The Henryk Niewodniczański
Institute of Nuclear Physics
Polish Academy of Sciences

Study of three-body Bose-Einstein correlations in LHCb and development of the reconstruction algorithms based on machine learning

---

Author:

Miłosz Zdybał

First supervisor:

Assoc. Prof. Marcin Kucharczyk,

Second supervisor:

Prof. Piotr Kulczycki

Thesis submitted for the degree of
Doctor of Philosophy in Physics

Kraków, November 2022

Abstract

This thesis describes studies on the Bose-Einstein correlations for the triplets of same-sign pions. The events used in the analysis were recorded by the LHCb experiment in proton-proton collisions at centre-of-mass energy of $\sqrt{s} = 7$ TeV. This is the first study of three-body Bose-Einstein correlations in forward region of pseudorapidity of the LHCb detector and the results are interpreted in the core-halo model for the first time in proton-proton collisions. The parameters of the core-halo model are determined in bins of charged particle multiplicity. A tendency is observed that with growing charged particle multiplicity the signs of coherent particle emission can be observed. The results are presented in the context of other analyses in this field.

The second part of this dissertation is devoted to growing needs for fast and efficient track reconstruction in high energy physics experiments that are addressed with the help of machine learning techniques. A proof-of-concept solution based on the deep neural network is proposed for the MUonE experiment and is promising in the search for New Physics in the sector of anomalous magnetic moment of a muon. The presented implementation meets the conditions of classical reconstruction and at the same time it constitutes a convenient basis for further research.

Streszczenie

Niniejsza rozprawa opisuje przeprowadzone badania korelacji Bosego-Einsteina dla trójek jednoimiennie naładowanych pionów. Wykorzystane przypadki zostały zarejestrowane przez eksperyment LHCb ze zderzeń proton-proton przy energii w układzie środka masy $\sqrt{s} = 7$ TeV. Jest to pierwsza analiza dotycząca trzyciałowych korelacji Bosego-Einsteina w obszarze dużych pseudopośpieszności, charakterystycznego dla detektora LHCb oraz jest to pierwsze zastosowanie modelu core-halo do interpretacji wyników w zderzeniach proton-proton. Parametry modelu core-halo są wyznaczone dla różnych zakresów krotności cząstek naładowanych. Wraz ze wzrostem krotności cząstek naładowanych, obserwowana zależność może wskazywać na częściowo koherentną emisję pionów. Otrzymane wyniki są zaprezentowane w kontekście innych badań dotyczących tego zjawiska.

Druga część rozprawy porusza temat rosnących wymagań na szybką i wydajną rekonstrukcję śladów w eksperymentach fizyki wysokiej energii oraz możliwości zastosowania technik uczenia maszynowego. Zaproponowane oryginalne rozwiązanie wykorzystuje głęboką sieć neuronową do rekonstrukcji śladów w eksperymencie MUonE dedykowanemu poszukiwaniom śladów Nowej Fizyki w sektorze anomalnego momentu magnetycznym muonu. Przedstawiona implementacja spełnia uwarukowania jakościowe klasycznego algorytmu w zakresie rekonstrukcji, jednocześnie stanowiąc dogodną podstawę do dalszych badań.

Contents

Chapter 1

Introduction

The nature of the multiparticle production within the process of hadronization has been investigated for nearly seven decades, but is still not well understood. Quantum interference effects in identical particle systems are a subject of studies for particle physicists since the 1950s, when the first correlations of same-sign pions were observed by Goldhaber, et al. in proton-antiproton annihilation [1]. Since then, many experiments have investigated aspects of the quantum interference and its dependence on various observables, such as charged particle multiplicity, transverse momentum, and hadron mass. All the Large Hadron Collider (LHC) experiments have contributed to the quantum interference studies involving proton-proton [2–6], proton-lead [6, 7] and lead-lead [6, 8–10] collisions.

The correlations are caused by the effects of quantum statistics and the final state interactions of strong and Coulomb origins. The space-time properties of the hadron emission volume can be studied through the parameters of the density function in the region of small four-momenta difference using the quantum interference effect between indistinguishable particles emitted by a finite-size source. The so-called small systems, i.e. proton-proton and proton-ion collisions, are characterized by significantly shorter lifetimes than their heavy-ion counterparts, giving a better experimental insight into the early system dynamics and the initial geometry [11, 12].

In the case of particle physics, the Hanbury Brown-Twiss (HBT) interference effect [13], known from radioastronomy, manifests itself in the Bose-Einstein Correlations (BEC) between identical bosons, or Fermi-Dirac Correlations (FDC) in the case of fermions. The Bose-Einstein correlations are the result of the Bose-Einstein statistics, which allow multiple particles to share the same quantum state and manifest themselves by the increased probability of observing identical bosons emerging from a close region in the phase-space. Those correlations are studied by measuring a correlation function in groups of two or more identical particles and they give an insight into the evolution of the hadron source.

The available accuracy and statistics of data in the operating high-energy physics experiments make it possible to study three-body correlations, which may reveal the nature of the hadronization stage and help to verify theoretical models [14–17]. The main interest in the three-particle correlations is hadron creation mechanisms beyond chaotic (thermal) emission. Study of two- and three-particle correlations can be used to experimentally measure thermalization

and coherence in the source, testing the limits of the core-halo model with the thermalized core.

There have been already some results on the three-particle correlations reported by the ALICE experiment at the LHC [10] with limited interpretation of the results. Sice recently, analyses of three-particle correlations have been using an interpretation within the core-halo model [17]. This model [16, 17] provides a set of parameters describing the expansion of the system after a collision and information about the hadronic source properties not available in the two-body analyses.

The current thesis presents a study of three-particle Bose-Einstein correlations for same-sign charged pions in the proton-proton collisions in the Large Hadron Collider beauty (LHCb) experiment with regard to the core-halo model. The analysis employs proton-proton (pp) collisions recorded at $\sqrt{s} = 7$ TeV [1] centre-of-mass energy. The study is performed in bins of the charged-particle multiplicity.

The LHCb detector [18], thanks to its form of a single arm spectrometer, gives a chance to look at those phenomena in the unique forward region of pseudorapidity of $2 < \eta < 5$. Full instrumentation in the forward direction provides efficient track and vertex reconstruction and excellent particle identification - features universally important in the heavy flavour physics and wide spectrum of other phenomena. The LHCb detector has a potential to help in investigation of interference effects at small angles with respect to the beam direction with the highest collision energies available at the Large Hadron Collider. It provides an exclusive view into the particle production process in the forward direction for the development of the theoretical models. The analysis of the three particle BEC effect in the pp system described in this thesis is already in the publication procedure at LHCb [19]. Furthermore, the experimental aspects of two- and three-particle Bose-Einstein correlations in regard of LHCb data analysis have been discussed in another article by the author of this thesis [20].

In the second part of this dissertation, the topic of growing computational requirements of high energy physics experiments is addressed. New and updated detectors produce the amount of data that is not feasible to be stored with the use of traditional trigger systems. A necessity to efficiently process data coming from the detectors requires offline-quality reconstruction algorithms to be applied online, i.e. before storing data on disks. This pushes for a move to full-software triggers where all events are processed in the software, without hardware-based level of filtering. This poses immense requirements on the performance, leading to the adaptation of new techniques that can accelerate computations, e.g. by the use of graphics processing units (GPUs) that provide a high level of parallelization. One of the groups of techniques that can greatly benefit from the use of GPUs is machine learning, particularly artificial neural networks.

A proof-of-concept event reconstruction algorithm is presented, that uses deep neural network trained on a simulated data sample corresponding to the

[1] If not stated differently, *natural units* are used in the dissertation, i.e. $c = \hbar = 1$. Energy, mass and momentum are given in GeV, while space and time in fm.

test run performed by the MUonE experiment in 2018 [21]. The MUonE experiment [22], planned to operate at the muon beam at the CERN SPS [23], is designed to search for the signs of New Physics by the measurement of the hadronic contribution to the anomalous muon magnetic moment a_μ. This is a promising frontier, as combined results from Fermilab E989/Muon *g-2* [24] and E821 in Brookhaven National Laboratory [25] show a discrepancy of 4.2σ between the Standard Model prediction and the measurement. The MUonE experiment intends to use a novel method related to the measurement of the hadronic part of running of the electromagnetic coupling constant in the space-like region through the scattering of high-energy muons on the atomic eletrons in a low-Z target through the elastic process of $\mu e \to \mu e$.

It is worth adding that the author of this thesis has been actively involved in the development of the track and μ-e elastic event reconstruction algorithms as well as the analysis of the data collected in the MUonE testbeam in 2018, published in [21]. Furthermore, the first results of the developed machine learning procedure dedicated to the pattern recognition and track reconstruction was presented by the author during the Doctoroal Consortium session of the 18th International Conference on Distributed Computing and Artificial Intelligence, 6th-8th October 2021, with the proceedings published in [26].

The thesis is organised as follows. In Chapter 2 an overview of the theory regarding intensity interferometry and quantum-mechanical correlations is given, followed by a description of the model used in this analysis, as well as previous experimental results. The LHCb experiment and its upgrades are described in Chapter 3. Chapter 4 is devoted to the performed analysis of three-particle Bose-Einstein correlations in proton-proton collisions intepreted within the core-halo model. The second part of this thesis begins with Chapter 5 describing the planned MUonE experiment. In Chapter 6 an overview of machine learning techniques is provided, followed by a description of the elaborated deep-neural-network based event reconstruction algorithm for the MUonE experiment with an outlook for future developments. The dissertation is summarized in Chapter 7.

Chapter 2

Theory overview

The HBT studies are one of the measurements that can be used to gain an insight into the process of multiparticle production and put constraints on the phenomenological models. This chapter provides an introduction to the theoretical concepts of the BEC effect, a model used to interpret it, and a description of the implementation of this phenomenon in the high-energy physics simulation frameworks. An overview of the experimental results in the HBT measurements is given in Sec. 2.5.

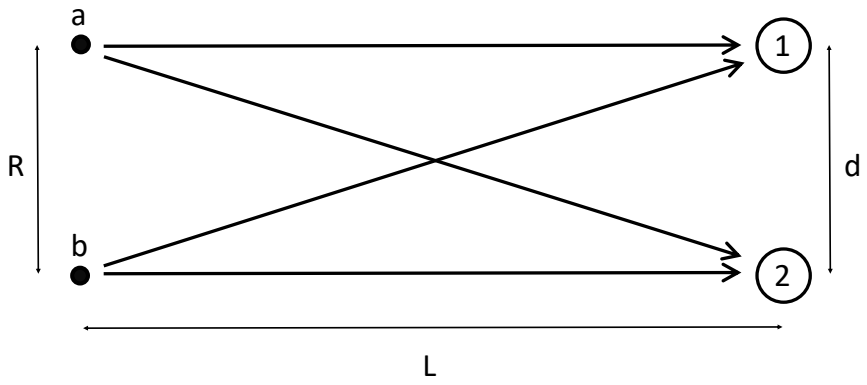
2.1 Intensity interferometry

Intensity interferometry was proposed in the mid-1950s by radio astronomers R. Hanbury Brown and R. Q. Twiss [13] and was later known as the HBT interferometry. This method was used to measure the apparent angular size of the distant radio-source – the Sirius star. A standard technique at the time was Michaelson interferometry which compared the amplitudes of the light at two separated points. In this method the size of the source is determined by analysing the structure of the diffraction pattern in function of distance between the points. The downside of this method was the requirement of a very large distance between the measuring points, going up to thousands of kilometres, to acquire the size of a very small source. A long distance between detectors was not a case in the method proposed by Hanbury Brown.

The measurement of the angular size of the Sirius star was performed using two optical telescopes with photomultiplier tubes serving as detectors, mounted on the trolleys that could be moved on the circular track, thus allowing to change the distance between them. A pair of photons emitted from the two different points of the source was observed in the detectors. The photons were correlated owing to the second order quantum interference effect and the angular size of the source was related to the shape of the interference pattern in a function of the detectors distance.

Methodology employing intensity interferometry can be pictured by two sources a and b separated by distance R and two detectors *1* and *2* separated by distance d with the distance between the sources and the detectors L much larger than R and d (Fig. 2.1). Source a emits a spherical electromagnetic wave of amplitude $\alpha e^{ik|\vec{r}-\vec{r_a}|+i\phi_a}/|\vec{r}-\vec{r_a}|$ and b emits a spherical electromagnetic wave of amplitude $\beta e^{ik|\vec{r}-\vec{r_b}|+i\phi_b}/|\vec{r}-\vec{r_b}|$, with ϕ_a and ϕ_b being random phases. The

FIGURE 2.1: Principle of the interferometry. Light emmited from two random sources *a* and *b* is represented by the arrows. Independent detectors *1* and *2* measure the intensity.



total amplitudes at the detectors *1* and *2* are:

$$A_1 = \frac{1}{L}(\alpha e^{ikr_{1a}+i\phi_a} + \beta e^{ikr_{1b}+i\phi_b}), \tag{2.1}$$

$$A_2 = \frac{1}{L}(\alpha e^{ikr_{2a}+i\phi_a} + \beta e^{ikr_{2b}+i\phi_b}), \tag{2.2}$$

where r_{1a} and r_{2a} are the distances between the source *a* and the detectors *1* and *2*, r_{1b} and r_{2b} are the distances between the source *b* and the detectors *1* and *2*, respectively.

The total intensities in the detectors *1* and *2* are as follows:

$$I_1 = \frac{1}{L^2}(|\alpha|^2 + |\beta|^2 + \alpha^*\beta e^{i[k(r_{b1}-r_{a1})+\phi_b-\phi_a]} + \alpha\beta^* e^{i[k(r_{b1}-r_{a1})+\phi_b-\phi_a]}), \tag{2.3}$$

$$I_2 = \frac{1}{L^2}(|\alpha|^2 + |\beta|^2 + \alpha^*\beta e^{i[k(r_{b2}-r_{a2})+\phi_b-\phi_a]} + \alpha\beta^* e^{i[k(r_{b2}-r_{a2})+\phi_b-\phi_a]}). \tag{2.4}$$

Averaging over random phases the exponential terms aim for zero leading to average intensities in the detectors *1* and *2*

$$\langle I_1\rangle = \langle I_2\rangle = \frac{1}{L^2}(\langle|\alpha|^2\rangle + \langle|\beta|^2\rangle). \tag{2.5}$$

The product of average intensities $\langle I_1\rangle\langle I_2\rangle$ does not depend on the distance between the detectors. The result of multiplying intensities before averaging includes additional term $1 + \cos(...)$ that does not vanish in the averaging over

the phases:

$$\begin{aligned}\langle I_1 I_2\rangle &= \langle I_1\rangle\langle I_2\rangle + \frac{2}{L^4}|\alpha|^2|\beta|^2\cos(k(r_{1a}-r_{2a}-r_{1b}+r_{2b}))\\ &= \frac{1}{L^4}[(|\alpha|^4+|\beta|^4)+2|\alpha|^2|\beta|^2\,(1+\cos(k(r_{1a}-r_{2a}-r_{1b}+r_{2b}))].\end{aligned} \tag{2.6}$$

In this way the correlation function of the intensities can be defined as:

$$C(\vec{d}) = \frac{\langle I_1 I_2\rangle}{\langle I_1\rangle\langle I_2\rangle} = 1 + 2\frac{\langle|\alpha|^2\rangle\langle|\beta|^2\rangle}{(\langle|\alpha|^2\rangle+\langle|\beta|^2\rangle)^2}\cos(k(r_{1a}-r_{2a}-r_{1b}+r_{2b})). \tag{2.7}$$

If the separation between the sources and the detectors is large ($L \gg R$) then $k(r_{1a}-r_{2a}-r_{1b}+r_{2b})$ can be estimated with $k\theta d$, where $\theta = R/L$ is the angular size of the source from the perspective of the detectors and θd is the wavelength of the light.

2.2 Quantum mechanical correlations

In the case of particle physics, the BEC or FDC effects are the result of the quantum statistics describing a given particle system, caused by the symmetrisation (antisymmetrisation) of the wavefunction describing a system of bosons (fermions). Correlations can be used to probe the space-time geometry of the hadronization region, enabling the determination of the size and shape of the particle source. The BEC or FDC correlations are studied in terms of the components of the particles' four-momenta difference. Studies of the non-spherical space-time evolution of the hadronization region are covered by the analyses of the quantum interference phenomena in three dimensions. The decomposition of the particles' four-momenta difference is commonly performed using the Longitudinal Centre-of-Mass System [27, 28]. The LCMS frame makes it possible to separate between the spatial and temporal dimensions. Transversal and longitudinal correlation parameters are acquired using decomposition of the correlation function into three directions in order to determine the quantum interference parameters in the LCMS coordinates. The study presented in this dissertation, however, is based on the assumption of static, spherically symmetric sources that can be characterised by univariate distributions, hence the theory discussion is narrowed down to this case. This assumption follows the conditions undertaken in many HBT analyses as it can easily relate the correlation radius to the effective size of the particle-emitting source.

The source of the single-particle emission can be described in the Wigner-function formalism using the source function $S(x,p)$ [16, 29]. This emission function is a covariant Wigner transform of the source density matrix, where $x = (t,r)$ is the four-vector in time-space, and $p = (E,p)$ is the four-momentum of the emitted particle.

Under the assumption that the source is entirely chaotic, the two-particle

Wigner function can be expressed as the symmetrized product of the one-particle source functions. It can be represented using Fourier-transformed emission function,

$$\tilde{S}(q,P) = \int d^4x S(x,P) e^{iq\cdot x}, \tag{2.8}$$

where

$$q = p_1 - p_2 \tag{2.9}$$

and

$$P = \frac{p_1 + p_2}{2}. \tag{2.10}$$

The singe-particle and two-particle invariant momentum distributions can be expressed as

$$N_1(p) = \int d^4x S(x,p) \tag{2.11}$$

and

$$N_2(p_1,p_2) = \int d^4x_1 d^4x_2 S(x_1,p_1) S(x_2,p_2) \left|\Psi_{p_1,p_2}(x_1,x_2)\right|^2, \tag{2.12}$$

where $\Psi_{p_1,p_2}(x_1,x_2)$ is the wavefunction of the two-particle system. The two-particle correlation function is defined as the ratio of the two-particle momentum distribution to the product of single-particle distributions,

$$C_2(p_1,p_2) = \frac{N_2(p_1,p_2)}{N_1(p_1)N_1(p_2)} = \frac{\int d^4x_1 d^4x_2 S(x_1,p_1) S(x_2,p_2) \left|\Psi_{p_1,p_2}(x_1,x_2)\right|^2}{\int d^4x_1 S(x_1,p_1) \int d^4x_2 S(x_2,p_2)}. \tag{2.13}$$

If we assume that the source is fully chaotic and neglect the final state and long range correlations, the symmetric wavefunction of the particle pair can be described using plane-wave approximation:

$$\Psi_{p_1,p_2}(x_1,x_2) = \frac{1}{\sqrt{2}}[e^{ip_1x_1+ip_2x_2} + e^{ip_1x_2+ip_2x_1}]. \tag{2.14}$$

The correlation function can be expressed using the momentum difference q:

$$C_2(p_1,p_2) = 1 + \frac{\left|\tilde{S}(q,P)\right|^2}{\tilde{S}(0,p_1)\tilde{S}(0,p_2)} \approx 1 + \frac{\left|\tilde{S}(q,P)\right|^2}{\left|\tilde{S}(0,P)\right|^2}, \tag{2.15}$$

where $\tilde{S}(Q,P)$ is the Fourier transformed source function $S(x,P)$

$$\tilde{S}(q,P) = \int d^4x S(x,P) e^{iqx}, \tag{2.16}$$

and P is the average momentum as in Eq. 2.10.

For the simplicity, let us assume the source to be one-dimensional and static.

The source function can be then factorized into spatial and momentum components [30]:

$$S(x,p) = f(x)g(p), \tag{2.17}$$

where x and p are one-dimensional coordinate and momentum variables, with spatial distribution normalized to 1 ($\int dx f(x) = 1$). With the normalization, the spectrum for the single particle is:

$$N_1(p) = \int dx f(x)g(p) = g(p) \tag{2.18}$$

With Eq. 2.12, the correlation function will take the form of:

$$C_2(p_1,p_2) = 1 + \left|\tilde{f}(q)\right|^2, \tag{2.19}$$

where q is the relative momentum as in Eq. 2.9. This means that the correlation function measures the absolute value squared Fourier transformed coordinate-space distribution function (usually called the *characteristic function*) of the particle emitting source, where

$$\tilde{f}(q) = \int dx f(x) e^{iqx}. \tag{2.20}$$

With further assumptions that $\tilde{f}(q)$ is analytic at $q = 0$ and can be expanded into a convergent series, the correlation function can be approximated as

$$C_2(q) \approx 1 + e^{-q^2R^2}, \tag{2.21}$$

where R is the width of the source emission function.

In BEC studies four-momenta difference Q is commonly used, defined as:

$$Q \equiv \sqrt{-(p_1 - p_2)^2} \equiv \sqrt{M^2 - 4\mu^2}, \tag{2.22}$$

where M is the two boson invariant mass, and μ is the boson rest mass.

In a more general manner, the source can be considered to be described by the Lévy stable distribution as its characteristic function. This results in the correlation function in the form of [30]:

$$C_2(Q) = 1 + e^{-|RQ|^\alpha}, \tag{2.23}$$

where α is the *Lévy index of stability* that can take values $(0, 2]$. There are two special cases to be considered: $\alpha = 2$ corresponds to the Gaussian distribution, and $\alpha = 1$ to the Lorentzian (or Cauchy) distribution. Measurements performed by the PHENIX collaboration show the value of α to be approximately 1.2 [31]. For compatibility with previous results, $\alpha = 1$ is used throughout this dissertation.

Parameter λ is referred to as the so-called *intercept parameter*, which corresponds to the extrapolated value of the correlation function at $Q = 0$ GeV [32]. This observable can be interpreted within the core-halo model which assumes that the particle emission can take place in a central core or in an extended halo

originating from long-lived resonance decays (see Sec. 2.4.2). This quantity is also often referred to as the correlation strength because, in reality, the value of the intercept parameter can be modified by various experimental effects. However, despite its sensitivity to the experimental methods, it can carry valuable physics information - not only at the level of (in)coherence of the particle emission in the core, but also on the fraction of the unresolved halo contribution, even in small systems such as proton-proton collisions.

The final form of the correlation function is:

$$C_2(Q) = N(1 + \lambda e^{-|RQ|}), \tag{2.24}$$

with N for normalization.

2.3 BEC modelling

Collision systems are historically divided into heavy-ion and small (particle) systems, with a proton-nucleus in between. The distinction is based on the use of hydrodynamic models to describe the system after a collision. With growing multiplicities of the particle collisions, many aspects related to *Quark-Gluon Plasma* (QGP) [33] formation in nucleus-nucleus and proton-nucleus collisions were successfully applied to the small systems [34], like proton-proton.

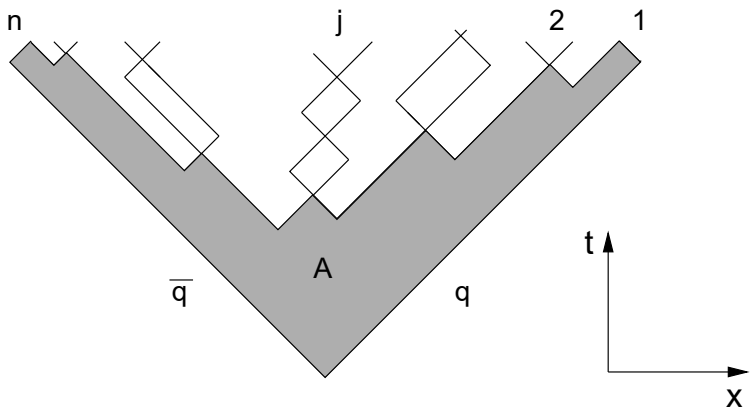
It was shown that similarities between pp and pPb collisions are present, including particle spectra at low p_T [35], particle ratios [36], radial flow [37] and HBT radii trends in like-sign charged-particle multiplicity N_{ch} and average transverse momentum of the pair k_T [2, 38], suggesting the presence of the collective flow [39]. The debate is ongoing whether the hydrodynamic description can be also applied to the small systems, as it is not limited by the theoretical lower bound on the size of a QGP droplet being $\simeq 0.15$ fm [40], compared to the proton size of 0.86 fm.

The use of hydrodynamical models to describe collision systems has a long history [41, 42], but did not attract popularity in mid-20th century. Traditionally, it was viewed that the application of hydrodynamic description requires at or near equilibrium state of the medium. This condition is not fulfilled in the highly inhomogeneous small systems or even larger systems of nuclear collisions [40, 43]. However, hydrodynamics can describe the evolution of the not even locally isotropic matter after the so-called *hydrodynamization time*, when hydrodynamic modes begin to dominate over nonhydrodynamic ones. The size of the system does not seem to matter in that case, as the observations show similar characteristics of hydrodynamic expansion over the range from heavy-ion to proton collisions [40].

2.3.1 Lund model of hadronization

It should be emphasized that the various stages of the system evolution can be described in different ways by appropriate phenomenological models, which is necessary because of the nonperturbative nature of the simulated processes. One of the models describing the process of particle production is the Lund

FIGURE 2.2: Space-time diagram of the string breakup in the Lund Model [44].



Model [44–47], used by a series of particle generators, e.g. PYTHIA [48, 49], used in the present analysis. Within this model the system evolution up to the hadronisation stage, without hydrodynamic expansion, is governed by the fragmentation of the so-called colour strings between the underlying quarks and antiquarks (partons).

Let us consider a pair of quark q and antiquark $\bar{q}$ propagating from one point in opposite directions in the centre-of-mass system of the pair. With increasing distance, the energy of the colour field between the quarks is growing. It is realized in the form of a massless relativistic string with quarks at its ends. The energy accumulated in the colour field is proportional to the string constant $\kappa \simeq 1$ GeV/fm. The string can break into smaller parts, creating new quark-antiquark pairs in the break-up points that become the new endpoints of the string fragments. The newly created quarks and antiquarks from adjacent vertices and the field between them can form a hadron. The process is shown in Fig. 2.2.

The first formulation of the model did not include gluons in the process, leading to the discrepancy when compared with experimental data. In [50] gluons are included as the internal excitations of the string. The probability of forming a set of hadrons from the given set of partons is described by the so-called *Area Law* and can be formulated as:

$$dP_n(\{p_j\}; P_{tot}) = \prod_{j=1}^{n} N_j d^2p_j \delta(p_j^2 - m_j^2) \delta(\sum_{j=1}^{n} p_j - P_{tot}) \exp(-bA) \qquad (2.25)$$

The probability amplitude of the event is proportional to

$$\exp[(i\kappa - b/2)A], \qquad (2.26)$$

with modulus squared of the amplitude corresponding to the probability density of the process proportional to

$$\exp(-bA). \qquad (2.27)$$

The amplitude of probability of producing n particles with given momenta is the sum of amplitudes $\mathcal{M}_\mathcal{P}$ corresponding to each of the $n!$ possible permutation of identical particles in the final state,

$$\mathcal{M} = \sum_\mathcal{P} \mathcal{M}_\mathcal{P}. \tag{2.28}$$

Then the probability of the final state is proportional to

$$|\mathcal{M}|^2 = \sum_\mathcal{P} |\mathcal{M}_\mathcal{P}|^2 [1 + \sum_{\mathcal{P}' \neq \mathcal{P}} \frac{2Re(\mathcal{M}_\mathcal{P} \mathcal{M}^*_{\mathcal{P}'})}{|\mathcal{M}_\mathcal{P}| + |\mathcal{M}_{\mathcal{P}'}|}], \tag{2.29}$$

where the second term can be called weight $w_\mathcal{P}$, so

$$|\mathcal{M}|^2 = \sum_\mathcal{P} |\mathcal{M}_\mathcal{P}|^2 w_\mathcal{P} \tag{2.30}$$

The interference effect in the Lund model is caused by the fact that different string fragmentations can lead to an identical final state. Let us consider two final states where the identical particles *1* and *2* are exchanged in order. The area in the two break-up schemes is different (see Fig. 2.3), hence the corresponding amplitudes will be different. If we mark the two permutations as $\mathcal{P}$ and $\mathcal{P}'$, we can define the area difference as:

$$\Delta A = A_\mathcal{P} - A_{\mathcal{P}'}. \tag{2.31}$$

Then the weight factor can be formulated as [28]:

$$w_\mathcal{P} = 1 + \sum_{\mathcal{P}' \neq \mathcal{P}} \frac{\cos(\frac{\Delta A}{2\kappa})}{\cosh(\frac{b\Delta A}{2} + \frac{\Delta(\sum p_\perp^2)}{2\kappa})}, \tag{2.32}$$

where the denominator of the second argument of the cosh is the total probability of producting a parton with the transverse momentum $p_\perp$.

The area difference ΔA depends on the momentum difference of the particles [27]:

$$\frac{\Delta A}{2\kappa} = \delta p \delta x, \tag{2.33}$$

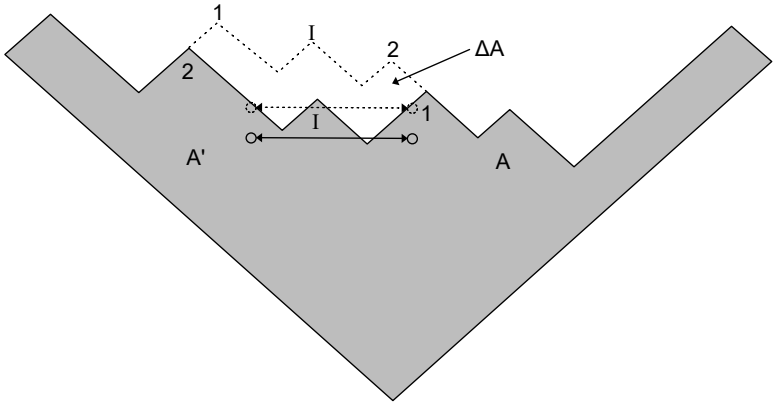
where $\delta p = p_2 - p_1$ is the momentum difference and $\delta x = (\delta t, 0, 0, \delta z)$ is the space-time difference between the production points of two identical bosons.

The contribution to the weight $w_\mathcal{P}$ coming from the configuration $\mathcal{P}'$ diminishes with an increasing area difference. This leads to the conclusion that the interference effect is the strongest for the exchanges of bosons with small differences in the momentum and location.

The correlation function in the Lund Model is defined as the ratio between two-particle probability density with and without the BEC weights:

$$C_2(Q) = \frac{\rho_2^w(q_1, q_2)}{\rho_2(q_1, q_2)} \tag{2.34}$$

FIGURE 2.3: Space-time diagram of two possible fragmentations leading to the same final state. ΔA shows a difference in the area of the break-up schemes.



There is no assumption of random phases and incoherence of the source in this model.

2.3.2 BEC in simulations

The Bose-Einstein effect between identical bosons produced in hadronic interaction cannot be easily included in the event generator programs for high-energy physics, as in such programs only a probability weight for an event with disregard for quantum mechanical effects may be produced. The PYTHIA generator [48, 49], which is used in the present analysis, is quoted as an example of implementing the BEC effect in the high-energy physics simulation. Its implementation is based on the Lund Model (described in Sec. 2.3.1). The BEC effect is introduced into the simulation after the phase of string fragmentation and modelling decays of short-lived resonances. After all particles are generated with their initial momenta, the BEC correction is applied, and the Q values are calculated for each pair of identical particles in the simulated event. It is assumed that the correlation function has Gaussian parametrization. The particles' momenta are shifted to achieve enhancement in the pairs closely separated. At the same time, depletion is simulated in particles with larger separation. Once the shifts for all pairs in an event are calculated, the momentum of each individual particle is modified according to the sum of the shifts determined for this particle. Such an approach does not conserve the total energy in an event, which is corrected by different algorithms in PYTHIA. After finishing this procedure, the simulation continues with modelling decays of long-lived particles.

It is worth noting that no quantum mechanical aspects of the interference effects are modelled, as it is highlighted by the authors in the documentation. Moreover, the HBT correlations are introduced with various assumptions, such as the one on the parametrization of the BEC effect. This makes the use of the simulated events with the BEC effect not suitable for any analysis involving

correlation studies. This is why only events simulated with PYTHIA with the BEC modelling inactivated are used in the present analysis.

2.4 Three-particle correlations in the framework of core-halo model

Bose-Einstein correlations of identical hadrons can be a source of information about hadron creation from the strongly interacting matter formed after a particle or nuclear collision. Measurements of this phenomenon may reveal the hadron creation mechanisms beyond chaotic (thermal) emission, allowing, among others, the limits of the core-halo model with a thermalized core to be tested. An analysis of two- and three-particle correlations can provide an experimental measure of coherence and thermalization in the source.

2.4.1 Core-halo model

The region of a pion emission after a proton-proton collision can be described as a system with a large halo [16]. Based on the characteristics of the hadron emission, the volume can be divided into two distinguishable parts – the centre and the halo. In the central part, the *core*, particles are created in direct processes of hydrodynamic evolution or particle production from excited strings (see Sec. 2.3.1) and following re-scattering. The *halo* around the core consists of pions coming from a decay of long-lived hadronic resonances, i.e. η, η', ω and K^0, where long-lived means a decay length over 20 fm/c. Those time scales are significantly (5-10 times) larger than for re-scattering of pions created in the primordial processes and short-lived resonances (ρ, $N*$, Δ) which are comparable to each other, thus both accounted for as the core production.

2.4.2 Three-body correlations

In the case of three-particle correlations, the correlation function is defined similarly to Eq. 2.13 as the ratio of the three-particle invariant momentum distribution to the one-particle distributions:

$$C_3(p_1, p_2, p_3) = \frac{N_3(p_1, p_2, p_3)}{N_1(p_1)N_1(p_2)N_1(p_3)}, \tag{2.35}$$

where the single-particle momentum distribution is defined as in Eq. 2.11 and the three-particle distribution is defined as:

$$N_3(p_1, p_2, p_3) = \int d^4x_1 d^4x_2 d^4x_3 S(x_1, p_1)S(x_2, p_2)S(x_3, p_3)\left|\Psi_{p_1,p_2,p_3}(x_1, x_2, x_3)\right|^2, \tag{2.36}$$

with Ψ_{p_1,p_2,p_3} being the three-particle wave-function.

Assuming properly symmetrized plane-waves for the wave functions and the Lévy distribution as the source function, the three-particle correlation function can be expressed as a convolution of two-particle correlation functions

for each pair of the particles from the triplet (denoted by subscripts, $Q_{mn} = \sqrt{-(q_m - q_n)^2}; m, n = 1, 2, 3$) and can be formulated as [17]:

$$C_3^{(0)}(Q_{12}, Q_{13}, Q_{23}) = 1 + \ell_3 e^{-0.5(|Q_{12}R|^\alpha + |Q_{13}R|^\alpha + |Q_{23}R|^\alpha)} + \ell_2(e^{-|Q_{12}R|^\alpha} + e^{-|Q_{13}R|^\alpha} + e^{-|Q_{23}R|^\alpha}), \tag{2.37}$$

The parameters ℓ_3 and ℓ_2 are the correlation strengths coming from three- and two-body correlations, respectively. The Lévy index of stability α will be set as 1 (see Sec. 2.2) and the parameter R is the Lévy scale, which corresponds to the radius of the source volume and can be calculated using two-pion correlations.

Additionally, two parameters can be defined – two- and three-particle correlation strenghts, being the extrapolated intercepts for the respective correlation functions:

$$\lambda_2 \equiv C_2(Q_{12} \to 0) - 1 = \ell_2, \tag{2.38}$$

$$\lambda_3 \equiv C_3(Q_{12} = Q_{13} = Q_{23} \to 0) - 1 = \ell_3 + \ell_2. \tag{2.39}$$

The three-particle correlation strength λ_3 and two-particle correlation strength λ_2 carry the information about hadron creation mechanisms.

2.4.3 Parameters of the core-halo model

The parameters of the core-halo system can be used to describe the properties of the hadron emission. The fraction of the core f_c describes the fraction of the particles that originated from the core. The emission function can be formulated as:

$$S(x, p) = f_c S_c(x, p) + (1 - f_c) S_h(x, p), \tag{2.40}$$

where S_c and S_h are the emission functions of the core and halo respectively.

In this model the emission from the core can be partly coherent, contrary to the assumption of a fully chaotic emission. The source function can be written as:

$$S_c(x, p) = S_c^p(x, p) + S_c^i(x, p), \tag{2.41}$$

where the upper index p stands for the part that is coherent and i for incoherent.

The invariant momentum spectrum of the total emission can be expressed as:

$$N(p) = \int d^4x S(x, p) = N_c(p) + N_h(p), \tag{2.42}$$

and core contribution as:

$$N_c(p) = \int d^4x S_c(x, p) = N_c^p(p) + N_c^i(p). \tag{2.43}$$

The partial coherence parameter p_c can be defined:

$$p_c = N_c^p(p)/N_c(p), \tag{2.44}$$

which is the fraction of the core particles that were emitted coherently.

The previously introduced fraction of the core f_c can be similarly expressed using invariant momentum distributions as:

$$f_c = N_c(p)/N(p) \tag{2.45}$$

The values of the correlation strengths (Eq. 2.38, 2.39) can be expressed with f_c and p_c [51]:

$$\lambda_2 = f_c^2[(1+p_c)^2 + 2p_c(1-p_c)], \tag{2.46}$$

$$\lambda_3 = 2f_c^3[(1-p_c)^3 + 3p_c(1-p_c)^2] + 3f_c^2[(1-p_c)^2 + 2p_c(1-p_c)]. \tag{2.47}$$

The value of λ_2 can also be calculated using two-body correlations, as in [52].

As it has been already mentioned, the measurement of λ_2, when combined with λ_3, may test the limits of the core-halo model with a thermalized core. For this purpose a new parameter can be introduced, κ_3, which is a function of λ_3 and λ_2. The parameter κ_3 can be used to determine if there are additional effects in the core, i.e. partial coherence or not fully thermalized core:

$$\kappa_3 = 0.5(\lambda_3 - 3\lambda_2)/\lambda_2^{3/2} = 1. \tag{2.48}$$

In that case the value of κ_3 would deviate from 1.

Experimentally, to be able to investigate the partial coherence within the limits of the core-halo model, the equations 2.46 and 2.47 should be solved for f_c and p_c using the measured values of the λ_2 and λ_3 parameters as input. To determine whether a solution exists, the overlap of the f_c as function of p_c for given λ_2 and λ_3 should be investigated. In other words, the second and the third order Bose-Einstein correlation functions consistency should be checked, and also the degree of partial coherence can be determined from the region of a potential overlap.

2.5 Overview of previous experimental results

Two particle Bose-Einstein and Fermi-Dirac correlations in various collision systems and energies were subjects of studies in many experiments using different accelerators, including (but not limited to): SPS (Super Proton Synchrotron) [53–55], LEP (Large Electron-Positron collider) [56–58], LHC (Large Hadron Collider) [2–10], RHIC (Relativistic Heavy Ion Collider) [31, 59, 60].

Three-particle correlation were recently studied by the PHENIX experiment at the RHIC and ALICE experiment at the LHC.

The first analysis by ALICE [10] regarded two- and three-particle correlations in PbPb collisions at $\sqrt{S_{NN}} = 2.76$ TeV. The study was focused on the search for the coherence using a correlation function based on the *three-particle cumulant* c_3 (three-pion correlations with two-pion correlations removed). The authors concluded, that two-body correlations rule out a spherical Gaussian fully chaotic source. The three-pion correlations suggested a coherent emission

for a low momentum (up to $23\% \pm 8\%$). Suppression of three-pion correlations only for a low momentum was consistent with the formation of the Bose-Einstein condensate.

In the later study, the ALICE experiment sought for a coherent emission in three- and four-particle correlations in a wide variety of the collision systems [6]. The data was divided into bins of the transverse momentum k_T of the triplet or quadruplet. Once again, the cumulants were extracted in the search for a suppression caused by the coherence. The core-halo model was mentioned in the context of correlations between the core and halo emissions, however not for the purpose of the results' interpretation. Three-body correlation functions fitted with an Edgeworth parametrization [61] for different collision systems are presented in Fig. 2.4. The results show suppression of the four-pion Bose-Einstein correlations in PbPb collisions that could be explained with $33\% \pm 3\%(\text{stat}) \pm 9\%(\text{syst})$ coherent fraction. In smaller systems no significant suppression was observed, but the unknown strength of the non-femtoscopic background prevents the authors from drawing the absolute conclusion.

The analysis performed by the PHENIX experiment using data from AuAu colissions at $\sqrt{S_{NN}} = 200$ GeV used a core-halo model for the interpretation of the results [17]. The values of the correlation radius R and Lévy index of stability α acquired in two-body BEC analysis [62] were used. The plots of the calculated correlation functions are presented in Fig. 2.5. The data points are fitted with a function described in Eq. 2.37. The value of the κ_3 parameter is shown in Fig. 2.6 in a function of the transverse mass of the pair of pions taken from the same-sign pion triplet, $m_T = \sqrt{m^2 + p_T^2}$, where m is the rest mass of a single particle in the pair, and k_T is the average pair transverse momentum. It slightly deviates from 1 in the region of $m_T \approx 450$ MeV, suggesting a sign of the coherent behaviour in the pion production. The authors emphasize the preliminary character of the results and hold off from conclusions.

2.5.1 Charged-particle multiplicity dependence

If not stated otherwise, the results described in this and the following sections (2.5.1 - 2.5.3) concern two-particle correlations, as the results for three-particle correlations are limited to the PHENIX and ALICE analyses mentioned above.

Bose-Einstein correlations are often studied in relation to the mean charged-particle multiplicity $\langle N_{ch} \rangle$ or to the mean local charged-particle multiplicity density in pseudorapidity $\langle dN_{ch}/d\eta \rangle$. Many experiments have shown that the correlation radius grows with multiplicity, as shown in the compilation of the results by the ALICE experiment in Fig. 2.7 for two-particle analyses. The collision systems are characterised by slightly different slopes, but the trend is alike in all cases, regardless of the collision system or energy [63–65]. This may suggest the similar evolution of the different systems, which is the case of the heavy-ion collisions associated with the hydrodynamic expansion [66]. The case of proton-proton collisions is not conspicuous in that manner, as there is no agreement on using a hydrodynamic description in small systems and

FIGURE 2.4: Three-body full (C_3^{QS}) and cumulant (c_3^{QS}) correlation functions for triplets of the same-sign pions, fitted with an Edgeworth parametrization [61]. Bottom panels show the ratio of the data to the fit. ALICE [6].

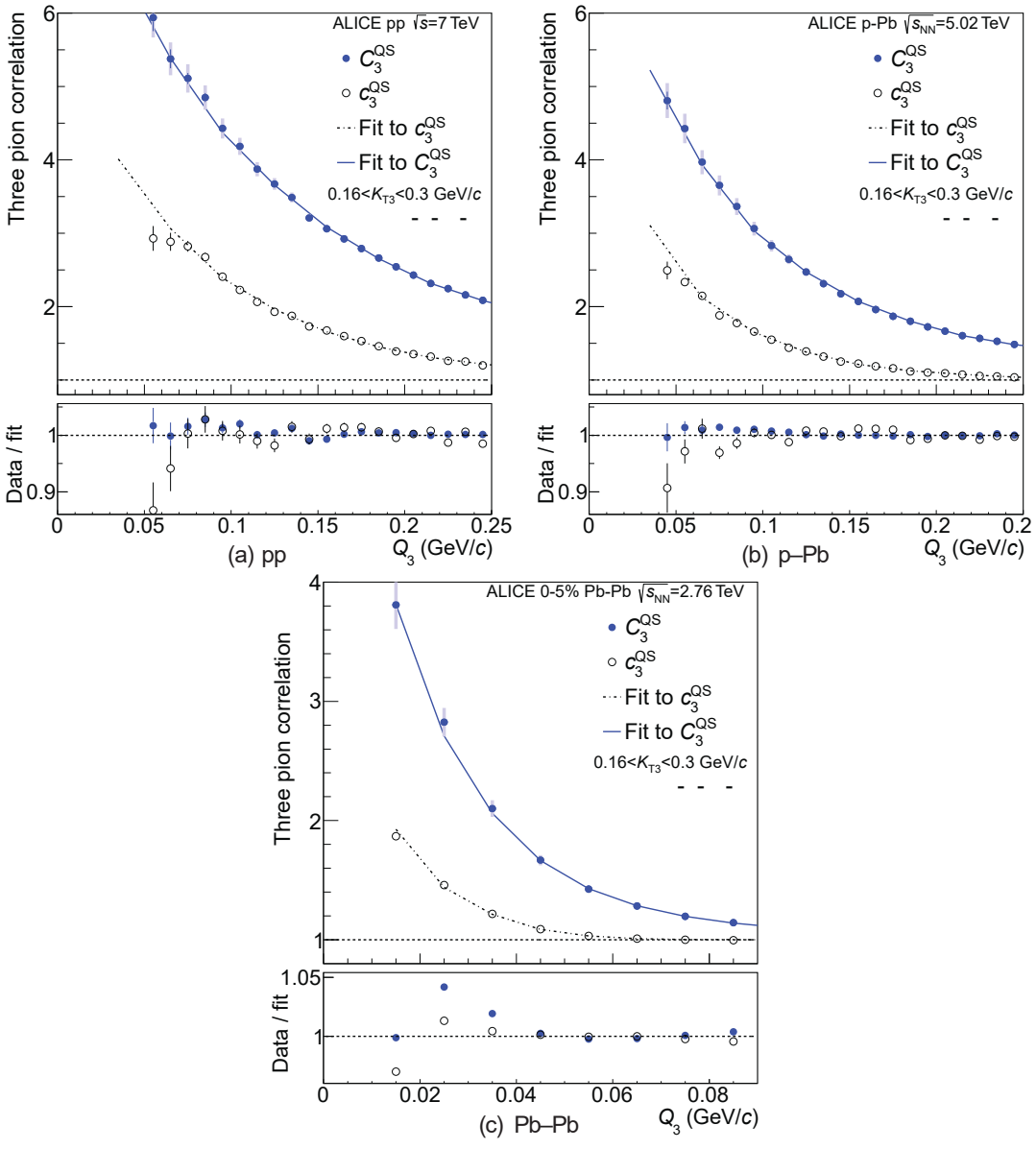


FIGURE 2.5: Three-body correlation function for triplets of the same-sign pions, PHENIX [17].

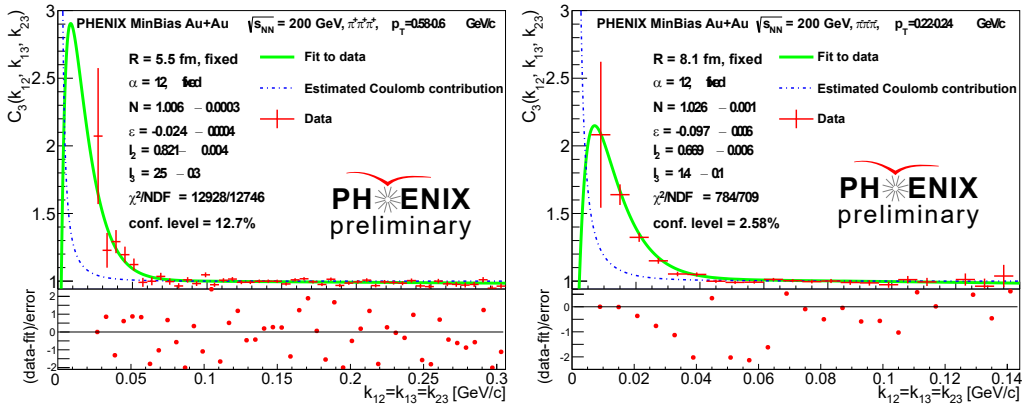


FIGURE 2.6: Value of the κ_3 parameter in the function on m_T, PHENIX [17].

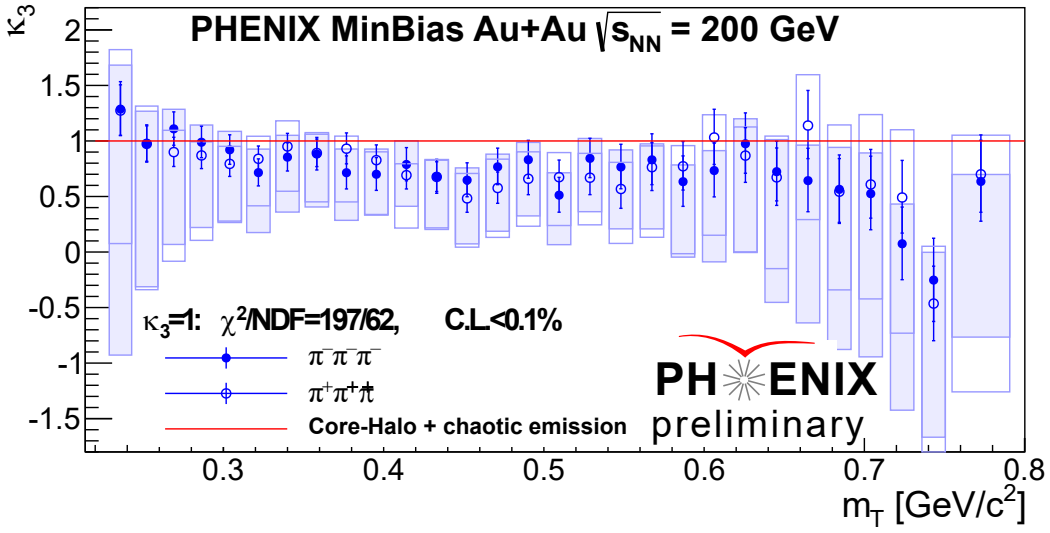
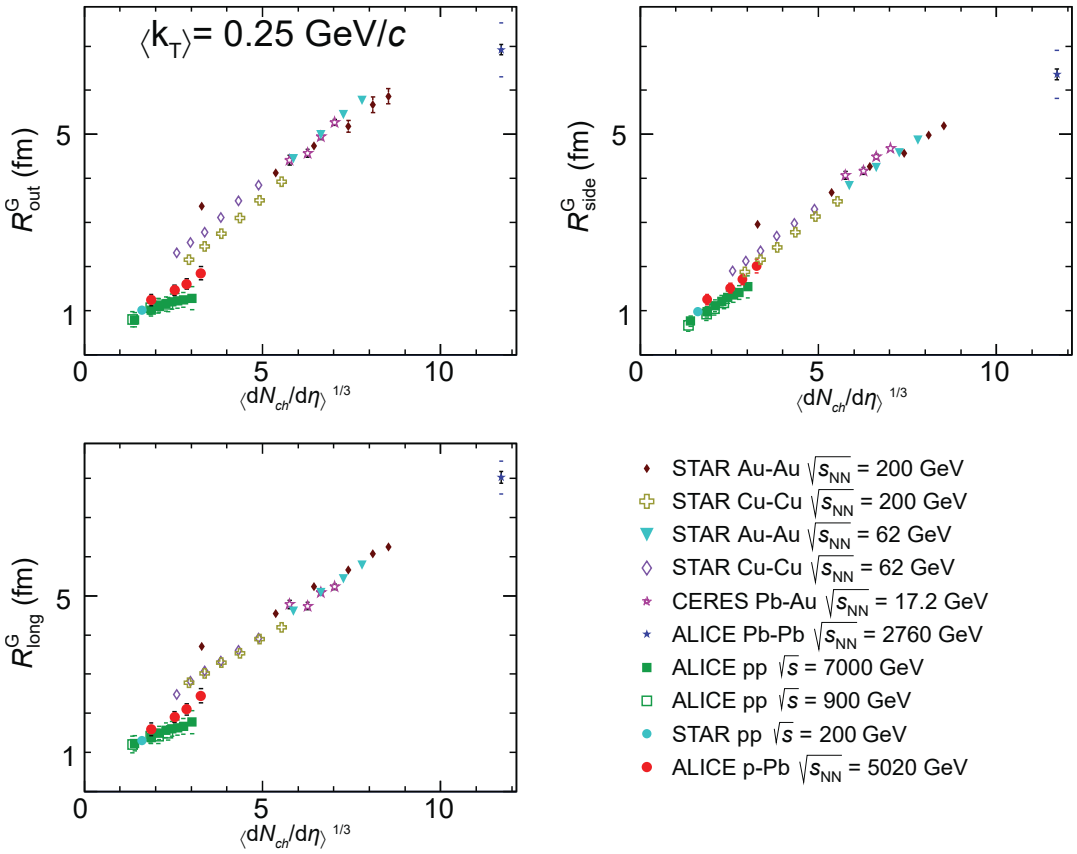


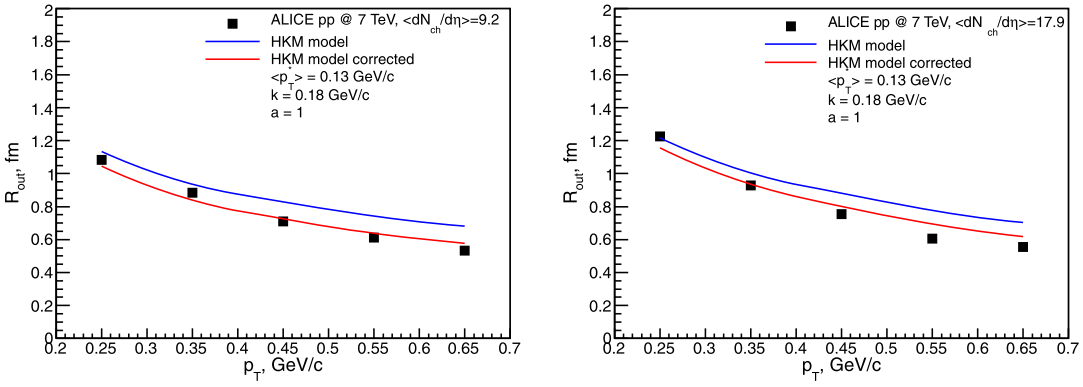
FIGURE 2.7: Dependence of the correlation radius R on the average charged-particle multiplicity density for different collision systems and energies. *Out*, *side* and *long* refers to the directions in the LCMS reference frame [27, 28]. ALICE [65].



the potential saturation of the correlation radius for high-multiplicity events in collision energy of $\sqrt{s} = 7$ TeV was reported by ATLAS [3].

Experimental results are not always easy to compare with theoretical models, as the measured values may not be the HBT radii, but geometrical dimensions of the particle source at a not well-defined stage of evolution. Despite this, an analysis of the qualitative trends can provide useful information about the dynamics of the system. Hydrodynamic [11, 67, 68] and *Color-Glass Condensate* (CGC) [11, 69, 70] frameworks predict the dependence of the correlation radius on the charged-particle multiplicity, also in small systems. The CGC model also predicts that after some threshold in the multiplicity the increase in the correlation radii should stop. As stated before, hydrodynamic representation predict the scaling of the radii with $\langle dN_{ch}/d\eta \rangle^{1/3}$. Both approaches were confronted by the CMS experiment in the analysis of the proton-proton collisions at $\sqrt{s} = 13$ TeV [71].

FIGURE 2.8: Dependence of the measured correlation radii on the transverse momentum for two different multiplicity densities, ALICE [68].



2.5.2 Dependence on the transverse momentum and transverse mass

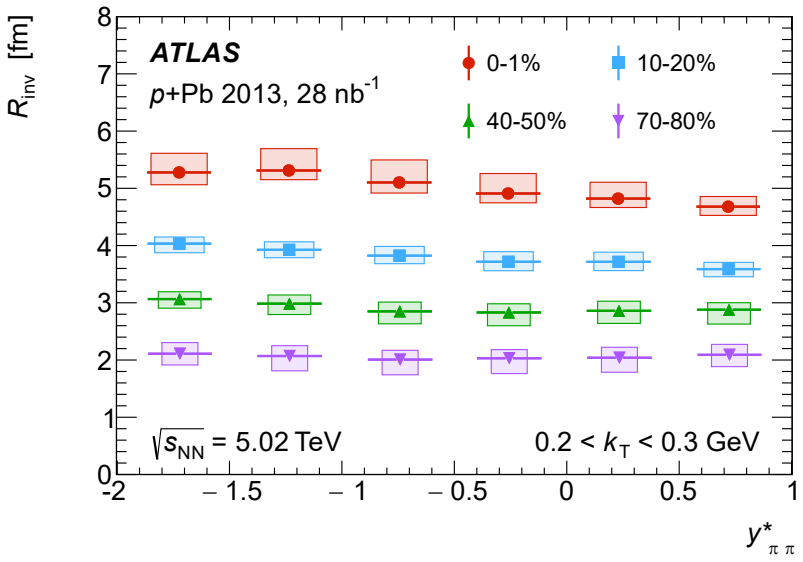
Analyses of the Bose-Einstein correlations are also often performed in dependence on the average pair transverse momentum k_T or transverse mass $m_T = \sqrt{m^2 + k_T^2}$, where m is the rest mass of a single particle. The measured correlation radii decrease for higher transverse momenta which is attributed to the hydrodynamic expansion of the system [68, 72] (see Fig. 2.8). Correlation radii for different pairs of particles in relation to the transverse mass were measured e.g. by the ALICE experiment for PbPb collisions [73].

2.5.3 Dependence on the particle pair rapidity

Measured not as often as previously mentioned is the dependence on the rapidity of the particle pair, calculated relatively to the centre-of-mass frame of the collision system. The ATLAS experiment performed a study of pion pairs produces in pPb collisions at $\sqrt{s_{NN}} = 5.02$ TeV [74]. The measurements were done only in the central rapidity available at the ATLAS detector.

This is why similar measurements for both two- and three-particle correlations in the forward rapidity region, available in the LHCb experiment, may give important information relevant for further development of theoretical models.

FIGURE 2.9: Dependence of the invariant correlation radius R_{inv} on the pair rapidity $y^*_{\pi\pi}$, ATLAS [74]



Chapter 3

LHCb experiment

Although the study of the quantum correlations is not the mainstream interest of the LHCb experiment, available instrumentation is well suited for the BEC analysis, with efficient track and vertex reconstruction and particle identification. The unique region of covered pseudorapidity widens the knowledge accumulated by other experiments.

The LHCb experiment has been collecting proton-proton collision data since 2009. In the first phase of data collecting, called Run 1 (2010-2013), the energy of the proton-proton collisions was 7 and 8 TeV. In the second phase, called Run 2 (2015-2018), the energy in the centre-of-mass of the proton-proton system increased to 13 TeV. Already in the Run 1 phase, most of the data was collected at the luminosity of $4 \times 10^{32}\text{cm}^{-2}\text{s}^{-1}$, exceeding the planned nominal value by a factor of two. In addition to proton-proton collisions, the LHCb experiment participates in a physics program related to ultra-relativistic heavy ion collisions. In the Run 1 phase, it collected data from proton-lead collisions with a centre-of-mass energy per nucleon equal to 5.02 TeV and 8.16 TeV, as well as from lead-lead collisions with a centre-of-mass energy per nucleon equal to 5.02 TeV. Another unique feature of the LHCb experiment is the ability to study hadron-nucleus collisions on a fixed target using the SMOG system. In Run 2 it was used to measure collisions of pHe, pNe and pAr with energies up to $\sqrt{s_{NN}} = 110$ GeV.

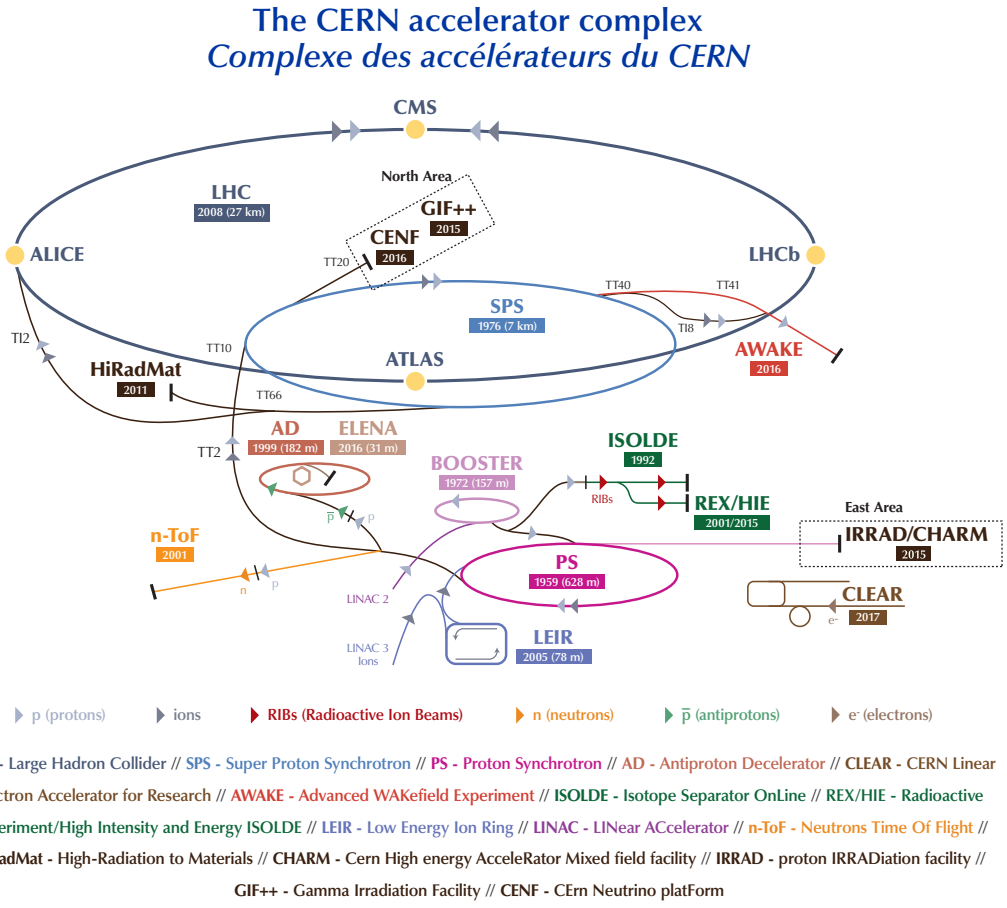
The description of the detector and software stack is related, among others, to Run 1 [75] of the LHC over the years 2009-2013, as the analysis uses data from this datataking period. An overview of the LHCb upgrades is given in Sec. 3.4.

3.1 The Large Hadron Collider

The LHC (Large Hadron Collider) accelerator [76] located at CERN is currently the most powerful machine of this type in the world. This two-ring superconducting accelerator and collider is located in the 27 km circumference tunnel on the French-Swiss border, north-west of Geneva, at a depth ranging from 50 to 175 meters. Various collision systems are available at the LHC, primarily proton-proton (pp), but also including lead nuclei in proton-lead (pPb) and lead-lead ($PbPb$).

The LHC is a part of the CERN accelerator complex [77] (Fig. 3.1), which consists of pre-accelerators and non-LHC experiments. Protons created in the

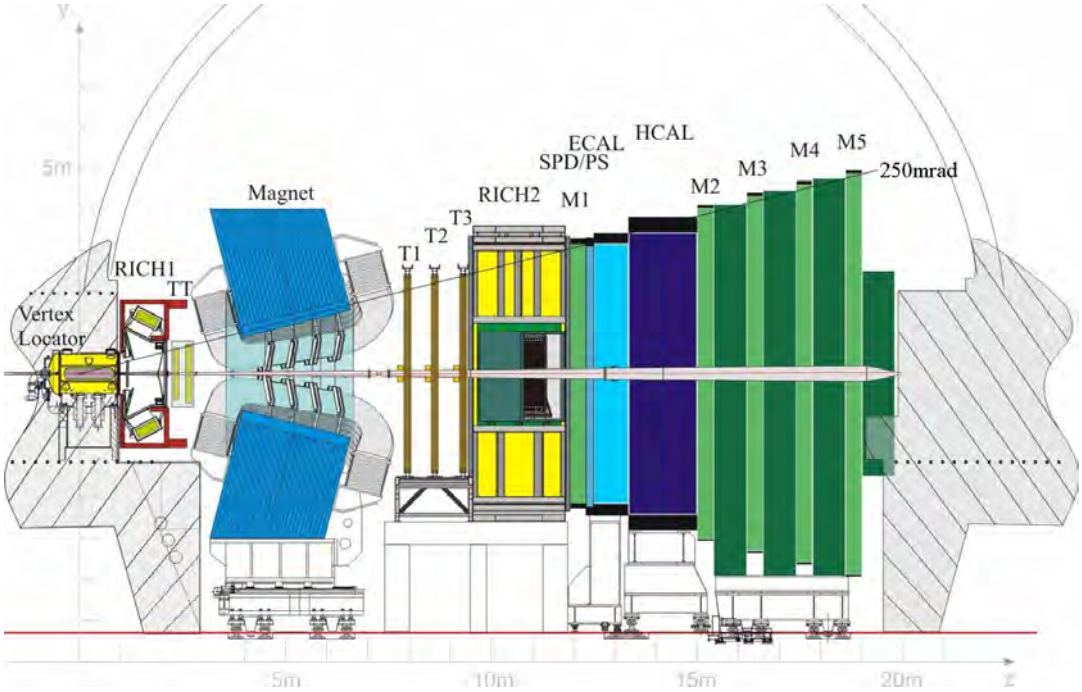
FIGURE 3.1: Schematic view of the CERN accelerator complex [77].



ionization of the hydrogen atoms in the electric field are accelerated by the Linac2 to the energy of 50 MeV and injected into a Booster to achieve 1.4 GeV. The next step is the Proton Synchrotron (PS), accelerating the beam to 25 GeV and pushing it to the last stage before the LHC, the Super Proton Synchrotron (SPS), where the protons gain an energy up to 450 GeV. The particles enter the separate beam pipes going in the opposite directions. The beam is accelerated in a series of 400 MHz radio-frequency (RF) cavities, gaining around 0.5 MeV per turn, while being bent using 8.3 T magnets. The bending magnets are superconducting and operating in the liquid helium environment, in contrast to the room-temperature iron magnets used in the other accelerators at CERN. There are four interaction points around the ring, where the detectors are located.

During *Run 1* with data taking period of 2011-2012, pp collisions where provided by the LHC at the bunch crossing rate of 40 MHz, with centre-of-mass energies of $\sqrt{s} = 7$ TeV in 2011 and $\sqrt{s} = 8$ TeV in 2012, with nominal luminosity of $1 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$.

FIGURE 3.2: Schematic view of the LHCb detector in Run 1, y-z plane [18].



3.2 LHCb spectrometer

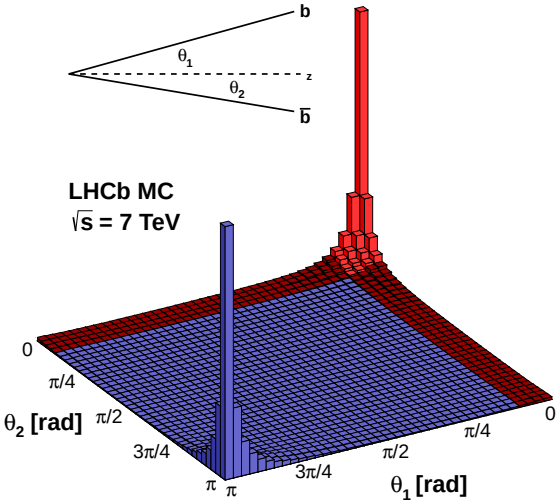
The LHCb experiment [18, 78] was designed to search for evidence of new physics in the *CP* violation through the study of rare decays of b (beauty, bottom) and c (charm) hadrons. This specific aim resulted in a design different from the other LHC experiments: ALICE [79], ATLAS [80], CMS [81].

In inelastic proton-proton collisions, b- and $\bar{b}$-hadrons are preferably produced in a narrow forward (or backward) cone, with small angles with respect to the proton beams. For this reason, the LHCb detector was built in the form of a single-arm forward spectrometer with an angular coverage of 15-300 mrad in the bending (horizontal) plane and 15-250 mrad in the non-bending (vertical) plane.

A schematic view of the LHCb detector is pictured in Fig. 3.2. The right-handed coordinate system is adopted, with the origin in the interaction point, the z-axis pointing downstream in the direction of the beam, and y-axis pointing vertically upwards. The geometry of the detector is adapted to the properties of the dominant $b\bar{b}$ production mechanism, where both hadrons are emitted in the same forward (or backward) direction. The pseudorapidity coverage of $2.0 < \eta < 5.0$ allows capturing about half of the produced $b\bar{b}$ pairs (Fig. 3.3).

The operating luminosity of the LHCb detector was designed at $2{\times}10^{32}\text{cm}^{-2}\text{s}^{-1}$, which is two orders of magnitude lower than the peak luminosity of the LHC. This led to the need for the local control of the luminosity using a luminosity levelling technique [82]. Limiting the number of the visible proton-proton interactions per bunch crossing (*pile-up*) minimizes the risk of the radiation damage to the vertex detectors the combinatorial background of the events.

FIGURE 3.3: Production of $b\bar{b}$ pairs with LHCb acceptance marked in red [83].



3.2.1 Tracking system

A precise tracking system and particle identification are crucial for studying the physics of heavy quarks. An expected fly distance of the b-hadron before decaying is in the range of several millimeters, because of the lifetimes and large boost. This creates secondary vertices (SV) that are displaced with respect to the beam axis.

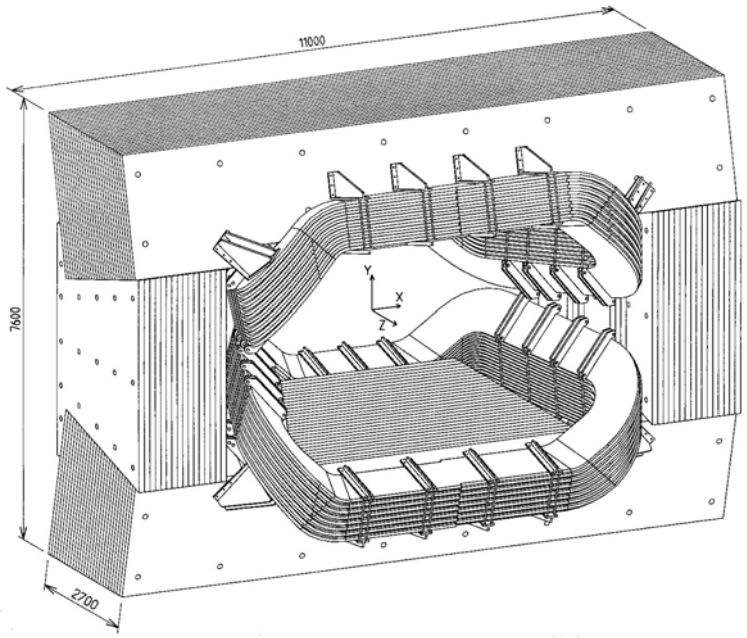
The LHCb tracking system consists of VErtex LOcator (VELO) and four planar tracking stations: Tracker Turicensis (TT) upstream of the dipole magnet and three downstream tracking stations (T1-T3). The downstream stations are further divided into the Inner Tracker (IT) and the Outer Tracker (OT).

Magnet

The magnetic field is introduced to enable measurement of the momentum of charged particles. It is provided by a warm dipole magnet (Fig. 3.4).

The magnet consists of two saddle-shaped coils (27-ton each) mounted symmetrically in an iron window-frame yoke (1450-ton). It covers the acceptance of ±250 mrad vertically and ±300 mrad horizontally. The horizontal plane is called *bending plane*, as the magnetic field is aligned mostly in the y direction, while the vertical plane is called *non-bending plane*. The magnet provides the field that peaks at 1.1 T.

FIGURE 3.4: Dipole magnet of the LHCb, dimensions in mm [84].



VELO

The VErtex LOcator (VELO) is a solid-state detector which provides precise measurement of the location of the primary vertex and secondary vertices of b-hadrons or c-hadrons. The majority of b-hadrons decay in the VELO, with the b-hadron daughter tracks converging to a point displaced from the primary vertex. A precise track reconstruction in this region is crucial to separate vertices. The VELO layout has been optimized to minimize the amount of material in the acceptance region while providing a good geometrical coverage. It is built around the collision region and detects particles using 300 µm thick silicon strip sensors. The VELO is made of a series of 21 stations arranged along the z-axis. Each station is composed of two half-moon shaped modules providing r and ϕ coordinates of the trajectory of charged particles. The r-sensors have concentric semicircles (4×512 strips) centered on the nominal LHC beam position. When closed, the sensors are placed at 8 mm from the beam axis. The spatial resolution on the primary vertex depends on the number of tracks, but its average value is about 42 µm in the z-axis direction and 10 µm in the transverse plane. Such high precision reduces the combinatorial background caused by the tracks assigned to the wrong decay vertex, which is a dominant type of background.

VELO is also used in the measurement of the *impact parameter* (IP, a transverse distance between the closest particle trajectory and the vertex). IP can be used to reject background coming from long-lived heavy flavour hadron decays. The IP resolution is < 35 µm for particles with $p_T > 1$ GeV [85].

FIGURE 3.5: Top: Cross-section of the VELO silicon sensors in the x-z plane at $y = 0$, corresponding the fully closed position. Bottom: front face view of the first modules in closed and open positions. [18].

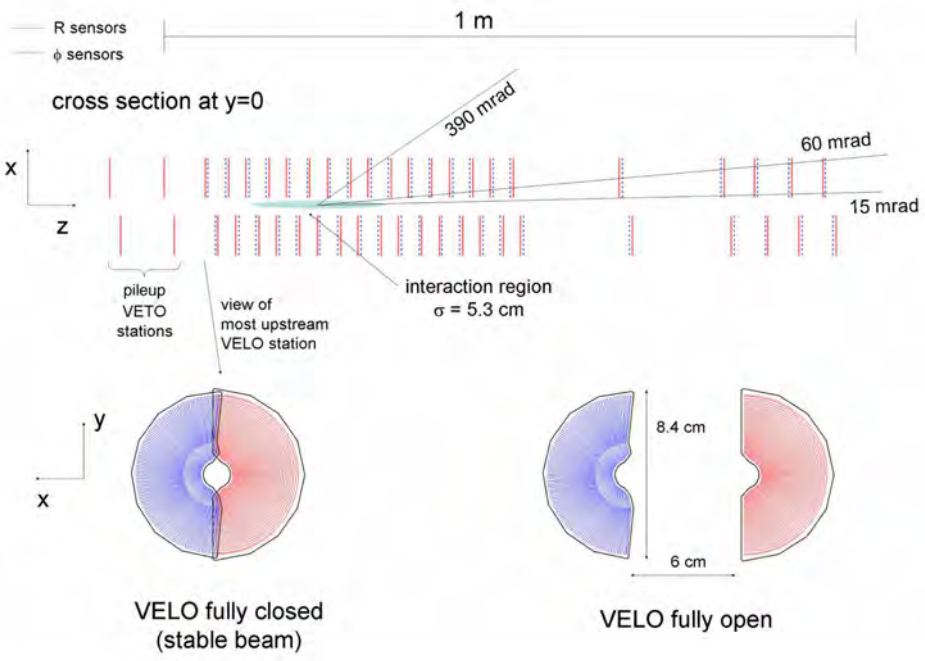
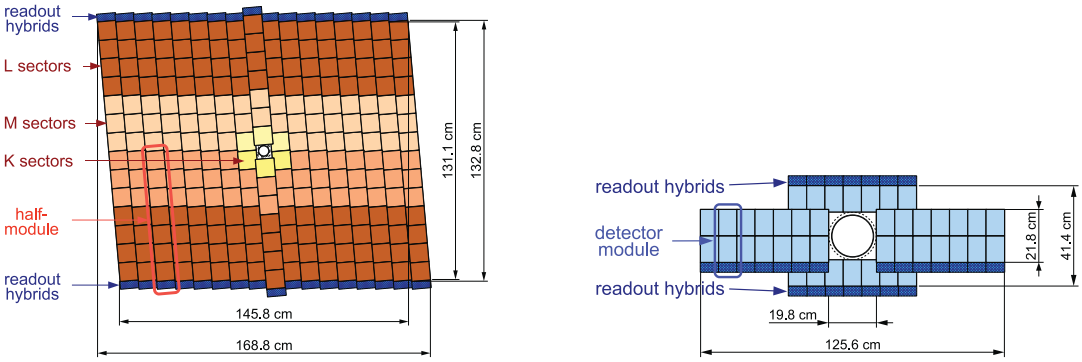


FIGURE 3.6: Left: schematic view of the third TT detection layer. Right: schematic view of an x detection layer in the second IT station [18].



TT, IT and OT

The Silicon Tracker (ST) is a collective name for the silicon microstrips tracking detectors: *Tracker Turicensis* (TT) located upstream of the magnet, and the *Inner Tracker* (IT) located downstream in the station T2. The spatial resolution of the TT is 50 µm and 57 µm for the IT. Each of the trackers consists of four detection layers. For improved spatial information and more precise resolution of the particle momentum, the second and third layers are rotated by an angle of -5° and +5° to the x-axis (the so-called *stereo* configuration), respectively.

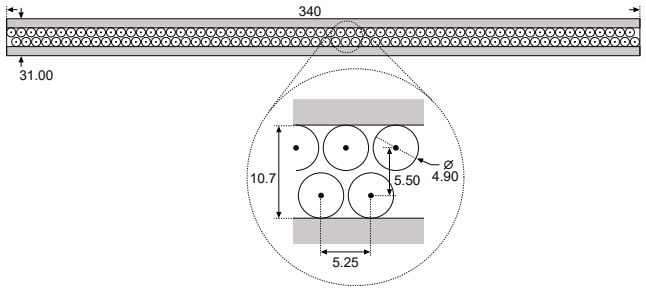
The arrangement of the sensors in the TT is illustrated in Fig. 3.6 (left), using an example of the third (rotated) layer. The sensors provide a full coverage of the nominal LHCb acceptance. The silicon sensors are 500 µm thick, 9.64 cm wide and 9.44 cm long, providing 512 readout strips.

By contrast, the cross-shaped IT covers only 1.3% of the LHCb acceptance; however, 20% of the charged particles are created in the forward region covered by the sensor. The silicon strip technology is the same as used in the TT, with a difference in the thickness. The arrangement of the sensor planes follows the *stereo* configuration used in the TT. The sensors are installed in four boxes mounted around the beam pipe, as shown in Fig. 3.6 (right).

The spatial hit resolution is 62 µm for TT and 58 µm for IT, with hit efficiency (an ability to detect charged particles) 99.3% and 99.7%, respectively.

The Outer Tracker (OT) is a drift-tube gas detector. The gas composition in the straw-tubes (the layout shown in Fig. 3.7) is a mixture of argon (70%), CO_2 (28.5%) and O_2 (1.5%). The detector provides spatial resolution of 200 µm and drift time below 50 ns. As in the previously described tracking stations, the modules are arranged in four layers with the two internal modules being in the *stereo* configuration.

FIGURE 3.7: Layout of the straw-tubes inside the OT module [18].



Track reconstruction

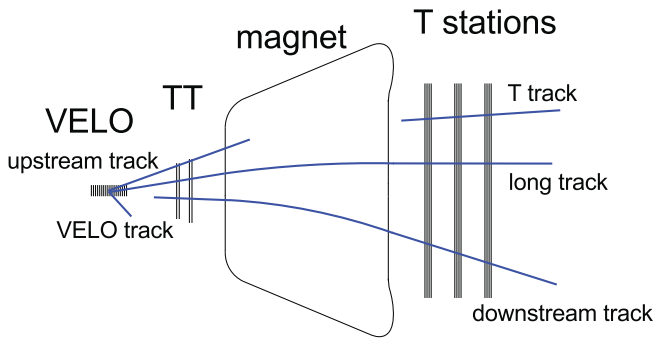
Not every particle is registered by all the tracking stations. By combining information from the tracking stations and the PID system, the trajectories of particles are reconstructed using specialized algorithms. Each track can fall into one of the categories visualized in Fig. 3.8:

- *VELO track* - registered hits are only in the VELO subdetector. All tracks are straight lines, as there is no magnetic field in the VELO. Those tracks are used mostly in the primary vertex reconstruction, but they are also useful e.g. to determine the charged particle multiplicity relevant to the BEC analyses;
- *Upstream track* - reconstructed using hits in VELO and the TT. In most cases associated with particles with a low momentum that are driven out of the detector by the magnet;
- *Downstream track* - reconstructed using information from the TT and the T-stations, but with no hits in the VELO. Tracks used in the reconstruction of long-lived particles that decay after passing the VELO;
- *T track* - registered hits are only in the T1-T3 tracking stations. Particles usually originate from the secondary interactions;
- *Long track* - most accurate and useful tracks, as they passed all the subdetectors.

Additionally, two classes of tracks that were wrongfully reconstructed because of imperfections in the reconstruction algorithms can be defined:

- *ghost* (*fake*) - tracks reconstructed from a set of unrelated hits, not representing any particle trajectory;
- *clone* - multiple tracks reconstructed from the hits of one particle.

FIGURE 3.8: Classes of reconstructed tracks [86].



3.2.2 Particle identification

The task of particle identification depends on the number of subdetectors providing information about different groups of particles, to compose a full picture.

RICH detectors

The first identifing detector on the particles' path is RICH1, the first of the two Ring Imaging CHerenkov (RICH) stations. The principle of this technology is to measure the angle of the cone of the Cherenkov radiation when a particle is going through the medium (the so-called *radiator*). Cherenkov radiation is emitted when a charged particle is moving with a velocity higher than the speed of light in this medium. Radiators with different optical properties are used to cover the full momentum spectrum of the particles. In the case of RICH1, aerogel is used for the particles with a momentum lower than 10 GeV and C_4F_{10} for momenta between 10 and 60 GeV. High momentum particles (15-100 GeV) are handled by RICH2 using CF_4 as a radiator.

The angle of the Cherenkov radiation cone can be used to distinguish between different kinds of hadrons. The dependence on the particle momentum in different radiators is presented in Fig. 3.9.

Both RICH detectors are built in a similar manner: Cherenkov light generated in the radiators is directed to the photon detectors using a set of spherical and flat mirrors. Schematic views of RICH1 and RICH2 are presented in Fig. 3.10. RICH1 is installed right after VELO and upstream the magnet, while RICH2 downstream the magnet and in front of the calorimeter. Acceptance is different for each of the detectors, in RICH1 it is full 25-300 (250) mrad in bending (non-bending) plane, while RICH2 covers a limited angle of 15-120 (100) mrad where most of the high momentum hadrons are found.

Calorimeters

The calorimeter system is composed of four subdetectors identifying photons, electrons and hadrons (including neutral ones, like a π^0). The energies and positions of the particles are measured, providing important input for the

FIGURE 3.9: Cherenkov angle as a function of particle momentum in different RICH radiators [18].

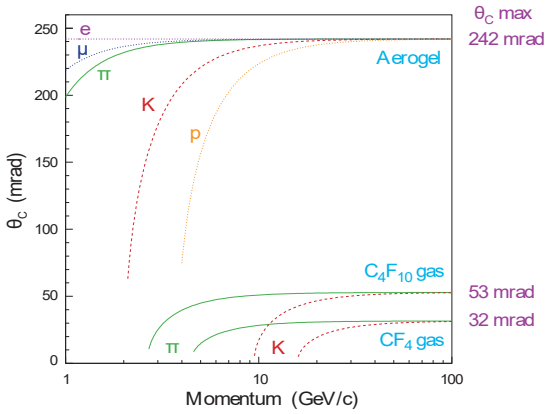
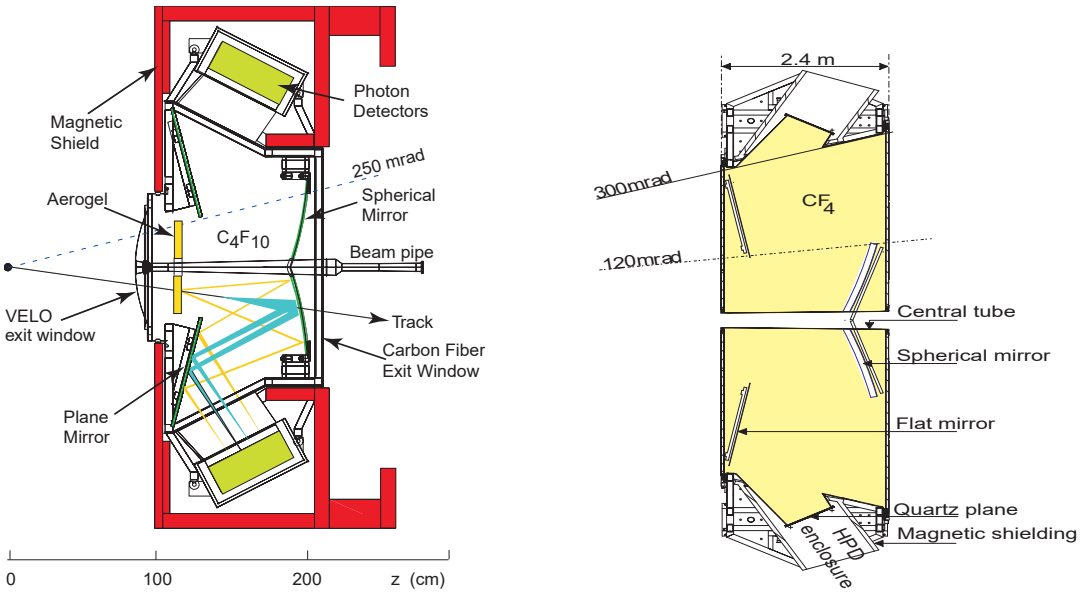


FIGURE 3.10: Schematic view of the RICH1 (left) and the RICH2 (right) [18].



hardware trigger system, especially by indicating hadrons with high transverse energy E_T.

The calorimeters work on the principle that in the absorber a particle loses its energy in a cascade of interactions. It produces electromagnetic and/or hadronic showers of secondary particles that are later measured to reconstruct the energy of the initial particle.

The SPD (scintillator pad) detector is used to select charged particles in the process of rejecting the high E_T π^0 background. The SPD detector is placed in front of the PS (preshower) detector with a thin lead converter in between. The detectors are equipped with high granularity scintillator pads with variable segmentation dependent on the particle flux in each part of the detector.

The ECAL (Electromagnetic CALorimeter) is built using 3300 detector modules of alternating layers of 4 mm scintillating tiles and 2 mm lead sheets. The HCAL (Hadronic CALorimeter) has 3 mm scintillating layers and 10 mm iron absorbers. The layers of the ECAL are perpendicular to the beam, while in the HCAL the layers are parallel. The constituents of a particle shower created in the detector interact with the scintillator material creating UV light that is transported using wavelength shifting (WLS) fibers to the photomultipliers and then to the readout electronics. The relative energy resolution provided by the ECAL is $(\frac{\sigma E}{E}) = 1\% \oplus \frac{10\%}{\sqrt{E}}$ and by HCAL is $(\frac{\sigma E}{E}) = 9\% \oplus \frac{69\%}{\sqrt{E}}$, E is given in GeV and $\oplus$ stands for adding contributions in quadrature.

Muon chambers

Muons are present in many final states of the b-hadron decays, which makes proper identification crucial. The muon detectors are located at the end of the LHCb detector, as muons are characterized by long lifetime and weak interaction with the detector material. One of the main objectives of the muon detectors is to provide the trigger with information about high p_T muons.

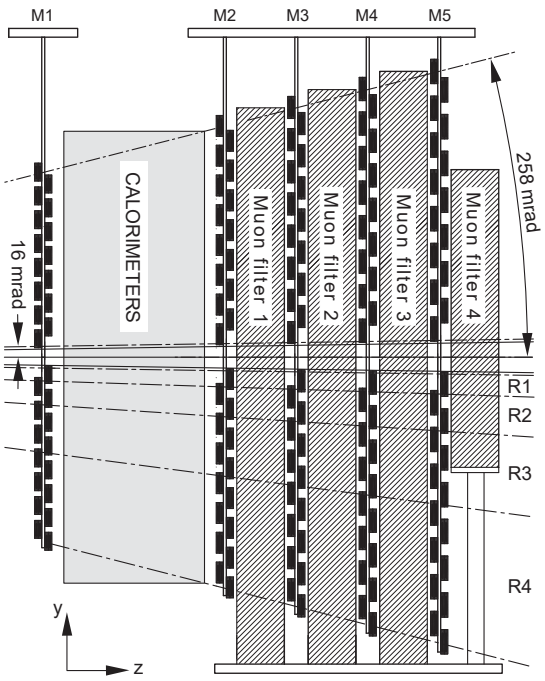
The first muon station (M1) is located in front of the calorimeters and the other four (M2-M5) are located downstream. Stations M2 to M5 are separated with 80 cm thick layers of iron. Each station is larger than the previous one to support coverage of 20-306 (16-258) mrad in bending (non-bending) plane. A side view of the muon system is shown in Fig. 3.11. Stations M1 to M3 have a high spatial resolution in the bending plane to measure particles' direction and p_T, while M4 and M5 have a limited resolution and are used mainly for identification of highly penetrating muons.

3.2.3 Trigger and stripping

Storing of all data coming from the detector is not possible because of the storage limitations. The *trigger* system was developed to save only the events characterized by the sets of required parameters making them potentially valuable for physics analysis. The system has three levels[1]: one hardware *Level-0* (L0) and two software *High Level Triggers* (HLT1 and HLT2). A decision to

[1]This architecture of the trigger was in place for Runs 1 and 2, and changes dramatically in the upgrade for Run 3, see Sec. 3.4.

FIGURE 3.11: Side view of the muon system [18].



save a given event is based on the inputs from the systems described in Sec. 3.2.1 and 3.2.2.

Hardware trigger

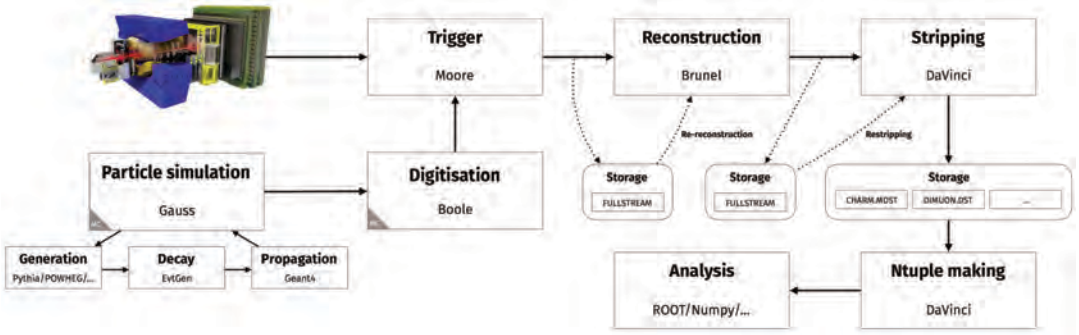
Hardware trigger, built using FPGAs (Field Programmable Gate Arrays), is responsible for the first stage of event filtering and reducing the event rate from 40 MHz of the bunch-crossing in the LHC to 1 MHz, at which the readout of the full detector is feasible. The data from the calorimeters and the muon system is being used to filter events containing electrons, hadrons and muons with high p_T or E_T. The events selected by the Level-0 trigger are stored in the temporary memory for further filtering.

Software trigger

Partial reconstruction is performed int the HLT1. Information from all subdetectors is used to find events with high p_T and high IP, which may suggest decays of long-lived particles containing b- and c-quarks. The HLT1 runs on an online farm in real-time, storing the events to local hard drives. It reduces the event rate to 30 kHz.

Spare resources are used by the HLT2 that writes final events to the offline tape storage. At this stage full reconstruction is performed and both inclusive and exclusive selections are applied. The reconstruction algorithms used in this process are mostly the same as those used in the offline processing with slight simplifications implemented to improve the performance. After the HLT2 the event rate is reduced to around 2 kHz.

FIGURE 3.12: Data flow and software components of the LHCb computing model [87].



Stripping selections

To simplify the selection of events for the analysis, the data are filtered in the process of *stripping*. In this offline process (*i.e.* performed on data stored on tapes), sets of loose preselection criteria are applied, creating *stripping lines*. If there are changes in the stripping lines or new lines are introduced, the process is re-run on the whole dataset (the so-called *re-stripping*). Similar selections are grouped into *stripping streams* that are centrally managed and available to users to avoid reprocessing of the full dataset. Versions of stripping and reconstruction are the main information used to identify the data used in physics analysis.

3.3 Data flow and software stack

A schematic view of the data flow is presented in Fig. 3.12, including the main software components used in the processing of real and simulated data. The software is built in the GAUDI software framework [88] and based on ROOT [89].

The first step in the data flow is the High Level Trigger (see Sec. 3.2.3) that is implemented in the MOORE software package.

The next step is the BRUNEL application that reconstructs the events, creating objects describing tracks and clusters, using triggered raw data. The output is saved in the *DST* format that stores raw data and reconstructed events.

The DAVINCI application is responsible for creating *streams* of files containing events with similar selections. At this stage the files can be stored in DST or, to save space, in *µDST* (micro-DST) format that have raw data of the event removed.

The DST and µDST files can be used for user analysis, but DAVINCI can prepare files in another format - *Ntuple*.

The data flow for the simulated data is virtually the same as for real data. The detector is replaced by the GAUSS application that employs other tools to generate particles (PYTHIA, POWHEG), simulate decays (EVTGEN [90]) and propagate resulting particles in the detector (GEANT4 [91]). The generated events undergo the process of *digitisation* in the BOOLE application, after which they are ready for the HLT and the following steps in the chain.

FIGURE 3.13: Instantaneous and integrated luminosity expectations [92].

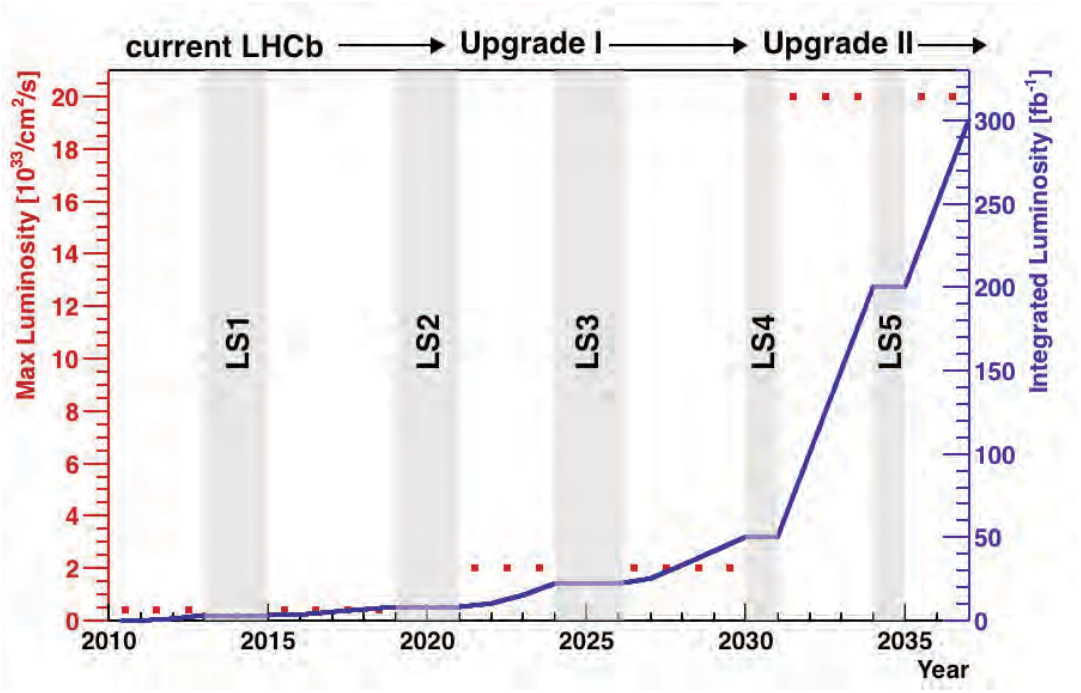
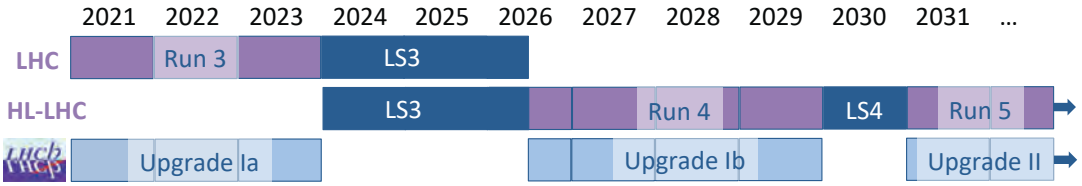


FIGURE 3.14: Upgrade schedule of the LHCb detector for the HL-LHC [92].

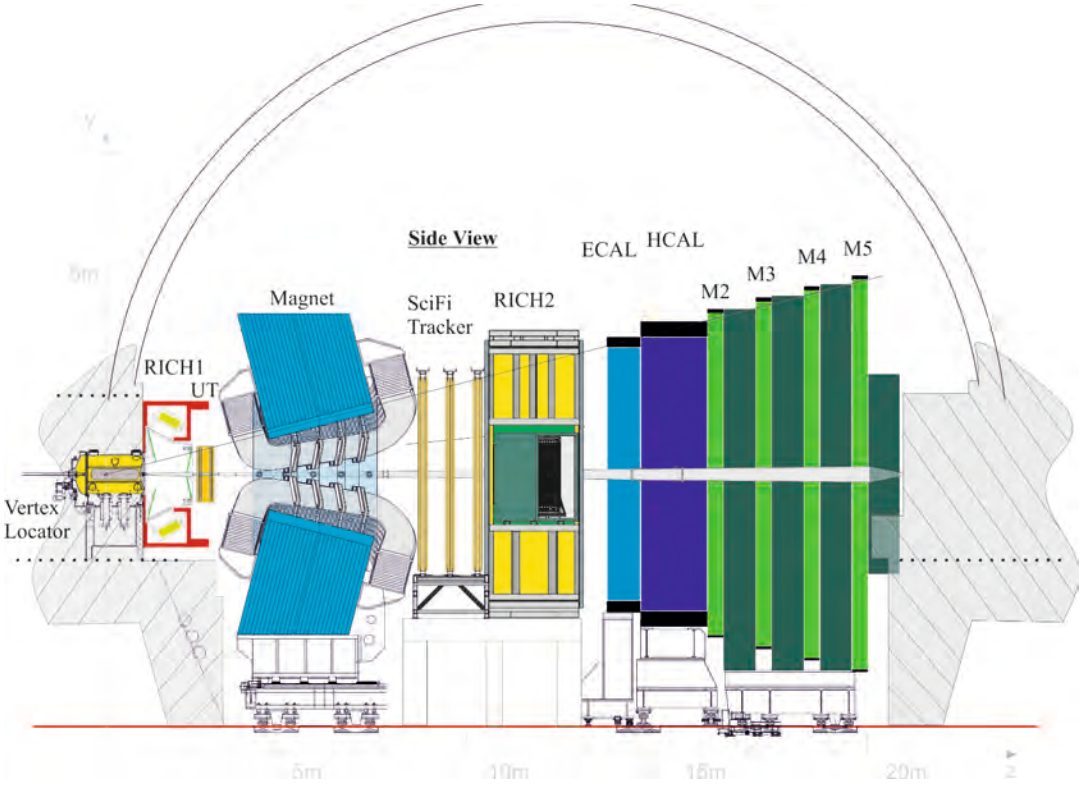


3.4 Detector upgrade

The search for the physics beyond the Standard Model requires an approach called an intensity frontier. The European Strategy for Particle Physics in 2013 acknowledged this requirement with a recommendation to proceed with the programme of upgrading the luminosity of the LHC machine - the *HL-LHC* (High luminosity LHC) project [93, 94]. The instantaneous luminosity (defined as the number of potential collisions per unit area per second) will increase by an order of magnitude, up to $\mathcal{L} = 2 \times 10^{34}$ cm^{-2} s^{-1}. The expected levels of instantaneous and integrated luminosity are presented on Fig. 3.13.

The LHCb detector requires appropriate upgrades to take the full advantage of the change in the beam luminosity. The planned upgrades in relation to the schedule of the LHC are presented in Fig. 3.14, including data-taking periods (*Run 3-5*) and long shutdowns (*LS3, LS4*).

FIGURE 3.15: Schematic view of the LHCb detector after Upgrade Ia [92].



3.4.1 Upgrade Ia

A major challenge in this step is the increased readout of the full detector to the LHC bunch-crossing rate of 40 MHz. The readout will be carried out using a fully software trigger system for the first time [95]. The software trigger will work in two tiers, HLT1 and HLT2, with HLT1 output being also used for real-time alignment and calibration (see Fig. 3.16). The data for the trigger system will be transferred to a computing farm located on the surface using a radiation-protected optical link. Placing all the trigger related computing hardware outside the experiment cavern will allow the commercially available off-the-shelf components (e.g., CPUs, GPUs - graphics processing units) to be used. The components do not need to be radiation-proof, thus lowering the cost and increasing the simplicity.

The components of the detector will see a major overhaul, letting in modernized sensors and readout electronics. More than 90% of the detector channels will be replaced. The updated VELO detector [96] will be able to perform vertex reconstruction with an increased pile-up. The tracking system [97] will receive 10%-20% momentum resolution improvement after the TT detector is replaced with the silicon *Upstream Tracker* (UT), and new *Scintillating Fibre* (SciFi) detector will be installed in the downstream tracking stations (see Fig. 3.15). The upgrade of the PID system [98] will include changes to the photodetectors and electronics in RICH. The muon station M1 will be replaced with a neutron shield in order to protect the SciFi detectors from the particles reflected off the ECAL.

3.4.2 Upgrade Ib

This upgrade will introduce two new subdetectors to the LHCb. The first of them will enhance the tracking system with additional scintillating-fibre detectors on the internal surface of the dipole magnet, dedicated to tracking of the low momentum particles. The second detector will be the *Time Of internally Reflected CHerenkov light* (TORCH) module. By providing time-of-flight information it will enhance the identification of low-momentum protons and kaons to the range below 10 GeV.

The upgrade will be also a chance to replace the modules working close to the beam that are subject to radiation damage. During the phase Ib parts of the SciFi and ECAL will be replaced. The substitution will be made with upgraded elements ready for use in Upgrade II conditions.

3.4.3 Upgrade II

At an expected luminosity of $\mathcal{L} = 2 \times 10^{34}$ cm^{-2} s^{-1}, the mean number of visible proton interactions per bunch crossing will be 56, yielding around 2500 charged particles within the detector acceptance [92]. Precise measurements in this environment pose a great challenge that needs to be addressed by hardware and software solutions. The subdetectors need to be upgraded to increase their granularity and also reduce the amount of the material in the detectors. A crucial aspect of improving particle tracking and identification is increased precision in the timing.

In the VELO detector the data rates will increase by an order of magnitude, with corresponding changes in the occupancy and the radiation levels. A new 4D hybrid pixel detector will be developed that will have enhanced rate and timing capabilities. Changes in the mechanical design will reduce the amount of material between the interaction point and the measuring point, and the sensors will be thinned and prepared for the periodic replacement caused by radiation damage. Changes in downstream tracking will follow the trend of increasing detectors' granularity to adapting them to higher occupancies. The use of timing information and the minimization of the detector material will reduce incorrectly matched track segments between the upstream and downstream tracks. The RICH system will be evolutionally upgraded over Upgrade I changes, with new photomultipliers and optimization of RICH 1 for the lower momentum tracks.

Detailed plans for changes in the subdetectors and R&D directions can be found in [99]. A schematic diagram of the LHCb detector after Upgrade II is presented in Fig. 3.17.

3.5 Real Time Analysis

One of the major changes in the computing model in Upgrade I is adopting a paradigm of the Real Time Analysis (RTA) - offline-quality reconstruction performed online using the fully software trigger [101, 102]. It is important to point out that *real time* is understood here as a period of time between the

FIGURE 3.16: Upgraded online dataflow and software architecture [100].

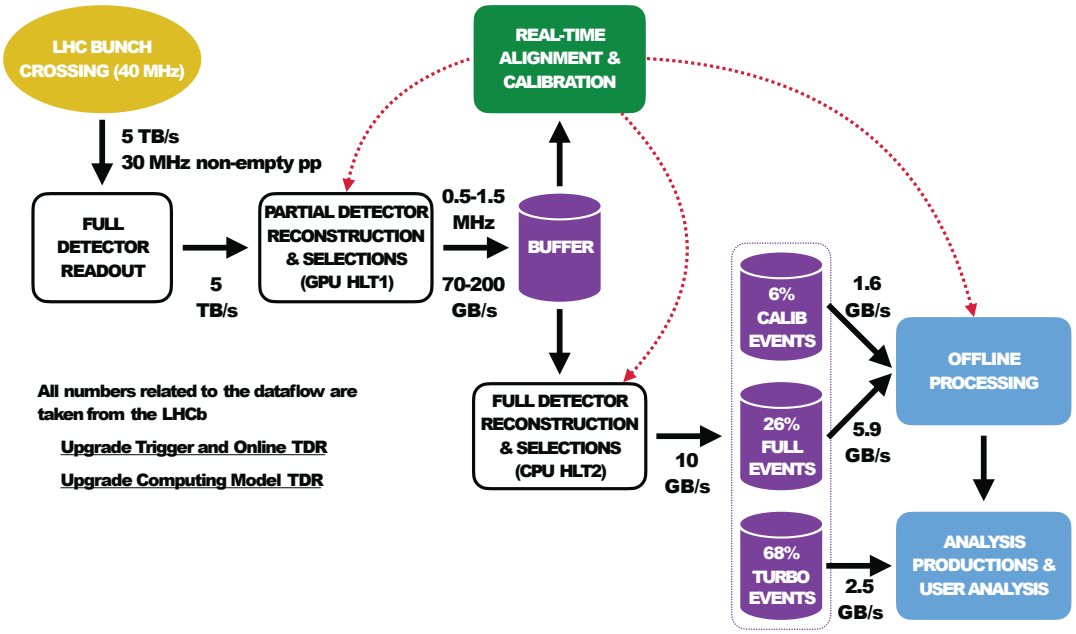
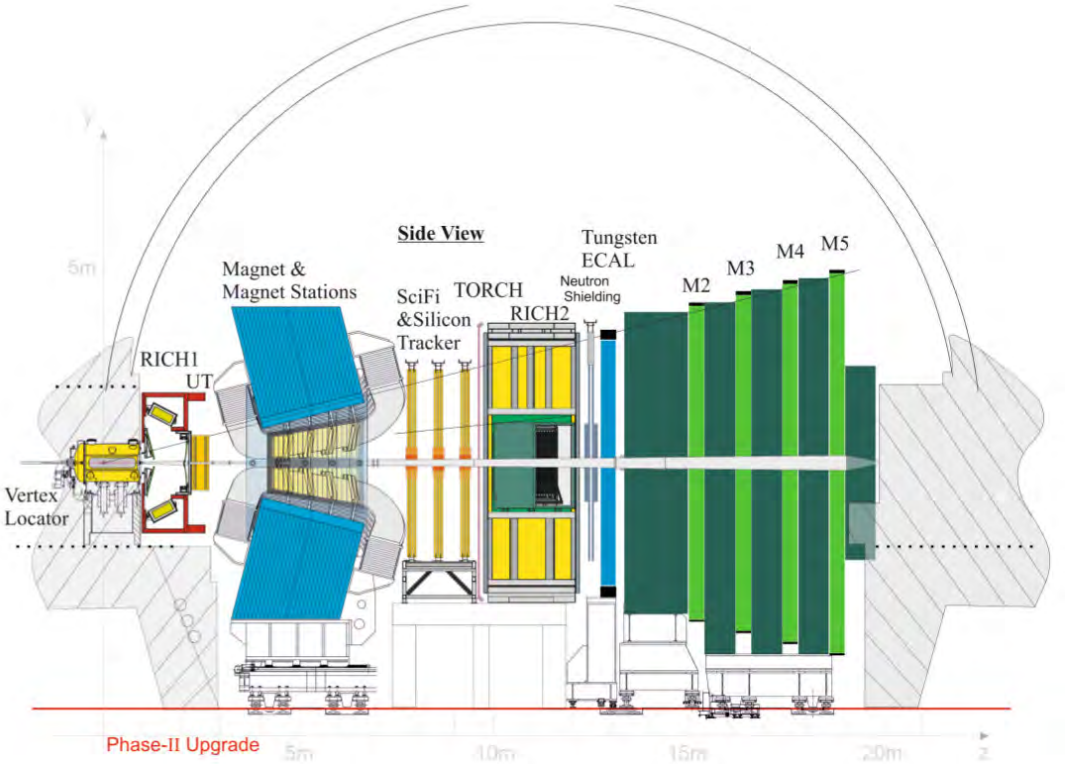


FIGURE 3.17: Schematic view of the LHCb detector after Upgrade II [92].



moment of collision and the decision of discarding the event forever or storing it offline in the permanent storage. In a synchronous trigger configuration, this time is limited to the interval between the collisions. The asynchronous nature of the multi-stage HLT trigger allows the 'real time' period to be prolonged up to two weeks, thanks to the 10 PB disk buffer. The reconstruction algorithms used in the HLT2 are of the same quality that those previously used in the offline reconstruction. In the new solution, the triggered data are ready for physics analyses.

This shift in paradigm puts even more stress on the software which has to be efficient and reliable as never before. To achieve this goal, complex and infallible testing and monitoring procedures need to be developed.

In 2019, the author of this dissertation joined the group responsible for preparing the development environment for monitoring the correct operation of the main data processing applications for the upgraded LHCb detector. This included monitoring of both detector simulation applications as well as data reconstruction and analysis applications. This monitoring was carried out by defining the metrics that allow the quality of registered data to be monitored and by controling the integration between the parts of the system, and meticulousness in the preparation of new releases of the software. One of the main tools in maintaining the interoperability of the components was integration testing. In such tests, many individual subsystems were executed to simulate the real-world workflow. The author of this thesis took part in the effort to develop tests that would verify the dataflow of the Monte-Carlo data samples. The applications were run in a chain (the output of one was the input of the next), starting from the digitisation process of the simulated data in the Boole, running the trigger on the digitised data (Moore), followed by the stripping procedure in the DaVinci. This simulated workflow is very similar to the processing of the real data (with the exception of the digitisation step).

For a period of time the author also participated in the optimization of the alignment algorithms for VELO and SciFi. The objective of such algorithms was to counter imperfections related to the production and installation of the detectors. Small adjustments had be implemented to correct the misalignment in the readings.

Chapter 4

Analysis of three-particle Bose-Einstein correlations in proton-proton collisions

In this chapter, an analysis of three-particle Bose-Einstein correlations for the same-sign charged pions at the LHCb is presented. The methodology used in the analysis is discussed in Sec. 4.1. The data samples used are described in Sec. 4.2, followed by the event selection in Sec. 4.3. The study of the BEC effect in the framework of the core-halo model is presented in Sec. 4.4, with a discussion of the results in Sec. 4.5.

4.1 Methodology

A description of the phenomenon of the quantum mechanical correlations between indistinguishable particles is provided in Sec. 2.2, while the core-halo model is introduced in Sec. 2.4.1. This section will focus on the experimental aspects of measurements of the correlations in proton-proton collisions.

4.1.1 Correlation function

Bose-Einstein correlations are commonly studied in the Lorentz-invariant variable Q, which is a difference of four-momenta of the particles (for definition see Eq. 2.22). Assuming properly symmetrized plane-waves for the wave functions and the Lévy distribution as the source function, the three-particle correlation function can be expressed as a convolution of two-particle correlation functions for each pair of the particles from the triplet (denoted by subscripts, $Q_{mn} = \sqrt{-(q_m - q_n)^2}; m, n = 1, 2, 3$) [17]:

$$\begin{aligned} C_3(Q_{12}, Q_{13}, Q_{23}) &= C_2(Q_{12})C_2(Q_{13})C_2(Q_{23}) \\ &= \frac{N(Q_{12})^{DATA}}{N(Q_{12})^{REF}} \frac{N(Q_{13})^{DATA}}{N(Q_{13})^{REF}} \frac{N(Q_{23})^{DATA}}{N(Q_{23})^{REF}}, \end{aligned} \tag{4.1}$$

where N^{DATA} and N^{REF} correspond to the number of signal and reference pairs, respectively. The four-momentum difference Q gives an invariant measure of the phase-space separation of the two-particle system. To be counted, all three particles must originate from the same primary vertex and have the same-sign

charge. In order to show the influence of the correlations, a reference sample should have all the properties and effects found in the signal, with the exception of BEC.

There are two main ways to create a reference sample for a three-particle correlation. The first way is to use a Monte Carlo sample, where the BEC effect is turned off. However, the main argument against the MC samples is a concern about an agreement between simulated and real data in distributions of crucial variables, such as particle momenta or an opening angle of the pair. In other approaches, a data sample is used, such as in the *event-mixing* method, where the triplet is composed of the same-sign pions, but coming from different events. In practice, reference triplets are created using particles originating from PVs from different events. The primary vertices are required to have the same VELO-track multiplicity to better reflect the properties of the Q distribution for the signal pairs. The main drawback of the event-mixed method is related to the fact that not only the BEC or FDC effect is removed from the data sample, but many other correlations can also be eliminated in this way. This effect, however, is taken into account by constructing the so-called double ratio (see Sec 4.1.2). The event-mixing method is used in the presented analysis. All selection requirements are the same in both samples, and only the events in the same charged particle multiplicity bin are chosen.

It is worth emphasizing, that in the first order, the correlation function is, by its construction, independent of effects related to the detector occupancy and material budget as well as of those related to the single-particle detection efficiency.

4.1.2 Double ratio

As already mentioned, the effects related, among others, to the imperfections of the construction of the reference sample, can be countered using the so-called *double ratio*, a technique commonly used in BEC studies. The ratio of the correlation functions for data and simulation was calculated as follows:

$$r_d \equiv \frac{C_2^{DATA}(Q)}{C_2^{MC}(Q)}. \tag{4.2}$$

The correlation function was constructed for the Monte Carlo sample with the BEC effect turned off.

This method, paired with an event-mixed reference sample, can be used to suppress some long range effects. It applies only to well simulated phenomena, and it is of great importance to maintain the same set of selection parameters for both samples. If applied correctly, the double ratio will result in a purified BEC signal, especially in the region of high Q up to 2 GeV. The introduction of the simulated sample and the choice of the generator and its parameters have an effect on the result that should be included in the study of systematic uncertainties.

4.1.3 Final state interactions and nonfemtoscopic background

There are additional effects present that can affect the correlation function, known as final state interactions (FSI) and related to strong and electromagnetic forces. The effect of the strong interactions in the case of pions is relatively small (see e.g. [103]) and is usually neglected in BEC studies. The most notable is Coulomb repulsion related to the same-sign electric charge of studied particles, affecting the effect of the correlations, especially in a low Q region. For point-like sources, the Coulomb interaction becomes equivalent to the so-called *Gamow penetration factor* [104–106], applying a weight $G_2(Q)$ to every two particle correlation function:

$$C_2^{corrected} = C_2(Q)G_2(Q), \tag{4.3}$$

where:

$$G_2(Q) = \frac{2\pi\zeta}{e^{2\pi\zeta} - 1}; \zeta = \pm\frac{\alpha m}{Q}, \tag{4.4}$$

where m is the particle rest mass, α is the fine-structure constant and the sign of ζ is positive for the same-sign particles, negative otherwise.

Long range correlations, being one of nonfemtoscopic effects related mostly to the energy-momentum conservation, are mostly reduced using the double ratio method and are accounted for with the introduction of the δ parameter with N for normalization:

$$\begin{aligned} &C_3(Q_{12}, Q_{13}, Q_{23}) = \\ &N(1 + \delta Q_{12})(1+\delta Q_{13})(1 + \delta Q_{23})G_3(Q_{12}, Q_{13}, Q_{23})C_3^{(0)}(Q_{12}, Q_{13}, Q_{23}), \end{aligned} \tag{4.5}$$

where $C_3^{(0)}$ is the correlation function that describes the pure BEC effect and G_3 stands for the correction of the Coulomb effect. According to the Riverside method [104], the Coulomb correction for the particle triplet can be approximated by a factorization of the corrections calculated for each of the pairs in the triplet:

$$G_3(Q_{12}, Q_{13}, Q_{23}) \approx G_2(Q_{12})G_2(Q_{13})G_2(Q_{23}). \tag{4.6}$$

It is important to note that the Coulomb effect is not present in the simulated data, hence the correction is applied only to the real data.

Cluster contribution (name proposed by CMS in [107]) is another prominent factor in the nonfemtoscopic background, related to the effects of minijet fragmentations and multibody resonance decays. It is harder to correct than long range correlations as it is present dominantly in the range of $Q < 0.5 - 1.0$ GeV that overlaps with the BEC signal. However, this contribution is not expected to be prominent in proton-proton collisions as it is in proton-ion and ion-ion. In proton-proton collisions, cluster contribution is at a negligible level, as the double-ratio distributions for opposite-sign pion pairs are virtually flat and close to unity, with the exception of the remaining structures related to resonance decays, which are not well represented in the simulation (see Sec. 4.4.1). This fact suggests that nonfemtoscopic background is properly modelled in the simulation and corrected for when the double-ratio is used. Potential discrepancies related

to the Monte Carlo generator used to simulate the same-sign and opposite-sign pairs were addressed in the systematic uncertainty studies.

4.1.4 Fitting of the three particle correlation function

As described in Sec. 2.4, the BEC effect for three-particle correlations in the framework of the core-halo model can be parametrised as (see Eq. 2.37):

$$C_3^{(0)}(Q_{12}, Q_{13}, Q_{23}) = 1 + \ell_3 e^{-0.5(|Q_{12}R|^\alpha + |Q_{13}R|^\alpha + |Q_{23}R|^\alpha)} + \ell_2(e^{-|Q_{12}R|^\alpha} + e^{-|Q_{13}R|^\alpha} + e^{-|Q_{23}R|^\alpha}).$$

As the long-range correlation and Coulomb effects must be also included (see Sec. 4.1.3), a full parametrisation for the experimentally measurable three-particle correlation function is given as:

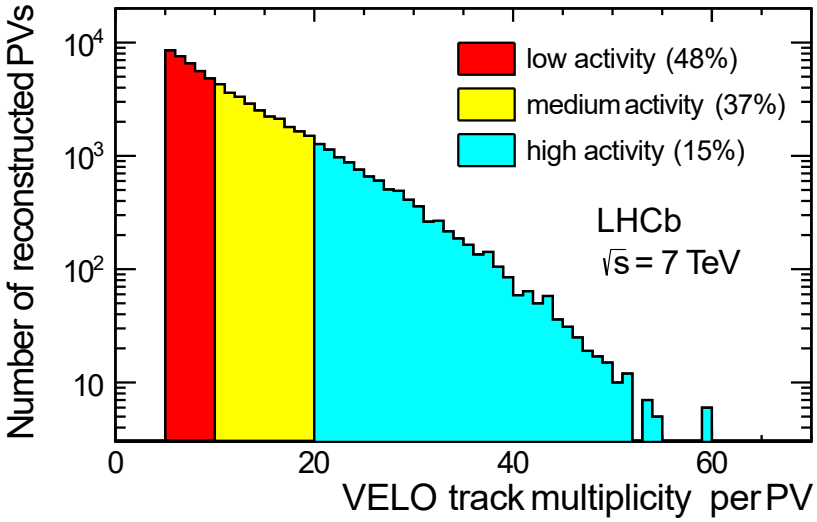
$$C_3(Q_{12}, Q_{13}, Q_{23}) = N(1 + \delta Q_{12})(1 + \delta Q_{13})(1 + \delta Q_{23}) G_2(Q_{12}) G_2(Q_{13}) G_2(Q_{23}) \\ (1 + \ell_3 e^{-0.5(|Q_{12}R|^\alpha + |Q_{13}R|^\alpha + |Q_{23}R|^\alpha)} + \ell_2(e^{-|Q_{12}R|^\alpha} + e^{-|Q_{13}R|^\alpha} + e^{-|Q_{23}R|^\alpha})), \tag{4.7}$$

where ℓ_2 and ℓ_3 are fit parameters related to the two-particle correlation strength λ_2 and three-particle correlation strength λ_3 (see Eq. 2.38 and Eq. 2.39), giving information about the hadron creation mechanism.

The parameters of the fit can be used to calculate the values of the parameters of the core-halo model (see Sec. 2.4.3). The fraction of the core f_c describes the fraction of particles that originated from the core of the emission volume. The partial coherence p_c stands for the fraction of the particles that were emitted coherently. Both parameters are linked to the correlation strengths in the way described in Eq. 2.46 and 2.47. The third parameter of this model, κ_3, indicating whether there is partial coherence in the core or not fully thermalized core, can also be expressed using correlation strengths, as shown in Eq. 2.48. The correlation radius R as well as the λ_2 parameter are known from the two-particle BEC analysis.

In the parametrisation shown in Eq. 4.7, Lévy-type source is assumed. It is commonly used in BEC analyses as it provides the best description in terms of the quality of the fit. The Lévy index of stability can take values in the range of $0 < \alpha < 2$. For two specific values of the α parameter of unity and two, the exponential and Gaussian parametrisations are reproduced, respectively. The general type of parametrisation usually provides better quality of the fits to the data than the simple cases. It was found in a study employing the Lévy parametrisation [31] that indeed the measured index of stability value is consistent neither with the Gaussian nor the exponential scenario, with the results at the level of $\alpha \sim 1.2$. However, if α is set free in the fit, the interpretation of the measured correlation radius is not straightforward. Usually, in analyses of the BEC effect performed by high-energy physics experiments, α is fixed to 1 in order to compare the results with other experiments and also to get the possibility to directly interpret the results. This is why the value of the Lévy index of stability is set to 1 in the present analysis.

FIGURE 4.1: Distribution of the VELO-track multiplicity per PV in the 2011 non-bias proton-proton sample [52].



4.2 Data and simulation samples

The data used in this analysis were collected by the LHCb experiment during the Run 1 datataking period in 2011, at a centre-of-mass energy of $\sqrt{s} = 7$ TeV. The recorded sample of pp collisions corresponds to an integrated luminosity of 1.0 fb^{-1}. The events were triggered using a minimum-bias trigger that requires at least one reconstructed VELO track. No-bias events were created in the unbiasing procedure that randomly removes one primary vertex and associated tracks from a minimum-bias event, providing a set of unbiased vertices. The average number of visible interactions[1] per bunch crossing is 1.4 [108]. The dataset contained 4×10^7 events.

For event generation, PYTHIA 8 [48, 49] was used with designated LHCb configuration [109]. Hadronic decays were simulated in EVTGEN [90], while final-state radiation - in PHOTOS [110]. An interaction of the generated particles with the detector material was simulated in GEANT4 [91]. The simulated sample contains 2×10^7 minimum-bias events. No special procedure to create a no-bias sample was used, as all the simulated events are unbiased.

Another sample generated with PYTHIA 6.4 [111] with PERUGIA0 tuning [112] was used in the systematic uncertainties study (Sec. 4.4.3). It contained 1×10^7 events.

The parameters of Bose-Einstein correlations in this analysis were calculated in the bins of charged-particle multiplicity. As the construction of the LHCb detector is unique among other experiments, the results in particle multiplicity bins cannot be directly compared with the results from other experiments. This is why *activity classes* were introduced, better reflecting the total multiplicity in the full solid angle. The classes were based on the VELO track multiplicity of the reconstructed primary vertices, which is a good analogue of the total

[1] A visible interaction is a primary vertex with at least five reconstructed VELO tracks.

multiplicity. The low activity class represented PVs with 5 to 10 reconstructed tracks (48% of all reconstructed PVs), the medium activity with 11 to 20 tracks (37%) and the high activity with more than 21 tracks (15%) (see Fig. 4.1). For each activity class, the VELO-track multiplicity N_{VELO} was reconstructed to the true charged-particle multiplicity N_{ch} in the process of *unfolding*. This method associated reconstructed multiplicities with those predicted by PYTHIA [109]. The values obtained using the Bayesian unfolding technique [113] were applied onto the data. The presented activity classes corresponded to charged-particle multiplicities $N_{ch} \in [8, 18]$ for low, $N_{ch} \in [19, 35]$ for medium, and $N_{ch} \in [36, 96]$ for high activity. These values are in agreement with the previous LHCb measurements [114].

4.3 Selection

A set of requirements was imposed on the data to ensure the quality of tracks used in the analysis. The requirements wre designed to properly select triplets of pions that may exhibit the Bose-Einstein correlations.

As the final results for the three-particle correlation measurement depend on the λ_2 and R parameters measured first for the two-particle correlations, an important issue in the present analysis was to determine such parameters for the pairs from particle triplets and check their compatibility with the previous LHCb results for two-particle Bose-Einstein correlations [52]. In this way the analysis procedure and event selection for two-body BEC analysis applied in the present analysis should be consistent with those applied in [52].

4.3.1 Preselection criteria

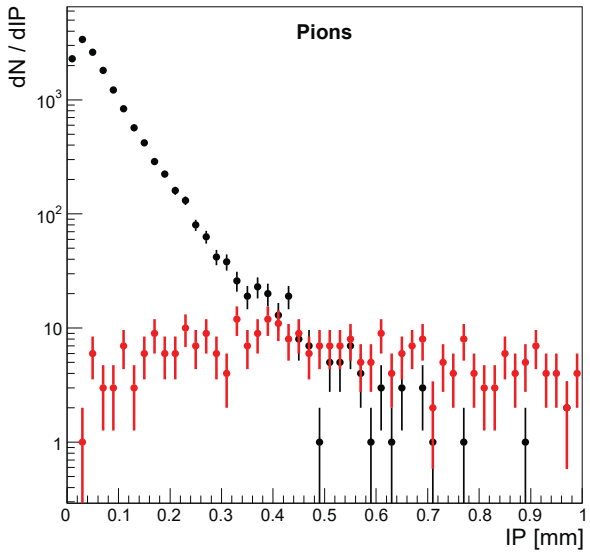
The BEC effect was analysed for particles with tracks meeting a set of conditions. The tracks had to be within the LHCb detector acceptance of $2.0 < \eta < 5.0$, tracked in the VELO and T stations downstream of the magnet (*long tracks*, see Sec. 3.2.1). At least five reconstructed tracks were required for each primary vertex. Muon contamination was prevented by the identification in the muon stations.

The preselection required at least one triplet of loosely identified pions, kaons or protons originating from the same primary vertex, where particles were assigned to PVs based on the minimum value of χ^2 and IP. The parameters and associated requirements are summarized in Table 4.1.

Multiple variables were taken into account in the process of selecting events for the analysis. Some of them were kinematic properties of a reconstructed particle, like the momentum p, transverse momentum p_T or impact parameter IP. The quality of the fit of a reconstructed track was related to its χ^2. Additional variables *ProbNN* related to the particle identification were calculated using the input from the multiple detectors and with the help of the artificial neural network (see Sec. 4.3.2).

To ensure the high quality of the reconstructed tracks used in the analysis, the requirements were imposed on the track p_T, track χ^2 and the probability of the track to be a ghost (*ProbNN(ghost)*). In order to keep as many low-p tracks

FIGURE 4.2: Impact parameter (IP) distribution for true pions from MC sample in logarithmic scale. Marked in black are true pions that were correctly assigned to the PV corresponding MC PV, while in red are marked missasigned PVs.



as possible, the requirement on the impact parameter had to be optimized, being related to the quality of the assignment of the track to the proper PV. In order to estimate the optimal value of the limit on IP, the distribution for MC true pion recognized to originate from the reconstructed PV corresponding to MC PV of their origin was compared with the distribution of true pions misassigned to the reconstructed PV corresponding to the MC PV they were originating from (see Fig. 4.2). The selection on IP was optimized to reject the region where the abundance of misassigned tracks became significant with respect to those properly assigned. This is why a loose limit on IP was chosen, i.e. IP < 0.4 mm.

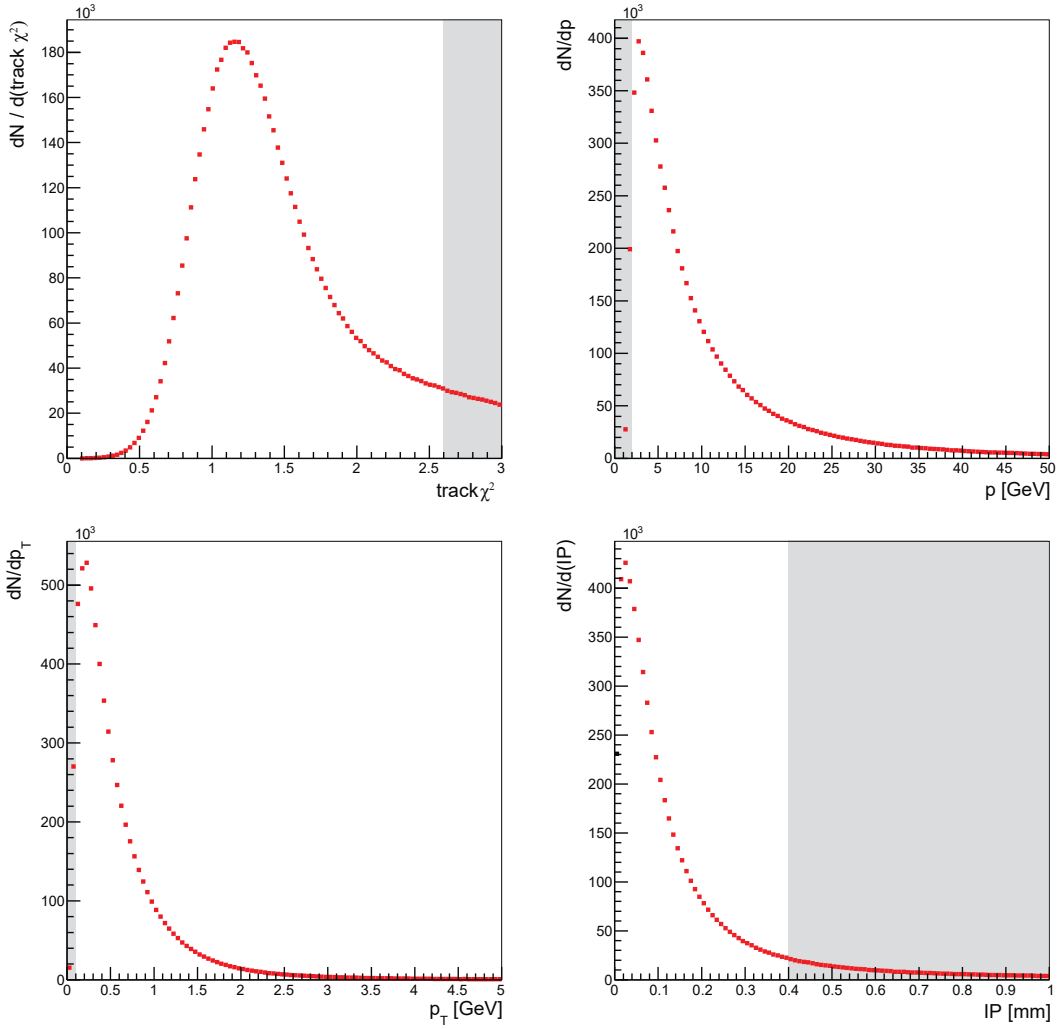
Applying the above preselection criteria resulted in a set of $\sim 5 \times 10^5$ events. The distributions of the variables with the preselection cuts are presented in Fig. 4.3-4.4.

4.3.2 Particle identification

The particle identification process was supported by the MC-trained artificial neural network that assigned a set of additional probabilities (*ProbNN*) to the reconstructed tracks. An example used in the preselection process is *ProbNN(ghost)*, the probability that a reconstructed track was in fact a ghost track (for the definition of a ghost track see Sec. 3.2.1). The other parameters were related to the particle type: pion, kaon, proton, muon, electron.

The simulation with regard to the particle identification is not perfect and *ProbNN* variables required calibration. In the process of *MC resampling*, PID

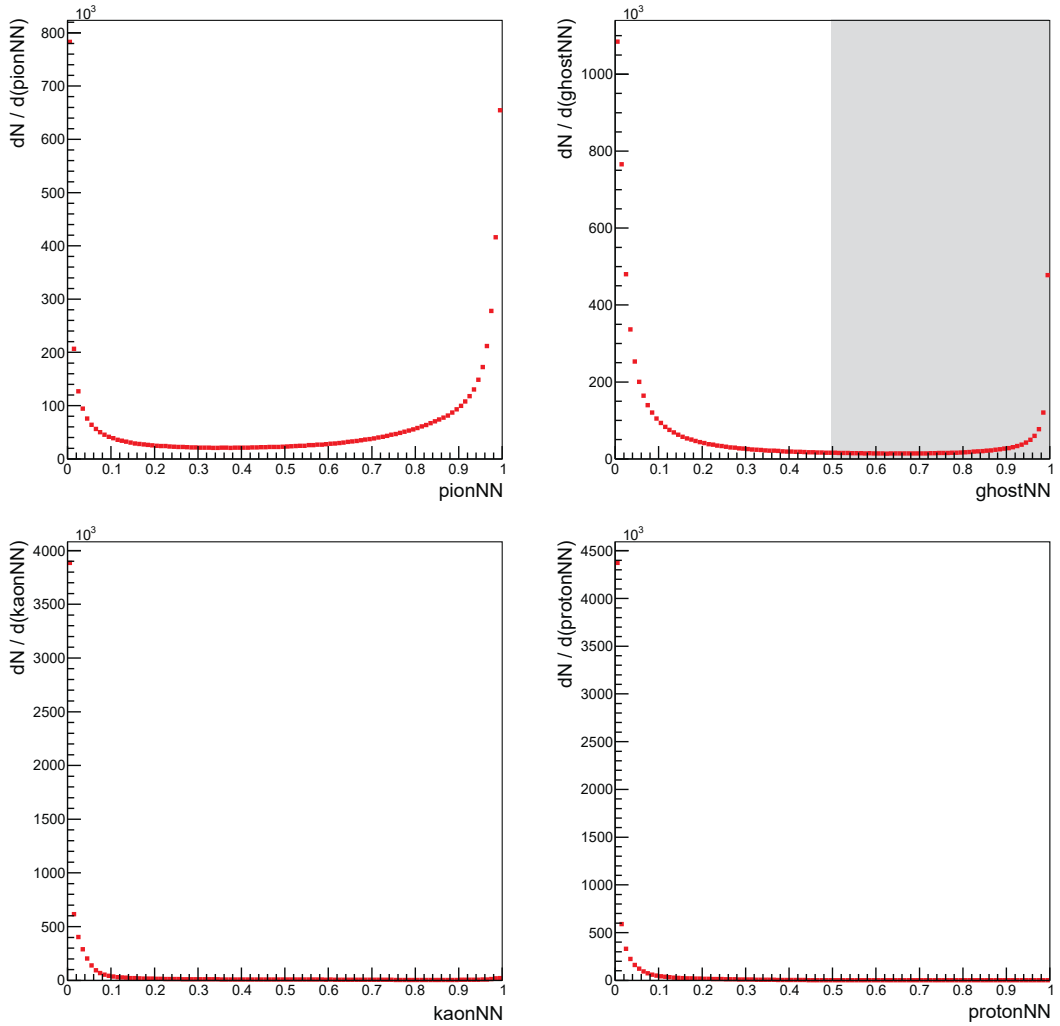
FIGURE 4.3: Distributions for 2011 NoBias data sample before preselection. Range rejected by the preselection cuts marked with a grey-shaded band.



Parameter	Requirement
track χ^2	< 2.6
track p	> 2.0 GeV
track p_T	> 0.1 GeV
track IP	< 0.4 mm
ProbNN(ghost)	< 0.5

TABLE 4.1: Track preselection criteria. *ProbNN(ghost)* is a probability of a ghost track (see Sec. 4.3.2).

FIGURE 4.4: Distributions for 2011 NoBias data sample before preselection. Range rejected by the preselection cuts (if applied) marked with a grey-shaded band.



variable distributions in MC were corrected to agree with the real data. A dedicated tool was developed for this task in the LHCb software, called PIDCalib [115]. It uses PID calibration samples that contain high statistics of particles based only on kinematic selections. A distribution of different types of particles is available (π, K, p, e, μ) in bins of p, η and N_{VELO}. As the type of generated particle is known, random PID variables values from a related calibration sample can be used. There is a small probability that for a given particle type and bin of p, η or N_{VELO}, the calibration bin will be empty. In such a case, the distributions are integrated over N_{VELO}, so only the p and η bins, are available. Empty bins for pions, kaons and protons make up below 0.1% of the total bins. This effect was taken into account in the systematic uncertainty (see Sec. 4.4.3).

4.3.3 Optimization of pion identification

There are several sources that can potentially distort a pure BEC signal. Pairs of kaons or protons could provide their own BEC and FDC, respectively, interfering with the signal coming from pions. Another potential issue comes from mixed pairs – pairs made of different particles (e.g. pion-kaon) and pairs involving wrongly reconstructed tracks. The former do not produce Bose-Einstein correlations, as they are not identical, the latter (described further in Sec. 3.2.1) are controlled with selection requirements.

To avoid possible second order distortions on the double ratio related to the differences between the data and simulation, the requirement on the probNNpi variable had to be chosen optimally. A high purity of the pion sample had to be ensured, but at the same time a strong requirement on the *probNN(pion)* could affect the signal region at low Q. As it can be seen in Table 4.2, an example for medium activity class, the purity of the pion sample increases with the value of the *probNN(pion)* limit; however, it varies only within 1%. The efficiencies of the requirement on the *probNN(pion)* for a single pion are summarized in Table 4.3. The effect of the cut on *probNN(pion)* on the shape of the correlation function for the data and Monte Carlo samples for medium activity class is shown in Fig. 4.5. A stronger cut on *probNN(pion)* has no influence on the shape of the correlation function for the Monte Carlo sample, except for in the first bin. Since in the low-Q region the separation in momentum between two particles is poor and the difference in data and MC reconstruction for the track pairs is significant, the region of $Q < 0.05$ GeV was excluded from further analysis (see Sec. 4.4.1). As the effect related to the cut on *probNN(pion)* is much stronger in the data, it means that a very strict requirement on *probNN(pion)* may reject the region of the phase-space where the signal pairs are expected. The optimal limit was chosen to *probNN(pion)* > 0.65, as for such a value the signal enhancement in the low-Q region for the data begins to saturate. Moreover, the purity of the pion sample for such a cut stays at acceptable level.

4.3.4 Cloned tracks

Contamination of a data sample with incorrectly reconstructed particles can influence the measured Bose-Einstein correlations effect. Cloned tracks are

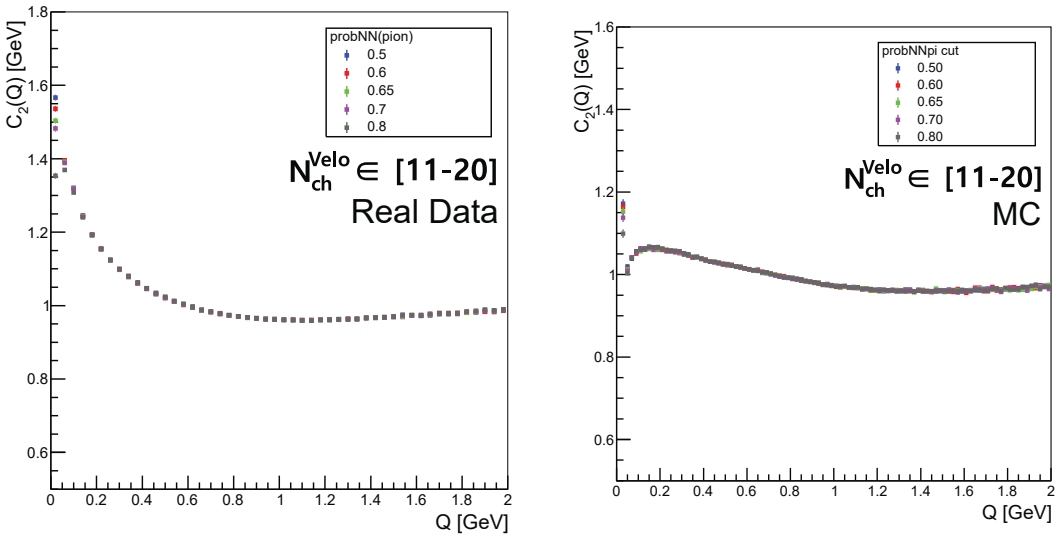
	[%]					
ProbNN(pion)	pions	kaons	protons	electrons	muons	ghosts
>0.50	97.56	0.76	0.46	0.20	0.19	0.83
>0.55	97.75	0.70	0.40	0.16	0.17	0.82
>0.60	97.94	0.62	0.35	0.13	0.15	0.80
>0.65	98.12	0.55	0.29	0.12	0.14	0.78
>0.70	98.30	0.48	0.24	0.10	0.12	0.76
>0.75	98.50	0.41	0.19	0.07	0.10	0.73
>0.80	98.68	0.34	0.15	0.06	0.08	0.69

TABLE 4.2: Fractions of the particles present in the true pion sample for different cuts on *ProbNN(pion)* in the second bin of VELO track multiplicity per *pp* collision $N_{ch}^{VELO} \in [11-20]$. Study performed on the MC data sample.

	Single pion efficiency [%]		
ProbNN(pion)	$N_{ch}^{VELO} \in [5-10]$	$N_{ch}^{VELO} \in [11-20]$	$N_{ch}^{VELO} \in [21-60]$
>0.50	95.63	95.32	94.75
>0.55	95.02	94.66	93.95
>0.60	94.37	93.92	93.02
>0.65	93.54	92.99	91.89
>0.70	92.24	91.65	90.38
>0.75	90.86	90.12	88.54
>0.80	88.77	87.92	86.01

TABLE 4.3: The efficiency of the requirement on *ProbNN(pion)* for different bins of VELO track multiplicity per *pp* collision. Study performed on the MC data sample.

FIGURE 4.5: Two-pion correlation function for different values of *ProbNN(pion)* for real data (left) and Monte Carlo (right).



especially detrimental as they are present mostly in the low-Q region, where the BEC signal is expected, appearing as a pair of almost identical, seemingly correlated, particles. Several cross-checks were made trying to understand the possible effect coming from the clones.

A study of the differences in the particles' slopes in x (t_x) and y (t_y)[2] was performed to assess how many clones remained after the final selection (i.e. applying the requirements on Δt_x and Δt_y). A plot of Δt_x and Δt_y distributions before and after the final selection (see Sec. 4.3.5) in the low-Q region $0.05 < Q < 0.1$ was performed, as illustrated in Fig. 4.6. On the left Δt_x was plotted with the requirement that $\Delta t_y < 0.5$ mrad, while on the right Δt_y was plotted with the requirement that $\Delta t_x < 0.5$ mrad. Such distributions were a motivation to apply a cut on the Δt_x and Δt_y (see Sec. 4.3.5).

An additional requirement on the number of shared VELO hits was applied to reduce the fraction of ghosts and cloned tracks, i.e. if there are two or more tracks that share all hits in the VELO detector, then the only one of them with the best track χ^2 is retained. Another cross-check was related to the observation that there are specific regions in the VELO detector which give rise to clone tracks that are not recognized by the Kullback-Liebler algorithm[116]. In the defult track reconstruction algorithm the track clone distance being the Kullback-Liebler distance to the closest track has to be $> (\log(5000) = 8.52)$. In the overlap regions between the two detector halves one can have particles giving hits in both halves and the reconstruction produces two slightly different tracks. These clone tracks are not recognized as sharing the same VELO hits, and their momenta might be also different enough that they are not affected by the cut on Kullback-Liebler distance. To check a possible effect, the double ratio without any additional assumptions was compared with the double ratio

[2]Where $t_x = p_x/p_z$ and $t_y = p_y/p_z$

Parameter	Requirement
track η	2.0 - 5.0
track χ^2	< 2.0
track p	> 2.0 GeV
track p_T	> 0.1 GeV
track IP	< 0.4 mm
ProbNN(pion)	> 0.65
ProbNN(kaon, proton)	< 0.5
ProbNN(ghost)	< 0.25
$\|\Delta t_x\|, \|\Delta t_y\|$	$\|\Delta t_x\| > 0.3$ mrad OR $\|\Delta t_y\| > 0.3$ mrad

TABLE 4.4: Final selection criteria.

constructed by removing the particle pairs where both particles have $80^\circ < \phi < 100^\circ$ or $260^\circ < \phi < 280^\circ$, as it is illustrated in Fig. 4.7. As it can be seen, the double ratios are the same in both cases. This cross-check was added as another contribution to the systematic uncertainty related to clones (see Sec. 4.3.5).

4.3.5 Final selection

The values of the parameters used in the requirements and their optimization are described in the previous sections.

Study of the correlations was limited to the Q range of 0.05 to 2.0 GeV. In the region of very low Q (< 0.05 GeV) the separation in the momentum between the particles was poor and the discrepancy between MC and data grew with $Q \to 0$. Furthermore, as it was checked for the MC sample, in the first bin of the distributions of the correlation function the fraction of pairs containing a ghost was about $\sim 25\%$ in all bins of VELO track multiplicity, while the fraction of pairs containing a clone varied from $\sim 8\%$ for the lowest multiplicity bin up to $\sim 15\%$ for the highest multiplicity bin. The upper value of the $Q = 2.0$ GeV was chosen for the final fits, as there was a proper simulation of the long-range correlations in MC as compared to the data up to $Q = 2.0$ GeV, i.e. the double ratio was approximately flat and close to 1 in this region.

The value of *ProbNN* for pion was set to >0.65 with IP <0.4 mm. Vetoes were imposed on kaons and protons with the requirement of respective *ProbNN*s < 0.5. The pseudorapidity of the outgoing tracks was limited to the LHCb acceptance of $2.0 < \eta < 5.0$.

Additional limits on the tracks' slopes were introduced, as small slope differences may indicate cloned tracks. Particle triplets were included in the analysis if the difference in Δt_x and Δt_y with respect to other particles from the triplet was larger than 0.3 mrad. Table 4.4 summarizes the final selection requirements on the parameters for the analysis.

FIGURE 4.6: Difference in t_x (left) and t_y (right) before and after final selection in three bins of VELO track multiplicity (top to bottom).

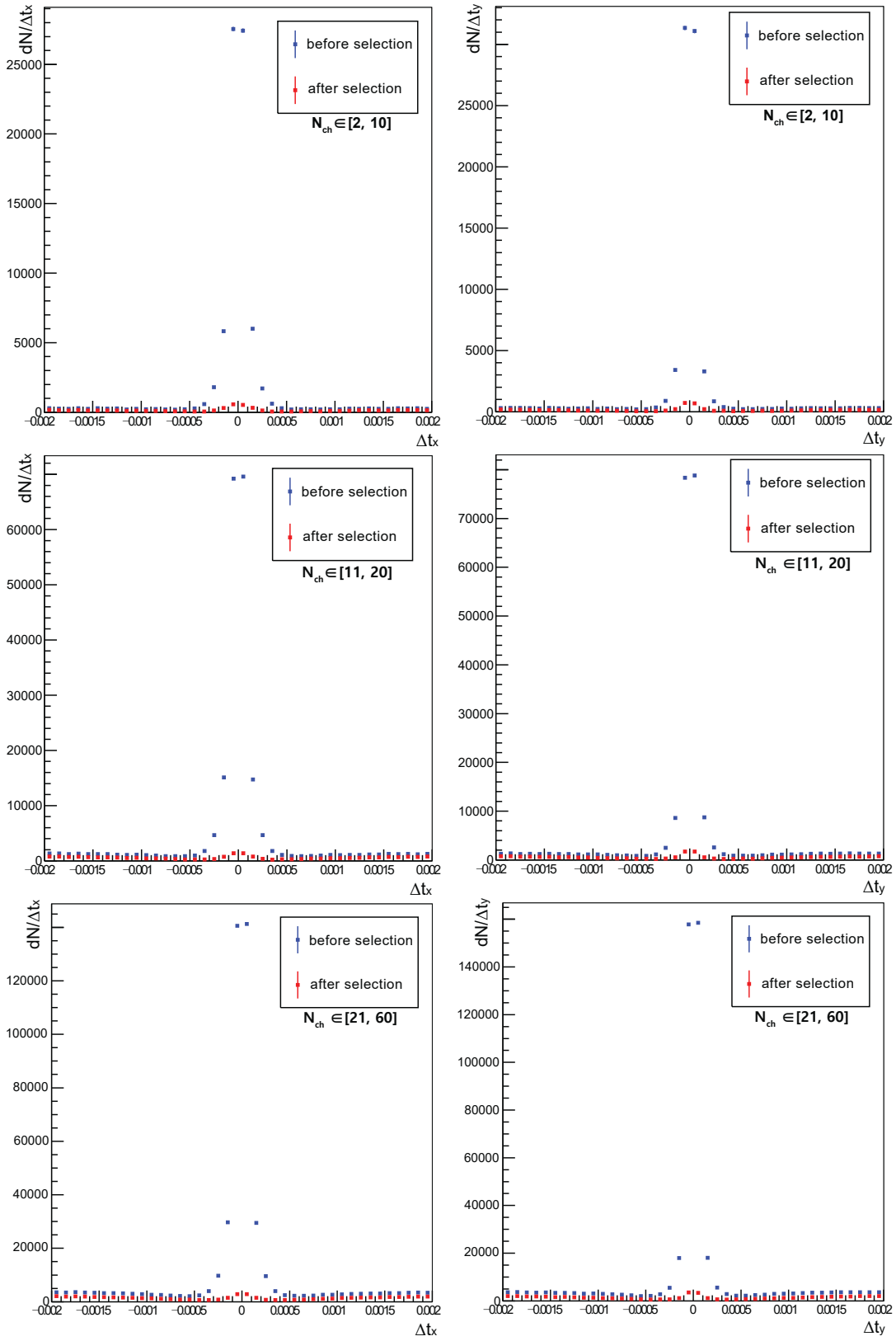


FIGURE 4.7: Comparison of double ratios in different bins of VELO-track multiplicity per PV before and after applying the $80° < \phi < 100°$ or $260° < \phi < 280°$ cut.

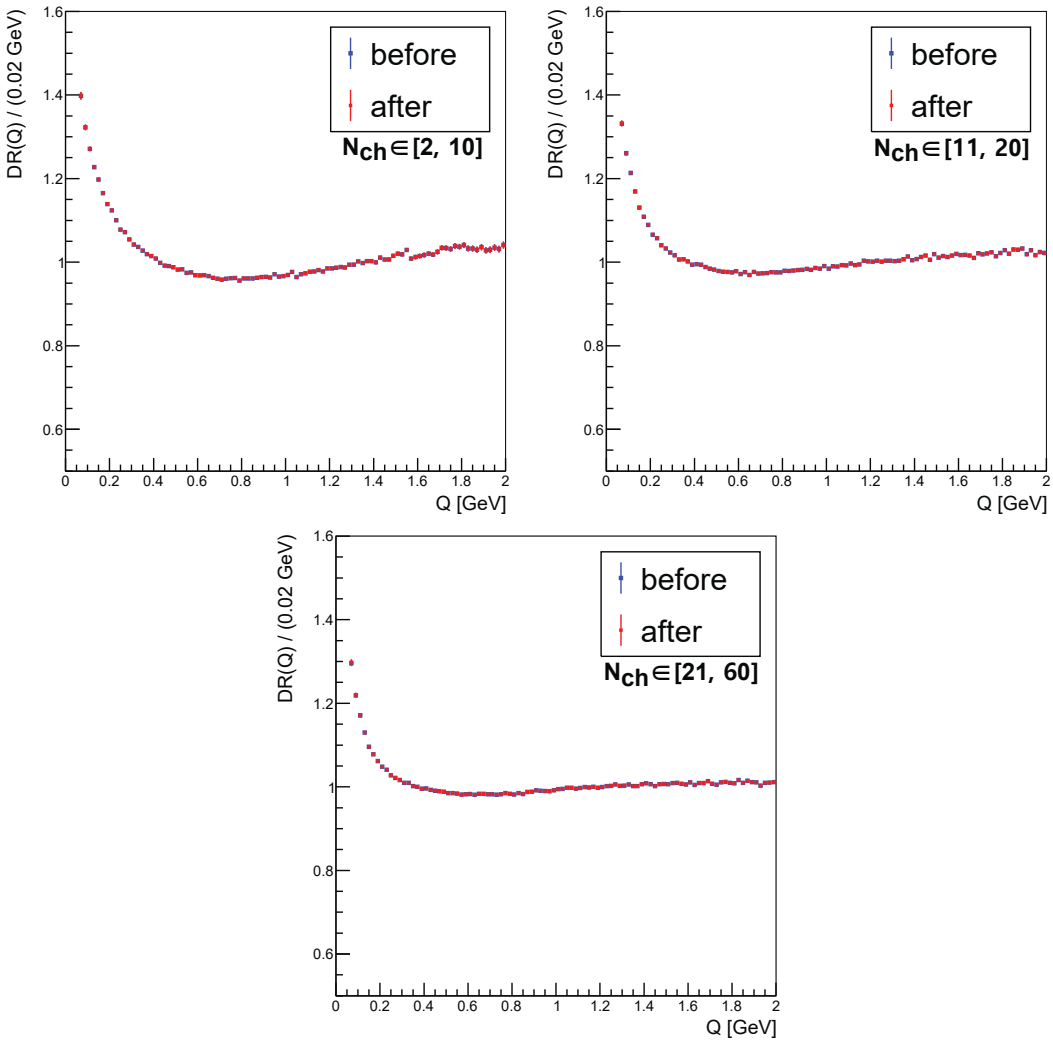
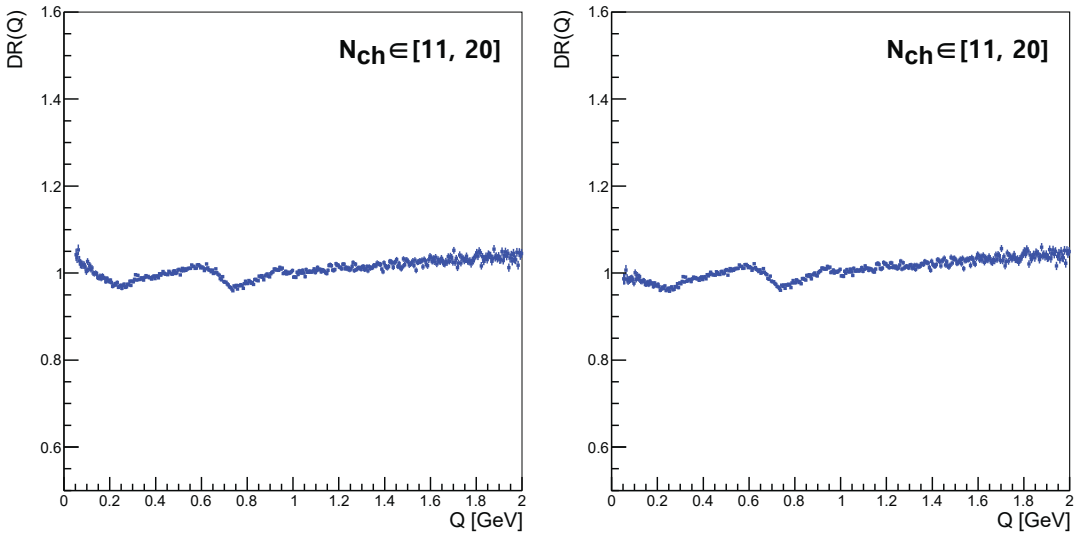


FIGURE 4.8: Double ratios for unlike-sign pion pairs for the second bin of VELO track multiplicity per *pp* collision. Coulomb effect present on the left, subtracted on the right.



4.4 Study of the BEC effect

4.4.1 Analysis procedure

The methods used in the analysis are described in Sec. 4.1. The reference sample was prepared using the event-mixing method (see Sec. 4.1.1). To ensure that the correlations were absent in reference triplets, each particle was chosen from a different event, from the primary vertex of the same charged particle multiplicity.

A correction for the Coulomb interactions between the same-signed particles was performed using the Gamow penetration factor (see Sec. 4.1.3). It was possible to use this simplification because of the small size of the pion source that can be treated more as point-like.

The correction procedure was applied to the correlation function for both the same-sign and the opposite-sign pion pairs in the data. An example of the result of this correction is given in Fig. 4.8, where the double-ratio distribution for the opposite-sign charged pion pairs in the medium-activity event class before and after the Coulomb correction is shown. It is visible that the correction for Coulomb interactions affects only the region of a very low Q. After the correction, the double-ratio distribution in this regime ($Q < 0.2$ GeV) became nearly flat, which is an expected behaviour when the Coulomb interactions are properly corrected for. A potential impact of the inaccuracy of this method on the final results was taken into account in the systematic uncertainty studies described in Sec. 4.4.3. The observed deviation from the straight line in the higher Q region results from the imperfect simulation of the $\pi^+\pi^-$ resonant structure which occurs for the unlike-sign pion MC sample.

Parameter	Published in [52]	Pairs in triplets
	N_{ch} 5-10	
R	5.10 ± 0.07	5.15 ± 0.09
λ_2	0.72 ± 0.01	0.70 ± 0.01
	N_{ch} 11-20	
R	7.52 ± 0.09	7.49 ± 0.09
λ_2	0.63 ± 0.01	0.63 ± 0.01
	N_{ch} 21-60	
R	9.15 ± 0.14	9.24 ± 0.14
λ_2	0.57 ± 0.01	0.55 ± 0.01

TABLE 4.5: Comparison of R and λ_2 in two-particle correlation functions fits between values published in [52] and calculated for pairs inside triplets. For triplets, average value of three pairs given. Uncertainties are statistical.

To account for nonfemtoscopic background and effects related to the selection requirements and construction of the reference sample, the double ratio technique was used (see Sec. 4.1.2).

To calculate parameters of the core-halo model, the double ratio of the correlation function was fitted using the parametrisation from Eq. 2.37.

4.4.2 Determination of the core-halo parameters

The calculation of the three-pion correlation function was based on the convolution of two-pion correlation functions, where the pion pairs must originate from the triplet of the same-sign pions assigned to the same PV (see Sec. 4.1.1). Correlation functions for a pair of pions in a triplet are shown in Fig. 4.9 for Q_{12} as an example. A signal enhancement related to the BEC effect is visible for low Q values. By contrast, the corresponding correlation function for the simulated sample in Fig. 4.10 shows no enhancement in pairs of identical pions with a small four-momentum difference.

In the first step the Coulomb corrected two-particle double ratios for Q_{12}, Q_{23} and Q_{13} were fitted using the formula for the two-particle correlations, with the correlation function given as:

$$C_2(Q) = N(1 + \lambda e^{-RQ}) \times (1 + \delta \cdot Q). \tag{4.8}$$

The results for the R and λ_2 parameters from such fits were then compared to the values published in [52] in order to check their compatibility. The values of R and λ_2 proved to be compatible in all bins of VELO track multiplicity, as shown in Tab. 4.5. In Fig. 4.11 the fit result for the two-particle double ratio in the middle bin of activity for Q_{12} is shown as an example.

Finally, the distributions of the double ratio of correlation functions for like-sign pion triplets, with the event-mixed reference sample, were fitted in three

FIGURE 4.9: Two-particle correlation functions for real data. Example for Q_{12} in three bins of VELO track multiplicity.

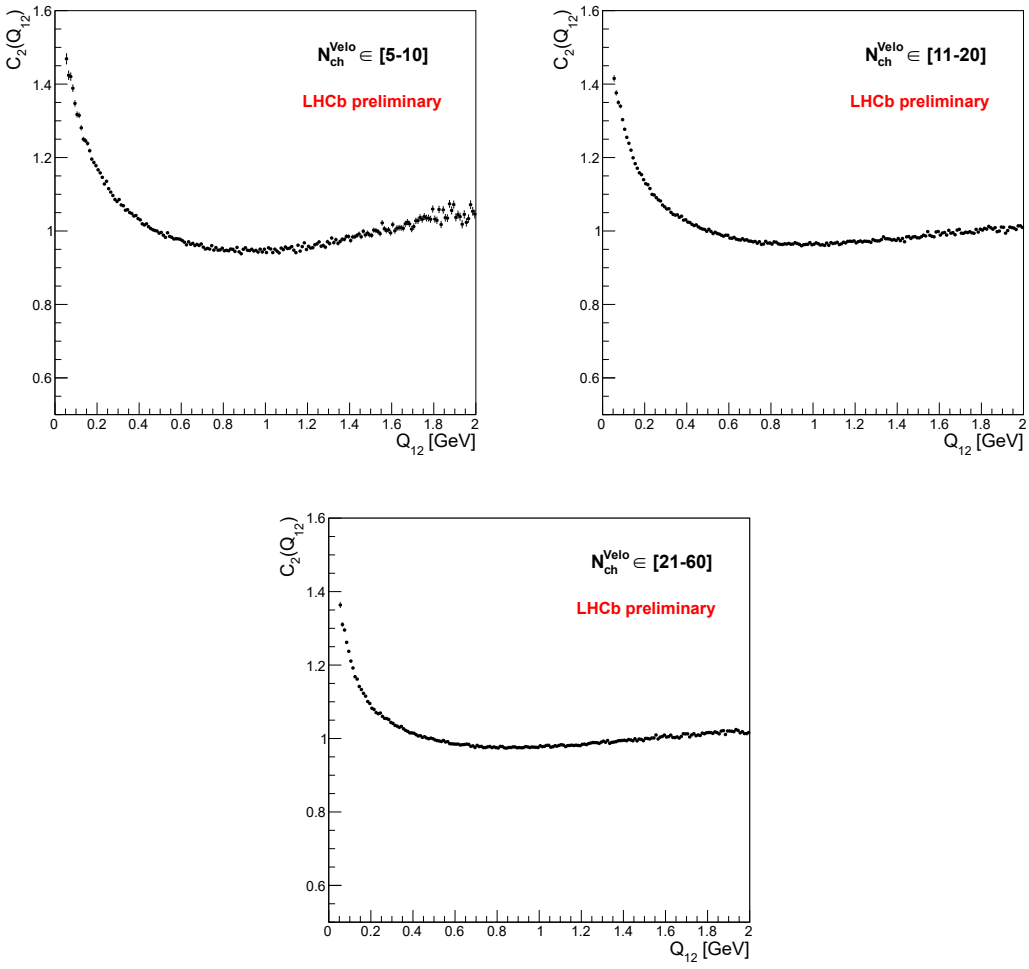


FIGURE 4.10: Two-particle correlation functions for Monte Carlo. Example for Q_{12} in three bins of VELO track multiplicity.

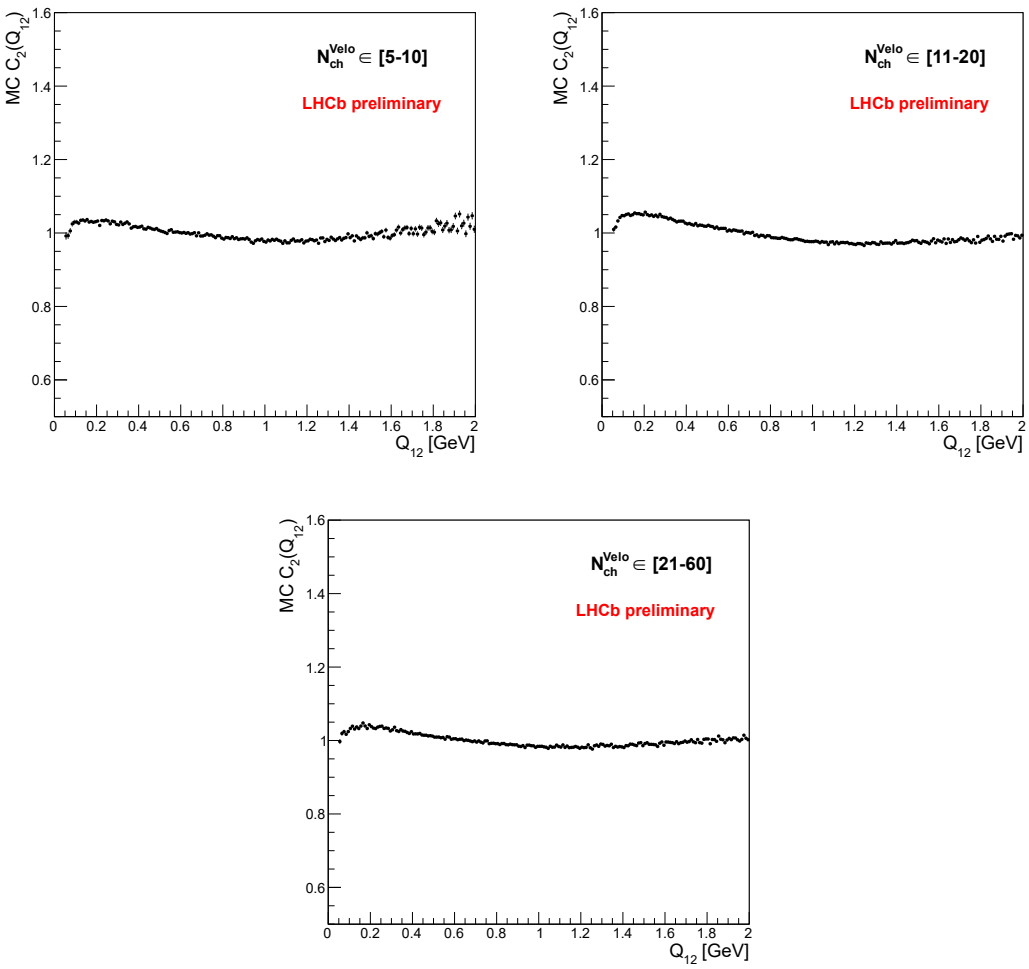
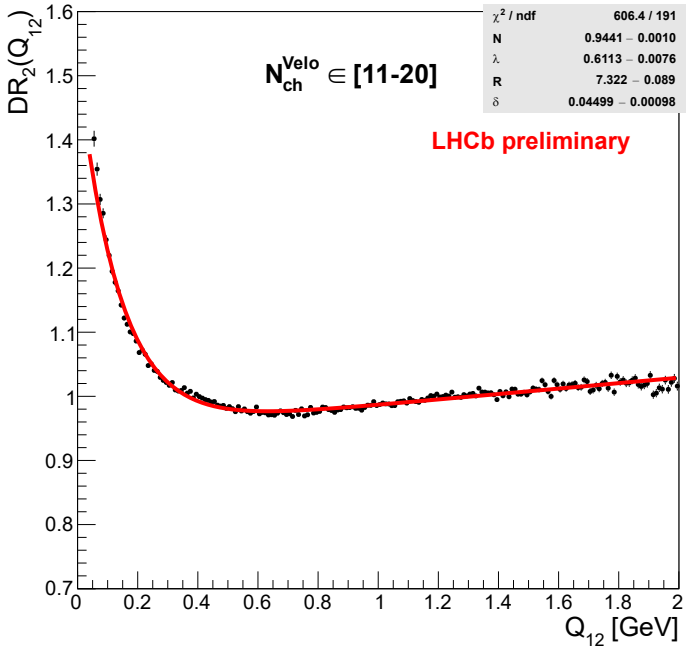


FIGURE 4.11: Fit to two-particle double ratio of correlation function for Q_{12} using Eq. 4.8.



different bins of VELO track multiplicity per pp collision using the parametrisation presented in Eq. 2.37. Fits to the double ratio are presented in Fig. 4.12. Included pull distributions show a good quality of the fit, only single points for the lowest Q values are outside the 3σ range. The Bose-Einstein correlation effect is visible as a signal enhancement for low values of Q.

Fit parameters were used to calculate the values of the parameters in the core-halo model according to the description in Sec. 2.4.3. In Table 4.6 the results are presented in three bins of VELO charged particle multiplicity with statistical and systematic uncertainties (described in detail in Sec. 4.4.3).

Dependence of the core-halo parameters on the VELO track multiplicity is shown in Fig. 4.13. The fraction of the core f_c decreases slightly for classes of higher activity. The partial coherence p_c increases significantly for higher multiplicities, which is reflected in the central value of κ_3 deviating from 1, but still staying within the error margin. The intercept parameter λ_3 shows steady decrease with growing multiplicity. With the exception of the λ_2 parameter, the systematic uncertainty is comparable to the statistical one.

4.4.3 Systematic uncertainties

Several sources of the systematic uncertainties were studied. The values calculated for each of the sources are summarized in Table 4.7, excluding the sources that proved to be negligible with an impact of $<0.1\%$. The results are

FIGURE 4.12: Results of the fit to double ratio (DR_3) for like-sign pion triplets with event-mixed reference sample and correction for the Coulomb effect in three bins of VELO track multiplicity for pp collision: 5-10 (top left), 11-20 (top right), 21-60 (bottom). Red line depicts the fit using the parametrisation described in the text. Diagonal visualization for $Q_{12} = Q_{13} = Q_{23}$ is depicted. Corresponding pull distributions are presented below the fits.

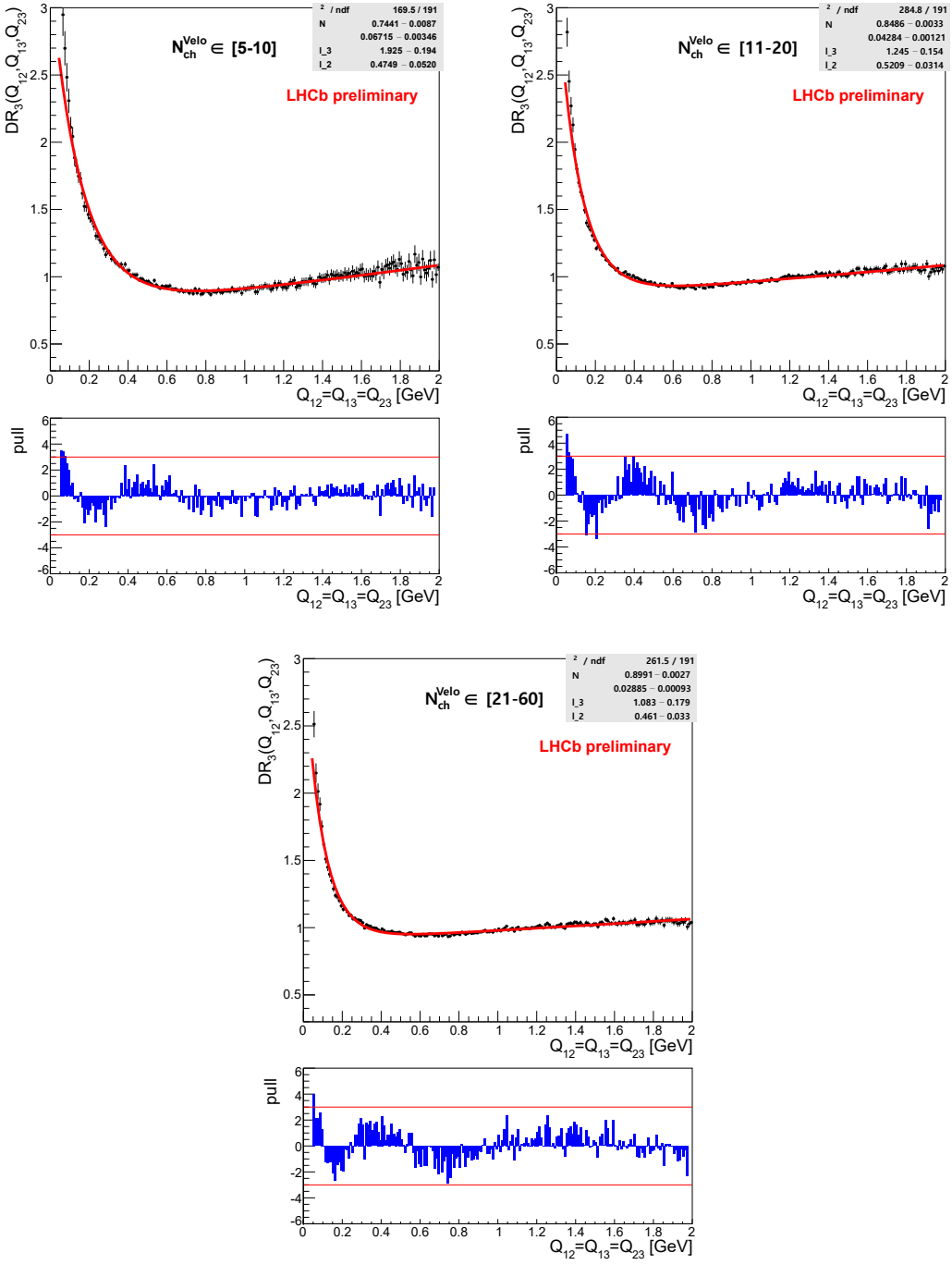
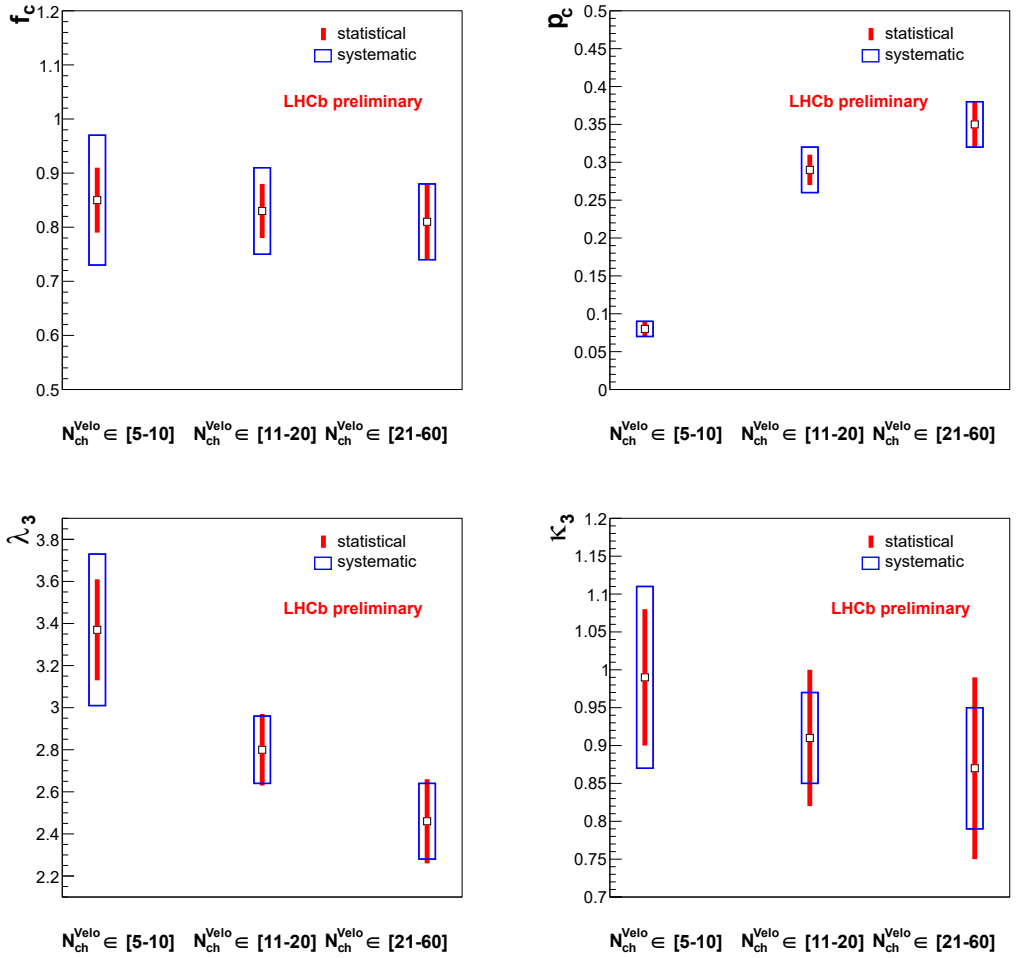


FIGURE 4.13: Values of the parameters of the core-halo model calculated for three bins of VELO track multiplicity. Statistical and systematic uncertainties marked with red and blue bars, respectively.



Parameter	N_{ch} 5-10
λ_2	$0.72 \pm 0.01(1.4\%) \pm 0.05(7.6\%)$
λ_3	$3.37 \pm 0.24(7.2\%) \pm 0.36(10.8\%)$
f_c	$0.85 \pm 0.06(7.3\%) \pm 0.12(13.8\%)$
p_c	$0.08 \pm 0.01(7.3\%) \pm 0.01(13.0\%)$
κ_3	$0.99 \pm 0.09(9.6\%) \pm 0.12(11.9\%)$
	N_{ch} 11-20
λ_2	$0.63 \pm 0.01(1.6\%) \pm 0.05(7.3\%)$
λ_3	$2.80 \pm 0.17(6.3\%) \pm 0.16(5.8\%)$
f_c	$0.83 \pm 0.05(6.5\%) \pm 0.08(9.1\%)$
p_c	$0.29 \pm 0.02(6.6\%) \pm 0.03(9.0\%)$
κ_3	$0.91 \pm 0.09(10.0\%) \pm 0.06(6.9\%)$
	N_{ch} 21-60
λ_2	$0.57 \pm 0.01(1.8\%) \pm 0.03(5.6\%)$
λ_3	$2.46 \pm 0.20(8.2\%) \pm 0.18(7.2\%)$
f_c	$0.81 \pm 0.07(8.4\%) \pm 0.07(8.7\%)$
p_c	$0.35 \pm 0.03(8.4\%) \pm 0.03(8.9\%)$
κ_3	$0.87 \pm 0.12(14.1\%) \pm 0.08(9.1\%)$

TABLE 4.6: Calculated values of core-halo parameters with statistical (first number) and systematic (second number) uncertainties, divided by the activity classes. Values of λ_2 are from [52].

	N_{ch} 5-10			
Source	λ_3	f_c	p_c	κ_3
MC generator	8.7%	10.4%	9.7%	10.4%
PV multiplicity	5.9%	8.6%	8.3%	4.7%
PV reconstruction	<1.0%	0.1%	0.1%	<0.1%
Fit binning	0.7%	0.1%	0.4%	1.2%
Fit low-Q range	0.3%	0.6%	1.0%	0.3%
Fit high-Q range	0.4%	0.5%	0.5%	0.3%
Ghost tracks	2.1%	2.4%	2.4%	1.8%
ProbNN(pion)	2.1%	2.7%	1.8%	3.3%
TOTAL	**11.0%**	**12.1%**	**14.0%**	**13.2%**
	N_{ch} 11-20			
Source	λ_3	f_c	p_c	κ_3
MC generator	3.6%	5.1%	5.0%	3.2%
PV multiplicity	2.1%	5.0%	5.0%	2.1%
PV reconstruction	1.1%	1.6%	1.6%	1.1%
Fit binning	1.1%	1.1%	1.1%	1.1%
Fit low-Q range	1.4%	1.5%	1.8%	2.5%
Fit high-Q range	1.8%	1.9%	1.8%	2.1%
Ghost tracks	1.8%	4.3%	4.3%	1.8%
ProbNN(pion)	2.1%	1.7%	1.4%	3.9%
TOTAL	**5.7%**	**9.0%**	**9.0%**	**6.8%**
	N_{ch} 21-60			
Source	λ_3	f_c	p_c	κ_3
MC generator	2.4%	2.8%	2.8%	0.4%
PV multiplicity	3.3%	5.3%	5.4%	3.7%
PV reconstruction	2.4%	2.5%	2.4%	2.8%
Fit binning	1.0%	0.4%	0.7%	2.0%
Fit low-Q range	1.8%	2.4%	3.0%	3.7%
Fit high-Q range	3.0%	3.4%	3.4%	3.3%
Ghost tracks	2.0%	2.1%	2.2%	2.4%
ProbNN(pion)	4.1%	3.9%	3.7%	5.7%
TOTAL	**7.5%**	**8.9%**	**9.1%**	**9.4%**

TABLE 4.7: Systematic uncertainties divided by sources and bins of VELO track multiplicity. Sources that proved to be negligible, omitted in the table, are described in Sec. 4.4.3.

divided by the activity class and detailed and total values are provided for each of the core-halo model parameters.

Selection cuts are treated here as the source of systematic uncertainty to assert their impact on the final result.

MC generator

Physics generator was a dominant source of systematic uncertainty. To study the effect of its choice on the results, a sample of $\sim 1 \times 10^7$ minimum bias events was locally produced using a PYTHIA 6.4 generator with PERUGIA0 tuning. This particular PYTHIA tuning was chosen as it was found that the long-range correlations are well modeled. Other generators were also tried, e.g. HERWIG++; however, they provided a poor description of the long range correlations.

PV multiplicity

In the case of multiple primary vertices in a single event, the correlation function could be affected because of the way the reference sample is constructed. Some residual correlations between the primary vertices of a single event could be preserved and included in the correlation function. This contribution to the systematic uncertainty was calculated as a difference between the central values of the fit results for the two extreme cases: (i) a single PV and (ii) three or four PVs in the event. The impact of the number of PVs in the event on the double ratio is shown in Fig. 4.14.

PV reconstruction

Because of a significant pile-up in 2011 data and MC, some inefficiency of the PV reconstruction can occur. To estimate the systematic impact of the PV reconstruction for both data and MC, randomly selected tracks were removed from the PV at rate of 10% of the primordial set of reconstructed tracks. The same reconstruction algorithm was then executed on the reduced set of tracks.

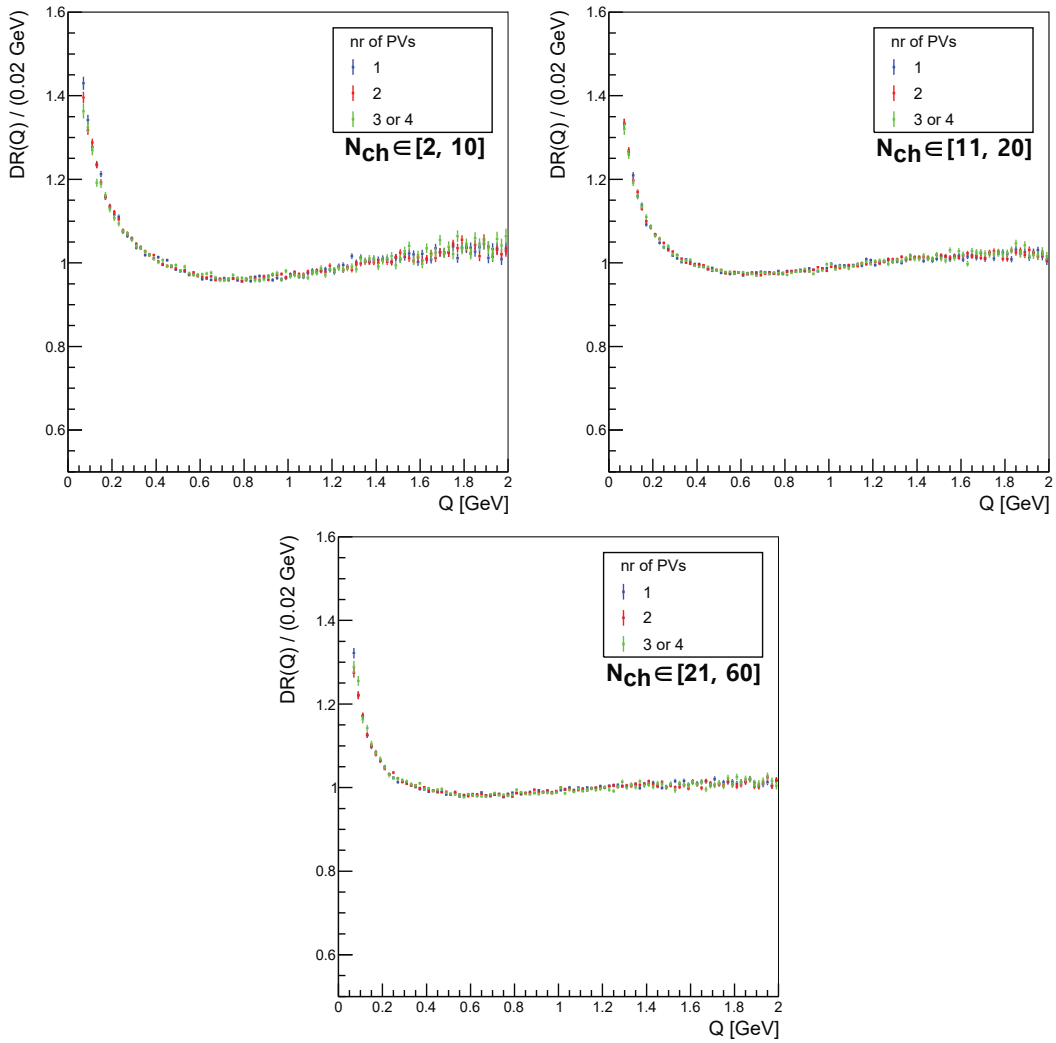
Fit binning

To assess the uncertainty coming from the adopted fit binning, the bin size was changed to 0.003 and 0.005 GeV. The uncertainty was found to be negligible for all bins of VELO track multiplicity.

Fit range in low-Q region

The systematic effect of the choice of the lower Q limit was tested by changing the limit by $\pm 20\%$ conservatively, i.e. to 0.04 and 0.06 GeV. The fits to the double ratio with two different ranges of Q were redone in three different bins of VELO track multiplicity.

FIGURE 4.14: Double ratio for like-sign pion pairs in different bins of VELO track mutiplicity for the events with 1 (blue), 2 (red) and 3-4 (green) PVs.



Cut on *ProbNN(ghost)*	N_{ch} 5-10	N_{ch} 11-20	N_{ch} 21-60
< 0.25	0.66%	0.88%	1.20%
< 0.50	0.75%	1.01%	1.37%

TABLE 4.8: The fractions of the ghosts after the final selection using MC truth information for three bins of VELO track multiplicity per *pp* collision for different cuts on *ProbNN(ghost)*.

Fit range in high-Q region

The systematic effect of the choice of the upper Q limit was tested by changing the limit by ±20% conservatively, i.e. to 1.8 and 2.2 GeV. The fits to the double ratio with two different ranges of Q were redone in three different bins of VELO track multiplicity

Ghost tracks

Most ghosts were already removed by the cuts on the track χ^2 and track probability to be a ghost as well as by applying a requirement on shared VELO hits, but there may be a discrepancy in ghost track ratios between the data and MC. To determine the systematic uncertainty related to the ghost tracks, the double ratio was refitted with a loose cut on the probability of a track being a ghost *ProbNN(ghost)* < 0.50. The fraction of the ghosts after the final selection using MC truth information for three different bins of VELO track multiplicity and for different cuts on ProbNN(ghost) < 0.25 and < 0.50 is gathered in Table 4.8.

ProbNN

A cut on the *ProbNN(pion)* adjusted the contamination of pions related to the misidentification. The systematic impact of the cut on the measured parameters was assessed using a looser cut (>0.35) that increased the fraction of misidentified pions by 50%.

Cloned tracks

The cloned tracks proved not to be influential, as only up to 1% of the simulated events contained them after the preselection. The systematic effect of the cloned tracks was determined by refitting the double ratio with the event-mixed reference sample for the cut on the logarithm of the Kullback-Liebler distance increased to 12.0. The observed change was insignificant in all bins of VELO track multiplicity.

A contribution to the systematic uncertainty coming from the tracks reconstructed incorrectly, due to hits in both VELO detector halves in the overlap region, was also tested and proved to be negligible. The detailed discussions on the misreconstructed tracks may be found in Sec. 4.3.4.

Coulomb correction

The Coulomb correction was performed using the Gamov factor (described in Sec. 4.1.3) which was only an approximation. However, the impact of changing the correction by ±20% on the final results was negligible.

Resolution of the Q variable

The impact of the resolution of the Q variable can be tested by smearing Q values with the Gaussian function with a standard deviation related to the target resolution. The impact caused by the slight resolution changes proved to be insignificant.

Particle identification

Particle identification plays a crucial role in the selection process and can directly influence the final results. The failure rate related to empty phase space bin (bin of p, η, N_{VELO}, where there are no resampling distributions from PIDCalib) stood for 2.2%, 0.4% and 0.2% for π, K and p, respectively. In such a case, the integrated distributions over track multiplicity were used. The systematic uncertainty coming from this effect was estimated by removing the particles pointing to the empty phase space bin from the analyzed sample and refitting the double ratios.

Residual acceptance effects

There is a possibility of differences in acceptance effects between the data and MC in the low-Q region. The possibility was assessed by comparing the signal double ratio with unlike-sign double ratio with Coulomb effect subtracted, where the BEC effect was not present (see Fig. 4.8(right)). A surplus of the signal related to the acceptance effects was then subtracted from the like-sign pairs double ratio and the fit redone. The change in the signal was found to be negligible.

4.5 Results and conclusions

The correlations between three indistinguishable pions were studied using the data collected by the LHCb in 2011 at a centre-of-mass energy of 7 TeV in proton-proton collisions. The fit applied to the three-pion double ratio of the correlation function using formula 2.37 provided parameters needed to interpret the results in the core-halo model. The values of the parameters are presented in Table 4.6, and are divided into three activity classes related to the charged particle multiplicity. The dependence of the core-halo parameters on the charged particle multiplicity is presented in Fig. 4.13. Statistical and systematic uncertainties are provided for all results.

This was the first application of the core-halo model to the three-particle correlations in the proton-proton collision system in the unique pseudorapidity acceptance of the LHCb experiment ($2.0 < \eta < 5.0$). The values of the partial

coherence parameter p_c suggested the presence of the partially coherent emission of pions, growing together with charged particle multiplicity. The value of the parameter κ_3 fell below unity for high N_{ch}, but stayed within statistical and systematic uncertainty, preventing the author from drawing distinct conclusions.

An analysis closest in its scope to the one described in this dissertation was performed by the PHENIX experiment at RHIC [17]. It used the core-halo model to interpret Bose-Einstein correlations for triplets of pions in AuAu colissions at $\sqrt{S_{NN}} = 200$ GeV in bins of transverse mass m_T. The values of κ_3 and correlation strength λ_3 were provided. Despite the differences in the experimental setup, the central values of κ_3 displayed a similar tendency that the central values being slightly below unity, which may suggest a partialy coherent emission. However, the central values were within the statistical and systematic uncertainties, so no strict conclusions could be drawn. The quoted values of three-particle correlation strengths were higher than those measured for proton-proton collisions. The conclusions in this field are limited, however, since the PHENIX analysis did not present the results in bins of particle multiplicity, so a direct comparison of the results for proton-proton and Au-Au is impossible.

This analysis shows potential for further studies to be conducted. The investigated collision systems could be expanded with pPb and PbPb. Further analyses could be performed additionally in the bins of k_T and η. Results in the unique forward acceptance region could help to develop new models and test the existing ones, such as the core-halo model.

Chapter 5

MUonE experiment

Precise and efficient event reconstruction plays a crucial role in every physics analysis, as it was shown in the BEC analysis. With new and updated experiments, the expected numbers of events and detectors occupancies will exceed the capabilities of the hardware to store incoming events on disks. This poses a new requirement to perform offline-quality reconstruction before storing. New techniques are required to be implemented to achieve the desired throughput, with machine learning being the most promising and gaining interest and adoption technique among experiments.

A good example of such an experiment is the MUonE project [22], where the crucial issue is the development of novel techniques in order to provide efficient online reduction of data and to maximize the statistical power of the final physics measurement. The experiment is designed to search for the signs of the New Physics by measuring the hadronic contribution to the anomalous muon magnetic moment a_μ. It is planned to operate at the M2 muon beam at the CERN SPS [23].

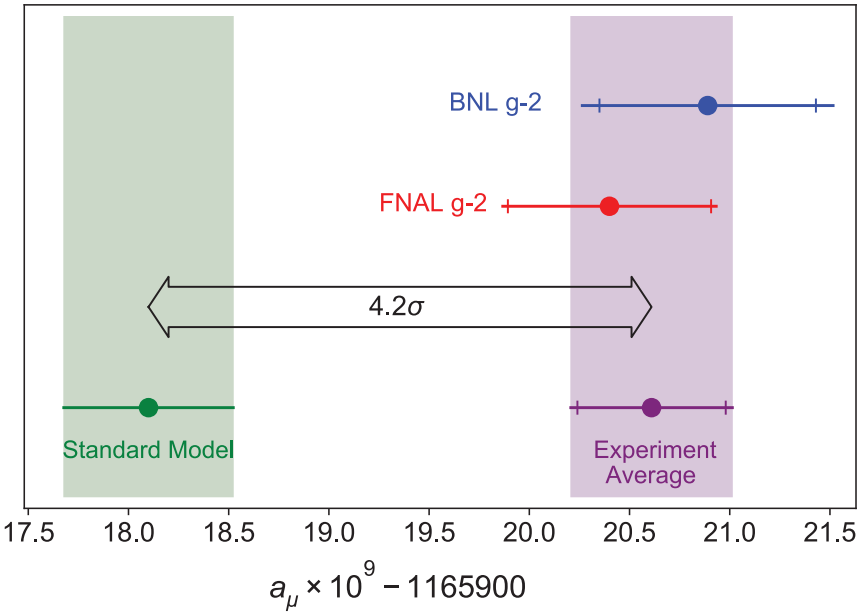
5.1 Physics motivation

The physics purpose of the MUonE experiment, a measurement of the hadronic contribution to the anomalous muon magnetic moment a_μ, is well motivated from the perspective of the Standard Model (SM) [117] which is very successful in describing experimental data. However, the SM is widely believed to be incomplete, and for this reason the search for the so-called New Physics (NP) phenomena beyond the Standard Model (BSM) at the TeV energy scale is being performed at the Large Hadron Collider. Despite the intensive research program undertaken at the LHC to address key issues related to the searches for the NP, up to date the signals for New Physics, if any, have still been hidden in the LHC data without any assumptions on what such New Physics signals might look like. Therefore, there are a number of alternative attempts to hunt for the objects whose behaviour goes beyond the Standard Model of particle physics. A very promising sector for the searches for the New Physics phenomena outside the LHC is related to the measurement of the muon's anomalous magnetic moment.

The muon's anomalous magnetic moment a_μ is defined as:

$$a_\mu = \frac{g-2}{2}, \tag{5.1}$$

FIGURE 5.1: Comparison of previous measurements of anoumalous muon magnetic moment a_μ with SM prediction [24].



where g is the gyromagnetic proportionality factor.

The results from the Fermilab E989/Muon *g-2* experiment [24] show the measured value $a_\mu^{exp} = 11659209.1(6.3) \times 10^{-10}$. Compared with the Standard Model value of $a_\mu^{SM} = 11659182.0(3.6) \times 10^{-10}$, the discrepancy is $a_\mu^{exp} - a_\mu^{SM} = 27.1(7.3) \times 10^{-10}$, which corresponds to 3.5-4 standard deviations. Those are in an agreement with previous measurements from E821 in Brookhaven National Laboratory [25], leading to the combined discrepancy of 4.2 σ between the experiment and theory. A summary of the previous results is presented in Fig. 5.1. In the following years the Fermilab experiment is expected to increase the precision by about a factor of 4, while another forthcoming experiment at J-PARC [118] should reach a similar precision. A serious limitation on increasing the significance of a possible discovery will be, however, the theory calculation, dominated by the leading order contribution from hadronic vacuum polarization $a_\mu^{HVP,LO}$, which cannot be determined using perturbative QCD (pQCD) methods. This is why a novel experiment has been proposed to measure the hadronic component of the running electromagnetic coupling in a momentum transfer region relevant to the calculation of the muon g-2 anomaly, allowing the sensitivity to the potential discovery of New Physics phenomena to be increased. The experiment is called MUonE [22] – it will enable a precise measurement of the hadronic contribution to the anomalous muon magnetic moment, employing the measurement of the shape of the differential cross-section for the $\mu e \to \mu e$ elastic process [119]. This method could reach a competitive precision below 0.5% on $a_\mu^{HVP,LO}$, under the condition that the systematic uncertainties are well controlled. This, together with the results from the BNL- E821, Fermilab-E989 and J-PARC g-2 experiments, will increase the significance of the observed discrepancy to the level of 7σ.

5.2 Theory predictions

The anomalous muon magnetic moment is composed of several contributions:

$$a_\mu = a_\mu^{QED} + a_\mu^{EW} + a_\mu^{QCD} + a_\mu^{NP}, \tag{5.2}$$

that correspond to quantum electrodynamics, electroweak, quantum chromodynamics and New Physics, respectively. An accurate determination of the first three contributions is necessary to define the impact of the New Physics.

5.2.1 QED contribution

The quantum electrodynamic term in Eq. 5.2 originates from the self-interaction process [120, 121]. It is caused by the quantum fluctuations related to the emission and absorption of virtual photons in a region of particle-field interaction leading to vacuum polarization by those photons into virtual particle-antiparticle pairs.

The quantum electrodynamic input is expected to be [122]:

$$a_\mu^{QED} = A_1 + A_2(m_\mu/m_e) + A_2(m_\mu/m_\tau) + A_3(m_\mu/m_e, m_\mu/m_\tau), \tag{5.3}$$

where m_μ, m_e, m_τ are the masses of muon, electron and taon, respectively. All A_i factors can be determined perturbatively:

$$A_i = \left(\frac{\alpha}{\pi}\right) A_i^{(2)} + \left(\frac{\alpha}{\pi}\right)^2 A_i^{(4)} + \left(\frac{\alpha}{\pi}\right)^3 A_i^{(6)} + \ldots, \tag{5.4}$$

where α is the fine-structure constant. Perturbative calculations were performed up to 10^{th} order [123, 124]. The most accurate value of the fine-structure constant was determined by the Cs atom interferometry experiment [125] to be:

$$\alpha^{-1}(Cs) = 137.035999046(27). \tag{5.5}$$

The final value of the calculated QED contribution to anomalous muon magnetic moment is:

$$a_\mu^{QED} = 116584718.931(104) \times 10^{-11}. \tag{5.6}$$

5.2.2 EW contribution

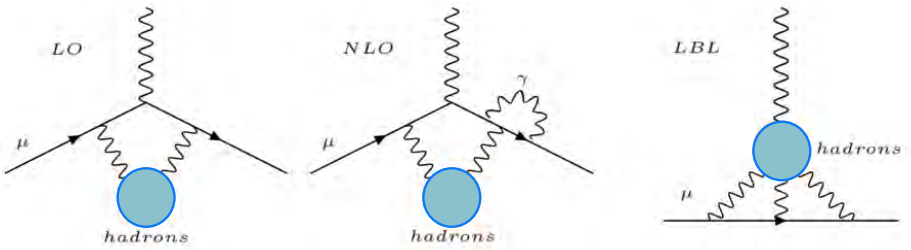
The electroweak term in Eq. 5.2 describes the loop contributions involving $W^\pm$, Z, Higgs bosons and neutrinos. It includes one-loop processes, as well as boson and fermion parts for two-loop processes:

$$a_\mu^{EW} = a_\mu^{EW}[\text{1-loop}] + a_\mu^{EW}[\text{2-loop,bos.}] + a_\mu^{EW}[\text{2-loop,ferm.}]. \tag{5.7}$$

The contribution from one-loop processes is [122]:

$$a_\mu^{EW}[\text{1-loop}] = 194.79(1.0) \times 10^{-11}, \tag{5.8}$$

FIGURE 5.2: Diagrams of the general types of QCD contributions, from left: Leading-Order, Next-to-Leading-Order, Light-by-Light.



while from two-loop processes are [126]:

$$a_\mu^{EW}[\text{2-loop,bos.}] = -18.42(0.1)\times10^{-11}; \quad a_\mu^{EW}[\text{2-loop,ferm.}] = -18.34(0.2)\times10^{-11} \tag{5.9}$$

Together they result in:

$$a_\mu^{EW} = 153.6(1.0) \times 10^{-11}. \tag{5.10}$$

The contributions from higher-loop processes proved to be negligible [127].

5.2.3 QCD contribution

Strong interactions constitute the smallest contribution to the total anomalous muon magnetic moment, but they introduce the most of the uncertainty. The QCD term consists of the HVP (Hadronic Vacuum Polarization) and HLbL (Hadronic Light-by-Light), see Fig. 5.2:

$$a_\mu^{QCD} = a_\mu^{HVP} + a_\mu^{HLbL}. \tag{5.11}$$

The HVP term originates from the hadronic loop vacuum polarization processes with significant contributions from the Leading-Order ($a_\mu^{HVP,LO}$), Next-to-Leading-Order ($a_\mu^{HVP,NLO}$) and Next-to-Next-to-Leading-Order ($a_\mu^{HVP,NNLO}$). There are two major ways to determine the HVP contributions. The first method is based on theory and lattice calculations [122], but it is model dependent, rendering results not directly comparable. The other methos uses experimental data to calculate dispersion relation, providing the most precise results used to determine the SM prediction. The Leading-Order contribution is provided by the dispersion integral:

$$a_\mu^{HVP,LO} = \frac{\alpha^2}{3\pi^2}\int_{4m_\pi^2c^4}^{\infty} ds \frac{K(s)}{s} R(s), \tag{5.12}$$

where $K(s)$ is the so-called kernel function, m_π is the rest mass of a pion, and $R(s)$ is the hadronic ratio:

$$R(s) = \frac{\sigma^0(e^+e^- \to hadrons(+\gamma))}{\sigma_{pt}}; \quad \sigma_{pt} = \frac{4\pi\alpha^2}{3s}, \tag{5.13}$$

where $\sigma^0(e^+e^- \to hadrons(+\gamma))$ is the cross-section of an electron-positron pair annihilation to hadrons. This dependence on the cross-section makes it possible to calculate the value of the dispersion integral using experimental data.

At the moment, the best results for HVP contributions are [128–130]:

$$a_\mu^{HVP,LO} = 6931(40) \times 10^{-11}, \tag{5.14}$$

$$a_\mu^{HVP,NLO} = -98.3(7) \times 10^{-11}, \tag{5.15}$$

$$a_\mu^{HVP,NNLO} = 12.4(1) \times 10^{-11}, \tag{5.16}$$

leading to total HVP contribution of:

$$a_\mu^{HVP} = 6845(40) \times 10^{-11}. \tag{5.17}$$

The HLbL contribution corresponds to the interaction of photons that produces additional photons [131]. In this case, an external on-shell photon interacts with three off-shell photons that couple to a muon. The HLbL contribution has to be calculated using lattice QCD or experimental data, as it is a non-perturbative process [122]. This leads to:

$$a_\mu^{HLbL} = 92(18) \times 10^{-11}. \tag{5.18}$$

The measurement based on the dispersive approach suffers from the resonances and threshold effects in the functional form of the s-channel cross-section of $e^+e^- \to hadrons$, which makes the dispersive approach difficult and finally not precise enough. Moreover, an alternative evaluation of the leading hadronic vacuum polarization term employing lattice QCD calculations is not yet to be conclusive. This is why a novel method was proposed to measure the hadronic component of the running electromagnetic coupling in a momentum transfer region relevant to the calculation of the muon g-2 anomaly, which is described in the following section.

5.3 Hadronic contribution determination with *μ-e* elastic scattering

The hadronic component of a_μ^{SM} is a major source of a theoretical uncertainty for the *g-2* prediction. Hadronic loops are the main limitation related to potential New Physics discovery. They originate mainly from the leading order

hadronic vacuum polarization term $a_\mu^{HVP,LO}$ that cannot be calculated pertubatively and is commonly calculated via a dispersion integral on the hadron production cross-section in the e^+e^- annihilation [132] (see Sec. 5.2.3).

It is possible to estimate the value of $a_\mu^{HVP,LO}$ using lattice QCD calculations [133–135], although the lattice determinations are still not conclusive.

The MUonE experiment is going to use an alternative way to evaluate $a_\mu^{HVP,LO}$ via a measurement of the effective electromagnetic coupling in the space-like region [136]. An independent determination of the hadronic vacuum polarization contribution to the $a_\mu^{HVP,LO}$ will help to assess a_μ more precisely.

The MUonE experiment will use the scattering of high-energy muons on the atomic electrons in a low-Z target through the elastic process of $\mu e \to \mu e$ to measure the hadronic part of running of the electromagnetic coupling constant in the space-like region [119].

Direct sensitivity to $a_\mu^{HVP,LO}$ is provided by the differential cross-section of the $\mu e \to \mu e$ process measured as the function of the squared momentum transfer in the space-like domain $t = q^2 < 0$ [136]. It is obtained by integrating the effective fine-structure constant:

$$a_\mu^{HVP,LO} = \frac{\alpha}{\pi} \int_0^1 dx(1-x)\Delta\alpha_{had}[t(x)], \tag{5.19}$$

where $\Delta\alpha_{had}$ is the hadronic contribution to the running of α, evaluated at squared momentum transfer

$$t(x) = \frac{x^2 m_\mu^2}{x-1} < 0. \tag{5.20}$$

The integrand on $a_\mu^{HVP,LO}$ (Eq. 5.19) is shown in Fig. 5.3(right).

Hadronic contribution $\Delta\alpha_{had}[t(x)]$ can be calculated using:

$$\alpha(t) = \frac{\alpha_0}{1 - (\Delta\alpha_{lep}(t) + \Delta\alpha_{had}(t))}, \tag{5.21}$$

with leptonic contribution $\alpha_{lep}(t)$ known form perturbative calculations. In Fig. 5.3(left), $\Delta\alpha_{lep}[t(x)]$ and $\Delta\alpha_{had}[t(x)]$ are shown as functions of x and t.

The dependence of the differential cross-section of the μ-e process is proportional to $|\alpha(t)/\alpha_0(t)|^2$:

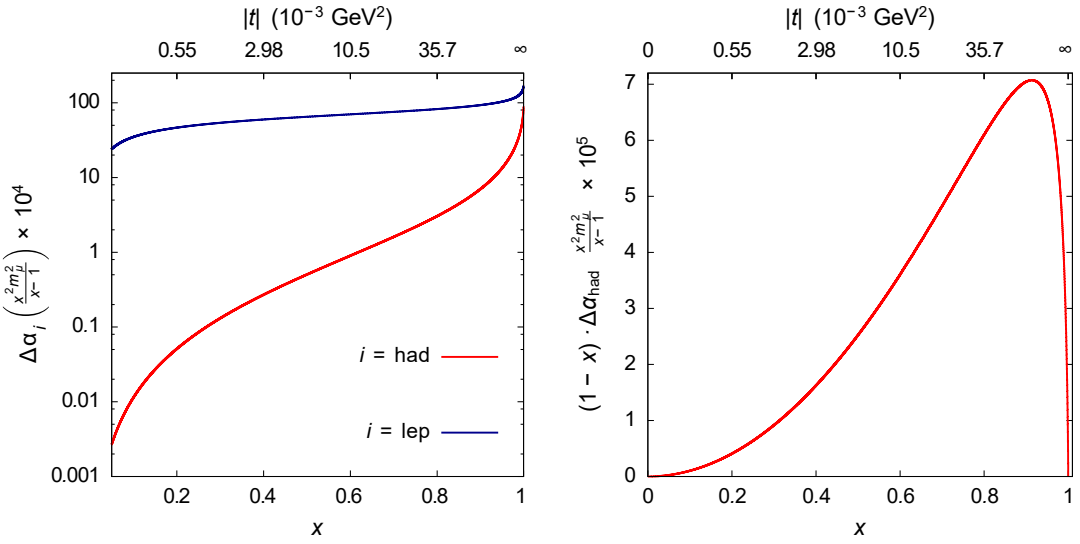
$$\frac{d\sigma}{dt} = \frac{d\sigma_0}{dt} \left| \frac{\alpha(t)}{\alpha_0(t)} \right|^2, \tag{5.22}$$

where $d\sigma_0/dt$ is the effective Born cross-section, including virtual and soft photons. The vacuum polarization effect is included in $\alpha(t)/\alpha_0(t)$.

With a known energy of an incoming muon and atomic electron being a fixed target, t is related to the energy of the scattered electron E_e^f or its angle θ_e^f:

$$t = (p_\mu^i - p_\mu^f)^2 = (p_e^i - p_e^f)^2 = 2m_e^2 - 2m_e E_e^f. \tag{5.23}$$

FIGURE 5.3: Left: $\Delta\alpha_{lep}[t(x)]$ and $\Delta\alpha_{had}[t(x)]$ as functions of x and t. Right: The integrand on $a_\mu^{HVP,LO}$ (Eq. 5.19) [119].



The energy and angle of the scattered electron can be expressed as:

$$E_e^f = m_e \frac{1 + r^2 \cos^2 \theta_e^f}{1 - r^2 \cos^2 \theta_e^f}, \tag{5.24}$$

$$\theta_e^f = \arccos\left(\frac{1}{r}\sqrt{\frac{E_e^f - m_e}{E_e^f + m_e}}\right), \tag{5.25}$$

where:

$$r \equiv \frac{\sqrt{(E_\mu^f)^2 - m_\mu}}{E_\mu^i + m_e}. \tag{5.26}$$

The angles of the scattered particles are correlated, as shown in Fig. 5.4. This helps to separate elastic events (close to the peak) from the background composed of radiative and inelastic processes. There is an ambiguity in an outgoing electron and muon for very low scattering angles, caused by similar momenta, that needs to be solved with proper particle identification (see Sec. 5.4.2). As the scattering angles are correlated to the differential cross-section, the measurement of the latter in the elastic scattering region will lead to the determination of the hadronic contribution to a_μ.

5.4 Experimental setup

The MUonE experiment will be located at the M2 muon beam in the CERN North Area, which provides muons with a momentum of $\sim$ 150 GeV/c and a rate of $\sim 1.3 \times 10^7 \mu/s$. The angles of the particles involved in the scattering of beam muons on a Beryllium target will be registered by the tracking system. The apparatus will consist of 40 stations, each of them being about 1 m long with

FIGURE 5.4: Correlation between scattering angles of electron and muon [119].

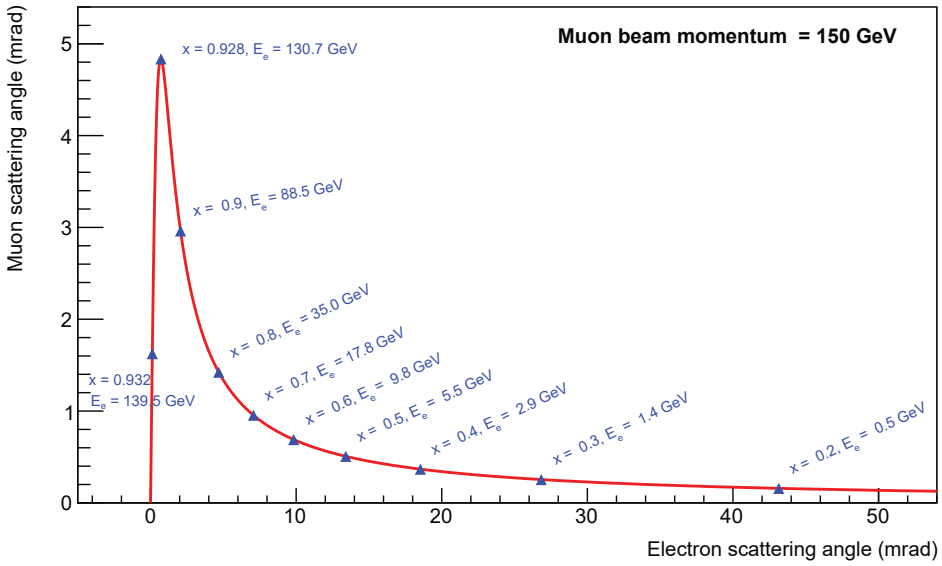
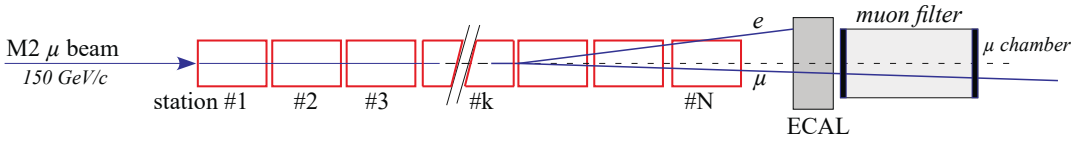


FIGURE 5.5: Schematic view of the MUonE experimental apparatus [22].



the 15 mm Beryllium target and tracking sensors. The use of the sequence of the target-tracking segments will allow the beam to be "reused" to increase the μe events statistics without the use of the Beryllium target of larger thickness that would lead to multiple scattering.

The tracking stations will be followed by the electromagnetic calorimeter (ECAL) used for the purpose of particle identification and measurement of electron energy. A downstream ECAL, muon filter will be installed. A schematic view of the apparatus is show in Fig. 5.5.

5.4.1 Tracking system

The tracking system is designed to precisely measure the scattering angles of the outgoing muon and electron, with respect to the incoming muon beam direction. Each tracking station consists of the target and three layers of the tracking modules, each built of two sensors. The tracking modules have 10×10 cm^2 active area that is sufficient to cover an expected angle range up to 30 mrad. A schematic view of the tracking modules is provided in Fig. 5.6. The drawing shows a scenario where tracking is performed between two targets.

FIGURE 5.6: Schematic view of one tracking station [22].

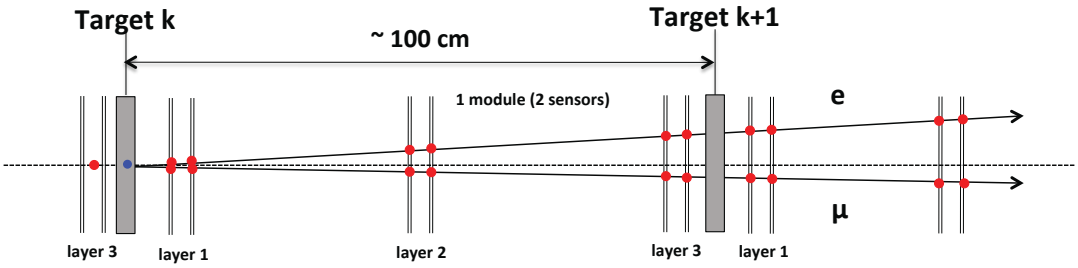
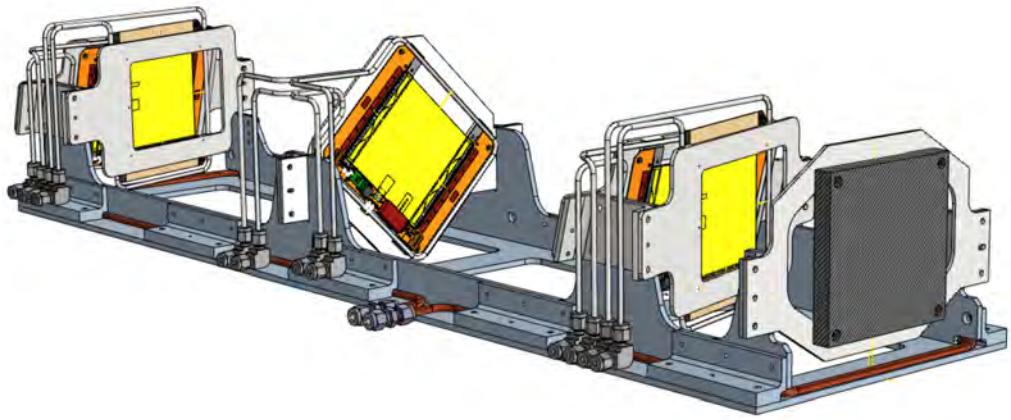


FIGURE 5.7: CAD image of the tracking station [137].



The tracking detector consists of three layers of pairs of detector planes. In each pair one sensitive layer measures x, and the second - y coordinate. As the distance between the layers is not insignificant, each measurement should be treated as a two-dimensional point with the second coordinate being z along the beam axis. Additionally, one layer is rotated around the beam axis (the so-called *stereo layer*) to solve the ambiguity in the track reconstruction. A CAD image of the single tracking module with the target and the tracking layers is presented in Fig. 5.7.

The state-of-the-art silicon strip sensors for the MUonE project are adopted from the CMS Tracker upgrade, characterized by a large active area sufficient to cover the full MUonE required acceptance, together with appropriate spatial resolution. They can also support a high readout rate of 40 MHz required for MUonE with their accompanying front-end electronics. The silicon sensors are 320 µm thick with n-in-p, being square sensors with an area of 10 cm $\times$ 10 cm. The strips are capacitively-coupled, with a pitch of 90 µm, and are segmented in two approximately 5 cm long strips. The DAQ is also adopted from the one developed for the CMS sensors for the HL-LHC upgrade.

FIGURE 5.8: Picture of the MUonE calorimeter.

5.4.2 Calorimeter and muon filter

To solve muon-electron ambiguity, downstream identifiers are planned to be installed They will consist of a calorimeter for electrons and a muon filter for muons. A homogenous electromagnetic calorimeter will be placed downstream all the tracker stations for particle identification, measurement of the electron energy and event selection. It will be built using lead tungstate ($PbWO_4$) crystals, similarly to the CMS electromagnetic calorimeter [138], as this material is characterized by a fast light scintillation emission time, good light yield and compact dimensions. The crystals will be arranged in an array, with a single crystal measuring 2.5×2.5 cm^2 of surface with a length of 23 cm (as shown in Fig. 5.8). Solid state sensors (SiPM and APD) will be used for readout. Muon-electron identification is ambiguous in the region of $\theta_e \leq 5$ mrad, which makes the required transverse dimensions of the calorimeter of the order of $1 \times 1\text{m}^2$. This will provide full containment of electrons with an energy $E \gtrsim 30$ GeV, with angular acceptance for $E \gtrsim 10$ GeV.

Because of their long lifetime, muons are able to penetrate through the whole detector. This makes the end of the detector the best location for installing a muon chamber, where all the other charged particles should be filtered. The chamber is designed to reject possible hadron contamination in the beam momentum.

5.4.3 Trigger system

The MUonE readout electronics will operate at a frequency of 40 MHz. This will result in an average rate of 1.25 of incoming muons during SPS spill duration. Such muons will traverse 40 sets of tracking stations following the Beryllium target where the elastic $\mu e \to \mu e$ scattering events will be searched for. The online selection is crucial for removing the majority of uninterested

events at 40 MHz while keeping the high efficiency of accepting genuine $\mu e \to \mu e$ scatterings. In most cases the stations will be traversed by a single or two muons from the incoming beam without any interaction. In the event of an interaction, the dominant sources of background will come from e^+e^- pair production and muon-nucleon collisions. In the MUonE experiment the trigger is assumed to consist of two separate software-based levels, the first of which is the processing of signals coming from the silicon microstrip sensors and will occur in FPGA, providing possible track segments composed of two consecutive hits in the sensors of one module - the so-called stubs. Next, the tracks will be formed out of the stubs. In the final step the track candidates will be checked to satisfy the conditions of coming from the $\mu e \to \mu e$ scattering. The crucial issue is to provide fast pattern recognition algorithms for such a track reconstruction and to check the efficiency of $\mu e \to \mu e$ scattering events and the reduction of the total output bandwidth. The required reduction down to 400 kHz has to be achieved. The pattern recognition algorithms will be first tested in real conditions during the planned Pilot Run in 2022-2023. The events accepted on the FPGA-based stage will then pass to the second level, based on the full event reconstruction. To overcome the bottleneck of limited disk/tape storage available at the offline stage, the high quality reconstruction and selection has to be moved from the offline to online stage. This way the experiments can profit from higher luminosity beams and extract high purity samples for the processes under study.

The use of machine learning techniques (described in Chapter 6) is foreseen at both FPGA-based trigger level and in the real-time event reconstruction.

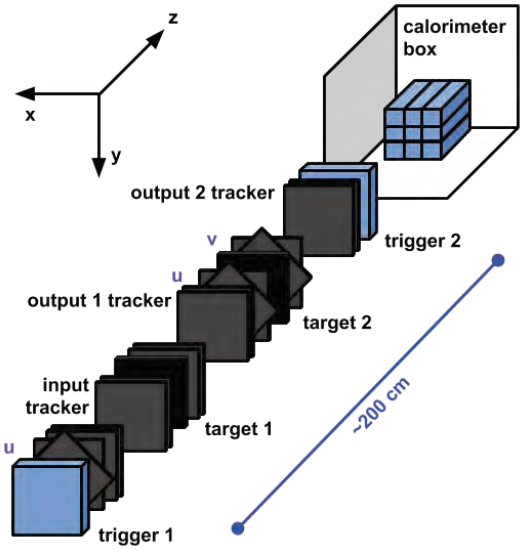
5.5 Testbeam in 2018

In 2018 a test run [21] was performed using high-intensity, high-emittance, high-momentum muons originating from the decays of the dumped pion beam used by the COMPASS experiment [139]. In this way a 187 GeV positive muon beam was obtained. The apparatus configuration for test beam is shown in Fig. 5.9.

The tracking system used silicon microstrip detectors, organized in 16 tracking planes. Each plane was equipped with a single-sided sensor with dimensions of $9.293 \times 9.293 \times 0.041$ cm, with 384 readout channels. Each of the layers measured one coordinate of the hit, x or y. Two *stereo-layers* rotated $\pm\frac{\pi}{4}$, providing the additional coordinates u and v. A graphite target, $10 \times 10 \times 0.8$ cm in size, was used. The calorimeter composed of BGO tampered crystals provided 15 mrad of angular acceptance.

The test beam period lasted for about 6 months, with datataking at the end of this time. After first filtering and applying a requirement of 6 hits upstream of the target, $\sim 2 \times 10^6$ events were selected. After the final requirements of the presence of an incoming track and at least two outgoing tracks, the number of the events was reduced to $\sim 94 \times 10^3$.

FIGURE 5.9: Schematic view of the apparatus used in the 2018 feasibility test [21].



5.5.1 Monte Carlo simulation

The 2018 test beam was accompanied by the production of a Monte Carlo sample corresponding to the apparatus configuration. Elastic μ-e scattering events were generated within the FairRoot framework [140] with the detector simulated in GEANT4. Simplification was introduced in the calorimeter modeling, using a single block instead of 9 separate crystals. The incoming muon beam was chosen to be a monoenergetic beam of 187 GeV, with the distributions of the x and y positions matching the ones measured during the test run. Events were generated using Leading Order (LO) calculations. In total, $\sim 100 \times 10^3$ events were generated.

5.5.2 Event reconstruction and selection

To reconstruct the events, a set of actions was performed in the order presented below.

Alignment

The process of alignment was performed for all the tracking layers, including stereo-layers. The position of the first layer was used as a reference, and all the consecutive layers were aligned using good quality reconstructed tracks with at least 10 hits assigned. The z position of a layer was taken from the measurements during a geometrical survey. The shift in the x-y plane and rotation around the z axis in relation to the first layer were determined in an iterative procedure. In each step, the layers were processed one-by-one: a track was refitted with the exclusion of the currently processed layer, a sum of residuals calculated for all tracks was minimized with respect to the shift and rotation of the current layer. The process was repeated until the changes were below

a given threshold. The resolutions of the individual layers were obtained from distributions of the residuals fitted with Gaussian distribution. The calculated resolutions were in a range from 15 to 37 µm, well below 40 µm specified for the used silicon sensors.

Track reconstruction

In the first step, the pairs were constructed from all hits in the x and y layers separately and 2D track finding was performed: a 2D line in x-z or y-z was determined for each pair of the hits. For each line, all hits within a window corresponding to 10 times the sensor resolution were collected and, after removing the outliers, a fit was performed. The acquired set of 2D track candidates was sorted by the number of hits and χ^2 of the fit. In the clone killing procedure the best tracks with the unique combinations of hits were accepted, with at least 3 hits in each projection required. Track candidates in 3D were combined from all pairs of the track candidates in x-z and y-z projections and an iterative fitting procedure was applied. In each iteration, the hits outside the 5σ radius from the fitted line were removed until there were no outliers.

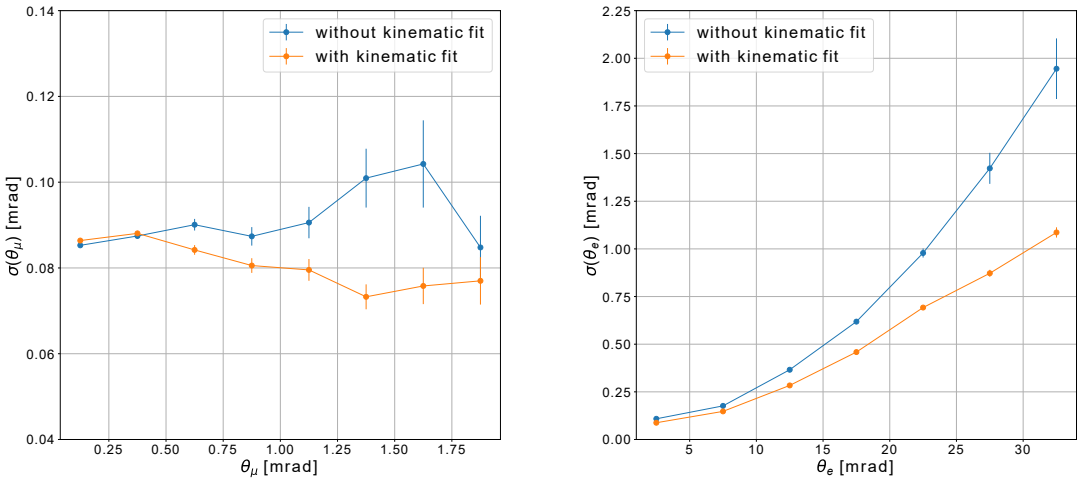
In the last step, the cloned tracks were removed. The tracks were sorted by the number of hits and χ^2/NDF. A track with the highest number of hits was chosen, and if more tracks had the same number of hits, the track was chosen with the χ^2/NDF criterion. If the track was chosen, the hits used by this track were not considered later. If the track had at least 3 hits in x and 3 hits in y, it was accepted and the procedure was repeated for the next track from the list. Thus a collection of the tracks using unique hits was constructed.

Reconstruction of the μ-e scattering event

To reconstruct the μ-e scattering events, all possible pairs of reconstructed tracks were analyzed to find the ones intersecting in the target. When a matching pair was found, a compatible incoming muon track was searched for. For the complete triplet of tracks, a dedicated vertex fit was performed. The momentum of the tracks was estimated to take into account multiple scattering (MS). The events with scattering angles < 2.5 mrad were assumed not to include MS. If a scattering angle was > 2.5 mrad, a track with a larger angle was assumed to be an electron. The scattering angle of the electron can be used to determine its momentum. The muons were assumed not to experience significant MS. The predicted MS was used to calculate the uncertainties of the hits assigned to the electron track. A kinematic fit of the vertex was performed, assuming the z position in the middle of the target. The fit was based on a constrained least square method, using the 3-d line slopes of the three tracks as well as the x and y vertex positions as free parameters. The total χ^2 used for minimization is the sum of the χ^2 contributions from all hits of the three tracks. This defines the total vertex χ^2/NDF, referred to as χ^2_{vtx}.

The angular resolution was calculated using MC simulation as the σ of the Gaussian fit to the difference between the true angle and the reconstructed angle. Resolution as a function of the true emission angle is presented in Fig. 5.10.

FIGURE 5.10: Angular resolution as a function of the scattering angle for simulated events (right: muon, left: electron) with and without dedicated kinematic fit of the scattering vertex [21].



For the muon it is relatively flat, but for the electron it shows a strong dependence on the angle, mainly as a result of multiple scattering.

Selection of μ-e scattering events

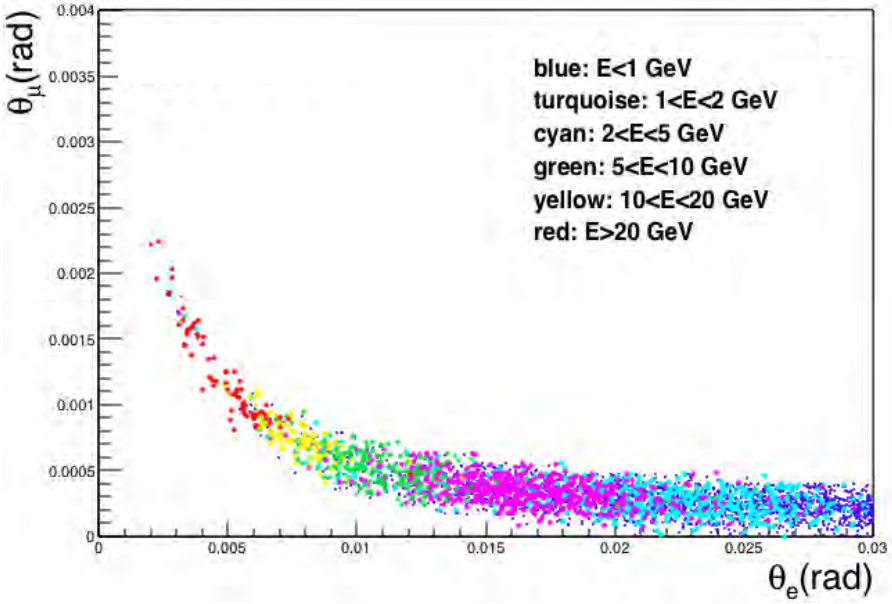
Before the analysis, two cuts were applied on the reconstructed μ-e events. The cut of $\chi^2_{vtx} < 10$ was the result of an analysis of the distribution of χ^2_{vtx}. The second cut puts a limit on the electron emission angle $\theta_e < 30$ to suppress low energy electrons. This was motivated by the fact that the cut on $E_e > 1$ GeV implied an angular cut at $\theta_e \approx 35$ mrad.

5.5.3 Results and conclusions

A precise measurement of the scattering angles of the outgoing particles in the $\mu e \to \mu e$ process is required by the MUonE experiment to evaluate the running fine-structure constant that will be used to calculate the hadronic contribution to the anomalous magnetic moment of a muon. The measurements performed during the test run provided a good quality sample of events. The kinematic correlation of the scattering angles is shown in Fig. 5.11. An additional cut on the minimum angular distance of the measured event to the expected theoretical kinematic curve, D_θ, calculated for a given incoming muon beam energy, was performed to reduce the background. Such a variable estimated the elasticity of a reconstructed event [141] and the cut on D_θ was optimized based on the simulation.

The main aim of the testbeam was to check An ability to select a clean sample of elastic scattering events. This proved to be possible, even with a resolution worse than planned for the final experiment, corroborating the correctness of the proposed experiment design. The study emphasized a necessity for an adequate calorimeter in order to understand the electrons emmited in the range of a few GeV and to determine a behaviour of the background. Another

FIGURE 5.11: Kinematical correlation of the outgoing muon and electron. Colors represent the energy deposit in calorimeter [21].



important aspect pointed out is an upgrade of the GEANT4 software, which will improve the simulation of the muon pair-production interaction $\mu \to \mu ee$ with the accurate angular distribution of the electrons. The use of GEANT4 version 10.7 has already been implemented in FAIRROOT and is undergoing testing.

It is worth noticing that the author of this thesis was actively involved in the development of the track and μ-e elastic event reconstruction algorithms as well as the analysis of the data collected in the MUonE testbeam in 2018.

Chapter 6

Machine learning based reconstruction algorithms

High energy physics experiments have a growing need for fast and efficient methods of particle track reconstruction. Sizes of collected data samples have reached the level that makes it not feasible to store them unprocessed, forcing more computations to be performed online (without storing the data on disks). A move from hardware triggers accompanied by software High Level Triggers to fully-software solutions (see also Sec. 3.4.1) puts high requirements on the software to produce high quality samples in very strict time constraints. To reduce the data footprint to an acceptable level, the performed reconstruction must have the quality of algorithms used previously in offline mode, but without time limits. To enable this, novel techniques need to replace older, central processing unit (CPU) based algorithms. Machine learning brings to the table high performance in pattern recognition tasks, together with powerful acceleration gained using graphics processing units. Although such techniques have not yet been applied on a large scale in any high energy physics experiments, they are being intensively developed and planned to be employed in the near future.

This chapter describes the logic and implementation of the track reconstruction algorithm based on machine learning techniques, designed as a proof-of-concept using a simulated data sample from the MUonE experiment.

6.1 AI, ML, ANN, DNN

There is a plethora of terms related to the development of systems that make decisions in a seemingly intelligent manner. The broadest one is artificial intelligence (AI) that does not describe any specific applications, but rather focuses on ideas and processes. AI studies intelligent entities that perceive their environment and perform autonomous actions to achieve goals. In a more data science oriented definition, it can be described as an ability to interpret and learn from external data to achieve a specific goal through flexible adaptation [142].

Machine learning (ML) is an umbrella term for a broad collection of techniques in which algorithms improve themselves (!) according to changing requirements and diverse phenomena. In the process of learning, called *training*, a *model* is created. The model itself may hold one of many structures that can be later used to make predictions, e.g.:

- artificial neural network (ANN) [143], further described in Sec. 6.1.1;
- decision tree [144], a tree-like structure, where acquired knowledge is described in decision nodes connected by branches and conclusions in the leaves;
- bayesian network [145], a directed acyclic graph connecting a set of variables, and their conditional dependencies;
- genetic algorithm [146], in which a solution is stored in a way analogous to a genotype and undergoes processes of natural selection, being subjected to crossovers and mutations.

6.1.1 Artificial neural networks

Artificial neural networks, or simply neural networks, date back to as far as 1944 [147]; however, no training mechanism for the network was presented. Supervised learning (see Sec. 6.1.2) was introduced in the 1950s and 1960s [148–150], followed by unsupervised learning (see Sec. 6.1.2) [151–153].

Artificial neurons

The structure of ANN is inspired by the biological neural network of the human brain, also adopting the nomenclature. A network contains artificial neurons (perceptrons) that are connected with each other, analogue to the synapses. Each input connection of a neuron has an assigned weight. A perceptron multiplies each of real-valued inputs by the corresponding weight and sums all the values. Then the *activation function* (described in the next section) is calculated on the sum and the response is sent to the following neurons.

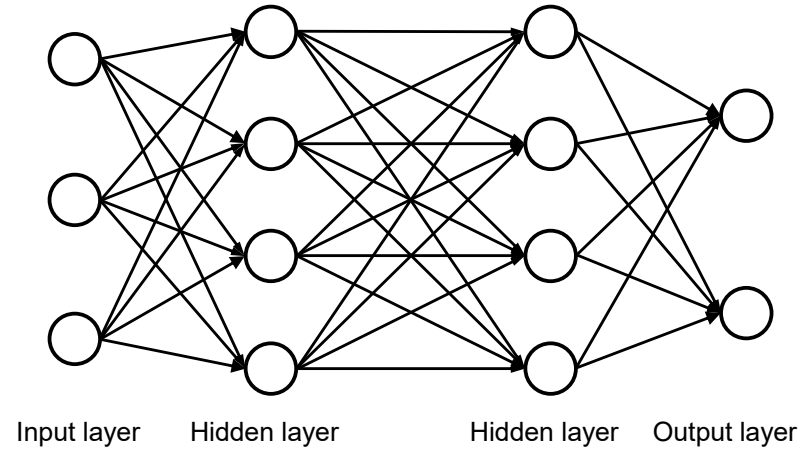
Neurons are organized in layers. In *feedforward neural networks* each neuron is connected with neurons in adjacent layers; by contrast, loops are present in *recurrent neural networks*. The first layer is an *input layer*, the last one is tan *output layer*, and the layers in between are called *hidden layers*. A network with one hidden layer is called a *shallow network*, while a *deep network* has multiple hidden layers. An example of a feedforward deep neural network with two fully connected hidden layers is shown in Fig. 6.1.

Activation function

The activation function, also called the *transfer function*, is a function that defines the answer of the neuron for the given input [154]. The most popular is the family of sigmoid functions [155]. Mostly they are bounded, differentiable, have non-negative derivative and one inflection point. Generally their graph is "S"-shaped. An example of a sigmoid function is the *logistic function* described with the formula:

$$S(x) = \frac{1}{1+e^x} = \frac{e^x}{e^x+1} = 1 - S(-x). \tag{6.1}$$

FIGURE 6.1: Schematic view of a fully-connected deep neural network.



A counterpart to the logistic function, like sigmoid, in multiple dimensions is the *Softmax* function [156]:

$$f_i(x) = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}}, \tag{6.2}$$

where $i = 1, 2, \ldots, n$ denotes the i-th coordinate in the n-dimensional space. The softmax function is usually applied in the output layers of multi-class models, as it provides probabilities for each class (returned values vary from 0 to 1 and sum up to 1). Another example of a popular sigmoid function is the *hyperbolic tangent*:

$$f(x) = \tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{e^{2x} - 1}{e^{2x} + 1}. \tag{6.3}$$

Another function used as an activation function is *Softsign* [157] that is mostly used in regression computation problems [158] and speech recognition [159]. It is defined as:

$$f(x) = \frac{x}{|x| + 1}. \tag{6.4}$$

One of the most commonly used activation functions is a *rectifier* or a Rectified Linear Unit (ReLU) that gained popularity as an alternative with respect to a sigmoid for deep networks, at the same time being more biologically justified [160]. It is defined as a positive part of the argument:

$$f(x) = x^+ = max(0, x). \tag{6.5}$$

The function is highly successful thanks to its low computational cost and allows for sparse representations (with random initialization with symmetrical

distribution, half of the neurons will be inactive at the start). Potential problems coming from a nonexistent derivative in 0 do not seem to negatively affect practical applications [161].

The popularity of the ReLU led to the appearance of variants of this function. *Leaky ReLU* (LReLU) [162] forces updates on all neurons by a small negative slope for negative arguments:

$$f(x) = \begin{cases} x, \text{ if } x > 0 \\ \alpha x, \text{ if } x \leq 0 \end{cases}, \tag{6.6}$$

where $\alpha \leqslant 0$. This function has a non-zero derivative except 0 that can be controlled by α, with relatively small $|\alpha|$ in practical applications. Compared to ReLU, a network with LReLU will lose its sparsity.

6.1.2 Training

The main objective of the training process is to generalize the model, so it can perform a task using data not seen before. There are several approaches to conducting the training process, the three main being *supervised training*, *unsupervised training* and *reinforcement training.*

Supervised training

In the supervised training, the so-called *training dataset* is provided in which every *feature vector* (a set of inputs) is *labeled*. The label plays a role of the *ground truth*, supplying the desired outcome during the training. The training is an iterative process: in each step an output from the model is compared with the label by the *loss function* (also called *cost function*) that provides a single value depicting how wrong[1] the response was - what constitutes its cost.

A widely used algorithm for training of feedforward networks is the *back-propagation* [154]. For each pair of the feature vector and the label, a gradient of the loss function with respect to the weights of the network is calculated. In each step weights are updated to minimize the loss function, using the gradient decent method. For high-dimensional networks, a related method of the stochastic gradient decent can be used for improved computational cost - in this case approximations are calculated instead of the actual gradient.

Unsupervised training

In unsupervised training the dataset is not labeled. The network is expected to find patterns in the data using networks internal representation. This process is similar to self-learning in the case of people. This self-organization manifests itself in capturing patterns as probability density distributions.

[1]The loss function itself is often constructed to grade how well the model performed, but the value is eventually negated for the sake of use in minimization.

Reinforcement training

In reinforcement training, intelligent agents are scored for their actions and try to maximize the total score. This method of learning can be used in the situations where there is no mathematical model of the problem. This kind of training is often utilized in tasks such as autonomous driving, artificial intelligence in video games, and swarm intelligence. The environment for the intelligent agents is often presented in the form of a Markov decision process [163] that provides a framework to solve decision problems where the outcome may be not entirely controlled.

6.2 Two-dimensional machine learning based event reconstruction for MUonE

In the MUonE experiment the algorithms of track finding based on machine learning techniques (Deep Neural Networks - DNN) are being developed and tested in order to speed up the reconstruction process. This may lead to the significant acceleration of the execution of pattern recognition algorithms at both FPGA-based trigger level as well as in the real-time event reconstruction algorithms. Moreover, it was demonstrated that the use of DNN techniques may also significantly improve both the reconstruction efficiency and the precision of measuring parameters that are crucial for final measurement.

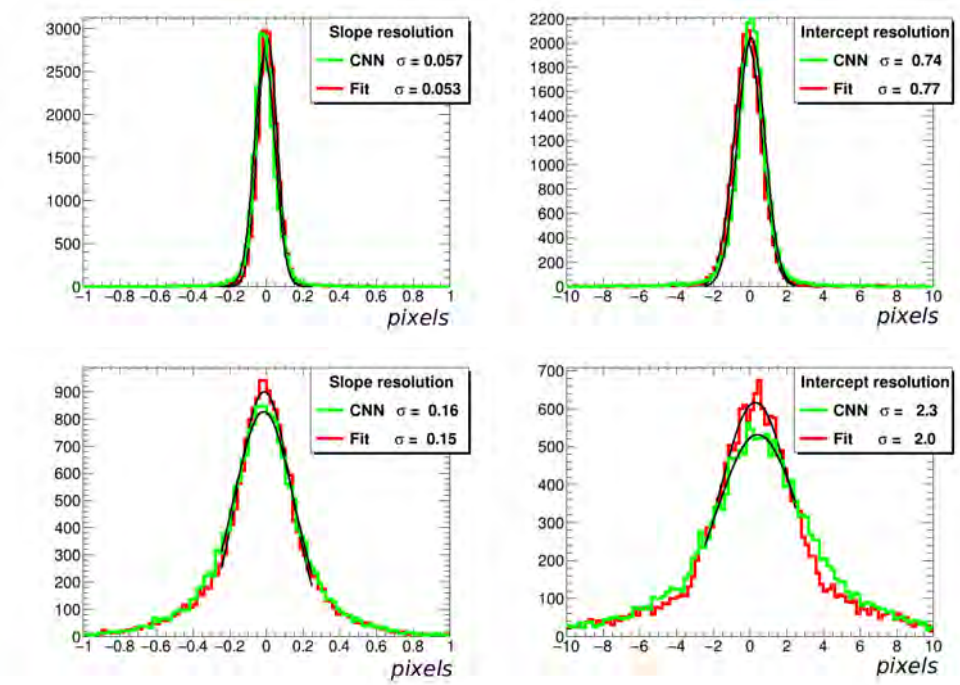
The first application of the machine learning techniques in the context of the MUonE experiment [164] used a convolutional neural network (CNN) [165].

The CNN technique can be used in both steps: track finding and event classification and selection. In the convolutional layer approach the layer's parameters consist of a set of filters called kernels that have a relatively small receptive field. During the forward pass, each filter is convolved across the width and height of the input volume, computing the dot product between the entries of the filter and the input, and producing an activation map of that filter. As a result, the network creates a set of filters that activate when they detect some specific type of feature at some spatial position in the input. This allows one single filter to find given features in the entire input image, even having a small receptive field.

A toy-model was used to generate training and testing samples of 20×10^3 events each, corresponding to the simulated μ-e elastic scattering signal events for the MUonE test beam 2018 (see Sec. 5.5). The events were represented with two-dimensional 28×28 pixel images containing one or more linear tracks and an optional noise.

The software library KERAS [166] with TensorFlow [167] backend was chosen for network implementation. A convolutional network was trained to return the parameters of a track (slopes and intercepts) in response to an image. The network consisted of two convolutional layers with a 3x3 convolution window followed by the MaxPooling layer and then by another two convolutional layers. The single dropout layer suppressed the overtraining, and the final regression was performed in the dense layer with 1024 nodes. In total the neural network had over 2 million trainable parameters. In the case of multi-track events,

FIGURE 6.2: Comparison of the distributions of a difference between the reconstructed and true track parameters: slope (left) and intercept (right), between CNN-based and classical track reconstruction. Noise level at 10% (top) and 30% (bottom).



long/short-term memory (LSTM) was introduced, based on the research by the HEP.TrkX project described in [168, 169]. A random noise and pixel inefficiency were added to make the model more realistic. The pixel efficiency was set to 70% and the noise varied in the range between 0 and 30%. The noise level was defined as the probability of a single pixel not belonging to the track to generate a false signal. Figure 6.2 shows the distribution of a difference between the reconstructed and true track parameters (slope and intercept) for events with 10% noise and with high, 30% noise for three-track events, using CNN-based as well as classical track reconstruction (the same as the one used in the MUonE 2018 testbeam data analysis). The Gaussian function was fitted to these distributions and its width was taken as a measure of the resolution.

The results from the neural network were used to find hits close to track candidates, which were then reconstructed using the linear robust fit [170].

The two-dimensional neural network algorithm was proven to be successful when applied to the toy-model corresponding to the simulated sample of MUonE testbeam 2018 data. It served as the base of the method that needed to be extended to the realistic three-dimensional case (see the section below).

6.3 Three-dimensional machine learning based event reconstruction for MUonE

A continuation of the work described in the previous section is the application of machine learning techniques to the full three-dimensional event reconstruction [26], both for the needs of the so-called Pilot Run of the MUonE experiment planned for 2022-2023 as well as for the final detector configuration.

The main goal of the designed neural network was first to reproduce slopes and offsets of two outgoing 3D tracks corresponding to two outgoing μ-e elastic scattering signal particles as the response to the set of hit coordinates.

6.3.1 Learning dataset

As the silicon sensors used in the 2018 testbeam (see Sec. 5.5) as well as the sensors planned to be used in the final detector configuration (see Sec. 5.4) can provide only two-dimensional information of the hit position, i.e. (x, z) or (y, z), it is not straightforward to expand the machine learning based track reconstruction to the full three-dimensional case. This is why the fundamental assumption was made that the neural network should be able to learn how to recognize (x, z), (y, z) and also stereo hits, and how to properly treat them in the pattern recognition as well as track reconstruction process. Such an assumption was positively verified using the DNN technique and then successfully applied as a pattern recognition procedure.

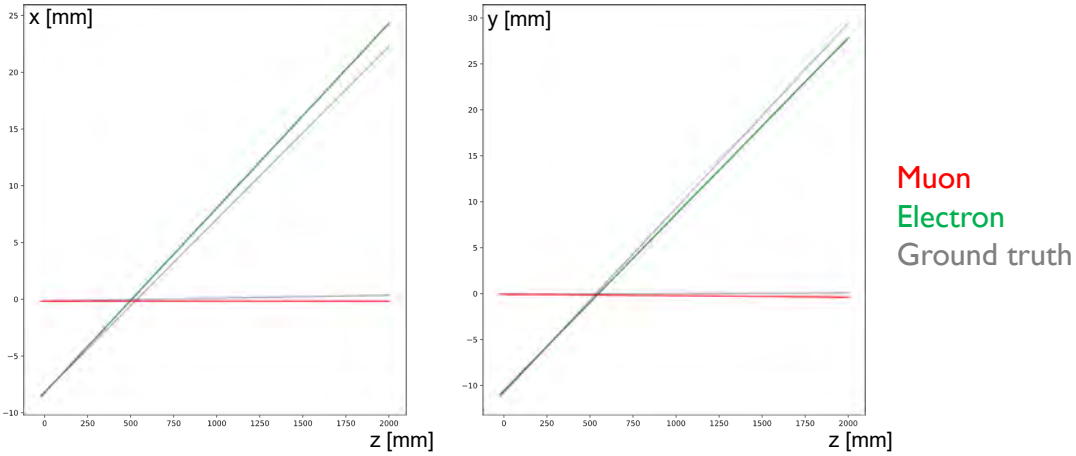
The dataset used for model training was generated with a leading-order event generator with the detector simulation in GEANT4. The simulated sample corresponded to the 2018 MUonE testbeam setup described in Sec. 5.5 and consisted of about 100×10^3 events. Each event was described with a set of 2D hit coordinates in x-z and y-z planes, where the z coordinates were related to the detector location and could be used to interpret the second coordinate as x or y. As the dataset was generated, ground truth was also provided in the form of slope and offset of each track. The samples were split between training and validation datasets in the 3:1 ratio.

6.3.2 Artificial neural network

PyTorch [171] was selected as the machine learning framework, as it is well documented and gained popularity in the scientific community. It offers a complete set of tools to handle data, training process and inference. During the training and inference, graphics processing units can be utilized to significantly accelerate computations. Those processes involve numerous matrix operations, and graphics cards are designed for this task having orders of magnitude more dedicated computing cores than CPUs.

The input vector was reduced from the verbatim event description by removing hits corresponding to the incoming muon and all z coordinates, as the latter were identical in all events. The vector consisted of 20 values representing hits registered by each silicon strip sensor, without a distinction between x, y

FIGURE 6.3: Example of the tracks reconstructed by the DNN, before applying further steps of the reconstruction algorithm. Deviation from the ground truth shown in grey is visible.



and *stereo* layers (see Sec. 5.4.1). The points were ordered by the ascending z coordinate.

The output vector contained 8 values, the slopes of two outgoing tracks in both x-z and y-z projections and corresponding offset values. This format was inherited from the ground truth for the ease of comparison.

The neural network was built using 4 fully connected layers, 1000 neurons each. This size was chosen as an acceptable balance between computation time and quality. During the development, the frequent retraining of the model was performed, putting constraints on the training time that would not be the case in the final application. As this solution is a proof-of-concept, no profound optimization study for the network size and shape was conducted as of that time.

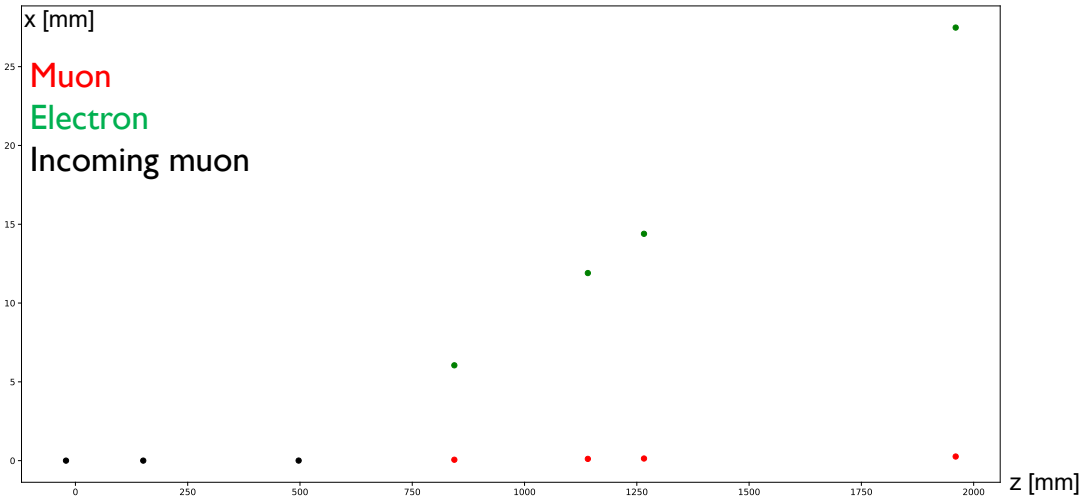
For the loss function, *MSELoss* (Mean Squared Error Loss) [172] from the PyTorch package was used. As the name suggests, it calculates the mean squared error between the output vector and the ground truth, which is minimized in the process of training.

The ReLU function (Eq. 6.5) was chosen as the activation function because of its good fit to use with deep networks. The use of the simplest of its generalized forms (Eq. 6.6) did not bring essential results.

6.3.3 Reconstruction algorithm

The predictions from DNN were visually compared with the ground truth. The tracks were relatively closely reconstructed (see Fig. 6.3), but did not provide a precision required by the experiment. This is why it was decided to use the DNN only in the pattern recognition procedure, as it is the most CPU time-consuming process as compared to the relatively fast linear track fitting procedure.

FIGURE 6.4: Example of the collections of hits corresponding to the μ-e elastic scattering signal tracks constructed based on the DNN-predicted track candidates. Points represent the hits in x-z projection with colours related to the true track.



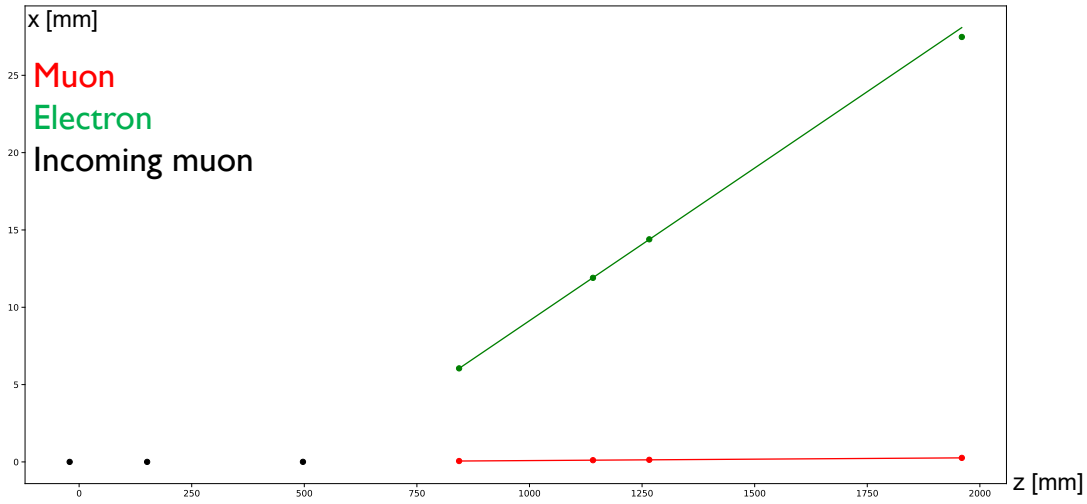
Deep neural network based pattern recognition

The pattern recognition procedure employed the DNN machine learning model where all the hits from the μ-e elastic scattering signal event were used to construct the three-dimensional DNN-predicted track candidates in a single go. In the next step the hits were assigned to the closest DNN-predicted track. In this way the collections of hits corresponding to the μ-e elastic scattering signal tracks were created in x-z and y-z projections (as the hits have only (x, z) or (y, z) coordinates), being then fitted using the RANSAC method (see the following section). An example event in x-z projection with colour-coded assignments is shown in Fig. 6.4.

Two-dimensional linear fit

The collections of hits corresponding to the μ-e elastic scattering signal tracks constructed based on the three-dimensional DNN-predicted track candidates were fitted using the RANSAC method separately in x-z and y-z projections because of the absence of the third coordinate. In the RANSAC iterative algorithm (*Random Sample Consensus*) [173], the minimal number of points required to describe a desired shape was randomly selected. For a line, two points were selected and the parameters of a line calculated. After that, all remaining points were classified as inliers or outliers depending on the distance between the point and the line being smaller or larger than the selected error threshold, respectively. A set of inliers is called the consensus set. The algorithm was repeated until the consensus set reached the required number of inliers. The parameters of this method included the maximum allowed iterations, error threshold for inlier qualification, and the required size of the consensus set. This method is capable of dealing with datasets with a large number of outliers if there are enough inliers.

FIGURE 6.5: Example of the result of the linear fit (solid lines) in x-z projection for an outgoing muon and electron from the μ-e elastic scattering signal event. The points represent the hits in x-z projection with colours related to the origin track.



During the fitting procedure, the residual threshold parameter (maximum residual to classify a point as an inlier) was set to 0.03, while the other parameters stayed on default values in Scikit-learn package implementation [174] (maximum 100 iterations with no explicit limit on minimum samples). As a result of such a linear fit, two 2D lines in x-z and y-z projections were established for each track. As an example, Fig. 6.5 shows the result of the linear fit in x-z projection for an outgoing muon and electron from the μ-e elastic scattering signal event.

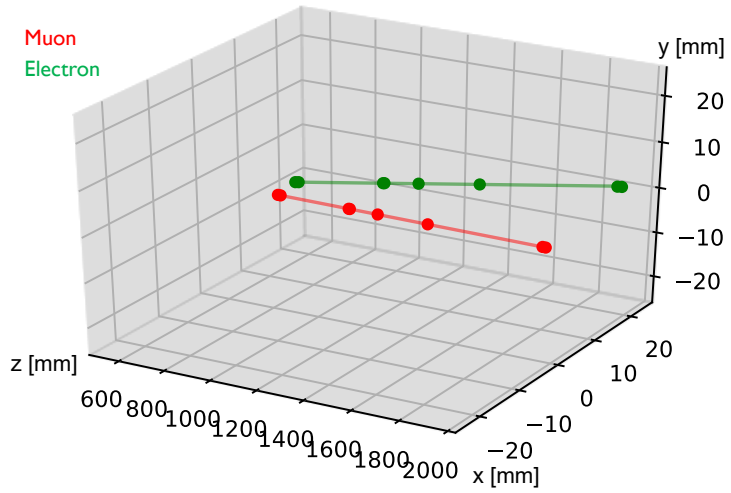
Final 3D track fit

As for every track two lines in x-z and y-z projections were found (see the fitting procedure described above), it was possible to calculate the missing coordinate for every hit from the collection of hits constructed based on the three-dimensional DNN-predicted track candidate. This provided the collections of 3D hits that were the basis for the reconstruction of three-dimensional tracks using a 3D linear RANSAC fit. An example of the 3D fit results is presented in Fig. 6.6. Double-points are visible where the detector layers for x-z and y-z are close to each other.

6.3.4 Results

The results of the track reconstruction algorithm described in Sec. 6.3.3 were used to determine the resolution in the track slopes, and to compare directly the results from the DNN-based track reconstruction and the classical algorithm (the same as used in the MUonE 2018 testbeam data analysis). The efficiencies understood as the percentages of the correctly reconstructed tracks, calculated

FIGURE 6.6: Example of the result of the 3D linear track fit (solid lines - red for muon and green for electron) for an outgoing muon and electron from the μ-e elastic scattering signal event. The points represent the hits with colours related to the origin track (red for muon and green for electron).



Particle	DNN based	Classical
Muon	100%	99.98%
Electron	99.66%	99.38%

TABLE 6.1: Efficiences of reconstruction algorithms.

for the threshold on the track slope at the level of 1×10^{-2}, are presented in Table 6.1. They are comparable for DNN-based and classical track reconstruction and slightly better for the DNN.

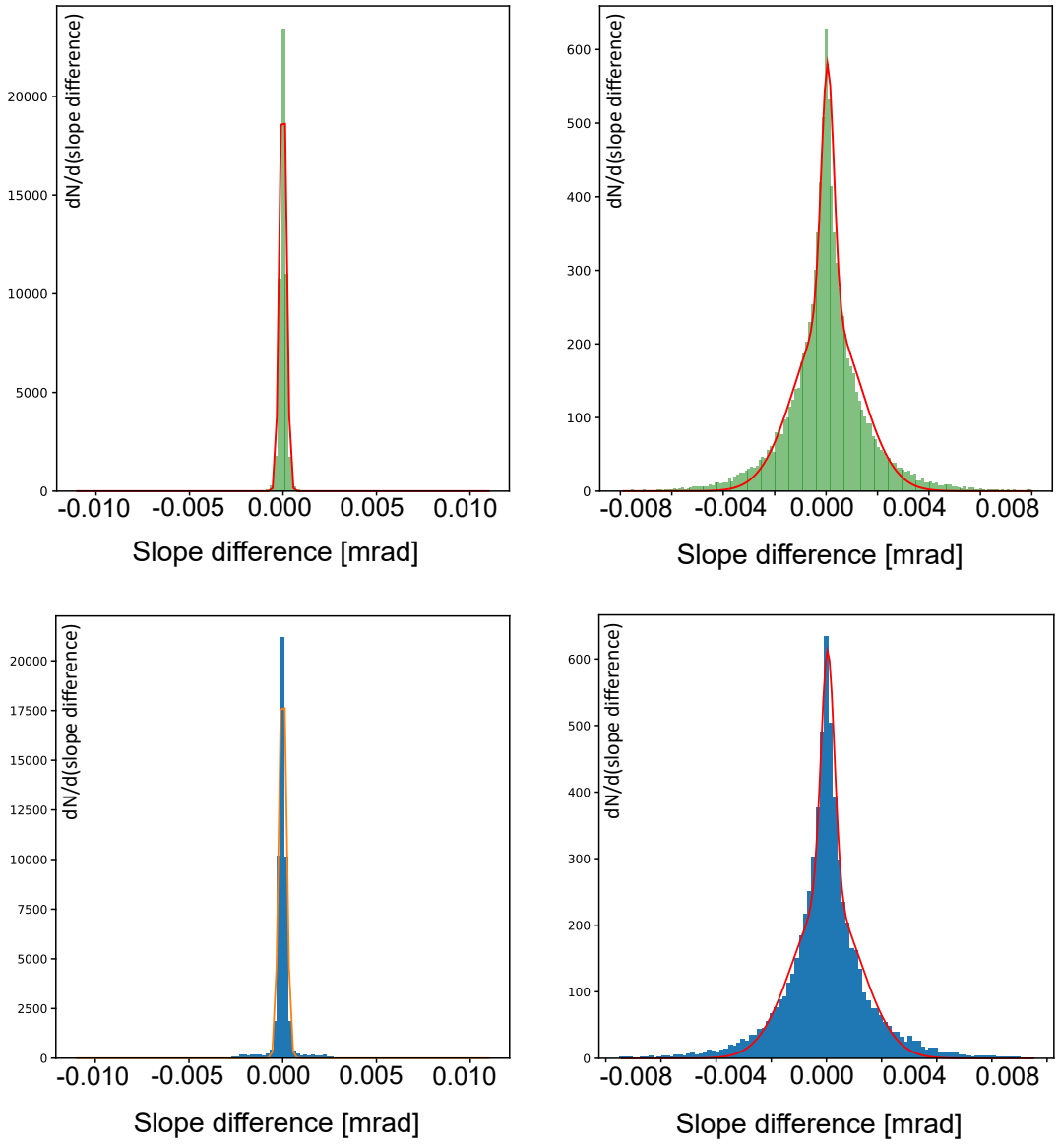
The distributions of the slope difference between the reconstructed track and the ground truth are presented in Fig. 6.7, with the muons fitted with single Gaussian, and the electrons with double Gaussian, as they are much more influenced by multiple scattering. They are compared with the results from the classical algorithm. The resolutions in the track slopes (values of standard deviations from the fits to the distributions in Fig. 6.7) are summarized in Table 6.2.

In general, the results of the DNN-based algorithm are comparable to the

Particle	DNN based	Classical
Muon	$\sigma = 0.000018$ mrad	$\sigma = 0.000019$ mrad
Electron	$\sigma_1 = 1.290$ mrad, $\sigma_2 = 0.245$ mrad	$\sigma_1 = 1.230$ mrad, $\sigma_2 = 0.244$ mrad

TABLE 6.2: Slope resolutions for an outgoing muon and electron.

FIGURE 6.7: Distributions of slope difference of reconstructed tracks (left for muons, right for electrons) in relation to the MC truth, for DNN-based algorithm (upper plots) and classical reconstruction (bottom plots).



classical reconstruction, although they are slightly better in some aspects. This concept has significant potential for progress and will be the basis for further research, which will be outlined in the next section.

6.4 Outlook

As mentioned above, the presented method is slightly better than the classical reconstruction and establishes a solid base for further development. It was possible to demonstrate that the use of DNN techniques in the event reconstruction may not only result in a significant acceleration of the execution of pattern recognition algorithms and event reconstruction algorithms, but also may improve both the reconstruction efficiency and the precision of measuring the parameters that are crucial for physics analysis. The most promising way to proceed would be to replace the DNN with the *Graph Neural Network* (GNN), which is getting much interest in the particle physics community [175–177]. GNN is a set of trainable functions which operate on graphs. The procedure is to use graphs instead of vectors or matrices as input data, and therefore – in the case of MUonE – the hit positions are treated as appropriate nodes. The output data are graphs with edges connecting hit points, where each edge has a weight being a probability to belong to a given track. GNN ensures the inductive bias, reduction of the number of parameters, deeper elaborated loss function, and above all much more natural data representation. The networks of this type are significantly more elastic in the case of events with missing or additional hits, when the static size of the input layer of the DNN may be a problem in real-world events, that are not as perfect as the simulated ones. The machine learning based track reconstruction procedure described in this chapter is assumed to be further developed and tested for the pattern recognition in the Pilot Run of the MUonE experiment that has started in 2022 and will continue in 2023, as well as for the final detector configuration.

New event reconstruction methods developed for the MUonE experiment based on the machine learning methods implemented on parallel GPU processing may become a standard approach in the future high energy physics experiments, facing an enormously tight execution time imposed by a fully-software trigger system and achieving a maximum possible event reconstruction efficiency and precision. Owing to these methods, the size of data, expected to increase quickly in the future experiments, will be efficiently reduced and also the statistical power of the final physics measurement will be maximized. Such algorithms, providing a significant reduction of the execution time as well as improvement in the efficiency and precision of measurements, will also be able to be applied in other future high-energy physics experiments.

Chapter 7

Summary and conclusions

In the present thesis a study of the Bose-Einstein correlations for triplets of the same-sign charged pions in proton-proton collisions with an energy of $\sqrt{s} = 7$ TeV recorded by the LHCb experiment has been described. For the first time, such a measurement was performed in the forward region at the LHC energies, and the results were interpreted within the core-halo model that provides an insight into the character of the hadron emission in terms of its coherency. The parameters of the core-halo model were determined in bins of reconstructed charged particle multiplicity, but unfolded multiplicity ranges were provided as well. The correlation functions were constructed with the event-mixing based reference sample, and fitted using the Lévy-type parametrization with the index of stability equal to unity. The double ratio technique was employed to correct for nonfemtoscopic background, together with the Gamow penetration factor utilized to account for electromagnetic final state interactions that can be factorized using the Riverside method.

The results of the analysis, especially in the partial coherence parameter p_c, which is increasing with the charged particle multiplicity, suggest a possibility of coherent emission of pions. At the same time, the κ_3 parameter which, by varying from unity, should also indicate a partial coherence, seems to be consistent with unity within statistical and systematic uncertainties. Such a behaviour of the κ_3 parameter is similar to the trend observed by the PHENIX experiment in gold-gold collisions in bins of transverse mass, but that analysis does not present the results in bins of particle multiplicity, thus a direct comparison is impossible. Further conclusions in this field are limited because of relatively large statistical and systematic uncertainties in the f_c and λ_3 parameters.

The results of the presented analysis provide a direct insight into the hadron creation processes from the strongly interacting matter formed after the particle collision, testing the limits of the core-halo model with a thermalized core. The results have also the potential to help to develop phenomenological models of multiparticle production providing an exclusionary input in the unique acceptance of the LHCb detector. The presented analysis demonstrates the potential to be further extended to higher proton-proton collision energies available at LHCb, the collision systems of pPb and PbPb, as well as to the bins of transverse momentum or transverse mass of the pair as well as bins of pseudorapidity, what would also help to directly compare the results with other experiments.

The second part of this dissertation is dedicated to the application of machine learning techniques to the track reconstruction for high energy physics experiments. With growing amounts of data captured by new and upgraded

experiments, storing of it all may not be feasible. Events need to be reconstructed and analyzed in real time to select and save only the interesting ones. This process requires new approaches that can solve this problem with the performance much higher than ever before.

The proof-of-concept track reconstruction algorithm based on the Deep Neural Network was elaborated and verified on the simulated sample corresponding to the 2018 test beam data of the MUonE experiment. As this experiment is dedicated to the search for signs of New Physics phenomena by measuring hadronic contribution to the anomalous muon magnetic moment, precise track reconstruction is crucial to ensure efficient online reduction of data and to maximize the statistical power of the final physics measurement. Such an algorithm is able to provide a significant reduction of the execution time and improve the efficiency and precision of the measurement. The preliminary results of the DNN based algorithm are slightly better than the classical reconstruction, achieving at the same time enormous reduction of the execution time for the pattern recognition process. The presented results have shown that machine learning techniques have great potential to be applied to track reconstruction. They have proved to be a viable option for the track reconstruction task and are planned to be further explored, especially by applying GNN machine learning models and by expanding the analysis with the background events. The algorithm developed is planned to be tested in the Pilot Run of the MUonE experiment in 2022-2023 as well as in the final detector configuration.

Bibliography

[1] G. Goldhaber et al. "Influence of Bose-Einstein Statistics on the Antiproton-Proton Annihilation Process". In: *Phys. Rev.* 120 (1 1960), pp. 300–312. DOI: 10.1103/PhysRev.120.300. URL: https://link.aps.org/doi/10.1103/PhysRev.120.300.

[2] [ALICE Collaboration] K. Aamodt et al. "Femtoscopy of *pp* collisions at $\sqrt{s} = 0.9$ and 7 TeV at the LHC with two-pion Bose-Einstein correlations". In: *Phys. Rev. D* 84 (11 2011), p. 112004. DOI: 10.1103/PhysRevD.84.112004. URL: https://link.aps.org/doi/10.1103/PhysRevD.84.112004.

[3] [ATLAS Collaboration] G. Aad et al. "Two-particle Bose-Einstein correlations in pp collisions at $\sqrt{\mathbf{s}} = \mathrm{s} = 0.9$ and 7 TeV measured with the ATLAS detector". In: *The European Physical Journal C* 75.10 (2015). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-015-3644-x. URL: http://dx.doi.org/10.1140/epjc/s10052-015-3644-x.

[4] [CMS Collaboration] V. Khachatryan et al. "First Measurement of Bose-Einstein Correlations in Proton-Proton Collisions at $\sqrt{s} = 0.9$ and 2.36 TeV at the LHC". In: *Phys. Rev. Lett.* 105 (3 2010), p. 032001. DOI: 10.1103/PhysRevLett.105.032001. URL: https://link.aps.org/doi/10.1103/PhysRevLett.105.032001.

[5] [CMS Collaboration] V. Khachatryan et al. "Measurement of Bose-Einstein correlations in pp collisions at $\sqrt{s} = 0.9$ and 7 TeV". In: *Journal of High Energy Physics* 2011.5 (2011). ISSN: 1029-8479. DOI: 10.1007/jhep05(2011)029. URL: http://dx.doi.org/10.1007/JHEP05(2011)029.

[6] [ALICE Collaboration] J. Adam et al. "Multipion Bose-Einstein correlations in pp, p-Pb, and Pb-Pb collisions at energies available at the CERN Large Hadron Collider". In: *Physical Review C* 93.5 (2016). ISSN: 2469-9993. DOI: 10.1103/physrevc.93.054908. URL: http://dx.doi.org/10.1103/PhysRevC.93.054908.

[7] P. Bozek and W. Broniowski. "Correlations from hydrodynamic flow in p-Pb collisions". In: *Phys. Lett. B* 718 (2013), pp. 1557–1561. DOI: 10.1016/j.physletb.2012.12.051. arXiv: 1211.0845 [nucl-th].

[8] J. Jia. "Collective phenomena in high-energy nuclear collisions". In: *Nucl. Phys. A* 931 (2014). Ed. by Peter Braun-Munzinger, Bengt Friman, and Johanna Stachel, pp. 216–226. DOI: 10.1016/j.nuclphysa.2014.08.045. arXiv: 1408.0066 [nucl-ex].

[9] J. Mercado. "Two-pion Bose-Einstein correlations in Pb-Pb collisions at 2.76 TeV with ALICE". In: *Journal of Physics G-nuclear and Particle Physics - J PHYS G-NUCL PARTICLE PHYS* 38 (July 2011). DOI: 10.1088/0954-3899/38/12/124056.

[10] [ALICE Collaboration] B. Abelev et al. "Two- and three-pion quantum statistics correlations in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV at the CERN Large Hadron Collider". In: *Phys. Rev. C* 89 (2 2014), p. 024911. DOI: 10.1103/PhysRevC.89.024911. URL: https://link.aps.org/doi/10.1103/PhysRevC.89.024911.

[11] B. Schenke and R. Venugopalan. "Eccentric Protons? Sensitivity of Flow to System Size and Shape in $p+p$, $p+\mathrm{Pb}$, and $\mathrm{Pb}+\mathrm{Pb}$ Collisions". In: *Phys. Rev. Lett.* 113 (10 2014), p. 102301. DOI: 10.1103/PhysRevLett.113.102301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.113.102301.

[12] P. Romatschke. "Light-heavy-ion collisions: a window into pre-equilibrium QCD dynamics?" In: *The European Physical Journal C* 75.7 (2015), p. 305. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-015-3509-3. URL: https://doi.org/10.1140/epjc/s10052-015-3509-3.

[13] R. Hanbury Brown and R. Q. Twiss. "Correlation between Photons in two Coherent Beams of Light". In: *Nature* 177.4497 (1956), pp. 27–29. ISSN: 1476-4687. DOI: 10.1038/177027a0. URL: https://doi.org/10.1038/177027a0.

[14] A. Bialas, W. Florkowski, and K. Zalewski. "Bose-Einstein correlations and thermal cluster formation in high-energy collisions". In: *Acta Phys. Polon. B* 45.9 (2014), pp. 1883–1902. DOI: 10.5506/APhysPolB.45.1883. arXiv: 1406.2499 [hep-ph].

[15] T. D. Gottschalk. "A simple phenomenological model for hadron production from low-mass clusters". In: *Nuclear Physics B* 239.2 (1984), pp. 325–348. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(84)90252-9. URL: https://www.sciencedirect.com/science/article/pii/0550321384902529.

[16] T. Csörgő, B. Lörstad, and J. Zimányi. "Bose-Einstein correlations for systems with large halo". In: *Zeitschrift für Physik C: Particles and Fields* 71.3 (1996), 491–497. ISSN: 1431-5858. DOI: 10.1007/bf02907008. URL: http://dx.doi.org/10.1007/BF02907008.

[17] Tamás Novák. "PHENIX Results of Three-Particle Bose-Einstein Correlations in $\sqrt{s_{NN}} = 200$ GeV Au+Au Collisions". In: *Universe* 4.3 (2018). ISSN: 2218-1997. DOI: 10.3390/universe4030057. URL: https://www.mdpi.com/2218-1997/4/3/57.

[18] [LHCb Collaboration] A. Augusto Alves et al. "The LHCb Detector at the LHC". In: *Journal of Instrumentation* 3.08 (2008), S08005–S08005. DOI: 10.1088/1748-0221/3/08/s08005. URL: https://doi.org/10.1088/1748-0221/3/08/s08005.

[19] M. Kucharczyk and M. Zdybal. "Three-pion Bose-Einstein correlations in pp collisions at 7 TeV". In: *LHCb-ANA-2020-074* (2020). URL: https://cds.cern.ch/record/2747321.

[20] M. Zdybał. "Particle Correlations at LHCb". In: *Acta Phys. Pol. B* 51 (2020), 1593–1597. 5 p. DOI: 10.5506/APhysPolB.51.1593. URL: http://cds.cern.ch/record/2799894.

[21] G. Abbiendi et al. "A study of muon-electron elastic scattering in a test beam". In: *Journal of Instrumentation* 16.06 (2021), P06005. DOI: 10.1088/1748-0221/16/06/p06005. URL: https://doi.org/10.1088%2F1748-0221%2F16%2F06%2Fp06005.

[22] G. Abbiendi et al. *Letter of Intent: the MUonE project*. Tech. rep. The collaboration has not yet a structure, therefore the names above are for the moment an indication of contacts. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2677471.

[23] N. T. Doble et al. "The upgraded muon beam at the SPS". In: *Nucl. Instrum. Methods Phys. Res., A* 343 (1993), 351–362. 34 p. DOI: 10.1016/0168-9002(94)90212-7. URL: https://cds.cern.ch/record/250676.

[24] [Muon g-2 Collaboration] B. Abi et al. "Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm". In: *Phys. Rev. Lett.* 126 (14 2021), p. 141801. DOI: 10.1103/PhysRevLett.126.141801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.126.141801.

[25] [Muon g-2 Collaboration] G. W. Bennett et al. "Final report of the E821 muon anomalous magnetic moment measurement at BNL". In: *Physical Review D* 73.7 (2006). DOI: 10.1103/physrevd.73.072003. URL: https://doi.org/10.1103%2Fphysrevd.73.072003.

[26] M. Zdybał, M. Kucharczyk, and M. Wolter. "DNN Based Prototype of the Track Reconstruction Algorithm for the MUonE Experiment". In: *Distributed Computing and Artificial Intelligence, Volume 2: Special Sessions 18th International Conference*. Ed. by Sara Rodríguez González et al. Cham: Springer International Publishing, 2022, pp. 202–205. ISBN: 978-3-030-86887-1.

[27] B. Andersson and M. Ringnér. "Bose-Einstein correlations in the Lund model". In: *Nuclear Physics B* 513.3 (1998), pp. 627–644. ISSN: 0550-3213. DOI: https://doi.org/10.1016/S0550-3213(97)00722-0. URL: https://www.sciencedirect.com/science/article/pii/S0550321397007220.

[28] B. Andersson and M. Ringnér. "Transverse and longitudinal Bose-Einstein correlations". In: *Physics Letters B* 421.1 (1998), pp. 283–288. ISSN: 0370-2693. DOI: https://doi.org/10.1016/S0370-2693(97)01593-1. URL: https://www.sciencedirect.com/science/article/pii/S0370269397015931.

[29] S. Chapman and U. Heinz. "HBT correlators - current formalism vs. Wigner function formulation". In: *Physics Letters B* 340.4 (1994), pp. 250–253. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(94)01277-6. URL: https://www.sciencedirect.com/science/article/pii/0370269394012776.

[30] T. Csörgő, S. Hegyi, and W. A. Zajc. "Bose-Einstein correlations for Lévy stable source distributions". In: *European Physical Journal C* 36 (July 2004), pp. 67–78. DOI: 10.1140/epjc/s2004-01870-9.

[31] [PHENIX Collaboration] A. Adare et al. "Lévy-stable two-pion Bose-Einstein correlations in $\sqrt{s_{NN}} = 200$ GeV Au + Au collisions". In: *Phys. Rev. C* 97 (6 2018), p. 064911. DOI: 10.1103/PhysRevC.97.064911. URL: https://link.aps.org/doi/10.1103/PhysRevC.97.064911.

[32] T. Csörgö. "Particle Interferometry from 40 Mev to 40 Tev". In: *Particle Production Spanning MeV and TeV Energies*. Ed. by W. Kittel, P. J. Mulders, and O. Scholten. Dordrecht: Springer Netherlands, 2000, pp. 203–257. ISBN: 978-94-011-4126-0. DOI: 10.1007/978-94-011-4126-0_8. URL: https://doi.org/10.1007/978-94-011-4126-0_8.

[33] E. V. Shuryak. "Quark-gluon plasma and hadronic production of leptons, photons and psions". In: *Physics Letters B* 78.1 (1978), pp. 150–153. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(78)90370-2. URL: https://www.sciencedirect.com/science/article/pii/0370269378903702.

[34] C. Loizides. "Experimental overview on small collision systems at the LHC". In: *Nuclear Physics A* 956 (2016), pp. 200–207. ISSN: 0375-9474. DOI: 10.1016/j.nuclphysa.2016.04.022. URL: http://dx.doi.org/10.1016/j.nuclphysa.2016.04.022.

[35] C. Andrei. "Light flavor hadron spectra at low pT and search for collective phenomena in high multiplicity pp, p-Pb and Pb-Pb collisions measured with the ALICE experiment". In: *Nuclear Physics A* 931 (2014). QUARK MATTER 2014, pp. 888–892. ISSN: 0375-9474. DOI: https://doi.org/10.1016/j.nuclphysa.2014.08.002. URL: https://www.sciencedirect.com/science/article/pii/S0375947414002516.

[36] [CMS Collaboration] V. Khachatryan et al. "Multiplicity and rapidity dependence of strange hadron production in pp, pPb, and PbPb collisions at the LHC". In: *Physics Letters B* 768 (2017), pp. 103–129. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2017.01.075. URL: http://dx.doi.org/10.1016/j.physletb.2017.01.075.

[37] T. Kalaydzhyan and E. Shuryak. "Collective flow in high-multiplicity proton-proton collisions". In: *Phys. Rev. C* 91 (5 2015), p. 054913. DOI: 10.1103/PhysRevC.91.054913. URL: https://link.aps.org/doi/10.1103/PhysRevC.91.054913.

[38] [ALLICE Collaboration] B. Abelev et al. "Freeze-out radii extracted from three-pion cumulants in pp, p-Pb and Pb-Pb collisions at the LHC". In: *Physics Letters B* 739 (2014), pp. 139–151. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2014.10.034. URL: https://www.sciencedirect.com/science/article/pii/S0370269314007576.

[39] Y. Hirono and E. Shuryak. "Femtoscopic signature of strong radial flow in high-multiplicity *pp* collisions". In: *Phys. Rev. C* 91 (5 2015), p. 054915. DOI: 10.1103/PhysRevC.91.054915. URL: https://link.aps.org/doi/10.1103/PhysRevC.91.054915.

[40] P. Romatschke. "Do nuclear collisions create a locally equilibrated quark-gluon plasma?" In: *The European Physical Journal C* 77.1 (2017). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-4567-x. URL: http://dx.doi.org/10.1140/epjc/s10052-016-4567-x.

[41] L. D. Landau. "On the multiparticle production in high-energy collisions". In: *Izv. Akad. Nauk Ser. Fiz.* 17 (1953), pp. 51–64.

[42] "On multiple production of particles during collisions of fast particles". In: *Collected Papers of L.D. Landau*. Ed. by D. Ter Haar. Pergamon, 1965, pp. 569–585. ISBN: 978-0-08-010586-4. DOI: https://doi.org/10.1016/B978-0-08-010586-4.50079-1. URL: https://www.sciencedirect.com/science/article/pii/B9780080105864500791.

[43] W. Florkowski, M. P. Heller, and M. Spaliński. "New theories of relativistic hydrodynamics in the LHC era". In: *Reports on Progress in Physics* 81.4 (2018), p. 046001. ISSN: 1361-6633. DOI: 10.1088/1361-6633/aaa091. URL: http://dx.doi.org/10.1088/1361-6633/aaa091.

[44] B. Andersson, S. Mohanty, and F. Sderberg. "The Lund fragmentation process for a multi-gluon string according to the area law". Undetermined. In: *The European Physical Journal C* 21.4 (2001), pp. 631–647. DOI: 10.1007/s100520100757.

[45] B. Andersson et al. "Parton fragmentation and string dynamics". In: *Physics Reports* 97.2 (1983), pp. 31–145. ISSN: 0370-1573. DOI: https://doi.org/10.1016/0370-1573(83)90080-7. URL: https://www.sciencedirect.com/science/article/pii/0370157383900807.

[46] B. Andersson and W. Hofmann. "Bose-Einstein correlations and color strings". In: *Physics Letters B* 169.4 (1986), pp. 364–368. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(86)90373-4. URL: https://www.sciencedirect.com/science/article/pii/0370269386903734.

[47] M.G. Bowler. "Elucidation of the relationship between two different treatments of Bose-Einstein correlations in the string model". In: *Physics Letters B* 185.1 (1987), pp. 205–208. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(87)91555-3. URL: https://www.sciencedirect.com/science/article/pii/0370269387915553.

[48] T. Sjöstrand et al. "An introduction to PYTHIA 8.2". In: *Computer Physics Communications* 191 (2015), pp. 159–177. ISSN: 0010-4655. DOI: https://doi.org/10.1016/j.cpc.2015.01.024. URL: https://www.sciencedirect.com/science/article/pii/S0010465515000442.

[49] T. Sjostrand, S. Mrenna, and P. Z. Skands. "PYTHIA 6.4 Physics and Manual". In: *JHEP* 05 (2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175.

[50] B. Andersson, S. Mohanty, and F. Söderberg. "The Lund fragmentation process for a multi-gluon string according to the area law". In: *European Physical Journal C* 21 (2001), pp. 631–647. DOI: 10.1007/s100520100757.

[51] T. Csörgö. "Particle Interferometry from 40 Mev to 40 Tev". In: Jan. 2000, pp. 203–257. ISBN: 978-0-7923-6432-0. DOI: 10.1007/978-94-011-4126-0_8.

[52] [LHCb Collaboration] R. Aaij et al. "Bose-Einstein correlations of same-sign charged pions in the forward region in pp collisions at $\sqrt{s} = 7$ TeV". In: *Journal of High Energy Physics* 2017.12 (2017). ISSN: 1029-8479. DOI: 10.1007/jhep12(2017)025. URL: http://dx.doi.org/10.1007/JHEP12(2017)025.

[53] H. Boggild et al. "Directional dependence of the pion source in high-energy heavy-ion collisions". In: *Physics Letters B* 349.3 (1995), pp. 386–392. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(95)00305-5. URL: https://www.sciencedirect.com/science/article/pii/0370269395003055.

[54] C. Albajar et al. "Bose-Einstein correlations in pp interactions at $\sqrt{s} =$ 0.2 to 0.9 TeV". In: *Physics Letters B* 226.3 (1989), pp. 410–416. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(89)91221-5. URL: https://www.sciencedirect.com/science/article/pii/0370269389912215.

[55] [CERES Collaboration] D. Adamová et al. "Beam energy and centrality dependence of two-pion Bose-Einstein correlations at SPS energies". In: *Nuclear Physics A* 714.1 (2003), pp. 124–144. ISSN: 0375-9474. DOI: https://doi.org/10.1016/S0375-9474(02)01369-6. URL: https://www.sciencedirect.com/science/article/pii/S0375947402013696.

[56] [ALEPH Collaboration] D. Decamp et al. "A study of Bose-Einstein correlations in e+e- annihilation at 91 GeV". In: *Zeitschrift fur Physik C Particles and Fields* 54.1 (1992), pp. 75–85. ISSN: 1431-5858. DOI: 10.1007/BF01881709. URL: https://doi.org/10.1007/BF01881709.

[57] [ALEPH Collaboration] D. Decamp et al. "Two-dimensional analysis of Bose-Einstein correlationsin hadronic Z decays at LEP". In: *The European Physical Journal C - Particles and Fields* 36.2 (2004), pp. 147–159. ISSN: 1434-6052. DOI: 10.1140/epjc/s2004-01884-3. URL: https://doi.org/10.1140/epjc/s2004-01884-3.

[58] [DELPHI Collaboration] P. Abreu et al. "Two-dimensional analysis of the Bose-Einstein correlations in e+e- annihilation at the Z0 peak". In: *Physics Letters B* 471.4 (2000), pp. 460–470. ISSN: 0370-2693. DOI: https://doi.org/10.1016/S0370-2693(99)01391-X. URL: https://www.sciencedirect.com/science/article/pii/S037026939901391X.

[59] [STAR Collaboration] J. Adams et al. "Three-Pion Hanbury Brown–Twiss Correlations in Relativistic Heavy-Ion Collisions from the STAR Experiment". In: *Phys. Rev. Lett.* 91 (26 2003), p. 262301. DOI: 10.1103/PhysRevLett.91.262301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.91.262301.

[60] [PHENIX Collaboration] S. S. Adler et al. "Bose-Einstein Correlations of Charged Pion Pairs in Au + Au Collisions at $\sqrt{s_{NN}} = 200$ GeV". In: *Phys. Rev. Lett.* 93 (15 2004), p. 152302. DOI: 10.1103/PhysRevLett.93.152302. URL: https://link.aps.org/doi/10.1103/PhysRevLett.93.152302.

[61] T. Csörgő and S. Hegyi. "Model independent shape analysis of correlations in 1, 2 or 3 dimensions". In: *Physics Letters B* 489.1 (2000), pp. 15–23. ISSN: 0370-2693. DOI: https://doi.org/10.1016/S0370-2693(00)00935-7. URL: https://www.sciencedirect.com/science/article/pii/S0370269300009357.

[62] D. Kincses. "PHENIX Results on Lévy Analysis of Bose–Einstein Correlation Functions". In: *Acta Physica Polonica B Proceedings Supplement* 10.3 (2017), p. 627. ISSN: 2082-7865. DOI: 10.5506/aphyspolbsupp.10.627. URL: http://dx.doi.org/10.5506/APhysPolBSupp.10.627.

[63] [The Axial Field Spectrometer Collaboration] T. Åkesson et al. "Bose-Einstein correlations in $\alpha\alpha$, pp and pp interactions". In: *Physics Letters B* 129.3 (1983), pp. 269–272. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(83)90859-6. URL: https://www.sciencedirect.com/science/article/pii/0370269383908596.

[64] M. A. Lisa et al. "Femtoscopy in relativistic heavy ion collisions: Two Decades of Progress". In: *Annual Review of Nuclear and Particle Science* 55.1 (2005), pp. 357–402. DOI: 10.1146/annurev.nucl.55.090704.151533. URL: https://doi.org/10.1146%2Fannurev.nucl.55.090704.151533.

[65] [ALICE Collaboration] J. Adam et al. "Two-pion femtoscopy in p-Pb collisions atsNN=5.02TeV". In: *Physical Review C* 91.3 (2015). DOI: 10.1103/physrevc.91.034906. URL: https://doi.org/10.1103%2Fphysrevc.91.034906.

[66] U. Heinz and R. Snellings. "Collective Flow and Viscosity in Relativistic Heavy-Ion Collisions". In: *Annual Review of Nuclear and Particle Science* 63.1 (2013), pp. 123–151. DOI: 10.1146/annurev-nucl-102212-170540. URL: https://doi.org/10.1146%2Fannurev-nucl-102212-170540.

[67] K. Werner et al. "Evidence for hydrodynamic evolution in proton-proton scattering at 900 GeV". In: *Physical Review C* 83.4 (2011). DOI: 10.1103/physrevc.83.044915. URL: https://doi.org/10.1103%2Fphysrevc.83.044915.

[68] V. M. Shapoval et al. "Femtoscopic scales in p+p and p+Pb collisions in view of the uncertainty principle". In: *Physics Letters B* 725.1-3 (2013), pp. 139–147. DOI: 10.1016/j.physletb.2013.07.002. URL: https://doi.org/10.1016%2Fj.physletb.2013.07.002.

[69] A. Bzdak et al. "Initial-state geometry and the role of hydrodynamics in proton-proton, proton-nucleus, and deuteron-nucleus collisions". In: *Physical Review C* 87.6 (2013). DOI: 10.1103/physrevc.87.064906. URL: https://doi.org/10.1103%2Fphysrevc.87.064906.

[70] L. McLerran, M. Praszalowicz, and B. Schenke. "Transverse momentum of protons, pions and kaons in high multiplicity pp and pA collisions: Evidence for the color glass condensate?" In: *Nuclear Physics A* 916 (2013), pp. 210–218. DOI: 10.1016/j.nuclphysa.2013.08.008. URL: https://doi.org/10.1016%2Fj.nuclphysa.2013.08.008.

[71] [CMS Collaboration] A. M. Sirunyan et al. "Bose-Einstein correlations of charged hadrons in proton-proton collisions at $$ \sqrt{s} $$ = 13 TeV". In: *Journal of High Energy Physics* 2020.3 (2020). DOI: 10.1007/jhep03(2020)014. URL: https://doi.org/10.1007%2Fjhep03%282020%29014.

[72] P. Bożek and W. Broniowski. "Size of the emission source and collectivity in ultra-relativistic p–Pb collisions". In: *Physics Letters B* 720.1-3 (2013), pp. 250–253. DOI: 10.1016/j.physletb.2013.02.014. URL: https://doi.org/10.1016%2Fj.physletb.2013.02.014.

[73] [ALICE Collaboration] J. Adam et al. "One-dimensional pion, kaon, and proton femtoscopy in Pb-Pb collisions atsNN=2.76TeV". In: *Physical Review C* 92.5 (2015). DOI: 10.1103/physrevc.92.054908. URL: https://doi.org/10.1103%2Fphysrevc.92.054908.

[74] [ATLAS Collaboration] M. Aaboud et al. "Femtoscopy with identified charged pions in proton-lead collisions at sNN=5.02 TeV with ATLAS". In: *Physical Review C* 96.6 (2017). DOI: 10.1103/physrevc.96.064908. URL: https://doi.org/10.1103%2Fphysrevc.96.064908.

[75] R Alemany-Fernandez et al. "Operation and Configuration of the LHC in Run 1". In: (2013). URL: https://cds.cern.ch/record/1631030.

[76] L. Evans and P. Bryant. "LHC Machine". In: *Journal of Instrumentation* 3.08 (2008), S08001–S08001. DOI: 10.1088/1748-0221/3/08/s08001. URL: https://doi.org/10.1088/1748-0221/3/08/s08001.

[77] E. Mobs. "The CERN accelerator complex - August 2018. Complexe des andaccélérateurs du CERN - Août 2018". In: (2018). General Photo. URL: https://cds.cern.ch/record/2636343.

[78] [LHCb Collaboration] R. Aaij et al. "LHCb detector performance". In: *International Journal of Modern Physics A* 30.07 (2015), p. 1530022. ISSN: 1793-656X. DOI: 10.1142/s0217751x15300227. URL: http://dx.doi.org/10.1142/S0217751X15300227.

[79] [ALICE Collaboration] K. Aamodt et al. "The ALICE experiment at the CERN LHC". In: *Journal of Instrumentation* 3.08 (2008), S08002–S08002. DOI: 10.1088/1748-0221/3/08/s08002. URL: https://doi.org/10.1088/1748-0221/3/08/s08002.

[80] [ATLAS Collaboration] G. et al. Aad. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[81] [CMS Collaboration] S. et al. Chatrchyan. "The CMS Experiment at the CERN LHC". In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004.

[82] F. Follin and D. Jacquet. "Implementation and experience with luminosity levelling with offset beam". In: (2014). Comments: 5 pages, contribution to the ICFA Mini-Workshop on Beam-Beam Effects in Hadron Colliders, CERN, Geneva, Switzerland, 18-22 Mar 2013, 5 p. DOI: 10.5170/CERN-2014-004.183. arXiv: 1410.3667. URL: http://cds.cern.ch/record/1955354.

[83] C. Elsässer. *$\bar{b}b$ production angle plots*. URL: {https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html}.

[84] [LHCb Collaboration] S. Amato et al. *LHCb magnet: Technical Design Report*. Technical design report. LHCb. Geneva: CERN, 2000. URL: http://cds.cern.ch/record/424338.

[85] [LHCb Collaboration] R. Aaij et al. "Performance of the LHCb Vertex Locator". In: *Journal of Instrumentation* 9.09 (2014), P09007–P09007. DOI: 10.1088/1748-0221/9/09/p09007. URL: https://doi.org/10.1088%2F1748-0221%2F9%2F09%2Fp09007.

[86] [LHCb Collaboration] R. Aaij et al. "Measurement of the track reconstruction efficiency at LHCb". In: *Journal of Instrumentation* 10.02 (2015), P02007–P02007. DOI: 10.1088/1748-0221/10/02/p02007. URL: https://doi.org/10.1088%2F1748-0221%2F10%2F02%2Fp02007.

[87] [LHCb Collaboration] R. Aaij et al. *Upgrade Software and Computing*. Tech. rep. Geneva: CERN, 2018. URL: http://cds.cern.ch/record/2310827.

[88] G. Barrand et al. "GAUDI — A software architecture and framework for building HEP data processing applications". In: *Computer Physics Communications* 140.1 (2001). CHEP2000, pp. 45–55. ISSN: 0010-4655. DOI: https://doi.org/10.1016/S0010-4655(01)00254-5. URL: https://www.sciencedirect.com/science/article/pii/S0010465501002545.

[89] I. Antcheva et al. "ROOT — A C++ framework for petabyte data storage, statistical analysis and visualization". In: *Computer Physics Communications* 180.12 (2009). 40 YEARS OF CPC: A celebratory issue focused on quality software for high performance, grid and novel computing architectures, pp. 2499–2512. ISSN: 0010-4655. DOI: https://doi.org/10.1016/j.cpc.2009.08.005. URL: https://www.sciencedirect.com/science/article/pii/S0010465509002550.

[90] David J. Lange. "The EvtGen particle decay simulation package". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 462.1 (2001). BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines, pp. 152–155. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(01)00089-4. URL: https://www.sciencedirect.com/science/article/pii/S0168900201000894.

[91] S. Agostinelli et al. "Geant4—a simulation toolkit". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(03)01368-8. URL: https://www.sciencedirect.com/science/article/pii/S0168900203013688.

[92] [LHCb Collaboration] R. Aaij et al. *Physics case for an LHCb Upgrade II - Opportunities in flavour physics, and beyond, in the HL-LHC era*. Tech. rep. ISBN 978-92-9083-494-6. Geneva: CERN, 2018. DOI: 10.17181/CERN.QZRZ.R4S6. arXiv: 1808.08865. URL: https://cds.cern.ch/record/2636441.

[93] European Strategy Group. *2020 Update of the European Strategy for Particle Physics*. Tech. rep. Geneva, 2020. DOI: 10.17181/ESU2020. URL: https://cds.cern.ch/record/2720129.

[94] I. Zurbano Fernandez et al. "High-Luminosity Large Hadron Collider (HL-LHC): Technical design report". In: 10/2020 (Dec. 2020). Ed. by I. Béjar Alonso et al. DOI: 10.23731/CYRM-2020-0010.

[95] J. Albrecht et al. "The upgrade of the LHCb trigger system". In: *Journal of Instrumentation* 9.10 (2014), C10026–C10026. ISSN: 1748-0221. DOI: 10.1088/1748-0221/9/10/c10026. URL: http://dx.doi.org/10.1088/1748-0221/9/10/C10026.

[96] [LHCb Collaboration] R. Aaij et al. *LHCb VELO Upgrade Technical Design Report*. Tech. rep. 2013. URL: http://cds.cern.ch/record/1624070.

[97] [LHCb Collaboration] R. Aaij et al. *LHCb Tracker Upgrade Technical Design Report*. Tech. rep. 2014. URL: http://cds.cern.ch/record/1647400.

[98] [LHCb Collaboration] R. Aaij et al. *LHCb PID Upgrade Technical Design Report*. Tech. rep. 2013. URL: http://cds.cern.ch/record/1624074.

[99] [LHCb Collaboration] R. Aaij et al. *Expression of Interest for a Phase-II LHCb Upgrade: Opportunities in flavour physics, and beyond, in the HL-LHC era*. Tech. rep. Geneva: CERN, 2017. URL: http://cds.cern.ch/record/2244311.

[100] [LHCb Collaboration] R. Aaij et al. "RTA and DPA dataflow diagrams for Run 1, Run 2, and the upgraded LHCb detector ". In: (2020). URL: https://cds.cern.ch/record/2730181.

[101] R. Aaij et al. "A comprehensive real-time analysis model at the LHCb experiment". In: *JINST* 14 (2019), P04006. 15 p. DOI: 10.1088/1748-0221/14/04/P04006. arXiv: 1903.01360. URL: https://cds.cern.ch/record/2665946.

[102] J. Albrecht et al. *Upgrade trigger & reconstruction strategy: 2017 milestone*. Tech. rep. Geneva: CERN, 2018. URL: http://cds.cern.ch/record/2310579.

[103] D. Kincses, M. I. Nagy, and M. Csanád. "Coulomb and strong interactions in the final state of Hanbury-Brown–Twiss correlations for Lévy-type source functions". In: *Phys. Rev. C* 102 (6 2020), p. 064912. DOI: 10.1103/PhysRevC.102.064912. URL: https://link.aps.org/doi/10.1103/PhysRevC.102.064912.

[104] D. Gangadharan. "Techniques for multiboson interferometry". In: *Phys. Rev. C* 92 (1 2015), p. 014902. DOI: 10.1103/PhysRevC.92.014902. URL: https://link.aps.org/doi/10.1103/PhysRevC.92.014902.

[105] M. Gyulassy, S. K. Kauffmann, and Lance W. Wilson. "Pion interferometry of nuclear collisions. I. Theory". In: *Phys. Rev. C* 20 (6 1979), pp. 2267–2292. DOI: 10.1103/PhysRevC.20.2267. URL: https://link.aps.org/doi/10.1103/PhysRevC.20.2267.

[106] S. Pratt. "Coherence and Coulomb effects on pion interferometry". In: *Phys. Rev. D* 33 (1 1986), pp. 72–79. DOI: 10.1103/PhysRevD.33.72. URL: https://link.aps.org/doi/10.1103/PhysRevD.33.72.

[107] [CMS Collaboration] A. M. Sirunyan et al. "Bose-Einstein correlations in pp, pPb, and PbPb collisions at $\sqrt{s_{NN}} = 0.9--7$ TeV". In: *Phys. Rev. C* 97 (6 2018), p. 064912. DOI: 10.1103/PhysRevC.97.064912. URL: https://link.aps.org/doi/10.1103/PhysRevC.97.064912.

[108] [LHCb Collaboration] R. Aaij et al. "Precision luminosity measurements at LHCb". In: *Journal of Instrumentation* 9.12 (2014), P12005–P12005. DOI: 10.1088/1748-0221/9/12/p12005. URL: https://doi.org/10.1088%2F1748-0221%2F9%2F12%2Fp12005.

[109] I. Belyaev et al. "Handling of the generation of primary events in Gauss, the LHCb simulation framework". In: *Journal of Physics: Conference Series* 331.3 (2011), p. 032047. DOI: 10.1088/1742-6596/331/3/032047. URL: https://doi.org/10.1088/1742-6596/331/3/032047.

[110] P. Golonka and Z. Was. "PHOTOS Monte Carlo: a precision tool for QED correctionsin Z and W decays". In: *The European Physical Journal C* 45.1 (2006), pp. 97–107. DOI: 10.1140/epjc/s2005-02396-4. URL: https://doi.org/10.1140%2Fepjc%2Fs2005-02396-4.

[111] T. Sjöstrand, S. Mrenna, and P. Skands. "PYTHIA 6.4 physics and manual". In: *Journal of High Energy Physics* 2006.05 (2006), pp. 026–026. DOI: 10.1088/1126-6708/2006/05/026. URL: https://doi.org/10.1088/1126-6708/2006/05/026.

[112] P. Z. Skands. *The Perugia Tunes*. 2009. DOI: 10.48550/ARXIV.0905.3418. URL: https://arxiv.org/abs/0905.3418.

[113] G. D'Agostini. "A multidimensional unfolding method based on Bayes' theorem". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 362.2 (1995), pp. 487–498. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(95)00274-X. URL: https://www.sciencedirect.com/science/article/pii/016890029500274X.

[114] [LHCb Collaboration] R. Aaij et al. "Measurement of charged particle multiplicities and densities in pp collisions at $\sqrt{s} = 7$ TeV in the forward region". In: *The European Physical Journal C* 74.5 (2014). DOI: 10.1140/epjc/s10052-014-2888-1. URL: https://doi.org/10.1140%2Fepjc%2Fs10052-014-2888-1.

[115] L. Anderlini et al. *The PIDCalib package*. Tech. rep. Geneva: CERN, 2016. URL: https://cds.cern.ch/record/2202412.

[116] M. Needham. *Clone Track Identification using the Kullback-Liebler Distance*. Tech. rep. Geneva: CERN, 2008. URL: https://cds.cern.ch/record/1082460.

[117] W. N. Cottingham and D. A. Greenwood. *An Introduction to the Standard Model of Particle Physics*. 2nd ed. Cambridge University Press, 2007. DOI: 10.1017/CBO9780511791406.

[118] M. Abe et al. *A New Approach for Measuring the Muon Anomalous Magnetic Moment and Electric Dipole Moment*. 2019. DOI: 10.48550/ARXIV.1901.03047. URL: https://arxiv.org/abs/1901.03047.

[119] G. Abbiendi et al. "Measuring the leading hadronic contribution to the muon g-2 via μe scattering". In: *The European Physical Journal C* 77.3 (2017). DOI: 10.1140/epjc/s10052-017-4633-z. URL: https://doi.org/10.1140%2Fepjc%2Fs10052-017-4633-z.

[120] F. H. Combley. "(g-2) factors for muon and electron and the consequences for QED". In: *Reports on Progress in Physics* 42.12 (1979), pp. 1889–1935. DOI: 10.1088/0034-4885/42/12/001. URL: https://doi.org/10.1088/0034-4885/42/12/001.

[121] F. Combley, F. J. M. Farley, and E. Picasso. "The CERN muon (g-2) experiments". In: *Physics Reports* 68.2 (1981), pp. 93–119. ISSN: 0370-1573. DOI: https://doi.org/10.1016/0370-1573(81)90028-4. URL: https://www.sciencedirect.com/science/article/pii/0370157381900284.

[122] T. Aoyama et al. "The anomalous magnetic moment of the muon in the Standard Model". In: *Physics Reports* 887 (2020), pp. 1–166. DOI: 10.1016/j.physrep.2020.07.006. URL: https://doi.org/10.1016%2Fj.physrep.2020.07.006.

[123] T. Aoyama, T. Kinoshita, and M. Nio. "Theory of the Anomalous Magnetic Moment of the Electron". In: *Atoms* 7.1 (2019), p. 28. DOI: 10.3390/atoms7010028.

[124] T. Kinoshita and M. Nio. "Tenth-order QED contribution to the lepton $g-2$: Evaluation of dominant α^5 terms of muon $g-2$". In: *Phys. Rev. D* 73 (5 2006), p. 053007. DOI: 10.1103/PhysRevD.73.053007. URL: https://link.aps.org/doi/10.1103/PhysRevD.73.053007.

[125] R. H. Parker et al. "Measurement of the fine-structure constant as a test of the Standard Model". In: *Science* 360.6385 (2018), pp. 191–195. DOI: 10.1126/science.aap7706. URL: https://doi.org/10.1126%2Fscience.aap7706.

[126] T. Ishikawa, N. Nakazawa, and Y. Yasui. "Numerical calculation of the full two-loop electroweak corrections to muon (g-2)". In: *Physical Review D* 99.7 (2019). DOI: 10.1103/physrevd.99.073004. URL: https://doi.org/10.1103%2Fphysrevd.99.073004.

[127] A. Czarnecki, W. J. Marciano, and A. Vainshtein. "Refinements in electroweak contributions to the muon anomalous magnetic moment". In: *Phys. Rev. D* 67 (7 2003), p. 073006. DOI: 10.1103/PhysRevD.67.073006. URL: https://link.aps.org/doi/10.1103/PhysRevD.67.073006.

[128] M. Davier et al. "A new evaluation of the hadronic vacuum polarisation contributions to the muon anomalous magnetic moment and to $\alpha(m_Z^2)$". In: *The European Physical Journal C* 80 (Mar. 2020). DOI: 10.1140/epjc/s10052-020-7792-2.

[129] B. E. Lautrup and E. de Rafael. "Calculation of the Sixth-Order Contribution from the Fourth-Order Vacuum Polarization to the Difference of the Anomalous Magnetic Moments of Muon and Electron". In: *Phys. Rev.* 174 (5 1968), pp. 1835–1842. DOI: 10.1103/PhysRev.174.1835. URL: https://link.aps.org/doi/10.1103/PhysRev.174.1835.

[130] A. Kurz et al. "Hadronic contribution to the muon anomalous magnetic moment to next-to-next-to-leading order". In: *Physics Letters B* 734 (2014), pp. 144–147. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2014.05.043. URL: https://www.sciencedirect.com/science/article/pii/S0370269314003529.

[131] "Evidence for light-by-light scattering in heavy-ion collisions with the ATLAS detector at the LHC". In: *Nature Physics* 13.9 (2017), pp. 852–858. DOI: 10.1038/nphys4208. URL: https://doi.org/10.1038%2Fnphys4208.

[132] N. Cabibbo and R. Gatto. "Electron-Positron Colliding Beam Experiments". In: *Phys. Rev.* 124 (5 1961), pp. 1577–1595. DOI: 10.1103/PhysRev.124.1577. URL: https://link.aps.org/doi/10.1103/PhysRev.124.1577.

[133] H. B. Meyer and H. Wittig. "Lattice QCD and the anomalous magnetic moment of the muon". In: *Progress in Particle and Nuclear Physics* 104 (2019), pp. 46–96. DOI: 10.1016/j.ppnp.2018.09.001. URL: https://doi.org/10.1016%2Fj.ppnp.2018.09.001.

[134] M. Della Morte et al. "The hadronic vacuum polarization contribution to the muon g - 2 from lattice QCD". In: *Journal of High Energy Physics* 2017.10 (2017). DOI: 10.1007/jhep10(2017)020. URL: https://doi.org/10.1007%2Fjhep10%282017%29020.

[135] T. Blum et al. "Calculation of the Hadronic Vacuum Polarization Disconnected Contribution to the Muon Anomalous Magnetic Moment". In: *Physical Review Letters* 116.23 (2016). DOI: 10.1103/physrevlett.116.232002. URL: https://doi.org/10.1103%2Fphysrevlett.116.232002.

[136] C. M. Carloni Calame et al. "A new approach to evaluate the leading hadronic corrections to the muon g -2". In: *Physics Letters B* 746 (2015), pp. 325–329. DOI: 10.1016/j.physletb.2015.05.020. URL: https://doi.org/10.1016%2Fj.physletb.2015.05.020.

[137] G. Abbiendi. "Status of the MUonE experiment". In: *Physica Scripta* 97.5 (2022), p. 054007. DOI: 10.1088/1402-4896/ac6297. URL: https://doi.org/10.1088%2F1402-4896%2Fac6297.

[138] C. Biino. "The CMS Electromagnetic Calorimeter: overview, lessons learned during Run 1 and future projections". In: *Journal of Physics: Conference Series* 587 (2015), p. 012001. DOI: 10.1088/1742-6596/587/1/012001. URL: https://doi.org/10.1088/1742-6596/587/1/012001.

[139] [COMPASS Collaboration] P. Abbon et al. "The COMPASS experiment at CERN". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 577.3 (2007), pp. 455–518. DOI: 10.1016/j.nima.2007.03.026. URL: https://doi.org/10.1016%2Fj.nima.2007.03.026.

[140] M. Al-Turany et al. "The FairRoot framework". In: *J. Phys. Conf. Ser.* 396 (2012). Ed. by M. Ernst et al., p. 022001. DOI: 10.1088/1742-6596/396/2/022001.

[141] S. R. Amendolia et al. "A measurement of the space-like pion electromagnetic form factor". In: *Nuclear Physics B* 277 (1986), pp. 168–196. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(86)90437-2. URL: https://www.sciencedirect.com/science/article/pii/0550321386904372.

[142] A. Kaplan and M. Haenlein. "Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence". In: *Business Horizons* 62.1 (2019), pp. 15–25. ISSN: 0007-6813. DOI: https://doi.org/10.1016/j.bushor.2018.08.004. URL: https://www.sciencedirect.com/science/article/pii/S0007681318301393.

[143] W. S. McCulloch and W. Pitts. "A logical calculus of the ideas immanent in nervous activity". In: *The bulletin of mathematical biophysics* 5.4 (1943), pp. 115–133.

[144] X. Wu et al. "Top 10 algorithms in data mining". In: *Knowledge and information systems* 14.1 (2008), pp. 1–37.

[145] "Bayesian Network". In: *Encyclopedia of Machine Learning*. Ed. by C. Sammut and G. I. Webb. Boston, MA: Springer US, 2010, pp. 81–81. ISBN: 978-0-387-30164-8. DOI: 10.1007/978-0-387-30164-8_65. URL: https://doi.org/10.1007/978-0-387-30164-8_65.

[146] T. Alam et al. "Genetic Algorithm: Reviews, Implementations, and Applications". In: (2020). DOI: 10.48550/ARXIV.2007.12673. URL: https://arxiv.org/abs/2007.12673.

[147] Frederic B. Fitch. "Warren S. McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. Bulletin of mathematical biophysics, vol. 5 (1943), pp. 115-133." In: *Journal of Symbolic Logic* 9.2 (1944), pp. 49–50. DOI: 10.2307/2268029.

[148] F. Rosenblatt. "The perceptron: a probabilistic model for information storage and organization in the brain." In: *Psychological review* 65 6 (1958), pp. 386–408.

[149] B. Widrow and M. E. Hoff. "Associative Storage and Retrieval of Digital Information in Networks of Adaptive "Neurons"". In: *Biological Prototypes and Synthetic Systems: Volume 1 Proceedings of the Second Annual Bionics Symposium sponsored by Cornell University and the General Electric Company, Advanced Electronics Center, held at Cornell University, August 30–September 1, 1961*. Ed. by E. E. Bernard and M. R. Kare. Boston, MA: Springer US, 1962, pp. 160–160. ISBN: 978-1-4684-1716-6. DOI: 10.1007/978-1-4684-1716-6_25. URL: https://doi.org/10.1007/978-1-4684-1716-6_25.

[150] Kumpati S. Narendra and M. A. L. Thathachar. "Learning Automata - A Survey". In: *IEEE Transactions on Systems, Man, and Cybernetics* SMC-4.4 (1974), pp. 323–334. DOI: 10.1109/TSMC.1974.5408453.

[151] S. Grossberg. "Some Networks That Can Learn, Remember, and Reproduce Any Number of Complicated Space-Time Patterns, I". In: *Journal of Mathematics and Mechanics* 19.1 (1969), pp. 53–91. ISSN: 00959057, 19435274. URL: http://www.jstor.org/stable/24902110 (visited on 05/16/2022).

[152] C. von der Malsburg. "Self-organization of orientation sensitive cells in the striate cortex". In: *Kybernetik* 14 (2004), pp. 85–100.

[153] D. J. Willshaw and C. Von Der Malsburg. "How Patterned Neural Connections Can Be Set Up by Self-Organization". In: *Proceedings of the Royal Society of London. Series B, Biological Sciences* 194.1117 (1976), pp. 431–445. ISSN: 00804649. URL: http://www.jstor.org/stable/77138 (visited on 05/16/2022).

[154] Y. LeCun et al. "Backpropagation Applied to Handwritten Zip Code Recognition". In: *Neural Computation* 1.4 (1989), pp. 541–551. DOI: 10.1162/neco.1989.1.4.541.

[155] C. Nwankpa et al. *Activation Functions: Comparison of trends in Practice and Research for Deep Learning*. 2018. DOI: 10.48550/ARXIV.1811.03378. URL: https://arxiv.org/abs/1811.03378.

[156] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. http://www.deeplearningbook.org. MIT Press, 2016.

[157] J. Turian, J. Bergstra, and Y. Bengio. "Quadratic Features and Deep Architectures for Chunking". In: *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Companion Volume: Short Papers*. Boulder, Colorado: Association for Computational Linguistics, June 2009, pp. 245–248. URL: https://aclanthology.org/N09-2062.

[158] P. Le and W. Zuidema. "Compositional Distributional Semantics with Long Short Term Memory". In: *Proceedings of the Fourth Joint Conference on Lexical and Computational Semantics*. Denver, Colorado: Association for Computational Linguistics, June 2015, pp. 10–19. DOI: 10.18653/v1/S15-1002. URL: https://aclanthology.org/S15-1002.

[159] W. Ping et al. *Deep Voice 3: Scaling Text-to-Speech with Convolutional Sequence Learning*. 2017. DOI: 10.48550/ARXIV.1710.07654. URL: https://arxiv.org/abs/1710.07654.

[160] D. Attwell and S. B. Laughlin. "An Energy Budget for Signaling in the Grey Matter of the Brain". In: *Journal of Cerebral Blood Flow & Metabolism* 21.10 (2001). PMID: 11598490, pp. 1133–1145. DOI: 10.1097/00004647-200110000-00001. eprint: https://doi.org/10.1097/00004647-200110000-00001. URL: https://doi.org/10.1097/00004647-200110000-00001.

[161] X. Glorot, A. Bordes, and Y. Bengio. "Deep Sparse Rectifier Neural Networks". In: *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*. Ed. by Geoffrey Gordon, David Dunson, and Miroslav Dudík. Vol. 15. Proceedings of Machine Learning Research. Fort Lauderdale, FL, USA: PMLR, 2011, pp. 315–323. URL: https://proceedings.mlr.press/v15/glorot11a.html.

[162] A. L. Maas, A. Y. Hannun, and A. Y. Ng. "Rectifier nonlinearities improve neural network acoustic models". In: *in ICML Workshop on Deep Learning for Audio, Speech and Language Processing*. 2013.

[163] R. Bellman. "A Markovian Decision Process". In: *Journal of Mathematics and Mechanics* 6.5 (1957), pp. 679–684. ISSN: 00959057, 19435274. URL: http://www.jstor.org/stable/24900506 (visited on 05/20/2022).

[164] M. Kucharczyk and M. Wolter. "Track Finding with Deep Neural Networks". In: *Computer Science* 20.4 (2019). DOI: 10.7494/csci.2019.20.4.3376. URL: https://journals.agh.edu.pl/csci/article/view/3376.

[165] Y. LeCun, Y. Bengio, and G. Hinton. "Deep Learning". In: *Nature* 521 (May 2015), pp. 436–44. DOI: 10.1038/nature14539.

[166] F. Chollet et al. *Keras*. https://keras.io. 2015.

[167] M. Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015. URL: https://www.tensorflow.org/.

[168] S. Farrell et al. "The HEP.TrkX Project: deep neural networks for HL-LHC online and offline tracking". In: *EPJ Web Conf.* 150 (2017), p. 00003. DOI: 10.1051/epjconf/201715000003. URL: https://doi.org/10.1051/epjconf/201715000003.

[169] S. Farrell et al. *Novel deep learning methods for track reconstruction*. 2018. DOI: 10.48550/ARXIV.1810.06111. URL: https://arxiv.org/abs/1810.06111.

[170] R. Brun and F. Rademakers. "ROOT — An object oriented data analysis framework". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 389.1 (1997). New Computing Techniques in Physics Research V, pp. 81–86. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(97)00048-X. URL: https://www.sciencedirect.com/science/article/pii/S016890029700048X.

[171] A. Paszke et al. *PyTorch: An Imperative Style, High-Performance Deep Learning Library*. 2019. arXiv: 1912.01703 [cs.LG].

[172] *MSELOSS*. URL: https://pytorch.org/docs/stable/generated/torch.nn.MSELoss.html.

[173] M. A. Fischler and R. C. Bolles. "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography". In: *Commun. ACM* 24.6 (1981), pp. 381–395. ISSN: 0001-0782. DOI: 10.1145/358669.358692. URL: https://doi.org/10.1145/358669.358692.

[174] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[175] S. Bhattacharya et al. *GNN-based end-to-end reconstruction in the CMS Phase 2 High-Granularity Calorimeter*. 2022. DOI: 10.48550/ARXIV.2203.01189. URL: https://arxiv.org/abs/2203.01189.

[176] [ATLAS Collaboration] D. T. Murnane et al. "Graph Neural Network Track Reconstruction for the ATLAS ITk Detector". In: (2022). URL: https://cds.cern.ch/record/2809518.

[177] X. Ju et al. *Graph Neural Networks for Particle Reconstruction in High Energy Physics detectors*. 2020. DOI: 10.48550/ARXIV.2003.11603. URL: https://arxiv.org/abs/2003.11603.