

# Denoising of Multicomponent Images Using Wavelet Least-Squares Estimators<sup>1</sup>

Steve De Backer<sup>a,\*</sup>, Aleksandra Pižurica<sup>b,2</sup> Bruno Huysmans<sup>b</sup>  
Wilfried Philips<sup>b</sup> Paul Scheunders<sup>a</sup>

<sup>a</sup>*Visionlab, Department of Physics, Antwerp University, Universiteitsplein 1, B-2610 Wilrijk, Belgium.*

<sup>b</sup>*Department for Telecommunications and Information Processing (Telin) Ghent University, Sint-Pietersnieuwstraat 41, B-9000 Gent, Belgium.*

---

## Abstract

In this paper, we study denoising of multicomponent images. The presented procedures are spatial wavelet-based denoising techniques, based on Bayesian least-squares optimization procedures, using prior models for the wavelet coefficients that account for the correlations between the spectral bands. We analyze three mixture priors: Gaussian scale mixture models, Bernoulli-Gaussian mixture models and Laplacian mixture models. These three prior models are studied within the same framework of least-squares optimization. The presented procedures are compared to Gaussian prior model and single-band denoising procedures. We analyze the suppression of non-correlated as well as correlated white Gaussian noise on multispectral and hyperspectral remote sensing data and Rician distributed noise on multiple images of within-modality magnetic resonance data. It is shown that a superior denoising performance is obtained when a) the interband covariances are fully accounted for and b) prior models are used that better approximate the marginal distributions of the wavelet coefficients.

*Key words:* multicomponent images, denoising, wavelets, Bayesian estimation, least squares estimators.

---

\* Corresponding author. tel.: +32-3-8202464; fax.: +32-3-8202245

*Email addresses:* `steve.debacker@ua.ac.be` (Steve De Backer), `sanja@telin.ugent.be` (Aleksandra Pižurica), `bhuysman@telin.ugent.be` (Bruno Huysmans), `philips@telin.ugent.be` (Wilfried Philips), `paul.scheunders@ua.ac.be` (Paul Scheunders).

<sup>1</sup> This work was supported by the Research Foundation - Flanders (FWO).

<sup>2</sup> A. Pižurica is a postdoctoral research fellow of FWO Flanders.

## 1 Introduction

With the evolution of imaging technology, an increasing number of imaging modalities becomes available. In this paper, we focus on what is called multicomponent or multiband imagery, which are images containing several image planes. The most well-known example is a color image, which generally consists of a red, green and blue component. In remote sensing for example, sensors are available that can generate multispectral or hyperspectral data, involving five to more than hundred bands. In medical imagery, distinct image modalities reveal different features of the internal body. Examples are MRI images, acquired by using different imaging parameters (T1, T2, proton density, diffusion, ...), different CT and nuclear medicine imaging modalities.

In these images, a trade-off exists between spatial and spectral resolution and signal-to-noise ratio (SNR), which makes noise handling important. Also, noise filtering and image enhancement can drastically facilitate the processing and analysis of multicomponent imagery. More in particular, this holds for the classification and segmentation of multicomponent images for identification purposes [1, 2].

Multiband image noise is usually treated as stochastic Gaussian distributed, where the noise in the different bands is not necessarily independent. Generally, in remote sensing multispectral noise removal is achieved by a transform, referred to as minimum noise fraction (MNF) [3, 4]. MNF contains two principal component transformations. The first one diagonalizes the noise covariance, the second one decorrelates the noise-whitened data. MNF only uses spectral information for denoising. On the other hand, spatial smoothing of each spectral band separately is also a common practice in multispectral noise reduction. Recently, two-step approaches have been suggested, where spectral decorrelation of the noise is combined with band-wise denoising [5, 6].

Here, we will concentrate on spatial denoising methods, although denoising of hyperspectral data can also be performed in the spectral direction. In the latter case, consecutive bands are considered as densely sampled continuous spectra. However, not all multicomponent images have a sufficient number of spectral bands for these methods to apply. Here, we are not assuming continuous spectra, but merely that there is some correlation between the bands. The presented methods are invariant to permutation of the bands, which would not be the case for denoising in the spectral direction.

With spatial denoising, one makes use of the uniformity in a spatially local neighborhood of a pixel. Selection of this neighborhood is dependent on the content of the image, to prevent it from smoothing over edges. Wavelet-based methods are the current state-of-the-art for spatial denoising. The wavelet

transform offers an efficient representation of spatial discontinuities within each spectral band [7–9]. It compresses the essential information of an image into a relatively few, large coefficients coinciding with the positions of image discontinuities. Such a representation naturally facilitates the construction of *spatially adaptive* denoising methods that can smooth noise without excessive blurring of image details. Typically, noise is reduced by *shrinking* the noisy wavelet coefficient magnitudes. Ideally, the coefficients that contain primarily noise (usually the smallest coefficients) are reduced to negligible values while the ones containing a “significant” noise-free component are reduced less [10]. Standard wavelet thresholding [10] treats the coefficients with magnitudes below a certain threshold as “non significant” and sets these to zero; the remaining, “significant” coefficients are kept unmodified (hard-thresholding) or reduced in magnitude (soft-thresholding). Shrinkage estimators can also result from a *Bayesian* approach [11–29], which imposes a prior distribution on noise-free data. Common priors for noise-free data include (generalized) Laplacians [11, 18, 21], alpha-stable models [20], double stochastic (Gaussian scale mixture) models [24, 25] and mixtures of two distributions [13–17] where one distribution models the statistics of “significant” coefficients and the other one models the statistics of “insignificant” data.

In this work, we will give a unifying wavelet-based Bayesian framework for multicomponent image denoising. While a few different Bayesian multicomponent methods existed before [30–35], we provide a general framework where all these methods fit as special cases with different prior models. This provides not only a better insight about these methods, but it also enables a fair comparison between these different models within the same framework. For this, the different compared techniques are all presented with a particular prior distribution, allowing for comparison of effectiveness of the suggested prior image model.

Recently, several wavelet based procedures for multicomponent images were proposed that account to some extent for the inter-component correlations, applying wavelet thresholding [36] or Bayesian estimation, using different prior models [30–35]. The purpose of this work is to present these in a Bayesian framework for wavelet-based denoising of multicomponent images to a) fully account for the inter-component covariances, and b) use different prior models that optimally approximate the marginal densities of the wavelet coefficients. Within this framework, 4 different prior models are reviewed: the Bernoulli-Gaussian model from the literature [30], and the Gaussian [31], the Gaussian Scale Mixture [32] and the Generalized Gaussian models [35], that have been investigated by the authors. The denoising performance of these techniques is compared to that of standard spectral denoising and spatial single-band techniques that do not or only partly account for the inter-component correlations. The methods are compared within two applications with different noise types: (1) remote sensing multispectral and hyperspectral data with non-correlated

as well as correlated white Gaussian noise and (2) magnetic resonance within-modality data with Rician distributed noise.

The outline of this paper is as follows: in the next section, we introduce the framework for spatial wavelet-based denoising on multicomponent images. The wavelet representation, the imaging model and the Bayesian estimation procedure are elaborated. In section 3, least-squares denoising procedures using 3 different priors are presented: the Gaussian Scale Mixture model, the Bernoulli-Gaussian model and the Laplacian mixture model. In section 4, the experiments are conducted and a discussion on the results follows.

## 2 The Bayesian multicomponent image denoising framework

For spatial denoising of multicomponent images, making use of inter-band correlations is essential. An image discontinuity (e.g., an edge, corner, line, point target,...) appearing in one image band is likely to appear in at least some of the remaining bands. Proper use of inter-band correlations facilitates discrimination between noise and image features, and even reveals image details that were “hidden” by noise in given bands or discards false structures generated by noise.

In general, designing a multicomponent denoiser involves making the following choices: the *image representation* (e.g., single-resolution or multiresolution); handling the *multicomponent character* of the data (e.g., vector-based processing or band-wise processing) and the *estimation approach* (e.g., choosing an optimization criterion). In the following subsections, we address each of these items and link them to the denoising framework presented in this paper.

### 2.1 Non-decimated wavelet transform

The wavelet transform reorganizes image content into a low-resolution approximation and a set of details of different orientations and different resolution scales. A fast algorithm for the discrete wavelet transform is an *iterative filter bank algorithm* of Mallat [8], where a pair of high-pass and low-pass filters followed by downsampling by two is iterated on the low-pass output. In a *non-decimated* wavelet transform that we consider in this paper, downsampling is excluded, and instead the filters are upsampled at each decomposition stage as explained later in the text. The outputs of the lowpass filter are the *scaling* coefficients and the outputs of the high-pass filter are the *wavelet* coefficients. At each decomposition level, the filter bank is applied sequentially to the rows and to the columns of the image. Low-pass filtering of both the rows and the

columns yields the low-pass LL subband and other combinations of low-pass and high-pass filtering yield the wavelet subbands at different orientations: High-pass filtering of rows and low-pass filtering of columns (HL) yields horizontal edges and the opposite combination (LH) yields vertical edges, while high-pass filtering of both the rows and the columns (HH) yields diagonal edges. The  $j$ th decomposition level yields the coefficients at the *resolution scale*  $2^j$ .

In image denoising, redundant wavelet transforms, like the non-decimated transform yield better results than the critically sampled one. A common alternative is using a critically sampled representation with *cycle spinning* [37] (denoising multiple cyclical shifts of the image and averaging over unshifted results) which yields a similar improvement over the critically-sampled case [11].

In this paper we use a non-decimated wavelet transform implemented with the algorithm *à trous* [38]. The algorithm inserts  $2^j - 1$  zeroes (i.e., “holes”, French *trous*) between the filter coefficients at the resolution level  $j$ . The size of each wavelet subband equals the size of the input image.

For compactness, denote the spatial position vector  $[m, n]$  by a single index  $l$ , and denote the scaling coefficients at the resolution level  $j+1$  by  $a_l^{(j+1)}$  and the wavelet coefficients at the corresponding scale in three orientation subbands by  $x_l^{(j+1,HL)}$ ,  $x_l^{(j+1,LH)}$  and  $x_l^{(j+1,HH)}$ . Let  $\mathbf{h}$  and  $\mathbf{g}$  denote, respectively, the low-pass and the high-pass filters associated with the wavelet function  $\psi$ , and let  $\bar{h}$  denote the conjugate complex of  $h$ , and  $\mathbf{h}^j$  an up-sampled filter, where  $2^j - 1$  zeroes are inserted between each two coefficients of  $\mathbf{h}$ . Denoting the discrete convolution by  $\star$ , the non-decimated wavelet decomposition is formally given by:

$$\begin{aligned} a_l^{(j+1)} &= \bar{h}^j \bar{h}^j_l \star a_l^{(j)} \\ x_l^{(j+1,HL)} &= \bar{g}^j \bar{h}^j_l \star a_l^{(j)} \\ x_l^{(j+1,LH)} &= \bar{h}^j \bar{g}^j_l \star a_l^{(j)} \\ x_l^{(j+1,HH)} &= \bar{g}^j \bar{g}^j_l \star a_l^{(j)}. \end{aligned} \tag{1}$$

Decomposing an image into  $J$  decomposition levels yields a wavelet image representation consisting of  $3J + 1$  subbands:  $[\mathbf{a}^{(J)}, \{\mathbf{x}^{j,HL}, \mathbf{x}^{j,LH}, \mathbf{x}^{j,HH}\}_{1 \leq j \leq J}]$ :

$$\begin{aligned} a_l^{(0)} &= h^J h_l^J \star a_l^{(J+1)} + \sum_{j=1}^J \left( g^j h_l^j \star x_l^{(j+1,HL)} + h^j g_l^j \star x_l^{(j+1,LH)} \right. \\ &\quad \left. + g^j g_l^j \star x_l^{(j+1,HH)} \right) \end{aligned} \tag{2}$$

where the input image is approximated by  $\mathbf{a}^{(0)}$  with a negligible error [38].

## 2.2 Wavelet processing of multicomponent data

A natural way of exploiting the interband correlations is by vector-based processing, operating on all the bands simultaneously. Let  $s_l^{(j,o,b)}$  denote the noise-free wavelet coefficient at *spatial* position  $l$ , *resolution* level  $j$ , *orientation* subband  $o$ , and *image band*  $b$ .  $x_l^{(j,o,b)}$  and  $n_l^{(j,o,b)}$  are the corresponding wavelet coefficients of the observed noisy image and the noise, respectively. A vector processing approach groups the wavelet coefficients  $x_l^{(j,o,b)}$  of all the  $B$  bands at a given spatial position, within a subband of a given orientation and resolution level into a  $B$ -dimensional vector:

$$\mathbf{x}_l^{(j,o)} = [x_l^{(j,o,1)}, \dots, x_l^{(j,o,B)}]^T. \quad (3)$$

Equivalent processing is typically applied to all the wavelet subbands, and hence we shall omit the indexes that denote the resolution level  $j$  and orientation  $o$ . In each wavelet subband, a multicomponent pixel obeys the additive noise model:

$$\mathbf{x} = \mathbf{s} + \mathbf{n} \quad (4)$$

where the probability density function (i.e., *density*) of  $\mathbf{n}$  is assumed to be a multivariate Gaussian of zero mean and covariance matrix  $\mathbf{C}_n$ :  $p(\mathbf{n}) = \phi(\mathbf{n}; \mathbf{C}_n)$ .  $\mathbf{C}_x$  denotes the covariance matrix of the noisy vector  $\mathbf{x}$  and  $\hat{\mathbf{C}}_s$  the estimate of the covariance matrix of the unknown noise-free vector  $\mathbf{s}$ .

The noise covariance in each wavelet subband is in general a scaled version of the input image noise covariance, where the scaling factors depend on the wavelet filter coefficients (see, e.g., [39]). With the orthogonal wavelet families [38], that we use in this paper, the noise covariance in all the wavelet subbands is equal to the input image noise covariance. In most cases, we will assume that the input image noise covariance is known. When conducting experiments with real noisy data, the noise covariance will be estimated separately. The signal covariance matrix is estimated as

$$\hat{\mathbf{C}}_s = \mathbf{C}_x - \mathbf{C}_n. \quad (5)$$

Since  $\hat{\mathbf{C}}_s$  is a covariance matrix, it needs to be semi-positive definite. This is assured by performing an eigenvalue decomposition and clipping possible negative eigenvalues to zero.

Throughout the whole paper,  $\mathbf{C}_x$  is determined over the entire image:  $\mathbf{C}_x = \langle \mathbf{x}_l \mathbf{x}_l^T \rangle_l$ , i.e., it is assumed that the interband correlations are the same for all wavelet coefficients of a subband.

### 2.3 Estimation approach and optimization criterion

Various linear and nonlinear (adaptive) methods can be applied for data denoising. We focus on the *Bayesian* approach, where a priori knowledge about the distribution of the noise-free data is assumed. In particular, we impose a multicomponent prior distribution (to be called hereafter *prior*) on the noise-free wavelet coefficients in a given subband and we differentiate between several approaches based on different specific priors. As an optimization criterion, we adopt *minimization of the mean squared error*, i.e., the Bayesian risk is a quadratic loss function. Estimation that uses this optimization criterion is referred to as least squares estimation.

The minimum mean squared error (MMSE) estimate is the posterior conditional mean

$$\mathbb{E}(\mathbf{s}|\mathbf{x}) = \int_{-\infty}^{\infty} \mathbf{s}p(\mathbf{s}|\mathbf{x})d\mathbf{x} = \frac{\int_{-\infty}^{\infty} \mathbf{s}p(\mathbf{x}|\mathbf{s})p(\mathbf{s})d\mathbf{s}}{\int_{-\infty}^{\infty} p(\mathbf{x}|\mathbf{s})p(\mathbf{s})d\mathbf{s}} = \frac{\int_{-\infty}^{\infty} \mathbf{s}\phi(\mathbf{x} - \mathbf{s}; \mathbf{C}_n)p(\mathbf{s})d\mathbf{s}}{\int_{-\infty}^{\infty} \phi(\mathbf{x} - \mathbf{s}; \mathbf{C}_n)p(\mathbf{s})d\mathbf{s}}. \quad (6)$$

Assuming e.g. a Gaussian prior for the noise-free signal  $p(\mathbf{s}) = \phi(\mathbf{s}; \mathbf{C}_s)$ , the above MMSE estimate becomes the Wiener filter:

$$\hat{\mathbf{s}} = \mathbb{E}(\mathbf{s}|\mathbf{x}) = \hat{\mathbf{C}}_s(\hat{\mathbf{C}}_s + \mathbf{C}_n)^{-1}\mathbf{x}. \quad (7)$$

## 3 Least squares denoisers using different priors

The standard Wiener result of the previous section is the result obtained using a multicomponent Gaussian prior model. It accounts for the multicomponent covariance, but it assumes that the marginal densities for the wavelet coefficients are Gaussian. It is well-known that this assumption is not justified, and that these marginals are symmetric and zero mean, but heavier tailed than Gaussians. Different other priors were proposed to better approximate the marginal densities. In this section we present three different multicomponent mixture priors and apply them within the MMSE estimation framework of the previous section.

### 3.1 The GSM model

The Gaussian scale mixture prior [25] models the probability density function  $p(\mathbf{s})$  by a mixture of Gaussians:

$$p(\mathbf{s}) = \int p(\mathbf{s}|z)p(z)dz \quad (8)$$

where  $p(z)$  is the mixing density, and  $p(\mathbf{s}|z)$  is a zero mean Gaussian with covariance  $\mathbf{C}_{\mathbf{s}|z}$ . Under the GSM model, the additive noise model (4) becomes:

$$\mathbf{x} = \mathbf{s} + \mathbf{n} = \sqrt{z}\mathbf{u} + \mathbf{n} \quad (9)$$

where both  $\mathbf{u}$  and  $\mathbf{n}$  are zero-mean Gaussians, with covariances given by  $\mathbf{C}_{\mathbf{u}}$  and  $\mathbf{C}_{\mathbf{n}}$  respectively. Then,  $\mathbf{C}_{\mathbf{s}|z} = z\mathbf{C}_{\mathbf{u}}$  or, by taking expectations over  $z$ , with  $\mathbb{E}(z) = 1$ :  $\mathbf{C}_{\mathbf{s}} = \mathbf{C}_{\mathbf{u}}$ .

GSM densities are symmetric and zero-mean and heavier tailed than Gaussians. These are known to better model the shape of the wavelet coefficient marginals than Gaussians. In [25], GSM's were applied to model local spatial neighborhoods of wavelet coefficients in greylevel images. In this work, we apply the GSM to model multicomponent wavelet coefficients. In this way, the prior fully accounts for the inter-band covariances.

The Bayes least squares estimate  $\mathbb{E}(\mathbf{s}|\mathbf{x})$  is given by:

$$\begin{aligned} \mathbb{E}(\mathbf{s}|\mathbf{x}) &= \int \mathbf{s}p(\mathbf{s}|\mathbf{x})d\mathbf{s} \\ &= \int \int_0^\infty \mathbf{s}p(\mathbf{s}, z|\mathbf{x})dzd\mathbf{s} \\ &= \int \int_0^\infty \mathbf{s}p(\mathbf{s}|\mathbf{x}, z)p(z|\mathbf{x})dzd\mathbf{s} \\ &= \int_0^\infty p(z|\mathbf{x})\mathbb{E}(\mathbf{s}|\mathbf{x}, z)dz. \end{aligned} \quad (10)$$

Since, using the GSM model  $\mathbf{s}$ , conditioned on  $z$  is Gaussian, the expected value within the integral is given by a Wiener estimate:

$$\mathbb{E}(\mathbf{s}|\mathbf{x}, z) = z\mathbf{C}_{\mathbf{u}}(z\mathbf{C}_{\mathbf{u}} + \mathbf{C}_{\mathbf{n}})^{-1}\mathbf{x}. \quad (11)$$

The posterior distribution of  $z$  can be obtained, using Bayes' rule:

$$p(z|\mathbf{x}) = \frac{p(\mathbf{x}|z)p(z)}{\int_0^\infty p(\mathbf{x}|\alpha)p(\alpha)d\alpha} \quad (12)$$

with  $p(\mathbf{x}|z) = \phi(\mathbf{x}; z\mathbf{C}_{\mathbf{u}} + \mathbf{C}_{\mathbf{n}})$ . In [25], the authors motivate the use of the so-called *Jeffrey's prior* [40] for the random multiplier  $z$ :  $p(z) \propto \frac{1}{z}$ . We also refer to [25] for further information about the practical implementation of the determination of (10).

### 3.2 Mixtures of two distributions

Another often used mixture prior in wavelet denoising is a mixture of two distributions, where one distribution models the "significant" (large) coeffi-

coefficients and the other distribution models “insignificant” (small) coefficients [12, 14, 16, 17, 19, 30, 41]. Examples are the mixture of two Gaussians [16], the mixture of a Gaussian and a point mass at zero [12, 14, 17, 30] and the mixture of a Laplacian and a point mass at zero [19, 41]. The mixing between the two distributions is controlled by a *Bernoulli* random variable  $q$  with  $P(q = 1) = \epsilon = 1 - P(q = 0)$ .

In general, under this mixture prior the probability density of the vector  $\mathbf{s}$  is

$$p(\mathbf{s}) = (1 - \epsilon)p(\mathbf{s}|q = 0) + \epsilon p(\mathbf{s}|q = 1) \quad (13)$$

and the substitution in the minimum mean squared error estimator yields

$$\mathbb{E}(\mathbf{s}|\mathbf{x}) = \mathbb{E}(\mathbf{s}|\mathbf{x}, q = 0)P(q = 0|\mathbf{x}) + \mathbb{E}(\mathbf{s}|\mathbf{x}, q = 1)P(q = 1|\mathbf{x}). \quad (14)$$

From the Bayes’ rule

$$P(q = 1|\mathbf{x}) = \frac{\epsilon p(\mathbf{x}|q = 1)}{(1 - \epsilon)p(\mathbf{x}|q = 0) + \epsilon p(\mathbf{x}|q = 1)} \quad (15)$$

where

$$\begin{aligned} p(\mathbf{x}|q = 0) &= \int \phi(\mathbf{x} - \mathbf{s}; \mathbf{C}_n)p(\mathbf{s}|q = 0)d\mathbf{s} \\ p(\mathbf{x}|q = 1) &= \int \phi(\mathbf{x} - \mathbf{s}; \mathbf{C}_n)p(\mathbf{s}|q = 1)d\mathbf{s} \end{aligned} \quad (16)$$

and the conditional means are

$$\begin{aligned} \mathbb{E}(\mathbf{s}|\mathbf{x}, q = 0) &= \frac{\int_{-\infty}^{\infty} \mathbf{s}\phi(\mathbf{x} - \mathbf{s}; \mathbf{C}_n)p(\mathbf{s}|q = 0)d\mathbf{s}}{\int_{-\infty}^{\infty} \phi(\mathbf{x} - \mathbf{s}; \mathbf{C}_n)p(\mathbf{s}|q = 0)d\mathbf{s}} \\ \mathbb{E}(\mathbf{s}|\mathbf{x}, q = 1) &= \frac{\int_{-\infty}^{\infty} \mathbf{s}\phi(\mathbf{x} - \mathbf{s}; \mathbf{C}_n)p(\mathbf{s}|q = 1)d\mathbf{s}}{\int_{-\infty}^{\infty} \phi(\mathbf{x} - \mathbf{s}; \mathbf{C}_n)p(\mathbf{s}|q = 1)d\mathbf{s}}. \end{aligned} \quad (17)$$

### 3.2.1 The Bernoulli-Gaussian prior model

Recently, in [30] a multicomponent Bernoulli-Gaussian prior model was proposed:

$$\begin{aligned} p(\mathbf{s}|q = 0) &= \delta(\mathbf{s}) \\ p(\mathbf{s}|q = 1) &= \phi(\mathbf{s}; \mathbf{C}_s). \end{aligned} \quad (18)$$

To our knowledge, this paper was the first work from the literature where a multicomponent prior was applied for spatial multicomponent image denoising. For this reason, we will adopt this prior in our work.

For this prior:  $p(\mathbf{x}|q = 0) = \phi(\mathbf{x}; \mathbf{C}_n)$  and  $p(\mathbf{x}|q = 1) = \phi(\mathbf{x}; \mathbf{C}_s + \mathbf{C}_n)$ . Besides, substituting (18) into (17) yields  $\mathbb{E}(\mathbf{s}|\mathbf{x}, q = 0) = 0$  and  $\mathbb{E}(\mathbf{s}|\mathbf{x}, q = 1) = \hat{\mathbf{C}}_s(\mathbf{C}_n + \hat{\mathbf{C}}_s)^{-1}\mathbf{x}$  and hence

$$\mathbb{E}(\mathbf{s}|\mathbf{x}) = \frac{\epsilon\phi(\mathbf{x}; \mathbf{C}_s + \mathbf{C}_n)}{(1 - \epsilon)\phi(\mathbf{x}; \mathbf{C}_n) + \epsilon\phi(\mathbf{x}; \mathbf{C}_s + \mathbf{C}_n)}\hat{\mathbf{C}}_s(\mathbf{C}_n + \hat{\mathbf{C}}_s)^{-1}\mathbf{x}. \quad (19)$$

In [30], three estimation procedures were described: a Maximum A Posteriori (MAP) technique, an MMSE approach and an extended SURE (ESURE) approach [42]. Using the MAP approach, the hidden variable is estimated for each coefficient, where it takes the value 1, leading to the Wiener result, or zero. For this estimation, the improvements over the Gaussian prior model seemed to be marginal. Using the MMSE approach,  $\epsilon$  is estimated as a continuous value using absolute moments. In the ESURE approach,  $\epsilon$  and  $\mathbf{C}_s$  will be free parameters, which are estimated by minimizing the risk. We will adopt the latter 2 approaches in our work.

### 3.2.2 The Laplacian mixture model

Another prior is a mixture of two truncated generalized Laplacian distributions [35], where a *Bernoulli* random variable controls switching between the central part of the distribution (describing the “insignificant” data) and its tails (describing the “signal of interest”). Related priors were used, e.g. in [27, 28] where the signal of interest is defined as a noise-free wavelet coefficient component that exceeds the noise standard deviation. Compared to the Bernoulli-Gaussian, this prior models more realistically the subband statistics, but is also more complex and no multicomponent version has been studied yet. We will adopt the band-wise procedure from [35] and describe it in the Bayesian multicomponent image denoising framework.

For noise-free images, marginal subband statistics is modeled by a generalized Laplacian (also called generalized Gaussian) density [8, 11, 18, 21]:

$$p(s) = \frac{\lambda\nu}{2\Gamma(\frac{1}{\nu})} \exp(-(\lambda|s|)^\nu) \quad (20)$$

where  $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$  is the Gamma function,  $\lambda > 0$  is the *scale* parameter and  $\nu$  is the *shape* parameter, which is typically  $\nu \in [0, 1]$  for natural images. The marginal conditional densities are:

$$p(s|q = 0) = \begin{cases} B_0 \exp(-(\lambda|s|)^\nu), & \text{if } |s| \leq \sigma \\ 0, & \text{if } |s| > \sigma \end{cases} \quad (21)$$

and

$$p(s|q=1) = \begin{cases} 0, & \text{if } |s| \leq \sigma \\ B_1 \exp(-(\lambda|s|)^\nu), & \text{if } |s| > \sigma \end{cases} \quad (22)$$

where  $\sigma$  is the noise standard deviation in the corresponding wavelet subband, and the normalization constants are [35] :  $B_0 = \lambda\nu \left( 2\Gamma(\frac{1}{\nu})\Gamma_{inc}\left((\lambda T)^\nu, \frac{1}{\nu}\right) \right)^{-1}$  and  $B_1 = \lambda\nu \left( 2\Gamma(\frac{1}{\nu}) \left[ 1 - \Gamma_{inc}\left((\lambda T)^\nu, \frac{1}{\nu}\right) \right] \right)^{-1}$ , where  $\Gamma_{inc}(x, a) = \frac{1}{\Gamma(a)} \int_0^x t^{a-1} e^{-t} dt$  is the *incomplete gamma function*.

To construct a multicomponent extension of the generalized Laplacian mixture prior, will adopt a simple solution from [35]. The estimator is applied *band-wise* and the correlation between the bands is taken into account by calculating a *band activity indicator* (BAI) from the different image bands. Define the BAI,  $z$ , at a given spatial position in a given wavelet subband as the wavelet coefficient magnitudes, averaged over all  $B$  image bands:

$$z = \frac{1}{B} \sum_{b=1}^B |x^{(b)}|. \quad (23)$$

The band-wise *ProbShrink* estimator, for band  $b$  (we omit the index  $b$ ) is given by:

$$\hat{s} = P(q=1|x, z)x = \frac{\eta(x)\xi(z)\mu}{1 + \eta(x)\xi(z)\mu}x \quad (24)$$

where

$$\eta(x) = \frac{p(x|q=1)}{p(x|q=0)}, \quad \xi(z) = \frac{p(z|q=1)}{p(z|q=0)} \quad \text{and} \quad \mu = \frac{\epsilon}{1 - \epsilon}. \quad (25)$$

The local MMSE estimator is  $\mathbb{E}(s|x, z) = P(q=1|x, z)\mathbb{E}(s|x, q=1) + P(q=0|x, z)\mathbb{E}(s|x, q=0)$ , but under the assumed mixture prior, the second term is close to zero and  $\mathbb{E}(s|x, q=1)$  is close to  $x$ . The prior probability is in [35] derived as  $P(q=1) = \epsilon = 1 - \int_{-\sigma}^{\sigma} p(s)ds$ , which for the generalized Laplacian prior yields  $\epsilon = 1 - \Gamma_{inc}\left((\lambda\sigma)^\nu, \frac{1}{\nu}\right)$ . The conditional densities  $p(x|q)$  ( $q=0, 1$ ) are derived from  $p(s|q)$  using a scalar version of (16). From these and (23),  $p(z|q)$  are calculated.

## 4 Experiments

In order to validate the proposed multicomponent noise reduction framework, we performed experiments on multicomponent data sets with simulated and real noise. In a first experiment, we simulated uncorrelated as well as correlated additive Gaussian noise on a Landsat multispectral image. In the second

experiment, real noisy data were used. For this, radiance data from an AVIRIS hyperspectral image in the near infra-red part of the spectrum was taken. The image contains spectral bands near the absorption window of atmospheric water vapor and carbon dioxide, where the signal to noise ratio tends to be low. In the second group of experiments, simulated MRI datasets were used; one containing images acquired with different MRI acquisition schemes and another containing diffusion tensor imaging data with 6 components.

For the wavelet decomposition, a four-level decomposition ( $J = 4$ ) of the non-decimated wavelet transform, described in Sec.2.1, is used. For the wavelet filters  $\mathbf{h}$  and  $\mathbf{g}$ , we use the Daubechies symmlets family [7]. Experiments show that these filters generally rendered the best results.

The proposed denoising techniques are compared to each other and to other denoising procedures, in the real and wavelet domain. In the experiments the following denoising procedures are compared:

- For the remote sensing experiments, the Minimum Noise Fraction (*MNF*) transform [3] is applied for comparison. MNF denoises the imagery by decorrelation of the noise followed by decorrelation of the spectral signal. By retaining only the decorrelated bands with highest signal to noise ratio, it is assumed that the noise is contained in the low SNR bands. By retaining the first  $N$  bands and then making the reconstruction, a denoised image is obtained. In the simulated noise case, we optimized the retained number of bands by maximizing the peak signal-to-noise ratio (PSNR).
- A first wavelet-based method applies denoising separately on each band  $b$ . Minimum mean squared error detail coefficient shrinkage is applied on each wavelet scale and orientation. Here, all detail coefficients are rescaled by

$$\hat{s}^b = \frac{\sigma_{x^b}^2 - \sigma_n^2}{\sigma_{x^b}^2} x^b$$

the ratio of signal variance over noise variance. It is clear that this *band-wise* procedure does not take any band correlation into account, not in the noise, nor in the signal.

- An extension of the band-wise approach is a *hybrid Gaussian* approach. Before employing the band-wise procedure, a forward MNF transform is applied. After band-wise denoising, the data is transformed back. In this procedure, interband correlations are taken into account by the use of MNF, which assures the decorrelation of the transformed bands.
- Then, the multiband MMSE denoising techniques that are proposed in this work are applied. First, a multidimensional *Gaussian* prior is assumed to describe the signal correlation between the bands. As a result, for each wavelet subband the vector shrinkage is given by (7), the vector Wiener filter.
- Next, the Gaussian scale mixture (*GSM*) prior is assumed, as presented in

section 3.1.

- Further on, we implemented the MMSE denoising, assuming the Bernoulli-Gaussian signal prior (noted by *BG-Moments*) from [30], and as described in section 3.2.1.
- Also, using the same Bernoulli-Gaussian prior, the extended E-sure (*BG-Sure*) from [30] is applied.
- Then, we apply *LM Probshrink*, using the Laplacian mixture prior model, as described in section 3.2.2. This approach assumes uncorrelated noise of equal intensity. Therefore, when applying this technique the noise covariance is first whitened, such that the noise covariance matrix becomes the unity matrix. After denoising, the data is then rotated back. As a matter of fact, using this prior, the interband correlations are only partly taken into account, and as such, this procedure should be regarded as a hybrid band-wise approach.

The performance of the different denoising techniques is quantitatively described by the PSNR (in dB), defined for 8 bit images of size  $N^2$ , with  $B$  bands as:

$$\begin{aligned} \text{PSNR}(\hat{\mathbf{S}}, \mathbf{S}) &= 10 \log \left( \frac{255^2}{\text{mse}(\hat{\mathbf{S}}, \mathbf{S})} \right) \\ \text{mse}(\hat{\mathbf{S}}, \mathbf{S}) &= \frac{1}{BN^2} \sum_{i,j=1}^N \sum_{b=1}^B (\hat{S}_{i,j}^{(b)} - S_{i,j}^{(b)})^2 \end{aligned} \quad (26)$$

#### 4.1 Experiment 1: Landsat dataset with simulated noise

In the first experiment, a Landsat 7 TM 7-band image acquired over the Winnipeg area, containing both urban and rural parts, was taken. All bands except for the thermal band (band 6) were used, obtaining a 6 band 8-bit image. On this image, additive Gaussian noise was simulated. We simulated both correlated and uncorrelated noise. In the uncorrelated case the covariance was  $\mathbf{C}_n = \sigma^2 \mathbf{I}$ , with  $\mathbf{I}$  the unity matrix. In the correlated case, the covariance was set to  $\mathbf{C}_n = \sigma^2 \mathbf{C}$ , with all values of  $\mathbf{C}$  set to .5, except for the diagonal elements, which were set to 1.  $\sigma$  was set to the values 5, 10, 15, 20 and 25. In this experiment,  $\mathbf{C}_n$  is assumed to be known. Here, we use noise variances which are rather high compared to the expected values in remote sensing. These values were taken arbitrarily to evaluate the denoising procedures.

Table 1 shows the results for the uncorrelated noise case. The results obtained for the image corrupted with correlated noise is shown in table 2. Figure 1 shows a detail of the denoising result for the red band, after corrupting the original with uncorrelated noise with  $\sigma = 15$ . Based on the obtained PSNR

<b>initial</b>	34.21	28.24	24.77	22.32	20.44
<b>MNF</b>	34.21	28.71	25.88	24.13	22.59
WAVELET BAND-WISE					
<b>Gaussian</b>	34.87	30.00	27.54	25.98	24.84
WAVELET HYBRID BAND-WISE					
<b>hybrid Gaussian</b>	34.87	31.28	29.07	27.60	26.50
<b>LM Probshrink</b>	35.23	30.85	28.60	27.11	25.95
WAVELET MULTICOMPONENT					
<b>Gaussian</b>	35.66	31.38	29.16	27.69	26.56
<b>GSM</b>	<b>37.70</b>	31.60	29.43	27.96	<b>26.82</b>
<b>BG-Moments</b>	35.81	<b>31.70</b>	<b>29.53</b>	<b>28.06</b>	26.50
<b>BG-Sure</b>	35.80	31.66	29.47	27.93	26.68

Table 1

PSNR values for Landsat experiments with simulated uncorrelated noise (in dB). The 5 columns show the results for values of  $\sigma$  given by 5, 10, 15, 20 and 25 respectively.

<b>initial</b>	34.22	28.24	24.77	22.32	20.43
<b>MNF</b>	34.22	28.24	25.04	22.81	21.14
WAVELET BAND-WISE					
<b>Gaussian</b>	34.87	29.99	27.54	25.95	24.83
WAVELET HYBRID BAND-WISE					
<b>hybrid Gaussian</b>	35.01	30.35	28.06	26.55	25.47
<b>LM Probshrink</b>	35.09	30.48	28.15	26.66	25.59
WAVELET MULTICOMPONENT					
<b>Gaussian</b>	35.15	30.52	28.20	26.66	25.55
<b>GSM</b>	<b>35.43</b>	30.91	28.57	27.00	25.85
<b>BG-Moments</b>	35.39	<b>30.93</b>	<b>28.65</b>	<b>27.10</b>	25.95
<b>BG-Sure</b>	35.38	30.90	28.60	27.05	<b>25.98</b>

Table 2

PSNR values for Landsat experiments with simulated correlated noise (in dB). The 5 columns show the results for values of  $\sigma$  given by 5, 10, 15, 20 and 25 respectively.

values of both experiments we can conclude the following:

- All wavelet-based approaches perform better than the traditional MNF transformation. The main reason for this is that the MNF transformation only uses spectral information. The wavelet based techniques also take spatial correlations into account. It has been demonstrated many times that the spatial decorrelation and the good energy concentration properties makes the wavelet domain suited for image and signal denoising.

- When comparing the multiband approaches to the band-wise approach, we can conclude that the use of information on the spectral band dependency of both noise and signal improves the denoising results. The hybrid approaches obtain a substantial improvement over the band-wise approach in the case of uncorrelated noise. A more modest improvement is obtained for correlated noise. Remark that treating LM Probshrink as a hybrid approach proves to be correct by these results.
- Applying the fully multiband approaches improves the results even further. The difference between band-wise and multiband techniques is larger in the uncorrelated noise case. Apparently, noise correlations between bands diminish the gain that can be obtained from the multiband approaches. This can be explained in the following way: assume noise correlation that is equal to the signal correlation; then the ratio in equation (7) becomes diagonal (because  $\mathbf{C}_s$  is just a scaled version of  $\mathbf{C}_s + \mathbf{C}_n$ ), meaning that all bands are treated separately, like in the band-wise case.
- The more accurate the description of the prior distributions is, the more improvement one obtains. The use of Bernoulli-Gaussian (BG-Sure, BG-Moments) and Gaussian scale mixtures (GSM) improves the denoising quality over the use of the Gaussian prior. In the correlated noise case there is barely a distinction between the techniques based on the BG prior and the GSM prior. This is also true in the uncorrelated case, except for low noise, where GSM seems to be superior.

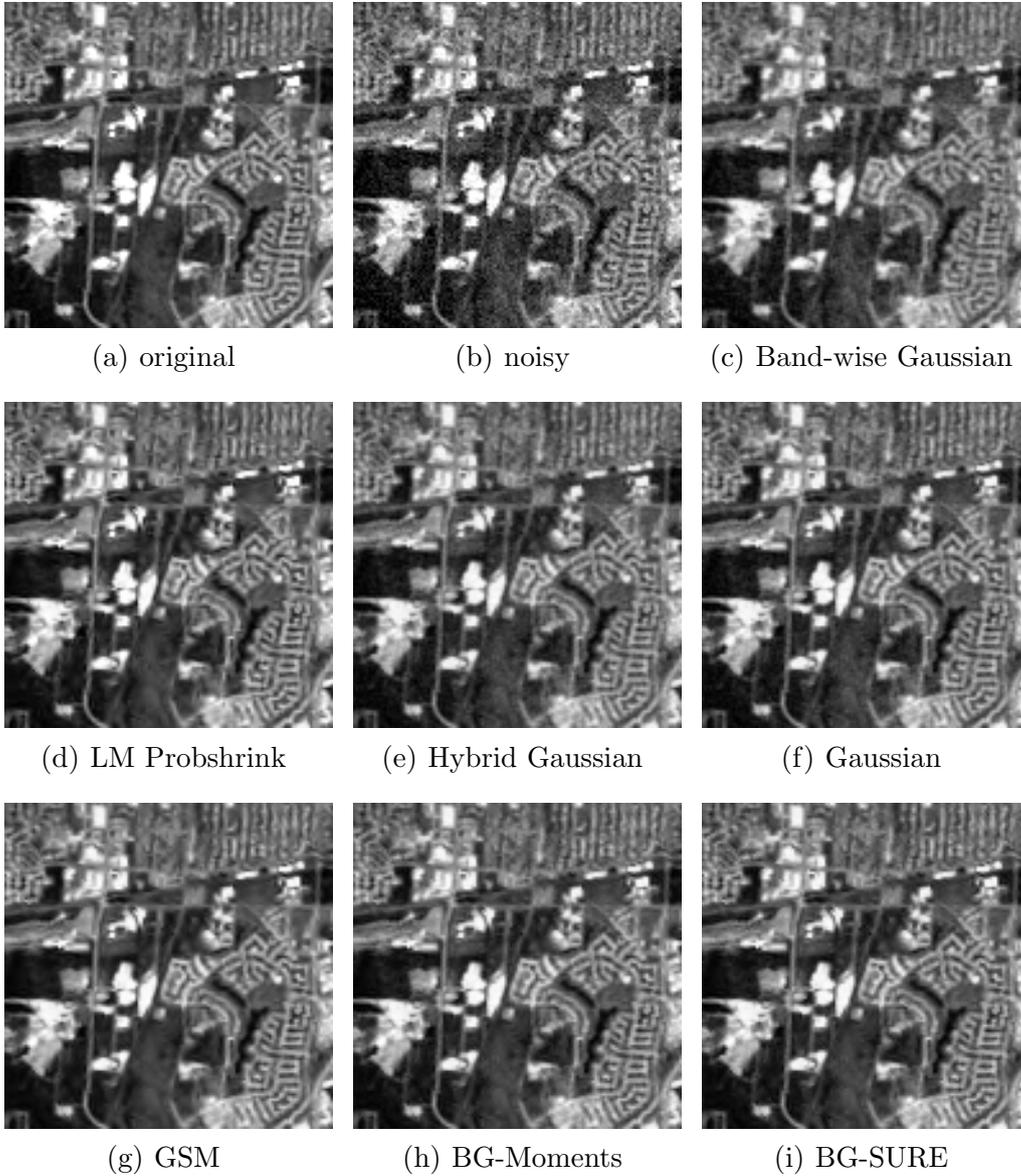


Fig. 1. a): detail of original red band of the Landsat image; (b): noise corrupted image; (c)-(i):results for the band after denoising with the presented wavelet based techniques.

#### 4.2 Experiment 2: AVIRIS dataset

In the second experiment an AVIRIS image over Cuprite, Nevada, was taken. Here, we concentrated on the infra-red (IR) part of the spectrum, containing noisy bands near the water vapor absorption band. For this experiment, 25 bands were uniformly selected in the  $2 - 2.5\mu m$  region. No additional noise was added. Figure 2 shows one band of the image near  $2.5\mu m$ .

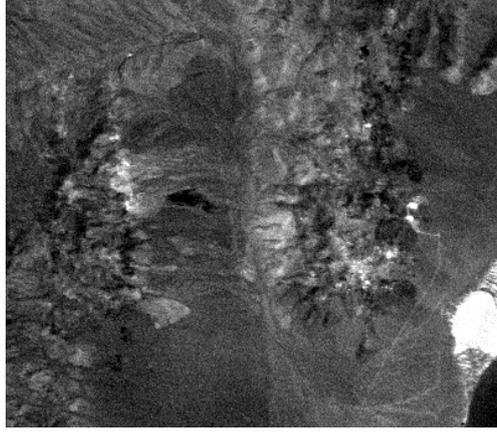


Fig. 2. AVIRIS  $2.479\mu m$  band of the cuprite image

In order to denoise this 25 band image, the noise covariance is required. The noise covariance was estimated as follows: let  $\mathbf{x}^{1,HH,b}$  denote the wavelet coefficients of the first resolution level and orientation subband  $HH$  from spectral band  $b$ . Then the diagonal elements of the noise covariance are estimated by the classical median estimator [42]:

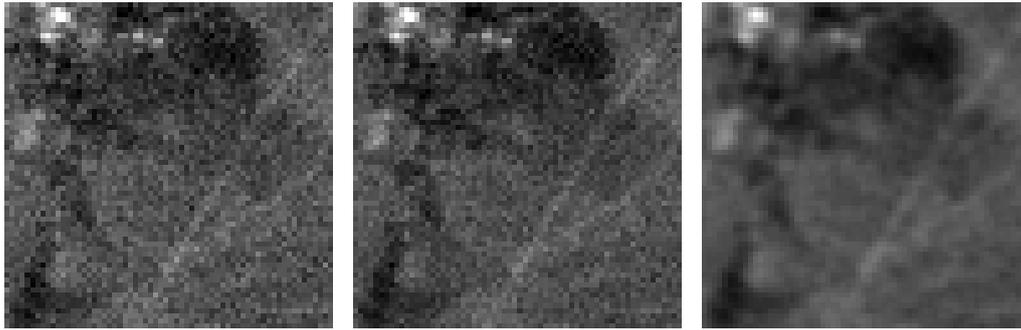
$$\sqrt{C_{bb}} = \frac{\text{median}(|x^{1,HH,b}|)}{0.6745}. \quad (27)$$

The off-diagonal elements between spectral bands  $b$  and  $b'$  are estimated, as in [30]:

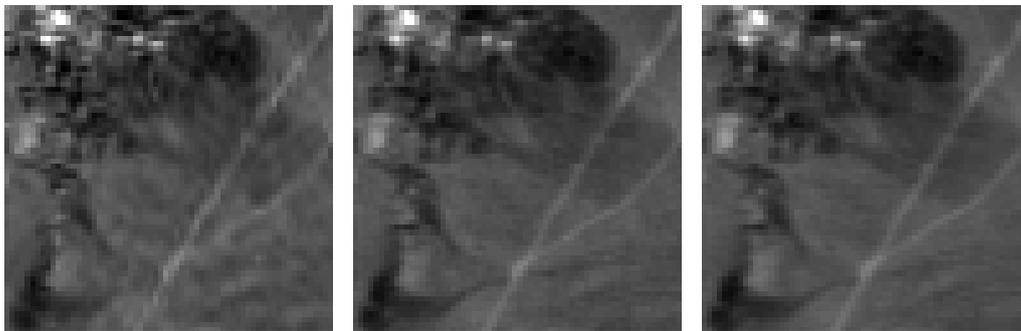
$$C_{bb'} = \frac{\sqrt{C_{bb}C_{b'b'}}}{4} \left( \left( \frac{\text{median}(|\frac{x^{1,HH,b}}{\sqrt{C_{bb}}} + \frac{x^{1,HH,b'}}{\sqrt{C_{b'b'}}}|)}{0.6745} \right)^2 - \left( \frac{\text{median}(|\frac{x^{1,HH,b}}{\sqrt{C_{bb}}} - \frac{x^{1,HH,b'}}{\sqrt{C_{b'b'}}}|)}{0.6745} \right)^2 \right) \quad (28)$$

All denoising techniques were applied on the image. For the MNF transform, 12 bands were retained to reconstruct the image. Again, for the hybrid band-

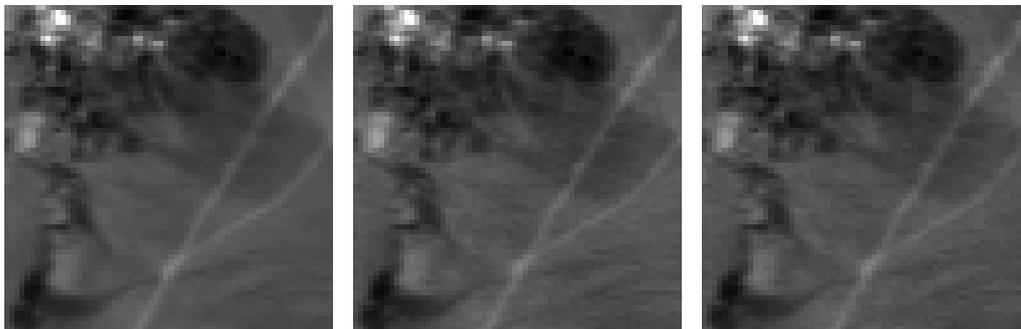
wise procedure a MNF transformation is used prior to denoising. Noise whitening is applied before the Probshrink procedure.



(a) original  $2.479\mu m$  band (b) Minimum Noise Fraction (c) Band-wise Gaussian denoising



(d) LM Probshrink (e) Hybrid Gaussian (f) Gaussian



(g) GSM (h) BG-Moments (i) BG-SURE

Fig. 3. (a): detail of original  $2.479\mu m$  band of the AVIRIS cuprite image; (b): denoising result from MNF transformation; (c)-(i): results for the  $2.479\mu m$  band after denoising with the presented wavelet based techniques.

In Figure 3, a detail of the denoising results for the AVIRIS image are shown. We can conclude:

- Using MNF, the noise in the spatial domain is hardly smoothed.
- From the wavelet-based approaches, visually, the band-wise approach retains most noise structure. This can be explained by the fact that no inter-band correlation is accounted for in this technique.

- The hybrid band-wise techniques already show an improvement in the reconstructed image. Yet, the result is visually smoother than for the other multiband techniques. In case of Probshrink, this result indicates that the local spatial indicator does not carry sufficient information of the interband correlations to denoise the image properly.
- Looking closer to the results of the full multiband techniques, one only sees minor differences: GSM seems to denoise the image more than the other techniques, rendering smoother regions.

### 4.3 Experiment 3: MRI simulation dataset

The MRI data set consists of 4 different simulated MRI acquisitions of the same subject (two T1, a T2 and a PD acquisition). The data set was obtained from the the Brainweb site [43]. On these images Rician distributed noise of different amplitude is simulated. The noise was independent for each component. The images are shown in Figure 4. In Table 3, the obtained PSNR values are shown, and Figure 5, the denoised  $T1$  image is shown, with  $\sigma = 15$ .

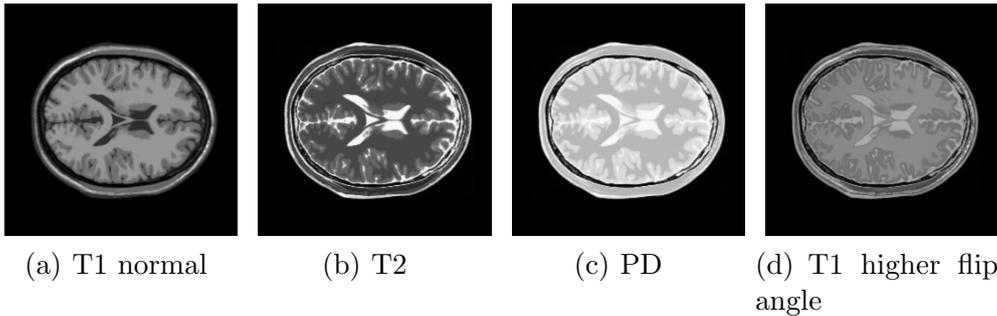


Fig. 4. Simulated MRI images using different acquisition schemes

The background pixels were not considered in the calculation of parameters used in the denoising procedures and to obtain the PSNR values. This would otherwise offset the estimation of the signal correlation, since it is calculated globally over the entire wavelet subband image. In the MRI experiment MNF did not improve the image quality, even not for high noise levels, and is not included in the table. In these experiments, the standard deviation of denoising performance over different instances of the noise is less than 0.1 dB. Similar conclusions as in the first experiment can be drawn.

<b>initial</b>	34.13	28.17	24.67	22.13	20.25
WAVELET BAND-WISE					
<b>Gaussian</b>	36.83	32.80	30.49	28.70	27.45
WAVELET HYBRID BAND-WISE					
<b>hybrid Gaussian</b>	37.58	33.32	30.94	29.09	27.78
<b>LM Probshrink</b>	39.07	35.21	32.99	31.17	29.93
WAVELET MULTICOMPONENT					
<b>Gaussian</b>	38.35	34.25	32.00	30.32	29.22
<b>GSM</b>	40.43	35.99	33.51	<b>31.53</b>	<b>30.22</b>
<b>BG-Moments</b>	<b>40.45</b>	<b>36.03</b>	<b>33.58</b>	25.92	30.14
<b>BG-Sure</b>	40.13	35.42	32.25	30.58	25.98

Table 3

PSNR values for Brainweb data experiments with simulated Rician noise (in dB). The 5 columns show the results for values of  $\sigma$  given by 5, 10, 15, 20 and 25 respectively.

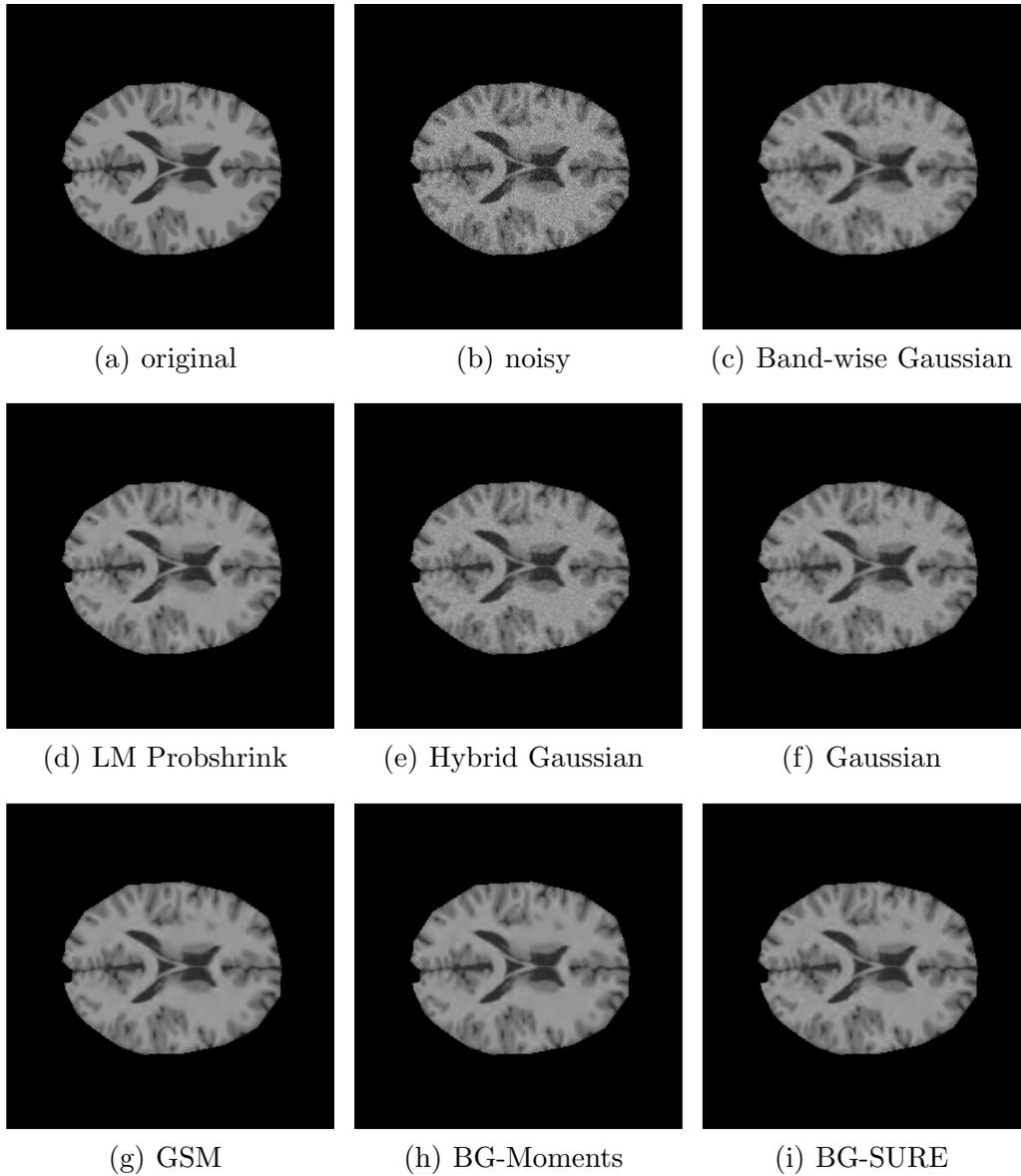


Fig. 5. a): original T1 image of four band set; (b): noise corrupted image; (c)-(i):results for the band after denoising with the presented wavelet based techniques.

#### 4.4 Experiment 4: MRI diffusion tensor imaging dataset

In the fourth experiment, denoising on another MRI dataset is performed. For this, we used a simulated MRI diffusion tensor imaging (DTI) data set (Figure 6), that was constructed using DTI simulation software that is built in-house. This set consists of 6 diffusion weighted images (DWI). In each image the grey values are related to the diffusion of water in different directions. Due to the nature of the acquisition, these DWI have a low SNR. As in the third experiment, we simulated Rician distributed noise, and applied the multicomponent denoising procedures on the 6 band image (the 6 DWI). Finally, the PSNR of the reconstructed bands is calculated and represented in table 4. Again, The standard deviation of denoising performance over different instances of the noise is less than 0.1 dB. In Figure 7, the results are shown for noise with  $\sigma = 15$ . Again, similar results as in the previous experiments are obtained. In this case, however, hybrid bandwise LM Probshrink is for higher noise levels slightly better than the multicomponent GSM model.

<b>initial</b>	34.21	28.11	24.58	22.11	20.18
WAVELET BAND-WISE					
<b>Gaussian</b>	36.48	32.47	30.29	28.76	27.35
WAVELET HYBRID BAND-WISE					
<b>hybrid Gaussian</b>	36.73	32.70	30.49	28.97	27.49
<b>LM Probshrink</b>	38.50	34.61	<b>32.45</b>	<b>30.87</b>	29.21
WAVELET MULTICOMPONENT					
<b>Gaussian</b>	37.04	33.24	31.32	29.97	28.75
<b>GSM</b>	<b>39.21</b>	<b>34.79</b>	32.43	30.72	29.18
<b>BG-Moments</b>	38.84	34.41	31.93	30.37	29.05
<b>BG-Sure</b>	38.47	34.19	31.70	30.64	<b>29.44</b>

Table 4

PSNR values for DTI data experiments with simulated Rician noise (in dB). The 5 columns show the results for values of  $\sigma$  given by 5, 10, 15, 20 and 25 respectively.

#### 4.5 Time complexity

Finally, the time complexity of the different techniques is studied. Table 5 shows the timings relative to Gaussian band-wise denoising which is set to one time unit. Here, only the time to process the wavelet coefficients is taken into account, since all techniques need the same computation of the wavelet and a inverse wavelet transform. One can clearly see the large range of computational complexity. The introduction of the multicomponent heavy tailed distributions increases the computational cost several orders. One exception is the BG-

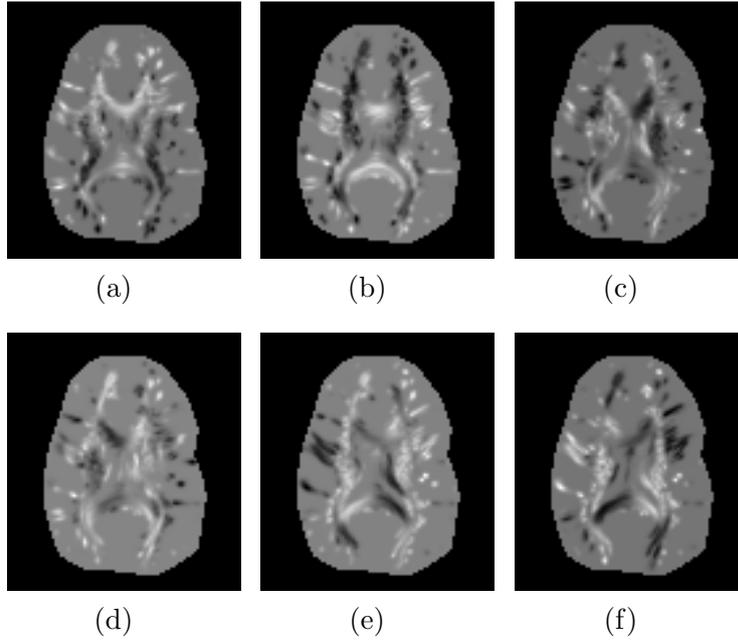


Fig. 6. Simulated MRI diffusion weighted images of different directional acquisition

Moments technique which seems preferable. This large difference in execution time between BG-Moments and the other multicomponent techniques can be explained by the existence of an analytic expression for  $\epsilon$  in (19) which makes a swift implementation possible. Compared to this, ESURE needs optimization of  $\epsilon$  and  $\mathbf{C}_s$ , while Probshrink and GSM need a numerical integral evaluation, operations which increase the computational burden.

BAND-WISE	
<b>Gaussian</b>	1.0
HYBRID BAND-WISE	
<b>LM Probshrink</b>	15
MULTICOMPONENT	
<b>Gaussian</b>	2.8
<b>GSM</b>	24
<b>BG-Moments</b>	4.5
<b>BG-Sure</b>	43
<b>WT and inverse WT</b>	5.1

Table 5

Relative time comparison for Winnipeg image experiment. The timings are without the wavelet transformations, for which the timing is separately indicated in last row

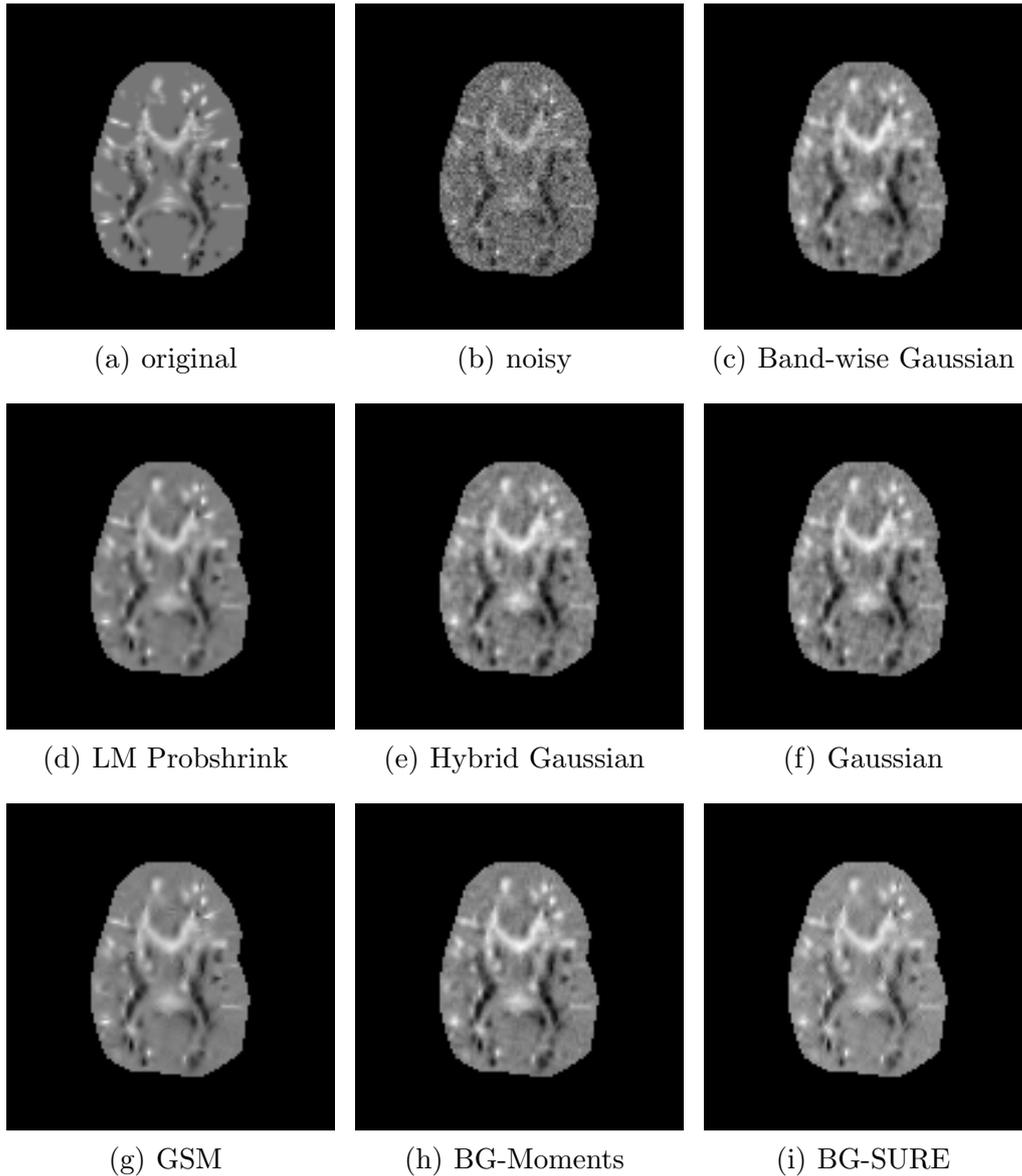


Fig. 7. a): one band of original DTI image; (b): noise corrupted image 24.7 dB; (c)-(i):results for the band after denoising with the presented wavelet based techniques.

## 5 Conclusions

In this paper, spatial wavelet-based denoising of multicomponent images was aimed at. We presented a Bayesian least-squares optimization framework, using different multicomponent prior models for the wavelet coefficients: Gaussian scale mixture models, Bernoulli-Gaussian mixture models and Laplacian mixture models. The presented procedures are compared to standard spectral denoising and single-band denoising procedures. We analyzed the suppression of non-correlated as well as correlated white Gaussian noise on Landsat

data and on MRI data with simulated noise. From the experimental results, we can conclude that superior denoising performance is obtained when a) applying spatial (wavelet-based) denoising, compared to spectral denoising; b) accounting for the interband correlations, compared to denoising on each band separately; c) using prior models that fully account for the interband covariances; and d) using prior models that better approximate the wavelet coefficients marginal densities.

## Acknowledgments

The authors wish to thank Dr. A. Leemans for making available the phantom DTI-MRI dataset.

## References

- [1] I. Thomas, V. Benning, N. Ching, *Classification of remotely sensed images*, Adam Hilger, Bristol, 1987.
- [2] C. Lee, D. Landgrebe, Analyzing high-dimensional multispectral data, *IEEE Trans. Geosci. Remote Sensing* 31 (4) (1993) 388–400.
- [3] A. A. Green, M. Berman, P. Switzer, M. D. Craig, A transformation for ordering multispectral data in terms of image quality with implications for noise removal, *IEEE Trans. Geosci. Remote Sensing* 26 (1) (1988) 64–74.
- [4] J. B. Lee, A. S. Woodyatt, M. Berman, Enhancement of high spectral resolution remote-sensing data by noise-adjusted principal components transform, *IEEE Trans. Geosci. Remote Sensing* 28 (3) (1990) 295–304.
- [5] J. L. Starck, P. Querre, Multispectral data restoration by the wavelet Karhunen-Loève transform, *Signal Processing* 81 (2001) 2449–2459.
- [6] H. Othman, S.-E. Qian, Noise reduction of hyperspectral imagery using hybrid spatial-spectral derivative-domain wavelet shrinkage, *IEEE Trans. Geosci. Remote Sensing* 44 (2) (2006) 397–408.
- [7] I. Daubechies, *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, Philadelphia, 1992.
- [8] S. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Trans. Pattern Anal. and Machine Intel.* 11 (7) (1989) 674–693.
- [9] M. Vetterli, J. Kovačević, *Wavelets and Subband Coding*, Prentice-Hall, 1995.

- [10] D. L. Donoho, De-noising by soft-thresholding, *IEEE Trans. Inform. Theory* 41 (1995) 613–627.
- [11] S. G. Chang, B. Yu, M. Vetterli, Adaptive wavelet thresholding for image denoising and compression, *IEEE Trans. Image Proc.* 9 (9) (2000) 1532–1546.
- [12] B. Vidakovic, Wavelet-based nonparametric Bayes methods, in: D. D. Dey, P. Müller, D. Sinha (Eds.), *Practical Nonparametric and Semiparametric Bayesian Statistics*, Vol. 133 of *Lecture Notes in Statistics*, Springer Verlag, New York, 1998, pp. 133–155.
- [13] B. Vidakovic, Nonlinear wavelet shrinkage with bayes rules and bayes factors, *J. of the American Statistical Association* 93 (1998) 173–179.
- [14] F. Abramovich, T. Sapatinas, B. Silverman, Wavelet thresholding via a Bayesian approach, *J. of the Royal Statist. Society B* 60 (1998) 725–749.
- [15] D. Leporini, J. C. Pasquet, H. Krim, Best basis representation with prior statistical models, in: P. Müller, B. Vidakovic (Eds.), *Lecture Notes in Statistics*, Springer Verlag, 1999, pp. 155–172.
- [16] H. A. Chipman, E. D. Kolaczyk, R. E. McCulloch, Adaptive Bayesian wavelet shrinkage, *J. of the Amer. Statist. Assoc* 92 (1997) 1413–1421.
- [17] M. Clyde, G. Parmigiani, B. Vidakovic, Multiple shrinkage and subset selection in wavelets, *Biometrika* 85 (2) (1998) 391–401.
- [18] E. P. Simoncelli, E. H. Adelson, Noise removal via Bayesian wavelet coring, in: *Proc. IEEE Internat. Conf. Image Proc. ICIP*, Vol. 1, Lausanne, Switzerland, 1996, pp. 379–382.
- [19] M. Hansen, B. Yu, Wavelet thresholding via MDL for natural images, *IEEE Trans. Inform. Theory* 46 (5) (2000) 1778–1788.
- [20] A. Achim, P. Tsakalides, A. Bezerianos, SAR image denoising via Bayesian wavelet shrinkage based on heavy-tailed modeling, *IEEE Trans. Geosc. and Remote Sensing* 41 (8) (2003) 1773–1784.
- [21] P. Moulin, J. Liu, Analysis of multiresolution image denoising schemes using generalized Gaussian and complexity priors, *IEEE Trans. Inform. Theory* 45 (1999) 909–919.
- [22] L. Şendur, I. W. Selesnick, Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency, *IEEE Trans. Signal Proc.* 50 (11) (2002) 2744–2756.
- [23] M. S. Crouse, R. D. Nowak, R. G. Baranuik, Wavelet-based statistical signal processing using hidden Markov models, *IEEE Trans. Signal Proc.* 46 (4) (1998) 886–902.
- [24] M. K. Mihçak, I. Kozintsev, K. Ramchandran, P. Moulin, Low-complexity image denoising based on statistical modeling of wavelet coefficients, *IEEE Signal Proc. Lett.* 6 (12) (1999) 300–303.

- [25] J. Portilla, V. Strela, M. J. Wainwright, E. P. Simoncelli, Image denoising using scale mixtures of gaussians in the wavelet domain, *IEEE Trans. Image Proc.* 12 (11) (2003) 1338–1351.
- [26] M. Malfait, D. Roose, Wavelet-based image denoising using a Markov random field a priori model, *IEEE Trans. Image processing* 6 (4) (1997) 549–565.
- [27] M. Jansen, A. Bultheel, Empirical Bayes approach to improve wavelet thresholding for image noise reduction, *J. Amer. Stat. Assoc.* 96 (454) (2001) 629–639.
- [28] A. Pižurica, W. Philips, I. Lemahieu, M. Acheroy, A joint inter- and intrascale statistical model for wavelet based Bayesian image denoising, *IEEE Trans. Image Proc* 11 (5) (2002) 545–557.
- [29] A. Pižurica, W. Philips, I. Lemahieu, M. Acheroy, A versatile wavelet domain noise filtration technique for medical imaging, *IEEE Trans. Medical Imaging* 22 (3) (2003) 323–331.
- [30] A. Benazza-Benyahia, J.-C. Pesquet, Building robust wavelet estimators for multicomponent images using Steins’ principle, *IEEE Trans. Image Proc.* 14 (11) (2005) 1814–1830.
- [31] P. Scheunders, J. Driesen, Least-squares interband denoising of color and multispectral images, in: *Proc. IEEE Internat. Conf. Image Proc. ICIP*, Vol. 2, Singapore, 2004, pp. 985–988.
- [32] P. Scheunders, S. De Backer, Wavelet denoising multicomponent images, using a Gaussian scale mixture model, in: *Proc. IEEE Internat. Conf. Pattern Recognition ICPR*, Vol. 3, Hong Kong, 2006, pp. 754–757.
- [33] A. Pižurica, B. Huysmans, P. Scheunders, W. Philips, Wavelet domain denoising of multispectral remote sensing imagery adapted to the local spatial and spectral context, in: *Proc. Int. Geosc. and Remote Sensing Symp. IGARSS*, Vol. 6, Seoul, Korea, 2005, pp. 4260–4263.
- [34] A. Pižurica, W. Philips, P. Scheunders, Wavelet domain denoising of single-band and multi-band images adapted to the probability of the presence of features of interest, in: *Proc of SPIE Conference Wavelets XI*, Vol. 5914, San Diego, California, USA, 2005, pp. 508–521.
- [35] A. Pižurica, W. Philips, Estimating the probability of the presence of a signal of interest in multiresolution single- and multiband image denoising, *IEEE Trans. Image Processing* 15 (3) (2006) 654–665.
- [36] P. Scheunders, Wavelet thresholding of multivalued images, *IEEE Trans. Image Proc.* 13 (4) (2004) 475–483.
- [37] R. R. Coifman, D. L. Donoho, Translation-invariant denoising, in: A. Antoniadis, G. Oppenheim (Eds.), *Wavelets and Statistics*, Springer Verlag, New York, 1995, pp. 125–150.
- [38] S. Mallat, *A wavelet tour of signal processing*, Academic Press, London, 1998.

- [39] M. Jansen, *Noise Reduction by Wavelet Thresholding*, Springer-Verlag, New York, 2001.
- [40] G. E. P. Box, C. Tiao, *Bayesian inference in statistical analysis*, Addison-Wesley, 1992.
- [41] I. Johnstone, B. Silverman, Empirical Bayes selection of wavelet thresholds, *Annals of Statistics* 33 (4) (2005) 1700–1752.
- [42] D. L. Donoho, I. M. Johnstone, Adapting to unknown smoothness via wavelet shrinking, *Journal of the American Statistical Association* 90 (432) (1995) 1200–1224.
- [43] R. K.-S. Kwan, A. C. Evans, G. B. Pike, Mri simulation-based evaluation of image-processing and classification methods, *IEEE Trans. Med. Imag.* 18 (11) (1999) 1085–1097.