Development and optimization of the highly granular silicon-tungsten electromagnetic calorimeter for the International Linear Collider

Kostiantyn Shpak

To cite this version:

Kostiantyn Shpak. Development and optimization of the highly granular silicon-tungsten electromagnetic calorimeter for the International Linear Collider. High Energy Physics - Experiment [hep-ex]. Université Paris-Saclay, 2018. English. NNT : . tel-02063211

HAL Id: tel-02063211

https://hal.science/tel-02063211

Submitted on 11 Mar 2019

université
PARIS-SACLAY

NNT : 2018SACLX016

Development and optimization of the highly granular silicon-tungsten electromagnetic calorimeter for the International Linear Collider

Thèse de doctorat de l’Université Paris-Saclay
préparée à l’Ecole polytechnique

École doctorale n°576:
Particules Hadrons Énergie et Noyau : Instrumentation, Image, Cosmos et Simulation (PHENIICS)
Spécialité de doctorat: physique des particules

Thèse présentée et soutenue à Palaiseau, le 15.01.2018, par

Kostiantyn SHPAK

Composition du Jury :

Imad LAKTINEH	IPNL, Lyon	Président du Jury
Yan BENHAMMOU	TAU, Tel Aviv	Rapporteur
Bertrand LAFORGE	LPNHE, Paris	Rapporteur
Henri VIDEAU	LLR, Palaiseau	Examinateur
Dirk ZERWAS	LAL, Orsay	Examinateur
Vladislav BALAGURA	LLR, Palaiseau	Directeur de thèse

Thèse de doctorat

Titre : Développement et optimisation d'un calorimètre électromagnétique en silicium et tungstène hautement granulaire pour le Collisionneur Linéaire International

Mots clés : ILC, SiW ECAL, haute granularité, prototype technologique, PFA, séparation des particules

Résumé : L'un des détecteurs les plus complexes dans les futurs calorimètres du Collisionneur Linéaire Internationale (ILC) est l'ECAL (calorimètre électromagnétique). Cette thèse couvre trois aspects à chaque niveau expérimental:

- Les performances en termes de pouvoir de séparation des particules entre gerbes électromagnétiques, et électromagnétique−hadronique à l'aide de 3 algorithmes de reconstruction de flux de particule (PFA): Arbor, Garlic, et Pandora. Les prototypes physiques CALICE, avec une granularité de 10×10 mm^2 pour le ECAL, et des simulations ILD, avec des cellules de 5×5 mm^2 ou 2.5×2.5 mm^2, ont été utilisés.
- La caractérisation du prototype technologique du SiW-ECAL: la linéarité de réponse en énergie de la puce frontale a été déterminée a l'aide de signaux injectés; la réponse à des faisceaux de haute énergie validée lors d'une campagne au CERN en 2015, qui a également permis de quantifier les événements "carrés" dus aux couplages capacitifs du capteur et d'effectuer une analyse de la dimension fractale des gerbes.
- La mesure de précision de la masse du W dans le canal de désintégration hadronique à l'ILC a été estimée en utilisant des événements simulés en géométrie ILD reconstruits avec PFA Pandora. Une précision statistique de 3 MeV pour 1 million d'événements collectés (100 fb^{-1}) est atteignable.

Une décision sur l'ILC est attendu en 2018.

Title : Development and optimization of the highly granular silicon-tungsten electromagnetic calorimeter for the International Linear Collider

Keywords : ILC, SiW ECAL, high granularity, technological prototype, PFA, particle separation

Abstract : Electromagnetic calorimeter (ECAL) is one of the most complex sub-detectors among all future calorimeters of the International Linear Collider (ILC). The current thesis is devoted to three independent tasks performed within CALICE/ILD SiW ECAL project:

- Electromagnetic-electromagnetic and electromagnetic-hadron shower separation study using three Particle Flow Algorithms (PFA) programs (Pandora, Garlic and Arbor) in CALICE physics prototypes with 10×10 mm^2 ECAL granularity and ILD simulations using standard ECAL with 5×5 mm^2 and finer 2.5×2.5 mm^2 cells.
- Recent SiW ECAL technological prototype tests. It comprises the results obtained with charge injection in the front end chip, and 2015 test beam data analysis including debugging of the data stream, square-events and fractal dimension analysis.
- Precision measurement of the W mass in hadronic decay channel at ILC using events simulated in ILD geometry and reconstructed with Pandora PFA. Statistical precision of 3 MeV is achievable after 1 million of W events (100 fb^{-1}).

Final decision on ILC is expected in 2018.

Résumé

Développement et optimisation d'un calorimètre électromagnétique en silicium et tungstène hautement granulaire pour le Collisionneur Linéaire International

L'un des détecteurs les plus complexes dans les futurs calorimètres du Collisionneur Linéaire Internationale (ILC) est l'ECAL (calorimètre électromagnétique). Cette thèse couvre trois aspects à chaque niveau expérimental :

- Les performances en termes de pouvoir de séparation des particules entre gerbes électromagnétiques, et électromagnétique-hadronique à l'aide de 3 algorithmes de reconstruction de flux de particule (PFA) : Arbor, Garlic, et Pandora. Les prototypes physiques CALICE, avec une granularité de 10×10 mm^2 pour le ECAL, et des simulations ILD, avec des cellules de 5×5 mm^2 ou 2.5×2.5 mm^2, ont été utilisés.
- La caractérisation du prototype technologique du SiW-ECAL : la linéarité de réponse en énergie de la puce frontale a été déterminée a l'aide de signaux injectés ; la réponse à des faisceaux de haute énergie validée lors d'une campagne au CERN en 2015, qui a également permis de quantifier les événements "carrés" dus aux couplages capacitifs du capteur et d'effectuer une analyse de la dimension fractale des gerbes.
- La mesure de précision de la masse du W dans le canal de désintégration hadronique à l'ILC a été estimée en utilisant des événements simulés en géométrie ILD reconstruits avec PFA Pandora. Une précision statistique de 3 MeV pour 1 million d'événements collectés (100 fb^{-1}) est atteignable.

Une décision sur l'ILC est attendu en 2018.

Mots-clés : ILC, SiW ECAL, haute granularité, prototype technologique, PFA, séparation des particules

Abstract

Development and optimization of the highly granular silicon-tungsten electromagnetic calorimeter for the International Linear Collider

Electromagnetic calorimeter (ECAL) is one of the most complex sub-detectors among all future calorimeters of the International Linear Collider (ILC). The current thesis is devoted to three independent tasks performed within CALICE/ILD SiW ECAL project:

- Electromagnetic-electromagnetic and electromagnetic-hadron shower separation study using three Particle Flow Algorithms (PFA) programs (Pandora, Garlic and Arbor) in CALICE physics prototypes with 10×10 mm^2 ECAL granularity and ILD simulations using standard ECAL with 5×5 mm^2 and finer 2.5×2.5 mm^2 cells.
- Recent SiW ECAL technological prototype tests. It comprises the results obtained with charge injection in the front end chip, and 2015 test beam data analysis including debugging of the data stream, square-events and fractal dimension analysis.
- Precision measurement of the W mass in hadronic decay channel at ILC using events simulated in ILD geometry and reconstructed with Pandora PFA. Statistical precision of 3 MeV is achievable after 1 million of W events (100 fb^{-1}).

Final decision on ILC is expected in 2018.

Key words: ILC, SiW ECAL, high granularity, technological prototype, PFA, particle separation

Acknowledgements

I would like to express my heartfelt gratitude to my PhD supervisor, Vladislav Balagura, for the many learning opportunities. Through your support and guidance, I have honed skills that I could apply in the real world beyond ILC and experimental particle physics.

My deepest gratitude to Vincent Boudry and Jean-Claude Brient for their notable support for the completion of my PhD. Special thanks to Henri Videau, cheers to your long conversations on non-physical and somehow physics-related topics in French, you helped me a lot to improve my French.

My humble thanks to the people who helped to become proficient with ILC software: Trong Hieu Tran, Daniel Jeans, Yacine Haddad, Manqi Ruan.

Special thanks to my LLR friends and colleagues, Stanislav, Floriana, Jonas, Iurii, Dan, Frédéric, Jérome, Artur and others for the time spent together in and out of LLR.

Also, I would like to thank LLR and École polytechnique for the support during 5.5 years that I was doing my Master and PhD.

Contents

Introduction

The High Energy Physics has started in the very end of 19th – beginning of 20th centuries from the discoveries of the electron in the cathode rays, the photon in the photo-effect, uranium decays and as X- and gamma-rays, atomic nucleus by the deflection at very high angles of alpha-particles bombarding a thin foil, the proton by identifying it with the hydrogen ion etc. In the middle of 20th century the exploration of the fundamental constituents of matter continued at the electromagnetic accelerators of particles. After more than one hundred years of intense experimental and theoretical efforts, we now know that the electromagnetic and the weak interactions are united and that the hadrons are made of quarks "glued" together by the strong interaction. We have a model surprisingly well describing all observed experimental effects which can be calculated in the theory. It is called the Standard Model and includes as elementary particles 6 leptons (e^-, ν_e, μ^-, ν_μ, τ^-, ν_τ), 6 quarks (u, d, c, s, t, b), their antiparticles, gauge bosons (γ, Z, $W^\pm$, g) of $\mathrm{U}(1)\times\mathrm{SU}(2)_L\times\mathrm{SU}(3)$ gauge interactions and the Higgs boson. The latter was predicted in a series of papers in 1964 [1–4] but was discovered only in 2012 at Large Hadron Collider in CERN [5,6] and became the last necessary ingredient of the Standard Model.

In spite of the triumph of the Standard Model which passed all stringent experimental tests made up to now, we know that it is not a complete theory. It does not include gravitation, has too many free parameters (19 even without the parameters describing massive oscillating neutrinos). There is a hierarchy problem: why the Higgs mass is so "low" and is not disturbed by potentially huge quantum loop corrections. Another arguments about the incompleteness of the Standard Model come from the cosmology. The Standard Model can not explain the composition of the cold dark matter and we also do not know what is the origin of the dark energy. Another mystery is the cosmic inflation which is probably required to have the isotropy and homogeneity of the visible universe over large distances.

pp ($p\bar{p}$) and e^+e^- accelerators

As the history of the High Energy Physics shows, many discoveries requiring the record energies have been made at hadron colliders. The top quark discovery at Tevatron ($p\bar{p}$ at $\sqrt{s} = 1.8$ TeV) in 1994–1995 is one recent example. It is becoming more and more difficult to increase the energies of the e^+e^- circular colliders because of the synchrotron radiation. From the relativistically generalized Larmor's formula it follows that the power loss of the particle in instantaneous circular motion is given by

$$P = \frac{e^2 c \beta^4 \gamma^4}{6\pi\epsilon_0 R^2},$$

where e, c and ϵ_0 are the proton charge, the speed of light and the electric constant, β, γ and R are relativistic parameters of the particle and the radius. This power should continuously be compensated by the accelerating systems. The difficulty to achieve the highest energies at e^+e^- collider is demonstrated by the Higgs boson hunting at LEP. In spite of some signs of the signal it was extremely difficult to reach the energies sufficient for the Higgs boson discovery. In the end it was decided to stop LEP and proceed with LHC, where the discovery was finally made.

Since the Higgs boson mass is measured to be relatively light, it is still possible to build the circular e^+e^- Higgs "factory" if its radius will be enlarged compared to LEP. The projects FCCee at CERN [7,8] and CEPC [9] in China study the future potential and physics cases of such machines. After the end of the e^+e^- program, such a collider can be transformed into a pp machine with the record energy. The other possibility is a linear collider, ILC [10–14], which is a topic of this thesis.

The big advantage of e^+e^- collider is much better signal over background ratio compared to the proton machine contaminated by strong QCD backgrounds. In addition, at LHC, to compensate a very small probability to create "interesting" particles in comparison with an overwhelming majority of minimum bias events, for example, CMS and ATLAS detectors were initially designed for the pile-up rates of 25, while after the HL-LHC upgrade it is going to be increased to 140–200. Another complexity of the hadron machine is that one needs to take into account the parton distribution functions (PDFs) in all calculations. All this limits the precision of the measurements.

Contrary to hadron machines, e^+e^- colliders have much cleaner signals and are much better suitable for the precision measurements. At "clean" ILC one will be able not only to measure the Higgs couplings at an excellent precision but also to apply the powerful missing momentum techniques eg. for $e^+e^- \to ZH$ reaction to "see" otherwise invisible Higgs decay channels. This should help to determine the Higgs branching fractions in a model independent way.

A similar interplay between the roles of hadron and e^+e^- colliders can be found in the history of $W^\pm$ and Z^0 bosons. They have been discovered at Sp$\bar{\text{p}}$S in 1983 in $p\bar{p}$ collisions at $\sqrt{s} = 540$ GeV and explored in detail with an excellent precision in e^+e^- reactions at LEP with maximal $\sqrt{s} = 209$ GeV and SLAC SLC.

The technology to build ILC accelerator already exists and is well proven eg. by XFEL machine at DESY [15]. To reduce the length and the cost of the accelerator but still retain much of its rich physics program, the currently proposed energy of ILC is 250 GeV. This should allow the precision measurement of the Higgs coupling constants and constrain the 6-dimensional Effective Field Theory [16, 17]. The possibility of the energy upgrade is also foreseen. If ILC will be approved, HL-LHC and ILC will be complementary in many respects and will drive the future of High Energy Physics.

ILC detector and PFA

For the ILC physics measurements it will be very important to precisely measure the 4-momenta of the hadronic jets. Hadronic decays dominate the decays of the Higgs boson, top quark, $W^\pm$ and Z^0 bosons. This is essentially because the quarks have three colors, and if the decays of the heavy particle to the quarks and leptons of the given generation are equally allowed, the "effective" number of quarks ("red", "green" and "blue") is 3 times larger than the number of leptons.

The pure calorimetric measurement of jets is limited by the measurement of strongly fluctuating hadronic showers. Therefore, the concept of Particle Flow Algorithm (PFA, [18, 19]) with "imaging" high granularity calorimeters have been proposed. Such calorimeters should be capable to separate individual showers from the particles in the jets. If the shower can be associated with the charged track, its energy is taken not from the calorimetric measurement, but from the high precision momentum measurement made by the tracker. More precisely, the momentum is determined from the curvature of the track in the magnetic field. The uncertainty of the tracker measurement typically comes from the hit resolution and the multiple scattering, but compared to the calorimetric uncertainties both can be safely neglected. According to the LEP measurements, charged hadrons carry in average $\sim 62\%$ of the jet energy. The advantage of PFA is that this large fraction, traditionally measured by the hadron calorimeter (HCAL), is now measured with the negligible error by the tracker. Only the remaining 10% of the hadronic energy carried by long-lived neutral hadrons is still measured by HCAL. The electromagnetic component having 27% of energy in average is measured by the electromagnetic calorimeter (ECAL). Finally, due to neutrinos 1.5% of energy in average escapes the detection. Following PFA approach to the jet reconstruction, one can achieve $\delta_E/E = 3-4\%$ energy resolution. This is almost twice better than in the traditional calorimetry. It will allow eg. to distinguish $W^\pm$ and Z^0 bosons decaying to two jets.

The PFA jet energy resolution at ILC will still be limited by the HCAL resolution for the jet energies below ~ 100 GeV. At higher energies the jet becomes strongly collimated due to strong boost. Then, the separation distance between the particles in the jet becomes so small, that the automatic "pattern recognition" can not always distinguish the individual showers. Sometimes, it fails to assign the right energy deposition to the right particle. In PFA terminology this is called "confusion". It degrades the jet energy resolution and dominates at the jet energies above 100 GeV.

This thesis is devoted to the analysis of the performance of ECAL and HCAL of the International Large Detector (ILD, [14, 20]) at ILC with the emphasis on ECAL. In particular, in the main part of the thesis we study how often the PFA programs make such a "confusion" in resolving two close electromagnetic or electromagnetic - hadronic showers. The failure rate is presented as a function of the distance between the showers. The study is performed both in ILD simulation and using the data accumulated by CALICE [21] physical prototypes, as it will be explained in detail later. All three currently existing PFA programs have been used and compared, they are called Pandora, Garlic and Arbor.

Silicon-tungsten ECAL

Since the PFA jet energy resolution is dominated either by HCAL or by the confusion but never by ECAL, the requirements for the ECAL intrinsic resolution are rather moderate at ILD. Essentially, the statistical term $\delta_E/\sqrt{E}$ less than 20% should be sufficient for the jets. Depending on the number of layers and the active detector thickness, the design value is in the range 15–20%. On the other hand, the requirements for the ECAL pattern recognition capabilities are stronger than for HCAL, the granularity should be as good as 5×5 mm^2.

This presents the main technological challenge for the technology. Depending on the chosen ILD size and the number of ECAL layers, the total number of ECAL channels is between 60 − 100 million. Commissioning, calibration and operation of the ECAL of a similar complexity has never been attempted before.

Probably, the most suitable ECAL technology for PFA is a silicon-tungsten sampling calorimeter. The tungsten absorber has a small Moliere radius, $R_M^{(W)} = 0.9327$ cm, this ensures the compactness of the electromagnetic showers and facilitates their separation. It also has a large ratio of nuclear interaction ($\lambda_I^{(W)} = 9.946$ cm) and radiation lengths ($X_0^{(W)} = 0.3504$ cm): $\lambda_I^{(W)}/X_0^{(W)} = 28.38$. This helps to separate hadronic and electromagnetic showers.

Silicon readout is expensive but offers numerous advantages. It is easily segmentable so that its granularity is limited only by the number of channels in the front-end electronics. It is reliable, stable in time, not sensitive to any environmental changes (eg. to the temperature) and can work in the strong magnetic field. The PIN diode sensors are perfectly linear. Since they have no intrinsic gain, when fully depleted, their response depends only on the thickness of the diode. The latter is constant across the pixels with high precision. Therefore, the detector should be easy to calibrate. The spread of the responses to the minimum-ionizing particle (MIP) is typically measured at the level of 5% and is dominated by the variation of the front-end electronics gain. The silicon price for ECAL is less than for the tracker devices, because the ECAL requirements are less stringent. In particular, as simulation shows, up to 10–15% of pixels can be not operational under the condition that they are distributed randomly in ECAL. Up to this level the jet energy resolution is almost unaffected. The price estimation for the ILD SiW ECAL from Hamamatsu in 2014 was about 2.5 EUR / cm^2. Still, this makes ECAL one of the most expensive parts of ILD. The absence of the intrinsic silicon gain also requires very low noise electronics. The modern ASIC front-end chips specially designed for ILD ECAL can deliver the required level of noise and the gain, however. Overall, with the SiW technology one can achieve the lowest level of ECAL systematic errors and obtain an excellent granularity for PFA.

SiW ECAL in CALICE collaboration and the content of this thesis

An alternative ILD ECAL option is based on the scintillator strip detectors readout by the Silicon Photomultipliers (SiPMs) [22]. They are less expensive but also have less granularity, because the SiPM price limits the number of channels. In the silicon ECAL with 5 × 5 mm^2 pixels it is the area which determines the price, the granularity matters less. The proposed scintillator strips have 9 times larger area (5 × 45 mm^2). This is compensated to some extent by alternating the orientations of the strips in ECAL layers and by forming “virtual” pixels of 5 × 5 mm^2 size as “intersections” of perpendicular strips. The energy in such “virtual” pixels is not measured but is estimated offline according to the energy sharing in the perpendicular strips in the adjacent layers.

SiW together with the scintillator-tungsten (ScW) ECALs are being developed within the CALICE collaboration [21]. It focuses on R&D of the calorimetry optimized for the PFA. In addition to two ECAL technologies, there is also an intensive research on three HCAL options. Namely, one with the scintillator tiles readout by SiPMs and two with the Resistive Plate Chambers (RPC). With the former technology the energy deposition in every 30 × 30 mm^2 tile is readout by 12 bit Analog-to-Digital Converters (ADCs). The granularity of RPCs is smaller, 10 × 10 mm^2, but the hits are readout either in the digital (yes or no) form or digitized using only 3 thresholds. The corresponding HCALs are called Analog (AHCAL, scintillator tiles), Digital (DHCAL, first RPC option) and Semi-Digital (SDHCAL, second RPC option).

Five main technologies developed by CALICE require modern, state-of-the-art solutions both in the active detector area and in the electronics. To utilize the synergy between different approaches, CALICE tries to unify the Data Acquisition (DAQ) electronics across different projects. In particular, the front-end ASICs are developed by the French Omega group for all calorimeters. They have different features but contain common blocks. The separate development of 4 different chips (for the scintillator, silicon and two for RPC) would require much more effort. In addition, DAQ electronics contains common cards for the readout: DIF and LDA (for SiW ECAL, LDA was later transformed to GDCC at LLR). They have different firmware and different features, but originally have been developed by the same English group which later left CALICE because of lack of funding for ILC in England.

The first SiW ECAL prototype was built in 2005–2007. It was called “physical” because its main purpose was to prove the physical principles of PFA with beam test data. The prototype had 30 layers of 18×18 cm^2 size with 1×1 cm^2 pixels. It was tested at CERN in 2006–2007 together with another CALICE physical prototype, the scintillator AHCAL. The last tests have been performed at Fermilab together with the DHCAL physical prototype. The results have proven the validity of PFA principles. CALICE has published several papers both related to PFA and to the calorimeter technologies. The accumulated data have even been used for the tuning of the Geant4 hadron interaction models. There was one paper especially important for PFA [26], with the analysis of the separation of the hadronic showers. It has confirmed the limits of the “confusion” in the PFA pattern recognition.

Both CALICE data samples mentioned above are also used in the main part of this thesis, as it will be discussed in detail later. In a way similar to the hadron separation study [26], in this PhD analysis we have determined the probability of the correct separation of two electromagnetic and electromagnetic–hadronic showers. This was done for the first time in CALICE. The method was changed and improved compared to [26], in particular, the probability as a function of the distance between the showers is now obtained as a "continuous" function, instead of a few points representing only the integrals over large areas. At the moment when [26] was written, there was only one PFA program called PandoraPFA. Now we have three programs which can separate the showers, PandoraPFA, Garlic and Arbor, and we study the performance of all three. In addition, we have obtained similar results in ILD simulation, both for the nominal 5×5 mm^2 SiW ECAL granularity and with 4 times smaller pixels of 2.5×2.5 mm^2. Before our study there was some indication that Pandora (the best PFA program) was not fully optimized for the granularity better than 5×5 mm^2. The thing was that the jet energy resolution did not improve (and even slightly degraded, though within statistical uncertainties) with smaller than 5×5 mm^2 pixels. To clarify this issue, we have studied also the granularity of 2.5×2.5 mm^2. Surprisingly, indeed, the performance obtained with this granularity was worse (except for Arbor where it was about the same). This result indicates that there is still some room for the improvements of the current PFA programs, which might affect the future ECAL optimization.

After successful completion of the physics prototype program, CALICE started to build and test the second generation prototypes. The problem with the physical prototypes was that the front-end electronics was placed alongside the detector. Therefore, they were not suitable for ILD. It was a difficult technological challenge to integrate the electronics into the active layer. For that, CALICE started to develop the so called "technological" prototypes. Eventually, they can become detector elements of a future ILD.

Currently, the SiW ECAL collaboration consists of five French laboratories (Omega group, LAL, LPNHE, LPSC and LLR) and two Japanese universities (Tokyo and Kyushu). It studies a wide range of R&D topics on the silicon sensor, the DAQ electronics, the mechanics of ILD ECAL, the cooling and the ECAL integration in ILD. An attempt has been made to describe at least some of these activities in this thesis. There is also a general discussion on the current status and the main results obtained with the new technological prototypes.

The original results obtained during the PhD are described in more detail. They consitute the second part of the thesis. The work in LLR has been started in 2014 from the determination of the characteristics of the SKIROC chip using the charge injection. To ensure precise measurements and to avoid any collective effects, every time the charge was injected only into one channel. The calibration was performed automatically at 11 input voltages for 256 channels (in 4 SKIROC chips) in each of three PCBs of the first technological prototype. This allowed to measure with high statistics the dynamic range of the chip, the variation of the gains and the spread of the saturation curves. In this and later tests of the technological prototype, a special attention has been paid to the debugging of the data stream used in the analysis. There are still inconsistencies in the data integrity in about 1% of events or less, unfortunately of not yet identified nature. Several bad channels have been found sharing the same position on several PCBs, which probably indicated that their larger pick-up noise was associated to the PCB routing. Another attempt has been made to debug the effect discovered in November 2015 beam tests at SPS, CERN, when the chip positioned in the core of $\sim$100 GeV electromagnetic shower often auto-triggered not at the right moment but in the next ILC bunch crossing which was retarded according to the current firmware settings by 400 nsec.

The silicon sensor is surrounded by the so-called guard ring. It is needed to reduce the high voltage gradient at the periphery, this helps to reduce the dark currents and to increase the high voltage stability of the sensor. The typical thickness of the guard ring is $\sim$500 μm, therefore, it can be connected to a fixed potential (somewhere between the ground and HV) only with the wire bonding. The latter is not required, however, for the connection of other parts of the sensor. The pixels are sufficiently large to be glued by the conductive epoxy. Therefore, to simplify the assembly process and to avoid the wire bonding, both in the physical and the technological prototypes the guard ring was not connected and its potential was left floating. This is sufficient for a good HV stability. However, there is one drawback in this solution. If a large energy deposition occurs close to the guard ring, through the capacitive coupling between the guard ring and the peripheral pixels it can trigger all boundary pixels. Many of such events have been visible in the physical prototype: they had little energy deposited inside the squared sensor, but all or many peripheral cells were fired. Such events are called "square" events. To reduce this effect, the design of the guard rings in the technological prototype has been improved. The guard ring is now segmented into two or even four independent parts. This reduces the capacitive coupling. Hamamtsu company also proposed and produced the sensors without the guard ring, their exact design is kept secret. The measurement of the probability of such "square" events in the new technological prototype with the improved guard rings have been one of the tasks during the PhD. This analysis, again using the data from November 2015 beam tests, is included in the thesis. Indeed, it was found that the new sensors considerably suppress "square" events. In the worst case of 150 GeV e^+ beam shooting between four sensors (ie. at their corners) after $8.4X_0$ of the tungsten absorber, the probability was measured at the level of $\leq 4 \cdot 10^{-4}$ which can be acceptable for ILD.

Another interesting study has been made to determine the fractal dimension of the electromagnetic and hadronic showers. The idea to use these two fractal dimensions to distinguish photons and hadrons was proposed by M. Ruan when he was in LLR. It was later published in [23–25] together with the estimates based on Monte Carlo simulation. To my knowledge, the fractal dimension was never determined using the real data. In November 2015 beam tests we had only three working ECAL layers, so we could not measure the full shower. In addition, at this stage of the prototype commissioning it was rather difficult to obtain clean data independent of the detector effects. Nevertheless, using an unprecedented ECAL granularity, a first attempt has been made to estimate the fractal dimension of the shower in the transverse plane at some longitudinal positions along the shower axis.

The results obtained during the PhD have been presented at several scientific conferences and published in their proceedings. The technological prototype was presented at 38th International Conference on High Energy Physics (ICHEP2016) in Chicago [28], at IEEE Nuclear Science Symposium and Medical Imaging Conference (IEEE 2016 NSS/MIC) in Strasbourg [28] and at Intrumentation for Colliding Beam Physics (INSTR 2017) in Novosibirsk [27]. The two shower separation analysis have been reported at the CALICE meetings, at ALCWS 2015 in Tokyo, ECFA LC Workshop 2016 in Santander, at Calorimetry for the High Energy Frontier (CHEF 2017) in Lyon and is published as publicly accessible CALICE Analysis Note CAN-057 [30].

This thesis is composed of the introduction, six chapters, conclusions, two Appendices, and the bibliography. The first chapter gives a brief introduction to the ILC project and describes the main physics measurements and the associated detector constraints. The second chapter is devoted to ILD and its PFA calorimetry. Also, it introduces the ILC simulation and three currently existing PFA programs: Pandora, Garlic, and Arbor. The third chapter gives historic overview of CALICE R&D. The fourth chapter is the main part of the thesis, it describes the shower separation analysis using both CALICE data and ILD simulation. The fifth chapter gives a general overview of the SiW ECAL technological prototype with an emphasis on the results obtained during this PhD which constitute the second part of the thesis: on the charge injection, DAQ debugging, "square" events and the shower fractal dimension. Appendices A.1 (B) contain some additional plots for the Chapters 4 (5).

Following the recommendations of the reviewers of this thesis, we also include here a preliminary analysis of $W^\pm$ mass measurement at ILD performed during last three months of PhD. It is not directly linked to the rest of the thesis but introduces another rather stringent test of PFA performance at ILD. One way to measure the $W^\pm$ mass is to perform the threshold energy scan. It may suffer from low statistics, however, as it was at LEP. The other way is to measure the invariant mass directly in the decay to the quark-antiquark pair. This approach is studied in the thesis. Two jets from W decay are reconstructed using Pandora PFA program. We analyze $e^+e^- \to W^+W^-$ reaction at 250 GeV which has a high cross section, and require that one W decays leptonically. This allows to avoid the color reconnection problem: if both W decay hadronically, the hadron interactions in the final state bias the two jet invariant mass spectrum. For simplicity, we consider only leptonic decays with $e^\pm$ and $\mu^\pm$, since for $\tau^\pm$ there are at least two neutrinos in the final state and this also complicates the analysis and the background suppression. This short term study is limited to the simple estimation of the statistical precision without any attempt to study the systematics or the backgrounds. However, we found a new possibility which was not explored before, to improve the precision. At LEP the similar $W^\pm$ mass measurement was significantly improved by constraining the decay product 4-momenta using the decay kinematics, since the energy of the beams was known with an excellent precision. Due to strong focusing of the beams at ILC and the associated beamstrahlung, the ILC center-of-mass is poorly known. Therefore, up to now it was generally believed that the LEP kinematic constraint can not be used at ILD. We found, however, that with the beam parameters described in ILC Technical Design Report, the beam energy constraint can still be useful. More specifically, we found that it transforms to the correlations between the true – measured mass difference and the total $W^\pm$ energy and momentum. These correlations allow to improve the precision by about a factor of two. The results are presented in the last sixth chapter.

Chapter 1

International Linear Collider

1.1 Brief overview of the ILC layout

Contrary to circular machines, the linear e^+e^- accelerator can collide the bunches only once. In addition, the acceleration to the highest energy should also be performed in “one go”. Since the length of the tunnel significantly contributes to the cost of the accelerator, it can not be very long and the aceleration should be performed with as high energy gradient as possible.

Without considering the plasma acceleration, currently there are two established acceleration technologies: using RF cavities and the electromagnetic field generated by another beam. The latter technology has been developed by the CLIC group at CERN [31–33]. It can provide higher gradients and can reach higher energies, up to a few TeV. It is also more expensive, however. Since the Higgs boson mass is relatively light, ILC adopted another technology based on the superconductive accelerating RF cavities shown in Fig. 1.1. Initially, the proposed ILC energy was 500–1000 GeV, but now it is decided to start from 250 GeV which is still sufficient for the precision measurements eg. in the Higgs sector. This can bring up to 40% cost reduction of the ILC project according to the ICFA statement on Nov 8 2017 concerning 250 GeV ILC operating as a Higgs boson factory. Since the parameters of the ILC at 250 GeV are currently “work in progress”, in the following we shall describe the old parameters of 500 GeV machine from the ILC TDR.

The average energy gradient of RF cavities for 500 GeV is 31.5 MeV. 7400 such cavities of ∼1 m length are required to accelerate $e^\pm$ from 15 to 250 GeV. The total length of the main accelerating linac for e^+ is 11 km and the same for e^-.

The full layout of 500 GeV ILC is shown in Fig. 1.2. It contains the following main subsystems:

- polarized electron and positron beam sources;
- 5 GeV electron and positron damping rings (DR) located in the same tunnel with 3.2 km circumference;
- two ∼15 km long e^+ and e^- beam transport systems from the damping rings to the main linacs;
- two stage e^+ and e^- bunch compressor systems before the injection into the main linacs;
- two 11 km long main linacs;

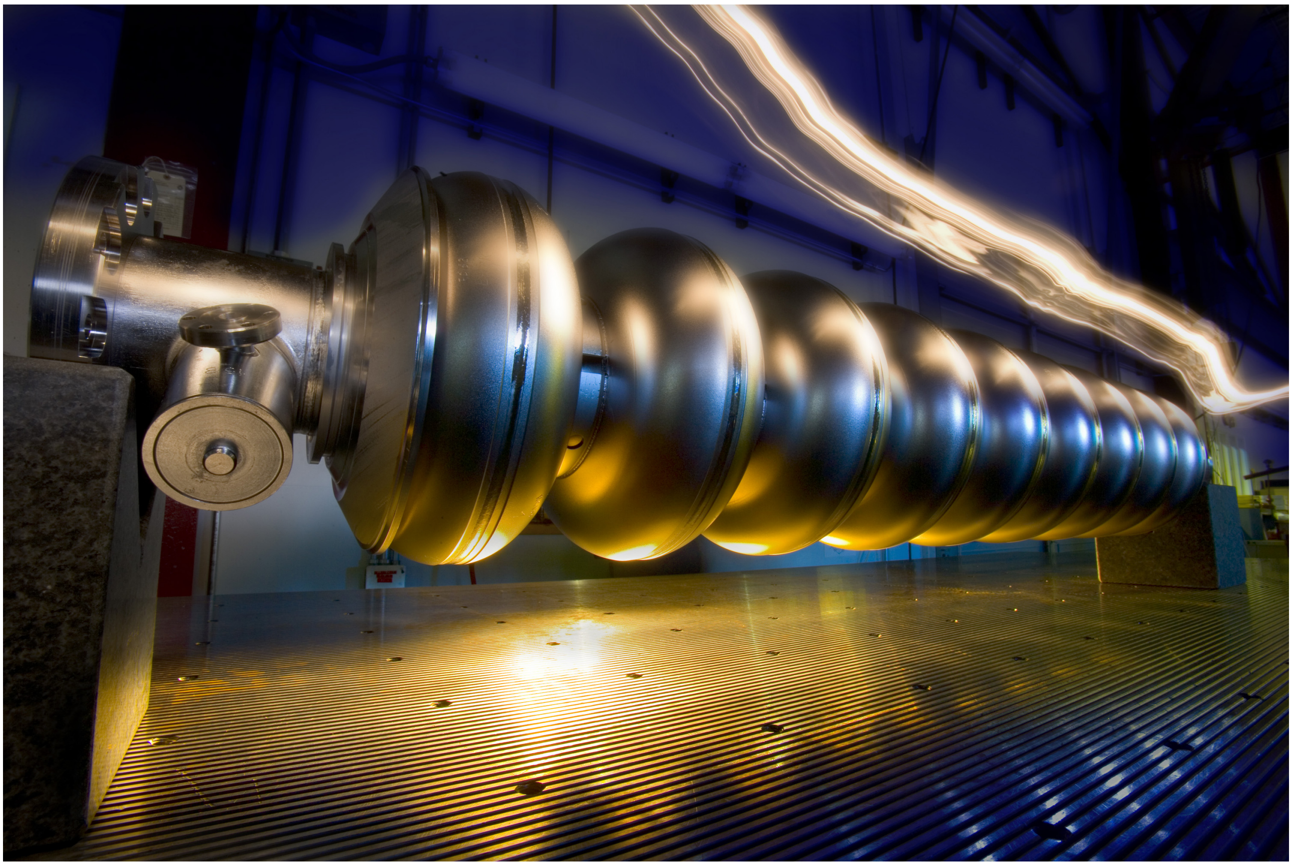


Figure 1.1 – ILC 1.3 MHz niobium superconducting RF cavity.

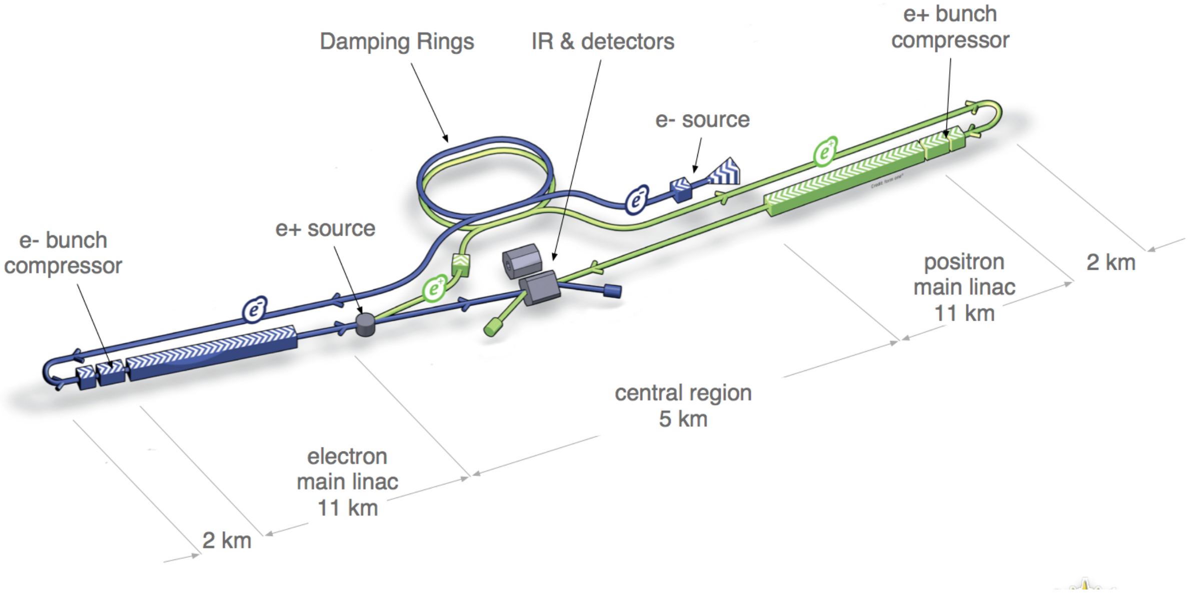


Figure 1.2 – Schematic layout of the ILC, indicating the major subsystems (not to scale).

- two 2.2 km long systems which deliver the e^+ and e^- bunches into collision with 14 mrad crossing angle;
- single interaction point, where there may be one or two detectors. In the latter case they are operated in a "push-pull" mode, ie. at a given moment either one or another detector can take data.

Main machine parameters from the ILC TDR for different operational scenarios are summarized in Table 1.1.

500 GeV collider will occupy a 31 km long underground tunnel. For 250 GeV machine it will be correspondingly shorter. The possibility to extend the tunnel to upgrade ILC to higher energies will be foreseen. The electron and positron sources and the damping rings are located near the interaction region (IR). The damping

			Baseline	500 GeV	machine	1st stage	L-upgrade	E_{CM}-upgrade
Center-of-mass energy	E_{CM}	GeV	250	350	500	250	500	1000
Collision rate	f_{rep}	Hz	5	5	5	5	5	4
Electron linac rate	f_{linac}	Hz	10	5	5	10	5	4
Number of bunches	n_b		1312	1312	1312	1312	2625	2450
Bunch population	N	$\times 10^{10}$	2.0	2.0	2.0	2.0	2.0	1.74
Bunch separation	Δt_b	ns	554	554	554	554	366	366
Pulse current	I_{beam}	mA	5.8	5.8	5.8	5.8	8.8	7.6
Main linac average gradient	G_a	MVm^{-1}	14.7	21.4	31.5	31.5	31.5	38.2
Average total beam power	P_{beam}	MW	5.9	7.3	10.5	5.9	21.0	27.2
Estimated AC power	P_{AC}	MW	122	121	163	129	204	300
RMS bunch length	σ_z	mm	0.3	0.3	0.3	0.3	0.3	0.25
Electron RMS energy spread	$\Delta p/p$	%	0.190	0.158	0.124	0.190	0.124	0.083
Positron RMS energy spread	$\Delta p/p$	%	0.152	0.100	0.070	0.152	0.070	0.043
Electron polarization	P_-	%	80	80	80	80	80	80
Positron polarization	P_+	%	30	30	30	30	30	30
Horizontal emittance	$\gamma\epsilon_x$	μm	10	10	10	10	10	10
Vertical emittance	$\gamma\epsilon_y$	nm	35	35	35	35	35	30
IP horizontal beta function	β_x^*	mm	13.0	16.0	11.0	13.0	11.0	22.6
IP vertical beta function	β_y^*	mm	0.41	0.34	0.48	0.41	0.48	0.25
IP RMS horizontal beam size	σ_x^*	nm	729.0	683.5	474	729	474	481
IP RMS vertical beam size	σ_y^*	nm	7.7	5.9	5.9	7.7	5.9	2.8
Luminosity	L	$\times 1034$ cm^{-2}sec^{-1}	0.75	1.0	1.8	0.75	3.6	3.6
Fraction of luminosity in top 1%	$L_{0.01}/L$		87.1%	77.4%	58.3%	87.1%	58.3%	59.2%
Average energy loss	δ_{BS}		0.97%	1.9%	4.5%	0.97%	4.5%	5.6%
Number of pairs per bunch crossing	N_{pairs}	$\times 10^3$	62.4	93.6	139.0	62.4	139.0	200.5
Total pair energy per bunch crossing	E_{pairs}	TeV	46.5	115.0	344.1	46.5	344.1	1338.0

Table 1.1 – Summary of the ILC $\sqrt{s} = 500$ GeV baseline parameters (columns 4-6). Column 7 presents a set of parameters for the $\sqrt{s} = 250$ GeV starting scenario (half of the original main linac). Parameters for the luminosity and energy upgrades are shown in columns 8 and 9 respectively. From ILC TDR [13].

rings are laterally shifted to avoid any interference with the detector hall. The electron and positron sources, the beam delivery systems and the main linacs share the same tunnel to reduce the excavation works during the construction.

According to the ILC TDR baseline design, the interaction region is shared by two detectors [14], the International Large Detector (ILD) and the Silicon Detector (SiD) working in the "push-pull" mode. While one detector is used for the data taking, the other is moved to the radiationally protected garage position where it can be maintained or upgraded.

1.1.1 Polarized particle sources

Electron source

The polarized electron beam is produced by a laser irradiating a strained Gallium-Arsenic photo-cathode in the DC gun, see Figure 1.3. The unbunched beam produced by the DC gun is transformed into the bunch trains with up to 90% polarization. Two separate DC gun and laser systems provide an excessive amount of electrons. Particles are passing through normal conducting structures for additional bunching and the first stage acceleration to 76 MeV. Then, the e^- beam is accelerated to 5 GeV in a superconducting linac. Before entering the damping rings, the e^- spin vector is rotated by superconducting solenoids and becomes vertical. The subsequent rotation to the longitudinal direction is performed before entering the main linac. A separate superconducting RF is used for bunch energy compression to match the acceptance of the damping rings. Further details on the ILC electron source are presented in [10, 12, 13].

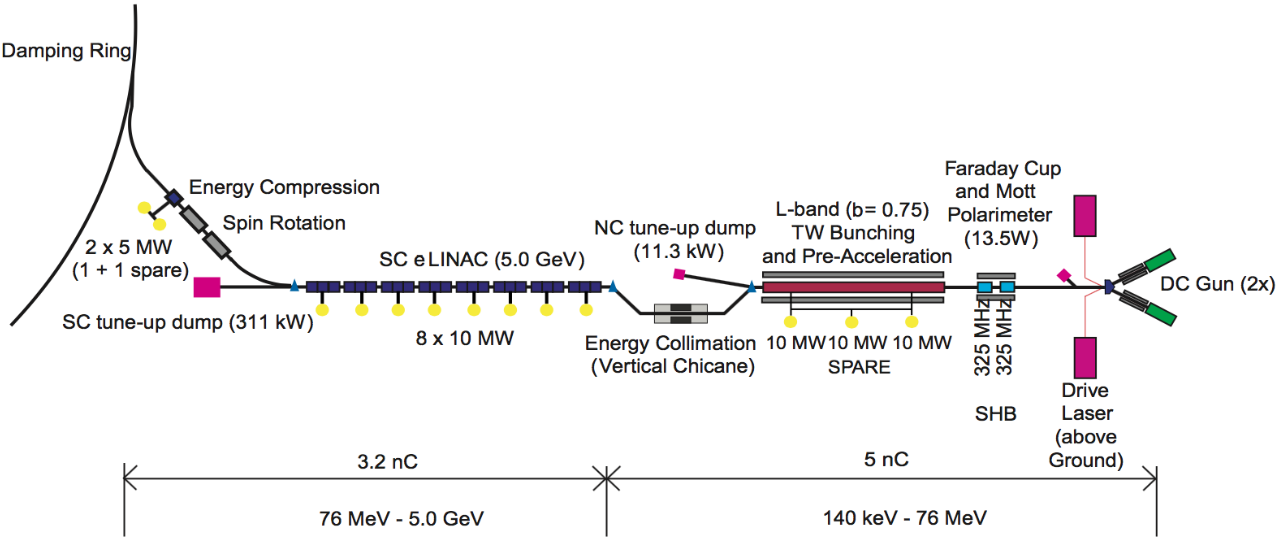


Figure 1.3 – Polarized electron particle source. From ILC TDR [13].

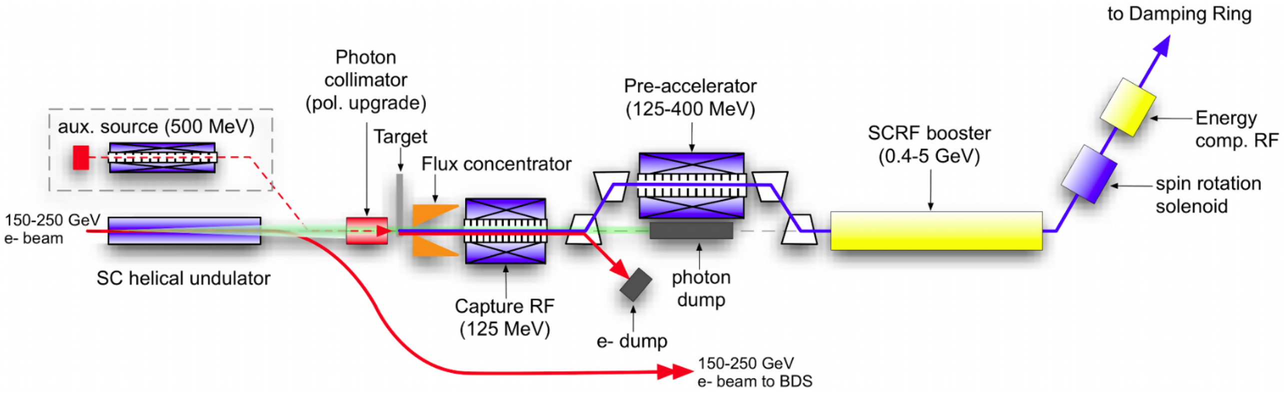


Figure 1.4 – Polarized positron particle source. From ILC TDR [13].

Positron source

The high intensity source of polarized positrons is much more challenging. A schematic view of its major elements is shown in Figure 1.4. Primary electron beam accelerated to ~ 150 GeV in the main linac is used for the positron production [36]. First, the e^- beam passes through the superconducting helical undulator and generates photons with the maximal energy of ~ 30 MeV. The 150 GeV electrons are then separated from photons by a low-emittance-preserving chicane. Photon beam is directed onto a $0.4X_0$ thin Titanium-alloy target, which produces a beam of electron-positron pairs. Electrons and remaining photons are dumped, while the positrons are captured and accelerated to 125 MeV. Then, they are accelerated to 400 MeV in the normal conducting linac with solenoidal focusing. Similarly to the electron beam, after that the positrons are accelerated to 5 GeV in a superconducting linac, their spin is rotated and becomes vertical and the bunch energy spread is compressed before entering the positron damping ring. According to the baseline, this system will produce the polarization of 30%. This is significantly lower than for electrons because of technological challenges. More details on the ILC positron source are given in [10, 12, 13].

1.1.2 Damping rings

Damping rings reduce the large transverse and longitudinal emittances of $e^{\pm}$ beams, ie. the average spread of $e^{\pm}$ in position-and-momentum phase space. This is required to reach the ILC luminosity goal. In the dumping rings the synchrotron radiation in the bending field reduces both the transverse and longitudinal momentum, the latter is then compensated by the precisely controlled acceleration in RF cavities. The net effect is the reduced spread of $e^{\pm}$ momenta in the beams. The radiation dumping occurs during 200 msec period between the bunch trains in 5 Hz ILC scenario. A special attention is paid to the injection and the extraction of the bunches from the rings to avoid emittance growth and the degraded stability of the bunches already present in the damping rings.

In the ILC baseline design e^+ and e^- damping rings have a circumference of 3.2 km and operate at 5 GeV. The rings are installed in the central ILC region one above the other in the same tunnel. Schematically the layout of the damping ring system is shown in Figure 1.5. The synchrotron dumping takes place in two arcs. They are connected to two straight sections each of 712 m length. One such section contains RF cavities, damping wigglers and a phase trombone, in the other there are injection and extraction systems and a circumference adjustment chicane.

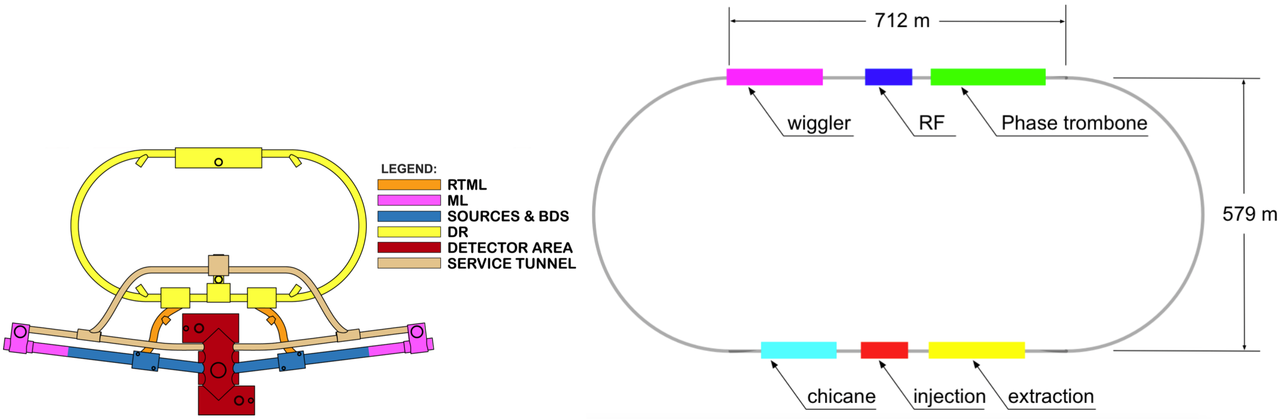


Figure 1.5 – Schematic view of the damping ring system in the central part of ILC.

1.1.3 Main Linacs

After the extraction from the damping rings, both e^+ and e^- beams are transferred to the corresponding main linacs via the Ring To Main Linac (RTML) sections. The electron and positron RTMLs are the longest continuous beam lines in ILC. Schematically the electron RTML system is shown in Figure 1.6, the positron RTML is identical. Each of them has

- ~ 15 km long 5 GeV transport line (ELTL);
- 180° turn-around ring (ETURN), which enables feed-forward beam stabilization;
- spin rotators (ESPIN) to orient the beam polarization to the desired direction;
- two stage bunch compressors (EBC1 & EBC2) which reduce the bunch sizes from several millimeters to a few hundred microns and accelerate $e^{\pm}$ from 5 Gev to 15 GeV.

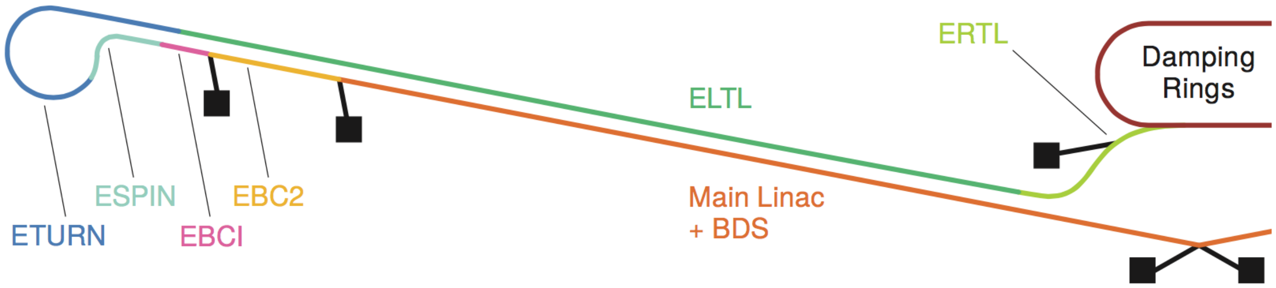


Figure 1.6 – Schematic view of the electron Ring To Main Linac (RTML) section.

The superconductive RF cavities in the main linacs operate at a frequency of 1.3 MHz. They are made of niobium and are kept at a temperature of 2 K. The cavities are assembled into the 12.65 m cryomodules shown in Figure 1.7. The cryomodule contains 9 RF cavities cooled with liquid helium, in every third one there is a quadrupole package. 850 such cryomodules are required in one linac.

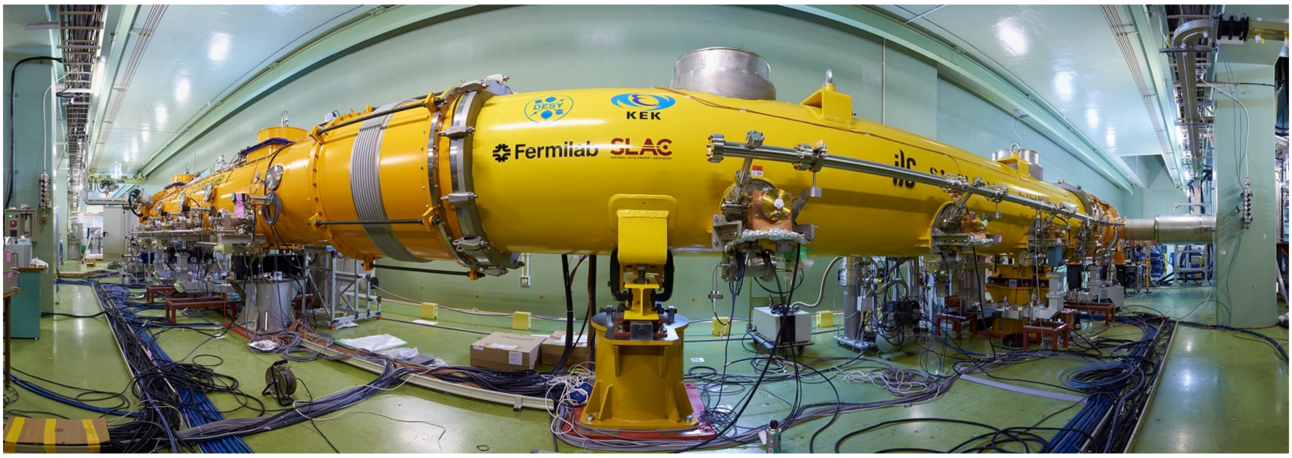


Figure 1.7 – Photo of the ILC cryomodule.

This superconducting RF technology is proven by the European X-Ray Free Electron Laser (XFEL) [15] 3.4 km facility operated in DESY (Hamburg, Germany) since 2017.

1.1.4 Beam delivery systems

After the main linacs, the beams are delivered to the Interaction Point (IP) via 2.2 km long beam delivery system (BDS). The latter focuses the beams to a few hundred nanometers horizontally and only a few nanometers vertically, see Table 1.1. This record focusing is necessary to reach the ILC luminosity goals. The beam crossing angle of 14 mrad simplifies the extraction of not collided particles to the main beam dumps. The BDS should also provide the characterization of the beam parameters such as the energy, the polarization etc. before and after the IP and remove the beam halo from the linac in order to minimize the beam background in the detector.

1.2 ILC physics potential

A detailed description of the ILC physics program is given in ILC TDR [11]. Here, we can only briefly summarize the main ILC physics program which includes the studies of the Higgs boson and the precise measurements of its couplings, the top quark physics, the precision measurements of the Standard Model electroweak parameters and the searches of Beyond the Standard Model (BSM) effects, see Table 1.2.

1.2.1 Higgs sector

The cornerstone of the ILC physics program is the detail study of the Higgs sector, ie. the measurement of the Higgs boson (or bosons) properties and the couplings. The Higgs-strahlung reaction $e^+e^- \to Zh$ provides the unique to the ILC way of measuring in a model independent way the Higgs boson width. This is performed by reconstructing the Z^0 decay products and the Higgs mass recoiling against them, without reconstructing the Higgs boson itself. The standard model can be extended to a 6-dimensional Effective Field Theory (EFT) with

Energy	Reaction	Physics Goal
91 GeV	$e^+e^- \to Z$	ultra-precision electroweak
160 GeV	$e^+e^- \to WW$	ultra-precision W mass
250 GeV	$e^+e^- \to Zh$	precision Higgs couplings
350-400 GeV	$e^+e^- \to t\bar{t}$	top quark mass and couplings
	$e^+e^- \to WW$	precision W couplings
	$e^+e^- \to \nu\bar{\nu}h$	precision Higgs couplings
500 GeV	$e^+e^- \to f\bar{f}$	search for Z'
	$e^+e^- \to t\bar{t}h$	Higgs coupling to top
	$e^+e^- \to Zhh$	Higgs self-coupling
	$e^+e^- \to \tilde{\chi}\tilde{\chi}$	search for supersymmetry
	$e^+e^- \to AH,\ H^+H^-$	search for extended Higgs states
700 - 1000 GeV	$e^+e^- \to \nu\bar{\nu}hh$	Higgs self-coupling
	$e^+e^- \to \nu\bar{\nu}VV$	composite Higgs sector
	$e^+e^- \to \nu\bar{\nu}t\bar{t}$	composite Higgs and top
	$e^+e^- \to \tilde{t}\tilde{t}^*$	search for supersymmetry

Table 1.2 – Main physics reactions of the ILC for different energies including BSM searches. From ILC TDR [11].

all possible operators of 6th dimension. EFT contains terms that modify the $e^+e^- \to Z^*$ vertex and therefore introduce corrections to the $e^+e^- \to Zh$ reaction. Such terms can still be constrained at ILD. In the framework of EFT the analysis becomes more complicated, but one can still achieve the model independence as discussed in [16,17].

The Higgs-strahlung cross section peaks around 250 GeV, see Figure 1.8. This is the reason of selecting this energy for the ILC start. Branching fractions of various Higgs boson decays will allow to measure its couplings to quarks and leptons of different flavors. The Higgs self-coupling requires higher energies, see Table 1.3. Figure 1.9 shows a comparison of the precision measurements that can potentially be achieved at High Luminosity LHC and at ILC running at $\sqrt{s} = 250$ GeV or 500 GeV.

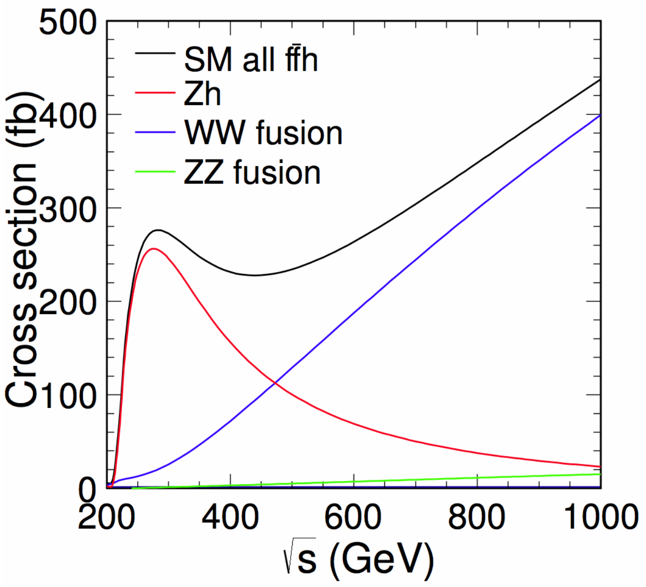


Figure 1.8 – Higgs production cross-section as a function of center-of-mass energy $\sqrt{s}$ for $M_h = 125$ GeV.

1.2.2 Top quark measurements

Another important part of the ILC physics program is the top quark studies. The reaction $e^+e^- \to t\bar{t}$ has a threshold at $\sqrt{s} = 350$ GeV, see Figure 1.10. The ILC threshold scan data will allow to measure the top quark mass and the width with the statistical precision of the order of 30 MeV [38,39]. The top quark is the heaviest particle in the Standard Model, so it has the highest coupling constant to the Higgs boson. The latter can be measured in $e^+e^- \to t\bar{t}h$ reaction which is open above $\sqrt{s} = 500$ GeV.

Topic	Parameter	Accuracy $\Delta X/X$	
Higgs	m_h	0.03%	$\Delta m_h = 35$ MeV, 250 GeV
	Γ_h	1.6%	250 GeV and 500 GeV
	$g(hWW)$	0.24%	
	$g(hZZ)$	0.30%	
	$g(hb\bar{b})$	0.94%	
	$g(hc\bar{c})$	2.5%	
	$g(hgg)$	2.0%	
	$g(h\tau^+\tau^-)$	1.9%	
	$BR(h \to invis.)$	$< 0.30\%$ (95% C.L.)	
	$g(ht\bar{t})$	3.7%	1000 GeV
	$g(hhh)$	26%	
	$g(h\mu^+\mu^-)$	16%	
Top quark	m_t	0.02%	$\Delta m_t = 34$ MeV, threshold scan
	Γ_t	2.4%	
	$\tilde{F}^{\gamma}_{1V}$	0.2%	500 GeV
	$\tilde{F}^{Z}_{1V}$	0.3%	
	$\tilde{F}^{Z}_{1A}$	0.5%	
	$\tilde{F}^{\gamma}_{2V}$	0.3%	
	$\tilde{F}^{Z}_{2V}$	0.6%	
W	m_W	0.004%	$\Delta m_W = 3$ MeV, threshold scan
	g_1	0.16%	500 GeV
	κ_γ	0.03%	
	κ_Z	0.03%	
	λ_γ	0.06%	
	λ_Z	0.07%	
H^0, A^0	m_H, m_A	1.5%	
	tan β	20%	
$\tilde{\chi}^+$	$m(\tilde{\chi}^+)$	1%	
	$m(\tilde{\chi}^0)$	1%	
$\tilde{t}$	$m(\tilde{t})$	1%	
	cos θ_t	0.4%	

Table 1.3 – Main Standard Model physics quantities that will be measured by ILC (taken from ILC TDR [10]).

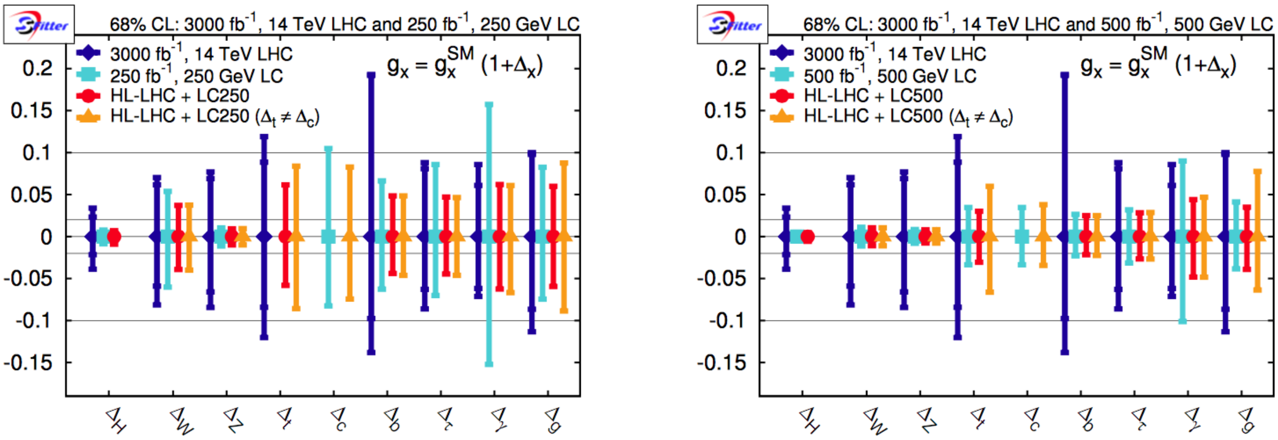


Figure 1.9 – Expected precision for Higgs coupling measurements at the High Luminosity LHC and ILC for $\sqrt{s} =$ 250 (left) and 500 GeV (right) scenarios [37]. The precision of the combined HL-LHC and ILC measurements are also shown.

1.2.3 Searches for the new physics

New particle search is probably the most interesting part of the ILC physics program [11,34]. As it was discussed in the Introduction, we know that the Standard Model is incomplete. ILC energy is much less than at LHC, but due to unique precision and a clean environment, it has better sensitivity than LHC in many BSM searches. The precision measurements of the Higgs boson properties and couplings at ILC are probably the most important

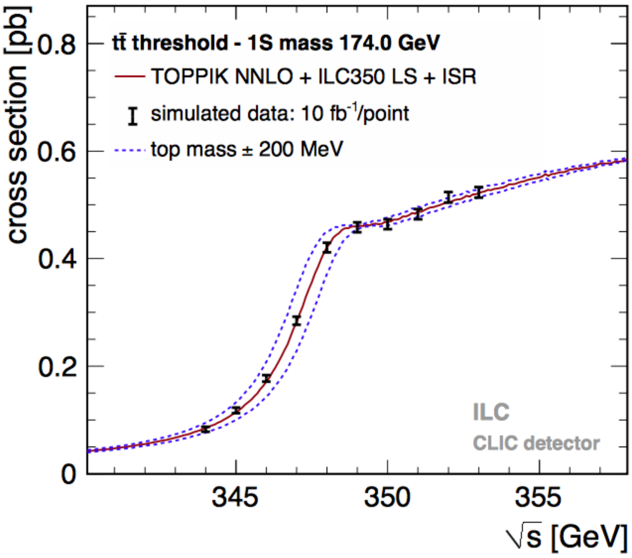


Figure 1.10 – The top quark background subtracted cross section measured at ILC with 10 fb^{-1} luminosity per point [39]. The top quark mass is chosen to be 174 GeV in the simulations (red line), ±200 MeV variation is shown by blue lines.

probe for new physics beyond the Standard Model. For example, new particles insensitive to the Standard Model interactions or hidden from the current experiments due to some other reason, can still couple to the Higgs boson. The deviations in the top sector can also signal new physics.

The BSM can also be discovered at the ILC in the direct searches. Eg. the ILC could be sensitive to the direct pair production of dark matter particles or other particles which are difficult to observe at LHC. e^+e^- annihilation even at 250 GeV can be sensitive to new boson resonances or to quark and lepton substructure.

1.3 The implication of the ILC physics program and machine environment on the ILC detectors design

Planned ILC physics program and machine environment place certain requirements on the detectors. The performances of the multi-purpose ILC detectors should be significantly higher than the current collider detectors, especially in terms of precision. Machine environment of the ILC, contrary to LHC, do not place strict requirements on the radiation hardness of the material or operation under very low temperatures, so advanced designs, and technologies can be realized within the detector. Data taking process will be in parallel with e^+e^- collision process. The most important requirements for the ILC detector design coming from the ILC physics program are summarized in Table 1.4 and briefly described further in this section. The description of the International Large Detector (ILD), one of the two ILC detector concepts and the Particle Flow concept of the event reconstruction, driven by ILC physics requirements, are presented in the next chapter.

1.3.1 Jet energy resolution (JER)

Many ILC e^+e^- collisions will be characterized by two or more jets in the final state. Detector should provide good jet energy resolution performance in order to distinguish $W^{\pm}$, Z and Higgs bosons - intermediate states of the $e^+e^- \rightarrow jets$ reactions, like $e^+e^- \rightarrow Zhh$ and other reactions mentioned in Table 1.4.

Detectors proposed for the ILC should be able to reconstruct the masses of the Z and $W^{\pm}$ vector bosons with a precision close to their natural decay width ($\sigma_m/m \approx \Gamma_{W^{\pm}}/m_{W^{\pm}} \approx \Gamma_Z/m_Z \approx 2.7\%$), see Figure 1.11. It means that in the 150 -350 GeV di-jet energy range ILC detector should provide $\sigma_E/E \leq 30\%/\sqrt{E(\mathrm{GeV})}$, more than two times better than for LEP experiments. This JER requirement places a number of the hardware and software challenges, which should be solved prior the ILC construction.

1.3.2 Tracker momentum resolution

For the realization of the ILC physics program, the precision $\Delta p_t/p_t^2$ of the momentum resolution of the charged particles should be $\sim 5\times10^{-5}$ (GeV/c)$^{-1}$, approximately ten times better the momentum resolution achieved in LEP experiments. High precision momentum measurements are relevant for the Higgs mass measurements from

Physics process	Measured quantity	Critical system	Physics magnitude	Required performance
Zhh $Zh \to q\bar{q}b\bar{b}$ $Zh \to ZWW^*$ $\nu\bar{\nu}W^+W^-$	Triple Higgs coupling Higgs mass $BR(h \to WW^*)$ $\sigma(e^+e^- \to \nu\bar{\nu}W^+W^-)$	tracker, calorimeter	Jet Energy Resolution $\Delta E/E$	3% to 4%
$Zh \to l^+l^-X$ $\mu^+\mu^-(\gamma)$ $Zh + h\nu\bar{\nu} \to \mu^+\mu^-X$	Higgs recoil mass Luminosity weighted E_{CM} $BR(h \to \mu^+\mu^-)$	μ detector, tracker	Charged particle momentum resolution $\Delta p_t/p_t^2$	$5 \times 10^{-5}(\frac{\mathrm{GeV}}{\mathrm{c}})^{-1}$
$Zh, h \to b\bar{b}, c\bar{c}, gg$	Higgs branching fractions, b-quark charge asymmetry	vertex detector	impact parameter	$5\mu\mathrm{m} \oplus \frac{10\mu\mathrm{m}}{p(\mathrm{GeV/c})\sin^{3/2}\theta}$
SUSY, eg. $\tilde{\mu}$ decay	$\tilde{\mu}$ mass	tracker, calorimeter, μ detector	momentum resolution, hermeticity	

Table 1.4 – Main requirements on the ILC detector performances coming from ILC physics program. From ILC TDR [14].

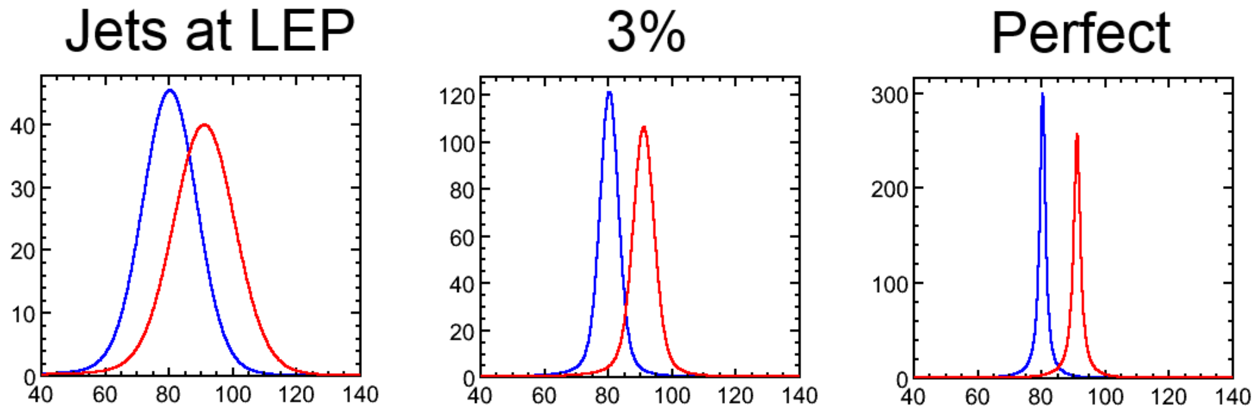


Figure 1.11 – From left to right: separation of the W/Z bosons for the jet resolution of LEP , 3% and perfect resolution.

the Higgs-strahlung reaction ($e^+e^- \to Zh$). The di-leptonic Z decay channel $Z \to \mu^+\mu^-$ allows measurements of the Higgs boson mass independently of the Higgs decay mode.

1.3.3 Vertex detector and impact parameter resolution

The studies of the electroweak symmetry breaking require a precise reconstruction of the Higgs decay particles, in order to distinguish the $h \to c\bar{c}$, $b\bar{b}$ and other $q\bar{q}$ decays. For example, the Higgs self-coupling constant can be measured in $e^+e^- \to Zhh$ reaction, allowed by the Standard Model and characterized by multi-jet final state. Precise quark tagging will significantly reduce the backgrounds. The vertex detector impact parameter resolution is required to be $5\mu\mathrm{m} \oplus 10\mu\mathrm{m}/(p(\mathrm{GeV/c})\sin^{3/2}\theta)$.

1.3.4 Other considerations

Energy and momentum resolutions are not the only requirements that are placed on the detector. Among the other requirements, the most important is detector hermeticity, it should be designed with minimal dead zones. Forward detectors in close to the beam pipe region will allow reconstruction of the invisible decays using the momentum conservation. Suitable pattern recognition techniques could be utilized for efficient particle identification and background suppression.

Chapter 2

Particle Flow approach to calorimetry and ILD detector

The final states of the e^+e^- collisions at ILC can be classified into two groups. The first one consists of events with high momentum leptons, like electrons and muons, which require excellent tracking performance of the detector using a strong magnetic field. Other physics processes will produce final states, that are composed of multiple jets (typically 2-6), frequently accompanied by charged leptons and neutrinos (missing energy). An accurate reconstruction of the jet invariant mass is a crucial point for precision physics measurements. Precise masses are needed for both reconstruction and identification of the events. It means that future ILC detectors should allow separation of the hadronic W and Z decays ($M_W = 80.385 \pm 0.015$ GeV$/c^2$, $M_Z = 91.1876 \pm 0.0021$ GeV$/c^2$) for di-jet final states. The achieved resolution for this measurements should be similar to the natural decay widths of the W and Z bosons ($\Gamma_W = 2.085 \pm 0.042$ GeV$/c^2$, $\Gamma_Z = 2.4952 \pm 0.0023$ GeV$/c^2$). It means that for the typical ILC jets of 50 to 500 GeV the target resolution σ_E/E should be smaller than $\sim 3.5\%$. This goal is unlikely to be achieved with a traditional approach to calorimetry [18]. A particle flow approach was proposed [19]. It imposes some constraints on the detector geometry and the reconstruction tools. These issues are described in this Chapter.

2.1 Classical vs Particle Flow approaches to calorimetry

Detailed jet fragmentation was measured at LEP (Large Electron-Positron Collider), which provides detailed information on jet particle composition, especially, the ones coming from Z decay [43]. In a typical jet, approximately 62% of energy is carried by charged particles (mainly hadrons), 27% is carried by photons, 10% by long-lived neutral hadrons and 1.5% by neutrinos. It is expected that ILC jets structure will be the same or very similar.

2.1.1 Classical calorimetry

According to the classical approach to calorimetry, jet energy will be measured as the sum of the energies deposited in ECAL and HCAL, see Figures 2.1 and 2.2 (left). For a typical ILC jet, it means that $\sim 72\%$ of the energy would be measured in HCAL. Due to the wide energy range of the particles in jets, it is hard to build a calorimeter with equal response on hadronic and electromagnetic showers (hardware compensation). Measurements of the hadronic showers are degraded by the significant shower fluctuations and uncertainties in the shower structure. Due to the poor HCAL resolution $\sigma_E/E \approx 55\%/\sqrt{E/GeV}$ the overall ILC jet energy resolution goal of $\sim 3.5\%$ to distinguish W and Z bosons, will be not achievable.

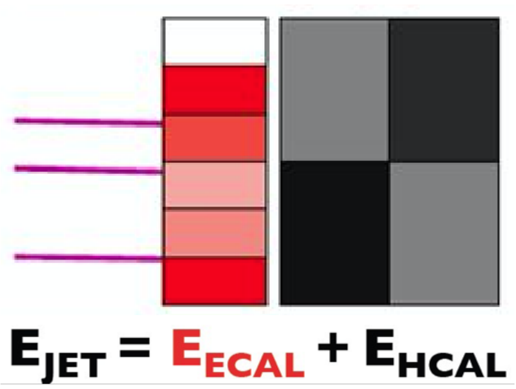


Figure 2.1 – Low segmented detectors are suitable only for traditional calorimetry.

2.1.2 Particle Flow calorimetry

In order to reach the ILC resolution goal, a new approach to the calorimetric measurements has been proposed. The best jet energy resolution can be achieved if it is possible to trace and reconstruct every single particle in a jet separately, collecting together the energy deposits left in each subdetector system, as illustrated in Figure 2.2 (right). This calorimetric approach is called Particle Flow. According to this concept, 62% of jet energy (carried by charged particles) can be measured with the tracking system, and the particle momentum can be estimated with extremely high precision using curvature of the tracks in the magnetic field. Charged particles should be properly associated with their contributions in calorimetric subdetectors. After that, the neutral component of the jet energy only should be measured with calorimeters, 27% of the jet energy (photons) should be measured with the ECAL with a typical resolution of $\sim 15\%/\sqrt{E/GeV}$, and only 10% of energy (long-lived neutral hadrons) should be measured with the HCAL with poor resolution characteristics ($\sigma_E/E \approx 55\%/\sqrt{E/GeV}$).

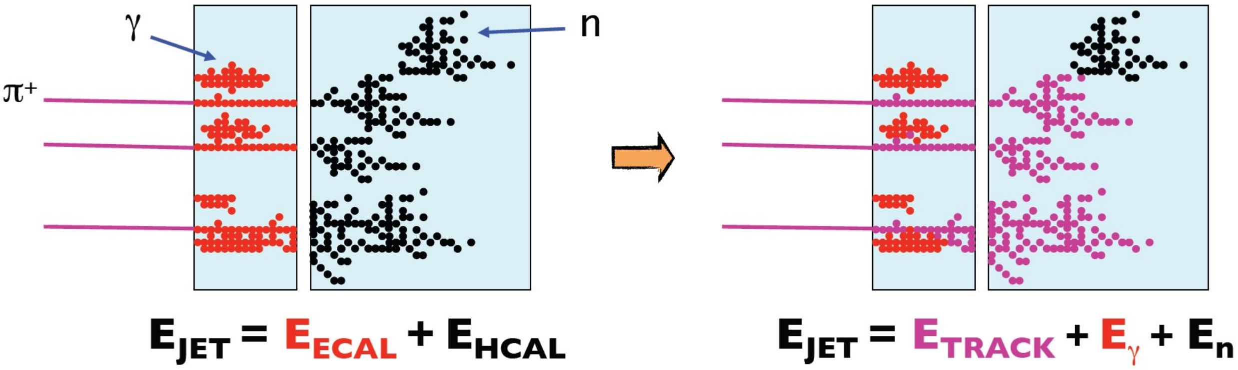


Figure 2.2 – The transition from traditional calorimetry to the fine granular Particle Flow calorimetry.

To implement this idea, each particle in the jet should be clearly distinguishable in the calorimeters and individually reconstructed later. The idea of imaging calorimeters has been proposed [19, 41, 42]. Imaging calorimeters require unprecedented segmentation in both transverse and longitudinal directions to record all shower interaction details. ILC detector is suitable for highly granular calorimetry, since e^+e^- interactions and final states are less complicated than pp. Also, ILC does not have high pile-up, the other major LHC disadvantage. The jet reconstruction in ILC detector needs to go beyond the simple summation of the contributions in ECAL and HCAL, which needs to be replaced by complex, sophisticated reconstruction algorithms. These algorithms are called Particle Flow reconstruction Algorithms (PFA).

2.1.3 Confusion

In practice the performance of PFA is limited by confusion [44], the impossibility to correctly associate detector energy deposits with particles. As a result, some energy deposits will be lost (missing energy) or double counted (energy overestimation). The typical ECAL related confusion comes from the impossibility to separate a photon shower from a charged particle shower correctly: MIP-like, interacting hadrons or electron, see Figure 2.3 (left). If part of the photon shower is merged with the charged particle, the final event energy is underestimated. If the opposite situation happens and part of the charged particle shower is combined with a photon shower, the event energy is overestimated. For the HCAL a common source of confusion is an incorrect distribution of hits between charged hadrons and long-lived neutrals, see Figure 2.3 (central, right). For the charged particles the energy measured with the calorimeter should be compatible with the momentum measured with the tracker. Special software compensation techniques should be applied to estimate the energy of the hadronic showers correctly [45]. A wrong distribution of energy between two neutral particles (photons) will not degrade the JER performance of the detector but can be crucial for some particular studies, like π^0 identification [46].

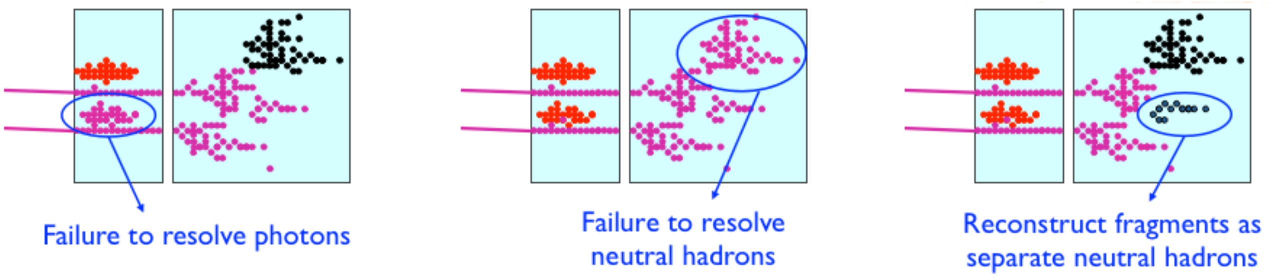


Figure 2.3 – Possible sources of confusion in a fine granular particle flow reconstruction.

Of course several confusions can compensate each other, but in general JER of the detector will be degraded. Confusion depends on the hardware design and software goodness. To minimize the confusion, following requirements are placed on the future ILC detector:

- ILC detector requires a good tracking system with high hermeticity placed into a strong magnetic field to measure the momenta of charged particles in jets;
- Highly granular ECAL and HCAL are needed to record detailed structure of the particle showers;
- Dedicated, sophisticated PFA reconstruction algorithms best optimized for the final detector geometry.

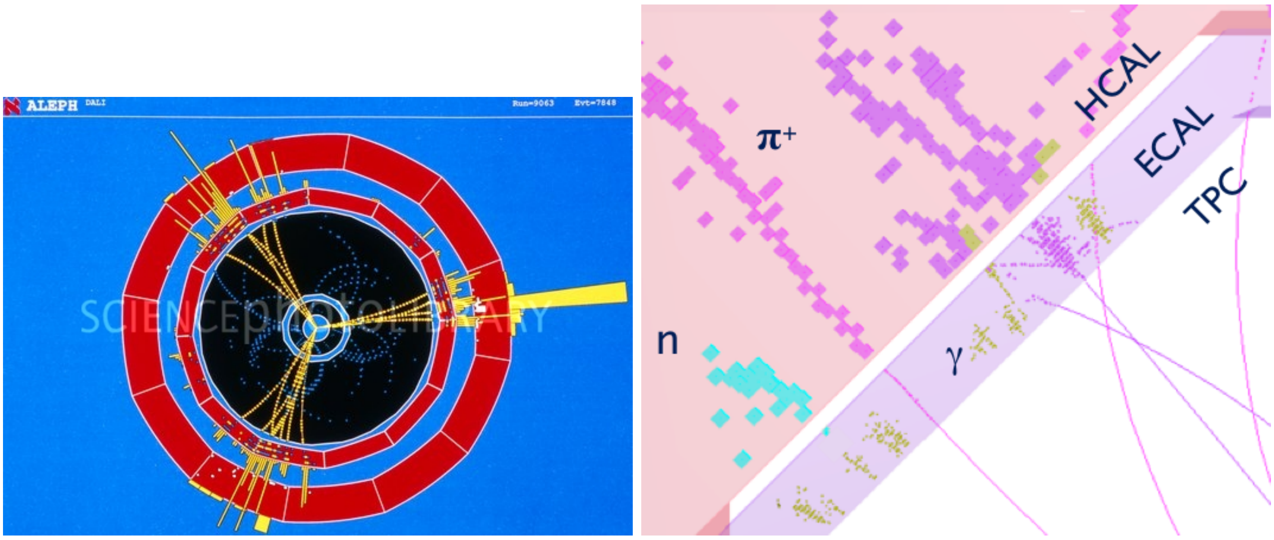


Figure 2.4 – Three-jet e^+e^- event detected at ALEPH (left). Typical example of 250 GeV ILD jet (right).

The International Large Detector (ILD) [14, 20] is one of the proposed detector concepts for the ILC. The ILD detector is designed for the application of PFA, and is composed of a tracker with a strong 3.5 T magnetic field and highly granular ECAL and HCAL. Schematically an example of the 250 GeV jet simulated in ILD geometry is presented in Figure 2.4 (right), where one can see detailed particle structure of the jet. These details were missing in the previous generation of e^+e^- experiments. An example of three-jet e^+e^- event detected at ALEPH [48] is shown in Figure 2.4 (left) [47]. Here, low transversal and longitudinal segmentation in the HCAL is not suitable for PFA application. ILD detector is described in the next section. PFA algorithms designed for the event reconstruction in ILD geometry are presented in section 2.3.2 of this chapter. Figure 2.5 shows an empirical dependence of the confusion term on the jet energy.

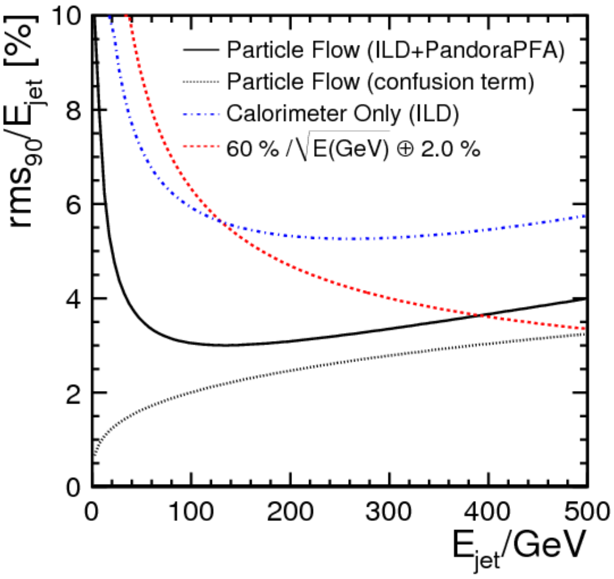


Figure 2.5 – The empirical functional form of the jet energy resolution obtained from Particle Flow calorimetry. Events are simulated in the ILD geometry and reconstructed with PandoraPFA (see later). The estimated contribution from the confusion term is shown with dotted line. From [18].

2.2 The International Large Detector for the International Linear Collider

The ILD detector concept is schematically shown in Figure 2.6. A basic description of the sub-detectors and their working principles is presented in this section. Some sub-detectors have several different technological solutions. All calorimetric sub-detectors are designed within the CALICE collaboration [21]. A short description of the CALICE activities is given in Chapter 3. According to the ILC Technical Design Report (TDR) [14], ILD is a nearly cylindrical multi-purpose detector for the high energy e^+e^- linear collider experiments. The detector is symmetrical with respect to the $x - y$ plane at the Interaction Point (IP). ILD ECAL has an octagonal barrel and two end-cap parts. Recently the ILD ECAL detector design was reviewed in order to validate all the parameters, all achievements are summarized in the SiW ECAL Technical Design Document [49]. Studies on cheaper detector options are also ongoing.

2.2.1 ILD tracking system

The ILD tracking system, see Figure 2.7, is composed of a high precision large volume time projection chamber (TPC), and silicon based vertex and tracking detectors. The ILD tracker together with ECAL and HCAL are placed inside a 3.5 T magnetic field.

ILD vertex system (VXT)

The vertex detector is needed for the reconstruction of the decay vertices of short-lived particles, such as D and B mesons. These decays are essential for the identification of heavy c and b quarks and τ leptons. The proposed six layer VTX detector should provide $\sim 3\mu$m spatial resolution close to the IP. Low material budget (below $0.15\% X_0$/layer) is required to prevent early showering of long-lived particles. The detector will work in extreme radiation conditions: the first layer is located 16 mm from the interaction point (IP), the sixth - at 60 mm. Additional details on VTX design and readout are presented in [14].

ILD silicon tracking system (SIT, SET, FTD)

Silicon tracking systems surround the ILD TPC (key ILD tracking detector, see later) and are called Silicon Envelope for the TPC [50]. It is composed of the Silicon Inner Tracker (SIT), the Silicon External Tracker (SET) and the Forward Tracking Detector (FTD).

Due to the high backgrounds close to the beam line, the TPC should be placed at least 30 cm away from it. Four layers of silicon strips (SIT detector) positioned in the radial gap between VTX and TPC improve the linking efficiency between the VTX and TPC. It improves also the reconstruction of long-lived stable

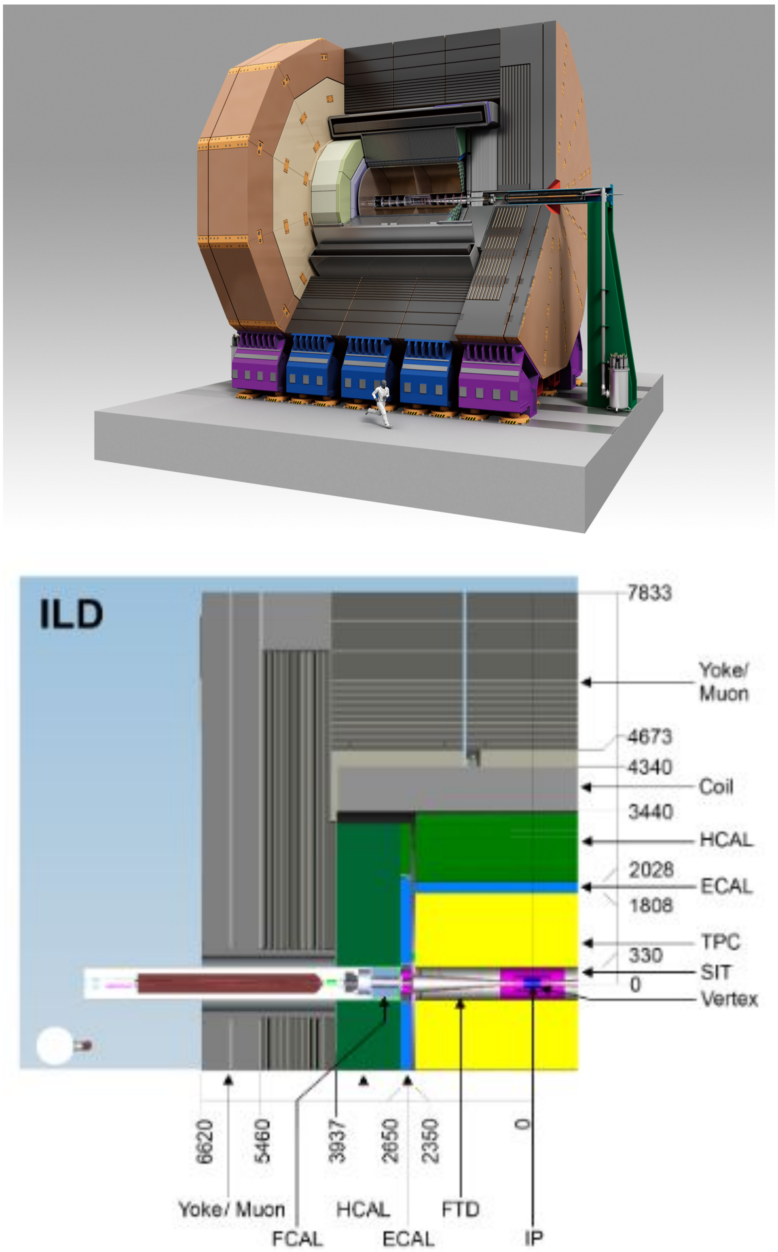


Figure 2.6 – Three-dimensional (top)/ quadrant (bottom) view of the ILD detector concept. Dimensions are in mm.

particles and provides additional track points for the low p_T charged particles (improvement of resolution and reconstruction). The SET detector is placed between the TPC and the barrel part of ECAL, and it is composed of two silicon strip layers, similar to SIT. SET improves the momentum measurements and links TPC tracks to the ECAL and, potentially, can be used for the clustering reconstruction algorithms. Also SET together with SIT provide precise time stamps which can be utilized for the alignment of TPC.

The TPC does not provide any coverage in the very forward detector region. The system of seven silicon tracking disks (FTD) with pixels for the two closest to the IP and strips for the others are installed between the beam pipe and the inner field cage of the TPC. FTD allows tracking of the particles passing at very small angles.

Detailed tracking sub-detector descriptions, leading construction challenges and alternative detector designs can be found in [51,52].

ILD TPC system

The Time Projection Chamber is the central tracking device in ILD. It provides almost continuous tracking with up to 224 points per track in (r, θ, φ) space, see Figure 2.8 (left). An additional advantage of the TPC is a minimum of material in the field cage (contrary to the silicon tracking), see Figure 2.8 (right). This will

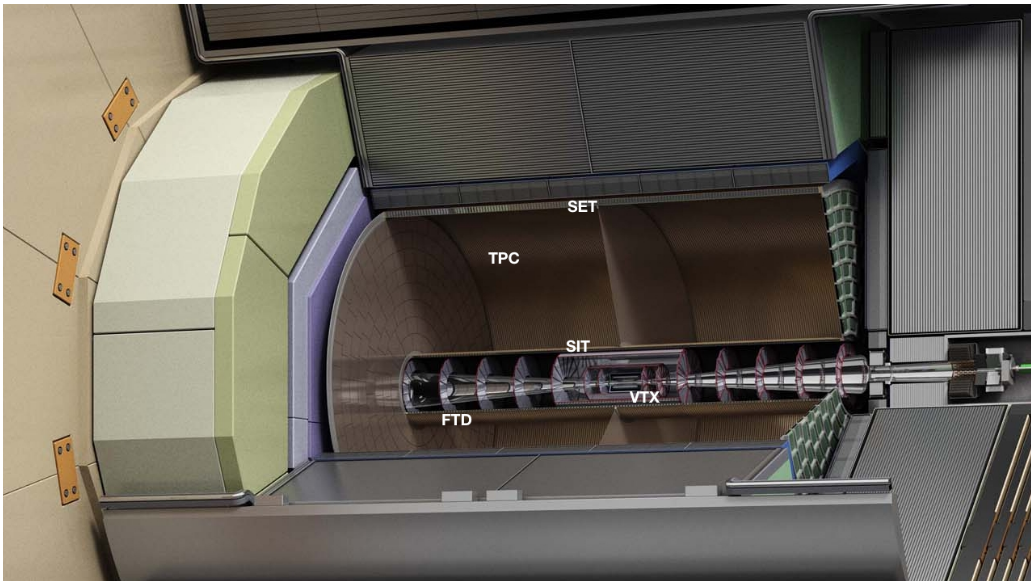


Figure 2.7 – A lateral view of the ILD tracking system. VTX, SIT, SET, FTD & TPC sub-detectors are clearly seen.

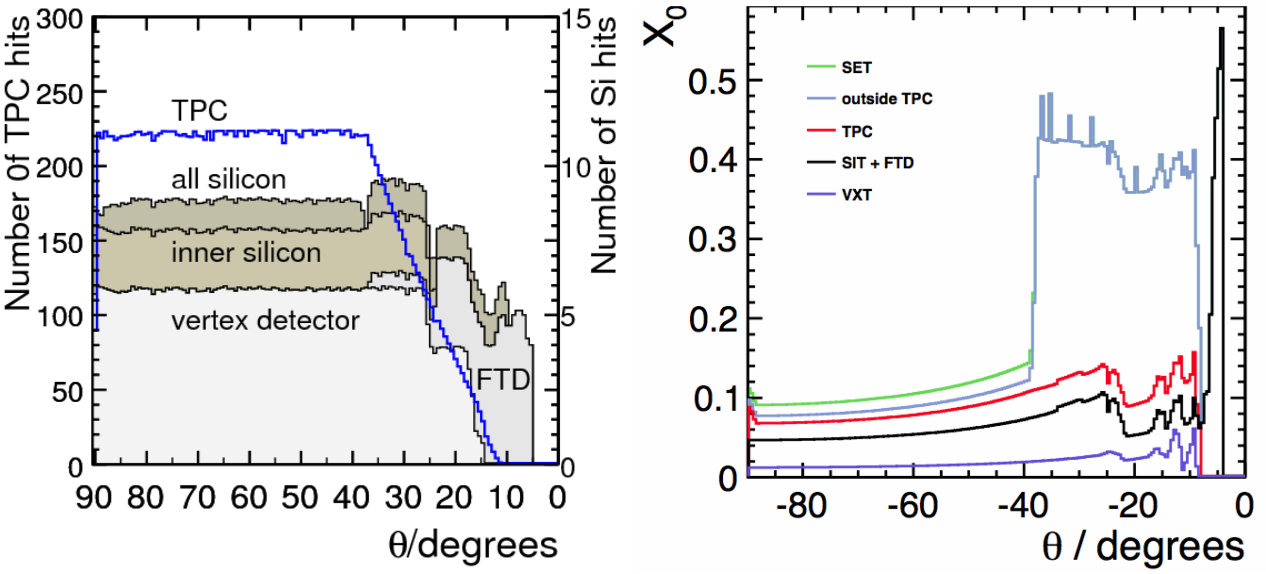


Figure 2.8 – An average number of hits for simulated 100 GeV muons as a function of polar angle θ (left). An average inner tracking detector material in radiation lengths X_0 versus polar angle θ (right). Figures are taken from [14].

allow identifying momenta of the charged particles via the curvature of the tracks. Figure 2.9 shows momentum resolution as a function of transverse momentum of particles. Also, it will be possible to identify back-scattering from the calorimeter, identify kinks, V_0, recover pair production and hadronic interactions. There are active ongoing studies on the TPC endplate, to improve readout electronics and cabling and decrease the material budget in front of the calorimeters. This is essential for the reconstruction of the showers. TPC also allows particle identification based on $\mathrm{d}E/\mathrm{d}x$, and this can be used for the $K^{\pm}\pi^{\pm}$ or $e^{\pm}\pi^{\pm}$ separation, especially for the low energy particles, when calorimeter based identification is difficult due to the small showers. By combining the information from all tracking subdetectors an efficiency greater than 99% is achieved for the track with $p_t > 1$ GeV/c.

All TPC related activities are done within the LCTPC collaboration [53]. Additional details on TPC design and physics performance can be found in [54].

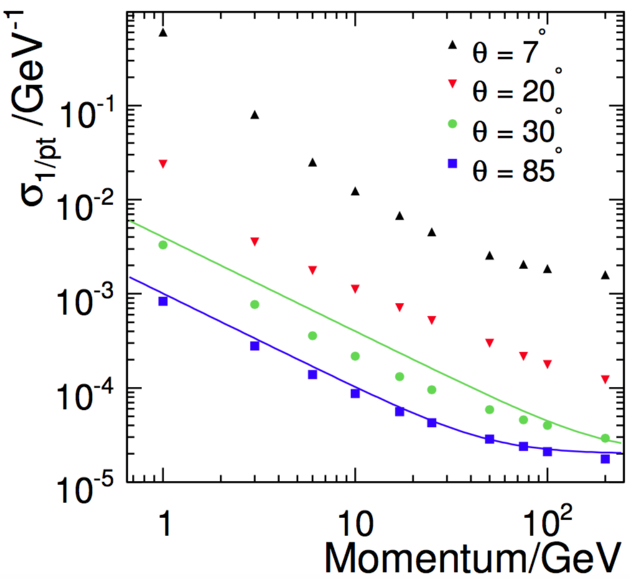


Figure 2.9 – Theoretically expected momentum resolution as a function of the transverse momentum of particles, for tracks with different polar angles: $\theta = 7°$, $20°$, $30°$ and $85°$. From [14].

2.2.2 ILD calorimetric system

Particle Flow approach requires individual reconstruction of particles in the detector, so that charged particles should be clearly separated from the neutrals using tracker and calorimeters. These requirements are achievable with sampling calorimeters with highly segmented active layers. The calorimetric system of ILD consists of a nearly cylindrical barrel system and two large end-caps. Both barrel and endcaps are segmented into electromagnetic and hadronic calorimeters (ECAL and HCAL). The detector hermeticity is achieved with the forward calorimeters, LumiCal and BeamCal, placed close to the beam line.

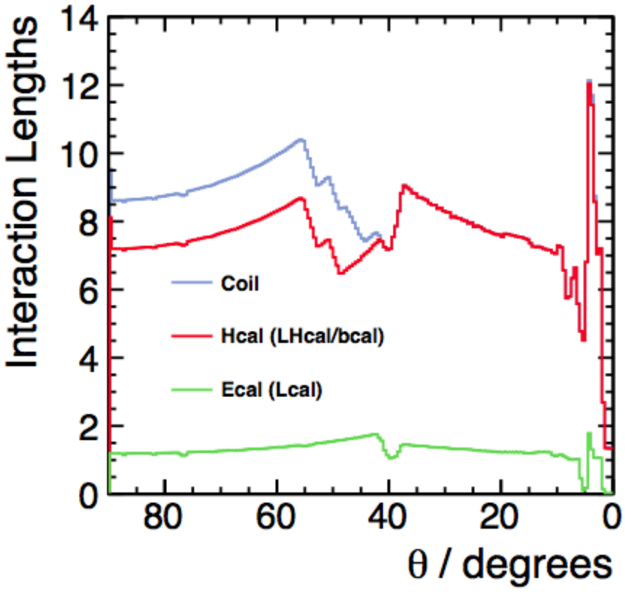


Figure 2.10 – Average total interaction length λ_n of the calorimetric sub-detectors and the coil of the ILD as a function of polar angle θ. From [14].

Electromagnetic calorimeter (ECAL)

For the ECAL, a compact calorimeter design, inspired by TESLA project [57], has been proposed. The calorimeter has a structure interleaving tungsten absorber plates with highly segmented active readout layers. The detector is composed of an octagonal barrel and two endcaps. Two active layer technologies are under development. The first one, the most advanced option, is based on highly segmented silicon with 5×5 mm^2 cells [55]. Another solution relies on 5×45 mm^2 scintillator strips [22], an effective 5×5 mm^2 granularity being achieved by changing the direction of the strips in odd and even layers. Both technologies have been tested within CALICE collaboration, see section 3.2 for additional details. Since this thesis is related to the SiW ECAL, further details on this detector technology are presented.

SiW ECAL is composed of 30 W absorber plates and 30 active silicon layers. Tungsten, as an absorber, is suitable for a compact detector construction. The overall ECAL thickness is ~ 23 cm, corresponding to $\sim 24X_0$, the tungsten radiation length being $X_0^W = 0.3504$ cm. The first 20 W absorber plates are 2.1 mm thick ($\sim 0.4X_0$ per layer), last 10 are 4.2 mm ($\sim 0.8X_0$). This ensures a good containment of the electromagnetic showers in ECAL ($< 2\%$ for 50 GeV e^+). The tungsten nuclear interaction length is $\lambda_I^W = 9.946$ cm, the ratio between interaction and radiational lengths $\lambda_I/X_0 = 28.38$ ensures that most of the hadrons will pass through ECAL just leaving $\mathrm{d}E/\mathrm{d}x$ hits without creating showers. These will facilitate hadron-photon separation in ECAL. The tungsten has a small Moliere radius ($R_M^W = 0.9327$ cm), this ensures compactness of the transverse dimensions of the showers and helps for the two photon separation.

Silicon as an active material is chosen due to the stability of the MIP response, as this facilitates calibration process. During the detector operation, silicon properties will degrade due to the aging, but detector can be easily recalibrated on a cell-per-cell level. For the simulations, a 5×5 mm^2 granularity is chosen, as it was shown that the detector performance improves even for cells smaller than the tungsten Moliere radius, see Figure 2.11. In total ECAL is composed of about 100 million channels. Currently, there is ongoing work on SiW ECAL technological prototype with an embedded very front-end electronics [27–29]. Current FEV10/11 boards are equipped with four 9×9 cm^2 square 325μm thick silicon sensors with 5.5×5.5 mm^2 cells. Silicon studies are ongoing, in particular, on sensor guard ring design to reduce so-called "square"-events (detailed description and related studies are presented in section 5.2.2). Silicon cells are read out with a dedicated 64-channel SKIROC2 ASIC readout chips [56]. More details on SiW ECAL technological prototype, SKIROC2 ASICs, and recent beam test results are presented in Chapter 5.

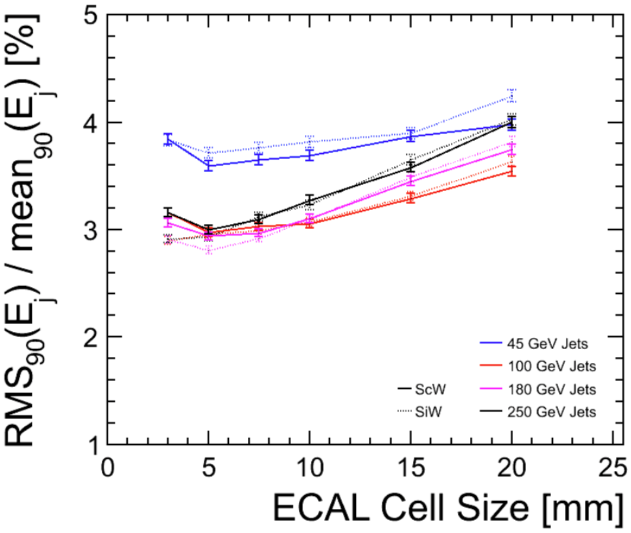


Figure 2.11 – Jet energy resolution as a function of the ECAL cell-size for both silicon and scintillator readouts. From [58].

Mechanical stability and rigidity of the ECAL are achieved with an alveolar mechanical structure. It is composed of carbon fiber reinforced epoxy (CRP) composite structure, which supports every second tungsten absorber plate in a well-defined position. The SiW ECAL mechanical structure is shown schematically on Figure 2.12. Octagonal barrel region is formed by eight inverted trapezoidal detector modules, so-called Videau geometry [20], see Figure 2.13. This design allows to minimize dead areas, but cause some difficulties for the reconstruction. For the end-caps, three different structures are needed. Up to 1.5 m (2.5 m) long barrel (endcap) detector elements, so called slabs, composed of one tungsten and two silicon readout layers are installed into each alveolus. Schematically the double-layer profile is shown in Figure 2.14. 15 double layer slabs are called an alveoli column, and a single ECAL barrel module is composed of 5 columns. The barrel consists of 8 staves made of 5 modules, 40 modules in total. This structure also can be seen as five octagonal wheels, but this does not correspond to assembling process. Each end-cap is subdivided into four quadrants, and each quadrant is subdivided into three modules, 12 modules in total. Each end-cap has a square hole for the beam line, for the hermeticity, the end-caps will be completed with end-cap rings (each is composed of two halves). This results in a very compact technology with minimal dead space. This technology can be readily adopted for the scintillator readout layers.

Due to the high prices of the silicon and tungsten, the SiW ECAL is the most expensive ILD sub-detector [14].

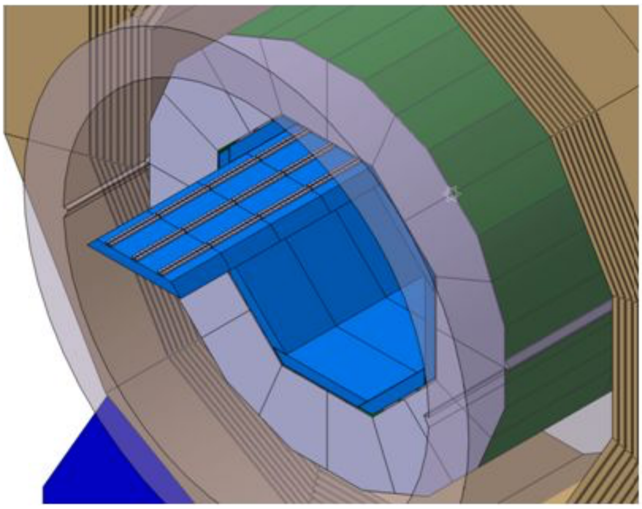
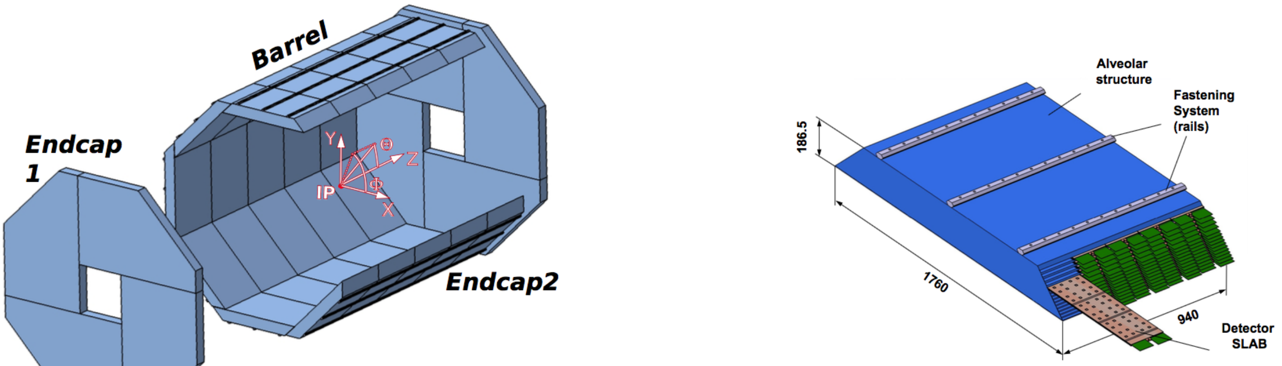


Figure 2.12 – SiW ECAL integration into ILD.



(b)

Figure 2.13 – ECAL barrel and two end-caps (left) and one ECAL barrel module (right).

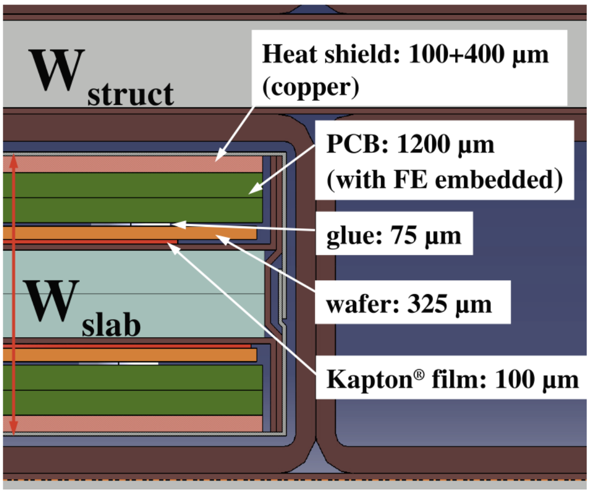


Figure 2.14 – Cross-section of the double-layer silicon slab inside the alveoli.

The main parameters of the ECAL to be optimized are the inner detector radius, its thickness (in X_0), the number of active detector layers, and the segmentation within the layers. Detector performance and its cost are significantly dependent on these parameters [49,58,59]. The main detector price reducing options under study are:

- Reducing the ECAL radius, see Figure 2.15. Detector price is proportional to the silicon surface area. The ECAL resolution is proportional to the radius and magnetic field as $\sigma_E/E \approx R^{-1}B^{-0.3}$. It means that higher magnetic field is needed for the smaller ECAL to prevent degradation of the resolution performance. Also, TPC, HCAL and outer detector radii should be reviewed.

- Reducing the number of detection layers without reducing the overall tungsten thickness. Several detector options are studied (10, 16, 20, 26 and 30 active layers), JER performance is investigated on single 45, 100, 180 250 GeV jet events generated with Geant4. The degradation of JER performance is shown in Figure 2.16.
- Reducing the ECAL radius and number of active layers in parallel.
- Implementing a so-called hybrid ECAL option, in which both silicon and scintillator layers are used for the detector construction. A typical solution is to equip front part of the ECAL with silicon and back part with the scintillator. An example of JER performance is shown in Figure 2.17. Three detector configurations were studied, 20 silicon + 8 scintillator layers, 14 Si + 14 Sc and 8 Si + 20 Sc, with $Z \rightarrow q\bar{q}$ events generated at $\sqrt{s}$=91, 200, 360, 500 GeV.

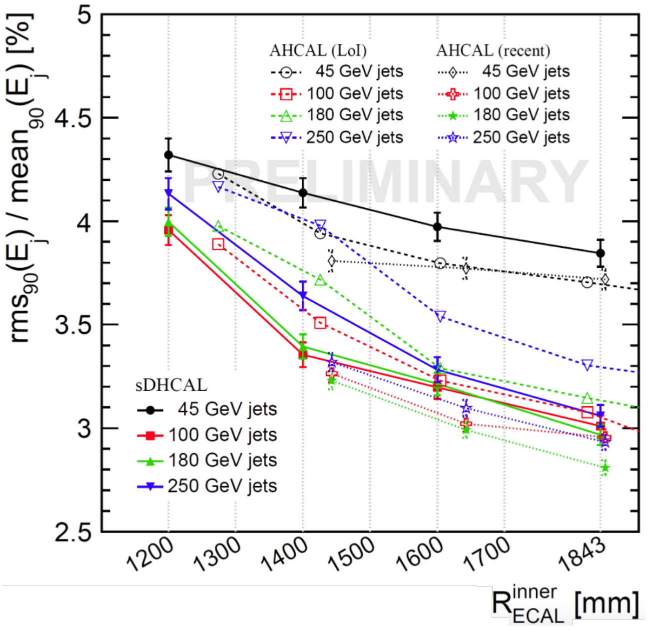


Figure 2.15 – Jet energy resolution in $Z \rightarrow q\bar{q}$ ($q = uds$) events at different center-of-mass energies as a function of the ECAL radius. Dashed (dotted) lines show JER performances for the ILD with ECAL and AHCAL calculated with Pandora v00-03 for ILD LoI [20] (Pandora v00-12 [58]). Solid lines show performance for the ILD with ECAL and SDHCAL with Pandora v00-09 not optimized for SDHCAL [59].

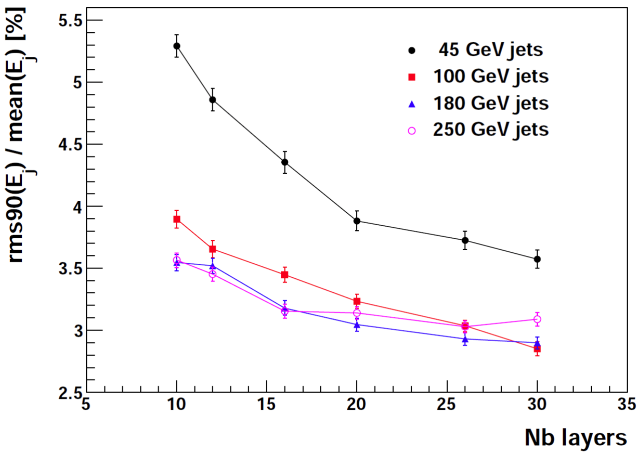


Figure 2.16 – Dependence of the relative jet energy resolution (rms_{90}/E_j) for single jets on the number of ECAL layers for events with $|\cos\theta_{q\bar{q}}| < 0.7$, for the ILD with silicon ECAL. The resolutions are shown for $e^+e^- \rightarrow Z \rightarrow u\bar{u},\ d\bar{d},\ s\bar{s}$ events at $\sqrt{s} = 91$, 200, 360 and 500 GeV. From [59].

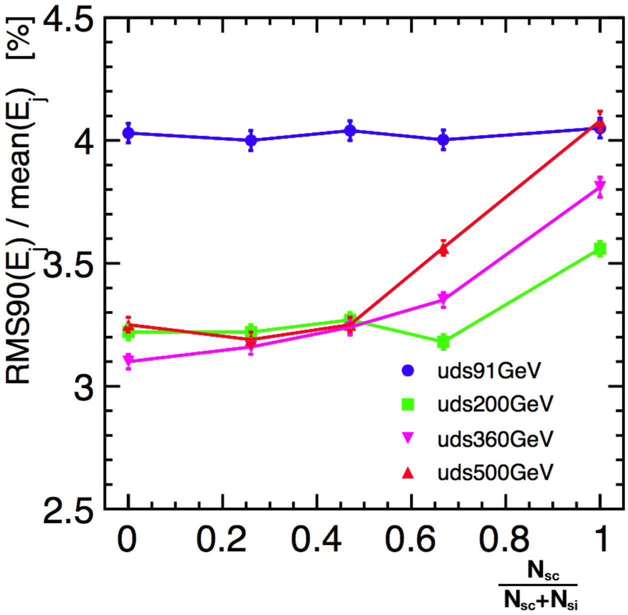


Figure 2.17 – Jet energy resolution in $Z \to q\bar{q}$ ($q = uds$) events at different center-of-mass energies, using a hybrid silicon/ scintillator ECAL. The JER is shown as a function of the fraction of scintillator layers (N_{sc}) in the ECAL, $\frac{N_{sc}}{N_{sc}+N_{si}}$. The total number of ECAL layers $N_{sc} + N_{si}$ is 28. From [14].

Hadronic calorimeter (HCAL)

According to the current proposal, HCAL follows the same octagonal ECAL structure, see Figures 2.18 and 2.19. The HCAL mechanical structure is also supporting the ECAL. Readout cables from tracker and ECAL are passing through the gap between HCAL barrel and endcaps. For the ILD HCAL, two highly segmented active layer technologies are under investigation. The first one, called AHCAL [60], uses 3×3 cm^2 scintillator square pads as an active material and Silicon photomultipliers (SiPM) for analog readout. The second technology, called SDHCAL [61], uses 1×1 cm^2 gaseous Glass Resistive Plate Chambers (GRPC) which provide 2-bit semi-digital readout. According to the ILC TDR, both technologies provide a 48-layer longitudinal segmentation. The Electromagnetic component of the jet is absorbed in the ECAL, so to reduce the costs both technologies are using stainless steel (for the pure iron $\lambda_I^{Fe} = 16.77$ cm, $X_0^{Fe} = 1.757$ cm) as an absorber, which is significantly cheaper than tungsten and is better for HCAL as shown on the AHCAL tests with tungsten []. Both barrel and endcaps of the HCAL are composed of 48 absorber plates of 2 cm each, corresponding to $6\lambda_I$. Figure 2.20 shows HCAL cell-size optimization studies for the AHCAL (left) and SDHCAL (right). It is clearly seen that both AHCAL and SDHCAL show better resolution performance with 1×1 cm^2 cell size. In the case of AHCAL it is possible to use coarser 3×3 cm^2 granularity without significant degradation of JER. Other detailed AHCAL and SDHCAL optimization studies are presented in [62, 63].

Description of the CALICE activities on physics prototypes of the AHCAL and SDHCAL is given in Section 3.3.

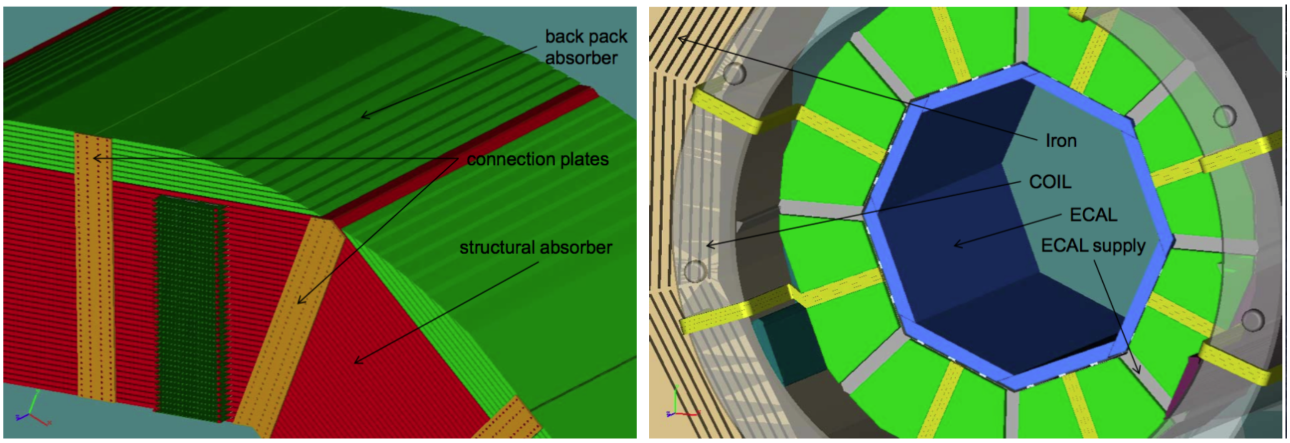


Figure 2.18 – Integration of the AHCAL structure.

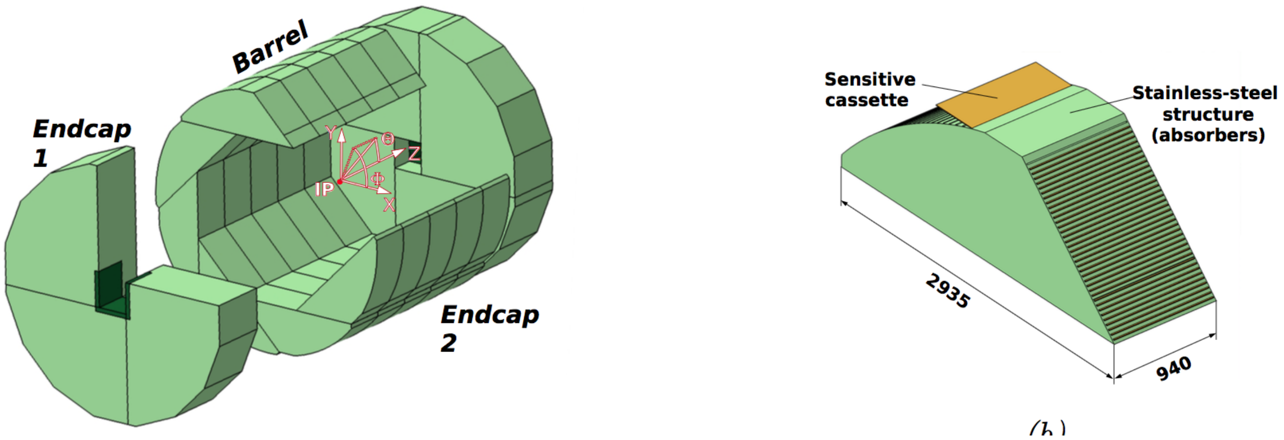


Figure 2.19 – Integration of the SDHCAL structure.

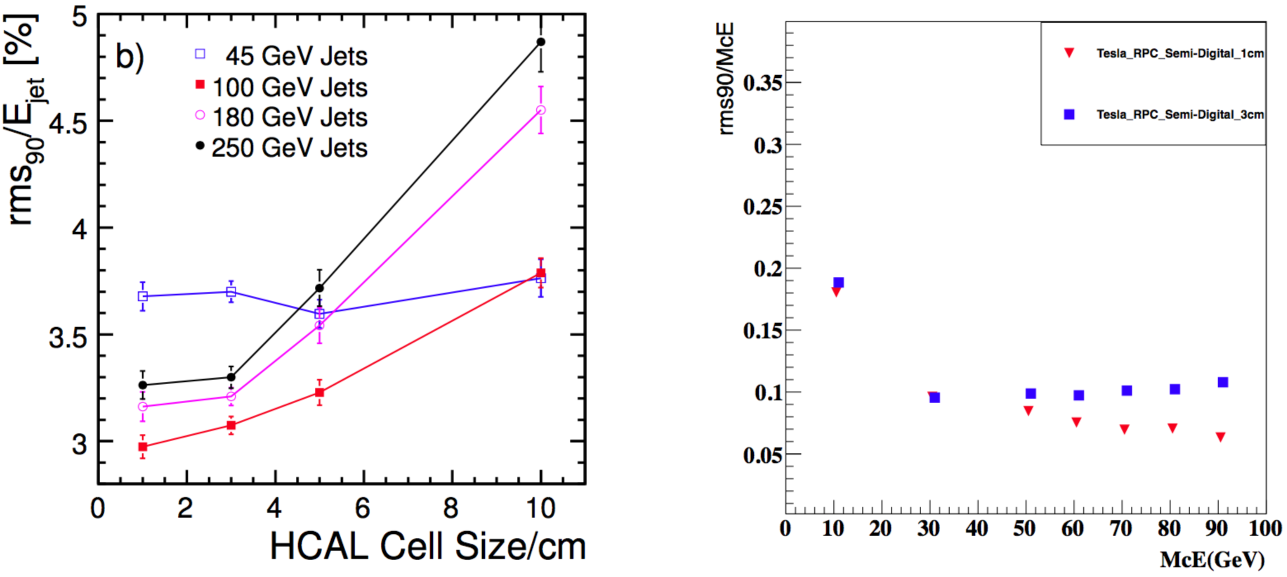


Figure 2.20 – Optimization of the analog and semi-digital hadron calorimeter cell sizes. Jet energy resolution obtained with PandoraPFA as a function of the AHCAL cell size for 45, 100, 180, 250 GeV jets (left). Single K_0 energy resolution in SDHCAL for 1×1 cm^2 and 3×3 cm^2 cells as a function of the K_0 energy (right). Figure are taken from [14].

Forward calorimeters (FCAL)

A system of three additional detectors (LumiCAL, BeamCAL, and LHCAL) will be installed in the very forward region of the detector close to the beam pipe, providing almost 4π detector hermiticity [64]. Schematically these forward calorimeter sub-detectors are shown in Figure 2.21. Bhabha scattering process will be used for the measurement of the luminosity with a relative precision better than 10^{-3} at $\sqrt{s} = 500$ GeV in the LumiCAL. This cylindrical calorimeter is centered on the outgoing beam. The apparatus is placed in a circular hole of the ECAL end-cap (inside ECAL ring), covering an angle between 31 mrad and 77 mrad. The BeamCAL provides a fast estimation of the relative changes in the luminosity on a bunch-to-bunch basis on beamstrahlung pairs. Similarly to the LumiCAL, this is a cylindrical electromagnetic calorimeter, and it covers the polar angle range 5 mrad and 40 mrad. This sub-detector is placed beyond the HCAL end-cap close to the beam pipe. Unlike LumiCAL and BeamCAL, the LHCAL is a hadronic calorimeter located in the region between the LumiCAL and BeamCAL inside the HCAL endcap. It extends the coverage of the HCAL to small polar angles and especially important for the very forward SUSY searches.

2.2.3 ILD outer detector

The outermost part of the ILD detector is composed of a superconducting coil providing 3.5 T magnetic field and an iron yoke instrumented to be used as muon tracker and tail catcher [65–67].

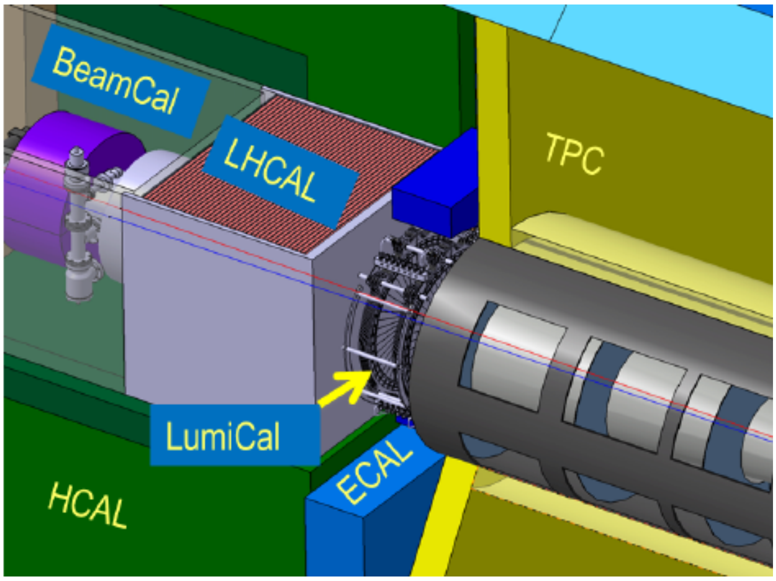


Figure 2.21 – The ILD very forward region. Support tube for the beam-pipe and final focusing carries LumiCal, BeamCal and LHCAL.

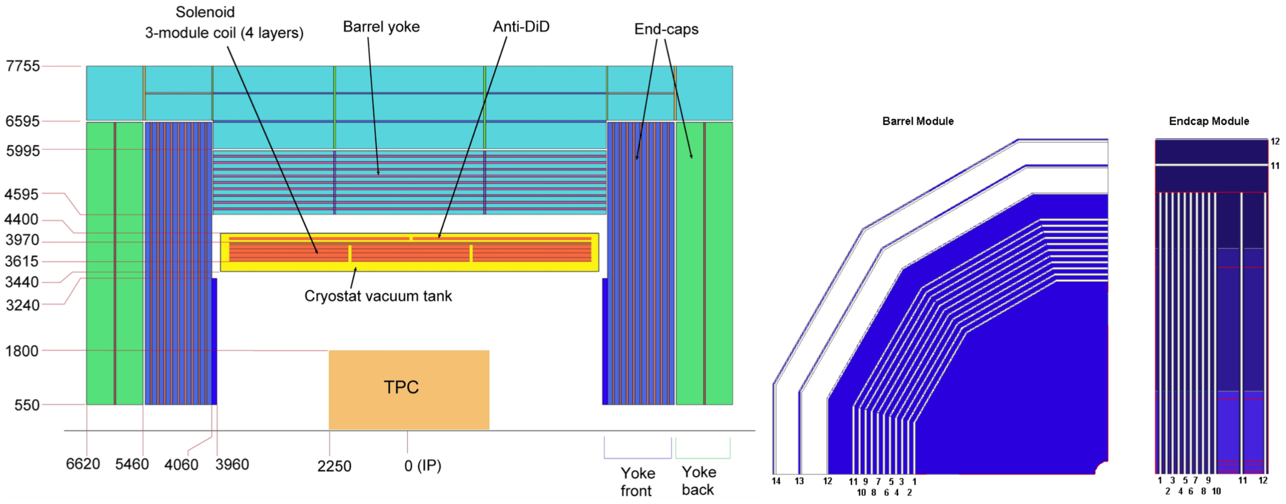


Figure 2.22 – Schematic cross section of the ILD magnet, all dimensions are in mm (left). Barrel and end-cap modules with sensitive layers of the ILD muon system/ tail catcher (right).

ILD coil and yoke system

For the successful application of the PFA, ILD tracking and calorimetric systems should be placed into a strong z-oriented magnetic field. The superconducting coil, see Figure 2.22 (left), surrounds TPC, ECAL, and HCAL, and creates an axial magnetic field of the nominal 3.5 T and up to maximal 4 T in a cylindrical volume of 6.88 m in diameter and 7.35 m in length. Within the TPC volume the integral field homogeneity of $|\int_0^{2.25m} \frac{B_r}{B_z} dz| \leq 10$ mm is required in the DBD. An anti-DID (Detector-Integrated-Dipole) is needed to minimize the background from the incoherent pairs from the beamstrahlung by driving them in the LumiCal hole. Precise magnetic field mapping in the TPC is required after the detector construction to achieve high precision tracking.

The iron yoke is a primary mechanical structure of the ILD. It supports all calorimetric and tracking subsystems of the detector. ILD muon system/ tail catcher is also installed on the yoke barrel and forward region. The coil is completely covered by the yoke to close the magnetic field and minimize stray fields. Besides, yoke and calorimeters provide detector self-shielding in terms of radiation protection.

ILD muon system/ tail catcher

The iron return yoke is instrumented with a muon tracker/ tail catcher. This system is needed for:

- efficient muon identification and hadron rejection, see Figure 2.23 (left);

- measurement of the hadron leakage from the HCAL to improve energy resolution, see Figure 2.23 (right).

The barrel part of the yoke is equipped with 14 sensitive layers. Active layers are installed between yoke iron plates. One layer is mounted in front of the yoke, ten layers are spaced 14 cm apart, and three layers are spaced 60 cm apart. Forward region of the yoke is equipped with 12 active layers, ten layers spaced by 14 cm are followed by two layers spaced by 60 cm outside the yoke. A schematic layout of the muon system/ tail catcher is shown in Figure 2.22 (right). Two options are considered for the readout, scintillator strips with SiPM or resistive plate chambers (RPC).

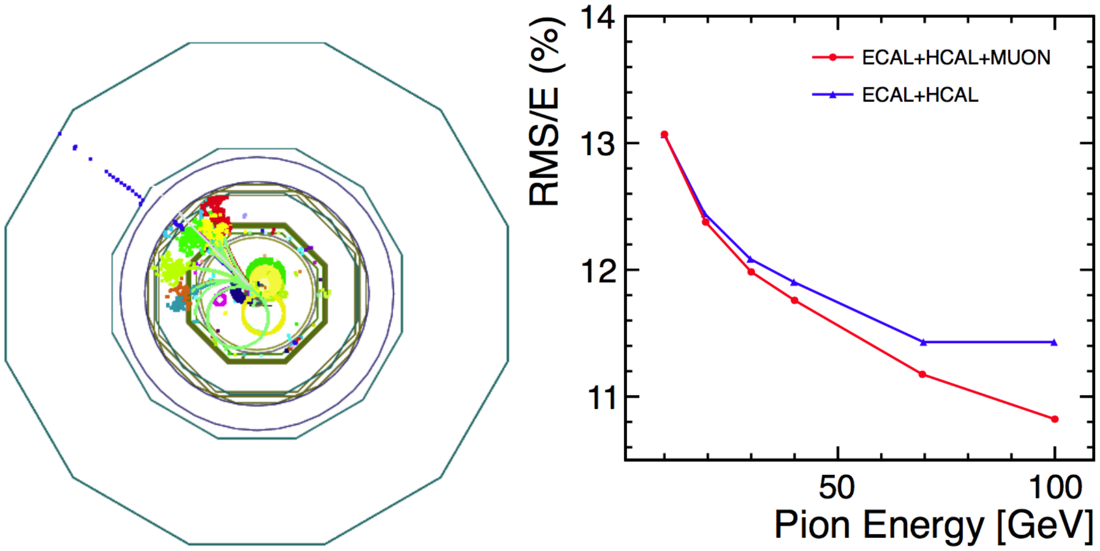


Figure 2.23 – An example of event display with a 50 GeV b-jet, the muon is seen in the muon system (left). Energy resolution of pions with and without the tail catcher as a function of incoming π energy (right). From [14].

2.3 ILD computing software and Reconstruction tools

In parallel to the ILD hardware R&D, there is ongoing activity on the software apparatus, called ILCsoft [68]. It provides data storage format standard for all subdetectors and simulation, analysis and reconstruction tools. Additional details are presented in this section.

2.3.1 ILC software tools

ILCsoft

ILCsoft is a set of software packages developed for the International Linear Collider. It provides core tools, such as LCIO, Gear, Mokka, and Marlin. This framework has been already used for the large MC production used for the ILD's Letter of Intent [20]. The ILCsoft framework is actively maintained and regular updates are released to provide more realistic simulations and improvements in the reconstruction of the ILD events.

LCIO [69] (Linear Collider Input/Output) is a persistence framework and data model that has been adopted as a standard by the ILC community. It defines an abstract event data model with hits, tracks, clusters, etc. and a particular file format to store the data.

Geant4 [70] (for GEometry ANd Tracking) is a toolkit for simulating the passage of particles through matter. It includes a complete range of functionalities including tracking of particle through matter, geometry description, physics models and hits. The physics processes cover a comprehensive range, including electromagnetic, hadronic and optical processes, a large set of long-lived particles, materials, and elements, over a wide energy range. It has been designed and constructed to work with various physics models, to handle complex geometries, and to be adaptable for optimal use in different sets of applications. Mokka [71] is a Geant4 based framework used for event simulation (uses event Generator, particle list and detector model as input). The output of Mokka is a LCIO file containing MCParticles and Hits deposited in the SubDetectors, and the Gear [72] (Geometry API for Reconstruction) XML file containing the geometry used for simulation. G4ParticleGun is a simple 1-particle ready-built Geant4 event generator which uses particle type, energy, position and direction as input. QGSP (Quark-Gluon String Precompound model) is a basic Geant4 physics list applying the quark gluon string model for high energy interactions of protons, neutrons, pions, kaons, and nuclei. Often, a high energy interaction creates an excited nucleus, which is passed to the code modeling the nuclear de-excitation.

QGSP_BERT is like QGSP, but using the Geant4 Bertini cascade model [73] for primary protons, neutrons, pions and kaons below ≈ 10 GeV. The Bertini model produces more secondary neutrons and protons than the Geant4 low energy parameterised (LEP) model, yielding a better agreement with experimental data. All MC simulations in this thesis are performed using QGSP_BERT physics list.

Marlin [74] (Modular Analysis and Reconstruction for the LINear collider) is a modular C++ application framework used to analyze LCIO data. Marlin provides a platform for the distributed development of ILC detector reconstruction and analysis algorithms. User written software packages can be easily configured inside this framework. The configuration of the application can be performed via an XML steering file, where one can change the default settings if needed.

CALICEsoft [75, 76] are software packages developed for the CALICE collaboration internal studies of the ILC project. CALICE software is built on top of the ILCsoft framework. CALICE test beam geometry is stored inside Mokka detectors database for the simulations. A set of software packages is implemented inside Marlin for the event reconstruction.

Detector models, simulation and digitization

All ILD sub-detectors are realistically implemented into the Mokka detector database to perform realistic physics studies. In particular, such engineering details as mechanical and support structures, cabling, electronics, dead materials, and cracks are adequately described. As much as possible the simulation models are independent of the readout technologies, if it is needed, different implementations are provided. Three stable Mokka models have been created for various detector configurations and readout technologies:

- ILD_o1_v05 Mokka model has SiW ECAL and AHCAL;
- ILD_o2_v05 Mokka model has SiW ECAL and SDHCAL;
- ILD_o3_v05 Mokka model has Sc ECAL and AHCAL.

ILD_o1_v05 and ILD_o2_v05 are the models used in the present document. A three-dimensional view of simulated ILD_o1_v05 geometry is shown in Figure 2.24. The ILD detector has detailed and realistic description:

- the tracking system (VTX, SIT, TPC & SET);
- the ECAL with W absorber and active layer based on silicon sensors with 5×5 mm^2 cells or 5×45 mm^2 scintillator strips;
- the HCAL with stainless steel absorber, two different mechanical structures and active layer designs for the AHCAL (layers filled with 3×3 cm^2 scintillator tiles are installed in TESLA-type mechanical structure) and SDHCAL (layers filled with 1×1 cm^2 GRPC are installed in Videau-type mechanical structure);
- the Forward calorimeters (LumiCAL, LHCAL, BeamCAL);
- the outer detector (final design is not frozen yet, so in the simulation models muon system is equipped with 3×3 cm^2 scintillator tiles, but in the ILC TDR 3×1 cm^2 strips are proposed).

Full event simulation is performed within Mokka framework. Mokka steering file allows implementing some changes into the simulated setup, like a modification of the number of active layers or absorber thickness, readout granularity, detector radius, etc. Any single change of the detector parameters causes recalculation of the entire detector to make it consistent with the proposed change. It means that not all changes are possible. In the present thesis simulations are done with the standard SiW ECAL with 5×5 mm^2 cells and with 2.5×2.5 mm^2 cells.

Recently the review of the ILD simulation models has been started in order to provide more realistic simulation setup. New simulation models and their description are available at [77].

Digitization of the ILD events can be done via dedicated C++ Marlin processors [78–82]. Hits in a tracking detector are digitized using a parametrization of the point resolution. Pixel detectors, like VTX (VXDPlanarDigiProcessor), first two layers of the FTD (FTDPixelPlanarDigiProcessor) and TPC (TPCDigiProcessor), provide 3D space points to the reconstruction algorithms. Silicon strip detectors, like SIT (SITPlanarDigiProcessor), SET (SETPlanarDigiProcessor) and 5 layers of the FTD (FTDStripPlanarDigiProcessor), provide only 1D TrackerHits, these 1D hits are converted into 3D space points by dedicated Marlin processors: SITSpacePointBuilder, SETSpacePointBuilder and FTDSpacePointBuilder. SiW ECAL and AHCAL information can be digitized with ILDCaloDigi or NewLDCCaloDigi processors. ECAL and HCAL calorimeter hit energies are multiplied by calibration factors according to the absorber sampling fraction, a threshold on the minimal hit energy can be applied. For the SDHCAL digitization, a SimDigital processor is needed. Digitized events can be used for the physics studies using PFA reconstruction tools.

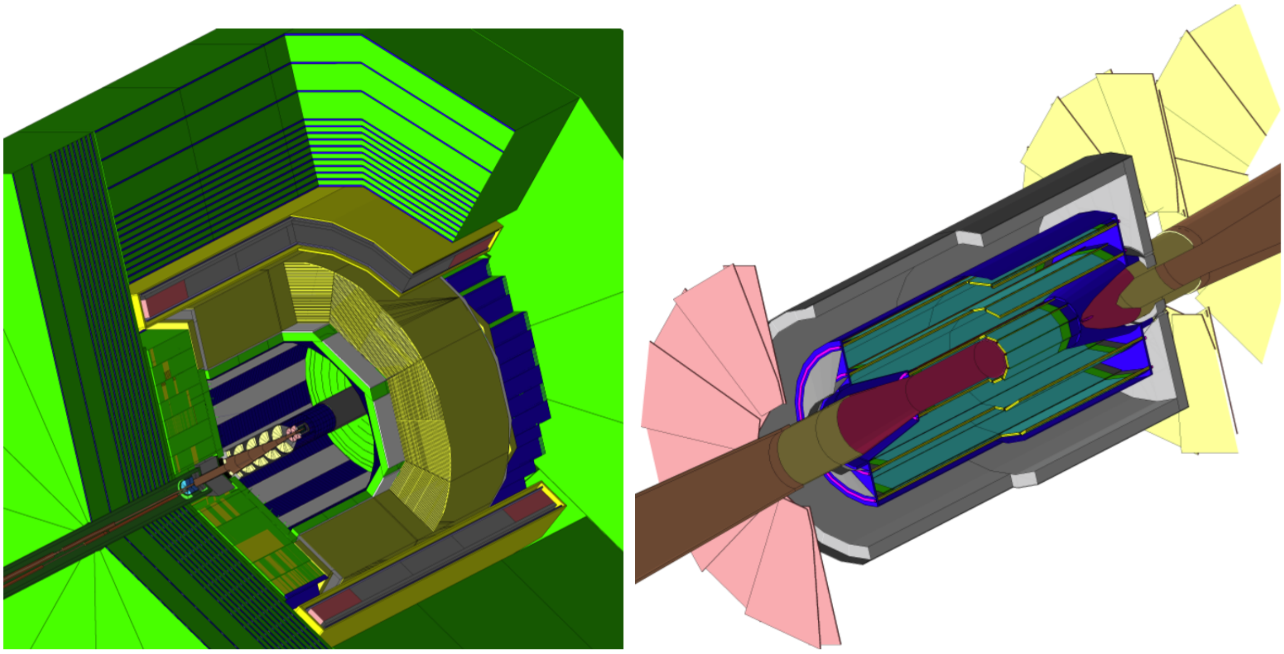


Figure 2.24 – Schematic 3D view of the ILD_o1_v05 Mokka geometry (left). Detailed view of inner tracking detectors (right). The following detectors are implemented into the simulations: VTX, SIT, TPC, SET, ECAL, HCAL, Coil, Yoke (from inside to outside) and FTD, LumiCAL, LHCAL, BeamCAL (forward direction).

2.3.2 PFA reconstruction tools

At the moment three different PFA reconstruction tools, PandoraPFA [18,83], Arbor [84,85], and Garlic [86,87] are available for the ILD event reconstruction. These algorithms are implemented as a set of Marlin packages (processors) working under the ILCsoft framework. All algorithms use the tracker and calorimeter hits from LCIO to perform event reconstruction. Tracking processors are the same for all three algorithms, but calorimeter clustering and particle identification mechanisms are different. XML steering files contain information about the Gear XML file with detector geometry to be used, the list of processors to be run and additional algorithms settings.

Tracking software

A set of tracking algorithms is implemented inside ILCsoft as Marlin processors [78,79]. For optimal performance, it is important to identify prongs, secondary vertices and particularly neutral ones, kinks and backscatters (see Figure 2.25). ClupatraProcessor provides the reconstruction of tracks in TPC. Track seeds are reconstructed with topological clustering methods. Tracks are reconstructed using Kalman Filter based extrapolations [88]. Optionally silicon tracking detectors can be included in this algorithms. ForwardTracking processor reconstructs tracks through the Forward Tracking Detectors (FTD). SiliconTracking_MarlinTrk processors is a dedicated pattern recognition algorithm for silicon detectors: VTX, SIT, SET, and FTD. TrackSubsetProcessor takes tracks provided by SiliconTracking_MarlinTrk and ForwardTracking as input and outputs them as one track collection. FullLDCTracking_MarlinTrk analyses all reconstructed track segments from all tracking processors and provides the final list of tracks for the PFA clustering algorithms. The track parameters are then extracted using the helical fit parameters.

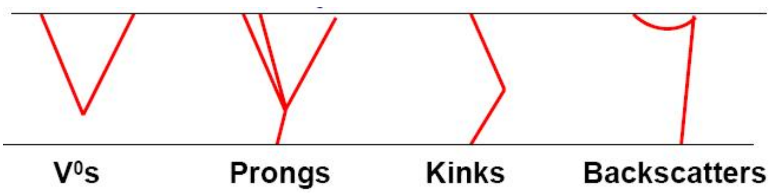


Figure 2.25 – Schematic view of possible charged particle decays in a tracker. Neutral vertices, prongs, kinks, and back-scatters have to be reconstructed.

PandoraPFA

PandoraPFA is a standard ILD reconstruction program, and it provides the best jet energy resolution among all PFA reconstruction programs. Pandora is implemented in ILCsoft as a set of C++ Marlin packages. Pandora uses tracks reconstructed by tracking processors and calorimeter hits as input for over 60 sophisticated clustering algorithms for particle reconstruction. XML steering file for Pandora requires providing Gear file with detector geometry and files with settings for particle identification. PandoraPFA is designed and best optimized for ILD geometry with SiW ECAL with 5×5 mm^2 cells and AHCAL with 3×3 cm^2 cells. An example of the τ jet ($\tau \to \nu + \pi^0(\to \gamma\gamma) + \pi^+$) reconstruction with PandoraPFA is shown in Figure 2.26. PandoraPFA program behavior is well-understood and has been documented in [18, 45]. Recently Pandora's codes were significantly rewritten to improve photon reconstruction, last versions of PandoraPFA programs are publicly available at Github [89]. Basics of the Pandora reconstruction are briefly schematically shown in Figure 2.27 and summarized below:

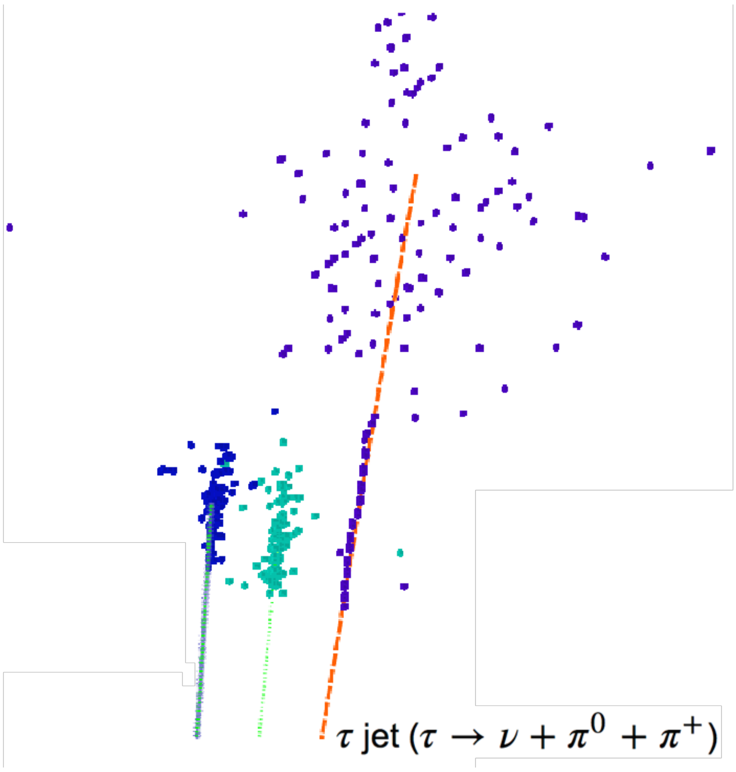


Figure 2.26 – An example of τ jet $\tau \to \nu + \pi^0(\to \gamma\gamma) + \pi^+$ reconstructed with PandoraPFA.

- Calorimeter hits are merged in clusters using a cone-based clustering algorithm. For the charged particles the ECAL entering points of the TPC track extrapolations are used as seeds for clustering. Cone-clustering works outwards the detector, starting from the first ECAL layer to the last HCAL layer.
- Clustering algorithm tends to over-split calorimetric contributions from individual particles to prevent accidental merging of particles at the early stage of the reconstruction. At the next step, these protoclusters are united into larger clusters by a set of topological algorithms. Fine granularity and excellent tracking performance of the calorimeters allow to separate nearby photon and charged hadron with very few mistakes, errors were observed in the case of two overlapping photons. See [30] and chapter 4 of this thesis for additional details.
- For the charged particles the inner detector tracks should be associated with calorimeter clusters. For the reconstructed particle, the inner detector track extrapolation (linear and helix fit) in the ECAL should point to the cluster, the particle energy measured with calorimeter should be, within errors, the same as the momentum provided by the tracker.
- If a mismatch between cluster energy and track momentum is observed, the re-clustering algorithms will try to regroup the hits in the event. If the track energy (E_{track}) is smaller than the cluster energy ($E_{cluster}$), some part of the initial cluster should be reconstructed as a neutral particle. In another case ($E_{track} > E_{cluster}$), the re-clustering algorithm will try to merge nearby hits with the cluster to increase its energy.

- Fragment-removal algorithms are seeking to resolve situations when parts of the charged clusters are reconstructed as neutral clusters. These algorithms are trying to associate neutral clusters with charged ones and are checking the consistency of the resulting track-cluster association.
- After these algorithms particle flow objects (PFOs) are formed. Properties of charged particles are extracted from tracks. Neutral particles have only calorimetric information.
- Particle identification algorithms flag PFO objects with PDG codes.

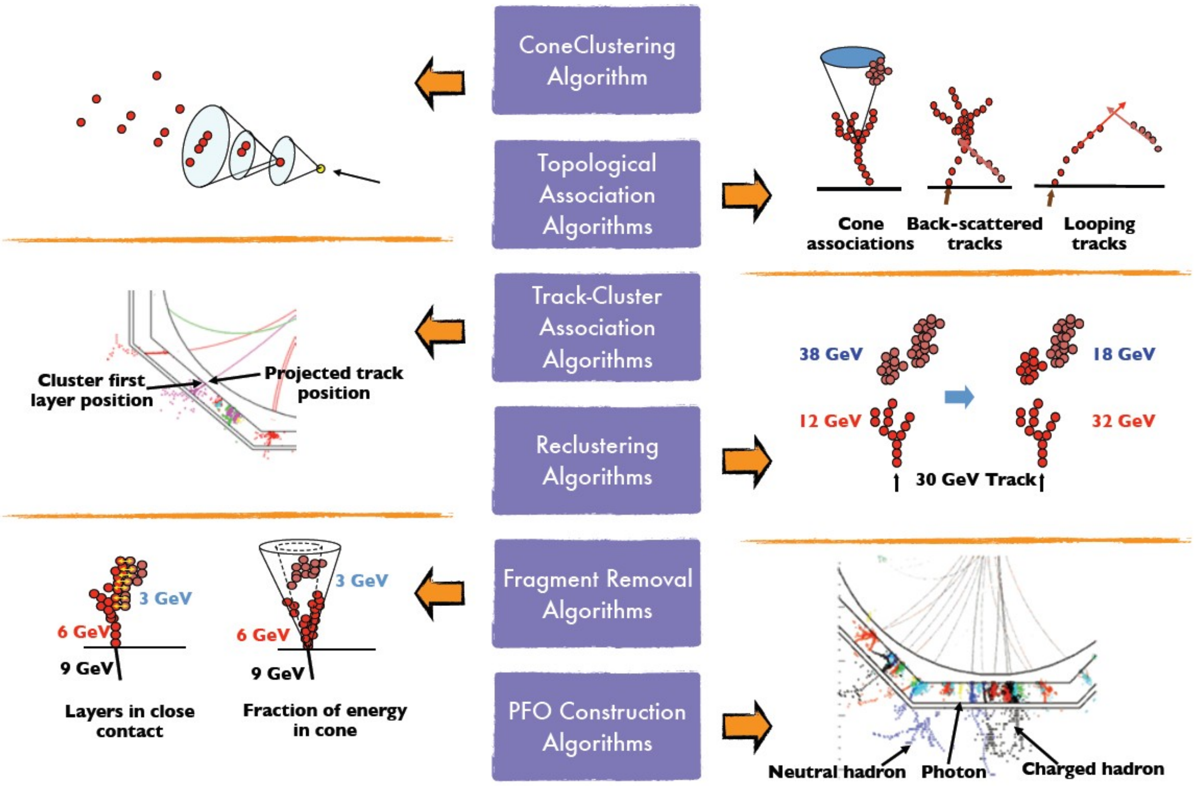


Figure 2.27 – Schematic illustration of the main PandoraPFA reconstruction algorithms.

Two versions of PandoraPFA program (v00-14 & v02-04) are used in this thesis. The first one is quite old and has significant problems with photon separation (see section 4.4 of this thesis). Photon reconstruction was improved, and v02-04 has been released. Several corrections in the PandoraPFA code were implemented because of findings in the present thesis:

- Some losses in two-photon reconstruction efficiency (PandoraPFA v02-03) for distances between 100 mm and 150 mm were observed [91], see Figure 2.28, part of the two-photon events were reconstructed as a single particle. The problem was related to the scales of the shower peak finding algorithm. This issue was fixed in the v02-04 release of the program.
- Standard PandoraPFA (v02-04) has problems in two-photon separation for the SiW ECAL with finer 2.5×2.5 mm^2 cells, due to the algorithm inefficiencies. Its performance is significantly worse than for larger 5×5 mm^2 cells. We got the suggestion to modify the limits of the PhotonFragmentRemoval algorithm [92]. Some details are presented in section 4.2. Significant improvements in separation performance are observed, but some additional improvements are still needed. It means that PandoraPFA is not entirely detector independent and some optimizations can be done.

The majority of the physics studies for ILC and CLIC experiments are done with PandoraPFA reconstruction program [93–98].

Arbor

Arbor (from Latin "arbor" - tree) is an alternative to PandoraPFA reconstruction program. This approach is based on the idea that hadronic shower development is very similar to a tree and its branch structure. For the first time, Arbor idea has been proposed by Henri Videau for the ALEPH collaboration. Recently this idea was

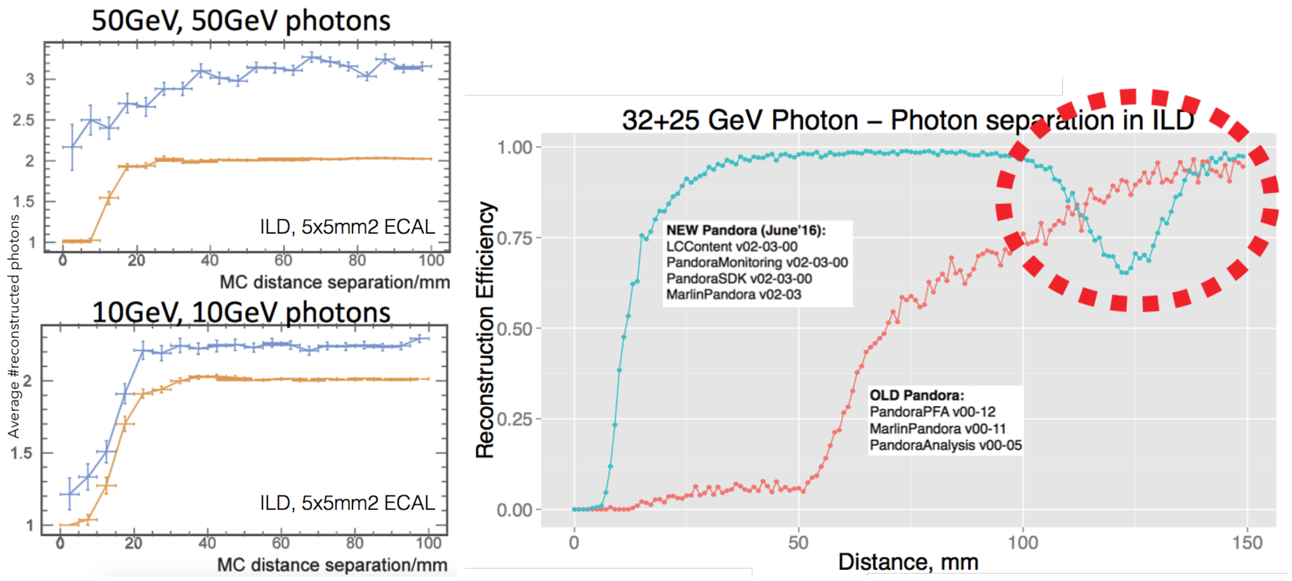


Figure 2.28 – Average number of reconstructed photons in a 2-photon sample as function of the distance between photons (left plot, from [90]), blue (yellow) curve corresponds to reconstruction with Pandora v00-12 (v02-00). Efficiency of 2-photon reconstruction with correct energies (±20%) and barycenters (±5 mm), significant fraction of photons for distances between 100 mm and 150 mm are reconstructed as single cluster (right plot, additional details can be find in chapter 4). This problem was fixed, see Figures A.33 - A.36 in Appendix A.1.

adopted for the ILD detector design [84]. In the present studies, Arbor (March'15 release) developed at IHEP (Beijing, China) is used. Another entirely independent implementation of the Arbor approach is proposed by the highly granular SDHCAL group at IPNL (Lyon, France) [85].

An example of a typical shower development in the detector is shown in Figure 2.29 (left). Here the charged hadron interacts with the absorber material, we see all main components of the shower: charged particles, neutral particles, electromagnetic and hadronic components. Physical observation of the similar interaction process in the highly granular sampling calorimeter is shown in Figure 2.29 (right). The tree-like topology of the hadronic shower is shown with black arrows.

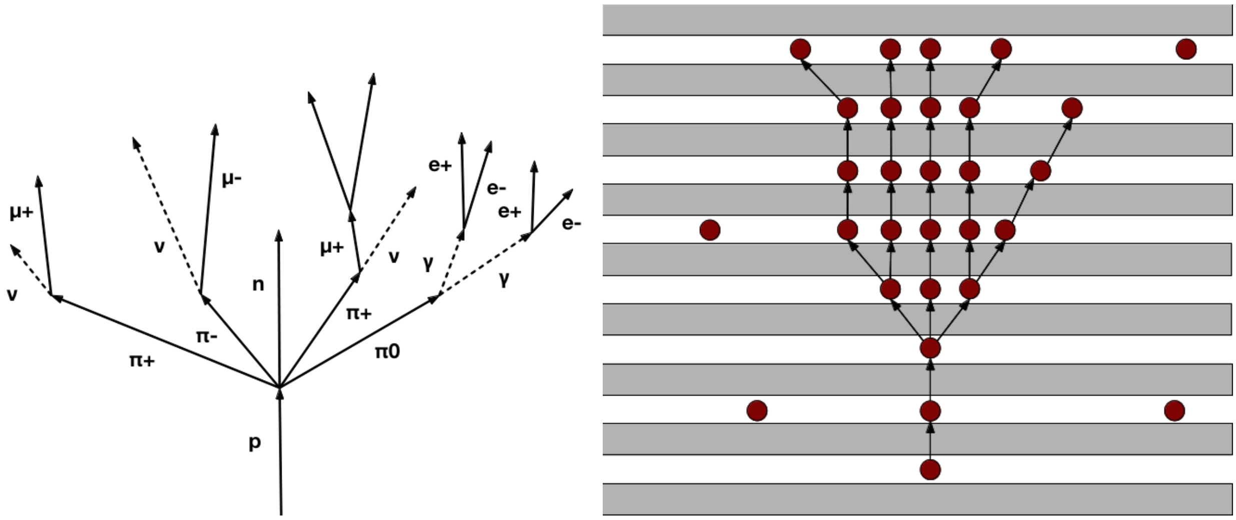


Figure 2.29 – Schematic view of a typical proton shower (left). An example of shower reconstruction with Arbor (right).

The Arbor algorithm is implemented as C++ Marlin processor in ILCsoft. It is best optimized for ILD with standard SiW ECAL (5×5 mm^2 cells) and SDHCAL (1×1 cm^2 cells). Currently, Arbor uses only geometry information of hits to reconstruct showers in calorimeters, i.e. energy information is not used. If a pair of hits is closer than the minimal distance threshold, these hits are connected with a connector (an arrow directed to the detector outwards). A connector cleaning procedure is performed to prevent the creation of loop structures. Several connectors can begin at a given hit (≥ 0), but at most one ends (only 0 or 1). Connectors are forming tree-like iterated structure with as long and smooth branches as possible. According to this tree-like topology, a hit is called a leaf if it has only one incoming connector (no outgoing connectors), a hit is called a seed if it has only outgoing connectors (≥ 1). From each leaf, there is a unique path to the seed. For a given seed the longest connector path is called a trunk, other paths that are traced to the same seed are called branches. Schematically a process of the Arbor tree reconstruction is shown in Figure 2.30. Additional merging algorithms are written for tree merging for neutral showers when several close seeds can be reconstructed in the same layer, this situation is illustrated in Figure 2.31. Reconstructed clusters are using shower fractal dimension analysis for the particle identification. Few examples of Arbor reconstruction are shown in Figure 2.32.

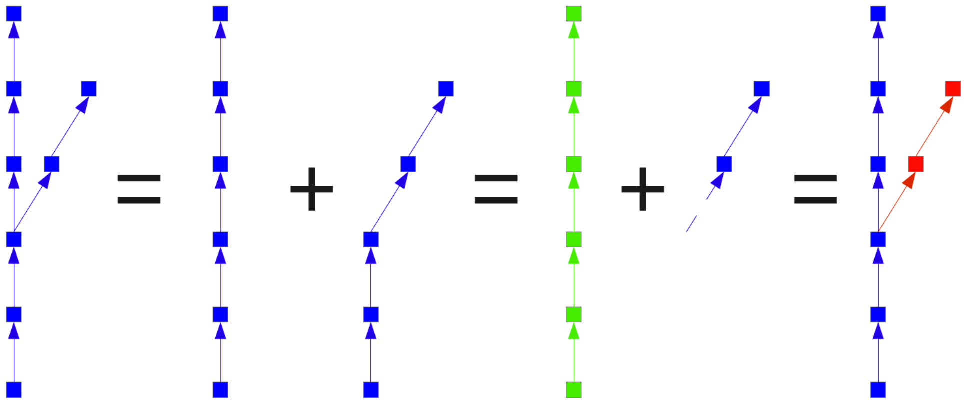


Figure 2.30 – Arbor mechanism of the tree structure decoupling into branches.

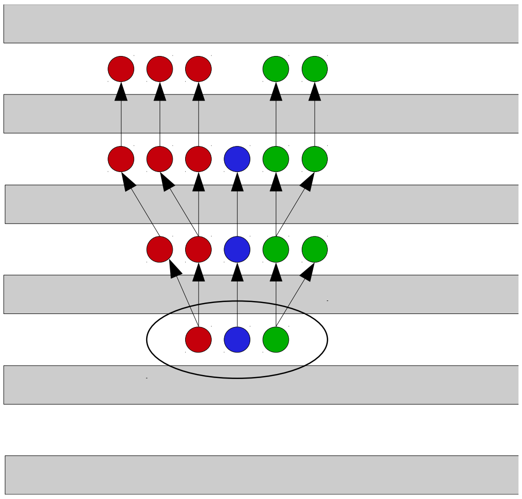


Figure 2.31 – Schematic view of a neutral tree merging algorithm.

The performance of Arbor was discussed with its developers. For the two-photon separation, the distribution of hits between reconstructed photons has been corrected. For this thesis Arbor (March'15 version) is used, but it tends to over-split particles in both ECAL and HCAL. Recently, a new version of Arbor was released [99], it has additional improvements in photon reconstruction and particle over-splitting. Some recent physics results with Arbor are available here [100].

Garlic

Garlic (GAmma Reconstruction at a LInear Collider experiment) is a third available PFA reconstruction program. It is designed to reconstruct photons from the interaction point (IP), simple e^+/e^- identification is also provided. Garlic algorithms make use of the narrow width and specific longitudinal profile of the electromagnetic showers. Garlic is implemented in the ILCsoft as C++ Marlin processor [86]. For the photon reconstruction, it uses tracks from the tracking processors and the ECAL calorimeter hits. It means that at the given stage full event reconstruction with Garlic is impossible, as it is not using the HCAL hits. Two examples of event reconstruction with Galic are shown in Figure 2.33. Potentially Garlic can be utilized as the photon identification algorithm for other PFA algorithms, like PandoraPFA. Basics of the Garlic reconstruction is briefly described below:

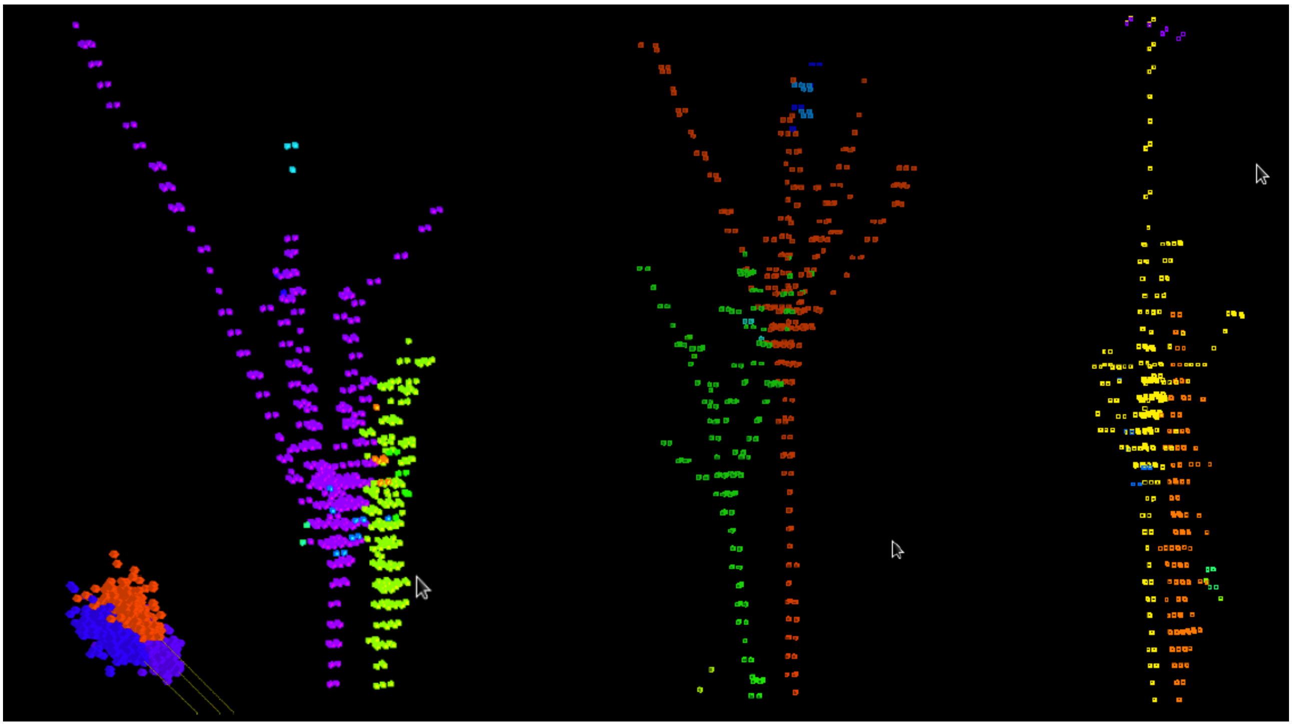


Figure 2.32 – The examples of nearby showers reconstructed by Arbor. Left event display corresponds to three nearby photons, other event displays to charged hadrons.

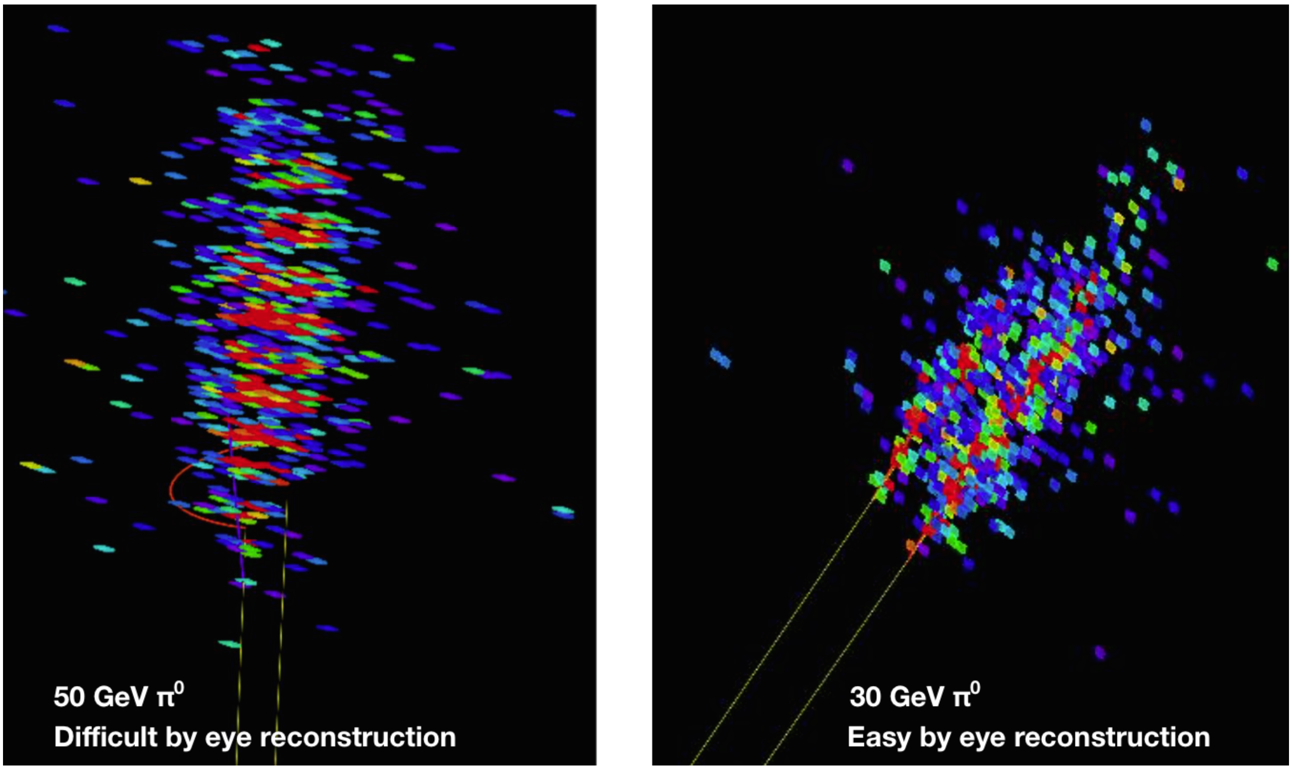


Figure 2.33 – Two examples of event reconstruction with Garlic: 30 and 50 GeV π^0 are decayed in $\gamma\gamma$.

- Garlic performs a simple e^+/e^- identification. For this, inner detector charged tracks are extrapolated to the front face of the ECAL defining the entrance point (seed) position. The algorithm forms a cluster core starting from this seed and reconstructs a cluster from ECAL hits. If the cluster energy is consistent with the corresponding track energy, this cluster will be identified as e^+ or e^- depending on the track curvature. These particles and their hits are not used for further photon cluster finding. It was observed that this algorithm needs additional tuning as charged hadrons with a large energy deposition in ECAL can be reconstructed as e^+/e^-.

- All hits closer than 10 mm (default cut) or 15 mm (suitable for the 10×10 mm^2 cells of the SiW ECAL physics prototype) to a TPC track extrapolated in ECAL are vetoed and excluded from further analysis to remove hits produced by charged non-showering particles.
- Remaining hits are grouped in pre-clusters using a primitive neighbor clustering to exclude isolated hits from the analysis.
- Seed-finding algorithm searches for the photon seeds (photons entrance point candidates). For this, all hits with signal higher than 2.5 MIP in the 12 first ECAL layers are projected on the plane perpendicular to the line between the IP and the pre-cluster center-of-gravity and passing through the intersection of the line with the ECAL front surface. Peaks of the distribution are considered as seeds.
- Cores are formed by the hits within a cylinder of radius 1.5 times the cell size around the seed. The cylinder direction is defined as IP to seed direction.
- Clusters are then iteratively built around the cores, with five iterations using a clustering distance of 1.4 times the cell size.
- Cluster merging algorithms are written to search for the cases in which a single photon has been reconstructed as several clusters. If at least 75% of the layers common to both clusters have close neighbor hits (closer than 1.5 times cell size), these clusters are merged.
- The next step is the photon identification procedure. Early versions of Garlic (v2.11 and earlier) are using neural networks to distinguish Garlic clusters originating from photons. Clusters are split into twelve classes: in six classes according to their energy (with energy boundaries at 0.2, 0.5, 1, 3, 5 and 10 GeV) and in 2 classes according to the distance between the cluster and the nearest track (closer or further than 100 mm). Neural networks are trained to identify photons in each of these 12 cluster cases. In the most recent Garlic version (v3.0.3) these neural networks are replaced with a list of cuts on the shower shape, profile, energy and other parameters that a reconstructed cluster should pass. Reconstructed photons are divided into three classes: tight, loose and very loose.

During my thesis, I had useful discussions with the Garlic developer to improve its photon reconstruction performance and implement some changes in the program code. Additional details can be found in the Section 4.2 of this thesis. Several ILC related physics studies have been performed with Garlic reconstruction algorithm. They are available here [102, 103].

Chapter 3

CALICE collaboration and the calorimeter prototypes

3.1 Overview of CALICE collaboration

CALICE collaboration (CAlorimeter for LInear Collider Experiment) [21] has been created to develop highly granular calorimeters [104, 105] optimized for PFA. Currently, the collaboration includes about 360 scientists from 61 institutions and 18 countries, see Figure 3.1. The collaboration develops a wide spectrum of various technologies schematically shown in Figure 3.2. They will be discussed in the following. The internal competition helps to constantly improve the calorimeter prototypes and to implement the modern state-of-the-art solutions. In addition, different technologies reuse common CALICE tools for data acquisition (DAQ).

Figure 3.1 – CALICE logo and 18 countries participating in the collaboration.

CALICE technologies can be used at ILC, CLIC [33], FCCee [7] or CEPC [9]. Recently, the silicon calorimeter technology has been approved for the HL-LHC CMS HGCAL [106, 107] phase II upgrade project. Scintillator tiles readout by Silicon Photomultipliers (SiPMs) can also be used in HGCAL, this is currently under discussion. The silicon sensors with low amplification of the order 10–20 are also proposed for ATLAS HGTD [108] upgrade project.

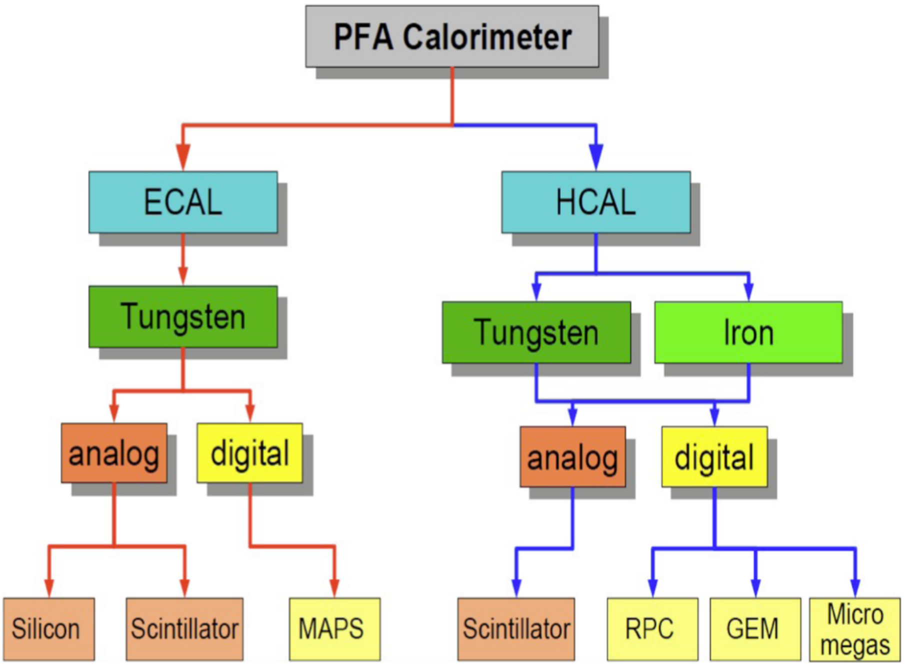


Figure 3.2 – Highly granular ECAL and HCAL technologies developed within CALICE collaboration.

3.2 ECAL technologies

Currently, CALICE collaboration is developing three different technologies for the highly granular ECAL, two with the silicon sensors and one with the scintillator strips [119]. The silicon sensors either have PIN diode pixels of about 5×5 mm^2 or MAPS pixels of about 50×50 μm^2 [128, 129]. In the latter case they are read out only in the digital yes – no mode. All technologies use tungsten as an absorber material, as it has a small Moliere radius, $R_M^{(W)} = 0.9327$ cm, and a large ratio of nuclear interaction ($\lambda_I^{(W)} = 9.946$ cm) and radiation lengths ($X_0^{(W)} = 0.3504$ cm): $\lambda_I^{(W)}/X_0^{(W)} = 28.38$, as it was explained in the Introduction.

Currently, the MAPS detectors are far from the realization in mass scale. In particular, their readout is too slow due to large number of channels. Another problem is the cooling. SiW ECAL proposes the best characteristics for PFA. The scintillator strips give less granularity but also are less expensive. The scintillator response is not uniform across the strip and across the SiPMs. The linearity of the latter also suffers from the saturation effects. SiPMs are sensitive to the temperature changes. Therefore, calibration of scintillator ECAL will require more efforts and may bring larger systematic errors.

3.2.1 Silicon-tungsten ECAL physics prototype

The first CALICE SiW ECAL physics prototype has been built in 2005–2006 with the last sensors completed in 2007 [111]. This is a full ECAL with 9720 channels, 30 layers and $\sim 24X_0$, see Figure 3.3. The transverse size is $18{\times}18$ cm^2. It is made of a carbon-fiber mechanical structure holding every second layer of the tungsten absorber. The structure has cavities called alveoli into which the active silicon detectors called slabs slide in as shown in Fig. 3.4. The slabs have other layers of tungsten in the middle and are equipped on both sides with the silicon sensors. The calorimeter contains in average 99.5% (98%) of the energy of the 5 GeV (50 GeV) e^+ shower. Absorber thickness varies across the calorimeter, the first 10 tungsten layers are 1.4 mm thick (corresponding to $0.4X_0$), next 10 in the middle are twice thicker, 2.8 mm ($0.8X_0$), and the last 10 layers have a thickness of 4.2 mm ($1.2X_0$).

One ECAL layer contains 3×3 silicon sensors, 3×2 in the upper and 3×1 in the lower slab. Each sensor has 6×6 pixels of 10×10 mm^2 size. The following development of the electronics in the technological prototypes allowed to reduce the granularity down to 5×5 mm^2. As it was discussed in the ECAL section in the previous chapter, the jet energy resolution is slightly better for 5×5 mm^2 for the jet energies around or above 100 GeV when the PFA confusion matters.

The silicon thickness in the physics prototype is 525 μm. The guard ring around the sensor which ensures high voltage stability and low dark currents, is shown in Figure 3.5, right. Since it was not segmented, there was a sizable fraction of "square" events when the significant energy deposition at the guard ring was propagated via the capacitive coupling to the peripheral pixels and fired them. Further details on this effect can be found in [112]. It will also be discussed in Chapter 5 where it will be shown that the guard ring optimization in the technological prototype allowed to significantly suppress this effect.

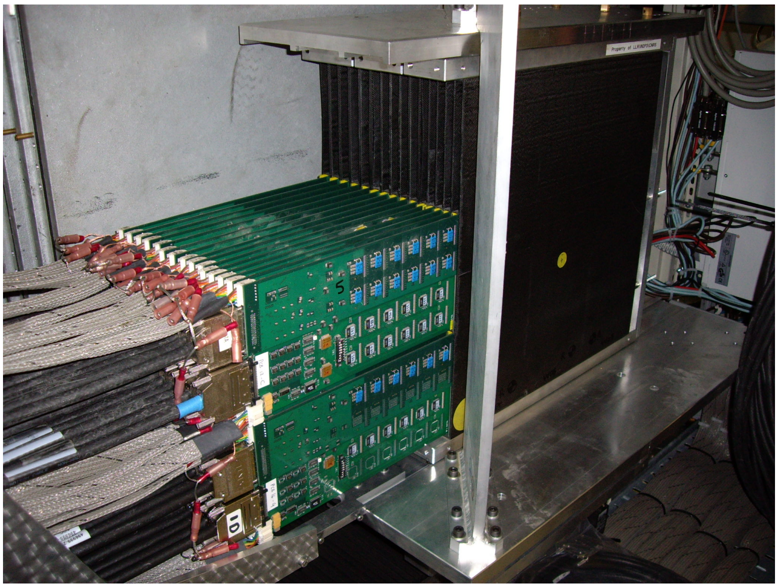


Figure 3.3 – CALICE SiW ECAL physics prototype.

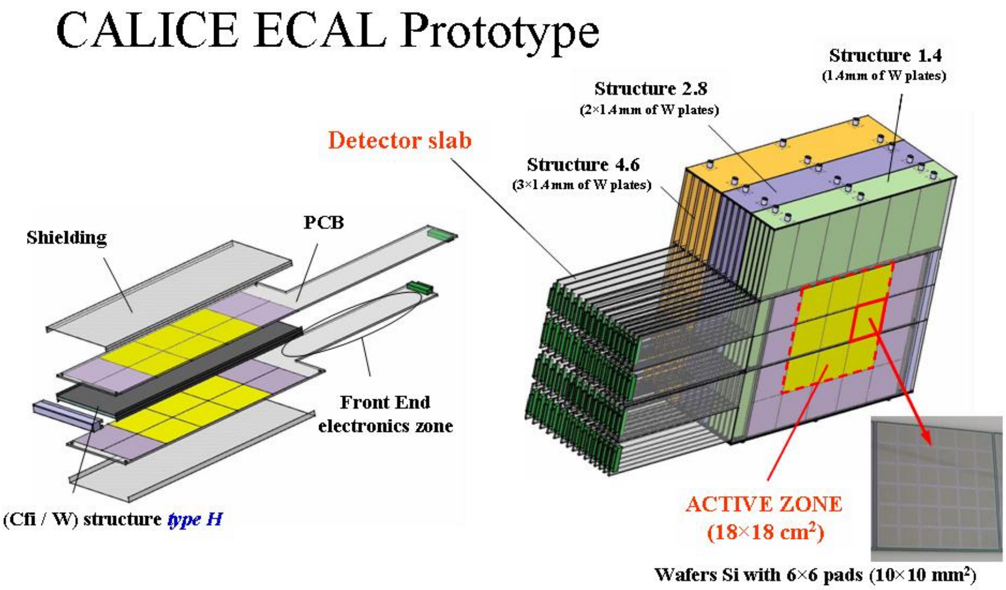


Figure 3.4 – Schematic view of SiW ECAL physics prototype, double-layer slab, and 6×6 cm^2 silicon sensor.

Both in the technological and physics prototypes, the silicon sensors are glued pixel-by-pixel to the corresponding pads of the printed circuit board (PCB) with the conductive epoxy. Contrary to the technological PCB which carries all front end electronics inside ECAL, in the physics prototype the electronics was not embedded but placed alongside the detector, see Figure 3.6. Two FLY_PHY3 ASIC chips each with 18 channels are used to read out one sensor. Their output analog signals are transmitted through the differential lines to a VME crate equipped with CALICE Readout Cards (CRC) which perform 16-bit analogue-to-digital conversion (ADC). Each CRC can read out up to 96 FLY_PHY3 chips or 5 layers, so 6 CRCs are needed in total.

An example of a 10 GeV electron event is shown in Figure 3.7, left. The distribution of the total energy measured by ECAL in the events with 20 GeV e^- showers is shown in Figure 3.7, right. Detector calibration was performed with the muons during all test beams at CERN, DESY and FNAL. The measured calibration constants were stable during 5 years of tests in 2006–2011 [116]. With 6–45 GeV electron beam at CERN the non-linearity was measured to be within 1% and the energy resolution to be $16.59\%/\sqrt{E(GeV)} \oplus 1.05\%$, see Figure 3.8. Further details on the physics prototype design and the test results can be found in [111, 113–115].

The technological prototypes of SiW ECAL will be discussed in detail in Chapter 5 of the thesis.

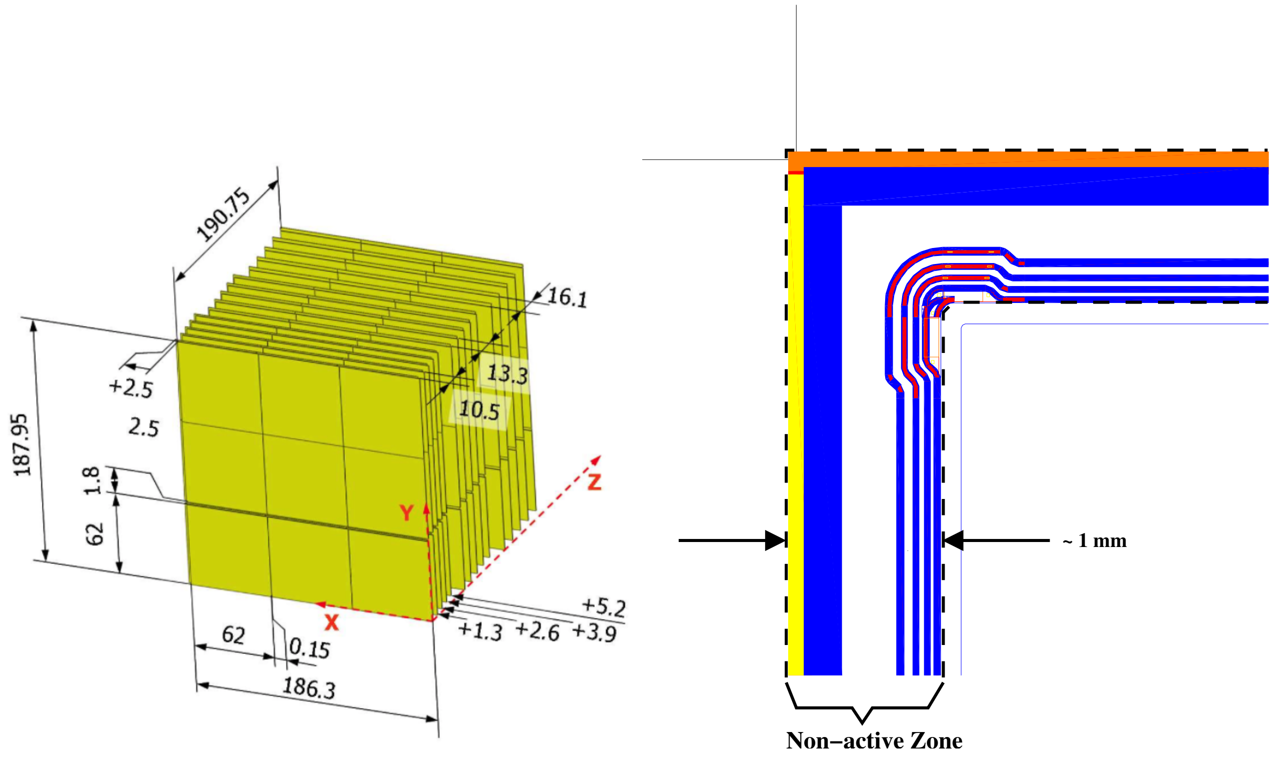


Figure 3.5 – SiW ECAL physics prototype active zone (left) and the design of the sensor guard ring (right).

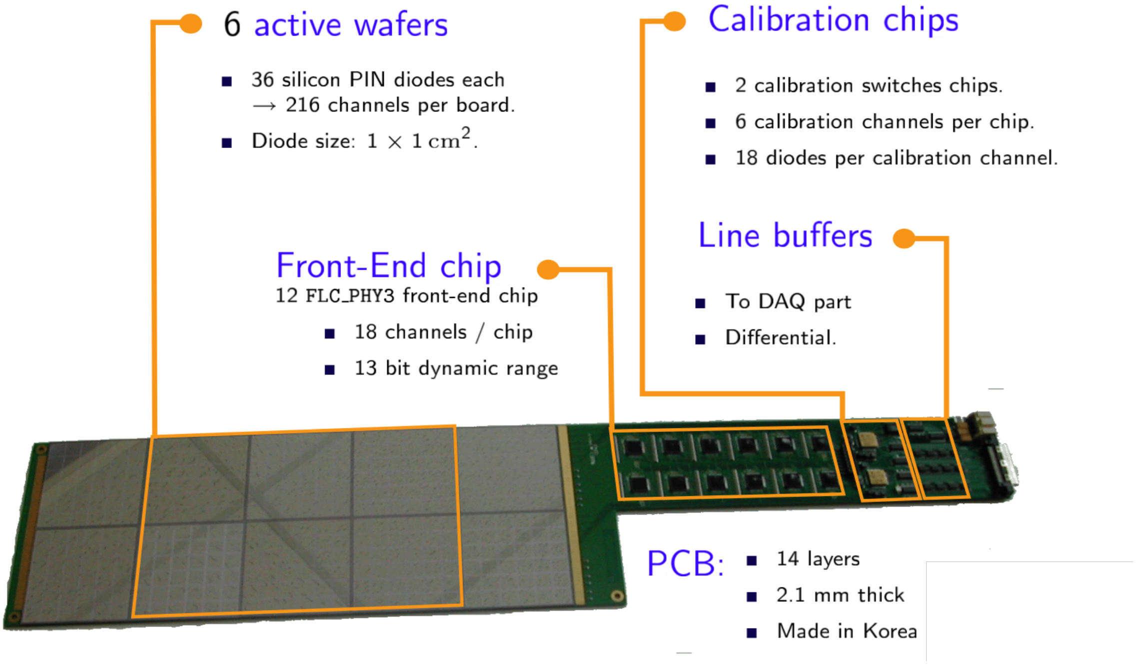


Figure 3.6 – SiW ECAL physics prototype PCB board. The front end electronics was not embedded inside ECAL.

3.2.2 Scintillator ECAL

The scintillator-tungsten ECAL (ScW ECAL) is made of scintillator 5×45 mm^2 strips readout by the SiPMs. To compensate less granularity, the strip orientation is alternated in ECAL layers. "Virtual" 5×5 mm^2 pixels are formed as "intersections" of perpendicular strips in the adjacent layers. The energy in such "virtual" pixels is not measured but is estimated offline according to the energy sharing in the perpendicular strips in the neighboring layers.

This solution allows to reduce the number of channels and the SiPMs, which otherwise dominate the detector cost. Optionally, the ILD ECAL can have both silicon and scintillator layers which is usually referred to as a

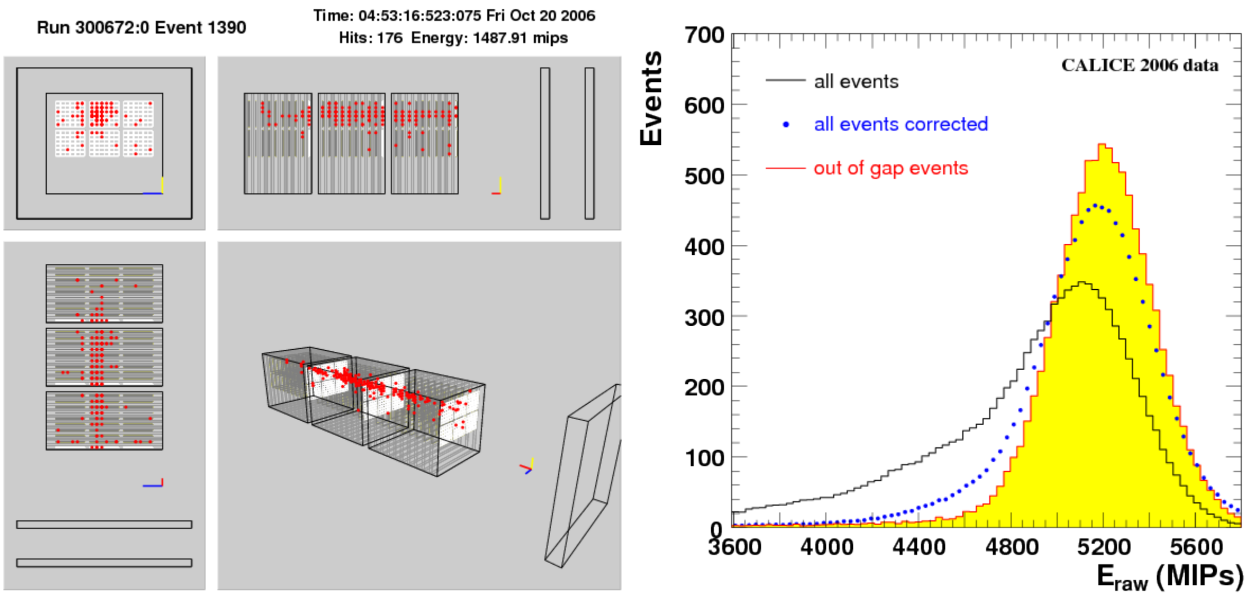


Figure 3.7 – A typical 10 GeV e^- event display in CALICE SiW ECAL (left). The distribution of the total ECAL energy measured in 20 GeV e^- shower events with and without correction for the losses in the gaps between the active sensor areas (right). Red curve shows the corresponding distribution not affected by dead areas. The histograms are normalized to the same number of entries. The Figure is taken from [117].

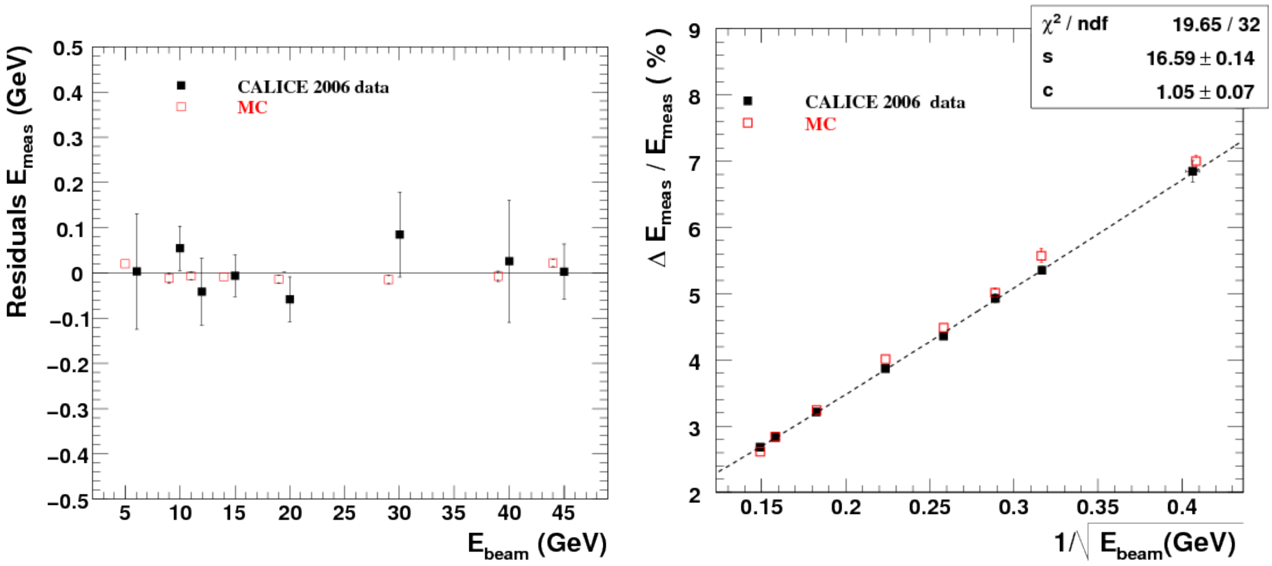


Figure 3.8 – The linearity of the SiW ECAL response as a function of the beam energy E_{beam} (left) and the relative energy resolution as a function of $1/\sqrt{E_{beam}}$ (right) [117].

“hybrid” ECAL. The optimal placement of the scintillator layers in the hybrid is still to be determined. One possibility is to have the scintillator in several last ECAL layers in the “transition” region to HCAL, especially if the scintillator technology is also used in the HCAL.

The difference between the HCAL scintillator is that the ECAL should be more precise as it measures 27% of the jet energy compared to 10% of neutral hadrons in HCAL. The ECAL constraints on the non-linearity and the dynamic range are more stringent because the electromagnetic showers are narrower and have larger energy concentration in the core. This requires SiPMs with more pixels. The scintillator photon yield should, on one hand, be sufficient for the minimum-ionizing-particle (MIP) calibration and, on the other hand, should not saturate the SiPM. Since the scintillator photon yield varies from strip to strip, there should be some margins on both ends. To increase the dynamic range, recently, the SiPMs with 10 000 pixels have been tried in the technological prototype. One pixel SiPM signal is given by the formula $C\Delta V$ where ΔV is the SiPM overvoltage and C is the pixel capacity. The latter is proportional to the pixel area, so with more pixels and less pixel size the signal becomes smaller. Currently, the ASIC chip dedicated to the SiPM readout and called SPIROC, has the characteristics optimized for the HCAL with about 6 times larger signals. Even its maximal gain does not match well to the 10 000 pixel SiPM signals.

The other thing is that the SiPMs should be precisely calibrated. Their responses vary significantly more than the signals from the silicon pixels, even across the SiPMs produced from the same wafer. The SiPM response also depends on the temperature, so one should either precisely measure it and correct offline or develop some temperature dependent circuitry automatically adjusting the overvoltage. The light collection efficiency is not constant across the scintillator strip, it is larger at the SiPM end. Therefore a special effort has been made to optimize the strip geometry, without introducing dead zones, and the coupling of the SiPM to the strip. The resulting non-uniformity can be within 10% or even less.

The first ScECAL physics prototype had 10×45 mm^2 strips assembled in 26 $9{\times}9$ cm^2 layers, 468 channels in total. It had 3.5 mm thick tungsten-cobalt absorber plates, the total radiation length was 18.5 X_0 [119, 120]. It was tested in March 2007 at DESY with $1-6$ GeV/c e^+ beams. The non-linearity was less than 1%, the energy resolution was found to be $13.45\%/\sqrt{E(GeV)} \oplus 2.87\%$ with the large constant term probably due to the non-uniformity of the strip responses and to the shower leakage.

To reduce the leakage, the second ScECAL prototype was made twice wider in each transverse direction, 18×18 cm^2, and contained two extra layers. The total radiation length has increased to 21.3 X_0. The prototype is shown in Figures 3.9 and 3.10. There were 72 readout channels per layer and 2160 channels in total. This prototype has been tested at FermiLab with electron and charged pion beams. The nonlinearity was measured to be 1.6% and the energy resolution to be $12.8\%/\sqrt{E(GeV)} \oplus 1.0\%$ for $2-32$ GeV electrons, see Figure 3.11. Further details and the test beam results can be found in [121].

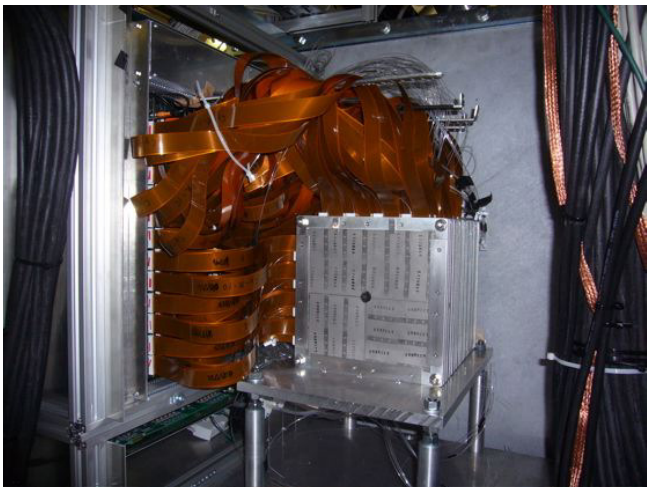


Figure 3.9 – ScECAL prototype in front of AHCAL.

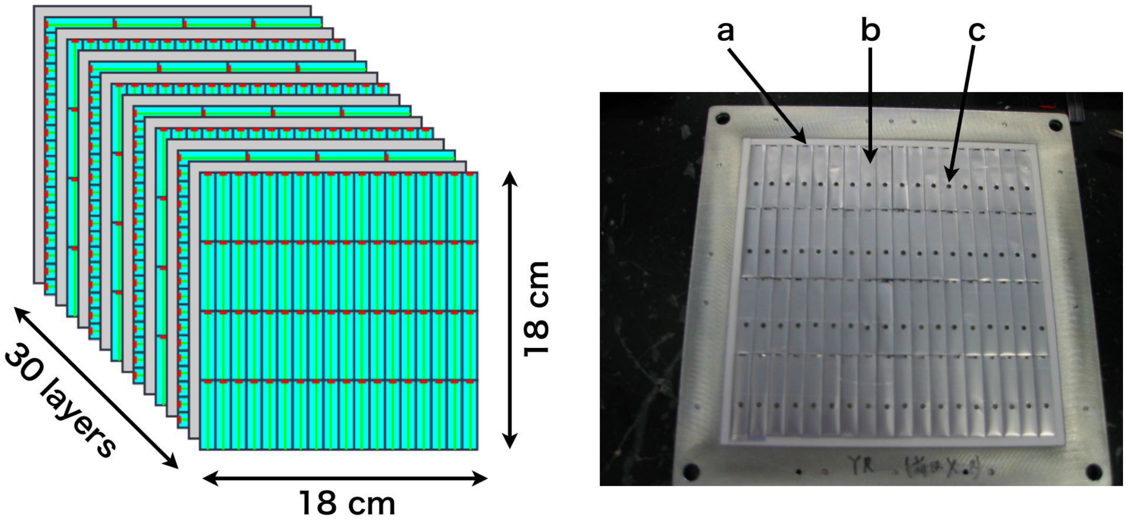


Figure 3.10 – The schematic structure of the ScECAL prototype (left). One layer of the prototype with a) the SiPM housing, b) the scintillator strip hermetically covered by the reflector foil with c) the holes for the LED calibration (right).

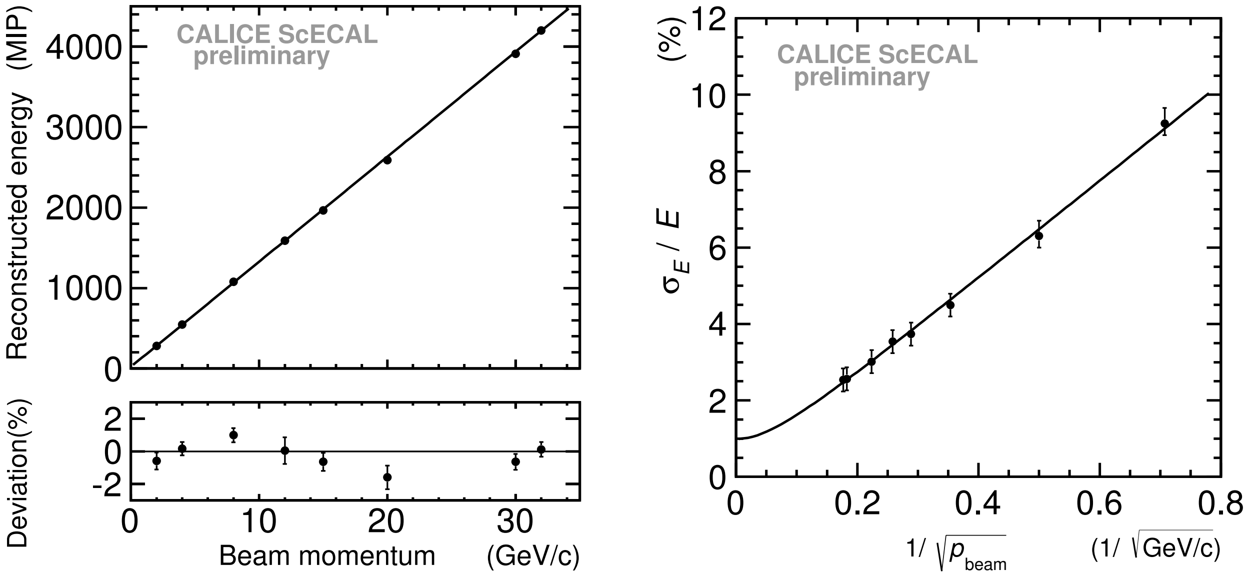


Figure 3.11 – The linearity (left) and the energy resolution (right) of the ScECAL prototype [121].

The first 18 × 18 cm^2 layer of the new technological ScECAL prototype with 144 5×45 mm^2 strips has been constructed and tested at DESY in autumn 2012. The readout electronics is embedded in the ECAL base board units (EBU) as shown in Figure 3.12. The results of the first tests of the technological prototype can be found in [122, 123].

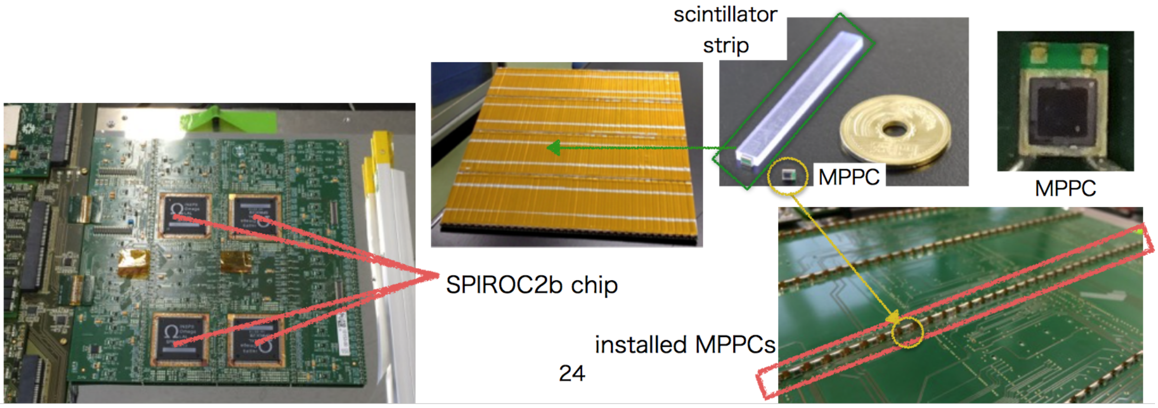


Figure 3.12 – ScECAL technological prototype detector layer and its main components.

3.2.3 Digital ECAL with MAPS

Digital ECAL is the third option of the highly granular ECAL. Its development has started within the CALICE collaboration and a crude prototype was even tested with beam [124], but now it is mainly continued for the ALICE Forward Calorimeter (FoCal) upgrade project. The recently constructed FoCal prototype [125, 126] is a first full calorimeter prototype. It will be very briefly described in the following. The prototype has record parameters: granularity of 30×30 μm^2 and the Moliere radius of about 11 mm, only 2 mm more than the Moliere radius of pure tungsten. It has 24 layers each with 0.97 X_0. The layer has a thickness of 4 mm, 3 mm of which consists of tungsten. The active area of one layer is 4×4 mm^2, the absorber size is 5×5 mm^2.

In the Monolithic Active Pixel Sensors (MAPS) technology each CMOS pixel provides yes – no binary information and the shower energy is determined from the number of hits. The pixelization at such an unprecedented level is needed in FoCal to separate individual photons from high energy π^0 decays in the forward ALICE region at LHC. The detection of direct photons in nuclear collisions can probe the parton distribution functions while the thermal photons are among the most interesting signals of quark-gluon plasma.

Though very promising, the MAPS technology is currently not seriously considered for ILD because of several technological challenges. The active pixel sensor dissipates typically 0.1 W·cm^{-2}, much more than the PIN diode. The maximal silicon area corresponding to 30 ECAL layers and the baseline ILD radius, is 2 600 m^2, which corresponds to the total MAPS power of 2.6 MW. Another problem is the MAPS readout time. In the FoCal prototype the PHASE2/MIMOSA23 [127] chip from IPHC is chosen as a sensor. Currently, this is the only large size MAPS sensor (with 640×640 pixel matrix) which allows the continuous readout of all pixels, thanks to four outputs at 160 MHz. Still, the 1 MHz rolling shutter corresponds to an integration time of 640 μsec. Such a low event rate is perfectly acceptable for the prototype but not suitable for ILD or LHC. Finally, the MAPS sensors are very expensive.

However, the first FoCal prototype results are very encouraging. The accumulated data allow to see and to measure the full electromagnetic showers with an unprecedented granularity. The MAPS technology is new and it may open new horizons in the future calorimetry.

3.3 HCAL technologies

CALICE develops two HCAL technologies, based on scintillator tiles readout by SiPMs and on Resistive Plate Chambers (RPC). They differ in granularity, 30×30 mm^2 for the former and 10×10 mm^2 for the latter, and in a way of measuring the energy depositions. Scintillator signals are currently digitized with 12 bits ADC, therefore the scintilator HCAL is called Analog HCAL (AHCAL). The RPC is readout either in a binary yes – no form, ie. with one threshold, or with three thresholds. In the former case the calorimeter is called Digital HCAL (DHCAL), in the latter case – Semi-Digital (SDHCAL). Historically, DHCAL appeared before SDHCAL, but with further development of the semi-digital electronics, SDHCAL becomes more preferable simply because it provides more information on the shower. DHCAL does not measure the shower energy directly, but it is approximated as a function of the number of hits. The SDHCAL provides three such numbers for three thresholds, and the shower energy is expressed as a function of these numbers. It is more precise for higher energy showers when DHCAL starts to saturate.

The baseline HCAL absorber is made of stainless steel. This is sufficient for the ILD. For the CLIC detector, to improve the containment of high energy showers and to reduce the detector size, it was proposed to use the tungsten absorber. CALICE has performed the tests of both absorbers. In the end the stainless steel was also chosen as the baseline for the CLIC detector.

3.3.1 Analog HCAL

AHCAL is a sampling calorimeter with the layers segmented into pads which are read out by SiPMs. In the first physics prototype the wave-length-shifting fibers were used for the light collection, in the recent prototypes they are removed and the SiPMs read out the scintillator light directly.

AHCAL physics prototype has a volume of ∼1 m^3, 38 layers and 7608 channels. It is a first detector which required “mass production” of SiPMs. The thickness of one layer of steel absorber is 17.4 mm, the total AHCAL thickness is 5.3 nuclear interaction lengths. In the ILD, AHCAL has a granularity of 3×3 cm^2. To reduce the number of SiPMs and to simplify the physics prototype, its segmentation is degraded at the boundaries and and at the back. Only 10×10 central tiles in the first 30 layers have the size 3×3 cm^2. Peripheral regions of the first 30 layers and all area of the eight last layers are covered by coarser 6×6 cm^2 or 12×12 cm^2 tiles. A photo of one layer is shown in Figure 3.13.

AHCAL physics prototype was operational in 2006–2011. It has demonstrated a good performance during all test beam campaigns in CERN, DESY, and FNAL. The time stability of SiPMs was sufficient for PFA [60]. Accumulated together with SiW ECAL data on hadronic interactions at unprecedented granularity level not only confirmed PFA principles [26, 30, 85], but were also used to tune Geant4 hadronic shower simulation models [109, 110].

The best AHCAL results can be achieved using so-called software compensation technique. Here, the total shower energy is calculated as a weighted sum of hit energies with the weights dependent on the local energy density around the hit. Smaller weights are assigned to the hits in the dense parts of the shower which often correspond to π^0 electromagnetic subshowers. This approach allows to compensate otherwise not $e : \pi$ compensated HCAL. Figure 3.14 shows the detector linearity and the energy resolution after applying the software compensation technique. Additional details on AHCAL physics prototype are published in [131–135].

Starting from 2011, a new technological AHCAL prototype is under development with electronics fully embedded in the active layer [136, 137].

3.3.2 Tail catcher/ muon tracker

Combined SiW ECAL and AHCAL physics prototypes contained 97% (92%) of the energy of 10 (80) GeV pion shower. To measure 3 (8)% leakage and also to detect muons, the setup was complete at the back by the so-called tail catcher/muon tracker (TCMT). It is made of scintillator strips and is similar to AHCAL but with much coarser granularity. TCMT has 16 layers shown in Figure 3.15. The steel absorber plates have transverse

Figure 3.13 – One scintillator layer of AHCAL physics prototype before assembly (left) and the readout electronics (right).

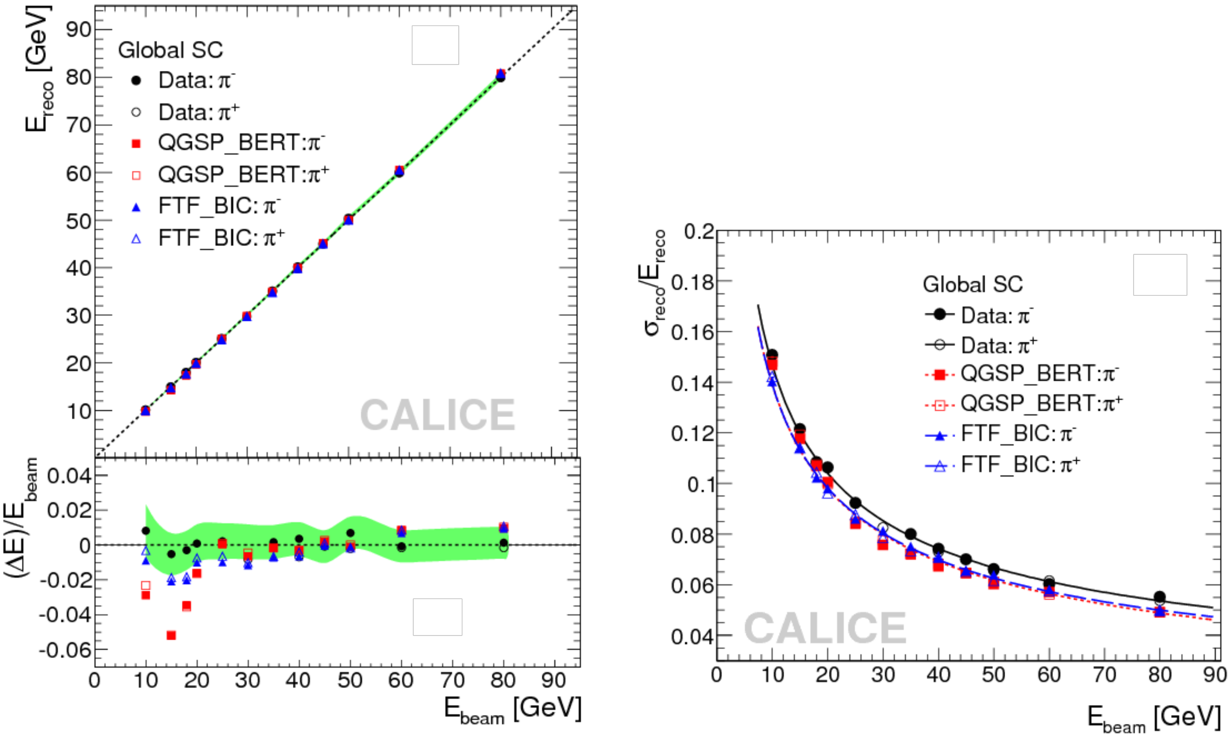


Figure 3.14 – AHCAL physics prototype response to pions (left) and its energy resolution (right) after applying the software compensation [130].

dimensions of 1168 × 1168 mm^2. The first eight are 19 mm thick, the rest have a thickness of 102 mm. The total thickness of TCMT is 5.5 nuclear interaction lengths. Scintillator strips in different layers have alternating orientations. Each 1 × 1 m^2 layer is segmented into 20 strips, there are 320 strips in total. Additional details on TCMT are presented in [67].

3.3.3 Digital and Semi-Digital HCAL

Both DHCAL and SDHCAL first prototypes had electronics embedded into the active layer. The design of the DHCAL layers was chosen to fit the CALICE AHCAL and the tail catcher structures containing steel absorber plates. The DHCAL and AHCAL volumes are therefore the same, ∼1 m^3.

The DHCAL prototype was built in 2008–2010. In total, 52 layers have been assembled, 38 were inserted to AHCAL structure and 14 to the tail catcher. One layer was composed of 3 RPC chambers. DCAL III ASIC chip was used for the readout. DHCAL prototype and the reconstructed image of the pion shower are shown in Figure 3.16. The prototype has been tested at CERN and FermiLab [139–141]. Additional information on DHCAL technologies can be found in [142–144].

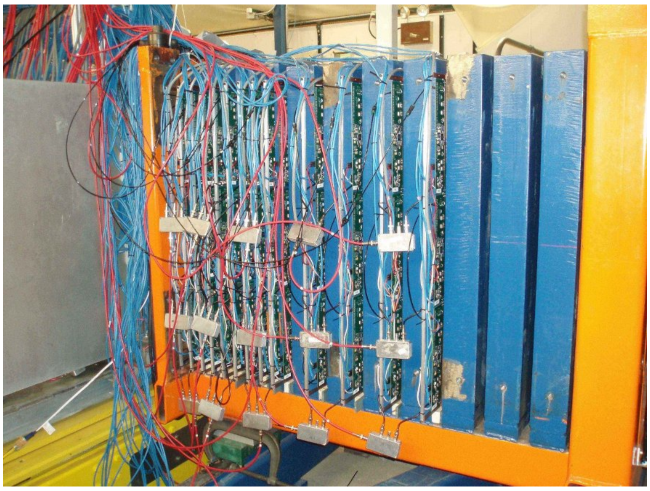


Figure 3.15 – CALICE tail catcher / muon tracker (TCMT).

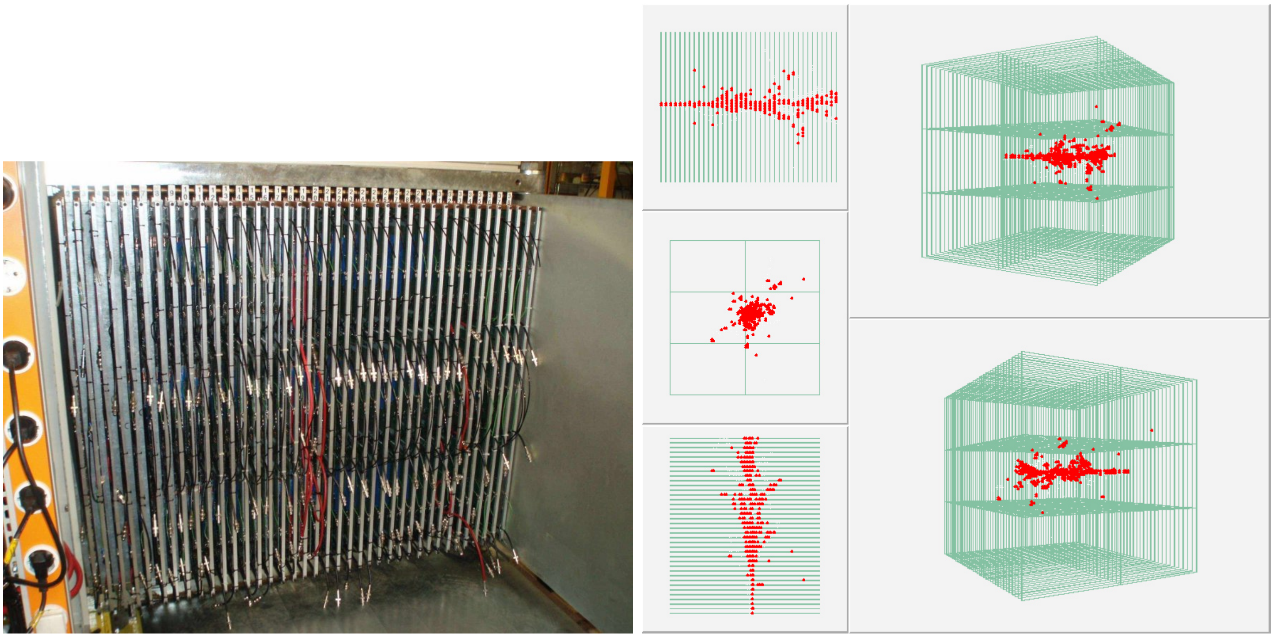


Figure 3.16 – DHCAL prototype with about 500 000 channels (left) and the reconstructed image of 10 GeV pion shower (right).

The SDHCAL prototype was built in 2011. It has 48 layers made of the Glass Resistive Plate Chambers (GRPC) with an area of about 1×1 m^2. The total volume is 1.3 m^3. The prototype and the image of one shower are shown in Figure 3.17 (left). HARDROC ASIC chip reads out the signals and compares them with 3 thresholds. The PCB equipped with 144 chips is shown in Figure 3.17 (right). The prototype was tested at CERN beams several times. Figure 3.18 shows the SDHCAL prototype linearity and the energy resolution. Additional details can be found in [61, 82, 138]. Instead of GRPC chambers one can also use gas electron multiplier (GEM) or micromegas chambers [145]. In principle, they can also be used in the analog mode and can stand higher rates than RPC. The disadvantage is higher cost.

DHCAL and SDHCAL prototypes have a record number of channels, about 500 000 each.

3.4 Beam tests of CALICE prototypes

Starting from 2005 CALICE physics prototypes are being tested in beams. The main tests of the physics and DHCAL prototypes (the latter had embedded electronics from the beginning) are summarized in Table 3.1.

For the shower separation study presented in this thesis in Chapter 4, we have used the data from the beam tests performed in 2007 at CERN SPS with SiW ECAL + AHCAL + TCMT (see Figure 3.19) and in 2011 at FNAL with SiW ECAL + DHCAL.

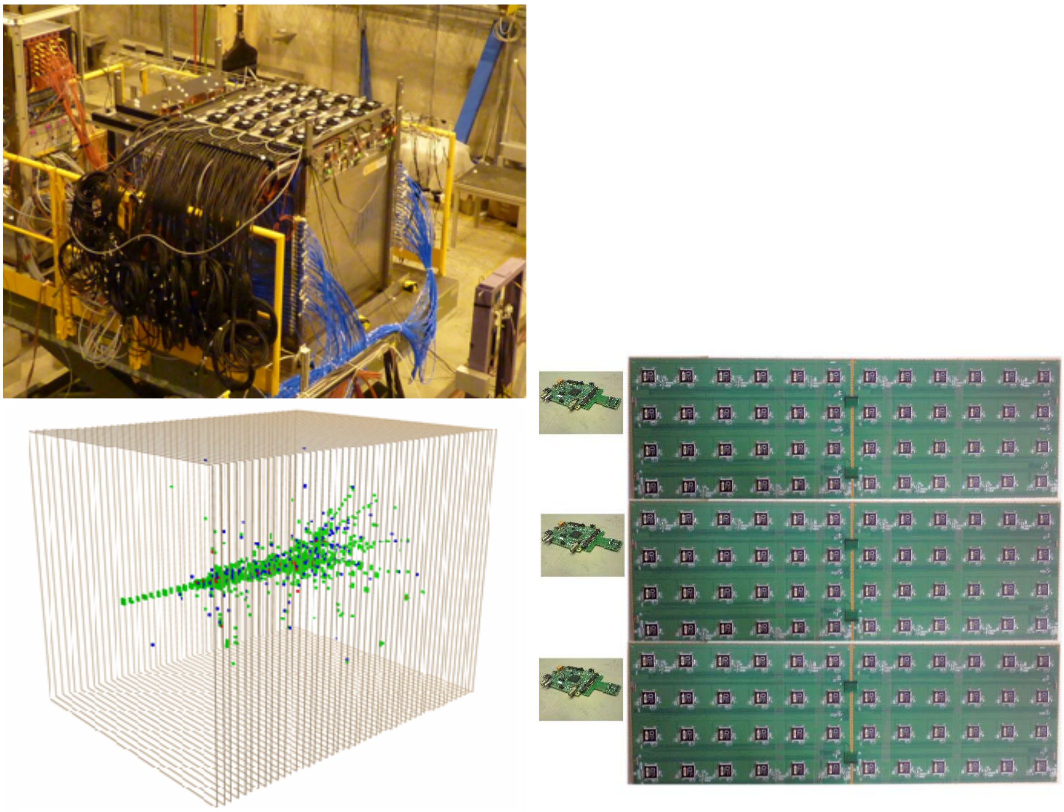


Figure 3.17 – SDHCAL prototype and one example of 30 GeV pion shower image (left), readout of 1 × 1 m^2 active layer by embedded 144 HARDROC chips (right).

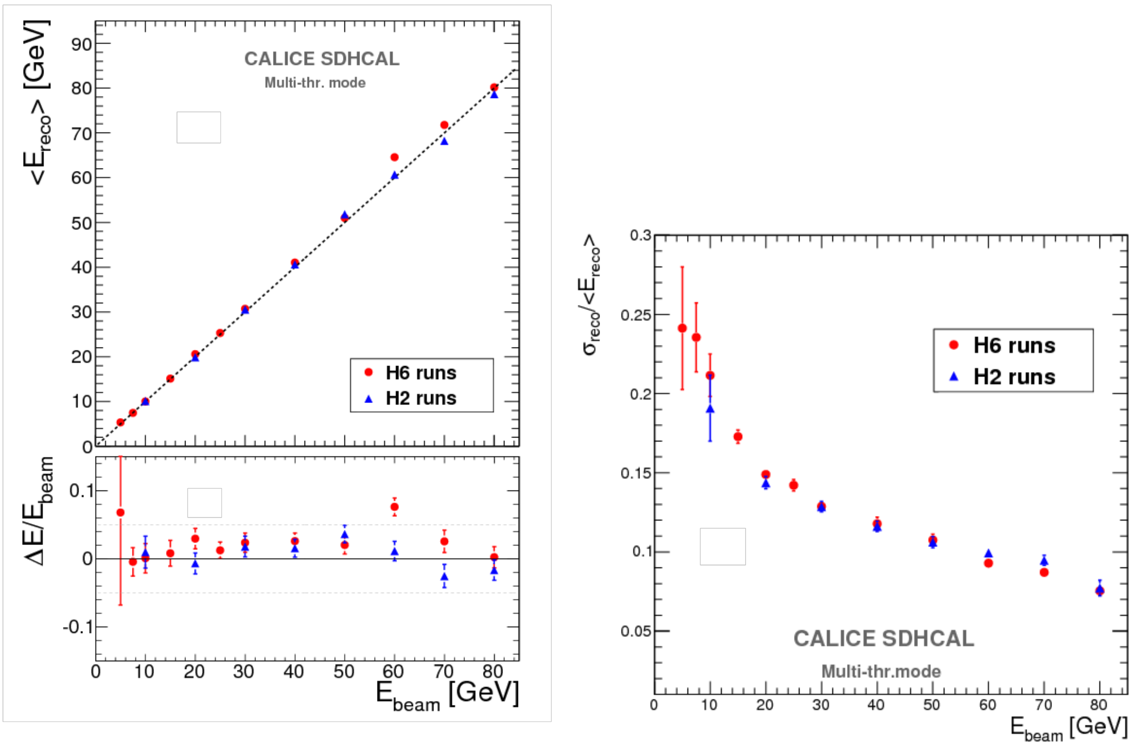


Figure 3.18 – Linearity of SDHCAL response to pion showers (left) and the relative resolution as a function of the pion beam energy (right) [138].

Year	Setup	TB facility
2005	SiW ECAL	CERN
2006	SiW ECAL + AHCAL + TCMT	CERN
2007	SiW ECAL + AHCAL + TCMT	CERN
2008	SiW ECAL + AHCAL + TCMT	FNAL
2009	Sc ECAL + AHCAL + TCMT	FNAL
2010	DHCAL	FNAL
2011	SiW ECAL + DHCAL	FNAL

Table 3.1 – CALICE beam test campaigns.

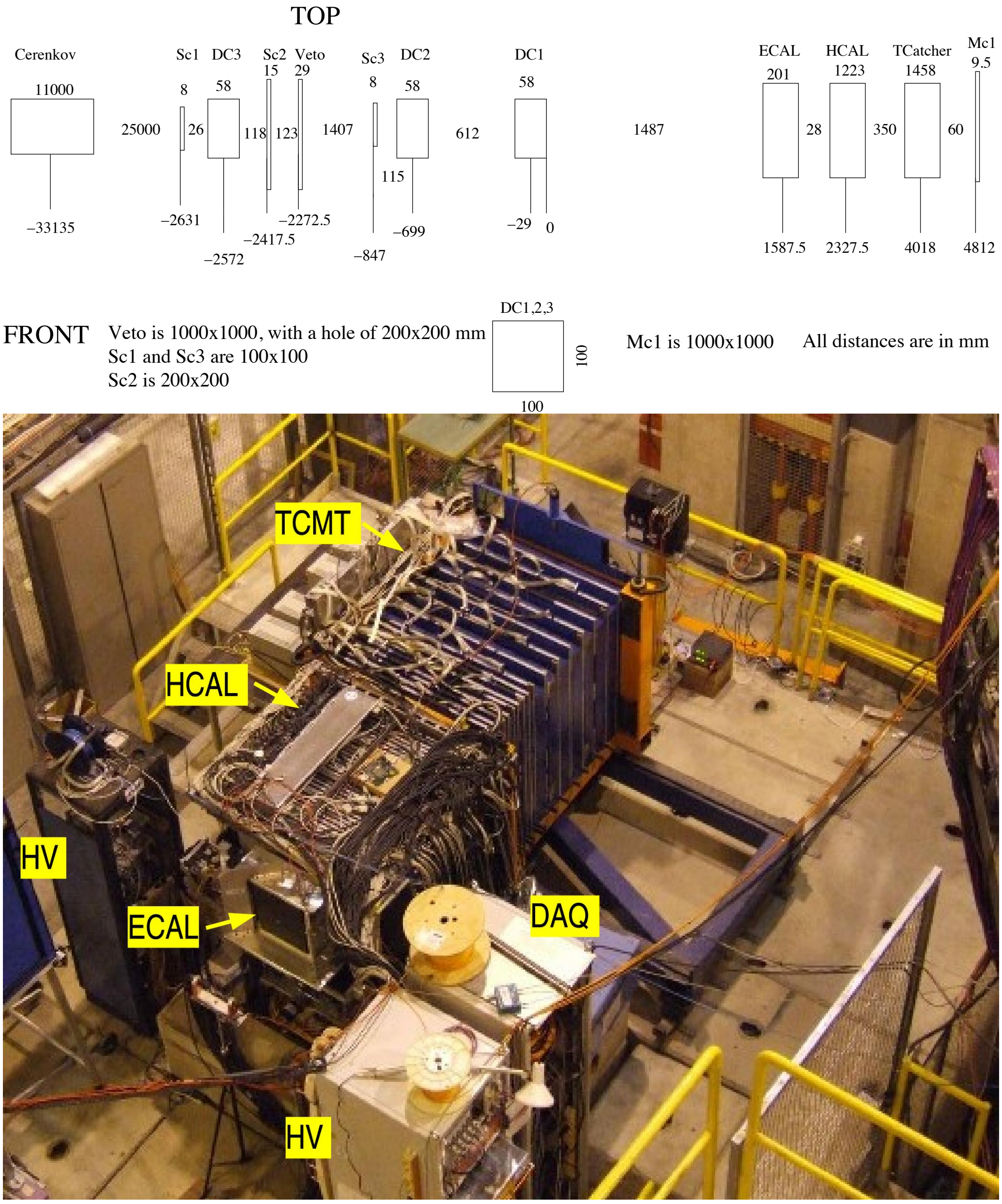


Figure 3.19 – CERN SPS beam line in 2007 (top) and the CALICE setup composed of SiW ECAL and AHCAL physics prototypes and TCMT (bottom).

Chapter 4

Separation of two overlapped electromagnetic or electromagnetic-hadronic showers using Pandora, Garlic and Arbor

4.1 Introduction

PFA jet energy resolution is dominated by the HCAL intrinsic resolution for the jet energies below ~100 GeV and by "confusion" in assigning the calorimeter deposits to the right particles at higher energies. The latter effect is important when the showers of highly boosted jet particles start to overlap in the calorimeters. In this case, the automatic "pattern recognition" sometimes fails to resolve two close showers or can reconstruct a fragment of the shower as a separate neutral hadron. In the former case, merging two photon showers into one does not degrade the jet resolution, since what matters for PFA is the total photonic energy. Merging of photon and hadron showers always degrades the resolution but to different extent, as schematically explained below:

- charged hadron + $\gamma \to \gamma$: leads to large loss of charged hadron energy;
- charged hadron + $\gamma \to$ charged hadron: degrades the resolution only slightly because of the non-compensation of the calorimeter, as the photon energy is counted as hadronic resulting to an overestimation of the jet energy;
- neutral hadron + $\gamma \to \gamma$ (neutral hadron): bias the jet energy down (up) again only due to the non-compensation.

Note that, for example, the separation of two close photon showers is important not only for the jets but eg. for π^0 reconstruction. One example is the classification of $\tau^\pm \to \pi^\pm n\pi^0\nu_\tau(\bar{\nu_\tau})$ decays (ie. determination of $n = 0, 1, 2$) from $h \to \tau^+\tau^-$. This is important for the statistical measurement of $\tau^\pm$ polarization and Higgs CP parity [102].

Confusion is a natural limit of PFA approach at high energies. It is important to verify that the present shower simulation models correctly describe it. The limit is fully dependent on the ability to separate individual showers, therefore CALICE has already studied the two hadron separation probability using the SiW ECAL and AHCAL physics prototype data [26]. The main topic of this thesis is a similar study of the separation of two electromagnetic or electromagnetic - hadronic showers. The difference with [26] is that this analysis is more ECAL oriented. Both [26] and the present study confirm that the results obtained with data are well reproduced by Geant4 simulation.

Another important aspect of the current study is to check the performance of various PFA programs and their versions in different conditions. We report the results for three currently existing PFA programs, Pandora [44], Garlic [86] and Arbor [84]. Pandora gives the best jet energy resolution and is used for many ILD optimization studies, Garlic is developed for the photon reconstruction in ECAL and Arbor is an alternative PFA algorithm which reconstructs showers as "trees". This study allows to compare the performance of these programs. When studying new version of Pandora we have observed a degradation of the performance at rather large separation distances of ∼10 cm. After contacting Pandora developers, they found and fixed a bug and the performance was returned to normal. This effect was impossible to find using only the overall jet energy resolution.

In addition to CALICE physics prototype data, we have studied the separation with the ILD simulation in the central barrel region. The study is performed both for the nominal 5×5 mm^2 ECAL granularity and with 2.5×2.5 mm^2. The latter should be better, but, surprisingly, we have found that this finer granularity degraded the probability of the separation except for Arbor, which demonstrated equal performance. This inconsistency means that eg. the best PFA program Pandora is tuned for the nominal 5×5 mm^2 granularity. There is still some room for PFA improvement for finer ECAL pixels and, possibly, some room for further ECAL optimization.

The content of this chapter is the following.

- First, we present the results on the separation of two overlapped electromagnetic showers in the CALICE SiW ECAL physics prototype alone;

- then, we extend the analysis to the lectromagnetic – hadronic shower separation using CALICE SiW ECAL and AHCAL physics prototype data;

- finally, both analyses are repeated for ILD Monte Carlo (MC) simulation with 5×5 mm^2 and 2.5×2.5 mm^2 ECAL granularity for the ILD baseline models with AHCAL and SDHCAL.

Physics prototype data have been collected in the beam tests (TB) at CERN in 2007 and at FermiLab in 2011. The TB data results are compared with TB Monte Carlo simulations.

The events with two showers are obtained by single shower event mixing. Before overlay, one event is shifted by an integer number of pixels and the separation probability is determined as a function of the distance between shower centers.

The photon showers are approximated as positron showers reconstructed as photons. This approach is compared with the pure photon showers in Monte Carlo. There was no magnetic field in the beam tests. Note, that we did not make any attempt to emulate the photon showers with positron hits excluding those produced by e^+ dE/dx losses in the first layer(s) before the first e^+ interaction. As one can see later, the separation power of Pandora and Garlic almost does not depend on those hits. For Arbor this is not the case, however. This is because we remove the multishower events by requiring only one hit in the first ECAL layer. This hit serves as a perfect seed for the Arbor tree. Therefore, the Arbor performance is overestimated on e^+ sample and this agrees with the MC e^+ simulation. The true Arbor performance on photons should be taken from MC γ sample.

4.1.1 Detectors and their software description

The TB experimental setups consist of the SiW ECAL physics prototype, the hadronic calorimeter (AHCAL physics prototype in CERN'07 tests and DHCAL prototype in FNAL'11) and the tail catcher and muon tracker (TCMT). Only ECAL and AHCAL data are used in this analysis. Full SiW ECAL prototype has 30 active silicon readout layers each with an active area of 18×18 cm^2, segmented into 3×3 diode matrices with 1×1 cm^2 diodes (9720 channels in total). During CERN'07 tests six front ECAL layers were not fully equipped with the sensors. The bottom 3×1 row was missing [113, 114, 153, 154]. This was modelled in the simulation. In FNAL'11 tests all layers were complete.

In the ILD simulation, the SiW ECAL is combined with either AHCAL or SDHCAL. We study only the central barrel part located above the interaction point. The baseline ILD ECAL pixel is 5×5 mm^2, four times smaller than in the CALICE physics prototype.

The ECAL trigger threshold in the ILD simulation is set at 0.5 MIP level. Note, that this cut affects $5{\times}5$ mm^2 and 2.5×2.5 mm^2 pixels slightly differently. The pixel can be fired by two or more energy depositions smaller than 0.5 MIP if their sum exceeds the 0.5 MIP threshold. For 2.5×2.5 mm^2 granularity this effect is reduced due to smaller pixel area. Therefore, the 0.5 MIP cut is effectively slightly stronger for smaller pixels. The difference is expected to be small, however, since the 0.5 MIP threshold is relatively low.

Since all PFA programs have been written for ILD, we can not apply them directly to the CALICE prototype data. Instead, all hits from the latter are transferred to the ILD hits collections, so that vertical layers of the prototype are mapped to the same horizontal part of the ILD barrel layers above the interaction point. A small mismatch between CALICE SiW ECAL prototype and ILD ECAL will be discussed in the following.

Overall, in the analysis we are using the data collected with 2 CALICE setups (ECAL or ECAL+AHCAL prototypes), 2 MC simulated setups and 4 different ILD models. The simulation of events is done in the Mokka [71] framework based on Geant4. The digitization, reconstruction and some part of the analysis are done inside the Marlin [74] framework for ILC. The final analysis is done using the R [146] computing language oriented on data analysis and statistics. In spite of the fact that it is rarely used in High Energy Physics, we find it very useful.

4.2 Data processing

Here is a brief description of the software used in this analysis:

LCIO [69] (Linear Collider Input/Output) is a persistence framework and data model that has been adopted as a standard by the ILC community. It defines an abstract event data model with hits, tracks, clusters etc. and a concrete file format to store the data.

Gear [72] (Geometry API for Reconstruction) is a geometry description toolkit for ILC reconstruction software. For the simulation of the detector response one needs a detailed description of the material distribution in space.

Simulation tools:

- Geant4 [70] (for GEometry ANd Tracking) is a toolkit for simulating the passage of particles through matter. It includes a complete range of functionalities including tracking, geometry, physics models and hits. The physics processes offered cover a comprehensive range, including electromagnetic, hadronic and optical processes, a large set of long-lived particles, materials and elements, over a wide energy range. It has been designed and constructed to expose the physics models utilised, to handle complex geometries, and to enable its easy adaptation for optimal use in different sets of applications.

- Mokka [71] is a Geant4 based framework used for event simulation (uses event Generator, particle list and detector model as input). The outputs of Mokka are a LCIO file containing MCParticles and Hits deposited in the SubDetectors, and the Gear XML file containing the geometry used for simulation.

- G4ParticleGun is a simple one particle Geant4 event generator which propagates through the detector the particle with the given initial position, energy, direction and type.

- QGSP (Quark-Gluon String Precompound model) is a basic Geant4 physics list applying the quark gluon string model for high energy interactions of protons, neutrons, pions, kaons and nuclei. The high energy interaction creates an excited nucleus, which is passed to the code modeling the nuclear deexcitation. QGSP_BERT is like QGSP, but using the Geant4 Bertini cascade model for primary protons, neutrons, pions and kaons below ≈ 10 GeV. The Bertini model produces more secondary neutrons and protons than the Geant4 low energy parameterised (LEP) model, yielding a better agreement to experimental data.

Analysis and reconstruction are done inside Marlin [74] (Modular Analysis and Reconstruction for the LINear collider), which is a modular C++ application framework for the analysis of LCIO data. It provides a platform for the distributed development of ILC detector reconstruction and analysis algorithms. A few specific Marlin processors were written for the current analysis. ILCsoft and CALICEsoft are implemented inside the Marlin framework:

- ILCsoft [68] is a set of software packages developed for the International Linear Collider.

- CALICEsoft [75] are software packages developed for the CALICE collaboration internal studies for ILC project.

To simulate the CERN'07 (FNAL'11) TB setup TBCERN0807_p0709 (TBFNAL0508_p1211) Geant4 (geant4-10-01-patch-03) Mokka model is used. The latter model contains SiW ECAL together with AHCAL, but only ECAL hits are used in the analysis. For ILD Geant4 (geant4-09-05-patch-02) simulation we use ILD_o1_v05 (ILD_o2_v05) models with AHCAL (SDHCAL). The generated events are digitized using standard CALICE TB and ILD digitizers (for the latter we use ILDCaloDigi for ECAL, AHCAL and SimDigital for SDHCAL).

Events are reconstructed using 3 different PFA reconstruction programs: Pandora, Garlic and Arbor. They are implemented in ILCsoft (v01-17-09) as individual Marlin processors. Every algorithm uses the same list of tracking processors : ClupatraProcessor, SiliconTracking_MarlinTrk, ForwardTracking, TrackSubsetProcessor, FullLDCTracking_MarlinTrk [78, 79]. A short description of the reconstruction tools is given below:

- Pandora [44] is a standard ILD reconstruction program, which is used for many ILD optimization studies and gives the best jet energy resolution. Pandora tries to associate the reconstructed clusters with tracks, to identify whether we have charged (with correct energy) or neutral particles. In case of a big mismatch between the track and the associated cluster energy, Pandora starts the reclustering algorithm (merging or splitting of clusters) to get maximally correct correspondence between track and cluster energy. Pandora is best optimized for standard ILD detector with 5×5 mm^2 ECAL and 3×3 cm^2 AHCAL. In the current analysis, it is used for the reconstruction of the standalone ECAL (FNAL'11), ECAL+AHCAL (CERN'07) and ILD model with AHCAL. We also compare the performance of old Pandora version 00-14 with the recent one 02-04 with improved photon reconstruction. In the baseline Pandora works only with AHCAL. For the current analysis we did not make any attempt to run it over SDHCAL.

- Arbor [84] is an alternative to Pandora jet reconstruction program. Arbor starts from the seeds in the first layers and builds tree-like clusters. It is best optimized for high granularity, like 5×5 mm^2 ECAL and 10×10 mm^2 SDHCAL. Currently, it cannot reconstruct AHCAL with 30×30 mm^2 or coarser pixels. March'15 version of Arbor is used for the analysis. This version of Arbor does not have PID for neutral particles: all photons and neutrons are stored in the same collection in LCIO.

- Garlic [86] (GAmma Reconstruction at a LInear Collider experiment) is a program designed to identify photons produced at the interaction point. It makes use of the characteristic form of electromagnetic showers in ECAL: the narrow width of the shower and the specific longitudinal profile. Currently Garlic does not provide full event reconstruction, it does not use HCAL hits for the analysis. Garlic algorithm searches for the cluster core and forms a cluster around this core. Then it tries to associate the cluster with tracks to make e^+/e^- identification. Further, to form γ clusters it excludes all the hits around track extrapolation in the ECAL, and builds the core and the cluster. Later, depending on the Garlic version, clusters pass a neural network selection (version 2.11) or additional cuts (version 3.0.3). The Garlic version 2.10 has a cut on the angle between the cluster main axis and its direction to the interaction point (IP). In our case, the TB electromagnetic clusters are transferred into ILD geometry such that they point initially to IP. After large shifts, however, they do not point to IP anymore. Therefore we contacted the Garlic developer who then released version 2.11 without the angle cut. Only this version was used for TB analysis. For ILD MC both v2.11 and v3.0.3 were used.

By default Pandora is optimized for 5×5 mm^2 ECAL granularity. After the discussion with Pandora group [92] the following changes have been made for 2.5×2.5 mm^2:

- the photon likelihood file was retrained using a sample of 10^4 $Z \to u, d, s$ decays ($\sqrt{s} = 500$ GeV);
- the following line has been replaced in the Pandora XML steering file:

  ```
  <algorithm type = "PhotonFragmentRemoval"/>
  ```

  with

  ```
  <algorithm type = "PhotonFragmentRemoval">
      <CloseHitDistance1> 5.0 </CloseHitDistance1>
      <CloseHitDistance2> 2.5 </CloseHitDistance2>
  </algorithm>    .
  ```

The latter changes the default `CloseHitDistance1` and `CloseHitDistance2` values, 10 and 5 mm, to 5 and 2.5 mm in the `PhotonFragmentRemoval` algorithm. This reduces the merging of photon showers and improves the separation performance for 2.5×2.5 mm^2 granularity.

R is a programming language and software environment for data analysis, statistical computing and graphics [146–150]. We decide to use it for the final analysis instead of more common tools, like ROOT, C++ or python. There are thousands of R packages built around its core. We use `data.table` package [147] which allows to analyze large structured arrays of data. The final performance plots shown in the text and in Appendix A.2, are created with the R `ggplot2` package [148].

4.3 Event collection, selection & mixing

4.3.1 Event collection

For the analysis we use the single particle events (positrons e^+ and pions π^+) collected during CALICE beam tests at CERN (2007) and FermiLab (2011). A significant difference between CALICE beam test setup and ILD detector (in addition to the reduced granularity: 10×10 mm^2 instead of 5×5 mm^2) is the absence of magnetic field in physics prototypes. Photons γ are emulated with TB positrons e^+ (EM clusters are reconstructed as photons). The main difference in ECAL between positrons e^+ and photons γ is that the positrons start leaving dE/dx hits in the ECAL immediately while the photon starts showering only after the first interaction. There is also a small difference in the longitudinal profiles, as the photon attenuation length λ is equal to 9/7 of the positron radiation length X_0. These differences have been neglected and we use the unmodified positron shower as the input for the photon reconstruction algorithms. Note that this is done for MC in the same way as for data. We use only ECAL information for photons and assume that photons are entirely absorbed in ECAL, neglecting any shower leakage to HCAL (note that in this way we are not sensitive to HCAL noise). Pion events are taken only from CERN'07 data with AHCAL. TCMT information is not taken into account, because the reconstruction algorithms, designed for ILD, are not able to work with it. There was no tracker device in the TB and we reconstructed the e^+ entry point from the energy barycenter of the shower. For π^+, the entry point is simply defined as the center of the ECAL pixel fired in the first layer. The fluctuations of the π^+ shower energy barycenter are larger than the uncertainty of this measurement. Table 4.1 lists the runs used in the analysis and the corresponding beam parameters.

To confirm the TB results, π^+, e^+ and γ events are simulated in the corresponding geometries, TBCERN0807_p0709 and TBFNAL0508_p1211 [153, 154] with QGSP_BERT physics list.

Geometry	Energy, particle	Other details
TB FNAL'11	4 GeV e^+ 12 GeV e^+ 25 GeV e^+ 32 GeV e^+	10×10 mm^2 ECAL, HCAL: not used
MC FNAL'11	4 GeV e^+ & γ 12 GeV e^+ & γ 25 GeV e^+ & γ 32 GeV e^+ & γ	10×10 mm^2 ECAL, HCAL: not used
TB CERN'07	10 GeV e^+ 25 GeV e^+ 30 GeV π^+	10×10 mm^2 ECAL, AHCAL only for π^+
MC CERN'07	10 GeV e^+ 25 GeV e^+ 30 GeV π^+	10×10 mm^2 ECAL, AHCAL only for π^+
MC ILD	4 GeV γ 12 GeV γ 25 GeV γ 32 GeV γ 6 GeV γ 25 GeV γ 20 GeV π^+ 30 GeV π^+	5×5 mm^2 or 2.5×2.5 mm^2 ECAL, AHCAL or SDHCAL

Table 4.1 – Data samples used in this analysis.

π^+ and γ showers with the same energies as in TB are simulated in the ILD MC for both ECAL granularities, 5×5 mm^2 and 2.5×2.5 mm^2. Pandora reconstruction is used for AHCAL with 3×3 cm^2 tiles and Arbor for SDHCAL with 1×1 cm^2 cells.

All information about the data samples and the MC simulations is summarized in Table 4.1.

4.3.2 Event selection

Due to the capacitive coupling of the silicon sensor guard ring with the peripheric pixels, a large energy deposition at the sensor boundary sometimes produces so-called "square"-events, where many boundary pixels are fired by the capacitive cross talk (Figure 4.1). Such events were clearly visible in the ECAL physics prototype data. The following optimizations of the guard ring design significantly improved the situation, as it will be discussed in the next chapter. To exclude “square” events, the following two cuts have been applied for both e^+ and π^+ samples:

- number of hits per wafer ≤ 30;
- number of border hits per wafer ≤ 15.

The cuts also remove the events with large noise in one of the ECAL layers. MC simulation shows that the real showers are unaffected by these cuts, see Figure 4.2.

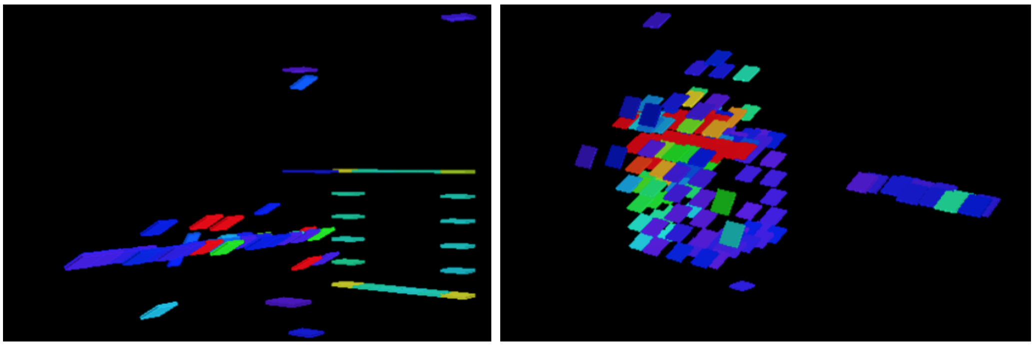


Figure 4.1 – Examples of “square” events in CALICE SiW ECAL physics prototype: 4 GeV pion (left) with only one affected sensor and 50 GeV event (right) with squares in 4 consecutive layers.

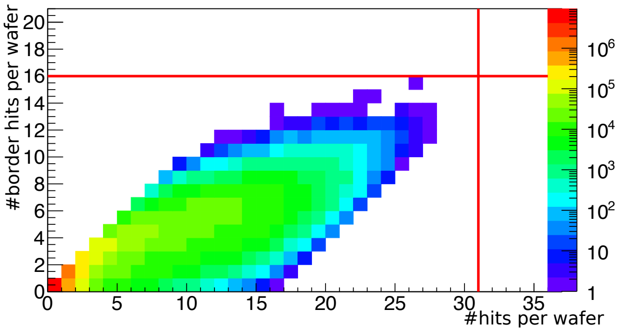


Figure 4.2 – The number of border hits per sensor (in total, each sensor has 20 border cells) versus the total number of sensor hits (there are 36 cells per sensor) in MC simulation for 32 GeV e^+ FNAL’11 sample. Red lines show the position of the cuts.

After that we applied the following cuts:

- Against multi-particle events: the first ECAL layer should have exactly one hit and its energy should not exceed the energy corresponding to 2 MIPs. These two cuts are more important for FNAL’11 data, especially, at lower energies, see Figures A.5 - A.12 in Appendix A.1, where the distributions are shown together with the cut values indicated by red vertical lines. Note that these cuts also suppress the early shower events with e^+ accompanied by the first e^+e^- pair in the first layer.
- To improve the transverse containment of the shower, to reduce the transverse shower leakage and the effect of the dead area at the silicon sensor boundaries due to the guard rings and inter-sensor gaps, the shower center is required to be within ±2 cm in both X and Y from the ECAL center, see Figures A.13 - A.16. Note, the front central ECAL sensor is 6×6 cm^2. The shower center is defined as the ECAL energy barycenter for e^+ and as the center of the fired pixel in the first ECAL layer for π^+.
- To further suppress events with noisy layers, we apply the energy dependent cut on the maximal ECAL layer energy for both e^+ and π^+, see Figures A.21 - A.24. For the former, additionally, the ECAL layer with the maximal energy is required to be in the range $4 \ldots 20$, see Figures A.17 - A.20. In FNAL’11 data the second and the third layers from the ECAL back were noisy, so the last three layers were excluded (out of 30). The same was done in FNAL’11 MC.
- To remove π^+ from the e^+ sample and muons from both π^+ and e^+ samples, a total energy cut is applied on the ECAL energy for e^+ and on ECAL+AHCAL for π^+, see Figures A.25 - A.28.

The full list of cuts is given in Tables 4.3 - 4.9, all corresponding distributions can be found in Appendix A.1. Note that all cuts are applied consecutively, i.e. each distribution corresponds to the data sample remaining after the cuts shown on previous plots.

MC TB cuts are exactly the same as for the TB data (in spite of the fact that e.g. there are no square events in the simulation). For photons we apply slightly different cuts than for positrons. Namely, we do not require one hit in the first layer and we double the energy limit for the first layer because the photon produces two charged particles after the first interaction. The cuts and the corresponding distributions for γ sample are shown in Figures A.5 - A.7.

There are no cuts for events simulated in the ILD except the events with the particles decaying before the ECAL are excluded.

For the cuts defined above and in the following analysis we use the two step energy calibration procedure for TB data. First, the pedestal subtracted signals (A_i) for every channel i are expressed in MIP-like signals from muons.

$$N_i^{ECAL,HCAL} = C_{MIP,i}^{ECAL,HCAL} A_i^{ECAL,HCAL} \ . \tag{4.1}$$

For CERN'07 data we use the standard CALICE calibrations. SiW ECAL calibration procedure is described in [116] and AHCAL calibration in [60]. For FNAL'11, the ECAL calibration was performed in [155] also using the muon data sample.

In the second step, we apply another global calibration factor (separately for CERN'07 and FNAL'11 data) to set the 30 GeV pion MIP-like signal to one. This equalizes CERN'07 and FNAL'11 calibrations. Note that, after that the "MIP" definition for ECAL corresponds to 30 GeV pion and is therefore slightly larger than the correct minimal value of dE/dx losses. This small difference affects only the cuts expressed in MIPs.

On top of that, MIP-to-GeV conversion factors for ECAL and HCAL are applied to place the peaks of the energy distributions for electrons and pions to the correct beam values. For the ECAL layers with double and triple tungsten thickness, the factor is doubled and tripled, respectively:

$$E^{e^+,\pi^+} = C_{ECAL}^{e^+,\pi^+}\left(\sum_{layer=1}^{10} N_{layer}^{ECAL} + 2\sum_{layer=11}^{20} N_{layer}^{ECAL} + 3\sum_{layer=21}^{30} N_{layer}^{ECAL}\right) + C_{HCAL}^{e^+,\pi^+}\sum_{layer=1}^{38} N_{layer}^{HCAL} \ , \tag{4.2}$$

where N_{layer} is the sum of all N_i in this layer.

For photons the same factors as for the electrons are used. Note, in [156] it was shown that for the best energy resolution the weights 1:2:3 are not optimal for the complex hadronic showers. Our study is devoted only to the particle separation which is almost insensitive to higher order corrections in the energy reconstruction, however. In addition, since a priori we have no information on the particle identification, we should apply the same weights for π^+ and $e^+(\gamma)$ hits. Therefore, following [26,152], the weights have been fixed at 1:2:3.

The MIP-to-GeV conversion factors $C_{ECAL,HCAL}^{e^+,\pi^+}$ are summarized in Table 4.2. For AHCAL the standard value of $C_{HCAL}^{e^+}$ is $0.0236 GeV/MIP$ [130,134]. Since we do not use AHCAL hits for e^+ we set $C_{HCAL}^{e^+}$ to zero in the formula above, however. AHCAL is not a compensated calorimeter, its response to e^+ is about 20% larger than to π^+ at the same energy [130]. In addition, the leakage to TCMT has been neglected. Due to both effects $C_{HCAL}^{\pi^+}$ is larger than $C_{HCAL}^{e^+}$ by about 23%, as measured with π^+ sample at 30 GeV, see Table 4.2. After ADC-to-MIP and MIP-to-GeV conversions applied consecutively, the precise definition of MIP does not matter for AHCAL. For ECAL it also does not influence the total shower energy, but only affects the cuts expressed in MIPs. The value of $C_{ECAL}^{e^+}$ was tuned to have the right energy for e^+ samples in our energy range. Due to non-compensation, the ECAL response to e^+ is also larger than to π^+. We neglected this difference, again because the reconstruction programs should not have an a priori information on the type of the shower (electromagnetic or hadronic). Instead, since photons dominate in ECAL, we assume that all hits originate from electromagnetic showers. For π^+ we use the same calibration constant as for e^+: $C_{ECAL}^{\pi^+} = C_{ECAL}^{e^+}$, see Table 4.2. The only place where the ECAL pion shower energy is needed is the Pandora reclustering algorithm where the track momentum is compared to the cluster energy and in the case of a big mismatch, the algorithm tries to make the clusterization differently. For e^+ in FNAL'11 TB data we have measured slightly larger $C_{ECAL}^{e^+} = 0.00440 GeV/MIP$ (instead of 0.00434 for CERN'07), possibly because we have excluded three last ECAL layers which were noisy, see second row in Tables 4.3 - 4.6.

Shower energies in ILD simulation are estimated with the default ILD calibrations.

Finally, for the mixing we select only events where the reconstruction program finds only one cluster. Therefore, the samples used by Pandora, Garlic and Arbor are slightly different. Namely, to study the separation power of the given program, we select only those electromagnetic shower events, where the corresponding program finds exactly one electromagnetic cluster. For Arbor we require exactly one neutral cluster, as it does not perform the particle identification. The selection of hadronic events is a little more complicated. For Pandora and Garlic we select only those events where Pandora finds exactly one hadronic cluster (as Garlic can

	ECAL	HCAL
e^+	0.00434	0.0236 / 0
π^+	0.00434	0.0290

Table 4.2 – MIP-to-GeV conversion factor $C^{e^+,\pi^+}_{ECAL,HCAL}$ from formula 4.2 in GeV/MIP. $C^{e^+}_{HCAL}$, though given, was set to zero in Eq. 4.2, as we used only ECAL to reconstruct e^+. For ECAL, MIP is defined as the signal from 30 GeV pion. For FNAL, $C^{e^+}_{ECAL} = 0.00440 GeV/MIP$, see the text.

not reconstruct hadrons). Arbor, unfortunately, often splits the single hadronic shower. Therefore, instead of one hadron shower, we require that Arbor does not find in π^+ event a *neutral* cluster with the energy barycenter positioned in ECAL. It is allowed, however, to have such clusters in HCAL.

4.3.3 CALICE TB to ILD geometry mapping

Since the PFA programs can not process TB data, the hits from the latter are transferred into the ILD barrel hits around $X = Z = 0$ in the ILD coordinates, ie. above the interaction point.

Both baseline ILD ECAL and ECAL physics prototype have 30 active layers, grouped by two and placed in alveolar mechanical structures. In the ILD TDR baseline, the first ECAL detector layer is placed in front of the absorber and plays a role of the preshower. In fact, though it is included in the simulation, it has never been used in the reconstruction.

The first ECAL physics prototype layer is behind 1.4 mm W absorber. After the requirement of only one hit in the first ECAL layer in TB, we do not use this layer anymore and do not transfer its hits to ILD. First, this allows to keep 1:1 matching between TB and ILD alveolar structures in the rest of the ECAL. On the other hand, we require at most one hit and the minimal energy deposition in the first layer, so that the shower effectively starts later. This makes the first TB layer partially equivalent to the preshower. So, since the reconstruction programs do not take the ILD preshower into account, this is the second reason to drop the “preshower” first layer hits in TB data as well.

The hits from all other 29 layers of TB ECAL are converted to the corresponding hits in 29 ILD layers (only the first layers in ILD and TB are not “visible” to PFA programs).

The procedure explained above is valid for Pandora and Garlic. For Arbor, however, we observe that in the absence of the first layer, if there are several hits in the second layer, they can be used by Arbor to start several independent showers. This results in a cluster “oversplitting”. In 32 GeV e^+ sample, up to 80% of events were split in two or more clusters. Therefore, we have made an exception for Arbor and mapped TB hits to ILD starting from the first layer, ie. for all 30 layers. The ILD preshower hits are normally included in the ECALBarrel PreShower collection and are invisible to PFA programs. Instead, here we include them to the ECALBarrel collection together with the other ECAL hits. In this way, Arbor can “see” them and process in a normal way. With the preshower layer taken into consideration only $\sim 25\%$ of events are oversplit.

After all selection cuts and geometry changes we have electromagnetic showers and pions in the ILD but without associated hits in the tracker. To complement π^+ showers with curved tracks required for PFA programs, the following procedure is used. A large sample of π^+ tracks with the given momentum is generated in ILD magnetic field with the endpoints at the ECAL face covering the region of the π^+ clusters. For every π^+ shower we select the first track having the transverse coordinates at SET tracker within ± 1 mm from π^+ hit in the first ECAL layer. The difference between the showers developed with and without magnetic field is neglected.

4.3.4 Event mixing

To form an event with two electromagnetic showers, the one from the first event is shifted along X with respect to its initial position by an integer number of pixels $(0, 1, 2, ..., 11$ cm) and overlayed with the second event. In case when the hits from two events overlap, their energies are summed. The procedure is illustrated in Figure 4.3.

For electromagnetic–π^+ case we shift the electromagnetic shower in ECAL. In this case we can not shift AHCAL pion hits by $0, 1, 2, ...$ ECAL pixels, because ECAL and AHCAL have different granularity (and, in addition, AHCAL granularity differs in the central and in the peripheral / backward regions).

The position of electromagnetic shower is taken as its ECAL energy barycenter. For π^+ cluster in TB data and MC, instead of the barycenter, we use the entrance point in the ECAL, i.e. the position of the hit in the first ECAL layer. Note, it is required that there is only one such hit. The same was done in the previous similar study on hadron separation [26]. For ILD π^+ data we use the track extrapolation from the tracker to define the ECAL entrance point.

Due to the finite beam width in the tests there is a spread in the shower positions (see Figures A.13 - A.16) before the shifts. The spread is larger than one ECAL pixel, therefore, after the shifts we obtain a continuous

Sample	TB 4 GeV e^+	MC 4 GeV e^+	MC 4 GeV γ
Vetoed layers	layers #27, 28, 29	layers #27, 28, 29	none
# Hits in 1st Layer	1 hit	1 hit	0 or 1 hit
1st layer energy	$<$ 2.0 MIP	$<$ 2.0 MIP	$<$ 4.0 MIP
Barycenter position X	-16 … 24 mm	10 … 50 mm	10 … 50 mm
Barycenter position Y	-20 … 20 mm	-20 … 20 mm	-20 … 20 mm
Layer max energy	layers #4 … 20	layers #4 … 20	layers #4 … 20
Energy (Layer max energy)	30 … 120 MIP	30 … 120 MIP	30 … 120 MIP
Energy cut	3.0 … 5.0 GeV	3.0 … 5.0 GeV	3.0 … 5.0 GeV
Run numbers	630065, 630067		
N evt. before / after cuts	266401 / 11307	50000 / 7063	50000 / 17686

Table 4.3 – ECAL cuts for 4 GeV e^+ and γ (FNAL'11).

Sample	TB 12 GeV e^+	MC 12 GeV e^+	MC 12 GeV γ
Vetoed layers	layers #27, 28, 29	layers #27, 28, 29	none
# Hits in 1st Layer	1 hit	1 hit	0 or 1 hit
1st layer energy	$<$ 2.0 MIP	$<$ 2.0 MIP	$<$ 4.0 MIP
Barycenter position X	-16 … 24 mm	10 … 50 mm	10 … 50 mm
Barycenter position Y	-20 … 20 mm	-20 … 20 mm	-20 … 20 mm
Layer max energy	layers #4 … 20	layers #4 … 20	layers #4 … 20
Energy (Layer max energy)	40 … 400 MIP	40 … 400 MIP	40 … 400 MIP
Energy cut	10.0 … 14.0 GeV	10.0 … 14.0 GeV	10.0 … 14.0 GeV
Run numbers	630061, 630062		
N evt. before / after cuts	41174 / 1164	50000 / 4987	50000 / 14683

Table 4.4 – ECAL cuts for 12 GeV e^+ and γ (FNAL'11).

Sample	TB 25 GeV e^+	MC 25 GeV e^+	MC 25 GeV γ
Vetoed layers	layers #27, 28, 29	layers #27, 28, 29	none
# Hits in 1st Layer	1 hit	1 hit	0 or 1 hit
1st layer energy	$<$ 2.0 MIP	$<$ 2.0 MIP	$<$ 4.0 MIP
Barycenter position X	-16 … 24 mm	10 … 50 mm	10 … 50 mm
Barycenter position Y	-20 … 20 mm	-20 … 20 mm	-20 … 20 mm
Layer max energy	layers #4 … 20	layers #4 … 20	layers #4 … 20
Energy (Layer max energy)	100 … 800 MIP	100 … 800 MIP	100 … 800 MIP
Energy cut	22.5 … 27.5 GeV	22.5 … 27.5 GeV	22.5 … 27.5 GeV
Run numbers	630037, 630038, 630039		
N evt. before / after cuts	263022 / 1677	50000 / 5518	50000 / 18983

Table 4.5 – ECAL cuts for 25 GeV e^+ and γ (FNAL'11).

Sample	TB 32 GeV e^+	MC 32 GeV e^+	MC 32 GeV γ
Vetoed layers	layers #27, 28, 29	layers #27, 28, 29	none
# Hits in 1st Layer	1 hit	1 hit	0 or 1 hit
1st layer energy	$<$ 2.0 MIP	$<$ 2.0 MIP	$<$ 4.0 MIP
Barycenter position X	-16 … 24 mm	10 … 50 mm	10 … 50 mm
Barycenter position Y	-20 … 20 mm	-20 … 20 mm	-20 … 20 mm
Layer max energy	layers #4 … 20	layers #4 … 20	layers #4 … 20
Energy (Layer max energy)	100 … 1000 MIP	100 … 1000 MIP	100 … 1000 MIP
Energy cut	29.0 … 35.0 GeV	29.0 … 35.0 GeV	29.0 … 35.0 GeV
Run numbers	630033, 630034, 630036		
N evt. before / after cuts	389355 / 1065	50000 / 5151	50000 / 18781

Table 4.6 – ECAL cuts for 32 GeV e^+ and γ (FNAL'11).

distribution of distances between two electromagnetic or electromagnetic-π^+ showers. They are sampled with 1 mm binning and the separation efficiency in every bin is determined (see for example Figure 4.5). The event with two electromagnetic showers has only ECALBarrel hits, without HCAL. In pion-photon events, all π^+ ILD

Sample	TB 10 GeV e^+	MC 10 GeV e^+	MC 10 GeV γ
Vetoed layers	none	none	none
# Hits in 1st Layer	1 hit	1 hit	0 or 1 hit
1st layer energy	$<$ 2.0 MIP	$<$ 2.0 MIP	$<$ 4.0 MIP
Barycenter position X	-25 ... 15 mm	-25 ... 15 mm	-25 ... 15 mm
Barycenter position Y	-31 ... 9 mm	-31 ... 9 mm	-31 ... 9 mm
Layer max energy	layers #4 ... 20	layers #4 ... 20	layers #4 ... 20
Energy (Layer max energy)	30 ... 800 MIP	30 ... 800 MIP	30 ... 800 MIP
Energy cut	8.0 ... 12.0 GeV	8.0 ... 12.0 GeV	8.0 ... 12.0 GeV
Run numbers	331308		
N evt. before / after cuts	250156 / 30237	50000 / 9815	50000 / 21368

Table 4.7 – ECAL cuts for 10 GeV e^+ and γ (CERN'07).

Sample	TB 25 GeV e^+	MC 25 GeV e^+	MC 25 GeV γ
Vetoed layers	none	none	none
# Hits in 1st Layer	1 hit	1 hit	0 or 1 hit
1st layer energy	$<$ 2.0 MIP	$<$ 2.0 MIP	$<$ 4.0 MIP
Barycenter position X	-25 ... 15 mm	-25 ... 15 mm	-25 ... 15 mm
Barycenter position Y	-31 ... 9 mm	-31 ... 9 mm	-31 ... 9 mm
Layer max energy	layers #4 ... 20	layers #4 ... 20	layers #4 ... 20
Energy (Layer max energy)	30 ... 800 MIP	30 ... 800 MIP	30 ... 800 MIP
Energy cut	22.5 ... 27.5 GeV	22.5 ... 27.5 GeV	22.5 ... 27.5 GeV
Run numbers	331313		
N evt. before / after cuts	209055 / 37305	50000 / 8021	50000 / 19732

Table 4.8 – ECAL cuts for 25 GeV e^+ and γ (CERN'07).

Sample	TB 30 GeV π^+	MC 30 GeV π^+
Vetoed layers	none	none
# Hits in 1st Layer	1 hit	1 hit
1st layer energy	$<$ 2.0 MIP	$<$ 2.0 MIP
1st Layer hit position X, Y	4×4 central cells of the central wafer	4×4 central cells of the central wafer
Layer max energy	layers #0 ... 29	layers #0 ... 29
Energy (Layer max energy)	0 ... 1000 MIP	0 ... 1000 MIP
Total ECAL energy	0.0 ... 40.0 GeV	0.0 ... 40.0 GeV
Total energy	20.0 ... 40.0 GeV	20.0 ... 40.0 GeV
Run numbers	331341	
N evt. before / after cuts	202681 / 56933	50000 / 23887

Table 4.9 – ECAL+AHCAL cuts for 30 GeV π^+ (CERN'07).

tracker hits are also included to the event. The muon detectors are never used.

In TB MC simulation of 4 GeV e^+ (30 GeV π^+) showers we have checked that the RMS of the deviation of the reconstructed ECAL energy barycenter (hit in the first ECAL layer) from the MC true position in X or Z ILD coordinates (transverse to the direction from IP) is 1.5 mm (2.9 mm). For the pions, the RMS is in good agreement with the expected value $10/\sqrt{12} \approx 2.9$ mm, where $\sqrt{12}$ assumes continuous uniform distribution within one ECAL cell. Therefore, all our final separation efficiency curves for e.g. two 4 + 4 GeV electromagnetic (10 + 30 GeV electromagnetic - hadronic) showers are effectively smeared by about $1.5 \cdot \sqrt{2} \approx$ 2.1 mm ($\sqrt{1.5^2 \cdot 4/10 + 2.9^2} \approx 3.1$ mm, assuming $1/\sqrt{E}$ dependence of the ECAL barycenter uncertainty).

In ILD MC the events are already generated with the large transverse spread in X and Z directions. In this case, the events for the overlay are just picked up from the large simulated sample without making any additional shifts. Everything else remains the same as for the TB data and TB MC.

A short summary about all overlayed events is presented in Table 4.10. In the next sections the reconstruction procedure will be discussed.

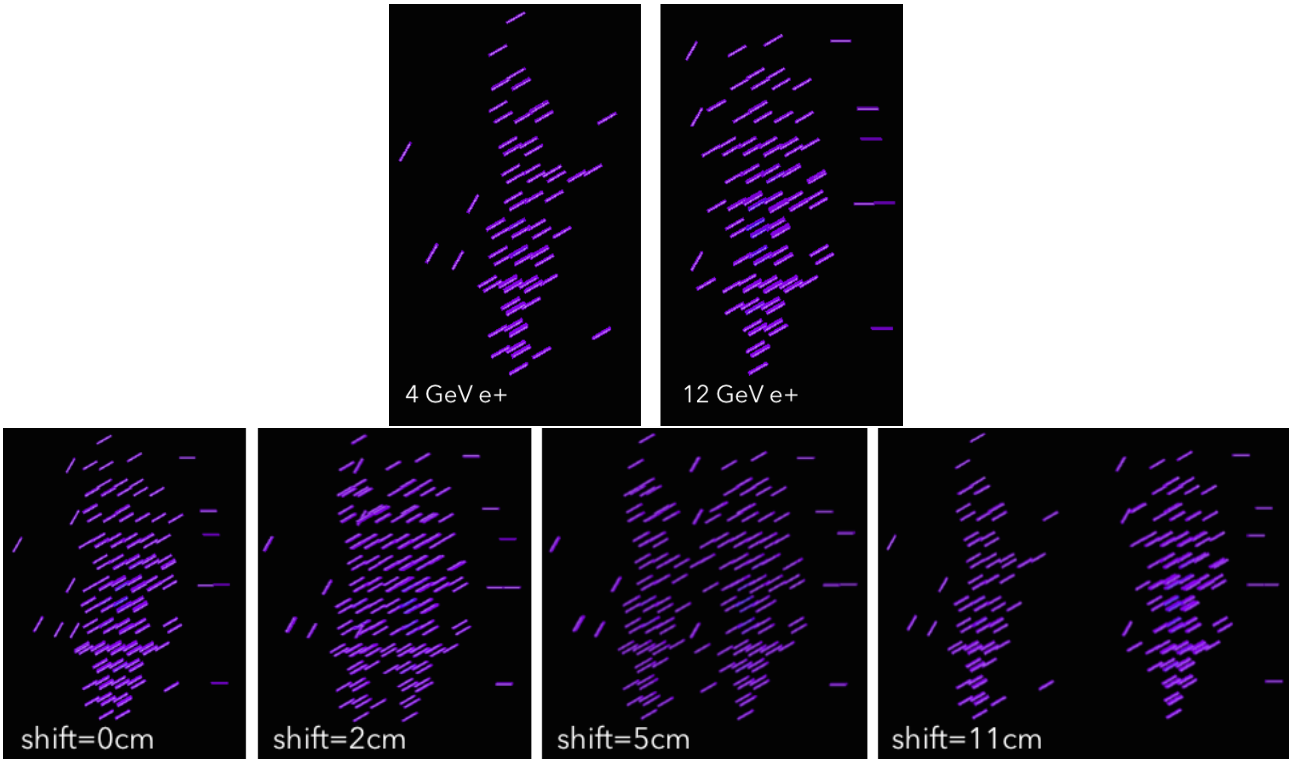


Figure 4.3 – Illustration of the event mixing procedure. Two initial showers are shown in the top. One of them is shifted by an integer number of pixels and overlayed with the second as shown in the bottom row for the displacement by 0, 2, 5 and 11 ECAL pixels.

2-particle	Energy	Geometry: particles
EM-EM	4+4 GeV	TB: e^+
	12+4 GeV	MCTB: e^+ & γ
	25+4 GeV	ILD: γ
	32+25 GeV	
Hadron-EM	30+10 GeV	TB: π^+ (ECAL+AHCAL) & e^+
	30+25 GeV	MCTB: π^+ (ECAL+AHCAL) & e^+/γ
Hadron-EM	20+ 6 GeV	ILD: π^+ & γ
	20+25 GeV	
	30+ 6 GeV	
	30+25 GeV	

Table 4.10 – List of samples obtained by event mixing.

4.4 Two shower events reconstruction

4.4.1 Separation of two electromagnetic showers in FNAL'11 data sample

The separation of two electromagnetic showers is studied using e^+ FNAL'11 data sample and e^+, γ samples of the corresponding TB MC, using Pandora (v00-14 & v02-04), Arbor and Garlic (v2.11). The mixed event is considered to be successfully reconstructed if

- it contains exactly two reconstructed electromagnetic showers;
- their energies and X, Z barycenter coordinates agree within ±20% and ±5 mm, respectively, with the energies and the coordinates reconstructed in the single shower events.

Figure 4.4 shows as an example two different e^+ clusters recorded in two different FNAL'11 events, overlayed with the shifts 0, 5 and 10 cm. The hits belonging to different reconstructed clusters have different colors. Arrows show the cluster direction. Red squares mark the Garlic seeds.

For 10 cm shift all programs correctly reconstruct the mixed events, while in the case of zero shift all fail. In the latter case, Pandora and Garlic do not find two clusters, while Arbor, thanks to two seeds in the preshower (shown in Figure 4.4 as the first layer), reconstructs two clusters but with wrong energies. For 5 cm shift, Garlic

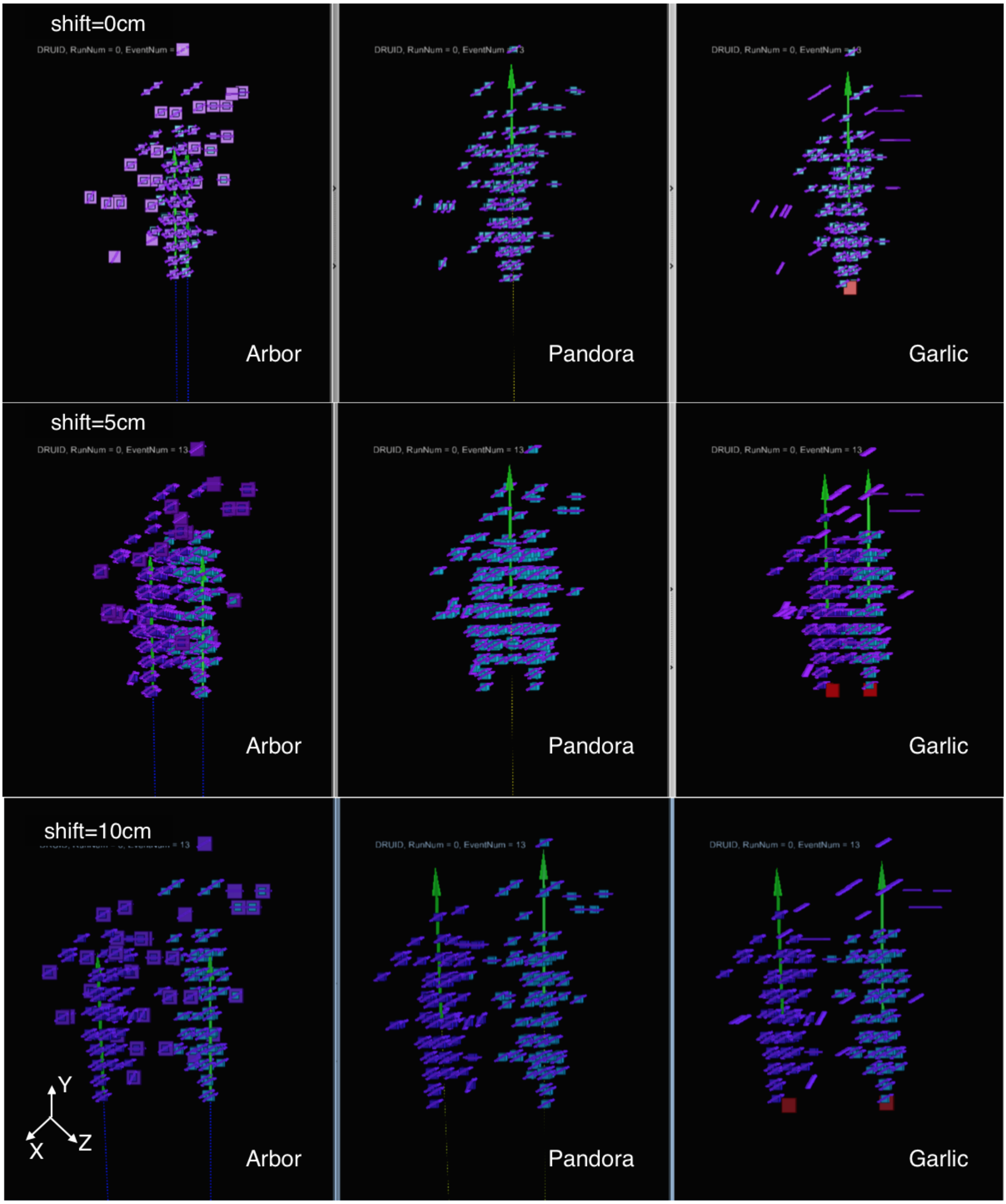


Figure 4.4 – Reconstruction of the mixed event with 4+4 GeV electromagnetic showers by Pandora, Garlic and Arbor (3 columns) for 0, 5 and 10 cm displacements (3 rows). The green arrows indicate the reconstructed shower.

and Arbor successfully reconstruct two clusters but Pandora fails. Note that Pandora is optimized for the jet energy reconstruction, not for the electromagnetic shower separation. Also, as we shall see later, it is optimized for the ILD ECAL granularity of 5×5 mm^2 and hence its separation power for 10×10 mm^2 and 2.5×2.5 mm^2 is worse. We have also tested the recent Pandora, version v02-04 and found that at higher energies it performed significantly better than the old v00-14 (see later), thanks to the dedicated photon finding algorithm.

The probability to reconstruct correctly the mixed event with 12 and 4 GeV electromagnetic showers is shown in Figure 4.5 as a function of the distance between the clusters. The other energy pairs of 4+4 GeV, 12+4 GeV, 25+4 GeV and 32+25 GeV and other plots with more details can be found in Figures A.29 - A.32 in Appendix A.2.

We see that generally there is a good agreement between TB results and e^+ MC simulation. Pandora and Garlic algorithms work similarly for e^+ and γ showers, so their MC curves are also similar. Pandora and Garlic seeds are found as a position of the maximal energy in the transverse projection of hits from a few first ECAL layers. New Pandora v02-04 has a special peak finding algorithm for the photon identification. This results in significantly better performance. Arbor separates positrons better than photons, because in the former case in the preselection we required one hit in the preshower. It serves as a perfect seed for the Arbor tree.

Some discrepancy between data and MC (especially at higher energies) can be attributed to the e^+ beam contamination by π^+'s which start showering in ECAL. Note also, that there is one problem in translating the TB geometry to ILD, i.e. the significantly displaced showers do not point back to the interaction point IP (coordinate origin in ILD coordinates). In TB geometry the particles are always at normal incidence to ECAL. Therefore, when viewed from IP, the shifted shower looks broader. This sometimes forces Pandora to reconstruct them as neutrons instead of photons, so the overall efficiency slightly degrades. In Garlic we modified the original version 2.10 and removed the requirement that the reconstructed shower should point to IP in version 2.11. Note, that both in Pandora and Garlic this requirement is used in the ILD reconstruction by default, while for some dedicated searches of new particles at ILC (like GMSB [157]) it should also be removed.

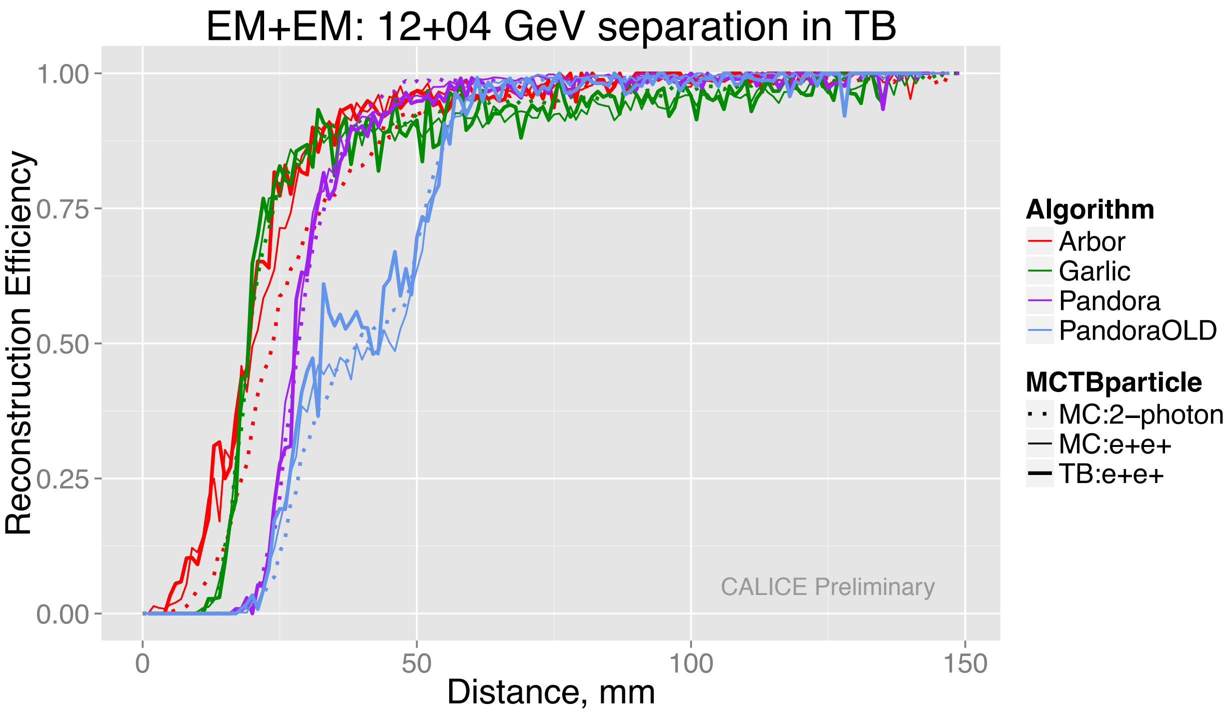


Figure 4.5 – The reconstruction efficiency of 12+4 GeV electromagnetic clusters versus the distance between them (FNAL'11). Other energy pairs may be found in Figures A.29 - A.32 in Appendix A.2.

4.4.2 Photon-photon separation in ILD

The separation of electromagnetic showers is further studied with ILD MC simulation where the ECAL cell sizes have been chosen to be either 5×5 mm^2 or 2.5×2.5 mm^2. The conditions of the successful separation and the energy pairs are the same as for FNAL TB. Reconstruction is performed with Pandora (v00-14 & v02-04), Arbor and Garlic (2.11 and 3.0.3). Results are presented in Figure 4.6 and in Figures A.33 - A.36 in Appendix A.2.

The finer 2.5×2.5 mm^2 granularity should improve the results, but as one can see this is not the case for Pandora and Garlic. Arbor is the only program which results do not degrade with finer granularity.

In Figure 4.7 and similar plots throughout the thesis, we classify the possible reasons of the inefficiency for Monte Carlo photons. Figure 4.7 is for 12+4 GeV photon pair reconstructed by Pandora. Similar plots for Garlic and Arbor may be found in Figures A.37 - A.43 in Appendix A.2. Three rows correspond to various granularities (2.5×2.5, 5×5 mm^2 for ILD and 10×10 mm^2 for TB MC). The black points show the same efficiency curve as in Figures 4.5, 4.6. The colored bands denote the fraction of events where 1, 2, 3, . . . neutral clusters are reconstructed, regardless of their energies and positions. For the correct reconstruction, we require

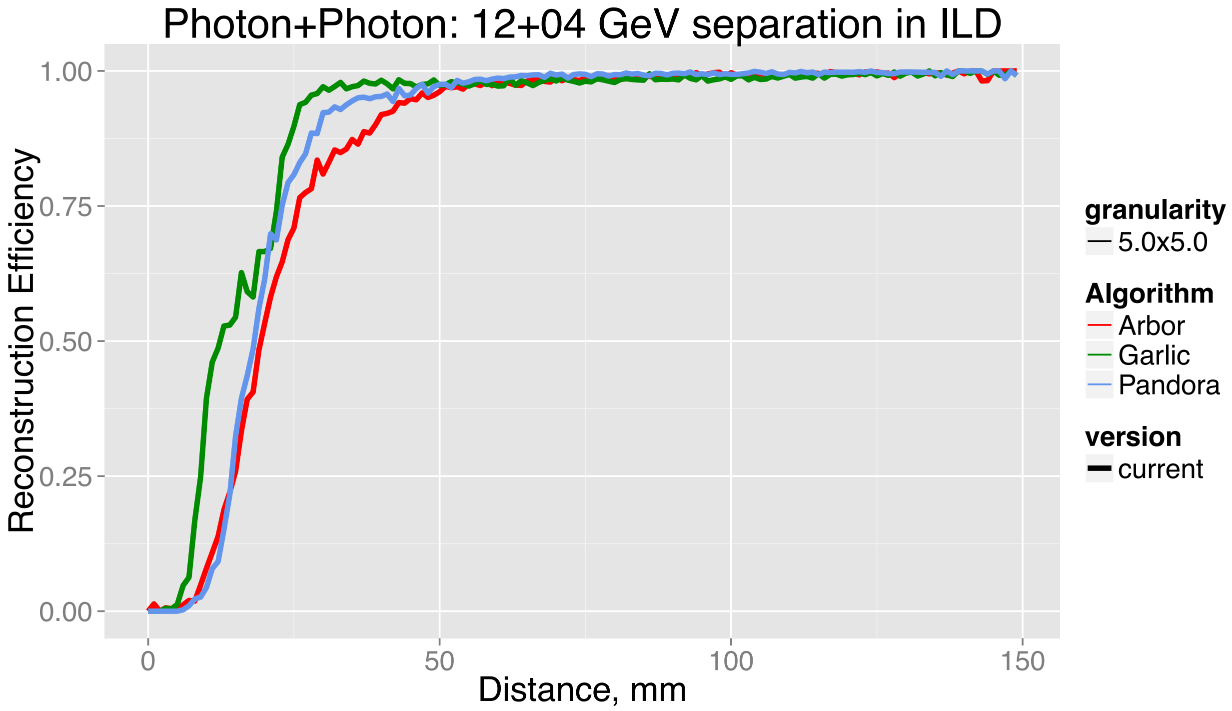


Figure 4.6 – The same as in Figure 4.5 but for ILD with 5×5 mm^2 ECAL pixels. Other energy pairs and the efficiency for 2.5×2.5 mm^2 granularity may be found in Figures A.33 - A.36 in Appendix A.2.

that the latter deviate by not more than ±20% and ±5 mm, respectively, from the corresponding values in the single shower events before mixing. Therefore, the black points are always inside or at the edge of the two clusters (orange) band.

Garlic identifies photons in two steps. In the first step it finds all neutral cluster candidates while in the second it checks their consistency with the photon hypothesis. Therefore, for Garlic we always present two plots (see Figures A.40 - A.43 in Appendix A.2). In the upper (lower) plots the bands are associated to the final number of reconstructed photons (neutral cluster candidates before the photon identification).

4.4.3 Hadronic - electromagnetic shower separation in CERN'07 data sample

Here, we use SiW ECAL together with AHCAL physics prototype. Since the Arbor algorithm is incompatible with coarse, 30×30 mm^2 or more, AHCAL granularity, the analysis is performed only with Garlic (v2.11) and Pandora (v00-14 & v02-04). Note that Garlic reconstructs only electromagnetic showers, so it uses internally only ECAL and tracker collections, but not HCAL. In Pandora the reclustering algorithm is switched on (this is the default). If the energy of the charged particle calculated with the tracker is significantly different from the energy deposited in the calorimeter, the Pandora reclustering algorithm tries to regroup hits in clusters to match the energies.

The criteria for the successful reconstruction of the electromagnetic shower in the mixed event are the same as before. Namely, the energy and X, Z barycenter coordinates of the electromagnetic clusters should agree within ±20% and ±5 mm, respectively, with the energy and the coordinates reconstructed in the single shower event. Here, we require that both Garlic and Pandora reconstruct exactly one photon. For Pandora we additionally require one reconstructed pion.

The efficiency of successful reconstruction as a function of distances between the showers is shown in Figure 4.8 and Figures A.44, A.45 in Appendix A.2. Garlic inefficiency is mainly associated with failures in photon identification. Pandora sometimes combines some pion HCAL hits with ECAL hits of the electromagnetic shower and forms a neutron shower.

4.4.4 Pion-photon separation in ILD

Similarly to photon–photon separation, we perform the ILD simulations with either 5×5 mm^2 or 2.5×2.5 mm^2 ECAL granularity. In addition, we study two HCAL options, AHCAL and SDHCAL. Pandora is used only for AHCAL and Arbor for SDHCAL. Garlic does not use HCAL hits.

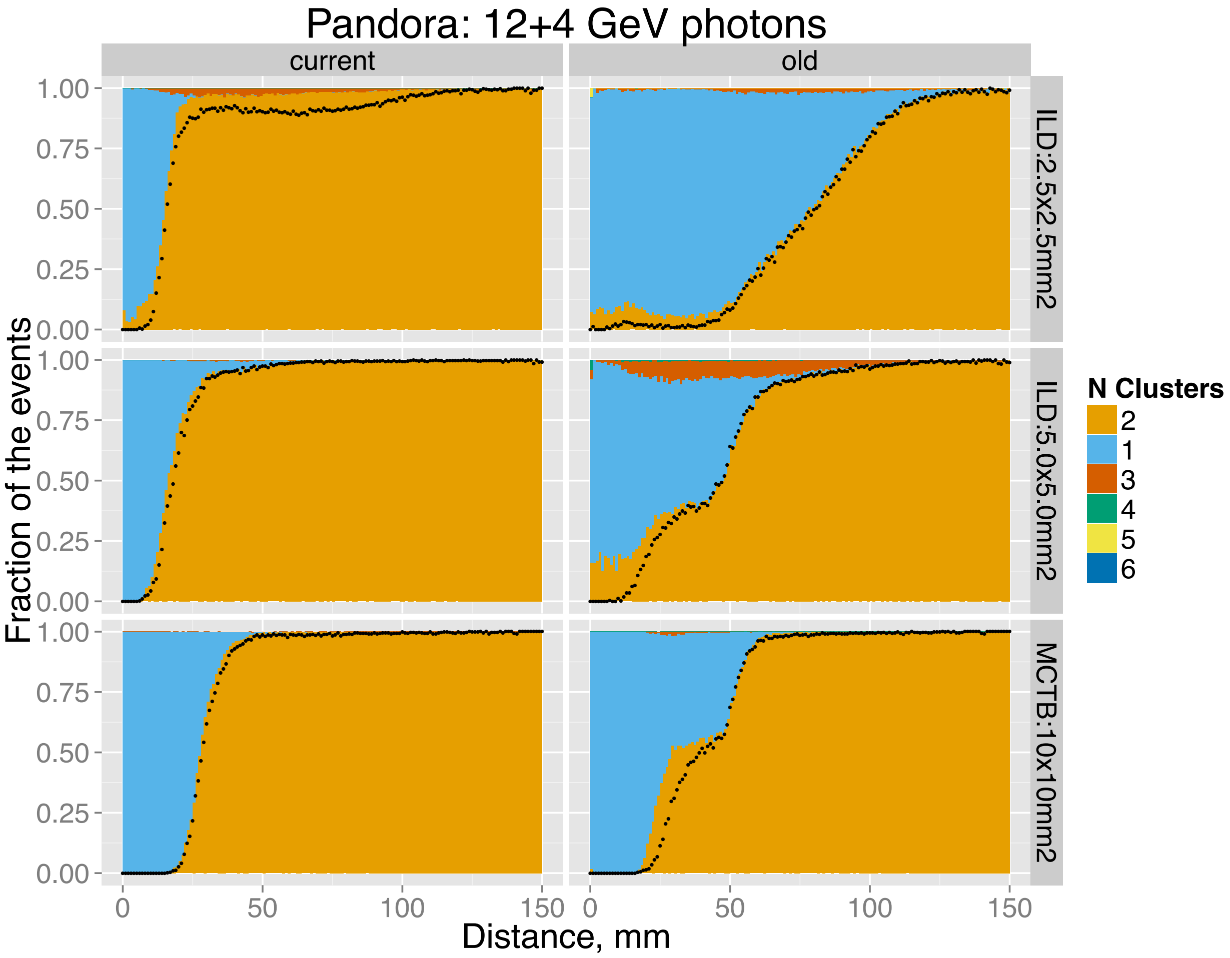


Figure 4.7 – Pandora reconstruction of 12+4 GeV two photon events: the colored bands show the fraction of events where 1, 2, 3, . . . neutral clusters are reconstructed, regardless of their energies and positions. The black points show the efficiency of reconstruction of exactly two clusters with the energies and the positions within ±20% and ±5 mm, respectively, from their values in single shower events, i.e. as in Figures 4.6 and 4.5.

The criteria for the correct reconstruction of the mixed event for Garlic and Pandora are the same as in CERN'07 TB analysis. The Arbor requirements are the same as for Pandora with one exception. Arbor currently does not make a particle identification and does not distinguish between photons and other neutrals. Therefore, instead of one pion and one photon, we require one charged and one neutral cluster reconstructed in the mixed event.

Results are presented in Figure 4.9 and in Figures A.46 - A.49 in Appendix A.2. Similarly to Figure 4.7, the probability to reconstruct the given number of neutral clusters is shown in Figure 4.10 (and Figures A.52 - A.59 in Appendix A.2) in the form of colored bands. At small distances between the clusters Pandora is more efficient than Garlic possibly due to the reclustering algorithm. New version of Pandora demonstrates better performance for 5 × 5 mm^2 ECAL pixels than the old one. Garlic internally ignores ECAL hits close to any extrapolated track from the tracker, this degrades its separation efficiency at small distances.

With 2.5 × 2.5 mm^2 ECAL granularity the performance of Pandora and Garlic degrades. With Arbor we observe another difficulty, it splits the pion shower (mostly in HCAL) and forms nonphysical neutral clusters. According to our definitions, this is counted as inefficiency, in spite of the fact that the neutral cluster energies are typically small. This is demonstrated in Figures A.50, A.51 in Appendix A.2, where the cluster energy is shown versus the radial distance of its energy barycenter from IP. Possibly, this source of inefficiency can be reduced in the future by better tuning Arbor parameters. In Figures 4.9, 4.10 and A.46 - A.49 we also show the Arbor performance when all such neutral clusters in HCAL are simply removed ($\text{Arbor}^{\text{HCAL neutrals}}_{\text{ignored}}$, orange line). In this case, the inefficiencies are associated only with the ECAL and the curve reaches 100% plateau.

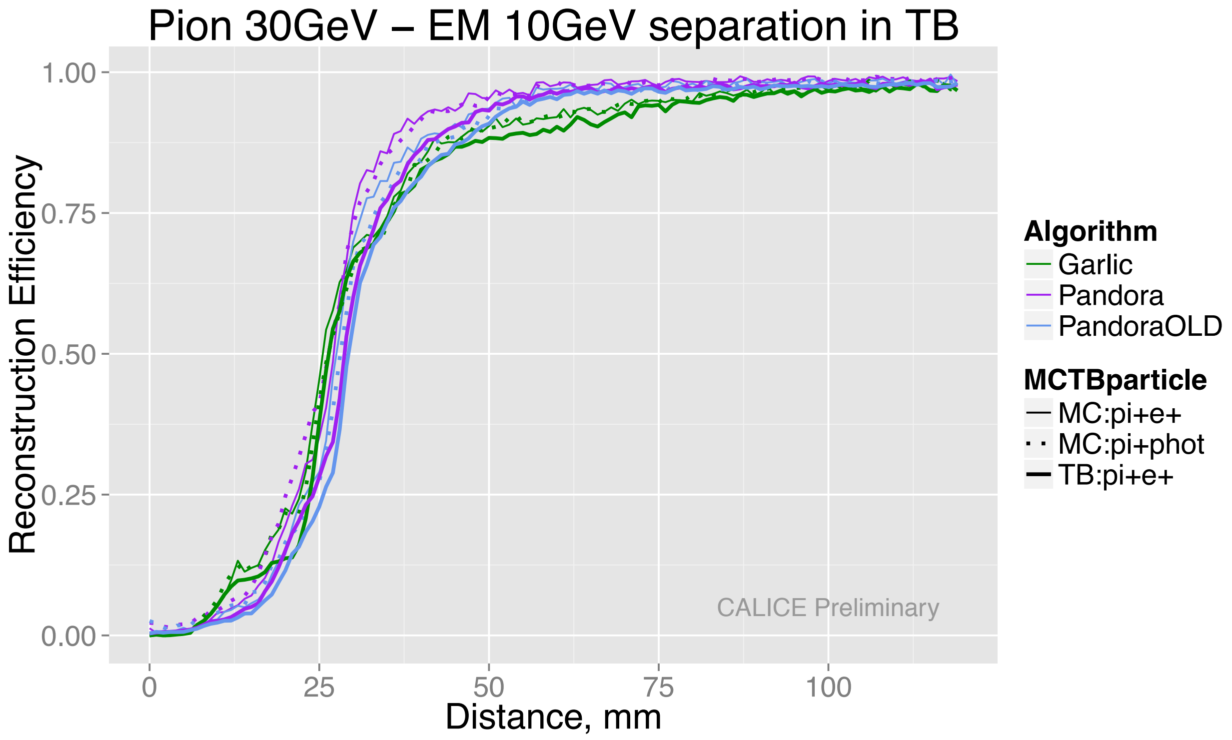


Figure 4.8 – The reconstruction efficiency of 30 GeV pion and 10 GeV electromagnetic shower versus the distance between them (CERN'07). Other energy pairs may be found in Figures A.44, A.45 in Appendix A.2.

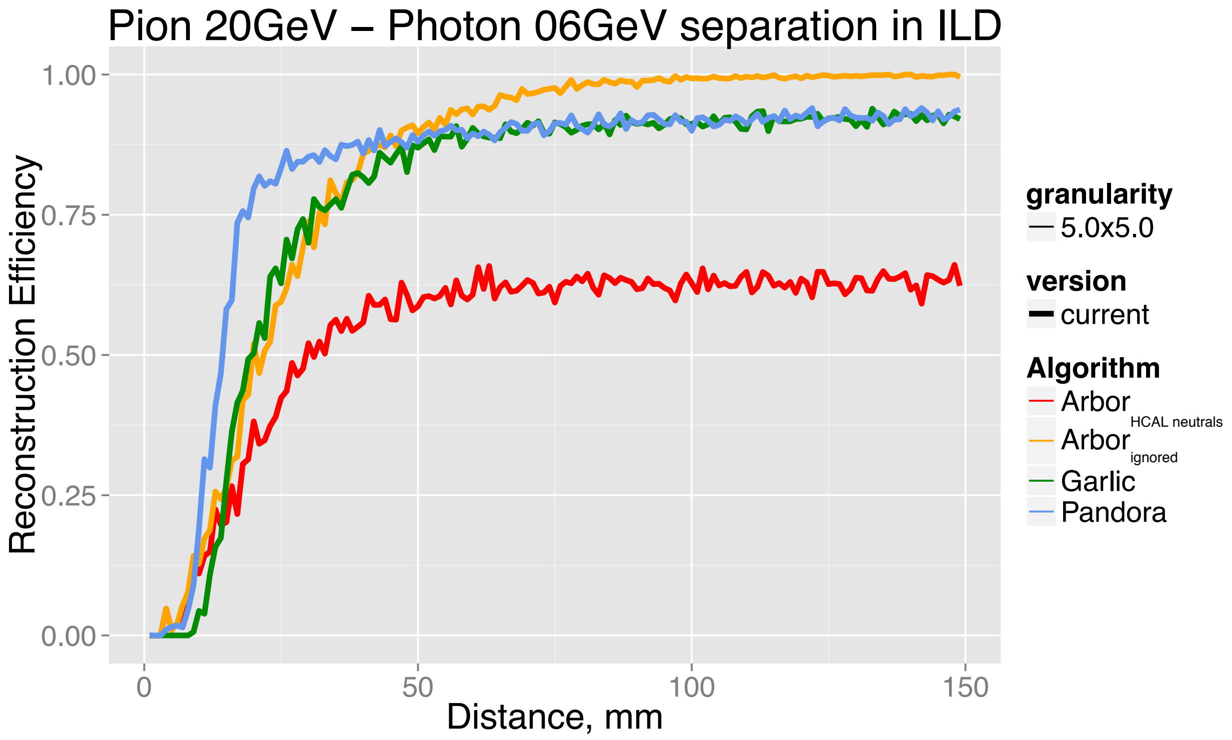


Figure 4.9 – The same as in Figure 4.8 but for the energy pair 20+6 GeV in ILD with 5×5 mm^2 ECAL pixels. Pandora reconstructs data with AHCAL, Arbor – with SDHCAL, and Garlic does not use HCAL information. Other energy pairs and the efficiency for 2.5×2.5 mm^2 granularity may be found in Figures A.46 - A.49 in Appendix A.2.

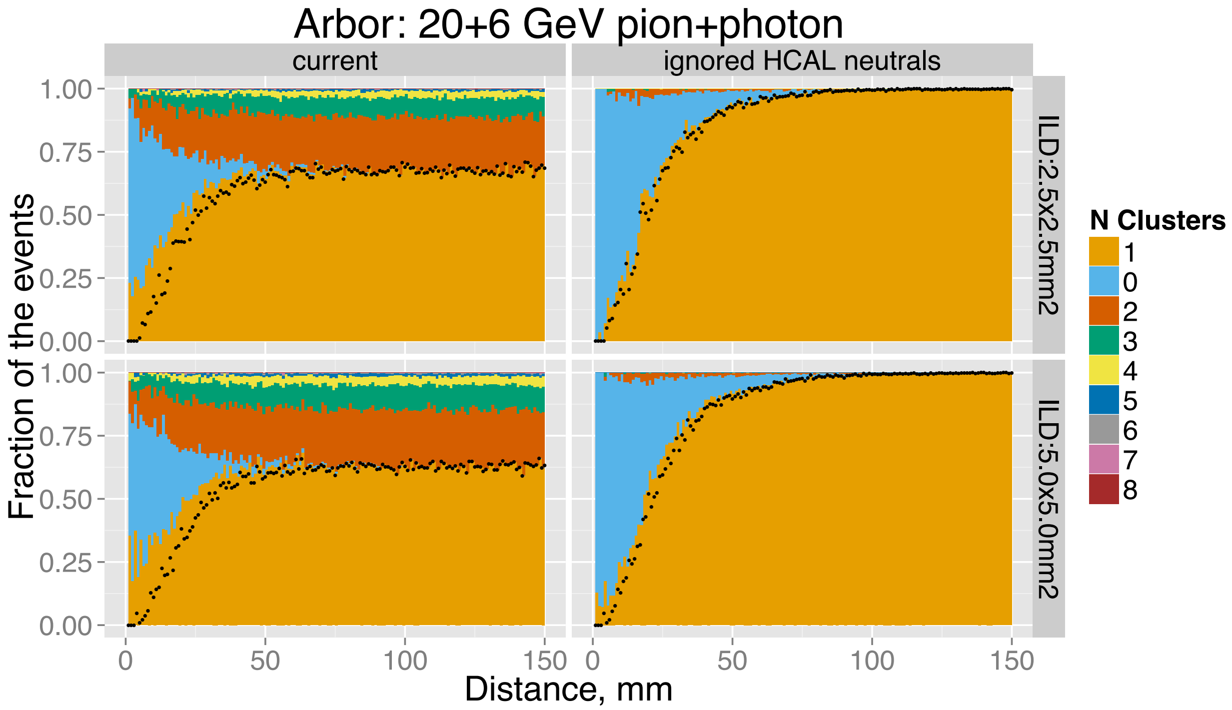


Figure 4.10 – Arbor reconstruction of 20 GeV pion and 6 GeV photon events with SiW ECAL + SDHCAL. The colored bands show the fraction of events with 0, 1, 2, ... reconstructed neutral clusters, regardless of their energies and positions. The black points show the same efficiency as in Figure 4.9. In the right column, all HCAL neutral clusters potentially created by Arbor are ignored. In this case, the inefficient events have either zero or two neutral clusters in ECAL.

4.5 Energy resolution of the simplest pion + photon jet

For the jet energy resolution (JER) it is not really important to reconstruct every particle individually. The goal is to estimate correctly the total energy of the jet. Eg. it does not matter how many photons are reconstructed and how they share the energy. What is important is the total amount of the photonic energy.

Note, however, that the photon resolution is very important eg. for the high energy π^0 reconstruction.

JER is degraded if part of the photon energy is associated to the hadron shower or vice versa. We can consider the pion-photon pair as the simplest jet and determine its energy. Currently only Pandora (for TB and ILD MC with AHCAL) and Arbor (for ILD MC with SDHCAL) can perform the full jet reconstruction. Garlic is intended only for the electromagnetic energy reconstruction in ECAL and, therefore, is not used in this study.

The confusion can be characterized by

$$\Delta E = E^{mixed} - E^{init}_{EM} - E^{init}_{\pi^+} \ , \tag{4.3}$$

Here, E^{init}_{EM} and $E^{init}_{\pi^+}$ are the energies before overlaying while E^{mixed} is the energy reconstructed in the mixed electromagnetic - hadronic event. We do not impose any requirement on the number of reconstructed clusters or their types.

We assume that $E^{init}_{\pi^+}$ is equal to the pion beam energy, as in PFA it should be precisely measured with the tracker. The same is done internally by Pandora and Arbor for the π^+ part of the reconstructed E^{mixed}. In other words, when the pion shower is linked to the ILD track (which we always add to the event), its energy is taken from the precise track energy. In this way, $E^{init}_{\pi^+}$ cancels out in Eq. 4.3 if the pion shower is reconstructed and correctly linked to the track. In this case ΔE in Eq. 4.3 measures only the mismatch between the reconstructed neutral energies. Note also, that in our preselection we always require that before the overlay the events contain only one reconstructed shower. At high separation distances between the showers, the same electromagnetic and hadronic clusters should be reconstructed before and after the overlay. Therefore, ΔE should asymptotically go to zero. The electromagnetic energy fluctuations due to ECAL intrinsic resolution asymptotically cancel out.

The ΔE average and its Root Mean Square (RMS) spread as a function of the distance between the showers are shown in Figure 4.11 for 20+6 GeV pion - photon pair in ILD MC. At low distances the photon hits are merged with the pion shower. The photon energy is lost which results in E_{mixed} underestimation. If large fraction of the photon energy is lost, ΔE becomes close to $-E^{init}_{EM}$ (as for Arbor curves at very low distances).

On the other hand, the photon shower energy variation, $RMS(\Delta E)$, becomes low. Therefore, to interpret the quality of the algorithm, $RMS(\Delta E)$ alone is not sufficient but it should be considered together with ΔE average. Arbor sometimes forms separate neutral cluster from pion hits, as it was mentioned in section 4.4.4. This results in E^{mixed} overestimation at higher distances. When such clusters are ignored ($\text{Arbor}^{\text{HCAL neutrals}}_{\text{ignored}}$, orange curve), E^{mixed} asymptotically approaches zero. Overall, new Pandora has the best performance. Its reclustering algorithm based on π^+ track momentum and shower energy match improves the reconstruction even up to low distances. Other energy pairs, pion-positron combinations with both ILD and TB results may be found in Figures A.60 - A.65 in Appendix A.2.

4.6 Summary of the analysis and conclusions

The overlapping of particle showers in the jets with energies above 100 GeV leads to the confusion and degradation of the jet energy resolution. In this analysis we study this effect by reconstructing two close showers obtained by event mixing. Three PFA reconstruction programs, Pandora, Garlic and Arbor, are studied. For the first two programs we study separately two versions ("old" and "new"). The separation efficiency is obtained for both electromagnetic and electromagnetic-hadronic shower pairs as a function of their distance. The data have been collected during two test beam campaigns in CERN'07 and FNAL'11 with the CALICE physics prototypes of SiW ECAL and AHCAL. To apply PFA programs, the hits from the test beam data have been converted to the central barrel part of the ILD calorimeter. For the electromagnetic showers we use positrons in the data and, additionally, photons in the Monte Carlo simulation. For the hadronic showers we use π^+. The experimental results are in good agreement with the Monte Carlo simulation (QGSP_BERT physics list, Geant4-10-01-patch-03 for CALICE and Geant4-09-05-patch-02 for ILD).

FNAL'11 SiW ECAL data have been used to study the separation of two electromagnetic showers. The energy pairs 4+4, 12+4, 25+4, 32+25 GeV have been analyzed and the results are presented in Figures 4.5, A.29 - A.32. Contrary to Pandora and Garlic which separate $\gamma - \gamma$ and $e^+ - e^+$ without notable differences, Arbor performs on $e^+ - e^+$ significantly better. This is an artifact of our preselection requiring one hit in the first layer which serves as a perfect seed for the Arbor tree. Arbor performance on MC $\gamma - \gamma$ is unbiased.

The π^+ – electromagnetic shower separation is studied with CERN'07 SiW ECAL + AHCAL data using energy pairs 30+10 and 30+25 GeV. Arbor is optimized for a fine HCAL granularity and is not suitable for 30×30 mm^2 or coarser AHCAL pixels. Therefore, we use here only Pandora and Garlic. The results are shown in Figures 4.8, A.44, A.45. Pandora reconstructs both particles and performs better, while Garlic is intended only for the ECAL, so it reconstructs only one photon and can not use AHCAL information.

We also compare the performance achieved with the physical prototypes with the simulation of the full ILD baseline detector in the same central barrel part, both with AHCAL (for Pandora) and semi-digital SDHCAL (for Arbor). As one can see from Figures 4.6, 4.9, A.33 - A.36, and A.46 - A.49, the results are significantly better. Note, that in the baseline ILD the ECAL pixels are 4 times smaller (5×5 mm^2) than in the physical prototype (10×10 mm^2). For both Pandora and Garlic, the "new" versions perform better than the "old" ones. We also simulate the ILD detector with 2.5×2.5 mm^2 ECAL pixels for a comparison. Surprisingly, Garlic and Pandora performance degrades (except for 4+4 GeV $\gamma - \gamma$ case for Pandora). This means that both programs are optimized for 5×5 mm^2. Arbor performance remains approximately the same. In reconstructing the pion, it often finds extra neutral clusters in SDHCAL, though with low energies. Overall, among the three programs, Arbor performance is the least dependent on ECAL granularity.

Finally, we study the jet energy resolution with the simplest artificial "jet" formed by π^+ and an electromagnetic particle. The confusion error associated with the shower overlap, defined by Eq. 4.3, is studied with CERN'07 data and the corresponding simulation (Pandora) and ILD simulation (Pandora and Arbor). The results are shown in Figures 4.11, A.60 - A.65. Again, there is an agreement between the test beam data and the simulation results. In ILD simulation there is no notable difference between 5 mm and 2.5 mm ECAL granularities.

Overall, the particle separation results obtained with CALICE physical prototypes confirm Monte Carlo expectations and support PFA principles.

The Pandora version v02-04 is significantly improved compared to the previous v00-14, especially for the higher energies. The only degradation is observed for low energy electromagnetic pairs 4+4 GeV (standard 5×5 mm^2 ILD ECAL granularity). This can be improved as demonstrated by the other programs. Pandora has the best performance for hadron-electromagnetic showers separation. Here, the reclustering algorithm based on the matching of the pion track momentum and its shower energy helps to separate the particles down to very low distances. The reclustering works best for the higher photon energies. At low energies the pion shower can absorb the photon hits without creating significant energy mismatch.

Garlic demonstrates the best performance in separating two electromagnetic clusters among the three programs. At small distances when clusters overlap significantly, their shapes are distorted (especially at higher energies) and sometimes not reconstructed as electromagnetic. This results in reconstruction of less than two

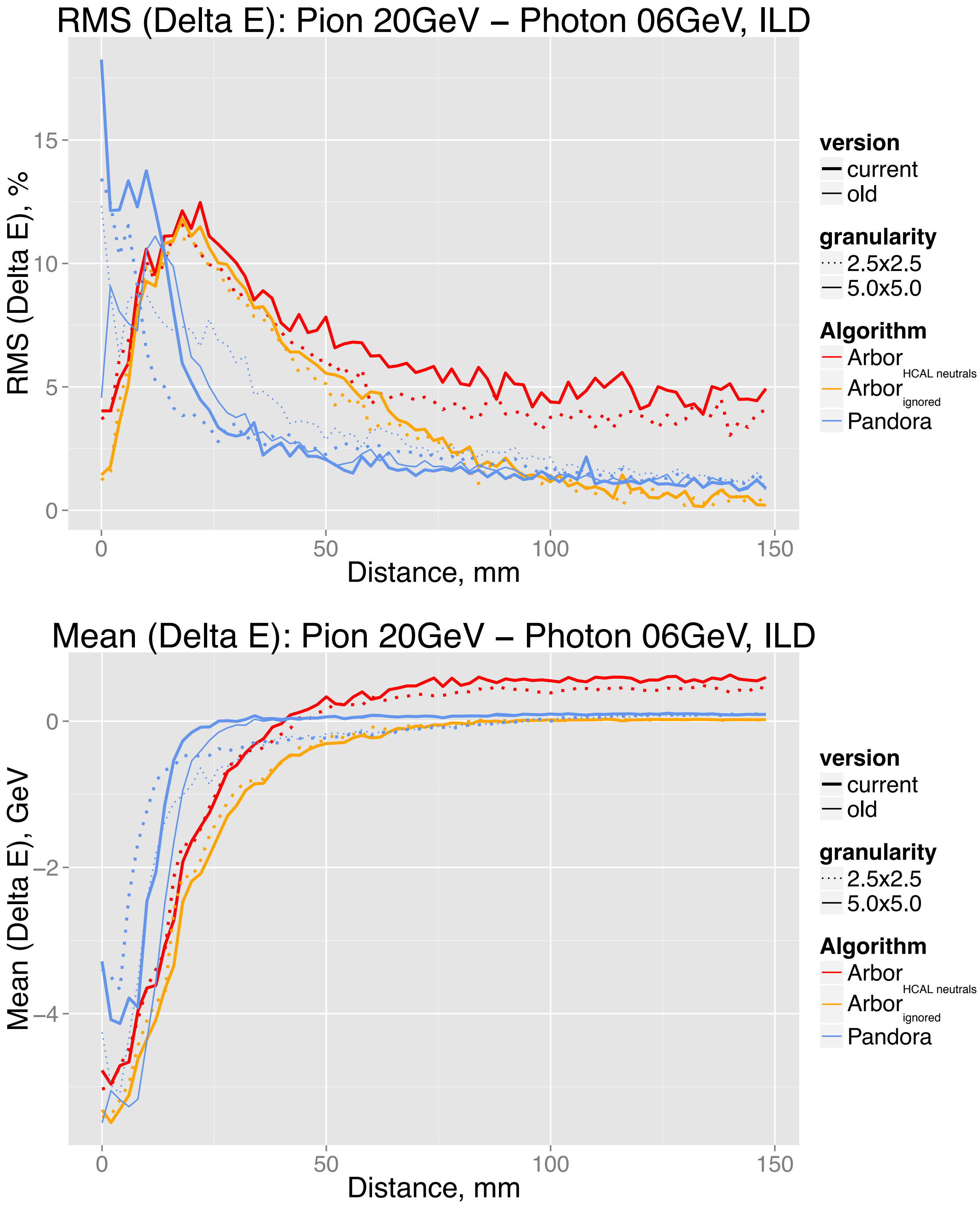


Figure 4.11 – The standard deviation (top) and the mean (bottom) of the confusion ΔE defined in Eq. 4.3 as a function of the distance between 20 GeV pion and 6 GeV photon showers in ILD model with SiW ECAL and AHCAL (SDHCAL) reconstructed by Pandora (Arbor).

photons in the mixed event. Garlic ignores ECAL hits around tracks, this could also prevent the photon shower to be reconstructed. Since it is designed for the electromagnetic reconstruction it does not consider HCAL hits. The new version v3.0.3 has better performance in separating electromagnetic showers, while the old v2.11 better separates hadrons from photons.

Arbor is the only program not specifically tuned for 5×5 mm^2 ECAL pixel size. Its performance for 2.5×2.5 mm^2 granularity is sometimes better than Garlic and Pandora. It tends to oversplit electromagnetic or hadronic showers. It does not have the algorithm similar to Pandora reclustering and therefore its separation efficiency between hadrons and photons drops at low distances.

Chapter 5

SiW ECAL technological prototype

The successful and stable operation of CALICE SiW ECAL physics prototype (2006–2011) has confirmed physics performance of the highly granular calorimetry. Due to not compact design and not embedded very front-end electronics, this technology cannot be scaled to ILD size detector. Starting from 2011, ILD / CALICE SiW ECAL group is working on the technological prototype, which will eventually grow to the ILD detector element and will be suitable for a mass production.

The ILC accelerator collides short bunch trains in two beams every 200 msec. One bunch train lasts only 1 msec and has ~ 0.3 μsec bunch spacing. The “idle” time of 199 msec is needed to perform the radiation dumping and to reduce the beam emittances in the dumping rings. This time structure allows to switch off the front-end electronics in ILC idle time. This mode of operation is called power pulsing. It allows to reduce the power consumption and to simplify the cooling. More specifically, the front-end electronics is switched ON $\sim$1 msec before the bunch collisions for electronics stabilization, stays ON during the collisions and then is switched OFF after complete readout of accumulated hits. With the power pulsing, the passive cooling of the detector layers should be sufficient. The SKIROC2 front-end chip [56] is designed to work in the power pulsing mode and the new SiW ECAL technological prototype is successfully operated in this mode from 2013 on.

The most recent SiW ECAL technological prototypes were tested in beam during three campaigns: in CERN (November 2015 & June 2016) and in DESY (June 2017). I took shifts in all three, participated in the data analysis which is described in the following, and made presentations on behalf of the CALICE collaboration and ILD SiW ECAL group at 38th International Conference on High Energy Physics in Chicago (ICHEP, 2016) and Institute of Electrical and Electronics Engineers Nuclear Science Symposium in Strasbourg (IEEE NSS 2016).

This chapter, first, briefly summarizes the recent achievements in SiW ECAL R&D, namely, the construction of the barrel alveolar mechanical prototype, SKIROC2 ASIC chip and front-end PCB development, advances in readout electronics, tests of the silicon sensors and results of the technological prototype beam tests. Then, it presents the original results obtained during the PhD:

- the study of SKIROC2 chip characteristics (a dynamic range, a spread of the gains and the saturated signals etc.) using the charge injection,

- study of the “square” events,
- preliminary longitudinal profile of 3 GeV e^+ showers,
- checks and debugging of the DAQ data stream integrity,
- preliminary determination of the shower fractal dimension in the transverse plane.

5.1 Very brief overview of recent SiW ECAL prototypes

5.1.1 Mechanical prototype of ILD barrel module with alveoli

In the current ILD design, SiW ECAL will be installed in the modular alveolar carbon reinforced epoxy (CRP) composite structures. They will ensure the mechanical stability, strength, rigidity, and integrity of the ECAL. The carbon fiber structure holds every second tungsten layer. It has empty space between the tungsten layers, so-called alveoli, into which the active detector elements, called slabs, slide in. The slabs have in the middle the other layers of tungsten wrapped in carbon-fiber. Both sides of the slab are equipped with the silicon matrices of PIN diode pixels and the front-end electronics. The silicon sensors are glued pixel-by-pixel to the readout PCBs with a conductive epoxy. PCBs are connected in-line and read out from one slab end.

This design gives a compact ECAL with well defined active silicon and absorber positions with minimal dead zones. In the baseline ILD TDR design the ECAL has 30 layers, the barrel is composed of 8 × 5 modules, each with 5 columns. Each column has 15 alveoli corresponding to 30 layers.

The full size mechanical prototype of the barrel module with 3 columns have already been fabricated during winter 2011/2012. The prototype and its dimensions are shown in Figure 5.1. The mechanical tolerances are within the limits. The slab insertion process has been demonstrated using the tungsten slab prototype without electronics. Additional details on the alveolar mechanical prototype and its tests can be found in [158].

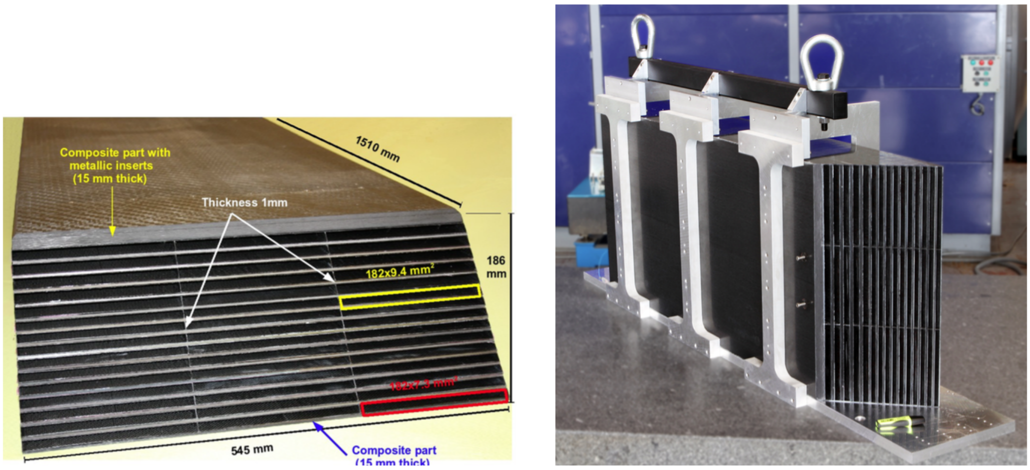


Figure 5.1 – Front view of the alveolar mechanical prototype (left) and its side view in the transportation frame (right).

5.1.2 Technological prototypes with 256 and 1024 channels

The work on SiW ECAL technological prototype has been started in 2011. The current detectors are equipped with one PCB. It is connected to the adapter card providing the interface to the data acquisition system (DAQ) composed of two other cards, Digital InterFace Receiver Board (DIF) and the data concentrator GDCC. The latter is connected to the computer.

The first PCB, called FEV8, carries one silicon sensor with 18×18 pixels of 5 × 5 mm^2 size. It has 256 channels and several of them are connected to multiple (2 or 4) pixels. The silicon pixels are glued by robot to the corresponding PCB pads. The readout is performed by 4 SKIROC2 chips, each with 64 channels. SKIROC2 currently works only in auto-trigger mode and has the ability to keep in analog memory, called SCA, up to 15 events. In every event both signals and pedestals from all 64 channels are stored.

The detectors with one PCB are called short slabs. 12 such slabs with FEV8 have been produced and tested in 2011–2013 [159–161]. 7 of them were operational. Figure 5.2 shows on the left one FEV8 short slab and on the right the setup composed of operational slabs during beam tests at DESY’12.

Starting from 2013 there is an ongoing R&D on the second generation of the technological prototype having FEV10/11 PCBs which can read out 4 sensors each with 256 pixels of 5.5 × 5.5 mm^2 size [162]. There are 1024

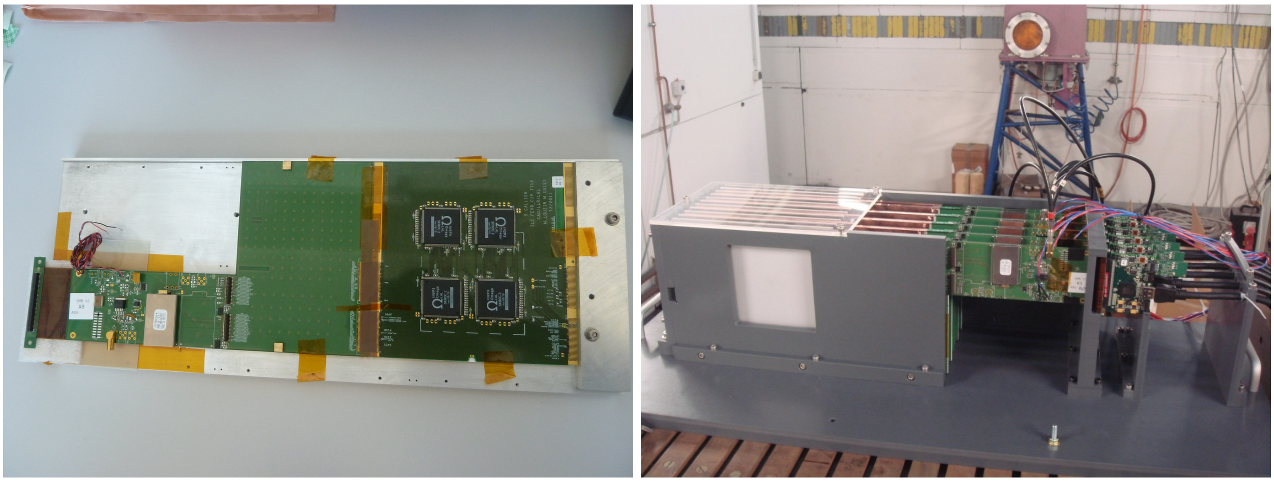


Figure 5.2 – FEV8 SiW ECAL short slab detector tested in 2012 (left) and the test beam setup at DESY (right).

channels and 16 SKIROC2 chips in the PCB in total. FEV11 short slab tests before installation in the beam area at SPS in CERN'15 are shown in Figure 5.3.

Figure 5.3 – Four FEV11 technological prototype slabs during November'15 tests at CERN.

Several tests have been performed with the technological prototype slabs:

- With cosmic muons all channels can be calibrated with 3% accuracy in 24 hours. This has been demonstrated with FEV8 slabs. The duty cycle in ILC is 1%. For cosmic tests it should be increased, eg. up to 85% at spill frequency of 5 Hz (200 msec) to allow 30 msec to read out all accumulated hits. The variation in MIP responses has been measured to be 4% in all channels. It is dominated by the differences between the chips (RMS across chip averages is 2.9%).

- 4 beam tests with FEV8 slabs in 2012–2013 in DESY [159–161] and 3 with FEV11 in 2015–2017 in CERN and DESY. In Nov'15 the first 3 FEV11 slabs were tested at CERN, in Jun'16 10 slabs (but only one operational) together with SDHCAL also at CERN and in Jun'17 7 operational slabs in DESY (see section 5.1.5).

- Direct charge injection in the front-end chips (see section 5.2.1). Injection scans are easy to perform thanks to the dedicated injection lines in the chip and in the boards. Normally, they are switched off to reduce the input noise, but they may be enabled per channel during initial chip configuration.

- Laser light tests [162]. First measurements of the cross-talk due to the electrical coupling of the sensor guard-ring and the peripheral pixels were performed with an infrared laser light ($\lambda = 1056$ nm) injected in silicon. "Square"-events are clearly reproduced. A typical fraction of the induced signal is $\sim 0.4...0.5\%$ times number of outer pixel sides (ie. ×2 for the corner pixel) with a few exceptions.

5.1.3 SKIROC2 front-end chip

The SKIROC2 (Silicon Kalorimeter Integrated Read-Out Chip) Very Front-End ASIC, designed by Omega lab [169], can readout silicon PIN diodes in AMS 0.35 μm SiGe technology [56, 160, 170, 171]. Compact design and 64 readout channels allow using this chip for the detectors with high transverse granularity. According to the current conservative design, the chips are mounted on the board in the BGA packaging. In future, potentially, the naked chips can be integrated on the board (Chip On Board), this will allow producing thinner detector. SKIROC2 ASIC has integrated acquisition chain, amplification, trigger decision, digitization, and readout. Each channel is made of a variable dual gain charge preamplifier (high and low gain) followed by a dual shaper to filter the charge measurement and a trigger chain. The collected charge is stored in a 15-depth buffer, 15 SCA memory slots, and sent to a 12-bit analog-to-digital converter (ADC) to be read out.

SKIROC2 can be power pulsed, it is power consumption is then reduced from 1.5 mW to 27 μW per channel.

The SKIROC2 analog part is designed for the charge depositions in the dynamic range between 0.4 fC and 10 pC. It is expected that one MIP corresponds to 4 fC, so the dynamic range can be expressed as 0.1 MIP to around 2500 MIPs.

The gain of the preamplifier depends on the feedback capacitor C_f as a function $1/C_f$. The slow line for the charge measurements and the fast line for the trigger decision follow the preamplifier. A band gap ensures the stability to the supply voltage and temperature for all the requested references in the analog core. To distinguish the events collected during one spill a slow clock with frequency up to 5 MHz is used (in prototypes we use 2.5 MHz). This produces an internal time stamp called Bunch Crossing Identifier (BCID or BX)

To limit the power consumption and data stream, ILC does not have an external trigger. Therefore, SKIROC2 has the capability of auto-triggering. If at least one channel detects signal above the threshold, the entire chip will be read out. The principal goal is to reach such a level of noises which allows 0.5 MIP threshold.

The fast line of the analog part, the SKIROC2 auto-trigger system, is marked in red in Figure 5.4. Auto-trigger consists of a high gain variable CRRC shaper, which is followed by a low offset discriminator to trigger down to 0.1 MIP. CRRC shaper adjustable minimal integration time is set to 30 nsec. Channel is marked as triggered if the fast integrated signal is above the threshold of the discriminator, then all channels in the chip are read out. The discriminator thresholds are controlled by one common 10 bit Digital-to-Analog Converter (DAC) and 64 DACs with 4 bits individually per channel. The trigger signal is sent to an 8-bit delay cell to provide the hold signal for the slow line. The delay time can be varied between 100 and 300 ns.

The slow line is made of a low gain and a high gain CRRC shapers and is shown as a green part in Figure 5.4. The signal is measured with track and a hold cell. 15 memory slots are used to store the measured signal from two slow shapers. The pedestals in non-triggered channels are readout at the first rising edge of the internal ASIC slow clock after the trigger signal (called OR64, as it is a logical OR between triggers in all channels). The charges stored in the 15-depth Switched Capacitor Arrays (SCA) cells are then converted by a 12-bit Wilkinson ADC, the yellow part in Figure 5.4, and sent to an integrated 4 Kbyte memory.

5.1.4 Silicon sensor studies

“Square” events

At the periphery, the silicon sensor is surrounded by the guard ring (GR). GR smooths the electrical field gradients at the sensor edges, ensures high voltage (HV) stability and low dark currents. Due to the conservative sensor design, physics prototype sensors have 1 mm wide guard ring. This is a dead area which requires corrections in the offline analysis. To equalize the response of the sensor in the center and at the borders, the energy depositions at the sensor periphery should be taken with higher weights to compensate the GR dead area.

There is another drawback associated to GR which was observed in physics prototype data. Due to the capacitive coupling of the silicon sensor GR with the peripheral pixels, a significant energy deposition at the sensor boundary sometimes results in “square” events when many boundary pixels are fired (see Figure 5.5) [111]. In principle, the problem can be solved by connecting GR to a fixed potential somewhere between ground and HV. Because GR is thin, such a connection requires wire bonding, however. This would introduce an extra assembly operation. Wire bonding is not required for anything else in the prototype since the pixels are wide enough to be glued. Therefore, in the current design the wire bonding is avoided and the GR potential is left floating.

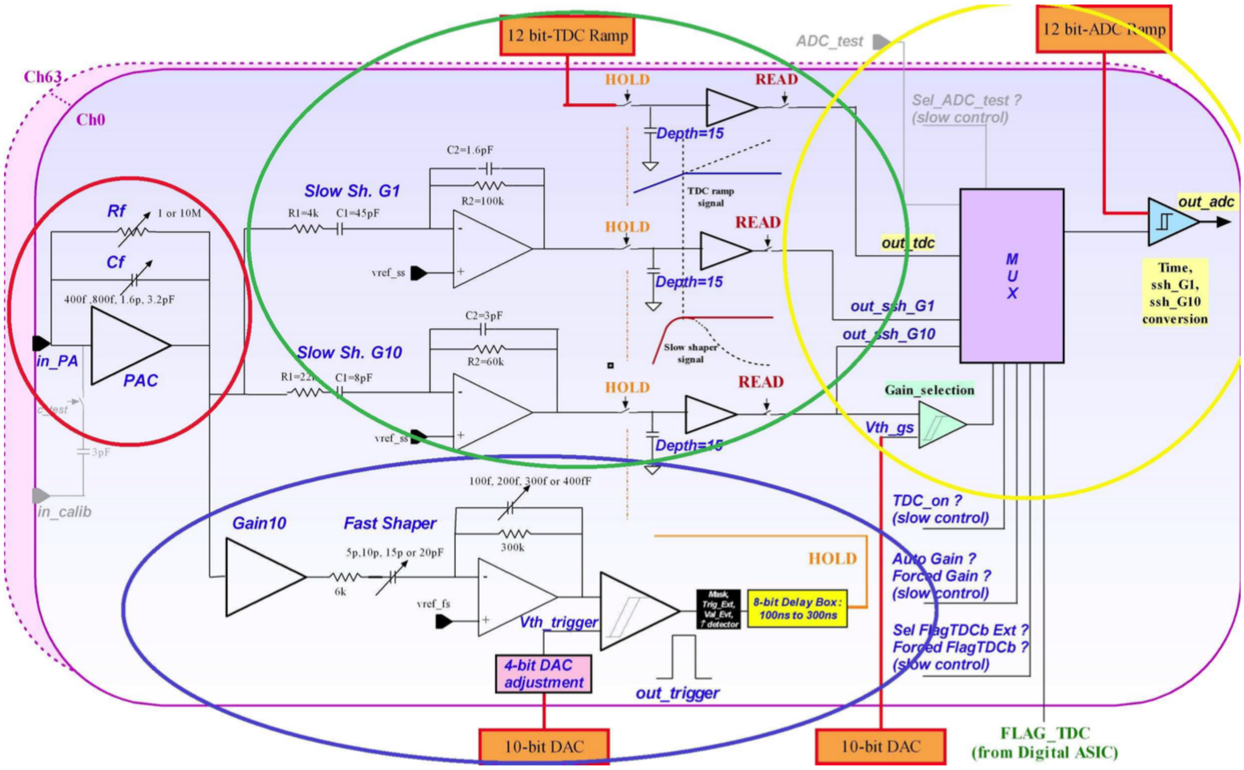


Figure 5.4 – Simplified schema of the SKIROC2 analog part.

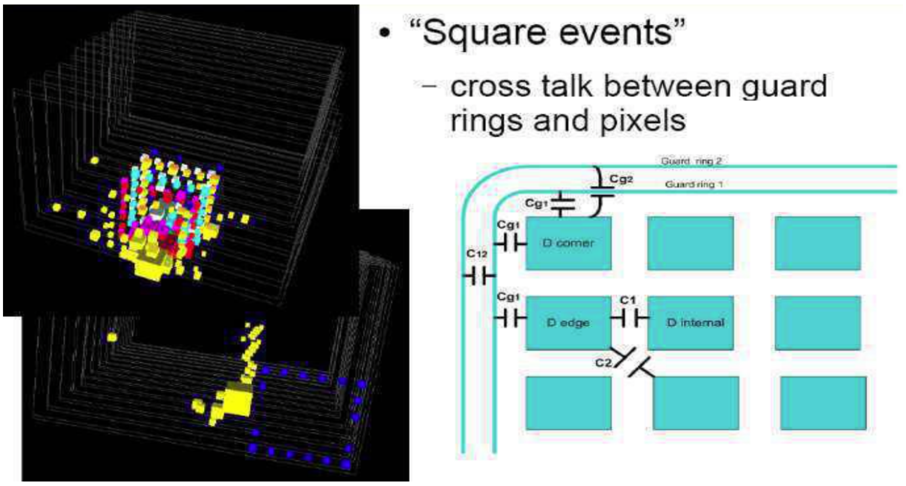

- “Square events”
  - cross talk between guard rings and pixels

Figure 5.5 – "Square"-event examples in SiW ECAL physics prototype (left). "Square"-event effect is caused by large energy deposition in the sensor GR and capacitive coupling between the sensor GR and peripheral cells (right).

The “square” events are suppressed in the technological prototype by optimizing the GR design. In the physics prototype GR has one segment. By splitting it into two or four, the capacitive coupling of GR to the peripheral pixels is reduced. Hamamatsu company also has “know-how” to produce the sensor without GR.

The rate of "square" events with the new sensors have been studied with infrared laser light ($\lambda = 1056 \pm 5$ nm) injected into 9×9 cm^2 silicon sensor of FEV8 technological prototype. The light produces electron-hole pairs in the silicon and can generate very large signals at high rates. In FEV8 detectors the silicon pixels are glued to the PCB with the conductive epoxy, but for this test a new mechanical setup was designed and fabricated in which the sensors can be connected and disconnected. The electrical contacts were made with little springs, see Figure 5.6 [162].

Infrared laser fires in the silicon sensor gap between aluminum contacts. Short <1 nsec pulses generate 700 MIP signals with 200 kHz frequency. Due to the bad grounding and not glued sensors, a high level of noises was observed. Only part of the springs was operational where clear signals were observed. As shown in Figure 5.7, a typical fraction of the induced signal in the peripheral pixels relative to the generated one is ≈ 0.4...0.5% per outer pixel side (eg. ×2 for the corner, ×2 in the case of 4 connected pixels).

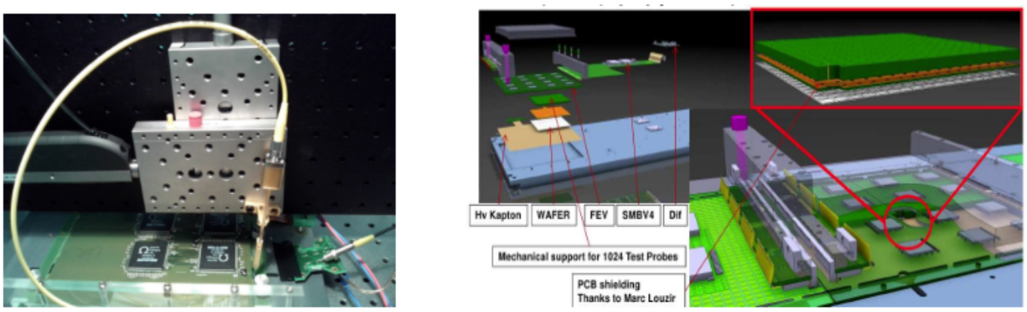


Figure 5.6 – Schematic view of the experimental setup used for laser injection in sensor guard-ring.

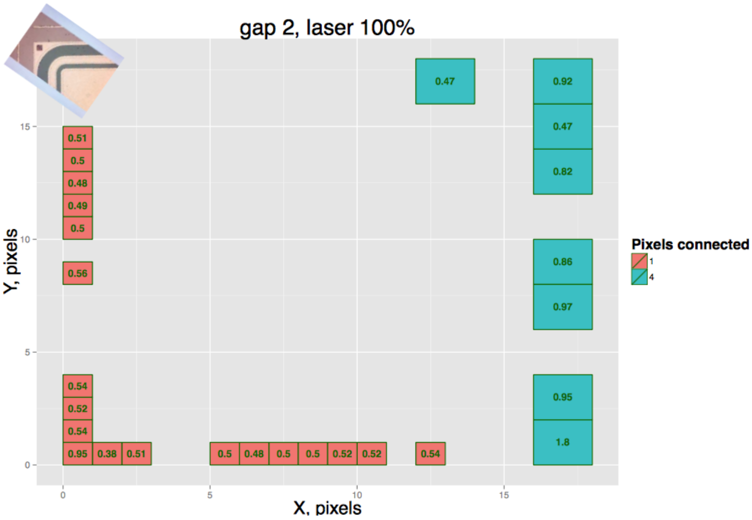
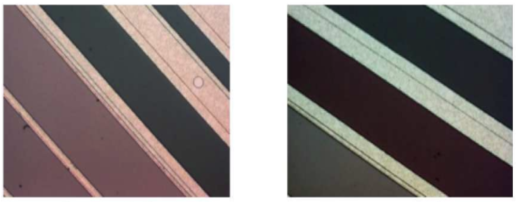
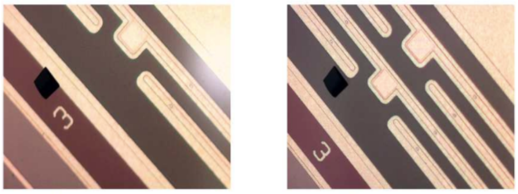


Figure 5.7 – Ring effects observed with the infrared laser in FEV8 sensors. Red squares correspond to the channels connected only to one pixel, blue squares - to the channels connected to four pixels. Ratio of the signal measured in channel to the total laser signal is written inside the squares (in %). From [162].

Later, similar laser injection studies were performed with small 3 × 3 and 4 × 4 sensors having different guard ring designs: 1, 2, 4 segments or "no-GR", see Figure 5.8 [173, 174]. Approximately 10% cross-talk between GR and peripheral cells is measured in the simplest 1-GR sensor. It is significantly suppressed in no-GR desing, as indicated in Figure 5.9, and similarly for 2 and 4 GR segments.



1 guard ring no guard ring



2 guard rings 4 guard rings

Figure 5.8 – Guard ring structures: the guard ring width is 80 μm (1 guard ring), or 20 μm (2, 4 guard rings).

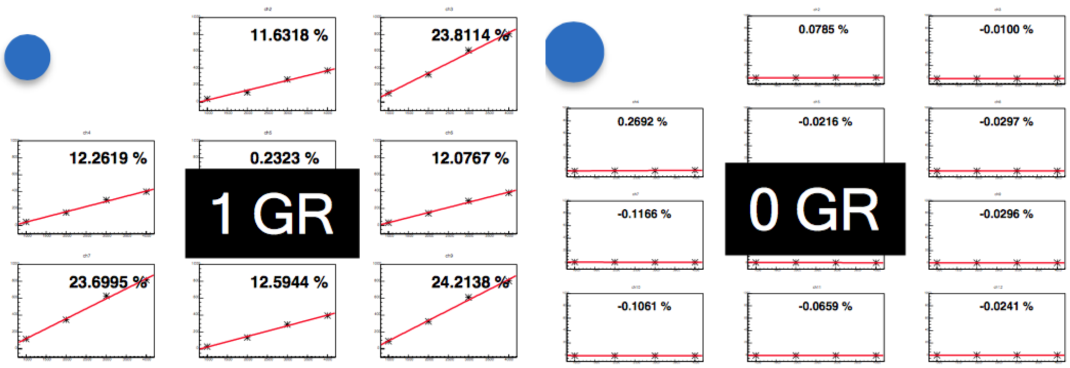


Figure 5.9 – With the laser light injection at the corner of the sensor (blue points) a significant cross-talk is seen only in 1-GR sensor (left). The effect is significantly suppressed for new no-GR design (right) [174]. The numbers indicate the measured fractions of the charge in peripheral pixels relative to the total charge generated by the laser. The points correspond to different laser intensities.

Silicon sensor irradiation studies

Under irradiation the performance of the silicon sensor degrades and its dark current increases. To study the effect, small silicon sensors have been irradiated by neutrons at Kobe tandem accelerator in Japan. The scheme of the experiment is shown in Figure 5.10. The sensors are placed behind the beryllium fixed target, the neutrons ($E_n < 7.8$ MeV) are emitted in the reaction

$$d(3\ MeV) + Be(target) \rightarrow B + n(Q = 4.36\ MeV). \tag{5.1}$$

The measured dark currents, compared with the expectations, are shown in Figure 5.11. This study demonstrates that the dark currents will be acceptable at least during ten years of ILC operated at $\sqrt{s} = 1$ TeV.

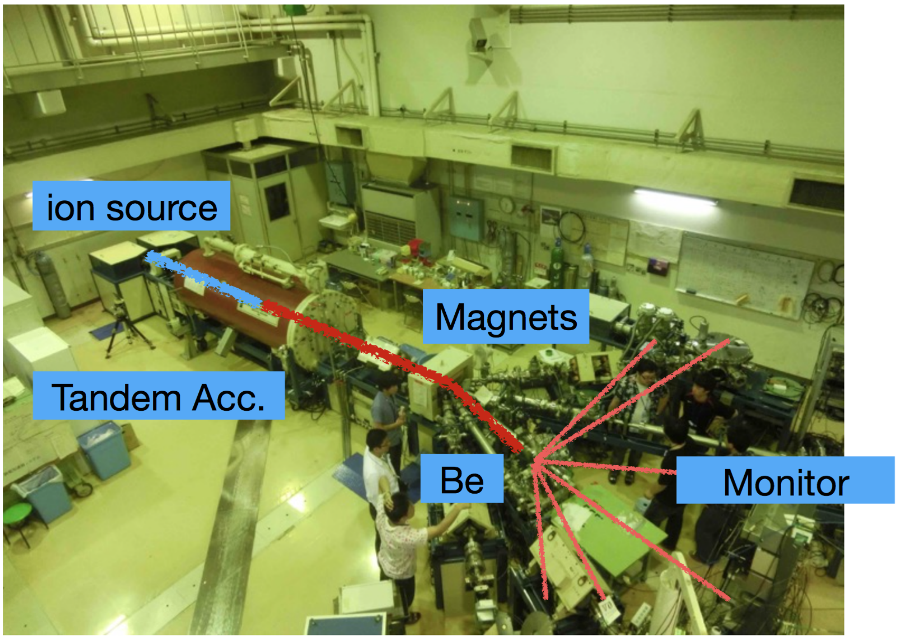


Figure 5.10 – Neutron irradiation tests at Kobe tandem accelerator.

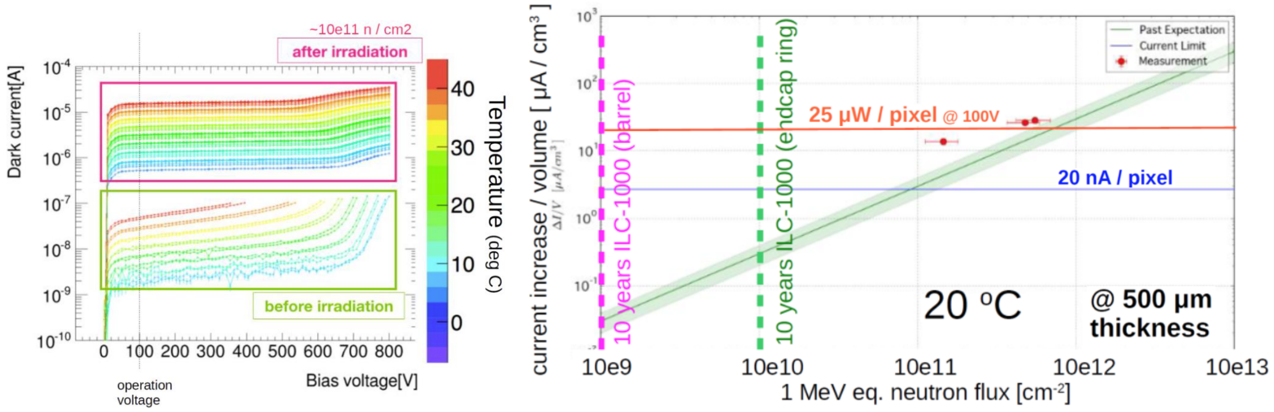


Figure 5.11 – I-V curves of silicon sensor measured at different temperatures before and after irradiation (left) and the extrapolation of the results to the dose acquired during 10 years of ILC operation (right) [174].

5.1.5 Technological prototype beam tests

The most recent FEV10/11-based technological prototypes have been tests in beams 3 times: at CERN SPS in November 2015 and in June 2016 and at DESY in June 2017. 3 layers out of 4 were operational in Nov'15, 1 out of 10 in Jun'16 and 7 out of 10 in Jun'17.

In Nov'15 the SPS beam time was kindly provided by CMS HGCAL. The second beam campaign was allocated by Semi-Digital HCAL (SDHCAL) for the combined tests with SiW ECAL. By that time 10 ECAL layers have been produced, seven were operational, but 6 of them suffered from high noises, only one layer was working as expected. Synchronization and the common DAQ of SiW ECAL and SDHCAL were tested, however. The third test beam campaign with standalone 7 SiW ECAL layers was in June 2017 at DESY. Before the tests, all slabs, interconnections etc. have passed quality tests to ensure good functionality, 3 slabs were rejected due to these checks.

The analysis of Nov'15 tests is the most advanced and will be discussed in the following.

First tests of 3 FEV10/11 slabs in November'15 at CERN SPS

Nov'15 setup with FEV10/11 boards is shown in Figure 5.12. There were 4 layers but due to high noises, one was switched off. The remaining 3 detectors were tested during two weeks with e^+, π^+, μ^+ beams in the broad energy range from 15 to 150 GeV, with different amount of absorber in front of the setup and between the layers and at 0°, $\sim 45°$ and 90° angles with respect to the beam. In total $\sim$ 85 Gb of data were taken. They are publicly available at [177]. Test beam e-log can be found in [178].

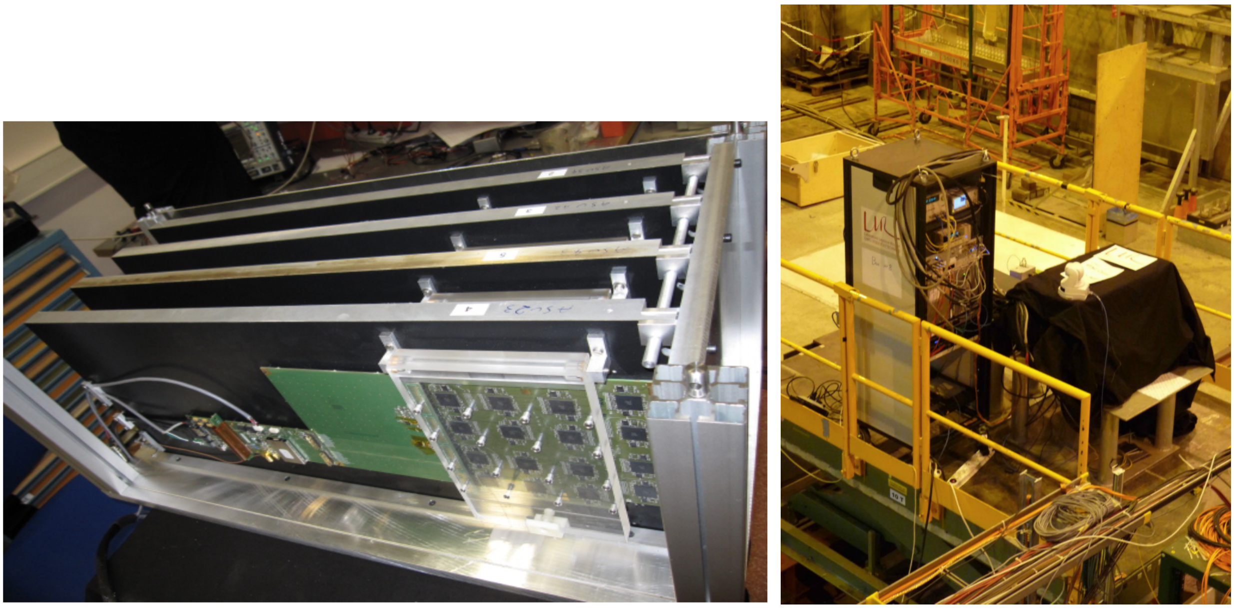


Figure 5.12 – Nov'15 SiW ECAL setup: 4 FEV10/11 boards each with 16 SKIROC2 chips.

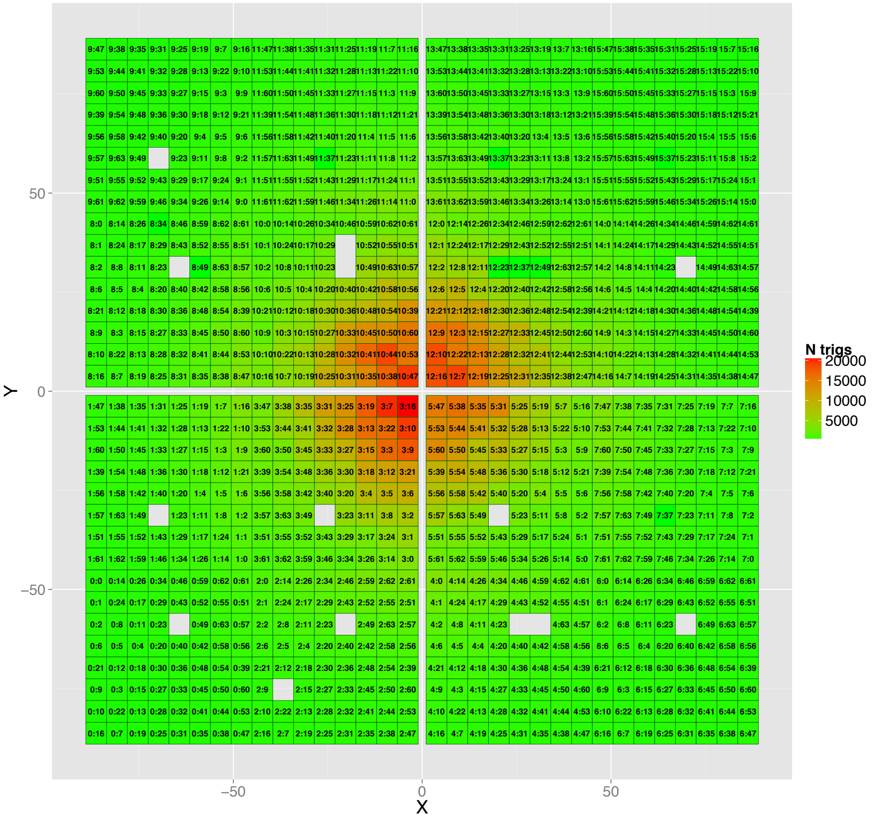


Figure 5.13 – Beam spot in one layer after $8.4X_0$ of W.

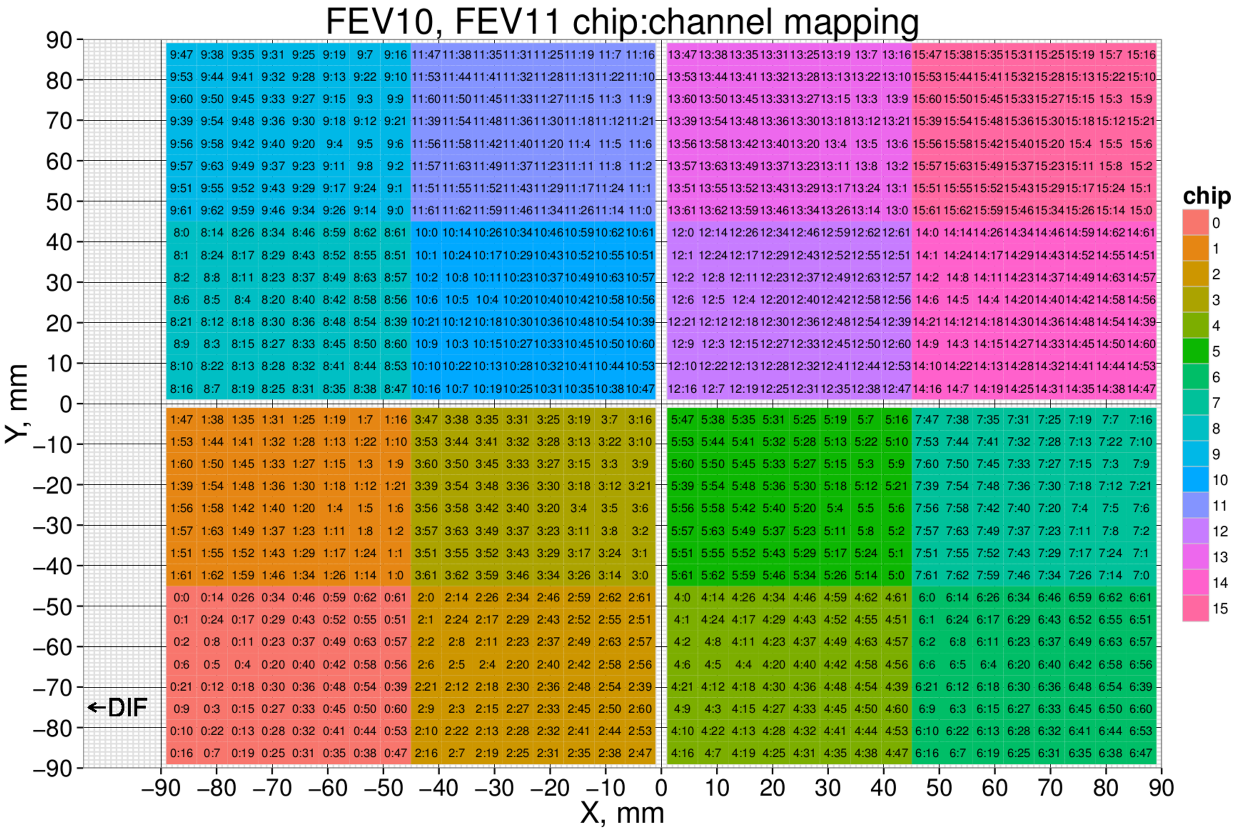


Figure 5.14 – FEV10/11 chip:channel mapping of pixels [180].

Beam spot

Clear beam spot was always seen in every layer, an example is shown in Figure 5.13. The chip:channel mapping of all pixels for the FEV10/11 board is shown in Figure 5.14. A typical transverse view of 150 GeV e^+ shower

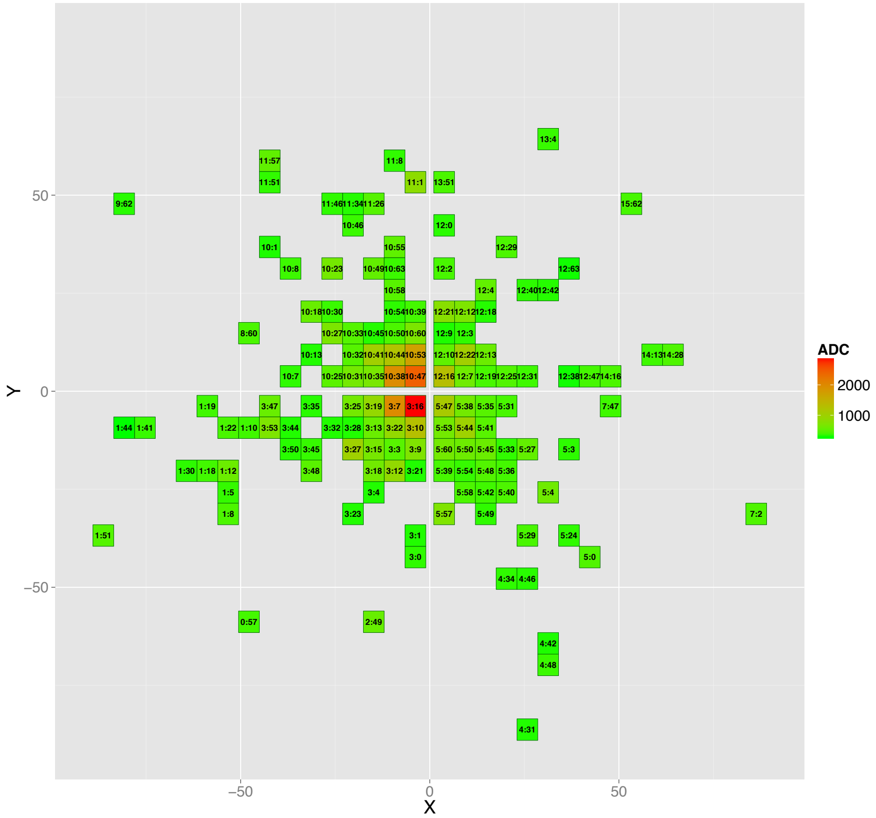


Figure 5.15 – 150 GeV e^+ shower in the layer after $8.4X_0$ of tungsten absorber.

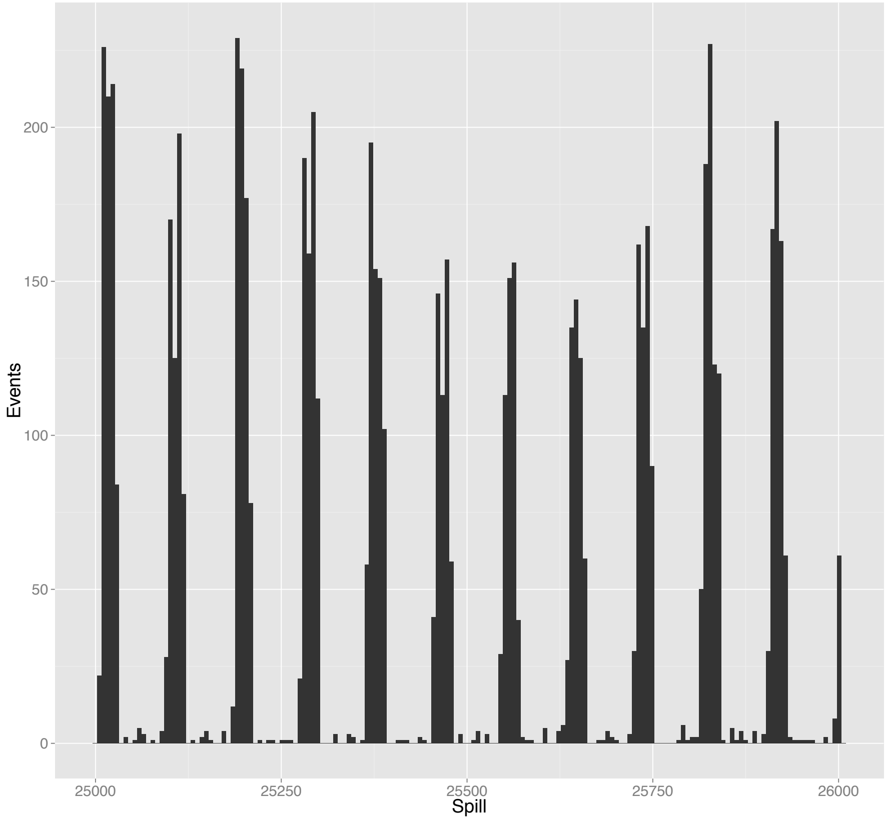


Figure 5.16 – Measured SPS spill structure.

visible after 8.4 X_0 of tungsten is shown in Figure 5.15. SKIROC2 trigger threshold was set by default to 230 for the entire beam period except for some runs in which the trigger efficiency was studied. 2.2% of channels were masked due to noises (in particular, the channel 37 was masked in every chip). The pattern of noisy channels was found using the “FindNoisy” algorithm. Figure 5.16 shows the typical SPS spill structure, more specifically, the full signal collected by one layer versus the acquisition number proportional to time. The low level of backgrounds is visible between the SPS spills. Optimal readout frequency was found to be 4 Hz. At 5 Hz, data for the last 16th chip were corrupted, further details on data debugging are presented in section 5.2.4.

Pedestals

The current version of SKIROC2 chip does not have pedestal suppression and every trigger causes the readout of all channels in the chip. The data from not triggered channels can be used for the pedestal position determination. Next generation of the chip, SKIROC3, will probably have automatic pedestal suppression to reduce the amount of data.

Pedestals should be calculated per channel and per SCA memory slot, $1024 \times 15 = 15360$ pedestals per one layer.

Pedestals can be biased if there are negative signals in the event (probably, due to pick up from the falling edges of digital signals). The fraction of such events is usually negligible, but still, they should be removed in the pedestal calculation.

An additional complication arises because of so-called retriggers. The physical event in one bunch crossing (BX) in some cases is followed by a series of unphysical events, retriggers, in BX+1, BX+2, BX+3,... The sequence of retriggers can be very long and can fill up the full memory of the chip. With low trigger threshold or in noisy environment the retriggers can appear spontaneously, without any visible reason. This happened during Jun’16 tests in 6 slabs and prevented to take data, because the chip memory was exhausted before the first physical event. In November 2015 the level of retriggers was acceptable for a stable detector operation at the nominal 230 ADC trigger threshold.

It was found that the physical events which are followed by retriggers had also biased pedestal values. Only pedestals are affected, the MIP signal position is unbiased. Note, that SKIROC2 pedestals are recorded later than the triggers, namely, at the rising edge of the next bunch clock. This might be related to the fact that the pedestals (MIPs) are biased by (independent of) the following retrigger. For a moment, the explanation of this effect is unknown.

To summarize, for the pedestal calculations only events without any negative signals and not followed by the retriggers are taken. Pedestals are calculated per channels and SCA memory slot. A small fraction of the channels (#45-47 in chips 1 and 9) has double peak pedestals, together with the problematic channel #37 they might need to be excluded from the analysis.

The remaining pedestal spectra are Gaussian, after pedestal subtraction they are shown by the red open histogram in Figure 5.17 for about 1000 channels in one layer. One can also see that the pedestal and MIP signals are well separated, the latter peaks at ~ 65. This is measured with the SKIROC2 gain corresponding to 1.2 pF preamplifier feedback capacitance. This was the default value during Nov’15 tests. Note, in ILD one should use 6 pF, ie. 5 times less gain to increase the SKIROC2 dynamic range. 1.2 pF was chosen to improve SKIROC2 signal-over-noise ratio, with 6 pF it is about twice worse.

It was checked that during the tests the pedestals were stable within $\pm 0.4\%$ of the MIP signal with one exception. All three slabs were power pulsed, but with the timing optimized in order to collect more data. Therefore, the acquisition time usually was much longer than in ILD, where it should be 1 msec only. To test ILD power pulsing, several runs have been accumulated with the ILD settings. It was observed that in this data in one out of three slabs the pedestals deviate by about 3.8% of the MIP from the values measured with the default settings. In the two other slabs the deviation was almost zero.

Currently, the electronics is switched on 1.4 msec before the moment when the first data can be taken. This delay is introduced to stabilize the electronics. The pedestal deviation in one of the slabs can indicate that the pedestal drift might not finish in 1.4 msec.

Muon detection efficiency and the gain uniformity across channels

Three ECAL layers are sufficient to determine the MIP efficiency: to study one layer one can require coincidence between the other two. 150 GeV muons, selected for this analysis, traverse ECAL almost perpendicularly. The analysis takes into account not perfect synchronization between the layers (±1 BX) and excludes the inefficiencies arising due to the limited SKIROC2 memory. To exclude muon showers, it is required that every trigger layer has only one hit.

The results for the first layer are shown in Figure 5.18. Every channel in 16 chips is represented by one point, a few outliers are shown in red. Similar plots for other two layers can be found in Appendix B.2, Figures B.13.

The efficiency averaged per chip with outliers excluded is shown in Figure 5.19 for all three layers. The inefficiency is mainly due to the trigger threshold (230 in SKIROC2 units). 2.9% of outliers are dominated by one chip (#14 in layer 3, contributes 2.1%). The average chip efficiency without outliers is 98—99%.

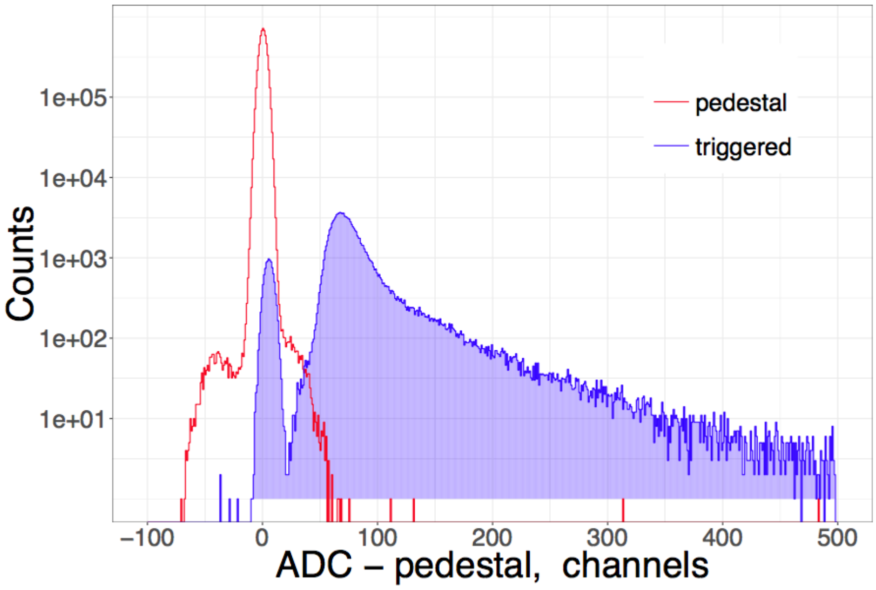


Figure 5.17 – The combined pedestal subtracted ADC spectra for all channels of one layer measured during two hours muon run [27]. Triggered (pedestal) data are shown in blue (read). A peak at zero for the triggers is caused by spontaneous retriggers.

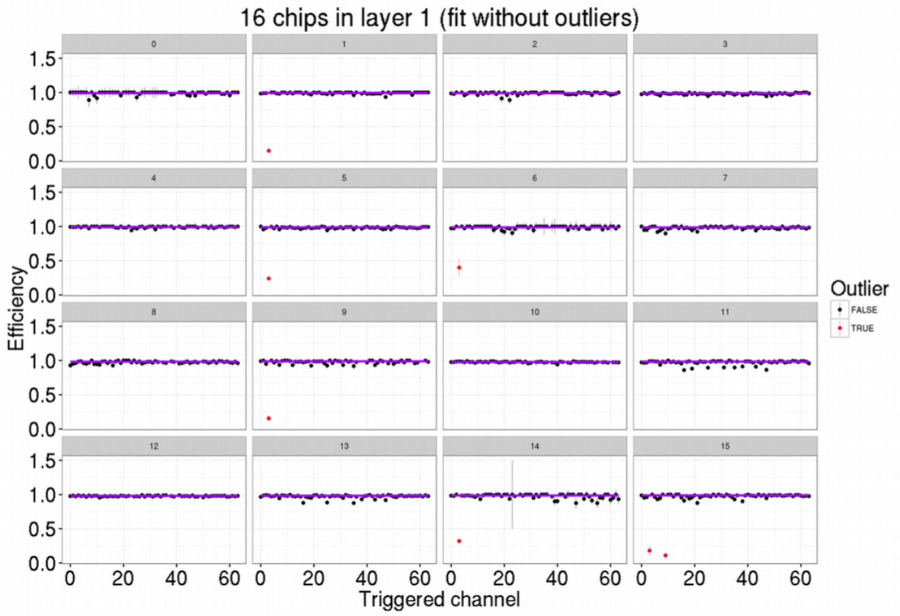


Figure 5.18 – Muon efficiency per chip and per channel for the first layer [181].

Figure 5.20 shows the dE/dx distributions of triggered signals recorded by 16 chips in the first layer.

Excellent gain uniformity across the pixels is one of the advantages of the silicon detector technology. It makes the calibration process easy and reduces the systematics. The measured spread of dE/dx most probable values for 150 GeV muon, is determined to be 6.4% before any calibration. The most probable value is calculated using the truncated mean method for the channels with sufficiently high statistics (83% of all channels).

5.2 Technological prototype studies performed during PhD

5.2.1 Charge injection study with FEV8 boards

Description of the setup

Three FEV8 SiW ECAL technological prototype slabs have been tested with direct charge injection in the front-end chips. Injection scans are made with FEV8 slabs 5, 8 and 10. Charge injection is done in each

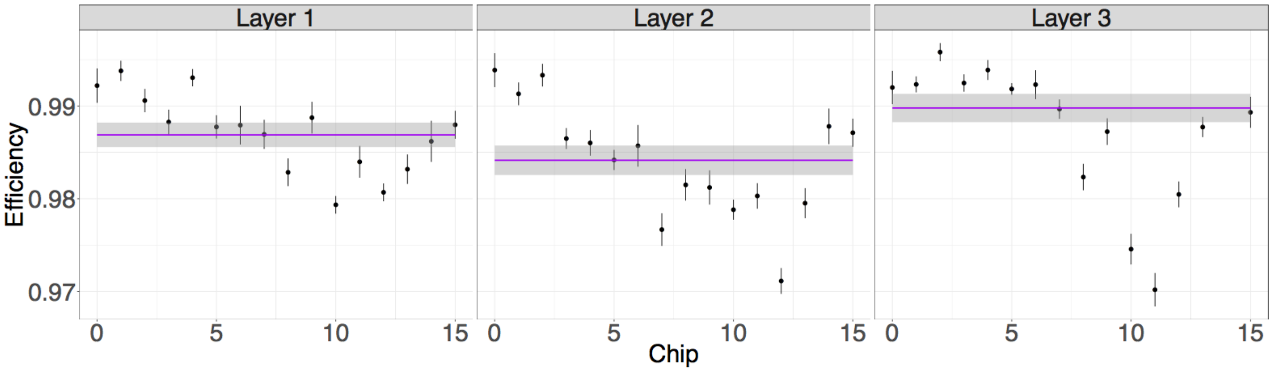


Figure 5.19 – Averaged muon efficiency per chip excluding 2.9% of outliers [27].

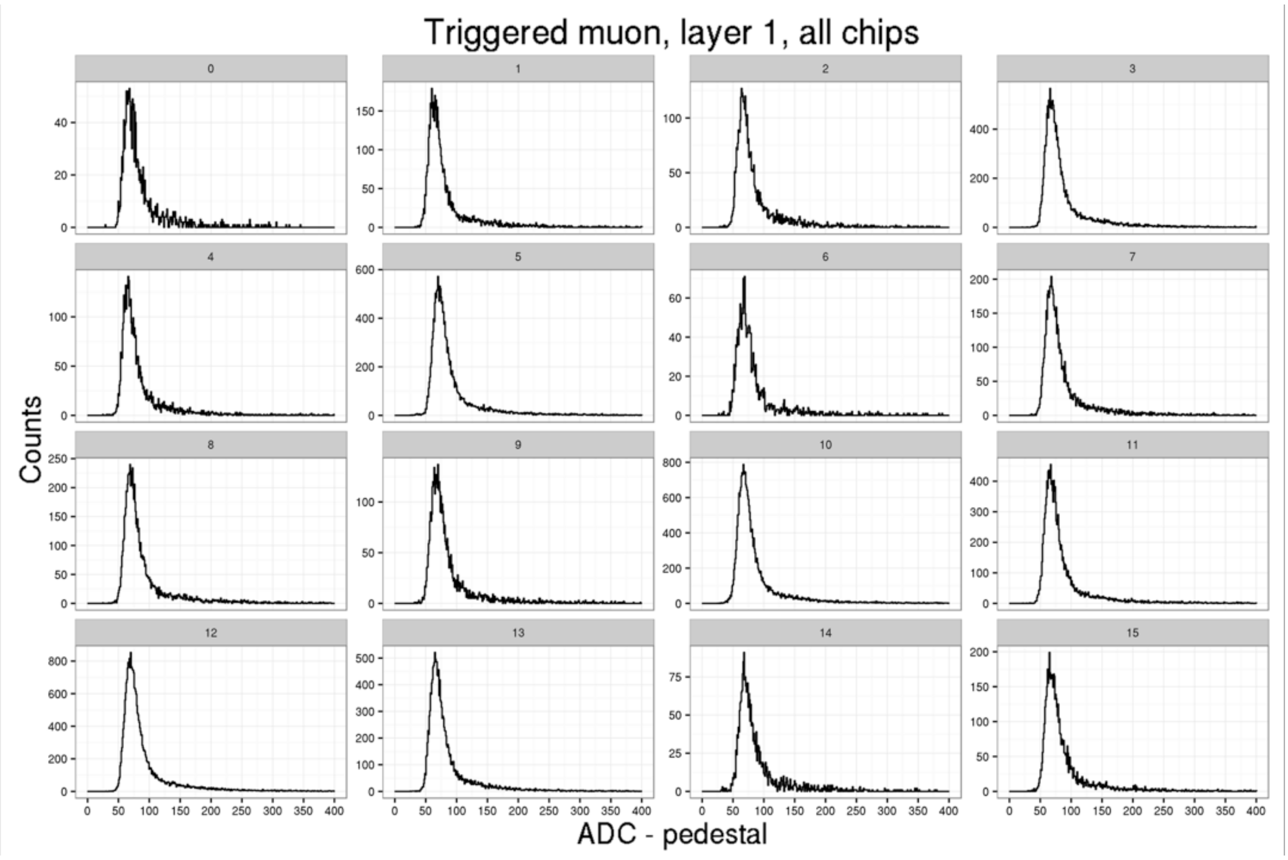


Figure 5.20 – dE/dx triggered signals from 150 GeV muon accumulated per chip in the first layer [27].

channel individually (of each chip and each slab). In all other channels preamplifier, injection and triggering were switched off. It is performed in power pulsing (PP) or continuous current (CC) modes. In slab 10 both modes (PP and CC) are used, in slab 8 only CC mode and in slab 5 only PP mode. Injections are made for 11 voltages of 0.01, 0.02, 0.05, 0.1, 0.15, 0.2, 0.3, 0.4, 0.6, 0.8, 1.0 Volts. It has been shown that some differences between CC and PP modes exist. The raw data is first converted to ASCII human readable format using C++. Then the data is analyzed with R programming language.

Pedestals

It is known [172] that pedestals should be measured per channel and per SCA memory. This has been correctly taken into account in all analyses. As an example, the measured pedestal ADC values are shown in Figure 5.21 for 64 channels of one chip. In total, there are 64 × 15 pedestals per chip.

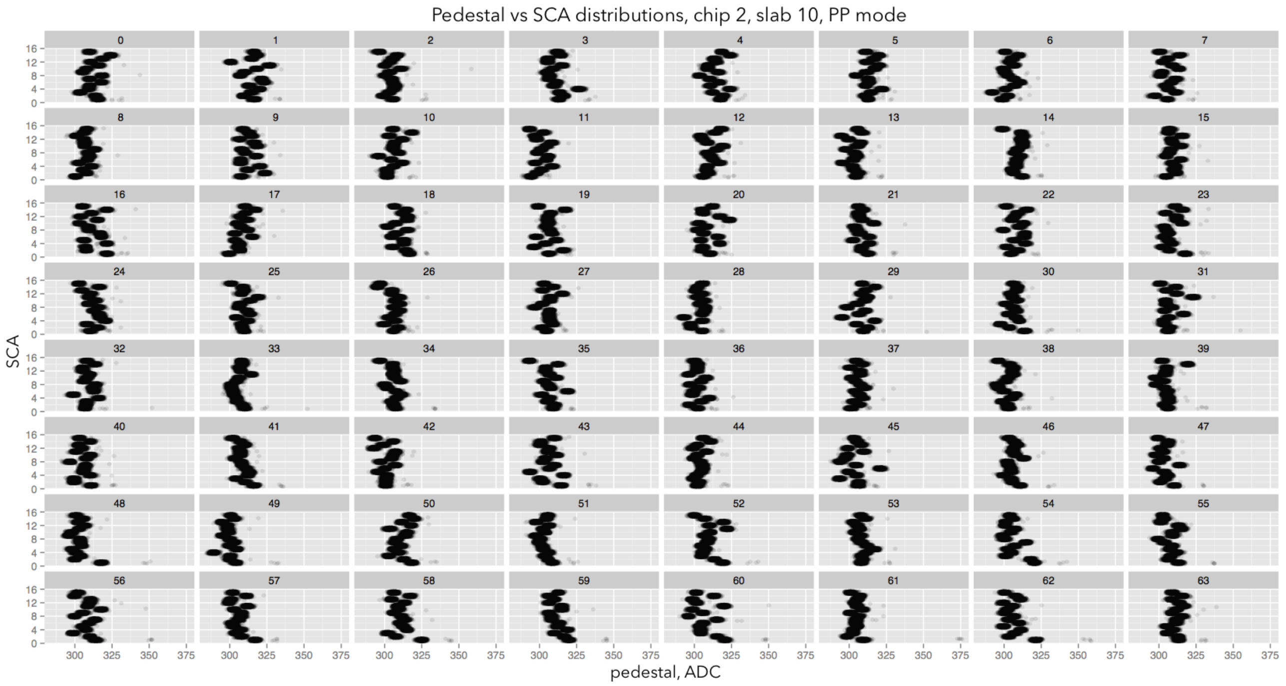


Figure 5.21 – Two-dimensional ADC-SCA distributions for pedestal events in 64 channels of chip 2, slab 10. SCA (Switched Capacitors Arrays) denotes one of 15 SKIROC2 memory slots.

Measured responses to injected charge

The channel response to the injected charge, or equivalently, pulse generator voltage is measured in analog-to-digital converter units, or ADC counts. Mean value of ADC is calculated for each channel and each voltage (pedestal is subtracted taking into account its SCA dependence). To reduce noise, only signals with correct BX (i.e. synchronized with the pulse generator) are taken into account.

An example of the measured response for one channel is shown in Figure 5.22. The curves for all channels are given in Figures 5.23, 5.24 for PP mode in slab 10. Similar plots for other slabs can be found in Appendix B (Figures B.1-B.8). Clearly, a few channels are not operational.

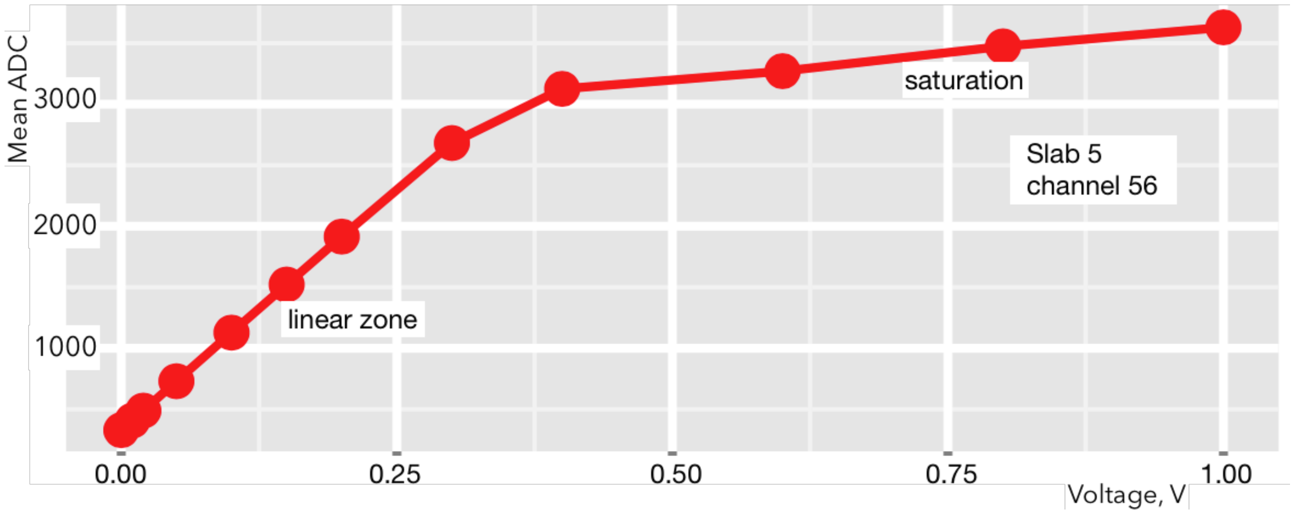


Figure 5.22 – An example of the average channel response as a function of injection voltage for channel 56, chip 0 of slab 5.

The curves have several parts as depicted in Figure 5.22:

- there is a small non linearity at the level of a few percent for low signals (of the order of one MIP),
- a linear zone for voltages between 0.05 and 0.3 V,
- the transition region from the linear zone to a saturation,

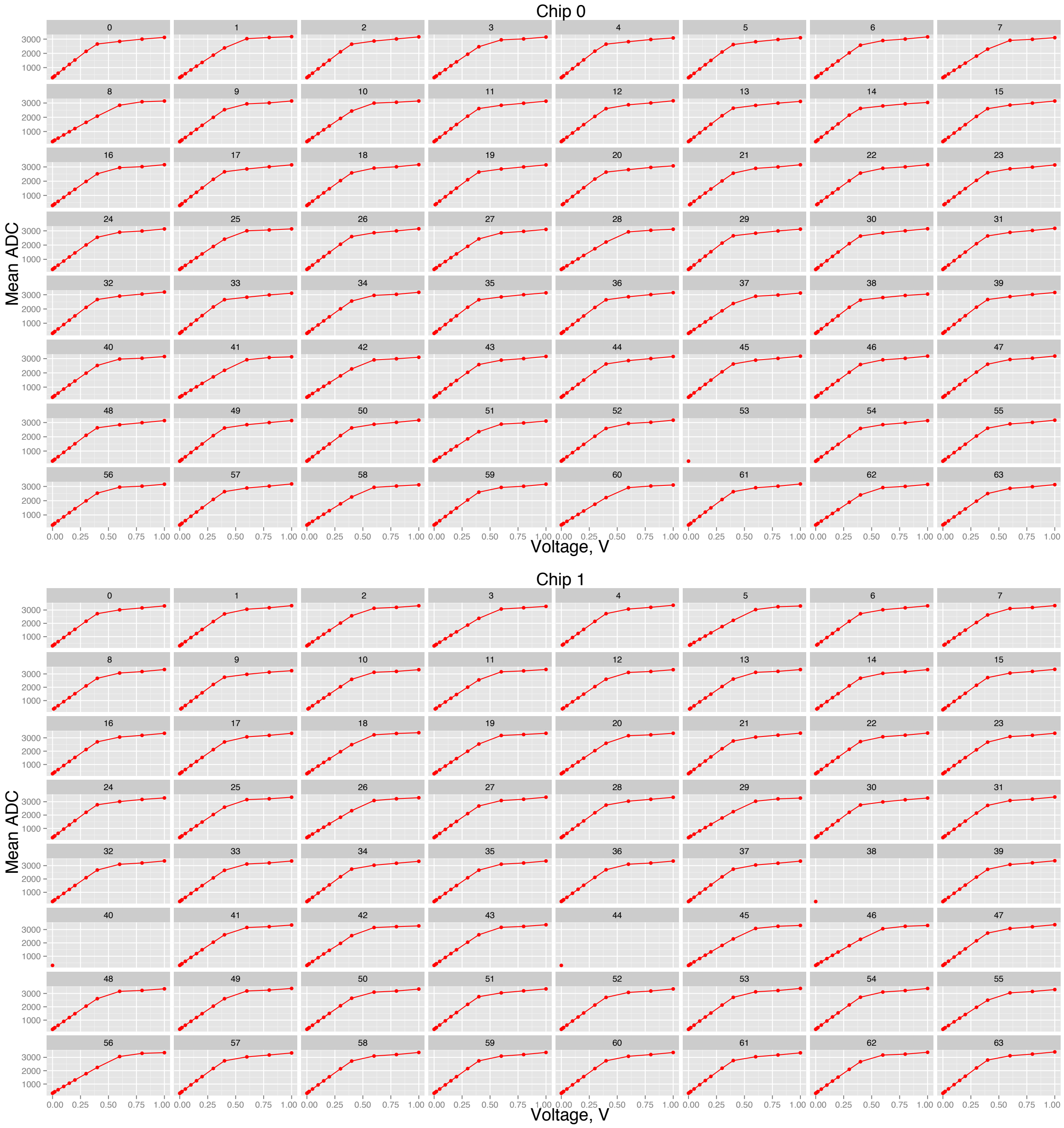


Figure 5.23 – Average channel response as a function of injection voltage for 64 channels in chips 0, 1 in slab 10, PP mode.

- the saturation for voltages >0.5 V.

Table 5.1 summarizes the statistics of linear slopes for voltages between 0.05 and 0.3 V. First four lines present information on chip average slopes. Next four lines - on RMS of slopes within each chip. Next two - RMS across all slab channels and RMS across chip averages. We see that difference between PP and CC modes in slab 10 is small. For all slabs, RMS for chips 2 and 3 is less than that of chips 0 and 1. Distributions of slopes for all channels for each slab are shown in Figures 5.25, 5.26.

Table 5.2 summarizes information about saturation at 1 V injection. First five lines give information about average ADC per chip and average ADC between chips. Next five lines present RMS of ADC values per chip and RMS between average chip ADC values. RMS of ADC values within chips are small, from 0.71% to 1.74%. RMS across chip averages is higher (from 3.07% to 3.68%), meaning that ADC values are grouped for each chip. The distributions of channel responses to 1 V injections in different slabs are shown in Figures 5.27, 5.28. Four peaks correspond to different chips.

This study demonstrates that gains and non-linearity in the real experiment can be calibrated per channel

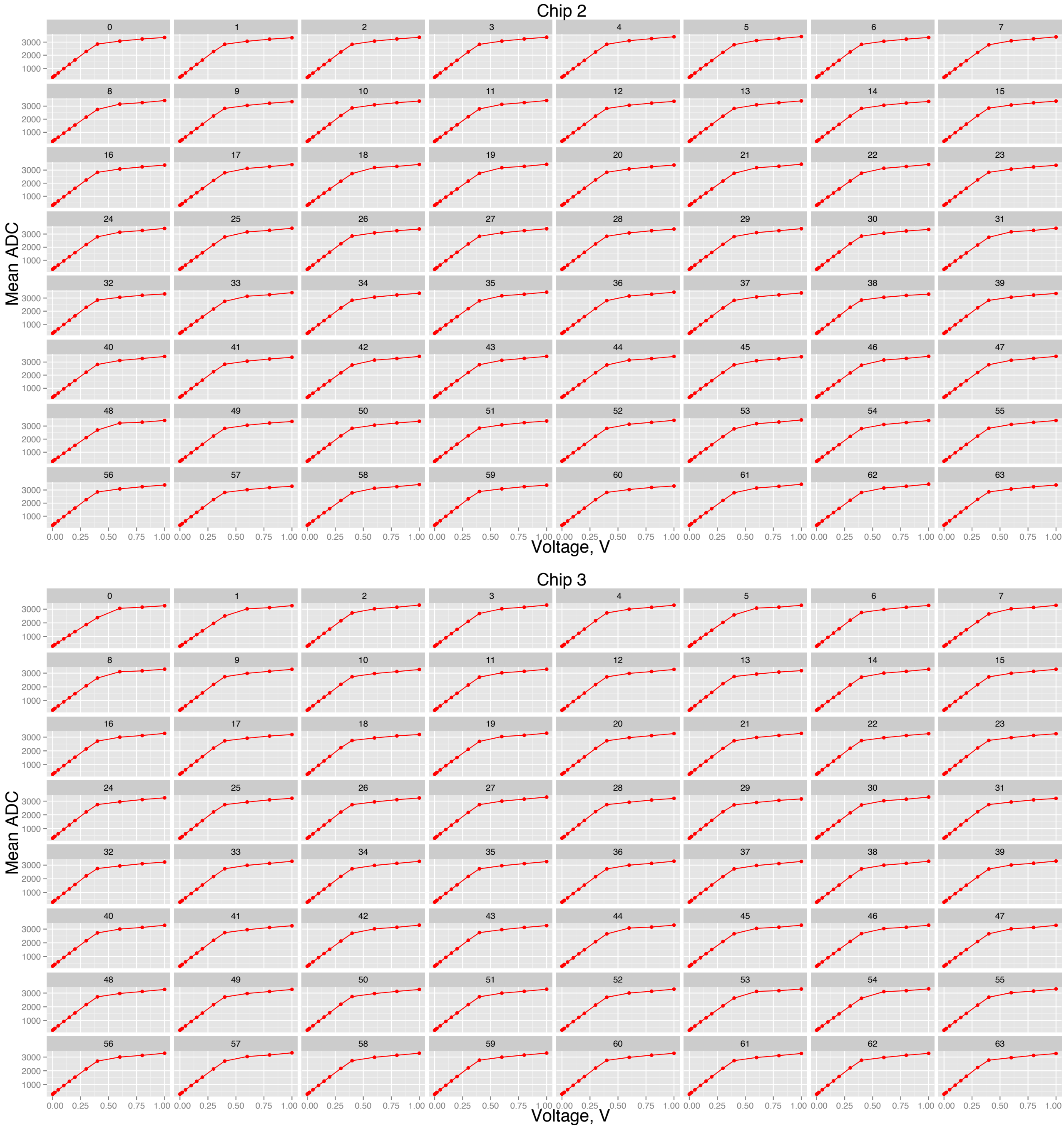


Figure 5.24 – Same as in the previous Figure 5.23, but for chips 2, 3 in slab 10.

Slab, mode	10, PP	10, CC	08, CC	05, PP
Average slope in chip 0	5704.3	5664.2	5888.6	7556.4
Average slope in chip 1	5829.6	5785.1	6014.1	7085.7
Average slope in chip 2	6362.6	6328.0	6547.4	7576.0
Average slope in chip 3	6151.0	6125.4	6611.5	7523.9
Slopes RMS within chip 0, %	7.29	7.44	7.12	6.55
Slopes RMS within chip 1, %	6.82	6.93	4.95	6.28
Slopes RMS within chip 2, %	2.31	2.42	2.34	2.36
Slopes RMS within chip 3, %	3.57	3.66	4.41	3.03
RMS across all slab channels, %	6.78	6.92	7.02	5.52
RMS across average chip slopes, %	4.99	5.11	5.86	3.15

Table 5.1 – Characteristics of linear slopes for voltages between 0.05 and 0.3 Volts.

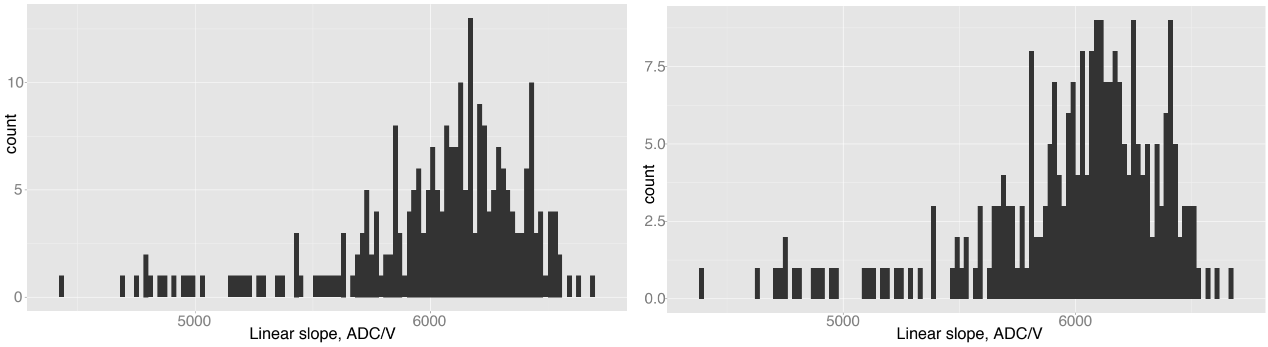


Figure 5.25 – Distribution of linear slopes in the injection voltage interval between 0.05 and 0.3 V for slab 10 in PP (left) and CC modes (right).

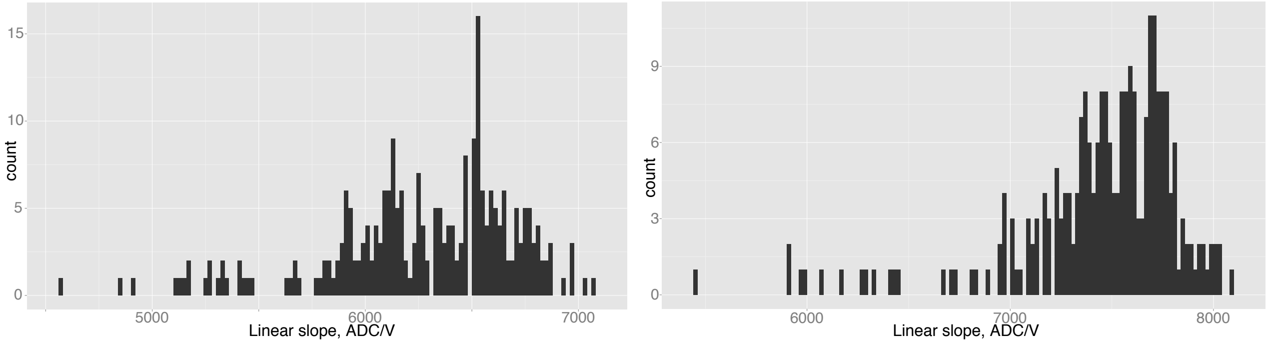


Figure 5.26 – Same as in the previous Figure 5.25, but for slab 8 in CC mode (left) and slab 5 in PP mode (right).

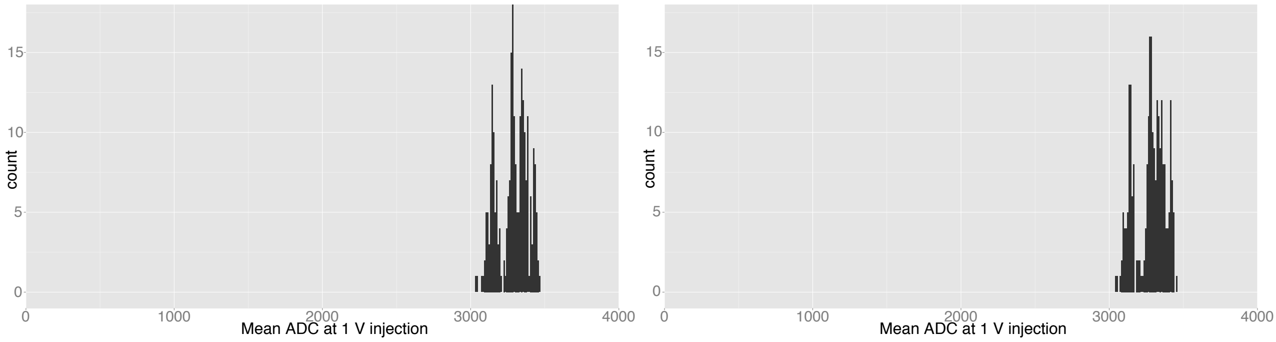


Figure 5.27 – Distribution of ADC values at 1 V injection, slab 10 in PP (left) and CC (right) modes. Four peaks correspond to different chips.

with the charge injection. The spread between the channels is at the level of a few percent. The charge inside the chip is injected through the small capacitors. Their values should also be uniform. The characteristics and

Slab, mode	10, PP	10, CC	08, CC	05, PP
Average ADC with 1 V injection in chip 0	3138.3	3130.8	3275.4	3599.0
Average ADC with 1 V injection in chip 1	3334.9	3324.1	3322.2	3364.5
Average ADC with 1 V injection in chip 2	3395.6	3387.7	3467.8	3455.5
Average ADC with 1 V injection in chip 3	3265.4	3262.8	3544.2	3385.4
Average ADC across chip averages with 1 V injection	3283.6	3276.4	3402.4	3451.1
RMS spread of ADCs with 1 V injection in chip 0, %	0.98	0.92	0.71	0.97
RMS spread of ADCs with 1 V injection in chip 1, %	0.93	0.89	1.27	1.19
RMS spread of ADCs with 1 V injection in chip 2, %	1.21	1.11	0.96	1.74
RMS spread of ADCs with 1 V injection in chip 3, %	1.00	0.91	0.91	1.60
RMS across average chip ADCs with 1 V injection, %	3.36	3.35	3.68	3.07

Table 5.2 – Characteristics of saturation at 1 V injection.

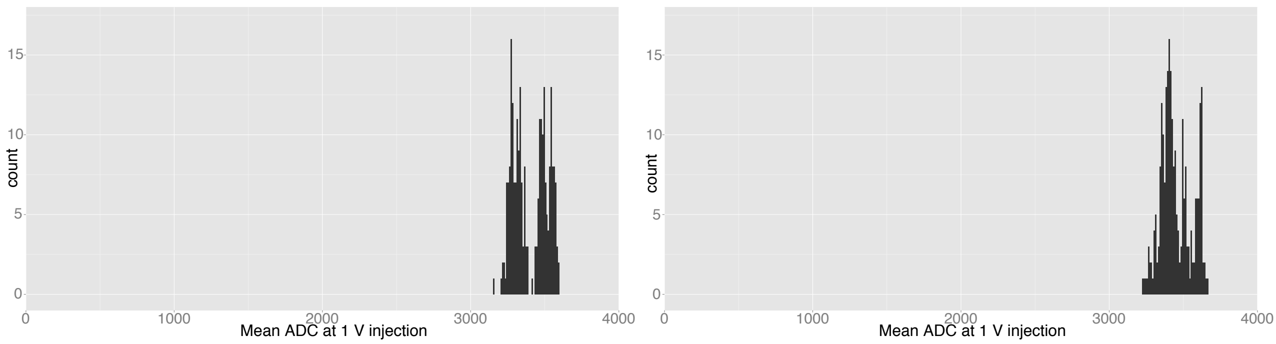


Figure 5.28 – Same as in the previous Figure 5.27, but for slab 8 in CC mode (left) and slab 5 in PP mode (right).

the uniformity of these capacitors are unknown, however, as they can not be measured externally. This question deserves a separate study in the future.

Small difference between PP and CC modes

A comparison between power pulsed (PP) and continuous (CC) operation modes is done per chip, per channel and per injection voltage for slab 10. As an example, Figure 5.29 shows the difference between the measured responses for the chip 2 of slab 10. Dependencies in all channels are similar. Relative difference between PP and CC (normalized to $(PP+CC)/2$, i.e. $2(PP-CC)/(PP+CC)$) is also shown in Figure 5.29. Similar plots for other slab 10 chips are shown in Appendix B (Figures B.9 - B.12). As it can be seen from columns 2, 3 in Table 5.1, the linear slopes in PP mode are slightly higher than in CC mode.

Problematic channels

As visible from Figures 5.23, 5.24 there are some channels with no signals. Such malfunctioning channels are the same in each slab: (chip, channel) = (0,53), (1,38), (1,40), (1,44). In addition, the channel (1,42) in slabs 5 and 8 is malfunctioning, but in slab 10 it is partially operational. Some unexpected signals are detected in other channels which are also mostly the same in every slab with little exceptions. Table 5.3 summaries all problematic channels with the bunch crossings (BX) where the wrong signals are detected and their average ADC position. The wrong signals appear in BX = 123, 248 suggesting that they are picked up from the digital signals in the boards having precise timing. The signals can be positive or negative, ie. below or above pedestal, depending on whether they are induced by the rising or falling edge of the digital signal. They do not depend on the injection. Note that the pedestal position varies from channel to channel and is also dependent on SCA memory slot, the average value is around 300.

Slab, mode		10, PP	10, CC		08, CC	05, PP
chip, channel	BX	wrong ADC	wrong ADC	BX	wrong ADC	wrong ADC
0, 47	248	319	324	123	297	372
0, 48				123		282
0, 52	248	235	240	123	248	
0, 53	248	239	240	123	334	272
0, 55	248	288	291	123		338
1, 31	248	354	358	123	352	356
1, 35	248	447	449	123	466	463
1, 38	248	637	637	123	639	682
1, 40	248	347	349	123	304	322
1, 42	many	spectrum	spectrum	123	321	348
1, 44	248	1185	1170	123	1265	1124
1, 47	248	311	313			
1, 48	248	213	217	123	262	
1, 49	248	230	233	123	252	238

Table 5.3 – Characteristics of problematic channels. Columns 3,4,6 and 7 are for slabs 10, 10, 8 and 6, respectively. PP and CC denote two modes of operation: power pulsed or continuous.

In total slab 10 has 13 problematic channels, slabs 8 and 5 each have 11. All problematic channels are in the chips 0 and 1, while the chips 3 and 4 are fully operational. This also proves that the problems are related to

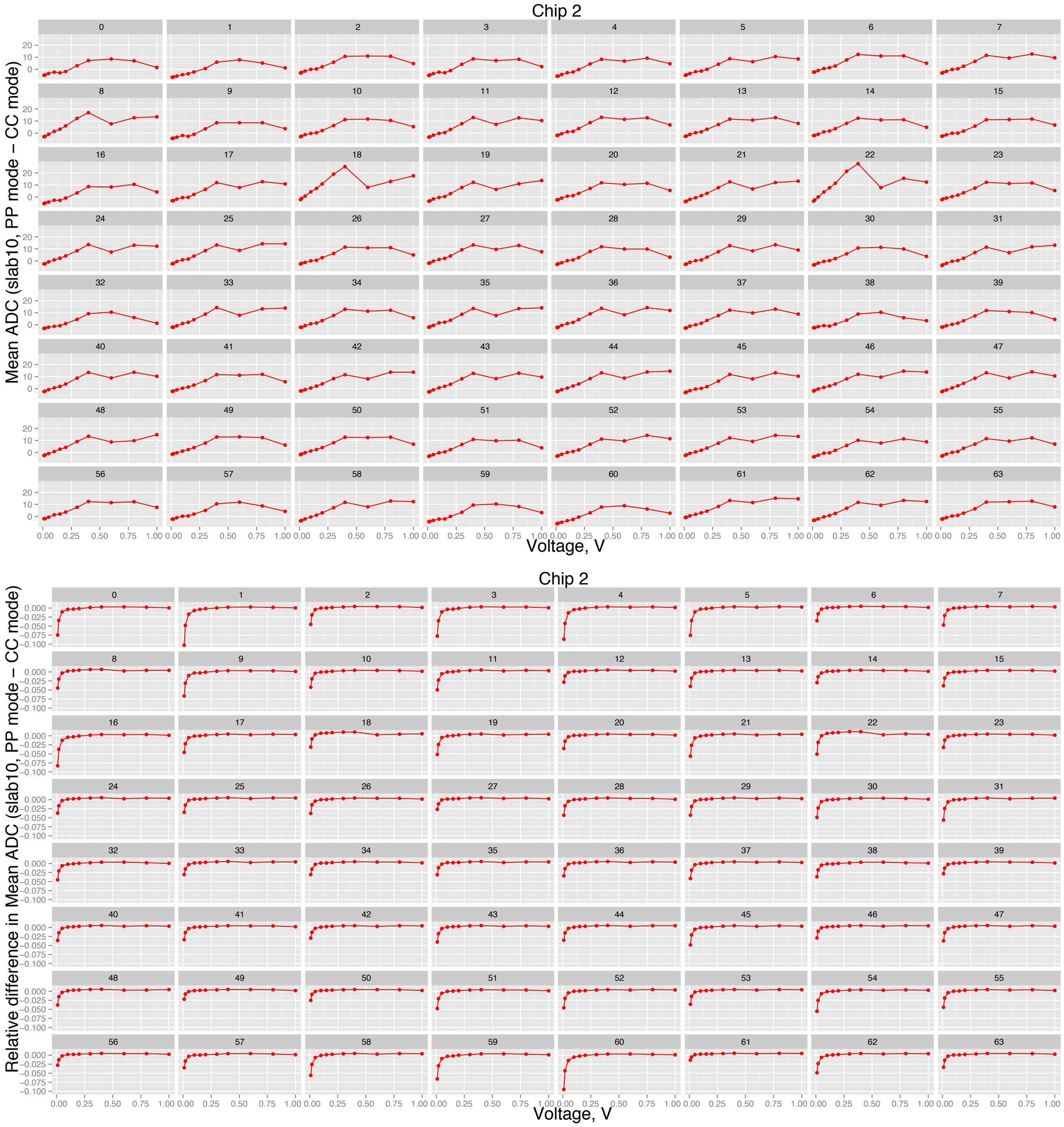
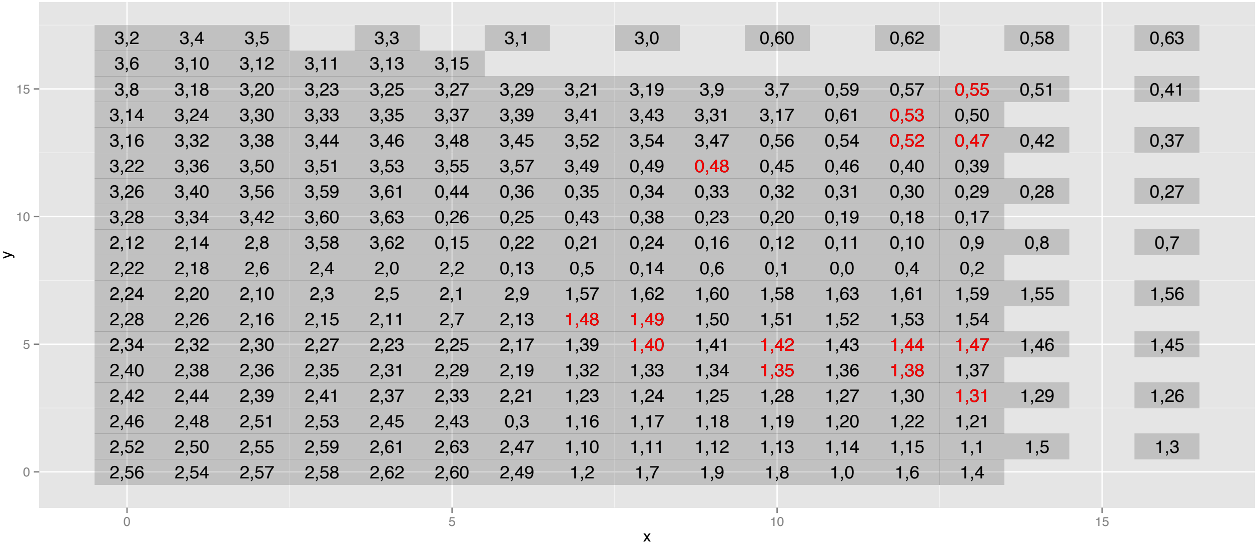


Figure 5.29 – Absolute (top) and relative (bottom) difference of the average channel response between PP and CC injection modes in slab 10 (chip 2).

the PCB boards and not to the chips. It is interesting that all problematic channels are grouped in two lines on PCB. This is visible from Fig. 5.30 which shows the geometrical placement of pixels connected to a given (chip, channel). The problematic channels are marked in red. Note, that at the top and on the right the boundary pixels are connected together and readout by groups of 2 or 4.

In the new boards FEV10/11 the situation is much improved. At the test beam the fraction of noisy channels was at the level of a few percent. The analysis of channel functionality based on charge injection in the new FEV10/11 goes beyond the scope of this thesis.



3,2	3,4	3,5		3,3		3,1		3,0		0,60		0,62		0,58	0,63
3,6	3,10	3,12	3,11	3,13	3,15										
3,8	3,18	3,20	3,23	3,25	3,27	3,29	3,21	3,19	3,9	3,7	0,59	0,57	0,55	0,51	0,41
3,14	3,24	3,30	3,33	3,35	3,37	3,39	3,41	3,43	3,31	3,17	0,61	0,53	0,50		
3,16	3,32	3,38	3,44	3,46	3,48	3,45	3,52	3,54	3,47	0,56	0,54	0,52	0,47	0,42	0,37
3,22	3,36	3,50	3,51	3,53	3,55	3,57	3,49	0,49	0,48	0,45	0,46	0,40	0,39		
3,26	3,40	3,56	3,59	3,61	0,44	0,36	0,35	0,34	0,33	0,32	0,31	0,30	0,29	0,28	0,27
3,28	3,34	3,42	3,60	3,63	0,26	0,25	0,43	0,38	0,23	0,20	0,19	0,18	0,17		
2,12	2,14	2,8	3,58	3,62	0,15	0,22	0,21	0,24	0,16	0,12	0,11	0,10	0,9	0,8	0,7
2,22	2,18	2,6	2,4	2,0	2,2	0,13	0,5	0,14	0,6	0,1	0,0	0,4	0,2		
2,24	2,20	2,10	2,3	2,5	2,1	2,9	1,57	1,62	1,60	1,58	1,63	1,61	1,59	1,55	1,56
2,28	2,26	2,16	2,15	2,11	2,7	2,13	1,48	1,49	1,50	1,51	1,52	1,53	1,54		
2,34	2,32	2,30	2,27	2,23	2,25	2,17	1,39	1,40	1,41	1,42	1,43	1,44	1,47	1,46	1,45
2,40	2,38	2,36	2,35	2,31	2,29	2,19	1,32	1,33	1,34	1,35	1,36	1,38	1,37		
2,42	2,44	2,39	2,41	2,37	2,33	2,21	1,23	1,24	1,25	1,28	1,27	1,30	1,31	1,29	1,26
2,46	2,48	2,51	2,53	2,45	2,43	0,3	1,16	1,17	1,18	1,19	1,20	1,22	1,21		
2,52	2,50	2,55	2,59	2,61	2,63	2,47	1,10	1,11	1,12	1,13	1,14	1,15	1,1	1,5	1,3
2,56	2,54	2,57	2,58	2,62	2,60	2,49	1,2	1,7	1,9	1,8	1,0	1,6	1,4		

Figure 5.30 – Geometrical position of pixels, connected to (chip, channel) pair. Noisy channels are shown in red. Channels at the top and on the right are connected to 2 or 4 pixels.

5.2.2 “Square” events

The most recent technological prototype is based on PCB FEV11. Every slab have 4 sensors each with 8×8 pixel matrices and 5.5×5.5 mm^2 pixels. The GR design of the sensors is improved.

The rate of “square” events in the new slab is measured in November’15 CERN SPS data with up to 150 GeV showers [175]. To maximize the effect, the beam was shooting in the PCB center corresponding to the corners of four sensors, 30 mm of tungsten absorber plates ($8.4X_0$) has been placed in front of the detector.

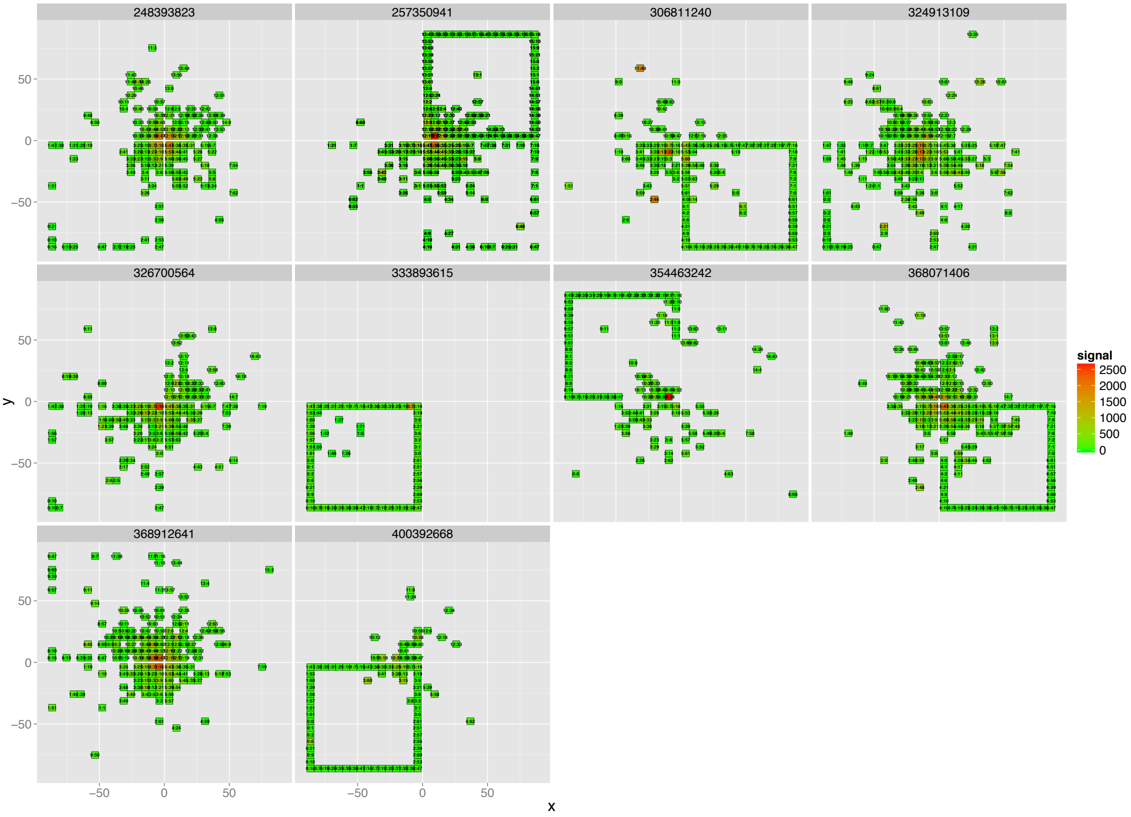


Figure 5.31 – "Square" events observed in one layer (DIF 1) in November’15 in 100 GeV e^+ sample (run 211). Detector layers are perpendicular to the beam.

Square-events are observed in every sensor. Some examples are shown in Figure 5.31, more can be found in

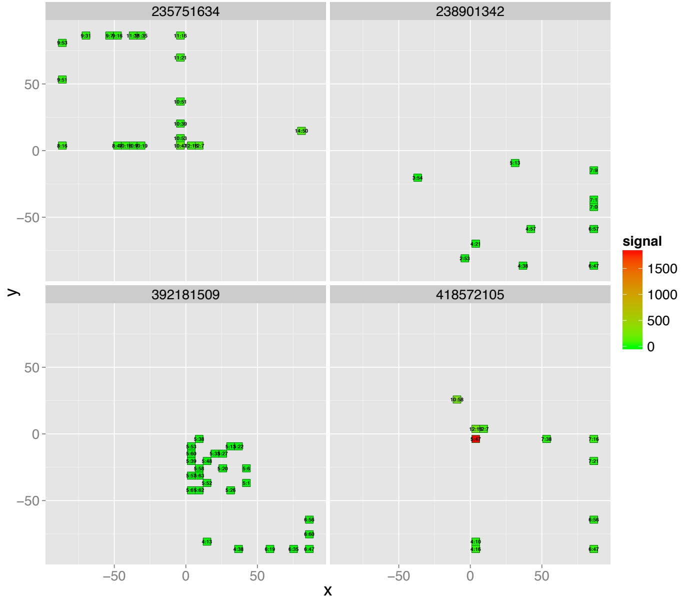


Figure 5.32 – Examples of "square" events which are difficult to observe if they overlap with a shower (run 211, DIF 2).

Figures B.14 in Appendix B. Table 5.5 summarizes the observed raw rate of the square events in three detector layers for e^+ and π^+ samples. As one can see, the rate is less than 0.04% even for the 150 GeV e^+ showers. Sometimes, the "squares" are not well visible, like in Figure 5.32, but the automatic algorithm still finds them. This is not always the case for the physics prototype, because with smaller 6×6 matrix of pixels it is less trivial to distinguish the "square" event from the shower.

The comparison of event rates in the physics and technological prototypes is presented in Figure 5.33 and is also summarized in Table 5.4 (Table 5.5) for the physics (technological) prototypes. In the physics prototype only the layer 15 was studied with similar amount of absorber ($8X_0$) in front. Note that the beam was shooting in the middle of the physics prototype, ie. in the middle of the central sensor. In the technological prototype it was shooting in the corners of 4 sensors to enhance the effect. Note also, the beam composition and the widths could be different in the two tests.

Energy, GeV	Run numbers	Beam	N SQ-events (layer 15)	Total N evt	SQ rate
4	630065(67)	π^+, e^+	3	$\sim$ 191000	0.0000157
12	630061(62)	π^+, e^+	3	$\sim$ 25500	0.0001176
25	630037(38,39)	π^+, e^+	54	$\sim$ 171000	0.0003158
32	630033(34,36)	π^+, e^+	123	$\sim$ 265000	0.0004642
50	630084	π^+	26	$\sim$ 54000	0.0004815

Table 5.4 – "Square" events in 15th layer ($\sim 8X_0$ of W) of the SiW ECAL physics prototype [175]. The 5th column lists the number of events with showers. Last column contains the fraction of "square" events relative to the numbers in the 5th column. The beam spots are shown in Figures A.13 and A.14 in Appendix A.1, the beam was shooting in the middle of the sensor.

The low probability of "square" events in the technological prototype demonstrates the advantage of the improved GR design.

To check that we are well sensitive to the "square" events, the effect was also studied in the data sample when the ECAL was turned around the vertical axis by 90°. It was positioned such that the beam entered the sensor at the edge, then passed (maximally) through ∼9 cm of silicon and exited from the opposite edge. After that the beam could pass in the same way the second sensor. Due to an angular dispersion of the beam and an imperfect alignment, such events were rare, but were present. The PCB was aligned vertically such that the beam passed through its central region, ie. at the sensor borders.

In this orientation, the particles could leave very big energy depositions in the peripheral sensor regions. Indeed, many "square" events are observed, see Figure B.15 in Appendix B. Figure 5.34 shows 150 GeV π^+ event with all four sensors affected.

In Jun'17 beam tests at DESY the fraction of "square" events was almost negligible because of relatively low e^- beam energy (1–6 GeV) [176].

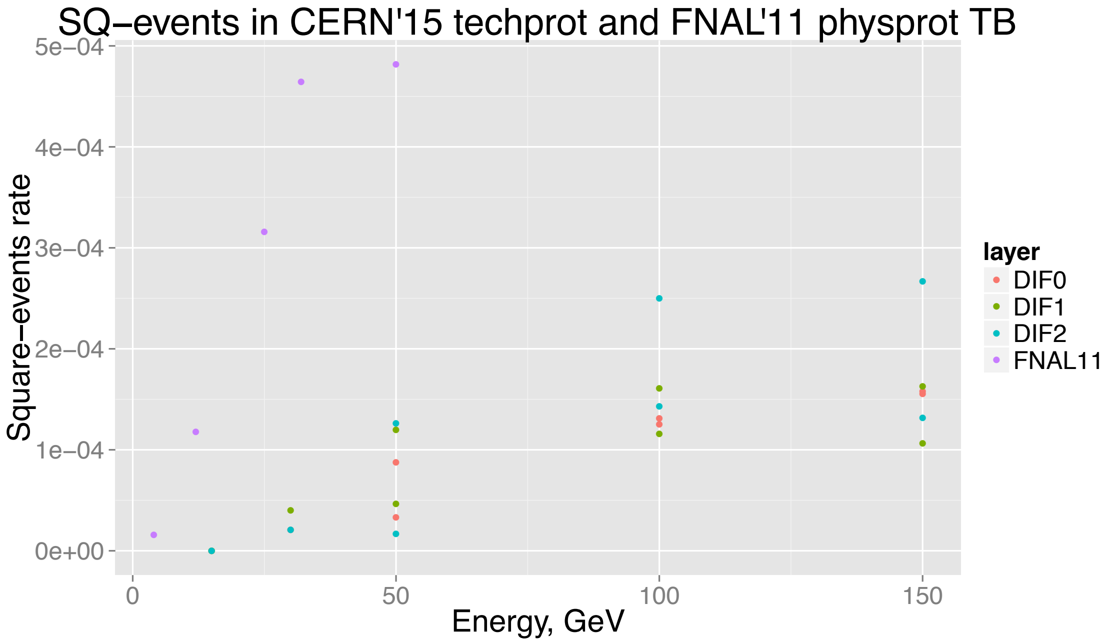


Figure 5.33 – Fraction of “square” events in SiW ECAL technological prototype tests in November’15 at CERN and in 15th layer of the physics prototype in FNAL’11 TB campaign. Note that for the physics prototype this effect is underestimated, as the beam was centered in the middle of the sensor and not in the corners of four wafers as for the technological prototype.

Energy, GeV	Run numbers	Beam	N DIF	N SQ-events	Total N evt	SQ rate
15	300	e^+	DIF 0	0	∼ 22000	0.0
			DIF 1	0	∼ 22000	0.0
			DIF 2	0	∼ 25000	0.0
30	301	e^+	DIF 0	1	∼ 48000	0.0000208
			DIF 1	2	∼ 50000	0.00004
			DIF 2	1	∼ 48000	0.0000208
50	302	e^+	DIF 0	2	∼ 60000	0.0000333
			DIF 1	4	∼ 86000	0.0000465
			DIF2	1	∼ 60000	0.0000167
100	204, 205, 206, 207, 210, 211	e^+	DIF 0	15	∼ 120000	0.000125
			DIF 1	24	∼ 150000	0.000161
			DIF 2	30	∼ 120000	0.00025
150	79, 82, 84, 86, 88, 114	e^+	DIF 0	6	∼ 38000	0.0001579
			DIF 1	7	∼ 43000	0.0001627
			DIF 2	5	∼ 38000	0.0001316
50	282	π^+	DIF 0	9	∼ 103000	0.0000874
			DIF 1	12	∼ 100000	0.00012
			DIF 2	13	∼ 103000	0.0001262
100	280, 281	π^+	DIF 0	18	∼ 137000	0.0001314
			DIF 1	16	∼ 138000	0.0001159
			DIF 2	20	∼ 140000	0.0001429
150	97, 99, 100, 102, 103	π^+	DIF 0	7	∼ 45000	0.0001556
			DIF 1	5	∼ 47000	0.0001064
			DIF 2	12	∼ 45000	0.0002667

Table 5.5 – “Square” events in the technological prototype [175]. The 5th column lists the number of events with showers. Last column contains the fraction of “square” events relative to the numbers in the 5th column. The beam was shooting in the corners of four sensors.

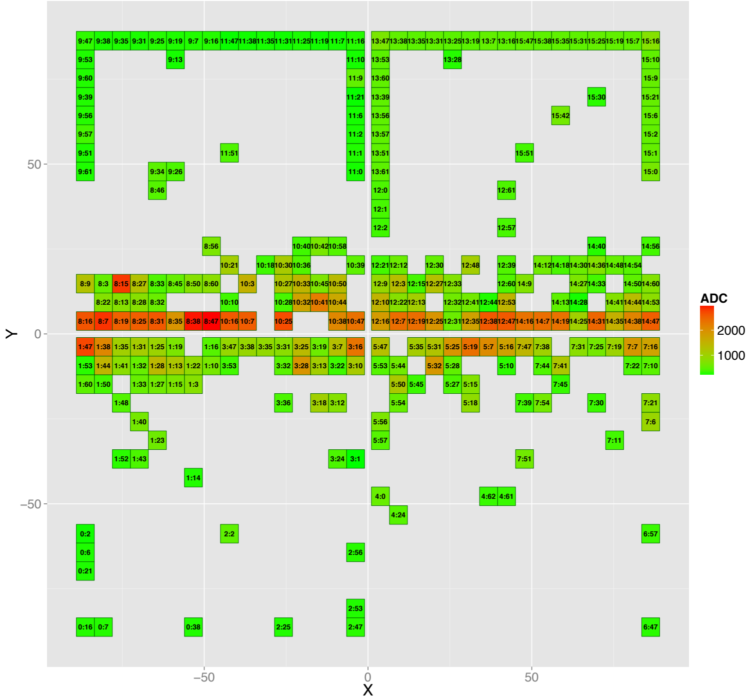


Figure 5.34 – An example of the 150 GeV π^+ event with 4 “squares”. The prototype was turned by 90° and put parallel to the beam.

5.2.3 Longitudinal shower profile

The analysis of Jun’17 beam tests at DESY with 7 slabs is ongoing. Here, the first very preliminary results are presented on the longitudinal shower profile. The data have been taken with 2.1, 2.1, 2.1, 4.2, 4.2, 4.2, 6.3 mm of tungsten installed before and between 7 ECAL layers.

Only events from those readout cycles are accepted where none of the chips reports more than 14 events. If the chip reports 15 events, there might be a possibility that there was 16th event which was not recorded because the SKIROC2 memory was exhausted. Note, all SKIROC2 chips take data independently and become full at different moments. By requiring maximally 14 events in all chips the problem of the limited memory is mitigated.

To take into account the possibility that some physical hits from the shower in bunch crossing BX are delayed until BX+1, both BX and BX+1 hits are merged if BX+1 is also triggered. The delayed triggers in BX+1 were visible in high energy showers at SPS, at lower DESY energies the effect should be smaller. Note, instead of physical hits there might be a retrigger in BX+1 and this introduces some bias. The retrigger hits are typically close to pedestals, therefore the bias is small, however.

The mismatch in the timing of different ECAL layers within ±1 BX has also been taken into account in reconstructing the full ECAL event.

The preliminary longitudinal shower profile for 3 GeV e^+ showers is shown in Figure 5.35. Every point is the average over sum of all pedestal subtracted ADC values in the given layer.

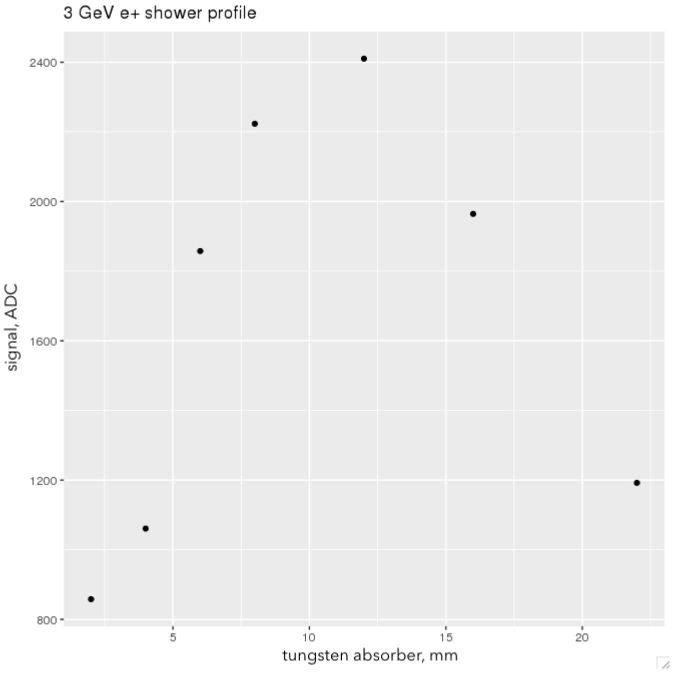


Figure 5.35 – Longitudinal profile of 3 GeV e^+ shower: total signal per layer versus the tungsten thickness in front of the layer (very preliminary).

5.2.4 Debugging of data corruption

The data from each slab are read out independently into separate data streams. Their format is shown schematically in Figures 5.36 and 5.37. The data is split into separate acquisition cycles (spills). In the data stream they are delimited by the header and the trailer containing the spill number. Inside these delimiters, there are CHIP blocks for every chip which triggers at least once during the acquisition cycle. The CHIP data are delimited by the CHIP header and the trailer containing the CHIP number. The detailed format of the spill and CHIP blocks can be consulted in Figure 5.36.

Every chip during one acquisition can record up to 15 events. All of them form the CHIP data block shown in Figure 5.37. It is send out by SKIROC2 and transparently propagated through the DAQ boards which only decorate it with the CHIP and the spill headers and trailers.

For future references, the mapping of FEV10/11 pixels to (chip:channel) pair is shown earlier in Figure 5.14.

After debugging the data stream from Nov'15 tests, the possible ways of data corruption were classified into several groups [184]. They are presented in the following. June'17 data analysis confirms these observations.

Too short spill readout time ($t = 10$ msec)

If the SKIROC2 readout time in the end of the spill is too short, two things may happen. They were observed in Nov'15 data when the readout time was set to t=10 msec

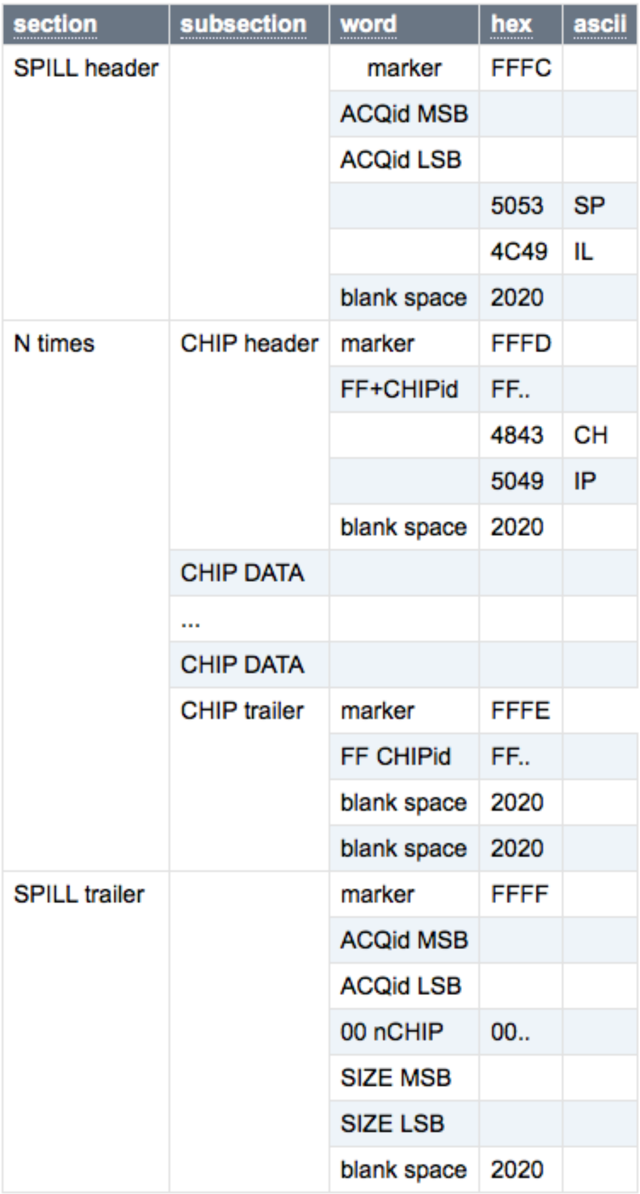
- One or several spills may disappear, as shown schematically in Figure 5.38 (left). The spill blocks before and after missing spills do not have any visible signs of corruption. Eg. in the run 114 $\sim$ 65% of spills are missing in all three layers.

- Alternatively, the header of the next spill ($N + 1$) may arrive before the trailer of spill N, as shown in Figure 5.38 (right), and overwrite the CHIP data (most probably, chip 0). It seems that data in CHIP blocks after the new spill are not affected, but the chip number counter is broken. Potentially it is possible to recover all chip blocks, except the one which is overwritten. In the run 114 three layers (DIF 0, 1, 2) have 833, 415 and 3196 corruptions of this type, respectively, in 11686, 11816 and 11723 recorded spills.

Fortunately, the data corruption was identified "online" and was fixed by setting the readout time to t=50 msec. The author of the thesis performed the offline analyses, presented above and in the following.

Too high spill frequency ($f = 5$ Hz)

At spill frequency $f = 5$ Hz (with readout time t=50 msec) the last chip (#15) has lower occupancy than its neighbours, see Figure 5.39. This also have been observed online and fixed by setting the spill frequency to



section	subsection	word	hex	ascii
SPILL header		marker	FFFC	
		ACQid MSB		
		ACQid LSB		
			5053	SP
			4C49	IL
		blank space	2020	
N times	CHIP header	marker	FFFD	
		FF+CHIPid	FF..	
			4843	CH
			5049	IP
		blank space	2020	
	CHIP DATA			
	...			
	CHIP DATA			
	CHIP trailer	marker	FFFE	
		FF CHIPid	FF..	
		blank space	2020	
		blank space	2020	
SPILL trailer		marker	FFFF	
		ACQid MSB		
		ACQid LSB		
		00 nCHIP	00..	
		SIZE MSB		
		SIZE LSB		
		blank space	2020	

Figure 5.36 – One slab data format [180].

4 Hz. The offline analysis has shown that chip 15 data block is often overwritten by the next spill header, as schematically shown in Figure 5.40. Sometimes the same data corruption is observed for chip 14, but with significantly lower rate. In total, eg. in the run 313 three layers (DIF 0, 1, 2) have 3275, 3274 and 3274 corruptions of these types in 30231 spills.

It is interesting that at even higher spill frequency $f = 10$ Hz (and readout time $t = 97.5$ msec, run 326), almost half of the chips are missing because next spill overwrites data blocks of many chips, see Figure 5.41 (left). The corruption rate remains about the same (in runs 326 and 313), namely, in 13863 spills there were 1528, 1494 and 1580 corruptions in three layers (DIF 0, 1, 2). With lower spill frequency $f = 4$ Hz (run 330) the corruptions disappear as visible from Figure 5.41 (right). This frequency value was set for the rest of Nov'15 tests.

Note, that in June'17 beam tests at DESY $f = 5$ Hz spill frequency was used without problems. This may be related to smaller data blocks for smaller particle rates and energies.

Delayed triggers in high energy showers

In Nov'15 beam tests it was found that the chips positioned in the core of high energy showers (100–150 GeV e^+) sometimes trigger not in the right bunch crossing BX but in BX+1. For a moment, the reason of this effect is unknown.

This is visible from the occupancy maps in BX, shown in Figure 5.42 (left) (150 GeV e^+, run 394). The rows correspond to the layers with DIF 0, 1 and 2. There were 2.4 X_0 of tungsten in front of the first layer (DIF 1)

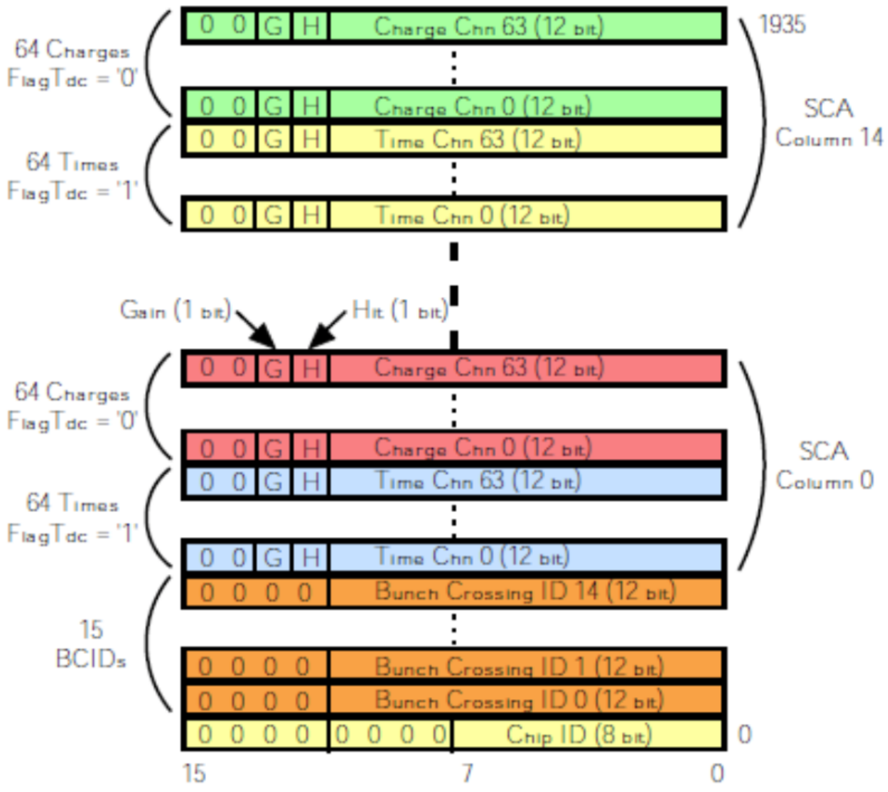


Figure 5.37 – SKIROC2 chip hexadecimal data format [180].

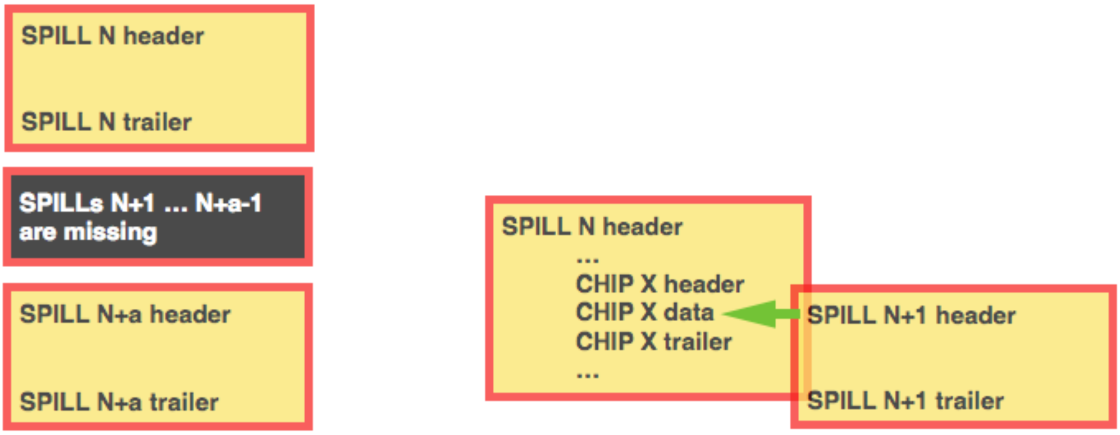


Figure 5.38 – Data corruption types observed in run 114 with too short readout time t=10 msec: one or many missing spills (left) and overwriting of CHIP X data with SPILL N+1 header (right).

and 4.2 X_0 between the first and the second (DIF 2) and the second and the third (DIF 0). The beam was shooting in chip 12 positioned in the lower left corner of the upper right sensor. As one can see, the occupancy there in the bottom and in the top plots (second and third layers, DIF 2, 0) is surprisingly low. In the first layer (DIF 1), shown in the middle, there is no visible drop of occupancy which means that the effect depends on the shower intensity: in the first layer after 2.4 X_0 of tungsten the shower is not sufficiently developed. After adding the hits in BX+1, the occupancy map becomes normal as shown in Figure 5.42 (right).

This effect was observed during the test beam online. In particular, it was checked that the effect does not depend on the particle rate in the beam.

In the offline analysis it was found that in addition to such delayed by 400 nsec triggers (400 nsec is the bunch clock period), there are events where part of the event is in BX and part is in BX+1. Quantitatively this effect could be estimated with Figure B.16 in Appendix B. In addition same effect is observed for other Nov'15 energies: 15, 30 and 50 GeV, but drop of occupancy, as in Figure 5.42 (left) for 100–150 GeV showers, is not distinguishable by eye.

It is interesting that one chip channel normally (in $>$ 95% of such events) triggers either in BX or in BX+1

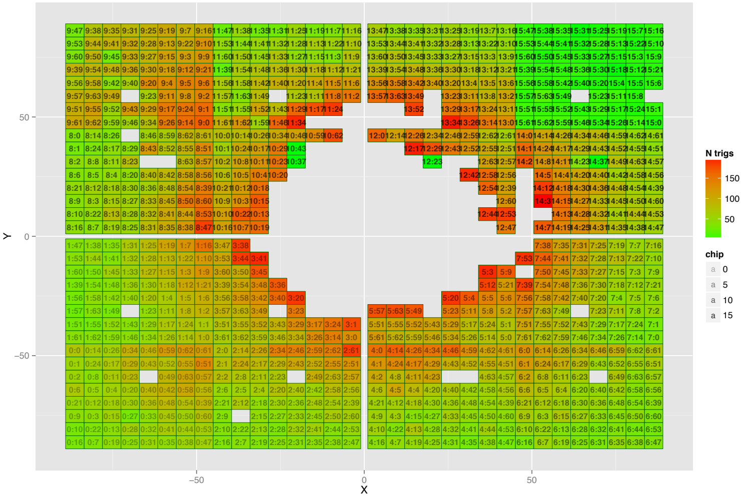


Figure 5.39 – Occupancy map of one layer (DIF 1, run 313). The central part with higher occupancy is excluded. Chip 15 in the upper right corner has lower occupancy due to data corruption caused by too high spill frequency (5 Hz) [183].

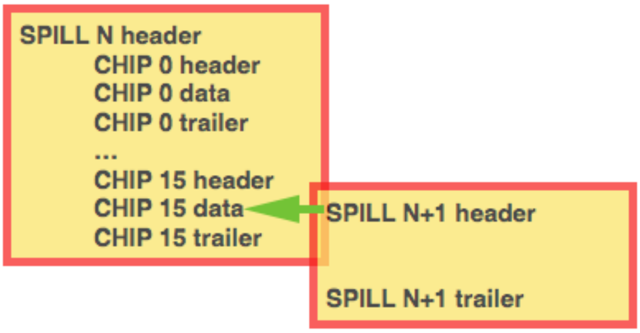


Figure 5.40 – Corruption of chip 15 data block by the next spill header at $f = 5$ Hz spill frequency.

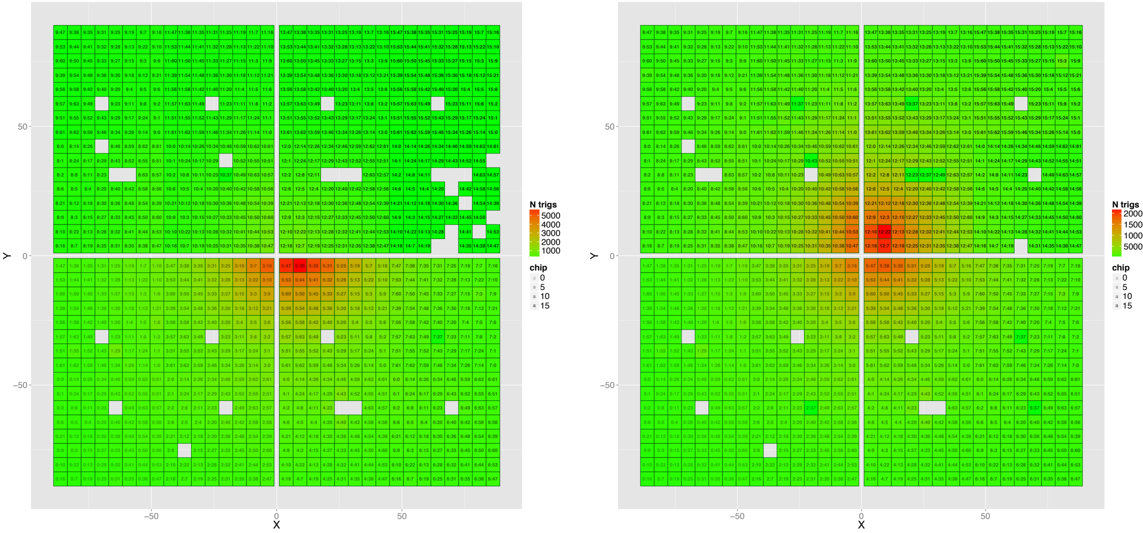


Figure 5.41 – Upper half of the layers has low occupancy due to data corruption at spill frequency $f = 10$ Hz (run 326, DIF 1, left plot). The occupancy map becomes normal at $f = 4$ Hz (run 330, same layer, right plot).

but not in both. Ie. there is almost no intersection between the triggered channels in BX and BX+1. The exceptions of this rule can be due to retriggers in BX+1 which are not always distinguishable from the shower hits appearing in BX+1. An example of such event is shown in Figure 5.43.

It is not clear, whether merging of the hits in BX and BX+1 gives the right image of the shower. The occupancy shown in Figure 5.42 (right) after this procedure looks normal. It is not evident, however, how to

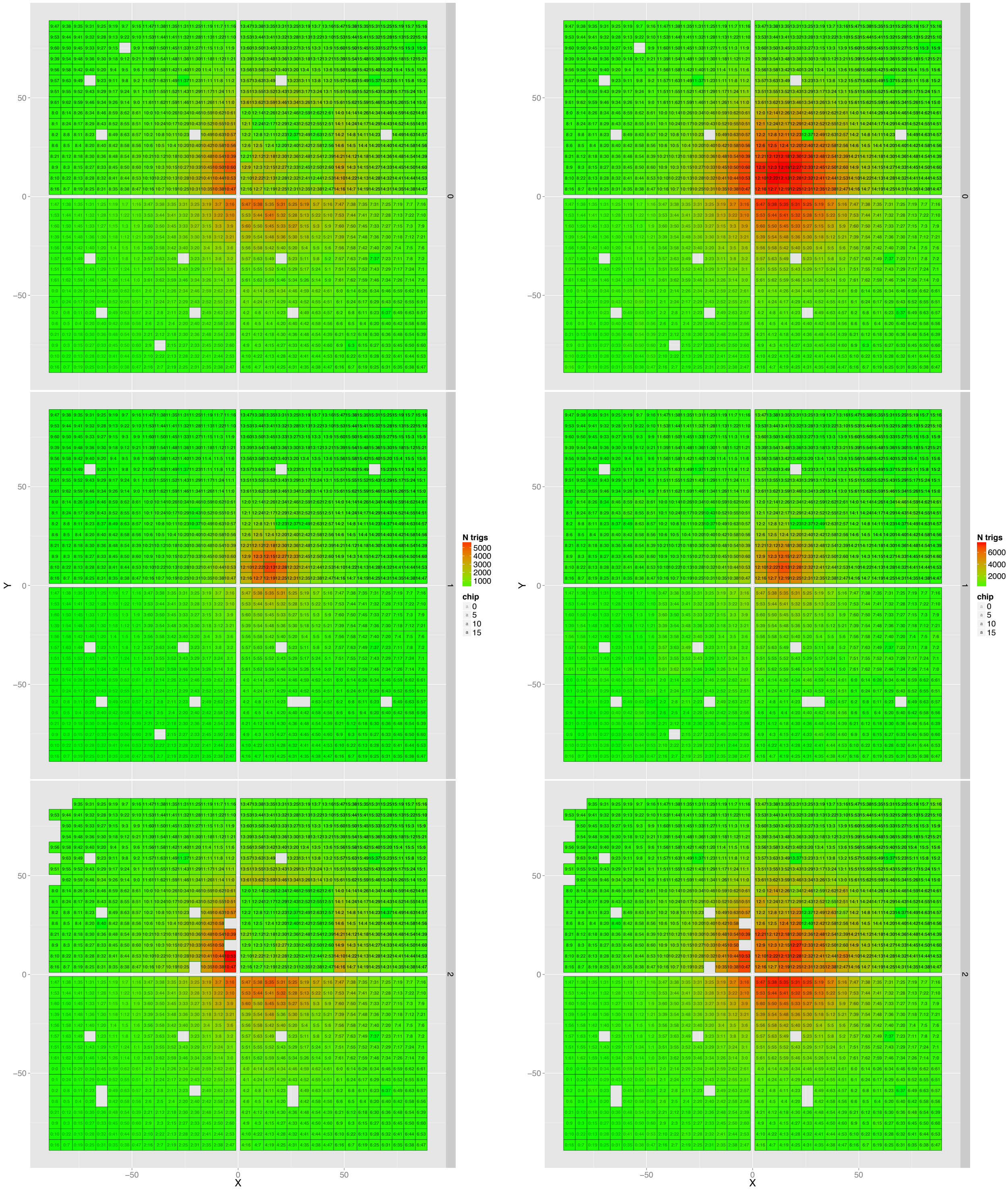


Figure 5.42 – Occupancy maps in three layers for 150 GeV e^+ sample (run 394). The amount of the tungsten absorber in front of the layers and their correspondence to the DIF numbers is 2.4 X_0, (first layer = DIF 1), 4.2 X_0, (second layer = DIF 2), 4.2 X_0, (third layer = DIF 0). The left column is for the trigger channels in the first BX detected by any chip, the right column shows the sum of occupancy maps in BX and BX+1.

get the correct energy estimation. Eg. in the events triggered both in BX and BX+1 sometimes there are not triggered channels with the signals significantly higher the trigger threshold. Clearly, this effect requires more studies before it is fully understood and fixed.

It is important for high energy showers. At smaller beam energies (1–6 GeV) in Jun'17 at DESY its rate was significantly lower.

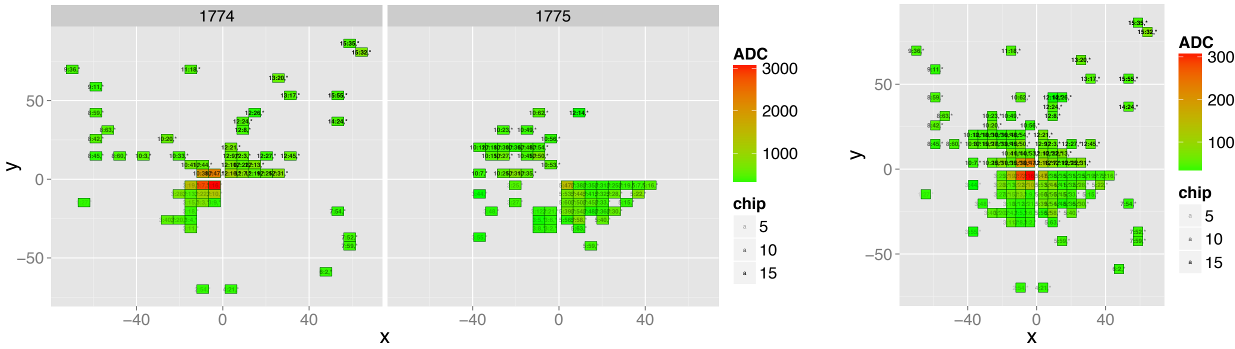


Figure 5.43 – An example of the event (100 GeV e^+ after $8.4X_0$ of W) with delayed trigger in BX and BX+1: left plot corresponds to hits triggered in BX=1774, middle – in BX+1=1775. Right plot corresponds to merged event. Chips 3, 8, 10, 12 are triggered in both BX and BX+1, while chip 5 only in BX+1.

Missing FF CHIPid number in CHIP header or trailer

The FF CHIPid number n_{chip} (sequential number of the CHIP DATA block among other chip blocks in the spill) can be missing in CHIP header or/and trailer block, while the corresponding CHIP DATA is correct, see Figure 5.44. Note, n_{chip} is not the same with chip number i. In case of this effect the chip is missing in the event display, as shown in Figure 5.45. Often this effect corrupts data in next $n_{chip}+1$ CHIP DATA block, eg. it may have $SCA > 15$ or $BX > 4095$ (bunch crossing number is coded in 12 bits). This data corruption type is rare. For example, in run 211 about 220, 250 and 750 chip data blocks are affected in DIF 0, 1, 2, respectively, out of $\sim 79\times10^3$, 78.8×10^3 and 206.5×10^3 in total.

Figure 5.44 – FF CHIPid number can be missing in header or/and trailer of the CHIP DATA block.

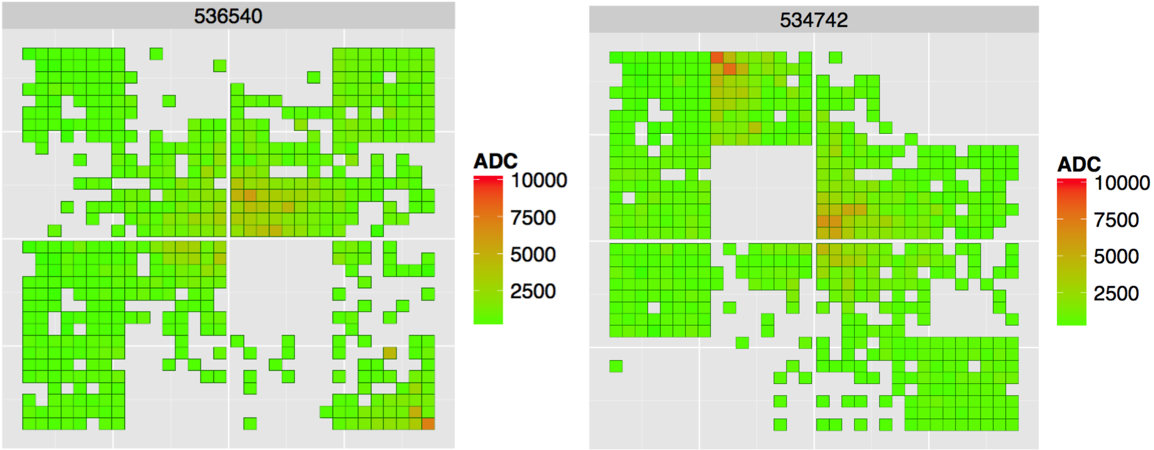


Figure 5.45 – Sum of triggered ADC counts in two different spills (given at the top of the plots) per pixel. Chips 5 (left) and 10 (right) are missing because the chip number is not present in the CHIP header or the trailer.

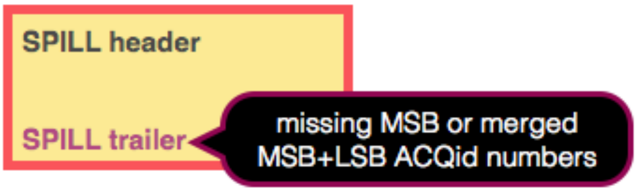


Figure 5.46 – Schema of the ACQid corruption in the SPILL trailer.

Wrong spill number in the SPILL trailer

The spill number in the SPILL trailer might be corrupted at the byte level, eg. the most significant byte (ACQid MSB) can be missing or it can be merged with the least significant byte (ACQid LSB), see Figure 5.46. Such corruptions are rare, for example in run 211 this problem was observed in 8, 12 and 142 spills in DIF 0, 1 and 2, respectively, out of 19259 spills in total.

5.3 Fractal dimension study

The measurement of the fractal dimension of the transverse shower image can help to distinguish electromagnetic and hadronic showers. This was proposed in [23–25].

The fractal dimension can be calculated from the number of hits in the same shower image obtained with two calorimeter granularities. In practice, one can degrade the granularity by combining $\alpha \times \alpha$ cells into one (starting eg. from the center of the slab). $\alpha = 1$ corresponds to the nominal granularity. As an alternative, here we consider $\alpha = 2$, ie. the calorimeter area is divided into squares comprising four pixels and they are considered as new effective "cells". The "cell" is hit if any of 4 pixels inside is hit. After denoting the number of hits as N_α, the fractal dimension can be derived from the comparison of granularities α and β as

$$FD_{\alpha\beta} = \frac{log(N_\alpha/N_\beta)}{log(\alpha)}. \tag{5.2}$$

and for $\alpha = 1$, $\beta = 2$:

$$FD_{12} = log_2(\frac{N_1}{N_2}) \tag{5.3}$$

Up to now, the fractal dimension of the showers was determined only in simulation. Thanks to the record granularity of the SiW ECAL technological prototype slabs, it can also be measured in Nov'15 data, in spite of the fact that the setup was equipped with only three layers. Even one layer is sufficient to obtain the transverse shower image. The measurements are complicated by the presence of instrumental effects, however, as the data have been accumulated during the commissioning phase of the prototypes. Therefore, they should be considered as preliminary and with the grain of salt.

For MIP-like particles leaving one hit in the layer, the fractal dimension is equal to 0. Positron showers have a dense, energetic core and loose periphery and exhibit the largest fractal dimension. It grows with the energy of the beam. Pion showers can be a composition of charged hadron tracks (with zero fractal dimension) and localized sub-showers. Their overall dimension is in average less than for the positrons of the same energy.

5.3.1 Event selection

Table 5.6 summarizes the data samples, beam particles and energies used in this analysis. About $8.4X_0$ of absorber material was in front of the prototype, while no tungsten plates were install between the layers. The layers were mounted on 2 cm aluminum plates, however, which act as an absorber between the layers (0.22 X_0.

The analysis procedure is described in the following.

- Only acquisition periods in SPS spills are selected, no-beam periods are excluded by requiring the minimal activity in the layer integrated over one acquisition.
- There should be coincidence between all three layers within the synchronization mismatch of ±1 BX.
- Hits in BX and BX+1 bunch crossing are merged. This is done to recover events with delayed trigger in BX+1, but also adds the background from the retrigger events.
- To remove events where the SKIROC2 memory becomes full, it is required that the chips should have at least one empty SCA in the end of acquisition with one exception: for the four central chips with delayed triggers, BX+1 is allowed to be recorded in the last SCA.

- Every chip should be triggered at least once during the acquisition to avoid missing CHIP block problems. This requirement potentially brings bias to the analysis.
- Note, no attempt to improve the purity of the beam and no ECAL calibration per channel was performed.

Energy, particle	Run number	N events
15 GeV, e^+	300	1513
30 GeV, e^+	301	2992
50 GeV, e^+	302	644
100 GeV, e^+	211	5876
150 GeV, e^+	114, 79, 86, 88	680, 142, 1156, 185
50 GeV, π^+	282	1222
100 GeV, π^+	281	640
150 GeV, π^+	247	444
150 GeV, e^+, π^+	97, 99, 100, 102, 103	212, 418, 20, 57, 509

Table 5.6 – Data samples used in the fractal dimension study.

5.3.2 Results

Figure 5.47 shows the transverse fractal dimension FD_{12} of 50 GeV e^+ showers measured in 3 detector layers. Similar distributions for 15, 30, 50, 100 and 150 GeV e^+ showers are given in Appendix B.5, Figure B.18. One can see that in average FD_{12} monotonically decreases from the first layer to the third $FD_{12}(\text{DIF1}) > FD_{12}(\text{DIF2}) > FD_{12}(\text{DIF0})$ (the order of layers to the beam corresponds to DAQ numbering DIF 1, 2, 0), probably mainly due to 2 cm aluminum plates and 5 cm air gaps between the layers (see Figure 5.12), where showers are losing their low energy periphery.

Figure 5.48, top row, shows the density distribution of FD_{12} for five e^+ energies (15, 30, 50, 100 and 150 GeV). One can see that in average FD_{12} is growing with energy. An even better separation between different energy samples can be obtained using two-dimensional distributions of FD_{12} and the total event energy (or the number of hits) in one layer, as shown in the middle (bottom) rows.

Figures B.17 and B.19 in Appendix B.5 show similar FD_{12} distributions for 50, 100, 150 GeV π^+ showers. In this case there is no clear separation between different energies, though FD_{12} decreases from the first layer to the third (order of the DIFs is 1, 2, 0) in average, as for e^+ data.

The fractal dimension can be used to separate e^+ and π^+ samples of the same energy, an example for 150 GeV beams is shown in Figure 5.49. Other energies can be found in Appendix B.5, Figures B.20 - B.22. This can be used to estimate the beam purity. For example, FD_{12} distributions for 150 GeV (runs 97, 99, 100, 102, 103) in the top of Figure 5.50 reveal the presence of both 150 GeV e^+ and π^+ particles in the beam. e^+ (π^+) component is visible as the narrow (broad) distribution on the right (left). The distributions of FD_{12} versus the total event energy or the number of hits in one layer for 150 GeV e^+/π^+ mixture (for the same runs) are shown in the middle and the bottom rows in Figure 5.50.

It will be important to repeat this preliminary analysis in the future when SiW ECAL technological prototype will be completed.

The fractal dimension can be used as an additional variable for the particle identification in PFA. This is already implemented in Arbor and Garlic.

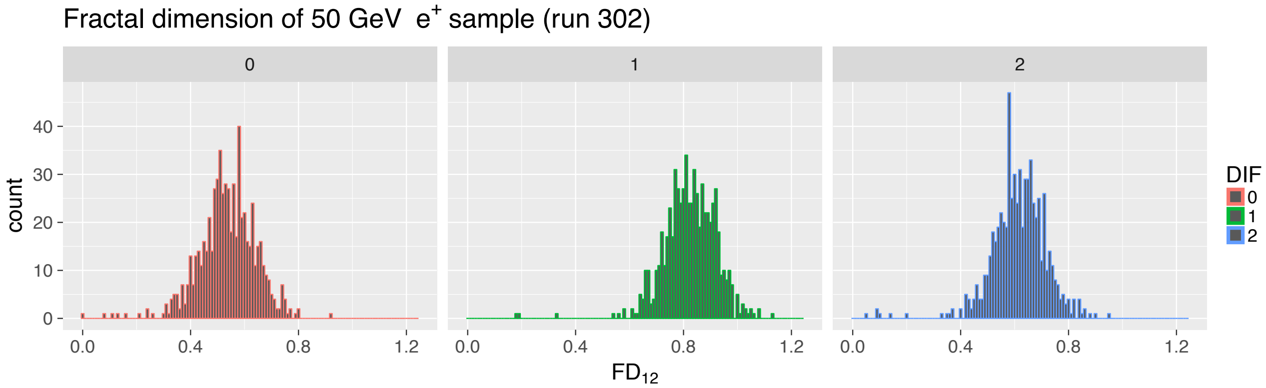


Figure 5.47 – FD_{12} distribution in three layers (DIF 0, 1, 2) for 50 GeV e^+ sample (run 302).

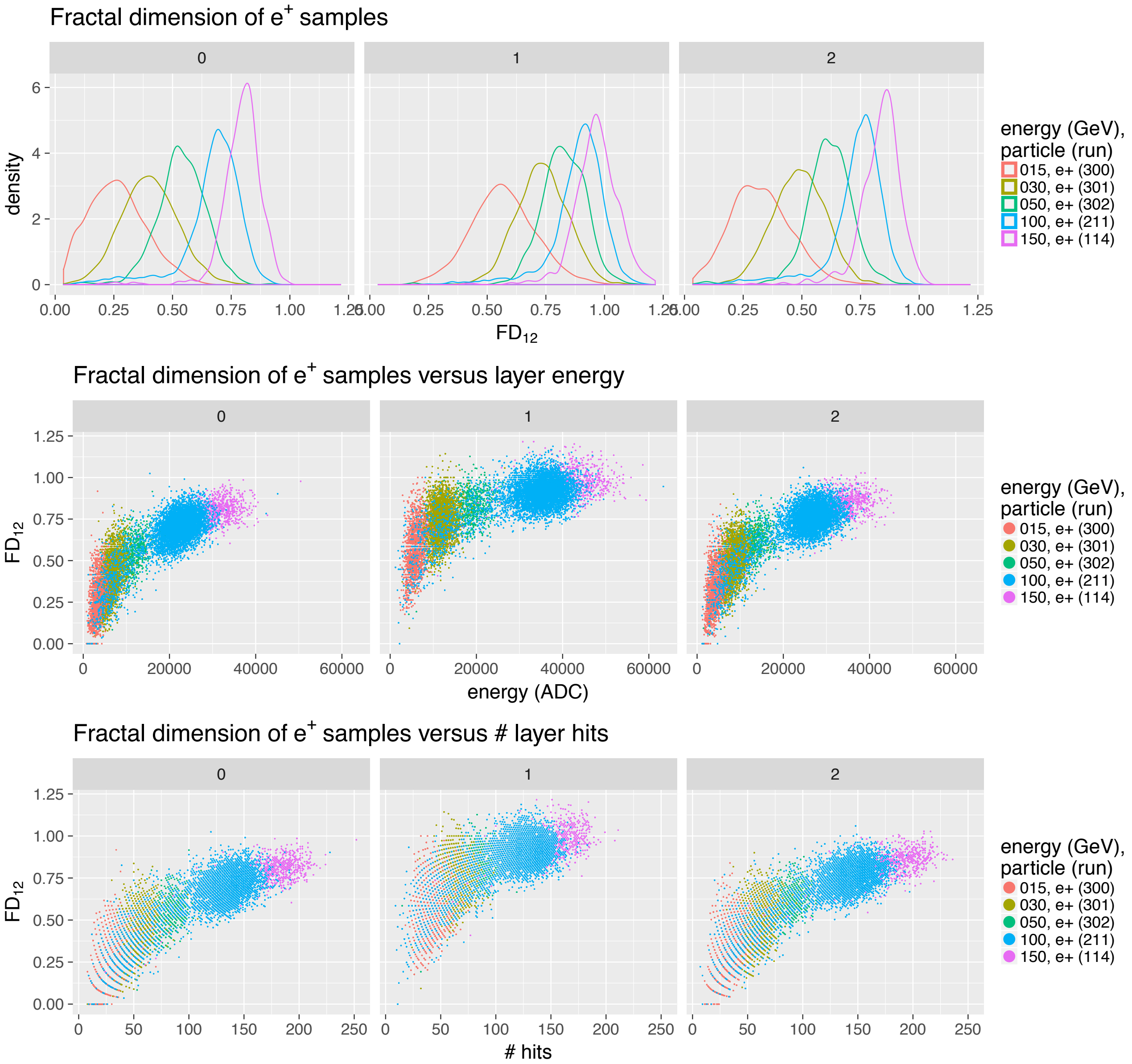
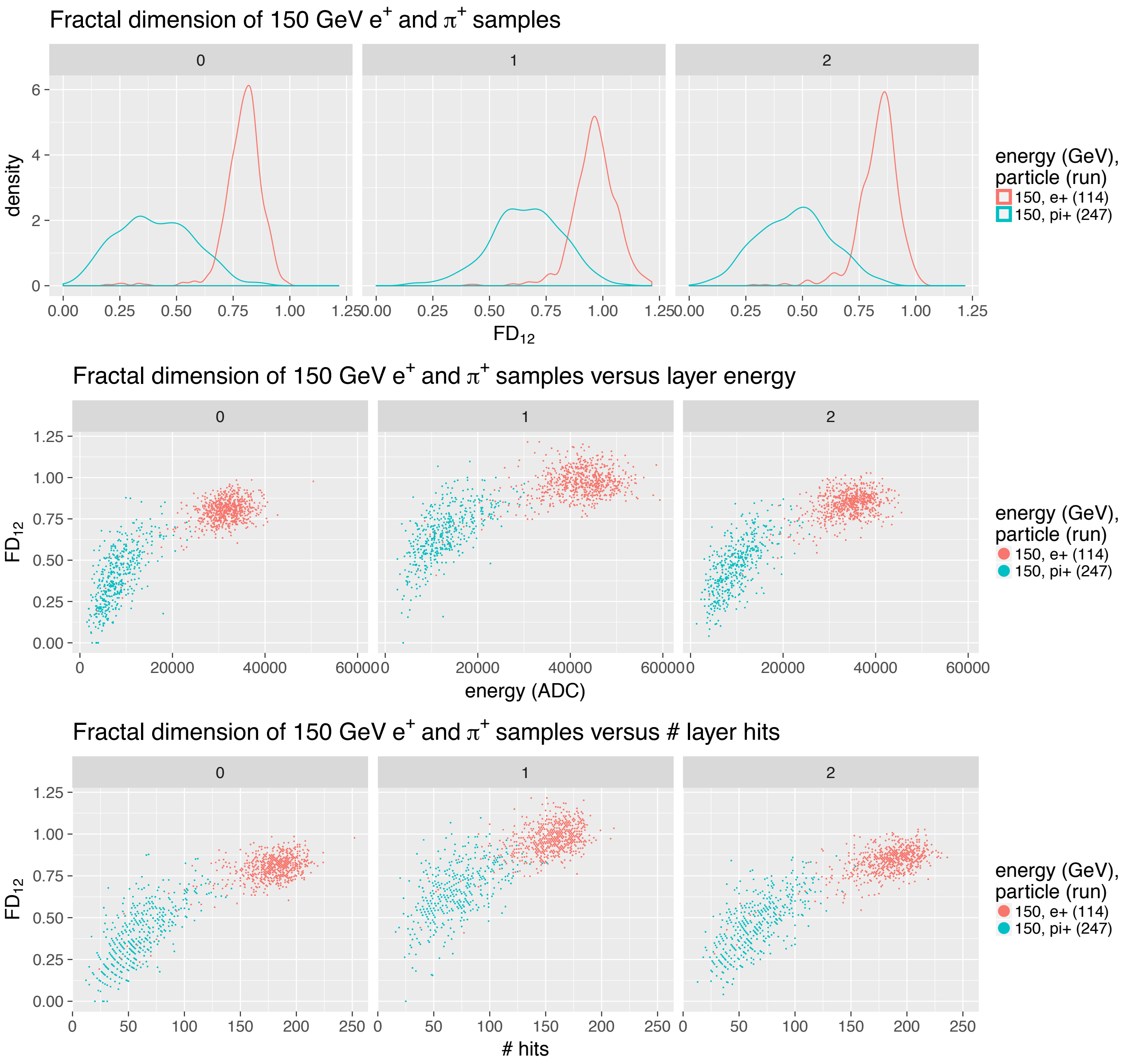


Figure 5.48 – FD_{12} density distributions for 15, 30, 50, 100, 150 GeV e^+ showers (top row, runs 300, 301, 302, 211, 114). FD_{12} versus the event energy or the number of hits (middle and bottom rows).



Fractal dimension of 150 GeV e^+ and π^+ samples

Fractal dimension of 150 GeV e^+ and π^+ samples versus layer energy

Fractal dimension of 150 GeV e^+ and π^+ samples versus # layer hits

Figure 5.49 – Same as in Figure 5.48, but for 150 GeV e^+ (run 114) and π^+ (run 247) showers.

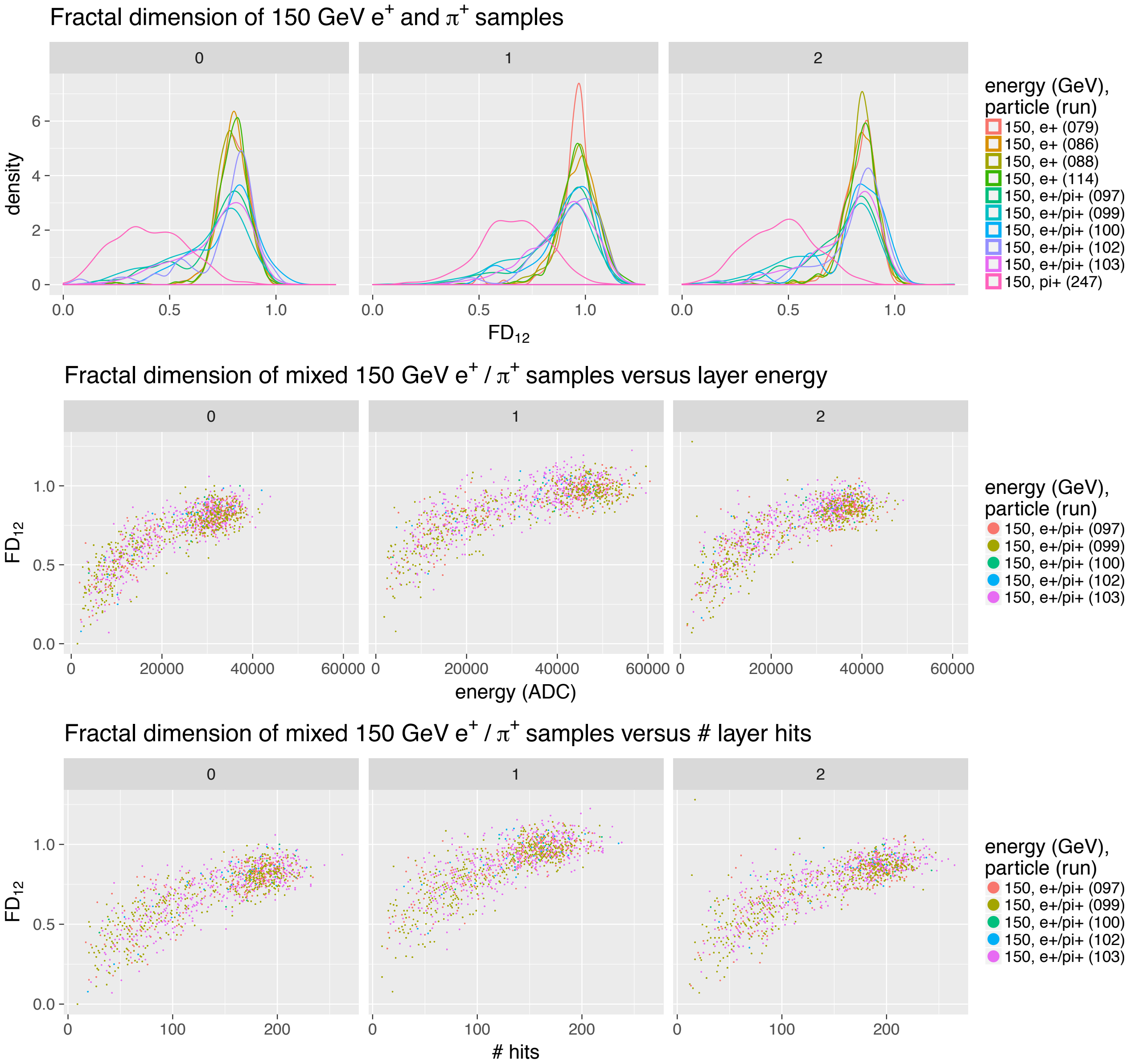


Figure 5.50 – Same as in Figure 5.48, but for 150 GeV e^+, π^+ samples or their mixtures. The distributions can be used to estimate the beam purity.

5.4 Conclusions

5.4.1 Conclusions on my SiW ECAL technological prototype activities

I have worked with both old FEV8 and current FEV10/11 versions of the SiW ECAL technological prototype.

With FEV8 prototype I have studied direct charge injection into the front-end chips of the slab. Data from three FEV8 slabs 5, 8, 10 in power pulsing and continuous current modes has been analyzed. Charge injection is done in each channel individually (of each chip and each slab). In all other channels preamplifier, injection and triggering were switched off. The amount of the injected charge is always known, so we can predict the channel response and perform calibration on channel level. 11 injection voltages of 0.01, 0.02, 0.05, 0.1, 0.15, 0.2, 0.3, 0.4, 0.6, 0.8 and 1.0 Volts cover linear working diapason, initial nonlinearity and the saturation region. It was observed that the pedestal values depend on SCA memory slot. It was found that variation across chips dominates the variation across channels, this can simplify calibrations in future. The difference between PP and CC modes is studied in a slab 10, the small difference between them is observed. Almost the same channels are found to be noisy and malfunctioning in each slab, their close to digital lines positions allow to conclude that these problems are caused by digital noise pickup and not by SKIROC2 structure.

I participated in three test beam campaigns in November'15 and June'16 at CERN SPS and in June'17 at DESY as a shifter and in the data analysis. With November'15 data I studied so-called "square"-event effects, caused by significant energy depositions into the sensor guard-ring. This effect is proportional to the size of the shower. For the 150 GeV e^+ beam shooting in the center of the slabs (corners of all sensors) after $8.4X_0$ of W absorber in front of the detector the observed "square"-event rate is smaller than 0.04%. The effect is significantly reduced in comparison with CALICE SiW ECAL physics prototype sensors. Also, I studied various data corruptions in the detector stream output. Some of them it is possible to correct offline later. Observed effect of a delayed trigger in BX and BX+1 in a single layer requires additional studies in the future. All features seen during November'15 TB were rediscovered during June'17 test beam.

Some physics analysis has been performed with e^+ and π^+ showers. Using November'15 CERN SPS data, 15 to 150 GeV e^+ and π^+ showers, the fractal dimension analysis has been performed. It was shown that for selected e^+ shower fractal dimension depends on the energy, it is growing with the energy of the e^+. The π^+ sample can be easily distinguished from the e^+ sample, potentially this can be used for beam purity calculations. Fractal dimension can be used as an additional variable for the particle identification.

The preliminary analysis of the longitudinal profile of 3 GeV e+ showers is performed using Jun'17 DESY data.

5.4.2 Technological prototype conclusions

Starting from 2011 there is ongoing activity on the SiW ECAL technological prototype. The carbon-epoxy alveolar mechanical prototype is produced and tested. It is scalable for entire ILD ECAL mechanical structure. Currently, all boards have been built in a semi-automatic mode, but technology should be fully automatized for the future ILC detector construction. Second generation FEV10/11 of the technological prototype boards are tested in the beam. The setups have shown stable behavior during November'15 and June'17 test beam campaigns. Achieved coincidence between the layers and signal-over-noise ratio $S/N \approx 19$ allows us to expect good results from this technology in future. Recently produced SKIROC2A ASIC chip will be mounted on next FEV12 version of the boards. Also, there is ongoing activity on long detector slab with several ASUs connected together. These long detector elements should be inserted into the alveolar mechanical structure. Additional efforts can be achieved in synergy with LHC collaborations, as recently they have shown their interest in CALICE silicon technologies.

Chapter 6

Precision measurement of the W boson mass in hadron decays at ILC

6.1 Introduction

W and Z boson masses are fundamental parameters of the Standard Model. Z boson mass is measured with high precision at LEP-I in e^+e^- collisions at Z-pole energy $\sqrt{s} = 91$ GeV. Current world average value [186] is:

$$M_Z^{world} = 91.1876 \pm 0.0021 \text{ GeV}. \tag{6.1}$$

Measured W boson mass has significantly lower precision. At LEP-II W mass is measured in $e^+e^- \rightarrow$WW-pair production process, which starts at $\sqrt{s} = 160$ GeV, with precision $M_W^{LEP} = 80.376 \pm 0.033$ GeV [187]. Two times better precision is achieved at Tevatron for W produced in pp collisions $M_W^{Tevatron} = 80.387 \pm 0.016$ GeV [188, 189]. The most recent measurement of W mass, published by ATLAS LHC collaboration in 2017, should be also mentioned here: $M_W^{ATLAS} = 80.370 \pm 0.019$ GeV [190]. Last result is not yet included in world average fit [186] which is dominated by Tevatron measurement:

$$M_W^{world} = 80.385 \pm 0.015 \text{ GeV}. \tag{6.2}$$

Achieved W mass precision is still lower than 7.8 MeV predicted by SM fit [191]. This fit prediction is based on contributions from theoretical (5.0 MeV) and experimental (6.0 MeV) uncertainties. The last ones are mainly driven by current uncertainties in $sin^2\theta^l_{eff} = 0.23152 \pm 0.00014$ and $m_t = 173.21 \pm 0.51 \pm 0.71$ GeV. More precise measurement of these experimental parameters and including deeper level Feynman diagrams will improve SM fit precision in future. It means that improving of M_W and $sin^2\theta^l_{eff}$ is very important for SM validation. Any deviations from SM predictions will give a possible indication of new physics.

It is expected that future ILC and application of PFA will improve the precision of the W and Z mass measurements [192]. Major W and Z production processes in e^+e^- collisions and their cross sections for unpolarized beams are shown in Figure 6.1. According to the studies performed for ILC TDR, it is expected that Z boson mass can be measured with a precision uncertainty of 1.6 MeV during high luminosity run at Z-pole $\sqrt{s} = 91$ GeV. W boson mass can be measured in WW-pair process using approaches similar to LEP [187]. First set of studies shows that WW-pair threshold scan with polarized beams around $\sqrt{s} = 161$ GeV can provide W mass precision uncertainty of ~ 5.2 MeV for 100 fb^{-1} of data [197]. Other study is done for $e\nu_e$W final states using $\sqrt{s}$ = 250 GeV and 500 GeV runs with polarized beams, at 250 GeV main contribution is coming from semileptonic decays of WW-pair (one W is decaying in two leptons and another in hadrons), while at 500 GeV main process is direct single-W production. W mass is reconstructed in hadronic decay channel. Expected statistical precision of W mass is 7.0 MeV (19.0 MeV) for $\sqrt{s}$ = 250 GeV (500 GeV) runs with 100 $\text{fb}^{-1}$of data [193].

According to the most recent ILC scenarios, the ILC will start its operation at $\sqrt{s}$ = 250 GeV without dedicated running periods at 91 GeV and 160 GeV. This will increase the relative importance of the WW channel as it will be possible to collect these events in parallel to ZH. This chapter presents a preliminary analysis of the precise W mass measurement in hadronic channels using 250 GeV full event simulation in ILD detector reconstructed with Pandora. According to the obtained results, after the initial period of ILC running, the error will be dominated by systematics.

6.2 W and Z production processes at ILC

ILC operation at $\sqrt{s}$ = 250 GeV means that processes with direct W and Z production will not be dominant during ILC operation. Major W and Z production processes and their cross-sections are shown in Figure 6.1 as a function of collision energy. At $\sqrt{s}$ = 250 GeV main source of W bosons is WW-pair production process, corresponding Feynman diagrams are shown in Figure 6.2. Branching ratios of the main W decay modes are listed in Table 6.1.

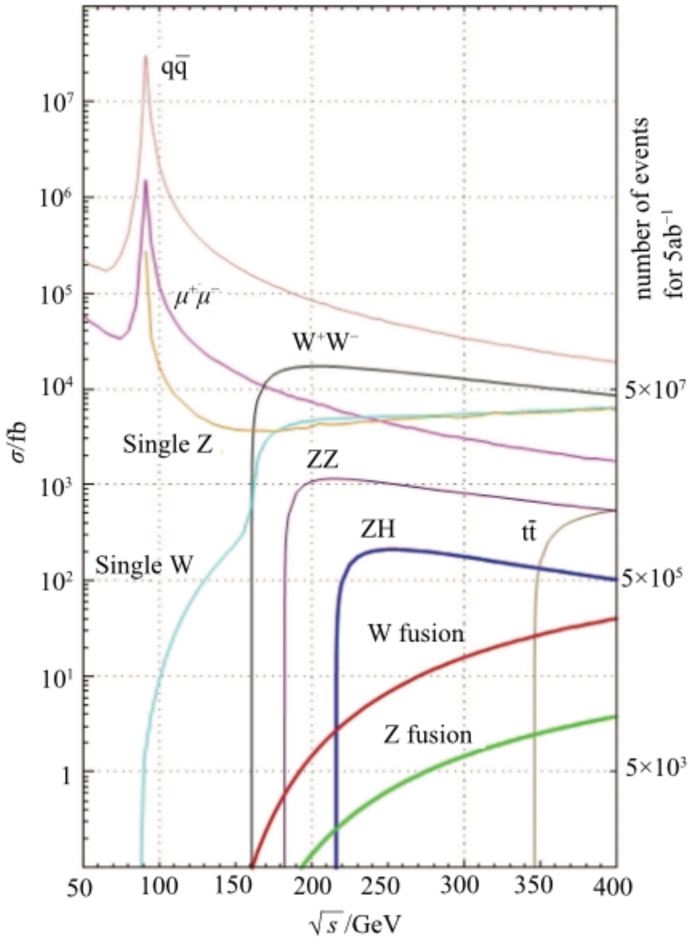


Figure 6.1 – Cross sections of the major e^+e^- SM processes with initial state radiation (ISR) effect taken into account. Figure is taken from [194], cross sections are shown for unpolarized beams.

To perform W mass analysis we are considering $e^+e^- \to W^+W^-$ pair production channel where one W is decaying into charged lepton and neutrino $l\nu_l$ ($l = e, \mu$ and $\nu_l = \nu_e, \nu_\mu$) and other W is decaying hadronically in jets. Precise identification of these two jets allows computing W mass with high precision. According to Table 6.1, total amount of events in chosen channels corresponds to ∼ 28.8% of all WW-pairs produced at ILC. In case of $l = e$ WW-pair process has same final state with single-W production process $e^+e^- \to e\nu_e W$, see Figure 6.4. This process should be added to the analysis, however, at $\sqrt{s}$ = 250 GeV the cross-section of this

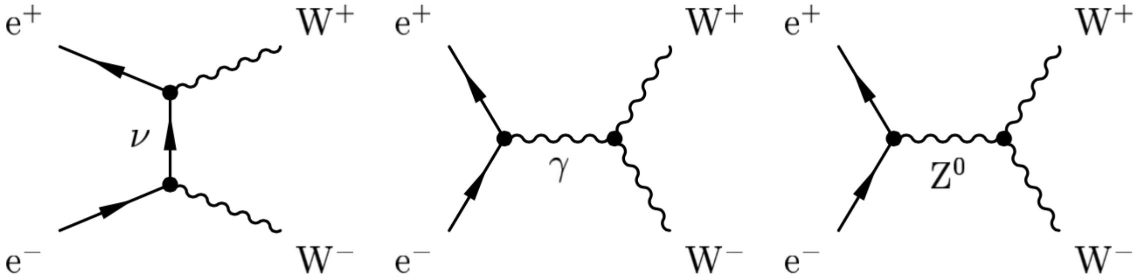


Figure 6.2 – Typical Feynman diagrams for the WW-pair process.

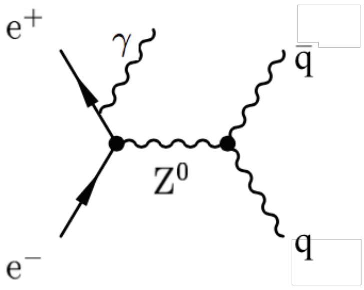


Figure 6.3 – Feynman diagram for γZ radiative return process.

W decay channel	Branching ratio, %
W $\rightarrow e\nu_e$	10.75 ± 0.13 %
W $\rightarrow \mu\nu_\mu$	10.57 ± 0.15 %
W $\rightarrow \tau\nu_\tau$	11.25 ± 0.20 %
W $\rightarrow$ hadrons	67.60 ± 0.27 %

Table 6.1 – The branching ratios of the main W decay channels according to PDG [186]. Branching ratio of W decay to hadrons is calculated assuming lepton universality.

process is quite low, as it is shown in Figure 6.5.

We are not considering backgrounds in the analysis. Backgrounds measured at LEP were at the level $< 15\%$ with respect to the signals ($e^+e^- \rightarrow q\bar{q}(\gamma, g)$, ZZ). According to the recent ILC study [193], the influence of the backgrounds on the W mass statistical error is almost negligible.

ILC, contrary to LEP, has higher uncertainties in the collision energy due to the beamstrahlung. However, it will be shown that beam constraint derived from ILC TDR parameters can still be useful for the analysis. W$\rightarrow \tau\nu_\tau$ channel has same statistics with $e\nu_e$ and $\mu\nu_\mu$ channels, but τ has short lifetime and many decay modes with at least two neutrinos. This results in higher backgrounds and loss of $\tau\nu_\tau$ channel efficiency. For simplicity, this channel is not considered for the current analysis.

The WW$\rightarrow l\nu_l l\nu_l$ decay is not used as it is limited statistically, eg. there is OPAL result on this channel, but with the precision of 0.41 GeV [187]. The $WW \rightarrow q\bar{q}q\bar{q}$ channel is also not considered as it has the largest systematics because of two effects caused by the strong interactions between hadrons from different W ("color reconnection") and Bose-Einstein coupling between identical hadrons from different W. Both effects modify the measured W mass and introduce the systematic bias which is difficult to estimate. At LEP the final systematics from leptonic-hadronic and purely hadronic final states was 36 MeV and 59 MeV, respectively [187], which gave 34 MeV from both methods.

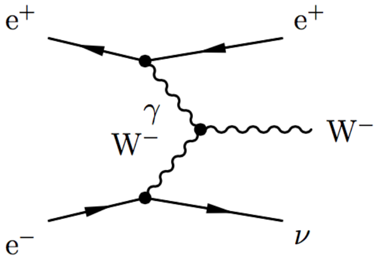


Figure 6.4 – Feynman diagram for the single-W process.

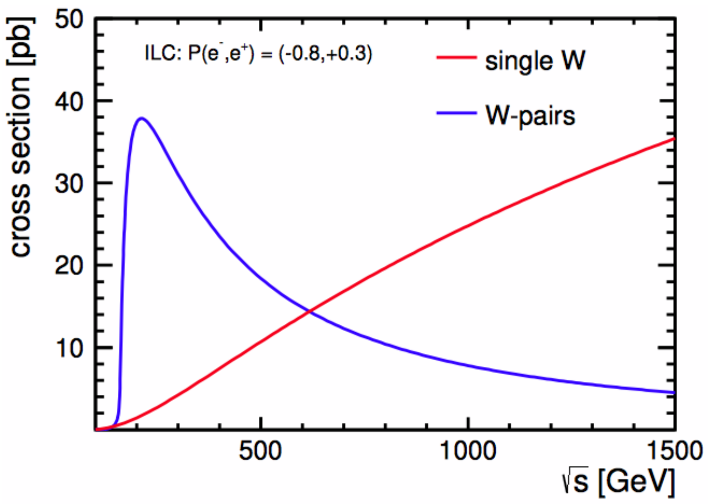


Figure 6.5 – Cross section for the single-W and WW-pair production processes e^-e^+ collisions with polarized $P(e^-, e^+) = (-80\%, +30\%)$ beams as a function of collision energy $\sqrt{s}$ [193].

Planned ILC will use polarized e^+e^- beams to increase the cross-sections and statistics. Figure 6.5 shows cross section of the WW and single-W channels in case of ILC polarization $P(e^-, e^+) = (-80\%, +30\%)$.

The energy scale of W mass can be calibrated with the mass of Z, produced via radiative return process $e^+e^- \to \gamma$Z, see Figure 6.3. Branching ratios of the Z decays are shown in Figure 6.2.

Z decay channel	Branching ratio, %
Z $\to e\nu_e$	3.363 ± 0.004 %
Z $\to \mu\nu_\mu$	3.366 ± 0.007 %
Z $\to \tau\nu_\tau$	3.370 ± 0.008 %
Z $\to$ invisible	20.00 ± 0.06 %
Z $\to$ hadrons	69.91 ± 0.06 %

Table 6.2 – The branching ratios of the main Z decay modes according to PDG [186].

6.3 Event processing and used samples

This section gives brief description of the event generation, simulation, digitization and reconstruction procedure in ILD detector and describes two samples used for the analysis.

6.3.1 Standard ILD procedure for MC-DBD data production

Whizard 1.95 [195] event generator is used to produce required final states of the WW and single-W processes (see section 6.3.2 for the details): $e\nu_e q\bar{q}$, $\mu\nu_\mu q\bar{q}$ and store them in StdHep format. Exact values of W mass M_W and width Γ_W used by Whizard are shown in Table 6.3. Possible beamstrahlung and initial state radiation (ISR) processes are taken into account during production. Events are grouped in four different samples corresponding to four possible combinations of pure e^+e^- polarizations $P(e^-, e^+) = (L, R)$, (R, L), (L, L) and (R, R), where L (R) corresponds to 100% left (right) handed polarization of the beam. Besides, Whizard calculates cross-section of the processes for each produced sample. This allows preparing a sample with any possible polarization by the proportional merging of samples according to required polarization and cross-sections of purely polarized samples. Quark hadronization processes are simulated using Pythia [196]. Particle composition in jets is tuned according to the data from OPAL experiment at LEP.

The procedure of simulation and reconstruction is standard for ILC and very similar to one described earlier in Chapter 4. Main steps of the used procedure are given below. StdHep generator files produced by Whizard are

generator data / particle	W	Z
mass M, GeV	80.418(9)	91.1881(9)
decay width Γ, GeV	2.048(9)	2.443(0)

Table 6.3 – W and Z masses $M_{W,Z}$ and decay widths used by Whizard for sample generation [198].

used as input for Mokka Geant4 (geant4-09-05-patch-01) with QGSP_BERT physics list to perform simulations in the standard ILD_o1_v05 Mokka geometry. This geometry corresponds to standard ILD detector with SiW ECAL with 5×5 mm^2 readout cells and AHCAL with 3×3 cm^2 cells. Simulated events are treated by a set of Marlin processors: BgOverlay processor to emulate backgrounds coming from the beam, digitization processors to emulate signal processing and get digitized output from each sub-detector and tracking processors to reconstruct particle trajectories in the tracker to use them as input for PFA reconstruction processors.

Comparison of the PFA algorithms is an important part of this thesis. Currently, full event reconstruction in ILD_o1_v05 geometry is possible only with Pandora PFA. Pandora returns a list of so-called “Particle Flow Objects" (PFOs) with information about reconstructed particles, their energy E^{reco}, momentum $\vec{p}^{reco}$, vertex point, etc. Two versions of the Pandora, v00-12 & v02-05, similar to the versions used for particle separation analysis in Chapter 4, are used for the event reconstruction. These versions are implemented in the standard ILCsoft v01-16-p10 and v01-17-11-p02, correspondingly. Beyond Pandora, later version of ILCsoft has more realistic digitization and some improvements in tracking.

6.3.2 Selected samples and event processing details

For the analysis two official 250 GeV ILD mc-dbd event samples produced for ILC TDR report are used:

- `4f_WW_semileptonic` sample with $e^+e^- \to$WW pair production process. One W is decayed in two leptons ($\mu\nu_\mu$ or $\tau\nu_\tau$) and other W into quarks $q\bar{q}$. Events with W$\to \tau\nu_\tau$ are excluded from the analysis.

- `4f_singleW_semileptonic` sample with a mix of WW-pair and single-W processes with $e^+e^- \to e\nu_e$W final state where W is decayed into two quarks $q\bar{q}$. Sample has such name due to historical reasons, but according to Figure 6.5 cross section of the single-W process at 250 GeV is significantly lower than WW-pair, so this sample mostly contains WW-pair events.

Summary of the characteristics of the selected samples is presented in Table 6.4: polarization, cross section σ, total number of simulated events.

MC sample, final state	Polarization	σ, fb	# Events	$\int Ldt$, fb^{-1}
`4f_WW_semileptonic`	e^-_L, e^+_R	18780.98	1919149	102.1
$\mu\nu_\mu q\bar{q}$, $\tau\nu_\tau q\bar{q}$	e^-_R, e^+_L	172.73	172733	1000.0
`4f_singleW_semileptonic`	e^-_L, e^+_R	9999.52	1927011	192.7
$e\nu_e q\bar{q}$	e^-_R, e^+_R	119.29	119289	999.9
	e^-_L, e^+_L	119.68	119683	1000.0
	e^-_R, e^+_L	85.62	85616	999.9

Table 6.4 – Cross-section σ, number of events and integrated luminosity $\int Ldt$ of used 250 GeV MC samples according to Whizard.

Cross section σ of any process in partially polarized beams $P(e^-, e^+) = (P_{e^-}, P_{e^+})$ can be derived from fully polarized cross-sections as:

$$\begin{aligned}\sigma(P_{e^-}, P_{e^+}) = \frac{1}{4}((1 - P_{e^-})(1 + P_{e^+})\sigma_{LR} + (1 + P_{e^-})(1 - P_{e^+})\sigma_{RL} + \\ (1 - P_{e^-})(1 - P_{e^+})\sigma_{LL} + (1 + P_{e^-})(1 + P_{e^+})\sigma_{RR}).\end{aligned} \tag{6.3}$$

Cross-sections corresponding to ILC with $P(e^-, e^+) = (-80\%, +30\%)$ polarization are obtained using Formula 6.3 and Table 6.4 and summarized in Table 6.5. WW$\to \mu\nu_\mu q\bar{q}$ ($\to e\nu_e q\bar{q}$) process has cross section 5499.96 fb (5898.17 fb).

$P(e^-, e^+) = (-80\%, +30\%)$ polarized samples in case of $e\nu_e q\bar{q}$, $\mu\nu_\mu q\bar{q}$ final states will mostly contain events from e^-_L, e^+_R collisions (more than 99% of events), so only events from e^-_L, e^+_R samples are used for the following analysis.

MC sample	σ, fb	σ_{LR}, fb	σ_{RL}, fb	σ_{LL}, fb	σ_{RR}, fb
`4f_WW_semileptonic`	10999.92	18780.98	172.73	—	—
`4f_singleW_semileptonic`	5898.17	9999.52	85.62	119.68	119.29

Table 6.5 – Cross-sections obtained from the generator. The second column corresponds to ILC with polarized beams $P(e^-, e^+) = (-80\%, +30\%)$, columns 3–6 - to fully polarized beams.

6.4 W di-jet event selection

Mass of W bosons decaying hadronically can be obtained by reconstruction of di-jets produced during hadronization of two quarks. Current performance of the standard jet finder algorithms, like FastJet [68], designed for Pandora and ILD is not suitable for individual jet reconstruction in case of di-jets. Instead of precise jet reconstruction, one can drop from the event everything that does not belong to hadronic W decay products, using advantages of the selected W production processes and beam environment at ILC. For the studied WW (single-W) physical channel it means that isolated lepton coming from the leptonic decay of W (vertex before W) should be tagged and removed from the di-jet particle candidates. Other backgrounds, like beamstrahlung or ISR photons, should be excluded or maximally suppressed in the analysis.

6.4.1 Beamstrahlung

At circular colliders, particles can pass through interaction region several times before interaction happens. Thus, at LEP it was possible to use beams which were focused below the level of beamstrahlung. Planned ILC will accelerate particles only once without a possibility to reuse them. It means that beams should be more focused to increase the collision rate. This would result in beamstrahlung photons radiated from e^+ or e^- before their actual collision and in uncertainty in the e^+e^- collision energy.

Beamstrahlung photons are mainly distributed in the very forward regions of the detector, as it is shown on the Figure 6.6. Most of these photons will leave detector without interactions trough the beam-pipe, but part of them will be detected in LimuCal and LHCAL, see Figures 6.7. Neutral clusters, reconstructed from beamstrahlung hits by Pandora in forward calorimeters, are removed from analysis according to $|\cos\theta| > 0.99$ cut where θ denotes opening angle between particle direction and detector main axis z.

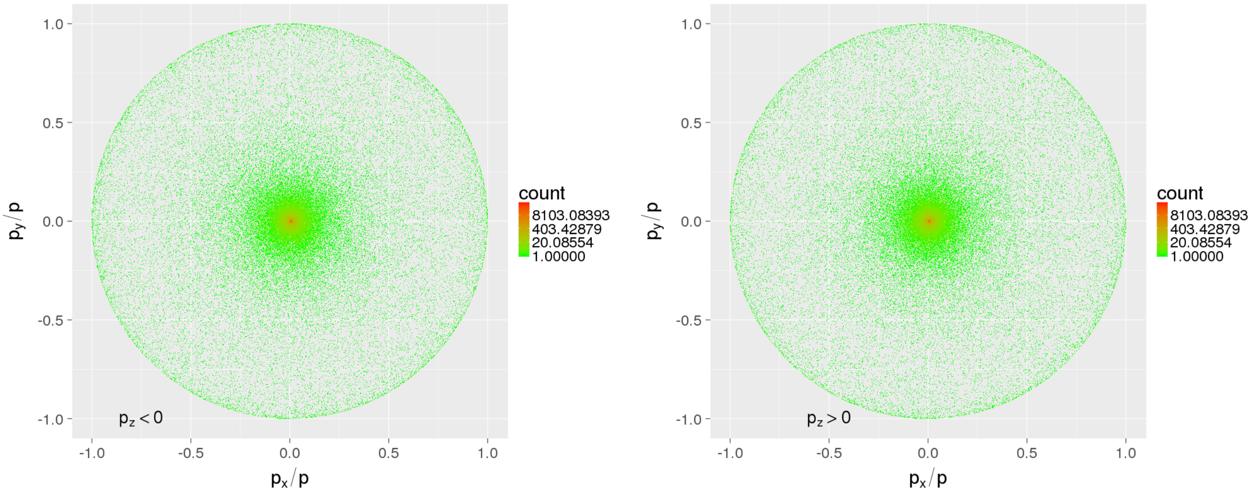


Figure 6.6 – Occupancy distributions of MC generated beamstralung and ISR photons in WW→ $\mu\nu q\bar{q}$ sample in $p_y/p \sim p_x/p$ coordinates ($\vec{p}$ is momentum of the photon) in left and right hemispheres of the detector. Detector main axis corresponds to (0,0) in this coordinates.

6.4.2 Isolated lepton tagging for $e\nu_e$W and $\mu\nu_\mu$W final states

High energy $e^\pm$ or $\mu^\pm$ leptons, coming from single-W ($e\nu_e$W) or leptonic decay of one W ($e\nu_e$ and $\mu\nu_\mu$) from WW-pair processes, should be identified and separated from jets coming from other W. Next tagging requirements are applied on the isolated lepton:

- $e^\pm/\mu^\pm$ lepton candidate should be reconstructed as a lepton, $e^\pm$ or $\mu^\pm$, depending on the decay channel.
- Its reconstructed energy E^{reco} should be above 15 GeV, energy distributions of MC-generated $\mu^\pm$ ($e^\pm$) coming from leptonic decay of W is shown on Figure 6.8 left (right) in green. Distributions of the reconstructed lepton energy in full events are shown in red.
- If two or more candidates are passing these requirements, the candidate with maximal energy is selected as isolated lepton and tagged.

After isolated lepton tagging, to prevent loss of the resolution and precision, it is essential to remove all other reconstructed neutrals associated with this lepton, such as:

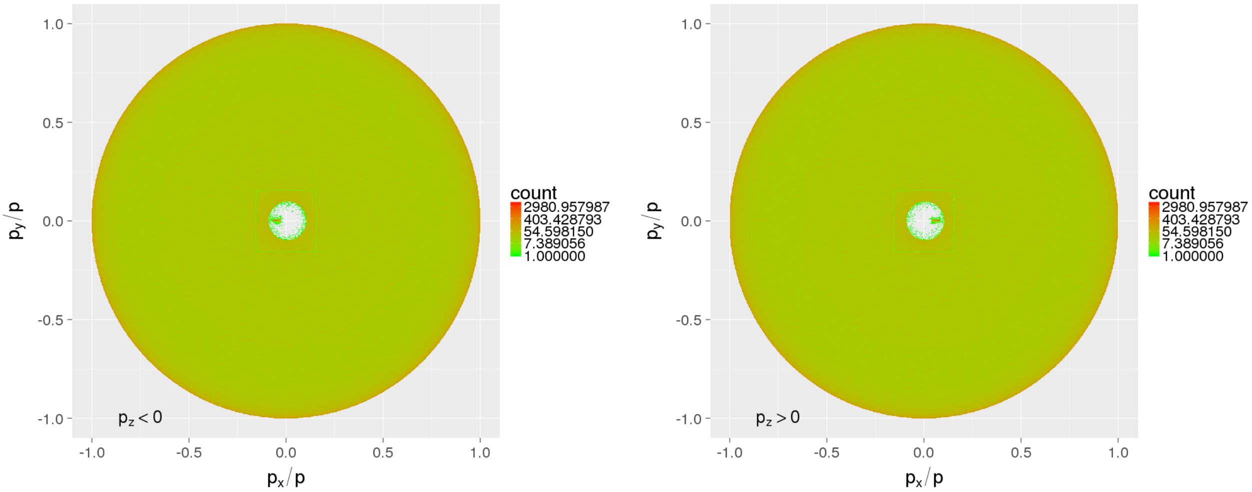


Figure 6.7 – Occupancy distributions of reconstructed photons in WW→ $\mu\nu q\bar{q}$ sample by Pandora v00-12 in $p_y/p \sim p_x/p$ coordinates in left and right hemisphere of the detector. Part of the beamstrahlung photons is registered in LumiCal and seen as a very narrow peak with high occupancy close to (0,0). These photons are removed from the analysis. Distributions for the other sample and Pandora v02-05 are similar.

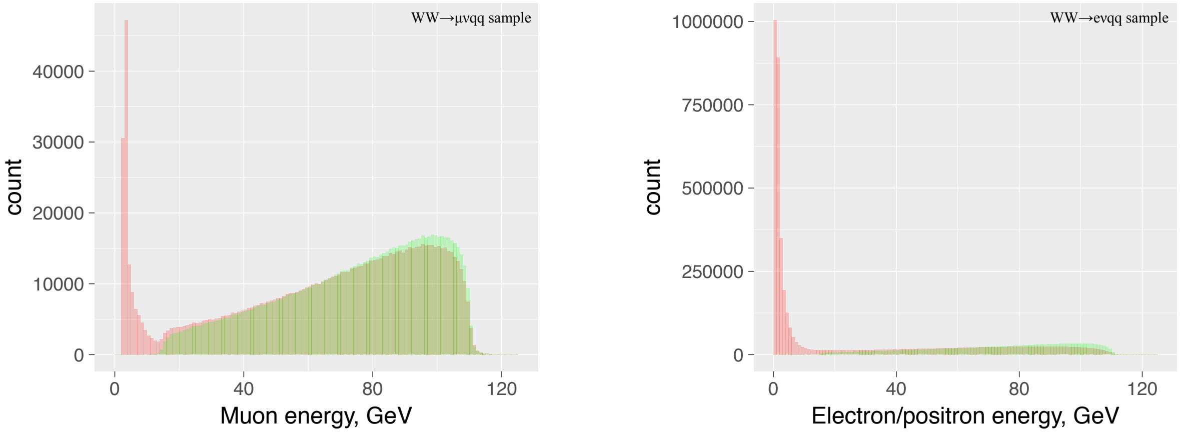


Figure 6.8 – Energy distribution of MC-generated leptons coming from W (in green). Distribution of all reconstructed leptons with Pandora v00-12 is shown in red. Left (right) plot corresponds to muons (electrons) from $\mu\nu_\mu q\bar{q}$ ($e\nu_e q\bar{q}$) sample. Cut $E^{reco} > 15$ GeV is applied on tagged leptons.

- bremsstrahlung photons, radiated between IP and calorimeter;
- confusion, when part of the tagged lepton is reconstructed as a neutral particle.

Distribution of the opening angles between tagged muons and its bremsstrahlung computed using MC-truth information is shown in Figure 6.9 (left). Confusion effect is estimated with reconstructed data because it cannot be derived at generator level. Distribution of the cosines of the opening angles $\theta^{open}_{(\mu,\gamma|n)}$ between tagged muons and each neutral reconstructed in the detector is shown in Figure 6.9 (right). Peak of the distribution close to $\cos\alpha^{open}_{\mu\ neurals} = 1$ corresponds to confusion and bremsstrahlung effects. Both effects should be excluded from the analysis, so all reconstructed neutrals within the cone $\cos\theta^{open}_{(\mu,\gamma|n)} > 0.98$ around tagged muons are removed from the events. The same cut is applied for tagged electrons.

All events with tagged leptons are selected for the W mass analysis, a total selection efficiency k_{eff} is shown in Table 6.4 for each version of Pandora.

MC sample Pandora version	k_{eff} v00-12	k_{eff} v02-05
WW→ $\mu\nu_\mu q\bar{q}$	0.967	0.965
WW→ $e\nu_e q\bar{q}$	0.915	0.911

Table 6.6 – Selection efficiency k_{eff} of the events.

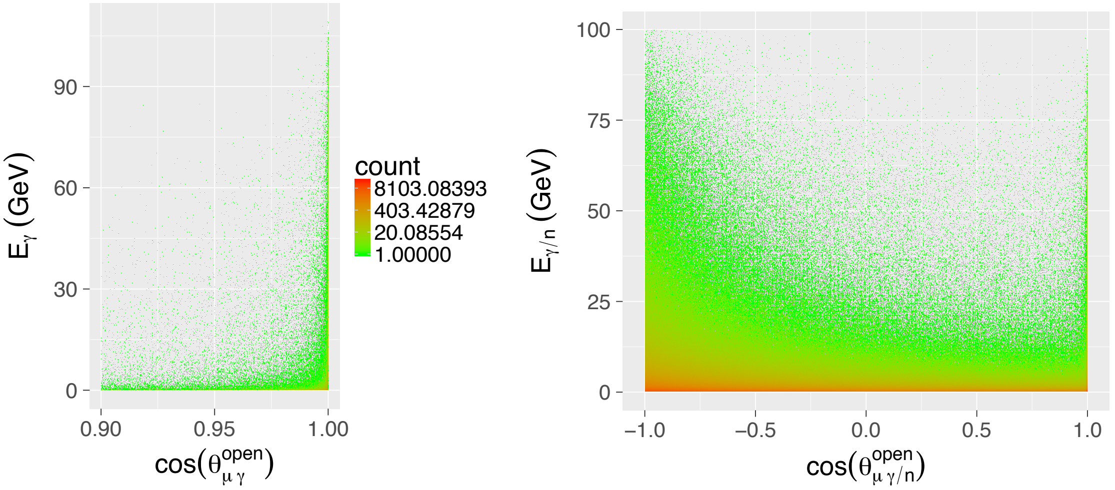


Figure 6.9 – Opening angule distribution (in $\cos\theta^{open}_{(\mu,\gamma)} > 0.9$ region) for bremstralung photons coming from μ (left plot, based on MC information for WW→ $\mu\nu q\bar{q}$). Right plot shows angular distribution between tagged muon and all reconstructed neutrals in event. Cut $\cos\theta^{open}_{(\mu,\gamma|n)} > 0.98$ at least partially removes bremstralung and confusion.

6.5 W mass reconstruction

6.5.1 Detector resolution

Invariant mass reconstruction in di-jets and Monte Calro di-quarks

After the event selection and cleaning procedure described in the previous section, we can assume that selected events have only reconstructed particles coming from W jets. Reconstructed W mass can be calculated as:

$$M_W^{reco} = \sqrt{\left(\sum_i E_i^{reco}\right)^2 - \left(\sum_i \vec{p}_i^{\,reco}\right)^2}, \tag{6.4}$$

where E_i^{reco} and $\vec{p}_i^{\,reco}$ are energy and momentum of the ith-particle in W di-jet event.

True Monte-Carlo mass of W, M_W^{MC}, is defined as:

$$M_W^{MC} = \sqrt{\left(E_1^{MC} + E_2^{MC}\right)^2 - \left(\vec{p}_1^{\,MC} + \vec{p}_2^{\,MC}\right)^2}, \tag{6.5}$$

where $E_{1,2}^{MC}$ ($\vec{p}_{1,2}^{\,MC}$) is true energy (momentum) of the two quarks coming from W.

Distributions corresponding to M_W^{reco} and M_W^{MC} are shown in Figure 6.10. One can see that reconstructed mass distributions M_W^{reco} are significantly broader than M_W^{MC}. Lineshapes of M_W^{MC} distributions can be modelled with relativistic Breit-Wigner shape $BW(m, M_W, \Gamma_W)$:

$$BW(m, M_W, \Gamma_W) = \frac{\Gamma_W^2 m^2/M_W^2}{\left(m^2 - M_W^2\right)^2 + \Gamma_W^2 m^4/M_W^2}, \tag{6.6}$$

where M_W (Γ_W) stands for mass (width) of W. M_W^{MC} lineshape does not depend on detector and PFA related effects, while lineshapes of the W reconstructed mass M_W^{reco} distributions are described by calorimetric PFA resolution $G(m)$ and relativistic Breit-Wigner $BW(m, M_W, \Gamma_W)$.

Improvement of calorimetric PFA resolution $G(m)$ using correlations between ΔM_W, reconstructed W energy E_W^{reco} and momentum p_W^{reco}

The difference between reconstructed M_W^{reco} and generated M_W^{MC} masses $\Delta M_W = M_W^{reco} - M_W^{MC}$ is shown in Figure 6.11, width of this resolution is determined only by PFA resolution $G(m)$.

It is generally believed that the powerful, at LEP, kinematic fit of the $e^+e^- \to WW \to l\nu_l q\bar{q}$ event which uses precisely known center-of-mass energy should not work at ILC. This is because of the beamstrahlung radiation which introduces too high uncertainty in the beam energies. In the course of this study, it was discovered, however, that this is not the case. Figure 6.12 shows the generated and reconstructed energies, E_W^{MC} and E_W^{reco} of the W boson hadronic final state. One can see that the generated distribution E_W^{MC} is relatively broad, but,

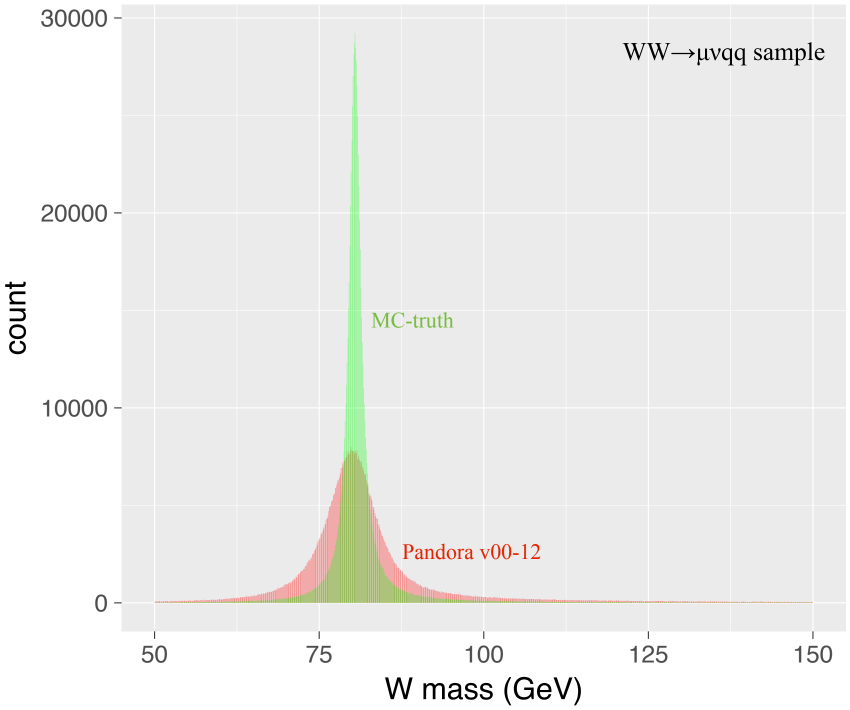


Figure 6.10 – Generated and reconstructed W mass distributions (M_W^{reco} distribution is shown in red, M_W^{MC} in green). Distribution for $e\nu_e q\bar{q}$ sample is similar.

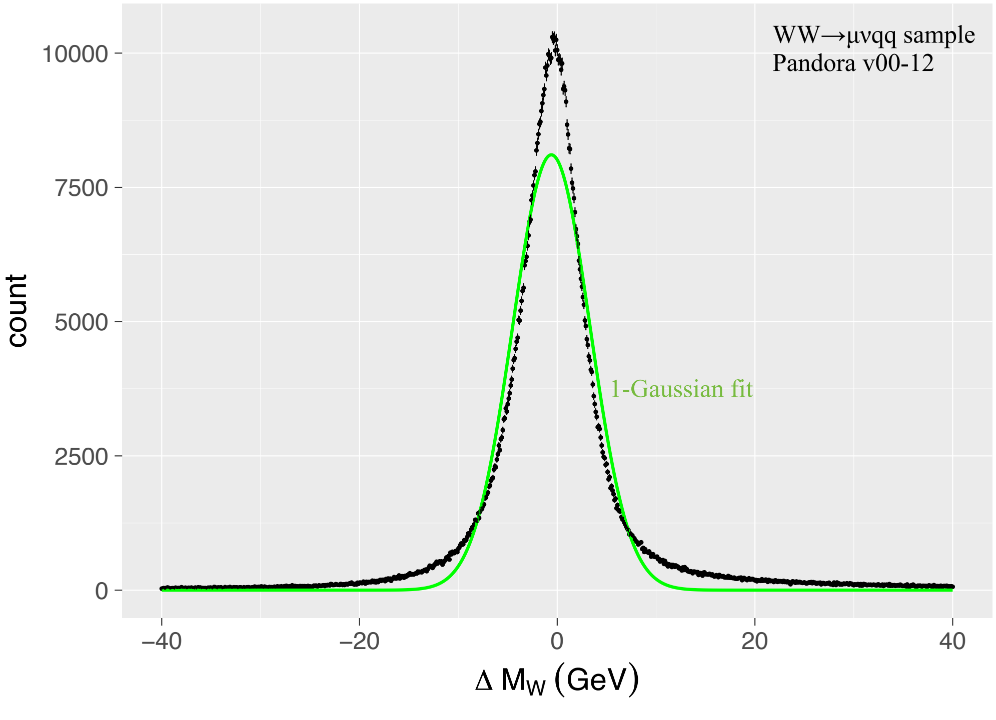


Figure 6.11 – Mismatch between generated and reconstructed masses of W decaying hadronically $\Delta M_W = M_W^{reco} - M_W^{MC}$ (Pandora v00-12, WW→ $\mu\nu_\mu q\bar{q}$). Distributions for other channels and Pandora v02-05 are similar. Green curve corresponds to the single Gaussian fit.

still, the reconstructed one E_W^{reco} is significantly wider. This means that by setting the reconstructed energy to exactly 125 GeV (we can not set it to the generated energy which should be ideal), one can potentially improve the accuracy of the jet momenta and W mass measurement.

The simplest way to check whether 125 GeV energy constraint can improve W mass resolution is to study the correlation between the (reconstructed - generated) mass mismatch ΔM_W and the energy of the reconstructed W boson E_W^{reco}, as it is shown in Figure 6.13. Indeed, the linear correlation is clearly visible and it allows to

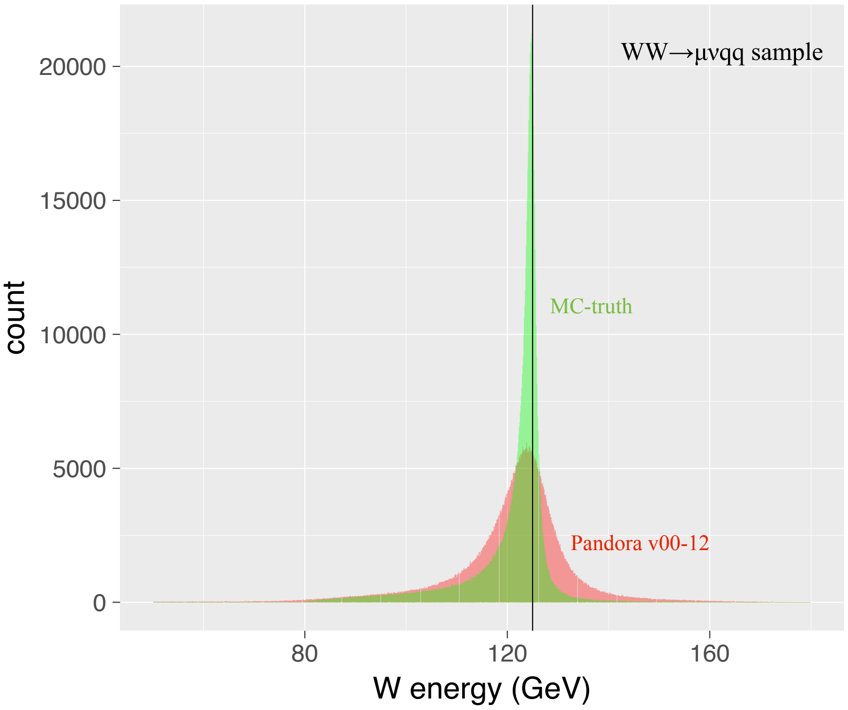


Figure 6.12 – Energy of the W boson decaying hadronically (E_W^{reco} distribution is shown in red, E_W^{MC} in green).

improve the mass resolution: after a little turn to make the line horizontal, its projection to the vertical axis should become narrower. In the same way the correlation with other variables have been checked, namely, with the reconstructed W momentum and with the energy and the momentum of the other $W \to l\nu_l$, which was reconstructed by assigning the missing momentum in the event to the neutrino from this decay. No sizable correlation was found with the variables of the other W. There is a slight correlation with the momentum of the W decaying to hadrons, however. It is not so well pronounced as the correlation with the energy, but still allows to improve the precision even further.

Both correlations can be understood in the following way. If PFA overestimates one of the jet energies, this increases both the reconstructed energy and, in average, the momentum and the mass. On the other hand, the energy and the momentum are not fully correlated (if W decayed at rest, its energy is the sum while the momentum is the difference between the two jet energies and momenta, respectively). Therefore, using the momentum after correcting for the correlation with the energy, still helps a little bit, as illustrated in Figure 6.14.

To ensure maximal effect from the correlations, the linear fit ΔM_W is applied on the E_W^{reco} and p_W^{reco} in parallel:

$$\mathrm{fit}(\Delta M_W) = \alpha_0 + \alpha_E E_W^{reco} + \alpha_p p_W^{reco}, \tag{6.7}$$

where α_0, α_E and α_p are the parameters of the linear fit. This procedure is done using R `lm` linear fit function from `stats` package [146]. ΔM_W^{corr} after correction caused by correlations is written as:

$$\Delta M_W^{corr} = \Delta M_W + 76.546668 - 1.316753 \cdot E_W^{reco} + 0.917709 \cdot p_W^{reco}. \tag{6.8}$$

These values are used for both available samples and versions of Pandora.

Fit of improved calorimetric PFA resolution

After correlation corrections, ΔM_W^{corr} distribution is narrower than ΔM_W and calorimetric PFA resolution $G^{corr}(m)$ can be obtained by fitting ΔM_W^{corr} distributions. Two independent fits are performed: with one and four Gaussian functions $C\exp\left(-(x-\mu)^2/2\sigma^2\right)$. The single Gaussian fit is useful, as it gives mass resolution and can be used for comparison and cross-check with a theoretical study on the minimal statistical error which can be achieved in the experiment [192]. Fit with four Gaussian functions gives a more precise profile of the resolution function and is used in the further analysis for the estimation of the W mass statistical precision in the next subsection:

$$G^{corr}(m) = \mathrm{fit}(\Delta M_W^{corr}(m)) = \sum_{i=1}^{4} C_i \exp\left(-\frac{(m-\mu_i)^2}{2\sigma_i^2}\right). \tag{6.9}$$

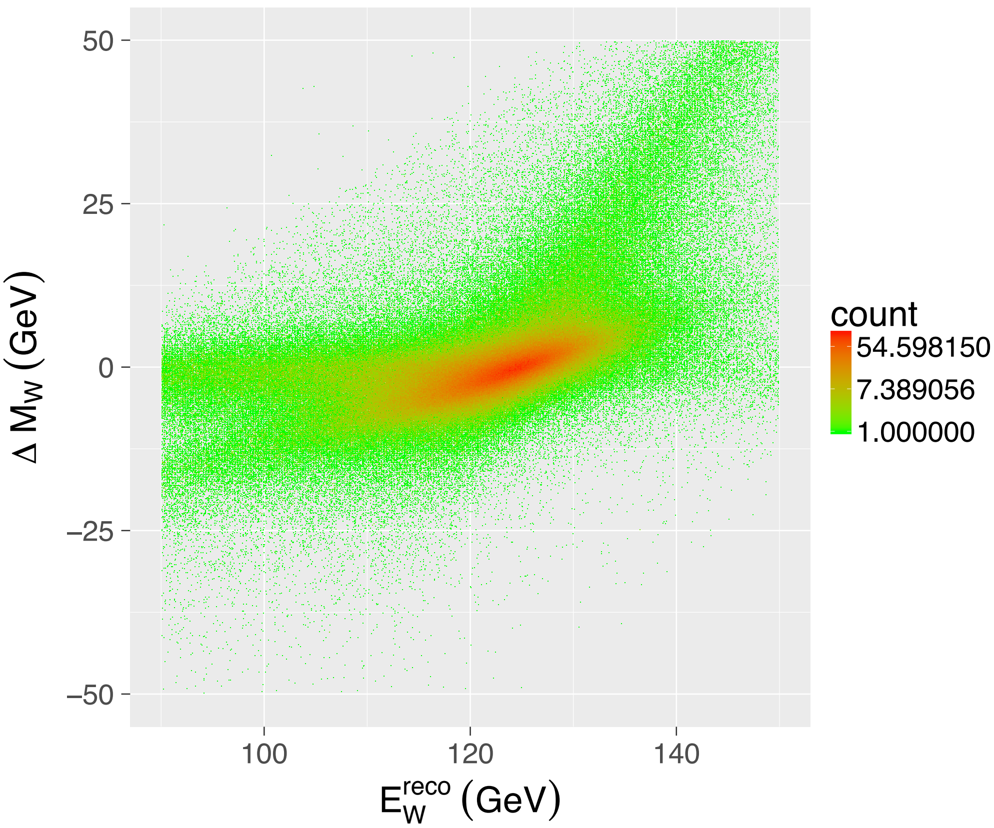


Figure 6.13 – Reconstructed energy of the W boson decaying hadronically versus the mismatch between reconstructed and generated masses ΔM_W (Pandora v00-12, WW→ $\mu\nu_\mu q\bar{q}$ channel).

Fit is processed using `nls` function from `stats` package for R [146]. Twelve parameters of four Gaussian fit (C_i, μ_i, σ_i, where i =1, 2, 3, 4) together with single Gaussian fit parameters (C, μ, σ) are shown in Tables 6.7, 6.8 for two used samples and two versions of Pandora.

An example of ΔM_W^{corr} distribution for WW→ $\mu\nu_\mu q\bar{q}$ sample reconstructed with Pandora v00-12 is shown in Figure 6.15. The blue curve shows fit to the normal distributions, sigma of the fit is ∼ 1.59 GeV. Without correlations it is 3.91 GeV, the corresponding fit is shown in Figure 6.11 (green curve). Red curve in Figure 6.15 shows four Gaussian fit.

	1-Gauss fit, Pandora v00-12			1-Gauss fit, Pandora v02-05		
MC sample	C	μ, GeV	σ, GeV	C	μ, GeV	σ, GeV
WW→ $\mu\nu_\mu q\bar{q}$	18956.3708	-0.1843	1.5936	18881.5994	-0.1986	1.5944
WW→ $e\nu_e q\bar{q}$	34478.2334	-0.1852	1.6379	34228.0166	-0.1976	1.6400

Table 6.7 – ΔM_W^{corr} single Gaussian fit parameters.

	4-Gauss fit, Pandora v00-12			4-Gauss fit, Pandora v02-05		
MC sample	C_i	μ_i, GeV	σ_i, GeV	C_i	μ_i, GeV	σ_i, GeV
WW→ $\mu\nu_\mu q\bar{q}$	18874.0934	-0.0900	0,9520	4799.9884	-0.5089	2.6037
	599.7652	-2.8525	8.3792	593.6529	-2.9127	8.4457
	49.2543	-20.6689	24.9917	47.5838	-21.2561	25.4445
	4869.3081	-0.4846	2.5845	18855.2401	-0.1012	0.9544
WW→ $e\nu_e q\bar{q}$	36436.9201	-0.0860	1.0305	37780.4335	-0.1131	1.0673
	323.1735	-8.1290	17.3473	496.8618	-6.9184	14.6436
	7300.1052	-0.6260	3.3172	5412.2048	-1.1515	3.8094
	576.7658	-9.6611	3.8290	1330.3710	2.5795	0.5789

Table 6.8 – ΔM_W^{corr} fit parameters with four Gaussians.

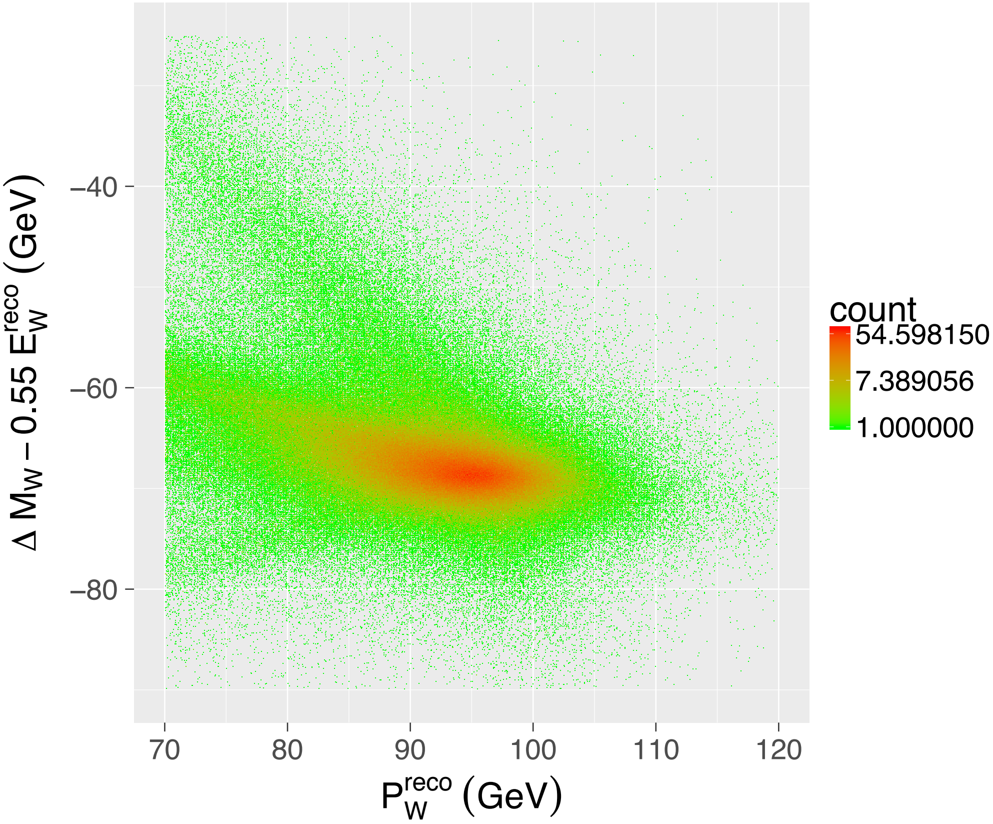


Figure 6.14 – Reconstructed momentum of the W boson decaying hadronically versus the mismatch between reconstructed and generated masses corrected for the correlation with the W energy (Pandora v00-12, WW→ $\mu\nu_\mu q\bar{q}$).

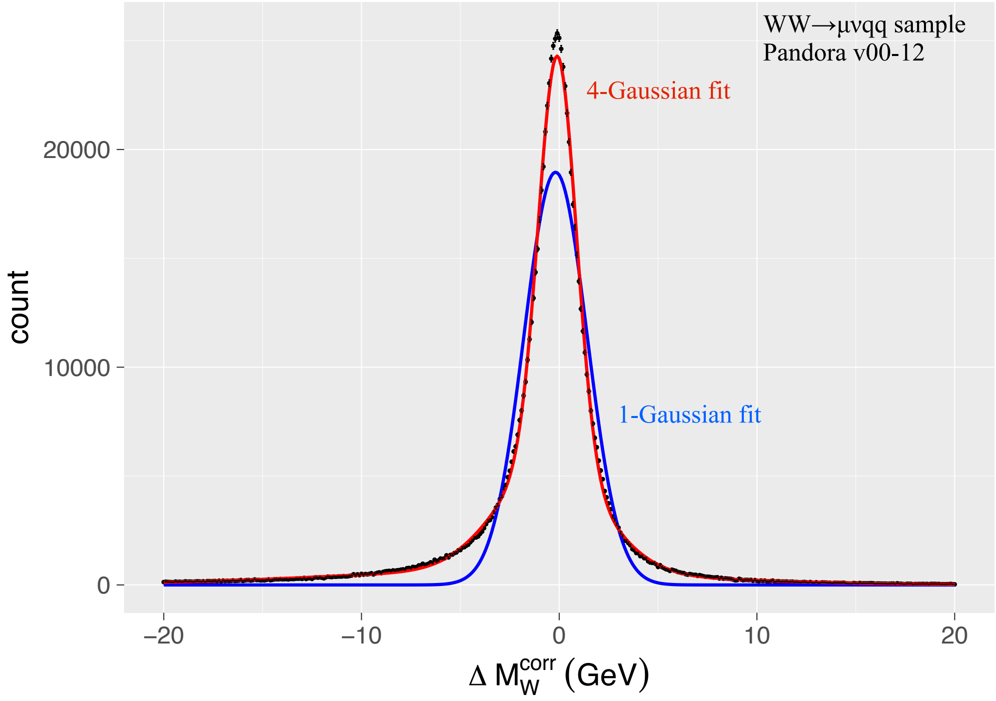


Figure 6.15 – PFA W mass resolution corrected for the correlations with the energy and the momentum. Red (blue) curve is the fit to the sum of four Gaussians (to one Gaussian).

6.5.2 Statistical precision of W mass measurement

Following the approach used in [187–189], W mass lineshape $GBW(m)$ can be obtained by convolution of relativistic Breit-Wigner distribution $BW(m, M_W, \Gamma_W)$, see Eq. 6.6, with parametrized PFA resolution $G^{corr}(m)$, see Eq. 6.9 and fit parameters in Tables 6.7 and 6.8:

$$GBW(m) = \int_{-\infty}^{+\infty} BW(t; M_W, \Gamma_W) G^{corr}(m-t)dt. \tag{6.10}$$

Direct analytical calculation of this integral is difficult, however, it can be computed numerically. To perform this calculation, continuous functions $BW(m)$ and $G^{corr}(m)$ are replaced by discrete arrays of numbers corresponding to m between -200 GeV and $+200$ GeV with step equal to 0.01 GeV. The entire procedure described in this section is also repeated with a step equal to 0.02 GeV to check convergence of the final results on W mass statistical precision. Convolution (Eq. 6.10) is calculated in two different ways independently to cross-check the results. Both methods, listed below, converge and the difference between them is significantly less than achieved precision:

- First method uses Fourier transform F and convolution theorem[1]. Integral definition of convolution, Eq. 6.10, can be rewritten as: $GBW(m) = F^{-1}\left(F(BW) \cdot F(G)\right)$. Here Fourier transfrom F is computed using function `fft` (Fast Discrete Fourier Transform) from R `stats` package [146].
- Second method is based on `convolute` function from `RcppFaddeeva` package [199]. It uses $BW(m)$ and $G^{corr}(m)$ arrays as inputs and its output is convolution $GBW(m)$.

Resulting $GBW(m)$ has only discrete values corresponding to m between -200 GeV and $+200$ GeV with step 0.01 GeV. Continuous density function $GBW^{cont}(m)$ of the W mass distribution, where $\int_{-\infty}^{+\infty} GBW^{cont}(m)dm = 1$, can be obtained by normalization and spline fit interpolation of $GBW(m)$ to all values of m between -200 GeV and $+200$ GeV. Normalization is realized as $GBW(m)/\sum_{\substack{m=-200\text{ GeV}\\ step=0.01\text{ GeV}}}^{m=+200\text{ GeV}} GBW(m)$. Spline fit with interpolation is done using `spline` and `approxfun` functions from `stats` package [146].

Each obtained $GBW^{cont}(m)$ density function is used for the generation of three W mass spectra with 10^5, 10^6 and 10^7 events with m distributed continuously between -200 GeV and $+200$ GeV. This corresponds to twelve studied W mass spectra, taking into account two studied WW channels (WW→ $\mu\nu_\mu q\bar{q}$ and $e\nu_e q\bar{q}$) and two versions of Pandora (v00-12 and v02-05). Generated W mass distributions are fitted using maximum likelihood method with continuous density functions $GBW^{cont}(m, M_W^{fit}, \Gamma_W^{fit})$ obtained as interpolation of the convolution of calorimetric PFA resolution $G^{corr}(m)$ with relativistic Breit-Wigner $BW(m, M_W^{fit}, \Gamma_W^{fit})$ where M_W^{fit} and Γ_W^{fit} are free parameters. Fit is realized with `fitdistr` function from `MASS` package [200], it uses generated W mass spectra and density functions $GBW^{cont}(m, M_W^{fit}, \Gamma_W^{fit})$ dependent on fit parameters M_W^{fit} and Γ_W^{fit} as input and for each spectrum it provides optimal M_W^{fit} and Γ_W^{fit} values together with their statistical precision. The obtained precisions for all studied spectra are given in Table 6.9. As one can see from this table, the difference between two versions of Pandora is negligible and both WW→ $\mu\nu_\mu q\bar{q}$ and $e\nu_e q\bar{q}$ channels give similar W mass accuracy.

The statistical precision should scale as $1/\sqrt{N}$, where N is number of events. This can be easily checked as we have 10^5, 10^6 and 10^7 samples for each WW decay channel and version of Pandora. Figure 6.16 shows achieved W mass statistical accuracy as a function of N events in the studied spectrum for each $GBW^{cont}(m)$. The differences in achieved statistical precision between two versions of Pandora and two studied WW decay channels are small. The combined fit of all data points gives $3.06 \cdot 10^3/\sqrt{N}$ (MeV) dependence of W mass accuracy as a function of N, see the yellow curve on Figure 6.16.

Correctness of the developed procedure for W mass statistical precision can be checked for the case of calorimetric resolution $G(m)$ represented by single Gaussian. Theoretical limit on the achievable experimental W mass statistical precision is studied in [192]. Updated results [201] of this study are presented in Figure 6.17 (left), where statistical precision of the W mass as a function of the calorimetric resolution width is shown. W mass spectra corresponding to 10^6 events are generated for single Gaussians with widths between 0 GeV and 5.0 GeV with 0.25 GeV step. Each spectrum is fitted with the so-called Voigtian fit, the convolution of single Gaussian with non-relativistic Breit-Wigner. The study is realized in C++ using RooFit Voigtian model [202].

My procedure, developed for calorimetric PFA resolution $G^{corr}(m)$ represented with four Gaussian fit of $\Delta M_W^{corr}(m)$, can reproduce results shown in Figure 6.17 (left) if $G^{corr}(m)$ is replaced with single Gaussian. W mass spectra with 10^6 events are generated for Gaussians with 0.5, 1.0,..., 4.5, 5.0 GeV widths using convolution of Gaussian with relativistic Breit-Wigner. Black dots on Figure 6.17 (right) show the statistical precision of W mass measurement as a function of the width of calorimetric resolution represented with single Gaussian. As one can see, Figures 6.17, left and right, are identical.

---

[1] Let f and g be two functions with convolution $f * g$. Let F denote Fourier transform operator, so $F(f)$ and $F(g)$ are Fourier transforms of f and g respectively. Then $F(f * g) = F(f) \cdot F(g)$ or using inverse Fourier transform F^{-1}: $f * g = F^{-1}\left(F(f) \cdot F(g)\right)$.

WW$\to \mu\nu_\mu q\bar{q}$, Pandora v00-12			
Stat. error/ N events	10^5	10^6	10^7
M_W, GeV	$9.547 \cdot 10^{-3}$	$3.026 \cdot 10^{-3}$	$9.556 \cdot 10^{-4}$
Γ, GeV	$18.506 \cdot 10^{-3}$	$5.880 \cdot 10^{-3}$	$18.570 \cdot 10^{-4}$

WW$\to \mu\nu_\mu q\bar{q}$, Pandora v02-05			
Stat. error/ N events	10^5	10^6	10^7
M_W, GeV	$9.572 \cdot 10^{-3}$	$3.028 \cdot 10^{-3}$	$9.575 \cdot 10^{-4}$
Γ, GeV	$18.655 \cdot 10^{-3}$	$5.888 \cdot 10^{-3}$	$18.617 \cdot 10^{-4}$

WW$\to e\nu_e q\bar{q}$, Pandora v00-12			
Stat. error/ N events	10^5	10^6	10^7
M_W, GeV	$9.777 \cdot 10^{-3}$	$3.107 \cdot 10^{-3}$	$9.785 \cdot 10^{-4}$
Γ, GeV	$19.030 \cdot 10^{-3}$	$6.036 \cdot 10^{-3}$	$19.012 \cdot 10^{-4}$

WW$\to e\nu_e q\bar{q}$, Pandora v02-05			
Stat. error/ N events	10^5	10^6	10^7
M_W, GeV	$9.768 \cdot 10^{-3}$	$3.100 \cdot 10^{-3}$	$9.817 \cdot 10^{-4}$
Γ, GeV	$18.997 \cdot 10^{-3}$	$6.030 \cdot 10^{-3}$	$19.156 \cdot 10^{-4}$

Table 6.9 – W mass and decay width error estimation for 10^5, 10^6 and 10^7 event samples for WW$\to e\nu_e q\bar{q}$ and WW$\to \mu\nu_\mu q\bar{q}$ channels, for Pandora v00-12 and v02-05. Calorimetric PFA resolution $G^{corr}(m)$ is obtained as a fit of $\Delta M_W^{corr}(m)$ with four Gaussians, see Table 6.8.

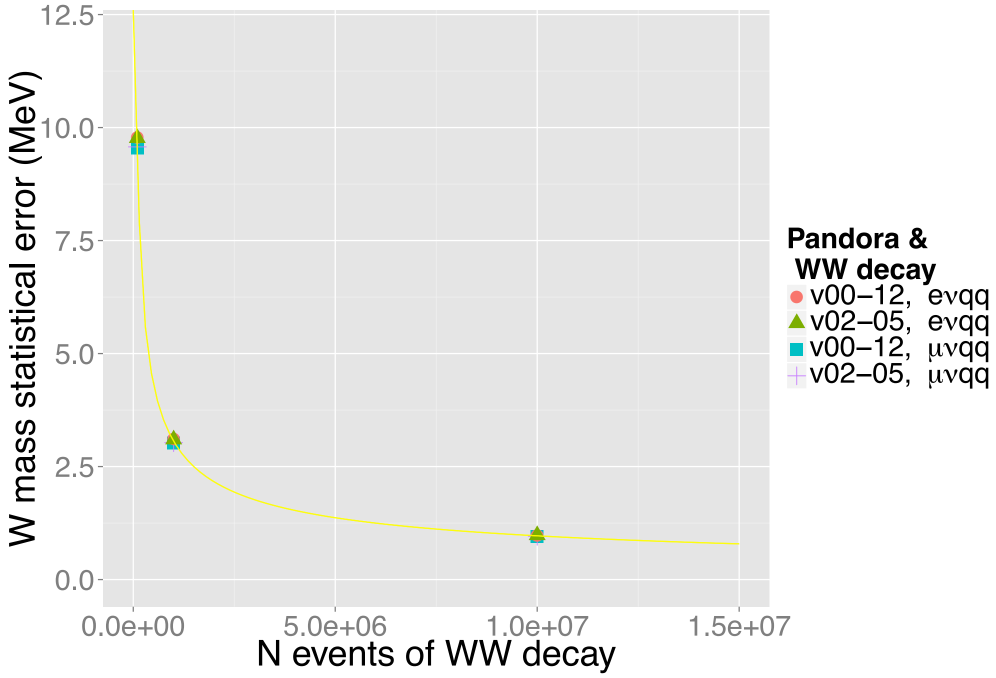


Figure 6.16 – Achievable W mass statistical error after 10^5, 10^6 and 10^7 WW$\to \mu\nu_\mu q\bar{q}$ and $e\nu_e q\bar{q}$ collected events reconstructed with Pandora v00-12 and v02-05. Yellow line, $3.06 \cdot 10^3/\sqrt{N}$ (MeV), shows fit of all obtained points.

Calorimetric PFA resolution obtained with single Gaussian fit provides a less realistic description of the reconstruction process, contrary to four Gaussian fit. This explains the fact that achievable statistical precision in studied WW channels is better for $G(m)$ represented as single Gaussian, results can be compared in Tables 6.9 and 6.10. One can see that, for example, for WW$\to \mu\nu_\mu q\bar{q}$ channel both versions of Pandora give ~ 2.68 MeV (~ 3.03 MeV) of W mass statistical error for 10^6 of W decays for single (four) Gaussian fit of calorimetric PFA resolution. For WW$\to e\nu_e q\bar{q}$ results are similar: ~ 2.72 MeV (~ 3.10 MeV) for 10^6 events for single (four) Gaussian fit. Colored dots on Figure 6.17 (right) show achieved precisions in studied channels with calorimetric resolution $G^{corr}(m)$ obtained as fit of ΔM_W^{corr}, see Figure 6.15 and Table 6.7, with single Gaussian with ~ 1.59 GeV (~ 1.64 GeV) resolution width for WW$\to \mu\nu_\mu q\bar{q}$ ($e\nu_e q\bar{q}$) channel. Results are in perfect agreement with theoretically predicted precisions, shown as black dots on Figure 6.17 (right), for calorimeter resolution represented as single Gaussians.

It was also checked that if one uses two Gaussians with $\sigma = 2.0$ GeV and 1.6 GeV instead of both relativistic

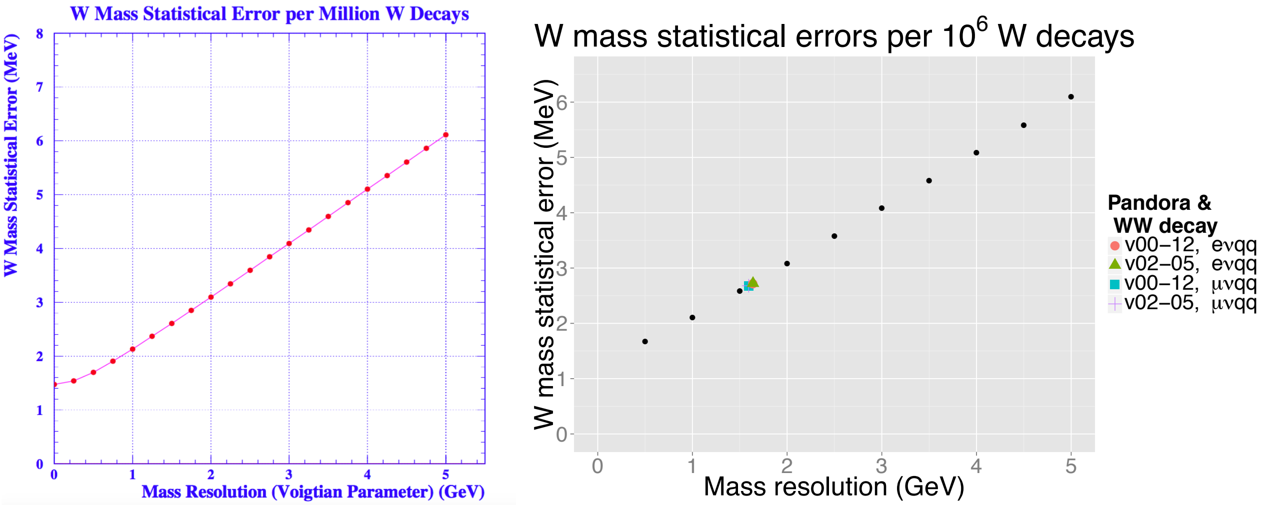


Figure 6.17 – Statistical precision of the W mass in 10^6 W decays as a function of calorimetric resolution represented as single Gaussian computed using fit of convolution of Gaussian with non-relativistic Breit-Wigner, so-called Voigtian fit (left plot, from [201]). The right plot is similar to the left, but Gaussian is convoluted with relativistic Breit-Wigner, both left and right plots show the same theoretical statistical precision of the W mass. Colored dots on the right plot represents W mass statistical precision achievable in 10^6 WW$\to e\nu_e q\bar{q}$ and WW$\to \mu\nu_\mu q\bar{q}$ events if calorimetric PFA resolution is obtained as a fit of $\Delta M_W^{corr}(m)$ with single Gaussian.

WW$\to \mu\nu_\mu q\bar{q}$, Pandora v00-12	
Stat. error/ N events	10^6
M_W, GeV	$2.678 \cdot 10^{-3}$
Γ, GeV	$4.543 \cdot 10^{-3}$

WW$\to \mu\nu_\mu q\bar{q}$, Pandora v02-05	
Stat. error/ N events	10^6
M_W, GeV	$2.679 \cdot 10^{-3}$
Γ, GeV	$4.544 \cdot 10^{-3}$

WW$\to e\nu_e q\bar{q}$, Pandora v00-12	
Stat. error/ N events	10^6
M_W, GeV	$2.719 \cdot 10^{-3}$
Γ, GeV	$4.587 \cdot 10^{-3}$

WW$\to e\nu_e q\bar{q}$, Pandora v02-05	
Stat. error/ N events	10^6
M_W, GeV	$2.724 \cdot 10^{-3}$
Γ, GeV	$4.598 \cdot 10^{-3}$

Table 6.10 – W mass and width statistical precision for 10^6 WW$\to e\nu_e q\bar{q}$ and WW$\to \mu\nu_\mu q\bar{q}$ events reconstructed with Pandora v00-12 and v02-05. Calorimetric PFA resolution $G^{corr}(m)$ is represented as single Gaussian (Table 6.7).

Breit-Wigner and calorimetric PFA resolution, the expected W mass error for 10^6 events, $\sqrt{2.0^2+1.6^2}\cdot 10^{-3} = 2.56125 \cdot 10^{-3}$ GeV, is reproduced: $2.561250056 \cdot 10^{-3}$ GeV ($2.558486607 \cdot 10^{-3}$ GeV) for one (two) parameter fit with M_W as free parameter and Γ_W fixed at 2.0 GeV (both M_W and Γ_W as free parameters).

The cross sections for $e^+e^- \to W^+W^- \to \mu\nu_\mu q\bar{q}$ and $\to e\nu_e q\bar{q}$ processes for various beam polarizations including the default ILC polarization are listed in Table 6.5. They are reported by the event generator. All branching ratios are included. Eg. for 2000 fb^{-1} in H20 scenario one can expect about 11 millions of $\mu\nu_\mu q\bar{q}$ events. Due to contribution from the single-W production, the cross sections of $e\nu_e q\bar{q}$ final states are slightly larger and ~ 11.8 millions of events are expected. The efficiency of the selection k_{eff} is mentioned in Table 6.6.

Taking it into account, the statistical precision for the combined sample with muon and electron decays of the other W, ie. $e^+e^- \to W^+W^- \to \mu\nu_\mu$+hadrons or $e\nu_e q$+hadrons, as a function of the ILC integrated luminosity $\left(\int Ldt\right)$ is shown in Figure 6.18. It is calculated as

$$3.06 \cdot 10^3 \Big/ \sqrt{(0.96 \cdot 5499.96 + 0.91 \cdot 5898.17)\left(\int Ldt\right)}, \text{ MeV.} \tag{6.11}$$

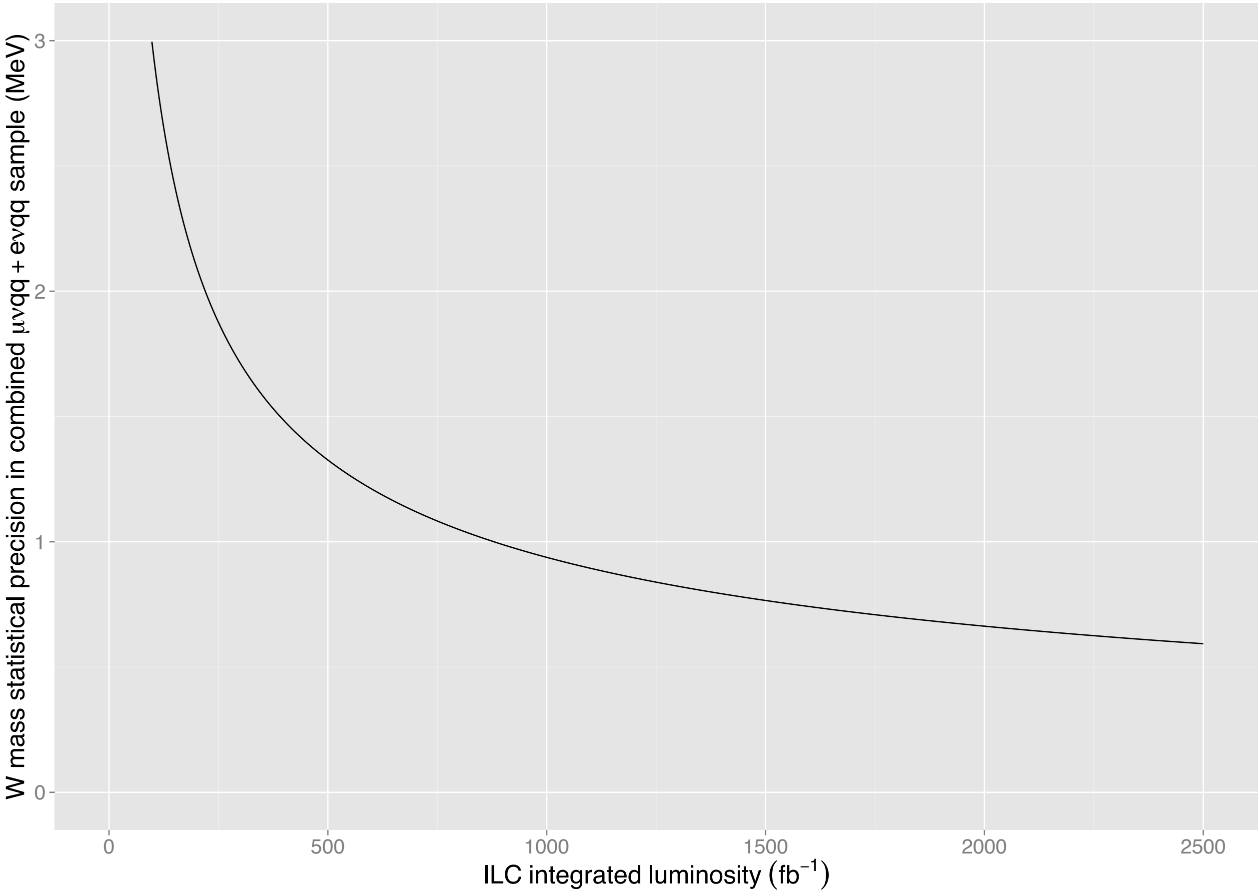


Figure 6.18 – Statistical precision of W boson mass measurement from $e^+e^- \to W^+W^- \to \mu\nu_\mu$+hadrons and $e\nu_e$+hadrons samples as a function of ILC integrated luminosity $\int Ldt$.

As one can see, the statistical precision will be at the current world average 15 MeV level already after the first 3.9 fb^{-1} accumulated at ILC.

Conclusions

International Linear Collider is the most advanced e^+e^- collider project. The alternatives to ILC are FCCee or CEPC. Currently, due to the cost optimization, the scenario with $\sqrt{s} = 250$ GeV center-of-mass energy (and possibly only one ILD detector) seems to be the most realistic. ILC physics program places the number of challenges for the physicists and engineers. ILC energies will be reachable with the superconducting RF cavities, the same technology is successfully used in the XFEL linear accelerator in DESY. This technology is scalable, in case of the ILC upgrade the increase of the energy of the beams can be achieved by an installation of additional SCRF cavities.

Regardless of the ILC energy, the proposed detector should provide the jet energy resolution of 3-4% for 45 - 250 GeV jet energies. This should give the ability to distinguish W and Z bosons, their mass reoslution will be comparable with their decay widths. This requirement can be realized only with the detector optimized for Particle Flow Algorithms. Reconstruction of each particle individually in the jet requires unprecedented transverse and longitudinal segmentation of the electromagnetic and hadronic calorimeters. Current studies show that $\sim 5 \times 5$ mm^2 ECAL granularity is sufficient to realize the ILC goals.

The highly granular silicon-tungsten technology is proposed for the ILD electromagnetic calorimeter. Silicon as an active material is more expensive than the alternative scintillator option. However, it offers many advantages, eg. better granularity, excellent linearity and uniformity across pixels, temperature independence, ease of calibration, time stability. Overall, the silicon technology will allow to achieve the minimal systematic errors.

Tungsten has small Moliere radius $R_M^{(W)} = 0.9327$ cm and a large ratio between the interaction and radiation lengths $\lambda_I^{(W)}/X_0^{(W)} \approx 28.38$. Small R_M ensures small transverse size of the electromagnetic shower thus facilitating the photon separation. Large λ_I/X_0 ratio ensures that large fraction of hadrons will pass through ECAL without showering facilitating hadron–photon separation. ILD SiW ECAL will have an unprecedented number of channels, $60 - 100$ million depending on the ILD size and the number of ECAL layers, and a silicon area up to ~ 2600 m^2.

The current R&D is focused on the construction, tests and improvements of the technological prototypes. They are developed within CALICE and ILD collaborations.

The previous physics prototype was successfully tested in 2005-2011 years at CERN, DESY, and Fermilab together with other CALICE HCAL prototypes. These data samples are used in the analysis of separation of two electromagnetic and electromagnetic – hadronic showers, presented in Chapter 4 of this thesis. Events are reconstructed with all three currently available PFA programs, Pandora, Garlic, and Arbor. Two shower events are obtained by the event mixing. The particle separation results obtained with the CALICE physics prototypes confirm Monte Carlo expectations. In addition, similar results are obtained for the ILD Monte Carlo, both for the nominal 5×5 mm^2 ECAL granularity and for 2.5×2.5 mm^2 pixels. Surprisingly, the latter results are worse for Pandora and Garlic proving that they are optimized for the nominal ECAL granularity and, strictly speaking, can not prove that 5×5 mm^2 granularity is optimal. The separation analysis is the main topic of this thesis.

The second part is devoted to the tests of the technological prototype. They are discussed in Chapter 5. First, this chapter very briefly summarizes the current R&D on this subject and then focuses on the results obtained during the PhD:

- the study of SKIROC chip characteristics (the dynamic range, the spread of the gains and the saturated signals etc.) using the charge injection,
- the study of the "square" events,
- the preliminary analysis of the longitudinal profile of 3 GeV e^+ showers accumulated in Jun'17 DESY tests,
- the debugging of the DAQ data stream integrity and
- the preliminary determination of the shower fractal dimension in the transverse plane using data from Nov'15 beam tests at CERN.

Finally, the third part of the thesis described in the last chapter contains the preliminary results on the precision which can be achieved at ILD in the measurement of the $W^{\pm}$ boson mass using the hadronic decay to two jets. Currently, the world average has the uncertainty of 15 MeV. It is expected that ILD will significantly improve it. The current uncertainty from the overall fit of Standard Model parameters is 7.8 MeV, therefore improved $W^{\pm}$ mass measurement will test the validity of the Standard Model. The PFA calorimeter performance is crucial for this task, both the jet energy resolution and the linearity of the jet energy reconstruction.

Appendix A

Appendix for Chapter 4 on particle separation

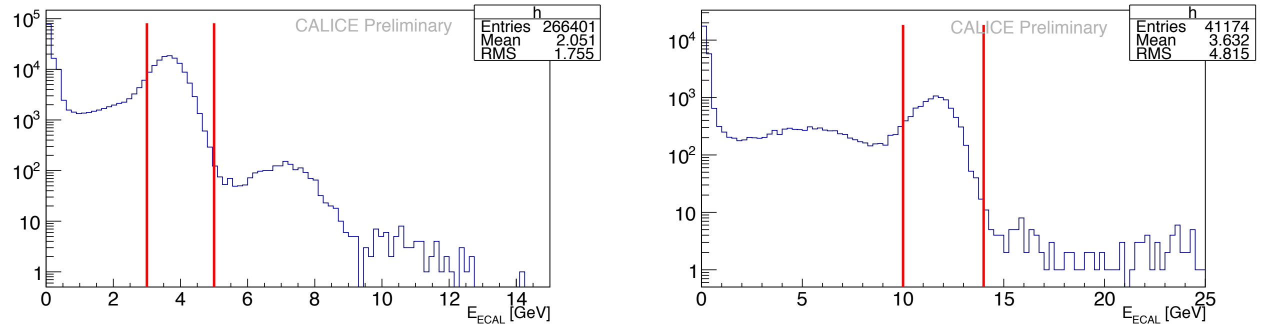
A.1 Distribution of variables and the corresponding cut values used in the data selection

For the analysis we use single particle events, e^+ and π^+, collected during CALICE beam tests at CERN (2007) and Fermilab (2011). Results are compared with γ, e^+ and π^+ MC simulations. All used CALICE event samples are listed in Table A.1. All events before the analysis has passed selection requirements described in section 4.3.2. The corresponding distributions are shown below. Note that all cuts are applied consecutively, i.e. each distribution is shown for the events satisfying the requirements from the previous plots.

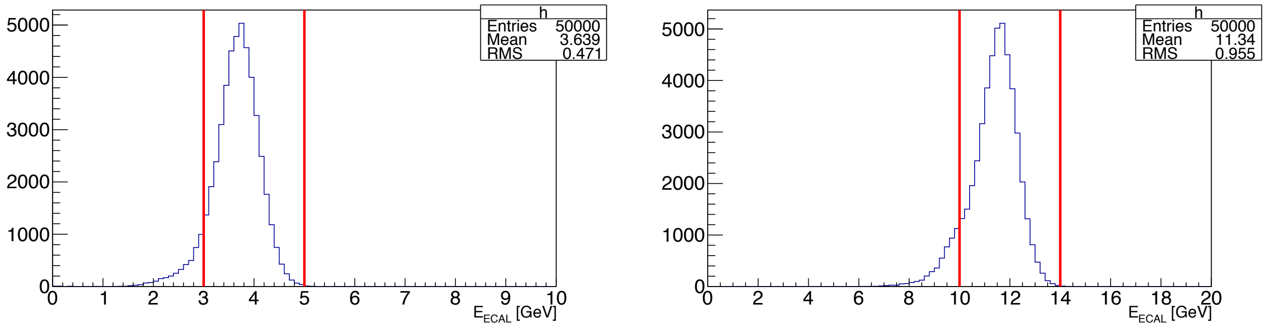
TB setup	Energy, particle
FNAL'11	4, 12, 25, 32 GeV e^+, γ (only MC)
CERN'07	10, 25 GeV e^+, γ (only MC) 30 GeV π^+

Table A.1 – List of CALICE particles used for the analysis.

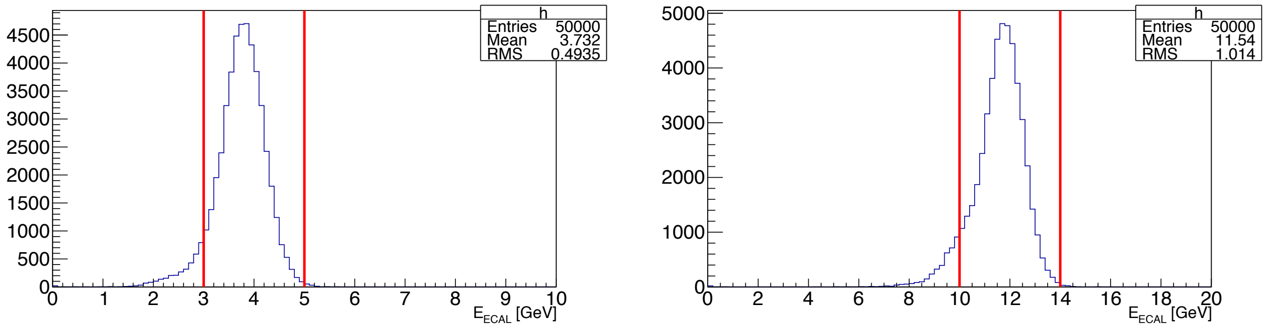
A.1.1 Full reconstructed energy before other cuts



(a) TB 4 GeV e^+ (b) TB 12 GeV e^+

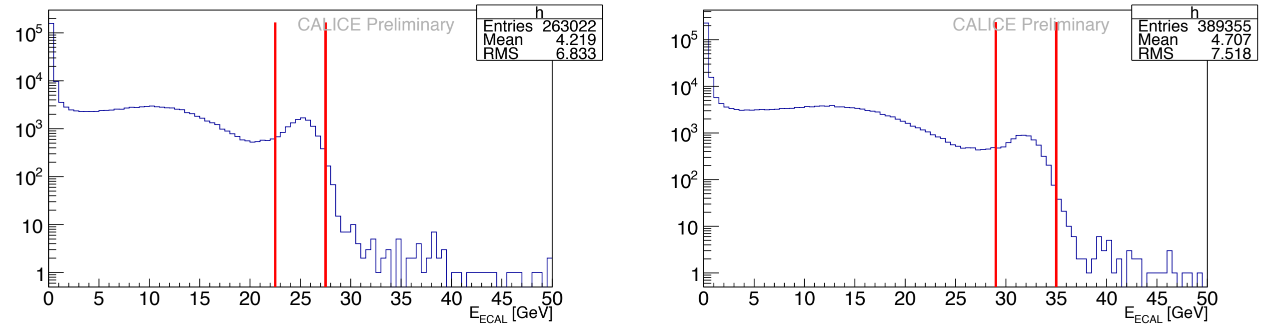


(c) MC 4 GeV e^+ (d) MC 12 GeV e^+

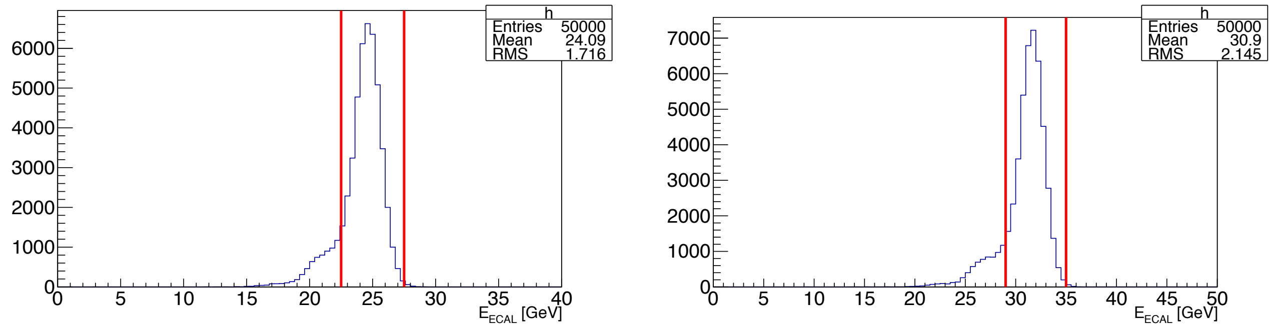


(e) MC 4 GeV γ (f) MC 12 GeV γ

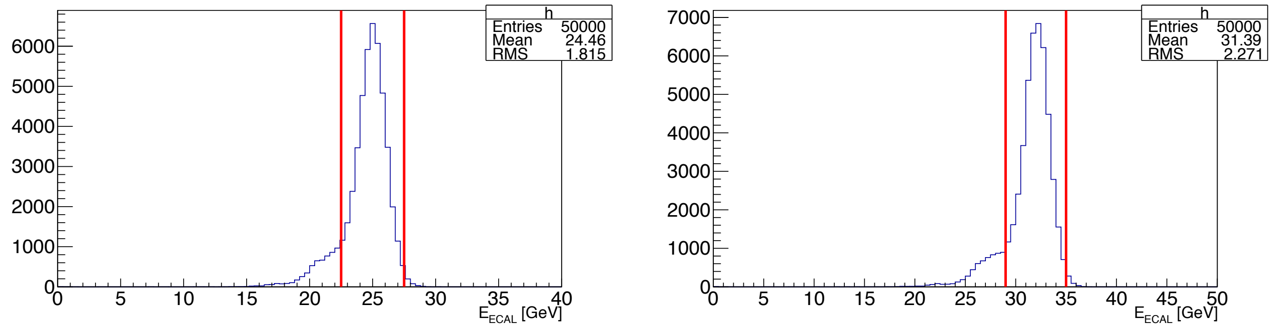
Figure A.1 – Full reconstructed energy (defined by Eq. 4.2) in ECAL for 4 and 12 GeV e^+, γ before applying any cut (FNAL'11). For the TB e^+ samples main backgrounds are pions (MIP-like and with developed showers in ECAL), muons (MIP-like), double entries (for 4 GeV e^+ sample one can clearly see the peak from double e^+ events, for other energies this double e^+ peak is not so clear, but, for example, double μ^+e^+ are still visible in the event display). Left tail in MC distributions mostly caused by shower leakage in the ECAL transverse direction.



(a) TB 25 GeV e^+ (b) TB 32 GeV e^+

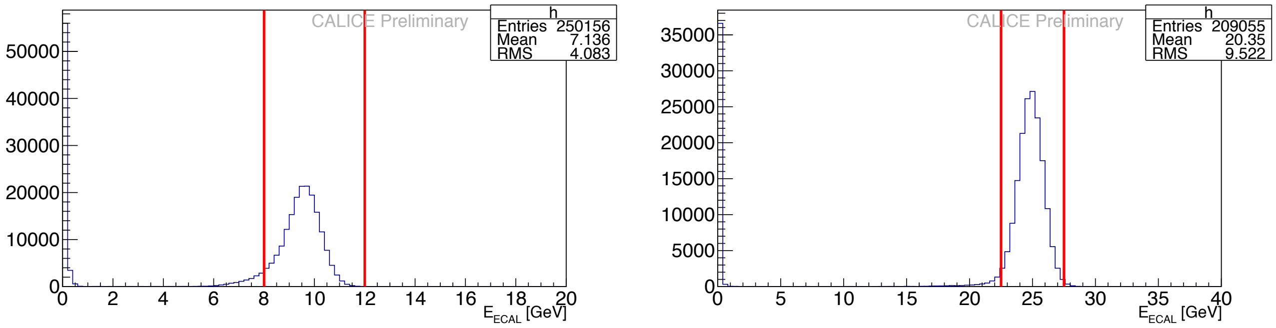


(c) MC 25 GeV e^+ (d) MC 32 GeV e^+

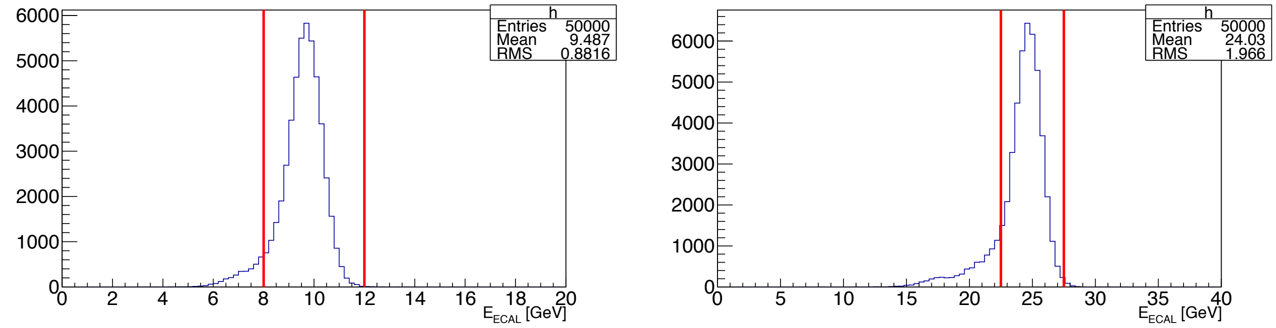


(e) MC 25 GeV γ (f) MC 32 GeV γ

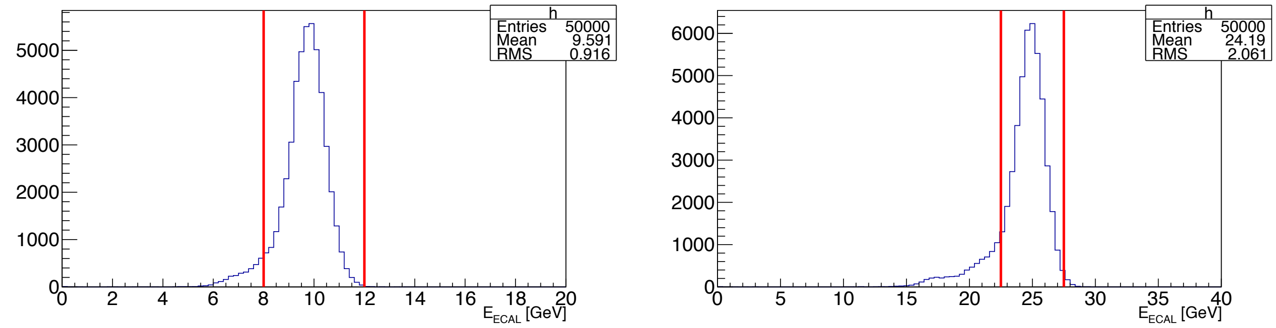
Figure A.2 – Same as in Figure A.1, but for energies 25 and 32 GeV (FNAL'11).



(a) TB 10 GeV e^+ (b) TB 25 GeV e^+

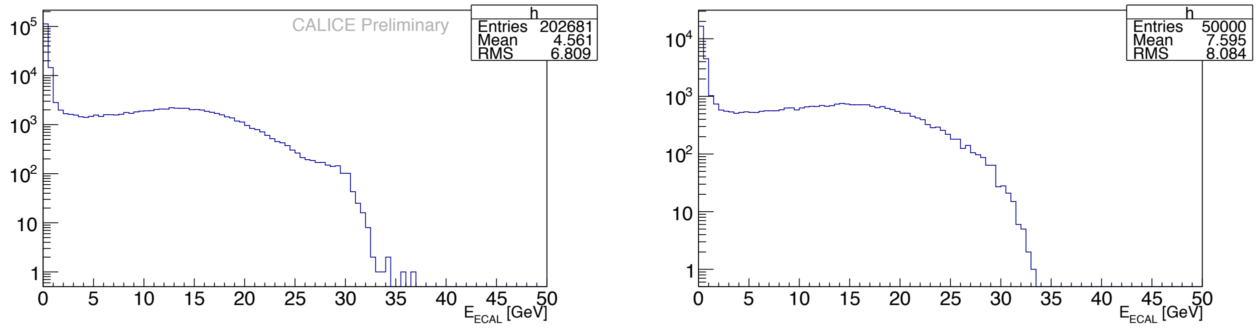


(c) MC 10 GeV e^+ (d) MC 25 GeV e^+



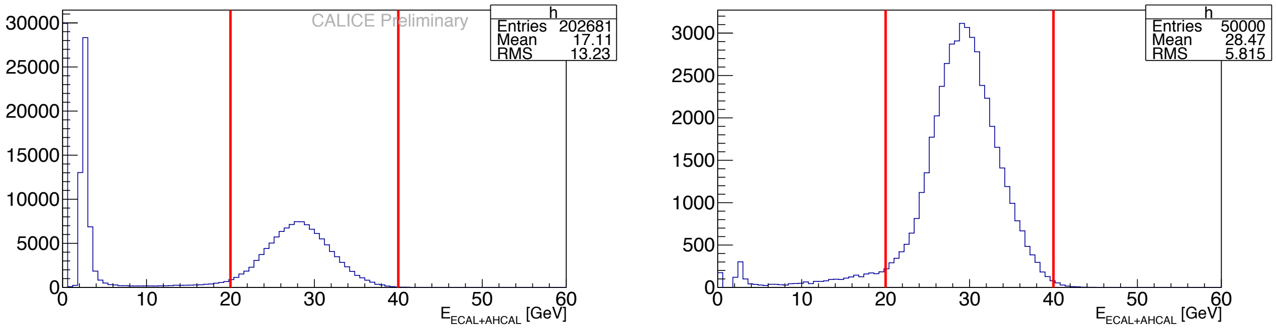
(e) MC 10 GeV γ (f) MC 25 GeV γ

Figure A.3 – Same as in Figures A.1 and A.2, for 10 and 25 GeV (CERN'07). One can see that CERN'07 beam has less contamination than FNAL'11 tests.



(a) TB 30 GeV π^+ ECAL

(b) MC 30 GeV π^+ ECAL

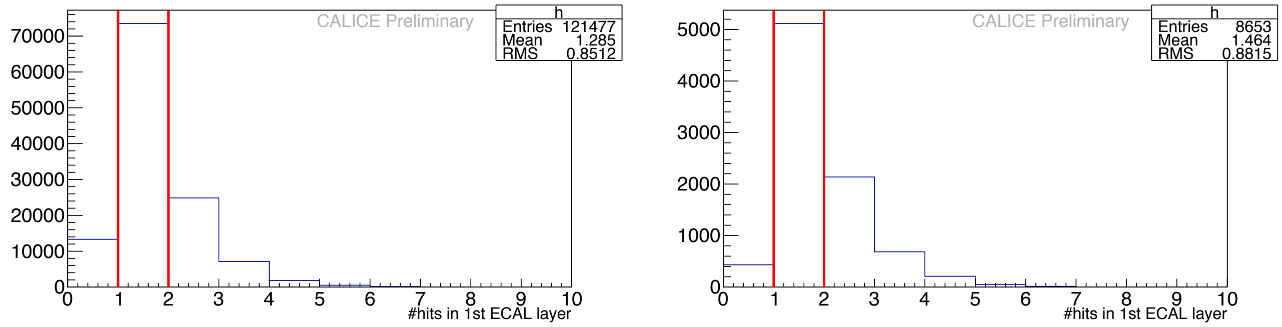


(c) TB 30 GeV π^+ ECAL+AHCAL

(d) MC 30 GeV π^+ ECAL+AHCAL

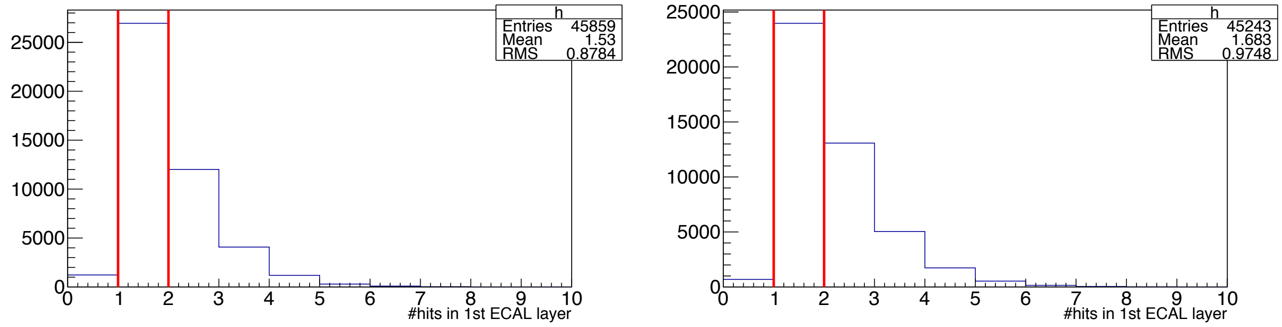
Figure A.4 – Same as in Figures A.1-A.3, for 30 GeV pions (CERN'07). The upper (lower) plots show the full reconstructed energy in ECAL (ECAL and AHCAL). Left peaks in the distributions correspond to muons. TB π^+ sample has small e^+ contamination.

A.1.2 Number of hits in the first layer of ECAL



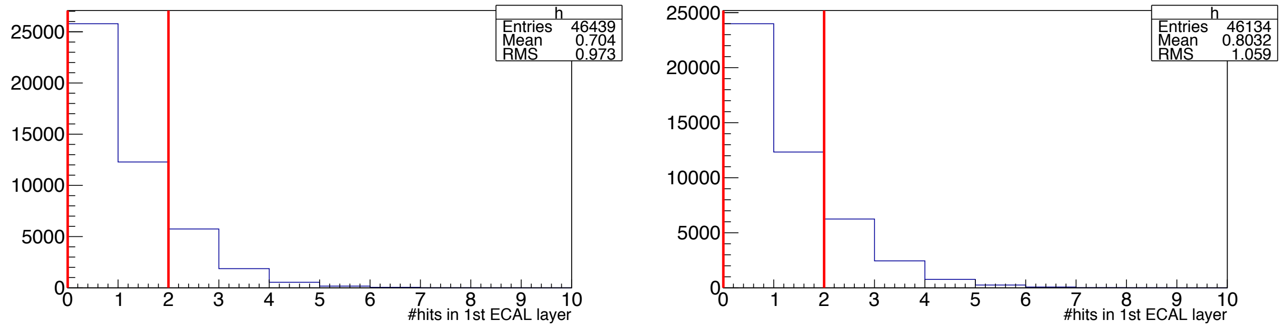
(a) TB 4 GeV e^+

(b) TB 12 GeV e^+



(c) MC 4 GeV e^+

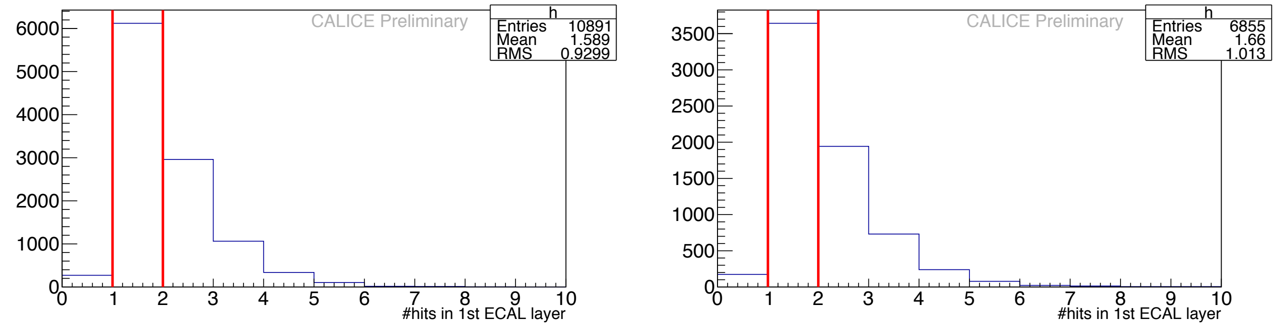
(d) MC 12 GeV e^+



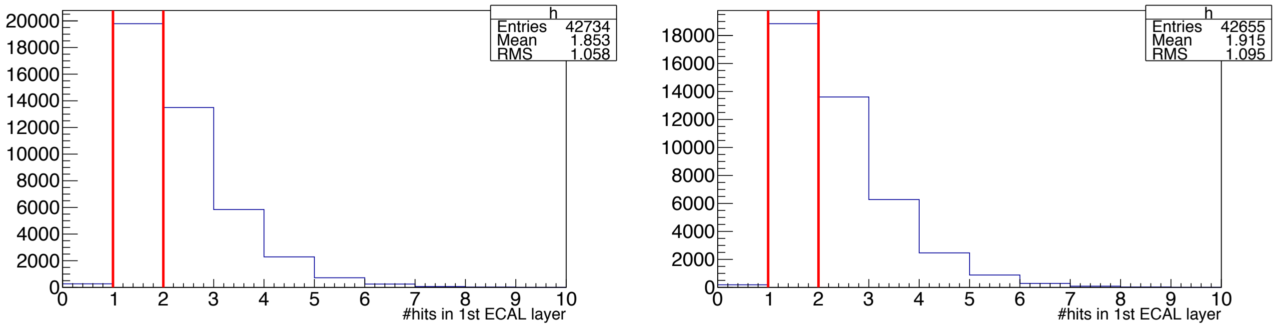
(e) MC 4 GeV γ

(f) MC 12 GeV γ

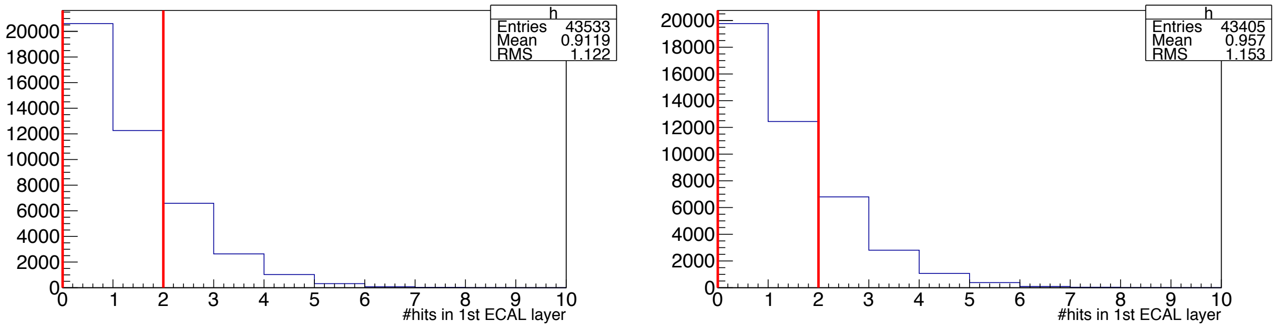
Figure A.5 – Number of hits in the first (# 0) layer of ECAL for 4 and 12 GeV (FNAL'11) after the cut on full ECAL energy from Figure A.1. One can see that for the TB the ratio between number of events with 1 and 2 hits in the 1st ECAL layer is smaller than for MC. It confirms the existence of double entries and the sample is cleaned from these events by requiring just 1 hit in the 1st ECAL layer. Of course, it means that some physical events with two close hits in the 1st layer are excluded from the analysis. For MC γ bin 0 is also accepted as most of the photons pass through the first ECAL layer without interaction.



(a) TB 25 GeV e^+ (b) TB 32 GeV e^+

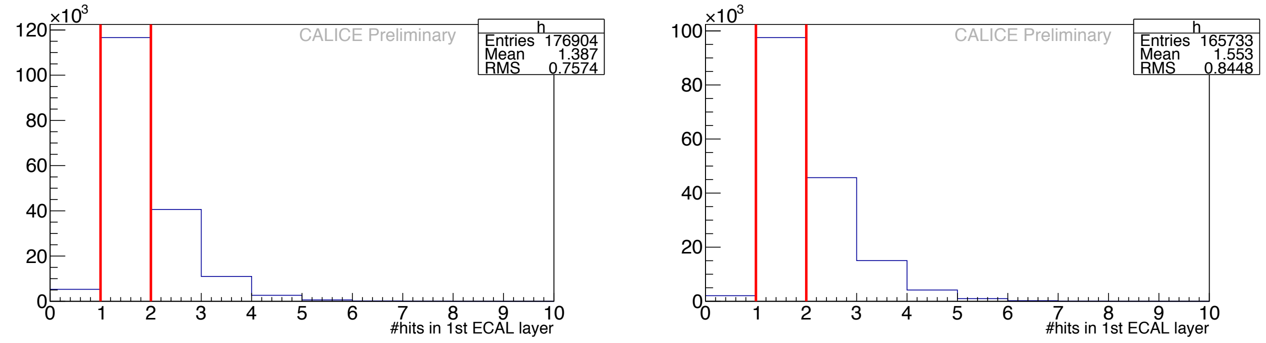


(c) MC 25 GeV e^+ (d) MC 32 GeV e^+



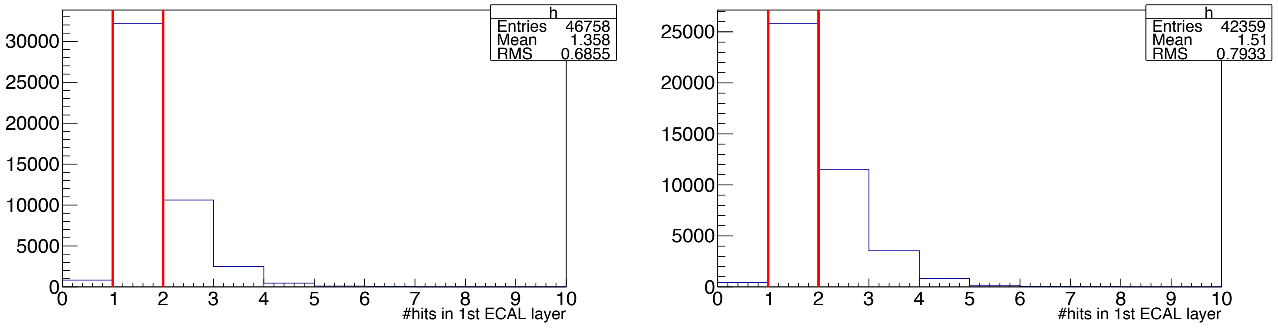
(e) MC 25 GeV γ (f) MC 32 GeV γ

Figure A.6 – Same as in Figure A.5, for 25 and 32 GeV (FNAL'11).



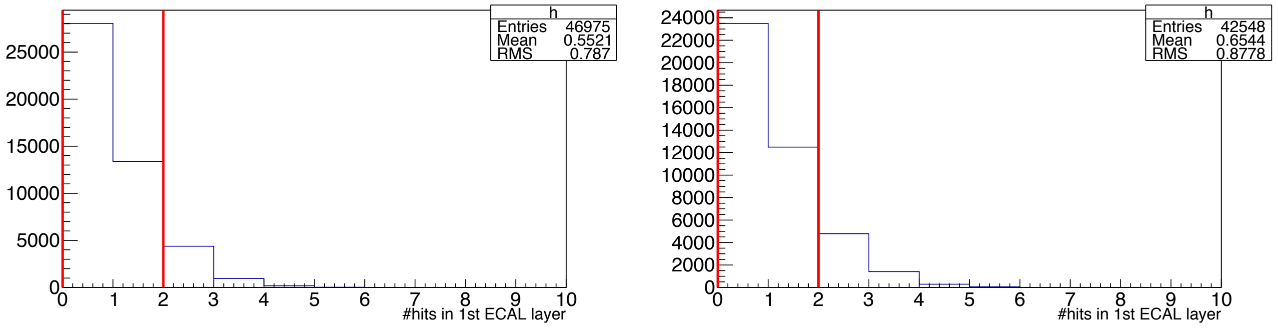
(a) TB 10 GeV e^+

(b) TB 25 GeV e^+



(c) MC 10 GeV e^+

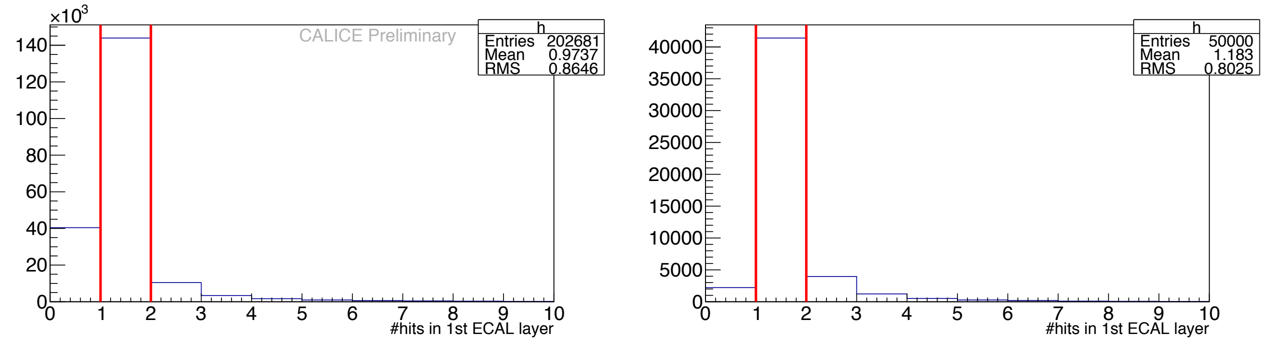
(d) MC 25 GeV e^+



(e) MC 10 GeV γ

(f) MC 25 GeV γ

Figure A.7 – Same as in Figures A.5 and A.6, for 10 and 25 GeV (CERN'07).

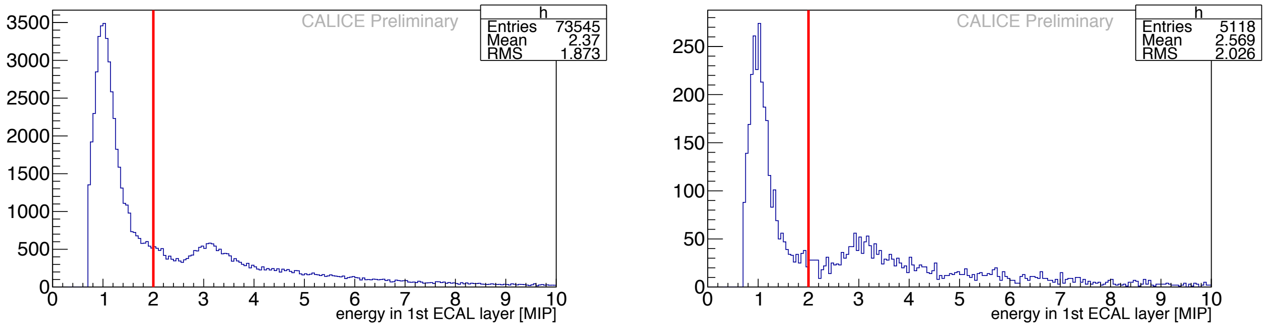


(a) TB 30 GeV π^+

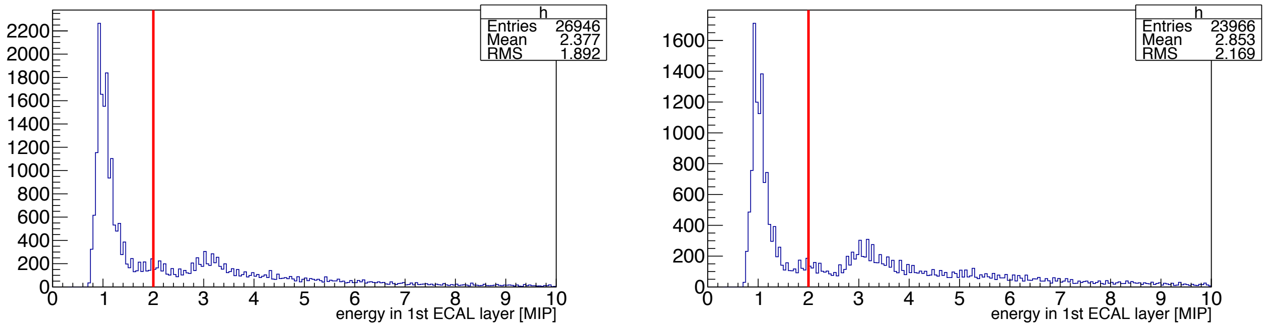
(b) MC 30 GeV π^+

Figure A.8 – Same as in previous Figures A.5 - A.7, for 30 GeV pions (CERN'07). We require exactly one hit in the first ECAL layer, as for e^+. The population of the bin zero is larger in data due to the noises.

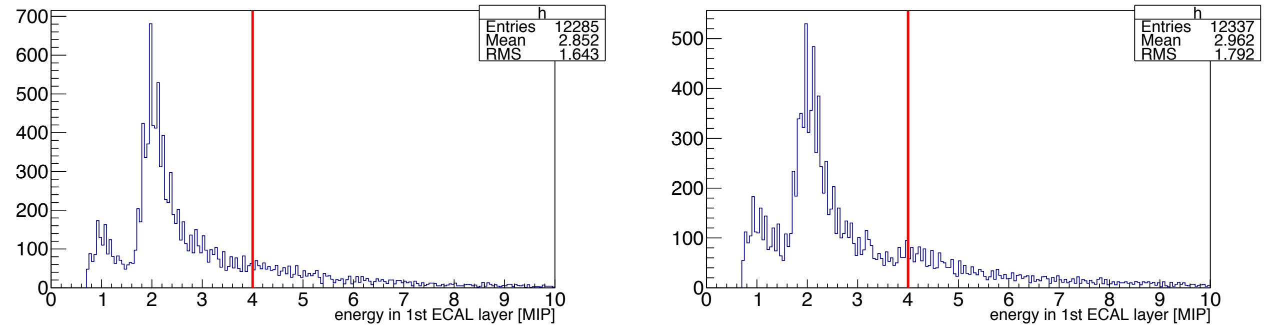
A.1.3 Total energy in the first layer of ECAL



(a) TB 4 GeV e^+ (b) TB 12 GeV e^+

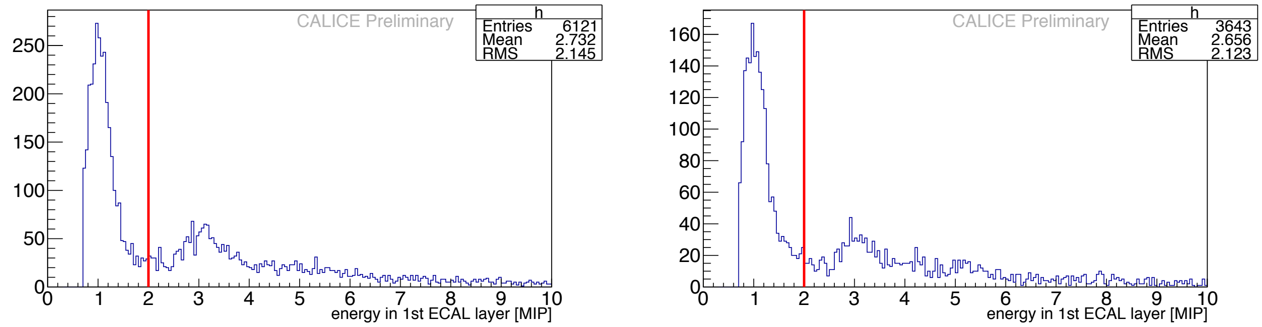


(c) MC 4 GeV e^+ (d) MC 12 GeV e^+

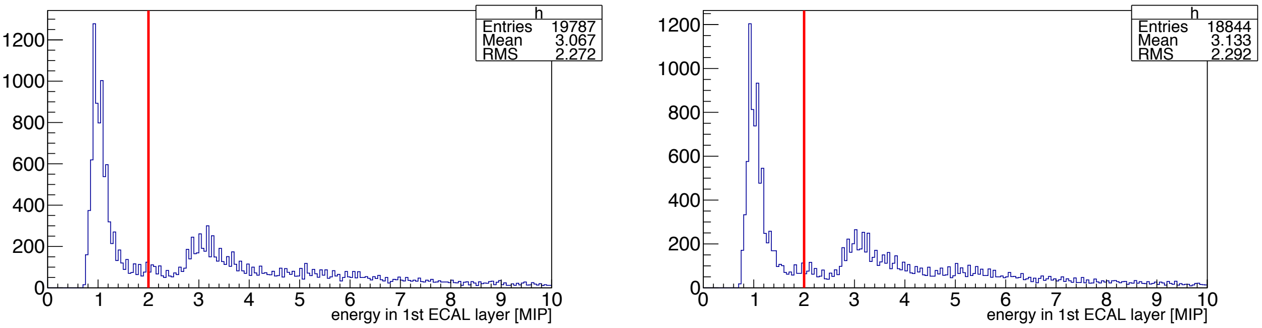


(e) MC 4 GeV γ (f) MC 12 GeV γ

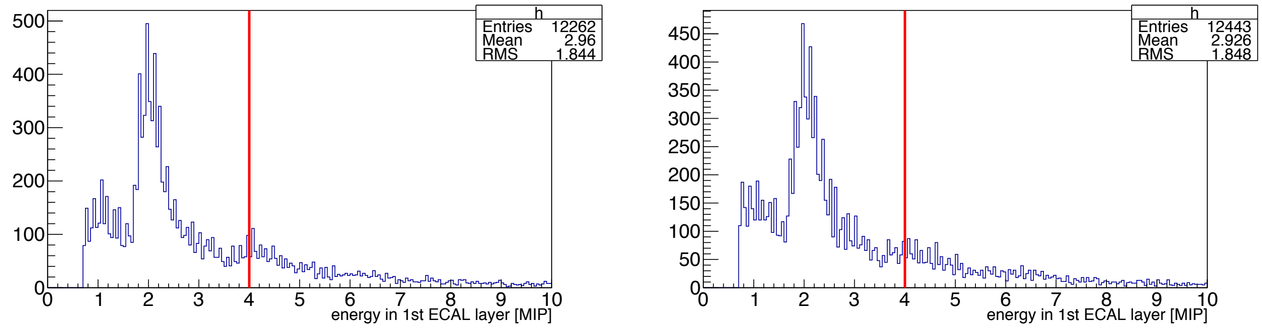
Figure A.9 – Total energy (in MIPs) in the first layer (# 0) of ECAL for 4 and 12 GeV (FNAL'11). For e^+ (γ) the peaks are at odd numbers 1, 3 (even 2, 4) because of the pair-wise production of e^+e^- in ECAL interactions. Positrons (photons) with energy smaller than 2 (4) MIPs are selected.



(a) TB 25 GeV e^+ (b) TB 32 GeV e^+

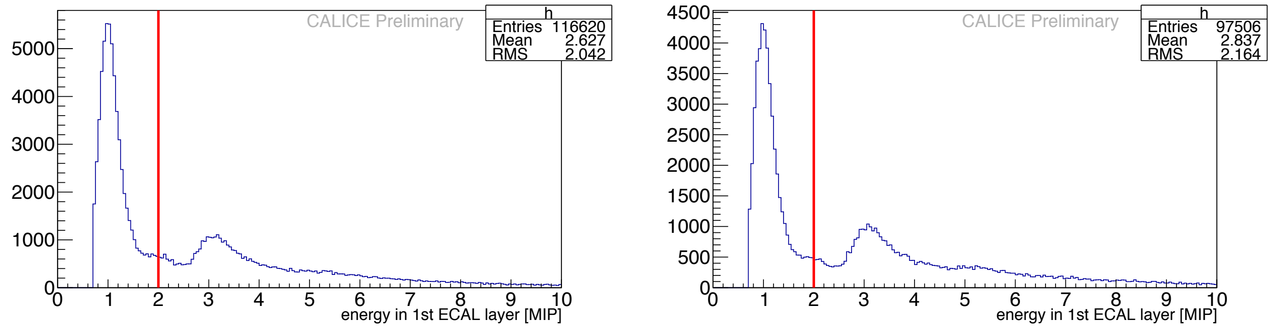


(c) MC 25 GeV e^+ (d) MC 32 GeV e^+

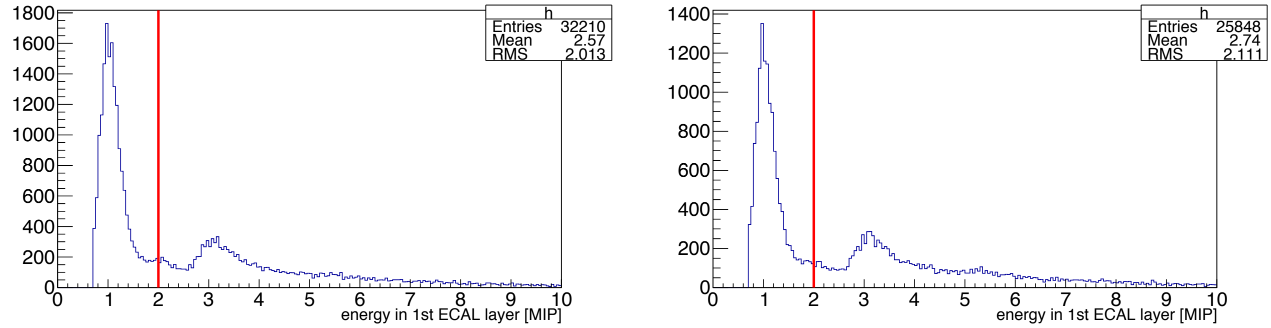


(e) MC 25 GeV γ (f) MC 32 GeV γ

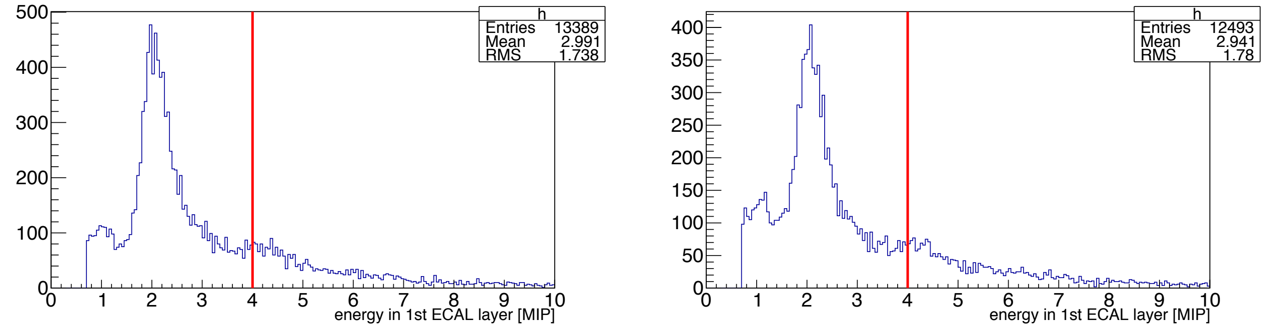
Figure A.10 – Same as in Figure A.9, for 25 and 32 GeV (FNAL'11).



(a) TB 10 GeV e^+ (b) TB 25 GeV e^+

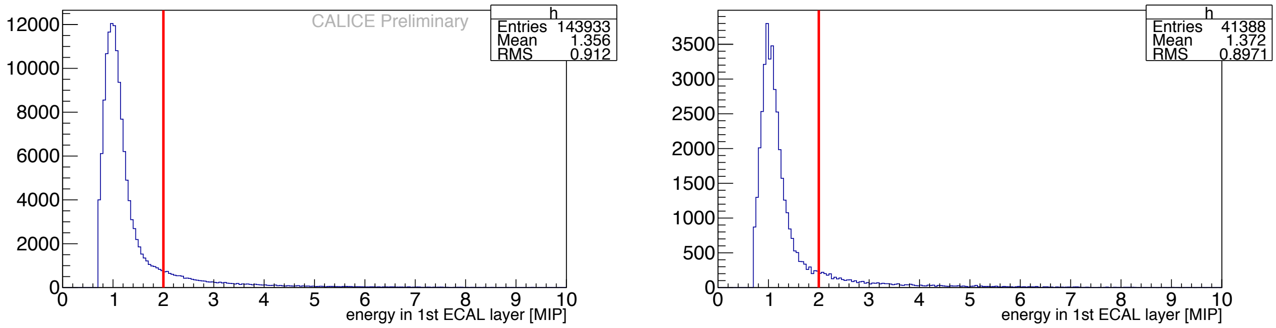


(c) MC 10 GeV e^+ (d) MC 25 GeV e^+



(e) MC 10 GeV γ (f) MC 25 GeV γ

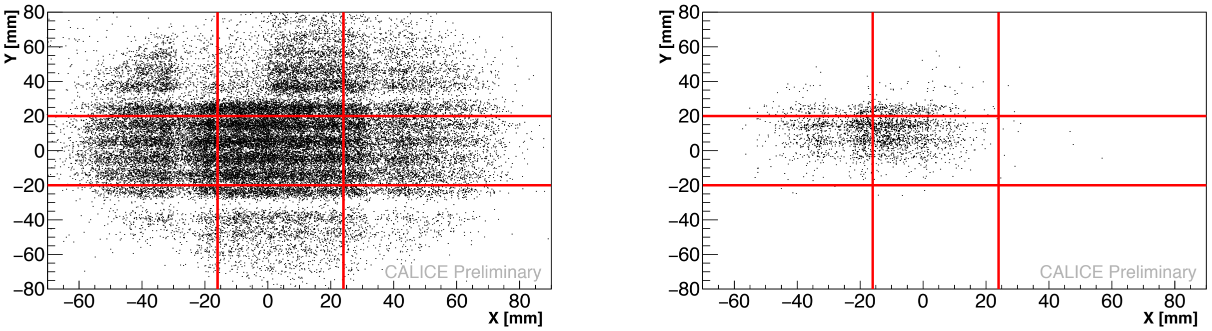
Figure A.11 – Same as in Figures A.9 and A.10, for 10 and 25 GeV (CERN'07).



(a) TB 30 GeV π^+ (b) MC 30 GeV π^+

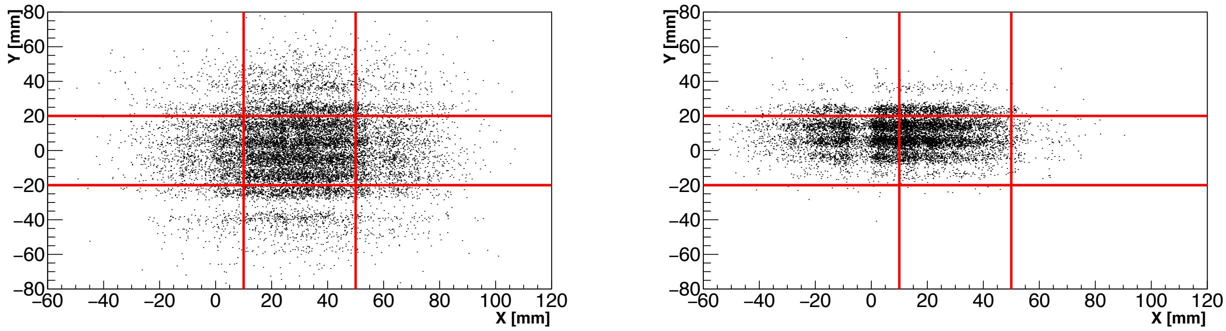
Figure A.12 – Same as is previous Figures A.9 - A.11, for 30 GeV pions (CERN'07).

A.1.4 Particle position in ECAL



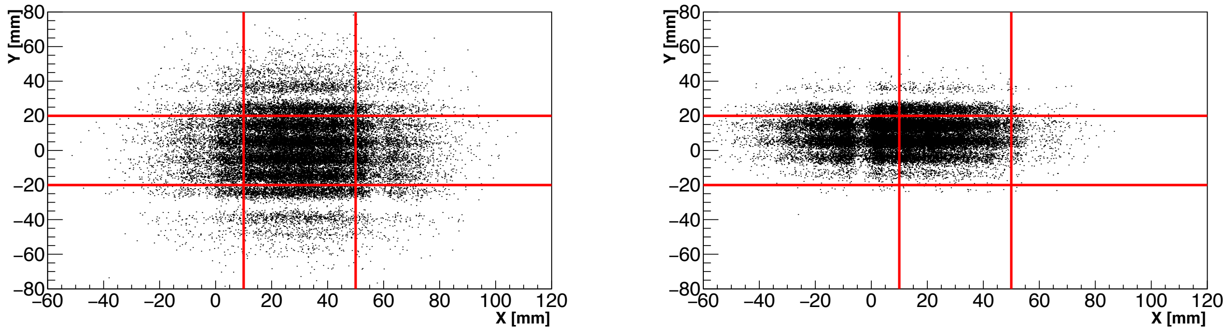
(a) TB 4 GeV e^+

(b) TB 12 GeV e^+



(c) MC 4 GeV e^+

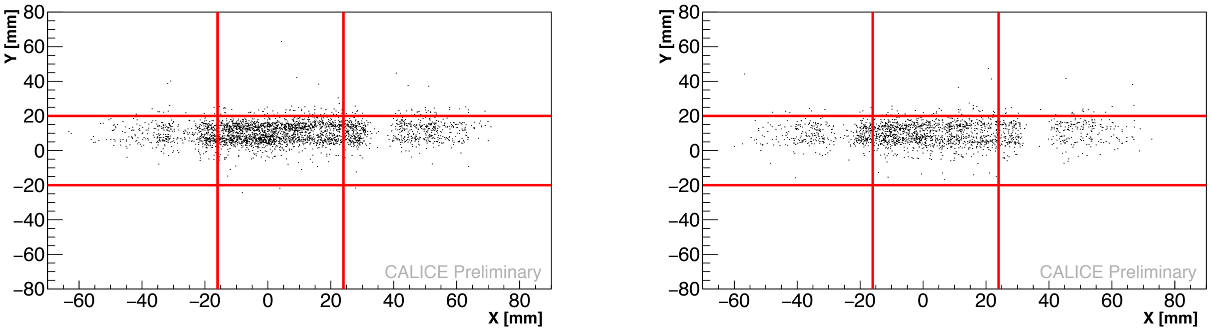
(d) MC 12 GeV e^+



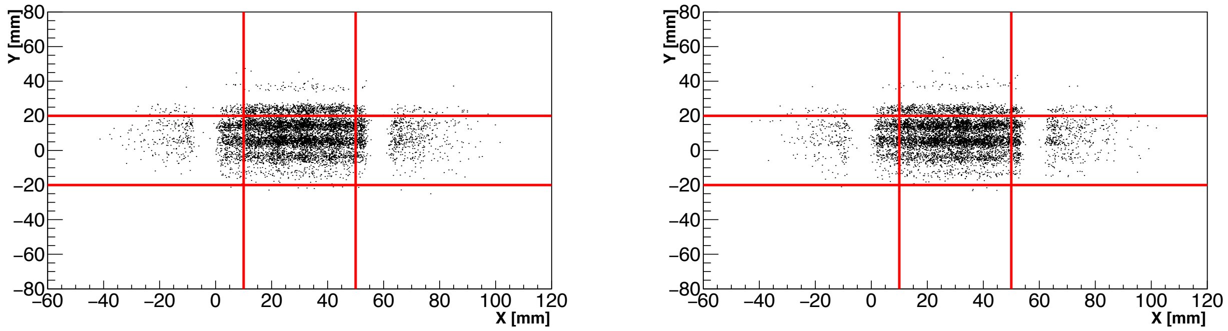
(e) MC 4 GeV γ

(f) MC 12 GeV γ

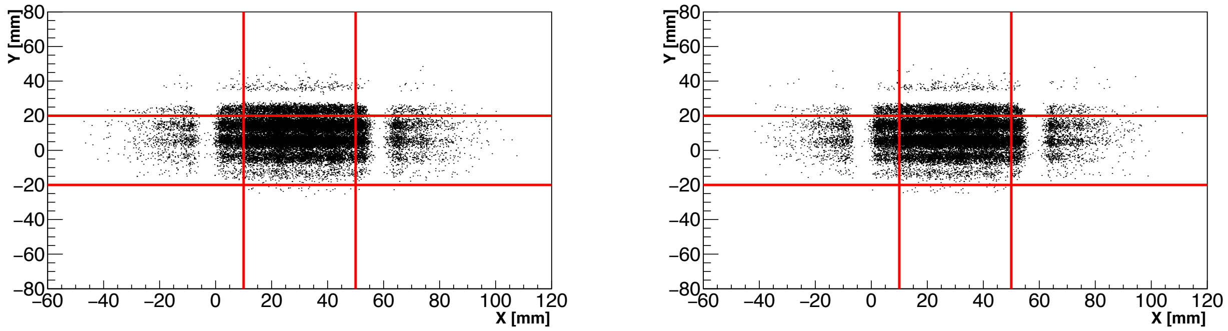
Figure A.13 – Transverse barycenter (energy center-of-gravity) position in ECAL for 4 and 12 GeV (FNAL'11). For the analysis it is important to reduce the relative importance of the dead zones between the sensors and to suppress events with significant transverse shower leakage. Therefore, we use events with energy barycenters in the central 4×4 cm^2 part of the central ECAL wafer.



(a) TB 25 GeV e^+ (b) TB 32 GeV e^+

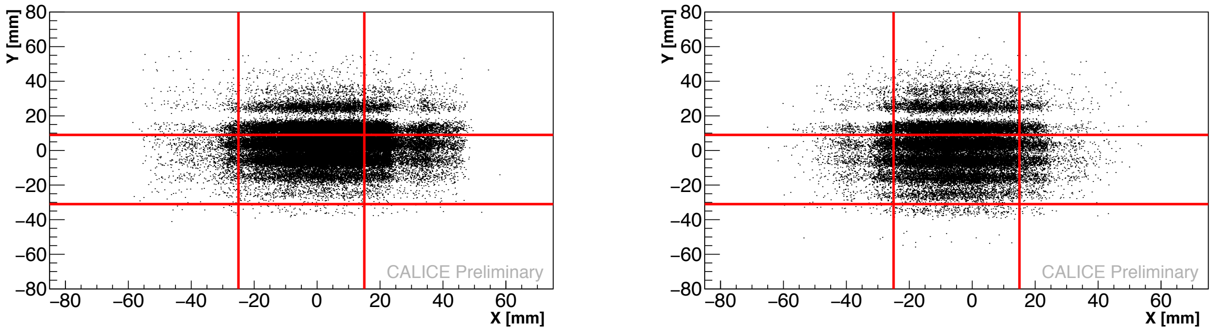


(c) MC 25 GeV e^+ (d) MC 32 GeV e^+

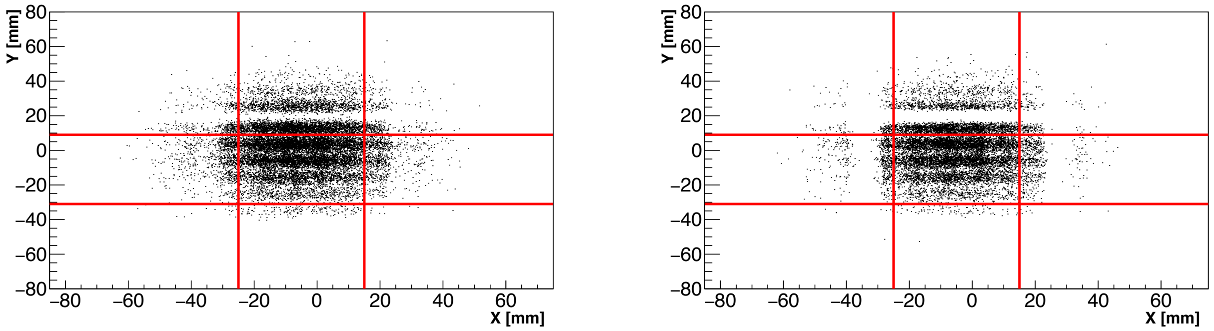


(e) MC 25 GeV γ (f) MC 32 GeV γ

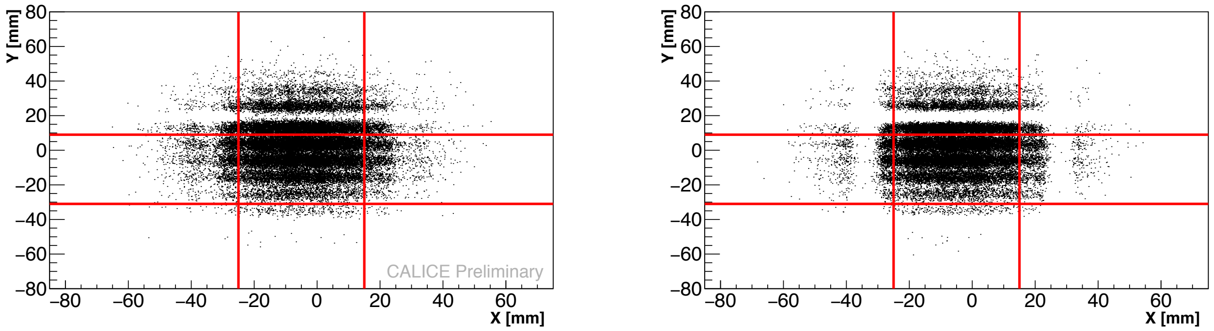
Figure A.14 – Same as in Figure A.13, for 25 and 32 GeV (FNAL'11).



(a) TB 10 GeV e^+ (b) TB 25 GeV e^+

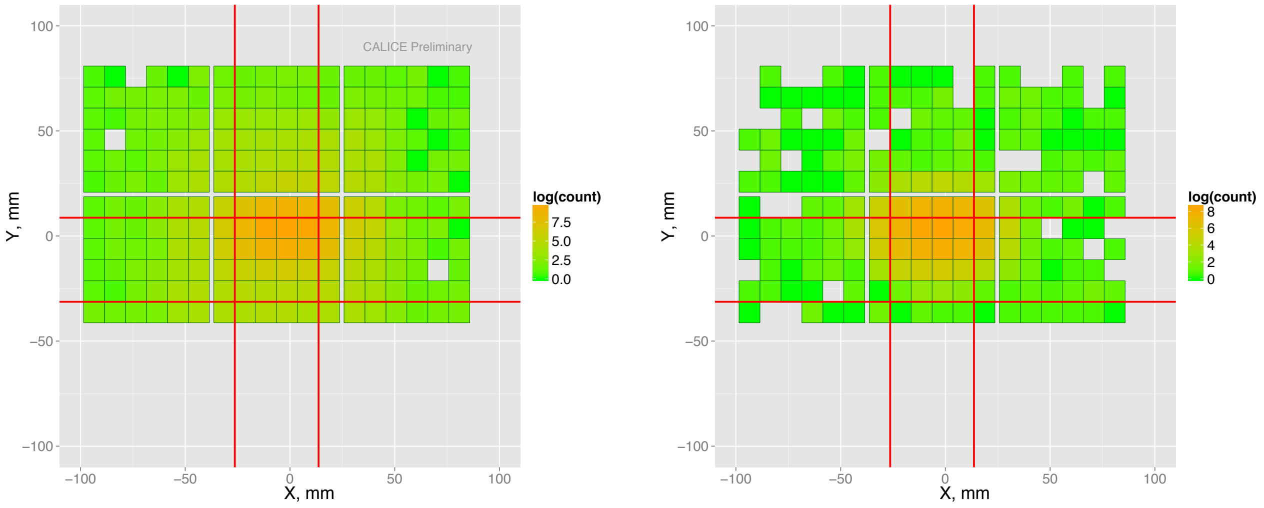


(c) MC 10 GeV e^+ (d) MC 25 GeV e^+



(e) MC 10 GeV γ (f) MC 25 GeV γ

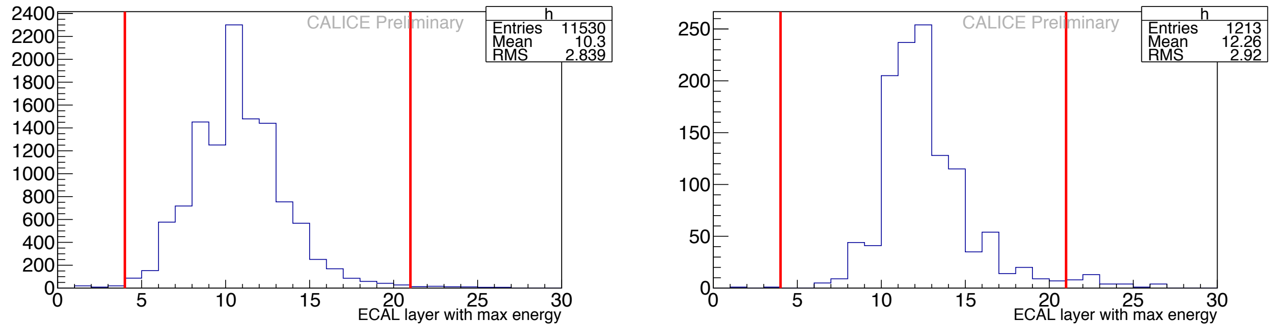
Figure A.15 – Same as in Figures A.13 and A.14, for 10 and 25 GeV (CERN'07).



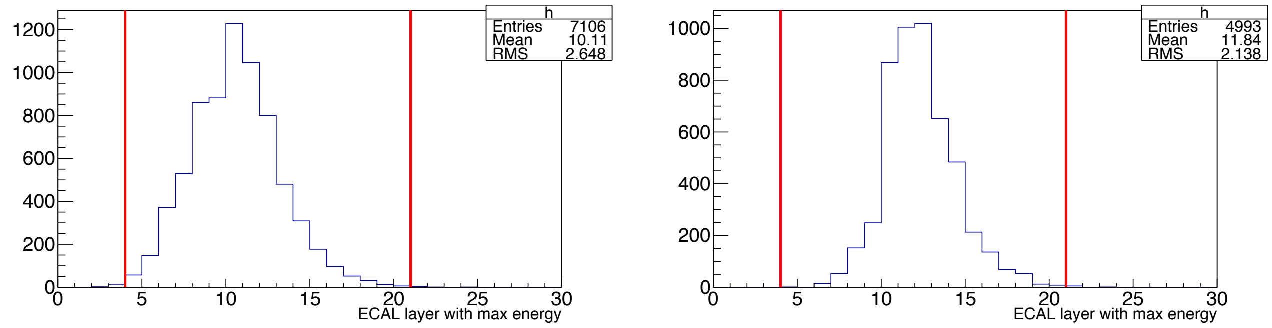
(a) TB 30 GeV π^+ (b) MC 30 GeV π^+

Figure A.16 – Occupancy map in the first ECAL layer for 30 GeV pions (CERN'07). One row of wafers in the six front-most layers was missing.

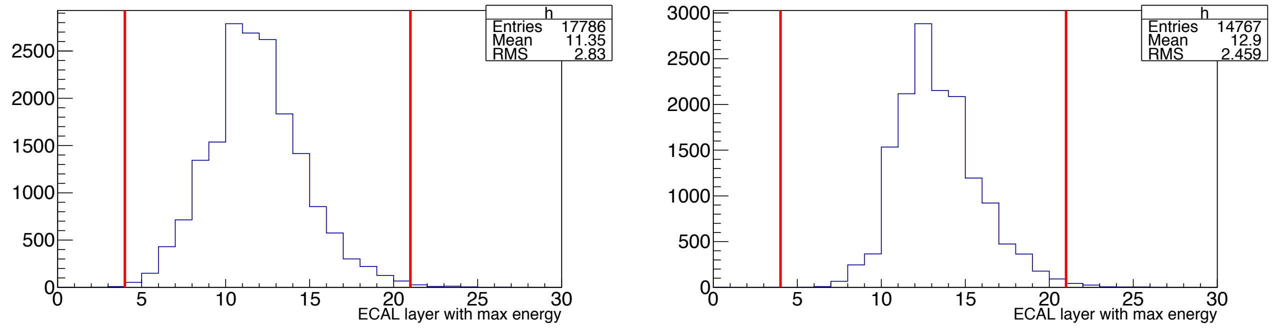
A.1.5 ECAL layer with maximal energy



(a) TB 4 GeV e^+ (b) TB 12 GeV e^+

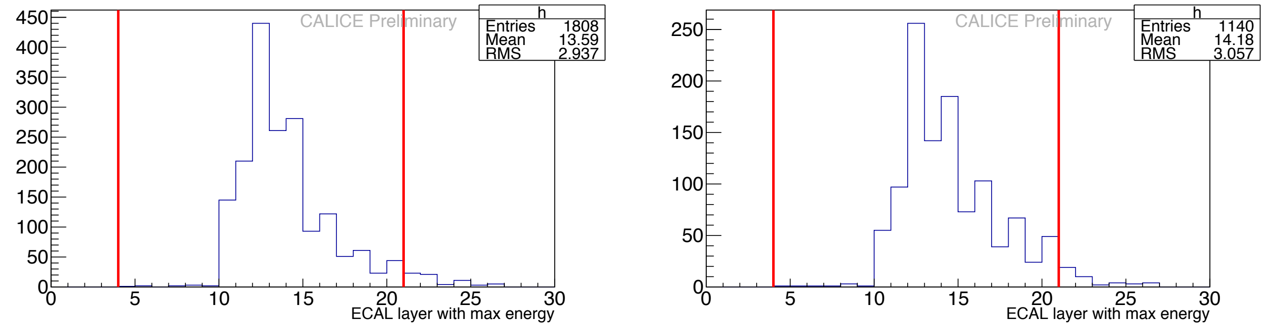


(c) MC 4 GeV e^+ (d) MC 12 GeV e^+

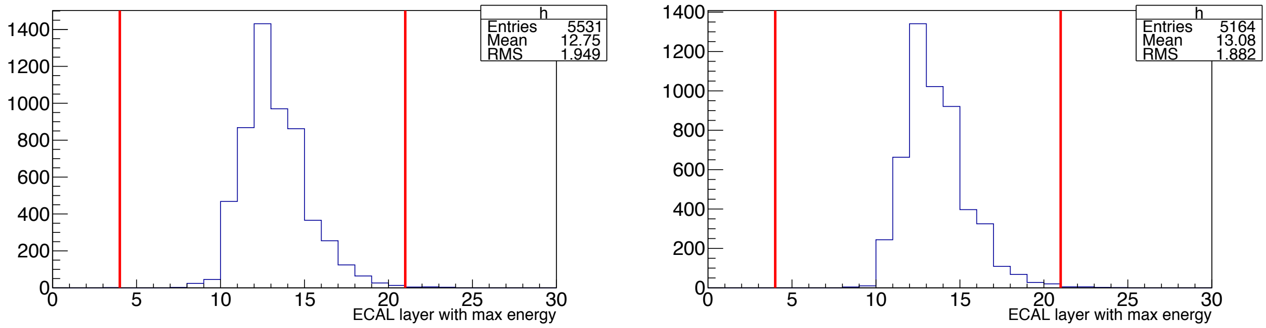


(e) MC 4 GeV γ (f) MC 12 GeV γ

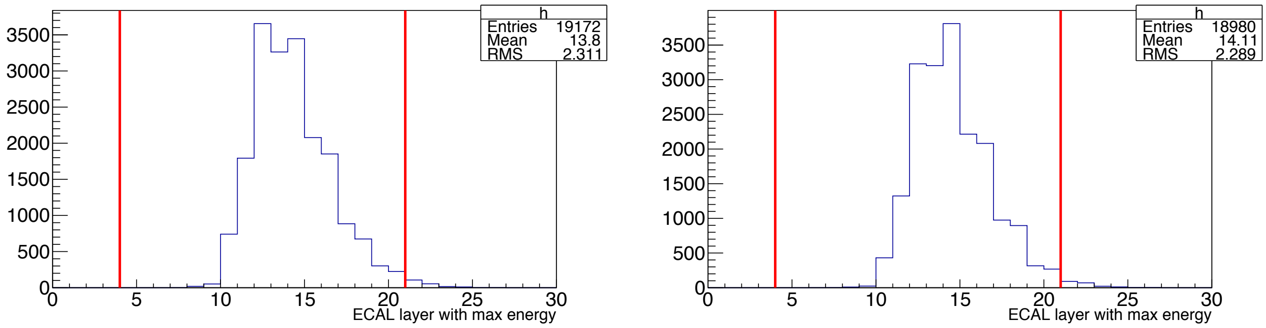
Figure A.17 – ECAL layer with the maximal energy for 4 and 12 GeV (FNAL'11). This cut is important against noisy layers and also to clean the sample from pions (where the layer with the highest energy deposition is close to the end of ECAL).



(a) TB 25 GeV e^+ (b) TB 32 GeV e^+

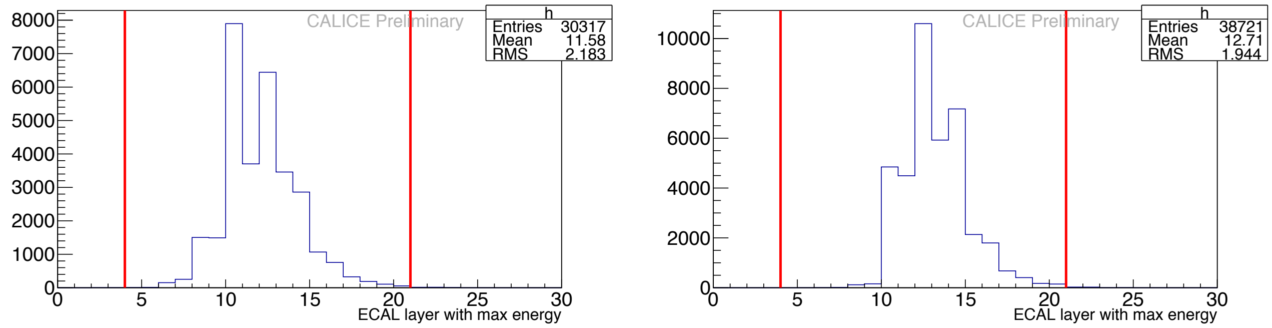


(c) MC 25 GeV e^+ (d) MC 32 GeV e^+

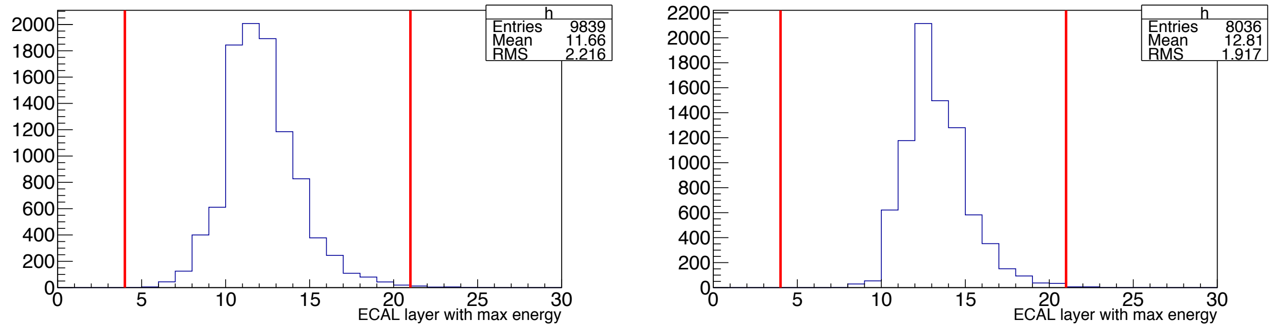


(e) MC 25 GeV γ (f) MC 32 GeV γ

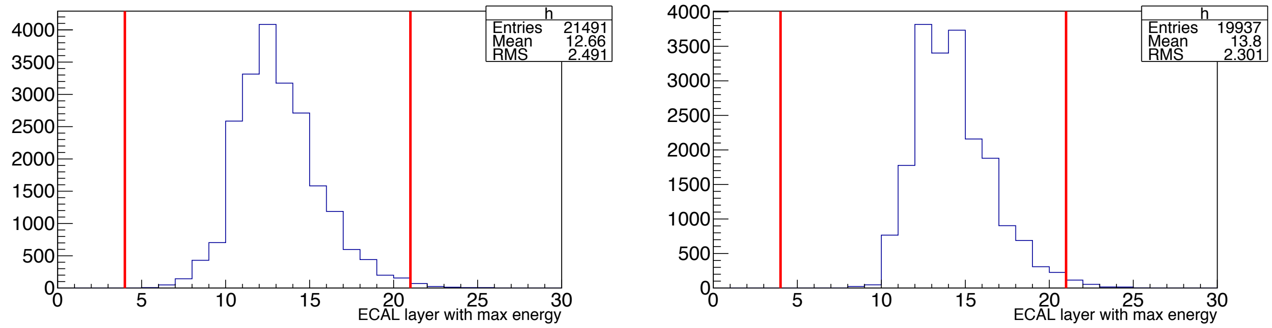
Figure A.18 – Same as in Figure A.17, for 25 and 32 GeV (FNAL'11).



(a) TB 10 GeV e^+ (b) TB 25 GeV e^+

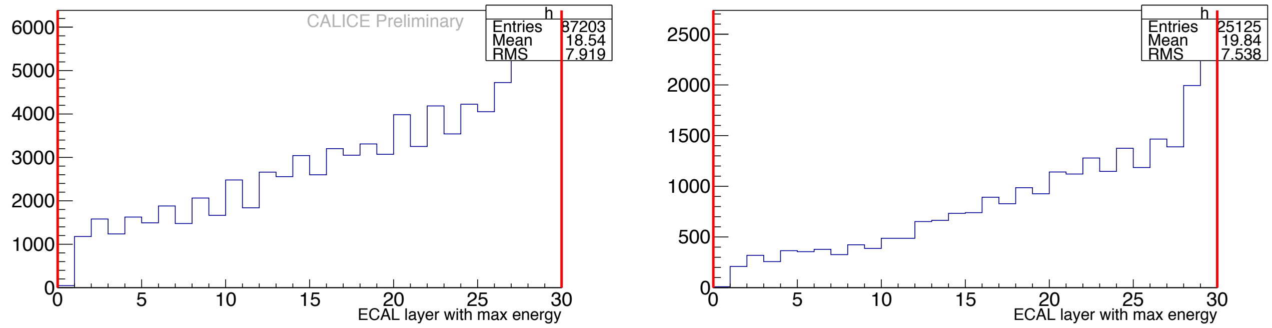


(c) MC 10 GeV e^+ (d) MC 25 GeV e^+



(e) MC 10 GeV γ (f) MC 25 GeV γ

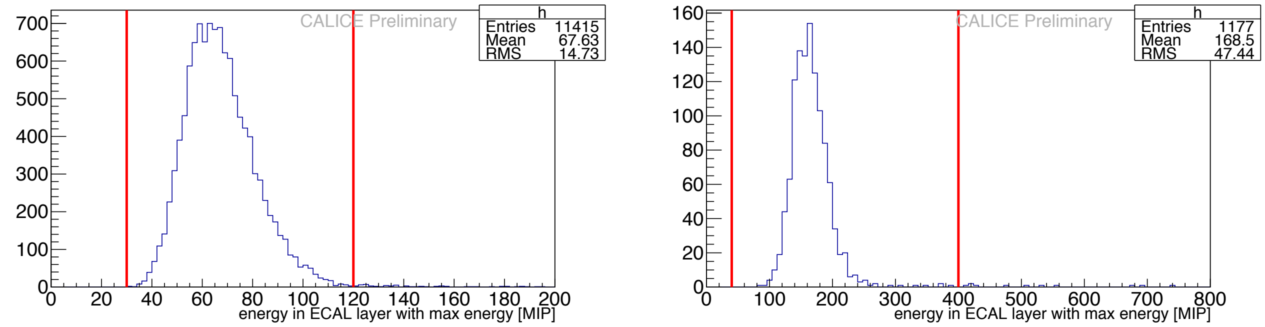
Figure A.19 – Same as in Figures A.17 and A.18, for 10 and 25 GeV (CERN'07).



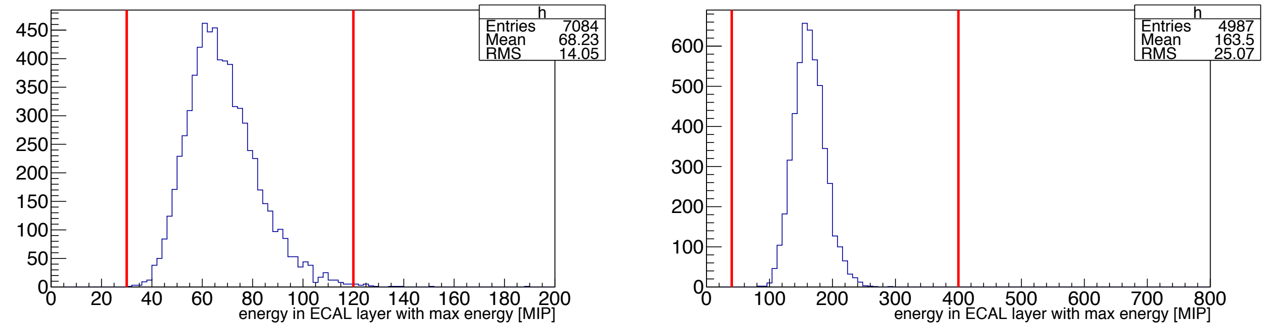
(a) TB 30 GeV π^+ (b) MC 30 GeV π^+

Figure A.20 – Same as in previous Figures A.17 - A.19, for 30 GeV pions (CERN'07). For π^+ every layer in ECAL can be layer with maximal energy deposition (no cut).

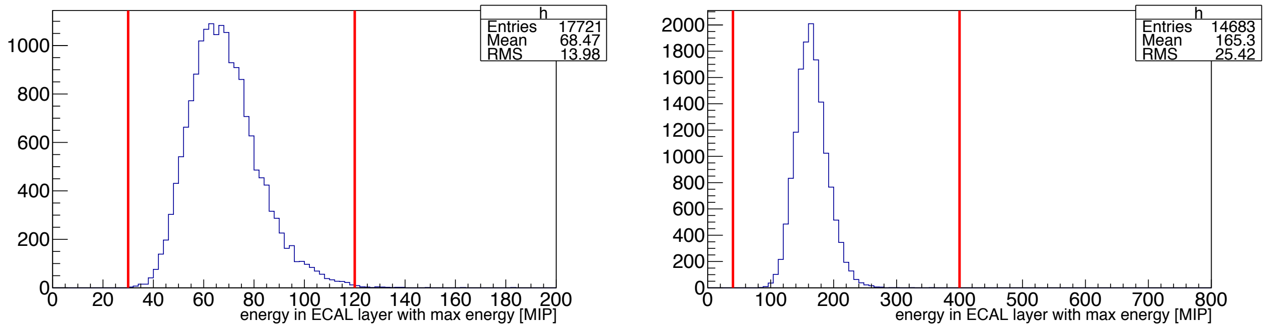
A.1.6 Energy in the ECAL layer with maximal energy



(a) TB 4 GeV e^+ (b) TB 12 GeV e^+

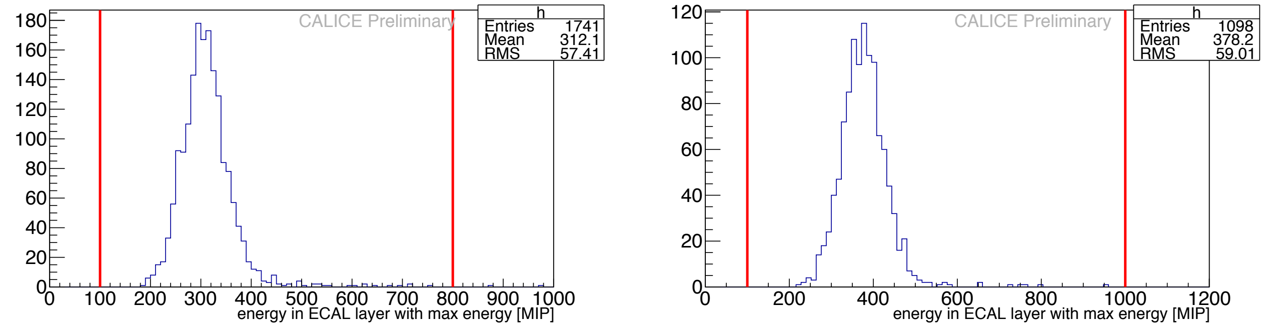


(c) MC 4 GeV e^+ (d) MC 12 GeV e^+

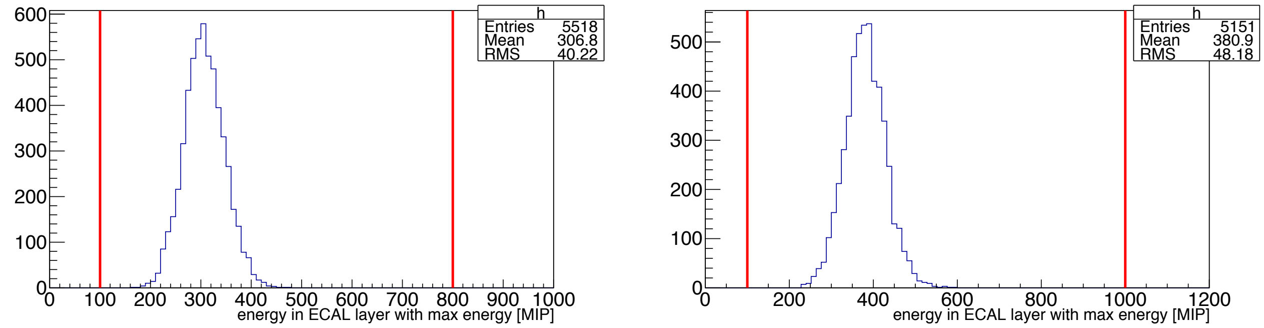


(e) MC 4 GeV γ (f) MC 12 GeV γ

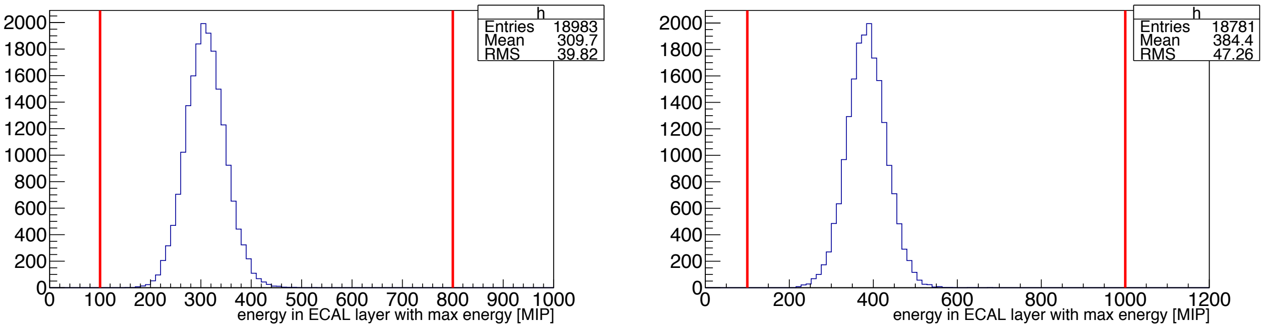
Figure A.21 – Energy (in MIPs) in the ECAL layer with maximal energy for 4 and 12 GeV (FNAL'11).



(a) TB 25 GeV e^+ (b) TB 32 GeV e^+

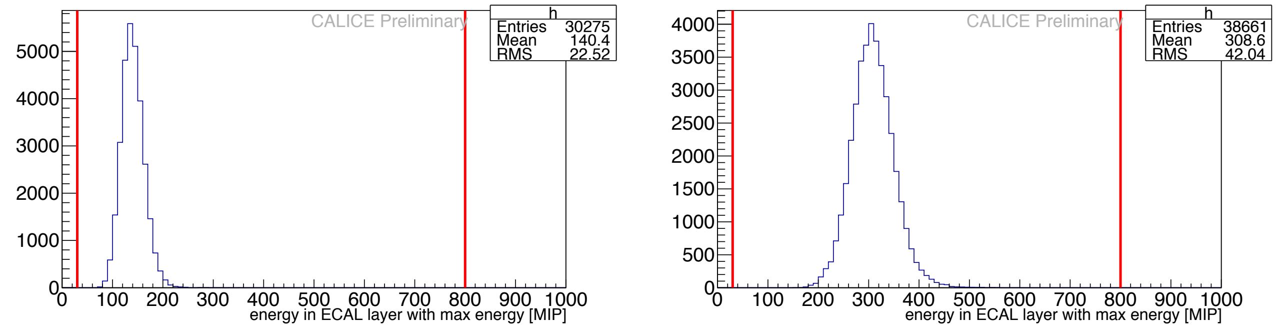


(c) MC 25 GeV e^+ (d) MC 32 GeV e^+

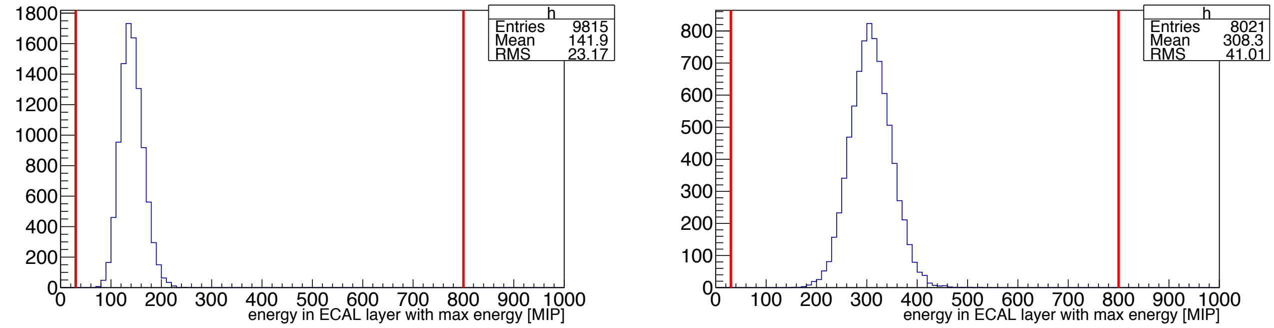


(e) MC 25 GeV γ (f) MC 32 GeV γ

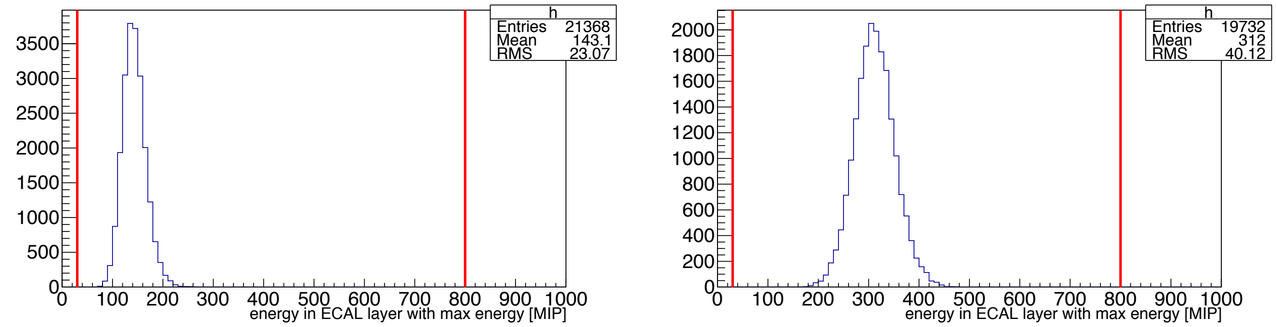
Figure A.22 – Same as in Figure A.21, for 25 and 32 GeV (FNAL'11).



(a) TB 10 GeV e^+ (b) TB 25 GeV e^+

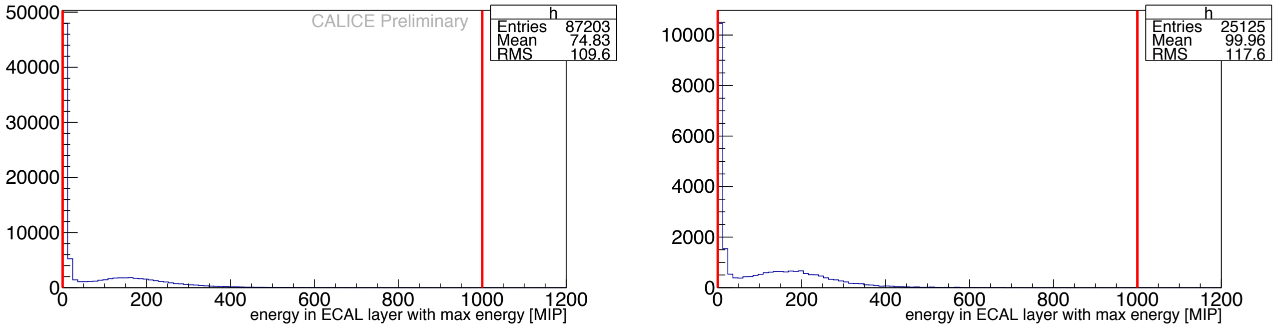


(c) MC 10 GeV e^+ (d) MC 25 GeV e^+



(e) MC 10 GeV γ (f) MC 25 GeV γ

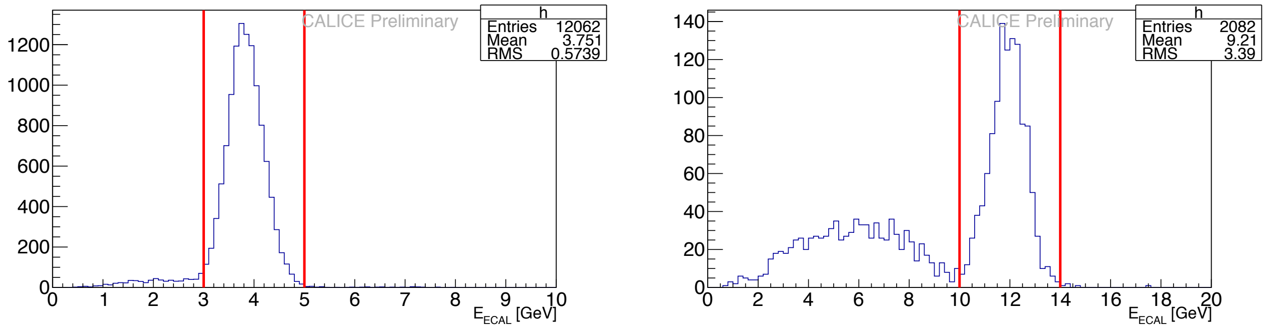
Figure A.23 – Same as in Figures A.21 and A.22, for 10 and 25 GeV (CERN'07).



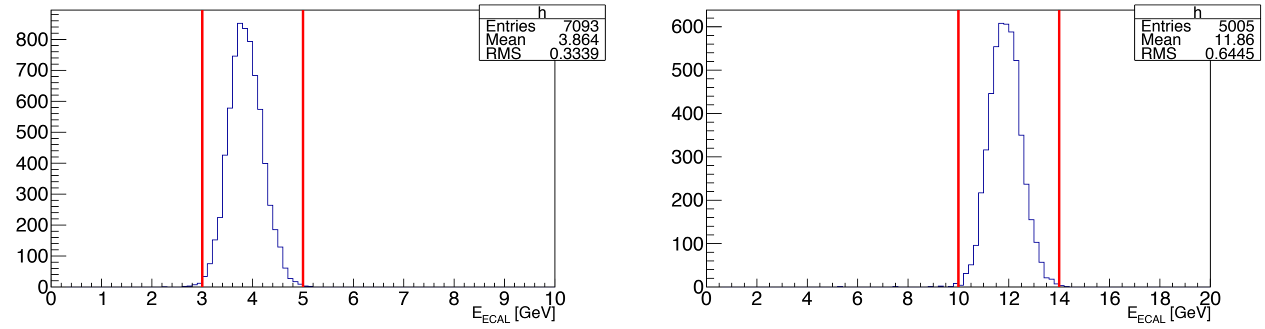
(a) TB 30 GeV π^+ (b) MC 30 GeV π^+

Figure A.24 – Same as in previous Figures A.21 - A.23, for 30 GeV pions (CERN'07). For π^+, the lower cut on the maximal layer energy is set to 0 to keep all MIP-like pions in ECAL.

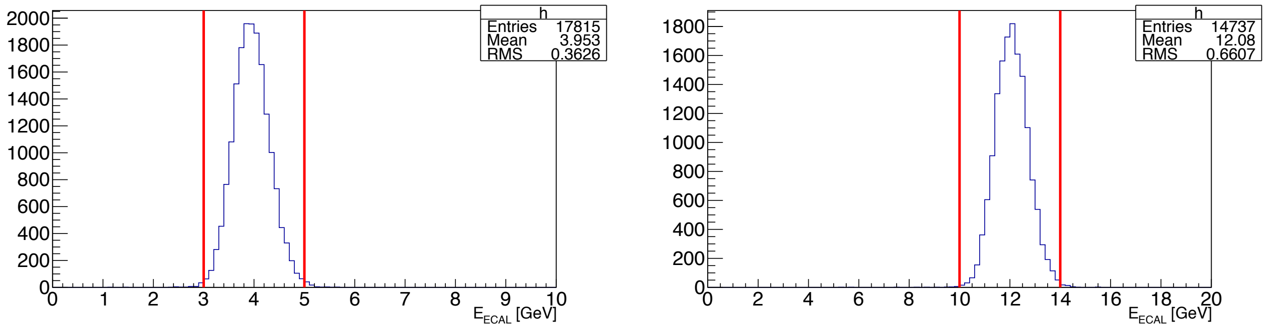
A.1.7 Reconstructed event energy after all cuts except the cut on the event energy itself



(a) TB 4 GeV e^+ (b) TB 12 GeV e^+

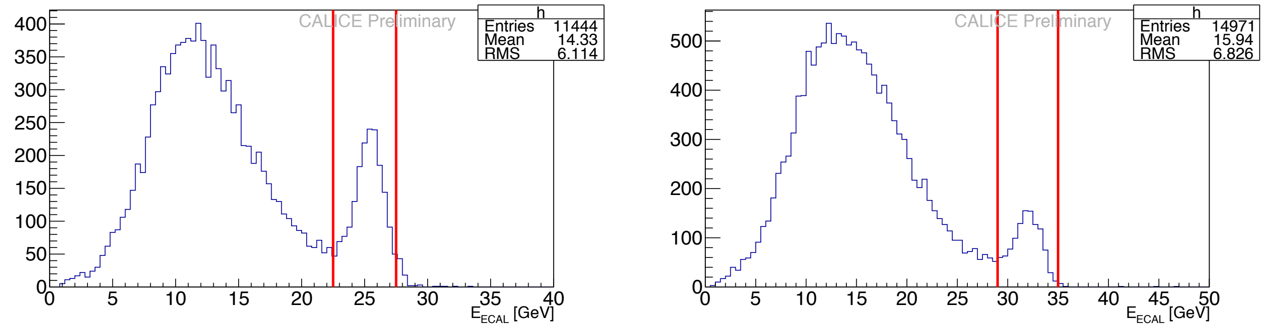


(c) MC 4 GeV e^+ (d) MC 12 GeV e^+

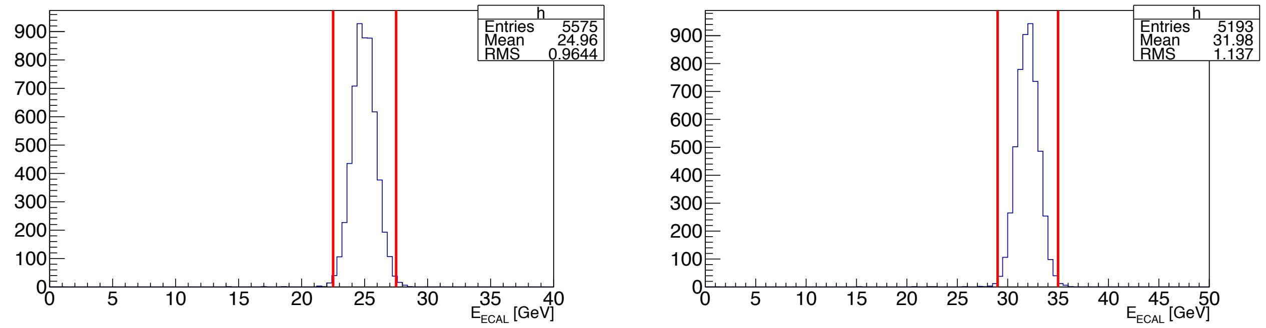


(e) MC 4 GeV γ (f) MC 12 GeV γ

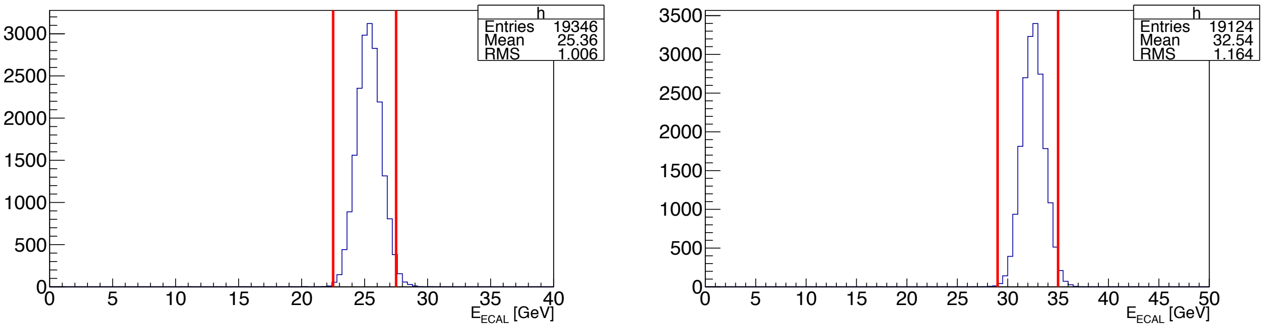
Figure A.25 – Reconstructed energy for events after all cuts except the cut on event energy for 4 and 12 GeV (FNAL'11). For the TB left peaks for 12, 25 and 32 GeV positrons (see also Figure A.26) correspond to pions with shower developed in the ECAL. After the energy cut shown by red lines, this background is heavily suppressed but a little contamination is left.



(a) TB 25 GeV e^+ (b) TB 32 GeV e^+

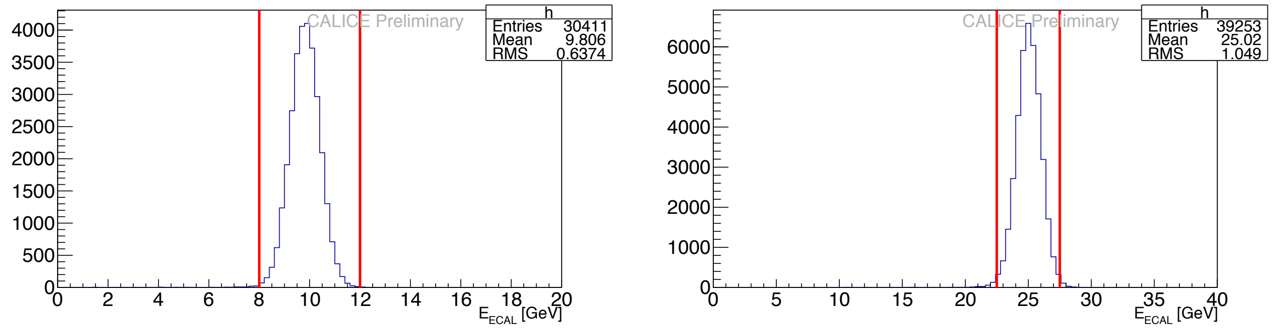


(c) MC 25 GeV e^+ (d) MC 32 GeV e^+

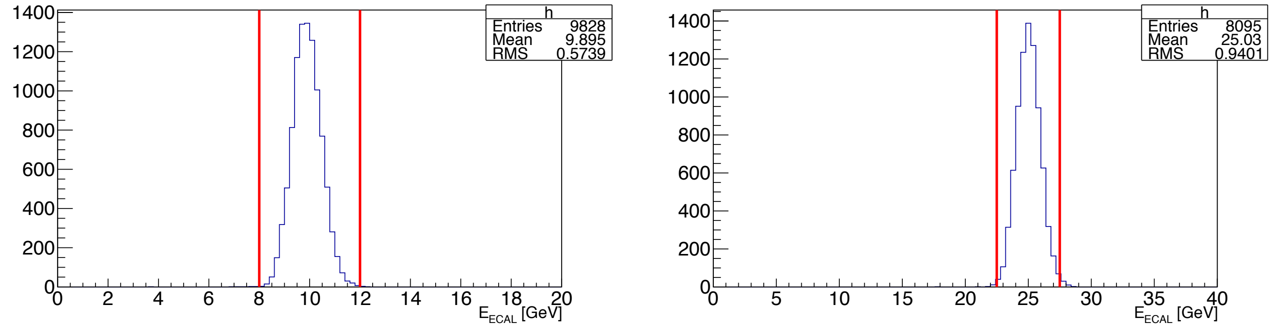


(e) MC 25 GeV γ (f) MC 32 GeV γ

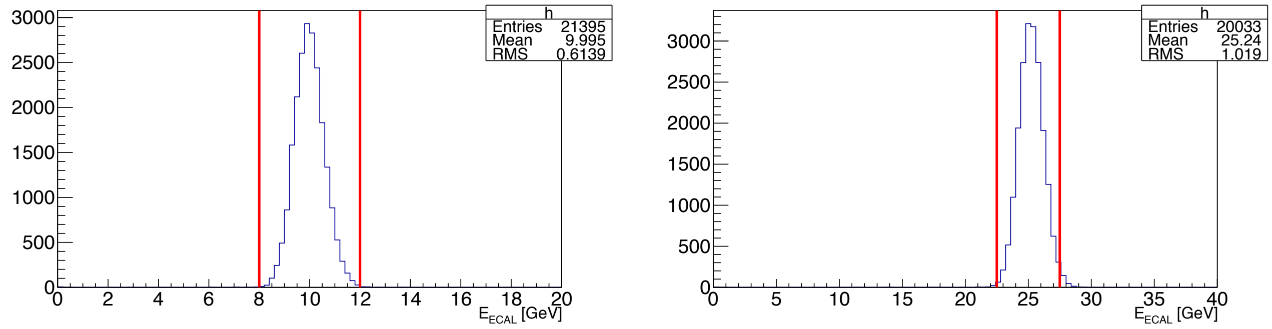
Figure A.26 – Same as in Figure A.25, for 25 and 32 GeV (FNAL'11).



(a) TB 10 GeV e^+ (b) TB 25 GeV e^+

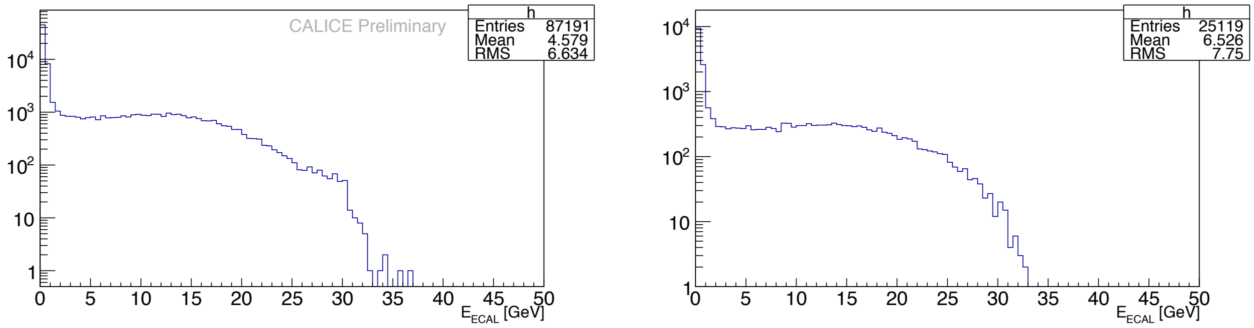


(c) MC 10 GeV e^+ (d) MC 25 GeV e^+



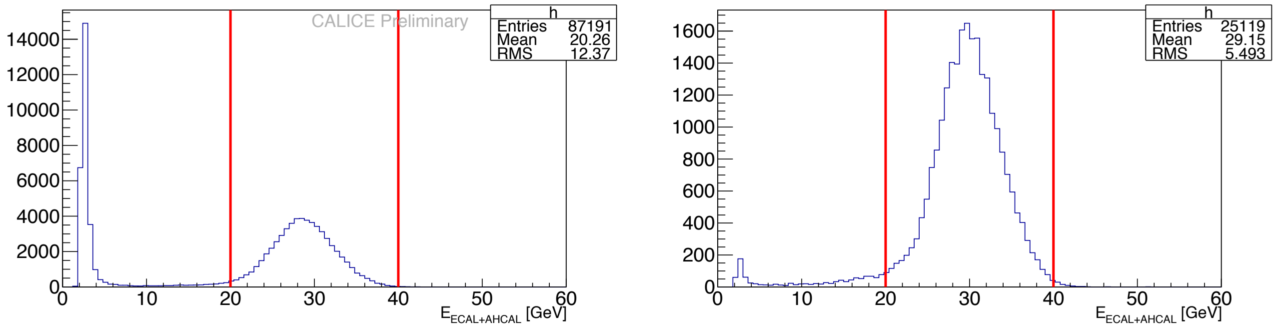
(e) MC 10 GeV γ (f) MC 25 GeV γ

Figure A.27 – Same as in Figures A.25 and A.26, for 10 and 25 GeV (CERN'07). One can see that the final e^+ samples almost do not have pion contamination (contrary to FNAL'11).



(a) TB 30 GeV π^+ ECAL

(b) MC 30 GeV π^+ ECAL



(c) TB 30 GeV π^+ ECAL+AHCAL

(d) MC 30 GeV π^+ ECAL+AHCAL

Figure A.28 – Same as in previous Figures A.25 - A.27, for 30 GeV pions (CERN'07). Left peak in energy distribution for TB corresponds to muons. Small e^+ contamination is present in the final TB π^+ sample.

A.2 Final plots for all energy pairs

Appendix A.2 contains results on particle separation for all studied pairs, see Table A.2. All analysis details could be found in sections 4.4 and 4.5.

2-particle	Energy, GeV	Geometry: granularity	PFA
e^+-e^+ (γ-γ)	4+4 12+4 25+4 32+25	TB FNAL'11: 10 × 10 mm^2 ECAL	Pandora (v00-14 & v02-04) Garlic (v2.11) Arbor (March'15)
γ-γ	4+4 12+4 25+4 32+25	ILD: 5 × 5 or 2.5 × 2.5 mm^2 ECAL	Pandora (v00-14 & v02-04) Garlic (v2.11 & v3.0.3) Arbor (March'15)
π^+-e^+(γ)	30+10 30+25	TB CERN'07: 10 × 10 mm^2 ECAL, 30 × 30 mm^2 AHCAL	Pandora (v00-14 & v02-04) Garlic (v2.11)
π^+-γ	20+6 20+25 30+6 30+25	ILD: 5 × 5 or 2.5 × 2.5 mm^2 ECAL, 30 × 30 mm^2 AHCAL or 10 × 10 mm^2 SDHCAL	Pandora (v00-14 & v02-04) Garlic (v2.11 & v3.0.3) Arbor (March'15)

Table A.2 – Final list of all studied energy pairs.

A.2.1 Separation of two electromagnetic clusters with FNAL’11 data

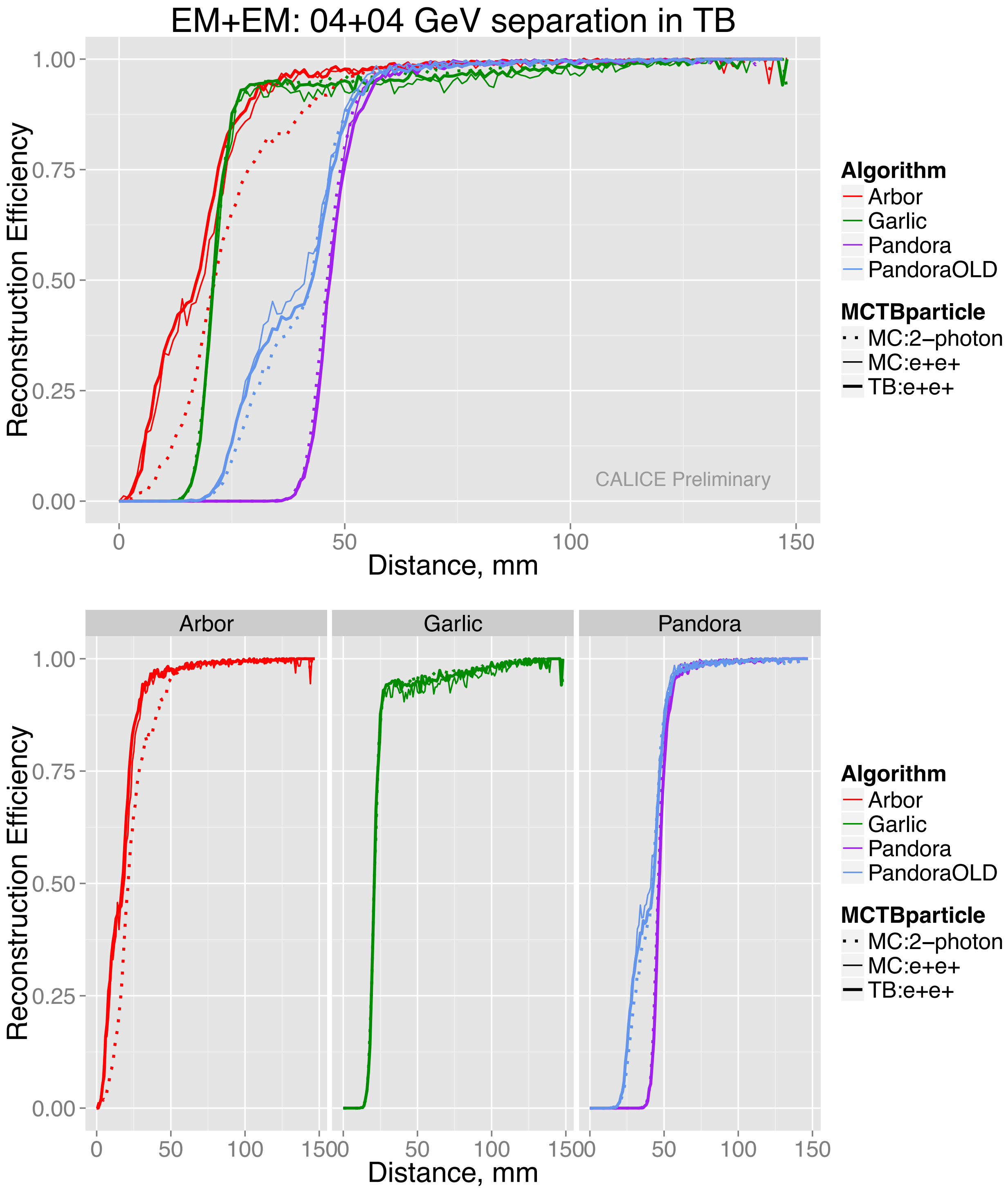


Figure A.29 – Efficiency of 4+4 GeV 2-EM clusters reconstruction versus the distance between showers (FNAL’11). For convenience all efficiency curves are mixed together on the top plot and shown individually for every PFA on the bottom row plots.

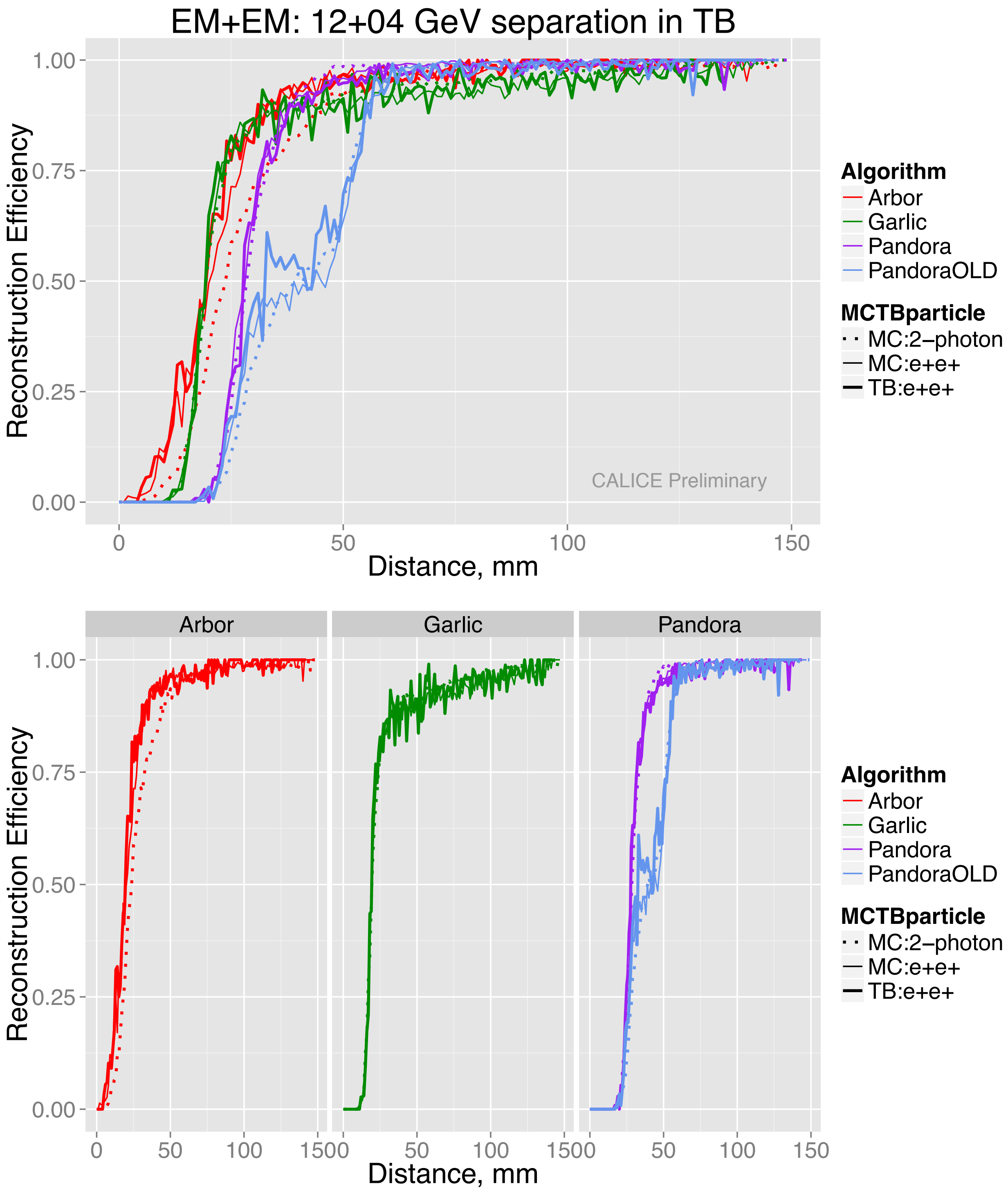


Figure A.30 – Same as in Figure A.29 but for 12+4 GeV energy pair.

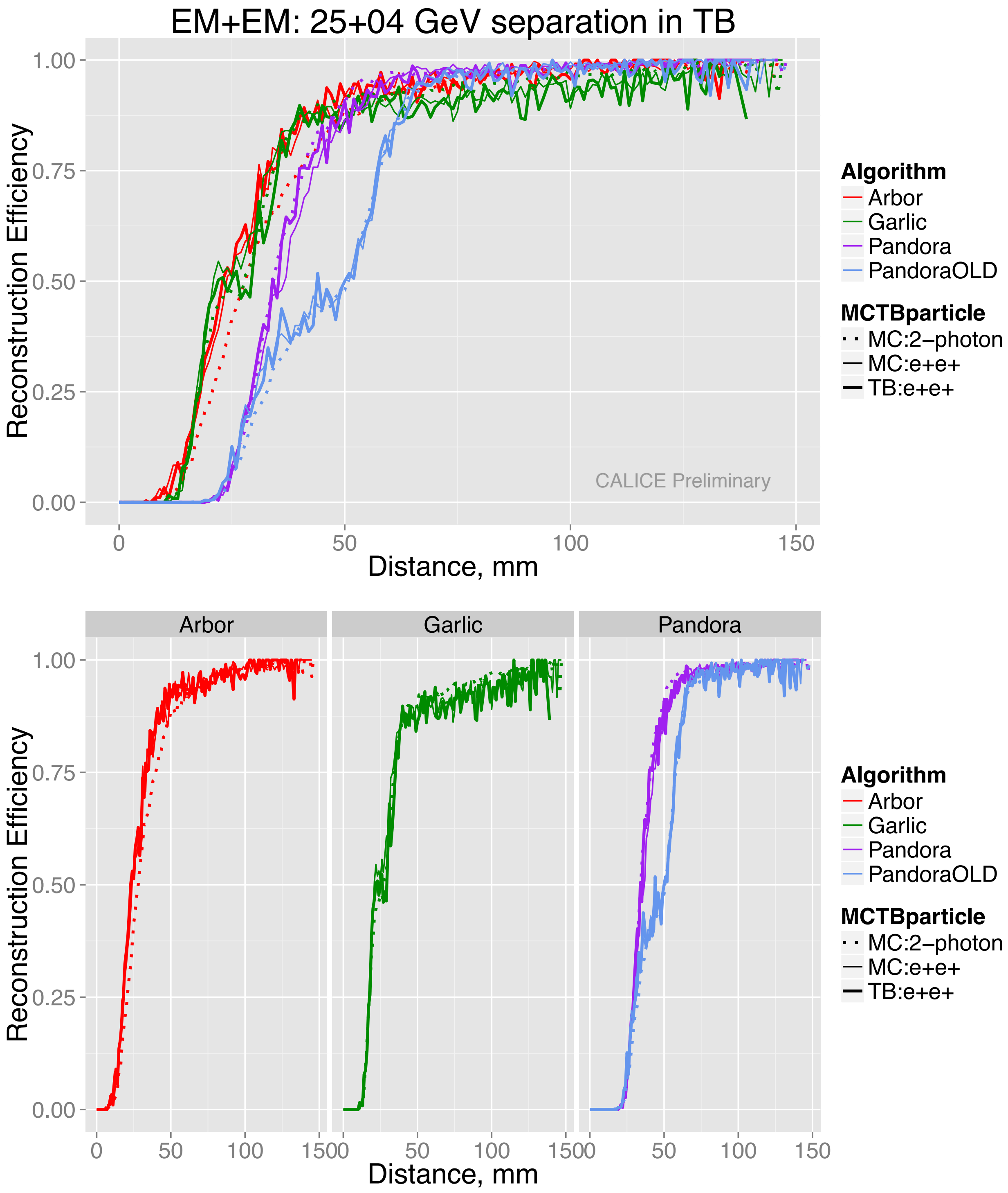


Figure A.31 – Same as in Figure A.29 but for 25+4 GeV energy pair.

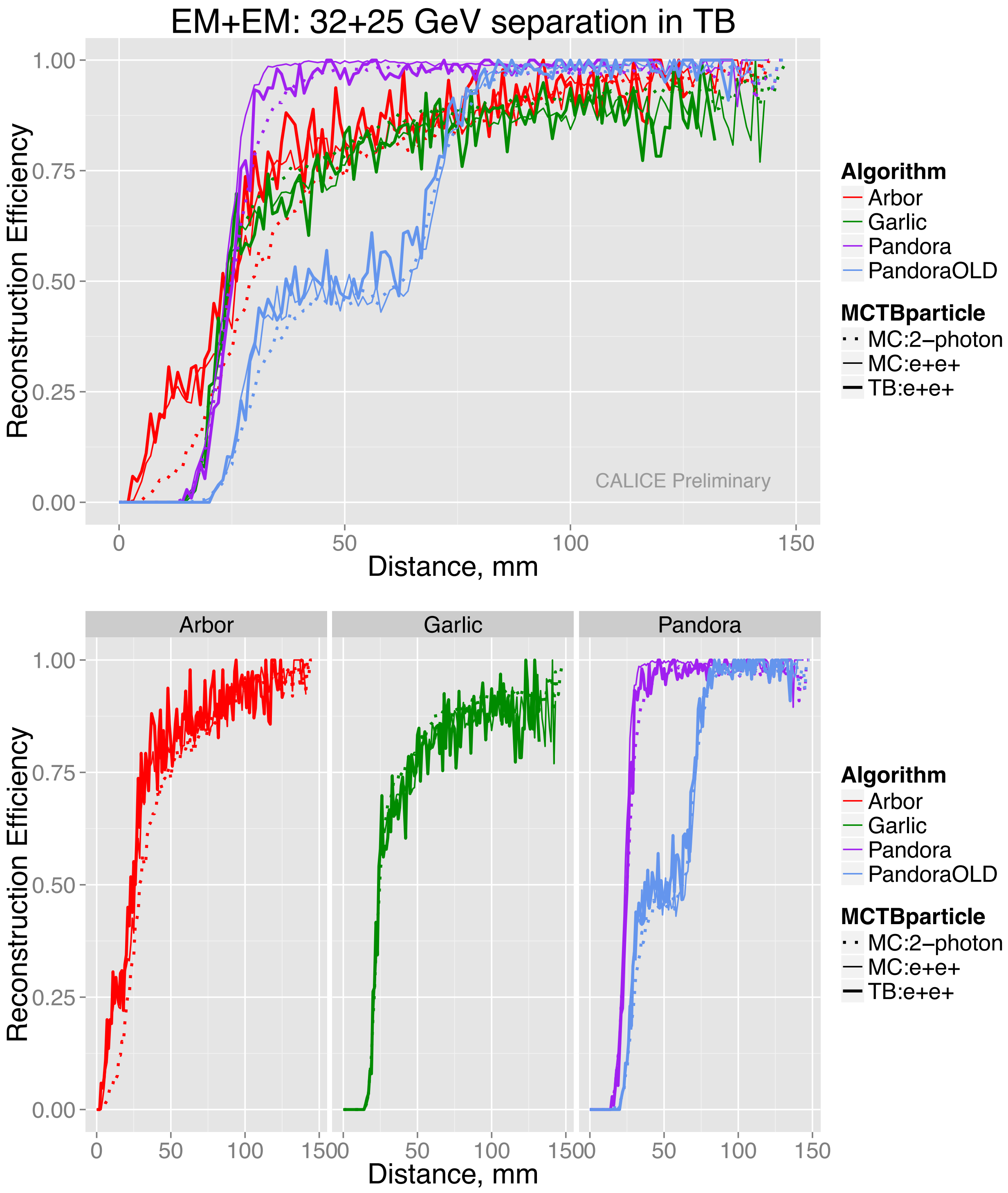


Figure A.32 – Same as in Figure A.29 but for 32+25 GeV energy pair.

A.2.2 Two photon separation in ILD

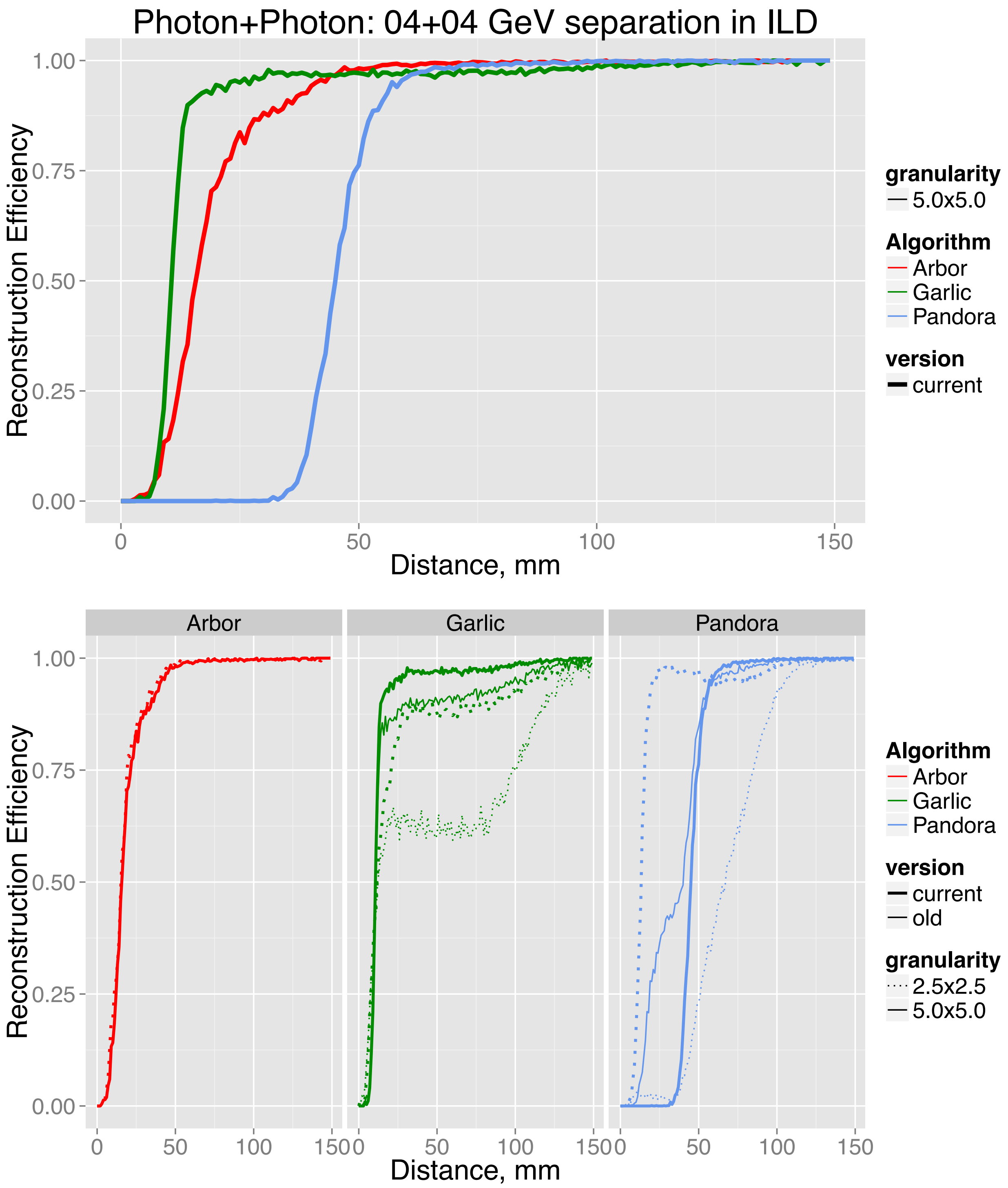


Figure A.33 – Efficiency of 4+4 GeV 2-γ reconstruction versus the distance between showers (ILD). On the top plot the reconstruction efficiencies of the new versions of PFA algorithms for the standard 5×5 mm^2 ECAL cells are shown. Bottom row plots have more information, old Pandora, Garlic and finer 2.5×2.5 mm^2 ECAL granularity plots are added.

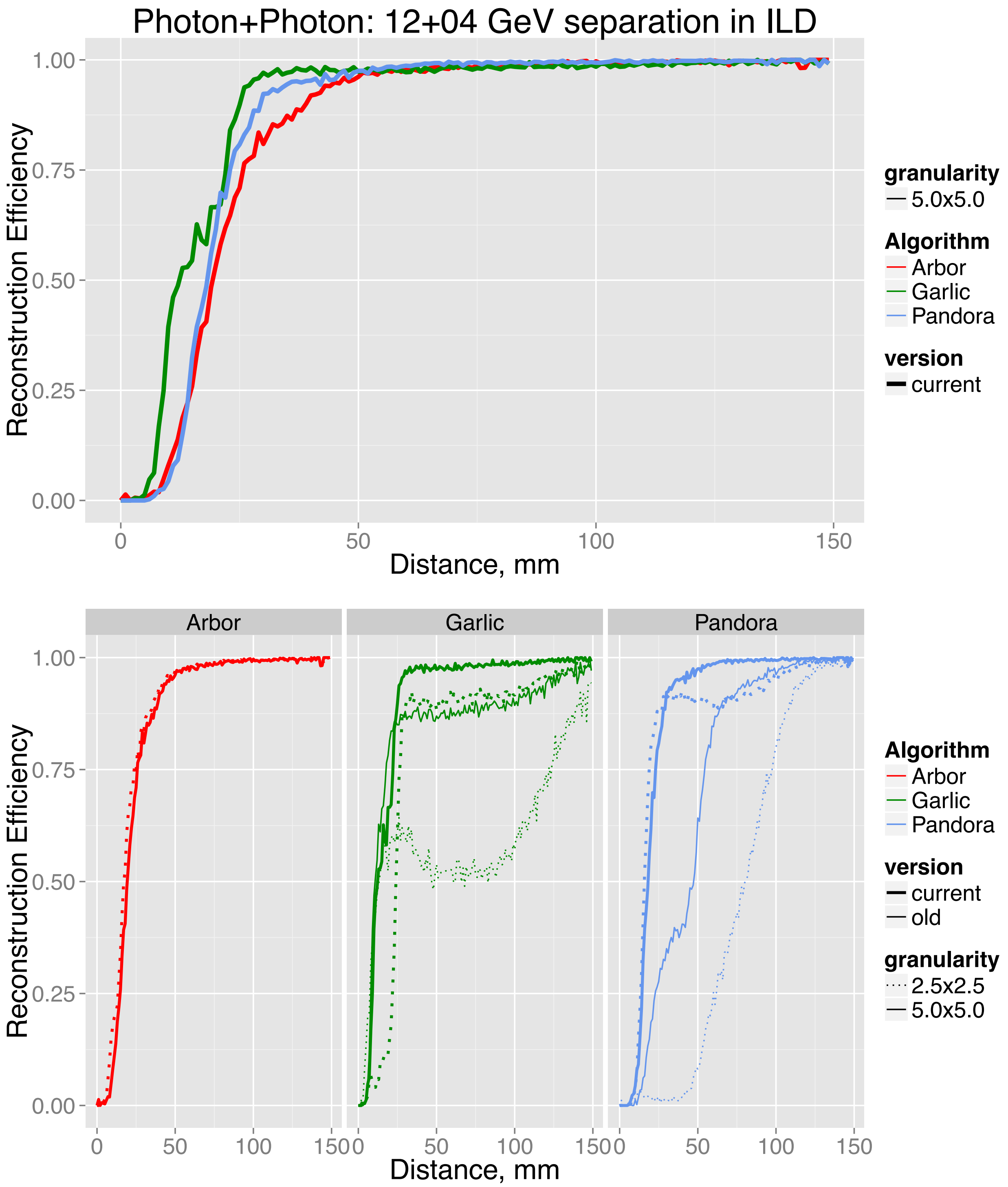


Figure A.34 – Same as in Figure A.33 but for 12+4 GeV energy pair.

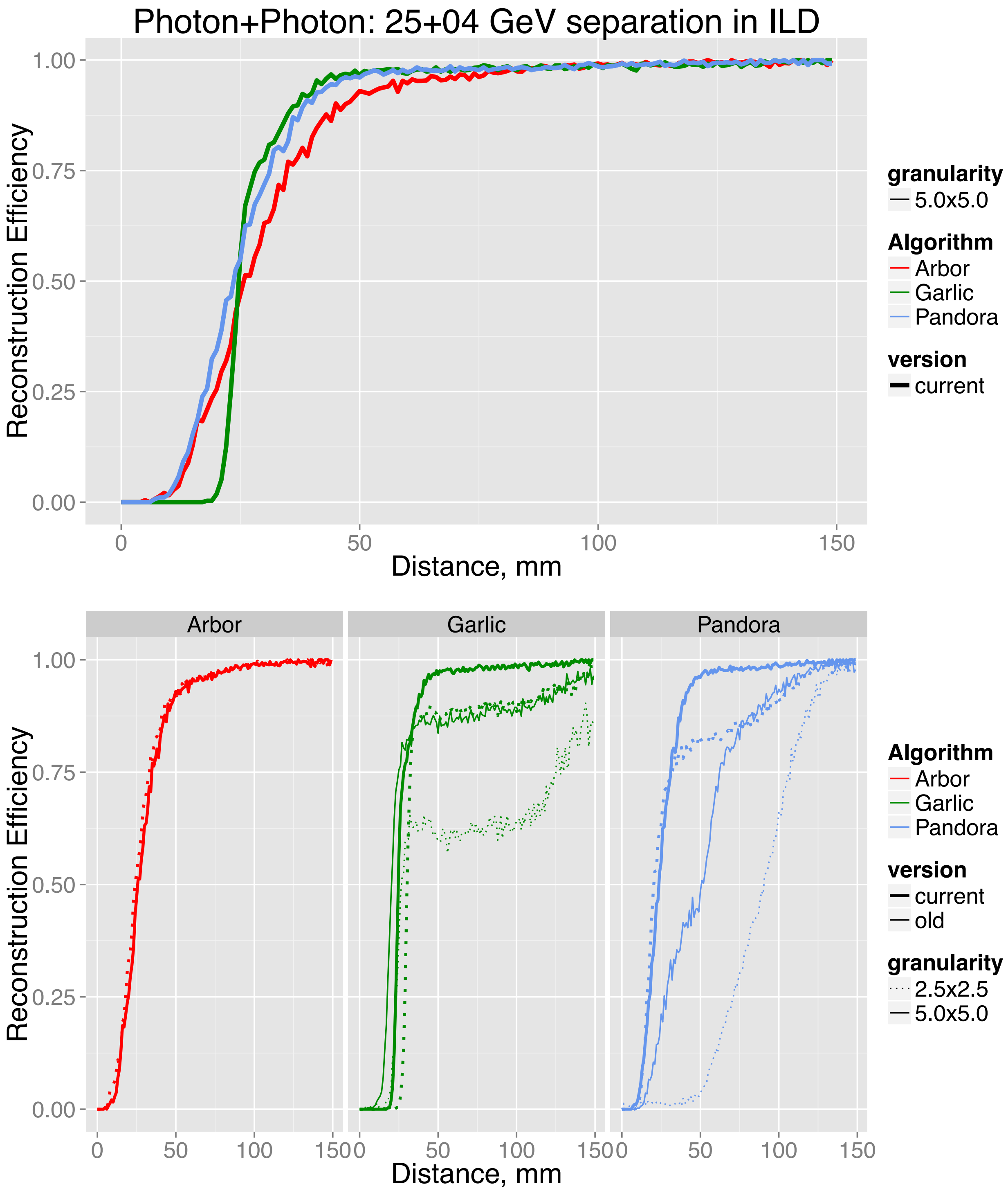


Figure A.35 – Same as in Figure A.33 but for 25+4 GeV energy pair.

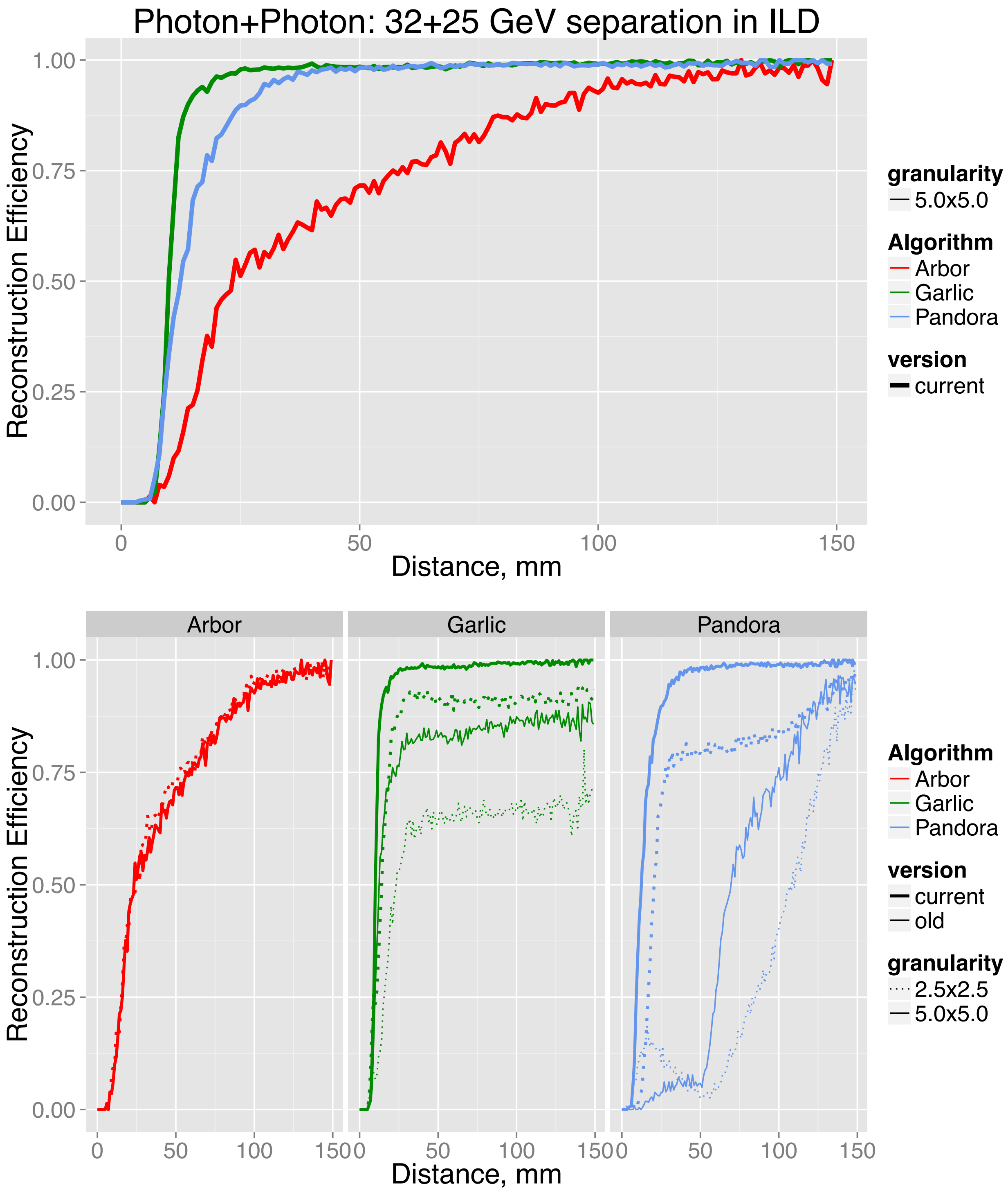


Figure A.36 – Same as in Figure A.33 but for 32+25 GeV energy pair.

A.2.3 Classification of inefficiencies in two-photon reconstruction

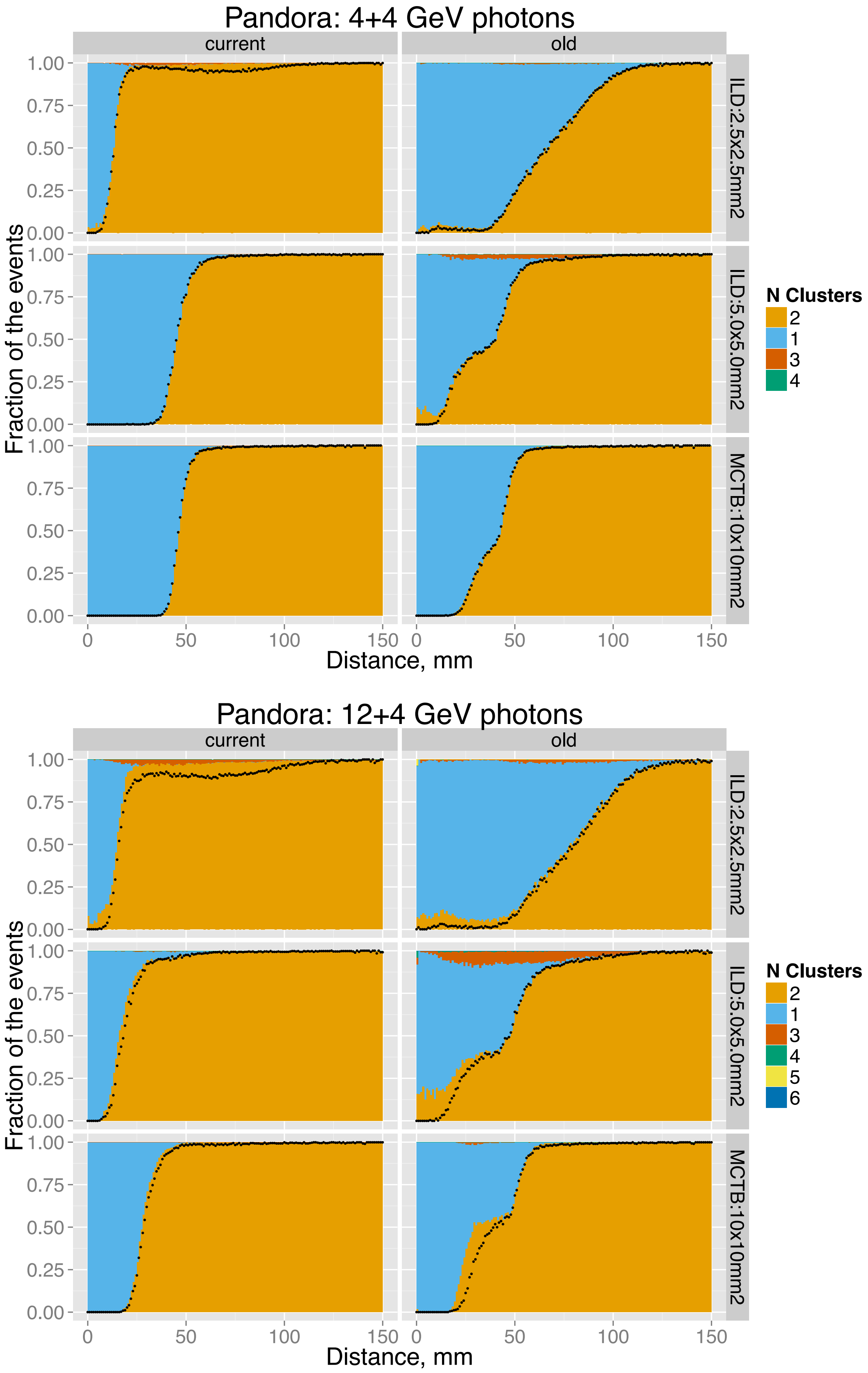


Figure A.37 – Pandora reconstruction of 4+4 (top), 12+4 (bottom) GeV two photon events: the colored bands show the fraction of events where 1, 2, 3, … neutral clusters are reconstructed, regardless of their energies and positions. The black points show the efficiency of reconstruction of exactly two clusters with the energies and the positions within ±20% and ±5 mm, respectively, from their values in single shower events, i.e. as in Figs. A.29, A.30, A.33 and A.34.

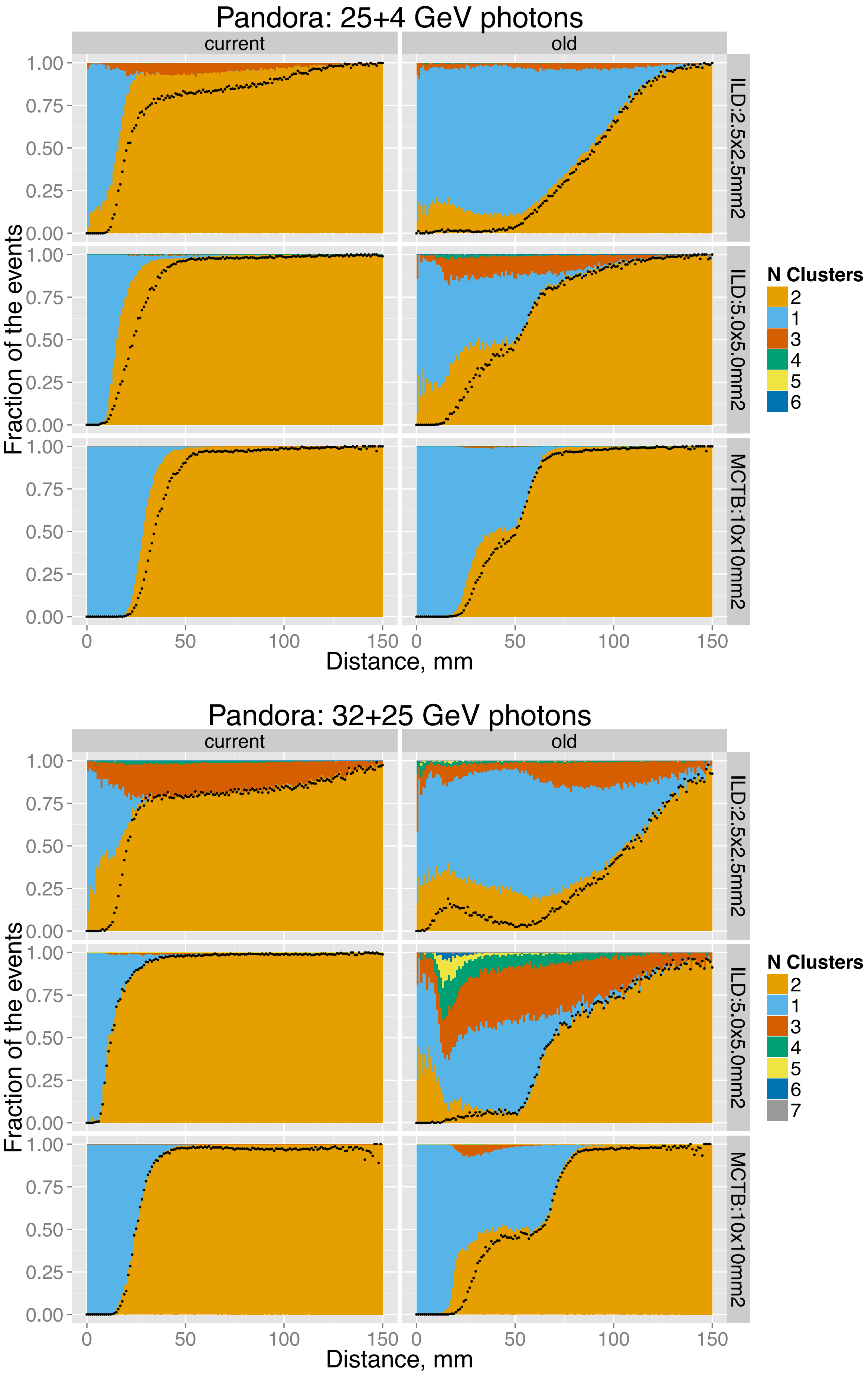


Figure A.38 – Same as in Figure A.37 but for 25+4 (top), 32+25 (bottom) GeV energy pairs.

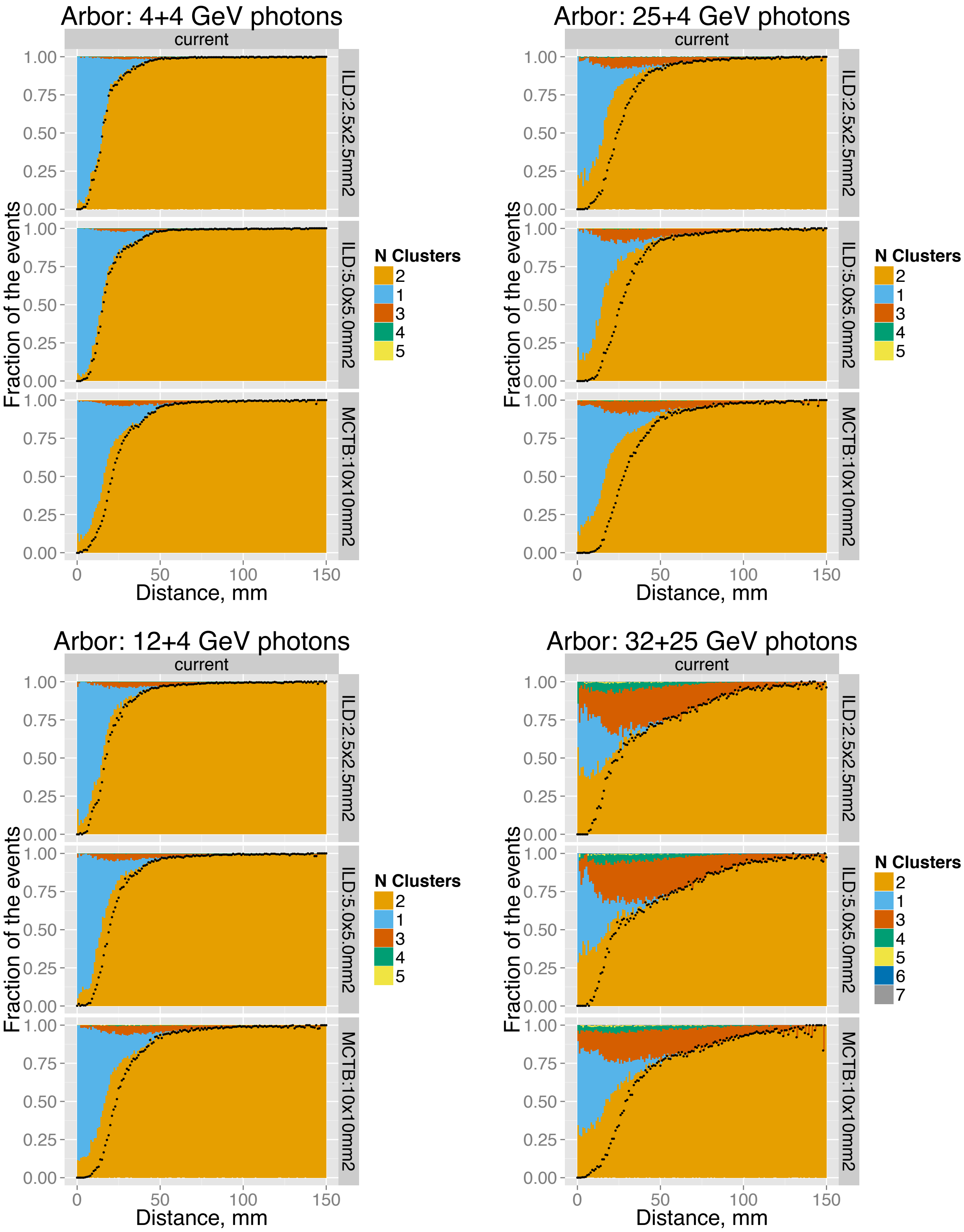
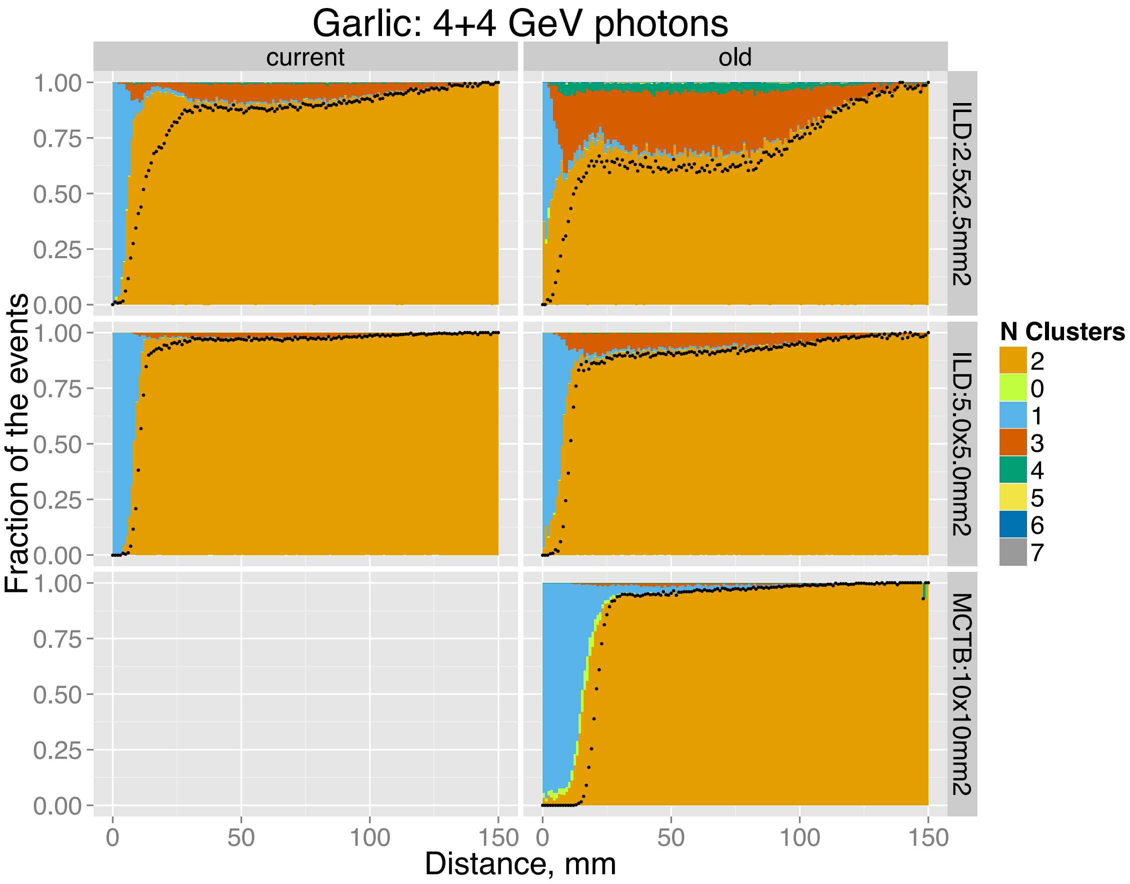


Figure A.39 – Arbor reconstruction of two photon events: the colored bands show the fraction of events where 1, 2, 3, … neutral clusters are reconstructed, regardless of their energies and positions. The black points show the efficiency of reconstruction of exactly two clusters with the energies and the positions within ±20% and ±5 mm, respectively, from their values in single shower events.



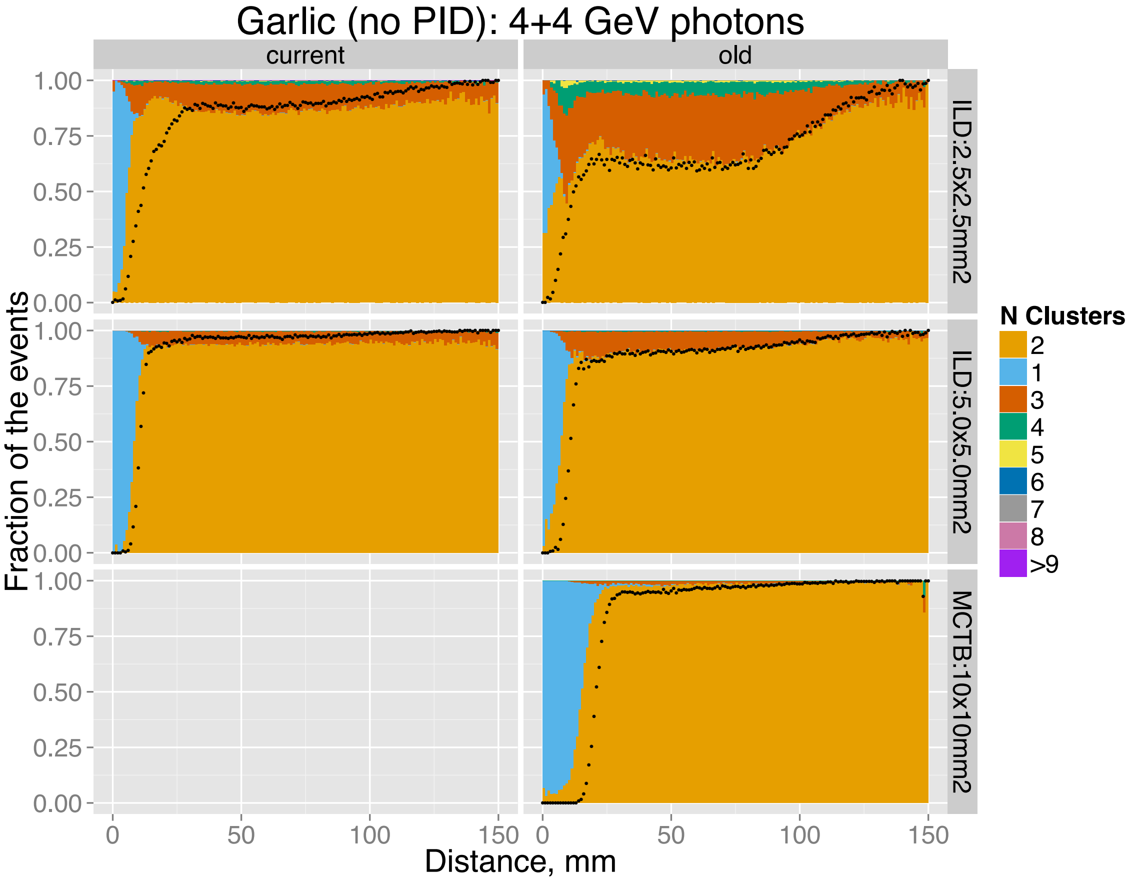
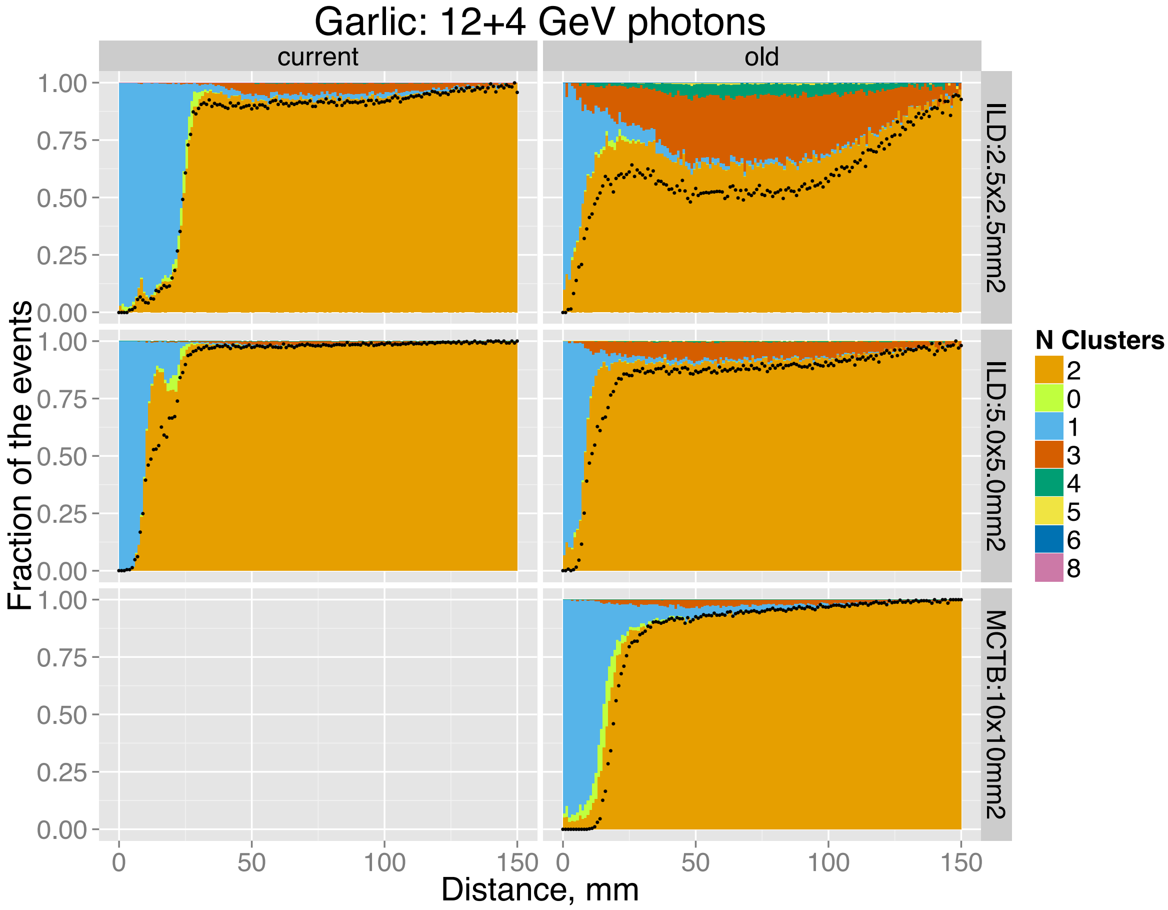


Figure A.40 – Garlic reconstruction of 4+4 GeV two photon events: the colored bands show the fraction of events where 1, 2, 3, . . . photons (top) or neutral cluster candidates before the photon identification (bottom) are reconstructed, regardless of their energies and positions. The black points show the efficiency of reconstruction of exactly two photons (after the photon reidentification) with the energies and the positions within ±20% and ±5 mm, respectively, from their values in single shower events, i.e. as in Figs. A.29 and A.33.



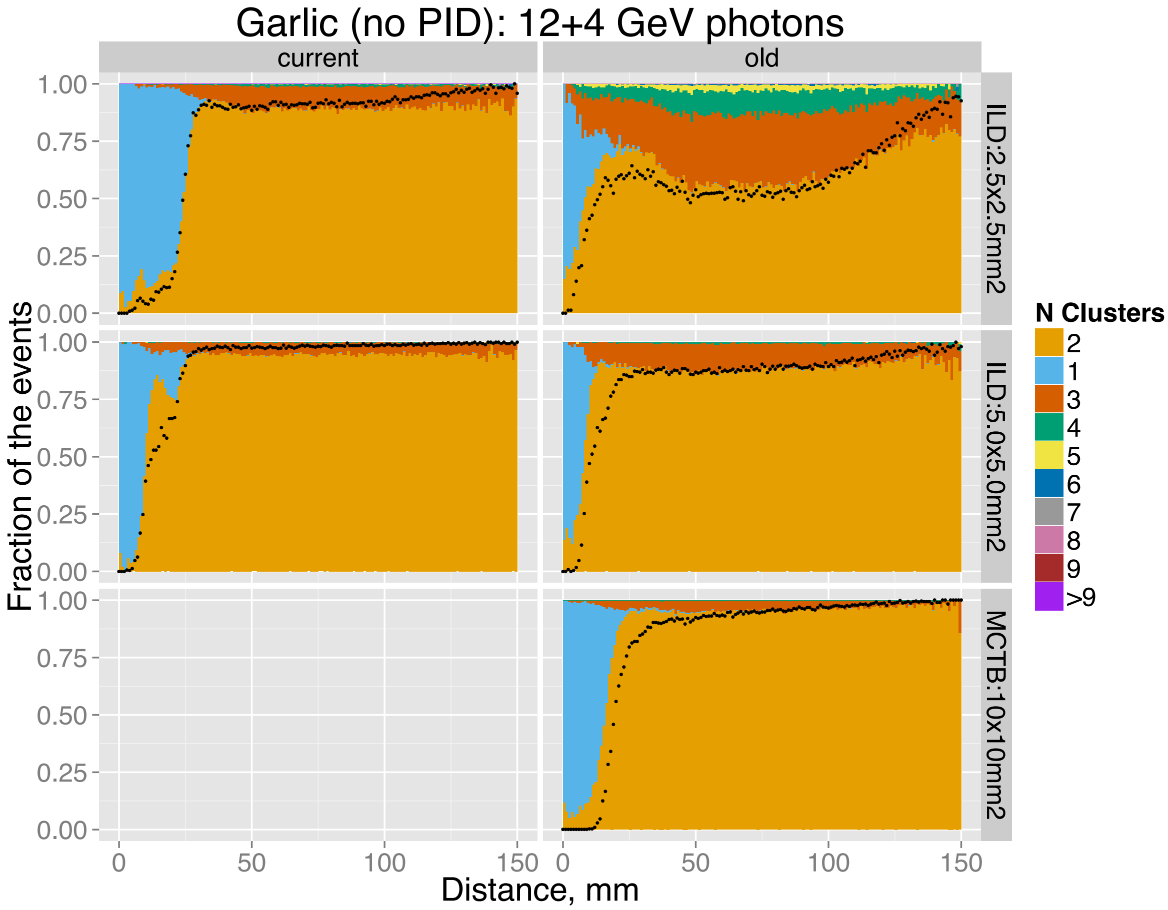


Figure A.41 – Same as in Figure A.40 but for 12+4 GeV energy pair.

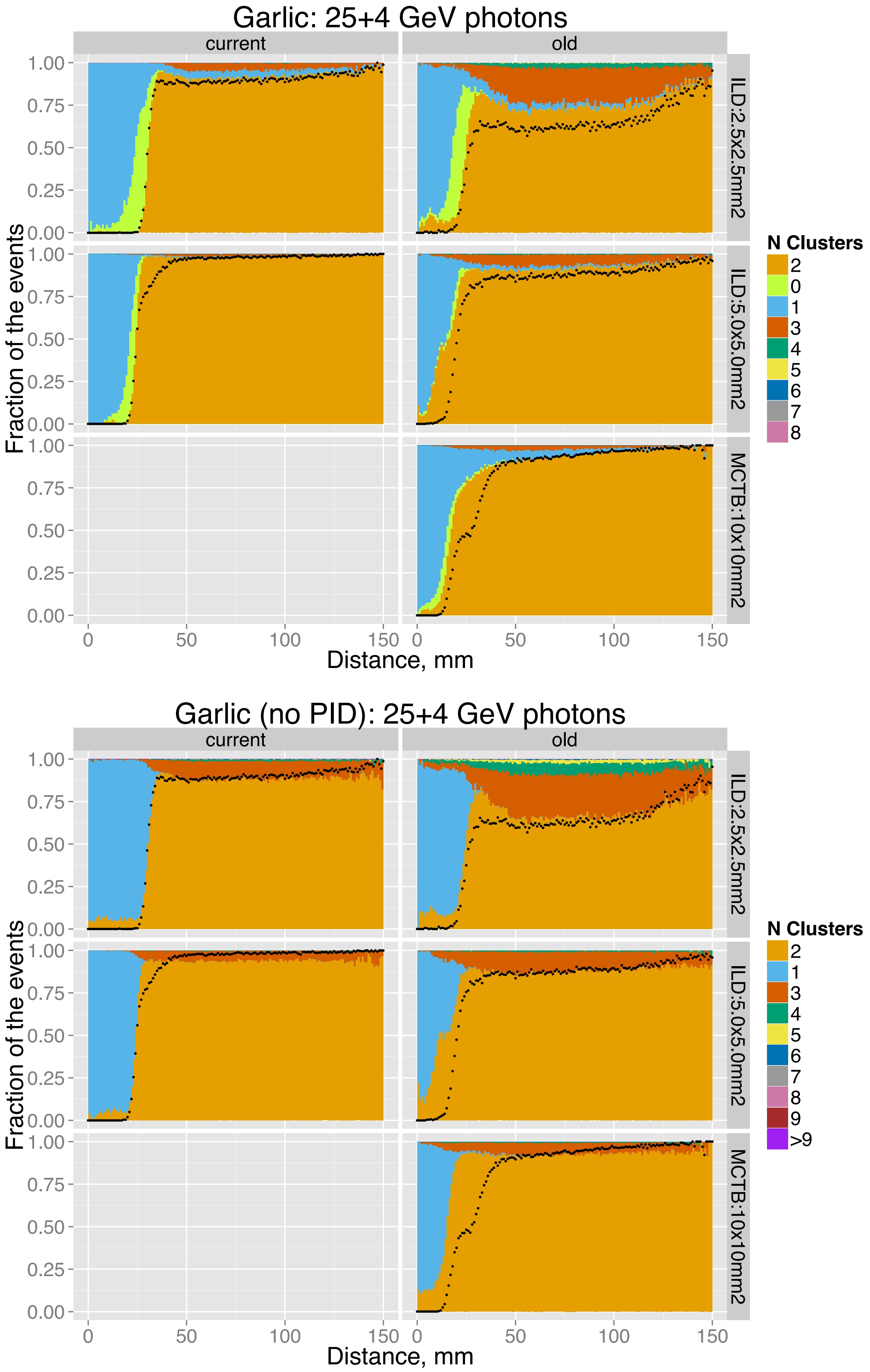


Figure A.42 – Same as in Figure A.40 but for 25+4 GeV energy pair.

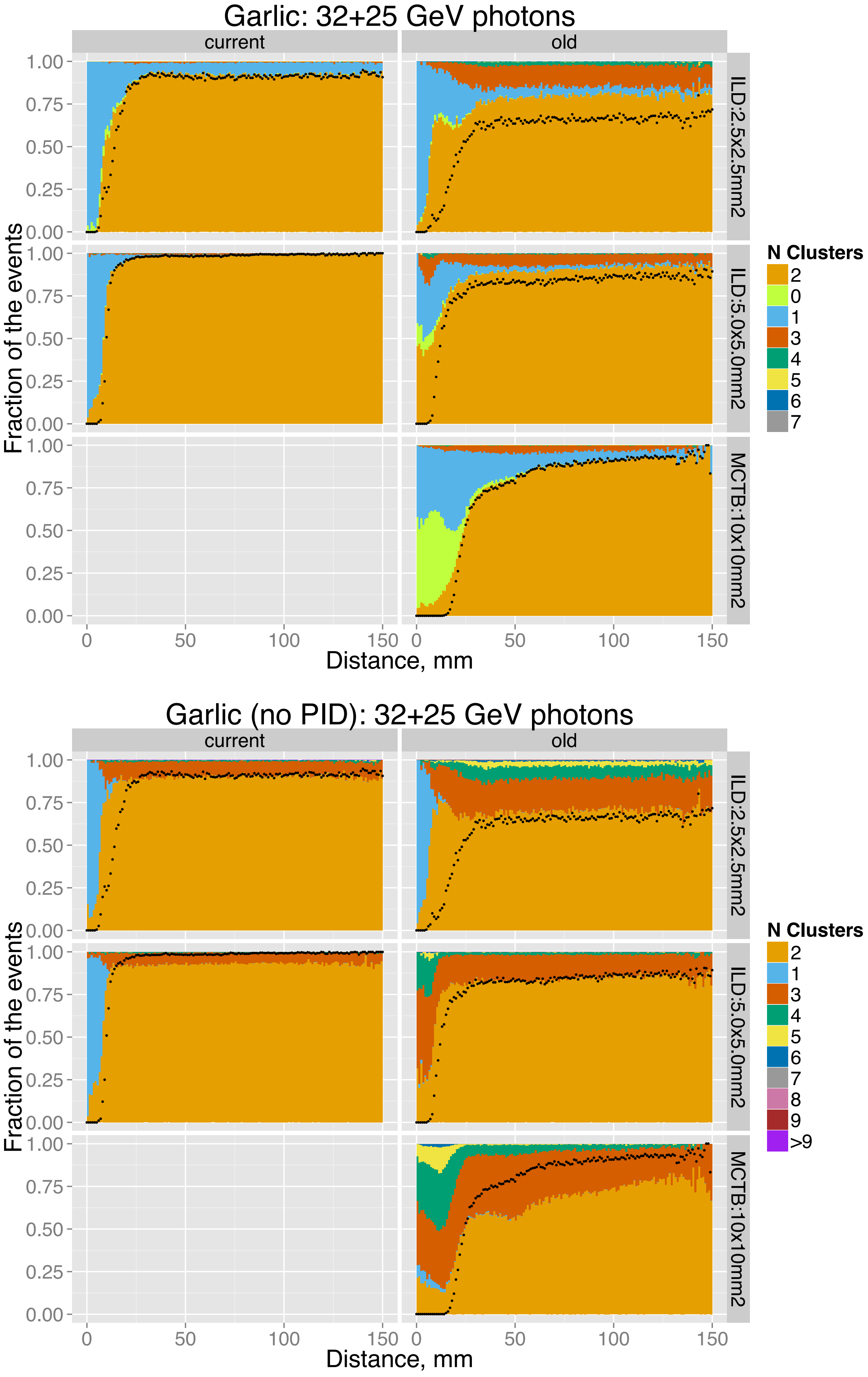


Figure A.43 – Same as in Figure A.40 but for 32+25 GeV energy pair.

A.2.4 Pion-EM separation with CERN'07 data

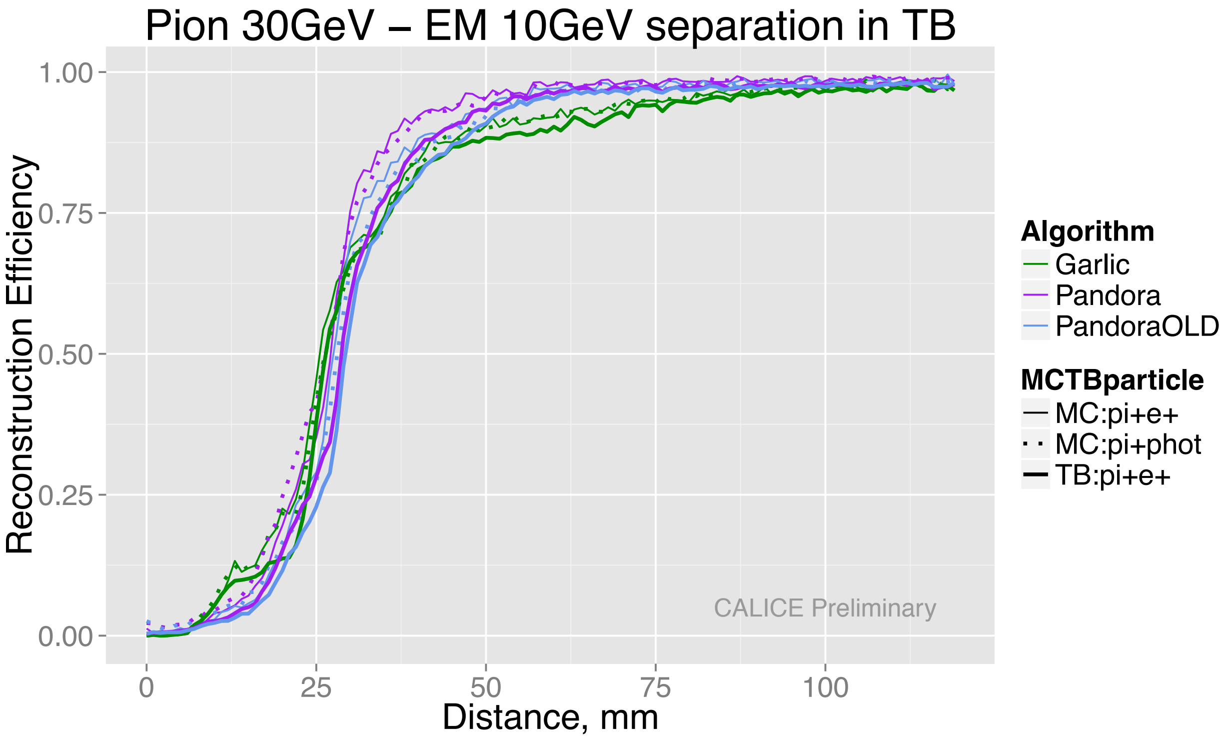


Figure A.44 – Efficiency of 30+10 GeV Pion-EM clusters (ECAL+AHCAL) reconstruction versus the distance between showers (CERN'07).

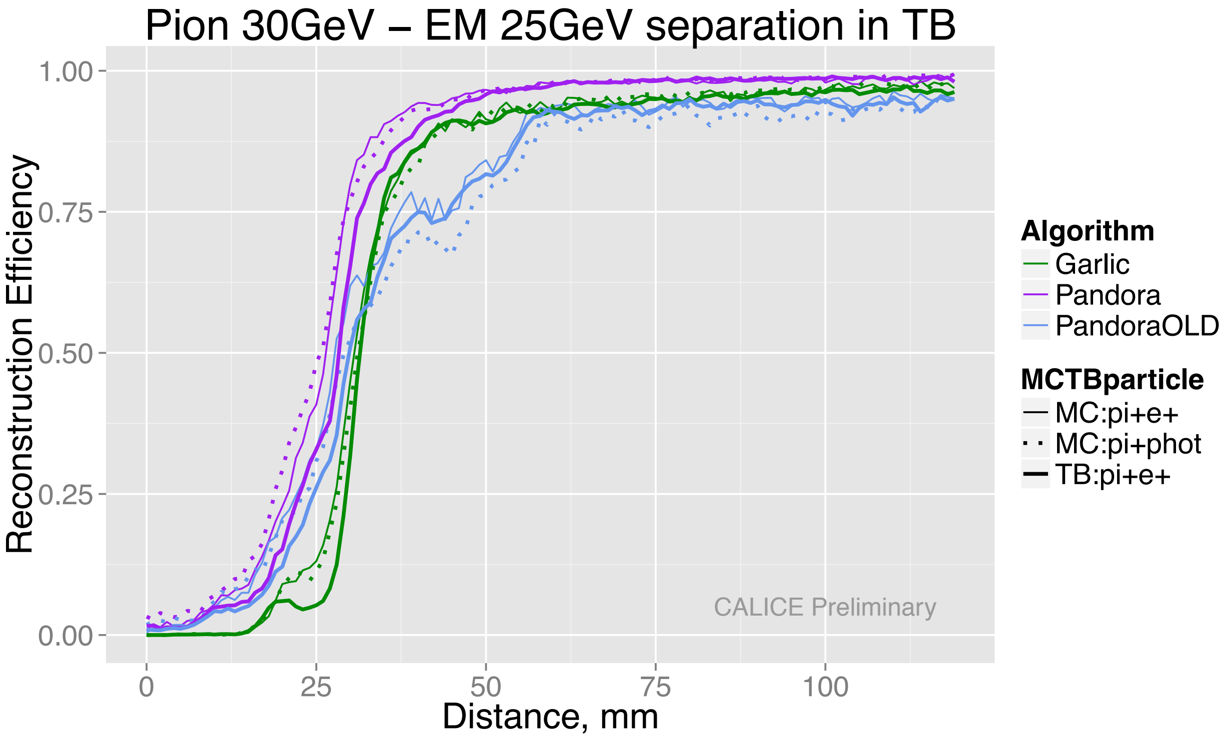


Figure A.45 – Same as in Figure A.44 but for 30+25 GeV energy pair.

A.2.5 Pion-photon separation in ILD

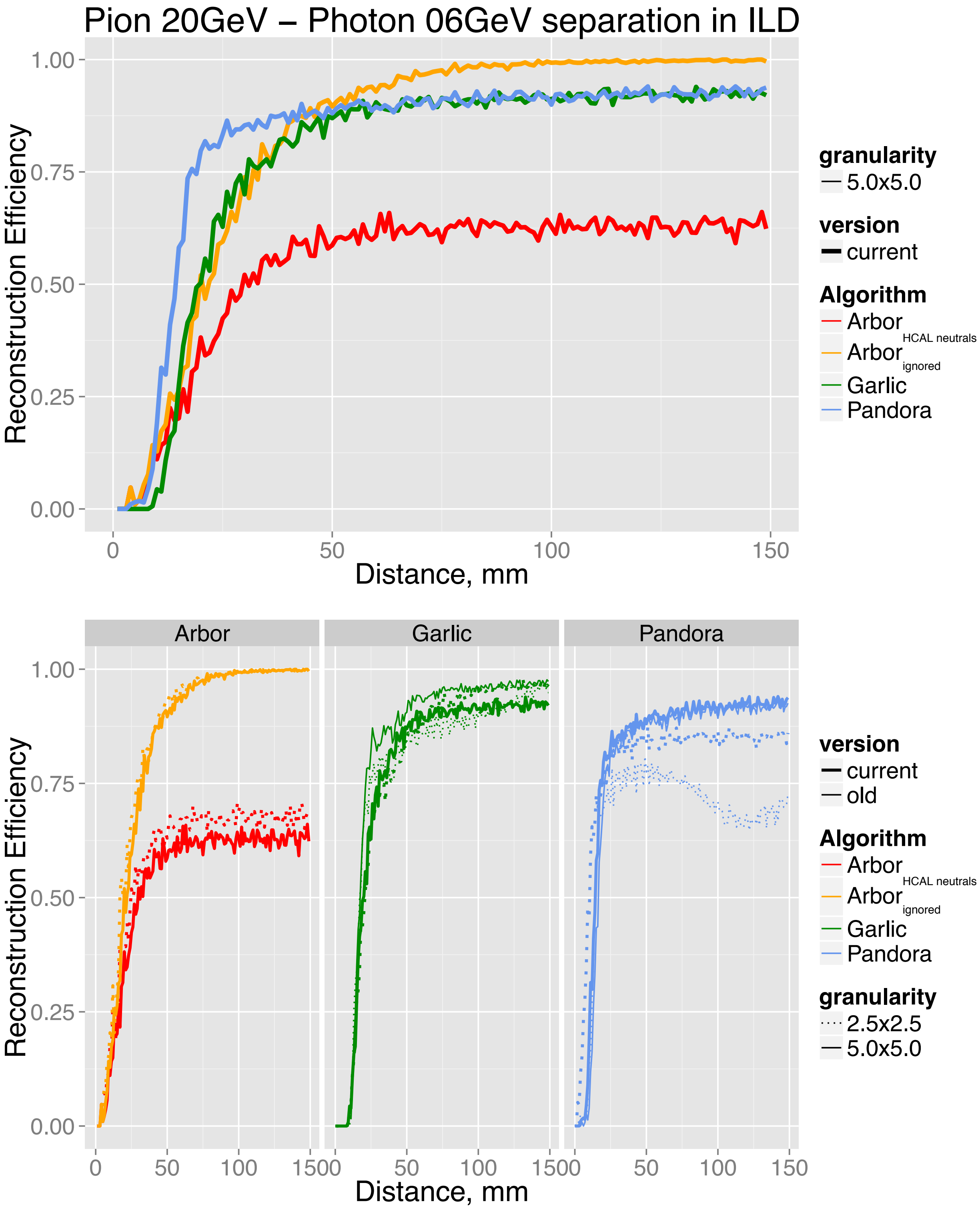


Figure A.46 – Efficiency of 20+6 GeV Pion-photon clusters reconstruction versus the distance between showers (ILD, Pandora - AHCAL, Arbor - SDHCAL, Garlic - no HCAL).

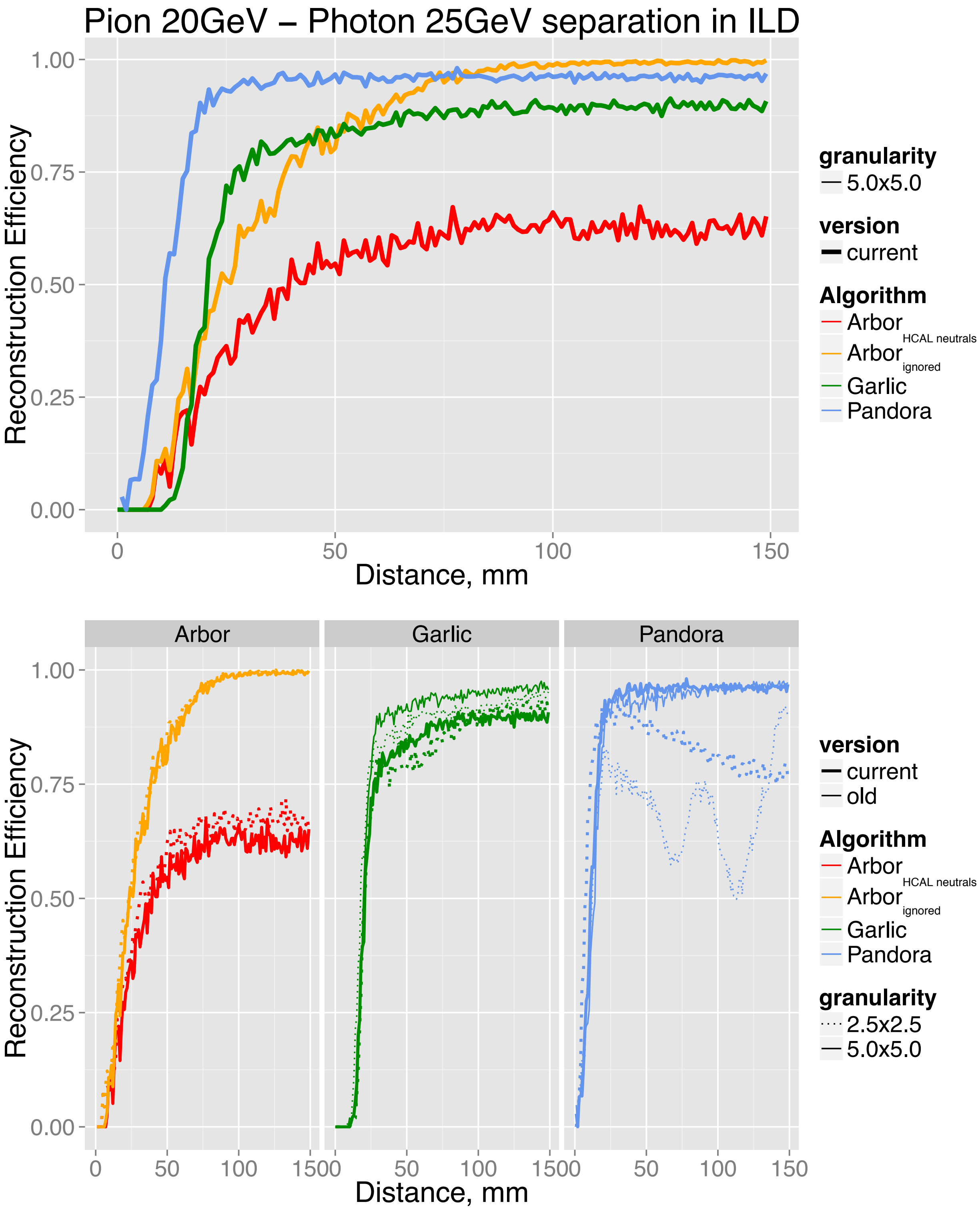


Figure A.47 – Same as in Figure A.46 but for 20+25 GeV energy pair.

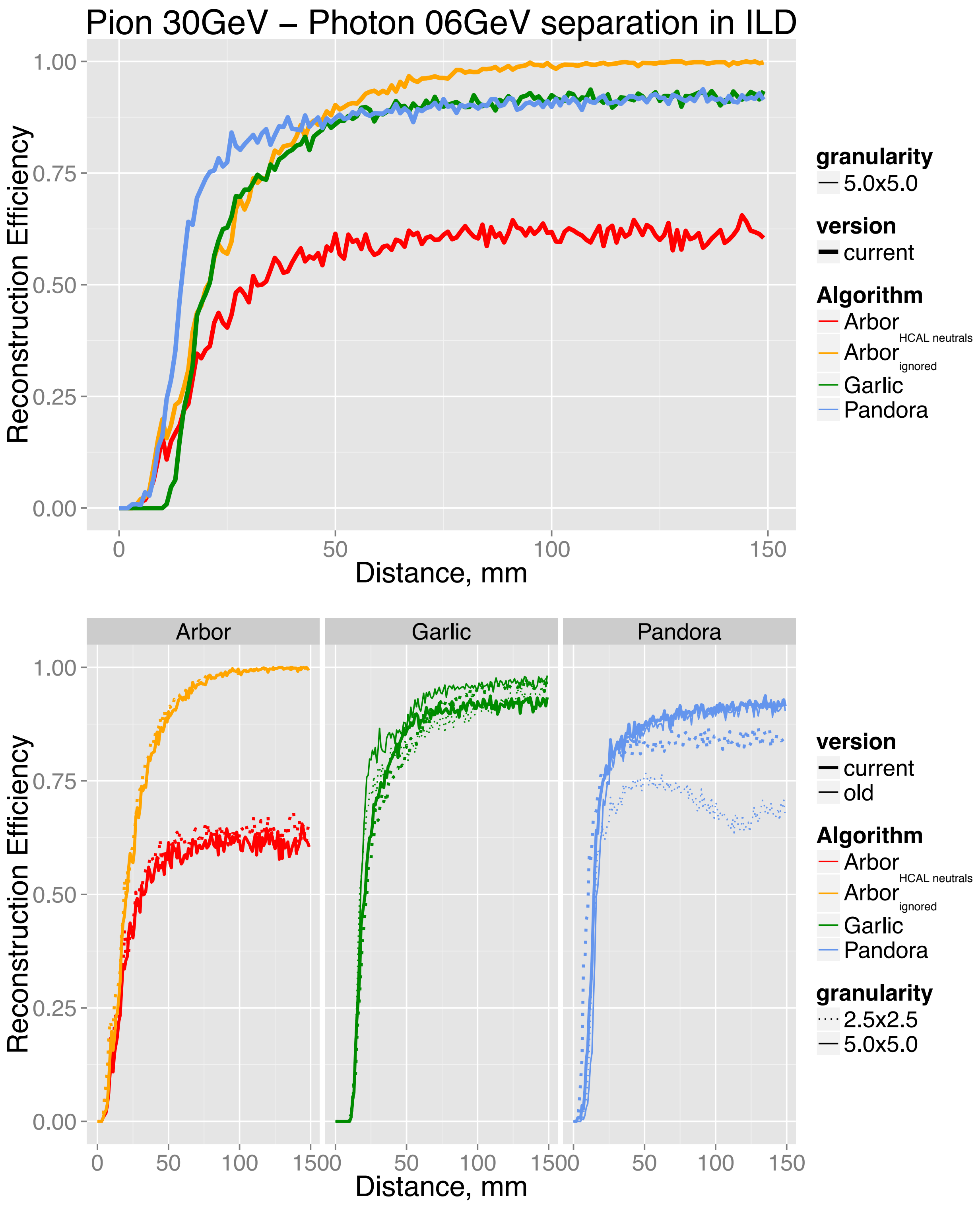


Figure A.48 – Same as in Figure A.46 but for 30+6 GeV energy pair.

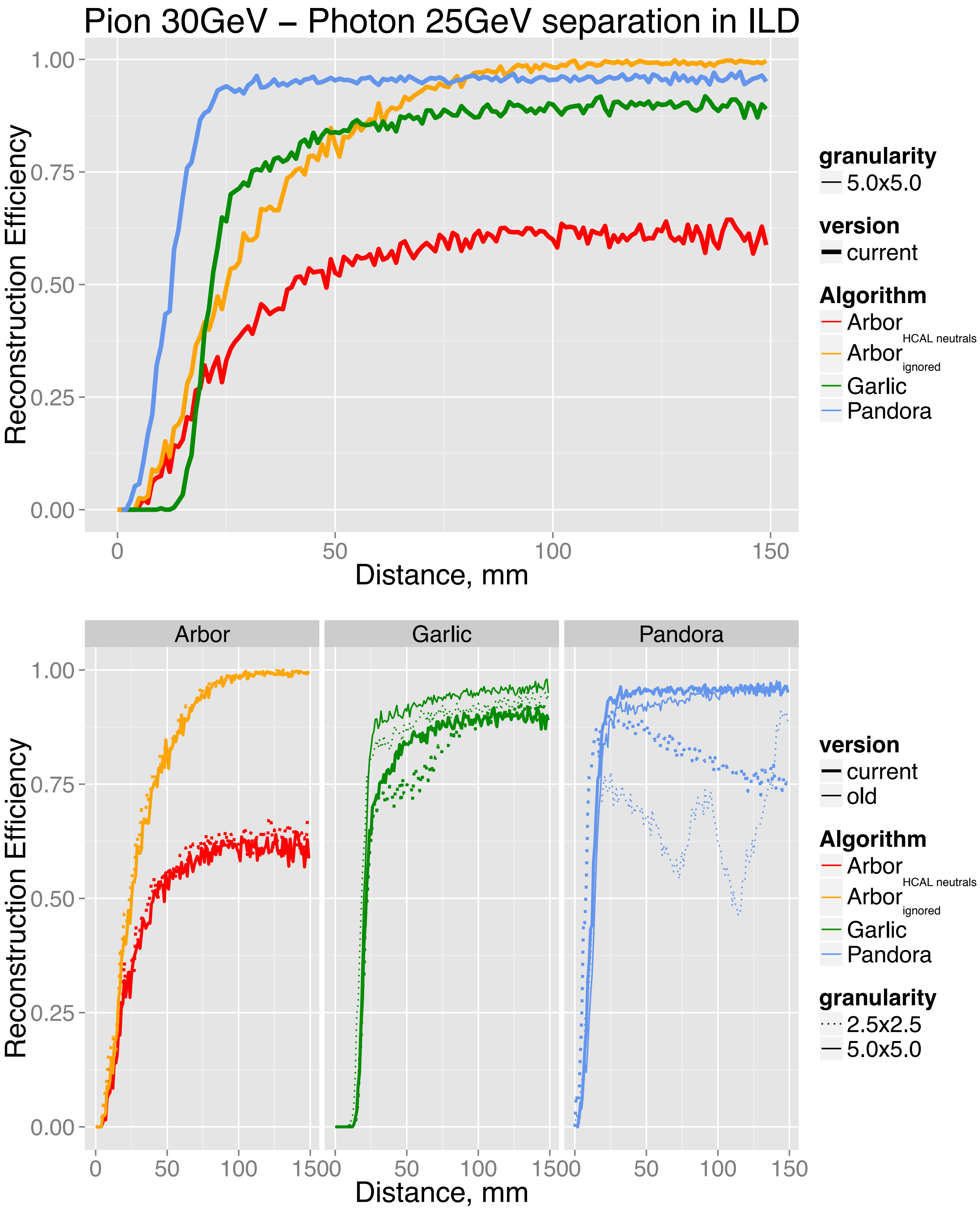


Figure A.49 – Same as in Figure A.46 but for 30+25 GeV energy pair.

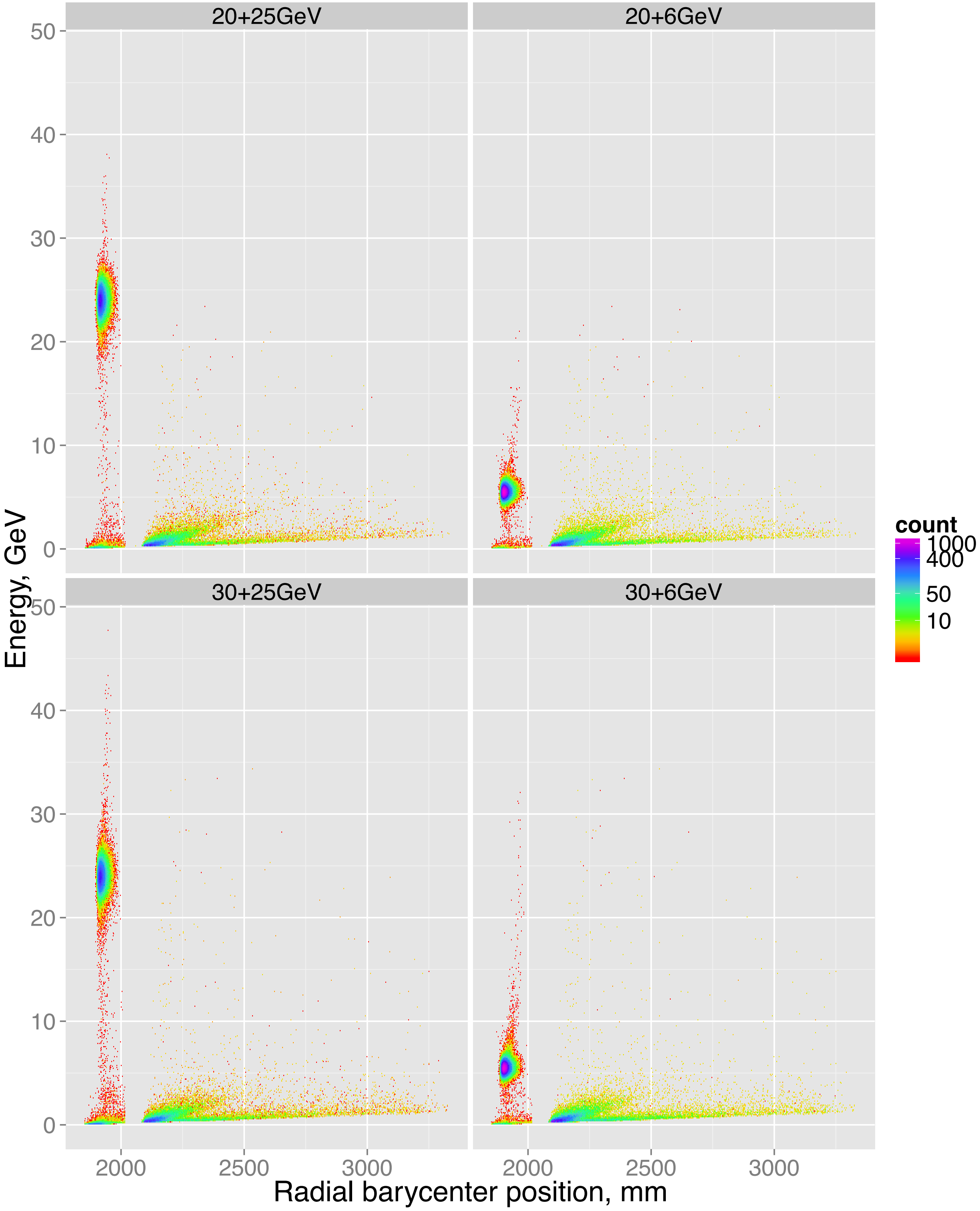


Figure A.50 – For this picture we select only neutral clusters reconstructed by Arbor in $\pi^+\gamma$ sample for ILD ECAL with 5×5 mm^2. The distribution of their energy versus the distance from their energy barycenter to IP in ILD radial direction is shown. ECAL and HCAL are separated by the gap at around 2050 mm. In addition to γ, many neutral unphysical clusters are visible in HCAL,though with low energies.

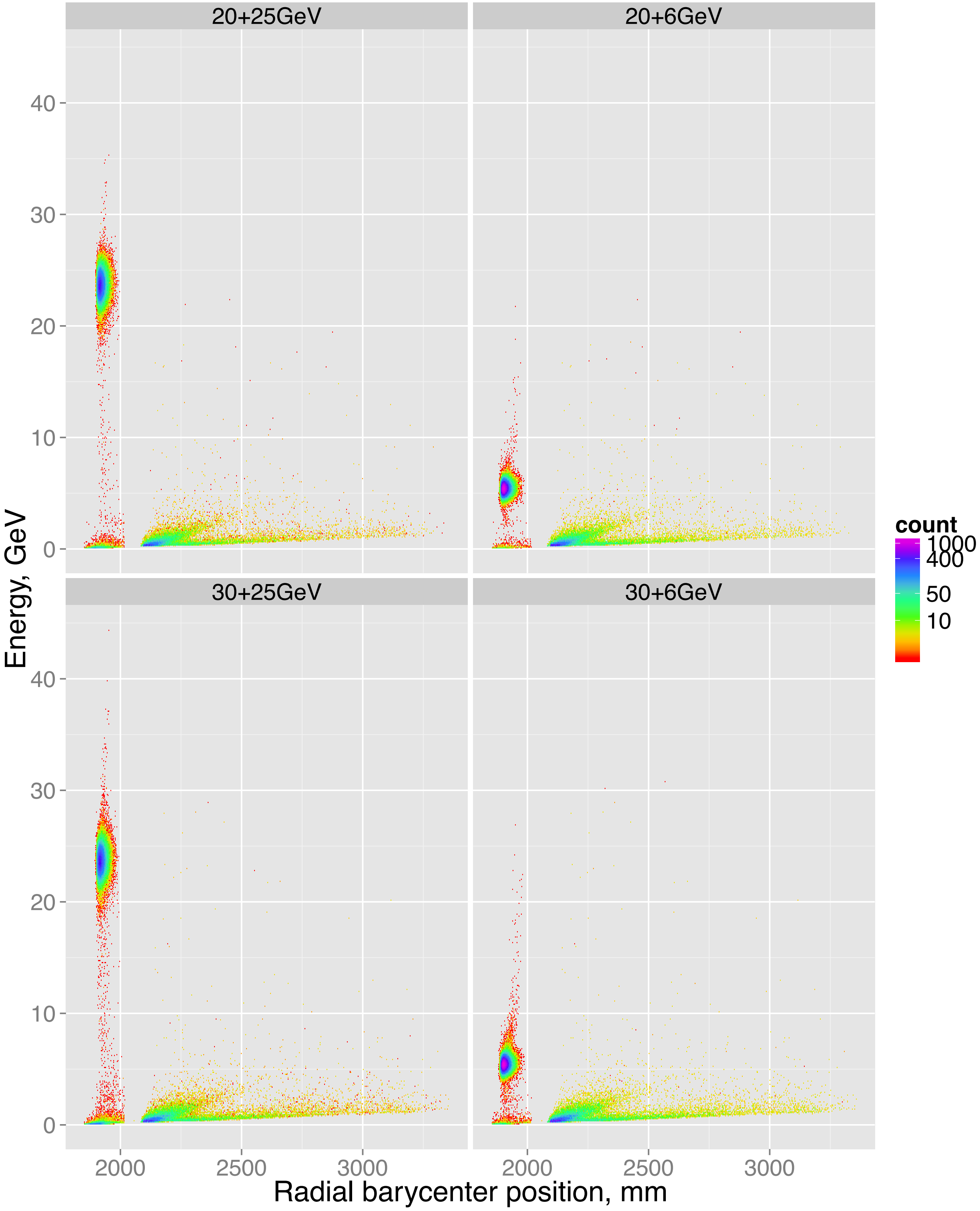
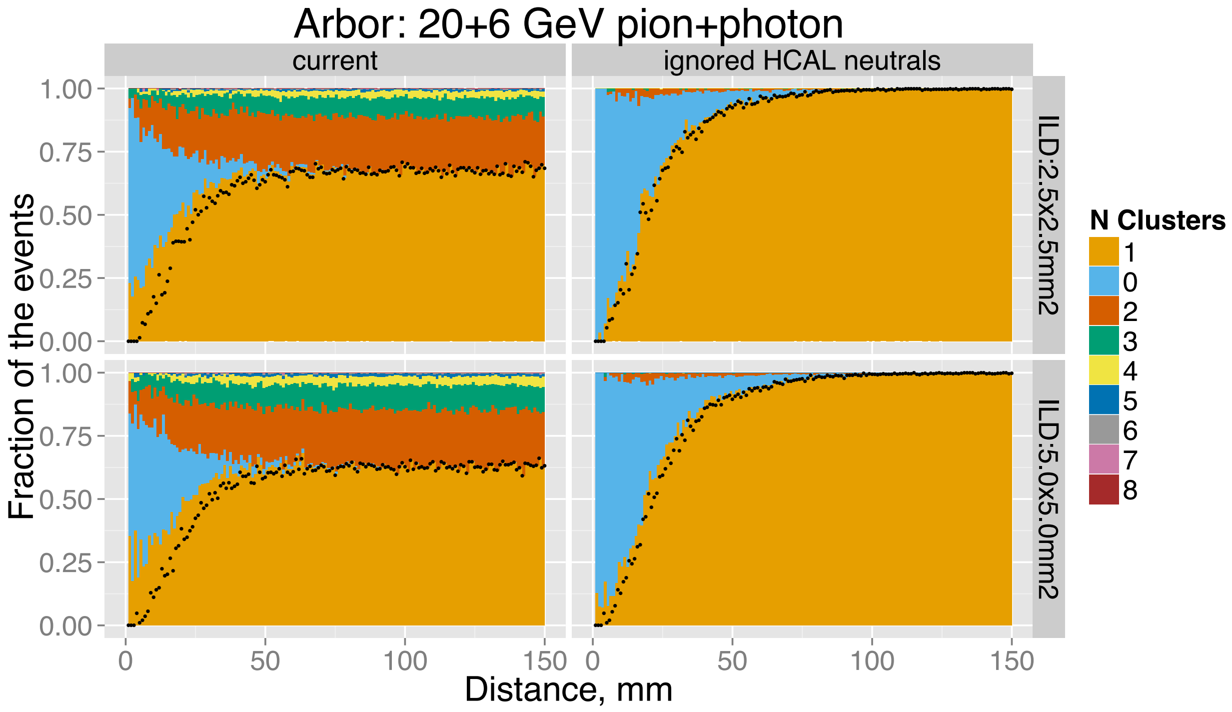


Figure A.51 – The same as in Figure A.50 but with 2.5×2.5 mm^2 ECAL pixel.

A.2.6 Classification of inefficiencies in pion-photon reconstruction



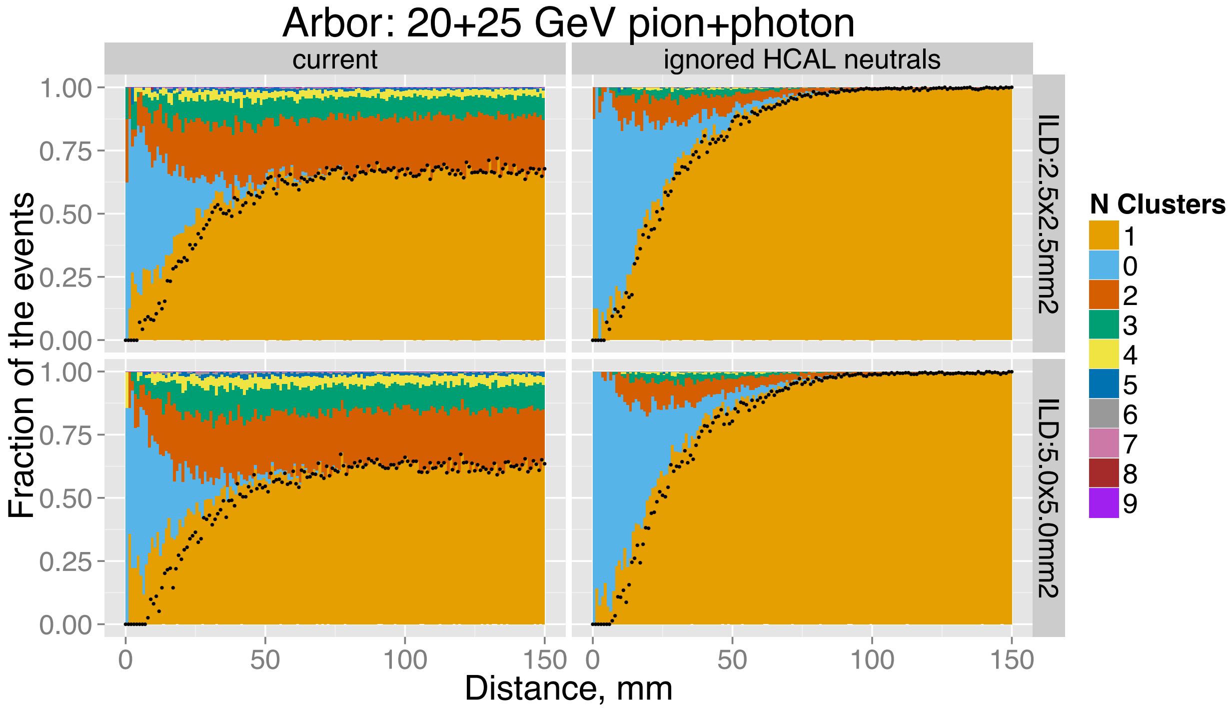
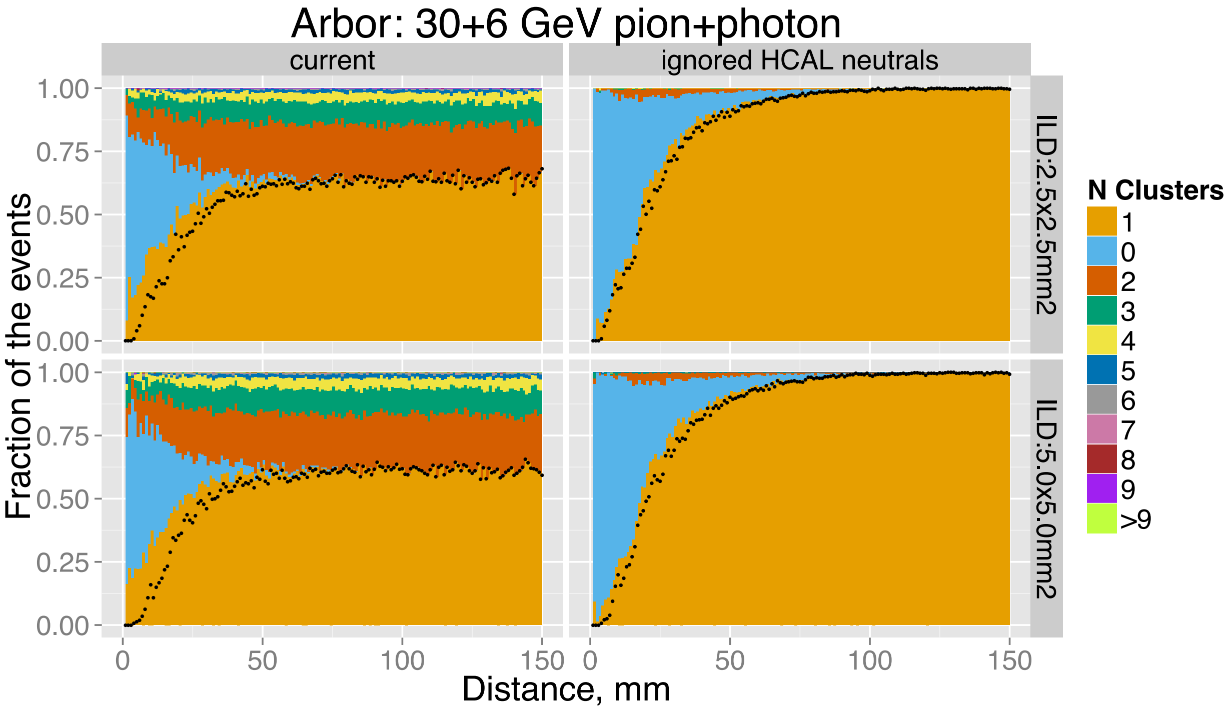


Figure A.52 – Arbor (SDHCAL) reconstruction of 20+6 (top), 20+25 (bottom) GeV pion-photon events: the colored bands show the fraction of events with 0, 1, 2, ... reconstructed neutral clusters, regardless of their energies and positions. The black points show the same efficiency as in Figures A.46, A.47. In the right column plots ("ignored HCAL neutrals") we ignore all HCAL neutral clusters, so that the inefficient events have either zero or two neutral clusters in ECAL.



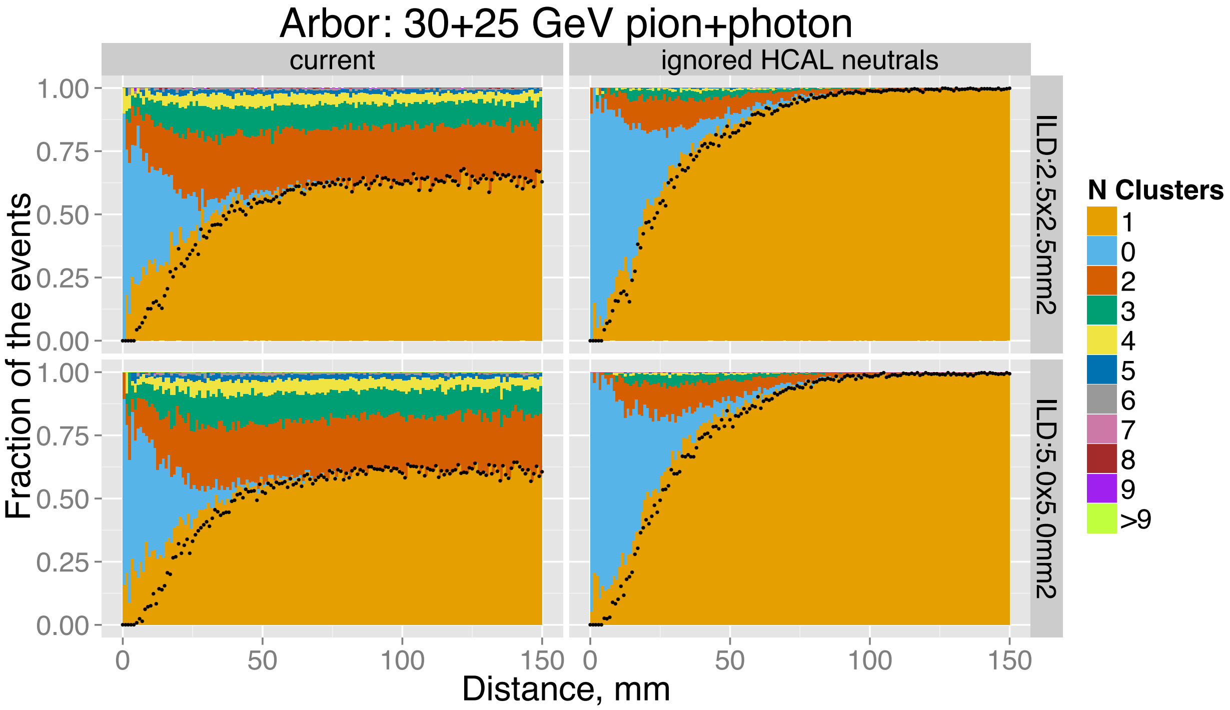
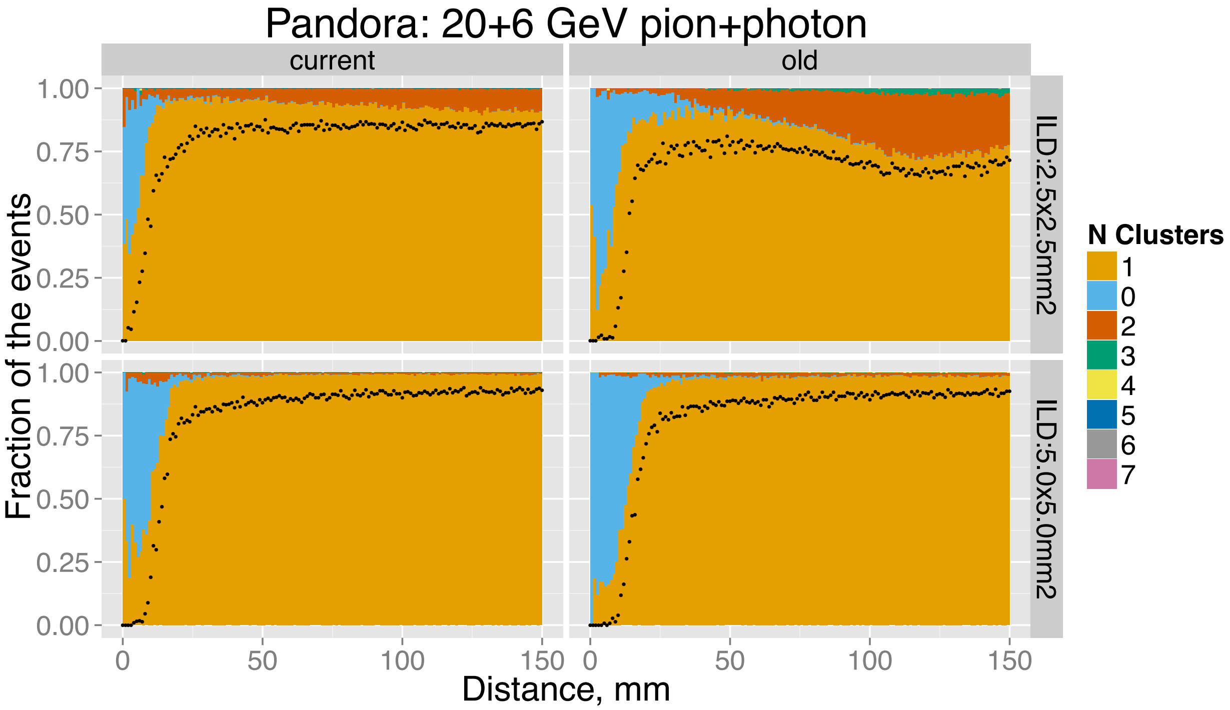


Figure A.53 – Same as in Figure A.52 but for 30+6 (top), 30+25 (bottom) GeV energy pairs.



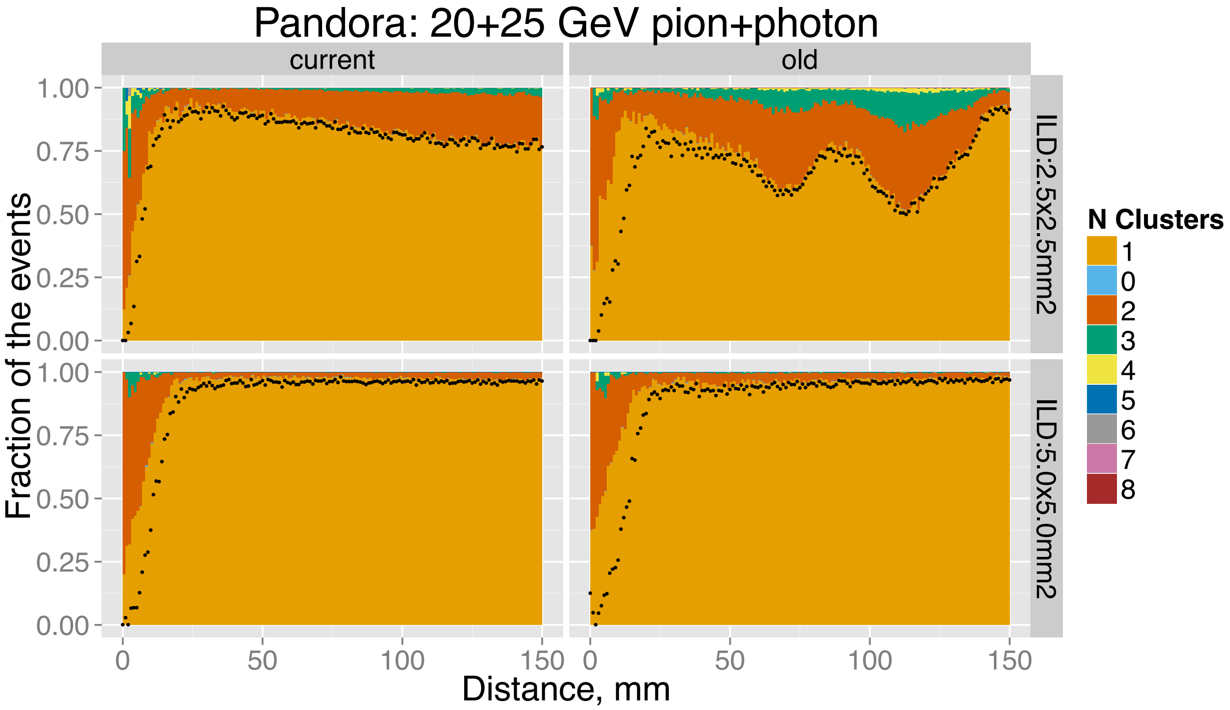
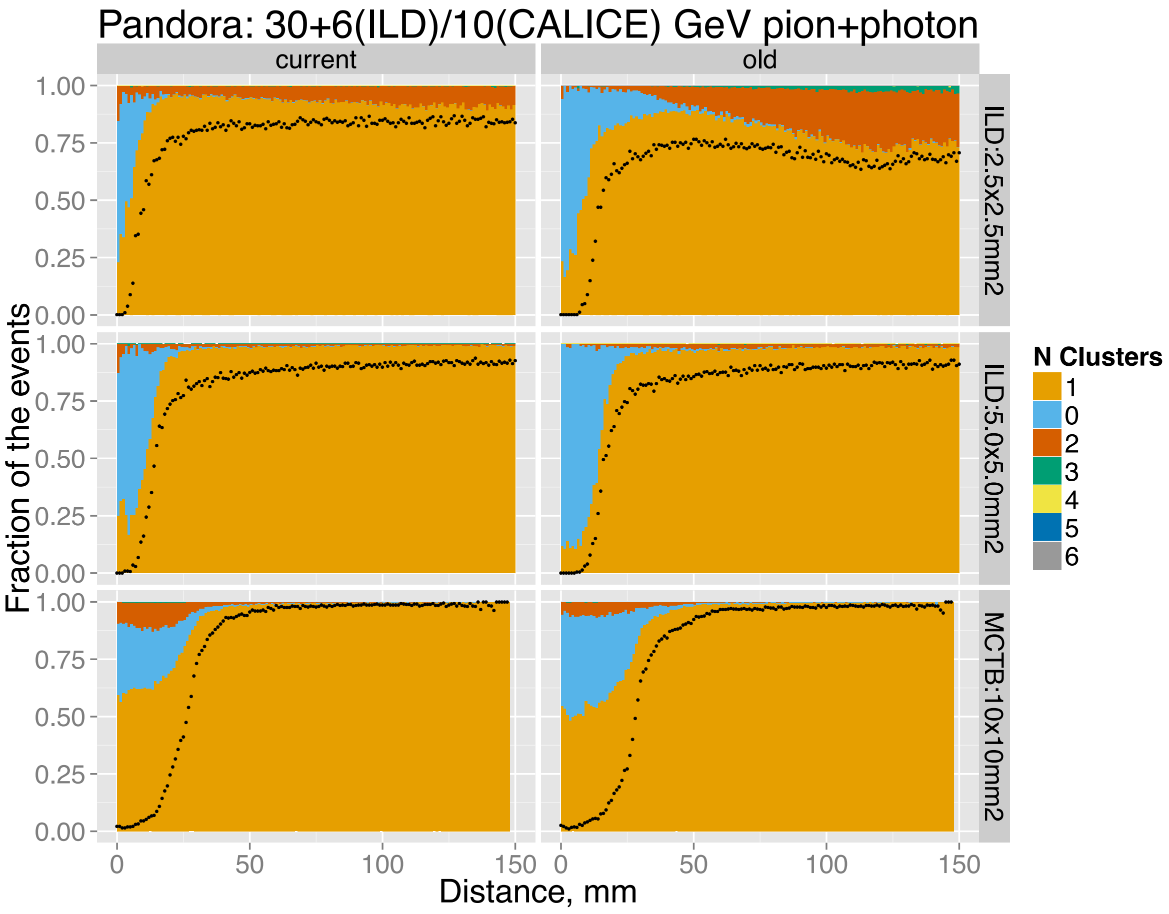


Figure A.54 – Pandora (AHCAL) reconstruction of 20+6 (top), 20+25 (bottom) GeV pion-photon events: the colored bands show the fraction of events with 0, 1, 2, . . . reconstructed neutral clusters, regardless of their energies and positions. The black points show the same efficiency as in Figures A.46, A.47.



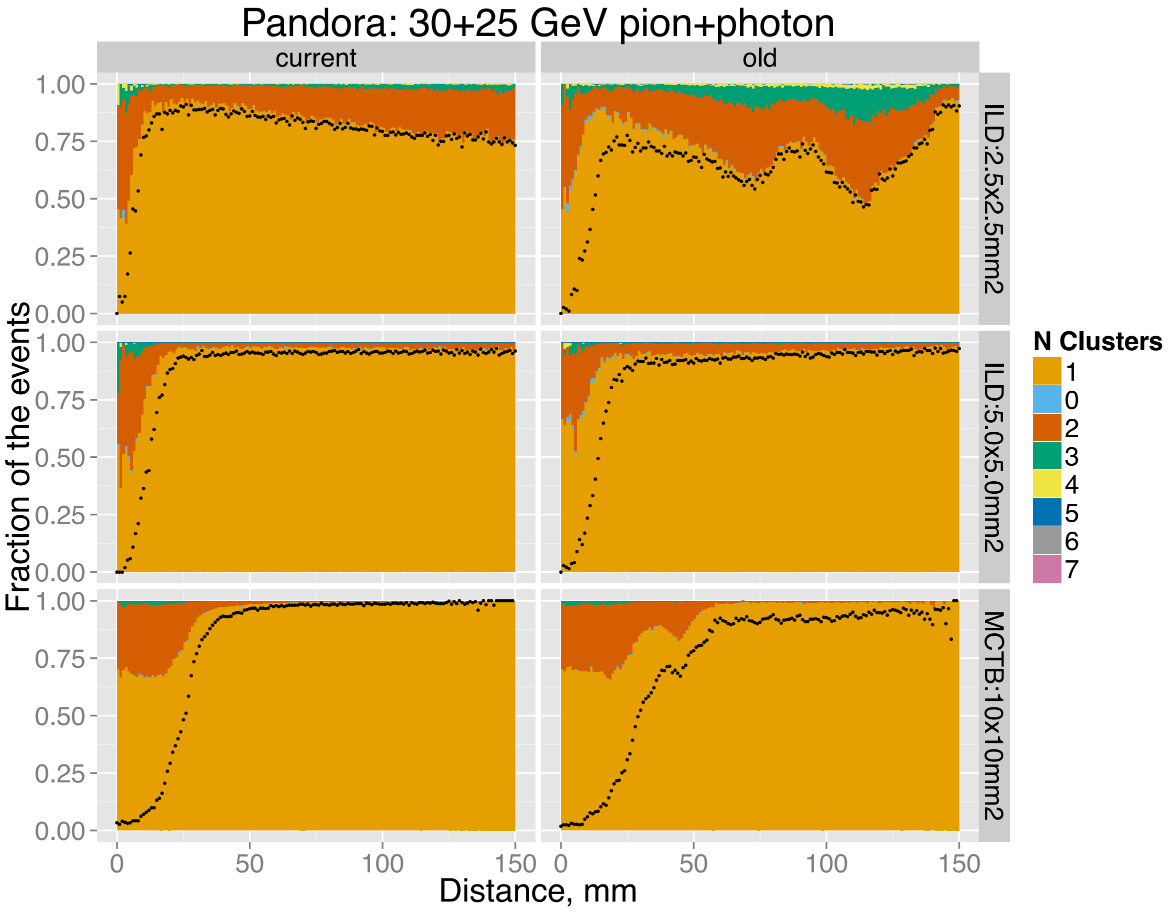
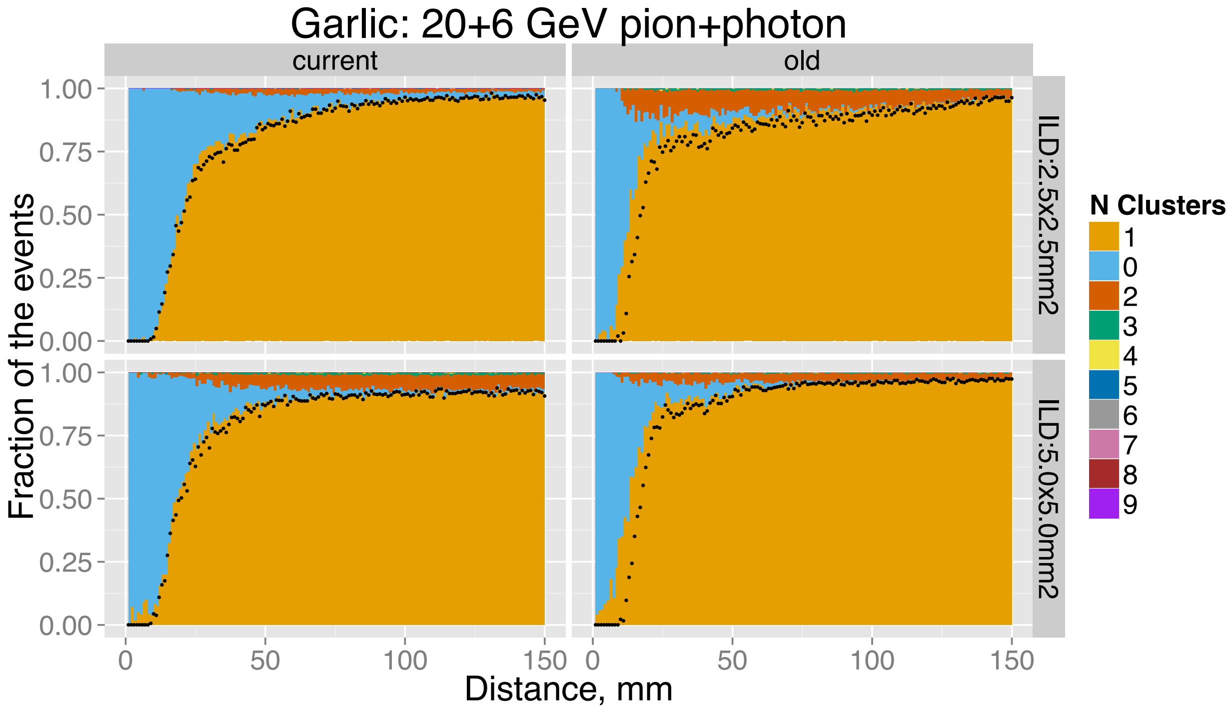


Figure A.55 – Same as in Figure A.54 but for 30+6 (ILD)/10 (CALICE) (top), 30+25 (bottom) GeV energy pairs. On the top plots we group together 30+6 GeV ILD (5 mm^2 and 2.5.5 mm^2 ECAL cells) and 30+10 GeV CALICE MC $\pi^+ - \gamma$ pairs. Bottom plots show performance for 30+25 GeV $\pi^+ - \gamma$ for ILD and CALICE.



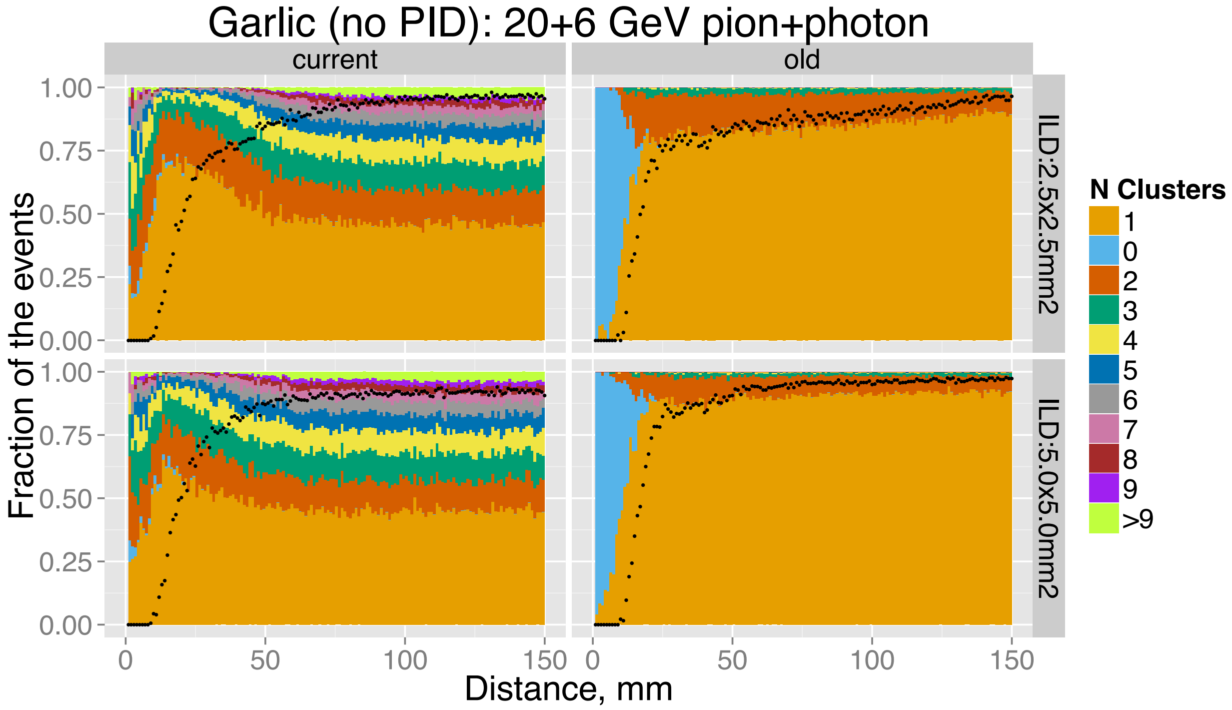


Figure A.56 – Garlic (no HCAL) reconstruction of 20+6 GeV pion-photon events: the colored bands show the fraction of events where 0, 1, 2, . . . photons (top) or neutral cluster candidates before the photon identification (bottom) are reconstructed, regardless of their energies and positions. The black points show the efficiency of reconstruction of exactly two photons (after the photon reidentification) with the energies and the positions within ±20% and ±5 mm, respectively, from their values in single shower events, i.e. as in Figure A.46. Note, that Garlic reconstructs only ECAL hits. The extra neutral clusters are formed by the pion shower hits not close to the pion track. The hits around the pion track are vetoed and not considered by Garlic.

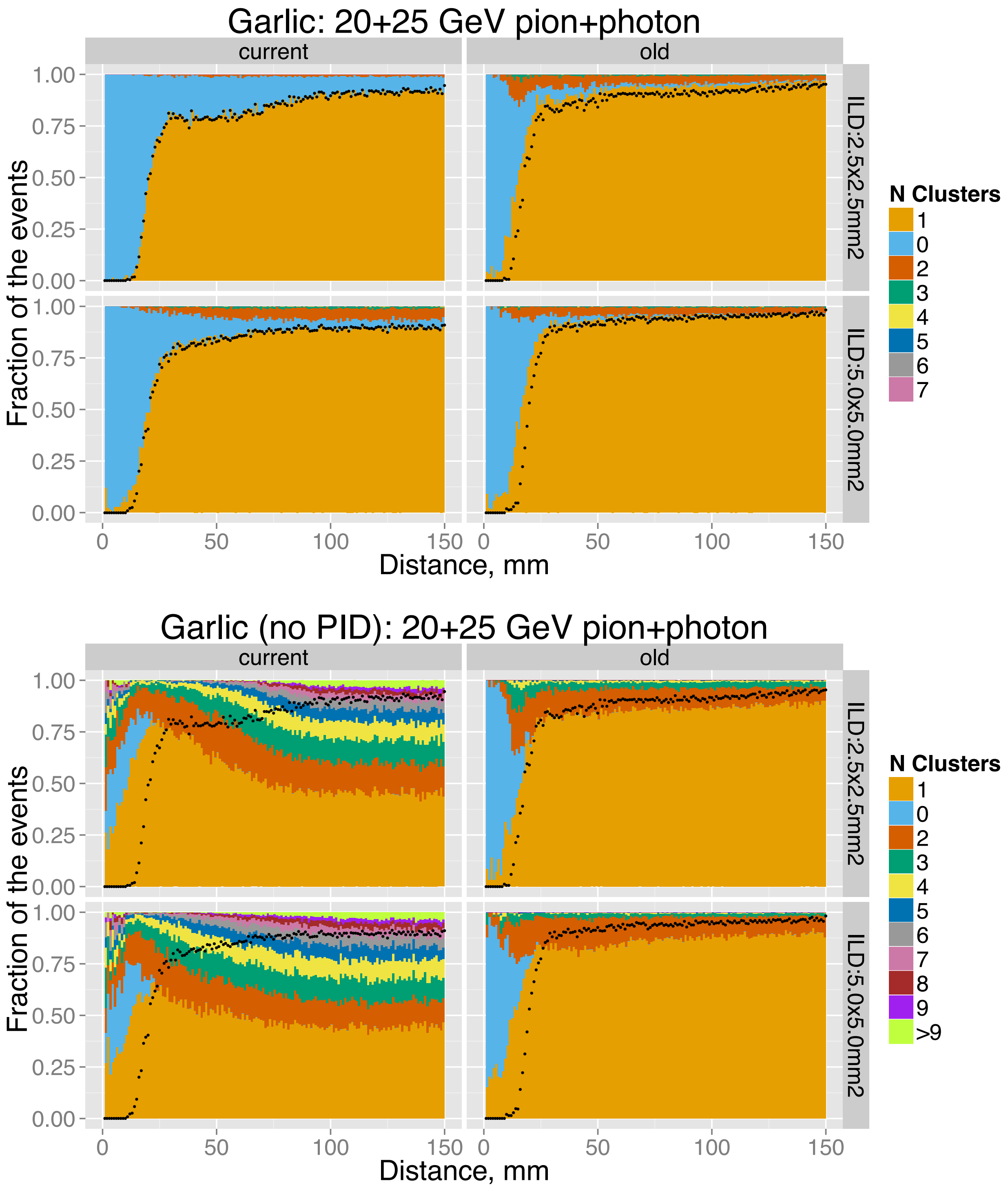
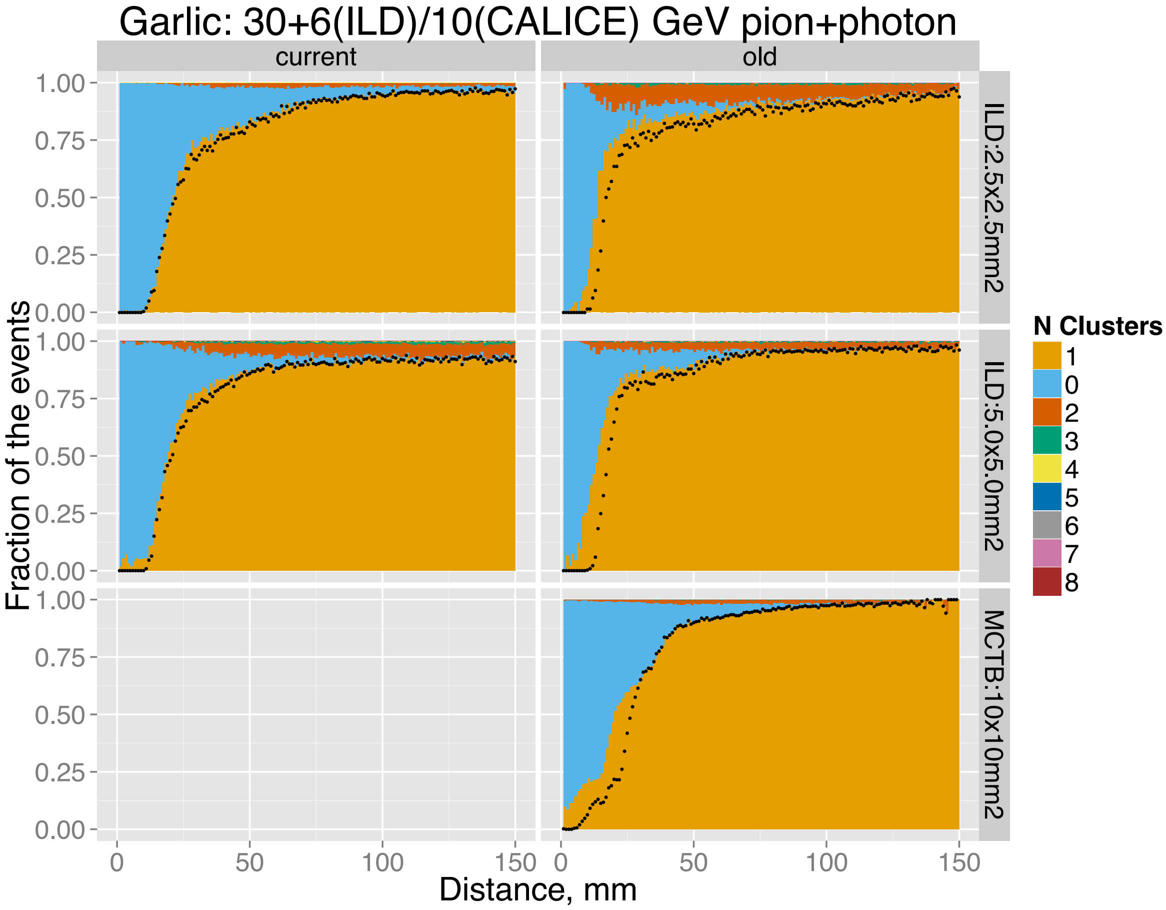


Figure A.57 – Same as in Figure A.56 but for 20+25 GeV energy pair.



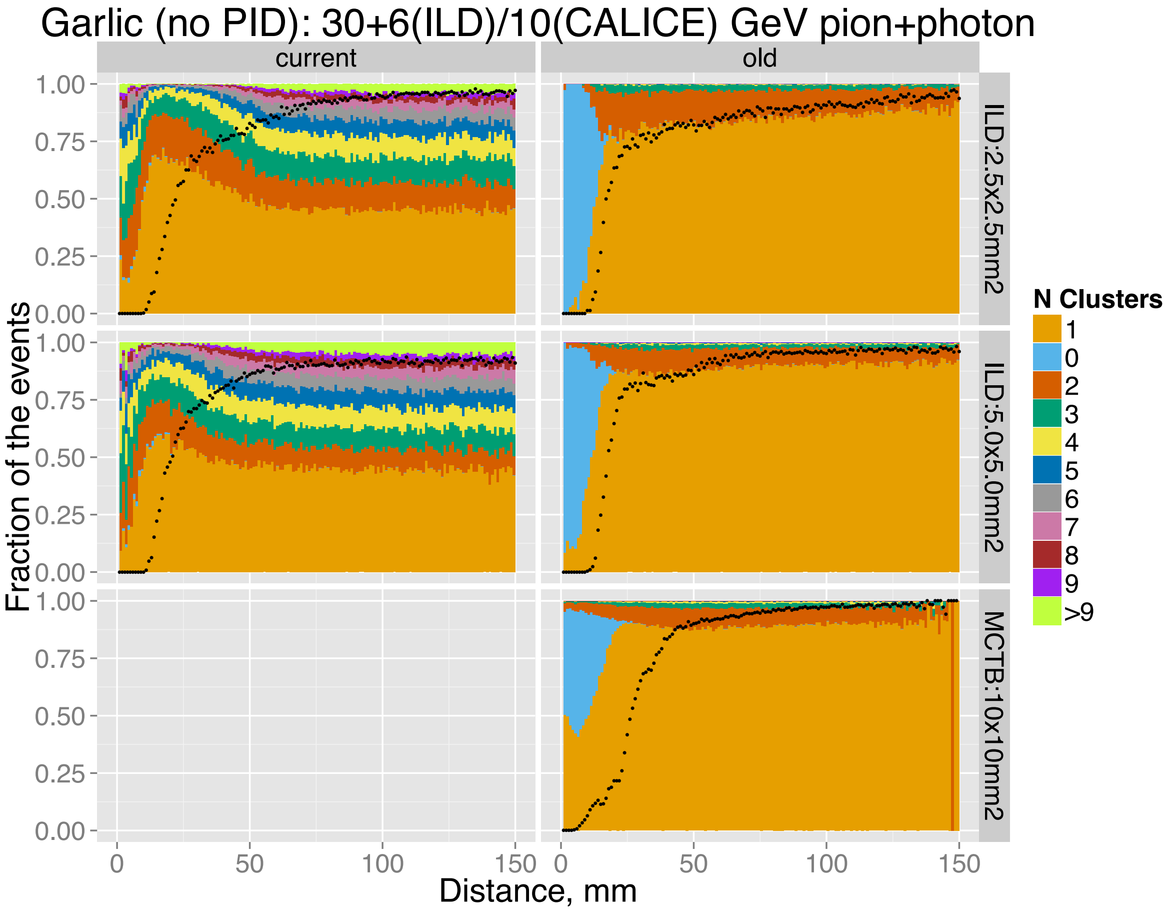
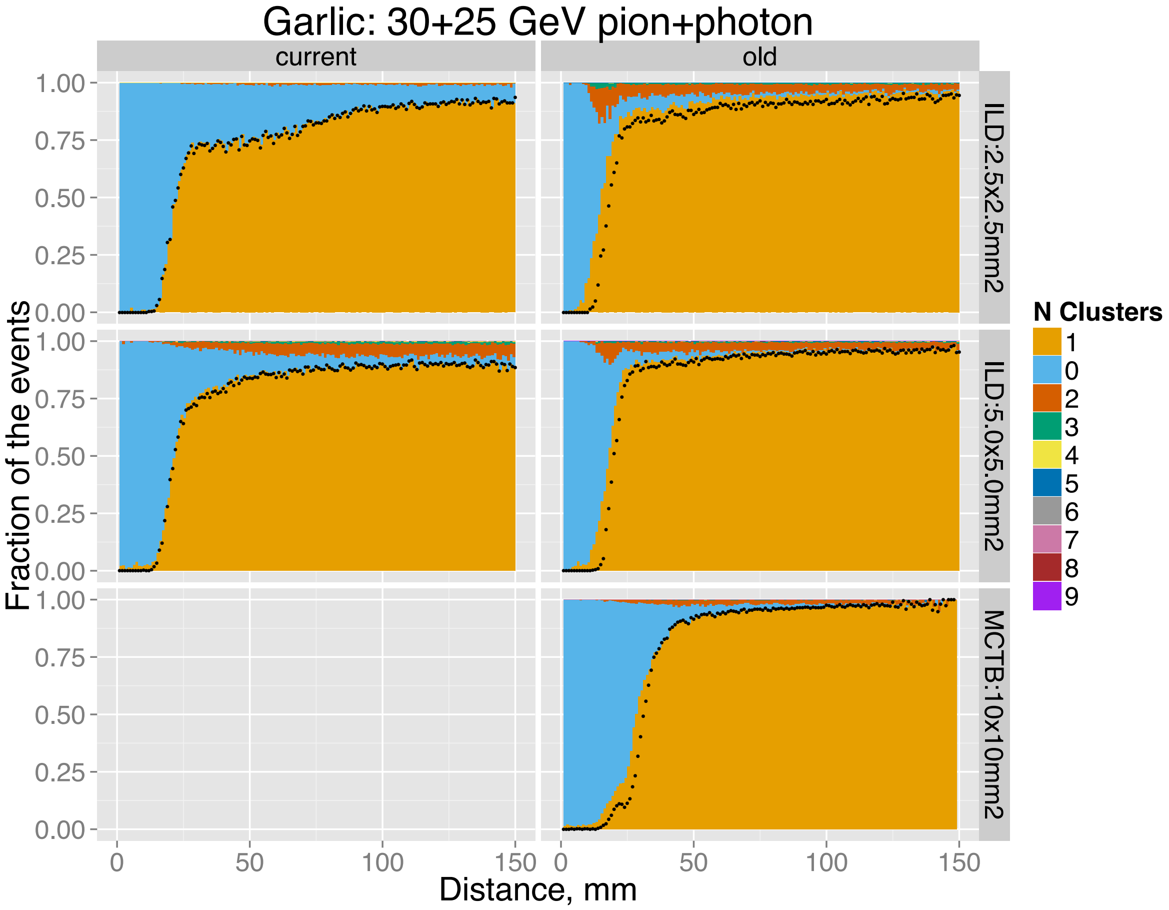


Figure A.58 – Same as in Figure A.56 but for 30+6 (ILD)/10 (CALICE) GeV energy pair.



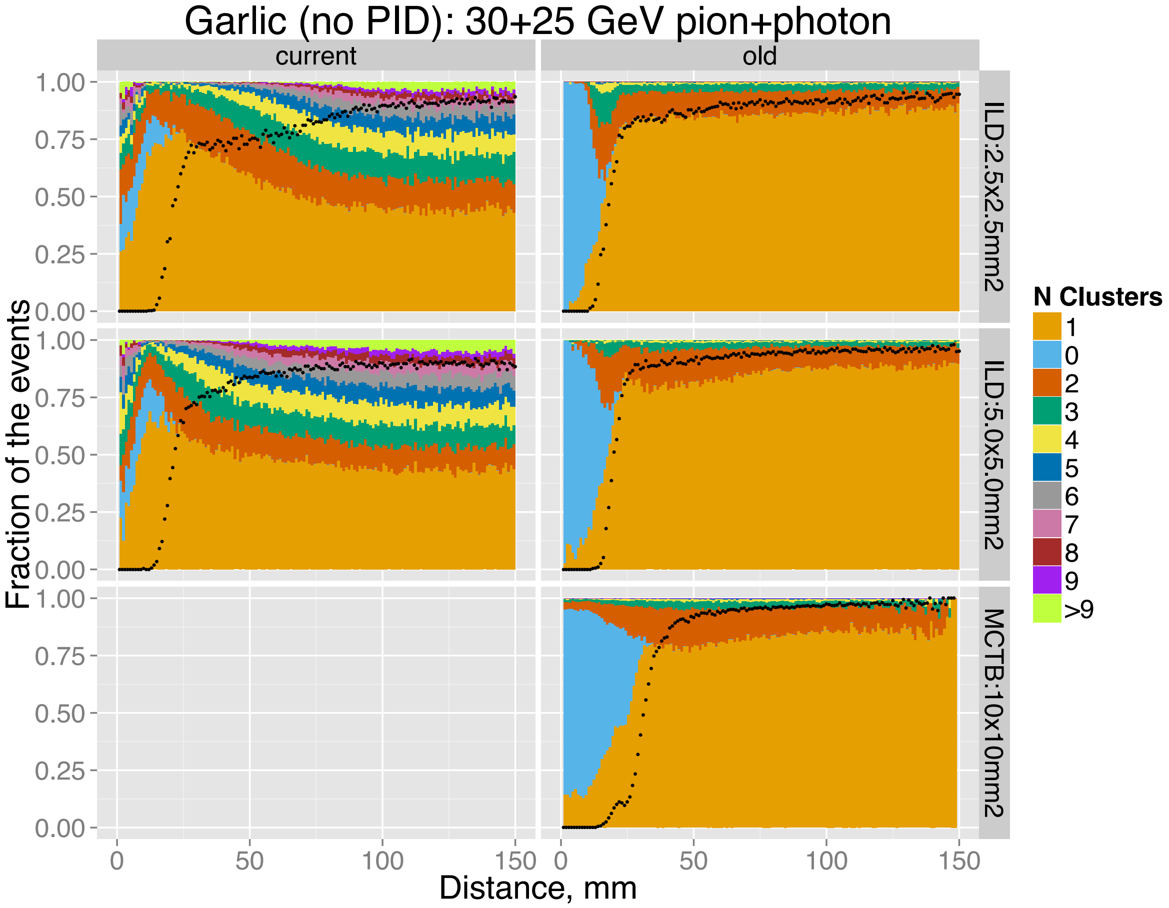


Figure A.59 – Same as in Figure A.56 but for 30+25 GeV energy pair.

A.2.7 JER of the simplest artificial pion+photon jet

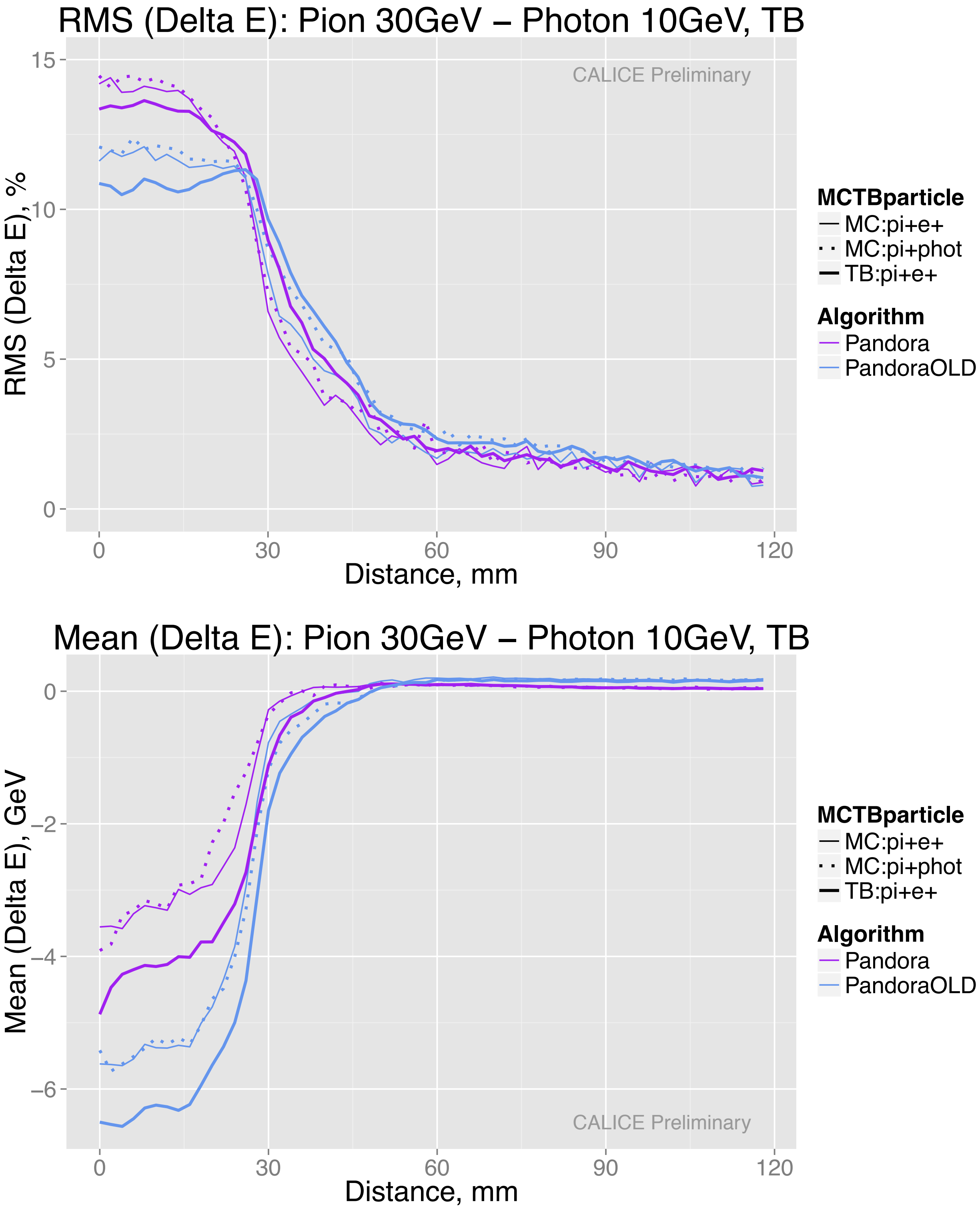
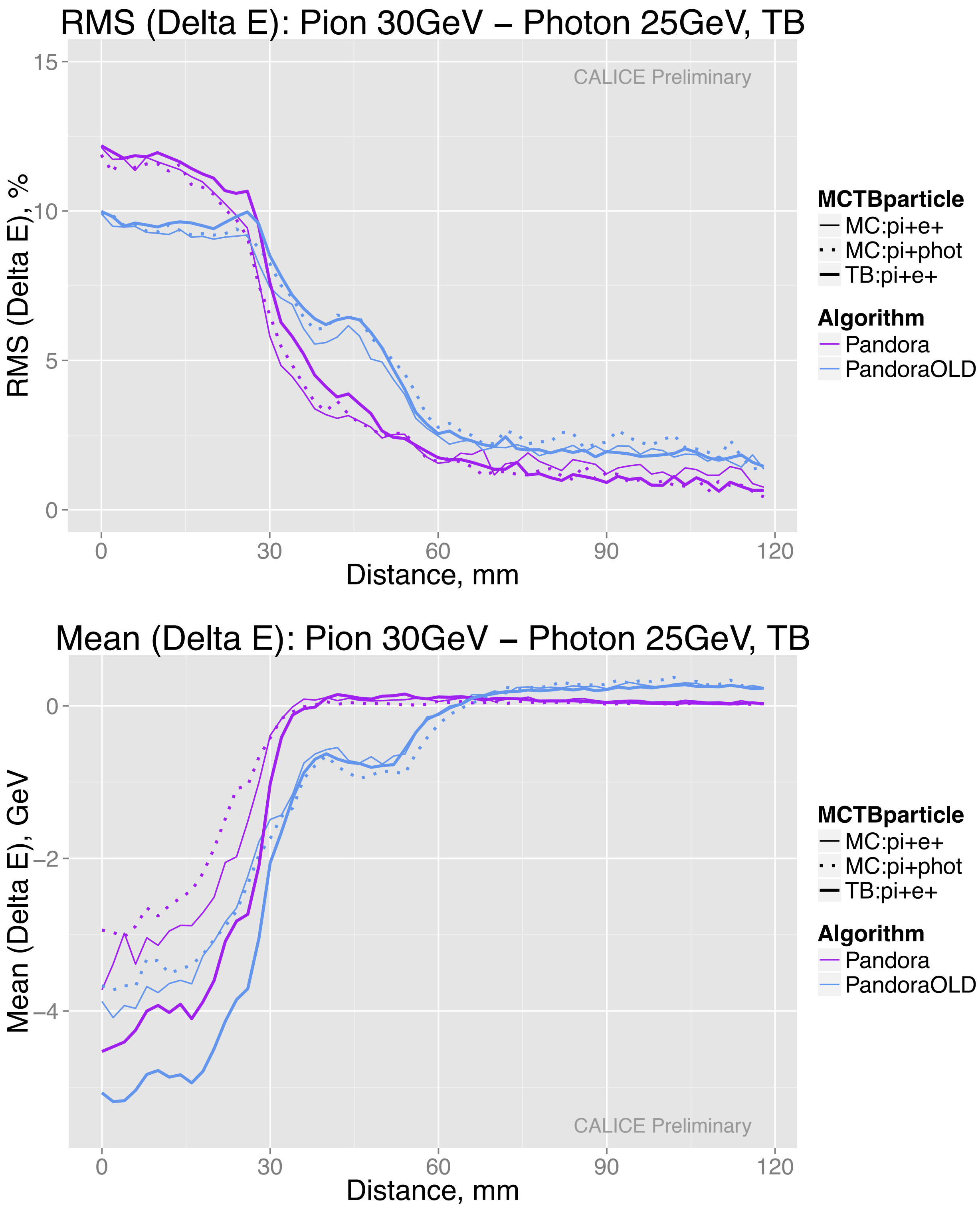
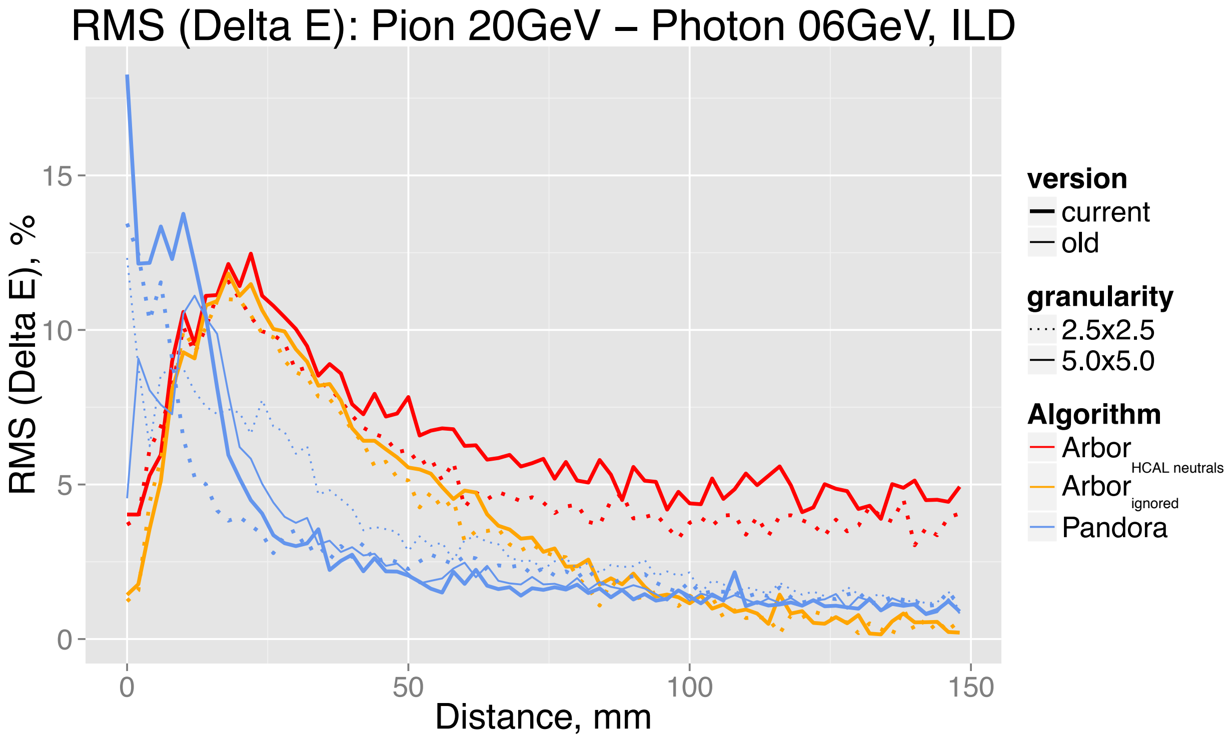


Figure A.60 – The standard deviation (top) and the mean (bottom) of the confusion energy ΔE defined in Eq. 4.3 as a function of the distance between 30+10 GeV pion-EM showers (CERN'07). Note, it is impossible to build the same plots for Garlic, due to the absence of PID for hadrons.


RMS (Delta E): Pion 30GeV – Photon 25GeV, TB

Mean (Delta E): Pion 30GeV – Photon 25GeV, TB

Figure A.61 – Same as in Figure A.60 but for 30+25 GeV energy pair.



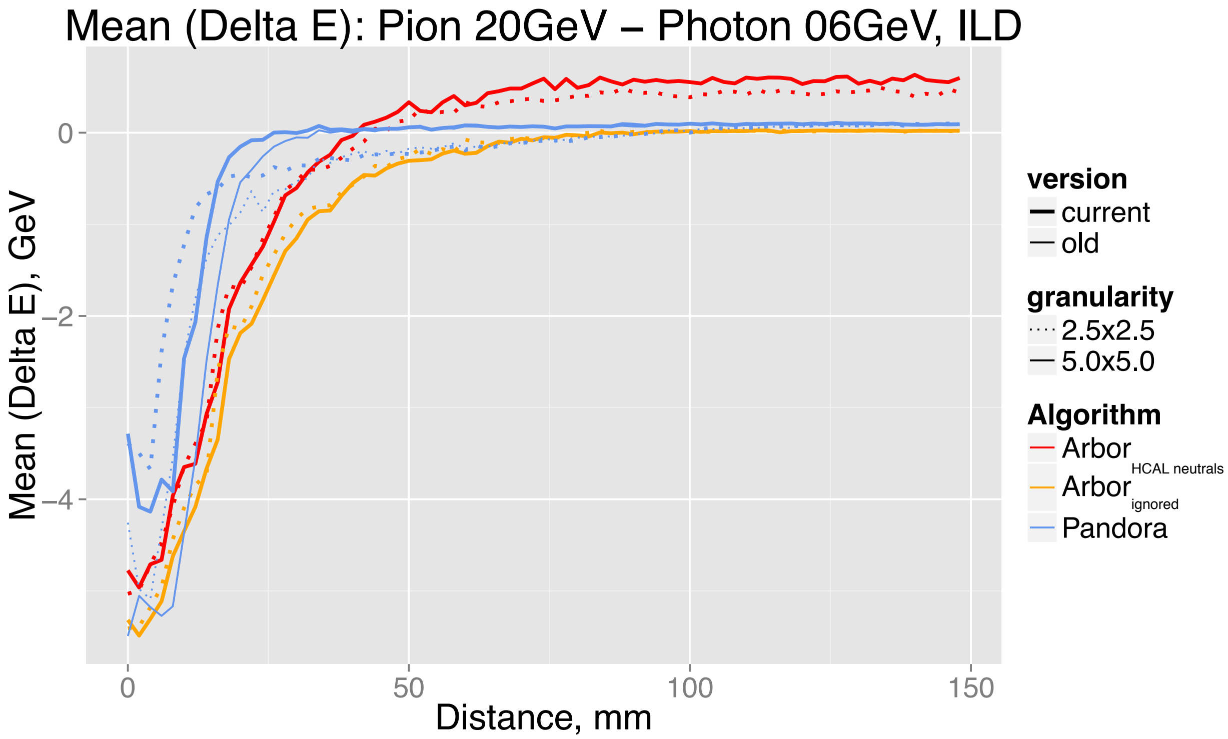


Figure A.62 – The standard deviation (top) and the mean (bottom) of the confusion energy ΔE defined in Eq. 4.3 as a function of the distance between 20+6 GeV pion-photon showers in ILD (Pandora - AHCAL, Arbor - SDHCAL).

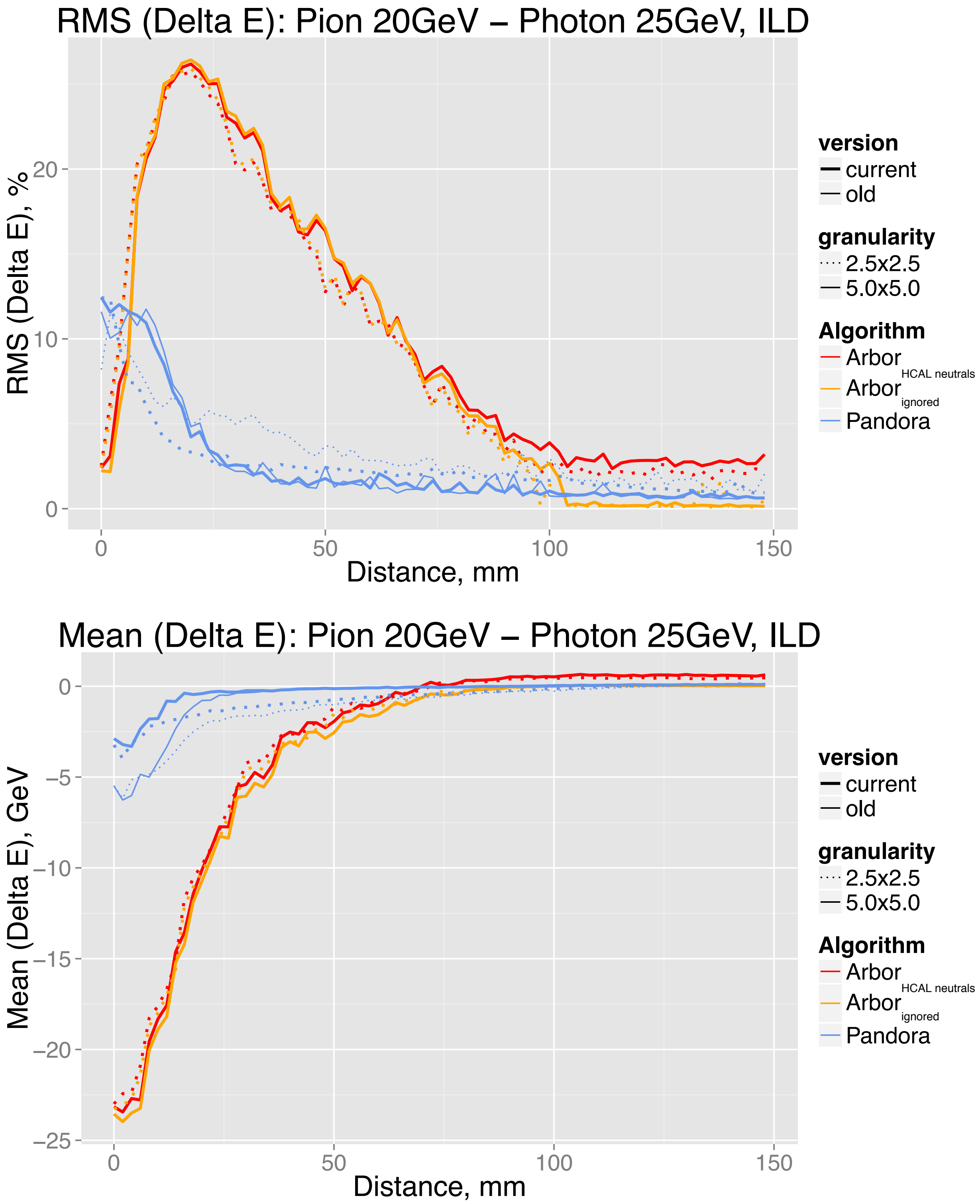


Figure A.63 – Same as in Figure A.62 but for 20+25 GeV energy pair.

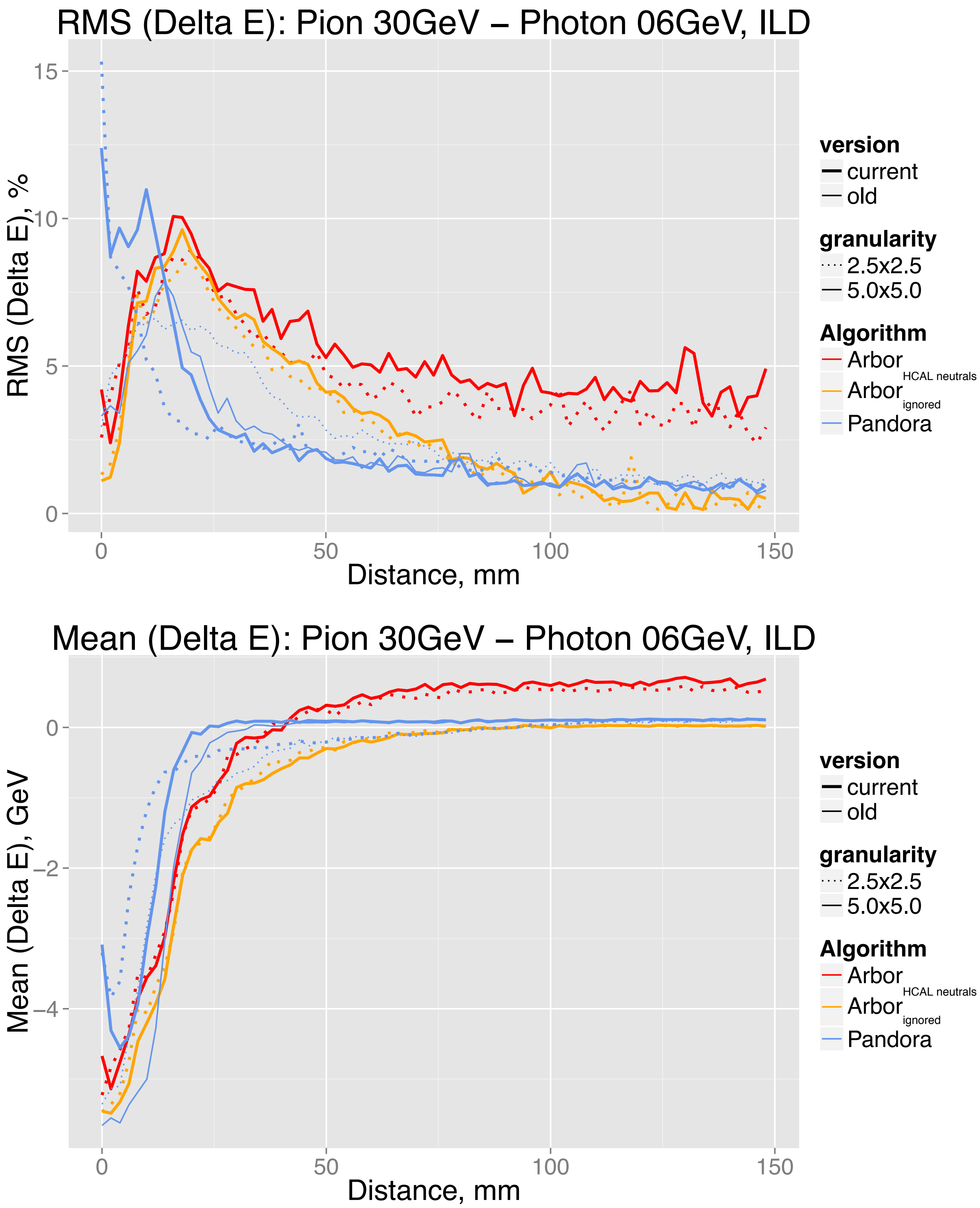


Figure A.64 – Same as in Figure A.62 but for 30+6 GeV energy pair.

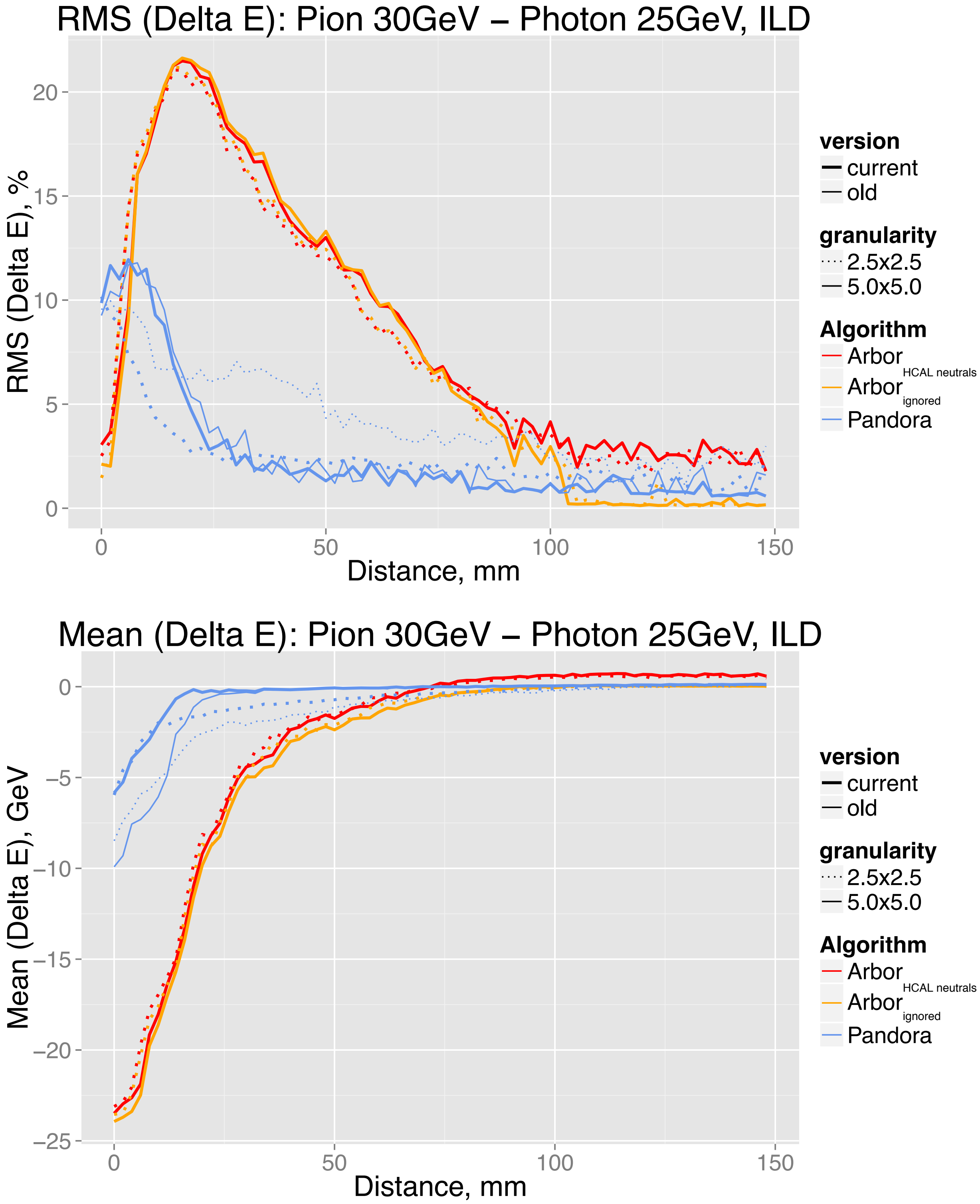


Figure A.65 – Same as in Figure A.62 but for 30+25 GeV energy pair.

Appendix B

Appendix for Chapter 5 on technological prototype

B.1 Charge injection studies with FEV8 boards

B.1.1 Average channel response in three slabs

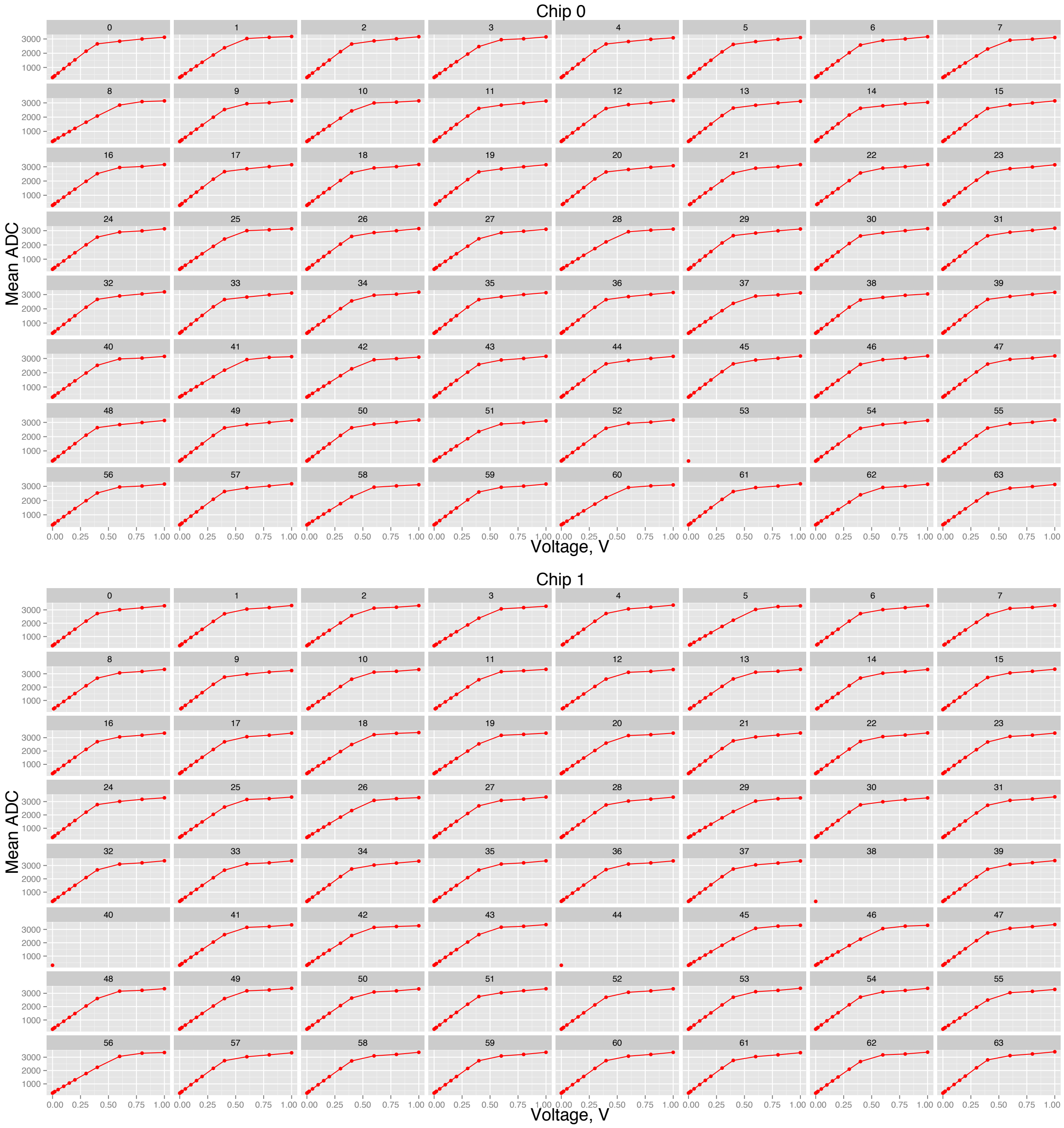


Figure B.1 – Average channel response as a function of injection voltage for 64 channels in chips 0, 1 in slab 10, PP mode.

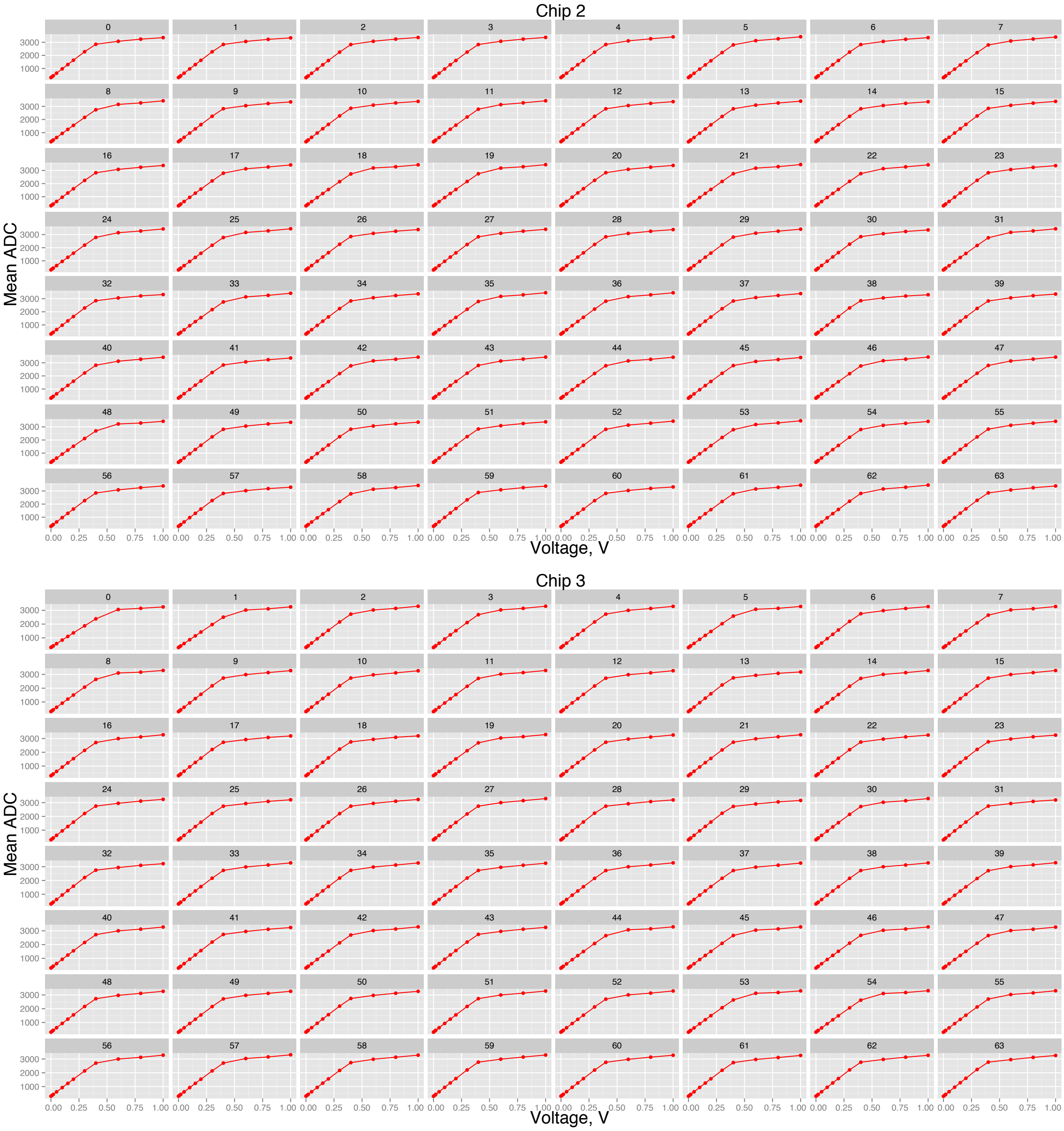


Figure B.2 – Same as in previous Figure B.1, but for chips 2, 3 in slab 10.

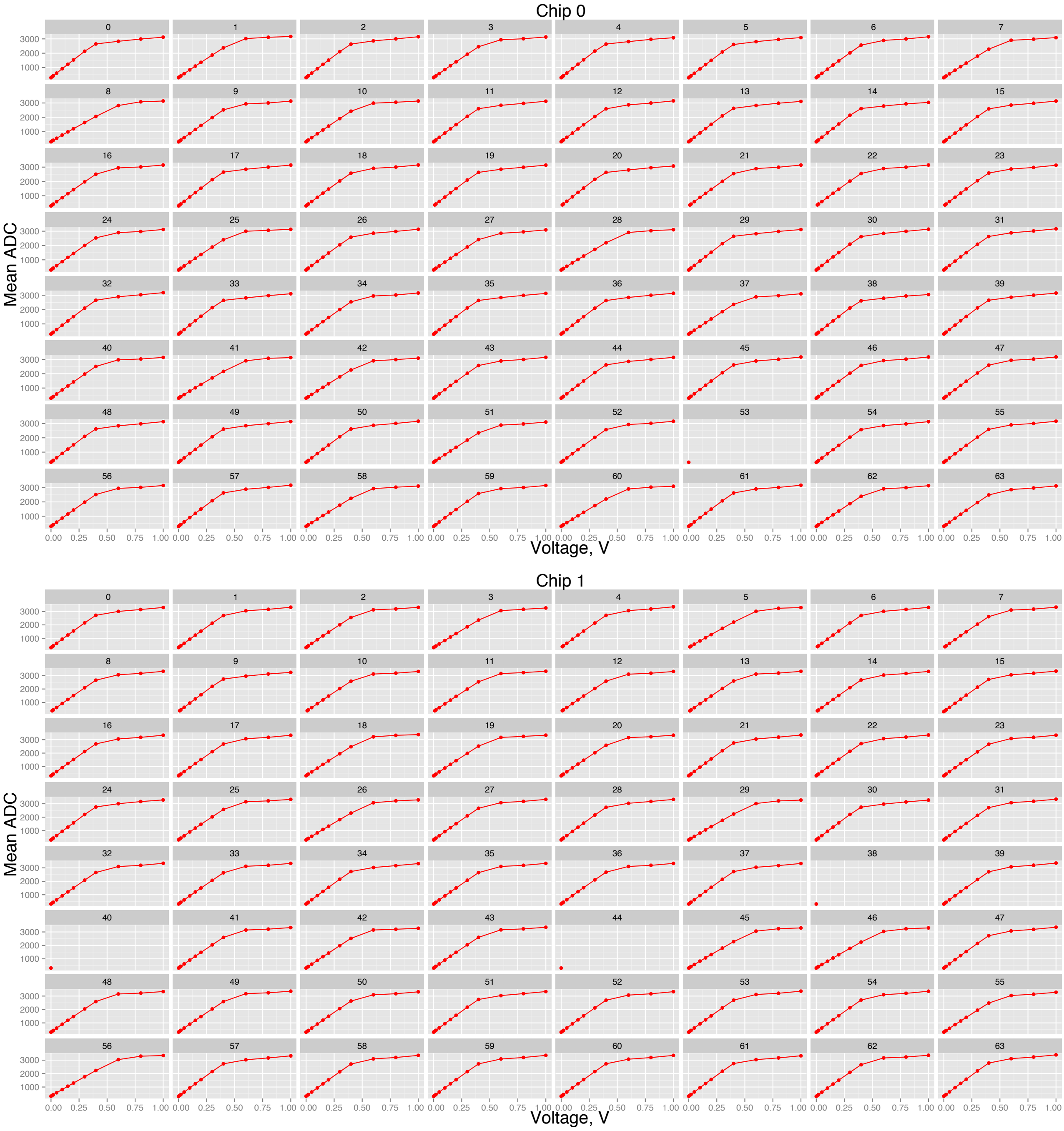


Figure B.3 – Average channel response as a function of injection voltage for 64 channels in chips 0, 1 in slab 10, CC mode.

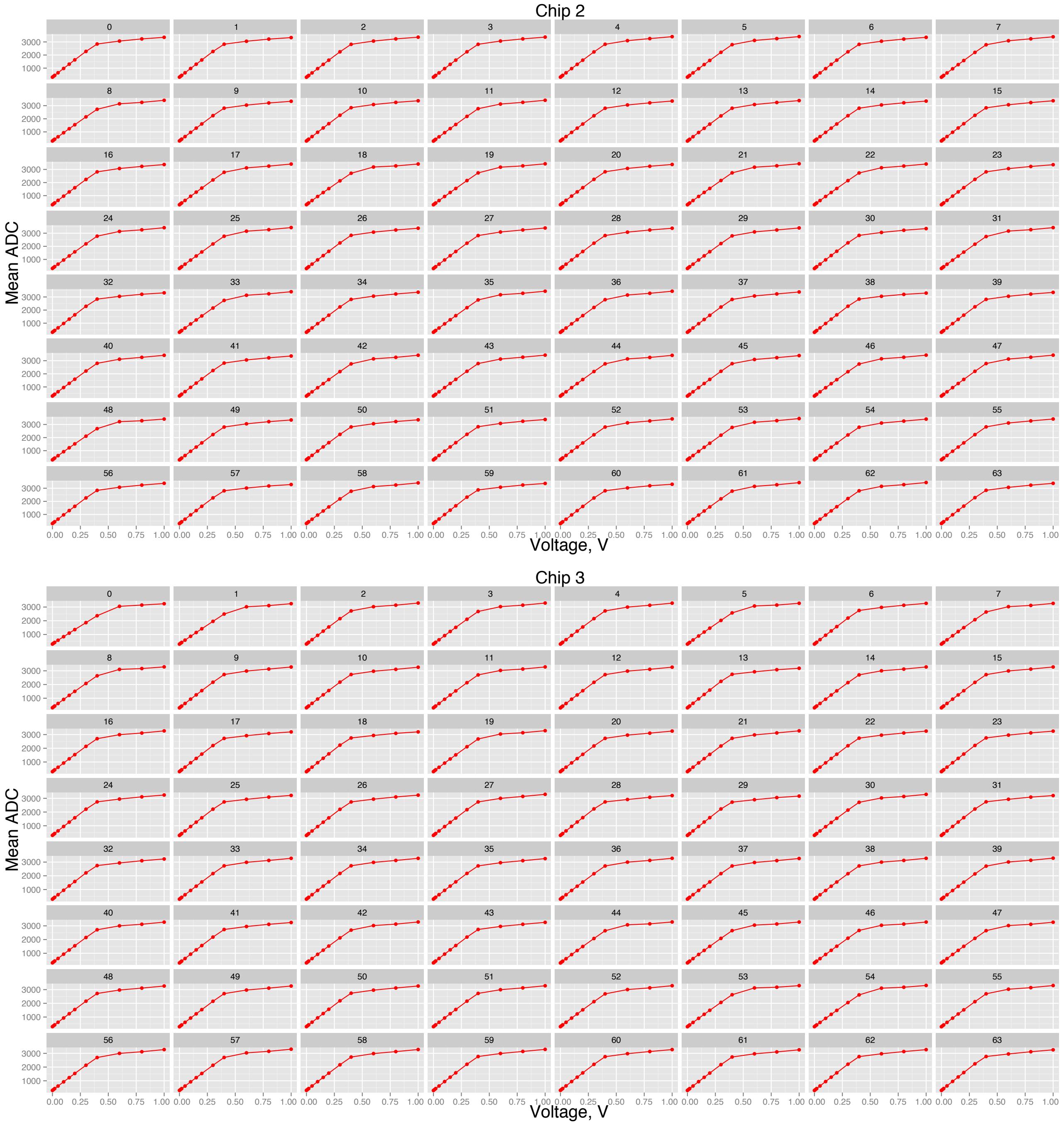


Figure B.4 – Same as in previous Figure B.3, but for chips 2, 3 in slab 10.

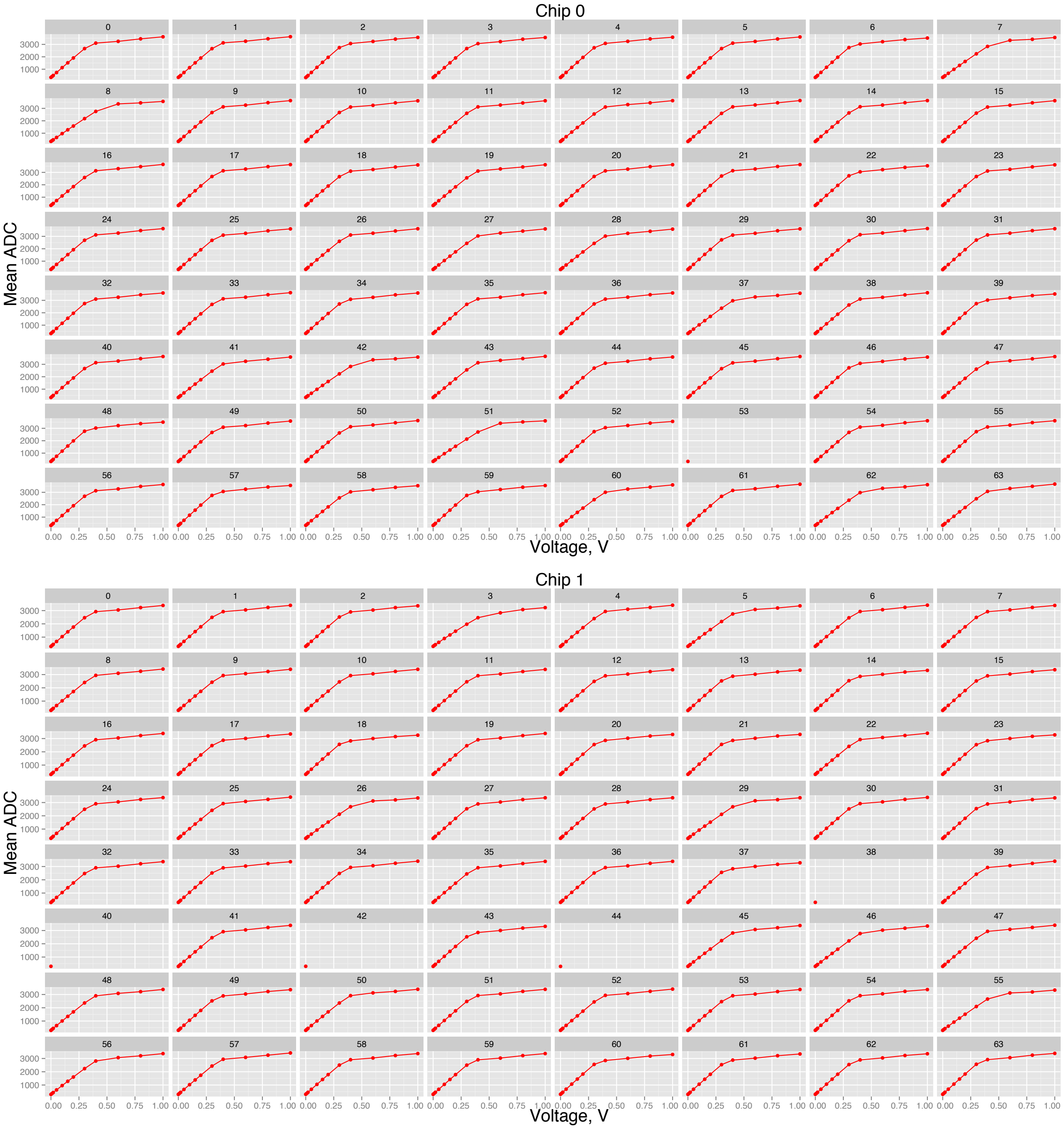


Figure B.5 – Average channel response as a function of injection voltage for 64 channels in chips 0, 1 in slab 5, PP mode.

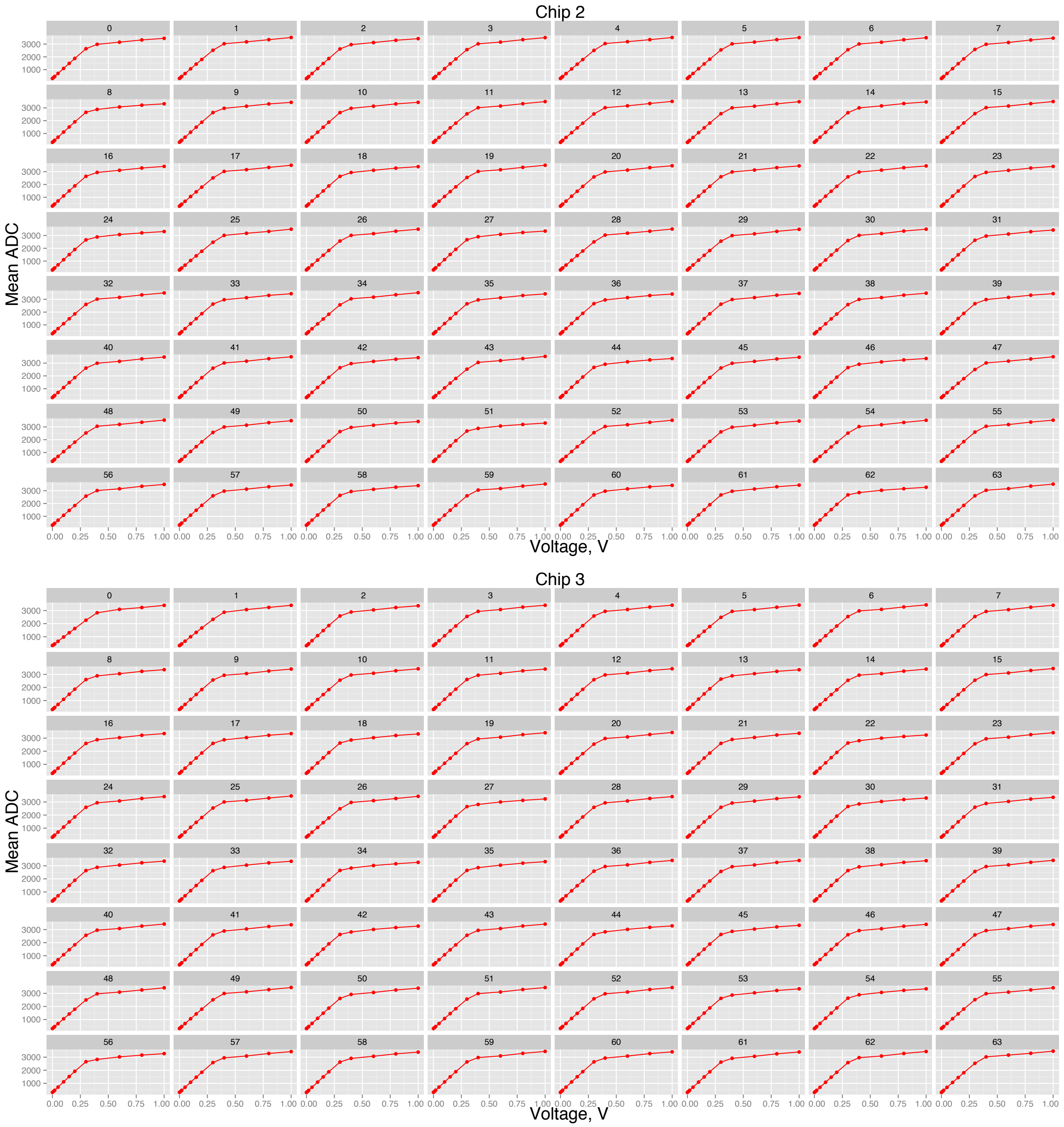


Figure B.6 – Same as in previous Figure B.5, but for chips 2, 3 in slab 5.

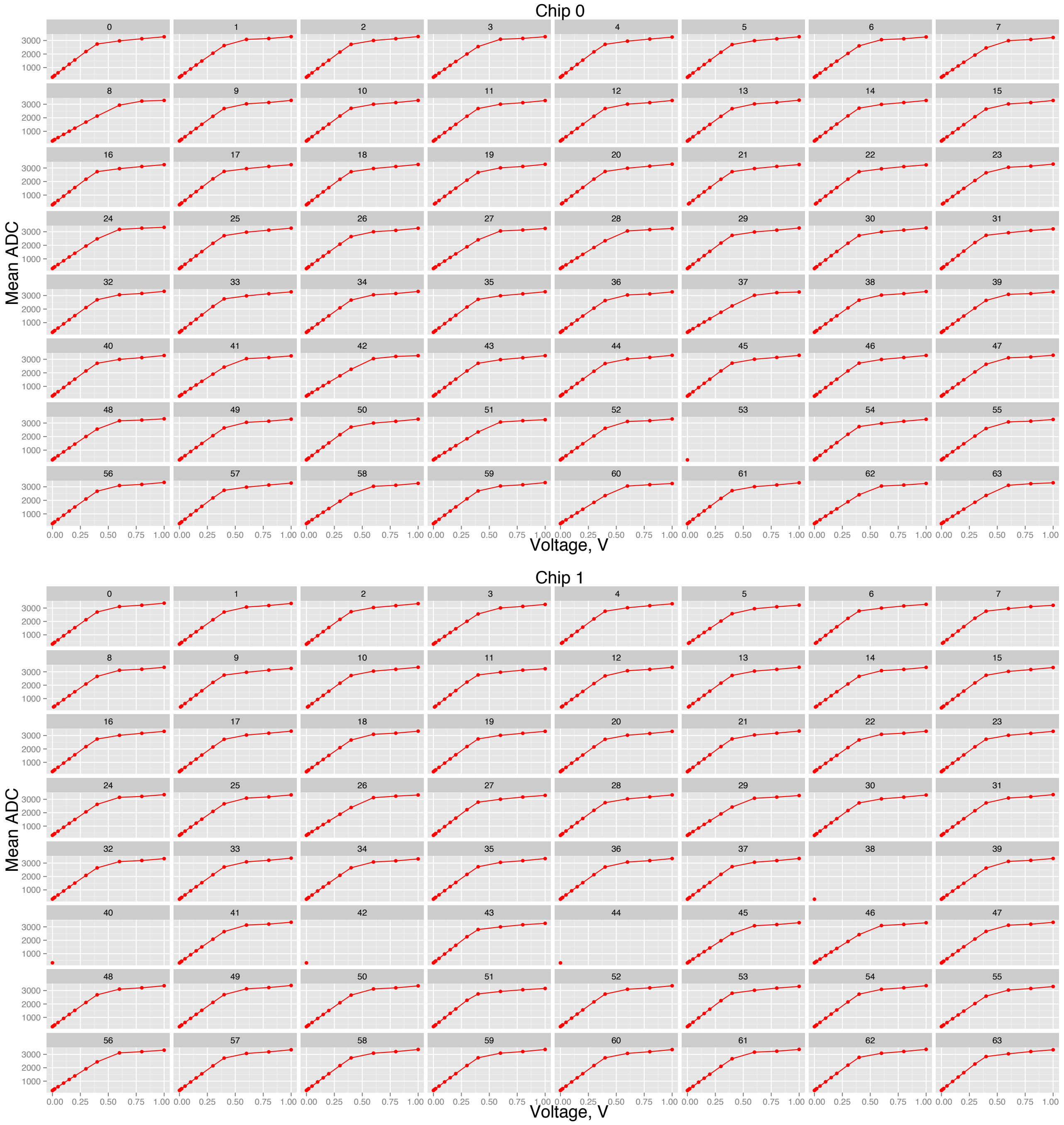


Figure B.7 – Average channel response as a function of injection voltage for 64 channels in chips 0, 1 in slab 8, CC mode.

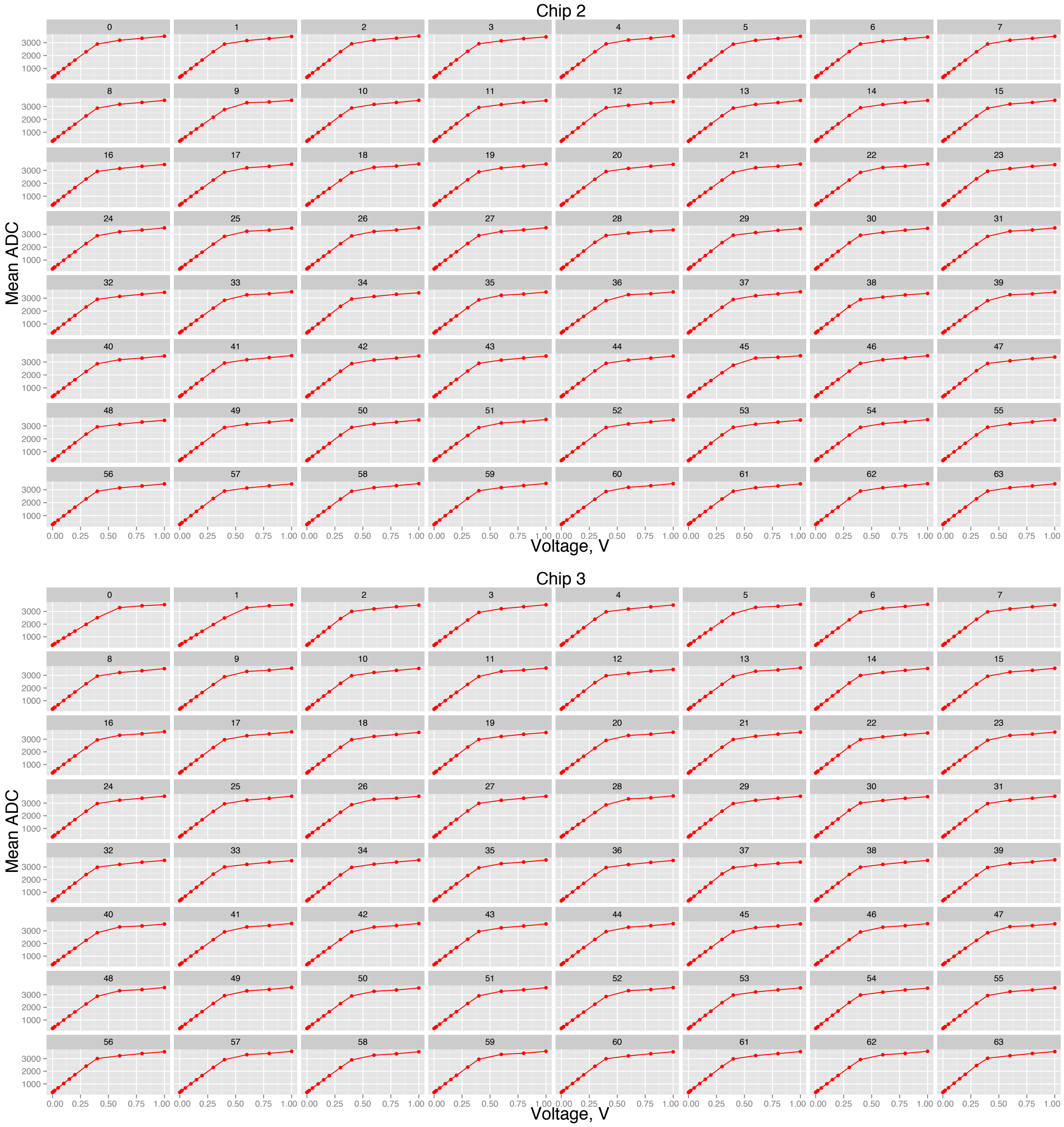


Figure B.8 – Same as in previous Figure B.7, but for chips 2, 3 in slab 8.

B.1.2 Difference between PP and CC injection modes in one slab

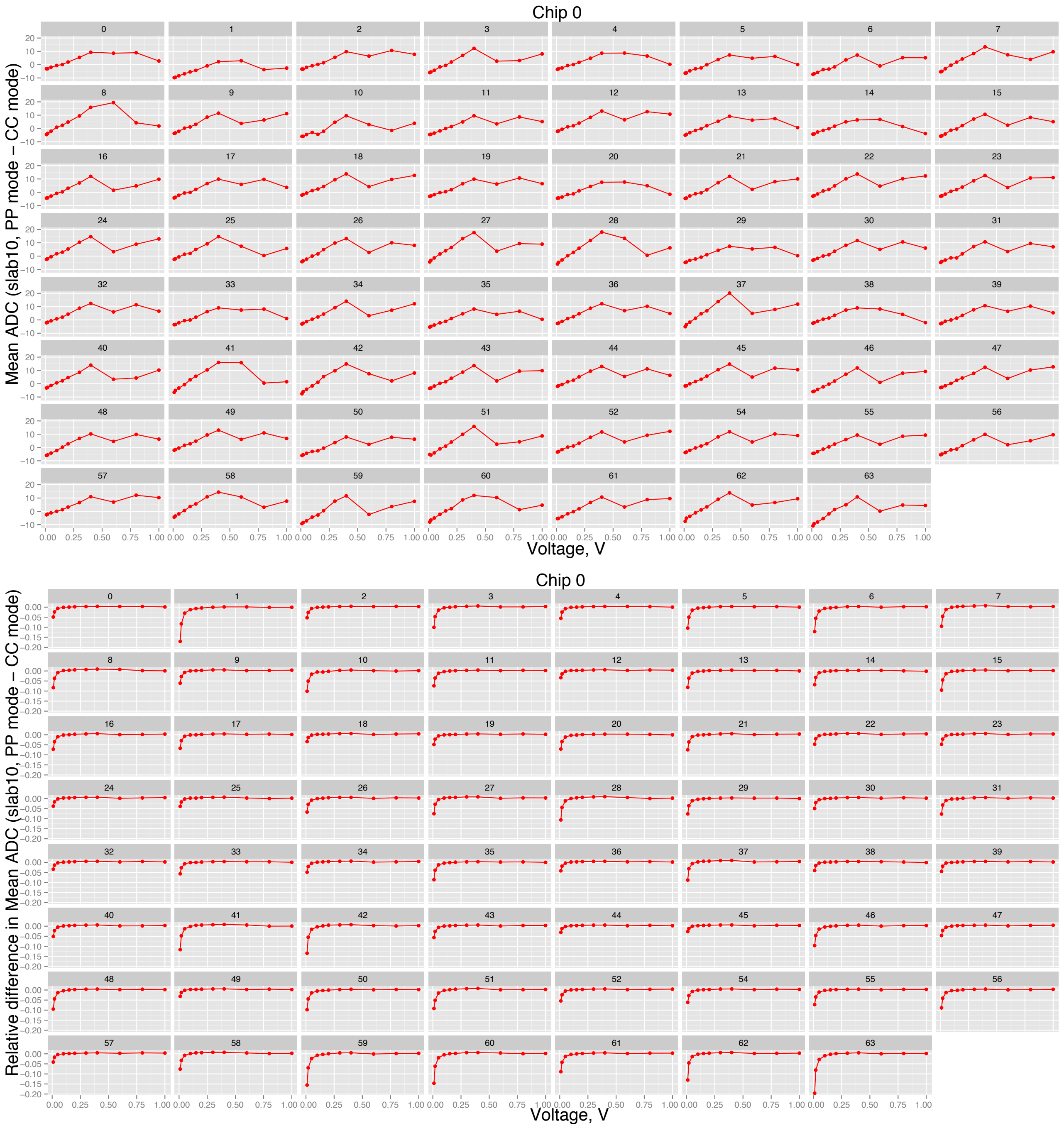


Figure B.9 – Absolute (top) and relative (bottom) difference of the average channel response between PP and CC injection modes in slab 10 (chip 0).

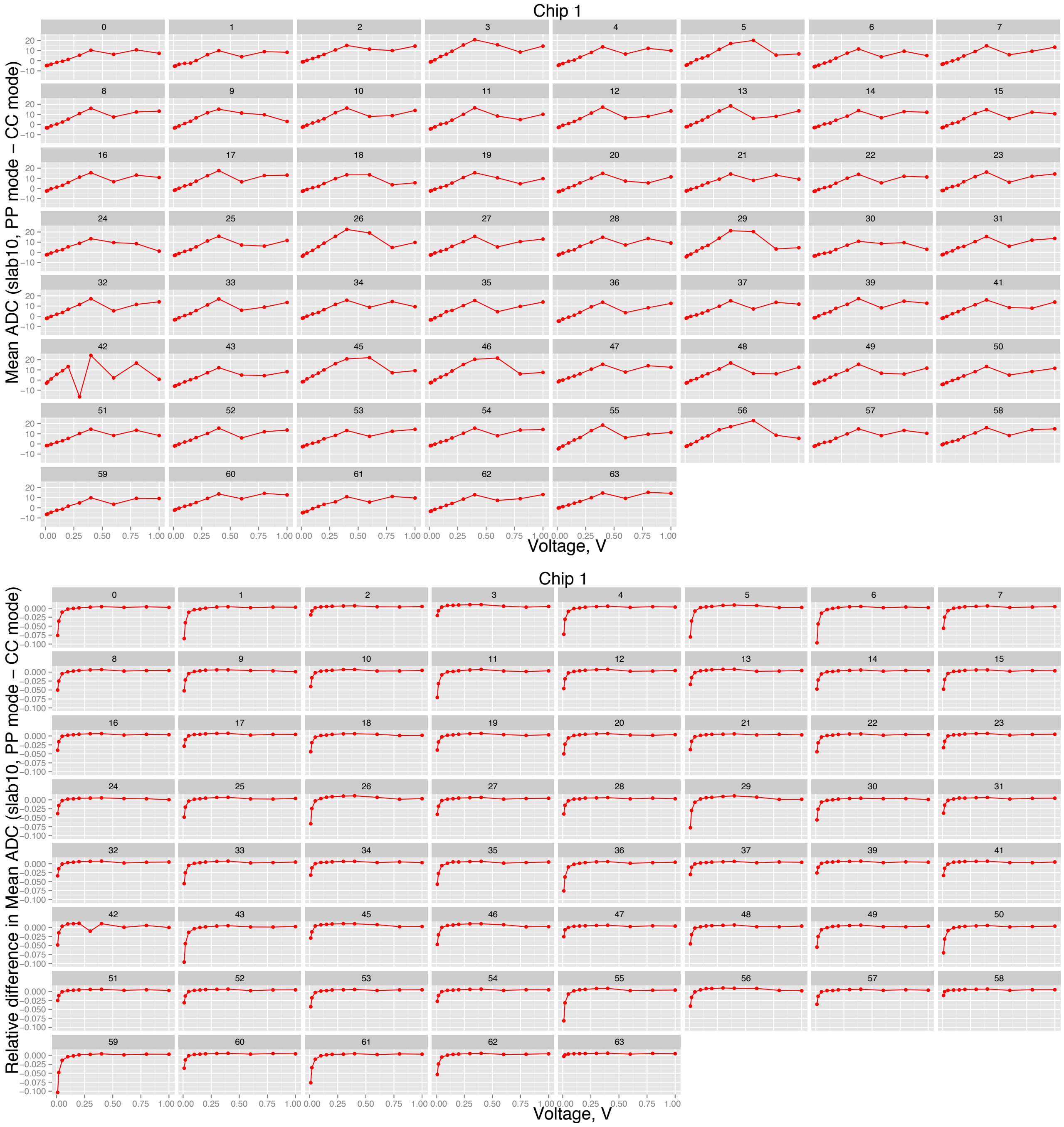


Figure B.10 – Same as in previous Figure B.9, but for chip 1 in slab 10.

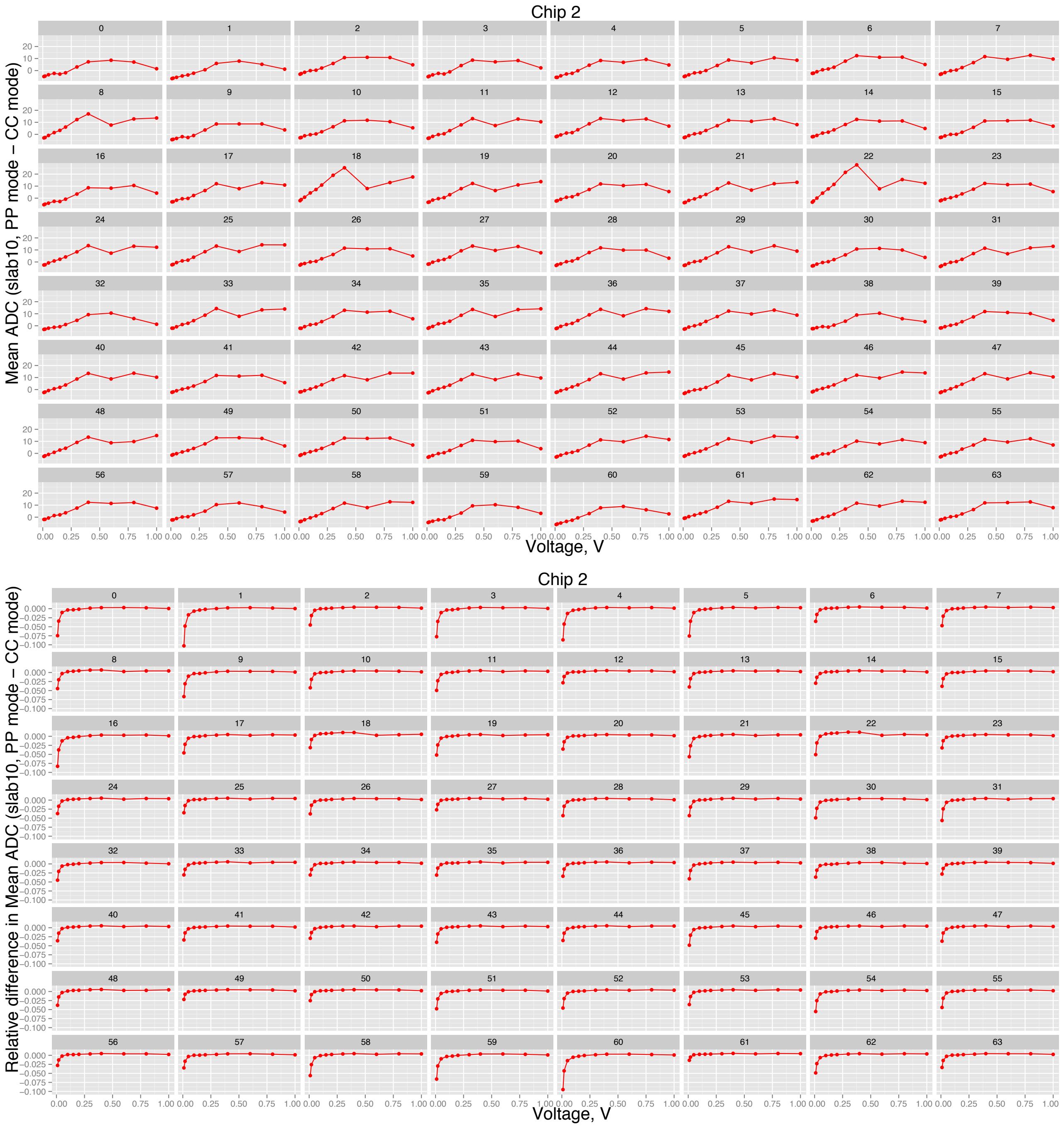


Figure B.11 – Same as in Figure B.9, but for chip 2 in slab 10.

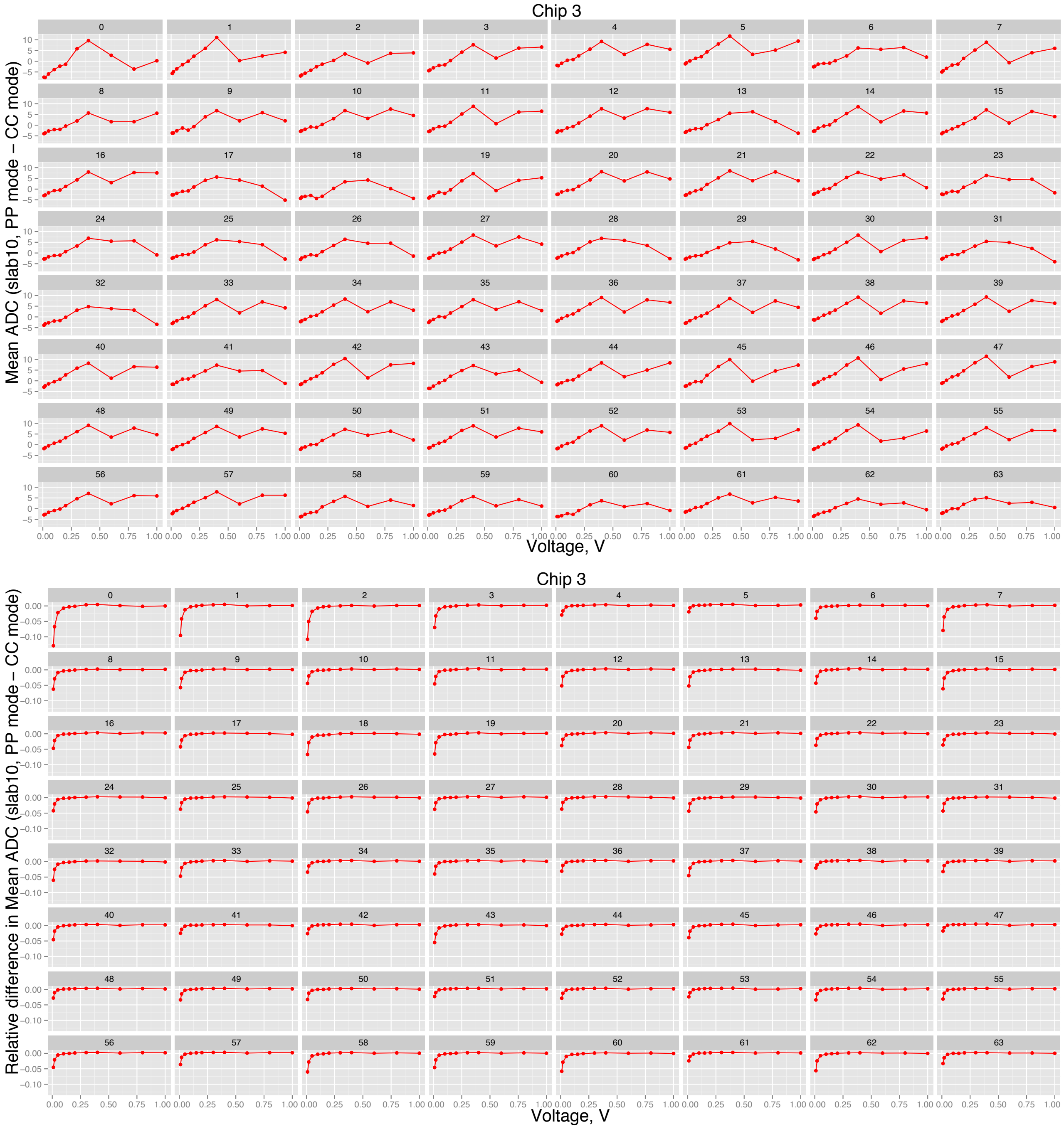


Figure B.12 – Same as in Figure B.9, but for chip 3 in slab 10.

B.2 Muon efficiency for three FEV10/11 layers

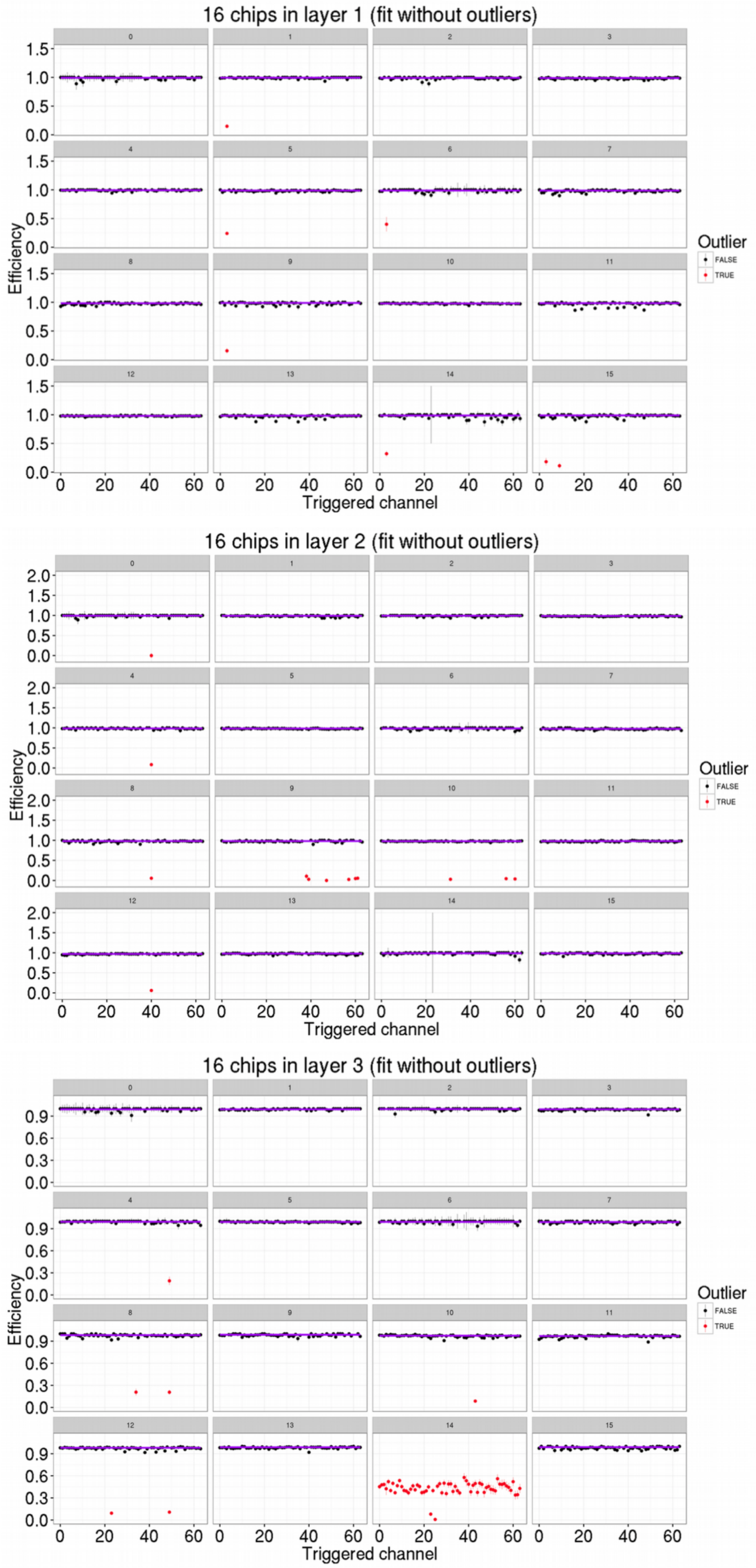


Figure B.13 – Muon efficiency per chip, channel for three FEV10/11 layers.

B.3 "Square"-events in FEV10/11 sensors

B.3.1 "Square"-events for the normal detector position

Figure B.14 – "Square"-events observed in DIF 0 (top) and 2 (bottom) during the November'15 TB setup in run 211 (100 GeV e^+ beam). Active detector layers are perpendicular to the beam.

B.3.2 “Square”-events for the detector parallel to the beam

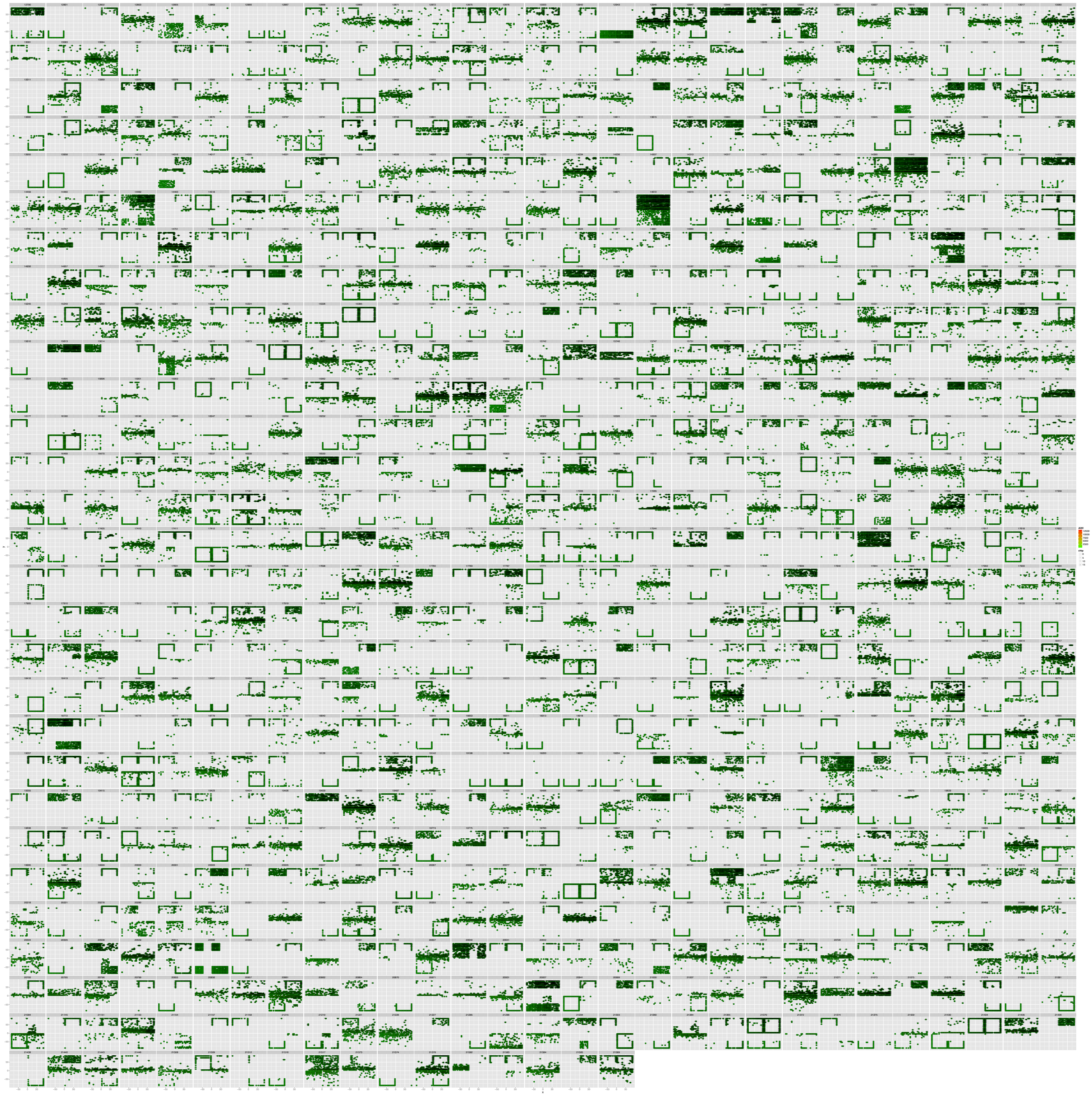


Figure B.15 – Increased rate of the “square”-event effect , when DIF 1 layer was turned by 90° (parallel to the beam), run 472 (150 GeV π^+ beam).

B.4 Delayed triggers in high energy showers

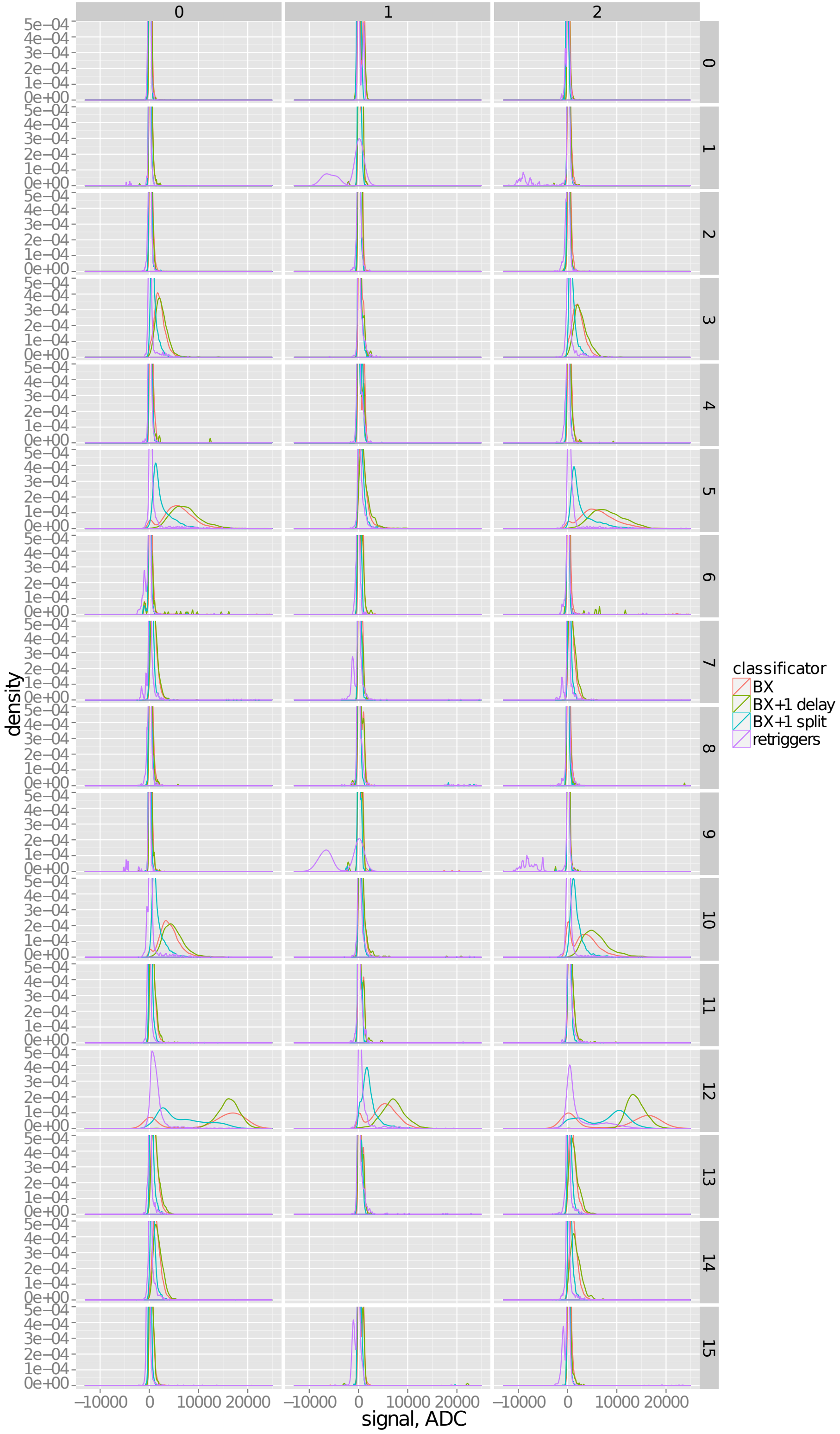


Figure B.16 – Density plot distributions of the collected signals by 16 chips (16 rows) in each layer (3 columns correspond to DIF 0, 1, 2) in run 394 (occupancy map is shown in Figure 5.42). Red lines show distributions of signals in the first physical BX, green lines - in BX+1 for chips with delayed triggers (chip is empty in BX), blue lines - in BX+1 for chips with the split trigger (chip is active in both BX and BX+1) and purple line - all other retriggers (BX+2,...). One can see that split and delayed triggers exist almost in each chip, but the effect rates are different. Maximal rates of these effects are observed in chip 12 in each DIF. DIF 1 and 2 also have significant effect rates in chips 3, 5 and 10. High rates of the split and delayed trigger effects are the reasons of the reduced occupancy in the chip 12 in DIF 1 and 2 where the beam is centered.

B.5 Fractal dimension study

B.5.1 Fractal dimension FD_{12} distributions for e^+ and π^+ samples

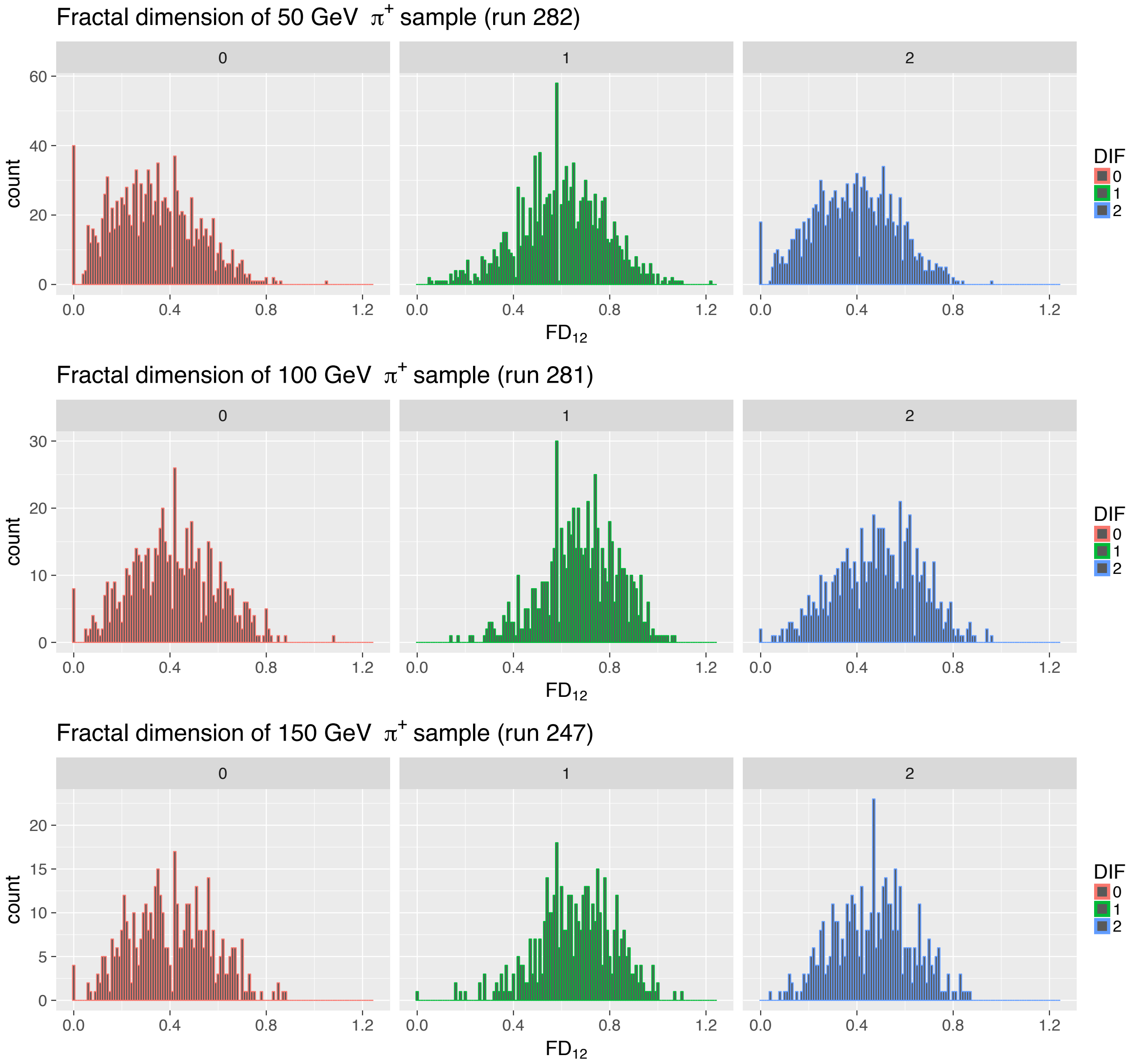


Figure B.17 – The FD_{12} distributions in three active layers (DIF 0, 1, 2) for 50, 100,150 GeV π^+ samples (runs 282, 281, 247).

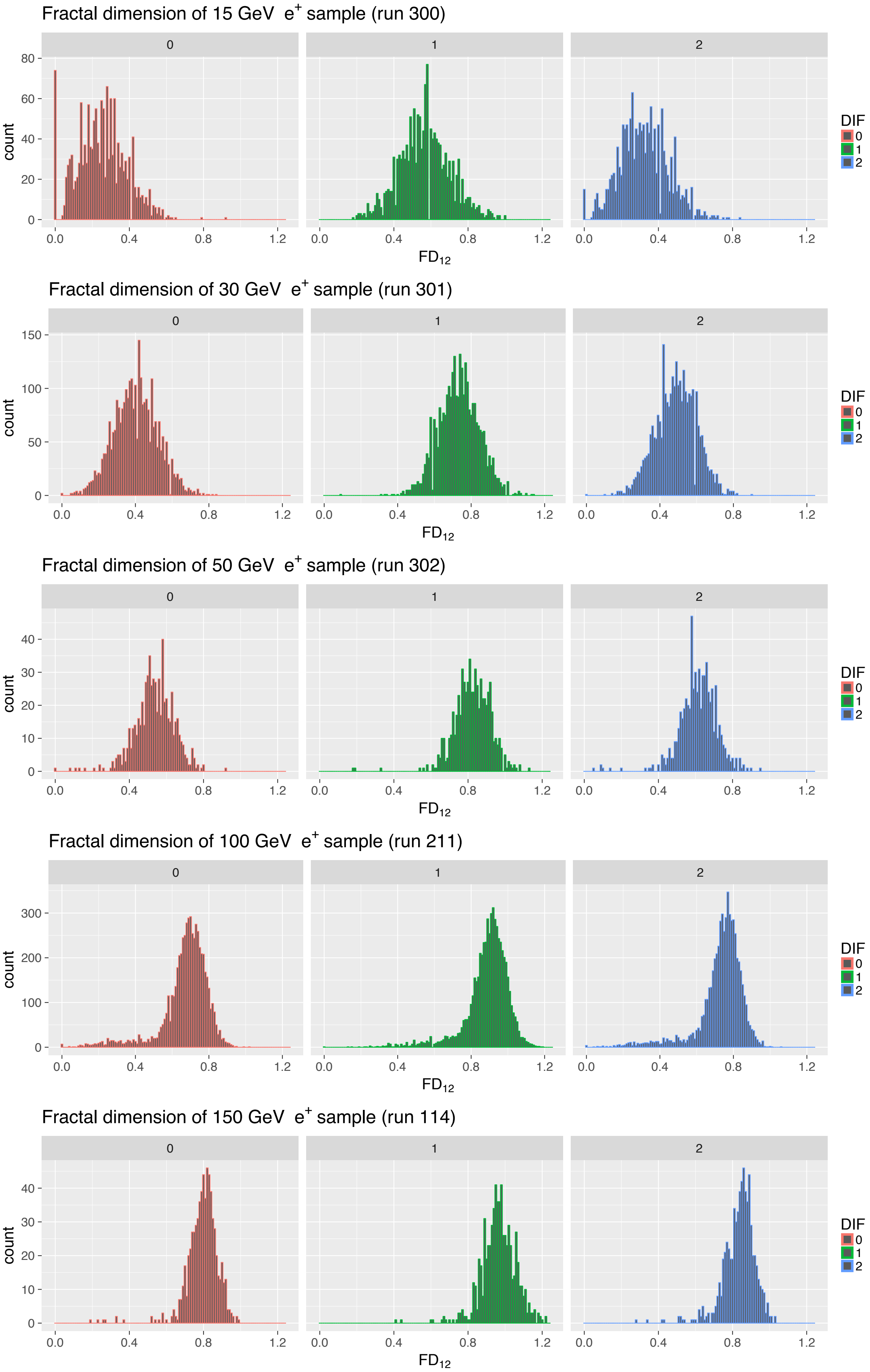


Figure B.18 – Same as in previous Figure B.17, but for 15, 30, 50, 100, 150 GeV e^+ samples (runs 300, 301, 302, 211, 114).

B.5.2 Distribution plots of the FD_{12} versus event energy and number of hits

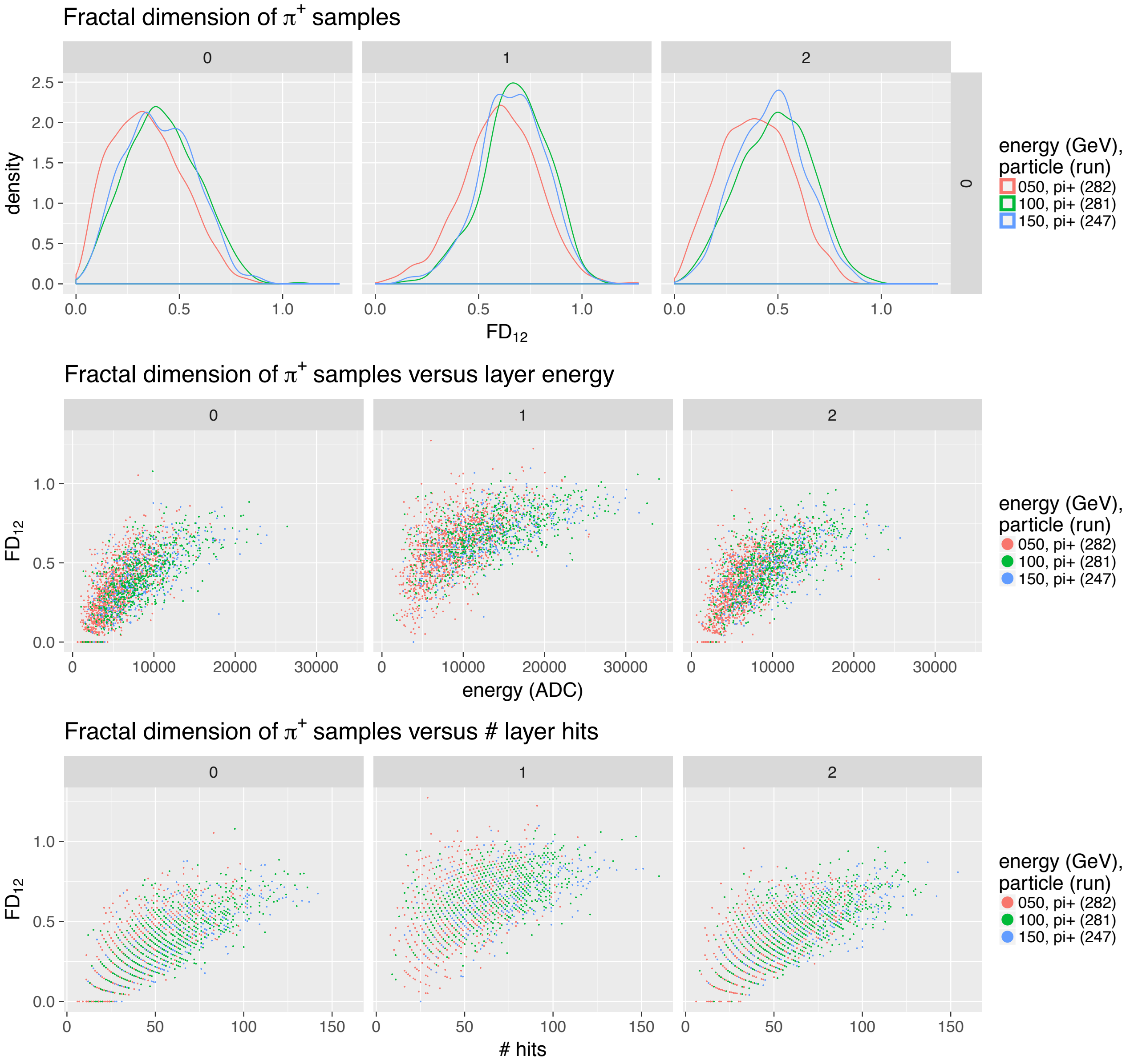
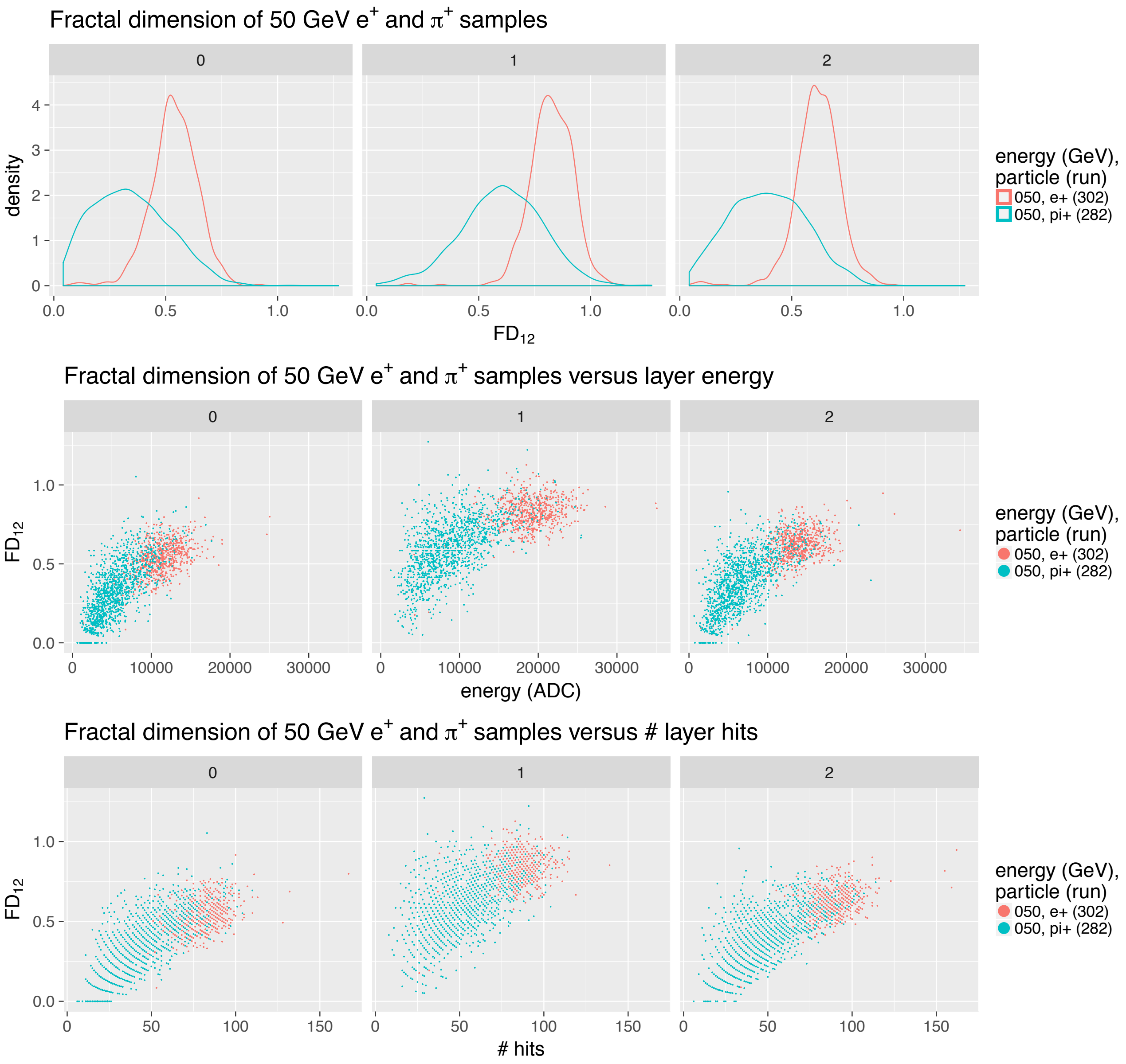


Figure B.19 – FD_{12} density plot distributions for 50, 100, 150 GeV π^+ samples (top row plots; runs 282, 281, 247). FD_2 versus event energy (N hits) distributions are shown in middle row plots (bottom row).



Fractal dimension of 50 GeV e^+ and π^+ samples

Fractal dimension of 50 GeV e^+ and π^+ samples versus layer energy

Fractal dimension of 50 GeV e^+ and π^+ samples versus # layer hits

Figure B.20 – Same as in previous Figure B.19, but for 50 *GeV* e^+ (run 302) and π^+ (run 282) samples.

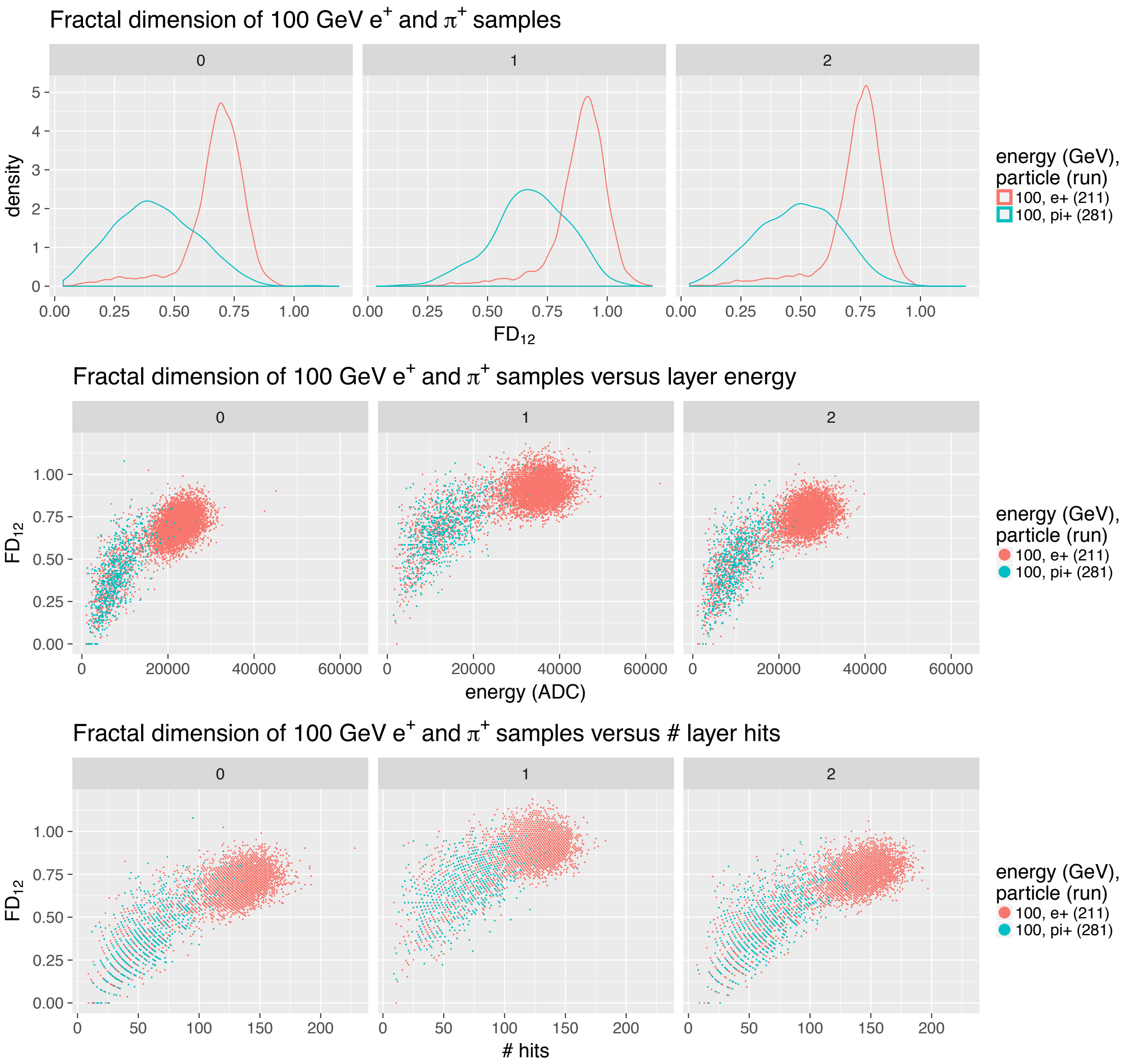


Figure B.21 – Same as in Figure B.19, but for 100 GeV e^+ (run 211) and π^+ (run 281) samples.

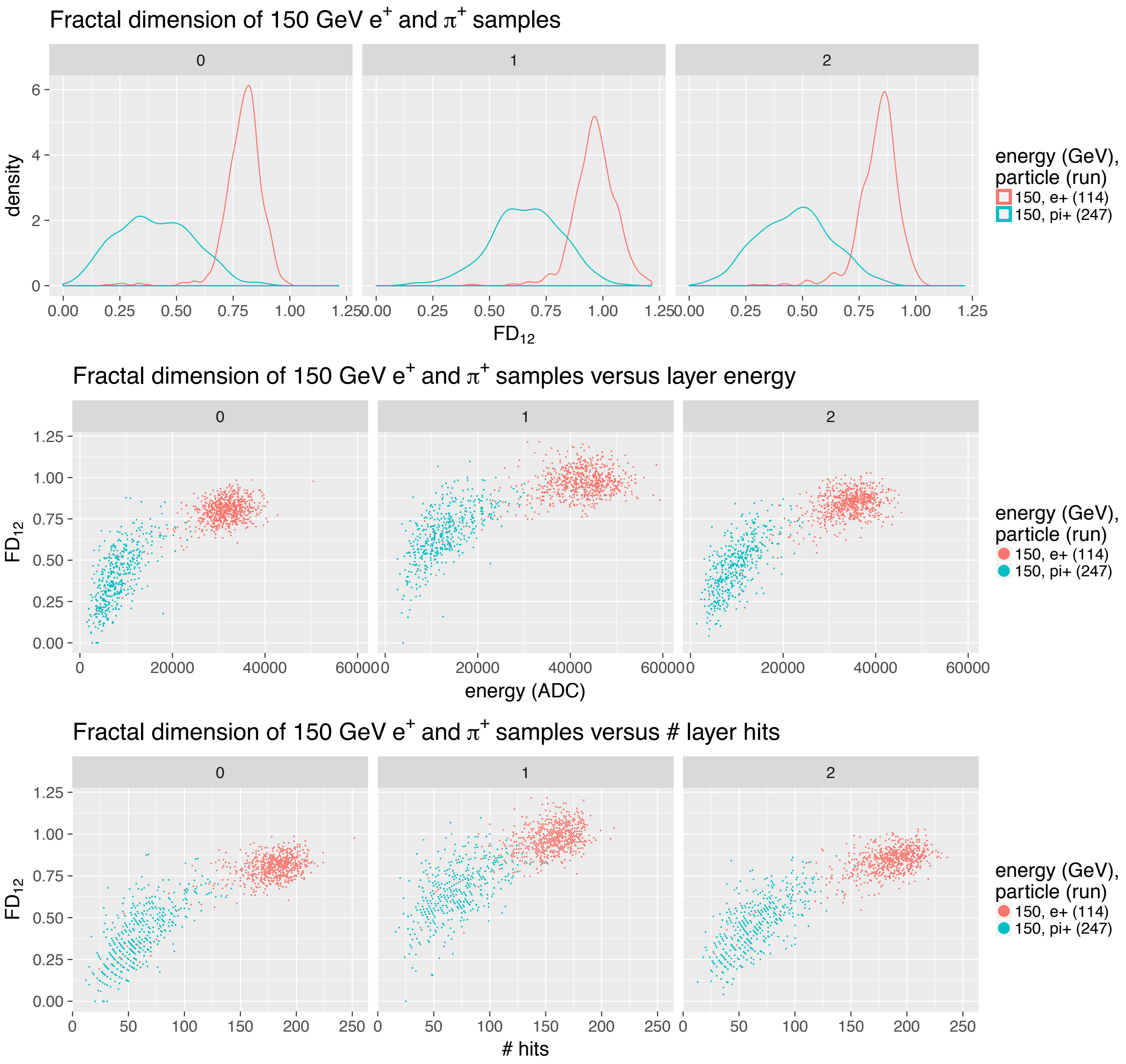


Figure B.22 – Same as in Figure B.19, but for 150 GeV e^+ (run 114) and π^+ (run 247) samples.

Bibliography

[1] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons", Phys. Rev. Lett. **13** 321 (1964).

[2] P. W. Higgs, "Broken symmetries, massless particles and gauge fields", Phys. Lett. **12** 132 (1964).

[3] P. W. Higgs, "Broken symmetries and the masses of gauge bosons", Phys. Rev. Lett. **13** 508 (1964).

[4] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global conservation laws and massless particles", Phys. Rev. Lett. **13** 585 (1964).

[5] S. Chatrchyan *et al.* [CMS Collaboration], "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," Phys. Lett. B **716**, 30 (2012) doi:10.1016/j.physletb.2012.08.021 [arXiv:1207.7235 [hep-ex]].

[6] G. Aad *et al.* [ATLAS Collaboration], "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," Phys. Lett. B **716**, 1 (2012) doi:10.1016/j.physletb.2012.08.020 [arXiv:1207.7214 [hep-ex]].

[7] FCC-ee project website, `http://tlep.web.cern.ch`

[8] F. Zimmermann [FCC team], "FCC-ee Machine Status," presentation at FCC Week 2015, `https://indico.cern.ch/event/340703/contributions/802104/attachments/668692/919126/FCC_ee_machine_overview_final.pdf`

[9] CEPC-SPPC Study Group, "CEPC-SPPC Preliminary Conceptual Design Report. 1. Physics and Detector," IHEP-CEPC-DR-2015-01, IHEP-TH-2015-01, IHEP-EP-2015-01.

[10] T. Behnke *et al.*, "The International Linear Collider Technical Design Report - Volume 1: Executive Summary," arXiv:1306.6327 [physics.acc-ph].

[11] H. Baer *et al.*, "The International Linear Collider Technical Design Report - Volume 2: Physics," arXiv:1306.6352 [hep-ph].

[12] C. Adolphsen *et al.*, "The International Linear Collider Technical Design Report - Volume 3.I: Accelerator & in the Technical Design Phase," arXiv:1306.6353 [physics.acc-ph].

[13] C. Adolphsen *et al.*, "The International Linear Collider Technical Design Report - Volume 3.II: Accelerator Baseline Design," arXiv:1306.6328 [physics.acc-ph].

[14] T. Behnke *et al.*, "The International Linear Collider Technical Design Report - Volume 4: Detectors," arXiv:1306.6329 [physics.ins-det].

[15] XFEL project website, `http://www.xfel.eu`

[16] T. Barklow *et al.*, "Improved Formalism for Precision Higgs Coupling Fits", arXiv:1708.08912.

[17] T. Barklow *et al.*, "Model-Independent Determination of the Triple Higgs Coupling at e+e- Colliders", arXiv:1708.09079.

[18] M. A. Thomson, "Particle Flow Calorimetry and the PandoraPFA Algorithm," Nucl.Instrum.Meth., A611:25–40, (2009) doi:10.1016/j.nima.2009.09.009 [arXiv:0907.3577v1 [physics.ins-det]].

[19] J. C. Brient and H. Videau, "The Calorimetry at the future e+ e- linear collider," eConf C **010630**, E3047 (2001) [hep-ex/0202004].

[20] T. Abe *et al.* [Linear Collider ILD Concept Group], "The International Large Detector: Letter of Intent," doi:10.2172/975166 arXiv:1006.3396 [hep-ex].

[21] CALICE collaboration website, `https://twiki.cern.ch/twiki/bin/view/CALICE/CaliceCollaboration`

[22] K. Kotera, D. Jeans and T. Takeshita, "Particle Flow Algorithm with a strip-scintillator ECAL," presentation and proceedings at CHEF2013, `https://inspirehep.net/record/1295987/files/CHEF2013_Katsushige_Kotera.pdf`

[23] M. Ruan, V. Boudry, J. C. Brient, D. Jeans and H. Videau, "Fractal dimension analysis in a highly granular calorimeter," J. Phys. Conf. Ser. **368**, 012038 (2012). doi:10.1088/1742-6596/368/1/012038

[24] M. Ruan, D. Jeans, V. Boudry, J. C. Brient and H. Videau, "Fractal Dimension of Particle Showers Measured in a Highly Granular Calorimeter," Phys. Rev. Lett. **112**, no. 1, 012001 (2014) doi:10.1103/PhysRevLett.112.012001 [arXiv:1312.7662 [physics.ins-det]].

[25] M. Ruan, "Shower fractal dimension analysis in a highly-granular calorimeter," J. Phys. Conf. Ser. **490**, 012227 (2014). doi:10.1088/1742-6596/490/1/012227

[26] C. Adloff *et al.* [CALICE Collaboration], "Tests of a particle flow algorithm with CALICE test beam data," JINST **6**, P07005 (2011) doi:10.1088/1748-0221/6/07/P07005 [arXiv:1105.3417 [physics.ins-det]].

[27] V. Balagura *et al.*, "SiW ECAL for future e^+e^- collider," JINST **12**, no. 07, C07013 (2017) doi:10.1088/1748-0221/12/07/C07013 [arXiv:1705.10838 [physics.ins-det]].

[28] K. Shpak, on behalf of the CALICE/ILD SiW ECAL group, "The ILD/CALICE Silicon-Tungsten Electromagnetic Calorimeter: status and potential," PoS(ICHEP2016)1085 `https://pos.sissa.it/282/1085/`

[29] K. Shpak, on behalf of the ILD SiW ECAL group, "Technological Prototype of the CALICE / ILD Silicon-Tungsten Electromagnetic Calorimeter," IEEE2016 NSS/MIC conference records, N44-2

[30] K. Shpak, V. Balagura, "Separation of two overlapped electromagnetic or electromagnetic-hadronic showers in CALICE highly granular physics calorimeter prototypes using Pandora, Garlic and Arbor Particle Flow Algorithms," CALICE Analysis Note CAN-057, `https://twiki.cern.ch/twiki/pub/CALICE/CaliceAnalysisNotes/CAN-057.pdf`

[31] M. Aicheler *et al.*, "A Multi-TeV Linear Collider Based on CLIC Technology : CLIC Conceptual Design Report," doi:10.5170/CERN-2012-007

[32] P. Lebrun *et al.*, "The CLIC Programme: Towards a Staged e+e- Linear Collider Exploring the Terascale : CLIC Conceptual Design Report," doi:10.5170/CERN-2012-005 arXiv:1209.2543 [physics.ins-det].

[33] L. Linssen, A. Miyamoto, M. Stanitzki and H. Weerts, "Physics and Detectors at CLIC: CLIC Conceptual Design Report," doi:10.5170/CERN-2012-003 arXiv:1202.5940 [physics.ins-det].

[34] K. Fujii *et al.*, "Physics Case for the 250 GeV Stage of the International Linear Collider", arXiv:1710.07621.

[35] D. M. Kaplan [MAP and MICE Collaborations], "Muon Colliders and Neutrino Factories," EPJ Web Conf. **95**, 03019 (2015) doi:10.1051/epjconf/20159503019 [arXiv:1412.3487 [physics.acc-ph]].

[36] G. Alexander *et al.*, "Undulator-Based Production of Polarized Positrons," Nucl. Instrum. Meth. A **610**, 451 (2009) doi:10.1016/j.nima.2009.07.091 [arXiv:0905.3066 [physics.ins-det]].

[37] M. Klute, R. Lafaye, T. Plehn, M. Rauch and D. Zerwas, "Measuring Higgs Couplings at a Linear Collider," Europhys. Lett. **101**, 51001 (2013) doi:10.1209/0295-5075/101/51001 [arXiv:1301.1322 [hep-ph]].

[38] S. Blyweert [ATLAS and CMS Collaborations], "Top-quark mass measurements at the LHC," arXiv:1205.2175 [hep-ex].

[39] K. Seidel, F. Simon and M. Tesař, "Prospects for the Measurement of the Top Mass in a Threshold Scan at CLIC and ILC," LCD-Note-2012-013, `https://cds.cern.ch/record/1498599/files/LCD-2012-013.pdf`

[40] J. Reuter, "The BSM Physics Case of the ILC," arXiv:1602.08352 [hep-ph].

[41] H. Videau and J. C. Brient, "Calorimetry optimised for jets,"

[42] V. L. Morgunov, "Calorimetry design with energy-flow concept (imaging detector for high-energy physics),"

[43] I. G. Knowles and G. D. Lafferty, "Hadronization in Z^0 decay," J. Phys. G **23**, 731 (1997) doi:10.1088/0954-3899/23/7/003 [hep-ph/9705217].

[44] J. S. Marshall and M. A. Thomson, "Pandora Particle Flow Algorithm," arXiv:1308.4537 [physics.ins-det].

[45] H. L. Tran, K. Krüger, F. Sefkow, S. Green, J. Marshall, M. Thomson and F. Simon, "Software compensation in Particle Flow reconstruction," arXiv:1705.10363 [physics.ins-det].

[46] R. M. Sternheimer, "Energy Distribution of γ Rays from π^0 Decay," Phys. Rev. **99**, 277 (1955) doi:10.1103/PhysRev.99.277.

[47] Three-jet event in ALEPH detector, CERN, http://www.sciencephoto.com/media/1568/view

[48] ALEPH experiment, https://home.cern/about/experiments/aleph

[49] M. Anduze, D. Grondin and H. Videau, "Silicon Tungsten electromagnetic calorimeter "ECal" Technical Design Document," https://llrbox.in2p3.fr/owncloud/index.php/s/S0WrOA6CzINiFwZ

[50] J. E. Augustin *et al.*, "A silicon envelope for the TPC," LC-DET-2003-013.

[51] A. Savoy-Navarro, "Large Area Silicon Tracking: New Perspectives," arXiv:1203.0736 [physics.ins-det].

[52] A. Savoy-Navarro, "Overview of the SiLC R and D Activities," arXiv:0903.4361 [physics.ins-det].

[53] LCTPC collaboration, https://www.lctpc.org

[54] D. Attié *et al.* [LCTPC Collaboration], "A Time Projection Chamber with GEM-Based Readout," Nucl. Instrum. Meth. A **856**, 109 (2017) doi:10.1016/j.nima.2016.11.002 [arXiv:1604.00935 [physics.ins-det]].

[55] D. Jeans [ILD SiW ECAL groups], "The silicon-tungsten electromagnetic calorimeter for ILD," presentation and proceedings at CHEF2013, https://inspirehep.net/record/1286395/files/CHEF2013_Daniel_Jeans.pdf

[56] S. Callier, F. Dulucq, C. de La Taille, G. Martin-Chassard and N. Seguin-Moreau, "SKIROC2, front end chip designed to readout the Electromagnetic CALorimeter at the ILC," JINST **6**, C12040 (2011). doi:10.1088/1748-0221/6/12/C12040

[57] T. Behnke *et al.*, "TESLA Technical Design Report - PART IV: A Detector for TESLA," http://tesla.desy.de/new_pages/TDR_CD/PartIV/partIV.pdf

[58] J. S. Marshall, "ECAL Optimisation Studies," presentation at CLIC Workshop (4th February 2014), https://indico.cern.ch/event/275412/contributions/1617692/attachments/498830/689108/ECAL_marshall_04.02.14.pdf,
J. S. Marshall, "ECAL Optimisation Studies," presentation at CLICdp Detector Optimisation Meeting (23rd April 2014), https://indico.cern.ch/event/309926/contributions/715740/attachments/592599/815653/ECAL_marshall_23.04.14.pdf

[59] SiW ECAL optimization studies:
T. H. Tran, "SiW ECAL + SDHCAL dimention performance optimisation," presentation at LCWS2013, https://agenda.linearcollider.org/event/6000/contributions/27852/attachments/23124/35995/tran_ehcal_opt.pdf,
T. H. Tran, "SiW ECAL optimisation in simulation," presentation at ILD meeting (September 2013), https://agenda.linearcollider.org/event/6113/contributions/28460/attachments/23639/36685/ECALopt_TrongHieuTRAN.pdf,
T. H. Tran, "ILD cost-performance optimisation," presentation at JCL2013, https://indico.cern.ch/event/274076/contributions/1615290/attachments/494286/683018/tran.JCL.ILDopt.pdf

[60] C. Adloff *et al.* [CALICE Collaboration], "Construction and Commissioning of the CALICE Analog Hadron Calorimeter Prototype," JINST **5**, P05004 (2010) doi:10.1088/1748-0221/5/05/P05004 [arXiv:1003.2662 [physics.ins-det]].

[61] G. Baulieu *et al.*, "Construction and commissioning of a technological prototype of a high-granularity semi-digital hadronic calorimeter," JINST **10**, no. 10, P10039 (2015) doi:10.1088/1748-0221/10/10/P10039 [arXiv:1506.05316 [physics.ins-det]].

[62] AHCAL optimization studies:
S. Green, J. S. Marshall, "PandoraPFA and AHCal Optimisation Studies," presentation at ILD Software and Optimisation Workshop (February 2016), https://agenda.linearcollider.org/event/6954/contributions/34302/attachments/28264/42734/ILDWorkshop_Pandora_Green_24-2-16.pdf,

K. Krüger, F. Sefkow, H. L. Tran "AHCAL Optimisation with Pandora," presentation at CALICE collaboration meeting (March 2016), http://pubdb.xfel.eu/record/316159/files/Lan_Optimisation.pdf,
H. L. Tran, "AHCAL Optimisation Study," presentation at ALCWS2015, https://agenda.linearcollider.org/event/6557/contributions/31731/attachments/26164/40108/Lan_HCALOpti.pdf,
H. L. Tran, "Software Compensation for AHCAL optimisation," FLC Long Talk presentation (30th November 2015) http://www.desy.de/~huonglan/FLCLongTalk_30Nov2015/Lan_FLCLongTalk.pdf

[63] SDHCAL optimization studies:
G. Grenier, "SDHCAL Optimisation strategy," presentation at ILD meeting (Sepember 2013), https://agenda.linearcollider.org/event/6113/contributions/28444/attachments/23626/36671/SDHCAL-GRPC-ILD2013.pdf,
I. Laktineh, "Semi-Digital Hadronic CALorimeter for ILD," presentation at ILD workshop (May 2011), https://agenda.linearcollider.org/event/4901/contributions/20346/attachments/16552/26957/ILD-SDHCAL-IL.pdf,
T. H. Tran, "SiW ECAL + sDHCAL dimension-performance optimisation," presentation at LCWS2013, https://agenda.linearcollider.org/event/6000/contributions/27852/attachments/23124/35995/tran_ehcal_opt.pdf

[64] H. Abramowicz *et al.*, "Forward Instrumentation for ILC Detectors," JINST **5**, P12002 (2010) doi:10.1088/1748-0221/5/12/P12002 [arXiv:1009.2433 [physics.ins-det]].

[65] V. D'Ascenso, N. Saveliev, "Mathematical modeling and Study of the ILD Muon System/ Tail Catcher," http://www-flc.desy.de/lcnotes. LC-DET-2012- xxx.

[66] F. Kircher, U. Schneekloth, *et al.*, "Conceptual Design of the ILD Detector Magnet System," LC-DET-2012-081 http://flc.desy.de/lcnotes/noteslist/localfsExplorer_read?currentPath=/afs/desy.de/group/flc/lcnotes/LC-DET-2012-081.pdf

[67] [CALICE Collaboration], "Construction and performance of a silicon photomultiplier/extruded scintillator tail-catcher and muon-tracker," JINST **7**, P04015 (2012) doi:10.1088/1748-0221/7/04/P04015 [arXiv:1201.1653 [physics.ins-det]].

[68] ILCsoft, http://ilcsoft.desy.de/portal, https://github.com/iLCSoft

[69] LCIO (Linear Collider I/O), http://lcio.desy.de

[70] Geant4 (for GEometry ANd Tracking), http://geant4.cern.ch

[71] Mokka: geometry driver developers, http://ilcsoft.desy.de/v01-17-09/Mokka/mokka-08-05-01/doc/Kernel/geometry_drivers.html

[72] Gear (GEometry Api for Reconstruction), http://ilcsoft.desy.de/portal/software_packages/gear/index_eng.html

[73] D. H. Wright and M. H. Kelsey, "The Geant4 Bertini Cascade," Nucl. Instrum. Meth. A **804**, 175 (2015). doi:10.1016/j.nima.2015.09.058

[74] Marlin (Modular Analysis & Reconstruction for the LINear collider), http://ilcsoft.desy.de/portal/software_packages/marlin/index_eng.html

[75] CALICEsoft, https://twiki.cern.ch/twiki/bin/view/CALICE/SoftwareMain

[76] M. Terwort *et al.* [CALICE Collaboration], "The CALICE Software Framework and Operational Experience," doi:10.1109/NSSMIC.2010.6036245 arXiv:1011.4765 [physics.ins-det].

[77] ILD geometry models for simulations, https://github.com/iLCSoft/lcgeo/tree/master/ILD/compact

[78] F. Gaede, S. Aplin, R. Glattauer, C. Rosemann and G. Voutsinas, "Track reconstruction at the ILC: the ILD tracking software," J. Phys. Conf. Ser. **513**, 022011 (2014). doi:10.1088/1742-6596/513/2/022011

[79] F. Gaede, "Tracking in iLCSoft: Focus on tracking for ILD," presentation at VTX Meeting (May 2017) https://agenda.linearcollider.org/event/7450/contributions/38569/attachments/31566/47551/gaede_ilcsoft_tracking.pdf

[80] D. Jeans, O. Hartbrich, "Realistic calorimeter hit digitisation in the ILDCaloDigi processor," LC-TOOL-2014-011 https://flc.desy.de/lcnotes/noteslist/localfsExplorer_read?currentPath=/afs/desy.de/group/flc/lcnotes/LC-TOOL-2014-011.pdf

[81] NewLDCCaloDigi digitization processor for Marlin, `http://polywww.in2p3.fr/~ruan/DHCALDigitization/DHCALDigi/include/RPCDHCALCaloDigi.h`

[82] Z. Deng *et al.* [CALICE Collaboration], "Resistive Plate Chamber Digitization in a Hadronic Shower Environment," JINST **11**, no. 06, P06014 (2016) doi:10.1088/1748-0221/11/06/P06014 [arXiv:1604.04550 [physics.ins-det]].

[83] B. Xu, "Improvement of photon reconstruction in PandoraPFA," arXiv:1603.00013 [physics.ins-det].

[84] M. Ruan and H. Videau, "Arbor, a new approach of the Particle Flow Algorithm," arXiv:1403.4784 [physics.ins-det].

[85] R. Ete, "Separation of nearby hadronic showers in the CALICE SDHCAL prototype detector using ArborPFA," CALICE Analysis Note CAN-054, `https://twiki.cern.ch/twiki/pub/CALICE/CaliceAnalysisNotes/CAN-054.pdf`

[86] D. Jeans, J. C. Brient and M. Reinhard, "GARLIC: GAmma Reconstruction at a LInear Collider experiment," JINST **7**, P06003 (2012) doi:10.1088/1748-0221/7/06/P06003 [arXiv:1203.0774 [physics.ins-det]].

[87] D. Jeans, "Photon and π^0 identification: recent GARLIC developments," presentation at HGC4ILD (February 2015), `https://indico.in2p3.fr/event/11192/session/2/contribution/40/material/slides/0.pdf`

[88] P. Zarchan, H. Musoff, "Fundamentals of Kalman Filtering: A Practical Approach," American Institute of Aeronautics and Astronautics, Incorporated. ISBN 978-1-56347-455-2.

[89] GitHub: "Multi-algorithm pattern recognition PandoraPFA," `https://github.com/PandoraPFA`

[90] J. S. Marshall, M. A. Thomson, "Pandora LC Reconstruction," presentation at LCWS2015, `https://twiki.cern.ch/twiki/pub/CLIC/LCWS2015/Marshall_Pandora_LCWS2015.pdf`

[91] K. Shpak, "Update on particle separation CAN057," presentation at CALICE collaboration meeting (September 2016), `https://agenda.linearcollider.org/event/7304/contributions/37325/attachments/30521/45658/CALICE_Arlington_Shpak_v3.pdf`

[92] Private communications with Boruo Xu from Cambridge Pandora group

[93] T. Suehara, "Analysis of Tau-pair process in the ILD reference detector model," arXiv:0909.2398 [hep-ex].

[94] T. Suehara and J. List, "Chargino and Neutralino Separation with the ILD Experiment," arXiv:0906.5508 [hep-ex].

[95] T. Suehara, A. Miyamoto, K. Fujii, N. Okada, H. Ito, K. Ikematsu and R. Yonamine, "Tau-pair Performance in ILD detectors," arXiv:0902.3740 [hep-ex].

[96] C. Bartels and J. List, "Model independent WIMP Searches in full Simulation of the ILD Detector," arXiv:1007.2748 [hep-ex].

[97] K. Ikematsu, A. Miyamoto and K. Fujii, "Feasibility study of the forward-backward asymmetry of the $e^+e^- \to t\bar{t}$ process in all-hadronic decay modes at $\sqrt{s} =$ 500-GeV with the ILD detector," arXiv:1005.4635 [hep-ex].

[98] O. Kittel, G. Moortgat-Pick, K. Rolbiecki, P. Schade and M. Terwort, "Measurement of CP asymmetries in neutralino production at the ILC," Eur. Phys. J. C **72**, 1854 (2012) doi:10.1140/epjc/s10052-011-1854-4 [arXiv:1108.3220 [hep-ph]].

[99] D. Yu, M. Ruan, V. Boudry and H. Videau, "Lepton identification at particle flow oriented detector for the future e^+e^- Higgs factories," arXiv:1701.07542 [physics.ins-det].

[100] M. Ruan, "Calorimeter optimization at the CEPC

[101] Simulation," presentation at CALICE collaboration meeting (March 2017), `https://agenda.linearcollider.org/event/7454/contributions/38805/attachments/31385/47212/CEPC-Opti.pdf`

[102] T. H. Tran, V. Balagura, V. Boudry, J. C. Brient and H. Videau, "Reconstruction and classification of tau lepton decays with ILD," Eur. Phys. J. C **76**, no. 8, 468 (2016) doi:10.1140/epjc/s10052-016-4315-2 [arXiv:1510.05224 [physics.ins-det]].

[103] D. Jeans, "Tau lepton reconstruction at collider experiments using impact parameters," Nucl. Instrum. Meth. A **810**, 51 (2016) doi:10.1016/j.nima.2015.11.030 [arXiv:1507.01700 [hep-ex]].

[104] C. Adloff [CALICE Collaboration], "CALICE Report to the DESY Physics Research Committee, April 2011," arXiv:1105.0511 [physics.ins-det].

[105] F. Sefkow, A. White, K. Kawagoe, R. Pöschl and J. Repond, "Experimental Tests of Particle Flow Calorimetry," Rev. Mod. Phys. **88**, 015003 (2016) doi:10.1103/RevModPhys.88.015003 [arXiv:1507.05893 [physics.ins-det]].

[106] D. Contardo, M. Klute, J. Mans, L. Silvestris and J. Butler, "Technical Proposal for the Phase-II Upgrade of the CMS Detector," CERN-LHCC-2015-010, LHCC-P-008, CMS-TDR-15-02.

[107] R. Rusack, "High Granularity Calorimeter for the CMS Endcap at HL-LHC," CMS-CR-2016-444

[108] K. Einsweiler, L. Pontecorvo, D. Fido, "ATLAS Phase-II Upgrade Scoping Document," CERN-LHCC-2015-020. LHCC-G-166

[109] C. Adloff *et al.* [CALICE Collaboration], "Validation of GEANT4 Monte Carlo Models with a Highly Granular Scintillator-Steel Hadron Calorimeter," JINST **8**, 07005 (2013) doi:10.1088/1748-0221/8/07/P07005 [arXiv:1306.3037 [physics.ins-det]].

[110] B. Bilki *et al.* [CALICE Collaboration], "Testing hadronic interaction models using a highly granular silicon–tungsten calorimeter," Nucl. Instrum. Meth. A **794**, 240 (2015) doi:10.1016/j.nima.2015.05.009 [arXiv:1411.7215 [physics.ins-det]].

[111] J. Repond *et al.* [CALICE Collaboration], "Design and Electronics Commissioning of the Physics Prototype of a Si-W Electromagnetic Calorimeter for the International Linear Collider," JINST **3**, P08001 (2008) doi:10.1088/1748-0221/3/08/P08001 [arXiv:0805.4833 [physics.ins-det]].

[112] R. Cornat "FKPPL technical activities - Si-W ECAL Silicon wafers status," presentation at FKPPL Workshop (February 2009), `https://indico.in2p3.fr/event/1658/contributions/20946/attachments/17085/`

[113] M. Reinhard, "SiW ECAL: the status before Fermilab," presentation at CALICE collaboration meeting (March 2008) `https://agenda.linearcollider.org/event/2553/contributions/7576/attachments/5732/9638/MReinhard_180308.pdf`

[114] M. Reinhard, "SiW ECAL," presentation at LCWS2008, `https://agenda.linearcollider.org/event/2628/contributions/7955/attachments/6058/10165/MReinhard_LCWS08_SiWECAL.pdf`

[115] M. Reinhard "CP violation in the Higgs sector with a next-generation detector at the ILC," PhD thesis, `https://twiki.cern.ch/twiki/pub/CALICE/CaliceTheses/MR_thesis_final.pdf`

[116] D. R. Ward "Study of the response of the CALICE silicon-tungsten electromagnetic calorimeter prototype to electrons," CALICE Analysis Note 001, `https://twiki.cern.ch/twiki/pub/CALICE/CaliceAnalysisNotes/CAN-001.pdf`

[117] C. Adloff *et al.* [CALICE Collaboration], "Response of the CALICE Si-W electromagnetic calorimeter physics prototype to electrons," Nucl. Instrum. Meth. A **608**, 372 (2009) doi:10.1016/j.nima.2009.07.026 [arXiv:0811.2354 [physics.ins-det]].

[118] P. Buzhan *et al.*, "Silicon photomultiplier and its possible applications," Nucl. Instrum. Meth. A **504**, 48 (2003). doi:10.1016/S0168-9002(03)00749-6

[119] K. Francis *et al.* [CALICE Collaboration], "Performance of the first prototype of the CALICE scintillator strip electromagnetic calorimeter," Nucl. Instrum. Meth. A **763**, 278 (2014) doi:10.1016/j.nima.2014.06.039 [arXiv:1311.3761 [physics.ins-det]].

[120] D. Jeans, "The scintillator ECAL beam test at DESY, 2007 - First results," CALICE analysis note CAN-005, `https://twiki.cern.ch/twiki/pub/CALICE/CaliceAnalysisNotes/CAN-005.pdf`
D. Jeans, "The scintillator ECAL beam test at DESY, 2007 - Update 1," CALICE analysis note CAN-006, `https://twiki.cern.ch/twiki/pub/CALICE/CaliceAnalysisNotes/CAN-006.pdf`
D. Jeans, "The scintillator ECAL beam test at DESY, 2007 - Update 2," CALICE analysis note CAN-007, `https://twiki.cern.ch/twiki/pub/CALICE/CaliceAnalysisNotes/CAN-007.pdf`
D. Jeans, "The scintillator ECAL beam test at DESY, 2007 - Update 3," CALICE analysis note CAN-012, `https://twiki.cern.ch/twiki/pub/CALICE/CaliceAnalysisNotes/CAN-012.pdf`

[121] K. Kotera, “First Stage Analysis of the Energy Response and Resolution of the Scintillator ECAL in the Beam Test at FNAL, 2008,” CALICE analysis note CAN-016, `https://twiki.cern.ch/twiki/pub/CALICE/CaliceAnalysisNotes/CAN-016.pdf`
A. Khan, “Addendum A: Temperature Correction on ADC/MIP Conversion Factors for September 2009 Data of ScECAL FNAL Test Beam Module,” CALICE analysis note CAN-016 Addendum A, `https://twiki.cern.ch/twiki/pub/CALICE/CaliceAnalysisNotes/CAN-016a.pdf`
K. Kotera, “Addendum B: Update of the Analysis of the Test Beam Experiment of the CALICE ScECAL Physics Prototype,” CALICE analysis note CAN-016 Addendum B, `https://twiki.cern.ch/twiki/pub/CALICE/CaliceAnalysisNotes/CAN-016b.pdf`
K. Kotera, “Addendum C: Update of the Analysis of the Test Beam Experiment of the CALICE ScECAL Physics Prototype,” CALICE analysis note CAN-016 Addendum C, `https://twiki.cern.ch/twiki/pub/CALICE/CaliceAnalysisNotes/CAN-016c.pdf`

[122] K. Kotera [CALICE Collaboration], “Scintillator Strip ECAL Optimization,” arXiv:1404.1672 [physics.ins-det].

[123] S. Uozumi, “R&D status of the Scintillator-strip based ECAL for the ILD,” presentation at LCWS2014, `https://agenda.linearcollider.org/event/6389/contributions/30707/attachments/25372/39052/satoru_LCWS14_v1.pdf`

[124] T. Price, “MAPS technology for Vertexing, Tracking, and Calorimetry”, Physics Procedia **37**, 932 (2012).

[125] C. Zhang *et al.* [ALICE FoCal Group], “FoCal – A high granularity electromagnetic calorimeter for forward direct photon measurements” Nucl. Instrum. Meth. A **845**, 542 (2017).

[126] G. Nooren *et al.* [ALICE FoCal Group], “The FoCal prototype – an extremely fine-grained electromagnetic calorimeter using CMOS pixel sensors”, arXiv:1708.05164 [physics.ins-det].

[127] A. Brogna *et al.*, “PHASE-1 User Manual”, http://www.iphc.cnrs.fr/IMG/PH1-UserMan-20080916.pdf.

[128] T. Price, “Status of DECAL,” presentation at CALICE collaboration meeting (September 2016), `https://agenda.linearcollider.org/event/7304/contributions/37340/attachments/30515/45647/tprice_decal_uta16.pdf`

[129] T. Price *et al.*, “Update on DECAL studies for future experiments,” presentation at CALICE collaboration meeting (March 2017), `https://agenda.linearcollider.org/event/7454/contributions/38741/attachments/31353/47155/pricet_decal_llr_2017.pdf`

[130] C. Adloff *et al.* [CALICE Collaboration], “Hadronic energy resolution of a highly granular scintillator-steel hadron calorimeter using software compensation techniques,” JINST **7**, P09017 (2012) doi:10.1088/1748-0221/7/09/P09017 [arXiv:1207.4210 [physics.ins-det]].

[131] C. Adloff *et al.* [CALICE Collaboration], “Electromagnetic response of a highly granular hadronic calorimeter,” JINST **6**, P04003 (2011) doi:10.1088/1748-0221/6/04/P04003 [arXiv:1012.4343 [physics.ins-det]].

[132] V. Andreev *et al.*, “A high granularity scintillator hadronic-calorimeter with SiPM readout for a linear collider detector,” Nucl. Instrum. Meth. A **540**, 368 (2005). doi:10.1016/j.nima.2004.12.002

[133] G. Eigen *et al.* [CALICE Collaboration], “Hadron shower decomposition in the highly granular CALICE analogue hadron calorimeter,” JINST **11**, no. 06, P06013 (2016) doi:10.1088/1748-0221/11/06/P06013 [arXiv:1602.08578 [physics.ins-det]].

[134] B. Bilki *et al.* [CALICE Collaboration], “Pion and proton showers in the CALICE scintillator-steel analogue hadron calorimeter,” JINST **10**, no. 04, P04014 (2015) doi:10.1088/1748-0221/10/04/P04014 [arXiv:1412.2653 [physics.ins-det]].

[135] C. Adloff *et al.* [CALICE Collaboration], “Track segments in hadronic showers in a highly granular scintillator-steel hadron calorimeter,” JINST **8**, P09001 (2013) doi:10.1088/1748-0221/8/09/P09001 [arXiv:1305.7027 [physics.ins-det]].

[136] K. Kruger [CALICE Collaboration] “Prototype tests for a highly granular scintillator-based hadron calorimeter,” J. Phys. Conf. Ser. **587**, no. 1, 012033 (2015). doi:10.1088/1742-6596/587/1/012033

[137] F. Sefkow, “The new AHCAL prototype,” presentation at CALICE collaboration meeting (March 2017), `https://agenda.linearcollider.org/event/7454/contributions/38713/attachments/31363/47175/AHCAL-LLR.pdf`

[138] V. Buridon *et al.* [CALICE Collaboration], "First results of the CALICE SDHCAL technological prototype," JINST **11**, no. 04, P04001 (2016) doi:10.1088/1748-0221/11/04/P04001 [arXiv:1602.02276 [physics.ins-det]].

[139] B. Bilki *et al.*, "Hadron Showers in a Digital Hadron Calorimeter," JINST **4**, P10008 (2009) doi:10.1088/1748-0221/4/10/P10008 [arXiv:0908.4236 [physics.ins-det]].

[140] B. Bilki *et al.*, "Measurement of Positron Showers with a Digital Hadron Calorimeter," JINST **4**, P04006 (2009) doi:10.1088/1748-0221/4/04/P04006 [arXiv:0902.1699 [physics.ins-det]].

[141] B. Freund *et al.* [CALICE Collaboration], "DHCAL with Minimal Absorber: Measurements with Positrons," JINST **11**, no. 05, P05008 (2016) doi:10.1088/1748-0221/11/05/P05008 [arXiv:1603.01652 [physics.ins-det]].

[142] B. Bilki *et al.*, "Calibration of a digital hadron calorimeter with muons," JINST **3**, P05001 (2008) doi:10.1088/1748-0221/3/05/P05001 [arXiv:0802.3398 [physics.ins-det]].

[143] C. Adams *et al.*, "Design, construction and commissioning of the Digital Hadron Calorimeter—DHCAL," JINST **11**, no. 07, P07007 (2016) doi:10.1088/1748-0221/11/07/P07007 [arXiv:1603.01653 [physics.ins-det]].

[144] L. Arazi *et al.*, "Beam Studies of the Segmented Resistive WELL: a Potential Thin Sampling Element for Digital Hadron Calorimetry," Nucl. Instrum. Meth. A **732**, 199 (2013) doi:10.1016/j.nima.2013.08.006 [arXiv:1305.1585 [physics.ins-det]].

[145] C. Adloff *et al.*, "Construction and test of a 1×1 m2 Micromegas chamber for sampling hadron calorimetry at future lepton colliders", Nucl. Instrum. Meth. A **729**, 90 (2013).

[146] R Core Team (2014). "R: A language and environment for statistical computing." R Foundation for Statistical Computing, Vienna, Austria. `http://www.R-project.org/`.

[147] M. Dowle, T. Short, S. Lianoglou, A. Srinivasan with contributions from R. Saporta and E. Antonyan (2014). "`data.table`: Extension of `data.frame`. R package version 1.9.4." `http://CRAN.R-project.org/package=data.table`

[148] H. Wickham, "`ggplot2`: elegant graphics for data analysis," Springer, New York, 2009.

[149] H. Wickham, "The Split-Apply-Combine Strategy for Data Analysis," Journal of Statistical Software, 40(1), 1-29, 2011. `http://www.jstatsoft.org/v40/i01/`.

[150] H. Wickham, "Reshaping Data with the reshape Package," Journal of Statistical Software, 21(12), 1-20, 2007. `http://www.jstatsoft.org/v21/i12/`.

[151] Mokka detector geometry database, `http://flc.desy.de/ldcoptimization/tools/mokkamodels.php`

[152] C. Adloff *et al.*, "Study of the interactions of pions in the CALICE silicon-tungsten calorimeter prototype," JINST **5** (2010) P05007 doi:10.1088/1748-0221/5/05/P05007 [arXiv:1004.4996 [physics.ins-det]].

[153] F. Salvatore, "Simulation of the CALICE Test Beams with MOKKA," presentation at CALICE collaboration meeting (September 2007), `https://agenda.linearcollider.org/event/1785/contributions/4432/attachments/3355/5633/MokkaSalvatore.pdf`

[154] K. Shpak, "CERN'07 CALICE SiW ECAL+AHCAL+TCMT TB setup & Mokka models," CALICE internal note CIN-026, `https://twiki.cern.ch/twiki/pub/CALICE/CaliceInternalNotes/CIN-026_r1.pdf`

[155] S. Chen, D. Jeans, "2-photon separation studies using CALICE SiW test beam data," presentation at CALICE collaboration meeting (Argonne, 2014), `https://agenda.linearcollider.org/event/6341/contributions/30062/attachments/24888/38387/2-photons_SiW.pdf`

[156] K. Seidel, F. Simon, "Initial Study of Hadronic Energy Resolution in the Analog HCAL and the Complete CALICE Setup," CALICE analysis note CAN-015, `https://twiki.cern.ch/twiki/pub/CALICE/CaliceAnalysisNotes/CAN-015.pdf`

[157] F. Santanastasio, "Search for Supersymmetry with Gauge-Mediated Breaking using high energy photons at CMS experiment," PhD thesis, `http://www.roma1.infn.it/cms/tesiPHD/santanastasio.pdf`

[158] M. Anduze *et al.*, "A large scale prototype for a SiW electromagnetic calorimeter for a future linear collider - EUDET Module," EUDET-Memo-2010-17, `https://www.eudet.org/e26/e28/e86887/e104848/EUDET-Memo-2010-17.pdf`

[159] J. Rouëné [CALICE Collaboration], "Construction and testing of a large scale prototype of a silicon tungsten electromagnetic calorimeter for a future lepton collider," Nucl. Instrum. Meth. A **732**, 470 (2013). doi:10.1016/j.nima.2013.05.129

[160] M. S. Amjad *et al.*, "Beam test performance of the SKIROC2 ASIC," Nucl. Instrum. Meth. A **778**, 78 (2015). doi:10.1016/j.nima.2014.12.011

[161] J. Rouëné, "A Highly Granular Silicon-Tungsten Electromagnetic Calorimeter and Top Quark Production at the International Linear Collider," PhD thesis, `https://tel.archives-ouvertes.fr/tel-01062136/document`

[162] V. Balagura, "SiW ECAL status and plans," presentation at CALICE collaboration meeting (Argonne, 2014), `https://agenda.linearcollider.org/event/6341/contributions/30063/attachments/24889/38388/summary_1_balagura.pdf`

[163] V. Balagura, "ANR report, task 1: cosmic tests,"

[164] L. Lavergne *et al.*, "Silicon sensor gluing for the SiW ECAL - R&D status at LPNHE," presentation at HGC4ILD workshop (February, 2015), `https://indico.in2p3.fr/event/11192/session/0/contribution/34/material/slides/0.pdf`

[165] M. Anduze, D. Bailey, R. Cornat, P. Cornebise, A. Falou*et al.*, "JRA3 Electromagnetic Calorimeter Technical Design Report," 2009, pp.1-39. <in2p3-00612932> `http://hal.in2p3.fr/in2p3-00612932/document`

[166] R. Pöschl, "News from FEV8_COB," presentation at CALICE collaboration meeting (March, 2017), `https://agenda.linearcollider.org/event/7454/contributions/38735/attachments/31346/47146/talk220317-cob.pdf`

[167] T. Frisson *et al.*, "SiW ECAL Technological Prototype Long slab and FEV_COB," presentation at CALICE collaboration meeting (September, 2013), `https://agenda.linearcollider.org/event/6050/contributions/28162/attachments/23355/36294/talk110913.pdf`

[168] J. Nanni, "Status of short & long SLABs," presentation at CALICE collaboration meeting (March, 2017), `https://agenda.linearcollider.org/event/7454/contributions/38734/attachments/31344/47142/CALICE_Collaboration_meeting_17March22-24.pdf`

[169] Omega lab, SKIROC Front-end Chip webpage, `http://omega.in2p3.fr/index.php/products/skiroc.html`

[170] T. Suehara *et al.*, "SKIROC2/2A study in Kyushu," presentation at CALICE collaboration meeting (March, 2017), `https://agenda.linearcollider.org/event/7454/contributions/38701/attachments/31336/47130/170322-calice-suehara-sk2.pdf`

[171] A. Lobanov, "Bench tests of single SKIROC2/2A at LLR," presentation at CALICE collaboration meeting (March, 2017), `https://agenda.linearcollider.org/event/7454/contributions/38733/attachments/31339/47134/SKbench_LLR_forCALICE.pdf`

[172] N. Seguin-Moreau *et al.*, "SKIROC2 Silicon Kalorimeter Integrated Read-Out Chip," presentation at TWEPP2011,

[173] H. Sumida, Y. Sudo, T. Suehara, T. Yoshioka, K. Kawagoe, "Silicon Sensor Study for ILD ECAL," presentation at CALICE collaboration meeting (March, 2016), `https://agenda.linearcollider.org/event/6892/contributions/33891/attachments/27918/42268/Silicon_sensor_study_for_ILD_ECAL_for_CALICE_meeting.pdf`

[174] T. Suehara, "Towards a Technological Prototype for a High-granularity Electromagnetic Calorimeter for Future Lepton Colliders," presentation at CALOR2016 (May, 2016), `https://indico.cern.ch/event/472938/contributions/1150740/attachments/1276160/1893406/suehara-calor.pdf`

[175] K. Shpak, "Square events," presentation at SiW ECAL BT2015 Analysis meeting (March 2016), `https://indico.in2p3.fr/event/12986/contributions/12453/attachments/10591/13119/Squarepresentation.pdf`

[176] SiW ECAL test beam elog, entry #1186 about square events in SiW ECAL sensors during DESY test beam, `https://llrelog.in2p3.fr/calice/1186`

[177] CERNBOX, https://cernbox.cern.ch,
to copy the November'15 CERN SPS SiW ECAL test beam data:
cd «your data space» (e.g. with sufficient storage, don't use your afs directory... typically a temporary space).
. /afs/cern.ch/project/eos/installation/user/etc/setup.sh
eos ls -l "eos/user/b/boudry/tb2015-pass1/*0086*" (for run0086)
xrdcp –verbose "root://eosuser.cern.ch//eos/user/b/boudry/tb2015-pass1/run_0086_by_dif1.raw.root" ./

[178] SiW ECAL test beam e-log, please contact Vincent Boudry (vincent.boudry@cern.ch) to get an access, https://llrelog.in2p3.fr/calice/

[179] SiW ECAL test beam 2015 run-log, https://indico.in2p3.fr/event/12532/contributions/10177/attachments/8440/10469/Presentation-2.pdf

[180] CERN twiki page for the 2015 SiW ECAL TB analysis, https://twiki.cern.ch/twiki/bin/view/CALICE/SiWEcalBT201511Analysis#Analysis

[181] V. Balagura, CNRS report "November'15 test beam analysis"

[182] J. b. Cizel, R. Cornat, "About (digital ?) noise and Scurves," presentation at CALICE collaboration meeting (September, 2016), https://agenda.linearcollider.org/event/7304/contributions/37334/attachments/30491/45602/calice_dallas_JBCizel_RC.pdf

[183] SiW ECAL test beam elog, entry #618 on reduced occupancy in chip 15, https://llrelog.in2p3.fr/calice/618

[184] K.Shpak, "Detected DIF output bugs," presentation at BT2015 Analysis meeting (April, 2016), https://indico.in2p3.fr/event/13077/contributions/12720/attachments/10769/13338/PresentationBUGs.pdf

[185] J.-Y. Hostachy, D. Grondin , J. Giraud, "CALICE Si/W ECAL Endcap structures and cooling system," presentation at CALICE collaboration meeting (March, 2017), https://agenda.linearcollider.org/event/7454/contributions/38739/attachments/31345/47143/LPSC_Meeting_LLR_22_03_2017-v2.pdf

[186] C. Patrignani *et al.* [Particle Data Group], Chin. Phys. C **40**, 100001 (2016).

[187] S. Schael *et al.* [ALEPH, DELPHI, L3, OPAL and LEP Electroweak Collaborations], "Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP", Phys.Rept. **532**, 119 (2013), [arXiv:1302.3415 [hep-ex]].

[188] V. M. Abazov *et al.* [D0 Collaboration], "Measurement of the W Boson Mass with the D0 Detector", Phys.Rev.Lett. **108** 151804 (2012).

[189] T. Aaltonen *et al.* [CDF Collaboration], "Precise Measurement of the W-Boson Mass with the CDF II Detector", Phys.Rev.Lett. **108** 151803 (2012).

[190] M. Aaboud *et al.* [ATLAS Collaboration], "Measurement of the W-boson mass in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector," arXiv:1701.07240 [hep-ex].

[191] M. Baak *et al.* [The Gfitter Group], "The global electroweak fit at NNLO and prospects for the LHC and ILC", Eur. Phys. J. **C 74** 3046 (2014).

[192] G. Moortgat-Pick (ed.) *et al.* [The Gfitter Group], "Physics at the e^+e^- Linear Collider", Eur.Phys.J. **C 75** 371 (2015) 371 [arXiv:1504.01726 [hep-ph]].

[193] K. Tsushimoto, "W boson in the ILC experiment. Study of the mass measurement using hadron decays", Master Thesis, Shinshu Univ. [in Japanese] (2015).

[194] X. Mo, G. Li, M. Q. Ruan and X. C. Lou, "Physics cross sections and event generation of e^+e^- annihilations at the CEPC," Chin. Phys. C **40**, no. 3, 033001 (2016) doi:10.1088/1674-1137/40/3/033001 [arXiv:1505.01008 [hep-ex]].

[195] The Wizard Event Generator, https://whizard.hepforge.org

[196] T. Sjostrand, S. Mrenna and P. Z. Skands, "PYTHIA 6.4 Physics and Manual," JHEP **0605**, 026 (2006) doi:10.1088/1126-6708/2006/05/026 [hep-ph/0603175].

[197] G. W. Wilson, “Updated Study of a Precision Measurement of the W Mass from a Threshold Scan Using Polarized e^- and e^+ at ILC,” arXiv:1603.06016 [hep-ex].

[198] Private communications with Akiya Miyamoto and Hiroaki Ono from KEK

[199] B. Auguie, D. Eddelbuettel and S. G. Johnson, “`RcppFaddeeva`: `Rcpp` Bindings for the `Faddeeva` Package,” R package version 0.1.0, 2015, `https://CRAN.R-project.org/package=RcppFaddeeva`

[200] W. N. Venables, & B. D. Ripley, “Modern Applied Statistics with S,” Fourth Edition, Springer, New York, 2002. ISBN 0-387-95457-0

[201] Private communications with Graham W. Wilson from Kansas University

[202] W. Verkerke and D. P. Kirkby, “The RooFit toolkit for data modeling,” eConf C **0303241**, MOLT007 (2003) [physics/0306116].