
Bayesian spectral and temporal feature inspection in magnetar giant flare SGR 1806-20

Andreas Koch



Munich 2018

Bayesianische Analyse von spektralen und temporalen Charakteristika im großen Magnetarausbruch SGR 1806-20

Andreas Koch

Bachelor Thesis
an der Fakultät für Physik
der Ludwig-Maximilians-Universität
München

vorgelegt von
Andreas Koch

München, den 14. Juni 2018

Erstgutachter: PD Dr. Torsten Enßlin
Tag der mündlichen Prüfung: 5. Juli 2018

Abstract

This bachelor's thesis presents a multidimensional, Bayesian analysis of the magnetar giant flare SGR 1806 – 20. Within the framework of information field theory, utilizing the D4PO algorithm[1], and closely modelling the observational process of the satellite's measuring device, the light curve of the event is reconstructed with regards to its time and energy dimension. Additionally, a reconstruction of the light curve's temporal power spectrum results in the detection of peaks relating to overtones of the neutron star's rotation frequency. In contrast to this, the reconstructed spectral power spectrum is smooth. Lastly, it is shown that each of the light curve's oscillation displays three distinct peaks with a spectral distribution similar to each other.

Zusammenfassung

Diese Bachelorarbeit präsentiert eine multidimensionale, bayesianische Analyse des großen Magnetarausbruchs SGR 1806 – 20. Im Rahmen der Informationsfeldtheorie wird mittels des D4PO Algorithmus[1] und eines genauen Modells des Beobachtungsprozesses des Satelliten die Lichtkurve des Ausbruchs bezüglich Zeit und Energie rekonstruiert. Zusätzlich ergibt die Rekonstruktion von dessen zeitlichem Leistungsspektrums mehrere Maxima, welche den Oberschwingungen der Rotationsfrequenz des Neutronensterns entsprechen. Im Kontrast dazu hat das rekonstruierte energetische Leistungsspektrum einen gleichmäßigen Verlauf. Als Letztes wird gezeigt, dass die spezifischen Maxima in den Oszillationen der Lichtkurve eine ähnliche Energieverteilung aufweisen.

Contents

1	Introduction	1
2	Neutron Stars	2
2.1	Structure	2
2.2	Characteristics	4
3	Signal inference of photon observations	5
3.1	Information Theory	5
3.1.1	Bayes Theorem	5
3.1.2	Information Field Theory	6
3.1.3	Maximum Entropy Principle	8
3.1.4	Many Small Additive Count Processes	10
3.2	D4PO	12
3.2.1	Likelihood	12
3.2.2	Prior Assumptions	13
4	Measurement and Data	18
4.1	Satellite	18
4.1.1	Resolution	18
4.1.2	Instrument Sensitivity	19
4.1.3	Limitations on the Observation	19
4.2	Dataset	21
5	Response Operator	24
5.1	Time Domain	24
5.2	Energy Domain	25
5.2.1	Mapping of Energy Intervals	25
5.2.2	Instrument Sensitivity	27
5.3	Conditions on the signal field	27
6	Validation	29
6.1	Creating Mock Data	29
6.2	Reconstruction	32
7	Results	34
7.1	Reconstruction	34
7.2	Energy Spectra of Features within Signal Map Oscillations	39

8 Conclusion	42
Bibliography	43

1 Introduction

On December 27, 2004, a highly energetic X-ray burst emitted by a neutron star 50,000 lightyears away could be observed as the brightest temporary extra-solar event ever recorded, even surpassing the full moon brightness for 0.2 seconds [2]. This event is known as the giant flare of SGR 1806 – 20. Its analysis utilizing the D4PO algorithm, an information theoretical method for reconstructing signals derived by Pumpe et al. [1] will be the topic for this thesis. The goal is to approximate the event’s light curve with regards to its energy and time, filtering out disturbances in the process.

Neutron stars are interesting objects as they allow us to study physics at extreme conditions. Being the preliminary stage to black holes, neutron stars have a mass density unprecedented by other visible objects in space. They provide us the unique opportunity to gain insight into an environment impossible to reproduce on Earth, thus enhancing our current understanding of physics.

As already shown by Palmer et al. [2], the neutron star’s light curve exhibits temporal oscillations at a distinct frequency. Furthermore, it is expected to send continuous synchrotron radiation displaying a smooth energy spectrum without major features. The final reconstruction proves to be consistent with these expectations. Additionally, each of the light curve’s oscillations displays three distinct peaks, which are going to be analyzed. It is found that the peaks’ spectral distributions are similar to each other.

After a brief introduction to neutron stars, the theoretical background for the D4PO algorithm will be explored. It will be continued by the concrete implementation of a model of the measurement device used for the observation, which is needed for the performed analysis. Subsequently, the validation of the reconstruction and the obtained results will be presented.

2 Neutron Stars

A star of mass larger than $5 M_{\odot}$ forms a neutron star once it abruptly starts fusing carbon or oxygen. Due to a rapid increase in temperature and fusion rate, it then erupts in a supernova. Its core collapses under the gravitational force until eventually it is predominantly made up of neutrons with a small fraction of protons and electrons. The collapse stops when an equilibrium between gravitational pull and neutron degeneracy pressure, which is caused by the Pauli exclusion principle, is reached [3]. Because of the sudden contraction of the core and conservation of angular momentum, the rotation of the core increases manifold and only slows down at a small constant rate over the lifetime of the neutron star. For its further development, the β^- decay and its inverse reaction,



are of importance. The reactants are neutrons n , protons p^+ , electrons e^- , neutrinos ν_e and anti-neutrinos $\bar{\nu}_e$. At first both reactions are likely to happen nearly in equilibrium inside a neutron star. As both processes lead to the creation of neutrinos, which exhibit almost no interaction with other matter, plenty of the neutron star's thermal energy is released in the form of neutrinos. The neutron star's initial temperature of $\approx 10^{11} K$ decreases to $\approx 10^9 K$ during the time interval of only one year [4]. When modeling the neutron star's interior by an ideal zero temperature Fermi gas, the reaction of a neutron to a proton and electron becomes energetically unfavorable and is prohibited. This leads to the dominance of neutrons inside a neutron star.

2.1 Structure

At the neutron star's surface there is much less pressure compared to the interior and thus the density allows for atoms to form. In this environment, the lowest energy state of matter is a collection of ^{56}Fe , which makes up the outer hull of a neutron star [5]. The gravity is immense and leads to the surface being very smooth, with mountains being only as high as several millimeters.

Going inwards, the density increases immensely so that the atoms dissolve to

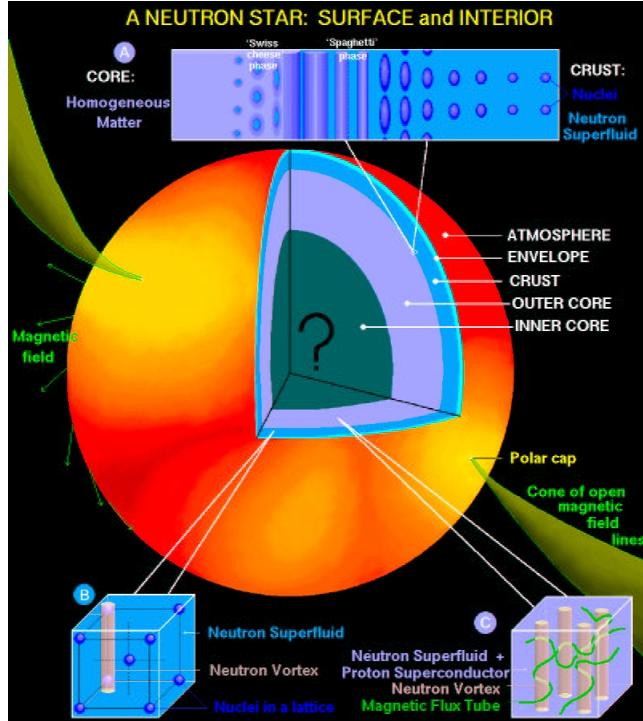


Figure 2.1: Structure of a neutron star displaying different layers that are the result of the increasing density. Its envelope is composed of ^{56}Fe atoms and the crust can be modelled by a neutron superfluid that is transitioning to homogenous matter in the outer core. Figure taken from [6].

form a plasma. As the density increases further, electrons are captured by protons to form neutrons as in reaction 2.2 and at a certain depth with a density found in atomic cores, neutrons start dripping out of the atomic cores. At this stage, the interior might be modelled by a superfluid of neutrons as can be seen in Figure 2.1. This means that the viscosity disappears and the fluid flows frictionless. In the outer core, protons are likely to start forming pairs displaying superconducting behaviour. The density in the inner core increases above the level of atomic cores probably leading to the deconfinement of quarks and other phenomena that are unknown to us. The behaviour of matter under such extreme densities is currently impossible to reproduce with particle accelerators on Earth.

In order to describe a star adequately, besides its pressure, mass, luminosity and temperature gradient, a relation between its pressure and density is needed. This relation is called the equation of state and is not known for neutron stars by reason of theoretical difficulties [7] that are related to the extreme conditions in a neutron star's core. Therefore, new insights gained from analyzing observations of neutron stars might lead to better constraints of the equation of state and

thus would help us to better understand neutron stars. Furthermore, these insights would be beneficial for research in particle physics as well.

2.2 Characteristics

The magnetic field of a neutron star is very important for its detection as a significant fraction of its electromagnetic radiation is emitted by physical processes involving the magnetic field. Charged particles get deflected by the Lorentz force and spiral around the magnetic field lines emitting continuous synchrotron radiation. They can only escape the neutron star along its magnetic field axis, which is aligned at an angle to its rotation axis. For this reason, the radiation is emitted in a cone around the magnetic field axis and is regularly registered as radio wave pulses.

A neutron star's magnetic field strength begins at $B \approx 10^8 G$. This magnitude is partly caused by the conservation of magnetic flux that is defined over the star's original surface area. Once the star's collapse is over, its surface area has decreased by several orders of magnitude and the magnetic flux increases accordingly. A very strong magnetic field of $B \approx 10^{12} - 10^{15} G$ classifies a neutron star as a magnetar [8]. These objects are likely to exhibit starquakes that are caused either by a sudden reconfiguration of its magnetic field or by a sudden slowdown of its rotation. Similar to an earthquake, the magnetar's crust cracks but results in the release of plasma from its core. The plasma gets accelerated by the magnetic field and its charged particles emit high energy photons[9]. Depending on how severe the starquake is, the respective radiation arriving at Earth is in the X-ray or γ -ray regime.

There are many objects that sporadically enter an active mode and emit many short X-ray bursts in succession. For this reason, these objects are referred to as Soft Gamma-ray Repeaters (SGR) and currently, they are explained by the magnetar model. Only three times since 1979, SGRs have been observed emitting giant flares, events with an energy of $\approx 10^3$ higher than regular X-ray bursts [2].

It has often been theorized that neutron stars exhibit periodicities in addition to their rotation frequency. If such periodicities were found, the additional information provided by their properties could be used to constrain models of the neutron star equation of state. Consequently, by adding these constraints the equation of state would become a better approximation to reality.

Now that a rough overview about the physics of a neutron star was given, the theoretical background of the method utilized for the analysis of the giant flare of magnetar SGR 1806 – 20 will be derived.

3 Signal inference of photon observations

3.1 Information Theory

At the heart of information theory lies the extraction of physically meaningful information from observations. By using mathematical measures to quantify this information, its storage and communication becomes possible. Actual applications are signal reconstruction and processing, data compression, error correction, and cryptographic algorithms, although information theory has been used in a wide variety of other fields. It relies on probability theory and statistics, of which the most important principles for this thesis will be explained in the following sections [10].

3.1.1 Bayes Theorem

In probability theory a fundamental concept is a conditional probability defined as

$$P(B|A) = \frac{P(A, B)}{P(A)}, \quad (3.1)$$

with A, B representing statements, $P(A, B)$ the probability of both statements being true, $P(A)$ the probability of only A and $P(B|A)$ the probability of B given A . This concept can be expressed by two basic rules, the sum rule,

$$P(B|A) = P(B, \bar{C}|A) + P(B, C|A), \quad (3.2)$$

and the product rule,

$$P(A, B|C) = P(A|B, C) P(B|C), \quad (3.3)$$

which is equivalent to the definition of a conditional probability. A, B, C are separate statements and \bar{C} is the complement of C , the statement "not C ". With these rules established, two basic mathematical operations can now be

applied to calculate conditional probabilities. The product rule can be used in each direction so that

$$P(A, B|C) = P(B|A, C) P(A|C), \quad (3.4)$$

is also true. Dropping condition C , equating the two expressions and solving it for $P(B|A)$ yields

$$P(B|A) = \frac{P(A|B) P(B)}{P(A)}. \quad (3.5)$$

This equation, known as Bayes Theorem, is the basis of statistical inference [11].

3.1.2 Information Field Theory

Until now, the introduced probabilities $P(x)$ had discrete variables x , whereas from now on s denotes a continuous signal field and $\mathcal{P}(s)$ represents the probability density function dependent on this field. In this context, the statement s implies that the variable s of the probability density function is actually the evaluated signal field, " $s^{var} = s^{val}$ ".

Furthermore, each component of Bayes Theorem is referred to with distinct names used in statistical inference,

$$\mathcal{P}(s|d) = \frac{\mathcal{P}(d|s) \mathcal{P}(s)}{\mathcal{P}(d)}, \quad (3.6)$$

where d is the respective discrete data vector, $\mathcal{P}(s|d)$ is called the posterior, $\mathcal{P}(d|s)$ the likelihood, $\mathcal{P}(s)$ the prior and $\mathcal{P}(d)$ is referred to as the evidence. The likelihood encompasses the way the signal is processed by the instrument and how the noise is accrued to end up with the data. There are several ways how this accrual of noise can happen. A simple data model can be

$$d = R s + n, \quad (3.7)$$

with the noise n being described by an additive term living in data space and a response operator R transforming the signal field to a vector in data space. In our case, the noise is generated by a Poisson process and its probability distribution will be derived later in this chapter. To obtain an adequate signal reconstruction, the noise needs to be filtered out.

The prior incorporates all a priori knowledge about the signal field s before it is observed and must therefore not be dependent on the data d .

In contrast, the evidence is a marginalization over all possible signal fields,

$$\mathcal{P}(d) = \int \mathcal{D}s \mathcal{P}(d, s), \quad (3.8)$$

and hence acts as a normalization factor that only depends on the data d . The pathintegral $\int \mathcal{D}s$ goes over all signal configurations and weights them with $\mathcal{P}(d, s)$ [12].

Put together, these distributions result in the posterior distribution. It gives a probability to every possible signal field realization that might have produced the data. Beginning with the posterior distribution of Bayes Theorem, equation (3.6) may be rewritten as

$$\mathcal{P}(s|d) = \frac{1}{Z(d)} e^{-\mathcal{H}(s,d)}, \quad (3.9)$$

where the information Hamiltonian $\mathcal{H}(s, d)$ and the partition function $Z(d)$ are introduced [13]. The information Hamiltonian is defined as

$$\mathcal{H}(s, d) = -\ln(\mathcal{P}(s, d)) = -\ln(\mathcal{P}(d|s) \mathcal{P}(s)) \quad (3.10)$$

with the partition function representing the evidence,

$$Z(d) = \mathcal{P}(d) = \int \mathcal{D}s e^{-\mathcal{H}(s,d)}. \quad (3.11)$$

By this transformation, many techniques originally developed for quantum mechanics and statistical mechanics may now be used.

The goal of an IFT signal reconstruction is to calculate moments of the signal field averaged over the posterior $\mathcal{P}(s|d)$, to get an estimate of its mean,

$$m = \langle s \rangle_{(s|d)} = \int \mathcal{D}s s \mathcal{P}(s|d) \quad (3.12)$$

and its covariance,

$$\mathbf{S} = \langle ss^\dagger \rangle_{(s|d)} = \int \mathcal{D}s s s^\dagger \mathcal{P}(s|d). \quad (3.13)$$

The mean is of importance as it is the best estimator for s when minimizing the error given by a quadratic loss function [14],

$$\mathcal{L}(m, s) = (m - s)^2. \quad (3.14)$$

If the posterior distribution is single peaked and symmetric, the maximum of the posterior can be used to obtain the posterior mean [1]. For convenience, the

information Hamiltonian can be minimized instead,

$$\frac{\partial \mathcal{H}(s, d)}{\partial s} = -\frac{\partial}{\partial s} \ln(\mathcal{P}(s|d) Z(d)) = -\frac{\partial}{\partial s} \ln(\mathcal{P}(s|d)) \stackrel{!}{=} 0, \quad (3.15)$$

which still yields the maximum of the posterior due to the nature of the negative logarithm. Normally the maximum of the posterior is a good approximation, however this is not always the case, as not every distribution has only one peak and is almost never evenly distributed around its maximum. In those cases, the maximum a posteriori is only in first order a good approximation.

Another option for figuring out a good posterior mean and covariance estimate is by approximating the posterior by a Gaussian distribution and validating its accuracy with an information theoretical measure of choice.

Now that the baselines for a signal reconstruction were established, the statistics for photon observations that lead us to the noise term mentioned earlier and an expression for the likelihood will be derived. For this purpose, an important concept of information theory will be made use of.

3.1.3 Maximum Entropy Principle

This concept is the maximum entropy principle and its goal is to update a known probability distribution $q(x)$ with the new knowledge of a constraint to give the new probability distribution $p(x)$. The entropy function is a measure for the relative information gain of p with respect to q and is defined as the negative information gain,

$$\mathcal{S}[p|q] = - \sum_x p(x) \ln \left(\frac{p(x)}{q(x)} \right). \quad (3.16)$$

As the optimal strategy of updating should imply that a minimum of extra information is incorporated in $p(x)$ and thus the information gain is at a minimum, the entropy needs to be maximized [15].

For this reason, it is beneficial to define the different states of knowledge in following statements:

- I = " $q(x) = P(x|I)$ " represents the already known information
- J = " $d = \langle f(x) \rangle = \sum_x f(x)p(x)$ " with $p(x) = P(x|I, J)$ refers to new information
- Normalization as a probability distribution: $\sum_x p(x) = 1$

With the new information J, the expectation value of an arbitrary function $f(x)$ is given as d and can be used as a constraint on the entropy. The normalization can be used as a constraint as well and the resulting entropy represents an

extreme value problem with given constraints. This problem can be solved by adding these constraints via the Lagrange multipliers λ and μ . The entropy function becomes

$$\mathcal{S}^*[p|q] = - \sum_x p(x) \left[\ln \left(\frac{p(x)}{q(x)} \right) - \lambda - \mu f(x) \right]. \quad (3.17)$$

Maximizing the entropy gives

$$\frac{\delta \mathcal{S}^*}{\delta p(x)} = - \ln \left(\frac{p(x)}{q(x)} \right) + \lambda + \mu f(x) - \frac{p(x)}{p(x)} \stackrel{!}{=} 0, \quad (3.18)$$

which gives an expression for $p(x)$,

$$p(x) = q(x) e^{\lambda-1} e^{\mu f(x)} = \frac{q(x)}{\mathcal{Z}(\mu)} e^{\mu f(x)} \quad (3.19)$$

with

$$\mathcal{Z}(\mu) = \sum_x q(x) e^{\mu f(x)}. \quad (3.20)$$

In the last step, λ was chosen to be $-\ln(\mathcal{Z}(\mu)) + 1$. As a result, μ can now be calculated using

$$d = \sum_x \frac{1}{\mathcal{Z}(\mu)} f(x) q(x) e^{\mu f(x)} = \frac{1}{\mathcal{Z}(\mu)} \frac{\partial \mathcal{Z}(\mu)}{\partial \mu} = \frac{\partial \ln \mathcal{Z}(\mu)}{\partial \mu} \quad (3.21)$$

and with known μ , the partition function $\mathcal{Z}(\mu)$ and the new probability distribution $p(x)$ that minimizes the information added can be calculated as well. Equations 3.16 – 3.20 have kindly been adopted from the lecture on information theory by Enßlin, T. [15]. With this way to update probability distributions given new information, a model to describe the observation process will be derived below.

Neutron stars emit photons with a continuous energy and time resolution. In contrast to this, a satellite's energy and time resolution is discrete, when the photons are summed to give each satellite bin's count. Suppose the photon count of one bin is split into N different processes each containing n_i photons with a known expectation value $\langle n_i \rangle = f$ for all i . The processes will be assumed to be independent as the observed photons usually do not depend on other photons. Additionally, the N different processes resemble the energy and time resolution of the neutron star. Naturally, this is only accurate in the limit for $N \rightarrow \infty$ as the resolution is continuous. This condition will be enforced at the end of the following model.

3.1.4 Many Small Additive Count Processes

First the maximum entropy principle will be applied to one of the N processes as the probability distribution for one process $P(n_i|f = \langle n_i \rangle)$ is desired, given the constraint that the expectation value of n_i is f and the original probability distribution is $q(n_i)$. n respresents the total number of counts and the expectation value of n is $\lambda = \langle n \rangle = fN$

Summing up these statements:

- N = "number of processes"
- $I = n_i \in \mathbb{N}$
- $n = \sum_{i=1}^N n_i = \text{"total number of counts"}$
- $q(n_i) = P(n_i|I) = \text{const.} = q$
- $J = \langle n_i \rangle_{(n_i|J)} = f \text{ for all } i \Rightarrow \langle n \rangle_{(n|J)} = \lambda \equiv fN$

1. calculate $\mathcal{Z}(\mu)$:

$$\mathcal{Z}(\mu) = q \sum_{n_i=0}^{\infty} e^{\mu n_i} = q \sum_{n_i=0}^{\infty} [e^{\mu}]^{n_i} = \frac{q}{1 - e^{\mu}}$$

2. determine μ :

$$\begin{aligned} \frac{\partial \ln \mathcal{Z}(\mu)}{\partial \mu} &= \frac{\partial}{\partial \mu} [\ln q - \ln(1 - e^{\mu})] = -\frac{1}{1 - e^{\mu}} \cdot (-e^{\mu}) \\ &= \frac{e^{\mu}}{1 - e^{\mu}} \stackrel{!}{=} f \\ \Rightarrow e^{\mu} &= \frac{f}{1 + f} \end{aligned}$$

Insert the result into $\mathcal{Z}(\mu)$:

$$\mathcal{Z}(\mu) = \frac{q}{1 - \frac{f}{1+f}} = (1 + f)q$$

3. calculate $p(n_i) = P(n_i|f = \langle n_i \rangle)$:

$$\begin{aligned} p(n_i) &= \frac{q(n_i) e^{\mu n_i}}{\mathcal{Z}(\mu)} = \frac{q \cdot \left(\frac{f}{1+f}\right)^{n_i}}{q(1 + f)} \\ &= \frac{1}{1 + f} \left(\frac{f}{f + 1}\right)^{n_i} = f^{n_i} (1 + f)^{-1-n_i} \end{aligned}$$

Now that the probability distribution for one process is known, the probability $P(n|\lambda, N)$ will be calculated by marginalizing over every n_i ,

$$P(n|\lambda, N) = \underbrace{\sum_{n_1=0}^{\infty} \dots \sum_{n_N=0}^{\infty}}_{\equiv \sum_{n=0}^{\infty}} P(n, n_1, n_2, \dots n_N | \underbrace{\lambda, N}_{I'=J, I}).$$

Considering the independence of processes this can be decomposed to

$$\begin{aligned} P(n|I') &= \sum_{n=0}^{\infty} P(n|n_1, n_2, \dots n_N, I') P(n_1|I') P(n_2|I') \dots P(n_N|I') \\ &= \sum_{n=0}^{\infty} \delta_{n, \sum_{i=0}^N n_i} \frac{1}{1+f} \left(\frac{f}{f+1} \right)^{n_1} \dots \frac{1}{1+f} \left(\frac{f}{f+1} \right)^{n_N} \\ &= \left(\frac{1}{1+f} \right)^N \sum_{n=0}^{\infty} \delta_{n, \sum_{i=0}^N n_i} \left(\frac{f}{1+f} \right)^{\sum_{i=0}^N n_i} \\ &= \left(\frac{1}{1+f} \right)^N \frac{N^n}{n!} \left(\frac{f}{1+f} \right)^n. \end{aligned}$$

For the latter step, the knowledge that there are N^n possibilities to distribute n counts on N processes and $n!$ possibilities to reorder the n counts was used [15].

$$\begin{aligned} P(n|I') &= \frac{f^n}{(1+f)^{N+n}} \cdot \frac{N^n}{n!} \\ &= \frac{(\lambda/N)^n}{(1+\lambda/N)^{N+n}} \cdot \frac{N^n}{n!} \\ &= \frac{\lambda^n}{n!} \cdot \left(1 + \frac{\lambda}{N} \right)^{-N} \left(1 + \frac{\lambda}{N} \right)^{-n} \end{aligned}$$

As the required limit to capture an infinite number of processes ($N \rightarrow \infty$, $\lambda = \text{fixed}$, $f = \lambda/N \rightarrow 0$) is taken, the probability distribution becomes a Poisson distribution:

$$P(n|\lambda, N \rightarrow \infty) = \frac{\lambda^n}{n!} \underbrace{\left(1 + \frac{\lambda}{N} \right)^{-N}}_{\rightarrow e^{-\lambda}} \underbrace{\left(1 + \frac{\lambda}{N} \right)^{-n}}_{\rightarrow 1} = \frac{\lambda^n e^{-\lambda}}{n!}.$$

The equations used in section 3.1.4 again have kindly been taken from the lecture on information theory by Enßlin, T. [15].

The probability distribution for the total number of positive counts of any independent count events evidently results in a Poisson distribution as long as the

number of processes is reasonably high. This is also the case for other distributions with independent processes and an expected rate for every process tending to zero while the number of processes points to infinity. The convergence to a Poisson distribution applies to the statistics of photon observations for a satellite's pixel as well and is the origin for noise in the data as will be explored in more detail in the following section.

3.2 D4PO

The D4PO algorithm provides the possibility to denoise, deconvolve and decompose multidomain photon observations into morphologically different component fields [1]. Furthermore, it can reconstruct fields which are defined over multiple domains. For the use case of this thesis only the reconstruction of a diffuse signal with two domains is needed as the neutron star's emission can only be resolved in the time and energy domains.

As already established, photon count statistics follows a Poisson distribution with its magnitude depending on the average photon count rate λ . When it drops for low count rates, the signal-to-noise ratio drops as well. This poses a limitation to the reconstruction of faint source's signals and has to be considered. Keeping in mind the goal of formulating the posterior distribution, all relevant components will be derived in the following sections.

3.2.1 Likelihood

The likelihood emulates the observational process and thus explains the transition between signal and data space. From now on, the signal field is going to be referred to as the logarithmic photon flux

$$s(t, E) = \ln \left(\frac{\rho(t, E)}{\rho_0} \right), \quad (3.22)$$

with $\rho = \rho(t, E)$ being the photon flux dependent on time and energy. In astronomy and other fields, photon fluxes are positive definite and might vary on several orders of magnitude. For that reason, the logarithmic photon flux will be reconstructed instead with ρ_0 chosen to center s by ensuring $\langle s \rangle_{(s)} = 0$.

Any imaging device recording the photon flux introduces a mapping of the photon flux to the data space, which might exhibit particular features. Those instrumental features need to be considered during a reconstruction and are expressed by a linear operator \mathbf{R} , called the response. How the response operator was built for this application will be explained in Chapter 5. Applying the

response to the photon flux gives the number of expected counts $\boldsymbol{\lambda} = \mathbf{R}e^s$ in case there was no noise. For a given signal s , $\boldsymbol{\lambda} = \langle \mathbf{d} \rangle$ is the expectation value for the data, the actual observed counts. The behavior of each pixel is expressed by

$$\lambda_i = \int_{\Omega_i} dz R_i(z) e^{s(z)}, \quad (3.23)$$

with $\mathbf{z} = z(t, E)$ as the joint time and energy coordinates, Ω being the combined time and energy space and i referring to one pixel. Coming back to Poisson shot noise in observations of photon counts, the observed counts d_i for each pixel corresponds to the total count number n for an additive count model and its probability is therefore $\mathcal{P}(d_i|\lambda_i) = \lambda_i^{d_i} e^{-\lambda_i} / (d_i!)$. For the expected photon count of one pixel λ_i , a random photon count d_i is drawn with the probability for each being determined by the probability distribution $\mathcal{P}(d_i|\lambda_i)$. These create deviations from the expected counts λ_i and thus can be identified as noise, $n_i = d_i - \lambda_i$. As the counts for each pixel is independent of other pixel's counts, the likelihood is the product over all pixels,

$$\mathcal{P}(\mathbf{d}|\boldsymbol{\lambda}) = \prod_i \mathcal{P}(d_i, \lambda_i) = \prod_i \frac{1}{d_i!} \lambda_i^{d_i} e^{-\lambda_i}. \quad (3.24)$$

With an expression for the likelihood, only the relevant prior assumptions remain to be introduced in order to obtain the information Hamiltonian.

3.2.2 Prior Assumptions

As was previously mentioned, the photon flux is positive definite and varies on several orders of magnitude. Additionally, it will be derived to be statistical homogeneous in the next section. Based on our knowledge about the rotation of neutron stars, the flux is expected to show temporal and spectral correlations. Spectral correlations are caused by the fact that synchrotron and other sources of radiation present in the neutron star exhibit correlations over several orders of magnitude in the spectral domain. Both correlation functions, the spectral and temporal, manifest themselves in the logarithmic photon flux's covariance. Consequently, the basic structures for the logarithmic photon flux are derived by applying the maximum entropy principle with known first and second moments to a real field. Further assumptions about the flux's fourier transformed covariance structure include its spectral parameters probability distribution to be logarithmically uniform and fairly smooth. These are the same assumptions as have been made by Pumpe et al. [1] and will be explained below.

3.2.2.1 Logarithmic Photon Flux

The first moment for the flux is known, as $\langle s \rangle_{(s)} = 0$ was already enforced. In contrast, its covariance matrix needs to be modelled by the multiplied covariance matrices of each domain. Essentially, an element in a covariance matrix is the expected product of the corresponding variances of random variables, $\Sigma_{ij} = E[(X_i - E[X_i])(X_j - E[X_j])]$. When using standardized random variables $\frac{X_i}{\sigma(X_i)}$ with $\sigma^2 = \text{var}(X)$, the covariance matrix becomes equivalent to the correlation function. Therefore, the covariance can be described by the product of both domain's correlation functions. For this purpose the coordinate transformation $y = \log(E/E_0)$ with energy E_0 for reference and the usage of combined coordinates $z = (t, y)$ are reasonable.

Thanks to correlation functions only depending on relative differences, as opposed to absolute locations, the statistical properties of any part of the underlying signal field is the same as of any other part. This property is called statistical homogeneity and allows both correlation functions to be diagonalizable in harmonic space using the Fourier transformation. Naming the correlation functions for time and energy $\mathcal{T}_{t,t'}^{(s)} = \mathcal{T}^{(s)}(|t - t'|)$ and $\mathcal{Y}_{yy'}^{(s)} = \mathcal{Y}^{(s)}(|y - y'|)$, the flux's covariance can be represented by

$$\mathcal{S}_{zz'} = \mathcal{T}^{(s)}(|t - t'|) \mathcal{Y}^{(s)}(|y - y'|). \quad (3.25)$$

Knowing both first and second moments of an arbitrary field, the maximum entropy principle yields a Gaussian as the prior distribution on s ,

$$\mathcal{P}(s) = \mathcal{G}(s, \mathcal{S}) = \frac{1}{\sqrt{|2\pi\mathcal{S}|}} \exp\left(-\frac{1}{2}s^\dagger \mathcal{S}^{-1} s\right). \quad (3.26)$$

In the following, the representation of the temporal correlation function $\mathcal{T}_{t,t'}^{(s)} = \langle s_t \bar{s}_{t'} \rangle_{(s)}$ in Fourier space will be calculated as was done during the lecture on information theory by Enßlin, T. [15]:

$$\begin{aligned}
\mathcal{T}_{q,q'} &= \langle s_q \overline{s_{q'}} \rangle_{(s)} = \left\langle (Fs)_q \overline{(Fs)_{q'}} \right\rangle_{(s)} = \left\langle (Fs)_q (Fs)_{q'}^\dagger \right\rangle_{(s)} \\
&= \left\langle (Fs)_q (s^\dagger F^\dagger)_{q'} \right\rangle_{(s)} = \left(F \langle s_t \overline{s_{t'}} \rangle_{(s)} F^\dagger \right)_{qq'} \\
&= (F S F^\dagger)_{qq'} = \left(F_{qt} S_{tt'} F_{t'q'}^\dagger \right) \Big|_{\text{Einstein sum}} \\
&= \int dt e^{iqt} \int dt' \mathcal{T}_{tt'} e^{-iq't'} \\
&= \int dt \int dt' e^{i(qt - q't')} \mathcal{T}^{(s)}(|t - t'|) \\
&\stackrel{\cong}{=} \int dt \int dr e^{i(qt - q'(t-r))} \mathcal{T}^{(s)}(|r|) \Big|_{t'=t-r} \\
&= \underbrace{\int dt e^{i(q-q')t}}_{\cong \delta(q-q')} \underbrace{\int dr e^{iq'r} \mathcal{T}^{(s)}(|r|)}_{P_s(q')} \\
&\stackrel{\cong}{=} \delta(q - q') P_s(q).
\end{aligned}$$

Here q and q' are temporal Fourier space coordinates, the Fourier transformation is represented by $F_{qt} = e^{iqt}$ and the Einstein notation was used to sum over repeated indexes (the coordinates t and t'). $P_s(q) = e^{\tau_{\mathcal{T}}(q)}$ is the Fourier transformed temporal correlation function, the so called temporal power spectrum and $\tau_{\mathcal{T}}(q)$ is called its spectral parameter.

Calculating the spectral power spectrum analogous gives,

$$\mathcal{Y}_{kk'} = F_{ky} \mathcal{Y}_{kk'} F_{k'y'}^\dagger = \delta(k - k') e^{\tau_{\mathcal{Y}}(k)}. \quad (3.27)$$

The power spectra are fully determined by the spectral parameters $\tau_{\mathcal{T}}(q)$ and $\tau_{\mathcal{Y}}(k)$ with harmonic space coordinates q and k . As they are unknown at the beginning of the reconstruction, $\tau_{\mathcal{T}}(q)$ and $\tau_{\mathcal{Y}}(k)$ need to be reconstructed along with the signal. In addition, their covariance structure will be derived by introducing further priors.

3.2.2.2 Priors on Power Spectra and their Covariances

Each power spectrum needs a logarithmically uniform prior since power spectra might vary on several orders of magnitude. For that reason, every element of each power spectrum will be assumed to follow the inverse gamma distribution,

resulting in

$$P_{\text{un}}(\tau | \alpha_l, q_l) = \prod_l \frac{q_l^{\alpha_l - 1}}{\Gamma(\alpha_l - 1)} \times e^{-(\alpha_l \tau_l + q_l e^{-\tau_l})} \left| \frac{d e^{\tau_k}}{d \tau_l} \right|, \quad (3.28)$$

with α_l and q_l being parameters that describe shape and scale of P_{un} and Γ being the gamma function. For $\alpha_l \rightarrow 1$ or $q_l \rightarrow 0$, the magnitudes of τ_l are not constrained as the inverse gamma distribution becomes asymptotically flat on a logarithmic scale [1].

Due to the Fourier transformation, a power spectrum that exhibits fluctuations on arbitrarily small scales in k-space corresponds to a signal field that exhibits correlations over arbitrarily large scales in position space [16]. For many natural processes, which do not exhibit distinct scales, a smooth logarithmic power spectrum can be assumed. This will be enforced by the prior contribution

$$P_{\text{sm}}(\tau | \sigma) \propto \exp \left(-\frac{1}{2\sigma^2} \tau^\dagger T \tau \right), \quad (3.29)$$

in which T is chosen such that it sums the second logarithmic derivative of τ over all Fourier modes and therefore penalizes high curvatures. The magnitude of the smoothing enforcing prior can be controlled by the parameter σ and needs to be figured out by testing on mock data, at which point the power spectrum picks up noise features. σ should then be chosen with a parameter that does not allow noise features in the power spectrum. This could be done to improve the accuracy of our method, since the parameter values have not been validated.

To summarize, the goal of this signal reconstruction is to give an estimate of the posterior mean. A good approximation to this can be the maximum of the posterior. For this reason, the information Hamiltonian consisting of the likelihood and priors will be minimized. All prior parameters that were introduced during this section and their relation to the signal field and the data are illustrated in Figure 3.1.

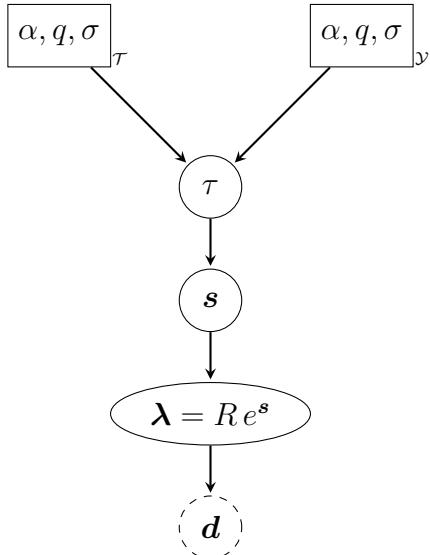


Figure 3.1: Hierarchical parameter model showing all required parameters of the prior assumptions. α and q correspond to the shape and scale of the prior on the power spectrum amplitudes and σ to the strength of the smoothness enforcing prior. Both priors are about spectral parameters τ that specify the covariance of the process that produced the signal field. The signal field is processed by the satellite with the response operator R to give the expected photon count λ . The data d is drawn from a Poisson distribution corresponding to this expectation and is depicted in a dashed circle. Parameters in boxes have to be specified by the user and black solid circles are quantities reconstructed by the algorithm. Figure adopted from [1].

4 Measurement and Data

4.1 Satellite

The Rossi X-ray Timing Explorer (RXTE) was a satellite to observe black holes and neutron stars. It had five photon counter modules onboard, each containing several layers of gas and anode wires as can be seen in Figure 4.2. Whenever a photon impacts a gas volume, several atoms in it get ionized in a chain ionization as the energy of X-ray photons are much higher than the difference of energy states in the gas atoms. That creates an ionization cloud which drifts to the anode wires where an electron avalanche is created. Once collected at the anode wires, the electron pulse is amplified, shaped and converted to a digital pulse height.

4.1.1 Resolution

The aforementioned conversion process takes approximately $18.2 \mu\text{s}$ for all photon counter units, and is essentially independent of energy [17].

In every photon counter unit, each photon is detected in a distinct energy bin at a specific time. The events of each counter unit are passed on to event analyzers that process the data. They are mapping the strength of the event to 256 energy channels in a specific way for each measurement device, utilizing different modes. For this observation, only the "GoodXenon" mode was used to analyze the data [9]. In this mode the data processing time resolution was $\sim 1 \mu\text{s}$ and the spectral resolution was $\sim 0.1 \text{ keV}$ covering the nominal range of $\sim 2 - 120 \text{ keV}$. In total the limiting time resolution is the conversion process mentioned earlier as it dominates the error. Even though the registration of photons is happening in separate energy bins, the effective energy resolution of the combined instrument can be assumed to be higher than the amount of available pixels of one detector because the different instrument's energy bins overlap. For computational reasons explained in section 5.3, the pixel size of the signal in its energy domain will be set to 256. The assumed limiting energy resolution is therefore $\sim 0.44 \text{ keV}$.

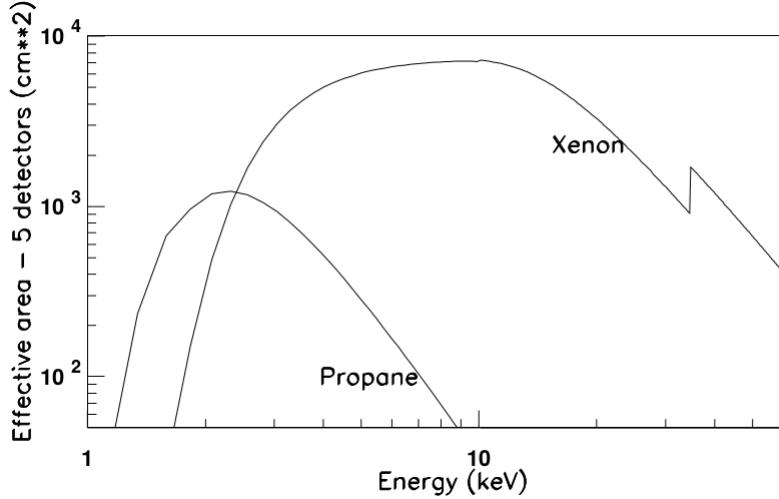


Figure 4.1: Effective area for the energy spectrum of the photon counters with a range of $\sim 1 - 50$ keV. Figure taken from [18]

4.1.2 Instrument Sensitivity

Every measurement device typically has a function which captures the instrument's sensitivity to turn the observed quantity into actual data. With photon observations, this would be done via an effective area function defined over the whole energy spectrum. It models the area over which the photon flux is observed. This function is measured frequently during the lifetime of the satellite. For the satellite RXTE the effective area is depicted in Figure 4.1, but only for a range of $\sim 1 - 50$ keV. This does not cover the whole energy spectrum of the satellite's instruments and thus will not be used for the reconstruction.

4.1.3 Limitations on the Observation

Although the satellite did not observe the source directly as it was inclined at ~ 30 degrees to it, the giant flare SGR 1806 – 20 was still detected. It is worth noting that a modification of the original spectrum likely happened while the photons propagated through the passive material in the satellite case [9]. Of further importance is that one photon counter unit lost its propane layer multiple years before the relevant event occurred. This leads to a general increase of unwanted detections of photons that originated from other astronomical objects and are called background events. In comparison to the dominant signal of the giant flare burst this effect becomes negligible and thus will not be considered.

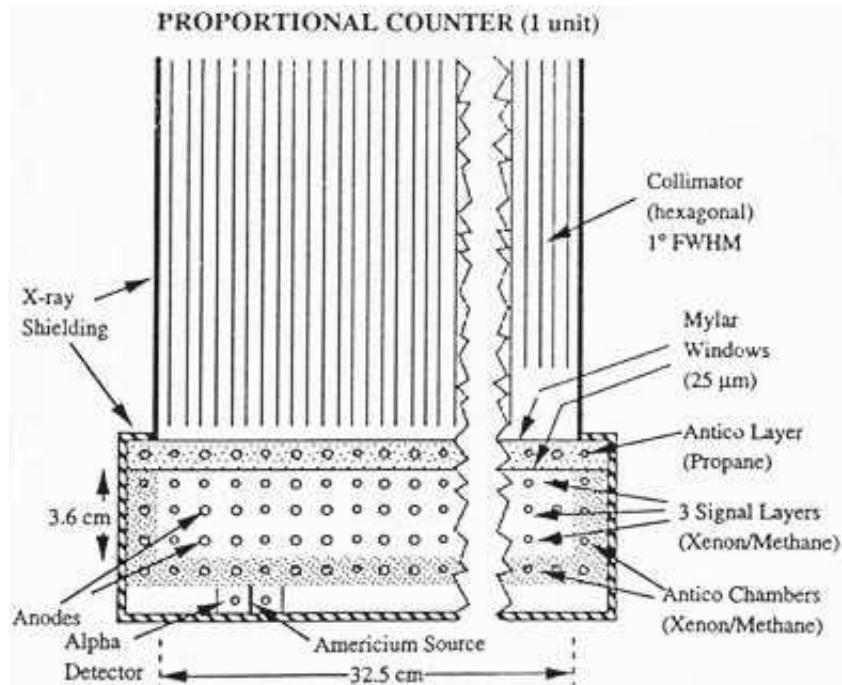


Figure 4.2: Photon Counter Module Cross Section View. The field of view is defined by the collimator which allows the alignment on the target. Below the collimator, there is a mylar window above and below a propane layer. The mylar is a gas that acts as a lower boundary for the observed photons at $\sim 2\text{ keV}$. The main Xenon-filled detector volume consists of 3 layers of signal anodes and a back layer of anodes. Together with the propane layer, they absorb and thus detect photons [17]. Figure applied from [19].

4.2 Dataset

Numerous characteristics of the dataset are of relevance. At the beginning of the giant flare, the instrument telemetry was saturated due to its limited capacity to register photons. Therefore, several time intervals which show no detection of any photons whatsoever exist and will be referred to as dead times. During the process of reconstruction this saturation manifested itself as an obstacle, because no features besides the first peak were identifiable. Hence, the analysis of the signal was restricted to start at ~ 10 s after the initial spike.

In total, $\sim 1,000,000$ photons were detected by three of the 5 instruments and the majority had energies up to ~ 70 keV. The distribution of the energy channel's events of each instrument can be seen in Figure 4.3. As each energy channel theoretically covers a spectral range of $\sim 0.43 - 0.45$ keV, the diagrams represent an approximation for each instrument's spectral distribution of events.

The characteristic energy distribution of each instrument also includes distinct upper energy limits, which differ for the relevant instruments. The upper energy limit for the reconstruction will be set to the lowest upper instrument limit so that no edges become visible due to suddenly using only fewer instruments and thus having less photon counts. An illustration of this effect can be found in Figure 4.4.

In addition, the uppermost energy bin of each instrument exhibits incongruities with the rest of the data. Its energy channels show a sharp increase in comparison to the preceding energy channels, possibly caused by the detection of high energetic photons that would be out of range otherwise. This results in the formation of an edge in the reconstructed signal, which is presented in Figure 4.5. For this reason the uppermost energy bin will also be neglected.

In conclusion, the utilized data will begin at ~ 10 s, lasting until the flare is over at ~ 400 s and its spectral range will be between 2 – 114.6 keV.

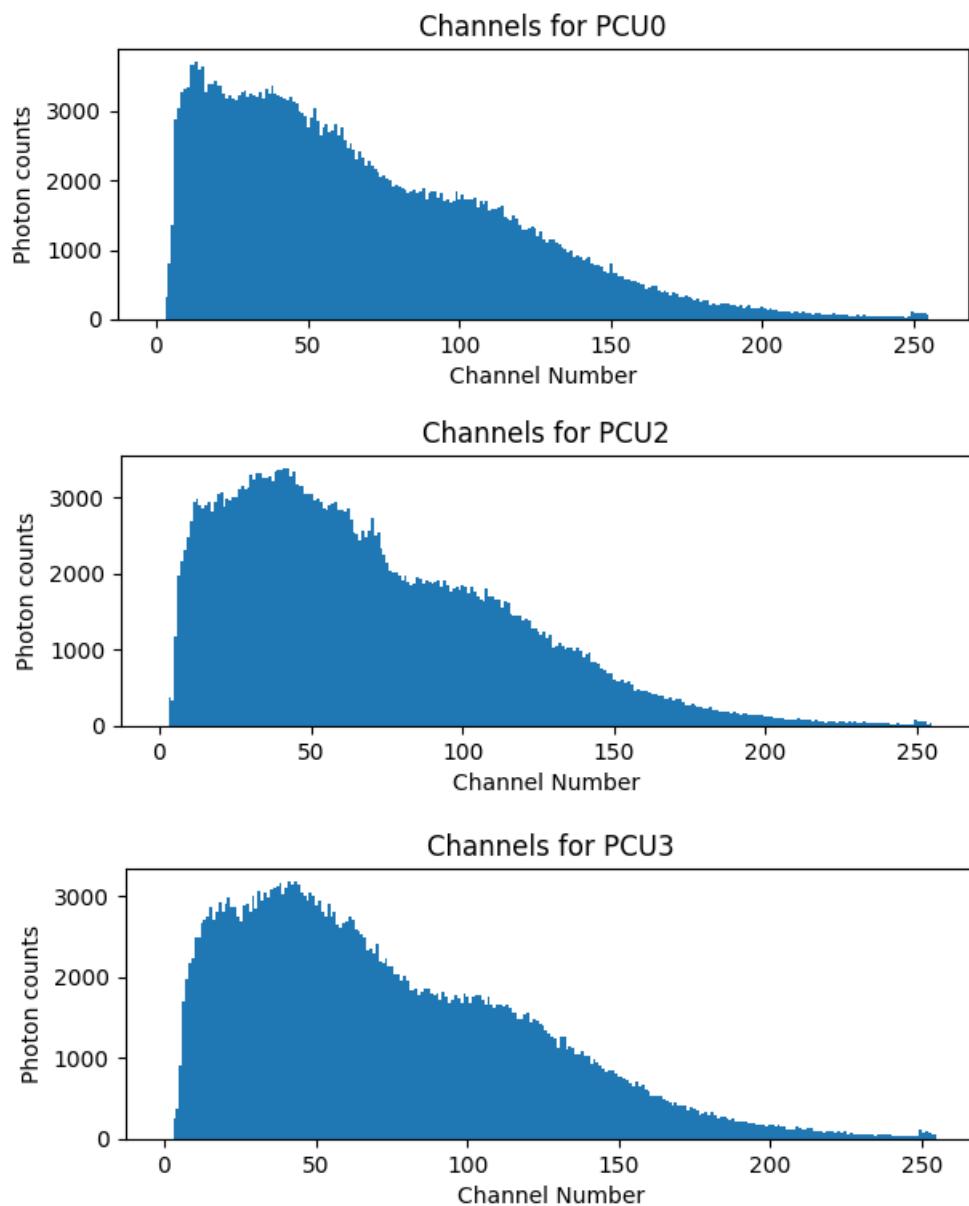


Figure 4.3: The distribution of photons over all energy channels for each instrument summed over the time dimension.

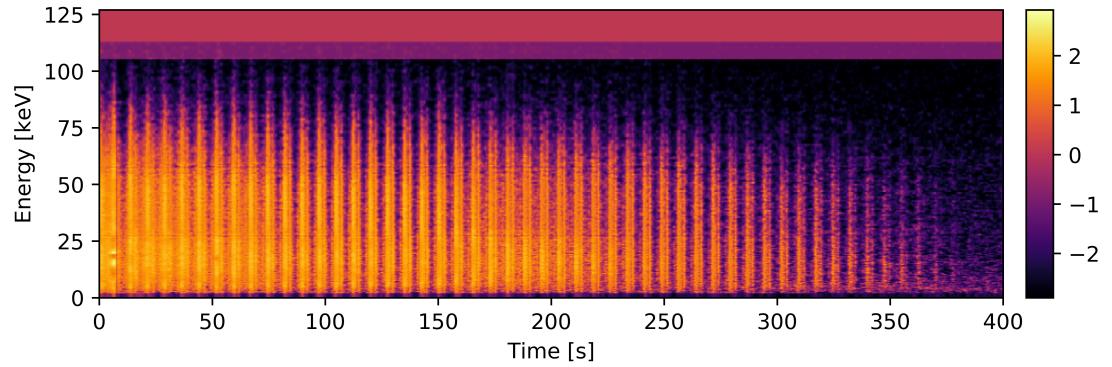


Figure 4.4: The reconstruction of the data without the last 6 energy bins yields clearly visible edges at ~ 105 keV and ~ 113 keV. The results are consistent with the limits at 105.19 keV for two instruments and one at 112.9 keV for the remaining instrument.

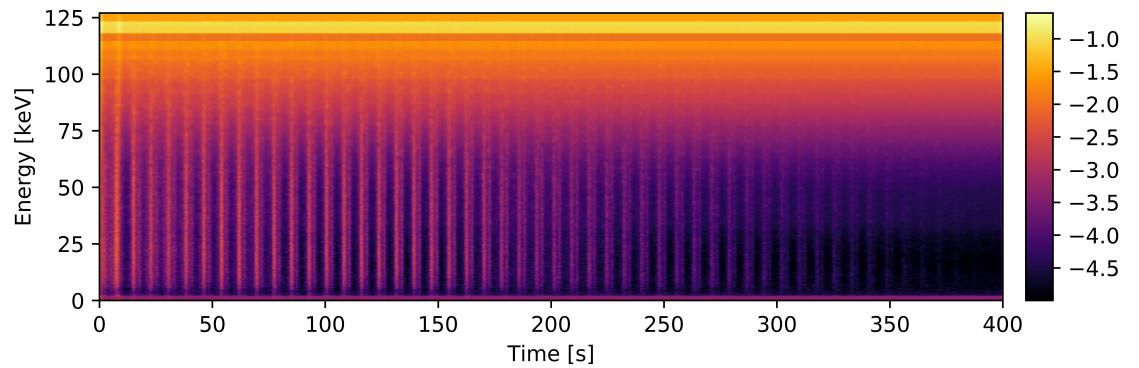


Figure 4.5: The reconstruction of the data showing the whole energy spectrum. In the upper energy spectrum, edges can be found at the borders of the last energy bin of two instruments, ranging from ~ 114.6 keV to ~ 117.9 keV. A similar edge is visible at the beginning of the remaining instrument's last energy bin.

5 Response Operator

The response is quite important for a signal reconstruction as it accurately emulates the measurement devices used for observation. In the previous chapter, it was discussed how the instruments work and what their limitations regarding resolution and range of data are. Considering the limitations, it can be concluded that the upper limit for the dimensionality of the signal map, which is reconstructed, is 256 pixels for its energy domain and $\sim 21,4 \times 10^6$ pixels for its time domain. A response operator maps every pixel of the photon flux ρ_i that is represented by a discrete signal map ρ , to every pixel of the expected observed counts λ_i . As they both have the same dimensionality, an explicit representation of the response operator as a matrix in computer memory becomes impossible with regular computers. For this reason, the response operations will be defined implicitly by functions that model the mapping efficiently as discussed below. The response will be implemented via a sparse matrix of the python package `scipy` [20]. It is built by defining each non-zero matrix component explicitly and memorizing them along with the associated indices. For this purpose, the respective fields need to be flattened. Once the response operator is applied, they will be unflattened to end up with the desired shape for the field. Every component that is not explicitly defined is automatically assumed to be zero and no memory will be allocated for them.

5.1 Time Domain

Earlier, the concept of dead times due to saturation of the instrument was mentioned. Time intervals with no recordings by the satellite could either be attributed to no photons arriving at the satellite or to having reached maximum capacity. As is already known, the saturation of the satellite only occurred during the first 10 s and shortly afterward. Accordingly, the longest time interval during this time will be used as a benchmark with an error margin of 10 percent. Using this method, only one dead time was detected. This feature will be incorporated in the response by an identity operator, called time mask. It features ones on its diagonal, with the dead time interval being the only exception using zeros instead. During the reconstruction, the signal map will be continued in congruence with the rest of the signal map, but with great uncertainty,

connecting both endpoints of the interval. This means that the signal map is reconstructed using the current reconstruction of its correlation function, e.g. its power spectrum [21].

As the measurement devices exhibit no influence on the data in the time domain besides its resolution, the mapping is done linear with no scaling factor. Therefore, an identity operator is enough to capture the temporal behaviour of the instrument. As the time mask is already an identity operator except for the dead times, it will be used as the temporal response operator.

5.2 Energy Domain

5.2.1 Mapping of Energy Intervals

In comparison to the time domain, mapping photons to their energy domain is more complex. An arbitrary photon could be detected in either of the three instruments. The energy bin it falls into is determined only by the photon's energy. Lastly, as the energy bins each relate to one or more energy channel, a photon could be detected in either energy channel related to its energy bin. As there is no knowledge available how to accurately distinguish between channels or instruments, their relative probabilities were calculated using the total available detections in the dataset.

The response operator can be derived by examination of each of the three processes mentioned in the last paragraph. The first process is captured by a multiplication of the discrete signal map with the probability for each instrument creating a new fork for each instrument,

$$\boldsymbol{\lambda} = \begin{pmatrix} \alpha_1 \mathbf{R}_1 \\ \alpha_2 \mathbf{R}_2 \\ \alpha_3 \mathbf{R}_3 \end{pmatrix} \boldsymbol{\rho}, \quad (5.1)$$

where α_i denotes the scalar probability for a photon to be recorded by instrument i and \mathbf{R}_i refers to each instrument's individual response matrix. As only the energy domain is regarded, the discrete signal map $\boldsymbol{\rho}$ can be identified with a vector that is multiplied with a matrix. The total matrix depicted above contains the individual matrices \mathbf{R}_i stacked vertically on top of each other.

The other two processes each are portrayed by a matrix. Every individual response matrix \mathbf{R}_i therefore contains two matrices. The first matrix \mathbf{R}_{i_b} is performing the iteration over all energy bins, comparing the signal map bins ρ_i

with the energy bins e_i . This is exemplified by

$$\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0.3 & 0 & 0 & 0 \\ 0 & 0.7 & 1 & 0.1 & 0 \\ 0 & 0 & 0 & 0.9 & 1 \end{pmatrix}}_{R_{ib}} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \\ \rho_5 \end{pmatrix}. \quad (5.2)$$

All counts of the discrete signal map's bins that are located in one energy bin will be added to the photon count of that energy bin. If a signal map bin is split by the edge of an energy bin, the corresponding fractions of the signal map bin's counts will be calculated and added to the corresponding energy bins. The same applies if a signal map bin were to be broader than an energy bin.

The second matrix \mathbf{R}_{ic} in the construction of the individual response matrix \mathbf{R}_i is representing the connection between energy bins e_i and channels c_i . The probabilities for each energy channel are calculated by a summation of all counts arriving in this channel and normalizing it by the total counts arriving in the respective energy bin. For this reason, there will be no overlap of two energy bins mapping to one energy channel, as can be seen in the next example,

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0.6 & 0 \\ 0 & 0 & 0.5 \\ 0 & 0 & 0.3 \\ 0 & 0 & 0.2 \end{pmatrix}}_{R_{ic}} \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}. \quad (5.3)$$

Put together, the preliminary response operator \mathbf{R}_{map} is represented by

$$\lambda = \underbrace{\begin{pmatrix} \alpha_1 \mathbf{R}_{1c} \mathbf{R}_{1b} \\ \alpha_2 \mathbf{R}_{2c} \mathbf{R}_{2b} \\ \alpha_3 \mathbf{R}_{3c} \mathbf{R}_{3b} \end{pmatrix}}_{\mathbf{R}_{map}} \rho, \quad (5.4)$$

with α_i referring to the scalar probability for a photon to be recorded by instrument i , \mathbf{R}_{ic} to the matrix for mapping energy bins to channels and \mathbf{R}_{ib} to the matrix that maps signal bins to energy bins.

5.2.2 Instrument Sensitivity

As there is no adequate effective area function available to us that gives information about the satellite's sensitivity at each energy channel, a smoothed version of the distribution of photons per energy channels summed over the whole burst, as seen in Figure 4.3, was used as an approximation of the sensitivity. This is comparable to the assumption that the source radiates at a constant magnitude at each point of the energy spectrum. As a constant vector, the energy sensitivity is not denoised by the D4PO algorithm. For this reason, the smoothing is necessary to counter the Poisson shot noise that is inherent to the distribution. To conclude, every component of the preliminary expected counts λ_j gained by application of \mathbf{R}_{map} to the signal map, with j referring to its energy domain is multiplied with the respective sensitivity factor g_j . Therefore, the final form of the spectral response operator is

$$\boldsymbol{\lambda} = \underbrace{\mathbf{g} \odot \mathbf{R}_{map}}_{\mathbf{R}} \boldsymbol{\rho} \quad (5.5)$$

where \mathbf{R}_{map} refers to the response operator derived in the last section, \mathbf{g} to the constant sensitivity vector and \odot is used to illustrate an element-wise multiplication. Now that every part of the final response operator \mathbf{R} was explained, further conditions that influence the shape of the discrete signal map $\boldsymbol{\rho}$ will be established.

5.3 Conditions on the signal field

The reconstruction algorithm makes use of the Fast Fourier Transformation (FFT) algorithm to transform the signal field's discrete representation to get its respective power spectra as explained earlier. It allows for a very efficient transformation between position and phase space calculating the Fourier transform in a recursive manner. Meanwhile, it divides the values of a discrete function by half at each recursion step. For this reason, the shape of the function to be Fourier transformed must be an exponent of 2 to utilize the FFT algorithm. With this knowledge, it becomes apparent why the signal map was chosen to have an energy dimensionality of $2^8 = 256$.

The Fourier transformation and its inverse used in the FFT algorithm can be represented by

$$\tilde{f}_k = \sum_{j=0}^{n-1} f_j e^{ijk2\pi/n} \quad (5.6)$$

and

$$f_j = \frac{1}{n} \sum_{k=0}^{n-1} \tilde{f}_k e^{-ijk2\pi/n}, \quad (5.7)$$

with f_j meaning component j of an arbitrary function, \tilde{f}_k the k th component of its Fourier transform, i being the imaginary unit and n being the number of pixels [22]. From these equations, it is evident that the first and last element of the function are equal as,

$$f_n = \frac{1}{n} \sum_{k=0}^{n-1} \tilde{f}_k \underbrace{e^{-ik2\pi}}_{=1} = f_0. \quad (5.8)$$

Consequently, the signal map ends up having periodic boundary conditions for both time and energy dimension as we use a Fourier transformation for both domains.

The signal map is framed by zero-padding areas with a thickness of half the signal map, in order to counteract these periodic boundary conditions. The response operator does not map any of the padded elements to the expected counts map λ . For this reason, the reconstruction algorithm continues the signal map from its borders through the padding areas. As a result the enforced boundary conditions do not influence the reconstructed signal map significantly.

6 Validation

6.1 Creating Mock Data

In order to use the method for reconstructing signals explained in the previous chapters, a validation using test data is necessary. We emulate a typical signal map we would expect from the observed source by constructing adequate power spectra with a peak in the temporal spectrum. This leads to the signal map exhibiting temporal oscillations. When the signal map is processed by the response, the so called mock data is drawn randomly from a Poisson distribution and emerges with the same temporal oscillations as the signal map. The mock power spectra are chosen to be

$$e^{\tau\tau(\omega)} = \frac{a_\omega}{(1 + (\omega/\omega_0)^2)^2} \quad (6.1)$$

$$e^{\tau y(k)} = \frac{a_k}{(1 + (k/k_0)^2)^2}, \quad (6.2)$$

with a_ω , a_k , ω_0 and k_0 being constants. The power spectra are shown in Figures 6.1 and 6.2.

The peak in the temporal power spectrum is constructed by multiplying one element at 0.02 [1/s] of $e^{\tau\tau(\omega)}$ with the factor 20. All parameters for building the mock power spectra, the temporal peak and thus the mock signal map as well, are chosen empirically. The peak was introduced at a frequency that produces a plausible oscillation in the signal map, which can be seen in Figure 6.3 along with the resulting data in Figure 6.4.

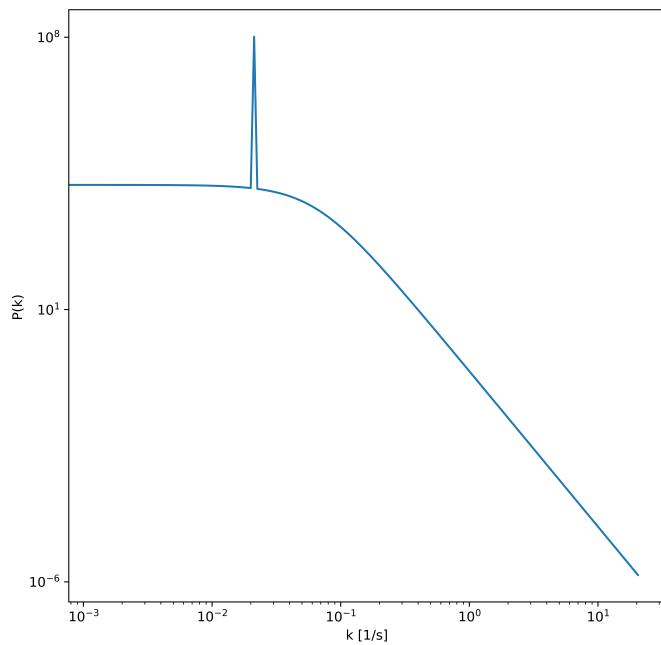


Figure 6.1: Temporal mock power spectrum with clearly visible peak at $0.02 [1/s]$. The power spectrum obeys equation 6.1 except for the peak.

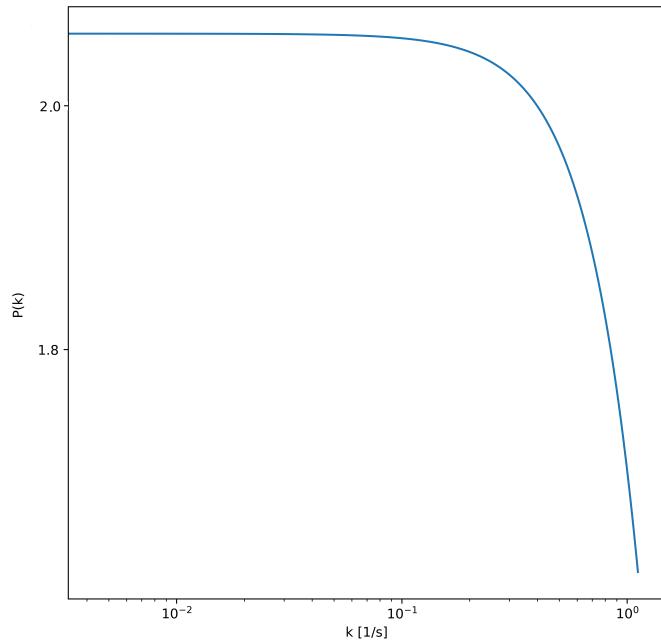


Figure 6.2: Energy mock power spectrum obeying equation 6.2.

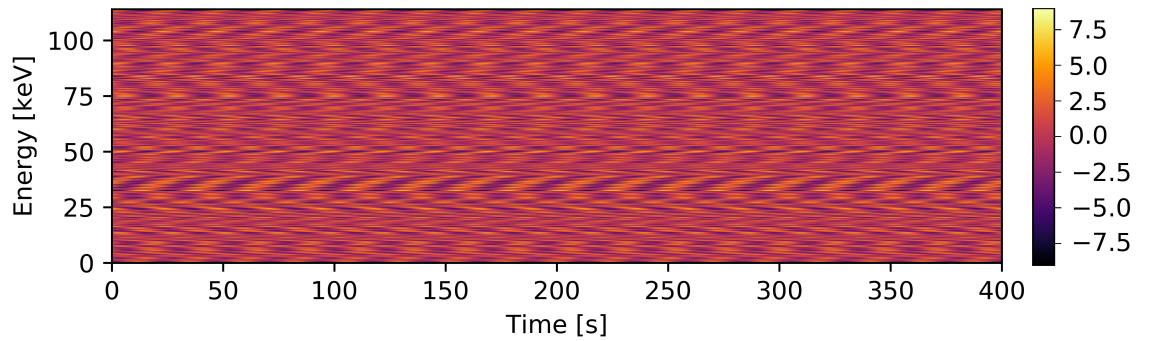


Figure 6.3: Mock signal map with apparent oscillation regarding its time dimension. These oscillations are the result of the introduced peak in the according temporal mock power spectrum.

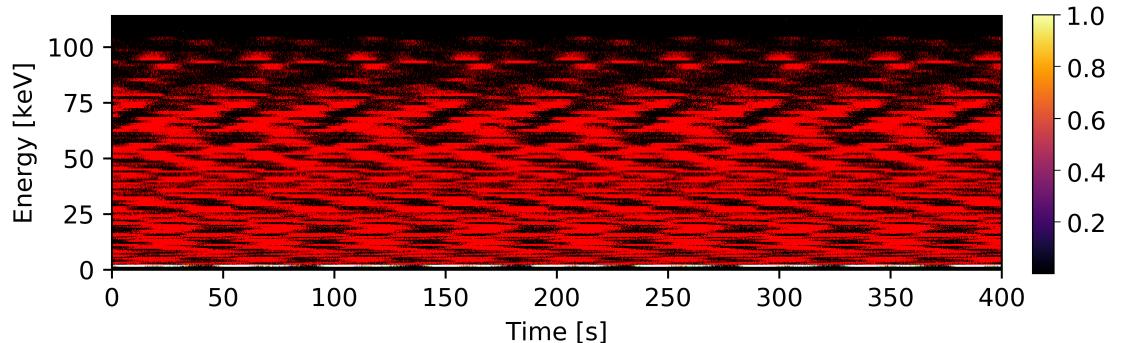


Figure 6.4: Mock data exhibiting partly detectable temporal oscillations.

6.2 Reconstruction

The goal of this chapter is to prove that correlations in the reconstructed signal map are correctly identified. If a feature were found in the power spectrum, a proof to its integrity would be vital. In the last section, the creation of mock data displaying temporal oscillations following the addition of a peak in the mock signal map's temporal power spectrum was discussed.

If there were no knowledge about this oscillation and thus the initial guess for the temporal power spectrum did not display any peak, the recovery of a peak at the exact location as it was placed in the temporal mock power spectrum would suggest the success of this reconstruction. Hence, our initial guesses for the power spectra are going to be equivalent to the mock power spectra, except for the peak. Since the recovery of features in the power spectra are more important for further analysis than the reconstructed signal map, the initial guess for it is going to be the mock signal map.

As a result of this test, the reconstructed temporal power spectrum shows a peak at position $\sim 0.02 [1/s]$ and is depicted in Figure 6.5. The location is consistent with the peak's original position and thus can be regarded as a proof of this method. It is evident that the range of the reconstructed power spectrum's power is physically incorrect as it would normally be expected to vary on several orders of magnitude. This characteristic can be attributed to an error in the employed reconstruction algorithm, possibly due to false volume factors used for normalization of power spectra. Power spectra are defined as

$$P_s(k) = \frac{\langle |\mathbf{s}_k|^2 \rangle}{V}, \quad (6.3)$$

with the Fourier transformed covariance $S_{kk'} = \langle |\mathbf{s}_k| \rangle$ and the volume factor V representing the signal map's volume. Although this trait confines the scope of the reconstruction, relative differences in the reconstructed signal map and power spectra should remain unaffected. Therefore, correlations of the signal map and features observable in the power spectra are still eligible for profound analysis.

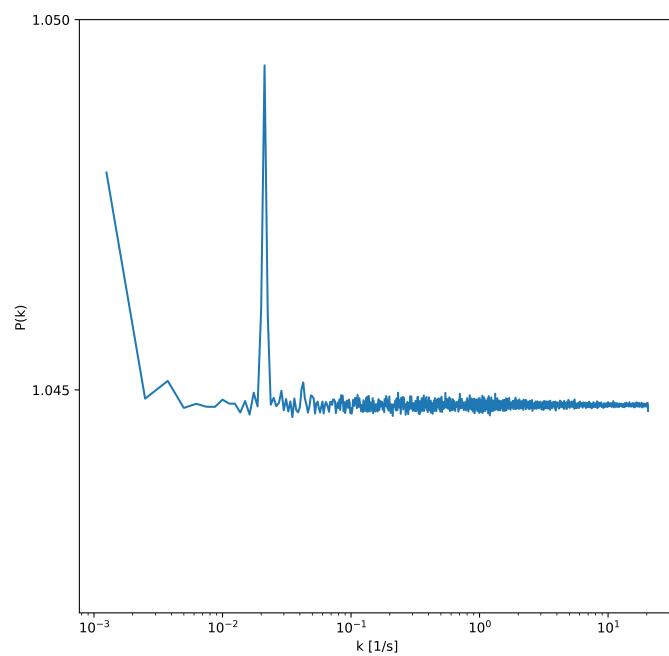


Figure 6.5: Reconstructed temporal power spectrum depicting an observable peak at location $\sim 0.02 [1/s]$. The range of the power spectrum's power is physically incorrect and is likely caused by false volume factors used as a normalization for the power spectra during the reconstruction.

7 Results

To obtain a proper reconstruction of the signal map and its power spectra, several approaches with different parameters have been made. These include the usage of the satellite energy channel sensitivity, which will be referred to as sensitivity mask from now on, and the smoothness parameter of the smoothness enforcing prior that affects the power spectra. The relevant differences of the signal map and power spectra emerging from these approaches will be discussed below during the first section. Afterwards follows an analysis about the energy spectra of features within oscillations of the signal map.

7.1 Reconstruction

With the satellite's resolution and range established in previous chapters, a signal map with the shape of 2^8 energy pixels and a temporal dimensionality of 2^{14} were selected. A higher resolution is possible, but leads to higher computational cost and thus to more time being needed for computation. The reconstructed signal map without the sensitivity mask and smoothness parameters of $\sigma_0 = 10^4$ and $\sigma_1 = 10^2$ is illustrated in Figure 7.1. The reconstruction shows clearly visible temporal oscillations and a smooth energy spectrum similar to the distribution in Figure 4.3. Comparable to the physically incorrect power range seen in Figure 6.5, the range of photon counts also diminishes. This can be attributed to the same cause of incorrect volume factors, as the reconstructed signal map's range is closely connected to the range of its power spectra.

The same reconstruction, but with the usage of the sensitivity mask, is depicted in Figure 7.2. This leads to blurring effects at the start of the energy spectrum and predominantly seen for higher energies, beginning at ~ 30 keV. An explanation for this behaviour might be that the sensitivity mask includes a lower sensitivity for the mentioned segments of the energy spectrum. Therefore, only a small amount of the signal map's magnitude is being recorded. As a result, a little expected photon count rate λ_i implies a much larger signal strength at the respective positions. For this reason, all segments with a lower sensitivity are amplified independent of them belonging to a temporal oscillation or a segment in between. Naturally, as the photon counts of pixels in oscillations demagnify at higher energies, their signal map strengths are on similar scales as the pixel's signal map strengths located in non-oscillations.

In Figures 7.3 and 7.4, the power spectra of the initial reconstructed signal without sensitivity mask are displayed. The temporal power spectrum includes several spiky features on small scale that would relate to correlations over great scale in the signal map. As none such correlations were observed, these features are likely artefacts that would have been compensated for if the reconstruction had continued until full convergence. Thus they are assumed to be of no relevance.

A more detailed depiction of the subsequent peaks are given in Figures 7.5 and 7.6, with Figure 7.6 being the result of a weaker smoothing by parameters $\sigma_0 = 10^5$ and $\sigma_1 = 10^4$. The location of every major peak is specified in harmonic space and its value is noted on the axis. The harmonic space of the power spectrum has a resolution of $\Delta q = 0.00125$ [Hz]. Consistent with our knowledge about the rotation frequency of the neutron star, which is at $q_0 = 0.13249$ [Hz] [21], the most dominant peak is recovered at $q = 0.131 \pm 0.00125$ [Hz]. Further peak positions q [Hz] are recorded in Table 7.1 and 7.2 as multiples of the rotation frequency, called overtones. Their uncertainty is $\sigma_q = \Delta q / q_0 \approx 0.01$ for both smoothing configurations. In harmonic space, the occurrence of overtones of oscillation frequencies is to be expected as any multiple fulfills the conditions set by the oscillation. The wave described by any overtone is at a maximum during every oscillation maximum and at a minimum for the oscillation's minima. Therefore, any peak that is a multiple of the rotation frequency can be attributed to the rotation as well. Apparently, every observed peak is an overtone of the rotation frequency and lies well within the uncertainty of the expected location.

Lastly, the energy power spectrum is smooth and shows no distinct features. In the next section the spectral distribution of the three subparts of the oscillations in the signal map are going to be discussed.

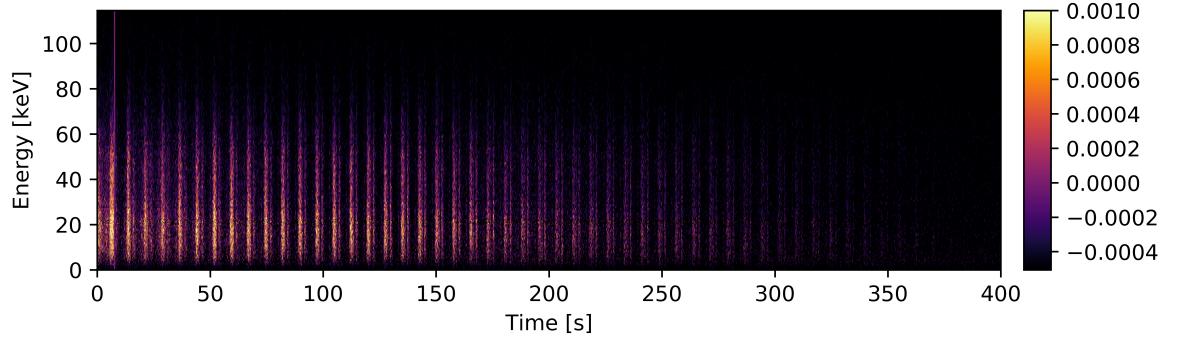


Figure 7.1: Reconstruction of signal map without sensitivity mask, temporal dimensionality of 2^{14} and smoothness parameters of $\sigma_0 = 10^4$ and $\sigma_1 = 10^2$. The temporal oscillation is evident as is an artefact at ~ 10 s. The artefact might be caused by an error in the response operator.

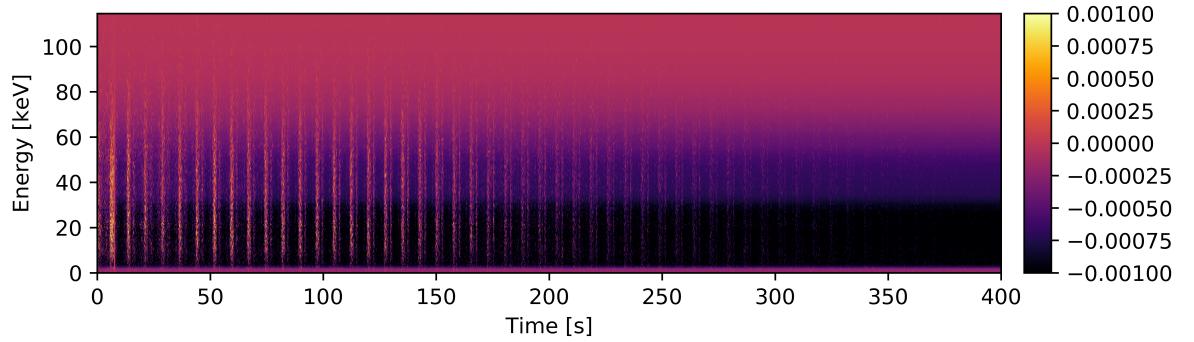


Figure 7.2: Reconstructed signal map with sensitivity mask, temporal dimensionality of 2^{14} and smoothness parameters of $\sigma_0 = 10^4$ and $\sigma_1 = 10^2$. The same artefact at ~ 10 s as in Figure 7.1 is apparent. Due to the sensitivity mask suppressing the observed energy spectrum at the lowest and highest energies, only a small amount of the signal map's magnitude is being recorded there. Thus the little expected photon counts that arrive in this part are magnified. This effect is likely the cause for the washed-out segments of the signal at higher energies.

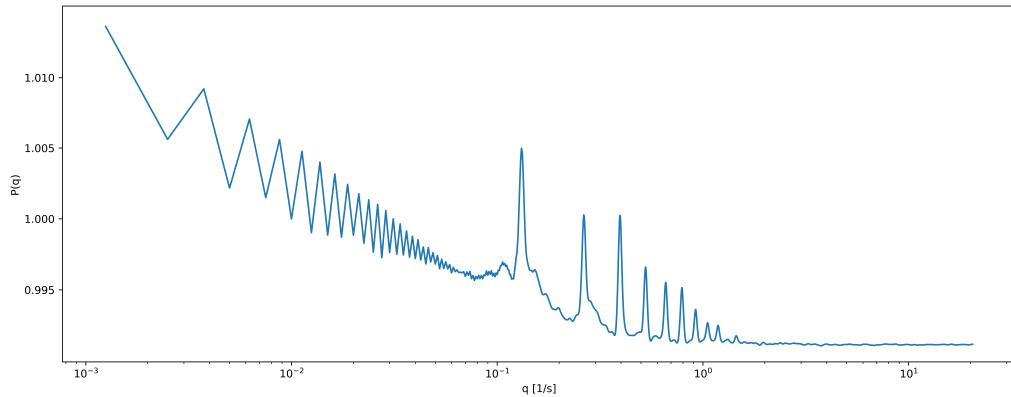


Figure 7.3: Reconstruction of temporal power spectrum with a temporal dimensionality of 2^{14} and smoothness parameters of $\sigma_0 = 10^4$ and $\sigma_1 = 10^2$. In Figures 7.5 and 7.6, the frequency peaks are analyzed.

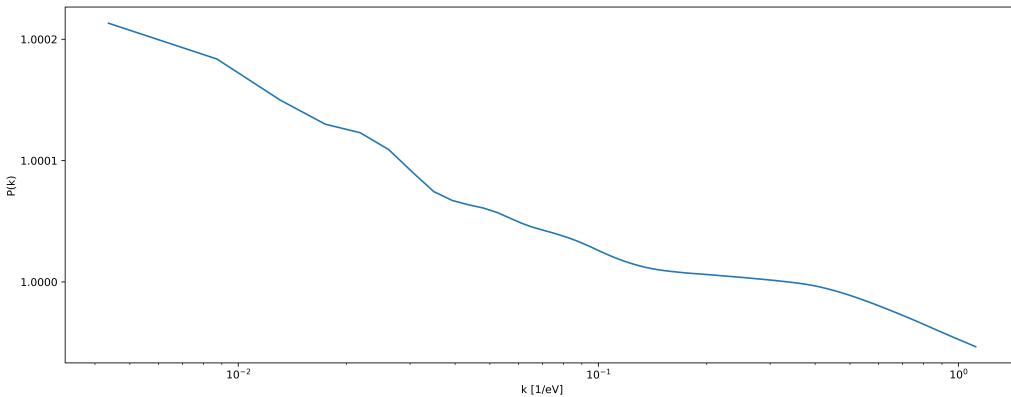


Figure 7.4: Reconstructed spectral power spectrum with a temporal dimensionality of 2^{14} and smoothness parameters of $\sigma_0 = 10^4$ and $\sigma_1 = 10^2$.

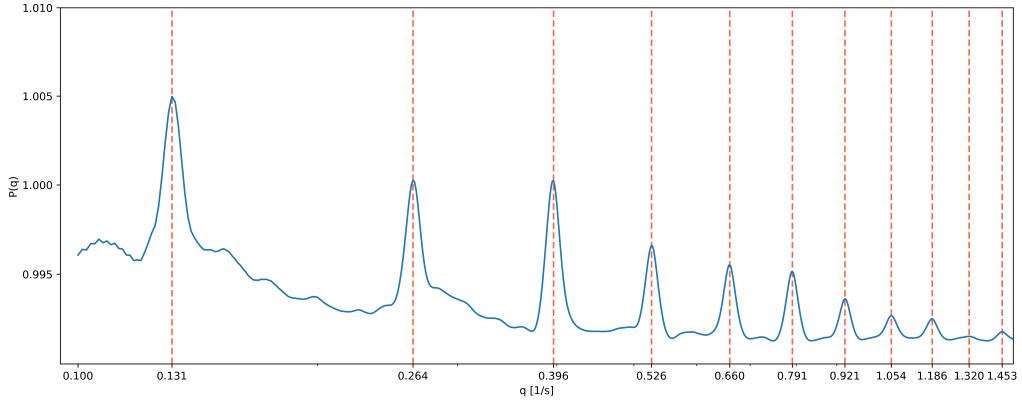


Figure 7.5: Reconstructtction of temporal power spectrum with a temporal dimensionality of 2^{14} and smoothness parameters of $\sigma_0 = 10^4$ and $\sigma_1 = 10^2$. The first peak is clearly visible at $q = 0.131$ [Hz]. Every peak is recorded as a multiple of the rotation frequency in Table 7.1.

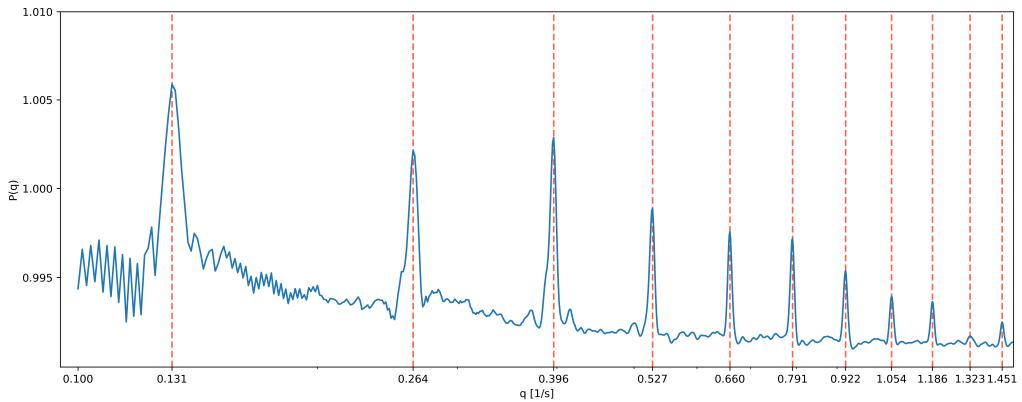


Figure 7.6: Reconstructed temporal power spectrum with a temporal dimensionality of 2^{14} and smoothness parameters of $\sigma_0 = 10^5$ and $\sigma_1 = 10^4$. This parametrization allows for a weaker smoothing of the power spectrum. Every peak is recorded as a multiple of the rotation frequency in Table 7.1.

q [Hz]	q/q_0	q [Hz]	q/q_0
0.264	1.99	0.264	1.99
0.396	2.99	0.396	2.99
0.526	3.97	0.527	3.98
0.660	4.98	0.660	4.98
0.791	5.97	0.791	5.97
0.921	6.95	0.922	6.96
1.054	7.96	1.054	7.96
1.186	8.95	1.186	8.95
1.320	9.96	1.323	9.99
1.453	10.97	1.451	10.95

Table 7.1: Frequencies q [Hz] of the reconstructed temporal power spectrums' peaks with their multiplicity of the rotation frequency q/q_0 . On the left side are the frequency peaks of the reconstruction using strong smoothing as in Figure 7.5 and on the right side the peaks using weak smoothing seen in Figure 7.6. Both reconstructions indicate the same frequencies.

7.2 Energy Spectra of Features within Signal Map Oscillations

In congruence with Pumpe et al. [21], the reconstructed signal map oscillations show distinct peaks when enhanced. As these might be the result of different physical processes on the neutron star, it is worth analyzing their spectral distribution. For this reason, one oscillation of the signal map with a time dimensionality of 2^{14} is split in three different segments as can be seen in Figure 7.7 for $100 - 140$ s and $0 - 40$ keV. The red lines split the oscillations apart with a frequency of 0.1318 Hz as this was a suitable frequency for the signal map. The green lines are used to mark the end of one peak in the oscillation. With every available peak resembling a fraction of the signal map, all oscillations are overlayed and summed to get each of the three peaks' means. These are depicted in Figure 7.8 for the signal map and in Figure 7.9 for the raw data. The three peaks' mean distributions are fairly similar, hinting at a common origin of these three emission regions within oscillations.

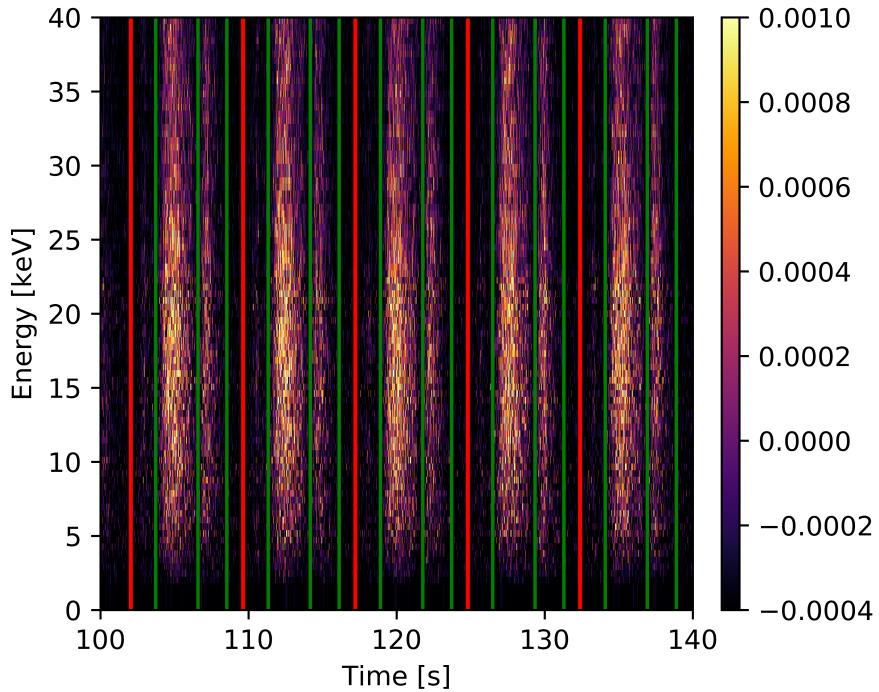


Figure 7.7: Signal Map extract for 100 – 140 s and 0 – 40 keV overlaying a fragmentation with an empirical rotation frequency of 0.1318 Hz. The red lines divide the oscillations and the green lines signal the end of one peak in the oscillations. The last section between the third green line and the red line is neglected as it is not clearly assignable to either the first or last peak and exhibits almost no magnitude.

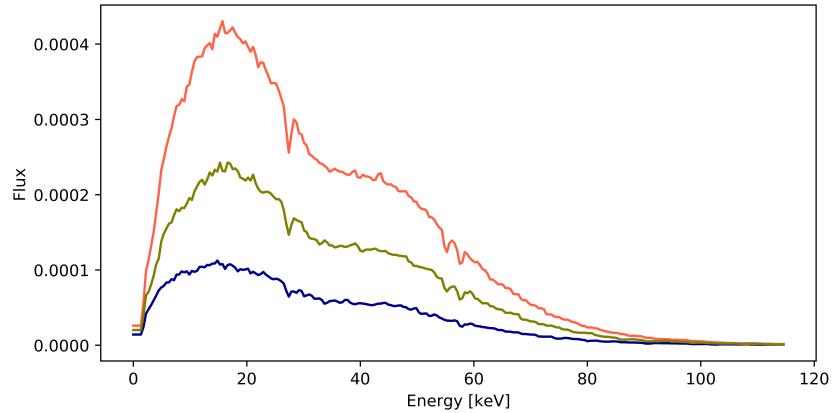


Figure 7.8: Mean spectral distribution of the three peaks for all oscillations of the signal map. The blue graph represents the first peak, the red graph the second peak and the green graph the third peak.

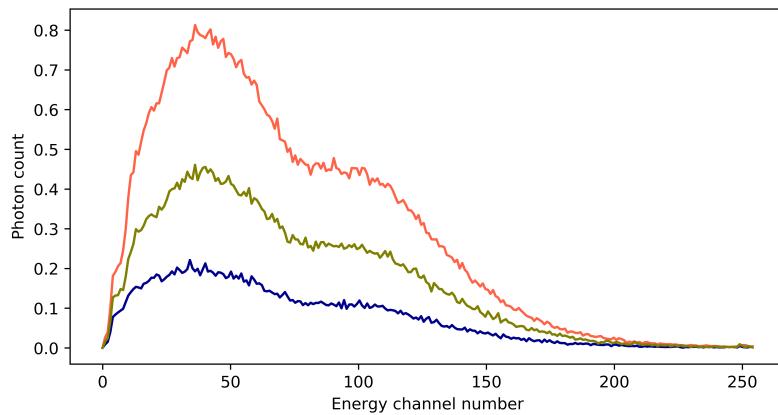


Figure 7.9: Mean energy channel distribution of the three peaks for all oscillations of the raw data. The blue graph represents the first peak, the red graph the second peak and the green graph the third peak.

8 Conclusion

At the beginning of this thesis, an overview over neutron stars and our current understanding of the cause for magnetar giant flares was given. By the usage of information field theory, the data of a magnetar giant flare of the neutron star SGR 1806–20 has been analyzed. For this purpose, the behaviour of the satellite that observed the event was modelled as a linear response operator. Adding the statistical phenomenon of Poisson shot noise, denoising the data and tracing it back to a signal map was achieved by the D4PO signal reconstruction algorithm [1]. During the last section, the resulting signal maps for different initial parameters and their power spectra were presented. The findings included distinct peaks in the temporal power spectrum that relate to overtones of the rotation frequency. Additionally, a smooth energy power spectrum was found. This was expected since synchrotron radiation should display a smooth energy distribution without strong features. Lastly, it was shown that each of the light curve’s oscillations displays three distinct peaks with a spectral distribution similar to each other. As there are no visible differences, a common origin of the emission peaks is suggested.

For SGR 1806 – 20, Israel et al. found additional periodicities to its rotation frequency, but only within certain time intervals of the data and Pumpe et al. used a higher temporal resolution for the reconstructed signal map to find one additional frequency in the temporal power spectrum [9][21]. These alterations as well as a thorough determination of the optimal smoothness parameters could be subjects for future works utilizing this thesis to survey any additional frequency peaks in the temporal power spectrum. As no additional periodicities were found in the temporal power spectrum with the reconstruction algorithm of this thesis, none of the previously recorded frequencies can be confirmed.

Neutron stars are one of the most interesting objects in space and provide us with the exclusive opportunity to study matter at extreme conditions. As these conditions are impossible to be created on Earth, neutron stars will remain an important topic for study in the years to come. For this purpose, the study of magnetar giant flares is appealing as new insights might lead to better constraints of the equation of state. With many telescopes focused on magnetars, data of future giant flares is certain to emerge and its analysis might provide us with a better understanding of neutron stars.

Bibliography

- [1] D. Pumpe, M. Reinecke, and T. A. Enßlin. “Denoising, Deconvolving and Decomposing multi-Dimensional Photon Observations- The D4PO Algorithm”. In: *ArXiv e-prints* (Feb. 2018). arXiv: 1802.02013 [astro-ph.IM].
- [2] David M Palmer et al. “A giant γ -ray flare from the magnetar SGR 1806–20”. In: *Nature* 434.7037 (2005), p. 1107.
- [3] Pankaj Jain. *An introduction to astronomy and astrophysics*. CRC Press, 2016.
- [4] W. Gebhardt. *Script to lecture on neutron stars at University of Regensburg*. 2007.
- [5] Thierry J-L Courvoisier. *High energy astrophysics: an introduction*. Springer Science & Business Media, 2012.
- [6] J. Lattimer. *Script to lecture on neutron stars at Stony Brook University, New York*. 2005.
- [7] JM Lattimer and M Prakash. “Neutron star structure and the equation of state”. In: *The Astrophysical Journal* 550.1 (2001), p. 426.
- [8] Robert C Duncan and Christopher Thompson. “Formation of very strongly magnetized neutron stars-Implications for gamma-ray bursts”. In: *The Astrophysical Journal* 392 (1992), pp. L9–L13.
- [9] GianLuca Israel et al. “The Discovery of Rapid X-Ray Oscillations in the Tail of the SGR 1806–20 Hyperflare”. In: *The Astrophysical Journal Letters* 628.1 (2005), p. L53.
- [10] John R Pierce. *An introduction to information theory: symbols, signals and noise*. Courier Corporation, 2012.
- [11] G Alastair Young and Richard L Smith. *Essentials of statistical inference*. Vol. 16. Cambridge University Press, 2005.
- [12] Niels Oppermann, Georg Robbers, and Torsten A Enßlin. “Reconstructing signals from noisy data with unknown signal and noise covariance”. In: *Physical Review E* 84.4 (2011), p. 041118.
- [13] Torsten A Enßlin, Mona Frommert, and Francisco S Kitaura. “Information field theory for cosmological perturbation reconstruction and nonlinear signal analysis”. In: *Physical Review D* 80.10 (2009), p. 105005.

- [14] Torsten A Enßlin and Mona Frommert. “Reconstruction of signals with unknown spectra in information field theory with parameter uncertainty”. In: *Physical Review D* 83.10 (2011), p. 105014.
- [15] T. Ensslin. *Script for Information Theory and Information Field Theory*. 2018.
- [16] N. Oppermann et al. “Reconstruction of Gaussian and log-normal fields with spectral smoothness”. In: *Phys. Rev. E* 87.3, 032136 (Mar. 2013), p. 032136. DOI: 10.1103/PhysRevE.87.032136. arXiv: 1210.6866 [astro-ph.IM].
- [17] Keith Jahoda et al. “Calibration of the Rossi X-ray timing explorer proportional counter array”. In: *The Astrophysical Journal Supplement Series* 163.2 (2006), p. 401.
- [18] A. Smale. “The XTE Technical Appendix”. In: () .
- [19] Janos Schmidt. “Beobachtung von quasi-periodischen Oszillationen bei jungen Neutronensternen”. PhD thesis. 2015.
- [20] Eric Jones, Travis Oliphant, Pearu Peterson, et al. *SciPy: Open source scientific tools for Python*. [Online; accessed *jtoday*]. 2001–. URL: <http://www.scipy.org/>.
- [21] D. Pumpe et al. “Search for quasi-periodic signals in magnetar giant flares. Bayesian inspection of SGR 1806-20 and SGR 1900+14”. In: *Astronomy and Astrophysics* 610, A61 (Feb. 2018), A61. DOI: 10.1051/0004-6361/201731800. arXiv: 1708.05702 [astro-ph.HE].
- [22] Uwe Schöning. *Algorithmik*. Spektrum, Akad. Verlag, 2001.

I would like to express my gratefulness to my tutor Dr. Daniel Pumpe for his earnest and patient oversight of this work, without which it would have not been possible.

Furthermore, I would like to thank my good friend and colleague Marvin Baumann as this thesis is a result of our collaboration that helped in enduring all hardships along the way.

Hiermit erkläre ich, die vorliegende Arbeit selbständig verfasst zu haben und keine anderen als die in der Arbeit angegebenen Quellen und Hilfsmittel benutzt zu haben.

München, den 14. Juni 2018