Astronomy & Astrophysics manuscript no. articlemc
October 10, 2011

Multichannel Poisson Denoising and Deconvolution on the Sphere : Application to the Fermi Gamma Ray Space Telescope

J. Schmitt[1], J.L. Starck[1], J.M. Casandjian[1], J. Fadili[2], and I. Grenier[1]

[1] CEA, Laboratoire AIM, CEA/DSM-CNRS-Université Paris Diderot,
CEA, IRFU, Service d'Astrophysique, Centre de Saclay, F-91191 Gif-Sur-Yvette cedex, France
[2] GREYC CNRS-ENSICAEN-Université de Caen, 6, Bd du Maréchal Juin, 14050 Caen Cedex, France

October 10, 2011

Abstract

A multiscale representation-based denoising method for spherical data contaminated with Poisson noise, the multiscale variance stabilizing transform on the sphere (MS-VSTS), has been recently proposed. This paper first extends this MS-VSTS to spherical 2D-1D, where the two first dimensions are longitude and latitude, and the third dimension is a meaningful physical index such as energy or time. Then we introduce a novel multichannel deconvolution built upon the 2D-1D MS-VSTS, which allows to get rid of both the noise and the blur introduced by the point spread function (PSF) in each energy (or time) band. The method is applied to simulated data from the Large Area Telescope (LAT), the main instrument of the Fermi Gamma-Ray Space Telescope, which detects high energy gamma-rays in a very wide energy range (from 20 MeV to more than 300 GeV), and whose PSF is strongly energy-dependent (from about 3.5° at 100 MeV to less than 0.1° at 10 GeV).

Key words. Keywords should be given

Methods: Data Analysis, Techniques: Image Processing.

Contents

1. Introduction

1.1. Literature overview

The gamma-ray sky is currently studied with unprecedented sensitivity and image capability thanks to the Large Area Telescope (LAT), the main instrument of the Fermi Gamma-Ray Space Telescope (Atwood et al. 2009), in an energy range between 20 MeV to greater than 300 GeV. The detection of gamma-ray point sources is made difficult by two main factors : the Poisson noise and the instrument's point spread function (PSF). The Poisson noise is due to the weakness of the fluxes of celestial gamma-rays, especially outside the galactic plane and far away from intense sources. The PSF width is strongly energy-dependent, it varies from about 3.5° at 100 MeV to less than 0.1° (68% containment) at 10 GeV. Owing to large-angle multiple scattering in the tracker, the PSF has broad tails, the 95%/68% containment ratio may be as large as 3.

An extensive literature exists on Poisson noise removal and the interested reader may refer to Schmitt et al. (2011); Starck et al. (2010) for a thorough review. Motivated by new X-ray and gamma-ray data challenges, several restoration methods have been recently released in astrophysics that are based on wavelets (Movit 2009; Starck et al. 2009; Faÿ et al. 2011; Schmitt et al. 2010) or the Bayesian machinery (Conrad et al. 2007; Norris et al. 2011). Wavelets have also been used for source detection in FERMI data (Abdo et al. 2010), and a first Poisson denoising algorithm for spherical data was proposed in Schmitt et al. (2010). Starck et al. (2009) developed a denoising approach that handles effectively multichannel data acquired on a Cartesian grid, and where the third dimension can be any physically meaningful index such as wavelength, energy or time. While traditional techniques would integrate all the third dimension in oder to improve the signal-to-noise ratio (SNR) of the sources, their approach allows to detect the sources while preserving their whole spectral information and without scarifying the sensitivity. Nonetheless, none of of these methods take into account the point spread function (PSF) of the instrument. In the other hand, deconvolution can prove very helpful in many situations such as sources identification or flux estimation in the low energy bands where the resolution degrades severely. To the best of our knowledge, no general method for both denoising and deconvolving multichannel on the sphere has been developed in the literature.

1.2. Contributions

In this paper, we propose a general framework for denoising and deconvolution that comply with all the above requirements: deal with (i) multichannel data, (ii) acquired on the sphere, and (iii) contaminated by Poisson noise. Our approach builds upon the concept of *variance stabilization* applied wisely to the spherical wavelet transform coefficients (Starck et al. 2006). Doing so, the transformed and variance stabilized coefficients can be treated as if they were contaminated by a zero-mean white Gaussian noise. The developed algorithms are validated on simulated Fermi HEALPix multichannel cubes (nside = 256) with energy bands ranging from 50 MeV to 50 GeV.

1.3. Paper organization

The rest of the paper is organized as follows. In section 2, we present a new wavelet transform for multichannel data on the sphere, and we show how the variance stabilization transform can be introduced in the decomposition. Section 3 details the way this transform can be used for Gaussian and Poisson noise removal. Section 4 describes our deconvolution algorithm on the sphere for both mono-channel and multichannel spherical data. In Section 5, we finally draw some conclusions and give possible perspectives of this work.

1.4. Notations

For a real discrete-time filter whose impulse response is $h[i]$, $\bar{h}[i] = h[-i]$, $i \in \mathbb{Z}$ is its time-reversed version. The discrete circular convolution product of two signals will be written $\star$, where the term circular stands for periodic boundary conditions. The symbol $\delta[i]$ is the Kronecker delta.

For the wavelet representation, the low-pass analysis filter is denoted h and the high-pass is taken as $g = \delta - h$ throughout the paper. We denote the up-sampled version of h as $h^{\uparrow j}[l] = h[l]$ if $l/2^j \in \mathbb{Z}$ and 0 otherwise. Define $h^{(j)} = \bar{h}^{\uparrow j-1} \star \cdots \star \bar{h}^{\uparrow 1} \star \bar{h}$ for $j \geqslant 1$ and $h^{(0)} = \delta$.

The scaling and wavelet functions used for the analysis (respectively, synthesis) are denoted ϕ (with $\phi(\frac{x}{2}) = \sum_k h[k]\phi(x-k), x \in \mathbb{R}$ *and* $k \in \mathbb{Z}$) and ψ (with $\psi(\frac{x}{2}) = \sum_k g[k]\phi(x-k), x \in \mathbb{R}$ *and* $k \in \mathbb{Z}$) (respectively, $\widetilde{\phi}$ and $\widetilde{\psi}$).

2. Multiscale Representation for Multichannel Spherical Data with Poisson Noise

2.1. Fast Undecimated 2D-1D Wavelet Decomposition/Reconstruction on the Sphere

Our goal is to analyze multichannel data acquired on a sphere with a non-isotropic 2D-1D wavelet, where the two first dimensions are spatial (longitude and latitude) and the third dimension is either the time or the energy. Since the dimensions do not have the same physical meaning, it appears natural that the wavelet scale along the third dimension (energy or time) should not be connected to the spatial scale. Hence, we define the wavelet function as

$$\psi(k_\theta, k_\varphi, k_t) = \psi^{(\theta\phi)}(k_\theta, k_\varphi)\psi^{(t)}(k_t), \tag{1}$$

where $\psi^{(\theta\phi)}$ is the spherical spatial wavelet and $\psi^{(t)}$ the 1D wavelet along the third dimension. Similarly to Starck et al. (2009), we consider only isotropic and dyadic spatial scales, and we build the discrete 2D-1D wavelet decomposition by first taking a spherical 2D undecimated wavelet transform for each k_t, followed by a 1D wavelet transform for each spatial wavelet coefficient along the third dimension.

Hence for a given multichannel data set on the sphere $Y[k_\theta, k_\varphi, k_t]$, applying first the 2D spherical undecimated wavelet transform, we have the reconstruction formula

$$Y[k_\theta, k_\varphi, k_t] = a_{J_1}[k_\theta, k_\varphi, k_t] + \sum_{j_1=1}^{J1} w_{j_1}[k_\theta, k_\varphi, k_t], \qquad \forall k_t, \tag{2}$$

where J_1 is the number of spatial scales, a_{J_1} is the (spatial) approximation subband and $\{w_{j_1}\}_{j_1=1}^{J_1}$ are the (spatial) detail subbands. To lighten the notations in the sequel, we replace the two spatial indices by a single index k_r which corresponds to the pixel index. Expression (2) reads now

$$Y[k_r, k_t] = a_{J_1}[k_r, k_t] + \sum_{j_1=1}^{J1} w_{j_1}[k_r, k_t], \qquad \forall k_t. \tag{3}$$

Then, for each spatial location k_r and for each 2D wavelet scale j_1, we apply a 1D wavelet transform along t on the spatial wavelet coefficients $w_{j_1}[k_r, \cdot]$ such that

$$w_{j_1}[k_r, k_t] = w_{j_1,J_2}[k_r, k_t] + \sum_{j_2=1}^{J_2} w_{j_1,j_2}[k_r, k_t], \quad \forall (k_r, k_t), \tag{4}$$

where J_2 is the number of scales along t. The approximation spatial subband a_{J_1} is processed in a similar way, hence yielding

$$a_{J_1}[k_r, k_t] = a_{J_1,J_2}[k_r, k_t] + \sum_{j_2=1}^{J_2} w_{J_1,j_2}[k_r, k_t], \qquad \forall (k_r, k_t). \tag{5}$$

Inserting (4) and (5) into (3), we obtain the 2D-1D spherical undecimated wavelet representation of Y:

$$Y[k_r, k_t] = a_{J_1,J_2}[k_r, k_t] + \sum_{j_1=1}^{J_1} w_{j_1,J_2}[k_r, k_t] + \sum_{j_2=1}^{J_2} w_{J_1,j_2}[k_r, k_t] + \sum_{j_1=1}^{J_1}\sum_{j_2=1}^{J_2} w_{j_1,j_2}[k_r, k_t]\ . \tag{6}$$

In this expression, four kinds of coefficients can be distinguished:

– Detail-Detail coefficients ($j_1 \leqslant J_1$ and $j_2 \leqslant J_2$):

$$w_{j_1,j_2}[k_r, \cdot] = (\delta - \bar{h}_{1\mathrm{D}}) \star (h_{1\mathrm{D}}^{(j_2-1)} \star a_{j_1-1}[k_r, \cdot] - h_{1\mathrm{D}}^{(j_2-1)} \star a_{j_1}[k_r, \cdot]). \tag{7}$$

– Approximation-Detail coefficients ($j_1 = J_1$ and $j_2 \leqslant J_2$):

$$w_{J_1,j_2}[k_r, \cdot] = h_{1\mathrm{D}}^{(j_2-1)} \star a_{J_1}[k_r, \cdot] - h_{1\mathrm{D}}^{(j_2)} \star a_{J_1}[k_r, \cdot]. \tag{8}$$

– Detail-Approximation coefficients ($j_1 \leqslant J_1$ and $j_2 = J_2$):

$$w_{j_1,J_2}[k_r,\cdot] = h_{\mathrm{1D}}^{(J_2)} \star a_{j_1-1}[k_r,\cdot] - h_{\mathrm{1D}}^{(J_2)} \star a_{j_1}[k_r,\cdot]. \tag{9}$$

– Approximation-Approximation coefficients ($j_1 = J_1$ and $j_2 = J_2$):

$$a_{J_1,J_2}[k_r,\cdot] = h_{\mathrm{1D}}^{(J_2)} \star a_{J_1}[k_r,\cdot]. \tag{10}$$

2.2. Multi-Scale Variance Stabilizing Transform on the Sphere (MS-VSTS)

Schmitt et al. (2010) proposed a multiscale variance stabilizing transform adapted for spherical data corrupted with Poisson noise. This transform was dubbed the Multi-Scale Variance Stabilizing Transform on the Sphere (MS-VSTS). In a nutshell, the MS-VSTS consists in plugging a variance stabilizing transform (VST) within a multi-scale transform–the Isotropic Undecimated Wavelet Transform on the Sphere (IUWTS)– in order to get (approximately) Gaussian zero-mean multiscale coefficients with constant variance.

The MS-VSTS scheme is defined recursively by inserting a (nonlinear) square-root VST within the IUWTS steps. That is,

$$\text{IUWTS} \left\{ \begin{array}{rcl} a_j &=& h_{j-1} \star a_{j-1} \\ d_j &=& a_{j-1} - a_j \end{array} \right. \Longrightarrow \begin{array}{c} \text{MS-VSTS} \\ \text{(VST + IUWTS)} \end{array} \left\{ \begin{array}{rcl} a_j &=& h_{j-1} \star a_{j-1} \\ d_j &=& T_{j-1}(a_{j-1}) - T_j(a_j) \end{array} \right. , \tag{11}$$

where T_j is the VST operator at scale j:

$$T_j(a_j) = b^{(j)} \mathrm{sign}(a_j + c^{(j)}) \sqrt{|a_j + c^{(j)}|}, \tag{12}$$

with the VST constants $b^{(j)}$ and $c^{(j)}$ that depend solely on the filter h and the scale level j. It has been shown in Zhang et al. (2008) that the MS-VSTS detail coefficients d_j on locally homogeneous parts of the underlying intensity signal follow asymptotically a zero-mean normal distribution with an intensity-independent variance that relies only on the filter h and the current scale j. Consequently, both the stabilized variances and the constants $b^{(j)}$ and $c^{(j)}$ can be pre-computed once and for all for a given h (Schmitt et al. 2010).

2.3. Multichannel MS-VSTS

This section extends the MS-VSTS machinery to the multichannel case. This amounts to wisely plugging the VST into the spherical 2D-1D undecimated wavelet transform introduced in Section 2.1. Again, this will give rise to four types of coefficients that take the following forms:

– Detail-Detail coefficients ($j_1 \leqslant J_1$ and $j_2 \leqslant J_2$):

$$w_{j_1,j_2}[k_r,\cdot] = (\delta - \bar{h}_{\mathrm{1D}}) \star \left(T_{j_1-1,j_2-1} \left(h_{\mathrm{1D}}^{(j_2-1)} \star a_{j_1-1}[k_r,\cdot] \right) - T_{j_1,j_2-1} \left(h_{\mathrm{1D}}^{(j_2-1)} \star a_{j_1}[k_r,\cdot] \right) \right). \tag{13}$$

– Approximation-Detail coefficients ($j_1 = J_1$ and $j_2 \leqslant J_2$):

$$w_{J_1,j_2}[k_r,\cdot] = T_{J_1,j_2-1} \left(h_{\mathrm{1D}}^{(j_2-1)} \star a_{J_1}[k_r,\cdot] \right) - T_{J_1,j_2} \left(h_{\mathrm{1D}}^{(j_2)} \star a_{J_1}[k_r,\cdot] \right). \tag{14}$$

– Detail-Approximation coefficients ($j_1 \leqslant J_1$ and $j_2 = J_2$):

$$w_{j_1,J_2}[k_r,\cdot] = T_{j_1-1,J_2} \left(h_{\mathrm{1D}}^{(J_2)} \star a_{j_1-1}[k_r,\cdot] \right) - T_{j_1,J_2} \left(h_{\mathrm{1D}}^{(J_2)} \star a_{j_1}[k_r,\cdot] \right). \tag{15}$$

– Approximation-Approximation coefficients ($j_1 = J_1$ and $j_2 = J_2$):

$$a_{J_1,J_2}[k_r,\cdot] = h_{\mathrm{1D}}^{(J_2)} \star a_{J_1}[k_r,\cdot]. \tag{16}$$

In summary, all 2D-1D wavelet coefficients $\{w_{j_1,j_2}\}_{j_1 \leqslant J_1, j_2 \leqslant J_2}$ are now stabilized, and the noise on all these wavelet coefficients is zero-mean Gaussian with known variance that depends solely on h on the resolution levels (j_1, j_2). As before, these variances can be easily tabulated.

3. Application to Multichannel Denoising

Let X be the noiseless data and Y their observed noisy version. For instance, for the additive zero-mean white Gaussian noise we have $Y \sim \mathcal{N}(X, \sigma^2)$, and for the Poisson noise $Y \sim \mathcal{P}(X)$. The goal of denoising is to estimate X from Y.

3.1. Warm up: Gaussian noise

We start with the simple and instructive case where the noise in Y is additive white Gaussian. As the spherical 2D-1D undecimated wavelet transform described in Section 2.1 is linear, the noise remains Gaussian in the transform domain. Therefore, the thresholding strategies which have been developed for wavelet Gaussian denoising can be applied to the spherical 2D-1D wavelet transform. Denoting TH the thresholding operator, the denoised 2D-1D estimate of X obtained by thresholding the wavelet coefficients in (6) reads

$$\widetilde{X}[k_r, k_t] = a_{J_1,J_2}[k_r, k_t] + \sum_{j_1=1}^{J_1} \mathrm{TH}(w_{j_1,J_2}[k_r, k_t]) + \sum_{j_2=1}^{J_2} \mathrm{TH}(w_{J_1,j_2}[k_r, k_t]) + \sum_{j_1=1}^{J_1} \sum_{j_2=1}^{J_2} \mathrm{TH}(w_{j_1,j_2}[k_r, k_t]). \tag{17}$$

A typical choice of TH is the hard thresholding operator parametrized by the scalar threshold $\tau \geqslant 0$, i.e.

$$\mathrm{TH}(x) = \begin{cases} 0 & \text{if } |x| < \tau, \\ x & \text{otherwise.} \end{cases} \tag{18}$$

The threshold τ is typically chosen between 3 and 5 times the noise standard deviation.

3.2. Poisson noise

The case of Poisson noise is much more involved than its Gaussian counterpart, the main reason being that the noise variance is equal to its mean. This is where the MS-VSTS comes into play. Indeed, the role of the MS-VSTS is precisely to get rid of this dependence of the variance on the mean by ensuring that the transformed coefficients are Gaussian with constant variance (without loss of generality, this variance can be assumed equal to 1). In other words, after the MS-VSTS, we are brought to a Gaussian denoising problem where standard thresholding approaches apply.

Having said this, denoising is nevertheless not straightforward because there is no explicit reconstruction formula available because of the form of the non-linear stabilization equations above. Formally, the stabilizing operators T_{j_1,j_2} and the convolution operators along the spatial and the third dimensions do not commute, even though the filter bank satisfies the exact reconstruction formula. To circumvent this difficulty, we propose to solve this reconstruction problem by advocating an iterative reconstruction scheme.

3.3. Iterative Reconstruction

Let $\mathbf{W}$ be the transform operator associated to the 2D-1D IUWTS described in Section 2.1, and $\mathbf{R}$ be its inverse transform. Let's define the multiresolution support which is determined by the set of significant coefficients detected among $\mathbf{W}Y$ at each scale (j_1, j_2) and location (k_r, k_t), i.e.

$$\mathcal{M} = \{(j_1, j_2, k_r, k_t) \big| (\mathbf{W}Y)_{j_1,j_2}[k_r, k_t] \text{ is significant}\}. \tag{19}$$

Let $\mathbf{M}$ be the orthogonal projector onto $\mathcal{M}$, i.e. $\forall d$

$$(\mathbf{M}d)_{j_1,j_2}[k_r, k_t] = \begin{cases} (\mathbf{W}Y)_{j_1,j_2}[k_r, k_t] & \text{if } (j_1, j_2, k_r, k_t) \in \mathcal{M}, \\ d_{j_1,j_2}[k_r, k_t] & \text{otherwise.} \end{cases} \tag{20}$$

Our goal is to seek a solution $\widetilde{X}$ that preserves the significant structures of the original data by reproducing exactly the same wavelet coefficients as those of the input data Y, but only at scales and positions where significant coefficients have been detected. Furthermore, as Poisson intensity functions are positive by nature, a positivity constraint is imposed on the solution. It is clear that there are many solutions satisfying the positivity and multiresolution support consistency requirements, e.g. Y itself. Thus, our reconstruction problem based solely on these constraints is an ill-posed inverse problem that must be regularized. Typically, the solution in which we are interested must be sparse by involving the lowest budget of wavelet coefficients. Therefore our

reconstruction is formulated as a constrained sparsity-promoting minimization problem over the transform coefficients d

$$\min_d \|d\|_1 \text{ subject to } \begin{cases} d_{j_1,j_2}[k_r,k_t] = (\mathbf{W}Y)_{j_1,j_2}[k_r,k_t], \forall (j_1,j_2,k_r,k_t) \in \mathcal{M}, \\ X \succcurlyeq 0. \end{cases} \quad (21)$$

and the intensity estimate $\widetilde{X}$ is reconstructed as $\widetilde{X} = \mathbf{R}\tilde{d}$, where $\tilde{d}$ is a global minimizer of (21). Recall that $\|\cdot\|_1 = \sum_{j_1,j_2,k_r,k_t} |d_{j_1,j_2}[k_r,k_t]|$ is the ℓ_1-norm playing the role of regularization and is well known to promote sparsity (Donoho 2004). This problem can be solved efficiently using the hybrid steepest descent algorithm (Yamada 2001) (Zhang et al. 2008), and requires about 10 iterations in practice. Transposed into our context, its main steps can be summarized as follows:

Require: Input noisy data Y, a low-pass filter h, multiresolution support $\mathcal{M}$ from the detection step, number of iterations $N_{\max}$.
1: Initialize $d^{(0)} = \mathbf{M}\mathbf{W}Y$.
2: **for** $n = 1$ to $N_{\max}$ **do**
3: $\bar{d}^{(n)} = \mathbf{M}d^{(n-1)}$;
4: $d^{(n)} = \mathbf{W}\mathbf{P}_+\left(\mathbf{R}\ \mathrm{ST}_{\beta_n}(\bar{d}^{(n)})\right)$;
5: Update the step $\beta_n = (N_{\max} - n)/(N_{\max} - 1)$.
6: **end for**
7: **return** $\widetilde{X} = \mathbf{R}d^{(N_{\max})}$.

where $\mathbf{P}_+$ is the orthogonal projector onto the positive orthant, ST_{β_n} is the entry-wise soft-thresholding operator with threshold β_n, i.e. for $x \in \mathbb{R}$, $\mathrm{ST}_{\beta_n}(x) = \max(0, 1 - \beta_n/|x|)x$.

The final multichannel MS-VSTS Poisson noise removal algorithm is summarized in the following steps:

Require: Input noisy data Y, a low-pass filter h, threshold level τ.
1: Multichannel spherical MS-VST: Apply the 2D-1D MS-VSTS to Y using (13)-(16).
2: Detection: Detect the significant coefficients that are above τ, and get the multiresolution support $\mathcal{M}$.
3: Reconstruction: Apply the above algorithm with $\mathcal{M}$ to get the denoised data $\widetilde{X}$.

3.4. Experiments

The multichannel MS-VSTS algorithm has been applied to a simulated Fermi data set, with 14 energy bands between 50 MeV and 1.58 GeV. Figures 1 and 2 depict the denoising results on two energy bands. The algorithm is able to recover most of the sources, even the faint ones, on each energy band. Even more importantly, the 2D-1D MS-VSTS denoising algorithm allows to recover the spectral information for each spatial position, as it can be seen from Figure 3.

4. Deconvolution of Spherical Data with Poisson Noise

In this section, we introduce a wavelet deconvolution approach for monochannel and multichannel data on the sphere with Poisson noise. The main idea underlying the method is to capitalize on the MS-VSTS described above. We first introduce the deconvolution problem and then describe how the MS-VSTS can be used to solve the deconvolution problem.

4.1. Problem statement

Many problems in signal and image processing can be cast as inverting the linear system:

$$Y = \mathbf{H}X + \varepsilon\ , \quad (22)$$

where $X \in \mathcal{X}$ is the data to recover, $Y \in \mathcal{Y}$ is the degraded noisy observation, ε is an additive noise, and $\mathbf{H} : \mathcal{X} \to \mathcal{Y}$ is a bounded linear operator which is typically ill-behaved since it models

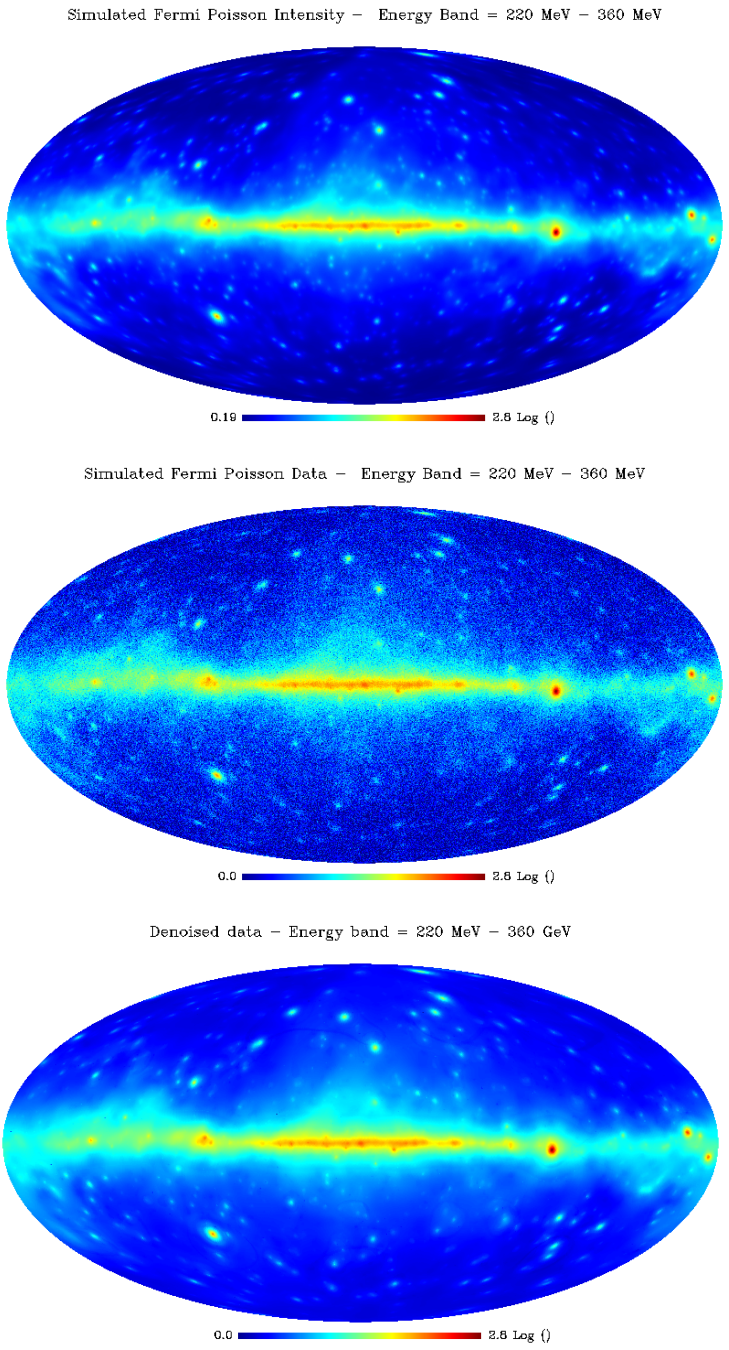

Figure 1. Result of the multichannel Poisson denoising algorithm on simulated Fermi data over the energy band 220 MeV - 360 MeV. Top: Simulated intensity skymap. Middle: Simulated noisy skymap. Bottom: denoised skymap. Maps are in logarithmic scale.

an acquisition process that encounters loss of information. When $\mathbf{H}$ is the identity, it is just a denoising problem which can be treated with the previously described methods. Inverting (22) is usually an ill-posed problem. This means that there is no unique and stable solution.

Our objective is to remove the effect of the instrument's PSF. In our case, $\mathbf{H}$ is the convolution operator by a blurring kernel (i.e. PSF) whose consequence is that Y lacks the high frequency content of X. Furthermore, since the noise is Poisson, ε has a variance profile $\mathbf{H}X$. The problem at hand is then a deconvolution problem in the presence of Poisson noise.

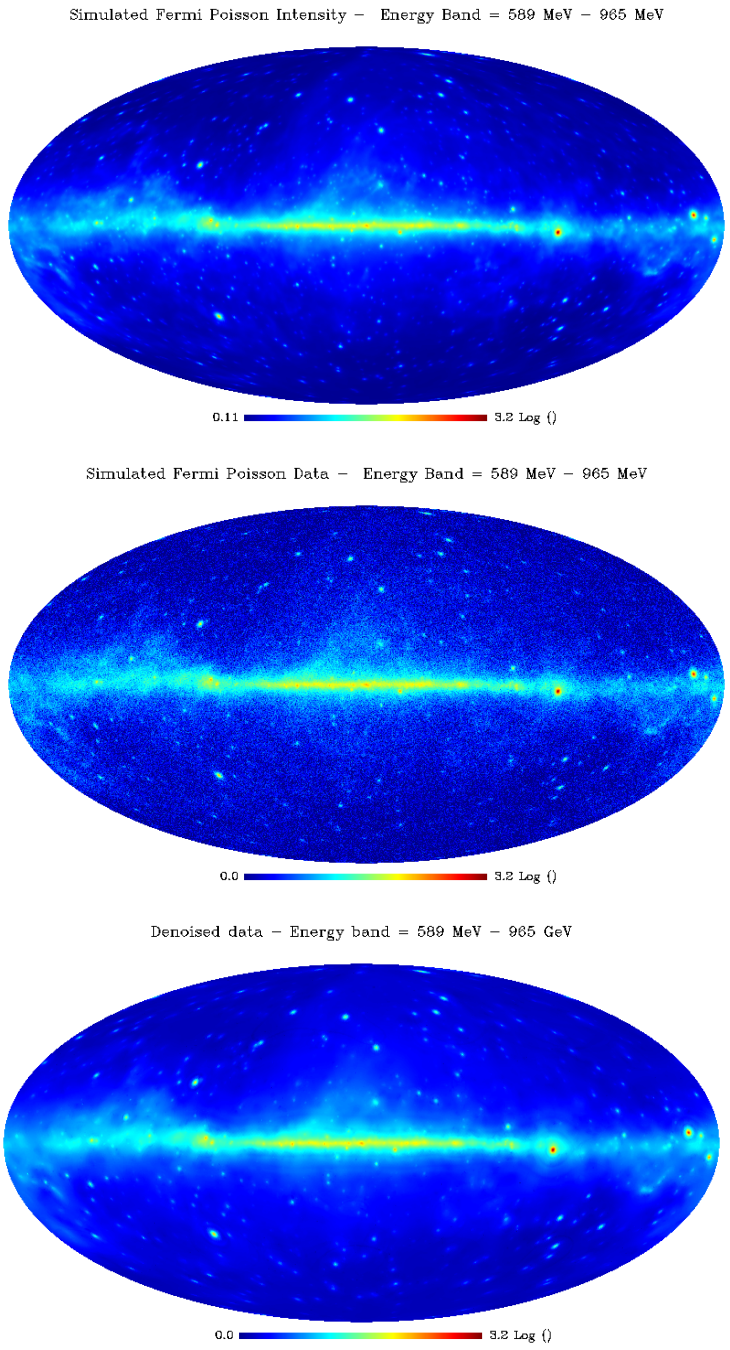

Figure 2. Result of the multichannel Poisson denoising algorithm on simulated Fermi data over the energy band 589 MeV - 965 MeV. Top: Simulated intensity skymap. Middle: Simulated noisy skymap. Bottom: denoised skymap. Maps are in logarithmic scale.

Therefore, we need to both regularize the solution and handle the Poisson statistics of the noise. In order to regularize such an inversion problem and reduce the space of candidate solutions, one has to add some prior knowledge on the typical structure of the original data X. This prior information accounts for the smoothness of the solution and can range from the uniform smoothness assumption to more complex knowledge of the geometrical structures of X.

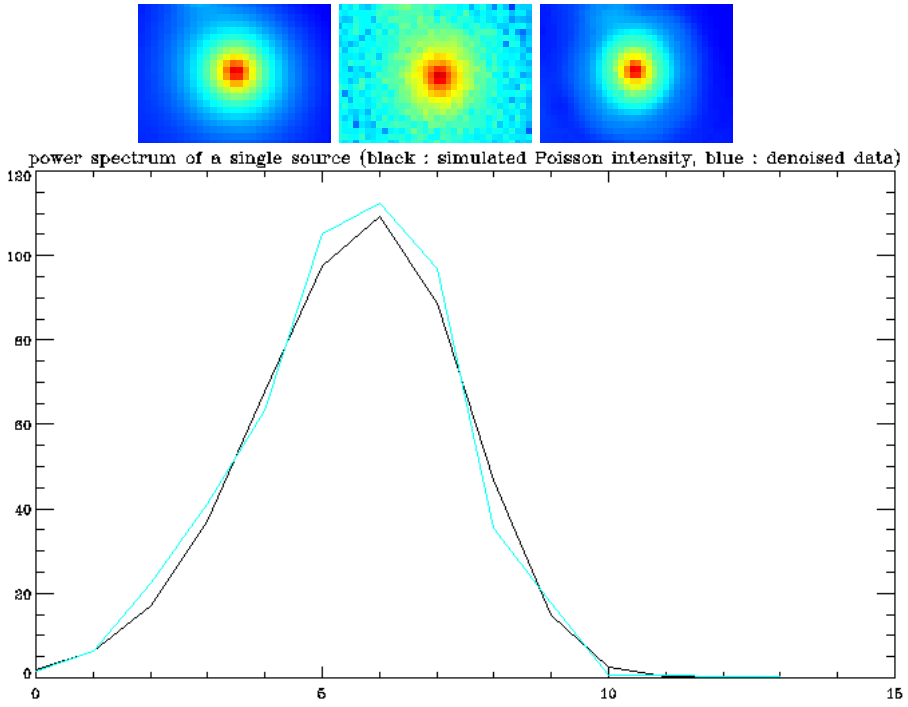

Figure 3. Spectrum of a single gamma-ray point source recovered using the multichannel MS-VSTS denoising algorithm. Top: Single gamma-ray source from simulated Fermi data integrated along the energy axis (left: simulated source; middle: noisy source; right: denoised source). Bottom: Spectrum of the center of the point source: intensity as a function of the energy band with 14 energy bands between 50 MeV and 50 GeV (black: true simulated spectrum; cyan: restored spectrum from our denoising algorithm.

In our LAT realistic simulations, the PSF width depends strongly on the energy, from 6.9° at 50 MeV to better than 0.1° at 10 GeV and above. Figure 4 shows the normalized profiles of the PSF for different energy bands.

4.2. Monochannel Deconvolution

Let's start with the single-channel case. In the literature, several algorithms have been proposed for image deconvolution on the Cartesian grid. The Richardson-Lucy algorithm is certainly the most famous in astrophysics. In this paper, we propose a regularized Richardson-Lucy algorithm for deconvolving data on the sphere data.

The Richardson-Lucy algorithm originates from a fixed-point equation obtained by maximizing the Poisson likelihood with respect to X while preserving positivity. This entails a multiplicative update rule, starting at $n = 0$ and $X^{(0)} = 1$ and iterating

$$X^{(n+1)} = X^{(n)} \otimes \left(\mathbf{H}^T (Y \oslash \mathbf{H} X^{(n)}) \right), \tag{23}$$

where $\otimes$ (resp. $\oslash$) stands for the element-wise multiplication (resp. division) between two vectors. $\mathbf{H}^T$ is the transpose of $\mathbf{H}$ whose action on an image consists in convolving it with the time-reversed version of the PSF associated to $\mathbf{H}$. However, it is well-known that because of the lack of regularization, the Richardson-Lucy algorithm tends to amplify the noise after a few iterations.

Let's define $R^{(n)}$ as the residual at iteration n

$$R^{(n)} = Y - \mathbf{H} X^{(n)}. \tag{24}$$

$R^{(n)}$ can be written as the sum of its IUWTS detail subband $\{d_j\}_{1 \leqslant j \leqslant J}$ and the last approximation subband a_J. That is,

$$R^{(n)}[k_r] = a_J[k_r] + \sum_{j=1}^{J} d_j[k_r], \quad \forall k_r. \tag{25}$$

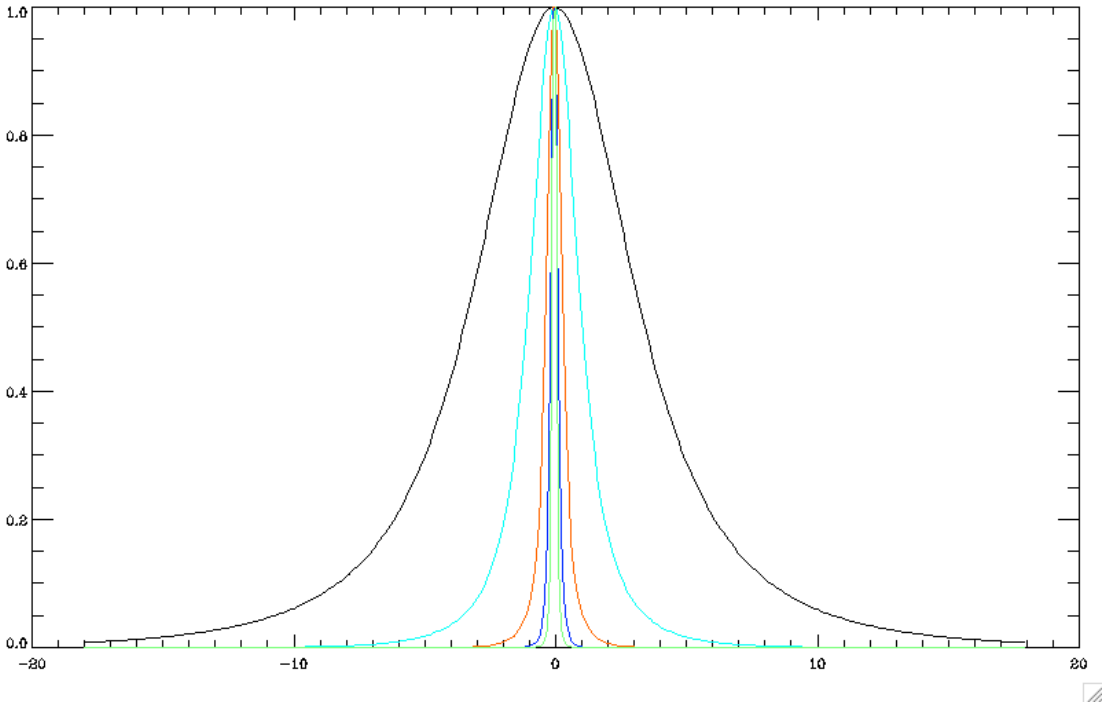

Figure 4. Normalized profile of the PSF for different energy bands as a function of the angle in degree. Black: 50 MeV - 82 MeV. Cyan: 220 MeV - 360 MeV. Orange: 960 MeV - 1.6 GeV. Blue: 4.2 GeV - 6.9 GeV. Green: 19 GeV - 31 GeV.

The wavelet transform provides a mean to extract only the significant structures from the residual at each iteration. In fact, with increasing iteration, a large part of the residual becomes statistically insignificant. The regularized significant residual is then, for a location k_r,

$$\bar{R}^{(n)}[k_r] = a_J[k_r] + \sum_{(j,k_r)\in\mathcal{M}} d_j[k_r] \;, \tag{26}$$

where $\mathcal{M}$ is the multiresolution support defined similarly to (19). The regularized Richardson-Lucy scheme then becomes

$$X^{(n+1)} = \mathbf{P}_+ \left(X^{(n)} \otimes \left(\mathbf{H}^T \left((\mathbf{H}X^{(n)} + \bar{R}^{(n)}) \oslash \mathbf{H}X^{(n)} \right) \right) \right) \;. \tag{27}$$

This algorithm is similar to Murtagh et al. (1995), except that the *à trous* wavelet transform is replaced by the undecimated wavelet transform. Next section shows how the same algorithm can be extended to multi-channel case.

4.3. Multichannel Deconvolution

As the PSF is channel-dependent, the convolution observation model is now

$$Y[\cdot, k_t] = \mathbf{H}_{k_t} X[\cdot, k_t] + \varepsilon[\cdot, k_t]$$

in each channel k_t, where $\mathbf{H}_{k_t}$ is the (spatial) convolution operator in channel k_t with known PSF.

The same recipe as in the monochannel case applies with the notable difference that the spherical 2D-1D MS-VSTS is used instead of its monochannel counterpart. The multichannel multiresolution support $\mathcal{M}$ is obtained after thresholding these coefficients.

Let now $\mathbf{H}$ be the multichannel convolution[1] operator, which acts on a 2D-1D multichannel spherical data set X by applying $\mathbf{H}_t$ on each channel $X[\cdot, k_t]$ independently[2]. The regularized

[1] Strictly speaking, this is a slight abuse of terminology since the kernel is not channel-invariant.

[2] If X were to be vectorized by stacking the channels in a long column vector, $\mathbf{H}$ would be a block-diagonal matrix whose blocks are the circulant matrices $\mathbf{H}_{k_t}$.

multichannel Richardson-Lucy scheme is then

$$X^{(n+1)} = \mathbf{P}_+ \left(X^{(n)} \otimes \left(\mathbf{H}^T \left((\mathbf{H}X^{(n)} + \bar{R}^{(n)}) \oplus \mathbf{H}X^{(n)} \right) \right) \right) , \tag{28}$$

where $\bar{R}^{(n)}$ is the regularized (significant) residual

$$\bar{R}^{(n)}[k_r, k_t] = a_{J_1,J_2}[k_r, k_t] + \sum_{(j_1,j_2,k_r,k_t)\in\mathcal{M}} w_{j_1,j_2}[k_r, k_t] . \tag{29}$$

4.4. Experiments

The algorithm was applied to the 7 energy bands (50 MeV-1.58 GeV) of our simulated Fermi data set. Figures 5 to 8 display the deconvolution results on 4 energy bands. Figure 9 shows the performance of the multichannel MS-VSTS deconvolution algorithm on a single point source. The deconvolution not only removes effectively the blur and recovers sharply localized point sources, but it also allows to restore the whole spectral information. To get a better visual impression of the performance of the deconvolution algorithm, Figure 10 depicts the result on a single HEALPix face covering the galactic plane. The effect of the deconvolution is strikingly good. Our MS-VSTS multichannel deconvolution algorithm manages to remove a large part of the blur introduced by the PSF.

Software

The software related to this paper, **MRS/Poisson**, and its full documentation will be included in the next version of ISAP (Interactive Sparse astronomical data Analysis Packages) via the web site:

`http://jstarck.free.fr/isap.html`

5. Conclusion

This paper extends the MS-VSTS framework to deal with monochannel deconvolution, multichannel denoising and multichannel deconvolution. Unlike the monochannel MS-VSTS, the multichannel MS-VSTS fully exploits the information in the 2D-1D data set and allows to recover the spectral information on the sources. As the PSF is strongly dependent on the energy, it is very important to have a multichannel method for deconvolution. Multichannel deconvolution using MS-VSTS removes a large part of the PSF blur and significantly improves the sharpness of the spatial localization of point sources.

Acknowledgements. Some of the results in this paper have been derived using the Healpix (Górski et al. 2005) and the MR/S software (Starck et al. 2006). This work was supported by the French National Agency for Research (ANR-08-EMER-009-01) and the European Research Council grant SparseAstro (ERC-228261).

References

Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, ApJS, 188, 405
Atwood, W. B., Abdo, A. A., Ackermann, M., et al. 2009, ApJ, 697, 1071
Conrad, J., Scargle, J., & Ylinen, T. 2007, in American Institute of Physics Conference Series, Vol. 921, The First GLAST Symposium, ed. S. Ritz, P. Michelson, & C. A. Meegan, 586–587
Donoho, D. 2004, For Most Marge Underdetermined Systems of Linear Equations, the minimal l^1-norm solution is also the sparsest solution, Tech. rep., Department of Statistics of Stanford University
Faÿ, G., Delabrouille, J., Kerkyacharyan, G., & Picard, D. 2011, ArXiv e-prints
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, Astrophysical Journal, 622, 759–771
Movit, S. 2009, in Bulletin of the American Astronomical Society, Vol. 41, American Astronomical Society Meeting Abstracts, 332.06
Murtagh, F., Starck, J.-L., & Bijaoui, A. 1995, Astronomy and Astrophysics, Supplement Series, 112, 179–189
Norris, J. P., Gehrels, N., & Scargle, J. D. 2011, in Bulletin of the American Astronomical Society, Vol. 43, American Astronomical Society Meeting Abstracts, 108.03
Schmitt, J., Starck, J. L., Casandjian, J. M., Fadili, J., & Grenier, I. 2010, Astronomy and Astrophysics, 517
Schmitt, J., Starck, J.-L., Fadili, J., & Digel, S. 2011, in Advances in Machine Learning and Data Mining for Astronomy, ed. M. Way, J. Scargle, K. Ali, & A. Srivastava (Chapman and Hall)
Starck, J., Fadili, J. M., Digel, S., Zhang, B., & Chiang, J. 2009, A&A, 504, 641
Starck, J.-L., Moudden, Y., Abrial, P., & Nguyen, M. 2006, Astronomy and Astrophysics, 446, 1191

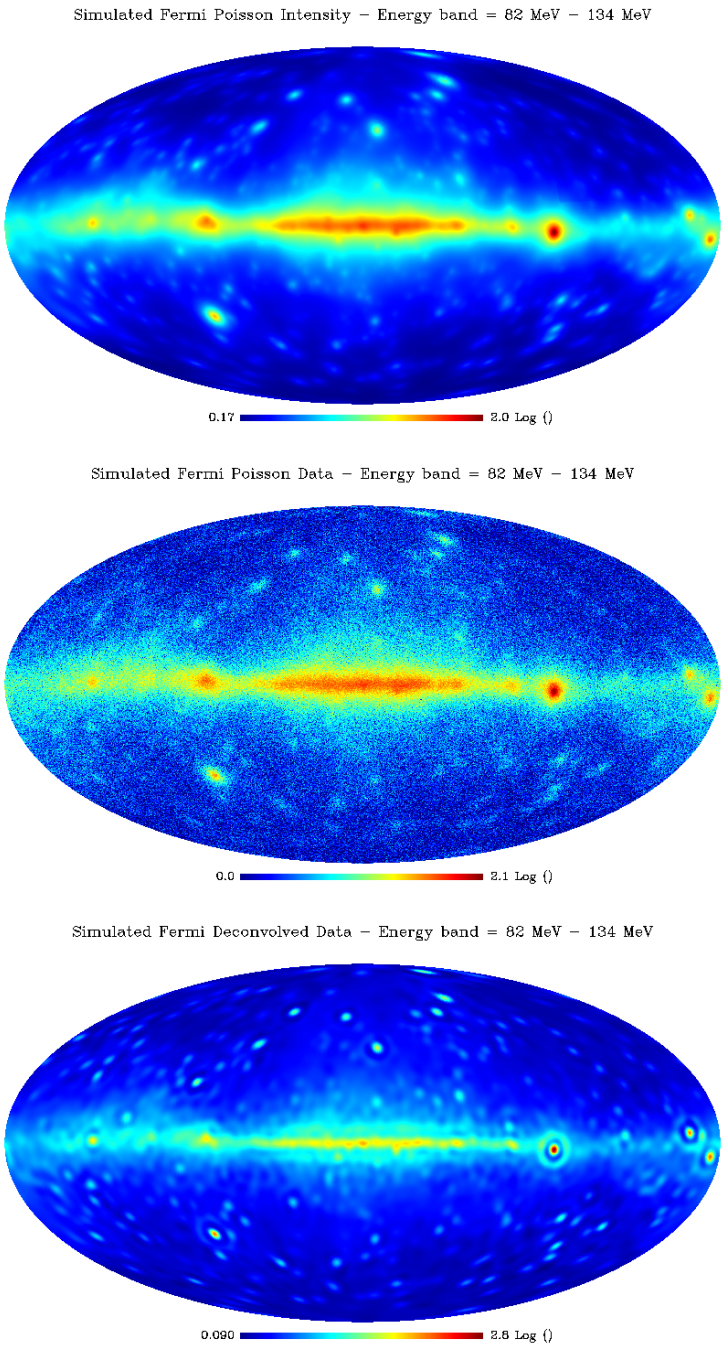

Figure 5. Result of the multichannel deconvolution algorithm on different energy bands. Top: Simulated (blurred) intensity skymap. Middle: Blurred and noisy skymap. Bottom: Deconvolved skymap. Energy band : 82 MeV - 134 MeV. Maps are in logarithmic scale.

Starck, J.-L., Murtagh, F., & Fadili, M. 2010, Sparse Image and Signal Processing (Cambridge University Press)
Yamada, I. 2001, in Inherently Parallel Algorithm in Feasibility and Optimization and their Applications (Elsevier), pp. 473–504
Zhang, B., Fadili, J., & Starck, J.-L. 2008, IEEE Transactions on Image Processing, vol. 11, no. 6, pp. 1093

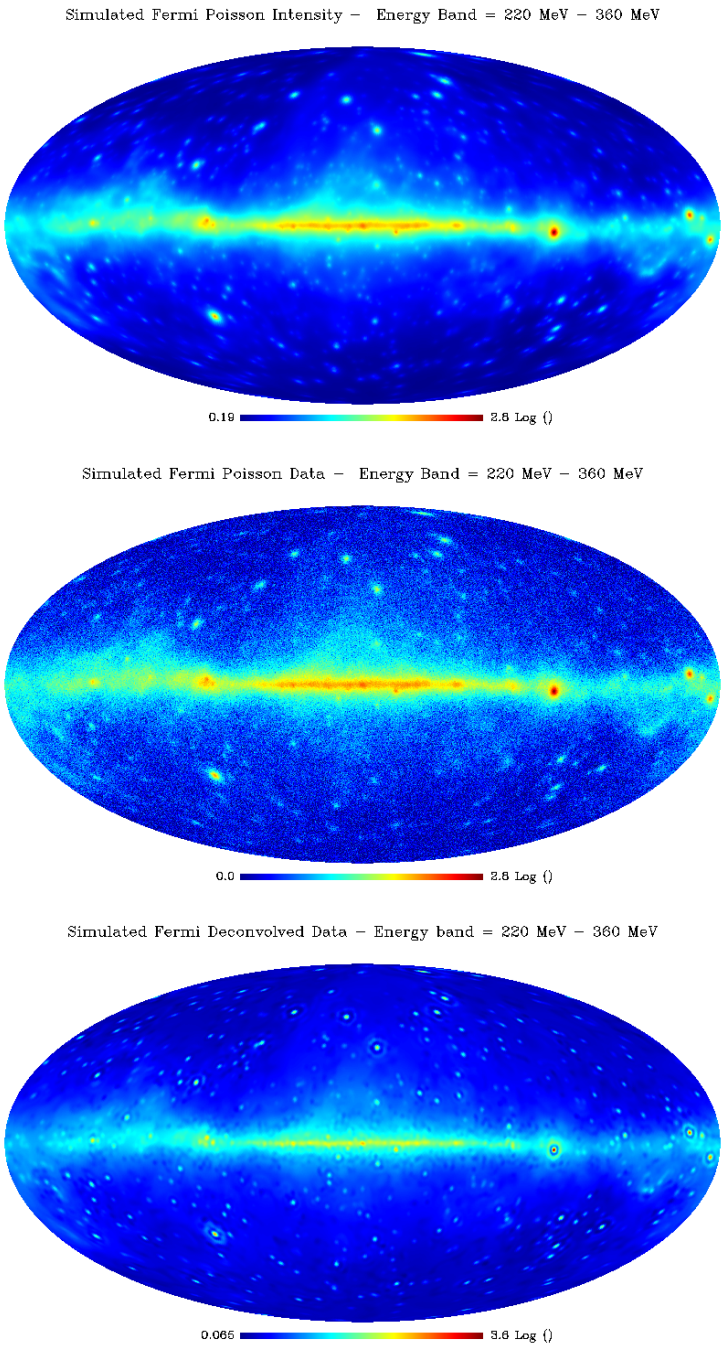

Figure 6. Result of the multichannel deconvolution algorithm on different energy bands. Top: Simulated (blurred) intensity skymap. Middle: Blurred and noisy skymap. Bottom: Deconvolved skymap. Energy band : 220 MeV - 360 MeV. Maps are in logarithmic scale.

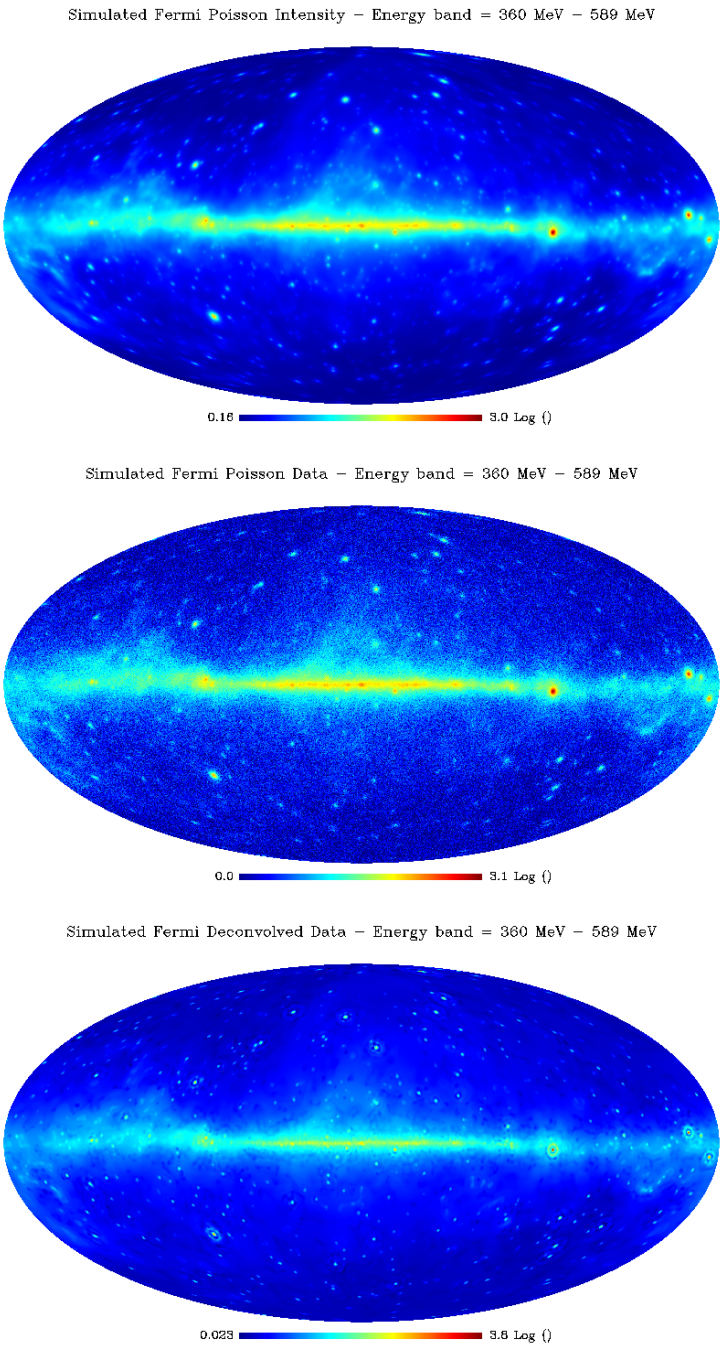

Figure 7. Result of the multichannel deconvolution algorithm on different energy bands. Top: Simulated (blurred) intensity skymap. Middle: Blurred and noisy skymap. Bottom: Deconvolved skymap. Energy band : 360 MeV - 589 MeV. Maps are in logarithmic scale.

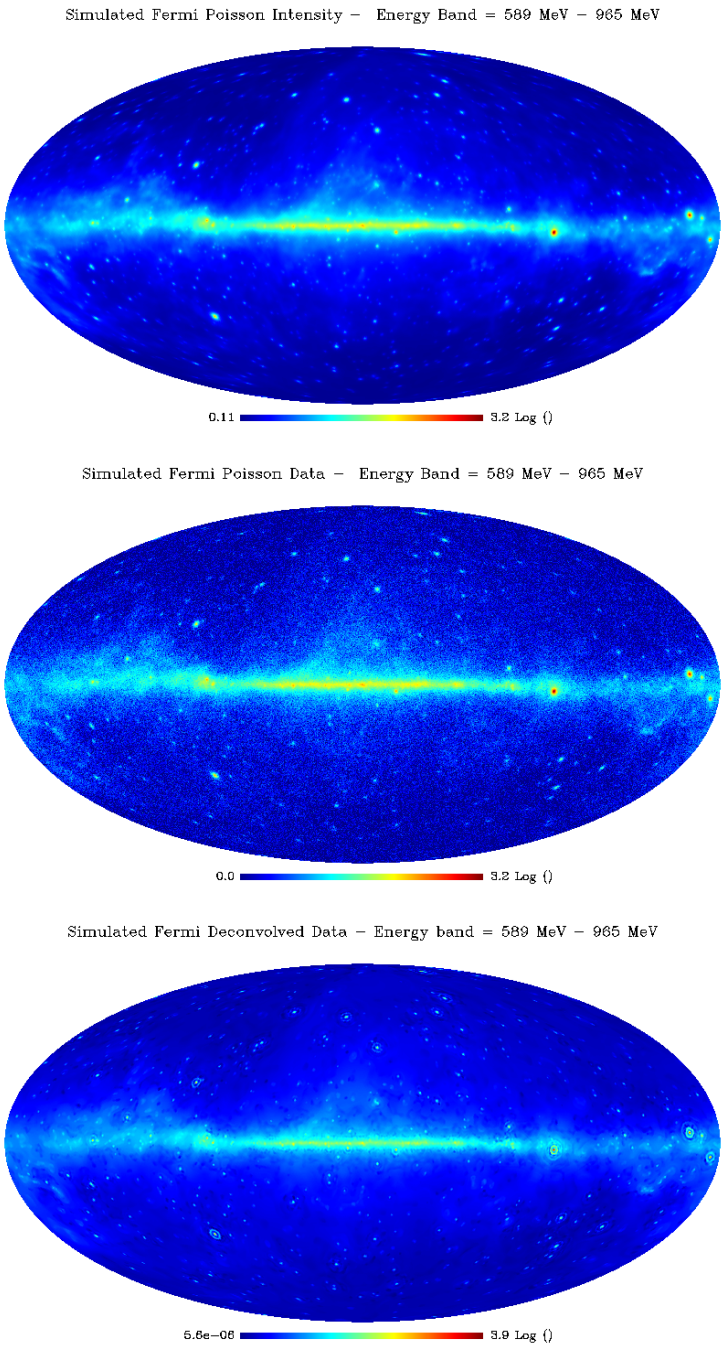

Figure 8. Result of the multichannel deconvolution algorithm on different energy bands. Top: Simulated (blurred) intensity skymap. Middle: Blurred and noisy skymap. Bottom: Deconvolved skymap. Energy band : 589 MeV - 965 MeV. Maps are in logarithmic scale.

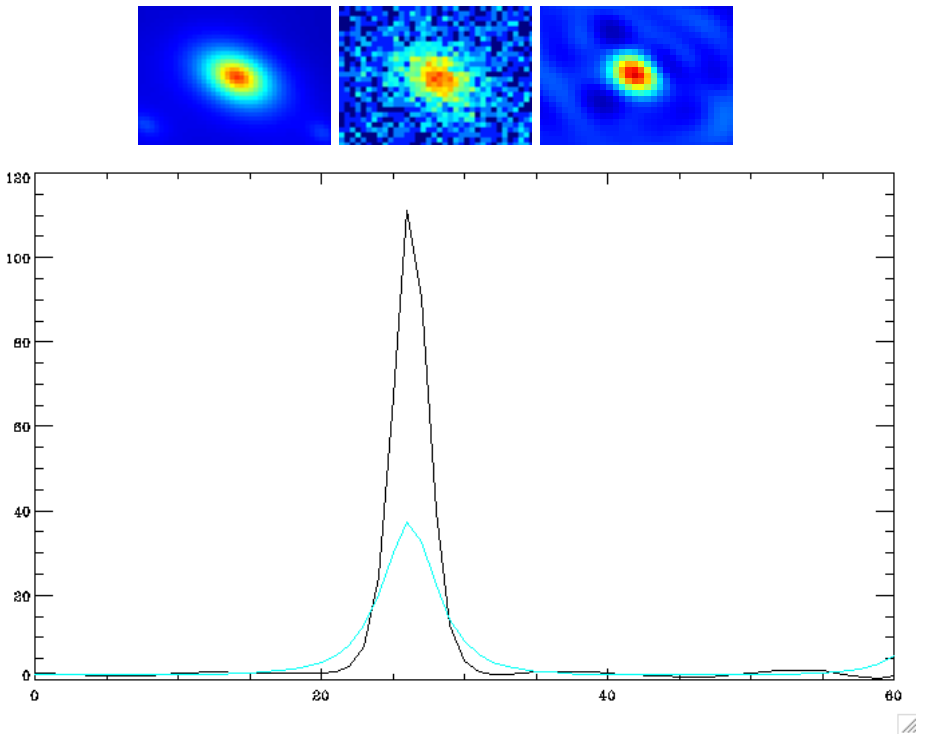

Figure 9. Profile of a single gamma-ray point source recovered using the multichannel MS-VSTS deconvolution algorithm. Top: Single gamma-ray point source on simulated (blurred) Fermi data (energy band: 360 MeV - 589 MeV) (left: simulated blurred source; middle: blurred noisy source; right: deconvolved source). Bottom: Profile of the point source (cyan: simulated spectrum; black: restored spectrum from the deconvolved source.

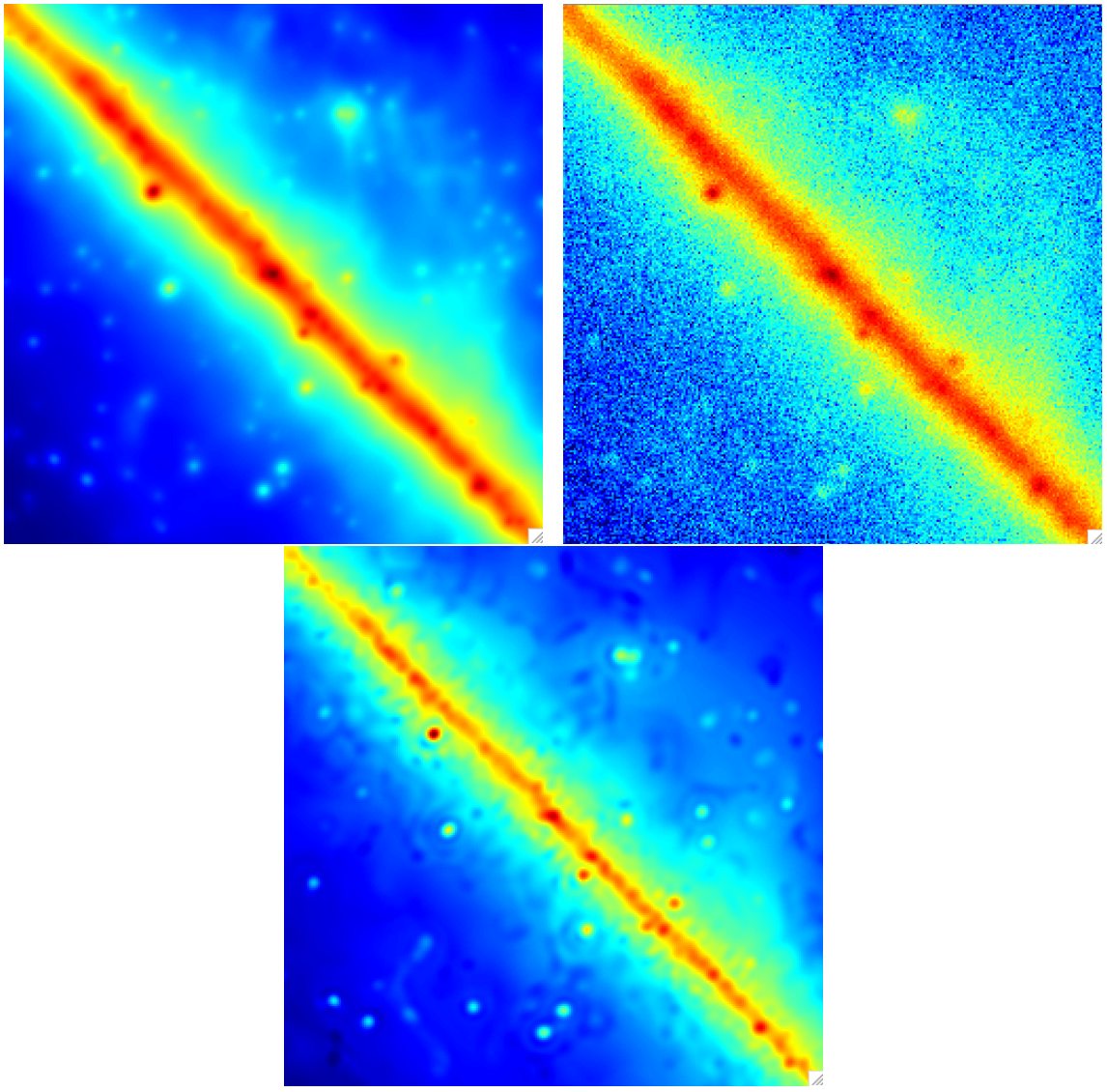

Figure 10. View on a single HEALPix face. Result of the deconvolution algorithm on the galactic plan. Top Left: Simulated Fermi Poisson intensity. Top Right: Simulated Fermi noisy data. Bottom: Fermi data deconvolved with multichannel MS-VSTS. Energy band: 360 MeV - 589 MeV. Pictures are in logarithmic scale.