



**HAL**  
open science

## Diffusion Weighted Image Denoising Using Overcomplete Local PCA

José V. Manjón, Pierrick Coupé, Luis Concha, Antoni Buades, Louis D.  
Collins, Robles Montserrat

► **To cite this version:**

José V. Manjón, Pierrick Coupé, Luis Concha, Antoni Buades, Louis D. Collins, et al.. Diffusion Weighted Image Denoising Using Overcomplete Local PCA. PLoS ONE, 2013, 8 (9), pp.e73021. 10.1371/journal.pone.0073021 . hal-00871777

**HAL Id: hal-00871777**

**<https://hal.science/hal-00871777v1>**

Submitted on 10 Oct 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.

# Diffusion Weighted Image Denoising using overcomplete Local PCA

José V. Manjón<sup>1</sup>, Pierrick Coupé<sup>2</sup>, Luis Concha<sup>3</sup>,  
Antonio Buades<sup>4,6</sup>, D. Louis Collins<sup>2</sup>, Montserrat Robles<sup>1</sup>

<sup>1</sup> Instituto de Aplicaciones de las Tecnologías de la Información y de las Comunicaciones Avanzadas (ITACA), Universidad Politécnica de Valencia, Camino de Vera s/n, 46022 Valencia, Spain.

<sup>2</sup> McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, Canada. University, 3801 University Street, Montreal, Canada H3A 2B4.

<sup>3</sup> Institute of Neurobiology, National Autonomous University of Mexico. Boulevard Juriquilla 3001, 76230, Querétaro, México.

<sup>4</sup> CMLA, ENS Cachan, 61 av. du président Wilson 94235 CACHAN, France

<sup>5</sup> Laboratoire Bordelais de Recherche en Informatique, Unité Mixte de Recherche CNRS (UMR 5800), 351, cours de la Libération F-33405 Talence cedex, France

<sup>6</sup> Departament de Matemàtiques, Universitat Illes Balears, Ctra Valldemossa km 7.5, 07122 Palma, España.

\* **Corresponding author:** José V. Manjón. Instituto de Aplicaciones de las Tecnologías de la Información y de las Comunicaciones Avanzadas (ITACA), Universidad Politécnica de Valencia, Camino de Vera s/n, 46022 Valencia, Spain.

Tel.: (+34) 96 387 70 00 Ext. 75275 Fax: (+34) 96 387 90 09.

E-mail address: jmanjon@fis.upv.es (José V. Manjón)

**Keywords:** Diffusion Weighted Imaging, Denoising, PCA

## **Abstract**

Diffusion Weighted Images (DWI) normally shows a low Signal to Noise Ratio (SNR) due to the presence of noise from the measurement process that complicates and biases the estimation of quantitative diffusion parameters. In this paper, a new denoising methodology is proposed that takes into consideration the multicomponent nature of multi-directional DWI datasets such as those employed in diffusion imaging. This new filter reduces random noise in multicomponent DWI by locally shrinking less significant Principal Components using an overcomplete approach. The proposed method is compared with state-of-the-art methods using synthetic and real clinical MR images, showing improved performance in terms of denoising quality and estimation of diffusion parameters.

## 1. Introduction

Magnetic Resonance imaging (MRI) has been intensively used to study the normal and pathological human brain. In the last decade, Diffusion Weighted Imaging (DWI) has gained a lot of attention due to its ability to measure the microscopic motion of water molecules within tissue, and hence lead to insight regarding cerebral white matter microstructure. One of the most prominent applications of DWI is Diffusion Tensor Imaging (DTI), where the directionality and the magnitude of water diffusion is estimated using a tensor model yielding images of normal and abnormal white matter fiber structure and maps of brain connectivity through fiber tracking [1][2].

Diffusion weighted images have an inherently low signal to noise ratio (SNR) due to low signal amplitude and pronounced thermal noise, which is more evident than in conventional MRI due to extremely fast echo-planar acquisition strategies. Such low SNR makes DWI analysis complicated and biases the estimation of quantitative diffusion parameters [3]. Moreover, this limited SNR makes automated processing of these images challenging and potentially misleading.

To increase the SNR, it is a common practice to average several acquisitions in order to reduce noise variance (but without removing noise driven bias). However, this approach is time consuming in terms of acquisition and therefore not adequate for typical clinical settings where patients cannot remain still for extended periods of time.

On the other hand, denoising techniques can be applied to improve data quality as a post-processing step, thereby not increasing the scanning time. Denoising techniques applied in DWI can be divided on three categories: first, the techniques that directly filter the acquired DW images [4]-[18], second, those that regularize the tensors after estimation [19],[20] and third those that use a regularization term during the estimation/inversion [21]-[23]. However, Jones and Basser [3] demonstrated that the underestimation of diffusion anisotropy due to noise cannot be corrected once the tensor parameters are determined.

Recently, several methods that directly operate over the DWI data have reported remarkable results. For example, Wiest-Daesslé et al. [17] proposed a modification of the well-known Non-local Means method to deal with Rician noise which was further studied by Descoteaux et al. [7]. More recently, Tristan et al. [18] proposed a Linear Minimum Mean Squared Error (LMMSE) rooted approach to deal with Rician noise and the multi-component nature of DW images.

From a different point of view, Principal Component Analysis (PCA) and related approaches have been previously used for noise reduction in images [24]-[26]. In this context, noise removal can be done by a) decomposing the signal into the local principal components, then b) shrinking the less relevant components, and finally c) reconstructing back the signal. The key idea of this process is the fact that image patterns can be represented as a linear combination of a small number of basis images while the noise, being not sparse will be spread over all available components. In this sense, an interesting approach is the two steps method proposed by Zhang et al. [27] where similar patches are grouped together before PCA decomposition takes place. Alternatively, PCA can be used to learn an orthogonal basis set from the noisy data and to use it to represent noisy patches but previously zeroing small coefficients [28]. This approach takes benefit from the intrinsic sparsity properties of the images.

PCA based denoising has already been applied for MRI filtering. In Manjón et al. [29] PCA was used as a postprocessing step to remove remaining noise after the application of a multicomponent non-local means filter for multimodal MRI. Also very recently, structure adaptive and edge constrained PCA related approaches has been proposed for DWI denoising [30,31].

In this paper, we propose a new denoising method based on local PCA designed to take into account the Rician nature of the noise present in DW images. The proposed filter takes advantage of the multi-directional nature of DW images by using local PCA decomposition to exploit the local signal profile redundancy in contrast with related PCA based methods that made use of local spatial pattern redundancy instead.

## 2. Material and methods

Principal component analysis is an orthogonal linear transformation that maps the data into a new coordinate system such that the greatest variance, by any projection of the data, comes to lie on the first axes (called the first principal component), the second greatest variance on the second coordinate, and so on.

PCA-based denoising can be achieved using global information of an image series (one component per image) [26] or locally using local image patches [24],[25],[29]. The first approach, although effective, requires the number of images to be higher than the number of significant components of the image resulting in a less sparse representation. This problem can be overcome by performing PCA decomposition over small local windows instead of the whole image what significantly produces sparser representation (in the extreme case, i.e. in homogeneous areas, it can result in the sparsest representation not requiring any component to represent the data but only the mean value of the region).

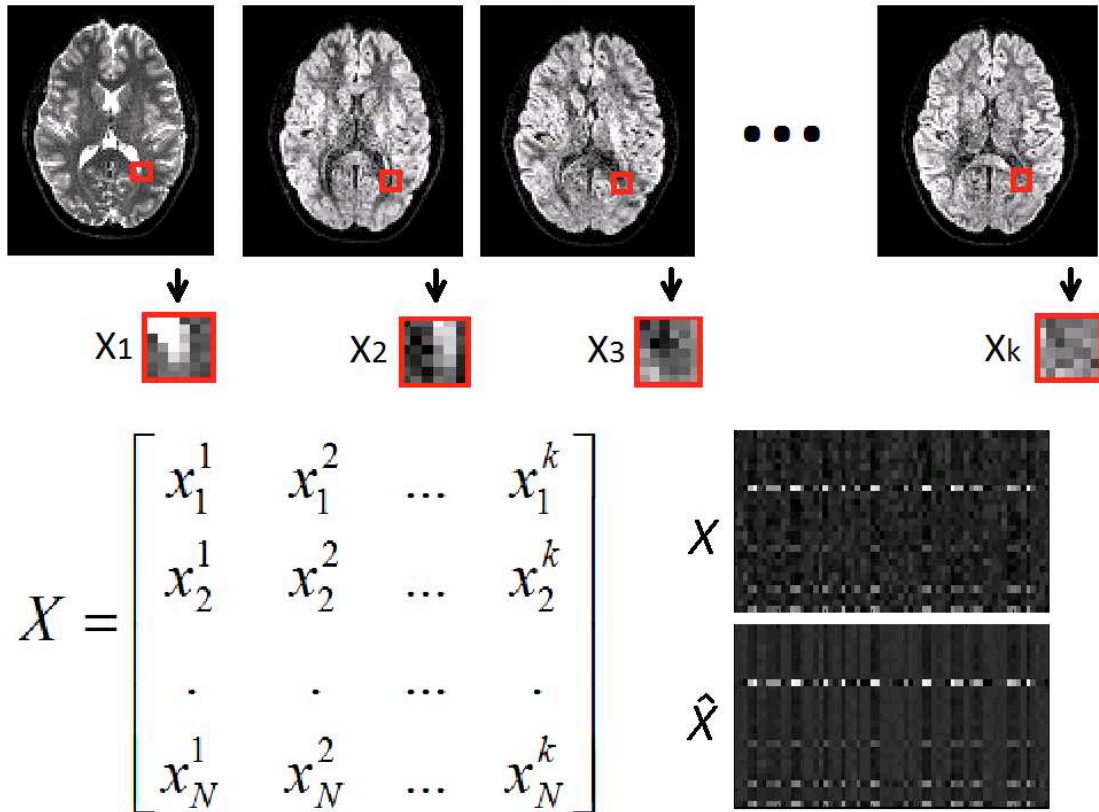
### 2.1. Local PCA Denoising (LPCA)

In our proposed approach, we apply a local PCA to exploit the multi-directional redundancy of DWI patterns rather than using local spatial image pattern redundancy as done by Muresan and Parks [24]. This has the benefit of not requiring a search for similar patches within the image [27] resulting in a much faster processing.

Our method assumes that the whole directional information of DW images can be locally represented by less than the original  $K$  components. For this reason, the image is analyzed using a local 4D sliding block and at each position a PCA decomposition of this block is applied in order to locally find the most reduced representation of this data. The cancelation of superfluous information in the transformed data reduces noise while preserving the main features of the images.

For each point  $x_i$  of the image domain  $\Omega \subset \mathfrak{R}^3$ , the 3D patches surrounding  $x_i$  in each directional image  $k$  are reordered as a column vector of a matrix  $X$  (see

figure 1).  $X$  is thus an  $N \times K$  matrix where  $N$  corresponds to the number of voxels of the 3D patch around the point of interest ( $N=64$  in our experiments which corresponds to  $4 \times 4 \times 4$  voxels included in the 3D patch), and  $K$  is the number of directional images (in DWI,  $K$  can range from 7 to the number of acquired directions). Therefore each row vector of this matrix represents the value of a voxel  $x_i$  across all  $K$  image directions.



**Figure 1.** Example of local matrix  $X$  formation from a DW image series. Each block is converted into a column of matrix  $X$ . At bottom-right, an example of a real matrix  $X$  and its filtered version  $\hat{X}$  is shown to highlight the high level of profile signal redundancy present in the local matrix  $X$ .

PCA is applied to the set of row vectors, considering  $N$  samples of dimension  $K$ . The set of row vectors of  $X$  are first normalized to have its barycenter in the origin (that is to center the columns of  $X$  by subtracting to each one its own average or equivalently to remove from each 3D block its mean). PCA transforms the sample vectors into a new coordinate system, where the few first

components represent the most part of the variation of the original data. The new PCA system is obtained by computing the eigenvectors of the covariance matrix of  $X$  being all the associated eigenvalues positive. These eigenvectors are stored as columns in a squared  $K \times K$  matrix,  $W$ . The associated eigenvalues correspond to the amount of variation of the new components and are stored in a diagonal matrix of size  $K \times K$ ,  $D = (d_{kk})$ . The new coordinates of the original data are computed by a simple matrix product (1) (see [46] and [47] for a more detailed introduction to eigenvalues decomposition and PCA).

$$Y = X W \quad (1)$$

Classical thresholding algorithms cancel coefficients of low magnitude. Instead of deciding to cancel a coefficient depending on its magnitude, we take the decision based on the associated eigenvalue. By doing that, we decide to cancel the  $k$ th coefficient of the  $i$ th vector depending on the value  $d_{kk}$ .

The use of this thresholding strategy has two advantages. First, since the new basis depends on the data the classical thresholding strategy would be less stable and second the proposed strategy is optimal for the whole set of  $N$  vectors. This thresholding strategy can be numerically applied by modifying the matrix  $D$  into  $\hat{D}$ . Each value in the diagonal of  $D$  is canceled if its magnitude is lower than a certain parameter  $\tau$ . After an exhaustive search for the optimum value of  $\tau$  parameter, it was set to  $(2.7\sigma)^2$ , where  $\sigma^2$  is the estimated local noise variance. This threshold is set depending on  $\sigma^2$  since each element of the diagonal matrix  $D$ ,  $d_{kk}$ , actually represents the variance of the different principal components.

Finally the denoised matrix is obtained by computing the inverse PCA transformation [46,47].

$$\hat{X} = YW^{-1}\hat{D} \quad (2)$$



where the inverse  $W^{-1}$  is equal to the transpose of  $W$  since the matrix is orthogonal and matrix  $\hat{D}$  is obtained by comparing the magnitude of the diagonal with a fixed threshold as described previously. Therefore, although the final values are obtained by a matrix product, the overall process is nonlinear.

The local PCA denoising is done in an overcomplete manner since overlapping patches are used to achieve the processing of each voxel of the image volume. Due to this patch overlapping, several estimations are obtained for a given voxel since all the voxels are processed independently. Each one of these estimations is obtained by a thresholding in a different basis. The advantages of such a multipoint-wise approach compared to point-wise method are described in Katkovnik et al [32].

Therefore, for each voxel all the local estimates  $\hat{x}_i(j)$  are combined from all the overlapping  $j$  blocks at position  $i$  using the following weighted average rule:

$$\hat{x}_i = \frac{\sum_{j=1}^V \theta_j \hat{x}_i(j)}{\sum_{j=1}^V \theta_j} \quad \theta_j = \frac{1}{1 + \|\hat{D}_j\|_0} \quad (3)$$

where  $V$  is the number of overlapping blocks contributing to  $\hat{x}_i$  and  $\theta_j$  is the weight of each block  $j$ , which is proportional to the inverse of the  $\hat{D}_j$   $L_0$  norm (i.e., the number of nonzero coefficients of the diagonal matrix  $\hat{D}$  at block  $j$  after the thresholding operation). This approach gives more weight to estimates with more null values after thresholding. Such averaging enables to remove more noise and minimizes Gibbs artifacts [31] in a similar manner to the translation invariant denoising proposed by Coifman and Donoho [34]. The only parameter in the proposed method is the threshold value  $\tau$ .

## 2.2. Adaptation to Rician noise

Noise in magnitude MR images is usually modeled with a Rician distribution [35]. The asymmetry of the Rician distribution results in a non-constant intensity bias as it depends on SNR. Therefore, all measurements based on intensity will be biased such as Apparent Diffusion Coefficient (ADC), Fractional Anisotropy (FA), etc.

To avoid such bias, some authors have proposed to remove the bias in the squared magnitude image [17], [35]. However, this approach cannot be applied within our framework since we are not reducing the noise using the averaging principle. In our case, due to the effect of PCA thresholding, the bias in the squared domain is not constant, but dependent on intensity. However, it can be estimated theoretically and inverted in the original domain using the properties of the first moment of a Rician distribution [36].

The proposed algorithm is used to recover the mean value of the Rice distribution  $R(v, \sigma)$  with parameters  $v$  and  $\sigma$ , being  $v$  the true value we want to recover and  $\sigma$  the noise standard deviation. This expected value writes as:

$$E[R(v, \sigma)] = \sigma \sqrt{\frac{\pi}{2}} \exp\left(-\frac{v^2}{2\sigma^2}\right) \left( \left(1 + \frac{v^2}{2\sigma^2}\right) I_0\left(\frac{v^2}{4\sigma^2}\right) + \left(\frac{v^2}{2\sigma^2}\right) I_1\left(\frac{v^2}{4\sigma^2}\right) \right) \quad (4)$$

where  $I_0$  and  $I_1$  are the modified Bessel functions of order zero and one, respectively.

Once we obtain our denoised value, we should compensate the bias by inverting previous expression and recovering the true value  $v$ . However, this is not possible analytically and a numerical inverse is computed. We observe that  $E[R(v, \sigma)]/\sigma$  can be written as a function of  $\phi = v/\sigma$ .

$$\frac{E[R(v, \sigma)]}{\sigma} = \sqrt{\frac{\pi}{2}} \exp\left(-\frac{\phi^2}{2}\right) \left( \left(1 + \frac{\phi^2}{2}\right) I_0\left(\frac{\phi^2}{4}\right) + \left(\frac{\phi^2}{2}\right) I_1\left(\frac{\phi^2}{4}\right) \right) \quad (5)$$

The inverse of this expression as a function of  $\phi$  can be stored into a Look Up Table (LUT) that we denote by  $\eta(\phi)$ .

The final algorithm corrects the estimated value  $x$  obtained by the denoising process to:

$$\hat{x} = \sigma\eta(x/\sigma) \quad (6)$$

### 2.3. Noise estimation in DWI

The proposed method requires the local noise level present in the images as the only input parameter. Although several methods have been proposed to estimate Rician noise level on MRI images [10],[37] few of them are able to estimate spatially varying noise patterns. Here, we propose two novel methods to provide local noise variance estimation of DW images depending on whether they have one or multiple  $b=0$  images.

#### Multiple $B=0$ image noise variance Estimator (MUBE)

If multiple non-DWI (i.e.,  $b=0$  s/mm<sup>2</sup>, termed  $b_0$  henceforth) images are available in the dataset we can use image differences to robustly estimate the noise pattern. Landman et al. [38] presented a method for local noise variance estimation using the differences of two non diffusion weighted images that robustly estimated the noise pattern assuming that Rician noise can be well approximated by a Gaussian distribution if the local SNR is high enough (SNR>5).

We propose a similar approach, albeit without assuming Gaussian distributed noise by using the analytical correction scheme proposed by Koay and Basser [36].

Given a set of  $n$   $b_0$  images ( $n \geq 2$ ), we first compute global PCA decomposition of image series (i.e. we perform a PCA decomposition of a matrix formed by as many columns as image voxels and as many rows as images, see Fig. 2). In this way, first components will be associated to both signal and noise while the last components will contain mainly noise contribution. Therefore, an initial noise field estimation can be obtained just calculating the local noise standard

deviation of the least significant component within regions of 3x3x3 voxels (figure 2).

However, since we are dealing with Rician and not Gaussian noise, the local standard deviation is underestimated at low SNR areas due to the asymmetric distribution of Rician noise. To correct such underestimation a correction factor was applied based on the local SNR as described by Koay and Basser [36] and used for MRI noise estimation in Coupe et al. [37].

$$\hat{\sigma}^2 = \sigma^2 / \xi(SNR) \quad (7)$$

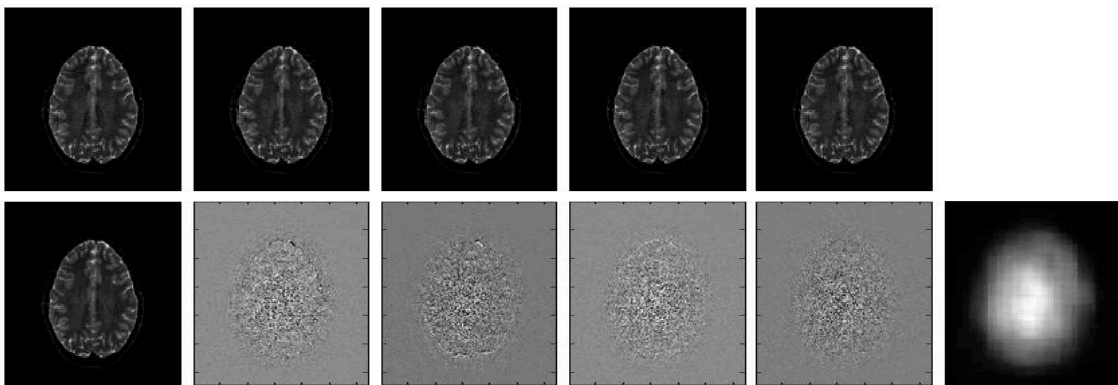
where  $\xi(SNR)$  is defined as follows:

$$\xi(SNR) = 2 + SNR^2 - \frac{\pi}{8} \times \exp\left(-\frac{SNR^2}{2}\right) \left( \left(2 + SNR^2\right) I_0\left(\frac{SNR^2}{4}\right) + SNR^2 I_1\left(\frac{SNR^2}{4}\right) \right) \quad (8)$$

$$SNR = \mu / \sigma$$

Here  $\mu$  is the local mean,  $\sigma^2$  is the local estimation of the noise variance and  $\hat{\sigma}^2$  is the corrected estimation in the Rician case.

Finally, a low-pass filter is applied to regularize the estimated noise field to provide a more regular noise field. We used a kernel size of 15 mm<sup>3</sup> which was experimentally determined.



**Figure 2.** Upper row, from left to right: first to 5th  $b=0$  images. Lower row, from left to right: first to 5th principal components and the regularized estimated noise field. Note that the least significant component clearly shows the spatially varying noise pattern while showing no anatomy at all.

### Single B=0 image noise variance Estimator (SIBE)

If only one non-DWI is available we need to follow a different strategy to obtain the initial noise field estimation. In this case, we can use the PCA decomposition of the gradient images instead of the b0 images (we should note here that the SNR of gradient images is significantly lower than b0 images invalidating the Gaussian assumption for all brain regions). Similarity to MUBE method, first components will be associated to both signal and noise while the last components will contain mainly noise. Thus a local noise estimation (using also a 3x3x3 voxel local regions) from this last component will enable to obtain an initial noise field estimation. Finally, this underestimated initial noise field is processed in the same manner of MUBE method to correct local noise underestimation.

### **3. Experiments and results**

To evaluate and compare the proposed methods a set of experiments were performed with simulated and in-vivo real datasets.

#### **3.1 Simulated dataset**

A numerical phantom consisting of a set of diffusion-weighted images was generated using the Numerical Fiber Generator (NFG) software package [39]. The simulated images were generated with a diffusion-weighted response function based on the diffusion tensor model with a fractional anisotropy of FA=0.8, apparent diffusion coefficient  $ADC=0.9 \times 10^{-3}$  mm<sup>2</sup>/s, 7 b=0 s/mm<sup>2</sup> images and 60 DWIs along 60 uniformly distributed diffusion-gradient directions (b =3000 s/mm<sup>2</sup>), with volume field of view dimensions of 100x100x100 voxels, with a voxel size of 2x2x2mm<sup>3</sup>.

#### **3.2 In-vivo real dataset**

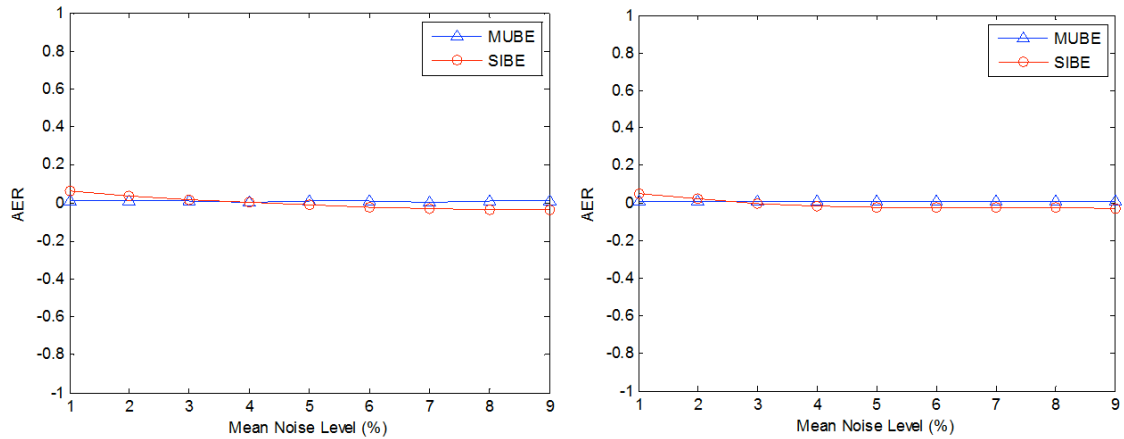
To compare the different methods on an in-vivo real case DWI data were acquired using a 3T MR scanner (Achieva, Philips Medical Systems) equipped with a gradient system providing a maximum gradient strength of 40 mT/m and with an 8-channel phased array head coil. A standard SENSE (factor=2) spin-echo EPI pulse sequence was used to acquire DWI data (b-value: 700 s/mm<sup>2</sup>; 21 diffusion directions) with the following parameters: TE/TR ~ 60/8000 ms; FOV = 215x215x85 mm<sup>2</sup>; matrix size = 172x172 with 68 slices and a spatial resolution at 1.2x1.2x1.2 mm<sup>3</sup>.

### 3.3 Noise estimation experiments

To validate the proposed noise estimation methods, an experiment consisting in estimating the level of noise after adding a known amount of Rician noise to the described synthetic phantom was performed. Both homogeneous and spatially inhomogeneous noise distributions were applied (the applied noise standard deviation ranged from 1% to 9% of the maximum b<sub>0</sub> image amplitude). We applied inhomogeneous noise field to simulate the noise patterns related to sensitivity maps. To measure noise estimation error, the Absolute mean Error Ratio (AER) was used:

$$\text{AER} = \frac{1}{|\Omega|} \sum_{v \in \Omega} |1 - (\sigma_i^e / \sigma_i^r)| \quad (9)$$

where  $\sigma_i^e$  is the local estimated noise standard deviation,  $\sigma_i^r$  is the local real noise applied,  $i$  is the local coordinate and  $\Omega$  represents the image volume. During this experiment, MUBE and SIBE methods were evaluated. The mean AER of MUBE was 0.0070, while it was 0.0276 for SIBE for homogeneous noise. The mean AER of MUBE was 0.0089 and 0.0233 for SIBE for inhomogeneous noise. As can be noticed, both methods obtained accurate results with estimation errors lower than 3% at the different noise levels (see figure 3).



**Figure 3.** In the left figure the AER results of the two proposed methods for stationary noise is shown. At the right, the same results for spatially inhomogeneous noise is presented.

### 3.4 Comparison of denoising methods

Our proposed LPCA method was compared to three recently proposed methods for MRI denoising (two of them were previously used for DWI denoising):

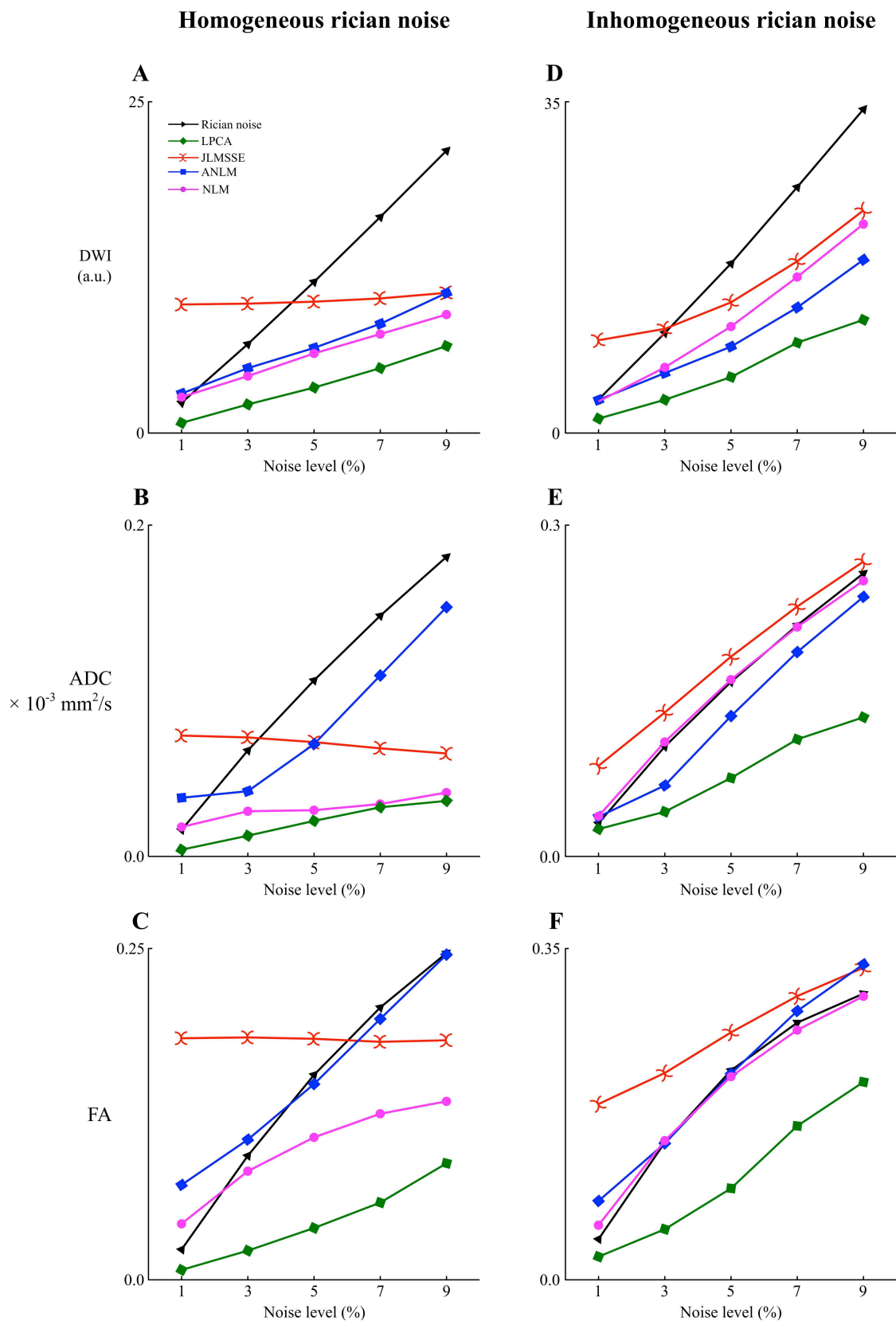
- Joint Linear Minimum Mean Squared Error (JLMMSE) [18]. In this method all the gradient images are filtered together, exploiting the first and second order information that they share.
- Non-local Means (NLM) with Rician bias correction [17]. This method is an adaptation of Optimized blockwise NLM method [40] to deal with Rician noise. This method was applied to each DWI gradient image separately, as suggested by Wiest-Daesslé et al. [17].
- Adaptive Non-local Means (ANLM) with Rician bias correction [41]. This method is an evolution of NLM method [42] that is able to deal with spatially varying noise patterns typically present on parallel imaging. Although we have no news about the use of this method for DWI denoising we decided to include this method in the comparison since this method is able to deal with spatially varying noise fields typically present on parallel acquired DWI. This method was also applied to each DWI gradient image separately.

All three methods were run with their default parameters. In the two NLM variants, we used our own implementation of the method (Matlab mex file) while in the second case we used the command JointLMMSE utility integrated in the Slicer 3.0 software [41]. The default routines available in "dtifit" (part of the FSL [44] software distribution, FMRIB, Oxford) were used to process the data. In particular, the tensor model was fitted using least squares [45]. The tensor was diagonalized and quantitative measures extracted (i.e., FA and ADC). Probabilistic tractography was performed using *bedpostx* (also part of FSL), which uses Markov Chain Monte Carlo sampling to build distributions of diffusion parameters, including that of the principal diffusion orientation, which is then used to repeatedly seed tracks that sample from such a distribution to follow particular trajectories in a stochastic fashion. 5000 tracks were initiated per seed voxel, and all defaults defined in the *bedpostx* and *probtrackx* programs were used.

### **3.5 Simulated data experiment**

The simulated phantom data was corrupted with different levels of Rician noise ranged from 1% to 9% of the maximum non-DWI signal amplitude (both stationary and spatially varying noise patterns were applied). To measure the different filter performances the root-mean-square-error (RMSE) measure was used. For the spatially stationary noise case all JLMMSE, NLM and our proposed method used the noise estimation provided by MUBE method while ANLM internally estimated the amount of noise. For the spatially varying noise case, both NLM and JLMMSE methods assume constant noise across the volume and therefore they were run using the noise level present at the background which is typically used to estimate noise level on MRI. Similarly to the stationary noise case ANLM method internally estimated the local noise level and finally our proposed method used the noise estimation provided by MUBE method. Figure 4 shows the results of the different methods for both stationary and spatially inhomogeneous noise patterns in terms of RMSE, ADC error and FA error. In all the cases the proposed LPCA method produced lower RMSE values than compared methods.

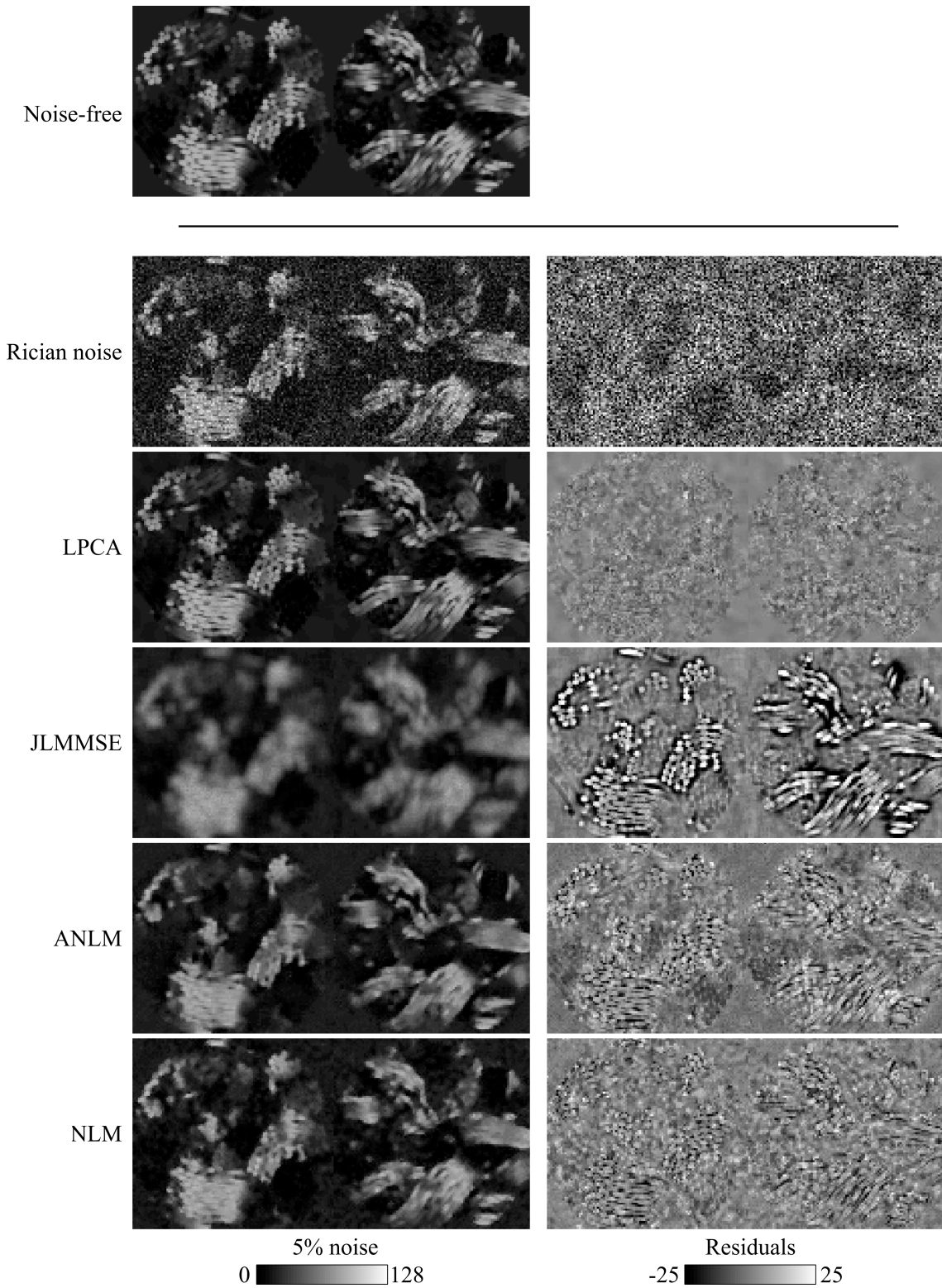




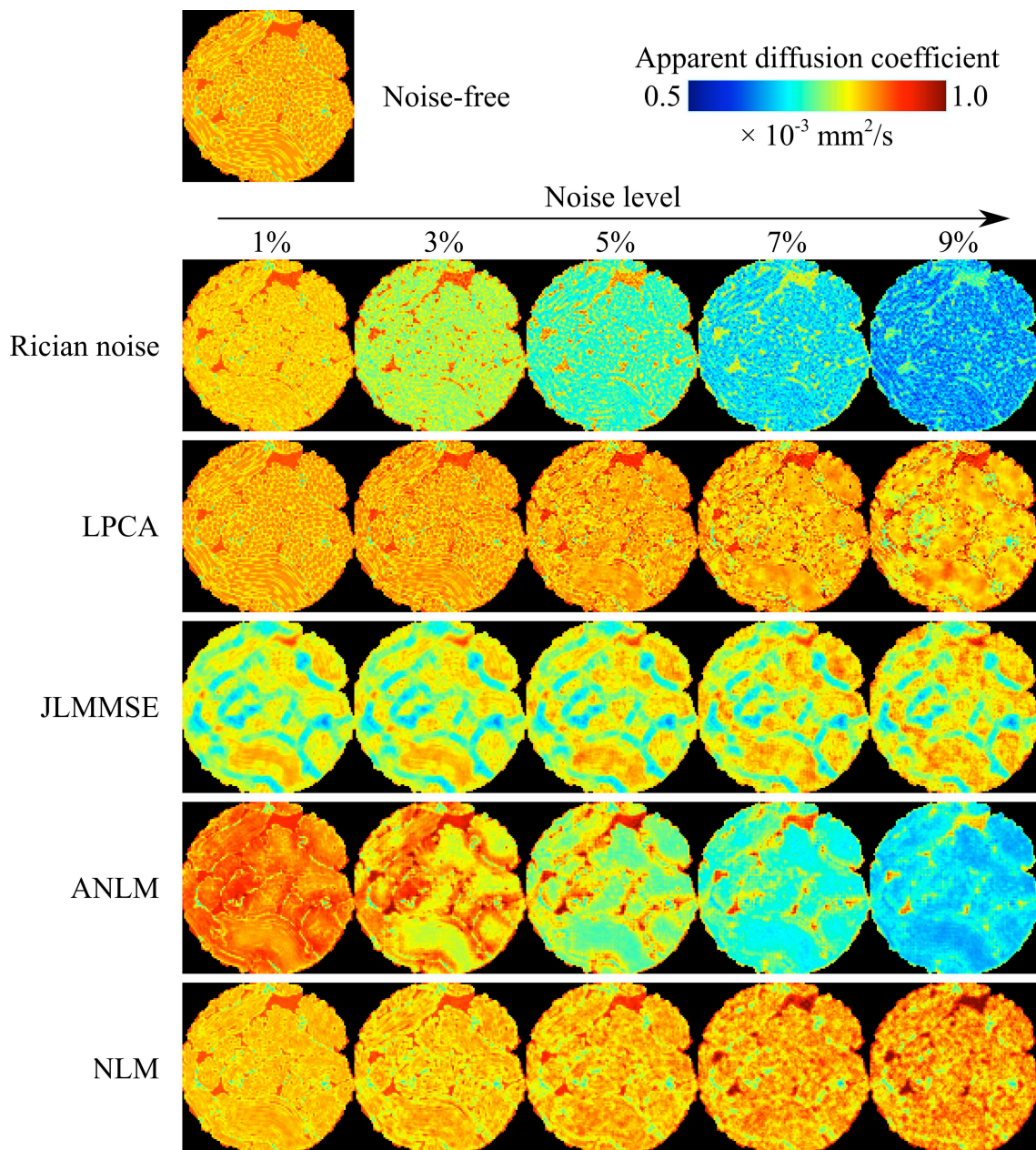
**Figure 4.** Root mean square errors obtained with the compared filters for DWI intensity (A and D), ADC (B and E) and FA (C and F). Left column: results for spatially homogeneous Rician noise. Right column: results for spatially inhomogeneous Rician noise.

In Figure 5, a visual inspection of the results is proposed for the different methods compared at 5% Rician noise. The LPCA method was able to significantly reduce noise in the images while minimally affecting the original signal. Both NLM based methods slightly over-smoothed the data while the JLMSE methods produced a more severe blurring, as it can be noticed in the image residuals (difference of the noise free and denoised images).

On the other hand, the proposed method improved the estimation of the diffusion parameters (ADC and FA) for all noise levels (Figure 4) demonstrating that the LPCA method is not only able to remove high frequency noise content but also to reduce the signal bias which significantly affects the estimated parameters. Figure 6 shows the impact of Rician noise on ADC computation. We can observe that at 9% of Rician noise, the ADC values estimated from noisy phantom decreased up to 50% compared to ADC values estimated on the noise-free phantom. This demonstrates the importance of correcting for this intensity bias introduced by Rician noise. In this visual example we can see how the proposed method is able to remove noise derived bias to recover ADC values closer to ground truth compared to other filters. JLMSE and NLM methods were able to reduce to some extent the Rician bias while ANLM method failed (probably due to an inaccurate noise estimation).



**Figure 5.** Synthetic phantom before and after denoising in the case of 5% of Rician noise. Notice the better preserved detail in LPCA-denoised images.



**Figure 6.** ADC results for the different methods compared. As can be noticed, the proposed method produces the less biased estimates of ADC values for all the studied noise levels.

### 3.6 In-vivo real data experiment

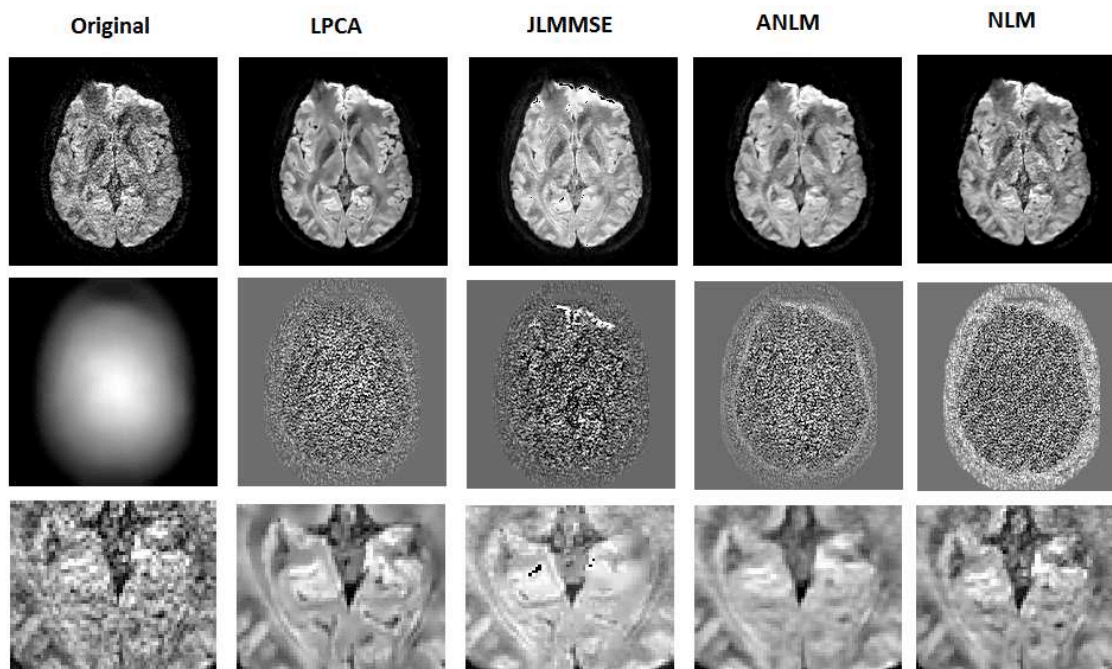
The in-vivo real dataset was filtered with the four methods compared and the resulting images were analyzed. Noise was estimated using the proposed SIBE method since only one b0 image was available for this dataset. As expected, the noise field estimated by SIBE showed a spatially varying noise pattern due to the use of a parallel imaging SENSE sequence (see Fig. 7). ANLM, NLM and JLMMSE methods were run with their default parameters. In NLM method the noise standard deviation parameter was estimated from the image background. ANLM and JLMMSE methods internally estimated the noise level.

As can be noticed in fig. 7, the proposed LPCA method produced a clear and contrasted image. The NLM method showed some residual noise due to the noise level underestimation at image central areas. In fact, this method does not take into account the spatially varying noise. The JLMMSE method showed some artifacts at frontal areas and a slightly change of contrast probably due to an inaccurate Rician noise bias correction. ANLM perform very well but seemed to over smooth some image details.

Diffusion measures were also performed to evaluate the performance of the four methods compared (fig. 8). It can be noted that LPCA and ANLM methods provided the greatest improvements, particularly evident in the FA and ADC maps. The deep brain structures in the FA map of the LPCA-denoised data still show slight noise, albeit considerably less than the original data. All denoising methods improved the diffusion measurements with the exception of JLMMSE (this could be caused by an inaccurate noise estimation since this method was not designed to deal with spatially varying noise patterns ).It is also worth to note that the uncertainty on the estimation of the principal diffusivity was drastically reduced by all denoising methods. The greatest reduction was obtained using the LPCA method.

Finally, tractography experiments were performed to evaluate the effect of the denoising algorithms on the directional information of water diffusion. Tractography of the genu of the corpus callosum, the posterior limb of the internal capsule and right crus of fornix were obtained. Tracts for the corpus

callosum and internal capsule result from seeding a single voxel, whereas the fornix was seeded in 40 voxels at the level of the body of the fornix, with an AND operator at the level of the hippocampus. The same tracking criteria were used for all data sets. Figure 9 shows tractography results of the different denoised data using the compared methods. As can be observed LPCA produced the most coherent results of all compared methods. JLMMSSE and ANLM methods produced spurious tracks not present on LPCA and NLM-denoised data.



**Figure 7.** Example results on a clinical dataset. First row-Left to right: Original noisy data, LPCA results, JLMMSSE results, ANLM results and NLM results. Second row-Left to right: Estimated noise field, LPCA residuals, JLMMSSE residuals, ANLM residuals and NLM residuals. Third row: close-up of the results of first row.

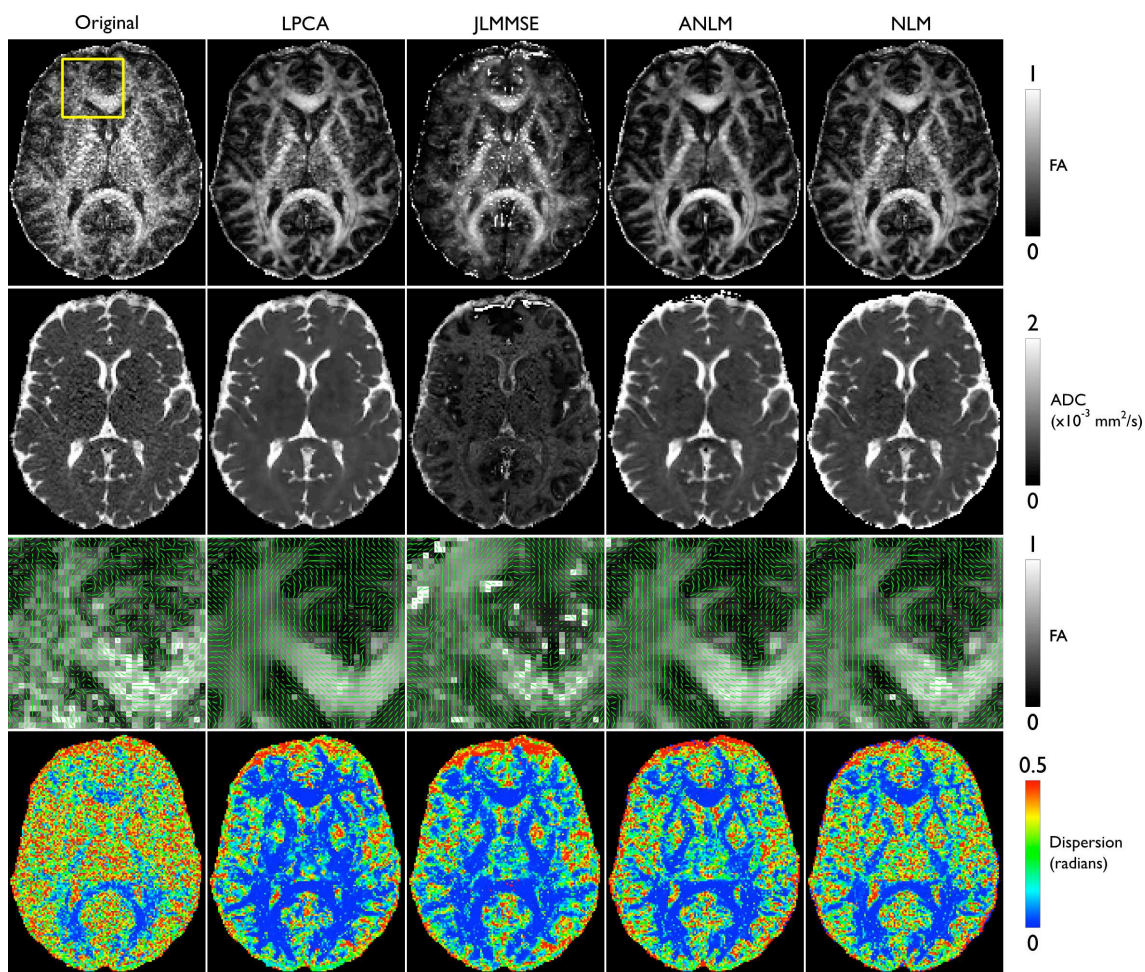


Figure 8. Diffusion tensor metrics. The original, low SNR, data yielded maps of Fractional Anisotropy (FA, first row) and Apparent Diffusion Coefficient (ADC, second row) of poor quality, which were improved by all denoising methods, with the exception of JLMSE (third column). The direction of the main diffusivity, shown as green quivers (row 3), is overlaid on the corresponding FA map, showing a much smoother and coherent pattern in the LPCA-denoised data. The uncertainty of the principal diffusivity (row 4) is drastically reduced by all denoising methods, with the greatest reduction using the LPCA filter.

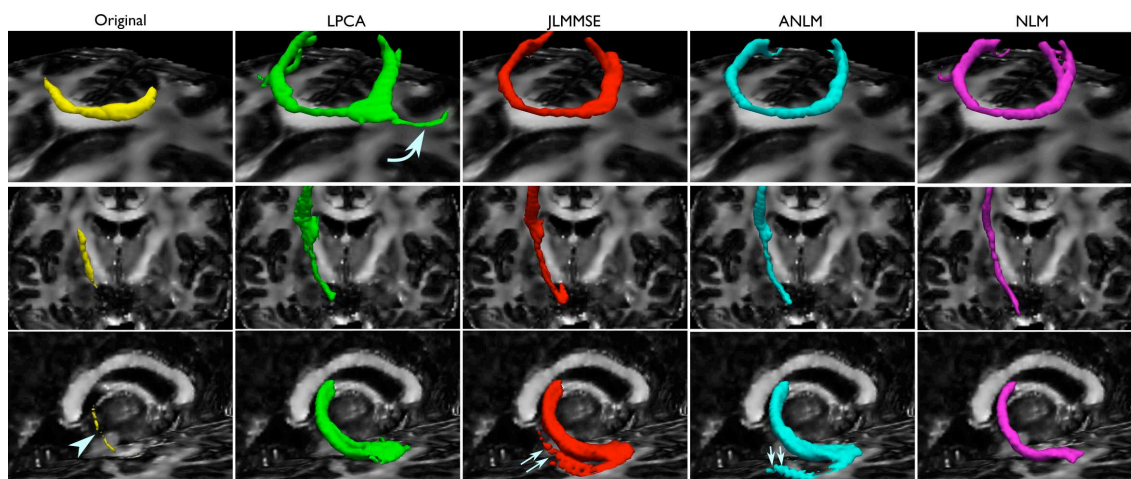


Figure 9. Tractography of the genu of the corpus callosum (first row), the posterior limb of the internal capsule (second row) and right crus of fornix (third row), overlaid on the FA map derived from the LPCA-denoised data. Isosurfaces represent those voxels that are intersected by at least 100 tracts in the case of the corpus callosum and internal capsule, and by at least 7 in the case of the fornix. Notice the limited propagation of all the tracts derived from the original, low SNR, data sets (first column), with the crus of the fornix barely visible (arrowhead). The corpus callosum was extended by all algorithms, but reached even farther in the LPCA-denoised data set (curved arrow). The crus of the fornix is extended and its width increased by all denoising methods, but JLMMSE and ANLM-denoised data sets produced spurious tracks (double arrows), which were not produced in the LPCA and NLM-denoised data sets.

Regarding the processing times of the compared methods, the fastest one was the JLMMSE with only 40 seconds, the second was the proposed LPCA method with around 3 minutes and finally NLM and ANLM methods took over 8 and 15 minutes correspondently. All experiments were run on an i7 intel machine with 16 GB RAM using Matlab 2011a and Windows 7.



## 4. Discussion

In this paper, a new method for DWI denoising is proposed and evaluated using both synthetic and real clinical data. In addition, a comparison with related state-of-the-art methods is provided. Despite its simplicity, we demonstrated improvements on DWI denoising and diffusion parameter estimation using our LPCA filter compared to previously proposed methods.

Our proposed filter removes noise in multi-directional DWI data by using an overcomplete local PCA-based decomposition. Such noise reduction capabilities can be understood due to high level of local profile redundancy which allows representing diffusion profiles using only a small number of components thus enabling to remove non signal related components (i.e. noise) efficiently. In addition, we have shown that our proposed filter not only reduces the noise present in the images, but also the bias induced by the Rician nature of the noise, thus producing diffusion parameters that better reflect the characteristics of the tissue, rather than noise-biased measurements. This has been clearly shown with the ADC and FA results where the lower parameter estimation error was coherent for all the noise levels tested. Moreover, tractography results were improved using our denoised method, considerably reducing the uncertainty of fiber directions and increasing the probability of connections between voxels within a given tract. Therefore, tracts can be reconstructed more reliably and thus the analyses of the tissue microstructure by means of tensor-derived metrics would be more sensitive.

The quantitative diffusion maps derived from the simulated data sets showed that the bias introduced by Rician noise in the estimation of quantitative diffusion parameters are minimized when denoising of the DWI is performed, with a clear advantage of the LPCA method compared to JLMME and NLM based algorithms. This was also true in the case of inhomogeneous noise, a scenario ever more commonly encountered, as parallel imaging methods are becoming commonplace. Moreover, the directional information of the diffusion data is not affected by our method, showing negligible deviation from the true principal diffusivity compared to the noise-free phantom and clearly

outperforming both the noisy phantom as well as the other denoising algorithms tested. For clinical dataset, this translated into better depictions of connectivity with a higher degree of confidence.

Our current method can be easily applied to any multi-directional DWI data set, independently of the diffusion analysis performed (e.g., DTI, HARDI, q-ball, etc...), potentially improving any quantitative measure derived from them. This has immediate benefits for the analysis of brain connectivity, as well as the study of tissue microstructure in the healthy and diseased brain.

### **Acknowledgements**

We are grateful to Dr Bassem Hiba for his help in the acquisition of the clinical DWI data and Dr. Antonio Tristan for his support running the JLMMSE method used in the comparison. This work has been supported by the Spanish grant TIN2011-26727 from Ministerio de Ciencia e Innovación. This work has been also partially supported by the French grant “HR-DTI” ANR-10-LABX-57 funded by the TRAIL from the French Agence Nationale de la Recherche within the context of the Investments for the Future program.

## References

[1] Sundgren, Q. Dong, D. Gomez-Hassan, S. K. Mukherji, P. Maly, and R. Welsh. (2004). Diffusion tensor imaging of the brain: Review of clinical applications. *Neuroradiology*, 46(5), 339–350.

[2] Johansen-Berg, H., and T. E Behrens. (2006). Just Pretty Pictures? What Diffusion Tractography Can Add in Clinical Neuroscience. *Curr Opin Neurol* 19, (4), 379-85.

[3] Jones, D.K., Basser, P.J., (2004). Squashing peanuts and smashing pumpkins: how noise distorts diffusion-weighted MR data. *Magnetic Resonance in Medicine* 52, 979–993.

[4] McGraw T, Vemuri BC, Chen Y, Rao M, Mareci T. (2004). DT-MRI denoising and neuronal fiber tracking. *Med Image Anal.*, 8(2), 95-111.

[5] McGraw T, Vemuri B, Ozarslan E., Chen Y. and Mareci T. (2009). Variational denoising of diffusion weighted MRI. *Inverse Probl. Imaging*, 3(4), 625–648.

[6] Wirestam R, Bibic A, Lätt J, Brockstedt S, Ståhlberg F. (2006). Denoising of complex MRI data by wavelet-domain filtering: Application to high-*b*-value diffusion-weighted imaging. *Magn. Reson. Med.*, 56(5), 1114–1120.

[7] Descoteaux M, Wiest-Daesslé N, Prima S, Barillot C, Deriche R. (2008) Impact of Rician adapted non-local means filtering on HARDI. *MICCAI 2008*; 11,122-130.

[8] Chen B and Hsu E. (2005). Noise removal in magnetic resonance diffusion tensor imaging. *Magn Reson Med*, 54, 393-407.

[9] Martin-Fernandez M, Muñoz-Moreno E, Cammoun L, Thiran JP, Westin CF, Alberola-López C. (2009). Sequential anisotropic multichannel Wiener filtering with Rician bias correction applied to 3D regularization of DWI data. *Med Image Anal.* 13(1), 19-35.

[10] Aja-Fernandez S., Niethammer M., Kubick M. Shenton E and Westin C.F. (2008). Restoration of DWI data using a Rician LMMSE estimator. *IEEE Trans Med Imaging*, 27,1389-1403.

[11] Basu S, Fletcher T, Whitaker R. (2006). Rician noise removal in diffusion tensor MRI. *MICCAI2006*, 9,117-25.

[12] Ding Z, Gore JC, Anderson AW. (2005). Reduction of noise in diffusion tensor images using anisotropic smoothing. *Magn Reson Med.*, 53(2),485-90.

[13] Iglesias Gonzalez J. E., Thompson Paul M., and Zhuowen Tu. (2011). Modeling diffusion-weighted MRI as a spatially variant Gaussian mixture: Application to image denoising. *Med. Phys.* 38, 4350.

[14] Hamarnesh G and Hradsky J. (2007). Bilateral filtering of diffusion tensor magnetic resonance images. *IEEE Transactions on Image Processing*, 16(10), 2463-2475.

[15] Xu Q , Anderson A, Gore J, and Ding Z. (2010). Efficient anisotropic filtering of diffusion tensor images. *Magnetic Resonance Imaging*, 28, 200-211.

[16] Parker G. J. M., Schnabel J. A., Symms M. R., Werring D. J., and Barker G. J. (2000). Nonlinear smoothing for reduction of systematic and random errors in diffusion tensor imaging. *J Magn Reson Imag*, 11(6), 702–710.

[17] Wiest-Daesslé N, Prima S, Coupé P, Morrissey SP, Barillot C. (2007). Non-local means variants for denoising of diffusion-weighted and diffusion tensor MRI. *MICCAI2007*, 10, 344-351.

[18] Tristán-Vega A, Aja-Fernández S. (2010). DWI filtering using joint information for DTI and HARDI. *Med Image Anal.*, 14(2), 205-218.

[19] Weickert J, Brox T. (2002). Diffusion and regularization of vector and matrix valued images. Saarland Department of Mathematics, Saarland University.

[20] Wang, Z., Vemuri, B.C., Chen, Y., Mareci, T., (2004). A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from DWI. *IEEE transactions on Medical Imaging*, 23, 930-939.

[21] Reisert M and Kiselev V.G. (2011). Fiber Continuity: An Anisotropic Prior for ODF Estimation. *IEEE Transaction on Medical Imaging*, 30(6),1274-1283

[22] Goh A, Lenglet C, Thompson P.M. and Vidal R. (2009). Estimating Orientation Distribution Functions with Probability Density Constraints and Spatial Regularity. MICCAI2009, 1, 877-885.

[23] Fillard, P., Pennec, X., Arsigny, V., Ayache, N., (2007). Clinical DT-MRI estimation, smoothing, and fiber tracking with log-Euclidean metrics. IEEE Transactions on Medical Imaging, 26, 1472–1482.

[24] Muresan D. and Parks, T.W. (2003). Adaptive principal components and image denoising. ICIP03, 101-104.

[25] Bydder M, Du J. (2003). Noise reduction in multiple-echo data sets using singular value decomposition. Magn Reson Imaging, 24(7), 849-856.

[26] Poon P.K., Wei-Ren Ng, Sridharan V. (2009). Image Denoising with Singular Value Decompositon and Principal Component Analysis. <http://www.u.arizona.edu/~ppoon/ImageDenoisingWithSVD.pdf>

[27] Zhang L, Dong W, Zhang D., Shi G. (2010). Two-stage image denoising by principal component analysis with local pixel grouping. Pattern Recognition, 43:1531-1549.

[28] Deledalle C, Salmon J, Dalalyan A. (2011). Image denoising with patch based PCA: local versus global. BMVC2011.

[29] Manjón JV, Thacker N, Lull JJ, Garcia-Martí G, Martí-Bonmatí L, Robles M. Multicomponent MR Image Denoising. (2009). International Journal of Biomedical imaging, Article ID 756897.

[30] Lijun Bao, Marc Robini, Wanyu Liu, Yuemin Zhu. 2013. Structure-adaptive sparse denoising for diffusion-tensor MRI. Medical Image Analysis, 17(4): 442–457.

[31] Fan Lam, S. Derin Babacan, Justin P. Haldar, Michael W. Weiner, Norbert Schuff, Zhi-Pei Liang. 2013. Denoising diffusion-weighted magnitude MR images using rank and edge constraints. Magnetic Resonance in Medicine, 69(6):1-13.

[32] Katkovnik V., Foi A., Egiazarian K., and Astola J. (2010). From local kernel to nonlocal Multiple-Model image denoising. Int. J. Comput. Vision, 86(1),1–32.

[33] Manjón JV, Coupé P, Buades A., Collins D.L., Robles M. (2012). New Methods for MRI Denoising based on Sparseness and Self-Similarity. Medical Image Analysis, 16(1), 18-27.

[34] Coifman R. and Donoho D. L. (1995). Translation Invariant Denoising, Wavelets and Statistics. Anestis Antoniadis, ed. Springer Verlag Lecture Notes.

[35] Nowak, R.D. (1999). Wavelet-based Rician noise removal for magnetic resonance imaging. *IEEE Transactions on Image Processing* 8(10), 1408–1419.

[36] Koay CG, Basser PJ. (2006). Analytically exact correction scheme for signal extraction from noisy magnitude MR signals. *J Magn Reson*, 179,317–322.

[37] Coupé P, Manjón J.V., Gedamu E., Arnold D., Robles M., Collins D. L. (2010). Robust Rician Noise Estimation for MR Images. *Medical Image Analysis*, 14(4), 483–493.

[38] Landman B, Bazin P.L., Prince J.L. (2009). Estimation and application of spatially variable noise fields in diffusion tensor imaging. *Magnetic Resonance Imaging*, 27(6), 741-751.

[39] Close, T.G., Tournier, J.-D., Calamante, F., Johnston, L.A., Mareels, I., Connelly, A., (2009). A software tool to generate simulated white matter structures for the assessment of fibre tracking algorithms. *Neuroimage* 47, 1288–1300.

[40] Coupé P, Yger P, Prima S, Hellier P, Kervrann C, Barillot C. (2008). An Optimized Blockwise NonLocal Means Denoising Filter for 3-D Magnetic Resonance Images. *IEEE Transactions on Medical Imaging*, 27(4),425–441.

[41] Manjón JV, Coupé p, Martí-Bonmatí I, Robles M, Collins L.(2010) Adaptive Non-Local Means Denoising of MR Images with Spatially Varying Noise Levels. *Journal of Magnetic Resonance Imaging*, 31,192-203, 2010



[42] Coupé P, Hellier P., Prima S., Kervrann C, Barillot C. (2008). 3D Wavelet Subbands Mixing for Image Denoising. *International Journal of Biomedical Imaging*.

[43] Gering D.T., Nabavi A., Kikinis R., Grimson W..E.L., Hata N., Everett P., Jolesz F.A., Wells III W.M. (1999). An Integrated Visualization System for Surgical Planning and Guidance using Image Fusion and Interventional Imaging. *MICCAI1999*, 2, 809-819.

[424] Smith S.M, Jenkinson M, Woolrich M.W., Beckmann C.F., Behrens T.E.J., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, 23(1):208-19.

[45] Basser PJ, Mattiello J, Le Bihan D. (1994). Estimation of the effective self-diffusion tensor from the NMR spin-echo. *J Magn Reson* ;103:247–254

[46] Strang, G. (1976). *Linear Algebra and Its Applications Academic*. New York,19802.

[47] Jolliffe, I. T. (1986). *Principal component analysis (Vol. 487)*. New York: Springer-Verlag.