Optimization of the Wavelet analysis to search for point sources of cosmic-rays

von

Ania Koob

Bachelorarbeit in Physik

vorgelegt der
Fakultät für Mathematik, Informatik und Naturwissenschaften
der
Rheinisch-Westfälischen Technischen Hochschule Aachen

im Juli des Jahres 2012

angefertigt am

III. Physikalischen Institut A

Erstgutachter und Betreuer

Prof. Dr. Thomas Hebbeker
III. Physikalisches Institut A
RWTH Aachen

Contents

List of Figures

1. Introduction

In order to gain further understanding about the universe, the arrival direction, composition and energy of cosmic rays are studied. While travelling through the universe, cosmic-rays are deflected by magnetic fields and no longer contain information about their original direction. However, the highest energy cosmic-rays are not significantly deflected by magnetic fields and can thus be used to determine sources of cosmic-rays. As the high energy particle interacts with the atmosphere, a cascade of secondary particles is produced. These secondary particles can be be measured by ground-based detectors and information about the primary particle can be gathered.

Large scale experiments as the Pierre Auger Observatory in Argentina measure (amongst other) the arrival direction of the most energetic cosmic rays. In order to distinguish a preferred arrival direction of ultra high energy cosmic rays (UHECR), one has to analyse the collected data. This analysis can be done using a wavelet analysis.

Wavelets are mathematical functions used to extract information from a given signal. In the early 1980s, the Wavelet analysis was used in geophysics to analyse seismic signals [1]. The Wavelet analysis has proven useful in astrophysical problems, it was used to study the cosmic microwave background data collected with the WMAP experiment [23], [24].

There are different kinds of wavelets, for example the Mexican Hat wavelet or the Needlet. In this thesis, the Needlet wavelet will be used.

Needlets are spherical wavelets with a lot of promising properties. They were successfully used to analyse Auger data collected from 01. January 2004 to 28. February 2011 by Matthias Plum [25].

In order to study the validity of wavelets for different kinds of probable sources of UHECRs, the Needlet analysis has been performed on simulated Monte Carlo data. A Gaussian signal with varying intensity and smearing was hidden in isotropic background. A Needlet analysis with different needlet parameters was performed, and a Signal to Noise Ratio was calculated to quantify the best parameters. Furthermore the effect of intensity and smearing of the source was studied, as well as the effect that different resolutions have on the Needlet analysis.

2. Cosmic Rays

During his balloon-flights in 1912, Victor Hess measured an unexpected increase in radiation at altitudes of roughly one kilometre from the ground. He realized that this radiation originates from outside our atmosphere and postulated cosmic rays [2]. In 1938, Pierre Auger performed coincidence measurements in the Swiss Alps and measured coincident signals at different positions. He concluded that these signals originated from one cosmic ray particle which, upon reaching the Earth's atmosphere and interacting with nitrogen, oxygen or argon, induced an extensive cosmic ray shower [3].

2.1 Air Showers

Air showers consists of different components, a hadronic, a muonic, a neutrino and an electromagnetic component [5]. Figure 2.1 shows an extensive air shower. As the extensive air shower propagates through the atmosphere, fluorescence- and cherenkov light as well as radio emission are produced.

2.2 Energy Spectrum of cosmic-rays

Cosmic-rays have been observed with energies from $10^{10}\,\mathrm{eV}$ up to over $10^{20}\,\mathrm{eV}$. Cosmic-rays are usually characterized by its flux I, which defines the number of cosmic particles arriving per square meter, angle and unit of time. The particle flux depends on the energy E by a simple power law:

$$\frac{dI}{dE} = E^{-\gamma} \tag{2.1}$$

γ is the spectral index. This dependency has been measured by different experiments over the years.

Figure 2.2 shows the flux as a function of the energy. The featured called 'knee' and 'ankle' describe a change of the spectral index γ. The 'knee' refers to a change in slope (steepening) between energies of $10^{15}\,\mathrm{eV}$ and $10^{16}\,\mathrm{eV}$. The 'ankle' characterizes the feature around $10^{18.5}\,\mathrm{eV}$, where the slope decreases [4].

2.3 Origin of cosmic-rays

Sources of high energy cosmic rays have not been confirmed yet. In 2007, the Pierre Auger Observatory showed a correlation between the arrival direction of cosmic rays with energies above $6 \cdot 10^{19}\,\mathrm{eV}$ and the positions of active galactic nuclei (AGN) lying within ~ 75 megaparsecs ($2.31 \cdot 10^{21}\,\mathrm{km}$) [7].

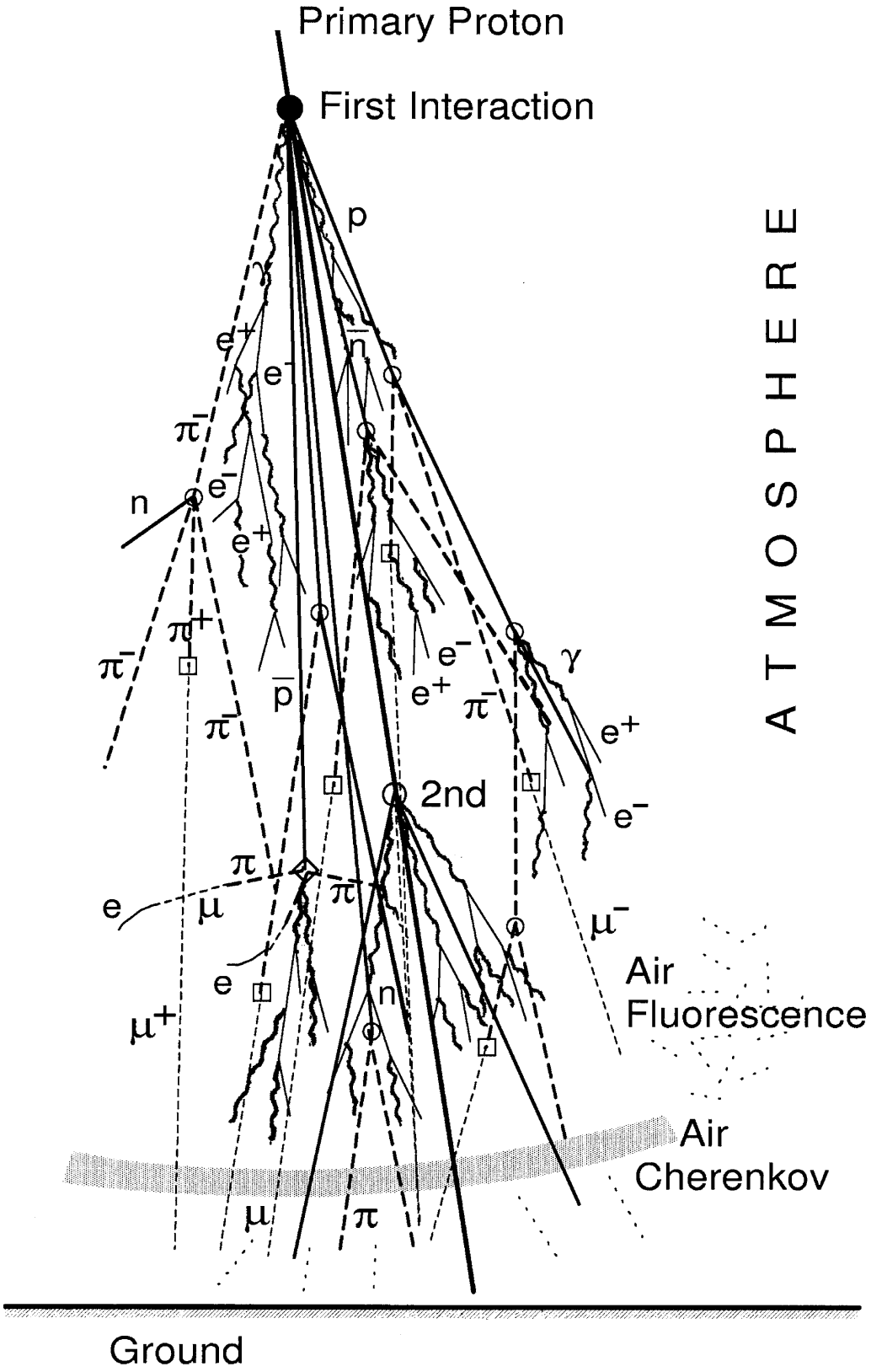


Figure 2.1: Composition of extensive air showers, [5]. Air showers consist of a hadronic, a muonic a neutrino and an electromagnetic component. Fluorescence- and cherenkov light, as well as radio emission are produced as the extensive air shower propagates through the atmosphere.

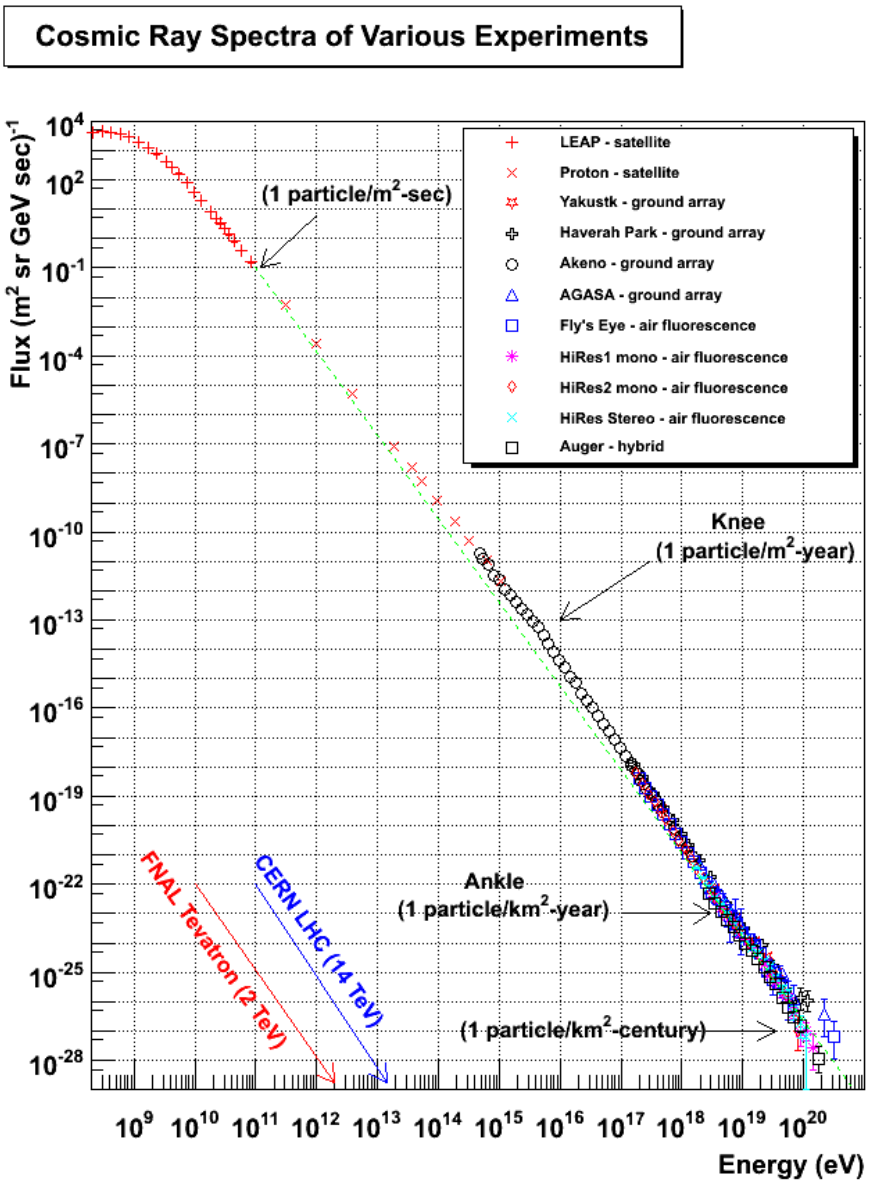


Figure 2.2: Spectrum of cosmic-rays [6]. The flux is shown as a function of the energy. The slope changes at different features of the spectrum, these featured called 'knee' (between 10^{15} eV and 10^{16} eV) and 'ankle' (around $10^{18.5}$ eV) describe a change of the spectral index γ.

2.4 Acceleration of cosmic rays

Cosmic particles can be accelerated by different features in the universe.

Acceleration up to 10^{20} eV can be provided by neutron stars with strong magnetic fields ($> 10^{12}$ G), jets of Active Galactic Nuclei (AGN), gamma-ray bursts (GRB), radio galaxies and clusters of galaxies. Figure 2.3 shows the magnetic field strength plotted against the size of astrophysical objects. The lines drawn in this figure show

protons that can be accelerated to a given energy for a shock front velocity β by astrophysical objects located above this line [9].

The acceleration of cosmic rays is described by the first and second order Fermi acceleration [8].

In the second order of Fermi acceleration, the particles find themselves in magnetised clouds, originating from jets of active galactic nuclei, or shock waves of supernovae remnants. The particles are deflected by magnetic inhomogeneities and gain energy. In the first order, the particles are accelerated in the propagating plasma of shock fronts of supernovae remnants (SNR). The particle travels from a shock front to unshocked plasma (which moves faster) and back to the shock front. A particle who travels through a shock front gains energy until the energy of the shock front is consumed. The particle then breaks free and may be detected on earth.

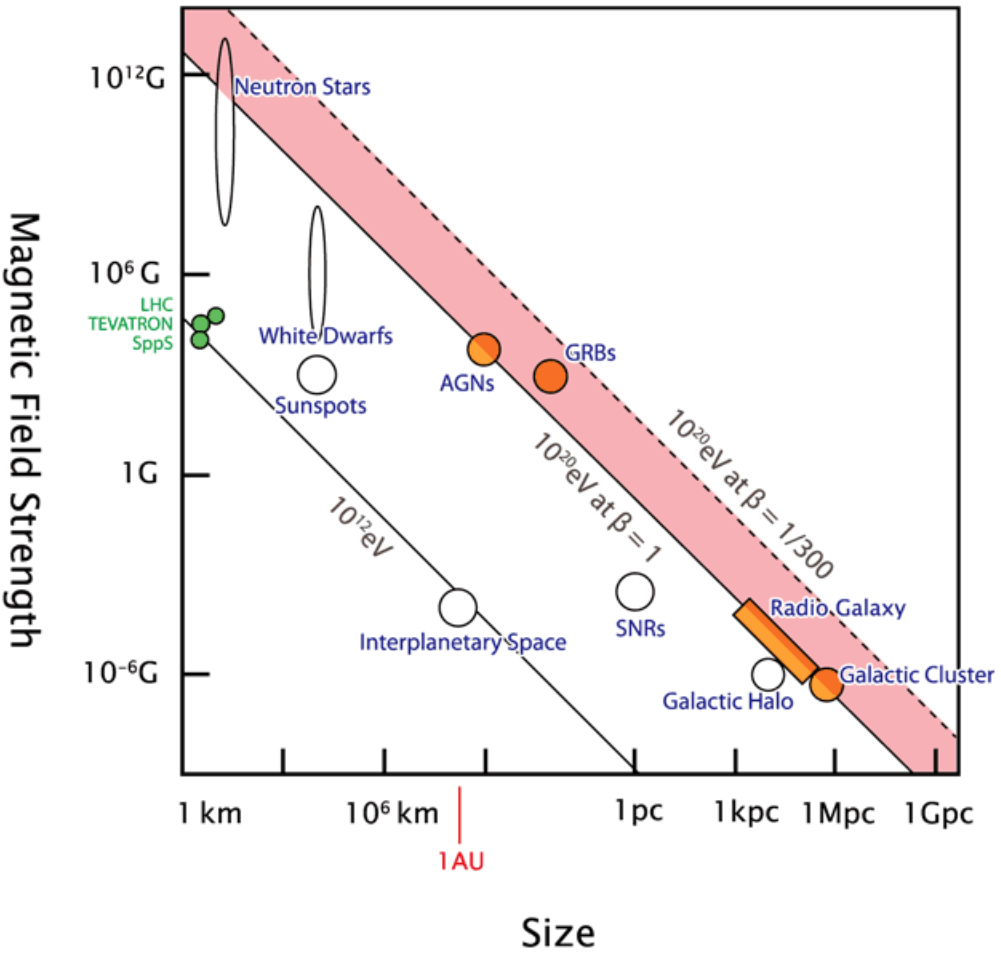


Figure 2.3: Hillas plot [9]. The magnetic field strength is plotted against the size of astrophysical objects. Astrophysical features that can accelerate protons to a given energy for a fixed shock front velocity β, are located above the corresponding line.

3. The Pierre Auger Observatory

The Pierre Auger Observatory is an international experiment which detects and studies cosmic rays. It allows the reconstruction of primary particles in an energy range from $10^{17}\,\mathrm{eV}$ to $10^{21}\,\mathrm{eV}$ [10]. As the flux of particles with energies above $10^{19}\,\mathrm{eV}$ is approximately one particle per square kilometre per year, a large array is needed to get statistical significant results. The Pierre Auger Observatory, shown in figure 3.1, is located in the Pampa Amarilla, province of Mendoza, in Argentina. It is an hybrid experiment, measuring air showers with both fluorescence telescopes (Fluorescence Detector FD) and Water-Cherenkov-detectors (Surface Detector SD) [11],[12]. Furthermore, radio emissions of extensive air showers are measured by the AERA experiment.

3.1 Surface Detector

1660 SD stations are arranged in an $3000\,\mathrm{km}^2$ array [14]. The Cherenkov-detectors are filled with $12\,\mathrm{t}$ of pure water. Particles travelling faster than light in pure water emit cherenkov light, which is collected by 3 photomultiplier tubes. As secondary particles of one air shower are measured in multiple SD stations, a difference in the detection time between the different detectors is used to calculated the trajectory of the incoming cosmic ray. Furthermore, the energy of the primary cosmic ray particle can be calculated based on the amount of light which is detected in the photomultipliers. The surface detector array has a 100% duty cycle.

3.2 Fluorescence Detector

The 24 Fluorescence Telescopes are placed at 4 different stations overlooking the SD array [15]. In moonless nights the development of a shower can be observed, as charged particles of the shower excite the nitrogen molecules, which emit ultra-violet fluorescence light. This light is collected by the telescopes and measured by photomultipliers. This provides information about the energy and the arrival direction of the primary particle.

3.3 Enhancements

In the last few years, three additional detectors were added to the Pierre Auger Observatory.

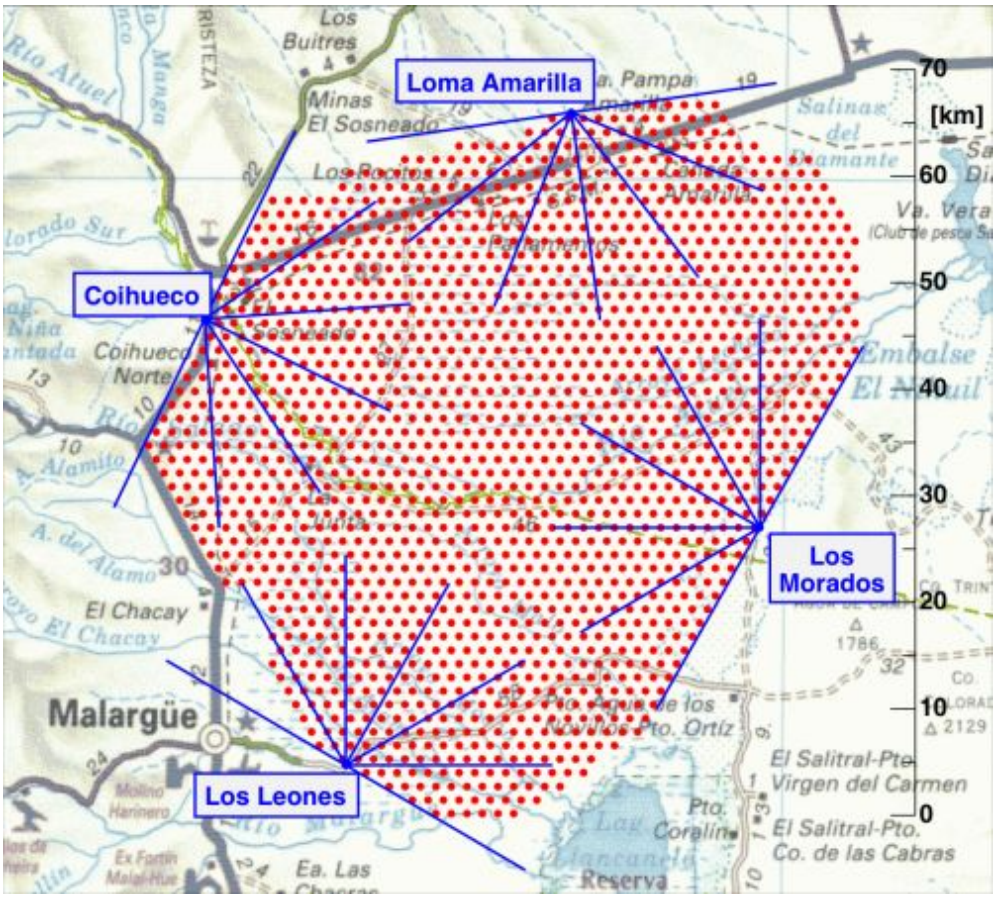


Figure 3.1: Auger Array in Argentina,[12]. The red dots represent the SD stations, with 4 telescope buildings (containing 6 telescopes each) situated at Coihueco, Loma Amarilla, Los Leones and Los Morados overlooking the array.

AERA

The Auger Engineering Radio Array (AERA) detects UHECRs with energies above 10^{17} eV by triggering on the radio emission produced in the associated air showers [13], [16].

AMIGA

AMIGA (Auger Muons and Infill for the Ground Array) is an enhancement for the Auger Observatory to study both the energy spectrum of cosmic rays down to 10^{17} eV and the muonic component of showers [17].

HEAT

To measure cosmic rays with energies down to 10^{17} eV, three High Elevation Auger Telescopes (HEAT) have been installed at Coihueco, with a viewing angle tilted 30° upwards compared to standard FD. HEAT provides the possibility to study the cosmic ray energy spectrum and mass composition in a very interesting energy range, where the transition from galactic to extragalactic cosmic rays is expected to happen [18].

4. Wavelet Analysis

4.1 Introduction

Wavelets are mathematical functions of the L^2 space, and can be seen as wave-like oscillating functions which are used to extract information from data. A signal is decomposed into wavelets without loss of information (i.e. reversible). These different wavelets are obtained from a mother wavelet by changing its dilatation and translation parameters [19], [20].

4.2 Multipole Expansion

The multipole expansion is a standard method for anisotropy studies. In spherical coordinates (θ, ϕ), the spherical harmonics are used.

$$Y_{lm}(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)}{(l+m)}} P_l^m(cos(\theta)) e^{im\phi} \tag{4.1}$$

P_l^m are the associated Legendre polynomials with coefficients l = 0,1,2,... and m = 0, $\pm$ 1, $\pm$ 2, ..., $\pm$ l. $l = 0$ describes a monopole, l=1 a dipole, l=2 a quadrupole and so on [21]. Spherical data $f(\theta,\phi)$ is expressed by the multipole expansion

$$f(\theta,\phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{l,m} Y_{l,m}(\theta,\phi) \tag{4.2}$$

with the multipole coefficients $a_{l,m}$ for the spherical harmonics $Y_{l,m}(\theta,\phi)$. The infinite sum defines the angular precision.

4.3 Healpix software

HEALPix is an acronym for Hierarchical Equal Area isoLatitude Pixelization of a sphere [22]. As astrophysical data is recorded in spherical coordinates, HEALPix provides a handy subdivision of a spherical surface in equal-size bins. The pixels are distributed on constant latitudes, which enables a faster analysis involving spherical harmonics. Figure 4.1 shows the pixelisation scheme for different resolutions. A HEALPix map consists of $N_{pix} = 12 \cdot N_{side}^2$ pixels, with an equal area of $\Omega_{pix} = \frac{4\pi}{12N_{side}^2}$ per pixel. N_{side} is the number of divisions along the side of a base-resolution pixel that is needed to reach a desired high-resolution partition [22].

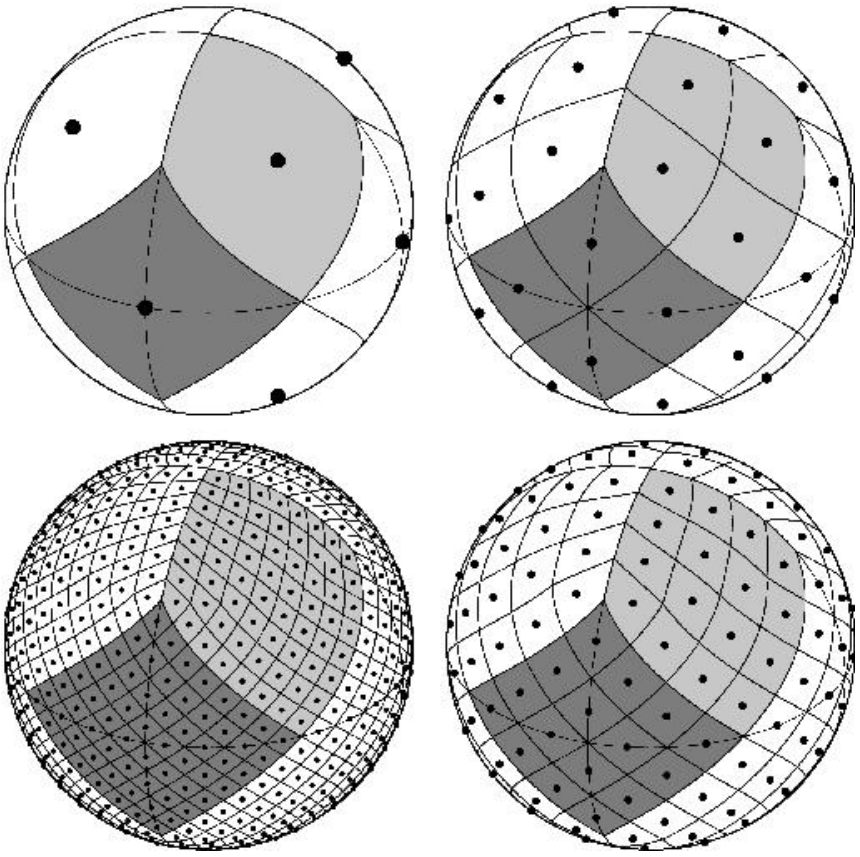


Figure 4.1: HEALPix pixelisation scheme for different resolutions (12, 48, 192, and 768 pixels). To obtain a higher resolution, one pixel is divided into four new pixels.

4.4 Wavelet Analysis

Wavelets have been used to extract non-Gaussian data of the Cosmic Microwave Background (CMB) recorded by WMAP [23] [24]. Wavelets are wave-like oscillations, they either have a finite length or are fast decaying. The decay is achieved by an exponentially decreasing term [21]. The wavelet transform consists of a combination of wavelet and unknown signal, in order to extract information from the signal. It is a convolution of the wavelet function Φ_R of different scales with the signal trace. Figure 4.2 shows the procedure of a wavelet analysis [25]. A conversion in spherical harmonics is applied to both the mother wavelet and the data (signal and noise) that is going to be analysed. Wavelet and signal are multiplied, this product is then transformed back into the spatial domain.

4.5 The Needlet

Needlets are a form of spherical wavelets which offer a lot of advantages over other wavelets [26]. They enjoy good localization properties in both spatial and frequency domain. For a full-sky analysis, the different needlet coefficients are, at a fixed angular position, asymptotically uncorrelated [27]. Needlets are also adapted to the HEALPix package.

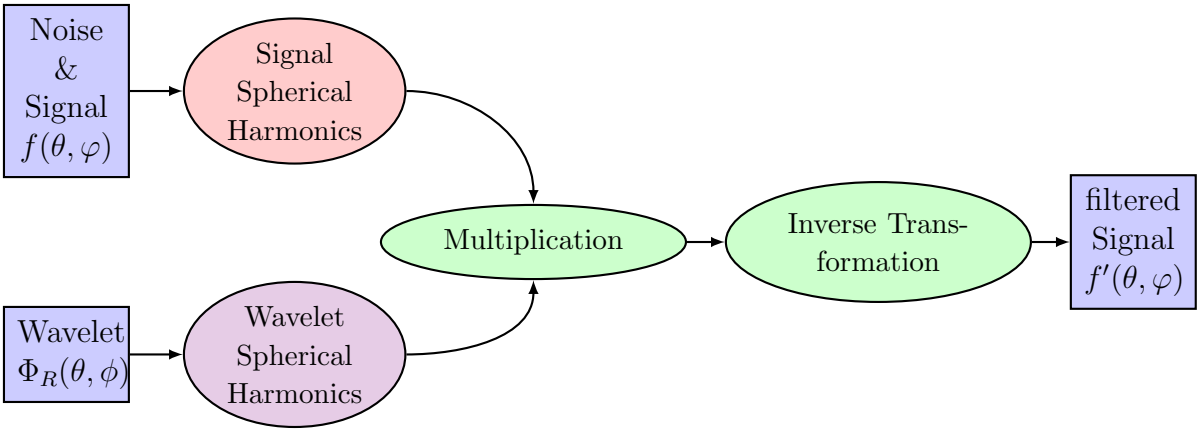


Figure 4.2: Flowchart of Wavelet analysis , [25]. The Wavelet as well as the data $f(\theta,\varphi)$ are transformed into spherical harmonics. They are then multiplied and transformed back into the spatial domain, which yields the filtered signal $f'(\theta,\varphi)$.

The range of the needlet is defined by the parameters j (needlet scale) and B (needlet width) [25]. As the needlet is localized on a finite number of multipole moments l, the multipole expansion (formula 4.2) can be truncated. The equatorial coordinate system is used. The spherical needlet is defined as [28]

$$\Psi_{jk}(\delta,\alpha) = \sqrt{\lambda_{jk}} \sum_{l} b\left(\frac{l}{B^j}\right) \sum_{m=-l}^{l} Y^*_{lm}(\delta,\alpha) Y_{lm}(\zeta_{jk}) \tag{4.3}$$

α is the right ascension, δ the declination. $Y_{lm}(\zeta_{jk})$ and Y^*_{lm} are the spherical harmonics, with $\sum_{m=-l}^{l} Y^*_{lm}(\theta,\phi) Y_{lm}(\zeta_{jk})$ being the spherical harmonics expansion. $\{\zeta_{jk}\}$ denotes coordinates of the pixel center of the Healpix scheme, k is the pixel number. ζ_{jk} is the pixel center in the HEALPix scheme. λ_{jk} is the scaling factor, and can be approximated by $1/N_{pix}$, with N_{pix} being the number of pixels for the chosen HEALPix resolution.

Needlets should be seen as a product of the projection operator $\sum_{m=-l}^{l} Y^*_{lm}(\theta,\phi) Y_{lm}(\zeta_{jk})$ with a suitably chosen window function, called needlet kernel function $b(\frac{l}{B^j})$ [25]. It corresponds to a filter with values between zero and one. The filter is based on the multipole moment l, the needlet width B and the needlet scale j. As seen on figure 4.3 and 4.4, higher multipole moments are used for increasing needlet scales and widths.

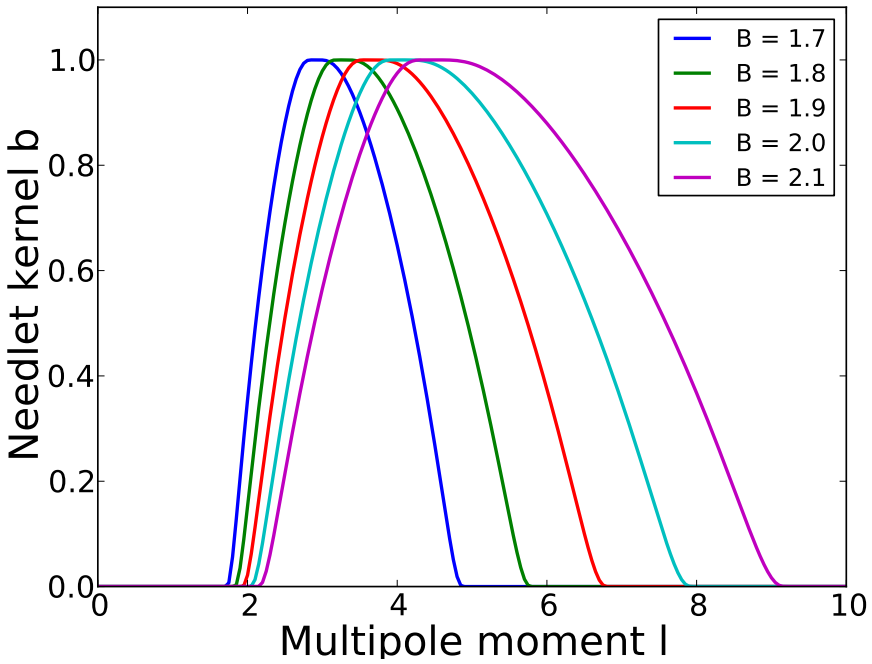


Figure 4.3: Needlet kernel function $b(\frac{l}{B^j})$ plotted against the multipole moment l from different needlet widths B at a fixed needlet scale j = 2. For increasing needlet width, higher and also more multipole moments are used.

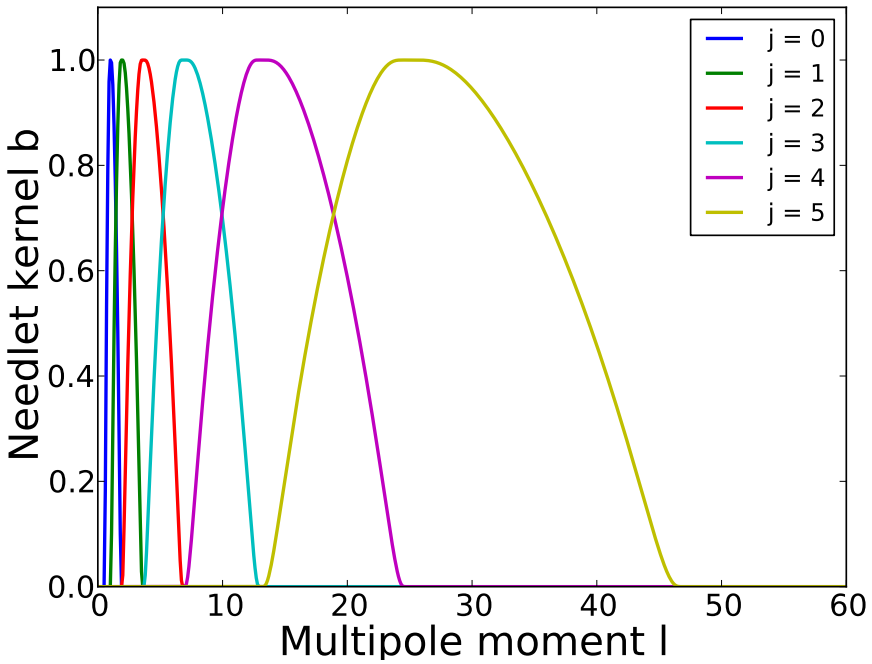


Figure 4.4: Needlet kernel function $b(\frac{l}{B^j})$ plotted against the multipole moment l from different needlet scales j at a fixed needlet width B = 1.9. For higher needlet scales additional multipole moments are used.

The angular resolution of a Needlet analysis depends on the N_{side} used. The angular resolution can be calculated for an N_{side} of 64 (corresponding to a map of 49152 pixels) as follows [22]:

$$\alpha \equiv \sqrt{\Omega_{pix}} = \sqrt{\frac{4\pi}{12N_{side}^2}} = 0.016\,\mathrm{rad} \tag{4.4}$$

The angle, converted into degrees is $\alpha = 0.916°$. As the angular resolution α depends on the range of the multipole moments, higher multipole moments can distinguish smaller structures. Equation 4.6 can be used to calculate the maximal multipole moment l_{max} considering Healpix resolution and limiting computing time.

$$\alpha \simeq \frac{180°}{l} \tag{4.5}$$

$$l_{max} \simeq \frac{180°}{\alpha} = 196.5 \tag{4.6}$$

For a map containing 49152 pixels, a maximum multipole of 250 has been used.

As for a lower resolution, a N_{side} of 32 (corresponding to 12288 pixels) leads to an angular resolution of $\alpha = 1.83°$ [22].

Transforming a signal $f(\delta, \alpha)$ into spherical harmonic coefficients a_{lm} leads to following formula for the needlet power value β_{jk}, where k is the HEALPix pixel and j the needlet scale.

$$\beta_{jk} := \sqrt{\lambda_{jk}} \sum_l b\left(\frac{l}{B^j}\right) \sum_{m=-l}^{l} a_{lm} Y_{lm} \tag{4.7}$$

5. Needlet Analysis of Monte-Carlo Data

In this chapter, the needlet analysis is applied to various Monte Carlo data. Different needlet widths B and maximum needlet scales jmax are used and the result of their respective analysis compared. The equatorial coordinate system is used.

5.1 Monte Carlo Simulation

The following analysis was done for a full-sky exposure, which means that no coverage mask was applied and the sphere was uniformly weighted.

30000 events were distributed in the HEALPix skymap containing 49152 pixels, this corresponds to an angular resolution of $\alpha = 0.916°$. The number 30000 was chosen because it corresponds to the number of anisotropy data collected by the Pierre Auger Observatory. 50, 100, 150 or 200 source events were randomly generated, following a Gaussian distribution. 29950, 29900, 29850 or 29800 background events respectively were isotropically simulated and added to the signal events.

The Monte Carlo maps were generated with a Gaussian source situated at 90° right ascension (α) and $-30°$ declination (δ). Different Gaussian source smearings were simulated, namely 3°, 6° and 12°. 6° corresponds to the size of Centaurus A, the nearest AGN, who is suspected to be a source of UHECRs. A smearing of 3°(12°) was chosen as it is half(double) of Centaurus A's size. As a different approach, a rougher binning of the HEALPix map was set up, namely 12288 pixel, which corresponds to an angular resolution of $\alpha = 1.83°$ [22]. Table 5.1 shows the different settings.

In addition, each analysis was done with thirteen different needlet Parameters. Needlet width B and maximum needlet scale jmax were varied. Table 5.2 shows the

number of pixels	number of signal entries	number of background entries	smearing
49152	50	29950	3°,6°,9°
49152	100	29900	3°,6°,9°
49152	150	29850	3°,6°,9°
49152	200	29800	3°,6°,9°
12288	150	29850	3°,6°,9°

Table 5.1: Different settings (resolution, intensity and smearing) used for a needlet analysis.

needlet width B	maximum needlet scale jmax	maximum multipole moment l
1.7	7	69
1.8	7	110
1.9	5	47
2.0	5	63
2.1	5	85
2.15	5	98
2.2	5	113
2.3	5	148
2.4	4	79
2.6	4	118
2.8	4	172
3.0	4	242
3.1	3	92

Table 5.2: Different B and jmax parameters used in the wavelet analysis, as well as maximum multipole moment used.

different B and j factors used, as well as the maximum multipole moment l used in the analysis. In appendix A.2, an additional table shows the different multipole moments used for each needlet scale.

Figure 5.1 shows the signal density distribution. A Gaussian source located at $\alpha = 90°$ (right ascension) and $\delta = -30°$ (declination) with a smearing of 6°. In this example, 150 signal events were randomly distributed based on this distribution, which is shown in figure 5.2. Figure 5.3 shows the same events projected in a two-dimensional histogram.

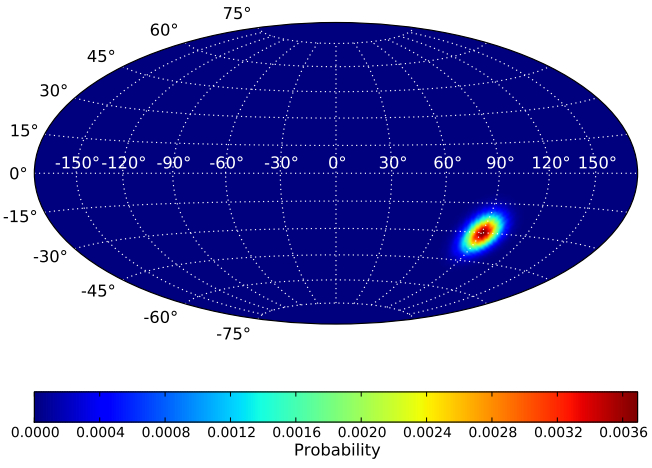


Figure 5.1: Signal density distribution for a Gaussian point source located at $\alpha = 90°$ (right ascension) and $\delta = -30°$ (declination) with a smearing of 6°.

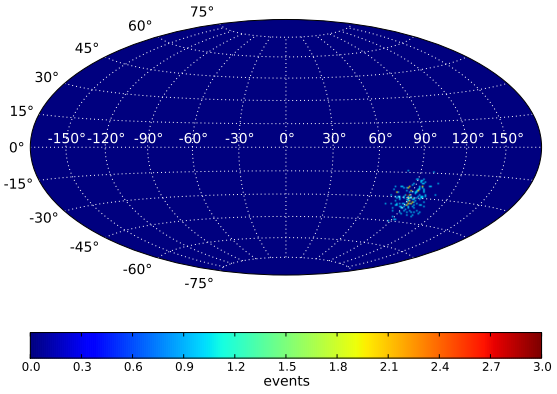


Figure 5.2: Monte Carlo signal skymap with 150 events and a 6° smearing.

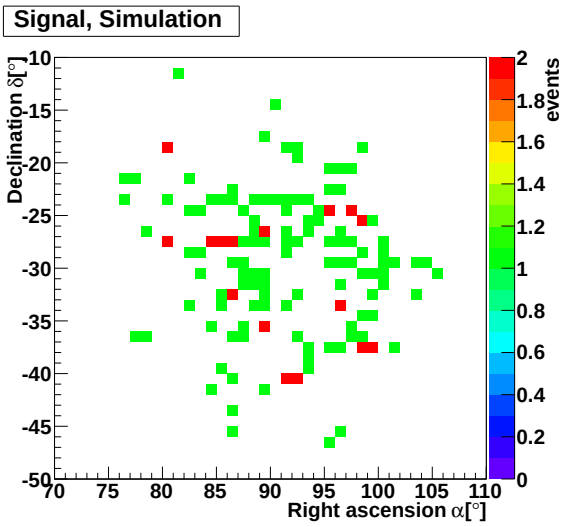


Figure 5.3: Monte Carlo signal with 150 events and a 6° smearing, projected in a two dimensional histogram.

To obtain a total of 30000 events in one map, 29850 background events were isotropically distributed. Figure 5.4 shows the obtained map containing signal and background entries, figure 5.5 the corresponding two-dimensional histogram.

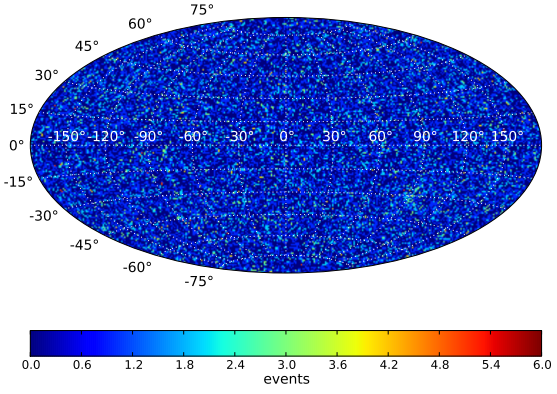


Figure 5.4: Monte Carlo skymap with 150 signal events, a 6° smearing and 29850 isotropic background events.

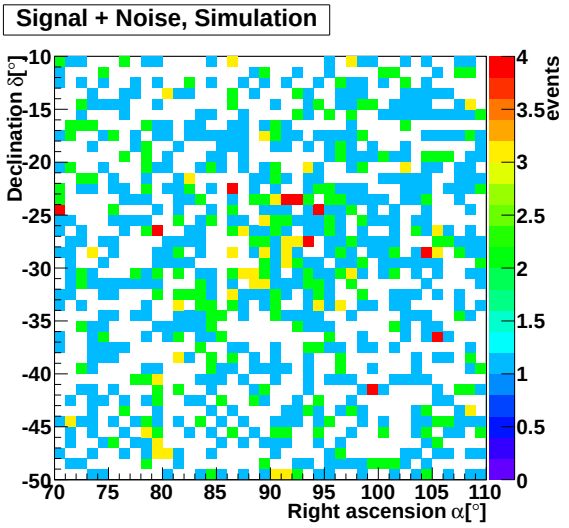
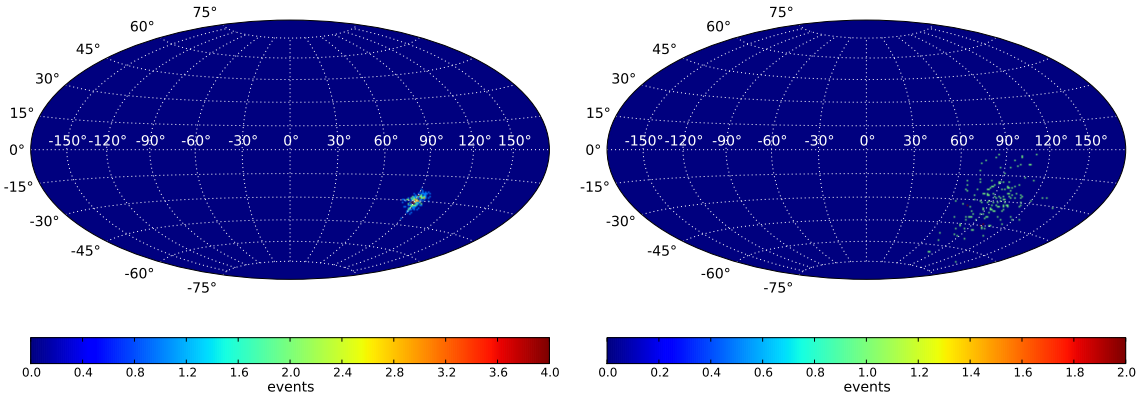


Figure 5.5: Monte Carlo with 150 signal events, a 6° smearing and 29850 isotropic background events, projected in a two dimensional histogram.

As a comparison, figure 5.6 shows 150 entries simulated following a Gaussian distribution with a smearing of 3° and 12°.



(a) 3° smearing. **(b)** 12° smearing.

Figure 5.6: 150 singal events simulated following a Gaussian distribution with a smearing of 3° and 12°.

5.2 Analysis

In a first step, the needlet analysis is applied to the skymap containing both signal and background events. As an example, the needlet values B = 2.3 and jmax = 5 were chosen. A separation in spherical harmonics of the data was done (as the needlet is already defined in the frequency domain), then signal and wavelet were multiplied and transformed back into the spatial domain. Figure 5.7 shows the procedure.

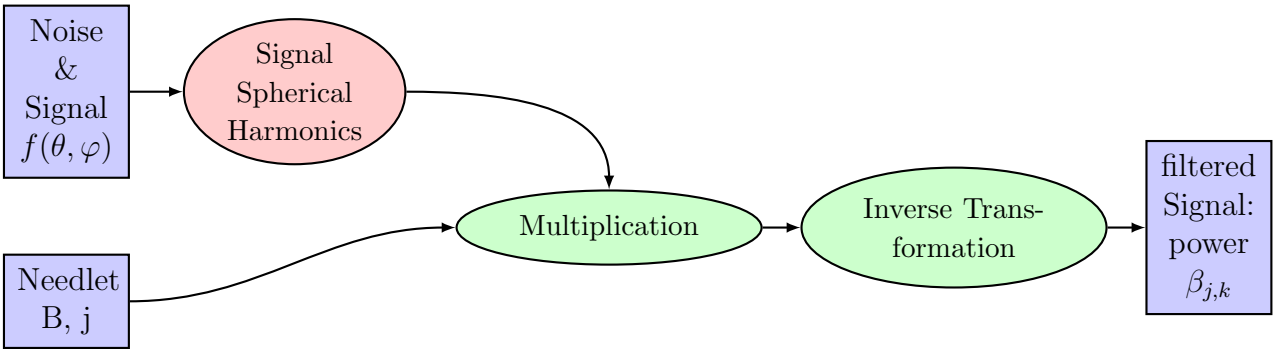


Figure 5.7: Flowchart of needlet analysis, [25]. The data $f(\theta, \varphi)$ is transformed into spherical harmonics, as the needlet is already defined in spherical harmonics. They are then multiplied and transformed back into the spatial domain, which yields the filtered signal $f'(\theta, \varphi)$.

During the spherical harmonics expansion, the skymap is split into the different needlet scales, which are shown in figure 5.8. The power β_{jk}, with j being the scale and k the pixel, is shown in arbitrary units (a.u.). The reconstructed source is not visible in every needlet scale. The scales j = 4 and j = 5 are sensitive to structures smaller than the one simulated, in these scales, only isotropic fluctuations are visible.

As for scales j = 0 and j =1, they are sensitive to structures larger than the one simulated. The source is clearly visible in the scales j = 2 and j = 3.

5.2.1 Significance

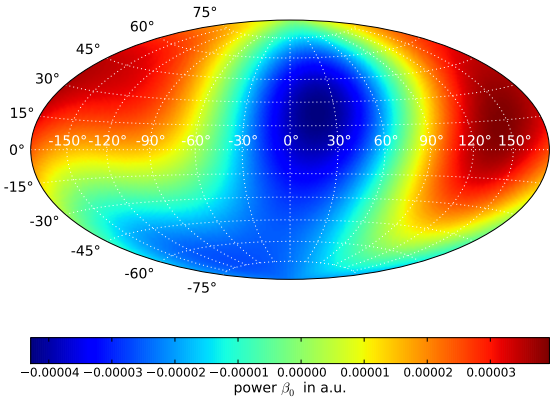
In a second step, the previously calculated power β_{jk} is converted into a significance S_{jk} by comparing it to isotropy. For this, 25000 MC-samples with a random number (between 0 and 50000) of isotropic background events are simulated and afterwards analysed with the different needlet parameters B and jmax used. For every needlet scale, 4 random pixels out of every of the 25000 maps were taken and filled in a two dimensional histogram in figure 5.9, histogram (a). As this analysis is done for a full-sky exposure, it is possible to select random pixels out of every map. For every number of entries bin, a one-dimensional Gaussian function is fitted to the power distribution in this particular bin. Histograms (b),(c) and (d) of figure 5.9 show mean $\langle\beta_{jk_{noise}}\rangle$, standard deviation $\sigma_{jk_{noise}}$ and $\frac{\chi^2}{ndf}$ respectively of the fit. As 30000 events were simulated in this study, only the bin corresponding to number of events = 30000 is relevant.

To recalculate the power β_{jk} of the needlet analysis (figure 5.8) into a significance, the standard deviation of the Gaussian fit (figure 5.9, histogram c)) is needed.

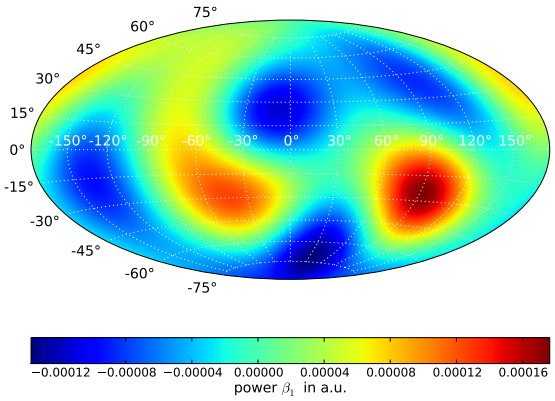
The Significance S_{jk} is calculated as follows [25]:

$$S_{jk} = \frac{|\beta_{jk} - \langle\beta_{jk_{noise}}\rangle|}{\sigma_{jk_{noise}}} \tag{5.1}$$

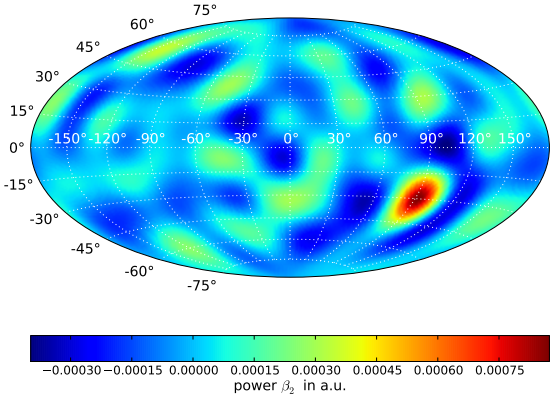
This leads, for every needlet scale, to a skymap normalized on the corresponding value $\sigma_{jk_{noise}}$.



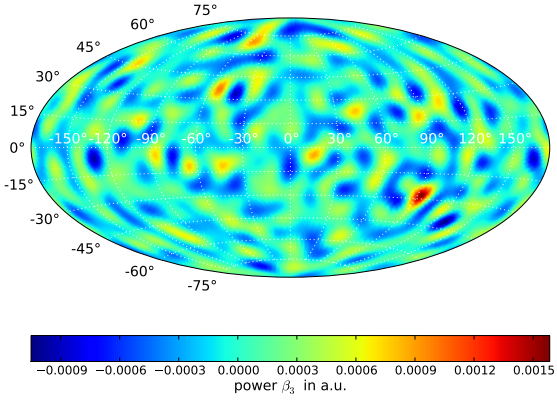
(a) Analysis result at needlet scale j = 0.



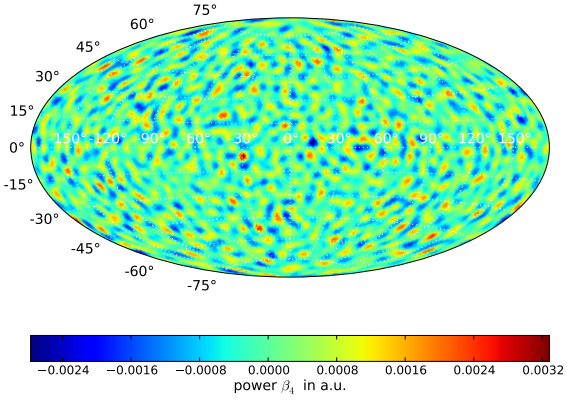
(b) Analysis result at needlet scale j = 1.



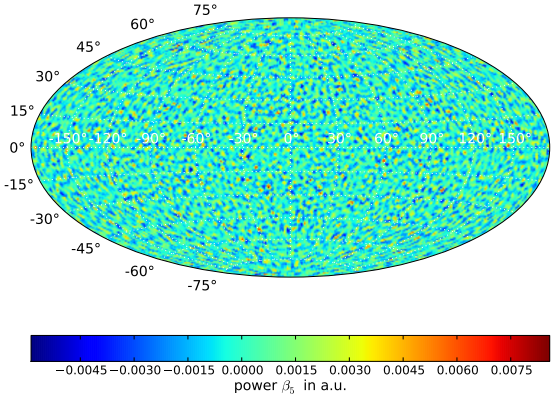
(c) Analysis result at needlet scale j = 2.



(d) Analysis result at needlet scale j = 3.

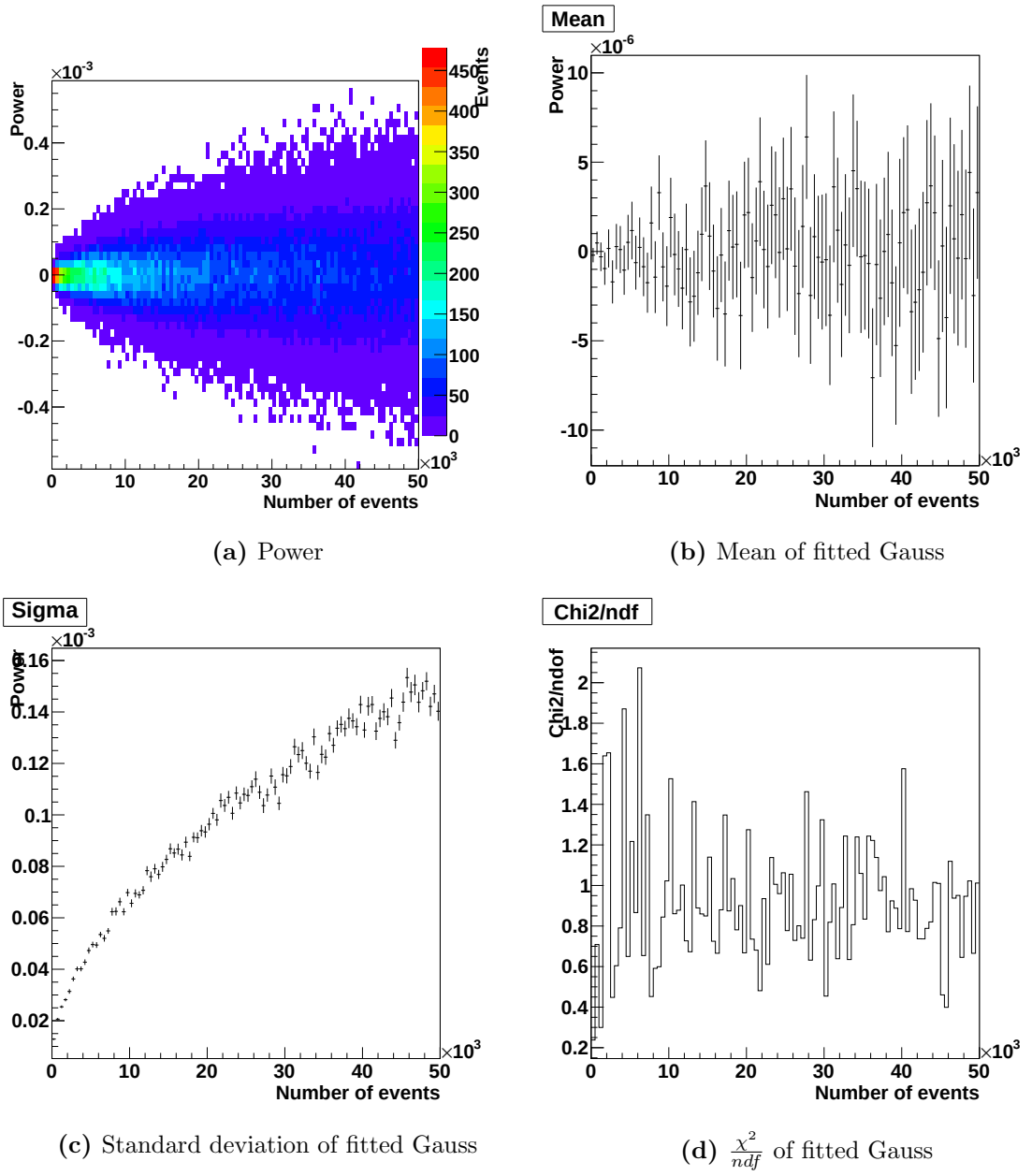


(e) Analysis result at needlet scale j = 4.



(f) Analysis result at needlet scale j = 5.

Figure 5.8: MC data sample analysed with needlet parameters B = 2.3 and jmax = 5, split into different needlet scales. The power β_{jk} is shown in arbitrary units. The different needlet scales are sensitive to different sizes of sources, the simulated source is clearly visible in the scales j = 2 and j = 3. The scales j = 0 and j = 1 are sensitive to structures larger than the one simulated, whereas the scales j = 4 and j = 5 are sensitive to smaller structures. In the needlet scales j = 4 and j = 5, only isotropic fluctuations are visible.



(a) Power

(b) Mean of fitted Gauss

(c) Standard deviation of fitted Gauss

(d) $\frac{\chi^2}{ndf}$ of fitted Gauss

Figure 5.9: 25000 isotropy maps were simulated, with varying number of events. In this study, only the bin corresponding to 30000 number of events is relevant. In histogram a), the entries of four random pixels out of one fixed needlet scale in every isotropy map were filled in. In every bin, a one-dimensional Gaussian χ^2 fit was performed, with histograms b), c) and d) showing the resulting mean, standard deviation and $\frac{\chi^2}{ndf}$. This example shows isotropy maps analysed with the needlet paramters B = 2.3, jmax = 5, zeroth needlet scale.

5.2.2 Threshold cut

As all of the needlet scales are normalized, a threshold cut is applied to select data which is not conform with isotropy. Only values above 3 standard deviations are kept, which (for a Gaussian distribution) corresponds to a 0.0027 probability that these values are conform with isotropy. The threshold cut is done using the standard deviation of the Gaussian fit on isotropic maps, as shown in figure 5.9.

The significance S^*_{jk} is determined by the following threshold condition

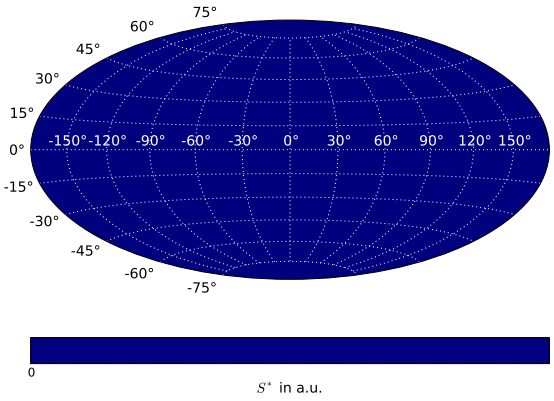
$$S^*_{jk} = 0 \text{ for } |S_{jk}| < 3 \tag{5.2}$$
$$S^*_{jk} = S_{jk} \text{ for } |S_{jk}| \geq 3 \tag{5.3}$$
$$\tag{5.4}$$

This is applied to every scale of figure 5.8, which leads to figures 5.10.

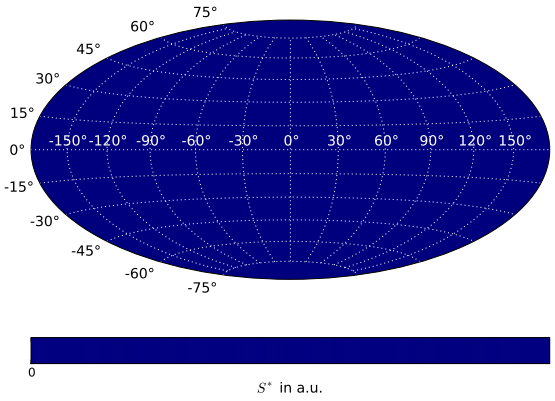
As this analysis is designed to search for point sources of cosmic-rays in Auger SD Data and we have no a priori information about the size of possible sources, all needlet scales have to be taken into account. This is done by summing up the different scales:

$$S^*_k = \sum_{j=0}^{jmax} S^*_{jk} \tag{5.5}$$

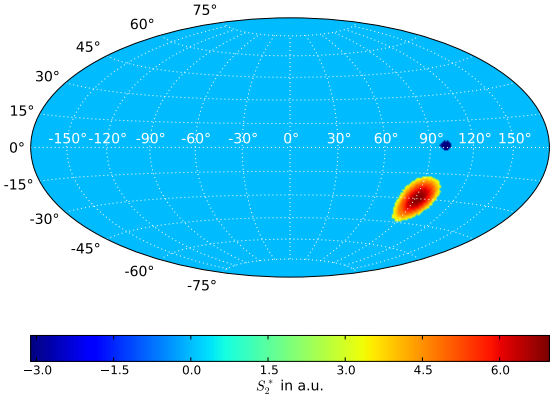
Figure 5.11 shows the final result of this analysis with the applied threshold cut for a map containing 150 signal events, 29850 noise events and a Gaussian smearing of 6°. Figure 5.12 and 5.13 show the result for a smearing of 3° and 12°. In this example, the needlet parameters used are B = 2.3, jmax = 5. For all the simulated smearing, the source was successfully reconstructed by the wavelet analysis.



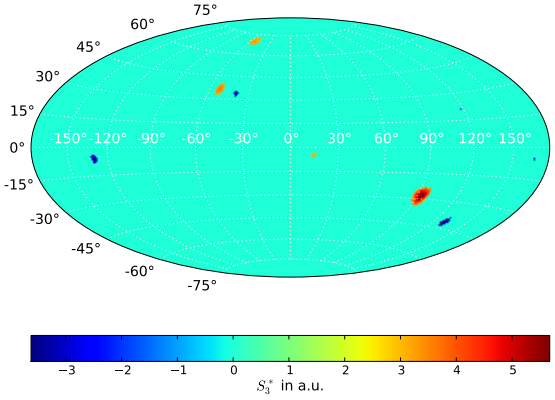
(a) Results after threshold cut,needlet scale j = 0.



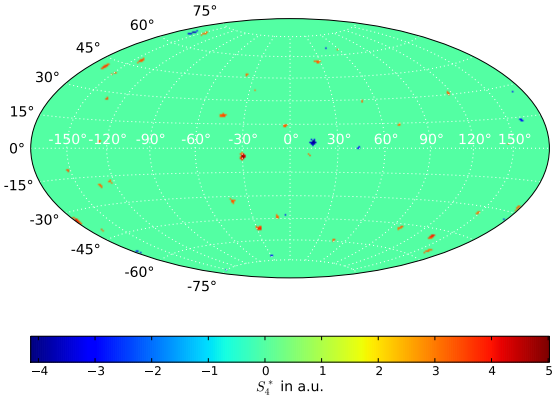
(b) Results after threshold cut,needlet scale j = 1.



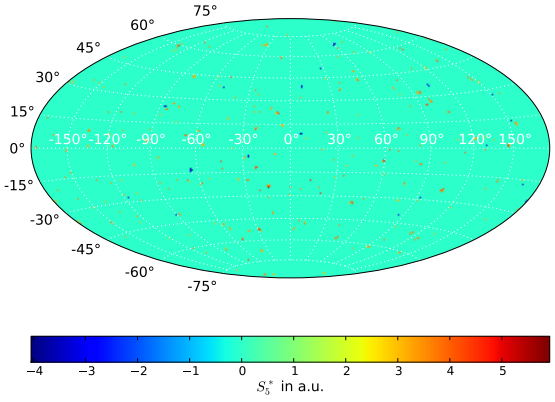
(c) Results after threshold cut,needlet scale j = 2.



(d) Results after threshold cut,needlet scale j = 3.



(e) Results after threshold cut,needlet scale j = 4.



(f) Results after thresholdcut,needlet scale j = 5.

Figure 5.10: MC data sample analysed with Needlet parameters B = 2.3, jmax = 5, split into different needlet scales, after performing a 3σ threshold cut. The Significance S^*_{jk} is shown in arbitrary units. In the first two needlet scales, all pixels are set to zero due to the threshold cut. In the scales j = 2 and j =3, the source has been clearly distinguished from isotropy, in the two last scales j = 4 and j = 5, isotropic fluctuations are still visible after the threshold cut.

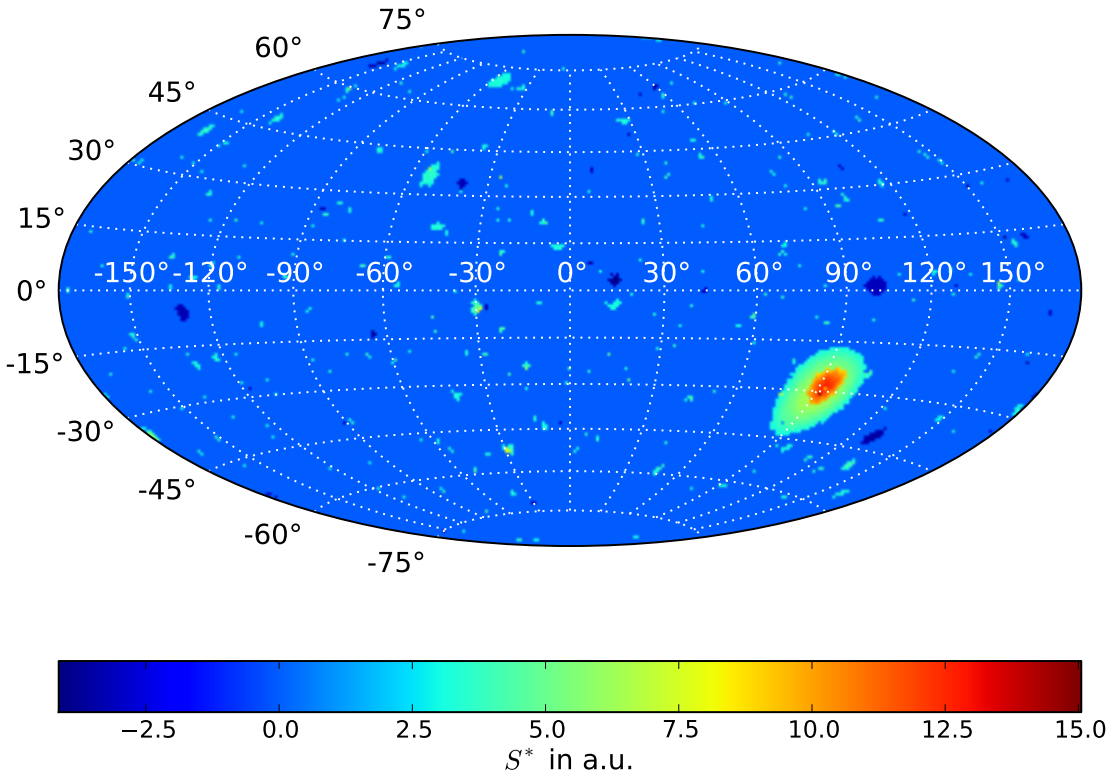


Figure 5.11: Result of a needlet analysis with an 3σ threshold cut. The sample contained 150 signal events (gaussially distributed at 90° right ascension and −30° declination with a smearing of 6°) and 29850 background events. The needlet parameters used for this analysis were B = 2.3, jmax = 5.

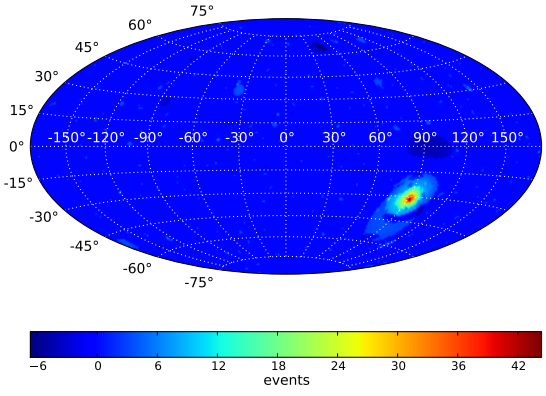


Figure 5.12: Result of a needlet analysis for a source with 3° smearing, simulated with 150 signal events, 29850 background events. The needlet parameters used for this analysis were B = 2.3, jmax = 5.

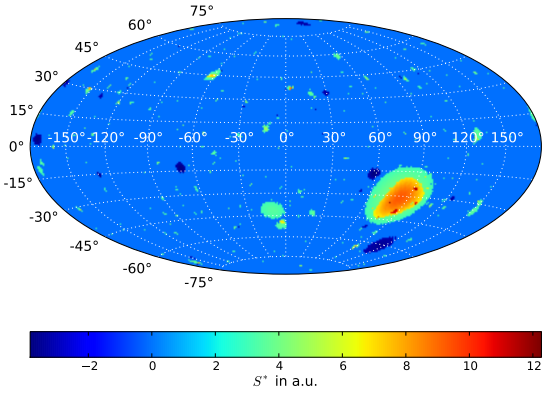


Figure 5.13: Result of a needlet analysis for a source with 12° smearing, simulated with 150 signal events, 29850 background events. The needlet parameters used for this analysis were B = 2.3, jmax = 5.

5.2.3 Log-Likelihood fit

Figure 5.14 shows the reconstructed signal in a two-dimensional histogram. A log-likelihood fit of a two dimensional Gaussian function was done to verify if the source was reconstructed at the right position. Table 5.3 shows the result of a Log-Likelihood Fit applied to the result of a needlet analysis with the parameters B = 2.3 and jmax = 5, compared with the simulated location and smearing of the source. The source has been successfully reconstructed at the simulated coordinates, the smearing also corresponds with the simulation within the uncertainties.

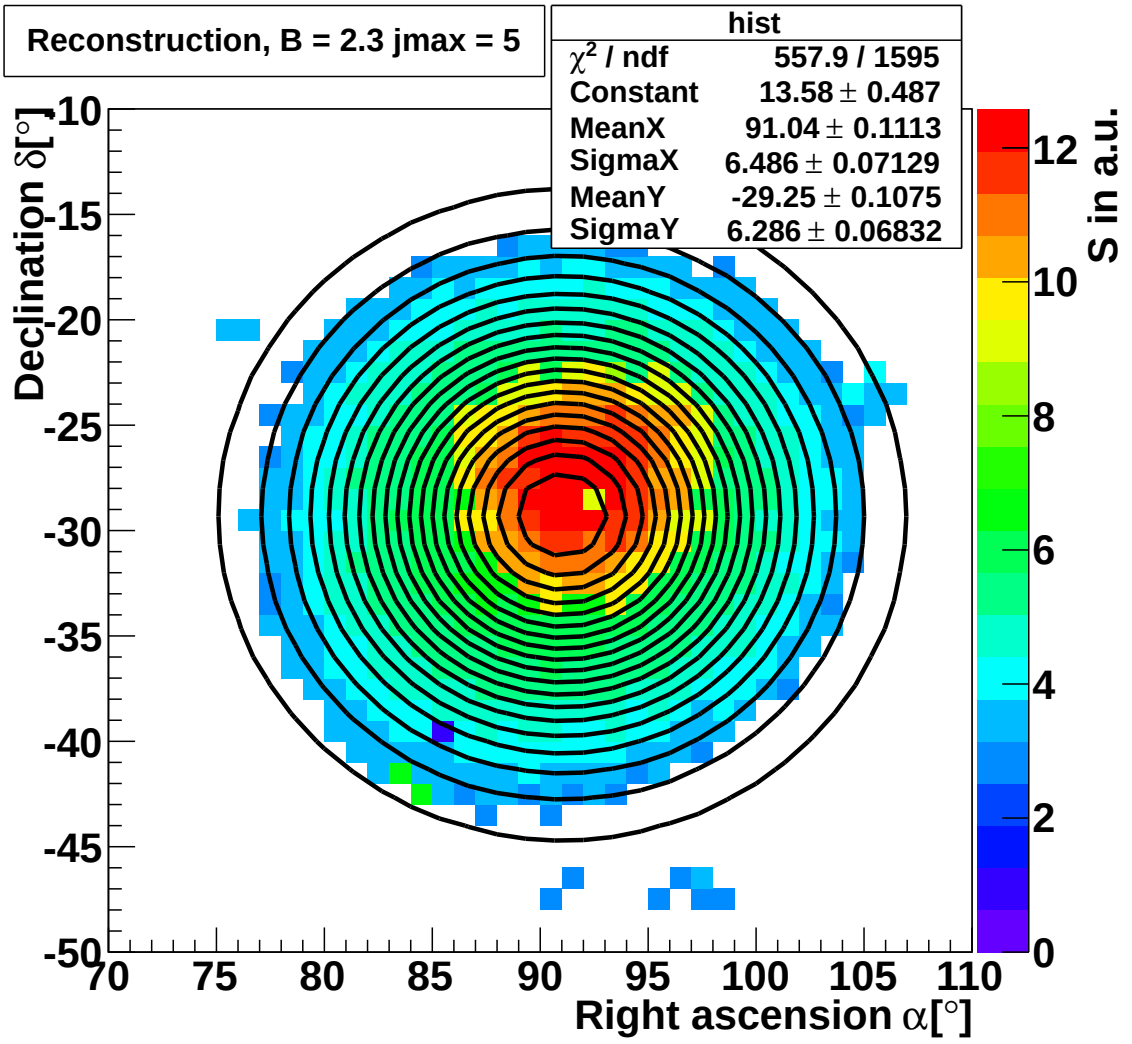


Figure 5.14: Reconstructed source with log-likelihood fit of a two dimensional Gauss, showing the result of a needlet analysis with an 3σ threshold cut. The sample contained 150 signal events (gaussially distributed at 90° right ascension and −30° declination with a smearing of 6°) and 29850 background events. The needlet parameters used for this analysis were B = 2.3, jmax = 5. The source has been successfully reconstructed at the simulated coordinates, with the expected smearing.

	Simulation	Reconstruction
right ascension	90°	$(91.04 \pm 0.29)°$
declination	−30°	$(-29.25 \pm 0.29)°$
smearing	6°	$(6.38 \pm 0.28)°$

Table 5.3: Simulation of a Gaussian source with 150 signal entries, 29850 noise entries, compared to a Log-Likelihood fit of a two-dimensional Gaussian of the analysed sample with the needlet parameters B = 2.3, jmax = 5 or a map containing 49152 pixels. The uncertainty includes the uncertainty from the fit and the angular resolution divided by $\sqrt{12}$.

5.3 Reconstruction of the simulated source position

The reconstruction of the source position for sources with different properties has been studied on examples. The position of the reconstructed source as well as the smearing is studied for different resolutions. The different resolutions (49152 pixels and 12288 pixels) were compared, as well as different smearings. An average over 500 analysed MC maps was calculated. For an example of 150 source events, 29850 noise events, figure 5.15 shows the right ascension of the reconstructed source. For both resolutions, the reconstructed right ascension corresponds well with the expected position of 90°. For the declination shown in figure 5.16, the reconstructed position also corresponds to the simulated position of −30°. For both coordinates, the reconstruction for both resolutions was successful.

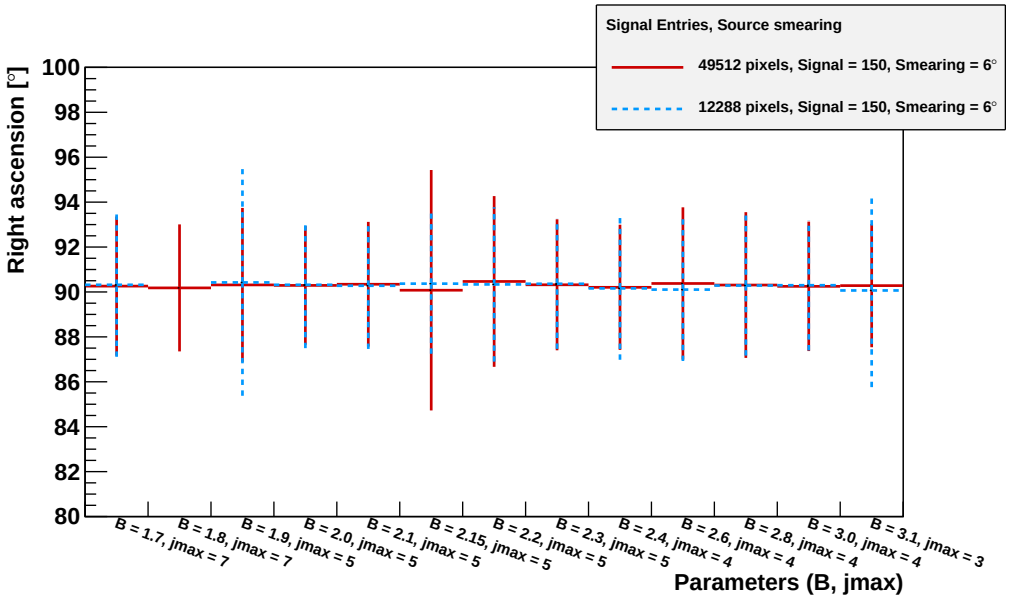


Figure 5.15: Reconstructed right ascension for different resolutions. For both resolutions (49152 and 12288 pixels), the reconstructed right ascension corresponds well with the expected position of 90°.

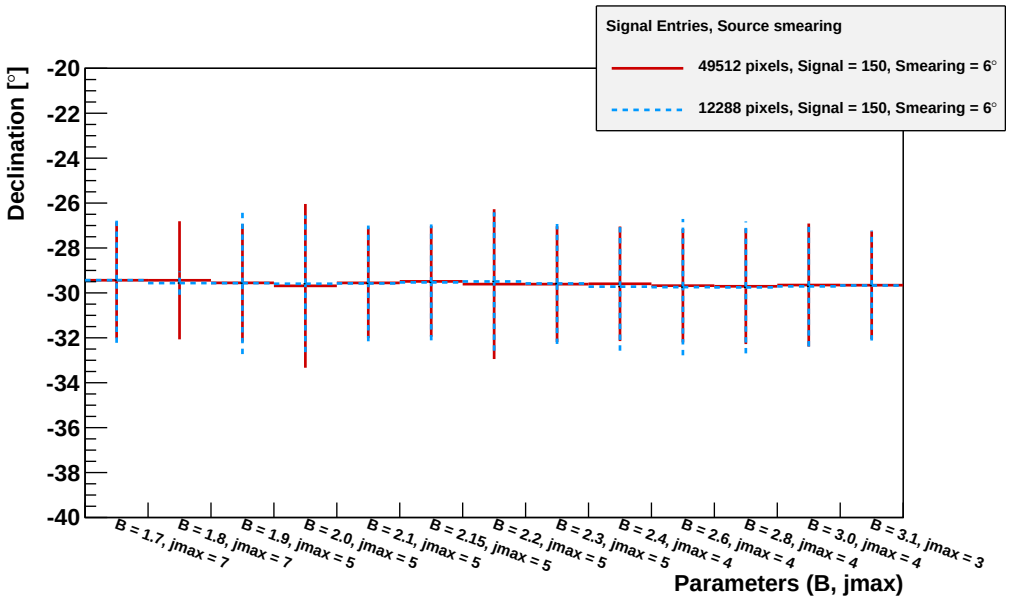


Figure 5.16: Reconstructed declination for different resolutions. For both resolutions (49152 and 12288 pixels), the reconstructed declination corresponds well with the expected position of $-30°$.

In figure 5.15 and 5.18, sources of different smearings have been successfully reconstructed at the expected position, although for a smearing of 12°, the spatial reconstruction is not as precise as for the the smearings of 3° and 6°. This can be explained by the fact that 150 signal events distributed along a Gaussian with a smearing of 12° leads to 150 events spread in a large area, which complicates a precise reconstruction.

Figures 5.19 and 5.20 show the reconstructed smearing of the Gaussian source. In figure 5.19, the expected smearing of 6° was not successfully reconstructed, but rather a smearing of 8° to 9°. Nevertheless, for the reconstruction of the smearing, no significant discrepancies between the two resolutions occurred. In figure 5.20, the reconstructed smearing of sources with a Gaussian smearing of 3°, 6° and 12° has been studied. Sources with a simulated smearing of 3° have been reconstructed as 7° to 8°, whereas sources with a simulated smearing of 12° have been reconstructed with approximately the expected size. The result of a wavelet analysis for a simulated source with a smearing of 12° is shown in figure 5.13. This figure shows that the Log-likelihood Fit might be over-sensitive to sources with a smearing 7° to 9°. For a simulated source with a smearing of 3°, the wavelet analysis clearly yields a small source, as can be seen in figure 5.12, so that it can be stated that the bad results for the reconstructed smearing of a small source as due to the Log-Likelihood fit, and not due to the wavelet analysis.

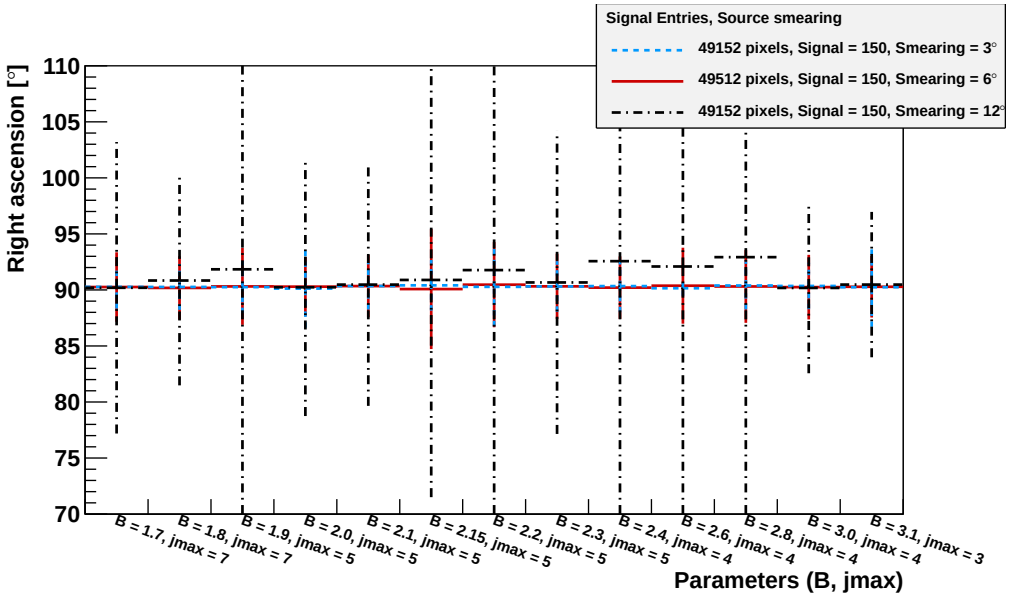


Figure 5.17: Successful reconstructed right ascension for smearings of 3° and 6°. For a smearing of 12°, the reconstructed position of the source has higher uncertainties and does not coincides perfectly with the simulated position.

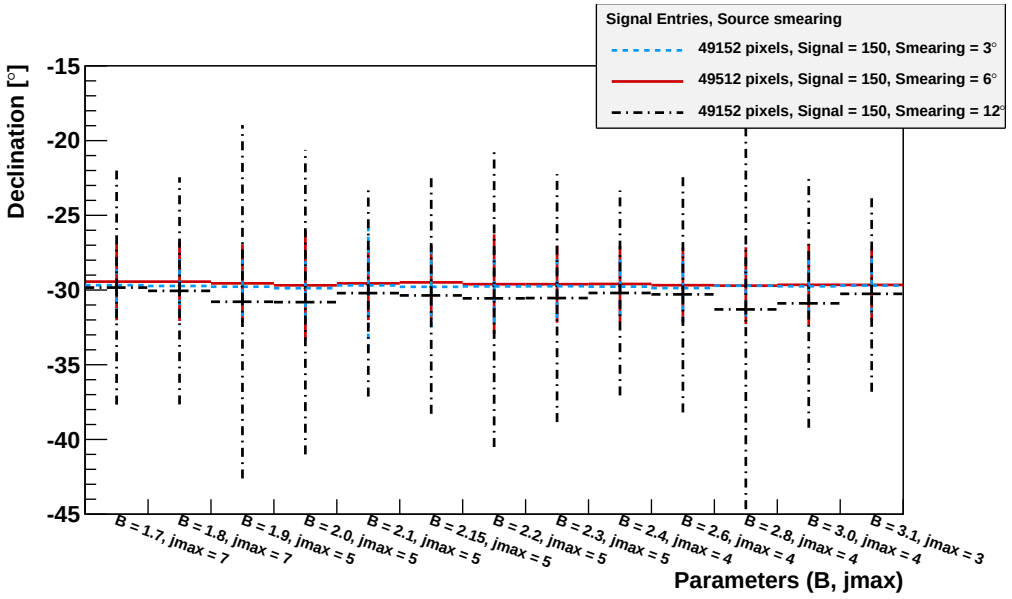


Figure 5.18: Successful reconstructed declination for the different smearings 3° and 6°, although for a smearing of 12°, the reconstructed position of the source is not as precise as for the other smearings and has a higher uncertainty.

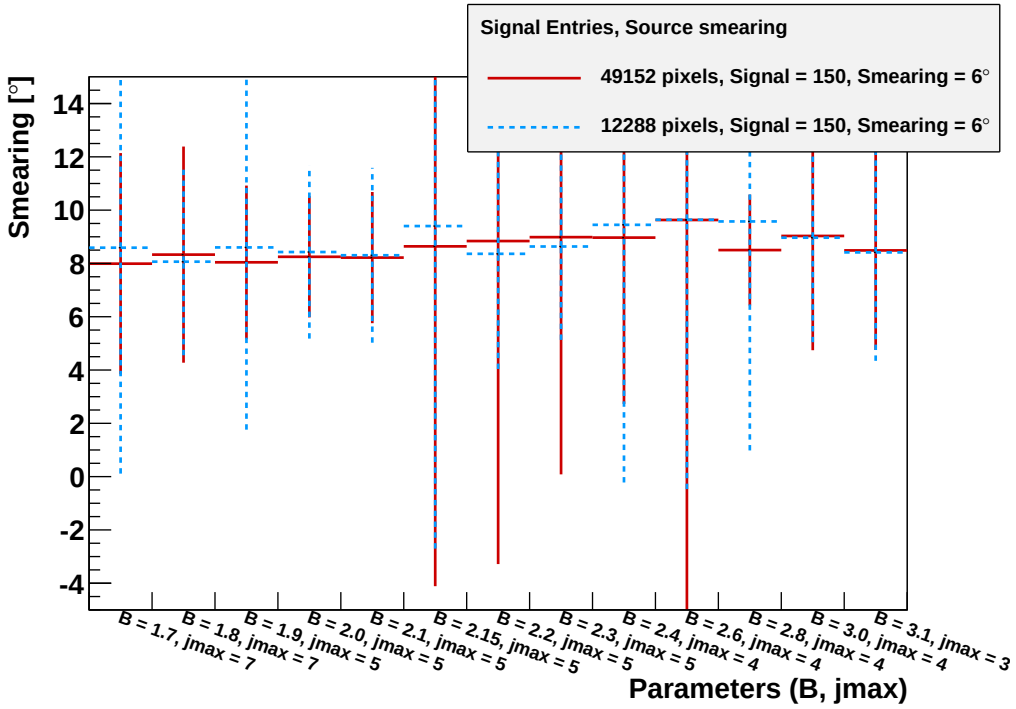


Figure 5.19: Reconstructed smearing for different resolutions. No difference in the reconstructed smearing between the two resolutions can be seen, although the reconstructed smearing is overestimated in all needlet scales.

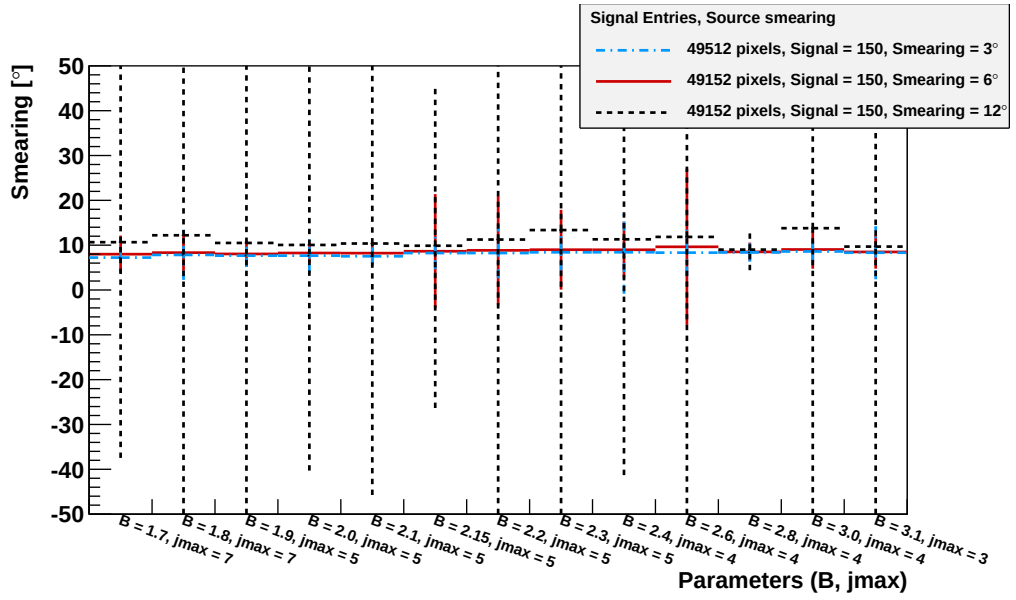


Figure 5.20: Reconstructed smearing for different smearings. For a source with a smearing of 3°, the reconstructed smearing has been severely overestimated in all the needlet scales, this is a feature of the Log-Likelihood fit, not of the wavelet analysis. The source simulated with a smearing of 12° has been reconstructed to approximately the right size, but with a large uncertainty.

5.4 Signal to Noise Ratio

The Signal to Noise Ratio (SNR) was used to compare the result of the wavelet analysis for sources with different properties as well as for the different parameters.

$$SNR = \frac{\langle S_{Bin_{Max}} \rangle}{\langle RMS \rangle} \tag{5.6}$$

with $\langle S_{Bin_{Max}} \rangle$ being the maximum significance in one bin associated to the source in the analysed map, $\langle RMS \rangle$ being the root mean square (RMS) of the analysed background map.

To calculate this RMS, a map containing only noise events was simulated. The exact analysis as done to a map containing a source was done. The root mean square of the obtained map was calculated. 500 of these maps were analysed and the mean with standard deviation was calculated.

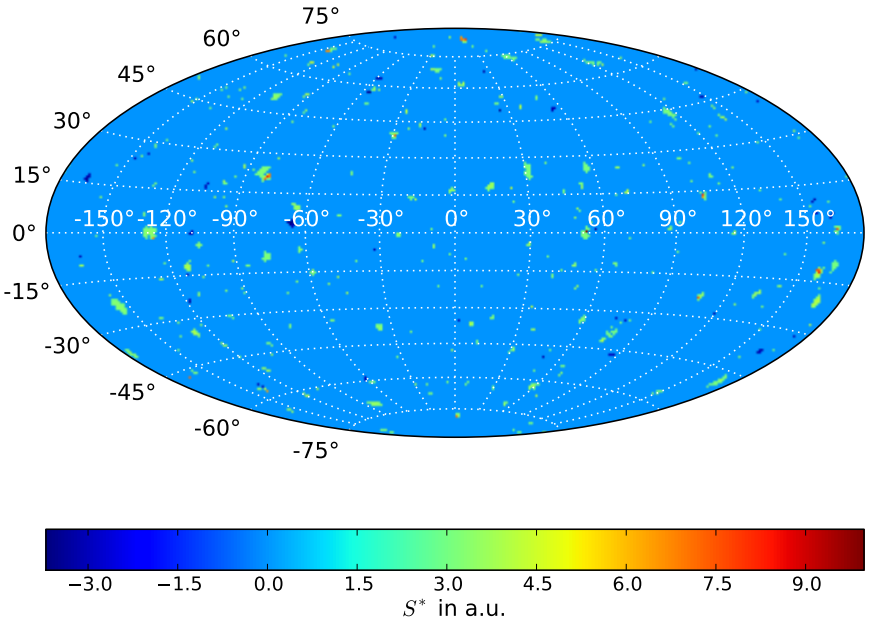


Figure 5.21: Noisemap after analysis and thresholdcut, RMS = 0.618. The isotropy map was analysed using the needlet parameters B = 2.3 and jmax = 5.

6. Results

The needlet analysis has been performed for Gaussian sources of different smearings and various numbers of signal events. The thus created Monte Carlo maps were analysed with different needlet coefficients B and jmax. Figure 6.1 shows the Signal to Noise Ratio plotted against the needlet parameters.

Figure 6.1 shows the results of the different numbers of signal entries used. The Signal to Noise Ratio is highest for maximum signal entries (200) and minimum smearing (3°), as naively expected.

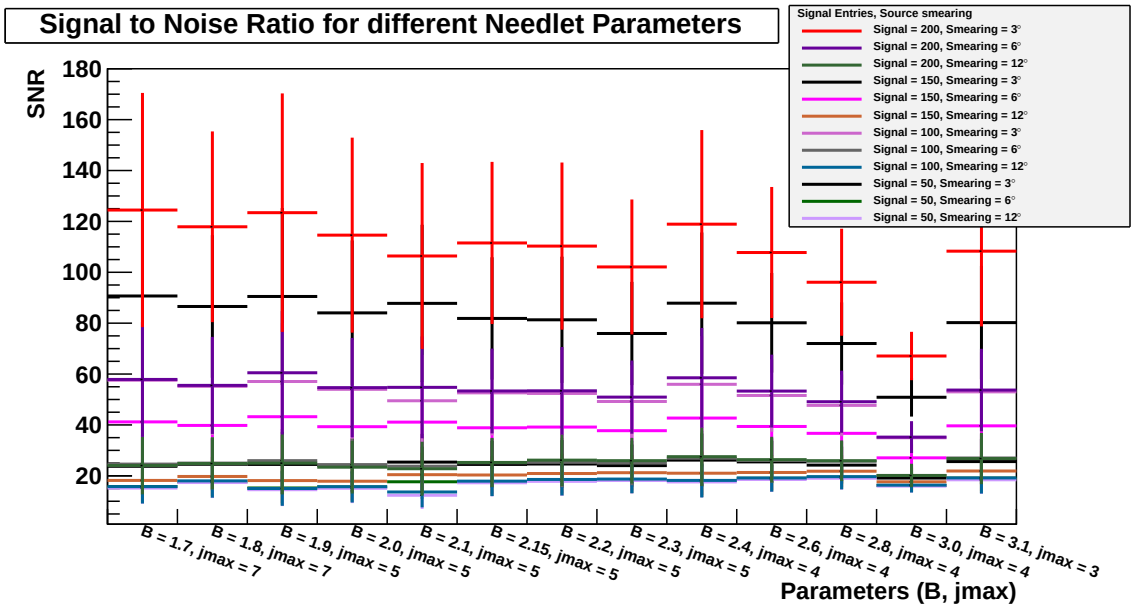


Figure 6.1: Signal to Noise Ratio for different needlet parameters B and jmax. The SNR is best for high number of signal events and small smearing. An overall feature can be observed for nearly all sources: for small needlet with (up to B = 2.6), the SNR is nearly in the same level. For B = 2.8, jmax = 4 and B = 3.0, jmax = 4, a clear decrease in the SNR can be observed. The SNR rises again for the last parameter set, B = 3.1, jmax = 3. The needlet scale j = 4 which has not been considered for the last parameter set, causes a lot of fluctuations in the analysed map, as well as in the analysed noise map. This leads to a higher $\langle RMS \rangle$ of the analysed noise map, and so to a lower SNR in the bins B = 2.8, jmax = 4 and B = 3.0, jmax = 4

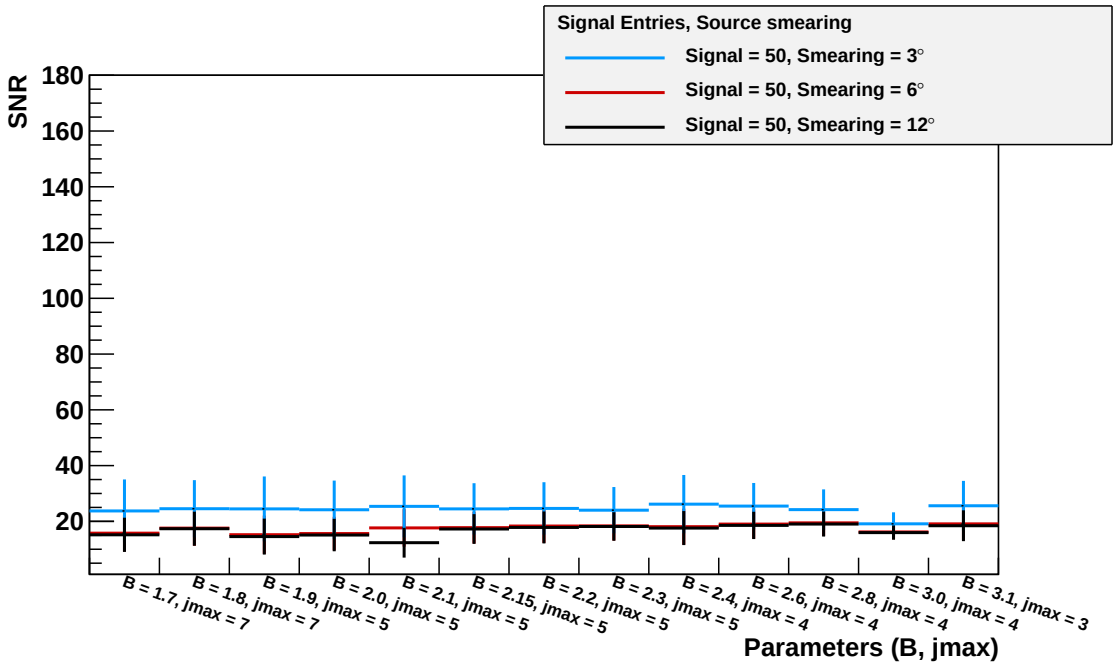


Figure 6.2: Signal to Noise Ratio for different needlet parameters B and jmax, for MC-maps simulated with 50 signal events. The SNR is best for small smearing, for a smearing of 6° and 12°, no significant difference regarding the SNR is visible.

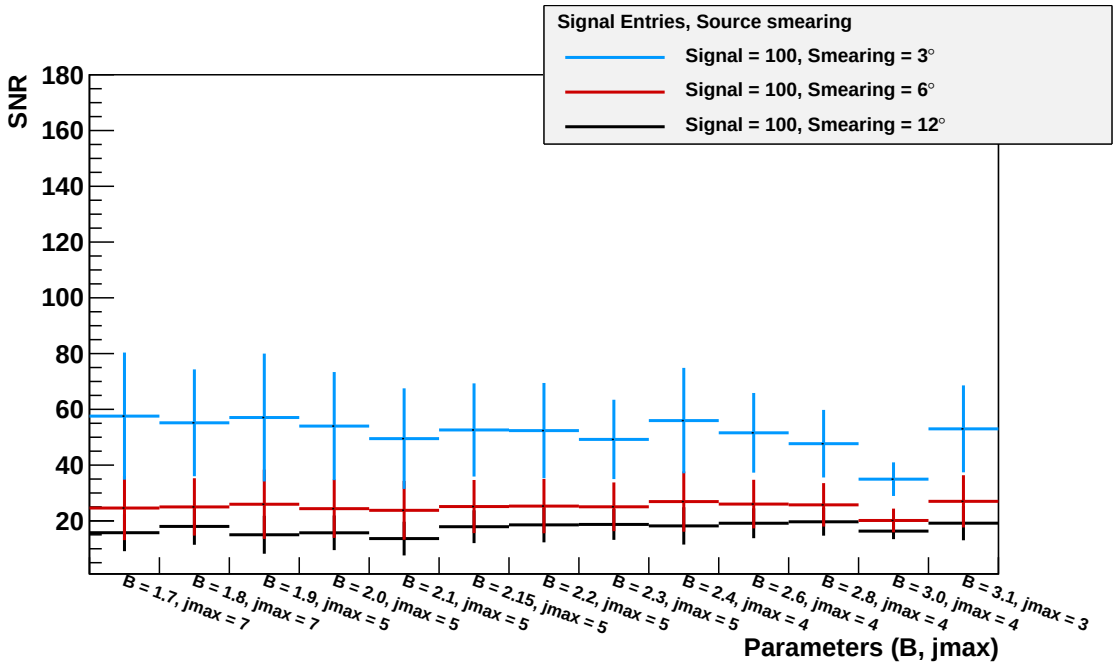


Figure 6.3: Signal to Noise Ratio for different needlet parameters B and jmax for MC-maps simulated with 100 signal events. Smaller smearing leads to a higher SNR.

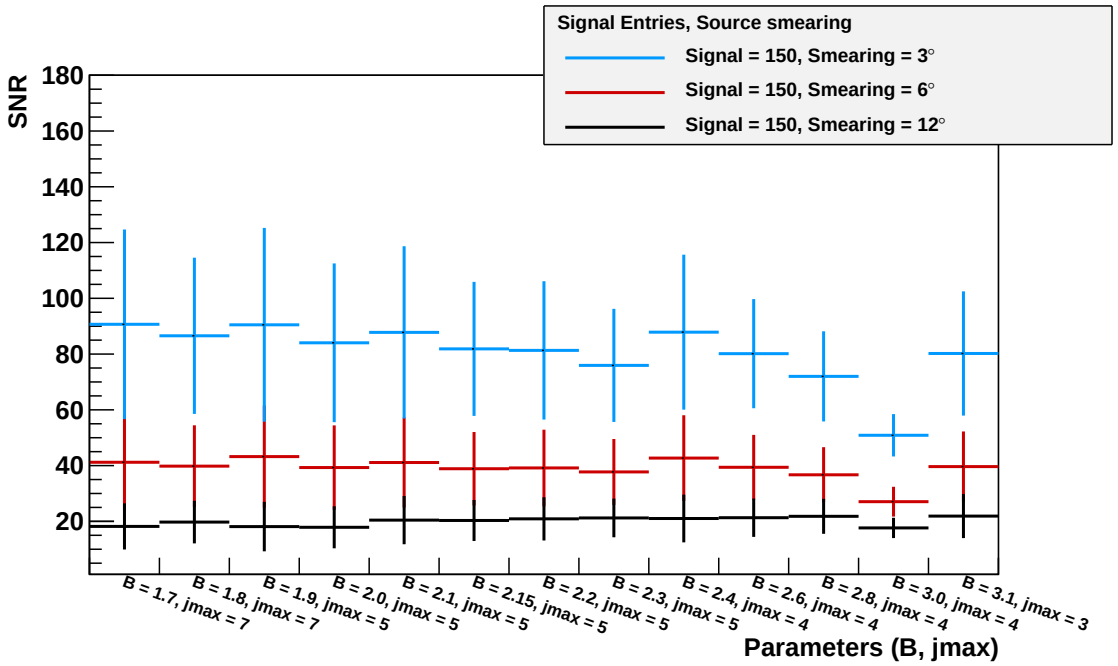


Figure 6.4: Signal to Noise Ratio for different needlet parameters B and jmax, for analysed MC-maps simulated with 150 signal events. As in figure 6.3, a smaller smearing leads to a higher SNR.

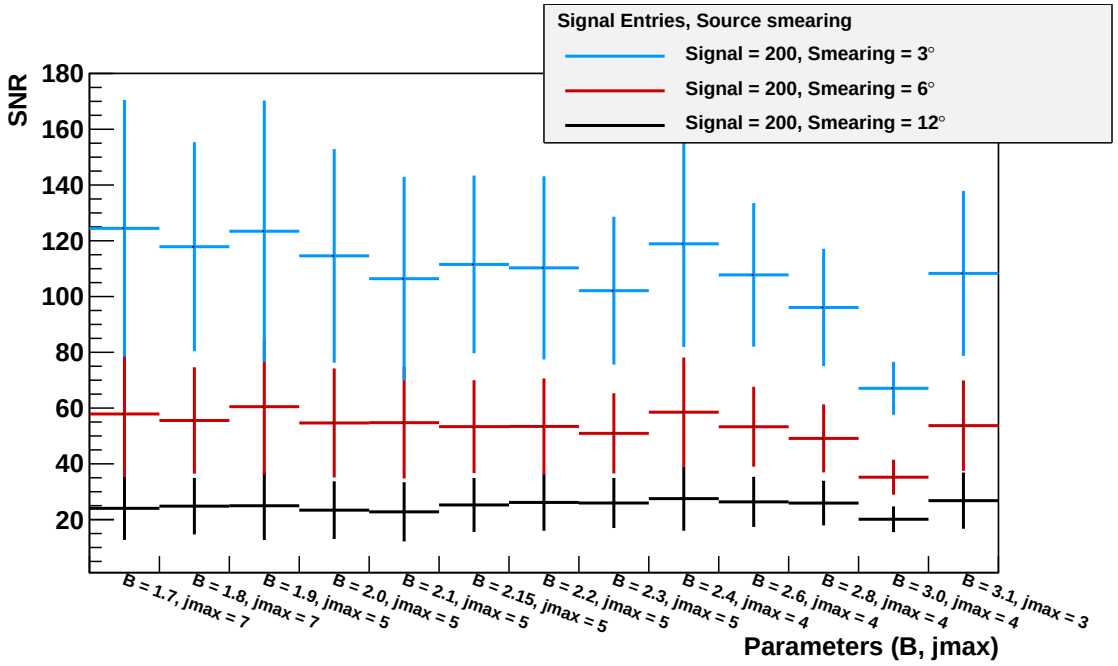
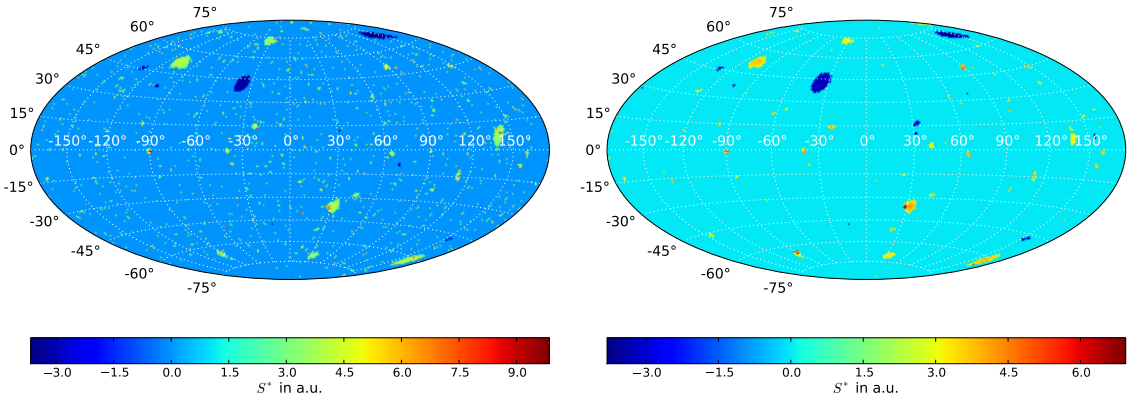


Figure 6.5: Signal to Noise Ratio for different needlet parameters B and jmax, for analysed MC-maps simulated with 200 signal events. A small smearing (3°) and the high number of signal events (200) leads to a maximum SNR.

For higher B values, one observes a decay of the SNR. It rises again for the last parameter set, B = 3.1, jmax = 3. The main difference between the second to last (B = 3.0, jmax = 4) and the last parameter set is that the needlet scale j = 4 has been omitted.

In figure 6.6, a noise sample is analysed by two different needlet parameter sets (B = 3.0, jmax = 4, in comparison to B = 3.1, jmax = 3). In the first figure, considerably more isotropic fluctuations are visible than in the second figure. This is due to the additional needlet scale, sensitive to very small anisotropies. A higher $\langle RMS \rangle$ has been calculated for the map containing 5 needlet scales (scale zero to scale 4) which leads to a lower SNR.



(a) Noise sample analysed with needlet parameters B = 3.0, jmax = 4, RMS = 0.708

(b) Noise sample analysed with needlet parameters B = 3.1, jmax = 3, RMS = 0.409

Figure 6.6: Noise sample after analysis and threshold cut. For the needlet parameters B = 3.0, jmax = 4, a RMS of 0.708 is calculated for this map, whereas for the parameters B = 3.1, jmax = 3, a lower RMS of 0.409 is calculated. The additional needlet scale (j = 4) causes additional isotropic fluctuations in the analysed noisemap, which leads to a higher RMS and thus a lower SNR.

Comparison of the Signal to Noise Ratio for different resolutions

Figure 6.7 shows the SNR for different needlet parameters at a fixed intensity (150 signal events) for different resolutions, 49152 pixels (angular resolution: 0.92°) in comparison to 12288 pixels (angular resolution: 1.83°). For all bins except the one for B = 3.0, jmax = 4, no significant SNR discrepancy is visible between the different resolutions. For B = 3.0, jmax = 4, the SNR for lower resolution is a lot better than for higher resolution. Figure 6.8 shows analysed noise maps of different resolution (49152 pixels and 12288 pixels). The RMS for the lower resolution is lower compared to the higher resolution, as a lot more fluctuations can be seen in the higher resolution map. This is due to the fact that different standard deviations are needed for the threshold cut of these maps. Figures 6.9 show the power β for two different resolutions. This power has been obtained out of isotropy MC-samples containing 30000 background events. The power of each pixel (for a needlet width

B = 2.3 and a needlet scale j = 2 shown in this example) has a Gaussian shape. A Gaussian χ^2 fit is performed on the histograms containing the power values. As the high resolution map has a lot more pixels, but the same amount of events as the lower resolution map, more pixels contain no events at all, which leads to a lot more entries in the bins around $\beta = 0$ bin in the histogram. This results in a smaller standard deviation for the isotropy MC-sample on a map with higher resolution. The calculated significance S_{jk} is higher for a lower standard deviation, therefore in the map containing 49152 pixels, more pixels survive the $S_{jk} \geq 3$ cut. This causes the fluctuations seen in figure 6.8 (a), which leads to a higher RMS and thus a lower SNR for a map containg 49152 pixels as comparison to 12288 pixels.

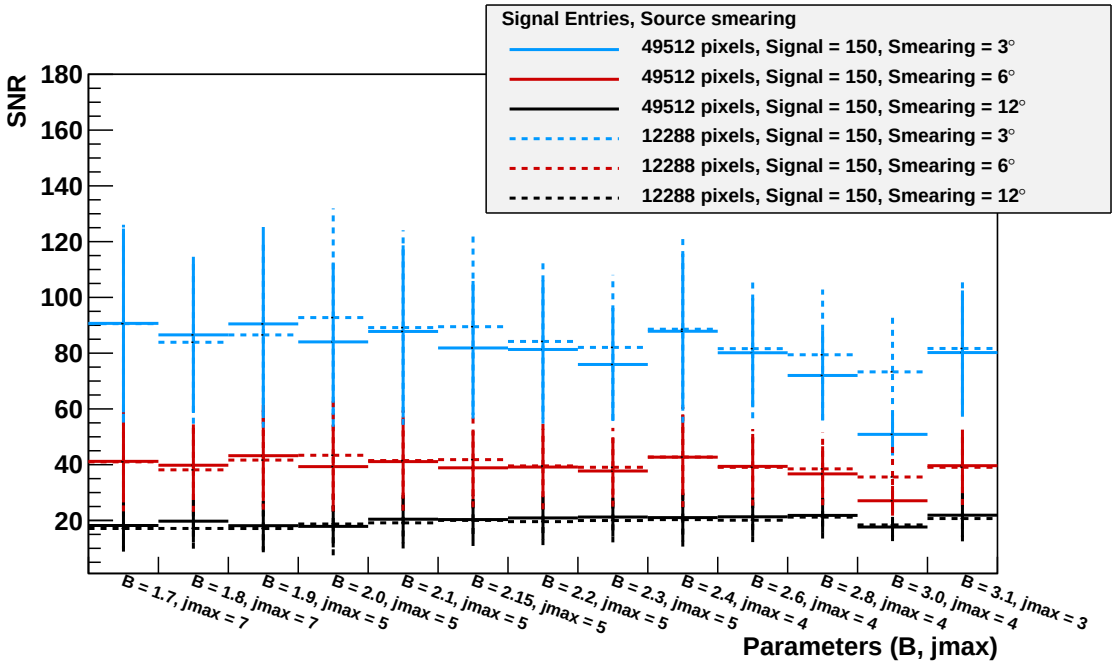
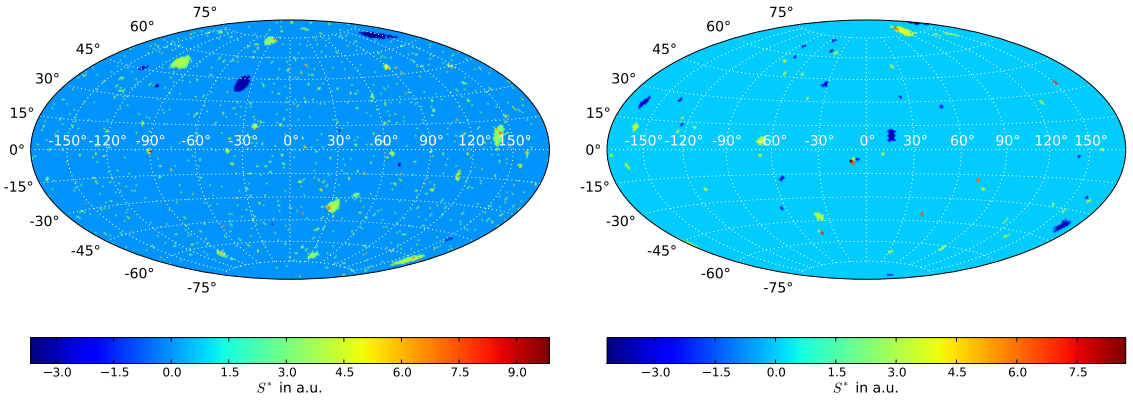


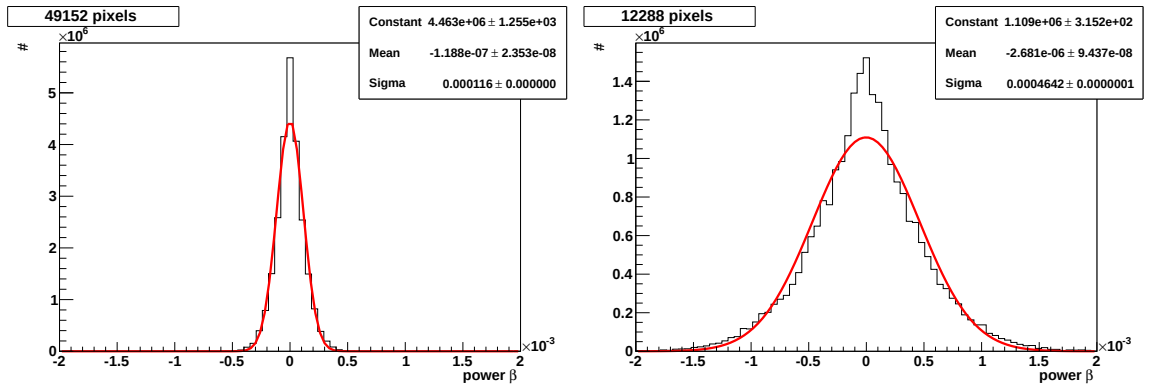
Figure 6.7: Signal to Noise Ratio for different resolutions. For all bins except the one for B = 3.0, jmax = 4, no significant SNR difference is visible. For B = 3.0, jmax = 4, the SNR for lower resolution is a lot better than for higher resolution.



(a) Noise sample containing 49152 pixels, RMS = 0.675

(b) Noise sample containing 12288 pixels, RMS = 0.499

Figure 6.8: Noise sample of different resolutions after analysis and threshold cut. For a higher resolution (49152 pixels), a higher RMS is calculated than for the lower resolution (12288 pixels). The standard deviation needed for the threshold cut of these maps differ for the two resolutions, this is explained in figure 6.9.



(a) Gaussian fit on MC-isotropy map containing 49152 pixels

(b) Gaussian fit on MC-isotropy map containing 12288 pixels

Figure 6.9: Gaussian fit on the power calculated out of MC-isotropy maps of different resolutions. Due to the fact that 30000 events were randomly distributed in the maps of different resolutions, the map containing 49152 pixels has a lot of empty pixels compared to the map containing 12288 pixels. This leads to many entries in the bin for a power of about zero in the histogram for a higher resolution and thus a smaller standard deviation in the fit. This leads to a higher significance S_{jk} in the analysed noise map an so to more pixels passing the $S_{jk} \geq 3$ condition. The RMS of a map containing 49152 pixels is higher than for one containing 12288 pixels, and thus the SNR is better for the low resolution map.

7. Conclusion

In this Bachelor thesis, the wavelet analysis was successfully used to find anisotropies in maps generated by Monte Carlo. One Gaussian source of varying intensity and smearing was simulated and isotropic noise was added, in maps with 49152 (angular resolution: 0.92°) or 12288 pixels (angular resolution: 1.83°). An analysis using the needlet wavelet was applied to the simulated maps, separating the MC-data into different needlet scales. The needlet analysis yielded a power β_{jk} in arbitrary units, which was converted into a significance S_{jk} by comparing the resulting power of the analysed map to the one of maps containing isotropy. Furthermore a threshold cut was applied to the significance, setting all pixels to zero whose entries are within 3σ conform with isotropy, yielding the significance after the cut, S^*_{jk}. As the different needlet scales represent sensitivities to anisotropies of different scales, and for a study on real data no a priori information about the size of the possible anisotropies is known, all needlet scales are added up. The simulated source was found by this analysis and could be clearly distinguished.

To be able to compare the result of the wavelet analysis for different parameters, the Signal to Noise Ratio was introduced. For simulated sources having a high intensity and a small smearing, the best Signal to Noise Ratio was calculated. As for the different needlet parameters, it was shown that the wavelet analysis works for all the tested parameters, whereas smaller needlet widths yield slightly better results. It is noteworthy that higher needlet scales cause isotropic fluctuations in the resulting map, as all needlet scales are added up and these high needlet scales are not sensitive to the size of simulated source.

A lower resolution of the map did not cause a distinguished change for the resulting map as well as for the Signal to Noise Ratio. For higher needlet widths and scales, the lower resolution enhances the wavelet analysis, as less isotropic fluctuations are present in the higher needlet scales.

In a further study, it might be interesting to look at the different needlet scales instead of summing them up. For a knows size of an astrophysical object suspected to emit ultra high energy cosmic rays, the one needlet scale sensitive to this size could only be considered.

Bibliography

[1] *Wavelet Analysis for Geophysical Applications*, Praveen Kumar and Efi Foufoula-Georgiou, Reviews of Geophysics, VOL. 35, NO. 4, PP. 385-412, 1997, doi:10.1029/97RG00427

[2] *Über Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten*, Victor F. Hess, *Physik. Zeitschr.* XIII,1912

[3] *Extensive Cosmic-Ray Showers*, Pierre Auger, P. Ehrenfest, R. Maze, J. Daudin, Robley and A. Fréon *Rev. Mod. Phys.*,11 1939

[4] K. Nakamura et al., *J. Phys.* G 37, 075021(2010) and 2011 partial update for the 2012 edition

[5] *Extensive Air Showers High Energy Phenomena and Astrophysical Aspects A Tutorial, Reference Manual and Data Book*, Peter K.F. Grieder, modified figure

[6] http://www.physics.utah.edu/ whanlon/spectrum.html visited on 11.07.2012

[7] *Correlation of the highest energy cosmic rays with nearby extragalactic objects*, The Pierre Auger Collaboration 2007, *arXiv:0711.2256v1 [astro-ph]*

[8] *Analysis of possible sources and acceleration mechanisms of ultra high energy cosmic rays*, Sarah Schmetkamp, Bachelor thesis, RWTH-Aachen,Institute IIIa, 2009

[9] http://jemeuso.riken.jp/en/about2.html visited on 14.06.12

[10] *Properties of the surface detector data recorded by the Pierre Auger observatory*, Stefan Grebe, Master thesis, Universität Siegen, 2008

[11] http://www.auger.org/observatory/ visited on 14.06.12

[12] http://www.auger.de visited on 14.06.12

[13] *AERA The Auger Engineering Radio Array*, Klaus Weidenhaupt, 2012, http://www.physik.rwth-aachen.de/institute/institut-iiia/mitarbeiter/personen/martin-erdmann/conference-talks/ visited on 24.07.2012

[14] http://www.auger.de/public/sd.html visited on 26.07.2012

[15] http://www.auger.de/public/fd.html visited on 26.07.2012

[16] *Data Acquisition, Triggering, and Filtering at the Auger Engineering Radio Array*, J. L. Kelleya, for the Pierre Auger Collaboration, May 2012, *arXiv:1205.2104v1 [astro-ph.IM]*

[17] *AMIGA, Auger Muons and Infill for the Ground Array*, Pierre Auger Collaboration A. Etchegoyen, *arXiv:0710.1646v1 [astro-ph]*

[18] *HEAT - a low energy enhancement of the Pierre Auger Observatory*, C. Meurer, N. Scharf on behalf of the Pierre Auger Collaboration, *arXiv:1106.1329v1 [astro-ph.IM]*

[19] *Wavelets applied to the Detection of Point Sources of UHECRs*, Rafael A. Batista, Ernesto Kemp, Rogerio M. de Almeida, Bruno Daniel, *arXiv:1012.3645v2 [astro-ph.IM]*

[20] *Detection of point sources in cosmic ray maps using the mexican hat wavelet family*, Rafael Alves Batista, Ernesto Kemp, Bruno Daniel, *arXiv:1201.2799v1 [astro-ph.IM]*

[21] *Large Scale Anisotrophy Studies of Ultra High Energy Cosmic Rays Using Date Taken with the Surface Detector of the Pierre Auger Observatory*, Marius Grigat, PhD thesis, RWTH Aachen, 2011

[22] *HEALPix - a Framework for High Resolution Discretization, and Fast Analysis of Data Distributed on the Sphere*, K.M. Gorski, E. Hivon, A.J. Banday, B.C. Wandlet, F.K. Hansen, M.Reinecke, M.Bartelman. *arXiv:astro-ph/0409513v1*

[23] *Wavelet analysis and the detection of nongaussianity in the CMB*, M. Hobson, A. Jones, and A. Lasenby, *astro-ph/9810200*

[24] *Detection of non-Gaussianity in the WMAP 1 - year data using spherical wavelets*, P. Vielva, E. Martinez-Gonzalez, R. Barreiro, J. Sanz, and L. Cayon, *arXiv:astro-ph/0310273v2*

[25] *Wavelet Analysis of the Arrival Directions of Ultra High Energy Cosmic Rays at the Pierre Auger Observatory*, Matthias Plum, Diploma thesis, RWTH Aachen, 2011

[26] *Spherical Needlets for CMB Data Analysis*, D. Marinucci, D. Pietrobon, A. Balbi, P. Baldi, P. Cabella G. Kerkyacharian, P. Natoli D. Picard, N. Vittorio, *arXiv:0707.0844v1 [astro-ph]*

[27] *Asymptotics for spherical needlets*, P. Baldi, G. Kerkyacharian, D. Marinucci, D. Picard, *arXiv:math/0606599v2 [math.ST]*

[28] *Needlet Detection of Features in WMAP CMB Sky and the Impact on Anisotropies and Hemispherical Asymmetries*, Davide Pietrobon, Alexandre Amblard, Amedeo Balbi, Paolo Cabella, Asantha Cooray, Domenico Marinucci, 2008, *arXiv:0809.0010v2 [astro-ph]*

A. Appendix

A.1 List of abbreviations

AGN	Active Galactic Nucleus
AERA	Auger Engineering Radio Array
AMIGA	Auger Muon-detectors and Infill for the Ground Array
CMB	Cosmic Microwave Background
FD	Fluorescence Detector
GRB	Gamma-Ray Burst
HEAT	High Elevation Auger Telescope
PMT	Photomultiplier tube
SD	Surface Detector
SNR	Signal to Noise Ratio
UHECR	Ultra-High Energy Cosmic Ray
WMAP	Wilkinson Microwave Anisotropy Probe

A.2 Multipole moments used for different needlet parameters

Parameters	j = 0	j = 1	j = 2	j = 3	j = 4	j = 5	j = 6	j = 7
B = 1.7 jmax = 7	1 to 1	2 to 2	2 to 4	3 to 8	5 to 14	9 to 24	15 to 41	25 to 69
B = 1.8 jmax = 7	1 to 1	2 to 3	2 to 5	4 to 10	6 to 18	11 to 34	19 to 61	35 to 110
B = 1.9 jmax = 5	1 to 1	2 to 3	2 to 6	4 to 13	7 to 24	14 to 47		
B = 2.0 jmax = 5	1 to 1	2 to 3	3 to 7	5 to 15	9 to 31	17 to 63		
B = 2.1 jmax = 5	1 to 2	2 to 4	3 to 9	5 to 19	10 to 40	20 to 85		
B = 2.15 jmax = 5	1 to 2	2 to 4	3 to 9	5 to 21	10 to 45	22 to 98		
B = 2.2 jmax = 5	1 to 2	2 to 4	3 to 10	5 to 23	11 to 51	24 to 113		
B = 2.3 jmax = 5	1 to 2	2 to 5	3 to 12	6 to 27	13 to 64	29 to 148		
B = 2.4 jmax = 4	1 to 2	2 to 5	3 to 13	6 to 33	14 to 79			
B = 2.6 jmax = 4	1 to 2	2 to 6	3 to 17	7 to 45	18 to 118			
B = 2.8 jmax = 4	1 to 2	2 to 7	3 to 21	8 to 61	22 to 172			
B = 3.0 jmax = 4	1 to 2	2 to 8	4 to 26	10 to 80	28 to 242			
B = 3.1 jmax = 3	1 to 3	2 to 9	4 to 29	10 to 92	10 to 92			

Figure A.1: Multipole moments used for the different needlet scales

Erklärung

Ich versichere, dass ich die Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie Zitate kenntlich gemacht habe.

Aachen, den 30. Juli 2012

Ania Koob