The influence of noise on radio signals from cosmic rays

Bachelor Thesis in Physics & Astronomy

Katharina Holland

Supervisor: Dr. Charles Timmermans

Institute for Mathematics, Astrophysics and Particle Physics

Radboud University Nijmegen

Preface

This thesis is the result of my Bachelor project. I have done my Bachelor project at the department of Experimental High Energy Physics of the Institute of Mathematics, Astrophysics and Particle Physics at the Radboud University of Nijmegen.
I would like to thank Jose Coppens, Antje Fitzner, Stefan Grebe, Harm Schoorlemmer and of course my supervisor Charles Timmermans for helping programming, answering questions and having a nice time.

Contents

1 Introduction

When a cosmic ray enters the atmosphere, an extensive airshower starts. Billions of particles “rain“ on the Earth. At the Pierre Auger Observatory, there are several possibilities to detect extensive air showers. One of these possibilities is radio detection of cosmic rays. Radio antennas (just as all other measurement systems) measure not only air showers but also noise. The noise originates from galactic background and signals of human devices. Many devices are small band emitters which mean that their signals only appear in a narrow frequency range, this noise can be filtered out. The galactic background is always measured. This noise superposes the signal of the air shower. It is important to know which effect the noise has on the signal, especially on the amplitude and the timing of the signal, as these are needed for a proper reconstruction of an air shower.
In this thesis the effect of noise on a signal is analysed. The following questions will be answered. At what signal to noise level is the timing affected? At what signal to noise level is the amplitude reliable, are there systematic effects is this determination? What is the systematic uncertainty on amplitude and timing due to noise? What is the effect on the reconstruction of the arrival direction of an air shower?

2 Cosmic rays & extensive air showers

Cosmic rays are high energetic particles which come from outer space and enter the atmosphere of the Earth. Most of these particles are protons (90 %). When one of these particles, with enough energy, hits the atmosphere an air shower starts. Through the interaction between the cosmic ray particle and one particle of the atmosphere, many particles are produced. These secondary particles create further particles through collision with other particles from the atmosphere. This process continues until the air shower particles no longer have enough energy to create new particles. The resulting is called extensive air shower (EAS). Cosmic rays may have energies ranging from of 10^8 to 10^{20} electron volts (eV). The particles are likely accelerated by magnetic fields, existing on the sun, in solar winds and in remnants of supernova explosions. EAS with a high energy arrive with a rate of one particle per square kilometer per year on the Earth. Due to the fact that these particles are rare, a measurement above the atmosphere is not possible and a measurement of secondary particles is necessary, to see the effect of these particles at all. The measured energy spectrum of cosmic rays can be seen in Figure 1.
In 1912 Victor Hess noticed that natural radiation comes from above. He did a series of measurements during balloon flights. He measured radiation from the ground level and up to an altitude of 5 km. His measurements showed, that the radiation decreases up to an altitude of 1 km, but increases above this level. His conclusion was that the radiation originates from above. The name ’cosmic rays’ was introduced by Robert Andrews Millikan in the 1920s.[1]

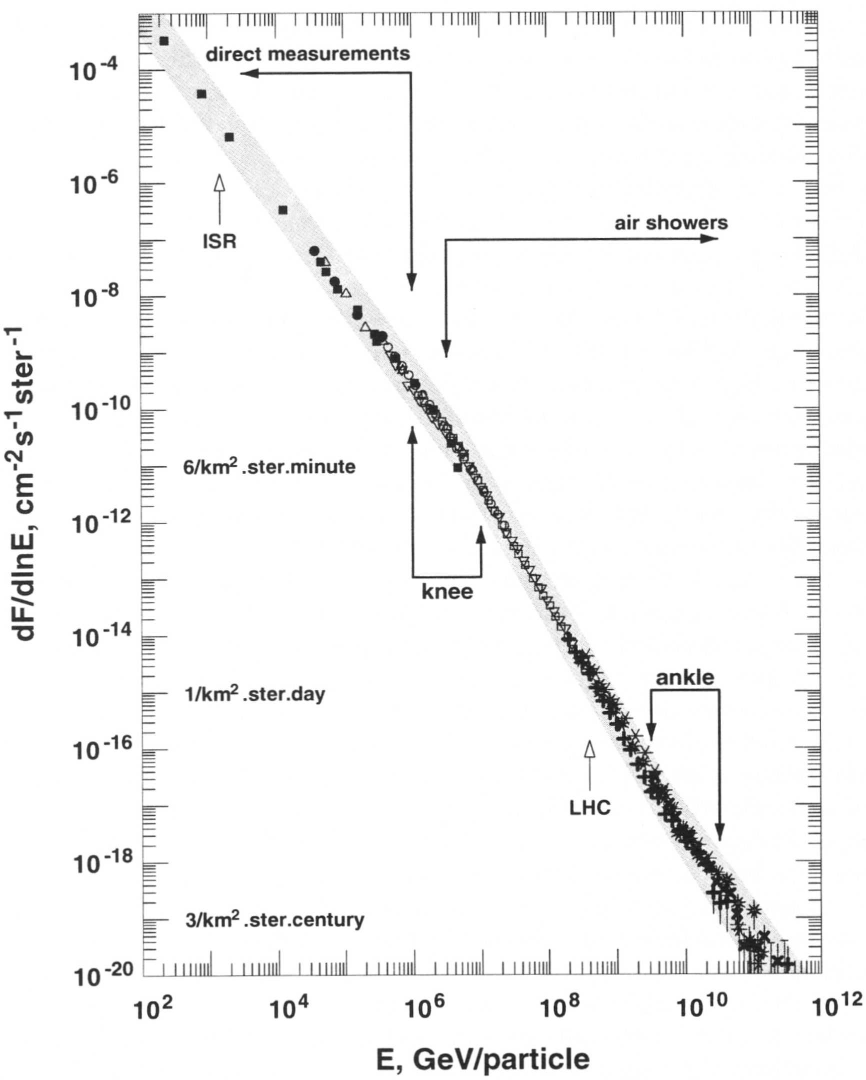

Figure 1: The energy spectrum of cosmic rays. [2]

3 Pierre Auger Observatory

The Pierre Auger Observatory (PAO) is an international observatory build to analyse the nature and origin of the highest energy cosmic rays through the detection of EAS with energies above $10^{18} eV$. The observatory is named after Pierre Victor Auger who studied EAS in the Alps, up to an altitude of $3500m$ in 1938. The PAO is split into two sites, one on the northern hemisphere and another one on the southern. The northern site is planned to be built in south-east Colorado, USA. The southern site is completely constructed and located near the town of Malargüe in Mendoza, Argentina. Its size is 3000 km^2. A map of the observatory is shown in Figure 2.

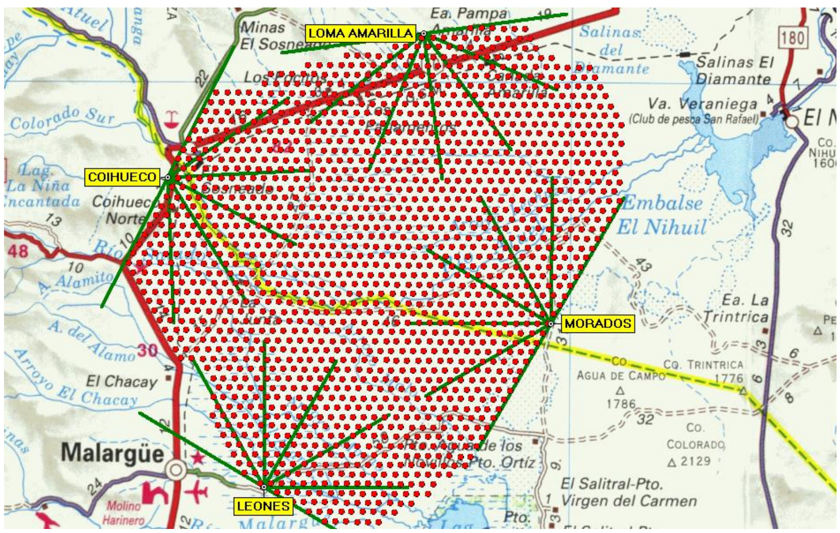

Figure 2: A map of the observatory. Each dot symbolises a surface detector stations, the lines indicate the field of view of the fluorescence detector telescopes. [3]

At the observatory several techniques are used to detect extensive air showers:

- Surface Detector (SD)
  Over the area, 1600 water tanks are distributed in a triangular structure with a distance of $1500m$ between the tanks. Each tank is filled with $12m^3$ water. The incident particles produce Cherenkov radiation which is measured in the tank by 3 photomultipliertubes. If more than three stations are triggered, the signals will be saved. The measured signals are used to reconstruct the shower parameters (energy, arrival direction,...) of the EAS. [4]

- Fluorescence Detector (FD)
  24 fluorescence detectors in total are located in four buildings at the border of the observatory. The light observed originate from nitrogen atoms that get into an excited state if an air shower passes them. By falling back into the ground state, they emit a photon. This emitted light can be measured with the fluorescence telescopes in clear moonless nights. [5]

Together the two detectors, as displayed in Figure 3, form a hybrid detector. With the FD a calorimetric measurement of the energy is possible, which is used to calibrate the SD.

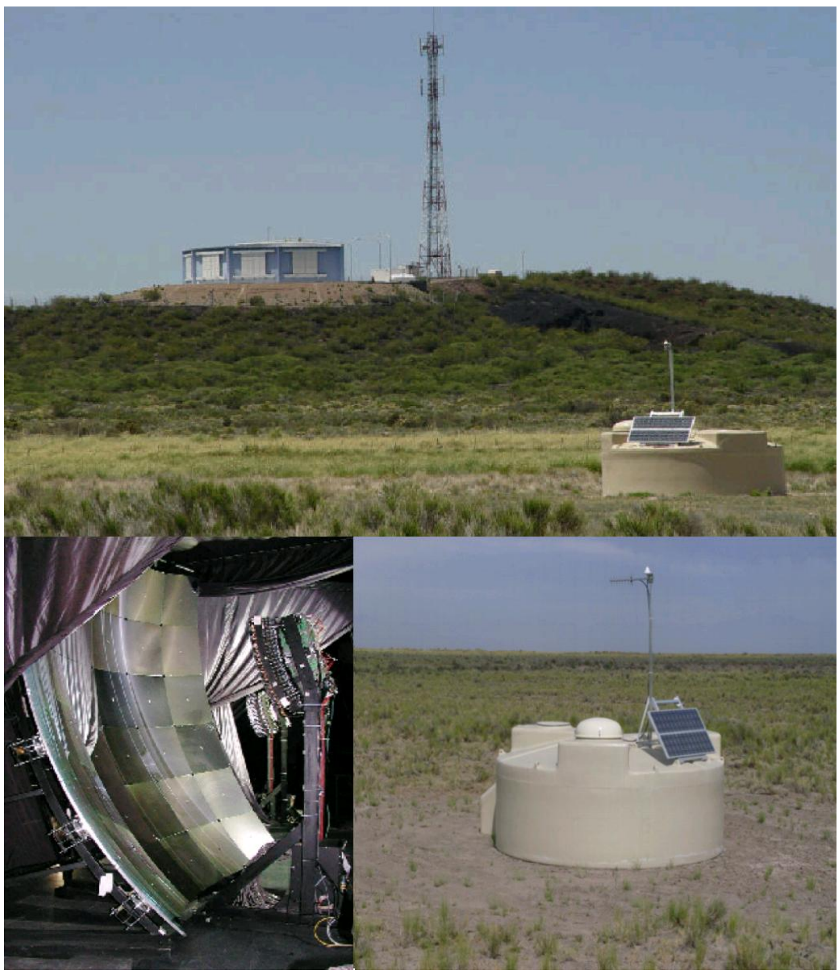

Figure 3: The top picture shows a FD-Station and a surface detector. The picture on the bottom left shows one of the fluorescence telescopes with mirror and camera, and the one on the bottom right a SD-station. [3]

4 Measurements in radio frequencies

4.1 Radio Detector

A different and new method to detect EAS makes use of radio detectors. The EAS emits electromagnetic radiation, as charged particles move in a magnetic field. The radiation is measured with antennas in the radio frequency domain. There are several possibilities to decide if a signal (at one station) is saved or not. At the PAO the decision can be made in two different ways. One is the self trigger method, which means that a signal is saved if the signal shows certain characteristics, like a minimal amplitude. Alternatively an external trigger can be used. The signal is saved if an external instrument, like scintillator panels, trigger. The signals of the different stations are saved if an EAS is measured with two or more stations. For a minimal reconstruction of an EAS are three stations needed. A picture of a radio antenna is shown in Figure 4.

4.2 Radiosignals

The antenna measures the projection of the E-field in the horizontal plane from the EAS in the East-West direction (x) and the North-South direction (y). The signals from the antennas are amplified and filtered. Due to the fact that the optimal frequency range is still unknown, different filters are tested. In the measured data set, signals of EAS have been recorded with two different filter settings, $25-75MHz$ and $50-75MHz$.

Figure 4: A radio antenna with two arms (North-South and East-West) and solar panels.[6]

5 Data

The data set that is used in this analysis consists of simulated signals, that are generated with the MGMR Monte-Carlo program [7]. For each shower, values of the E- field in x, y and z direction are given as function of time, with a time resolution of $1ns$. In order to match the measured data properly these events are changed into a signal with 4000 samples / $10\mu s$. At the PAO the signals are filtered between 25-75 MHz. To obtain a simulated signal in the same frequency domain, a rectangular filter is applied to the simulated data. Furthermore, just like in data, a Hilbert envelope is created, which has only positive amplitudes. The maximum amplitude of a trace can now be determined easier and more precise. A Hilbert envelope is an envelope that is produced with the help of a Hilbert transformation. The amplitude of the Hilbert envelope is the square root of the original amplitude squared plus the Hilbert transformation squared thus

$$Amplitude_{Hilbertenvelope} = \sqrt{Amplitude^2 + Amplitude^2_{Hilberttransformation}}$$

The result can be seen in Figure 5. The signal can be characterised by the position and the amplitude of the maximum. Unless otherwise specified, in the further analysis a signal [8] in the x - direction is used. The signal after passing through a band pass filter is called OS and has a maximum amplitude of $1037.6\mu V/m$ at sample 2001. The signal after passing through a band pass filter and creating a Hilbert envelope is called OSH and shows a maximum amplitude of $1391.8\mu V/m$ at sample 2004.

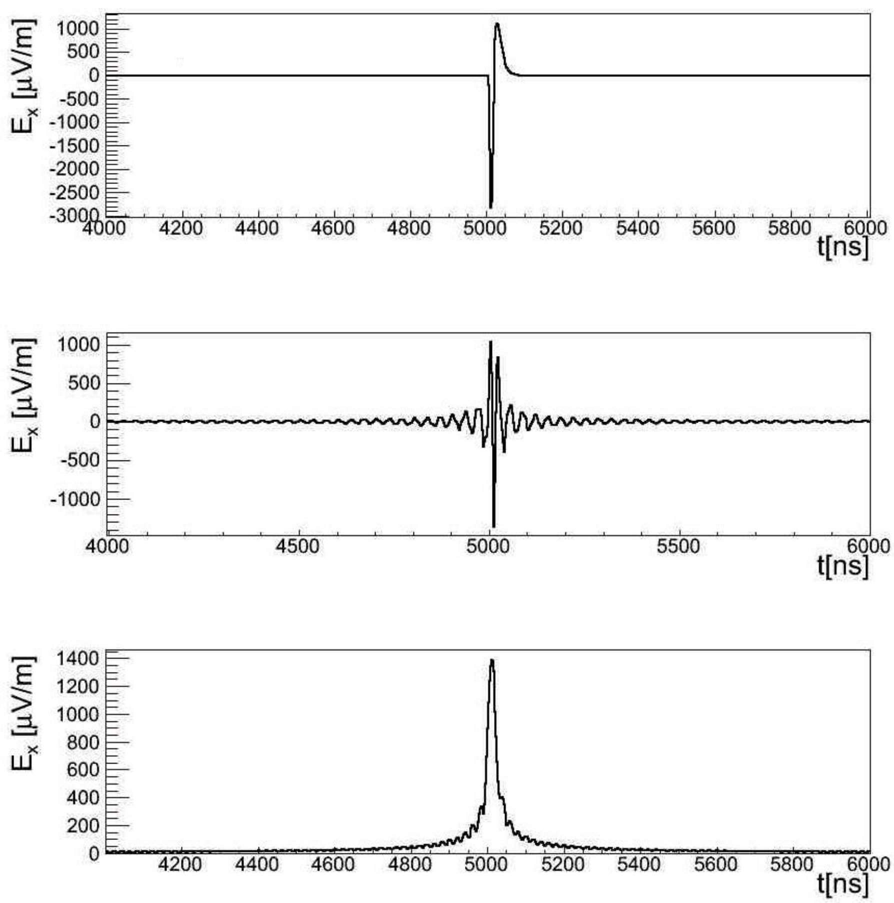

Figure 5: Simulated electric field of a cosmic ray induced air shower without a bandpass filter (top), with a bandpass filter (middle) and the Hilbert envelope of the filtered signal (bottom) in the time domain.

6 Analysis

6.1 Noise generation

Noise can be generated in many ways. It is chosen to simulate noise by first defining a frequency spectrum, and adding a random[1] phase to each frequency. Due to the fact that the radio antennas are sensitive between 25 and $75MHz$, noise is simulated only in this range. The spectrum used is a falling spectrum, that has a linear dependence on the frequency. The effect of the slope of the noise spectrum is studied by modifying the slope parameter, but leaving the total noise level, the area under the spectrum, constant. Also a flat spectrum is used. On all spectra's is a fluctuation of 10% of the amplitude of the flat spectrum. The phase generation (ranging from $-\pi$ to π) is needed in order to be able to perform an inverse Fourier transformation, which in turn transforms the generated noise in the frequency domain into a noise trace in the time domain (see Figure 6).

The amount of noise is characterised by the Root Mean Square (RMS) of the trace in the time domain. This is a measure of the variation of the trace around zero.

6.2 Generated noise added to a signal

Next, noise is added to the signal, having the same band pass filter applied to both of them. For a comparison with the OSH a Hilbert envelope is created for this new series (signal plus noise). There are now two times a signal + noise trace. One time trace contains the Hilbert envelope and one time trace the simulated signal plus noise.
In order to compare different simulations, or simulated data it is useful to determine a Signal to Noise Ratio (SNR) as:

$$SNR = \frac{A}{RMS_{Noise}} \tag{1}$$

with A the maximum amplitude of the signal in the time domain and the RMS of the noise trace in the time domain without Hilbert envelope.
The SNR can be seen as a function of the noise level in Figure 7, where the amplitude is obtained from the Hilbert envelope. The SNR falls off as $\frac{A}{x}$, for a constant signal.

[1]produced with Root r.Uniform

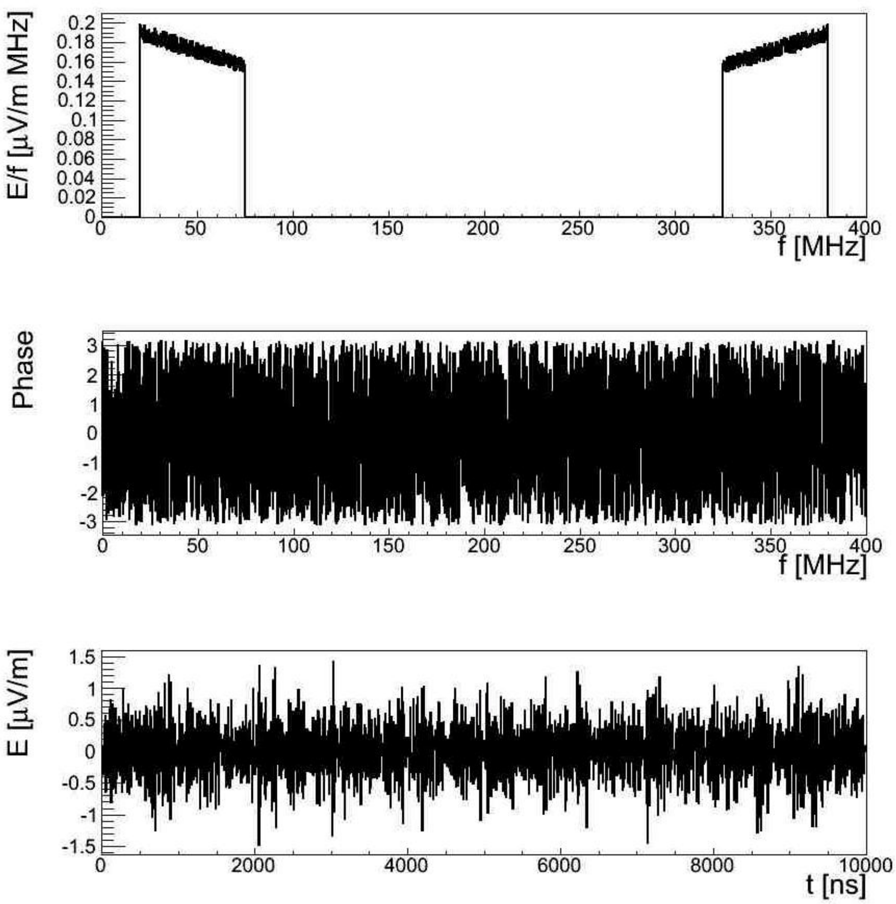

Figure 6: The generated amplitude spectrum can be seen in the top graph (in frequency domain). In the middle the random phase between π and $-\pi$ are displayed. The bottom figure shows the produced noise in the time domain.

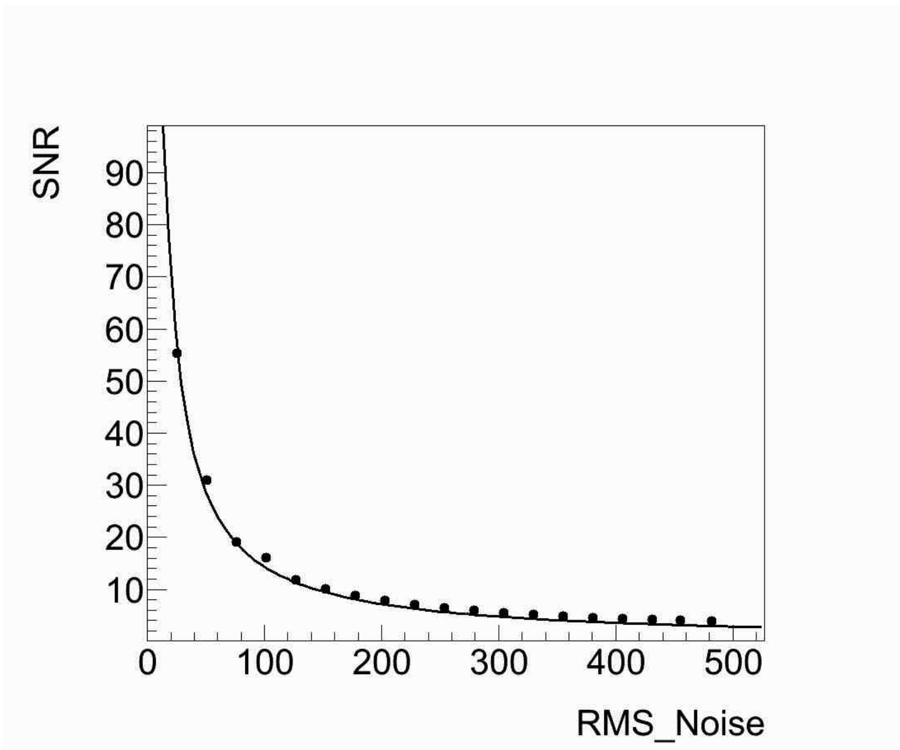

Figure 7: The SNR depending of the noise magnitude. The drawn line shows an A/x fit to the points.

6.3 The variation of the original maximum by adding noise

The original signal is generated with a maximum at sample 2001 and an amplitude of $1037.6\mu V/m$. The amplitudes found at this same sample 2001 after adding noise can be seen in Figure 8. The data is sampled every $2.5ns$ and the signal is added 500 times to a noise trace. The noise level of these traces is such that the corresponding SNR is 33.6. The mean of sample 2001 is found to be $1035\mu V/m$, which is in agreement with the original amplitude. This is expected as noise has a mean amplitude of zero.
From this kind of histograms a mean and a RMS are derived. This analysis is done for different SNR. The result of that analysis is displayed in Figure 9.

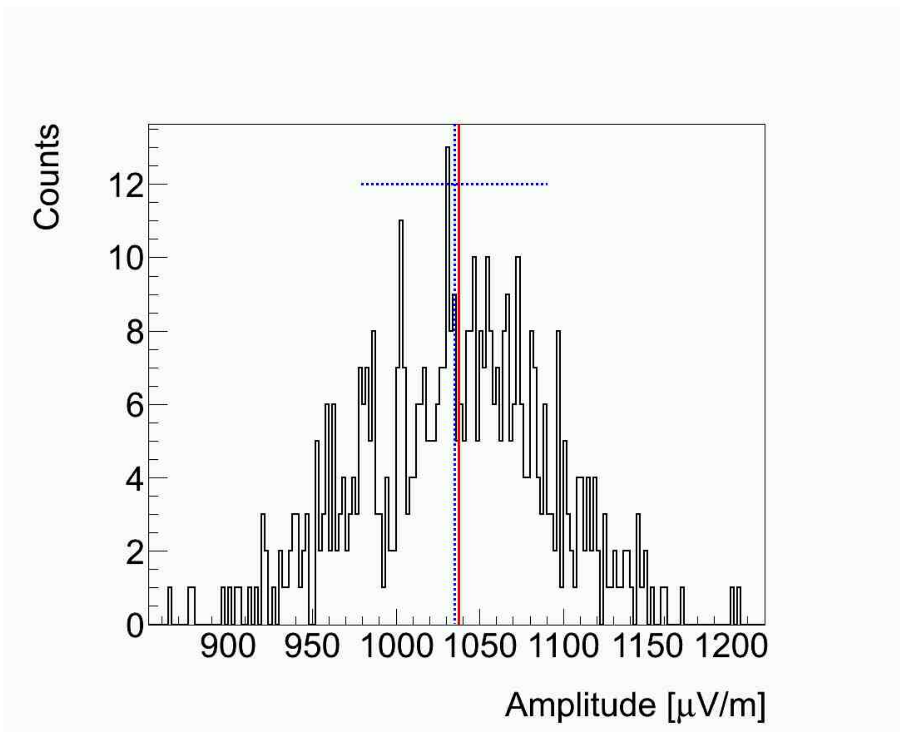

Figure 8: Variation of the amplitude of the maximum of a bandpass filtered radio pulse. Noise with a strength that corresponds to a SNR of 33.6 has been added 500 times to the signal. The (red) line shows the values of the original maximum, the dotted lines the mean and RMS of the distribution.

Furthermore it is possible to make the same plot for the Hilbert envelopes. The maximum amplitude changes to $1391.8.6\mu V/m$ and is at sample 2004. The result is shown in Figure 10. When using an envelope, the amplitude increases for a decreasing SNR. This is a feature of the Hilbert envelope. When calculating

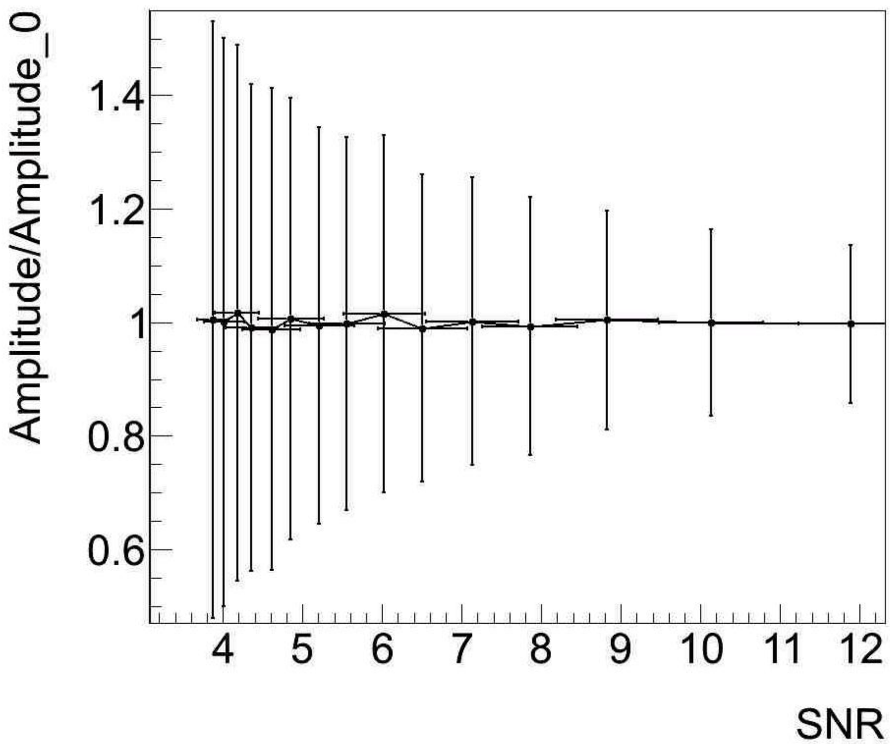

Figure 9: The average amplitude at the time of the original maximum. The amplitude is averaged over 500 iterations and normalised by the amplitude of the OS. The error bar shown is the RMS of the value obtained in these 500 iterations.

the Hilbert envelope the Hilbert transformation of the amplitude is added in quadrature to the square of the original amplitude. This leads in average to a higher amplitude. Which is seen in Figure 10.
Many of the histograms are produced by using a SNR of around 33.6 or 36.6. This SNR does not appear in the comparison given, when SNR is displayed on the x-axis. This is due to the fact that this is not an interesting SNR. However, it serves to illustrate the method used.

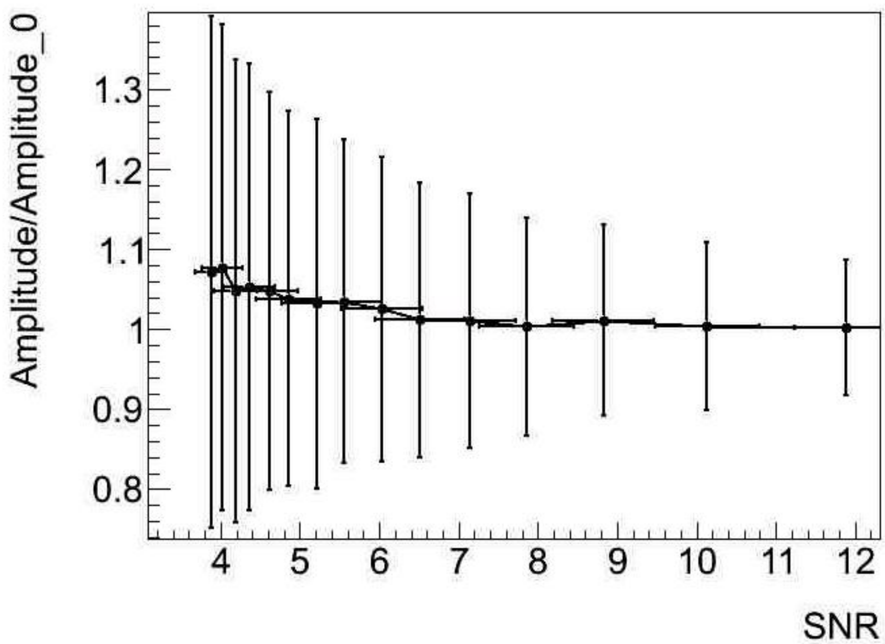

Figure 10: The average amplitude of the Hilbert envelope at the sample from the originally maximum is shown. The amplitude is averaged over 500 iterations and normalised to the Hilbert envelope of the original maximum amplitude.

6.4 The effect of noise on the maximum

The previous analysis shows the effects of noise on a specific sample. When locating a signal, one need to find the maximum of a trace. Therefore it is important to know how the noise influences the complete trace. This is studied in the following analysis.
After adding noise to a time series, the amplitude and position of the maximum is determined. The variation of the position and amplitude of the maximum can be seen in Figure 11 for the raw time series and in Figure 12 for the Hilbert envelopes. When creating these histograms the signal is added to 500 different noise traces with a SNR round 33.6 for a signal without envelope and a SNR round 36.6 for a signal with envelope. The whole time series, 4000 samples, are scanned for the maximum.
The upper histogram of Figure 11 shows the amplitude distribution. The average amplitude is around the original maximum, at $1037.6\mu V/m$. The amplitude distribution follows a Gaussian shape. When amplitudes are bigger than the original, most likely a positive number is added to the old maximum or to a sample beside the maximum, therefore the amplitude of this samples increases. When a negative noise amplitude is added to the largest (signal) amplitudes a maximum less than the original may occur. From this histogram the mean and the RMS are determined. The mean amplitude is $1036\mu V/m$ and the RMS $54.64\mu V/m$. The difference between the mean of the distribution and the original maximum is $1.6\mu V/m$.
The bottom histogram shows that most maxima's are found at the original position, sample 2001. There are some maxima's next to the original sample. Again the mean and the RMS of the histogram are calculated. On average the maximum is found at sample 2002, with a variation of 0.717, less than one sample.

The distribution of envelopes, in Figure 12, shows that most maximum values are found around $1400\mu V/m$, when the maximum of the original signal is at $1391.8\mu V/m$. Again the mean and the RMS of the distribution are determined. The mean is at $1422\mu V/m$ and the variation is $34.19\mu V/m$. The difference between mean and original value is larger than before, which is due to the fact that the envelope increases the signal (see section 5). The mean is displaced to higher amplitudes.
From the bottom histogram it can be seen that most maxima's are found at the original position, sample 2004. There are many maxima's directly besides sample 2004, at sample 2005. From this histogram the mean and the RMS are also calculated. On average the maximum is found at sample 2005, with a variation of 0.765, less than one sample.

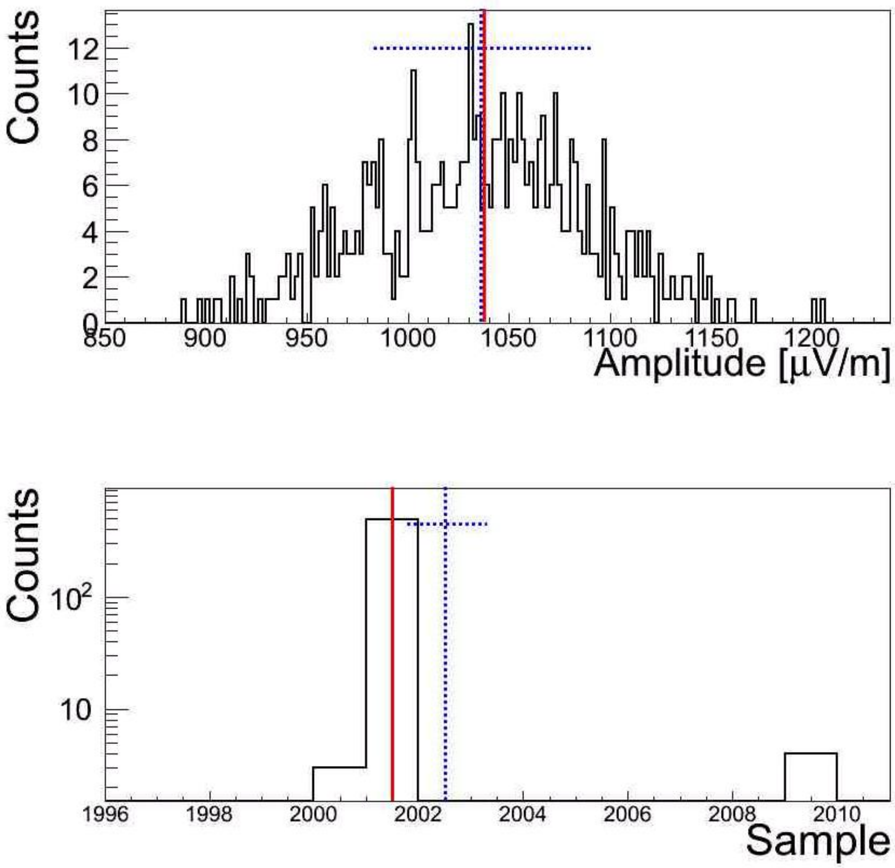

Figure 11: The amplitude and the position of the maximum from 500 simulations, with a SNR round 33.6 by scanning the whole trace. The amplitude of the original maximum is $1037.6\mu V/m$ at sample 2001 (line). The mean and RMS are symbolised by the dotted line.

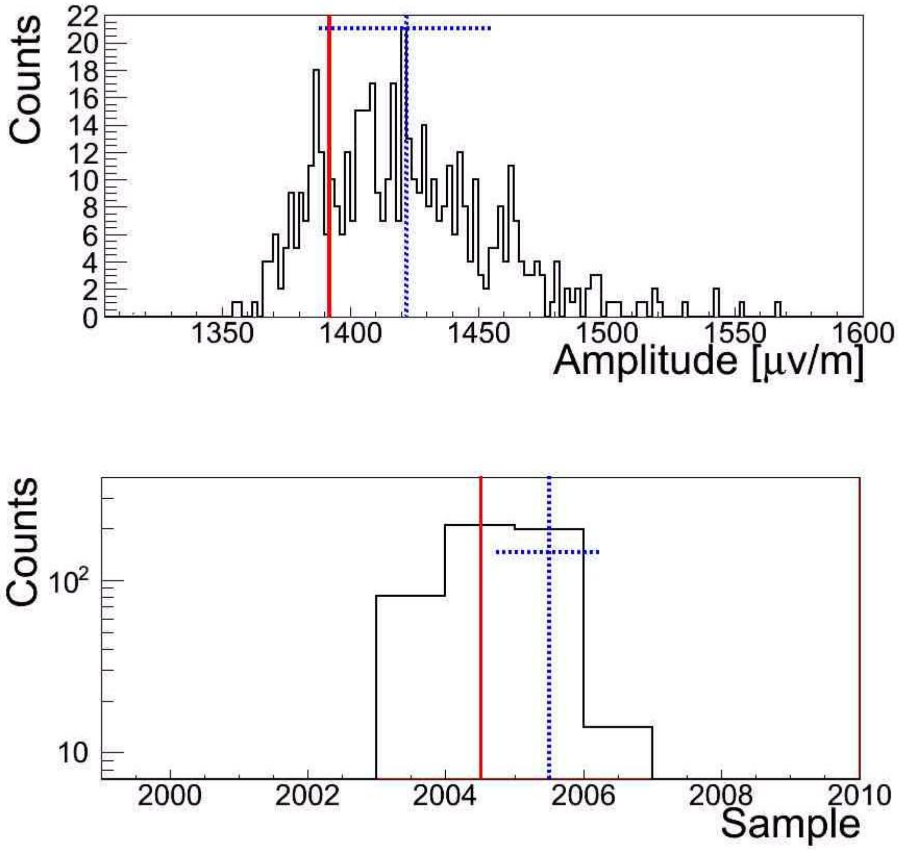

Figure 12: The amplitude and the position of the maximum of the Hilbert envelopes from 500 simulations, with a SNR round 36.6 by scanning the whole trace. The amplitude of the original maximum is $1391.8\mu V/m$ at sample 2004 (line). The mean and RMS of the distribution are symbolised by the dotted line.

6.5 The average and variation of the maximum as a function of SNR

From histograms like the one shown in Figure 12 the mean and the RMS have been determined. These values are now plotted against the different SNR-values in Figure 13 and 14. For these graphs a time window, in which the maximum is determined, is defined to lie between sample 1500-2500.

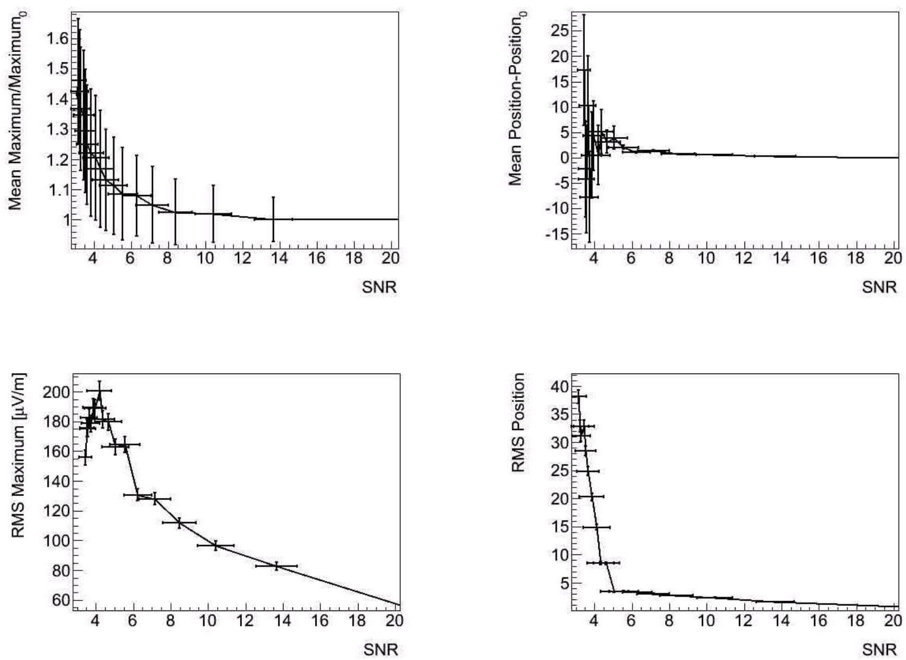

Figure 13: The mean and RMS from the distribution of maximum amplitudes and their positions from 500 simulations per SNR. The trace is scanned from sample 1500-2500.

In the diagrams on the top left, the mean of the maximal distribution divided by the original maximum can be seen. If the noise increases (SNR decreases), the mean of the maximum increases too. This originates from the fact that it is searched for a maximum within 1000 samples, when the noise is relatively high compared to the signal, there is a large probability that at least one of the samples shows a higher amplitude than the original one. On average the signals are reconstructed with an amplitude that is to large.
In the diagrams on the bottom left the variation of the maximum amplitude can be seen. At low noise (high SNR) values the RMS increases for increasing noise, but for a SNR < 5 the RMS decreases. The uncertainly on the variations is small, only a few $\mu V/m$.
In the top right diagram the average distance between the position of the max-

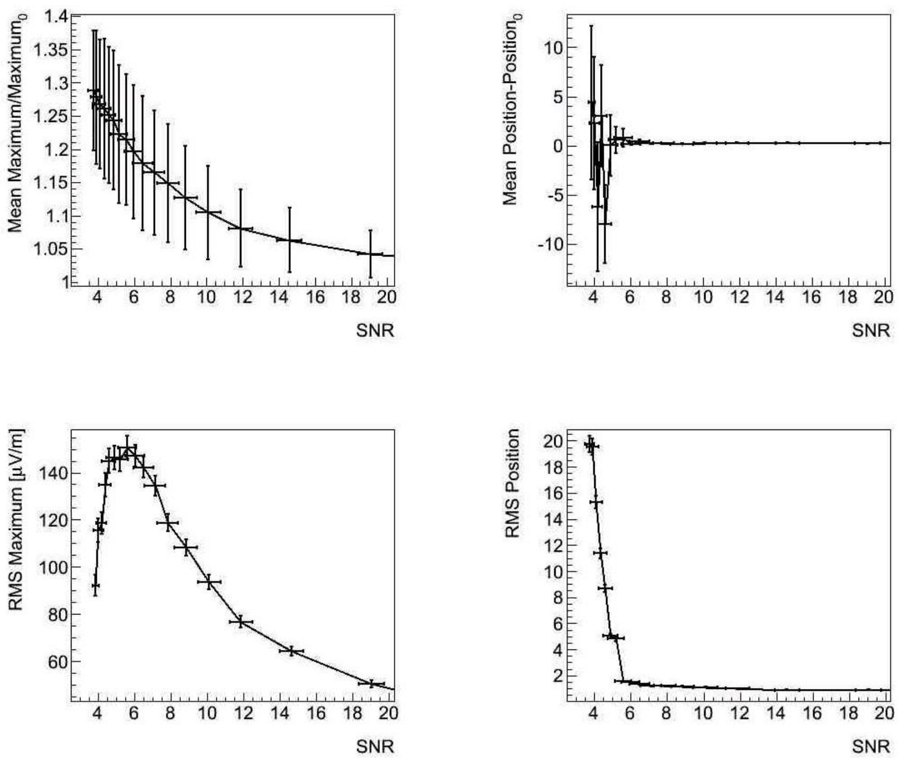

Figure 14: The mean and RMS from the distribution of the maximum amplitude and their positions. Here the Hilbert envelopes are used. (500 simulations per SNR, search window 1000 samples around original maximum).

imum with and without noise is shown. The average distance is almost around zero, however the error bars are high for SNR smaller than 6. For high noise values, there is a good change that the noise is higher than the original amplitude somewhere in the search window. This causes the position of the maximum to be almost evenly distributed within the search window. As this window is placed around the original sample, the mean location of the signal does not change. Furthermore, the procedure of picking the maximum random noise leads to a smaller variation in the amplitude. In short a decreasing RMS of the amplitudes and an increasing RMS of the position can be expected. The location of the maximum position is more often far beside the original position.
This uncertainty can also be seen in the bottom right diagram, the uncertainty is almost zero for SNR bigger than 6, however for a SNR <6 uncertainty increases. For proper signal timing a SNR > 6 is necessary.
The same procedure has been applied to the Hilbert envelopes (see Figure 14). An enhancement of the maximal values is found for large noise values. Again

the variations on this maximum are smaller than found in the times traces.
A further study is performed on the results of the maximal value of the Hilbert envelopes.
Figure 15 shows a diagram of the ratio of the mean maximum found when noise is added to the original trace (top left diagram of Figure 14. The function $A(SNR) = 1 + p0 * e^{-p1*SNR}$ is fitted.

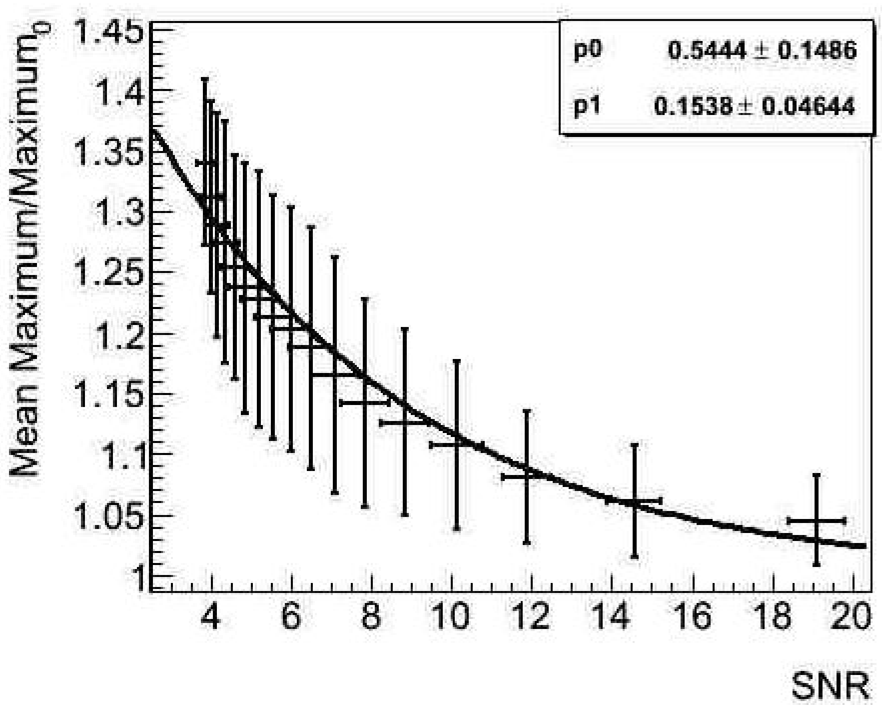

Figure 15: The mean over the original amplitude with fit. The function used is $A(SNR) = 1 + p0 * e^{-p1*SNR}$.

6.6 Varying the size of the searching window

In the previous section the trace is scanned between samples 1500-2500 to find the position and amplitude of the maximum. If the position is well known it is possible to use a smaller window size. An example of this based, on an analysis of the Hilbert envelopes, can be seen in Figure 16. Here, the trace is scanned for 100 points around the original maximum position.
The differences can be seen in the RMS and not in the mean. The diagram on the left bottom 16 shows that the RMS does not decrease to a value below $100\mu V/m$. The reason for this is the fact that now only a few samples are scanned so that the probability of the noise amplitude exceeding the signal is much smaller. This results in a smaller decrease of the variation at high noise levels. From the bottom right diagram it can be seen that the uncertainly on the position is below one for a SNR of 6 or higher, for smaller SNR the RMS grows to 10, which implies that the uncertainly is 10% of the scanned window size. The RMS reflects the fact that the window size is smaller. Naturally, the search window has to reflect the initial uncertainty on the position of the signal. The better this is known, the more reliable the results will be.
A diagram with more window sizes in displayed in Figure 17. It can be seen that the variation (RMS)of the position decreases for smaller search window sizes. The window size has been varied between 100 to 1000 samples around the original maximum.

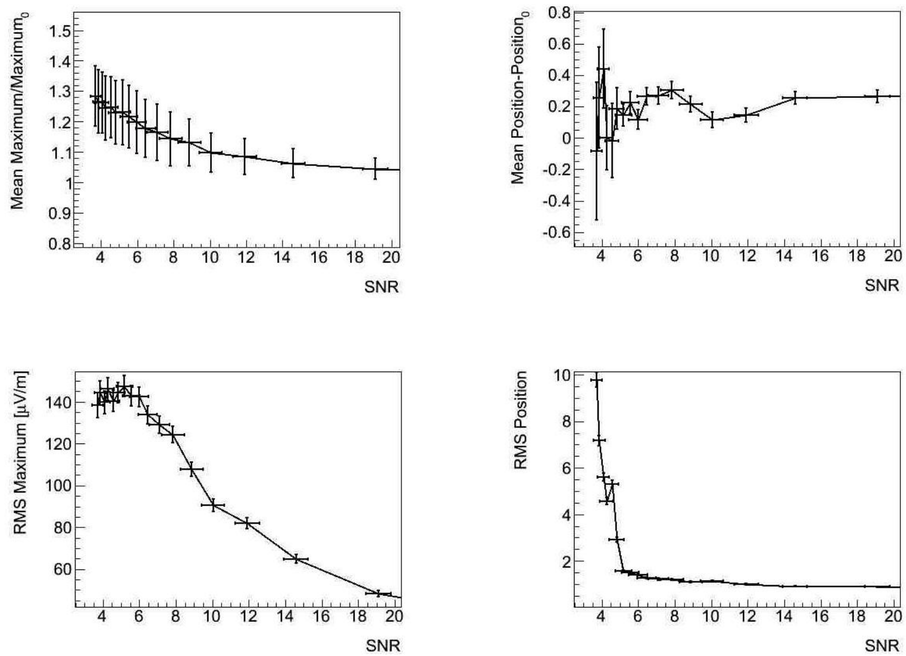

Figure 16: The mean and RMS from the distribution of the found maximum amplitudes and their positions. The Hilbert envelopes are used. (500 simulations per SNR, search window 100 samples around original maximum).

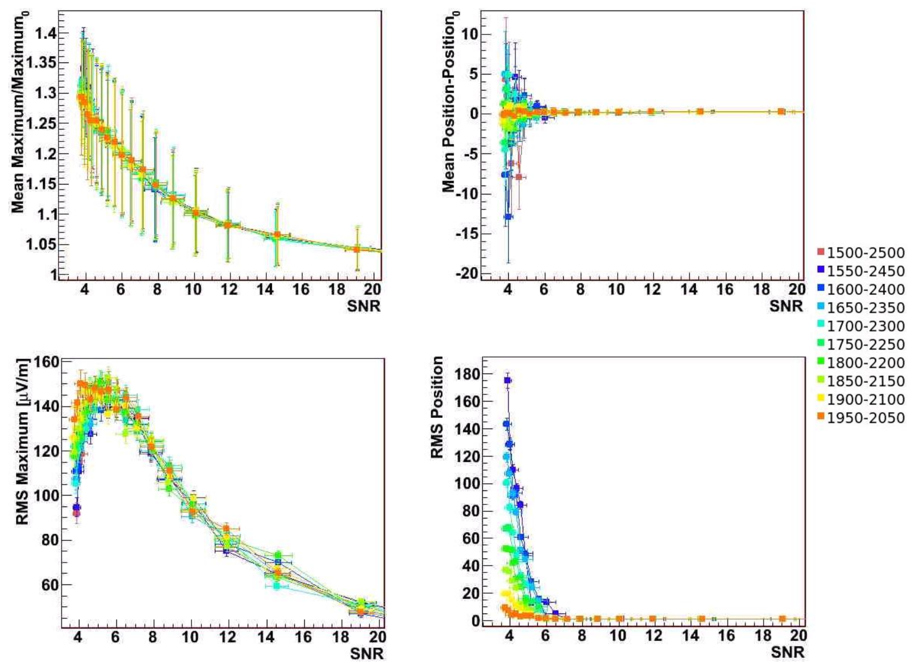

Figure 17: The dependence of the Mean and RMS of the maximum amplitude and the position. Each colour symbolises a different window size. Here the Hilbert envelopes are used.

6.7 Different noise generation

In section 5.1 it is explained how the noise is generated. It is possible to change the slope of the spectrum. It is chosen to investigate 6 different spectra's. A spectrum with a constant amplitude A (no slope) and spectra's with a slope of:

$$s = \frac{x * A}{f}$$

with A the amplitude of the constant spectrum, f the frequency difference and x equal 0.2, 0.4, 0.6, 0.8 or 1. In the previous analysis a slope with x=0.2 is used.
The spectrum's are constructed such that the area is constant, thus creating a constant noise level. The results of the analysis are shown in Figure 18.
It can be seen, that the different spectral shapes hardly influence the results. The differences in the results are too small to be seen in this scale and are certainly all within the uncertainties. The plot it self looks like the result in Figure 14. It is clearly visible that only the power of the noise is important. In other words, a constant power/area yields to noise with the same power/RMS.

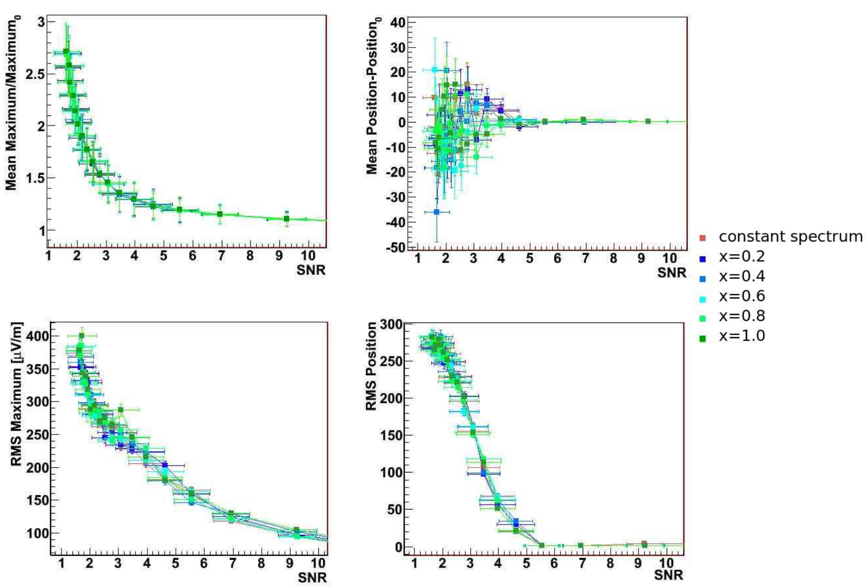

Figure 18: The dependence of the mean and RMS of the amplitude and position for different noise spectra by the creation from noise. Each colour represents an other slope.

6.8 Arrival direction reconstruction

The results of the analyses presented so far is used to determine the uncertainty on the determination of the arrival direction of cosmic rays due to noise in the measurement. If the measured or simulated data from three or more stations is available, the arrival direction of the EAS can be reconstructed. For such a reconstruction the arrival time of the EAS measured at each station, which is obtained from the position of the signal in the time trace or the location of the maximum amplitude, is necessary. With this data it is possible to reconstruct the zenith and azimuth angle, which are defined in Figure 19.
The reconstruction of the arrival direction of the simulations without noise is compared to the same reconstruction of a simulation with noise added. The amount of noise added corresponds to a SNR of 6. The differences in the obtained zenith and azimuth angles are shown in Figure 20. The top histogram shows a distribution of the opening angle. This angle can only be positive and has a ranges from $0° - 4°$. The middle histogram shows the difference between the zenith angles. The average differences is zero and the maximum deviation found is less than five degrees. The azimuth angle shows a similar effect, where the maximum deviation is four degrees. When noise is added to a signal with a SNR of about six, the uncertainty on the arrival directions is in the same order of magnitude as the uncertainty form the SD-reconstruction (about three degrees) [1].

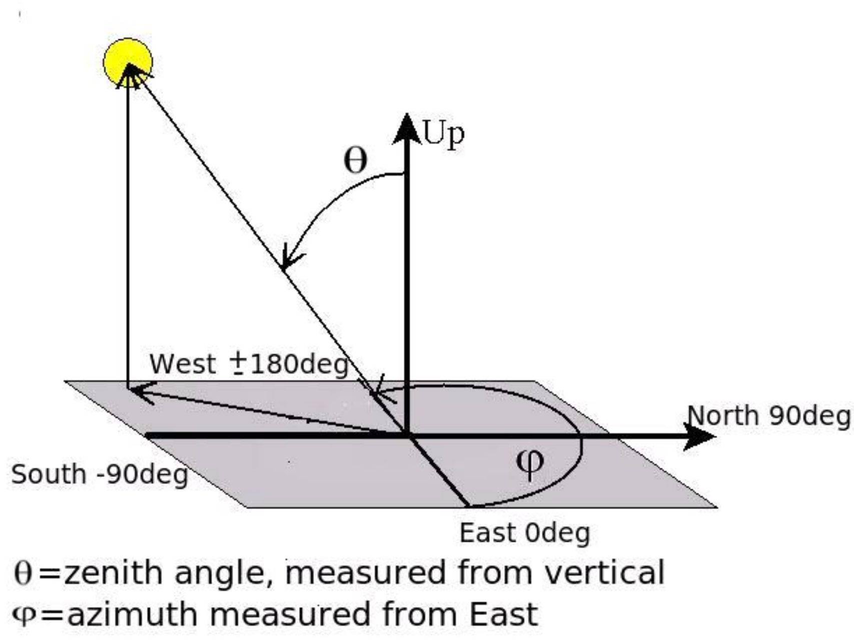

Figure 19: θ =Zenith angle measured from vertical and ϕ = Azimuth angle measured anticlockwise from the east [9].

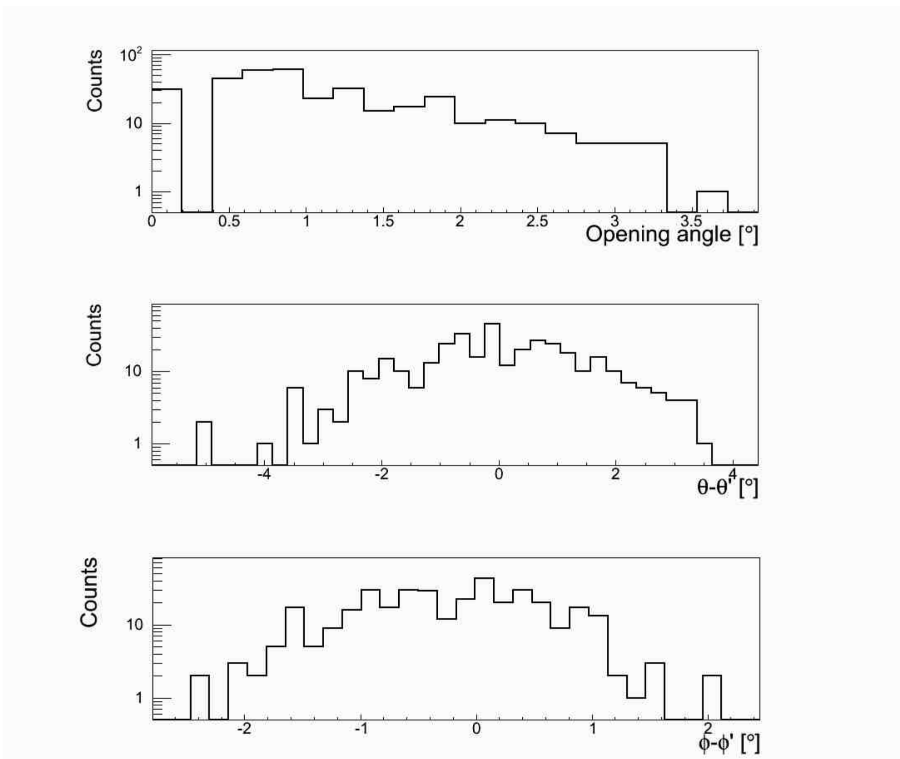

Figure 20: The opening angle (top plot), and the difference in θ (zenith) (middle plot) and ϕ (azimuth) (bottom) angle from a reconstruction with and without noise. The signals from the three stations are added 500 times to noise with SNR=6.

7 Conclusion

In this thesis the effect of adding noise on position and amplitude of a signal is analysed for different signal to noise ratios. There a difference when the analysis is performed on the time traces with or without Hilbert envelope. The analysis on the envelope leads to higher reconstructed maximum amplitude and it is necessary to correct for this effect, see Figure 14, Figure 16 and the fit in Figure 15. The reconstruction on the envelope has a smaller uncertainty, which means that the variation in the results obtained is smaller. Therefore, a study of the different search window is only performed on the Hilbert envelopes. The same holds for the change in the noise generation and the angular reconstruction.

For a SNR of 6, noise adds on average about 25% of the maximum signal amplitude. For lower SNR this effect is even lager. For a SNR of 10 the average amplitude is 10% higher. This increase in amplitude makes a correction essential.

To find the correct position in time of the maximum a SNR of 6 or more is required as can be seen in Figure 14. The uncertainty of the position increases for lower SNR.

A change of the spectral index, makes nearly no difference, as long as the total noise power remains constant, see Figure 18.

The noise influences the arrival direction with an uncertainly of $4°$, at a SNR of 6. This is comparable with the reconstruction of the arrival direction measurements with surface detectors.

References

[1] A.A.Watson "Cosmic rays of the highest energies", Contemporary Physics, 2002, volume 43, number 3, pages 181-195

[2] T.Stanev Springer 2004, 2nd ed., High energy cosmic rays, page 10.

[3] Pictures from www.auger.org

[4] I. Allekotte *et al.* [Pierre Auger Collaboration], "The Surface Detector System of the Pierre Auger Observatory," Nucl. Instrum. Meth. A **586** (2008) 409 [arXiv:0712.2832 [astro-ph]].

[5] J. A. Abraham *et al.* [The Pierre Auger Collaboration], "The Fluorescence Detector of the Pierre Auger Observatory," arXiv:0907.4282

[6] Picture from Stefan Grebe

[7] K. D. de Vries, O. Scholten and K. Werner, "Modeling coherent geomagnetic radiation from cosmic ray induced air showers", Proceedings of the 31st ICRC, Łódź 2009

[8] MGRM-3388350.mgrm; pole1

[9] http://www.esrl.noaa.gov/gmd/grad/solcalc/azelzen.gif with changes

A Abbreviations

- EAS - Extensive Air Shower
- FD - Fluorescence Detector
- OS - Original Signal (signal without noise) without Hilbert envelope
- OSH - Hilbert envelope of the Original Signal
- PAO - Pierre Auger Observatory
- RMS - Root Mean Square
- SD - Surface Detector
- SNR - Signal to Noise Ratio