Gamma/neutron separation in the Martian radiation environment

Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von
Jan Köhler

Kiel im November 2011

Referent/in :	Robert F. Wimmer-Schweingruber
Korreferent/in :	Wolfgang J. Duschl
Tag der mündlichen Prüfung :	26.01.2012
Zum Druck genehmigt :	26.01.2012

Der Dekan

Abstract

The Mars Science Laboratory is a NASA Mars-rover mission which was launched in November 2011, and will land on Mars in August 2012. One part of the rover is the Radiation Assessment Detector (RAD) which will measure a broad spectrum of energetic particles at the surface of Mars. The instrument consists of a silicon detector telescope stack followed by two scintillators which are enclosed by an anticoincidence. The high-Z cesium-iodide scintillator and the proton-rich plastic scintillator are used for detecting charged as well as neutral particles. In contrast to charged particles, neutral particles can pass the anti-coincidence without triggering a signal and can therefore be distinguished from charged particles. The cesium-iodide scintillator has a high sensitivity to gammas while the plastic scintillator has a high sensitivity to neutrons. However, since cesium-iodide is sensitive to neutrons just as plastic is sensitive to gammas as well, separating gammas and neutrons is a challenging task. The measurement of neutral particles is further complicated by the fact that neutral particles often deposit only a random fraction of their energy in a detector. Therefore, the spectra measured in the scintillators do not reflect the incoming particle spectra.

The scope of this work is to provide methods to separate gamma and neutron measurements of the RAD instrument. In this context, a model was developed to calculate the propagation of solar energetic particles, which can induce an increase of the Martian neutral particle radiation. Based on maximum likelihood estimations, an inversion technique to calculate the unknown Martian neutral particle spectra from measurements and detector response function was developed. A detector response function was calculated with a detailed Monte-Carlo model. The inversion technique was evaluated for different models of underlying measurement statistics, validated with neutron beam calibration measurements and tested for artificial generated measurements of the Martian neutral particle spectra.

The developed inversion techniques are applicable to a wide range of particle detectors and will be used for future evaluation of the RAD neutral particle measurements.

Zusammenfassung

Das Mars Science Laboratory ist eine NASA Rover-Mission, die im November 2011 gestartet ist und voraussichtlich im August 2012 auf dem Mars landen wird. Ein Teil des Rovers ist der Radiation Assessment Detector (RAD), der das Spektrum energetischer Teilchen auf der Marsoberfläche messen wird. Das Instrument besteht aus einem Teleskop aus Siliziumdetektoren gefolgt von zwei Szintillatoren, die von einer Antikoinzidenz eingeschlossen sind. Neutralteilchen können die Antikoinzidenz passieren, ohne ein Signal zu erzeugen und können damit von geladenen Teilchen unterschieden werden. Der Cäsiumiodid-Szintillator hat durch seine hohe Kernladungszahl eine hohe Sensitivität für Gamma-Teilchen, der Plastikszintillator hat eine hohe Protonendichte und somit eine hohe Sensitivität für Neutronen. Da aber Cäsiumiodid ebenfalls sensitiv für Neutronen und Plastik ebenso sensitiv für Gamma-Teilchen ist, ist das Trennen von Gamma-Teilchen und Neutronen eine herausfordernde Aufgabe. Zudem wird die Messung von Neutralteilchen dadurch erschwert, dass diese häufig nur einen zufälligen Anteil ihrer Energie im Detektor deponieren. Die Messung der Szintillatoren enthält also nur indirekte Informationen über das Neutralteilchenspektrum.
Ziel dieser Arbeit ist es, Methoden zur Trennung der Gamma- und Neutronenmessungen von RAD zu entwickeln. In diesem Kontext wurde ein Modell entwickelt, welches die Ausbreitung solarer energiereicher Teilchen beschreibt, die ein wichtiger Faktor für das Teilchenspektrum auf dem Mars sind. Weiter wurde eine Maximum-Likelihood Inversionsmethode entwickelt, mit der das unbekannte Teilchenspektrum aus der Messung und der Detektorantwortfunktion berechnet werden kann. Die Detektorantwortfunktion wurde anhand eines Monte-Carlo-Modells des Detektors erstellt. Die Inversionsmethode wurde für verschiedene Statistikmodelle verglichen, anhand von Kalibrationsmessungen verifiziert und für künstlich erzeugte Messungen der Mars-Teilchenspektren getestet.
Die entwickelten Methoden sind auf eine Vielzahl von Teilchendetektoren anwendbar und werden zur Auswertung der RAD Neutralteilchen Messungen verwendet.

Contents

Chapter 1

Introduction

In this work, methods to calculate the neutral particle spectra from the measurements of the Radiation Assessment Detector (RAD) are developed. RAD is an instrument onboard the National Aeronautics and Space Administration (NASA) Mars Science Laboratory (MSL) mission. In the first part of this work, a model to calculate the propagation of Solar Energetic Particles (SEP), which are an important factor for the Martian neutral particle spectra on Mars, is developed. In the second part of this work, two different inversion approaches, which calculate the neutral particle spectra from the given RAD measurement, are developed and compared.

In this chapter an overview over the MSL mission, the RAD instrument and an introduction to the Martian radiation environment are given, finally the scope of this work is presented.

1.1 Mars Science Laboratory

The MSL is a NASA rover which was launched on November 26, 2011. The main part of the mission is the surface rover Curiosity (Fig. 1.1) which is expected to land in the Gale Crater on Mars in August 2012, where it will autonomously explore the Martian surface. In contrast to previous comparatively small Mars rovers, Curiosity has a mass of 900 kg and is therefore too large for an airbag landing which has been used in previous missions. Instead, the rover is lowered from a hovering platform by the "Sky Crane" system. MSL features a large number of instruments to investigate geology, chemistry, and radiation on the Martian surface and to asses Mars' habitability, i.e. finding out if the Martian environment supports or has ever been able to support microbial life.

The rover uses a Radioisotope Thermoelectric Generator (RTG) and, therefore, does not depend on solar power. However, as discussed in this work, this advantage in power supply comes at the cost of a neutron background caused by the RTG, which interferes with the neutral particle measurements of the RAD instrument.

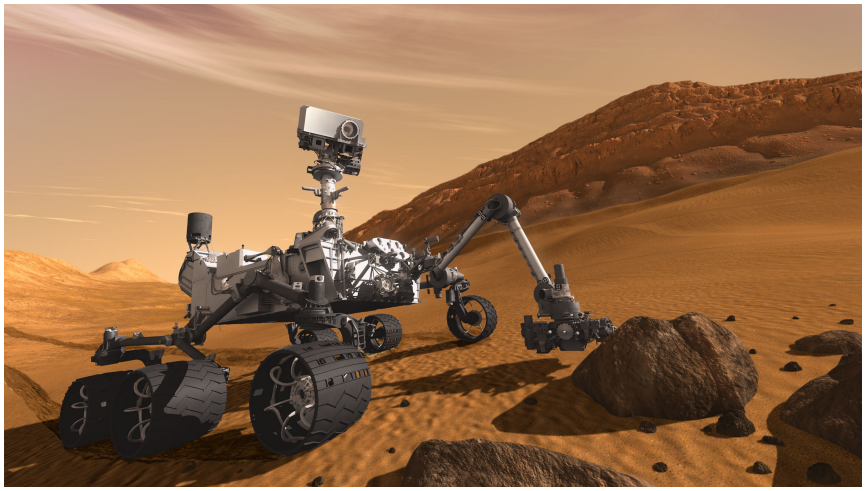

Figure 1.1: Artist's concept of NASA's Mars Science Laboratory rover. The mast or "head" of the rover rises to about 2.1 meters above ground level. The arm can extend up to 2 meters to drill into rocks and collect soil samples. The rover itself has a length of about 3 meters. Provided by *NASA photojournal*.

1.1.1 Scientific goals

To support the overall strategy, *Following the Water*, of NASA's Mars Exploration Program, four main science goals were formulated [NASA MSL Website, 2011]:

- Determine whether life ever arose on Mars
- Characterize the climate of Mars
- Characterize the geology of Mars
- Prepare for human exploration

These four general science goals are realized by the following specific science objectives:

1. Determine the nature and inventory of organic carbon compounds
2. Inventory the chemical building blocks of life (carbon, hydrogen, nitrogen, oxygen, phosphorous, and sulfur)
3. Identify features that may represent the effects of biological processes
4. Investigate the chemical, isotopic, and mineralogical composition of the Martian surface and near-surface geological materials
5. Interpret the processes that have formed and modified rocks and soils

Figure 1.2: The RAD instrument before being integrated into the MSL rover. The red protective cover on top of the instrument was removed before flight. Provided by *NASA photojournal*.

6. Assess long-timescale (i.e., 4-billion-year) atmospheric evolution processes
7. Determine present state, distribution, and cycling of water and carbon dioxide
8. Characterize the broad spectrum of surface radiation, including galactic cosmic radiation, solar proton events, and secondary neutrons

1.2 Radiation Assessment Detector

The RAD instrument (Fig. 1.2), which measures the energetic particle spectrum, onboard the MSL rover was developed and built by a collaboration between the Christian-Albrechts-Universität zu Kiel and the Southwest research Institute in San Antonio.

1.2.1 Scientific goals

The primary goal of RAD is to fully characterize the radiation environment at the surface of Mars. Within this investigation RAD addresses five specific science objectives [Hassler, 2004]:

1. Characterize the energetic particle spectrum at the surface of Mars, including direct and indirect radiation created both in the atmosphere and regolith

2. Determine the radiation dose rate and equivalent dose rates for humans on the surface of Mars

3. Validate Mars atmospheric transmission models and radiation transport codes

4. Determine the radiation hazard and mutagenic influences to life, past and present, at and beneath the Martian surface

5. Determine the chemical and isotopic effects of energetic particles on the Martian surface and atmosphere

The main point addressed in this work is the neutral component of the energetic particle spectrum, for which methods to calculate the spectra from the ambiguous neutral particle measurements of RAD are developed.

1.2.2 Instrument concept

The RAD instrument can be divided into the RAD Electronics Box (REB), which has been built by Southwest Research Institute, and the RAD Sensor Head (RSH) which has been designed and developed by the Christian-Albrechts-Universität zu Kiel.

As schematically shown in Fig. 1.3, the sensor head consists of a telescope of silicon detectors (**A**, **B**, **C**) followed by a cesium iodide scintillator (**D**) and a plastic scintillator (**E**). Both, **D** and **E** are enclosed by a plastic Anti-Coincidence (AC) (**F1**, **F2**). The readout of the **D** and **E** detector is each done by three photodiodes (the white trapezoid shape on the **D** scintillator in Fig. 1.3).

Charged particles are detected with a telescope formed by the **A**, **B**, **C** and, depending on their energy, the **D** and **E** detector. Neutral particles are detected in the **D** and **E** detector in AC with **C** and **F1**, **F2**. The **D** detector is mainly sensitive to gammas and the **E** detector is mainly sensitive to neutrons. However, both detectors are also sensitive to neutrons respectively γ-rays which, among other effects, complicates the neutral particle measurement. Calculating the incoming gamma and neutron spectra from the neutral particle measurements in **D** and **E** is one of the main aims of this thesis. The initial preamplification of the detector signal is done in the sensor head. To minimize noise, the preamplifiers were put as close as possible to the detectors (see green electronic boards between **A** and **B** in Fig. 1.3). Further signal processing is done in the REB, a Pulse-Height Analysis (PHA) is obtained, coincidence conditions are evaluated, data is stored, put into histogram format and sent back to Earth over the spacecraft's communication bus. Due to constraints in power consumption, RAD has a duty cycle with 15 minutes of measurement every hour.

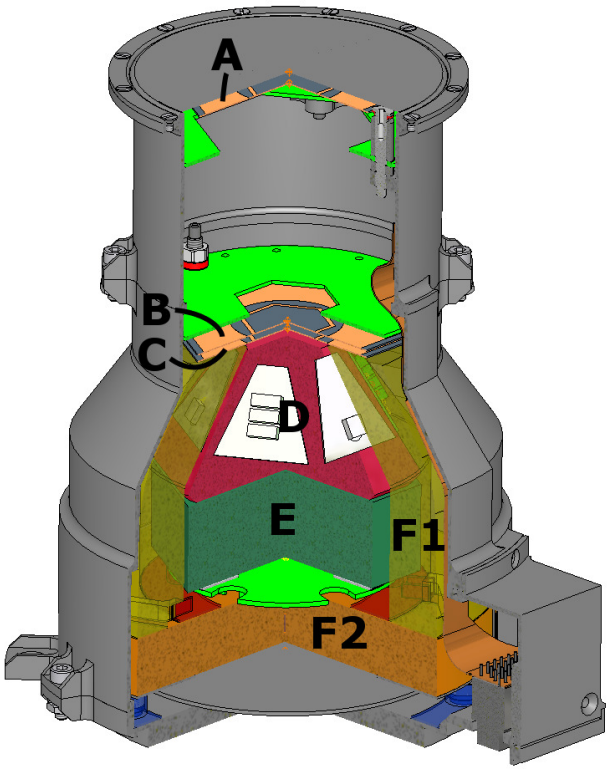

Figure 1.3: Cutaway view of the RAD sensor head. CAD graphic by L. Seimetz, edited by O. Kortmann.

1.3 The Martian radiation environment

Although some regions on Mars are magnetized due to crustal remanence of a former magnetic field, present-day Mars possesses no planetary magnetic field comparable to the one found on Earth. Because of the absence of a significant magnetic field, the Galactic Cosmic Rays (GCR) and SEPs can enter the Martian atmosphere unhindered. Although the Martian atmosphere is comparatively thin ($\sim 2 \cdot 10^{-5}$ g/cm^3 at the surface), energetic particles entering the Martian atmosphere create a cascade of secondary particles, which form a major part of the radiation environment observed on the Martian surface. Another major part is given by secondary particles generated in the Martian soil. The composition of these particles depends on the soil composition, e.g. subsurface water-ice has a large influence on neutron particles.

A schematic example of the particle cascades in atmosphere and soil is shown in Fig. 1.4, where incident proton particles create a cascade of charged and neutral particles. In this work especially the neutral particle component, i.e. neutron and gamma spectra, for which an inversion approach is developed, is of interest.
Understanding the transport of radiation through the Martian atmosphere and estimating the influence of atmospheric density and of soil composition is a challenging task and is

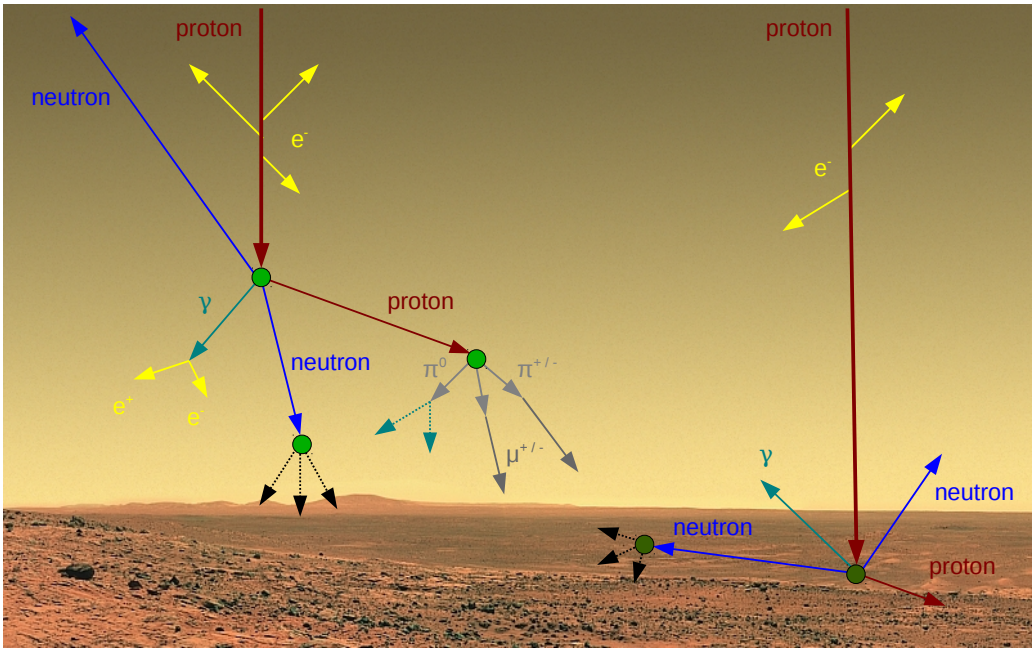

Figure 1.4: Schematic overview of the secondary particle production in the Martian atmosphere and soil. Provided by *NASA photojournal*, edited by B. Ehresmann.

not treated in this work. However, the doctoral thesis Ehresmann [2012] discusses those processes in detail and provides data which are used in this thesis.

The two main sources for the Martian particle radiation are GCRs and SEPs.

1.3.1 Galactic cosmic rays

The GCRs consist mainly of high energetic protons (87%) followed by ionized helium (12%) and ionized heavier elements (1%). The density of the GCRs is far lower than the density of the solar wind particles, however, the average GCR-particle energy is with 7 GeV by several magnitudes larger than the energy of solar-wind particles or SEPs. An overview of the GCR spectrum is shown in Fig. 1.5. The origin of this radiation are believed to be acceleration by shock fronts of supernovae explosions and jets from black holes and pulsars.

One remarkable aspect of the GCRs is the temporal variation of its intensity, which is anti-correlated with the solar cycle. During low solar activity the GCR intensity increases, while it decreases during high solar activity. Among others, this effect is caused by the stronger scattering of the GCRs by the irregular solar magnetic field [Prölss, 2003].

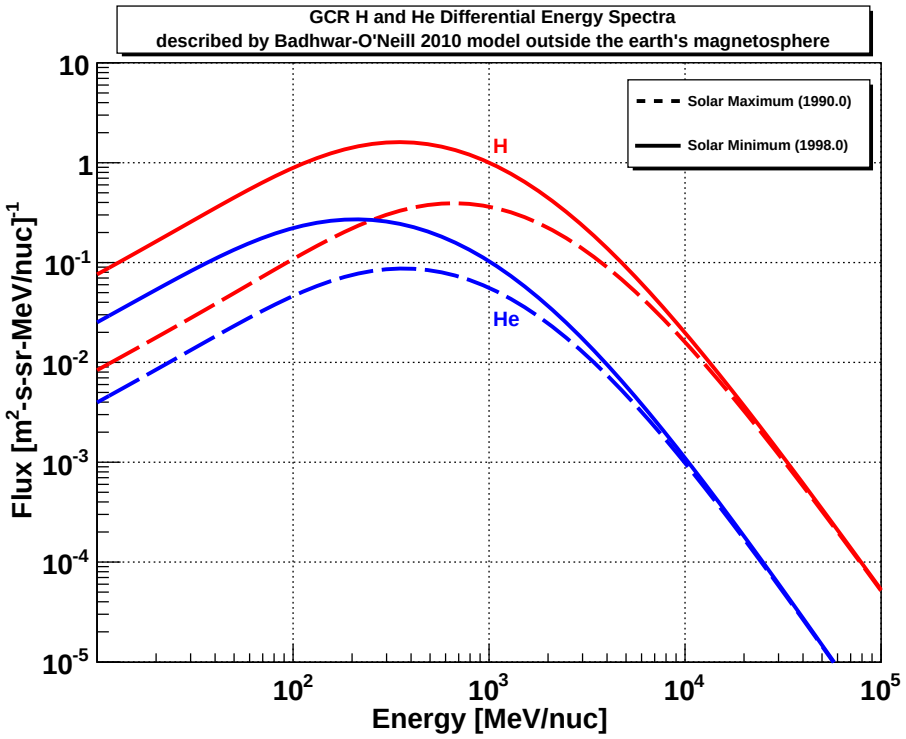

Figure 1.5: Energy spectrum of the proton and helium components of the GCRs for solar minimum and maximum, as described by the Badhwar O'Neill 2010 model. Figure by A.I. Mrigakshi.

1.3.2 Solar Energetic Particles

Solar energetic particles are mainly generated in solar flares and Coronal Mass Ejections (CME). In a solar flare, an unstable magnetic field configuration on the solar surface collapses and evolves through changing and reconnection into a more stable state. This change in the magnetic structure is accompanied by an energy release in the form of plasma heating and particle acceleration. CMEs are bursts of plasma and magnetic fields, which are released from the solar corona into interplanetary space. Similar as in a flare, an unstable magnetic structure collapses and a part of the magnetic field is separated, through reconnection of field lines, from the solar magnetic field. This unconnected magnetic structure expands outwards into space and can drive a shock front accelerating particles. Both flares and CMEs can accelerate particles up to energies of several 100 MeV/nuc. Those particles propagate along the solar magnetic field lines, which can be approximated as Archimedian spirals, see Fig. 1.6.

The propagation of SEPs is strongly influenced by scattering with magnetic waves and fluctuations in the solar magnetic field. If a particle oscillates with the same frequency as a magnetic fluctuation, it is in resonance and its pitch angle can be scattered. Pitch-

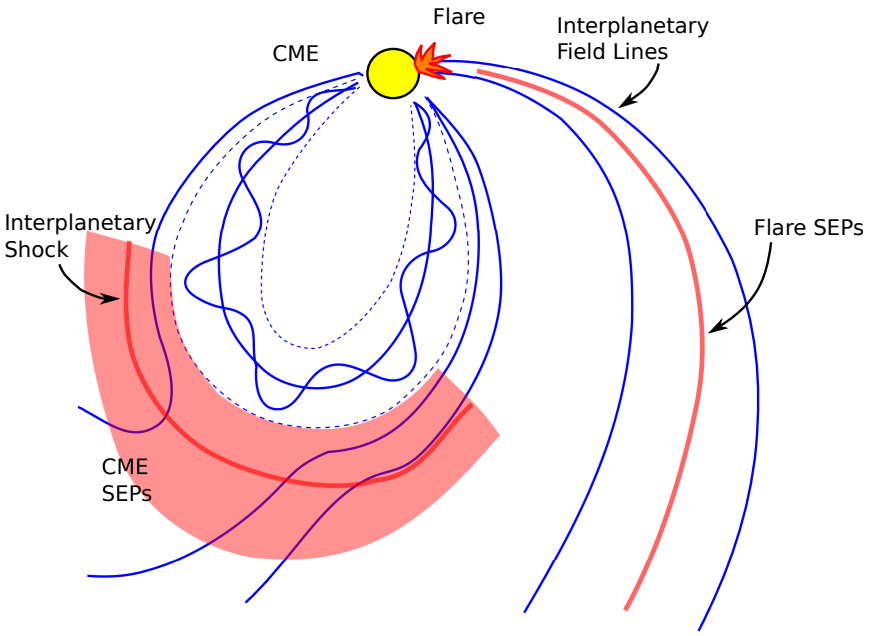

Figure 1.6: Schematic example of SEP generation by flares and CMEs.

angle scattering by magnetic fluctuations affects intensity, propagation velocity, and other factors of SEP events. The composition of the magnetic fluctuation in the solar wind is, therefore, important for describing the propagation of SEPs.

Energetic particles reaching Earth, can cause aurorae, disrupt radio transmissions, damage satellites, cause power outages, or even increase the level of radiation. Because of Mars' missing magnetosphere, the increase of radiation caused by a SEP event is far higher than on Earth.

1.4 Scope of this work

This work is organized in two main parts. In the first part, a model describing the propagation of SEPs from Sun to Mars is derived. In the second part, a method to calculate the incident gamma and neutron spectra from the MSL/RAD measurement is developed. The gap between propagation of SEPs from Sun to Mars and the measurement of neutral particles on the Martian surface, i.e. the production of secondary particles in the Martian atmosphere, is treated in Ehresmann [2012].

In the first part of the second chapter, a modified wavelet analysis is developed, which is used to analyze magnetic field data for the occurrence of short waves or fluctuating pulses. This analysis evaluates amplitude, width, and rate of occurrence of waves over a wide range of frequencies. Additional informations, such as polarization of the waves, can be extracted as well. The modified wavelet analysis is applied to ACE/MAG data for fast and slow solar wind and for solar minimum and maximum. In addition, Helios and Ulysses data

are used to investigate the radial dependence of the mean amplitudes of the fluctuations for solarcentric distances from 0.3 - 1.0 AU (Helios) and from 1 - 5 AU (Ulysses). A model which creates a set of artificial fluctuations with the same statistical properties as found in the solar wind is verified and used as a base for test-particle simulations.

An ab-initio code, based on the Lorentz force, is developed and tested for scattering of energetic particles on one magnetic fluctuation. This model is extended to include a complete magnetic-field line from sun to Mars, which exhibits the same statistical properties as found in the solar wind data. The influence of large amplitude waves on the particle propagation from Sun to Mars is modeled for magnetic fluctuations with several different intensities.

In the third chapter, an approach for inverting neutron and gamma spectra from a measurement with two scintillators is developed. Two different statistical approaches are considered, the least squares approach, based on Gaussian statistics and the Poisson approach based on Poisson statistics. Both approaches connect a measurement vector with the spectrum of incoming particles via a Detector Response Function (DRF). A maximum likelihood estimation is employed and the resulting negative log likelihood functions are minimized numerically. The ability of the NNLS and Poisson method to invert measurements for different levels of statistics and their sensitivity to initial guesses is discussed in detail for artificial generated measurements. Several strategies for optimizing the initial guess to find the optimal result are developed. Finally, both methods are compared in terms of their ability to handle errors in the DRF. Several DRFs for the **D** and **E** detector of the RAD instrument are derived and verified by inverting calibration measurements.

In the following chapter, the DRFs are used to generate and invert artificial measurements of the expected Martian neutron and gamma spectra. The inversion procedures are tested and compared for realistic artificial measurements with count rates as expected on Mars. The influence of the RTG is calculated and methods to correct the measurement are tested. Finally, the inversion approach is extended for a generalized measurement with coincidence conditions. This approach is applied to the 2D neutral particle histograms of MSL/RAD, to improve the results of the inversion.

In a conclusion, the most important findings are summarized and an outlook for future applications is given. In the appendix, a more detailed description of the inversion algorithm and the optimization procedure for the initial guess is given. The derivation of the final DRFs for RAD are discussed in detail. Some additional insight into the simulation of the RAD instrument is given within a detailed analysis of bismuth calibration runs and the thermal behavior of the instrument.

Chapter 2

Propagation of solar energetic particles from Sun to Mars

The main part of the Martian particle spectra the RAD instrument will measure is caused by GCRs, which generate a cascade of secondary particles in the Martian atmosphere. Another significant contribution to the Martian radiation spectra is given by SEPs which are generated by coronal mass ejections or flares, travel along the solar magnetic field and possibly reach Mars. Just as the GCRs, SEPs create a cascade of secondary particles in the Martian atmosphere. An example for the gamma and neutron spectra resulting from an SEP event [Lovell *et al.*, 1998] is presented in Gurtner *et al.* [2005], where the interaction of SEPs is calculated with an extended `GEANT4`-simulation code. Understanding the propagation of SEPs on their way to Mars is important for calculating their influence on the Martian radiation environment and, therefore, estimating the hazards for future manned missions on Mars.

The common way to describe the propagation of energetic particles in the turbulent interplanetary magnetic field is the focused transport equation [Roelof, 1969],

$$\frac{\partial f}{\partial t} + v\mu\frac{\partial f}{\partial z} + \frac{1-\mu^2}{2L}v\frac{\partial f}{\partial \mu} = \frac{\partial}{\partial \mu}D_{\mu\mu}\frac{\partial f}{\partial \mu}, \tag{2.1}$$

where $f(z, v, \mu, t)$ is the phase space density per unit magnetic line length, which means f has to be divided by the magnetic field strength $B(z)$ to compare theoretical with observed intensities, z is the coordinate along the mean magnetic field, v the particle velocity and μ the pitch-angle cosine. The second term on the left-hand side represents the motion along the mean magnetic field, where $v\mu$ corresponds to the guiding center velocity, the third term on the left-hand side describes the pitch-angle focusing effect through the magnetic gradient which is characterized by the focusing length $L = -B(z)/(\partial B/\partial z)$. The right-hand term describes the pitch-angle scattering by resonant magnetic fluctuations. For a combination of Quasi-Linear Theory (QLT) and slab model, which is known as the "standard model" [Schlickeiser and Jaekel, 1992], the pitch-angle diffusion coefficient was derived by Jokipii [1966] and Hasselmann and Wibberenz [1968] for a power spectrum $P(k) \propto k^{-q}$ of the

fluctuating field as

$$D_{\mu\mu} = \frac{\nu}{2}(1-\mu^2). \tag{2.2}$$

The scattering frequency ν is given by

$$\nu(\mu) = \frac{v}{2\lambda_{\parallel}(4-q)(2-q)}|\mu|^{q-1}, \tag{2.3}$$

where $\lambda_{\parallel}$ is the parallel mean free path for a nearly isotropic distribution function [Wibberenz and Beeck, 1986]. Typical values of $\lambda_{\parallel}$ range from below 0.1 to 1 AU [Palmer, 1982]. Note that $\lambda_{\parallel}$ solely influences the magnitude of the diffusion coefficient and is not an actual distance between two scattering events. The diffusive approximation assumes that a particle is permanently influenced by resonant fluctuations. However, it is unlikely that resonant waves are permanently present, e.g. Bruno and Bavassano [1989] showed that only $15-41\%$ of the observed intervals contain Alfvénic fluctuations on hourly scales. The rate of occurrence of Alfvénic intervals varies with flow speed, heliocentric distance and solar activity [Bavassano and Bruno, 1991]. In periods with a weak background and rare strong scattering events, the concept of a diffusion approximation is not justified.

In the first part of this chapter, a modified wavelet analysis is developed to investigate the amplitude and temporal occurrence of magnetic fluctuations. In the second part of this chapter, an ab-initio simulation is developed which uses the obtained wave statistics to calculate the propagation of test particles in the turbulent solar magnetic field.

2.1 Analysis of magnetic fluctuations in the solar wind

The propagation of energetic particles through the heliosphere is strongly influenced by fluctuations in the magnetic field. These fluctuations have been measured by several spacecraft such as Helios, Ulysses or the Advanced Composition Explorer (ACE). The properties of those fluctuations have been analyzed for various aspects, e.g. Mariani *et al.* [1978] calculated the radial dependence of the Pythagorean variance, Denskat and Neubauer [1982] and Bavassano *et al.* [1982] studied the radial evolution of the power spectra.The suggestion of separating inward from outward propagating fluctuations has been made by Grappin *et al.* [1983] and can be achieved via the Elsässer variables [Elsässer, 1950]. The magnetic fluctuations can be separated in fluctuations perpendicular and parallel to the magnetic field, where pitch-angle scattering is mainly caused by perpendicular fluctuations [Shalchi and Schlickeiser, 2004]. Perpendicular magnetic fluctuations can be separated in waves which are transverse and axially symmetric with wave vectors parallel to the mean magnetic field (slab geometry) and 2D fluctuations (magnetostatic structures which move past the observer). The fraction of parallel propagating fluctuations in the solar wind is given by ~ 0.2 [Bieber *et al.*, 1996].

The properties of the magnetic fluctuations are factor for determining the pitch-angle scattering of SEPs. For the transport of SEPs, the interplanetary magnetic field can

usually be described as a Parker field with superimposed irregularities. The motion of the energetic particles is described as an adiabatic motion along the smooth Parker field and pitch-angle scattering by magnetic fluctuations. In the QLT, the particle's pitch angle is mainly scattered by fluctuations, with wave vectors parallel or antiparallel to the mean magnetic field, which are in resonance with the particle's gyro motion. In the slab model for the QLT, charged particles, with velocity v and gyration time T, are influenced by fluctuations with wavenumber k when the resonance condition

$$v_{\parallel} T = 2\pi/k \tag{2.4}$$

is fulfilled [Tsurutani and Lakhina, 1997]. In contrast, 2D fluctuations contribute very little to particle scattering [Shalchi and Schlickeiser, 2004].

Assuming that the irregularities superposed on the average field are small, the QLT describes pitch-angle scattering as a diffusive process. The diffusion of the particle's pitch angle can be described by the diffusion coefficient $D_{\mu\mu}(\mu)$ [Dröge, 2000], which is directly determined by the spectral index q of the magnetic fluctuations [Jokipii, 1966; Hasselmann and Wibberenz, 1968]. However, large amplitude fluctuations can lead to nonlinear wave-particle interactions and make the concept of a superposition of small-amplitude irregularities rather questionable [Dröge, 1994]. Often, only the power-spectral density is shown and is believed to be sufficient to characterize magnetic fluctuations. However, if the magnetic fluctuations consist of few large waves in a weak background, the power-spectral density may not be sufficient to fully characterize the pitch-angle scattering through the diffusion coefficient.

In the following sections, a modified wavelet analysis is developed to investigate the temporal occurrence of magnetic fluctuations, which is then used to analyze amplitude and rate of occurrence of magnetic fluctuations in Helios, ACE and Ulysses data for several heliocentric distances, different latitudes, and solar activities.

2.1.1 Wavelet analysis

If the phase information of a signal is neglected, a lot of information is lost. Even if the complex spectrum is considered, the nature of the original signal can not be interpreted easily. Fig. 2.1 (top) shows an oscillation with a rising frequency and two short oscillating pulses. For the first example, the Fourier transformation gives the power per frequency, but yields no easily accessible information about the shape of the signal. For the second example, the power and frequency of the oscillation can be determined, but it is not possible to deduce whether it originates from one strong or many weak pulses.

However, there exist several tools to determine the temporal evolution of a frequency spectrum. For the short-time Fourier transformation, the signal is multiplied with a short window which moves along the time axis. The transformation is defined by

$$X(\tau,\omega) = \int_{-\infty}^{\infty} x(t) w(t-\tau) e^{-j\omega t}\, dt, \tag{2.5}$$

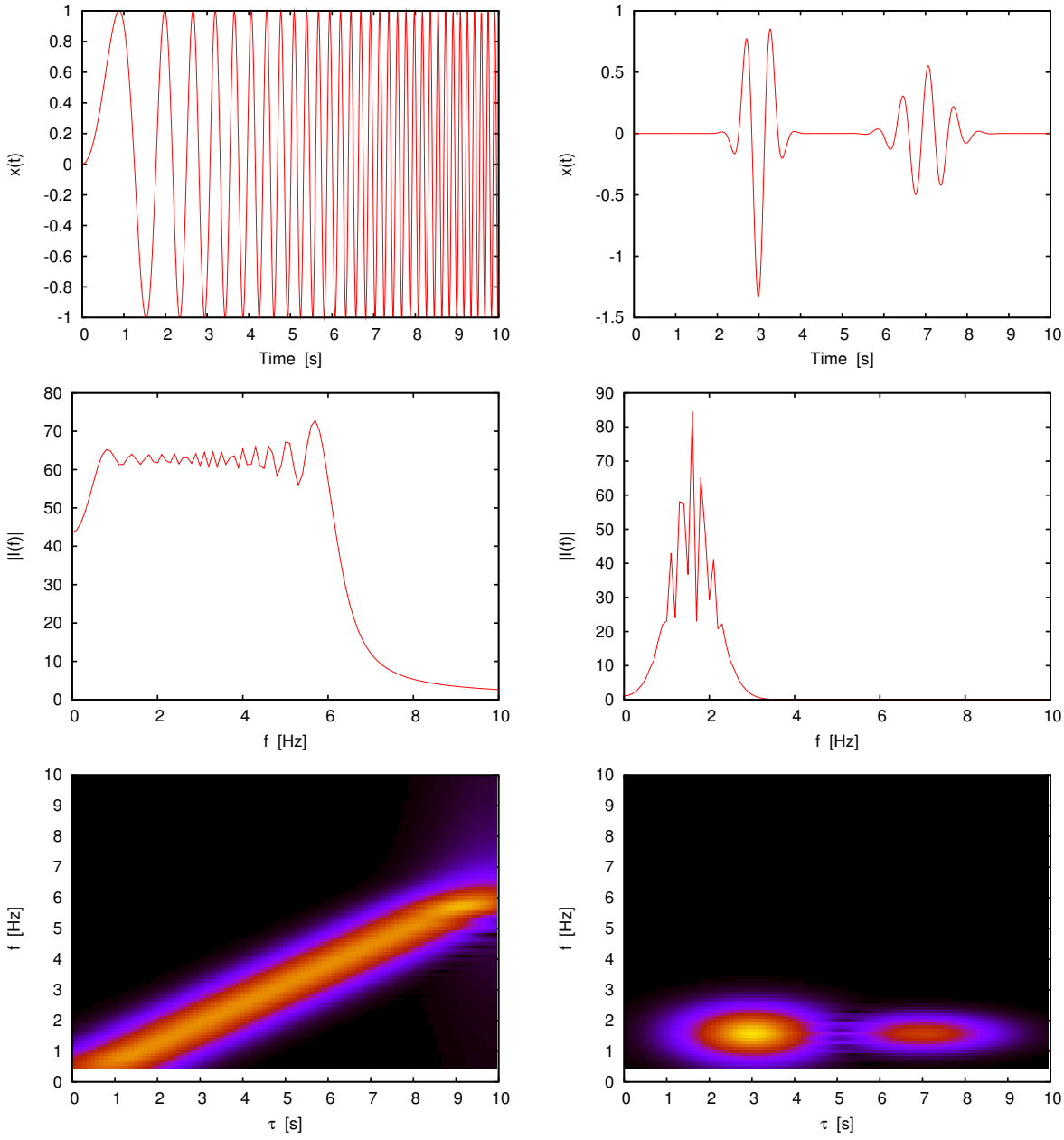

Figure 2.1: Top: An oscillation with rising frequency and two short oscillation pulses. Middle: The corresponding Fourier transformations. Bottom: The corresponding short-time Fourier transformations for a Gaussian window function with standard deviation 1.

where $w(t)$ is a window function, e.g., a zero centered Gaussian function. This yields an intensity spectrum $I(\tau, f)$, where τ is the position of the window. While this provides information on the temporal evolution of the spectrum, a major disadvantage is the fixed window width.For a given window width d, frequencies below $1/d$ can not be resolved because their wavelength is larger than the length of the window. On the other hand, for frequencies $\gg 1/d$ there are many wavelengths in the window, which leads to a poor temporal resolution.

Another widely used tool, which analyses the temporal occurrences of frequencies in a signal, is the wavelet transformation. The Continuous Wavelet Transformation (CWT) is defined as [Percival and Walden, 2000]

$$W(\lambda, \tau) := \int_{-\infty}^{\infty} x(t) \, \frac{1}{\sqrt{\lambda}} \psi \left(\frac{t - \tau}{\lambda} \right) \tag{2.6}$$

which is essentially the integral of the signal $x(t)$ multiplied with the wavelet $\psi(t)$ whose frequency is a function of λ and whose position is given by τ. As the name suggests, a wavelet is a "small wave", which means it grows and decays over a short period of time. Two basic properties of a real valued wavelet defined over the real axes $(-\infty, \infty)$ are:

1. $\psi(t)$ integrates to zero.
2. $\psi(t)^2$ integrates to unity.

In addition there must be a $T > 0$ for which $\psi(t)$ only deviates insignificantly from zero for $t \notin [-T, T]$.
The most common wavelets are:

- The Haar wavelet (see Fig. 2.2)

$$\psi(t) = \begin{cases} -1/\sqrt{2}, & -1 < t \leq 0, \\ 1/\sqrt{2}, & 0 < t \leq 1, \\ 0, & \text{otherwise} \end{cases} \tag{2.7}$$

  which is probably the oldest wavelet and named after A. Haar. Among others, it is used to monitor signals for abrupt changes.

- The Mexican hat wavelet (see Fig. 2.2)

$$\psi(t) = \frac{2}{\pi^{1/4}\sqrt{3\sigma}} \left(1 - \frac{t^2}{\sigma^2} \right) e^{-t^2/2\sigma^2} \tag{2.8}$$

  which is the normalized negative second derivative of a Gaussian function. One of its applications is blob detection in computer vision.

- The Morlet wavelet (see Fig. 2.2)

$$\psi(t) = C e^{-i2\pi f_0 t} \left(e^{-t^2/2} - \sqrt{2} e^{-(2\pi f_0)^2/4} e^{-u^2} \right), \tag{2.9}$$

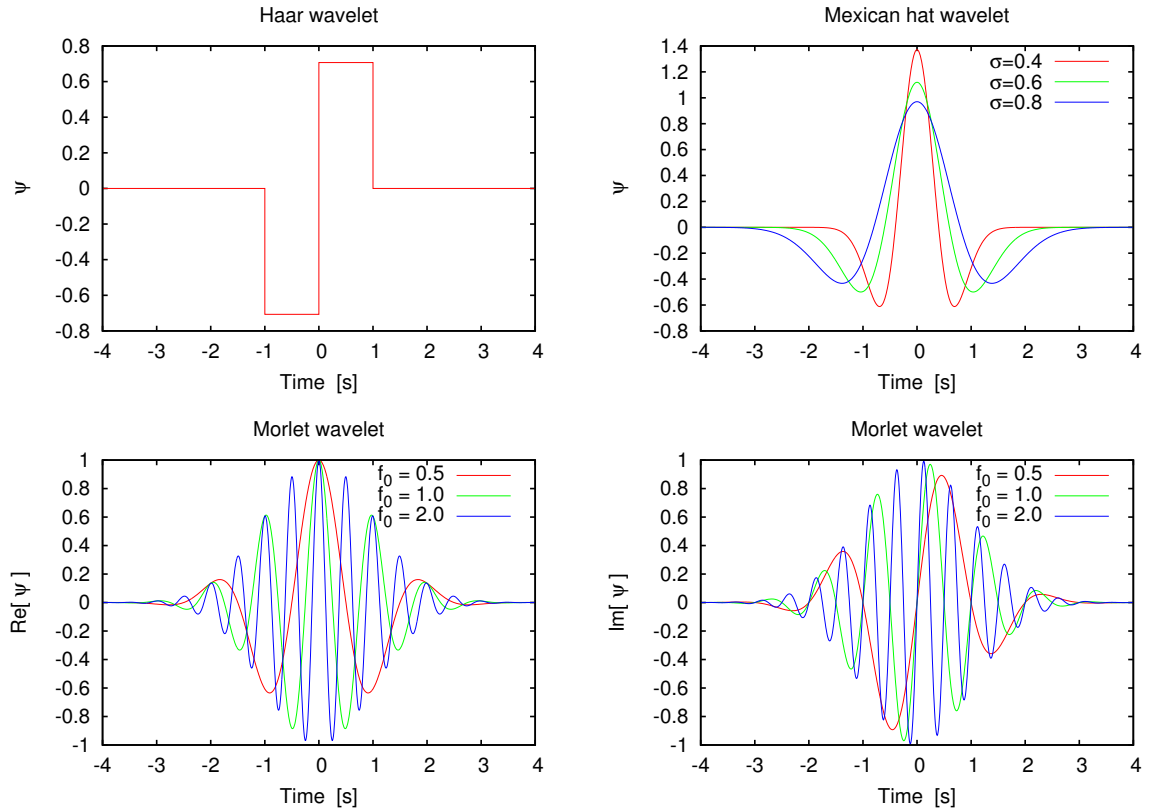

Figure 2.2: Three wavelets. Top, left: The Haar wavelet (Eq. 2.7). Top, right: The Mexican hat wavelet (Eq. 2.8) for various σ values. Bottom: The Morlet wavelet (Eq. 2.10), real (left) and imaginary (right) part for several values of f_0.

which can by approximated by

$$\psi(t) = \pi^{-1/4} e^{-i2\pi f_0 t} e^{-t^2/2}. \tag{2.10}$$

Note that f_0 is a constant which determines the shape of the wavelet and not the analyzed frequency. As obvious from Eq. 2.10, the wavelet consists of a complex oscillation multiplied with a Gaussian shaped envelope. This means a CWT with the Morlet wavelet behaves very similar to a Fourier transformation of a signal multiplied with a Gaussian shaped window function, i.e. a short-time Fourier transformation. The key difference compared to the short-time Fourier transformation is that the size of the window adapts with the analyzed frequency. The width of the wavelet σ_ψ, which is given by $f/f_0 = \lambda^{-1}$, depends on the analyzed frequency f. In contrast to the short-time Fourier transformation, the window width adapts with the frequency and the Gaussian shaped envelope will always contain the same number of wavelengths.

Since the aim is to analyze magnetic field data for the occurrence of magnetic fluctuations, a wavelet with a shape similar to the magnetic fluctuations is the best choice. Therefore, the Morlet wavelet is chosen, which is basically an oscillation with a Gaussian shaped envelope. The frequency of the wavelet is given by $f = f_0/\lambda$; f_0 determines the number of wavelengths in the Gaussian envelope. With increasing width of the Morlet wavelet the temporal resolution of the CWT decreases while the frequency resolution increases. E.g., for $f_0 \to \infty$ the CWT approaches a Fourier transformation which has no temporal resolution, while for $f_0 \to 0$ the shape of the wavelet approaches a peak which can give no useful information about occurring frequencies. Making a compromise between temporal and frequency resolution, f_0 is in the following chosen as $f_0 = 1$ Hz.

In Fig. 2.3 we show the wavelet analysis of the same signal as in Fig. 2.1. The peaks of $|W|_f(t)$ clearly show occurrences of the analyzed frequency. An example for a wavelet analysis of ACE MAG data with a wide range of analyzed frequencies is shown in Fig. 2.4. Unlike for the short-time Fourier transformation, high and low frequencies can be resolved.

The modified wavelet analysis

Similar as for a Fourier transformation, the square modulus $|W(\lambda,\tau)|^2$ is proportional to the energy of fluctuations with frequency $f = f_0/\lambda$ at position τ [Carbone and Bruno, 2005]. To analyze time series for occurring wave amplitudes, instead of for energy of the fluctuations, the CWT needs to return the correct amplitude of an oscillation. This means it needs to fulfill the condition

$$|W(\lambda = f_0/f, \tau)| = A \tag{2.11}$$

for the analyzed signal

$$x(t) = A \cdot \sin(2\pi f t + \phi). \tag{2.12}$$

Calculating the wavelet analysis for $x(t)$,

$$\begin{aligned} |W(\lambda = f_0/f, \tau)| &= \left| \int dt\, A \sin(2\pi f t + \phi)\, e^{-i2\pi f(t-\tau)} \frac{e^{-\frac{(t-\tau)^2}{2(f_0/f)^2}}}{\sqrt{\sqrt{\pi}(f_0/f)}} \right| \\ &\approx \int dt\, \frac{1}{2} A \cdot \frac{e^{-\frac{(t-\tau)^2}{2(f_0/f)^2}}}{\sqrt{\sqrt{\pi}(f_0/f)}} \\ &= A \cdot \sqrt{\sqrt{\pi}(f_0/f)/2}, \end{aligned}$$

yields a scaling factor of $\sqrt{\frac{2f}{\sqrt{\pi} f_0}}$, which allows us to obtain the correct amplitude of a signal. Fig. 2.5 shows the corrected wavelet analysis for a constant oscillation and for a short Gaussian shaped oscillation. While the amplitude of the constant oscillation is calculated correctly, the amplitude of the oscillating pulse is decreased while the width is increased.

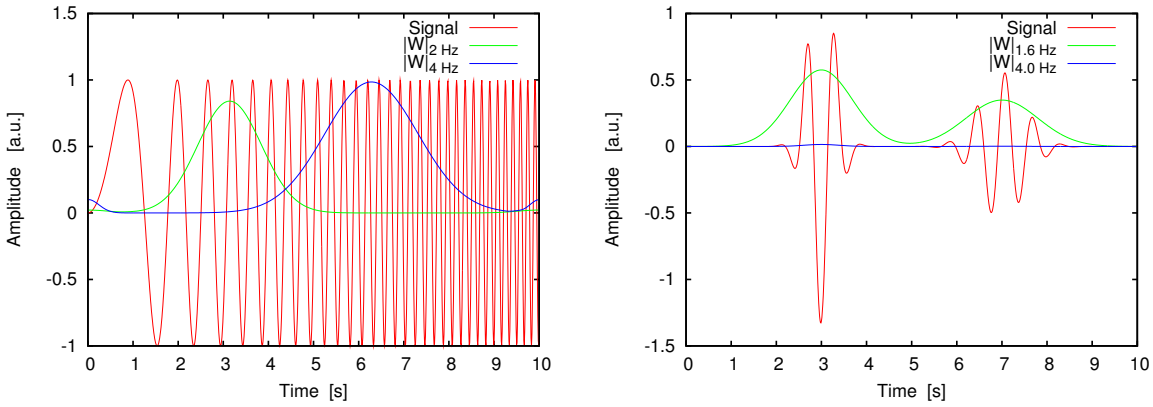

Figure 2.3: Wavelet transformations for the same signals as in Fig. 2.1. The frequencies of the CWT are chosen arbitrarily, a thorough analysis of the signal would evaluate a broad range of frequencies.

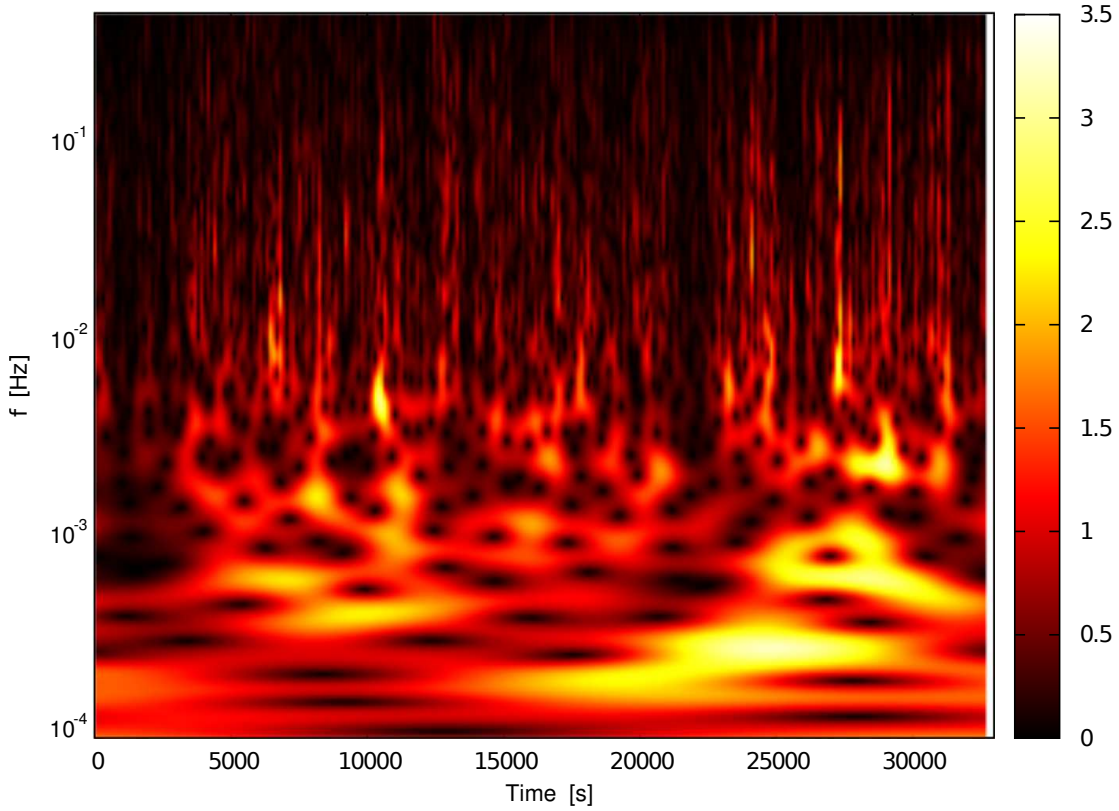

Figure 2.4: Example of a wavelet analysis for an interval of ACE MAG data (2000, day of year 1.0 - 1.38). A broad range of evaluated frequencies is shown in a 2D density-plot representation. The colors give the intensity $|W|_f(t)$ over time and frequency.

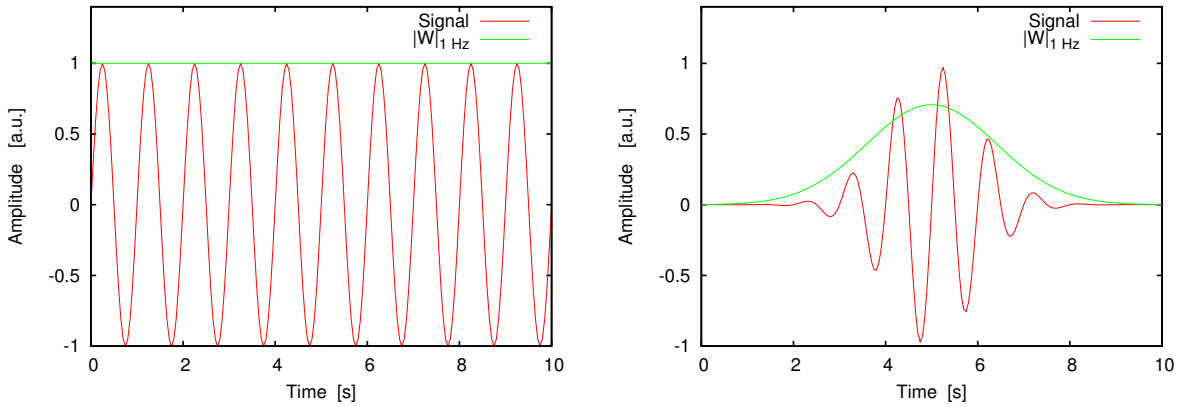

Figure 2.5: Examples for wavelet analysis with applied scaling factor. Left: $|W|$ displays the correct amplitude of a constant oscillation. Right: Although the scaling factor was applied, amplitude and shape of $|W|$ do not match the envelope of the Gaussian oscillation.

This can be understood if one considers that Eq. 2.6 is basically the convolution of signal and wavelet, in this case the convolution of two Gaussian functions. The convolution of two Gaussian functions results in a Gaussian function whose standard deviation is the geometric sum of both standard deviations. From the knowledge of the standard deviation of the wavelet and of the transformed signal, the envelope of the original signal can be reconstructed. The standard deviation of the wavelet is given by $\sigma_\psi = \frac{f_0}{f}$, where f is the frequency of the wavelet. An arbitrary Gaussian-shaped oscillation is given by

$$x(t) = \frac{A}{\sqrt{2\pi}\sigma_x} e^{-\frac{t^2}{2\sigma_x^2}} \cdot \sin(2\pi f t). \tag{2.13}$$

We define the amplitude of the wave $x_{max} := \frac{A}{\sqrt{2\pi}\sigma_x}$. The resulting CWT then yields

$$|W| \approx \int dt \, \frac{A}{\sqrt{2\pi}\sigma_x} e^{-\frac{t^2}{2\sigma_x}} \cdot \frac{1}{\sqrt{2\pi}\sigma_\psi} e^{-\frac{(\tau-t)^2}{2\sigma_\psi}} \tag{2.14}$$

$$= \frac{A}{\sqrt{2\pi}\sigma_x} e^{-\frac{\tau^2}{2\sigma_W^2}}. \tag{2.15}$$

σ_W and $W_{max} := \frac{A}{\sqrt{2\pi}\sigma_W}$ can be easily obtained from any transformed signal. Therefore, the original signal can be reconstructed via

$$\sigma_x = \sqrt{\sigma_W^2 - \sigma_\psi^2} \tag{2.16}$$

$$x_{max} = W_{max} \frac{\sigma_W}{\sigma_x}. \tag{2.17}$$

Fig. 2.6 shows the modified wavelet transformation of a short Gaussian pulse and the reconstructed signal which matches the envelope of the pulse.

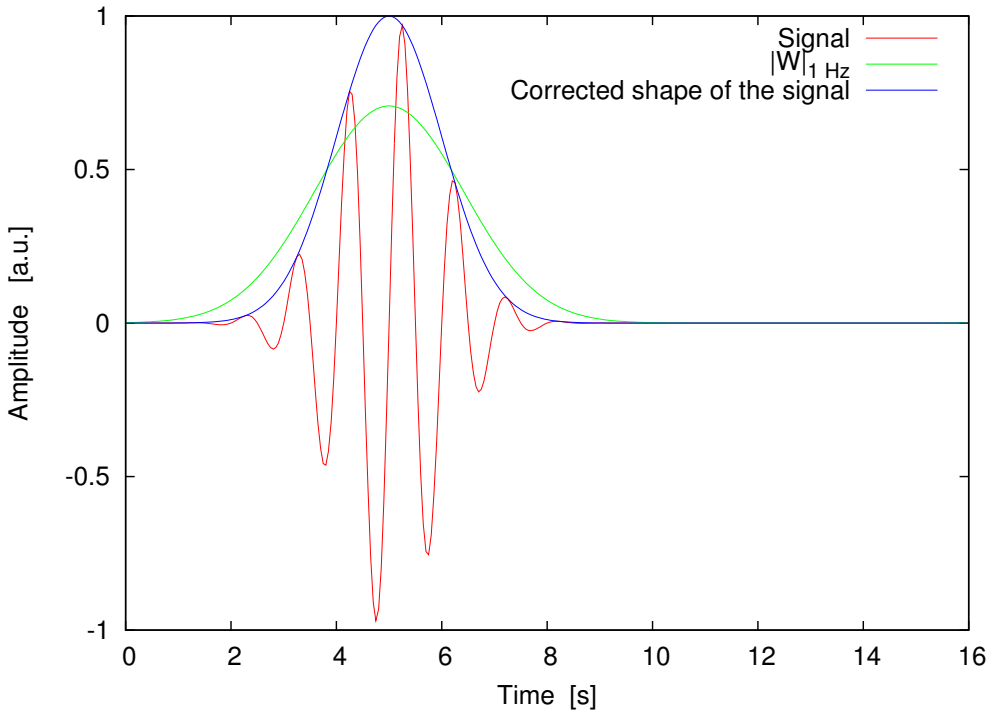

Figure 2.6: A 1 Hz Gauss-shaped oscillation (red), the corresponding wavelet transformation (green), and the corrected transformation (blue) are shown.

Determining amplitude and width of fluctuations

In the turbulent interplanetary magnetic field with a superposition of many non-Gaussian waves, the situation is more complex. In these situations, the exact reconstruction, which was shown for a single Gaussian wave, becomes inaccurate. The non-Gaussian-shaped oscillations and the superposition of many waves with various frequencies can create situations where $\sigma_W \approx \sigma_\psi$, resulting in very large amplitudes or even complex values for σ_x. Therefore a simple case differentiation is used, which describes the amplitude x_{max} and width $2\sigma_x$ of the oscillations in turbulent magnetic field data very well.

$$\sigma_x = 0.75\,\sigma_W \tag{2.18}$$
$$x_{max} = 1.5\,W_{max} \tag{2.19}$$

for $\sigma_W \leq 2/f$; and for $\sigma_W > 2/f$

$$\sigma_x = \sigma_W \tag{2.20}$$
$$x_{max} = W_{max}, \tag{2.21}$$

where W_{max} and σ_W are height and standard deviation of a detected peak in the CWT. This method allows us to count for each analyzed frequency f the occurring amplitudes

of the fluctuations x_{max} and the width of the waves $2\sigma_x$. Considering the width $2\sigma_W$ of waves, it is important to mention the dependence on the wavelet width. A wavelet with a large width has a poor temporal resolution and smooths out short neighboring waves which then would be detected as a single wave. It can therefore be said that a statistic of wave width can be obtained, but depends on the width of the wavelet, i.e. the chosen temporal resolution.

Consideration for the analyzed frequencies

The maximum analyzed frequency f_{max} is obviously given by the Nyquist frequency, the minimum frequency f_{min} is chosen so that the length of the analyzed interval contains at least 20 wavelengths $1/f_{min}$. Eq. 2.6 can be understood as a convolution of wavelet and signal. The convolution theorem states that a Fourier transformation of a convolution is the point-wise product of the Fourier transformations. One can interpret a CWT as transformation into frequency space, selecting those frequencies of the signal which are given by the wavelet, and then transforming the frequencies back. Therefore, one needs to select the frequencies of the wavelets (and therefore their position in frequency space) in a way that all frequencies are analyzed, but not multiple times by different wavelets. The Fourier transformation of a Morlet wavelet, which analyzes the frequency f_i, is given by

$$\hat{\psi}(f) \propto e^{-0.5(2\pi(f-f_i)^2)\frac{f_0^2}{f_i^2}}. \tag{2.22}$$

Note that its width is $\propto f_i$. For the frequency range $[f_{min}, f_{max}]$ and $f_c = 1$ Hz, the analyzed frequencies are given by

$$f_j = f_{max}/1.5^i \qquad \text{with } f_i > f_{min}. \tag{2.23}$$

This way, the frequencies of ψ_{f_j} and $\psi_{f_{j+1}}$ do not overlap significantly, and there is no large gap between f_i and f_{i+1}.

Reversibility

For an analyzed signal, a set of wave amplitudes $x_{f_i,j}$ and wave widths $2\sigma_{f_i,j}$ can be obtained. The detected wave amplitudes should contain all information about the signal's power-spectral density. Therefore, an artificial signal, filled with wave amplitudes from the original signal, should contain the same power-spectral density.

To verify this, an interval of magnetic fluctuations from ACE MAG measurements [Smith *et al.*, 1998] is analyzed. For each detected wave amplitude $x_{f_i,j}$, a Gaussian wave packet with standard deviation $\sigma_{f_i,j}$, is randomly inserted in an artificial signal. The artificial signal is generated by

$$x(t) \;=\; \sum_i \sum_j x_{f_i,j} \sin\left(2\pi f \cdot t + \phi_{f_i,j}\right) \cdot e^{-\frac{(t-t_{f_i,j})}{2\sigma^2_{f_i,j}}}, \tag{2.24}$$

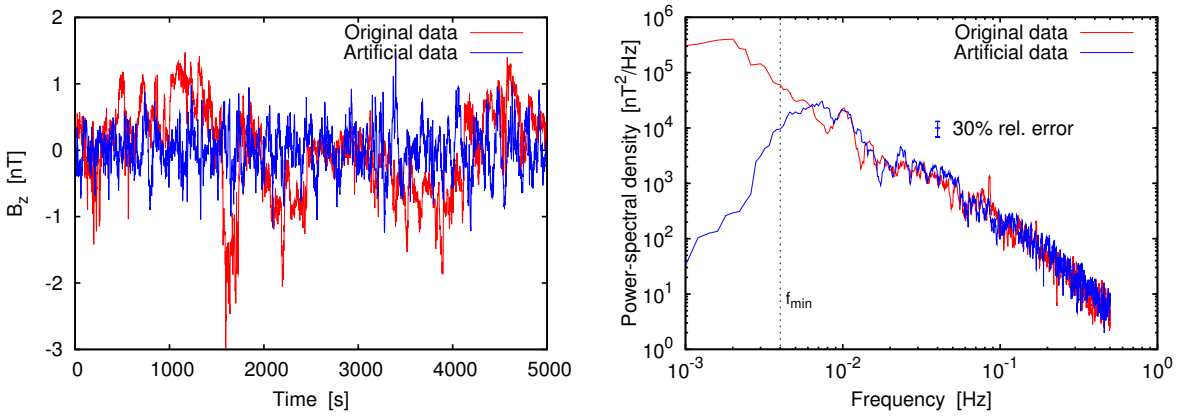

Figure 2.7: Left: Comparison of original ACE MAG B_z-data (2006, day of year 2.45 - 2.55) and artificial, via Eq. 2.24, created data. The fluctuations in the artificial signal are randomly distributed, therefore, we do not expect an signal but a similar signal. Right: The corresponding (smoothed) power-spectral densities agree for a wide range of frequencies. Only for frequencies below f_{min}, the artificial signal shows a decreased intensity.

where the first term is the sum over all analyzed frequencies, the second term is the sum over all detected waves with frequency f_i. $\phi_{f_i,j}$ and $t_{f_i,j}$ are the randomly assigned phase and position of each wave.

Fig. 2.7 shows original and artificial data for a 5000 s interval of magnetic field data. Amplitude and width of the peaks of the CWTs are automatically evaluated and corrected via Eq. 2.16. The analyzed frequencies are given by Eq. 2.23 with $f_{max} = 0.5$ Hz and $f_{min} = 0.004$ Hz. Because of the random position and phase of the artificial waves, original and artificial data are not identical. Nevertheless, one can get an impression of the similarity between original and artificial data. The artificial signal recreates the shape of the power-spectral density for a wide range of frequencies. Only for $f < f_{min}$, the artificial signal fails to recreate the power-spectral density.

Several signals with exotic power-spectral densities were tested as well, e.g. box shaped power spectral densities, $I \propto f$ or $I \propto f^{-1}$. In all cases, the artificial signal recreated the original power-spectral density well. One can therefore assume that the statistics about amplitude, width and frequency of the waves contain the information to recreate a similar signal with the same power-spectral density.

Determining the polarization of waves

Using two perpendicular components of electromagnetic fluctuations, the CWT can be used to identify the mode of polarization. The CWT with a Morlet wavelet returns a complex signal. For two given perpendicular signals $x(t)$, $y(t)$ the angle in the complex

plane between their two CWTs

$$\phi := \arctan\left(\frac{\Im(W_{x(t)}/W_{y(t)})}{\Re(W_{x(t)}/W_{y(t)})}\right) \tag{2.25}$$

states their mode of polarization. For example, the angle between the wavelet transformations of

$$\begin{aligned} x(t) &= \cos(\omega t) \\ y(t) &= \cos(\omega t + \phi) \end{aligned}$$

is given by ϕ. Linear polarized waves will, therefore, be identified by $\phi = 0, \pm\pi$, while circular polarized waves are given by $\phi = \pm\pi/2$, where $\pm$ states the sense of rotation. Other ϕ values can consequently be attributed to elliptical polarizations.

2.1.2 Magnetic fluctuations at 1 AU in ACE MAG data

ACE resides at the first Lagrangian point and the MAG instrument provides 1-second averaged magnetic field data since 1999 [Smith *et al.*, 1998], providing data from both solar minimum and solar maximum. Hence, wave amplitudes can not only be compared for slow and fast solar wind but also for solar minimum and solar maximum conditions. For a simplified selection of fast and slow wind intervals, intervals according to the solar wind speed measured by the SWEPAM instrument are selected. Solar wind velocities below 420 km/s are classified as slow solar wind and solar wind velocities above 450 km/s are classified as fast solar wind. In each interval, the mean magnetic field is calculated to determine the components of the magnetic fluctuations parallel and perpendicular to the mean magnetic field. The years 1999 - 2003 were analyzed representing solar maximum, the years 2005 - 2008 were analyzed representing solar minimum conditions.

Amplitude distribution

To get an impression of the amplitude distribution, several intervals of slow and fast solar wind during solar minimum as well as solar maximum are analyzed for occurring wave amplitudes $B_\perp$ and width $2\sigma_B$ for fluctuations perpendicular to the mean magnetic field, as described in the previous chapter. The amplitude distributions $p_f(B_\perp)$ is created from a histogram of detected amplitudes, where each detected amplitude is scaled with the corresponding wave width.

Fig. 2.8 shows the amplitude distribution $p_f(B_\perp)$ for magnetic fluctuations perpendicular to the mean magnetic field, for three different intervals of slow solar wind. One can clearly see an exponential distribution of wave amplitudes for all analyzed frequencies and solar wind intervals. For the three analyzed intervals, one finds large differences in the slopes of the amplitude distributions, but in each interval the mean wave amplitude increases with decreasing frequency. The same behavior and the same exponential distributions can be found for the vast majority of all intervals of fast and slow solar wind as well as for solar minimum and maximum. Therefore, we assume an exponential distribution of amplitudes for each analyzed frequency f

$$p_f(B_\perp) = \frac{1}{\bar{B}_\perp} \exp\left(-\frac{B_\perp}{\bar{B}_\perp}\right), \tag{2.26}$$

where $\bar{B}_\perp$ is the mean amplitude found in the analyzed interval at the observed frequency. According to Fig. 2.8, there is a strong variation of $\bar{B}_\perp$ for different solar wind intervals in addition to the systematic increase with decreasing frequency.

To get an impression of the variance of the mean amplitude for different solar wind intervals, Fig. 2.9 shows the mean amplitude of the waves perpendicular to the mean magnetic field $\bar{B}_\perp$ for all slow wind intervals from 1999. Only those intervals which are sufficiently long to create appropriate amplitude statistics are considered. For a given frequency, $\bar{B}_\perp$ shows a high variability with a standard deviation $\sigma_{\bar{B}_\perp}$ in the same order of magnitude as its mean value $\langle \bar{B}_\perp \rangle$. A systematic analysis yields the relation $\sigma_{\bar{B}_\perp} \approx 0.5 \langle \bar{B}_\perp \rangle$.

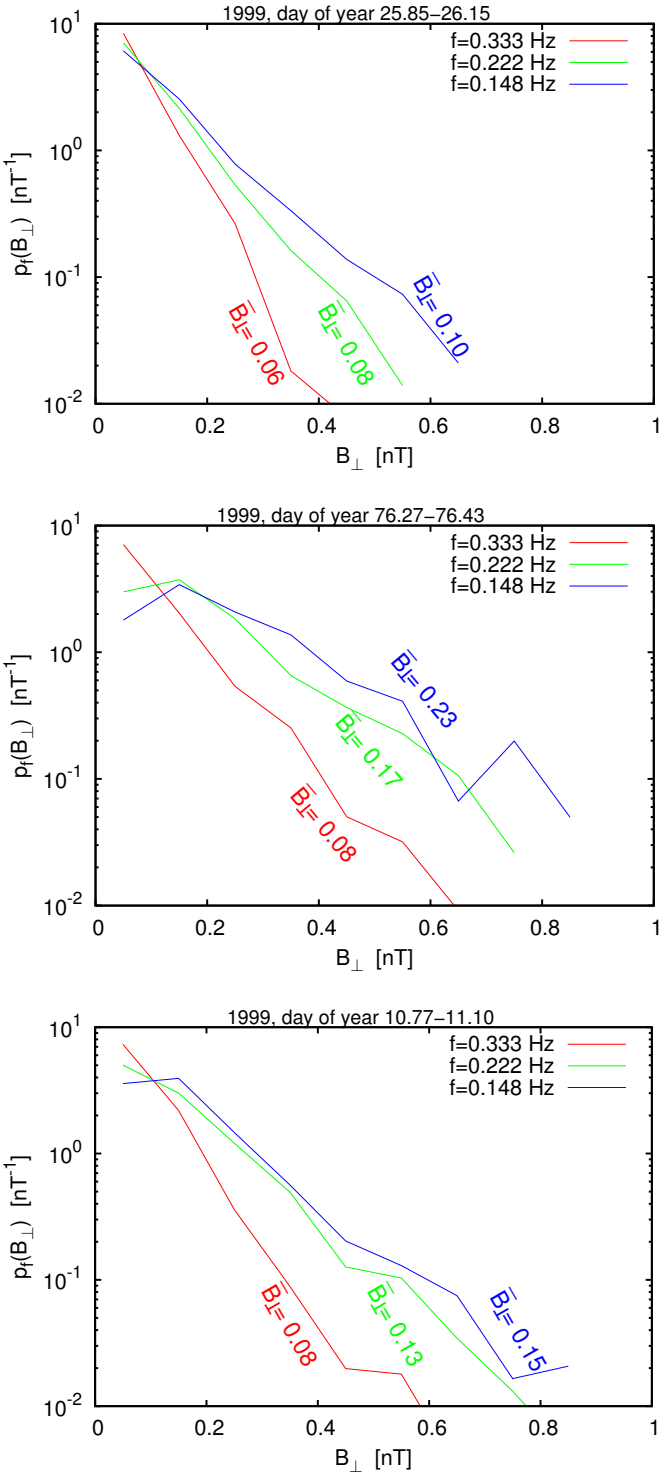

Figure 2.8: Examples of amplitude distributions for magnetic fluctuations perpendicular to the mean magnetic field for three different intervals of slow solar wind. For each distribution the mean amplitude $\bar{B}_\perp$ is specified in nT. The amplitudes are exponentially distributed. Although all analyzed intervals are in slow solar wind, the slope of the distribution is different for each interval. In each example the slope (and therefore $\bar{B}_\perp$) increases with decreasing frequency.

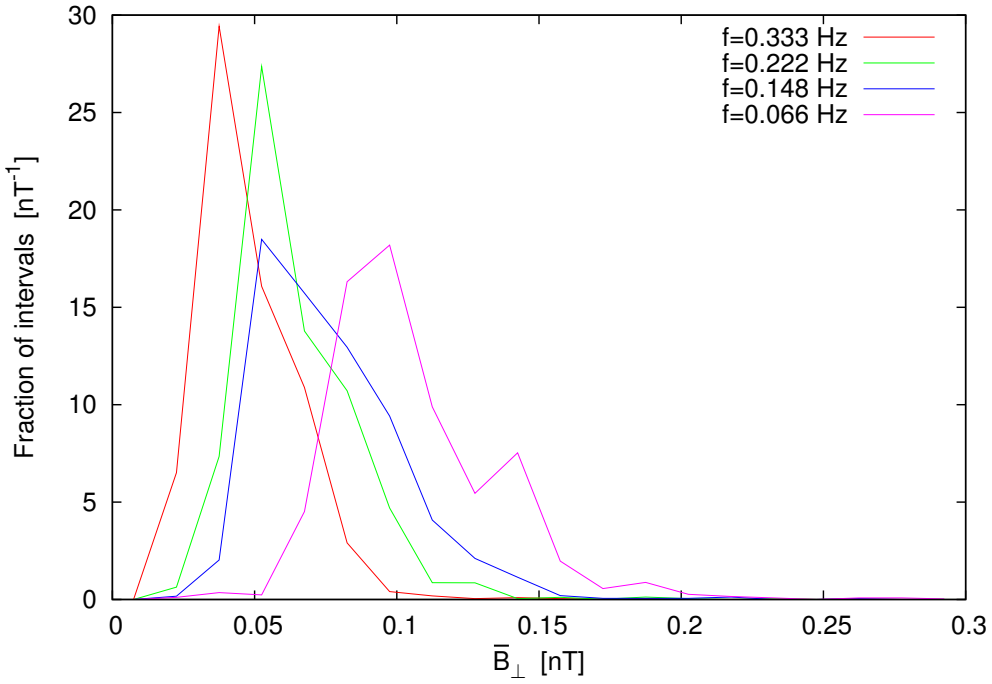

Figure 2.9: The calculated mean amplitudes perpendicular to the mean magnetic field $\bar{B}_\perp$ for all slow wind intervals in 1999. For the analyzed frequencies, there is a high variability of $\bar{B}_\perp$ values, whose deviation is in the same order of magnitude as their mean value. The average $\bar{B}_\perp$ value increases with decreasing frequency.

As already indicated in Fig. 2.9, the mean $\bar{B}_\perp$ values increase with decreasing frequencies, which means that for low frequencies there is an increased occurrence of large amplitude waves. In a next step, the mean amplitude $\bar{B}_\perp$ is systematically calculated for each slow and fast solar wind interval during solar minimum 2005 - 2008 and during maximum 1999 - 2003.

Fig. 2.10 shows an increase of the average mean amplitude $\langle \bar{B}_\perp \rangle$ with decreasing frequency and a higher value of $\langle \bar{B}_\perp \rangle$ during fast wind intervals. The steep decrease of the wave amplitudes above 0.2 Hz is due to the 1 second resolution of the magnetic field data. The increased wave activity during solar maximum is likely explained by the increased solar activity. The increase during fast solar wind can be explained considering that fluctuations in fast wind spent less time traveling from Sun to 1 AU and therefore the decay of the waves through the turbulent cascade is less advanced. This can also be observed in the power-spectral density [Tu and Marsch, 1995]. Fluctuations parallel to the mean magnetic field can be analyzed using the same approach. Fig. 2.11 shows the ratio of fluctuations parallel and perpendicular to the mean magnetic field. Fluctuations parallel to the magnetic field show almost the same behavior as fluctuations perpendicular to the magnetic field, but have significantly smaller amplitudes $\langle \bar{B}_\parallel \rangle / \langle \bar{B}_\perp \rangle \approx 0.8$.

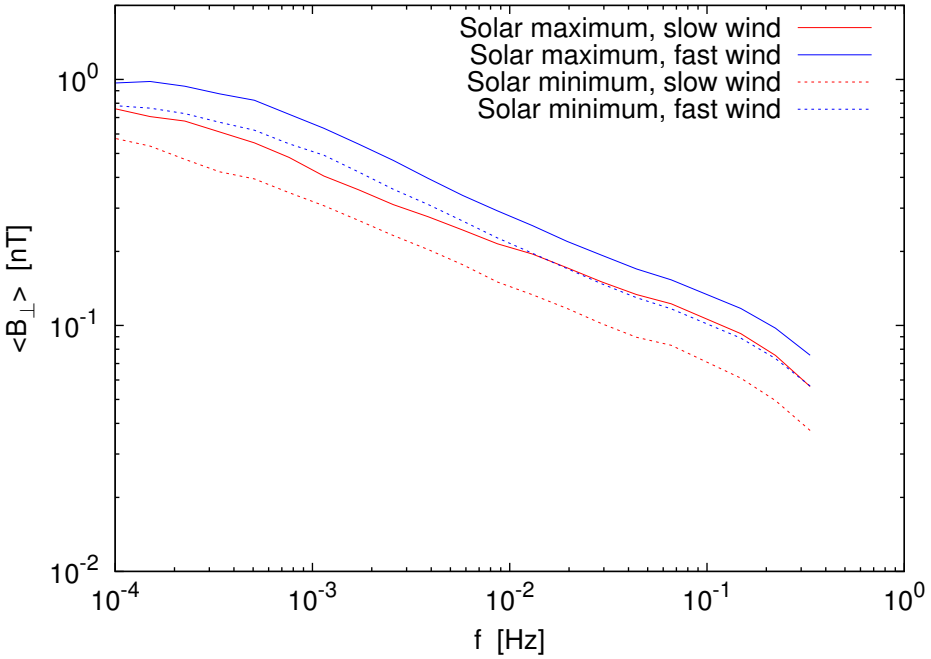

Figure 2.10: The averaged $\bar{B}_\perp$ values during solar minimum (1999 - 2003) and during solar maximum (2005 - 2008) for fast and slow solar wind. The corresponding standard deviations are roughly given by $0.5\langle\bar{B}_\perp\rangle$. The wave amplitudes increase for low frequencies and are higher during solar maximum. $\bar{B}_\perp$ increases towards low frequencies as well as for slow solar wind and during solar maximum.

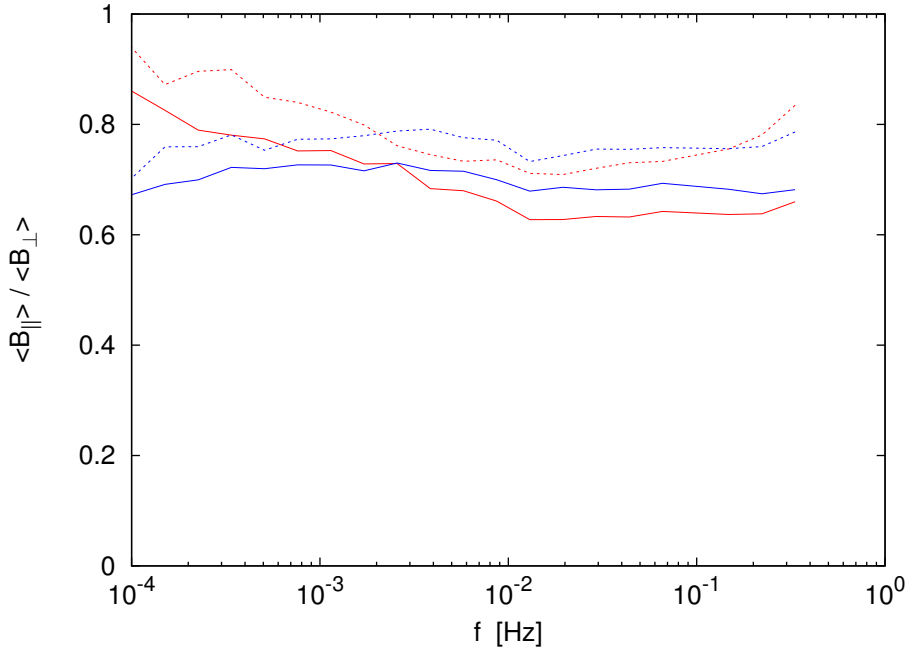

Figure 2.11: The ratio of the averaged $\bar{B}$ for amplitudes parallel and perpendicular to the mean magnetic field. Amplitudes parallel to the mean magnetic field show the same behavior as shown in Fig. 2.10, but are significantly smaller $\langle\bar{B}_\parallel\rangle/\langle\bar{B}_\perp\rangle \approx 0.8$.

Verifying results

Although not shown here, one can generate (via Eq. 2.24) an interval of artificial fluctuations with $\langle \bar{B} \rangle(f) \propto f^{-0.3}$ and amplitudes distributed via Eq. 2.26, as displayed in Fig. 2.10. The power-spectral density of the artificial data exhibits the same spectral index of $-5/3$ (i.e. Kolmogorov turbulence) as the measured data.

To understand why the relation $\langle \bar{B} \rangle(f) \propto f^{-0.3}$ leads to a power-spectral density with a spectral index of $-5/3$, the power-spectral density of a set of N short oscillating pulses with frequency ω_0, amplitude A, and positions t_i is examined. A single pulse is given by

$$x(t) = Ae^{-\frac{(t-t_i)^2}{2\sigma^2} + i\omega_0(t-\phi_i)} \tag{2.27}$$

where σ is given by $f_0^{-1} = 2\pi/\omega_0$. The Fourier transformation of a single wave is given by

$$\hat{x}(\omega) = A\sigma e^{-0.5\sigma^2(\omega-\omega_0)^2} \cdot e^{i\omega_0\phi_i + i(\omega_0-\omega)t_i}. \tag{2.28}$$

For waves in a sufficiently large interval, the power of the signal increases linear with the number of waves N.

$$\begin{aligned} P(\omega) &= \int d\omega |\hat{x}_{1,\ldots,N}(\omega)|^2 && (2.29)\\ &\approx N \int d\omega |\hat{x}_1(\omega)|^2 && (2.30) \end{aligned}$$

One can now compare the power-spectral density for an interval with length L, which is filled with $N = L/(4\sigma) = L/4 \cdot f_0$ oscillations, for different wave frequencies f_0:

$$\begin{aligned} P(\omega) &\approx N|\hat{x}_1(\omega)|^2 && (2.31)\\ &= L/4 \cdot f_0 \left| A\sigma e^{-0.5\sigma^2(\omega-\omega_0)^2} \cdot e^{i\omega_0\phi_i + i(\omega_0-\omega)t_i} \right|^2 && (2.32)\\ &= L/4 \cdot f_0 A^2\sigma^2 \cdot e^{-\sigma^2(\omega-\omega_0)^2} \left| e^{i\omega_0\phi_i + i(\omega_0-\omega)t_i} \right|^2 && (2.33)\\ &= L/4 \cdot A^2/f_0 \cdot e^{-\sigma^2(\omega-\omega_0)^2}, \quad \text{with } \sigma = f_0^{-1}. && (2.34) \end{aligned}$$

For a given wave frequency f_0, $P(\omega)$ is Gaussian shaped, however, the height of the Gauss peak depends on f_0. Since the maximum of the exponential term equals 1, its height decreases with f_0^{-1}. For an interval of the length L which is filled with $N_j = L/4 \cdot f_j$ waves for the frequencies $f_j = a^j f_0$, this will result in a power-spectral density with a spectral index of -1. If the amplitudes are assigned according to Fig. 2.26 and $A_j \propto f_j^{-0.3}$, the power-spectral density for a single wave frequency f_j is given by

$$\begin{aligned} P(\omega) &\propto (f_j^{-0.3})^2/f_0 |e^{-0.5\sigma^2(\omega-\omega_j)^2}|^2 && (2.35)\\ &\propto f_j^{-1.6} |e^{-0.5\sigma^2(\omega-\omega_j)^2}|^2, && (2.36) \end{aligned}$$

resulting in an overall spectral index of $-1.6 \approx 5/3$, which is the spectral index of turbulent magnetic fluctuations in the solar wind at 1 AU.

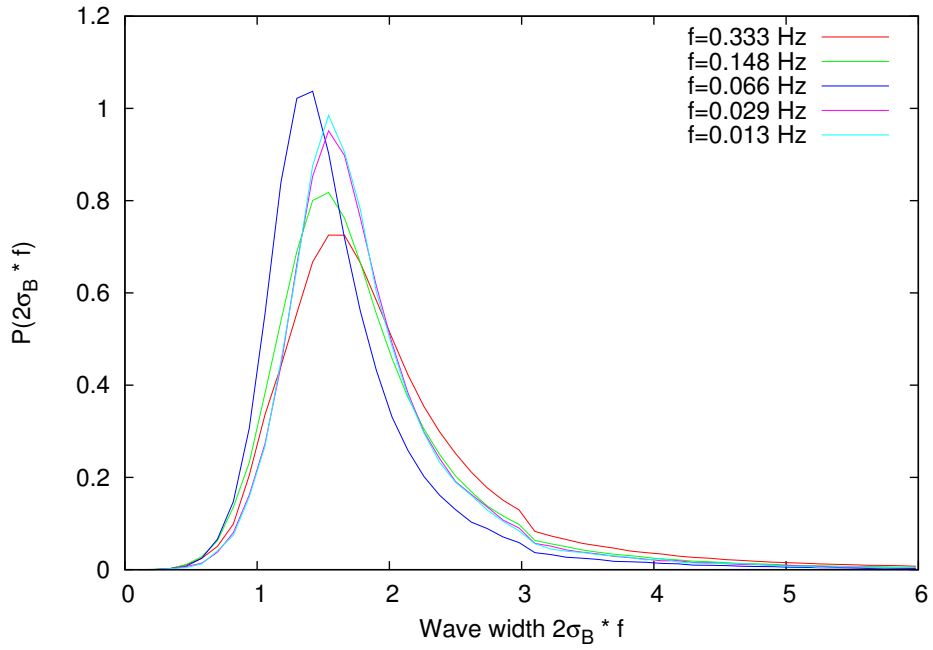

Figure 2.12: Distribution of wave width $2\sigma_B$ for slow wind intervals in 1999 scaled with the analyzed frequency f shown for various frequencies. The mean width of the waves is given by $2\sigma_B \cdot f \approx 2$, independently of the analyzed frequency.

Wave width

As mentioned before, the width of the detected waves $2\sigma_B$ depends on the width of the wavelet. Nevertheless, Fig. 2.12 shows the distribution of wave widths, scaled with the analyzed frequency for the slow wind intervals in 1999. For each frequency the distribution shows a clear maximum which is located near $2\sigma_B \cdot f = 2$, i.e. 2 oscillations in $2\sigma_B$. For the range of analyzed frequencies the distributions look very similar and are independent of solar wind speed or solar cycle.

Polarization of waves

Fluctuations which propagate parallel to the mean magnetic field can be classified by their sense of rotation with respect to the mean magnetic field. L-/R-mode waves rotate in a left/right-hand direction with respect to the mean magnetic field. Analyzing two perpendicular components of the magnetic fluctuations which are also perpendicular to the mean magnetic field, Eq. 2.25 can be used to calculate the mode for individual waves.

Fig. 2.13 shows examples of the distribution of wave polarizations, calculated as described in Sec. 2.1.1. For the analyzed frequency range, clear maxima can be found at $\phi = 0, \pm\pi$, independent of solar wind type or solar cycle,. The major part of the magnetic

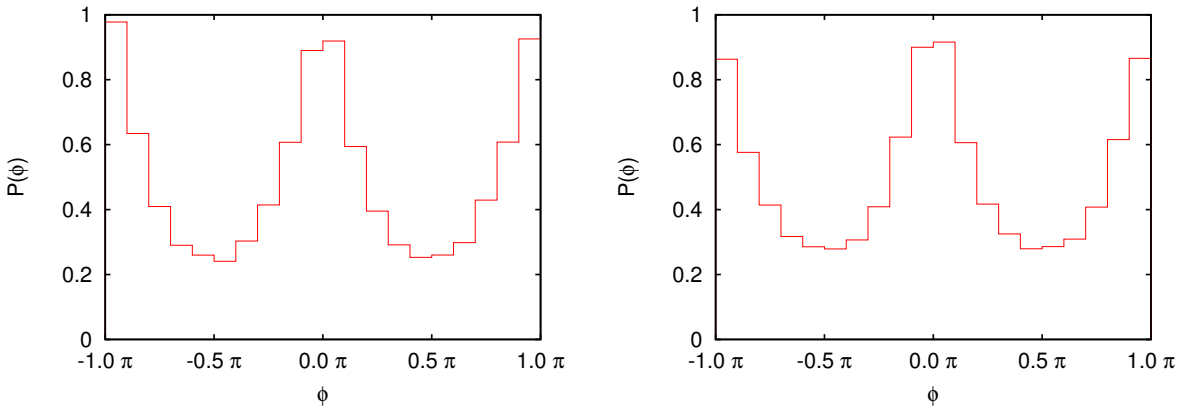

Figure 2.13: Examples of wave polarization distribution for waves in slow solar wind with $f = 0.0087$ Hz (left) and for fast solar wind with $f = 0.029$ Hz (right) in 1999. ϕ states the phase between the two perpendicular components of the wave. The maxima at $\phi = 0, \pm\pi$ are caused by linear polarized waves and 2D structures.

fluctuations in the solar wind is given by 2D structures, which show wave vectors nearly perpendicular to the mean magnetic field and can be understood as static structures convected past the observer. Therefore, the 2D structures do not display a circular polarization and contribute mainly to $\phi = 0, \pm\pi$. Systematically comparing $\langle \bar{B}_\perp \rangle$ values for circular ($0.25\ \pi < \phi < 0.75\ \pi$ or $-0.25\ \pi > \phi > -0.75\ \pi$) and linear ($-0.25\ \pi < \phi < 0.25\ \pi$, $\phi > 0.75\ \pi$ or $\phi < -0.75\ \pi$) polarized waves for different frequencies and solar conditions, yields a fraction of $\approx 30\%$ circular polarized waves. The mean amplitudes for circular and linear polarized waves show a ratio of $\langle \bar{B}_\perp \rangle_{circ} / \langle \bar{B}_\perp \rangle_{lin} \approx 0.8$.

2.1.3 Magnetic fluctuations between 0.3 and 1 AU in Helios data

The Helios mission performed in-situ measurements of the interplanetary medium (especially the interplanetary magnetic field) in the inner heliosphere. The orbits of the two identical spacecraft Helios 1 and Helios 2 ranged from aphelion 1 AU to perihelion 0.29 and 0.31 AU. Their heliographic latitudinal range covered ±7.25 degrees and lies therefore nearly in the ecliptic plane. The spacecraft were launched on December 10, 1974 and January 15, 1977 and their missions ended on September 4, 1985 and March 9, 1980. A general overview over the spacecraft mission strategy and the scientific instruments on board of Helios 1/2 is given in Porsche [1977, 1984]. The highly elliptical orbits in the ecliptic plane of the spacecraft yield the possibility to study magnetic field fluctuations for a wide range of heliocentric distances.

Slow and fast solar wind intervals were classified as described in the previous section. The solar wind speed was obtained from the Plasma instrument. In contrast to ACE data, solar minimum and maximum conditions were not considered. The sampling frequency in the Helios magnetic field measurements is irregular, time between two data points ranges from 45 seconds to several minutes. Therefore, for each analyzed frequency an interval with an adequate sampling frequency and length is selected. To convert an interval with irregular sampling to a fixed sampling frequency, a cubic interpolation is used. Data is selected from the entire mission time from both spacecraft, Helios 1 and Helios 2.

Amplitude distribution

After verifying that the amplitude distribution in the Helios magnetic field data exhibits the same exponential distribution and variation as the ACE data at 1 AU, the mean wave amplitude $\langle \bar{B}_\perp \rangle$ can be evaluated for several heliocentric distances. Fig. 2.14 (top) shows $\langle \bar{B}_\perp \rangle$ for slow (solid) and fast (dashed) solar wind and for the distances 0.3 - 0.4 AU, 0.5 - 0.6 AU, and 0.9 - 1.0 AU. For all three distances, the mean amplitudes show the same frequency dependence as found in the ACE data. As expected, there is an increase of the mean amplitudes with decreasing heliocentric distance. Fig. 2.14 (bottom) compares the Helios data from 0.9 - 1.0 AU with the ACE data at 1 AU. Although the mean amplitudes agree reasonable well, the Helios data show an increased $\langle \bar{B}_\perp \rangle$ for low frequencies.

Since the mean amplitudes increase with decreasing heliocentric distance, it is expected that at 0.9 - 1 AU the mean amplitudes are increased compared to those in ACE data. The dependence of the mean amplitude on the heliocentric distance is visualized in Fig. 2.15. $\langle \bar{B} \rangle_f(r)$ is shown for several frequencies of slow and fast solar wind for the distances 0.3 - 1.0 AU. The mean amplitudes clearly show the dependency $\langle \bar{B}_\perp \rangle_f(r) \propto r^{-2}$. The heliospheric magnetic field in the ecliptic can be approximated by [Prölss, 2003]

$$B = B(r_0)\sqrt{1 + \left(\frac{\Omega_s r}{u_{SW}}\right)^2}\left(\frac{r_0}{r}\right)^2, \tag{2.37}$$

where Ω_0 is the solar rotation and u_{SW} is the solar wind speed. Therefore, the heliospheric magnetic field deceases by $B \propto \sqrt{r^{-4} + cr^{-2}}$, but for $r < 1$ AU the dominant term in

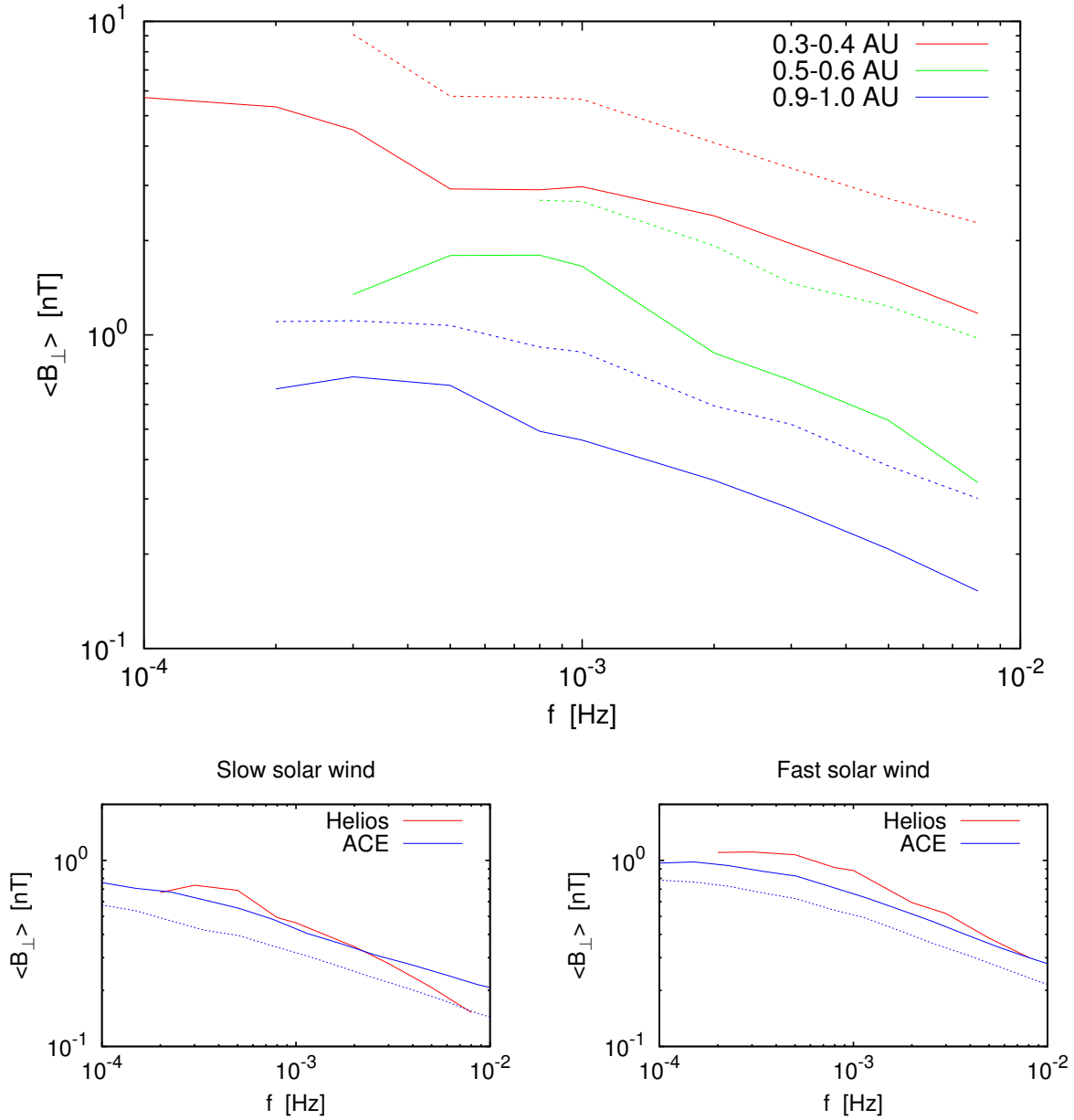

Figure 2.14: Top: $\langle \bar{B}_\perp \rangle$ for slow (solid) and fast (dashed) solar wind from Helios 1/2 data, for the distances 0.3 - 0.4 AU, 0.5 - 0.6 AU, and 0.9 - 1.0 AU. For all distances, $\langle \bar{B}_\perp \rangle$ shows the same behavior as in Fig. 2.10. $\langle \bar{B}_\perp \rangle$ values with insufficient statistics were omitted. Bottom: A direct comparison of Helios data from 0.9 - 1.0 AU and ACE data at 1 AU. Solar maximum solid, solar minimum dashed.

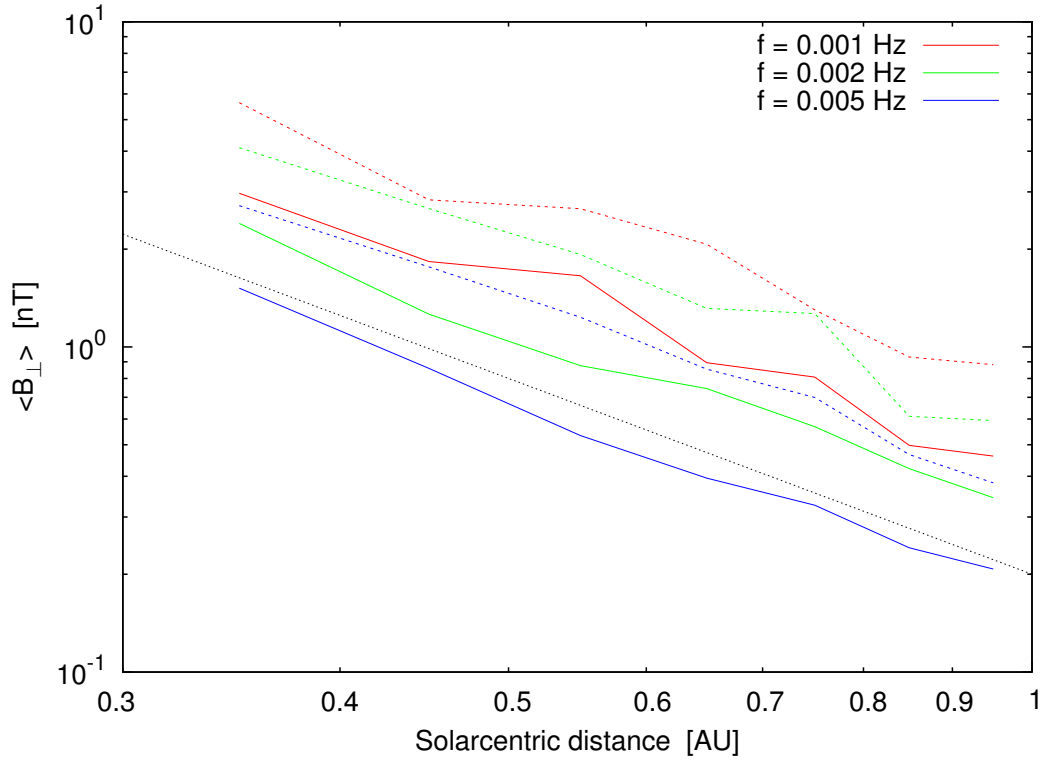

Figure 2.15: $\langle \bar{B}_\perp \rangle$ for several frequencies of slow (solid) and fast (dashed) solar wind for the distances 0.3 - 1.0 AU in steps of 0.1 AU. For comparison a r^{-2} power law is plotted as well (black curve).

Eq. 2.37 decreases with r^{-2}. This means that for $r < 1$ AU the mean amplitudes of the fluctuations decrease with approximately the same rate as the heliospheric magnetic field.

2.1.4 Magnetic fluctuations between 1 and 5 AU in Ulysses data

The Ulysses mission studied the inner heliosphere in three dimensions. The spacecraft was placed in a high-inclination, heliocentric orbit with aphelion at 5.4 AU and perihelion at 1.3 AU. Within this orbit Ulysses reached latitudes up to 80° [Wenzel *et al.*, 1992]. The mission started on October 6, 1990 and ended on July 1, 2008 when the power output from the RTG had decreased so far that there was insufficient power for the internal heaters.

The magnetic field measurements on Ulysses were performed by the Vector Helium Magnetometer and the Fluxgate Magnetometer [Balogh *et al.*, 1992] and provide the three components of the magnetic field vector in 1 second resolution. The high-inclination, elliptical orbit does not only provide the possibility to continue the investigation of the radial dependence of $\langle \bar{B} \rangle$, but also to study the latitudinal dependence of the mean amplitudes. Slow and fast solar wind intervals were classified as described in the previous sections. The solar wind speed was obtained from the SWOOPS instrument.

Amplitude distribution

The dependence of the mean amplitude of the magnetic fluctuations on frequency and heliocentric distance in the ecliptic plane at 0.3 - 1.0 AU has been evaluated in the previous sections for frequencies ranging from 10^{-4} - 0.5 Hz. To continue the investigation of the radial evolution of the magnetic fluctuations, Ulysses magnetic field data in the ecliptic plane, i.e. lat.$\pm 10°$, were analyzed. In addition, the time interval was restricted to the years 1990 - 1999, where Ulysses traveled from 1 to 5 AU in the ecliptic plane. Using this comparatively short time minimizes the effects of long term variations like the solar cycle dependency. Fig. 2.16 (top) shows $\langle \bar{B} \rangle$ for slow and fast solar wind and for the distances 1 - 2 AU, 2 - 3 AU, 3 - 4 AU, and 4 - 5 AU. For all distances the mean amplitudes show the same frequency dependence as in the ACE and Helios data. As expected from the Helios measurements, there is an increase of the mean amplitudes with decreasing heliocentric distance. In Fig. 2.16 (bottom) the Ulysses data from 1 - 2 AU is compared to the ACE data at 1 AU. The mean amplitudes in the Ulysses data are decreased due to the increased heliocentric distance but agree very well with shape and intensity of the mean amplitudes in ACE data.

Radial dependence of the amplitude distributions

The dependence of the mean amplitude on the heliocentric distance is visualized in Fig. 2.17. $\langle \bar{B} \rangle_f(r)$ is shown for several frequencies of slow and fast solar wind for the distances 1 - 5 AU. While for $r < 1$ AU the mean amplitudes show the dependence $\langle \bar{B} \rangle_f(r) \propto r^{-2}$, an analysis of the radial evolution of the mean amplitudes for all frequencies yields the relation $\langle \bar{B} \rangle_f(r) \propto r^{-(1.5 \pm 0.15)}$ for $r > 1$ AU. As discussed above, the magnitude of the heliospheric magnetic field decreases with $B \propto \sqrt{r^{-4} + cr^{-2}}$, i.e. a superposition of an r^{-2} and an r^{-1} decrease. For distances above 1 AU, the r^{-1} part becomes dominant. This means that for $r > 1$AU the mean amplitudes of the fluctuations decrease significantly

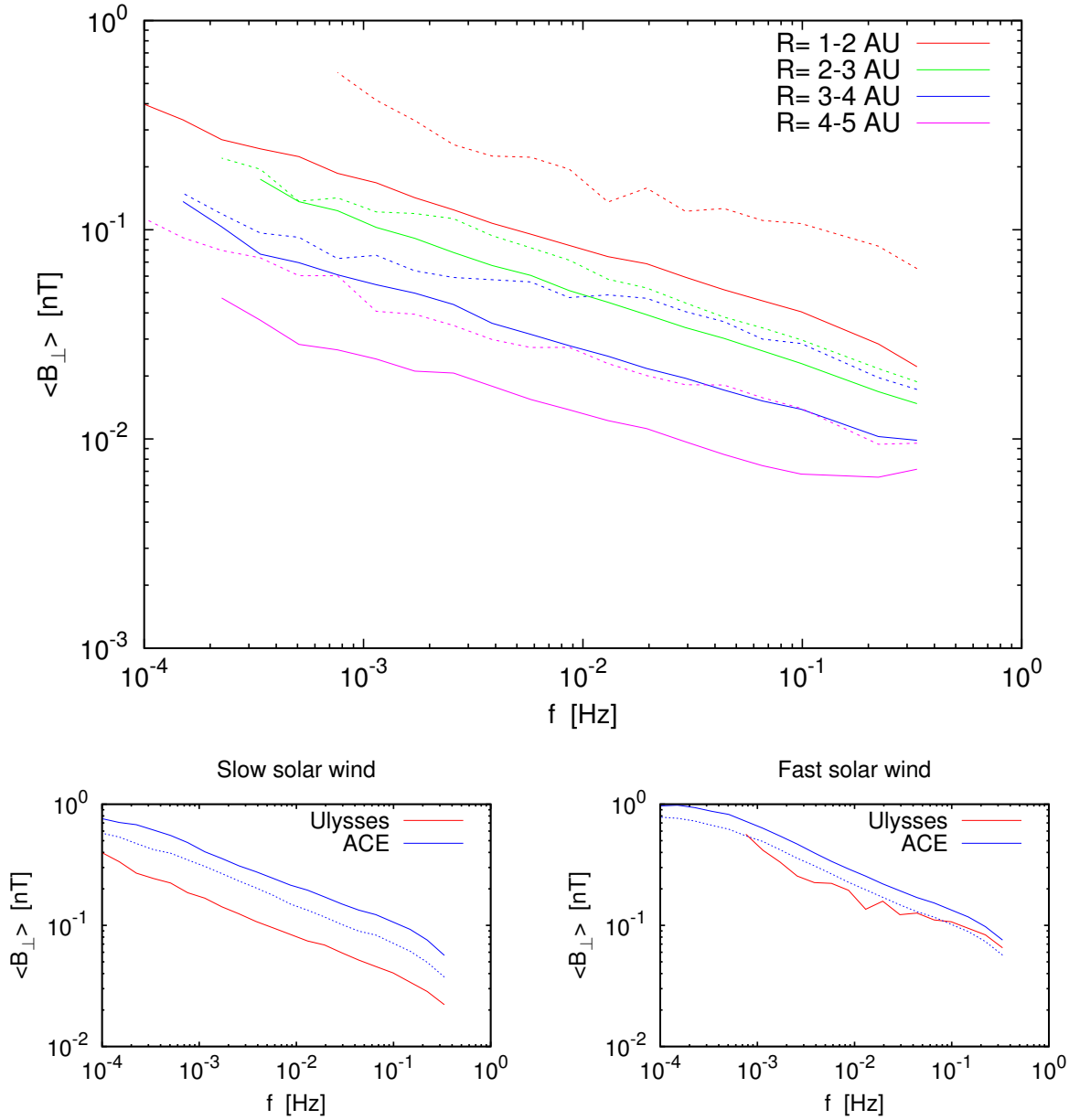

Figure 2.16: Top: $\langle \bar{B} \rangle$ for slow (solid) and fast (dashed) solar wind Ulysses data for distances 1 - 5 AU. For all distances, $\langle \bar{B} \rangle$ shows the same behavior as ACE and Helios data. $\langle \bar{B}_\perp \rangle$ values with insufficient statistics were omitted. Bottom: A direct comparison of Ulysses data from 1 - 2 AU and ACE data at 1 AU. Solar maximum solid, solar minimum dashed.

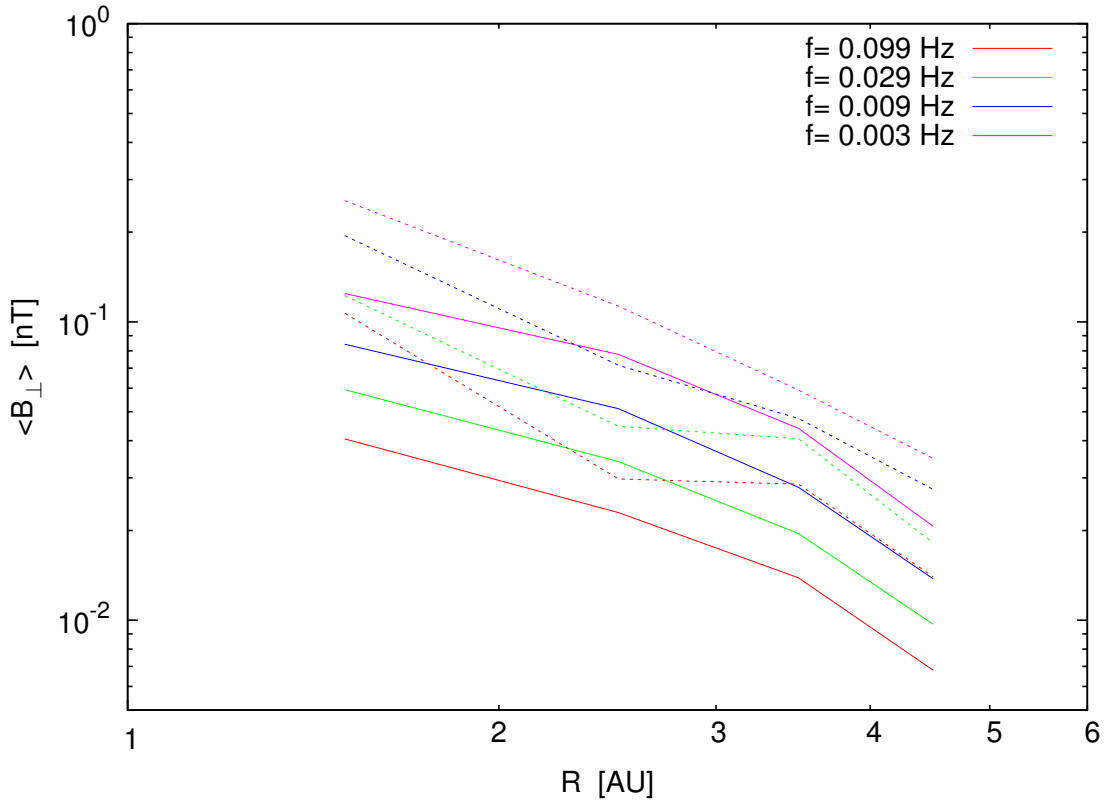

Figure 2.17: $\langle \bar{B} \rangle$ for several frequencies of slow (solid) and fast (dashed) solar wind for the distances 1 - 5 AU in steps of 1 AU.

faster than the heliospheric magnetic field, which is caused by dissipation of the magnetic fluctuations [Marsch, 1991].

Latitudinal dependence of the amplitude distributions

Since the orbit of Ulysses leaves the ecliptic plane, it offers the possibility to study not only the radial but also the latitudinal dependency of the amplitude distributions. To evaluate the latitudinal dependency at a certain radial distance, the interval 1 - 2 AU is chosen, in which the Ulysses spacecraft covers latitudes from 0° to 80° ("fast latitude scan").

A first analysis shows that for all observed frequencies the mean amplitudes decrease with latitude ϕ. From the ecliptic plane to the near polar region $\langle \bar{B} \rangle$ decreases by a factor of ≈ 0.35. However, the same decrease can be found for the magnitude of the magnetic field and is due to the small but still significant increase of the heliocentric distance from the ecliptic plane (1.3 AU) to the near polar region (2 AU). Compensating this effect by scaling the mean amplitudes with the magnitude of the magnetic field results in a constant distribution of mean amplitudes over latitude.

2.2 Particle propagation in the solar magnetic field

The focused transport equation, which is the standard approach for calculating the propagation of SEPs, models the guiding center motion of particles and calculates pitch-angle scattering via a diffusive term. The approach presented in this work takes a step back and calculates the exact particle trajectory in a background field superimposed with magnetic fluctuations. The particle motion is given by the Lorentz force

$$\dot{\vec{p}} = q\left(\vec{E}_A + \vec{v} \times (\vec{B}_0 + \vec{B}_A)\right) \tag{2.38}$$

where $\vec{v}$, $\vec{p}$, q are the particle velocity, momentum and charge, $\vec{B}_0$ is the ambient magnetic field given by the Parker spiral and $\vec{E}_A$, $\vec{B}_A$ are the superposed electromagnetic fluctuations. Instead of lavishly modeling the Parker field in three dimensions around the Sun, the field is modeled as a straight field line in z-direction, where the position z along the field line and the heliocentric distance r are linked via

$$\delta z = \delta r \sqrt{1 + \left(\frac{\Omega}{u_{\mathrm{sw}}}\right)^2 r^2}, \tag{2.39}$$

where u_{sw} is the solar wind speed and Ω is the solar sideral rotation rate. The strength of the Parker field is given by

$$|\vec{B}(r)| = B_0 \left(\frac{r_0}{r}\right)^2 \sqrt{1 + \left(\frac{\Omega}{u_{\mathrm{sw}}}\right)^2 r^2}, \tag{2.40}$$

where B_0 is the strength of the magnetic field at r_0 [Agueda, 2008]. To fulfill the constraint $\nabla \cdot \vec{B} = 0$, the individual components of the magnetic field (in cylindrical coordinates $B_z(z),\ B_\rho(z),\ B_\phi(z)$) are linked via

$$-\frac{\partial}{\partial z} B_z(z) \cdot \rho = B_\rho(z) \tag{2.41}$$

and $B_\phi = 0$. Using Eq. 2.40 and 2.41, the magnetic field along z can be constructed in a Cartesian coordinate system. Detailed statistics of the magnetic fluctuations were obtained in the previous section. As described in Eq. 2.24, the fluctuations are Monte-Carlo (MC) generated as a set of discrete Gaussian wave packets which are transverse and axially symmetric, with wave vectors parallel to the mean magnetic field $\vec{B}_0$. Since the velocity of the particles is large compared to the temporal evolution of the waves, the fluctuations are treated as magneto static, which results in $\vec{E}_A = 0$. The effect of neglecting the temporal evolution has been studied for simple wave configurations, and was found to be negligible.

Because this is an ab-initio calculation, systematic processes, such as pitch-angle focusing and drifts which are explicitly considered in the diffusion approximation Eq. 2.1, are already included. Like in the slab model, particles with gyration time T will be influenced by fluctuations with wavenumber k when the resonance condition Eq. 2.4 is fulfilled. However, these resonant wave-particle interactions will be calculated ab-initio by the Lorentz force.

2.2.1 Wave-particle interactions with a single wave

Before the full transport model from Sun to Mars is applied, pitch-angle scattering on one Gaussian shaped wave is analyzed. This allows us to test the model for a far simpler field geometry and to get an impression of the expected wave-particle interactions. The magnetic field is given by a homogeneous field along the z-axis which is superposed with a circular polarized Gaussian shaped wave with field vectors in x and y direction. The magnetic field in the simulation is then given by:

$$\begin{aligned} B_x(z) &= A \cdot \cos(f/u_{sw} \cdot z) \cdot \exp\left(0.5 \, \frac{z^2}{(f/u_{sw})^2}\right) \\ B_y(z) &= \pm A \cdot \sin(f/u_{sw} \cdot z) \cdot \exp\left(0.5 \, \frac{z^2}{(f/u_{sw})^2}\right) \\ B_z(z) &= B_0. \end{aligned} \tag{2.42}$$

The solar wind speed is selected as $u_{sw} = 400$ km/s, the magnitude of the magnetic field is set to $B_0 = 4$ nT, the width of the wave is given by f/u_{sw}, the temporal evolution of the wave was neglected. Through the $\pm$ sign the wave's sense of rotation can be adjusted, i.e. an l- or r-mode wave. Depending on the sense of rotation, the resonance condition can or can not be fulfilled. If the $\pm$ sign is set to $+$, the wave and a positive charged particle with $\mu > 0$ have the same sense of rotation and can therefore interact resonantly. To evaluate the wave-particle interaction, protons with energy E are injected at $z = -10 \cdot f/u_{sw}$ with a homogeneous distribution of pitch-angle cosines ($\mu > 0$). In Fig. 2.18, the particle's pitch angles before and after passing the wave are compared for various wave frequencies and amplitudes. Unlike Eq. 2.4 suggests, the resonance condition is not only fulfilled for one pitch-angle (for a given energy and wave frequency) but also for a wide range of pitch angles. This can be explained by the frequency distribution of a Gaussian shaped wave, which is also a Gaussian shape and becomes narrower with increasing width of the Gauss shape. Consequently, for an increased wave width the resonance areas in Fig. 2.18 become narrower.
As expected, the change in the pitch angle increases with wave amplitude. Although the wave amplitudes are comparatively small ($B_0 = 4$ nT, $A = 0.1$ nT, compare ACE data), the resulting pitch-angle scattering can be observed up to $\Delta\mu$ =0.2, which does not agree with the idea of small waves causing small changes in the pitch angle.

The density of the points in Fig. 2.18 can be interpreted as probability distribution for $\Delta\mu$. For a resonant wave-particle interaction with small wave amplitudes, the probability density of $\Delta\mu$ shows maxima for the largest $|\Delta\mu|$ values and a minimum at $\Delta\mu = 0$, similar to the probability density for a pendulum, e. g. Fig. 2.18 ($E = 1$ MeV, $f = 5 \cdot 10^{-4}$ Hz, $A = 0.01$ nT for $\mu = 0.3$). For large amplitudes a folded structure arises, which shows multiple maxima, e. g. Fig. 2.18 ($E = 1$ MeV, $f = 5 \cdot 10^{-4}$ Hz, $A = 0.1$ nT for $\mu = 0.3$). The scattering is not only influenced by the resonance condition, but also by the phase between wave and particle. For a phase difference of 0, i.e. $(v_x, v_y) \parallel (B_x, By)$, the wave does not influence the particle (at least for a plain, perfect wave), for a phase difference of $\pm\pi/2$, i.e. $(v_x, v_y) \perp (B_x, By)$, the influence becomes maximum. This shows that there are

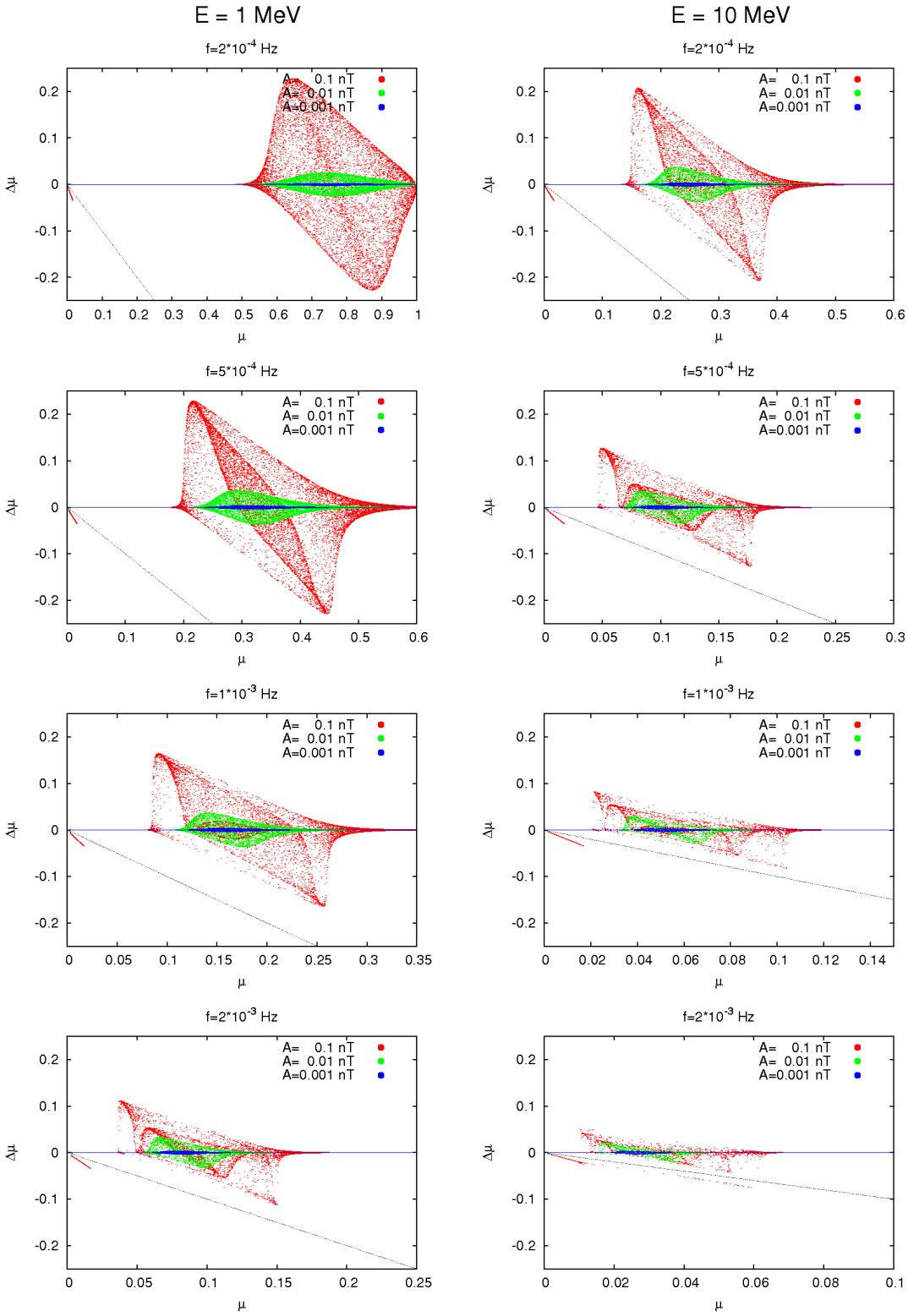

Figure 2.18: Scatter plot of pitch-angle change $\Delta\mu$ over the proton's pitch angle μ for one wave-particle interaction with a Gaussian shaped wave with amplitude A and frequency f. Compared are 1 MeV (left) and 10 MeV (right) protons. Note the different ranges for the μ-axis. Values below $\Delta\mu = -\mu$ (black line) are reflected particles.

phase-regimes with slow and with fast scattering. Depending on the initial phase difference between wave and particle, the resulting pitch angle ends up in one of the different maxima of the probability distribution.

For small μ, one can observe a sharp line of negative $\Delta\mu$ values. This line is caused by reflected particles. The reflection is caused by the increase of the magnetic field strength (Eq. 2.42), not by resonant wave-particle interactions. In the QLT, pitch-angle scattering through $\mu = 0$ is problematic, because of the resonance gap in the diffusion coefficient $D_{\mu\mu}$ [Dröge, 2003]. The scattering frequency (Eq. 2.3) becomes zero for μ with $q \neq 1$. This can be solved by introducing an additional ϵ value to fill this resonance gap. In this ab-initio simulation, the particle is not scattered through $\mu = 0$ by a resonant wave-particle interaction, but by a reflection on a magnetic fluctuation. Introducing an artificial process to transport particles through $\mu = 0$ is, therefore, not necessary.

2.2.2 Test-particle propagation from Sun to Mars

Computation time is an extremely crucial factor for this ab-initio calculation. When numerically solving Eq. 2.38, each step of the particle in the magnetic field has to be calculated, i.e. the magnetic field and the fluctuations need to be calculated. Therefore, the calculation of the magnetic fluctuations should be as simple as possible. This is achieved by generating the magnetic field before the simulation as a discrete set of data points along the z-axis. In the simulation the discrete magnetic field is linearly interpolated. The field data contains

- The coordinate along the z-axis, as a set of discrete points with $dz = 100$ km.
- The magnetic field in z direction, which is generated from a Parker field (Eq. 2.40) for a solar wind speed of 400 km/s.
- The magnetic field perpendicular to z in the form $\frac{\partial}{\partial z}B_z(z)$, which used is in the simulation with Eq. 2.41 to calculate the field vectors B_x, B_y.
- The magnetic fluctuations in x and y direction. The statistics for generating magnetic fluctuations via Eq. 2.24, in slow solar wind during solar minimum, were obtained in the previous section. The frequency range from 10^{-4} to 1 Hz was implemented, frequencies above 1 Hz do not contribute significantly to pitch-angle scattering because they are above the proton gyrofrequency. As shown in the previous section, the amplitudes of the magnetic fluctuations in different intervals show a high variability. Different magnitudes of fluctuations are realized by scaling the fluctuations with a factor a. Note that only parallel propagating waves were generated.

The field line is generated with a length of 2.5 AU and spans from a heliocentric distance of 0.3 AU to 1.93 AU. Mars is assumed to be at 1.524 AU, the field line spans up to 1.93 AU to allow SEPs to be scattered back towards Mars after they passed Mars. In a second step, the generated magnetic field is used for the ab-initio test particle propagation via Eq. 2.38. 10 MeV protons with an isotropic distribution of pitch-angle cosines ($\mu > 0$) are injected

at 0.3 AU and propagate along the field line, i.e. the z-axis. Their time-intensity profile at Mars is shown in Fig. 2.19 for the fluctuation scales $a = 0.5$, 1.0, 2.0,and 4.0. The different fluctuation scales were selected to account the high variability of the fluctuation intensity for different solar wind conditions. As expected, with an increase in a the particles arrive later at Mars. For $a = 0.5$, the intensity shows a clear peak at $t = 6100$ s, for $a = 2.0$, the intensity shows a maximum at $t = 8000$ s and decreases only very slowly. For $a = 0.5$, the pitch-angle scattering is relatively weak, therefore the pitch angles are focused towards $\mu = 1$ and the particles propagate fast along the field line. For $a = 2.0$, the pitch-angle scattering is strong and counteracts the pitch-angle focusing effect.

This can directly be seen in the pitch-angle distribution, which is shown in Fig. 2.20 for Mars at $t = 5000$ - 10000 s. The anisotropy decreases with increasing a, while for $a = 0.5$ almost all pitch angles are at $\mu = 1$, for $a \geq 1$ the pitch-angle distributions show even negative μ values. The anisotropy for $a = 2.0$ appears to be larger than for $a = 4.0$. Both distributions show the same time interval, however, for $a = 4.0$ the first particles arrive several 100 s later than for $a = 2.0$. Naturally, the first particles which arrive at Mars have large pitch angles, particles with small or negative pitch angles can be found only later in an event. In addition, Fig. 2.19 and 2.20, show the results for three generated field lines for each a. Although the power-spectral density of the magnetic fluctuations are the same, each generated field line has a unique composition of fluctuations. Both, the time-intensity profile and the pitch-angle distribution show a strong dependence on this composition. Although the power-spectral density is for each configuration the same, position and even shape of the time-intensity peak change for different field configurations. The corresponding pitch-angle distributions show different shapes as well. This effect is especially pronounced for small wave amplitudes ($a = 0.5$), where the effect of an average fluctuation is weak, but few large fluctuations have a huge influence.

Propagation and pitch angle of the particles for $a = 0.5$ are shown in detail in Fig. 2.21. Before $t = 3000$ s, the particles behave as expected. A large fraction shows pitch angles near $\mu = 1$ and particles with large μ propagate faster along the field line. At $t = 4000$ s, one can see a large irregularity, where one large wave scatters a significant fraction of particles from $\mu \approx 1$ to smaller pitch angles. This irregularity becomes less visible over the time, however, its effect can be seen in Fig. 2.19 and 2.20, comparing the blue, the red, and the green curves. The intensity reaches its maximum at a later time and the pitch-angle distribution is less anisotropic.

2.3 Conclusions

In the first part of this chapter, the composition of the magnetic fluctuations in the solar wind has been studied by using a modified wavelet analysis. Amplitudes and widths of the magnetic fluctuations were analyzed for a wide range of frequencies, for different heliocentric distances and solar wind conditions. The wave amplitudes show a high variability for different solar wind intervals. The amplitude distributions for defined frequencies in single intervals are given by an exponential behavior, where the mean amplitude increases with

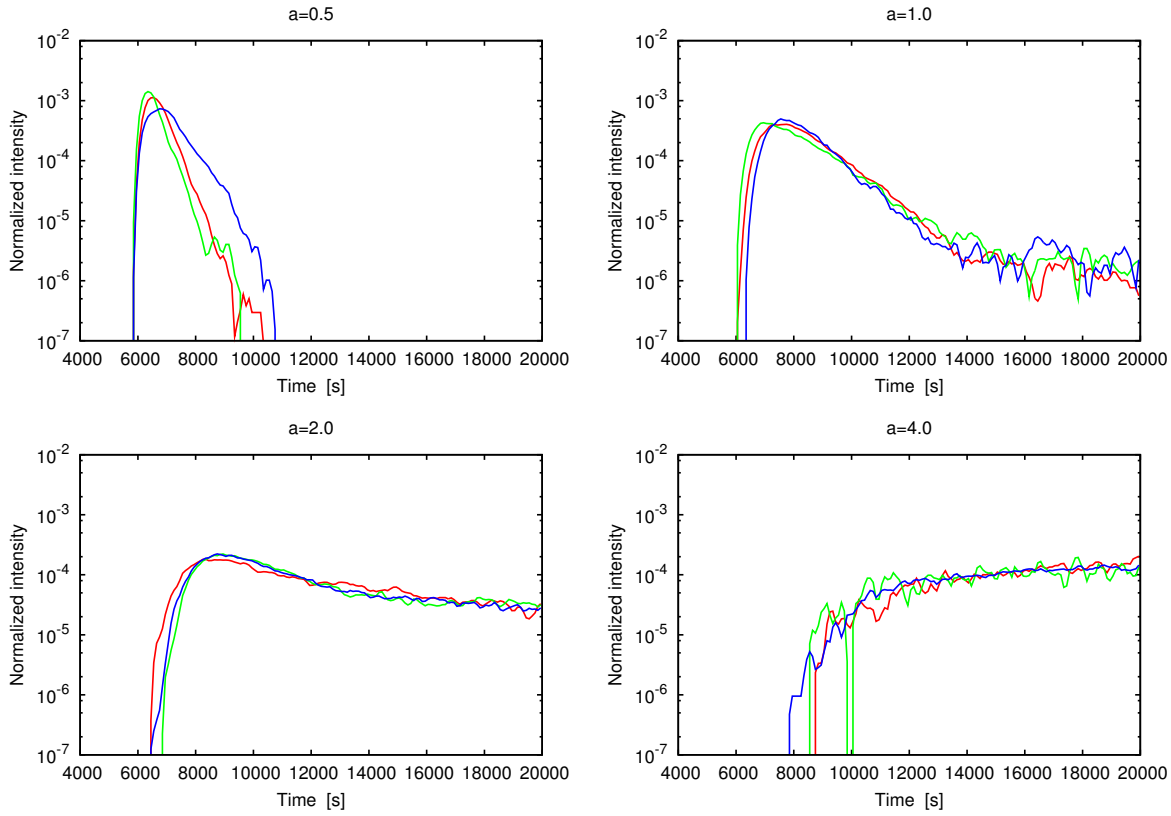

Figure 2.19: Time-intensity profile at Mars for 10 MeV SEP protons injected at 0.3 AU. The magnetic fluctuations are scaled with $a = 0.5$, $a = 1.0$, $a = 2.0$, and $a = 4.0$. The intensity is normalized to unity. Each plot shows the results for three generated field lines, which have the same power-spectral density, but different compositions of wave pulses.

decreasing frequency. The exponential distribution of wave amplitudes shows that even in solar wind intervals with weak magnetic fluctuations, there is a possibility of encountering some large amplitude waves.

In the second part of this chapter, an ab-initio model, which calculates a particle trajectory via the Lorentz force, was developed. Using single oscillating pulses with amplitudes as observed in the solar wind, it was shown that even comparatively small fluctuations can lead to large changes in the particle's pitch angle. The observed large changes of the pitch angles do not agree with the assumptions of the QLT that the irregularities of the magnetic field are sufficiently small and that the changes of the energetic particle's pitch angle during a single gyration are small [Dröge, 1994, 2003].

To investigate the influence of this non diffusive large pitch-angle scattering, the propagation of SEPs from Sun to Mars was calculated using the acquired magnetic-fluctuation data. A field line, based on the Parker field, was superposed with a set of magnetic fluctuations, recreating the power-spectral density observed in the solar wind. Although the

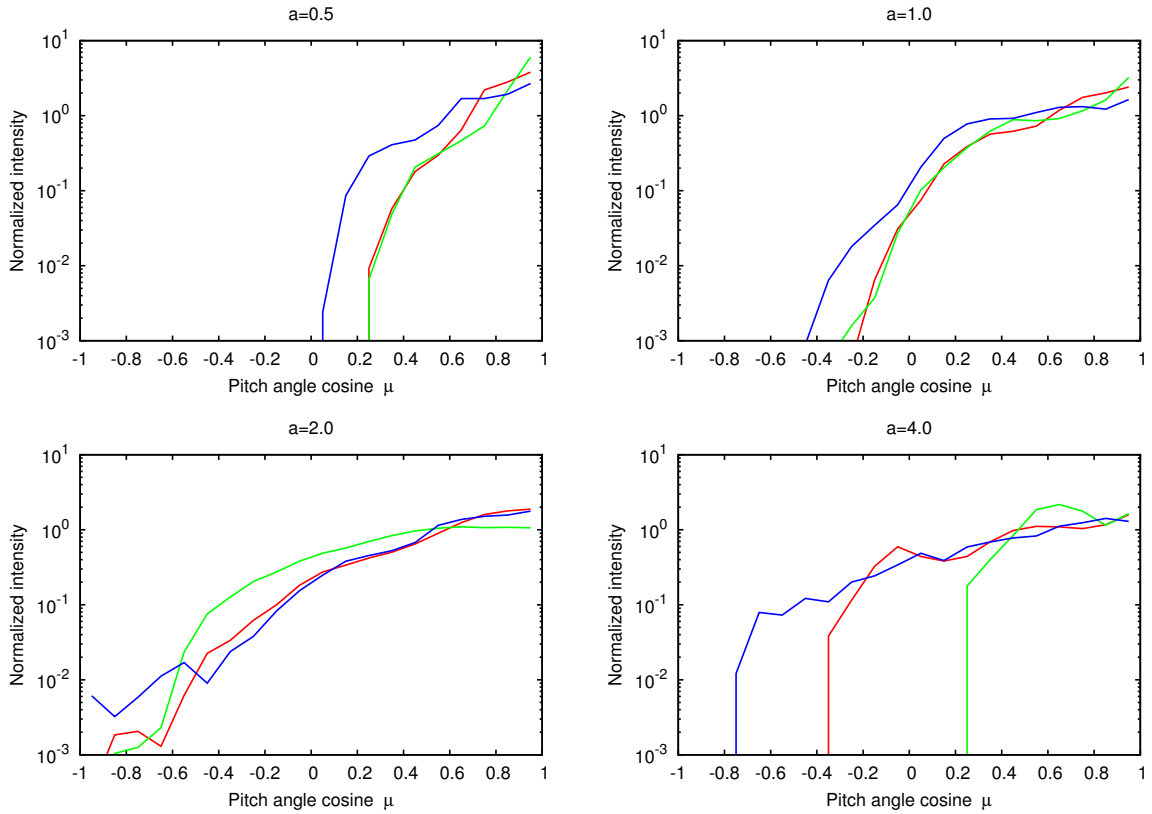

Figure 2.20: Pitch-angle distribution at Mars at $t = 5000$ - 10000 s for the same field lines as in Fig. 2.19.

same power-spectral density was used, time-intensity profiles and pitch-angle distributions vary for different configurations of the magnetic fluctuations. Especially for weak magnetic fluctuations, where the particles propagate more or less unperturbed, the presence of a few large fluctuations can change the particle propagation considerably.

For an exact description of wave-particle interactions, the knowledge of the spectrum is not sufficient. The power-spectral density does not contain any information about the rate of occurrence of fluctuations with a certain frequency and amplitude. A more detailed analysis of the magnetic fluctuation which takes the amplitude and rate of occurrence into account, enables us to model scattering events correctly even for rare strong scattering events where a diffusive approximation breaks down.
This also demonstrates the need to include the effects of wave particle scattering by large-amplitude waves in the focused transport equation (Eq. 2.1) for energetic particles. One possible approach is to include jump processes in addition to the usual diffusive pitch-angle scattering, as for example in the differential Chapman-Kolmogorov Equation [Gardiner, 2004].

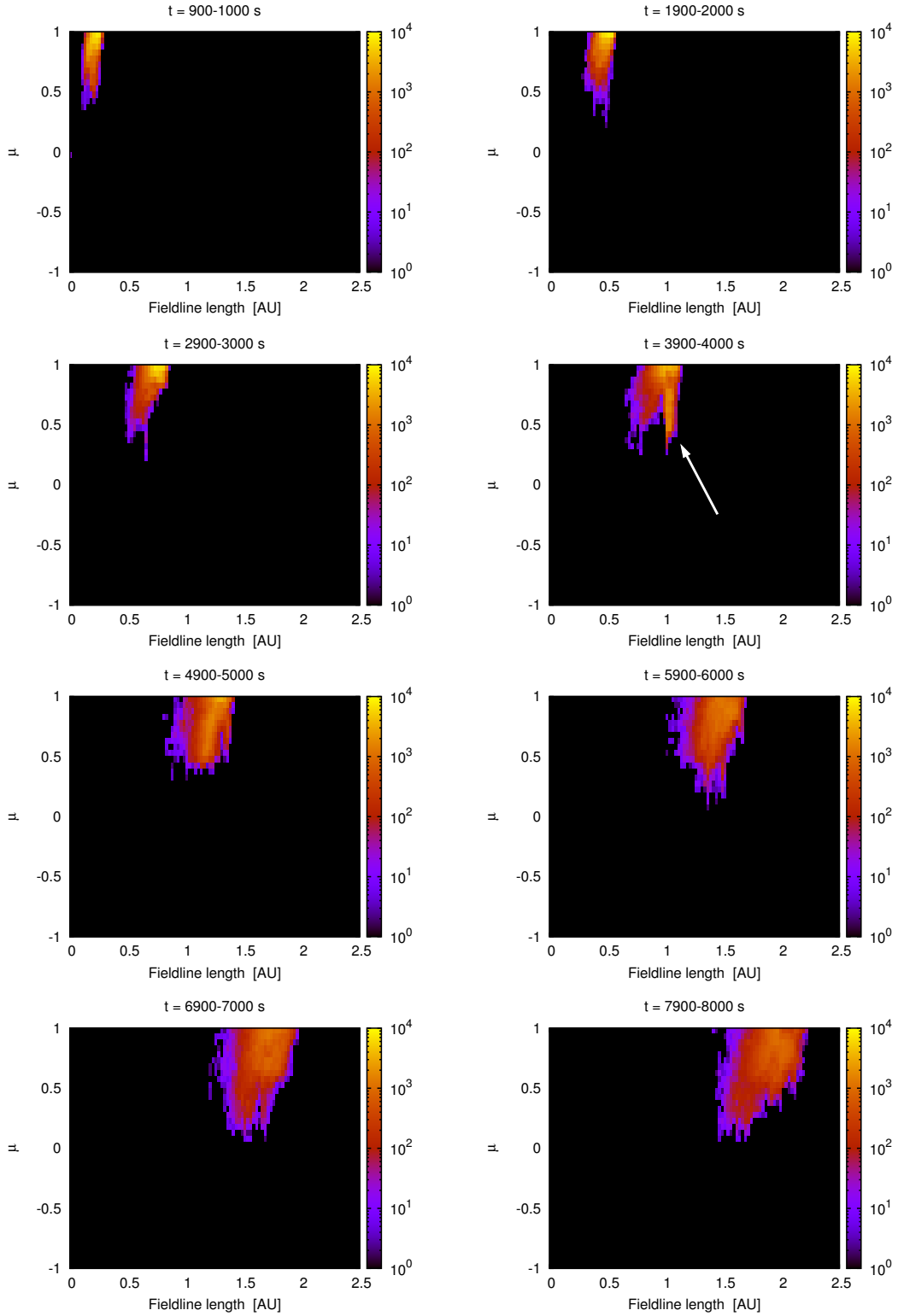

Figure 2.21: The particle's pitch angle and position along the field line at different times for the event shown in Fig. 2.19 and 2.20 ($a = 0.5$, blue curve). The field line ranges from 0.3 AU to 1.93 AU. Mars, which has a heliocentric distance of ∼1.524 AU, is located at field line length ∼ 1.685 AU. The arrow in t = 3900 - 4000 s points at an irregularity, caused by one strong scattering event.

Chapter 3

Gamma/neutron inversion for scintillator measurements

The RAD instrument includes two scintillators, the anorganic **D** detector, which is composed of CsI(Tl) and is highly sensitive to gammas, and the organic **E** detector, composed of BC-432m which is hydrogen-rich and therefore sensitive to neutrons. For a schematic of the RSH see Fig. 3.1. The surrounding Anti-Coincidence (AC), consisting of the plastic scintillators **F1**, **F2** and the silicon detector **C**, prevents a neutral particle measurement from being polluted by charged particles. In addition to the background radiation from the RTG, which interferes with the measurement, there are a number of factors which complicate the readout of the detectors. Although **D** is mainly sensitive to gammas, it is also somewhat sensitive to neutrons. The same holds true for **E** and gammas. In contrast to stopping charged particles, which deposit their complete energy, neutral particles interact only indirectly with matter and deposit their energy through secondary particles.

The main interactions of gamma particles in matter are the photoelectric effect, Compton scattering and pair production, which all generate secondary electrons which will then deposit their energy in the scintillator. For Compton scattering only a part of the gamma's energy is transferred to the electron (e.g. see Fig. 3.1). The same is true for electron-photon showers, resulting from pair production, where electrons and gammas may escape the detector, thus only a part of the incident gamma's energy is deposited in the detector. The main interaction for neutrons in the **E** detector is given by elastic scattering. In addition, inelastic scattering, radiative neutron capture processes and other nuclear reactions may occur [Leo, 1994]. Through elastic scattering in the hydrogen-rich **E** detector, recoil protons are produced. The probability-density distribution of their energies is flat up to the energy of the incoming neutron (see e.g. Kortmann [2010]). In addition to incomplete energy deposit of neutral particles, the light collection of the photodiodes depends on the location of the energy deposits in the scintillator [Kortmann, 2010]. Those effects lead to ambiguous measurements, which do not reflect the incoming particle spectra. This creates the necessity to develop an inversion procedure which calculates the real spectrum of the incoming neutron and gamma particles from the measured energy deposits.

Let us consider a spectrum of incoming neutron or gamma particles $\vec{f}$, which is binned

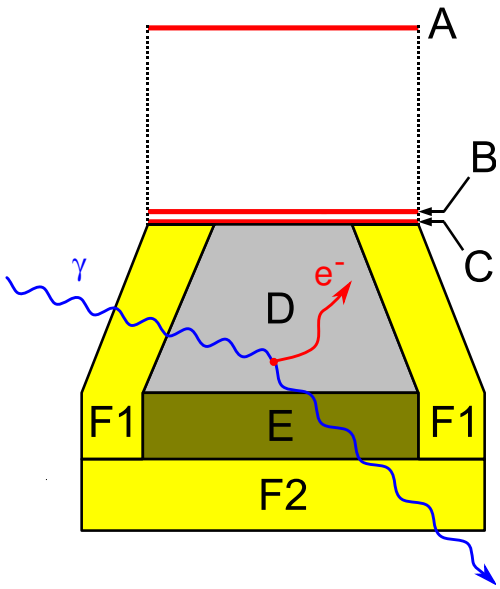

Figure 3.1: Schematic side view of the RAD instrument with the example of a detected gamma ray. The gamma ray enters the instrument through the AC **F1**, is Compton scattered in the **D** detector and leaves the instrument through the **E** detector and the AC **F2**. The gamma itself deposits no energy, but some of its energy is transferred to an electron via Compton scattering, which than deposits its energy in **D**.

in the energy bins $\vec{E} = (E_1, ..., E_n)$. A mapping between a spectrum of incoming particles and the measured detector response $\vec{z}$ can be performed by

$$\vec{z} = \mathbf{A} \cdot \vec{f}, \tag{3.1}$$

where the matrix **A** describes the DRF and defines the probability that a particle with a certain energy is detected in an energy bin. In a perfect detector the light output of the scintillator would be directly proportional to the penetrating particle's energy, thus $A_{i,j}$ would be 0 for $i \neq j$ and $A_{i,i}$ would be determined by the geometric factor for the energy E_i. In a more realistic scenario the matrix would be "smeared out" by Gaussian noise, which would result in an **A** where not only the diagonal, but also the neighboring, elements are $\neq 0$.

In addition to a modification through noise, the shape of the DRF for MSL/RAD must account for the effects of incomplete energy deposits which occur, e.g., for Compton scattering or elastic neutron scattering. Schematic DRFs, which give an example for a measurement with a charged-particle detector and with a neutron detector, are shown in Fig. 3.2. While the noisy detector matrix has a simple, smeared out, diagonal shape, the schematic neutron detector matrix has a triangular shape, which means that the deposited energy for an incoming neutron is randomly distributed between zero and its initial energy. Because the **D** detector is also somewhat sensitive to neutrons and the **E** detector is

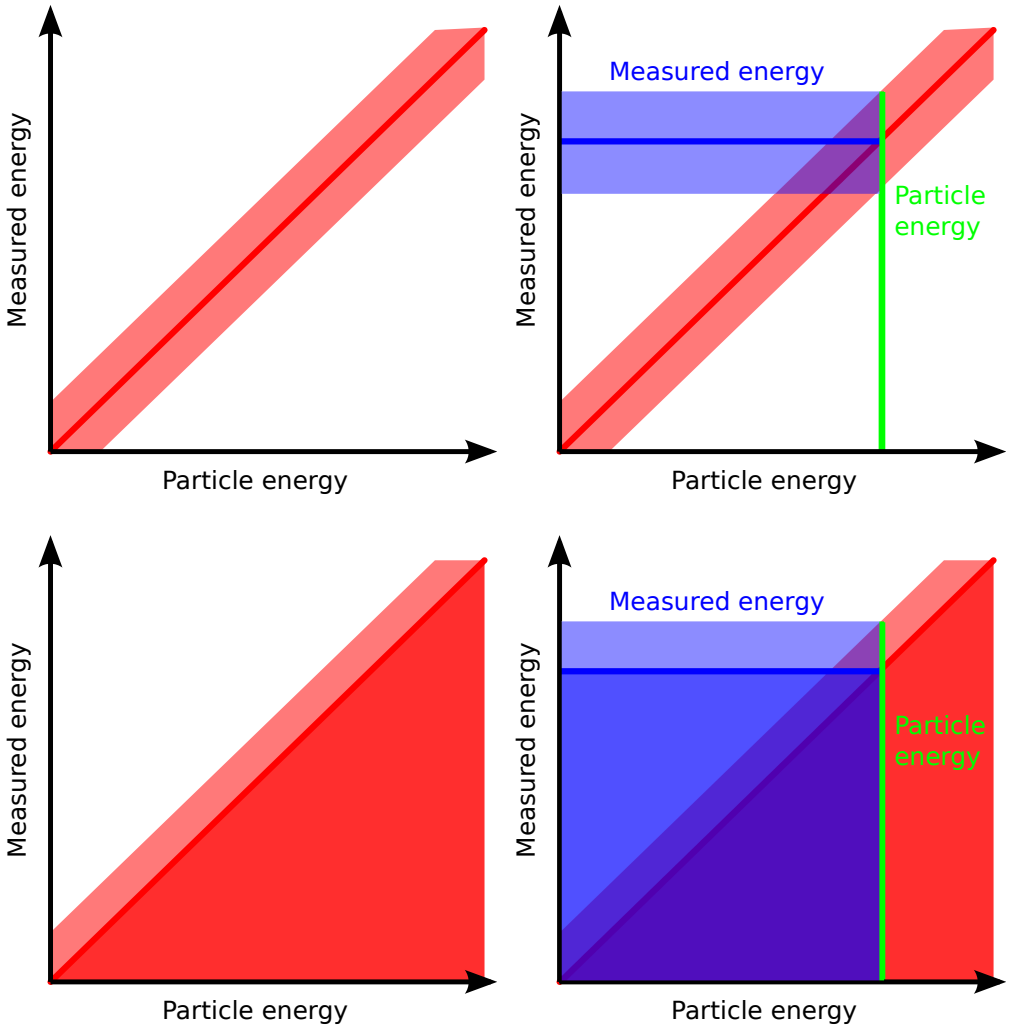

Figure 3.2: Top: Schematic DRF for a noisy charged-particle detector where detected energy is directly proportional to the particle's energy. Without noise the DRF would be strictly diagonal (dark red), the noise smears out the diagonal line (light red). Bottom: Schematic DRF for a noisy detector similar to **E** for neutrons. An incident neutron creates a secondary proton with a flat energy distribution ranging from zero up to the incident neutron's energy, therefore the triangular shape of the matrix. The additional noise smears out this shape. An example of an incident particle (green) is given on the right side. The probability with which the particle is detected in a measurement bin is shown in blue.

somewhat sensitive to gammas, the measurement in **D** and **E** can not be treated separately. The combined gamma and neutron measurements in **D** and **E** can be described by

$$\vec{z_D} = \mathbf{A}_{D,\gamma} \cdot \vec{f_\gamma} + \mathbf{A}_{D,N} \cdot \vec{f_N} \tag{3.2}$$

$$\vec{z_E} = \mathbf{A}_{E,\gamma} \cdot \vec{f_\gamma} + \mathbf{A}_{E,N} \cdot \vec{f_N}, \tag{3.3}$$

where $\mathbf{A}_{D/E,\gamma/N}$ is the DRF for gammas/neutrons in **D**/**E**, $\vec{f}_{\gamma/N}$ is the spectrum of incoming gammas/neutrons, and $\vec{z}_{D/E}$ is the measurement in **D**/**E**. Let us define the stacked vector $\vec{f} := \vec{f_\gamma} \oplus \vec{f_N} = (f_{1\gamma}, ..., f_{n\gamma}, f_{1N}, ..., f_{nN})$ and $\vec{z} := \vec{z}_D \oplus \vec{z}_E$. The mapping between incoming particles and measurement can then be performed by Eq. 3.1, where **A** consists of the gamma and neutron matrices for the **D** and **E** detector

$$\mathbf{A} := \begin{pmatrix} \mathbf{A}_{D,\gamma} & \mathbf{A}_{D,N} \\ \mathbf{A}_{E,\gamma} & \mathbf{A}_{E,N} \end{pmatrix},$$

as shown in Fig. 3.3 for a highly simplified general example of a gamma/neutron detector. Although this linear system of equations looks simple, it is often singular and is in general an ill posed problem. To obtain the spectrum of incoming particles $\vec{f}$ from the measurement $\vec{z}$, two different Maximum Likelihood (ML) methods are investigated: a constrained Non-Negative Least Squares (NNLS) and a ML estimator with underlying Poissonian statistics.

3.1 Inversion techniques

To obtain the incoming-particle spectra $\vec{f}$ from a measurement vector $\vec{z}$, Eq. 3.1 needs to be solved for $\vec{f}$. Since **A** is often singular, the inverse matrix $\mathbf{A}^{-1}$ often can not be obtained. However, there exist several methods to bypass this problem, e.g. obtaining a pseudo-inverse matrix via singular value decomposition, which has been done by Böhm *et al.* [2007]. Although this approach would enable us to mathematically solve Eq. 3.1, a simple matrix-times-vector multiplication does not consider the stochastic nature of the measurement process.

Consider a single measured particle, which creates an event in one measurement channel of the detector. Eq. 3.1 would suggest that even one measured particle is distributed over a wide range of channels. For a very large number of measured particles, the stochastic nature of a measurement becomes negligible and Eq. 3.1 can be solved simply by $\vec{f} = \mathbf{A}^{-1} \cdot \vec{z}$. For low count rates this effect will produce unphysical results, such as negative count rates, even if **A** is not singular. Further problems, even for large count rates, arise from inaccuracies in **A**. The DRF will always be known only with limited accuracy. This means we need to solve Eq. 3.1, but all we know about the exact system

$$\vec{z} = \tilde{\mathbf{A}} \cdot \vec{f},$$

is that $||\tilde{\mathbf{A}} - \mathbf{A}|| < h$, where $h > 0$. To obtain $\vec{f}$, with the constraints of non-negative intensities ($f_i > 0$), Eq. 3.1 can be replaced by a constrained optimization problem, i.e. fitting a set of parameters to maximize the likelihood that a spectrum $\vec{f}$ creates the measurement $\vec{z}$.

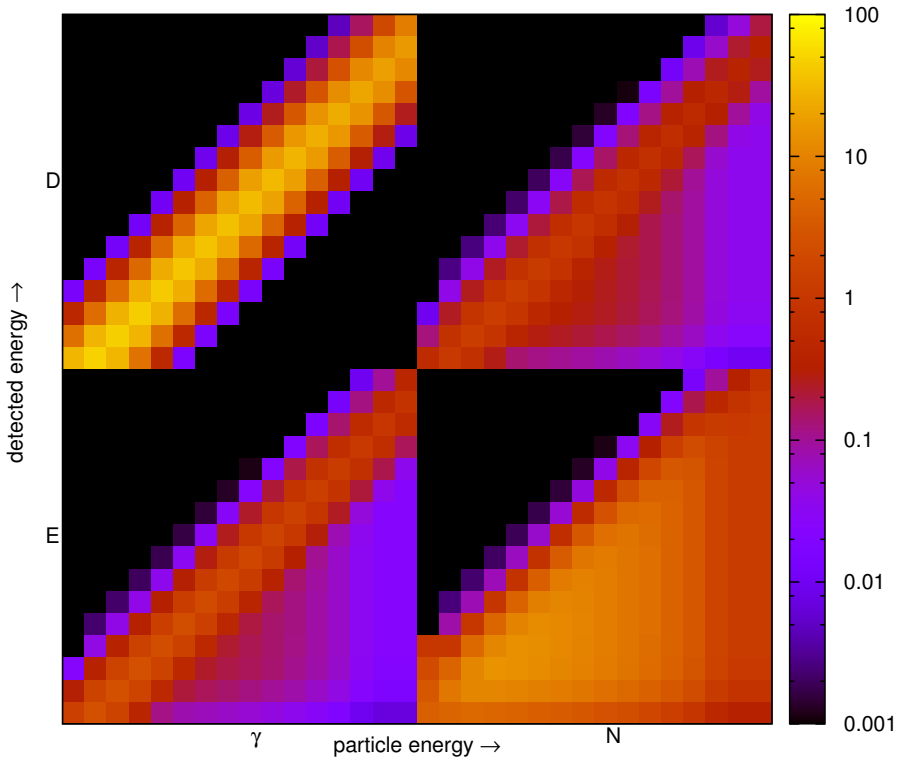

Figure 3.3: A highly simplified DRF for gammas (left) and neutrons (right) for detector **D** (top) and detector **E** (bottom). Each submatrix contains the energy range 1 - 200 MeV with 16 logarithmically spaced bins.

3.1.1 Maximum-likelihood estimations

For an observed data set and a statistical model, which describes this data set, a ML estimation selects the model parameters, which give the observed data the highest probability. Consider a model $z(x, \vec{a})$, which depends on the parameters $\vec{a} = (a_1,\ a_2,\ ...)$, and a corresponding set of independent measurements $(x_i,\ z_i)$ with normal distributed errors in z. The likelihood that the parameters $\vec{a}$ result in the measurements $(x_i,\ z_i)$ is given by

$$L(\vec{a}) = \prod_i p(z_i - z(x_i, \vec{a})), \tag{3.4}$$

where $p(z)$ is a normal distribution with zero mean and standard deviation σ. Usually, instead of maximizing Eq. 3.4, its negative logarithm is minimized.

$$l(\vec{a}) = \sum_i \frac{(z_i - z(x_i, \vec{a}))^2}{\sigma^2} \tag{3.5}$$

Obviously, Eq. 3.5 becomes minimal if the parameters $\vec{a}$ are set in a way that the squared difference between measurements z_i and corresponding expectation values $z(x_i, \vec{a})$ be-

come minimal. E.g. a model describing a linear dependence between z and x is given by $z(x_i, a_0, a_1) = a_0 + a_1 x_i$. As long as the model is linear, it can be fitted relatively easy by solving the partial derivatives of the parameters a_j

$$0 = \sum_j \frac{\partial}{\partial_{a_j}} l(\vec{a}).$$

If a model is nonlinear, a method which can deal with a set of nonlinear equations must be used, for instance the algorithm by Levenberg and Marquardt [Press, 2007].

3.1.2 Non-negative least squares

To obtain the input spectrum from measurements in **D** and **E**, Eq. 3.1 can be replaced by a ML estimation with underlying Gaussian statistics. The measurements are given by $\vec{z} = \vec{z}_D \oplus \vec{z}_E$, the free parameters are given by $\vec{f} = \vec{f}_\gamma \oplus \vec{f}_N$, and the model for the expectation value is given by $\vec{z}(\vec{f}) = \mathbf{A} \cdot \vec{f}$. The spectrum, which is the most likely to result in the observed measurement, is then given by

$$\min \sum_i \left(\sum_j a_{ij} f_j - z_i \right)^2 . \tag{3.6}$$

Considering that negative intensities are unphysical results, Eq. 3.6 can be formulated as a non-negative least squares (NNLS) problem,

$$\min \sum_i \left(\sum_j a_{ij} f_j - z_i \right)^2 , \quad \text{with } f_i \geq 0. \tag{3.7}$$

NNLS methods have been considered in Bjõrck [1996]; Lawson and Hanson [1987]. As is true with other methods, the least-squares problem always has a solution, but it is non-unique if the rank of the matrix **A** is less than its dimension n.

A similar method is given by the minimum chi-square method, which is not considered in this work because it behaves similar as the NNLS method. A description and evaluation of the chi-square method can be found in Köhler *et al.* [2011].

3.1.3 Maximum-likelihood estimation with underlying Poissonian statistics

In the examples above, the ML estimation is used with underlying Gaussian statistics. However, often the measurement errors are not normal distributed, but follow another distribution, e.g. the Lorentz or the Poisson distribution. Fig. 3.4 shows an example of a counting experiment, which is based on Poissonian statistic. For low count rates the measured values obviously do not agree with Gaussian statistics. The probability density

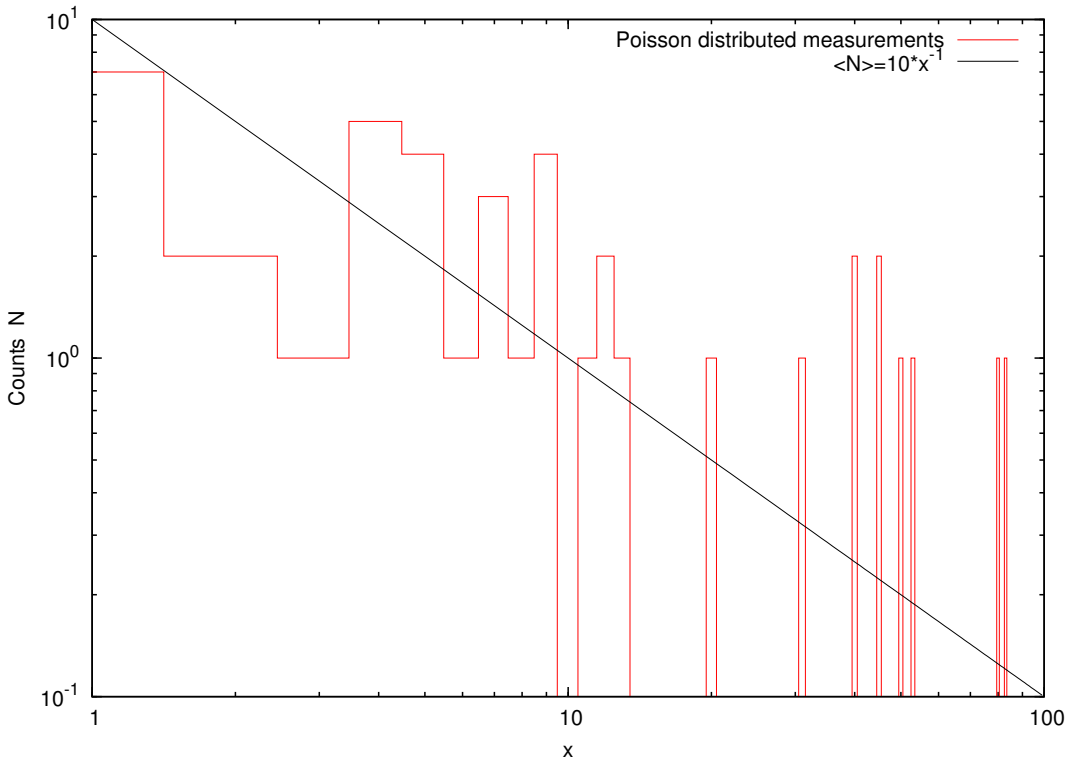

Figure 3.4: Example for an arbitrary counting experiment with an underlying power law $\langle N \rangle(x) = 10 \cdot x^{-1}$. For large count rates ($x < 10$) the counts are more or less Gaussian distributed around the expectation value $\langle N \rangle(x)$. But for low count rates ($x > 10$) the counts are either 0, 1, or in few cases 2, 3 ... , which is obviously not the result of an underlying Gaussian process. A comparison of the probability density function for large and low expectation values is shown in Fig. 3.5.

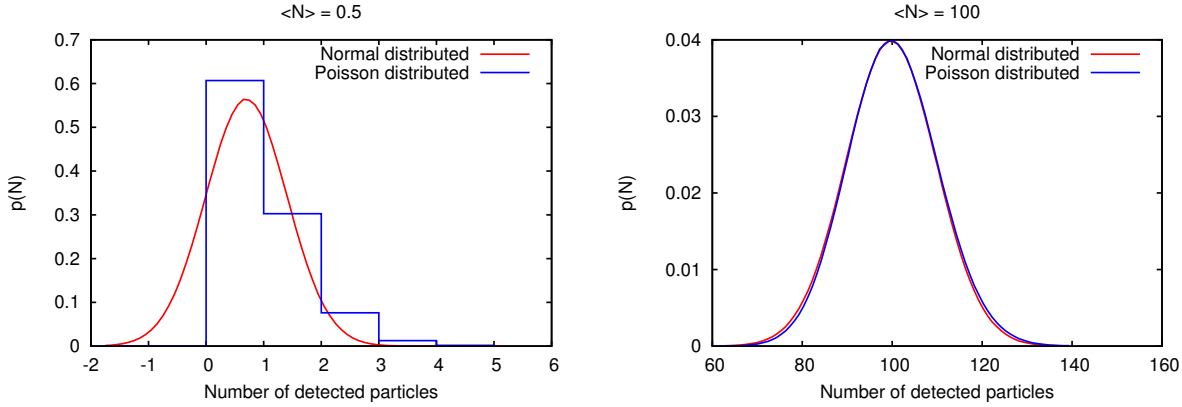

Figure 3.5: Gaussian and Poissonian probability density functions for low ($\langle N \rangle = 0.5$) and large ($\langle N \rangle = 100$) expectation values. While for $\langle N \rangle = 100$ the resulting probability densities are nearly the same, for $\langle N \rangle = 0.5$ the Gaussian function exhibits unphysical (negative) values, which do not agree with counting experiments.

functions for Poisson and Gaussian statistics are compared in Fig. 3.5. While the counting experiment produces only discrete positive integer values, for low count rates mostly 0 and 1, the Gaussian distribution falsely assumes non-integer and even negative values. In such situations, the assumption of Gaussian statistics can give far too much or too few weight to bins with 0 or 1 counts, and only the assumption of Poissonian statistics will give correct results in a ML estimation.

Since the detector measurement is a counting experiment, it is Poisson distributed. ML estimations with underlying Poissonian statistics have been considered by Hauschild and Jentschel [2001] and Hannam and Thompson [1999].
According to the Poisson distribution, the probability that a measurement, with an expectation value of $\lambda_i = (\mathbf{A} \cdot \vec{f})_i$, results in z_i counts is given by

$$p(z_i, \lambda_i) = \frac{\lambda_i^{z_i} e^{-\lambda_i}}{z_i!}.$$

The corresponding probability function $P(\vec{z}, \vec{f})$ that includes all parameters f_i and all independent measurements z_i, which maximizes the likelihood function $L(\vec{f})$, is given by

$$L(\vec{f}) = P(\vec{z}, \vec{f}) = \prod_{i=1}^{n} p(z_i, \lambda_i).$$

The corresponding negative log-likelihood function has the form [Vogel, 2002]

$$l(f) \quad = \quad \sum_i (\lambda_i - z_i \ln \lambda_i) + c, \tag{3.8}$$

where $c = \sum_i \ln(zi!)$. For the estimation of the spectrum $\vec{f}$ the constrained minimization problem

$$\min \sum_i \left(\lambda_i - z_i \ln \lambda_i\right), \text{ with } f_i \geq 0, \tag{3.9}$$

needs to be solved. Eq. 3.7 and Eq.3.9 are often called merit functions.

In this work, the optimization problems Eq. 3.7 and Eq. 3.9, with the constraints $f_i > 0 \, \forall f_i$, are solved using the SciPy [Jones *et al.*, 2001] implementation of the `L-BFGS-B` algorithm [Zhu *et al.*, 1997; Byrd *et al.*, 1994].

3.2 Comparing NNLS and Poissonian methods for gamma/neutron measurements

In this section NNLS and Poisson method are compared and explored for the general example of a gamma/neutron detector, which consists of two scintillators, one with a high

sensitivity for gammas and one with a high sensitivity for neutrons. Using artificially generated measurements, the described methods are analyzed for their performance and robustness. Via MC modeling, the matrix $\mathbf{A}$ is used to create measurements from given input spectra. The inversion methods will then use the artificial measurements to calculate the input spectra. Since the real input spectrum is always known, differences between calculated and real spectra can be analyzed in detail systematically. This way, both methods can be tested and compared without undergoing extensive measurement campaigns with the real instrument.

3.2.1 Producing artificial measurements

Since a particle measurement is a discrete process, $\vec{z}$ consists only of positive integers. A single particle which hits one of the detectors will be detected as exactly one event, not as a vector of detection probabilities, which is given by $\mathbf{A} \cdot f$. For generating a measurement for a hit of a single neutron with energy E_j, the j-th column of $\mathbf{A}$ defines the probabilities that the neutron is detected in a certain energy bin and detector. The probability distribution for detection of a particle with energy E_j in measurement bin i is given by

$$p_i = \frac{\mathbf{A}_{i,j}}{\sum_i \mathbf{A}_{i,j}}. \tag{3.10}$$

For an arbitrary input spectrum $C_0\, f(E_j)$ and a number of measured particles N, each particle is distributed via Eq. 3.10. The C_0, which results in N measured particles, can be determined by

$$N = \sum_i z_i = \sum_i \sum_j \mathbf{A}_{ij} C_0\, f\,(E_j) \tag{3.11}$$

$$\Longrightarrow C_0 = \frac{N}{\sum_i \sum_j \mathbf{A}_{ij} f(E_j)}. \tag{3.12}$$

To create a measurement vector of N events, for each energy E_j

$$N_j = \sum_i A_{ij} C_0\, f\,(E_j)$$

particles are distributed with probability density from Eq. 3.10. To account for the stochastic nature of the input spectrum, the number of particles per energy bin N_j are Poisson distributed, with an expectation value of $\langle N_j \rangle = \sum_i A_{ij} C_0 f(E_j)$. Hence, $N = \sum_j N_j$ is merely an expectation value. Using the above described MC method, one can generate artificial measurements for arbitrary input spectra.

3.2.2 Determining errors for an inverted spectrum

Because of the stochastic nature of the measurement process, the inverted spectra do not match the input spectra perfectly - even if $\mathbf{A}$ is perfectly known. Estimating the errors,

caused by the Poissonian nature of the measurement, can be done via bootstrap MC. For bootstrap MC a number of synthetic datasets, based on the actual data set, are generated. In each synthetic dataset, a set the size of the original data set is drawn with replacement from the original data [Press, 2007].

For a given measurement vector $\vec{z}$, the measurement is usually already processed in histogram form. In this case, the bootstrap MC is realized by Poisson-generated synthetic histograms $\vec{z}^{(j)}$, with expectation values given by the original histogram. I.e., each bin of a synthetic measurement $z_i^{(j)}$ is generated from a Poisson distribution with expectation value $\lambda = z_i$. The resulting spectrum and the corresponding errors are given by

$$\vec{f} = \left\langle \vec{f}^{(j)} \right\rangle_j \tag{3.13}$$

$$\Delta\vec{f} = \left\langle \left\langle \vec{f}^{(j)} \right\rangle \right\rangle_j, \tag{3.14}$$

where $\vec{f}^{(j)}$ is the inverted spectrum for synthetic measurement $\vec{z}^{(j)}$.

3.2.3 Comparing inversion results

Using artificially generated data, the performance of the Poissonian ML method can be compared to the NNLS method. As an input spectrum a simple powerlaw

$$f_{\gamma/N}(E_j) = E_j^{\mathrm{s}},$$

with $s = -2$, is chosen for the gamma and neutron spectrum. Fitting a functional behavior, the quality of the different methods can be compared directly by comparing the few fitted parameters, e.g. spectral index and intensity. In this case, the number of free parameters is too large for a direct comparison. Therefore, the relative errors of the bootstrap MC obtained results (Eq. 3.13) are used to measure the fit quality. However, this method assumes that the bootstrap MC obtained errors are correct, and that the results do not systemically deviate from the input spectrum. Therefore, when using the mean relative error as a measure, one needs to verify that mean and errors of the inverted spectrum agree with the input spectrum.

3.2.4 Obtaining a first initial guess

One crucial factor for fitting algorithms is the initial guess. Since a good detector should have an $\mathbf{A}$ which is close to diagonal, the initial guess is determined with this assumption

$$f_{i,\mathrm{init}} = \frac{z_i}{\sum_j A_{ji}}. \tag{3.15}$$

For a detector with a geometric factor $A_{ij} = 0$ for $i \neq j$, Eq. 3.15 is the solution of $z = \mathbf{A} \cdot f$. The more the detector response deviates from a diagonal shape, the more inaccurate Eq. 3.15 becomes.

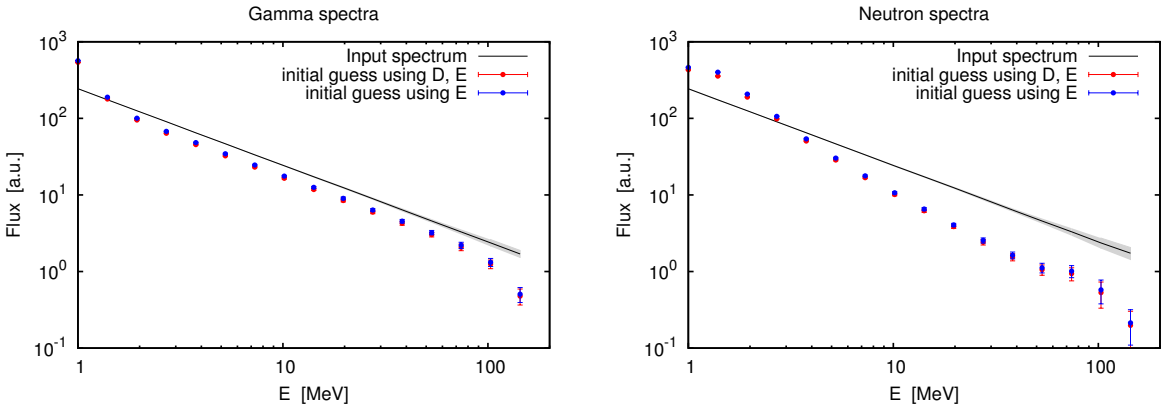

Figure 3.6: Input spectrum (black with standard deviation in gray) and initial guess for gammas (left) and neutrons (right) calculated for $N = 10^5$ events. The initial guess was calculated via Eq. 3.15, using the full matrix **A** (red) and using only the sub-matrices **D** for gammas and **E** for neutrons (blue). Since the **D** detector is mainly sensitive for gamma particles and the **E** detector is mainly sensitive for neutrons, the initial guess can be calculated using only the sub-matrices. However, this brings no clear improvement of the initial guess. While the initial guess for the gammas is astonishingly accurate, the initial guess for the neutrons shows an increased spectral index.

Fig. 3.6 shows the performance of the initial guess for a measurement with $N = 10^5$ events. While Eq. 3.15 is astonishingly accurate for $\vec{f}_\gamma$, $\vec{f}_N$ shows an increased slope of the power-spectral density. Considering the shape of the DRF, which has been used (Fig. 3.3), one finds that the gamma matrices are close to diagonal, while the neutron matrices deviate from this clear diagonal shape. This can be understood using an even more oversimplified $\mathbf{A}_{E,N}$ matrix, with A_{ij}=1 for $j \geq i$ and A_{ij}=0 for $j < i$. For an uniform input spectrum with $f_i = 1$, $\vec{f}_{\text{init}}$ is then given by

$$\vec{f}_{\text{init}} = (N,\ (N-1)/2,\ ...,\ (N-1)/(i+1), ...,\ 1/N)\,,$$

which shows the same effect as for the initial neutron guess in Fig. 3.6. Continuing this consideration for steeper spectra, one finds that this effect becomes less pronounced, which can also be observed for the correct $\mathbf{A}_{E,N}$.

Although Eq. 3.15 deviates from the neutron and gamma spectra, it still produces spectra with similar shape and similar intensities $I_{\gamma/N} = \sum_i f_{\gamma/N,i}$. The intensities of initial guess and input spectra were compared for various spectral shapes, relative intensities I_γ/I_N, and DRFs. Intensities of initial guess and input spectra deviated only by factors of $< 20\%$. Even though Eq. 3.15 provides only a crude guess, it can still provide useful information about intensities of neutron I_N and gamma I_γ spectra.

If some information about the shape of the spectra is available, useful initial guesses

can be obtained from

$$\vec{f}_{\text{init}} = I_\gamma \vec{f}_\gamma \oplus I_N \vec{f}_N, \tag{3.16}$$

where I_γ, I_N are obtained from Eq. 3.15 and $\vec{f}_\gamma$, $\vec{f}_N$ are the normalized shapes of the expected spectra. Trying different shapes, scaled with I_γ, I_N, yields much better initial guesses than using $\vec{f}_{\text{init}}$ from Eq. 3.15. Although this approach uses some knowledge of the underlying functional behavior, it does not restrict the solution to a given shape.

3.2.5 Inversion of artificial measurements

Using the initial guess from Eq. 3.15, the different inversion techniques are compared for $N = 200,\ 2000,\ 20000$ measured particles. Fig. 3.7 shows the mean input spectra and the mean inverted spectra for neutrons and gammas. The error bars show the errors, determined via bootstrap MC, of the inverted spectra. The variability of the input spectra, which is due to its stochastic generation, is shown in gray. The mean inverted spectra for the Poisson and the NNLS method both agree well with the input spectra, which shows that the methods do not create any systematic errors. The error bars, however, give information about the quality of the inverted spectra. Especially for the gammas, an increase of the relative errors, with decreasing number of particles per energy bin, is very clearly visible. While both methods show similar errors for low energies, where count rates are high, the Poisson method shows much smaller error bars for high energies, where count rates are low, i.e. the Poissonian nature of the measurement process becomes relevant. In Fig. 3.8 the relative errors $\langle \Delta \vec{f} \rangle / \langle \vec{f} \rangle$ are compared for different N. Astonishingly the Poisson method not only produces the better results for small count rates, it also shows slightly decreased (bootstrap MC) errors for high count rates, e.g. for $N = 2 \cdot 10^6$ in the 2 - 10 MeV energy range.

It should be noted that the Poisson method has a less steep merit function. For high count rates, the numerically calculated gradient can therefore be influenced by numerical noise. This problem can be circumvented by using larger step-widths for the calculation of the merit function's gradient. The approach is explained in detail in Sec. A.1.1.

While the Poisson method shows less variability in the resulting spectra, it appears to be more sensitive to the initial guess. To investigate this effect, both methods are compared for two different initial guesses, one given by Eq. 3.15 and one given by the input spectrum - the perfect guess. To compare both methods for a wide range of N, the quality of the resulting spectra $\vec{f}$ is determined via the mean relative error. In Fig. 3.9 it can clearly be seen that the relative error decreases with increasing number of particles. For all values of N the Poisson method performs better than the NNLS method.

Since the input spectrum $\vec{f}_{\text{input}}$ is the desired result, we can assume that it is the optimal initial guess. Comparing the two different initial guesses in Fig. 3.9, one finds that the NNLS method can not be improved by providing a more accurate initial guess. The Poisson method, however, shows minor improvements for small N and very clear improvements for $N > 5000$. Depending on the initial guess, the fitting algorithm finds different local minima

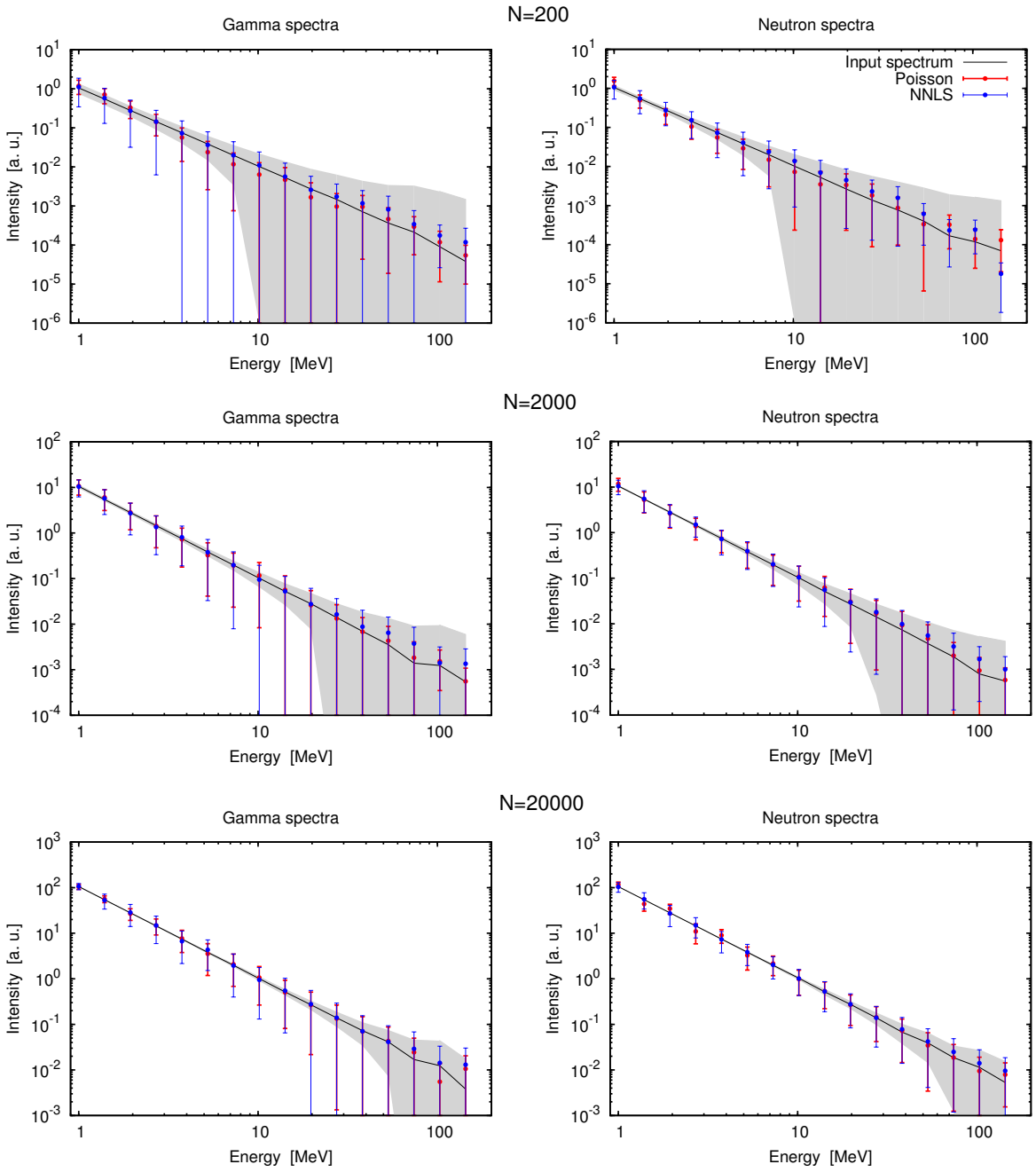

Figure 3.7: Mean input spectra (black with standard deviation in gray) and mean of 1000 inverted spectra, and the corresponding errors for artificial measurements with $N = 200,\ 2000,\ 20000$ measured particles. The standard deviation of the input spectrum is due to its stochastic generation.

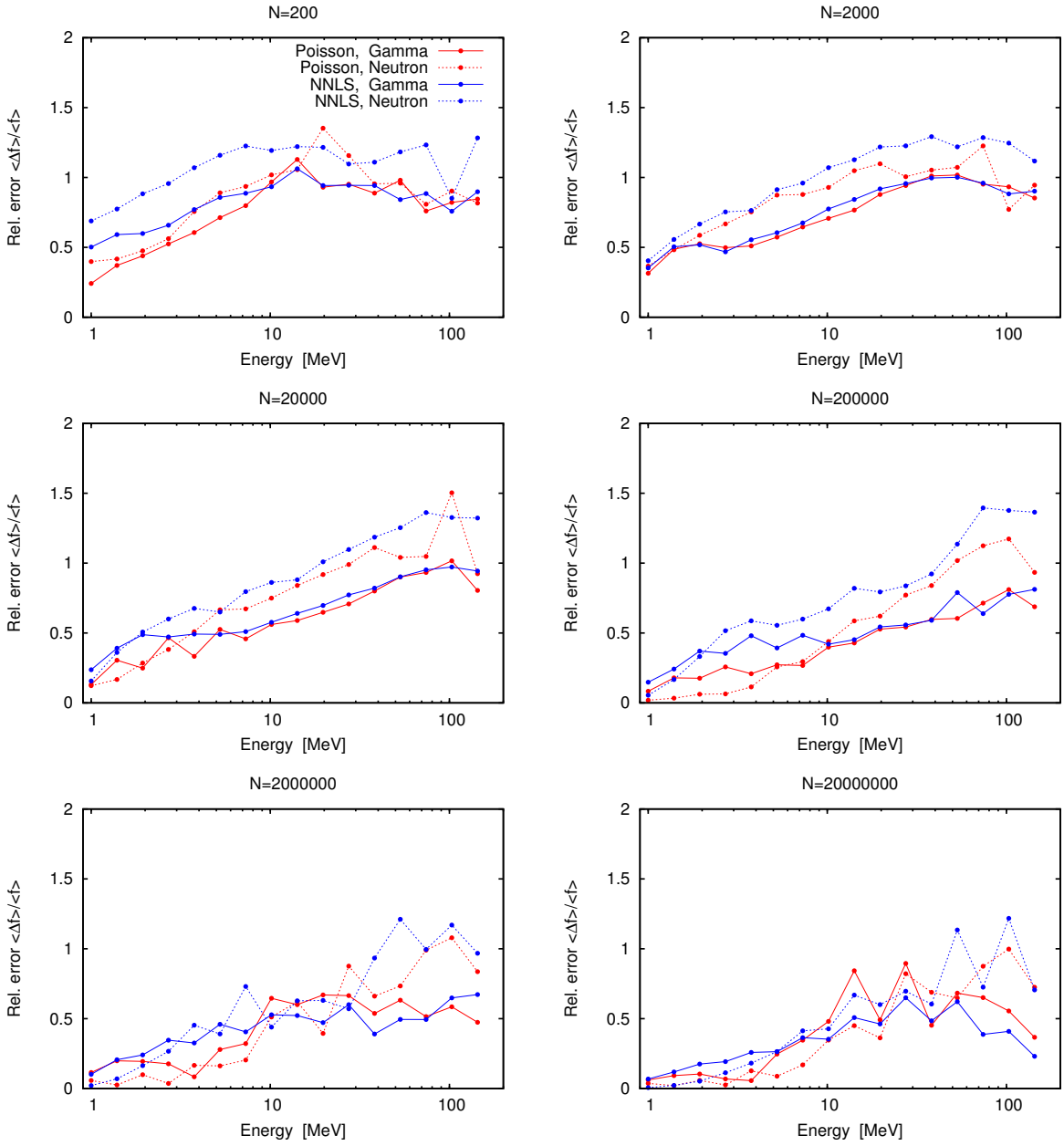

Figure 3.8: Relative errors $\langle \Delta \vec{f} \rangle / \langle \vec{f} \rangle$ for different N. For $N = 200$ the relative errors of the Poisson method are significantly smaller up to high energies, where the relative error is ≈ 1. For $N = 2000$ the relative errors are similar for low energies, but the Poisson method again produces better results for higher energies. For $N = 20000$ the Poisson method shows only slightly better results then the NNLS method.

of the merit function. Optimizing the initial guess to find a minimum as close as possible to the global minimum is, therefore, an important part of the fitting procedure.

3.2.6 Strategies finding the global minimum

As shown in Fig. 3.9, an improved initial guess improves the resulting spectra significantly. The choice of $\vec{f}_{\text{init}}$ is a crucial factor for finding not a local, but the global minimum of the merit function. For an inversion with a large number of free parameters, finding the global minimum via brute force variation of $\vec{f}_{\text{init}}$ is nearly impossible. There exist several methods to search the parameter space for a global minimum, e.g. the Gelfand algorithm [James, 1972] which determines new initial guesses from two previous calculated results. However, all those methods require fine tuning and adaption to the particular inversion problem. To understand the Poisson method and identify typical problems in finding the global minimum, a single measurement is investigated in detail. Fig. 3.10 shows input and inverted spectra for one single inversion. Comparing the result for $\vec{f}_{\text{init}} = \vec{f}_{\text{input}}$ to the input spectrum $\vec{f}_{\text{input}}$, one still finds a large deviation between $\vec{f}_{\text{input}}$ and the inverted spectrum $\vec{f}$. This means that $\vec{f}$ is more likely to produce the measurement than $\vec{f}_{\text{input}}$. This also becomes clear by comparing the artificial measurement $\vec{z}$ with $\mathbf{A} \cdot \vec{f}_{\text{input}}$ and $\mathbf{A} \cdot \vec{f}$, where $\mathbf{A} \cdot \vec{f}$ is much closer to $\vec{z}$ than $\mathbf{A} \cdot \vec{f}_{\text{input}}$. This is caused by the Poissonian nature of the measurement process. Without imposing any additional constraints on the fit, e.g. $\vec{f}$ is monotone decreasing, the resulting spectrum can not be improved very much.

However, the resulting spectrum from $f_{\text{init}} =$ Eq. 3.15 shows typical pronounced minima which are accompanied by overestimated neighboring points. E.g., the gamma spectrum at $E = 3$ MeV shows a minimum, while the neighboring values $E = 2$ MeV and $E = 4$ MeV are overestimated. The same minima and overestimations occur for most initial guesses and also for the NNLS method. This behavior can be replicated with a smeared out unit matrix. For a perfect unit matrix f_i would be given by z_i. In a smeared out matrix, an f_i can be zero - and the contributions to z_i are achieved by f_{i-1}, f_{i+1}. A highly simplified example can be constructed for

$$\begin{pmatrix} 4 \\ 4 \\ 4 \\ 4 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 2 & 2 \end{pmatrix} \cdot \vec{f}. \tag{3.17}$$

Both, $\vec{f} = (1, 1, 1, 1, 1)$ and $\vec{f} = (2, 0, 2, 0, 2)$ solve the equation above. In contrast to the solutions of the minimization problem, both $\vec{f}$ are exact solutions of the equation. Nevertheless, it demonstrates the over/underestimation effect which occurs in the minimization problem. This effect is know to be caused by a too small bin size. In fact, one can easily see that Eq. 3.17 would not display this effect for a larger bin size. However, this effect can also be caused by poor initial guesses and can be avoided by supplying more exact inital guesses, or by repeating the fit with a corrected initial guess based on a first solution. In

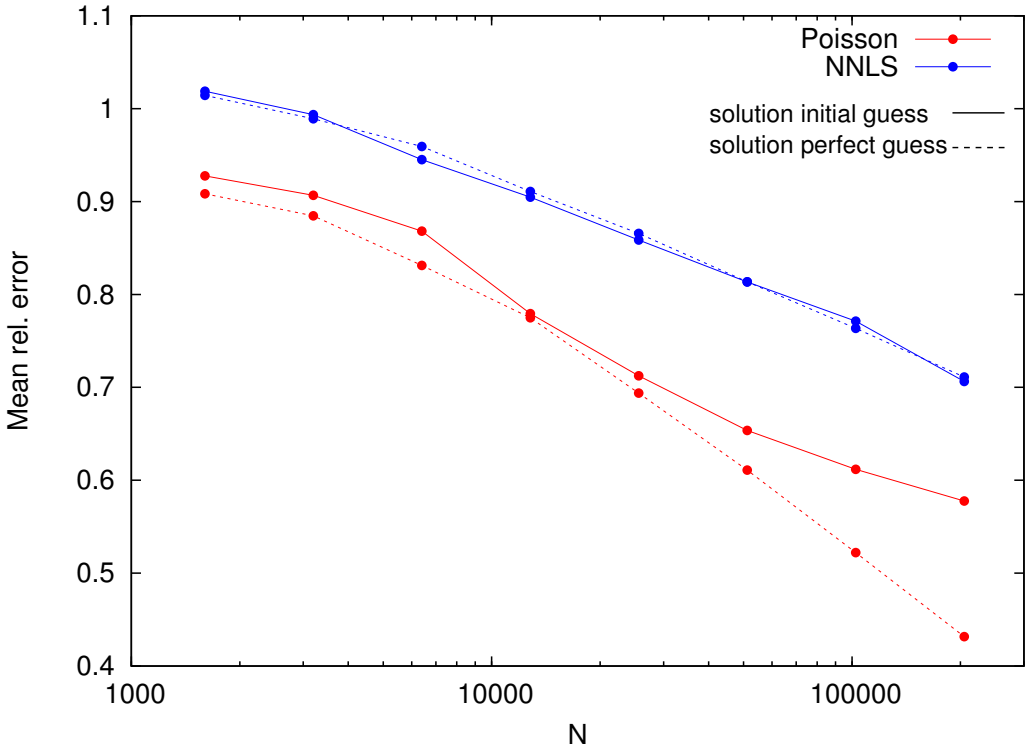

Figure 3.9: Mean relative error over number of measured events for the ML method with underlying Poisson statistics and for the NNLS method. For each method two initial guesses are compared, the initial guess from Eq. 3.15 and the perfect guess. While the NNLS method does not improve for an improved initial guess, the relative errors for Poisson method can be reduced by providing an improved initial guess.

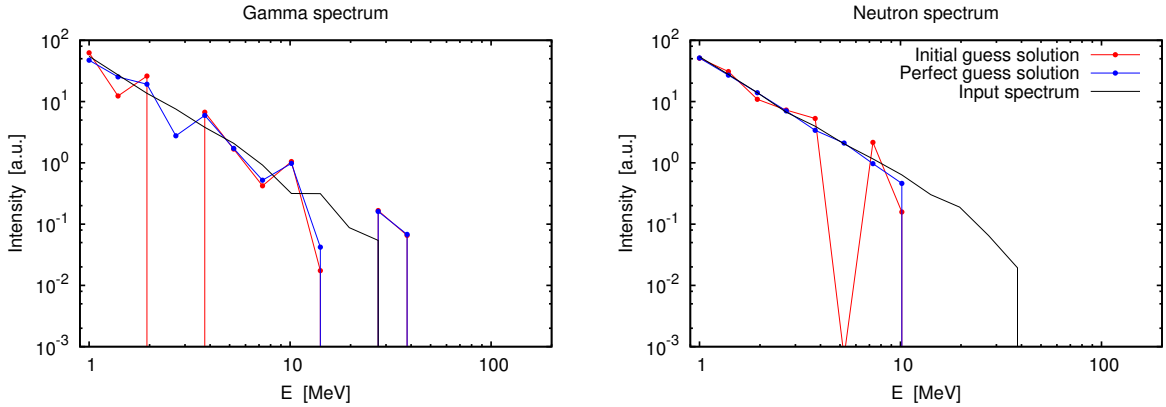

Figure 3.10: Example for an artificial measurement with $N = 10000$ events, the Poisson-inverted spectra for f_{init}=Eq. 3.15, and $f_{\text{init}} = f_{\text{input}}$, and f_{input} are shown. For $f_{\text{init}} =$ Eq. 3.15 the resulting spectra show a typical pattern, where an underestimated value is accompanied by two neighboring overestimated values. Because of the stochastic nature of the measurement, even the resulting spectrum for $f_{\text{init}} = f_{\text{input}}$ does not agree with f_{input} and shows a similar but weaker over/underestimation pattern.

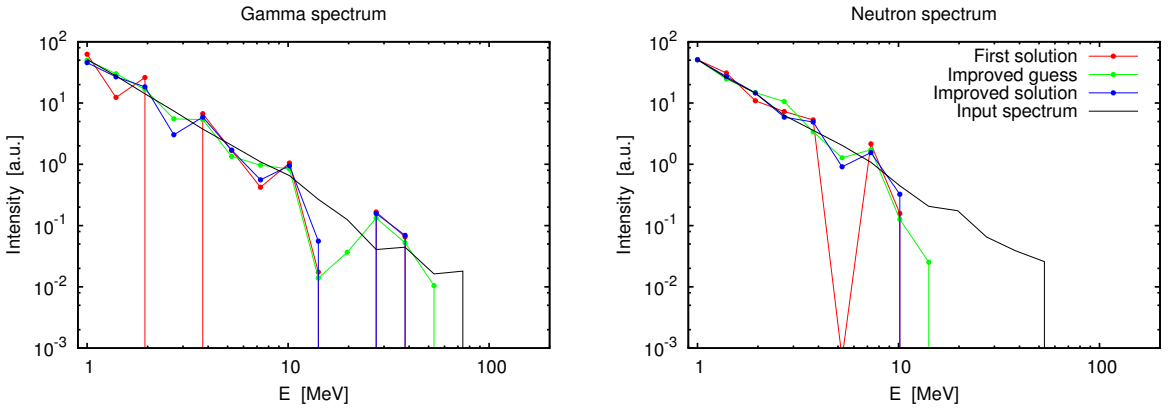

Figure 3.11: Example for an artificial measurement with $N = 10000$ events as shown in Fig. 3.10, in addition the corrected initial guess and the resulting spectra are shown. The improved initial guess reverses the over/underestimation of the first result, the resulting spectrum is much closer to the input spectrum.

Fig. 3.10 most of the over/underestimation pattern vanishes for $f_{\text{init}} = f_{\text{input}}$, which means that a large part of this pattern is caused by the poor initial guess.

Avoiding extreme over/underestimations

Considering this effect, an improved initial guess can be obtained by using the initial result, and smoothing the occurring pronounced minima. Let a pronounced minimum be a value f_i with $f_{i-1} > a \cdot f_i < f_{i+1}$, where $a > 1$ can be adjusted according to expected input spectra and quality of the results. Further, let $\mathbb{M}_a$ be the set of i for which the minimum condition is fulfilled. A simple correction of the over/underestimations can be achieved by beginning with $\vec{f}_{\text{init}} = \vec{f}$, where $\vec{f}$ is the solution for the first initial guess. For each $i \in \mathbb{M}_a$ the guess is adjusted to

$$f_{\text{init},i} = f_i + b \cdot (f_{i-1} + f_{i+1}) \tag{3.18}$$

$$f_{\text{init},i\pm1} = (1 - b) \cdot f_{i\pm1}, \tag{3.19}$$

where $0 < b < 1$. The quality of the resulting spectra has systematically been compared for $N = 10000$ events, in the parameter range $a \in [1, 10]$, $b \in [0, 1]$. Best results were obtained for $a = 1.2$ and $b = 0.2$. Fig. 3.11 shows the same artificial measurement as in Fig. 3.10, but also using the improved initial guess. The resulting spectra show less pronounced minima and are much closer to the input spectra. This improvement can also be seen in the corresponding value of the merit function. From several improved initial guesses with different a, b values the one with the best (smallest) merit value can therefore be chosen as the best solution.

In the bootstrap MC error-determination method, each synthetic measurement is inverted several times with improved initial guesses, and the best solution is chosen. As

before, from the best solutions of all synthetic measurements the mean and error is calculated. As already visible in Fig. 3.9, the NNLS method can not be improved by using $\vec{f}_{\text{init}} = \vec{f}_{\text{input}}$, consequently, applying improved guesses yields no improvement either.

On the one hand, a general inversion algorithm should make as few assumptions as possible about the shape of the spectrum, on the other hand, some assumptions need to be made to improve the initial guess. The method demonstrated above implicitly assumes a smooth continuous spectrum. For an input spectrum with peaks or gaps, the method would falsely try to correct those peaks or gaps. But since the values of the merit function, for the initial solution and the "improved" solution, can be used to select the best solution, this approach can only improve the result.

Randomized initial guesses

A direct approach, which yields good results, is repeating the fit for several randomized initial guesses. For each initial guess the fit yields a local minimum of the merit function, of which the smallest minimum, i.e. the best solution, is selected. A randomized initial guess is obtained from a previous solution $\vec{f}$ by

$$f_{\text{init, i}} \;=\; f_i \cdot (1 + c\,\xi), \tag{3.20}$$

where $0 < c < 1$ and ξ is a uniform distributed random value in $[-1,\ 1]$.

Using a large set of reasonable initial guesses

Without a basic knowledge of the input spectrum an initial guess can not be easily estimated from the measurement vector $\vec{z}$ (e.g. via Eq. 3.15). In most cases a basic knowledge of the input spectrum is available, and since the intensity of the input spectra can be estimated fairly well, several initial guesses can be constructed from a set of possible input spectra. Promising results were produced by using Eq. 3.15 and a set of possible input spectra (normalized to the correct intensity), and using those as a source for randomized initial values via Eq. 3.20. Good intensities for an initial guess can be obtained iteratively, with initial intensities I_N, I_γ from Eq. 3.15, and then adjusted with the intensities of the following inversions

$$\begin{aligned}
\vec{f}_{\gamma,\,\text{init},\,1} &= \vec{f}_{\gamma,\,\text{init}} \cdot \frac{|\vec{f}_{\gamma,\,\text{init,\,guess}}|}{|\vec{f}_{\gamma,\,\text{init}}|} \\
\vec{f}_{N,\,\text{init},\,1} &= \vec{f}_{N,\,\text{init}} \cdot \frac{|\vec{f}_{N,\,\text{init,\,guess}}|}{|\vec{f}_{N,\,\text{init}}|} \\
&\vdots \\
\vec{f}_{\gamma,\,\text{init},\,i} &= \vec{f}_{\gamma,\,\text{init}} \cdot \frac{|\vec{f}_{\gamma,\,i-1}|}{|\vec{f}_{\gamma,\,\text{init}}|} \\
\vec{f}_{N,\,\text{init},\,i} &= \vec{f}_{N,\,\text{init}} \cdot \frac{|\vec{f}_{N,\,i-1}|}{|\vec{f}_{N,\,\text{init}}|},
\end{aligned} \tag{3.21}$$

where $\vec{f}_{\gamma/N,\,\mathrm{init},\,i}$ is the i-th initial guess, with the resulting spectrum $\vec{f}_{\gamma/N,\,i}$, and $\vec{f}_{\gamma/N,\,\mathrm{init,\,guess}}$ is the initial guess from Eq. 3.15. Fitting a measured -2 powerlaw with several guess shapes (powerlaws with spectral indices -1.6, -1.8 -2.2, -2.4), adjusting the intensity as described above, and selecting the results with the highest likelihood value, yields results with the same quality as $\vec{f}_{\mathrm{init}} = \vec{f}_{\mathrm{input}}$.

Finding the global minimum for mathematical measurements

To obtain an estimation of the method's ability to find a global minimum, a mathematical measurement $\vec{z} = \mathbf{A} \cdot \vec{f}$ is generated, for which both methods are compared without the large influence of the Poissonian measurement process. The intensity of $\vec{f}$ is chosen to fulfill $z_i > 10$. This way, the probability density functions for normal and Poisson distributed events are similar. For $z_i < 10$ a Poisonian probability distribution is very inaccurate for describing a behavior which is not based on counting statistics, like this mathematical measurement. Fig. 3.12 shows the inversion of a mathematical measurement, for both Poisson and NNLS method. The different curves show the results for various initial guesses. For both methods, the different solutions are obtained for the initial guess obtained from Eq. 3.15. Based on the intensity of this first solution, two powerlaws with spectral index -3 are used as an initial guess. The following initial guess is given by two powerlaws with spectral index -2.1, which is very close to the real value of -2. Using this solution, several via Eq. 3.20 randomized solutions are then used as initial guesses to improve the previous result.

For the initial guess from Eq. 3.15, both methods find a solution which shows some energy bins with clear deviations from the original power law shape. Using the poor initial guess (spectral index -3) the solutions slightly improve, but still show deviations from the input spectrum. The more accurate guess of powerlaws with spectral index of -2.1 results in a clear improvement of the solutions, with only minor deviations from the input spectrum. By randomizing this solution and using it as an initial guess (Eq. 3.20), the previous solution can be improved further. The results of the NNLS method improve further in this step, small deviations vanish and at least in the figure there is no difference visible between input spectra and solution. However, the results of the Poisson method becomes worse, although the value of their merit function decreases. This may seem paradox, but considering that the Poisson method wrongly assumes an underlying Poissonian process, it is reasonable that the most likely solution is not the correct solution. This demonstrates again the importance of the applied statistical model. Even without any measurement errors or counting statistics, the choice of initial guesses is important for finding the best solution, i.e. a minimum close to the global minimum of the merit function.

3.2.7 Errors in A

The $\mathbf{A}$ used for the inversion is only an approximation to the real unknown $\tilde{\mathbf{A}}$. It is obtained from `GEANT4` simulations [Agostinelli, 2003] and calibration, and should therefore be of a similar shape, with only small deviations. To investigate the effect of those inaccuracies,

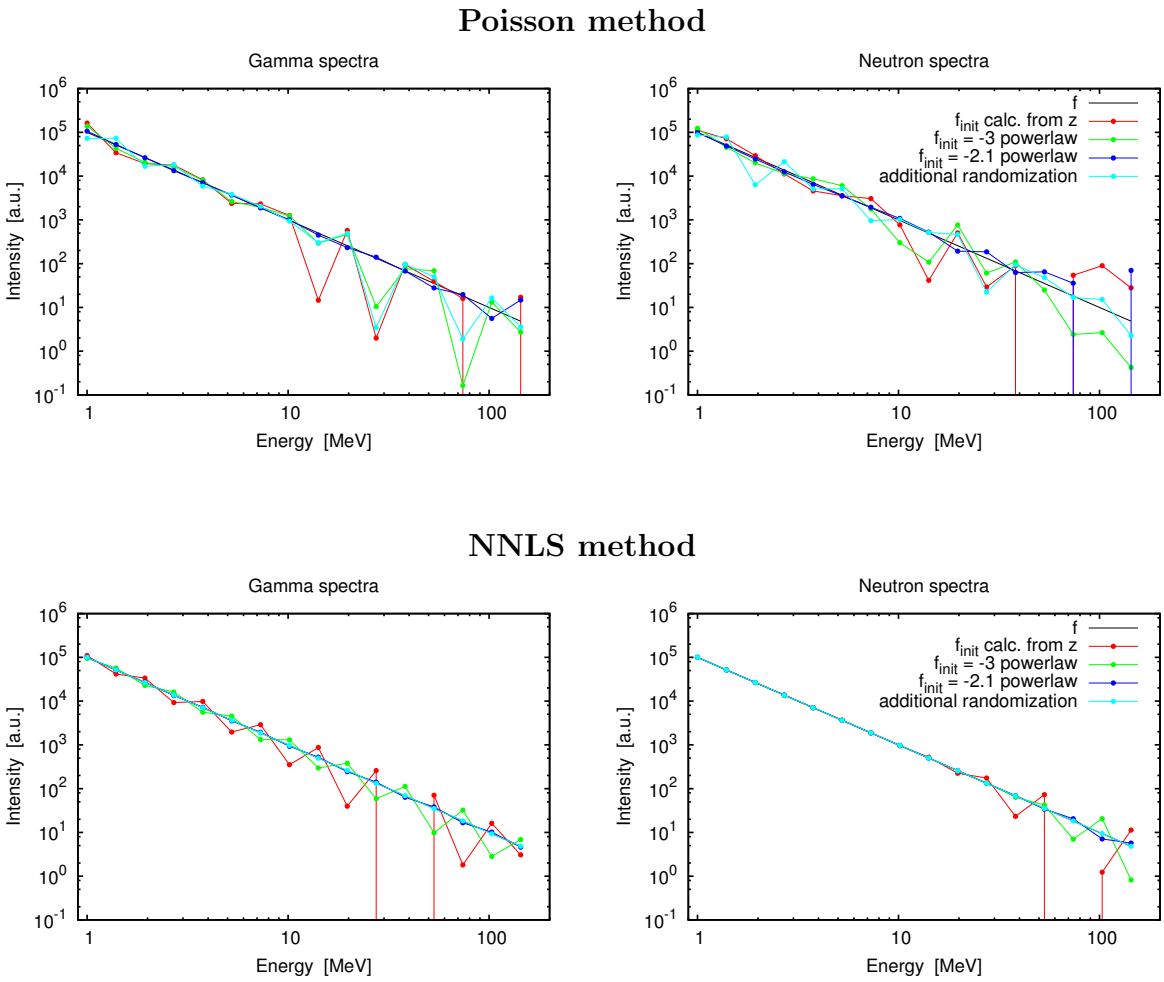

Figure 3.12: Inversions for different initial guesses for a mathematical measurement ($\vec{z} = \mathbf{A} \cdot \vec{f}$) of gamma/neutron spectra with powerlaw shapes with a spectral indices of -2 (black curve). For Poisson (top) and NNLS (bottom) method the different solutions are obtained for the initial guesses given by Eq. 3.15 (red). Based on the intensity of this first solution a powerlaw with spectral indices -3, -3 is used as an initial guess (green). The following initial guess is given by powerlaws with spectral index -2.1 (blue), which is very close to the real shape -2. Using this solution, several via Eq. 3.20 randomized initial guesses are then used as starting points (light blue).

the DRF from Fig. 3.3 is used to generate the artificial measurement, and an inversion is done with a flawed matrix. The flawed matrix $\mathbf{A}$ is generated by

$$A_{ij} = \tilde{A}_{ij} \cdot (1 + a\xi), \tag{3.22}$$

where ξ is a Gaussian distributed random variable with zero mean and unit standard deviation and a is a scaling factor for the error. Since A_{ij} is obviously positive, negative values are set to 0. This means the real unknown geometric factors are given by $\tilde{\mathbf{A}}$, the known geometric factors, determined by simulations and experiments, are given by $\mathbf{A}$. This way the inversion techniques can be compared for different error factors a.

In this case the mean relative errors can not be used to measure the method's robustness to errors in $\tilde{\mathbf{A}}$, because the bootstrap MC generated errors for one measurement are all based on the same flawed $\mathbf{A}$. The errors do not consider systematic errors, which may occur as a result of a wrong DRF, only errors due to the stochastic nature of the measurement. Instead the deviation

$$\mathrm{Dev}(\vec{f}) \quad = \quad \frac{1}{n} \sum_{i=1}^{n} \frac{(f_i - f_{\mathrm{input},i})^2}{f_{\mathrm{input},i}} \tag{3.23}$$

of the inverted spectrum from the input spectrum is calculated. Fig. 3.13 shows the deviation of both methods. The Poisson method shows a remarkable robustness to errors in $\tilde{\mathbf{A}}$. The NNLS method, however, shows a much larger increase of deviations for increasing a. Even for an error scaling factor of 0.5, the deviation from the input spectrum for the Poisson method is below 10, while the NNLS method produces deviations above 20 for $a \geq 0.3$.

One should be aware that Eq. 3.22 only produces statistical errors or inaccuracies in the DRF, it does not include systematic errors. A systematic error may very well lead to unphysical results. However, Fig. 3.13 shows that the Poisson method is much better in handling small deviations in the DRF than the NNLS method.

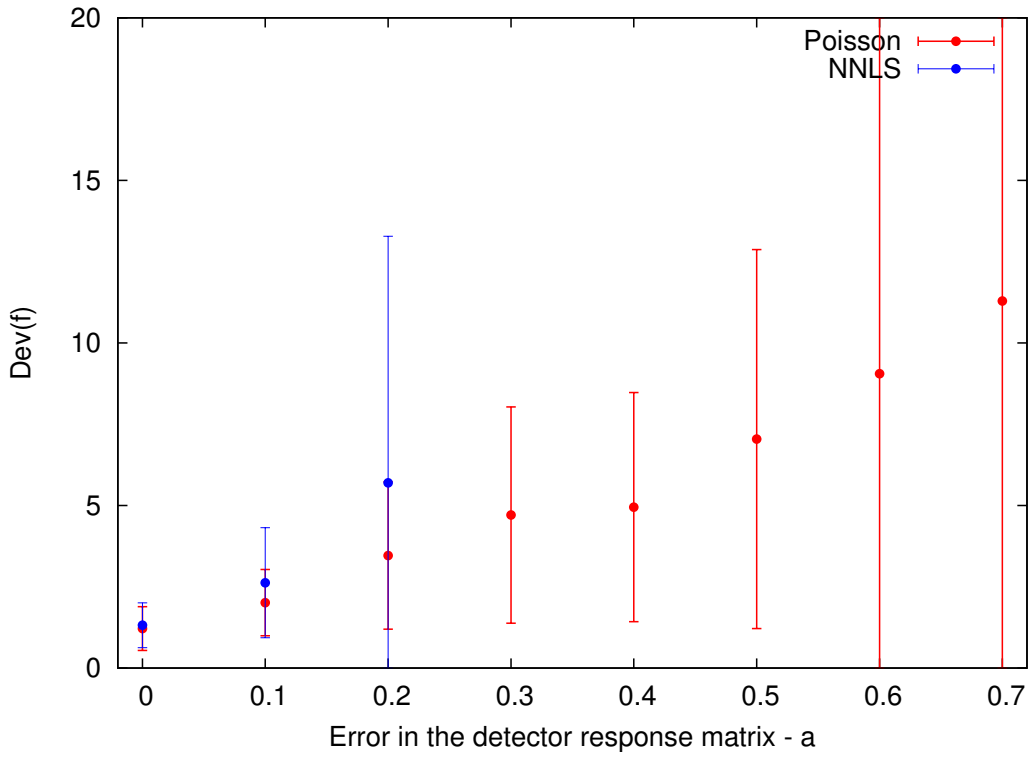

Figure 3.13: Performance of the different inversion techniques for various error scaling factors a for MC generated measurements with $N = 10000$ events. The values for NNLS are out of range for $a > 0.3$.

3.3 Gamma/neutron inversion for the MSL/RAD instrument

In the previous section, a highly simplified DRF (Fig. 3.3) was used. The matrix was used as general example for exploring inversion techniques for gamma/neutron scintillator measurements and does not claim to reproduce the correct behavior of the RAD instrument. In this section, DRFs for the gamma and neutron response of the **D** and **E** detector are derived, which aim at reproducing the behavior of sensor head and electronics correctly.

Two geometric factor matrices for different tasks are calculated. For energies below 20 MeV, linear-spaced matrices with 1 MeV bins are calculated, which are used to invert calibration measurements. For the full energy range, log-spaced matrices are derived, whose binning corresponds to the format of the MSL/RAD gamma/neutron histograms.

3.3.1 Calculating the detector response matrices

The response of the detector to incident particles is calculated with the `GEANT4` toolkit [Agostinelli, 2003]. A detailed model, based on Computer-Aided Design (CAD) data of the RAD instrument, is used to calculate the energy deposits in the scintillators. The submatrices are calculated as follows. For each energy bin $[E, E + \Delta E]$, N particles are shot from a spherical surface at the detector. Directions and starting points of the primary particles are isotropically distributed. In a first step, the direct energy deposit in **D** and **E** is calculated for gammas and for neutrons. The number of particles in the energy range $[E, E + \Delta E]$, which deposit energy in $[E_{dep}, E_{dep} + \Delta E]$, is calculated as a 2D histogram of primary particle energy over deposited energy. To obtain the corresponding geometric factor, the energy bins are scaled with

$$g = 4\pi r^2 \cdot \frac{2}{N}, \tag{3.24}$$

where r is the radius of the spherical surface from which the particles are shot, N is the number of primary particles per energy bin, and the factor 2 denotes the fact that only one half of the primary particles are shot into the sphere. The geometric factor matrices for the direct, unquenched energy deposits are shown in Fig. 3.14 (top).

In a next step, the effect of quenching is considered. Following Kortmann [2010], the quenching of the energy deposit ΔE per step Δx is calculated with the Birks formula [Craun and Smith, 1970]

$$\frac{\Delta E_q}{\Delta x} = \frac{\frac{\Delta E}{\Delta x}}{1 + kB\frac{\Delta E}{\Delta x}}, \tag{3.25}$$

where the kB are given by

$$kB_{\mathbf{D}} = 1.52\frac{\mu m}{MeV} \tag{3.26}$$

$$kB_{\mathbf{E}} = 0.60 \cdot 10^{-2}\frac{cm}{MeV}. \tag{3.27}$$

The Birks constant for CsI was obtained from Avdeichikov *et al.* [2002], for the plastic scintillator the value from Kortmann [2010] was used, which was adapted to explain the calibration measurements, and differs from the literature value of 1.31 $\frac{g}{cm^2 MeV}$. Note that in contrast to Craun and Smith [1970], Eq. 3.25 states the energy deposit in the scintillator, not the light output. The geometric factor matrices for quenched energy deposits are shown in Fig. 3.14 (bottom). Comparing quenched and unquenched energy deposits in Fig. 3.14, one can clearly see the reduced energy deposit due to quenching in $\mathbf{A}_{E,N}$. In a last step, the PhotonMC, which was introduced in Kortmann [2010], is applied to calculate the light output of the scintillator. In addition, electronic noise is added.

The PhotonMC is an optical model of the light distribution in the scintillators. In the first step, the `GEANT4` simulation creates an output with energy deposit for each step a particle moves through the scintillator, in the second step the PhotonMC traces photons

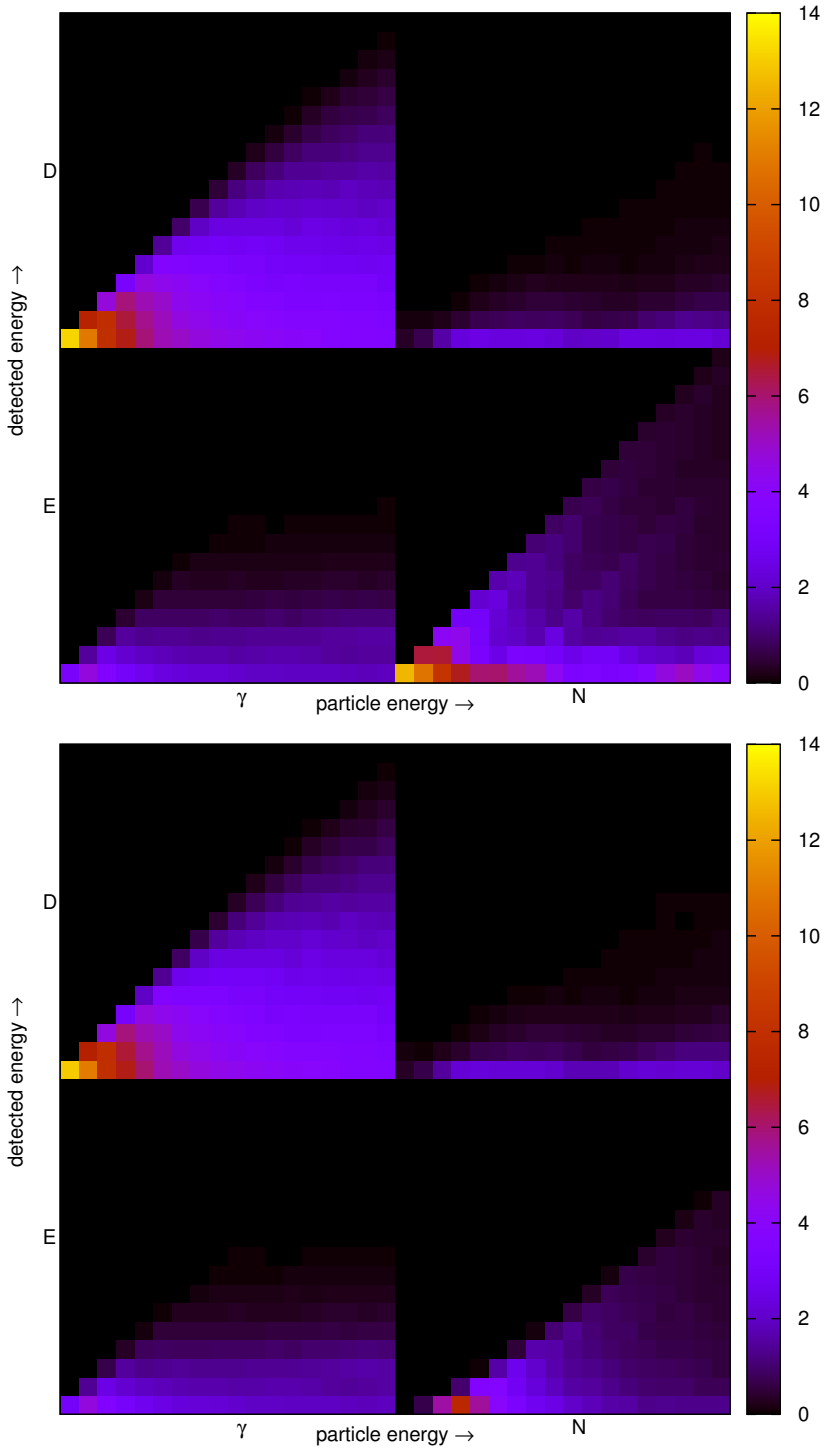

Figure 3.14: Geometric factor matrices calculated with `GEANT4`. The colors denote the geometric factor [cm^2 sr] for detecting a gamma/neutron with an energy E in a certain energy bin. Each matrix consists of the four submatrices $\mathbf{A}_{D,\gamma}$, $\mathbf{A}_{D,N}$, $\mathbf{A}_{E,\gamma}$ and $\mathbf{A}_{E,N}$. For each submatrix primary particle energy [1-19 MeV] versus deposited energy [1-19 MeV] is plotted. Top: Energy deposit without quenching. Bottom: Quenching was applied via Eq. 3.25.

from position of the energy deposit to a photodiode. The simulation mainly addresses the influence of the particle's path through the scintillator on the amount of light reaching a photodiode. This is a significant factor for RAD because there are three different photodiodes per scintillator. For example, a particle passing just in front of one photodiode creates a large signal in this diode and a small in the other two diodes, while a particle passing through the center of the scintillator creates equal signals in all diodes. Part of this effect can be compensated by using a weighted mean of all three photodiodes, see Fig. B.1 and Eq. B.1. But, if the energy deposit is small, as for gamma/neutron measurements, the signal in the low-gain diode is dominated by noise, leaving only the medium- and high-gain diode for calculating the weighted mean. To complicate things further, the medium-gain diode of the **D** scintillator of the flight spare model is defect, leaving only the high-gain diode.

Therefore, two different sets of matrices are generated, one for calibration measurements in the low energy range, using only the **DU** and **EU** channel, and one for the final flight configuration, with full energy range and the correct calculation of the weighted mean of the photodiodes.

3.3.2 Instrument response for particles below 20 MeV

The `GEANT4` toolkit calculates elastic and inelastic scattering of neutron's with various models. Depending on the neutrons energy, `GEANT4` choses between different theoretical and empirical models. The choice of the correct models and their boundaries is essential for describing the detector behavior correctly. Below 20 MeV, `GEANT4` uses empirical high precision scattering models. Considering the model transition and the insecurities in the theoretical models, DRFs for low energy measurements are generated for an energy range 1-19 MeV. The 1-19 MeV matrices are shown in Fig. 3.15.

The neutral particle histograms, which will be generated by the RAD instrument in its final configuration, will reject silicon hits, i.e. particles creating a direct energy deposit in one of the photodiodes. The discrimination between scintillator and silicon hits is realized by comparing the signals of the three photodiodes attached to the **D** and **E** detector, respectively. If an event shows similar light distribution between the three diodes, it is counted as a valid event. If an event creates a large signal in only one photodiode, it is rejected as a silicon hit. An asymmetry in light distribution, shown by the PhotonMC model in Kortmann [2010], will complicate a silicon hit rejection, especially since the medium-gain photodiode in **D** can not be used. Therefore, Fig. 3.15 includes silicon hits and will be used for the inversion of calibration measurements and comparisons between different inversion techniques. For completeness, Fig. 3.16 shows the low-energy response function with silicon hit rejection.

3.3.3 Instrument response for the full energy range

In its final configuration, RAD will process neutral particle events onboard and will send particle histograms for the **D** and **E** detector back to Earth. The histograms will both

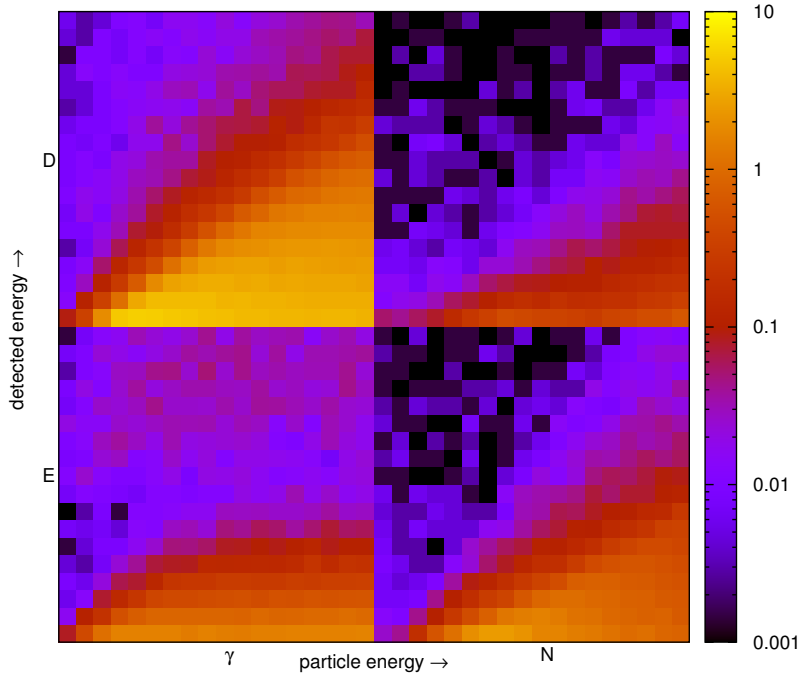

Figure 3.15: Geometric factor matrices calculated with GEANT4 and the PhotonMC model. The colors denote the geometric factor [cm^2 sr] for detecting a gamma/neutron with an energy E in a certain measurement channel. Each matrix consists of the four submatrices $\mathbf{A}_{D,\gamma}$, $\mathbf{A}_{D,N}$, $\mathbf{A}_{E,\gamma}$ and $\mathbf{A}_{E,N}$. For each submatrix primary particle energy [1-19 MeV] versus detector response [measurement channel 2500-7000] is plotted.

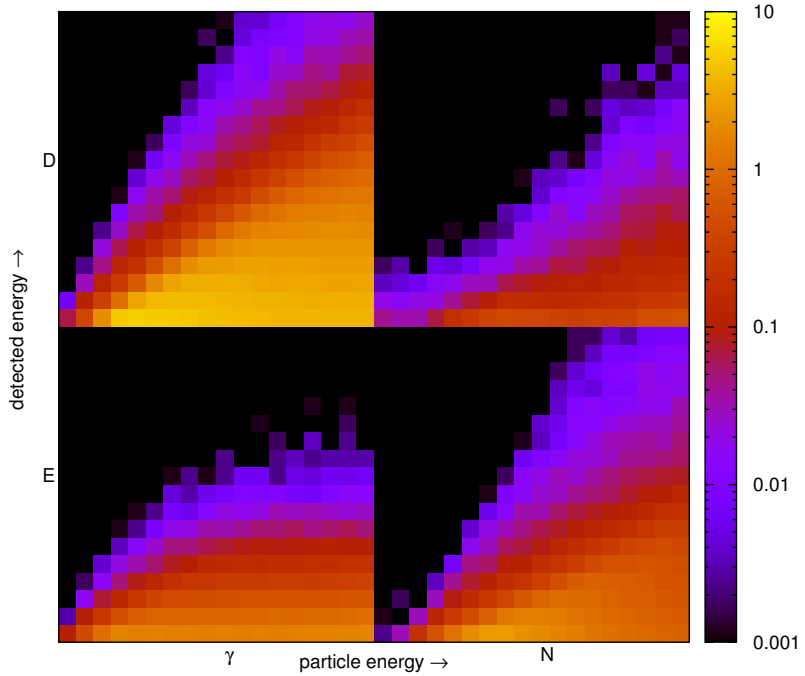

Figure 3.16: Same matrices as in Fig. 3.15, but with applied silicon hit rejection.

consist of 48 log-spaced energy bins, which cover the full energy range of the detector [Southwest Research, 2010]. Considering the given histogram formats, the corresponding DRFs are generated with 48×48 log-spaced bins. The energy range of the input spectra is given by 3 - 300 MeV for gamma particles and by 4.5 - 300 MeV for neutron particles.

The ranges of the measurement vectors are given by 4-256 MeV for both detectors, for a detailed description see Appendix B. In contrast to the 19 MeV matrix from Fig. 3.15, the full DRF only considers particles which do not trigger the AC, given by **F1**, **F2** and **C**. The silicon hit rejection and the E_{sum} process is replicated by simulating the light output for all photodiodes, and calculating the resulting weighted mean as described in Appendix B.1. To reduce the number of free parameters, two additional set of matrices with 48×24 and 48×12 bins are generated as well. The 48×48 matrices are shown in Fig. 3.17.

3.3.4 PTB calibration measurements

The neutron beam calibration measurements performed at the Physikalisch-Technische Bundesanstalt (PTB) in Braunschweig provide an opportunity to test the inversion method and the geometric factor matrix from Fig. 3.15, not only for artificially generated, but also for real measurements. The sensor head was exposed to 5, 14.8 and 19 MeV neutron beams [Nolte *et al.*, 2004]. The 5 MeV neutron-beam measurement showed large systematic deviations from the expected behavior [Kortmann, 2010] and is therefore excluded from this analysis. The inversion of the PTB data is also described in Köhler *et al.* [2011], where in addition to the Poisson and NNLS method, the chi-square method is tested. Before applying inversion methods, the ability of the `GEANT4` MC and the PhotonMC code to model the sensor head is tested.

Fig. 3.18 compares the calibration runs with `GEANT4` simulations of monoenergetic 14.8 and 19 MeV neutron beams which enter the instrument from above. Measurements and simulations agree very well for both **D** and **E** detector. Because the trigger threshold is not included in the simulation, it does not agree with the measurement for channels ≤ 2500. We can therefore assume that the DRF is able to describe both the 14.8 and 19 MeV run correct. However, one should note that, among other measurements, the PTB measurements were used to create the PhotonMC model in Kortmann [2010]. An additional 5 MeV measurement could not be reconstructed. As already shown in Kortmann [2010], this measurement does not agree with the `GEANT4` simulations.

As a more challenging task, a merged measurement from the 14.8 and 19 MeV neutron beams is inverted as well. The measurement is created from both measurement vectors as $\vec{z} = 0.25\,\vec{z}_{14.8\mathrm{MeV}} + \vec{z}_{19\mathrm{MeV}}$, where the factor 0.25 was introduced to obtain similar intensities for both beams. Using the results from Sec. 3.2.6, initial guesses are calculated as follows.

1. A set of prototype shapes is generated. In this case 18 neutrons beams with different beam energies and 16 powerlaws with spectral indices 0, -1, -2, -3 with all possible combinations for gamma and neutron spectral indices.

2. For each prototype shape several inversions are made (Eq. 3.21)

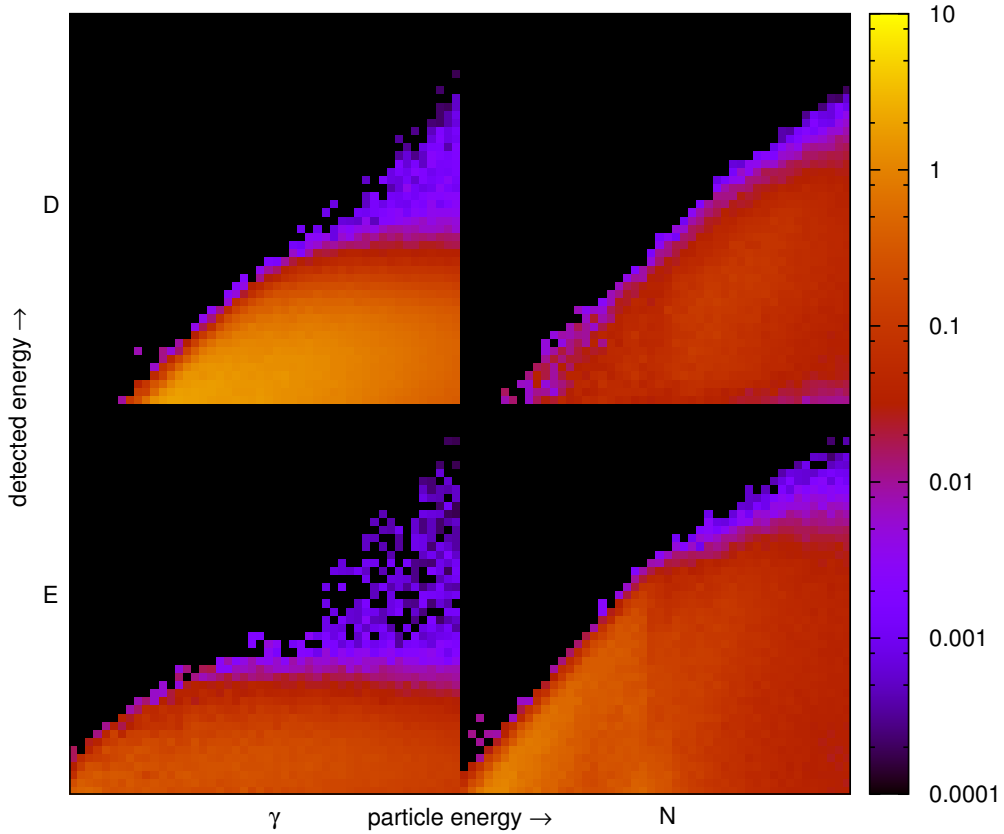

Figure 3.17: The complete 96×96 geometric factor matrices for the full energy range of RAD. The colors denote the geometric factor [cm^2 sr] for detecting a gamma/neutron with an energy E in a certain ADC channel. The matrix consists of the four submatrices $\mathbf{A}_{D,\gamma}$, $\mathbf{A}_{D,N}$, $\mathbf{A}_{E,\gamma}$ and $\mathbf{A}_{E,N}$. For each submatrix primary particle energy versus detector response is shown in log-spaced bins. The measured energy range is for both **D** and **E** given by 4-256 MeV. The energy range of the input spectrum is given by 3 - 300 MeV for gamma particles and by 4.5 - 300 MeV for neutrons. Here, each submatrix has 48 energy bins for the input spectrum. Depending on available statistics, the number of bins can be reduced to 24 or 12 energy bins. The number of energy bins of the measurement vectors is determined by Southwest Research [2010].

- The inversion for the initial guess from Eq. 3.15 is calculated to get a first estimation of the intensities for f_γ and f_N.
- The inversion for the prototype shape is calculated with intensities adjusted from the first estimation.
- The resulting intensities are used to adjust the intensities of the prototype shape for another inversion attempt.
- Intensities of the prototype shape are iteratively adjusted until the merit function improves no further. Resulting spectra and merit value are stored.

3. Determine the solution with the minimum merit.
4. Find the 2, 3, 4 ... best matching prototype shapes via $\frac{\vec{f}_{\text{prot}} \cdot \vec{f}_{\text{solution}}}{|\vec{f}_{\text{prot}}||\vec{f}_{\text{solution}}|}$.
5. Create merged spectra from the 2, 3, and 4 best matching prototype shapes weighted with $\frac{\vec{f}_{\text{prot}} \cdot \vec{f}_{\text{solution}}}{|\vec{f}_{\text{prot}}||\vec{f}_{\text{solution}}|}$.
6. For each merged spectrum iteratively adapt intensities (as described in step 2).
7. Select result with minimum merit from all prototype and merged shapes.

In step 4 the best solution is copied with few prototype spectra. If all prototype spectra are weighted and merged to a new initial guess, the initial guess would be identical to the previous solution (if the prototype spectra provide a sufficient base of vectors for the parameter space). Since the initial guess starts in the local minimum from a previous attempt, no improvement will be made. Using very few prototype spectra will recreate the fundamental shape of the solution while neglecting glitches. For the simple neutron beam spectrum, a correction of over/underestimations, as described in Sec. 3.2.6, was neglected. Step 3 can be improved by selecting several prototype spectra with the smallest resulting merit and execute step 4 for each of them. A general description of the initial guess selection is given in Appendix A.2.

Fig. 3.19 shows the resulting inversions for the 14.8 MeV and the 19 MeV neutron beam. As described above, an additional combined measurement vector from the 14.8 MeV and 19 MeV measurement $\vec{z} = 0.25\vec{z}_{14.8\text{MeV}} + \vec{z}_{19\text{MeV}}$ was inverted as well, the results are shown in the lower panel. For all methods, the neutron beams can be identified remarkably well. However, a gamma background is also found. Unfortunately, the spectrum of the induced gamma radiation is not known, therefore we have no way of determining the accuracy of the inverted gamma spectrum.

For the 14.8 MeV neutron beam, the Poisson method shows a clear beam in the 14.5 MeV energy bin, while the NNLS inversion finds most neutrons in the 15.5 MeV bin and only few neutrons in the 14.5 MeV bin. For the 19 MeV neutron beam, all methods clearly identify the 19 MeV neutrons and do not show any significant deviation in intensities. Although not as clear as for the 14.8 and 19 MeV measurement, the combined neutron beams are clearly resolved. As before the NNLS method finds the maximum of the 14.8

MeV beam in the 14.5 MeV bin. The inverted beam intensities differ for the two methods. This is especially visible for the 19 MeV beam in the combined measurement. While the Poisson method reproduces the total number of counts in the two separate neutron beams, the NNLS method overestimates the 19 MeV intensity. The results, shown in Fig. 3.19, were obtained using the randomization of initial guesses described in Eq. 3.20. As shown before for MC measurements, this improves the Poisson method considerably, but has a much lesser effect on the NNLS method.

3.4 Conclusions

In this chapter an inversion approach for gamma/neutron scintillator measurements was developed. The inversion is based on a maximum likelihood estimation with underlying Gaussian or Poissonian statistics. For the example of an unspecified gamma/neutron scintillator measurement, both methods were compared for artificially generated measurements with different count rates. As expected, the Poisson method produces better results for low count rates. For the error estimation of the statistical errors of an inverted measurement, the bootstrap MC approach was applied. The choice of the initial guess was discussed in detail, and an algorithm which supplies and optimizes initial guesses for a given measurement was developed. The influence of errors in the DRF was evaluated for both methods. In contrast to the NNLS method, the Poisson method showed a remarkable robustness to errors.

Based on the `GEANT4` and PhotonMC codes, two DRFs for MSL/RAD were derived, one matrix for the low energy range, one for the full energy range. Using the low-energy-range DRF for RAD, both methods were tested with PTB neutron beam measurements. For 14.8 MeV and 19 MeV neutron beams, both method produced very good results, as well as for a combined measurement with two neutron beams.

It could be shown that the Poisson method produces significantly better results for low count rates, where the Poissonian nature of the measurement is dominant. However, when minimizing the log likelihood value, it is much more sensitive to the choice of the initial guess. While the NNLS method finds its global minimum (or a local minimum close to it) even for poor initial guesses, the Poisson method needs a good initial guess to find the global minimum. By developing a method which supplies, selects and optimizes initial guesses, and by performing multiple inversions per measurement, the ability of the Poisson method to find the global minimum could be improved. Another problem that occurred was that the calculation of the numerical gradient for the Poisson-log-likelihood function is error-prone to numerical noise. This could be solved by adjusting the step width for the calculation of the numerical gradient. The developed strategies are available within a Python library, which inverts a measurement with a given DRF using a set of supplied initial guesses and optionally applies initial guess optimization or step width adaption for the numerical gradient.

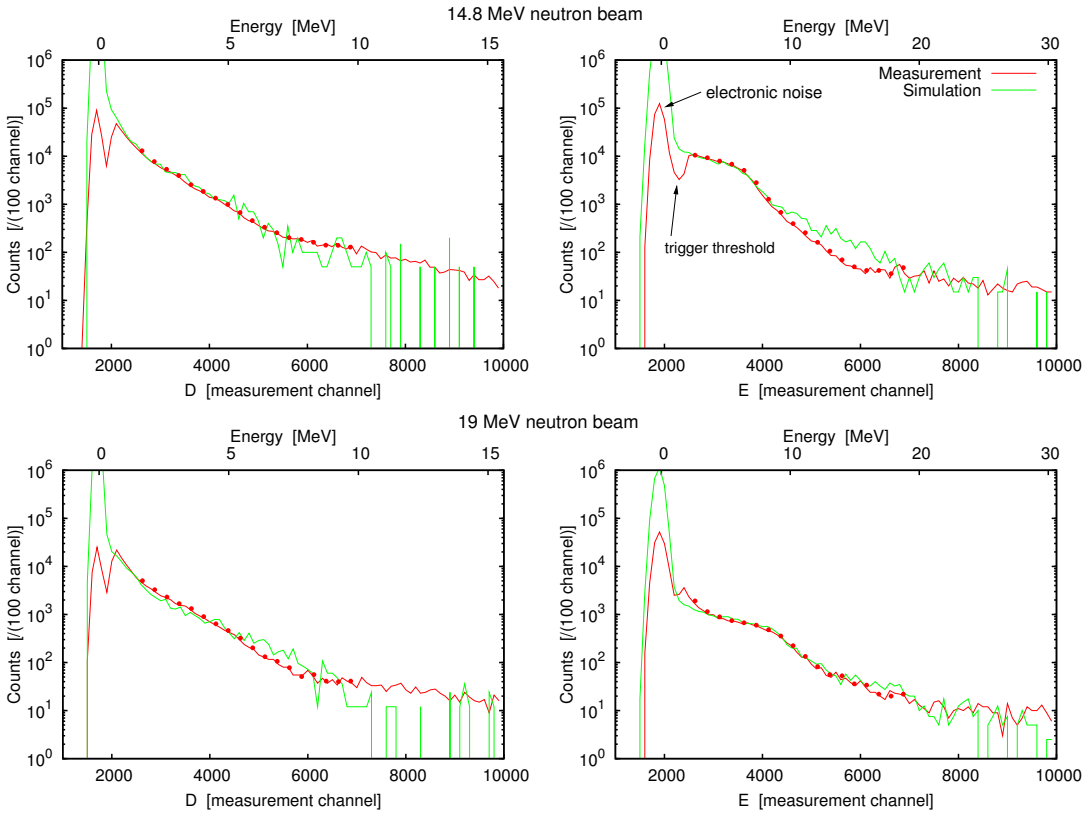

Figure 3.18: PTB calibration measurement (red) and simulation (green) for the 14.8 MeV (top) and 19 MeV (bottom) neutron beam for the **D** (left) and **E** detector (right). Measurements are compared with a 10^6 particle `GEANT4` simulation. Measurements and simulations match very well for both detectors. Because the trigger threshold is not included in the simulation, we do not expect agreement in measurement channels ≤ 2500. The red dots show the corresponding measurement vectors $\vec{z}_D$, $\vec{z}_E$, which are binned to match the DRF from Fig. 3.15. To allow better comparison, their intensities are reduced by a factor of 4, corresponding to the increased bin size.

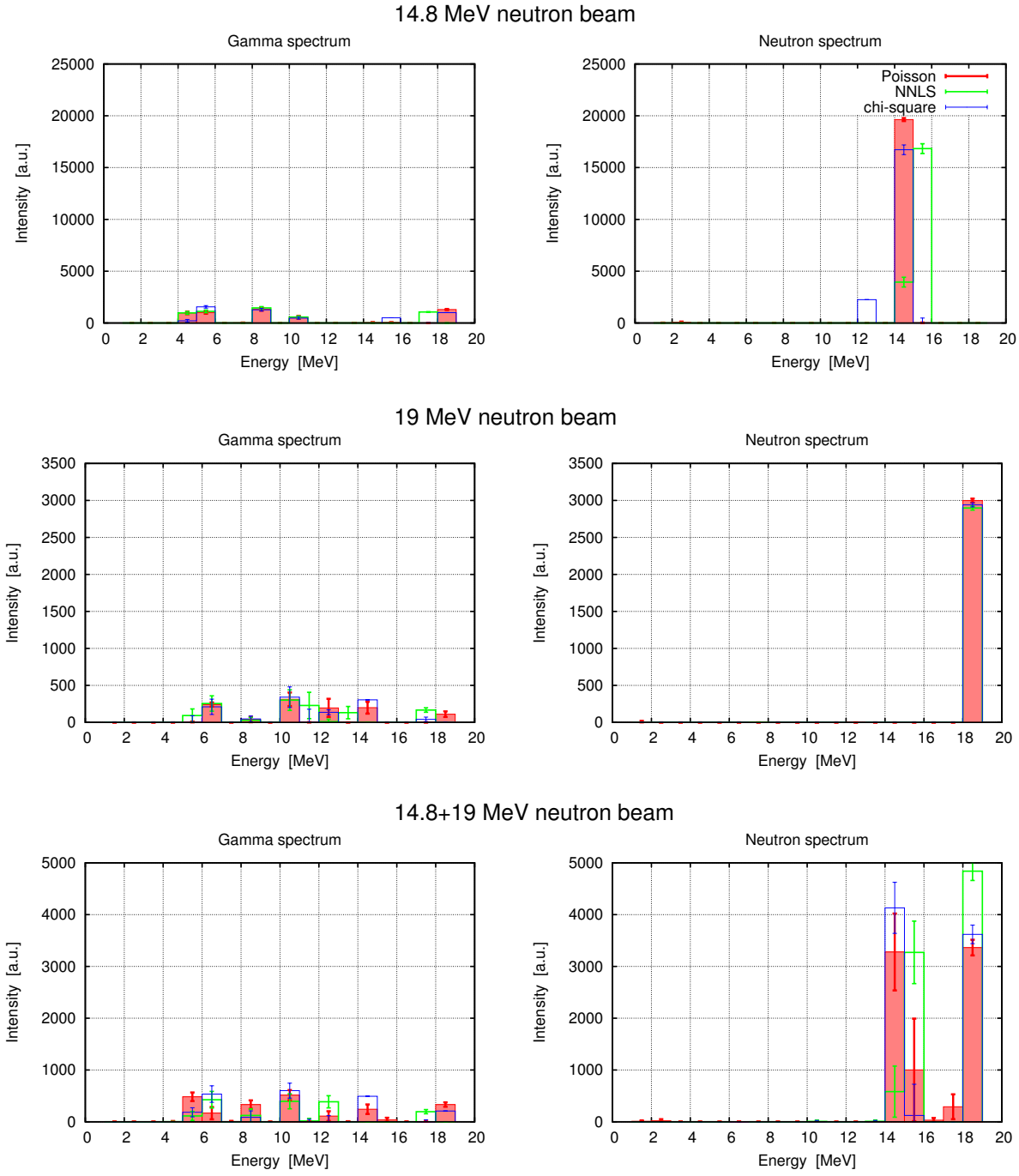

Figure 3.19: Inverted gamma and neutron spectra for 14.8 MeV and 19 MeV neutron beam measurements. The bottom panel shows an artificial measurement, where the 14.8 MeV and the 19 MeV measurement were merged into a new measurement vector $\vec{z} = 0.25\,\vec{z}_{14.8\mathrm{MeV}} + \vec{z}_{19\mathrm{MeV}}$. Compared are the results of the Poisson, the NNLS and the chi-square method, which is described in Köhler *et al.* [2011].

Chapter 4

Gamma/neutron measurements on Mars

In contrast to the PTB measurements in the previous chapter, the RAD neutral particle measurements on Mars will be processed into histograms before they are sent back to Earth. In Appendix B, the boundaries of the neutral particle histograms $\vec{z}_D$, $\vec{z}_E$ are determined to fit the restriction of the maximal allowed count rate and the ranges of the input spectra $\vec{f}_\gamma$, $\vec{f}_N$ are chosen to create an optimal DRF. The resulting DRF is shown in Fig. 3.17.

In addition to the 1D histograms, a 8×8 bin 2D histogram will be generated, which stores neutral particle events which generate an energy deposit in both **D** and **E**. This occurs for example, if a recoil proton leaves the **E** detector and deposits energy in **D** as well, or if an electron created by a gamma escapes **D** and deposits some energy in **E**. The 2D histogram provides additional information, which can be used to improve the inversion. In this chapter, the inversion methods are tested for the expected Martian neutral particle spectra and the influence of the RTG counts is estimated. Finally, a method to invert measurements with coincidence conditions is derived, which is then used to improve the results for the expected Martian measurements by including the 2D neutral particle histogram.

4.1 Gamma and neutron spectra on Mars

In the previous section, it was shown that the initial guess is a crucial factor for the quality of the resulting spectra. Without basic knowledge of the input spectrum an initial guess can not be easily estimated from the measurement vector $\vec{z}$. In most cases basic knowledge of the input spectrum is given. Since RAD will measure particles in the Martian radiation environment, the instrument behavior for the expected Martian gamma/neutron spectra is investigated. In Ehresmann [2012], the radiation in the Martian atmosphere, induced by GCR protons and helium, is calculated for several altitudes and atmospheric densities with `Planetocosmics` [Desorgher *et al.*, 2006]. As an estimation for RAD, the gamma/neutron spectra at 0 km altitude during solar minimum condition are selected.

	a	b
Gamma spectrum	3.85 ± 0.98	-1.57 ± 0.10
Neutron spectrum	1.78 ± 0.50	-1.39 ± 0.12

Table 4.1: Resulting parameter set for fitting Eq. 4.1 to the Martian gamma/neutron spectra from Fig. 4.2.

Fig. 4.1 shows the omnidirectional fluxes for neutrons and gammas, i.e. the sum of upward and downward directed fluxes without any consideration of angular dependence, for the full `Planetocosmics`-calculated energy range. For a wide range of energies, neutron and gamma flux are in the same order of magnitude. A clear functional dependence, e.g. a power law, can not be seen. In Fig. 4.2, the energy range relevant for the RAD instrument is shown in detail. In the plotted energy range (1 - 400 MeV), both spectra can roughly be approximated by power law shapes. The spectra show some deviations, e.g. the neutron evaporation peak at 1 MeV, which is also present in the GCR induced neutron spectrum of Earth [Hess *et al.*, 1959]. Nevertheless, both spectra are approximated by a power law

$$I \;=\; a \cdot E^b, \tag{4.1}$$

which can at least serve as an initial guess for an inversion. The fitted parameters are shown in Tab. 4.1.

4.1.1 Different energy resolutions

To compare the three matrices (48, 24 and 12 energy bins) for the expected Martian particle spectra, MC measurements with $N = 10^7$ particles are generated, which corresponds to ~ 77 hours of measurement. Using this large number of events ensures that the stochastic nature of the measurement process is not the dominant factor for the inversions.
Poisson and NNLS method are compared in Fig. 4.3 and 4.4. Each figure shows an example for a single inverted measurement, the average inverted spectra, and the relative errors for each energy bin.

In the top panel, one can clearly see for a single inverted gamma/neutron spectrum that the errors determined via bootstrap MC appropriately describe the inverted spectrum. The inverted spectrum, together with its error bars, is well in range of the real input spectrum. The average behavior is shown in the middle panel. For gamma and neutron spectra at low energies one finds a systematic discrepancy for the 48×48 and the 48×24 matrices, which is in the range of the given error bars. This discrepancy does not occur for the 48×12 matrix. Because in log space the given error bars appear comparatively small, the relative errors are shown in the bottom panel to give a better comparison. As expected, the relative errors decrease with a decreasing number of energy bins. However, comparing the mean relative errors for 12, 24, and 48 bins, one finds an increase which is not due to statistics alone. From counting statistics one would expect an increase by a factor of $\sqrt{2}$

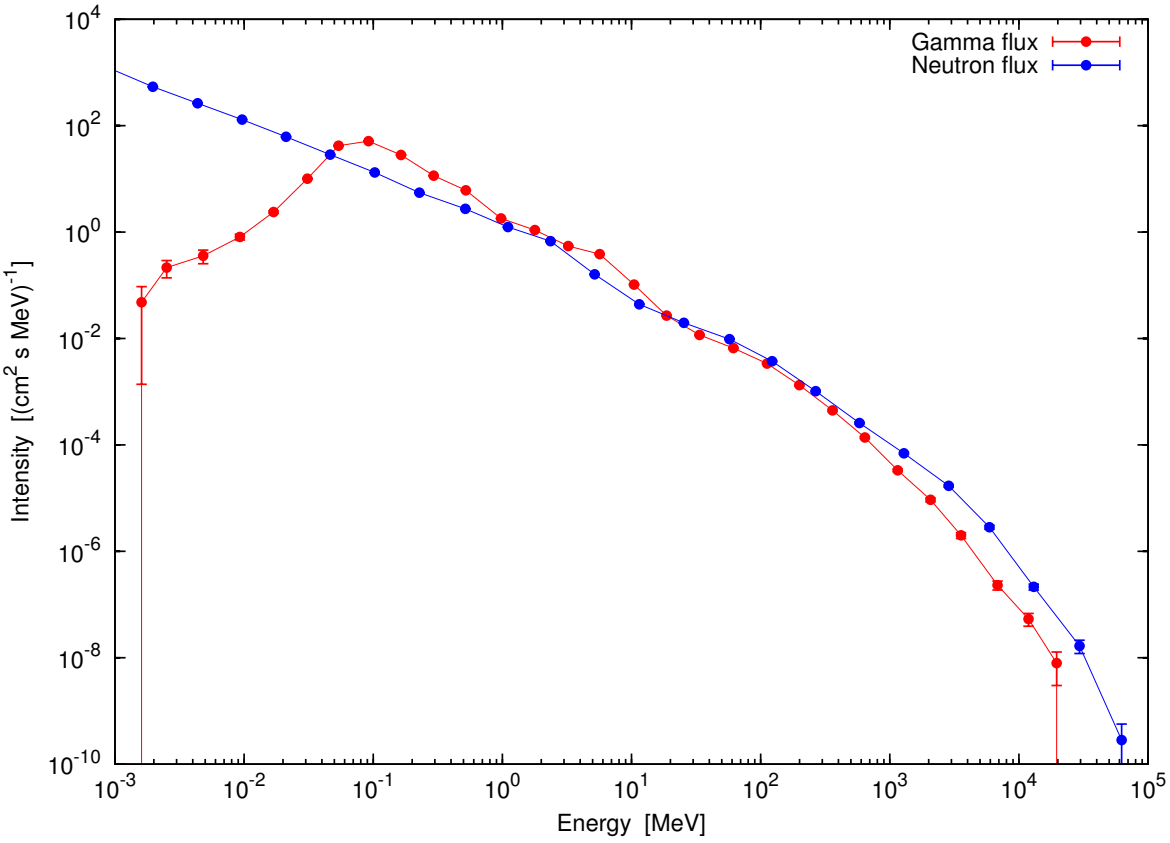

Figure 4.1: GCR proton and helium induced gamma/neutron flux at the Martian surface during solar minimum as calculated with `Planetocosmics`. Angular dependences are not considered, shown here is the sum of up- and downward directed flux.

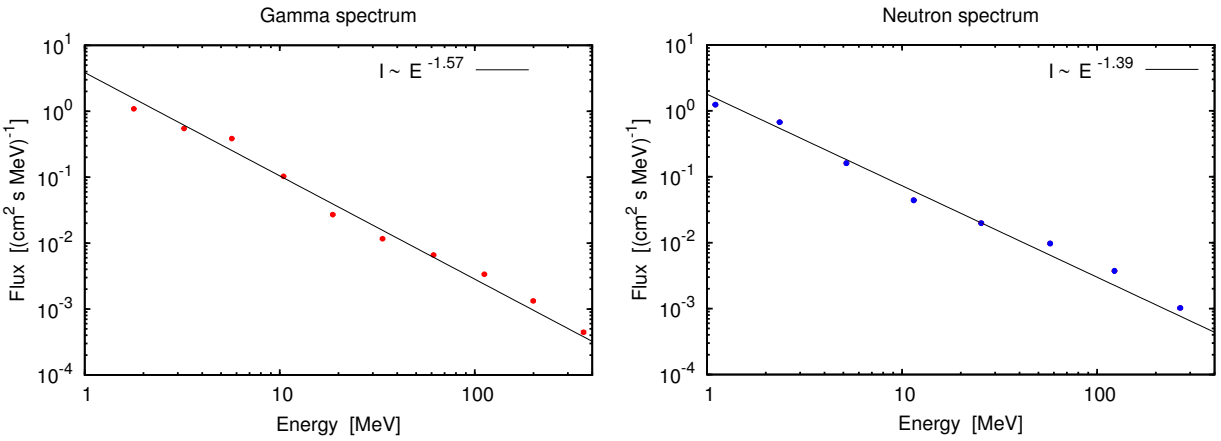

Figure 4.2: The gamma/neutron spectra from Fig. 4.1 in the energy range relevant for the RAD measurements (1 - 400 MeV). Both spectra can roughly be approximated with a power law. The fitted parameters can be found in Tab. 4.1.

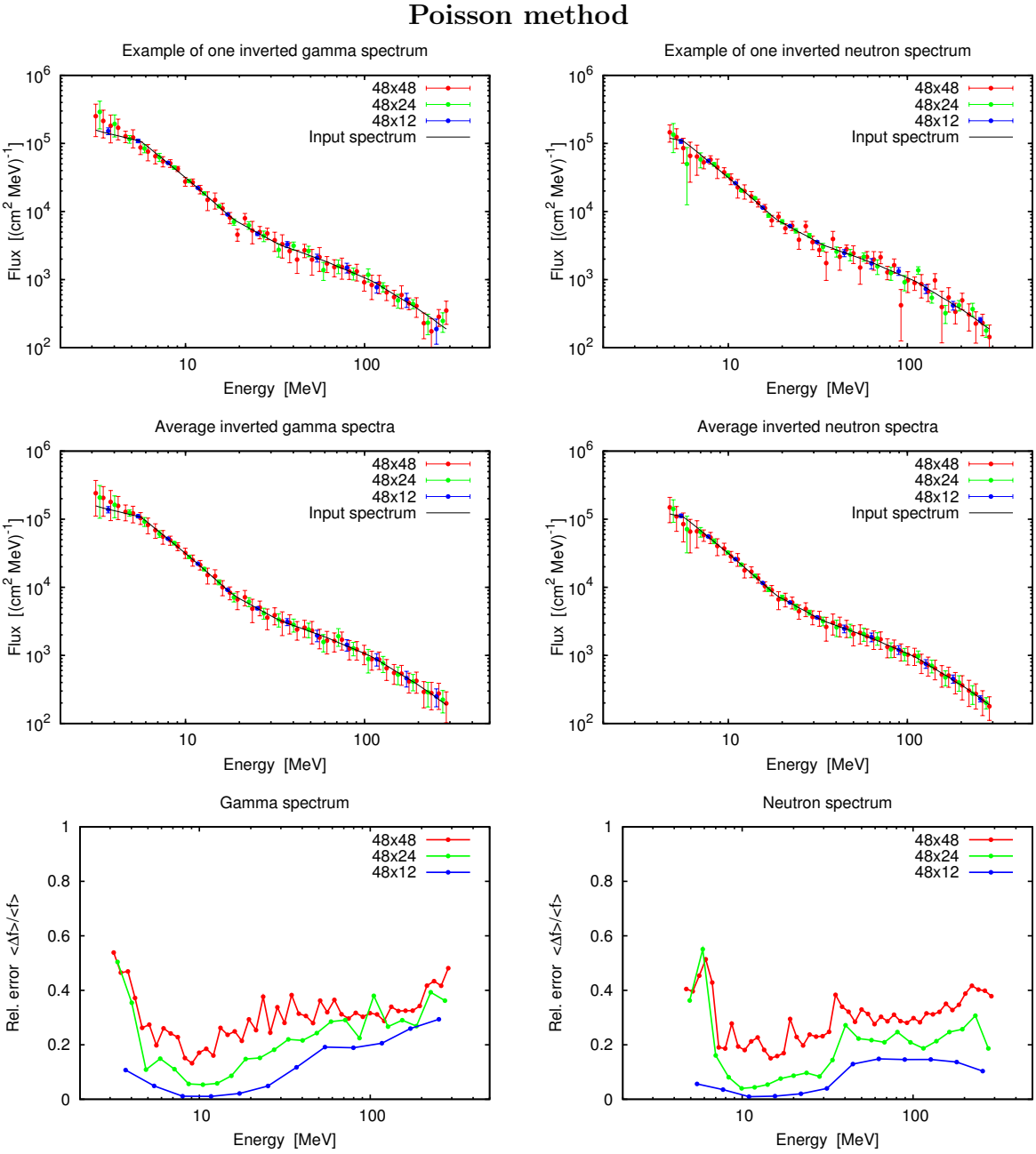

Figure 4.3: Inversion with underlying Poissonian statistics for estimated Martian gamma/neutron spectra with $N = 10^7$ MC generated events. From top to bottom: Inverted spectra for one MC measurement, average of 500 inverted spectra and the average relative errors. Each plot compares the different DRFs with 48, 24 and 12 energy bins. The corresponding figures for the NNLS method are shown in Fig. 4.4.

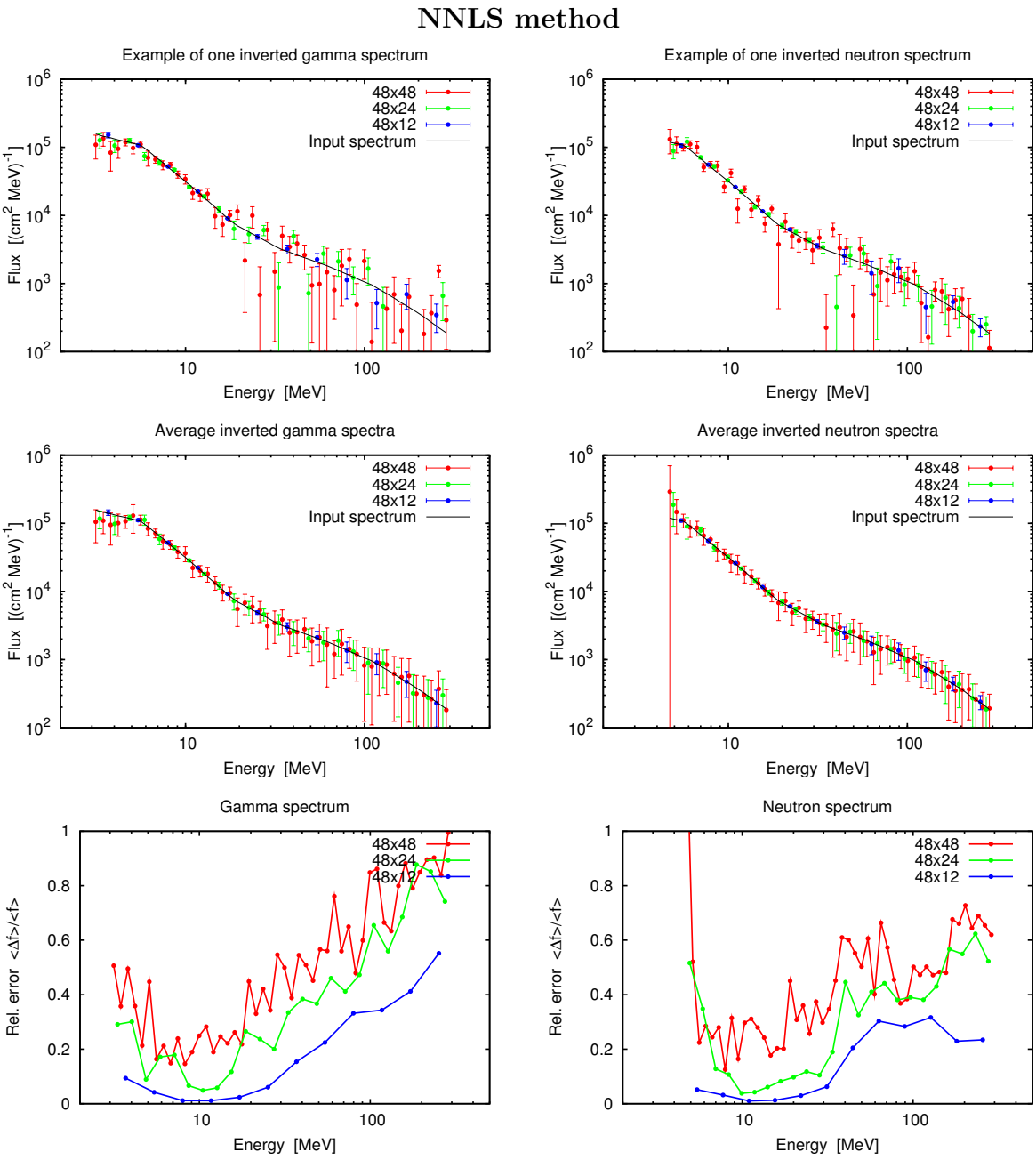

Figure 4.4: Inversion with underlying Gaussian statistics for estimated Martian gamma/neutron spectra with $N = 10^7$ MC generated events. From top to bottom: Inverted spectra for one MC measurement, average of 500 inverted spectra and the average relative errors. Each plot compares the different DRFs with 48, 24 and 12 energy bins.

in the relative error for a doubled number of bins, i.e. half the number of counts per bin. As expected from counting statistics, the increase of the relative errors is $\sim \sqrt{2}$ for 24 - 48 bins. However, for 12 - 24 bins the increase of the relative errors is $\sim$2, which shows that the binning in the DRF is a crucial factor for the inversion and an increased number of bins can not always simply be compensated by an increased number of counts. The mean relative errors are shown in Tab. 4.3.

Considering the NNLS method (Fig. 4.4), one finds a strong dependence on the number of energy bins. For 48 energy bins, the inversion produces no usable results and only for 24 and 12 energy bins the relative errors are similar to the Poisson method. As for the Poisson inversion, the NNLS method produces a systematic error for low energies.

The direct comparison of Fig. 4.3 and 4.4 shows that even for high count rates the Poisson method produces significantly better results. Therefore, it is selected for further evaluation.

4.1.2 Realistic measurements

For the given Martian gamma/neutron spectra, $N = 10^7$ events correspond to $\sim$ 77 hours of continuous measurement. Since RAD will take a 15 minute measurement every hour, 77 hours of "continuous" measurement equals 308 hours. Therefore, more realistic N are given by $N = 32400,\ 64800,\ 129600$ and 259200 events, which are the expectation values for 15, 30, 60, and 120 minutes of measurement. One should note that the DRF assumes that no neutral particle is falsely rejected (false negatives). However, depending on the final AC settings, the rate of false negatives may be as high as 50%, which will reduce the count rates by a factor of 2.

Fig. 4.5 shows the relative errors for the Poisson inversion for 15, 30 and 60 minutes of measurement and DRFs with 24 and 12 energy bins. The 48×48 response matrix has been neglected because it produced only poor results even for $N = 10^7$. The relative errors show basically the same shapes as in Fig. 4.3. As expected, the inversion for 12 energy bins shows significantly better results.The inversions with 24 energy bins produce large relative errors and do not seem to be suitable for short measurement times.
The gamma spectra show the smallest relative errors around 10 MeV, the neutron spectra show the smallest relative errors at 10 - 30 MeV. For the 15 minutes measurement, the relative errors of the gamma spectrum exceed 1 for energies above $\sim$ 50 MeV, the relative errors of the neutron spectrum have a maximum of $\sim$ 0.8. For a 120 minute measurement, the maximum of the relative errors of the gamma spectrum is < 0.8 and for the neutron spectrum below 0.5. While a single 15 minute measurement gives only very poor results, a 60 minute measurement already produces adequate results.

4.2 Gamma/neutron background from the RTG

The Radioisotope Thermoelectric Generator (RTG) onboard MSL supplies power for the rover. In the RTG the heat is generated mainly through α-decays of ^{238}Pu. The pluto-

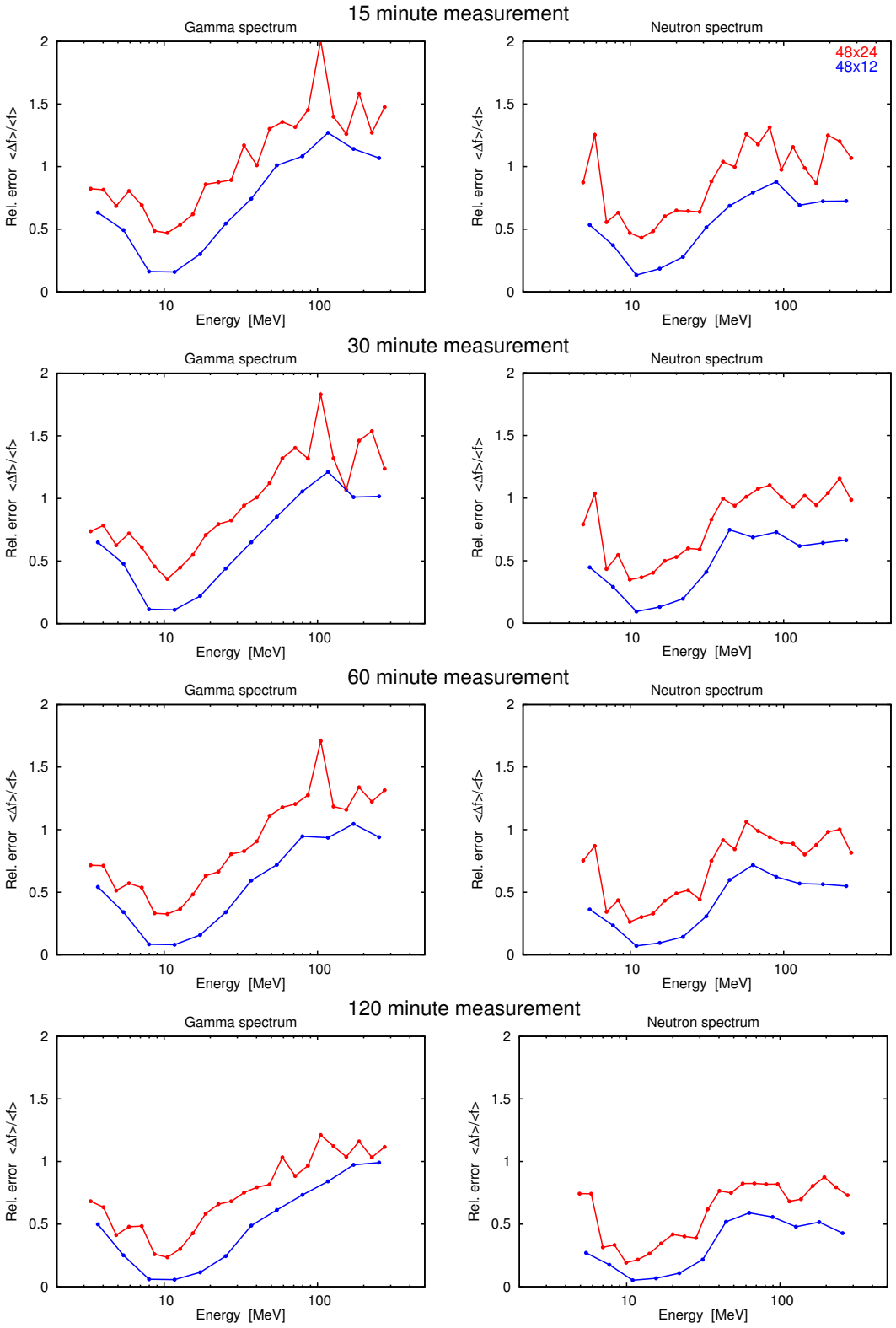

Figure 4.5: Relative errors of the inverted Martian gamma/neutron spectra for 15, 30, 60 and 120 minutes of continuous measurement, which corresponds to $N = 32400$, 64800, 129600 and 259200 particle events. Shown are the results for Poisson method for a DRF with 24 and 12 energy bins.

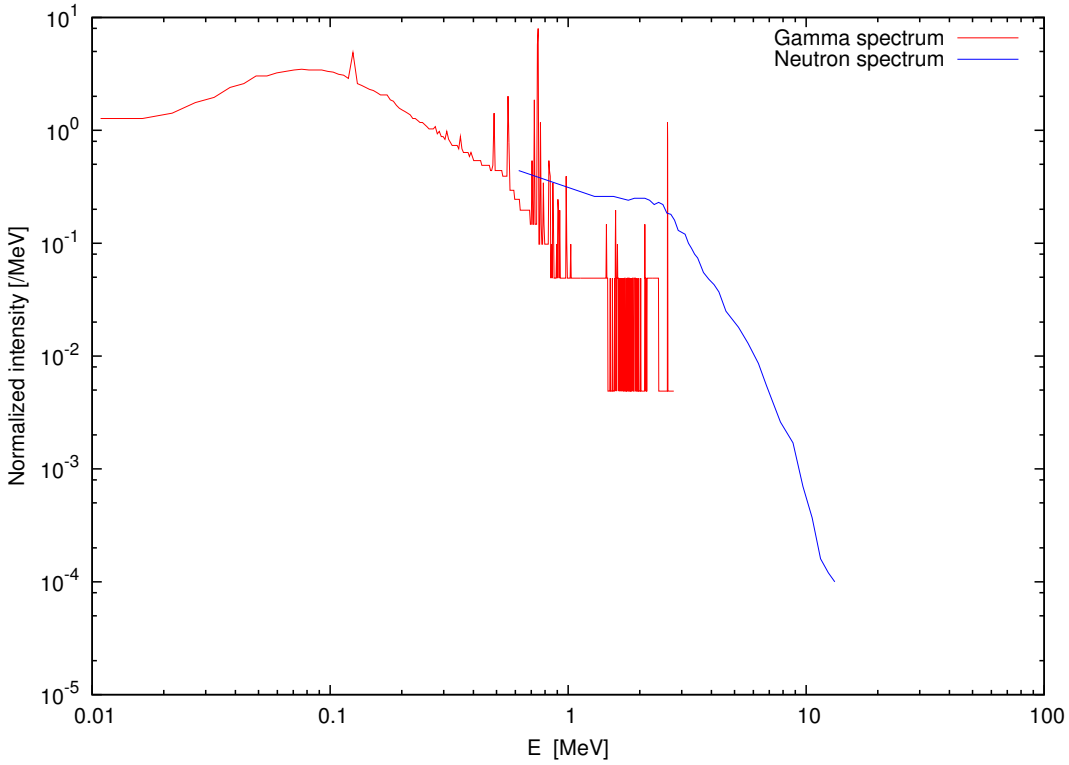

Figure 4.6: Gamma and neutron spectra from radiation measurements on the Ulysses qualification RTG. Intensities of both spectra are normed to unity. No significant gamma flux is expected above 3 MeV. The neutron spectrum can be extrapolated with an exponential behavior.

nium decays with a half time of $T_{1/2}$ =87.7 y [NuDat2.5, 2011]. Within the decay chain fast neutrons up to ∼19 MeV and gammas up to ∼3 MeV are produced. Especially the neutron flux of the RTG will create a strong background which will reduce the effective sensitivity for neutral particle measurements. Therefore, detailed knowledge about the RTG's gamma/neutron spectra is needed to estimate its influence on the gamma/neutron measurements.

The exact amount of the RTG's plutonium, as well as detailed information about its position or the source's shielding, are not known. However, the gamma/neutron spectra for the Ulysses RTG, which uses ^{238}Pu as well, are available [Hawley, 1984] and can be used to estimate the shape of the RTG spectra. The spectra, obtained from the datasets in Hawley [1984], are shown in Fig. 4.6. Since no information about the relative intensity is available, both spectra are normed to unity.

Neither the spectra, nor the exact position and shielding of the RTG are known, but RAD measured the RTG spectra at the Idaho National Laboratory (INL) for several distances and different shielding between RTG and the instrument.

A general problem for the INL run is the defective medium-gain photodiode in **D**. The

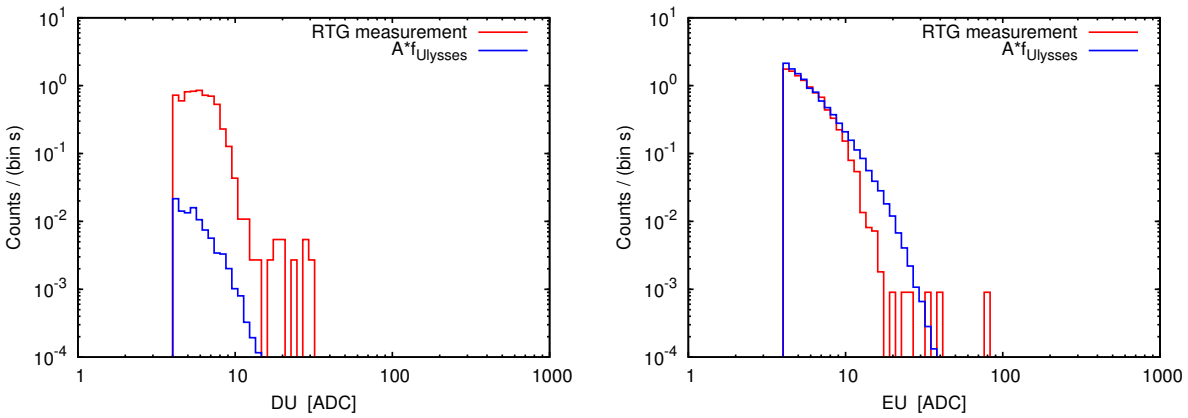

Figure 4.7: Original and matched artificial RTG measurements for the `inl_edc_36_12_20100218_16_51_0` run. For the intensity $I_N = 10$ the measurement in the **E** detector can be recreated very well, the measurements in **D** suffer from the defect medium-gain photodiode and the disabled AC **F1**. A silicon hit rejection, as applied for **E**, is therefore nearly impossible. This means charged particles, entering the instrument from the side, are shown as well. The RTG gammas have only low energies and therefore do not appear in the **D** or **E** histogram.

silicon hit rejection is only based on the low- and high-gain diode and leads to unsatisfactory results. In addition, the **F1** AC was disabled, therefore charged particles could not be rejected. Another possible problem is the high particle flux of the RTG, which produces more counts than the electronics can process. This can in principal be solved by a rate correction. Due to these problems, a direct approach of inverting the measurements yields no results which would be at least similar to the expected spectra. However, using the neutron spectra from the Ulysses RTG as the input spectrum, its intensity I_N can be fitted to the measurement. The best attempt of replicating the measurements is shown in Fig. 4.7. For the intensity $I_N = 10$, the measurement in the **E** detector can be recreated very well. Unsurprisingly, the measurement in the **D** detector can not be recreated. Its high intensity could be due to charged particles, problems with the silicon hit rejection, errors in the rate correction, or caused by secondary particles.

As already mentioned, position and shielding of the RTG are not the same as in the INL measurement, therefore I_N and the shape of the spectrum may change considerably. Nevertheless, the neutron spectrum of the Ulysses RTG, scaled with the INL intensities, will be used to investigate the influence of the RTG and the measurements on Mars. Pre-launch and cruise phase measurements of the RTG spectra will give some reliable gamma/neutron measurements.

A direct comparison of Martian and RTG neutron spectra is shown in Fig. 4.8. One can clearly see that the RTG spectrum is dominant for neutrons up to 9 MeV. Considering the huge uncertainties of the RTG intensities, this value may vary by some MeV.

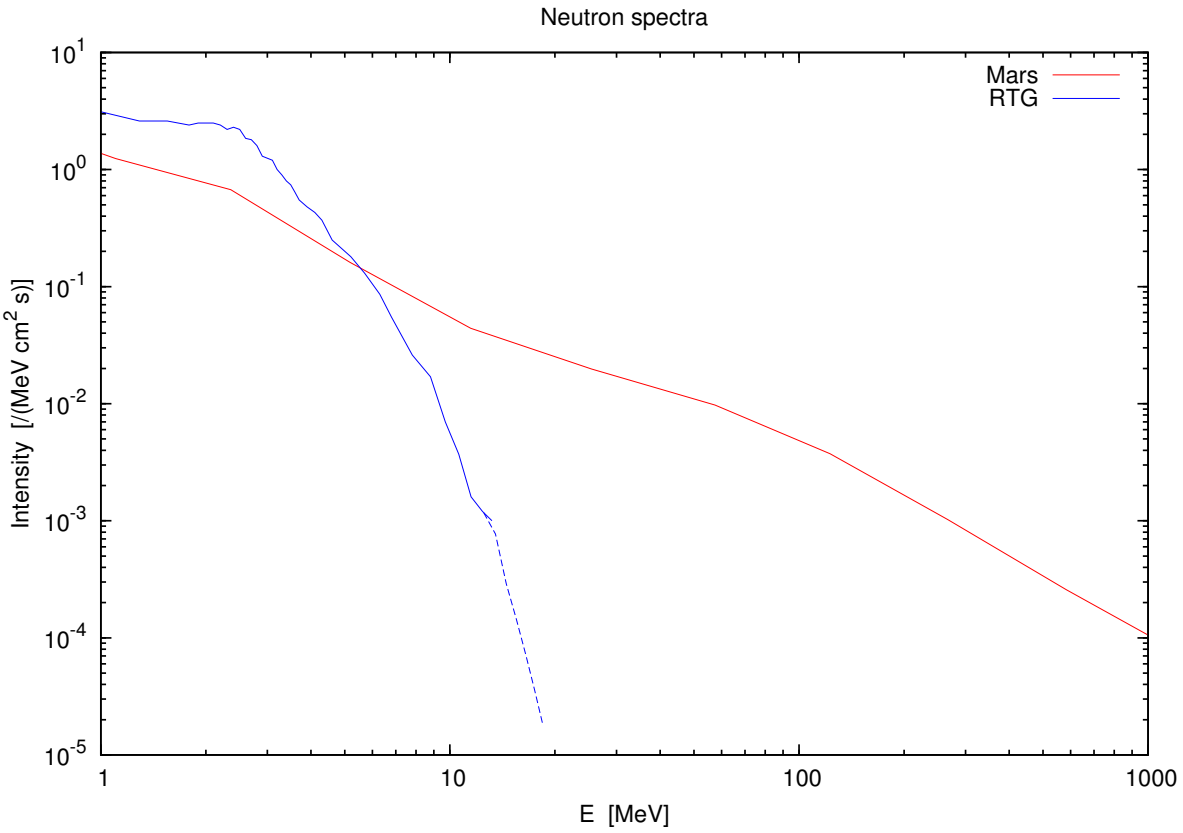

Figure 4.8: Comparison of Martian and expected RTG neutron spectrum. The intensity of the Ulysses RTG spectrum was fitted to match the INL measurement (Fig. 4.7).

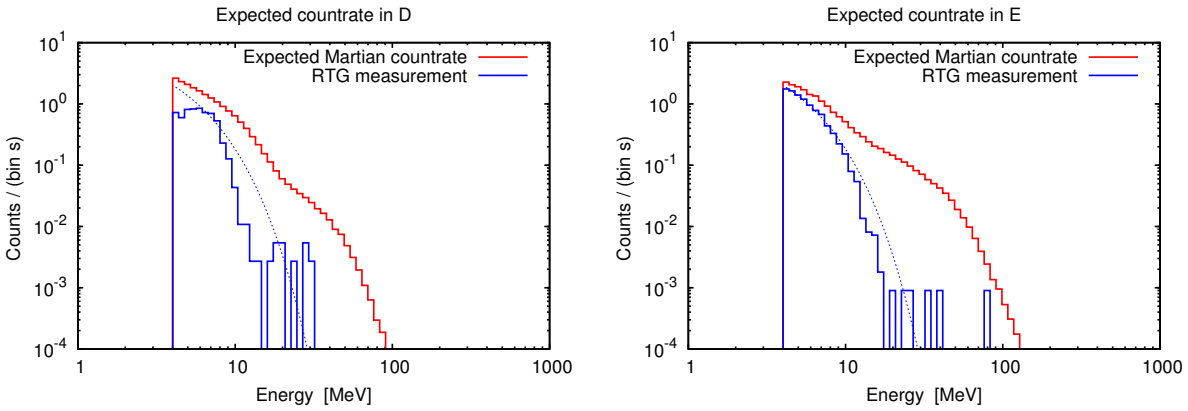

Figure 4.9: Expected Martian gamma/neutron counts for a 1 s measurement and counts from RTG measurement `inl_edc_36_12_20100218_16_51_0` with intensity normalized to 1 s. The measurements are binned to match the DRF from Fig. 3.17. A functional approximation (Eq. 4.2) of the RTG measurement is plotted dashed.

4.2.1 Removing the RTG counts from measurement

Depending on the assumptions about the RTG spectrum, there are several approaches to distinguish between RTG and Martian neutral particle spectra. In principle it is possible to subtract the RTG spectrum from the resulting spectrum after the inversion procedure. However, the shape of the spectrum is not known. Even if the MSL RTG spectrum had the same shape as the Ulysses RTG, shielding and backscattered particles would modify the resulting spectrum. Those effects can not easily be estimated. However, the count rate in **D** and **E**, caused by the RTG, can be measured directly before launch and during cruise phase. Therefore, the most promising approach is to subtract the expected RTG counts from the measurement, leaving the remaining counts to be of Martian origin. The intensity of the RTG spectra will decrease over time, and since the half life of ^{238}Pu is known, it can be used to account for the decrease.

The expected count rates for RTG and Martian neutral radiation are shown in Fig. 4.9. Although the neutron spectrum in Fig. 4.8 shows more RTG neutrons than Martian neutrons for low energies, the count rate induced by Martian radiation is higher than the RTG induced count rate for each energy bin. This is due to the fact that the count rate at low energy bins is not only caused by low energy neutral particles but also by high energy particles which deposit only little energy. The measurements that are, especially for **D**, only a rough approximation of the expected RTG counts on Mars can be approximated by

$$f_\gamma(E) = \exp\left(-\frac{E}{2.5\text{MeV}}\right) \cdot 10/(\text{bin s}) \tag{4.2}$$

$$f_N(E) = \exp\left(-\frac{E}{2.5\text{MeV}}\right) \cdot 10/(\text{bin s}). \tag{4.3}$$

Since there are less RTG than Martian counts in each measurement bin, it should in principle be possible to subtract the expected RTG counts from the measurement $\vec{z}_{Mars} = \vec{z}_{\text{measurement}} - \vec{z}_{RTG}$ before the inversion, without generating negative count rates due to Poissonian noise.

To account for the Poissonian nature of the measurement, several inversions are done. For each inversion, Poisson distributed random values, with expectation values given by the expected RTG count rate $\vec{z}_{RTG}$, are subtracted from the measurement. The final result is given by the mean of all inversions. This procedure is done in the same step as the error determination via bootstrap MC. The errors resulting from the RTG noise are therefore included in the bootstrap MC determined error bars.

Fig. 4.10 shows the resulting relative errors for inversions with different RTG intensities I, where the number of events from the RTG are given by $I \cdot z_{\gamma/N}(E)$. The increase of the relative errors is surprisingly small, i.e. the mean relative error for the expected RTG signal increases only by a factor of 0.08 (0.1) for an inversion with 12 (24) energy bins and 10^7 Martian gamma/neutron events. The increase of the relative errors due to the RTG is small compared to the relative errors already present due to the Poisson noise. The increase is mainly at low energy bins where, the influence of the RTG is the strongest (Fig. 4.9). An overview over the increase of the mean relative errors is given in Tab. 4.2.

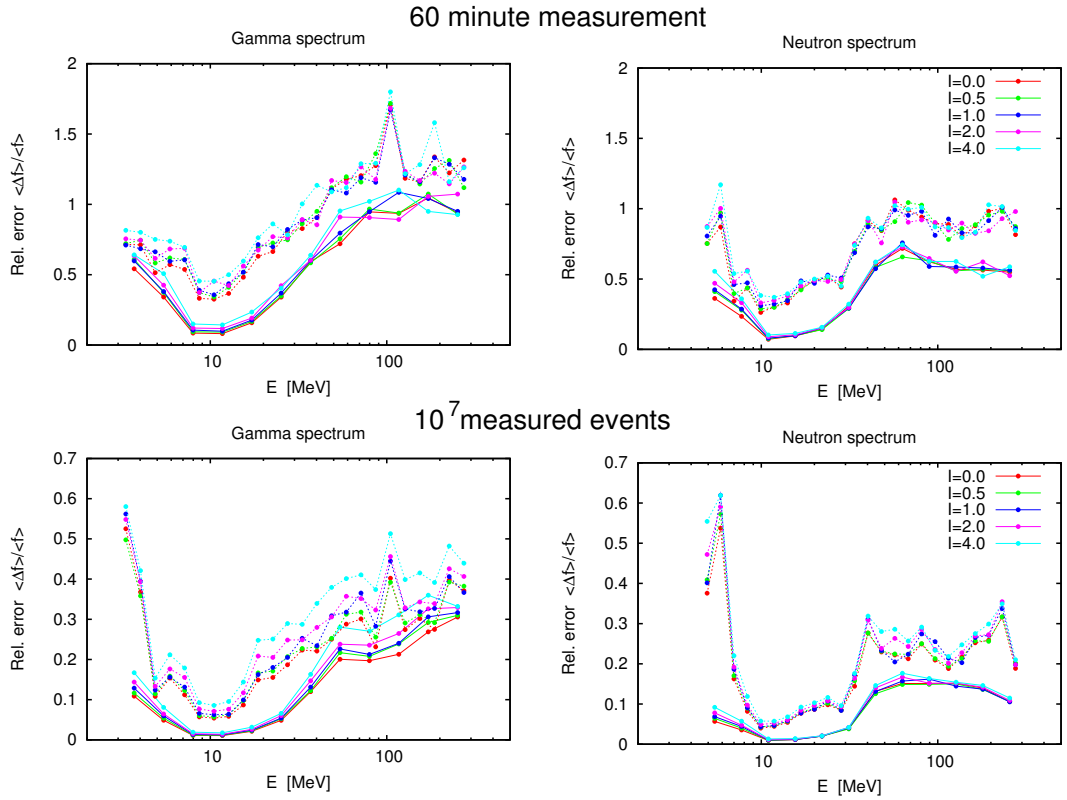

Figure 4.10: The relative errors for an inversion of the expected Martian spectra under the influence of the RTG. Shown are a 60 minute (∼129600 events) measurement (top) and a measurement with 10^7 events (bottom) for 12 energy bins (solid) and 24 energy bins (dashed). The expected RTG counts and the corresponding Poisson noise were subtracted before each inversion procedure.

N	energy bins	$I = 0$	$I = 0.5$	$I = 1$	$I = 2$	$I = 4$
129600	12	0	0.019	0.049	0.080	0.124
129600	24	0	0.014	0.013	0.030	0.091
10^7	12	0	0.039	0.079	0.149	0.267
10^7	24	0	0.033	0.095	0.155	0.293

Table 4.2: Relative increase of the mean relative error for an RTG signal (Eq. 4.2) with intensity I. Compare Fig. 4.10.

4.3 Inverting measurements with coincidence conditions

The approach in the previous chapters describes the measurement of neutral particles in an instrument consisting of several detector elements. Each neutral particle is assumed to deposit energy in either **D** or **E**, not in both detectors. However, high energy neutral particles can create energy deposits in both detectors. The approach developed in the previous chapter does not consider this coincidence.

A typical measurement concept for a measurement with coincidence conditions is given by a telescope of silicon detectors. Charged particles are detected in coincidence and a particle's species and energy can be determined by comparing the energy deposits in the different detectors. Viewing the measurement for each detector as a one-dimensional vector $\vec{z}$, which contains counts per measurement channel, neglects the information about the coincidence for energy deposits belonging to one event. To include this information, a measurement can be viewed as an N-dimensional matrix $\mathbf{z}$, where N is the number of detectors. For a telescope of two silicon detectors, $\mathbf{z}$ is the typical 2D histogram which shows the number of events over energy deposit in the first versus energy deposit in the second detector.

For a spectrum of incoming particles $\vec{f} = (f_1, ..., f_n)$, where the component f_i is the intensity in th i-th energy bin E_i, and a N-dimensional measurement $\mathbf{z}$, a measurement can be mathematically described by

$$\mathbf{z} = \underline{\mathbf{A}} \cdot \vec{f} \tag{4.4}$$

$$= \begin{bmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_n \end{bmatrix} \cdot \vec{f}, \tag{4.5}$$

where $\underline{\mathbf{A}}$ is a stack of N-dimensional matrices $\mathbf{A}_i$ $(i = 1, ...n)$, and $\mathbf{A}_i$ is the response of the detectors to a particle from the i-th energy bin. The multiplication between the stack of response matrices and the spectrum of incoming particles is defined as

$$\underline{\mathbf{A}} \cdot \vec{f} := \sum_{i=1}^{n} \mathbf{A}_i \; f_i. \tag{4.6}$$

4.3.1 Reduction to a one-dimensional measurement vector

Instead of using N-dimensional measurements and stacked DRFs, the measurement matrix and the detector response stack can be flattened to a one dimensional vector or a two dimensional matrix, respectively. This can be achieved by

$$z_i = [\mathbf{z}]_{i_1,..., \, i_N}$$
$$a_{j,i} = [\mathbf{A}_j]_{i_1,..., \, i_N},$$

where z_i are the components of the flattened measurement vector and $a_{i,j}$ are the components of the corresponding two dimensional DRF. The indices $i_1, ..., i_N$ are given by

$$\begin{aligned} i_N &= i \mod n \\ i_{N-1} &= i/n \mod n \\ &\vdots \\ i_{N-j} &= i/n^j \mod n \\ &\vdots \\ i_1 &= i/n^N. \end{aligned}$$

Eq. 4.4 is therefore reduced to the inverse problem described by Eq. 3.1, which can be solved by numerous methods. However, especially for two dimensional measurements the stacked DRFs and the measurement are far more easy to interpret than the corresponding flattened DRF and measurement. In this work, the minimization is achieved by a quasi-Newtonian method, i.e. the algorithm minimizes the free parameters for the given merit function without any explicit knowledge of the DRF. Since both methods are mathematically equivalent and the high dimensional measurement is more easy to interpret, the unflattened detector response stack is used in the following sections.

4.3.2 Multiple particle species

Considering multiple particle species $p_1, ..., p_m$ can be achieved by using a stacked vector as an input spectrum $\vec{f} = \vec{f}_{p_1} \oplus ... \oplus \vec{f}_{p_m}$ and extending the stacked DRFs

$$\mathbf{z} = \begin{bmatrix} \mathbf{A}_{p_1,1} \\ \vdots \\ \mathbf{A}_{p_1,n} \\ \vdots \\ \mathbf{A}_{p_m,1} \\ \vdots \\ \mathbf{A}_{p_m,n} \end{bmatrix} \cdot \vec{f}. \tag{4.7}$$

For the example of the gamma/neutron measurement with MSL/RAD, $\mathbf{z}$ is a one dimensional matrix, containing the counts per measurement bin in the $\mathbf{D}$ and $\mathbf{E}$ detector, the $\mathbf{A}_i$ are given by the lines of the corresponding DRF (Fig. 3.3), and $\vec{f} = \vec{f}_\gamma \oplus \vec{f}_N$ is the stacked vector of gamma and neutron spectrum. Eq. 4.7 is then equivalent to Eq. 3.1.

For the example of stopping charged particles in a silicon telescope, this approach yields no significant benefits since the particles energy can be calculated directly from the measurement. But for a combination of stopping and penetrating particles and different particle species in one measurement, i.e. particles with different species and energies are measured in the same bins of $\mathbf{z}$, this approach can be used to calculate the particle spectra.

It should be carefully considered how many detectors, i.e. the dimension of N, are used for an inversion. For example, one could describe the MSL/RAD measurements as 32-dimensional matrices (one dimension for each channel) considering all particle species at the same time. This would most likely lead to statistical problems and in general to uninterpretable data. In most cases a cut logic, which discriminates between particles species, penetrating and stopping particles and sets the AC condition for neutral particles, is much more easy to handle than a high-dimensional measurement.

4.3.3 Inversion with coincidence conditions for RAD

Besides the two neutral particle histograms for **D** and **E**, RAD generates a 2D histogram for neutral particles which deposit energy in **D** and **E**. An event which deposits energy in **D** and **E** can, e.g., occur when a neutron-induced recoil proton leaves the **E** detector and deposits some of its energy in **D**. Although neutral particles depositing energy in both scintillators are comparatively rare, the 2D histogram yields some additional information, which can be used to assist in the inversion of the neutral particle histograms.
A measurement for the 2D neutral particle histogram is given by

$$\mathbf{z}^{2D} \;=\; \underline{\mathbf{A}} \cdot \vec{f}, \tag{4.8}$$

where $\mathbf{z}^{2D}$ is the 8×8 measurement histogram and $\vec{f}$ is the same input spectrum as for the gamma/neutron inversion with 1D histograms. The stacked DRF for the 2D histogram consists of

$$\underline{\mathbf{A}} = [\mathbf{A}_{\gamma,1}, ..., \mathbf{A}_{\gamma,n}, \mathbf{A}_{N,1}, ..., \mathbf{A}_{N,n}].$$

Each $\mathbf{A}_{\gamma/N,k}$ is a 8×8 matrix, where an element $a_{i,j} = \mathbf{A}_{\gamma/N,k,(i,j)}$ represents the geometric factor for a gamma/neutron from energy interval f_k, for depositing energy in measurement bin $z^{2D}_{i,j}$. The combined merit function is given by the sum of the merit functions for the 1D and 2D histogram inversion. With underlying Poissonian statistics, it is given by

$$l(f) \;=\; \sum_i (\lambda_i - z_i \ln \lambda_i) + \sum_{i,j} (\lambda^{2D}_{i,j} - z^{2D}_{i,j} \ln \lambda^{2D}_{i,j}), \tag{4.9}$$

where $\lambda^{2D}_{i,j} = (\underline{\mathbf{A}} \cdot \vec{f})_{i,j}$.
Several examples of the $\mathbf{A}_{\gamma,i}$ and $\mathbf{A}_{N,i}$ for the same $\vec{f}$, as used for the RAD DRF from Fig. 3.15, are shown in Fig. 4.11.

4.3.4 Using the 2D histograms to improve the results

The additional information from the 2D histograms can be used to improve the inversion, i.e. decrease the relative errors of the neutral particle measurement on Mars. In Fig. 4.12, the relative errors from Fig. 4.3 are compared with the relative errors of an inversion including the 2D histograms. The merit functions are given by Eq. 3.8 and 4.9.

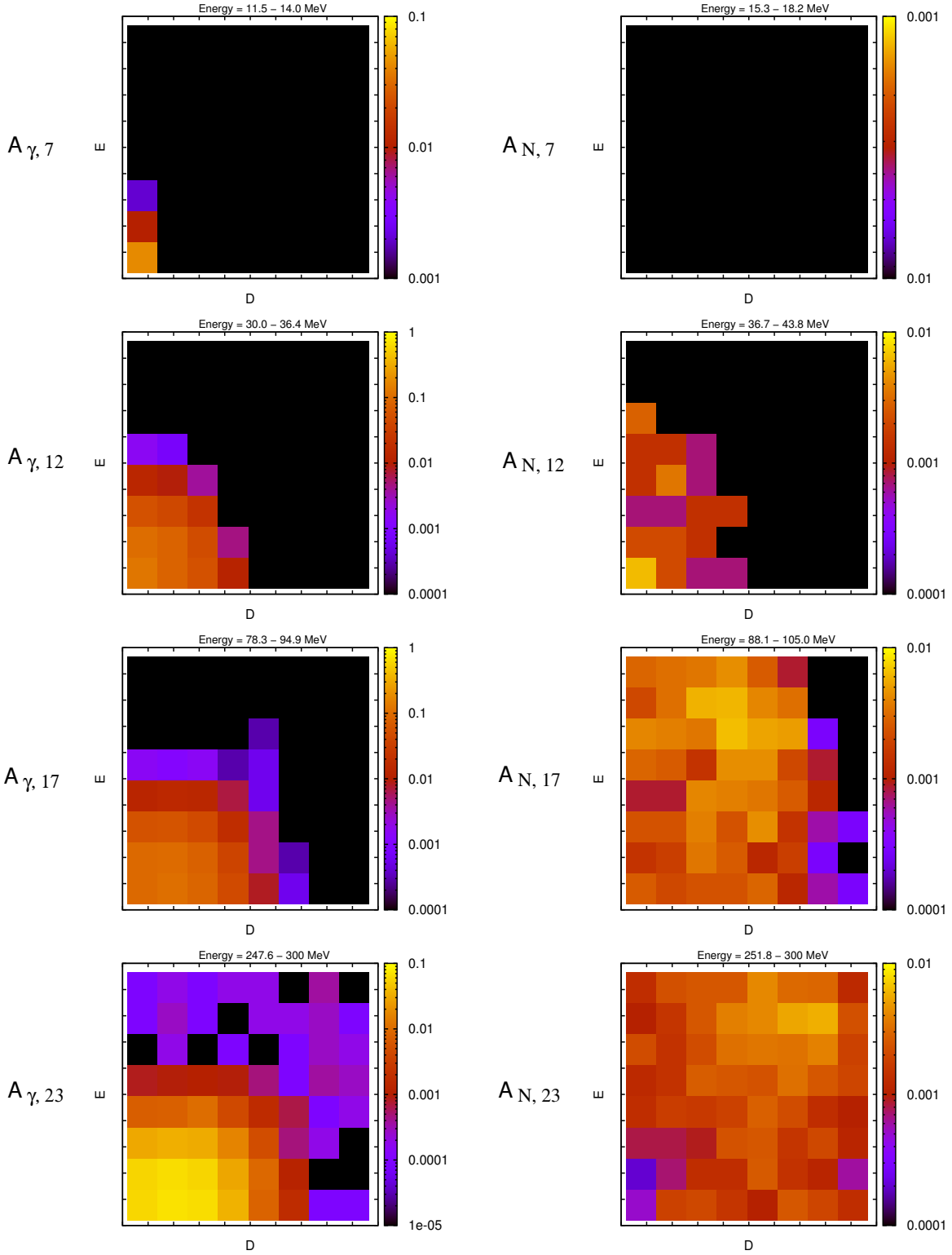

Figure 4.11: Examples of the DRFs $\mathbf{A}_{\gamma/N,i}$ for the 2D measurement. The color denotes the geometric factor in [sr cm^2] for a gamma/neutron with energy $f_{\gamma/N,i}$ to deposit its energy in a measurement bin of $\mathbf{z}^{2D}$. The measurement bins along the $x(y)$-axis give the energy deposit in the **D** (**E**) detector. The energy spectra $\vec{f}_\gamma$, $\vec{f}_N$ are the same as for the 1D inversion approach, above each $\mathbf{A}_{\gamma/N,i}$ the corresponding gamma/neutron energy is given.

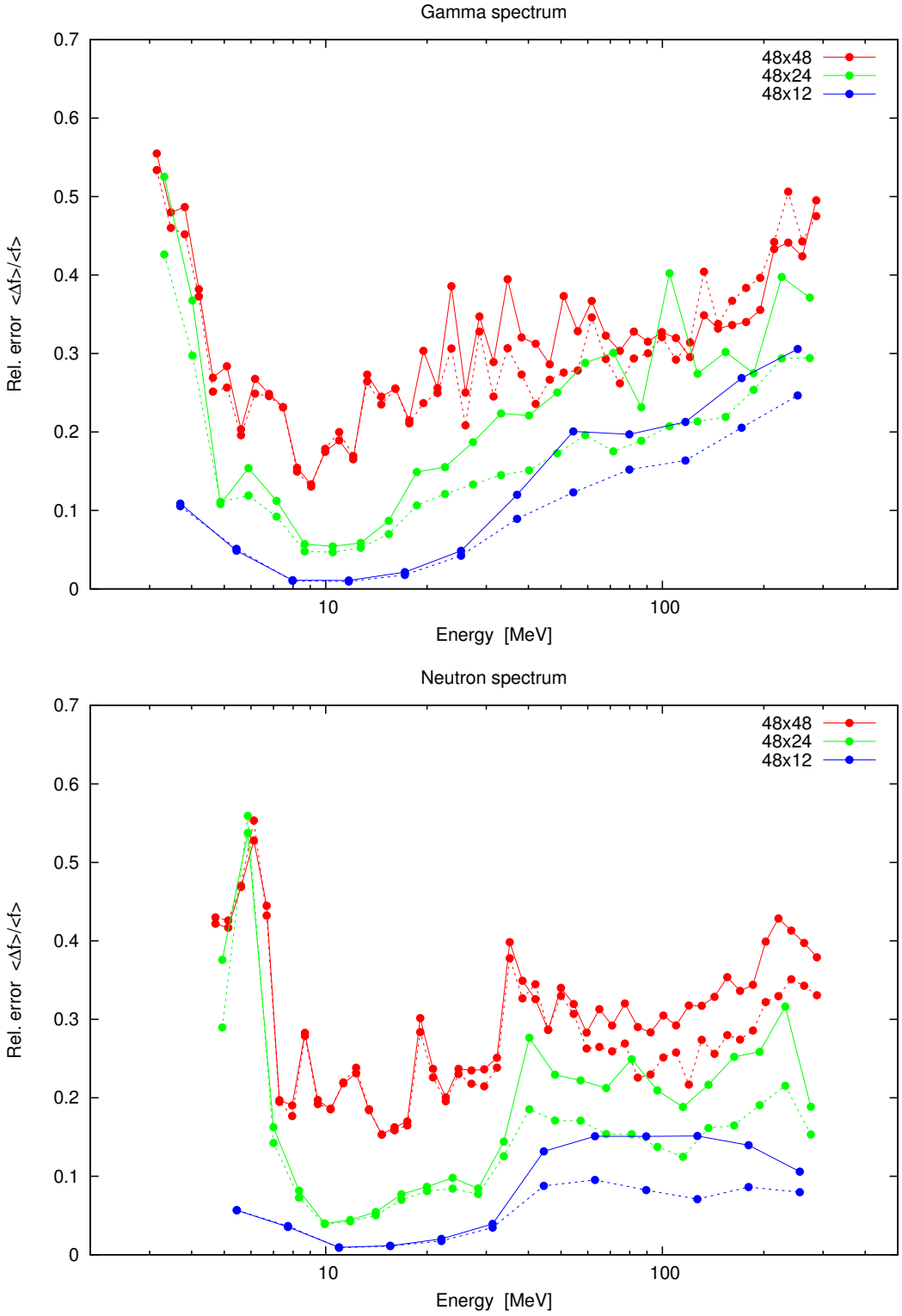

Figure 4.12: Relative errors for an inversion with underlying Poissonian statistics for estimated Martian gamma/neutron spectra with $N = 10^7$ MC generated events. Compared are the relative errors for inversions using only the 1D histograms (solid) and for inversions using the 1D and 2D histograms (dashed) for DRFs with 48, 24 and 12 energy bins.

For both neutron and gamma spectra, the relative errors decrease by a factor of ~ 0.3 for high energies and show a minor decrease at low energies. The minor decrease (at low energies) can be understood, considering that low energy gammas and neutrons have only a low probability of creating a signal in both detectors and therefore do not contribute to the 2D histogram. This can be seen in Fig. 4.11, where the low energy $\mathbf{A}_{\gamma/N,i}$ have no or very few entries. As shown in Tab. 4.3, the mean relative error decreases by 20% for the 48×12 and 48×24 matrices and by 10% for the 48×48 matrices. This improvement comes at the cost of an increased runtime. Depending on the number of stacked matrices, the runtime increases by a large factor (compare Tab. 4.3). For a flattened $\underline{\mathbf{A}}$, as described in Sec. 4.3.1, the runtime increases merely by a factor of $\sim$2. This effect is caused by the inefficient Python implementation of the stacked-matrix multiplication.

Matrix	**1D measurements**		**1D + 2D measurements**	
	mean rel. error	runtime	mean rel. error	runtime
48×12	0.107	1	0.079	45
48×24	0.212	1	0.161	65
48×48	0.309	1	0.289	82

Table 4.3: Mean relative errors and runtimes compared for the different DRFs as shown in Fig. 4.12. The runtime is given in units of runtime for 1D measurements.

4.4 Conclusions

In this chapter, the inversion approach, developed in the previous chapter, was applied to the expected Martian gamma/neutron spectra using the full DRF for RAD. A direct comparison of NNLS and Poisson method shows that even for high count rates the Poisson method produces significantly better results. Therefore, it was chosen for further evaluation. Since RAD only measures 15 minutes per hour, the expected measurements for 15, 30, 60, 120 minutes of continuous measurement (equaling 1, 2, 4, and 8 hours) were inverted. As expected, short measurement times result in large errors and useful results can only obtained for several hours of measurement. Best results, in terms of low relative errors, were obtained in the energy range of $\sim$10 - 50 MeV for neutrons and gammas.

One crucial factor for the neutral particle measurements is the influence of the RTG. Based on RTG measurements at INL, the expected count rate was estimated. With the assumption that the correct RTG measurements will be known, the real neutral particle measurement can be obtained simply by removing the RTG counts. To account for the Poisson nature of the measured RTG counts, this subtraction is realized by removing a Poisson generated RTG measurement from each bootstrap MC inversion. For the intensity obtained from the INL measurement, the relative errors increase by a fraction of less than 10%. Even for four times this intensity, the relative errors increase only by $\sim$ 25%. How-

ever, in order to apply this correction, RTG measurements with RTG and RAD mounted on MSL in their final configuration is required. Those measurements will be conducted before launch and during cruise phase.

In the last part of this chapter, an inversion approach, which includes multidimensional measurements, was presented. In addition to the 1D histograms of **D** and **E**, the 2D histograms (**D** over **E**) can be included in the inversion approach. Events with energy deposit in **D** and **E** occur at high energies, therefore the relative errors for those energies can be reduced by using the information of the 2D histograms. In a direct comparison of 1D and 1D+2D inversion, the mean relative error of the 1D+2D inversion decreases by $\sim$10-20%.

Chapter 5

Conclusions and outlook

As a first step towards a propagation model for SEPs towards Mars, the magnetic fluctuations in the solar wind were analyzed for wave amplitudes in dependence of their frequency and solar wind conditions. The observed exponential distribution of wave amplitudes shows that there is always a finite possibility for the occurrence of large amplitude waves. The effect of pitch-angle scattering for the observed wave amplitudes was investigated with an ab-initio model, where particles were scattered by a single wave pulse. Even for moderate amplitudes, strong pitch-angle scattering occurred, which can not be modeled by a diffusive approximation. The propagation of SEP particles was investigated for a complete magnetic field line from Sun to Mars. Even for this comparatively long particle propagation, the influence of one strong scattering event had an influence on the time-intensity profile and pitch-angle distribution of an SEP event.

One advantage of the developed ab-initio model is that it makes few assumptions and approximations about pitch-angle scattering or focusing. All processes and scattering events are calculated ab-initio. For a given magnetic field, the propagation of a single particle can be calculated via the Lorentz force with any order of accuracy. This comes however at the price of high computation times, especially compared to the focused transport equation. Another problem is the unknown magnetic field in solar events. Intensity and spectral shape of the magnetic fluctuation may be known, but not the complete distribution of fluctuation along the magnetic field line. A promising approach to include the effects of large amplitude waves is, e.g., extending the pitch-angle diffusion in the focused transport equation with jump processes, based on the observed behavior for pitch-angle scattering with large amplitude waves in this ab-initio model.

For the general concept of gamma/neutron measurements, featuring one high-density and one high-proton-content scintillator, an inversion approach, based on maximum likelihood estimations, was developed. The DRF for RAD, which is a crucial part of the inversion method, was calculated based on `GEANT4` simulation. The existing detailed model of the detector geometry by O. Kortmann, which is based on CAD data, was adapted and used as a base for further `GEANT4` simulations. The PhotonMC code was used to model light distribution in the scintillators and the readout of the photodiodes. In contrast to

the previously available simulation software, the detector model was implemented without any library dependences except for `GEANT4`, which simplifies further instrument modeling for the RAD team. One major modification of the simulation code is the implementation of the neutron cross sections in the physics list. Cross sections and interaction models for neutrons above 20 MeV are still being discussed, and will hopefully be determined by future calibration measurements.

Another issue for the DRF is the definition of the trigger thresholds for which no definite value has been selected yet. Since the trigger thresholds may change several times during the mission, Python scripts to generate DRFs for arbitrary energy ranges or different physics lists were made available.

The inversion of Martian gamma/neutron spectra, based on the `Planetocosmics` calculations by B. Ehresmann, was tested with artificial measurements. Best results were obtained for a maximum likelihood estimation based on Poissonian statistics. The results could be further improved by extending the inversion technique to include the neutral particle 2D histograms. This was achieved by developing an inversion approach for detector measurements with an arbitrary number of coincidence conditions, which can not only be applied to RAD but to various instruments which measure particles in multiple detectors in coincidence, such as the Kiel electron telescope.

RTG measurements from INL were used as an estimation for the RTG induced count rate during the mission. It could be shown that if the RTG induced count rate is accurately known, the RTG counts can be removed without any strong decrease of the inversion quality. The available INL measurements may give a good estimation of the RTG count rates. However, shielding by the MSL rover and scattering in the Martian soil will change the RTG induced count rate. RTG measurements with RTG and RAD both mounted on board the MSL rover will probably be performed before launch and during cruise phase. However, those measurements will not include the backscattered RTG-neutrons from the Martian soil and the cruise-phase measurements will contain GCR-induced gamma particles. A promising approach would be to calculate the influence of Martian soil and GCR-induced gammas with `GEANT4`, using a volume filled with Martian soil and a rudimentary model of the spacecraft, respectively.

For future data evaluation for the MSL/RAD team, the inversion method is implemented within a Python library. The library inverts a given measurement with a selectable DRF. It offers options such as the statistical model, step width of the numerical gradient, set of initial guesses and several optimization methods for the initial guesses.

Appendix A

Detailed description of the inversion approach

In this chapter, the methods for fitting gamma/neutron measurements are described in detail. The inversion methods developed in this work are implemented in the `inversion` library. They can be applied to an arbitrary merit function, but are optimized for the inversion of gamma/neutron measurements with the MSL/RAD instrument.

A.1 Minimizing a given merit function

To obtain the minimum of the merit functions, with the constraint of positive intensities ($f_i > 0 \forall f_i$), the SciPy [Jones *et al.*, 2001] implementation of the `L-BFGS-B` algorithm [Zhu *et al.*, 1997; Byrd *et al.*, 1994] is used. The `L-BFGS-B` algorithms is a version of the **B**royden-**F**letcher-**G**oldfarb-**S**hanno algorithm which is a quasi-Newtonian method for solving nonlinear minimization problems. It implements a **L**ow memory extension and the option of setting **B**ound constraints.

A.1.1 Numerical gradient

Since an analytical expression for the gradient of the merit function can often not be obtained, the `L-BFGS-B` algorithm automatically calculates a numerical approximation. Even for a small number of parameters to be minimized, and a very simple merit function with only one minimum, the algorithm may fail to find this minimum. For the numerical calculation of a derivative, a finite ϵ value needs to be selected (e. g. $\partial/\partial_x f(x) \approx \frac{f(x)-f(x-\epsilon)}{\epsilon}$). The SciPy implementation of the `L-BFGS-B` algorithm uses a default value of $\epsilon = 10^{-8}$. Especially for high count rates and a merit function given by the Poissonian ML estimation this value can be too low, which means numerical noise can be larger than the gradient.

The higher the count rates (large free parameters) the higher the probability that the numerical gradient is influenced by numerical noise. To address this problem the `gnSolver.solution` object provides the possibility to define multiple ϵ values $[\epsilon_1, \epsilon_2, ...\epsilon_n]$

with $\epsilon_i > \epsilon_{i+1}$, which will be used in the given order.

1. Obtain $\vec{f}_1$ from initial guess $\vec{f}_{init}$ using the `L-BFGS-B` with $\epsilon = \epsilon_1$
2. Obtain $\vec{f}_2$ from initial guess $\vec{f}_1$ using the `L-BFGS-B` with $\epsilon = \epsilon_2$

   ⋮

n. Obtain $\vec{f}_n$ from initial guess $\vec{f}_{n-1}$ using the `L-BFGS-B` with $\epsilon = \epsilon_n$

The solution $\vec{f}$ is given by $\vec{f}_n$.
`gnSolver.solution.solve(f_init)` invokes this procedure, the ϵ values can be set via `gnSolver.solution.epsilon = [1e-6, 1e-8]`.

For high count rates and a merit function given by the Poissonian ML estimation, a set of useful ϵ values, which was sufficient for all analyzed measurements, is given by $[10^{-2}, 10^{-4},\ 10^{-6},\ 10^{-8}]$. For the NNLS method, $\epsilon = 10^{-8}$ was always sufficient. Some analyzed merit functions with only 2 free parameters showed that the gradient for NNLS-merit function is usually much larger than for the Poisson-merit function, therefore the NNLS method did not need any adapted ϵ values to reduce the influence of numerical noise.

A.2 Providing initial guesses

In most cases, one has at least a vague idea of the shape of the incoming particle spectra, e.g. a powerlaw with some spectral index, or a neutron beam at some position. Since computational time becomes increasingly cheaper, minimizing a merit function for multiple initial guesses is no problem.

We define a set of appropriate initial guesses for the gamma $F_\gamma = \{f_{\gamma,1},\ ...f_{\gamma,n}\}$ and neutron spectra $F_N = \{f_{N,1},\ ...f_{N,m}\}$. The full set of initial guesses for the gamma/neutron spectra is then given by

$$F_{I_\gamma, I_N} = \{I_\gamma \cdot \vec{f}_\gamma \ \oplus\ I_N \cdot \vec{f}_N \mid \forall\ \vec{f}_\gamma \in F_\gamma,\ \vec{f}_N \in F_N\}, \tag{A.1}$$

where $I_\gamma,\ I_N$ are the adjustable intensities of the gamma and neutron spectra. Spectra can be generated by extending the `gnSpectra.spectra` class, ore are already available as predefined objects in `gnSpectra`.

A.3 Finding the optimal solution for one initial guess

For one initial guess $\vec{f}_{I_\gamma, I_N} \in F_{I_\gamma, I_N}$ the optimal solution is calculated in several steps.

1. The initial intensities $I_\gamma,\ I_N$ are obtained from the solution of the initial guess given by Eq. 3.15. The solution $\vec{f}$ is obtained as described in Sec. A.1.1. $I_\gamma,\ I_N$ are then used for the initial guess $\vec{f}_{I_\gamma, I_N}$.

2. The solution $\vec{f}$ for the initial guess $\vec{f}_{I_\gamma, I_N}$ is obtained as described in Sec. A.1.1.

3. The solution provides corrected intensities $I_\gamma,\ I_N$ for the next initial guess

4. Proceed from step 2 until the solution of the merit function did not improve for $m > 0$ iterations (i.e. no smaller minimum of the merit function was found for m iterations).

5. Optionally, try to improve the result starting from step 2 with sets of random values for $I_\gamma,\ I_N$.

Matching the intensities in such an iterative way may seem overly complicated at first. However, one hast to consider, that every entry of $\vec{f}_{I_\gamma, I_N}$ is a free parameter in the optimization process. Even if the correct shape for the initial guess is given, but false intensities are assumed, the resulting $\vec{f}$ will differ completely from the initial guess and the correct solution.

A.4 Finding the solution for a given set of initial guesses

The solution strategy for one initial guess is explained in Sec. A.3. Using the set of initial guesses F_{I_γ, I_N}, or any reasonable subset, this algorithms is applied for each initial guess. The best solution is then chosen in terms of the smallest value of the merit function. The following procedures are called by executing the call function of the `gnSolver.iterGuess` object.

A.4.1 Merging of initial guesses

Eq. A.1 may contain a large number of shapes, but considering the computation time, this number should be constrained to a reasonable amount. This means, that more complex shapes, e.g. a combination of several powerlaws or several neutron beams at different energies, are general not feasible to use in Eq. A.1.

For the example of two neutron beams at different energies, a reasonable set of initial guesses would be a set of single neutron beams with different energies. None of the initial guesses would be very close to the input spectrum, and the best solution of all initial guesses is often far away from the input spectrum. However, often the best solution shows some distinct features of the input spectrum. For the example of two neutron beams, the best solution is often obtained by one initial guess with a beam at one of the correct positions. The corresponding solution may show on very clear beam and one smeared out beam, both at the correct position, but only one beam with the correct shape. Having an idea about the input spectrum (some neutron beams) this effect can be used to generate a new, merged, initial guess from the original set of initial guesses.

To select the best m candidates for an initial guess, the initial guesses with the highest similarity with the best solution are chosen. The similarity between $\vec{f}$ and $\vec{f}_{I_\gamma, I_N}$ is

calculated via

$$a(\vec{f}, \vec{f}_{I_\gamma,I_N}) = \frac{\vec{f} \cdot \vec{f}_{I_\gamma,I_N}}{|\vec{f}| \cdot |\vec{f}_{I_\gamma,I_N}|}.$$

Choosing the m initial guesses with the highest similarity ($\vec{f}_{I_\gamma,I_N,1}, \vec{f}_{I_\gamma,I_N,2}, ...\vec{f}_{I_\gamma,I_N,m}$), $m-1$ new initial guesses are generated:

$$\vec{f}_1 = a(\vec{f}, \vec{f}_{I_\gamma,I_N,1})\vec{f}_{I_\gamma,I_N,1} + a(\vec{f}, \vec{f}_{I_\gamma,I_N,2})\vec{f}_{I_\gamma,I_N,2} \tag{A.2}$$

$$\vdots$$

$$\vec{f}_{m-1} = a(\vec{f}, \vec{f}_{I_\gamma,I_N,1})\vec{f}_{I_\gamma,I_N,1} + \ldots + a(\vec{f}, \vec{f}_{I_\gamma,I_N,m})\vec{f}_{I_\gamma,I_N,m} \tag{A.3}$$

For each merged initial guess, the solution is calculated via Sec. A.3. If the previous best solution can be improved by the merged initial guesses, a new best solution is selected.

This approach can significantly improve results for some spectra (e.g. two neutron beams), but increases computation time for simple spectral shapes. Therefore, it can optionally be enabled via `iterGuess.merge=True`, the number of initial guesses to merge (2, 3, ... m) can be set via `iterGuess.nrRerge= [2, 3, 4, 5]`.

A.4.2 Using randomized solutions as initial guesses

Assuming that a good solution, somewhere near the optimal solution, has already been found, further improvement can often be achieved by using randomized initial guesses. The randomization is achieved by

$$f_{i,\text{init}} = f_i(1 + c\xi), \tag{A.4}$$

where $0 < c < 1$, ξ is a uniform distributed random variable in the interval $[-1, 1]$, and f_i are the components of the best solution $\vec{f}$. For a set of n values $[c_1, \ldots c_n]$ with $c_i > c_{i+1}$, new initial guesses are generated and corresponding solutions are calculated via Sec. A.1.1

1. Try m times to improve the best solution with initial guess from Eq. A.4 with $c = c_1$
2. Try m times to improve the best solution with initial guess from Eq. A.4 with $c = c_2$

   $\vdots$

n. Try m times to improve the best solution with initial guess from Eq. A.4 with $c = c_n$

This procedure improves the best solution especially for inaccurate initial guesses. m and $[c_1, \ldots c_n]$ can be adjusted according to the situation. Good results were often obtained for $m = 10$ and the c values [1, 0.5, 0.1]. This method can be enabled via `iterGuess.Randomize=True`. m and the c values can be set via `iterGuess.nrRand=10` and `iterGuess.randScale = [1, 0.5, 0.1]`.

A.4.3 Removing extrema

To remove the occurrence of alternating extreme minima and maxima, as described in Sec. 3.2.6, new initial guesses can be calculated from the best solution via Eq. 3.18.

Extrema minima are defined as $f_{i-1} > f_i/a < f_{i+1}$. A list of a values can be set via `iterGuess.invertMinimaScale = [0.1, 0.01]`. For each given a a new initial guess is generated and the corresponding solution is calculated as described in Sec. A.1.1. If the previous best solution can be improved by the new initial guesses, a new best solution is selected. Good results were obtained for the a values [0.1, 0.01, 0.001]. This method can be enabled via `iterGuess.invertMinima = True`.

Appendix B

Determining the optimal detector response function

For the success of the inversion procedure, the choice of the number of measurement and input-spectra bins is a crucial factor. The measurement events are processed onboard the MSL rover, the neutral particle histograms are sent back as histograms with 48 log-spaced bins for the **D** and **E** detector [Southwest Research, 2010]. The total number of measurement bins is therefore given by 96. Another important aspect is the range in which the bins are distributed, e.g. the smallest measurement bin should be above the noise peak, columns in **A** with only 0 entries should be avoided.

In this chapter a general description of the neutral particle measurement for RAD is given, the boundaries for the measurement bins $[z_{\mathrm{min}},\ z_{\mathrm{max}}]$ are set to satisfy the constraints of the onboard electronics, and the range of the input spectra $[E_{\mathrm{min}},\ E_{\mathrm{max}}]$ is determined.

B.1 Neutral particle measurements

Both the **D** and **E** detector are read out by three photodiodes with different gain factors, low, medium, and high. Each diode represents two measurement channels with two different gain factors, which sums up to six readout channels for both **D** and the **E**: **L**, **N**, **M**, **I**, **H**, **U** (increasing gain from left to right). To obtain one measurement value from the energy calibrated channels of one photodiode, a weighted mean is calculated [Böttcher, 2008]. If both channels of a photodiode are not in over or underflow, i.e. neither near the noise peak nor above the maximum ADC value, a weighted mean is calculated via

$$E_{sum} \quad = \quad E_h + k \cdot (E_l - E_h), \tag{B.1}$$

where E_h is the channel with the higher gain, E_l is the channel with lower gain and $k = 0.0,\ 0.125,\ 0.25,\ 0.5$. If one channel is in over/underflow, E_{sum} is given by the other channel.

B.1.1 Silicon hit rejection

In a next step the E_{sum} values for the different photodiodes are compared and a new E_{sum} is calculated, first low-gain with medium-gain photodiode, then for the resulting E_{sum} and the high-gain photodiode (see Fig. B.1). If one photodiode is in over/underflow, E_{sum} is given by the other photodiode. If one channel has a significantly higher energy than the other one, it is probably a silicon hit, i.e. an event were one of the photodiodes has been hit directly. This is tested with the normalized difference of the E_{sum} values

$$N = 64 \cdot \frac{E_l - E_h}{E_l - E_h}, \tag{B.2}$$

where E_l is the lower gain E_{sum} value and E_h is the higher gain E_{sum} value. The result N is compared to a threshold value. If the one of the photodiodes shows a silicon hit, the other one is selected, otherwise the resulting E_{sum} value is calculated via Eq. B.1.

Low energies may not be above the noise peak of the low-gain channels, but will be calculated in the high-gain channels. High energies may be in overflow in the high-gain channels, but can be processed in the low-gain channels.

B.1.2 Calibration values

The final calibration values for **D** and **E** are still to be decided. Therefore, the noise peak positions were determined individual for each run. For a preliminary energy calibration, the values

$$\begin{aligned} \mathbf{DU} &= 2.4 \cdot 10^{-3}\mathrm{MeV/ADC} \\ \mathbf{EU} &= 6.9 \cdot 10^{-3}\mathrm{MeV/ADC} \end{aligned}$$

were used. The relative gains for the other **D**, **E** channels and other calibration values were taken from Martin [2008].

B.1.3 Anticoincidence thresholds

The basic condition for neutral particle detection is, that the there is no energy deposit in the surrounding AC. Events which fulfill the condition

$$\neg(\mathbf{F} \wedge \mathbf{C} \wedge \mathbf{C2})$$

are accepted as neutral particle events, where **C** and **C2** are combined segments of the **B** and **C** silicon detector, see Kortmann [2010] (Appendix H). Since the measurements are influenced by noise, an appropriate cut condition, a threshold, has to be selected. If the cut threshold is selected to low, noise will lead to a large fraction of falsely rejected neutral particles - false negatives. If the cut threshold is selected to high, energy deposits

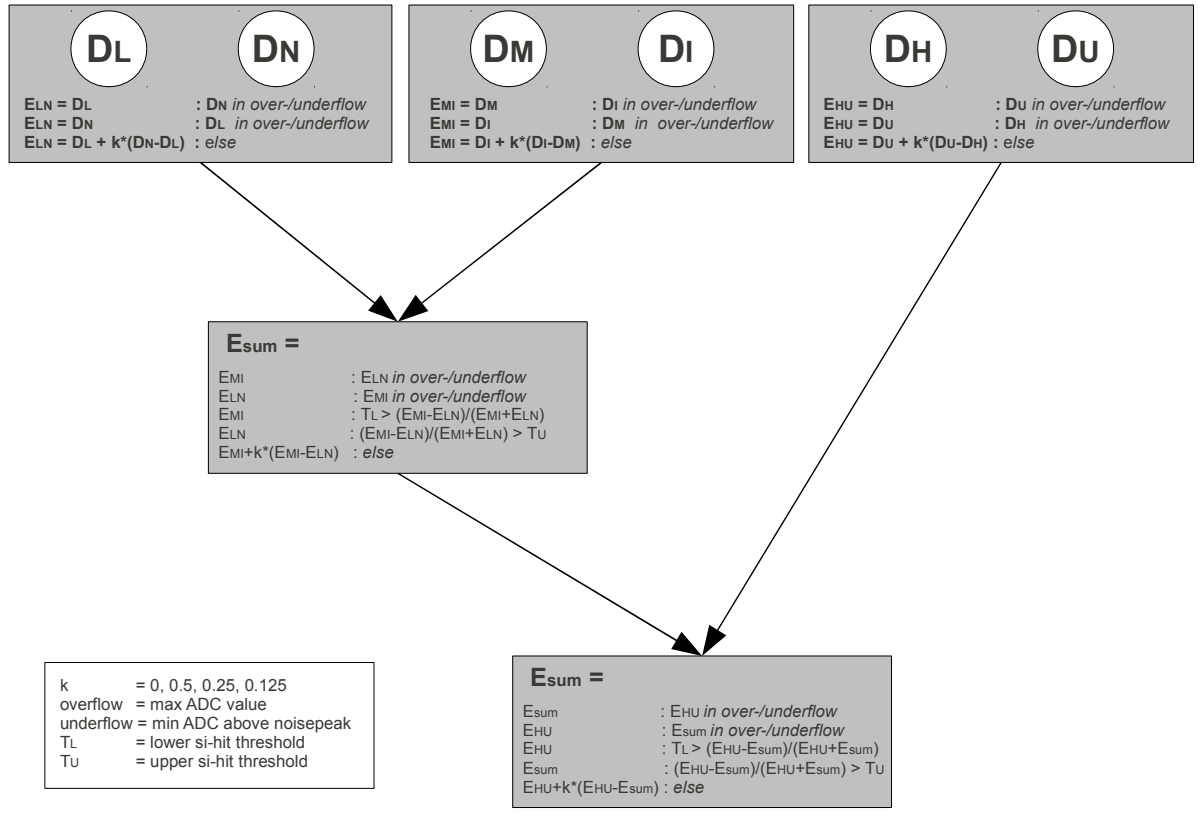

Figure B.1: E_{sum} calculation for the **D** detector with applied silicon hit detection. Each of the three photodiodes has two readout channels, which are grouped as **DL-DN**, **DM-DI** and **EH-EU**. For each photodiode a weighted mean of its two readout channels is calculated. Beginning with **DL-DN**, **DM-DI** the photodiodes are checked for silicon hits and a weighted mean is calculated. This weighted mean is than compared to **DH-DU**, checked for silicon hits, and the final weighted mean is calculated.

of charged particles may not be detected and a large fraction of charged particles are falsely accepted as neutrals - false positives.

While the noise peak in the silicon channels **C**, **C2** is clearly separated from any minimally ionizing particles, the noise is much more dominant in the **F** scintillators. At this point, the RAD software supports two different options for **F**. Setting cut values for **F1**, **F2**, and **F=F1+F2** above the noise peaks, or setting the AC threshold for **F** to zero. Although the latter option results in a false negative rate of 50% it has some promising advantages [Kortmann, 2010]. For example, a 0 keV threshold results in very few false positives and *exactly* 50% false negatives. In contrast to trigger thresholds above 0 keV, the rate of false negatives will not change if the noise increases or decreases due to temperature or other effects (e.g. see Appendix C.5).

Although the 0 keV threshold is more likely to be applied, the DRFs will be generated for a threshold > 0 keV. Note that in the `GEANT4` model the AC is not influenced by noise, and is assumed to operate with false negative rate of 0%. Therefore, the DRF for the 0 keV threshold can be obtained by a scaling **A** with 0.5.

B.1.4 Trigger thresholds for D and E

Unfortunately the choice of the trigger thresholds E^D_{min} and E^D_{min}, which also determine the minimum energy in the neutral particle histograms, can not be determined by considerations of the Martian neutral particle spectra, but are determined by restrictions of the detector electronics. Considering that RAD can process 1 - 2 kHz of events and most of the occurring events will be due to charged particles, the neutral particle channels are required to trigger less than 100 - 200 events per second. In addition to the Martian neutral particle spectra, the RTG contributes to a large part of the count rate as well.

An estimation of the RTG count rates is obtained from measurements at INL, where RAD was placed in a distance of 1.0 m from the RTG. During the measurement, various shieldings were tried, as well as an undefined larger distance between RTG and RAD. In contrast to the Ulysses measurements [Hawley, 1984], the RTG was not placed above a pit, therefore a lot of backscattered particles are to be expected. Considering the unknown shielding on board the MSL rover, the measurements give a reasonable estimation, but by far no exact prediction, of the expected RTG spectra. Fig. B.2 shows the count rates for various shieldings and distances between RAD and RTG. One can clearly see the increased count rate for the runs without shielding, as well as an inverse relationship between count rate and distance. For a trigger threshold of 2 MeV, the **D** and **E** detector would count $\sim$ 200 and 30 particle/s only from the RTG.

To determine the optimal values for E^D_{min}, E^D_{min}, the expected Martian gamma and neutron field is simulated with `GEANT4` and the expected count rates are calculated. The expected Martian neutral particle spectra are taken from Ehresmann [2012], where the radiation in the Martian atmosphere, induced by GCR protons and helium, is calculated for several altitudes and atmospheric densities with `Planetocosmics`. As an estimation for RAD, the gamma/neutron spectra at 0 km altitude during solar minimum condition are selected. Fig. 4.1 shows the omnidirectional flux for neutrons and gammas, i.e. the sum of

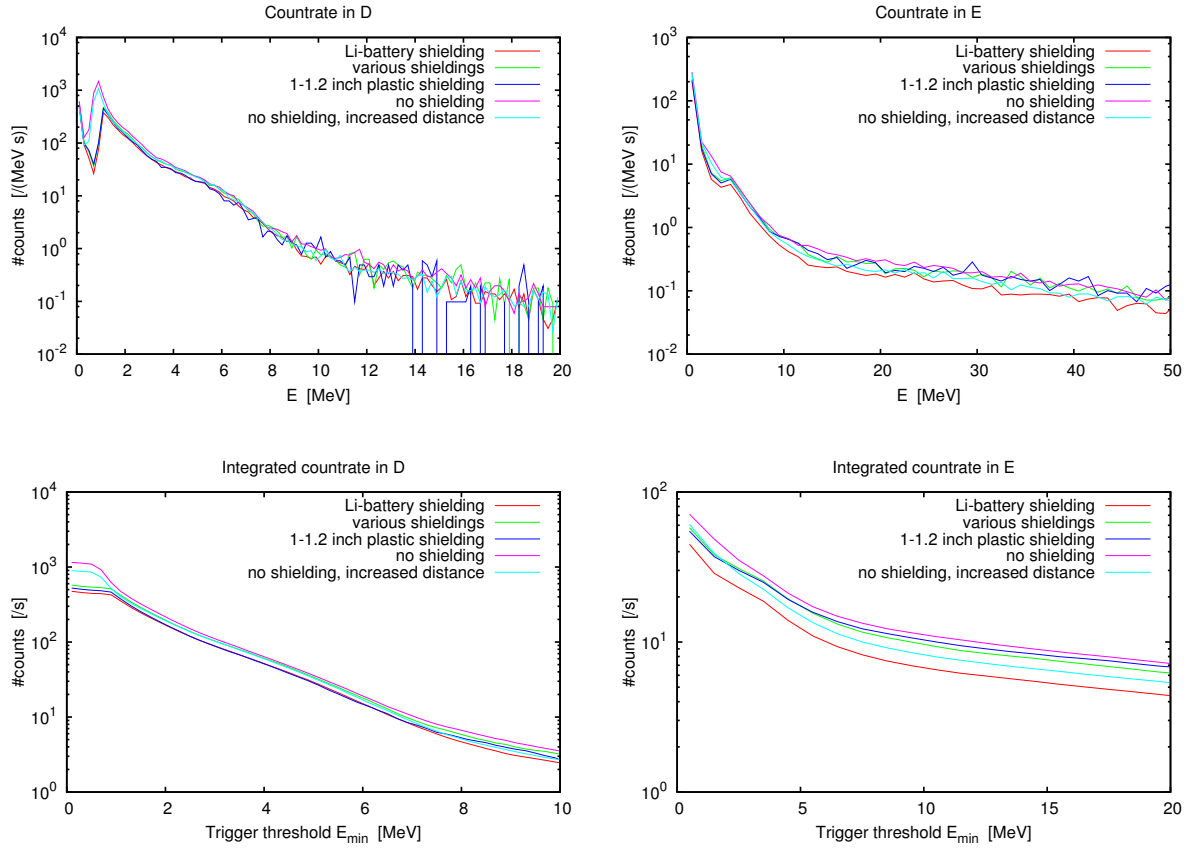

Figure B.2: RTG count rate in **D** (left) and **E** (right) for the measurements `inl-edc-36-(8,9,10,11,12)` (red to cyan). Top: the count rate [/(MeV s)]. Bottom: the expected integrated count rate in dependence of the trigger threshold $E_{\min}$. Correction for silicon hits or an AC was not applied. One should note, that different (lower) trigger thresholds were used in run 11 and 12, which results in an strongly increased count rates at low energies.

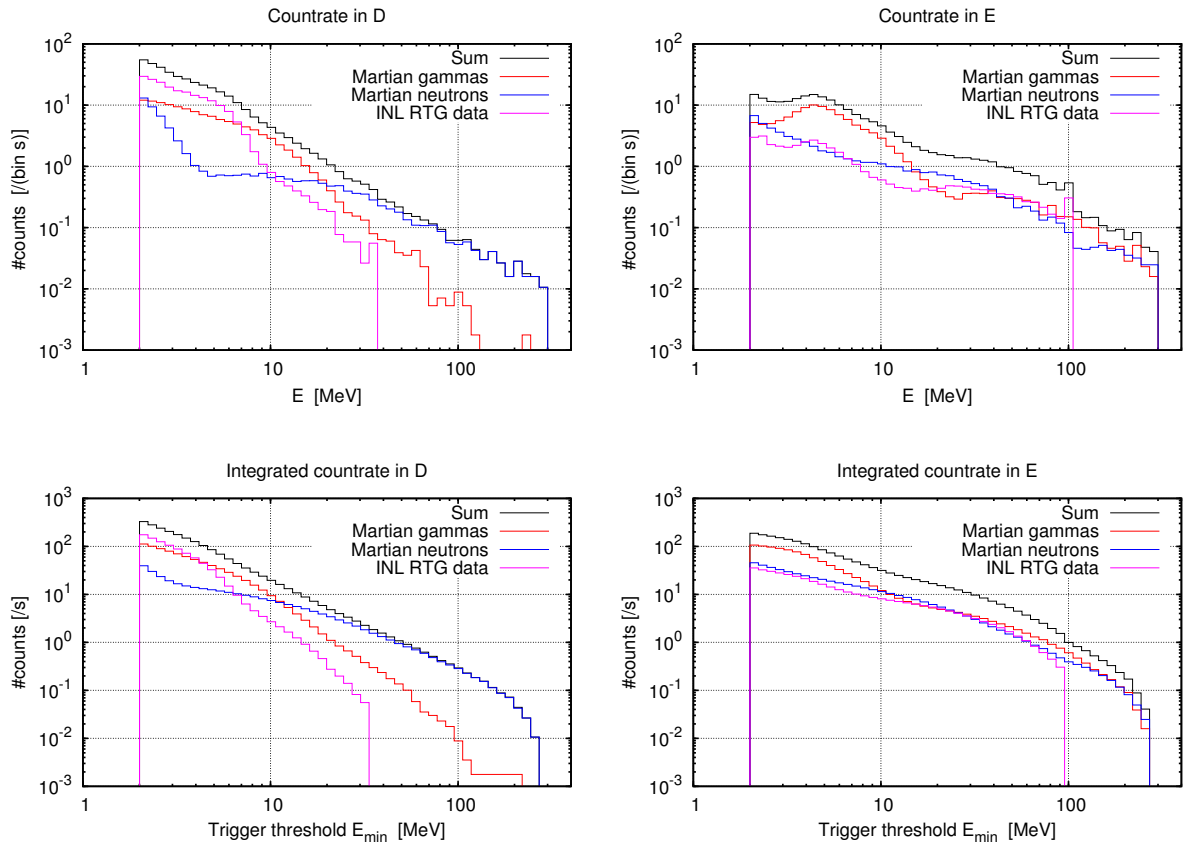

Figure B.3: Expected neutral particle count rate for the **D** (left) and **E** (right) detector. Shown are `GEANT4` calculated Martian gamma (red) and neutron (blue) counts, as well as RTG count rate from the INL measurements. Correction for silicon hits and AC was not applied, the corrected measurements are shown in Fig. B.4. The large neutron count rate at low energies in **D** is due to low energy neutrons (See Fig. B.6).

upward and downward directed flux without any consideration of angular dependence, for the full `Planetocosmics`-calculated energy range.

The simulated Martian and the measured RTG count rates are shown in Fig. B.3. The measurement histograms each consist of 48 log spaced bins, i.e. the format defined in Southwest Research [2010]. For an estimation of the measured and accepted events, the simulated Martian and the measured RTG count rates, are shown in Fig. B.4, with silicon hit rejection and AC applied. For trigger thresholds at 4 MeV, both detectors are in the desired range of $\sim$ 100 counts/s. However, this includes only the counts from Martian and RTG neutral particles. Charged particles, which are detected as false positives, are not considered.

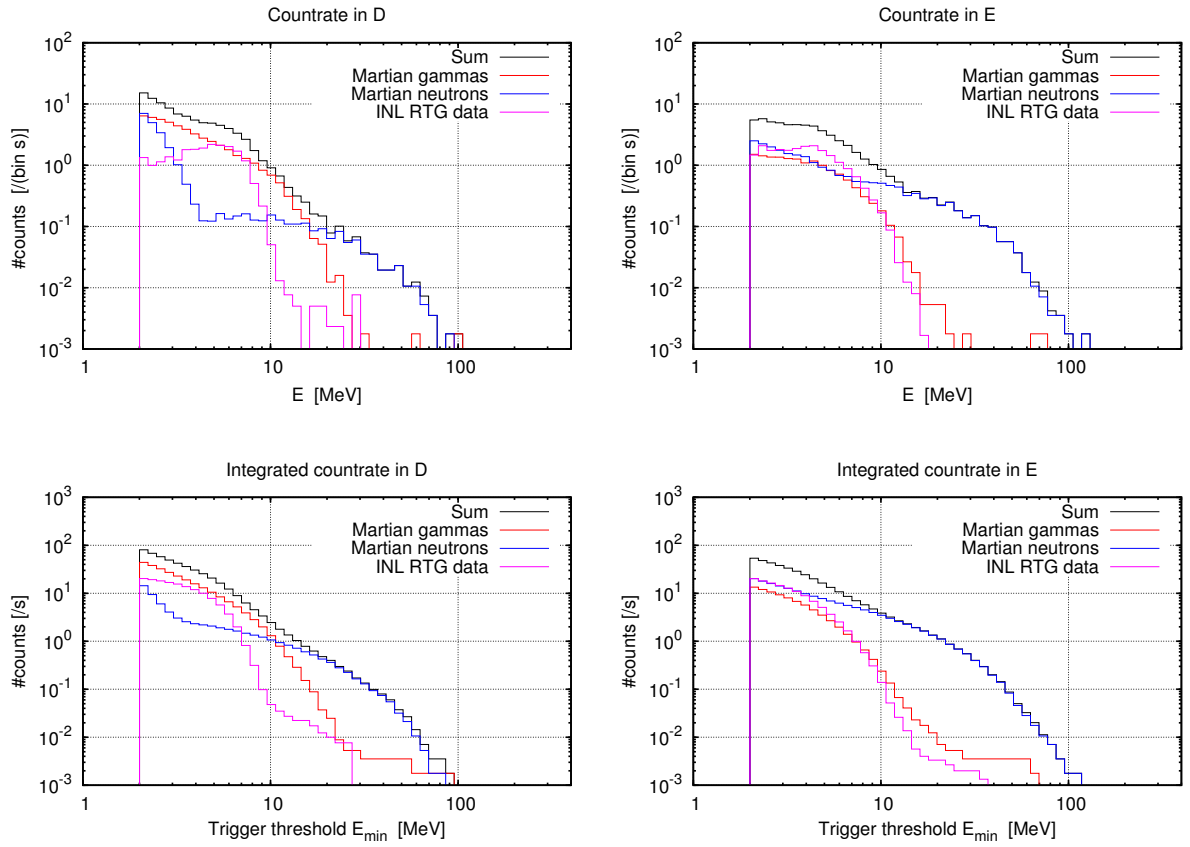

Figure B.4: Expected neutral particle count rate for the **D** (left) and **E** (right) detector. Shown are `GEANT4` calculated Martian gamma (red) and neutron (blue) spectra, as well as RTG count rate from the INL measurements. For `GEANT4` simulated Martian spectra, silicon hits were simulated for each photodiode and then filtered via condition Eq. B.1. For the RTG measurements, silicon hits were filtered via condition Eq. B.1, but since the medium-gain photodiode for **D** was defective, with poor results. AC thresholds were set to 20 keV for **C**, **C2** and to $3\sigma_{\text{noise}}$ for **F2**. **F1** was disabled for the RTG measurement, therefore, charged particles can not be rejected and the real count rate might be decreased by some factor.

B.1.5 Neutral particle histogram ranges

Because of restrictions of the signal processing, reasonable values for the log-spaced histograms are 8 or 4 bins per octave. For 48 bins this results in 6 and 12 octaves. A given $E_{\min}$ defines therefore $E_{\max} = E_{\min} \cdot 2^6$ and $E_{\max} = E_{\min} \cdot 2^{12}$. Keeping in mind that the energy deposit of gammas and neutrons saturates with increasing primary particle energy (see Fig. B.5), $E_{\max} = E_{\min} \cdot 64 = 256$ MeV is sufficient to contain the major part of all occurring neutral particle events, compare expected count rate in Fig. B.4 and energy deposit over primary particle energy in Fig. B.5. For both neutral particle histograms the boundaries are chosen to be

$$\begin{aligned} E_{\min} &= 4 \text{ MeV} \\ E_{\max} &= 256 \text{ MeV}. \end{aligned}$$

In addition to the 1D histograms, a 2D histogram of the **D** and **E** channel is defined. The 2D histogram counts neutral particles, which deposit energy in **D** and **E**, e.g. a neutron creates a recoil proton in **E** which does not stop in **E** and enters detector **D**. The histogram consists of 8×8 log spaced bins, with one axis the energy deposit in **D**, the other axis the energy deposit in **E**. $E_{\min}$ is again given by the trigger threshold of **D** and **E** detector, $E_{\max}$ depends on the number of bins / octave. For 8, 4, 2, 1 bins / octave $E_{\max}$ is given by 8, 16, 64, 1024 MeV. Fig. B.7 shows a **D** over **E** scatter plot for both neutrons and gammas. Comparing the deposited energies with the possible $E_{\max}$ values, $E_{\max} = 64$ MeV is selected.

B.2 Full detector response function for RAD

In the previous section the ranges of the neutral particle histograms was determined. The range and binning of the measurement $\vec{z} = \vec{z}_D \oplus \vec{z}_E$ is therefore given. The remaining free parameter are the range and binning of the input spectrum $\vec{f} = \vec{f}_\gamma \oplus \vec{f}_N$. Together $\vec{f}$ and $\vec{z}$ unambiguously determine the range and binning of **A**.

Mathematically, as long as **A** is invertible, the exact shape of **A** is of minor importance. However, considering singular matrices and the effect of measurement errors and Poissonian statistics, the shape of **A** needs to be carefully determined. A general guideline should be, that the closer the matrix is to diagonal shape, the more easy an inversion becomes. For the RAD instrument, a diagonal shape con not be achieved, but the DRF is close to a triangular shape. Empty columns in **A** ($\mathbf{A}_{i,j} = 0,\ \forall i$ for a given j) should be avoided, otherwise the corresponding bin of the input spectrum f_i can not be minimized.

Beside those general guidelines, the DRF influences the inversion in many unforeseeable ways. Therefore, based on the guidelines described above, a large set of different matrices is generated and for each DRF several inversion attempts of the same spectrum are performed. The one with the best results is then chosen for further studies. The geometric factor matrices are calculated from a `GEANT4` simulation with 10^8 gammas and neutrons with isotropically distributed positions and a uniform distribution of energies. An AC is applied

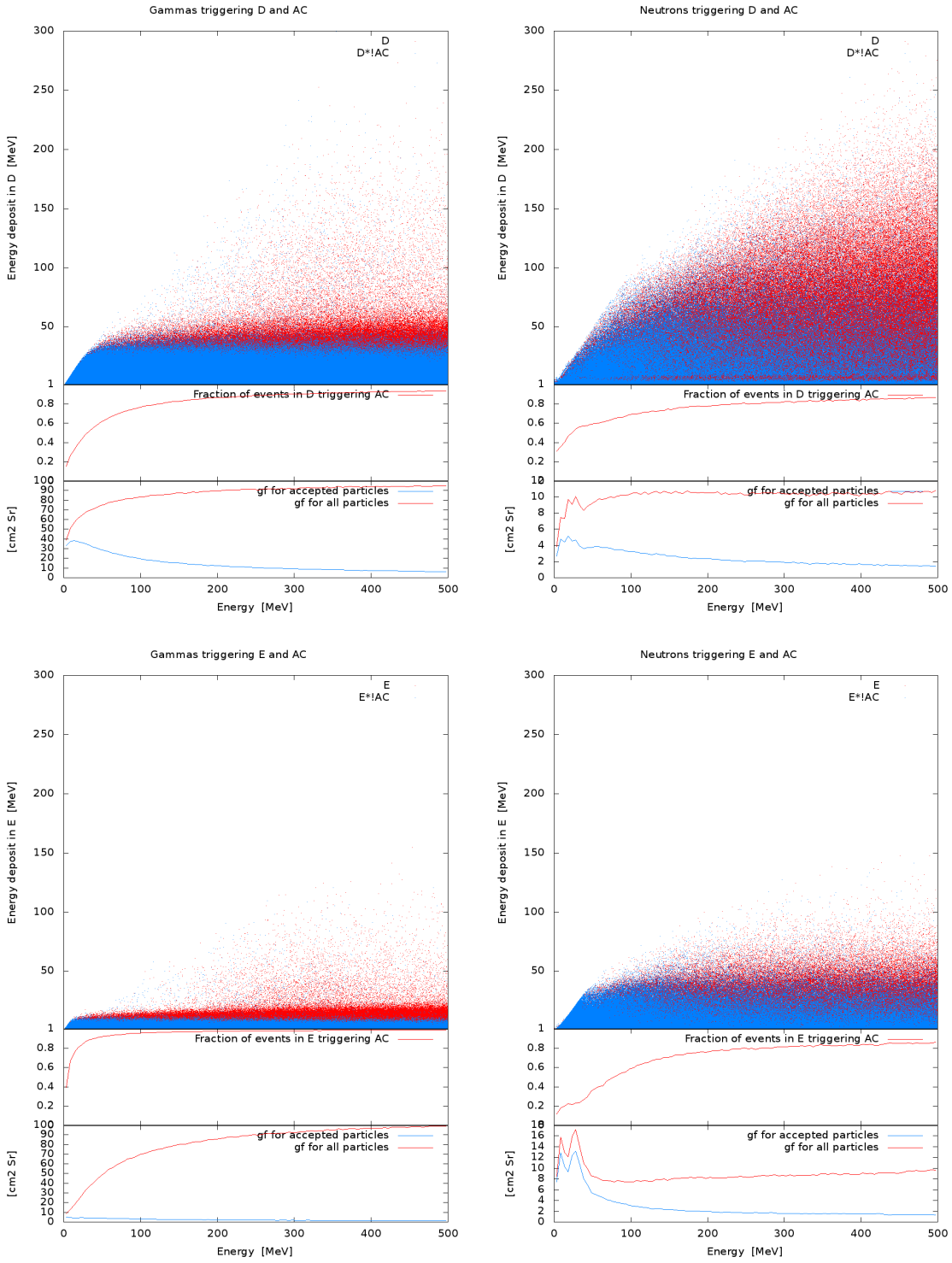

Figure B.5: GEANT4 simulations of the response of the **D** (top) and **E** (bottom) detector to gammas (left) and neutrons (right). Each plot consists of three subplots, beginning with a scatterplot of primary particle energy versus deposited energy (all events (red) and events passing the AC (blue)), followed by the fraction of rejected particles plotted over primary particle energy, and the geometric factor for all detected events (red) and for events passing the AC (blue). Fig. B.6 shows the same figures for low-energy neutrons.

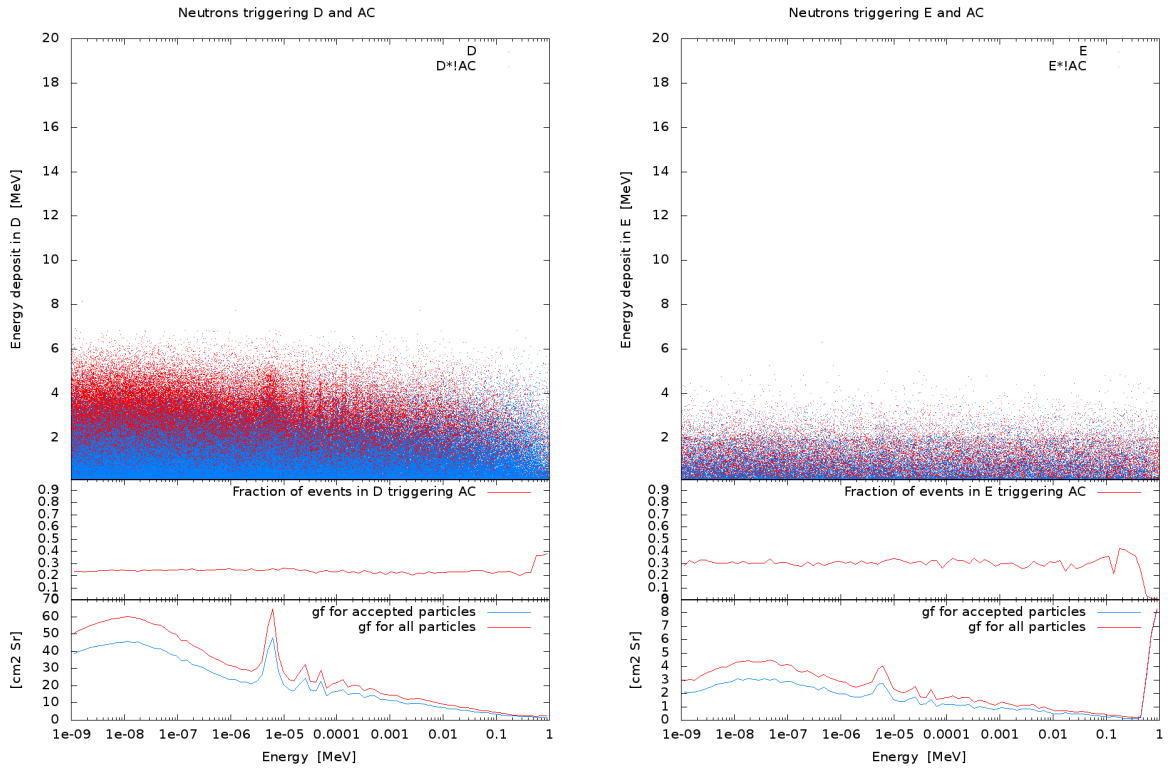

Figure B.6: Same as in Fig. B.5 for low energy neutrons. The most frequent process causing the high energy deposits is the neutron capture $n + ^{133}\mathrm{Cs} \rightarrow ^{134}\mathrm{Cs}$, where the excited nucleus emits several gammas. The half-life of ^{134}Cs is given by $t_{1/2} = 2.0648$ a, and therefore contributes significantly to the count rate in **D**.

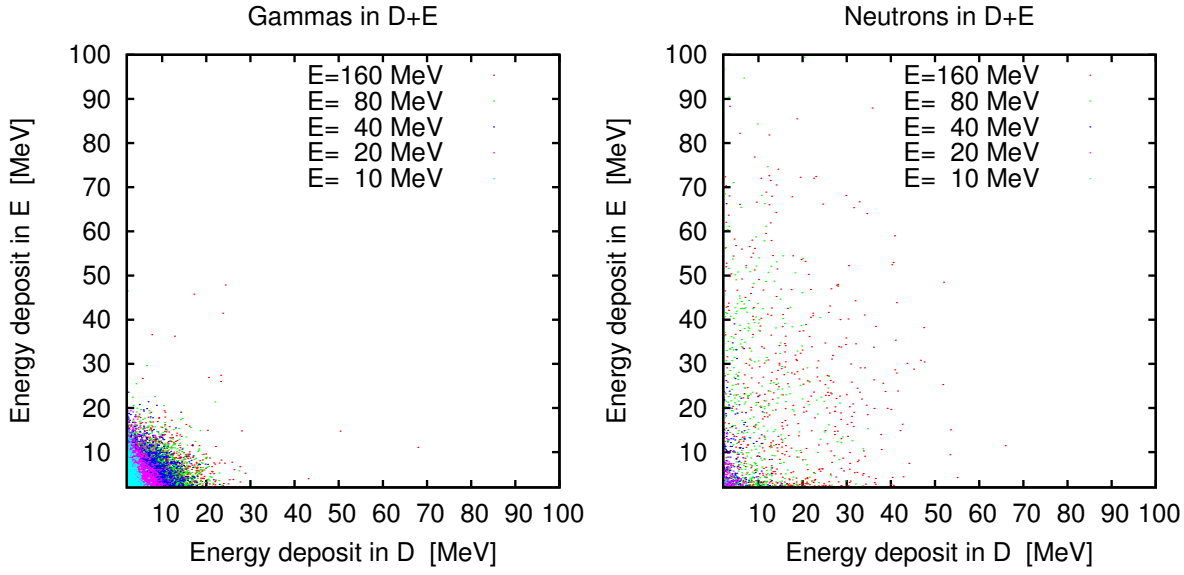

Figure B.7: `GEANT4` simulations of energy deposit in **D** over **E** for gammas and neutrons with isotropic distribution, exemplary shown for several energies.

as described above, the deposited energy is calculated from the weighted mean of the photodiodes. The input spectra $\vec{f_\gamma}$, $\vec{f_N}$ are binned logarithmically with each 24 bins. The minimal ($f_{\gamma,1}$, $f_{N,1}$) and maximal ($f_{\gamma,n}$, $f_{N,n}$) energy of the gamma and neutron spectrum is varied, to find an optimal $\mathbf{A}$. For each DRF, 100 inversions for a MC measurement of the expected Martian input spectrum are done. Errors are determined via bootstrap MC, i.e. each generated measurement is inverted 10 times. To measure the inversion quality, the mean relative error is calculated. In addition, to get a more detailed impression, the minimal and the maximal relative error are shown as well.

An overview of the evaluated matrices is shown in Tab. B.1. The best values for the minimal energies were found to be $f_{\gamma,1} = 3$ MeV, $f_{N,1} = 4.5$ MeV. For the maximal energy $f_{\gamma/N,n} = 300$ MeV was selected. A similarly small relative error was also obtained for $f_{N,1} = 5.5$ MeV. Nevertheless, $f_{N,1} = 4.5$ MeV was chosen because it yields the larger energy range.

For $f_{\gamma,1} = 3$ MeV, $f_{N,1} = 4.5$ MeV the low energy part of $\mathbf{A}_{E,\gamma}$ and $\mathbf{A}_{E,N}$ are very close to triangular shape, for larger values of $f_{\gamma,1}$, $f_{N,1}$ the triangular shape is shifted (see Fig. B.8). Comparing the maximal energies 300 and 500 MeV in Fig. B.8, one can see that for 300 MeV the submatrices have a (deformed) triangular shape up to high input-spectrum energies, while for 500 MeV the slope of the triangular shape decreases at high input-spectrum energies. This decreased slope means, that for high energies the detector response to an incident particle looks more or less the same, therefore one can not distinguish between different particle energies. Ignoring input energies above 300 MeV seems acceptable, if one considers the neutron and gamma spectra and the geometry factor for high energies (Fig B.5). I.e., the neutral particles in energy interval 300 - 500 MeV contribute with less than 2% to the count rate.

$f_{\gamma,1}$	$f_{N,1}$	$f_{\gamma,n}$	$f_{N,n}$	mean rel. err.	std rel. err.	min rel err.	max rel. err.
3.0	3.0	300	300	0.48	0.24	0.11	1.11
3.0	3.5	300	300	0.47	0.23	0.10	0.89
3.0	4.0	300	300	0.47	0.23	0.10	0.91
3.0	4.5	300	300	**0.46**	0.22	0.10	0.91
3.0	5.0	300	300	0.47	0.22	0.10	0.99
3.0	5.5	300	300	0.46	0.22	0.10	0.88
3.5	3.0	300	300	0.48	0.26	0.10	1.14
3.5	3.5	300	300	0.48	0.25	0.10	0.96
3.5	4.0	300	300	0.48	0.24	0.11	0.91
3.5	4.5	300	300	0.47	0.23	0.11	0.94
3.5	5.0	300	300	0.47	0.23	0.11	0.98
3.5	5.5	300	300	0.48	0.25	0.11	0.97
4.0	3.0	300	300	0.50	0.26	0.11	1.24
4.0	3.5	300	300	0.49	0.26	0.11	0.99
4.0	4.0	300	300	0.48	0.24	0.10	0.93
4.0	4.5	300	300	0.48	0.24	0.12	0.92
4.0	5.0	300	300	0.48	0.24	0.11	0.98
4.0	5.5	300	300	0.48	0.25	0.12	0.95
4.5	3.0	300	300	0.50	0.30	0.09	1.63
4.5	3.5	300	300	0.50	0.26	0.10	0.96
4.5	4.0	300	300	0.49	0.25	0.11	0.98
4.5	4.5	300	300	0.49	0.25	0.10	0.96
4.5	5.0	300	300	0.49	0.25	0.10	1.00
4.5	5.5	300	300	0.48	0.25	0.10	0.96
5.0	3.0	300	300	0.52	0.29	0.16	1.51
5.0	3.5	300	300	0.50	0.26	0.14	1.09
5.0	4.0	300	300	0.51	0.25	0.17	0.96
5.0	4.5	300	300	0.49	0.24	0.14	0.96
5.0	5.0	300	300	0.49	0.24	0.16	0.93
5.0	5.5	300	300	0.49	0.26	0.13	0.97
5.5	3.0	300	300	0.52	0.31	0.13	1.79
5.5	3.5	300	300	0.49	0.26	0.13	0.97
5.5	4.0	300	300	0.50	0.25	0.13	0.95
5.5	4.5	300	300	0.48	0.25	0.13	0.95
5.5	5.0	300	300	0.49	0.25	0.13	1.02
5.5	5.5	300	300	0.49	0.26	0.13	0.98
3.0	3.0	400	400	0.49	0.27	0.07	1.33
3.0	3.5	400	400	0.50	0.26	0.08	0.98
3.0	4.0	400	400	0.49	0.26	0.08	0.95
3.0	4.5	400	400	0.49	0.24	0.08	0.95
3.0	5.0	400	400	0.48	0.25	0.08	0.94
3.0	5.5	400	400	0.49	0.26	0.08	1.01
3.5	3.0	400	400	0.51	0.29	0.10	1.59
3.5	3.5	400	400	0.49	0.24	0.10	0.93
3.5	4.0	400	400	0.48	0.25	0.10	1.03
3.5	4.5	400	400	0.49	0.25	0.10	0.94
3.5	5.0	400	400	0.48	0.24	0.10	0.92
3.5	5.5	400	400	0.49	0.26	0.10	0.91
4.0	3.0	400	400	0.50	0.29	0.09	1.50
4.0	3.5	400	400	0.49	0.27	0.09	0.93
4.0	4.0	400	400	0.49	0.26	0.09	0.99
4.0	4.5	400	400	0.49	0.26	0.09	0.99
4.0	5.0	400	400	0.48	0.26	0.09	0.98
4.0	5.5	400	400	0.48	0.27	0.09	0.99
4.5	3.0	400	400	0.50	0.26	0.12	1.16
4.5	3.5	400	400	0.50	0.26	0.12	1.02
4.5	4.0	400	400	0.49	0.25	0.13	1.03

$f_{\gamma,1}$	$f_{N,1}$	$f_{\gamma,n}$	$f_{N,n}$	mean rel. err.	std rel. err.	min rel err.	max rel. err.
4.5	4.5	400	400	0.50	0.25	0.12	0.98
4.5	5.0	400	400	0.49	0.26	0.13	0.96
4.5	5.5	400	400	0.49	0.27	0.09	1.02
5.0	3.0	400	400	0.53	0.29	0.12	1.34
5.0	3.5	400	400	0.51	0.27	0.12	1.10
5.0	4.0	400	400	0.51	0.26	0.13	1.00
5.0	4.5	400	400	0.51	0.26	0.13	1.07
5.0	5.0	400	400	0.50	0.27	0.12	1.05
5.0	5.5	400	400	0.50	0.28	0.10	0.99
5.5	3.0	400	400	0.52	0.28	0.16	1.19
5.5	3.5	400	400	0.51	0.27	0.14	1.02
5.5	4.0	400	400	0.49	0.25	0.16	1.02
5.5	4.5	400	400	0.51	0.26	0.14	0.96
5.5	5.0	400	400	0.50	0.27	0.13	1.00
5.5	5.5	400	400	0.51	0.29	0.11	1.08
3.0	3.0	500	500	0.52	0.28	0.09	1.38
3.0	3.5	500	500	0.52	0.26	0.09	1.02
3.0	4.0	500	500	0.50	0.25	0.08	0.90
3.0	4.5	500	500	0.50	0.25	0.09	0.91
3.0	5.0	500	500	0.51	0.25	0.09	0.93
3.0	5.5	500	500	0.50	0.25	0.09	0.95
3.5	3.0	500	500	0.51	0.28	0.08	1.19
3.5	3.5	500	500	0.50	0.26	0.08	1.00
3.5	4.0	500	500	0.49	0.26	0.08	0.92
3.5	4.5	500	500	0.50	0.26	0.09	0.94
3.5	5.0	500	500	0.49	0.26	0.08	0.94
3.5	5.5	500	500	0.48	0.25	0.09	0.87
4.0	3.0	500	500	0.51	0.28	0.10	1.37
4.0	3.5	500	500	0.49	0.25	0.10	0.91
4.0	4.0	500	500	0.49	0.27	0.10	0.93
4.0	4.5	500	500	0.49	0.26	0.10	0.94
4.0	5.0	500	500	0.49	0.26	0.10	0.93
4.0	5.5	500	500	0.50	0.27	0.10	0.96
4.5	3.0	500	500	0.51	0.30	0.12	1.56
4.5	3.5	500	500	0.50	0.27	0.11	0.98
4.5	4.0	500	500	0.50	0.26	0.12	0.91
4.5	4.5	500	500	0.49	0.26	0.12	0.90
4.5	5.0	500	500	0.49	0.27	0.11	0.96
4.5	5.5	500	500	0.50	0.27	0.11	0.97
5.0	3.0	500	500	0.53	0.31	0.12	1.65
5.0	3.5	500	500	0.51	0.26	0.12	0.93
5.0	4.0	500	500	0.51	0.27	0.12	1.01
5.0	4.5	500	500	0.51	0.27	0.12	0.98
5.0	5.0	500	500	0.51	0.28	0.13	0.98
5.0	5.5	500	500	0.51	0.28	0.09	1.01
5.5	3.0	500	500	0.53	0.31	0.13	1.53
5.5	3.5	500	500	0.52	0.27	0.13	1.04
5.5	4.0	500	500	0.51	0.28	0.12	0.98
5.5	4.5	500	500	0.51	0.27	0.13	0.97
5.5	5.0	500	500	0.51	0.28	0.11	0.99
5.5	5.5	500	500	0.51	0.27	0.11	0.93

Table B.1: Inversion results for geometric factor matrices with different minimal ($f_{\gamma,1}$, $f_{N,1}$) and maximal ($f_{\gamma,n}$, $f_{N,n}$) energies for gammas and neutrons. The quality of the inversion of the expected Martian input spectrum is given as the mean relative error. To get a more detailed impression, standard deviation, minimum and maximum of the relative error are shown as well.

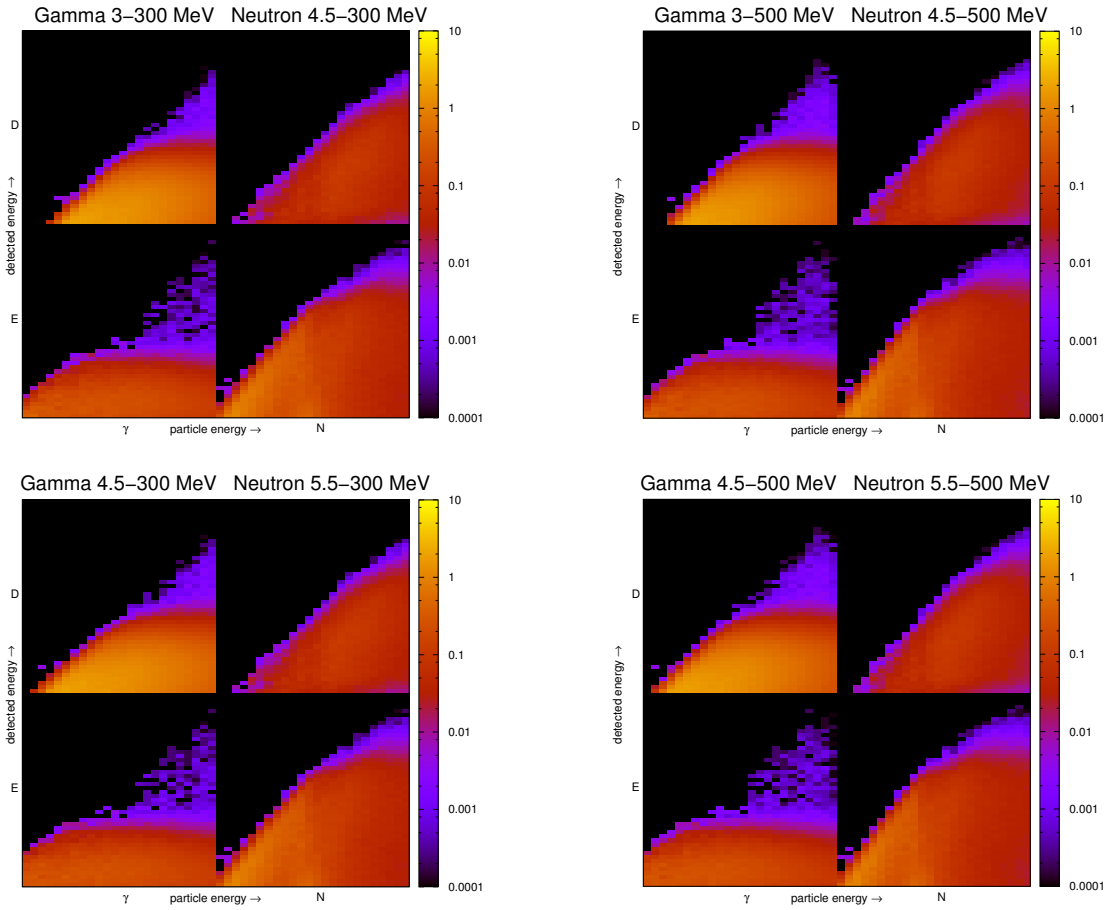

Figure B.8: Examples of the evaluated DRFs. The color denotes the geometric factor in [sr cm^2]. The range for the detected energies in **D** and **E** is given by 4 - 256 MeV. While for a maximum incoming-particle energy of 300 MeV the shapes of the submatrices look more or less triangular, for 500 MeV the slope of the triangular shape decreases at higher energies.

Appendix C

Simulation of ^{207}Bi calibration runs for MSL/RAD

C.1 Introduction

To calibrate the silicon channels of RAD and to investigate the temperature dependence of the calibration values, several measurements with a ^{207}Bi source on top of the instrument were performed. In addition to gammas, a ^{207}Bi source emits conversion electrons with the main energies 481, 975, 1047 keV and is therefore a convenient tool for calibrating particle detectors in the low energy range.

In terms of understanding RAD, a simulation recreating the ^{207}Bi calibration runs is done. In the calibration runs a ^{207}Bi electron source was placed directly on top the protection foil of the RAD instrument. Source data is available on `ixion.boulder.swri.edu: /home/rad/cal_data/` for both the FlightRAD (FM02) and the CalRAD (FM01) unit. Since the CalRAD data is more recent, the CalRAD calibration runs are chosen for an detailed analysis. In addition to the energy spectra in the detectors, the particle tracks and secondary production are analyzed, and unclear parameters in position and geometry of the ^{207}Bi source are estimated. A comparison between calibration run and simulation shows very good agreement. The dose rate in the **B** and **E** detector is calculated. Finally the temperature dependence of the FM02 sensorhead is investigated for ^{207}Bi thermal calibration runs.

C.2 Simulation setup

The simulation of the ^{207}Bi calibration run is done via the `GEANT4` toolkit. The geometry is set up using the BSPSolid class [Kortmann, 2010], which creates `GEANT4` solids directly from STL files generated from CAD models. The model of the RAD sensorhead (Fig. 1.3) implements:

- detector housing (magnesium)

- all silicon detectors, scintillators and their corresponding readout diodes
- protection foil (aluminum and kapton layers)
- downholder for the protection foil (aluminum)

Exact dimensions of the model are available as STL files. The **A** detector is separated in an inner (**A2**) and in an outer (**A1**) segment. The geometry of the detectors segments, **A2**, **A1**, **B** and **C**, is defined in Kortmann [2010] (Appendix H). The model of the ^{207}Bi source consists of

- aluminum guard and support rings
- platinum foil (thickness 5 μm)
- acrylic foil (thickness 0.1 μm)
- ^{207}Bi spot between platinum and acrylic foil (diamater 5.08 mm, thickness 0.1 μm)

Exact dimensions of the source are available as STL files. The manufacturer denoted a thickness of approximately 50 μm for the platinum foil. However, in the simulation, a 50 μm platinum foil can not reproduce the spectra of the calibration runs, while the estimation of 5 μm leads to much more accurate results, see Fig. C.5, C.7.

Electron, gamma and X-ray levels were taken from NuDat2.5 [2011]. To create decay events as realistic as possible, a MC process steps through the branches of the decay scheme and creates multiple primary particles per event. Fig. C.1 shows the ^{207}Bi decay scheme which was used in the simulation. In the particle generation process, the occurring X-ray radiation is considered as well. This decay process leads to an average of 2.9 particles per event, consisting of 0.12 electron, 1.79 gamma particles and 1.0 x-rays.

C.3 Energy deposits in the detectors

The energy spectra in the silicon detectors are shown in Fig. C.2 (top). Count rates in **A2**, **A1** are significantly higher than those in the **B** and **C** detector. Most of the primary electrons stop in the **A** detector and therefore do not create a signal in **B** or **C**. The conversion electrons can clearly be identified by the peaks at $\sim$ 400 and 1000 keV in the **A2** detector, the maximum at 100 keV is caused by minimal ionizing particles, the sharp lines below 100 keV are caused by the occurring X-rays.

Fig. C.2 (bottom) shows the energy spectra in the scintillators. The energy deposits are taken straightforward from the `GEANT4` simulations without any consideration of scintillator properties, readout diodes, or electronic noise. The energies of the primary gammas are visible as sharp lines only in the simulated CsI (**D**) scintillator, which has the largest cross section for gamma particles. The corresponding Compton edges are visible in all scintillators.

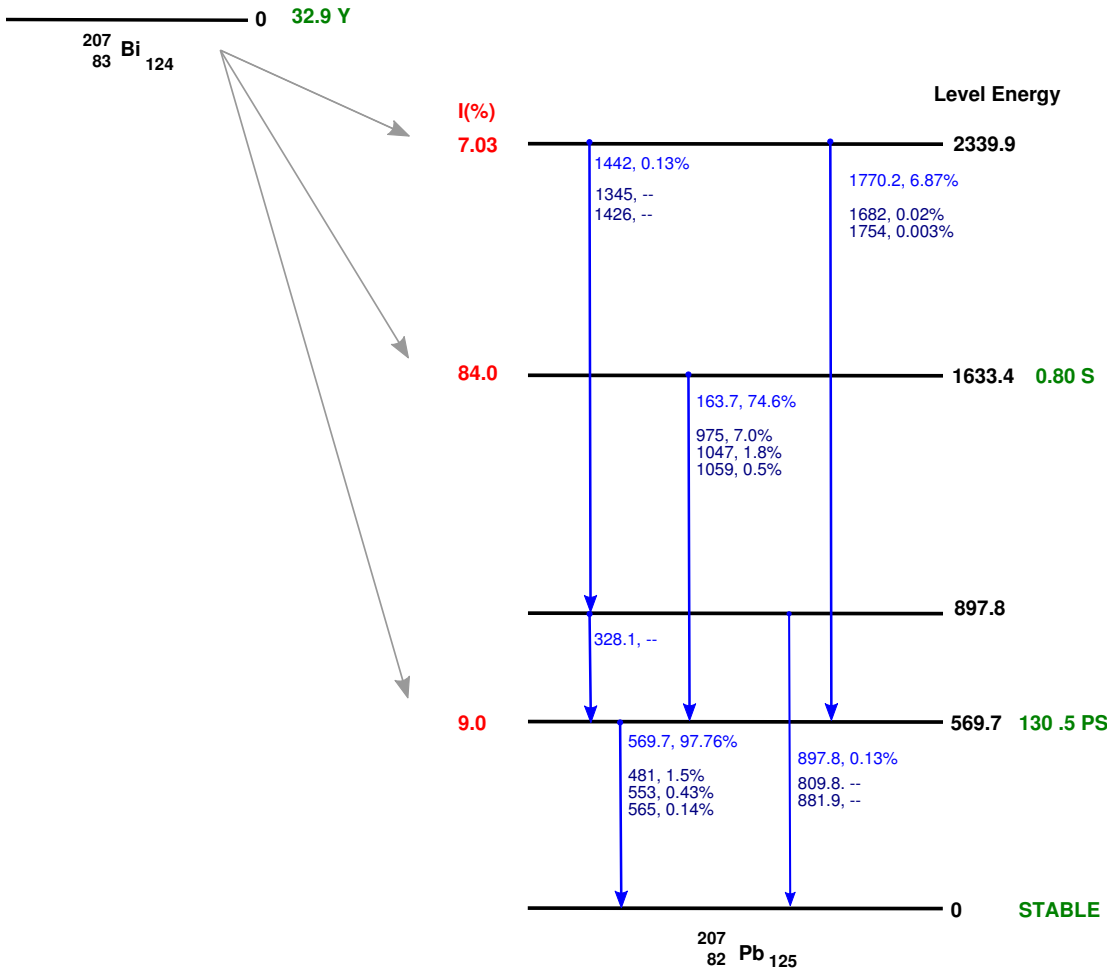

Figure C.1: The ^{207}Bi decay scheme. Level energy (black), gamma energy (blue) and conversion electron energy are plotted in keV. The half-life is plotted in green. The particle-generation gun in the simulation uses a MC process to step through the different branches and generate the correct distribution of primary particles per decay. E.g., allowed primary particle generations are: 1770.2 keV gamma + 481 keV electron or 975 keV electron + 569.7 keV gamma.

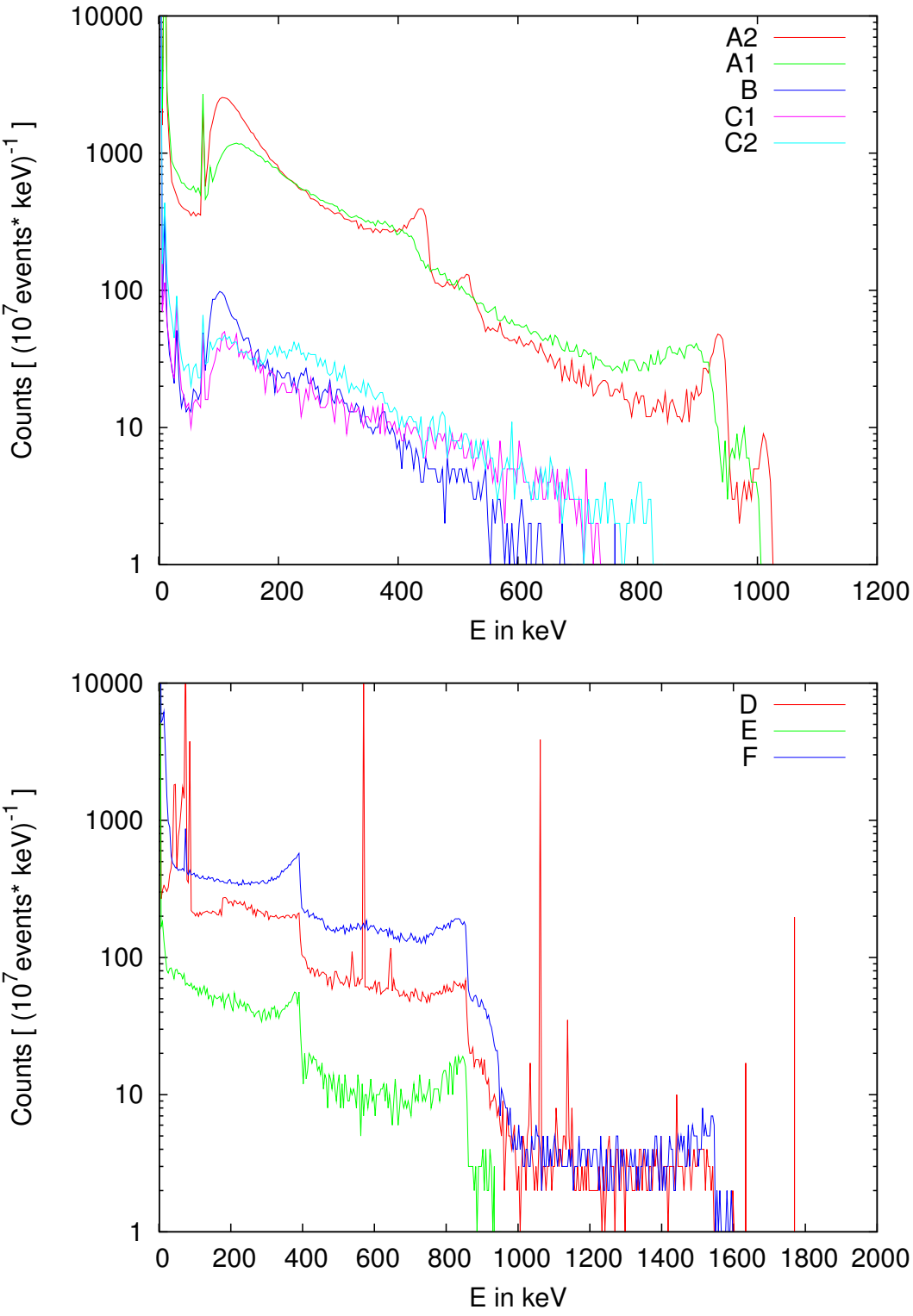

Figure C.2: Energy spectra for a run with 10^7 decay events. Top: Energy deposits in the silicon detectors. Bottom: Energy deposits in the scintillators.

	γ	e^-	e^+
A	0.38%	6.4%	0.001%
B	0.02%	0.128%	0 %
C	0.02%	0.07%	0 %
D	2.24%	4.58%	0.001 %
E	0.004%	0.1%	0 %
F	0.04%	0.92%	0 %
Detector housing	0.25%	1.35%	0 %
Protection foil	0.8%	9.5%	0%
Downholder	0.01%	0.36%	0%
^{207}Bi platinum foil	2.1%	18.9 %	0%
^{207}Bi housing	0.09%	1.35 %	0%

Table C.1: Probability of secondary production per decay event in different detector elements for a run with 10^5 decay events.

C.3.1 Particle tracks and secondary production

To investigate the influence of the various sensorhead and source elements on the propagation of particles, the particle tracks are evaluated. The primary particle tracks, produced secondaries, and visited volumes are read from the `GEANT4` tracking output, and are processed as trees of primary and secondary particles. By evaluating the event trees, statistical information about particle paths, produced secondaries, and occurring processes, can be obtained. The methods are available as a python library at `asterix.ieap.uni-kiel.de:/home/asterix/koehler/particle_cacade`.

The path of the primary electrons is plotted in Fig. C.3 (top). Starting from the ^{207}Bi element, one half of the electrons passes through the acrylic foil, the other half through the platinum foil. Following the paths through the acrylic foil, 34.6% of the primary electrons reach the **A** detector directly. Following the paths through the platinum foil, one finds that 11.4% of the electrons are reflected back into the ^{207}Bi element, and 7.5% of the produced electrons reach the **A** detector after being reflected by the platinum foil. In contrast to the electrons which reach **A** directly, electrons, which are backscattered by the platinum foil, loose some of their energy and, therefore, do not appear as sharp lines in the **A** detector. For a platinum foil with a thickness of 50 μm, the ratio of backscattered electrons, which reach **A**, increases to 22%. This leads to a decrease of the peaks in the energy spectra of the **A** detector, which does not agree with the callibration run, see Fig. C.7. Fig. C.3 (bottom) shows the paths of the primary gamma particles. As expected the graph shows straight paths without any backscattering.

The production of secondaries is another factor which can influence the energy spectra in the detectors. Table C.1 shows the probability per decay event for the generation of secondaries. The major part of all secondaries in generated in the **A** and **D** detector, the protection foil, and platinum foil. Tracking the paths of the secondaries with same visual-

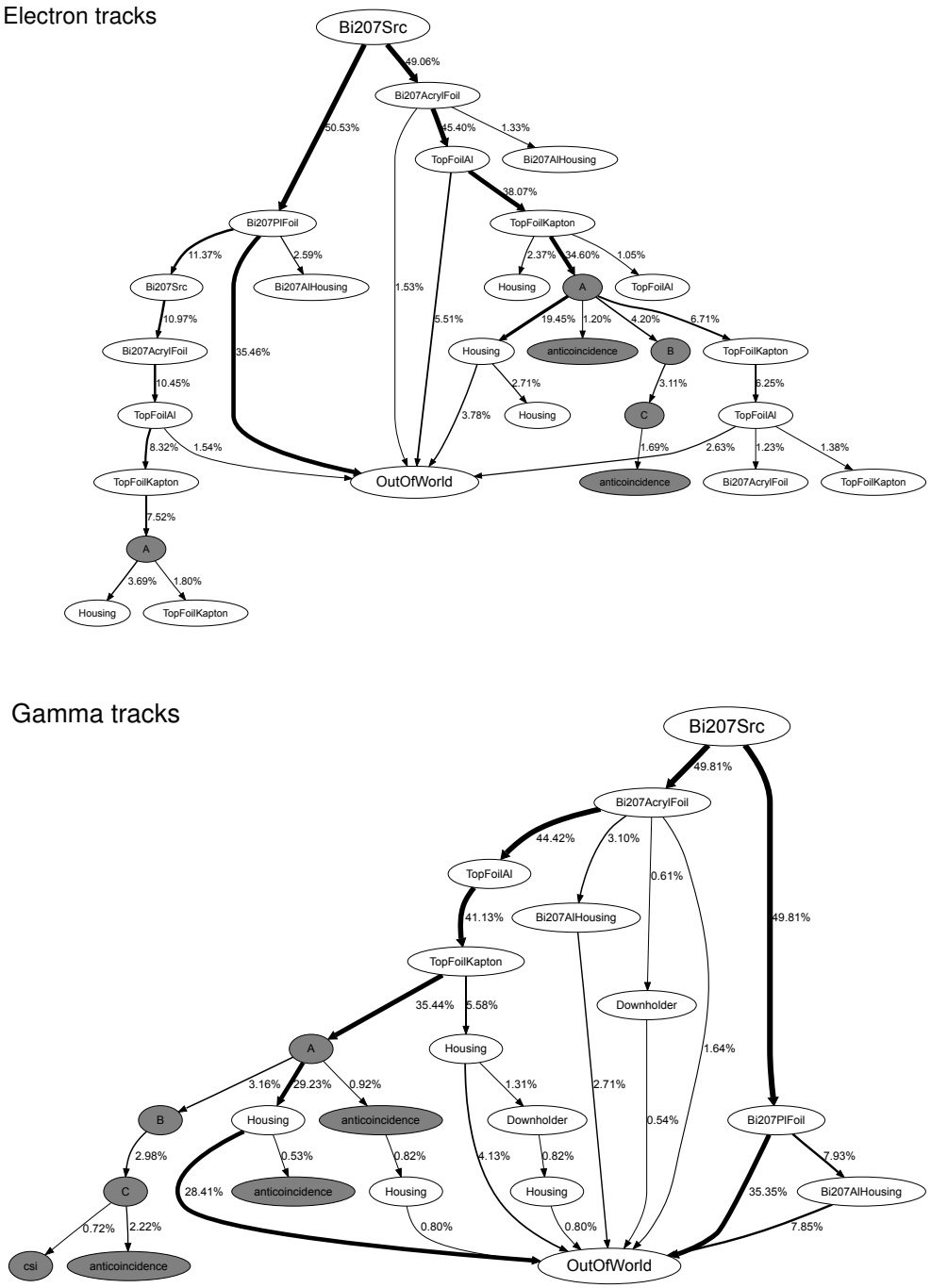

Figure C.3: Graphs of the particle tracks. The world volume is not shown, maximum length of one path is restricted to 8 volumes, intensities below 1% resp. 0.5% are excluded. Top: Tracks of the primary electrons. Following the paths through the platinum foil, 11.4% of the electrons are reflected back into the ^{207}Bi element and 7.5% of those electrons reach the **A** detector. Following the direct paths through the acrylic foil 34.6% of the primary electrons reach the **A** detector directly. Bottom: Tracks of the primary gamma particles. Tracks reaching the **E** detector exist, but their intensity is below the 0.5% visualization threshold.

ization method as for the primary particles, one finds that the majority of the generated secondaries do not reach any detector. An exception are the electrons generated in the platinum foil, of which 0.5% reach the **A** detector, and the secondaries generated in the protection foil, of which 1.1% reach the **A** detector. However, the influence of those secondaries is small compared to the influence of the backscattered electrons from the platinum foil.

C.3.2 Source position

During the calibration run, the ^{207}Bi source was placed directly on top of the sensorhead. However, since it was placed on the protection foil manually, a ± 1 mm inaccuracy can be assumed. Another source of inaccuracy is the distance between protection foil and ^{207}Bi spot, which can not be reconstructed clearly from the source description. Although small changes in the source position do not cause any significant changes in the count rate of the **A** detector, they can cause significant differences in the count rates of **A1** and **A2**. A deviation from the center position will increase the count rate in **A1** and decrease the count rate in **A2**. Likewise, an increased distance between protection foil and the ^{207}Bi spot will decrease the count rate in **A2** but can increase the count rate in **A1**. The influence of the sources position can be determined by calculating the ratio of the count rates of **A2** and **A1**, as a function of the displacements in x and z direction. For $(dx, dz) = (0, 0)$, the source is placed exactly at the center of the protection foil, where the distance between the ^{207}Bi element and the protection foil is 1.1 mm.

Fig. C.4 shows the ratio for several values of dx, dz and can be used to estimate the source position in the calibration runs. The influence of the ± 1 mm inaccuracy of the x position is small compared to the influence of a displacement in z-direction. The same effect is also present in the segments of the **B** and **C** detector, but far less pronounced then for the **A** detector. Depending on the source position, the conversion electrons can become clearly visible in the **A1** detector as well. E.g., $dz >5$ mm results in clearly visible electron peaks in the **A1** detector.

C.3.3 Comparison with calibration run data

To compare simulation with calibration runs, the histogram of the simulation data is fitted to the histogram of the calibration run `CalRAD_Bi207_20090127_15_48_0`, available at `ixion.boulder.swri.edu/home/rad/cal_data/CalRAD/`. Since the ^{207}Bi electrons have comparatively low energies and appear mainly in the upper silicon detectors (Fig. C.3), only the ultra-high-gain channels (**U**/**Hr**) of the **A2**, **A1**, **B**, and **C** detectors are fitted.

Determining source position

The dependence of the count rate **A2**/**A1** on the position of the ^{207}Bi source is shown in Fig. C.4. Matching the count rates of the calibration runs, the source position is estimated at $dx = 0$ mm, $dz = 2.8$ mm. Since the influence of the uncertainty in dx direction is

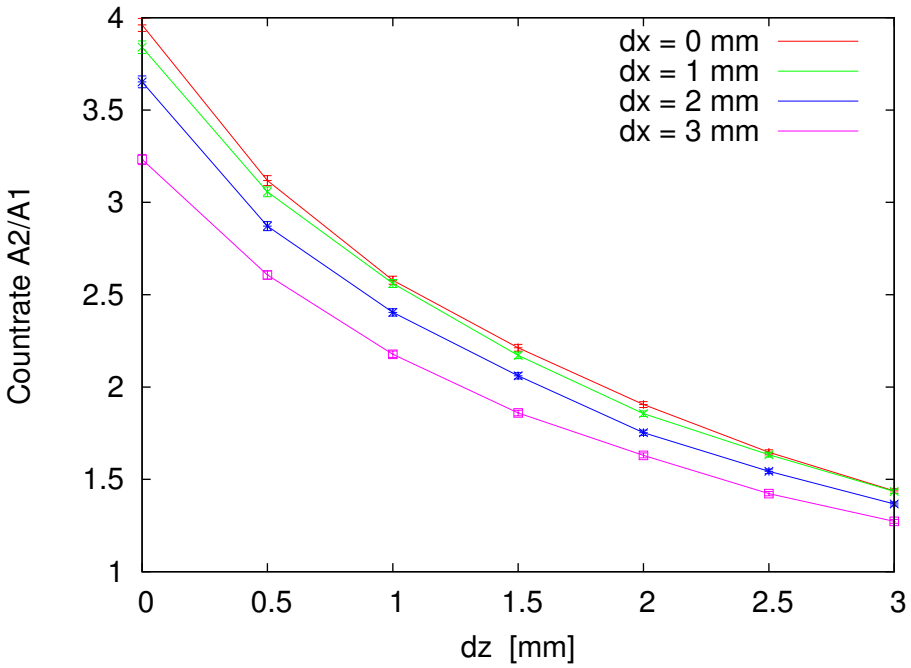

Figure C.4: The ratio of the count rates in **A2/A1** for different source positions dx, dz. The displacement in x direction accounts for the positioning of the source by hand, the z displacement accounts for the unclear geometry of the source.

much smaller than the influence of the uncertainty in dz direction, any possible shift in dx is neglected.

Fitting routine

Let $\vec{H}(\mathbb{A})$ be the histogram of a data set $\mathbb{A}$. The simulation data $\mathbb{S}$ is fitted to the measured data $\mathbb{M}$ via the parameters o - the offset of the calibration data in ADC, s - the conversion factor in keV/ADC, σ - the electronic noise in ADC, and I - which normalizes the intensity of the simulation-data histogram to the intensity of the calibration-data histogram.

To estimate the electronic noise σ, a Gaussian is fitted via the χ^2 method to the noise peak of the calibration run data $\vec{H}(\mathbb{M})$. $I \cdot \vec{H}(\hat{\mathbb{S}}(o, s, \sigma))$ is then fitted to $\vec{H}(\mathbb{M})$ with a χ^2-minimizing fit via the parameters o, s and I, where

$$\hat{\mathbb{S}}(o, s, \sigma) := \{o + x/s + \xi(\sigma) \mid \forall x \in \mathbb{S}\}, \tag{C.1}$$

and $\xi(\sigma)$ is a normal distributed random variable with zero mean and σ variance. The noise peak is excluded from the fit, because the simulation triggers all decay event and therefore shows a much large noise peak.

The function

$$f(o, s, I) \;:=\; \sum_i \left(\frac{H_i(\hat{\mathbb{S}}) - H_i(\mathbb{M})}{E_i} \right)^2 \tag{C.2}$$

is minimized, where

$$E_i := \max\left(1,\, \sqrt{H_i(\mathbb{M})}\right). \tag{C.3}$$

To estimate errors, the fit is repeated several times, for different electronic-noise random values, for several bootstrap MC [Press, 2007; Kortmann, 2010] generated measurements, and for several initial guess values selected from a uniform distribution centered around the best fit.

Fig. C.5 and C.6 show the fitted energy spectra for the evaluated channels, the fitted parameters are shown in Table C.2. The shape of simulation and calibration spectra agree very well, the I parameter matches for all channels within the estimated errors, which means the intensity of simulation and calibration spectra match as well.

Note that the offset o is not the maximum of the noise peak. This effect is probably caused by the peak-detector of the Voltage Input Electronics for Nuclear Applications (VIRENA). If an event was triggered in one channel, the PHA determines the maximum of the signal in the timespan Δt. The signal is given by offset + pulse + noise. Without an energy deposit in the detector, the pulse is 0 and the PHA finds the maximum noise value in the time interval Δt. The noise creates an additional shift of the signal. For a nonzero energy deposit in the detector the pulse is Gaussian shaped, with a hight proportional to the deposited energy. The maximum of the signal is given by offset + pulse + poise, but in contrast to the pulse=0 case, the maxima of the noise only contribute near the pulse maximum. The shift due to noise is therefore less than for the pulse=0 case.

Channel	o [ADC]	o' [ADC]	s [keV/ADV]	σ [ADC]	I
A2U	1945 ±4	1965 ±1	0.306 ±0.01	11.8 ±0.1	3.0 ±0.1
A1U	1560 ±3	1584 ±1	0.316 ±0.02	25.5 ±0.1	2.9 ±0.1
BU	1846 ±2	1858 ±1	0.328 ±0.05	10 ±0.1	3.0 ±0.1
CU	1635 ±2	1652 ±1	1.61 ±0.01	5.4 ±0.1	3.0 ±0.1

Table C.2: Table of the best fit parameters for the `CalRAD_Bi207_20090127_15_48_0` calibration run. o is the 0 keV position, o' the position of the noisepeak.

Estimation of the platinum foil thickness

As already mentioned in the previous sections, the manufacturer denoted the thickness of the platinum foil with 50 μm, with which the simulation can not reproduce the measured spectra. As shown in Fig. C.3, the platinum foil reflects a large fraction of the electron,

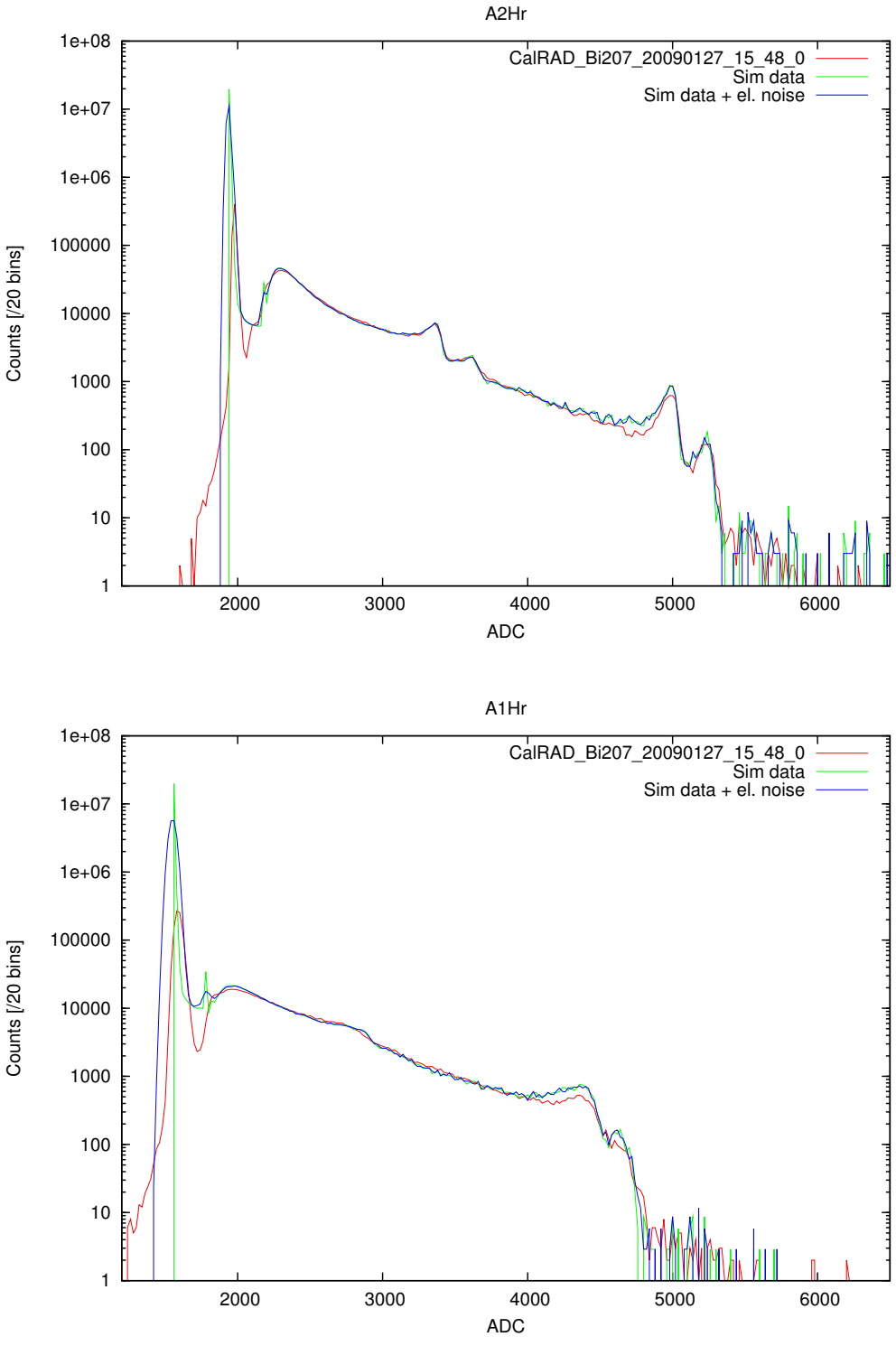

Figure C.5: Selected best fits for the **A2U** and the **A1U** channel. The Noise peak can be seen in the left part of the spectrum, followed by the peak from the minimum ionizing particles. In the **A2U** detector, the peaks of the conversion electron can clearly be identified. In the **A1U** detector the peaks can barely be identified.

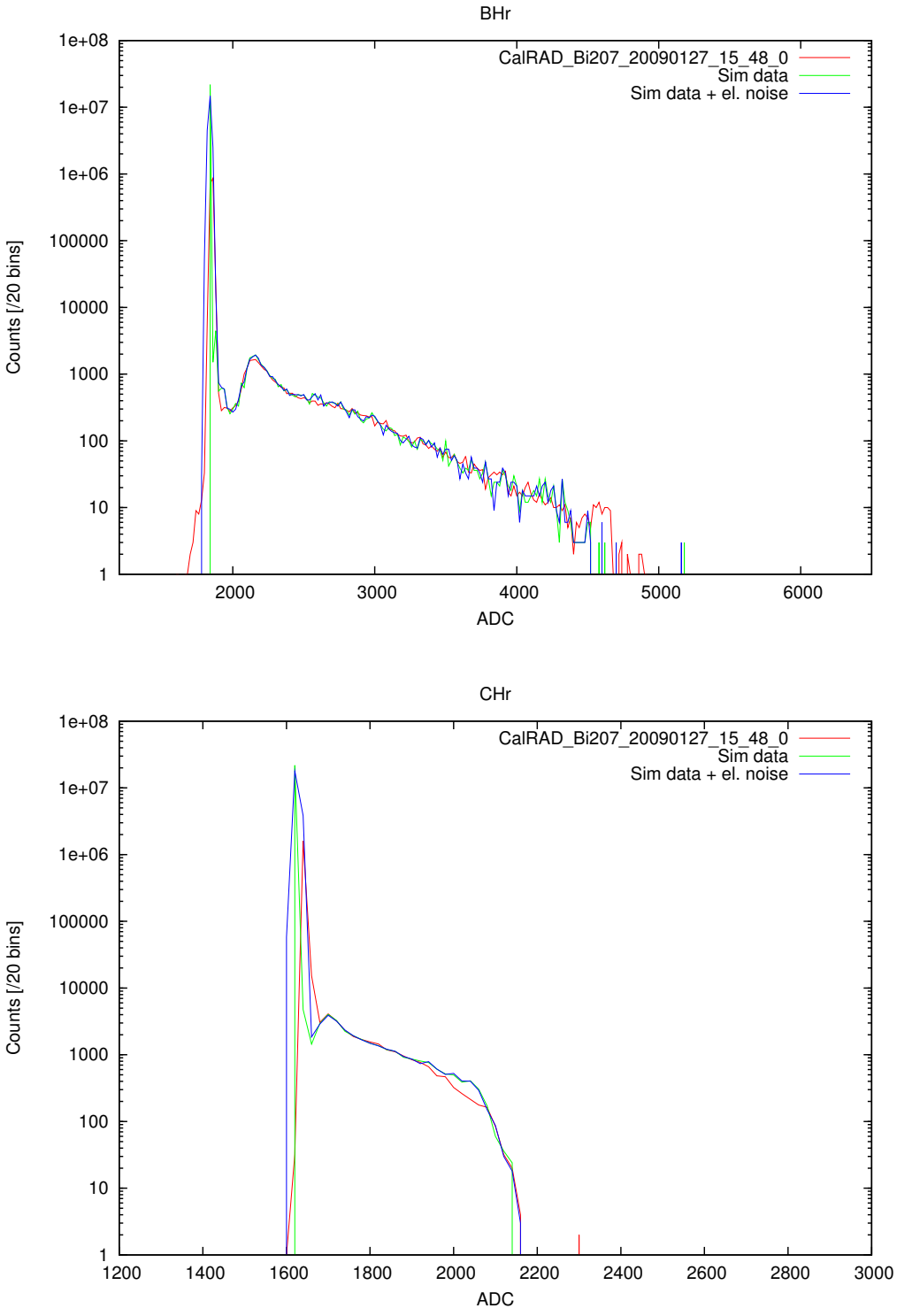

Figure C.6: Selected best fits for the BU and the CU channel. Unlike in the **A2U**, **A1U** channels (Fig. C.2), the peaks of the conversion electrons can not be identified anymore.

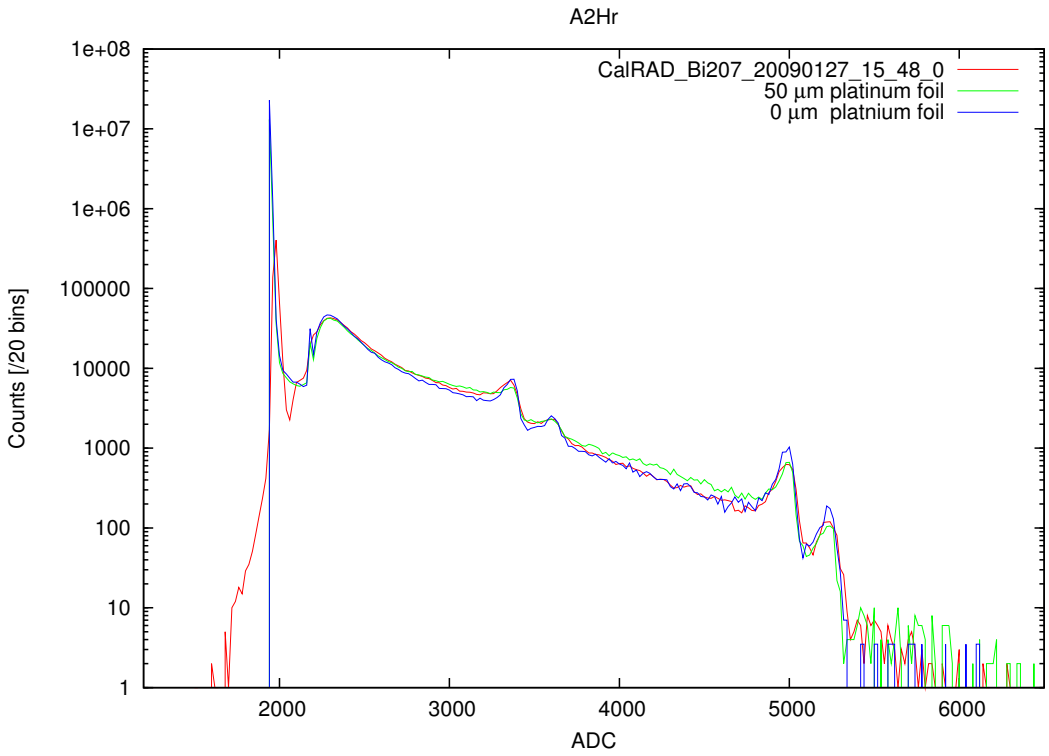

Figure C.7: Comparison of measurement and simulations for 0 and 50 μm platinum foil. Fit parameters were taken from Table C.2, the I parameter had to be adjusted to match the changed intensity due to the backscattered electron, $I_{0\mu m} = 4$, $I_{50\mu m} = 2$. Without the platinum foil the peaks of the conversion electron are much sharper then in the measurements, with 50μm platinum foil the peaks from the simulation are less pronounced then in the measurement.

which appear as noisy electron in the **A** detector. An increase of the foils thickness increases the ratio of reflected electrons as well. Since determining the real value via simulation requires a lot of computation time, the value was determined by bisecting the thickness of the platinum foil. The values 50 μm, 0 μm, 25 μm, 10 μm, 5 μm have been calculated. 5μm has been chosen as the best fitting value, with an error estimate of ± 3 μm.
For completeness, Fig. C.7 shows the **A2U** channel is shown for 0 and 50 μm.

C.4 Absorbed dose rate in the B and E detector

The absorbed dose rate is defined as the absorbed energy per mass and time. For the `GEANT4` simulation, the absorbed dose rate for a detector d can be calculated via

$$\dot{D}_d = \frac{a}{N} \sum_{i=1}^{N} E_{d,i}/m_d, \tag{C.4}$$

where

- $\dot{D}_d$ is the absorbed dose rate in detector d,
- $E_{d,i}$ is the energy deposit in detector d for the ith event,
- N is the number of simulated decay events,
- m_d is the mass of detector d and,
- a is the activity of the ^{207}Bi electron source.

Using a source activity of 37 kBq, a platinum foil thickness of 5 μm, and a detector mass of $m_B = 0.13$ g, $m_E = 35.1$ g leads to an absorbed dose rate of

$$\dot{D}_B = 15.2 \pm 0.2 \text{ nGy/s}, \tag{C.5}$$
$$\dot{D}_E = 0.183 \pm 0.002 \text{ nGy/s}. \tag{C.6}$$

Using a 50 μm platinum foil for the ^{207}Bi source, the absorbed dose rate increases due to backscattered electrons. However, this increase is not present in the **E** detector, in which the energy deposits are mainly caused by gamma particles.

$$\dot{D}_B = 19.5 \pm 0.2 \text{ nGy/s}, \tag{C.7}$$
$$\dot{D}_E = 0.180 \pm 0.002 \text{ nGy/s}. \tag{C.8}$$

File	T [°C]
FTP_30_01_20081213_23_47_0.bin	-42.6 ± 0.3
EDC_10_02_20081214_21_50_0.bin	-42.01 ± 0.03
EDC_10_04_20081215_01_16_0.bin	-31.8 ± 0.1
EDC_10_06_20081215_03_50_0.bin	-21.2 ± 0.2
EDC_10_09_20081215_07_00_0.bin	-11.68 ±0.04
EDC_10_08_20081215_06_16_0.bin	-10.6 ± 0.1
EDC_10_12_20081215_09_43_0.bin	-0.6 ± 0.3
EDC_10_14_20081215_12_55_0.bin	9.4 ± 0.3
EDC_10_16_20081215_15_53_0.bin	19.85 ± 0.04
EDC_10_18_20081215_18_20_0.bin	30.1 ± 0.1
FTP_26_01_20081211_17_26_0.bin	35.7 ± 0.1
FTP_26_2_20081211_17_48_0.bin	36.2 ± 0.2
EDC_10_20_20081215_20_24_0.bin	40.6 ± 0.3
EDC_10_22_20081215_22_56_0.bin	49.3 ± 0.2
EDC_10_24_20081216_01_36_0.bin	58.9 ± 0.4
FTP_27_01_20081212_00_57_0.bin	60.24 ± 0.07
FTP_28_01_20081212_21_14_0.bin	60.30 ± 0.07

Table C.3: Calibration files from ixion.boulder.swri.edu: /home/rad/cal_data/FlightRAD/2008/FM2_thermal and REB temperatures used for the analysis.

C.5 Analysis of thermal datasets

Since RAD will operate in an environment with varying temperatures, the temperature dependence of gain, offset, and noise were investigated for a wide range of temperatures. As shown in the previous section, the peaks of the conversion electron are only visible in the **A** detector. Using the peak positions, E. Boehm was able to determine the temperature dependence of the **A2U**, **A1U** channels [Böhm, 2010]. Fitting the shapes of the spectra, the temperature dependence of the **BU** and **CH** can be obtained as well.

For the ^{207}Bi calibration, the FM2 sensor head and the FlightREB FM1 were used, which is the same configuration as on MSL. The data files are available at ixion.boulder.swri.edu: /home/rad/cal_data/FlightRAD/2008/FM2_thermal. Tab. C.3 contains a list of files which were used for the analysis. Files with insufficient statistics were excluded. Through the different files, the apparent position of the ^{207}Bi source and the intensity of the spectra both vary. This means, although the fitting method works for this calibration data, a lot of fine tuning for each data file is necessary. Noise peak position o' and σ can easily be fitted for each data file. For offset o and shift s, best results were obtained by choosing one data file with sufficient statistics, calculating gain and offset by fitting GEANT4 simulation data as described above - then calculating the relative change of the gain, by comparing the spectra from the data files to the calibrated data file. EDC_10_12_20081215_09_43_0.bin

Channel	o [ADC]	o' [ADC]	s [ADC/MeV]	σ [ADC]
A2U	1614 ±4	1634.03 ±0.03	0.331 ±0.001	10.52 ±0.02
A1U	1692 ±3	1715.1 ±0.1	0.35 ±0.02	24.8 ±0.1
BU	1596 ±4	1615.9 ±0.3	0.321 ±0.003	15.0 ±0.6
CH	1775 ±10	1786.93 ±0.03	1.67 ±0.03	2.5 ±0.1

Table C.4: Table of the best fit parameters for `EDC_10_12_20081215_09_43_0.bin`. o' denotes the position of the noisepeak, o denotes the 0 keV position.

($T = -0.51$°C) is used for the energy calibration, the fitted parameters are presented in Tab. C.4. Temperature dependence of offset and noise are shown in Fig. C.8 and C.9, the relative change of the gain $s(T)/s(T = -0.51°\text{C})$ is shown in Fig. C.10. Since the ^{207}Bi source creates no clear signal in the **D**, **E** and **F** channels, the temperature dependence of the gain can not be obtained. However, o' and σ_{noise} can easily be calculated.

With the exception of $T > 40$°C for the **A2U** channel, the noisepeak position always displays a linear dependence on temperature. Although there are some deviations, the temperature dependence of σ_{noise} can be approximated by a linear equation. For **A2U**, **CH** and **DH** there are some systematic deviations near 0°C, which are probably not caused by temperature effects.

The the relative change of the gain was validated by `GEANT4` calibrations of the -42.6°C and 48.9°C datasets. Without any consideration of underlying physics, the functions

$$o'(T) = x_o + b_o \cdot T \tag{C.9}$$

$$\sigma(T) = a_\sigma + b_\sigma \cdot T \tag{C.10}$$

$$s(T)/s(T = -0.51°C) = a_s + b_s \cdot T + c_s \cdot T^2 + d_s \cdot T^3, \tag{C.11}$$

are fitted with χ^2 minimization, errors were determined via bootstrap MC. Extreme deviations were excluded, e.g. the noise peak positions for **A2U** above 40°C. Best fit parameters are shown in Tab. C.5.

C.6 Conclusions

The simulation of the ^{207}Bi source can fully reproduce the spectra from the calibration run. Fitting simulation to measurement, one finds that the noise peak from the measurements does not correspond to the offset $o\hat{=}0$ keV, but is biased toward positive values. The simulation clearly shows that only an insignificant fraction of the electrons reaches the scintillators and only a small fraction reaches the **B** and **C** detector. Clear peaks in the spectra, which are used to identify the conversion electrons, are only visible in the **A2** and, depending on the source position, in the **A1** detector. Nevertheless, the measurement spectra can be fitted to the `GEANT4` spectra to obtain calibration values for **B** and **C** as well. The source position, which is not clearly defined due to the unclear geometry, can be estimated by matching the ratio of the count rates in **A2**/**A1**.

	A2U	**A1U**	**BU**	**CH**
a_o [ADC]	1635 ±1	1714 ±1	1628± 4	1787 ±1
b_o [ADC/°C]	-0.24 ±0.06	-0.12 ±0.04	0.76± 0.08	-0.27 ±0.05
a_σ [ADC]	10.1 ±0.2	24.0 ±0.4	8.9 ±0.1	2.51 ±0.03
b_σ [ADC/°C]	0.033 ±0.005	0.12 ±0.01	0.027 ±0.005	0.003 ±0.001
a_s [ADC]		1.0008 ± 0.0008		
b_s [ADC/°C]		$-2.36 \cdot 10^{-3} \pm 0.03 \cdot 10^{-3}$		
c_s [ADC/°C^2]		$4.53 \cdot 10^{-5} \pm 0.07 \cdot 10^{-5}$		
d_s [ADC/°C^3]		$-2.8 \cdot 10^{-7} \pm 0.1 \cdot 10^{-7}$		

Table C.5: Table of the best fit parameters for the temperature dependence of noise peak position o', noise σ and normalized gain $s(T)/(T = -0.51°C)$. The temperature of the normalized gain is assumed to be identical for all channels.

	DH	**EH**	**F1**	**F2**
a_o [ADC]	1811 ±1	1874 ±1	1902± 4	1794 ±32
b_o [ADC/°C]	-0.45 ±0.04	-0.77 ±0.05	1.4± 0.1	-0.63 ±0.06
a_σ [ADC]	8.4 ±0.8	5.5 ±0.4	93 ± 2	78 ±2
b_σ [ADC/°C]	0.01 ±0.02	0.01 ±0.01	0.6 ±0.1	0.39 ±0.03

Table C.6: Table of the best fit parameters for the temperature dependence of noise peak position o' and noise σ.

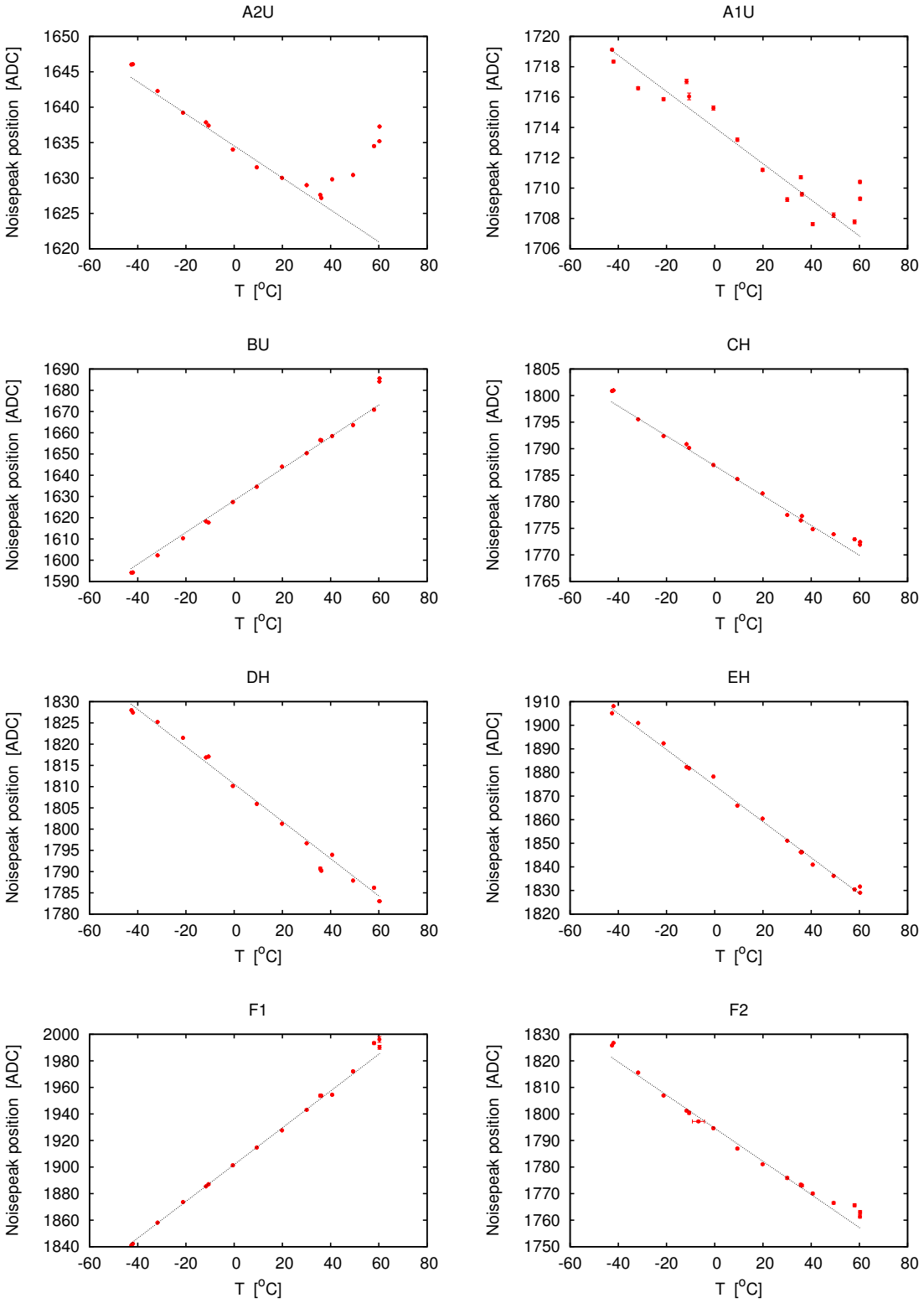

Figure C.8: Temperature dependence of the noise peak position o' for the (ultra) high-gain channels. The parameters of the fitted black curves (Eq. C.9) can be found in Tab. C.5, C.6.

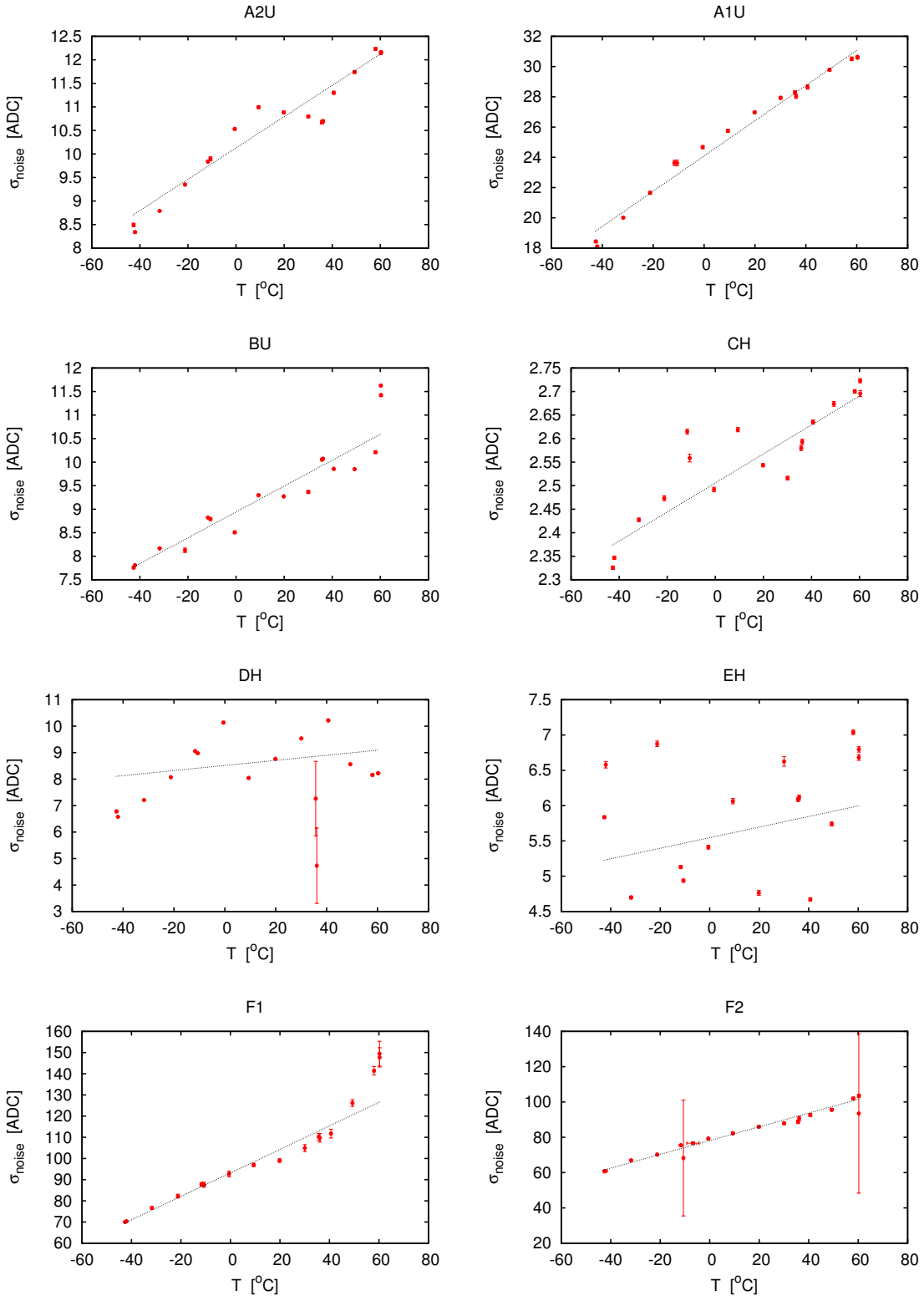

Figure C.9: Temperature dependence of the noise σ for the (ultra) high-gain channels. The parameters of the fitted black curves (Eq. C.10) can be found in Tab. C.5, C.6.

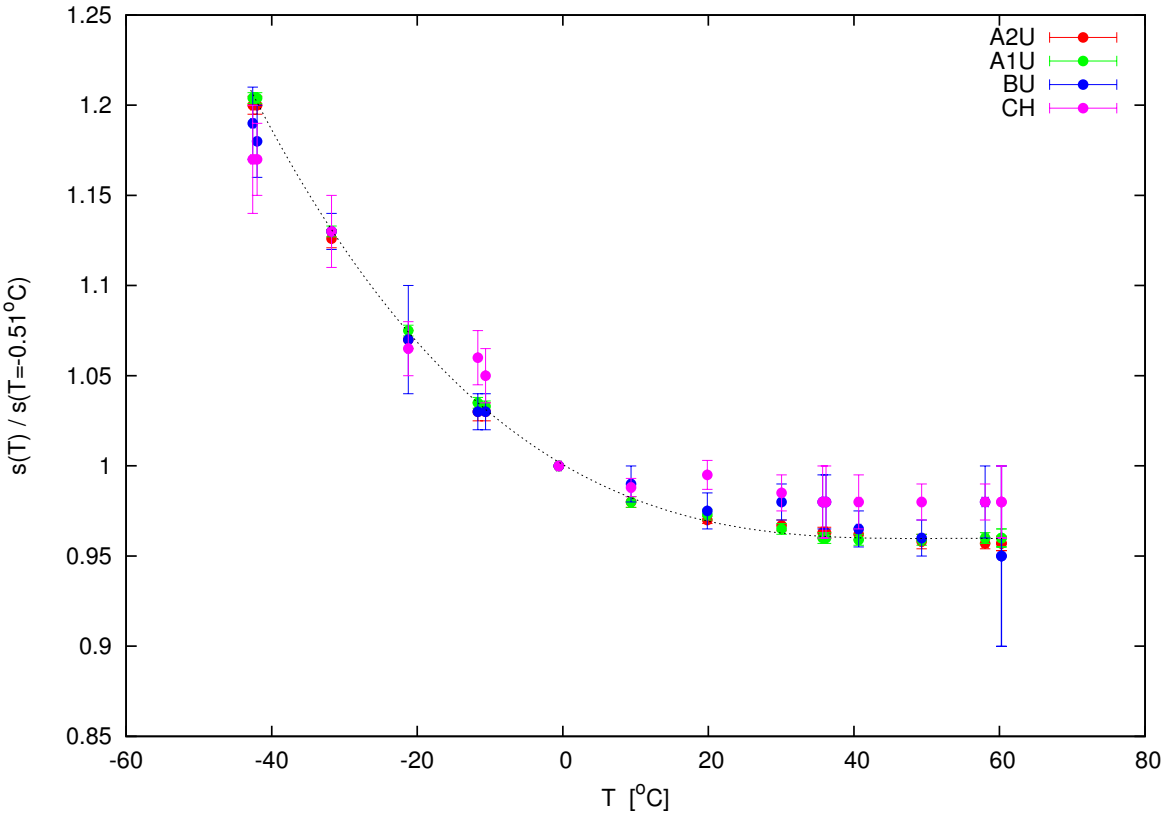

Figure C.10: Temperature dependence of the normalized gain $s(T)/(T = -0.51°C)$ for **A2U**, **A1U**, **BU**, **CH**. The parameters of the fitted black curve (Eq. C.9) can be found in Tab. C.5.

One of the most remarkable aspects, is the large influence of the platinum foil, on which a large fraction of electron are backscattered before they reach the **A** detector, hence, producing additional noise and increasing count rate and absorbed dose rate in the silicon detectors. A ^{207}Bi source without a platinum foil would create much sharper electron peaks in the spectra and, therefore, allow a more precise calibration.

Thermal datasets of ^{207}Bi calibration runs were analyzed, and temperature dependence of noise offset and σ were calculated directly from the datasets. The relative change of the gain was calculated, the absolute gain was only determined for selected datasets with sufficient statistics. Temperature dependence of noisepeak position and σ_{noise} can be approximated by a linear function, the relative gain can by approximated by a third-order polynomial.

Bibliography

S. Agostinelli. GEANT4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.

N. Agueda. *Near-relativistic electron events. Monte Carlo simulations of solar injection and interplanetary transport.* PhD thesis, Universitat de Barcelona, 2008.

V. Avdeichikov, B. Jakobsson, V. Nikitin, P.V. Nomokonov, and A. Wegner. Systematics in the light response of BGO, CsI(Tl) and GSO(Ce) scintillators to charged particles. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 484(1-3):251–258, 2002.

A. Balogh, T.J. Beek, R.J. Forsyth, P.C. Hedgecock, R.J. Marquedant, E.J. Smith, D.J. Southwood, and B.T. Tsurutani. The magnetic field investigation on the ULYSSES mission - Instrumentation and preliminary scientific results. *Astronomy and Astrophysics Supplement Series (ISSN 0365-0138)*, 92:221–236, 1992.

B. Bavassano and R. Bruno. Solar Wind Fluctuations at Large Scale: A Comparison Between Low and High Solar Activity Conditions. *Journal of Geophysical Research*, 92, 1991.

B. Bavassano, M. Dobrowolny, F. Mariani, and N.F. Ness. Radial Evolution of Power Spectra of Interplanetary Alfvénic Turbulence. *Journal of Geophysical Research*, 87(A5):3617–3622, 1982.

J.W. Bieber, W. Wanner, and W.H. Matthaeus. Dominant two-dimensional solar wind turbulence with implications for cosmic ray transport. *Journal of Geophysical Research*, 101:2511–2522, 1996.

A. Bjõrck. *Numerical Methods for Least Squares Problems.* SIAM: Society for Industrial and Applied Mathematics, 1996.

E. Böhm, A. Kharytonov, and R.F. Wimmer-Schweingruber. Solar energetic particle spectra from the SOHO-EPHIN sensor by application of regularization methods. *Astronomy and Astrophysics*, 473(2):673–682, 2007.

E. Böhm. Analysis of FM2 Thermal dataset of year 2008. Technical report, 2010.

S. Böttcher. Signal Processing in the Radiation Assessment Detector for MSL. Technical report, 2008.

R. Bruno and B. Bavassano. Large-scale solar wind fluctuations in the inner heliosphere at low solar activity. *Journal of Geophysical Research*, 94:168–176, 1989.

R.H. Byrd, P. Lu, J. Nocedal, and C. Zhu. A Limited-Memory Algorithm for Bound Constrained Optimization, 1994.

V. Carbone and R. Bruno. The Solar Wind as a Turbulence Laboratory. *Living Rev. Solar Phys*, 2(4), 2005.

R. Craun and D. Smith. Analysis of response data for several organic scintillators. *Nuclear Instruments and Methods*, 80(2):239–244, 1970.

K.U. Denskat and F.M. Neubauer. Statistical Properties of Low-Frequency Magnetic Field Fluctuations in the Solar Wind from 0.29 to 1.0 AU During Solar Minimum Conditions: HELIOS 1 and HELIOS 2. *Journal of Geophysical Research*, 87(A4):2215–2223, 1982.

L. Desorgher, E.O. Flückiger, and M. Gurtner. The PLANETOCOSMICS Geant4 application. *36th COSPAR Scientific Assembly*, 36:2361, 2006.

W. Dröge. Transport of solar energetic particles. *The Astrophysical Journal Supplement Series*, 90:567, 1994.

W. Dröge. Particle scattering by magnetic fields. *Space Science Reviews*, 93:121–151, 2000.

W. Dröge. Solar Particle Transport In A Dynamical Quasi-Linear Theory. *The Astrophysical Journal*, 589:1027–1039, 2003.

B. Ehresmann. *The Martian Radiation Environment - Early Mars and Future Measurements with the Radiation Assessment Detector*. PhD thesis, 2012.

W. Elsässer. The Hydromagnetic Equations. *Physical Review*, 79(1):183–183, 1950.

C. Gardiner. *Handbook of Stochastic Methods: for Physics, Chemistry and the Natural Sciences (Springer Series in Synergetics)*. Springer, 2004.

R. Grappin, J. Leorat, and A. Pouquet. Dependence of MHD turbulence spectra on the velocity field-magnetic field correlation. *Astronomy and Astrophysics (ISSN 0004-6361)*, 126:51–58, 1983.

M. Gurtner, M.R. Moser, E.O. Flückiger, and L. Desorgher. Simulation of the interaction of space radiation with the Martian atmosphere and surface. *Advances in Space Research*, 36(11):2176–2181, 2005.

M. Hannam and W. Thompson. Estimating small signals by using maximum likelihood and Poisson statistics. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 431(1-2):239–251, 1999.

K. Hasselmann and G. Wibberenz. Scaterring of charged particles by random electromagnetic fields. *Zeitschrift fuer Geophysik*, 34:353–388, 1968.

D.M. Hassler. MSL – Radiation Assessment Detector (RAD) Proposal. Technical report, 2004.

T. Hauschild and M. Jentschel. Comparison of maximum likelihood estimation and chi-square statistics applied to counting experiments. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 457(1-2):384–401, 2001.

E.A. Hawley. Radiation Measurements on the Qualification RTG. Technical report, 1984.

W.N. Hess, H. Patterson, R. Wallace, and E. Chupp. Cosmic-Ray Neutron Energy Spectrum. *Physical Review*, 116(2):445–457, 1959.

F. James. MINUIT Tutorial. *CERN Computing and Data Processing School*, (3), 1972.

J.R. Jokipii. Cosmic-Ray Propagation. I. Charged Particles in a Random Magnetic Field. *The Astrophysical Journal*, 146:480, 1966.

E. Jones, T. Oliphant, P. Peterson, and Others. SciPy: Open source scientific tools for Python, 2001.

J. Köhler, B. Ehresmann, C. Martin, E. Böhm, A. Kharytonov, O. Kortmann, C. Zeitlin, D.M. Hassler, and R.F. Wimmer-Schweingruber. Inversion of neutron/gamma spectra from scintillator measurements. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 269(22):2641–2648, 2011.

O. Kortmann. *Scintillator performance investigation for MSL/RAD.* PhD thesis, Christian-Albrechts-Universität zu Kiel, 2010.

C.L. Lawson and R.J. Hanson. *Solving Least Squares Problems (Classics in Applied Mathematics).* Society for Industrial Mathematics, 1987.

W.R. Leo. *Techniques for Nuclear and Particle Physics Experiments: A How-to Approach.* Springer, 1994.

J.L. Lovell, M.L. Duldig, and J.E. Humble. An extended analysis of the September 1989 cosmic ray ground level enhancement. *Journal of Geophysical Research*, 103(A10):23733–23742, 1998.

F. Mariani, N. F. Ness, L. F. Burlaga, B. Bavassano, and U. Villante. The large-scale structure of the interplanetary magnetic field between 1 and 0.3 AU during the primary mission of HELIOS 1. *Journal of Geophysical Research*, 83:5161–5166, November 1978.

E. Marsch. Kinetic Physics of the Solar Wind Plasma. pages 45–133. 1991.

C. Martin. Some aspects on the problem of the energy calibra- tion for the MSL/RAD flight units. Technical report, 2008.

NASA MSL Website. http://marsprogram.jpl.nasa.gov/msl/, 2011.

R. Nolte, M.S. Allie, R. Böttger, F.D. Brooks, A. Buffler, V. Dangendorf, H. Friedrich, S. Guldbakke, H. Klein, J.P. Meulders, D. Schlegel, H. Schuhmacher, and F.D. Smit. Quasi-monoenergetic neutron reference fields in the energy range from thermal to 200 MeV. *Radiation protection dosimetry*, 110(1-4):97–102, 2004.

NuDat2.5. http://www.nndc.bnl.gov/nudat2/, 2011.

I.D. Palmer. Transport Coefficients of Low-Energy Cosmic Rays in Interplanetary Space. *Reviews of Geophysics*, 20(2):335–351, 1982.

D.B. Percival and A.T. Walden. *Wavelet Methods for Time Series Analysis*. Cambridge University Press, 2000.

H. Porsche. General aspects of the mission Helios 1 and 2. Introduction to a special issue on initial scientific results of the Helios Mission. *Journal of Geophysics Zeitschrift Geophysik*, 42:551–559, 1977.

H. Porsche. 10 Jahre HELIOS 10 Years HELIOS. In *Publication Celebrating the 10th Anniversary of the Launch of HELIOS on December 10, 1974, DFVLR Oberpfaffenhofen*. 1984.

W.H. Press. *Numerical recipes: the art of scientific computing*. Cambridge University Press, 2007.

G. Prölss. *Physik des erdnahen Weltraums: Eine Einführung (German Edition)*. Springer, 2003.

E. Roelof. *Lectures in High-Energy Astrophysics, ed Oegelman, H. and Wayland, J.R.*, volume 111. 1969.

R. Schlickeiser and U. Jaekel. The Fokker-Planck coefficients of cosmic ray transport in random electromagnetic fields. *Journal of Physics G Nuclear Physics*, 18:1089–1118, 1992.

A. Shalchi and R. Schlickeiser. The Parallel Mean Free Path of Heliospheric Cosmic Rays in Composite Slab/Two-dimensional Geometry. I. The Damping Model of Dynamical Turbulence. *The Astrophysical Journal*, 604(2):861–873, 2004.

C.W. Smith, J L Heureux, and N.F. Ness. THE ACE MAGNETIC FIELDS EXPERIMENT C. W. SMITH, J. LHEUREUX and N. F. NESS. *Space Science Reviews*, pages 613–632, 1998.

Southwest Research. USERS MANUAL FOR THE RADIATION ASSESSMENT DETECTOR OF THE MARS. Technical report, 2010.

B.T. Tsurutani and G. S. Lakhina. Some basic concepts of wave-particle interactions in collisionless plasmas. *Reviews of Geophysics*, 35:491–501, 1997.

C.Y. Tu and E. Marsch. MHD structures, waves and turbulence in the solar wind: Observations and theories. *Space Science Reviews*, 73(1):1210, 1995.

C. R. Vogel. *Computational methods for inverse problems, Volume 10*. SIAM, 2002.

K.P. Wenzel, R.G. Marsden, D.E. Page, and E.J. Smith. The ULYSSES Mission. *ASTRON. AND ASTROPHYS. SUPPL. V.92*, 1992.

G. Wibberenz and J. Beeck. Pitch angle distributions of solar energetic particles and the local scattering properties of the interplanetary medium. *Astrophysical Journal*, 311:437–450, 1986.

C. Zhu, R.H. Byrd, P. Lu, and J. Nocedal. Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM Transactions on Mathematical Software*, 23(4):550–560, 1997.

Acronyms

AC Anti-Coincidence.

ACE Advanced Composition Explorer.

CAD Computer-Aided Design.

CME Coronal Mass Ejections.

CWT Continuous Wavelet Transformation.

DRF Detector Response Function.

GCR Galactic Cosmic Rays.

INL Idaho National Laboratory.

MC Monte-Carlo.

ML Maximum Likelihood.

MSL Mars Science Laboratory.

NASA National Aeronautics and Space Administration.

NNLS Non-Negative Least Squares.

PHA Pulse-Height Analysis.

PTB Physikalisch-Technische Bundesanstalt.

QLT Quasi-Linear Theory.

RAD Radiation Assessment Detector.

REB RAD Electronics Box.

RSH RAD Sensor Head.

RTG Radioisotope Thermoelectric Generator.

SEP Solar Energetic Particles.

VIRENA Voltage Input Electronics for Nuclear Applications.

Danksagung

Zunächst möchte ich Herrn Prof. Dr. Robert F. Wimmer-Schweingruber für die Vergabe des Themas und für die sehr gute Betreuung danken. Er schaffte es, mich auch in schwierigen Phasen zu motivieren und war stets bereit, auftretende Probleme zu diskutieren.

Weiterhin gilt mein Dank allen Mitarbeitern der Gruppe von Prof. Dr. Wimmer und Prof. Dr. Heber. Zusätzlich zu vielen fachlichen Hinweisen und Diskussionen herrschte stets ein angenehmes Arbeitsklima. Insbesondere danke ich Dr. Lars Berger, Dipl. Phys. Bent Ehresmann, Dipl. Phys. Nina Dresing, Dipl. Phys. Christian Drews, Dipl. Phys. Jan Giesler, Dipl. Phys. Lauri Panitsch, Dipl. Phys. Thies Peleikes und Dipl. Phys. Christoph Terasa für Hinweise, Korrekturen und für zahlreiche Diskussionen zu meiner Arbeit.

Für ihre fachliche Unterstützung beim MSL/RAD Projekt danke ich Dr. César Martin, Dr. Eckart Boehm, Dr. Stephan Böttcher, Dipl. Ing. Lars Seimetz und natürlich ganz besonders Dr. Onno Kortmann, dessen Arbeit die Grundlage für die MSL/RAD Simulationen ist.
Weiterhin danke ich dem Deutschen Zentrum für Luft- und Raumfahrt das diese Arbeit im Rahmen des Projektes #50QX0702 gefördert hat.

Zum Schluss danke ich meinen Eltern und meiner Schwester für ihre Unterstützung und natürlich auch Tine, die sowohl die Höhen als auch Tiefen meiner Arbeit miterlebt hat.

Eidesstattliche Versicherung

Hiermit versichere ich an Eides Statt, dass ich die vorliegende Arbeit abgesehen vom Rat meiner akademischen Lehrer ohne fremde Hilfe und lediglich unter der Verwendung der angegebenen Literatur sowie den bekannten Nachschlagewerken der Naturwissenschaften angefertigt habe, und sie nach Inhalt und Form meine eigene ist. Diese Arbeit ist unter Einhaltung der Regeln guter wissenschaftlicher Praxis entstanden.
Des weiteren versichere ich, dass diese Arbeit weder ganz noch teilweise an anderer Stelle zur Prüfung vorlag. Frühere Promotionsversuche wurden von mir nicht vorgenommen.

Kiel, den

_____

Jan Köhler