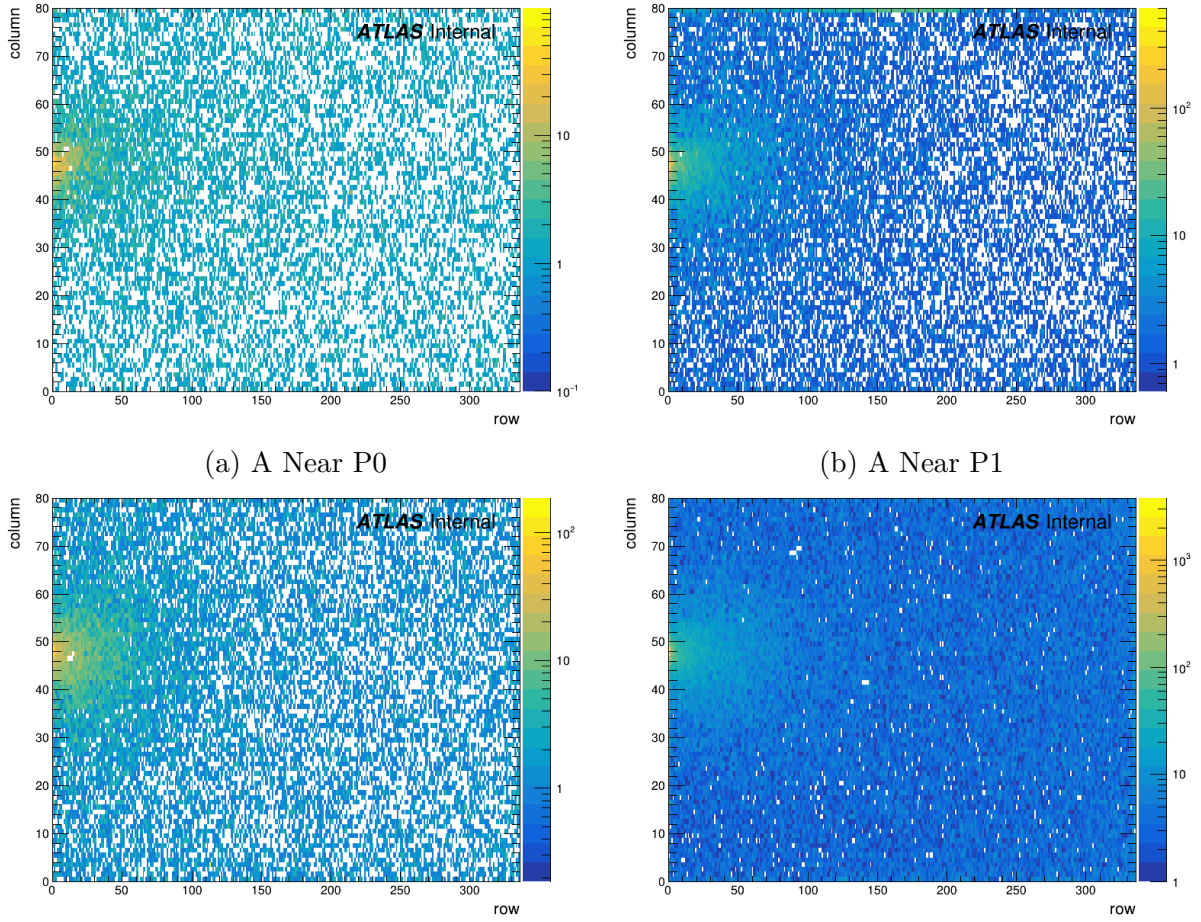


Figure A.16: Part of hit map of station A Far plane 0 (**left**) and plane 2 (**right**) in run 331020.

A.5.3 Noisy Pixels – Events without a Reconstructed Track

Another way of searching for noisy pixels is to analyse events in which no track is reconstructed but the fired pixels are still present. Since noisy pixels are expected to fire when there is no passing proton such selection should allow the detection of noisy pixels. Figure A.17 contains results for described method. The number of events where a pixel fired

without a reconstructed track is significantly smaller in comparison to the number of all events without a reconstructed track (for one station the number varied from 14 069 273 to 20 630 689 events). Only the pixels on the edge of the diffractive pattern and the detector edge have higher multiplicity, but such behaviour is expected and those should not be considered noisy. Therefore it seems that this method will not be useful for noisy pixel identification.



(a) A Near P0

(b) A Near P1

(c) A Near P2

(d) A Near P3

Figure A.17: Pixels that fired in the events without reconstructed track for Station A Near.

A.6 Track Method

Another way to distinguish the dead and LEFF pixels is based on the reconstructed tracks. In principle, it can be assumed that protons are passing thru all 4 planes in each station. The pixels connected to the track should be the same ones which gave the signal. The dead pixels in this method will never give a signal when the track points at them, whereas the

LEFF pixel will be active only in some percentage of events. In this study, the tracks were reconstructed using an algorithm from AFPAnalysisToolbox.

In the analysis, only events with maximally two hits in each plane were considered. This is due to the geometry of the Silicon tracker (see [63]), where in most cases, two pixels will give a signal. Additionally, the track must have been reconstructed and if two pixels fired they had to be neighbours. This condition excluded events in which one pixel was fired by a passing proton and another one was triggered by noise.

For such events, it was checked whether the pixels indicated by the reconstructed track are the ones that fired. If the pixel position did not match then the pixel indicated by the track was marked as "faulty" in this particular plane.

The mentioned event selection results with four possible scenarios of hits in a plane, shown In Figure A.18. In the first case, a proton may go through the detector at such an angle that only one pixel will be hit and the reconstructed track will point to this pixel (Figure A.18 situation in P0). However, in most cases, the track will be pointing at both of them (Figure A.18 situation in P1). It is worth mentioning that the pixel will fire (see 4) when a minimal amount of charge, usually $2000e^{-}$, is induced in the pixel. It was found that this amount of the charge corresponds to about 30 μm of the effective length in the pixel. Thus, the calculated pixels from the track had to have at least 30 μm of the track length to be counted as predicted from this track (since the shorter ones would cause the pixel to not fire, due to too little charge left in the singular pixel). In the case a pixel is faulty, there can be no hits in a given plane (Figure A.18 situation in P2) or there will be one hit instead of two (Figure A.18 situation in P3).

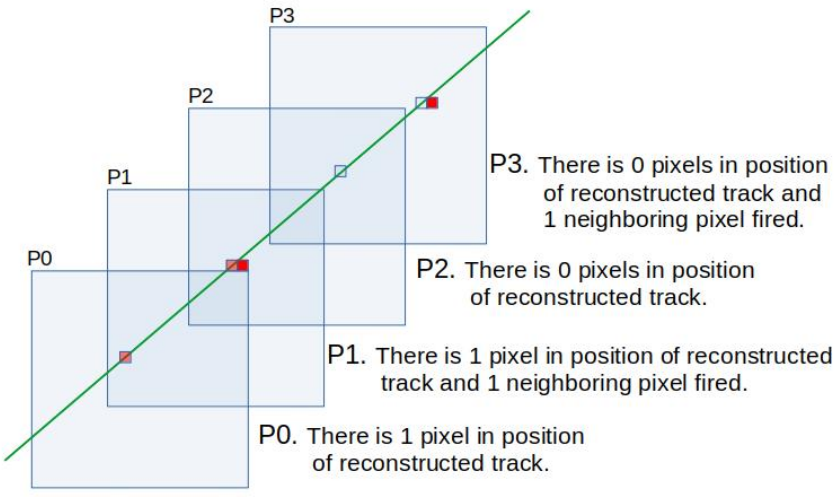
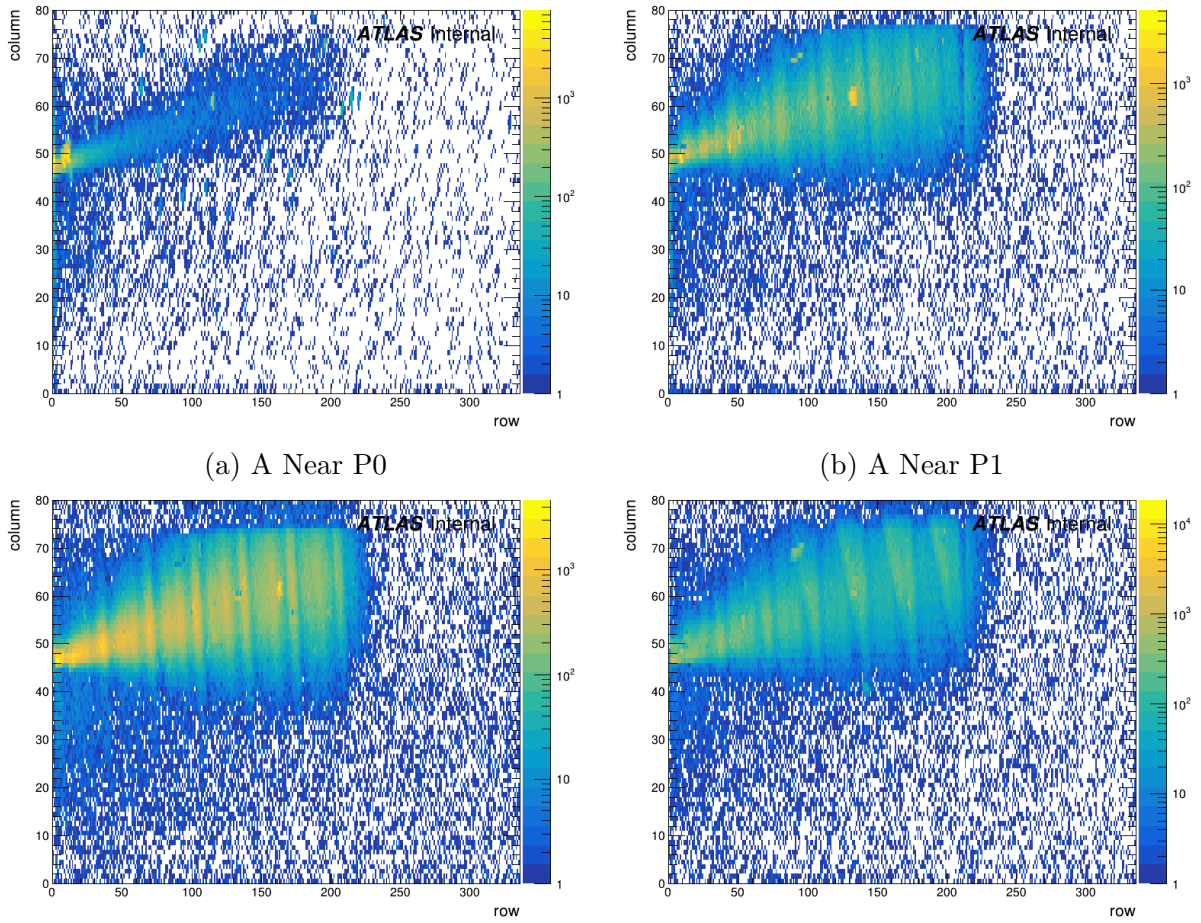


Figure A.18: A scheme showing possible scenarios of hits in AFP planes for the reconstructed track.

One should note that the track method is especially sensitive to the local alignment.

As the multiple hits may worsen the track reconstruction and, in consequence, method efficiency. The results are presented in Figure A.19. For the results presented here, the pixel was marked as dead if in at least 90% of events it did not give a signal in the plane but the track was passing through it. The pixel was marked as LEFF when the percentage of events was within the range of 30%–90%. The comparison of this method with other used methods will be discussed in the next section.



(a) A Near P0

(b) A Near P1

(c) A Near P2

(d) A Near P3

Figure A.19: Results of the track method for station A NEAR in run 331020. On the z-axis, there is the number of events in which the pixel did not fire but the track was pointing to it.

A.7 Summary

A.7.1 Dead Pixels

This work focused on the identification of defective pixels in AFP SiT detectors. For all algorithms (the dead and LEFF pixel identification) the tests were performed using artificial hit maps created using different amounts of data. One can consider looking for defective pixels during the data taking, In such a case, smaller hit maps would be preferred since the reaction time for potential problems should be quicker. For the offline analysis, larger samples may be used which will assure more accurate results.

The most important limitation of the dead pixel finding method is the range in which the method is effective for smaller hit maps. The tests give perfect results (all dead pixels found with no incorrectly identified pixels) for both smaller (within the range) and larger hit maps.

This algorithm was able to find all masked pixels[6] with few additional dead ones. Additionally, another method for the dead pixel finding was introduced – the track method discussed in A.6. This method was able to identify some groups of dead pixels but it was not efficient for singular ones.

In Table A.2 the number of dead pixels either found by both algorithms or by only one of them is listed in the first column. The numbers in the table are given for each plane in station A Near and similar numbers were obtained for other AFP stations. The algorithm for finding the dead pixels finds all of the masked pixels in all planes plus a couple more. The tracking method, however, does not perform as well as this algorithm and there is a large discrepancy between those two methods.

Table A.2: Number of dead pixels found either by one of the methods listed in the first column or by both of them. For station A Near.

Algorithms	Number of pixels marked as dead by both algorithms				Number of pixels marked as dead by only I algorithm				Number of pixels marked as dead by only II algorithm			
	P0	P1	P2	P3	P0	P1	P2	P3	P0	P1	P2	P3
I dead pixel finding alg. II track method	0	0	1	0	126	20	19	21	2	1	18	29
I dead pixel finding alg. II masking procedure	125	13	16	16	1	7	4	5	0	0	0	0
I track method II masking procedure	0	0	1	0	2	1	18	29	125	13	15	16

[6]Pixels which were set off during the calibration procedure and their list is well known.

A.7.2 LEFF Pixels

The next analysed group of faulty pixels are the LEFF pixels. Here only four methods (2_COL, 2_ROW, 4_PIX and 8_PIX) were used for tests with the hit maps created using 5 LB of the data. It was found that the combination with the 2_ROW method is providing the best results for the LEFF pixel identification (highest efficiency and smallest possible number of incorrect pixels). In the tests using larger hit maps (100 LB), the 2_ROW method identified incorrectly very few pixels for the threshold of 0.9 (corresponding to finding pixels with efficiencies 90% or lower). The other algorithms could be improved using the FIT method, which can lower the number of incorrectly identified pixels. However, this method is very sensitive to the type of the provided hit map and it is not recommended. Lastly, the performance of the algorithms was checked assuming different efficiencies of the pixels (varied between 10% and 80%). In this study, the 2_COL method (apart from 2_ROW method) was maintaining the best LEFF pixel identification. Therefore in the experimental data studies, the 2_ROW and 2_COL methods with the threshold value of 0.9 were chosen.

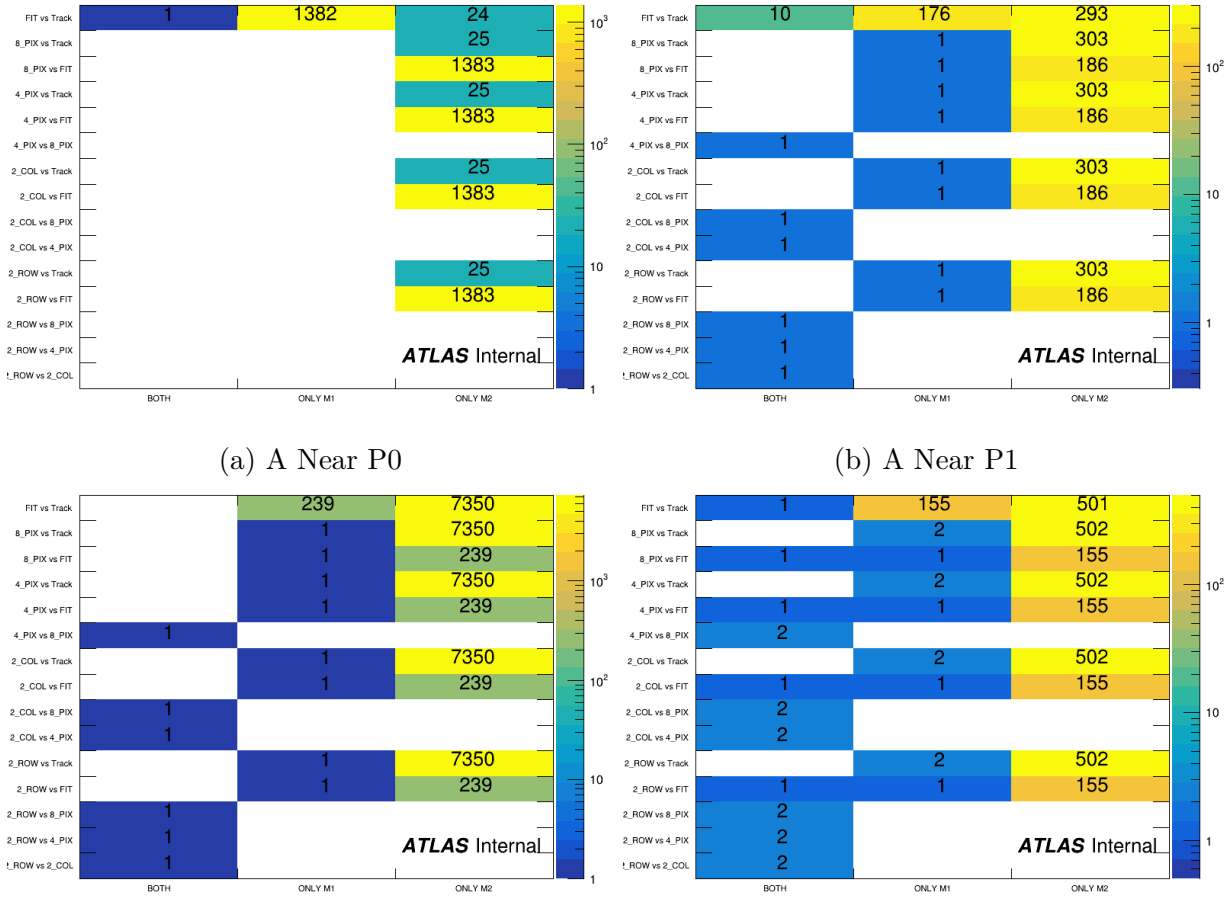
The LEFF finding was also performed using the track method. The pixel was marked as LEFF when in more than 30% of events the pixel did not give the signal when the track was pointing at it. Similarly, for the dead pixel studies, this method often was finding a lot of pixels which were not visible as LEFF on the hit map.

To conclude the comparison of all methods' performance is presented in Figure A.20. It shows the number of pixels found by both, only the first or the second of the compared methods. Here, all methods except for the track and FIT methods were finding few LEFF pixels.

A.7.3 Noisy Pixels

The last group of faulty pixels are the noisy ones. To find them the same methods were applied as in the case of the LEFF pixels. The 2_COL, 2_ROW, 4_PIX and 8_PIX methods provided much better results than the FIT method. The number of the incorrectly identified pixels was much smaller and for the threshold of 1.4, there were no such pixels found. Whereas the FIT method was still identifying incorrectly tens of the pixels. For the four methods using the neighbouring pixels, the use of the threshold value of 1.4 allowed observation of pixels which were set to give a signal 50% more than expected (the efficiency of 150% total).

Similarly to the LEFF pixels the artificial hit maps were introducing some bias and using only one method was not sufficient. In the experimental data analysis, two methods 2_COL and 2_ROW were used. An interesting observation was made. It was found that the noisy pixels will appear near the LEFF pixels because of the charge sharing. Due to this fact, the LEFF pixels were excluded during the calculation of the average for these



(a) A Near P0

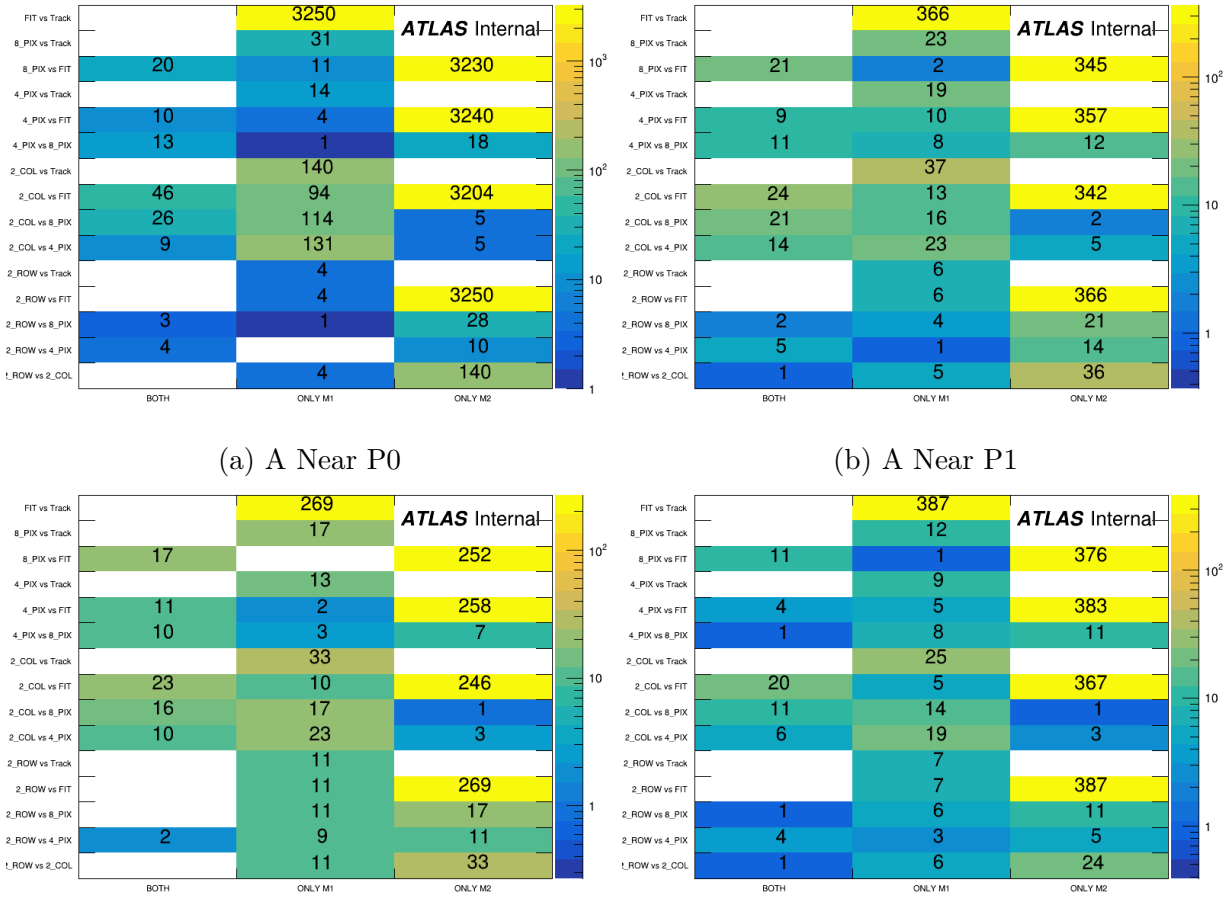
(b) A Near P1

(c) A Near P2

(d) A Near P3

Figure A.20: Number of pixels which were found by: both methods (marked as "BOTH" on the y-axis), only first method ("ONLY M1") or only second ("ONLY M2"). On the y-axis Method M1 vs. Method M2 is marked.

methods.



(a) A Near P0

(b) A Near P1

(c) A Near P2

(d) A Near P3

Figure A.21: Number of pixels which were found by: both methods (marked as "BOTH" on the y-axis), only the first method ("ONLY M1") or only the second ("ONLY M2"). On the y-axis, Method M1 vs Method M2 is marked.

The last method of noisy pixel identification is to use events in which the track was not reconstructed but the pixel fired. Only the pixels at the edge of the diffractive pattern and pixel detector have higher multiplicity but it was anticipated and those should not be considered noisy.

Lastly, the comparison of all methods of noisy pixel identification is presented in Figure A.21. The results for dead, LEFF and Noisy pixels in 2017 low μ runs are presented in ATLAS internal note: [77].

Appendix B

Validation of the AFP HLT Exclusive Jet Trigger

The amount of data produced at the Large Hadron Collider puts scientists in a very challenging position. During Run 2, the data were produced at the peak crossing rate of 40 MHz. Furthermore, considering the size of the raw data event (around 1 MB), a very large amount of data would have to be stored for further analysis. Fortunately, many events are not needed to be stored. For this purpose dedicated electronics and software algorithms, called triggers, were created. More information about the ATLAS trigger system can be found in Section 3.6.

For a given physics analysis dedicated L1 and HLT triggers are usually needed. The following is devoted to one of the HLT trigger algorithms selecting the exclusive jet production processes. Exclusivity means that all final state particles can be measured. This determines strong kinematic constraints between the central state and the forward protons.

The centrally produced jets can be detected by the ATLAS main detector, but the forward protons are out of its range and can be registered by the AFP detectors (see Section 4).

B.1 Backgrounds

The main backgrounds to the exclusive jet production are non-diffractive and single diffractive jets overlaid with forward protons originating from the pile-up.

As was shown in [76] the highest background rate comes from the non-diffractive jet production. These are triggered at the ATLAS L1 by a dedicated algorithm described in [78]. The efficiency of the single-jet triggers as a function of transverse momentum of the offline jets for different L1 triggers is shown in Figure B.1. As one can conclude, a fully efficient L1 jet algorithm for the offline jets with p_T >150 GeV is L1_J20 and L1_J50 (red and blue line). L1_J75 (purple line) starts to be fully effective for jets with p_T >180 GeV

and L1_J100 (green line) for jets with p_T >230 GeV.

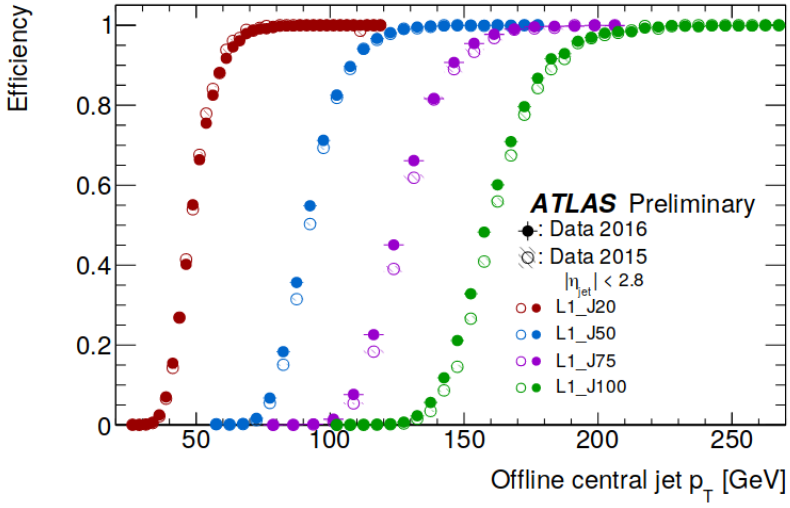


Figure B.1: Efficiency of the single-jet triggers as a function of the offline jet p_T for different L1 triggers in the central region. From [79].

The lowest not prescaled trigger during the standard 2017 and 2018 data taking was L1_J100. Furthermore, this threshold may be even higher in Run3. Therefore, for an efficient selection of exclusive jets, the threshold at the L1 must be lowered. This issue may be solved by having a dedicated L1 topo algorithm.However, since this is not a subject of this work, it will not be discussed further.

At the HLT, in the Run 2 menu, the lowest not prescaled di-jets trigger had a threshold of about 400 GeV [80] – much higher than the one needed for the studies of the exclusive jet production. Therefore, there was a need to create dedicated HLT algorithms.

B.2 Exclusive Jet Trigger Algorithms

A complete trigger algorithm for the exclusive jets consists of two parts at the HLT level:

1. comparison of the proton position predicted from di-jet system kinematics to the registered proton position in AFP,

2. comparison of the jet vertex reconstructed in ATLAS and the one reconstructed using the AFP Time-of-Flight system.

B.2.1 Comparison of the Proton Position Predicted from the Di-jet System Kinematics to that of the Proton Registered in the AFP

The idea is to match the position of "the proton" calculated from the jet kinematics to that of the proton registered in the AFP station. The features of jets (p_T, energy, pseudorapidity and azimuthal angle) are determined from information delivered by the ATLAS central detector.

From all jets in an event, the two with the highest p_T are taken. At this step, the common vertex is not required as it is expected that the high-p_T jets are relatively rare and should not originate from the pile-up. Using the jet kinematics, the mass, (M_{jj}), and rapidity, (y_{jj}), of the di-jet system are calculated. Next, the anticipated relative energy loss of the forward proton is calculated from the jet system taking into account the A/C ATLAS hemispheres:

$$\xi_{dijet}^{A/C} = \exp(\pm y_{jj}) \frac{M_{jj}}{2E_{beam}}.$$

One can also calculate the energy of the jet system in the A/C ATLAS hemisphere:

$$E_{dijet}^{A/C} = E_{beam}(1 - \xi_{dijet}^{A/C}).$$

Using the di-jet system energy it is possible to predict the forward "proton" positions in the AFP station. This is achieved with help of a special transport tool. The tool takes as input information on the proton energy, the position of the interaction vertex (fixed at (0,0,0) in the trigger algorithm) and the slope in the x and y directions (also fixed at (0,0) in the trigger algorithm). Also, the details of the LHC optics have to be delivered in the form of the TWISS files. The transport tool requires also dedicated parameterisation files (cf. [81]) for each settings of the LHC optics.

A diagram of the trigger algorithm is presented in Figure B.2. The first requirement for an event to pass the trigger conditions is the jet reconstruction in the central detector (TrigJetRec). Then, the jet container is passed further through a trigger item (HLTJETHypo). The second condition is the AFP track reconstruction (Trig_AFPSthe iTrkReco). It starts with the Raw Data reconstruction (AFP_DataProviderTool), followed by their digitisation (AFP_Raw2DigTool) and reconstruction of the AFtracks (AFP_SidLocRecoTool). When these two conditions are met, the AFP tracks and jet container are used as an input for TrigAFPJetAllTE where the first part of the trigger algorithm, the proton position matching (using AFP_ProtonTransportTool), is performed. This part was implemented into the ATLAS software. Studies of the properties of this algorithm are discussed below.

Currently, there are three possible conditions in the exclusive jet trigger which can be used for the event selection. Those are:

- $|x_{dijet} - x_{proton}| < 2.5$ mm,

- $|y_{dijet} - y_{proton}| < 2.5$ mm,
- $R = \sqrt{(x_{dijet} - x_{proton})^2 + (y_{dijet} - y_{proton})^2} < 2$ mm,

with the coordinates $x(y)_{dijet}$ $x(y)_{proton}$ determined from the di-jet and the AFP data. In the current default settings, only the third condition is efficiently used.

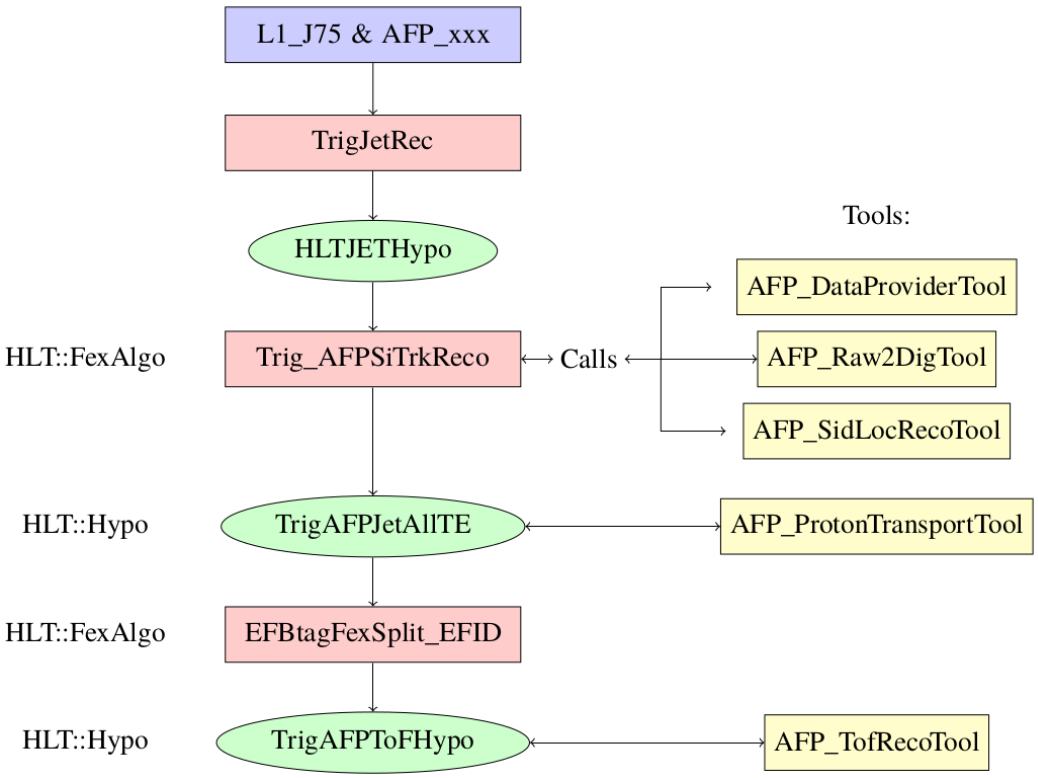


Figure B.2: Diagram of the HLT exclusive jet trigger algorithm.

B.2.2 ToF – ATLAS Vertex Matching

The second part of the exclusive jet trigger uses vertices reconstructed in the central (ATLAS) and forward (AFP) detectors. In order to reconstruct the vertex using the AFP, the difference in the time of flight of the scattered protons has to be computed:

$$V_p = (t_A - t_C)\frac{c}{2},$$

where t_A (t_C) is the time of flight measured at the AFP A (C) side. The time of flight is measured by the ToF detectors, see Sec. 4.

This trigger will use the tracking and vertexing algorithms, which will be called by the AFP trigger chain, where information is propagated via the Trigger Element. The Time-

of-Flight detector information will be reconstructed via a tool executed by the hypothesis algorithm (AFP_TofRecoTool). The cuts will be based on the difference in the positions of the reconstructed vertices. This part of the algorithm is still in the implementation phase.

B.2.3 Influence of the LHC Optics on the Trigger Efficiency

The forward proton trajectories strongly depend on the LHC optics. Since the trigger efficiency may depend on the proton position in the AFP stations, it is crucial to understand how different settings of the machine optics will influence it. In order to do so, a Stand Alone ($STAA^{p1}_{sig}$, naming convention is explained in Sec. B.3.1) MC sample was prepared. It was generated using the FPMC [82] generator and contains exclusive jets (process 16001). For the event generation, the central mass energy was set to 13 TeV, the jet rapidity was limited to (-6,6) and the minimum p_T of a leading quark/gluon was set to 150 GeV. An event was considered when it contained a proton on both sides, A and C, with $0.02 < \xi_{proton} < 0.2$, and had at least two jets with p_T >130 GeV reconstructed with either radius R=0.4 or R=1.0. The proton position matching algorithm applied in the trigger uses the data gathered by the ATLAS central detector and AFP. For analysis purposes, the AFPAnalysisToolbox was used (see Section 4.4). Studied MC samples contained information about the scattered protons and jets in ATLAS. From both of those objects, the positions of protons in the AFP stations can be calculated [1].

B.3 First Look at the Optics Properties

In order to study the levelling process influence on the reconstruction of the proton position two β^* values, typically used during Run 2, were considered: $\beta^* = 30$ cm and 40 cm. It was checked that the resulting proton positions for a given crossing angle, θ_C, are exactly the same. Therefore, all further analyses were done only for $\beta^* = 30$ cm.

The expected proton position in the AFP station for a given θ_C value is shown in Figure B.3. With the increase of the θ_C value, the y-position is getting smaller. This effect is more and more visible with decreasing value of x-AFP. Worth noticing is the point $x \sim 4$ mm, where the proton position is independent of the crossing angle. This is an expected and well-known property of collision optics and it is used for example in the dynamic alignment method [69].

In the trigger algorithm, the θ_C value is fixed. This means that due to the luminosity levelling the trigger efficiency may change. To simulate such conditions, the position of the proton predicted from di-jet (x_{dijet} and y_{dijet}) was reconstructed with the fixed value of $\theta_C =$ 140 μrad and the position of the produced proton (x_{proton} and y_{proton}) was reconstructed with $\theta_C = 140$ μrad, 160 μrad or 180 μrad.

[1] The vertex was set to (0, 0, 0).

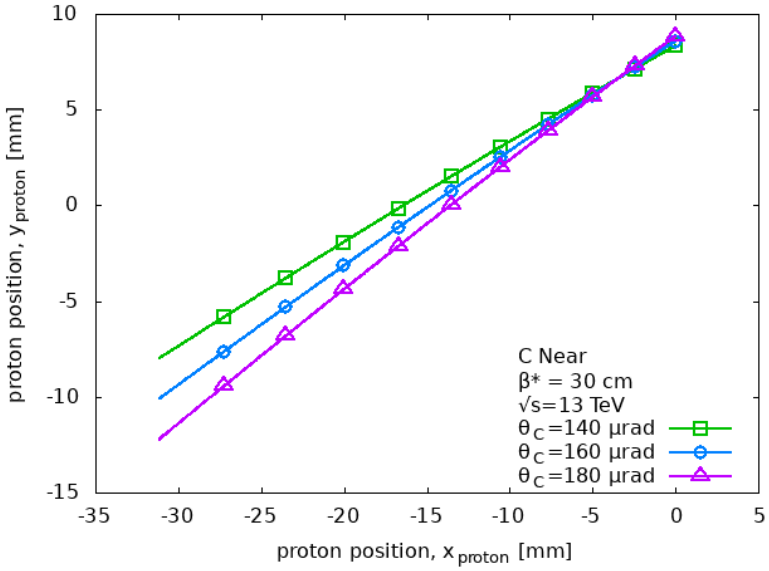


Figure B.3: Proton position in the AFP detector for β^*=30 cm and various values of the crossing angle: 140, 160 and 180 μrad.

Figure B.4 (B.5) shows the correlation between x_{proton} (y_{proton}) and x_{dijet} (y_{dijet}) for various crossing angles and anti-kt radii. The red lines represent the values of the default cuts used in the algorithm (see Section B.2.1). When θ_C starts to differ (the luminosity levelling) from the value fixed in the trigger algorithm, the correlation gets weaker for the y direction (c.f. the plots in the centre and on the bottom). This is not observed for x as the reconstruction of the x position does not depend on the vertical crossing angle. Additionally, it seems that the bigger radius (right) for the jet reconstruction provides better results than the smaller one (left).

The percentage of events which passed the trigger conditions is shown in Table B.1. When the difference in θ_C between the trigger set-up and the real LHC conditions is increasing, the trigger efficiency decreases however the changes are within the error ranges. Therefore, it can be concluded that the θ_C levelling has a minimal effect on the trigger algorithm. The performance of the algorithm with a larger radius is better by about 4%. In the case of a standalone sample without the pile-up, this is somewhat expected due to having more signal particles composing a jet.

Table B.1: Efficiency of the exclusive jet trigger for different θ_C settings and anti-kt radii.

	θ_C=140 μrad	θ_C=160 μrad	θ_C=180 μrad
R=0.4	89.4 ±1.3 [%]	88.8±1.3 [%]	87.4±1.3 [%]
R=1.0	93.7 ±1.4 [%]	93.1±1.4 [%]	91.9±1.4 [%]

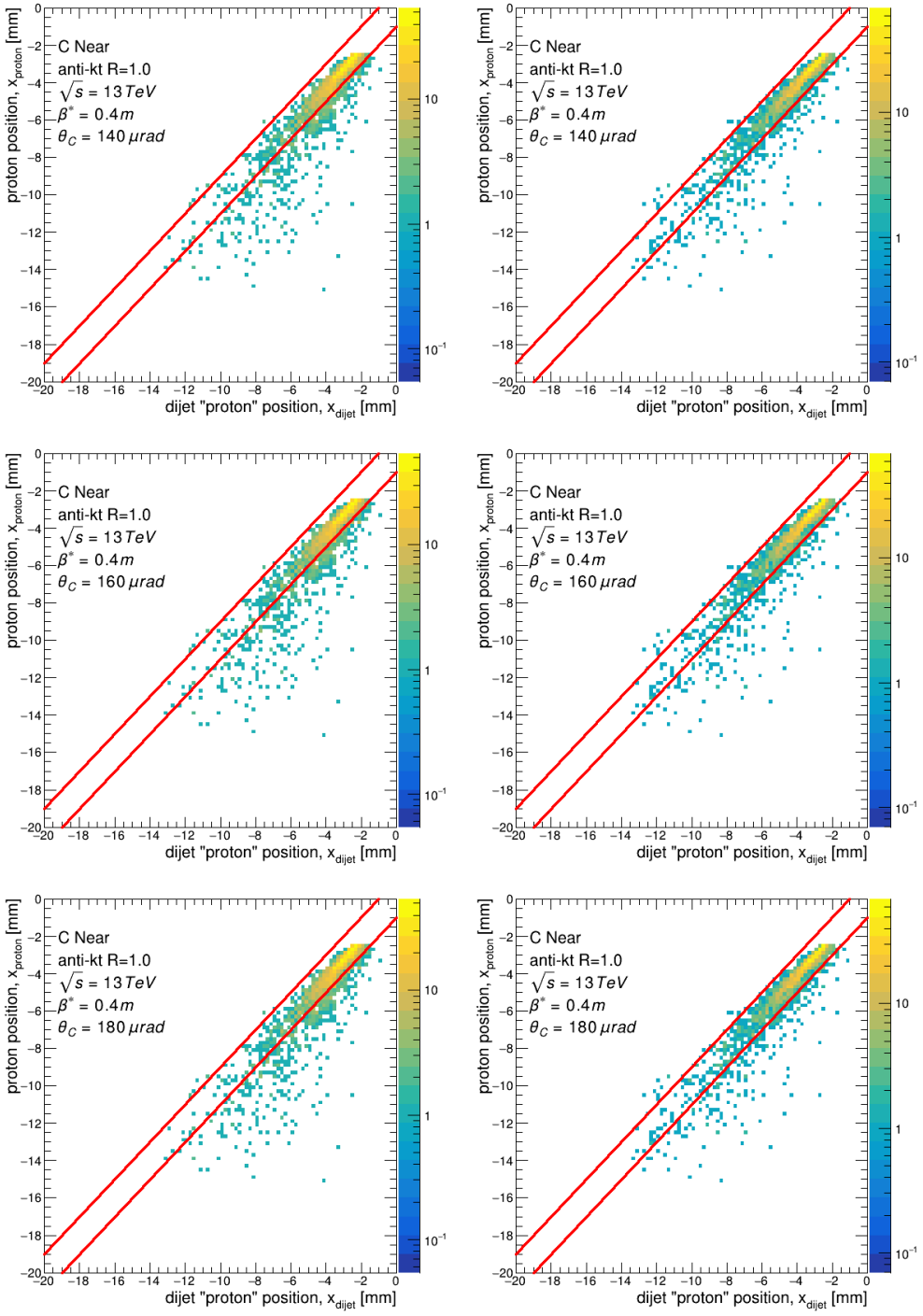


Figure B.4: Correlation between the x-position at the AFP Near station reconstructed from the di-jet (x_{dijet}) system and from the forward proton (x_{proton}). From top to bottom, the crossing angle for the reconstructed forward proton was set to 140, 160 and 180 μrad (**top**, **centre** and **bottom** plots respectively). The **left** plots are for the jets reconstructed with R=0.4, and the **right** ones for R=1.0. The lines represent the default trigger cut, R = 2 mm.

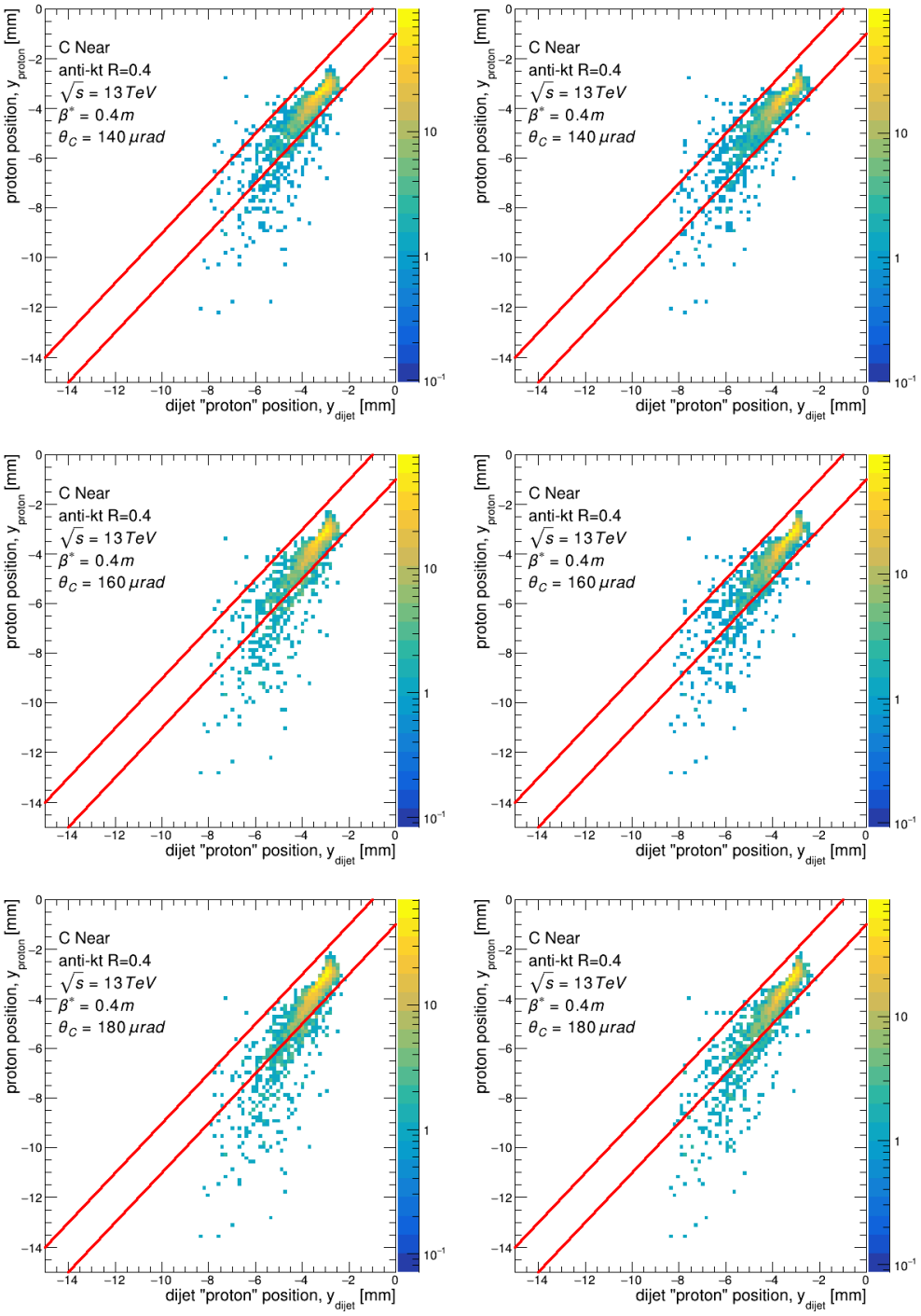


Figure B.5: Correlation between the y-position at the AFP Near station reconstructed from the di-jet system (y_{dijet}) and the forward proton (y_{proton}). From left to right the crossing angle for the reconstructed forward proton was set to 140, 160 and 180 μrad (**top**, **centre** and **bottom** plots respectively). The **left** plots are for the jets reconstructed with R=0.4, and the **right** ones for R=1.0. The lines represent the default trigger cut, R = 2 mm.

It is also worth checking whether there are regions of proton energy loss or position x/y where the trigger is more efficient and whether one can consider some additional cuts to reduce the backgrounds. The ξ regions were chosen as follows:

$$\xi = (0.02 + i * 0.01,\ 0.03 + i * 0.01), \text{for } i \in \{0, 1, ..., 9\}.$$

The x and y coordinates were matched to these regions by calculating the proton position for a given ξ value (p_T=0). The obtained values are listed in Table B.2.

Table B.2: Range of proton energy loss, position x_{proton} and y_{proton} for a given region (Ri).

		R0	R1	R2	R3	R4	R5	R6	R7	R8	R9
θ_C	ξ	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1	0.11	0.12
all	x	-2.45	-3.71	-5.01	-6.35	-7.71	-9.11	-10.55	-12.02	-13.54	-15.09	-16.69
140		-2.86	-3.49	-4.14	-4.81	-5.50	-6.21	-6.95	-7.71	-8.49	-9.31	-10.15
160	y	-2.79	-3.51	-4.25	-5.02	-5.81	-6.62	-7.47	-8.34	-9.24	-10.17	-11.13
180		-2.71	-3.52	-4.36	-5.22	-6.11	-7.03	-7.98	-8.97	-9.98	-11.03	-12.11

The efficiency shown in Figure B.6 was calculated as the number of events accepted by a trigger with a forward proton in a given region of ξ, x and y to all events having a proton in this region. The numbers in red above the x-axis show the percentage of events in a given region w.r.t. all visible in the AFP. The forward proton was reconstructed using the information about its transverse momentum which should make the proton position reconstruction more precise than the one using only the di-jet system-related information. Using the momentum in the reconstruction will occasionally cause the forward proton to be in a different x (y) region than the one corresponding to the ξ region. This can be noticed when studying the percentage of events in a given ξ/x/y region (red numbers above each bin in Figure B.6). Without migration, the numbers in each bin should be the same. In addition, it is worth mentioning that the areas further away from the beam centre tend to contain fewer and fewer protons. It is noticeable that for all regions the reconstruction using R=1.0 provides better results than the one using R=0.4 (comparison of first/second and third/fourth column).

B.3.1 Detector and Pile-up Influence on the Trigger Efficiency

Till now, the discussed trigger efficiency was based on the assumption that there is a pure signal event present in the ATLAS/AFP. In this part, the effect of pile-up is considered. It can affect the trigger efficiency due to:

- presence of additional protons in the AFP station: this may increase the chance for an event to be accepted,

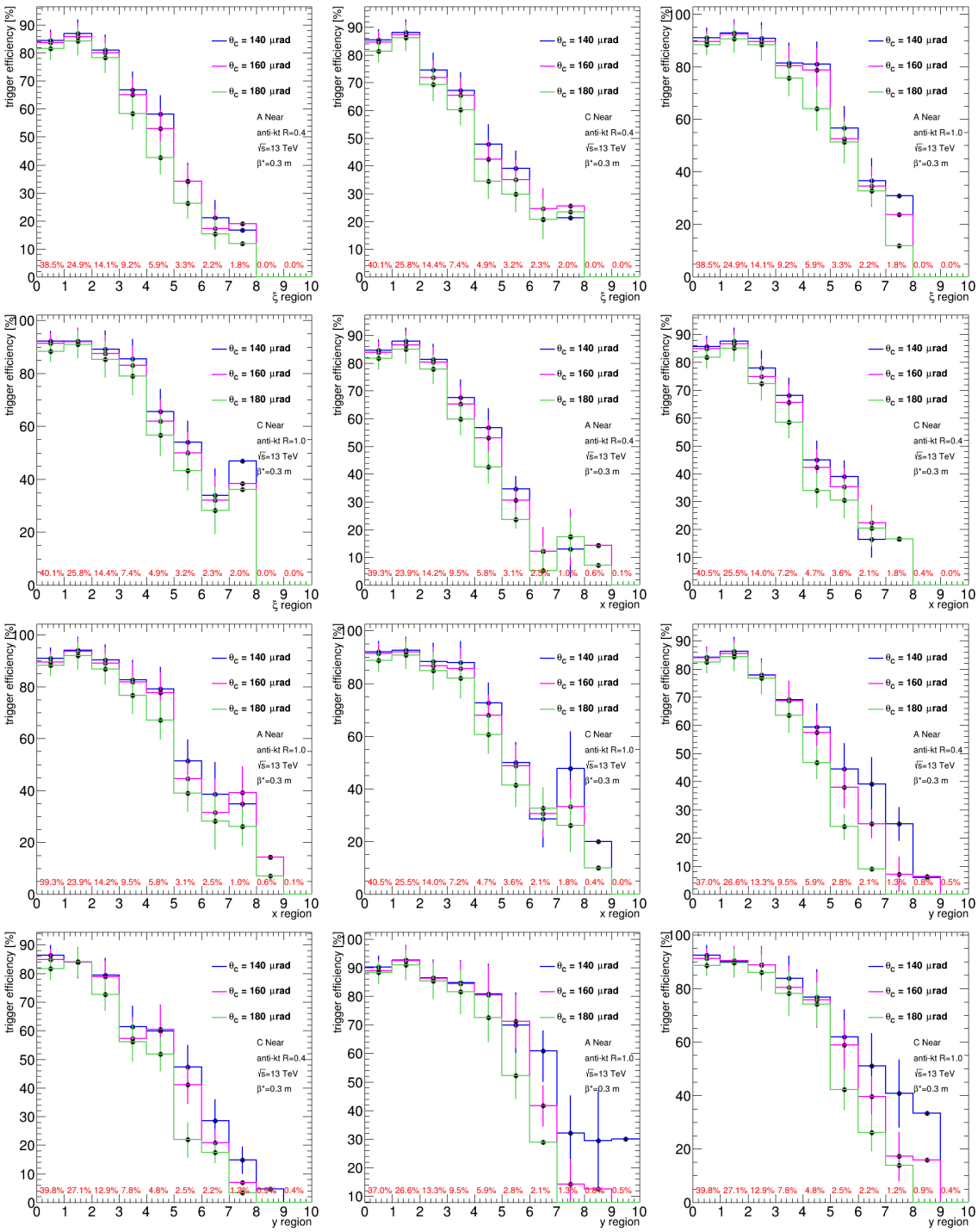


Figure B.6: The trigger efficiency in different regions of ξ, x and y (**top**, **centre** and **bottom**) for both R=0.4 (**first** and **second** column) and R=1.0 (**third** and **fourth** column) anti-kt radii and the sides A (**first** and **third** column) and C (**second** and **fourth** column). The crossing angle in the trigger algorithm was fixed to 140 μrad in all the cases whereas the crossing angle used in the reconstruction of p_{proton} was set to 140 μrad (blue line), 160 μrad (magenta line) or 180 μrad (green line). The numbers above the x-axis show the percentage of events in a given region w.r.t. all visible in the AFP.

- presence of additional particles in the central detector: this affects the jet reconstruction, worsening the predictions based on the di-jet system.

Another interesting issue to study is the influence of the detector effects. The reconstruction of the particle properties may differ from the real values either due to imperfection of the measurement or reconstruction algorithms. To make a firm conclusion concerning the trigger algorithm, it is crucial to understand the ATLAS reconstruction and its influence on trigger efficiency.

Monte Carlo Samples with Pile-up

The following analysis was based on several MC samples which have three significant set-ups. Each sample will be referred to by using those set-ups' values: $RECO^{pX}_{type}$, where:
- RECO – ATLAS for full ATLAS simulation or STAA for the stand-alone sample (particle level, jet reconstruction),
- type – sig (signal sample – exclusive jets) or bg (background sample – non-diffractive jets),
- X – value of the pile-up.

In order to study the influence of the detector effects on the signal events, the ATLAS MC sample of 5000 events and pile-up of 23 was used ($ATLAS^{pis}_{sig}$). The generation was based on the signal sample described in Chapter B.2.3. It contains the jet containers described in Sec. 3.5.3. Information about the forward protons was retrieved from the TruthProtcontainer of the signal events.

In addition, two STAA MC samples were used. First sample, $STAA^{p1}_{sig}$, was arleady described in Chapter B.2.3. The second sample contained the pile-up: $STAA^{p23}_{sig}$. In this sample, each signal event was joined with 22 SoftQCD/HardQCD events generated by Pythia 8.185 [83]. These pile-up events were merged accordingly to their cross-section. Following the Pythia recommendation, the SoftQCD and HardQCD processes were generated separately. For both, the centre of mass energy was set to 13 TeV. In addition, for the HardQCD event, a minimum invariant mass was set to 6 GeV. It is worth mentioning that all $STAA^{pX}_{sig}$ samples contained exactly the same exclusive jet events as the ATLAS samples (same seed). For example, the first event of $ATLAS^{p23}_{sig}$ contains the same exclusive process as the first event in $STAA^{p23}_{sig}$ which is helpful in the studies of the detector and pile-up effects on the signal events.

In a similar manner, two background STAA MC samples were created. Here, instead of the exclusive jets, events with at least two non-diffractive jets with $p_T > 130$ GeV were used. Furthermore, two forward protons were added (in $RECO^{pX}_{sig}$ this modification was not needed since the request was already fulfilled by the exclusive process). For a sample with pile-up=23 ($STAA^{p23}_{bg}$) a HardQCD or SoftQCD event containing a proton on side A/C (with $0.02 < \xi_{proton} < 0.2$) was added instead of usual random Hard/SoftQCD event.

Then, the rest of the pile-up events was added without any further requirements on their properties. In $STAA^{p1}_{bg}$ two protons were added to each event and no additional minimum-bias particles were included. As will be discussed later, such an approach allows emulation of the ATLAS pile-up suppression algorithms. The summary of the contents of the MC samples is presented in Table B.3.

Table B.3: Content of MC samples.

Sample	**Content**
$ATLAS^{p23}_{sig}$	1 exclusive jet process in each event 22 processes from pileup (standard ATLAS simulation)
$STAA^{p1}_{sig}$	1 exclusive jet process in each event
$STAA^{p23}_{sig}$	1 exclusive jet process in each event 22 SoftQCD/HardQCD processes
$STAA^{p1}_{bg}$	1 process with non-diffractive jets with $p_T > 130$ GeV 1 proton on A side with $0.02 < \xi_{proton} < 0.2$ no min-bias particles in central detector 1 proton on C side with $0.02 < \xi_{proton} < 0.2$ no min-bias particles in central detector
$STAA^{p23}_{bg}$	1 process with non-diffractive jets with $p_T > 130$ GeV 1 SoftQCD/HardQCD processes with proton on A side with $0.02 < \xi_{proton} < 0.2$ 1 SoftQCD/HardQCD processes with proton on C side with $0.02 < \xi_{proton} < 0.2$ 20 SoftQCD/HardQCD processes without cuts on protons

Jet Reconstruction

In this section the comparison of the MC samples: $STAA^{p23}_{sig/bg}$, $ATLAS^{p23}_{sig}$ and $STAA^{p1}_{sig/bg}$ is presented. Figure B.7 shows the number of jets with p_T greater than 130 GeV for each sample and for different radii of the jet reconstruction.

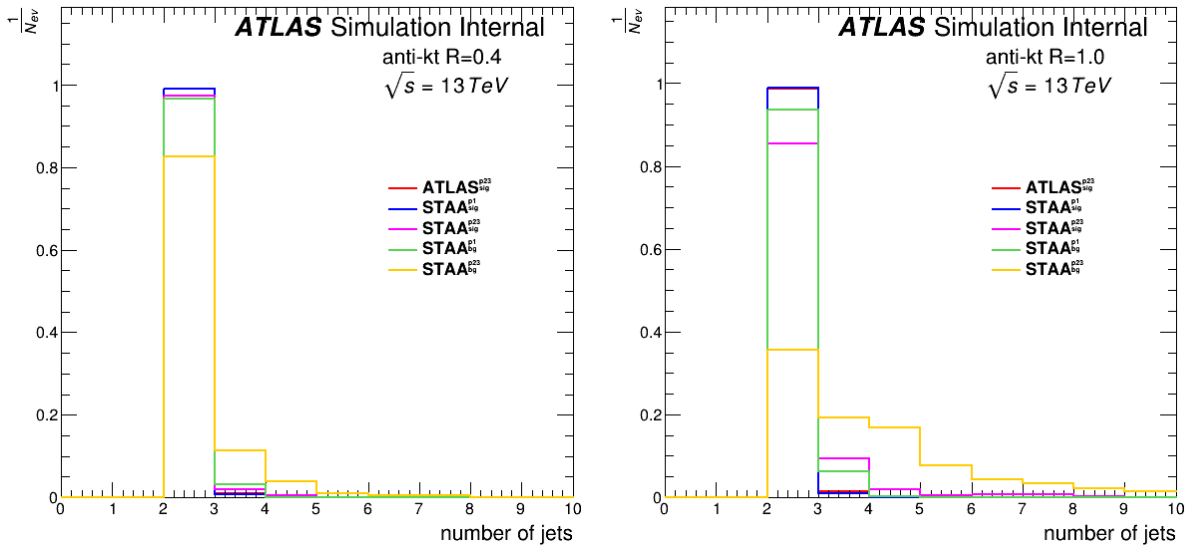


Figure B.7: Number of jets with $p_T > 130$ GeV in MC samples: $STAA^{p1}_{sig}$ (blue), $STAA^{p23}_{sig}$ (magenta), $STAA^{p23}_{bg}$ (yellow), $ATLAS^{p23}_{sig}$ (red) and $STAA^{p1}_{bg}$ (green). The **left** figure is for R=0.4 and the **right** is for R=1.0.

Samples $ATLAS^{p23}_{sig}$ and $STAA^{p1}_{sig}$ are very similar in both cases. It means that the ATLAS pile-up suppression algorithms are working efficiently. For sample $STAA^{p23}_{sig}$, there are some dissimilarities with respect to $STAA^{p1}_{sig}$ which come from the presence of the pile-up particles in the event – more jets are visible. The biggest dissimilarity is visible for the background sample ($STAA^{p23}_{bg}$), where especially for larger anti-kt radius the number of jets increases. The presence of the pile-up particles results in more reconstructed "mini-jets". One should observe that with more "mini-jets", higher p_T jets can be produced. This is not visible for $STAA^{p1}_{bg}$ as there are only non-diffractive jets and two protons.

In Figure B.8 the transverse momentum of the leading jet was plotted for all five samples. In both cases of different radii, the $STAA^{p1}_{sig}$ sample contains the leading jets of smaller transverse momentum values with respect to other samples. It is understandable because this sample contains only a signal event, no effects of the pile-up are included *i.e.* no additional particles in the jet reconstruction are present. Again, the $ATLAS^{p23}_{sig}$ MC sample is similar to $STAA^{p1}_{sig}$ and $STAA^{p1}_{bg}$ in contrast to $STAA^{p23}_{sig}$ and $STAA^{p23}_{bg}$. This is understandable since in the last two samples the pile-up events contribute to jets without any suppression.

As was mentioned in Sec. B.2.1, to accept an event by the trigger the energy loss has to be calculated using the di-jet system kinematics. Its value translates into the proton position in the AFP station. In Figure B.9 left (right) the ξ_{dijet} calculated using the di-jet system for sides A (C) is presented. For both sides and anti-kt radii, the energy loss for the background sample ($STAA^{p23/p1}_{bg}$, yellow/green line) populates the low ξ region. The protons with $0.1 \lesssim \xi$ are out of the range of the detector and will not be visible.

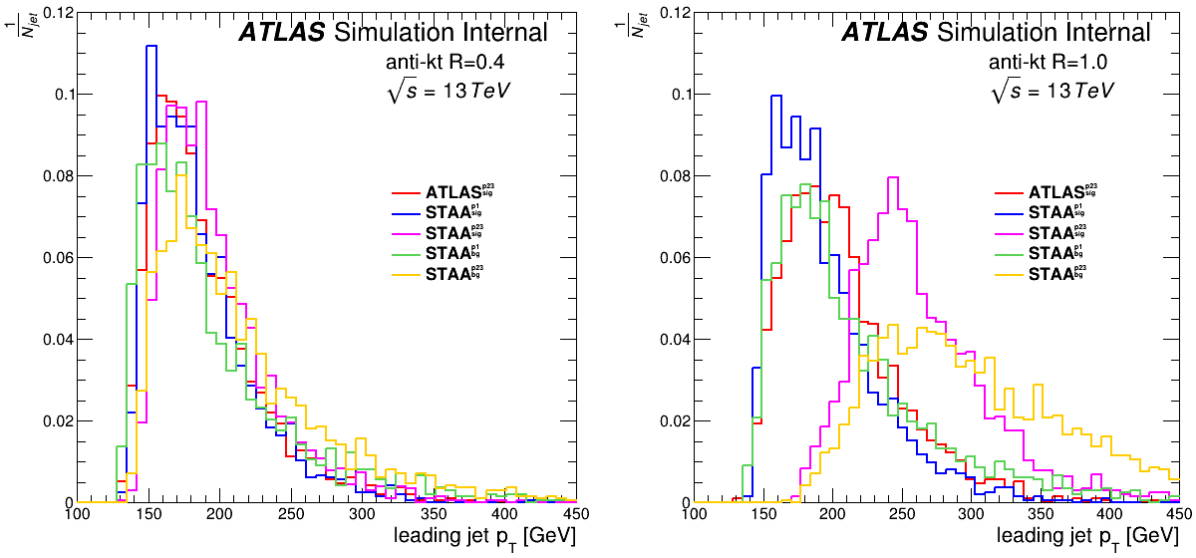


Figure B.8: Distribution of the leading jet transverse momentum in MC samples: $STAA^{p1}_{sig}$ (blue), $STAA^{p23}_{sig}$ (magenta), $STAA^{p1}_{bg}$ (green), $STAA^{p23}_{bg}$ (yellow) and $ATLAS^{p23}_{sig}$ (red). The **left** plot is for R=0.4 and the **right** one is for R=1.0.

Plots presented in Figure B.10 show a comparison between the energy of the forward proton and ξ calculated using the di-jet system for each discussed MC sample. A good correlation is visible in all signal samples. For the $STAA^{p1}_{sig}$ sample (top left) the events have always $\xi_{dijet} \leq \xi_{proton}$. It is because there are no additional pile-up-originating particles in the jet reconstruction. Therefore, the jets can only be reconstructed as smaller than in reality. The $ATLAS^{p23}_{sig}$ sample (top right) contains some events where $\xi_{dijet} > \xi_{proton}$, but in most cases, the jet reconstruction algorithm provides events with smaller ξ_{dijet} value. The correlation in the $STAA^{p23}_{sig}$ sample (bottom left) is worse than that in the $ATLAS^{p23}_{sig}$ one as there is no pile-up suppression. Finally, for the background sample (bottom right) there is no correlation, which is expected due to the fact that the jets and protons come from different processes.

It was mentioned that events in signal samples are the same. It is well visible in the top left plot of Figure B.11, where ξ_{proton} is exactly the same for $STAA^{p1}_{sig}$ and $ATLAS^{23}_{sig}$ samples. For correlation with $STAA^{p23}_{sig}$ (top right), there are few cases where the pile-up proton was closer than the one from the signal event (top right). ATLAS MC sample shows also a good correlation with the $STAA^{p1}_{sig}$ sample reconstructed using the anti-kt R=0.4 (centre left). In the case of R=1.0, the correlation (bottom right) becomes weaker. The pile-up presence in the MC sample weakly influences the correlation when the anti-kt R=0.4 algorithm is used (centre right). For R=1.0 the jet system in the standalone sample provides much higher ξ than the corresponding ones from the ATLAS simulation (bottom right).

The trigger accepts an event when the prediction is close enough to the detected proton,

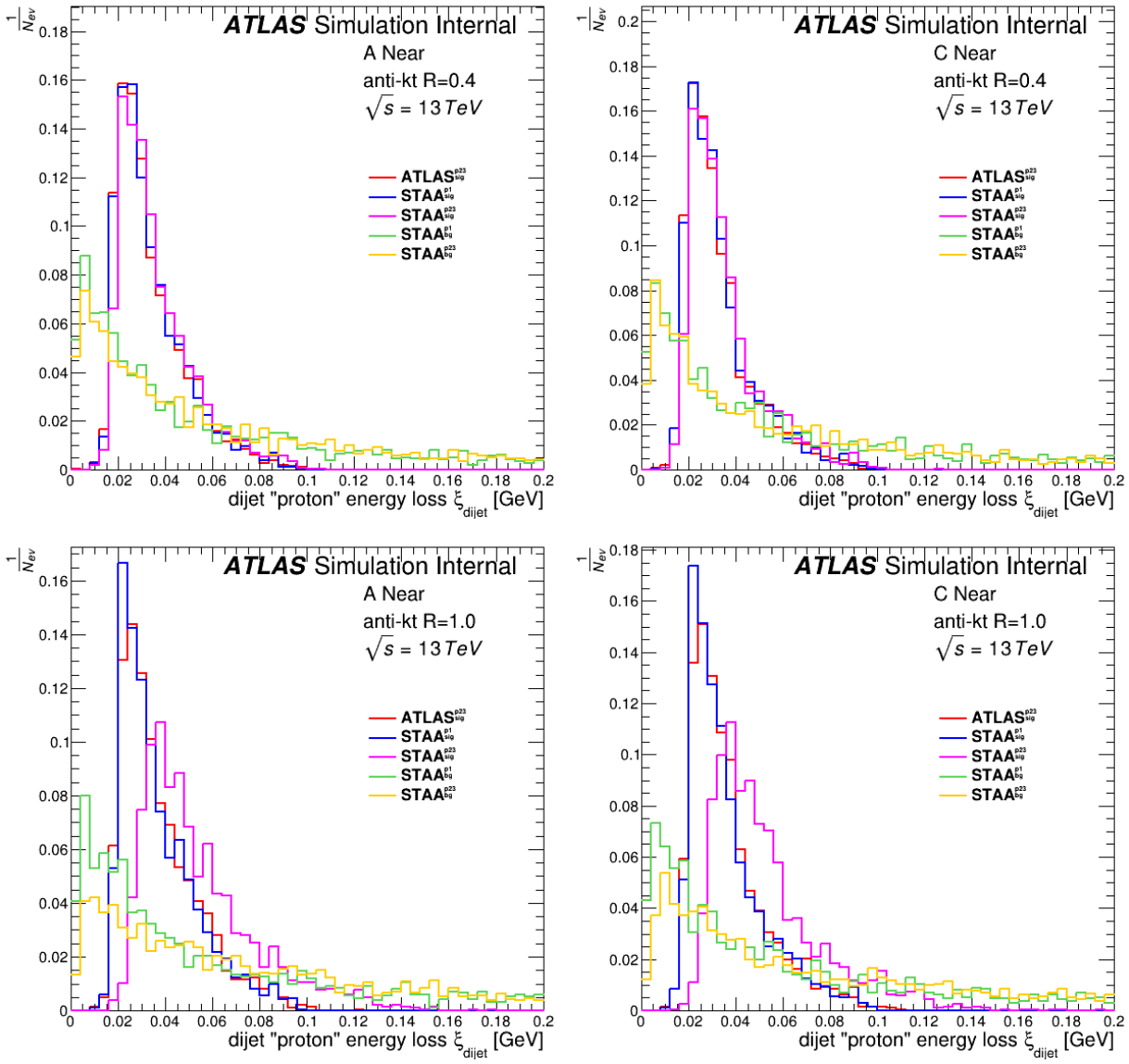


Figure B.9: The energy loss of a proton as calculated from the di-jet system kinematics for the MC samples: $STAA^{p1}_{sig}$ (blue line), $STAA^{p23}_{sig}$ (magenta), $STAA^{p1}_{bg}$ (green), $STAA^{p23}_{bg}$ (yellow) and $ATLAS^{p23}_{sig}$ (red). The jets are reconstructed using the anti-kt algorithm with R=0.4 (**top**) and R=1.0 (**bottom**). The **left** plots are for protons on side A, the **right** ones for protons on side C.

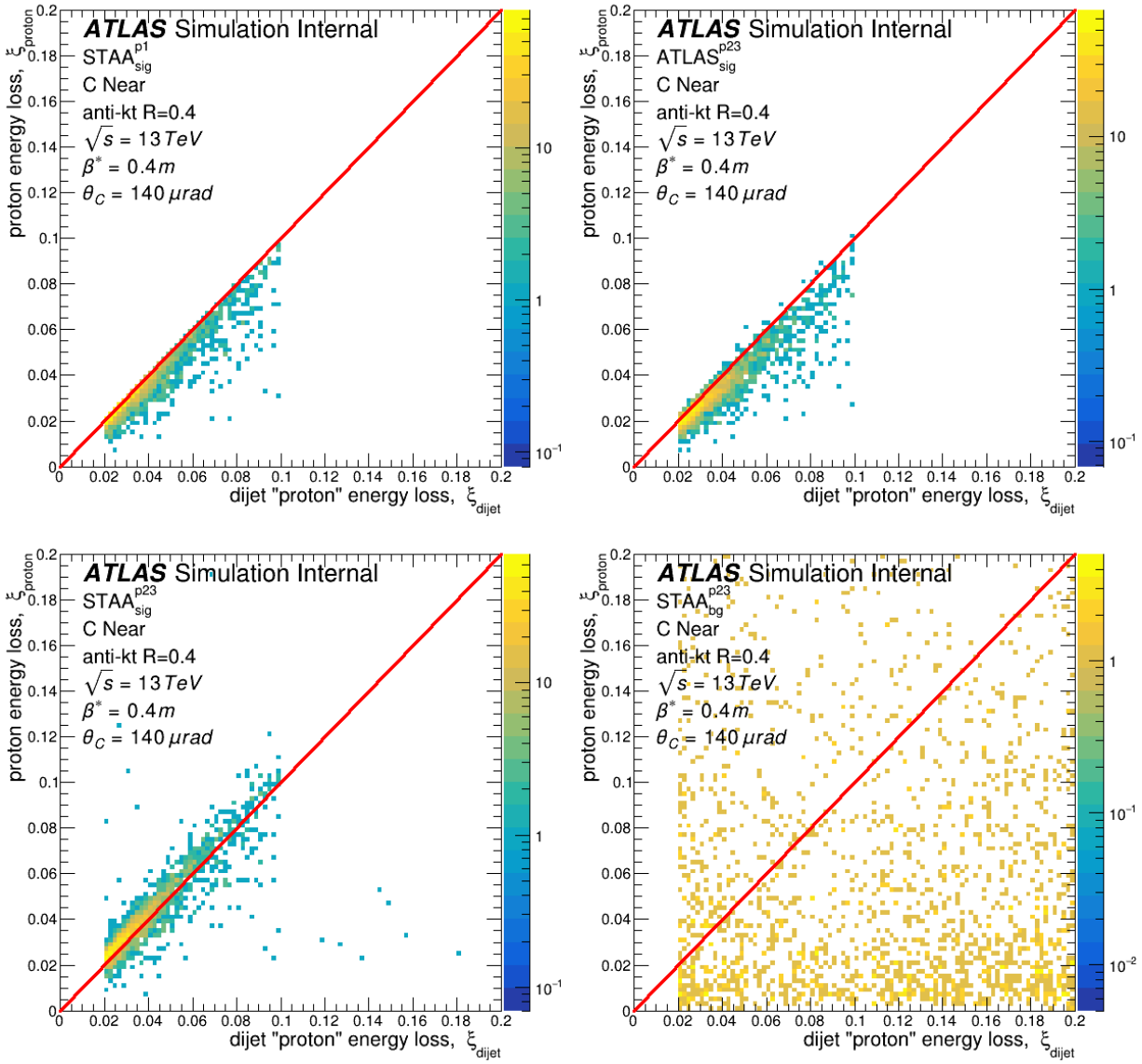


Figure B.10: Correlations between ξ calculated using the di-jet system and that of the proton registered in the AFP station for different MC samples: $STAA^{p1}_{sig}$ (**top left**), $ATLAS^{p23}_{sig}$ (**top right**), $STAA^{p23}_{sig}$ (**bottom left**) and $STAA^{p23}_{bg}$ (**bottom right**).

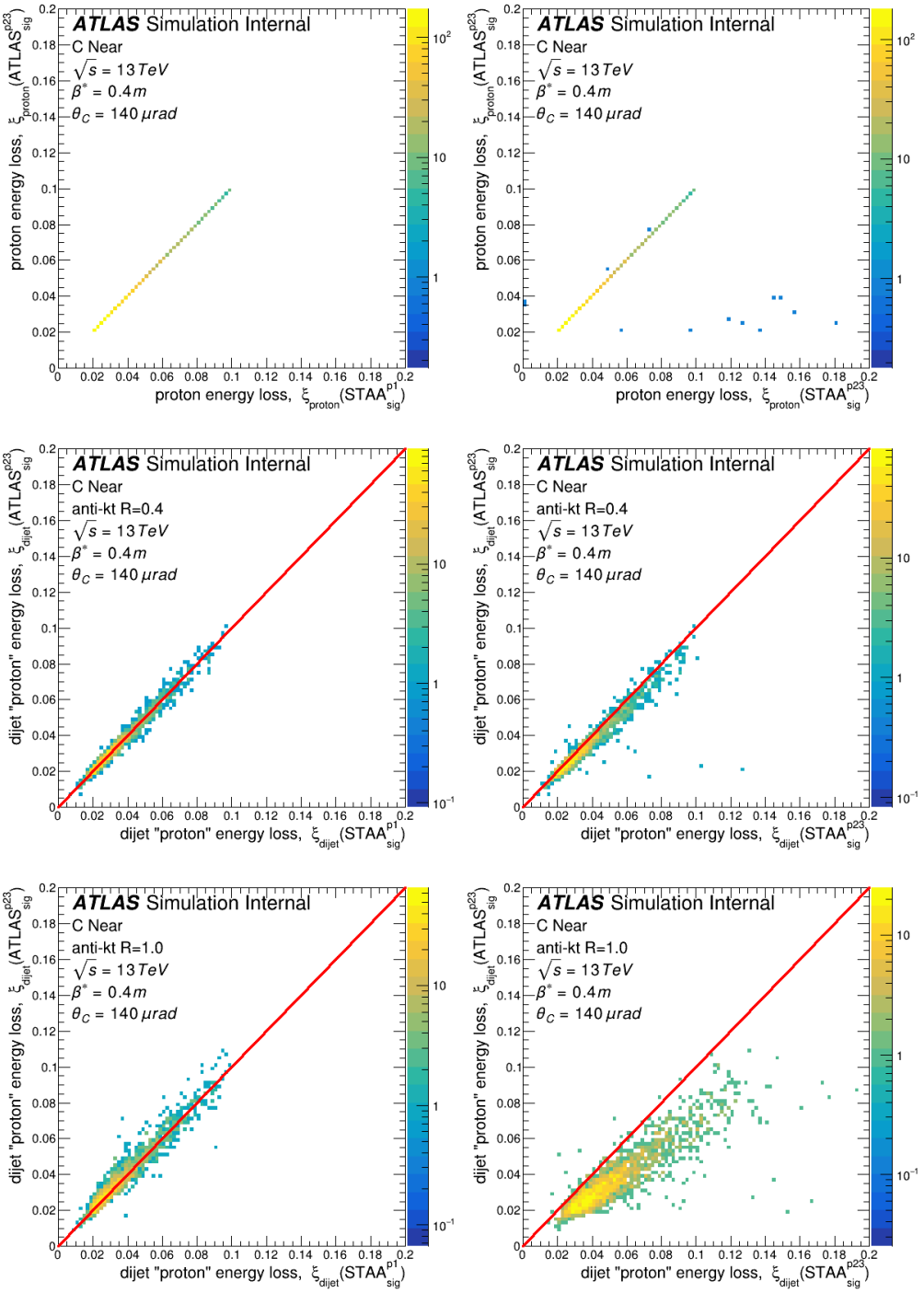


Figure B.11: Correlations between the relative energy loss of a proton in different MC samples. **Left** plots are for $STAA^{p1}_{sig}$ and $ATLAS^{p23}_{sig}$, **right** for $STAA^{p23}_{sig}$ and $ATLAS^{p23}_{sig}$. **Top** plots: ξ calculated from the forward protons. **Centre** plots: ξ calculated from the di-jet system with jets reconstructed with radius R=0.4. **Bottom** plots: ξ calculated from the di-jet system with jets reconstructed with radius R=1.0.

i.e. it has to fulfil the conditions described previously in Sec. B.3. Figure B.12 shows the distance R between the di-jet prediction and the registered forward proton. Position predicted from di-jet was always reconstructed using θ_C=140 μrad for the jet radius of 0.4, whereas the position of the reconstructed forward proton was calculated with θ_C=140 μrad (top), 160 μrad (centre) and 180 μrad (bottom). This represents the situation when the LHC optics settings would change during the run (the luminosity levelling) while the trigger settings remain fixed at θ_C=140 μrad.

As one can conclude from these plots, with the default settings most of the signal events would be accepted by the trigger. The exception is the $STAA^{p23}_{sig}$ sample when the jets are reconstructed with R=1.0 and the effect of the pile-up is more visible. The maximum is at about 2 mm. As it was expected, in the background events no maxima are visible as there is no correlation between the jets and protons.

The trigger contains also the conditions concerning the distances in the x and y coordinates. Those distributions are plotted in Figures B.13 and B.14 with the dashed red lines representing the default trigger cuts. Similarly to the R-distance distributions, the background events tend to be distributed in the whole spectrum of the x and y positions, whereas the signal events are mostly confined to the trigger acceptance region. Moreover, it seems that for the x/y position it is worth considering the asymmetrical cuts keeping in mind that the efficiency may worsen for different θ_C values. More discussion will be held in the next chapter.

Finally, it was checked whether the difference in ξ_{dijet} between the ATLAS and standalone samples depends on the forward proton energy loss (ergo the proton position in the AFP). No such correlation is visible (c.f. Figure B.15) neither when jets were reconstructed with radius R=0.4 nor R =1.0.

B.3.2 Trigger Efficiency

The exclusive jet trigger algorithm has three conditions one can choose from to accept an event. These concern the distance between the proton positions in x, y directions and radius R. Up till now, the performance of the trigger for the default values of these cuts was presented and discussed. Here, a discussion of the trigger efficiency improvement following the change of those conditions will be given.

Existing Trigger Conditions

The trigger efficiencies for different R/x/y values are presented in Figure B.16 (jets reconstructed with anti-kt R=0.4) and B.17 (jets reconstructed with anti-kt R=1.0). Only one selection criterion (either x or y or R) was used at the time, whereas the others were omitted. Different lines represent $ATLAS^{p23}_{sig}$, $STAA^{p1}_{sig}$ and $STAA^{p1}_{bg}$ MC samples. The left plots are for a situation when the forward proton was reconstructed with θ_C= 140 μrad, the right

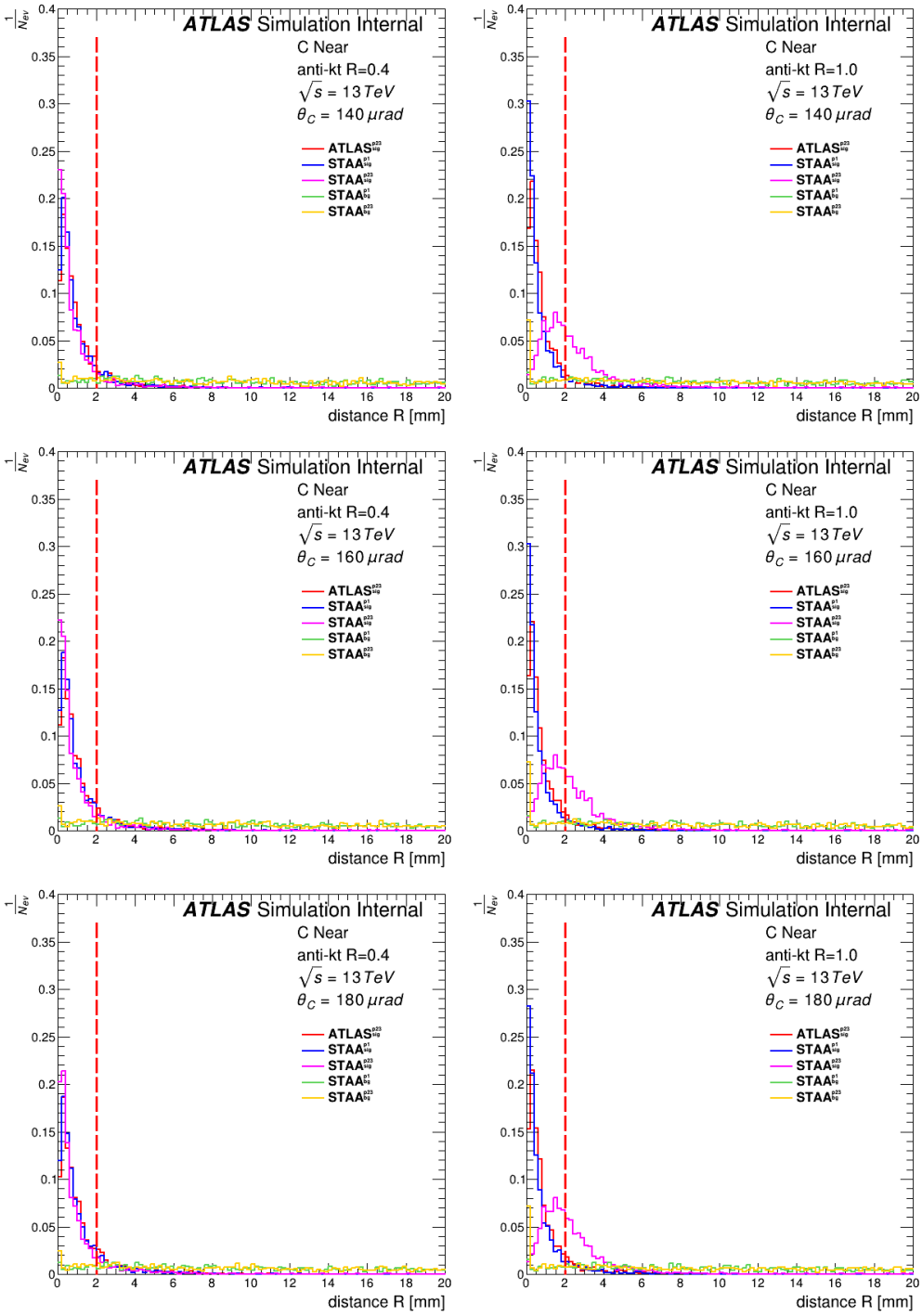


Figure B.12: On plots there is the distance R between prediction from the di-jet system and produced forward proton with different settings of crossing angle: θ_C=140 μrad (**top**), θ_C=160 μrad (**centre**), θ_C=180 μrad (**bottom**). Jets are reconstructed with anti-kt with R=0.4 (**left**) and R=1.0 (**right**). Solid lines represent different MC samples: $STAA^{p1}_{sig}$ (blue), $STAA^{p23}_{sig}$ (magenta), $STAA^{p23}_{bg}$ (yellow), $ATLAS^{p23}_{sig}$ (red) and $STAA^{p1}_{bg}$ (green). The dashed line represents the value of the default cut.

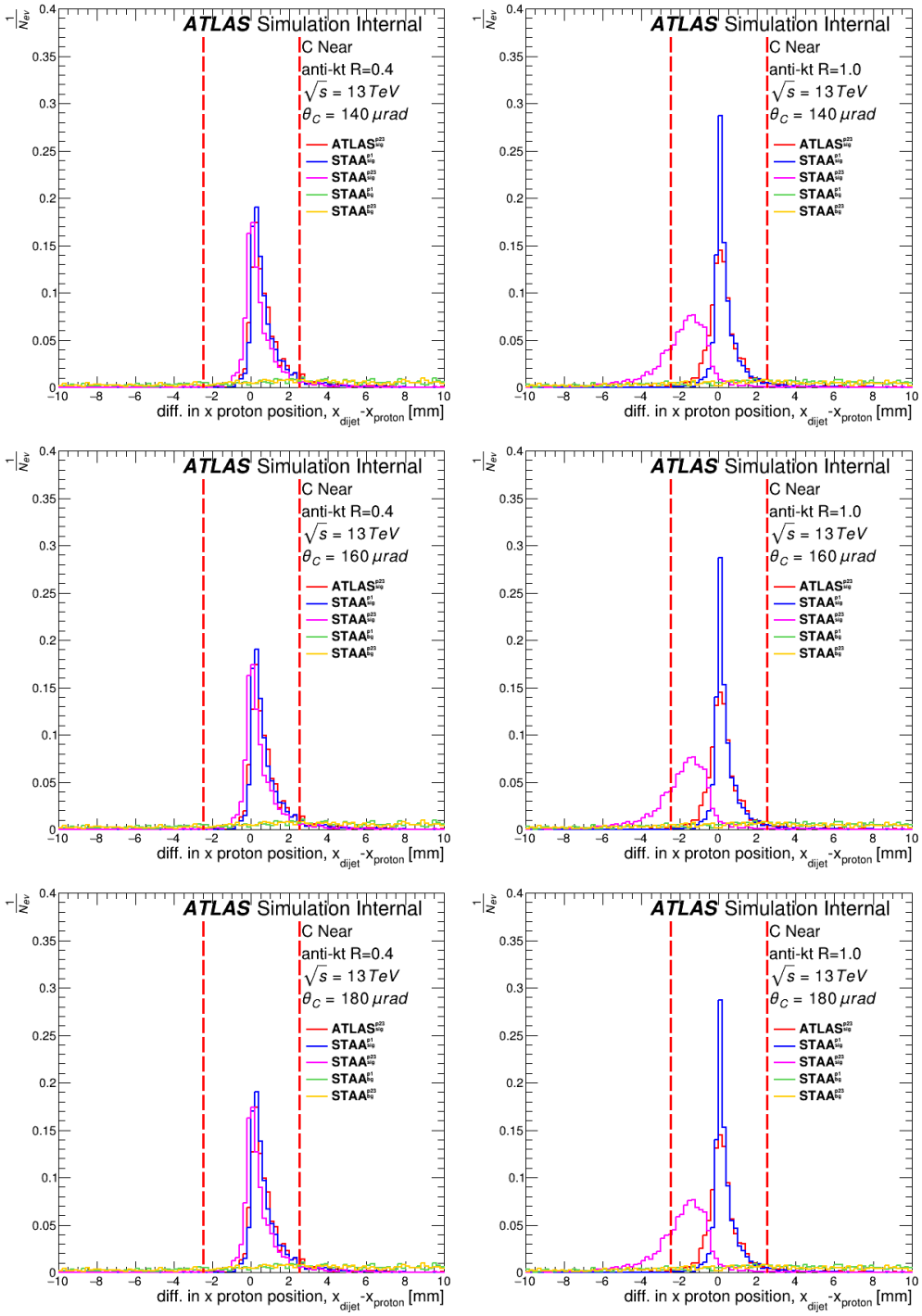


Figure B.13: On plots there is the distance R between prediction from the di-jet system and produced forward proton with different settings of crossing angle: θ_C=140 μrad (**top**), θ_C=160 μrad (**centre**), θ_C=180 μrad (**bottom**). Jets are reconstructed with anti-kt with R=0.4 (**left**) and R=1.0 (**right**). Lines represent different MC samples: $STAA^{p1}_{sig}$ (blue)the , $STAA^{p23}_{sig}$ (magenta), $STAA^{p23}_{bg}$ (yellow) and $ATLAS^{p23}_{sig}$ (red) and $STAA^{p1}_{bg}$ (green). The dashed line represents the value of the default cut.

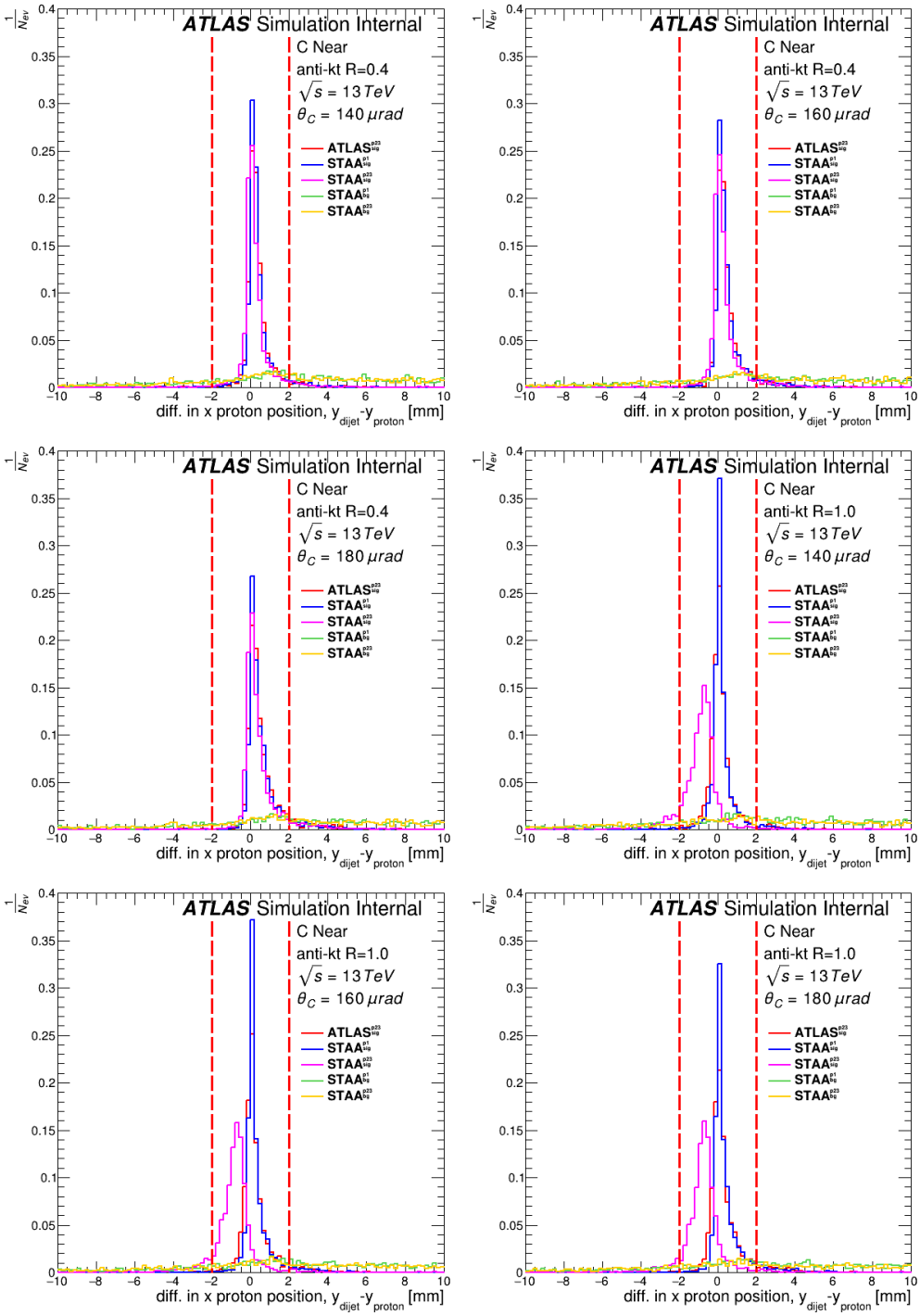


Figure B.14: On plots there is the distance R between prediction from the di-jet system and produced forward proton with different settings of crossing angle: θ_C=140 μrad (**top**), θ_C=160 μrad (**centre**), θ_C=180 μrad (**bottom**). Jets are reconstructed with anti-kt with R=0.4 (**left**) and R=1.0 (**right**). Lines represent different MC samples: $STAA^{p1}_{sig}$ (blue), $STAA^{p23}_{sig}$ (magenta), $STAA^{p23}_{bg}$ (yellow) and $ATLAS^{p23}_{sig}$ (red) and $STAA^{p1}_{bg}$ (green). The dashed line represents the value of the default cut.

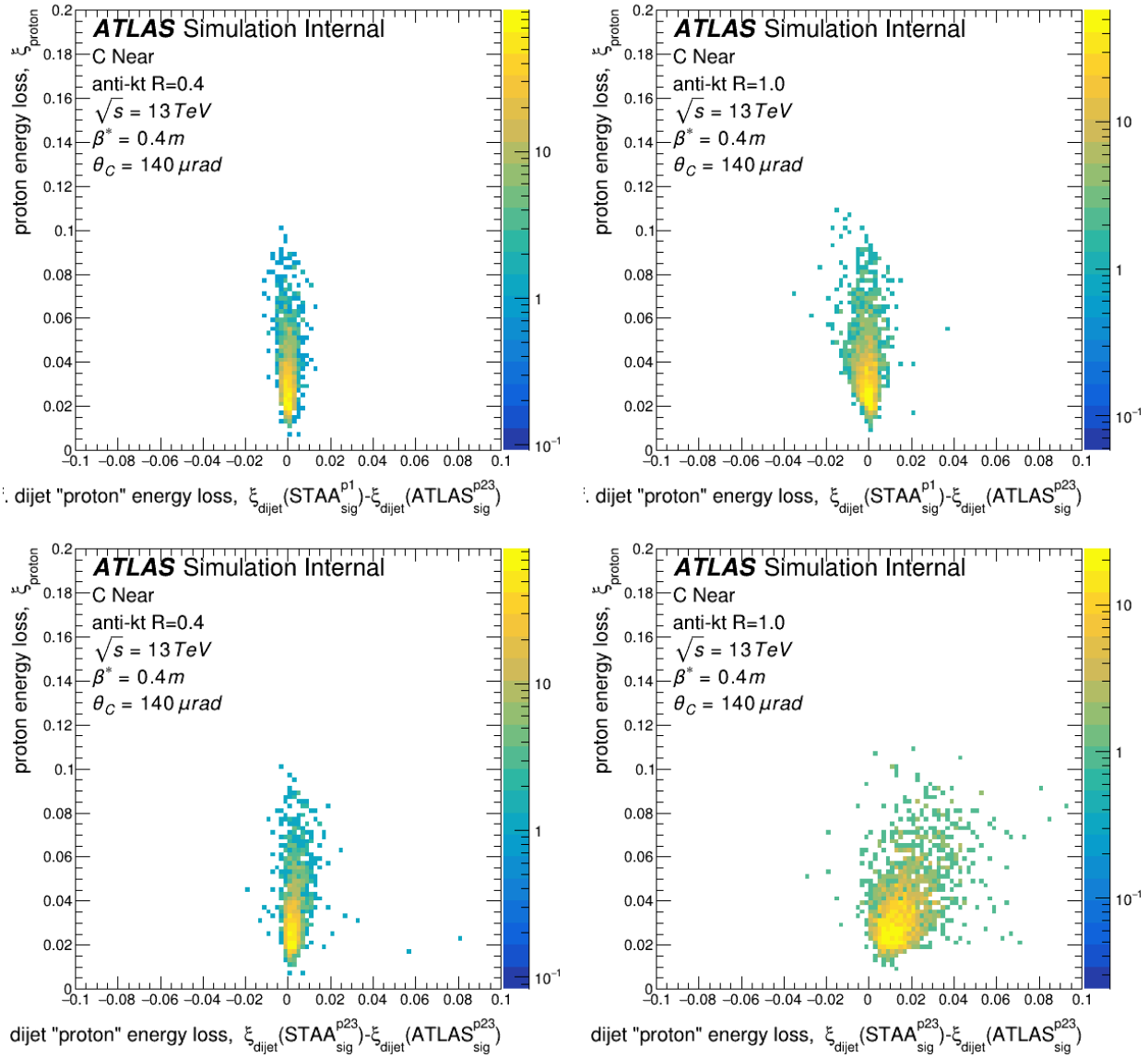


Figure B.15: The ξ_{proton} for the signal event versus the difference between ξ_{dijet} in $ATLAS^{p23}_{sig}$ and $STAA^{p1}_{sig}$ (**top**) and $STAA^{p23}_{sig}$ (**bottom**). The jets were reconstructed with R=0.4 (**left**) and R=1.0 (**right**).

ones for θ_C= 180 μrad (the prediction from the di-jet system was computed assuming θ_C= 140 μrad in both cases). It is visible that the ATLAS MC sample behaves like $STAA^{p1}_{sig}$, giving good results for default trigger cuts whereas efficiency on the background sample stays at a very low level. The values listed in Table B.1 are visible on the plots, presented as points at the default values (dotted red line) for $STAA^{p1}_{sig}$ sample (solid blue line). For a better comparison, in Table B.4 the values of x, y and R cut are listed when the trigger efficiency reached a certain level for different MC samples.

Table B.4: The trigger efficiency for asymmetrical x/y and R cuts for different MC samples with forward proton reconstructed with various θ_C and prediction from di-jet computed using θ_C=140 μrad.

		Trigger Efficiency					
		θ_C = 140 μrad			θ_C = 180 μrad		
Sample	Cut [mm]	90%	95%	99%	90%	95%	99%
anti-kt R=0.4							
$ATLAS^{p23}_{sig}$	R	3.25	4.46	7.35	3.65	5.25	7.75
	x	2.00	2.80	5.20	2.00	2.80	5.20
	y	1.25	1.85	3.30	1.70	2.50	4.40
$STAA^{p1}_{sig}$	R	2.15	3.05	5.95	2.40	3.50	6.50
	x	1.20	1.85	4.20	1.20	1.85	4.20
	y	0.65	1.15	2.65	0.90	1.55	3.55
anti-kt R=1.0							
$ATLAS^{p23}_{sig}$	R	2.60	3.65	6.50	2.95	4.20	7.15
	x	1.55	2.25	4.55	1.55	2.25	4.55
	y	1.05	1.60	2.85	1.35	2.20	3.85
$STAA^{p1}_{sig}$	R	1.50	2.30	4.80	1.70	2.70	5.40
	x	1.75	1.25	3.15	1.75	1.25	3.15
	y	0.50	0.90	2.30	0.65	1.25	3.20

Before claiming whether the default values for the trigger conditions are sufficient it is necessary to study the rates of background events. Such rates are presented in Figure B.18. As before, the x-axis presents different values of the trigger conditions (R, x and y). For the default value of R=2 mm, the rate for a sample with the pile-up of μ=23 is about 100 Hz and for μ=46 it is about 200 Hz. There are no differences between the radii used for the jet reconstruction for the background events but the efficiency of the trigger is higher for a larger radius which seems to be a slightly better choice for the trigger algorithm. This assumes a perfect pile-up suppression algorithm. If the trigger condition would be bigger by 0.5 mm the efficiency would go up by a few per cent but the background rate would

increase by 50%.

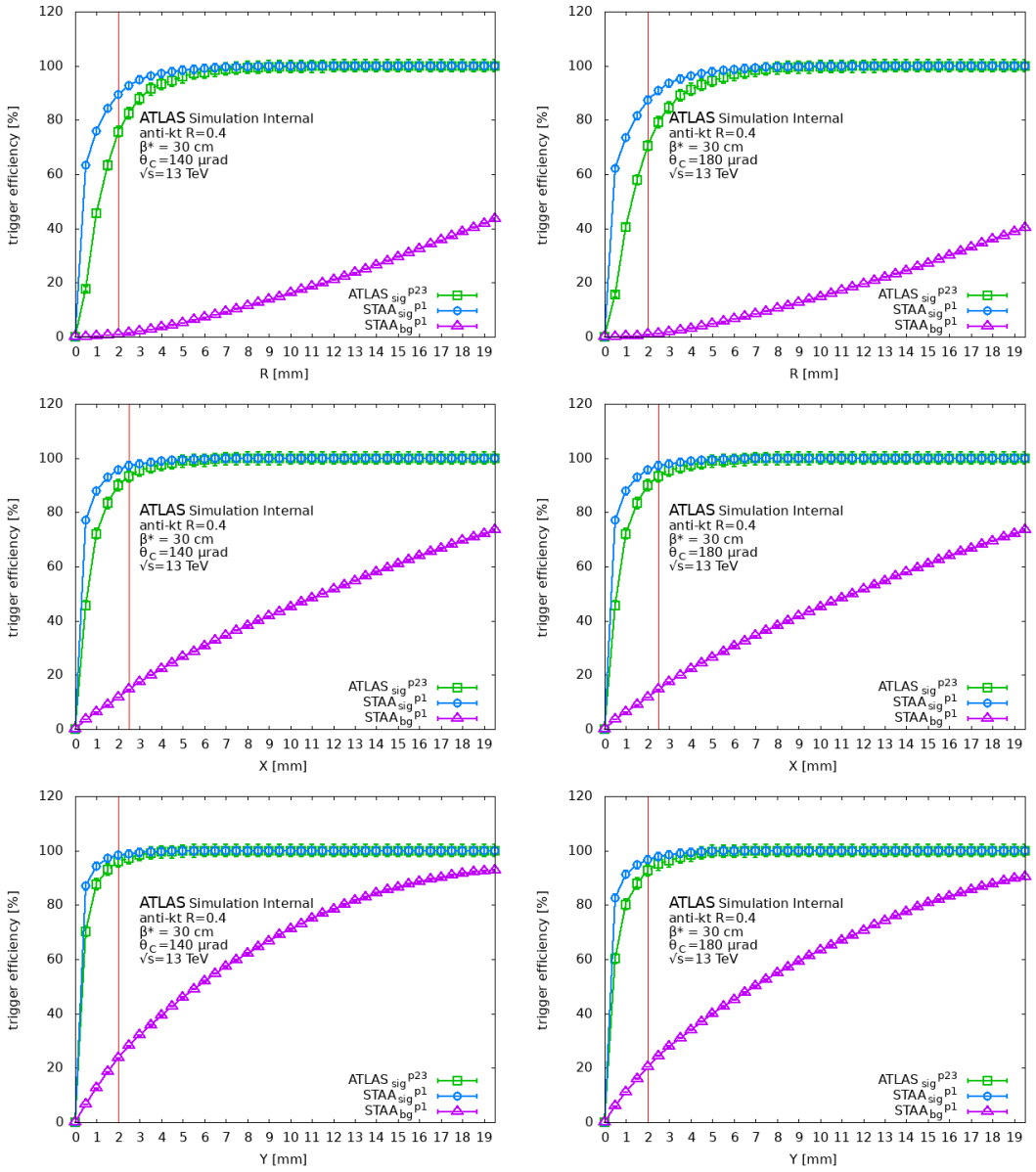


Figure B.16: The trigger efficiency for the different trigger conditions: R (**top**), x (**centre**) and y (**bottom**) and different MC samples: $ATLAS_{sig}^{p23}$ (green), $STAA_{sig}^{p1}$ (blue) and $STAA_{bg}^{p1}$ (purple). All plots are for the jets reconstructed with a radius of 0.4. For the **left** plots both the position predicted from di-jet and the forward proton position were reconstructed assuming θ_C= 140 μrad. For the **right** ones the position predicted from the di-jet system was reconstructed with θ_C= 140 μrad but the forward proton position was reconstructed with θ_C= 180 μrad. The dotted red line represents the initial trigger setup.

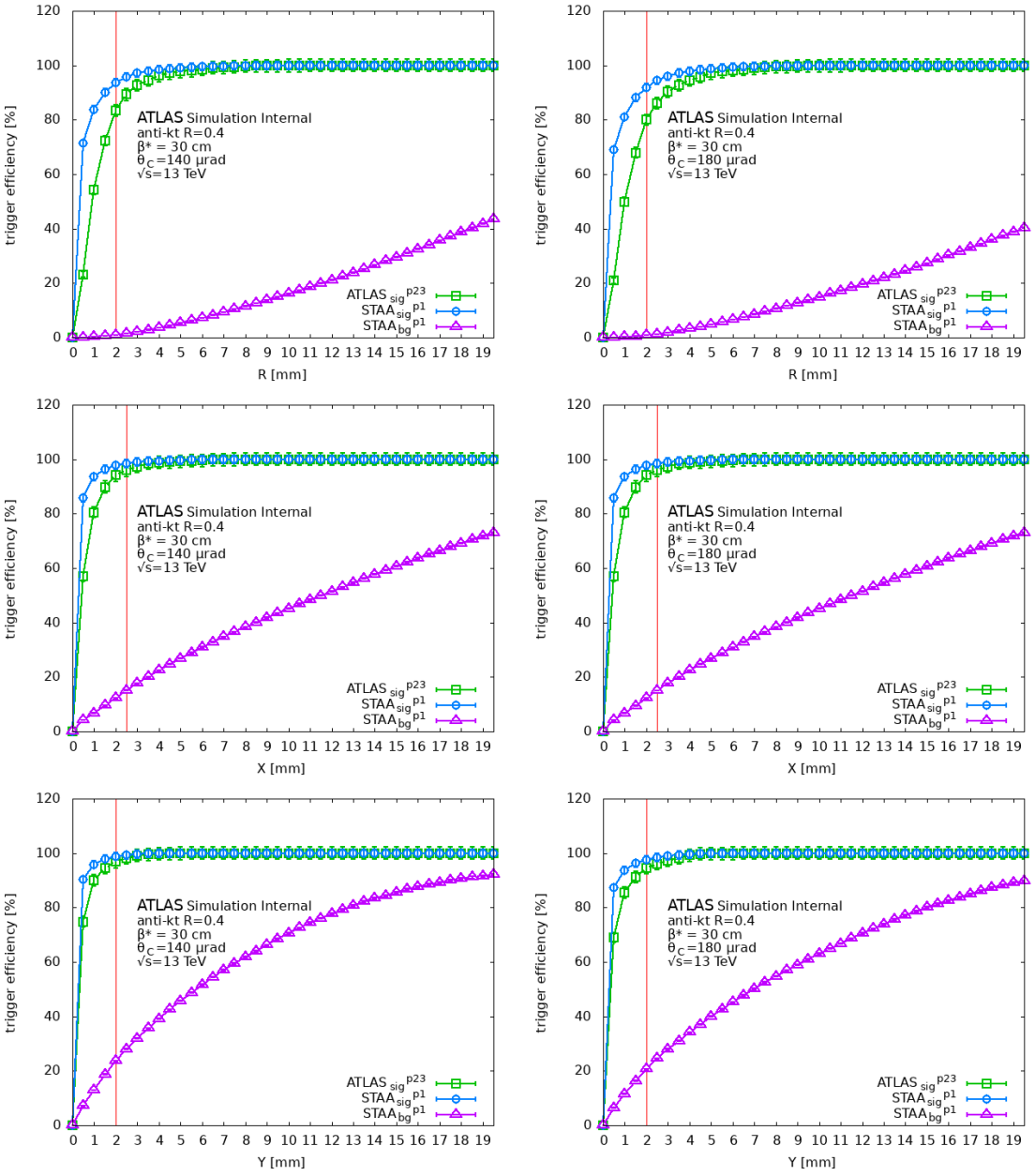


Figure B.17: The trigger efficiency for the different trigger conditions: R (**top**), x (**centre**) and y (**bottom**) and different MC samples: $ATLAS_{sig}^{p23}$ (green), $STAA_{sig}^{p1}$ (blue) and $STAA_{bg}^{p1}$ (purple). All plots are for the jets reconstructed with a radius of 1.0. For the **left** plots both positions predicted from the di-jets and from the forward proton were reconstructed assuming θ_C= 140 μrad. For the **right** ones the position predicted from the di-jets was reconstructed with θ_C= 140 μrad but the forward proton position was reconstructed with θ_C= 180 μrad. The dotted red line represents the initial trigger setup.

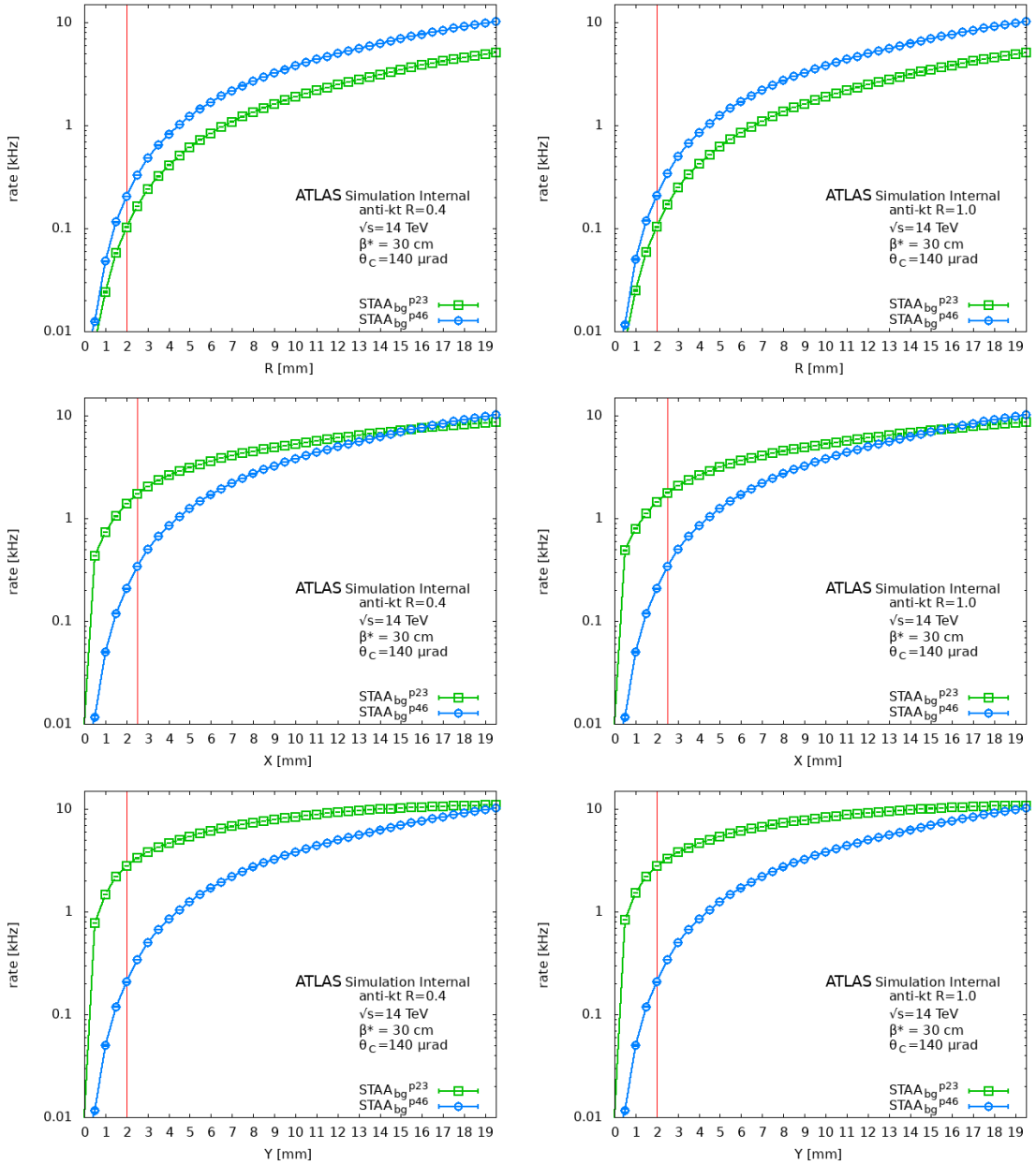


Figure B.18: The rates of the background events accepted by a trigger with pile-up μ=23 (green) and 46 (blue) with the assumption that the pile-up suppression algorithm works perfectly. The **left** plots are for the jets reconstructed with the anti-kt using the radius R=0.4, and the **right** ones are for R=1.0. Different trigger conditions were considered: R (**top**), x (**centre**) and y (**bottom**).

Elliptical and Asymmetrical Trigger Conditions

Figures B.13 and B.14 show the difference in the x and y positions. Those differences are not centred at zero. This suggests the use of different cuts for the trigger. In this section, elliptical and asymmetrical x/y cuts are studied. The elliptical trigger condition accepted an event if the registered proton is within the ellipse titled by 45° with respect to the longer edge of the AFP detector. The ellipse has certain values of a and b radii and it is centred at the position of a proton reconstructed from the di-jet system. The asymmetrical cuts were considered separately for the x and y directions. Here forward proton had to be within the range (x_{min}, x_{max}) or (y_{min}, y_{max}). The idea of new cuts and the default ones is presented in Figure B.19.

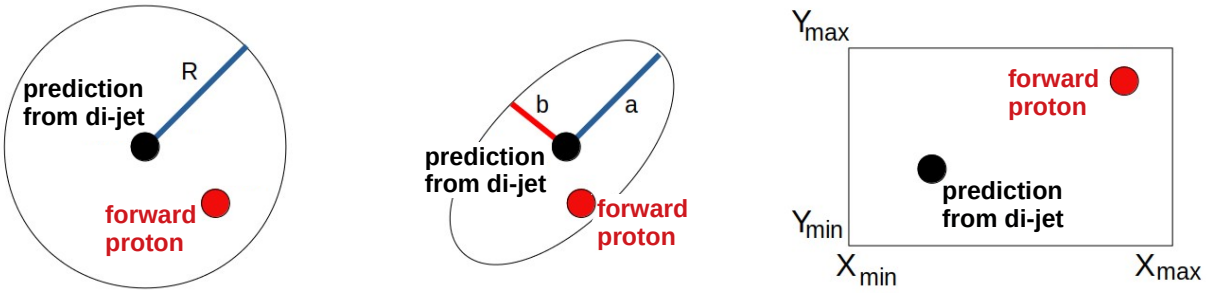


Figure B.19: A scheme of the trigger conditions: default circular (**left**), elliptical (**centre**) and asymmetrical x/y (**right**).

The trigger efficiencies for the $ATLAS_{sig}^{p23}$ and $STAA_{bg}^{p23}$ samples for various values of a and b in the elliptical cut and x_{min} (y_{min}) and x_{max} (y_{max}) were calculated. The results for different radii for the jet reconstruction and crossing angles used in the proton transport are presented in Figures B.20–B.25. All the values were changed in steps of 0.2 mm. The number in each bin represents the percentage of the accepted events (left plots) or the background rate (right plots) for a given cut. The efficiencies and rates for the asymmetrical cuts are much higher than for the elliptical ones because the ellipse limits the event acceptance both in the x and y directions, whereas for the x/y asymmetrical cut, the other direction was unlimited.

In cases of the y and elliptical cuts (see Figure B.20 and B.21), there is a drop in the trigger efficiency and background rate when the crossing angle used in reconstructing the forward proton differs from the one used in the di-jet prediction in the trigger. There is no such difference in the x cut because the change in θ_C does not influence the reconstructed x value.

The changes of the ellipse b radius, x_{min} and y_{min} result in smaller differences in the efficiency and rate than the radius, x_{max} and y_{max}. This effect was already visible on plots B.13 and B.14 where the asymmetry could be observed. Additionally, in all cases, the

change of the radius in the jet reconstruction increases the efficiency and rate, however, for the elliptical cut the increase of the rate seems to be smaller than that for the x/y cuts.

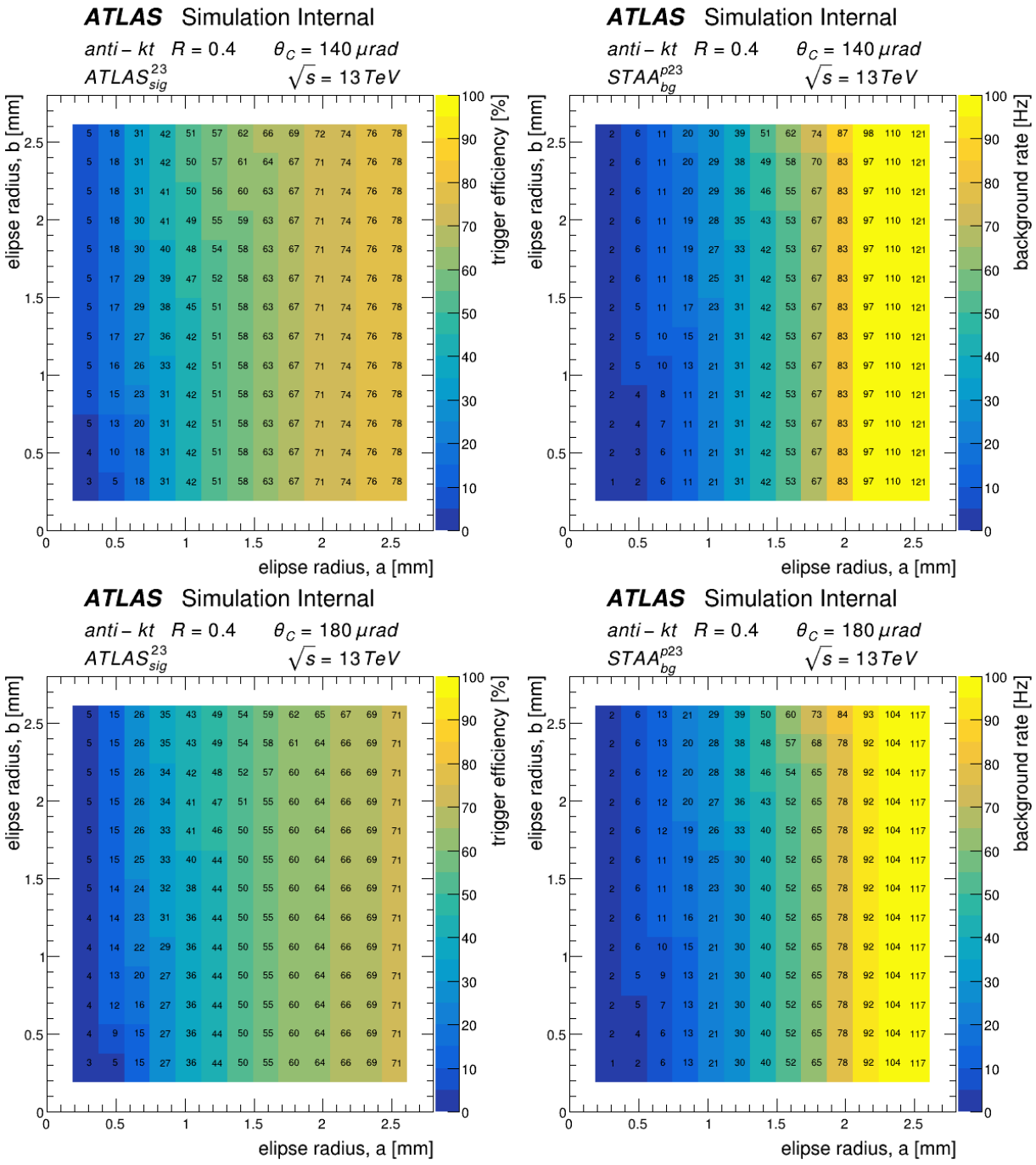


Figure B.20: The trigger efficiency (**left**) and background rates (**right**) for different values of the ellipse radii. The jets were reconstructed using the anti-kt with R=0.4 whereas forward proton position was calculated using θ_C = 140 μrad (**top**) or 180 μrad (**bottom**).

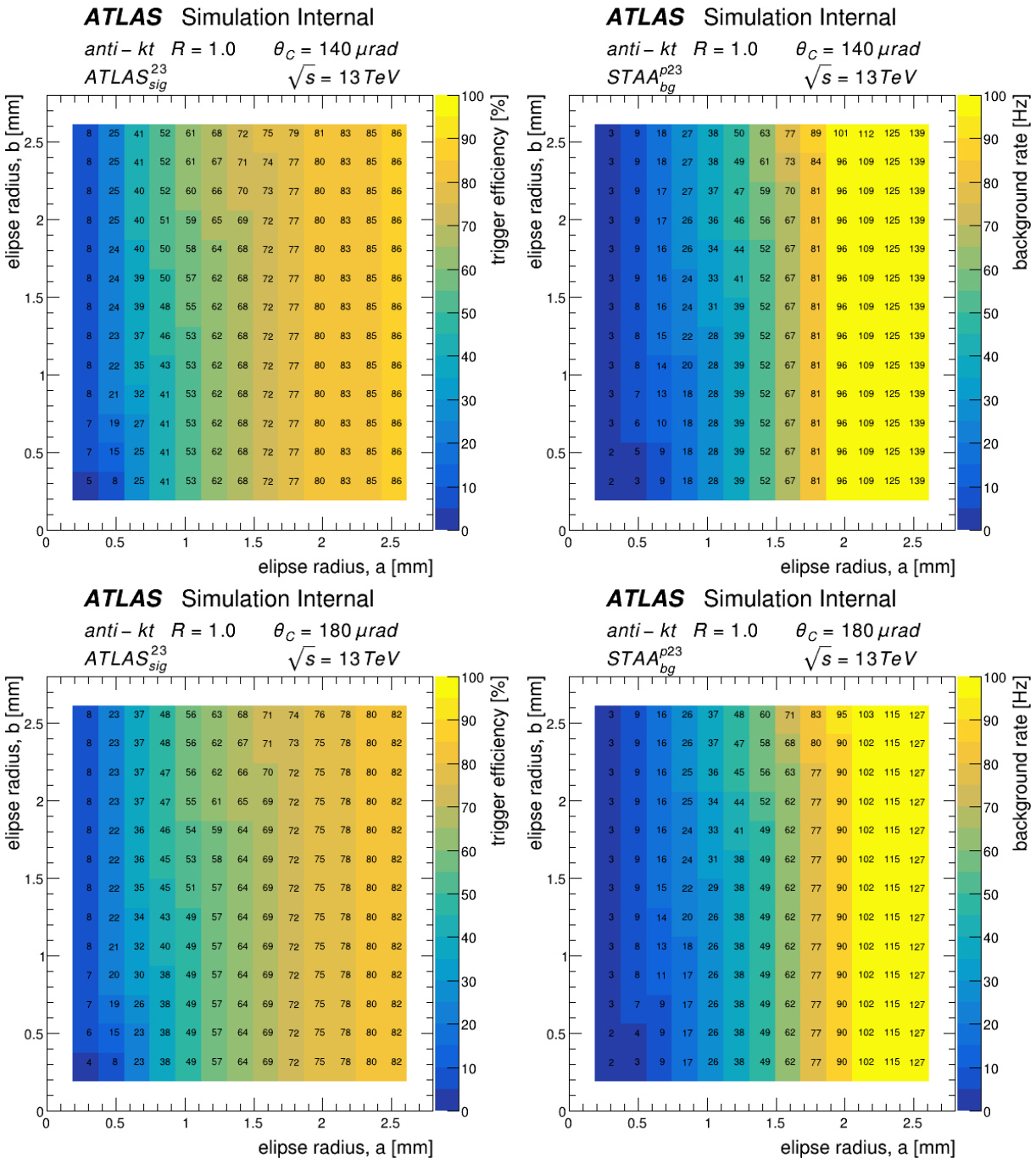


Figure B.21: The trigger efficiency (**left**) and background rates (**right**) for different ellipse radii. For jets reconstruction the anti-kt algorithm with R=1.0 was used whereas the forward proton position was calculated using $\theta_C = 140$ μrad (**top**) or 180 μrad (**bottom**).

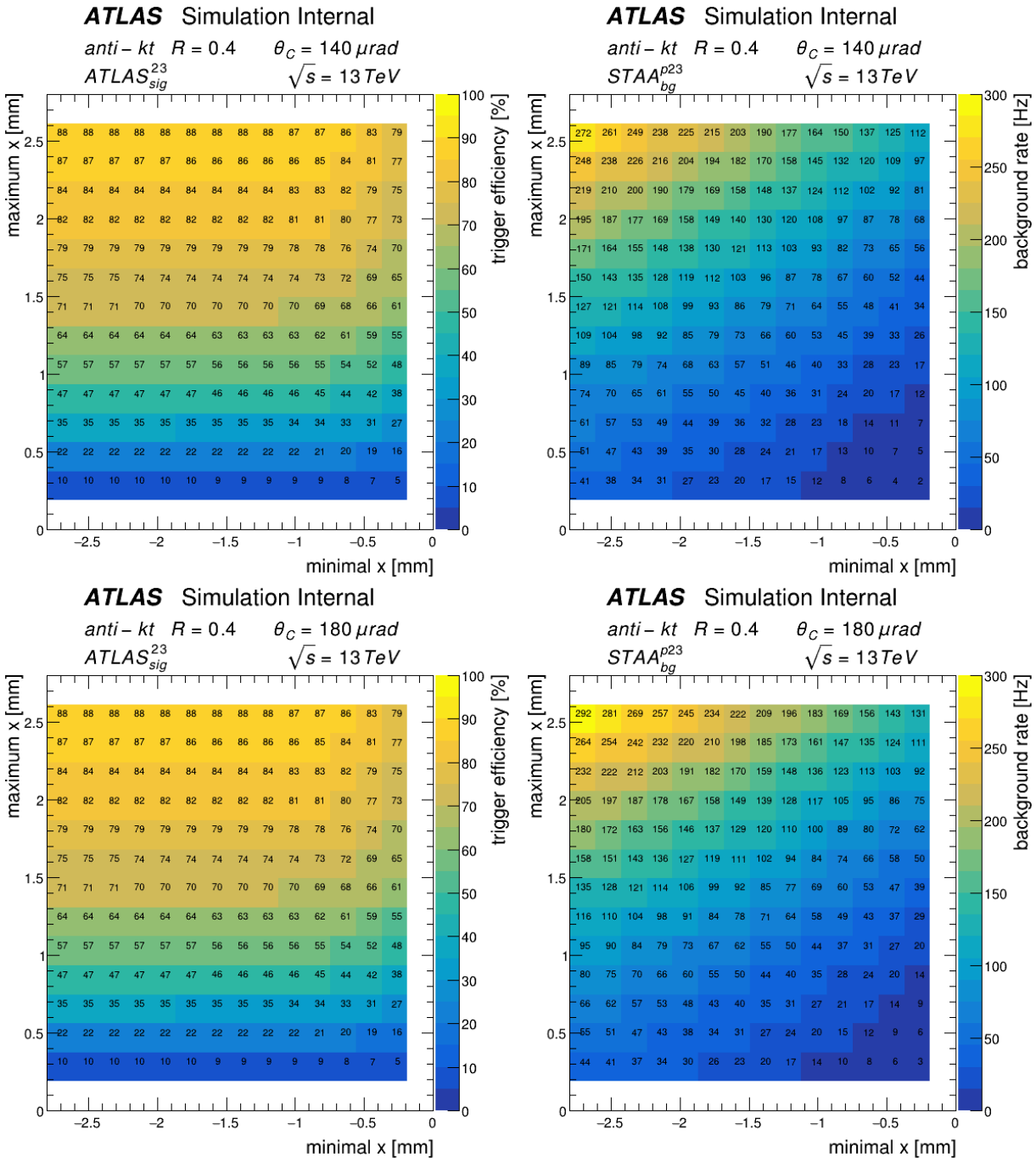


Figure B.22: The trigger efficiency (**left**) and background rates (**right**) for different values of the asymmetrical x cut. The jets were reconstructed using the anti-kt with R=0.4 whereas the forward proton position was calculated using $\theta_C = 140$ μrad (**top**) or 180 μrad (**bottom**).

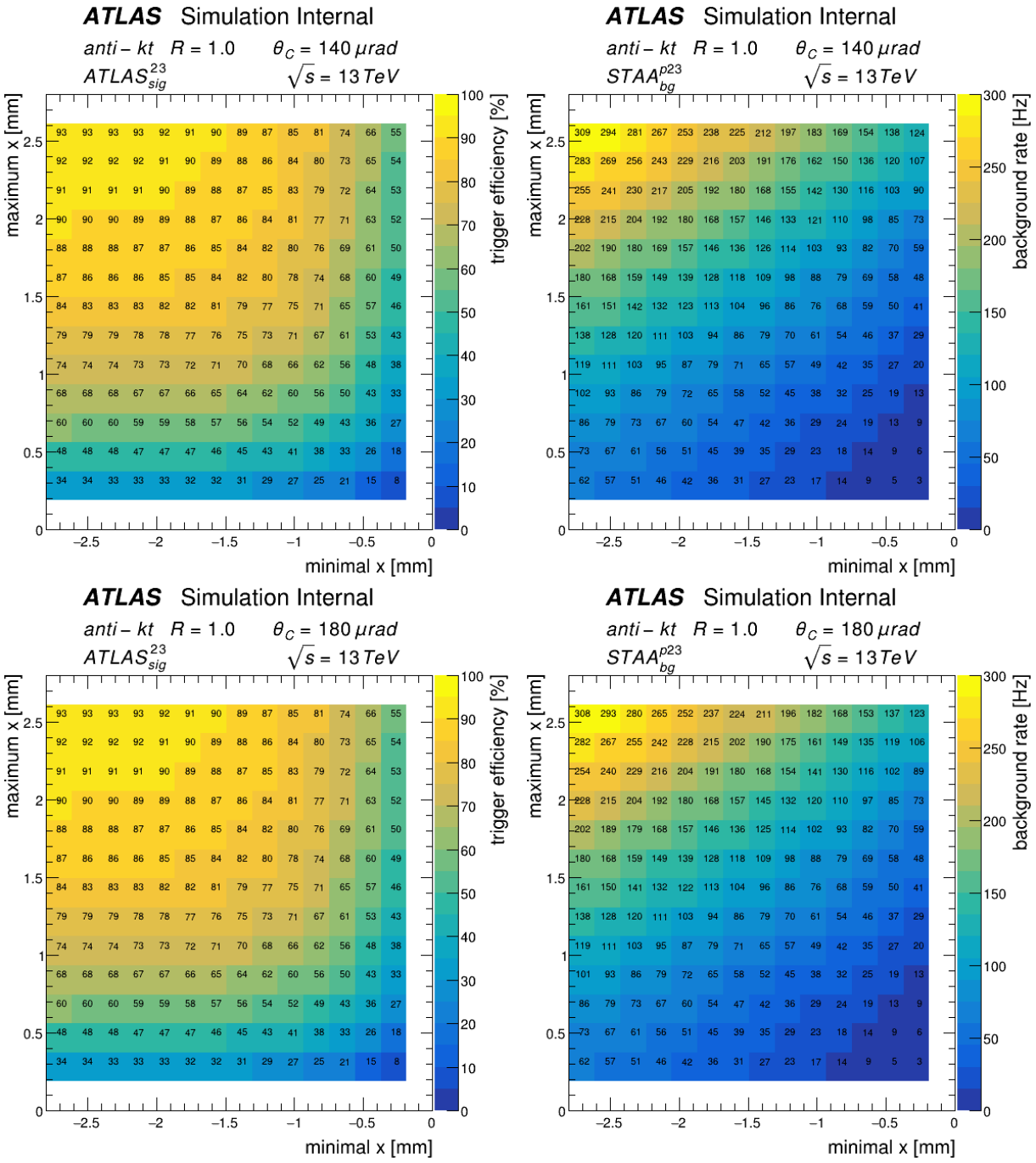


Figure B.23: The trigger efficiency (**left**) and background rates (**right**) for different values of the asymmetrical x cut. The jets were reconstructed using the anti-kt with R=1.0 whereas the forward proton position was calculated using $\theta_C = 140$ μrad (**top**) or 180 μrad (**bottom**).

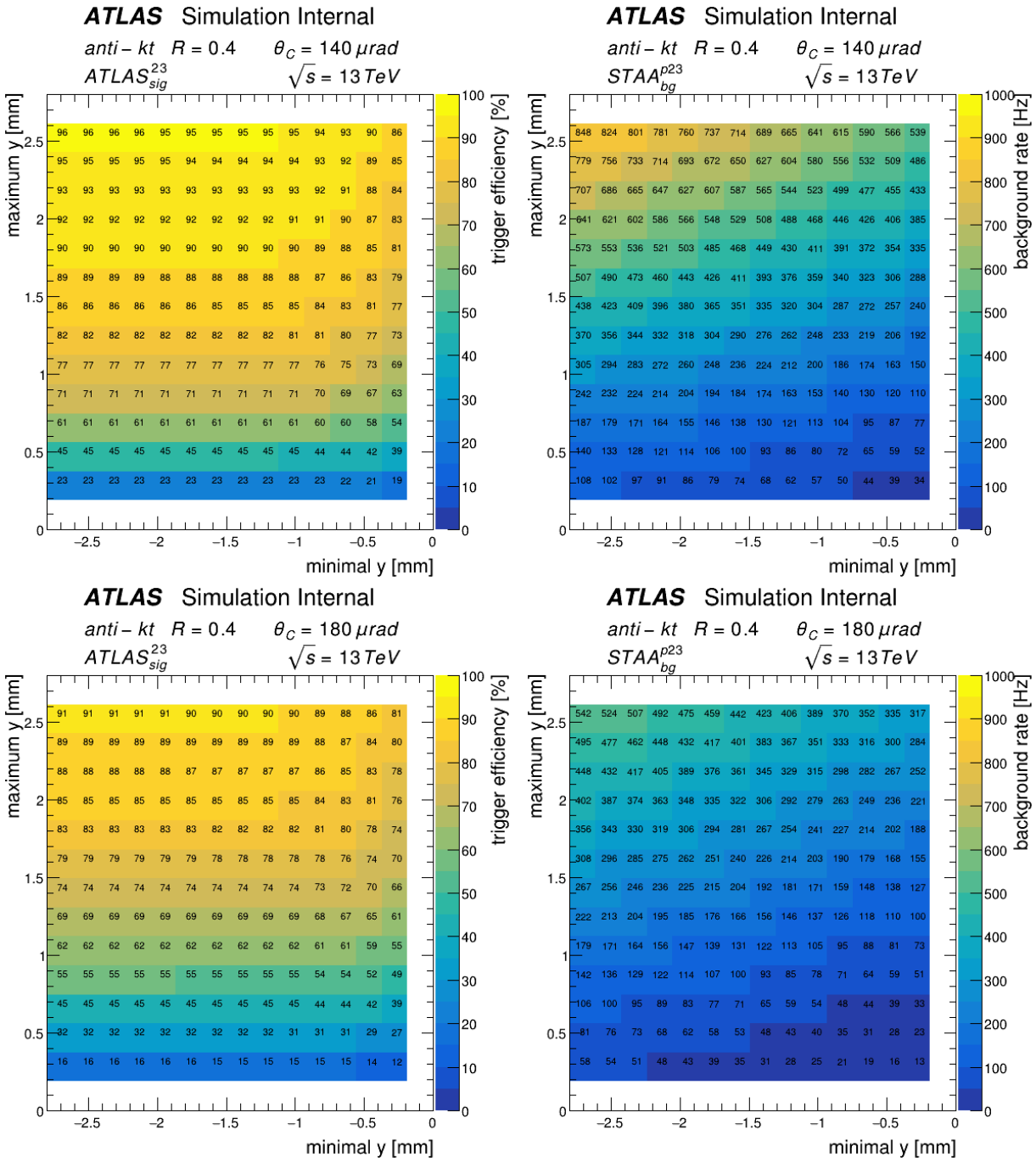


Figure B.24: The trigger efficiency (**left**) and background rates (**right**) for different values of the asymmetrical y cut. The jets were reconstructed using the anti-kt with R=0.4 whereas the forward proton position was calculated using $\theta_C = 140$ μrad (**top**) or 180 μrad (**bottom**).

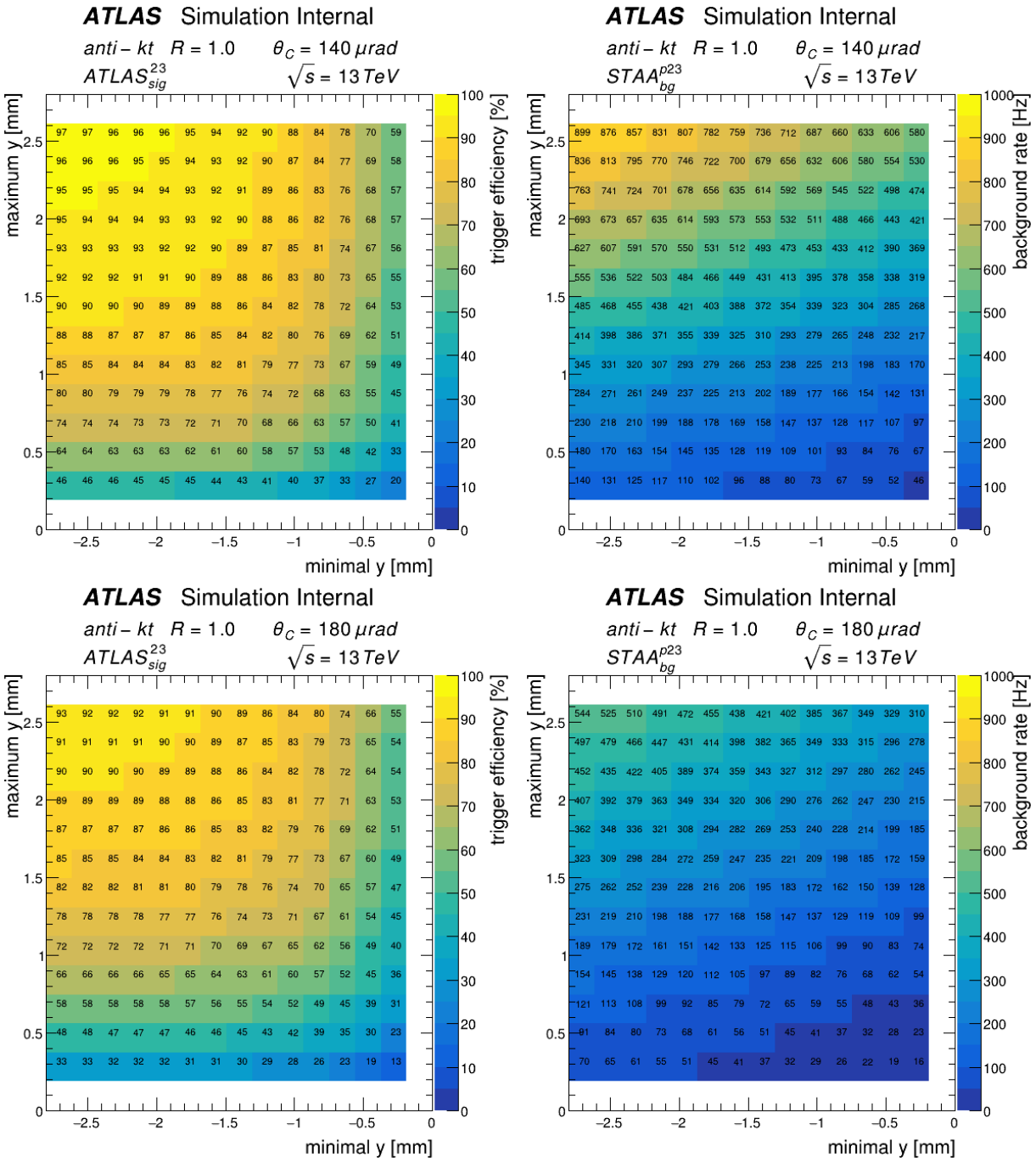


Figure B.25: The trigger efficiency (**left**) and background rates (**right**) for different values of the asymmetrical y cut. The jets were reconstructed using the anti-kt with R=1.0 whereas the forward proton position was calculated using $\theta_C = 140$ μrad (**top**) or 180 μrad (**bottom**).

B.3.3 Summary

The goal of this study was an investigation of the HLT exclusive trigger algorithm performance and its possible improvements. First, the dependence on the LHC optics setting was discussed as it is subject to changes during the data-taking. It was shown that there is no dependence on β^* value for the range foreseen in the levelling procedure and the change in the crossing angle has a small impact on the trigger efficiency.

Next, the pile-up influence on jet reconstruction was studied. The comparison between the STAA and ATLAS MC samples for the signal and background events was presented. It was concluded that the ATLAS pile-up suppression algorithms work well and the $ATLAS^{p23}_{sig}$ can be compared to $STAA^{p1}_{sig}$. Since these two samples had similar properties it was assumed that also the background standalone sample $STAA^{p1}_{bg}$ will behave similarly as the ATLAS simulation. Using these, the trigger efficiency and the expected background rate as a function of trigger cuts were estimated.

Finally, the trigger efficiencies and background rates were calculated for newly proposed trigger conditions: asymmetrical x and y and elliptical cuts. It is worth adding these extra conditions to the existing software, as they may provide a better background reduction. However, it should be noted that at the beginning of the data-taking, the AFP detectors must be properly aligned. Without such information available in the trigger software the fine-tuning of the cuts will not be possible.

The studies of the LHC optics resulted in a very small difference in efficiency due to the change of crossing angle. However, it should be also kept in mind that the real LHC conditions may result in some shifts in the proton positions. Therefore, it is suggested to use the cuts applying some safety margin, *e.g.* 0.5 mm.

It should be also pointed out that for effective studies of the exclusive jet properties, a full trigger chain has to be ready. This includes the L1 Topo algorithm, ToF trigger (first two trains), the discussed HLT algorithm and an additional HLT algorithm based on ToF – the vertex reconstruction. In addition, the jets seeded at L1 by $\sim$60 GeV object, should not contribute by more than 10 Hz at the HLT level rate.

Appendix C

AFP Jet Triggers Efficiencies

In all elementary particle studies the triggers take an important part in the analysis. They are the first step in the event selection procedure (*cf.* Section 3.6). This selection of events plays a crucial role as an inaccurate selection may lead to false results and conclusions. It is essential to choose a suitable trigger(s) for each analysis. This work is focusing on processes that have both the forward proton(s) and the central jets, thus the trigger requirements should take this fact into account. The studies dedicated to understanding the efficiencies of triggers considered in the main analysis were performed and are described in this Appendix.

C.1 AFP Jet Triggers Menu

The existence of a forward proton can be checked by looking at events with signals in the AFP stations. In Run 2 the AFP triggers were present at the first level trigger, L1. In each station, three planes were predefined as the ones that may trigger an event. The station reports the "ok" status when 2 out of 3 planes give a signal. Information about the station status is stored by dedicated triggers (see Table C.1). Those triggers were made to not trigger an event by itself since it would cause a high rate, however, they are useful in the trigger studies. Information about their association with a specific station and side is helpful and preset (accessible) in events saved due to the other triggers. The BGRP0 (Bunch Group 0) in the name of those triggers means that all filled bunches were taken into consideration.

The requirement of jets is usually satisfied by requesting at least one jet with a predefined minimum transverse momentum. This can be done either at the L1 or HLT level. In this analysis three jet triggers are studied: HLT_j10, HLT_j20 and L1J12. Chosen jet triggers are available in a configuration with the AFP triggers.

Table C.1: List of the AFP triggers present in Run 2 data-taking period and the list of stations sending a signal needed for an event to be accepted by the trigger.

Trigger Name	Stations Used	Used in Physics?
L1AFP_FSA_SIT_BGRP0	station A Far	No
L1AFP_NSA_BGRP0	station A Near	No
L1AFP_NSC_BGRP0	station C Near	No
L1AFP_FSC_SIT_BGRP0	station C Far	No
L1AFP_A	stations A Far and A Near	Yes
L1AFP_C	stations C Far and C Near	Yes
L1AFP_A_OR_C	either side A or C	Yes
L1AFP_A_AND_C	both sides A and C	Yes

C.2 Methods for Calculating Efficiencies

The trigger efficiencies can be estimated using various methods. This work will focus on one of them, the so-called "minimum-bias" method. This method uses the minimum bias trigger as a monitor trigger[1]. It is unbiased since the monitor trigger is independent of the signal trigger. However, it has one major disadvantage: low statistics. This is due to separate prescales (see Section 3.6) applied to the AFP/jet and minimum bias triggers.

The efficiencies of the L1 (L1Trigger) and HLT (HLTTrigger) triggers are calculated as

$$\mathbf{EFF_{L1Trigger}} = \frac{\mathbf{Nev}_{\text{ L1Trigger(BP) and L1MonitorTrigger(AP) and AdditionalCut}}}{\mathbf{Nev}_{\text{ L1MonitorTrigger(AP) and AdditionalCut}}},$$

$$\mathbf{EFF_{HLTTrigger}} = \frac{\mathbf{Nev}_{\text{HLTTrigger(AP) and HLTMonitorTrigger(AP) and AdditionalCut}}}{\mathbf{Nev}_{\text{ HLTMonitorTrigger(AP) and AdditionalCut}}},$$

where:

- L1Trigger is one of the triggers of interest: L1AFP_A, L1AFP_C, L1AFP_A_OR_C, L1AFP_A_AND_C, L1AFP_A_OR_C_J12, L1AFP_A_AND_C_J12, L1J12, L1AFP_FSA_SIT_BGRP0, L1AFP_NSA_BGRP0, L1AFP_NSC_BGRP0, L1AFP_FSC_SIT_BGRP0.
- HLTTrigger is either HLT_j10 or HLT_j20.
- L1MonitorTrigger means the decision of HLT_mb_sptrk.
- HLTMonitorTrigger means the decision of L1MBTS_2.
- AdditionalCut is the requirement of the reconstructed proton on the AFP side/s, the reconstructed track in the AFP station or the presence of a jet with minimum p_T of 12 GeV, depending on which trigger efficiency is studied.

[1]Only events that passed the minimum bias trigger were considered.

◦ BP (AP) means that the trigger fired before (after) prescale.

C.3 Level 1 Triggers

As was mentioned in Section 3.6 the trigger chain starts with the L1 trigger. The analysis of the efficiencies of such triggers was done including the time, the leading jet p_T and the proton energy loss dependencies.

C.3.1 Trigger Efficiency – The Time Dependence

In the first step, the influence of the time flow on the triggers' efficiency will be checked. In Figures [C.1-C.5] the efficiency of different triggers as a function of the LB is shown for various runs. One should note that only LBs which had the status "ready for physics" and had the AFP inserted into the beam pipe were used and the first bin on those plots represents the first LB fulfilling these conditions.

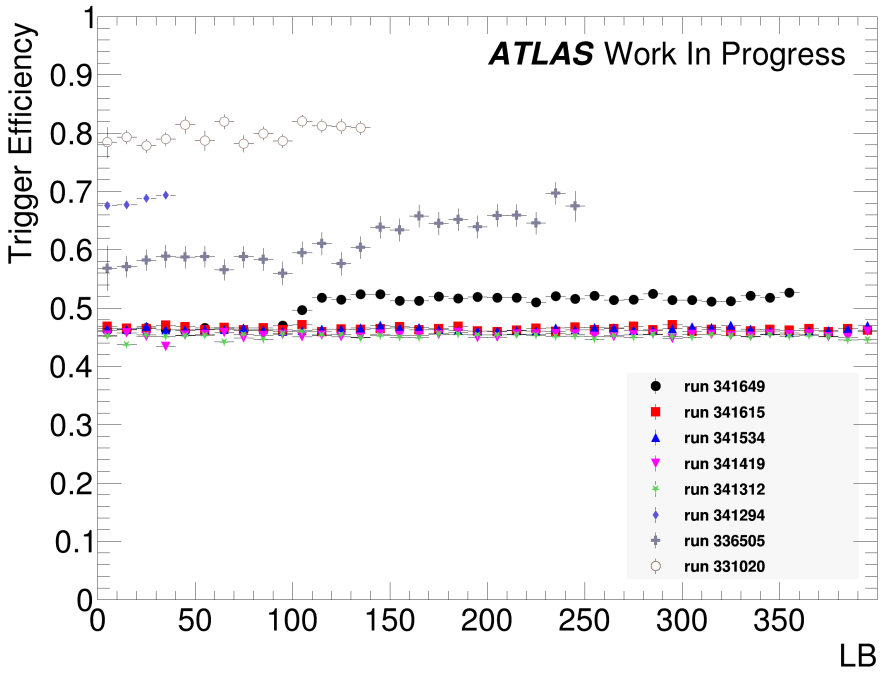


Figure C.1: Efficiencies of the trigger L1AFP_A_OR_C as a function of LBs with the AFP inserted and "ready for physics" status.

Some interesting features can be observed studying the efficiency of the L1AFP_A_OR_C trigger presented in Figure C.1. The runs 341312 (green stars), 341419 (magenta triangles) and 341615 (red squares) are behaving in the same way, keeping a constant efficiency at the level of 46%. For these runs the L1AFP_A trigger efficiency (Figure C.2) is much smaller

than that of L1AFP_C (Figure C.3). The efficiency is lowered due to the performance of the A Far station (see Figure C.4a). It is the only station which had only 3 planes present (2 in run 331020) hence the lower efficiency.

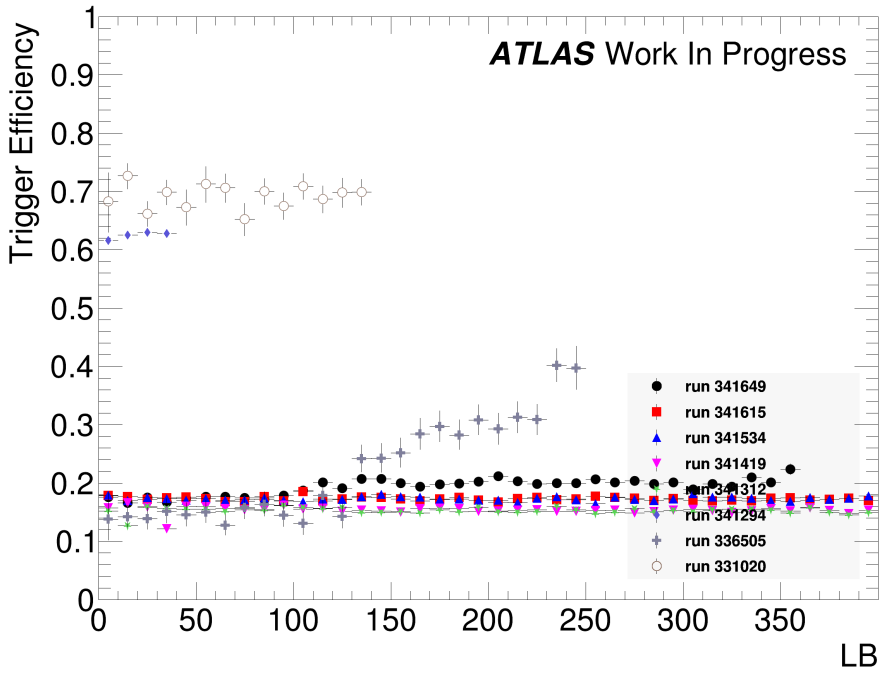


Figure C.2: Efficiency of the trigger L1AFP_A as a function of LBs with the AFP inserted and "ready for physics" status.

The jump in the L1AFP_A_OR_C trigger efficiency can be observed for run 341649 (black circles). One can notice a 10% increase in the middle of the run. However, such behaviour is somehow expected as the pile-up change from μ ~2.0 to μ ~1.0 in this particular run. Also, observation of the triggers for singular sides and stations brings interesting facts. For the C side, the jump is clearly visible, however, in the L1AFP_A (Figure C.2) case it is not so apparent. Further analysis reveals that this jump is not present for the A Far station (Figure C.4a).

The last efficiency increase can be observed for run 336505 (grey crosses), where in the middle of the run the efficiency of L1AFP_A_OR_C continuously grows by about 10%. It was found that for that LB range, the noisy pixels start to appear resulting in higher trigger rates. It is expected that for this run the L1AFP_A_OR_C trigger rate would be the highest since this run pile-up is the lowest (μ ~0.04). This is not the case as the highest efficiency is observed for the earliest run 331020 (not filled circles) with μ ~1.0. The more detailed investigation reviles that the unexpected lower performance is caused solely by station A Far (Figure C.4 and C.5), which efficiency is much higher for run 331020 (this is caused probably by a different triggering scheme). For the rest of the stations, the efficiency is as expected highest for the run 336505.

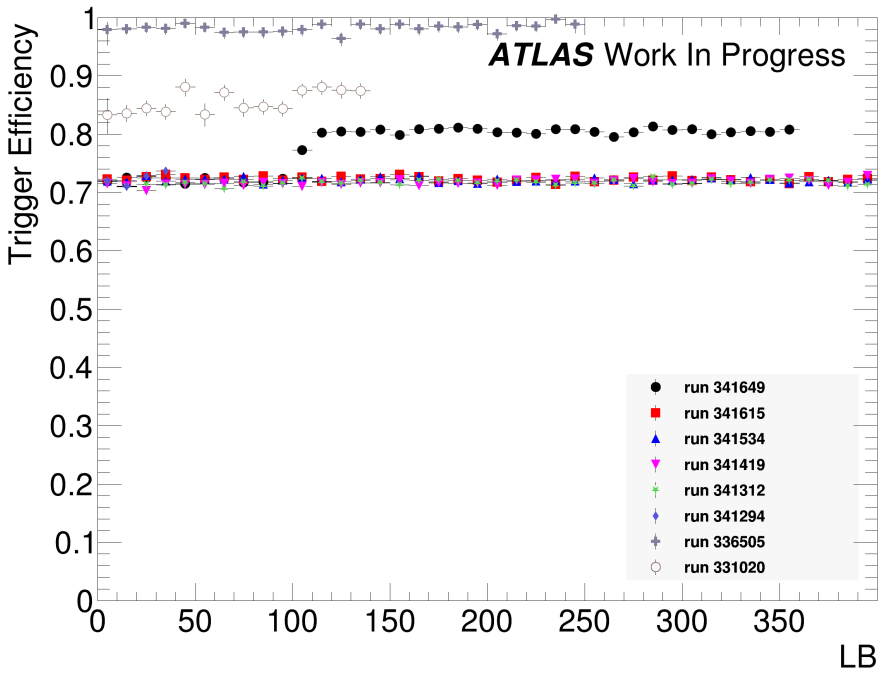


Figure C.3: Efficiency of the trigger L1AFP_C as a function of LBs with the AFP inserted and "ready for physics" status.

The second highest efficiency is seen for the shortest run 341294 (purple diamonds). This is once again caused by the A Far station and the presence of noisy pixels.

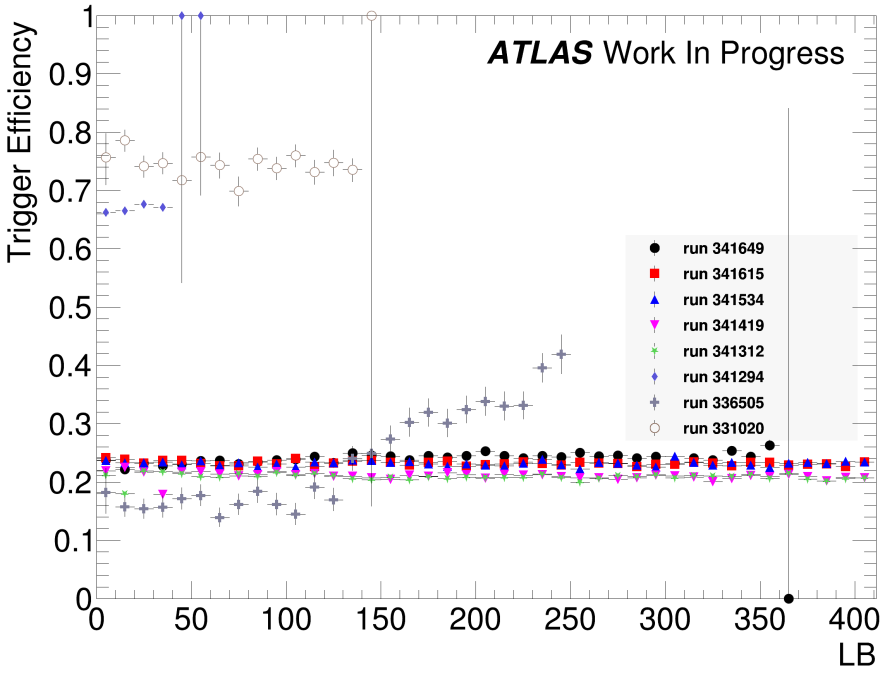
The observed results imply that the triggering using a plane with the turned-off HV (Station A Far Plane 2) significantly lowers the efficiency and makes the trigger unstable since the noisy pixels are more likely to spawn in the AFP detector.

C.3.2 Dead Time Effect

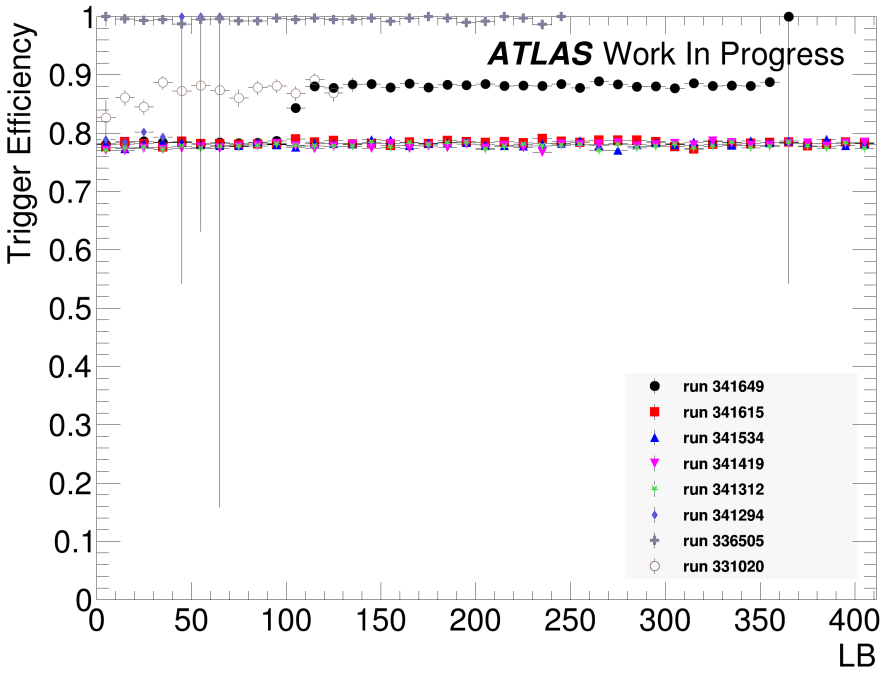
The efficiency dependence on a single bunch in a train was studied. It is expected that the AFP SiT-based triggers will show variations of efficiency due to the so-called dead time effect. The dead time is a period of time during which a trigger is inactive after triggering an event. During this time a trigger cannot provide another signal.

The bunch structure is important in this part of studies as it may affect the efficiency. The medium-μ runs (with the exception of 341294) have the same structure. The first train has 10 consecutive bunches instead of a typical number of 8. The additional two bunches in the first train are visible in Figures [C.6 - C.8], in the 9th and 10th bin.

The trigger efficiencies as a function of the bunch number with the first bin representing the first bunch in the train are shown, cf. plots C.6 and C.8. In all medium-μ runs, the dead time effect is clearly visible (e.g. see Fig. C.6). The efficiency drops by ~20% in the first bunches and after the 7 bunches-long-pause, the trigger recovers to a slightly lower

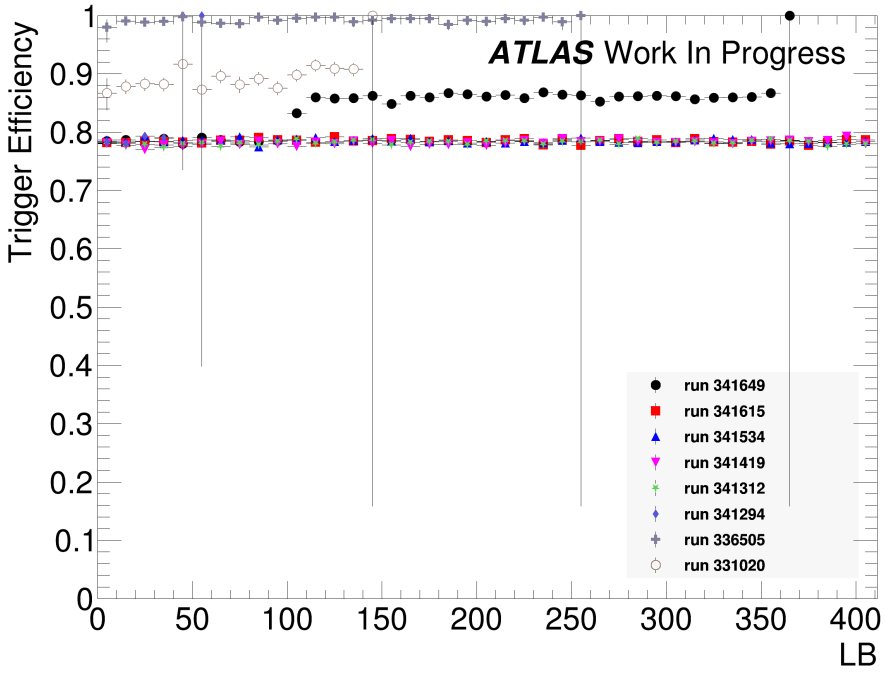


(a) L1AFP_FSA_SIT_BGRP0

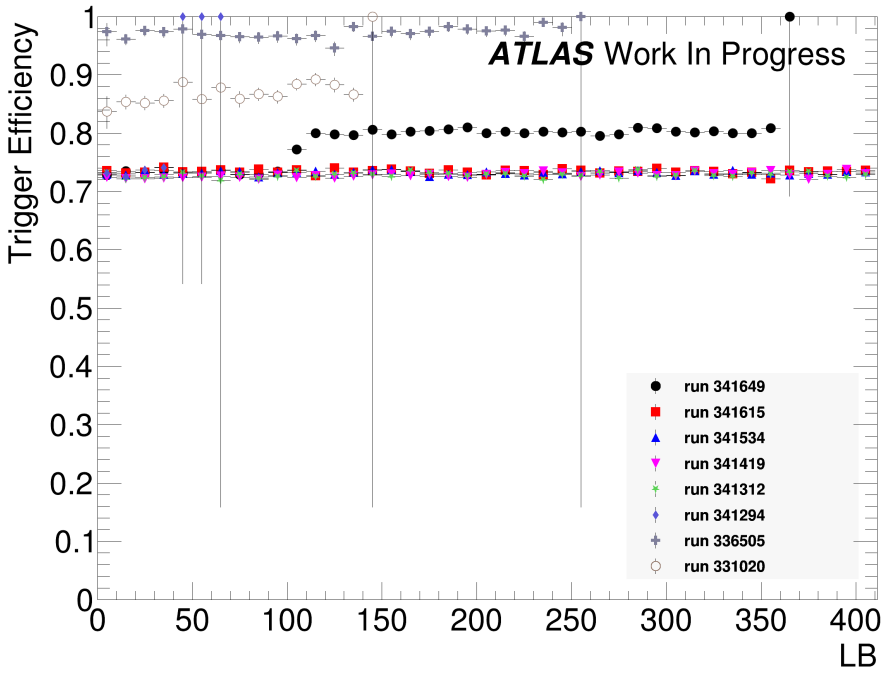


(b) L1AFP_NSA_BGRP0

Figure C.4: Efficiencies of the triggers as a function of the LBs with the AFP inserted and "ready for physics" status.



(a) L1AFP_NSC_BGRP0



(b) L1AFP_FSC_SIT_BGRP0

Figure C.5: Efficiencies of the triggers as a function of the LBs with the AFP inserted and "ready for physics" status.

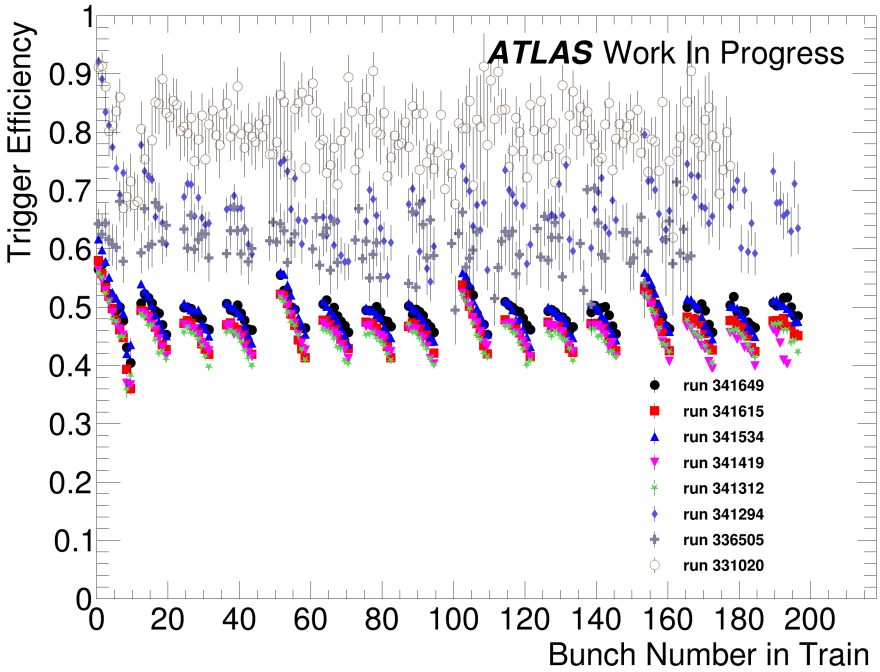


Figure C.6: Trigger L1AFP_A_OR_C efficiency as a function of the bunch number in a train.

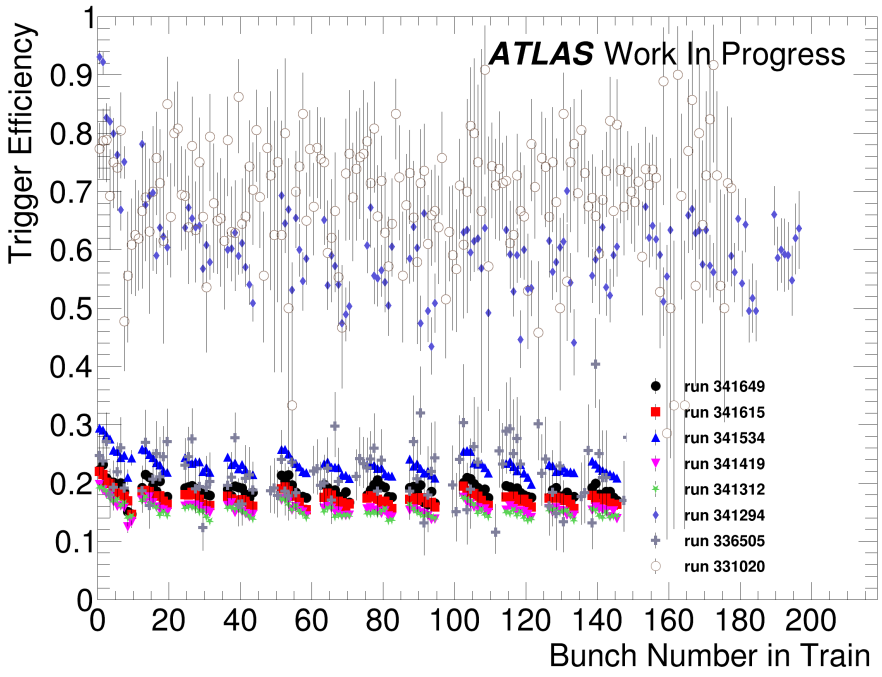
value than the starting one and the rate decreases again. Smaller, about 10%, drops in the efficiency are visible after each shorter pause of 4 bunches when the trigger did not have enough time to fully recover. The same structure is visible in all triggers for each side and station (Figures C.7-C.9).

Run 331020 is the only run with a bunch structure different from the other runs. Unfortunately, its statistics is too low to determine if the dead-time effect is present, not only for the L1AFP_A_OR_C trigger but also for the triggers tagging on a singular side or station.

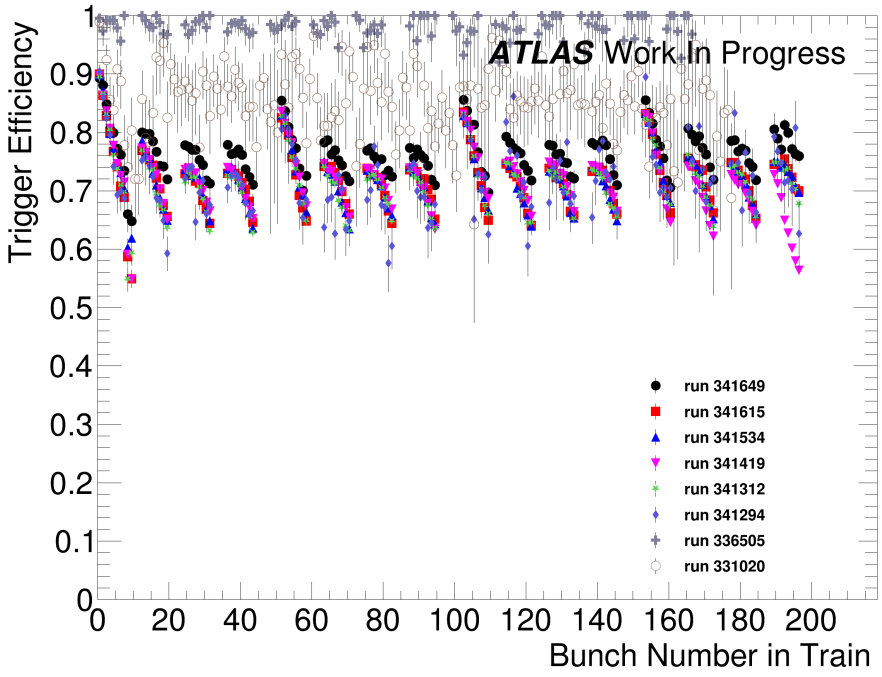
The dead time effect is also not visible for run 336505. However, it is not due to its bunch structure. It is similar to that of the electro-weak runs where the dead time effect is clearly visible. The reason for this observation is the very low pile-up ($\mu \sim 0.04$). In this run, the efficiencies for stations on side C and station Near A are very close to 100%.

C.3.3 Trigger Efficiency – Leading Jet p_T and Proton ξ Dependence

Lastly, the dependence on the proton ξ and the leading jet p_T is studied. The AFP triggers performance as a function of the scattered proton energy loss. is presented in Figure C.10. Such a dependence was plotted for the L1AFP_A_OR/AND_C and L1AFP_A/C triggers. It can be observed that the efficiency within the AFP acceptance is flat for both

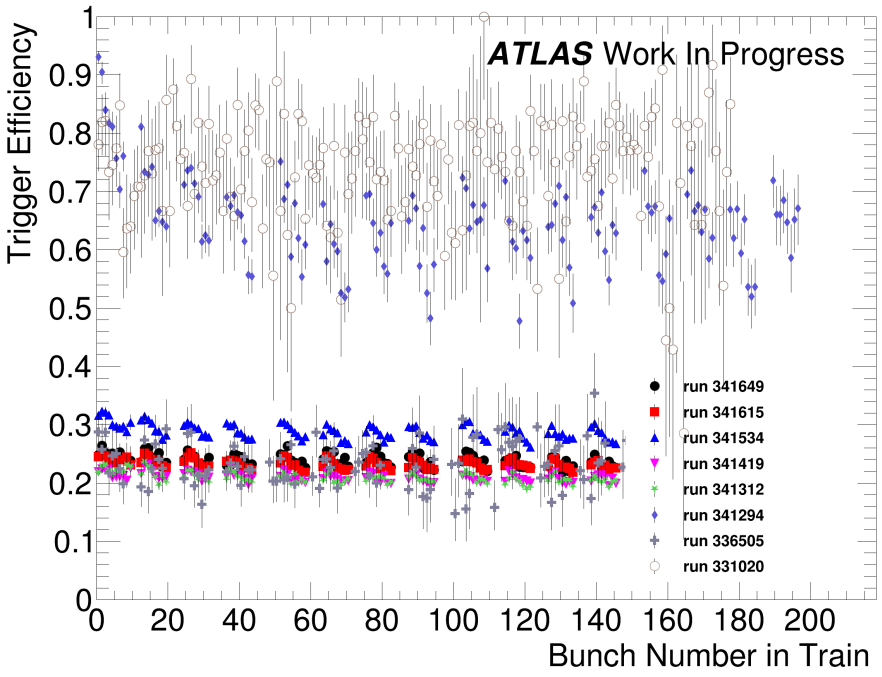


(a) L1AFP_A

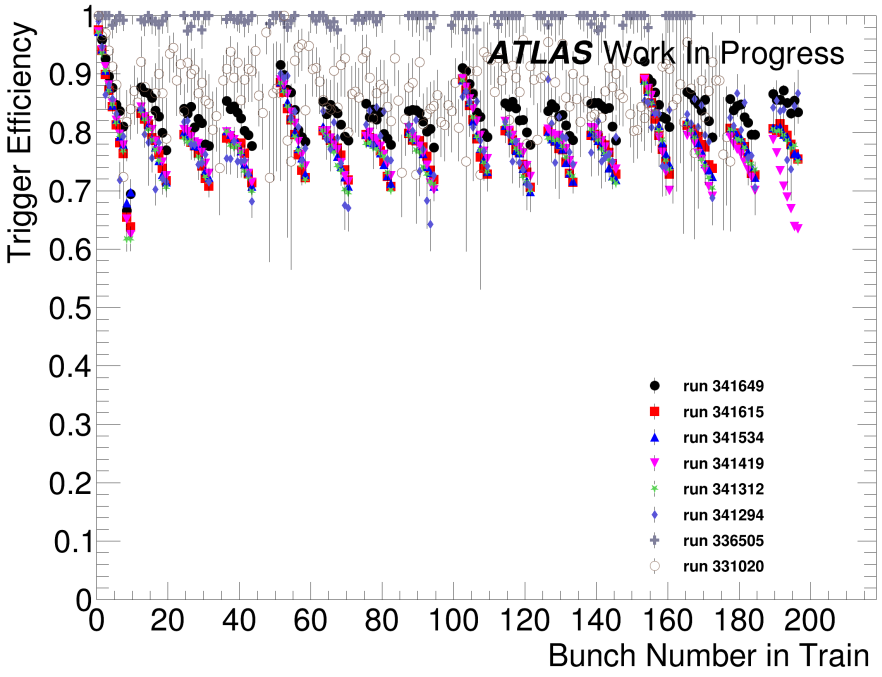


(b) L1AFP_C

Figure C.7: Trigger efficiency as a function of the bunch number in a train.

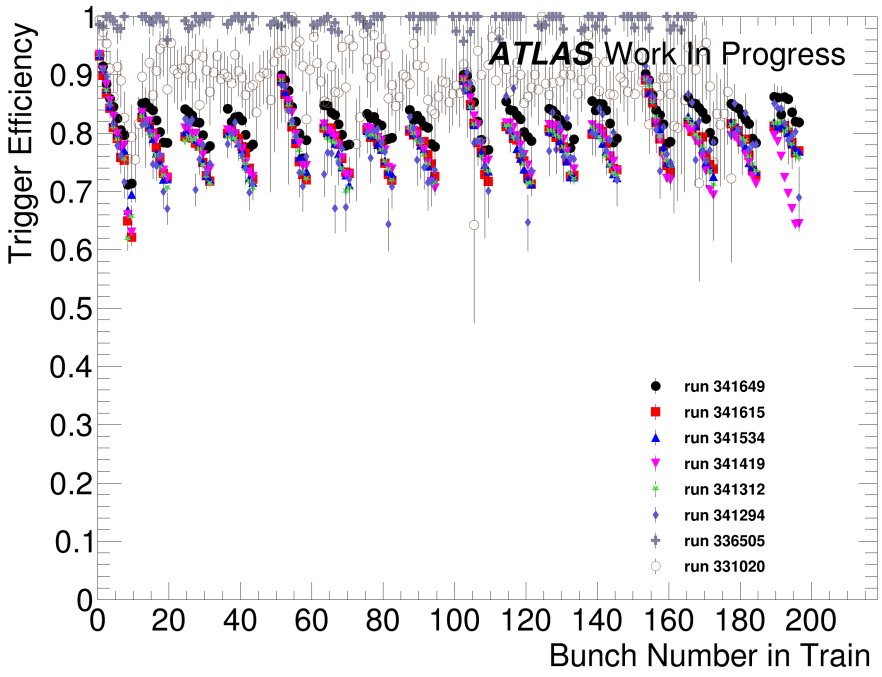


(a) L1AFP_FSA_SIT_BGRP0

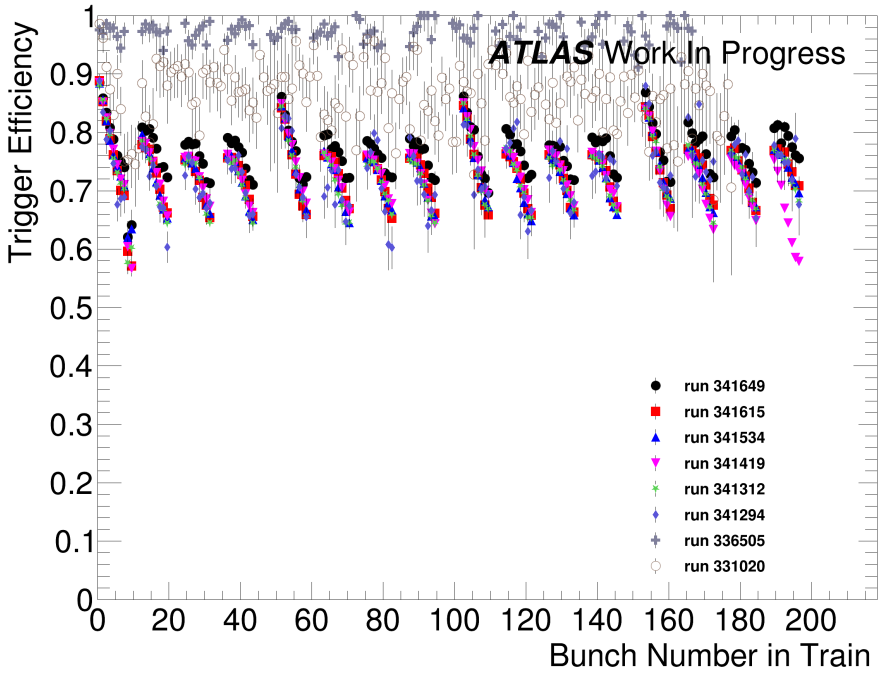


(b) L1AFP_NSA_BGRP0

Figure C.8: Trigger efficiency as a function of the bunch number in a train.



(a) L1AFP_NSC_BGRP0



(b) L1AFP_FSC_SIT_BGRP0

Figure C.9: Trigger efficiency as a function of the bunch number in a train.

L1AFP_C and L1AFP_A_AND_C, except for a small energy loss region of ξ <0.05. The L1AFP_A_OR_C trigger is influenced by the L1AFP_A efficiency which increases for larger ξ. As previously, this different behaviour is related to station A Far and for fully working stations one should expect a constant trigger efficiency as a function of the scattered proton energy loss.

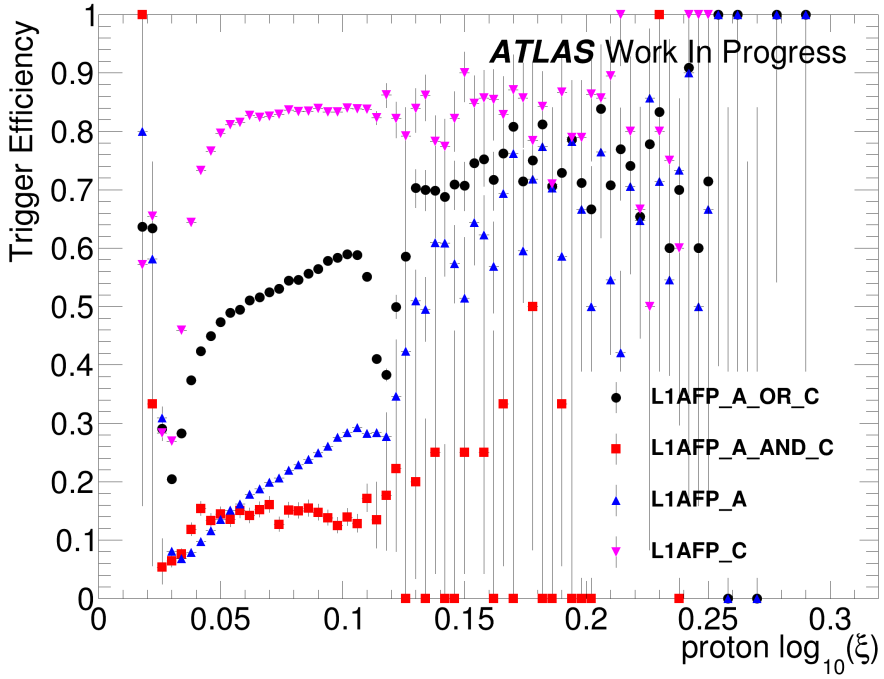
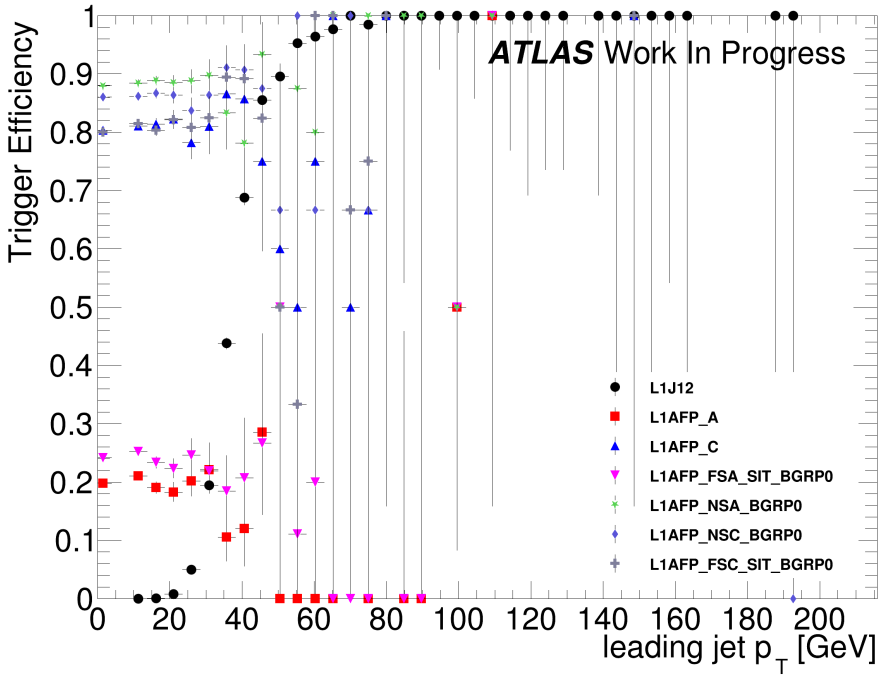


Figure C.10: Efficiency of the AFP triggers in run 341649 as a function of the proton energy loss.

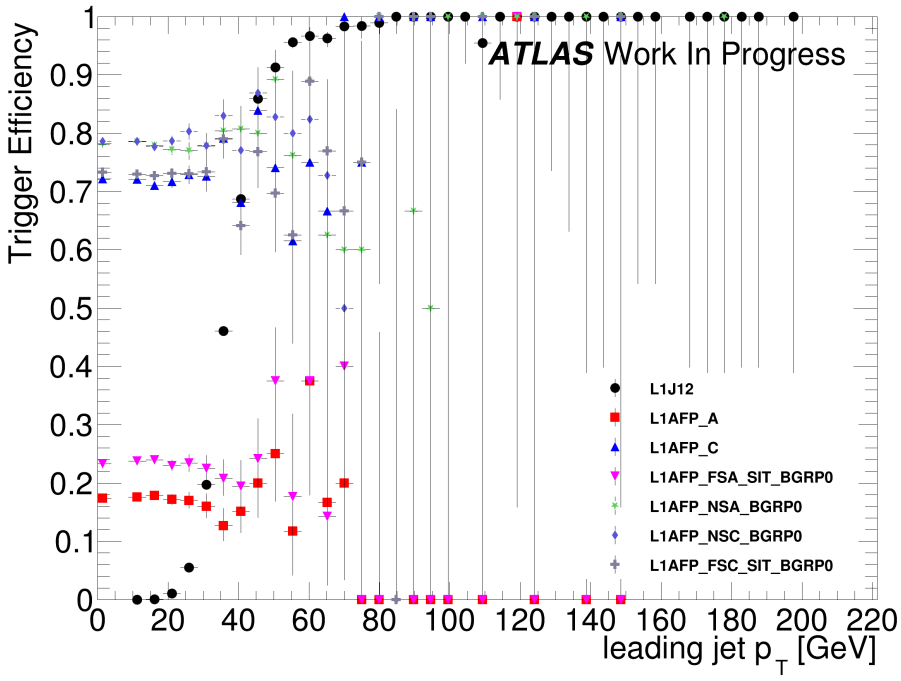
In Figure C.11 efficiencies of the triggers as a function of the leading jet p_T are presented for the pile-up values ∼1.0 (Figure C.11a) and ∼2.0 (Figure C.11b). The AFP triggers do not depend on the leading jet p_T (constant efficiency visible on plots). It is worth noticing that events accepted by the AFP triggers are dominated by the small p_T jets. In the L1J12 trigger case, the efficiency stays the same despite the pile-up change. This trigger is not efficient for the lower transverse momentum jets and reaches full efficiency only for jets with $p_T > 60$ GeV. Therefore, in the analysis considering jets with very small transverse momentum, it is especially important to include this trigger efficiency.

C.4 HLT Triggers

The last thing to discuss is the efficiency of the HLT jet triggers, which are used along with AFP triggers. In Figure C.12 the efficiency of the HLT_10/20 triggers as a function of the leading jet p_T is presented. Both triggers start to be fully effective for $p_T \sim 40$ GeV.



(a) run 341649 with μ <1.5



(b) run 341649 with μ >1.5

Figure C.11: Efficiency of the triggers as a function of the leading jet p_T in run 341649 with μ ∼2.0 and 1.0 (**top** and **bottom**).

Similarly to the L1J12 trigger, both trigger efficiencies need to be considered in analysis using low p_T jets.

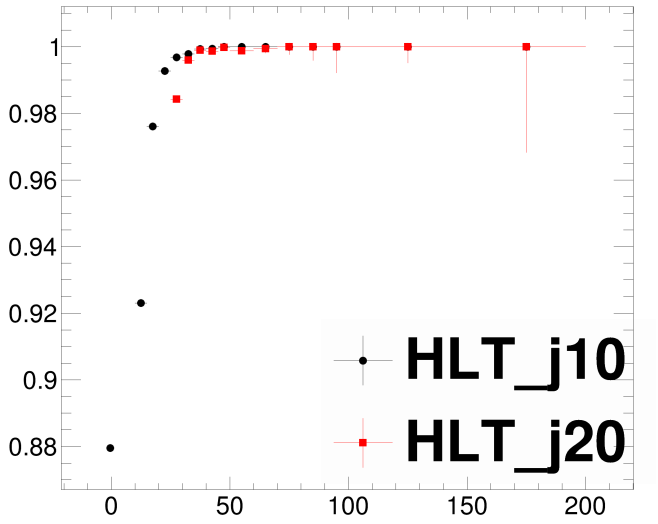


Figure C.12: Efficiency of the HLT triggers as a function of the leading jet transverse momentum in run 341649.

C.5 Summary

The AFP triggers performance was studied taking into account different aspects that could influence their efficiencies.

It was found that during the data-taking with fully working stations, the AFP triggers should perform at a constant level, being sensitive only to the changing pile-up. It is expected that for the pile-up close to 0, the AFP triggers should maintain the efficiency very close to 100%, with the efficiency dropping to about 90% for μ ~1.0 and 80% for μ ~2.0. It is also expected for the Far stations' efficiency to be smaller than those of the Near stations by ~5% since more noise is expected in these stations. The singular station efficiency will affect the full efficiency of the AFP trigger as it uses information from both Near and Far stations, triggering on one or both sides, not on a singular station.

On the other hand, the stations which are not fully functional demonstrate a drop in the efficiency of the triggers. Additionally, the triggering on the planes with the HV turned off may result in some periods of time with higher trigger rates as such planes are prone to have hot pixels.

In this study, the Dead Time Effect was also analysed. While the effect was not visible for run 331020 due to the low statistics and run 336505 due to a very low pile-up and efficiency being close to 100%, it was present in all electro-weak runs. In those runs, the

efficiency was dropping to the same minimum during the filled bunches of a train and was recovering during the empty ones. The space of seven empty bunches was sufficient for the triggers to recover to their starting efficiency.

Last but not least the dependence on the transverse momentum of the leading jet was investigated. The AFP triggers were found to not depend on p_T^{jet} and to maintain good performance even for jets with smaller values of p_T^{jet} for which the jet triggers did not perform well.

Finally, the influence of the reconstructed proton energy loss on the AFP trigger efficiencies was discussed. From the edge of the detector, the efficiency quickly grows and maintains a stable level with increasing energy loss of a proton.

Bibliography

[1] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716:1, 2012.

[2] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett. B*, 716:30, 2012.

[3] R. L. Workman et al. Review of Particle Physics. *PTEP*, 2022:083C01, 2022.

[4] J I Friedman and H W Kendall. Deep inelastic electron scattering. *Annual Review of Nuclear Science*, 22(1):203–254, 1972.

[5] Richard E. Taylor. Deep inelastic scattering: The early years. *Rev. Mod. Phys.*, 63:573–595, Jul 1991.

[6] Henry W. Kendall. Deep inelastic scattering: Experiments on the proton and the observation of scaling. *Rev. Mod. Phys.*, 63:597–614, 1991.

[7] F. D. Aaron et al. Combined Measurement and QCD Analysis of the Inclusive e+- p Scattering Cross Sections at HERA. *JHEP*, 01:109, 2010.

[8] Richard P. Feynman. Very high-energy collisions of hadrons. *Phys. Rev. Lett.*, 23:1415–1417, 1969.

[9] Graeme Watt. MSTW PDFs and impact of PDFs on cross sections at tevatron and LHC. *Nuclear Physics B - Proceedings Supplements*, 222-224:61–80, jan 2012.

[10] Hung-Liang Lai et al. New parton distributions for collider physics. *Phys. Rev. D*, 82:074024, Oct 2010.

[11] P. Jimenez-Delgado and E. Reya. Dynamical next-to-next-to-leading order parton distributions. *Phys. Rev. D*, 79:074023, Apr 2009.

[12] P. D. B. Collins. *An Instroduction to Regge Theory & High Energy Physics*. Cambrigde University Press, 1977.

[13] A. Donnachie and P. V. Landshoff. Total cross-sections. *Phys. Lett. B*, 296:227–232, 1992.

[14] K. Goulianos. Diffraction in hadron-hadron interactions. *Nuclear Physics B - Proceedings Supplements*, 99(1-2):9–20, apr 2001.

[15] Sandy Donnachie, Günter Dosch, Peter Landshoff, and Otto Nachtmann. *Pomeron Physics and QCD*. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 2002.

[16] G. Ingelman and P.E. Schlein. Jet structure in high mass diffractive scattering. *Physics Letters B*, 152(3):256–260, 1985.

[17] R. Bonino et al. Evidence for transverse jets in high-mass diffraction: Ua8 experiment. *Physics Letters B*, 211(1):239–246, 1988.

[18] A Brandt et al. Cross section measurements of hard diffraction at the SS-collider. *Physics Letters B*, 421(1-4):395–404, mar 1998.

[19] T. Affolder and et. all Akimoto. Diffractive dijets with a leading antiproton in pp collisions at $\sqrt{s} = 1800$ gev. *Physical Review Letters*, 84(22):5043 – 5048, 2000.

[20] A.B. Kaidalov, V.A. Khoze, A.D. Martin, and M.G. Ryskin. Probabilities of rapidity gaps in high energy interactions. *The European Physical Journal C*, 21(3):521–529, jul 2001.

[21] A. Breskin, European Council for Nuclear Research, and R. Voss. *The Cern Large Hadron Collider: Accelerator and Experiments*. Number Bd. 12 in The Cern Large Hadron Collider: Accelerator and Experiments. CERN, 2009.

[22] LHC Study Group. and European Organization for Nuclear Research. *Design study of the Large Hadron Collider (LHC) : a multiparticle collider in the LEP tunnel / The LHC Study Group*. CERN Geneva, 1991.

[23] J R J Bennett, C J Densham, and P V Drumm. Test of a high power target design. Technical report, CERN, Geneva, 1996.

[24] Michael Barnes et al. *The PS complex as proton pre-injector for the LHC: design and implementation report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2000.

[25] J P Riunaud et al. An Antiproton Decelerator in the CERN PS Complex. page 5 p, 1996.

[26] n-TOF Technical Design Report on Experimental Area. Technical report, CERN, Geneva, 2000.

[27] A. Brandt et al. Cross-section measurements of hard diffraction at the SPS collider. *Phys. Lett. B*, 421:395–404, 1998.

[28] Klaus G. Steffen. *High energy beam optics*, volume 17 of *Interscience monographs and texts in physics and astronomy*. Interscience Publishers, New York, 1965.

[29] ATLAS Collaboration. ATLAS Forward Proton Phase-I Upgrade: Technical Design Report. ATLAS-TDR-024; CERN-LHCC-2015-009, 2015.

[30] ATLAS Collaboration. ATLAS Forward Detectors for Measurement of Elastic Scattering and Luminosity: Technical Design Report. ATLAS-TDR-18; CERN-LHCC-2008-004, 2008.

[31] M. Trzebiński. Machine Optics Studies for the LHC Measurements. *Proc. SPIE Int. Soc. Opt. Eng.*, 9290:929026, 2014.

[32] K. Aamodt et al. The ALICE experiment at the CERN LHC. *JINST*, 3:S08002, 2008.

[33] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.

[34] CMS Collaboration. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[35] LHCb Collaboration. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008. Also published by CERN Geneva in 2010.

[36] The TOTEM Collaboration. The TOTEM experiment at the CERN large hadron collider. *Journal of Instrumentation*, 3(08):S08007–S08007, aug 2008.

[37] The LHCf Collaboration. The LHCf detector at the CERN large hadron collider. *Journal of Instrumentation*, 3(08):S08006–S08006, aug 2008.

[38] James Pinfold et al. Technical Design Report of the MoEDAL Experiment. Technical report, 2009.

[39] ATLAS Collaboration. ATLAS Inner Detector: Technical Design Report, Volume 1. ATLAS-TDR-4; CERN-LHCC-97-016, 1997.

[40] ATLAS Collaboration. ATLAS Pixel Detector: Technical Design Report. ATLAS-TDR-11; CERN-LHCC-98-013, 1998.

[41] ATLAS Collaboration. ATLAS Insertable B-Layer: Technical Design Report. ATLAS-TDR-19; CERN-LHCC-2010-013, 2010.

[42] ATLAS Collaboration. Operation and performance of the ATLAS semiconductor tracker. *JINST*, 9:P08009, 2014.

[43] Adrian Vogel. Atlas transition radiation tracker (trt): Straw tube gaseous detectors at high rates. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 732:277–280, 2013. Vienna Conference on Instrumentation 2013.

[44] ATLAS Collaboration. ATLAS Detector and Physics Performance: Technical Design Report, Volume 1. ATLAS-TDR-14; CERN-LHCC-99-014, 1999.

[45] ATLAS Collaboration. ATLAS Liquid Argon Calorimeter: Technical Design Report. ATLAS-TDR-2; CERN-LHCC-96-041, 1996.

[46] ATLAS Collaboration. ATLAS Tile Calorimeter: Technical Design Report. ATLAS-TDR-3; CERN-LHCC-96-042, 1996.

[47] ATLAS Collaboration. ATLAS Muon Spectrometer: Technical Design Report. ATLAS-TDR-10; CERN-LHCC-97-022, CERN, 1997.

[48] G. Avoni et al. The new lucid-2 detector for luminosity measurement and monitoring in atlas. *Journal of Instrumentation*, 13(07):P07017, jul 2018.

[49] ATLAS Collaboration. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. *Eur. Phys. J. C*, 77:490, 2017.

[50] ATLAS Collaboration. Charged-particle multiplicities inppinteractions measured with the atlas detector at the lhc. *New Journal of Physics*, 13(5):053033, May 2011.

[51] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. The anti-kt jet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063–063, apr 2008.

[52] T Hryn'ova and K Nagano. Trigger Menu Strategy for Run 2. Technical report, CERN, Geneva, 2014.

[53] ATLAS Collaboration. Jet reconstruction and performance using particle flow with the ATLAS Detector. *Eur. Phys. J. C*, 77:466, 2017.

[54] ATLAS Collaboration. Operation of the ATLAS trigger system in Run 2. *JINST*, 15:P10004, 2020.

[55] M. Kocian. Readout and trigger for the afp detector at atlas experiment. *Journal of Instrumentation*, 12(1), 1 2017.

[56] J. Lange et al. Beam tests of an integrated prototype of the ATLAS Forward Proton detector. *JINST*, 11(09):P09005, 2016.

[57] ATLAS Collaboration. Performance of the ATLAS Forward Proton Time-of-Flight Detector in 2017. ATL-FWD-PUB-2021-002, 2021.

[58] L. Deniau (editor), H. Grote, G. Roy and F. Schmidt. Mad-X User's Guide.

[59] R. Ciesielski and K. Goulianos. MBR Monte Carlo Simulation in PYTHIA8, 2012.

[60] GEANT4 Collaboration, S. Agostinelli, et al. GEANT4 – a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250, 2003.

[61] R. Staszewski, M. Trzebiński, and J. Chwastowski. Dynamic alignment of the forward-proton detectors at the LHC. *Advances in High Energy Physics*, 2012:1–8, 2012.

[62] R. Staszewski and J. Chwastowski. Transport Simulation and Diffractive Event Reconstruction at the LHC. *Nucl. Instrum. Meth. A*, 609:136–141, 2009.

[63] L Adamczyk et al. AFP: A Proposal to Install Proton Detectors at 220 m around ATLAS to Complement the ATLAS High Luminosity Physics Program. Technical report, CERN, Geneva, 2011. AFP technical proposal.

[64] M Trzebinski and R Staszewski. Measurement of Central Difftractive Jet Production with the ATLAS Forward Proton Detectors (Feasibility Study). Technical report, CERN, Geneva, 2013.

[65] L Adamczyk et al. Technical Design Report for the ATLAS Forward Proton Detector. Technical report, 2015.

[66] M. Trzebiński. Hard diffraction with proton tagging at the LHC. *Acta Physica Polonica B*, 46(8):1499, 2015.

[67] K. Akiba et al. LHC forward physics. *Journal of Physics G: Nuclear and Particle Physics*, 43(11):110201, oct 2016.

[68] ATLAS AFP Collaboration. ATLAS Forward Protons Silicon Tracker Performance in 2017 low pile-up runs. Technical report, CERN, Geneva, 2022.

[69] R. Staszewski, J. Chwastowski, K. Korcyl, and M. Trzebiński. Alignment-related effects in forward proton experiments at the LHC. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 801:34–43, nov 2015.

[70] ATLAS Collaboration. The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie–Landshoff diffractive model. ATL-PHYS-PUB-2016-017, 2016.

[71] Richard D. Ball et al. Parton distributions with LHC data. *Nuclear Physics B*, 867(2):244–289, feb 2013.

[72] Maciej Trzebiński. Summary of 2016 and 2017 Data Taking. Technical report.

[73] Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector. Technical report, CERN, Geneva, 2015.

[74] ATLAS Collaboration. ATLAS Pythia 8 tunes to 7 TeV data. ATL-PHYS-PUB-2014-021, 2014.

[75] M. Trzebiński. Exclusive jet measurement in special LHC runs — feasibility studies. *Acta Physica Polonica B*, 47(6):1745, 2016.

[76] Maciej Trzebiński, Rafał Staszewski, and Janusz Chwastowski. On the possibility of measuring the single-tagged exclusive jets at the LHC. *The European Physical Journal C*, 75(7), jul 2015.

[77] Paula Agnieszka Erland and Maciej Trzebinski. Dead and Low Efficiency Pixels in AFP Silicon Detector. Technical report, CERN, Geneva, 2021.

[78] C. Leonidopoulos and A. R. Martinez. Trigger menu for 2018 data taking for approval.

[79] Theodor Herwig. ATLAS jet trigger performance in 2016 data. *PoS*, ICHEP2016:854, 2017.

[80] K. Jakobs. Atlas status of the experiment. (CERN-RRB-2018-073).

[81] M. Trzebiński, R. Staszewski, and J. Chwastowski. LHC high- runs: Transport and unfolding methods. *ISRN High Energy Physics*, 2012:1–10, sep 2012.

[82] M. Boonekamp, A. Dechambre, V. Juranek, O. Kepka, M. Rangel, C. Royon, and R. Staszewski. Fpmc: a generator for forward physics, 2011.

[83] Peter Skands, Stefano Carrazza, and Juan Rojo. Tuning pythia 8.1: the monash 2013 tune. *The European Physical Journal C*, 74:1–39, 2014.