A&A, 684, A155 (2024)
https://doi.org/10.1051/0004-6361/202347750
© The Authors 2024

Astronomy & Astrophysics

The first spatio-spectral Bayesian imaging of SN1006 in X-rays

M. Westerkamp[1,2], V. Eberle[1,2], M. Guardiani[1,2], P. Frank[1], L. Scheel-Platz[1,2,3,4,5], P. Arras[1], J. Knollmüller[6,7], J. Stadler[1,6], and T. Enßlin[1,2,6]

[1] Max Planck Institute for Astrophysics, Karl-Schwarzschild-Str. 1, 85748 Garching, Germany
e-mail: margret@mpa-garching.mpg.de
[2] Faculty of Physics, Ludwig-Maximilians-Universitaet Muenchen (LMU), Geschwister-Scholl-Platz 1, 80539 Munich, Germany
[3] Institute of Biological and Medical Imaging, Helmholtz Zentrum Muenchen, Ingolstaedter Landstr. 1, 85764 Neuherberg, Germany
[4] Institute of Computational Biology, Helmholtz Zentrum Muenchen, Ingolstaedter Landstr. 1, 85764 Neuherberg, Germany
[5] Technical University of Munich, School of Medicine, Einsteinstr. 25, 81675 Munich, Germany
[6] Excellence Cluster ORIGINS, Boltzmannstr. 2, 85748 Garching, Germany
[7] Technical University of Munich, TUM School of Natural Sciences, Boltzmannstr. 2, 85748 Garching, Germany

Received 17 August 2023 / Accepted 15 January 2024

ABSTRACT

Supernovae (SNs) are an important source of energy in the interstellar medium. Young remnants of supernovae (SNRs) exhibit peak emission in the X-ray region, making them interesting objects for X-ray observations. In particular, the supernova remnant SN1006 is of great interest due to its historical record, proximity, and brightness. Thus, it has been studied with a number of X-ray telescopes. Improving X-ray imaging of this and other remnants is an important but challenging task, as it often requires multiple observations with different instrument responses to image the entire object. Here, we use *Chandra* observations to demonstrate the capabilities of Bayesian image reconstruction using information field theory (IFT). Our objective is to reconstruct denoised, deconvolved, and spatio-spectral resolved images from X-ray observations and to decompose the emission into different morphologies, namely, diffuse and point-like. Further, we aim to fuse data from different detectors and pointings into a mosaic and quantify the uncertainty of our result. By utilizing prior knowledge on the spatial and spectral correlation structure of the diffuse emission and point sources, this method allows for the effective decomposition of the signal into these two components. In order to accelerate the imaging process, we introduced a multi-step approach, in which the spatial reconstruction obtained for a single energy range is used to derive an informed starting point for the full spatio-spectral reconstruction. We applied this method to 11 *Chandra* observations of SN1006 from 2008 and 2012, providing a detailed, denoised, and decomposed view of the remnant. In particular, the separated view of the diffuse emission ought to provide new insights into the complex, small-scale structures in the center of the remnant and at the shock front profiles. For example, our analysis reveals sharp X-ray flux increases by up to two orders of magnitude at the shock fronts of SN1006.

Key words. methods: statistical – techniques: image processing – ISM: supernova remnants – X-rays: general – X-rays: ISM

1. Introduction

In the year 1006, observers on Earth were able to see the light of a bright "new star", which eventually faded after a few months. This observation is now attributed to a type 1a supernova (SN1a) event that produced a remnant (SNR), now known as SN1006 or SNR G327.6+14.6. It is the brightest stellar event ever recorded and its historical record (Stephenson 2010) is one of the reasons why this remnant was an interesting target for several observational campaigns. SN1006 is notable for being a relatively unobscured supernova remnant (SNR; Katsuda et al. 2013) that is large in angular size due to its proximity to Earth (Winkler et al. 2003). All these points have made SN1006 a good object for studying SN1a events and has led to an impressive research history.

In particular, X-ray observations of the remnant have provided important information about the dynamics and energies of the supernova explosion and the surrounding interstellar medium. When a supernova explodes, it creates a rapidly expanding shell of ejected material that compresses and aggregates up the surrounding interstellar medium (ISM). The collision between the expanding shell and the ISM creates a shock wave that propagates into the ISM and heats it up so that it emits thermal and non-thermal X-rays (Seward & Charles 2010). In young SNRs, both thermal and non-thermal emission have a peak in the 0.5–10 keV energy range (Li et al. 2015), making current X-ray telescopes perfect for studying these objects. An important observation was made by Koyama et al. (1995), who detected synchrotron X-ray emission in the envelope of SN1006, supporting the theoretical expectation that the shock wave of SNRs accelerates particles to extremely high energies. This is believed to be a major production process of cosmic rays (CRs). Accordingly, SNRs are one important source of energy for the ISM via cosmic rays. This observation led to many subsequent spectral (Helder et al. 2012; Reynolds 2008) and spatio-spectral analyses of SN1006 (Bamba et al. 2003; Winkler et al. 2014; Li et al. 2015) to study the spatially varying X-ray production processes in SN1006. In addition, supernovae are known to produce heavier elements from lighter ones during the explosion, which are ejected into the ISM and enable the formation of new stars and planetary systems, making them very important for the Galactic metabolism. Winkler et al. (2014) and Li et al. (2015) studied the spatial distribution of elements in the remnant. Long-term observations of SN1006 allowed Winkler et al. (2014) and

Open Access article, published by EDP Sciences, under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
This article is published in open access under the Subscribe to Open model. Open access funding provided by Max Planck Society.

Katsuda et al. (2013) to study its proper motion and thereby have offered insights into the dynamics of the explosion, the evolution of the remnant, and its interaction with the interstellar medium. Despite the extensive previous studies of SN1006 and other SNRs, there are still a number of aspects that are not well understood. Among them are the details of particle acceleration at shock fronts (Vink 2011).

In recent years, there have been significant advances in X-ray astronomy aimed at studying such high-energy phenomena in the universe. These advances have been driven in large part by the development of new X-ray satellite missions such as *Chandra*, *XMM-Newton* and *Suzaku*, which have provided unprecedented spatial and spectral resolution. However, any technological advance in space-based astronomical instruments must be accompanied by advances in imaging methodology in order to exploit the full potential of these instruments. Here, we focus on the development of such an imaging method, capable of denoising, deconvolving, and decomposing the data, and apply it to *Chandra* observations of SN1006: the highest resolution data of this SNR to date. The aim is to obtain a more detailed view of the small-scale structures of the remnant and, thus, to allow a more detailed study of the open questions in the field of supernovae and their remnants as well as to challenge and benchmark the imaging method.

To obtain an accurate and meaningful reconstruction of the true flux from the given X-ray data, there are a number of challenges that need to be overcome. X-ray telescopes such as *Chandra* record the data from these high-energy phenomena as photon count events accompanied by information about the photon's arrival direction, time, and energy. In the present work, the events are categorized into spatial and energy bins, which yields independent Poisson statistics for each pixel. In particular, X-ray observations often have low count rates, which poses a challenge because of the resulting poor signal-to-noise ratio. Accordingly, a major task in X-ray imaging is the denoising of the corresponding data. In addition, there is an instrument-specific response to the observed X-ray flux, which complicates the relation between the sky and data, particularly the exposure and the point spread function (PSF). A coherent representation and application of the response is a complex problem, as the instrumental properties of X-ray instruments tend to change with off-axis angle, energy, and time. Ultimately, one of the goals of X-ray imaging is to discriminate between noise, background, and extended and point sources. So far, most imaging techniques are designed to extract either the point sources or the diffuse flux, but lack the ability to reconstruct both simultaneously. In the study of such extended sources as SN1006 in particular, this separation of components is essential when investigating the spectra and, thus, the emission properties of the remnant at each location.

The study presented here aims to address these challenges in X-ray imaging. In particular, we use information field theory (IFT; Enßlin 2019) as a versatile mathematical framework for reconstructing the signal from large and noisy data sets by combining information theory, statistical physics, and probability theory. Along with the numerical IFT algorithms implemented in the software package NIFTy (Arras et al. 2019), it provides an excellent tool for denoising, deconvolving and decomposing the image. This capability has already been demonstrated for Poisson data (Selig & Enßlin 2015; Pumpe et al. 2018; Scheel-Platz et al. 2022). The basis of IFT is Bayes theorem applied to the problem of reconstructing fields. In our case, the sky photon flux is regarded to be a field, which we subsequently refer to as the signal field. It is inferred given prior knowledge on its configuration and the measurement data, which is interpreted using a separate model for the measurement response. The instrument description, including its noise statistics determines the so called likelihood; in other words, the probability to observe specific data given a sky flux configuration. By combining the prior and the likelihood into the posterior distribution, we obtain not only an estimate of the actual sky photon flux as its posterior mean, but also an estimate of the uncertainty via the posterior variance.

During inference, the prior model guides the separation of the signal into different components, such as point-like and diffuse structures. Therefore, we need to carefully encode our knowledge of the different components into our prior model to give the inference the chance to discriminate their contributions to the observed photon counts. To this end, we have modeled the signal field as a superposition of different physical fluxes: the emission from point-like and extended sources. Assigning a different correlation structure to the diffuse emission from extended sources, which is assumed to be spatially correlated, and the point sources, which are assumed to be spatially uncorrelated, makes it possible to distinguish between these components. A spatio-spectral prior allows the reconstruction of the emissivity as a function of energy and spatial position. Further knowledge about the different spectra of the components improves their separation.

The instrumental description encoded in the likelihood drives the deconvolution of the data from the PSF, image denoising, and exposure correction. Specifically for *Chandra*, there are two different ACIS X-ray imagers: ACIS-I and ACIS-S. The majority of the chips in ACIS-I and ACIS-S are front-illuminated. However, ACIS-S also contains two chips that are back-illuminated leading to a significant number of non-astronomical photon events in these regions. To account for the latter, we added a further model component of a non-astronomical, spatially varying but temporally constant background that is present in regions of the back-illuminated chips. An additional challenge is the fact that *Chandra*'s field of view (FOV) is small compared with the extent of SN1006. It is therefore not possible to capture the entire remnant in a single ACIS-I or ACIS-S image. Instead, mosaicking is required (Winkler et al. 2014), which can be effectively implemented, even for varying instrument responses, by combining the corresponding likelihoods.

Overall, the spatio-spectral inference of the sky flux is associated with a significantly higher computational complexity than an inference that only considers the spatial direction. Therefore, we introduce a multi-step model, which considers two different priors: a purely spatial one and a spatio-spectral one. First, we performed a spatial reconstruction using the spatial prior. The result of this spatial reconstruction was mapped onto the entire spatio-spectral sky. The mapped sky with multiple energy bins added was then used as the initial guess for the subsequent spatio-spectral reconstruction. This allowed us to perform parts of the reconstruction and especially of the component separation in a smaller parameter space.

This multi-step model, which we call the transition model, and the reconstruction results on SN1006 are presented and discussed in this paper. In Sect. 2, we present current methods used in X-ray imaging and their application results on SN1006 data thus far. We also review state-of-the-art approaches to photon count data in the field of IFT. An introduction to the imaging of photon data with IFT is given in Sect. 3. The explicit structure of the algorithm and in particular of the transition model is given in Sect. 4. Section 5 focuses on the corresponding prior description and Sect. 6 explains the instrument model and the *Chandra* observations of SN1006. In Sect. 7, we present a reconstruction

from synthetic data to validate the method, before finally presenting and discussing the reconstruction results on SN1006 in Sect. 8. The conclusion and outlook for future research is given Sect. 9.

2. Related work

This section is devoted to a review of previous studies and state-of-the-art developments in X-ray imaging. Previous investigations in high-energy astrophysics, with a focus on X-ray studies, are highlighted in three parts. First, we have a discussion of previous and current X-ray imaging techniques, followed by an explanation of the results of these techniques applied to SN1006 and an introduction to previous imaging techniques with IFT, which is the basis for the reconstruction presented here.

2.1. State-of-the-art X-ray imaging

The study of X-ray phenomena in the universe began in the 1960s and it is still a relatively new field of astrophysics due to the inability of ground-based telescopes to observe X-rays from astronomical sources. However, there have been many technical developments since then, discussed in greater detail in Seward & Charles (2010). Here, we focus on the imaging techniques that have been developed and are in use, with a non-exclusive focus on the *Chandra* X-ray Observatory. A more comprehensive summary of recent developments in X-ray analysis for X-ray telescopes *XMM-Newton*, *Suzaku*, and *Chandra* can be found in Seward & Charles (2010).

Among others, Seward & Charles (2010) have offered insights into the steps and techniques in the widely used *Chandra* data processing pipeline[1]. The corresponding methods and further data imaging and response tools have been implemented in the software tool *Chandra* Interactive Analysis of Observations (CIAO; Fruscione et al. 2006), developed by the *Chandra* X-ray Center (CXC).

Overall, there are some standards for extended sources, such as SN1006, which have been applied in recent publications. One of these is the reduction of background from the data, which can obscure the signal from the source of interest. A disadvantage of this approach is that the subtraction of the background comes at the cost of eliminating real X-ray events. Another tool, applied in particular for extended sources, is mosaicking. This allows for the analysis of sources that have a greater extent than *Chandra*'s FOV. In CIAO, mosaicking is implemented by transforming the raw count images, the effective area, and the background maps into a single coordinate system. Reconstructing an image from these mosaics has its difficulties, as there are often several PSFs and response matrix function (RMF)s for one source. So far, this problem has been overcome by calculating and using the weighted average of the PSF and RMF for the data patches, as suggested by Broos et al. (2010).

One of the final steps, which depends on the object of interest, is source detection and extraction. The aim is to separate the X-ray source of interest from the background. For this purpose, three well-known algorithms have been implemented in CIAO: the sliding cell algorithm (Calderwood et al. 2001), wavelet detection algorithm (Freeman et al. 2008) and Voronoi tessellation and percolation algorithm (Ebeling & Wiedenmann 1993). The sliding cell algorithm, previously used for Einstein and ROSAT, searches for sources by summing the counts in a square cell that slides over the image. For comparison, the counts in a cell assigned to the background are taken. From the ratio of the counts in the cell to the counts in the background, the cell might be assigned to a source. Wavelet detection, on the other hand, decomposes the signal into a series of wavelets. By analyzing the coefficients of the wavelets, patterns of different scales can be detected in the data. Finally, data cleaning and source extraction techniques differ for point sources and diffuse emission. This involves additional work as the pipeline needs to be run several times to fully extract point source and diffuse emission information.

There have been other approaches to source decomposition that fall into the category of blind source separation. In general, the goal of blind source separation is to automatically decompose observations into features maximizing their statistical separation. In Warren et al. (2005), a principal component analysis (PCA) approach was presented to determine the location of the contact discontinuity and the shock wave, and thus find evidence for cosmic ray acceleration in the SNR Tycho. In particular, sparse blind source separation aims to compress the signal and extract its essence in this way. One application of sparse blind source separation to *Chandra* data was recently presented in Picquenot et al. (2019). Moreover, a generalized morphological component analysis Bobin et al. (2016) was performed to X-ray data of Cassiopeia A by Bobin et al. (2020) to decompose the spectrum into its components such as thermal and non-thermal emission. Here, the generalized morphological component analysis models the source as a linear combination of a fixed number of morphological components and solves the according blind source separation problem, while putting sparsity constraints on the morphological components.

In addition, Bayesian and machine learning approaches have been applied for source separation, model comparison or point source characterization. Guglielmetti et al. (2009) analyzed Bayesian techniques for the joint estimation of sources and background, while Cruddace et al. (1988) implemented a maximum likelihood algorithm for the calculation of certain parameters of the detected sources, which has also been used for *XXM-Newton*. In Ellien et al. (2023), different components of the spectrum were modeled for *Chandra* data of five thin bands around Tycho and different one-, two-, and three-component models were analyzed via a Bayesian model comparison. Recently, a machine learning approach was published by Kumaran et al. (2023), with the aim of using it as an automated source classifier. The approach is based on supervised learning and allows point sources to be assigned to specific classes.

2.2. Previous studies of SN1006 in the X-ray range

The supernova remnant SN1006 has an exciting scientific record. As mentioned above, the remnant is of great scientific interest in the study of Type 1a supernovae and their remnants for many reasons: its proximity, low obscuration, and large size. In particular, X-ray observations of the remnant provide an opportunity to study its evolution. Accordingly, there have been intensive studies of SN1006 in this energy range, starting with observations by ROSAT (Willingale et al. 1996) and ASCA (Koyama et al. 1995). The ASCA data on SN1006 were analyzed by Koyama et al. (1995) and Dyer et al. (2003), which led to the confirmation of theoretical predictions that cosmic rays are accelerated at the shock fronts of the remnant. It was also discovered that there are several processes in the supernova remnant that are responsible for the X-ray emission. In fact, it was found that the northeast (NE) and southwest (SW) of SN1006 are dominated by non-thermal, synchrotron emission, while the northwest (NW)

[1] https://cxc.harvard.edu/ciao/dictionary/sdp.html

and southeast (SE) edges are less distinct and are attributed to thermal emission. Accordingly, Dyer et al. (2003) analyzed non-thermal and thermal models on the ASCA data.

The new technologies of the X-ray telescopes *XMM-Newton* and *Chandra* have led to an unprecedentedly high resolution of X-ray sources and, thus, to improved data for SN1006. Bamba et al. (2003) published the first spatio-spectral study of *Chandra* ACIS-S data from the NE shell of SN1006, followed by ACIS-I mosaic data from the analysis of Cassam-Chenaï et al. (2008). Here, we want to highlight the publication of Winkler et al. (2014), as their reconstructed image ought to be the main point of comparison for ours. In Winkler et al. (2014), the standard *Chandra* pipeline (described in Sect. 2.1) was used and point sources were extracted using the wavelet detection algorithm. To use multiple data patches, the observations were merged using CIAO. A more recent study of SN1006 was carried out by Li et al. (2015) on *XMM-Newton* data using the SAS software. The data were preprocessed in a similar way to the *Chandra* pipeline and wavelet detection was used. However, point source detection was only possible at high energies because of the risk of misidentifying small-scale structures in the low energy regime as point sources.

2.3. Previous work on high energy count data with IFT

High-energy astronomical data, including X-ray and gamma-ray data, are recorded in photon counts. So far there have been no applications of IFT to *Chandra* X-ray data, but there have been studies on gamma-ray data and methodological research on the reconstruction and component separation of such count data. First, the algorithm D3PO by Selig & Enßlin (2015) based on IFT implemented the denoising, deconvolution, and decomposition of count data. Building on this, D4PO by Pumpe et al. (2018) allows D3PO to work on fields that have spectral and temporal coordinates in addition to spatial coordinates. Finally, Scheel-Platz et al. (2022) built a model of the gamma-ray sky and applied a variant of D4PO in a spatio-spectral setting.

In this work, we adapt a similar model as presented in Scheel-Platz et al. (2022) to describe the X-ray sky. As such, this is the first application of IFT imaging to X-ray data. Further, we introduce a method to fuse several data sets with different detector characteristics, pointing directions, and noise levels into a mosaic. We demonstrate how the imaging can be accelerated and improved by a multi-step model, which is presented in Sect. 4.

3. Image reconstruction with IFT

In X-ray imaging, we are dealing with finite, incomplete, and noisy data. Here, we use IFT (Enßlin 2019), an information theory for fields, to infer the X-ray sky as a continuous field from this finite data, d. In general, a physical field, $s : \Omega \to \mathbb{R}$, assigns a value to each point in the space Ω, which describes a continuous physical quantity such as temperature, pressure, intensity, and so on; in our case, this is the X-ray flux. Given the data, d, we obtain constraints on the field of interest, which we call the signal field. Since the data provide only a finite number of constraints on the signal field, there could have been an infinite number of signals that have produced the data, even if we completely neglect noise. For this reason, prior assumptions about the field are needed to sufficiently constrain the signal field, s. Given the likelihood, $\mathcal{P}(d|s)$, which describes the measurement, and a statistical description of the prior, $\mathcal{P}(s)$, the posterior probability of the signal given the data can be calculated via Bayes theorem:

$$\mathcal{P}(s|d) = \frac{\mathcal{P}(d|s)\mathcal{P}(s)}{\mathcal{P}(d)}. \tag{1}$$

Through IFT, we aim to inspect this posterior probability, as it allows us to draw posterior samples and thereby calculate any important posterior quantity such as the posterior mean,

$$m = \langle s \rangle_{(s|d)} \equiv \int \mathcal{D}s \; s \, \mathcal{P}(s|d), \tag{2}$$

where $\int \mathcal{D}s$ denotes the path integral over all possible field configurations and a measure of uncertainty via the covariance of the posterior probability,

$$D = \langle (m - s)(m - s)^{\dagger} \rangle_{(s|d)}. \tag{3}$$

The expectation value over the posterior probability is denoted by $\langle\rangle_{(s|d)}$ and $\dagger$ gives the adjoint of the corresponding field. Therefore, the statistical treatment of the fields of interest in IFT creates an important advantage, as we may not only present a point estimate of the field, but also quantify its reliability at each position.

A more detailed description of the likelihood and prior model is given in Sect. 4. Here, we describe image reconstruction with IFT given a general measurement equation. Accordingly, we consider a measurement as a function f that maps a field from its continuous space to a discrete data space. This function is determined by the response, $R(s)$, of the instrument and some statistical noise, n, in the measurement,

$$d = f(R(s), n). \tag{4}$$

Given this generic measurement equation we can calculate the likelihood by marginalizing over the measurement noise,

$$\mathcal{P}(d|s) = \int \mathrm{d}n \, \mathcal{P}(d, n|s) = \int \mathrm{d}n \, \mathcal{P}(d|n, s) \, \mathcal{P}(n|s), \tag{5}$$

$$= \int \mathrm{d}n \, \delta(d - f(R(s), n)) \, \mathcal{P}(n|s), \tag{6}$$

$$= \mathcal{P}\left(f^{-1}(R(s),\ d)|s\right) \left|\frac{\partial f(R(s), n)}{\partial n}\right|^{-1}, \tag{7}$$

where f^{-1}, is the inverse of the measurement function with respect to the second argument, n, and $|\partial f/\partial n|$ is the functional determinant. When combining the likelihood with a prior distribution to obtain the posterior, the main difficulty lies in normalizing the posterior, namely, in computing the evidence: $\mathcal{P}(d) = \int \mathcal{D}s \, \mathcal{P}(d|s)\mathcal{P}(d)$. To circumvent the problem of analytically intractable normalization, we approximate the posterior via variational inference (VI), where a possibly complex posterior distribution $\mathcal{P}(s|d)$ is approximated by a simpler one, $\mathcal{Q}(s|d)$. Mathematically, the Kullback-Leibler divergence (KL; Kullback & Leibler 1951) is the measure that needs to be optimized to find the optimal approximation,

$$\mathcal{D}_{\mathrm{KL}}(\mathcal{Q}(s|d)|\mathcal{P}(s|d)) = \int \mathcal{D}s \, \mathcal{Q}(s|d) \ln\left(\frac{\mathcal{Q}(s|d)}{\mathcal{P}(s|d)}\right). \tag{8}$$

Here, we use the VI version of the KL divergence; when minimized for the parameters of $\mathcal{Q}(s|d)$, it ensures that in the approximation the least amount of spurious information is introduced. The expectation propagation (EP) version of the KL

Table 1. Number of hyper-parameters in each model per component.

Model	s	s_d	s_p	s_b
SF	24	11	2	11
MF	53	19	15	19

Table 2. Number of latent parameters in each model per component.

Model	s	s_d	s_p	s_b
SF	3.4×10^6	1.2×10^6	1.0×10^6	1.2×10^6
MF	1.4×10^7	4.4×10^6	5.0×10^6	4.4×10^6

divergence, $\mathcal{D}_{\mathrm{KL}}(\mathcal{P}(s|d)|\mathcal{Q}(s|d))$, would be more conservative, as it would just ensure that a minimum of information is lost, but none have been introduced. However, EP requires integration over the intractable posterior $\mathcal{P}(s|d)$, while VI only requires integration over a conveniently chosen function $\mathcal{Q}(s|d)$ (e.g., a Gaussian) and, therefore, this is feasible. As a consequence, uncertainty estimates obtained from the VI approximation are known to be a bit too optimistic, which should be kept in mind. However, those are nevertheless well informative about the structure of the uncertainties. For further details, we refer to Frank et al. (2021).

4. Algorithm overview of Bayesian inference of the X-ray sky

4.1. Structure of the reconstruction algorithm

The measure of interest in our reconstruction of SN1006 is the sky flux s as a function of space and energy. In other words, the signal field, s, lives on a space consisting of a two-dimensional (2D) position space and a one-dimensional log-energy space, denoted by $z = (\boldsymbol{x}, y) \in \Omega = \mathbb{R}^2 \times \mathbb{R}$, where $y = \log(E/E_0)$ and E_0 is the reference energy. In order to guide the inference in the latent space and to reduce computational complexity, we introduce a multi-step model, which we call the transition model. The transition model divides the actual reconstruction into three parts, with three different inference problems, which are solved by VI. First, we aim to reconstruct the sky at a single energy level. Here, we perform a purely spatial reconstruction of the signal of interest. This part of the reconstruction is called the single frequency (SF) reconstruction. Its results are used to determine an informed starting position for the spatio-spectral reconstruction, subsequently called the multifrequency (MF) reconstruction. The standard reconstruction algorithm for the SF and MF model are further described in Sect. 4.2. We model the mapping from the SF image to the MF image space as an inference problem, whose solution constitutes step two (introduced in Sect. 4.3). In the third step, we solve the MF reconstruction using the starting point provided by step two. A similar model was previously used by Arras et al. (2022) to move from a spatial domain to a spatio-temporal domain.

Using the transition model, we can solve significant parts of the reconstruction problem, including the separation of point source and diffuse emission, in the SF setting, which has less model and computational complexity. Table 1 shows the number of hyper-parameters for each model, SF and MF, and its subcomponents, reflecting the model complexity. Table 2 presents the number of latent parameters in the model as a measure for the computational complexity. Figures B.1 and B.2 show a quantitative comparison regarding computational complexity and reconstruction error for the transition model presented here versus a pure MF reconstruction. A schematic overview of the described reconstruction algorithm can be seen in the diagram in Fig. 1. The MF and SF prior models themselves are discussed in Sect. 5.

4.2. Variational inference and generative models

As mentioned in Sect. 3, we approximate the posterior given information on the prior model (as described in Sect. 5) and the likelihood description (see Sect. 6) via VI. Two approaches to VI of posteriors within the current NIFTy package are metric Gaussian variational inference (MGVI; Knollmüller & Enßlin 2020) and geometric variational inference (geoVI; Frank et al. 2021). They are designed to approximate high-dimensional and complex posterior probability distributions via optimization of the cross entropy term of the KL in Eq. (8). Both approaches perform the KL optimization in a coordinate space of the problem, in which the prior is a standard Gaussian. In particular, the signal field is described by a generative model $s = s(\xi)$ given a set of latent parameters ξ with a standard Gaussian prior $\mathcal{P}(x_i) = \mathcal{G}(\xi_i, \mathbb{1})$. The generative model encodes all prior knowledge on the corresponding field. To this end, the likelihood is formulated as a function of the latent parameters, $\mathcal{P}(d|\xi)$ and the posterior $\mathcal{P}(\xi|d)$ can be inferred via VI. In this work, geoVI is used, which optimizes the KL for the parameters of a non-linear coordinate transformation in which the posterior becomes an approximate standardized Gaussian. Thereby, geoVI allows for the representation of non-Gaussian signal posteriors. The detailed implementation can be found in Frank et al. (2021). In any case, we need to define generative prior models for both, the SF and the MF model, given the corresponding latent parameters $s_m = s_m(\xi_m)$, $m \in \{\mathrm{SF}, \mathrm{MF}\}$. The detailed explanation of these models is part of the prior description in Sect. 5. The according posterior approximations for each model m are denoted by $\mathcal{Q}_m$.

4.3. Transition model

The indirect encoding of fields in generative models complicates the transition from one model (e.g., the SF model) to another (e.g. the MF model) as the corresponding generative function is in generally not invertible; in other words, its inverse is not unique. Thus, our objective is to determine a mapping function, T, that plausibly maps the parameters of the SF model, ξ_{SF}, to their corresponding MF parameters ξ_{MF},

$$T : \xi_{\mathrm{SF}} \rightarrow \xi_{\mathrm{MF}}. \tag{9}$$

As the transition model is intended to be flexible and adaptable to a range of initial and final models, we implement it as an inference problem. Given the posterior signal space mean $m_{\mathrm{SF}} = \langle s_{\mathrm{SF}}(\xi) \rangle_{\mathcal{Q}_{\mathrm{SF}}}$ and signal space variance $\sigma^2_{\mathrm{SF}} = \langle (s_{\mathrm{SF}}(\xi_{\mathrm{SF}}) - m_{\mathrm{SF}})^2 \rangle_{\mathcal{Q}_{\mathrm{SF}}}$ of the SF reconstruction, we infer the corresponding latent space parameters of the MF model, which we take as the starting point $\xi_{I,\mathrm{MF}}$ for the MF reconstruction. The according virtual measurement equation is:

$$d_{\mathrm{T}} = m_{\mathrm{SF}} = \mathcal{R} s_{\mathrm{MF}}(\xi_{I,\mathrm{MF}}) + n, \; n \curvearrowleft \mathcal{G}(n, N), \tag{10}$$

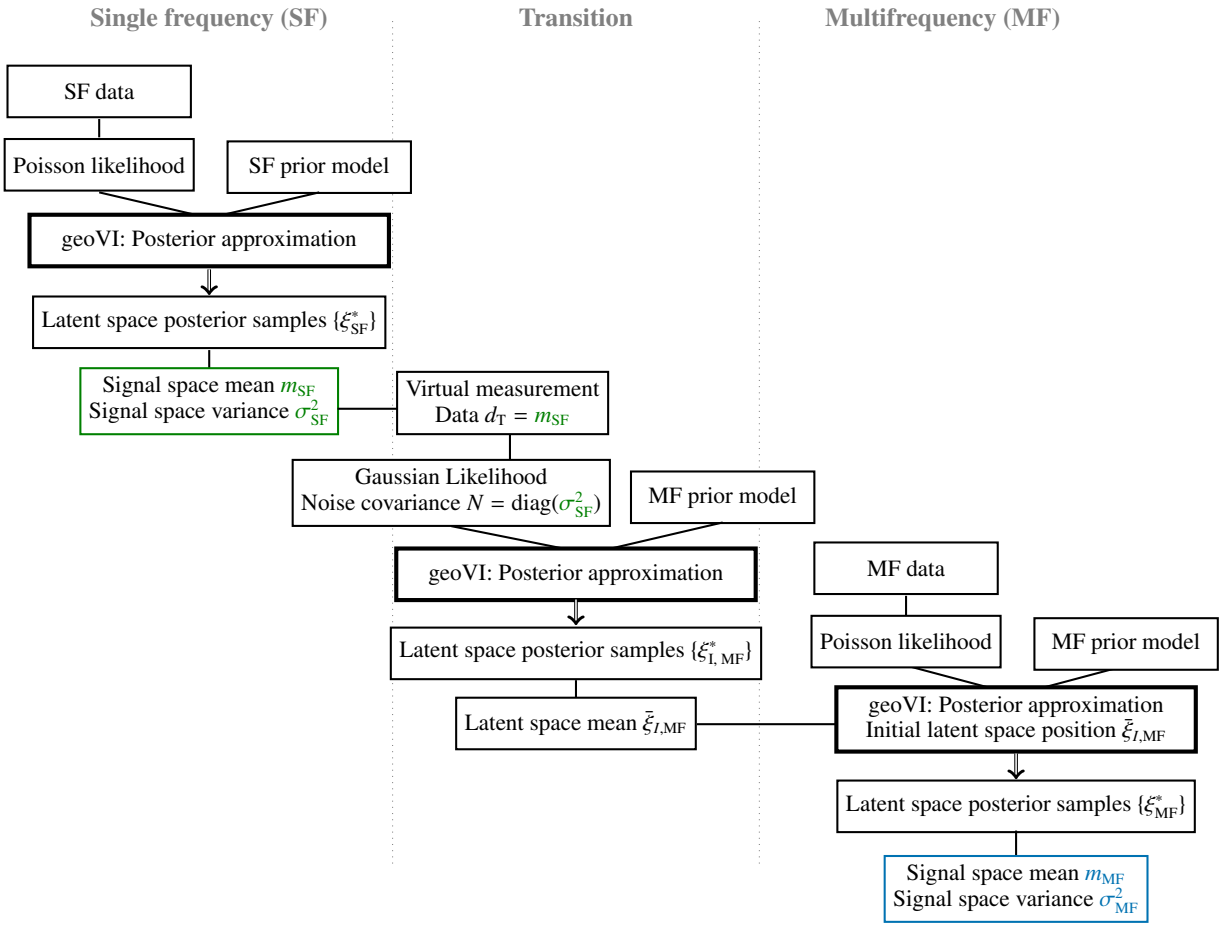


Fig. 1. Structure of the transition model given the generative prior models for the SF reconstruction $s_{\mathrm{SF}} = s_{\mathrm{SF}}(\xi_{\mathrm{SF}})$ and for the MF reconstruction $s_{\mathrm{MF}} = s_{\mathrm{MF}}(\xi_{\mathrm{MF}})$, which transform the according latent parameters ξ_{SF} and ξ_{MF} from the latent space into the signal space. Here, $\xi_{\mathrm{I,\,MF}}$ denotes the initial position of the MF reconstruction in latent space.

where $N = \mathrm{diag}(\sigma^2_{\mathrm{SF}})$. The transition response $\mathcal{R}$ is a linear operator that can be chosen adaptively according to the problem under consideration. In the present analysis, $\mathcal{R}$ is an operator that extracts the highest energy bin from the spatio-spectral field s_{MF}, generating a 2D field, s_{SF}. The likelihood of the mapping inference problem is given by the linear measurement equation in Eq. (10) and the likelihood derivation in Sect. 3 is described by a Gaussian:

$$\mathcal{P}(d_{\mathrm{T}}|\xi_{I,\mathrm{MF}}, \sigma^2_{\mathrm{SF}}) = \mathcal{G}(d_{\mathrm{T}} - \mathcal{R}s_{\mathrm{MF}}(\xi_{I,\mathrm{MF}}), \mathrm{diag}(\sigma^2_{\mathrm{SF}})). \tag{11}$$

The posterior for the initial latent parameters in the MF model $\mathcal{P}(\xi_{I,\mathrm{MF}}|d_{\mathrm{T}}, \sigma^2_{\mathrm{SF}})$ is approximated by $\mathcal{Q}_{\mathrm{T}}(\xi_{I,\mathrm{MF}}|d_{\mathrm{T}}, \sigma^2_{\mathrm{SF}})$, with geoVI as described in Sect. 4.2. We chose the posterior mean of the transition $\langle \xi_{I,\mathrm{MF}} \rangle_{\mathcal{Q}_{\mathrm{T}}}$ as the initial position for the subsequent MF reconstruction. This results in an overall algorithm that starts with a high-energy slice and uses this reconstruction as a starting point for the subsequent spatio-spectral reconstruction. The flow of reconstructions in this approach is illustrated in Fig. 1. The decision to start with the high energy slice was deliberate, as this particular energy range has a more consistent effective area for *Chandra*. Since the transition result is used only as an initial guess, we assume a consideration of a diagonal transition noise covariance, N, to be sufficient. The possibly underestimated noise level is corrected in the subsequent spatio-spectral reconstruction steps.

5. Prior models for the X-ray sky

5.1. Prior composition

As described in Sect. 4.1, we consider two different prior models, one for the SF reconstruction and one for the MF reconstruction. The fields in the SF reconstruction are defined in the spatial domain, $s_{\mathrm{SF}} : \Omega_{\mathrm{SF}} = \mathbb{R}^2 \to \mathbb{R}^+$, whereas the MF fields have an additional spectral dimension, $s_{\mathrm{MF}} : \Omega_{\mathrm{MF}} = \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}^+$. Regardless of the model, we assume that the X-ray sky consists of two possible sources: point sources and diffuse sources. Different prior models represent fluxes of different morphologies, each shaped by their physical production processes. The flux in diffuse structures should vary smoothly over position space. In other words, field values in the vicinity of a location are similar to that, which is best represented by the correlation structure of the field. In contrast, point sources are spatially uncorrelated and therefore best represented by spatially independent and sparsity enforcing priors. We discuss the specific use of either component in more detail below. In the following, the validity of the assumptions made for $s \in \{s_{\mathrm{SF}}, s_{\mathrm{MF}}\}$ is assumed to hold for both the SF and MF sky.

We represent the flux signal, s, as a superposition of point sources, s_{p}, and diffuse sources, s_{d}. In addition, we add a background component, s_{b}, which in our case accounts for the different backgrounds in front-illuminated (FI) and back-illuminated (BI) chips, which are further discussed in Sect. 6.

Correspondingly, we denote the latent space sub-vectors, which parametrize these individual components, as ξ_p, ξ_d, and ξ_b, which altogether form the total latent space vector of the model, $\xi = (\xi_p, \xi_d, \xi_b)$. The according generative prior model is given by,

$$s(\xi) = \underbrace{s_p(\xi_p) + s_d(\xi_d)}_{\text{sky flux}} + \underbrace{R' s_b(\xi_b)}_{\text{BI background}} . \tag{12}$$

Here, R' denotes a mask, which assures that the additional background field is added in BI chip regions only. By expressing the transformation into the standardized coordinate system as a function s_i with $i \in$ d, p, b, we obtain a generative model for each component as described in Enßlin (2022).

A set of prior samples can be seen in the first row of the synthetic data generation diagram in Fig. A.1. Furthermore, the according prior samples and synthetic data allow us to choose the model hyper-parameters correctly. Here, we perform the search in two steps. First, we ensure mathematically via a coarse adjustment of the parameters that the order of magnitude in the counts of the data in Fig. 2 is the same as the order of magnitude in the expected counts of the pixel-wise product of the exposure and the prior samples. Second, we look at the corresponding synthetic data, as shown in Fig. A.1 and fine-tune hyper-parameters such that the components in the actual data and the synthetic data are morphologically similar, in order to improve the convergence of the algorithm.

5.2. Correlated components

Correlated components correspond to a flux that can vary over several orders of magnitude and exhibit spatial correlation. In this sense, the diffuse sky emission and the background are represented by correlated components. Their morphology is implemented by representing the signal for a correlated component as a log-normal processes:

$$s = e^{\tau} \text{ with } \mathcal{P}(\tau|T) = \mathcal{G}(\tau, T), \tag{13}$$

with an unknown covariance, T, describing the correlation structure of the correlated signal component. Since the correlation structure is not known a priori, we infer it concurrently by incorporating the correlated field model from Arras et al. (2022). Using the reparametrization trick introduced by Kingma et al. (2015), we describe the logarithmic sky flux as a generative process:

$$\tau = A\xi_\tau \text{ with } T = AA^\dagger. \tag{14}$$

We went on to model the correlations in space and energy separately and assumed a priori statistical homogeneity and isotropy of the correlated logarithmic sky flux components in each of the subspaces $\Omega^{(k)}$, where $\Omega = \bigotimes_k \Omega^{(k)}$. Thus, according to the Wiener-Khinchin theorem, the corresponding covariances for the space $\Omega^{(k)}$, $T^{(k)}$ are diagonal and defined by the power spectrum $p_{T^{(k)}}$.

To learn the correlation structure of the correlated component, the power spectrum was modeled non-parametrically by representing the logarithmic power spectrum by an integrated Wiener process according to Arras et al. (2022). In particular, the mean and uncertainty of the parameters resulting from the chosen representation, such as the slope of the logarithmic power spectrum, its offset, and the fluctuations around the described power law, are learned from the data by modeling them as generative processes. This introduces further latent parameters to describe the generative model for the correlation structure. In the following, we refer to this prior model as the correlated field.

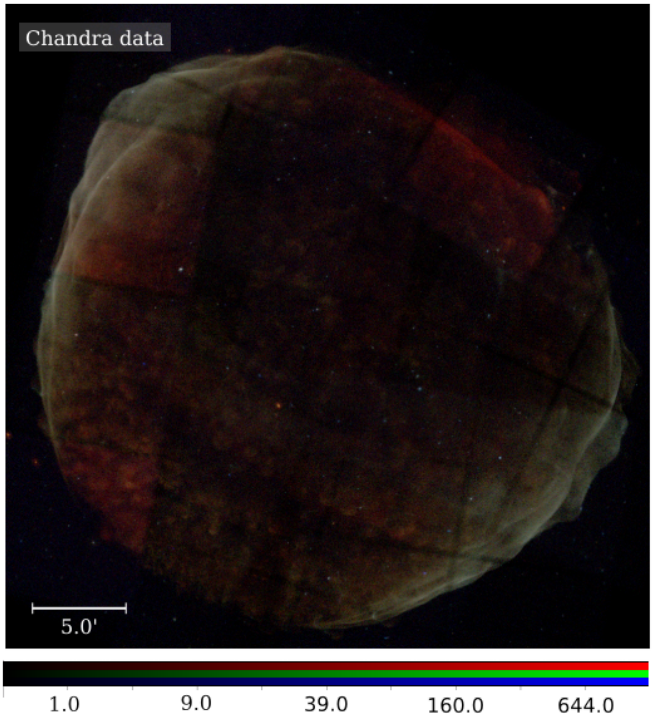


Fig. 2. Visualization of the photon count data used for the reconstruction (Table 3) with right ascension on the x-axis and declination on the y-axis: red = 0.5–1.2 keV, green = 1.2–2.0 keV, and blue = 2.0–7.0 keV.

In the case of the SF model, we only considered the spatial correlations. Accordingly, the generative model for the spatially correlated components in the SF reconstruction is defined via $s_{SF}(\xi_{SF}) = e^{\tau_{SF}(\xi_{SF})}$ with $\Omega = \Omega^{(k)} = \mathbb{R}^2$. In the MF model, we combined the power spectra for the independent spatial and spectral domain via a tensor product and define $s_{MF}(\xi_{MF}) = e^{\tau_{MF}(\xi_{MF})}$. A further description of this generative model and its normalization can be found in Arras et al. (2022).

The diffuse and background components are represented by these spatially and spectrally correlated components. The number of derived hyper-parameters as well as latent parameters per component in each model is given in Tables 1 and 2. For the correlation structure of the diffuse and background components, we made different prior assumptions to ensure an adequate separation of these components. In particular, we assumed that the spatial power spectrum of the diffuse sky structures has a slightly declining slope, allowing for small-scale structures in this component; meanwhile the spectrum of the background is assumed to be steep, allowing only for smooth background noise in the back-illuminated chips.

5.3. Point-like components

Point-like components appear local without any spatial correlation structure, due to their extreme distances. Consequently, we assume that the sky fluxes from point sources are spatially independent and, thus, their prior is factorized in a spatial direction:

$$\mathcal{P}(s_p) = \prod_{\boldsymbol{x}} \mathcal{P}(s_p(\boldsymbol{x}, y)).$$

In Selig & Enßlin (2015), different functional forms of possible point source luminosity priors were analyzed. Since the reconstruction of SN1006 requires a point source prior capable of modeling a few very bright point sources, we chose the inverse-gamma prior for the spatial direction according to Guglielmetti et al. (2009):

$$\mathcal{P}(s_p|q,\alpha) = \prod_x \frac{(q)^\alpha}{\Gamma(\alpha)} \left(\frac{1}{s_p^{(x)}}\right)^{\alpha+1} \exp\left(\frac{-q}{s_p^{(x)}}\right), \tag{15}$$

where α is the shape parameter of the inverse-gamma distribution and q is the corresponding lower flux cutoff. The inverse-gamma prior behaves as a power law for fluxes much larger than the cutoff value, which matches the behavior observed for the luminosity functions of high-energy astrophysical point sources. Intuitively, it encodes the assumption that with increasing luminosity, the set of point sources exceeding it becomes increasingly sparse. We model the inverse-gamma prior in standardized coordinates via inverse transform sampling leading to the generative model:

$$s_{\mathrm{SF}} = s_p(\xi_p), \tag{16}$$

where ξ_p is drawn from a standard Gaussian. Accordingly, s_p encodes the entire complexity of the inverse-gamma distribution and $s_p(\xi_p)$ is drawn according to Eq. (15). In the SF model, Eq. (16) describes the accurate generative model for the point sources.

For the MF model, we need to consider the spectral axis as well, by modeling the point source flux as spatially independent functions of the logarithmic energy y, according to Scheel-Platz et al. (2022). We assume that each point in the spatial subdomain has non-negligible correlations in the energy direction, as described by the correlated field component. In particular, we want to obtain a power-law dependence in the energy direction, defined by the spectral index a, and add fluctuations around it by a correlated field τ_p,

$$(s_{\mathrm{MF}})_p(x,y) = (s_{\mathrm{SF}})_p(x) \frac{\mathrm{e}^{\tau_p^{(x)}(y)+a_p^{(x)}y}}{C}, \tag{17}$$

where C is the normalization. Here, not only the correlated field is described by a generative model, but also the spectral index $a_p^{(x)}$ at every location x is learned. Thus, the additional energy axis introduces a number of new hyper- and latent parameters. The exact numbers of hyper- and latent parameters for the point sources in each model are given in Tables 1 and 2.

6. *Chandra* instrument and data description

In Bayesian X-ray imaging, the prior model (Sect. 5) is responsible for decomposing the components, whereas the denoising and deconvolution is controlled by the likelihood model, which describes the measurement process. In general, an X-ray telescope provides photon counts that are statistically binned into pixels. This stochasticity is modeled by Poisson noise. The Poisson distribution gives the probability of the actual number of photon counts per bin, given the expected number of events, λ,

$$\mathcal{P}(\boldsymbol{d}|\boldsymbol{\lambda}) = \prod_i \mathcal{P}(d_i|\lambda_i) = \prod_i \frac{1}{d_i!}\lambda_i^{d_i}\mathrm{e}^{-\lambda_i}. \tag{18}$$

In the end, we want to know the photon flux at each point in position and energy space. To do this, we need to model the response function R, which in a first step transforms the continuous flux field into a pixel-wise vector of expected photon counts λ_i, given a sky and BI background model. The response function includes all aspects of the instrument specific measurement, which are described in more detail below. Given the response, R (see Sect. 6.1), the number of expected counts at each pixel, λ_i, is calculated via $\boldsymbol{\lambda}(z) = R(s)(z)$. Because the *Chandra* FOV is small compared to the extent of SN1006, multiple observations were taken to cover the whole SNR. For each of several data patches, j, we get the data, $\boldsymbol{d}_j$, and the response, R_j, which need to be fused. Here, we introduce a mechanism that accounts for differences in the exposure and the PSF between the patches. By assuming that each patch is observed independently, we can write the log-likelihood of the full mosaic as the sum over individual patches (Eq. (18)) for each data patch, $\boldsymbol{d}_j$, and the corresponding expected counts, $\boldsymbol{\lambda}_j$, calculated from the response, R_j,

$$\ln \mathcal{P}(\boldsymbol{d}|\boldsymbol{\lambda}) = \sum_j \ln \mathcal{P}(\boldsymbol{d}_j|\boldsymbol{\lambda}_j). \tag{19}$$

6.1. Chandra instrument response

We consider the data taken by the Advanced CCD Imaging Spectrometer (ACIS; Garmire et al. 2003), which is able to determine the energy of each incoming photon by using charge-coupled devices (CCDs). In particular, we consider the energy range 0.5 keV to 7.0 keV, which we bin in accordance with Winkler et al. (2014) into three energy bins (0.5–1.2 keV, 1.2–2.0 keV, and 2.0–7.0 keV). *Chandra* carries two different kinds of ACIS detectors, ACIS-I, used for imaging, and ACIS-S, used for imaging and spectral analysis. According to their application, ACIS-I and ACIS-S differ in the chips they are built of. In particular, ACIS-I is constructed out of FI chips only, which means that the incidental X-ray photons have to pass through the metal wiring until they reach the light-receiving surface. In contrast, ACIS-S also contains BI chips, where the CCD is flipped, such that the gate structure and channel stops do not face the X-ray-illuminated side. Accordingly, the BI chips are more sensitive to soft X-rays, so they are well suited for spectral analyses. However, they have a lower high-energy quantum efficiency and a worse resolution due to increased noise (Keith Arnaud & Siemiginowska 2011). The exact layout of ACIS-I and ACIS-S can be found in Chandra X-ray Center (2021).

We use version 4.14 of CIAO tool (Fruscione et al. 2006) designed by the CXC to extract information on the response ingredients such as the PSF and the exposure as well as on the event files itself for each patch. Here, we have made use of tools from the "data manipulation" category for extracting and binning the data and from the "response tool" category to generate the ingredients of the instrument response.

The exposure map is a key component in the process of converting raw X-ray data into scientifically useful data products, such as images and spectra. The exposure map combines information from the instrument map, which characterizes the instrument sensitivity such as the effective area and aspect solution (McDowell 2006), which describes the spacecraft pointing and roll to create a map of the total observing time (or exposure) for each pixel in the field of view.

In Evans et al. (2010), the effective area as a function of energy is shown for the different chips. As mentioned above already, the FI chips are much less sensitive to low energy X-ray photons than the BI chips. On the other hand, the BI chips have more background flux. The exposure maps for the FI and BI chips can be seen in Fig. 3. In order to account for the higher

Table 3. Information on the *Chandra* ACIS observations for the used data of SN1006 according to Winkler et al. (2014).

ObsID	Instrument	RA	Dec	Date
9107	ACIS-S	15:03:51.5	–41:51:19	24.06.2008
13737	ACIS-S	15:02:15.9	–41:46:10	20.04.2012
13738	ACIS-I	15:01:41.8	–41:58:15	23.04.2012
14424	ACIS-I	15:01:41.8	–41:58:15	27.04.2012
13739	ACIS-I	15:02:12.6	–42:07:01	04.05.2012
13740	ACIS-I	15:02:40.7	–41:50:21	10.06.2012
13741	ACIS-I	15:03:48.0	–42:02:53	25.04.2012
13742	ACIS-I	15:03:01.8	–42:08:27	15.06.2012
13743	ACIS-I	15:03:01.8	–41:43:05	28.04.2012
14423	ACIS-I	15:02:50.9	–41:55:25	25.04.2012
14435	ACIS-I	15:03:42.5	–41:54:49	08.06.2012

Notes. Observations taken by the instrument ACIS-S are followed by the ACIS-I observations.

noise in the BI chips, we introduce an additional BI background field in Sect. 5.

In general, for any of the considered chip types, it is evident that the chips are more sensitive to higher energies, leading us to the decision to take the highest energy bin (2.0–7.0 keV) as a starting point for the transition model. The PSF was simulated using the Model of AXAF Response to X-rays (MARX; Davis et al. 2012). MARX is a software developed by the CXC (amongst others) to simulate the response (i.e., the PSF) of the *Chandra* X-ray Observatory, taking into account the telescope optics, pointing, and aspect of the telescope. We generated the response for each dataset and for each dataset, we used a homogeneous, spatially invariant PSF – but different PSFs for the different patches. The consequences of this approximation are addressed further in the quantitative discussion of the results in Sect. 8.1. For a further analysis of spatially variant PSFs, we refer to Eberle et al. (2023).

6.2. Chandra data of SN1006

The photon count data taken by the instrument is translated into an event table. Each event has information on time, energy, and position. Here, the data are binned into 1024×1024 spatial pixels and three energy bins. Moreover, the pointing direction of *Chandra* varies in time. Thus, the aspect correction is necessary, that is, taking into account the pointing direction of the telescope as a function of time. The data itself is an event list, which specifies the position in the chip coordinates and the arrival time of each photon. In McDowell (2001) the calculation of the sky coordinates of the photon given this event list is specified.

The latest *Chandra* observations of SN1006 according to Winkler et al. (2014) was chosen for the reconstruction presented here. Information on the data is given in Table 3. The aim of the study of Winkler et al. (2014) was to measure the motion of the remnant and to get a more detailed view on its fine scale structures. Thus, the data observations were designed by Winkler et al. (2014) in order to match former observations, to be able to measure the expansion and using a longer exposure time in order to get a more detailed picture. The previous observations, which were taken as first-epoch images by Winkler et al. (2014), comprised a study of the non-thermal NE rim and the thermal NW rim with the ACIS-S array (Long et al. 2003; Katsuda et al. 2009) and the whole remnant with a mosaic of the ACIS-I observation

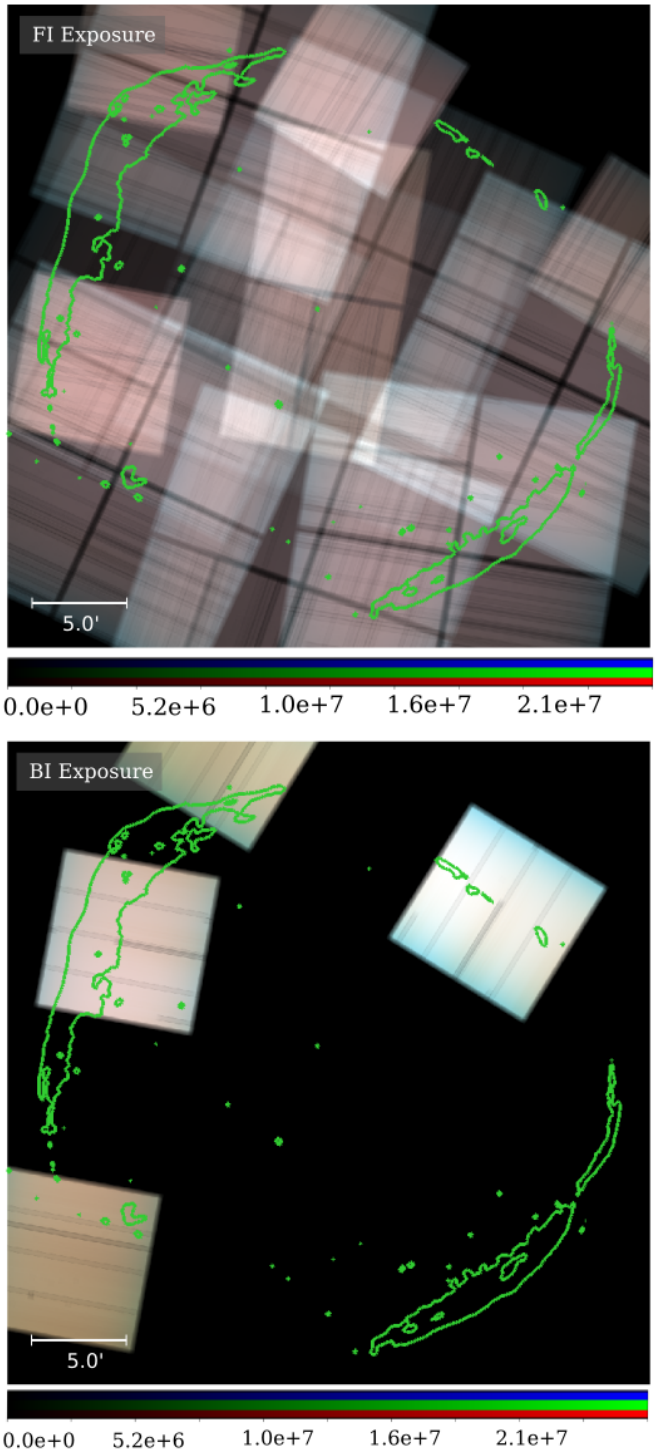


Fig. 3. Visualization of the exposures used for the reconstruction in [s cm^2] (Table 3). Top: full exposure for all FI chips. Bottom: full exposure for all BI chips.

(Cassam-Chenaï et al. 2008). Accordingly, the reconstruction deals with data from the BI and FI chips.

7. Validation of the algorithm using synthetic data

To demonstrate the performance of the developed algorithm, we perform the inference described above on a realistic but simulated dataset. Such a reconstruction based on synthetic data is useful not only for validating the reconstruction method, but also

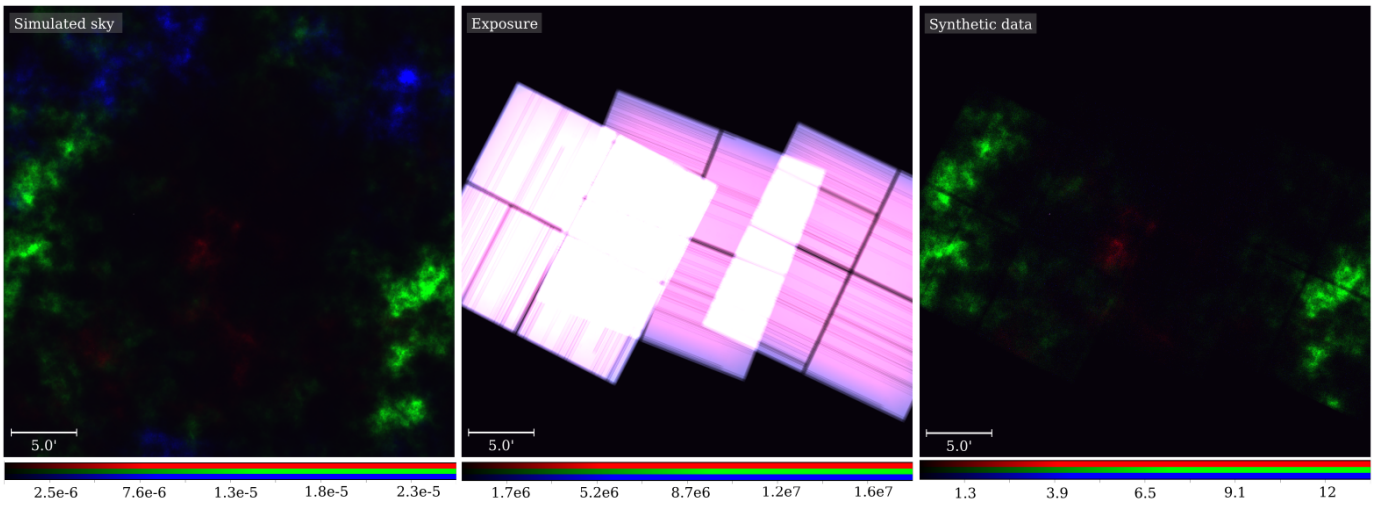


Fig. 4. Generation of synthetic data. Left: sky sample generated for the validation experiment. Center: *Chandra* exposure, modeled by combining three patches (14423, 14424, 14435; see Table 3). Right: synthetic data corresponding to the sky sample, obtained by convolving the sky sample with the PSF and drawing a pixel-wise Poisson sample from the resulting detector flux prediction.

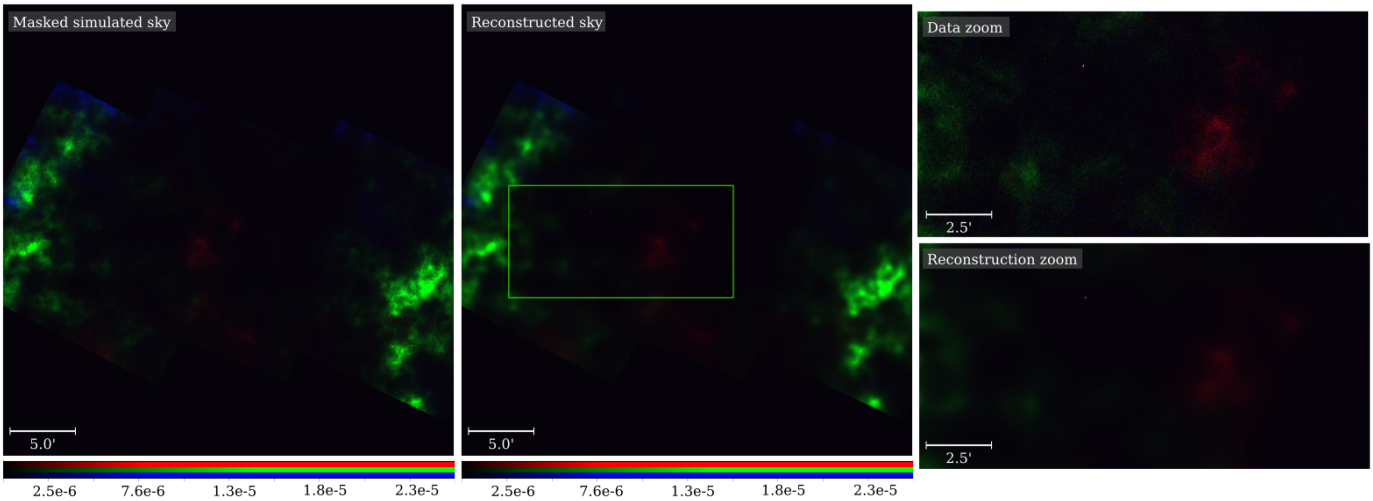


Fig. 5. Reconstruction results on synthetic data. Left: sky sample generated for this study masked by the extent of the exposure patches (14423, 14424, 14435; see Table 3). Center: reconstruction result, i.e., the posterior mean, of the imaged sky masked by the extent of the same exposure patches. Right: zoomed-in regions of the data on top and of the reconstructed image below. The shown cutout region is marked in the center image.

for determining certain parameters of the actual reconstruction, such as the sky flux detection limit. By constructing our model as a generative model, we were able to draw realistic sky samples that are similar to the region of the X-ray sky considered in this study. Given a sample of the sky, we were able to apply the response to it and mimic the Poisson noise. As a result, we obtained the synthetic data. The process of generating synthetic data is illustrated in the Appendix A. For simplicity, we considered only three of the ACIS-I exposure patches for the synthetic reconstruction, rather than the whole set. The relevant synthetic data are shown in Fig. 4, together with the actual simulated sky sample and the considered exposure map.

The resulting spatio-spectral reconstruction of the synthetic data in Fig. 5 shows that the structures of the simulated sky are well captured. The denoising and response corrections are clearly visible when compared to the data shown in Fig. 4. In particular, the right-hand side Fig. 5 shows an enlarged version of the data and reconstruction to illustrate the denoising and deconvolution.

As mentioned before, a particular strength of the X-ray imaging method presented here is that we not only get an expectation of the signal, but also a corresponding standard deviation to this estimate. The corresponding pictures of the standard deviation for the individual energy bins can be seen in Fig. C.2. As expected, higher mean values exhibit greater variability in the flux, which in turn leads to a higher absolute uncertainty. However, given the standard deviation, σ_s, the reconstruction mean, m_s, and the fact that we know the signal ground truth, s_{gt}, itself from our generative model, we can calculate even more interesting validation measures, such as the uncertainty weighted residuals (UWRs) per energy bin, i:

$$(\epsilon_{\mathrm{UWR}})_i = \frac{(m_s)_i - (s_{\mathrm{gt}})_i}{\sigma_i}. \tag{20}$$

In Fig. C.1, we show the UWRs as well as the residuals,

$$r = (m_s - s_{\mathrm{gt}}). \tag{21}$$

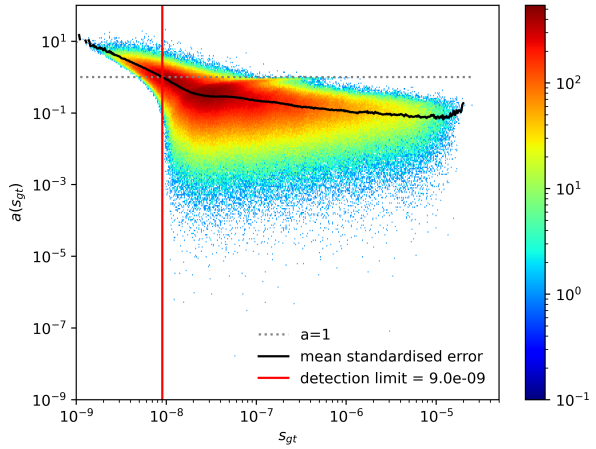


Fig. 6. Visualization of the 2D histogram for the sample averaged relative distance of the posterior sky flux samples vs. the ground truth sky flux (Eq. (22)). The detection limit is determined via the intersection of the line showing the mean standardized error $\bar{a}$ (Eq. (23)) with the a=1 line.

Areas with many counts show higher and more correlated residuals than those with low counts. Accordingly, the UWRs show that these pixels with high counts have higher uncertainty weighted residuals, due to a relatively small uncertainty. Overall, the simulated reconstruction demonstrates that the method developed is internally consistent. Therefore, we used this synthetic reconstruction to set the threshold above which we can no longer distinguish noise from signal, which we refer to as the detection limit. The detection limit is used as a plotting lower limit in the actual reconstruction. Below this lower limit flux values are not shown in the image.

The posterior approximation gives us the opportunity to draw posterior samples $s^* \hookleftarrow \mathcal{Q}(s|d)$, which we can use to calculate the sample averages. In order to determine the detection limit, we define the standardized error, a, between the ground truth, s_{gt}, and each of these samples, s^*, as a function of the ground truth flux,

$$a(s_{\mathrm{gt}}) = \left| \frac{s^* - s_{\mathrm{gt}}}{s_{\mathrm{gt}}} \right|. \tag{22}$$

In Fig. 6, the sample-averaged 2D histogram $a(s_{\mathrm{gt}})$ as a function of the ground truth flux, s_{gt}, is shown. For each value of $s_{\mathrm{gt,\,i}}$. we calculate the mean standardized error of the histogram bins along the a-axis, $\bar{a}(s_{\mathrm{gt,\,i}})$, where $n(a_j, s_{\mathrm{gt,\,i}})$ is the number of counts for each bin (i, j) in the 2D histogram:

$$\bar{a}(s_{\mathrm{gt,\,i}}) = \frac{\sum_j a_j * n(a_j, s_{\mathrm{gt,\,i}})}{\sum_j n(a_j, s_{\mathrm{gt,\,i}})}. \tag{23}$$

Figure 6 reflects the expectation that the standardized error increases with smaller flux. To establish a detection threshold for low fluxes, we defined a limit beyond which we cannot confidently ascertain the presence of a signal in our observations. In this study, we determined the detection threshold in such a way that the mean standardized error is less than 1. In other words, the detection threshold is set via the intersection point of the mean standardized error $\bar{a}$ and the line $a = 1$, leading to a detection limit of $9.0\mathrm{e}^{-9}[\mathrm{s}^{-1}\ \mathrm{cm}^{-2}]$. We only performed this synthetic analysis on three of the data patches. Therefore, this threshold is conservative for the reconstruction with all patches.

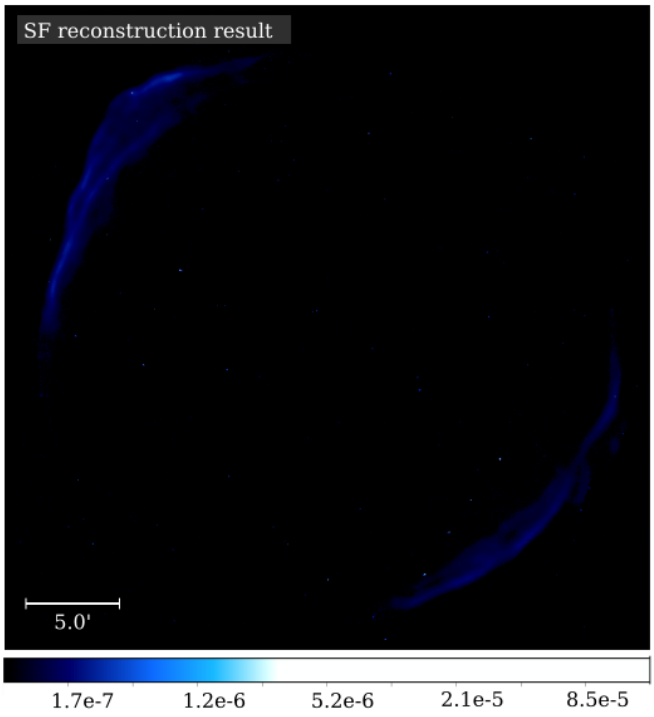


Fig. 7. Spatial reconstruction result for the highest energy bin in $[\mathrm{s}^{-1}\ \mathrm{cm}^{-2}]$.

8. Results and analysis of inference performance

In this section, we discuss the results of the sky flux reconstruction. The additional background from the BI chips according to Eq. (12) is removed in the reconstruction process. In Fig. 7, we show the intermediate result of the SF reconstruction of the highest energy bin. This is taken as the initial condition of the subsequent MF reconstruction of SN1006, whose reconstruction results are shown in Fig. 8 given the data shown in Table 3 using Bayesian imaging and the transition model introduced in Sect. 4.1. The reconstruction was visualized using the SAOImage DS9 imaging application (Joye & Mandel 2003). The according results for each energy bin including the according color bars can be found in Fig. D.2.

As noted above, we are not only reconstructing the sky flux itself, but also its correlation structure in its correlated components. Accordingly, the posterior mean of the power spectrum in the spatial direction of the diffuse, extended sky component was also reconstructed and is shown for each energy bin in Fig. D.1.

8.1. Quantitative discussion

Figure 8 displays the final results of applying our reconstruction algorithm to SN1006: the reconstructed sky and its separated components. The overall separation of diffuse emission and point sources succeeded. The point sources are clearly identified and the PSF deconvolution is particularly evident for the point sources. The diffuse structures are almost free of point source contributions. The effect of component separation can be seen more clearly in Fig. 9, which shows a zoom-in on the NE quarter of the remnant and its components. It can also be seen that some of the additional background noise from the BI chip is partially absorbed into the background model.

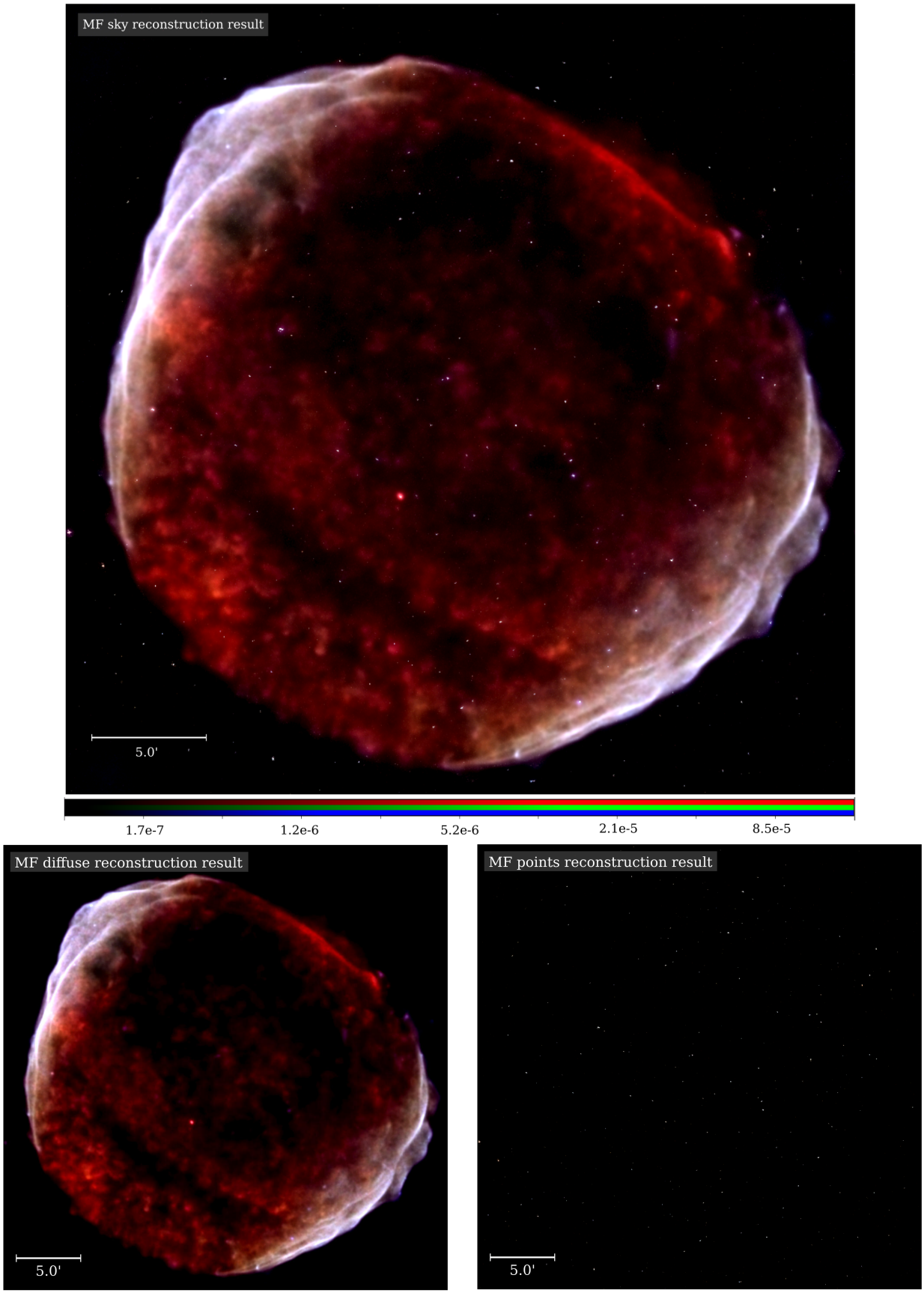


Fig. 8. Reconstruction results for the flux in [s^{-1} cm^{-2}] (red = 0.5–1.2 keV, green = 1.2–2.0 keV, and blue = 2.0–7.0 keV. The corresponding color bars for each of these energy bins can be found in Fig. D.2): top: full sky reconstruction mean. Bottom left: reconstruction mean for diffuse emission. Bottom right: reconstruction mean for point sources.

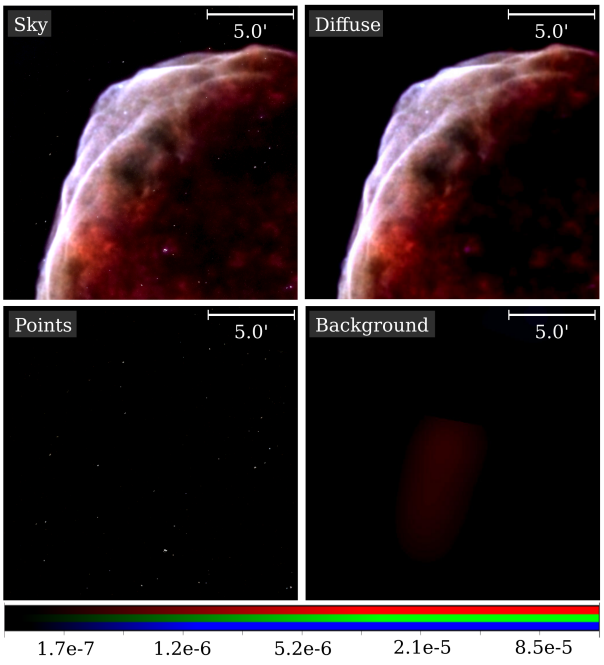


Fig. 9. NE quarter of SN1006 and its components. From left to right: total sky, diffuse emission, point sources, and BI background.

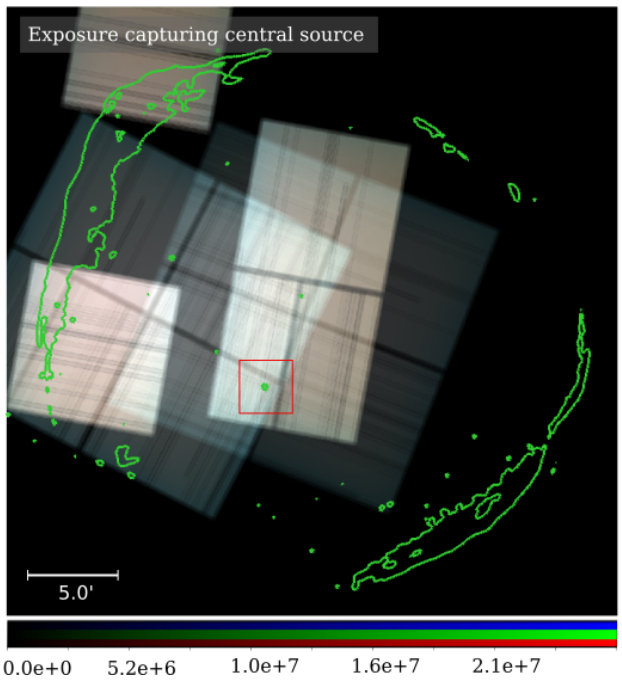


Fig. 10. Exposures that capture the not well separated point source (marked red).

One soft X-ray point source in the center of the remnant was not well separated from the diffuse emission. We believe this was caused by PSF mismodeling in the outer pixels of the detectors, where the source is located in all observations considered, due to the assumption of invariant PSFs within each observation. Figure 10 shows the exposure maps of the observations that covered this source and the position of the point source within these exposure maps. The pointing of ACIS-I and ACIS-S described in Chandra X-ray Center (2021) suggests large

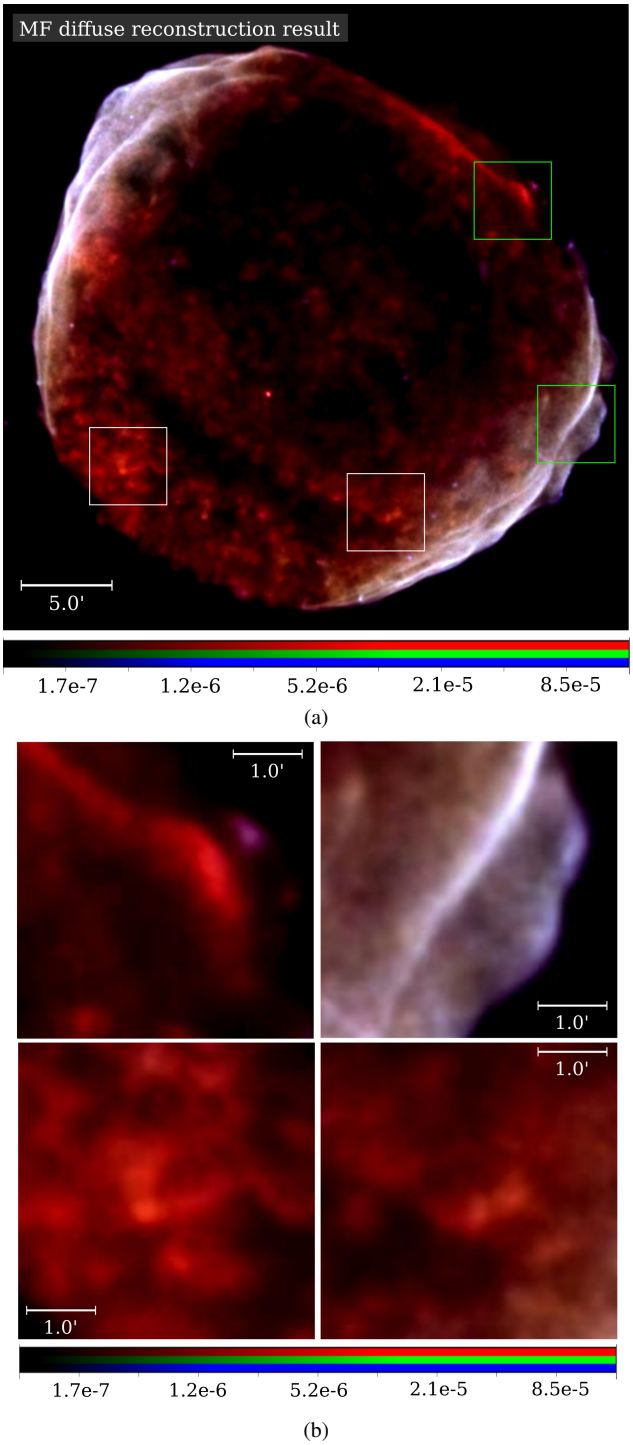


Fig. 11. Zoom-in on the reconstruction results in the diffuse component. (a) The whole diffuse reconstruction and the locations of the zoomed-in areas; (b) the top panels represent the green areas marked in the remnant and show zoom-ins on the denoised shell of the remnant. The lower panels are represented by the white boxes in (a) and show structures in the inner soft X-ray emission of the remnant.

deviations of the actual PSF from the PSF model used in the positions of the mismodeled point source in detector coordinates. Dealing with position dependent PSFs will be addressed in a future publication.

Figure 11 shows the reconstruction of diffuse emissions from the remnant in detail. In order to study the remnant effectively,

it is crucial to get a detailed view of it. To improve the clarity of our results, we present four close-up images of the remnant, highlighting its small-scale structures. The analysis shows that the shell is denoised both in the NW region (where we expect thermal emission) and in the SW region (where we expect non-thermal emission). We can also see that the denoising has improved the resolution of the small-scale structures in the inner X-ray emission of the remnant with respect to the data and also in comparison to the previous study of Winkler et al. (2014).

Due to the statistical approach presented in this study, we obtained an estimate of the sky flux via the mean of the posterior probability, but also an uncertainty estimate via its standard deviation. The corresponding standard deviations for each energy bin are shown in Fig. D.2. The top row of the figure shows the different energy bins of the posterior mean for better comparison. It can be seen that, as expected, the standard deviation is higher for regions of higher flux.

Thanks to the probabilistic approach we adopted, we gain the ability to draw posterior samples from the inferred distribution. Such posterior samples, s^*, have allowed us to compute the posterior mean, the standard deviation or any other statistical quantity of interest. Correspondingly, we can calculate the absolute noise-weighted residual (NWR) $(\epsilon_{\mathrm{NWR}})_j$ for each data patch, j, as another interesting outcome of our results:

$$(\epsilon_{\mathrm{NWR}})_j = \left\langle \frac{|d_j - \lambda_j(s)|}{\sqrt{\lambda_j(s)}} \right\rangle_{Q(s|d)} \approx \frac{1}{N} \sum_{i=1}^{N} \frac{|d_j - \lambda_j(s_i^*)|}{\sqrt{\lambda_j(s_i^*)}}. \tag{24}$$

Here, λ_j describes the reconstructed expected number of counts for each pixel in the data patch, j. The absolute NWRs provide a way to quantify the difference between a measured data point and its reconstruction, and help to distinguish between true deviations of the data from the reconstruction and deviations that are simply due to Poisson noise. The plots of the absolute NWRs are shown in Table D.1 for each energy bin and data patch. These plots can be used as a sanity check on the correctness of the reconstruction presented here, as they allow us to point out systematically unmodeled effects in the likelihood and the prior. We can see that the NWRs are close to one in most regions, which implies a well-fitting model and reconstruction. In particular, in regions around point sources or at strong edges, we find higher NWRs, which we attribute to the well-functioning deconvolution in these regions, leading to deviations of the reconstructed signal from the data.

8.2. Analysis and comparison with previous studies

As mentioned above, SN1006 has been studied extensively using a variety of instruments. In particular, several studies using X-ray telescopes have produced images of the SNR. These studies have provided important insights into its structure and evolution, thus advancing our understanding of supernova explosions and their aftermath. Corresponding imaging approaches to SN1006 can be found in Winkler et al. (2003), Li et al. (2015), Bamba et al. (2003) and also in Fruscione et al. (2006), which demonstrated the fidelity of the *Chandra* data processing pipeline for SN1006. In particular, in this study we have focused on the data and energy regions used by Winkler et al. (2014) and compare our reconstructions with their results. In Winkler et al. (2014), a comparison was made between the X-ray image and a Hα image of SN1006 from CTIO. The comparison shows that there are several thin arcs of Balmer emission in the southern part of the remnant, which lie just in front of small-scale structures in the X-ray emission. In Fig. 11, we show these central parts of the remnant, which are dominated by soft X-rays. We show the enlarged cut-outs of these regions for the extracted remnant. Compared to previous studies, small-scale structures have an improved resolution due to the denoising and deconvolution and are well disentangled from any point sources in the background. Accordingly, the presented reconstruction provides a more detailed view of the inner part of the remnant, enabling a more accurate study of its small-scale structures.

As noted in Li et al. (2015), the different energy bands show spatial variations in the remnant SN1006 with respect to each other. Figure 8 shows these differing spatial variations of intensity with different X-ray energy bands. In particular, in Figs. 9 and 11, parts of the hard X-ray lobe (the non-thermal regions in the SW and NE of the remnant) are well resolved and denoised, without any point source contribution. Soft X-rays in the NW shell are shown in Fig. 11, which shows the shell of the thermal emission. Koyama et al. (1995) presented the first observational evidence that supernova shocks produce cosmic rays. However, the details of the acceleration mechanism of the particles is still an open question. Therefore, the study of the shock fronts is important to gain further insight into the acceleration mechanism and the dynamics of the shock front. The separation of the diffuse emission from the remnant allows us to visualize long intensity profiles along the remnant. Figure 12 shows such radial intensity profiles of the SNR in eight equidistantly space orientations. These denoised and deconvolved profile lines can be very useful to search for halos in front of the non-thermal regions and to get insights into the magnetic field strength according to Helder et al. (2012). Here, the profile lines show strong and sharp X-ray flux increases by up to two orders of magnitude at the shock front in the non-thermal regions in the SW and NE of SN1006. Notwithstanding, we defer the analysis of the structure of the remnant based on our reconstructions to future works.

9. Conclusion

In summary, we present a technique for obtaining an estimate of the true sky photon flux from *Chandra* X-ray event data. By using the sky flux as a generative process, this method allows us to infer not only the flux itself, but also its correlation structure in its extended components. Based on IFT and the application of Bayes' theorem, this method approximates the posterior probability of a signal given the data via geoVI in a problem-adapted latent space. This allows us to draw posterior samples in order to compute the expected sky flux, posterior uncertainty, and further validation and diagnostic measures.

Modeling the different correlation structures of point sources, diffuse emission, and background in the BI chips, we were able to separate point-like, extended sources and the additional noise in the BI chips from the sought-after signal. Compared to previous separation and source extraction techniques, which are usually specified to extract either point sources or extended sources, the inference based on IFT accounts for both components jointly. In particular, we built a spatio-spectral model for the sky flux based on the D4PO algorithm implemented by Scheel-Platz et al. (2022) and used it for the spatio-spectral reconstruction of the X-ray sky. Since the spatio-spectral reconstruction is computationally expensive for a large number of pixels, we introduced an accelerated inference model, called the transition model. In the transition model, we first performed a spatial reconstruction in a single energy band, which has almost one order of magnitude less degrees of freedom as the spatio-spectral reconstruction, making it computationally less complex. The result of the spatial reconstruction, which already contains a lot of information about the sky flux in an energy bin and about

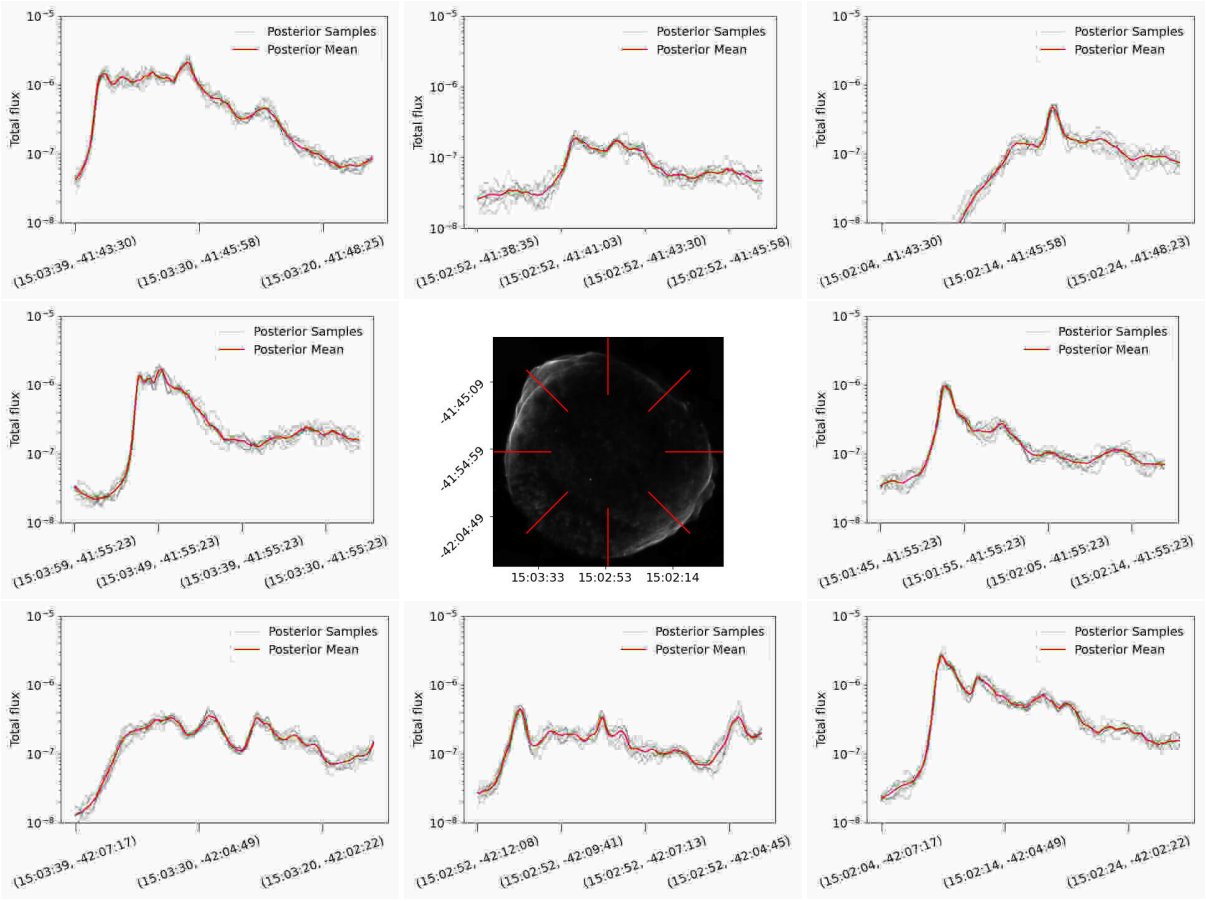


Fig. 12. Flux intensity profiles in [s^{-1} cm^{-2}]. The center image shows the location of the lines along which we present the intensity profile in pixel coordinates. The corresponding intensity profiles are plotted next to the line. The posterior mean of the reconstructed flux is plotted in red and the corresponding posterior samples are plotted in grey. The profiles are shown from left to right from the outsides to the insides of the SNR.

the component separation, was used as an initial condition for the spatio-spectral reconstruction.

A benefit of the here presented analysis is the ability to build mosaics of different observations via the sum of logarithmic likelihoods. Each likelihood has its own description of the instrument response. This approach solves the problem of modelling different PSFs for the same source in different data patches.

We applied the spatio-spectral reconstruction to the latest *Chandra* observations of SN1006, presented by Winkler et al. (2014). The resulting image is a denoised, deconvolved and decomposed image, which provides a detailed view of the small-scale structures of SN1006. We reconstructed a separate image of the point sources present in the considered datasets, which can be compared with point source catalogs and, more importantly, which allows us to study the X-ray emission of SN1006 without point source contributions. The different energy ranges in NE and SW – dominated by synchrotron emission – and the rest of the remnant – dominated by thermal emission – are clearly visible. The intensity profiles at the shell of the remnant are denoised and not visibly affected by point source contributions. We also show other diagnostics such as a simulated data reconstruction, uncertainties, and noise-weighted residuals as a check for systematic errors.

Taking this work as a starting point for spatio-spectral Bayesian imaging of X-ray data, we have pointed out the need for further methodological improvements. One is based on the use of a spatially varying PSF. Alexander et al. (2003) already showed the actual spatial variability of the PSF in the *Chandra* image of the Deep Field North. We expect that the separation of the central point source in SN1006 to improve by the implementation of a spatially varying PSF. However, this is not a trivial task, as an invariant PSF can be applied via multiplication in Fourier space, whereas a spatially varying PSF cannot. Methods are currently being developed to solve this problem in the language of IFT, including a neural network approach recently presented by Eberle et al. (2023). In addition, a line model capable of modelling lines in the thermal emission will help to further resolve the energy direction. An interesting option here, already mentioned in Seward & Charles (2010), would be to define different models for synchrotron emission and bremsstrahlung, with the goal of eventually decomposing the diffuse emission of the remnant into its thermal and non-thermal components. In general, we aim to further improve the reconstruction speed and reduce its computational cost to enable studies of more data sets, larger regions, and with higher resolutions in the spatial and spectral dimensions. In addition, we want to further optimize our hyper-parameter search to enable a faster convergence of the algorithm.

Acknowledgement. M. W., V. E. and M. G. acknowledge support for this research through the project Universal Bayesian Imaging Kit (UBIK, Förderkennzeichen 50OO2103) funded by the Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR). P. F. acknowledges funding through the German Federal Ministry of Education and Research for the project ErUM-IFT: Informationsfeldtheorie

für Experimente an Großforschungsanlagen (Förderkennzeichen: 05D23EO1). J. S. and J. K. acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2094 – 390783311.

References

Alexander, D. M., Bauer, F. E., Brandt, W. N., et al. 2003, AJ, 126, 539
Arras, P., Baltac, M., Enßlin, T. A., et al. 2019, Astrophysics Source Code Library [record ascl:1903.008]
Arras, P., Frank, P., Haim, P., et al. 2022, Nat. Astron., 6, 259
Bamba, A., Yamazaki, R., Ueno, M., & Koyama, K. 2003, ApJ, 589, 827
Bobin, J., Sureau, F., & Starck, J.-L. 2016, A&A, 591, A50
Bobin, J., Hamzaoui, I. E., Picquenot, A., & Acero, F. 2020, IEEE Trans. Image Process., 29, 9429
Broos, P. S., Townsley, L. K., Feigelson, E. D., et al. 2010, ApJ, 714, 1582
Calderwood, T., Dobrzycki, A., Jessop, H., & Harris, D. E. 2001, ASP Conf. Ser., 238, 443
Cassam-Chenaï, G., Hughes, J. P., Reynoso, E. M., Badenes, C., & Moffett, D. 2008, ApJ, 680, 1180
Chandra X-ray Center 2021, The Chandra Proposers' Observatory Guide, version 23.0
Cruddace, R. G., Hasinger, G. H., & Schmitt, J. H. 1988, in European Southern Observatory Conference and Workshop Proceedings, 28, 177
Davis, J., Bautz, M., Dewey, D., et al. 2012, Proceedings of SPIE – The International Society for Optical Engineering, 8443
Dyer, K. K., Reynolds, S. P., & Borkowski, K. J. 2003, ApJ, 600, 752
Ebeling, H., & Wiedenmann, G. 1993, Phys. Rev. E, 47, 704
Eberle, V., Frank, P., Stadler, J., Streit, S., & Enßlin, T. 2023, Entropy, 25, 652
Ellien, A., Greco, E., & Vink, J. 2023, ApJ, 951, 103
Enßlin, T. A. 2019, Ann. Phys., 531, 1800127
Enßlin, T. 2022, Entropy, 24, 374
Evans, I., Primini, F., Glotfelty, K., et al. 2010, ApJS, 189, 37
Frank, P., Leike, R., & Enßlin, T. A. 2021, Entropy, 23, 853
Freeman, P., Kashyap, V., Rosner, R., & Lamb, D. 2008, ApJS, 138, 185
Fruscione, A., Mcdowell, J. C., Allen, G. E., et al. 2006, Proc. SPIE, 6270, 62701V
Garmire, G. P., Bautz, M. W., Ford, P. G., Nousek, J. A., & Ricker, G. R. Jr. 2003, SPIE Conf. Ser., 4851, 28
Guglielmetti, F., Fischer, R., & Dose, V. 2009, MNRAS, 396, 165
Helder, E. A., Vink, J., Bykov, A. M., et al. 2012, Space Sci. Rev., 173, 369
Joye, W. A., & Mandel, E. 2003, ASP Conf. Ser., 295, 489
Katsuda, S., Petre, R., Long, K. S., et al. 2009, ApJ, 692, L105
Katsuda, S., Long, K. S., Petre, R., et al. 2013, ApJ, 763, 85
Keith Arnaud, R. S., & Siemiginowska, A. 2011, Handbook of X-ray Astronomy, Cambridge Observing Handbooks for Research Astronomers (Cambridge University Press)
Kingma, D. P., Salimans, T., & Welling, M. 2015, arXiv e-prints [arXiv:1506.02557]
Knollmüller, J., & Enßlin, T. A. 2020, arXiv e-prints [arXiv:1901.11033]
Koyama, K., Petre, R., Gotthelf, E. V., et al. 1995, Nature, 378, 255
Kullback, S., & Leibler, R. A. 1951, Ann. Math. Stat., 22, 79
Kumaran, S., Mandal, S., Bhattacharyya, S., & Mishra, D. 2023, MNRAS, 520, 5065
Li, J.-T., Decourchelle, A., Miceli, M., Vink, J., & Bocchino, F. 2015, MNRAS, 453, 3953
Long, K. S., Reynolds, S. P., Raymond, J. C., et al. 2003, ApJ, 586, 1162
McDowell, J. 2001, Coordinate Systems for Analysis of On-Orbit Chandra Data Paper I: Imaging, CXC Internal Document
McDowell, J. C. 2006, ASP Conf. Ser., 351, 47
Picquenot, A., Acero, F., Bobin, J., et al. 2019, A&A, 627, A139
Pumpe, D., Reinecke, M., & Enßlin, T. A. 2018, A&A, 619, A119
Reynolds, S. P. 2008, ARA&A, 46, 89
Scheel-Platz, L. I., Knollmüller, J., Arras, P., et al. 2022, A&A, 680, A2
Selig, M., & Enßlin, T. A. 2015, A&A, 574, A74
Seward, F. D., & Charles, P. A. 2010, Exploring the X-ray Universe, 2nd edn. (Cambridge University Press)
Stephenson, F. R. 2010, Astron. Geophys., 51, 5.27
Vink, J. 2011, A&A Rev, 20, 49
Warren, J. S., Hughes, J. P., Badenes, C., et al. 2005, ApJ, 634, 376
Willingale, R., West, R. G., Pye, J. P., & Stewart, G. C. 1996, MNRAS, 278, 749
Winkler, P. F., Gupta, G., & Long, K. S. 2003, ApJ, 585, 324
Winkler, P. F., Williams, B. J., Reynolds, S. P., et al. 2014, ApJ, 781, 65

Appendix A: Synthetic data generation

Given the generative models, we can construct prior samples of the individual components and of the imaged sky composed of them, as described in Sect. 5. Three of these prior samples are shown at the top of Fig. A.1. They illustrate how the prior samples are converted into simulated data using the instrument response and mimicking Poisson noise. As mentioned in Sect. 5, we use the prior samples and the simulated data to fine-tune the hyper-parameters prior to reconstruction. As we can see by comparison, the chosen hyper-parameters ensure that the order of magnitude of the data in Fig. 2 is the same as the order of magnitude of the simulated data. Moreover, and more importantly, the simulated data allow us to perform the validation of the algorithm, as described in Sect. 7.

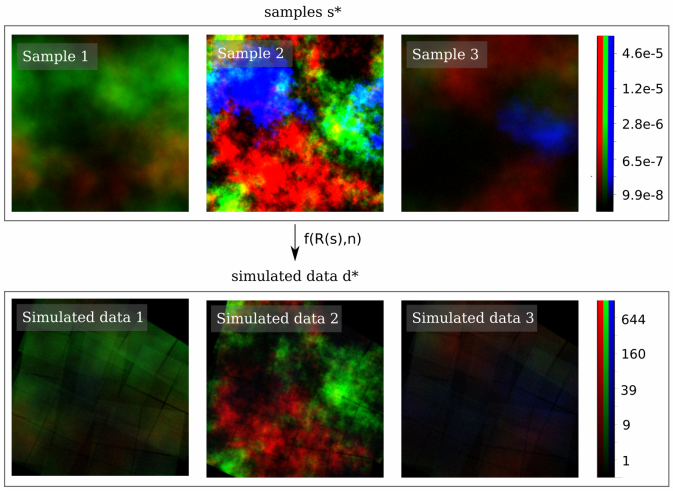


Fig. A.1: Illustration of generation of simulated data for three prior samples, showing the variance in intensity and correlation structure permitted by the prior.

Appendix B: Computational analysis

In this section, we present a comparison of the introduced algorithm including the transition model and a pure MF reconstruction in terms of computational time and reconstruction error. In case of the transition model, we started with a SF reconstruction and use the corresponding result as an initial condition for the MF reconstruction, as described in Sect. 4. In the other case, we started the reconstruction on the whole MF parameter space from the beginning. We considered four different spatial resolutions, from 64×64 to 512×512 pixels, for which we generated simulated data and perform the corresponding reconstruction on a single core for the transition model and the pure MF model. This allowed us to compare, for each problem size, the time complexity at each iteration and the reconstruction error as a function of time.

As mentioned in Sect. 4, the parameter space for the SF reconstruction is much smaller (Table 2), leading to higher computational time for each iteration in the MF reconstruction. This effect is also shown in Fig. B.1. The time complexity of each iteration is accordingly higher in the MF reconstruction than in the SF reconstruction, resulting in an overall lower time complexity for the transition model reconstruction. It can be seen that similarly to the duration of each iteration in the reconstruction, the transition time also increases with the growing parameter space.

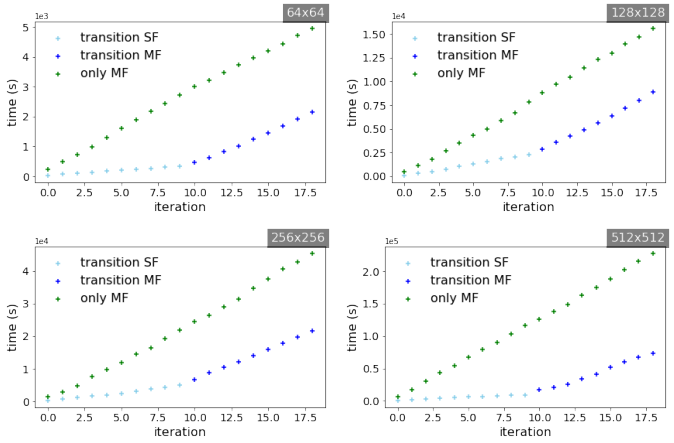


Fig. B.1: Time complexity (top-left: 64×64 spatial pixels; top-right: 128×128 spatial pixels; bottom-left: 256×256 spatial pixels; bottom-right: 512×512 spatial pixels). The time complexity is plotted for the different models. In green, we show just the MF reconstruction times per iteration. Light blue: duration for each iteration in the SF model before the transition. Dark blue: duration of the MF model iterations after the transition. The first dark blue marker also includes the transition time.

However, the increased transition time is not significant compared to the overall time savings.

What is more important for the analysis is the notion of how the reconstruction error of the reconstruction behaves over time. This is shown for the different components in Fig. B.2. As a measure of the reconstruction error, we computed the posterior sample mean for N samples of the Frobenius norm of the sample residuals $r^* = (s^* - s_{\mathrm{gt}})$ according to Eq. (21) for each component:

$$\langle \|r\|_F \rangle_{s^*} = \frac{1}{N} \sum_{n=0}^{N} \left(\sum_{i,j,k} (r_n^*)^2_{i,j,k} \right)^{\frac{1}{2}}, \tag{B.1}$$

where i, j, k are the corresponding spatial and spectral pixel indices. Due to the consideration of a smaller space in the iterations of the SF model, we would typically expect a smaller reconstruction error in terms of small Frobenius norm. It can be seen that the computational advantage of the transition model approach increases as the problem size increases in terms of a higher number of spatial pixels. This is especially true for the diffuse component, which is constructed from an outer product of correlated fields.

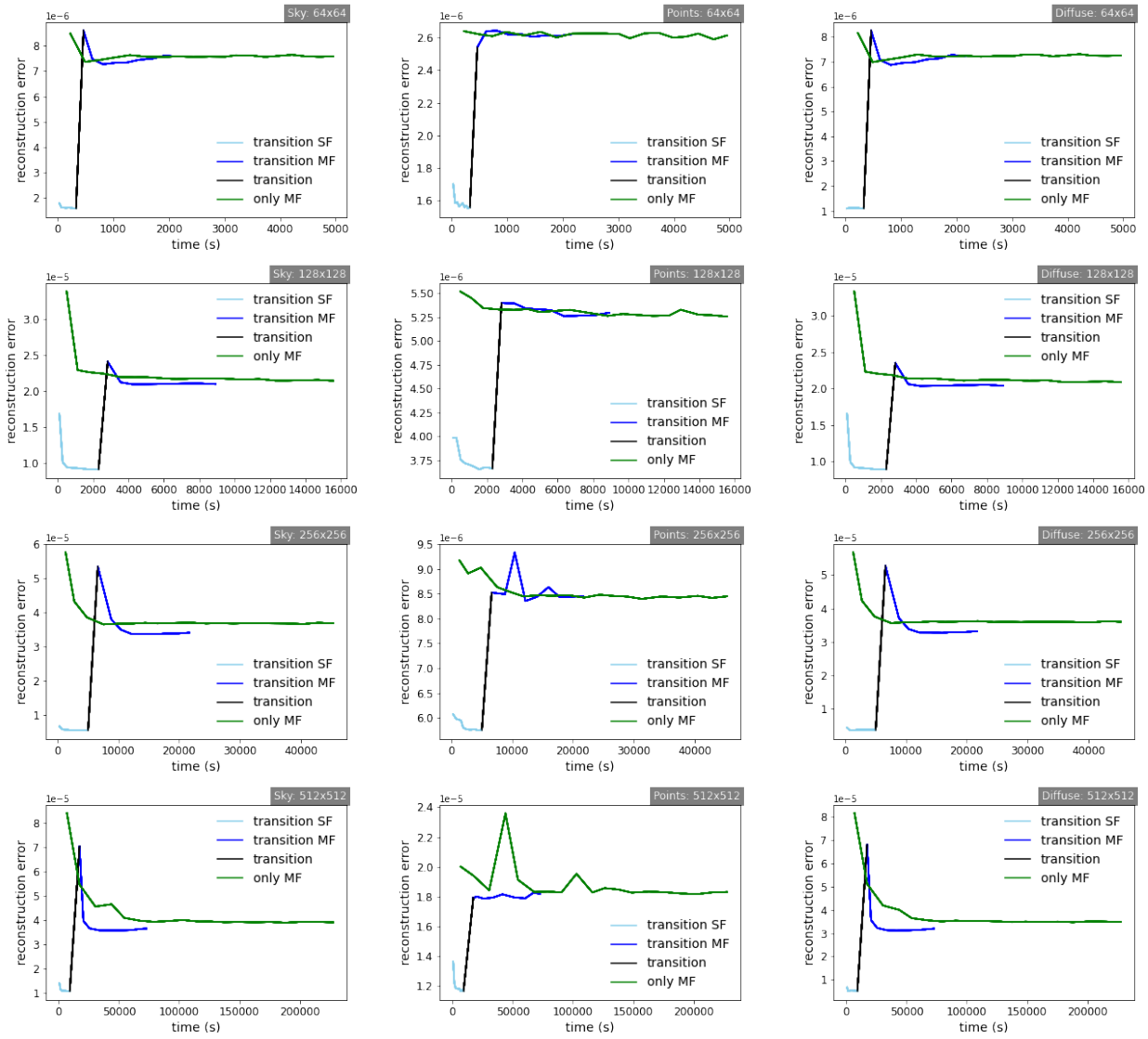


Fig. B.2: Reconstruction error in terms of the Frobenius norm (Eq. (B.1). From top to bottom: for 64×64, 128×128, 256×256, and 512×512 spatial pixels. From left to right: for the imaged sky: the point sources and the diffuse component. The green line marks the reconstruction error as a function of time for the pure MF reconstruction. The light blue line marks the reconstruction error of the SF reconstruction as part of the transition model and, correspondingly, the black line marks the transition and the blue line marks the subsequent transition model MF reconstruction. In the iterations of the SF model, we typically anticipate lower reconstruction error in terms of small Frobenius norm. This expectation is attributed to the model's consideration of a smaller space.

Appendix C: Further diagnostics for synthetic data reconstruction

In this section, we present further diagnostic plots for sanity checks on the simulated data reconstruction in Sect. 7. The analysis of these plots can be found in the according sections. We show the UWRs and residuals for the simulated data case in Fig. C.1. Figure C.2 shows the reconstruction results for the simulated data case for each energy bin together with the associated uncertainty.

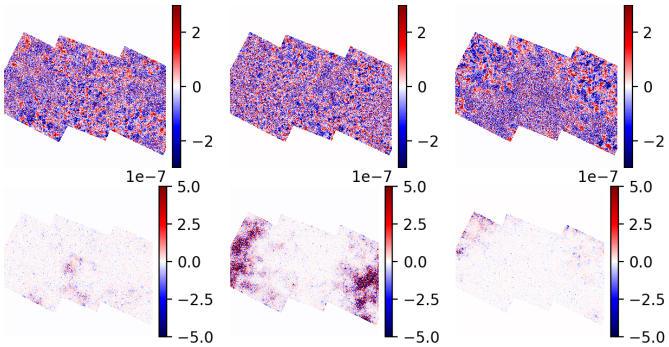


Fig. C.1: Synthetic data reconstruction UWRs (top row) and residuals (bottom row) for the individual energy bins (left: 0.5-1.2 keV, center: 1.2-2.0 keV, right: 2.0-7.0 keV) according to Eq. (20).

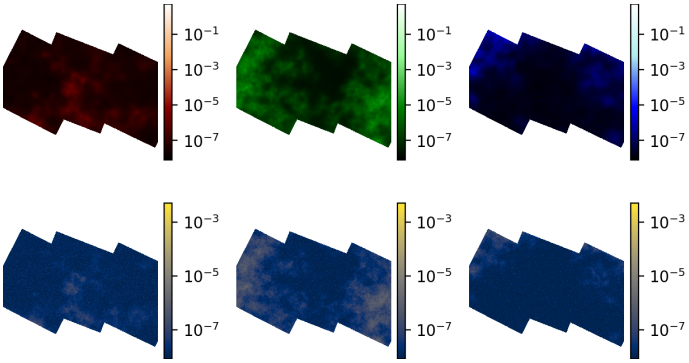


Fig. C.2: Synthetic data reconstruction uncertainties for the individual energy bins (left: 0.5-1.2 keV, center: 1.2-2.0 keV, right: 2.0-7.0 keV. Top row: Reconstruction results for the flux in [s^{-1} cm^{-2}] for the individual energy bins. Bottom row: Uncertainty maps for the individual energy bins.

Appendix D: Further diagnostics for SN1006 reconstruction

Here, we show more diagnostic plots for the analysis of the reconstruction results presented in Sect. 8. First, the reconstruction mean and posterior samples of the spatial power spectrum are shown in Fig. D.1. For further analysis of the reconstruction of the sky flux of the remnant SN1006, we present the posterior standard deviation separately for each energy bin and accompanied by the corresponding color bars in Fig. D.2. Finally, Table D.1 shows the NWRs according to Eq. (24) for each energy bin and dataset.

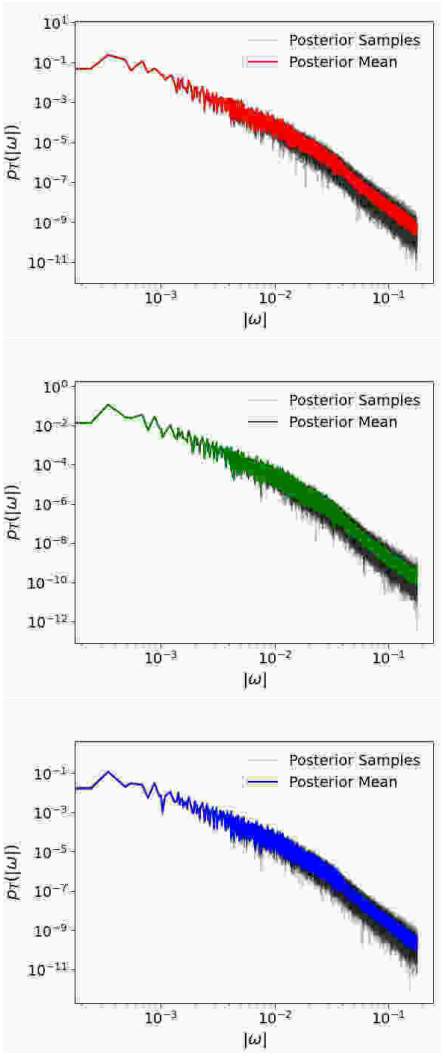


Fig. D.1: Spatial power spectra of the posterior mean and samples in the diffuse component for each energy bin (red: 0.5-1.2 keV, green: 1.2-2.0 keV, blue: 2.0-7.0 keV)

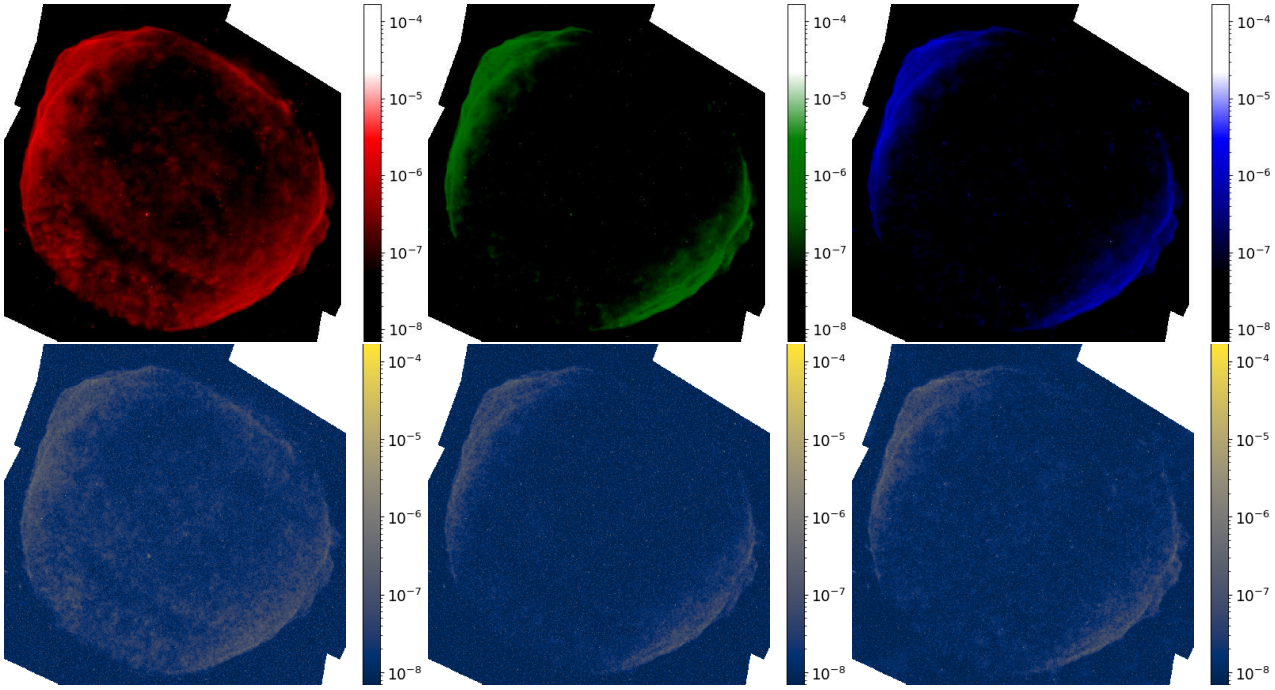
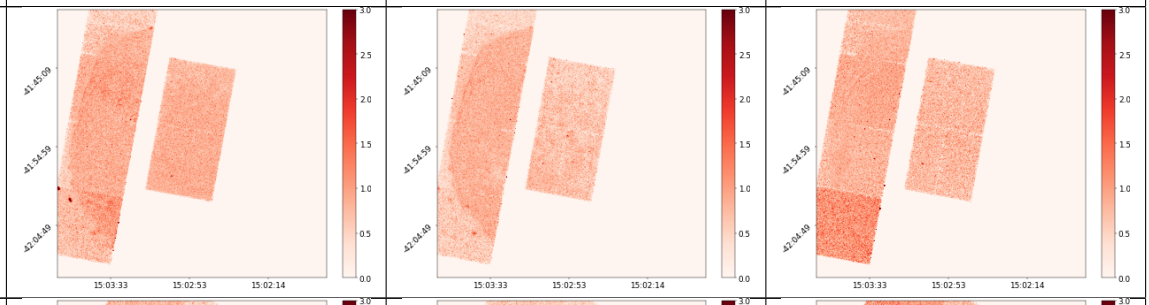
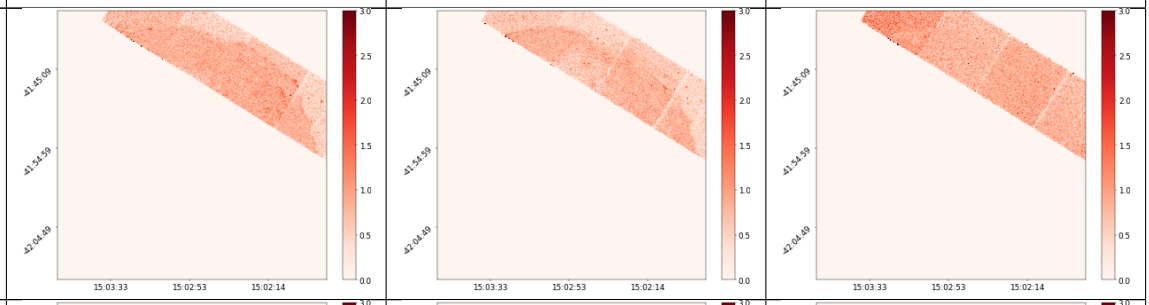
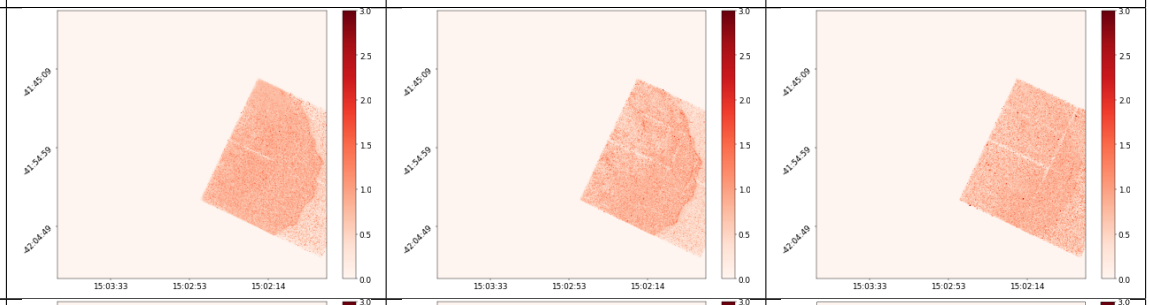
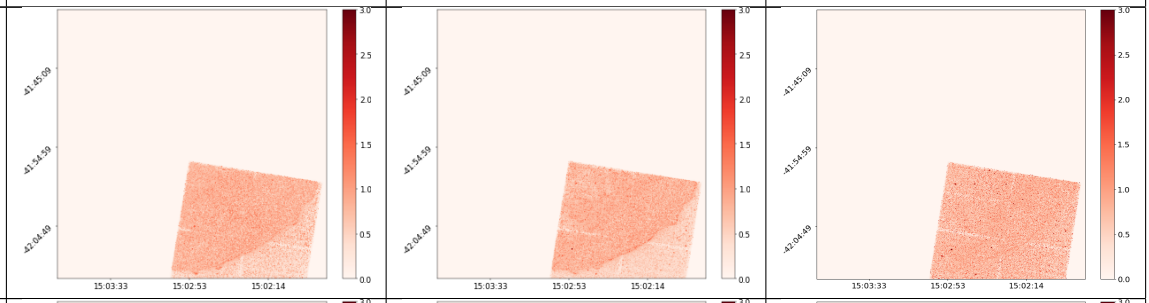
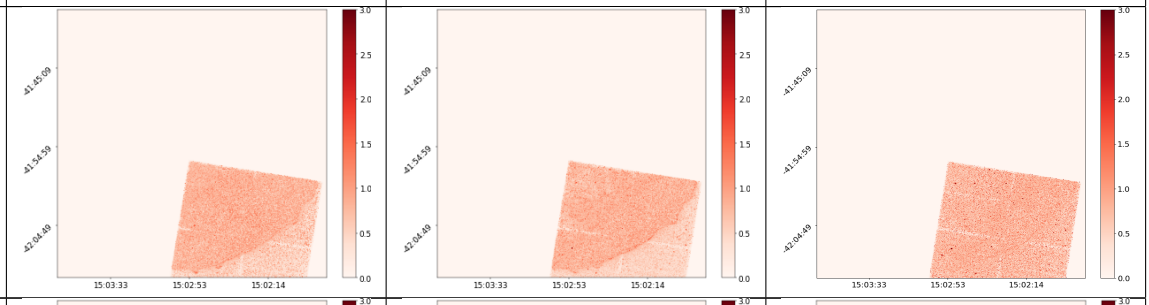
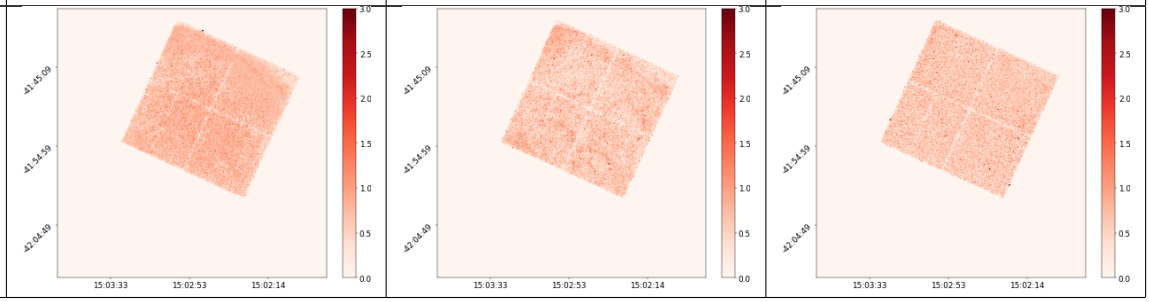


Fig. D.2: Posterior means and standard deviations for each energy bin in [s^{-1} cm^{-2}]: Top row: Posterior means (red: 0.5-1.2 keV, green:1.2-2.0 keV, blue:2.0-7.0 keV). Bottom row: Posterior standard deviations (left: 0.5-1.2 keV, center:1.2-2.0 keV, right:2.0-7.0 keV).

Patch ID	0.5-1.2keV	1.2-2.0keV	2.0-7.0keV
9107			
13737			
13738			
13739			
13739			
13740			

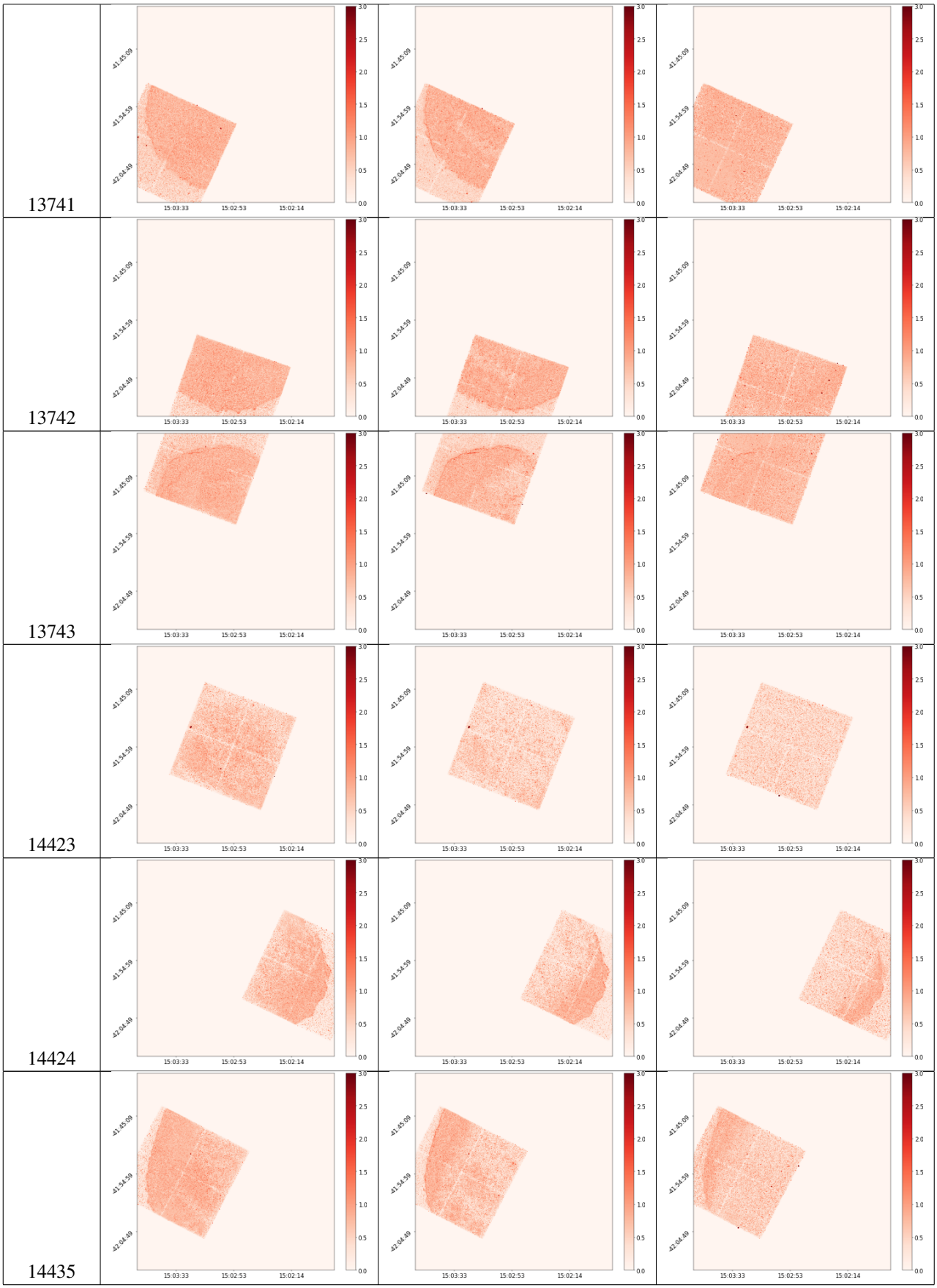


Table D.1. NWR (Eq. (24)) for each dataset in Table 3 and energy bin.