Universiteit Utrecht

Electromagnetic Shower Recognition with a Forward Calorimeter for the ALICE Experiment

-

Alexandru-Ionut Babeanu

July 24, 2013

Master Thesis
Report number: UU(SAP) 13-7

Supervisors:
Prof. Dr. Thomas Peitzmann
MSc. Davide Lodato

Utrecht University
Faculty of Science/Department of Physics and Astronomy
Institute for Subatomic Physics
Buys Ballot laboratory
PO BOX 80 000, 3508 TA Utrecht, The Netherlands

Abstract

This work deals with discriminating between electromagnetic showers induced by high-energy single photons and those induced by the double photon decays originating from high-energy neutral pions. The preliminary design of a hybrid detector is employed, which is intended for use as an electromagnetic calorimeter in the forward region of the ALICE experimental setup at CERN. The calorimeter combines conventional, low-granularity sensors with non-conventional high-granularity sensors. Results from simulations are systematically studied, revealing that the high-granularity technology is essential for separating photon pairs originating from neutral pions and for identifying these decays based on invariant mass analysis. Further improvements enabled by shower shape analysis for unseparated photon pairs are explored. The study shows that a specific procedure for processing the response of the high-granularity component is required in order for its effective use in shower shape analysis. The performance of the detector is evaluated in the context of proton-proton collisions, emphasizing the advantage of high-granularity technology, the interplay between invariant mass analysis and shower shape analysis, as well as the complementarity of low-granularity and high-granularity information for shower shape analysis.

Contents

1 Physics Motivation

The ALICE collaboration [1] aims at studying effects associated to the quark-gluon plasma [2] (QGP) produced in high-energy collisions of nuclei, which should exist as a consequence of the quantum chromodynamics (QCD) sector of the Standard Model of Particles and Interactions – see chapter 9 of Ref. [3] and references therein. Hadrons are bound states of fermionic particles called quarks, which are kept together by continuous exchange of bosonic partices called gluons. Quarks and gluons, referred to as partons, are thus seen as the fundamental constituents of hadronic matter, which both carry a so-called color charge. While quarks carry the quantum numbers of hadrons, most of the mass of individual hadrons arises dynamically from the gluon binding. On the other hand, the momentum of rapidly moving hadrons is carried in approximately equal fractions by quarks and gluons, which was first observed in deep inelastic scattering experiments (DIS) [4]. In high-energy hadronic collisions, the hadrons would then be broken into these fundamental constituents, while more quarks and gluons can be generated from the vacuum. The quark-gluon plasma can then roughly be understood as the totality of quarks and gluons produced immediately after a high-energy collision, together with its collective behaviour. Such effects can for instance be studied via elliptic flow [5] and jet-quenching [6] in lead-lead collisions versus proton-proton collisions. Proton-proton collisions are used as a baseline, where QGP effects should be much less obvious due to the smaller number of quarks and gluons present in the initial state, as well as due to the small volume occupied by these particles, which is comparable to the typical confinement volume. Essentially, such measurements rely on high-energy elementary particles loosing energy by multiple interactions with the QGP. The addition of FoCal to the ALICE experimental setup would greatly enhance, in high-energy (multi)hadronic collisions, the access to the state preceding the formation of the QGP, a state which should be strongly dominated by gluon saturation.

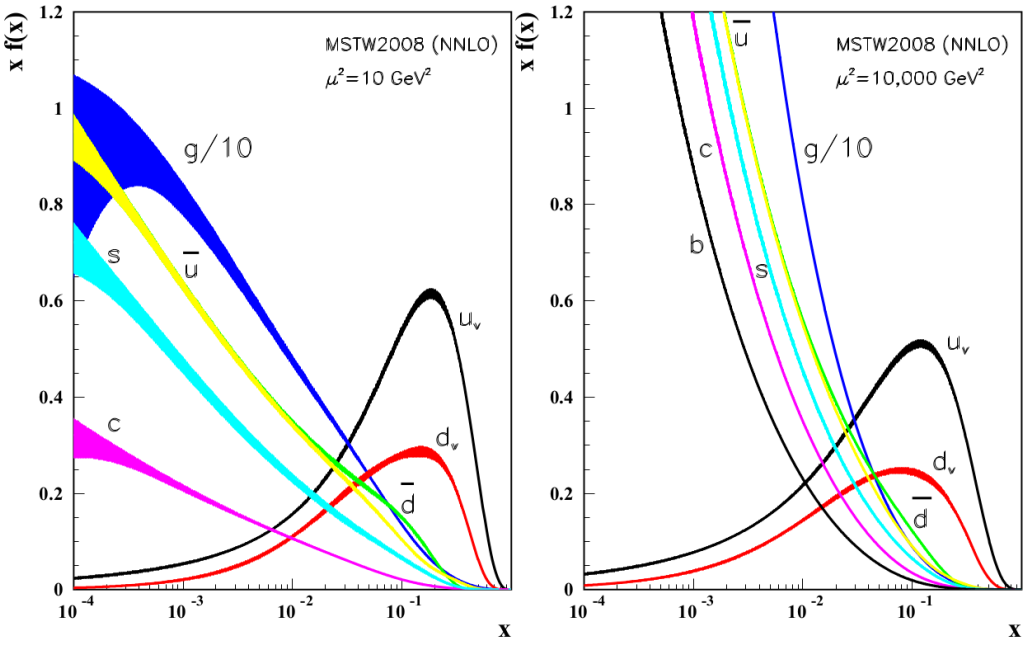

Figure 1: Proton PDFs for valence up quarks u_v, valence down quarks u_d, down antiquarks $\bar{d}$, up antiquarks $\bar{u}$, strange quarks s, charm quarks c and gluons g, as a function of momentum fraction, for a momentum transfer $Q^2 = 10\text{GeV}$ (left) and $Q^2 = 10^5\text{GeV}$ (right). – copied from Chap. 16 of [3].

QCD essentially describes interactions between quarks and gluons. The capability of QCD to predict hadronic scattering cross-sections relies on a-priori knowledge about the structure of the interacting hadrons. This knowledge is encoded by parton distribution functions (pdf's) and is typically gained from DIS experiments, where electrons are usually used for probing the interior of the hadronic target, via exchanges of virtual (space-like) photons described by leading-order quantum electrodynamic (QED)

processes. The discussion is restricted here to protonic targets. There exists one pdf for every type of parton possibly contained in the proton. Mathematically, a pdf is defined with respect to:

- Q^2 – the (positive) squared norm of the 4-momentum of the exchanged (virtual) photon,
- x – the fraction of the 3-momentum of the proton carried by the respective type of parton, with respect to the reference frame where the energy entry of the 4-momentum of the exchanged (virtual) photon has a value of 0 (the Breit frame).

At a certain value of Q^2, each pdf xf provides, as a function of x, the probability that the respective type of parton is probed with the respective momentum fraction – between x to $x + dx$. Figure 1 shows a recently obtained set of pdf's, for $Q^2 = 10$ GeV and $Q^2 = 10^5$ GeV, with $x \in [10^{-4}, 1]$, the most important ones for the purpose of this work being those of the valence up quarks xu_v, valence down quarks xd_v and gluons xg, where the latter is downscaled by a factor of 10 for visual purposes. At this point, it is worth noticing that gluons xg are increasingly more likely to be probed inside protons with decreasing momentum fraction x.

One can think of the pdf's as encoding the large distance (low momentum) behavior that cannot be calculated in perturbative QCD. However, QCD does predict how these distributions (including the gluonic ones) evolve with Q^2, via the DGLAP evolution equations [7], given that measurements for some Q^2 and x are available. This procedure takes into account the fact that (anti)quarks may split into quark-gluon pairs before being probed, together with a few other (not all) 3-leg interactions involving quarks and gluons allowed by QCD. The procedure does not take into account the gluon fusion process allowed by QCD, according to which two gluons merge into one gluon. The rate of this process is proportional to the gluon pdf $xg(x, Q^2)$ and to the gluon fusion cross section $\sigma_{gg\to g} \approx \frac{\alpha_S}{Q^2}$. As can be seen in Fig. 1, $xg(x, Q^2)$ strongly increases with decreasing x. At low x, it also decreases with increasing Q^2, however not enough to compensate for the decrease of $\sigma_{gg\to g}$ with Q^2, implying that the gluon fusion process becomes significant for low x and low Q^2, when $xg(x, Q^2)\sigma_{gg\to g} \geq 1$, beyond which the gluon saturation (or color glass condensate) regime is expected to set in [8].

For such values of x and Q^2, the failure of the DGLAP due to the unaccounted gluon fusion processes would be manifested by a taming of the increasing tendency of $xg(x, Q^2)$ with decreasing x mentioned before – see Chap. 16 of Ref. [3] and references therein, which would otherwise entail a violation of the unitarity condition for the hadron-hadron cross section [9], [10]. Verifying the existence of the gluon saturation regime while probing the gluonic pdf at low x and low Q^2 constitutes the main motivation behind the FoCal detector.

In the context of capabilities at LHC, Investigating gluon saturation can be achieved by measuring its effect on the rates of prompt photon production via mixed QED-QCD processes, in proton-proton versus proton-lead collisions. The notion of prompt photon refers to photons produced via the 2 processes shown on the left hand side of Fig. 2, showing the quark-gluon Compton process and the quark-antiquark annihilation, which are variations of the same Feynman diagram. For most of the x interval, the former is much more frequent than the latter, since gluons are much more abundant in nucleons/nuclei than antiquarks, as can be roughly understood from Fig. 1 by comparing the xg curve against the $x\bar{u}$ and $x\bar{d}$ curves. Experimentally accessible information about the prompt photon provides partial information about the gluon involved in the interaction [1]. The two diagrams on the right hand side of Fig. 2 show two important sources of background for the prompt photon. The first one is called bremsstrahlung (also referred to as quark fragmentation) [2], describing the interaction between a quark from one nucleon/nucleus and one from the other nucleon/nucleus, with a photon being emitted by one of the two. The second one

[1] Systematically measuring the jet produced by the other quark/gluon in the final state provides more complete information about the gluonic pdf, via the so called photon-jet coincidence measurements. This aspect is not treated in the present study

[2] In analogy to the similar process completely governed by QED, where the gluon line is replaced by a photon line and the two quark lines can be replaced by any quark or other electromagnetically charged lepton

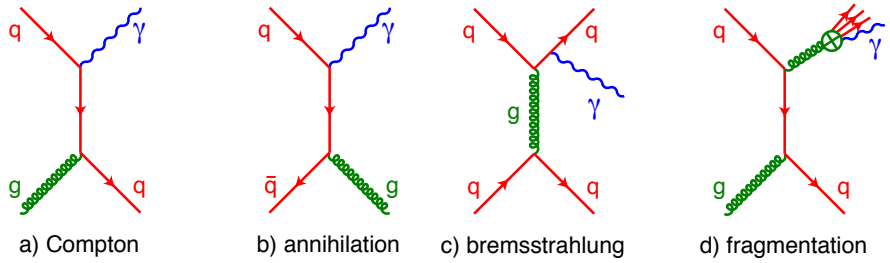

Figure 2: Main Feynman diagrams contributing to photon production. At higher order, one deals with quark-gluon Compton process (a) and quark-antiquark annihilation (b). At next-to-leading-order, one has non-isolated photon emission from bremsstrahlung of a quark on another quark via the gluonic field (c), and from gluon fragmentation process (d). In these diagrams, the legs on the left hand denote the partons in the initial state, with each parton originating from one of the nucleons/nuclei involved in the collision, while the legs on the right side denote the partons in the final state – copied with permission from [11]

is called fragmentation, according to which a photon is produced from the jet resulting from a gluon in the final state. At least for the present study, the term "direct" is used for referring to photons generated via any of the four processes depicted in Fig. 2, while the attribute "prompt" is solely used for the two leading-order processes on the left. Detecting prompt photons and measuring the associated yields motivates the use of an electromagnetic calorimeter – further information related to calorimetry is given in section 2.5.

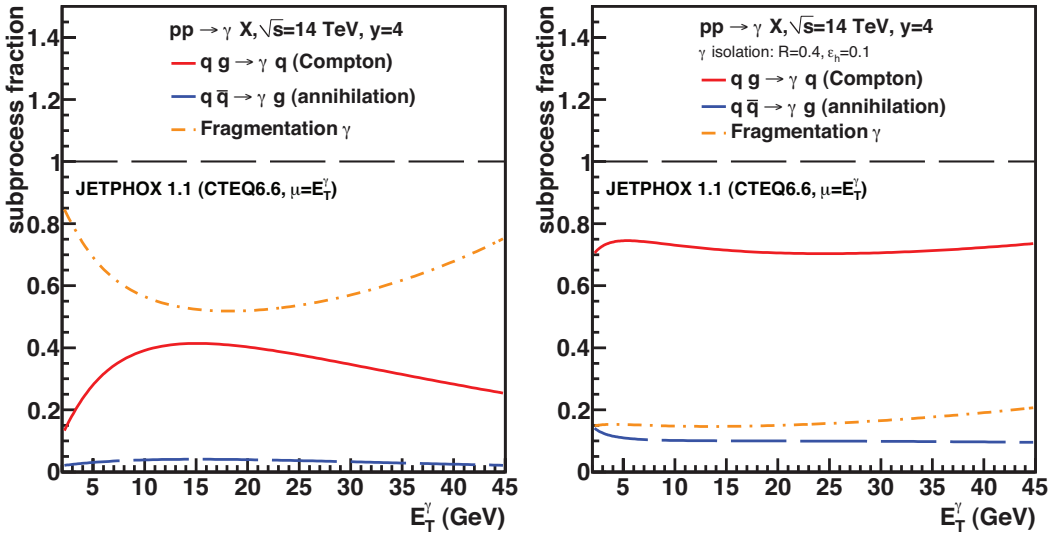

Figure 3: Relative contribution to NLO prompt photon production of the individual subprocesses shown in Fig. 2 as a function of photon transverse energy, at center-of-mass energy $\sqrt{s} = 14$TeV and rapidity $y = 4$, without an isolation cut (left) and with an isolation cut (right), computed with the JETPHOX program – copied from [12]

For the above two sources of background, much of the jet produced together with the direct, non-prompt photon can usually be found within a narrow cone around the direction of the photon. This implies that an isolation cut may effectively be employed for eliminating most of the background. Fig. 3 provides quantitative information about such a cut, based on proton-proton simulations produced with the JETPHOX program. There, the red and blue lines show, as a function of transverse photon energy, the relative contributions to the production of direct photons via the prompt photon channels, namely Compton and annihilation respectively. The dashed orange line shows the combined contribution of

the direct, non-prompt photon channels referred to in Fig. 2 as bremsstrahlung and fragmentation – bremsstrahlung is understood in Fig. 3 as a partial contribution to fragmentation photons. The left hand side shows the pure profiles as obtained from the simulation, while the right hand side shows the curves after applying an isolation cut on the energy content within a narrow cone around the direct photon. One can clearly see that the relative contribution of the direct, non-prompt photon is dramatically reduced. The transverse energy range and the rapidity window $y = 4$ are adapted for forward measurements at the LHC, which are compatible with the geometry of the proposed FoCal detector – see Sec. 2.

In addition, the other main source of background is represented by decays of neutral mesons π_0 and η, which both mainly decay in two photons. For the purpose of this study, only the π_0 background is considered, given that its contribution is much larger. At high energies, double photon decays are highly boosted in the laboratory frame compared, such that at high rapidities the directions of the two photons are separated by small angles and usually both hit the detector. Such background can thus be rejected by performing invariant mass analysis, with a veto on shower pairs with an associated reconstructed invariant mass within a window around the known mass of the neutral pion $m_{\pi_0} = 0.135\text{GeV}/c$. For higher energies, when the two photons hit the detector so close to each other that only one shower is identified, shower shape analysis can be used for rejecting part of this background. More details concerning kinematic aspects are given in section 2.2. Performing such studies at the TeV scale motivates the use of a calorimeter with good energy resolution and excellent transverse position granularity.

From the perspective of calorimetry, electromagnetic showers very similar to those caused by high-energy photons can also be caused by high-energy electrons or positrons. However, in proton-proton or proton-lead collisions, the direct production of high-energy (hundreds of GeV) electrons or positrons has a very low rate compared to direct photon. Electron-positron production rates comparable to the rate of prompt photon production are achieved via decays of the J/Ψ meson. In principle, this background source can also be reduced via invariant mass analysis. For the purpose of the present study, the J/Ψ background is neglected, as neutral pions are much more abundant. This is also a consequence of phase-space considerations, due to the much higher mass of the J/Ψ meson ($3.097\text{GeV}/c^2$) compared to the mass of the neutral pion ($0.135\text{GeV}/c^2$).

For an interaction between two partons contained by nucleons belonging to opposing beams, one can show that there exists a simple kinematic correspondence linking the momentum fractions of the two partons in the initial state to experimentally accessible kinematic variables of the 2 two particles in the final state. The momentum fraction x_1 of the more energetic parton, the momentum fraction x_2 of the less energetic parton and the momentum transfer Q^2 between the two relate to the rapidities γ_3, γ_4 of the two outgoing particles and their transverse momentum p_T in the following way:

$$x_1 = \frac{p_T}{\sqrt{s}}(e^{y_3} + e^{y_4}) \tag{1}$$

$$x_2 = \frac{p_T}{\sqrt{s}}(e^{-y_3} + e^{-y_4}) \tag{2}$$

$$Q^2 = \frac{1}{4}x_1 x_2 s \tag{3}$$

where $\sqrt{s}$ is the center of mass energy, given that the initial partons travel along the beam axis in opposite directions. As explained above, one is interested in probing the low Q^2 and low x region of the gluonic pdf, Based on the three equations above, one can easily check that both Q^2 and x_2 are minimized for large $y_3 \approx y_4 > 0$, when both particles in the final state exit the interaction on the same $z > 0$ side [3], under approximately equal polar angles $\theta_1 \approx \theta_2$ (when their masses are negligible, such that their rapidities approximately match their pseudorapidities). For this scenario, one can simply write:

[3] Where z is the Cartesian coordinate along the beam axis, as explained in the following section.

$$x_2 = x_T e^{-y} \tag{4}$$

$$Q^2 = (\frac{1}{2} x_2 e^y \sqrt{s})^2 = p_T^2 \tag{5}$$

for $y := y_3 = y_4$, where $x_T = \frac{2p_T}{\sqrt{s}}$. On the other hand, the gluonic and anti-quark pdf's strongly increase with decreasing x, making the rate of direct photon production maximal for the large $\gamma_3 \approx \gamma_4 > 0$ scenario. In turn, this implies that, for measurements of inclusive, isolated photons, after the invariant-mass and shower shape cuts, on the $z > 0$ side of the interaction point, equations 4 and 5 can be used as reasonable approximations for x_2 (momentum fraction of the lower-momentum parton) and Q^2, by identifying y and p_T with the measured rapidity and transverse momentum of the photon – without measuring the jet produced from the second particle in the final state, thus without enforcing the $y_3 = y_4$ constraint.

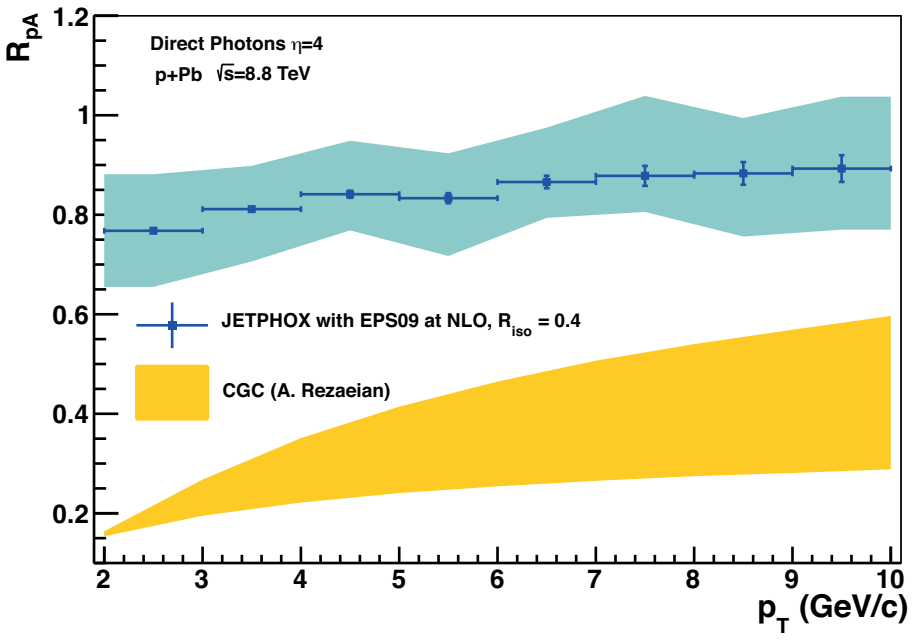

Figure 4: Predictions of γ_{dir} nuclear modification factor R_{pPb} as a function of transverse momentum p_T in proton-Lead collisions with a center-of-mass energy of $\sqrt{s} = 8.8$TeV, for a pseudorapidity $\eta = 4$, from perturbative QCD at NLO (blue) and a Color Glass Condensate mode (yellow), the latter being based on a model from [13] – copied with permission from [11].

Based on these arguments, the presence of gluon saturation may be observed by studying the nuclear modification factor R_{pA} for direct photons as a function of (low) transverse momentum p_T, for a specific window of high rapidity γ (low x_2), for low transverse momenta. The presence of gluon saturation would significantly reduce the production of direct photons in pA collisions compared to pp collisions, after taking into account the increased number of nucleons involved, due to higher partonic densities in atomic nuclei. In this manner, the measured nuclear modification factor R_{pA} would be significantly lower than the value predicted by perturbative QCD, an effect which should increase with decreasing p_T (decreasing Q^2). This expected effect is shown in Fig. 4, where the pQCD prediction is compared to a prediction of a Color Glass Condensate model.

The nuclear modification factor R_{pA} is an experimental observable typically used in high-energy physics for quantifying non-linear effects related to the presence of more than one nucleon-nucleon pair in the initial state, compared to the baseline scenario involving only one nucleon-nucleon pair. When evaluated via measurements of direct photons γ_{dir}, this quantity can be mathematically expressed as:

$$R_{\mathrm{pA}} = \frac{\frac{1}{p_T}\frac{dN_{\mathrm{pA}}^{\gamma_{\mathrm{dir}}}(b)}{dp_T}}{\langle N_{\mathrm{coll}}(b)\rangle \frac{1}{p_T}\frac{dN_{pp}^{\gamma_{\mathrm{dir}}}}{dp_T}} \tag{6}$$

Here, $\frac{1}{p_T}\frac{dN_{\mathrm{pA}}^{\gamma_{\mathrm{dir}}}(b)}{dp_T}$ is the measured invariant yield of γ_{dir} production in pA collisions, as a function of centrality b. Similarly, $\frac{1}{p_T}\frac{dN_{pp}^{\gamma_{\mathrm{dir}}}}{dp_T}$ is the measured invariant yield in pp collisions, while $\langle N_{\mathrm{coll}}(b)\rangle$ is the number average number of nucleon-nucleon collisions. In the context of the LHC, Lead ions would be used as a specific type of heavy nucleon A := Pb, thus leading to measurements of R_{pPb}.

According to the argumentation above, measuring direct photons yields at high rapidities and low transverse momentum provides access to gluon saturation effects. A successful measurement relies on separating the γ_{dir} signal from the main source of background after applying an isolation cut, which is represented by neutral pions. Separating the electromagnetic showers of the two γ's produced by a π_0 motivates performance studies of a calorimeter with fine granularity layers, understanding the improvement granted by such layers and perfecting shower-shape analysis tools, solely based on simulations. These are the aspects that most of this work is focusing on.

As an outline for the rest of this manuscript, Sec. 2 describes the current FoCal design as used here, in the context of the ALICE experimental setup and in the context of electromagnetic calorimetry. Sec. 3 explains the simulation chain, from the level of detector response to the level of analysis. Sec. 4 explores γ vs π_0 discrimination in single particle simulations, via invariant mass analysis and shower shape analysis, allowing for certain decisions to be made between different competing analysis procedures and observables. With such decisions being made, Sec. 5 further explores the impact of several analysis scenarios on the detector performance in simulated proton-proton collisions. Finally, Sec. 6 concludes the study.

2 Detector Concept

This section presents the design and underlying detection principles of the current FoCal concept. Specific design details mentioned here, such as dimensions or number of layers are in accordance with actual values used for producing simulations described in Section 3. Although these are kept fixed for the present work (and document), they are subject to change, as the R&D process is still ongoing.

2.1 The ALICE experiment

The ALICE detector [14] was designed to provide access to experimental observables relevant to the study of the QGP in high-energy collisions. It essentially consists of a central barrel with a solenoid magnetic field, aligned with the Z-axis, and a muon arm with a dipolar magnetic field aligned with the X-axis. The central barrel has a onion-like structure, surrounding the interaction point, each cylindrical layer being a different sub-detector. The muon arm consists of an absorber and several tracking stations for muon. In both cases, tracking and measurements of particle momenta are allowed by the existence of the magnetic fields bending the trajectories in the XY and YZ planes respectively. The Z-axis points along the beam against the muon arm. Figure 5 shows the overall setup of the ALICE experiment, with the muon arm towards the right.

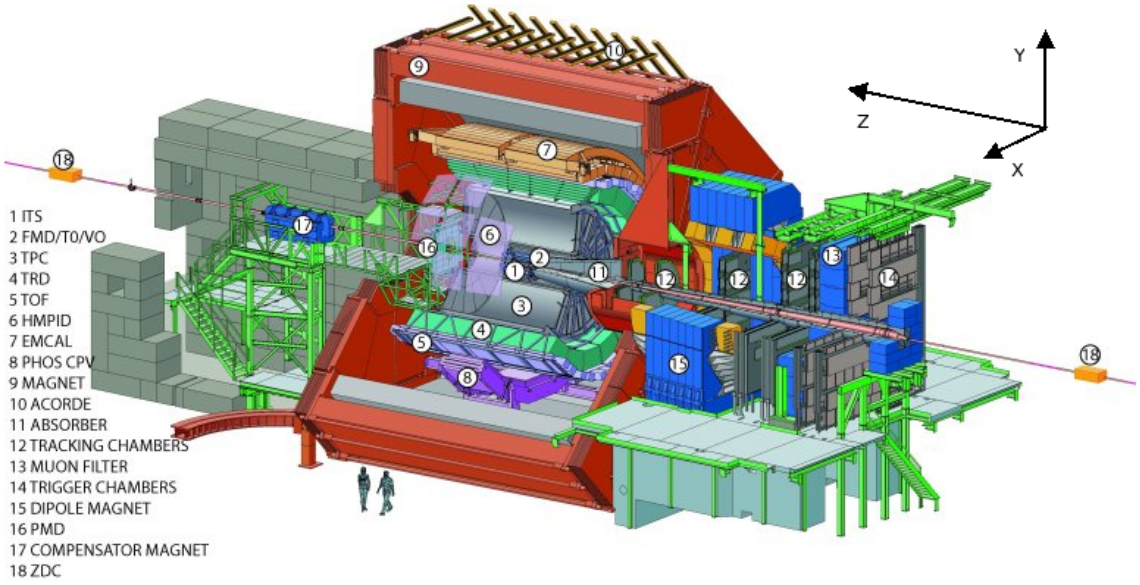

Figure 5: Schematic view of the ALICE experimental setup, with the abbreviations of its main subcomponents and the commonly agreed Cartesian axes, with the Z-axis pointing along the beam axis, away from the muon arm. The origin of the Cartesian system is assigned to the interaction point, the polar angle θ is measured with respect to the Z-axis, while the azimuthal angle ϕ is measured with respect to the X-axis, according to the right-hand rule along the Z-axis – adapted from [1].

2.2 Kinematic Considerations

The main task for the FoCal detector is the measurement of prompt photons produced in proton-proton and proton-lead collisions delievered by LHC to the interaction point of the ALICE experiment. This measurement has to be done at high absolute values of pseudorapidity $|\eta|$, over a large interval of transverse momentum p_T. For this reason, FoCal has to be positioned around the beam pipe, perpendicular

to the beam (Z-)axis.

The pseoudorapidity range covered by FoCal is directly determined by the distance between its center and the interaction point, by its inner and outer radii. In particular, if the inner and outer radii are considered fixed, as they are constrained to a great extent by the radius of the beam-pipe and by weight and cost limits respectively, pseudorapidity ranges are:

$$\eta_{\text{min}} = -\log\left(\tan\left(\frac{\theta_{\text{min}}}{2}\right)\right), \quad \eta_{\text{max}} = -\log\left(\tan\left(\frac{\theta_{\text{max}}}{2}\right)\right), \tag{7}$$

with the polar range given by:

$$\theta_{\text{min}} = \tan\left(\frac{r_{\text{min}}}{z}\right), \quad \theta_{\text{max}} = \tan\left(\frac{r_{\text{max}}}{z}\right), \tag{8}$$

where z is the distance from the interaction point (along the beam axis), r_{min} and r_{max} are the inner and outer radii.

Figure 6 illustrates the correspondences between z, r_{min}, r_{max}, η_{min} and η_{max}. It can be seen that η_{max} increases with z and decreases with r_{min}, while η_{min} increases with z and decreases with r_{max}.

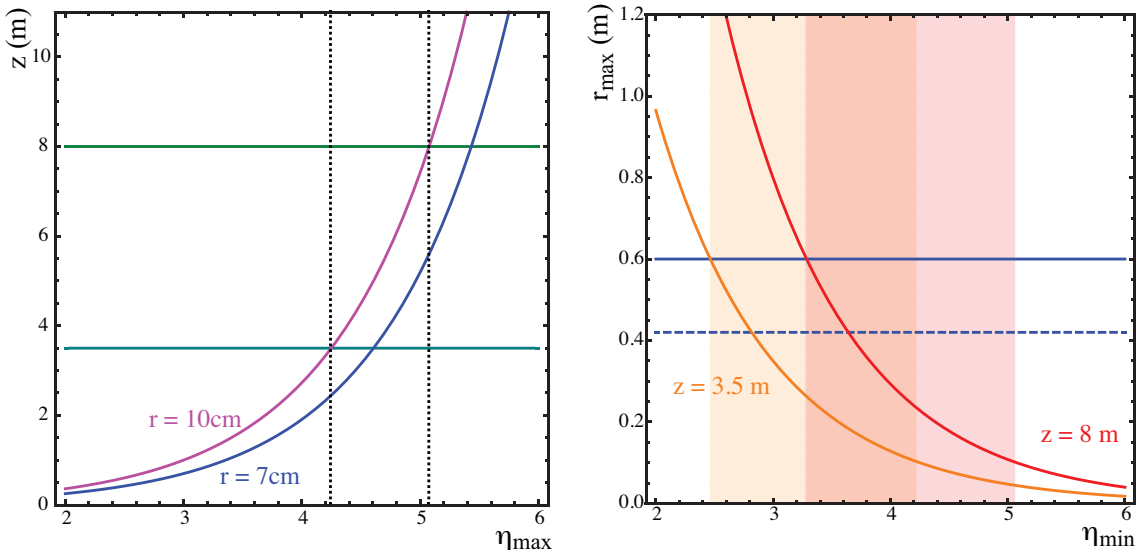

Figure 6: Relationship between z-coordinate of FoCal and upper pseudorapidity limit η_{max} for two values of the inner radius r_{min} (left) and between the outer radius r_{max} and the lower pseudorapidity limit η_{min}, for the two possible FoCal locations along the Z-axis – copied with permission from [11].

Given that the most important source of background for direct photons γ_{dir} are decay photons γ_{dec} generated from decays of neutral pions π_0, it is imperative that such decays are identified and rejected. When the π_0 energy E_{π_0} is large (in the order of hundreds of GeV), the strong Lorentz boost translating the decay from the π_0 center of mass frame to the lab frame yields very small opening angles between the two γ_{dec} (typically less than 1 degree). This makes the π_0 reconstruction difficult, due to the tendency of the two electromagnetic showers to overlap. Fig. 7 shows this effect more quantitatively, where the decay asymmetry is defined as:

$$\alpha^{\gamma-\gamma} = \frac{|E_{\gamma_1} - E_{\gamma_2}|}{|E_{\gamma_1} + E_{\gamma_2}|} \tag{9}$$

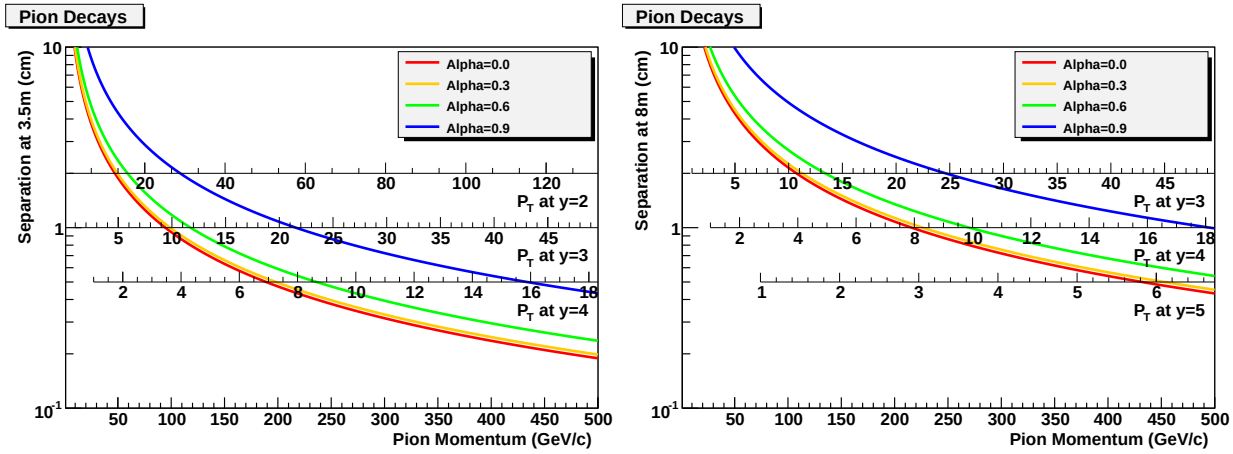

Figure 7: Distance between the incidence positions of the two photons originating from a π^0 decay, as a function of the total pion momentum, for different decay asymmetries α, for FoCal at 3.5 m (Left) and at 8 m (Right). The additional scales show, for several values of rapidity y, the transverse momentum p_T corresponding to the respective total momentum value p – copied with permission from [11].

where E_{γ_1} and E_{γ_2} are the energies of the two γ_{dir}. Instead of the opening angle, the figure shows the separation between the two decaying photons $d_{\gamma-\gamma}$ within the detector plane, for two possible detector locations along the z-axis. It is obvious that the separation decreases dramatically with the pion momentum p_{π_0} – thus with its energy, since $E_{\pi_0} = \sqrt{p_{\pi_0}^2 + m_{\pi_0}^2}$, where m_{π_0} is the invariant mass of the pion. Furthermore, the separation also increases with decreasing α, such that symmetrical decays are most difficult to separate, for a given pion momentum. Finally, the separation increases with increasing Z, which makes large z distance more attractive from this point of view. The curves were obtained using only kinematic considerations related to the π_0 decay, such as conservation of 4-momentum and a Lorentz boost along the initial π_0 flight direction. In order to compute the separation, the simplifying assumption that the π_0 decays along the Z-axis has been used – this is a stronger condition, given that misaligned π_0 would lead to somewhat larger separations. Under this assumption, the separation as a function of momentum p_{π_0}, decay asymmetry $\alpha^{\gamma-\gamma}$ and detector position z is given by:

$$D^{\gamma-\gamma} \approx \frac{2 m_{\pi_0} z}{\sqrt{\left(p^2 + m_{\pi_0}^2\right)\left(1 - \alpha^2\right)}}, \tag{10}$$

For completion, the transverse momentum p_T corresponding to the indicated value of momentum p for a certain rapidity y is also shown, although the implied angular displacement from the Z-axis is not being taken into account for computing the separation (which essentially assumes $y = \infty$). Rapidity is defined similarly to pseudorapidity, except that it also depends on the invariant mass m of the respective particle, besides the polar angle θ, whereas the latter does not. However, in the limit of $m << E$, which is the case for pions ($m_{\pi_0} \approx 0.14\text{GeV}/c^2$) in the range of $E \approx 100\text{GeV}$, the notions of rapidity and pseudorapidity coincide $y \approx \eta$. It is worth noting that the π^0 decays are uniform with respect to α, such that 10% of the π^0 decay with $\alpha < 0.1$, 10% of the π^0 decay with $0.1 < \alpha < 0.2$ etc. This is a direct consequence of the isotropy of (back-to-back) π_0 decays in the center of mass frame.

2.3 Detector Location

In the context of the ALICE experimental setup, the geometrical requirements mentioned can be conveniently satisfied by two locations. Both these locations take advantage of the relatively free space around the beam axis on the side opposite from the muon arm, at positive Z values (what is commonly referred to as the A-side). The first possibility is to install FoCal at 3.60 m from the interaction point (this is

called here the 4m position), attached to the main frame of the central barrel, replacing of the Photon Multiplicity Detector (PMD). This possibility is schematically shown in Fig. 8. The second possibility is to place FoCal at 8m from the interaction point, sustained by the support designed for the compensator magnet, as shown in Fig. 9, after displacing the latter to higher z values. Throughout the rest of this document, the two positions will be referred to simply as the 4m and the 8m positions. Fig. 6 shows η acceptance ranges provided by both these positions, while Fig. 7 shows the corresponding separations between two γ_{dec} originating from a π_0.

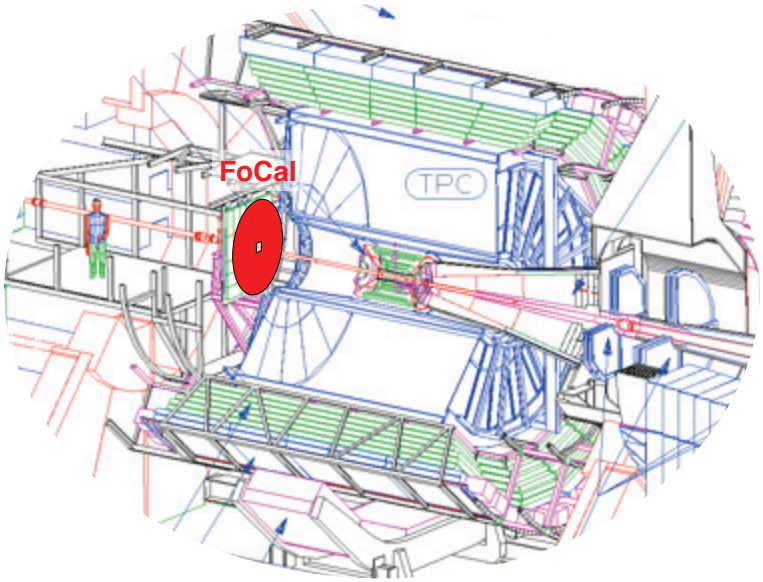

Figure 8: Sketch of the 4m FoCal position within the ALICE experimental setup – copied with permission from [11].

The 8m position is technically more challenging than the 4m position, as it requires drastic changes of the beam pipe and auxiliary equipment opposite to the muon arm. However, the 8m positions has several important advantages. First, the value $\eta_{\max} \approx 5.3$ at 8m is much larger than $\eta_{\max} \approx 4.4$ at 4m. According to section 1, this gives better access to the kinematic region where gluon saturation effects are expected to be significant. Moreover, particle densities at the detector level would increase with increasing separation from the interaction point, leading to improvements in the analysis. Second, the supporting structure currently holding the compensator magnet at the 8m position would also be able to withstand the weight of the hadronic calorimeter component of FoCal, which would greatly enhance the performance of isolation cuts. For these reasons, the 8m location is preferred by the ALICE-FoCal collaboration. Nonetheless, all studies presented in this thesis make use of Monte Carlo simulations generated for the 4m positions, due to pragmatical considerations concerning the simulation time.

2.4 Electromagnetic showers

When interacting with dense matter, both photons and electrons may give rise to what is known as an electromagnetic cascade, also referred to as an electromagnetic shower. At low energies, both photons and electrons interact with matter primarily via ionization, which essentially takes away energy from the incoming particles and gives it to electrons in the material. Through ionization, energy is essentially deposited in a very confined space along the incoming particle's trajectory. At higher energies however, other processes become important, which entail the production of new energetic particles. In the presence of the nuclei's strong electromagnetic fields, photons produce electron-positron pairs via pair production, while electrons give birth to secondary photons via bremsstrahlung. Pair production and bremsstrahlung are both QED processes described by variations of the same Feynman diagram. Depending on their energies, secondary particles generated in this manner produce further particles via these two processes,

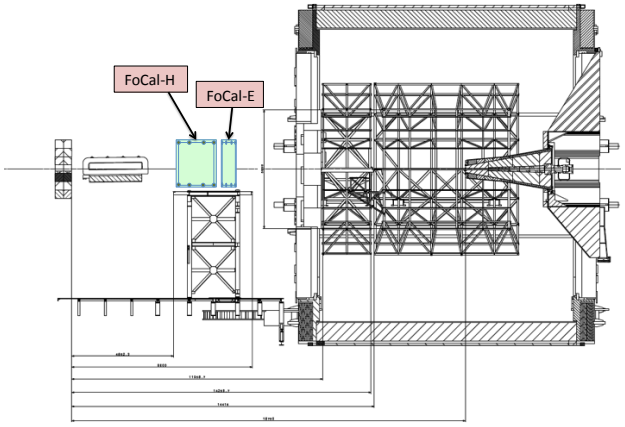

Figure 9: Sketch of the 8m FoCal position within the ALICE experimental setup – copied with permission from [11].

thus giving rise to the electromagnetic cascade or shower. Typically, showers develop within cone-like regions of space, with the vertex at the first interaction of the incoming particle. The opening angle of the cone decreases with increasing energy of the original particle, while the depth increases with increasing energy, given that secondary particles are in turn generated with more energy and thus higher likelihood of generating more particles. The energy below which particles within the cascade no longer give rise to further particles is referred to as the critical energy, more precisely defined below.

The notion of radiation length is typically used as a measure of the distance required for significant changes to appear in an electromagnetic shower. In particular, the radiation length X_0 is defined as the average distance that a high-energy electron typically travels through a specific material in order to be left only with $1/e$ of its energy, due to bremsstrahlung. It is also $7/9$ of the average distance that a high-energy photon has to travel through a specific material before producing an electron-positron pair [3, 15]. This is a material constant, depending on the atomic number, mass number and density of absorber, when measured in cm. It is common in literature to use the notion of radiation length for the product $X_0\rho$, measured in g/cm^2, where ρ is the density of the absorber, with X_0 depending on density. Taking out the density dependence, this definition is convenient, for instance, when minimizing detector thickness and detector weight together, while making a choice between different absorbing materials. Here, radiation length uses the former definition with units of cm. For example the radiation length is $X_0(\mathrm{W}) = 0.3504$ cm for Tungsten, $X_0(\mathrm{Pt}) = 0.3051$ cm for Platinum and $X_0(\mathrm{Pb}) = 0.5612$ cm for Lead [16]. The similar notion of interaction length λ_I is used for characterizing hadronic showers. It is defined the same as radiation length is defined for electrons, by replacing the concept of bremsstrahlung with that of inelastic nuclear interaction. The interaction length is $\lambda_I(\mathrm{W}) = 9.946$ cm for Tungsten, $\lambda_I(\mathrm{Pt}) = 9.125$ cm for Platinum and $\lambda_I(\mathrm{Pb}) = 17.59$ cm for Lead [16]. Note that the interaction length is much larger than the radiation length for all these materials, which implies that hadronic showers start developing much later than electromagnetic ones.

The critical energy E_c is another essential concept for understanding electromagnetic cascades. According to the definition by Rossi [17], it is the electron energy satisfying the condition: $\left|\frac{dE}{dx}\right|_{\mathrm{ioniz}} = \frac{E}{X_0}$. It is also an approximation of the energy at which the electron energy loss via bremsstrahlung overcomes the energy loss via ionization. Below this energy, an electron typically no longer contributes to the development of the electromagnetic cascade. The critical energy tends to be slightly lower for positrons than for electrons. Quantitatively, the electron's critical energy is $E_c(\mathrm{W}) = 7.97$ MeV for Tungsten, $E_c(\mathrm{Pt}) = 7.59$ MeV for Platinum and $E_c(\mathrm{Pb}) = 7.43$ MeV for Lead [16].

For describing the lateral development of electro-magnetic showers, the notion of Moliere radius is commonly used. More precisely, the Moliere radius R_M is defined as [3, 18, 19]:

$$R_M = \frac{X_0 E_s}{E_c} \tag{11}$$

where the scale energy $E_s = 4\pi e m_e c^2 \sqrt{2\pi\epsilon_0 \hbar c} = 21.2052$ MeV is a combination of universal constants, E_c is the critical energy and X_0 is the radiation length. The Moliere radius attains values like $R_M(\mathrm{W}) = 0.9327$ cm for Tungsten, $R_M(\mathrm{Pt}) = 0.8527$ cm for Platinum and $E_c(\mathrm{Pb}) = 1.602$ cm for Lead [16]. Approximately 90% of the energy deposited by an electromagnetic shower is contained in a cylinder with radius R_M around the original track of the particle.

2.5 Detection Principles

Conceptually, calorimetry is a specific approach for detecting particles based on the energy deposited by their showers, which have to be contained in the detector as much as possible. Depending on the type of shower for which it is designed, a calorimeter can be either electromagnetic or hadronic [4]. Depending on whether it consists entirely or partly of sensitive material, it can be either a homogeneous or a sampling calorimeter.

Optimally (see Section 2.3), the Focal detector would have one electromagnetic and one hadronic component, referred to as FoCal-E and FoCal-H [5]. They would be placed one after the other in this order, since electromagnetic showers tend to develop much earlier than hadronic ones – the radiation length is much smaller than the interaction length. This is a typical design choice also made, for instance, by the LHCb collaboration [20] Both components would be sampling calorimeters. Typically, a sampling calorimeter makes use of metallic, absorbing plates with sensitive plates sandwiched in between. The signal is generated based on energy deposited in the sensitive plates. This provides longitudinal information about shower development. Transverse information is provided by transverse segmentation of the sensitive plates. The capability of providing information about the energy of the showering particle relies on a preliminary calibration of the calorimeter to particles of known energies.

Along the longitudinal direction, the thickness of a slice is usually chosen to match the radiation length of the absorber, in the case of electromagnetic calorimeters and the interaction length in the case of hadronic calorimeters. The granularity for transverse segmentation is typically chosen to match the shower half width, which for electromagnetic calorimeters is approximately given by the Moliere radius. In this manner, most showers are contained by a few transverse segments. For more details concerning calorimetry see [21].

The crucial aspect for FoCal-E is the possibility of distinguishing direct photons from decay photons. Decay photons are mainly produced in pairs, via decays of neutral mesons π_0 and η, the former being produced much more frequently. As discussed in Section 2.2, the angle between the two decaying photons in the lab frame decreases with increasing π_0 energies, thus causing an increasing overlap between the two showers. As shown if Fig. 6, the separation between the two photons of a decaying pion in the detector plane easily reaches and even goes below typical values of the Moliere radius mentioned in section 2.4. At that point, disentangling the two photon showers using a transverse granularity in the order of the Moliere radius is highly difficult. For this reason, high granularity sensitive layers, with transverse size going much below the Moliere radius, are required for separating and analyzing electromagnetic showers in detail. Digital silicon pixels sensors with transverse size in the order of tens to hundreds of microns can be employed for this task. However, building an entire electromagnetic calorimeter from high-granularity

[4] This section mostly focuses on concepts related to electromagnetic calorimetry, as they are much more important for the present work

[5] At the time of writing this thesis, only the FoCal-E design had been discussed in detail

pixels alone, with the size requirements of FoCal, would also produce much higher amounts of raw data than by employing sensors with larger size but with analogue readout.

Having these considerations in mind, the FoCal collaboration has reached a hybrid calorimeter design, employing two kinds of sensitive components, thus leading to two kinds of layers:

- Low Granularity Layers (LGL), with lateral segmentation chosen to approximately match the absorber's R_M and with analogue readout
- High Granularity Layers (HGL), with lateral segmentation 100 times smaller than the absorber's R_M and with digital readout

The binary, high granularity information would be further integrated electronically within the layer, such that a pseudo-analogue output is obtained – see Sections 2.6 and 3.2. HGL data would then be used mainly for obtaining information about transverse shower position and for separating neighboring showers, while LGL data would be used for obtaining information about shower energy. It is clear that sophisticated clustering and shower-reconstruction algorithms are required for effectively combining LGL and HGL data. Such aspects are treated in detail in Sections 3.3 and 3.4, together with the Appendix [6]

In terms of materials, the FoCal-E makes use of silicon (Si) as an active material, due to the availability of multiple silicon technologies with various granularities and of tungsten (W) as an absorber, due to its relatively small Moliere radius (R_M) and radiation length (X_0) [16]:

$$R_M = 9\text{mm}$$
$$X_0 = 3.5\text{mm}$$

2.6 Segmentation

The current detector design, as used in simulations described in Section 3, consists of 20 parallel slices of tungsten of equal thickness (0.35cm), measuring one radiation length each. Behind each tungsten slice there is a thin region containing the silicon detectors and associated electronics. The notion of a layer refers to the tungsten slice and the active, silicon region following it, the entire detector thus consisting of 20 layers: 2 HGL and 18 LGL.

In terms of transverse segmentation, the LGL make use of 1cm $\times$ 1cm silicon pads with analogue readout. The HGL make use of silicon pixels (MAPS) of 0.01cm $\times$ 0.01cm with digital readout. Each FoCal layer makes use of only one of these technologies, pixels and pads not being mixed within any layer.

Within the HGL, the binary signal from pixels is electronically integrated within high granularity cells (HGC) of 0.1cm $\times$ 0.1cm with a resulting pseudo-analogue output. Within the LGL, in the XY-plane low granularity cells (LGC) of 1cm $\times$ 1cm are used, matching the actual pads in terms of size, with an analogue output. However, low granularity layers are combined into low granularity segments, within which signal is electronically integrated along the Z-axis, such that the output of every LGC consists of the summed response of 4 or 5 consecutive pads. Conventionally each HGL is also assigned to one high granularity segment. Data processing is ultimately conducted using the output of segment-level cells:

- square HGC of 1x1 mm with integer amplitudes spanning between 0 to 100, making use of energy deposited over 1 X_0

[6] It turns out that although position information from LGL may be used when matched HGL information is not available, energy information from HGL may not be effectively used when matched, LGL information is not available. This relates to the difficulty of calibrating the response of the HGL.

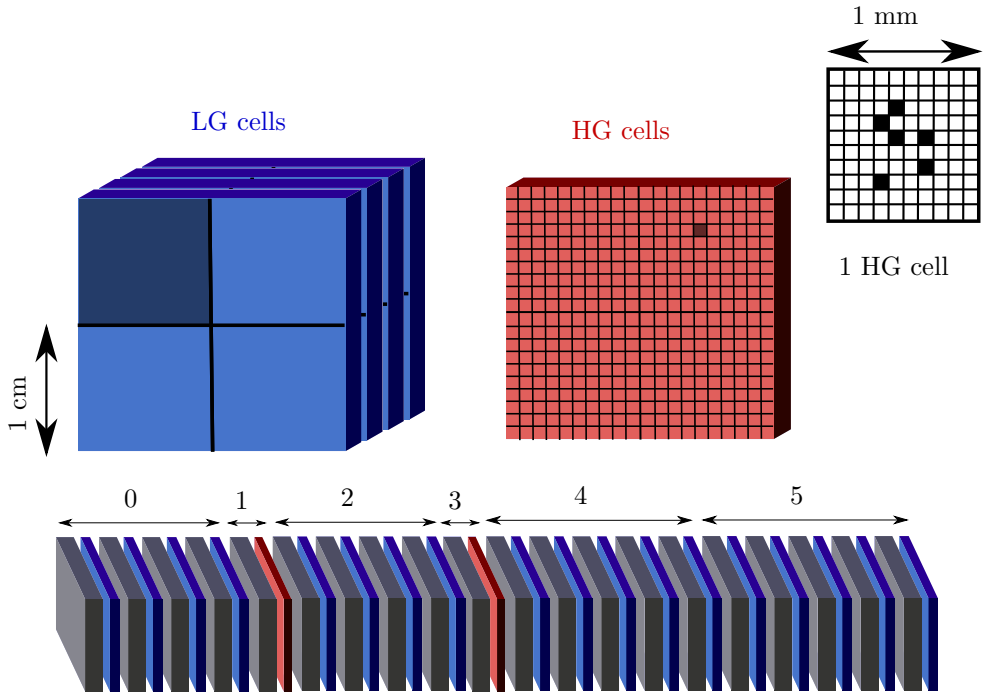

Figure 10: Sketch of current FoCal-E design, with Z-axis pointing horizontally from left to right, with absorbing components shown in grey, low-granularity sensitive components shown in blue and high-granularity sensitive components shown in red.

- square LGC of 1x1 cm with real amplitudes measured in keV, making use of energy deposited over 4 or 5 X_0 (see below)

In the current FoCal-E desing, there are 6 segments (counted between 0 to 5) composed of 4 LGL, 1 HGL, 4 LGL, 1 HGL, 5 LGL and 5 LGL in this order, for increasing values of z. Thus, segments 0,2,4,5 make use of LGC, while segments 1 and 3 make use of HGC. This is also summarized in figure 10.

In the actual detector, the HGC will most probably be built from silicon based Monolithic Active Pixel Sensors (MAPS). These employ the complementary metal-oxide-semiconductor (CMOS) technology, which incorporates, for each pixel, a component sensitive to ionizing radiation together with front-end electronics [22]. The MIMOSA chips, developed by IPHC Strasbourg, provide a good example of this technology being implemented. On the other hand, conventional silicon pad technology would be employed for realizing the LGL cells.

3 Simulation and Data Management

The study presented here is entirely based on results obtained from simulations. Before diving into details, a summary of the entire simulation chain is worth presenting. Each step in this summary is followed by a reference to the actual section where more details are given. First, depending on the purpose, either high-energy single particles originating at the interaction point or particles resulting from proton-proton collisions are simulated, followed by the simulation of the interaction between resulting particles and the material belonging to the ALICE experimental setup, including detectors (Sec. 3.1). Information about the energy deposited in the sensitive layers of FOCAL is then converted into digits (Sec. 3.2). The resulting digits are used for segment-level clustering (Sec. 3.3). Clusters from all segments are used for reconstructing detector-level showers (Sec. 3.4), with the collective digit amplitude information from the low granularity segments being converted to particle energy values, via a calibration function (Sec. 3.6). Detector-level information associated to these showers is used in the higher-level analysis (Sec. 3.8).

3.1 Event and Detector Simulation

Detector simulations were conducted using the ALIROOT v4-20-Rev32 environment [25], which combines the root data analysis framework together with packages containing information about the actual ALICE experimental configuration and with the capabilities of GEANT3 [24] for simulating the interactions between particles and material. Proton-proton collision event generation made use of Pythia 6 [23], while single particle events were generated directly in ALIROOT. The final state particles are then propagated in the space around the interaction point, with bending in magnetic field being taken into account for charged particles. These interactions usually involve further decays and showering of particles originating at the interaction point. Information about the simulated energy deposited in the sensitive parts of the layers used by the simulated FOCAL detector is then saved in the form of hits distributed along the particle tracks. In practice, these hits are simply real-valued numbers encoding the deposited energy in eV.

Volume	Material	Thickness [cm]
Wolfram	Pure W	0.35
Glue	G10	0.05
SiPad	Silicon	0.05
Glue	G10	0.05
Wiring	Cu	0.01
Gap	Air	0.05
Total		0.56

Table 1: Structure and materials used for a generic FoCal layer GEANT3 implementation.

Although the FoCal design makes use of low granularity and high granularity layers, the GEANT3 implementation is the same for all layers, the differentiation between LGL and HGL being only artificially imposed at the level of digitization – Sec. 3.2. Table 2 summarizes the material structure of such a layer used for this implementation. The tungsten absorber Pure W and the sensitive Silicon are the essential ingredients. For an electromagnetic shower, each of these layers has an effective thickness of one radiation length X_0.

Although 50 such layers are simulated, only hits recorded in the first 20 layers are used in the analysis, which implies that effectively an electromagnetic calorimeter of $20X_0$ is used in simulations. The sensitive part of the first layer is positioned at $z_{\text{min}} = 360.985$cm from the interaction point, the sensitive part of the last used layer is at $z_{\text{max}} = 372.185$cm and the sensitive part of the last unused layer is at $z'_{\text{max}} = 388.985$cm. $z_{\text{FCP}} = 374$cm (Focal Center Plane), located approximately in the middle of the 50-layered sandwich (thus behind the used 20 layers) denotes the XY plane with respect to which information

from the different segments used by FOCAL (involving the 20 used layers only) is combined during the shower-reconstruction algorithm – Sec. 3.4. Another plane used for defining certain quantities is: $z'_{\mathrm{FCP}} = 366$cm, located in between the used layers.

3.2 Digitization

This subsection describes how the hits mentioned in Sec. 3.1, generated within the sensitive parts of the layers of FoCal's, are converted to segment level digits. As explained in Sec. 2, the 20 layers of FoCal are grouped into several segments. Let S be the set of simulated detector segments, with cardinality $N_S := |S|$, every segment being labeled by a natural number smaller than N_S: $S = \{s/ \in \overline{0, N_S - 1}\}$, with fixed number of segments $N_S = 6$ for this entire study. Segments are of two types:

- High Granularity Segments (HGS), using High Granularity Layers (HGL)
- Low Granularity Segments (LGS), using Low Granularity Layers (LGL)

Let HGS and LGS denote the subsets of natural numbers in S labeling the respective types of segments, such that: $\mathrm{HGS} \cup \mathrm{LGS} = S$, $\mathrm{HGS} \cap \mathrm{LGS} = \emptyset$. The current FoCal design makes use of 2 HGS and 4 LGS, arranged as shown in Fig. 10, implying that $\mathrm{HGS} = \{1, 3\}$ and $\mathrm{LGS} = \{0, 2, 4, 5\}$. Both HG segments make use of one HGS each. LG segments 0 and 2 make use of 4 LG layers each, while LG segments 4 and 5 make use of 5 LG layers each. All simulations shown here have been carried out using this design, with the detector placed at the 4m position.

The differentiation between the two types of layers is realized by imposing virtual square lattices to each layer, the spacing of the lattice being significantly larger for LGL than for HGL. The two lattice spacings $l^{\mathrm{LGS}} = 1$cm and $l^{\mathrm{HGS}} = 1$mm were kept constant throughout this study.

For the HGL, each cell of its associated virtual lattice is further divided into pixels – 100 identical square pixels are used for every cell of every LGL. The 1mm cells employed by the HGL may also be referred to as "macropixels". The energy of every hit recorded within the space enclosed by such a pixel is first multiplied by a relative thickness coefficient (here kept to a constant value of 0.04), which accounts for the fact that the HG chips are expected to be thinner than the LG pads. Then, if the resulting value is converted to a binary variable based on a digitization threshold (here kept to a constant value of 4000 eV), thus producing a digital output. Finally, for every event, all binary variables associated to every HG are added up to produce an integer-valued signal (here between 0 to 100). In this manner, one HGS is effectively one HGL together with its pseudo-analogue output at the level of every HG cell, thus covering 1 X_0.

On the other hand, sequences of 4-5 consecutive LGL are grouped together to form LGS, while signal from overlapping LG cells are longitudinally integrated within the respective segment. In this manner, the energies of hits occurring in the sensitive parts of any of the 4-5 overlapping cells are added up without any further processing, besides a conversion from eV to keV. In this manner, one LGS is effectively a group of 4-5 LGL together with its analogue output at the level of every longitudinally integrated LG cell, thus only covering 4-5 X_0.

The output provided by the cells of the HGS and LGS are generically referred to as digits. Let D^s be the set of digits recorded during a specific event [7], in segment s, with cardinality $N_D^s := |D^s|$, with individual elements labeled as d_i^s:

$$D^s = \{d_i^s / i \in \overline{0, N_D^s - 1}\}, \tag{12}$$

[7] For the rest of this section, all description is valid for every single event, as individual simulated events do not influence each other in any way. All algorithms and computations below are carried out for every simulated event.

Let ϵ_i^s be the amplitude of digit d_i^s, with $\epsilon_i^s \in \{0, 1, ..., 100\}, \forall s \in$ HGS and $\epsilon_i^s \in \mathbb{Q}_+, \forall s \in$ LGS. The coordinates of the digit $x(d_i^s)$, $y(d_i^s)$ are simply the coordinates of the center of the cell where the respective digit is recorded.

3.3 Clustering Algorithm

The clustering algorithm presented here relies on a specific method for keeping track of relative distances between detector cells potentially carrying digit information, at the level of each individual segment. This method is referred to as "ringer", and essentially associates to every possible pair of cells the integer value which is closest to the real-valued distance, in units of the lattice spacing in the respective segment. Let "ringer distance" be the name for this distance approximation and ρ^s its formal, generic notation used for segment s. Formally, such a scheme is useful for providing for a given cell with coordinates (x_1, y_1) the coordinates (x_i, y_i) of all other cells (the ring) separated by the same (integer) ringer distance $\rho^s_{1,i}$ within segment s, namely those for which the quantity $\sqrt{(x_i - x_1)^2 + (y_i - y_1)^2}$ is closest to $\rho^s_{1,i}$, where all quantities in this paragraph are in units of the lattice spacing in segment s. This is simply a geometrical-computational tool speeding up the calculations. The implied correspondence between the relative lattice position and ringer distance is shown in Fig. 11

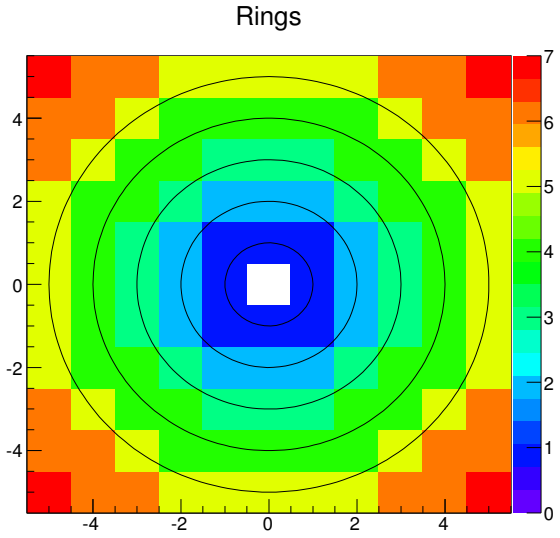

Figure 11: Assignment of ringer distances to cells in the vicinity of a specific cell – copied with permission from Ref. [26].

The main task of the clustering algorithm is to group the set of digits D^s generated in segment s during a certain event into several clusters, depending on their 2D spatial distribution. Digits are allowed to contribute to multiple clusters. One can understand a cluster as a 2D section in an electromagnetic shower. The idea is to maximize the percentage of incoming photons for which a cluster is reconstructed while minimizing the number of artifact clusters.

The clustering process requires a set of seeds based on which clusters would be grown. For the present algorithm, a digit d_j^s may be promoted to the status of a seed if certain criteria are met. Besides the digit amplitude ϵ_j^s, another property of interest is represented by the notion of seed energy of digit d_j^s, denoted by $\tilde{\epsilon}_j^s$, which is simply the sum of the amplitudes of all digits within the small lattice neighborhood:

$$\tilde{\epsilon}_j^s = \sum_{i}^{\rho_{i,j}^s \leq \rho_{\min}^s} \epsilon_i^s \tag{13}$$

where ρ_{min}^s is a first array of algorithm parameters, containing the minimal ring radius within each segment s. In practice, only two parameters are used: $\rho_{\min}^{\text{HGS}}$ and $\rho_{\min}^{\text{LGS}}$ for all HG and LG segments respectively.

During the clustering algorithm, digits are upgraded and downgraded to seeds. Let Δ^s be a dynamical set of seeds formed in segment s, with cardinality $N_\Delta^s := |\Delta^s|$, with individual seeds denoted by δ_i^s:

$$\Delta^s = \{\delta_i^s / i \in \overline{0, N_\Delta^s - 1}\}, \tag{14}$$

Let $\text{dig} : \Delta^s \to D^s$ be an injective but non-surjective function mapping seeds to digits for segment s. Similarly, let $\text{seed} : \{d_i^s \in D^s / (\exists \delta_{i'}^s \in \Delta^s).\text{dig}(\delta_{i'}^s) = d_i^s\} \to \Delta^s$ be the inverse function mapping digits to seeds for segment s. Having a seed $\delta_{i'}^s$ mapped to a digit d_i^s in this manner implies that the seed $\delta_{i'}^s$ is created from digit d_i^s (the digit is promoted as a seed), and allows, formally speaking, for the properties of digit – amplitude ϵ_i^s, seed energy $\tilde{\epsilon}_i^s$ and position $x(d_i^s)$, $y(d_i^s)$ – to also be available at the level of the seed – as $\epsilon_{i'}^s$, $\tilde{\epsilon}_{i'}^s$, $x(d_{i'}^s)$ and $y(d_{i'}^s)$ respectively.

Let W_{ij}^s denote a weight encoding the relative "influence" of seed δ_j^s on digit d_i^s. Effectively, W_{ij}^s is a cluster-formation weight [8], encoding the fraction of amplitude contribution of digit d_i^s to cluster c_j^s which may be formed from seed δ_j^s, normalized at the level of digit d_i^s: $\sum_j W_{ij}^s = 1, \forall j, s$. The formula for computing these cluster-formation weights is given by:

$$W_{ij}^s = \frac{f_W^s(d_i^s, \delta_j^s)}{\sum_{j'} f_W^s(d_i^s, \delta_{j'}^s)} \tag{15}$$

with function f_W^s differentiating between the digit from which the seed has been created in the first place, the digits that may contribute to the cluster and digits that may not contribute to the cluster, due their relative distances (either too large or too small):

$$f_W^s(d_i^s, \delta_j^s) = \begin{cases} g_W^s(\tilde{\epsilon}_j^s, 0), & \text{if } \text{dig}(\delta_j^s) = d_i^s \\ g_W^s(\tilde{\epsilon}_j^s, \rho_{i,j}^s), & \text{if } (\neg\text{dig}(\delta_j^s) = d_i^s) \wedge (\rho_{\min}^s < \rho_{i,j}^s < \rho_{\max}^s) \\ 0, & \text{else} \end{cases} \tag{16}$$

where $\rho_{\max}^s$ is another array of segment-level algorithm parameters, the maximal ring radius, which effectively sets a sharp, long distance threshold. Effectively, only digits separated by a ringer distance below $\rho_{\max}^s$ from seed δ_j^s contribute towards the cluster possibly formed from this seed. The contributions of digits with a ringer distance below $\rho_{\min}^s$ are already accounted for via the seed energy $\tilde{\epsilon}_j^s$. In Eq. (16), $g_W^s : \mathbb{Q}_+^* \times \mathbb{N} \to \mathbb{Q}_+$ is a function describing the expected shower profile, increasing with the first argument (seed energy) and decreasing with the second argument (ringer distance from the seed). The function employed during this study reads:

$$g_W^s(\epsilon, \rho) = \frac{\epsilon}{(l_s \rho / r_1^s) + e^{\left(l_s \rho / r_2^s\right)}} \tag{17}$$

[8] In the sense that they are needed for constructing the clusters in the first place, while being normalized at the level of digits; they are used below for constructing the cluster-description weights, which are normalized at the level of the already formed clusters, needed for fine tunning the definitions of shower properties.

where r_1^s and r_2^s are two parameters associated to the expected cluster width, in cm, for segment s, while $\rho_{i,j}^s$ is again the discrete, ringer distance between digit d_i^s and seed δ_j^s. The same radially decaying profile encoded by (17) is also used for rejecting some of the seeds that are identified in the vicinity of other, more energetic seeds – see first line of step 3(b)iB of the clustering algorithm. One can think of using different functions for optimizing the algorithm, which has been attempted in Ref. [26]. This showed that, although the function in Eq. (17) does not provide a good description of the cluster profile, changing the function does not significantly improve the efficiency.

Let C^s be a dynamical set of segment-level clusters formed in segment s, with cardinality $N_C^s := |C^s|$, with individual elements labeled as c_i^s:

$$C^s = \{c_i^s / i \in \overline{0, N_C^s - 1}\}, \tag{18}$$

The notation cluster(δ_j^s) simply encodes the creation of a new cluster at a level of segment s, having δ_j^s as a seed. The seed acts like a "center" with respect to which the relative amplitude contributions of neighboring digits are computed according to Eq. (15), which are further required for computing the cluster energy and further cluster properties. Effectively all seeds surviving until the end of the clustering algorithm are promoted to clusters. Computing further cluster properties may employ further weighting the amplitude contributions of digits $\epsilon_i^s W_{ij}^s$. Let w_{ij}^s denote a weight encoding the relative "importance" of digit d_i^s for cluster c_j^s, normalized at cluster level: $\sum\limits_i w_{ij}^s = 1, \forall j, s$. There are multiple ways of defining such cluster-description weights. Three of them are described and used during this study:

The first, standard procedure is linear weighting:

$$w_{ij}^s := \frac{\epsilon_i^s W_{ij}^s}{\sum\limits_{i'} \epsilon_{i'}^s W_{i'j}^s} \tag{19}$$

which identifies the cluster-description weight with the ratio between the fraction of digit-to-cluster amplitude contribution and the sum of all such contributions from digits towards the same cluster.

The second procedure involves logarithmic weighting according to:

$$w_{ij}^s := \max\left(0, \left(w_0 + \ln \frac{\epsilon_i^s W_{ij}^s}{\sum\limits_{i'} \epsilon_{i'}^s W_{i'j}^s}\right)\right) \tag{20}$$

which provides an increasing, concave function with respect to the digit/cluster energy ratio: $\epsilon_i^s W_{ij}^s \Big/ \sum\limits_{i'} \epsilon_{i'}^s W_{i'j}^s$. This means that the importance of low-amplitude digits is emphasized in comparison to linear weighting. The free parameter w_0 acts as a threshold on low energy ratios, threshold which eliminates the contributions of more digits with decreasing w_0. Logarithmic weighting is employed for enhancing the cluster position resolution and cluster width resolution of the LGS. This procedure was introduced in Ref. [27], where an optimal $w_0 := 4.5$ is found from simulations with a calorimeter with granularity approximately matching the Moliere radius, which is also the case for FoCal's LG component. An analogous optimization is presented in Sec. 4.1.1, leading to an optimal value of $w_0 = 4.2$.

The third procedure involves power-law weighting according to:

$$w_{ij}^s := \left(\frac{\epsilon_i^s W_{ij}^s}{\sum\limits_{i} \epsilon_{i'}^s W_{i'j}^s}\right)^p \tag{21}$$

which provides an increasing, convex function with respect to the digit/cluster energy ratio: $\epsilon_i^s W_{ij}^s \Big/ \sum_{i'} \epsilon_{i'}^s W_{i'j}^s$. This means that the importance of high-amplitude digits is emphasized in comparison to linear weighting. The free parameter p controls the strength of this tendency, with $p = 1$ matching the linear weighting in (19). Power-law weighting is employed for enhancing the cluster position resolution and cluster width resolution of the HGS. To the author's knowledge, this procedure has not been studied before. The free parameter p is also optimized in Sec. 4.1.2, leading to a value of $p := 2.2$.

Note that in the three equations above the quantity $\epsilon(c_j^s) = \sum_i \epsilon_i^s W_{ij}^s$ can be already interpreted as the raw (uncalibrated) energy content of cluster c_j^s to be formed around seed δ_j^s – this quantity is also listed as a cluster property in via equation (37).

Above, the notions of digits, seeds, and clusters are defined as elements of unordered sets. For the algorithm itself, the order of elements within a specific set often matters. For this purpose, the notations $\vec{D}^s$, $\vec{\Delta}^s$ and $\vec{C}^s$ are used for referring to the ordered versions of the sets of digits, seeds and clusters at the level of segment s. The (re)ordering criteria are specified in the pseudo-code below whenever necessary, within "()" following a "." following a set assignment operation indicated by "$\leftarrow$", the latter also being enclosed by "()", as these sets change while the different steps of the algorithm are executed. Whenever elements are added to or removed from the sorted versions of the sets $\vec{D}^s$, $\vec{\Delta}^s$ and $\vec{C}^s$, as indicated by the pseudo-code, it is implied that the same operation is applied to the unsorted versions D^s, Δ^s and C^s, and vice-versa. This also holds for modifying certain properties of specific elements. Individual elements are referred to using the same notations: d_i^s, δ_j^s and c_j^s respectively [9], with the indices indicating the position in the sorted sets, if such an ordering is already present before the respective step of the algorithm.

The algorithm starts with the sets of digits Δ^s obtained from the digitization procedure explained in Sec. 3.2, while the other sets are initially empty: $\Delta^s = \emptyset, C^s = \emptyset$. For segments $s \in \{3, 4, 5\}$, the clustering algorithm makes use of preseeds, as indicated below, meaning that information associated to clusters found in earlier segments is used for finding clusters in these, later segments. This makes sense since showers starting early might not have a clearly visible seed in the later segments. The pseudo-code description of the (segment-level) clustering algorithm is shown below, followed by further clarifications and explanations in text. Variable and set assignments are denoted by $\leftarrow$. The symbol "." is used with the meaning of "with the property", namely to signal the fact that a defining property of a variable used immediately before this sign is introduced immediately after this sign.

for every $s \in \vec{S}$:

1. $(\vec{D}^s \leftarrow D^s).(\epsilon_i^s \geq \epsilon_j^s \ \forall i < j)$
2. **for every** $d_i^s \in \vec{D}^s$:
   - compute $\tilde{\epsilon}_i^s$ according to Eq. (13)
3. (a) **if** $s \in$ PS:
   - **if** $s \in$ HGS: $s' \leftarrow s_{\text{seed}}^{\text{HGS}}$
     else: $s' \leftarrow s_{\text{seed}}^{\text{LGS}}$
   - $\Delta^s \leftarrow \Delta^s \cup \text{convSeeds}(s', s, C^{s'})$

   (b) **for every** $(d_i^s \in \vec{D}^s).(\epsilon_i^s > \epsilon_{\text{min}})$:
   i. **if** $\neg\text{rej}(d_i^s)$
   A. $\Delta^s \leftarrow \Delta^s \cup \text{seed}(d_i^s)$, $\text{dig}(\text{seed}(d_i^s)) \leftarrow d_i^s$
   B. **for every** $(d_j^s \in \vec{D}^s).(\rho_{j,i}^s < \rho_{\text{max}})$
   - **if** $\rho_{j,i}^s < \rho_{\text{min}}$: $\text{rej}(d_j^s) \leftarrow 1$

[9] Of course, the character used for the index may change.

- **if** $\frac{g^s_W(\tilde{\epsilon}^s_j,0)}{g^s_W(\tilde{\epsilon}^s_i,\rho^s_{j,i})} < R^s_R$: $\mathrm{rej}(d^s_j) \leftarrow 1$

(c) **for every** $d^s_i \in D^s$, **for every** $\delta^s_j \in \Delta^s$:

- compute W^s_{ij} according to Eq. (15)

4. $(\vec{\Delta}^s \leftarrow \Delta^s).(\tilde{\epsilon}^s_i \leq \tilde{\epsilon}^s_j \ \forall i < j)$ **for every** $\delta^s_j, \delta^s_i \in \vec{\Delta}^s$:

(a) **if** $\sum_i \epsilon^s_i W^s_{ij} < E^s_{\mathrm{min}}$: $\Delta^s \leftarrow \Delta^s - \{\delta^s_j\}$

(b) **if** $\sum_i W^s_{ij} < \mathrm{nDig}^s_{\mathrm{min}}$: $\Delta^s \leftarrow \Delta^s - \{\delta^s_j\}$

5. **for every** $\delta^s_j \in \Delta_s$:

(a) $c^s_j \leftarrow \mathrm{cluster}(\delta^s_j)$

(b) compute cluster properties of c^s_j described by Eq. (37), (27) - (36), using either Eq. (19), (20) or (21) for computing w_{ij}

(c) $C^s \leftarrow C^s \cup \{c^s_j\}$

In the description above, the notation $\mathrm{rej}(d^s_i)$ encodes an auxiliary boolean variable telling whether the digit d^s_i has been rejected (1) or not (0) as a seed during step 3 of the clustering algorithm. The notation $\mathrm{convSeeds}(s', s, C^{s'})$ states that the XY positions of the clusters found in segment s' are extrapolated along the z to XY positions the plane of segment s, where digits are looked for in the vicinity of each of these positions, within a neighborhood of ρ^s_{min} rings. If present at all, the d^s_i digit with the highest amplitude ϵ_i is associated to each extrapolated position and promoted to a seed δ^s_i. This is the preseeding procedure mentioned above.

$\mathrm{PS} \subset S$ denotes the subset of segments s for which this procedure is employed ("preseeded" segments). $s^{\mathrm{HGS}}_{\mathrm{seed}}$ and $s^{\mathrm{LGS}}_{\mathrm{seed}}$ are, respectively, the high granularity and low granularity segments s' from which the preseed cluster information is extracted (preseeding segments) – HG clusters are used for preseeding HG segments and LG clusters for preseeding LG segments. Given that clustering is carried out over each segment at a time, in the order of increasing z and increasing $s \in \vec{S}$, it is required that the preseeding segments are placed before the preseeded segments: $s^{\mathrm{HGS}}_{\mathrm{seed}} < s'.(s' \in \mathrm{PS} \cap \mathrm{HGS})$ and $s^{\mathrm{LGS}}_{\mathrm{seed}} < s'.(s' \in \mathrm{PS} \cap \mathrm{LGS})$. Here $\mathrm{PS} = \{3, 4, 5\}$, $s^{\mathrm{HGS}}_{\mathrm{seed}} = 1$ and $s^{\mathrm{LGS}}_{\mathrm{seed}} = 2$.

In step 1 the digits are sorted according to decreasing amplitudes. In step 2 the seed energies of all digits are calculated. Step 3a is concerned with the preseeding procedure, for segments requiring it. The ordered set of digits is looped over in step 3b, until the first digit with amplitude beneath the seed threshold is found. Within each step of the loop, the digit is promoted to seed unless already rejected, then the digits in its lattice neighborhood are looked at, for ringer distances smaller than ρ^s_{max}. Digits in this neighborhood are rejected as seeds either if they fall below ρ^s_{min} with respect to the just formed seed or if their amplitude is not high enough compared to the expected cluster profile around the just formed seed. The latter rejection threshold is controlled by another segment-dependent algorithm parameter R^s_R In step 3c the cluster formation weights from all digits to all surviving seeds are computed according to (15), most digit $\rightarrow$ seed combinations typically leading to the value of 0. Further rejection is carried out in step 4, based on (W-weighted) energy content 4a and (W-weighted) digit content 4b, with the corresponding upper rejection thresholds controlled by parameters E^s_{min} and $\mathrm{nDig}^s_{\mathrm{min}}$ respectively. Finally, the surviving seeds are promoted to (segment-level) cluster whose properties are then computed in step 5.

3.4 Shower Reconstruction Algorithm

The clustering algorithm essentially provides the set C^s of segment-level cluster for every segment s. Within the shower reconstruction algorithm, clusters generated from the HG segments (belonging to

$C^s, s \in \text{HGS}$) are combined (or merged) to produce the set C^H of (detector level) semi-final HG clusters (or showers), with cardinality $N_C^H := |C^H|$, with individual clusters labeled as c_i^H:

$$C^H = \{c_i^H / i \in \overline{0, N_C^H - 1}\}, \tag{22}$$

Similarly, clusters generated from the LG segments (belonging to $C^s, s \in \text{HGS}$) are combined (or merged) to produce the set C^L of (detector-level) semi-final LG clusters (or showers), with cardinality $N_C^L := |C^L|$, with individual clusters labeled as c_i^L:

$$C^L = \{c_i^L / i \in \overline{0, N_C^L - 1}\}, \tag{23}$$

Finally, semi-final HG and LG clusters are combined to generate the set C^F of (detector level) final clusters (or showers), with cardinality $N_C^F := |C^F|$, with individual clusters labeled as c_i^F:

$$C^F = \{c_i^F / i \in \overline{0, N_C^F - 1}\}, \tag{24}$$

The above 3 sets C^H, C^L and C^F are dynamical, as elements may be added and removed while executing the shower-reconstruction algorithm. Within the description of the algorithm it is implied the associated 3 cardinalities $N_C^H - 1$, $N_C^L - 1$ and $N_C^F - 1$ are consistently changed whenever the respective sets are changed.

The idea is that every semi-final LG cluster c_j^L makes use of information from at most one LG cluster c_i^s from every LG segment $s \in \text{LG}$, while every semi-final HG cluster c_j^H makes use of information from at most one HG cluster c_i^s from every HG segment $s \in \text{HG}$. In turn, every final c_k^F cluster makes use of information from exactly one LG semi-final cluster and at most one HG semi-final cluster. In terms of the algorithm, a final cluster must have well defined and reasonably accurate energy information as well as transverse position information. In this sense, a final cluster must have an associated LG semi-final cluster in order to exist, since no reliable energy information would be available otherwise. On the other hand, a final cluster may exist without an associated HG semi-final cluster, since position information is also provided by the LG semi-final cluster, although significantly less accurate.

In order to formally keep track of the merging operations within the description of the shower shape reconstruction, as well as for when computing shower properties described in Section 3.7, the following four notations are used:

- $\text{merL}(j, s)$ gives, for the semi-final, LG cluster c_j^L the index j^{sL} of the corresponding (LG) segment level cluster $c_{j^{sL}}^s$ of segment s that contributes to it
- $\text{merH}(j, s)$ gives, for the semi-final, HG cluster c_j^H the index j^{sH} of the corresponding (HG) segment level cluster $c_{j^{sH}}^s$ of segment s that contributes to it
- $\text{merFL}(j)$ gives, for the final cluster c_j^F the index j^{FL} of the corresponding semi-final, LG cluster $c_{j^{FL}}^L$ that contributes to it
- $\text{merFH}(j)$ gives, for the final cluster c_j^F the index j^{FH} of the corresponding semi-final, HG cluster $c_{j^{FH}}^H$ that contributes to it

where the first two relate semi-final clusters to segment level clusters and the last two relate final clusters to semi-final clusters. In a similar spirit, the notation $\text{hasHG}(c_j^F)$ is a boolean variable encoding whether the respective final cluster makes use of of a HG semi-final cluster (1) or not (0):

$$
\begin{aligned}
&(C^F \leftarrow C^F \cup \{c_j^F\}) \rightarrow (\text{hasHG}(c_j^F) \leftarrow 0) \\
&(\text{merFH}(j) \leftarrow k.(c_k^H \in C^H)) \rightarrow (\text{hasHG}(c_j^F) \leftarrow 1)
\end{aligned} \tag{25}
$$

being initialized to 0 when the final cluster c_j^F is created and assigned a value of 1 when a HG semi-final cluster c_k^H is merged as part of c_j^F. Finally, the notation $\text{mer}(c_i^s)$ is a boolean auxiliary variable (used only within the description of the algorithm itself) encoding whether the respective segment level cluster has been merged (1) or not (0) during step 1 or 2 of the shower reconstruction algorithm:

$$
\begin{aligned}
&(C^s \leftarrow C^s \cup \{c_i^s\}) \rightarrow (\text{mer}(c_i^s) \leftarrow 0), s \in \text{HGS} \\
&(\text{merH}(j,s) \leftarrow i) \rightarrow (\text{mer}(c_i^s) \leftarrow 1), s \in \text{HGS} \\
&(C^s \leftarrow C^s \cup \{c_i^s\}) \rightarrow (\text{mer}(c_i^s) \leftarrow 0), s \in \text{LGS} \\
&(\text{merL}(j,s) \leftarrow i) \rightarrow (\text{mer}(c_i^s) \leftarrow 1), s \in \text{LGS}
\end{aligned} \tag{26}
$$

which is initialized to 0 when (LG or HG) segment-level cluster c_i^s is created and assigned to a value of 1 when merged as part of a semi-final (LG or HG) cluster. In the shower reconstruction algorithm below, it is implied that the values of hasHG and mer are set automatically according to the logical statements above, whenever the values of merL, merH or merFH are set.

The notation $\text{dZFC}(c^a, c^b)$ denotes the distance between two (generic) clusters c^a and c^b, after being projected from their original plane to the z_{FCP}, under the assumption that the particle producing them originated at the interaction point.

This algorithm is strictly valid only if the minimal and maximal ringer distances are the same in all HG and LG segments respectively: $\rho_{\text{min}}^s = \rho_{\text{min}}^{\text{LGS}}, \forall s \in \text{LGS}$ and $\rho_{\text{min}}^s = \rho_{\text{min}}^{\text{HGS}}, \forall s \in \text{HGS}$.

Below, the formal description of the shower reconstruction algorithm follows, assuming the existence of the sets of segment level clusters C^s provided by the clustering algorithm described in Section 3.3. On the other hand, the sets of semi-final and final clusters are initially empty: $C^H = \emptyset, C^L = \emptyset, C^F = \emptyset$.

1. **for every** $s_a \in \text{LGS}$, starting with $s_{\text{seed}}^{\text{LGS}}$:
   - **for every** $s_b \in \text{LGS} - \{s_a\}$:
     - **for every** $c_i^{s_a} \in C^{s_a}$:
       (a) **if** $\neg\text{mer}(c_i^{s_a})$:
           * $k \leftarrow N_{C^L}$
           * $C^L \leftarrow C^L \cup \{c_k^L\}$
           * $\text{merL}(k, s_a) \leftarrow i$
           else:
           * $k \leftarrow k'.(\text{merL}(k', s^a) = i)$
       (b) **for every** $(c_j^{s_b} \in C^{s_b}).(\neg\text{mer}(c_j^{s_b}))$:
           * **if** $\text{dZFC}(c_i^{s_a}, c_j^{s_b}) < \rho_{\text{min}}^{\text{LGS}}$:
             · $\text{merL}(k, s_b) \leftarrow j$
2. **for every** $s_a \in \text{HGS}$, starting with $s_{\text{seed}}^{\text{HGS}}$:
   - **for every** $s_b \in \text{HGS} - \{s_a\}$:
     - **for every** $c_i^{s_a} \in C^{s_a}$:
       (a) **if** $\neg\text{mer}(c_i^{s_a})$:
           * $k \leftarrow N_{C^H}$

* $C^H \leftarrow C^H \cup \{c_k^H\}$
* $\mathrm{merH}(k, s_a) \leftarrow i$

else:

* $k \leftarrow k'.(\mathrm{merH}(k', s^a) = i)$

(b) **for every** $(c_j^{s_b} \in C^{s_b}).(\neg\mathrm{mer}(c_j^{s_b}))$:

* **if** $\mathrm{dZFC}(c_i^{s_a}, c_j^{s_b}) < \rho_{\min}^{\mathrm{HGS}}$:
  · $\mathrm{merH}(k, s_b) \leftarrow j$

3. • **for every** $c_j^L \in C^L$ compute shower properties described by Eq. (40), (42), (44),(46), (48) and (50) (the latter using Eq. (38) for energy calibration)
   • **for every** $c_j^H \in C^H$ compute shower properties described by Eq. (39), (41), (43),(45), (47) and (50)

4. **for every** $c_j^H \in C^H$:

   (a) $\mathrm{NHC} \leftarrow \{(c_i^L \in C^L).\ \mathrm{dZFC}(c_j^H, c_i^L) < \rho_{\max}^{\mathrm{LGS}}\}$

   (b) **if** $\mathrm{NHC} \leftarrow \emptyset$
   • $C^H \leftarrow C^H - \{c_j^H\}$

   else
   • $k \leftarrow i.(c_i^L \in \mathrm{NHC} \wedge \mathrm{dZFC}(c_j^H, c_i^L) \leq \mathrm{dZFC}(c_j^H, c_{i'}^L) \forall c_{i'}^L \in \mathrm{NHC})$
   • $q \leftarrow N_{C^F}$
   • $C^F \leftarrow C^F \cup \{c_q^F\}$
   • $\mathrm{merFH}(q) \leftarrow j$
   • $\mathrm{merFL}(q) \leftarrow k$

5. **for every** $(c_j^L \in C^L).\nexists(c_k^F \in C^F).\mathrm{merFL}(k) = j$:

   (a) $q \leftarrow N_{C^F}$
   (b) $C^F \leftarrow C^F \cup \{c_q^F\}$
   (c) $\mathrm{merFL}(q) \leftarrow j$

6. **for every** $c_j^F \in C^F$ compute shower properties described by (52), (53), (54), (55), (58), (59), (60), (61) and (62)

Step 1 of the algorithm creates the list of semi-final LG clusters C^L. It does that via two nested loops over the LG segments. For every LG segment s_b picked by the outer loop and for every LG segment s_a picked by the inner loop, clusters $c_i^{s_a}$ at the level of the latter segment are iterated. Those clusters $c_i^{s_a}$ that have not yet been merged within a semi-final LG clusters are used to create a semi-final LG cluster. For every cluster $c_i^{s_a}$ in segment s_a, the cluster $c_i^{s_b}$ (if any) in segment s_b that is within the minimal LG ringer (transverse) distance from $c_i^{s_a}$, in terms of the projected clusters' positions in the FOCAL center plane, and that has not been merged yet is merged within the same LG semi-final cluster as $c_i^{s_a}$. Step 2 creates the list of semi-final HG clusters C^H according to the same procedure, except that the minimal HG ringer distance is used instead of the minimal LG ringer distance.

Step 3 is only concerned with computing the variables associated to semi-final LG and HG clusters, according to equations in Section 3.7. In step 4 each semi-final HG cluster c_j^H is assigned to nearest semi-final LG cluster c_i^L positioned within the maximal LG ringer distance – comparisons again done in terms of distances between projected cluster positions in the FOCAL center plane. A final cluster c_q^F is created from semi-final HG cluster c_j^H and from semi-final LG cluster c_i^L and added to the set of final clusters C^F. If there is no semi-final LG cluster c_i^L satisfying the required condition, the respective semi-final HG cluster c_j^H is discarded. The remaining semi-final LG clusters not combined with semi-final HG clusters are promoted alone to final clusters in step 5. Finally, step 6 is concerned with computing the variables describing all obtained final clusters, according to the indicated equations from Section 3.7.

s	seg. type	ρ^s_{min}	ρ^s_{max}	ϵ^s_{min}	E^s_{min}	$\text{nDig}^s_{\text{min}}$	r^s_1	r^s_2	R^s_R
0	LGS	1	5	0	5000	2.0	0.9	0.5	10.0
1	HGS	2	40	0	7	3.0	0.14	0.5	3.5
2	LGS	1	5	0	10000	3.0	0.8	0.5	10.0
3	HGS	2	40	0	10	10	0.18	0.5	3.5
4	LGS	1	5	0	15000	6	1.7	0.5	10.0
5	LGS	1	5	0	8500	5.0	2.85	0.5	10.0

Table 2: Parameters used by the Clustering and Shower-Shape reconstruction algorithms.

Table 2 summarizes the segment-dependent parameters used by the clustering and/or shower-reconstruction algorithm, with the associated values that have been used for obtaining the results presented in Sections 4 and 5. The first two columns list the segments and their type $s \in S$. This combination of values has been reached while trying to increase as much as possible the efficiency of photon identification firstly, and secondly the efficiency the identification of pairs of photons originating from π_0. These optimizations have been conducted by members of the ALICE-FoCal collaboration prior to this study, and the commonly agreed upon values are simply used here without further attempts to optimize them. For the reader's convenience, the meaning and relevance of these parameters is shortly summarized below:

- ρ^s_{min} is the minimal ringer radius in segment s, measured in integers of lattice ringer distance (see beginning of Section 3.3), used for:
  - computing seed energies $\tilde{\epsilon}^s_j$ according to Eq. (13), in step 2 of the clustering algorithm
  - rejecting seeds located very close to higher energy seeds within the same segment, in step 3(b)iB of the clustering algorithm
  - merging clusters formed within segments of the same type in steps 1b and 2b of the shower-reconstruction algorithm
- ρ^s_{max} is the maximal ringer radius in segment s, measured in integers of the lattice ringer distance (see beginning of Section 3.3), used for:
  - setting a high transverse distance threshold on digits that may be considered part of a clusters formed around a specific seed, via Eq. (16) in step 3c of the clustering algorithm
  - merging HG semi-final clusters with LG semi-final clusters, in step 4 of the shower-reconstruction algorithm
- ϵ^s_{min} is a lowest amplitude digit-to-seed promotion threshold on digits in segment s, in step 3b of the clustering algorithm; measured in the same units as the amplitudes of digits in the respective segment (keV for $s \in$ LGS, integer amplitudes for $s \in$ HGS); currently has a trivial value in all segments
- E^s_{min} a lowest amplitude seed-to-cluster promotion threshold on seeds in segment s, based on the weighted amplitude contributions from all digits, in step 4a of the clustering algorithm
- $\text{nDig}^s_{\text{min}}$ a lower number-of-cells seed-to-cluster promotion threshold on seeds in segment s, based on the weighted number of digits contributing to the possible cluster, in step 4b of the clustering algorithm
- r^s_1 is the first parameter controlling the expected radial cluster profile encoded by Eq. (17), measured in units of cm
- r^s_2 is the second parameter controlling the expected radial cluster profile encoded by Eq. (17), measured in units of cm
- R^s_R is the rejection ratio parameter (dimensionless), used for rejecting seeds located between ρ^s_{min} and ρ^s_{max} of higher-energy seeds in the same segment s, based on the expected radial cluster profile, in step 3(b)iB of the clustering algorithm

3.5 Cluster Properties

This section lists the formulas associated to properties of segment-level clusters that are computed while executing the clustering algorithm. These properties are defined for every cluster c_j^s in segment s. These properties are further used (Sec. 3.7) for constructing higher-level variables that can possibly be used for discriminating between photon showers on one hand and double photon showers on the other hand (as well as showers triggered by hadrons). First, one computes the transverse (X and Y) cluster position, based on a weighted average of positions of digits within the same segments:

$$\langle x \rangle (c_j^s) = \frac{\sum\limits_{i}^{\rho_{i,j}^s \leq 3} x(d_i^s) w_{ij}^s}{\sum\limits_{i}^{\rho_{i,j}^s \leq 3} w_{ij}^s}, \;\; \langle y \rangle (c_j^s) = \frac{\sum\limits_{i}^{\rho_{i,j}^s \leq 3} y(d_i^s) w_{ij}^s}{\sum\limits_{i}^{\rho_{i,j}^s \leq 3} w_{ij}^s} \tag{27}$$

where the cluster description weights w_{ij}, giving the relative importance of digit d_i^s towards cluster c_j^s, are computed according to Eq. (19), (20) or (21), depending on whether linear, logarithmic or power-law weighting is used respectively. Note that only for (27) the additional constraint is imposed that the ringer distance $\rho_{i,j}^s$ between the digit d_i^s contributing to the average and the position of the seed δ_j^s out of which the cluster is formed is smaller or equal to 3 (for both HG and LG segments). Otherwise, large distance fluctuations may lead to cluster positions that are inconsistent positions of the respective seeds. However, when used in equations (30), (31) and (36) below, this ringer distance constraint for computing $\langle x \rangle$ and $\langle y \rangle$ is asbandoned. Similarly, the weighted averages of the squares of the X and Y positions are also calculated:

$$\left\langle x^2 \right\rangle (c_j^s) = \frac{\sum\limits_{i} x^2(d_i^s) w_{ij}^s}{\sum\limits_{i} w_{ij}^s}, \;\; \left\langle y^2 \right\rangle (c_j^s) = \frac{\sum\limits_{i} y^2(d_i^s) w_{ij}^s}{\sum\limits_{i} w_{ij}^s} \tag{28}$$

together with the weighted off-diagonal covariance:

$$\langle xy \rangle (c_j^s) = \frac{\sum\limits_{i} x(d_i^s) y(d_i^s) w_{ij}^s}{\sum\limits_{i} w_{ij}^s} \tag{29}$$

without imposing the upper ringer distance threshold any longer, since these quantities are only used further for calculating cluster and shower shapes. The results of equations (27) – without the ringer distance constraint – and (28) are used for computing the squared standard deviations:

$$\sigma_{x^2}(c_j^s) = \left\langle x^2 \right\rangle (c_j^s) - \langle x \rangle^2 (c_j^s), \;\; \sigma_{y^2}(c_j^s) = \left\langle y^2 \right\rangle (c_j^s) - \langle y \rangle^2 (c_j^s) \tag{30}$$

while the results of (27) – without the ringer distance constraint – and (29) are used for computing the mixed quantity:

$$\sigma_{xy}(c_j^s) = \langle xy \rangle (c_j^s) - \langle x \rangle \langle y \rangle (c_j^s), \tag{31}$$

which are in turn used for computing

$$d(c_j^s) = \sigma_{y^2}(c_j^s) - \sigma_{x^2}(c_j^s) \tag{32}$$

and:

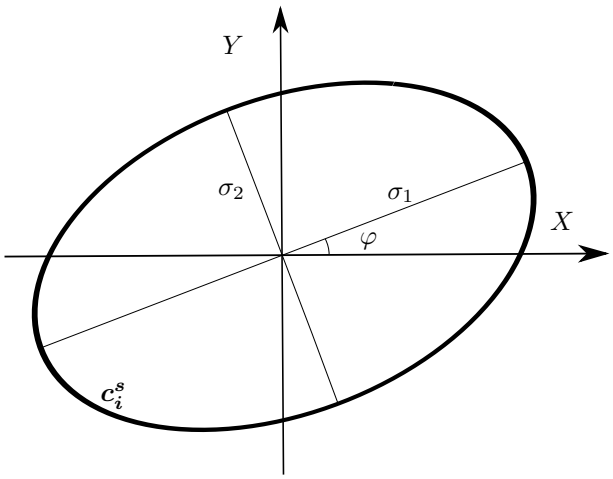

Figure 12: Assumed elliptical profile of clusters generated at the level of any segment.

$$s(c_j^s) = \sqrt{d^2(c_j^s) + 4\sigma_{xy}^2(c_j^s)} \tag{33}$$

Equations (30), (31), (32) and (33) are employed together for computing the semi-major cluster width:

$$\sigma_1(c_j^s) = \sqrt{\frac{\sigma_{x^2}(c_j^s) + \sigma_{y^2}(c_j^s) + s(c_j^s)}{2}} \tag{34}$$

the semi-minor cluster width:

$$\sigma_2(c_j^s) = \sqrt{\frac{\sigma_{x^2}(c_j^s) + \sigma_{y^2}(c_j^s) - s(c_j^s)}{2}} \tag{35}$$

and the orientation of the semi-major width of the cluster with respect to the X-axis of the lattice:

$$\varphi(c_j^s) = \arctan \frac{[d+s]\langle y \rangle + 2\sigma_{xy}\langle x \rangle}{2\sigma_{xy}\langle y \rangle - [d-s]\langle x \rangle}(c_j^s) \tag{36}$$

where the angle $\varphi(c_j^s) \in [-\pi/2, \pi/2]$ by definition spans only two quadrants of the trigonometric circle. The quantities defined in equations (34), (35) and (36) are exact when the cluster is an elliptically Gaussian distribution in 2 dimensions, with the calculation procedure taken from [28] and briefly illustrated by Fig. 12

In addition to the cluster-shape variables above, the raw, uncalibrated energy content of a segment level cluster is given by:

$$\epsilon(c_j^s) = \sum_i \epsilon_i^s W_{ij}^s \tag{37}$$

where W_{ij}^s are the weights defined in (15) (normalized to 1 per digit d_i^s).

3.6 Calibration

The summed up raw energy of every semi-final, low granularity cluster created during the shower-reconstruction algorithm is converted to realistic particle energy information via the following calibration function:

$$\Gamma(\epsilon) = t_0(\epsilon + t_1)^{t_2} + t_3 \tag{38}$$

making use of 4 parameters t_0, t_1, t_2 and t_3, which are found from a preliminary fitting to true particle energy versus raw detector response information, obtained from a large set of single photon simulations. Such a calibration is only used for the LG, analogue signal. Calibrating the response of HG segments turned out to be much more difficult. For this reason, the HG layers are only used for extracting position information.

The values used during this study are: $t_0 = 9.375 \cdot 10^{-5}, t_1 = 8.532 \cdot 10^3$, $t_2 = 1.0$ and $t_2 = 3.89 \cdot 10^{-1}$, given that in equation (38) Γ is in units of GeV while ϵ is in units of keV.

3.7 Shower Properties

When segment-level clusters are merged to form (detector-level) semi-final clusters during the shower-reconstruction algorithm, information contained by the relevant properties (or variables) describing the segment level clusters are combined in order to produce the properties (or variables) describing the semi-final clusters.

First, the reconstructed transverse position of a semi-final cluster is given by:

$$x(c_j^H) = \frac{\sum\limits_{s\in\text{HGS}}^{} x(c_{j^{sH}}^s)\epsilon(c_{j^{sH}}^s)}{\sum\limits_{s\in\text{HGS}}^{} \epsilon(c_{j^{sH}}^s)}, \quad y(c_j^H) = \frac{\sum\limits_{s\in\text{HGS}}^{} y(c_{j^{sH}}^s)\epsilon(c_{j^{sH}}^s)}{\sum\limits_{s\in\text{HGS}}^{} \epsilon(c_{j^{sH}}^s)} \tag{39}$$

for a HG semi-final clusters and by:

$$x(c_j^L) = \frac{\sum\limits_{s\in\text{LGS}}^{} x(c_{j^{sL}}^s)\epsilon(c_{j^{sL}}^s)}{\sum\limits_{s}^{} \epsilon(c_{j^{sL}}^s)}, \quad y(c_j^L) = \frac{\sum\limits_{s\in\text{LGS}}^{} y(c_{j^{sL}}^s)\epsilon(c_{j^{sL}}^s)}{\sum\limits_{s}^{} \epsilon(c_{j^{sL}}^s)} \tag{40}$$

for a LG semi-final cluster, where the notations $j^{sH} := \text{merH}(j, s)$ and $j^{sL} := \text{merL}(j, s)$ are used in (39) and (40) respectively, with merH, merL notations defined themselves at the beginning of Section 3.4. These notations are used with the same meaning in all equations up to (50). Equations (39) and (40) simply state that the transverse positions of semi-final HG and LG clusters respectively are (uncalibrated) energy-weighted averages over the positions of the segment-level HG and LG clusters respectively composing them, which are in turn defined by (27) above.

Similarly, the semi-major and semi-minor widths of a semi-final HG cluster are also (uncalibrated) energy-weighted averages over the semi-major and semi-minor widths of the merged segment-level HG clusters:

$$\sigma_1^H(c_j^H) = \frac{\sum\limits_{s\in\text{HGS}}^{} \sigma_1(c_{j^{sH}}^s)\epsilon(c_{j^{sH}}^s)}{\sum\limits_{s\in\text{HGS}}^{} \epsilon(c_{j^{sH}}^s)}, \quad \sigma_2^H(c_j^H) = \frac{\sum\limits_{s\in\text{HGS}}^{} \sigma_2(c_{j^{sH}}^s)\epsilon(c_{j^{sH}}^s)}{\sum\limits_{s\in\text{HGS}}^{} \epsilon(c_{j^{sH}}^s)} \tag{41}$$

with $j^{sH} := \text{merH}(j,s)$ used with the same meaning as in (39), while the semi-major and semi-minor widths of a semi-final LG cluster are also (uncalibrated) energy-weighted averages over the semi-major and semi-minor widths of the merged segment-level LG clusters:

$$\sigma_1^L(c_j^L) = \frac{\sum\limits_{s\in\text{LGS}}^{} \sigma_1(c_{j^{sL}}^s)\epsilon(c_{j^{sL}}^s)}{\sum\limits_{s\in\text{LGS}}^{} \epsilon(c_{j^{sL}}^s)}, \quad \sigma_2^L(c_j^L) = \frac{\sum\limits_{s\in\text{LGS}}^{} \sigma_2(c_{j^{sL}}^s)\epsilon(c_{j^{sL}}^s)}{\sum\limits_{s\in\text{LGS}}^{} \epsilon(c_{j^{sL}}^s)} \tag{42}$$

with $j^{sL} := \text{merL}(j,s)$ used with the same meaning as in (40).

In an analogous manner, the average orientation of semi-final clusters is defined:

$$\varphi^H(c_j^H) = \frac{1}{2}\text{atan2}\left(\frac{\sum\limits_{s\in\text{HGS}}^{} \sin(2\varphi(c_{j^{sH}}^s))\epsilon(c_{j^{sH}}^s)}{\sum\limits_{s\in\text{HGS}}^{} \epsilon(c_{j^{sH}}^s)}, \frac{\sum\limits_{s\in\text{HGS}}^{} \cos(2\varphi(c_{j^{sH}}^s))\epsilon(c_{j^{sH}}^s)}{\sum\limits_{s\in\text{HGS}}^{} \epsilon(c_{j^{sH}}^s)} \right) \tag{43}$$

for HG semi-final clusters and:

$$\varphi^L(c_j^L) = \frac{1}{2}\text{atan2}\left(\frac{\sum\limits_{s\in\text{LGS}}^{} \sin(2\varphi(c_{j^{sL}}^s))\epsilon(c_{j^{sL}}^s)}{\sum\limits_{s\in\text{LGS}}^{} \epsilon(c_{j^{sL}}^s)}, \frac{\sum\limits_{s\in\text{LGS}}^{} \cos(2\varphi(c_{j^{sL}}^s))\epsilon(c_{j^{sL}}^s)}{\sum\limits_{s\in\text{LGS}}^{} \epsilon(c_{j^{sL}}^s)} \right) \tag{44}$$

for LG semi-final clusters. The trigonometric functions sin, cos and atan2 are conveniently combined in order to account for angular periodicity, while the factors of 2 and 1/2 account for the fact that the orientations are defined as angles belonging to $[-\pi/2, \pi/2]$ and not to the full trigonometric circle. More explicitly, the calculation procedure encoded by equations (43) and (44) is as follows:

- the orientations of the contributing segment-level clusters $\varphi(c_{j^{H/L}}^s)$ are multiplied by 2, such that they cover the entire $[-\pi, \pi]$ interval instead of the $[-\pi/2, \pi/2]$ interval
- the sines and cosines of the resulting angles are separately averaged over, generating Y and X values respectively, both distributed within $[-1, 1]$
- the arctan2(Y,X) function is evaluated on the Y and X values above, which provides an angle between $[-\pi, \pi]$
- finally, the resulting angle is divided by 2 and thus mapped back to the interval $[-\pi/2, \pi/2]$, which is the angular average

The segment-level cluster orientation variables defined in (36) are also used for constructing transverse cluster alignment variables:

$$r_\varphi^H(c_j^H) = \sqrt{\left(\frac{\sum\limits_{s\in\text{HGS}}^{} \cos(2\varphi(c_{j^{sH}}^s))\epsilon(c_{j^{sH}}^s)}{\sum\limits_{s\in\text{HGS}}^{} \epsilon(c_{j^{sH}}^s)} \right)^2 + \left(\frac{\sum\limits_{s\in\text{HGS}}^{} \sin(2\varphi(c_{j^{sH}}^s))\epsilon(c_{j^{sH}}^s)}{\sum\limits_{s\in\text{HGS}}^{} \epsilon(c_{j^{sH}}^s)} \right)^2} \tag{45}$$

for HG semi-final and:

$$r_\varphi^L(c_j^L) = \sqrt{\left(\frac{\sum\limits_{s\in\text{LGS}}^{} \cos(2\varphi(c_{j^{sL}}^s))\epsilon(c_{j^{sL}}^s)}{\sum\limits_{s\in\text{LGS}}^{} \epsilon(c_{j^{sL}}^s)} \right)^2 + \left(\frac{\sum\limits_{s\in\text{LGS}}^{} \sin(2\varphi(c_{j^{sL}}^s))\epsilon(c_{j^{sL}}^s)}{\sum\limits_{s\in\text{LGS}}^{} \epsilon(c_{j^{sL}}^s)} \right)^2} \tag{46}$$

for LG semi-final clusters. These variables measure the extent to which the ellipses matched to the contributing HG and LG segment-level clusters are aligned with respect to each other in the transverse XY plane. Both the LG and HG alignment variables belong to the interval $[0, 1]$, where 0 signals a very poor alignment while 1 signals a very good alignment. This translates to uniform vs peaked distribution of segment-level cluster orientations within the $[-\pi/2, \pi/2]$ space, after angular periodicity is being accounted for – the alignment variables can also be understood as inverse angular standard deviations. The computational procedure encoded by (45) and (46) is very similar to the one encoded by (43) and by (44) described above, the first two steps being identical. For the former equations, the last two steps consist of:

- evaluating the squares of the Y and X values obtained from the averaged sin and cos values of the doubled segment-level orientations
- taking the square root of the sum of the squares obtained above, which gives a variable between $[0, 1]$

For further details about angular statistics see for instance [30].

The average orientations at the level of the semi-final clusters (43) and (44), together with the orientations of the contributing segment-level clusters (36) and their associated semi-major and semi-minor widths (42) and (41) are used for constructing vectorial-weighted semi-major and semi-minor widths:

$$\hat{\sigma}_1^H(c_j^H) = \frac{\sum\limits^{s\in\text{HGS}} \sigma_1^P(c^s_{j^{sH}}, c_j^H)\epsilon(c^s_{j^{sH}})}{\sum\limits^{s\in\text{HGS}} \epsilon(c^s_{j^{sH}})}, \quad \hat{\sigma}_2^H(c_j^H) = \frac{\sum\limits^{s\in\text{HGS}} \sigma_2^P(c^s_{j^{sH}}, c_j^H)\epsilon(c^s_{j^{sH}})}{\sum\limits^{s\in\text{HGS}} \epsilon(c^s_{j^{sH}})}, \tag{47}$$

for HG semi-final clusters and:

$$\hat{\sigma}_1^L(c_j^L) = \frac{\sum\limits^{s\in\text{LGS}} \sigma_1^P(c^s_{j^{sL}}, c_j^L)\epsilon(c^s_{j^{sL}})}{\sum\limits^{s\in\text{LGS}} \epsilon(c^s_{j^{sL}})}, \quad \hat{\sigma}_2^L(c_j^L) = \frac{\sum\limits^{s\in\text{LGS}} \sigma_2^P(c^s_{j^{sL}}, c_j^L)\epsilon(c^s_{j^{sL}})}{\sum\limits^{s\in\text{LGS}} \epsilon(c^s_{j^{sL}})}, \tag{48}$$

for LG semi-final clusters, which are essentially energy-weighed averages of σ_1^P and σ_2^P, which in turn are functions acting on every contributing segment-level cluster and on the semi-final cluster containing it, defined by:

$$\begin{aligned}
\sigma_1^P(c^s, c) &= \left(\sqrt{\frac{\cos^2(\varphi(c^s) - \varphi(c))}{\sigma_1^2(c^s)} + \frac{\sin^2(\varphi(c^s) - \varphi(c))}{\sigma_2^2(c^s)}} \right)^{-1} \\
\sigma_2^P(c^s, c) &= \left(\sqrt{\frac{\sin^2(\varphi(c^s) - \varphi(c))}{\sigma_1^2(c^s)} + \frac{\cos^2(\varphi(c^s) - \varphi(c))}{\sigma_2^2(c^s)}} \right)^{-1}
\end{aligned} \tag{49}$$

where c is the semi-final cluster and c^s is a contributing segment-level cluster. Regardless of whether c^s and c are HG or LG segment-level and semi-final clusters respectively (although both should be of the same type, either HG or LG), $\sigma_1^P(c^s, c)$ provides the semi-width of the ellipse matched to c^s along the direction given by the average orientation of c – orientation of its semi-major width, thus giving a contribution of segment-level cluster c^s towards the vectorial-weighted semi-major width of the semi-final cluster c containing it. Similarly, $\sigma_2^P(c^s, c)$ provides the semi-width of the ellipse matched to c^s along the direction perpendicular to the direction given by the average orientation of c – orientation of its

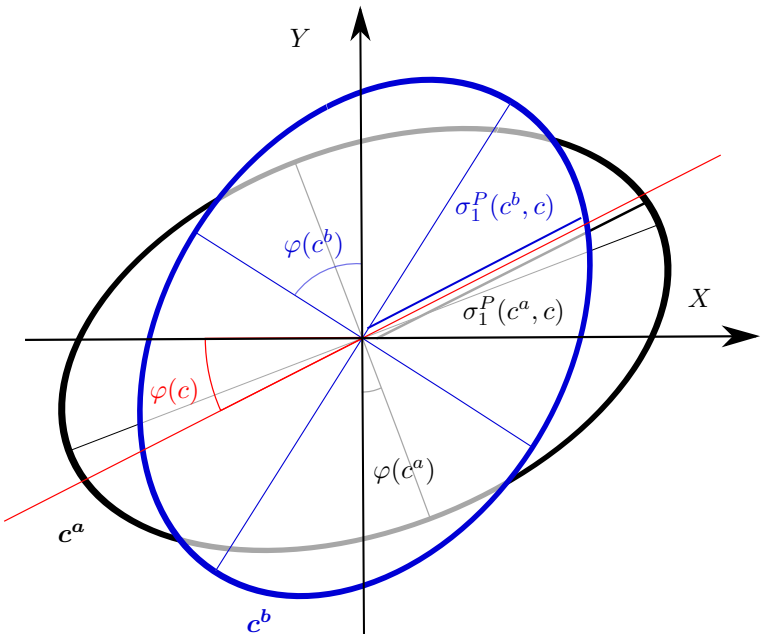

Figure 13: Sketch of calculation of semi-widths contributions from two, misaligned segment level clusters (shown in black and blue) towards computing the semi-major, vectorial width of the containing semi-final cluster, along the average cluster direction (shown in red).

semi-minor width, thus giving a contribution of segment-level cluster c^s towards the vectorial-weighted semi-minor width of the semi-final cluster c containing it. This calculation procedure is partly illustrated by Fig. 13.

In addition to the above shape variables, the energy content of semi-final clusters is given by:

$$E(c_j^H) = \sum^{s\in\mathrm{HGS}} \epsilon(c_{j^{sH}}^s), \quad E(c_j^L) = \Gamma\left(\sum^{s\in\mathrm{LGS}} \epsilon(c_{j^{sL}}^s)\right), \tag{50}$$

for HG (left) and LG (right) semi-final clusters. The energy of high-granularity semi-final clusters is just a sum of the amplitude content of the contained segment-level clusters. It does not provide reliable information about the original energy of the particle. On the other hand, the energy of low-granularity semi-final clusters is the calibrated sum of the amplitude content of the contained segment-level clusters, where the calibration function (38) is being used.

Another variable describing the longitudinal shower development is introduced at the level of semi-final LG clusters:

$$\Lambda(c_j^L) = \frac{\epsilon(c^5_{\mathrm{merL}(j,5)})}{\epsilon(c^2_{\mathrm{merL}(j,2)}) + \epsilon(c^4_{\mathrm{merL}(j,4)})} \tag{51}$$

which is simply the ratio between the energy content in (LG) segment 5 and the sum of the energy contents of (LG) segments 2 and 4, for semi-final LG cluster c_j^L. This variable is intended for discriminating hadronic showers from electromagnetic ones, the former typically starting to develop later in an electromagnetic calorimeter, thus having larger values of Λ.

From here on, all equations describe properties of final clusters, resulting either from the combination of an LG semi-final cluster and a HG semi-final cluster, or just from an LG semi-final cluster alone. The

notations $j^{FH} := \text{merFH}(j)$ and $j^{FL} := \text{merFL}(j)$ will be consistently used throughout the remaining equations for indicating correspondences between semi-final and final clusters, with merFH and merHL defined in Section 3.4.

The transverse coordinates of a final cluster:

$$X(c_j^F) = x(c_{j^{FH}}^H)\text{hasHG}(c_j^F) + x(c_{j^{LH}}^L)(1 - \text{hasHG}(c_j^F)) \tag{52}$$

and:

$$Y(c_j^F) = y(c_{j^{FH}}^H)\text{hasHG}(c_j^F) + y(c_{j^{LH}}^L)(1 - \text{hasHG}(c_j^F)) \tag{53}$$

are given by the transverse coordinates of the contributing HG semi-final cluster, if this exists, otherwise by the transverse coordinates of the contributing LG semi-final cluster, as encoded by the hasHG notation introduced in Section 3.4. The transverse coordinates of the semi-final clusters are computed via (39) and (40) above.

The semi-major/minor HG width of a final cluster:

$$\sigma_{1/2}^H(c_j^F) = \sigma_{1/2}^H(c_{j^{FH}}^H)\text{hasHG}(c_j^F) - (1 - \text{hasHG}(c_j^F)) \tag{54}$$

is given by the semi-major/minor width of the contributing HG semi-final cluster, if this exists, otherwise it is set to the value of -1, which is outside the meaningful range. On the other hand, the semi-major/minor LG width of a final cluster:

$$\sigma_{1/2}^L(c_j^F) = \sigma_{1/2}^L(c_{j^{FL}}^L) \tag{55}$$

is always given by the semi-major/minor width of the contributing LG semi-final cluster, as every final cluster must have a contributing semi-final LG cluster. The semi-widths of semi-final clusters are computed via (41) and (42) above.

Similarly, the HG transverse alignment of a final cluster:

$$r_\varphi^H(c_j^F) = r_\varphi^H(c_{j^{FH}}^H)\text{hasHG}(c_j^F) - (1 - \text{hasHG}(c_j^F)) \tag{56}$$

is given by the transverse alignment of the contributing HG semi-final cluster, if this exists, otherwise it is set to the value of -1, which is outside the meaningful range. On the other hand, the LG transverse alignment:

$$r_\varphi^L(c_j^F) = r_\varphi^L(c_{j^{FL}}^L) \tag{57}$$

is always given by the transverse alignment of the contributing LG semi-final cluster, as every final cluster must have a contributing semi-final LG cluster. The transverse alignment of semi-final clusters are computed via (45) and (46) above.

The same logical scheme is followed by the HG transverse orientation of a final cluster:

$$\varphi^H(c_j^F) = \varphi(c_{j^{FH}}^H)\text{hasHG}(c_j^F) - 2(1 - \text{hasHG}(c_j^F)) \tag{58}$$

with a trivial value of -2, together with the LG transverse orientation of the final cluster:

$$\varphi^L(c_j^F) = \varphi^L(c_{j^{FL}}^L) \tag{59}$$

The transverse orientations of semi-final clusters are computed via (43) and (44) above.

Finally, the vectorial weighted semi-major/minor width of a final clusters are given by:

$$\hat{\sigma}^H_{1/2}(c^F_j) = \hat{\sigma}^H_{1/2}(c^H_{j^{FH}})\text{hasHG}(c^F_j) - (1 - \text{hasHG}(c^F_j)) \tag{60}$$

when using the high-granularity information and by:

$$\hat{\sigma}^L_{1/2}(c^F_j) = \hat{\sigma}^L_{1/2}(c^L_{j^{FL}}) \tag{61}$$

when using the low-granularity information. The vectorial widths of semi-final clusters are computed via (47) and (48) above.

In addition to the above shower-shape variables, the calibrated (GeV) energy of a shower (or final cluster) is given by:

$$E(c^F_j) = \left(\frac{E(c^H_{j^{FH}})}{\sum\limits_{l}^{l^{FL}=j^{FL}} E(c^H_{l^{FH}})} \text{hasHG}(c^F_j) - (1 - \text{hasHG}(c^F_j)) \right) E(c^L_{j^{FL}}) \tag{62}$$

which essentially says that energy of the final cluster is simply the energy of the contributing LG semi-final cluster, if no HG semi-final cluster is contributing, or by a fraction of the energy of the contributing LG semi-final cluster, if such a HG semi-final cluster exists. This fraction is simply given by the ratio of the amplitude of the associated HG semi-final cluster and the sum of all other HG semi-final clusters merged with the contributing LG semi-final cluster – if there are no other HG semi-final clusters merged with the contributing LG semi-final cluster, the fraction is 1. This procedure is crucial for deciding about the energy content of overlapping electromagnetic showers that can only be separated by the HG segments and not by the LG segments, such as those produced by decay photons originating from the same neutral pion.

The longitudinal energy ratio of the final cluster is simply the longitudinal energy ratio of the associated semi-final LG cluster:

$$\Lambda(c^F_j) = \Lambda(c^L_{j^{FL}}) \tag{63}$$

which is defined via (51).

3.8 Higher-Level Analysis

The term "higher-level analysis" refers here to the process of combining information associated to different reconstructed showers within single events, possibly also using Monte Carlo truth information about actual particle tracks, in order to create data structures that allow for the figures in Sections 4 and 5 to be generated. This procedure is carried out by making use of the results of the shower reconstruction algorithm explained in Section 3.4. It can use, for every event, either the set of final clusters C^F or set of semi-final LG clusters C^L, depending on whether the HG information is considered or ignored – in Sections 4 and 5, this distinction is often specified by "full clustering" vs "coarse clustering", or by "HG+LG clustering" vs "LG clustering".

The higher-level analysis procedure depends to a great extent on the type of events studied. Here, only single particle events and proton-proton collision events are studied.

For single-particle events, shower information is stored for every event and within every event for every generated secondary track, if any, or for the primary track otherwise. The intersection of every such track with the z_{FCP} plane is computed, where the positions of the showers (final clusters or semi-final LG clusters) are also projected. Every track is matched to the closest shower position, within the z_{FCP} plane. The match is considered valid for further analysis if the distance between the intersection of the track and the shower position D_M is smaller than a certain threshold. Every track can be assigned to at most one shower and every shower can be assigned to at most one track. All variables defined at the level of the matched shower are kept at the level of the respective track, together with further Monte Carlo truth information concerning the actual particle traveling along that track. In addition, for every single π_0 for which two showers are reconstructed by the detector, the invariant mass $\text{IM}^{\gamma-\gamma}$ is calculated according to (64). Single-particle simulated events are mostly used as a test ground for discriminating variables, which allows for such heavy use of Monte Carlo truth information.

On the other hand, proton-proton collision events are mainly used for evaluating the performance of the simulated detector, together with its clustering algorithm, analysis algorithm and higher-level cuts that are imposed on discriminating variables. This requires a data structure that is much more detector oriented. For this reason, in the final data structure that is used for creating the relevant plots, information is stored shower-by-shower, regardless of the event where the shower is produced (every shower is, again, either a final cluster or a semi-final LG cluster). However, before putting all showers together, information from multiple showers is combined within the same event, with the purpose of computing shower-level variables which depend on other showers produced within the same event. These variables are the invariant mass in Eq. (65) and the transverse momentum within the isolation cone in Eq. (66). At the same time, truth information concerning tracks and particles propagated along them is accessed. For every shower, information is saved about the track intersecting the z_{FCP} plane at the smallest distance D_M from the projected shower position.

For two showers c^a and c^b (either final clusters or LG semi-final clusters), the invariant mass calculation is carried out according to the following formula:

$$\text{m}^{a,b} = \sqrt{E_a E_b - 2\vec{p}_a \cdot \vec{p}_b} \tag{64}$$

which is valid under the assumption that each of the two showers is produced by a photon. E_a and E_b are the calibrated shower energies and $\vec{p}_a$ and $\vec{p}_b$ are the associated 3-momenta. The direction of each 3-momentum is extracted from shower position information, under the simplifying assumption that the particle originates at the interaction point, which is reasonable enough for photons decaying from a π_0, since the latter is very short lived. The magnitude of each 3-momentum is again given by the cluster energy, which again makes use of the assumption that particle responsible for a shower is a photon.

For a (final / LG semi-final) cluster $c_j^{F/L} \in C^{F/L}$ obtained from a specific proton-proton event, the associated invariant mass is given by:

$$\text{IM}(c_j^{F/L}) = m^{j,k}. \left(|m^{j,k} - m^{\pi_0}| = \min \left\{ |m^{j,q} - m^{\pi_0}| \Big/ c_q^{F/L} \in C^{F/L} \right\} \right) \tag{65}$$

with $m^{j,k}$, $m^{j,q}$ computed according to (64). Equation (4.2) simply states that the invariant mass value associated to shower $c_j^{F/L}$ is the one that is closest to the invariant mass of a neutral pion and that can be realized by pairing $c_j^{F/L}$ with any other shower $c_q^{F/L}$ produced during the same event. This implies that the higher-level analysis requires at least $N_C^{F/L} \times N_C^{F/L}$ steps for every proton-proton event, where $N_C^{F/L}$ is the number of showers saved for that event, in order for all possible shower-shower pairs to be evaluated in terms of the resulting invariant mass

For a (final / LG semi-final) cluster $c_j^{F/L} \in C^{F/L}$ obtained from a specific proton-proton event, the associated transverse momentum content of its isolation cone is computed according to:

$$p_T^{\text{iso}}(c_j^{F/L}) = \sum_{k}^{R^{j,k}<R^{\text{iso}}} p_T(c_k^{F/L}) \tag{66}$$

where $p_T(c_k^{F/L})$ is simply the transverse projection of the 3-momentum associated to shower $c_k^{F/L}$, computed in the same manner as $\vec{p}_a$ and $\vec{p}_b$ employed by (64) (see explanations below the equation). R^{iso} is an dimensionless analysis parameter controlling the opening of the isolation cone. $R^{j,k}$ is the distance between showers $c_j^{F/L}$ and $c_k^{F/L}$ in the pseudorapidity-azimuthal space (η, ϕ) (k iterates over all other showers in the event) computed according to:

$$R^{a,b} = \sqrt{(\eta_a - \eta_b)^2 + (h(\phi_a - \phi_b))^2} \tag{67}$$

where η_a, η_b and ϕ_a, ϕ_b are the pseudorapidities and azimuthal angles of the tracks supposed to have generated the two showers involved, geometrically obtained from the shower positions, using again the assumption that the particles originate at the interaction point. The function h above is just accounting for the periodicity of the azimuthal angle.

4 Neutral Pion Recognition in Single-Particle Simulations

This section is mainly concerned with exploring the two main ways of recognizing π_0 particles, namely invariant-mass analysis and shower-shape analysis. The former is treated in Sec. 4.2 while the latter is treated in Sec. 4.3. At the same time, the advantage of using high-granularity information is pointed out, first at the level of the clustering and shower-reconstruction algorithm and second at the level of computing shower-shape variables. This is relevant for the invariant mass analysis and for the shower shape analysis respectively. It is implied that invariant-mass analysis can only be used to reject electromagnetic showers generated by π_0 decays that are resolved (separated or disentangled) by the clustering and shower-reconstruction algorithm, while the shower-shape analysis is mainly relevant for those that are not resolved [10]. Before that, certain optimizations affecting the quality of shower position and shape reconstruction are treated in Sec. 4.1. The analysis sequence consisting of the clustering and shower-shape reconstruction will be often collectively referred to as "the LG/(HG+LG) clustering (algorithm)" for simplicity, with the "LG" or "LG+HG" being used to indicate whether high-granularity information is either accessed or not (sometimes, this is also indicated by "semi-final LG clusters" vs "final clusters").

For the studies conducted in this section, 6 single particle data samples are being used, 3 with single γ simulations and 3 with single π_0 simulations. All of them are carried out with the same detector configuration with FoCal placed at the 4m position. In addition, for all 6 data samples the original particles follow the same pseudorapidity distribution with $3.5 < \eta < 4.5$, with a uniform θ distribution – implying that more particles are generated at high η than at low η. The notations associated to the 6 sample for further use here, along with further kinematic and statistical characteristics are listed below:

- samp^A_γ, with uniform transverse momentum $0 < p_T(\text{GeV}/c) < 6$, 10000 events
- $\text{samp}^A_{\pi_0}$, with uniform transverse momentum $4 < p_T(\text{GeV}/c) < 10$, 10000 events
- samp^B_γ, with almost constant momentum (and energy) $p = E/c \approx 300\text{GeV/c}$, 1000 events
- $\text{samp}^B_{\pi_0}$, with almost constant momentum (and energy) $p \approx E/c \approx 300\text{GeV/c}$, 4000 events
- samp^C_γ, with almost constant momentum (and energy) $p = E/c \approx 500\text{GeV/c}$, 1000 events
- $\text{samp}^C_{\pi_0}$, with almost constant momentum (and energy) $p \approx E/c \approx 500\text{GeV/c}$, 2000 events

with the transverse momentum ranges of samples samp^A_γ and $\text{samp}^A_{\pi_0}$ having been originally chosen such that the efficiency of the clustering algorithm in detecting single γs and in separating single π_0s to be studied in parallel for clusters of comparable transverse momentum, given that the reconstructed transverse momentum of a neutral pion is divided among the two (supposedly identified) clusters. However, such aspects are not the main topic of Section 4.

For the studies conducted here, for events in the γ samples are selected whenever the true particle track intersects the $z_{\text{min}} \approx 361\text{cm}$ plane within a radius $8\text{cm} < r < 80\text{cm}$ from the $z = 0$ axis surrounded by the beam pipe, which delimits an area fully covered by the detector. In parallel, events in the π_0 samples are selected whenever both tracks associated to the two decay photons satisfy these conditions. In practice, most events pass this selection.

For further use below, it is useful to define the notion of match distance (position mismatch or track-to-cluster distance) denoted by D^M. For single-particle events, this is a variable which uses Monte Carlo (MC) truth information, defined at the level of every photon track (primary track, for single γ events,

[10] Seldom, based on LG shower variables, the latter may be useful in rejecting π_0 showers that are only disentangled in the HG segments but not in the LG segments, for which the reconstructed invariant mass falls outside the imposed window around the π_0 peak, due to errors originating in dividing the energy content of the LG semi-final cluster between the two final clusters generated in combination with the merged semi-final HG clusters (with "merged" used in the sense of combining HG and LG information, as in Sec. 3)

secondary for single π_0 events), indicating the distance, in the $z_{\rm FCP}$ plane between the true impact position of the respective particle and the position of the closest reconstructed cluster, if such a cluster exists without having been matched to another track associated to a higher-energy particle.

4.1 Optimizing Shower Position and Shape Reconstruction

This section is concerned with optimizing the position and shape reconstruction capabilities of the clustering and shower-reconstruction algorithms described in Sections 3.3 and 3.4 respectively, together with the cluster and shower-shape calculation procedure described in Sections 3.5 and 3.7 respectively. Section 4.1.1 focuses on information from the LG segments while Sec. 4.1.2 focuses on information from the HG segments. Although the optimizations made in these two Sections, translated to certain values of the parameters used for weighting the amplitudes of digits, are used in Sections 4.2.1, 4.2.2, 4.3.1 and 4.3.2, they have been investigated in the first place because of certain unsatisfactory results which are explained in Sections 4.3.1 and 4.3.2 respectively.

4.1.1 Low Granularity Segments (Logarithmic Weighting of Digit Amplitudes)

Here, the process of optimizing the shower position and shower shape reconstruction is illustrated, when using only the LG segment information. Thus, only information associated to semi-final low granularity (LG) clusters is being used for producing the plots, thus using the results of LG clustering. The study is restricted to the samp_γ^A sample mentioned above, while focusing on the improvements that can be achieved by employing logarithmic weighting instead of linear weighting for the amplitudes of digits generated in the LG segments.

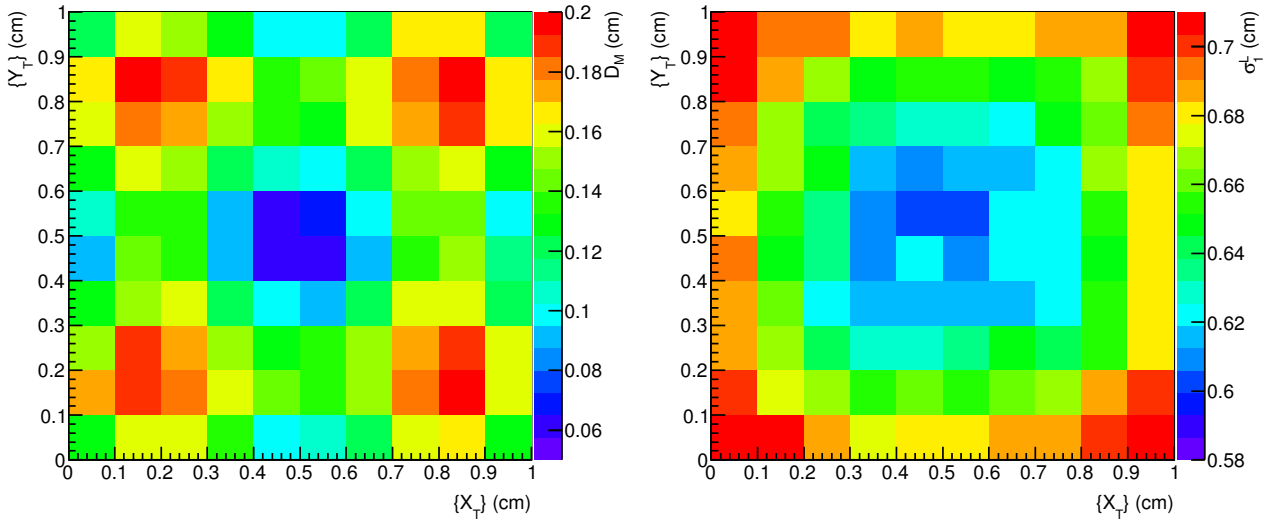

Figure 14: Bin averaged relative distance D_M between the γ impact point and reconstructed LG semi-final cluster position (left) and (bin averaged) reconstructed LG semi-final cluster semi-major width σ_1^L (right), as a functions of the true fractional coordinates $\{X_T\}$ and $\{Y_T\}$ in the z'_{FCP}, using linear LG weighting.

Cluster reconstruction artifacts are probably best illustrated in figure 14. On the left-hand side, it shows how the (bin averaged) mismatch between reconstructed cluster position and true position depends on

the true, relative position with respect to the coarse lattice of the LG segments. On the right-hand side, it shows how the (bin averaged) reconstructed cluster semi-major width depends on the same true, relative position in the lattice. One can notice that the position mismatch has clear maxima in the vicinity of the four corners of the square lattice, and a clear minimum in the center of the lattice. This implies that the positions of those particles hitting a cell close to the center are well reconstructed, while the positions of photons hitting the detector in the proximity of a corner of a LG cell is poorly reconstructed. This tendency of worsening position reconstruction is however alleviated when getting very close to the edge of a cell. On the other hand, the semi-major width of reconstructed clusters is low when the photon hits the a cell close to the center and increases as the true position gets closer to one of the edges.

The true, relative positions with respect to the LG lattice $\{X_T\} := X_T - [X_T]$, $\{Y_T\} := Y_T - [Y_T]$ are simply the fractional parts of the true X_T and Y_T coordinates of the intersection between the true particle track and the z'_{FCP} transverse plane, given that LG lattice would have a corner exactly at $X = 0$ and $Y = 0$ if the detector was to also cover that region (which is actually occupied by the beam-pipe).

The effects described above are in agreement with those reported for instance by Ref. [27] and can be easily understood by considering how the energy of a photon is distributed among the LG cells in the vicinity of the true photon impact position. When the photon hits the lattice close to the center of a cell, energy is equally distributed among neighboring cells, with a much lower amount than in the respective cell. In this manner, there is no systematic bias in reconstructing the position, as the neighboring cells have approximately equal contributions, while the reconstructed width has relatively low values, on average, since neighboring cells have relatively low contributions. When the photon hits very close to the edge of a cell, energy is equally distributed among the two cells separated by the respective edge. The position is reconstructed very close to the edge as well thus leading to very low bias, while the width (or at least its projection along the direction perpendicular to the respective edge) is reconstructed with systematically larger values, since there are two cells with approximately equal contributions. When the photon hits far from the center but also far from the edge, there is a systematic bias in position, since the energy deposited in the cell following that respective edge does not grow linearly with decreasing separation of the position of incidence from that edge. The width on the other hand is reconstructed with values in between those of the former two situations.

This lattice dependence of position and shape reconstruction, is removed when employing logarithmic weighting of the amplitude contributions of digits towards clusters, as suggested by Ref. [27]. This procedure is employed since it deals with all artifacts discussed here with a single formula, while also improving the γ vs π_0 discriminating power of the main LG discriminating variable, as shown in Sec. 4.3.1 below. Instead, analytical corrections could also be considered, but should have been applied independently to position and width reconstruction and they typically lead to poorer results when reconstructing clusters originating from non-perpendicular particle impacts. Moreover, such corrections would not be valid for combined π_0 showers that are not separated by the clustering algorithm, since two photon incidence positions are involved in such cases. Using logarithmic instead of linear weighting translates to employing (20) instead of (19) for computing the cluster description weights introduced in Section 3.3. Equation (20) also makes use of a free parameter w_0 which should be optimized. Here, this is simply done by studying how w_0 affects the average match distance $\langle D_M \rangle$ and the average cluster asymmetry $\langle A^L \rangle$. It is desirable that both these quantities are minimized, meaning that the the positions of the clusters are on average closest to the photon impact positions and that their shapes are as close as possible to being circular. The effect of different values of w_0 on reducing lattice reconstruction artifacts is then illustrated.

Figure 15 shows the distribution of the clusters produced by all single-γ events with respect to match distance D_M (on the left hand side) and with respect to cluster asymmetry A_L (on the right hand side), for linear weighting and for logarithmic weighting with two w_0 parameters. The asymmetry of a semi-final coarse cluster is defined as:

$$A^L := A(c_j^L) = \frac{\sigma_1(c_j^L) - \sigma_2(c_j^L)}{\sigma_1(c_j^L) + \sigma_2(c_j^L)} \tag{68}$$

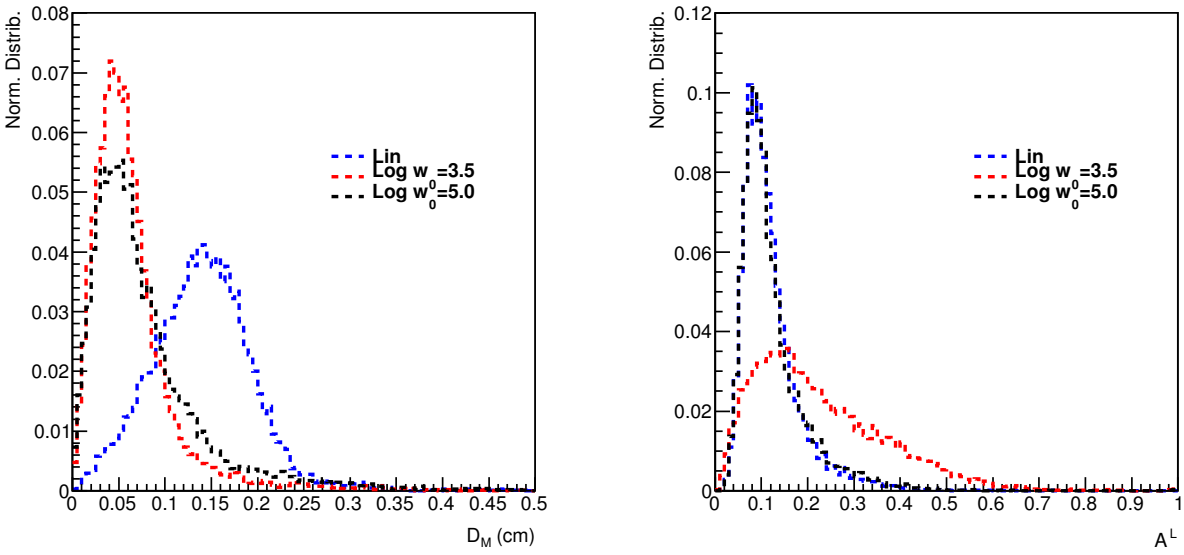

Figure 15: Distribution of LG semi-final clusters with respect to match distance D_M (left) and with respect to cluster asymmetry A^L (right), for linear LG weighting and for logarithmic LG weighting with parameter $w_0 = 3.5$ and $w_0 = 5.0$.

with semi-widths σ_1 and σ_2 computed according to (42). Note that the asymmetry is positive $A^L \geq 0$ since the condition $\sigma_1^L \geq \sigma_2^L$ is always enforced.

The plot on the left of Fig. 15 shows that logarithmic weighting dramatically improves position reconstruction, while the plot on the right shows that it may significantly alter cluster shape reconstruction for certain w_0 parameters. Optimizing w_0 thus translates to a trade-off between decreasing the average match distance as much as possible, while still keeping the average asymmetry reasonably low.

This optimization process is best illustrated in figure 16, which shows how the average match distance $\langle D_M \rangle$ (left) and the average cluster asymmetry $\langle A^L \rangle$ depend on the w_0 parameter employed by the logarithmic weighting procedure. The horizontal lines show the values obtained via linear weighting. Note that the lowest average match distance is obtained for $w_0 \approx 4.0$, which is more then two times lower than the value provided by linear weighting. On the other hand, the lowest asymmetry value is obtained for $w_0 \approx 4.5$, which matches the value obtained when using linear weighting. Based on this information, the value $w_0 = 4.3$ is taken as optimal and used for later studies, as it provides a good trade-off between position reconstruction and shape reconstruction.

Figure 17 shows how logarithmic weighting removes the lattice artifacts pointed out in Fig. 14. Different curves correspond to different type of weightings (either linear or logarithmic with different values of w_0), with colors used consistently on the two sides. On both sides of Fig. 17, the horizontal axis shows the true, relative photon impact position with respect to the LG lattice, along the X-axis. For both plots, a selection of events is made on the true relative position along the Y-axis, such that only those photons hitting the detector within the central, horizontal band of a LG cell are considered. In this sense, the two plots of Fig. 17 can be understood as X-axis projections of the central, horizontal bands of the two corresponding two plots in Fig. 14. In contrast to the Z-axes of Fig. 14, the Y-axes of 17 show the (bin averaged) mismatch distance along the X-axis (on the right-hand side), which is allowed to take negative values, together with the cluster semi-width along the X-axis (on the left-hand side), rather than the (bin averaged) absolute mismatch distance and semi-major width respectively. All these transformations are employed in order to isolate the non-uniformities present along the X-axis. The X semi-width σ_x^L of

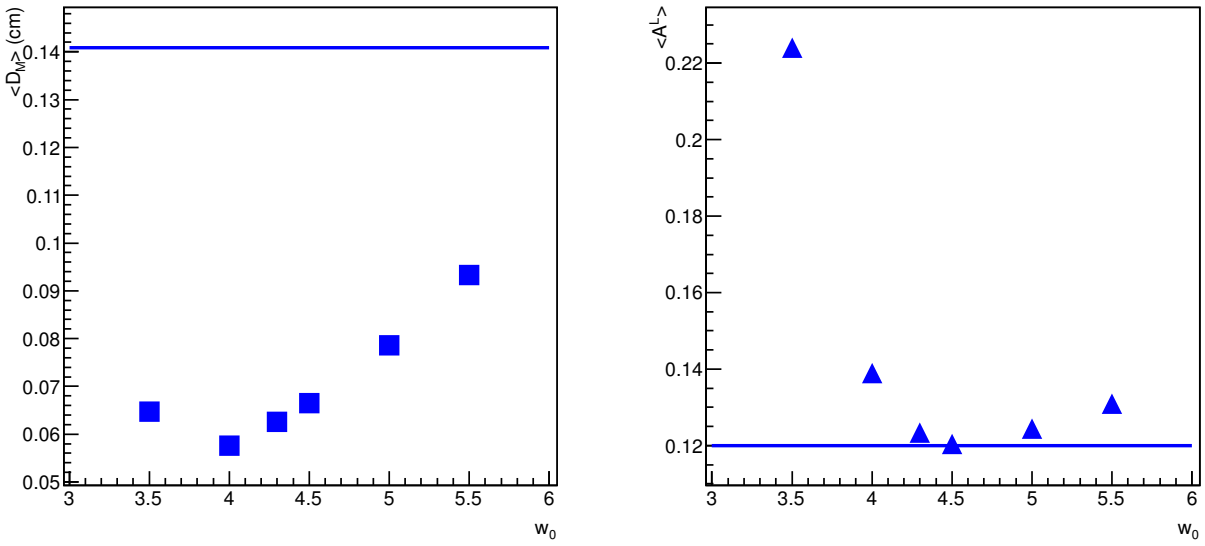

Figure 16: Overall average match distance D_M for reconstructed LG semi-final cluster position (left) and overall average reconstructed LG semi-final cluster asymmetry A^L (right) as a functions of logarithmic LG weighting parameter w_0; in parallel, the horizontal lines show the values obtained via linear LG weighting

.

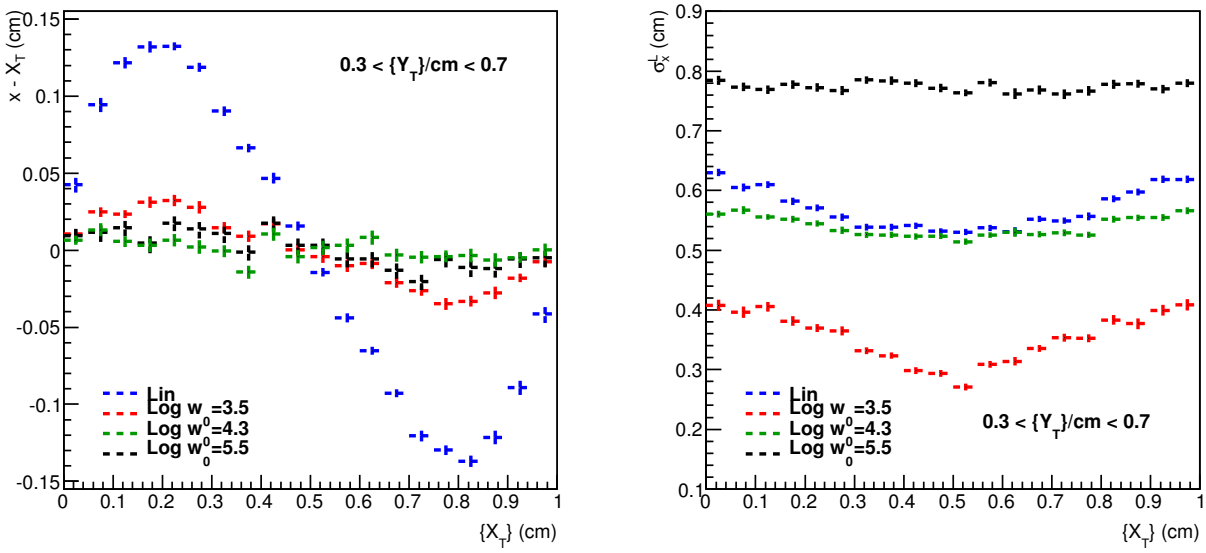

Figure 17: Bin averaged relative X-missmatch $X_M(c_j^L)$ for reconstructed LG semi-final cluster positions (left) and (bin averaged) reconstructed LG semi-final cluster X-width σ_x^L (right), as a functions of the true horizontal fractional coordinate $\{X_T\}$, for a constrained vertical fractional coordinate $0.3 < \{Y_T\} < 0.7$ in the z'_{FCP}, for linear weighting and for logarithmic LG weighting with different values for parameter w_0.

semi-final LG clusters plotted on the right-hand side of Fig. 17 is recovered according to:

$$\sigma_x^L := \sigma_x(c_j^L) = \left(\sqrt{\frac{\cos^2(\varphi(c_j^L))}{\sigma_1^2(c_j^L)} + \frac{\sin^2(\varphi(c_j^L))}{\sigma_2^2(c_j^L)}} \right)^{-1} \tag{69}$$

where the corresponding semi-major width σ_1^L, semi-minor width σ_2^L and orientation φ^L are given by equations (42) and (44).

On the left-hand plot of Fig. 17, it can be seen that photons hitting close to the center of an LG cell, as well as those photons hitting very close to the edge between two LG cells, have small (absolute) position mismatch, while the others have large mismatch, when linear weighting is used. The fact that $x - X_T$ takes positive values for $\{X_T\} < 0.5$ and negative values for $\{X_T\} > 0.5$ signals that cluster centers tend to be systematically reconstructed closer to cell centers than the actual impact positions. The right-hand plot shows that the X-spread of clusters is larger close to the edges than in the center when linear weighting is used. Both these effects are removed or at least significantly reduced by logarithmic weighting with parameter $w_0 = 4.3$, which has been found optimal above. Note, on the right-side, that the average semi-width along the X direction increases with increasing w_0 for logarithmic weighting, which is understandable since w_0 also acts as a threshold on lower-amplitude contributions from distant cells, and the larger the w_0 the more these contributions are included and emphasized, thus leading to larger cluster widths. When w_0 is too low, the non-uniformity of σ_x^L with respect to $\{X\}$ is actually enhanced, as higher-amplitude contributions from neighboring cells start being discarded. By symmetry, the same behavior would be observed if everything was plotted with respect to $\{Y\}$ while defining the selection with respect to $\{X\}$, namely by looking along central, vertical bands in the LG cells instead of central horizontal bands. A similar behavior is observed when looking at the combined statistics from the complementary two bands in $\{Y\}$ or $\{X\}$, with artifacts becoming more visible for logarithmic weighting with $w_0 = 3.5$ (red curve).

At this point, one might wonder how the observations made in relation to the right-hand plots of Figures 14 and 17 can be compatible with the observation made in relation to the right-hand plot of Figure 16. The latter suggests that linear weighting is better or as good as logarithmic weighting in terms of cluster asymmetry, while the former suggests that linear weighting introduces artifacts in reconstructing cluster shapes, which can removed by logarithmic weighting to a great extent. The two aspects are partly reconciled when realizing that asymmetry is defined in terms of the semi-major an semi-minor axes, which are not necessarily aligned with the LG lattice. In this manner, the observed increase (Fig. 17) in shower width along the direction perpendicular to the edge when the photon hits close to the edge, which is also responsible for the observed increase in the semi-major width (Fig. 14) also leads to an increase in the semi-minor width (not shown here directly), which keeps the asymmetry small. However, this argument is weakened by the observation made below in relation to Fig. 18, which shows that a bias in the reconstructed cluster orientation is also introduced by linear weighting, which is such that the orientation towards the direction perpendicular to the edge next to the photon impact position is preferred. This aspect has not been completely explored nor understood here, and could be more easily studied using approaches such as the one employed by [29], according to which 2D profiles resembling clusters induced by electromagnetic decays in a calorimeter are directly generated using a simple Monte Carlo methods, with the same width and position reconstruction method being applied on the "digits" resulting from integrating the "signal" within a coarse 2D lattice.

The influence of the square LG lattice on cluster reconstruction also becomes apparent via its biasing effect on the cluster orientation. Fig. 18 shows the distribution of LG semi-final clusters with respect to the transversal orientation angle φ^L, computed according to (44). Again, different curves correspond to different kinds of weightings (either linear or logarithmic with different values for the w_0 parameter). It becomes apparent that both linear weighting and logarithmic weighting with low w_0 entail a distribution that is highly peaked around the values of $\varphi^L = 0$ and $\varphi^L = \pm\frac{\pi}{2}$, instead of a uniform distribution. This implies that the ellipses matched to the reconstructed clusters tend to be aligned either with the X-axis or with the Y-axis much more than they should. The effect is directly related to the width over-estimation

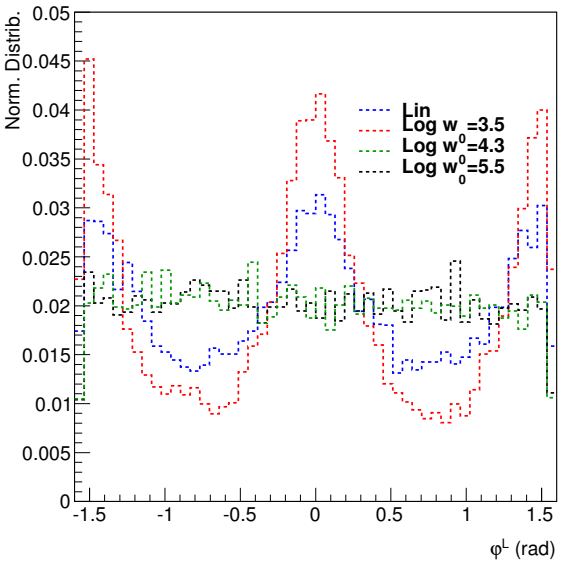

Figure 18: Distribution of LG semi-final clusters with respect to average LG cluster orientation φ^L, for LG linear weighting and for logarithmic LG weighting with different values of w_0.

in the vicinity of a lattice edge pointed out in relation to the right hand side of Fig. 17. Specifically, photons hitting the detector in the vicinity of an Y-aligned lattice edge tend to generate clusters with over-estimated X-widths, thus determining a preferred X-alignment. Conversely, photons hitting the detector in the vicinity of an X-aligned lattice edge tend to generate clusters with over-estimated Y-widths, and thus with a preferred Y-alignment This angular bias thus also vanishes for the optimal $w_0 = 4.3$, as can be seen in Fig. 18.

4.1.2 High Granularity Segments (Power-Law Weighting of Digit Amplitudes)

In this section, the process of optimizing the shower position and shower shape reconstruction is illustrated, when using the HG segment information. Information associated with final clusters is being used for producing the plots, thus using the results of HG+LG clustering. Although the final clusters also contain LG information, the position is given solely by the information in the HG layers. This assumes that every final cluster has an associated semi-final HG cluster contributing to it, which is reasonable since the study is restricted to the samp_γ^A sample mentioned above. The study focuses on the improvements that can be achieved by employing power-law weighting instead of linear weighting for the amplitudes of digits generated in the HG segments. Investigating power-law weighting of amplitudes in the HG segments was motivated by an unexpected lack of discriminating power of the HG shower-shape variables when linear weighting is employed, which is here described at the beginning of Section 4.3.2. When dealing with the HG segments, it does not make sense to show reconstructed final cluster position and width with respect to the fractional transverse coordinates, as done in Sec. 4.1.1 with semi-final clusters, since photon tracks (which are mostly not aligned with the Z-axis) impacting one HG segment at one position relative to the HG lattice would impact the other HG segment at a different position, due to the small (1mm) spacing of the HG lattice.

Figure 19 shows the distribution of final clusters generated by single photon events with respect to the mismatch distance between the true photon impact position and the position of the reconstructed cluster in the z'_{FCP} transversal plane (left) together with the distribution with respect to the asymmetry of the associated HG semi-final cluster (right). Different curves correspond to different procedures of weighting the amplitudes of digits in the HG segments. Either linear (19) or power-law (21) weighting

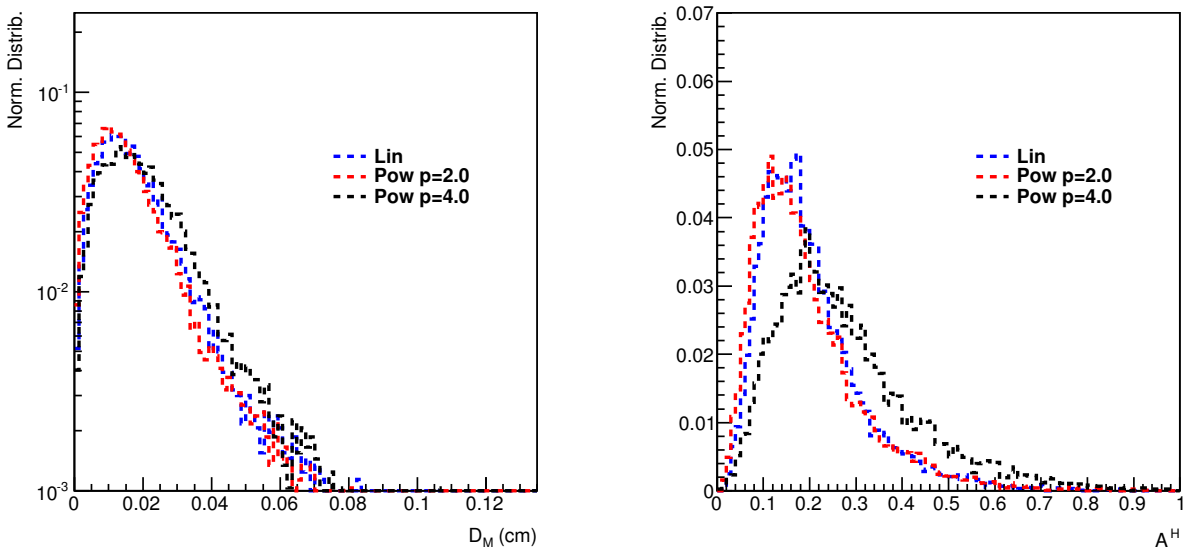

Figure 19: Distribution of final clusters with respect to match distance D_M (left) and with respect to cluster HG asymmetry A^H (right), for linear HG weighting and for power-law HG weighting with parameter $p = 2.0$ and $p = 4.0$.

is employed, where the latter makes use either of parameter $p = 2.0$ or $p = 4.0$. The plots show that power-law weighting with $p = 2.0$ slightly shifts both distributions towards lower values, compared to linear weighting, while power-law weighting with $p = 4.0$ shifts both distributions towards higher values. The HG asymmetry values used on the right side of Fig. 19 are defined as:

$$A^H := A^H(c_j^F) = \frac{\sigma_1^H(c_j^F) - \sigma_2^H(c_j^F)}{\sigma_1^H(c_j^F) + \sigma_2^H(c_j^F)} \tag{70}$$

where σ_1^H and σ_2^H are computed according to (54).

Figure 20 illustrates an optimization procedure similar to the one in Fig. 16. On the left-hand side, the average mismatch distance is shown as a function of the parameter governing the power-law weighting in the HG segments. On the right-hand side, the average asymmetry is shown in parallel. The horizontal line shows the corresponding values for the case when linear weighting is employed. It becomes apparent that a value of $p = 2.0$ is optimal in the sense of minimizing the match distance, while a value of $p = 2.5$ is optimal in the sense of minimizing the HG asymmetry. The value of $p = 2.2$ is chosen as a convenient compromise and employed for further studies. Note that the effect in improving position reconstruction in the HG segments introduced by this optimization procedure is much smaller than the one introduced by the optimization procedure for the LG segments – compare the right-hand plots of Fig. 20 and of Fig. 16. This is partly due to the fact that when computing the transverse positions of the semi-final HG clusters, much of the fluctuations in tails of the clusters, which are attenuated by the power-law weighting when computing the shape variables, are already eliminated by the cutoff of the ringer distance with respect to the positions of the associated seeds, as explained in Sec. 3.5 in relation to Eq. 27.

Figure 21 shows that larger values of the parameter p introduces lattice reconstruction artifacts in the HG segments. On the left-hand side, the distribution of final clusters with respect to the orientation of the contributing segment-level cluster in (HG) segment $s = 1$ is shown, for linear weighting and for power-law weighting with different p values. The plot on the right-hand side shows the same thing for (HG) segment $s = 3$. It becomes clear that power-law weighting with $p = 3.0$, $p = 4.0$ biases the distributions towards

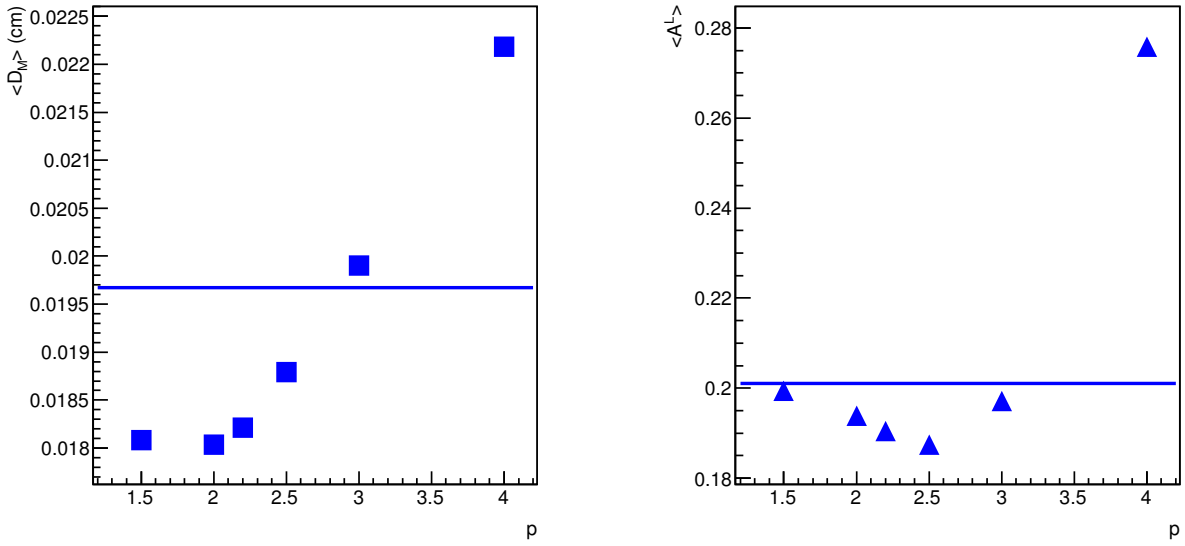

Figure 20: Overall average match distance D_M for reconstructed final cluster positions (left) and (overall) average reconstructed final cluster asymmetry A^H (right), as a functions of power-law HG weighting parameter p; in parallel, the horizontal lines show the values obtained via linear HG weighting

.

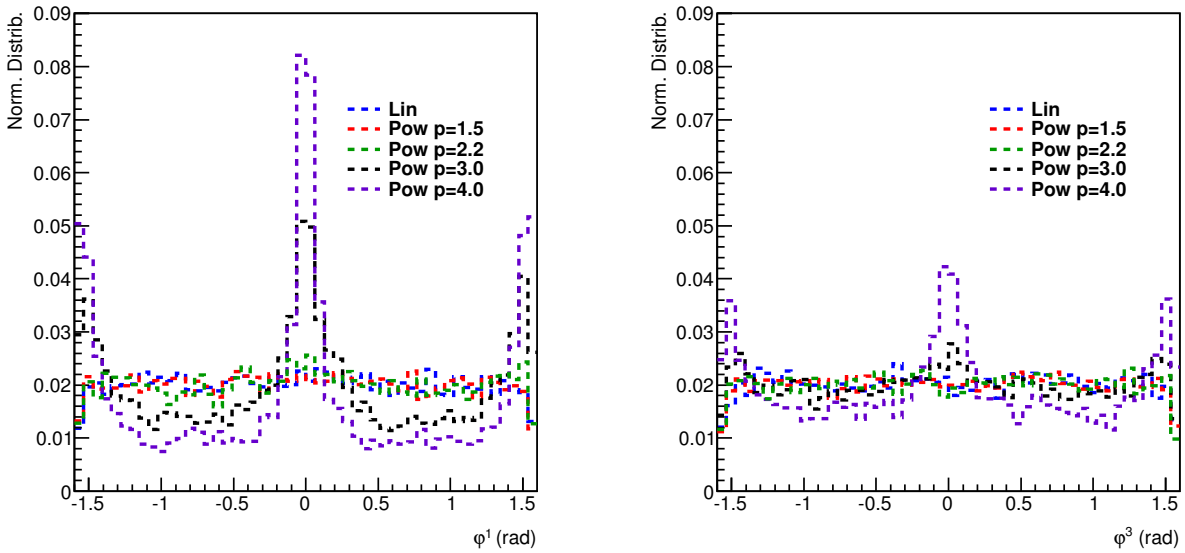

Figure 21: Distribution of final clusters with respect to HG cluster orientation in (HG) segment 1 $\varphi^1 := \varphi(c^1_{j^{1L}})$ (left) and with respect to HG cluster orientation in (HG) segment 3 $\varphi^3 := \varphi(c^3_{j^{3L}})$ (right) for linear HG weighting and for power-law (HG) weighting with different values of p.

the values $\varphi^{1/3} = 0$, $\varphi^{1/3} = \pm\frac{\pi}{2}$. The effect is similar to the one shown in Fig. 18 in the context of LG orientations, only that in the latter the orientation is averaged over the multiple LG segments, while Fig. 21 shows the more primitive orientation at the level of the two HG segments. If the average values were used here, the biasing would no longer be visible, since tracks producing clusters biased with respect to orientation in one HG segment don't necessarily produce clusters biased in the same way in the other HG segment. This is directly related to the lack of correlation between the relative intersection points of the track in the different HG segments, due to smaller size of the spacing of the HG lattice.

4.2 Invariant Mass Variable

This section is concerned with studying the efficiency of the invariant mass analysis for rejecting π_0 decays for which for which the algorithm is able to separate the two showers. Each separated π_0 event has to be selected in the first place, according to the specifications at the beginning of Section 4. In addition, it is required that two showers are found and that each shower is matched to one true track of one of the two impacting photons. The showers are required to be within a match distance of $D_M < 0.135$cm for the higher energy photon and $D_M < 0.15$cm for the lower energy photon, when HG+LG clustering is being used and within a match distance of $D_M < 0.5$cm for the higher energy photon and $D_M < 0.55$cm for the lower energy photon, when LG clustering is being used. The study is conducted using semi-final LG clusters obtained from LG clustering in Sec. 4.2.1 and using final clusters obtained from LG+HG clustering in Sec. 4.2.2.

Two important variables, extracted from Monte Carlo truth information, are being used for plots shown in this section, as well as for certain plots in Sec. 4.3. These are the π_0 decay asymmetry $\alpha^{\gamma-\gamma}$ defined by equation (9) in Sec. 2, and the track-to-track distance $D^{\gamma-\gamma}$, which is the true distance between the intersection points of the two decay γ tracks with the z'_{FCP} plane. Due to kinematic aspects related to the decay of a relativistic π_0, the track-to-track distance is a monotonously increasing function of the decay asymmetry, for a constant π_0 energy, assuming that the trajectory of the latter is perpendicular to the detector. The distribution of asymmetry is uniform, while the distribution of track-to-track distance is peaked towards low distances, with a minimum distance for completely symmetric decays. This minimal distance decreases with increasing π_0 energy.

4.2.1 Low Granularity Clustering

This section focuses on understanding the effectiveness of the clustering (+shower-reconstruction) algorithm in separating showers induced by single π_0 decays, when using only low granularity information, together with the extent to which invariant mass analysis can be employed for rejecting such decays. The study is restricted to the $\text{samp}^{A}_{\pi_0}$ data sample described above, which is divided into 3 sub-samples, based on the true energy carried by the π_0, in order to understand how the efficiency of the invariant-mass procedure changes with increasing energy.

Figure 22 illustrates the efficiency of the (LG) clustering algorithm in correctly identifying two clusters in π_0 events, before any invariant mass analysis is performed. The plots show, with respect to the (true) energy asymmetry of the π_0 decay (left) and with respect to the (true) distance between the two impact position of the photons in the z'_{FCP} plane, the ratio of events for which two (semi-final LG) clusters are reconstructed within $D_M < 0.135$cm of the true impact positions (matched) – these events are further referred to as "separated". The ratios are defined with respect to the events selected according to the procedure described at the beginning of Section 4.2. Different colors in the two plots correspond to different (true) π_0 energy intervals. It can be noticed that π_0 cluster separation efficiency decreases with increasing energy, which is what one expects due to decreasing $\gamma-\gamma$ angles in the lab frame as the Lorentz boost gets stronger Furthermore, decays with large asymmetry $\alpha^{\gamma-\gamma}$ and distance $D^{\gamma-\gamma}$ are more often separated by the (LG) clustering algorithm.

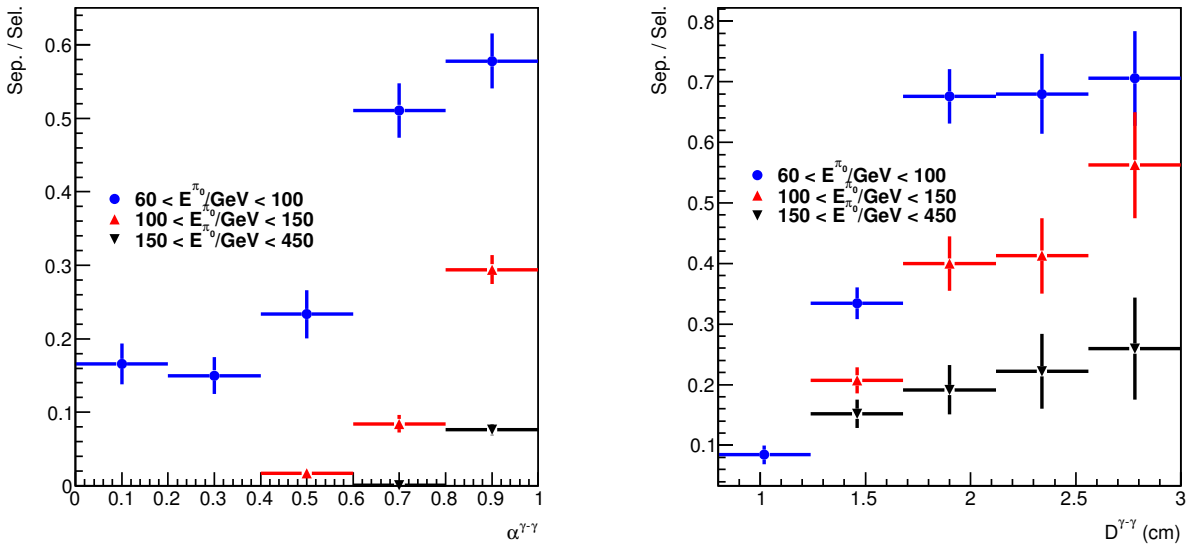

Figure 22: Ratio between separated and selected π_0 events, as a function of true decay asymmetry $\alpha^{\gamma-\gamma}$ (left) and as a function of true track-to-track distance $D^{\gamma-\gamma}$ in the $z'_{\rm FCP}$ plane (right), for three π_0 true energy (E^{π_0}) intervals, after employing LG clustering with logarithmic weighting ($w_0 = 4.3$).

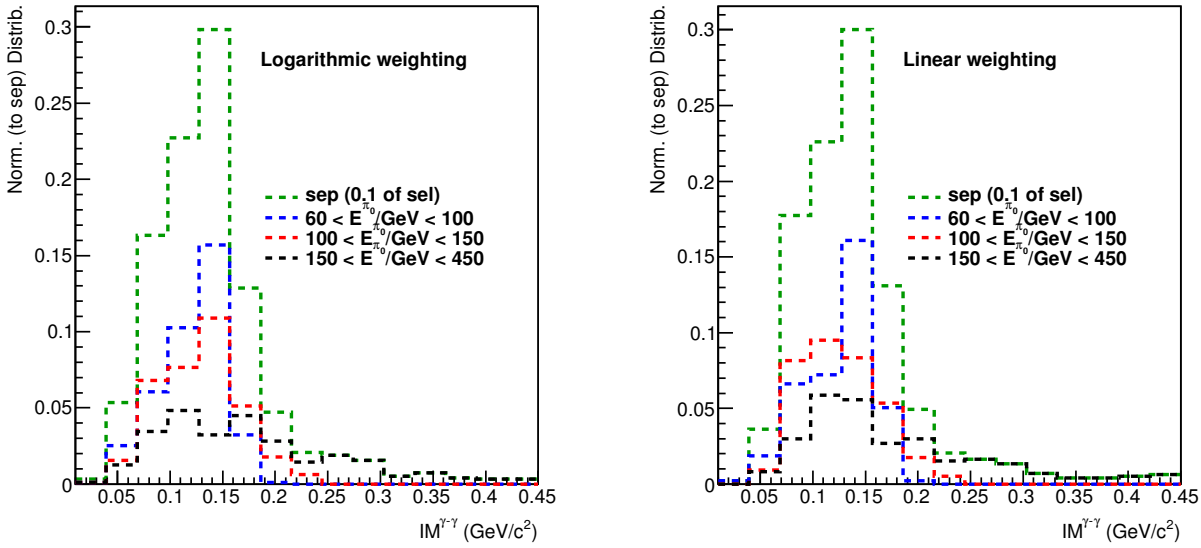

Figure 23: Distribution of separated $\pi-0$ events with respect to the invariant mass reconstructed from the two reconstructed (and matched) LG semi-final clusters, for 3 E^{π_0} intervals, as well as for the cumulative sample, using the results of LG clustering with logarithmic weighting ($w_0 = 4.3$) (left) and the results of LG clustering with linear weighting (right).

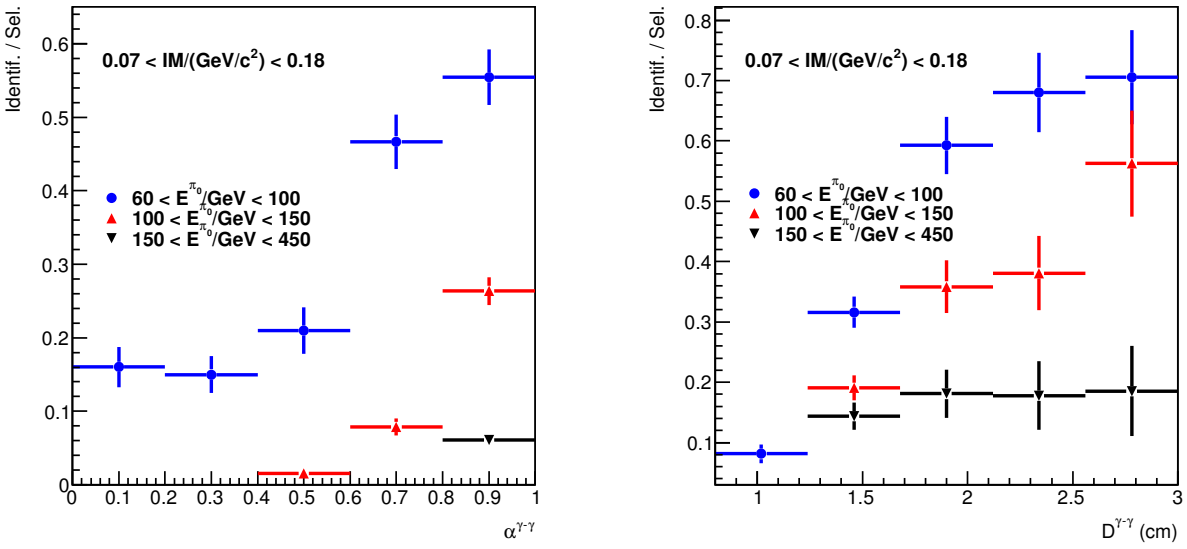

Figure 24: Ratio between identified and selected π_0 events, based on the invariant mass reconstructed from the two (reconstructed and matched) semi-final LG clusters, as a function of true decay asymmetry $\alpha^{\gamma-\gamma}$ (left) and as a function of true track-to-track distance $D^{\gamma-\gamma}$ in the z'_{FCP} plane (right), for three π_0 true energy (E^{π_0}) intervals, after employing LG clustering with logarithmic weighting ($w_0 = 4.3$).

Figure 23 shows, for those π_0 events which are correctly separated by the clustering algorithm (as explained in relation to Fig. 22), their distribution with respect to the invariant mass reconstructed from the respective pairs of clusters. Different curves (blue, red and black) indicate the same true energy data split as in Fig. 22. The green curve corresponds to the entire sample of separated events (thus the sum of the other 3 distributions). One can notice the coarse peaks around the neutral pion mass $m_{\pi_0} \approx 0.135\mathrm{GeV}/c^2$, for different π_0 energy intervals. The calculations on the left-hand side were performed using the results of the (LG) clustering with logarithmic weighting with $w_0 = 4.3$ (also valid for all the other plots in this subsection), while those on the right-hand side were performed using the results of the (LG) clustering with linear weighting. The strong similarity between the two plots shows that introducing logarithmic weighting does not significantly improve the invariant mass analysis. Since the invariant mass calculation makes use of position information and energy information from the paired clusters, while Section 4.1.1 showed that position reconstruction is improved by logarithmic weighting, this indicates that the main problems in reconstructing the invariant mass from semi-final LG clusters originate in assigning the cluster energy content.

Figure 24 shows the ratio of correctly identifying π_0 particles based on invariant mass analysis, with respect to true decay asymmetry (left) and true photon-photon distance in the z'_{FCP} transversal plane (right). The plots are very similar to those in Fig. 22, the ratios also being defined with respect to the set of all selected events, only that an additional constraint (cut) is imposed specifying that the invariant mass should be found within the $0.07(\mathrm{GeV}/c^s) < \mathrm{IM} < 0.18(\mathrm{GeV}/c^s)$. Note that the profiles are almost identical to those in Fig. 22, with a marginal lowering of certain bins (by definition, bins in Fig. 24 can only move downwards with respect to those in Fig. 22). This implies that most of the π_0s that are correctly separated by the (LG) clustering algorithm are also rejected via an invariant mass cut, the main problem being their separation in the first place. However, this is also due to the fact that the invariant mass cut has been chosen in such a way that most of the reconstructed π_0 peak is covered by the rejection window. For realistic pPb, PbPb collisions, the width of this window would be constrained by the amount of pairs of independent showers, possibly originating from prompt photons, falsely rejected under the π_0 hypothesis (combinatorial background for background rejection).

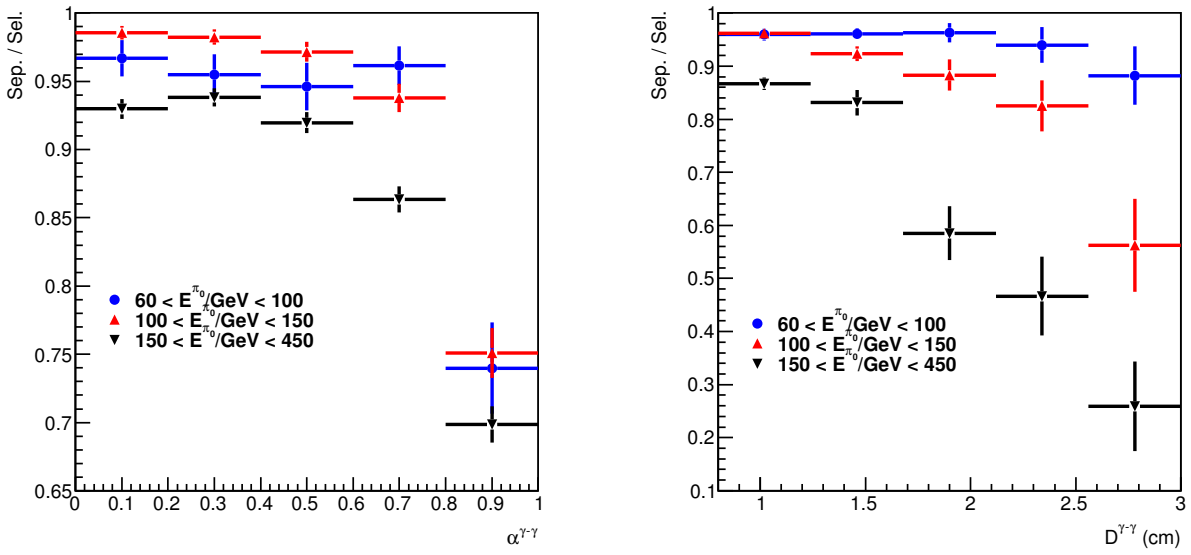

Figure 25: Ratio between separated and selected π_0 events, as a function of true decay asymmetry $\alpha_T^{\gamma-\gamma}$ (left) and as a function of true track-to-track distance $D_T^{\gamma-\gamma}$ in the $z'_{\rm FCP}$ plane (right), for three π_0 true energy (E^{π_0}) intervals, after employing HG+LG clustering with logarithmic LG weighting ($w_0 = 4.3$) and power-law HG weighting ($p = 2.2$).

Overall, the poor performance of the invariant-mass analysis after using the clustering algorithm only on the low-granularity segments is a first, strong indication that information from the high-granularity segments is needed for separating the clusters induced by the two photons originating from neutral pion decays. It is obvious that the need of this information increases with increasing pion energy.

4.2.2 High Granularity and Low Granularity Clustering

Having a structure very similar to Sec. 4.2.1, this section focuses on understanding the effectiveness of the clustering (+shower-reconstruction) algorithm in separating showers induced by single π_0 decays, when using information from the entire detector (HG and LG segments), together with the extent to which invariant mass analysis can be employed for rejecting such decays. The study is restricted to the samp$^A_{\pi_0}$ data sample described above, with the same true π_0 energy split as in Sec. 4.2.1 for similar reasons.

Figure 25 illustrates the efficiency of the (LG+HG) clustering algorithm in correctly identifying two clusters in π_0 events, before any invariant mass analysis is performed. The plots show, with respect to the (true) energy asymmetry of the π_0 decay (left) and with respect to the (true) distance between the two impact position of the photons in the z'_{FCP} plane, the ratio of events for which two (final) clusters are reconstructed within $D_M < 0.135$cm of the true impact positions (matched) – these events are further referred to as "separated". The ratios are defined with respect to the events selected according to the procedure described at the beginning of Section 4.2. Different colors in the two plots correspond to different (true) π_0 energy intervals. It is now clear that the separation efficiency of the clustering algorithm is much larger when also using the information from the HG segments, compared to the situation when only the LG segments are employed, illustrated by Fig. 22. Qualitatively similar to Fig. 22, π_0 cluster separation efficiency decreases with increasing energy, which is again due to decreasing $\gamma - \gamma$ angles in the lab frame as the Lorentz boost gets stronger. The effect is however significantly weaker than in the previous section, which becomes apparent when comparing the asymmetry plots on the left sides of Figures 25 and 22. On the other hand, for a specific π_0 energy interval, the cluster separation efficiency no

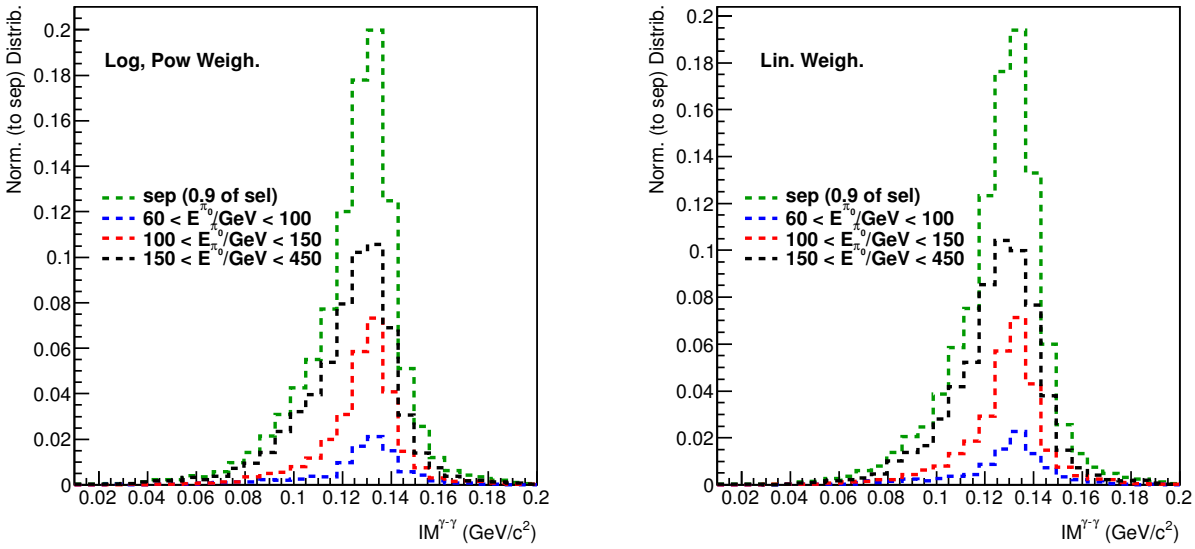

Figure 26: Distribution of separated π_0 events with respect to the invariant mass reconstructed from the two reconstructed (and matched) final clusters, for three E^{π_0} intervals, as well as for the cumulative sample, using the results of HG+LG clustering with logarithmic LG weighting ($w_0 = 4.3$) and power-law HG weighting ($p = 2.2$) (left) and the results of HG+LG clustering with linear HG and LG weighting (right).

longer increases with increasing asymmetry $\alpha^{\gamma-\gamma}$ and distance $D^{\gamma-\gamma}$, but decreases with these quantities.

The above observations imply that the HG+LG clustering algorithm is relatively less efficient in separating the clusters from asymmetric π_0 decays then it is in separating the clusters from symmetric π_0 decays, whereas the converse is true for the LG clustering algorithm, although the former performs better than the latter for all types of decays. Generally speaking, extremely asymmetric decays are challenging since the low-energy cluster is likely be lost in the tail of the high-energy cluster, through the rejection in step 3(b)iB (first bullet point) of the clustering algorithm explained in Section 3.3 [11]. On the other hand, extremely symmetric decays are difficult to separate due to the proximity of the two impact points throughout the sensitive layers of the calorimeter, which makes it likely that only one cluster is assigned to the combined energy deposition – in the clustering algorithm, this is manifested via seed rejection in the immediate vicinity of another seed in step 3(b)iB (second bullet point).

Figure 26 shows, for those π_0 events which are correctly separated by the clustering algorithm (as explained in relation to Fig. 25), their distribution with respect to the invariant mass reconstructed from the respective pairs of clusters. Different curves (blue, red and black) indicate the same true energy data split as in Fig. 25. The green curve corresponds to the entire sample of separated events (thus the sum of the other 3 distributions). One can notice that all curves are clearly peaked around the neutral pion mass $m_{\pi_0} \approx 0.135\text{GeV}/c^2$ 23. The calculations on the left-hand side were performed using the results of the (HG+LG) clustering with logarithmic LG weighting with $w_0 = 4.3$ and power-law HG weighting with $p = 2.2$ (also valid for all the other plots in this subsection), while those on the right-hand side were performed using the results of the (HG+LG) clustering with linear weighting. The strong similarity between the two plots shows once more that introducing different kinds of weightings does not significantly

[11] When considering how the invariant mass and shower shape analyses are combined here, there can be no π_0 decay that would be rejected when using the LG clustering algorithm but not when using the LG+HG clustering algorithm. Nonetheless, in realistic pp, pPb and PbPb collisions, the HG clustering would lead to significantly larger amounts of showers being reconstructed, which can increase the amount of shower pairs falling within the π_0 rejection window, possibly increasing the extent of signal rejection. Such effects have not been investigated in detail here

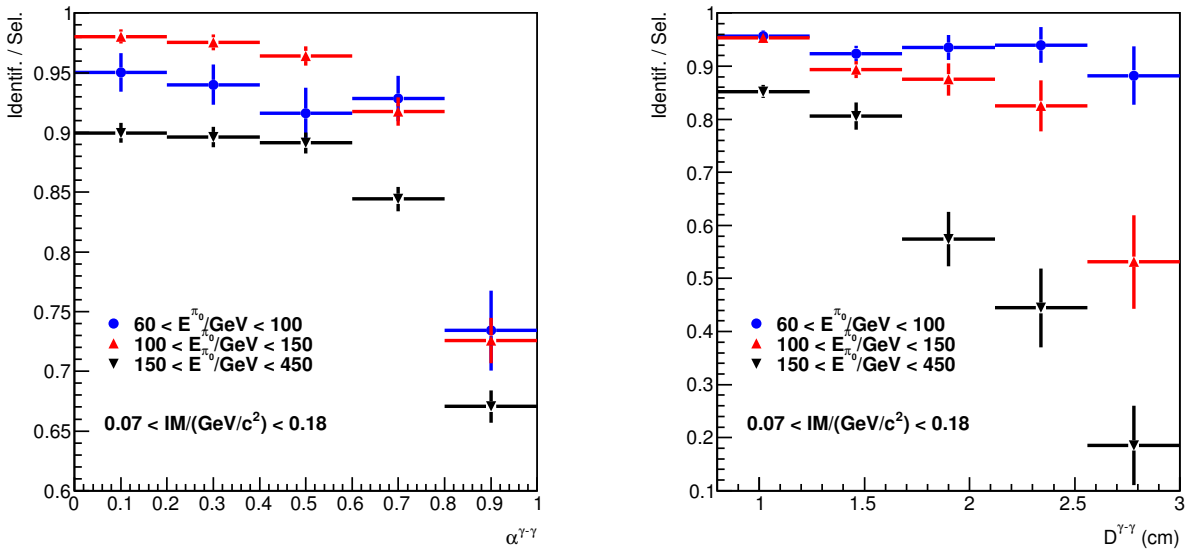

Figure 27: Ratio between identified and selected π_0 events, based on the invariant mass reconstructed from the two (reconstructed and matched) final HG clusters, as a function of true decay asymmetry $\alpha_T^{\gamma-\gamma}$ (left) and as a function of true track-to-track distance $D_T^{\gamma-\gamma}$ in the $z'_{\rm FCP}$ plane (right), for three π_0 true energy (E^{π_0}) intervals, after employing HG+LG clustering with logarithmic LG weighting ($w_0 = 4.3$) and power-law HG weighting ($p = 2.2$).

affect the invariant mass analysis.

Figure 27 shows the ratio of correctly identified π_0 particles based on invariant mass analysis, with respect to true decay asymmetry (left) and true photon-photon distance in the z'_{FCP} transversal plane (right). Once more, the plots are very similar to those in Fig. 25, the ratios also being defined with respect to the set of all selected events, only that an additional constraint (cut) is imposed specifying that the invariant mass should be found within the $0.07(\text{GeV}/c^s) < \text{IM} < 0.18(\text{GeV}/c^s)$. The profiles are almost identical to those in Fig. 25, with a marginal lowering of certain bins (by definition, bins in Fig. 27 can only move downwards with respect to those in Fig. 25). This implies that most of the π_0s that are correctly separated by the (LG) clustering algorithm are also rejected via an invariant mass cut, as was already apparent from Fig. 26.

4.3 Shower Shape Variables

This section is concerned with studying the efficiency of the shower-shape analysis for rejecting π_0 decays for which only one shower is obtained after applying the clustering and shower-shape reconstruction algorithms. These events are referred to as unseparated π_0 events. Each unseparated π_0 event has to be selected in the first place, according to the specifications at the beginning of Section 4. In addition, only one shower needs to be found and matched to the higher energy photon within $D_M < 0.135$cm when HG+LG clustering is being used and within $D_M < 0.5$cm when LG clustering is being used. The study is conducted using (mostly) semi-final LG clusters obtained from LG clustering in Sec. 4.3.1 and using (exclusively) final clusters obtained from LG+HG clustering in Sec. 4.3.2.

4.3.1 Low Granularity Variables

This section explores the γ vs π_0 discriminating power of different shower level low-granularity variables (also referred to as observables), with mathematical expressions given in Sec. 3.7. The study mostly focuses on comparisons between the samp_{γ}^{A} vs $\text{samp}_{\pi_0}^{A}$ data samples, although samples samp_{γ}^{B} vs $\text{samp}_{\pi_0}^{B}$ and samples $\text{samp}_{\pi_0}^{C}$ vs $\text{samp}_{\pi_0}^{C}$ are also looked at for comparisons with higher-energy intervals. Only π_0 events that are not separated by the algorithm (marked as "Non-Sep." in the plots) are being used. Data samples samp_{γ}^{A} and $\text{samp}_{\pi_0}^{A}$ are split in two sub-samples based on reconstructed shower energy, in order to observe the variation of discriminating power with energy, if any. The results of the LG clustering algorithm are mostly used, although the results of the HG+LG algorithm are also used for certain comparisons.

Figure 28 illustrates the power of discriminating between single γ (red) and single π_0 (blue) events introduced by the semi-major (left) and semi-minor (right) LG widths, computed at the level of semi-final clusters according to equation (42) in Section 3.7. Only non-separated π_0 events are being considered. The results of LG clustering with linear weighting are being used. The distributions are shown for two intervals of reconstructed shower energy, namely $50\text{GeV} < E^{sh} < 125\text{GeV}$ (top) and $125\text{GeV} < E^{sh} < 200\text{GeV}$ (bottom), since one might expect the widths to be somewhat energy-dependent. Note the strong discriminating power of the semi-major σ_1^L width, which was expected based on the consideration that transverse profiles of unseparated π_0 decays should have much more pronounced elliptical features compared to those induced by single γs. The strong discriminating power in σ_1^L is present for both energy intervals, although to a greater extent at lower energies, which is consistent with the fact that the π_0 decay angle decrease with increasing π_0 energy in the lab frame. By placing a background rejection cut on $\sigma_1^L > 0.745\text{cm}$ for the lower energy interval, one rejects approximately 4% of single γ's (signal) and 71% of the π_0's (background). In the higher energy interval, a cut on $\sigma_1^L > 0.73\text{cm}$ rejects approximately 5% of single γ's and 41% of the π_0's.

However, one would expect that the σ_2^L distribution should be similar for both γ and π_0 events. This expectation is based on the assumption that within in each segment the semi-minor width is measured along the direction approximately perpendicular to the line passing through the points where the particle tracks intersect the segment, while σ_1^L would be measured along this line. The right-hand plots of Fig. 28 show that in practice, with the given computation procedure, this expectation is not completely valid: Although σ_2^L is systematically lower than σ_1^L, for both π_0 and γ events take separately, σ_2^L is still on average somewhat larger for π_0 unseparated showers than for γ unseparated showers.

In other words, the σ_2^L variable still carries some unexpected discriminating power. Although not shown explicitly here, detailed studies conducted for this project showed that, when using the results of LG clustering with linear weighting, the sum $\sigma_1^L + \sigma_2^L$ effectively combines the discriminating power of both variables, behavior which is unexpected and counter-intuitive. A study conducted in parallel and described in [29] suggests that this is actually an effect of the low-granularity lattice combined with the specific calculation procedure used for obtaining σ_1^L and σ_2^L (which assumes Gaussian distribution of energy within the transverse sections of shower), while employing linear weighting for the amplitudes of the digits. As explained in Section 4.1.2 in relation to Fig. 18, the low-granularity lattice biases the orientation of the reconstructed clusters towards the horizontal and vertical axes, thus introducing a difference between the actual and the reconstructed shower transverse orientation. This invalidates to some extent the assumption that σ_1^L and σ_2^L are measured along and perpendicular to the line passing through the two impact points of the decay photons. Based on these considerations, one would expect that replacing the linear weighting procedure with the logarithmic weighting one would remove this lattice artifact, while transferring the discriminating power of the semi-minor width to the semi-major one.

This expectation is confirmed by figure 29, which shows the γ and π_0 distributions with respect to the same semi-minor (left) and semi-major (right) semi-major width of semi-final LG clusters, for low (top) and high (bottom) cluster energy intervals, this time using the results of logarithmic LG weighting with the parameter considered optimal in Section 4.1.1. When looking back at the corresponding distributions

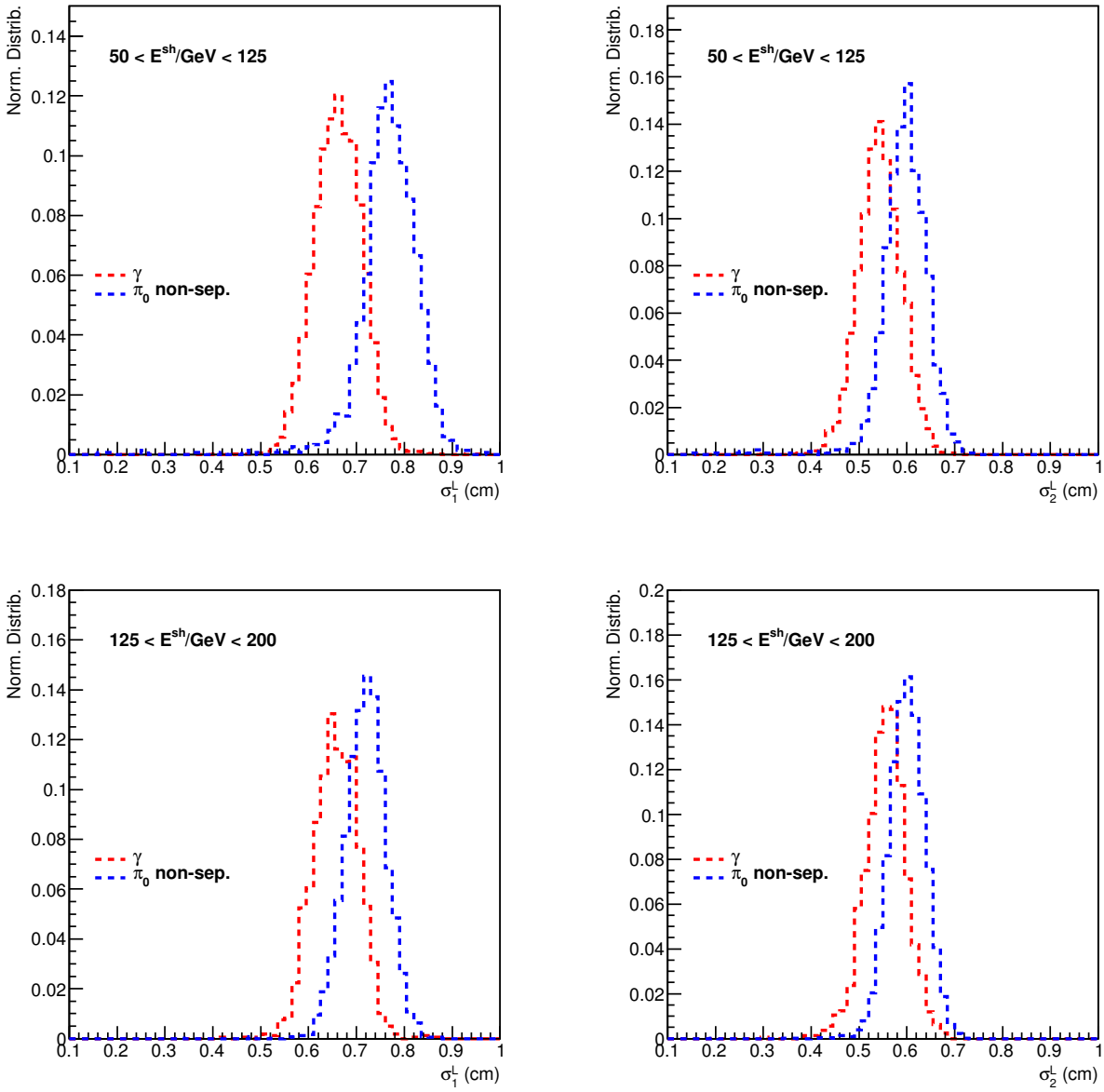

Figure 28: Distribution of single γ events (red) vs the distribution of unseparated single π_0 events (blue), for two shower energy intervals $E^{sh} := E(c_j^L)$ (top vs bottom) with respect to the LG semi-major σ_1^L (left) and LG semi-minor σ_2^L (right) reconstructed cluster widths, for semi-final LG clusters, after linear LG weighting.

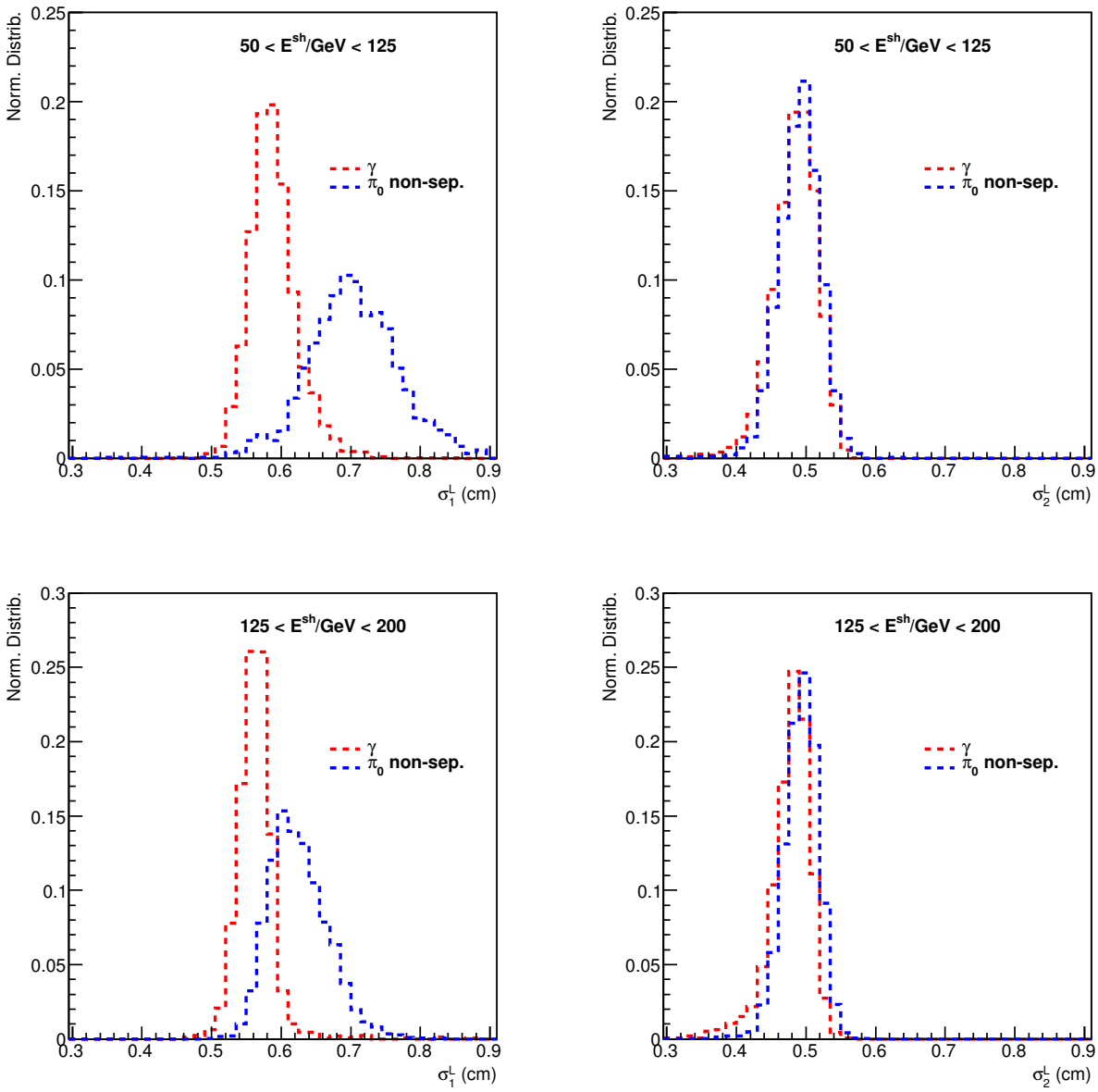

Figure 29: Distribution of single γ events (red) vs the distribution of unseparated single π_0 events (blue), for two shower energy intervals $E^{sh} := E(c_j^L)$ (top vs bottom) with respect to the LG semi-major σ_1^L (left) and LG semi-minor σ_2^L (right) reconstructed cluster widths, for semi-final LG clusters, after logarithmic LG weighting with $w_0 = 4.3$.

in figure 28, it becomes clear that the discriminating power of σ_1^L increases while that of σ_2^L is completely removed, thus validating the expectation that the semi-minor, transverse width of γ and unseparated π_0 clusters should be the same. Using logarithmic weighting removes the cluster orientation bias introduced by the LG lattice and allows for the intrinsic elliptical features of π_0-induced clusters to be better accounted for. Imposing a background rejection cut on $\sigma_1^L > 0.655$cm for the lower energy interval, leads to approximately 4% single γ's rejection and 79% π_0's rejection. In the higher energy interval, a cut on $\sigma_1^L > 0.61$cm rejects approximately 3% of single γ's and 60% of the π_0's. By comparing these numbers to those reported in relation to Fig. 28, it becomes clear that logarithmic weighting significantly improves the discriminating power of the semi-major width.

The observation above points out that a cut on the σ_1^L variable would be very effective for separating signal from background when the results of LG clustering are used. At this point, one may wonder whether such a cut is of any use when information from the HG segments is also taken into account during the clustering process, since most of the π_0 decays that are non-separated by LG clustering are actually separated by HG+LG clustering and may very likely be rejected via invariant mass analysis. In order to understand this, figure 30 shows the distribution with respect to the same semi-major and semi-minor widths and for the same energy intervals as in Fig. 29, only that now final clusters produced by via HG+LG clustering are being used, using logarithmic LG weighting and power-law HG weighting. On the left-hand side, one can notice the irregularities appearing in the non-separated π_0 distribution due to the much poorer statistics. Still, it is clear that the σ_1^L variable retains some of its discriminating power. In addition, a cut on σ_1^L would also reject some of the (few) separated π_0 decays which fall out of the π_0 invariant mass peak and are not rejected by the invariant mass analysis, since final clusters mapped to the same semi-final LG cluster directly take their σ_1^L from the latter – see Section 3.7. In parallel, the photon and neutral pion distributions with respect to the semi-minor width σ_2^L retain their strong similarity, recovered when introducing logarithmic weighting above.

In order to study the discriminating power of more advanced LG variables for the rest of this subsection, the LG clustering results are used again. This is done firstly for enhanced statistics and secondly for providing a good understanding of how the detector can perform without its HG segments.

Figure 31 shows the distribution of semi-final LG clusters induced by γ and unseparated π_0 events with respect to the transverse LG shower alignment variable r_φ^L, computed according to equation (46), for the same two intervals of reconstructed shower energy as before. This observable was conceived as an attempt of taking into account information about segment-level cluster orientation, which is completely ignored when computing the σ_1^L and σ_2^L variables according to (42). One would expect that segment-level clusters contributing to a photon induced shower (in this case a semi-final LG cluster) show on average a lower extent of overall transverse alignment than those contributing to a neutral pion induced shower. The two decay photons originating from the latter already provide a well defined direction in the transverse plane, thus breaking the rotational symmetry that is characteristic for a photon induced shower. The left hand-side of Fig. 31 illustrates this effect, as non-separated π_0 events are obviously peaked towards $r_\phi^L = 1$, which corresponds to high alignment, in contrast to γ events which almost uniformly distributed within $r_\varphi^L \in [0, 1]$. However, the discriminating power of r_φ^L is strongly diminished by a (background rejection) cut on the semi-major width $\sigma_1^L > 0.6$cm, as can be noticed on the right-hand side of Fig. 31, where the events surviving the cut (satisfying $\sigma_1^L < 0.6$cm) are shown for both distributions. Rather than investigating combined cuts in the σ_1^L vs r_φ^L space, other shower variables are investigated, which combine the width and orientation information within a single formula.

The more sophisticated variables mentioned above are referred to here as the vectorial average semi-major and semi-minor LG widths $\hat{\sigma}_1^L$ and $\hat{\sigma}_2^L$, computed according to Eq. (48) in Sec. 3.7. In effect, these variables should provide an approximation of the reconstructed semi-major and semi-minor widths, if the energy deposited in all LG segments and contained by the reconstructed shower was first summed up before conducting any Gaussian statistics, but after taking into account transverse displacement when going from one segment to another. Figure 32 shows the distributions of γ and π_0 events with respect to these two variables, for the same to bins of shower energy. It is clear that the vectorial semi-minor width (right) shows the same distribution for photons and pions, while a stronger discriminating power in

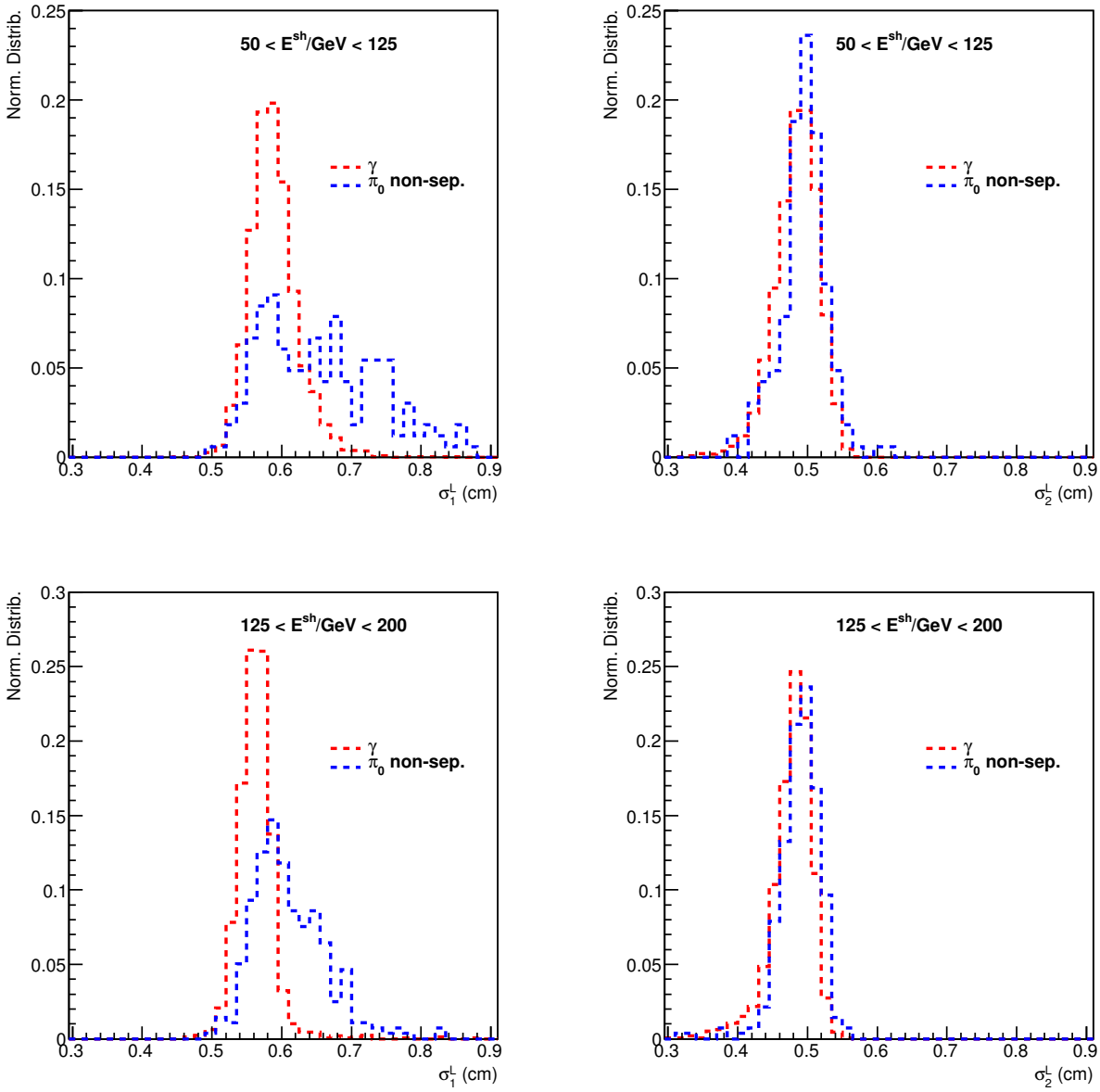

Figure 30: Distribution of single γ events (red) vs the distribution of unseparated single π_0 events (blue), for two shower energy intervals $E^{sh} := E(c_j^F)$ (top vs bottom) with respect to the LG semi-major σ_1^L (left) and LG semi-minor σ_2^L (right) reconstructed cluster widths, for final clusters, after logarithmic LG weighting with $w_0 = 4.3$ and power-law HG weighting with $p = 2.2$.

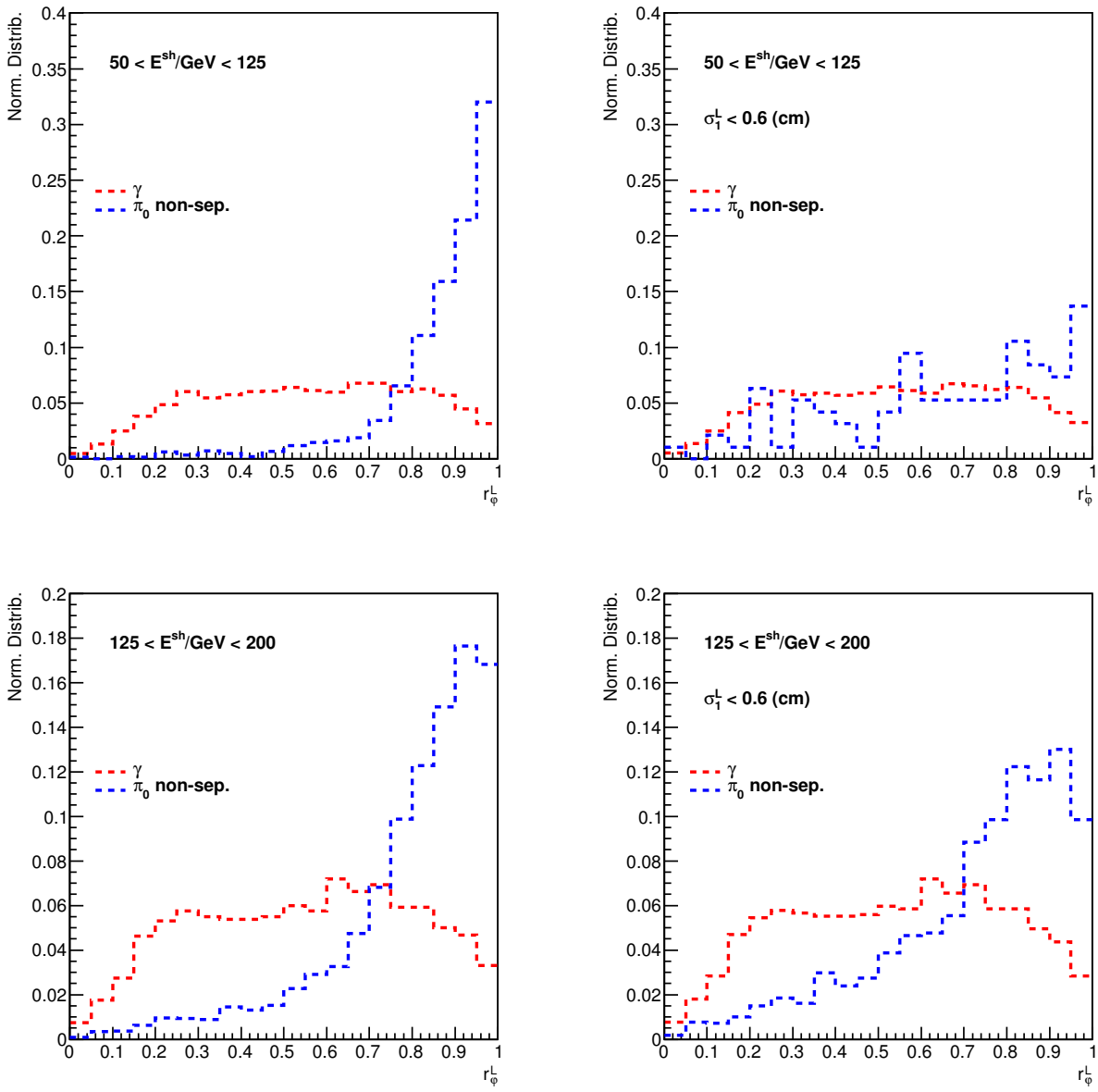

Figure 31: Distribution of single γ events (red) vs the distribution of unseparated single π_0 events (blue), for two shower energy intervals $E^{sh} := E(c_j^F)$ (top vs bottom) with respect to the LG reconstructed transversal shower alignment r_φ^L, with no additional cut (left) and after an additional (background-rejection) cut on the semi-major width σ_1^L (right), for semi-final LG clusters, after logarithmic LG weighting with $w_0 = 4.3$.

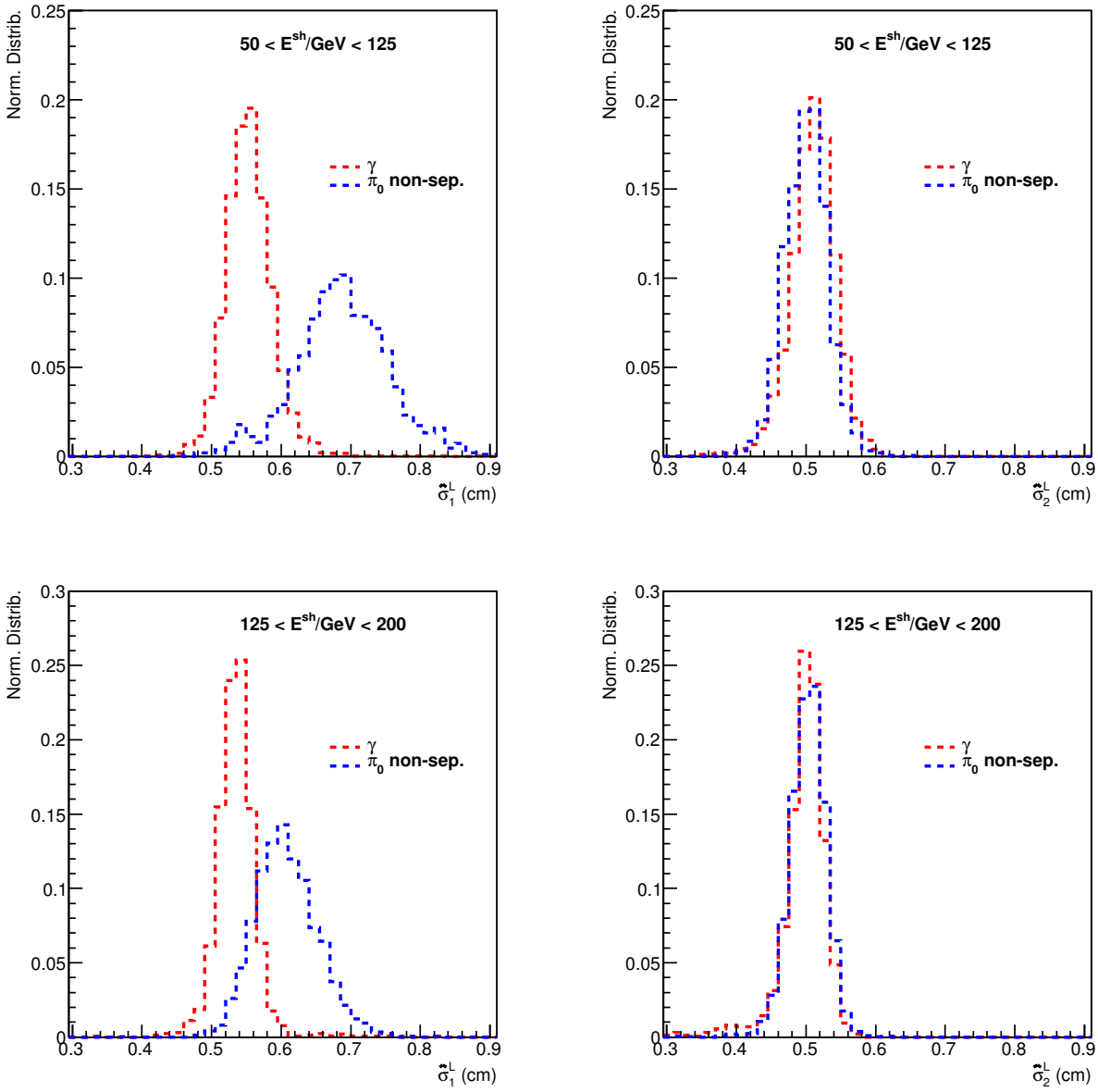

Figure 32: Distribution of single γ events (red) vs the distribution of unseparated single π_0 events (blue), for two shower energy intervals $E^{sh} := E(c_j^F)$ (top vs bottom), with respect to the semi-major $\hat{\sigma}_1^L(c_j^F)$ (left) and semi-minor $\hat{\sigma}_2^L(c_j^H)$ (right) reconstructed vectorial cluster widths, for semi-final LG clusters, after logarithmic LG weighting with $w_0 = 4.3$.

the semi-major width (left) becomes apparent, when comparing to the same plots in Fig. 29. Imposing a background rejection cut on $\hat{\sigma}_1^L > 0.61$cm for the lower energy interval, leads to approximately 5% single γ's rejection and 89% π_0's rejection. In the higher energy interval, a cut on $\hat{\sigma}_1^L > 0.58$cm rejects approximately 4% of single γ's and 72% of the π_0's. When looking back at the percentages reported in relation to Fig. 29, it becomes clear that making use of segment-level angular information for computing the semi-major LG vectorial shower width $\hat{\sigma}_1^L$ further improves discriminating power, at least when logarithmic weighting is being employed.

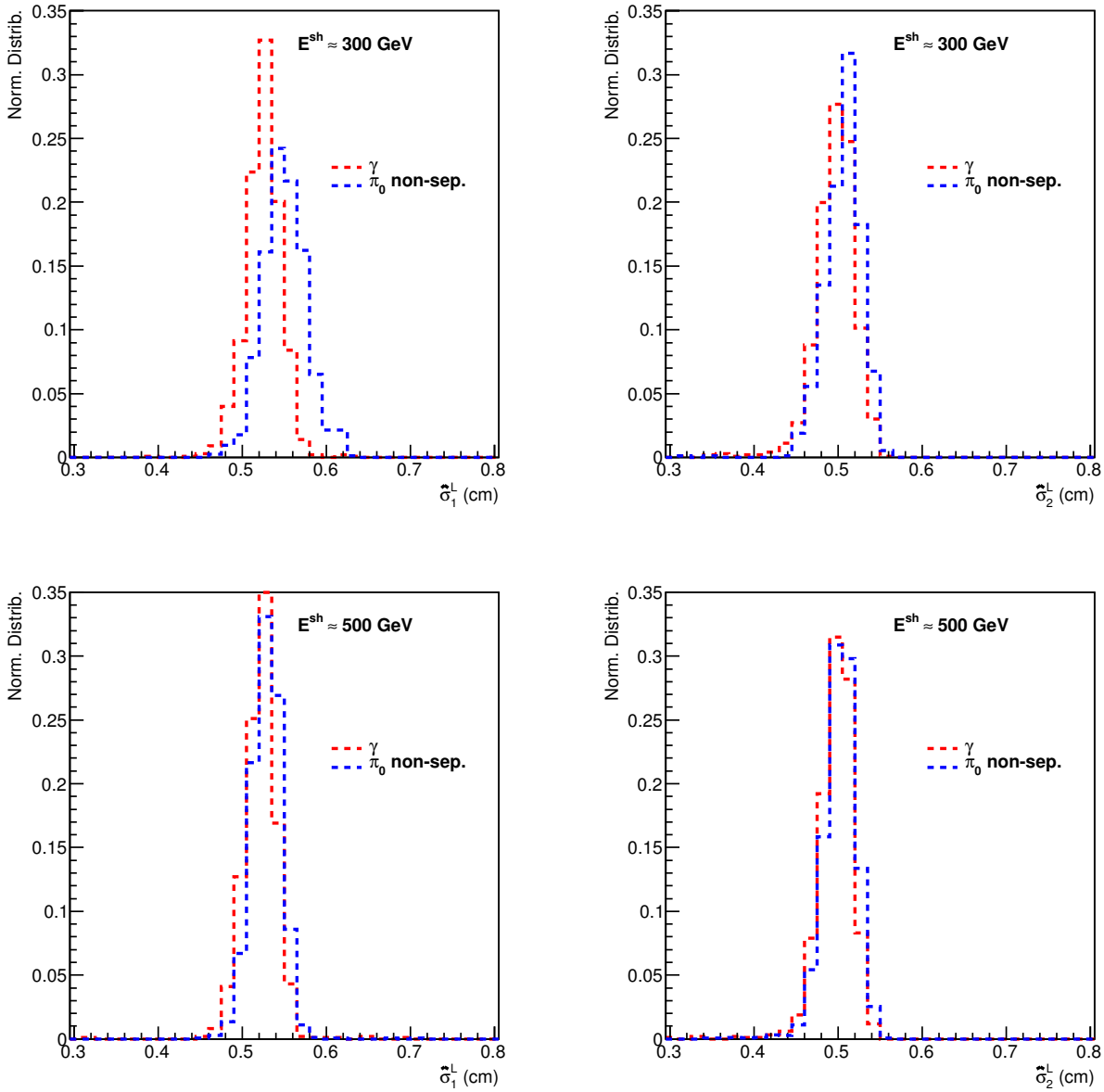

Figure 33: Distribution of single γ events (red) vs the distribution of unseparated single π_0 events (blue), for two shower (higher) energy intervals $E^{sh} := E(c_j^F)$ (top vs bottom) with respect to the semi-major $\hat{\sigma}_1^L$ (left) and semi-minor $\hat{\sigma}_2^L$ (right) reconstructed vectorial cluster widths, for final clusters, after logarithmic LG weighting with $w_0 = 4.3$ and power-law HG weighting with $p = 2.2$.

It is interesting to check the behavior of $\hat{\sigma}_1^L$ and $\hat{\sigma}_2^L$ for showers of higher energies. Figure 33 shows

the distributions with respect to these 2 variables, on the left and right side respectively, for photon and neutral pion events within two higher narrow energy windows, namely 300 GeV (top) and 500 GeV (bottom). Here, the clustering algorithm makes use again of information from the entire detector (both HG and LG segments), since non-separated π_0 are more abundant due to the decreasing efficiency of the clustering algorithm with higher energy. As expected, the overlap between the distributions with respect to $\hat{\sigma}_2^L$ is still present. On the other hand, the discriminating power with respect to $\hat{\sigma}_1^L$ is poor at 300 GeV and already absent at 500 GeV, when using HG+LG clustering. It is also worth noticing that imposing one of the cuts suggested above on $\hat{\sigma}_1^L$ intended to eliminate the non-separated π_0 background at low energies would reject less signal or no signal at all at higher energies. Although not explicitly shown here, the $\hat{\sigma}_1^L$ and $\hat{\sigma}_2^L$ distributions at low energies with LG+HG clustering are similar to those in Fig. 30, with discriminating power still present in the former but with poor statistics in the π_0 sample.

This section shows that advanced shower-shape analysis for rejecting π_0 showers which are not separated by the clustering algorithm is possible and desirable, based on information from LG segments, both in the LG clustering and LG+HG clustering situations. It is most convenient to apply a single cut on the $\hat{\sigma}_1^L$ variable explained above after employing logarithmic weighting in the LG segments. In addition, for discriminating non-separated showers of higher energies, examining similar shower-shape variables based on information from the HG segments is required, which is being done in the following section.

4.3.2 High Granularity Variables

This section explores the γ vs π_0 discriminating power of different shower level high-granularity variables (also referred to as observables), with mathematical expressions given in Sec. 3.7. The study mostly focuses on comparisons between the samp_γ^B vs $\text{samp}_{\pi_0}^B$ data samples and between the samp_γ^C vs $\text{samp}_{\pi_0}^C$ data samples since HG variables are of most relevance at high energies, where π_0 decays are less likely separated by the clustering algorithm. Samples samp_γ^A and $\text{samp}_{\pi_0}^A$ are also looked at for comparisons with lower-energy intervals. Only π_0 events that are not separated by the algorithm (marked as "Non-Sep." in the plots) are being used. Data samples samp_γ^A and samp_γ^A are split in two sub-samples based on reconstructed shower energy, in order to observe the variation of discriminating power with energy, if any. The studies are only performed on the results of the LG+HG clustering algorithm.

Figure 34 shows the distribution of single γ events and that of single, unseparated π_0 events with respect to the semi-major (σ_1^H) and semi-minor (σ_2^H) HG widths, for one narrow window of (high) energy. The calculations make use of the results of the HG+LG clustering algorithm, with logarithmic weighting of digits in the LG segments and linear weighting of digits in the HG segments. Although the average σ_2^H is significantly lower on average than σ_1^H for both distributions, it is obvious that there is no γ vs π_0 discriminating power present in either of the two variables. At this point, this result seems rather unexpected, since the left-hand side of Fig. 7 in Sec. 2 shows that the minimal separation (for minimal asymmetry) between the incident positions of the two photons at 3.5m from the interaction point is 2mm, for an energy of 500GeV. Given that the HG segments have a lattice size of 1mm, one would thus expect that the HG semi-major width σ_1^H would allow some separation between the two distributions. One hypothesis possibly explaining this lack of separation is that in the HG segments the Gaussian statistics is sensible to statistical fluctuations at the periphery of the cluster, while the essential information concerning transverse shape might be contained closer to the core of the cluster.

Power-law weighting in the HG segments has been employed here in order to give more emphasis on the relatively high-amplitude digits, likely found close to the center of the cluster, than to relatively low amplitude digits, likely found close to the periphery of the cluster. The formula is given in Eq. (21). Figure 35 shows the effect of power-law weighting on unseparated π_0 and γ distributions within two narrow windows of energy. Fig. 34 should be directly compared with the lower part of Fig. 35. At 500 GeV, it becomes clear that power-law weighting restores the expected discriminating power in the σ_1^H variable, while leaving the distributions with respect to σ_2^H identical. Note that all values are systematically lower compared to Fig. 34, which is the direct consequence of the fact that, while computing the respective

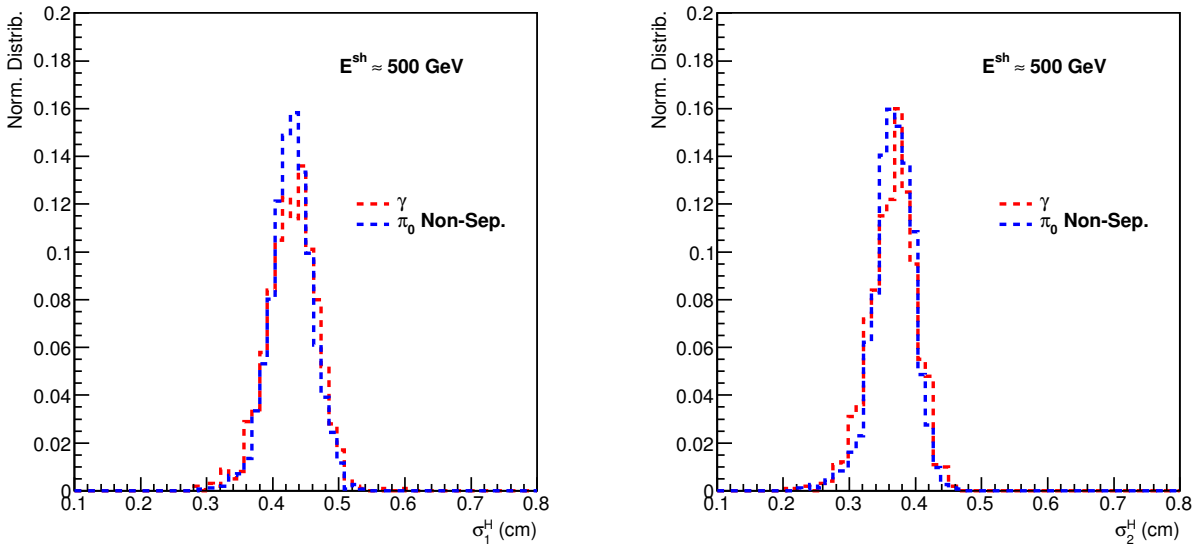

Figure 34: Distribution of single γ events (red) vs the distribution of unseparated single π_0 events (blue), for one narrow window of (high) cluster energy $E^{sh} := E(c_j^F)$ with respect to the HG semi-major σ_1^H (left) and HG semi-minor σ_2^H (right) reconstructed cluster widths, for final clusters, after logarithmic LG weighting with $w_0 = 4.3$ and linear HG weighting.

observables, less emphasis is placed on lower-amplitude digits, which are located further from cluster cores. The semi-major HG width also separates well 300 GeV π_0 showers from 300 GeV γ showers, as can be seen in the upper-left plot. At this energy, a double-peak structure becomes apparent in the π_0 distribution with respect to σ_1^H. This effect is explained towards the end of this subsection. In parallel, some difference between the two distributions with respect to σ_2^H become visible at 300 GeV, which reminds of the similar, unexpected discriminating power in σ_2^L explained in Section 4.3.1, removed by logarithmic weighting. The effect is not is not explored further in the current study. According to the distributions shown in Fig. 35, imposing a background rejection cut on $\sigma_1^H > 0.123$cm for the lower energy interval, leads to approximately 3.5% single γ's rejection and 68.0% π_0's rejection. In the higher energy interval, a cut on $\sigma_1^H > 0.127$cm rejects approximately 4.3% of single γ's and 77.6% of the π_0's.

Figure 36 shows the distributions with respect to the transverse alignment variable r_φ^L, for the same narrow windows of energy used in Fig. 35. This variable is computed according to Eq. (45), indicating high alignment between the segment-level HG clusters for $r_\varphi^L = 1$ and low alignment for $r_\varphi^L = 0$. One would expect that final clusters induced by non-separated π_0 particles would be more aligned than those induced by γ particles. This effect is visible in the plots on the left-hand side of Fig. 36, however with small discriminating power. For the events surviving a cut on σ_1^H, the discriminating power is even smaller, as shown in the plots on the right-hand side of Fig. 36. Making use of the same approach as in Sec. 4.3.1, the distributions with respect to the vectorial semi-major and semi-minor widths are studied next, instead of studying more sophisticated, combined cuts on the standard semi-major width σ_1^H and the alignment r_φ^H.

These distributions are shown in Fig. 37, which has the same format as Fig. 35 above. Note the strong similarity between the distributions in the two figures. According to the distributions shown in Fig. 37, imposing a background rejection cut on $\hat{\sigma}_1^H > 0.120$cm for the lower energy interval, leads to approximately 3.0% single γ's rejection and 68.5% π_0's rejection. In the higher energy interval, a cut on $\hat{\sigma}_1^H > 0.123$cm rejects approximately 5.7% of single γ's and 82.4% of the π_0's. Comparing these numbers with those reported in relation to Fig. 37 shows that the vectorial semi-major HG width does

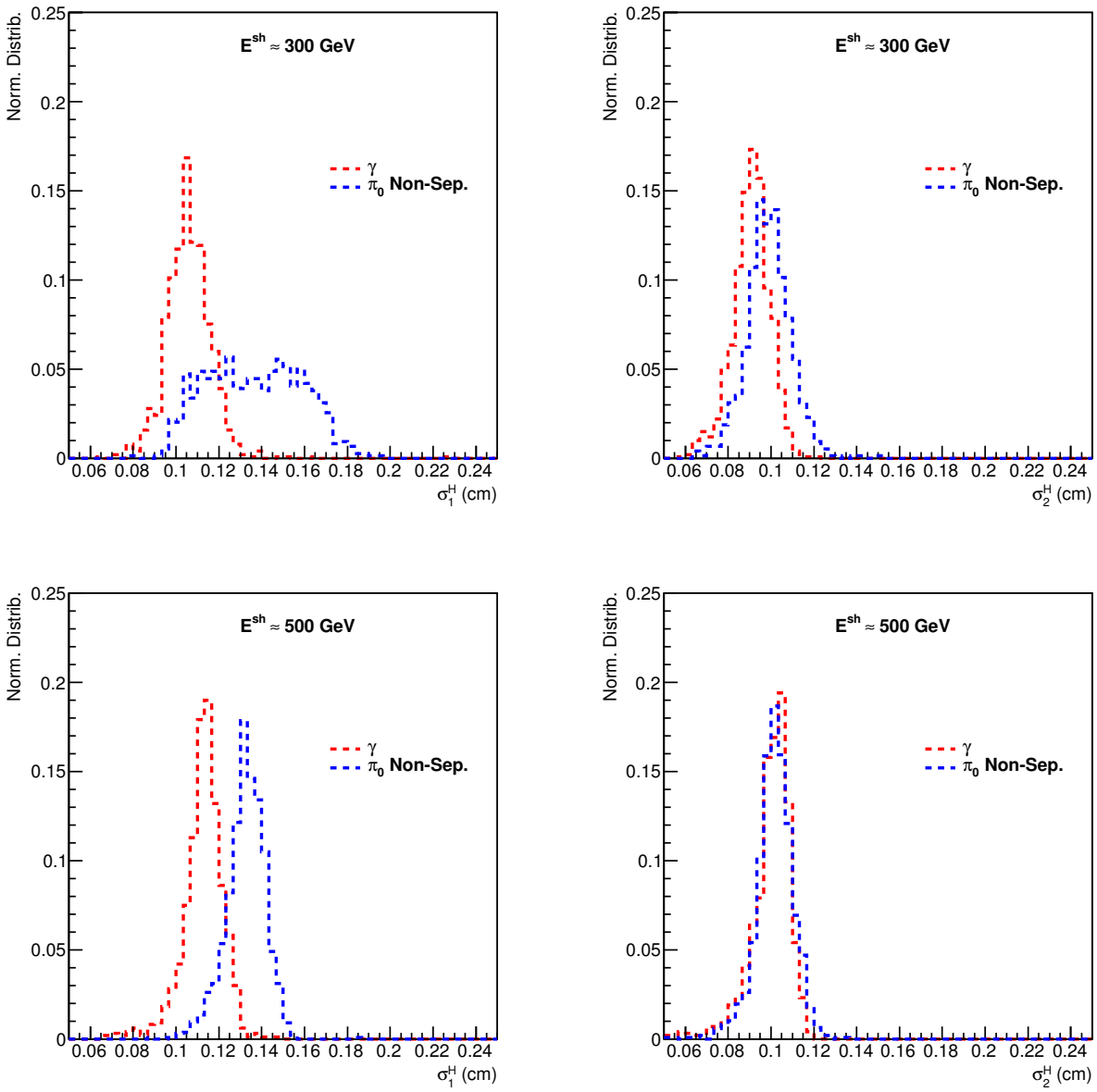

Figure 35: Distribution of single γ events (red) vs the distribution of unseparated single π_0 events (blue), for two narrow windows of cluster energy $E^{sh} := E(c_j^F)$ (top vs bottom) with respect to the HG semi-major σ_1^H (left) and HG semi-minor σ_2^H (right) reconstructed cluster widths, for final clusters, after logarithmic LG weighting with $w_0 = 4.3$ and power-law HG weighting with $p = 2.2$.

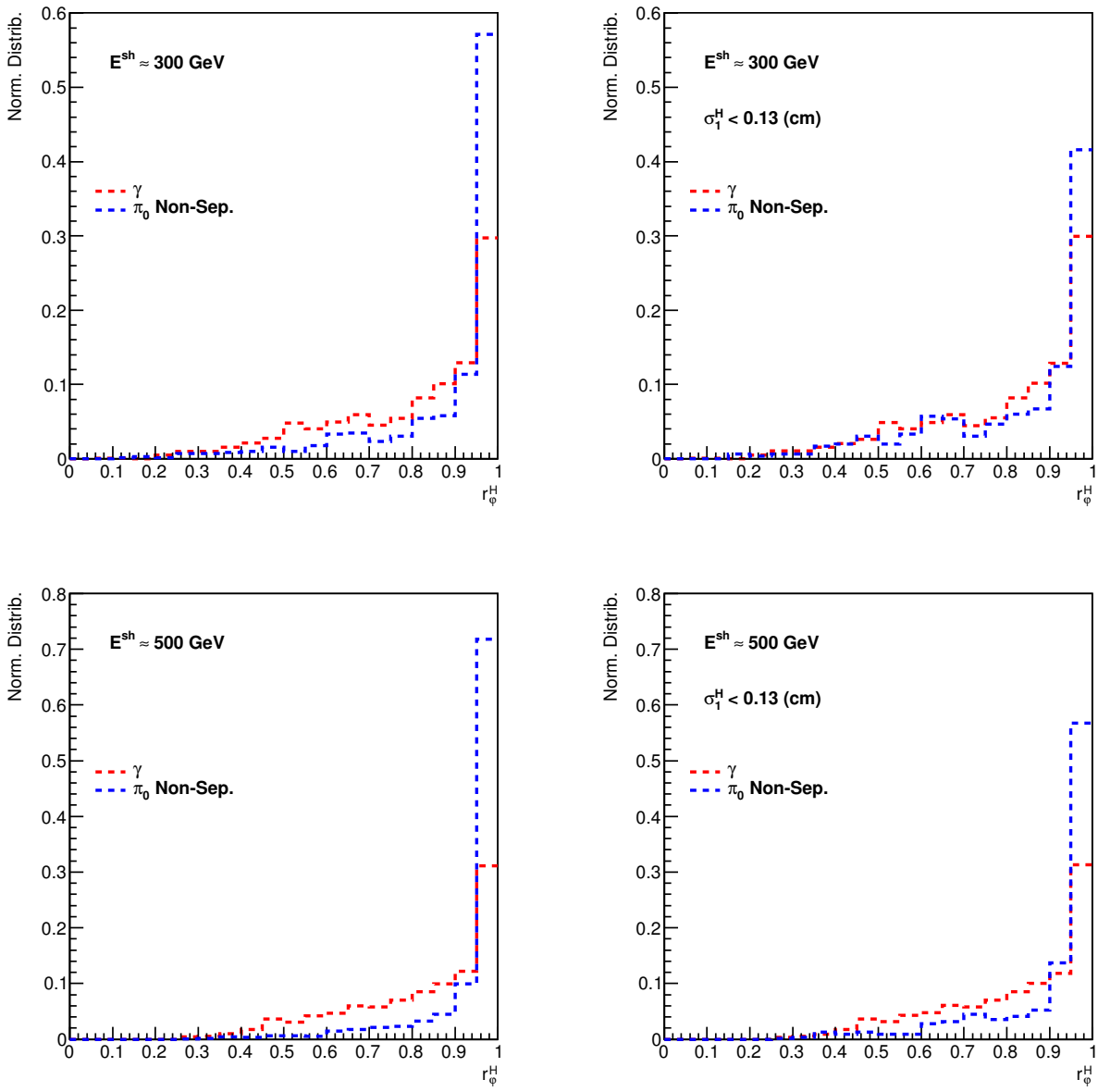

Figure 36: Distribution of single γ events (red) vs the distribution of unseparated single π_0 events (blue), for two narrow windows of shower energy $E^{sh} := E(c_j^F)$ (top vs bottom), with respect to the HG reconstructed transversal shower alignment r_φ^H, without any further additional constraint (left) and after a (background-rejection) cut on the semi-major HG width σ_1^H (right), for final clusters, after logarithmic LG weighting with $w_0 = 4.3$ and power-law HG weighting with $p = 2.2$.

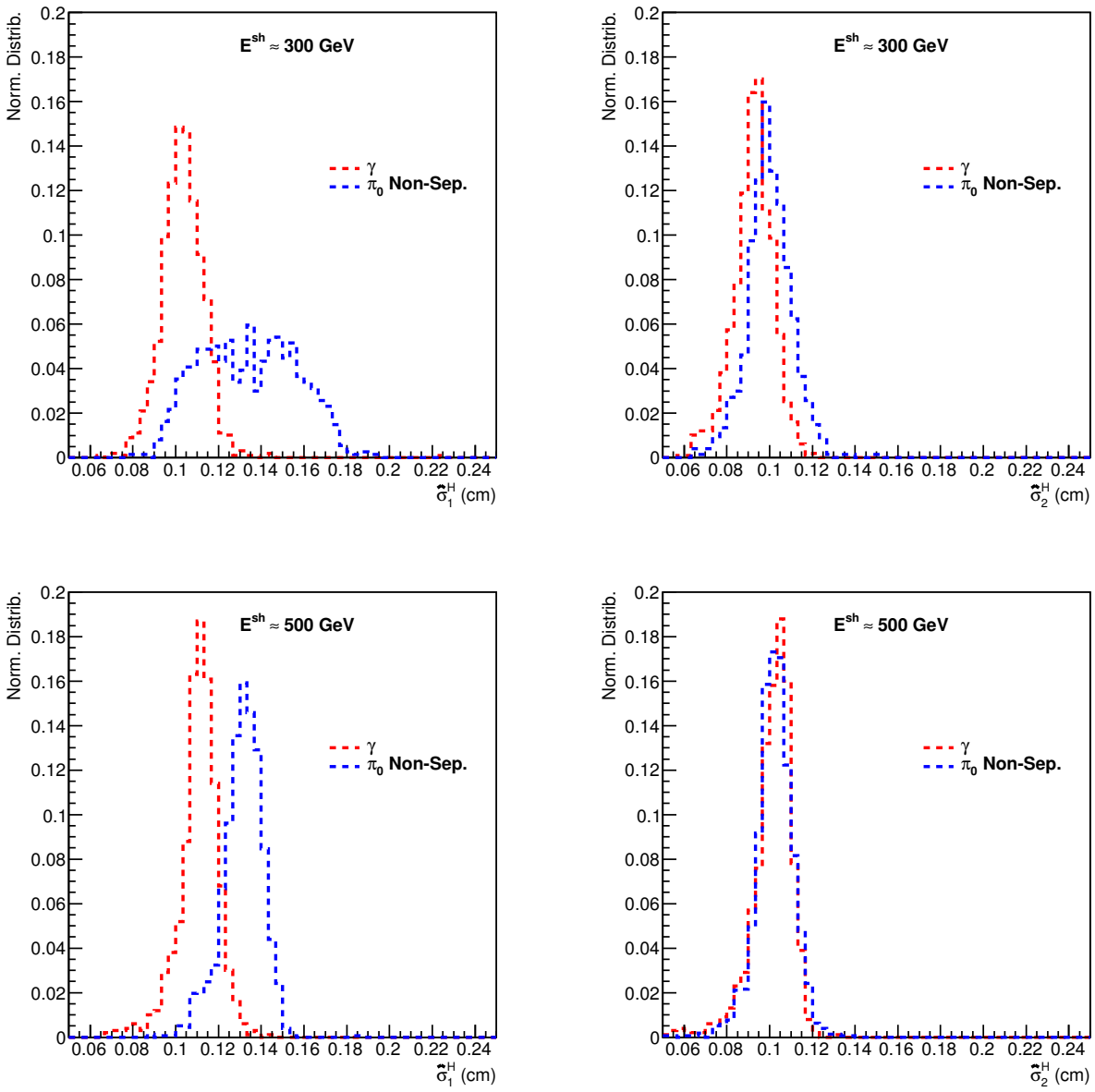

Figure 37: Distribution of single γ events (red) vs the distribution of unseparated single π_0 events (blue), for two narrow windows of shower energy $E^{sh} := E(c_j^F)$ (top vs bottom) with respect to the semi-major $\hat{\sigma}_1^H$ (left) and semi-minor $\hat{\sigma}_2^L$ (right) reconstructed vectorial HG widths, for final clusters, after logarithmic LG weighting with $w_0 = 4.3$ and power-law HG weighting with $p = 2.2$.

provide slightly more discriminating power than the standard one. The improvement brought by angular information when computing the detector-level widths is smaller for the HG widths than for the LG widths. In the relevant energy ranges, the HG transverse alignment variable showed less discriminating power than the LG transverse alignment variable, which is visible when comparing Fig. 37 to 31. This can be related to the smaller number of HG segments in use and to their relatively small shower-depth coverage of two radiation lengths.

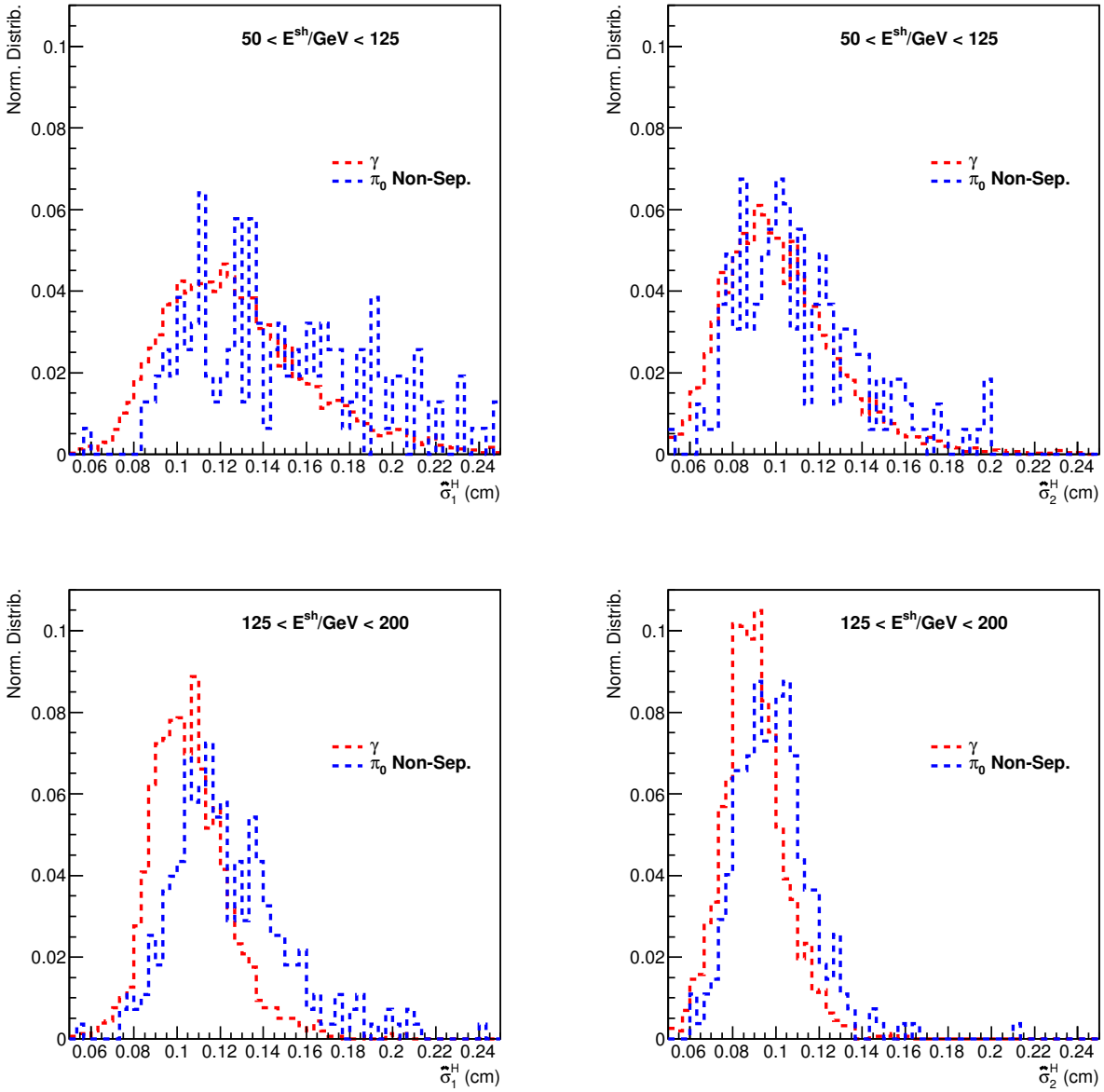

Figure 38: Distribution of single γ events (red) vs the distribution of unseparated single π_0 events (blue), for two cluster (lower) shower energy intervals $E^{sh} := E(c_j^F)$ (top vs bottom), with respect to the semi-major $\hat{\sigma}_1^H$ (left) and semi-minor $\hat{\sigma}_2^H$ (right) reconstructed vectorial cluster HG widths, for final clusters, after logarithmic LG weighting with $w_0 = 4.3$ and power-law HG weighting with $p = 2.2$.

Figure 38 shows the behavior of the γ and unseparated π_0 distributions, with respect to the vectorial widths σ_1^H and σ_2^H, when going back to lower energies, while still using the full LG+HG clustering

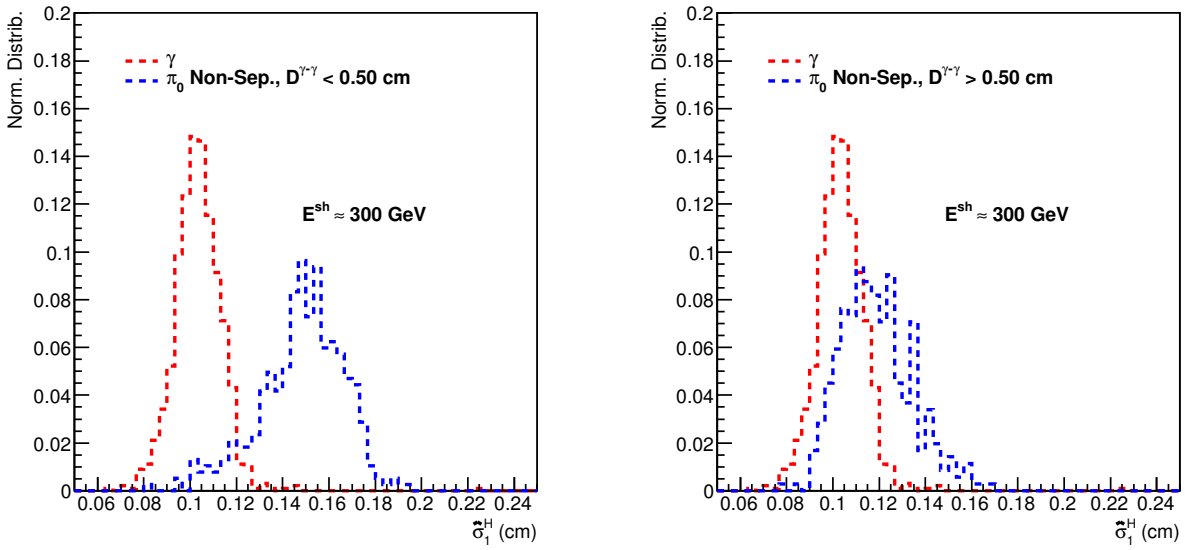

Figure 39: Distribution of single γ events (red) vs the distribution of unseparated single π_0 events (blue), for one narrow window of shower energy $E^{sh} := E(c_j^F)$, with respect to the semi-major $\hat{\sigma}_1^H$ reconstructed vectorial cluster width, events with small (left) vs large (right) track-to-track distance in the z'_{FCP} plane, for final clusters, after logarithmic LG weighting with $w_0 = 4.3$ and power-law HG weighting with $p = 2.2$.

with the same weighting procedure. The reduced statistics in the π_0 sample becomes obvious through the irregularities in the respective distributions. It becomes clear that the distributions with respect to σ_1^H overlap more with decreasing energy. As explained below, this is related to the tendency of the clustering algorithm to perform worse with decreasing energy in terms of relative efficiency of separating asymmetric π_0 decays compared to symmetric π_0 decays. It is also important to notice that placing a cut for rejecting showers with $\hat{\sigma}_1^H > 0.12$cm, as suggested by the numbers reported in relation to Fig. 37, which is convenient for high energies, will reject a very large signal fraction at lower energies. This is important for the performance aspects treated in Sec. 5, suggesting that such a cut should only be enabled above a certain lower energy cutoff.

It is now worth turning back to the double-peak effect mentioned in relation to the unseparated π_0 distribution with respect to σ_1^H in the upper-left plot of Fig. 35. The effect remains present in the distribution with respect to $\hat{\sigma}_1^H$ in the upper-left plot of Fig. 37. The latter variable is solely used here for studying this effect, since it has more discriminating power and will also be used in Sec. 5. It is interesting to understand whether the two peaks correspond to asymmetric vs symmetric π_0 decays. This hypothesis is tested in Fig. 39, which shows the same distributions as in the upper-left plot of Fig. 37, only that symmetric π_0 decays (left) are separated from asymmetric π_0 decays (right) by a threshold on the (true) track-to-track distance in the z'_{FCP}. It is clear that the distribution of the former is peaked towards low $\hat{\sigma}_1^H$ while the distribution of the latter is peaked towards high $\hat{\sigma}_2^H$. This is a counter-intuitive effect, since one might expect that asymmetric decays would generate larger semi-major widths due to larger distances between the two incidence positions in the transverse plane. However, the low-energy photon of an asymmetric decay would mostly generate low-amplitude digits within the joint shower, with a small relative contribution towards width calculation, relative contribution which is further lowered by power-law weighting. Nonetheless, for small decay asymmetries, the semi-major width should still increase with increasing asymmetry (and thus with track-to-track distance in the z'_{FCP} plane, when dealing with the same, narrow energy interval, which is the case here), which indicates that a turning point should be present for some intermediate value of asymmetry (and thus track-to-track distance).

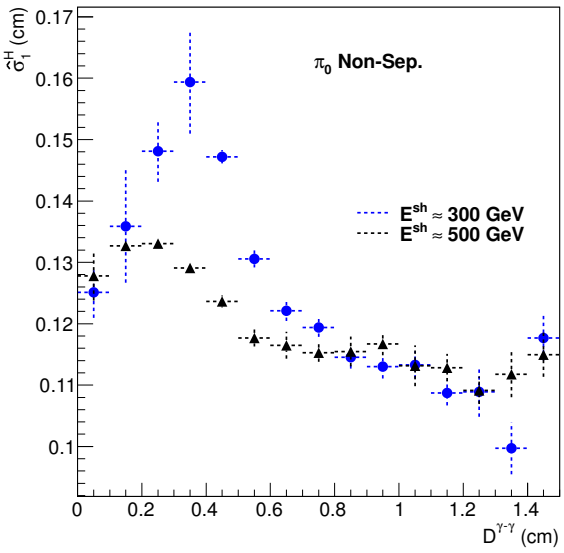

Figure 40: Bin averaged vectorial semi-major HG width $\hat{\sigma}_1^H$ as a function of track-to-track separation in the $z'_{\rm FCP}$ plane $D^{\gamma-\gamma}$, for single π_0 events in two narrow windows of shower energy $E^{sh} := E(c_j^F)$ (blue vs black), for final clusters, after logarithmic LG weighting with $w_0 = 4.3$ and power-law HG weighting with $p = 2.2$.

This turning point is shown in Fig. 40, which shows how the semi-major, vectorial HG width σ_1^H changes with changing track-to-track distance in the $z'_{\rm FCP}$ plane $D^{\gamma-\gamma}$. The behavior is illustrated for both energy windows. For unseparated π_0 events at 300GeV, the turning point is achieved for $D^{\gamma-\gamma} \approx 0.3$cm, while for those at 500GeV, the turning point is achieved for $D^{\gamma-\gamma} \approx 0.2$cm. This suggests that the critical track-to-track distance decreases with increasing energy. There is no obvious explanation for this behavior at this point.

As a further exploration of this effect, Fig. 41 shows the efficiency in rejecting unseparated π_0 decays via a cut on $\sigma_1^H > 0.123$cm. The left-hand plot shows the efficiency as a function of decay asymmetry while the right-hand plot shows the efficiency as a function of track-to-track distance in the $z'_{\rm FCP}$ plane, for both energy intervals. The figure shows that the efficiency is significantly lower for asymmetric π_0 decays with $0.8 < \alpha^{\gamma-\gamma} < 1.0$ then for the rest of the asymmetry range, for both energy intervals. For 300 GeV, the efficiency also suffers a small decrease for symmetric decays, which is in agreement with Fig. 40, which shows a steeper 300 GeV curve for low $D^{\gamma-\gamma}$ distance.

It is desirable to gain some knowledge about the extent to which the poor efficiency of the the cut on σ_1^H is relevant for the overall capability of the entire procedure for rejecting decay photons. Such information is provided by Fig. 42, using the 300 GeV π_0 sample on the left and the 500 GeV π_0 sample on the right. On each side, the distribution of selected π_0 events with respect to decay asymmetry is shown in blue, normalized to 1. The black and red curves show the distributions of non-separated π_0 decays and of non-identified π_0 decays respectively, both of them being normalized with respect to the blue curve, such that the respective bin contents show percentages out of all selected π_0 events. The non-identified are those unseparated π_0 decays that also survive the cut on $\hat{\sigma}_1^H$ variable, namely those that cannot be rejected as background based on neither invariant-mass analysis nor shower-shape analysis. The first interesting aspect is that the (HG+LG) clustering algorithm performs relatively better in separating symmetric decays than asymmetric decays at 300GeV, while at 500 GeV it performs better for asymmetric decays than for symmetric decays. In both cases however, a large fraction of non-identified π_0 decays are concentrated towards high decay asymmetry values. It becomes apparent that this fraction increases with decreasing energy, together with the fraction of asymmetric, non-separated decays, relative to all

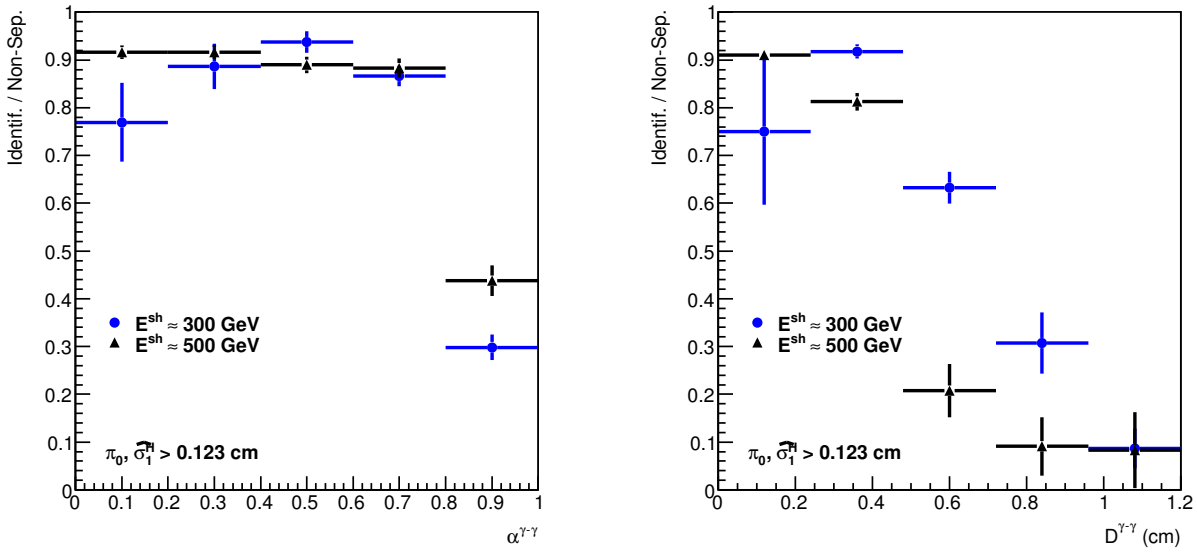

Figure 41: Ratio between identified and non-separated π_0 events, based on the indicated shower-shape cut on $\hat{\sigma}_1^H$, as a function of true decay asymmetry $\alpha^{\gamma-\gamma}$ (left) and true track-to-track distance in the $z'_{\rm FCP}$ plane (right), for single π_0 events in two narrow windows of shower energy $E^{sh} := E(c_j^F)$ (blue vs black), for final clusters, after logarithmic LG weighting with $w_0 = 4.3$ and power-law HG weighting with $p = 2.2$.

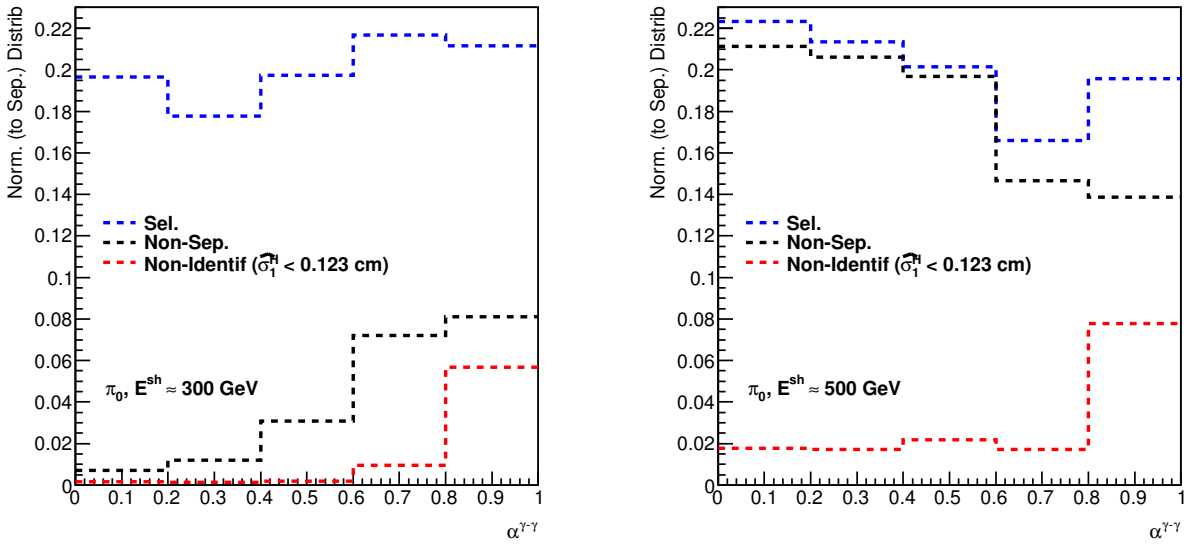

Figure 42: Distribution of selected (blue), non-separated (black) and non-identified (red) single π_0 events, for two shower energy intervals $E^{sh} := E(c_j^F)$ (left vs right), with respect to the true decay asymmetry $\alpha^{\gamma-\gamma}$ (left) and true track-to-track distance in the $z'_{\rm FCP}$ plane (right), for final clusters, after logarithmic LG weighting with $w_0 = 4.3$ and power-law HG weighting with $p = 2.2$.

non-separated decays. Combined with the observation above, according to which asymmetric decays tend to have low values of $\hat{\sigma}_1^H$, this also explains the lack of discriminating power in $\hat{\sigma}_1^H$ at lower energy, pointed out in relation to Fig. 38.

These observations suggest that, if improvements are to be made in relation to the clustering algorithm and/or shower-shape analysis, these improvements should aim specifically at rejecting asymmetric π_0 decays with higher efficiency. More quantitatively, Fig. 42 suggests that such decays represent around 8% of the π_0 sample. They also represent 50% or more of the surviving (non-identified) π_0 decays, meaning that a successful improvement in this direction could reduce the remaining background by one half. The trends explained above in relation to Fig. 42 suggest that these numbers should be approximately valid also for lower energy, while there is no reason to expect significant changes in proton-proton or proton-lead events. Suggestions about such possible improvements are given in Sec. 6.

5 Performance in Proton-Proton Events

This section focuses on studying the FoCal performance in proton-proton collisions in relation to performing prompt photon yield measurements, using the analysis procedure developed in Sections 3 and 4. The study makes use of two data samples storing information about clusters (either final or LG semi-final) generated in proton-proton simulated events at a center of mass energy $\sqrt{s} = 14$ TeV, in the manner described in section 3.8. Two data samples are used, one containing a large fraction of direct-photon induced clusters (let this be called the signal sample), while the other contains clusters induced by all kinds of particles (let this be called the background sample). All cluster entries originate from events produced by running minimum-bias Pythia and saving only the information associated to interesting events in each case. More specifically, for producing the signal sample only events that produce a direct photon in the kinematic acceptance window of FoCal are used. For producing the background sample only the events producing at least one final-state particle in the kinematic acceptance window of FoCal are used. The MC truth information concerning tracks matched to clusters in the signal sample is accessed at all times in order to only take into account the actual clusters induced by direct photons. Although the signal sample contains information about clusters generated by 5000 pp events while the background sample contains information about clusters generated by 100000 pp events, in effect the number of direct photons in the signal sample is enhanced by a factor of 39 when considering the content of the background sample, in terms of relative cross sections. This signal enhancement factor matters when computing a signal/(background+signal) figure of merit later on.

A further kinematic constraint is imposed on all clusters used in this study: only those clusters reconstructed within a pseudorapidty window $3.0 < \eta < 4.0$ are taken into account. This does not affect the invariant mass analysis for those π_0 decays with one cluster induced outside this window and one inside, since the invariant mass variable is computed for all cluster pairs before imposing this cut and to each cluster the lowest possible invariant mass is associated as a cluster level-variable, as explained in Sec. 3.8.

During this section, either LG clustering or HG+LG clustering are being employed, as explained below. In either case, logarithmic weighting with $w_0 = 4.3$ is used for LG clustering and power-law weighting with $p = 2.2$ is used for HG weighting, with parameters settled upon as being optimal in Sec. 4. Both Sections 5.1 and 5.2 below employ cuts on the invariant mass variable IM described by Eq. (65) of Sec. 3.8, on the isolation cone transverse momentum variable p_T^{iso} described by Eq. (66) of Sec.3.8 and on the longitudinal shower variable Λ described by (63) in Sec. 3.7. The first aims at rejecting decay π_0 separated by the clustering algorithm, the second at rejecting fragmentation photons and the third at rejecting hadronic background (mainly charged pions $\pi^{\pm}$) which tend to induce showers which start developing only late in a electromagnetic calorimeter. All these cuts are kept fixed, using the same values as in [11]. Section 5.1 introduces and optimizes the cut on the vectorial, energy-averaged, LG semi-major width $\hat{\sigma}_1^L$, studied in Sec. 4.3.1 in the context of single particle simulations. Section 5.2 introduces and optimizes the cut on the vectorial, energy-averaged, HG semi-major width $\hat{\sigma}_1^H$, studied in Sec. 4.3.2 in the context of single particle simulations. The notion of "decay rejection" (denoted in plots as "dec rej") refers to the combination of cuts on IM, Λ, $\hat{\sigma}_1^L$ and possibly σ_1^H, while "isolation" ("iso") refers to a cut on p_T^{iso}. All cuts shown in plots of this section are acceptance cuts, meaning that only (final or semi-final LG) clusters satisfying the respective condition are being accepted as signal, while the others are rejected as background. For all labels employed by plots below, IM is in units of GeV/c^s, p_T in units of GeV/c, while Λ, $\hat{\sigma}_1^L$ and $\hat{\sigma}_1^H$ are in units of cm. All error shown here are computed using binomial statistics.

5.1 Relevance of High-Granularity Clustering

Here, the performance of FoCal is evaluated for the scenario when only LG clustering is being employed versus the scenario when HG+LG clustering is being employed, in order to provide a final evaluation of the impact of the HG segments. In terms of shower-shape analysis, only the LG information is being used in either case, as this section also serves the secondary purpose of independently optimizing the cut

on the LG $\hat{\sigma}_1^L$ variable in the HG+LG clustering scenario, before introducing and optimizing the cut on the HG σ_1^H in Sec. 5.2.

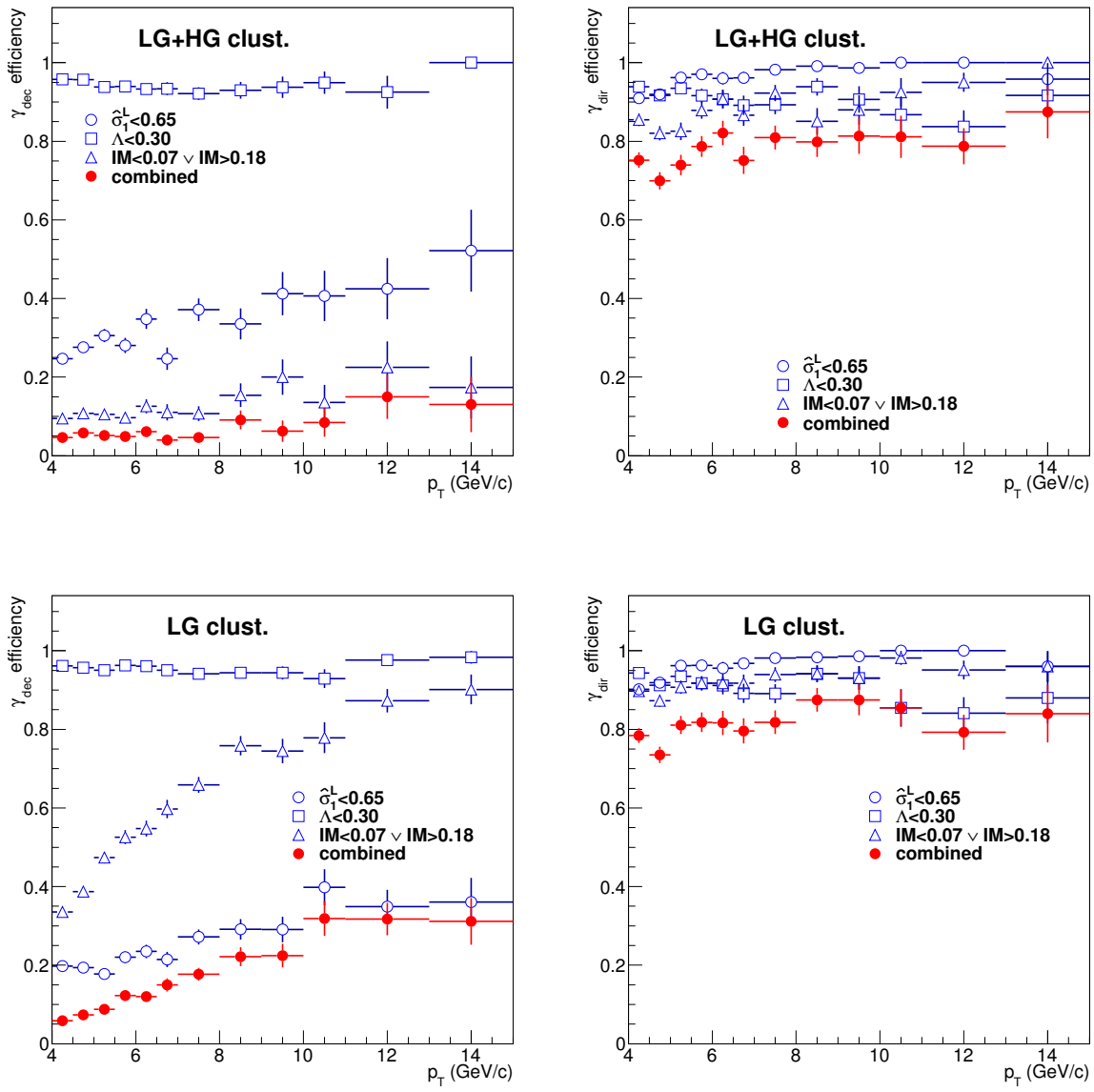

Figure 43: Rate of acceptance as direct photon, as a function of transverse momentum, for showers produced by the LG+HG (top) and LG (bottom) clustering, matched to decay (left) or direct (right) photon tracks, passing a transverse shower-shape cut (open circles), a longitudinal shower development cut (open squares), an invariant-mass cut (open triangles) and all 3 cuts (red, full circles).

Figure 43 illustrates the impact of the HG segments of reducing the rate of decay γ acceptance, while keeping the rate of signal acceptance at almost the same level, when only applying the cut on the invariant mass variable, on the semi-major LG width, and on the longitudinal variable. Note, in the two plots on the left hand side, that when LG clustering is employed (bottom), most of the background rejection is achieved via the shower-shape cut on the σ_1^L variable, while the invariant mass does most of the rejection in when LG+HG clustering is employed (top). This is mainly a consequence of the fact that a much larger fraction of the π_0 decays are separated by the algorithm, as discussed in Sec. 4. It can also be noticed that when moving from LG+HG clustering to LG clustering, the γ_{dec} rejection rate due to the

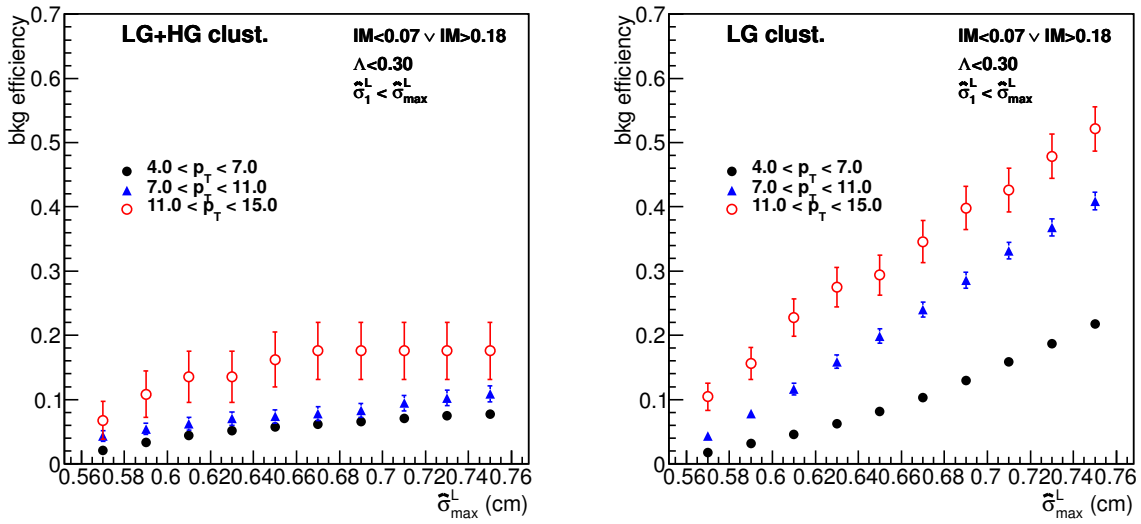

Figure 44: Rate of background acceptance after decay rejection as a function of maximal $\hat{\sigma}_1^L$ defining the transverse shower-shape cut, within 3 different transverse momentum intervals, for LG+HG clustering (left) and for LG clustering (right).

cut on the semi-major width σ_1^L increases, although the showers that are no longer separated due to the absence of HG information have the same semi-major LG width as the final clusters that are formed in pairs out of them when the HG information is present. The increase in rejection rate is just an effect due to the shifting of the cluster p_T spectrum, which is a decreasing function of p_T, from lower to higher p_T when abandoning HG information, since cluster pairs are effectively combined together with their reconstructed energy (and thus reconstructed transverse momentum). In this manner, more clusters induced by decay photons fall within the $4 < p_T(\mathrm{GeV}/c) < 15$ range, out of which most are also rejected by the cut on $\hat{\sigma}_1^L$, thus leading to an overall increase in rejection rate. The effect can also be understood as a shifting towards higher p_T of the curve depicted by blue circles when switching from the upper left plot to the lower left plot. The same cut $\hat{\sigma}_1^L < 0.65$cm is employed both situations, which results from the optimization process described below. On the right-hand plots of this figures, when switching from LG clustering to LG+HG clustering, one can also notice a slight decrease in γ_{dir} acceptance via the invariant mass cut. This is due to the increased number of shower pairs, with one shower induced by a direct photon, falling within the π_0 rejection window, which is directly related to the increased number of reconstructed showers.

Figures 44, 45 and 46 illustrate the process of optimizing the (combined) decay rejection by systematically varying the cut on the semi-major LG width $\hat{\sigma}_1^L$. For this reason, the p_T range used above is divided in three intervals, which correspond to the 3 type of labels used in these figures, as indicated by the legends. Fig. 44 shows, for each p_T interval, the fraction of background clusters surviving the decay rejection cut, for different values of the cut on $\hat{\sigma}_1^L$, showing the in parallel the LG+HG and the LG clustering scenarios. Following the same scheme, Figures 45 and 46 show the fraction of surviving signal and the ratio between surviving signal and all surviving clusters respectively. The latter is computed by taking into account the relative signal enhancement discussed at the beginning of Sec. 5 and can be seen as a figure of merit.

It is clear from Fig. 44 that the exact value of the cut on $\hat{\sigma}_1^L$ plays a much more important role in deciding the amount of rejected background in the context of LG clustering than in the context of HG clustering. On the other hand, its influence of signal rejection is the same for LG and LG+HG clustering, as shown in Fig. 45. Note that with increasing tightness of the cut, the signal acceptance is first affected in the lower p_T interval, than in the intermediate p_T interval and finally in the higher p_T interval. Finally, tightening this cut increases the signal/all ratio under both LG and LG+HG clustering scenarios, while also increasing the associated uncertainties. Also note that the lowest p_T bin has the highest signal/all

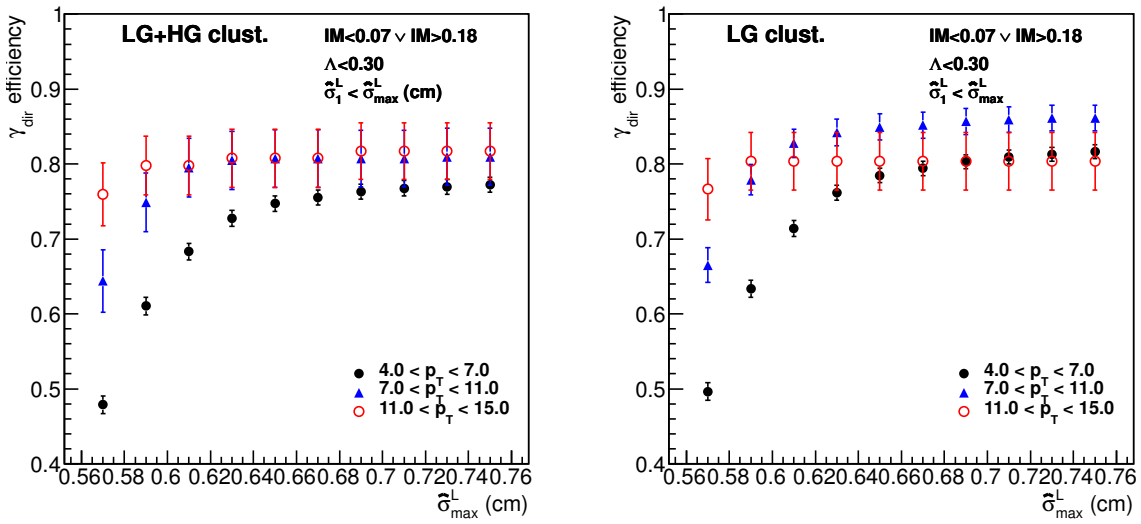

Figure 45: Rate of signal (clusters induced by direct photons) acceptance after decay rejection as a function of maximal $\hat{\sigma}_1^L$ defining the transverse shower-shape cut, within 3 different transverse momentum intervals, for LG+HG clustering (left) and for LG clustering (right).

ratio in the LG clustering scenario but the lowest signal/all ratio in the LG+HG clustering scenario.

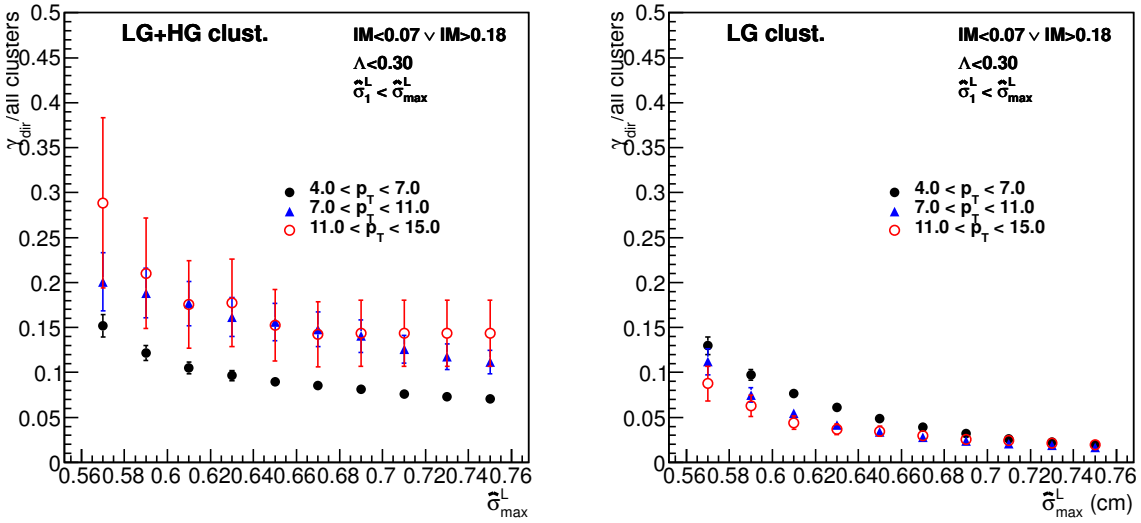

Figure 46: Ratio between signal clusters and all clusters surviving the decay rejection cut, as function of maximal $\hat{\sigma}_1^L$ defining the transverse shower-shape cut, for 3 different transverse momentum intervals, for LG+HG clustering (left) and for LG clustering (right).

Following the approach of giving more priority to signal acceptance than to background rejection indicates an optimal cut on $\hat{\sigma}_1^L = 0.65$cm for both scenarios, for which signal acceptance is kept on the edge of the high plateaus visible in Fig. 45.

For this chosen cut, Fig. 47 illustrates the p_T dependence of the signal and background acceptance rates for the combined decay rejection cut (studied above), the isolation cut and for both these cuts combined, again contrasting the LG and the LG+HG clustering scenarios (bottom vs top). In terms of background

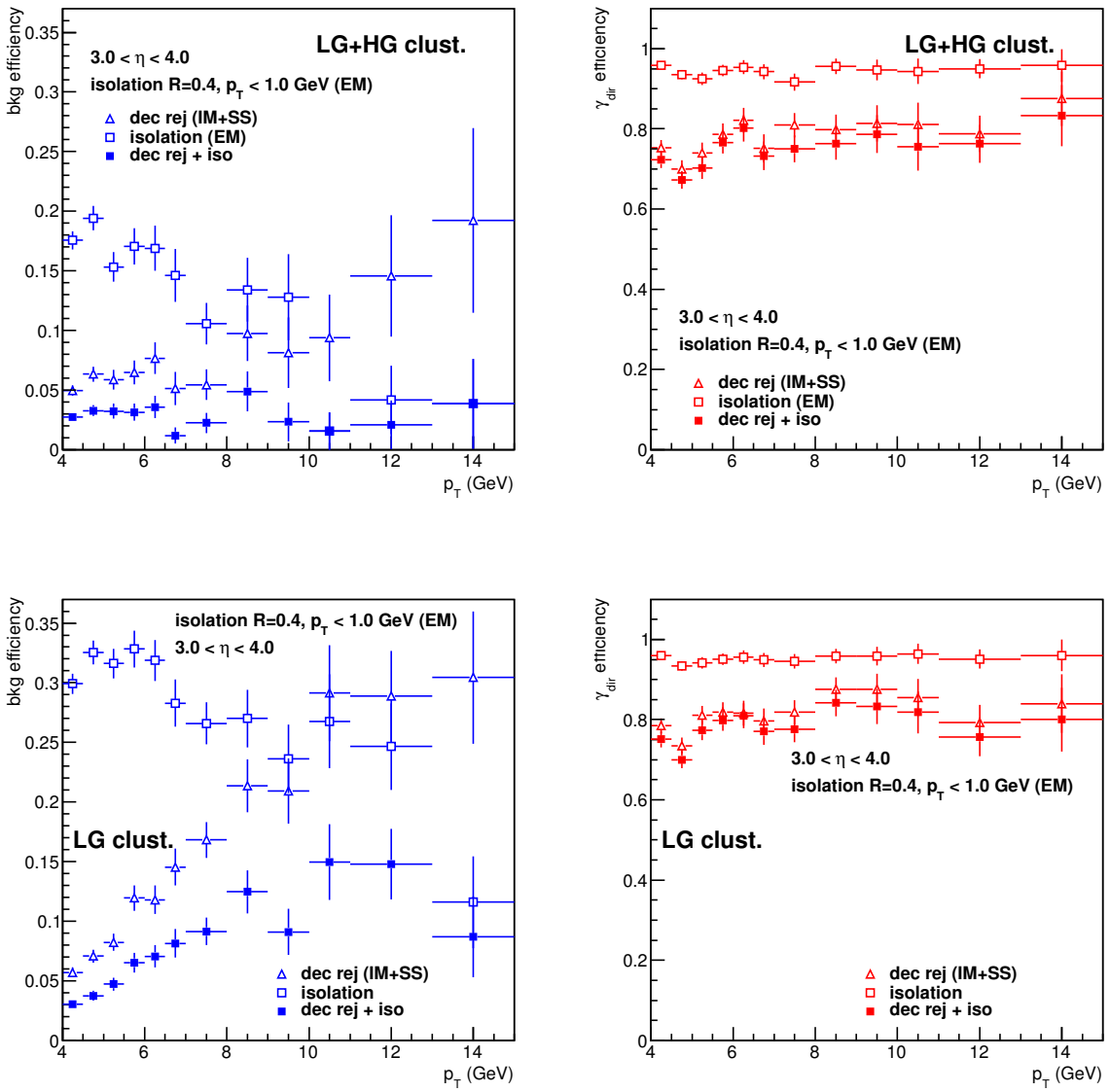

Figure 47: Background acceptance (left) and signal acceptance (right) efficiencies, as a function of transverse momentum, using LG+HG clustering (top) and LG clustering (bottom), due to the decay rejection cut (open triangles), the isolation cut (open squares) and both the isolation and decay rejection cuts combined (full squares).

rejection, it is clear that both the decay rejection cut (which is generally more efficient for high p_T) and the isolation cut (which is generally more efficient for low p_T) are dramatically improved when introducing the HG segments. Overall, their combination reduces the upper limit on background acceptance from approximately 15% to approximately 5% when introducing HG clustering. In parallel, signal acceptance remains at approximately the same values when introducing the HG segments.

Figure 48 shows the dependence of the ratio between clusters induced by signal and clusters induced by signal and background. The two plots correspond to the two clustering scenarios. Different curves show the same ratio after applying the same two cuts as in Fig. 47 and their combination. In addition to Fig. 47, the signal/all ratio is also shown for the situation when no cut is being applied. It is instructive to first notice the overall increase in the signal/all curve for no applied cuts, when switching from LG+HG

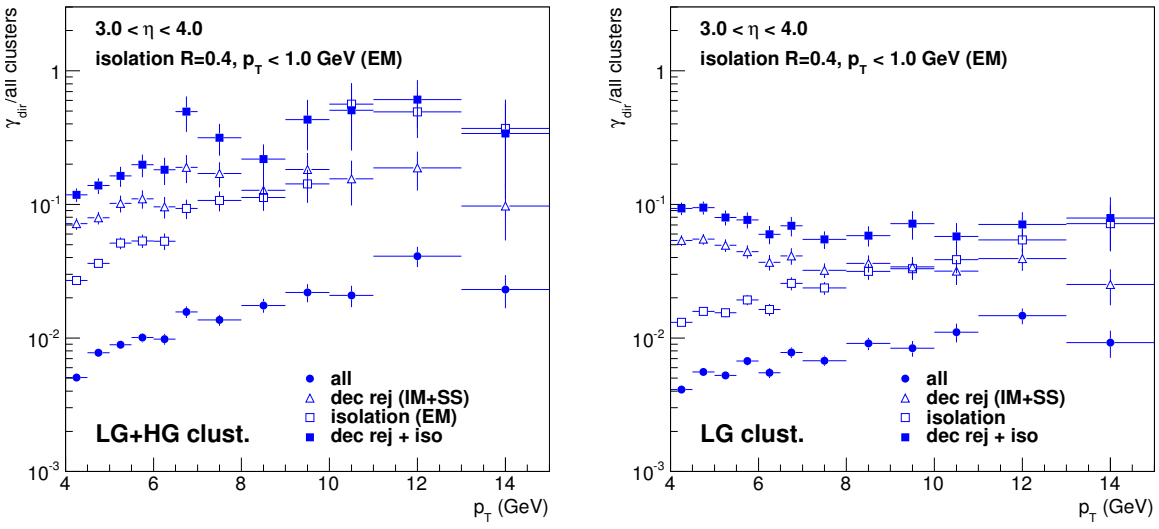

Figure 48: Ratio between signal and signal+background as a function of transverse momentum, for LG+HG clustering (left) and for LG clustering (right), for all clusters (full circles), for clusters passing the decay rejection cut (open triangles), for clusters passing an isolation cut (open squares) and for clusters passing both the isolation and decay rejection cuts (full squares).

to LG clustering, together with the steeper increase with p_T at low p_T in the former scenario. This is again related to an overall shift towards the right of the p_T spectrum associated to clusters induced by decay photons, due to lack of separation when only using the LG information during. This increases the amount of background clusters in the $4 < p_T(\mathrm{GeV}/c) < 15$ window, thus reducing the signal/all ratio. When the decay rejection is employed, the ratio shows a clear increase with increasing transverse momentum when HG information is used, whereas a clear decrease is visible with decreasing transverse momentum when HG information is not used. When both cuts are applied (upper-most curve), the ratio stays within 0.1 to 0.5 when HG information is used, but below 0.1 when HG information is not used. This again indicates the clear advantage introduced by the HG segments at the level of the cluster-finding process.

5.2 Relevance of High-Granularity Variables

This section explores the performance improvements entailed by introducing a cut on the HG semi-major width $\hat{\sigma}_1^H$, in addition to the cuts mentioned above, solely in the context of HG+LG clustering. Together with the cut on $\hat{\sigma}_1^L$ discussed in Sec. 5.1, this would constitute a more sophisticated transverse shower shape cut, which in turn would be part of an improved decay rejection cut. As pointed out in section 4.3.2, a cut on $\hat{\sigma}_1^H$ should only be enabled for shower energies higher than a certain threshold, which is denoted here by $E_{\mathrm{min}}^{\mathrm{sh}}$. Otherwise, the HG cut would eliminate too much signal.

First, Fig. 49 illustrates the effect of the cut on the HG semi-major width on decay-photons and direct photon, in parallel with the other two cuts employed by the decay rejection cut. The transverse shower-shape cut denoted by the blue circles states that, in addition to passing the $\hat{\sigma}_1^L$ cut also used before, a shower also needs to have a reconstructed energy lower than 150 GeV or a semi-minor HG width $\hat{\sigma}_1^H$ smaller than 0.14 cm. Note the decreased decay photon acceptance rate indicated by this curve at high p_T when comparing to the upper-left plot of Fig. 43, as well as the overall decrease in the red curve. Given that the direct photon curves do not significantly change when introducing this cut, which can be seen by comparing the right plot of Fig. 49 with the upper-right plot of Fig. 43, this is a first, clear

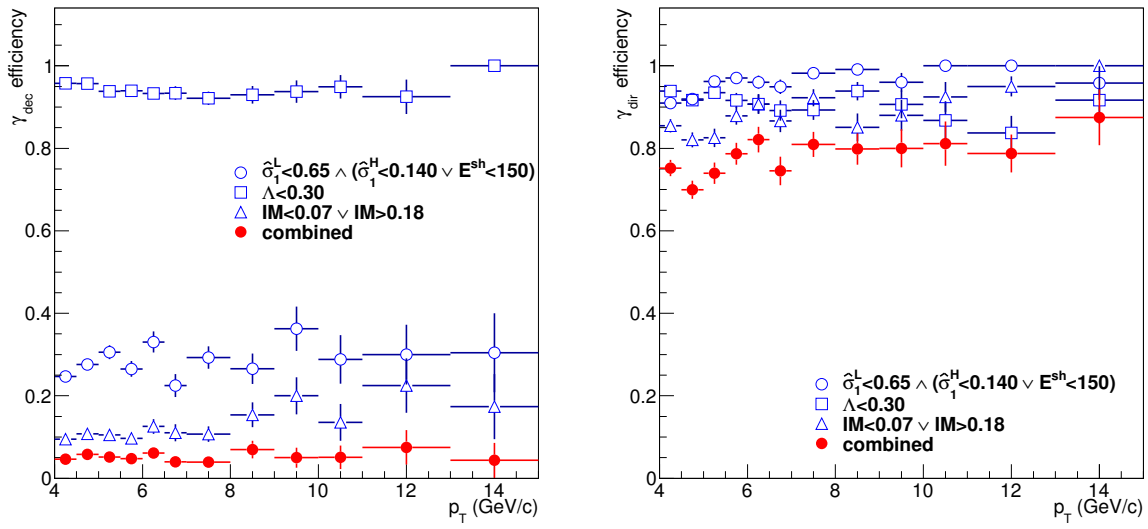

Figure 49: Rate of acceptance as direct photon, as a function of transverse momentum, for showers produced by the LG+HG clustering, matched to decay (left) or direct (right) photon tracks, passing a transverse shower-shape cut (open circles), a longitudinal shower development cut (open squares), an invariant-mass cut (open triangles) and all 3 cuts (red, full circles).

indication of the HG variable improving the performance towards high p_T.

The specific value of the $\hat{\sigma}_1^H$ cut and that of the associated energy threshold $E_{\rm min}^{\rm sh}$ were chosen via an optimization procedure similar to the one employed in Sec. 5.1. This is illustrated in Fig. 50, which shows in parallel the background acceptance rate and signal acceptance rate after applying the decay rejection, for the 3 p_T bins, for 4 different values of the energy threshold and for a multitude of values of the cut on the semi-major HG width. Background acceptance rate is shown on the left-hand side while signal acceptance rate is shown on the right-hand side, with the three horizontal lines corresponding to the three p_T bins and the four colors corresponding to the four values of the energy threshold. In all cases, the horizontal axis spans different values of the cut on the HG semi-major width $\hat{\sigma}_1^H$. The figure shows that for multiple combinations of values of the HG width and the associated energy threshold, the background is reduced for intermediate and high transverse momentum. For the lowest value of the energy threshold $E_{\rm min}^{\rm sh} = 75$GeV the signal acceptance rate is also drastically reduced in the lower p_T bin, even for relatively loose cuts on the width, which is not desirable. The combination given by $\hat{\sigma}_1^H = 0.14$cm and $E_{\rm min}^{\rm sh} = 150$GeV is of most convenience, as it keeps the signal acceptance rate high in all p_T bins while reducing the background acceptance rate for intermediate and high transverse momentum. This is all in agreement with the findings explained in section 4, both the decreased background acceptance towards high p_T and the decreased signal acceptance at low p_T when an energy cutoff is not specified. There, the energy cutoff had been suggested due to the tail appearing in the γ distribution with respect to σ_1^H at low energies. For a certain window of pseudorapidity that is kept fixed here, transverse momentum roughly increases with energy, thus explaining the correspondence.

Figure 51 further illustrates the improvement determined by the additional cut on the semi-major HG width in the context of overall signal and background acceptance rates as a function of p_T. The improvements can be seen in the high p_T bins of the "dec rej" curve in the plot on the right-hand side, when comparing it to the upper-left plot of Fig. 51, as well as in the 2nd largest p_T bin of the "dec rej + iso" curve, where all background is essentially removed.

The same improvement can be seen in the context of the signal/all ratios shown in Fig. 52 relative to the

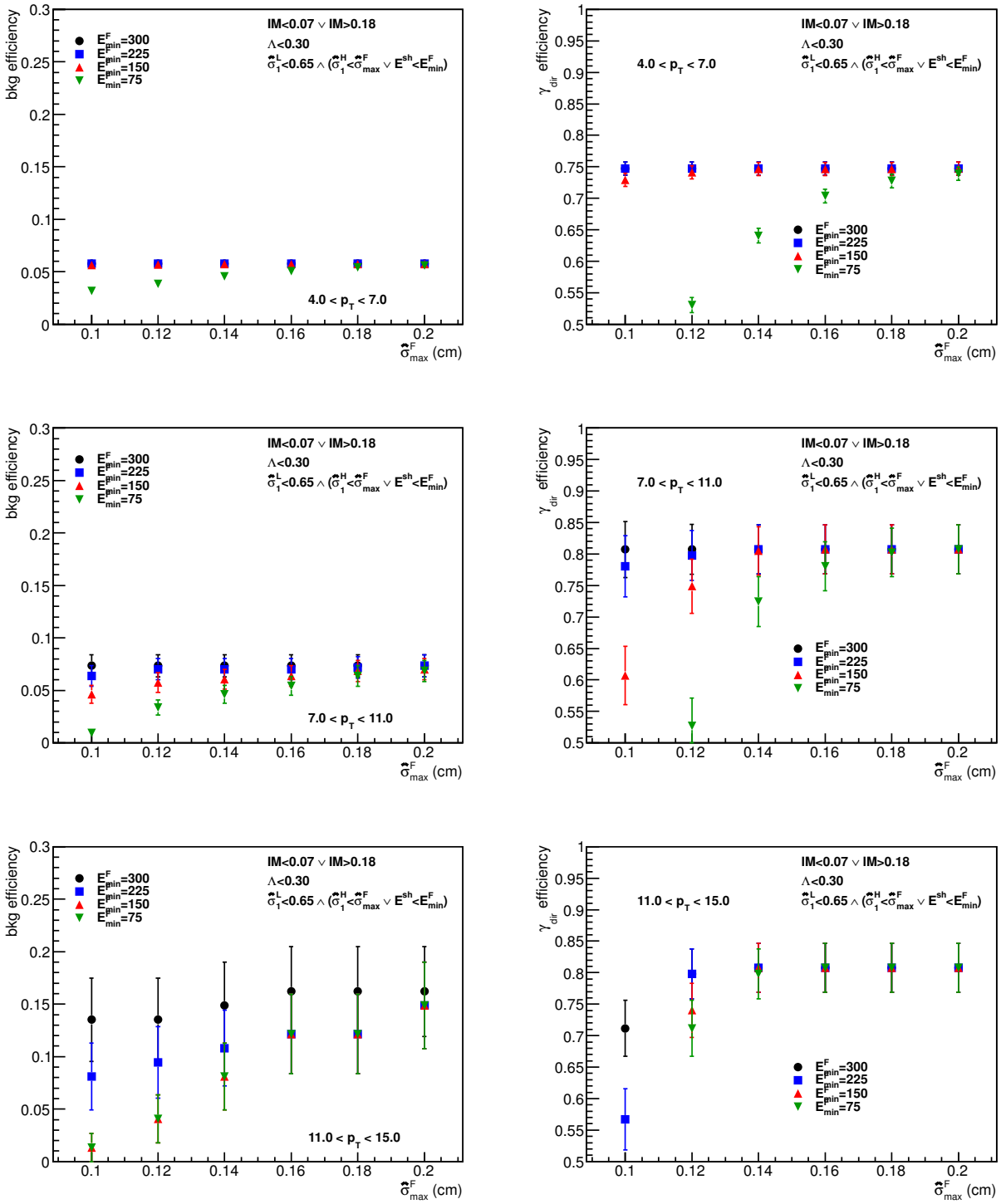

Figure 50: Rate of background acceptance (left) vs rate of signal acceptance (right) after decay rejection as a function of maximal $\hat{\sigma_1}^H$ defining the HG component of the transverse shower-shape cut, within 3 different transverse momentum intervals (top vs center vs bottom), for LG+HG clustering, for different values of the shower energy treshold $E^{\rm sh}_{\rm min}$.

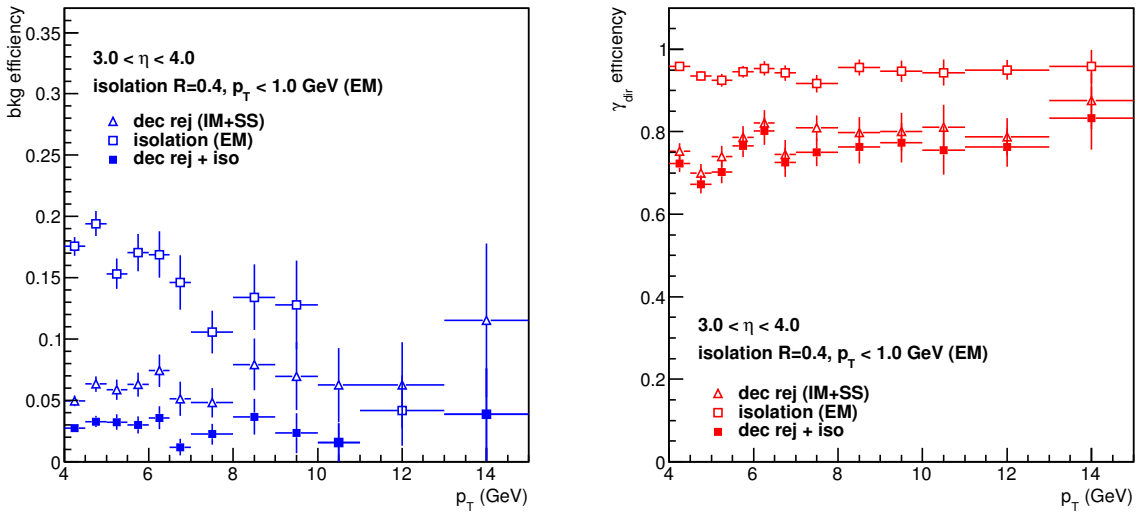

Figure 51: Background acceptance (left) and signal acceptance (right) efficiencies, as a function of transverse momentum, using LG+HG clustering, due to the decay rejection cut (open triangles), the isolations cut (open squares) and both the isolation and decay rejection cuts combined (full squares).

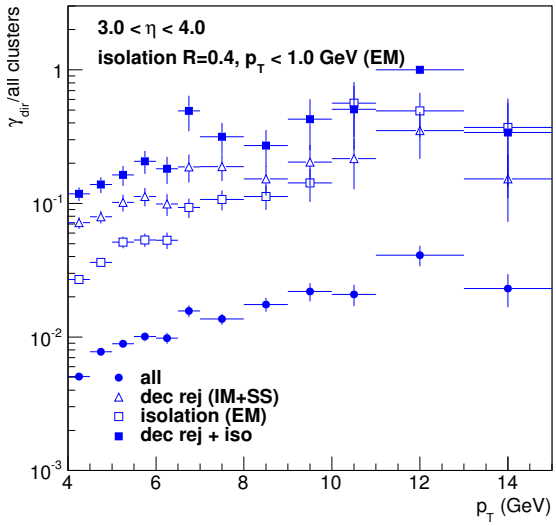

Figure 52: Ratio between signal and signal+background as a function of transverse momentum, for LG+HG clustering (left) and for LG clustering (right), for all clusters (full circles), for clusters passing the decay rejection cut (open triangles), for clusters passing an isolation cut (open squares) and for clusters passing both the isolation and decay rejection cuts (full squares).

analysis scenario	signal/all for: $8\text{GeV}/c < p_T < 15\text{GeV}/c$
LG clust. and LG cut	0.064 ± 0.007
LG+HG clust. and LG cut	0.345 ± 0.064
LG+HG clust. and LG+HG cut	0.406 ± 0.078

Table 3: Value and error of signal/all ratio for showers with a transverse momentum higher than $8\text{GeV}/c$ passing the combined decay rejection and isolation cuts, for the 3 different analysis scenarios.

left plot of Fig. 48. the signal/all ratio after decay rejection is somewhat higher for high p_T, while the reaching the largest possible value of 1 after both decay rejection and isolation cuts, for the 2nd largest p_T bin.

As a summary of Sec. 5 in terms of high p_T performance (where the HG cut has some influence), table 3 shows the ratio between the number of signal showers and the number of all showers passing the combined decay rejection and isolation cuts for the 3 different situations studied here. The first line shows the values associated to LG clustering with the LG semi-major width cut, the second line shows the values associated to LG+HG clustering with the LG semi-major width cut and the third line shows the values associated to LG+HG clustering with LG and HG semi-major width cuts. The numbers are essentially obtained by combining the data contained by the higher p_T bins of the "dec rej + iso" curve (denoted by blue squares) of Fig. 48-right, Fig. 48-left and Fig. 52 respectively. It is clear that introducing the HG segments at the level of the clustering and shower reconstruction algorithms has a high overall improvement in terms of signal/all ratio within the respective p_T range. Introducing the cut on the HG variable further improves the figure, although the effect is much smaller and barely visible if the uncertainty ranges are taken into account.

6 Discussion, Conclusions and Outlook

In section 1, the manuscript started with motivating why double electromagnetic showers, mostly caused by neutral pion decays, need to be recognized and rejected as background of single electromagnetic showers, mostly induced by direct/prompt photons. Such recognition capabilities are needed in order to perform measurements of the nuclear modification factor R_{pPb} of prompt photon yields, which can provide information about low x and low Q^2 gluon distribution functions in the nucleons, which might show gluon saturation effects. While providing brief experimental details in Sec. 2, it was argued that high-granularity layers/segments, which are included in the currently agreed upon FoCal design, are needed for successfully recognizing such double photon decays. Section 3 explained the simulation chain, focusing on the shower-finding algorithm (which consists of a transverse clustering algorithm and a longitudinal shower reconstruction algorithm) developed by the ALICE-FoCal collaboration and used here, as well as on formal aspects related to constructing various shower-level discriminating variables. At the level of the algorithm, it is essential that low-granularity (LG) information is processed separately from high-granularity (HG) information, the two components being integrated only at the end. It is thus easy to also run the entire analysis procedure while ignoring HG information, which allows for an assessment of the value of the HG segments. At the level of the observables, the emphasis was placed on developing reliable measures of the semi-major widths of showers, which allows one to recognize combined photon showers that are not separated by the shower-finding algorithm and thus cannot be rejected based on invariant mass considerations.

Restricting the study to the 4m position scenario, Sec. 4 offered a detailed exploration of the efficiency of the shower-finding algorithm, of the invariant mass analysis and shower-shape analysis procedures, comparing and contrasting the 2 situations of FoCal with HG segments and FoCal without HG segments. The invariant mass analysis can only (correctly) reject neutral pion decays that are separated by the shower-finding algorithm, while the shower shape analysis is mostly useful for rejecting neutral pion decays which are not separated by the shower-finding algorithm. It became clear that the HG segments allows the shower-finding algorithm to separate a much larger fraction of neutral pion decays, most of which are then rejected via invariant mass analysis. For instance, for pion energies between 100 to 150 GeV, the HG information takes the average separating efficiency from around 30% to around 90%. On the other hand, effective shower shape analysis relies on specific procedures of weighting the amplitude of the digits. The discriminating power in the semi-major shower width computed from LG information is enhanced when employing logarithmic weighting in the LG segments, which is mostly relevant for energies below 250 GeV. Discriminating power in the semi-major shower width computed from HG information is achieved when employing power-law weighting in the HG segments, which is mostly relevant for energies above 150 GeV. Section 4.1 describes the process of fine-tunning the parameters employed by these two weighting methods. While the former method had been used before, the latter one was first introduced during this study, up to the author's knowledge. Sec. 5 provided more realistic performance figures in the context of proton-proton simulated collisions at 14 TeV, starting with a comparison between the two situations characterized by the access versus lack of access to HG information at the level of the shower-finding algorithm and continuing with an evaluation of the improvements enabled by the use of HG discriminating variables.

Towards the beginning of section 4 it was argued that the increased discriminating power in the semi-major LG width when using logarithmic weighting is directly related to the removal of certain lattice artifacts. Since the lattice width of the LG segments approximately matches the Moliere Radius, certain biases in the orientation of the reconstructed showers are present for certain particle impact positions relative to the LG lattice. For non-separated neutral pion showers, this invalidates the assumption that the semi-major width of the shower is measured along the line passing through the two impact points of the two decay photons with the transverse plane. In parallel, the discrimination power of the semi-major HG width is restored by power-law weighting, since this places more emphasis on shower cores and less on shower tails, the latter being very sensitive to fluctuations in the HG segments. In terms of exact shower variables, towards the end of section 4 it was shown that the most convenient LG and HG transverse variables for use in shower-shape analysis are the energy weighted semi-major LG and HG vectorial

widths, which take into account cluster widths at the level of each segments as well as their orientations in the transverse plane. These variables are constructed in such a way that they effectively provide information about showers compressed along their longitudinal directions, which would be available if the deposited energies were first integrated longitudinally, although information is actually hierarchically processed and integrated, first in the transverse plane of each segment and then at the level of the entire detector. It was also suggested that a cut on the respective HG variable should only be enabled for high shower energies.

In the context of HG shower-shape analysis methods explored towards the end of Sec. 4, an interesting effect was found in relation to the distribution of unseparated π_0 showers with respect to the HG semi-major width. For unseparated showers induced by neutral pions with the same energy (say 300 or 500 GeV), the semi-major HG width increases with increasing decay asymmetry, due to the increasing distance between the impact positions of the two photons. However, there exists a critical asymmetry value beyond which this tendency is reversed, as the semi-major HG width suffers a sudden decrease and than remains at low values for larger asymmetries (lower semi-major width values than for symmetric decays). Beyond this critical asymmetry value, which depends on the pion energy, the digits induced by the lower-energy photon are to small to influence the value of the reconstructed semi-major width. This implies a systematic lack of discriminating power for asymmetric π_0 decays. When using the HG information in the shower-finding algorithm, with increasing energy an increasing fraction of the unseparated π_0 showers are asymmetric, thus avoiding both the invariant mass cut and the shower-shape cut. At least for energies above 300 GeV, such showers make up more than 50% of the surviving π_0 background, although they only represent at most 10% of the original π_0 background. More details are given in the last part of Sec. 4.3.2. In might thus be worth considering further research in the direction of improving asymmetric π_0 decay separation at the level of the clustering algorithm or at least recognition at the level of shower-shape analysis. The former approach would aim at improving the criteria used for distant seed rejection within the influence zone of a higher energy seed at the level of segment-level clustering – see Sec. 3.3, second line of step 3(b)iB of the clustering algorithm. Certain improvements systematically studied in [26] did not significantly increase the π_0 separation efficiency, while keeping the single γ identification efficiency at the same level. However, this study remained within the hierarchical approach of only using information from within the respective segment, whereas correlations among relative seed positions and energies among multiple segments of the same type could lead to better results. Following the latter approach, instead of adjusting the rejection criterion, one may simply save, at the level of each segment-level cluster, information about seeds rejected in the proximity of the cluster, related to relative energy and orientation (at least information about the largest such seed). When combining segment-level clusters into semi-final clusters, such information can be combined into another shower-level variable, for use in the shower-shape analysis. All these are of interest mostly in the context of the HG segments. However, given the close proximity of the first HG segment to the the front of the detector, the procedure might be precluded by branches of certain double showers not starting to develop at that point. In this sense, adding a third or possibly even a fourth HG segment to the current FoCal design, at larger depths in the detector would probably provide opportunities for better improvements along these lines. Thus, the two aspects should be studied together during future research and not separately.

In the context of proton-proton simulated collisions, Section 5 further elaborated on the comparison of the LG versus the LG+HG clustering scenarios, as well as on understanding the effect of the cuts on the LG and HG semi-major vectorial widths introduced in Sec. 3 and studied in Sec. 4. The first aspect was to a great extent motivated by the existence of another calorimeter within approximately the same pseudorapidity window as FoCal at 4m, namely the LHCb electromagnetic calorimeter. Being placed at larger distances from the interaction point while also having higher granularities than the LG segments of the proposed ALICE-FoCal detector, the FoCal without HG segments can be seen as a good approximation for estimating its performance. It turns out that the performance under the LG clustering scenario is significantly poorer, in terms of fraction of rejected background and in terms of ratio between the signal induced clusters and all clusters. It is estimated that due to the higher background acceptance rate, combined with larger systematic errors originating in the decay rejection cut, a low granularity calorimeter would not provide small enough systematic errors associated to the overall direct/prompt photon yield measurement for reliably measuring the R_{pPb} with the precision required for verifying the

existence of gluon-saturation effects. Further on, it was shown that placing an appropriate cut on the HG semi-major width further improves the performance within the HG+LG clustering scenario, although this improvement is marginal for the transverse momentum interval studied here. It is also worth noticing that the improvement brought about by the introduction of the HG cut is very specific to the 4m position, which was solely studied here. In the context of the 8m position (which is currently preferred by the ALICE-FoCal collaboration), double photon decays would be better separated in the plane of the detector, thus decreasing the fraction of unseparated π_0 showers subject to shower-shape analysis and shifting the relevance of HG shower-shape variables to significantly higher energies.

Mostly focusing on performance, analysis and algorithm aspects, this has been a successful, simulation-oriented study of electromagnetic shower recognition using the proposed FoCal detector for the ALICE experiment at CERN.

Acknowledgments:

The author is grateful to his supervisor, Thomas Peitzmann, for this challenging research topic and for skillful guidance and supervision throughout the entire academic year. The author is thankful to Davide Lodato for his comprehensive introduction to the simulation chain and the associated computational tools and for his systematic supervision, especially during the initial part of this research. Many thanks to Marco van Leeuwen for his guidance, for his patient explanations, for his trust and for sharing his plotting tools and templates, which were effectively used, to a higher or lower extent, for producing most of the figures shown in the 2nd half of this manuscript. Cenek Zach is acknowledged for his high-quality and user-friendly simulation framework, consistently used for producing the results shown here. Martijn Reicher, Gert-Jan Nooren and Terry Awes are acknowledged for constructive comments and criticism. The author is also thankful to Gordian Zomer and Tom Bannink for a productive collaboration and to Alexandru Dobrin for his help with technical issues. Last but not least, the author is thankful to his office colleagues: Naghmeh, Jasper, Philip, Alexandros and Alexandru for maintaining a pleasant work atmosphere.

References

[1] ALICE Collaboration, *The ALICE Experiment*, Public website "`http://aliceinfo.cern.ch/Public/en/Chapter2/Chap2InsideAlice-en.html`".

[2] G. M. García, *Advances in Quark Gluon Plasma*, arXiv:1304.1452.

[3] N. Nakamura et al. (Particle Data Group), Journ. of Phys. G**37**, 075021 (2010).

[4] M. Breidenbach et al., *Observed Behavior of Highly Inelastic Electron-Proton Scattering*, Phys. Rev. Lett. **23** (16) 930-934 (1969).

[5] R. Snellings, *Elliptic flow: A brief review*, New Journal of Phys. **13** (5) 055008 (2011).

[6] K. Adkox et al., *Suppression of Hadrons with Large Transverse Momentum in Central Au+Au Collisions at* $\sqrt{s_{NN}} = 130 GeV$, Phys. Rev. Lett. **88** 022301 (2001).

[7] V.N. Gribov, L.N.Lipatov, Sov. J. Nucl. Phys. **15**, 438 (1972); G. Altarelli and G. Parisi, Nucl. Phys. B **126**, 298 (1977); Yu. L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977).

[8] L. McLerran and R. Venugopalan, *Gluon distribution functions for very large nuclei at small transverse momentum*, Phys. Rev. D **49**, 3352 (1994).

[9] L.V. Gribov, E.M. Levin and M.G. Ryskin, *Semihard Processes in QCD*, Phys. Rep. **100** 1 & 2, 1-150 (1983).

[10] E. Iancu, A. Leonidov, L. McLerran, *The Color Glass Condensate: an Introduction*, arXiv:hep-ph/0202270.

[11] ALICE FoCal Collaboration, *A Forward Calorimeter (FoCal) for the ALICE Experiment*, Letter of Intent (2013).

[12] R. Ichou and D. d'Enterria, *Sensitivity of isolated photon production at TeV hadron colliders to the gluon distribution in the proton*, Phys. Rev. D **82**, 014015 (2010).

[13] A.H. Rezaeian, *CGC predictions for p+A collisions at the LHC and signature of QCD saturation*, Phys. Lett. B **718**, 1058 (2013).

[14] ALICE Collaboration, Technical Design Reports.

[15] E. Segré, *Nuclei and Particles*, Benjamin (1964) p. 65 ff.

[16] Particle Data Group, *Atomic and Nuclear Properties of Materials for more than 300 materials*, Public website "`http://pdg.lbl.gov/2011/AtomicNuclearProperties/`".

[17] B. Rossi, *High Energy Particles*, Prentice-Hall, New Jersey (1952).

[18] W.R. Nelson et al, *Electron-Induced Cascade Showers in Copper and Lead at 1 GeV*, Phys. Rev. **149**, 201 (1966).

[19] G. Bathow et al, *High Energy Particles*, Nucl. Phys. **B20**, 592 (1970).

[20] LHCb Collaboration, Technical Design Reports, Calorimeters (2000).

[21] R. Wigmans, *Calorimetry*, Clarendon Press, Oxford (2000).

[22] R. Turchetta et al, *A monolithic active pixel sensor for charged particle tracking and imaging using standard VLSI CMOS technology*, Nucl. Instrum. Meth. **458**, 677 (2001).

[23] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 Physics and manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175; T. Sjöstrand, S. Mrenna, and P. Skands, *A Brief Introduction to PYTHIA 8.1*, arXiv:0710.3820.

[24] *GEANT - Detector Description and Simulation Tool*, CERN Program Library Long Writeup W5013.

[25] R. Brun, P. Buncic, F. Carminati, A. Morsch, F. Rademakers, K. Safarik on behalf of the ALICE collaboration, *The AliRoot framework, status and perspectives*, Proceedings to *Computing in High Energy and Nuclear Physics* (2003).

[26] T. Bannink, *Improving clustering methods for the FoCal detector*, Bachelor Thesis, Utrecht University (2013).

[27] T.C. Awes et al, *A simple method of shower localization and identification in laterally segmented calorimeters*, Nucl. Instrum. Meth. Phys. Res. **A311** 1 & 2, 130-138 (1992).

[28] D.J. Fegan et al, *γ/hadron separation at TeV energies*, J. Phys. G: Nucl. Part. Phys. **23** 1013 (1997).

[29] G. Zomer, *FoCal particle separation*, Bachelor Thesis, Utrecht University (2013).

[30] N.I. Fisher, *Statistical Analysis of Circular Data*, Cambridge University Press (1993).