



Poisson noise removal in multivariate count data

Jalal M. Fadili, Jean-Luc Starck, B. Zhang, Seth Digel

► To cite this version:

Jalal M. Fadili, Jean-Luc Starck, B. Zhang, Seth Digel. Poisson noise removal in multivariate count data. SPIE Wavelets XIII, Aug 2009, San Diego, United States. 7446, 74461B-11p., 2009, <10.1117/12.825063>. <hal-00813572>

HAL Id: hal-00813572

<https://hal.archives-ouvertes.fr/hal-00813572>

Submitted on 15 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.

Poisson noise removal in multivariate count data

M.J. Fadili^a, J.-L. Starck^b, B. Zhang^c and S. Digel^d

^a GREYC CNRS-ENSICAEN-Univ., 14050 Caen Cedex, France.

^b LAIM CEA/DSM-CNRS-Univ. Paris Diderot, Saclay, F-91191 Gif-Sur-Yvette, France.

^c Medisys Research Lab, Philips Healthcare, 92156 Suresnes Cedex, France.

^d SLAC & Kavli Institute for Particle Astrophysics and Cosmology, Stanford, CA 94075, USA.

ABSTRACT

The Multi-scale Variance Stabilization Transform (MSVST) has recently been proposed for 2D Poisson data denoising.¹ In this work, we present an extension of the MSVST with the wavelet transform to multivariate data—each pixel is vector-valued—, where the vector field dimension may be the wavelength, the energy, or the time. Such data can be viewed naively as 3D data where the third dimension may be time, wavelength or energy (e.g. hyperspectral imaging). But this naive analysis using a 3D MSVST would be awkward as the data dimensions have different physical meanings. A more appropriate approach would be to use a wavelet transform, where the time or energy scale is not connected to the spatial scale. We show that our multivalued extension of MSVST can be used advantageously for approximately Gaussianizing and stabilizing the variance of a sequence of independent Poisson random vectors. This approach is shown to be fast and very well adapted to extremely low-count situations. We use a hypothesis testing framework in the wavelet domain to denoise the Gaussianized and stabilized coefficients, and then apply an iterative reconstruction algorithm to recover the estimated vector field of intensities underlying the Poisson data. Our approach is illustrated for the detection and characterization of astrophysical sources of high-energy gamma rays, using realistic simulated observations. We show that the multivariate MSVST permits efficient estimation across the time/energy dimension and immediate recovery of spectral properties.

Keywords: Poisson noise, multivariate data, wavelets, multi-scale variance stabilization.

1. INTRODUCTION

A host of estimation methods have been proposed in the literature for non-parametric Poisson noise removal. Major contributions consist of variance stabilization: this is a classical approach which consists in preprocessing the data by applying a variance stabilizing transform (VST) such as the Anscombe transform.² It can be shown that the transformed data are approximately stationary, independent and Gaussian. Once we are brought to the Gaussian denoising problem, standard approaches are used before the VST is inverted to get the final estimate. Haar-Fisz transform is another widely used VST,^{3,4} which combines the Fisz transform⁵ within the Haar transform. Jansen⁶ introduced a conditional variance stabilization (CVS) approach which can be applied in any wavelet domain resulting in stabilized coefficients. However, these transformations are only valid for a sufficiently large number of counts per pixel (and of course, for a larger still number of counts, the Poisson distribution becomes Gaussian with equal mean and variance).

Hypothesis testing has also been proposed for Poisson noise removal. Kolaczyk⁷ first introduced a Haar domain threshold, which implements a hypothesis testing procedure based on a user-specified false positive rate (FPR). Zhang et al.⁸ extended the hypothesis testing to the biorthogonal Haar domain, leading to a more regular reconstruction for smooth intensities. Bijaoui and Jammal⁹ derived the probability density function (*pdf*) of any wavelet coefficient allowing hypothesis testing in a general wavelet domain, but is computationally complex than Haar-based methods due to the lack of closed-form *pdf* expressions. Kolaczyk¹⁰ proposed “corrected” versions of the usual Gaussian-based thresholds for Poisson data. However, the asymptotic approximation adopted by¹⁰ may not allow reasonable solutions in low-count situations.

Further author information: (Send correspondence to M.J. Fadili) E-mail: Jalal.Fadili@greyc.ensicaen.fr, Telephone: 33 231 45 29 22.

Empirical Bayesian estimators for Poisson noise removal were studied in.^{6,11–13} The low-intensity case apart, Bayesian approaches generally outperform the direct wavelet filtering.^{14,15} Poisson denoising has also been formulated as a penalized maximum likelihood estimation problem^{16–18} within wavelet, wedgelet and platelet dictionaries.

In a recent paper, Zhang et al.¹ have proposed to combine variance stabilization and multiscale transforms, leading to the Multi-Scale Variance Stabilization Transform (MSVST). MSVST can be seen as a generalization of the Anscombe transform; see section 2 for details. This new approach 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 propose a multivalued extension of MSVST, which allows us to remove the Poisson noise in multivariate 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 this multivalued extension can be used advantageously for approximately Gaussianizing and stabilizing the variance of a sequence of independent Poisson random vectors. This approach is shown to be fast and very well adapted to low-count situations. We use a hypothesis testing framework in the wavelet domain to denoise the Gaussianized and stabilized coefficients, and then apply an iterative reconstruction algorithm to recover the estimated vector field of intensities underlying the Poisson data. Our approach is illustrated for the detection and characterization of astrophysical sources of high-energy gamma rays, using realistic simulated observations. We show that the multivariate MSVST permits efficient estimation across the time/wavelength/energy dimension and recovery of spectral properties. Section 2 reviews the MSVST method relative to the isotropic undecimated wavelet transform and section 3 shows how it can be extended to the 2D-1D case. Section 4 presents some experiments on realistic simulated astronomical data. Conclusions are drawn in section 5.

2. THE 2D MULTISCALE VARIANCE STABILIZATION TRANSFORM (MSVST)

In this section, we review the MSVST method,¹ restricted here to the wavelet transform. Indeed, the MSVST can use other transforms such as the ridgelet or the curvelet transforms; see.¹

2.1 VST of a filtered Poisson process

Let's start with a simple 1D case. Given \mathbf{X} a sequence of n independent Poisson random variables $X_i, i = 1, \dots, n$, 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 $k = 1, 2, \dots$.

If $h = \delta$, then we recover the Anscombe VST² 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 shown¹ that for a filter h , the transform $Z = b \operatorname{sign}(Y + c) \sqrt{|Y + c|}$ with b and c defined as

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

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 in¹ have also proved that Z is asymptotically normally distributed with mean $b\sqrt{|\tau_1|\lambda}$ and unit variance.

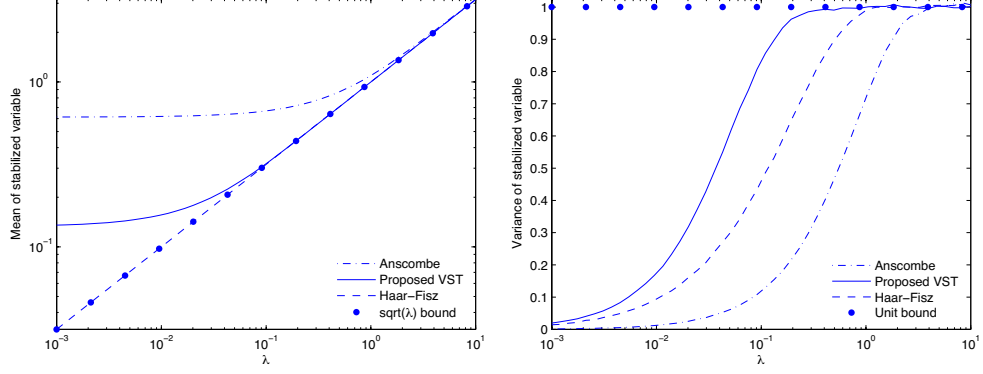


Figure 1. Behavior of the expectation $\mathbb{E}[Z]$ (left) and variance $\text{Var}[Z]$ (right) as a function of the underlying intensity, for the Anscombe VST, 2D Haar-Fisz VST, and out VST with the 2D B_3 -Spline filter as a low-pass filter h .

Fig.1 shows the Monte-Carlo estimates of the expectation $\mathbb{E}[Z]$ (left) and the variance $\text{Var}[Z]$ (right) obtained from $2 \cdot 10^5$ Poisson noise realizations of \mathbf{X} , plotted as a function of the intensity λ for both Anscombe² (dashed-dotted), Haar-Fisz (dashed)¹⁹ 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 $\text{Var}[Z]$ 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 $\lambda \gtrsim 10$, using Haar-Fisz for $\lambda \gtrsim 1$, and using out VST (after low-pass filtering with the chosen h) for $\lambda \gtrsim 0.1$.

2.2 MSVST with the Wavelet Transform

2.2.1 The 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 $W = \{w_1, \dots, w_J, a_J\}$, where w_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” algorithm²⁰

$$\begin{aligned} a_{j+1}[l] &= (\bar{h}^{\uparrow(j)} \star a_j)[l] = \sum_k h[k] a_j[l + 2^j k], \\ w_{j+1}[l] &= (\bar{g}^{\uparrow(j)} \star a_j)[l] = \sum_k g[k] a_j[l + 2^j k]. \end{aligned} \quad (2)$$

where $h^{\uparrow(j)}[l] = h[l]$ if $l/2^j \in \mathbb{Z}$ and 0 otherwise, $\bar{h}[l] = h[-l]$ is the time-reversed version of h , and “ \star ” denotes discrete circular convolution. The reconstruction is given by

$$a_j[l] = \frac{1}{2} \left((\tilde{h}^{\uparrow(j)} \star a_{j+1})[l] + (\tilde{g}^{\uparrow(j)} \star w_{j+1})[l] \right).$$

The filter bank $(h, g, \tilde{h}, \tilde{g})$ needs to satisfy the so-called exact reconstruction condition^{21,22}.

The so-called Isotropic UWT (IUWT)²³ 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 applications²⁴ and biomedical imaging²⁵ to detect isotropic objects.

2.2.2 MSVST with the UWT

Now the VST can be combined with the UWT to stabilize the wavelet coefficients of a standard separable UWT. The VST is applied to the approximation coefficients $(a_j)_j$, leading to the following scheme:

$$\text{UWT} \begin{cases} a_j &= \bar{h}^{\uparrow(j-1)} \star a_{j-1} \\ w_j &= \bar{g}^{\uparrow(j-1)} \star a_{j-1} \end{cases} \implies \begin{matrix} \text{MS-VST} \\ + \\ \text{UWT} \end{matrix} \begin{cases} a_j &= \bar{h}^{\uparrow(j-1)} \star a_{j-1} \\ w_j &= \bar{g}^{\uparrow(j-1)} \star \mathcal{A}_{j-1}(a_{j-1}) \end{cases} \quad (3)$$

where

$$\mathcal{A}_j(a_j) = b^{(j)} \text{sign}(a_j + c^{(j)}) \sqrt{|a_j + c^{(j)}|}. \quad (4)$$

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

$$c^{(j)} = \frac{7\tau_2^{(j)}}{8\tau_1^{(j)}} - \frac{\tau_3^{(j)}}{2\tau_2^{(j)}}, \quad b^{(j)} = 2\sqrt{\frac{|\tau_1^{(j)}|}{\tau_2^{(j)}}}. \quad (5)$$

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 . Since these constants scale-dependent, so is the VST, hence the name MSVST.

In the case of the isotropic UWT (IUWT), given the structure of the filter bank, the stabilization procedure is given by

$$\text{IUWT} \begin{cases} a_j &= \bar{h}^{\uparrow(j-1)} \star a_{j-1} \\ w_j &= a_{j-1} - a_j \end{cases} \implies \begin{matrix} \text{MSVST} \\ + \\ \text{IUWT} \end{matrix} \begin{cases} a_j &= \bar{h}^{\uparrow(j-1)} \star a_{j-1} \\ w_j &= \mathcal{A}_{j-1}(a_{j-1}) - \mathcal{A}_j(a_j) \end{cases}, \quad (6)$$

where \mathcal{A}_{j-1} is defined similarly to \mathcal{A}_j in (4).

2.2.3 Asymptotic Distribution of the Detail Coefficients

The following result shows the asymptotic normality of the MSVST detail coefficients under the null hypothesis H_0 that the underlying true intensity λ is locally homogeneous. The asymptotic normal distribution has an intensity-independent variance which relies solely on the filter h and the current scale.

THEOREM 2.1.

(i) *MSVST+IUWT: let $b^{(j)} := \text{sign}(\tau_1^{(j)})/\sqrt{|\tau_1^{(j)}|}$. If λ is constant within the support of the filter $h^{(j)}[k - \cdot]$, then*

$$w_j[k] \xrightarrow[\lambda \rightarrow +\infty]{\mathcal{D}} \mathcal{N}\left(0, \frac{\tau_2^{(j-1)}}{4\tau_1^{(j-1)^2} + \frac{\tau_2^{(j)}}{4\tau_1^{(j)^2}} - \frac{\langle h^{(j-1)}, h^{(j)} \rangle}{2\tau_1^{(j-1)}\tau_1^{(j)}}\right), \quad (7)$$

where $\langle \cdot, \cdot \rangle$ represents the scalar product.

(ii) *MSVST+UWT: let $b^{(j)} := 2\sqrt{|\tau_1^{(j)}|/\tau_2^{(j)}}$, and $h^{(j)} = \bar{h}^{\uparrow(j-1)} \star \dots \star \bar{h}^{\uparrow(1)} \star \bar{h}$ for $j \geq 1$ and $h^{(0)} = \delta$. If λ is constant within the support of the filter $(\bar{g}^{\uparrow(j-1)} \star h^{(j-1)})[k - \cdot]$, then $w_j[k] \xrightarrow[\lambda \rightarrow +\infty]{\mathcal{D}} \mathcal{N}(0, \sigma_j^2)$, where*

$$\sigma_j^2 = \frac{1}{\tau_2^{(j-1)}} \sum_{m,n} \bar{g}^{\uparrow(j-1)}[m] \bar{g}^{\uparrow(j-1)}[n] \sum_k h^{(j-1)}[k] h^{(j-1)}[k+m-n] \quad (8)$$

Again, the above values of $b^{(j)}$, $c^{(j)}$, $\tau_k^{(j)}$ and σ_j can all be pre-computed once (h, g) has been chosen.

3. 2D-1D MSVST POISSON DENOISING

3.1 2D-1D Wavelet Transform

In the previous section, we have seen how a Poisson noise can be removed from 2D image using the wavelet transform and the MSVST. Extension to a q D data sets is straightforward by separable tensor product, and the denoising will perform very well 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, wavelength or energy, we are clearly not in this situation, and the naive analysis of a 3D wavelet does not make sense. Therefore, we want to analyze the

data with a wavelet transform, where the scale along the non-spatial dimension is not connected to the spatial scale. Hence, an ideal wavelet function would be defined by:

$$\psi(x, y, z) = \psi^{(xy)}(x, y) \psi^{(z)}(z), \quad (9)$$

where $\psi^{(xy)}$ is the spatial wavelet and $\psi^{(z)}$ is the wavelet along the non-spatial dimension.

In the following, for the sake of clarity and without loss of generality, we will consider only the IUWT filter bank in the spatial domain. We note j_1 the spatial resolution index (i.e. scale = 2^{j_1}), j_2 resolution index along the non-spatial dimension. 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 , and then, for each spatial location (k_x, k_y) and for each 2D wavelet scale scale j_1 , to apply a 1D wavelet transform along z on the spatial wavelet coefficients.

Therefore, given the properties of the IUWT filter bank, the 2D-1D IUWT of the input data S is such that

$$S[k_x, k_y, k_z] = a_{J_1, J_2}[k_x, k_y, k_z] + \sum_{j_1=1}^{J_1} w_{j_1, J_2}[k_x, k_y, k_z] + \sum_{j_2=1}^{J_2} w_{J_1, j_2}[k_x, k_y, k_z] + \sum_{j_1=1}^{J_1} \sum_{j_2=1}^{J_2} w_{j_1, j_2}[k_x, k_y, k_z]. \quad (10)$$

In this expression, we distinguish four kinds of coefficients:

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

$$w_{j_1, j_2}[k_x, k_y, k_z] = (\delta - \bar{h}) \star \left(h^{(j_2-1)} \star a_{j_1-1}[k_x, k_y, \cdot] - h^{(j_2-1)} \star a_{j_1}[k_x, k_y, \cdot] \right), \quad (11)$$

where h is the low-pass defining the IUWT filter bank.

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

$$w_{J_1, j_2}[k_x, k_y, k_z] = h^{(j_2-1)} \star a_{J_1}[k_x, k_y, \cdot] - h^{(j_2)} \star a_{J_1}[k_x, k_y, \cdot]. \quad (12)$$

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

$$w_{j_1, J_2}[k_x, k_y, k_z] = h^{(J_2)} \star a_{j_1-1}[k_x, k_y, \cdot] - h^{(J_2)} \star a_{j_1}[k_x, k_y, \cdot]. \quad (13)$$

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

$$a_{J_1, J_2}[k_x, k_y, k_z] = h^{(J_2)} \star a_{J_1}[k_x, k_y, \cdot]. \quad (14)$$

3.2 2D-1D MSVST

Putting all pieces together, we are now ready to plug the MSVST into the 2D-1D IUWT. The idea again is to first stabilize the (low-pass) approximation subbands, and then compute the stabilized detail coefficients. As in section 3.1, we distinguish the previous four types of coefficients for which the 2D-1D MSVST takes the following forms:

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

$$w_{j_1, j_2}[k_x, k_y, k_z] = (\delta - \bar{h}) \star \left(\mathcal{A}_{j_1-1, j_2-1} \left[h^{(j_2-1)} \star a_{j_1-1}[k_x, k_y, \cdot] \right] - \mathcal{A}_{j_1, j_2-1} \left[h^{(j_2-1)} \star a_{j_1}[k_x, k_y, \cdot] \right] \right). \quad (15)$$

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

$$w_{J_1, j_2}[k_x, k_y, k_z] = \mathcal{A}_{J_1, j_2-1} \left[h^{(j_2-1)} \star a_{J_1}[k_x, k_y, \cdot] \right] - \mathcal{A}_{J_1, j_2} \left[h^{(j_2)} \star a_{J_1}[k_x, k_y, \cdot] \right]. \quad (16)$$

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

$$w_{j_1, J_2}[k_x, k_y, k_z] = \mathcal{A}_{j_1-1, J_2} \left[h^{(J_2)} \star a_{j_1-1}[k_x, k_y, \cdot] \right] - \mathcal{A}_{j_1, J_2} \left[h^{(J_2)} \star a_{j_1}[k_x, k_y, \cdot] \right]. \quad (17)$$

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

$$c_{J_1, J_2}[k_x, k_y, k_z] = h^{(J_2)} \star a_{J_1}[k_x, k_y, \cdot]. \quad (18)$$

Now, all 2D-1D wavelet coefficients w_{j_1, j_2} are stabilized. In the same vein as in Theorem 2.1, it can be shown that under the null hypothesis that the 2D-1D intensity Λ is locally homogeneous within the support of the 2D-1D wavelet at scales (j_1, j_2) , the noise on the stabilized coefficients is Gaussian with known scale-dependent variance that depends solely on h . Thus one can detect significant detail coefficients by binary hypothesis testing or hard thresholding at a given critical threshold. The appeal of a binary hypothesis testing approach is that it allows quantitative control of statistical significance.

3.3 Iterative Reconstruction

Following the detection step, we have to invert the MSVST scheme to reconstruct an estimate of the underlying intensity Λ . This is however not straightforward because there is no explicit reconstruction formula available. The formal reason is that 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 reformulate the reconstruction as a sparsity-promoting convex optimization problem. In the sequel, as the noise on the stabilized coefficients is Gaussian, we assume without loss of generality that it is standardized to a unit variance.

We define the multiresolution support \mathcal{M}_{j_1, j_2} which is determined by the set of detected significant coefficients at each scale $j_1 \leq J$ and $j_2 \leq J_2$ and location (k_x, k_y, k_z) , i.e.,

$$\mathcal{M}_{j_1, j_2}[k_x, k_y, k_z] = \begin{cases} 1 & \text{if } w_{j_1, j_2}[k_x, k_y, k_z] \text{ is significant,} \\ 0 & \text{otherwise.} \end{cases} \quad (19)$$

We denote \mathbf{W} the 2D-1D undecimated wavelet transform described above, \mathbf{R} the inverse transform and S the observed Poisson noisy data set. We are seeking an estimate that preserves the significant structures in the original data by reproducing exactly the same coefficients as the wavelet coefficients of the input data S , but only at scales and positions where significant signal has been detected. 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 estimated intensity data Λ . Therefore our reconstruction is formulated as a constrained sparsity-promoting minimization problem over the transform coefficients w

$$\min_w \|w\|_1 \quad \text{subject to} \quad \begin{cases} \mathcal{M}w = \mathcal{M}\mathbf{W}S \\ \text{and } \mathbf{R}w \geq 0 \end{cases}, \quad (20)$$

and the intensity estimate $\hat{\Lambda}$ is reconstructed as $\hat{\Lambda} = \mathbf{R}\hat{w}$, where \hat{w} is a minimizer of (20). Note that the set of minimizers is not empty by coercivity. Moreover, this problem is a convex optimization problem which can be cast as a Linear Program (LP) and solved using interior-point methods. However, the computational complexity of the LP solver increases dramatically with the size of the problem. Here we propose an alternative based on the hybrid steepest descent (HSD).²⁶ The HSD approach allows minimizing convex functionals over the intersection of fixed point sets of nonexpansive mappings. It is much faster than LP, and in our problem, the nonexpansive mappings do have closed forms. Transposed into our context, its main steps can be summarized as follows:

Inputs: noisy data S ; a low-pass filter h ; multiresolution support \mathcal{M} from the detection step; number of iterations N_{\max} .

Initialization: $w^{(0)} = \mathcal{M}\mathbf{W}S = \mathcal{M}w_S$,

for $t = 1$ **to** N_{\max} **do**

$$1. \tilde{w}_{j_1, j_2} = \mathcal{M}_{j_1, j_2} w_{S_{j_1, j_2}} + (1 - \mathcal{M}_{j_1, j_2}) w_{j_1, j_2}^{(t-1)}, \forall j_1, j_2.$$

$$2. w^{(t)} = \mathbf{W}P_+(\mathbf{R} \text{ST}_{\beta_t}[\tilde{w}]),$$

3. Update the step $\beta_t = (N_{\max} - t)/(N_{\max} - 1)$.

Output: $\hat{\Lambda} = \mathbf{R}w^{(N_{\max})}$.

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 . A careful convergence analysis of this algorithm is given in.¹

In summary, 2D-1D wavelet Poisson denoising with the MSVST involves the following three main steps:

1. **Transformation** : Compute the 2D-1D-MSVST using (15)-(18).
 2. **Detection** : Detect significant detail coefficients by hypothesis testing. Here, we take benefit from the asymptotic Gaussianity of the noise in the stabilized data. We can then detect the significant coefficients by classical binary hypothesis testing (corrected or not for multiple testing and dependencies).
 3. **Reconstruction** : Reconstruct the denoised data using the knowledge of the detected coefficients in the algorithm above.
-

4. EXPERIMENTAL RESULTS

4.1 MSVST-2D-1D versus MSVST-2D

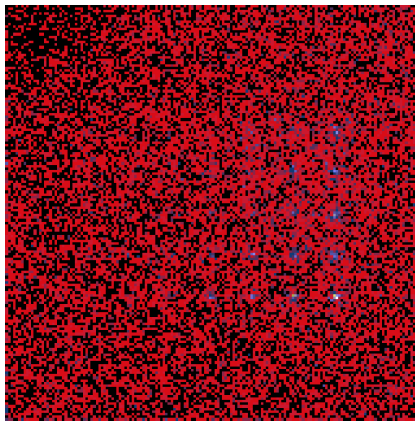


Figure 2. Image obtained by summing along the z -axis of the noisy data set.

We have simulated a 2D-1D data set containing seven columns and five rows of 2D Gaussian-shaped sources. Each source has a different power-law spectrum (i.e. the z -axis is energy). The data set size is $161 \times 161 \times 31$, with an average of 0.032 photons per pixel. As the number of photon counts is very low, for visual convenience, we show in Fig. 2 the 2D image obtained after summing the noisy data set along the z -axis. Fig. 3 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 data, and summing afterward the denoised frames along the z -axis. Fig. 3(a)-(b)

show denoising results for the 2D-MSVST with respectively critical threshold values $\tau = 3$ and $\tau = 5$, and Fig. 3(c)-(d) depict the results for the 2D-1D-MSVST using respectively $\tau = 4$ and $\tau = 6$ critical thresholds*. The reason for using a higher threshold level for the 2D-1D data 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.²⁷ As the number of coefficients is much higher with the whole 2D-1D data set, 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. 3, the results are very close. This means that applying a 2D-1D denoising on the whole data set instead of a 2D denoising on the integrated image preserves the sensitivity of the MSVST. The chief advantage of the 2D-1D-MSVST is the fact that we recover the energy information for each spatial position.

Fig. 4(a)-(b) display respectively a frame of the input noisy data set and the same frame after the 2D-1D-MSVST denoising. Fig. 5 displays the obtained spectra at the centers of two distinct sources. This result is rather impressive given that most of the counts in the 2D-1D data set are at most 1-2 photons (shown in white in 4(a)).

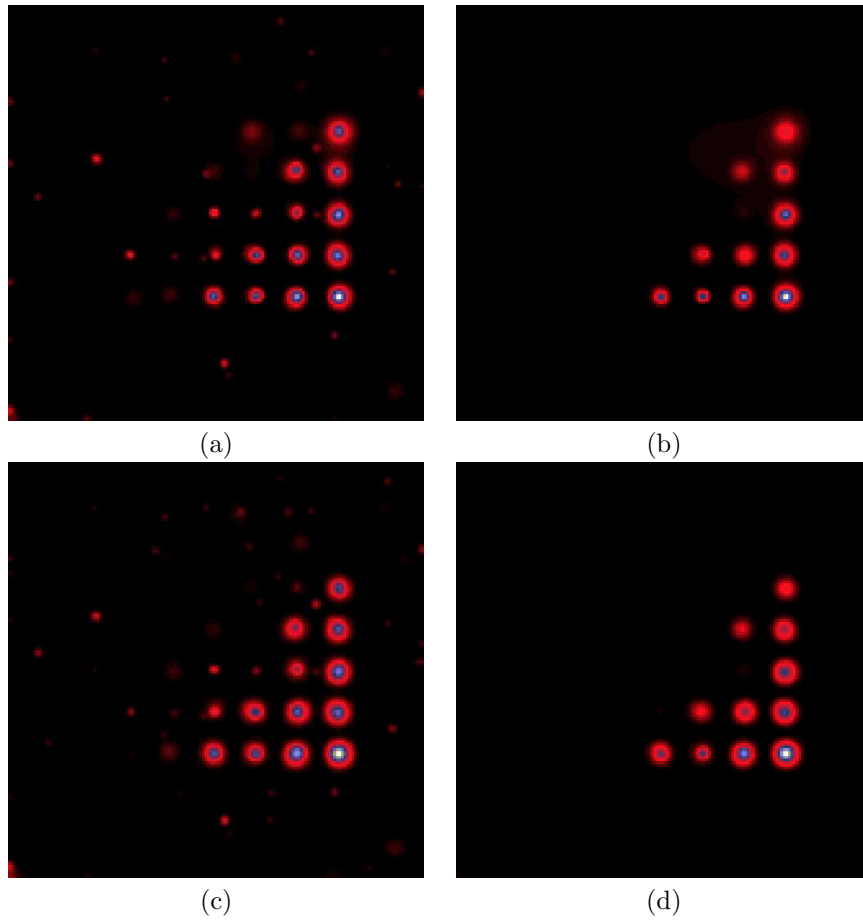


Figure 3. (a)-(b): 2D-MSVST filtering on the integrated image with critical thresholds of respectively $\tau = 3$ and $\tau = 5$. (c)-(d): image of frames summed along z -axis after a 2D-1D-MSVST denoising of the noisy data set, with respectively $\tau = 4$ and $\tau = 6$.

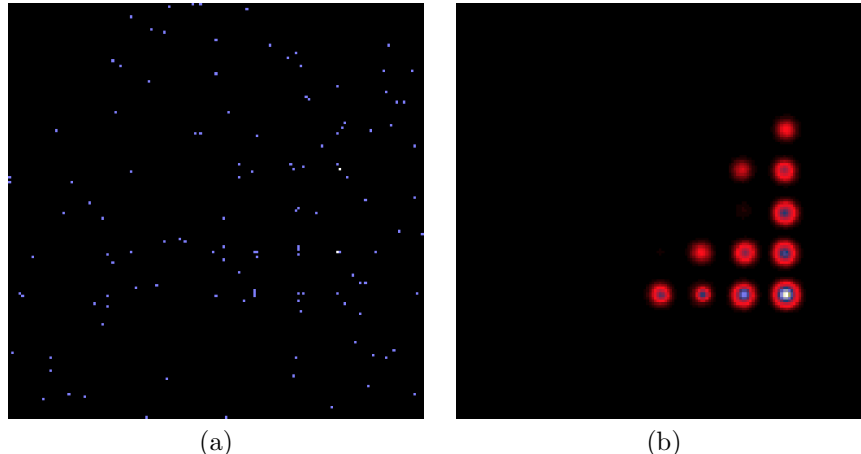


Figure 4. (a): a frame from the noisy data set and (b): the same frame after the 2D-1D-MSVST denoising at threshold $\tau = 6$.

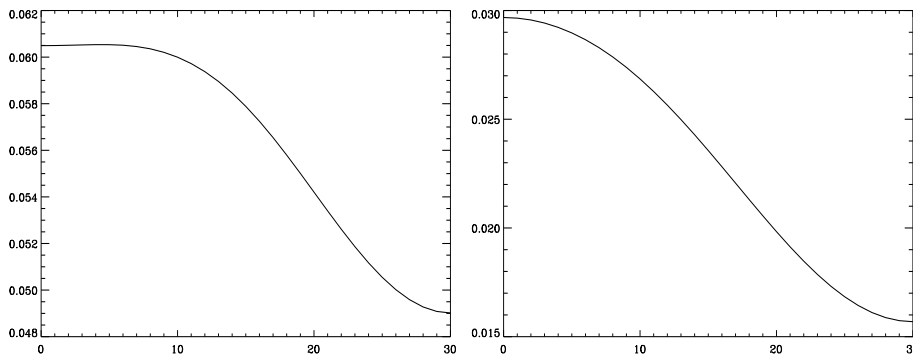


Figure 5. Pixel spectra at two different spatial locations after the 2D-1D-MSVST filtering.

4.2 Application to astronomical data

We applied our method to realistic simulated observations with the Large Area Telescope (LAT). The LAT was launched in June 2008 by NASA on the Fermi Gamma-ray Space Telescope mission. In this experiment, we generated a simulated $720 \times 360 \times 128$ data set using the Galprop code that has a model of the diffuse gamma-ray emission of the Milky Way. 128 energy planes are logarithmically spaced from 30 MeV to 50 GeV. A six months LAT data set was created by multiplying the simulated data set with the exposure (6 months), and by convolving each energy band with the point spread function of the LAT instrument. Finally we have created the noisy observations assuming a Poisson noise distribution. Fig. 6 shows from top to bottom the original simulated data, the noisy data and the denoised data for energy band 171-181 MeV.

5. CONCLUSION

In this paper, we propose a multivalued extension of MSVST associated with an appropriate 2D-1D wavelet transform, which proved very efficient to denoise Poisson count data. The proposed algorithm performs as well as the 2D-MSVST applied to summed frames. But unlike 2D denoising, 2D-1D extension fully exploits the information in the whole data set and recovers the information along the z -axis which is of paramount importance in many science fields such as hyperspectral imaging.

*Recall that the Gaussian noise after the 2D-1D-MSVST is assumed have a unit variance.

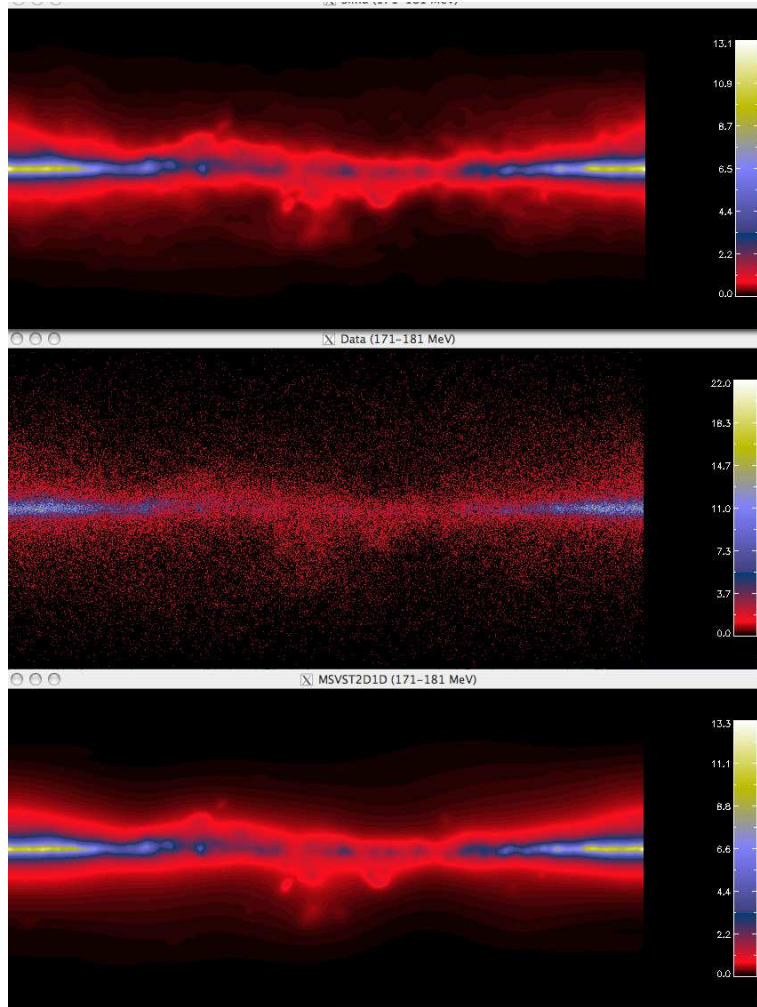


Figure 6. Top to bottom: simulated data of the diffuse gamma-ray emission of the Milky Way in energy band 171-181 Mev, noisy simulated data and denoised data using the 2D-1D-MSVST.

REFERENCES

- [1] Zhang, B., Fadili, J., and Starck, J.-L., “Wavelets, ridgelets and curvelets for poisson noise removal,” *IEEE Transactions on Image Processing* **17**(7), 1093–1108 (2008).
- [2] Anscombe, F., “The transformation of Poisson, binomial and negative-binomial data,” *Biometrika* **15**, 246–254 (1948).
- [3] Fryzlewicz, P. and Nason, G. P., “A Haar-Fisz algorithm for Poisson intensity estimation,” *J. Comp. Graph. Stat.* **13**, 621–638 (2004).
- [4] Fryzlewicz, P., Delouille, V., and Nason, G. P., “GOES-8 X-ray sensor variance stabilization using the multiscale data-driven Haar-Fisz transform,” *J. Roy. Statist. Soc. ser. C* **56**(1), 99–116 (2007).
- [5] Fisz, M., “The limiting distribution of a function of two independent random variables and its statistical application,” *Colloquium Mathematicum* **3**, 138–146 (1955).
- [6] Jansen, M., “Multiscale Poisson data smoothing,” *J. Roy. Statist. Soc. ser. B* **68**(1), 27–48 (2006).
- [7] Kolaczyk, E. D. and Dixon, D. D., “nonparametric estimation of intensity maps using Haar wavelets and Poisson noise characteristics,” *Astrophysical Journal* **534**, 490–505 (May 2000).
- [8] Zhang, B., Fadili, J., Starck, J.-L., and Digel, S. W., “Fast Poisson Noise Removal by Biorthogonal Haar Domain Hypothesis Testing,” *Statistical Methodology* **5**(4), 387–396 (2008).

- [9] Bijaoui, A. and Jammal, G., “On the distribution of the wavelet coefficient for a Poisson noise,” *Signal Processing* **81**, 1789–1800 (2001).
- [10] Kolaczyk, E. D., “Wavelet shrinkage estimation of certain Poisson intensity signals using corrected thresholds,” *Statist. Sinica* **9**, 119–135 (1999).
- [11] Kolaczyk, E. D., “Bayesian multiscale models for Poisson processes,” *J. Amer. Statist. Ass.* **94**, 920–933 (Sept. 1999).
- [12] Timmermann, K. E. and Nowak, R. D., “Multiscale Modeling and Estimation of Poisson Processes with Application to Photon-Limited Imaging,” *IEEE Trans. Inf. Theo.* **45**, 846–862 (Apr. 1999).
- [13] Nowak, R. D. and Kolaczyk, E. D., “A statistical multiscale framework for Poisson inverse problems,” *IEEE Transactions on Information Theory* **46**, 1811–1825 (Aug. 2000).
- [14] Nowak, R. D. and Baraniuk, R. G., “Wavelet-Domain Filtering for Photon Imaging Systems,” *IEEE Transactions on Image Processing* **8**, 666–678 (May 1999).
- [15] Antoniadis, A. and Sapatinas, T., “Wavelet shrinkage for natural exponential families with quadratic variance functions,” *Biometrika* **88**, 805–820 (2001).
- [16] Sardy, S., Antoniadis, A., and Tseng, P., “Automatic smoothing with wavelets for a wide class of distributions,” *J. Comput. Graph. Stat.* **13**, 399–421 (June 2004).
- [17] Willett, R. M. and Nowak, R. D., “Fast, Near-Optimal, Multiresolution Estimation of Poisson Signals and Images,” in [*EUSIPCO*], (2004).
- [18] Willett, R., “Multiscale Analysis of Photon-Limited Astronomical Images,” in [*SCMA IV*], (2006).
- [19] Fryźlewicz, P. and Nason, G. P., “A Haar-Fisz algorithm for Poisson intensity estimation,” *J. Comp. Graph. Stat.* **13**, 621–638 (2004).
- [20] Holschneider, M., Kronland-Martinet, R., Morlet, J., and Tchamitchian, P., “A Real-Time Algorithm for Signal Analysis with the Help of the Wavelet Transform,” in [*Wavelets: Time-Frequency Methods and Phase-Space*], 286–297, Springer-Verlag (1989).
- [21] Mallat, S., [*A Wavelet Tour of Signal Processing*], Academic Press (1998).
- [22] Starck, J.-L. and Murtagh, F., [*Astronomical Image and Data Analysis*], Astronomical image and data analysis, by J.-L. Starck and F. Murtagh. Astronomy and astrophysics library. Berlin: Springer, 2006 (2006).
- [23] Starck, J.-L., Fadili, M., and Murtagh, F., “The Undecimated Wavelet Decomposition and its Reconstruction,” *IEEE Transactions on Image Processing* **16**(2), 297–309 (2007).
- [24] Starck, J.-L., Murtagh, F., and Bijaoui, A., [*Image Processing and Data Analysis: The Multiscale Approach*], Cambridge University Press (1998).
- [25] Olivo-Marin, J. C., “Extraction of spots in biological images using multiscale products,” *Pattern Recognition* **35**(9), 1989–1996 (2002).
- [26] Yamada, I., “The hybrid steepest descent method for the variational inequality problem over the intersection of fixed point sets of nonexpansive mappings,” in [*Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications*], Butnariu, D., Censor, Y., and Reich, S., eds., Elsevier (2001).
- [27] Benjamini, Y. and Hochberg, Y., “Controlling the false discovery rate - a practical and powerful approach to multiple testing,” *J. R. Stat. Soc. B* **57**, 289 (1995).