# Influence of signal fluctuations on the determination of the core position for the Auger infill array

Tijs Karskens
Centro Atómico Bariloche, Argentina and
Radboud Universiteit Nijmegen, the Netherlands

May 30, 2012

**Abstract**

This study takes a closer look at the precision of the reconstruction of events recorded by the infill array at the Pierre Auger Observatory. The goal is to evaluate the influence of signal fluctuations in individual detectors using the CDAS reconstruction algorithm, on uncertainties in core position, energy and direction as well as possible reconstruction biases. An example of the latter might be that the core of an event is pushed away from the nearest detector - or the other way around. The main conclusion will be that the uncertainties provided by the CDAS reconstruction algorithm are incomplete, and an additional uncertainty on the core position for both $X-$ and $Y-$ coordinate of up to 60 meters should be added.

# 1 Introduction

## 1.1 Introduction

At every moment of every day, the Earth is bombarded by small particles coming from outer space. These particles are mostly protons, but can also be larger nuclei. Upon reaching our atmosphere, these particles will strike molecules in the atmosphere, and thus create a shower of secondary, lower energy particles. The Pierre Auger Observatory detects showers at the Earth's surface, using surface detectors. Surface detectors rely on the Cherenkov effect to detect incoming particles. This effect can be explained as follows: In water, light travels at roughly 75% of its speed in vacuum. When a particle passes through water at a speed greater than this, it produces a cone of light [1]. The surface detectors measure the amount of Cherenkov-light and convert this into a number of particles passing through the detector.

## 1.2 The Array

The entire Pierre Auger Observatory, located in western Argentina, spans a surface area of about 3000 km$^2$. There are roughly 1660 detectors in total. The main area, called the regular array, has detectors positioned in a triangular grid at distances of 1500 meters. In a certain small area of 23.4 km$^2$ detectors are located at a 750 meter distance. This is called the Infill array.

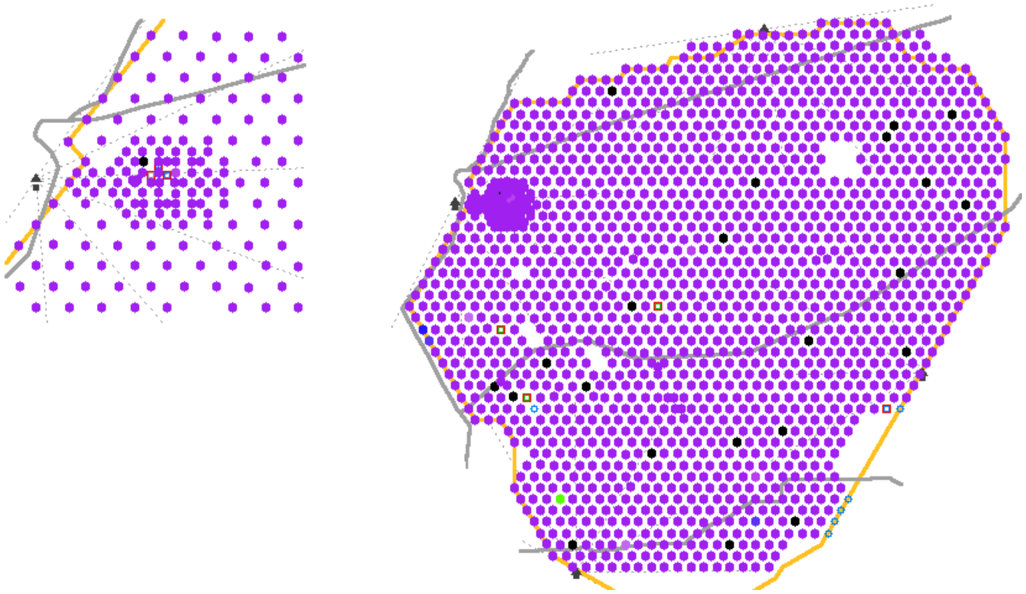


Figure 1: Array of the surface detectors.
Left the infill array, right the regular array.

### 1.3 Event detection

When a particle hits our atmosphere, a shower of particles of lower energy is created. The shower expands perpendicular to its axis in the atmosphere, as secondary particles strike molecules, by which they create tertiary particles, and so on. Upon reaching the Earth's surface there is an area in which the particle density is high enough that it can be measured with our setup. The size of this area depends mainly on the energy of the original particle, and the zenith angle of the incoming shower. Particles with higher energy result in a larger shower size, and vice versa. As particles pass trough a detector, three Photomultiplier Tubes (PMT's) measure the light emitted through Cherenkov radiation. The detectors involved in the event are classified, depending on the distance to the detector with the largest signal. The detector with the largest signal gets labeled 'hottest'. Detectors 750 m (1500 m for the regular array) away from that are labeled 'first crown'. Detectors at 1300 m (2600 m for the regular array) are the second crown. All detectors further away are in the third crown. To analyze the incoming cosmic rays, data gathered from detectors triggered by the event are combined. The CDAS reconstruction algorithm [2] is a computer program that uses this combined data to reconstruct values such as the position, direction and energy of the shower.

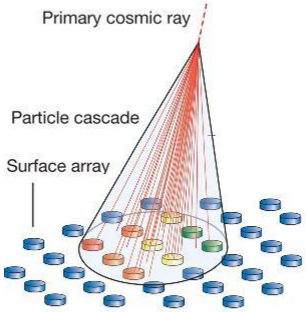


Figure 2: Detection of a shower by the surface detector array

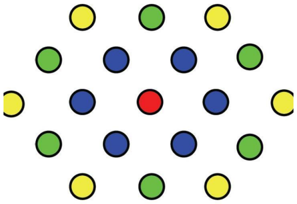


Figure 3: Layout of the array. Hottest detector in red, $1^{st}$ crown in blue, $2^{nd}$ crown in green and $3^{rd}$ crown in yellow

# 2 Method

In this note, the effect of the measurement uncertainty on determining the core position of events is analyzed. This is done in two steps. In the first step a few key values of events are analyzed. The second step will look at fluctuations in the calculated core position and energy by varying the signals detected in the infill array. Before making calculations, a few assumptions have been made. The first is that all detectors, and events, are at the same altitude; no fluctuations in the surface nor curvature of the Earth's surface is taken into account [3]. Furthermore, biases caused by the border of the array are neglected [4]. This is important for events close to the border, where part of the shower remains undetected.

## 2.1 Phase one

First, each event is reconstructed with CDAS. Second, the distance from the shower to a detector is calculated for every detector involved in the reconstruction. The distance calculation is done twice: once ignoring possible angles at which the shower arrives, and therefore only calculating the distances from detector to shower core on the ground. From now on this will be called the 2D distance. Afterwards, the angles are taken into account, and so the perpendicular (3D) distance is calculated. These results are compared with simulations using approximately 1.000.000 randomly generated events. In these simulations the arrival direction (azimuthal angle $\phi$ and zenith angle $\theta$) and coordinates are drawn within the allowed values. Figure 4 gives an overview of the position of the detectors and the allowed area for the simulated events. During this phase, possible reconstruction biases might be detected. Besides needing to know of any reconstruction biases in the next step, it might also supply possible data selection criteria that can be used to avoid biases.

The variables are distributed as follows:

- $0° \leq \theta \leq 60°$
- $P(\theta) = \frac{\cos(\theta)\cdot\sin(\theta)}{\int\cos(\theta)\cdot\sin(\theta)d\theta}$
- $-180° \leq \phi \leq 180°$
- $-375 \text{ m} \leq x \leq 375 \text{ m}$
- $-650 \text{ m} \leq y \leq 650 \text{ m}$
- $\phi, x, y$ are distributed uniformly within their intervals

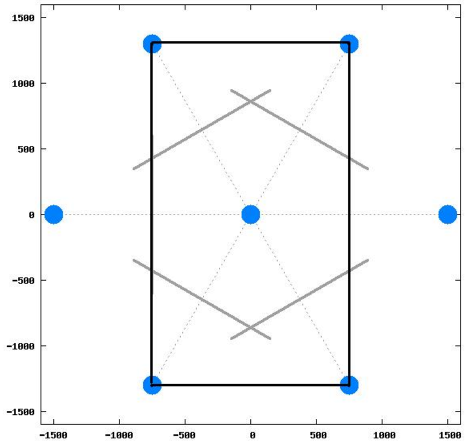


Figure 4: "First Brillouin zone" and detectors surrounding it. The black outlining is the area where simulated events are randomly distributed

## 2.2 Phase two

In a second part, the signal strength detected by the Cherenkov detectors is varied within their uncertainty range. This is done as follows: First, an event is reconstructed to find original values such as the core position, energy, angles, and signals in all the detectors. Next, the signal measured by all involved detectors is modified. Fluctuations are taken randomly from a Gaussian distribution, with as mean the original signal, and as standard deviation $\sigma(Signal) = 1.06 \cdot \sqrt{Signal}$ [5]. However, the PMT's in detectors have a certain threshold, which can be as low as 0.7 VEM (Vertical Equivalent Muon). If a detector measures a signal below this value, it is not flagged as triggered. For that reason, signals fluctuating below 1 VEM are filtered out and redone. To have a fair distribution of signals and avoid a systematic energyshift as much as possible, only events with initial signal (the original signal, before modification) $\geq$ 2 VEM for all triggered detectors are reconstructed. This procedure is repeated until a total of 650 good simulations are created for each recorded event. These altered reconstructions are compared with the original one to see differences in location, direction and energy. A similar analysis has been presented in GAP 2008-048 [6] for the regular array.

# 3 Used data

## 3.1 Data

This study used version v4r8 of the CDAS reconstruction software, all with standard reconstruction settings ("eGeomOptimal", LDF with optimal $\beta$). For phase one, all of the infill events detected by the Pierre Auger Observatory between the 01-01-2008 and 31-12-2010 are used. Phase two uses all data between 01-04-2010 and 31-12-2010. Both phase one and phase two use only events with a zenith angle $\theta \leq 60^\circ$. This is because the CDAS software is not configured for events with a larger zenith angle.

## 3.2 Quality cuts

In any and all reconstructions several quality cuts are imposed; not every signal will reconstruct. For example background noise can trigger a single detector at any time. One important constraint is that an event is only reconstructed when it triggers at least three adjacent stations. That is because directions and location are impossible to reconstruct with fewer triggered detectors.

- The ICRCT5 cut requires the event to fall within a triangle of active detectors, and that 5 of the 6 detectors directly surrounding the hottest (5 out of 6 of the first crown) are active.
- The strict-T5 cut means that all six detectors in the first crown are active.

| Cut | Description | Number of events |
|---|---|---|
| 0 | All reconstructed events are included | 269898 |
| 1 | Events with energy $\leq 0.3$ EeV are filtered out | 8106 |
| 2 | Only strict-T5 events are taken into account | 6094 |

Table 1: Overview of the used cuts in phase one, along with the amount of events left after that particular cut. Every cut includes all previous cuts.

# 4 Results

## 4.1 Phase one

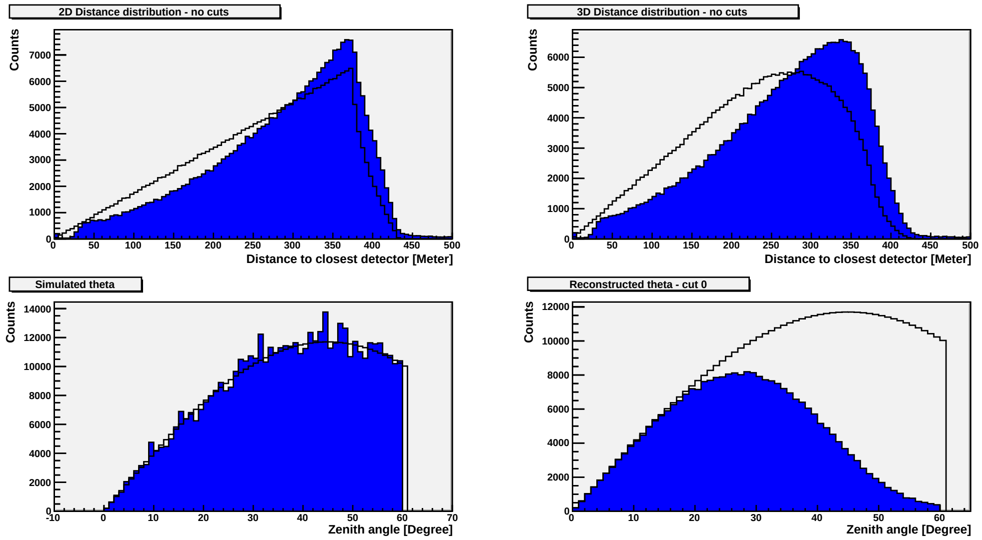


Figure 5: Reconstructed distributions - no cuts.
Top: Distance of the event to the nearest detector for 2D and 3D (blue), compared to simulations (black).
Bottom: Simulated theta (left, blue) and Reconstructed theta (right, blue) compared to the expected distribution (black).

In contrast to the reconstructions, the simulations do comply with expectations for the actual reconstructions. Position and $\phi$ are flatly distributed within their allowed values, and the generated $\theta$ distribution follows a $\cos(\theta) \cdot \sin(\theta)$. The distribution of the shortest distance from shower to a detector is as expected (the size of a ring with radius r increases linearly with r in 2 dimensions) and the result is similar for the 3 dimensional reconstructions. But quite clearly, there are differences between simulations and actual results. Events at the lowest distances appear to be heaped up at a slightly larger distance, in both the 2D and 3D reconstructions. Also, in contrast to the expectations, the number of events does not increase linearly as a function of distance. The graph for $\theta$ consistently does not have a $\cos(\theta) \cdot \sin(\theta)$ shape (Figure 5). As a comparison, official reconstruction values have been used as well. These give a similar distribution. There are several explanations for this:

1. Events are only reconstructed when 3 or more detectors are triggered. That means that events of the lowest energy detected will always be at a distance about equal to all three detectors in its active triangle. Higher zenith angles give a lower chance to trigger 3 detectors.
2. Reconstruction of events with saturated detectors will compensate for saturation. It is however unclear how much exactly, so under- or over-compenstation could exist.

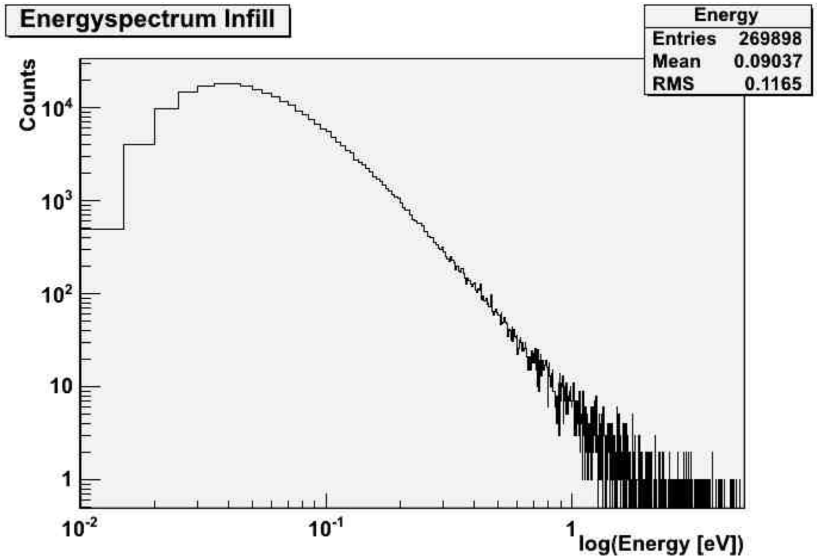


Figure 6: Detected energyspectrum for the infill array, using a strict-T5 cut.

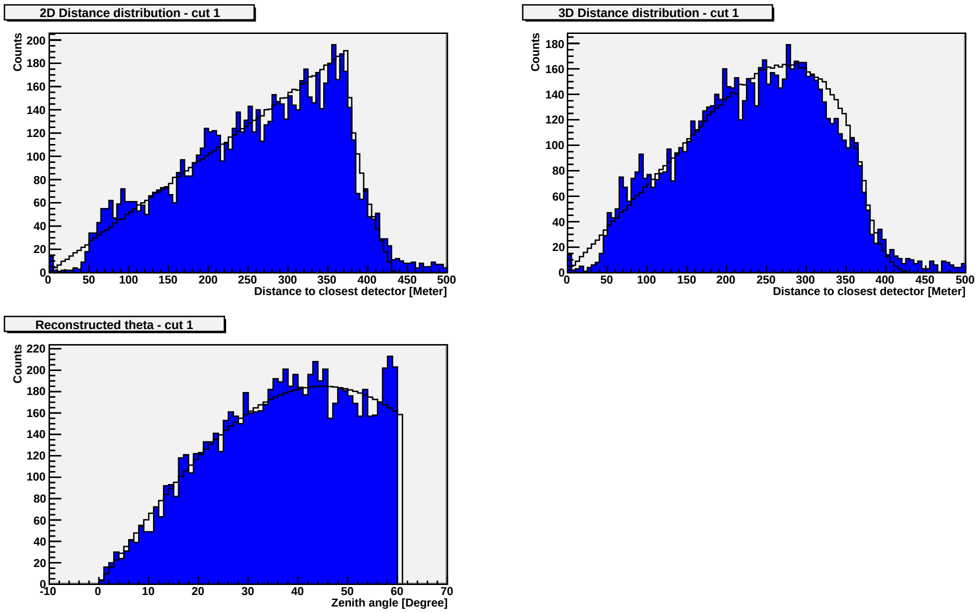


Figure 7: Reconstructions with E $\geq$ 0.3 EeV. Distribution of 2D distance (top left), 3D distance (top right) and theta (bottom left). Reconstructed values are in blue, compared to simulations in black.

Full acceptance of the infill array starts at an energy of roughly 0.2 EeV[1] ([7] and figure 6), so implementing a safe energy cut of 0.3 EeV filters out the largest cause of deviation from the expected results. Mainly the theta distribution follows expectations, and the distance distributions gain a more linear slope (figure 7). However, below 50 meter a bias can clearly be seen. There is a small peak at $r = 0$, while there are almost no events between the ranges of 10 to 40 meters. This can probably be attributed to saturation effects, and lack of knowledge of the Lateral Distribution Function (LDF) at short distances.

---

[1] 1 EeV = $10^{18}$ eV

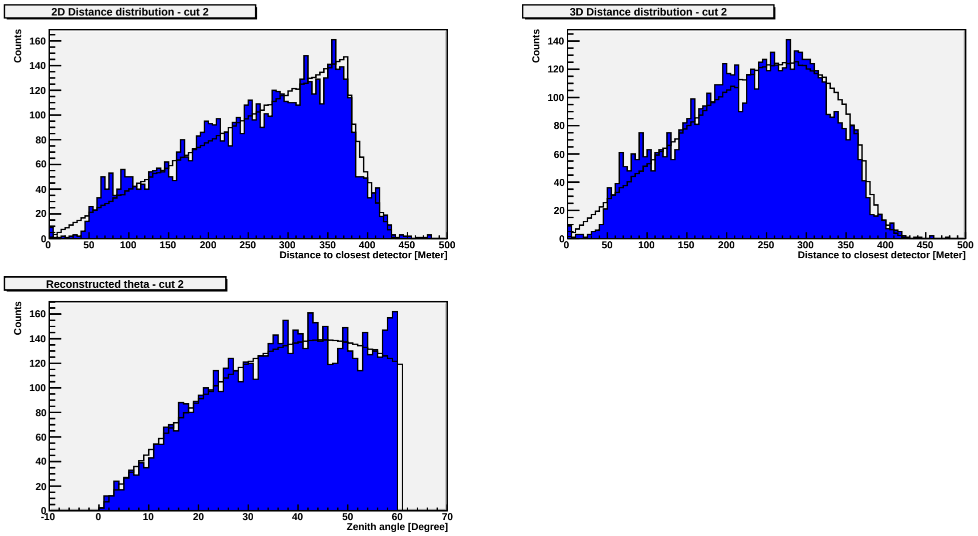


Figure 8: Reconstructions with E $\geq$ 0.3 EeV and strict T5. Distribution of 2D distance (top left), 3D distance (top right) and theta (bottom left). Reconstructed values are in blue, compared to simulations in black.

The last cut is to take into account only the strict-T5 events, since these events are reconstructed with the highest precision. This filters out a few last bad events with a distance to the nearest detector $\geq$ 450 meter. Apart from that no significant changes appear (figure 8).

| | Theta | | 2D Distances | | 3D Distances | |
|---|---|---|---|---|---|---|
| | Mean | RMS | Mean | RMS | Mean | RMS |
| Simulations | 36.56 | 14.39 | 263.1 | 94.01 | 233.7 | 88.13 |
| Cut 0 | 27.58 | 12.01 | 286.9 | 90.93 | 268.0 | 87.82 |
| Cut 1 | 36.74 | 14.42 | 262.4 | 96.62 | 234.3 | 91.01 |
| Cut 2 | 36.79 | 14.49 | 260.0 | 94.25 | 229.9 | 86.68 |

Table 2: Mean and RMS values after every cut.

Finally, I check the influence of saturated events. In these events it is possible that the core of the shower was closer to the hottest detector than reconstruction claims, since part of the cherenkov light is unnoticed. CDAS is supposed to compensate for this effect, but it is unclear how much. Thus in the distance distribution it could give a shift away from the very shortest distances to a distance somewhat smaller or larger. This would result in a peak either near the detector ($r = 0$), or at a larger radius. CDAS has a function[2] to check whether an event is saturated. As seen in figure 9, saturated events peak around a distance between 90-100 meters.

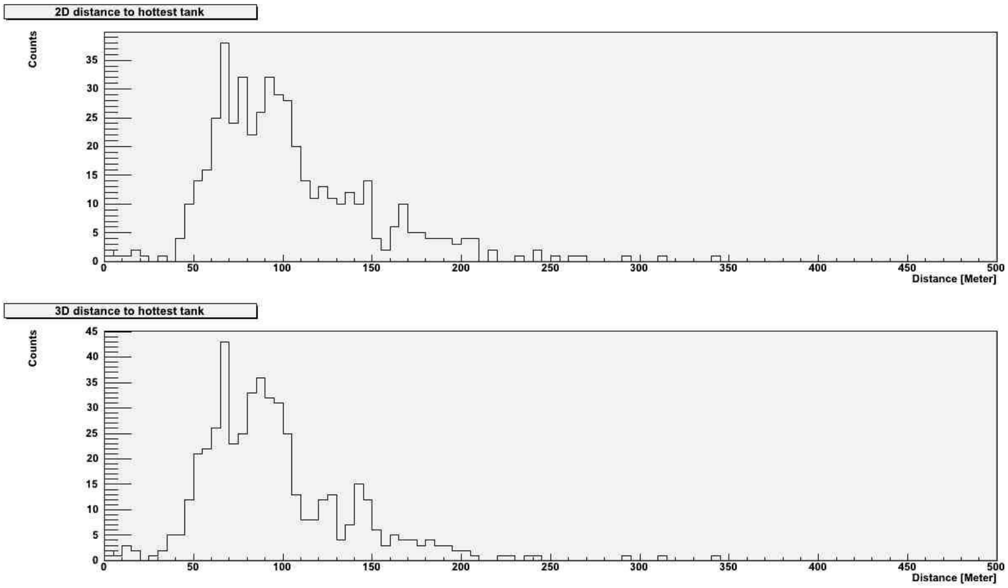


Figure 9: Distance to nearest detector for saturated events.

---

[2] event.fSelectedStations[i]->fSat

## 4.2 Phase two

Every event is originally reconstructed using the measured signal in every detector, and its time of detection. The angles are calculated mostly using the time of detection, while core position and energy are determined using the signal strength. For this reason, fluctuations of the signal strength only gives clear shifts for these two quantities. In this phase, events with strict-T5 cut, without saturated PMT's and energies $\geq 0.3$ EeV are used. Figures 10 and 11 show a typical output, these for the event with Auger ID 9662750. Figure 11 shows how in this method the signals are fluctuated in the detectors, to check that the algorithm is functioning OK. Most signals are Gaussian distributed, but there is a slight upward shift for a few detectors. This originates from the fact that detectors with a signal $\leq 1$ VEM are flagged as bad signal, and are re-calculated. The upward shift can also be seen in the energyshift (figure 10), which is not completely symmetric. Figures 12 and 13 show how the core position is shifted in the direction of the majority of the participating Cherenkov detectors, when fluctuating signals in that event.

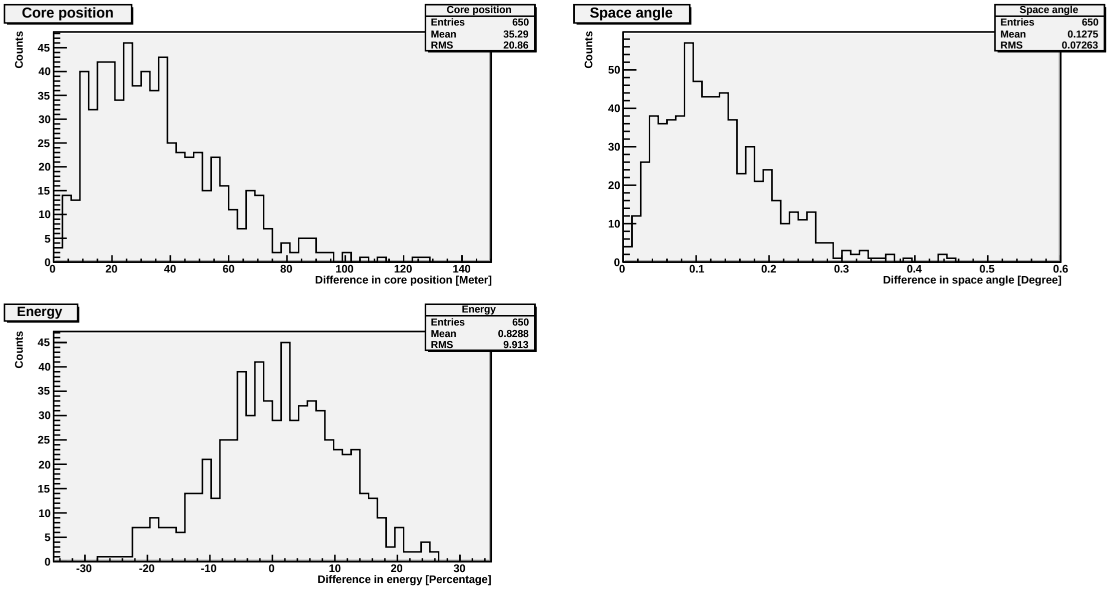


Figure 10: Output for event with Auger ID 9662750

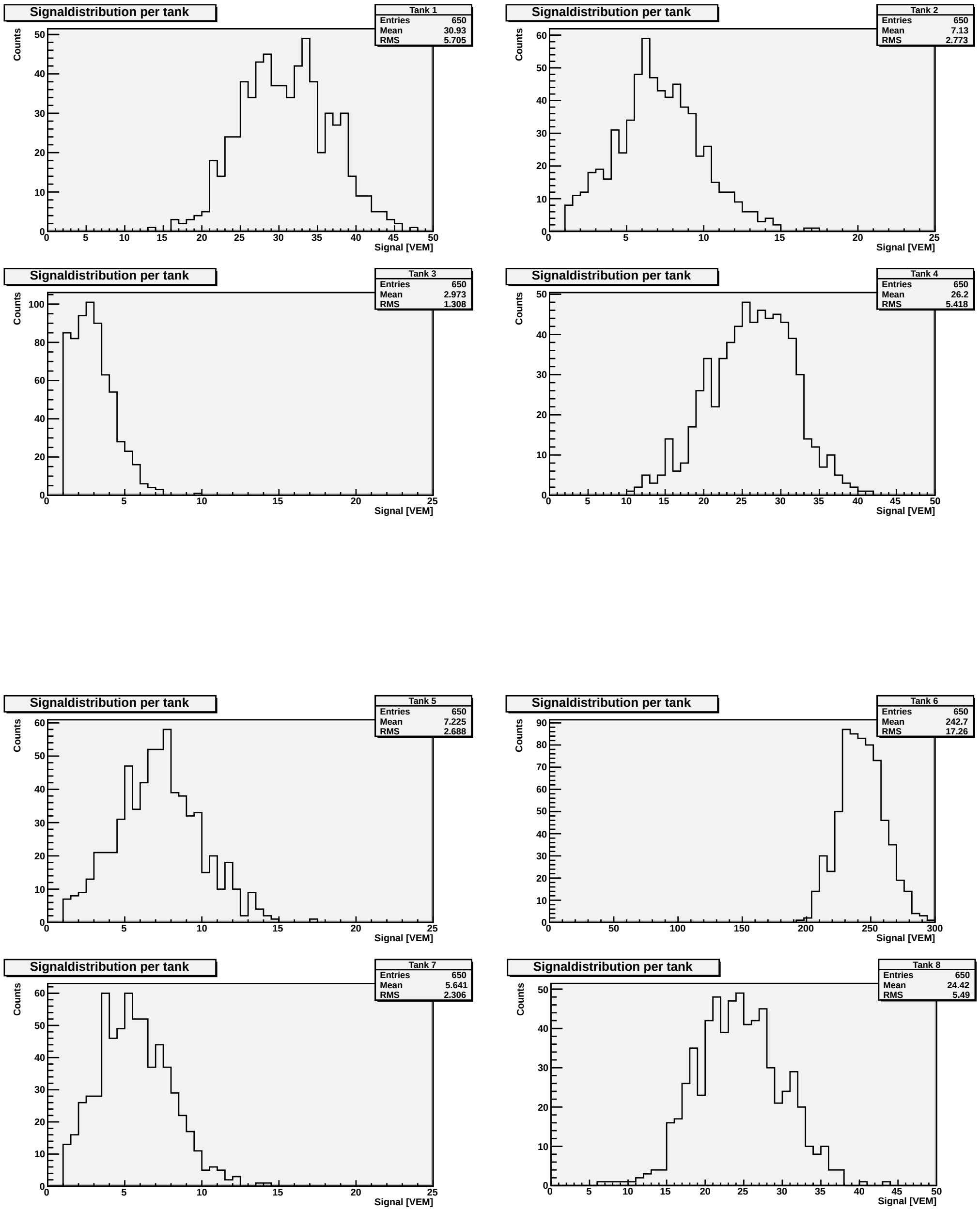
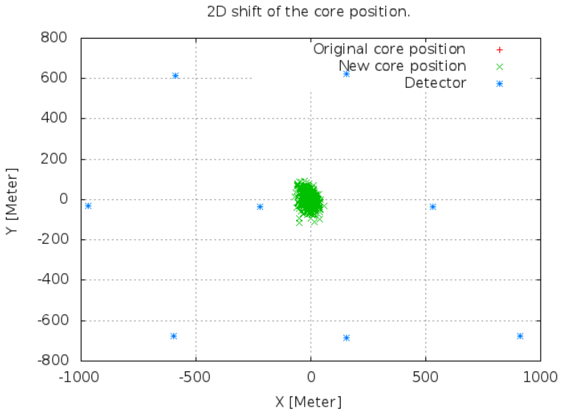


Figure 11: Fluctuated signals for event with Auger ID 9662750



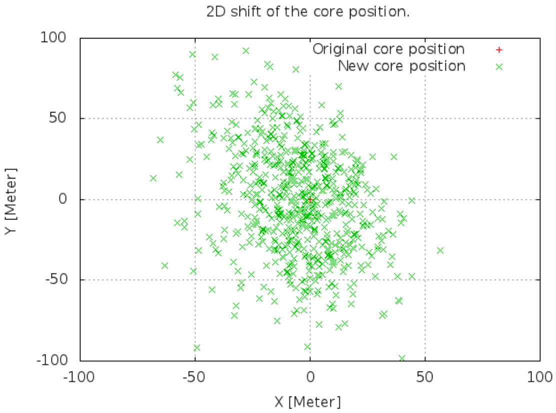


Figure 12: Overview of the shifting of core position for event with Auger ID 9662750 in two dimensions. Coordinates are normalized to the original core position of the event.

Figure 13: Zoomed overview over the shifting of core position for event with Auger ID 9662750 in two dimensions. Coordinates are normalized to the original core position of the event.

A possible influence in how much the core position shifts is its original distance to the hottest detector, and the number of detectors used in the reconstruction. Take for example an event which falls in the middle of a triangle of the only three triggered detectors. If the signals in the detectors are almost equal (which would be most probable in this scenario), fluctuations in the signals can even cause differences in the assignment of the hottest detector, and could pull the position from one detector to the other. To see how the stability of the reconstruction varies, the original 2D distance of the event to the hottest detector is plotted against the average shift in core position. The average shift is calculated by taking the mean of the lowest 68% of the simulated values for each event, which are mostly Gaussian distributed. After that, the events are separated into different groups according to the number of detectors that were triggered in the reconstruction. These groups are; 3-4 detectors, 5-8 detectors and 9 or more detectors. To see how much they can vary, the RMS values of core positions are plotted as a function of the number of detectors used in the reconstruction. Finally, the RMS value of energy is plotted vs the original energy of the event, also separated per group.

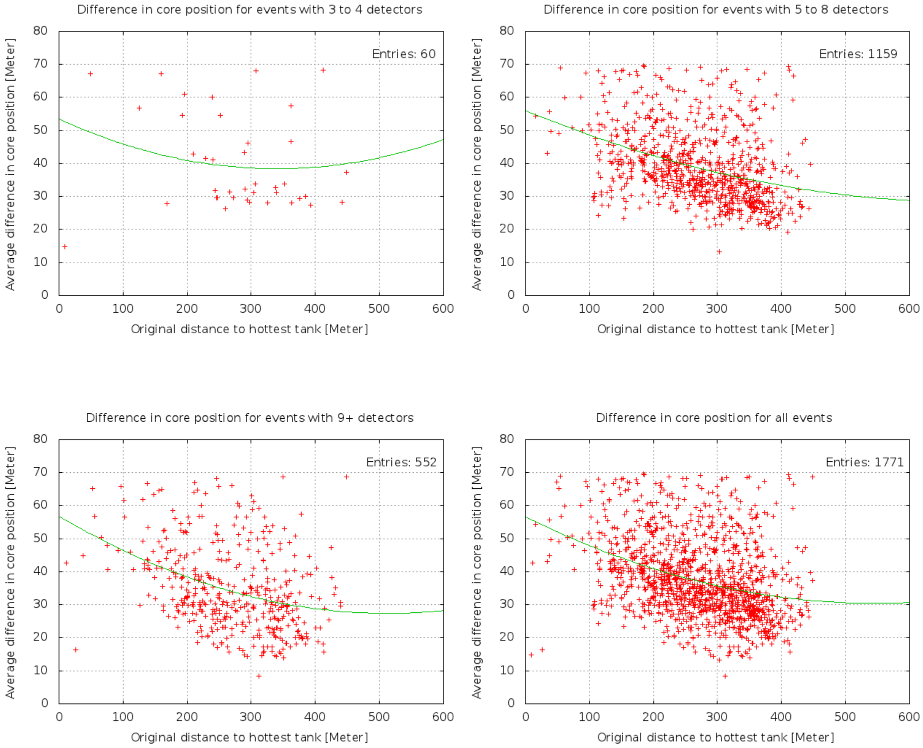


Figure 14: Difference in core position vs original distance of event to the hottest detector, separated per group of detectors. The green line shows the best fit through the data.

Figure 14 shows that the core position shifts, on average, more for events that are close to the hottest detector. There is, however, no significant difference in core position shift between the groups of different numbers of detectors. Calculation of the energy has relatively less spread for larger groups; The absolute RMS values stay about the same, while larger groups of detectors measure showers with higher energy. The energy-RMS also shrinks with distance to the hottest detector (figure 16). Between the four groups, the spread in core position shift varies. While there is a small peak for events with 5 detectors, the average spread grows less for events with more detectors used in the reconstruction (figure 15). This peak might be explained with the fact that for $\leq 5$ detectors the CDAS algorithm uses a fixed slope of the shower, instead of the free slope for $> 5$ detectors.

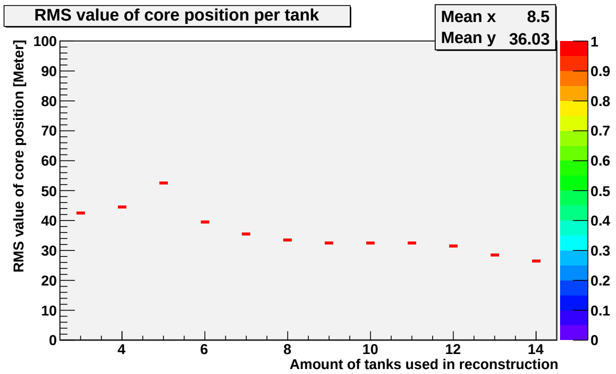


Figure 15: RMS values of the shift in core position for every event, separated per number of detectors used in the reconstruction.

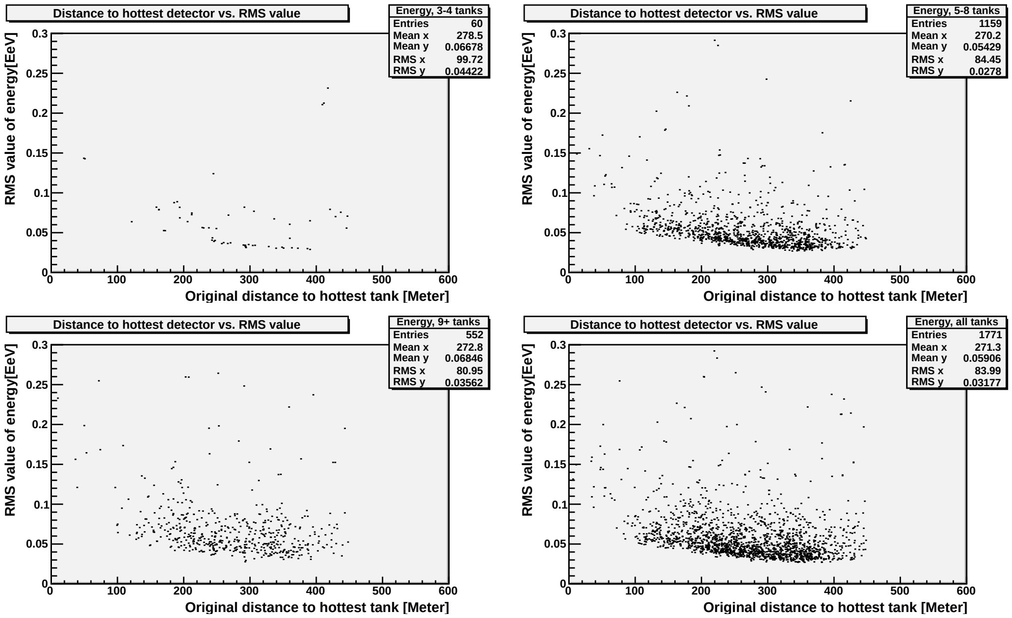


Figure 16: RMS values of the energy as function of the original energy, separated per group of detectors.

## 4.3 Core position uncertainty

The influence of the fluctuating signals on core position should be part of the complete description of the reconstruction uncertainty. This is checked by looking at the original uncertainty on the core position from CDAS, which gives an uncertainty on the X- and Y-position. Because these differ per event, the average uncertainty will be used. All infill-events from the first two months of 2010 are used, with the following cuts:

1. E ≥ 0.3 EeV.
2. No saturated detectors.
3. Only strict-T5 events

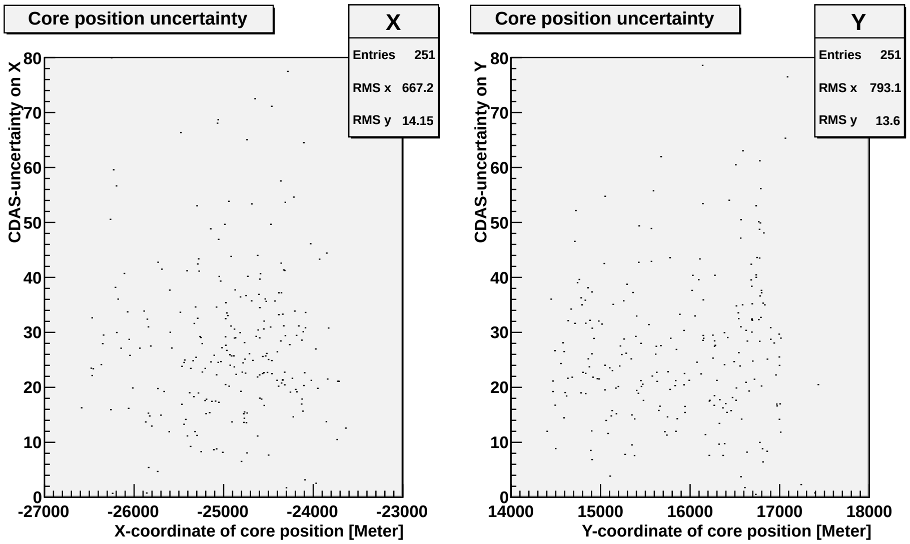


Figure 17: Scatterplots of the X- and Y-coordinates of events versus the uncertainty determined by CDAS.

This results in an uncertainty of ±14 meter for both the x- and y-coordinate. However, these values are determined empirically out of a part of all data. For that reason, this uncertainty should only be used as an approximation.

# 5 Conclusion

The infill array measures events in a wide range of energies. Most of these are below full acceptance, which lies at least a decade below the regular array's energy threshold of 3 EeV [8]. For events $\geq 0.3$ EeV, saturated events play a clear role in determination of the core position of events. Figure 9 shows that the number of saturated events peaks at a distance of 90 to 100 meters, but it is unclear if an actual shift of the position occurs. Taking that into account, the distribution of events above the energy threshold are according to expectations.

The CDAS uncertainty on the x- and y-coordinate of the core position is approximately ±14 meter. Fluctuating signals in the detectors provide an additional uncertainty between ±30 and ±60 meters. This means that calculations of the uncertainty in CDAS are incomplete, and do not take fluctuating signals into account. While the additional uncertainty is up to four times as much as the original uncertainty, it is important to note that this does not change with the number of detectors per event. Reconstructions using only a few detectors give, on average, the same shift in the core position as reconstructions using a lot more detectors, although be it with a larger spread. Calculation of the energy, on the other hand, does grow more precise with a larger number of triggered detectors.

## 5.1 Follow-up research

After doing this study, it is clear that here are a few points of interest for follow-up research. In this study, signals have been fluctuated using the average signalstrength per detector. However, quality cuts for events have very complex trigger systems for the measured signals. This goes a lot deeper than was possible in this study, by looking at individual PMT's and the shape of the measured signal by those PMT's. Looking into that would make it possible to exclude detectors from an event if their signal is too low. Also, silent detectors are not taken into account. Silent detectors are detectors that do not give a signal, but are still used in reconstruction (getting no signal in a detector also tells you something).

Third, because fluctuating all the signals takes a very long time, only a part of all acquired data was used. At the moment, there are too few results for events using only 3 or 4 detectors to draw conclusions about them. Hopefully, running more data will give more statistically significant results.

# 6 Acknowledgements

I would like to thank everyone in the Bariloche group for welcoming me so warm-heartedly, and their endless support during my stay. Particular thanks goes out to Ingo Allekotte, for his superb guidance throughout the project, and Hernán Asorey, for his continuing (technical) support and never-ending good mood. In Nijmegen, I'd like to thank Charles Timmermans for his long-distance supervision and help, and Stefan Grebe and Harm Schoorlemmer for all their patience during my start-up.

# References

[1] J. D. Jackson. *Classical Electrodynamics, 3rd Ed.* (Wiley, 1999).

[2] Website of the CDAS software. *http://www.auger.org.ar/CDAS*

[3] X. Bertou. *Effect of ground altitude on SD energy determination.* GAP-2008-001

[4] Stefan Grebe. *Robustness of the CDAS reconstruction algorithm*. GAP-2008-112

[5] The Pierre Auger Collaboration. *Statistical and systematic uncertainties in the event reconstruction and S(1000) determination by the pierre auger surface detector.*

[6] Alessio Gorgi et al. *On the SD reconstruction uncertainties and trigger efficiency*. GAP-2008-048

[7] D. Ravignani et al. *Observation of the spectrum with the AMIGA infill.* GAP-2011-010

[8] H. Asorey, I. Allekotte, X. Bertou, M. Gómez Berisso. *Determining the acceptance of the Pierre Auger Surface Detector with the Infill Array.* GAP-2009-112