Source Detection Using a 3D Sparse Representation: Application to the Fermi Gamma-ray Space Telescope

Jean-Luc Starck, Jalal M. Fadili, Seth Digel, Bo Zhang, Jim Chiang

▶ **To cite this version:**

Jean-Luc Starck, Jalal M. Fadili, Seth Digel, Bo Zhang, Jim Chiang. Source Detection Using a 3D Sparse Representation: Application to the Fermi Gamma-ray Space Telescope. Astronomy and Astrophysics - A&A, EDP Sciences, 2009, 504 (2), pp.641-652. <10.1051/0004-6361/200811388>. <hal-00813995>

HAL Id: hal-00813995

https://hal.archives-ouvertes.fr/hal-00813995

Submitted on 16 Apr 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Astronomy & Astrophysics manuscript no. aa11388-08
August 25, 2009
© ESO 2009

Source detection using a 3D sparse representation: application to the Fermi gamma-ray space telescope

J.-L. Starck[1], J. M. Fadili[2], S. Digel[3], B. Zhang[4], and J. Chiang[3]

[1] CEA, IRFU, SEDI-SAP, Laboratoire Astrophysique des Interactions Multi-échelles (UMR 7158), CEA/DSM-CNRS-Universite Paris Diderot, Centre de Saclay, 91191 Gif-Sur-Yvette, France
e-mail: jstarck@cea.fr
[2] GREYC CNRS UMR 6072, Image Processing Group, ENSICAEN 14050, Caen Cedex, France
[3] Stanford Linear Accelerator Center & Kavli Institute for Particle Astrophysics and Cosmology, Stanford, CA 94075, USA
[4] Quantitative Image Analysis Unit URA CNRS 2582, Institut Pasteur, 25–28, Rue du Docteur Roux, 75724 Paris Cedex 15, France

Received 20 November 2008 / Accepted 25 February 2009

ABSTRACT

The multiscale variance stabilization Transform (MSVST) has recently been proposed for Poisson data denoising (Zhang et al. 2008a). This procedure, which is nonparametric, is based on thresholding wavelet coefficients. The restoration algorithm applied after thresholding provides good conservation of source flux. We present in this paper an extension of the MSVST to 3D data—in fact 2D-1D data— when the third dimension is not a spatial dimension, but the wavelength, the energy, or the time. We show that the MSVST can be used for detecting and characterizing astrophysical sources of high-energy gamma rays, using realistic simulated observations with the Large Area Telescope (LAT). The LAT was launched in June 2008 on the Fermi Gamma-ray Space Telescope mission. Source detection in the LAT data is complicated by the low fluxes of point sources relative to the diffuse celestial foreground, the limited angular resolution, and the tremendous variation in that resolution with energy (from tens of degrees at $\sim$30 MeV to $\sim$0.1° at 10 GeV). The high-energy gamma-ray sky is also quite dynamic, with a large population of sources such active galaxies with accretion-powered black holes producing high-energy jets, episodically flaring. The fluxes of these sources can change by an order of magnitude or more on time scales of hours. Perhaps the majority of blazars will have average fluxes that are too low to be detected but could be found during the hours or days that they are flaring. The MSVST algorithm is very fast relative to traditional likelihood model fitting, and permits efficient detection across the time dimension and immediate estimation of spectral properties. Astrophysical sources of gamma rays, especially active galaxies, are typically quite variable, and our current work may lead to a reliable method to quickly characterize the flaring properties of newly-detected sources.

Key words. methods: data analysis – techniques: image processing

1. Introduction

The high-energy gamma-ray sky will be studied with unprecedented sensitivity by the Large Area Telescope (LAT), which was launched by NASA on the *Fermi* mission in June 2008. The catalog of gamma-ray sources from the previous mission in this energy range, EGRET on the Compton Gamma-Ray Observatory, has approximately 270 sources (Hartman et al. 1999). For the LAT, several thousand gamma-ray sources are expected to be detected, with much more accurately determined locations, spectra, and light curves.

We would like to reliably detect as many celestial sources of gamma rays as possible. The question is not simply one of building up adequate statistics by increasing exposure times. The majority of the sources that the LAT will detect are likely to be gamma-ray blazars (distant galaxies whose gamma-ray emission is powered by accretion onto supermassive black holes), which are intrinsically variable. They flare episodically in gamma rays. The time scales of flares, which can increase the flux by a factor of 10 or more, can be minutes to weeks. The duty cycle of flaring in gamma rays is not well determined yet, but individual blazars can go months or years between flares and in general we will not know in advance where on the sky the sources will be found.

The fluxes of celestial gamma rays are low, especially relative to the $\sim$1 m^2 effective area of the LAT (by far the largest effective collecting area ever in the GeV range). An additional complicating factor is that diffuse emission from the Milky Way itself (which originates in cosmic-ray interactions with interstellar gas and radiation) makes a relatively intense, structured foreground emission. The few very brightest gamma-ray sources will provide approximately 1 detected gamma ray per minute when they are in the field of view of the LAT. The diffuse emission of the Milky Way will provide about 2 gamma rays per second, distributed over the $\sim$2 sr field of view.

For previous high-energy gamma-ray missions, the standard method of source detection has been model fitting – maximizing the likelihood function while moving trial point sources around in the region of the sky being analyzed. This approach has been driven by the limited photon counts and the relatively limited resolution of gamma-ray telescopes. However, at the sensitivity of the LAT, even a relatively "quiet" part of the sky may have 10 or more point sources close enough together to need to be modeled simultaneously when maximizing the (computationally expensive) likelihood function. For this reason and because of the need to search in time, non-parametric algorithms for detecting sources are being investigated.

Literature overview for Poisson denoising using wavelets

A host of estimation methods have been proposed in the literature for non-parametric Poisson noise removal. Major contributions consist of variance stabilization: a classical solution is to preprocess the data by applying a variance stabilizing transform (VST) such as the Anscombe transform (Anscombe 1948; Donoho 1993). It can be shown that the transformed data are approximately stationary, independent, and Gaussian. However, these transformations are only valid for a sufficiently large number of counts per pixel (and of course, for even more counts, the Poisson distribution becomes Gaussian with equal mean and variance) (Murtagh et al. 1995). The necessary average number of counts is about 20 if bias is to be avoided.

In this case, as an alternative approach, a filtering approach for very small numbers of counts, including frequent zero cases, has been proposed in (Starck & Pierre 1998), which is based on the popular isotropic undecimated wavelet transform (implemented with the so-called à trous algorithm) (Starck & Murtagh 2006) and the autoconvolution histogram technique for deriving the probability density function (pdf) of the wavelet coefficient (Slezak et al. 1993; Bijaoui & Jammal 2001; Starck & Murtagh 2006). This method is part of the data reduction pipeline of the XMM-LSS project (Pierre et al. 2004) for detecting of clusters of galaxies (Pierre et al. 2007). This algorithm is obviously a good candidate for *Fermi* LAT 2D map analysis, but its extension to 2D-1D data sets does not exist. It is far from being trivial, and even if it were possible, computation time would certainly be prohibitive to allow its use for *Fermi* LAT 2D-1D data sets. Then, an alternative approach is needed. Several authors (Kolaczyk 1997; Timmermann & Nowak 1999; Nowak & Baraniuk 1999; Bijaoui & Jammal 2001; Fryźlewicz & Nason 2004; Zhang et al. 2008b) have suggested that the Haar wavelet transform is very well-suited for treating data with Poisson noise. Since a Haar wavelet coefficient is just the difference between two random variables following a Poisson distribution, it is easier to derive mathematical tools for removing the noise than with any other wavelet method. Starck & Murtagh (2006) study shows that the Haar transform is less effective for restoring X-ray astronomical images than the à trous algorithm. The reason is that the wavelet shape of the isotropic wavelet transform is much better adapted to astronomical sources, which are more or less Gaussian-shaped and isotropic, than the Haar wavelet. Some papers (Scargle 1998; Kolaczyk & Nowak 2004; Willet & Nowak 2005; Willett 2006) proposed a spatial partitioning, possibly dyadic, of the image for complicated geometrical content recovery. This dyadic partitioning concept is however again not very well suited to astrophysical data.

The MSVST alternative

In a recent paper, Zhang et al. (2008a) have proposed to merge a variance stabilization technique and the multiscale decomposition, leading to the Multi-Scale Variance Stabilization Transform (MSVST). In the case of the isotropic undecimated wavelet transform, as the wavelet coefficients w_j are derived by a simple difference of two consecutive dyadic scales of the input image (see Sect. 3.2), $w_j = a_{j-1} - a_j$, the stabilized wavelet coefficients are obtained by applying a stabilization on both a_{j-1} and a_j, $w_j = \mathcal{A}_{j-1}(a_{j-1}) - \mathcal{A}_j(a_j)$, where $\mathcal{A}_{j-1}$ and $\mathcal{A}_j$ are non-linear transforms that can be seen as a generalization of the Anscombe transform; see Sect. 3 for details. This new method is fast and easy to implement, and more importantly, works very well at very low count situations, down to 0.1 photons per pixel.

This paper

In this paper, we present a new multiscale representation, derived from the MSVST, which allows us to remove the Poisson noise in 3D data sets, when the third dimension is not a spatial dimension, but the wavelength, the energy or the time. Such 3D data are called 2D-1D data sets in the sequel. We show that it could be very useful to analyze *Fermi* LAT data, especially when looking for rapidly time varying sources. Section 2 describes the *Fermi* LAT simulated data. Section 3 reviews the MSVST method relative to the isotropic undecimated wavelet transform and Sect. 4 shows how it can be extended to the 2D-1D case. Section 5 presents some experiments on simulated *Fermi* LAT data. Conclusions are given in Sect. 6.

Definitions and 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. For the sake of clarity, the notation $h[i]$ is used instead of h_i for the location index. This will lighten the notation by avoiding multiple subscripts in the derivations of the paper. The discrete circular convolution product of two signals will be written $\star$, and the continuous convolution of two functions $*$. The term circular stands for periodic boundary conditions. The symbol $\delta[i]$ is the Kronecker delta.

For the octave band wavelet representation, analysis (respectively, synthesis) filters are denoted h and g (respectively, $\tilde{h}$ and $\tilde{g}$). 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, $\tilde{\phi}$ and $\tilde{\psi}$). We also define the scaled dilated and translated version of ϕ at scale j and position k as $\phi_{j,k}(x) = 2^{-j}\phi(2^{-j}x - k)$, and similarly for ψ, $\tilde{\phi}$ and $\tilde{\psi}$. A function $f(x, y)$ is isotropic if it is constant along all points (x, y) that are equidistant from the origin.

A distribution is stabilized if its variance is made constant, typically equal to 1, independently of its mean. A transformation applied to a random variable is called a variance stabilizing transform (VST), if the distribution of the transformed variable is stabilized and is approximately Gaussian.

Glossary

WT	WaveletTransform
DWT	Discrete(decimated)WaveletTransform
UWT	UndecimatedWaveletTransform
IUWT	IsotropicUndecimatedWaveletTransform
VST	VarianceStabilizationTransform
MSVST	Multi − ScaleVarianceStabilizationTransform
LAT	LargeAreaTelescope(LAT)
FDR	FalseDiscoveryRate

2. Data description

2.1. Fermi Large area telescope

The LAT (Fig. 1) is a photon-counting detector, converting gamma rays into positron-electron pairs for detection. The trajectories of the pair are tracked and their energies measured in order to reconstruct the direction and energy of the gamma ray.

The energy range of the LAT is very broad, approximately 20 MeV−300 GeV. At energies below a few hundred MeV, the

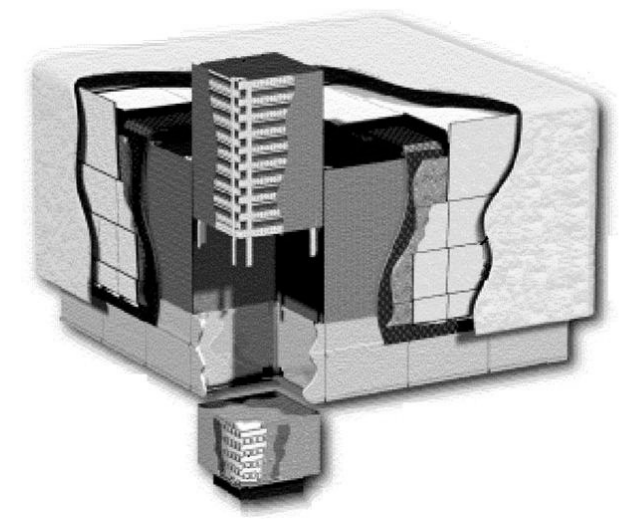

Fig. 1. Cutaway view of the LAT. The LAT is modular; one of the 16 towers is shown with its tracking planes revealed. High-energy gamma rays convert to electron-positron pairs on tungsten foils in the tracking layers. The trajectories of the pair are measured very precisely using silicon strip detectors in the tracking layers and the energies are determined with the CsI calorimeter at the bottom. The array of plastic scintillators that cover the towers provides an anticoincidence signal for cosmic rays. The outermost layers are a thermal blanket and micrometeoroid shield. The overall dimensions are $1.8 \times 1.8 \times 0.75$ m.

reconstruction and tracking efficiencies are lower, and the angular resolution is poorer, than at higher energies. The point spread function (PSF) width varies from about 3.5° at 100 MeV to better than 0.1° (68% containment) at 10 GeV and above. 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.

Wavelet denoising of LAT data has application as part of an algorithm for quickly detecting celestial sources of gamma rays. The fundamental inputs to high-level analysis of LAT data will be energies, directions, and times of the detected gamma rays. (Pointing history and instrument live times are also inputs for exposure calculations.) For the analysis presented here, we consider the LAT data for some range of time to have been binned into "cubes" $v(x, y, t)$ of spatial coordinates and time or, $v(x, y, E)$ of spatial coordinates and energy, because, as we shall see, the wavelet denoising can be applied in multiple dimensions, and so permits estimation of counts spectra. The motivations for filtering data with Poisson noise in the wavelet domain are well known—sources of small angular size are localized in wavelet space.

2.2. Simulated LAT data

The application of MSVST to problems of detection and characterization of LAT sources was investigated using simulated data. The simulations included a realistic observing strategy (sky survey with the proper orbital and rocking periods) and response functions for the LAT (effective area and angular resolution as functions of energy and angle). Point sources of gamma rays were defined with systematically varying fluxes, spectral slopes, and/or flare intensities and durations. The simulations also included a representative level of diffuse "background" (celestial plus residual charged-particle) for regions of the sky well removed from the Galactic equator, where the celestial diffuse emission is particularly intense. The denoising results reported in Sect. 5 use a data cube obtained according to this simulation scenario.

3. The 2D multiscale variance stabilization transform (MSVST)

In this section, we review the MSVST method (Zhang et al. 2008a), restricted to the Isotropic Undecimated Wavelet Transform (IUWT). Indeed, the MSVST can use other transforms such as the standard three-orientation undecimated wavelet transform, the ridgelet or the curvelet transforms; see (Zhang et al. 2008a). In our specific case here, only the IUWT is of interest.

3.1. VST of a filtered Poisson process

Given $\mathbf{X}$ a sequence of n independent Poisson random variables $X_i, i = 1, \cdots, n$, each of mean λ_i, let $Y_i = \sum_{j=1}^{n} h[j]X_{i-j}$ be the filtered process obtained by convolving the sequence $\mathbf{X}$ with a discrete filter h. Y denotes any one of the Y_i's, and $\tau_k = \sum_i (h[i])^k$ for $k = 1, 2, \cdots$.

If $h = \delta$, then we recover the Anscombe VST (Anscombe 1948) of Y_i (hence X_i) which acts as if the stabilized data arose from a Gaussian white noise with unit variance, under the assumption that the intensity λ_i is large. This is why the Anscombe VST performs poorly in low-count settings. But, if the filter h acts as an "averaging" kernel (more generally a low-pass filter), one can reasonably expect that stabilizing Y_i would be more beneficial, since the signal-to-noise ratio measured at the output of h is expected to be higher.

Using a local homogeneity assumption, i.e. $\lambda_{i-j} = \lambda$ for all j within the support of h, it has been shown (Zhang et al. 2008a) that for a non-negative filter h, the transform $Z = b\sqrt{Y + c}$ with $b > 0$ and $c > 0$ defined as

$$c = \frac{7\tau_2}{8\tau_1} - \frac{\tau_3}{2\tau_2}, \quad b = 2\sqrt{\frac{\tau_1}{\tau_2}} \tag{1}$$

is a second order accurate variance stabilization transform, with asymptotic unit variance. By second-order accurate, we mean that the error term in the variance of the stabilized variable Z decreases rapidly as $O(\lambda^{-2})$. From (1), it is obvious that when $h = \delta$, we obtain the classical Anscombe VST parameters $b = 2$ and $c = 3/8$. The authors in (Zhang et al. 2008a) have also proved that Z is asymptotically distributed as a Gaussian variate with mean $b\sqrt{\tau_1\lambda}$ and unit variance. A non-positive h with a negative c could also be considered; see (Zhang et al. 2008a) for more details.

Figure 2 shows the Monte-Carlo estimates of the expectation $\mathbb{E}[Z]$ (left) and the variance $\mathrm{Var}[Z]$ (right) obtained from 2×10^5 Poisson noise realizations of $\mathbf{X}$, plotted as a function of the intensity λ for both Anscombe (Anscombe 1948) (dashed-dotted), Haar-Fisz (dashed) (Fryźlewicz & Nason 2004) and our VST with the 2D B_3-Spline filter as a low-pass filter h (solid). The asymptotic bounds (dots) (i.e. 1 for the variance and $\sqrt{\lambda}$ for the expectation) are also shown. It can be seen that for increasing intensity, $\mathbb{E}[Z]$ and $\mathrm{Var}[Z]$ approach the theoretical bounds at different rates depending on the VST used. Quantitatively, Poisson variables transformed using the Anscombe VST can be reasonably considered to be unbiased and stabilized for $\lambda \gtrapprox 10$, using Haar-Fisz for $\lambda \gtrapprox 1$, and using out VST (after low-pass filtering with the chosen h) for $\lambda \gtrapprox 0.1$.

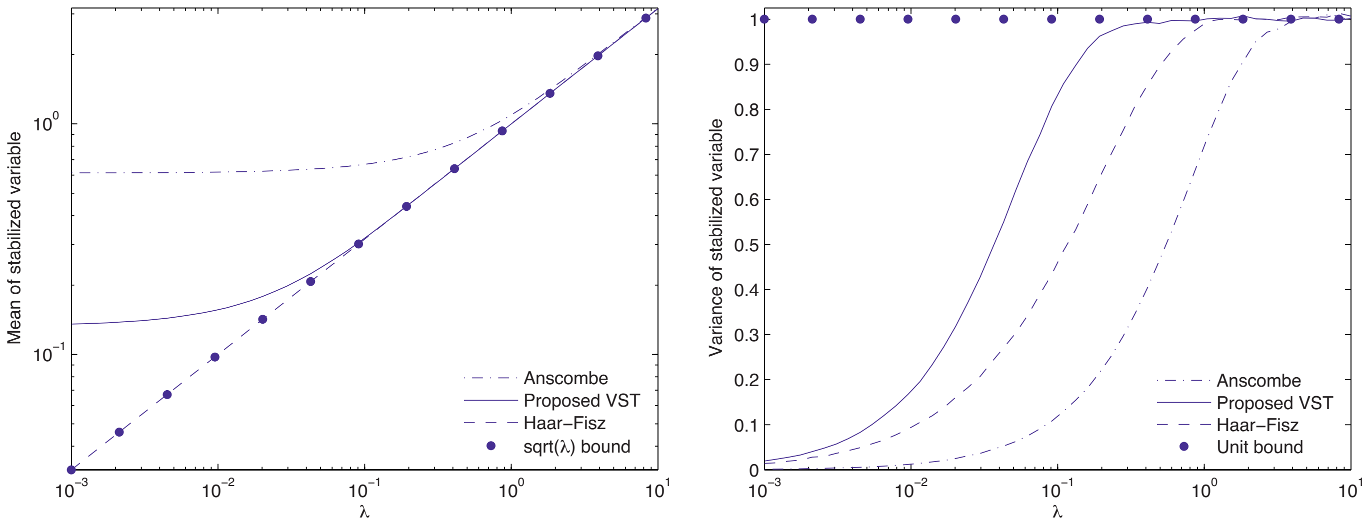

Fig. 2. Behavior of the expectation $\mathbb{E}[Z]$ (*left*) and variance $\mathrm{Var}[Z]$ (*right*) as a function of the underlying intensity, for the Anscombe VST, 2D Haar-Fisz VST, and out VST with the 2D B_3-Spline filter as a low-pass filter h.

3.2. The isotropic undecimated wavelet transform

The undecimated wavelet transform (UWT) uses an analysis filter bank (h, g) to decompose a signal a_0 into a coefficient set $W = \{d_1, \ldots, d_J, a_J\}$, where d_j is the wavelet (detail) coefficients at scale j and a_J is the approximation coefficients at the coarsest resolution J. The passage from one resolution to the next one is obtained using the "à trous" algorithm (Holschneider et al. 1989; Shensa 1992)

$$a_{j+1}[l] = (\bar{h}^{\uparrow j} \star a_j)[l] = \sum_k h[k] a_j[l + 2^j k], \tag{2}$$

$$w_{j+1}[l] = (\bar{g}^{\uparrow j} \star a_j)[l] = \sum_k g[k] a_j[l + 2^j k], \tag{3}$$

where $h^{\uparrow j}[l] = h[l]$ if $l/2^j \in \mathbb{Z}$ and 0 otherwise, $\bar{h}[l] = h[-l]$, and "$\star$" denotes discrete circular convolution. The reconstruction is given by $a_j[l] = \frac{1}{2}\left[(\tilde{\bar{h}}^{\uparrow j} \star a_{j+1})[l] + (\tilde{\bar{g}}^{\uparrow j} \star w_{j+1})[l]\right]$. The filter bank $(h, g, \tilde{h}, \tilde{g})$ needs to satisfy the so-called exact reconstruction condition (Mallat 1998; Starck & Murtagh 2006).

The Isotropic UWT (IUWT) (Starck et al. 2007) uses the filter bank $(h, g = \delta - h, \tilde{h} = \delta, \tilde{g} = \delta)$ where h is typically a symmetric low-pass filter such as the B_3-Spline filter. The reconstruction is trivial, i.e., $a_0 = a_J + \sum_{j=1}^{J} w_j$. This algorithm is widely used in astronomical applications (Starck et al. 1998) and biomedical imaging (Olivo-Marin 2002) to detect isotropic objects.

The IUWT filter bank in q-dimension $(q \geq 2)$ becomes $(h_{q\mathrm{D}}, g_{q\mathrm{D}} = \delta - h_{q\mathrm{D}}, \tilde{h}_{q\mathrm{D}} = \delta, \tilde{g}_{q\mathrm{D}} = \delta)$ where $h_{q\mathrm{D}}$ is the tensor product of q 1D filters $h_{1\mathrm{D}}$. Note that $g_{q\mathrm{D}}$ is in general non-separable.

3.3. MSVST with the IUWT

Now the VST can be combined with the IUWT in the following way: since the filters $\bar{h}^{\uparrow j}$ at all scales j are low-pass filters (so have nonzero means), we can first stabilize the approximation coefficients a_j at each scale using the VST, and then compute in the standard way the detail coefficients from the stabilized a_j's. Given the particular structure of the IUWT analysis filters (h, g), the stabilization procedure is given by

$$\text{IUWT} \quad \begin{cases} a_j = \bar{h}^{\uparrow j-1} \star a_{j-1} \\ w_j = a_{j-1} - a_j \end{cases}$$
$$\Longrightarrow \quad \begin{matrix}\text{MSVST}\\ + \\ \text{IUWT}\end{matrix} \quad \begin{cases} a_j = \bar{h}^{\uparrow j-1} \star a_{j-1} \\ w_j = \mathcal{A}_{j-1}(a_{j-1}) - \mathcal{A}_j(a_j). \end{cases} \tag{4}$$

Note that the VST is now scale-dependent (hence the name MSVST). The filtering step on a_{j-1} can be rewritten as a filtering on $a_0 = \mathbf{X}$, i.e., $a_j = h^{(j)} \star a_0$, where $h^{(j)} = \bar{h}^{j-1} \star \cdots \star \bar{h}^1 \star \bar{h}$ for $j \geq 1$ and $h^{(0)} = \delta$. $\mathcal{A}_j$ is the VST operator at scale j

$$\mathcal{A}_j(a_j) = b^{(j)} \sqrt{a_j + c^{(j)}}. \tag{5}$$

Let us define $\tau_k^{(j)} = \sum_i \left(h^{(j)}[i]\right)^k$. Then according to (1), the constants $b^{(j)}$ and $c^{(j)}$ associated to $h^{(j)}$ must be set to

$$c^{(j)} = \frac{7\tau_2^{(j)}}{8\tau_1^{(j)}} - \frac{\tau_3^{(j)}}{2\tau_2^{(j)}}, \qquad b^{(j)} = 2\sqrt{\frac{\tau_1^{(j)}}{\tau_2^{(j)}}}. \tag{6}$$

The constants $b^{(j)}$ and $c^{(j)}$ only depend on the filter h and the scale level j. They can all be pre-computed once for any given h. A schematic overview of the decomposition and the inversion of MSVST+IUWT is depicted in Fig. 3.

In summary, IUWT denoising with the MSVST involves the following three main steps:

1. **Transformation**: compute the IUWT in conjunction with the MSVST as described above.
2. **Detection**: detect significant detail coefficients by hypothesis testing. The appeal of a binary hypothesis testing approach is that it allows quantitative control of significance. Here, we take benefit from the asymptotic Gaussianity of the stabilized a_j's that will be transferred to the w_j's as it has been shown by (Zhang et al. 2008a). Indeed, these authors have proved that under the null hypothesis H_0:$w_j[k] = 0$ corresponding to the fact that the signal is homogeneous (smooth), the stabilized detail coefficients w_j follow asymptotically a centered normal distribution with an intensity-independent variance; see (Zhang et al. 2008a, Theorem 1) for details.

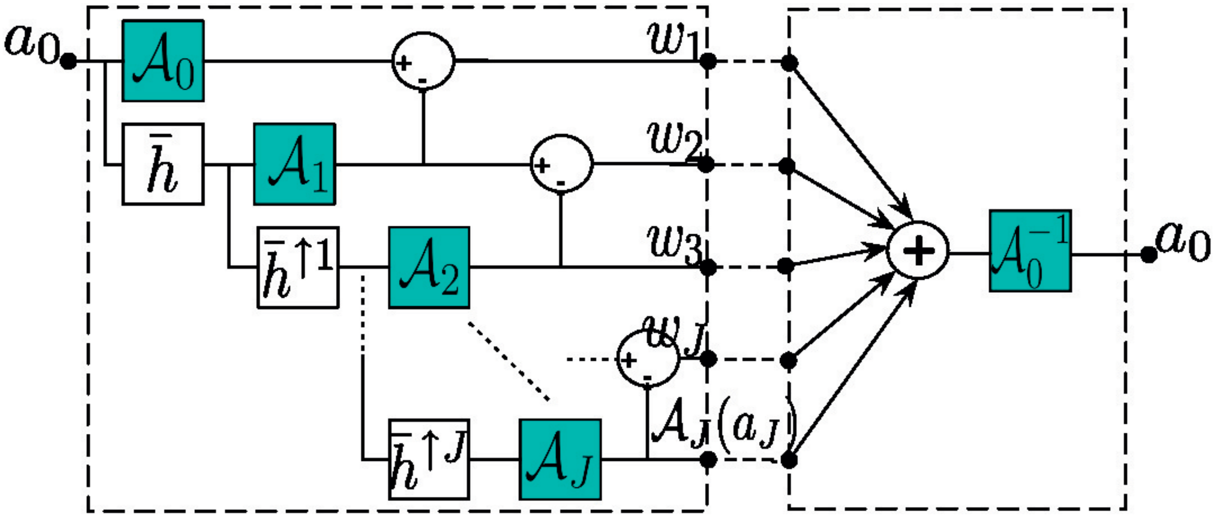

Fig. 3. Diagrams of the MSVST combined with the IUWT. The notations are the same as those of (4) and (7). The left dashed frame shows the decomposition part. Each stage of this frame corresponds to a scale j and an application of (4). The right dashed frame illustrates the direct inversion (7).

This variance depends only on the filter h and the current scale, and can be tabulated once for any h. Thus, the distribution of the w_j's being known (Gaussian), we can detect the significant coefficients by classical binary hypothesis testing.

3. **Estimation**: reconstruct the final estimate using the knowledge of the detected coefficients. This step requires inverting the MSVST after the detection step. For the IUWT filter bank, there is a closed-form inversion expression as we have

$$a_0 = \mathcal{A}_0^{-1}\left[\mathcal{A}_J(a_J) + \sum_{j=1}^{J} w_j\right]. \tag{7}$$

3.3.1. Example

Figure 4 upper left shows a set of objects of different sizes and different intensities contaminated by a Poisson noise. Each object along any radial branch has the same integrated intensity within its support and has a more and more extended support as we go farther from the center. The integrated intensity reduces as the branches turn in the clockwise direction. Denoising such an image is challenging. Figure 4, top-right, bottom-left and right, show respectively the filtered images by Haar-Kolaczyk (Kolaczyk 1997), Haar-Jammal-Bijaoui (Bijaoui & Jammal 2001) and the MSVST.

As expected, the relative merits (sensitivity) of the MSVST estimator become increasingly salient as we go farther from the center, and as the branches turn clockwise. That is, the MSVST estimator outperforms its competitors as the intensity becomes low. Most sources were detected by the MSVST estimator even for very low counts situations; see the last branches clockwise in Fig. 4 bottom right and compare to Fig. 4 top right and Fig. 4 bottom left.

4. 2D-1D MSVST denoising

4.1. 2D-1D wavelet transform

In the previous section, we have seen how a Poisson noise can be removed from 2D image using the IUWT and the MSVST. Extension to a qD data sets is straightforward, and the denoising will be nearly optimal as long as each object belonging to this q-dimensional space is roughly isotropic. In the case of 3D data where the third dimension is either the time or the energy, we are clearly not in this configuration, and the naive analysis of a 3D isotropic wavelet does not make sense. Therefore, we want to analyze the data with a non-isotropic wavelet, where the time or energy scale is not connected to the spatial scale. Hence, an ideal wavelet function would be defined by:

$$\psi(x, y, z) = \psi^{(xy)}(x, y)\psi^{(z)}(z), \tag{8}$$

where $\psi^{(xy)}$ is the spatial wavelet and $\psi^{(z)}$ is the temporal (or energy) wavelet. In the following, we will consider only isotropic and dyadic spatial scales, and we note j_1 the spatial resolution index (i.e. scale $= 2^{j_1}$), j_2 the time (or energy) resolution index. Thus, define the scaled spatial and temporal (or energy) wavelets

$$\psi^{(xy)}_{j_1}(x, y) = \frac{1}{2^{j_1}}\psi^{(xy)}\left(\frac{x}{2^{j_1}}, \frac{y}{2^{j_1}}\right) \text{ and}$$

$$\psi^{(z)}_{j_2}(z) = \frac{1}{\sqrt{2^{j_2}}}\psi^{(z)}\left(\frac{z}{2^{j_2}}\right).$$

Hence, we derive the wavelet coefficients $w_{j_1,j_2}[k_x, k_y, k_z]$ from a given data set D (k_x and k_y are spatial index and k_z a time (or energy) index). In continuous coordinates, this amounts to the formula

$$\begin{aligned} w_{j_1,j_2}[k_x, k_y, k_z] &= \frac{1}{2^{j_1}}\frac{1}{\sqrt{2^{j_2}}}\iiint_{-\infty}^{+\infty} D(x, y, z) \\ &\quad \times \psi^{(xy)}\left(\frac{x-k_x}{2^{j_1}}, \frac{y-k_y}{2^{j_1}}\right)\psi^{(z)}\left(\frac{z-k_z}{2^{j_2}}\right)\mathrm{d}x\mathrm{d}y\mathrm{d}z \\ &= D * \bar{\psi}^{(xy)}_{j_1} * \bar{\psi}^{(z)}_{j_2}(x, y, z), \end{aligned} \tag{9}$$

where $*$ is the convolution and $\bar{\psi}(x) = \psi(-x)$.

Fast undecimated 2D-1D decomposition/reconstruction

In order to have a fast algorithm for discrete data, we use wavelet functions associated to filter banks. Hence, our wavelet decomposition consists in applying first a 2D IUWT for each frame k_z. Using the 2D IUWT, we have the reconstruction formula:

$$D[k_x, k_y, k_z] = a_{J_1}[k_x, k_y] + \sum_{j_1=1}^{J_1} w_{j_1}[k_x, k_y, k_z],\ \forall k_z, \tag{10}$$

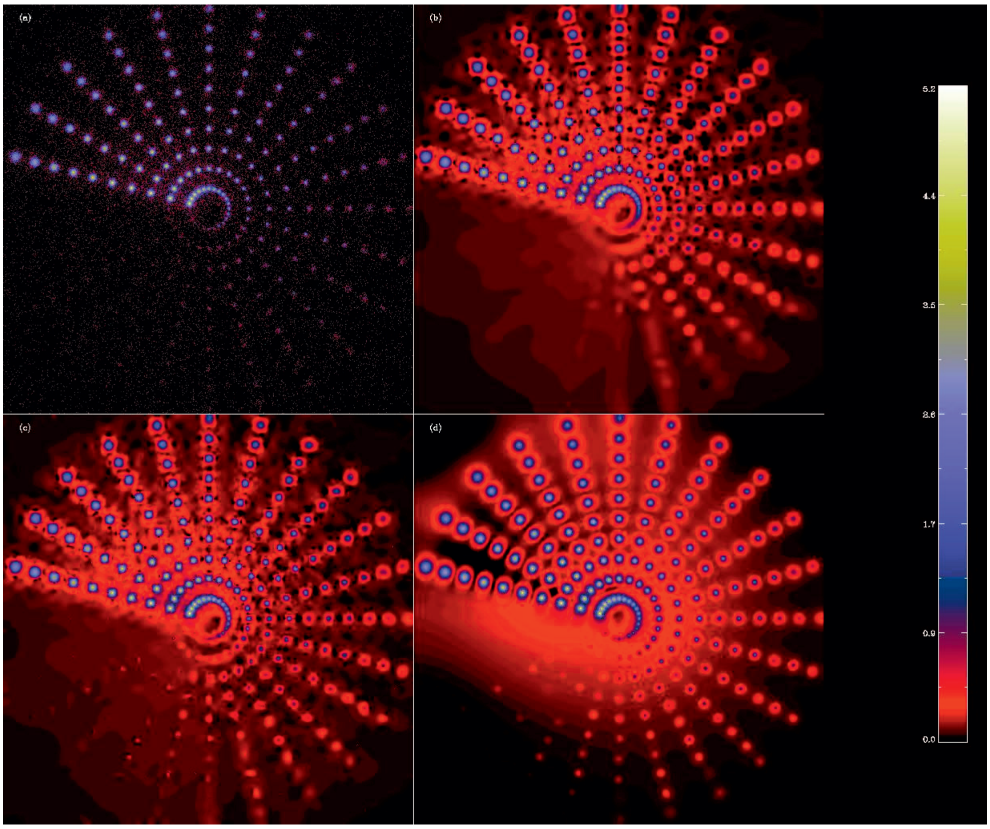

Fig. 4. *Top*, XMM simulated data, and Haar-Kolaczyk (Kolaczyk 1997) filtered image. *Bottom*, Haar-Jammal-Bijaoui (Bijaoui & Jammal 2001) and MSVST filtered images. Intensities logarithmically transformed.

where J_1 is the number of spatial scales. Then, for each spatial location (k_x, k_y) and for each 2D wavelet scale scale j_1, we apply a 1D wavelet transform along z on the spatial wavelet coefficients $w_{j_1}[k_x, k_y, k_z]$ such that

$$w_{j_1}[k_x, k_y, k_z] = w_{j_1,J_2}[k_x, k_y, k_z] + \sum_{j_2=1}^{J_2} w_{j_1,j_2}[k_x, k_y, k_z], \ \forall (k_x, k_y), \tag{11}$$

where J_2 is the number of scales along z. The same processing is also applied on the coarse spatial scale $a_{J_1}[k_x, k_y, k_z]$, and we have

$$a_{J_1}[k_x, k_y, k_z] = a_{J_1,J_2}[k_x, k_y, k_z] + \sum_{j_2=1}^{J_2} w_{J_1,j_2}[k_x, k_y, k_z], \ \forall (k_x, k_y). \tag{12}$$

Hence, we have a 2D-1D undecimated wavelet representation of the input data D:

$$D[k_x, k_y, k_z] = a_{J_1,J_2}[k_x, k_y, k_z] + \sum_{j_1=1}^{J_1} w_{j_1,J_2}[k_x, k_y, k_z] + \sum_{j_2=1}^{J_2} w_{J_1,j_2}[k_x, k_y, k_z] + \sum_{j_1=1}^{J_1} \sum_{j_2=1}^{J_2} w_{j_1,j_2}[k_x, k_y, k_z]. \tag{13}$$

From this expression, we distinguish four kinds of coefficients:

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

$$w_{j_1,j_2}[k_x, k_y, k_z] = (\delta - \bar{h}_{1\mathrm{D}}) \star \left(h_{1\mathrm{D}}^{(j_2-1)} \star a_{j_1-1}[k_x, k_y, .] - h_{1\mathrm{D}}^{(j_2-1)} \star a_{j_1}[k_x, k_y, .] \right). \tag{14}$$

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

$$w_{J_1,j_2}[k_x, k_y, k_z] = h_{1\mathrm{D}}^{(j_2-1)} \star a_{J_1}[k_x, k_y, .] - h_{1\mathrm{D}}^{(j_2)} \star a_{J_1}[k_x, k_y, .]. \tag{15}$$

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

$$w_{j_1,J_2}[k_x, k_y, k_z] = h_{1\mathrm{D}}^{(J_2)} \star a_{j_1-1}[k_x, k_y, .] - h_{1\mathrm{D}}^{(J_2)} \star a_{j_1}[k_x, k_y, .]. \tag{16}$$

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

$$a_{J_1,J_2}[k_x, k_y, k_z] = h_{1\mathrm{D}}^{(J_2)} \star a_{J_1}[k_x, k_y, .]. \tag{17}$$

As the 2D-1D undecimated wavelet transform just described is fully linear, a Gaussian noise remains Gaussian after transformation. Therefore, all thresholding strategies which have been developed for wavelet Gaussian denoising are still valid with the 2D-1D wavelet transform. Denoting TH the thresholding operator, the denoised cube in the case of additive white Gaussian noise is obtained by:

$$\tilde{D}[k_x, k_y, k_z] = a_{J_1,J_2}[k_x, k_y, k_z] + \sum_{j_1=1}^{J_1} \mathrm{TH}(w_{j_1,J_2}[k_x, k_y, k_z]) + \sum_{j_2=1}^{J_2} \mathrm{TH}(w_{J_1,j_2}[k_x, k_y, k_z]) + \sum_{j_1=1}^{J_1} \sum_{j_2=1}^{J_2} \mathrm{TH}(w_{j_1,j_2}[k_x, k_y, k_z]). \tag{18}$$

A typical choice of TH is the hard thresholding operator, i.e. $\mathrm{TH}(x) = 0$ if $|x|$ is below a given threshold τ, and $\mathrm{TH}(x) = x$ if $|x| \geq \tau$. The threshold τ is generally chosen between 3 and 5 times the noise standard deviation (Starck & Murtagh 2006).

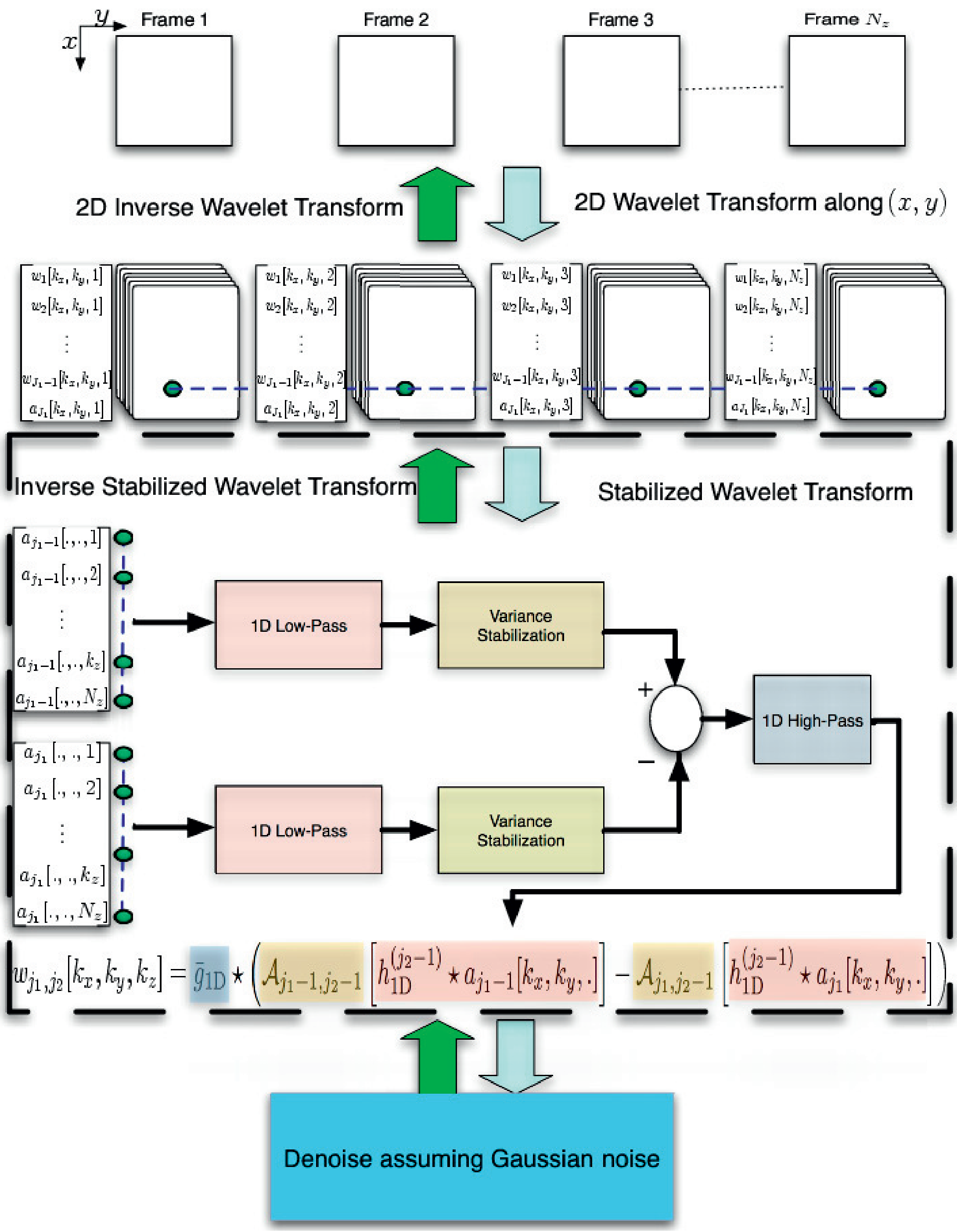

Fig. 5. Overview of MSVST with the 2D-1D IUWT. The diagram summarizes the main steps for computing the detail coefficients w_{j_1,j_2} in (19). The notations are exactly the same as those of Sect. 4.2 with $\bar{g}_{1D} = \delta - \bar{h}_{1D}$.

4.2. Variance stabilization

Putting all pieces together, we are now ready to plug the MSVST into the 2D-1D undecimated wavelet transform. Again, we distinguish four kinds of coefficients that take the following forms:

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

$$w_{j_1,j_2}[k_x,k_y,k_z] = (\delta - \bar{h}_{1D}) \star \left(\mathcal{A}_{j_1-1,j_2-1}\left[h_{1D}^{(j_2-1)} \star a_{j_1-1}[k_x,k_y,.]\right] - \mathcal{A}_{j_1,j_2-1}\left[h_{1D}^{(j_2-1)} \star a_{j_1}[k_x,k_y,.]\right]\right). \quad (19)$$

The schematic overview of the way the detail coefficients w_{j_1,j_2} are computed is illustrated in Fig. 5.

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

$$w_{J_1,j_2}[k_x,k_y,k_z] = \mathcal{A}_{J_1,j_2-1}\left[h_{1D}^{(j_2-1)} \star a_{J_1}[k_x,k_y,.]\right] - \mathcal{A}_{J_1,j_2}\left[h_{1D}^{(j_2)} \star a_{J_1}[k_x,k_y,.]\right]. \quad (20)$$

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

$$w_{j_1,J_2}[k_x,k_y,k_z] = \mathcal{A}_{j_1-1,J_2}\left[h_{1D}^{(J_2)} \star a_{j_1-1}[k_x,k_y,.]\right] - \mathcal{A}_{j_1,J_2}\left[h_{1D}^{(J_2)} \star a_{j_1}[k_x,k_y,.]\right]. \quad (21)$$

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

$$c_{J_1,J_2}[k_x,k_y,k_z] = h_{1D}^{(J_2)} \star a_{J_1}[k_x,k_y,.]. \tag{22}$$

Hence, all 2D-1D wavelet coefficients w_{j_1,j_2} are now stabilized, and the noise on all these wavelet coefficients is Gaussian with known scale-dependent variance that depends solely on h. Denoising is however not straightforward because there is no explicit reconstruction formula available because of the form of the stabilization equations above. Formally, the stabilizing operators $\mathcal{A}_{j_1,j_2}$ and the convolution operators along (x,y) and z 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 defining the multiresolution support (Murtagh et al. 1995) from the stabilized coefficients, and by using an iterative reconstruction scheme.

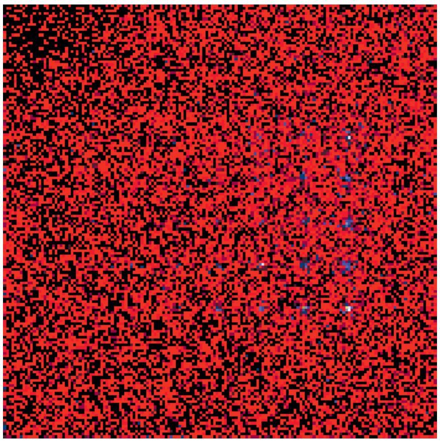

Fig. 6. Image obtained by integrating along the z-axis of the simulated data cube.

4.3. Detection-reconstruction

As the noise on the stabilized coefficients is Gaussian, and without loss of generality, we let its standard deviation equal to 1, we consider that a wavelet coefficient $w_{j_1,j_2}[k_x,k_y,k_z]$ is significant, i.e., not due to noise, if its absolute value is larger than a critical threshold τ, where τ is typically between 3 and 5.

The multiresolution support will be obtained by detecting at each scale the significant coefficients. The multiresolution support for $j_1 \leq J$ and $j_2 \leq J_2$ is defined as

$$M_{j_1,j_2}[k_x,k_y,k_z] = \begin{cases} 1 & \text{if } w_{j_1,j_2}\,[k_x,k_y,k_z] \text{ is significant,} \\ 0 & \text{otherwise.} \end{cases} \tag{23}$$

In words, the multiresolution support M indicates at which scales (spatial and time/energy) and which positions, we have significant signal. We denote $\mathcal{W}$ the 2D-1D undecimated wavelet transform described above, $\mathcal{R}$ the inverse wavelet transform and Y the input noisy data cube.

We want our solution X to preserve the significant structures in the original data by reproducing exactly the same coefficients as the wavelet coefficients of the input data Y, but only at scales and positions where significant signal has been detected (i.e. $M\mathcal{W}X = M\mathcal{W}Y$). At other scales and positions, we want the smoothest solution with the lowest budget in terms of wavelet coefficients. 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 that can be written as follows

$$\min_X \; \| \mathcal{W}X \|_1 \quad \text{subject to} \quad \begin{cases} M\mathcal{W}X = M\mathcal{W}Y \\ \text{and } X \geq 0, \end{cases} \tag{24}$$

where $\|\,.\,\|_1$ 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. 2008a), 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 M from the detection step; number of iterations $N_{\max}$.
Initialize $X^{(0)} = M\mathcal{W}Y = Mw_Y$,
for $t = 1$ to $N_{\max}$ **do**
$\tilde{d} = Mw_Y + (1-M)\mathcal{W}X^{(t-1)}$,
$X^{(t)} = P_+\left(\mathcal{R}\,\mathrm{ST}_{\beta_t}[\tilde{d}]\right)$,
Update the step $\beta_t = (N_{\max} - t)/(N_{\max} - 1)$.
end for

where P_+ is the projector onto the positive orthant, i.e. $P_+(x) = \max(x,0)$. ST_{β_t} is the soft-thresholding operator with threshold β_t, i.e. $\mathrm{ST}_{\beta_t}[x] = x - \beta_t \mathrm{sign}(x)$ if $|x| \geq \beta_t$, and 0 otherwise.

4.4. Algorithm summary

The final MSVST 2D-1D wavelet denoising algorithm is the following:

Require: Input noisy data Y; a low-pass filter h; threshold level τ,
1: _2D-1D-MSVST_: apply the 2D-1D-MSVST to the data using (19)–(22).
2: _Detection_: detect the significant wavelet coefficients that are above τ, and compute the multiresolution support M.
3: _Reconstruction_: reconstruct the denoised data using the algorithm above.

5. Experimental results and discussion

5.1. MSVST-2D-1D versus MSVST-2D

We have simulated a data cube according to the procedure described in Sect. 2.2. The cube contains several sources, with spatial positions on a grid. It contains seven columns and five rows of LAT sources (i.e. 35 sources) with different power-law spectra. The cube size is $161 \times 161 \times 31$, with a total number of photons equal to 25 948, i.e. an average of 0.032 photons per pixel. Figure 6 shows the 2D image obtained after integrating the simulated data cube along the z-axis. Figure 7 shows a comparison between 2D-MSVST denoising of this image, and the image obtained by first applying a 2D-1D-MSVST denoising to the input cube, and integrating afterward along the z-axis. Figure 7 upper

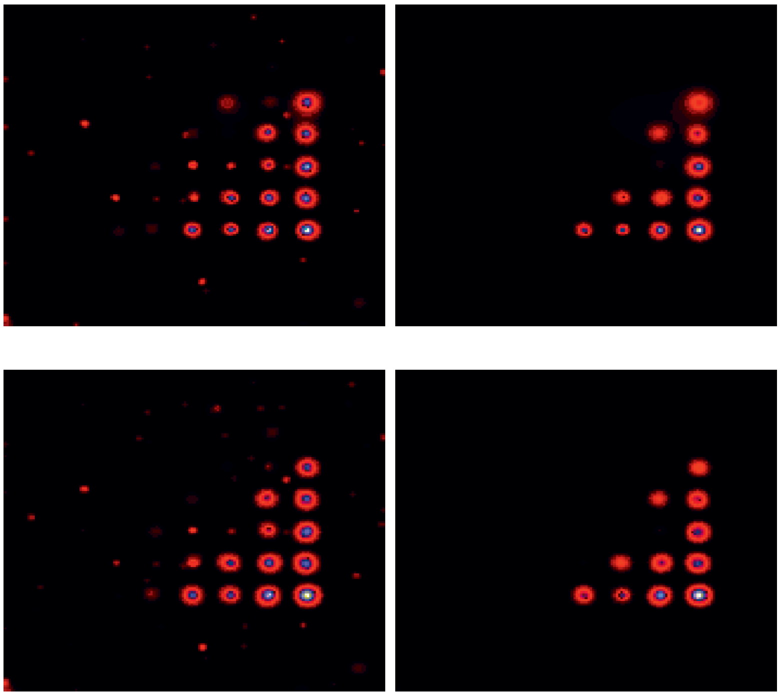

Fig. 7. *Top*, 2D-MSVST filtering on the integrated image with respectively a $\tau = 3$ and a $\tau = 5$ detection level. *Bottom*, integrated image after a 2D-1D-MSVST denoising of the simulated data cube, with respectively a $\tau = 4$ and a $\tau = 6$ detection level.

left and right show denoising results for the 2D-MSVST with respectively threshold values $\tau = 3$ and $\tau = 5$, and Fig. 7 bottom left and right show the results for the 2D-1D-MSVST using respectively $\tau = 4$ and $\tau = 6$ detection levels. The reason for using a higher threshold level for the 2D-1D cube is to correct for multiple hypothesis testings, and to get the same control over global statistical error rates. Roughly speaking, the number of false detections increases with the number of coefficients being tested simultaneously. Therefore, one must correct for multiple comparisons using e.g. the conservative Bonferroni correction or the false discovery rate (FDR) procedure Benjamini & Hochberg (1995). As the number of coefficients is much higher with the whole 2D-1D cube, the critical detection threshold τ of 2D-1D denoising must be higher to have a false detection rate comparable to the 2D denoising. As we can clearly see from Fig. 7, the results are very close. This means that applying a 2D-1D denoising on the cube instead of a 2D denoising on the integrated image does not degrade the detection power of the MSVST. The main advantage of the 2D-1D-MSVST is the fact that we recover the spectral (or temporal) information for each spatial position. Figure 8 shows two frames (frame 16 top left and frame 25 bottom left) of the input cube and the same frames after the 2D-1D-MSVST denoising top right and bottom right. Figure 9 displays the obtained spectra at two different spatial positions $(112, 47)$ and $(126, 79)$ which correspond to the centers of two distinct sources.

5.2. Time-varying source detection

We have simulated a time varying source in a cube of size $64 \times 64 \times 128$. The source has a Gaussian shape both in space and time. It is centered in the middle of the cube at $(32, 32, 64)$; i.e. its brightest point is at this location. The standard deviation of the Gaussian is 1.8 in space (pixel unit), and 1.2 along time (frame unit). The total flux of the source (i.e. spatial and temporal integration) is 100. We have added a background level of 0.1. Finally, Poisson noise was generated. Figure 10 shows respectively from left to right an image of the original source, the flux per time frame and the integration of all noisy frames along the time axis. As it can be seen, the source is hardly detectable in Fig. 10 right. By running the 2D-MSVST denoising method on the time-integrated image, we were not able to detect it. Then we applied the 2D-1D-MSVST denoising method on the noisy 3D data set. This time, we were able to restore the source with a threshold level $\tau = 6$. Figure 11 left depicts one frame (frame 64) of the denoised cube, and Fig. 11 right shows the flux of the recovered source per frame (dotted line). The solid and thick-solid lines show respectively the flux per time frame after background subtraction in the noisy data and the original noise-free data set. We can conclude from this experiment that the 2D-1D-MSVST is able to recover rapidly time-varying sources in the spatio-temporal data set, whereas even a robust algorithm such as the 2D-MSVST method will completely fail if we integrate along the time axis. This was expected since the co-addition of all frames mixes the few frames containing the source with those which contain only the noisy background. Co-adding followed by a 2D detection is clearly suboptimal, except if we repeat the denoising procedure with many temporal windows with varying size. We can also notice that the 2D-1D-MSVST is able to recover very well the times at which the source flares, although the source is slightly spread out on the time axis and the flux of the source is not very well estimated, and other methods such as maximum likelihood should be preferred for a correct flux estimation, once the sources have been detected.

5.3. Diffuse emission of the Galaxy

In this experiment, we have simulated a $720 \times 360 \times 128$ cube using the Galprop code Strong et al. (2007) that has a model of the diffuse gamma-ray emission of the Milky Way. The units

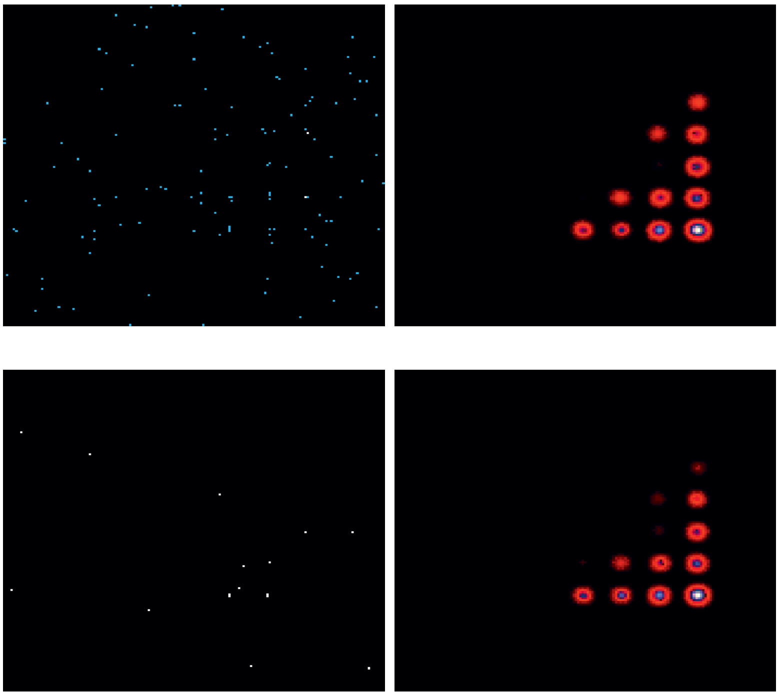

Fig. 8. *Top*, frame number 16 of the input cube and the same frame after the 2D-1D-MSVST filtering at 6σ. *Bottom*, frame number 25 of the input cube and the same frame after the 2D-1D-MSVST filtering at 6σ.

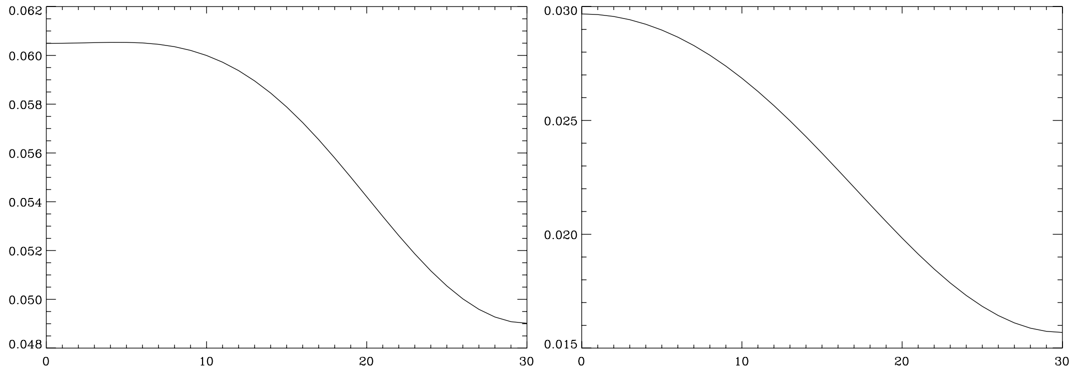

Fig. 9. Pixel spectra at two different spatial locations after the 2D-1D-MSVST filtering.

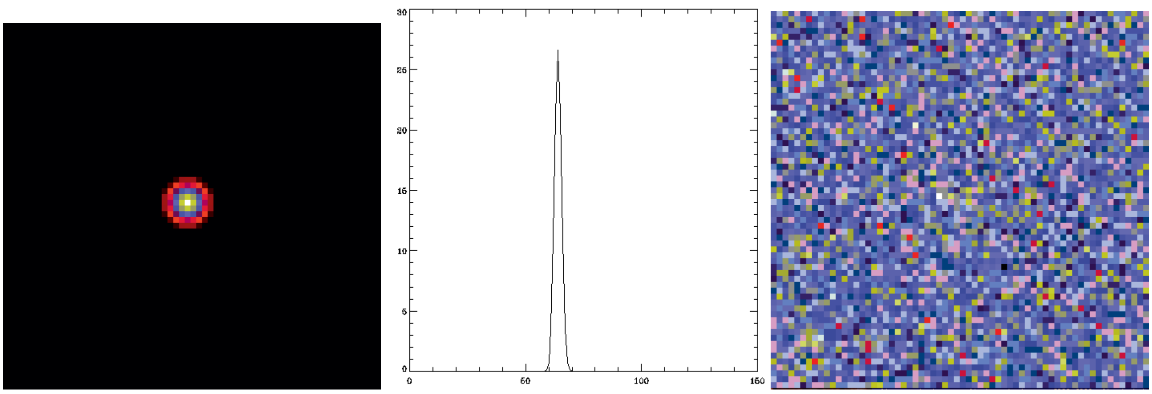

Fig. 10. Time-varying source. *From left to right*, simulated source, temporal flux, and co-added image along the time axis of noisy data cube.

of the pixels are photons $cm^{-2}\ s^{-1}\ sr^{-1}\ MeV^{-1}$. The gridding in Galactic longitude and latitude is 0.5 degrees, and the 128 energy planes are logarithmically spaced from 30 MeV to 50 GeV. A six months LAT data set was created by multiplying the simulated cube with the exposure (6 months), and by convolving each energy band with the point spread function of the LAT instrument.

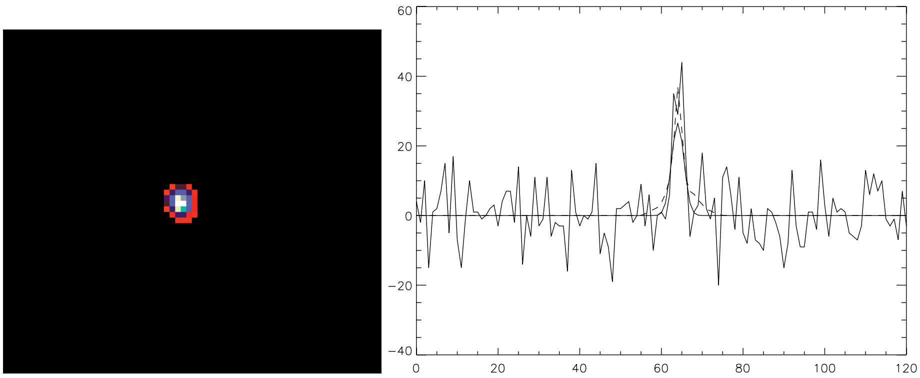

Fig. 11. Recovered time-varying source. *Left*, one frame of the denoised cube. *Right*, flux per time frame for the noisy data after background subtraction (solid line), for the original noise-free cube (thick-solid line) and for the recovered source (dashed line).

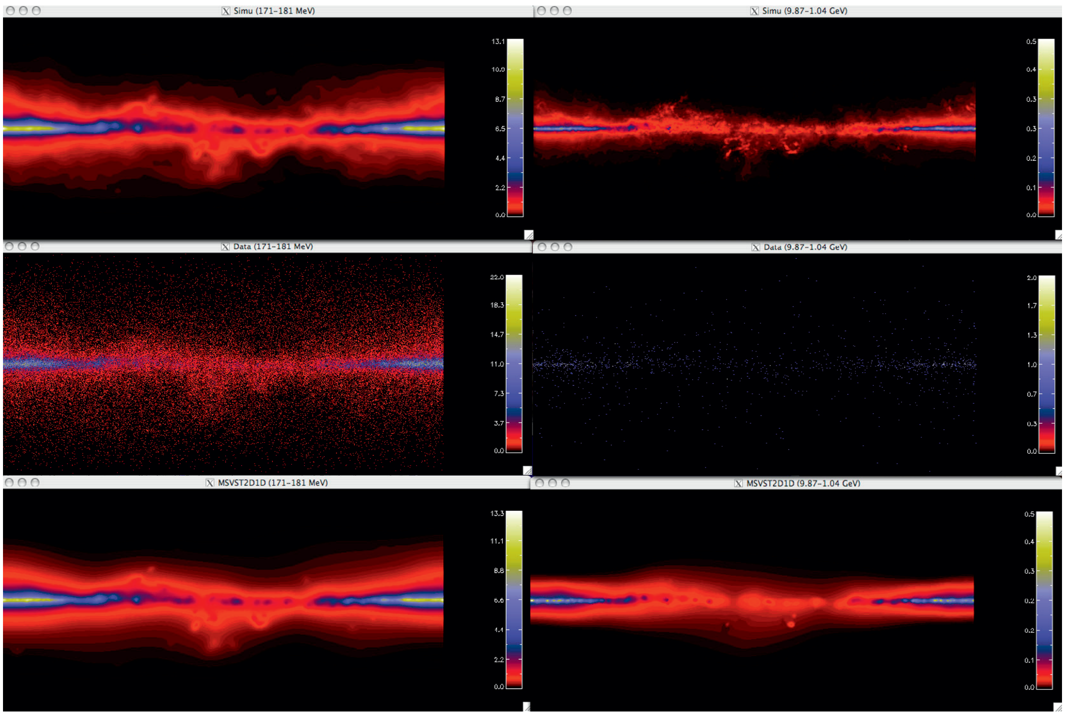

Fig. 12. *Left, from top to bottom*, simulated data of the diffuse gamma-ray emission of the Milky Way in energy band 171−181 MeV, noisy simulated data and filtered data using the MSVST. *Right*, same images for energy band 9.87−1.04 GeV.

The PSF strongly varies with the energy. Finally we have created the noisy observations assuming a Poisson noise distribution.

Figure 12 left shows from top to bottom the original simulated data, the noisy data and the filtered data for the band at energy 171−181 Mev. The same figures for the band 9.87−1.04 GeV are shown in Fig. 12 right.

6. Conclusion

The motivations for a reliable nonparametric source detection algorithm to apply to *Fermi* LAT data are clear. Especially for the relatively short time ranges over which we will want to study sources, the data will be squarely in the low counts regime with widely varying response functions and significant celestial foregrounds. In this paper, we have shown that the MSVST, associated with a 2D-1D wavelet transform, is a very efficient way to detect time-varying sources. The proposed algorithm is as powerful as the 2D-MSVST applied to co-added frames to detect a source if the latter is slowly varying or constant over time. But when the source is rapidly varying, we lose some detection power when we co-add frames having no source and those containing the sources. Our approach gives us an alternative to frame-co-adding and outperforms the 2D algorithms on the co-added frames. Unlike 2D denoising, our method fully exploits the information in the 3D data set and allows to recover the source dynamics by detecting temporally varying sources.

Acknowledgements. We thank Jean-Marc Casandjian for providing us the simulated data set of the diffuse emission of the Galaxy and Jeff Scargle for his helpful comments and critics. This work was partially supported by the French National Agency for Research (ANR -08-EMER-009-01).

References

Anscombe, F. 1948, Biometrika, 15, 246
Benjamini, Y., & Hochberg, Y. 1995, J. R. Stat. Soc. B, 57, 289
Bijaoui, A., & Jammal, G. 2001, Signal Processing, 81, 1789

Donoho, D. L. 1993, Proc. Symp. Applied Mathematics: Different Perspectives on Wavelets, 47, 173

Donoho, D. L. 2004, For Most Large Underdetermined Systems of Linear Equations, the minimal ℓ^1-norm solution is also the sparsest solution, Tech. rep., Department of Statistics of Stanford Univ.

Fryźlewicz, P., & Nason, G. P. 2004, J. Comp. Graph. Stat., 13, 621

Hartman, R. C., Bertsch, D. L., Bloom, S. D., et al. 1999, VizieR Online Data Catalog, 212, 30079

Holschneider, M., Kronland-Martinet, R., Morlet, J., & Tchamitchian, P. 1989, in Wavelets: Time-Frequency Methods and Phase-Space (Springer-Verlag), 286

Kolaczyk, E. 1997, ApJ, 483, 349

Kolaczyk, E., & Nowak, R. 2004, Ann. Stat., 32, 500

Mallat, S. 1998, A Wavelet Tour of Signal Processing (Academic Press)

Murtagh, F., Starck, J.-L., & Bijaoui, A. 1995, A&AS, 112, 179

Nowak, R., & Baraniuk, R. 1999, IEEE Transactions on Image Processing, 8, 666

Olivo-Marin, J. C. 2002, Pattern Recognition, 35, 1989

Pierre, M., Valtchanov, I., Altieri, B., et al. 2004, Journal of Cosmology and Astro-Particle Physics, 9, 11

Pierre, M., Chiappetti, L., Pacaud, F., et al. 2007, MNRAS, 382, 279

Scargle, J. D. 1998, ApJ, 504, 405

Shensa, M. J. 1992, IEEE Transactions on Signal Processing, 40, 2464

Slezak, E., de Lapparent, V., & Bijaoui, A. 1993, ApJ, 409, 517

Starck, J.-L., & Pierre, M. 1998, A&AS, 128

Starck, J.-L., & Murtagh, F. 2006, Astronomical Image and Data Analysis, Astronomical image and data analysis, ed. J.-L. Starck, & F. Murtagh (Berlin: Springer), Astronomy and astrophysics library

Starck, J.-L., Murtagh, F., & Bijaoui, A. 1998, Image Processing and Data Analysis, The Multiscale Approach (Cambridge University Press)

Starck, J.-L., Fadili, M., & Murtagh, F. 2007, IEEE Transactions on Image Processing, 16, 297

Strong, A. W., Moskalenko, I. V., & Ptuskin, V. S. 2007, Annual Review of Nuclear and Particle Science, 57, 285

Timmermann, K. E., & Nowak, R. 1999, IEEE Transactions on Signal Processing, 46, 886

Willet, R., & Nowak, R. 2005, IEEE Transactions on Information Theory, submitted

Willett, R. 2006, SCMA IV, in press

Yamada, I. 2001, in Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications, ed. D. Butnariu, Y. Censor, & S. Reich (Elsevier)

Zhang, B., Fadili, M., & Starck, J.-L. 2008a, IEEE Transactions on Image Processing, 17, 1093

Zhang, B., Fadili, M. J., Starck, J.-L., & Digel, S. W. 2008b, Stat. Methodol., 5, 387