Reconstruction Algorithms for the FADC DAQ System of the KASCADE-Grande Experiment

Diplomarbeit
zur Erlangung des akademischen Grades
Diplom-Physiker

dem Fachbereich Physik der
Universität Siegen

vorgelegt von
George Victor Andrei

Juli 2004

Contents

Chapter 1

Introduction and Motivation

KASCADE-Grande is an extensive air shower experiment to study the cosmic ray energy spectrum and chemical composition in the energy range 100 TeV – 1 EeV. It is a ground-based multi-detector set-up that measures simultaneously the electromagnetic, muon and hadronic components of the extensive air showers. The Grande array of the KASCADE-Grande experiment is sensitive to the charged component of the air shower. The current Grande data aquisition system provides only information about the energy deposit and the arrival time of the first particle in a detector station. No information about the time development of the signal is provided. The recording of a triggered shower event generates a dead time in which no further event can be recorded.

A new FADC DAQ system, currently under development, is going to be installed, in addition, in the Grande array of the KASCADE-Grande experiment. It is a free-running and self-triggering system, which will improve significantly the data quality of the experiment. It provides digitized photomultiplier pulse shapes containing information about the energy deposits and the time development. In order to profit from this improvement, the development of fast and efficient reconstruction algorithms for the FADC data is mandatory. A rigorous study of the impact, that the FADC characteristics, e.g. resolution and sampling frequency, have on the accuracy of the measurements is the aim of this thesis.

A brief introduction of the physics of cosmic rays and extensive air showers is given in chapter 2. The KASCADE-Grande experiment, together with its aims and components, is described in chapter 3. The chapter 4 presents the concepts and the features of the new FADC DAQ system.

Comprehensive investigations of different reconstruction methods require a large collection of pulse shapes. A Monte Carlo tool, developed to simulate pulse shapes as recorded by a Grande station, is described in chapter 5.

The energy deposit can be reconstructed by integrating the pulse shapes. The precision and accuracy of different integration methods, for equally spaced samples, are investigated. Time-base errors and amplitude fluctuations are added to the ideal pulses and their contribution to the reconstruction of the energy deposit is investigated in chapter 6.

Chapter 2

Cosmic Rays

2.1 Introduction

Cosmic rays are high energy corpuscular and electromagnetic radiations, coming from our galaxy or even from beyond, that permanently hit the Earth's atmosphere. They mainly consist of protons ($\approx 85\%$), alpha particles ($\approx 12\%$) and nuclei of different elements with $Z \geq 3$ ($\approx 3\%$). Electrons constitute only about 1% of the number of primary protons. Neutrinos and photons have been also observed. Antiparticles are very rarely found among the primary cosmic ray particles. The fraction of antiprotons with energies higher than 10 GeV in respect to the flux of protons amounts to 10^{-4}, while the flux of positrons represents approximately 10% of the number of electrons at energies around 10 GeV [Gru00, Lon92].

In the early 1900's, T.R. Wilson noticed that an electroscope, at that time a device widely used to measure levels of radioactivity, discharges even in the absence of a radioactive source. This leakage of charge was believed to be due to the presence of a background radiation, mainly coming from the ground. Therefore, in their attempt to reduce this effect, physicists performed experiments at high altitudes, mainly on the top of the highest buildings or constructions. An example is the experiment of Thomas Wulf on the Eiffel Tower in 1910, which provided a reduction of the leakage of about 42%. It was Victor Hess, an Austrian physicist, who discovered in 1912 that in fact, the radiation was penetrating downward, coming from beyond the atmosphere. He used a balloon to ascend to higher altitudes, while repeatedly measuring the leakage. He found that the leakage was first reducing, but after 1500 m altitude it was more intense than at the ground level. This was the first evidence of cosmic rays, and for his discovery, Hess was awarded the Nobel Prize in 1936 [Lon92].

Later, in 1926, Robert Millikan conducted several experiments with balloons which confirmed Hess' discovery and, moreover, he showed that the incoming radiation collides with the atoms from the atmosphere and produces secondary particles. Millikan named the extraterrestrial radiation as *cosmic rays*. That was the beginning of a new era in the field of particle physics. As a result of cosmic rays researches, Carl Anderson discovered the evidence of the positron in a cloud chamber in 1932, and, together with Seth Neddermeyer, discovered the muon in 1936. For the positron discovery he shared the Nobel prize with Victor Hess in 1936. Because the experiments with balloons were expensive and dangerous, most of the experiments moved to high altitudes of the mountains. In 1947, Cecil Powell and his group discovered the $\pi - meson$, exposing photographic plates to cosmic rays at 3000 m altitude, and, in the same year, Rochester and Butler discovered the *kaon* in tracks produced by cosmic rays in a cloud chamber. By the middle 1950's, particle accelerators replaced cosmic rays in high energy particle experiments, but even today cosmic rays provide the most energetic particles ever observed.

In 1938, following the experiments of Bruno Rossi, Pierre Auger discovered the *extensive air showers*, observing the time coincidence between detectors, placed at large distances. Based on his measurements, he concluded that the energy spectrum of the cosmic rays reaches up to 10^{15} eV. Today it is known that the energy spectrum of the primary cosmic rays extends over many decades of energy. The highest energetic cosmic ray event observed so far has 320 EeV (i.e. Fly's Eye experiment) [Rao98]. The energy spectra can be represented as the *differential flux*, as number of particles in the energy range $(E; E+dE)$, or as the *integral flux*, i.e. as number of particles with energies exceeding or equal to E. From various experiments, it was established that the former can be represented by a power law distribution:

$$\frac{dN}{dE} = E^{-\gamma} \quad , \tag{2.1}$$

where γ represents the *spectral index* that determines the slope of the distribution [Lon92]. Figure 2.2 shows the differential energy spectra of the primary cosmic rays, as they were measured by different experiments. In the PeV region the energy spectrum flattens, the spectral index γ changing from 2.7 to 3.1. This is usually referred to as the *knee* of the energy spectrum. In the region of 10 EeV the energy spectrum becomes steeper again, determining the so called *ankle*. The spectral index changes back to its initial value.

Both the *knee* and the *ankle* were confirmed by many experiments at different atmospheric depths, but the cause of the sudden change in the exponent of the power law energy spectrum has not been clearly understood yet. There are several theoretical models that try to explain the *knee*, based either on astrophysical reasons, like the effects of the cosmic ray transport through the interstellar medium, or on the assumption of a change of the hadronic interaction mechanisms.

The experimental results pointed to a change in the chemical composition of the primary spectrum. It has been observed that, above the *knee*, the flux of light particles decreases such, that at high energies the spectrum is dominated by heavy particles. It is expected that the *knee energy* is rigidity dependent [Rot03]:

$$E^{A}_{knee} = Z \cdot E^{proton}_{knee} \quad . \tag{2.2}$$

Since the flux of protons is the largest among the cosmic ray particles, the knee they cause is distinct, at about 5 PeV. Therefore, the *iron knee* is expected to be found around 0.1 EeV.

Cosmic rays can be detected in several ways, depending on the energy of the primaries. Since low energy cosmic rays are absorbed in the Earth's atmosphere, airborne cosmic ray detectors are carried in balloons at high altitudes, where they can remain aloft for days. Experiments like CAPRICE (1998) measured fluxes of particles between $3 \div 350$ GeV [Moc03]. The space based detection technique involves cosmic ray detectors mounted on research space stations to record events, while the station orbits the Earth. The ground-based detection techniques are the most widely used. Cosmic rays with energies beyond 100 TeV are studied via the extensive air shower they produce in the atmosphere. Large detector arrays are used to sample the secondary particles generated by the collision of the primary cosmic ray with the atmospheric nuclei. KASCADE-Grande is such an experiment and more details about its aims and results are provided in chapter 3. Using large telescopes, it is possible to detect the Cherenkov radiation emitted by the electromagnetic showers, like e.g. the HESS experiment does [Hin04].

2.2 Extensive Air Showers

When the primary cosmic ray undergoes nuclear collisions, high in the atmosphere, a large number of secondary particles is produced. These particles can decay or further interact with

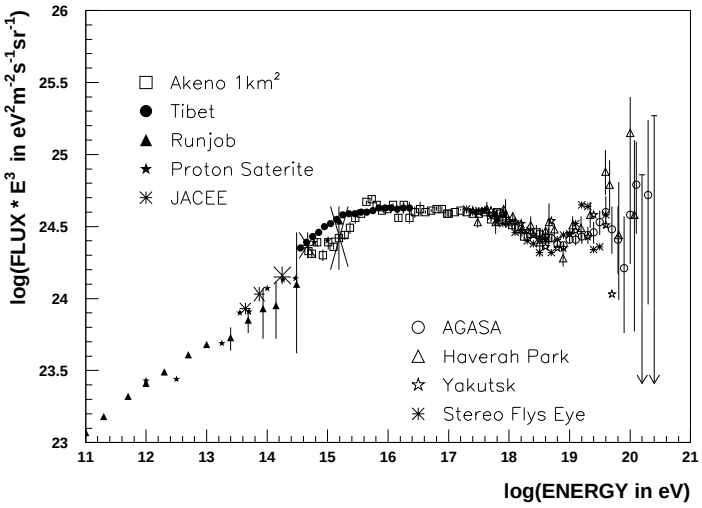


Figure 2.1: The differential energy spectra of primary cosmic ray particles for energies exceeding 0.1 TeV, as measured by several experiments. The *knee* occurs in the PeV region, while the *ankle* is observed in the 10 EeV region. Both scales are logarithmic and the flux is multiplied with E^3 [Nag00].

other atmospheric nuclei, such that a cascade of particles develops in the atmosphere. The number of particles produced in an extensive air shower (EAS) is directly related to the energy of the primary cosmic ray particle and keeps increasing with the propagation of the shower front until the shower development reaches a maximum.

The shower development depends on several competitive processes, like particle production, decay, ionization loss and absorption. The mean energy of the particles in the shower decreases as the shower develops in the atmosphere, i.e the mean energy decreases with the increase of the depth. When the energy falls below the threshold for production of additional particles, the energy will gradually be lost by ionisation and other processes. If the primary cosmic ray is a very energetic photon, then the shower develops mainly into an electromagnetic component, of which dominat effects are the pair production and bremsstrahlung. The average total number of electrons is given as a function of the atmospheric depth (t), measured in radiation lengths:

$$N_e(E_0,t) = \frac{0.31}{\sqrt{\beta_0}}\, exp\left[t\left(1-\frac{3}{2}\ln s\right)\right] \qquad , \tag{2.3}$$

where E_0 is the energy of the primary photon, $s \simeq \frac{3t}{t+2\beta_0}$ and $\beta_0 = ln(E_0/\varepsilon_0)$. The quantity ε_0 defines a critical energy (84.2 MeV) for the electrons in air, at which the ionization process becomes the predominant interaction for the electrons [Gre56].

The lateral distribution of the extensive air shower increases as a function of depth, and depends upon the energy and the type of the primary particle. The following formula, known as Nishimura-Kamata-Greisen (NKG) function, describes the particle density $\Delta(N_e,r)$ per m^2, at a certain distance r (in a plane perpendicular to the shower axis):

$$\Delta(N_e,r) = \frac{N_e}{2\pi r_0^2}\frac{\Gamma(4.5-s)}{\Gamma(s)\Gamma(4.5-2s)}\left(\frac{r}{r_0}\right)^{(s-2)}\left(1+\frac{r}{r_0}\right)^{s-4.5} \qquad , \tag{2.4}$$

where Γ represents the Gamma function, N_e is the shower size, s is the age parameter, and r_0 is the *Molière radius*. The age s is a parameter that characterizes the status of the shower development ($s = 0$ at the first interaction, $s = 1$ at the shower maximum), while the Molière radius determines the lateral spread of particles in the shower, as a result of the multiple collisions they undergo in the atmosphere. The number of electrons and muons produced in the shower can be obtained by using the NKG function to fit the experimental data.

Figure 2.2 shows a schematic representation of an extensive air shower developing in the atmosphere. The particles of the shower travel almost with the speed of light down to Earth, within a *pancake*-like shaped volume with a thickness of a few meters. The shower front mainly consists of muons and electrons, while the hadrons arrive delayed due to their larger mass. The zenith angle (θ) is measured from the time delay between the different detectors [Lon92].

2.2.1 The Hadronic Component

The hadronic component of the extensive air shower consists of pions, kaons, protons, neutrons and nuclear fragments. The most abundant are the charged and neutral pions ($\pi^{\pm}$,π^0). They are high energetic and unstable, with short mean life times, of $2.5 \cdot 10^{-8}$s ($\pi^{\pm}$) and $1.8 \cdot 10^{-16}$s (π^0) [Lon92]. The neutral pions decay almost immediately after they are produced. In this way, a large part of the *hadronic* energy is transferred to the electromagnetic component (see equation 2.9). The longer life time of the charged pions gives them the opportunity to undergo further nuclear interactions and to increase the size of the hadronic component. But, most of them are likely to decay into muons and the corresponding neutrinos (see equation 2.5), therefore reducing the number of hadrons present in the shower to 1% from the total number of particles created.

2.2.2 The Muon Component

The muon component of the air shower is the most penetrating. It results mostly from the decay of the high energy charged π-mesons and kaons, and with a very small probability from the photonuclear processes:

$$\pi^{\pm} \rightarrow \mu^{\pm} + \nu_{\mu}(\overline{\nu}_{\mu}) \,, \tag{2.5}$$

$$K^{\pm} \rightarrow \mu^{\pm} + \nu_{\mu}(\overline{\nu}_{\mu}) \,. \tag{2.6}$$

Since muons do not undergo strong interactions, but only lose part of their energy by ionization, the muon component will not suffer significant attenuation during the shower expansion. The low energy muons may decay into electrons, positrons and muon neutrinos:

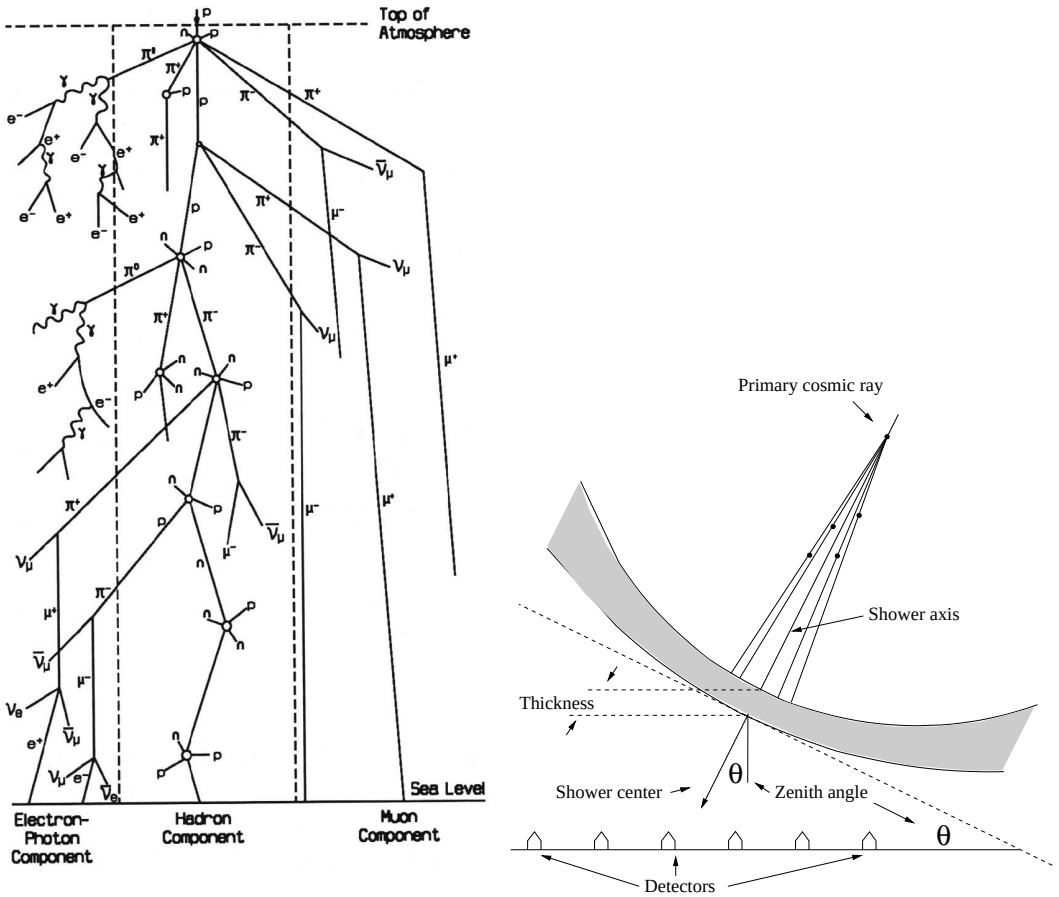
$$\mu^{+} \rightarrow e^{+} + \nu_e + \overline{\nu}_{\mu} \,, \tag{2.7}$$

$$\mu^{-} \rightarrow e^{-} + \overline{\nu}_e + \nu_{\mu} \,. \tag{2.8}$$

Due to the very small cross section, the flux of neutrinos is not at all attenuated in the atmosphere.

2.2.3 The Electromagnetic Component

The electromagnetic component of an extensive air shower is given by the electrons, positrons and γ-ray photons. The pair production is the dominant effect, due to the energies involved. If



(a) (b)

Figure 2.2: Schematic representation of an extensive air shower: the separation of the shower components (*a*) [Gri01], and the shower development in the atmosphere (*b*) (modified from [Bra01]).

the particle initiating the shower has hadronic nature, then the electromagnetic component is mainly generated by the decay of the neutral π-meson into two photons:

$$\pi^0 \rightarrow \gamma + \gamma \ . \tag{2.9}$$

The two high energy photons cause pair productions in the fields of the atmospheric nuclei or electrons. The created electrons and positrons either lose their energy by ionizing interactions, or undergo *bremsstrahlung* and emit photons. The photon can further produce another $e^- e^+$ pair, if its energy is sufficient (at least twice the rest mass of the electron). Therefore, the result is a cascade of electrons, positrons and photons which increases in size as it propagates down in the atmosphere. The photoelectric and Compton effects may appear only at energies lower than a few MeV, therefore their contribution is minimal.

The lateral distribution of the electromagnetic component is mainly due to multiple Coulomb scattering encountered by electrons and photons with the air nuclei. Also, the pair production and bremsstrahlung contribute to the lateral spread of the secondary particles with respect to the shower axis, but this contribution is rather small.

Chapter 3

The KASCADE-Grande Experiment

KASCADE-Grande is an extensive air shower experiment, located on site of the Forschungszentrum Karlsruhe at latitude 49° north, longitude 8° east and at 110 m above sea level. It is a ground-based multi-detector setup, comprising the KASCADE (**KA**rlsruhe **S**hower **C**ore and **A**rray **DE**tector), Grande and Piccolo arrays, the central detector and the muon tunnel. The aim of the experiment is to measure the cosmic ray primary energy spectrum and chemical composition in the energy range $100\,\mathrm{TeV} - 1\,\mathrm{EeV}$, especially in the *knee* region of the spectrum. The KASCADE experiment, in operation and taking data since 1996, is sensitive in the energy range $100\,\mathrm{TeV} - 10\,\mathrm{PeV}$. According to the data observed by KASCADE, the position of the knee shifts to higher energies with the mass of the primary. A separation of the primary particle spectrum into its chemical composition gives rise to the expectation that the *iron knee* may be identified at $E_{knee} \approx 0.1\,\mathrm{EeV}$.

Since the flux of primary particles decreases at higher energies, the active detection area has to be increased accordingly. The extension of the KASCADE experiment to the present KASCADE-Grande experiment, provides the opportunity to observe data in the energy range $10\,\mathrm{PeV} - 1\,\mathrm{EeV}$. This allows to discover the possible existence of a break in the iron spectrum (see e.g. [Hau03]).

3.1 The KASCADE Setup

The KASCADE setup consists of three main parts: the *detector array*, the *central detector* and the *muon tracking detector*. They allow simultaneous measurements of the electromagnetic, muon and hadronic components of the extensive air showers.

The detector array consists of 252 detector stations distributed over an area of $200 \times 200\,\mathrm{m}^2$. It is organized in clusters of 16 stations, the four *inner* clusters comprising only 15 stations (see figure 3.2). Each cluster acts like an autonomous unit which can trigger and digitize data independently. Except the stations from the inner clusters, each station is equipped with two liquid scintillator detectors to ensure the measurement of the electromagnetic component of the extensive air shower. Light guides collect and trace the scintillation light to photomultipliers. In order to prevent oxidation, the detector is filled with argon. The detectors from the *inner* clusters have two additional liquid scintillators installed and no muon detector.

Additionally, below the liquid scintillators, a lead/iron absorber and a plastic scintillator are employed to detect the muon component (see figure 3.1). The energy threshold for vertical incidence of the electromagnetic component is 5 MeV, while the absorber corresponds to $20\,\mathrm{X}_0$ (radiation length) and demands an energy threshold of 230 MeV. The mean energy deposit of

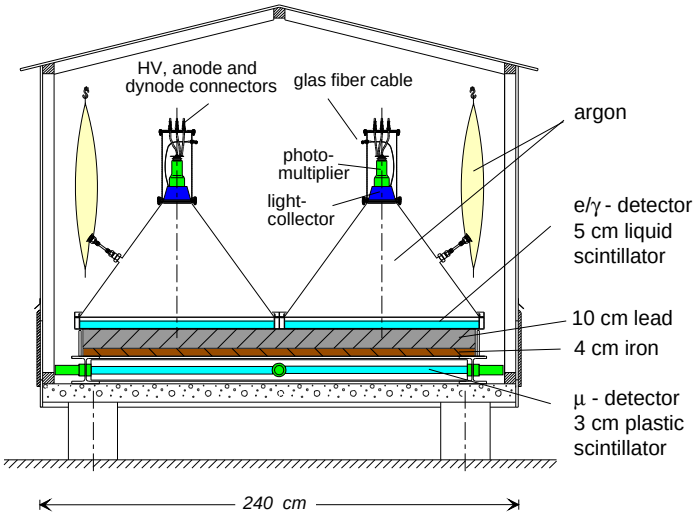


Figure 3.1: A profile view of a KASCADE detector station. The *outer* clusters have two e/γ detectors installed on the top of a lead/iron absorber and a muon detector. The *inner* stations have four e/γ detectors and no muon detector [Eng03].

the minimum ionizing particle (m.i.p.) was found to be 12 MeV, and energy resolution measurements achieved about 8% for the m.i.p. deposit [Eng03].

The central detector, positioned in the middle of the KASCADE array, comprises several types of detectors. The *hadron calorimeter* measures and provides data about the hadronic component of the extensive air showers.

Detector	Particle	Total area (m^2)	Threshold
KASCADE:			
Array, liquid scintillators	e/γ	490	5 MeV
Array, plastic scintillators	μ	622	230 MeV
Muon tracking detector, streamer tubes	μ	128×4 layers	800 MeV
central detector:			
Calorimeter, liquid ionization chambers	h	304×8 layers	50 GeV
Trigger layer, plastic scintillators	μ	208	490 MeV
Top cluster, plastic scintillators	e/γ	23	5 MeV
Top layer, liquid ionization chambers	e/γ	304	5 MeV
Multi-wire proportional chambers	μ	129×2 layers	2.4 GeV
Limited streamer tubes	μ	250	2.4 GeV
Grande	e/μ	370	3 MeV
Piccolo	e/μ	80	5 MeV

Table 3.1: Detector components of KASCADE-Grande, their total sensitive areas and thresholds for vertical particles [Eng03, Kam03] .

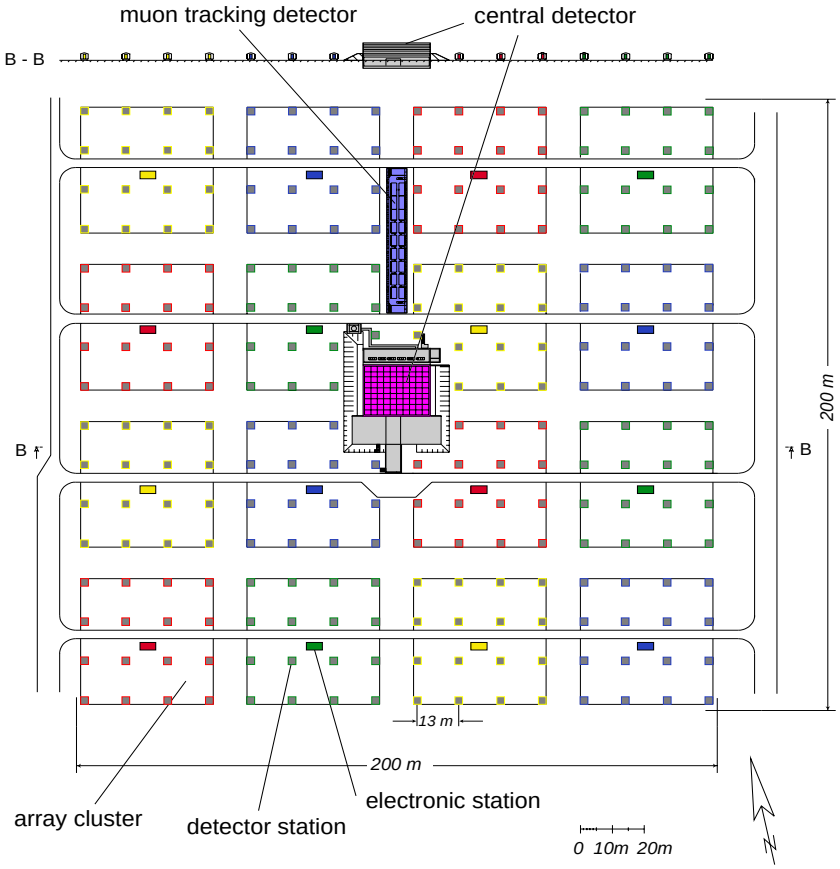


Figure 3.2: Schematic layout of the KASCADE array. The central detector is placed in the middle of a 252 detector station array. The array is arranged in 12 *outer* and 4 *inner* independent clusters. An electronic station, placed in the middle of each cluster, supplies the high voltage for the electronics of the detectors, collects and digitizes the data from each detector station of the cluster [Eng03].

It mainly consists of one lead absorber layer (filters the electromagnetic component), and 8 layers of iron interspaced with 9 layers of warm-liquid ionization chambers. A *trigger plane* consisting of plastic scintillators, is placed below the third absorption layer in order to measure the arrival time distributions of the muons. Below the calorimeter, a system of multi wire proportional chambers (MWPC) and limited streamer tubes (LST) are employed to measure the positions and the angles of the incoming muons with energies higher than 2.4 GeV.

The muon tracking detector has an effective area of $128\,\mathrm{m}^2$ and is installed underground, north of the central detector. It consists of three layers of streamer tubes, which allow to reconstruct the tracks of charged particles that cross the detector. Therefore, the size and the lateral distribution of the muon component can be determined.

A summary of the detector components of the KASCADE array is presented in table 3.1.

3.2 The Grande Array

The Grande array consists of 37 detector stations from the former EAS-TOP cosmic ray experiment (running between 1987 and 2000 at Campo Imperatore, Gran Sasso Laboratories, Italy). It covers a larger detection area than the KASCADE array, of about $700 \times 700\,\mathrm{m}^2$. Each Grande

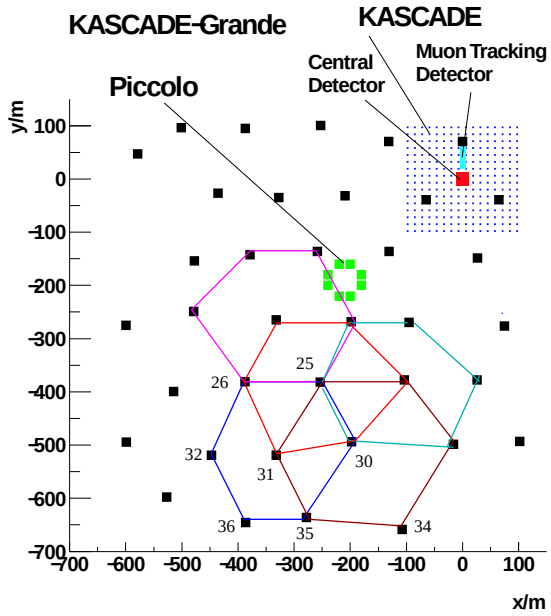


Figure 3.3: Schematic layout of the three components of the KASCADE-Grande experiment. The Piccolo array is placed near the center of the Grande array to create a trigger signal for the Grande and KASCADE components. The Grande array is arranged in 18 overlapping hexagonal clusters to ensure the 4 or 7-fold coincidences. The figure gives the example of five overlapping clusters (modified from [Hau03]).

station is equipped with 16 NE102A plastic scintillators ($80 \times 80\,\mathrm{cm}^2$ active detection area each, and 4 cm thickness), arranged in a 4×4 array. To each of these 16 scintillators, a pyramidal optical guide is attached to trace the light to a photomultiplier of the type XP3462. The photomultipliers operate at high gain mode, in a voltage range from 1500 V to 2500 V, and ensure the timing and low particle density measurements. The 16 photomultiplier pulses are added up by means of a mixer device. The resulting pulse is split in two different signals: one is discriminated in the detector station and used to trigger the station (*logic run signal*), and the other is sent through shielded cables to the Grande DAQ, where is digitized in peak sensing ADCs.

Additionally, each of the four central scintillators is equipped with a second photomultiplier of the same type, operating at low gain mode, in a voltage range from 1200 V to 1600 V. Their aim is to ensure the high particle density measurements. Measurements performed in the Grande stations resulted in a mean energy loss of about 10.3 MeV, for muons passing through the detector under a mean incidence angle of 36.9°. Multiplying this mean energy with the cosine of the considered angle, an 8.2 MeV mean energy loss per vertical minimum ionizing particle (m.i.p.) is obtained [Agl93]. The dynamic range achieved is 0.3 to 750 m.i.p./10 m^2 for the high gain channel, and from about 12 to 30,000 mips/10 m^2 for the low gain channel [Ber01].

For trigger purposes, the Grande array is arranged in 18 hexagonal clusters, with an average distance between stations of about 137 m. Each cluster comprises 7 detector stations, i.e. one station positioned in the center of the hexagon and the other six around it. Another feature of this arrangement is that each station is part of up to seven clusters (see figure 3.3).

When a Grande station detects a particle in the detector, a single particle event is triggered, and a logic run signal is sent to the Grande DAQ station. If more stations detect particles, in coincidence, then the Grande array triggers an air shower event. Apart from the situation when the array may be externally triggered, i.e by Piccolo component (see section 3.3), a cluster of the Grande array triggers only in the following two cases:

- a *four-out-of-seven* (4/7) coincidence: coincidence between the logic run signals of the central station of the cluster and three neighbouring peripheral stations, such that a parallelogram is formed. Taking the example given in figure 3.3, the cluster with the station number 31 as central station, a 4/7 coincidence is available, if it is triggered by the stations 31, 30, 25 and 26, but not by the stations 31, 30, 25 and 32.
- a *seven-out-of-seven* (7/7) coincidence: coincidence between all the logic run signals of the cluster.

Also, the 7/7 coincidence will be an external trigger for the other components of the experiment and the data will be sent to the KASCADE DAQ station. Events with only 4/7 coincidence will be locally stored on the Grande DAQ PC's. Apart from its internal trigger, the Grande data aquisition is also triggered by the Piccolo array and the other components of the KASCADE-Grande experiment.

3.3 The Piccolo Trigger Array

The Piccolo array consists of 8 detector stations (see figure 3.3) from the former KARMEN neutrino experiment. They form an octagon placed at about 20 m from each other. Each station is equipped with 12 scintillator plates with $10\,m^2$ detection area. The Piccolo array is positioned between the Grande and KASCADE detection centers, and its main aim is to provide a fast external trigger signal, in order to ensure common data taking for the two arrays. This increases the trigger efficiency and the density measurements for each shower.

Chapter 4

The FADC System

With the current Grande data acquisition system, a 500 μs dead time occurs after each shower event. When a coincidence condition is fulfilled (see chapter 3), a trigger is generated, and the read-out of the data from each Grande station starts. During this procedure, no other event can be recorded. On the other hand, the data recorded for each detector station contains no information about the energy development in time. First, the photomultiplier signal is integrated by a shaping amplifier during a period of 8 μs, and a pulse is generated with an amplitude proportional to the integrated charge of the signal. Then the signal is sent to the Grande DAQ station, where a peak ADC extracts the value of the amplitude, which is proportional to the energy deposit during the event. In addition to this information, a TDC hit is included, which provides information only about the time, when the signal has crossed the threshold.

A new FADC DAQ system, that improves the data quality of the KASCADE-Grande experiment, is currently under development. It is a free-running and self-triggering system, based on digitizer boards in each Grande station. The mixed photomultiplier pulses are digitized with zero dead time and sent via an optical fiber to the Grande DAQ station. A system of storage boards receives and writes the data to a ring buffer, from where a master PC selects of the single particle events, that takes place in coincidence. The new FADC DAQ system will be installed in parallel with the current system in the Grande array. It will provide digitized pulse shapes, that contain information about the energy deposits and their development in time, and give the opportunity to reconstruct other parameters of the air showers. The components of the FADC DAQ system are given in the following:

- 37 KGEMD (KASCADE-Grande Electromagnetic Detector Digitizer) boards, which are going to equip the Grande stations,
- 5 KGEMS(KASCADE-Grande Electromagnetic Detector Storage) boards, which will be installed in the Grande DAQ station,
- 5 first level PCs, which will be also installed in the Grande DAQ station,
- 5 KGEMP (KASCADE-Grande Electromagnetic Detector PCI Interface) cards, to be mounted in the first level PCs as the interfaces to the KGEMS boards,
- 1 master PC, also to be installed in the Grande DAQ station,
- 1 KGEMT card (KASCADE-Grande Electromagnetic Detector Trigger Receiver) card, to be mounted in the master PC.

4.1 The Digitizer Board

Each of the 37 Grande stations will be equipped with one digitizer board (KGEMD), that processes the high and low gain signals and transmits the generated data packets to the central DAQ station. As mentioned in chapter 3, each of the 16 scintillators of a Grande station is viewed by a photomultiplier, operating in high gain, for timing and low particle measurements. Each of the four central scintillators are viewed by an additional photomultiplier, operating in low gain, for timing and high particle measurements. The individual photomultiplier signals are summed to one high gain and one low gain signal [Chi03]. The resulting signals feed two separate inputs of the KGEMD board, by which they are digitized with an effective sampling frequency of 250 MHz.

The digitization is performed by eight FADCs[1] such, that each input signal is digitized by four cascaded FADCs. Each FADC has a resolution of 12 bits and a sampling rate of 62.5 MHz, and is driven by an onboard quartz oscillator[2]. The analog input signal is split and distributed to the four FADCs. The sampling frequency ensures that each FADC samples the input signal every 16 ns. To achieve an effective sampling rate of 250 MHz, the four FADCs are interleaved in time. The 62.5 MHz clock signal is split into four channels which are delayed such, that the input signal is digitized by one FADC every 4 ns.

The architecture of a 12-bit FADC (4096 channels) employs 4095 comparators, which simultaneously compare the analog input signal with an equally distributed threshold voltage at their inverting inputs. The output of the comparators is collected by a digital encoder that determines the digital output code, which corresponds to the ADC channel in which the sampled analog voltage is binned. This ensures a fast conversion of the analog input.

The FADCs digitize the input signals continuously, but the read-out is started only when the high gain signal exceeds a certain threshold value. This restricts the transmission of data, only to that being of interest. The digitization period lasts for 63 clock cycles, which corresponds to 1.008 μs. If the analog input signal exceeds the threshold value in the last 200 ns of the digitization period, the data acquisition is prolonged for another microsecond. The data from each FADC is buffered in a FIFO such, that the digitization and the data transmission are independent processes. A new event can be digitized, while the previous one is still being transmitted. This ensures a free running system with zero dead time. The multistage pipelined architecture of the FADCs allows to store samples during a certain amount of clock cycles. This permits to recover samples, digitized before the high gain signal which triggered the read-out, that are necessary for the determination of the pedestal (see figure 4.1).

The data is transmitted in packets to the Grande DAQ station via an optical link. Each data packet has a fixed size and consists of 512 words, which results in a total size of 1,024 bytes. The structure of a data packet consists of:

- 2 *magic words*, headers which define the beginning of the data package,
- 1 word defining the Grande station number,
- 252 words containing the digitized samples of the high gain signal,
- 252 words containing the digitized samples of the low gain signal,
- 5 words containing the timestamp information.

[1] Analog Devices, AD9238
[2] EPSON, SG-8002JC series

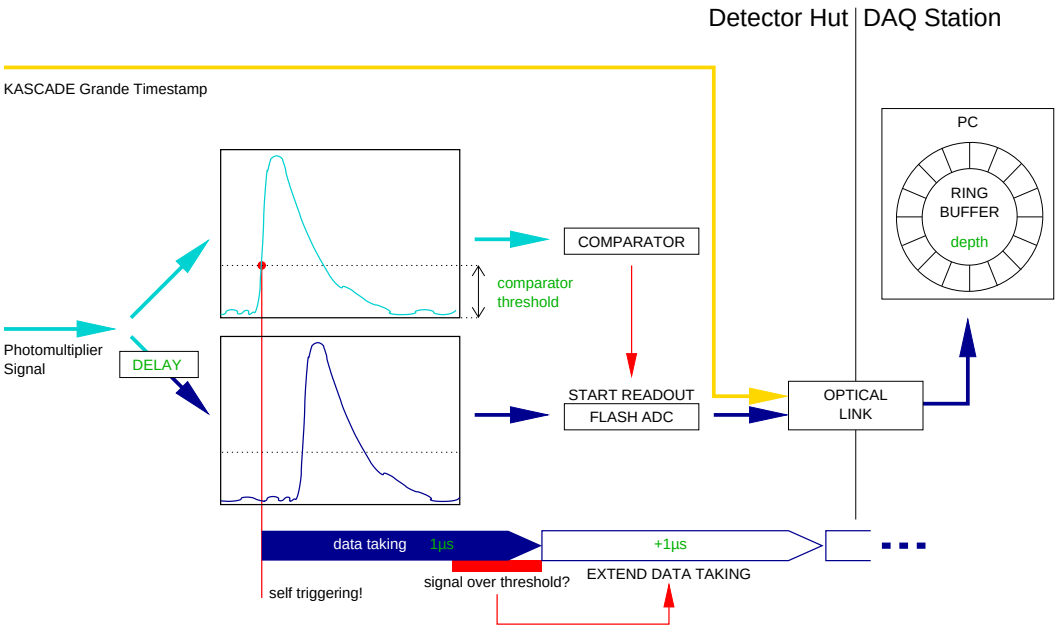


Figure 4.1: Schematic view of the FADC system. The high gain signal exceeds a certain threshold value and read-out is enabled. The digitization period lasts for 1.008 μs. The pipelined architecture allows to recover samples digitized before the threshold to be crossed. The data taking is extended if the high gain signal exceeds again the threshold in the last 200 ns of the digitization period. The digitized data are transmitted via optical link to the Grande DAQ station [Ove04].

The time stamp information consists of one word containing the 1 Hz counter value, two words containing the 5 MHz counter value and another two words containing the 62.5 MHz counter value. The 1 Hz and 5 MHz signals are incremented with every cycle of the central timestamp distributor, as received by the KGEMD board via an optical link. The 62.5 MHz is incremented by the internal clock of the KGEMD board. Both the 5 MHz and 62.5 MHz counters are reset at the reception of the 1 Hz signal.

4.2 The Storage Board

As previously mentioned, the central Grande DAQ station will be equipped with five VME storage boards (KGEMS). Their role is to receive and multiplex the data streams, packet-by-packet, from the 37 Grande detector stations, and to transfer the resulting data packets to the five first level PCs. The KGEMS board does not change or add information to the received data, since all the necessary information is already included in the data packets sent by the KGEMD board.

Each board has 8 input channels (optical link receivers) which ensure the reception of data packets from up to 8 detector stations (see figure 4.2). The data streams, received from the input channels, are buffered in a FIFO, which can store a maximum of 32 data packets. Then, the data is sent to an FPGA (Field Programmable Gate Array) that multiplexes the data streams. The resulting data stream is buffered into one common output FIFO, with a larger buffering capacity (128 data packets), and then it is sent to the first level PCs, via an LVDS cable connection.

Also, for test or debugging purposes, a VME bus interface is available on the KGEMS board.

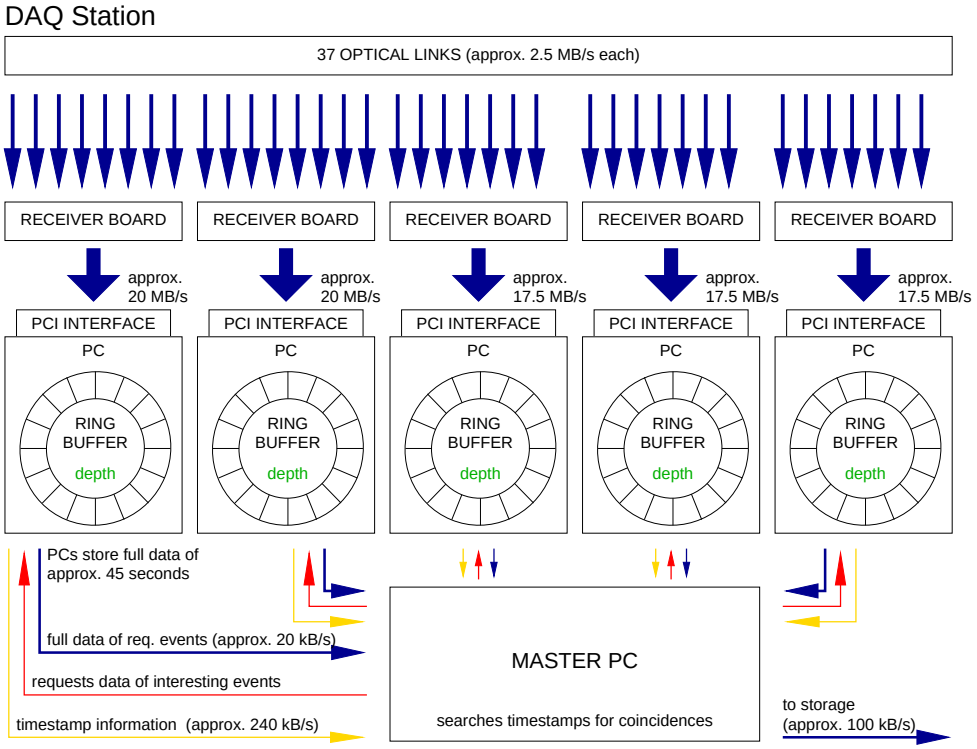


Figure 4.2: Schematic view of the receiver chain of the FADC system. Five storage boards receive data packets from 37 detector stations. The multiplexed data packets are send, via PCI interface to the first level PCs. These PCs collect and transmit the timestamp information to the master PC, which scans for coincidences and requests data about the interesting events [Ove04].

4.3 The First Level PCs and the PCI Interface Card

Five first level PCs collect, via PCI interface cards (KGEMP), the data transmitted by the storage boards. The PCs act as ring buffers, writing data to their internal memory. Each first level PC is going to be equipped with 1 GB of memory, of which only 128 MB will be used by the operating system. The remaining 896 MB are used to temporarily store the data. The first level PCs collect the information from the received data packets, i.e. the timestamps, the station numbers and the ring buffer positions of the data packets, and transmit it to the master PC. Also, the online software installed on each first level PC, performs the integration of the received pulse shapes, in order to obtain the energies of every single pulse, which will be filled into histograms to provide the energy spectra.

The KGEMPs are commercial PCI interface prototyping cards[3], which are connected through custom-made cards to the KGEMS boards. Their task is to transfer the data, via direct memory access, to the ring buffers, which are later accessed by the online software. A schematic view of the receiver chain is shown in figure 4.2.

4.4 The Master PC

The first level PCs store the full data for approximatively 45 seconds. The master PC checks the timestamps, as received from the first level PCs, in order to detect coincidences. A coincidence

[3] HK Meßsysteme GmbH, PCI-Proto-Lab

is defined by a certain number of single station events during a constant time window. When these requirements are fulfilled by certain events, the master PC requests the corresponding data packets from the first level PCs, and writes the data of air shower events to storage.

4.5 The Trigger Receiver Card

The KGEMT card is a custom-made trigger receiver PCI card, which is going to be installed in the master PC. Its task is to provide the interface to the central KASCADE-Grande trigger distributor. The master PC checks the trigger signal as part of the event finding conditions. This allows the FADC system to provide data for events triggered by other components of KASCADE-Grande, such as Piccolo or the KASCADE array. The KGEMT card provides inputs for the 1 Hz and 5 MHz signals of the KASCADE-Grande timestamp distributor, as well as for the central trigger signal. Also, it is equipped with a Global Positioning System (GPS) receiver, in order to set an exact time for the 1 Hz counter.

4.6 Conclusions

The new FADC DAQ system has the following advantages:

- it provides digitized photomultiplier pulse shapes containing information about the time development and the energy deposit. Additional parameters of the air shower event may be extracted from the pulse shapes;
- it is a free running system, ensuring zero dead time (longer events are completely sampled);
- it is a self-triggering system (in each station), with no hard-wired trigger conditions. The current Grande data acquisition system transmits data through 700 m long cables, favoring the noise interference;
- full data is available in the ring buffers of the first level PCs for 45 seconds, for the search of correlated single station events;
- it collects each single detector event and performs fast acquisition of the calibration spectra.

The only disadvantage of the FADC DAQ system is that it cannot provide a hardware trigger signal to the central KASCADE-Grande trigger distributor. Since the event selection is done by software, no electrical trigger signal can be generated at the time the event has been created.

Chapter 5

Monte Carlo simulation

Cosmic rays with energies beyond $5 \cdot 10^{14}$ eV can be studied via the extensive air showers (EAS) they induce in the atmosphere. Large ground-based detector arrays are used to sample the secondary particles generated by the collision of the primary cosmic ray with atmospheric nuclei. The detectors allow the measurement of signal shapes, whose properties are related to the physical observables of interest. Therefore, the characteristic properties of the observed extensive air showers can be reconstructed. By extracting the number of particles, that are incident on the array, and the corresponding energy deposits in the detectors, the number of particles created in the air shower and the primary cosmic ray energy can be estimated. Also, the arrival times of the particles in the detector can be used to determine the inclination of the shower.

The FADC system improves the data quality of the KASCADE-Grande experiment significantly (see chapter 4). In order to profit from this improvement, the development of fast and efficient reconstruction algorithms for the FADC data is mandatory. A rigorous study of the impact, that the FADC characteristics, e.g. resolution and sampling frequency, have on the accuracy of the measurements is the aim of this thesis. Comprehensive investigations of different reconstruction methods require a large collection of pulse shapes.

One possibility to create such a library of pulse shapes is to use *CORSIKA* (**CO**smic **R**ay **SI**mulations for **KA**scade) [Hec98] for a complete shower simulation and *CRES* (**C**osmic **R**ay **E**vent **S**imulation) for a detector response simulation. This full simulation chain consumes a large amount of time. To simulate showers, generated by primaries with energies of 10^{18} eV, up to one week per event is required. Therefore, it was decided to develop a faster tool, a Monte Carlo simulation that can create similar pulse shapes to those simulated by the the previously mentioned full simulation chain. Due to the lateral shower development, the particles arrive within several hundreds of meters from the shower axis. Thus the density of particles, that strike a detector, varies. The shapes of the measured pulses will vary accordingly, from a very high and narrow peak at close range, to an erratic form at larger distances. Therefore, three classes of signals depending on the position of a detector with respect to the shower core, were simulated: *close to the shower core*, *far from the shower core* and *at relative distance*. A library of pulses, consisting of 20000 events in the *far from the shower core* class and 10000 events in each of the other two classes, was produced. An event represents a recorded pulse during an acquisition time interval of 1 μs.

The present Monte Carlo tool simulates pulse shapes as response of only one detector, i.e. one Grande station, and not of the whole array. The intention is to concentrate on the pulse shape reconstruction and not on the reconstruction of the full air shower event. Furthermore, the individual characteristics of the electronics in the different Grande stations influence the data quality. The later analysis will be based on a model, that ranges from ideal to very noisy

situations and covers some of the particular effects of each Grande station (see chapter 6).

The Monte Carlo simulation will be described in detail in the following sections.

5.1 Underlying Signal and Single Particle Event Simulations

In general, Monte Carlo algorithms are stochastic methods employed to describe a real process. In order to develop a successful Monte Carlo simulation the processes involved have to be well known. Earlier measurements performed at the Forschungszentrum Karlsruhe helped to define a norm function for the signal. An experimental set-up consisting of a scintillator *paddle* (Grande-like scintillator) placed between two scintillating plates (to ensure the trigger condition) was used to measure the signal of incoming vertical muons [Bad03]. The area of the scintillating plates is about 20–30 cm^2 and they were placed such, that the total arrangement was not taller than 1 m. This ensured a small solid angle for the measurement of the incoming muons. A light guide and a photomultiplier tube (PM) were attached at one of the paddle's lateral ends.

When a particle interacts with the scintillator material, an amount of scintillation light proportional to the energy deposited in the detector is produced. This light is optically guided to the photomultiplier, which produces a current pulse proportional to the number of photons. In this way, the signal properties are related to the physical processes that are measured. By integrating the current pulse with an electric circuit with a time constant τ, a voltage pulse can be obtained. During the measurements, about 200 such photomultiplier signals fulfilling the coincidence condition were recorded with a digital oscilloscope of 2 ns resolution. The obtained photomultiplier signals were averaged and the resulting data points were described by the following parameterization:

$$U(t) = -\frac{k \cdot E_{dep}}{\tau - \tau_s} \left(e^{-\frac{t}{\tau_s}} - e^{-\frac{t}{\tau}} \right), \tag{5.1}$$

where k is proportional to the PM gain, τ_s is the scintillator decay constant and $\tau = RC$ (for a simple *RC* integrator).

Since the incoming vertical particles do not have the same momenta, the energy deposited in the paddle fluctuates. This leads to different measured pulse shapes. By averaging these pulses and taking the detector thickness (4 cm) into account, an 8 MeV mean energy deposit was assumed. This value corresponds to the minimum energy that an incoming vertical particle is able to deposit in the scintillator. Both the parameterization and the energy deposit (8 MeV) are presently implemented in the *CRES* detector response simulation.

Figure 5.1 shows the fit results of this parameterization to the averaged data points. The function does not describe the data and especially the upper points very well. The fit results have large errors. Theoretically, the decay constant of the NE102A plastic scintillator, used in the Grande station, is 2.4 ns [Leo93], while the fit returns a four times larger value. Besides these fit errors, detailed studies [Hus03] have shown that, applying fluctuations and using the function from equation 5.1 to fit the data points, in some cases one of the exponentials dominated the fit and the parameterization was reduced to a simple exponential which obviously failed to describe the data well.

Therefore, this parameterization was replaced by the Landau distribution. Later studies proved, that the Landau distribution is stable and describes the data better (see figure 5.2). This holds for the data points on the leading edge as well as for the upper points which are better described. In these simulations, the Landau function implemented in the ROOT physics analysis package [Bru97] has been used. This function is characterized by the most probable value and the width. It has been adapted from the CERNLIB routine G110 DENLAN [Köl93].

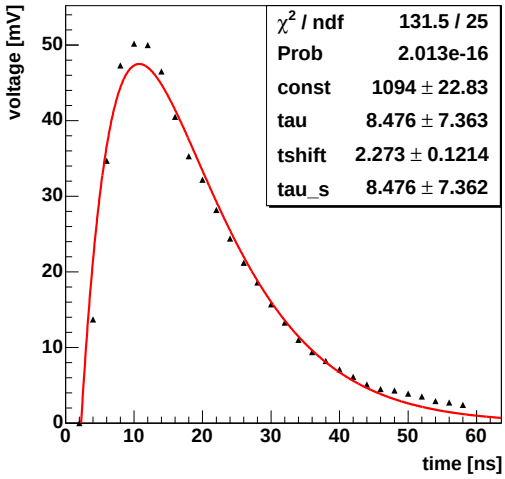


Figure 5.1: Fit of the experimental data points with the parameterization given in equation 5.1. The fit parameter *const* represents the scale factor, while *tshift* is a parameter introduced to ensure the fit of the data points on the leading edge.

Adopting the new parameterization, a Monte Carlo was developed to simulate photomultiplier pulses generated by single vertical muons. Assuming that the Landau distribution is a perfect description of the pulse of a single particle, and neglecting energy fluctuations, the FADC performance can be estimated.

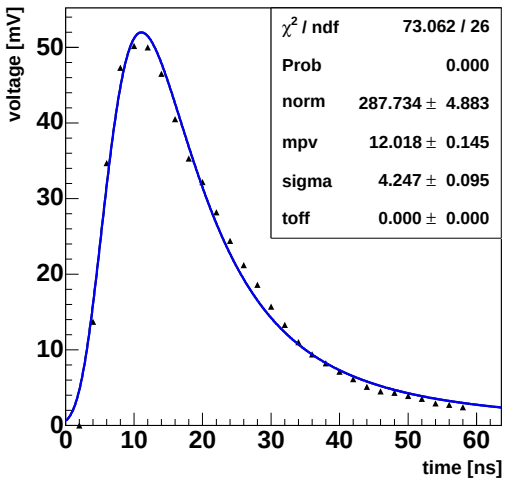


Figure 5.2: Fit of the experimental data points with the Landau function.

First, the measured data points were fitted with the Landau distribution. Then, the fit parameters were extracted and used to define the basic *true signal*. In order to simulate the lack of a correlation between the arrival time of the photomultiplier pulse and the clock signal, that drives the A/D converter, the simulated signals were shifted in time with random offsets. This time uncertainty is smaller than the sampling interval of the A/D converter. Assuming the characteristics of the FADC described in the previous chapter, e.g. an effective sampling frequency of 250 MHz, the offset was randomly generated within a time interval of 4 ns.

The effect of white noise, equally distributed over the whole frequency spectrum, was also considered. First, the baseline of the signals was shifted in amplitude by a constant *pedestal* in order to avoid negative amplitudes when white noise is applied. Then, because the amplitude of the white noise is described by a normal distribution, the constant pedestal was smeared in amplitude with this distribution. Any such a *realistic signal* is a hypothetical single particle event.

The signals were evaluated at every 4 ns and, considering the 12-bit resolution of the FADC, the corresponding amplitudes were binned into 4096 discrete intervals. The resulting data points were fitted with a parameterization defined by a Landau function and a constant pedestal. The fit results represent the *reconstructed* values and the comparison with the *true* values is done by dividing their difference by the *true* values, which results in the *R*-values:

$$R = \frac{reconstructed - true}{true} \quad . \tag{5.2}$$

Now the impact of the FADC characteristics on the accuracy of the measurement can be studied. Assuming that the area below the pulse describes the energy deposit, a *true energy* is considered by integrating the *true signal*. For each generated single particle event a *reconstructed energy* is defined by integrating the pulse and subtracting the constant pedestal. By applying equation 5.2, the quality of the energy reconstruction can be estimated.

The *time resolution* is another criterion for the quality of a measurement. It defines the precision with which the time characteristics of a signal could be determined. In these simulations, the *true time* of the pulse, i.e. the arrival time, is considered to be defined by the most probable value of the Landau function. The digitized values are fitted with a parameterization defined by the Landau function and a constant pedestal. From the fit results, the most probable value can be extracted. A *reconstructed time* can be obtained by subtracting the randomly generated phase, between the arrival time of the pulse and the FADC clock signal, from the most probable value of the fit. Then, the residuals are obtained by subtracting the *true time* from the *reconstructed time*. The time resolution of the measurement is defined as the standard deviation of the residuals.

The four histograms in figure 5.3 show the distributions of the time residuals and the *R* values (see equation 5.2) of the integral, width and pedestal variation. An energy resolution of about 2% and a time resolution of $(0.1255 \pm 4 \cdot 10^{-4})$ ns were achieved (figure 5.3 *a*,*b*). In these simulations the integration was performed with the adaptive Gaussian quadrature method implemented in the ROOT framework. This numerical method employs unequally spaced intervals. In the subsequent chapter the efficiencies of the integration methods, that employ equally spaced samples, are investigated.

5.2 The Multi-Particle Event Simulations

In the previous section only single particle events were discussed. It is necessary to extend the analysis to a scenario, in which many particles distributed over a certain time window penetrate

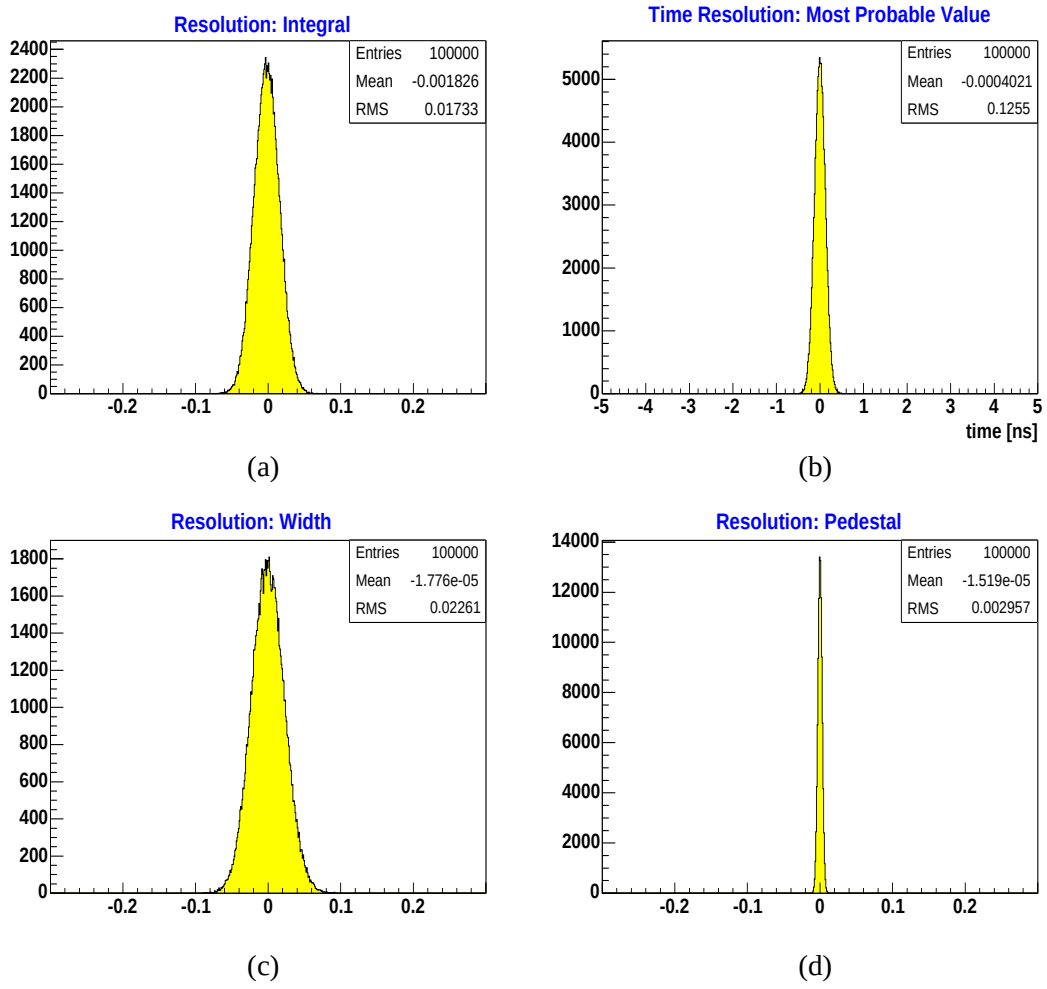


Figure 5.3: The time residuals (*b*) and the *R = (reconstructed-true)/true* values of the integral (*a*), width (*c*) and pedestal (*d*) in the single particle event simulations.

the scintillator (see chapter 2). This implies to pile up the single particle pulses given by the parameterization such, that the combined pulse shape describes the sum of the individual energy deposits. Applying successively this algorithm, a large collection of pulse shapes may be created.

As already mentioned, there is another possibility to create these pulse shapes, by using the *CORSIKA-CRES* simulation chain. One of the reasons to develop a separate tool is the large time consumption of the simulation chain. Furthermore, the *CRES* detector response simulation uses the described parameterization (see equation 5.1), which did not prove to be a reliable solution. Due to these reasons, a faster tool was developed, that can simulate pulse shapes looking similar to those created by *CRES* (fig 5.4).

This tool does not simulate real phenomena in the atmosphere or detector. But, using a simple model to generate a realistic output, the initial goals can be nevertheless achieved (see chapter 1). The Monte Carlo does not differentiate between particles, e.g. muons, electrons, photons, etc., but simulates randomly total energy deposits as a consequence of their interactions with the detector. The amount of overall energy deposits is related with the particle density and with their individual energy deposits. The former is randomly generated from an uniform distribution. Apportioning the overall energy, individual energy deposits, and hence, the number of particles are created. Additionally, a reference arrival time is assigned to each particle, and

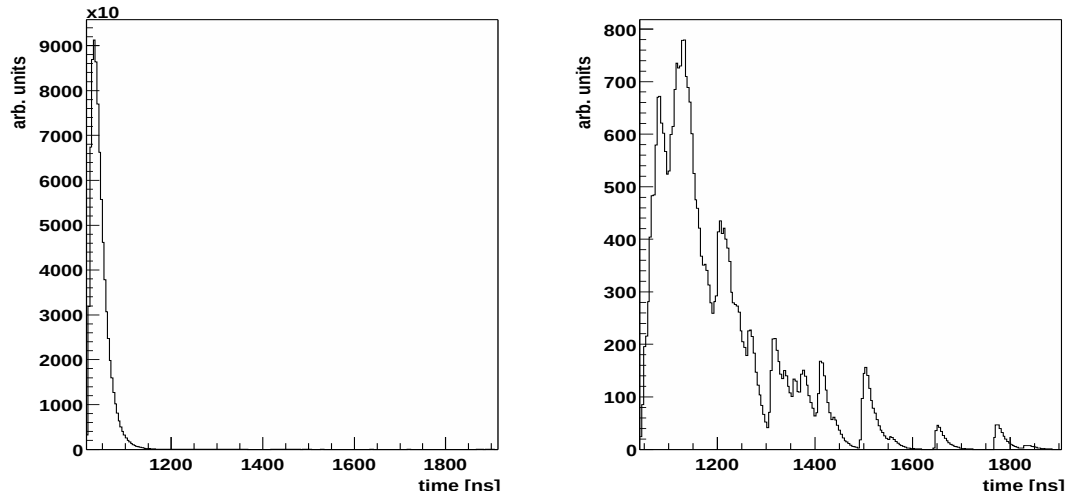


Figure 5.4: Examples of pulses created by the *CRES* detector response simulation. The shape of the signals is strongly influenced by the density of particles that hit the detector. The sharp peak in the *left* plot is a result of thousands of particles interacting with the detector within a short time window. The *right* plot shows the contribution of several particles distributed over a larger time window.

the individual contributions are described with the help of a parameterization. When the single particle signals are piled up, pulse shapes looking like those in figure 5.4 are expected. The shape of these combined pulses is strongly influenced by the particle density and their distribution in time. In conclusion, the Monte Carlo is separated in three classes with specific particle densities and arrival times. The *close to the shower core* class contains pulses generated by thousands of individual contributions distributed over a short time window, while the *far from the shower core* describes the other extreme, a few or several particles distributed in a much larger time window. The *relative distance* class is an intermediate step between the other two classes.

The following subsection presents the Monte Carlo method in more detail and discusses the obtained results.

5.2.1 The Monte Carlo Description

One of the main features of the photomultiplier tube is its fast time response. This is the time, in which the anode signal rises from 10% to 90% of the peak amplitude, when the photocathode is illuminated by a light pulse of a few tens of picoseconds width. In order to ensure a good description of the photomultiplier signal, the leading edge of the Landau function was replaced with the tangent in the leading edge's inflection point, i.e. the point where the function changes concavity. In figure 5.5 the normal Landau function is plotted with a *solid* line overlayed with the *dashed* line for the new parameterization. The differences on the right hand side of the peak are minimal (figure 5.5 *a*), but the rising edge is much better described and the rise time of the signal is 0.5 ns smaller (figure 5.5 *b*). The *vtp* parameter in the statistics box defines the slope transition point x_{tp} on the leading edge, i.e. the turning point where the truncation was applied:

$$x_{tp} = mpv - width \cdot vtp \quad . \tag{5.3}$$

The most probable value (*mpv*) and the width are given in the same statistics box. Extracting these fit parameters, a *norm function* was defined. This is the function, that is going to be used

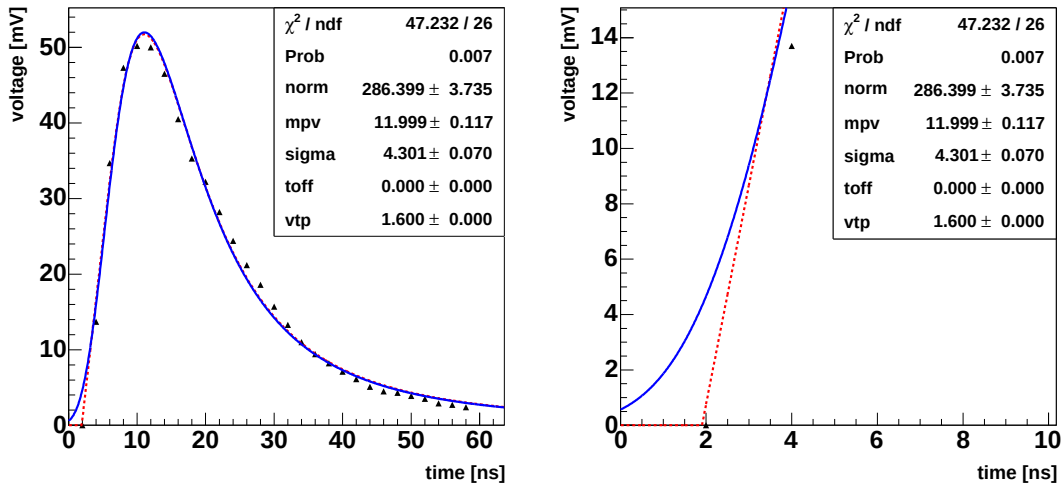


Figure 5.5: Fit of the experimental data points with both the Landau function (*solid* shape) and the new parameterization (*dashed* shape). The rising edge is better described by the new parameterization (plot *b*).

in the further simulations. The area below the shape was considered to describe the norm energy deposit E_{1p} of a single particle.

The classification of the Monte Carlo requires different particle densities to be considered in each class. The total energy deposit E_{dep}^{total}, recorded in the detector during an 1μs event, is a quantity strongly related with the particle density. In order to generate it, a basic number of particles N_{part}^{basic} specific to every class was considered. This number is randomly chosen from an uniform distribution within certain limits. The values implemented in the Monte Carlo simulation for this parameter are given in the table 5.1.

Considering the norm energy, the overall deposited energy is:

$$E_{dep}^{total} = E_{1p} \cdot N_{part}^{basic}. \tag{5.4}$$

Monte Carlo class	N_{part}^{min}	N_{part}^{max}
far from the shower core	1	200
at relative distance	150	5000
close to the shower core	2500	10^4

Table 5.1: Input values of the basic number of particles N_{part}^{basic} .

This overall energy is then randomly distributed such, that the sum of the individual energy deposits equals the total energy created according to equation 5.4. By randomly choosing an *energy fraction* (f_E) from a Gaussian distribution of $mean = 1$ and $width = 2$, the individual energy deposits were generated as follows:

$$E_{dep}^{indiv} = 10 \cdot f_E \cdot E_{1p}. \tag{5.5}$$

In order to avoid situations, in which one particle gets most of the total deposited energy E_{dep}^{total}, the individual energy deposit E_{dep}^{indiv} was limited to $10 \cdot E_{1p}$. This is a technical implementation rather than a simulation of a real phenomenon in the detector. The idea was to simulate pulses looking similar to those created by the CRES detector response simulation. On the other hand, the energy deposit in a detector at a small distance from the core of an air shower is very large. The measured photomultiplier signal is very high and narrow and represents the contribution of thousands of particles striking the detector in a short time window. Therefore a situation, in which E_{dep}^{total} is distributed in a very large percentage to only one particle, had to be avoided.

The algorithm, which distributes the overall energy E_{dep}^{total}, works like this. At first $E_{dep,\,1}^{indiv}$ generated by applying equation 5.5 is subtracted from E_{dep}^{total} such, that a remaining energy $E_{dep,\,1}^{rest}$ is obtained. If this quantity is positive and non-zero, another $E_{dep,\,2}^{indiv}$ is created and subtracted from $E_{dep,\,1}^{rest}$. The algorithm continues until the $E_{dep,\,n}^{rest}$ is either negative or zero. In these cases, the generated value of $E_{dep,\,n}^{indiv}$ is replaced by the previous $E_{dep,\,n-1}^{rest}$. Counting the number of generated individual energy deposits, a final number of particles N_{part}^{final} is obtained:

$$E_{dep}^{total} = \sum_{k=1}^{N_{part}^{final}} E_{dep,\,k}^{indiv} \quad . \tag{5.6}$$

The distribution of the resulted N_{part}^{final} in all 3 classes is shown in figure 5.6. Plot (*a*) shows the contribution of the *far from the shower core* class. The number of generated pulse shapes specific to this class is two times larger, when compared with the other two classes. The range from which N_{part}^{basic}, and hence, N_{part}^{basic}, is generated is short (see again table 5.1). Therefore, the probability to generate many times the same number, is large. The probability to have identical pulse shapes decreases with the increase of the number of particles, that contribute to the pulse. Except in the case, when only a few particles contribute to the final pulse, individual pulse shapes are assured. The individual energy deposits are randomly created, which is reflected in the shapes of the corresponding single pulses. Moreover, to every single pulse a random *reference time* is assigned, which shifts the signal along the time axis.

The *reference time* t_k, corresponding to the *k*th particle of the final pulse shape, is the time, when the energy contribution of the particle is detected by the photomultiplier. This time is intrinsically related to the real arrival time of the particle in the detector. Due to the atmospheric shower development most of the particles produced along its axis are located in the rim of the shower disk pointing into the direction of the Earth. Therefore most of these particles arrive in the detector early, while the others arrive with the rest of the shower plane. On the other hand, the particle density depends on the relative position of the detector with respect to the shower core. These effects are reflected in the shape of the pulses. At large distances a few particles are distributed over a considerable time extent, while at small distances the particle density is larger and the arrival time is distributed over a much shorter time window. In order to simulate this phenomena, an exponential distribution was chosen to generate the random reference times.

Additionally, two generation parameters α and τ_c are introduced. The parameter α represents the mean of the exponential distribution (see appendix A, section A.1). It ensures that the particles are distributed over different time extents, while the τ_c parameter defines a time limit for the spread of the individual pulses such, that:

$$t_{rndm} \leq \tau_c, \tag{5.7}$$

where t_{rndm} is the time randomly generated from the exponential distribution. Both α and τ_c are not fixed parameters, but are randomly generated within certain limits (equation 5.7) and their

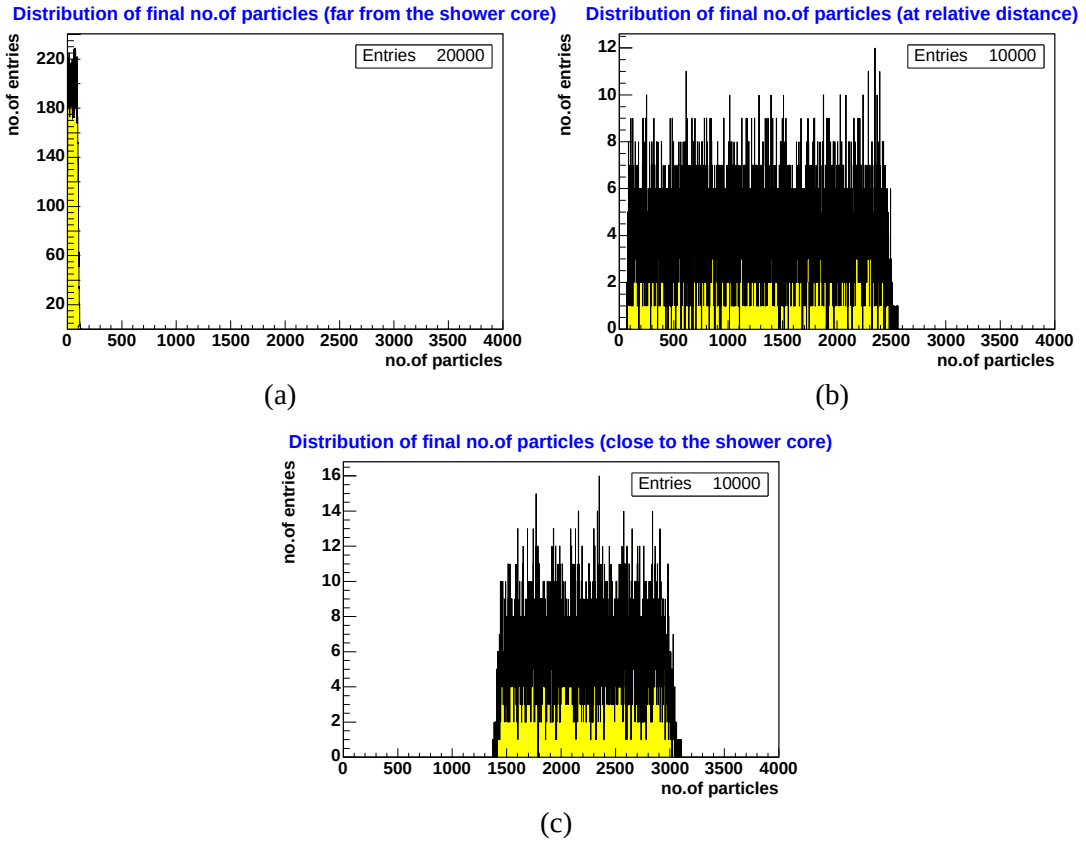


Figure 5.6: The distribution of the number of particles per event simulated by the Monte Carlo algorithm for each pulse class: *far from the shower core* (*a*), *relative distance* (*b*), *close to the shower core* (*c*).

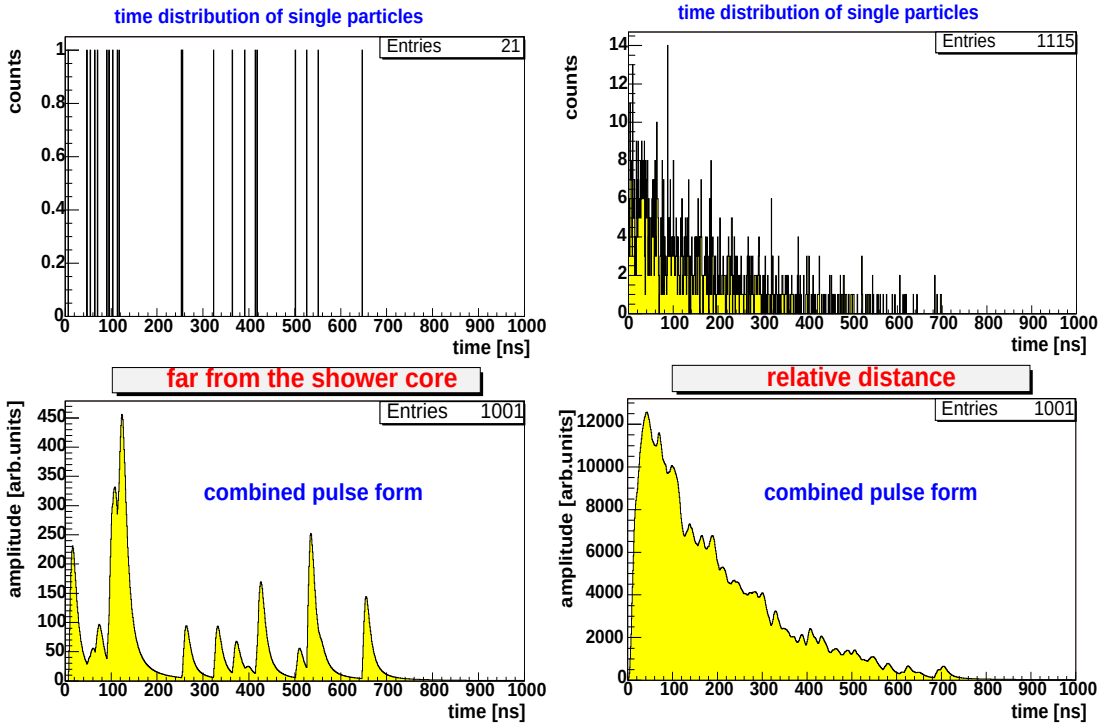
limits are distinctive from class to class (see table 5.2).

Monte Carlo class	α_{min}	α_{max}	τ_c^{min} [ns]	τ_c^{max} [ns]
far from the shower core	350	750	600	850
at relative distance	150	350	350	750
close to the shower core	0.1	150	20	350

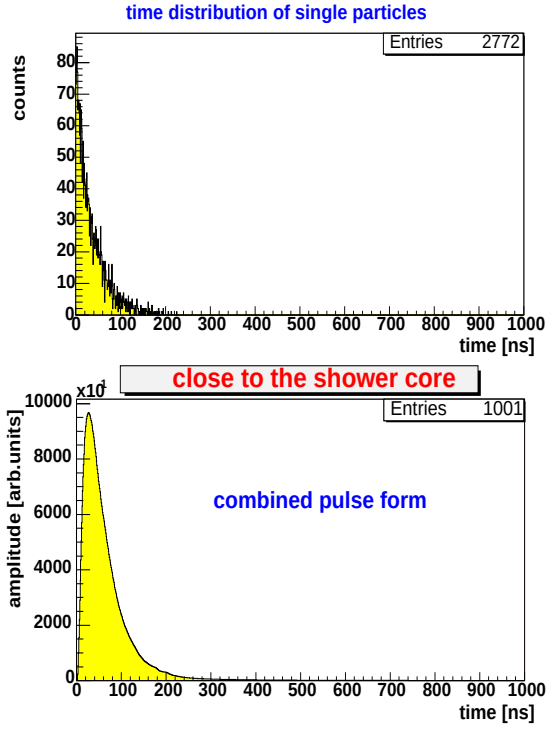
Table 5.2: Input values of the generation parameters α and τ_c.

These parameters do not have a clear physical meaning, since in reality a significant contribution of particles may be observed even at 900 ns. But, as it was mentioned in chapter 4, the FADC system can deal with this kind of events by extending the aquisition time to 2 μs. The analysis of the obtained pulses will be similar to that of the pulses detected in 1 μs, the main difference being in the double number of samples.

By using the norm function, the contribution of each particle to the signal can be described. The function is scaled up or down in amplitude according to the $E_{dep,k}^{indiv}$. A scale factor can be



(a) (b)



(c)

Figure 5.7: Examples of pulse shapes created with the Monte Carlo tool, and the corresponding time distributions of the single particles, that form the combined pulse shapes. The number of entries in the time distribution histograms represents the number of particles, while in the pulse histograms it denotes the number of bins.

defined as:

$$S_k = \frac{E_{dep,k}^{indiv}}{E_{1p}} \quad . \tag{5.8}$$

Shifting the norm function on the time axis with a certain amount t_k, a part of its Landau-like tail is cut away at the 1μs limit. The function parameters, i.e. the most probable value and the width, are not affected by this cut, but the area below the shape is changed. The losses are minimal, e.g. 0.5% for 500 ns time shift. In order to ensure a correct description of the $E_{dep,k}^{indiv}$, the E_{1p} should be replaced by the integration of the shifted pulse. First the norm function is shifted with a random t_k and the area below the shape is evaluated by integrating the shifted pulse over the 1μs interval. Then the pulse is scaled with the factor:

$$S_k = \frac{E_{dep,k}^{indiv}}{I_k}, \tag{5.9}$$

where I_k stands for the value of integration. In this way the individual signals describe the corresponding energy deposits of the single particle. The next step in the Monte Carlo is the pile-up of the individual pulses. The pulses are summed and the combined pulse shapes are saved in ROOT files together with the individual functions and their own parameters (constant, mean, width, reference time and scale factor). This allows to store continuous functions and to evaluate them in any point when retrieved. For a better reconstruction analysis the generation parameters ($E_{dep,k}^{indiv}$, α and τ_c) are stored as well.

5.2.2 Results and Conclusions

Figure 5.7 shows examples of simulated events specific to each class of the Monte Carlo tool. In each group, the upper plots represent the t_k *reference time* distribution of the single particles, that form the combined pulse shapes in the lower plots. Each event is simulated in a 1μs time window. The arbitrary units of the amplitude mean that no calibration was considered.

In the *far from the shower core* example (figure 5.7 *a*), the main characteristics of this class, the low particle density, is well described in both graphs. Several particles, each with a different contribution in energy, are distributed over a large time extent and the resulting pulse has an erratic shape. In the *relative distance* example (figure 5.7 *b*), the particle density is much higher and the time distribution is not as flat, but tends to follow the underlying exponential distribution. The shape of the combined pulse is influenced by this characteristic. The last example, *close to the shower core* class (figure 5.7 *c*), shows a high and narrow signal corresponding to a large number of particles striking the detector in a short time window. These three examples emphasize the sensitivity of the detector to different shower core positions.

A large collection, consisting of 20000 specific pulses for the *far from the shower core* class and 10000 specific pulses each for the *relative distance* and *close to the shower core* classes, was created. The time consumption is very small when compared with the requirements of the *CORSIKA-CRES* simulation chain. Practically, using the present Monte Carlo less than 24 hours are needed to generate 40000 pulses.

The following chapters are dedicated to an analysis based on this collection of pulse shapes. Implementing time-base errors and amplitude fluctuations, the impact of the FADC characteristics on the overall time and energy resolution is investigated.

Chapter 6

Energy Reconstruction

Ground-based detector arrays allow measurements of the extensive air showers, simultaneously in many points, over a large detection area. The Grande array of the KASCADE-Grande experiment is sensitive to the charged component of the air shower. Muons and electrons represent the dominant flux of charged secondary particles that contribute to the energy deposit in the detectors. When these charged particles interact with the plastic scintillators, the energy transferred will be partially dissipated as heat and mostly absorbed, by the atoms of the detection material. This latter process rises the atoms into excited states. Neglecting the amount of energy lost through lattice vibrations, the decay of the excited atoms into a lower energy state is accompanied by the release of a photon carrying the excess energy. These photons, if not reabsorbed, i.e. their energy is smaller than the energy needed to produce another excited state, are optically guided to the photomultipliers. Electrical signals proportional with the number of photons is then produced. The size of the photomultiplier signals is strongly influenced by the energy deposit in the detector.

Since the energy deposit is a valuable information for the shower reconstruction [Gla03], it is very important to know how accurate this information can be retrieved from the photomultiplier pulses. In this chapter, a study of the energy reconstruction is presented. The ideal pulse shapes created with the Monte Carlo tool (see chapter 5) are used as input for the investigations. The area below the pulse shape was considered to define the energy deposit. In order to retrieve it, the pulse shapes are integrated, such that the result is a quantity in arbitrary units of amplitude multiplied by time. The precision of the reconstruction and the time consumption of different integration methods are investigated.

Apart from the integration methods, there are other effects that can affect the quality of the reconstruction. The digitization of the signals involves inherent quantization losses. The level of noise, arising from the electronics of the Grande stations or from the surrounding environment, sets a low limit of accuracy for the quantity to be measured.

6.1 Numerical Integration Methods

As described in chapter 4, the FADC system digitizes the signals every 4 ns, which corresponds to an effective sampling rate of 250 MHz. When the amplitude of the high gain signal exceeds a threshold value, the read-out is enabled and 252 digitized samples, corresponding to a digitization period of 63 clock cycles, are transmitted to the central DAQ station. The photomultiplier pulses, received from each station, have to be integrated in order to estimate the energy deposit in the single station. Therefore, the precision and accuracy of different integration methods needs to be investigated. Since the data acquisition software of the FADC DAQ system has to process

more than 90,000 single station events per second [Ove04], the CPU time consumption is an important evaluation criterion.

As previously mentioned, the pulse shapes created with the Monte Carlo tool are used as input for this study. They are *ideal pulses*, undistorted by any time-base error or amplitude fluctuations. It is convenient to apply and to interpret the results of different integration methods applied to ideal pulses. The implementation of any noise characteristic can induce the appearance of a bias, and the analysis can result in a misjudgement. Unlike the FADC system, which outputs 252 digitized values, the simulated pulses ensure only 251 points for sampling and digitization.

The evaluation of the signals at discrete points, with a fixed time interval, restricts the investigation to numerical integration methods for equally spaced samples. For this study, the following classical *closed extended* formulas were considered:

$$\int_{x_1}^{x_N} f(x)dx = h\left[\frac{1}{2}f_1 + f_2 + f_3 + \ldots + f_{N-2} + f_{N-1} + \frac{1}{2}f_N\right] \quad , \tag{6.1}$$

$$\int_{x_1}^{x_N} f(x)dx = h\left[\frac{1}{3}f_1 + \frac{4}{3}f_2 + \frac{2}{3}f_3 + \frac{4}{3}f_4 \ldots + \frac{2}{3}f_{N-2} + \frac{4}{3}f_{N-1} + \frac{1}{3}f_N\right] \quad , \tag{6.2}$$

$$\int_{x_1}^{x_N} f(x)dx = h\left[\frac{5}{12}f_1 + \frac{13}{12}f_2 + f_3 + \ldots + f_{N-2} + \frac{13}{12}f_{N-1} + \frac{5}{12}f_N\right] \quad , \tag{6.3}$$

$$\int_{x_1}^{x_N} f(x)dx = h\left[\frac{3}{8}f_1 + \frac{7}{6}f_2 + \frac{23}{24}f_3 + f_4 \ldots + \frac{23}{24}f_{N-2} + \frac{7}{6}f_{N-1} + \frac{3}{8}f_N\right] \quad . \tag{6.4}$$

They are *closed* formulas because they use the value of the function at the endpoints, i.e. f_1 and f_N for a function defined within the range $[x_1, x_N]$, and *extended* because they are built up from multiple copies of lower order. They belong to a large family of numerical integration methods, called *Newton-Cotes formulas* [Pre02].

The first equation is the *extended trapezoidal rule*. It employs straight lines as interpolating polynomials between two successive points. The formula can be easily derived by applying Euclidean geometry. In each interval the two data points are connected by a straight line such, that the points and their projections on the x-axis form a trapezium. Then the integral is given by the area of the trapezium:

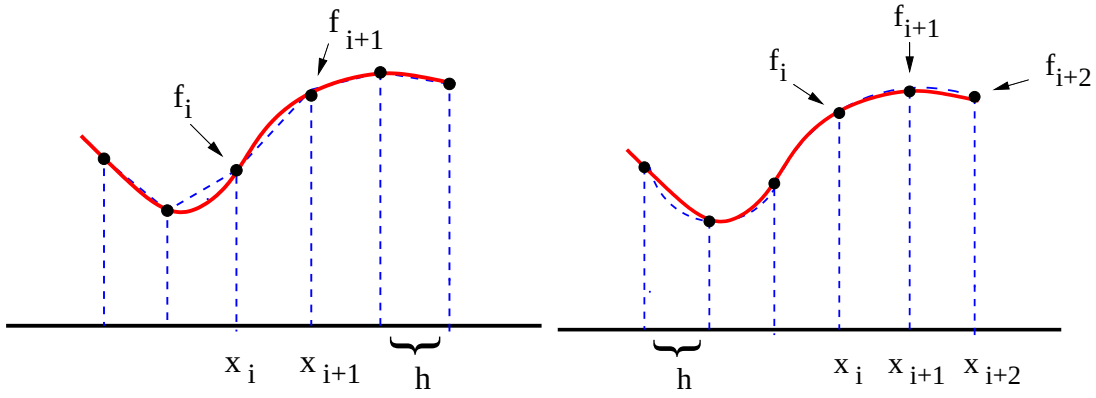
$$\int_{x_i}^{x_{i+1}} f(x)dx = h\ \frac{f_i + f_{i+1}}{2} \quad , \tag{6.5}$$

where f_i and f_{i+1} represent the evaluations of the function in the points x_i and x_{i+1}, and h is the width of the interval (see figure 6.1 *a*). This formula is often quoted as the *2-point rule*. Applying equation 6.5 for each interval and using the additive property of the integrals, equation 6.1 is obtained.

The equation 6.2 is the *extended Simpson's rule* [Pre02]. It is an extension to N intervals of the classical *Simpson's one-third rule* (or the *3-point rule*), which approximates the integrand with parabolic arcs (see figure 6.1 *b*). Thus, the interpolation is performed using three successive points:

$$\int_{x_i}^{x_{i+2}} f(x)dx = \frac{h}{3}\left[f_i + 4f_{i+1} + f_{i+2}\right] \quad . \tag{6.6}$$

The extended formula requires an even number of intervals. This condition is fulfilled by the 251 digitized samples, corresponding to a digitization period of 1μs. Summing the integrations in each successive pair of intervals, the extended formula is obtained.



(a) Trapezium rule (b) Simpson's rule

Figure 6.1: Graphical representations of the *trapezoidal* and *Simpson's one-third* rules. The function is evaluated at discrete points, equally spaced by the interval h. Trapezium rule employs straight lines as interpolating polynomials between two successive points, while Simpson's rule approximates the integrand with parabolic arcs between three successive points.

Equations 6.3 and 6.4 are built up from the combination of different low order rules. Equation 6.3 is obtained by applying the Simpson's one-third rule for each pair of intervals within the range $[x_2, x_{N-1}]$, and the trapezium rule in the intervals $[x_1, x_2]$ and $[x_{N-1}, x_N]$, such that:

$$\int_{x_1}^{x_N} f(x)dx = h\left[\frac{1}{2}f_1 + \frac{5}{6}f_2 + \frac{4}{3}f_3 + \frac{2}{3}f_4 + \ldots + \frac{2}{3}f_{N-3} + \frac{4}{3}f_{N-2} + \frac{5}{6}f_{N-1} + \frac{1}{2}f_N\right]. \quad (6.7)$$

Averaging this formula with the extended Simpson's rule, the middle terms will have unity coefficients, and the equation 6.3 – referred in this work to as *Averaged Simpson* – will be obtained. Equation 6.4 – referred to as *Modified Simpson* – is constructed by fitting cubic polynomials through successive groups of four points [Pre02].

The precision of reconstruction of these four integration methods is investigated by applying the sampling procedure. The simulated pulse shapes are evaluated every 4 ns over the 1 μs interval. The 251 data points define the input for the integration method. The result of the integration represents the *reconstructed* energy deposit, and the comparison with the *true* energy deposit, generated by the Monte Carlo simulation, is done by dividing the energy differences by the *true* energy deposit:

$$R = \frac{E_{dep}^{reconstr} - E_{dep}^{true}}{E_{dep}^{true}} \quad . \quad (6.8)$$

The relative energy differences (the R values in the above equation) are filled into histograms and the expected result should be distributed according to a normal distribution with the mean value centered at zero. The obtained histograms are fitted with a Gaussian function.

Figures 6.2 and 6.3 show the relative values for all three classes defined in the Monte Carlo algorithm. The *extended Simpson's rule* and the trapezium method provide slightly more accurate results than the other two integration methods. The latter methods show a tendency to overestimate the *true* values for pulses defined in the *close to the shower core* class. The shape plays an important role when the signal is evaluated and integrated with a finite number of discrete points. The distributions of relative values from the *far from the shower core* class show the largest spread, i.e. in the order of 0.1%. These deviations are the consequence of the erratic forms that characterize the pulses from this class. In the *relative distance* class, the signals

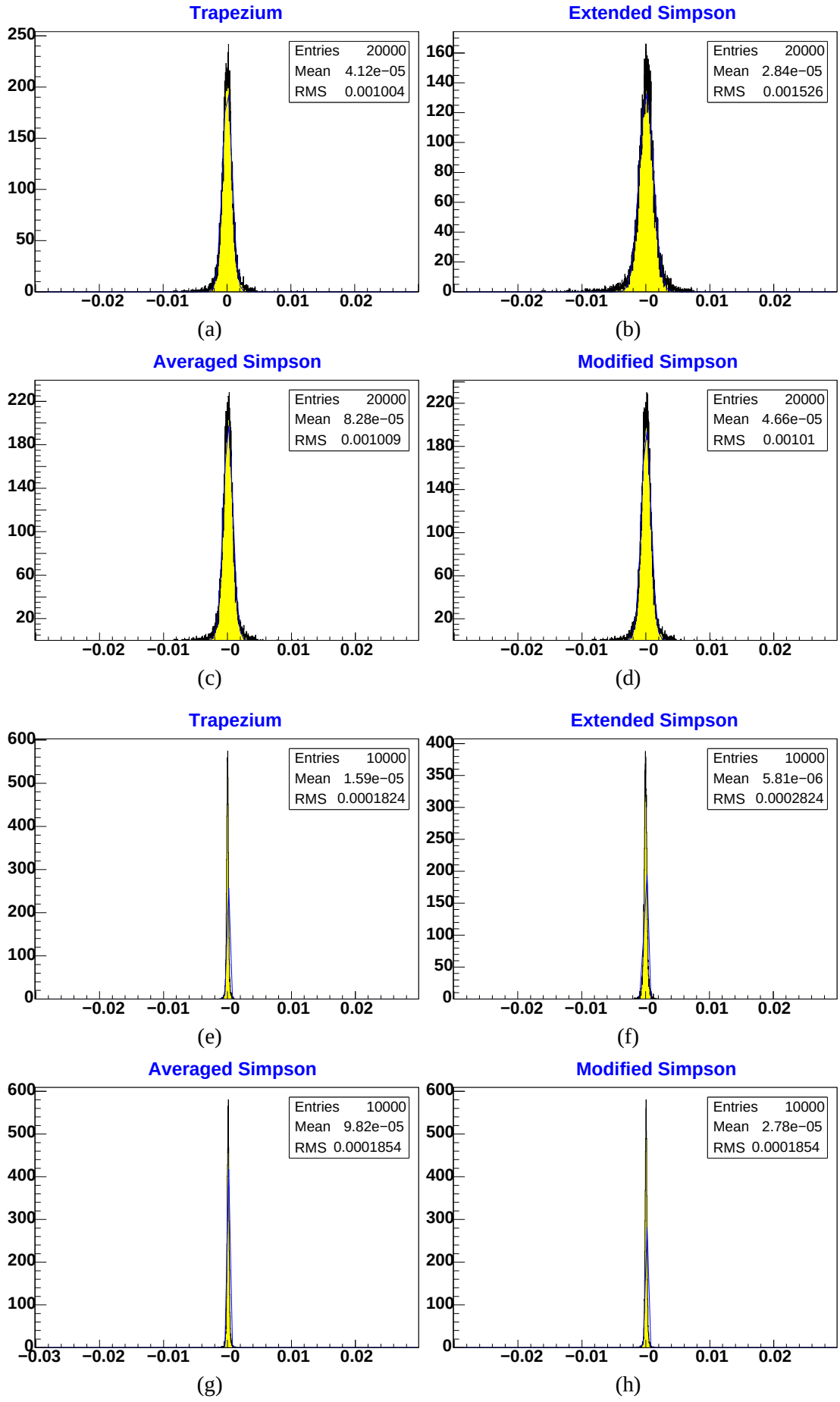


Figure 6.2: The relative energy differences in the *far from the shower core* (*a*, *b*, *c*, *d*) and *at relative distance* (*e*, *f*, *g*, *h*) classes, as results of the four integration methods.

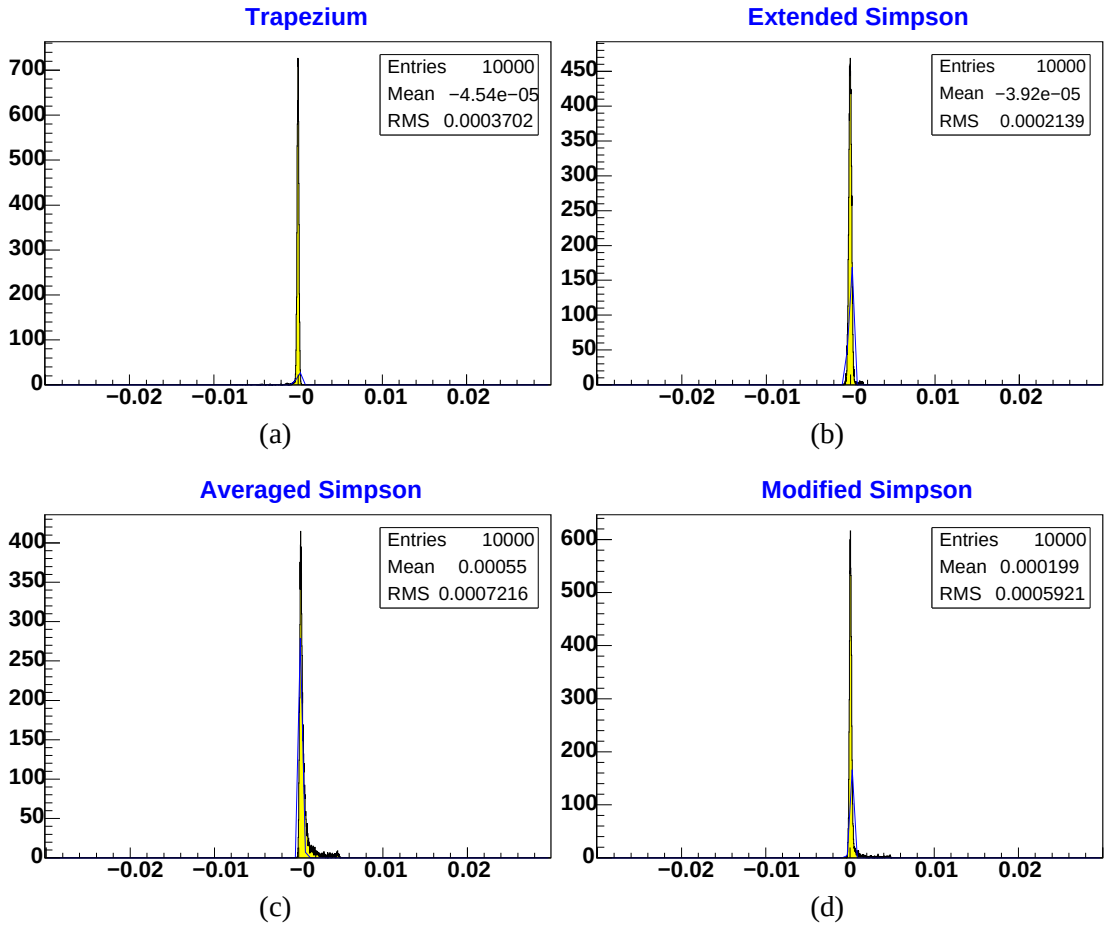


Figure 6.3: The relative energy differences for the *close to the shower core* class, as results of the four integration methods.

tend to have a more defined shape and the spread is one order of magnitude smaller. The pronounced Landau-like tail, that is a characteristic of the pulses from the *close to the shower core* class, is responsible for the systematic deviation shown by the last two integration methods. The Gaussian fit provides a better information about these deviations (see results in appendix B, section B.1). The percentage of outliers for *averaged Simpson* and *modified Simpson* integration methods is large, about 9.6% and 5.3%, respectively. Small deviations, due to the same tail, are as well observed for the Simpson and trapezium rules, but their effect is minimal.

As mentioned before, the CPU time consumption represents another evaluation criterion for these integration methods. This implies to determine the amount of time the CPU takes to perform the integration, assuming that the process makes exclusive and uninterrupted use of the CPU. Since the amount of time needed to perform a single integration is too small to be determined, each pulse was integrated 10,000 consecutive times for each integration method separately. The resulting amount of time was then averaged over the total number of pulses. There is no strong reason to split this analysis into different classes of pulses, since the integration depends rather on the number of calculations than of the shapes. The results for each integration method are given in table 6.1. As expected, the trapezium rule is the fastest method, due to the simple mathematical operations involved. The extended Simpson's rule is quite slow when compared with the other methods, i.e. about 8 times slower than the trapezium rule. This is explained by the use of conditional steps in the alternation of the odd and even terms.

Integration method		CPU time consumption *(milliseconds / 10^4 integrations)*
Trapezium	*(equation 6.1)*	2.9
Extended Simpson	*(equation 6.2)*	24.2
Averaged Simpson	*(equation 6.3)*	3.1
Modified Simpson	*(equation 6.4)*	3.1

Table 6.1: Averaged CPU time consumption for the four integration methods .

This study showed that, in principle, any of the integration methods can be used. In all presented situations there are no cases of relative energy differences larger than 2%. However, the trapezium method was chosen due to a consistent behavior and little CPU time consumption. The Simpson's extended formula, otherwise showing similar results as the trapezium method, has the disadvantage of a large CPU time consumption. The *averaged Simpson* and *modified Simpson* integration methods showed a limited performance in integrating the pulses specific to *close to the shower core* class, and a slightly larger CPU time consumption than the trapezium method.

6.2 Digitization and Quantization Errors

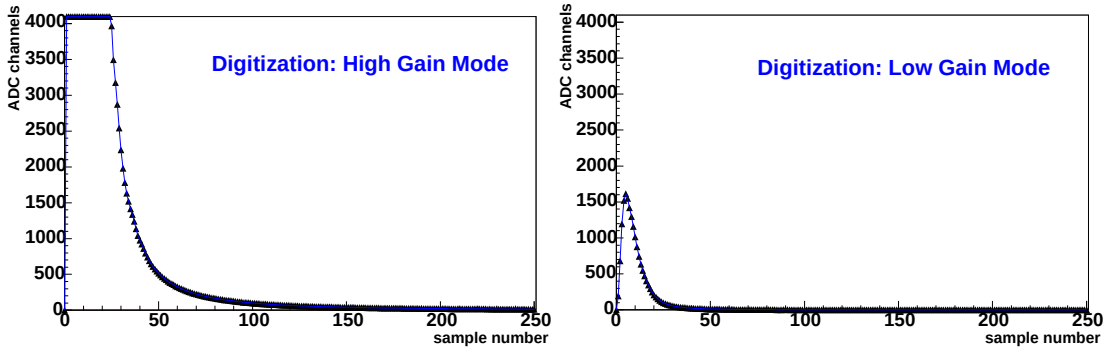
The *digitization,* or *quantization,* represents the conversion of the sampled analog signal amplitudes into digital values. It plays an important role in the signal processing and, due to the inherent errors that it generates, it should be carefully studied. These errors represent the difference between the real and the digitized values, and they are generally quoted as *quantization errors*.

This section is dedicated to a study of the impact that the 12-bit FADC caracteristics have on the accuracy of the energy reconstruction. The pulse shapes, generated with the Monte Carlo tool, are sampled and digitized, and the 251 resulting values are used as input for the pulse integration.

6.2.1 Digitization

A flash analog-to-digital converter (FADC) transforms an input signal from a continuous shape to discrete samples. First, the signal is evaluated at equally spaced discrete points and then the analog-to-digital conversion transforms the sampled values into corresponding digital output codes. These digital codes represent the quantization levels of the ADC. The difference between two adjacent levels is quoted as the *least significant bit* (LSB).

Unlike the successive approximation converter, that compares the input signal to only one reference voltage, the FADC architecture employs a set of $2^N - 1$ cascaded comparators to measure the analog signal to a resolution of N bits. A resistive divider with 2^N resistors supplies the reference voltage to each comparator, such that the reference voltage at the input of a comparator is one LSB greater than the reference voltage at the input of the previous comparator. Each comparator produces a logic "1", when the analog input voltage is higher than its input reference voltage, and a logic "0", when the analog input voltage is lower than the same reference voltage. The output of the comparators is collected by a digital encoder that determines, through a digital output code, the ADC channel in which the sample is binned.



(a) (b)

Figure 6.4: Example of a pulse shape, created by the Monte Carlo simulation, digitized in the high and low gain modes. The pulse amplitude is larger than the input range defi ned for the high gain mode (*a*) and the analysis is switched to the low gain mode (*b*).

The architecture of a 12-bit FADC employs 4095 comparators and 4096 resistors. The characteristics of such a 12-bit FADC were implemented in the simulations. A set of 4096 discrete levels, separated by 1 LSB, simulates the reference voltages at the input of comparators. Each sampled value is compared with the set of discrete values, and a variable *bit* is incremented each time the sample is larger than one of the references. At the end of the loop, the ADC channel, corresponding to the sampled amplitude value, is returned by the variable *bit*.

Apart from the FADC characteristics, the analysis takes the high and low gain channels, provided by the electronics of the Grande stations, into consideration. As mentioned in chapter 3, each Grande detector consists of 16 individual scintillators arranged in a 4×4 array. Each of these scintillators is viewed by a photomultiplier operating in high gain mode. The four central scintillators are viewed by 4 additional photomultipliers operating in low gain mode. The photomultiplier signals are multiplexed, separately for each gain mode, and the resulting high and low gain signals feed the two inputs of the KGEMD board. As mentioned in chapter 4, only the high gain signal can trigger the read-out of the FADC data. There is no trigger condition implemented in the simulations, because all pulses created with the Monte Carlo tool were assumed to fulfill the trigger condition. Nevertheless, they are first analyzed in high gain mode and if the amplitude of the pulses is larger than the input range of the high gain mode, the analysis is switched to the low gain mode.

It is necessary to define these two gain levels and the corresponding LSB values in the context of the present simulations. First of all, the amplitude of the pulse shapes is expressed in arbitrary units, since no calibration is being considered. Assuming an input range of 1V for the FADC, a maximum amplitude A^{LG}_{max}, in arbitrary units, was chosen to describe the upper limit of the input range, in the low gain mode. This amplitude is large enough to ensure that none of the pulses exceed it. Dividing A^{LG}_{max} into 4096 discrete intervals, 80 arbitrary units correspond to one LSB ($\approx 0.25\,$mV). The dynamic range achieved in the Grande stations, i.e. from 0.3 to 750 mips[1]/10 m^2 (high gain channel) and from 12 to 30,000 mips/10 m^2 (low gain channel) [Ber01], indicates a gain factor $f_G = 40$ between the two gain modes. Therefore, the LSB value, corresponding to the high gain mode, was considered to be 40 times lower, i.e. 2 arbitrary units. This automatically implies a 40 times lower input range A^{HG}_{max} for the high gain mode.

As previously mentioned, the pulses are first resolved in the high gain mode. If the amplitude

[1] units of minimum ionizing particle

of the high gain signal is greater than A_{max}^{HG}, the pulse is considered to be in *saturation*. This defines an *overflow* situation. In this case, an erroneous integration result is provided, since one or more amplitude values are underestimated. Therefore, the result is rejected and the pulse is analyzed in the low gain mode, with settings as described above.

The implementation has a technical character. In reality, the FADC system ensures the same maximum scale, i.e. 1V, for both gain modes. But, in the simulation it is much easier to apply different input ranges to each gain mode than to rescale the ideal pulses. Moreover, it is convenient, for the analysis of the results, to keep the pulses in their quasi-continuous form, as defined in the Monte Carlo algorithm. Figure 6.4 shows the example of a pulse from the *close to the shower core* class, which is digitized in both high and low gain modes. Due to the its high amplitude value, the pulse is saturated in the high gain mode (*a*) and the analysis is switched to the low gain mode (*b*). In the simulation code, the switch between the modes is applied before the pulse is digitized. Each sampled value is compared successively with the input limit A_{max}^{HG}. If any of the samples exceeds this limit, then the situation is flagged and the pulse is digitized in low gain mode by applying the previously defined factor $1/f_G$.

6.2.2 Quantization Errors

As described in the previous section, in the FADC architecture, the analog input voltage is simultaneously compared to each reference voltage at the input of comparators. This implies that, for a given sampled analog amplitude V_{in}, all the comparators with the input reference voltage smaller than V_{in} will output a logic ”1”, while the others will output a logic ”0”. The output of the comparators will look like a *thermometer code*: 00011111, for a 3-bit FADC example. The most important estimation of the analog signal is given by the comparator which produces the last logic ”1”, i.e. the 5th digit, from right to left, in the 3-bit FADC example. At this point, the difference between the analog signal and the reference voltage defines the *quantization error*. Any given sampled analog amplitude will be underestimated by this difference. The error will be different for each sample, but not larger than 1 LSB. Therefore, it can be said that the digitized sample is equivalent with the analog value plus the quantization error.

If no systematic effect interferes, then the quantization errors are uniformly distributed around + 0.5 LSB, with a standard deviation of $1/\sqrt{12}$ LSB (≈ 0.29 LSB). This means that the quantization errors are random and uncorrelated. In the present simulations, the quantization errors are estimated by subtracting the sampled value from the digital code, using the appropriate conversion.

6.2.3 Results of the Simulation

Applying the settings described in the subsection 6.2.1, each digitized pulse is integrated with the trapezium method. The integration result represents the *reconstructed* energy deposit. This value is compared with the *true* energy deposit, generated in the Monte Carlo simulation, by applying equation 6.8. The relative energy differences are filled into histograms. Figure 6.5 shows the digitization results for each class of pulse shapes.

The quantization of the pulses from *far from the shower core* class (see figure 6.5 *c*) yields the poorest results, in the order of 1.8%. There are two main effects that affect the quality of the energy reconstruction. First, the large values of the relative differences are the consequence of the digitization applied to pulses of low amplitudes. These pulses represent the contribution of a few or several particles that arrive with a considerable delay. Their amplitude is usually ranging between 25-60 ADC channels, which means that the system is viewing these pulses with only 1.5% of its full scale. Therefore, the quantization errors are large with respect to

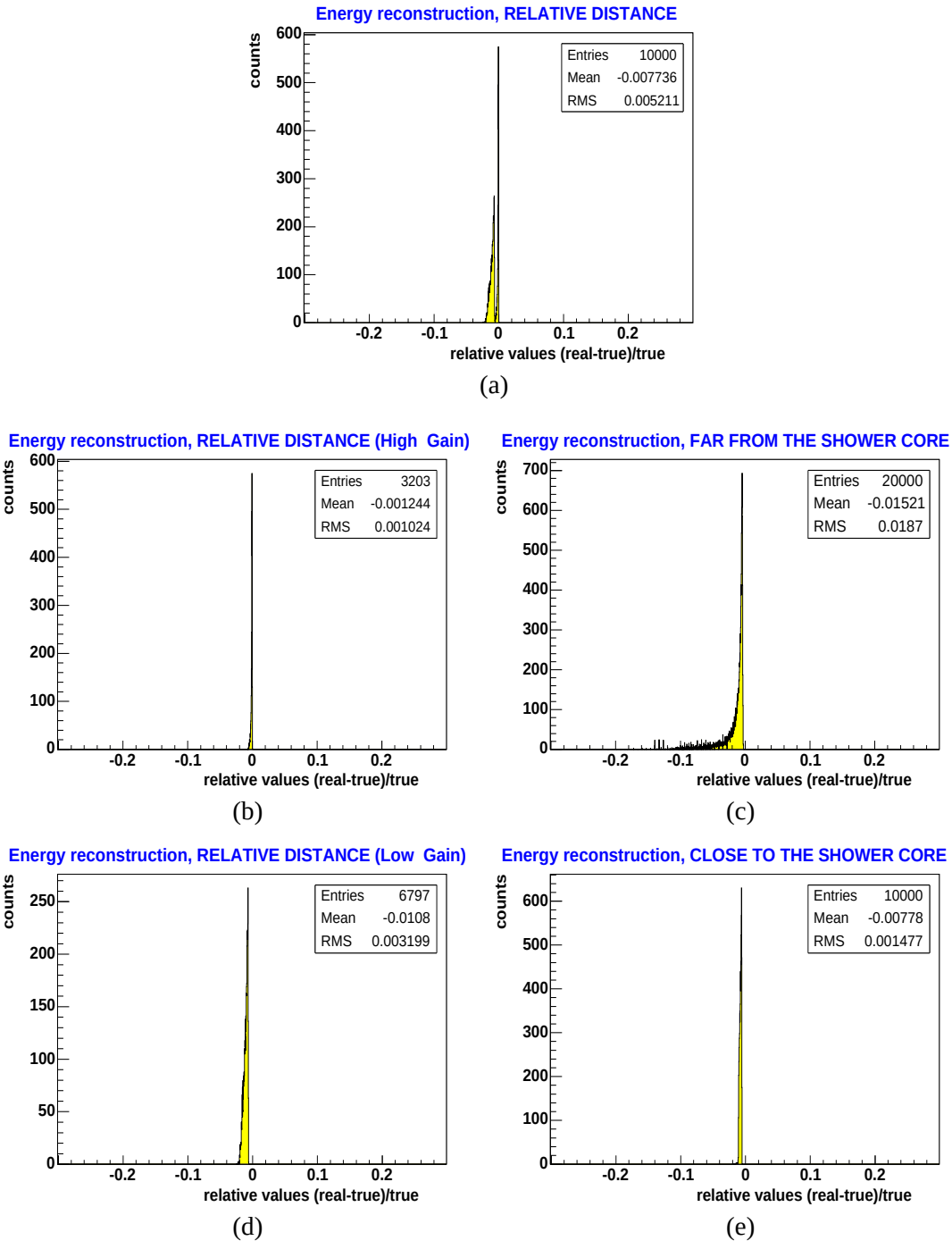


Figure 6.5: The quantization effect on the energy reconstruction. Each ideal pulse is digitized and integrated with the trapezium method. General results are given for all three classes (plots *a*, *c* and *e*). The graphs *b* and *d* depict the analysis for two gain modes of the pulse shapes specifi c to the *relative distance* class.

the pulse amplitude, and the energy deposit is largely underestimated. In fact, the pulses from this class are characterized by low amplitudes. The highest values were observed at about 1050 ADC channels, which represents only 25% of the 12-bit FADC full scale. The second effect, that lowers the quality of reconstruction, is induced by the quantization of the Landau-like tail of the pulses. This yields a systematic shift of the relative energy differences. About 12% of the total sampled amplitudes, most of them from the tail, are smaller than 1 LSB. Thus, for these samples a digital code 0 is determined, and the integration returns a *reconstructed* value smaller than the *true* value, by this underestimated quantity from the tail. The quantization errors are as well affected by the bias of the Landau-like tail (see appendix B, figure B.2 *a*, *b*). The form of the distribution suggests the interference of a systematic effect. The high peak in the plot B.2 *a* is due to the model developed in the Monte Carlo tool. The particles are randomly distributed in the 1μs interval, and their individual contributions are piled up. The resulting combined pulse form is strongly dependent on the random arrival time of the particles, such that the signal starts from the position indicated by the first arrived particle. This causes an amount of nil samples in the front of the signal (about 2% per total), which are responsible for the peak from plot B.2 *a*. These samples do not influence the integration result, but change the distribution of the quantization errors. If their contribution is excluded, then a much clearer picture of the systematic effect introduced by the tail is obtained (plot B.2 *b*). Most of the samples smaller than 1 LSB are localized in the tail (plot B.2 *c*) and the special form of the distribution (plot B.2 *b*) is mainly generated by the samples from the end of the signal (plot B.2 *d*). Plot B.2 *e* shows that the errors are dominant at small energy deposits, due to the reduced number of individual contributions.

The quantization of the pulse shapes from the *close to the shower core* class yields errors in the order of 0.15% (see figure 6.5 *e*). Due to their very high amplitudes with respect to A_{max}^{HG}, all these signals saturate the high gain channel (see the example in figure 6.4). Therefore, the entire analysis is performed in the low gain mode, by applying the factor f_G. In some cases, the high amplitudes range up to 96% of the low gain full scale. That ensures a good reconstruction of the energy deposit. But, the overall quality of the reconstruction is affected by the bias induced by the Landau-like tail of the pulses. The main characteristic of this class is the large number of particles arriving in a very short time window. This generates high pulse shapes with longer tails than the pulses from the *far from the shower core* class. Therefore, the number of samples taken from the tail is larger. The distribution of the quantization errors is modelled by this systematic effect (see figure B.1). Plots B.1 *c* and B.1 *d* prove that the dominant peak observed in the distribution of the quantization errors (plot B.1 *a*) is generated by the digitization of the sampled amplitudes from the tail, especially at the end of the signal. The samples from this region are responsible for the position and the shape of the peak that appears in plot B.1 *a*. Though most of the errors still arise at small energy deposits, a spread towards high energies is observed (plot B.1 *b*). This shows that tail increases with the number of individual contributions.

The quantization of the pulses from the *relative distance* class necessitates a special analysis. Due to the intermediary status of this class, the pulse amplitudes vary within the extremes defined by the other two classes. Therefore, about 32% of the pulses can be analyzed in the high gain mode (see figure 6.5 *b*), but for the rest, a low gain analysis is necessary (see figure 6.5 *d*). The combined effect of this two different analysis is depicted in figure 6.5 *a*. The energy reconstruction for the pulses quantized in the high gain mode provides the best results, in the order of 0.1%. This is explained by the high amplitudes of the pulses, many of them being close to the upper limit of the input range A_{max}^{HG}. As in the *close to the shower core* class, this ensures a good reconstruction of the energy deposit. Apart from this, the digitization shows practically no systematic influence, and the quantization errors are uniformly distributed (see figure B.3,

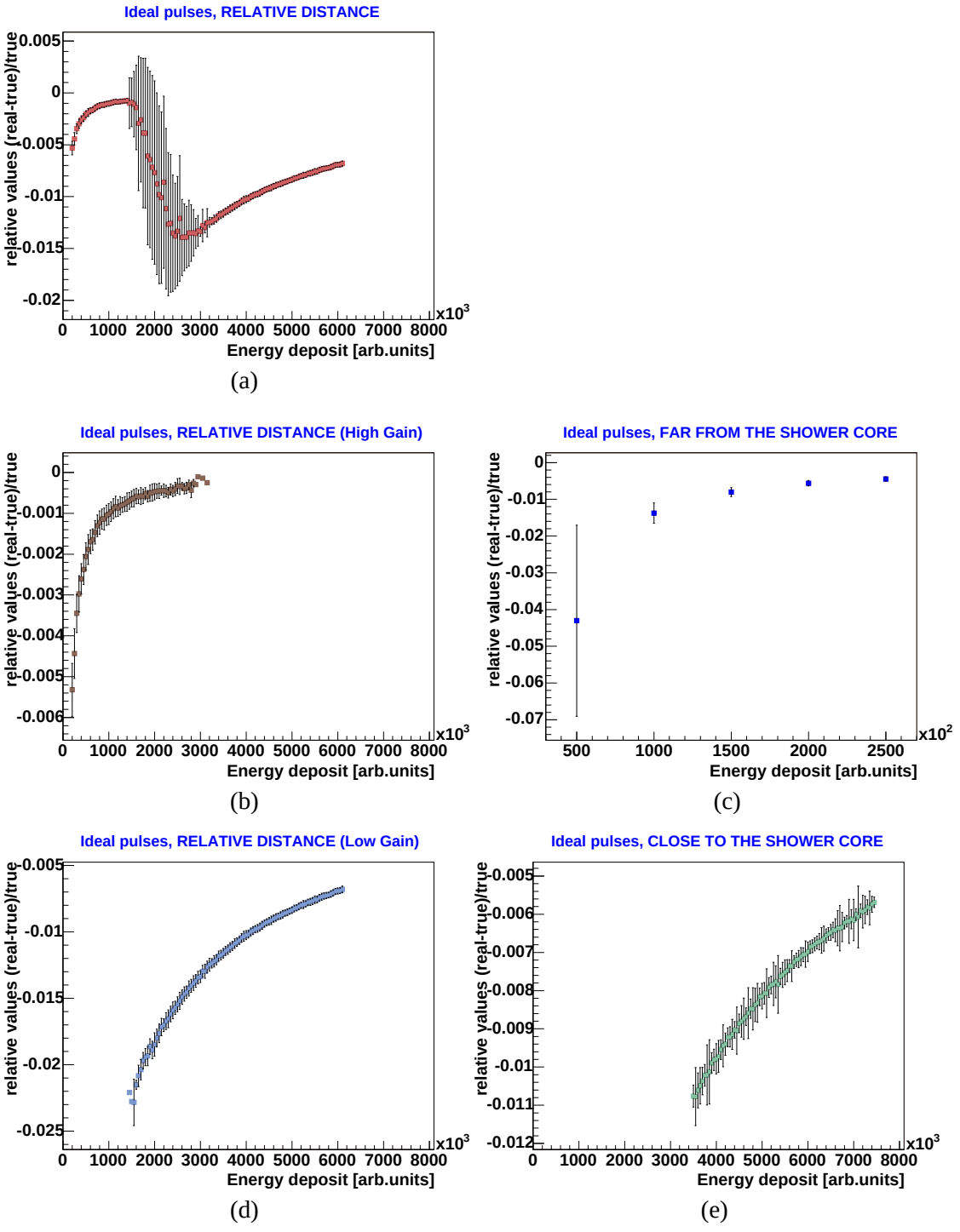


Figure 6.6: The digitization effect for a sliced energy deposit spectrum. The spectrum is divided in slices of energy of 50,000 arbitrary units. The relative values are fi lled into histograms. The mean and the width of each distribution are extracted and ploted as functions of the corresponding slice of energy.

a). The spike in the plot B.3 *c* shows that there is still a small amount of samples, i.e. 0.4%, of which value is zero, and their contribution is neglected in plot B.3 *a*. Anyway, the tail is still present and a large amount of the quantization errors arise from the sampled amplitudes in the tail (plot B.3 *d*).

The pulses quantized in the low gain mode provide poorer results, in the order of 0.3%. The larger deviations are due to pulses that exceed the high gain input range with only a few samples. This forces them to be digitized in the low gain mode, with a lower resolution in respect to the full scale. Apart from this, the systematic effect observed in the other two classes influences the quality of the reconstruction. The low gain analysis of the pulses from *relative distance* class presents similarities with the digitization of the signals from the *close to the shower core* class (see figure B.4, *a*). The pulses tend to a shape characterized by a long tail and its contribution generates the peak from plot B.4 *a*.

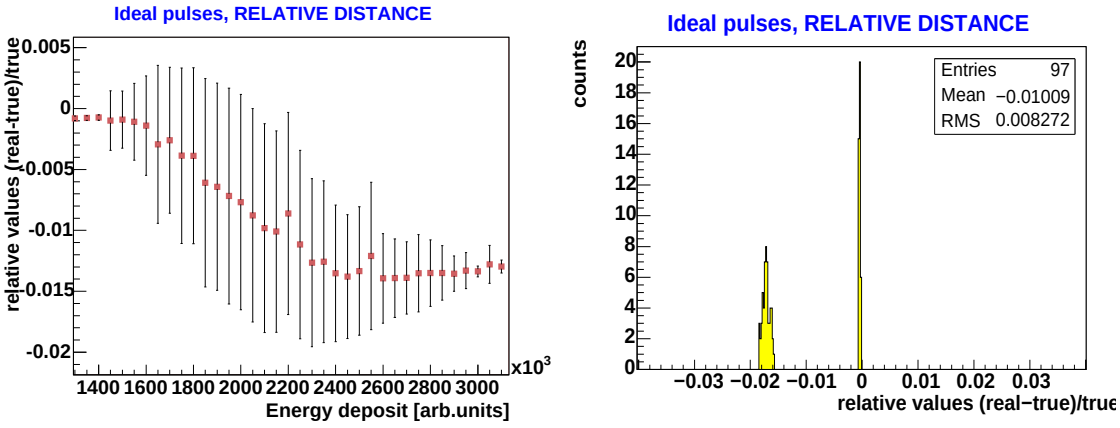
In order to understand the individual *signature* of each class better, a more detailed investigation was carried out. The energy spectrum was divided in slices of 50,000 arbitrary units, and for each *slice* a histogram was filled with the corresponding relative values. The resulting mean and width (RMS) values were extracted and plotted as a function of the energy slice (see figure 6.6). Plot 6.6 *c* emphasizes once again that the low amplitude pulses of the *far from the shower core* class display the largest systematic and statistical errors, in the order of 4% and 3%, respectively. The plots 6.6 *a*), 6.6 *b* and 6.6 *d* show that it is necessary to treat the *relative distance* pulses separately, due to their different digitization in the high and low gain modes. The energy overlap suggests that the amplitude of the pulses is not only dependent on the amount of energy deposit, but also on the arrival time of the particles. Pulses with high amplitudes, but describing lower energy deposits, saturate the high gain mode and force the analysis to the low gain mode. When the pulses are treated separately, the magnitude of the statistical errors is, in average, about 0.02% for the pulses digitized in the high gain, and 0.04% for those digitized in the low gain mode. The effect of the energy overlap and the larger systematic errors, provided by the latter, are responsible for the magnitude of the statistical errors, shown in the plot 6.6 *a*, i.e. of about 0.1%.

Figure 6.7 depicts the example of a *slice* ($2 \cdot 10^6$ arbitrary units) from the region where the energies overlap. The plot (*b*) shows the individual distributions of the relative values obtained in high (*right*) and low gain (*left*). The two distributions are generated by sets of completely different relative values. Therefore, when the distributions are separately considered, the statistical errors are small, about 0.01% for high gain, and 0.06% for low gain. The combination of the two sets of relative values into one common distribution, causes large statistical errors, about 0.08%, and a mean shifted accordingly.

This study shows, that the overall digitization effect, induced by the 12-bit FADC characteristics, on the reconstruction of the energy deposit, is smaller than 2%. The largest errors are caused by the pulses of which amplitudes are located at the low limit of the FADC resolution. The Landau-like tail induces a systematic shift, which affects the accuracy of the reconstruction and the distribution of the quantization errors. On the other side, the chosen input ranges for the high an low gain modes, i.e. A_{max}^{HG} and A_{max}^{LG}, bring their influence to the quality of the reconstruction, since they are arbitrarily chosen.

6.3 Noise Analysis

The term *noise* refers to a random, persistent and undesired disturbance, that accompanies or interferes with a certain signal. In electric devices, the noise is associated with small fluctuations in the voltage or in the current that flow through the circuit. These oscillations may be generated



(a) (b)

Figure 6.7: The overlaping energy region between the high and low gain analysis of the pulses from the *relative distance* class (plot *a*). The large statistical errors of the relative values, showed in plot *a*, are due to the individual distributions generated by each mode (plot *b*).

by an external source, e.g. electromagnetic coupling between an electric circuit and the *ac* power lines, or may be internally generated, as a result of the electrical charge transport. There are three main types of electrical noise: *thermal noise*, *shot noise* and *low frequency noise*.

Thermal noise results from the stochastic motion of the charge carriers in a conductor or semiconductor material. The electrons inside the material carry amounts of energy that vary with the temperature of the material. The random agitation of the electrons produces small fluctuations in energy, which are sufficient to cause significant noise potentials within the material [Mil72]. This phenomenon was first observed by J.B. Johnson in 1927, but a theoretical analysis was developed by H. Nyquist in 1928. He showed that the instantaneous voltage across any resistor R at a temperature T has a power spectral density:

$$\frac{dE_{th}^2}{df} = 4k_B T R \quad , \tag{6.9}$$

where E_{th} is the RMS value of the noise voltage, k_B is the Boltzmann's constant ($1.38 \cdot 10^{-23}$ J/K) [Mot93]. Thermal noise is often quoted as *Johnson noise* or *Nyquist noise*.

Shot noise is associated with the current flow across a potential barrier. Therefore it is present in transistors, diodes or vacuum tubes, but not in conductors. In transistors the shot noise is due to the random diffusion of the carriers through the *base-emitter* junction. Also, the random generation and recombination of the *hole-electron* pairs in a *p-n* junction results in small current variations, which are referred to as shot noise. In vacuum tubes, the shot noise arises from the random emission of electrons from the cathode [Mot93, Ott76].

The phenomenon was first theoretically investigated by W. Schottky in 1918, who called it *Schroteffekt*. He showed that spectral power density of the shot noise current fluctuations is:

$$\frac{dI_{sh}^2}{df} = 2q \cdot I \quad , \tag{6.10}$$

where I_{sh} is the RMS value of the noise current and q is the electron charge ($1.6 \cdot 10^{-19} C$) [Ott76].

Low frequency noise, also called $1/f$ noise or *flicker* noise, is present in all active devices. The spectral density of the noise voltage increases as the frequency decreases:

$$S_f(f) = \frac{E_f^2}{f^\alpha} \quad , \tag{6.11}$$

where E_f is the RMS value of the corresponding noise voltage and the exponent $\alpha \approx 1$. Thus, unlike the thermal and shot noises, the low frequency noise has spectral densities that are frequency dependent. In fact, if $\alpha = 0$, then equation 6.11 would describe a noise mechanism similar with the thermal and shot noises [Mot93].

In the following subsections, general characteristics of the above presented noise mechanisms are implemented in the simulations. Noise content is added to the ideal pulses and methods for the reconstruction of the energy deposit are investigated.

6.3.1 Time Jitter

For an analog-to-digital converter (ADC) it is theoretically assumed that the samples are taken at a regular time interval. In practice, different sources of errors produce small variations in the time accuracy and the result is an evaluation of the amplitude at the wrong instant. This instability of the time interval duration it is often quoted as *time jitter*. Although it is not really a noise, but rather a time-base error, the time jitter is mainly due to the noise which arises in the clock source of the ADC. When present in excess, the time jitter can degrade the performance of the conversion system.

Crystal oscillators are usually used as clock sources due to their high accuracy and frequency stability, and the low time jitter they feature. *Thermal noise* inside the oscillation circuit and the *phase modulation* of the oscillator frequency are the main contributors to the short-term instability in the clock signal. There are different ways to measure and express the time jitter. One of the most used terms is the *cycle-to-cycle* jitter (J_{cc}). It defines the time variation that occurs between two successive periods of a clock waveform:

$$J_{cc} = T_k - T_{k+1} \quad , \tag{6.12}$$

where T_k and T_{k+1} are two adjacent clock periods and k is the period index (see figure 6.8).

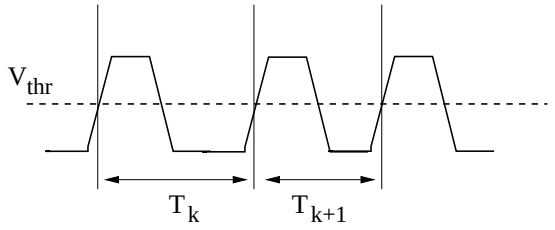


Figure 6.8: Schematic view of the cycle-to-cycle jitter. A threshold value V_{thr} defi nes a reference point on the rising edge of the clock for the time periods T_k and T_{k+1}.

Another important term in the jitter analysis is the *peak-to-peak* jitter (J_{pp}). It is defined as the difference between the lowest and highest deviation of the clock signal from its ideal position (see figure 6.9). Assuming a Gaussian distribution to describe the random jitter, the peak-to-peak value is used to define the limits that bound the jitter. The standard deviation of the Gaussian distribution defines another term, the *RMS* jitter.

In this study, it is assumed that the random jitter lies within ± 3.3 *RMS* jitter of a Gaussian distribution centered at zero. Therefore, the peak-to-peak jitter value is 6.6 *RMS* value. This means that the random jitter will be generated with a probability of 99.99% within the limits defined by J_{pp}. The values introduced in the simulation in order to generate the random jitter are given in table 6.2. The first row represents the technical specifications of the crystal oscillator implemented on the KGEMD board (see chapter 4). At this step, the J_{cc} jitter is considered, by restricting the difference between two consecutive random jitters to be larger than the indicated

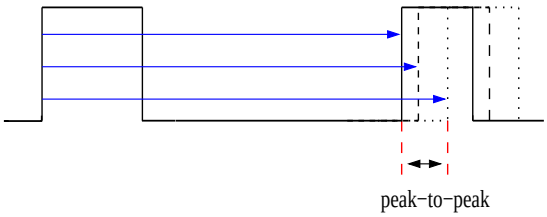


Figure 6.9: Schematic view of the peak-to-peak jitter. The difference between the lowest and the highest deviation of the clock signal, from its ideal position, defi nes the peak-to-peak jitter.

J_{cc} value. The rest of simulations use only the J_{pp} jitter, and hence, the *RMS* jitter. The time jitter analysis is performed in the following steps. For each J_{pp} value, a Gaussian distribution, of width *RMS* jitter, is defined. Then, at every 4 ns, a random jitter, generated from the Gaussian, is added to the time value. The pulse is evaluated at the jittered points and the resulting sampled amplitudes are digitized. Applying the trapezium integration and using equation 6.8, the relative energy differences are obtained and filled into histograms.

Peak-to-peak jitter *(nanoseconds)*	Cycle-to-cycle jitter *(nanoseconds)*
0.20	0.25
0.50	none
0.75	none
1.00	none
2.50	none

Table 6.2: Jitter values used in simulations .

Figure 6.10 depicts the results obtained for each class of pulses, by applying different jitter values. The mean value of each distribution is plotted as a function of the J_{pp}, and the width of the distribution is given as error bar. For a better understanding of the errors introduced by the time jitter, the results for $J_{pp} = 0$ (ideal case) are plotted as well in each graph. The conclusion is that the jitter contribution is, in general, negligible. The largest errors arise at the sampling and the digitization of the pulses from the *close to the shower core* class (plot *d*). The errors increase with the jitter value, from 0.14% in the ideal case, to 0.4% for $J_{pp} = 2.5\,\text{ns}$. The fact that the mean value stays constant, shows that the jitter affects the sampling of high amplitudes, rather than the sampling of the tail. The voltage in the tail decreases slowly and any sample taken at the wrong instant would not bring a significant change to the overall energy reconstruction.

Also, there are small jitter contributions to the digitization in high gain of the pulses from the relative distance class (plot 6.10 *a*. The errors increase very slow with the J_{pp} value, from 0.1% in the ideal case, to 0.14% for $J_{pp} = 2.5\,\text{ns}$. Similar to the previous case, the errors are generated when sampling the high amplitudes, rather than sampling the tail. This is proved by the results shown in the plots 6.10 *b* and *c*. The digitization of these pulses has shown results influenced by the low amplitude characteristic and the large systematic effect, induced by the sampled amplitudes from the tail (see section 6.2.3). Applying different values of time jitter does not change significantly the result of the reconstruction. The jitter mainly influences high pulses with large gradient, like those from the *close to the shower core* class, where the sampling is sensitive to any small time-base error.

The complete numerical results of the time jitter analysis are given in the tables B.4, B.5

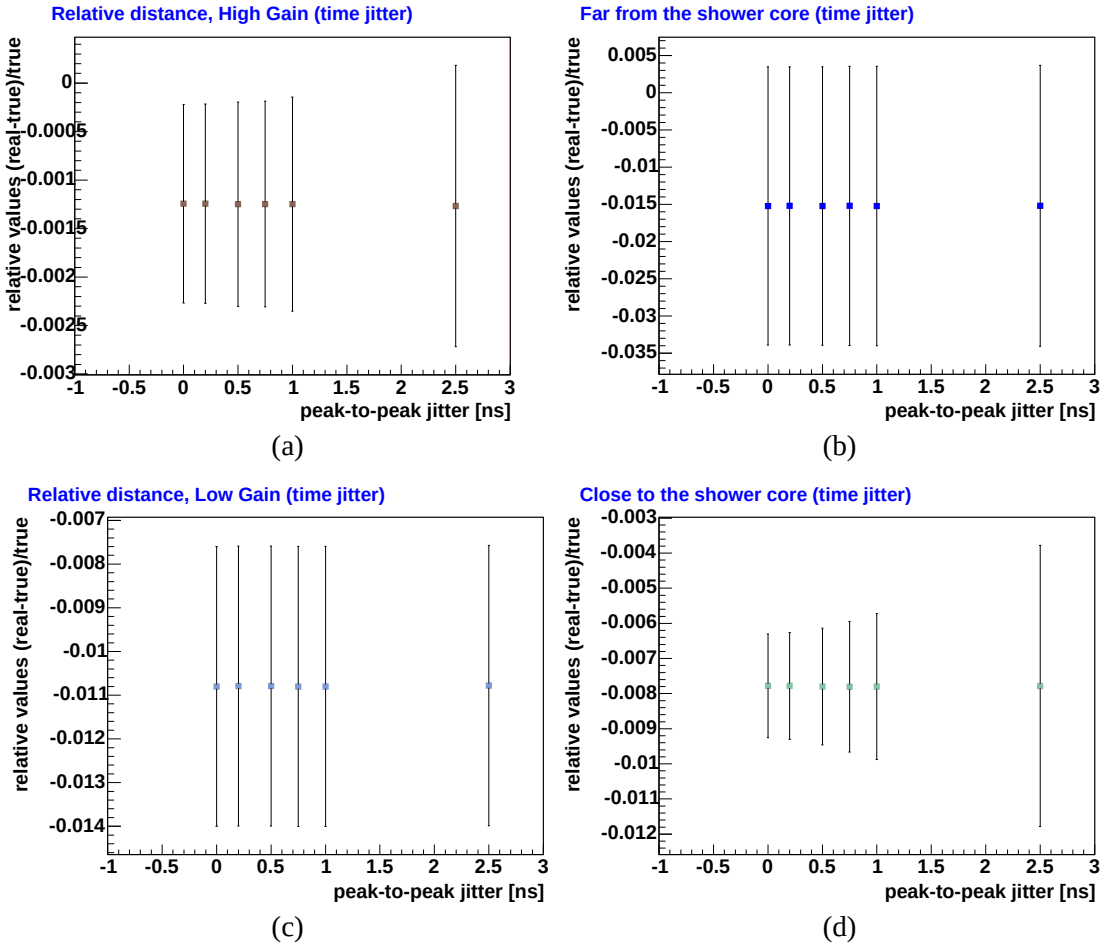


Figure 6.10: Jitter contribution to the energy reconstruction. The mean values of the distributions are plotted as a function of different peak-to-peak (J_{pp}) jitter values. The widths of the distributions are plotted as error bars.

and B.6 (see appendix B).

6.3.2 White Noise

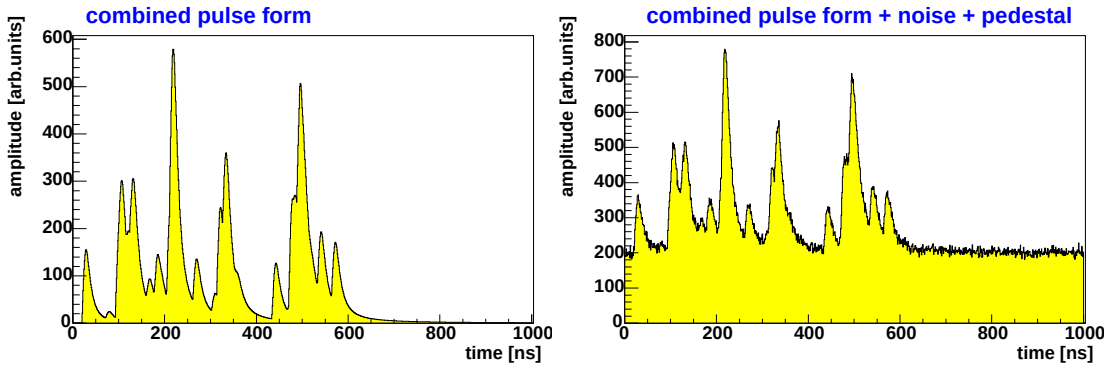
The *white noise* is the noise with equal power over the entire frequency spectrum. The thermal and the shot noises, described with equations 6.9 and 6.10, are sources of white noise. When measured in time, their amplitudes follow a Gaussian distribution.

In this section the white noise contribution to the reconstruction of energy is investigated. A Gaussian distribution, centered in zero, and with a certain width σ, is chosen to generate a random white noise. The following values, expressed in units of ADC channels, are chosen for the standard deviation (σ) of the Gaussian distribution:

0.15, 0.303, 1, 1.5, 2, 3, 4, 5, 10, 15.

To a good approximation, the noise is considered to be bounded within $\pm 3.3\sigma$. In order to avoid negative amplitudes, when the white noise is added to the ideal pulses, the baseline of the signals is arbitrarily shifted by 100 ADC channels (constant pedestal). This contribution has later to be estimated and subtracted from the integration result.

Another important implementation is related to the noise amplitude in the high and low



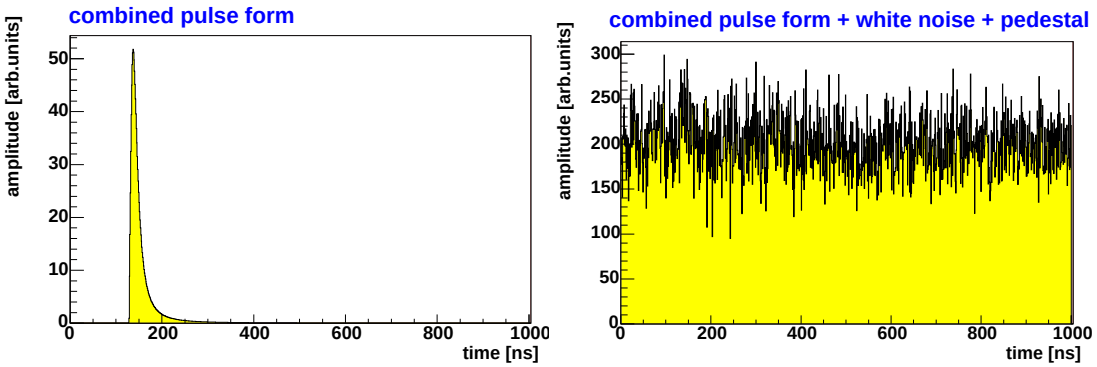
(a) (b)

Figure 6.11: Example of an ideal pulse which is smeared in amplitude by the white noise contribution. The pulse first is shifted in amplitude by a constant pedestal, and then smeared by the Gaussian noise. The pedestal has 100 ADC channels, while 2 arbitrary units correspond to 1 LSB (high gain analysis). In these example, the width of the noise is 4 ADC channels.

gain mode. In these simulations, it is assumed that the noise contribution is equal in both gain channels, i.e. the main amount of noise appears after the photomultiplier signals are mixed. This means that the shape of a pulse, resulted from the addition of noise to an ideal pulse in high gain mode, will not change when the analysis is switched to the low gain. As mentioned in section 6.2.1, the analysis in the low gain mode implies the change of the input range, from A_{max}^{HG} to A_{max}^{LG}, and not the rescaling of the pulse. Thus, the sampled amplitudes keep their values, expressed in arbitrary units, but the corresponding ADC channel will differ accordingly. That is to say, the *signal to noise* ratio is constant in both gains.

At every 4 ns, a random white noise value is generated and added to the sampled ideal amplitude, together with the constant pedestal. Figure 6.11 shows the example of an ideal pulse (plot *a*), from the *far from the shower core* class, which is shifted in amplitude by the constant pedestal and by the white noise contribution (plot *b*). In order to reconstruct the energy deposit of such a pulse, the contribution of the pedestal has to be estimated, and then subtracted from the integration result. For these simulations, a correction method that estimates the spread of the pedestal, both at the beginning and at the end of the signal, has been chosen. The first and the last four digitized samples are averaged separately, and the resulting pedestal levels are connected with a first order polynomial line. The 1μs time window, the polynomial line and the two pedestal levels form an area, the integral of which is subtracted from the integration result.

Since the pulses generated by the Monte Carlo tool are not a direct result of any threshold condition, the signals have a random starting position with respect to the 0 reference of the 1μs digitization interval. Therefore, all or part of the first four samples of the signal can be equal to zero, i.e the signal starts later, or can be positioned on the leading edge of the pulse. In the latter case, the averaged value of the four samples deviates significantly from the baseline such, that it introduces large errors, when the pedestal correction is applied. Due to the internal delay in the FADC system, a certain amount of samples, before the threshold to be crossed, are available for read-out, and hence, for analysis (see figure 4.1). This amount of samples is not consistently ensured by the Monte Carlo tool, but rather randomly. In order to eliminate this drawback, four leading zero samples are prepended to each ideal signal. Like the other 251 samples, they are weighted with the constant pedestal (100 ADC channels) and the random noise. It should be emphasized, that these leading samples are not used as input for the pulse integration, but only to determine the spread of the pedestal at the beginning of the signal.



(a) (b)

Figure 6.12: White noise effect on the low amplitude pulses. The height of the pulse is about 26 ADC channels. The RMS of the noise is 15 channels, but the noise amplitudes can be larger than 45 ADC channels. The pulse is *drowned* in the noise and the energy deposit is reconstructed with huge errors.

Integrating the pulse with the 251 digitized values and subtracting the estimated pedestal contribution, a *reconstructed* energy deposit is obtained and compared with the corresponding *true* energy deposit, by applying equation 6.8. The relative energy differences are filled into histograms.

Figure 6.13 shows the results obtained for each class of pulses, by applying different white noise values. The mean of each distribution is plotted as a function of the chosen RMS values of the white noise, and the width of each distribution is given as error bar. In order to understand the errors introduced by the white noise better, the ideal case (white noise RMS = 0) is plotted as well. As expected, in all cases, the errors increase with the level of noise. The largest effect is shown by the pulses from the *far from the shower core* class (plot *b*), i.e. about 50% error for an RMS of the white noise of 15 ADC channels. These huge errors are caused by the low amplitude pulses which are effectively *drowned* in the noise. Figure 6.12 shows the example of such a pulse, covered completely by the level of noise. In the ideal case, the pulse has an amplitude of about 26 ADC channels (for the high gain analysis, it is considered that 2 arbitrary units correspond to 1 ADC channel). The RMS of the white noise is, in this example, 15 ADC channels, but the noise can have amplitudes larger than 45 ADC channels. The energy deposit is reconstructed with very large errors, since only the noise is sampled and integrated.

The reconstruction of the pulses from the other two classes shows large errors as well, but much smaller when compared with the results of the *far from the shower core* class. For a RMS value of the white noise of 15 ADC channels, the errors range from 2%, in the case of the *relative distance* pulses analyzed in high gain mode (plot 6.13 *a*), to about 20%, in the case of the *relative distance* pulses analyzed in low gain mode (plot 6.13 *c*). The former shows a deviation of the mean from the ideal case of about 3%, which suggests that the reconstruction is influenced by a systematic effect. The reconstruction of the pulses from the *close to the shower core* class provides medium errors, up to 6%.

In order to understand these effects better, an analysis, similar to that introduced in section 6.2.3, was carried out. The spectrum of energy deposits, generated with the Monte Carlo tool, was divided into slices of 50,000 arbitrary units. For each *slice*, a histogram was filled with the corresponding relative values. The resulting mean and width values are extracted and plotted as a function of the corresponding *slice* of energy. The set of figures B.5 ÷ B.9 (see appendix B, section B.4) depicts the results of this analysis, carried out for different RMS values of white

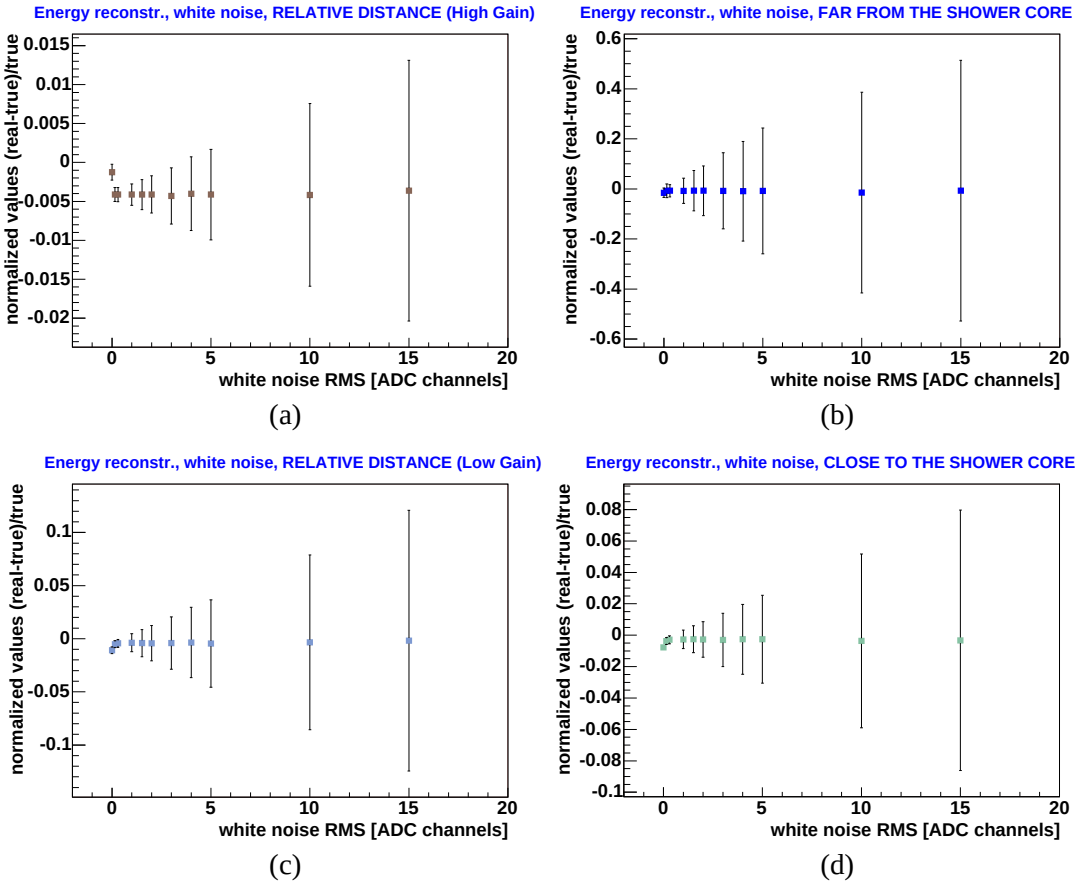


Figure 6.13: White noise contribution to the energy reconstruction. The mean values of the distributions are plotted as a function of the different RMS values of the white noise. The widths of the distributions are plotted as error bars.

noise.

The largest errors are provided by the reconstruction of the low amplitude pulses from the *far from the shower core* class (plots *b* and *f*, in each figure). As shown in figure 6.12, these pulses are completely covered by noise and the reconstruction provides very large errors, i.e. more than 100% for the RMS of the noise of 15 ADC channels. If this contribution is cut away, then the reconstruction provides much smaller errors, between 0.4% and 14%, from the lowest to the highest RMS value of the noise. The horizontal alignment of the mean values shows that the noise attenuates the Landau-like tail effect (see section 6.2.3). In fact, the white noise produces a *dithering* effect, by improving the digitization of the small and slowly varying tail signal. This means that the noises causes the digitized signal to alternate between adjacent quantization levels, such that the digitization result provides more information about the tail signal. It can be noticed that even a small amount of noise, e.g. RMS = 0.15 ADC channel (or LSB, as shown in the plots), can cause such an effect. Therefore, the quantization errors are now uniformly distributed. Also, the applied pedestal correction brings a significant contribution to the observed horizontal alignment. When the RMS value of the white noise becomes very large, the tail is *drowned* in the noise and the main contribution to the final result is brought by the pedestal correction method. In all cases, the mean values show a deviation smaller than 1%.

The reconstruction results provided by the other two pulse classes are modelled by the same effects, the digitization of the low amplitude pulses, the *dithering* effect and the pedestal cor-

rection. The reconstruction of the pulses from the *close to the shower core* (plots *d* and *h*, in each figure) shows such a characteristic. The poorest results arise from the digitization of the low amplitude pulses. In these cases, the errors range from 0.3% (RMS = 0.15 LSB) to 12% (RMS = 0.15 LSB), but the highest pulses are reconstructed with errors up to 4%. The *dithering* effect is slower, since the Landau-like tail of these pulses is longer and higher than the tail of the pulses from the *far from the shower core* class.

A special effect is provided by the analysis in high gain mode of the pulses from the *relative distance* class. At high energy deposits, the mean values show a displacement from the horizontal alignment, especially at low noise contributions. This *break* occurs in the region where the spectra of energy deposits of the pulses analyzed in high and low gain modes overlap. Part of the pulses from this region are characterized by a large number of particles, distributed in a large time extent, and by high amplitudes, at the limit of A_{max}^{HG}. The spread induces a significant tail signal, which biases the pedestal correction. The level detected at the end of the signal is significantly higher than the level detected by the four leading digitized samples. Therefore, the deviation showed in the plot, represents the error introduced by this effect and the correction method. Again, for low RMS values, the white noise does not produce a *dithering* effect on high tail signal. When the noise level is increased, an attenuation of the effect is observed, because the amplitudes of these pulses are *pushed* beyond the input range A_{max}^{HG}. Thus, they are analyzed in the low gain mode. For the rest of the pulses from this region, the attenuation is produced by the *dithering* effect and the applied pedestal correction.

This study shows that the white noise contribution to the reconstruction of the energy deposit is more than significant. The magnitude of the reconstruction errors increases with the RMS value of the white noise. Also, the signal to noise ratio is determinant for the quality of the reconstruction. The low amplitude pulses analyzed in the high gain mode are distorted or even completely covered by the noise content, and therefore, the reconstruction fails. The analysis of the higher amplitude pulses, in the same gain mode, provides the smallest errors, i.e. up to 2% for a RMS value of the noise of 15 ADC channels. As for the high gain mode, the reconstruction performed in the low gain mode provides largest errors in the analysis of the low amplitude pulses, i.e. in the order of $20-30\%$, for the same maximum simulated RMS value of the noise. This result brings an important conclusion. Due to the large noise content, high amplitude pulses, which normally provide a good reconstruction result in the high gain mode, are *pushed* in saturation, and their analysis in the low gain mode returns large reconstruction errors.

Also, it was observed that the white noise attenuates the systematic effect introduced by the Landau-like tail of the pulses. A *dithering* effect occurs, by alternating the tail signal between adjacent quantization levels, and thus, providing more information about the tail content. At large RMS values of the white noise, the pedestal correction brings the main contribution to the final result, since the tail signal is *drowned* in the noise.

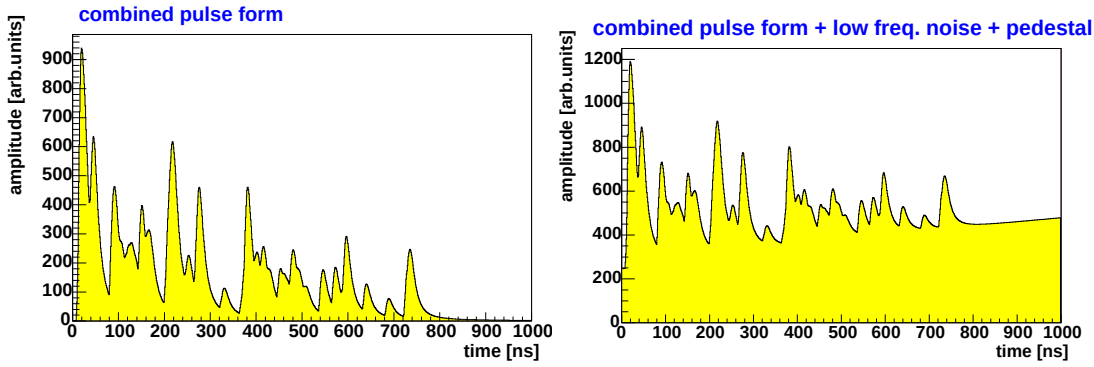
6.3.3 Low Frequency Noise

As previously mentioned, the noise power has spectral densities that vary as power of the inverse frequencies, i.e. $1/f^{\alpha}$. When the exponent $\alpha = 1$, the noise is referred as to *low frequency noise*. Since it is present in all active devices, an analysis of its contribution to the reconstruction of the energy deposit was carried out.

In these simulations, the following parameterization has been chosen to describe the low frequency noise:

$$F(t) = A \cdot sin(2\pi\nu t) \quad , \tag{6.13}$$

where A and ν represent the amplitude and respectively, the frequency of the noise. Two fre-



(a) (b)

Figure 6.14: The low frequency noise contribution. The ideal pulse (*a*) is shifted in amplitude by the constant pedestal (100 ADC channels) and by the low noise contribution (*b*). In this example, the noise frequency is 100 kHz, while the amplitude has 50 mV.

quency values are considered in this study: 15 kHz and 100 kHz. Apart from these values, the contribution of the noise generated by an electrical source, i.e. 50 Hz, was investigated. The conclusion was that the effect is negligible, since for 1 V amplitude of the noise, the maximum contribution in 1 μs is about 0.3 mV. This is slightly larger than the corresponding voltage value in 1 ADC channel (≈ 0.25 mV), and thus, the reconstruction is not affected by this type of noise.

The amplitude of the low frequency noise is simulated according to the following values: 5 mV, 10 mV, 20 mV, 30 mV, 50 mV, 100 mV, 150 mV, 200 mV. The voltages are converted in arbitrary units, by considering that, the 1 V input range for the FADC corresponds to the maximum amplitude A^{HG}_{max} (or A^{LG}_{max}) expressed in arbitrary units.

In order to ensure different low frequency noise contributions to each pulse, a time variable T is introduced in equation 6.13, such that:

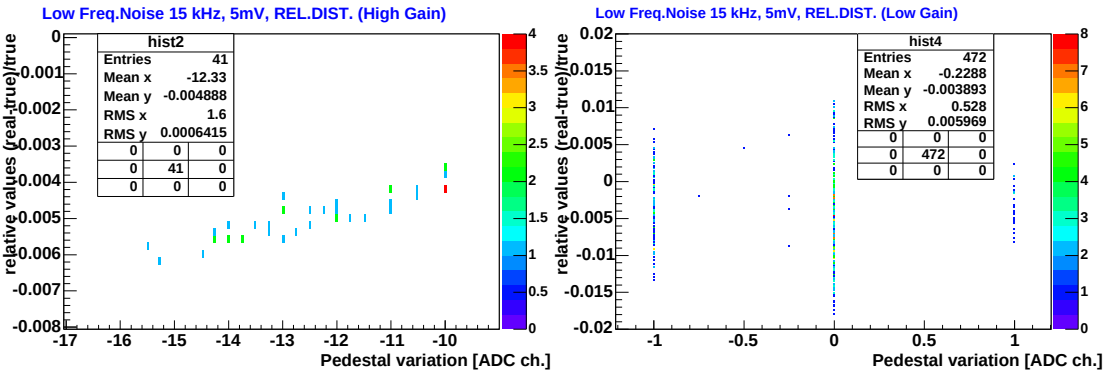
$$F(t+T) = A \cdot sin[2\pi\nu(t+T)] \quad . \tag{6.14}$$

The variable T is randomly generated from an uniform distribution such, that it takes different values at each event. This means, that the pulses will have different positions, with respect to the phase of the sine functions, and hence, the noise contribution will differ. The interval of the uniform distribution is chosen to be $[0; 2/\nu]$.

Like in the white noise analysis, at every 4 ns, the low frequency noise contribution and the constant pedestal (100 ADC channels) are added to the sampled ideal amplitude. Figure 6.14 shows the example of an ideal pulse (plot *a*), from the *far from the shower core* class, which is shifted in amplitude by the constant pedestal and the low frequency noise contribution. In order to apply pedestal corrections, four leading samples are prepended to each pulse. Then, the first and the last four samples are averaged separately, and the resulting pedestal levels are connected with a first order polynomial line. The two pedestal levels, the 1μs time window and the polynomial line form an area, the integral of which is subtracted from the integration result.

Apart from the white noise analysis, in these simulations the pedestal value is not always sufficient to avoid negative amplitudes. The noise contribution is, in some cases, large enough to determine negative sampled amplitudes. This defines an *underflow* situation. As in the white noise analysis, the noise contribution is considered equal for both high and low gain channels. Therefore, the *underflow* will exclude the pulse from any further analysis, since the *signal to noise* ratio will have the same value in the low gain mode.

Integrating the pulse with the 251 digitized values and subtracting the estimated pedestal



(a) High Gain (b) Low Gain

Figure 6.15: Landau-like tail contribution to the energy reconstruction, in the case of the pulses from the *relative distance* class. The pedestal variation is defined as the difference between the estimated level of noise at the beginning of the signal and that estimated at the end of the signal. The low frequency noise contribution is smaller than 1 ADC channel.

contribution, a *reconstructed* energy deposit is obtained. The relative values are computed with equation 6.8 and filled into histograms. Figure 6.16 depicts the reconstruction results obtained for each class of pulses, by applying 15 kHz (plots *a*, *b*, *c*, *d*) and 100 kHz (plots *a*, *b*, *c*, *d*) low frequency noises, with different amplitudes. The resulting mean values are plotted as a function of the noise amplitudes, and the width values are given as error bars. The ideal case (noise amplitude zero) is plotted as well.

In the case of 15 kHz low frequency noise, the reconstruction in high gain mode of the pulses from *relative distance* class (plot *a*), shows a distinct result, when compared with the other cases. The mean values of the relative energy differences deviate by about 0.3% from the ideal case, showing a consistent tendency to underestimate the *true* values. In the other cases (plots *b*, *c* and *d*), the mean values deviate from the ideal case towards the *true* values. In order to understand these effects better, a similar analysis to those described in sections 6.2.3 and 6.3.2, was carried out. The spectrum of energy deposits was divided into slice of 50,000 arbitrary units, and, for each *slice*, a histogram was filled with the corresponding relative values. The resulting mean and width values were extracted and plotted as a function of the corresponding *slice* of energy. The case of 0 mV (ideal case) is plotted as well. The results are shown in the set of figures B.10 ÷ B.13.

The reconstruction results, in the high gain mode, of the pulses from the *relative distance* class (plots *a* and *e*, in the mentioned set of figures) are influenced, like in the white noise analysis, by the low amplitude pulses and the significant tail signal, that characterizes the high amplitude pulses. The latter effect is depicted in the example from figure 6.15. The pedestal variation is plotted against the relative values, for the pulses from the *relative distance* class analyzed in high gain (plot *a*) and low gain (plot *b*) modes, in the case of 15 kHz low frequency noise, with 5 mV amplitude. In these examples, the pedestal variation is defined as the difference between the estimated level of noise at the beginning of the signal, and the level estimated at the end of the signal. These differences are larger in the high gain mode than in the low gain mode. Since the contribution during 1 μs of the 15 kHz low frequency noise, with 5 mV amplitude, is less than 1 ADC channel, in both gain modes, the differences between the estimated levels of noise are due to the tail in the end of the signal and the applied pedestal correction. This picture is a proof of the previous statement, that the tail signal in the high gain is quite significant, and influences the reconstruction results, in the sense that the pedestal correction will always cut, in

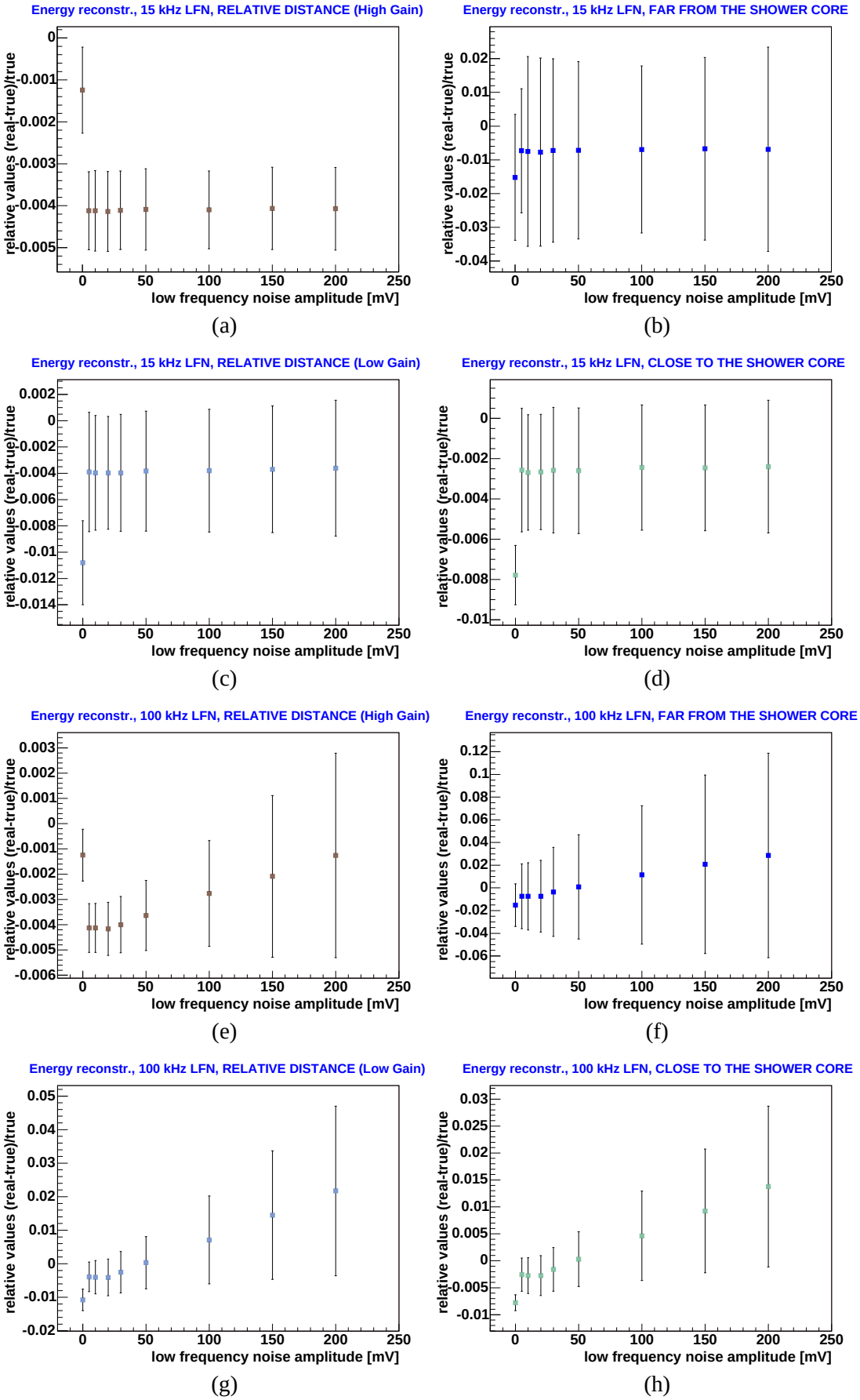


Figure 6.16: The 15 kHz (plots *a*, *b*, *c*, *d*) and 100 kHz (plots *e*, *f*, *g*, *h*) low frequency noises contribution to the energy reconstruction. The mean of the relative values is plotted as a function of the noise amplitudes, and the width of the relative values is given as error bars.

such cases, a significant part of the signal. Therefore, the reconstruction will underestimate the *true* energy deposit. The largest deviations from the *true* values, i.e. in the order of 0.6%, occur for large differences between the estimated levels of noise, and are reduced with the lowering of the same differences. Once the amplitude of the low frequency noise is increased, part of the pulses exceed the input range A_{max}^{HG}, and are analyzed in the low gain mode, or they are excluded from any analysis, when the amplitude of the noise pulls the signal amplitudes towards negative values. Nevertheless, the reconstruction, in the high gain mode, of the pulses from the *relative distance* class provides the lowest reconstruction errors, up to 0.6%.

In the low gain mode (plots *c* and *g* in the set of figures B.10 ÷ B.13), the tail contribution is much smaller, due to the fact that the input range A_{max}^{LG} is 40 times larger, and thus, the corresponding LSB value is larger by the same amount. The noise has a constant contribution in both gain modes and this is almost linear. This is due to the fact that the period of the 15 kHz low frequency noise is about 66 μs, much larger than the 1 μs digitization period. Since the tail contribution in the low gain analysis is small, and this tail usually represents a long signal, the pedestal correction will mostly overestimate the *true* energy deposits. This explains the better relative values achieved in the low gain analysis, for 15 kHz low frequency noise, than those achieved in the ideal case. This is as well a characteristic of the pulses from the *far from the shower core* (plots *b* and *f* in the set of figures B.10 ÷ B.13) and *close to the shower core* classes (plots *d* and *h* in the same set of figures), of which pulses are characterized by small and long tail signal. The magnitude of the errors is dominated by the poor reconstruction of the low amplitude pulses from each class, i.e. about 6% for the *far from the shower core* class, and, respectively, 0.8% for *close to the shower core* class.

The analysis performed for the 100 kHz low frequency noise shows the same characteristic seen in the previous analysis (see figure 6.16, plots *e*, *f*, *g*, *h*). At low noise amplitudes, the results of the analysis, in the low gain mode, of the pulses from the *relative distance* class (plot *e*), show a deviation of the mean values, which is distinct from the results provided by the other classes. But, as the amplitude of the low frequency noise increases, the mean values are shifted towards the *true* values. This effect models the results of all categories of pulse shapes. On the other side, the contribution, during 1 μs digitization period, of the 15 kHz low frequency noise with 200 mV amplitude, is equal to the contribution of the 100 kHz low frequency noise with 30 mV amplitude, i.e. 18.84 mV, which corresponds to about 77 ADC channels. This explains why the results are almost similar, up to this noise level. Above 30 mV amplitude values, the contribution of the 100 kHz low frequency noise determines the shift of the mean values, as shown in corresponding plots of figure 6.16.

In order to understand these effects, the same analysis, in which the energy spectrum is divided in slices of energy, is performed. The set of figures B.14 ÷ B.17 depicts the results of the analysis. Above 30 mV amplitude, as it was expected from the experience of the previous analyses, the deviation of the mean values, from the plots of the figure 6.16, is mainly due to the digitization of the low pulse amplitude. The largest errors are provided by the analysis of pulses from the *far from the shower core* class, i.e. 60%, while the lowest are provided by the analysis, in high gain mode, of the pulses from the *relative distance* class, i.e. in the order of 1%. But, these large errors mask two effects that can be observed in the plots of the new analysis. First, at higher energies, the mean values have a tendency to decrease. This shows that the contribution of the 100 kHz low frequency noise induces an underestimation of the respective pulses. On the other hand, all the mean values are shifted towards higher values, as the amplitude of the low frequency noise increases.

First of all, the period of the 100 kHz low frequency noise is 10 μs. This means that the *sine* characteristic is more emphasized in the digitization interval of 1 μs, and the noise contribution

is more significant, when compared with the characteristics of the 15 kHz low frequency noise. As a first effect, the negative amplitudes of the noise will pull many pulses into *underflow*, and, therefore, exclude them from any analysis. The positive amplitudes of the noise and the emphasized sine characteristic contribute to the reconstruction of the energy deposit such, that the pedestal correction mainly overestimates the *true* values. The contribution increases with the noise amplitude, and this explains why all mean values shift to higher values when the noise amplitude increases. At 200 mV noise amplitude, most of the reconstructed values are positive, while at 5 mV amplitude only a few were showing such a characteristic. On the other side, the Landau-like tail of the pulses contributes to the lower reconstruction results at higher energies. The tail contribution induces a higher estimated noise level at the end of the signal such, that the pedestal correction reduces the noise contribution. Since the tail signal increases, for all categories of pulse shapes, with the energy deposit, the mean values drop towards higher energy deposits.

This study shows how different noise contributions influence the reconstruction of the energy deposit. In the first case, 15 kHz low frequency noise, the noise contribution is almost linear and the largest errors are in the order of 2%. In the second case, 100 kHz low frequency noise, the noise contribution approaches a sine characteristic and the largest errors are in the order of 60%. For the latter case, the first order polynomial correction shows limitations, since the applied noise is nonlinear.

Chapter 7

Conclusions

In this thesis, reconstruction algorithms for the FADC DAQ system of the KASCADE-Grande experiment have been investigated. A Monte Carlo tool was developed in order to ensure a large collection of pulse shapes. The pulses were simulated as response of one Grande station, and they were organized in three classes, according to the distance from the shower core: *far from the shower core, at relative distance* and *close to the shower core*. Random energy deposits were created for each event, and fractions of these deposits were assigned to individual pulses. These pulses were scaled and piled up such, that the area below each combined shape equals a corresponding *true* energy deposit. The Monte Carlo turned out to be a useful tool to study many questions related to the pulse shape reconstruction. The resulting collection of pulse shapes was used as input for several studies.

In order to retrieve the energy deposit, the pulse shapes were integrated such, that the result is a quantity in arbitrary units of amplitude multiplied by time. This result defines the *reconstructed* energy deposit, which is compared with the corresponding *true* energy deposit. The precision and accuracy of the reconstruction and the time consumption of different integration methods, for equally spaced samples, were investigated. The trapezium method was chosen due to a consistent behavior and little CPU time consumption.

Implementing the 12-bit FADC characteristics, the digitization and quantization errors have been studied. The ideal pulses, simulated with the Monte Carlo tool, were sampled every 4 ns and digitized in the high gain mode. Whenever the amplitude of the pulses was larger than the input range of the high gain mode, the analysis was switched to the low gain mode. The resulting digitized samples were used as input for the trapezium integration. The study showed, that the energy deposit can be reconstructed with errors smaller than 2%. The largest errors were due to the digitization of the low amplitude pulses, which are located at the low limit of the FADC resolution. Another effect, that influences the reconstruction result, is the Landau-like tail of the pulses. The tail signal decreases slowly, and the digitization underestimates the energy contained in the tail. This effect models the distribution of the quantization errors, which normally should be distributed uniformly.

Time base-errors and amplitude fluctuations were implemented in the analysis, in order to simulate the individual characteristics of the electronics in each Grande station. The short-term instability in the clock signal (jitter) was investigated, by sampling the pulses at a wrong instant, and digitizing the resulting sampled amplitudes. The *reconstructed* energy deposits were obtained by integrating the digitized pulses with the trapezium method, and the results have been compared with the corresponding *true* energy deposits. This analysis showed, that the jitter contribution to the reconstruction of the energy deposit is negligible. The largest contribution is given by the analysis of the high amplitude pulses. These signals are characterized by a large

gradient, which makes the sampling sensitive to any small time-base error. The Landau-like tail brings no effect, since this signal decreases slowly.

The white noise contribution to the reconstruction of energy deposit was investigated, by applying different noise content to the ideal pulses. The white noise was generated from a Gaussian distribution, and different RMS values of the noise, i.e. from 0.15 ADC channels up to 15 ADC channels, were used to define the white noise content. Additionally, in order to avoid the negative amplitude values, the baseline of the pulses was arbitrarily shifted by 100 ADC channels. The resulting amplitudes were sampled and digitized, and trapezium integration was applied. In order to reconstruct the *true* energy deposit, the level of noise is estimated and subtracted from the integration result, by applying a first order polynomial line. The results showed that the magnitude of the reconstruction errors increases with the amplitude of the applied white noise. The largest errors were provided by the low amplitude pulses, which showed to be distorted or completely covered by the white noise content. Therefore, the reconstruction fails. The smallest error were provided by the analysis of the higher amplitude pulses in the high gain mode, i.e. up to 2% for an RMS of the white noise of 15 ADC channels. The largest errors obtained in the low gain analysis, i.e. $20-30\%$, were due to low amplitude pulses. This showed that, due to the large noise content, pulses which normally provide good results in the high gain analysis, exceed the input range of the high gain and are analyzed with large errors in the low gain mode. Also, a *dithering* effect was observed for small RMS values of the white noise. The noise forced the amplitude of the tail to alternate between adjacent quantization levels such, that the information from the tail can be reconstructed.

In the last analysis, low frequency noise was applied to the ideal pulses by means of a parameterization, based on a sine function. Two frequency values were considered, 15 kHz and 100 kHz, with different amplitudes between 5 mV and 200 mV. Due to the chosen frequency values, and hence, the resulting periods, the two contributions are different. The 15 kHz frequency noise brings a linear contribution, while the 100 kHz low frequency noise shows a characteristic that approaches the underlying sine function. These contributions were applied to the ideal pulses, and the baseline of the pulses was shifted by 100 ADC channels. In order to reconstruct the energy deposit, the same linear pedestal correction was applied, and the estimated noise level was subtracted from the integration result. Due to the linear contribution, the 15 kHz low frequency noise analysis has provided small errors, up to 2%. In the second case, the linear pedestal correction overestimates the sine-like contribution of the 100 kHz low frequency noise, and the error increase, i.e. up to 60%, with the amplitude of the applied noise.

Appendix A

Probability Distributions and Random Number Generation

A.1 The Exponential Distribution

The common way to generate random numbers from an exponential distribution is to use the inverse of the cumulative function. If the exponential distribution function has the form:

$$f(x) = \frac{1}{\lambda} \cdot e^{-\frac{x}{\lambda}}, \qquad for \quad x \geq 0 \ , \tag{A.1}$$

where λ is the mean of the distribution, then the cumulative distribution is:

$$F(x) = \int_0^x f(x)dx = 1 - e^{-\frac{x}{\lambda}} \ . \tag{A.2}$$

Applying the inverse of the cumulative distribution function, any variable *x* is generated from an exponential distribution:

$$x = F^{-1}(\xi) = -\lambda ln(1-\xi) = -\lambda ln(\xi\text{`}) \ , \tag{A.3}$$

where ξ is uniformly distributed in the unit interval, and $\xi\text{`} = 1 - \xi$.

A.2 The Uniform Distribution

The uniform distribution describes a constant probability density on an interval [a,b] and zero outside it:

$$f(x) = \frac{1}{b-a}, \qquad for \quad a \leq x \leq b \ . \tag{A.4}$$

The cumulative distribution is:

$$F(x) = \begin{cases} 0 & if \quad x \leq a \\ \frac{x-a}{b-a} & if \quad a \leq x \leq b \\ 1 & if \quad b \leq x \end{cases} \tag{A.5}$$

Any variable *x* is generated from an uniform distribution if:

$$x = F^{-1}(\xi) = (b-a)\xi + a \ , \tag{A.6}$$

where ξ is uniformly distributed in the unit interval.

A.3 The Gauss Distribution

The Gauss distribution:

$$f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \quad , \tag{A.7}$$

where μ is the mean and σ is the standard deviation of the distribution. There are several methods to obtain a random number from a Gaussian distribution. The method used for the present work is the method implemented in the ROOT physics analysis package, which is similar with the *Box-Muller transformation*. If y and z are uniformly distributed in the unit interval, then the variable x is distributed following a Gaussian with mean μ and width σ, if:

$$x = \mu + \sigma \cdot sin(2\pi z)\sqrt{-2 \cdot log(y)} \quad . \tag{A.8}$$

Appendix B

Plots and Tables

B.1 Numerical Integrations and Gaussian Fit Results

The relative energy differences given by equation 6.8 are filled into histograms and the resulting distributions are fitted with a Gaussian function. The mean and the width of the distributions ($Mean_{Data}$, RMS_{Data}) and of the Gauss fit ($Mean_{Fit}$, RMS_{Fit}), are given in the tables below. The last column in each table show how often the relative values larger than $6 \cdot RMS_{Fit}$ occur.

Integration Method	$Mean_{Data}$	RMS_{Data}	$Mean_{Fit}$	RMS_{Fit}	Outside $6 \cdot RMS_{Fit}$
Trapezium rule	$4.11 \cdot 10^{-5}$	$1.004 \cdot 10^{-3}$	$4.74 \cdot 10^{-5}$	$7.54 \cdot 10^{-4}$	0.36 %
Extended Simpson	$2.84 \cdot 10^{-5}$	$1.525 \cdot 10^{-3}$	$6.87 \cdot 10^{-5}$	$1.12 \cdot 10^{-5}$	0.51 %
Averaged Simpson	$8.28 \cdot 10^{-5}$	$1.009 \cdot 10^{-3}$	$9.01 \cdot 10^{-5}$	$7.58 \cdot 10^{-5}$	0.36 %
Modified Simpson	$4.65 \cdot 10^{-5}$	$1.0102 \cdot 10^{-3}$	$5.18 \cdot 10^{-5}$	$7.62 \cdot 10^{-5}$	0.35 %

Table B.1: Results of the Gaussian fit applied to the distribution of relative energy differences (*far from the shower core* class) .

Integration Method	$Mean_{Data}$	RMS_{Data}	$Mean_{Fit}$	RMS_{Fit}	Outside $6 \cdot RMS_{Fit}$
Trapezium rule	$1.59 \cdot 10^{-5}$	$1.82 \cdot 10^{-4}$	$1.44 \cdot 10^{-5}$	$1.43 \cdot 10^{-4}$	0.25 %
Extended Simpson	$5.81 \cdot 10^{-6}$	$2.82 \cdot 10^{-4}$	$6.78 \cdot 10^{-6}$	$2.14 \cdot 10^{-4}$	0.27 %
Averaged Simpson	$9.82 \cdot 10^{-5}$	$1.85 \cdot 10^{-4}$	$9.64 \cdot 10^{-5}$	$1.43 \cdot 10^{-4}$	0.27 %
Modified Simpson	$2.78 \cdot 10^{-5}$	$1.85 \cdot 10^{-4}$	$2.60 \cdot 10^{-5}$	$1.45 \cdot 10^{-4}$	0.27 %

Table B.2: Results of the Gaussian fit applied to the distribution of relative energy differences (*relative distance* class) .

Integration Method	$Mean_{Data}$	RMS_{Data}	$Mean_{Fit}$	RMS_{Fit}	Outside $6 \cdot RMS_{Fit}$
Trapezium rule	$-4.54 \cdot 10^{-5}$	$3.70 \cdot 10^{-4}$	$-7.42 \cdot 10^{-6}$	$1.05 \cdot 10^{-4}$	1.36 %
Extended Simpson	$-3.92 \cdot 10^{-5}$	$2.14 \cdot 10^{-4}$	$-4.98 \cdot 10^{-5}$	$1.65 \cdot 10^{-4}$	0.75 %
Averaged Simpson	$5.52 \cdot 10^{-4}$	$7.21 \cdot 10^{-4}$	$3.12 \cdot 10^{-4}$	$1.83 \cdot 10^{-4}$	9.6 %
Modified Simpson	$1.99 \cdot 10^{-4}$	$5.92 \cdot 10^{-4}$	$6.93 \cdot 10^{-5}$	$1.24 \cdot 10^{-4}$	5.3 %

Table B.3: Results of the Gaussian fi t applied to the distribution of relative energy differences (*close to the shower core* class) .

B.2 Digitization and Quantization Errors

B.2.1 Results of the *close to the shower core* and *far from the shower core* classes

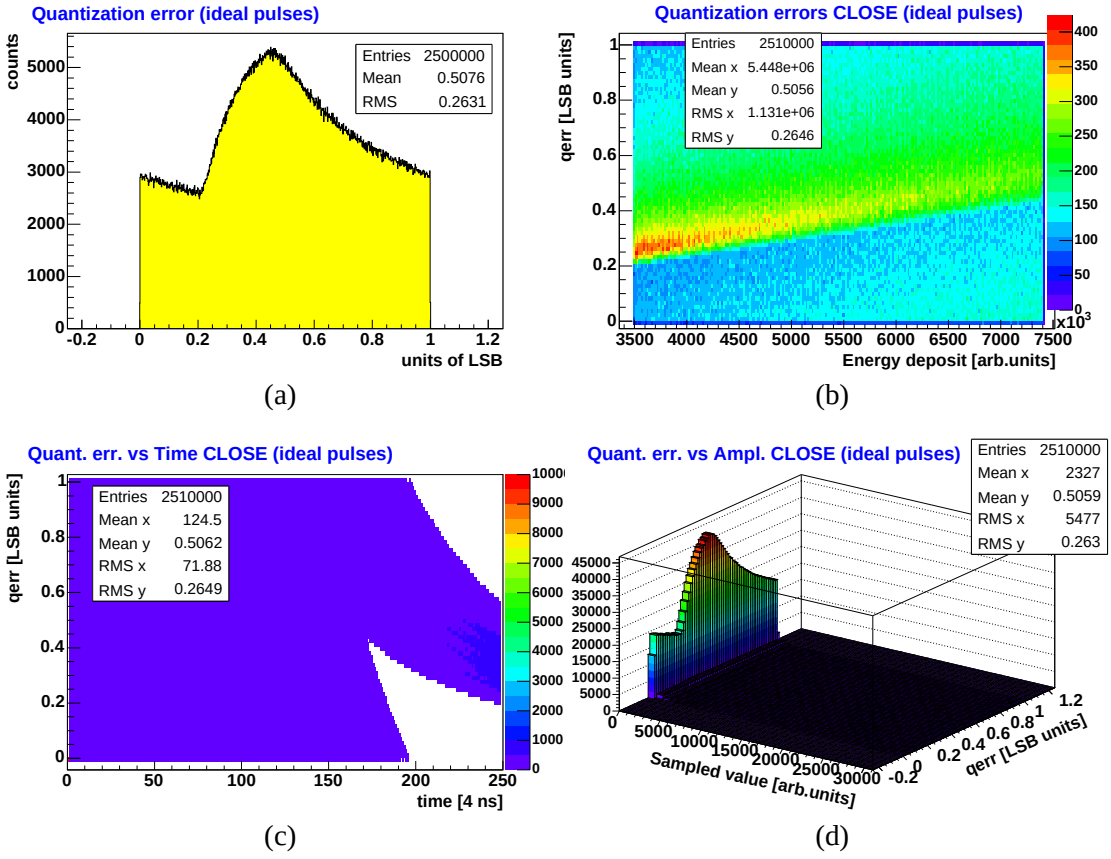


Figure B.1: Distributions of the quantization errors for the pulses from *close to the shower core* class. The dominant peak observed in the distribution of the quantization errors (plot *a*) is generated by the digitization of the sampled amplitudes from the tail (plot *d*), especially at the end of the signal (plot *c*). The errors spread towards high energies (plot *b*).

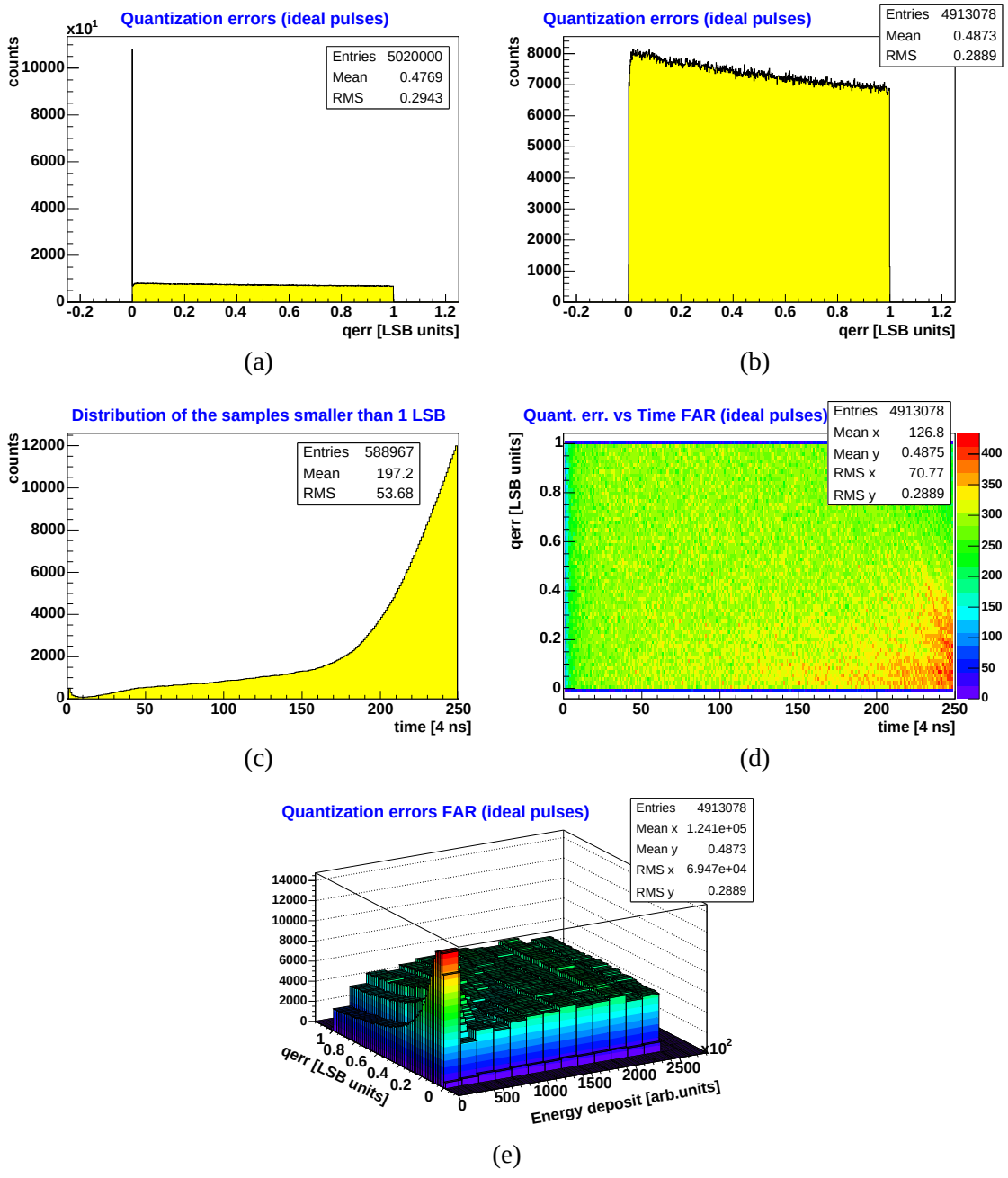


Figure B.2: Distributions of the quantization errors for the pulses from *far from the shower core class*. About 2% of the digitized samples are zero ((plot *a*) and their contribution is excluded from the distribution of the quantization errors (plot *b*). About 12% of the sampled amplitudes are smaller than 1 LSB, and most of them are taken from the tail (plot *c*). The errors are mostly generated by the digitization of the sampled amplitudes from the end of the signal (plot *d*), and are predominant at small energy deposits (plot *e*).

B.2.2 Results of the *relative distance* class, high and low gain analysis

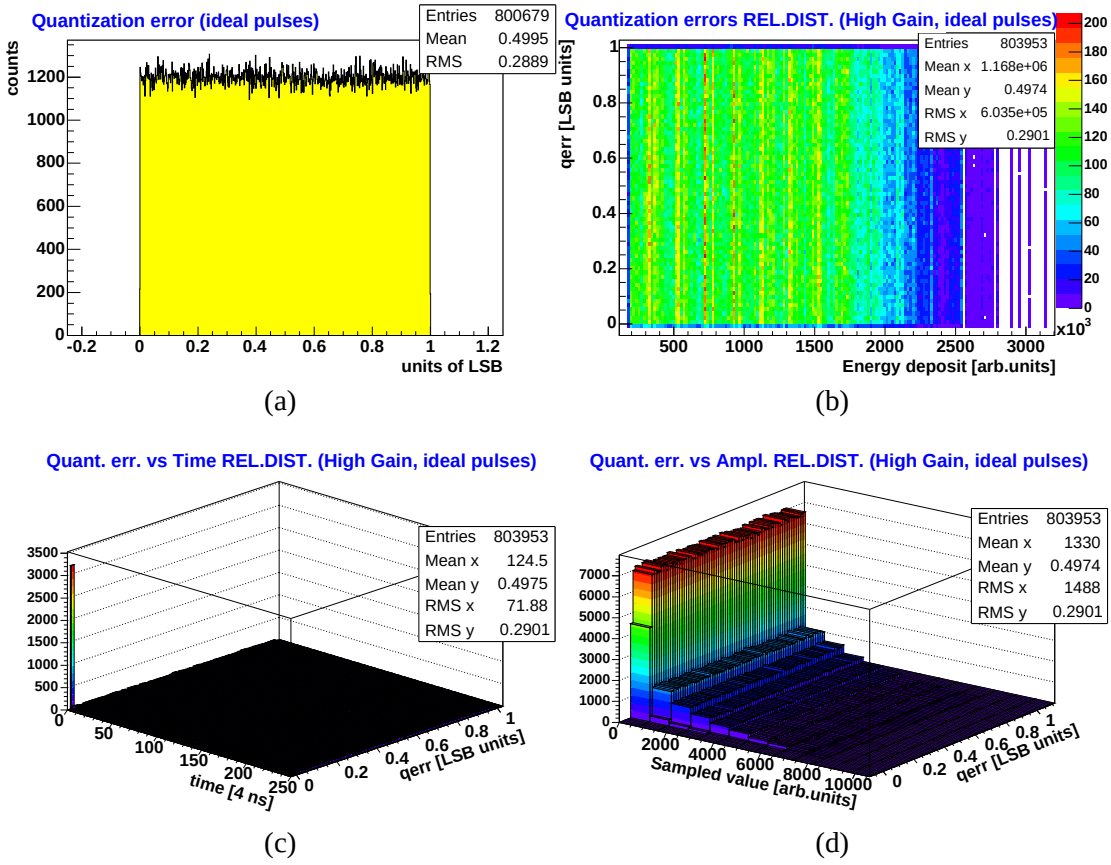


Figure B.3: Distributions of the quantization errors for the pulses from the *relative distance* class, digitized in the high gain mode. The quantization errors are uniformly distributed (plot *a*). A negligible amount of samples ($\sim 0.4\%$) are smaller in amplitude than 1 LSB (plot *c*). The contribution of the tail is still signifi cant (plot *d*), but it is uniformly distributed on the energy spectrum (plot *b*).

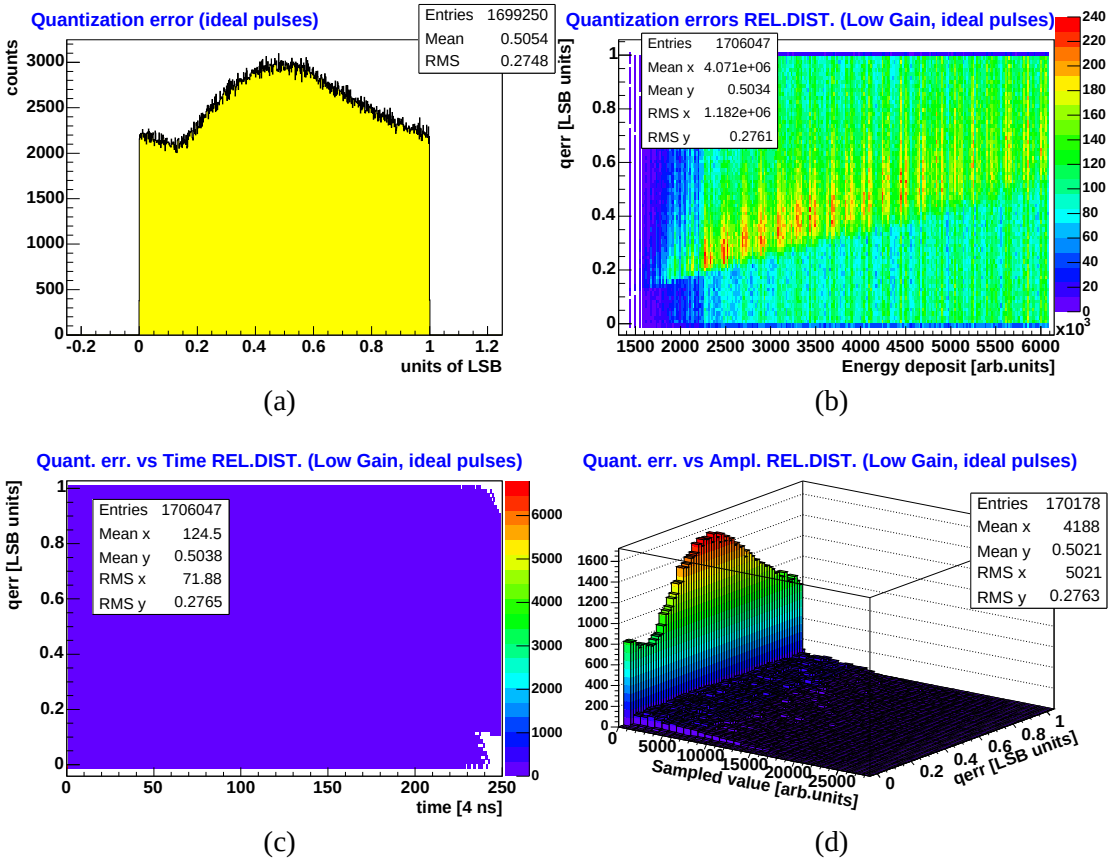


Figure B.4: Distributions of the quantization errors for the pulses from the *relative distance* class, digitized in the low gain mode. The shape of the distribution (plot *a*) is similar with that produced by the digitization of the pulses from *close to the shower core* class. The systematic effect induced by the tail influences the quality of the energy reconstruction.

B.3 Time Jitter Results

Monte Carlo class	$J_{pp} = 0\,\text{ns}$		$J_{pp} = 0.25\,\text{ns}$	
	Mean	*RMS*	*Mean*	*RMS*
far from the shower core	$-1.521 \cdot 10^{-2}$	$1.87 \cdot 10^{-2}$	$-1.52 \cdot 10^{-2}$	$1.869 \cdot 10^{-2}$
relative distance (in high gain)	$-1.244 \cdot 10^{-3}$	$1.024 \cdot 10^{-3}$	$-1.242 \cdot 10^{-3}$	$1.027 \cdot 10^{-3}$
relative distance (in low gain)	$-1.08 \cdot 10^{-2}$	$3.199 \cdot 10^{-3}$	$-1.079 \cdot 10^{-2}$	$3.203 \cdot 10^{-3}$
close to the shower core	$-7.78 \cdot 10^{-3}$	$1.477 \cdot 10^{-3}$	$-7.781 \cdot 10^{-3}$	$1.522 \cdot 10^{-3}$

Table B.4: Time jitter analysis results for $J_{pp} = 0\,\text{ns}$ and $J_{pp} = 0.25\,\text{ns}$.

Monte Carlo class	$J_{pp} = 0.5\,\text{ns}$		$J_{pp} = 0.75\,\text{ns}$	
	Mean	*RMS*	*Mean*	*RMS*
far from the shower core	$-1.521 \cdot 10^{-2}$	$1.871 \cdot 10^{-2}$	$-1.52 \cdot 10^{-2}$	$1.875 \cdot 10^{-2}$
relative distance (in high gain)	$-1.249 \cdot 10^{-3}$	$1.054 \cdot 10^{-3}$	$-1.247 \cdot 10^{-3}$	$1.061 \cdot 10^{-3}$
relative distance (in low gain)	$-1.079 \cdot 10^{-2}$	$3.202 \cdot 10^{-3}$	$-1.08 \cdot 10^{-2}$	$3.203 \cdot 10^{-3}$
close to the shower core	$-7.799 \cdot 10^{-3}$	$1.663 \cdot 10^{-3}$	$-7.805 \cdot 10^{-3}$	$1.86 \cdot 10^{-3}$

Table B.5: Time jitter analysis results for $J_{pp} = 0.5\,\text{ns}$ and $J_{pp} = 0.75\,\text{ns}$.

Monte Carlo class	$J_{pp} = 1\,\text{ns}$		$J_{pp} = 2.5\,\text{ns}$	
	Mean	*RMS*	*Mean*	*RMS*
far from the shower core	$-1.521 \cdot 10^{-2}$	$1.877 \cdot 10^{-2}$	$-1.519 \cdot 10^{-2}$	$1.888 \cdot 10^{-2}$
relative distance (in high gain)	$-1.248 \cdot 10^{-3}$	$1.103 \cdot 10^{-3}$	$-1.268 \cdot 10^{-3}$	$1.45 \cdot 10^{-3}$
relative distance (in low gain)	$-1.08 \cdot 10^{-2}$	$3.204 \cdot 10^{-3}$	$-1.078 \cdot 10^{-2}$	$3.275 \cdot 10^{-3}$
close to the shower core	$-7.8 \cdot 10^{-3}$	$2.079 \cdot 10^{-3}$	$-7.783 \cdot 10^{-3}$	$4 \cdot 10^{-3}$

Table B.6: Time jitter analysis results for $J_{pp} = 1\,\text{ns}$ and $J_{pp} = 2.5\,\text{ns}$.

B.4 White Noise Analysis

The energy deposit spectrum is divided in slices of energy of 50,000 arbitrary units. The relative values, obtained by applying equation 6.8, are filled into histograms. The mean and the width of each distribution are extracted and plotted as functions of the corresponding slice of energy. The following figures depict such plots, for different RMS values of the white noise.

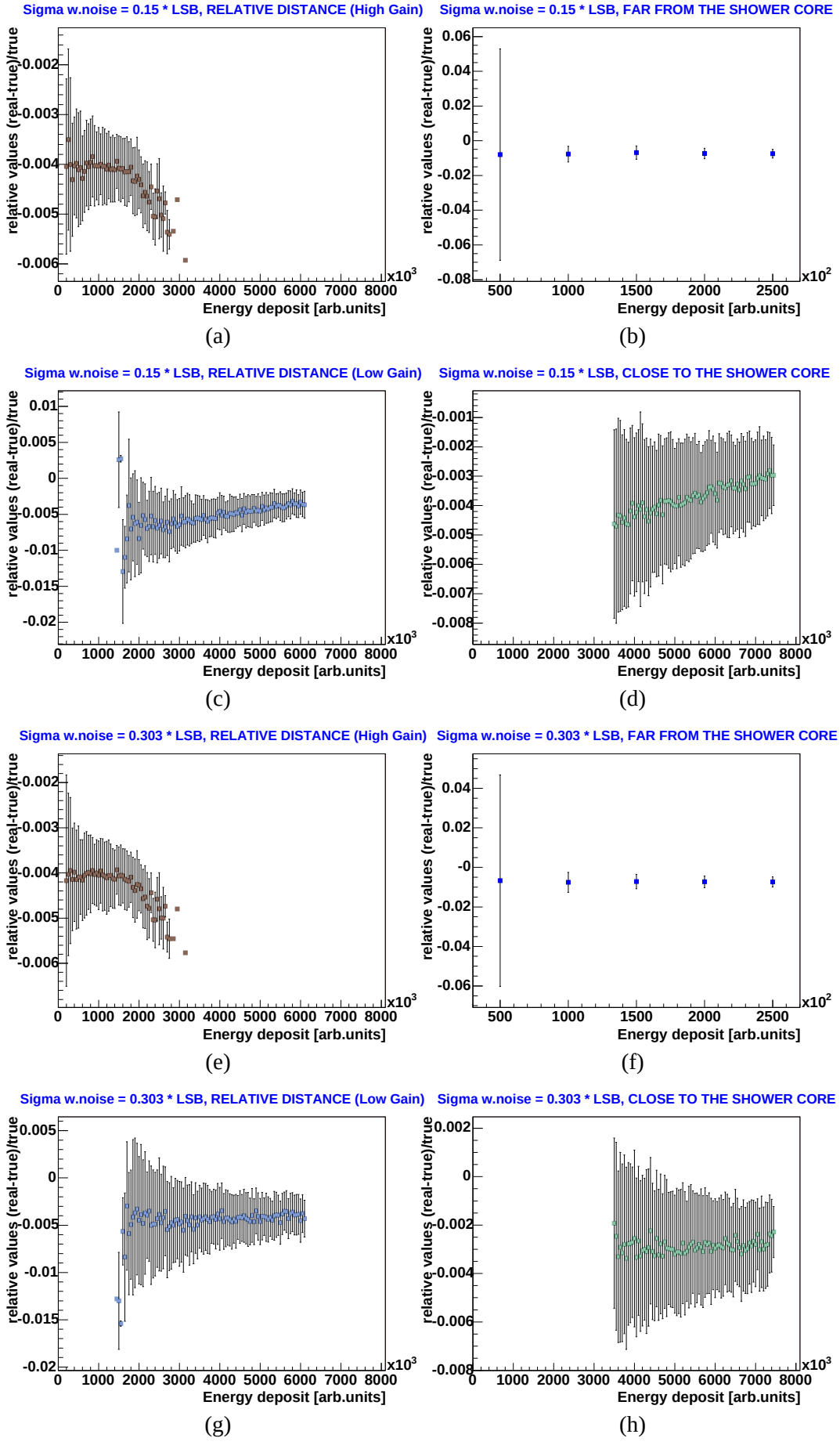


Figure B.5: The reconstruction results for RMS values of the white noise of 0.15 and 0.303 ADC channels.

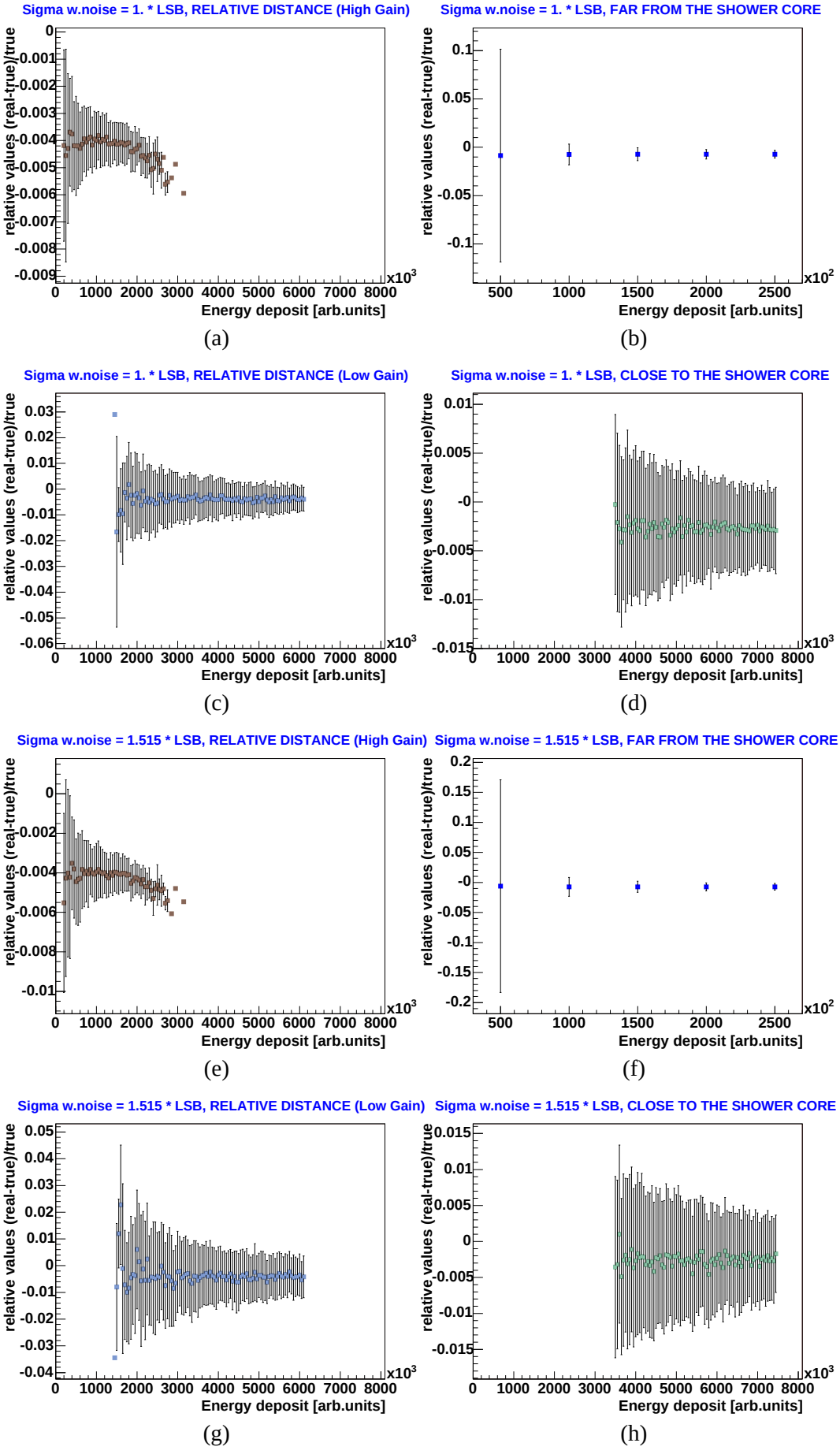


Figure B.6: The reconstruction results for RMS values of the white noise of 1 and 1.515 ADC channels.

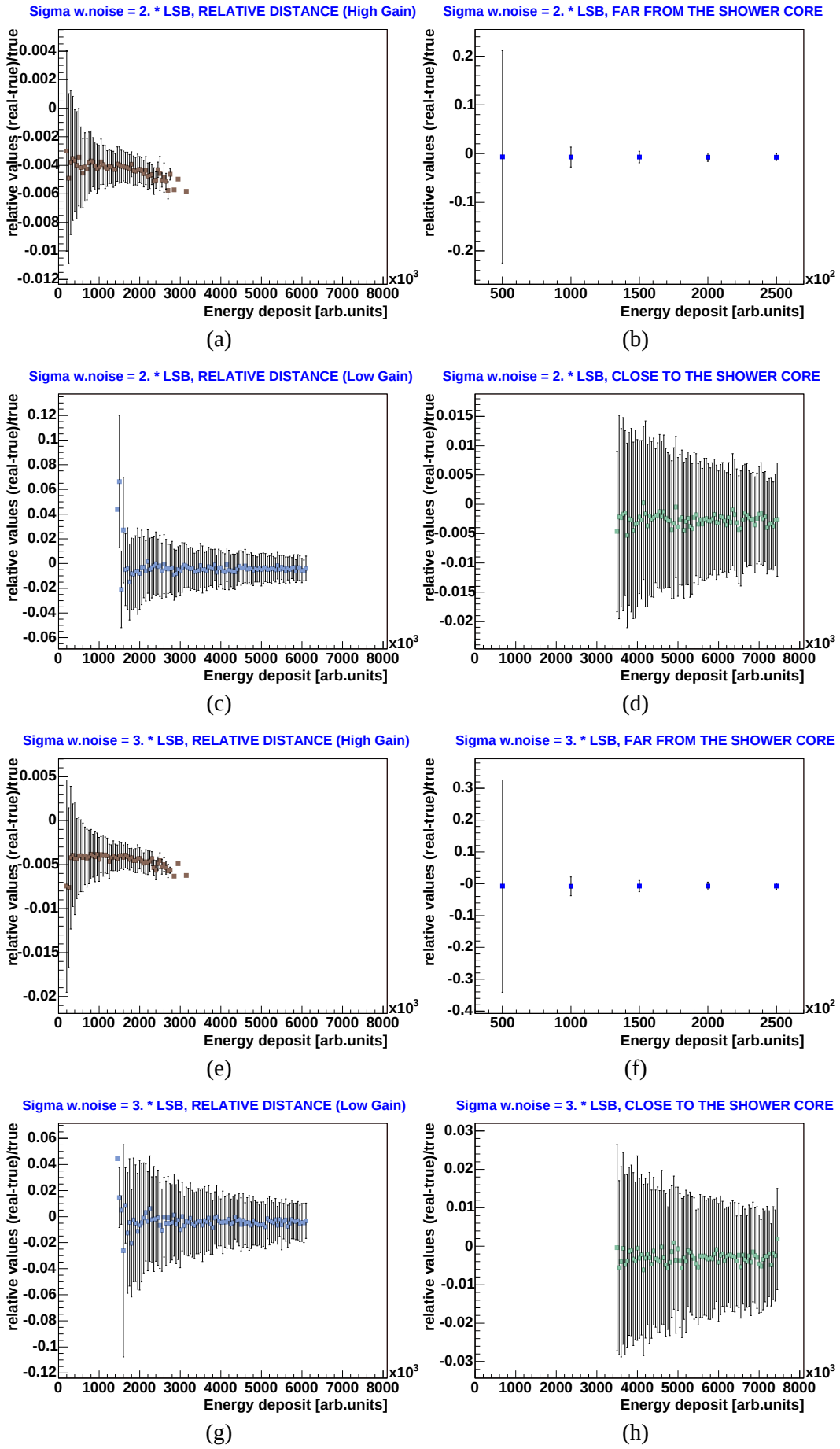


Figure B.7: The reconstruction results for RMS values of the white noise of 2 and 3 ADC channels.

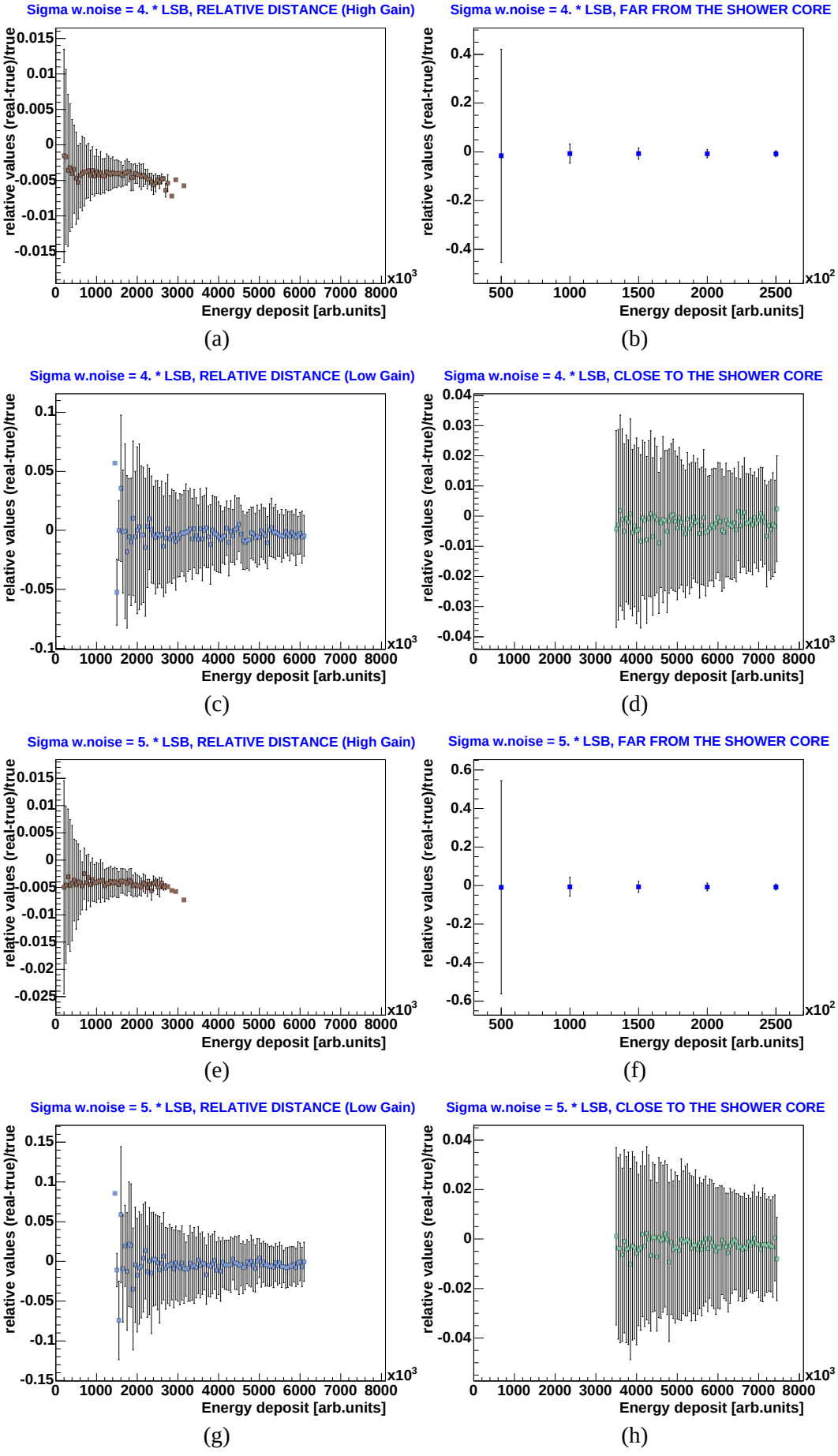


Figure B.8: The reconstruction results for RMS values of the white noise of 4 and 5 ADC channels.

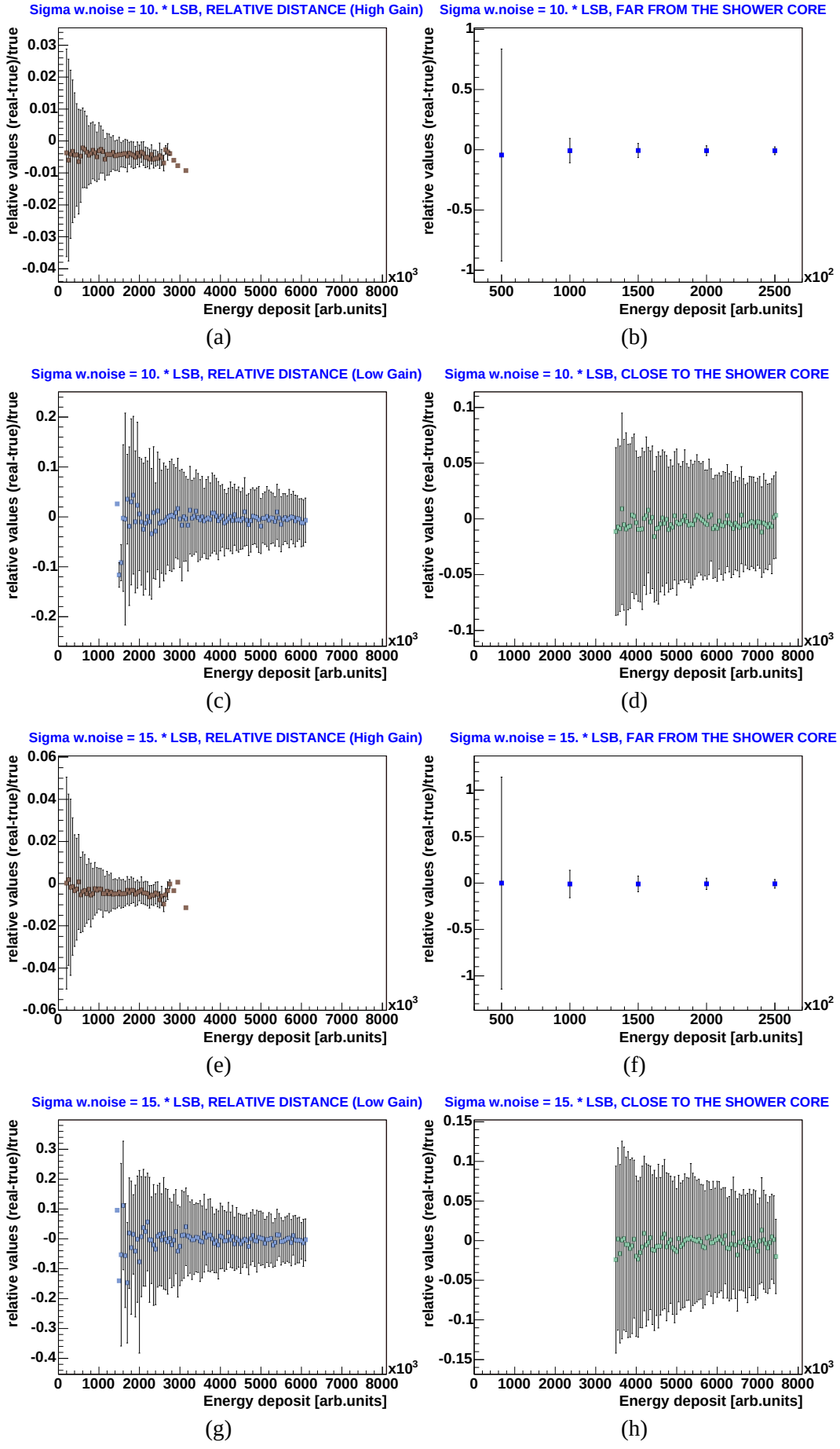


Figure B.9: The reconstruction results for RMS values of the white noise of 10 and 15 ADC channels.

B.5 Low Frequency Noise Analysis

The following tables give the reconstruction results achieved by the low frequency noise analysis. The relative values, obtained by applying equation 6.8, are filled into histograms, and the resulting mean and width (RMS) are given for 15 kHz (see subsection B.5.1) and for 100 kHz (see subsection B.5.2). The figures depict the relative values for the analysis in which the energy deposit spectrum is divided in slices of energy of 50,000 arbitrary units.

B.5.1 Low Frequency Noise, 15 kHz

Monte Carlo class	$A = 5\,\mathrm{mV}$		$A = 10\,\mathrm{mV}$	
	Mean	*RMS*	*Mean*	*RMS*
far from the shower core	$-7.31 \cdot 10^{-3}$	$1.84 \cdot 10^{-2}$	$-7.51 \cdot 10^{-3}$	$2.82 \cdot 10^{-2}$
relative distance (in high gain)	$-4.12 \cdot 10^{-3}$	$9.30 \cdot 10^{-4}$	$-4.12 \cdot 10^{-3}$	$9.59 \cdot 10^{-4}$
relative distance (in low gain)	$-3.90 \cdot 10^{-3}$	$4.53 \cdot 10^{-3}$	$-3.96 \cdot 10^{-3}$	$4.36 \cdot 10^{-3}$
close to the shower core	$-2.57 \cdot 10^{-3}$	$3.07 \cdot 10^{-3}$	$-2.69 \cdot 10^{-3}$	$2.87 \cdot 10^{-3}$

Table B.7: The relative values obtained for 15 kHz low frequency noise, with 5 mV and 10 mV noise amplitudes .

Monte Carlo class	$A = 20\,\mathrm{mV}$		$A = 30\,\mathrm{mV}$	
	Mean	*RMS*	*Mean*	*RMS*
far from the shower core	$-7.73 \cdot 10^{-3}$	$2.79 \cdot 10^{-2}$	$-7.23 \cdot 10^{-2}$	$2.72 \cdot 10^{-2}$
relative distance (in high gain)	$-4.14 \cdot 10^{-3}$	$9.53 \cdot 10^{-4}$	$-4.11 \cdot 10^{-3}$	$9.36 \cdot 10^{-4}$
relative distance (in low gain)	$-3.97 \cdot 10^{-3}$	$4.29 \cdot 10^{-3}$	$-3.97 \cdot 10^{-3}$	$4.44 \cdot 10^{-3}$
close to the shower core	$-2.67 \cdot 10^{-3}$	$2.87 \cdot 10^{-3}$	$-2.58 \cdot 10^{-3}$	$3.12 \cdot 10^{-3}$

Table B.8: The relative values obtained for 15 kHz low frequency noise, with 20 mV and 30 mV noise amplitudes .

Monte Carlo class	$A = 50\,\mathrm{mV}$		$A = 100\,\mathrm{mV}$	
	Mean	*RMS*	*Mean*	*RMS*
far from the shower core	$-7.17 \cdot 10^{-3}$	$2.63 \cdot 10^{-2}$	$-6.93 \cdot 10^{-3}$	$2.48 \cdot 10^{-2}$
relative distance (in high gain)	$-4.09 \cdot 10^{-3}$	$9.70 \cdot 10^{-4}$	$-4.10 \cdot 10^{-3}$	$9.27 \cdot 10^{-4}$
relative distance (in low gain)	$-3.83 \cdot 10^{-3}$	$4.56 \cdot 10^{-3}$	$-3.80 \cdot 10^{-3}$	$4.67 \cdot 10^{-3}$
close to the shower core	$-2.60 \cdot 10^{-3}$	$3.12 \cdot 10^{-3}$	$-2.44 \cdot 10^{-3}$	$3.11 \cdot 10^{-3}$

Table B.9: The relative values obtained for 15 kHz low frequency noise, with 50 mV and 100 mV noise amplitudes .

Monte Carlo class	$A = 150\,\mathrm{mV}$		$A = 200\,\mathrm{mV}$	
	Mean	*RMS*	*Mean*	*RMS*
far from the shower core	$-6.73 \cdot 10^{-3}$	$2.71 \cdot 10^{-2}$	$-6.89 \cdot 10^{-3}$	$3.03 \cdot 10^{-2}$
relative distance (in high gain)	$-4.07 \cdot 10^{-3}$	$9.81 \cdot 10^{-4}$	$4.07 \cdot 10^{-3}$	$9.85 \cdot 10^{-4}$
relative distance (in low gain)	$-3.70 \cdot 10^{-3}$	$4.82 \cdot 10^{-3}$	$-3.61 \cdot 10^{-3}$	$5.17 \cdot 10^{-3}$
close to the shower core	$-0.002456 \cdot 10$	$0.003119 \cdot 10$	$-0.002397 \cdot 10$	$0.003295 \cdot 10$

Table B.10: The relative values obtained for 15 kHz low frequency noise, with 150 mV and 200 mV noise amplitudes .

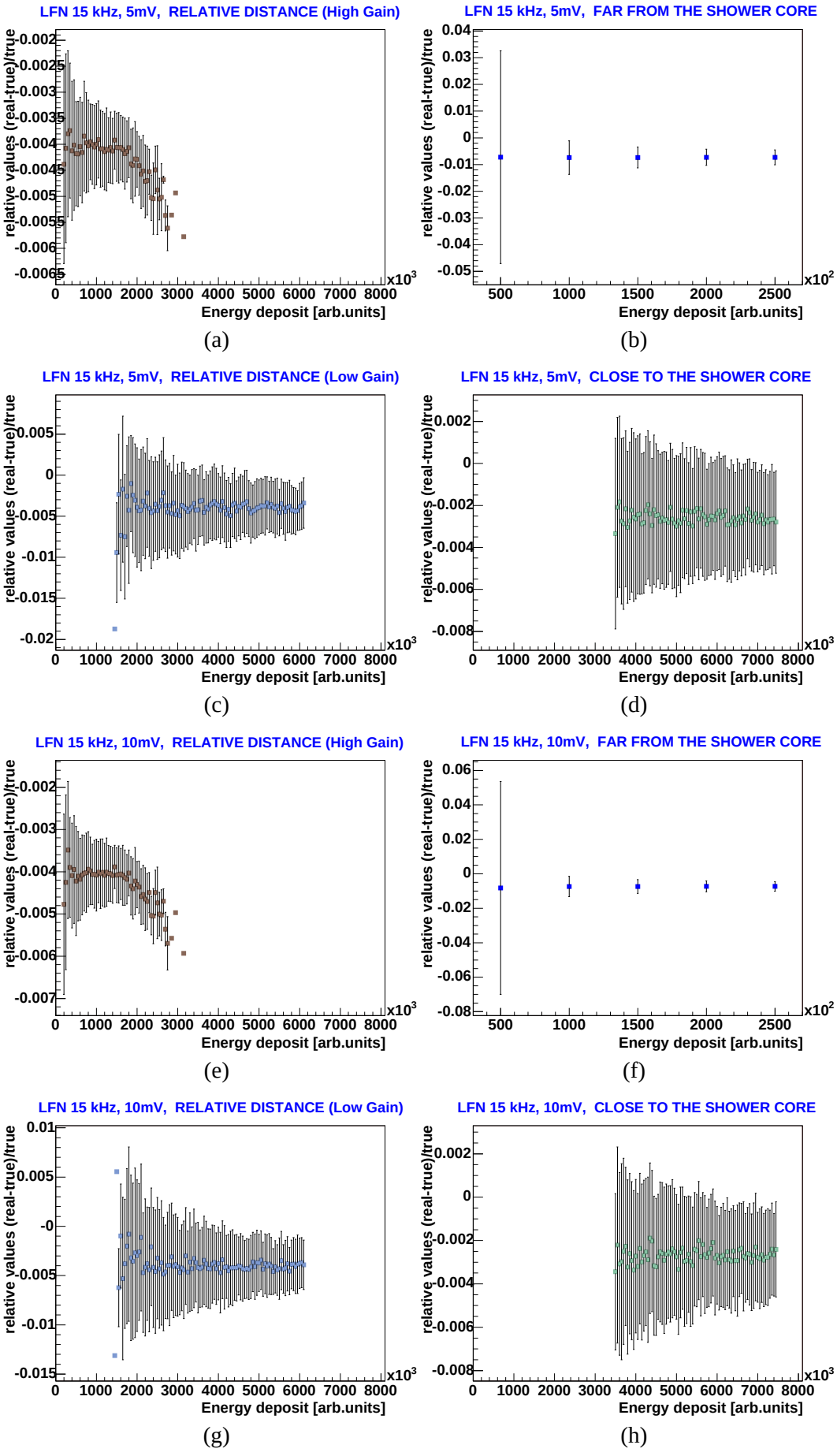


Figure B.10: The reconstruction results for 15 kHz low frequency noise with 5 mV and 10 mV noise amplitudes.

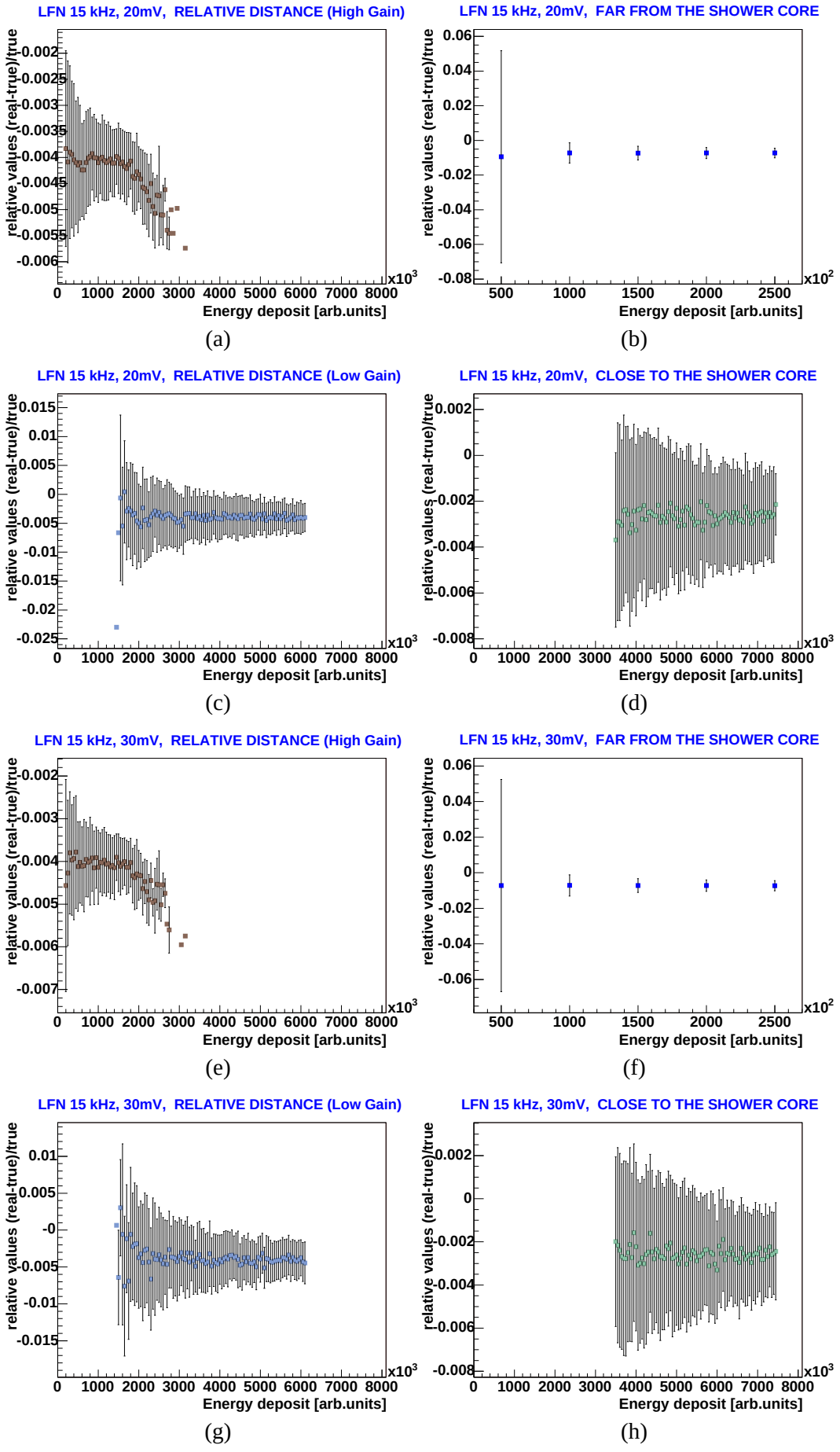


Figure B.11: The reconstruction results for 15 kHz low frequency noise with 20 mV and 30 mV noise amplitudes.

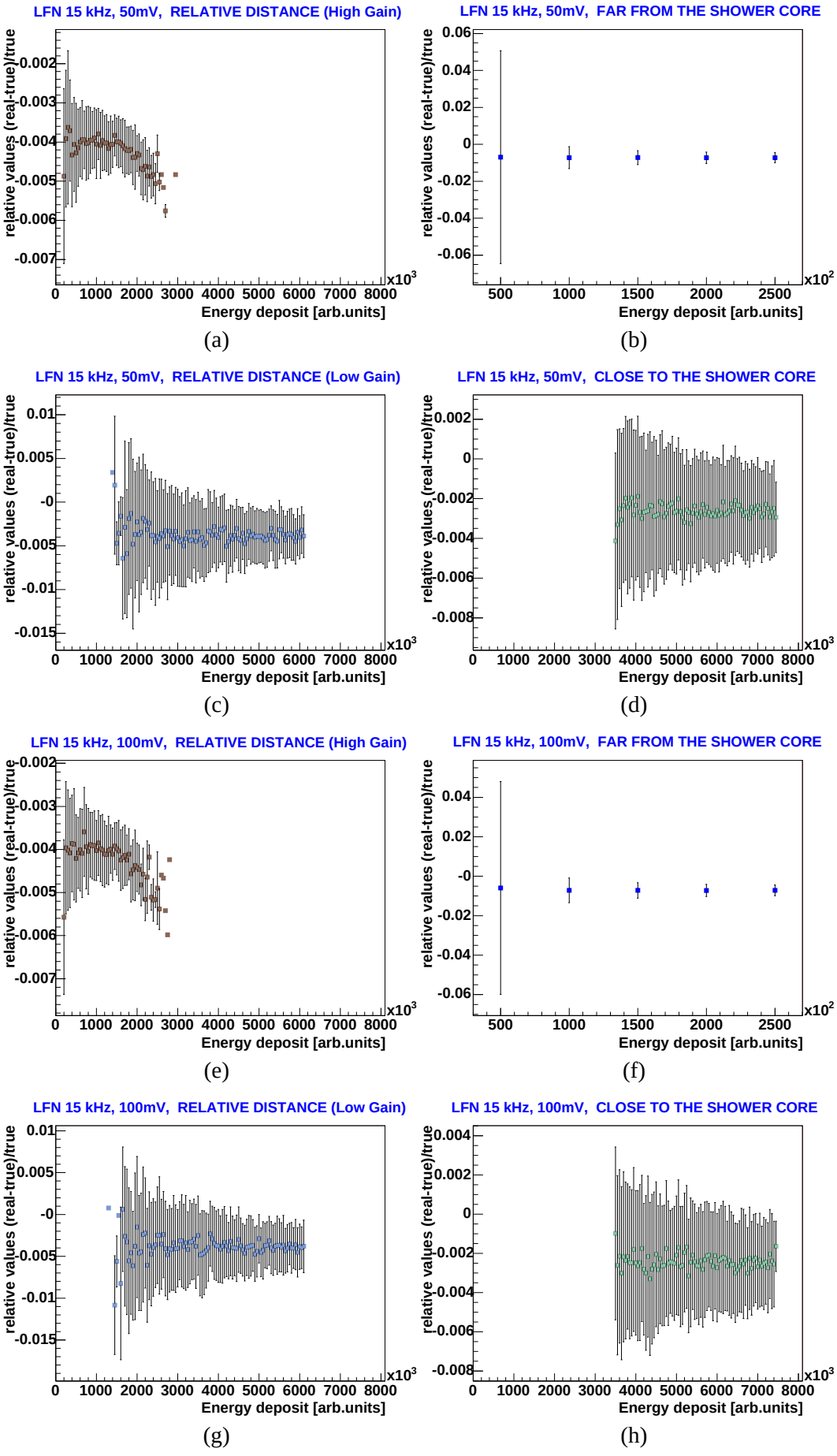


Figure B.12: The reconstruction results for 15 kHz low frequency noise with 50 mV and 100 mV noise amplitudes.

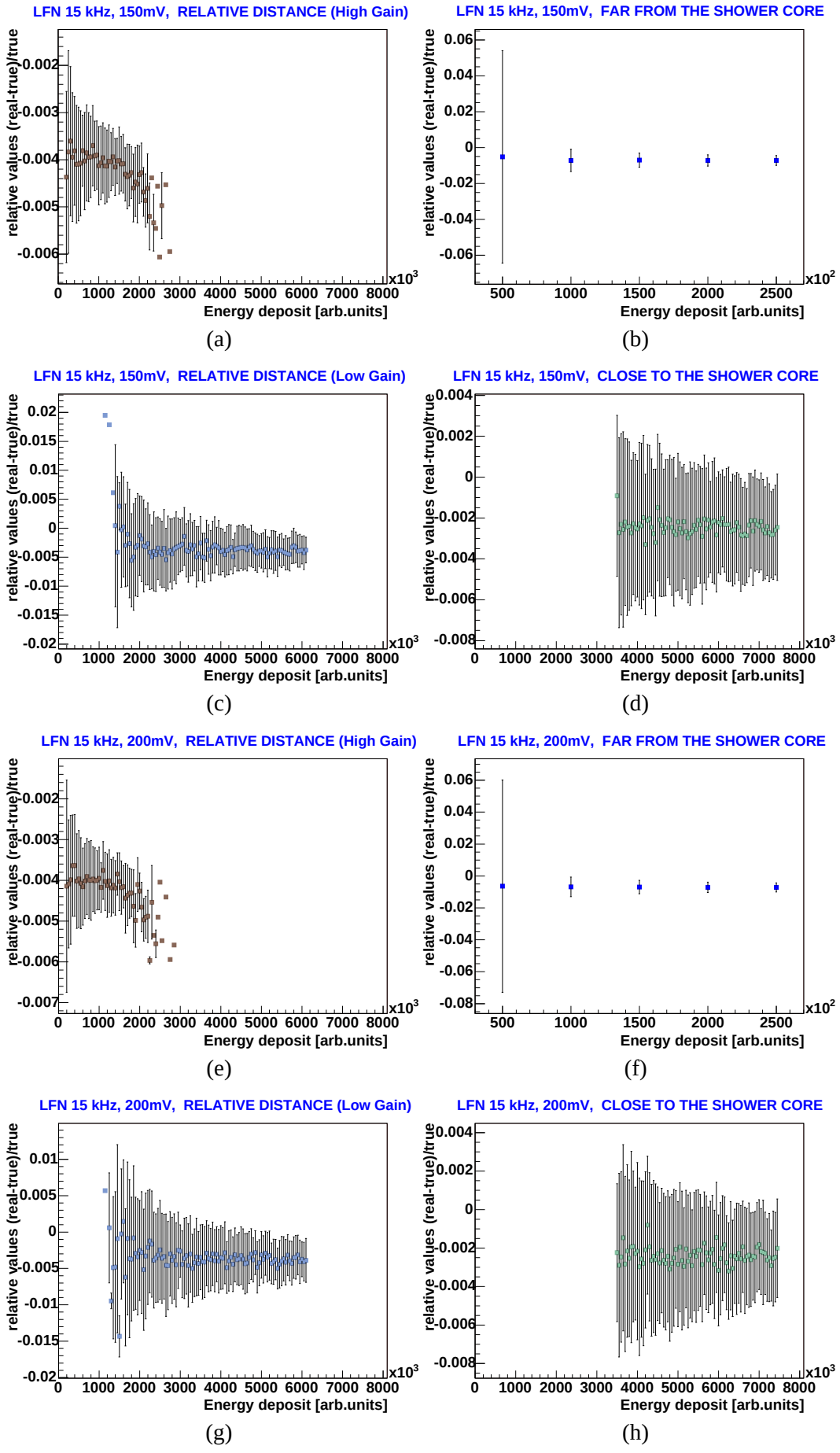


Figure B.13: The reconstruction results for 15 kHz low frequency noise with 150 mV and 200 mV noise amplitudes.

B.5.2 Low Frequency Noise, 100 kHz

Monte Carlo class	$A = 5\,\mathrm{mV}$		$A = 10\,\mathrm{mV}$	
	Mean	*RMS*	*Mean*	*RMS*
far from the shower core	$-7.38 \cdot 10^{-3}$	$2.86 \cdot 10^{-2}$	$-7.46 \cdot 10^{-3}$	$2.965 \cdot 10^{-2}$
relative distance (in high gain)	$-4.13 \cdot 10^{-3}$	$9.63 \cdot 10^{-4}$	$-4.12 \cdot 10^{-3}$	$9.70 \cdot 10^{-4}$
relative distance (in low gain)	$-3.90 \cdot 10^{-3}$	$4.41 \cdot 10^{-3}$	$-4.02 \cdot 10^{-3}$	$4.96 \cdot 10^{-3}$
close to the shower core	$-2.56 \cdot 10^{-3}$	$3.08 \cdot 10^{-3}$	$-2.74 \cdot 10^{-3}$	$3.33 \cdot 10^{-3}$

Table B.11: The relative values obtained for 100 kHz low frequency noise, with 5 mV and 10 mV noise amplitudes .

Monte Carlo class	$A = 20\,\mathrm{mV}$		$A = 30\,\mathrm{mV}$	
	Mean	*RMS*	*Mean*	*RMS*
far from the shower core	$-7.32 \cdot 10^{-3}$	$3.16 \cdot 10^{-2}$	$-3.53 \cdot 10^{-3}$	$3.92 \cdot 10^{-2}$
relative distance (in high gain)	$-4.16 \cdot 10^{-3}$	$1.05 \cdot 10^{-3}$	$-3.99 \cdot 10^{-3}$	$1.11 \cdot 10^{-3}$
relative distance (in low gain)	$-4.07 \cdot 10^{-3}$	$5.47 \cdot 10^{-3}$	$-2.53 \cdot 10^{-3}$	$6.16 \cdot 10^{-3}$
close to the shower core	$-2.76 \cdot 10^{-3}$	$3.70 \cdot 10^{-3}$	$-1.60 \cdot 10^{-3}$	$4.05 \cdot 10^{-3}$

Table B.12: The relative values obtained for 100 kHz low frequency noise, with 20 mV and 30 mV noise amplitudes .

Monte Carlo class	$A = 50\,\mathrm{mV}$		$A = 100\,\mathrm{mV}$	
	Mean	*RMS*	*Mean*	*RMS*
far from the shower core	$8.97 \cdot 10^{-4}$	$4.60 \cdot 10^{-2}$	$1.15 \cdot 10^{-2}$	$6.10 \cdot 10^{-2}$
relative distance (in high gain)	$-3.63 \cdot 10^{-3}$	$1.38 \cdot 10^{-3}$	$-2.77 \cdot 10^{-4}$	$2.09 \cdot 10^{-3}$
relative distance (in low gain)	$3.15 \cdot 10^{-4}$	$7.77 \cdot 10^{-3}$	$7.10 \cdot 10^{-3}$	$1.31 \cdot 10^{-2}$
close to the shower core	$3.07 \cdot 10^{-4}$	$5.09 \cdot 10^{-3}$	$4.627 \cdot 10^{-3}$	$8.30 \cdot 10^{-2}$

Table B.13: The relative values obtained for 100 kHz low frequency noise, with 50 mV and 100 mV noise amplitudes .

Monte Carlo class	$A = 150\,\mathrm{mV}$		$A = 200\,\mathrm{mV}$	
	Mean	*RMS*	*Mean*	*RMS*
far from the shower core	$2.08 \cdot 10^{-2}$	$7.87 \cdot 10^{-2}$	$2.86 \cdot 10^{-2}$	$9.01 \cdot 10^{-2}$
relative distance (in high gain)	$-2.09 \cdot 10^{-3}$	$3.20 \cdot 10^{-3}$	$-1.26 \cdot 10^{-3}$	$4.05 \cdot 10^{-3}$
relative distance (in low gain)	$1.45 \cdot 10^{-2}$	$1.92 \cdot 10^{-2}$	$2.17 \cdot 10^{-2}$	$2.53 \cdot 10^{-2}$
close to the shower core	$9.25 \cdot 10^{-3}$	$1.15 \cdot 10^{-2}$	$1.38 \cdot 10^{-2}$	$1.50 \cdot 10^{-2}$

Table B.14: The relative values obtained for 100 kHz low frequency noise, with 150 mV and 200 mV noise amplitudes .

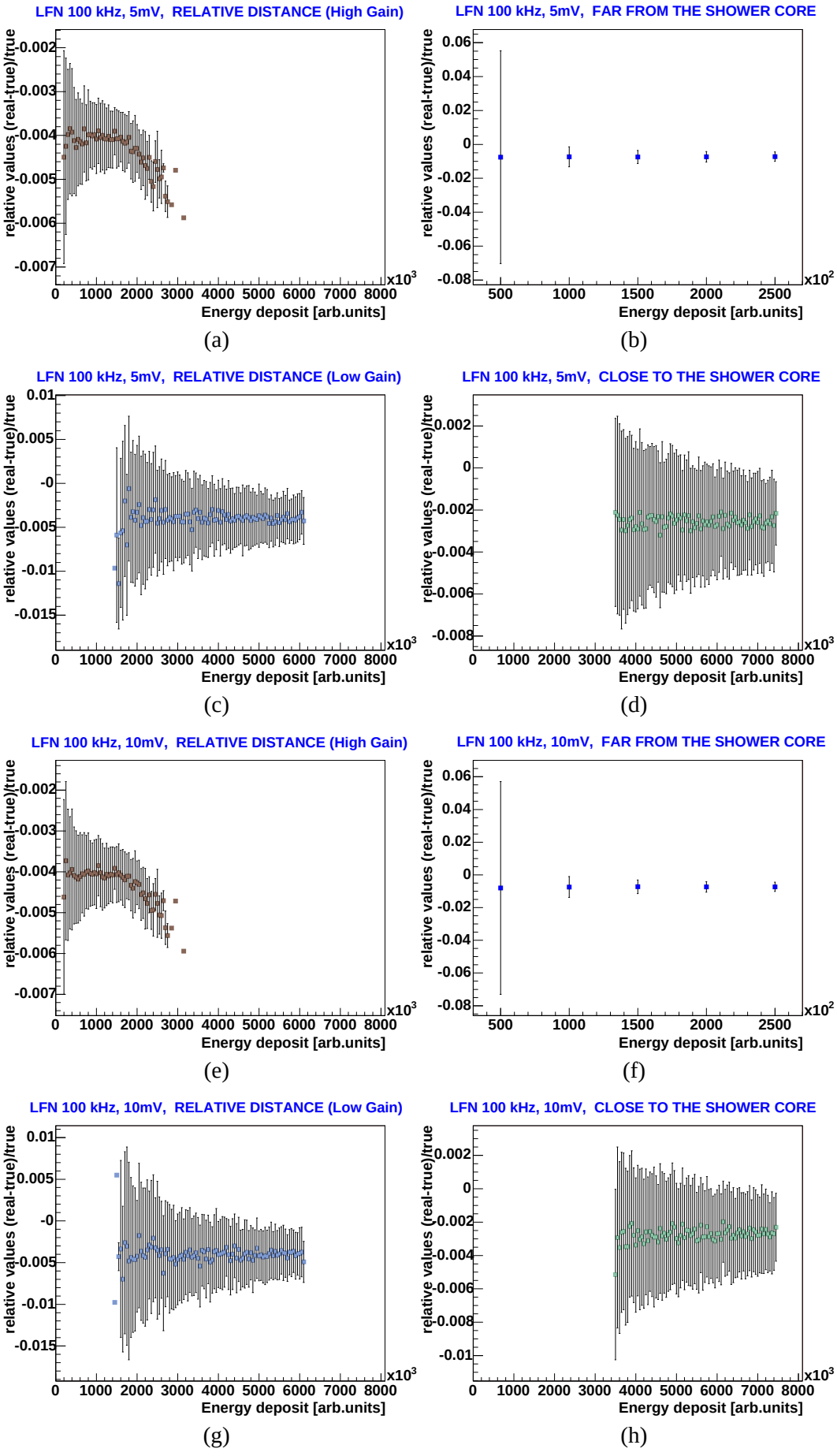


Figure B.14: The reconstruction results for 100 kHz low frequency noise with 5 mV and 10 mV noise amplitudes.

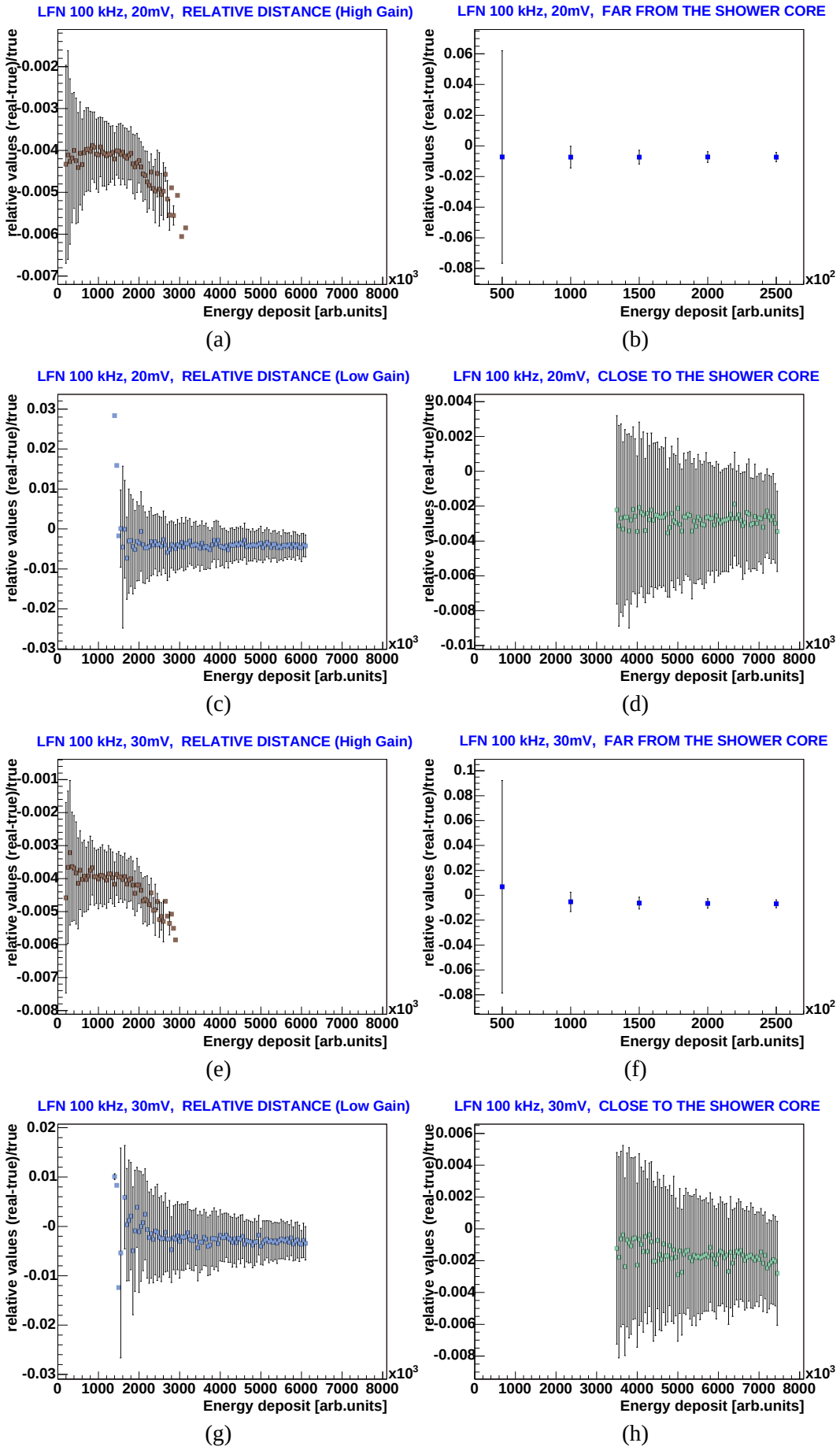


Figure B.15: The reconstruction results for 100 kHz low frequency noise with 20 mV and 30 mV noise amplitudes.

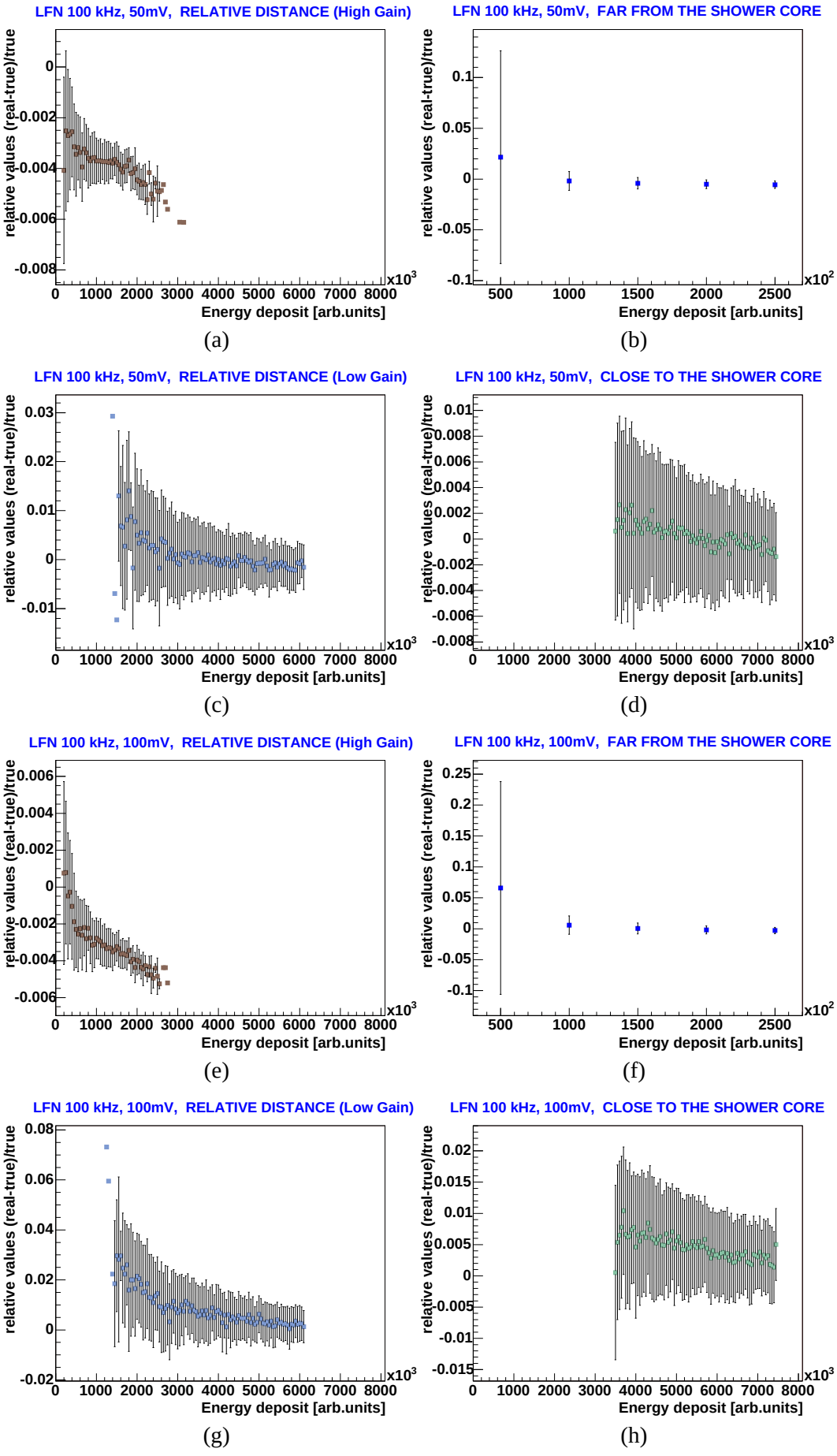


Figure B.16: The reconstruction results for 100 kHz low frequency noise with 50 mV and 100 mV noise amplitudes.

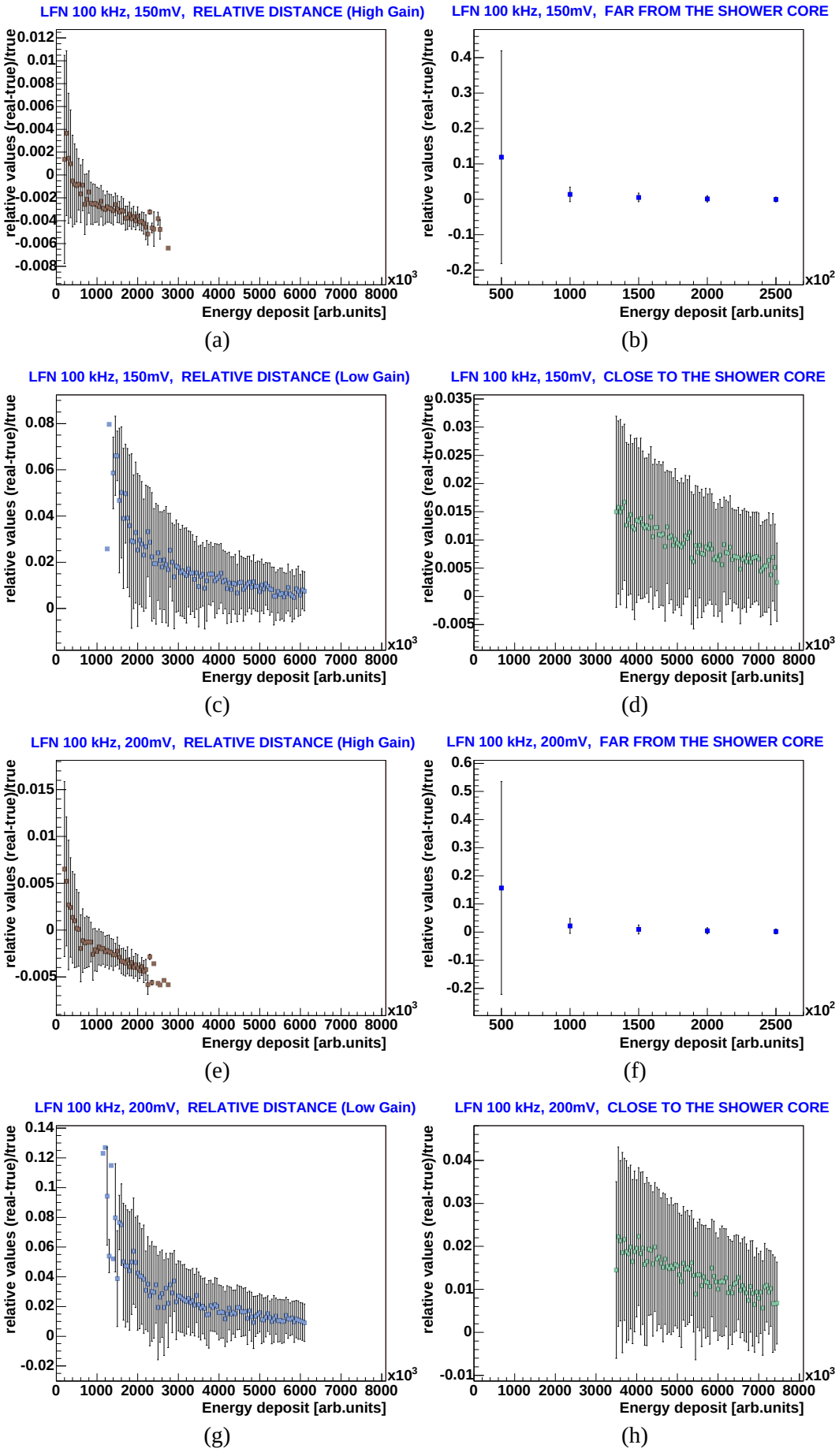


Figure B.17: The reconstruction results for 100 kHz low frequency noise with 150 mV and 200 mV noise amplitudes.

List of Figures

List of Tables

List of Acronyms

Acronym	Explanation
CORSIKA	COsmic Ray SImulations for KAscade
CPU	Central Processing Unit
CRES	Cosmic Ray Event Simulation
EAS	Extensive Air Shower
FPGA	Field Programmable Gate Array
FZK	Forschungszentrum Karlsruhe
FADC	Flash analog-to-digital converter
FIFO	First-In First-Out
KASCADE	KArlsruhe Shower Core and Array DEtector
KGEMD	KASCADE-Grande Electromagnetic Detector Digitizer Board
KGEMS	KASCADE-Grande Electromagnetic Detector Storage Board
KGEMP	KASCADE-Grande Electromagnetic Detector PCI Interface Card
KGEMT	KASCADE-Grande Electromagnetic Detector Trigger Receiver Card
LVDS	Low Voltage Differential Signal
m.i.p	minimum ionizing paricle
NKG	Nishimura-Kamata-Greisen (function)
RMS	Root Mean Square
VME	Versa Module Eurocard

Bibliography

[Agl93] M. Aglietta et al., *UHE cosmic ray event reconstruction by the electromagnetic detector of EAS-TOP*, Nucl. Instrum. Meth. **A336** (1993), 310–321.

[Bad03] F. Badea, Forschungszentrum Karlsruhe, private communication, October 2003.

[Ber01] M. Bertaina et al., *KASCADE-Grande: A conclusive experiment on the knee*, Proc. 27th International Cosmic Ray Conference (ICRC 2001) (K.-H. Kampert, G. Heinzelmann, and C. Spiering, eds.), vol. 2, 2001, pp. 792–795.

[Bra01] M. Brancus and H. Rebel, *Time structure of the extensive air shower muon component measured by the KASCADE experiment*, Astroparticle Physics **15** (2001), 149–165.

[Bru97] R. Brun and F. Rademakers, *ROOT - An Object Oriented Data Analysis Framework*, Nucl. Inst. Meth. in Phys. Res. **A 389** (1997), 81–86.

[Chi03] A. Chiavassa et al., *KASCADE-Grande: the Grande Array*, Proc. 28th International Cosmic Ray Conference (ICRC 2003), vol. 2, 2003, pp. 989–992.

[Eng03] J. Engler et al., *The cosmic-ray experiment KASCADE*, Nucl. Inst. Meth. in Phys. Res., Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **513** (2003), 490–510.

[Gla03] R. Glasstetter et al., *Shower Reconstruction Performance of KASCADE-Grande*, Proc. 28th International Cosmic Ray Conference (ICRC 2003), vol. 2, 2003, pp. 781–784.

[Gre56] see e.g. references in: Greisen, K., *Progress in Cosmic Ray Physics*, North-Holland Publishing Company, 1956.

[Gri01] P. K. F. Grieder, *Cosmic rays at earth: Researcher's reference, manual and data book*, Amsterdam, Netherlands: Elsevier (2001) 1093 p.

[Gru00] see e.g. references in: C. Grupen, *Astroteilchenphysik*, Vieweg, Braunschweig/Wiesbaden, 2000.

[Hau03] A. Haungs et al., *The KASCADE-Grande experiment*, Proc. 28th International Cosmic Ray Conference (ICRC 2003) (T. Kajita, Y. Asaoka, A. Kawachi, Y. Matsubara, and M. Sasaki, eds.), vol. 2, 2003, pp. 985–988.

[Hec98] D. Heck, G. Schatz, T. Thouw, J. Knapp, and J. N. Capdevielle, *CORSIKA: A Monte Carlo code to simulate extensive air showers*, FZKA-6019.

[Hin04] J. A. Hinton (The HESS Collaboration), *The status of the HESS project*, New Astron. Rev. **48** (2004), 331–337, astro-ph/0403052.

[Hus03] U. Husemann, University of Siegen, private communication, September 2003.

[Kam03] K.-H. Kampert et al., *Status of the KASCADE-Grande experiment*, Nucl. Phys. Proc. Suppl. **122** (2003), 422–426, astro-ph/0212347.

[Köl93] K. Kölbig, G110: Landau Distribution, 1985 (last revised 15.03.1993), http://wwwasdoc.web.cern.ch/wwwasdoc/shortwrupsdir/g110/top.html.

[Leo93] see e.g. references in: Leo, W.R., *Techniques for Nuclear and Particle Physics Experiments*, Springer-Verlag, 1993.

[Lon92] see e.g. references in: M. S. Longair, *High energy astrophysics, Volume 1, 2nd edition*, Cambridge Univ. Press, 1992.

[Mil72] see e.g. references in: Millman,J. and C. Halkias, *Integrated Electronics: Analog and Digital Circuits and Systems*, McGraw-Hill, 1972.

[Moc03] E. Mocchiutti, *Atmospheric and Interstellar Cosmic Rays Measured with the CAPRICE98 Experiment*, Ph.D. thesis, Kungl Tekniska Högskolan, Stockholm, 2003.

[Mot93] see e.g. references in: Motchenbacher, C.D. and J.A. Connely, *Low-Noise Electronic System Design*, Wiley-Interscience, 1993.

[Nag00] M. Nagano and A. A. Watson, *Observations and implications of the ultrahigh-energy cosmic rays*, Rev. Mod. Phys. **72** (2000), 689–732.

[Ott76] see e.g. references in: Ott, H., *Noise Reduction Techniques in Electronic Systems*, Wiley-Interscience, 1976.

[Ove04] S. Over, *Development and commissioning of data acquisition systems for the KASCADE-Grande experiment*, Diplomarbeit, University of Siegen, June 2004.

[Pre02] see e.g. references in: Press, W.H., Teukolsky S.A., Vetterling W.T. and Flannery B.P., *Numerical Recipes in C++*, Cambridge University Press, 2002.

[Rao98] see e.g. references in: M. V. S. Rao and B. V. Sreekantan, *Extensive Air Showers*, World Scientific Publishing, 1998.

[Rot03] M. Roth et al., *Energy Spectrum and Elemental Composition in the PeV Region*, Proc. 28th International Cosmic Ray Conference (ICRC 2003), 2003.

Acknowledgements

I am extremely grateful to the many people who have helped me in the study for and preparation of this thesis.

To Prof. Dr. Peter Buchholz for giving me the opportunity to work on such an interesting topic. I am grateful to him for many valuable advices and discussions on my analysis, as well as for his support in difficult times. I want to sincerely thank him for the tough lessons he gave me about experimental physics and not only. Also, I am grateful for the particular understanding he has shown accepting to receive last-minute text over the past few weeks.

I am grateful to Prof. Dr. Claus Grupen for kindly accepting to read and examine this thesis.

I am grateful to Marc Brüggemann for his valuable hints, suggestions and encouragements on my analysis, during *tons* of hours of discussions. He was a great help when nothing seemed to work well. I want to thank him for the great receptivity and patience he had whenever I was stuck in a problem, and for being daily a cheerful person.

I am also grateful to Dr. Wolfgang Walkowiak for the big help he gave me developing an object-oriented package, necessary for my simulations. For many useful discussions about the FADC DAQ system, I am indebted to all my colleagues from the KASCADE-Grande group (University of Siegen), but especially to Yuri Kolotaev, Sven Over and Dirk Zimmermann. Also, many thanks to Dr. Andreas Haungs, Dr. Florin Badea, Ralph Glasstetter, Dr. Harald Schieler and Dr. Jurgen Wochele (from the Forschungszentrum Karlsruhe) for presenting me the physics of the KASCADE-Grande experiment and providing me with useful suggestions about my analysis.

Also, I acknowledge Dr. Adrian Niculae, Dr. Wolfgang Walkowiak, Marc Brüggemann, Florin Maciuc and Sven Over for casting a professional eye on this thesis.

I would like to thank all my friends and colleagues from our working group for the enjoyable time we have spent together: Matthias Böcker, Marc Brüggemann, Tatiana Goldmann, Nadir Omar Hashim, Ulrich Husemann, Roxana Lixandru, Florin Maciuc, Adrian Niculae, Sven Over, Valentin Sipica, Rodrigo Vasquez and Dirk Zimmermann.

My special thanks and love to Maria Georgiadou, for her tremendous help, love, affection and encouragements during the last two years and a half. If everyone of us has a guardian angel, then she is my one.

Last but not least, I want to dedicate this thesis to my dear parents and my sister, that have been through very difficult times in the past months.

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Diplomarbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt, sowie Zitate und Ergebnisse Anderer kenntlich gemacht habe.

. . . .

(Ort) (Datum) (Unterschrift)