A wavelet-based regularized reconstruction algorithm for SENSE parallel MRI with applications to neuroimaging

Lotfi Chaâri[a,c], Jean-Christophe Pesquet[a], Amel Benazza-Benyahia[b], Philippe Ciuciu[c]

[a] *LIGM (UMR-CNRS 8049), Université Paris-Est, Champs-sur-Marne, 77454 Marne-la-Vallée cedex, France.*
{chaari,pesquet}@univ-mlv.fr.
[b] *Ecole Supérieure des Communications de Tunis (SUP'COM-Tunis), URISA, Cité Technologique des Communications, 2083, Tunisia.*
benazza.amel@supcom.rnu.tn
[c] *CEA/DSV/I^{2}BM/Neurospin, CEA Saclay, Bat. 145, Point Courrier 156, 91191 Gif-sur-Yvette cedex, France.*
philippe.ciuciu@cea.fr

Abstract

To reduce scanning time and/or improve spatial/temporal resolution in some Magnetic Resonance Imaging (MRI) applications, parallel MRI acquisition techniques with multiple coils acquisition have emerged since the early 1990's as powerful imaging methods that allow a faster acquisition process. In these techniques, the full FOV image has to be reconstructed from the resulting acquired undersampled k-space data. To this end, several reconstruction techniques have been proposed such as the widely-used SENSitivity Encoding (SENSE) method. However, the reconstructed image generally presents artifacts when perturbations occur in both the measured data and the estimated coil sensitivity profiles. In this paper, we aim at achieving accurate image reconstruction under degraded experimental conditions (low magnetic field and high reduction factor), in which neither the SENSE method nor the Tikhonov regularization in the image domain give convincing results. To this end, we present a novel method for SENSE-based reconstruction which

[☆] Part of this work was presented at the ISBI conference in 2008 [1].
This work was supported by Région Ile de France, and the Agence Nationale de la Recherche under grant ANR-05-MMSA-0014-01.

proceeds with regularization in the complex wavelet domain by promoting sparsity. The proposed approach relies on a fast algorithm that enables the minimization of regularized non-differentiable criteria including more general penalties than a classical ℓ_1 term. To further enhance the reconstructed image quality, local convex constraints are added to the regularization process. In vivo human brain experiments carried out on Gradient-Echo (GRE) anatomical and Echo-Planar Imaging (EPI) functional MRI data at 1.5 Tesla indicate that our algorithm provides reconstructed images with reduced artifacts for high reduction factors.

Key words: Parallel MRI, SENSE, Reconstruction, Regularization, Sparsity, Wavelet Transform, Convex Optimization, Proximal Methods.

1. Introduction

Reducing global imaging time is of primary interest in many Magnetic Resonance Imaging (MRI) applications like neuroimaging [2], cardiac [3] and abdominal [4] imaging since it allows clinicians to limit the patient's exposition to the MRI environment. To achieve this goal without degrading image quality significantly, parallel MRI (pMRI) systems have been developed in the 1990's. In such systems, multiple receiver surface coils with complementary sensitivity profiles located around the underlying object are employed to simultaneously collect MRI data in the frequency domain (i.e the k-space). In order to speed up the acquisition, the data are sampled at a frequency rate R times lower than the Nyquist sampling rate along at least one spatial direction, i.e. usually the phase encoding direction. A reconstruction step is then performed to build up a full Field of View (FOV) image by unfolding the undersampled coil-specific data. This pMRI reconstruction is a challenging task because of the low Signal to Noise Ratio (SNR) due to three main artifact sources: aliasing artifacts related to the undersampling rate, acquisition noise and also errors in the estimation of coil sensitivity maps. Using parallel imaging may allow one to reduce the global acquisition time, increase the spatial resolution while maintaining the Time-of-Repetition (TR) to a fixed value, or improve the acquisition SNR using multishot acquisitions with fewer TRs per image. Hence, in imaging sequences involving a single volume acquisition, the global imaging time is directly impacted: acquiring a high resolution T_1-weighted image takes about 9mn using convential settings and only 5mn (respectively, 3mn) using $R = 2$ (resp., $R = 4$) in pMRI. In

dynamic imaging like functional MRI (fMRI), the TR reduction is usually exploited in another direction, namely the improvement of spatial or temporal resolution instead of the decrease of the global scanning time. One underlying reason is that the Echo Planar Imaging (EPI) sequence on which fMRI data acquisition relies has low spatial resolution in non-parallel imaging (eg, 64×64 pixels by slice). In that case, pMRI enables high resolution imaging up to 128×128 for the same TR provided that a high enough R-factor is used.

The Simultaneous Acquisition of Spatial Harmonics (SMASH) [5], introduced by Sodickson and Manning in 1997, was the first reconstruction method operating in the k-space. It uses a linear combination of pre-estimated coil sensitivity maps to generate the missing phase encoding steps. Some other k-space based reconstruction techniques have also been proposed like GRAPPA (Generalized Autocalibrating Partially Parallel Acquisitions) [6] or generalizations of the SMASH method (AUTO-SMASH) [7, 8]. The specificity of SMASH is that it needs a separate coil sensitivity map estimation in contrast with these most recent generalizations, which proceed by acquiring a few additional k-space lines to derive coil sensitivity maps without a reference scan. This explains why these extensions are referred to as *autocalibrated*.

Indeed, GRAPPA and its alternatives may be preferred to non-autocalibrated methods when accurate coil sensitivity maps may be difficult to extract. This occurs when reference scans are difficult to acquire either because of limited global imaging time or spatial resolution changes. For instance, in lung and abdomen imaging, the numerous inhomogeneous regions with a low spin density make inaccurate the estimation of the sensitivity information. Reference scans may also appear inappropriate in other circumstances, typically when they are not able to account for artifacts that will appear later during the parallel imaging sequence. This is especially critical in dynamic imaging in case of motion artifacts. However, all the reported methods may suffer from phase cancellation problems, low SNR during the acquisition process and limited reconstruction quality.

In [9], an alternative reconstruction method called SENSE has been introduced. In its simplified form, which relies on Cartesian k-space sampling, SENSE is a two-step procedure involving first a reconstruction of reduced FOV images and second a spatial unfolding technique, which amounts to a weighted least squares estimation. This technique requires a precise estimation of coil sensitivity maps using a reference scan (usually a 2D Gradient-Echo (GRE)). Note that for non-Cartesian sampling schemes, like the spiral

one, SENSE reconstruction is more complicated and requires iterative reconstruction [10].

To the best of our knowledge, in actual clinical daily routines, only GRAPPA and SENSE-like algorithms are available on scanners: Siemens provides GRAPPA and mSENSE techniques which roughly correspond to ASSET and ARC methods on General Electric Healthcare machines. For a general overview of reconstruction methods in pMRI the reader may refer to [11].

In this paper, we focus on the SENSE imaging with a uniform Cartesian sampling. As reported hereabove, in addition to neuroimaging, many other applications like cardiac imaging benefit from the enhanced image acquisition abilities of this method. SENSE as well as alternative methods are usually supposed to achieve perfect reconstruction in case of noiseless data and perfect coil sensitivity maps knowledge. However, in practice, inaccuracies in the estimation of coil sensitivity maps (especially at the center of the FOV) and mainly at high reduction factors make the reconstruction problem ill-posed and the use of SENSE limited at conventional magnetic field strengths (1.5 T). To overcome this limitation, some alternatives have been proposed like the optimization of the coil geometry [12] and the improvement of coil sensitivity profile estimation [13]. However, at low magnetic fields, the maximal reduction factor usually employed is $R = 2$ because the reconstructed images using larger R values are affected by severe aliasing artifacts. If one wants to further improve the spatial or temporal resolution, or simply reduce the global imaging time keeping the same acquistion parameters, it would be necessary to increase factor R. To compensate for the intrinsic degradation of the image quality, it is crucial to regularize the reconstruction process which becomes severely ill-posed. To this end, several contributions have been proposed in the recent pMRI literature [14, 15, 16, 17, 18, 2, 1], most of them operating in the image domain to better estimate full FOV images. For an introductory survey to linear inverse problems with a special emphasis to pMRI reconstruction, the interested reader is referred to [19].

Standard regularization can improve the reconstructed image quality when the experimental conditions are not too degraded. However, under a low magnetic field intensity with a high reduction factor, the problem becomes much more difficult. In this paper, we focus on such extreme conditions and develop alternative strategies to classical regularization schemes. In short, we no longer assume that the sought image is Gaussian a priori, as done implicitly in Tikhonov regularization. The empirical histogram of an MRI

image may actually be multimodal thus departing largely from Gaussianity. Akin to [1], we make use of the wavelet transform because of the sparse representations they generate for regular images. Then, we derive sparsity promoting priors in the wavelet domain making feasible the regularization in this space before reconstructing the full FOV image by applying the inverse wavelet transform. We show that the ensuing regularized reconstruction method removes stronger artifacts and thus remains applicable even in severe experimental conditions (L_{12}UWR-SENSE method[1]). This is really possible when the artifact features are exploited in the sense that convex constraints are introduced to bound artifact values. We will also show that the choice of wavelet representations has been motivated by the ability to employ tractable statistical models and the emergence of fast convex non differentiable optimization methods.

The methodological contributions of this paper are twofold:

- the design of a novel wavelet-based inversion technique allowing us to incorporate useful constraints on the images to be reconstructed (L_{12}CWR-SENSE method[2]),

- the proposition of an efficient convex optimization algorithm that minimizes the related regularized criterion (Forward-Backward algorithm with Douglas-Rachford sub-iteration). This algorithm has been adapted to the pMRI framework (complex valued data) and requires the computation of some parameters, noticeably a Lipschitz constant.

The rest of this paper is organized as follows. In Section 2, we give a brief overview of the SENSE method and its regularized version in the image domain. Section 3 describes the methodology of our regularization scheme in the wavelet domain and provides the reader with details on efficient convex optimization algorithms we use for minimizing the regularized criteria. Results on real anatomical and functional data are reported in Section 4. In particular, a fair comparison of our L_{12}UWR-SENSE and L_{12}CWR-SENSE methods with SENSE and L_2R-SENSE (Tikhonov regularization) is provided both in terms of qualitative and quantitative evaluation criteria. It is shown

[1]This acronym stands for the Unconstrained version of the ℓ_{12} Regularization in the Wavelet domain applied to SENSE imaging.

[2]This acronym stands for the Constrained version of the ℓ_{12} Regularization in the Wavelet domain applied to SENSE imaging.

that we successively improve the reconstruction quality by upgrading from SENSE to L_{12}UWR-SENSE and from L_{12}UWR-SENSE to L_{12}CWR-SENSE at the expense of a higher numerical complexity. In Section 5, we discuss the pros and cons of the proposed framework and propose future research lines to investigate how it could be extended to: *i)* automatically tune the hyper-parameters from the original data; *ii)* handle the 2D-SENSE extension and *iii)* perform 3D+time or 4D reconstruction in case of fMRI acquisitions. Perspective on clinical applications are also mentioned.

2. Background

For two-dimensional MRI sequences, the acquisition is performed slice by slice either in sequential or interleaved order. Hence, parallel MRI also proceeds with a slice by slice acquisition to get all the 3D volume. In the SENSE imaging context, this 2D reconstuction is often referred to as 1D-SENSE, as opposed to 2D-SENSE which involves parallel reconstruction in two directions (e.g., encoding phase and partition directions). In what follows, we are therefore only interested in reconstructing a given 2D slice.

2.1. Basic 1D-SENSE reconstruction

An array of L coils is employed to measure the spin density $\overline{\rho}$ into the object under investigation.[3] The signal $\widetilde{d}_\ell$ received by each coil ℓ $(1 \leq \ell \leq L)$ is the Fourier transform of the desired 2D field $\overline{\rho} \in \mathbb{R}^{Y \times X}$ on the specified Field-of-View (FOV) weighted by the coil sensitivity profile s_ℓ, evaluated at some locations $\boldsymbol{k}_r = (k_y, k_x)^\top$ in the k-space:

$$\widetilde{d}_\ell(\boldsymbol{k}_r) = \int \overline{\rho}(\boldsymbol{r}) s_\ell(\boldsymbol{r}) e^{-\imath 2\pi \boldsymbol{k}_r^\top \boldsymbol{r}} \, d\boldsymbol{r} + \widetilde{n}_\ell(\boldsymbol{k}_r) \tag{1}$$

where $\widetilde{n}_\ell(\boldsymbol{k}_r)$ is an additive zero-mean Gaussian noise which is correlated between coils at a given location $\boldsymbol{k}_r$, but independent and identically distributed (iid) in the k-space, and $\boldsymbol{r} = (y, x)^\top \in Y \times X$ is the spatial position in the image domain[4]. The size of the acquired data $\widetilde{d}_\ell$ in the k-space clearly depends on the sampling scheme. For the sake of simplicity, a Cartesian

[3] The overbar is used to distinguish the "true" data from a generic variable.

[4] Y and X denote the size of the FOV along the phase and frequency encoding directions, respectively.

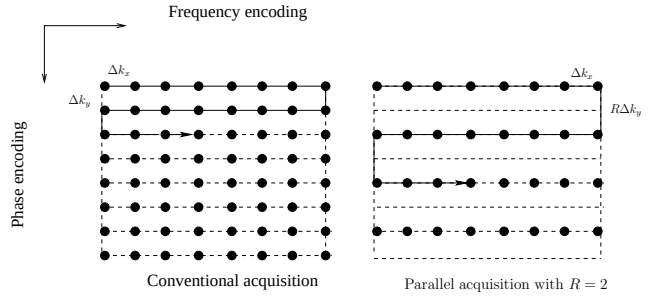

Figure 1: Sampling schemes for conventional and parallel acquisitions on a Cartesian grid. Here $R = 2$ indicates a subsampling along the phase encoding direction (Y-axis) by a factor of two.

coordinate system is generally adopted in the neuroimaging context. In this case, Fig. 1 clearly illustrates the differences between conventional and parallel acquisitions in terms of k-space sampling. In parallel MRI, the sampling period along the phase encoding direction is R times larger than the one used for conventional acquisition, $R \leq L$ being the reduction factor.

In its simplest form, SENSE imaging amounts to solving a one-dimensional inversion problem due to the separability of the Fourier transform. Note however that this inverse problem admits a two-dimensional extension in 3D imaging sequences like Echo Volume Imaging (EVI) [2] where undersampling occurs in two k-space directions. Let y and x be the positions in the image domain along the phase and frequency encoding directions, respectively. A 2D inverse Fourier transform allows us to recover the measured signal in the spatial domain. By accounting for the undersampling of the k-space by R, the inverse Fourier transform gives us the spatial counterpart of Eq. (1) in matrix form:

$$\boldsymbol{d}(\boldsymbol{r}) = \boldsymbol{S}(\boldsymbol{r})\overline{\boldsymbol{\rho}}(\boldsymbol{r}) + \boldsymbol{n}(\boldsymbol{r}), \tag{2}$$

where

$$\boldsymbol{S}(\boldsymbol{r}) \triangleq \begin{pmatrix} s_1(y,x) & \dots & s_1(y+(R-1)\frac{Y}{R},x) \\ \vdots & \vdots & \vdots \\ s_L(y,x) & \dots & s_L(y+(R-1)\frac{Y}{R},x) \end{pmatrix},$$

$$\overline{\boldsymbol{\rho}}(\boldsymbol{r}) \triangleq \begin{pmatrix} \overline{\rho}(y,x) \\ \overline{\rho}(y+\frac{Y}{R},x) \\ \vdots \\ \overline{\rho}(y+(R-1)\frac{Y}{R},x) \end{pmatrix}, \quad \boldsymbol{d}(\boldsymbol{r}) \triangleq \begin{pmatrix} d_1(y,x) \\ d_2(y,x) \\ \vdots \\ d_L(y,x) \end{pmatrix} \tag{3}$$

$$\text{and} \quad \boldsymbol{n}(\boldsymbol{r}) \triangleq \begin{pmatrix} n_1(y,x) \\ n_2(y,x) \\ \vdots \\ n_L(y,x) \end{pmatrix}.$$

In Eq. (2), each coil-dependent noise n_ℓ is a spatially iid circular zero-mean complex-valued Gaussian process defined over the $Y \times X$ FOV. However, the sequence $\boldsymbol{n}$ is identically distributed and spatially independent, but it is not independent at a given location $\boldsymbol{r}$ in the sense that its components $(n_\ell(\boldsymbol{r}))_\ell$ are correlated, which also reads $\boldsymbol{n}(\boldsymbol{r}) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Psi})$, where $\boldsymbol{\Psi}$ is the between-coil $L \times L$ covariance matrix [9, 20]. In practice, $\boldsymbol{\Psi}$ is estimated by acquiring L images $(\underline{\mathrm{d}}_\ell)_{1 \leq \ell \leq L}$ from all coils without radio frequency pulses, and its generic entry $\Psi(\ell_1, \ell_2)$ corresponding to the covariance between the two coils ℓ_1 and ℓ_2 is given by:

$$\Psi(\ell_1, \ell_2) = \frac{1}{Y \times X} \sum_{(y,x)} \underline{\mathrm{d}}_{\ell_1}(y,x) \underline{\mathrm{d}}^*_{\ell_2}(y,x), \ \forall (\ell_1, \ell_2) \in \{1, \ldots, L\}^2 \tag{4}$$

where $(\cdot)^*$ stands for the complex conjugate.

Therefore, the reconstruction step consists in inverting Eq. (2) and recovering $\overline{\boldsymbol{\rho}}(\boldsymbol{r})$ from $\boldsymbol{d}(\boldsymbol{r})$ at spatial positions $\boldsymbol{r} = (y,x)^\top$. Note that the coil images $(d_\ell)_{1 \leq l \leq L}$ as well as the sought image $\overline{\rho}$ are complex-valued, although $|\overline{\rho}|$ is only considered for visualization purposes.

The standard 1D-SENSE reconstruction method [9] is nothing but the Maximum Likelihood (ML) estimate, which amounts to minimizing a Weighted Least Squares (WLS) criterion from a deterministic viewpoint given the Gaussian noise assumption. The minimizer $\widehat{\boldsymbol{\rho}}_{\mathrm{WLS}}(\boldsymbol{r})$ at each spatial location $\boldsymbol{r}$ reads:

$$\begin{aligned} \widehat{\boldsymbol{\rho}}_{\mathrm{WLS}}(\boldsymbol{r}) &= \underset{\boldsymbol{\rho}(\boldsymbol{r}) \in \mathbb{C}^L}{\arg\min}\, \mathcal{J}_{\mathrm{WLS}}(\boldsymbol{\rho}(\boldsymbol{r})) = \underset{\boldsymbol{\rho}(\boldsymbol{r}) \in \mathbb{C}^L}{\arg\min} \parallel \boldsymbol{d}(\boldsymbol{r}) - \boldsymbol{S}(\boldsymbol{r})\boldsymbol{\rho}(\boldsymbol{r}) \parallel^2_{\boldsymbol{\Psi}^{-1}} \\ &= \left(\boldsymbol{S}^{\mathsf{H}}(\boldsymbol{r})\boldsymbol{\Psi}^{-1}\boldsymbol{S}(\boldsymbol{r})\right)^\sharp \boldsymbol{S}^{\mathsf{H}}(\boldsymbol{r})\boldsymbol{\Psi}^{-1}\boldsymbol{d}(\boldsymbol{r}), \end{aligned} \tag{5}$$

where $(\cdot)^{\mathsf{H}}$ (resp. $(\cdot)^\sharp$) stands for the transposed complex conjugate (resp. pseudo-inverse) and, $\parallel \cdot \parallel_{\boldsymbol{\Psi}^{-1}} = \sqrt{(\cdot)^{\mathsf{H}}\boldsymbol{\Psi}^{-1}(\cdot)}$ defines a norm on $\mathbb{C}^L$. In

practice, the performance of the 1D-SENSE method is limited because of the presence of *i)* distortions in the measurements $\boldsymbol{d}(\boldsymbol{r})$, *ii)* the putative ill-conditioning of $\boldsymbol{S}(\boldsymbol{r})$ specifically at locations $\boldsymbol{r}$ close to the image center and *iii)* the presence of errors in the estimation of $\boldsymbol{S}(\boldsymbol{r})$ mainly at brain air interfaces. These undesirable effects are illustrated in Fig. 3 (1D-SENSE reconstruction), which shows some aliasing artifacts in the reconstructed images for two values of the reduction factor: $R = 2$ and $R = 4$.

2.2. Tikhonov regularization

To improve the robustness of the solution to this ill-posed problem, Tikhonov or quadratic regularization is usually applied. As shown in [15, 14, 21], it consists in computing $\widehat{\boldsymbol{\rho}}_{\mathrm{PWLS}}(\boldsymbol{r})$ as the minimizer of a Penalized Weighted Least Squares (PWLS) criterion:

$$\begin{aligned}\widehat{\boldsymbol{\rho}}_{\mathrm{PWLS}}(\boldsymbol{r}) &= \underset{\boldsymbol{\rho}(\boldsymbol{r})\in\mathbb{C}^L}{\arg\min}\, \mathcal{J}_{\mathrm{PWLS}}(\boldsymbol{\rho}(\boldsymbol{r})) \\ &= \underset{\boldsymbol{\rho}(\boldsymbol{r})\in\mathbb{C}^L}{\arg\min}\big[\mathcal{J}_{\mathrm{WLS}}(\boldsymbol{\rho}(\boldsymbol{r})) + \kappa \parallel \boldsymbol{\rho}(\boldsymbol{r}) - \boldsymbol{\rho}_{\mathrm{r}}(\boldsymbol{r}) \parallel^2_{\boldsymbol{I}_R}\big], \end{aligned} \tag{6}$$

where $\boldsymbol{I}_R$ is the R-dimensional identity matrix and $\kappa > 0$. Interestingly, $\widehat{\boldsymbol{\rho}}_{\mathrm{PWLS}}$ matches the Maximum A Posterior (MAP) estimator in the Bayesian framework provided that the following separable complex circular Gaussian prior is considered on $\boldsymbol{\rho}(\boldsymbol{r})$: $\mathcal{N}(\boldsymbol{\rho}_{\mathrm{r}}(\boldsymbol{r}), \frac{1}{\sqrt{\kappa}}\boldsymbol{I}_R)$.

The regularization parameter κ ensures a balance between the closeness to the data and the penalty term, which controls the deviation from a given reference vector $\boldsymbol{\rho}_{\mathrm{r}}(\boldsymbol{r})$. The solution $\widehat{\boldsymbol{\rho}}_{\mathrm{PWLS}}(\boldsymbol{r})$ is given by:

$$\widehat{\boldsymbol{\rho}}_{\mathrm{PWLS}}(\boldsymbol{r}) = \boldsymbol{\rho}_{\mathrm{r}}(\boldsymbol{r}) + \big(\boldsymbol{S}^{\mathsf{H}}(\boldsymbol{r})\boldsymbol{\Psi}^{-1}\boldsymbol{S}(\boldsymbol{r}) + \kappa\boldsymbol{I}_R\big)^{-1}\boldsymbol{S}^{\mathsf{H}}(\boldsymbol{r})\boldsymbol{\Psi}^{-1}\big(\boldsymbol{d}(\boldsymbol{r}) - \boldsymbol{S}(\boldsymbol{r})\boldsymbol{\rho}_{\mathrm{r}}(\boldsymbol{r})\big).$$

Note that the accuracy of $\widehat{\boldsymbol{\rho}}_{\mathrm{PWLS}}(\boldsymbol{r})$ depends on the reference $\boldsymbol{\rho}_{\mathrm{r}}(\boldsymbol{r})$ and the choice of the regularization parameter κ. On the one hand, if κ tends to zero, the solution $\widehat{\boldsymbol{\rho}}_{\mathrm{PWLS}}(\boldsymbol{r})$ converges to $\widehat{\boldsymbol{\rho}}_{\mathrm{WLS}}(\boldsymbol{r})$. On the other hand, for large κ values the penalty term dominates in $\mathcal{J}_{\mathrm{PWLS}}$ and we get $\widehat{\boldsymbol{\rho}}_{\mathrm{PWLS}}(\boldsymbol{r}) \to \boldsymbol{\rho}_{\mathrm{r}}(\boldsymbol{r})$. To select a relevant value for κ, some works have proposed to resort to the discrepancy principle or the L-curve technique [2, 22, 23, 16, 17, 18] in order to optimize the trade-off between noise removal and contrast reduction. Other contributions have promoted the ML estimator using an Expectation-Maximization (EM) algorithm [24, 25].

Quadratic regularization is actually known to produce smoothing effects when a non-diagonal prior covariance or weighting matrix is involved in the

penalization. However, a separable quadratic penalization like in Eq. (6) may also induce such effects depending on the reference vector $\boldsymbol{\rho}_{\mathrm{r}}(\boldsymbol{r})$. To overcome this limitation and preseve tiny structures, one usually resorts to edge-preserving penalty terms [26, 27, 28, 29, 30, 31, 32].

2.3. Total variation regularization

Generally, edge-preserving penalizations take place in the image domain and make the regularization more efficient by limiting blurring effects and preserving the image boundaries. One typical example of such penalization is the Total Variation (TV) penalty. TV regularization has already been used in the MRI literature, especially for removing artifacts occuring in fast acquisition schemes [33] or parallel imaging [30, 31, 32, 34, 35, 18]. The ensuing global TV-regularized criterion reads as follows:

$$\begin{aligned}\widehat{\rho}_{\mathrm{TV}} &= \underset{\rho\in\mathbb{C}^{Y\times X}}{\arg\min}\, \mathcal{J}_{\mathrm{TV}}(\rho) \\ &= \underset{\rho\in\mathbb{C}^{Y\times X}}{\arg\min}\, \mathcal{J}_{\mathrm{L}}(\rho) + \kappa\|\rho\|_{\mathrm{TV}}, \end{aligned} \tag{7}$$

where $\|\cdot\|_{\mathrm{TV}}$ is the TV norm [31, 35] and

$$\mathcal{J}_{\mathrm{L}}(\rho) = \sum_{\mathbf{r}\in\{1,\ldots,Y/R\}\times\{1,\ldots,X\}} \mathcal{J}_{\mathrm{WLS}}(\boldsymbol{\rho}(\mathbf{r})). \tag{8}$$

Note that unlike in Eq. (6), the regularization term here is no longer separable in $\boldsymbol{r}$.

It has been reported in some recent works in the regularization literature that TV performs well especially on piecewise smooth images. It has also been observed that TV regularization may sometimes lead to some noise increase. Hence, under severe experimental conditions (at 1.5 T and $R = 4$) as shown in Section 4.3, the most important artifacts in our experiments appear as rings with sharp boundaries. In such circumstances, image-based edge-preserving priors may not appear well-adapted to regularize the reconstruction of the full FOV image. Therefore, in what follows, we make use of the Wavelet Transform (WT) [36] to improve artifact localization both in space and scales (or frequencies) and introduce some adequate priors in the wavelet space that promote sparsity of such a decomposition. We then derive efficient optimization algorithms that are able to cope with convex but non-differentiable criteria. As illustrated later, it will be shown that our penalization is better suited than its image-based TV counterpart.

3. Regularization in the WT domain

3.1. Motivation

As mentioned earlier, the presence of sharp artifacts makes the basic SENSE reconstruction inefficient under severe experimental conditions. An image representation where these artifacts can be easily detected and hence attenuated seems therefore necessary. In this respect, the WT has been recognized as a powerful tool that enables a good space and frequency localization of useful information [36]. In the literature, many wavelet decompositions and extensions have been reported offering different features in order to provide sparse image representations. For instance, decompositions onto orthonormal dyadic wavelet bases [37] including the Haar transform [38] as a special simple case or decompositions onto biorthogonal dyadic wavelets [39], M-band wavelet representations [40] and wavelet packet representations [41] have been extensively investigated in image denoising [42, 43, 44, 45, 46, 47] and deconvolution [46, 47, 48, 49]. In medical imaging, wavelet decompositions have also been widely used for image denoising [50, 26, 51], coil sensitivity map estimation and encoding schemes [52, 53, 54] in MRI, activation detection in fMRI [55, 56, 57, 58], tissue characterization in ultrasound imaging [59] and tomographic reconstruction [60]. An appealing property of the resulting decomposition is that the statistical distributions of the approximation and detail wavelet coefficients can be easily modelled in a realistic way. Hence, the Bayesian framework can be adopted to capture relevant information in the data through the likelihood definition, derive appropriate priors and select an efficient estimator to perform reconstruction in the wavelet domain.

3.2. Definitions and notations

In what follows, T stands for the WT operator and corresponds to a discrete decomposition onto a separable 2D M-band wavelet basis performed over $j_{\max}$ resolution levels. The full FOV image $\overline{\rho}$ of size $Y \times X$ can be seen as an element of the Euclidean space $\mathbb{C}^K$ with $K = Y \times X$ endowed with the standard inner product $\langle \cdot \mid \cdot \rangle$ and norm $\|\cdot\|$. As mentioned above, we are only interested in reconstructing one slice (2D image) for solving the 1D-SENSE problem. Hence, only 2D WT operators are investigated. In this context, the following notations are introduced.

Definition 3.1
Let $(e_k)_{1 \leq k \leq K}$ be the considered discrete wavelet basis of the space $\mathbb{C}^K$. The

wavelet decomposition operator T is defined as the linear operator:

$$\begin{aligned} T\colon \mathbb{C}^K &\to \mathbb{C}^K \\ \rho &\mapsto (\langle \rho \mid e_k \rangle)_{1\leq k\leq K}. \end{aligned} \tag{9}$$

The adjoint operator T^* serving for reconstruction purposes is then defined as the bijective linear operator:

$$\begin{aligned} T^*\colon \mathbb{C}^K &\to \mathbb{C}^K \\ (\zeta_k)_{1\leq k\leq K} &\mapsto \sum_{k=1}^{K} \zeta_k e_k. \end{aligned} \tag{10}$$

The resulting wavelet coefficient field of a target image function ρ is defined by $\zeta = \left(\boldsymbol{\zeta}_a, (\boldsymbol{\zeta}_{o,j})_{o\in\mathbb{O},1\leq j\leq j_{\max}}\right)$ where $\boldsymbol{\zeta}_a = (\zeta_{a,k})_{1\leq k\leq K_{j_{\max}}}$ and $\boldsymbol{\zeta}_{o,j} = (\zeta_{o,j,k})_{1\leq k\leq K_j}$, $K_j = KM^{-2j}$ being the number of wavelet coefficients in a given subband at resolution j (by assuming that Y and X are multiple of $M^{j_{\max}}$). The coefficients have been reindexed in such a way that $\zeta_{a,k}$ denotes an approximation coefficient at resolution level $j_{\max}$ and $\zeta_{o,j,k}$ denotes a detail coefficient at resolution level j and orientation $o \in \mathbb{O} = \{0,\ldots,M-1\}^2 \setminus \{(0,0)\}$. In the dyadic case ($M = 2$), there are three orientations corresponding to the horizontal, vertical or diagonal directions. Note that, when an orthonormal wavelet basis is considered, the adjoint operator T^* reduces to the inverse WT operator T^{-1} and the operator norm $\|T\|$ of T is equal to 1.

3.3. Wavelet-based regularized 1D-SENSE reconstruction

An estimate of the full FOV image $\overline{\rho}$ will be generated through the reconstruction wavelet operator T^*. Let $\overline{\zeta}$ be the unknown wavelet coefficients such that $\overline{\rho} = T^*\overline{\zeta}$. We aim at building an estimate $\widehat{\zeta}$ of the vector of coefficients $\overline{\zeta}$ from the observations $\mathbf{d} = (\boldsymbol{d}(\boldsymbol{r}))_{\mathbf{r}\in\{1,\ldots,Y/R\}\times\{1,\ldots,X\}}$.

To this end, we derive a Bayesian approach relying on suitable priors on the wavelet coefficients.

3.3.1. Likelihood

Given the observation model in Eq. (2) and the assumptions regarding the noise, the likelihood function factorizes over pixels lying in the $Y \times X$

FOV:

$$p(\mathbf{d} \mid T^*\zeta) = \prod_{\mathbf{r}\in\{1,\ldots,Y/R\}\times\{1,\ldots,X\}} p(\boldsymbol{d}(\mathbf{r}) \mid \boldsymbol{\rho}(\mathbf{r})) \propto \prod_{\mathbf{r}\in\{1,\ldots,Y/R\}\times\{1,\ldots,X\}} \exp\left(-\mathcal{J}_{\mathrm{WLS}}(\boldsymbol{\rho}(\mathbf{r}))\right)$$
$$\propto \exp\left(-\mathcal{J}_{\mathrm{L}}(T^*\zeta)\right) \tag{11}$$

with $\rho = T^*\zeta$ and $\mathcal{J}_{\mathrm{L}}$ is defined in Eq. (8).

3.3.2. Prior

Let f be the prior probability density function (pdf) of the image in the wavelet domain. By analyzing the correlation between the real and imaginary parts of the wavelet coefficients, very low correlations in pMRI images were found. Table 1 reports the values of the correlation coefficients computed over approximation and detail coefficients at different resolution levels using *Daubechies* wavelets of length 8. For the detail coefficients, the given values correspond to the average values over all orientations.

Table 1: Correlation coefficient between real and imaginary parts of approximation and detail coefficients over $j_{\max} = 3$ resolution levels. For the detail coefficients the reported values correspond to the average over all orientations (horizontal, vertical and diagonal).

		Approximation	Detail (Average)
Slice 5	$j = 1$	-0.004	-0.139
	$j = 2$	-0.031	-0.140
	$j = 3$	-0.026	-0.153
Slice 9	$j = 1$	-0.111	-0.077
	$j = 2$	-0.117	-0.032
	$j = 3$	-0.122	-0.031

These observations have motivated the choice of independent priors for the real and imaginary parts of the wavelet coefficients, the marginal distributions of which can be separately studied. For the sake of simplicity, we also assume that the real (resp. imaginary) parts of the wavelet coefficients are iid in each subband. Their statistical characteristics may however vary between two distinct subbands. Furthermore, by looking at the empirical distributions of the real and imaginary parts of the considered wavelet coefficients, we have noticed that their empirical histograms are well-fitted by a *Generalized Gauss-Laplace* (GGL) distribution. The histograms present

a single mode and their shape vary between the Gaussian and Laplacian densities. The corresponding pdf reads:

$$\forall \xi \in \mathbb{R}, \quad f(\xi; \alpha, \beta) = \sqrt{\frac{\beta}{2\pi}} \frac{e^{-(\alpha|\xi| + \frac{\beta}{2}\xi^2 + \frac{\alpha^2}{2\beta})}}{\operatorname{erfc}(\frac{\alpha}{\sqrt{2\beta}})}, \tag{12}$$

where $\alpha \in \mathbb{R}_+$ and β in $\mathbb{R}_+^*$ are hyper-parameters to be estimated. Fig. 2 illustrates the empirical histograms of real and imaginary parts of the horizontal detail subband at the first resolution level using the dyadic ($M = 2$) wavelet decomposition with *Daubechies* filters of length 8. This figure shows also that the adopted GGL distribution better fits the empirical histogram than a Generalized Gaussian (GG) pdf. This fact has been confirmed by applying a Kolmogorov-Smirnov goodness-of-fit test. At the coarsest reso-

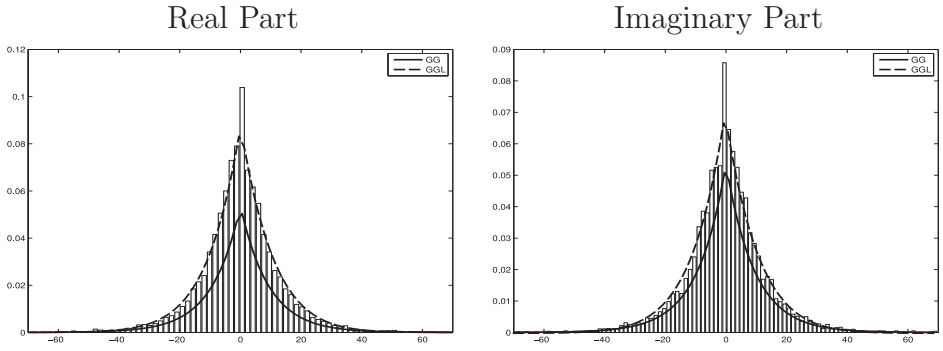

Figure 2: Example of normalized empirical histograms of wavelet coefficients and associated pdfs using GG (solid line) and GGL (dashed line) distributions, the hyperparameters being estimated using the ML estimator.

lution level $j_{\max}$, as often used in the wavelet literature, the distributions of both the real and imaginary parts of the approximation coefficients can be modelled by a Gaussian distribution since they belong to a low frequency subband.

3.3.3. Bayesian inference

Based on the prior and the likelihood given hereabove, the MAP estimator is computed by maximizing the full posterior distribution and finding

$$\widehat{\zeta}^{\mathrm{MAP}} = \arg\max_{\zeta \in \mathbb{C}^K} \big(\ln f(\zeta) + \ln p(\mathbf{d} \mid T^* \zeta) \big),$$

or equivalently by minimizing the following criterion:

$$\begin{aligned}\widehat{\zeta}^{\mathrm{MAP}} &= \underset{\zeta\in\mathbb{C}^K}{\arg\min}\, \mathcal{J}_{\mathrm{WT}}(\zeta) \\ &= \underset{\zeta\in\mathbb{C}^K}{\arg\min}\big(\mathcal{J}_L(T^*\zeta) + \mathcal{J}_P(\zeta)\big) \end{aligned} \tag{13}$$

$$\text{with} \quad \mathcal{J}_P(\zeta) = \sum_{k=1}^{K_{j\mathrm{max}}} \Phi_a(\zeta_{a,k}) + \sum_{o\in\mathbb{O}}\sum_{j=1}^{j_{\mathrm{max}}}\sum_{k=1}^{K_j} \Phi_{o,j}(\zeta_{o,j,k}) \tag{14}$$

and

$$\Phi_a(\zeta_{a,k}) = \frac{(\mathrm{Re}(\zeta_{a,k}) - \mu^{\mathrm{Re}})^2}{(\sqrt{2}\sigma^{\mathrm{Re}})^2} + \frac{(\mathrm{Im}(\zeta_{a,k}) - \mu^{\mathrm{Im}})^2}{(\sqrt{2}\sigma^{\mathrm{Im}})^2}, \tag{15}$$

$$\begin{aligned}\Phi_{o,j}(\zeta_{o,j,k}) = \alpha_{o,j}^{\mathrm{Re}}|\mathrm{Re}(\zeta_{o,j,k})| + \frac{\beta_{o,j}^{\mathrm{Re}}}{2}|\mathrm{Re}(\zeta_{o,j,k})|^2 + \alpha_{o,j}^{\mathrm{Im}}|\mathrm{Im}(\zeta_{o,j,k})| \\ + \frac{\beta_{o,j}^{\mathrm{Im}}}{2}|\mathrm{Im}(\zeta_{o,j,k})|^2. \end{aligned} \tag{16}$$

Hereabove, $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ (or $\cdot^{\mathrm{Re}}$ and $\cdot^{\mathrm{Im}}$) stand for the real and imaginary parts, respectively. The prior parameters $\boldsymbol{\alpha}_{o,j} = (\alpha_{o,j}^{\mathrm{Re}}, \alpha_{o,j}^{\mathrm{Im}}) \in (\mathbb{R}_+^*)^2$, $\boldsymbol{\beta}_{o,j} = (\beta_{o,j}^{\mathrm{Re}}, \beta_{o,j}^{\mathrm{Im}}) \in (\mathbb{R}_+^*)^2$, $\boldsymbol{\mu} = (\mu^{\mathrm{Re}}, \mu^{\mathrm{Im}}) \in \mathbb{R}^2$ and $\boldsymbol{\sigma} = (\sigma^{\mathrm{Re}}, \sigma^{\mathrm{Im}}) \in \mathbb{R}_+^2$ are unknown and need to be estimated. To this end, the Bayesian framework offers several concurrent strategies to compute hyper-parameter estimates: ML, marginal MAP or Posterior Mean (PM) estimators. The selected one usually depends on two ingredients: the full FOV image estimator and the algorithm employed for its computation. In the present case, the wavelet-based MAP estimator has been chosen as the solution to the 1D-SENSE regularized problem and its computation necessarily involves an optimization procedure. Thus, we propose to rely on the ML estimator for hyper-parameter estimation (see Subsection 3.4) prior to computing the full FOV image MAP estimator. Regarding the latter computation, since $\mathcal{J}_{\mathrm{WT}}$ is convex, its minimizers are global and hence global convergence can be guaranteed whatever the initialization. Nonetheless, the optimization cannot be performed by conventional descent algorithms like the pseudo-conjugate gradient method because $\mathcal{J}_P$ is not differentiable even if $\mathcal{J}_L$ is differentiable with a Lipschitz-continuous gradient. This difficulty is frequently encountered in inverse problems involving sparsity promoting priors [61, 62, 63, 64]. Therefore, we propose to apply a generalized form of the iterative optimization procedure developed in [46, 49], which is based on the Forward-Backward (FB) algorithm.

3.4. Maximum likelihood hyper-parameter estimation

The proposed reconstruction approach constitutes a useful alternative to quadratic regularization in SENSE imaging because it is able to provide an unsupervised setting to all hyper-parameters in the ML sense. Maximum-likelihood estimation for hyper-parameters is a very common procedure, which is currently used in various fields of signal and image processing when dealing with a small number of unknown but deterministic hyper-parameters $\boldsymbol{\Theta}$ (see for instance [65, 66]).

In wavelet-based pMRI image reconstruction, the ML estimator of hyper-parameters $\boldsymbol{\Theta} = \left(\boldsymbol{\mu}, \boldsymbol{\sigma}, \left(\boldsymbol{\alpha}_{o,j}, \boldsymbol{\beta}_{o,j}\right)_{o\in\mathbb{O}, 1\leq j\leq j_{\max}}\right)$ should be computed by maximizing the following *integrated* likelihood:

$$\widehat{\boldsymbol{\Theta}} = \arg\max_{\boldsymbol{\Theta}} p(\mathbf{d}; \boldsymbol{\Theta}) = \arg\max_{\boldsymbol{\Theta}} \int p(\mathbf{d} \,|\, T^{*}\zeta)\, f(\zeta; \boldsymbol{\Theta})\, d\zeta,$$

which requires to integrate out the sought image decomposition ζ and to iterate between image reconstruction and hyper-parameter estimation using the intensive EM algorithm [67]. In such circumstances, to alleviate the computational burden, akin to [68] we may proceed differently by assuming that a reference full FOV image $\overline{\rho}$ is available, and so is its wavelet decomposition $\overline{\zeta} = (T^{*})^{-1}\overline{\rho}$. In practice, our reference image $\overline{\rho}$ is obtained using 1D-SENSE reconstruction at the same R value.

The alternative ML estimation procedure consists of assuming that this reference image is a realization of the full prior distribution and thus in fitting $\boldsymbol{\Theta}$ directly on it. This procedure can be decomposed in two independent steps, the first one involving the setting of the Gaussian prior parameters $(\boldsymbol{\mu}, \boldsymbol{\sigma})$ attached to the approximation coefficients $\overline{\boldsymbol{\zeta}}_{a}$, and the second one being related to the estimation of the GGL prior parameters $(\boldsymbol{\alpha}_{o,j}, \boldsymbol{\beta}_{o,j})_{o\in\mathbb{O}, 1\leq j\leq j_{\max}}$ from the corresponding detail coefficients $(\overline{\boldsymbol{\zeta}}_{o,j})_{o\in\mathbb{O}, 1\leq j\leq j_{\max}}$. On the one hand, ML estimators $(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\sigma}})$ are explicitely given by the empirical mean and standard deviation:

$$\widehat{\mu}^{\mathrm{Re}} = \frac{1}{K_{j_{\max}}} \sum_{k=1}^{K_{j_{\max}}} \mathrm{Re}(\overline{\zeta}_{a,k}),\ \widehat{\sigma}^{\mathrm{Re}} = \sqrt{\frac{1}{K_{j_{\max}}} \sum_{k=1}^{K_{j_{\max}}} \left(\mathrm{Re}(\overline{\zeta}_{a,k}) - \widehat{\mu}^{\mathrm{Re}}\right)^{2}}$$

$$\widehat{\mu}^{\mathrm{Im}} = \frac{1}{K_{j_{\max}}} \sum_{k=1}^{K_{j_{\max}}} \mathrm{Im}(\overline{\zeta}_{a,k}) \text{ and } \widehat{\sigma}^{\mathrm{Im}} = \sqrt{\frac{1}{K_{j_{\max}}} \sum_{k=1}^{K_{j_{\max}}} \left(\mathrm{Re}(\overline{\zeta}_{a,k}) - \widehat{\mu}^{\mathrm{Im}}\right)^{2}}.$$

For each resolution level j and orientation o, $\widehat{\alpha}_{o,j}^{\mathrm{Re}}$ and $\widehat{\beta}_{o,j}^{\mathrm{Re}}$ are estimated from $\overline{\boldsymbol{\zeta}}_{o,j}$ as follows (we proceed similarly to estimate $\widehat{\alpha}_{o,j}^{\mathrm{Im}}$ and $\widehat{\beta}_{o,j}^{\mathrm{Im}}$ by replacing $\mathrm{Re}(\cdot)$ by $\mathrm{Im}(\cdot)$):

$$
\begin{aligned}
(\widehat{\alpha}_{o,j}^{\mathrm{Re}}, \widehat{\beta}_{o,j}^{\mathrm{Re}}) &= \underset{(\alpha,\beta)\in\mathbb{R}_+\times\mathbb{R}_+^*}{\arg\max}\; f(\mathrm{Re}(\overline{\boldsymbol{\zeta}}_{o,j}); \alpha, \beta) \\
&= \underset{(\alpha,\beta)\in\mathbb{R}_+\times\mathbb{R}_+^*}{\arg\max} \sum_{k=1}^{K_j} \log f(\mathrm{Re}(\overline{\zeta}_{o,j,k}); \alpha, \beta) \\
&= \underset{(\alpha,\beta)\in\mathbb{R}_+\times\mathbb{R}_+^*}{\arg\max} \Big\{ \alpha \sum_{k=1}^{K_j} |\mathrm{Re}(\overline{\zeta}_{o,j,k})| + \frac{\beta}{2} \sum_{k=1}^{K_j} \mathrm{Re}(\overline{\zeta}_{o,j,k})^2 + \frac{K_j \alpha^2}{2\beta} \\
&\quad - \frac{K_j}{2} \log(\frac{\beta}{2\pi}) + K_j \log\big(\mathrm{erfc}(\frac{\alpha}{\sqrt{2\beta}})\big) \Big\} \qquad (17)
\end{aligned}
$$

This two-dimensional minimization problem does not admit a closed form solution. Hence, we compute the ML parameters using the zero-order Powell optimization method [69]. Alternative solutions based on Monte Carlo methods [70, 71] or the Stein principle [72] can also be thought of, at the expense of the computational burden.

3.5. Forward-Backward optimization algorithm

The minimization of $\mathcal{J}_{\mathrm{WT}}$ is performed by resorting to the concept of proximity operators [73], which was found to be fruitful in a number of recent works in convex optimization [49, 74, 75] and an efficient alternative to the graduated non-differentiability algorithm [76, 77]. In what follows, we give a brief overview of this key tool for solving our optimization problem.

Definition 3.2 [73]

Let $\Gamma_0(\chi)$ be the class of lower semicontinuous convex functions from a separable real Hilbert space χ to $]-\infty,+\infty]$ and let $\varphi \in \Gamma_0(\chi)$. For every $\mathsf{x} \in \chi$, the function $\varphi + \|\cdot - \mathsf{x}\|^2/2$ achieves its infimum at a unique point denoted by $\mathrm{prox}_\varphi \mathsf{x}$. The operator $\mathrm{prox}_\varphi : \chi \to \chi$ is the proximity operator of φ.

Here is an example of a proximity operators which will be used in our approach.

Example 3.1
Consider the following function:

$$\begin{aligned} \varphi\colon \mathbb{R} &\to \mathbb{R} \\ \xi &\mapsto \alpha|\xi-\mu| + \frac{\beta}{2}(\xi-\mu)^2 \end{aligned} \tag{18}$$

with $\alpha \in \mathbb{R}_+$, $\beta \in \mathbb{R}_+^*$ and $\mu \in \mathbb{R}$. The associated proximity operator is given by

$$\forall \xi \in \mathbb{R}, \qquad \mathrm{prox}_{\varphi}\xi = \frac{\mathrm{sign}(\xi-\mu)}{\beta+1}\max\{|\xi-\mu|-\alpha, 0\} + \mu \tag{19}$$

where the sign function is defined as follows:

$$\forall \xi \in \mathbb{R}, \qquad \mathrm{sign}(\xi) = \begin{cases} +1 & \text{if } \xi \geq 0 \\ -1 & \text{otherwise.} \end{cases}$$

In this work, as the observed data are complex-valued, we generalize the definition of proximity operators to a class of convex functions defined for complex-valued variables. For the function

$$\begin{aligned} \Phi\colon \mathbb{C}^K &\to]-\infty, +\infty] \\ x &\mapsto \phi^{\mathrm{Re}}(\mathrm{Re}(x)) + \phi^{\mathrm{Im}}(\mathrm{Im}(x)), \end{aligned} \tag{20}$$

where ϕ^{Re} and ϕ^{Im} are functions in $\Gamma_0(\mathbb{R}^K)$ and $\mathrm{Re}(x)$ (resp. $\mathrm{Im}(x)$) is the vector of the real parts (resp. imaginary parts) of the components of $x \in \mathbb{C}^K$, the proximity operator is defined as

$$\begin{aligned} \mathrm{prox}_{\Phi}\colon \mathbb{C}^K &\to \mathbb{C}^K \\ x &\mapsto \mathrm{prox}_{\phi^{\mathrm{Re}}}(\mathrm{Re}(x)) + \imath\mathrm{prox}_{\phi^{\mathrm{Im}}}(\mathrm{Im}(x)). \end{aligned} \tag{21}$$

An example of proximity operator for a function of a complex-valued variable is given below.

Example 3.2
Consider the following function:

$$\begin{aligned} \Phi \;:\; \mathbb{C} &\to \mathbb{R} \\ \xi &\mapsto \alpha^{\mathrm{Re}}|\mathrm{Re}(\xi-\mu)| + \frac{\beta^{\mathrm{Re}}}{2}\big(\mathrm{Re}(\xi-\mu)\big)^2 \\ &\quad + \alpha^{\mathrm{Im}}|\mathrm{Im}(\xi-\mu)| + \frac{\beta^{\mathrm{Im}}}{2}\big(\mathrm{Im}(\xi-\mu)\big)^2 \end{aligned} \tag{22}$$

with $(\alpha^{\mathrm{Re}}, \alpha^{\mathrm{Im}}) \in (\mathbb{R}_+)^2$, $(\beta^{\mathrm{Re}}, \beta^{\mathrm{Im}}) \in (\mathbb{R}_+^*)^2$ and $\mu \in \mathbb{C}$. The associated proximity operator is:

$$\begin{aligned}\mathrm{prox}_{\Phi}\xi = \frac{\mathrm{sign}(\mathrm{Re}(\xi-\mu))}{\beta^{\mathrm{Re}}+1}\max\{|\mathrm{Re}(\xi-\mu)|-\alpha^{\mathrm{Re}},0\} \\ + \imath\frac{\mathrm{sign}(\mathrm{Im}(\xi-\mu))}{\beta^{\mathrm{Im}}+1}\max\{|\mathrm{Im}(\xi-\mu)|-\alpha^{\mathrm{Im}},0\}+\mu. \quad (23)\end{aligned}$$

Based on these definitions, and by extending the algorithm in [49] to the complex case, a minimizer of $\mathcal{J}_{\mathrm{WT}}$ can then be iteratively computed according to Algorithm 1. Note that in this algorithm, the expressions of $\mathrm{prox}_{\gamma_n\Phi_a}$ and $\mathrm{prox}_{\gamma_n\Phi_{o,j}}$ at each iteration n are provided by Example 3.2. It can also be noticed that λ_n and γ_n correspond to relaxation and step-size parameters, respectively.

Again, we iterate Algorithm 1 also called the L_{12}UWR-SENSE reconstruction method over all slices to perform 3D volume reconstruction as required for anatomical data. Regarding functional data, we also iterate over volumes separately to get the reconstructed series of EPI full FOV volumes. Interestingly, the description of Algorithm 1 allows us to realize that the computation of the solution can be parallelized over resolution levels since the proximity operators can be computed independently.

3.6. Convergence of Algorithm 1

For every $\mathbf{r} \in \{1,\dots,Y/R\} \times \{1,\dots,X\}$, let $\theta_{\mathbf{r}} \geq 0$ be the maximum eigenvalue of the Hermitian positive semi-definite matrix $\boldsymbol{S}^{\mathsf{H}}(\mathbf{r})\boldsymbol{\Psi}^{-1}\boldsymbol{S}(\mathbf{r})$ and let $\theta = \max_{\mathbf{r}\in\{1,\dots,Y/R\}\times\{1,\dots,X\}}\theta_{\mathbf{r}} > 0$. To guarantee the convergence of Algorithm 1, the step-size and relaxation parameters have to meet the following conditions:

Assumption 3.3

(i) $\inf_{n>0}\gamma_n > 0$ *and* $\sup_{n>0}\gamma_n < \frac{1}{\theta\|T\|^2}$,

(ii) $\inf_{n>0}\lambda_n > 0$ *and* $\sup_{n>0}\lambda_n \leq 1$.

More precisely, the following result can be shown:

Proposition 3.4 *Under Assumption 3.3, the sequence $(\zeta^{(n)})_{n>0}$ generated when iterating Steps 3 to 9 of Algorithm 1 converges linearly to the unique minimizer $\widehat{\zeta}$ of $\mathcal{J}_{\mathrm{WT}}$.*

Algorithm 1 L_{12}**UWR-SENSE**: 2D-slice wavelet-based regularized reconstruction

Let $(\gamma_n)_{n>0}$ and $(\lambda_n)_{n>0}$ be sequences of positive reals.

1: Initialize $\zeta^{(1)}$. Set $n = 1$, $\varepsilon \geq 0$ and $\mathcal{J}^{(0)} = 0$.
2: **repeat**
3: Reconstruct the image by setting $\rho^{(n)} = T^*\zeta^{(n)}$.
4: Compute the image $u^{(n)}$ such that:
$\forall \mathbf{r} \in \{1, \ldots, Y/R\} \times \{1, \ldots, X\}$,
$\boldsymbol{u}^{(n)}(\mathbf{r}) = 2\boldsymbol{S}^{\mathsf{H}}(\mathbf{r})\boldsymbol{\Psi}^{-1}\left(\boldsymbol{S}(\mathbf{r})\boldsymbol{\rho}^{(n)}(\mathbf{r}) - \boldsymbol{d}(\mathbf{r})\right)$,
where the vector $\boldsymbol{u}^{(n)}(\mathbf{r})$ is defined from $u^{(n)}$ in the same way as $\overline{\boldsymbol{\rho}}(\mathbf{r})$ is defined from $\overline{\rho}$ (see Eq. (3)).
5: Determine the wavelet coefficients $\upsilon^{(n)} = Tu^{(n)}$ of $u^{(n)}$.
6: Update the approximation coefficients of the reconstructed image $\rho^{(n+1)}$:
$\forall k \in \{1, \cdots, K_{j_{\max}}\}$, $\zeta_{a,k}^{(n+1)} = \zeta_{a,k}^{(n)} + \lambda_n\Big(\mathrm{prox}_{\gamma_n\Phi_a}(\zeta_{a,k}^{(n)} - \gamma_n\upsilon_{a,k}^{(n)}) - \zeta_{a,k}^{(n)}\Big)$.
7: Update the detail coefficients of the reconstructed image $\rho^{(n+1)}$:

$$\forall o \in \mathbb{O}, \forall j \in \{1, \ldots, j_{\max}\}, \forall k \in \{1, \cdots, K_j\},$$
$$\zeta_{o,j,k}^{(n+1)} = \zeta_{o,j,k}^{(n)} + \lambda_n\Big(\mathrm{prox}_{\gamma_n\Phi_{o,j}}(\zeta_{o,j,k}^{(n)} - \gamma_n\upsilon_{o,j,k}^{(n)}) - \zeta_{o,j,k}^{(n)}\Big).$$

8: Compute $\mathcal{J}^{(n)} = \mathcal{J}_{\mathrm{WT}}(\zeta^{(n)})$.
9: $n \leftarrow n + 1$
10: **until** $|\mathcal{J}^{(n-1)} - \mathcal{J}^{(n-2)}| \leq \varepsilon\mathcal{J}^{(n-1)}$
11: **return** $\rho^{(n)} = T^*\zeta^{(n)}$

Proof: See Appendix A.

The results we obtained using Algorithm 1 are discussed in Section 4. Since only artifacts of moderate size have been removed using Algorithm 1, we subsequently present an extension which accounts for additional constraints leading to a better reconstruction quality and cancellation of more severe artifacts.

3.7. Constrained wavelet-based regularization

We propose to extend our approach by incorporating an additional constraint in the method described hereabove in order to better regularize artifact regions. The resulting algorithm is called L_{12}CWR-SENSE method hereafter.

3.7.1. New optimality criterion

Here, we set local lower and upper bounds on the image intensity values in artifact areas, regardless of their shape and/or location. These bounds define the nonempty closed convex set:

$$C = \{\rho \in \mathbb{C}^K \mid \forall \boldsymbol{r} \in \{1, \ldots, Y/R\} \times \{1, \ldots, X\}, \rho(\boldsymbol{r}) \in C_{\boldsymbol{r}}\} \tag{24}$$

where the constraint introduced on the range values at position $\boldsymbol{r} \in \{1, \ldots, Y/R\} \times \{1, \ldots, X\}$ is modelled by:

$$C_{\boldsymbol{r}} = \{\xi \in \mathbb{C} \mid \mathrm{Re}(\xi) \in \mathbb{I}_{\boldsymbol{r}}^{\mathrm{Re}}, \mathrm{Im}(\xi) \in \mathbb{I}_{\boldsymbol{r}}^{\mathrm{Im}}\}, \tag{25}$$

with $\mathbb{I}_{\boldsymbol{r}}^{\mathrm{Re}} = [\mathrm{I}_{\mathrm{min},\boldsymbol{r}}^{\mathrm{Re}}, \mathrm{I}_{\mathrm{max},\boldsymbol{r}}^{\mathrm{Re}}]$ and $\mathbb{I}_{\boldsymbol{r}}^{\mathrm{Im}} = [\mathrm{I}_{\mathrm{min},\boldsymbol{r}}^{\mathrm{Im}}, \mathrm{I}_{\mathrm{max},\boldsymbol{r}}^{\mathrm{Im}}]$.

When taking into account the additional constraints defined in Eq. (25), the optimized criterion in Eq. (13) becomes:

$$\mathcal{J}_{\mathrm{CWT}}(\zeta) = \mathcal{J}_{\mathrm{WT}}(\zeta) + i_{C^*}(\zeta), \tag{26}$$

where

$$C^* = \{\zeta \in \mathbb{C}^K \mid T^*\zeta \in C\}$$

and i_{C^*} is the indicator function of the closed convex set C^* defined by:

$$\forall \zeta \in \mathbb{C}^K, \qquad i_{C^*}(\zeta) = \begin{cases} 0 & \text{if } \zeta \in C^* \\ +\infty & \text{otherwise.} \end{cases}$$

Hence, the constrained MAP estimator satisfies:

$$\widehat{\zeta}^{\mathrm{MAP}} = \arg\min_{\zeta \in C^*} \mathcal{J}_{\mathrm{WT}}(\zeta) = \arg\min_{\zeta \in \mathbb{C}^K} \mathcal{J}_{\mathrm{CWT}}(\zeta) \tag{27}$$

An open question is how to choose the reference image for deriving the convex set C^*. For simplicity, the pixelwise constraints have been computed on the reconstructed SENSE image as illustrated in Section 4. This computation is straightforward to perform and does not require any extra parameter estimation.

3.7.2. Computation of the constrained MAP estimator

Theoretically, to solve the minimization problem in Eq. (27), the FB iteration (see Eq. (40) in Appendix A) has to be modified into:

$$\zeta^{(n+1)} = \zeta^{(n)} + \lambda_n \Big(\mathrm{prox}_{\gamma_n \mathcal{J}_P + i_{C^*}} \big(\zeta^{(n)} - \gamma_n \nabla \mathcal{J}_L(\zeta^{(n)}) \big) - \zeta^{(n)} \Big). \tag{28}$$

The main difficulty here lies in the fact that the proximity operator of $\gamma_n \mathcal{J}_P + i_{C^*}$ does not admit a closed form. However, from its definition, we get:

$$\forall \zeta \in \mathbb{C}^K, \qquad \mathrm{prox}_{\gamma_n \mathcal{J}_P + i_{C^*}}(\zeta) = \arg\min_{\zeta' \in \mathbb{C}^K} \gamma_n \mathcal{J}_P(\zeta') + \mathcal{J'}_{\zeta}(\zeta') \tag{29}$$

where

$$\mathcal{J'}_{\zeta}(\cdot) = \frac{1}{2} \| \cdot - \zeta \|^2 + i_{C^*}(\cdot). \tag{30}$$

Although $\mathrm{prox}_{\gamma_n \mathcal{J}_P + i_{C^*}}$ does not take a simple expression, the proximity operator of $\gamma_n \mathcal{J}_P$ is given by Eq. (39) (see Appendix A) and the proximity operator of $\mathcal{J'}_{\zeta}$ is easily determined. Indeed, it is quite straightforward to show that

$$\forall \zeta' \in \mathbb{C}^K, \qquad \mathrm{prox}_{\mathcal{J'}_{\zeta}}(\zeta') = P_{C^*} \left(\frac{\zeta' + \zeta}{2} \right) \tag{31}$$

where P_{C^*} is the projection onto the convex set C^*. In turn, provided that the considered wavelet basis is orthonormal, the projection onto C^* of $\zeta' \in \mathbb{C}^K$ is obtained by performing the wavelet decomposition of the projection of $\rho' = T^* \zeta'$ onto C. The latter projection reads:

$$P_C(\rho') = \big(P_{C_{\boldsymbol{r}}}(\rho'(\boldsymbol{r})) \big)_{\boldsymbol{r} \in \{1,\ldots,Y/R\} \times \{1,\ldots,X\}} \tag{32}$$

where, for every $\boldsymbol{r} \in \{1, \ldots, Y/R\} \times \{1, \ldots, X\}$,

$$\forall \xi \in \mathbb{C}, \qquad \mathrm{Re}\big(P_{C_{\boldsymbol{r}}}(\xi)\big) = \begin{cases} \mathrm{I}^{\mathrm{Re}}_{\min,\boldsymbol{r}} & \text{if } \mathrm{Re}(\xi) < \mathrm{I}^{\mathrm{Re}}_{\min,\boldsymbol{r}} \\ \mathrm{I}^{\mathrm{Re}}_{\max,\boldsymbol{r}} & \text{if } \mathrm{Re}(\xi) > \mathrm{I}^{\mathrm{Re}}_{\max,\boldsymbol{r}} \\ \xi & \text{otherwise,} \end{cases} \tag{33}$$

a similar expression being used to calculate $\mathrm{Im}\big(P_{C_{\boldsymbol{r}}}(\xi)\big)$.

Knowing $\mathrm{prox}_{\gamma_n \mathcal{J}_P}$ and $\mathrm{prox}_{\mathcal{J}'_\zeta}$, $\mathrm{prox}_{\gamma_n \mathcal{J}_P + i_{C^*}\zeta}$ can be iteratively computed by solving the optimization problem in Eq. (29) using the Douglas-Rachford algorithm [78, 79]. More precisely, we apply the following proposition:

Proposition 3.5
Set $\eta^{(0)} \in \mathbb{C}^K$ and construct for all $m \in \mathbb{N}$:

$$\begin{cases} \eta^{(m+\frac{1}{2})} = \mathrm{prox}_{\mathcal{J}'_\zeta} \eta^{(m)} \\ \eta^{(m+1)} = \eta^{(m)} + \tau\big(\mathrm{prox}_{\gamma_n \mathcal{J}_P}(2\eta^{(m+\frac{1}{2})} - \eta^{(m)}) - \eta^{(m+\frac{1}{2})}\big), \end{cases} \tag{34}$$

where $\tau \in]0, 2[$. Then, $(\eta^{(m+\frac{1}{2})})_{m \in \mathbb{N}}$ converges to $\mathrm{prox}_{\gamma_n \mathcal{J}_P + i_{C^*}\zeta}$.

Inserting this extra iterative step in the FB algorithm and using the expressions of $\mathrm{prox}_{\gamma_n \mathcal{J}_P}$ and $\mathrm{prox}_{\mathcal{J}'_\zeta}$ in Eqs. (31)-(33) lead to Algorithm 2 hereafter called the L_{12}CWR-SENSE method. At iteration n, M_n is the number of times the Douglas-Rachford step is run. According to the convergence analysis conducted in [78, Prop. 4.2], if M_n is chosen large enough and Assumption 3.3 holds, iterating Steps 3 to 16 of the L_{12}CWR-SENSE method guarantees the convergence to the unique solution of $\mathcal{J}_{\mathrm{CWT}}$.

Note however that [78, Prop. 4.2] does not provide a practical guideline for chosing M_n. The practical rule we chose is explained in Section 4.1. The improvements resulting from the use of the L_{12}CWR-SENSE algorithm are illustrated in the next section, of course at the expense of the computation time.

4. Experimental results

Experiments have been conducted on real data sets comprising $256 \times 256 \times 14$ Gradient-Echo (GE) anatomical and $64 \times 64 \times 30$ fMRI GE-EPI images with respectively $0.93 \times 0.93 \times 8$ (mm^3) and $3.75 \times 3.75 \times 3$ (mm^3)

Algorithm 2 L_{12}**CWR-SENSE**: Constrained 2D-slice wavelet-based regularized reconstruction.

Let $(\gamma_n)_{n>0}$ and $(\lambda_n)_{n>0}$ be sequences of positive reals, let $(M_n)_{n>0}$ be a sequence of positive integers and set $\tau \in]0, 2]$.

1: Initialize $\zeta^{(1)}$. Set $n = 1$, $\varepsilon \geq 0$ and $\mathcal{J}^{(0)} = 0$.
2: **repeat**
3: Same as for Algorithm 1
4: Same as for Algorithm 1
5: Same as for Algorithm 1
6: Initialize the Douglas-Rachford algorithm by setting $\eta^{(n,0)} = \zeta^{(n)} - \gamma_n \upsilon^{(n)}$.
7: Douglas-Rachford iterations:

- **for** $m = 0$ to $M_n - 1$ **do**

9: Compute $\eta^{(n,m+\frac{1}{2})} = P_{C^*}\Big(\dfrac{\eta^{(n,m)} + \zeta^{(n)}}{2}\Big)$;
10: Update the approximation components of $\eta^{(n,m)}$:
$\forall k \in \{1, \cdots, K_{j_{\max}}\}$,
$\eta_{a,k}^{(n,m+1)} = \eta_{a,k}^{(n,m)} + \tau\big(\mathrm{prox}_{\gamma_n \Phi_a}(2\eta_{a,k}^{(n,m+\frac{1}{2})} - \eta_{a,k}^{(n,m)}) - \eta_{a,k}^{(n,m+\frac{1}{2})}\big)$,
11: Update the detail components of $\eta^{(n,m)}$:
$\forall o \in \mathbb{O}, \forall j \in \{1, \ldots, j_{\max}\}, \forall k \in \{1, \cdots, K_j\}$,
$\eta_{o,j,k}^{(n,m+1)} = \eta_{o,j,k}^{(n,m)} + \tau\big(\mathrm{prox}_{\gamma_n \Phi_{o,j}}(2\eta_{o,j,k}^{(n,m+\frac{1}{2})} - \eta_{o,j,k}^{(n,m)}) - \eta_{o,j,k}^{(n,m+\frac{1}{2})}\big)$;
12: If $\eta^{(n,m+1)} = \eta^{(n,m)}$, goto 14.

- **end for**

14: Update the wavelet coefficients of the reconstructed image: $\zeta^{(n+1)} = \zeta^{(n)} + \lambda_n(\eta^{(n,m+\frac{1}{2})} - \zeta^{(n)})$.
15: Compute $\mathcal{J}^{(n)} = \mathcal{J}_{\mathrm{CWT}}(\zeta^{(n)})$.
16: $n \leftarrow n + 1$.
17: **until** $|\mathcal{J}^{(n-1)} - \mathcal{J}^{(n-2)}| \leq \varepsilon \mathcal{J}^{(n-1)}$
18: **return** $\rho^{(n)} = T^* \zeta^{(n)}$

spatial resolution. GE anatomical images were acquired with TE/TR = 10/500 (ms) and BW=31.25 (kHz), while $N = 30$ EPI images were recorded

using TE/TR = 30/2400 (ms) and BW=125 (kHz) during a resting state session (subject lying on the scanner's bed with his eyes closed). Note also that these images have been acquired using acceleration factors $R = 2$ and $R = 4$ on a Signa 1.5 Tesla GE Healthcare scanner with an eight-channel head coil. Interestingly, the scanning time of anatomical data lasted 5 mn in non-parallel imaging, while acquisition duration was decreased to 3mn10s and 2mn20s in parallel imaging with $R = 2$ and $R = 4$, respectively.

For validation purpose, we also acquired a full FOV GE anatomical image ρ_{True} using a non accelerated acquisition ($R = 1$). This provides us with the ground truth for assessing the reconstruction accuracy of the different algorithms.

4.1. Results on anatomical data

A comparison between the basic-SENSE reconstruction and the L_2R-SENSE method is shown in Fig. 3. Note that in Tikhonov regularization, the reference image ρ_{r} was chosen as a mean image based on the basic-SENSE reconstruction, which contains the mean value of the signal of interest within the brain mask. The regularization parameter κ was manually fixed and different settings were tested in order to retain the best choice in terms of SNR since a reference image acquired in non-parallel imaging was available in this experiment. As expected, the aliasing artifacts in the basic-SENSE reconstructed image are smoothed by the L_2R-SENSE approach.

For the WT regularization, dyadic ($M = 2$) *Symmlet* orthonormal wavelet bases [80] associated with filters of length 8 have been used over $j_{\max} = 3$ resolution levels. Regarding the wavelet coefficients, the priors described in Subsection 3.3.3 have been employed. As outlined in Section 3.4, the related hyper-parameters $\boldsymbol{\Theta}$ have been estimated in the ML sense based on the reference image ρ_{True}.[5]

Full FOV image reconstruction was then performed using our L_{12}UWR-SENSE method. For the sake of simplicity, constant values of relaxation and step-size parameters (λ_n and γ_n, respectively) have been adopted along the algorithm iterations: first, we experimentally observed that $\lambda_n \equiv 1$ is the optimal value of the relaxation parameter in terms of convergence rate (see

[5] A couple of hyper-parameters is fitted for real/imaginary parts in each subband, i.e.for each approximation/detail coefficient sequence at each resolution level and orientation.

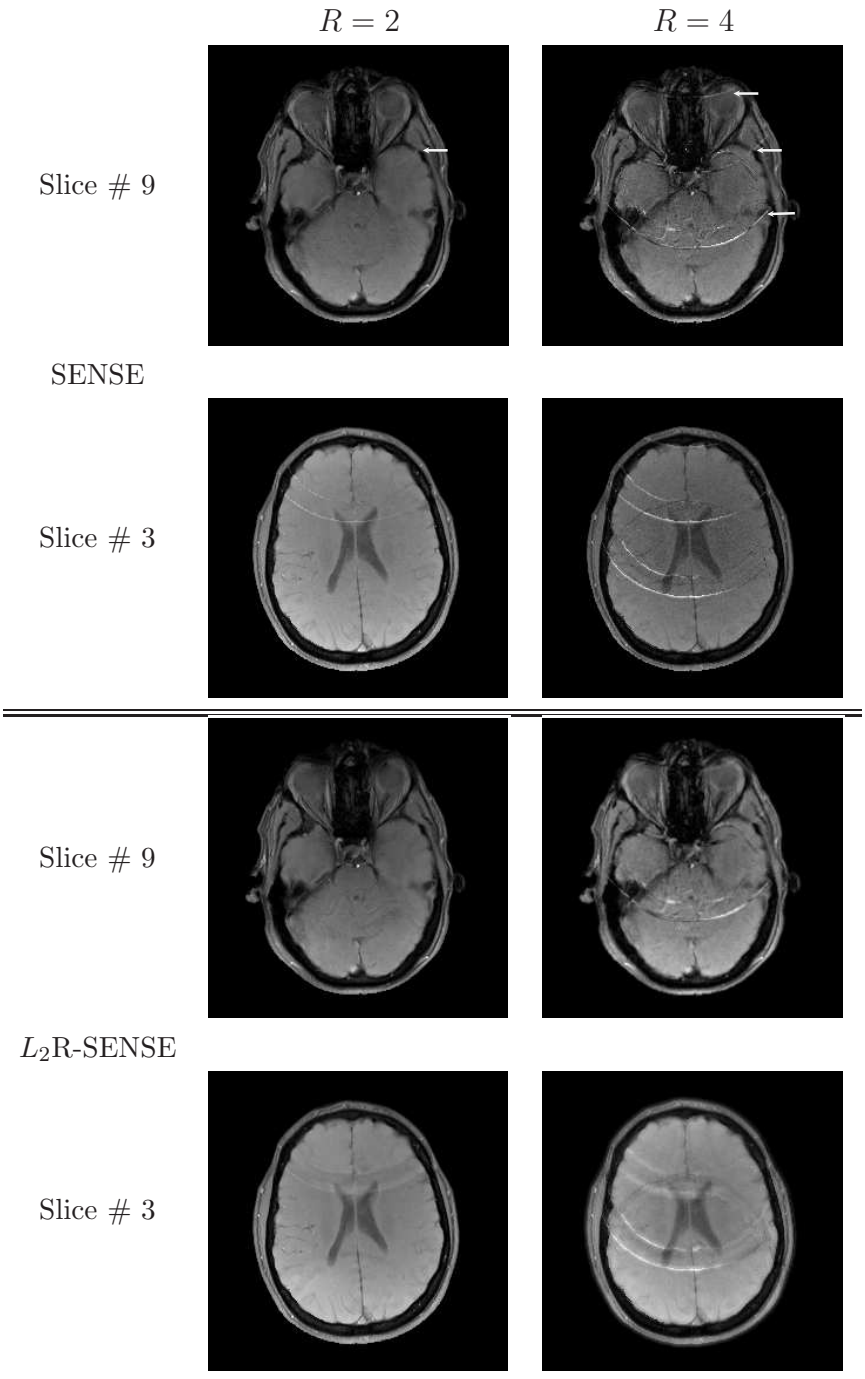

Figure 3: Two reconstructed slices using SENSE and Tikhonov regularization for $R = 2$ and $R = 4$. Artifacts of SENSE reconstruction are pointed out using white arrows on top panels.

Fig. 4). Second, a step-size parameter γ_n closed to the allowed maximum value in Assumption 3.3 provides the fastest convergence rate. After computing the constant θ related to the considered sensivity map, γ_n was thus chosen equal to $0.99/\theta = 12.83$.

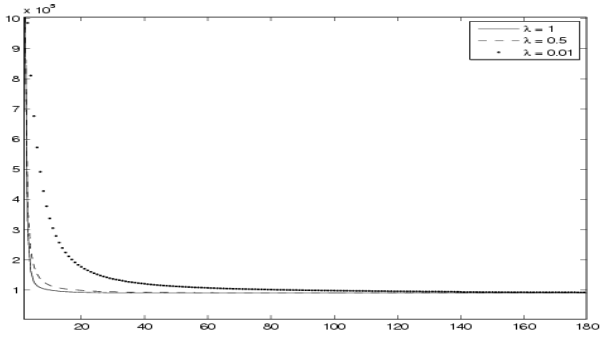

Figure 4: Convergence speed of the optimization algorithm w.r.t. the choice of the relaxation parameter λ for $j_{\max} = 3$ (criterion value w.r.t. iteration number).

The algorithm was stopped when $\mathcal{J}_{\mathrm{WT}}$ no longer significantly varies, by choosing $\varepsilon = 10^{-4}$ in Algorithm 1. For different values of λ, Fig. 4 illustrates the evolution of the optimized criterion $\mathcal{J}_{\mathrm{WT}}$ w.r.t. the iteration number for a 2D-slice reconstruction. In Fig 4, it is clear that after about 20 iterations the minimizer $\widehat{\zeta}^{\mathrm{MAP}}$ is reached. In terms of computation time, the approach was implemented using the C language and took 6 seconds on such data when running on an Intel Core 2 (2.26 GHz) architecture. Note that accelerated algorithms such as TWIST or FISTA have been recently proposed in the literature [81, 82] for minimizing a similar optimality criterion. In Fig. 5, we compared the different algorithms and their convergence speed on the pMRI application. With respect to the number of iterations required to achieve convergence, we observed no improvement in terms of convergence speed using FISTA instead of the proposed FB algorithm. Also, our FB implementation as well as FISTA perform faster than TWIST.

Our L_{12}CWR-SENSE constrained algorithm presented in Section 3.7 was applied to the anatomical data with the *Symmlet 8* wavelet basis. The parameters λ_n and γ_n have been set to the same values as for the L_{12}UWR-SENSE, while we chose $\tau = 1.99$ for the underlying Douglas-Rachford iterations as it was practically observed that this value gives the best convergence rate. In

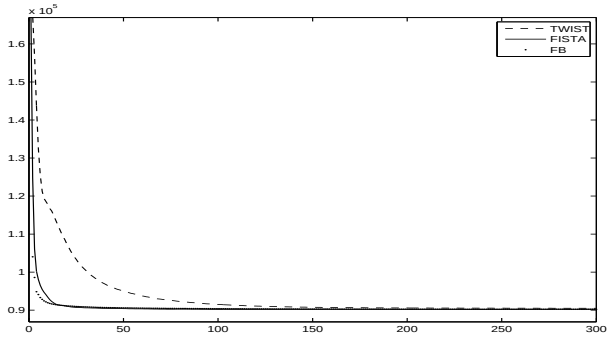

Figure 5: Convergence speed comparison for the TWIST, FISTA and FB algorithms criterion value w.r.t. iteration number.

practice, the value of M_n was defined as the minimal integer value such that

$$|\frac{\eta^{(n,M_n-1)} - \eta^{(n,M_n-2)}}{\eta^{(n,M_n-2)}}| < 10^{-4},$$

which results in about 4 iterations of the Douglas-Rachford algorithm. A morphological gradient [83] was used to detect artifact regions on which we apply the additional convex constraints. The upper and lowed bounds that define the convex sets $C_{\boldsymbol{r}}$ in Eq. (25) were computed using a morphological opening and closing operations applied to the basic-SENSE reconstruction in order to discard very low and high intensities. The L_{12}CWR-SENSE reconstruction steps are illustrated in Fig. 6. Clearly, the accuracy of the

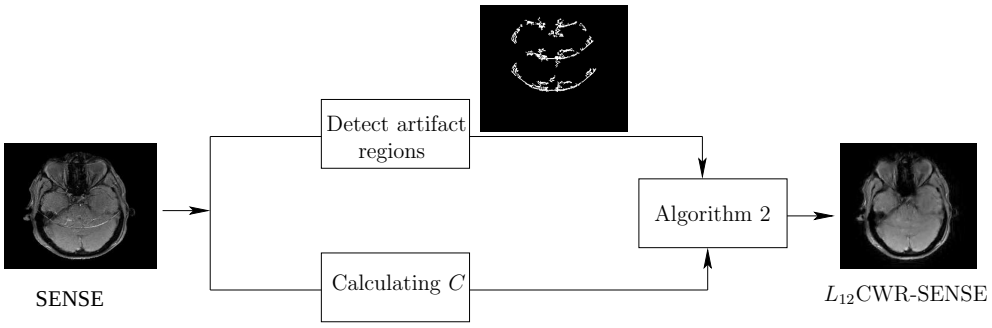

Figure 6: L_{12}CWR-SENSE reconstruction steps.

artifact region detection (see the mask image on the top of Fig. 6) impacts the performance of our L_{12}CWR-SENSE method since it defines the areas

on which the convex constraints are applied as lower and upper thresholds on the image intensities.

Fig. 7 shows reconstructed full FOV anatomical images using the proposed approaches (L_{12}UWR-SENSE and L_{12}CWR-SENSE) with $R = 2$ and $R = 4$. The smoothing effects observed in Fig. 3 with Tikhonov regularization do no longer exist in the WT regularized images in Fig. 7, where a quite accurate reconstruction is performed within the brain mask. Usual inspection of the L_{12}CWR-SENSE reconstruction results shows that most of the surviving artifacts in Fig. 3 have now been removed. For $R = 4$, the remaining artifacts are visible either inside or outside the artifact region. In the former case, these artifacts have not been completely removed because the fixed upper/lower intensity bounds are rather adapted to the rest of the artifact region, which is spatially more extended. In the latter case, the artifacts are simply due to some false negatives in the artifact region detection. This emphasizes the importance of the artifact detection step. For illustration purpose, Fig. 8 (left) shows the difference image between the reference and the reconstructed one using the L_{12}CWR-SENSE algorithm for Slice # 3 and $R = 4$, while Fig. 8 (right) shows the corresponding detected actifact region. Clearly speaking, adaptively fixing the upper/lower intensity bounds within the artifact region would lead to more accurate reconstruction. Moreover, false negatives in the artifact detection can also limit the L_{12}CWR-SENSE algorithm performance. On the other hand, false positives may lead to oversmoothed parts in the reconstructed images. In this context, a suitable trade-off has to be made during the artifact detection step in order to avoid very high false positive or negative rates.

From a quantitative point of view, significant improvements are achieved by our L_{12}UWR-SENSE and L_{12}CWR-SENSE algorithms in comparison with basic-SENSE and Tikhonov reconstructions. The quantitative measure we used on anatomical data is the SNR in dB computed as follows: $\mathrm{SNR} = 20\log_{10}\left(\|\rho_{\mathrm{True}}\|/\|\rho_{\mathrm{True}} - \widehat{\rho}\|\right)$, where ρ_{True} stands for the reference image and $\hat{\rho}$ is the reconstructed image obtained with any algorithm. Hence, the denominator gives us the amount of residual obtained by any reconstructor. Table 2 reports the SNR values corresponding to the basic-SENSE, the L_2R-SENSE and the proposed L_{12}UWR-SENSE and L_{12}CWR-SENSE techniques for different slices shown in Figs. 3-7.

SNR values are given for different slices for which the SENSE performance strongly varies. The selected slices concern the middle of the brain volume for which full FOV images contain a large area of signal of interest

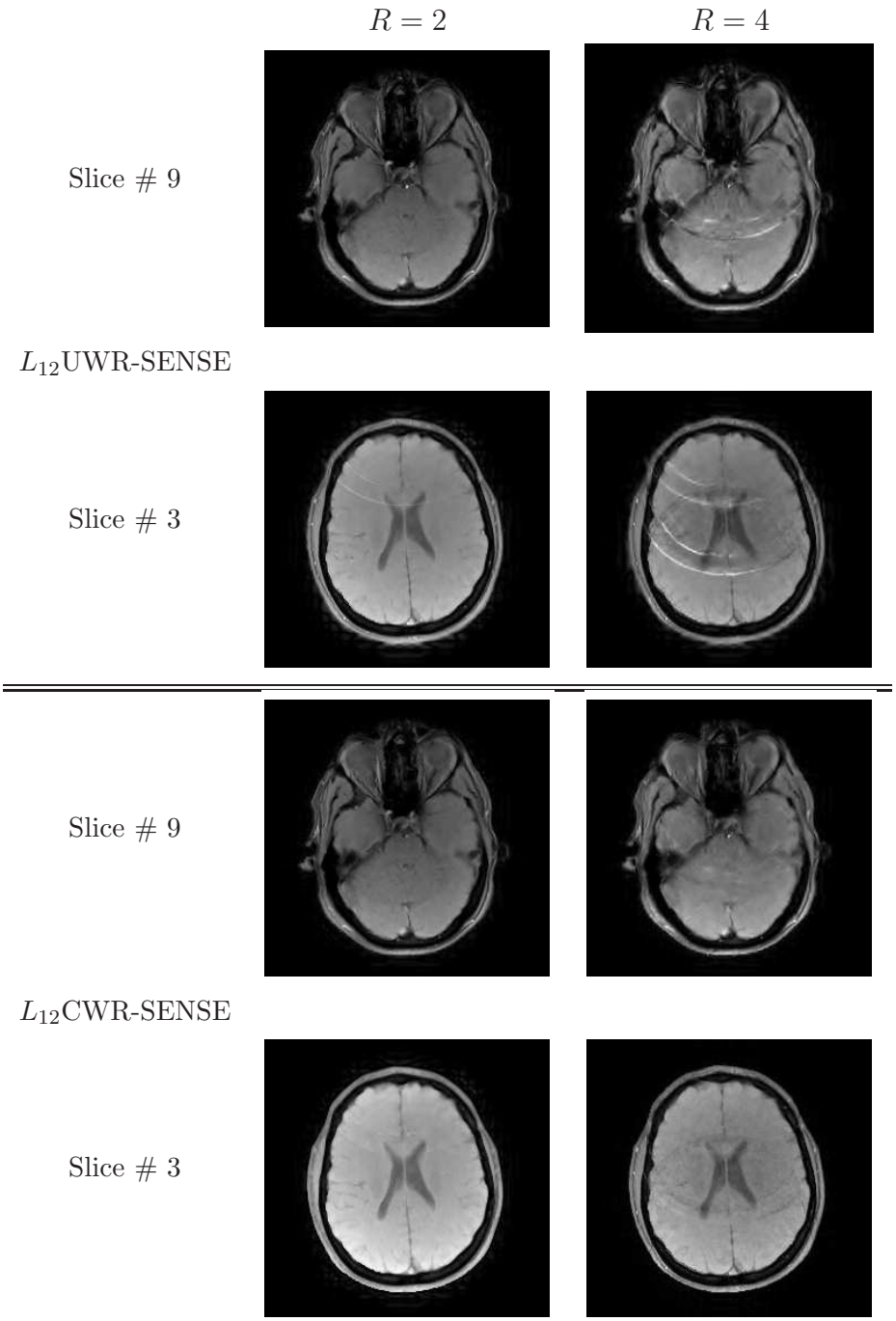

Figure 7: Two reconstructed slices using Algorithm 1 and Algorithm 2 for $R = 2$ and $R = 4$.

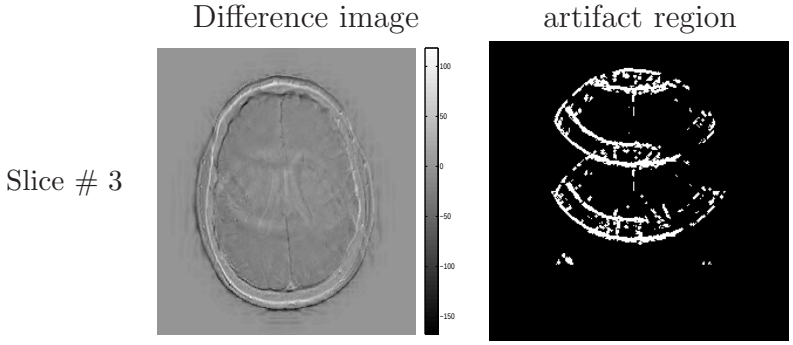

Figure 8: Left: difference image ($R = 4$) between the reference and reconstructed image using L_{12}CWR-SENSE; Right: detected artifact region using a morphological gradient.

Table 2: SNR (in dB) evaluation for reconstructed images using different methods for $R = 4$.

	SNR (dB)			
	SENSE	L_2R-SENSE	L_{12}UWR-SENSE	L_{12}CWR-SENSE
Slice #1	14.34	14.48	14.67	15.53
Slice #2	11.53	11.73	13.72	15.13
Slice #3	12.96	13.37	14.02	14.12
Slice #4	9.22	9.58	10.16	13.10
Slice #5	11.49	11.88	12.06	12.25
Slice #6	9.67	9.80	10.11	11.22
Slice #7	11.04	11.26	11.52	12.00
Slice #8	12.19	12.60	12.92	13.60
Slice #9	13.74	13.28	14.29	15.66
Volume average	11.79	11.99	12.60	13.62

w.r.t. the image background. Interestingly, Table 2 shows that both methods i.e. the L_{12}UWR-SENSE and L_{12}CWR-SENSE algorithms outperform the basic SENSE and L_2R-SENSE reconstructions whatever the slice location. SNR improvement can reach up to 3.6 dB and 3.4 dB when comparing our constrained reconstruction with basic-SENSE and L_2R-SENSE techniques, respectively. The maximal gain we measured between the unconstrained and constrained versions is equal to 2.94 dB as outlined in Table 2. Note also that the SNR improvement is much more significant where SENSE fails to

reconstruct accurate images.

4.1.1. Choice of the wavelet basis

Here, we study how the choice of the wavelet basis may influence the reconstruction performance. For comparison purposes, we present the results obtained with four different wavelet bases: dyadic *Symmlet* 8, dyadic *Daubechies* 8, dyadic *Haar* and *Meyer* with $M = 4$ bands [84]. In Fig. 9, reconstructed anatomical images using the different wavelet bases with $j_{\max} = 3$ are displayed. Some boundary effects appear in the reconstructed images, but with low intensity level. Moreover, they do not affect the brain mask when using the *Symmlet*, *Daubechies* and *Haar* wavelet bases. These additional artifacts become much more important with the *Meyer* 4-band wavelet basis because of its large spatial support. Hence they drastically decrease the SNR of the reconstructed full FOV image. Note also that the *Haar* wavelet basis introduces some blocking effects caused by its discontinuities that do not occur using alternative bases. Among the latters, *Symmlet* 8 gives slightly more accurate regularized results than *Daubechies* 8 but none of them generates significant additional artifacts.

4.1.2. Choice of the maximum resolution level

We focus on the effect of the choice of the maximum resolution level $j_{\max}$ in terms of reconstruction quality. The impact on reconstructed full FOV images can be emphasized through the difference between reconstructed images using 1, 2 and 3 resolution levels. Fig. 10 illustrates the difference between anatomical reconstructed images using one and two (left), or two and three (right) resolution levels.

The difference between anatomical reconstructed images at different resolution levels is significant since it can reach a value of 40 within the intensity range $[0, 255]$. Moreover, the intensity difference seems correlated with the presence of distortions since it appears particularly important in distorted areas. Hence, the higher the maximum resolution level, the better regularized the artifacts are. However, only slight improvements are obtained beyond three resolution levels (results not shown).

Note also that by increasing the number of resolution levels, boundary effects become more visible but they do not affect the brain. Clearly, $j_{\max} = 3$ appears as a fair compromise to achieve an acceptable reconstruction quality. Similar tests have been conducted on EPI images of size 64×64, and led to the same conclusions.

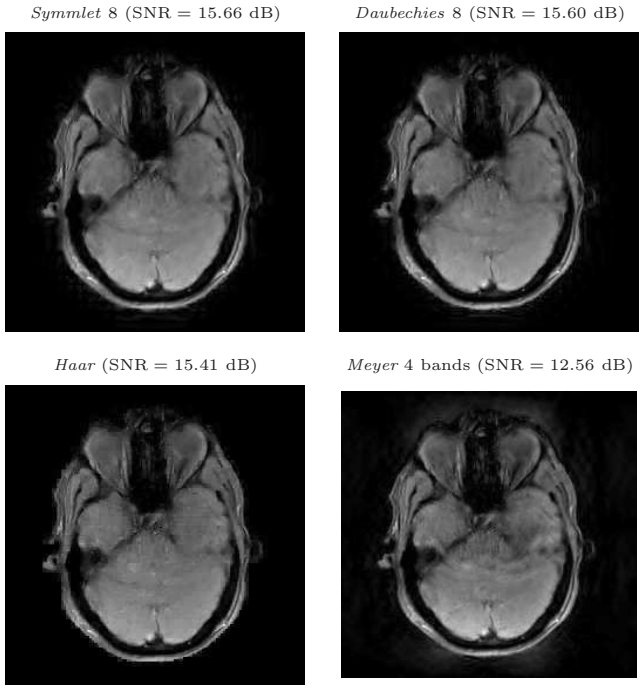

Figure 9: Influence of the wavelet basis on the anatomical reconstructed images using our L_{12}CWR-SENSE method.

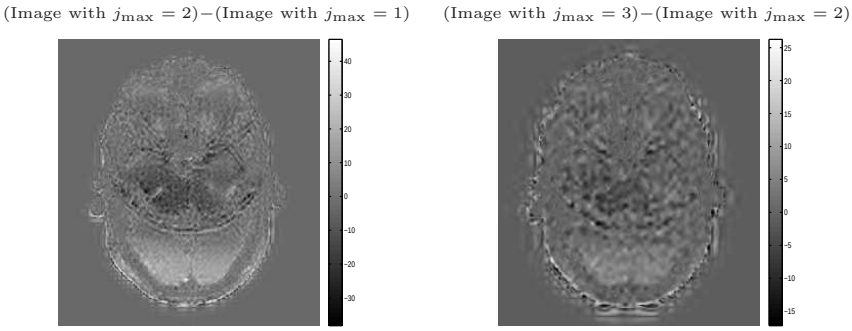

Figure 10: Influence of the number of resolution levels on the anatomical reconstructed images using our L_{12}CWR-SENSE method.

4.2. Results on functional EPI data

This experiment on functional EPI data of size 64×64 was conducted using the same wavelet basis and priors. Algorithm parameters (i.e., relax-

ation λ_n and step-size γ_n parameters) have been adjusted according to the same rules as for anatomical data: $\lambda_n = 1$ and $\gamma_n = 0.99/\theta = 20.63$ were chosen after deriving the θ constant. These EPI fMRI data have been acquired during a short resting state session of 1mn12s.
Fig. 11 illustrates two reconstructed full FOV slices using the four algorithms (basic SENSE, L_2R-SENSE quadratically regularized and our two L_{12}U/CWR-SENSE methods). It can be shown that many defective pixels were corrected using the proposed WT regularization in L_{12}UWR-SENSE without introducing additional artifacts, in contrast with Tikhonov regularization or L_2R-SENSE method. Reconstructed EPI images obtained using L_{12}CWR-SENSE show that the very high intensities were more smoothed than those retrieved using basic-SENSE and L_2R-SENSE methods. Residual defective pixels belonging to distorted areas in SENSE reconstruction have also been removed due to the convex constraints introduced in these areas. Note that the same approach as that applied to anatomical data has been adopted to detect artifact regions and compute the upper and lower bounds defining the convex sets $C_{\boldsymbol{r}}$.

4.3. Comparison with total variation regularization in the image domain

This section aims at comparing the proposed framework with the state-of-the art in edge-preserving regularization, especially with TV regularization performed in the image domain as briefly exposed in Section 2.3. TV regularization has been applied to anatomical and functional EPI data using moderate and large regularization levels. The TV reconstruction results at $R = 4$ are presented in Fig. 12 on anatomical and functional images corresponding to the nineth anatomical and twenty-seventh functional slices, respectively.

Figs. 12(a)-(b) show that some ringing artifacts have been smoothed but the strongest ones still exist. Moreover, the reconstructed images show some staircase effects. This effect is much more visible on functional EPI data at lower spatial resolution. To limit these effects, stronger TV penalization can be used at the expense of a loss in the information content of the reconstructed images as shown in Figs. 12(c)-(d). The comparison of the TV reconstruction results with the proposed unconstrained and constrained approaches for anatomical and functional data indicates that the proposed constrained algorithm L_{12}CWR-SENSE method outperforms TV regularization in the image domain.

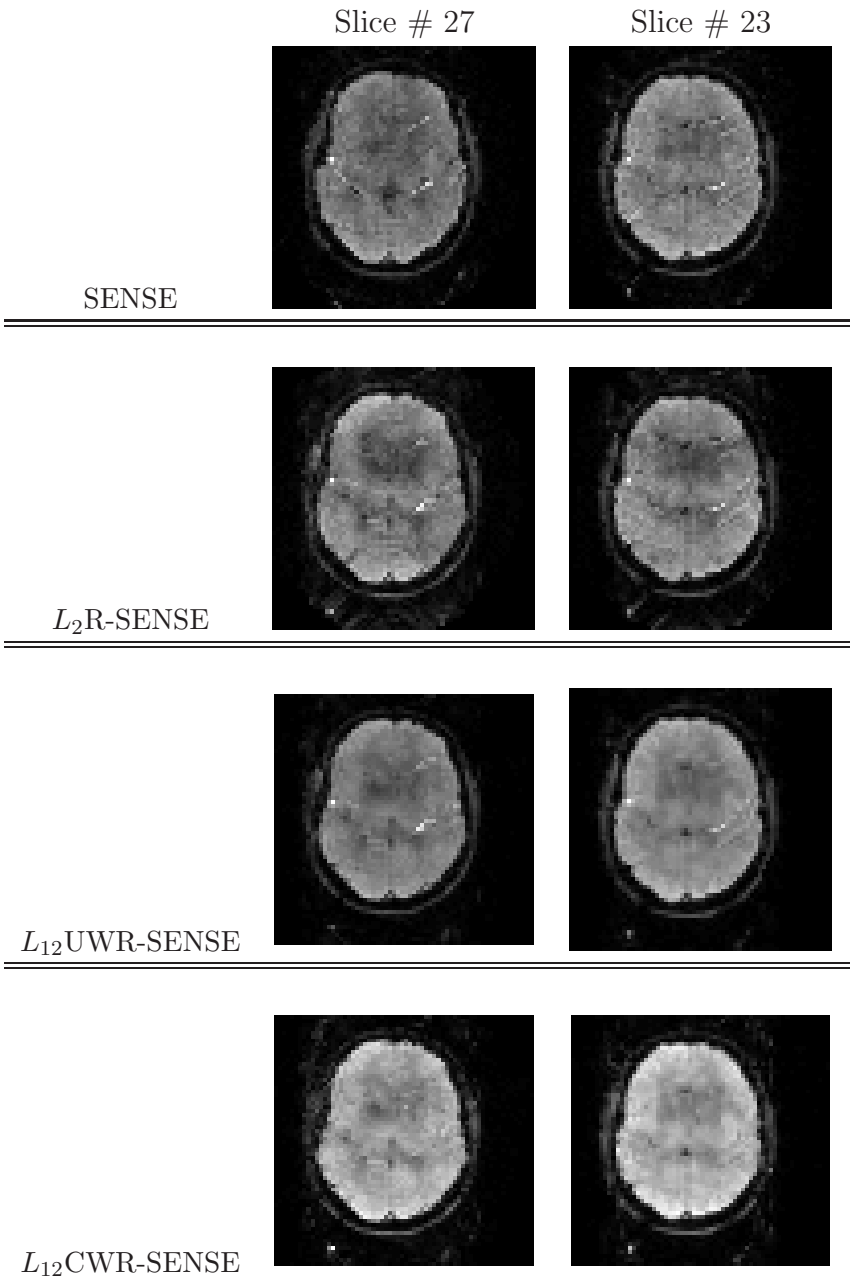

Figure 11: Two EPI reconstructed slices using basic SENSE, L_2R-SENSE, and our L_{12}UWR-SENSE/L_{12}CWR-SENSE for $R = 4$.

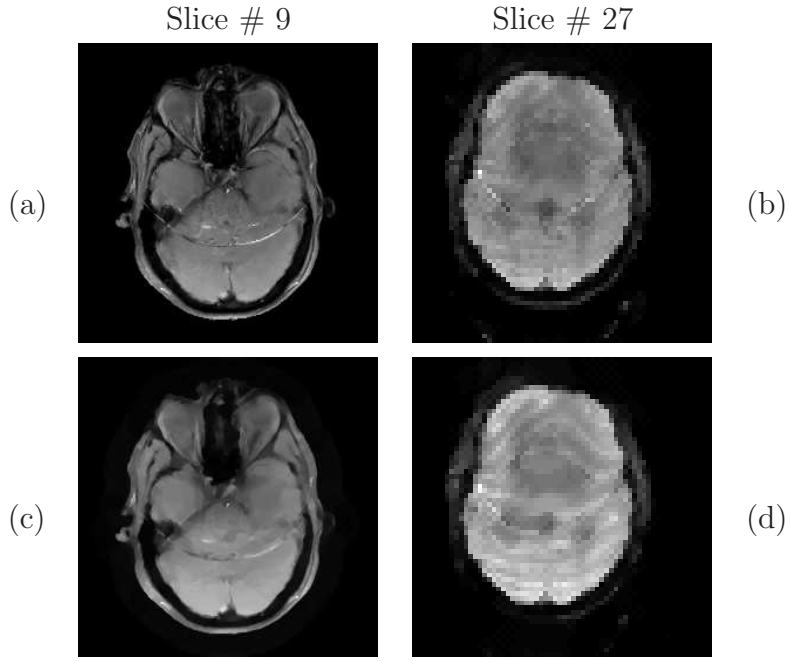

Figure 12: Reconstructed anatomical (left) and functional EPI (right) images using TV regularization in the image domain with moderate (a)-(b) and strong (c)-(d) regularization levels ($R = 4$).

5. Discussion and concluding remarks

In this paper, we proposed and tested a novel approach for SENSE reconstruction based on a regularization in a 2D orthonormal wavelet basis. This method reduces aliasing artifacts in the acquired data, which become critical when using high reduction factors at low magnetic field intensity. Our results on data acquired at 1.5 Tesla show strong improvements in reconstructed anatomical and functional images compared with basic SENSE and standard regularization methods (Tikhonov, Total Variation) that operate in the image domain. Different wavelet bases and settings have been tested and their impact has been evaluated on reconstruction performance. Among our two contributions, the unconstrained algorithm proceeds fastly and performs *unsupervised* reconstruction since all its parameters are automatically tuned from the data or from a reference image, which can result from SENSE reconstruction. Of course, optimal hyper-parameter estimation could be addressed jointly with image reconstruction in the promoted

Bayesian framework although the estimation method may be intensive [71]. In this paper, we just considered a fair compromise between fully automatic regularized reconstruction and efficiency by proposing a method remaining quite robust to hyper-parameter fluctuation. However, our unconstrainted reconstruction may appear inadequate in case of sharp artifacts.
A constrained algorithm has then been designed to cope with this reconstruction issue. It achieves more accurate reconstruction both qualitatively and quantitatively at the expense of the numerical complexity. A Douglas-Rachford subiteration is actually embodied in the optimization procedure to account for additional convex constraints. The crucial point in this method lies in the accurate detection of artifact regions. For doing so, we resort to pre-processings based on the computation of morphological gradients in a reference image resulting from SENSE reconstruction.
Also, some boundary effects have been observed outside the brain mask in wavelet-based reconstructed images. These artifacts may be easily removed by imposing a TV constraint on the background region, or by applying a similar convex contraints to this region (setting the upper and lower intensity bounds close to zero).

Developing iterative solutions for practical problems is a timely topic since an important improvement of computer speed and parallel computing has been reached during the last years. This trend has already been noticed in pMRI reconstruction problems through recent works cited along this paper. The novel methods described in this paper belongs to this research activity since it brings iterative but enhanced pMRI reconstruction solutions to the MRI community. Given the retained Forward-Backward optimization algorithm, one slice of 256×256 anatomical data was reconstructed in 6 seconds and one slice of 64×64 functional data was computed in 0.5 second only using three resolution levels, which indicates that the numerical complexity increases at a sublinear rate with spatial resolution. The proposed method is therefore relatively fast even if it remains almost four times slower than the basic SENSE algorithm. In the context of 1D-SENSE, further reduction of computation time can be achieved using multithreading over slices and parallel computing over volumes in fMRI, i.e. by treating independent data on different CPUs.

Future work will be devoted to a more sophisticated detection of artifact regions since this step directly impacts the result quality. An extension to 3D wavelet decompositions may also be envisaged in two different context: first, to account between-slice spatial dependencies during 3D volume recon-

struction and second, to deal with 2D-SENSE imaging in which the data are actually acquired using 3D imaging sequences [2].
Complementary data analysis on anatomical and functional images acquired at 3 Tesla will also be necessary to assess the robustness of the proposed methodology and its potential interest at higher magnetic fields as well as in the fMRI context where data are recorded during an experimental paradigm. To enlarge the impact of our approach, a special attention will be paid to demonstrate the influence of the reconstruction algorithm in pMRI on brain activity detection and thus on the improvement of the statistical sensitivity/specificity trade-of. A last crucial point will consist of comparing our SENSE-based reconstruction with autocalibrated alternatives like GRAPPA, specifically in the context of brain activity detection from fMRI data acquired during an experimental paradigm where motion artifacts may have a dramatic impact on the reconstruction accuracy and thus, where additional k-space lines acquired in autocalibrated methods may help to limit these effects.
Finally, combined with recent advances in fMRI data analysis [85, 86, 87], our contribution in pMRI reconstruction could help to decrease the required minimum number of subjects to derive neuroscience results at the population level using high resolution imaging.

APPENDIX A
Proof of Proposition 3.4

From Eqs. (14)-(16), it can be seen that $\mathcal{J}_P$ is a convex function such that

$$\forall \zeta \in \mathbb{C}^K, \quad \mathcal{J}_P(\zeta) \geq \frac{\vartheta_1}{2}\|\zeta\|^2 - \vartheta_0$$

where

$$\vartheta_0 = \frac{K_{j_{\max}}}{2}\Big((\frac{\mu^{\mathrm{Re}}}{\sigma^{\mathrm{Re}}})^2 + (\frac{\mu^{\mathrm{Im}}}{\sigma^{\mathrm{Im}}})^2\Big)$$
$$\vartheta_1 = \min\{(\sqrt{2}\sigma^{\mathrm{Re}})^{-2}, (\sqrt{2}\sigma^{\mathrm{Im}})^{-2}, (\beta^{\mathrm{Re}}_{o,j})_{o\in\mathbb{O},1\leq j\leq j_{\max}}, (\beta^{\mathrm{Im}}_{o,j})_{o\in\mathbb{O},1\leq j\leq j_{\max}}\}.$$

This means that $\mathcal{J}_P$ is a strongly convex function with modulus ϑ_L. [6] Since $\mathcal{J}_L$ is a finite convex function, $\mathcal{J}_{\mathrm{WT}}$ also is strongly convex. It is thus strictly convex and coercive (i.e. $\lim_{\|\zeta\|\to+\infty} \mathcal{J}_{\mathrm{WT}}(\zeta) = +\infty$) and, from standard result in convex analysis [88], it can be deduced that $\mathcal{J}_{\mathrm{WT}}$ has a unique minimizer $\widehat{\zeta}$.

In addition, $\mathcal{J}_L$ is a differentiable function and we have

$$\forall \zeta \in \mathbb{C}^K, \qquad \nabla\mathcal{J}_{LT}(\zeta) = \frac{\partial \mathcal{J}_{LT}(\zeta)}{\partial \mathrm{Re}(\zeta)} + \imath \frac{\partial \mathcal{J}_{LT}(\zeta)}{\partial \mathrm{Im}(\zeta)} = T\,\nabla\mathcal{J}_L(T^*\zeta), \qquad (35)$$

where $\mathcal{J}_{LT}(\zeta) = \mathcal{J}_L(T^*\zeta)$. Set $\rho = T^*\zeta$ and $u = \nabla\mathcal{J}_L(\rho)$. It can be then deduced from Eq. (11) that

$$\forall \mathbf{r} \in \{1,\ldots,Y/R\}\times\{1,\ldots,X\}, \qquad \boldsymbol{u}(\mathbf{r}) = 2\boldsymbol{S}^{\mathsf{H}}(\mathbf{r})\boldsymbol{\Psi}^{-1}\left(\boldsymbol{S}(\mathbf{r})\boldsymbol{\rho}(\mathbf{r}) - \boldsymbol{d}(\mathbf{r})\right).$$

where the vector $\boldsymbol{u}(\mathbf{r})$ is defined from u in the same way as $\overline{\boldsymbol{\rho}}(\mathbf{r})$ is defined from $\overline{\rho}$ in Eq. (3). Furthermore, for every $\zeta' \in \mathbb{C}^K$,

$$\|\nabla\mathcal{J}_{LT}(\zeta) - \nabla\mathcal{J}_{LT}(\zeta')\| \leq \|T\|\|u - u'\| \qquad (36)$$

[6] A function $f\colon \chi \to]-\infty,+\infty]$, where χ is a Hilbert space, is said *strongly convex* on χ with modulus $\vartheta_1 > 0$ if there exists some $g \in \Gamma_0(\chi)$ such that $f = g + \frac{\vartheta_1\|\cdot\|^2}{2}$.

where $u' = \nabla \mathcal{J}_L(\rho')$ and $\rho' = T^* \zeta'$. We have then

$$\begin{aligned}
\|u - u'\|^2 &= \sum_{\mathbf{r} \in \{1,\dots,Y/R\} \times \{1,\dots,X\}} \|\boldsymbol{u}(\mathbf{r}) - \boldsymbol{u}'(\mathbf{r})\|^2 \\
&= 4 \sum_{\mathbf{r} \in \{1,\dots,Y/R\} \times \{1,\dots,X\}} \|\boldsymbol{S}^{\mathsf{H}}(\mathbf{r}) \boldsymbol{\Psi}^{-1} \boldsymbol{S}(\mathbf{r}) \left(\boldsymbol{\rho}(\mathbf{r}) - \boldsymbol{\rho}'(\mathbf{r})\right) \|^2 \\
&\leq 4 \sum_{\mathbf{r} \in \{1,\dots,Y/R\} \times \{1,\dots,X\}} \theta_{\mathbf{r}}^2 \|\boldsymbol{\rho}(\mathbf{r}) - \boldsymbol{\rho}'(\mathbf{r})\|^2 \\
&\leq 4\theta^2 \|\rho - \rho'\|^2 \\
&\leq 4\theta^2 \|T\|^2 \|\zeta - \zeta'\|^2. \qquad (37)
\end{aligned}$$

Altogether, Eq. (36) and Eq. (37) yield

$$\|\nabla \mathcal{J}_{LT}(\zeta) - \nabla \mathcal{J}_{LT}(\zeta')\| \leq 2\theta \|T\|^2 \|\zeta - \zeta'\| \qquad (38)$$

which shows that $\mathcal{J}_{LT}$ has a Lipschitz continuous gradient with constant $2\theta\|T\|^2$.

Based on these observations and the fact that,

$$\begin{aligned}
&\forall \zeta = \big((\zeta_{a,k})_{1 \leq k \leq K_{j_{\max}}}, (\zeta_{o,j,k})_{1 \leq j \leq j_{\max}, 1 \leq k \leq K_j}\big), \\
&\quad \operatorname{prox}_{\gamma_n \mathcal{J}_P} \zeta = \big((\operatorname{prox}_{\gamma_n \Phi_a} \zeta_{a,k})_{1 \leq k \leq K_{j_{\max}}}, (\operatorname{prox}_{\gamma_n \Phi_{o,j}} \zeta_{o,j,k})_{1 \leq j \leq j_{\max}, 1 \leq k \leq K_j}\big), \qquad (39)
\end{aligned}$$

the sequence $(\zeta^{(n)})_{n>0}$ built by Algorithm 1 can be rewritten under the more classical Forward-Backward iterative form [49]:

$$\zeta^{(n+1)} = \zeta^{(n)} + \lambda_n \Big(\operatorname{prox}_{\gamma_n \mathcal{J}_P} \big(\zeta^{(n)} - \gamma_n \nabla \mathcal{J}_{LT}(\zeta^{(n)}) \big) - \zeta^{(n)} \Big) \qquad (40)$$

and, due to the Lipschitz differentiability of $\mathcal{J}_{LT}$, the convergence of the algorithm is secured under Assumption 3.3 (see [49, 74]). Furthermore, since $\mathcal{J}_P$ is strongly convex with modulus ϑ_1, we have (see [78] and references therein)

$$\forall n > 0, \qquad \|\zeta^{(n)} - \widehat{\zeta}\| \leq \Big(1 - \frac{\underline{\lambda} \underline{\gamma} \vartheta_1}{1 + \underline{\gamma} \vartheta_1}\Big)^{n-1} \|\zeta^{(1)} - \widehat{\zeta}\| \qquad (41)$$

where $\underline{\gamma} = \inf_{n>0} \gamma_n$ and $\underline{\lambda} = \inf_{n>0} \lambda_n$. This proves that $(\zeta^{(n)})_{n>0}$ converges linearly to $\widehat{\zeta}$.

References

[1] L. Chaâri, J.-C. Pesquet, Benazza-Benyahia, P. Ciuciu, Autocalibrated parallel MRI reconstruction in the wavelet domain, in: IEEE International Symposium on Biomedical Imaging, Paris, France, 2008, pp. 756–759.

[2] C. Rabrait, P. Ciuciu, A. Ribès, C. Poupon, P. Leroux, V. Lebon, G. Dehaene-Lambertz, D. Le Bihan, F. Lethimonnier, High temporal resolution functional MRI using parallel echo volume imaging, Magnetic Resonance Imaging 27 (4) (2008) 744–753.

[3] M. Weiger, K. P. Pruessmann, P. Boesiger, Cardiac real-time imaging using SENSE, Magnetic Resonance in Medicine 43 (2) (2000) 177–184.

[4] Y. Zhu, C. J. Hardy, D. K. Sodickson, R. O. Giaquinto, C. L. Dumoulin, G. Kenwood, T. Niendorf, H. Lejay, C. A. McKenzie, M. A. Ohliger, N. M. Rofsky, Highly parallel volumetric imaging with a 32-element RF coil array, Magnetic Resonance in Medicine 52 (4) (2004) 869–877.

[5] D. K. Sodickson, W. J. Manning, Simultaneous acquisition of spatial harmonics (SMASH): fast imaging with radiofrequency coil arrays, Magnetic Resonance in Medicine 38 (4) (1999) 591–603.

[6] M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, A. Haase, Generalized autocalibrating partially parallel acquisitions GRAPPA, Magnetic Resonance in Medicine 47 (6) (2002) 1202–1210.

[7] P. M. Jakob, M. A. Griswold, R. R. Edelman, D. K. Sodickson, AUTO-SMASH: a self-calibrating technique for SMASH imaging, Magnetic Resonance Materials in Physics, Biology and Medicine 7 (1) (1998) 42–54.

[8] R. M. Heidemann, M. A. Griswold, A. Haase, P. M. Jakob, VD AUTO-SMASH imaging, Magnetic Resonance in Medicine 45 (6) (2001) 1066–1074.

[9] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, P. Boesiger, SENSE : sensitivity encoding for fast MRI, Magnetic Resonance in Medicine 42 (5) (1999) 952–962.

[10] K. P. Pruessmann, M. Weiger, P. Börnert, P. Boesiger, Advances in sensitivity encoding with arbitrary k-space trajectories, Magnetic Resonance in Medicine 46 (4) (2001) 638–651.

[11] W. S. Hoge, D. H. Brooks, B. Madore, W. E. Kyriakos, A tour of accelerated parallel MR imaging from a linear systems perspective, Concepts in Magnetic Resonance 27A (1) (2005) 17–37.

[12] M. Weiger, K. P. Pruessmann, C. Lessler, P. Roschmann, P. Boeisiger, Specific coil design for SENSE: a six-element cardiac array, Magnetic Resonance in Medicine 45 (3) (2001) 495–504.

[13] M. Blaimer, F. Breuer, M. Mueller, R. M. Heidemann, M. A. Griswold, P. M. Jakob, SMASH, SENSE, PILS, GRAPPA: How to choose the optimal method, Magnetic Resonance in Medicine 15 (4) (2004) 223–236.

[14] Z. P. Liang, R. Bammer, J. Ji, N. J. Pelc, G. H. Glover, Making better SENSE: wavelet denoising, Tikhonov regularization, and total least squares, in: International Society for Magnetic Resonance in Medicine, Hawaï, USA, 2002, p. 2388.

[15] L. Ying, D. Xu, Z.-P. Liang, On Tikhonov regularization for image reconstruction in parallel MRI, in: IEEE Engineering in Medicine and Biology Society, San Francisco, USA, 2004, pp. 1056–1059.

[16] F.-H. Lin, K. K. Kwong, J. W. Belliveau, L. L. Wald, Parallel imaging reconstruction using automatic regularization, Magnetic Resonance in Medicine 51 (3) (2004) 559–567.

[17] L. Ying, B. Liu, M. Steckner, G. Wu, M. Wu, S.-J. Li, A statistical approach to SENSE reconstruction with arbitrary trajectories, Magnetic Resonance in Medicine 60 (2) (2008) 414–421.

[18] K. T. Block, M. Uecker, J. Frahm, Undersampled radial MRI with multiple coils. iterative image reconstruction using a total variation constraint, Magnetic Resonance in Medicine 56 (7) (2007) 1086–1098.

[19] A. Ribes, F. Schmitt, Linear inverse problems in imaging, IEEE Signal Processing Magazine 25 (2) (2008) 84–99.

[20] J. Sijbers, D. Poot, A. J. D. Dekker, W. Pintjens, Automatic estimation of the noise variance from the histogram of a magnetic resonance image, Physics in Medicine and Biology 52 (5) (2007) 1335–1348.

[21] A. Tikhonov, Tikhonov regularization of incorrectly posed problems, Soviet Mathematics Doklady 4 (1963) 1624–1627.

[22] A. Griesbaum, B. Kaltenbacher, B. Vexler, Efficient computation of the Tikhonov regularization parameter by goal-oriented adaptive discretization, Inverse Problems 24 (2), 20 pp.

[23] H. W. Engl, Discrepancy principles for Tikhonov regularization of ill-posed problems leading to optimal convergence rates, Journal of Optimization Theory and Applications 52 (2) (1987) 209–215.

[24] P. Ciuciu, J. Poline, G. Marrelect, J. Idier, C. Pallier, H. Benali, Unsupervised robust non-parametric estimation of the hemodynamic response function for any fMRI experiment, IEEE Transactions on Medical Imaging 22 (10) (2003) 1235–1251.

[25] P. Ciuciu, J. Idier, A. Roche, C. Pallier, Outlier detection for robust region-based estimation of the hemodynamic response function in event-related fMRI, in: IEEE International Symposium on Biomedical Imaging, Arlington, USA, 2004, pp. 392–395.

[26] Y. Wang, Z. Haomin, Total variation wavelet-based medical image denoising, International Journal of Biomedical Imaging 6 pp.

[27] S. L. Keeling, Total variation based convex filters for medical imaging, Applied Mathematics and Computation 139 (1) (2003) 101–119.

[28] O. Coulon, D. C. Alexander, S. Arridge, Diffusion tensor magnetic resonance image regularization, Medical Image Analysis 8 (1) (2004) 47–67.

[29] A. Raj, K. Thakur, Fast and stable Bayesian image expansion using sparse edge priors, IEEE Trans. on Image Process. 16 (4) (2007) 1073–1084.

[30] A. Raj, G. Singh, R. Zabih, B. Kressler, Y. Wang, N. Schuff, M. Weiner, Bayesian parallel imaging with edge-preserving priors, Magnetic Resonance in Medicine 57 (1) (2007) 8–21.

[31] B. Liu, L. Ying, M. Steckner, J. Xie, J. Sheng, Regularized SENSE reconstruction using iteratively refined total variation method, in: IEEE ISBI, Washington DC, USA, 2007, pp. 121–124.

[32] B. Liu, K. King, M. Steckner, J. Xie, J. Sheng, L. Ying, Regularized sensitivity encoding (SENSE) reconstruction using Bregman iterations, Magnetic Resonance in Medicine 61 (1) (2008) 145–152.

[33] M. Lustig, D. L. Donoho, P. J. M., Sparse MRI: The application of compressed sensing for rapid MR imaging, Magnetic Resonance in Medicine 58 (6) (2007) 1182–1195.

[34] L. Y. B. Liu, Y. M. Zou, SparseSENSE: application of compressed sensing in parallel MRI, in: IEEE International Conference on Technology and Applications in Biomedicine, Shenzhen, China, 2008, pp. 127–130.

[35] D. Liang, B. Liu, L. Ying, Accelerating sensitivity encoding using compressed sensing, in: IEEE Engineering in Medicine and Biology, Vancouver, Canada, 2008, pp. 1667–1670.

[36] S. Mallat, A wavelet tour of signal processing, Academic Press, New York, 1999.

[37] I. Daubechies, Orthonormal bases of compactly supported wavelets, Communications on Pure and Applied Mathematics 41 (7) (1988) 909–996.

[38] A. Haar, Zur theorie der orthogonalen funktionen systeme, Mathematische Annalen 69 (1910) 331–371.

[39] A. Cohen, I. Daubechies, J.-C. Feauveau, Biorthogonal bases of compactly supported wavelets, Communications on Pure and Applied Mathematics 45 (5) (1992) 485–560.

[40] P. Steffen, P. N. Heller, R. A. Gopinath, C. S. Burrus, Theory of regular M-band wavelet bases, IEEE Transactions on Signal Processing 41 (12) (1993) 3497–3511.

[41] R. R. Coifman, W. V. Wickerhauser, Entropy-based algorithms for best basis selection, IEEE Transactions on Information Theory 38 (2) (1992) 713–718.

[42] D. L. Donoho, I. M. Johnstone, Adapting to unknown smoothness via wavelet shrinkage, Journal of the American Statistical Association 90 (432) (1995) 1200–1224.

[43] D. Leporini, J.-C. Pesquet, Bayesian wavelet denoising: Besov priors and non-Gaussian noises, Signal Processing 81 (1) (2001) 55–67.

[44] P. Mueller, B. Vidakovic, MCMC methods in wavelet shrinkage: Non-equally spaced regression, density and spectral density estimation, Bayesian Inference in Wavelet-Based Models (1999) 187–202.

[45] G. Heurta, Multivariate Bayes wavelet shrinkage and applications, Journal of Applied Statistics 32 (5) (2005) 529–542.

[46] I. Daubechies, M. Defrise, C. DeMol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, Communications on Pure and Applied Mathematics 57 (11) (2004) 1413–1457.

[47] l. Daubechies, G. Teschke, Variational image restoration by means of wavelets: simultaneous decomposition, deblurring and denoising, Applied and Computational Harmonic Analysis 19 (1) (2005) 1–16.

[48] C. Vonesch, M. Unser, A fast thresholded Landweber algorithm for wavelet-regularized multidimensional deconvolutions, IEEE Transactions on Image Processing 17 (4) (2008) 539–549.

[49] C. Chaux, P. Combettes, J.-C. Pesquet, V. Wajs, A variational formulation for frame-based inverse problems, Inverse Problems 23 (4) (2007) 1495–1518.

[50] J. B. Weaver, X. Yansun, D. M. Healy, L. D. Cromwell, Filtering noise from images with wavelet transforms, Magnetic Resonance in Medicine 21 (2) (1991) 288–295.

[51] A. Pizurica, A. M. Wink, E. Vansteenkiste, W. Philips, J. Roerdink, A review of wavelet denoising in MRI and ultrasound brain imaging, Current Medical Imaging Reviews 2 (2) (2006) 247–260.

[52] F.-H. Lin, Y.-J. Chen, J. W. Belliveau, L. L. Wald, A wavelet-based approximation of surface coil sensitivity profiles for correction of image intensity inhomogeneity and parallel imaging reconstruction, Human Brain Mapping 19 (2) (2003) 96–111.

[53] N. Gelman, L. M. Wood, Wavelet encoding for 3D gradient-echo MR imaging, Magnetic Resonance in Medicine 36 (4) (1996) 613–619.

[54] M. Wendt, M. Busch, G. Lenz, J. L. Duerk, J. S. Lewin, R. Seibel, D. Grnemeyer, Dynamic tracking in interventional MRI using wavelet-encoded gradient-echo sequences, IEEE Transactions on Medical Imaging 17 (5) (1998) 803–809.

[55] U. E. Ruttimann, M. Unser, R. R. Rawlings, D. Rio, N. F. Ramsey, V. S. Mattay, D. W. Hommer, J. A. Frank, D. R. Weinberger, Statistical analysis of functional MRI data in the wavelet domain, IEEE Transactions on Medical Imaging 17 (2) (1998) 142–154.

[56] F. G. Meyer, Wavelet-based estimation of a semiparametric generalized linear model of fMRI time-series, IEEE Transactions on Medical Imaging 22 (3) (2003) 315–322.

[57] D. Van De Ville, T. Blu, M. Unser, Integrated wavelet processing and spatial statistical testing of fMRI data, Neuroimage 23 (4) (2004) 1472–1485.

[58] D. Van De Ville, T. Blu, M. Unser, Surfing the brain - An overview of wavelet-based techniques for fMRI data analysis, IEEE Engineering in Medicine and Biology Magazine 25 (2) (2006) 65–78.

[59] A. Mojsilovic, M. Popovic, S. Markovic, M. Krstic, Characterization of visually similar diffuse diseases from B-scan liver images using nonseparable wavelet transform, IEEE Transactions on Medical Imaging 17 (4) (1998) 541–549.

[60] T. Olson, Limited-angle tomography via multiresolution analysis and oversampling, in: IEEE International Symposium on Time-Frequency and Time-Scale Analysis, Philadelphia, USA, 1992, pp. 204–216.

[61] G. Harikumar, Y. Bresler, A new algorithm for computing sparse solutions to linear inverse problems, in: IEEE International Conference on Acoustics, Speech, and Signal Processing, Atlanta, USA, 1996, pp. 1331–1334.

[62] J. A. Tropp, Just relax: convex programming methods for identifying sparse signals in noise, IEEE Transactions on Information Theory 52 (3) (2006) 1030–1051.

[63] M. Elad, B. Matalon, M. Zibulevsky, Coordinate and subspace optimization methods for linear least squares with non-quadratic regularization, Applied and Computational Harmonic Analysis 23 (3) (2007) 346–367.

[64] A. Zymnis, S. Boyd, E. Candes, Compressed sensing with quantized measurements, IEEE Signal Processing Letters 17 (2) (2009) 149–152.

[65] F. Champagnat, Y. Goussard, J. Idier, Unsupervised deconvolution of sparse spike trains using stochastic approximation, IEEE Transactions on Signal Processing 44 (12) (1996) 2988–2998.

[66] S. S. Saquib, C. A. Bouman, K. D. Sauer, ML parameter estimation for Markov random fields with applications to Bayesian tomography, IEEE Transactions on Image Processing 7 (7) (1998) 1029–1044.

[67] A. Dempster, A. Laird, D. Rubin, Maximum likelihood from incomplete data via the EM algorithm (with discussion), Journal of the Royal Statistical Society, Series B 39 (1977) 1–38.

[68] A. Jalobeanu, L. Blanc-Féraud, J. Zerubia, Hyperparameter estimation for satellite image restoration using a MCMC maximum likelihood method, Pattern Recognition 35 (2).

[69] D. P. Bertsekas, Nonlinear programming, Second Edition, Athena Scientific, Belmont, USA, 1995.

[70] S. Ramani, T. Blu, M. Unser, Monte-Carlo SURE: A black-box optimization of regularization parameters for general denoising algorithms, IEEE Transactions on Image Processing 17 (9) (2008) 1540–1554.

[71] L. Chaâri, J.-C. Pesquet, J.-Y. Tourneret, P. Ciuciu, A. Benazza-Benyahia, A hierarchical Bayesian model for frame representation, in: IEEE International Conference on Acoustics, Speech, and Signal Processing, Dallas, USA, 2010, pp. 4086–4089.

[72] C. Chaux, L. Duval, A. Benazza-Benyahia, J.-C. Pesquet, A nonlinear Stein based estimator for multichannel image denoising, IEEE Transactions on Signal Processing 56 (8) (2008) 3855–3870.

[73] J.-J. Moreau, Proximité et dualité dans un espace hilbertien, Bulletin de la Société Mathématique de France 93 (1965) 273–299.

[74] P. L. Combettes, V. R. Wajs, Signal recovery by proximal forward-backward splitting, Multiscale Modeling and Simulation 4 (4) (2005) 1168–1200.

[75] P. L. Combettes, J.-C. Pesquet, Proximal splitting methods in signal processing, `http://arxiv.org/pdf/0912.3522`.

[76] P. Ciuciu, J. Idier, J.-F. Giovannelli, Regularized estimation of mixed spectra using a circular Gibbs-Markov model, IEE Trans. on Signal Processing 49 (10) (2001) 2202–2213.

[77] P. Ciuciu, J. Idier, A half-quadratic block-coordinate descent method for spectral estimation, Signal Processing 82 (7) (2002) 941–959.

[78] C. Chaux, J.-C. Pesquet, N. Pustelnik, Nested iterative algorithms for convex constrained image recovery problems, SIAM Journal on Imaging Sciences 82 (2) (2009) 730–762.

[79] P. L. Combettes, J.-C. Pesquet, A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery, IEEE Journal of Selected Topics in Signal Processing 1 (1) (2007) 564–574.

[80] I. Daubechies, Ten lectures on wavelets, Society for Industrial and Applied Mathematics, Philadelphia, USA, 1992.

[81] J. M. Bioucas-Dias, M. A. T. Figueiredo, A new TwIST: Two-step Iterative Shrinkage/Thresholding algorithms for image restoration, IEEE Transactions on Image Processing 16 (12) (2007) 2992–3004.

[82] A. Beck, M. Teboulle, A Fast Iterative Shrinkage-Thresholding Algorithm for linear inverse problems, SIAM Journal on Imaging Sciences 2 (1) (2009) 183–202.

[83] J. Serra, Image analysis and mathematical morphology, Academic Press, London, 1982.

[84] L. Chaux, L. Duval, J.-C. Pesquet, Image analysis using a dual-tree M-band wavelet transform, IEEE Transactions on Image Processing 15 (8) (2006) 2397–2412.

[85] A. Makni, J. Idier, T. Vincent, B. Thirion, G. Dehaene-Lambertz, P. Ciuciu, A fully Bayesian approach to the parcel-based detection-estimation of brain activity in fMRI, Neuroimage 41 (3) (2008) 941–969.

[86] T. Vincent, L. Risser, P. Ciuciu, Spatially adaptive mixture modeling for analysis of fMRI time series, IEEE Transactions on Medical Imaging 29 (4) (2010) 1059 – 1074.

[87] L. Risser, T. Vincent, F. Forbes, J. Idier, P. Ciuciu, Min-max extrapolation scheme for fast estimation of 3D potts field partition functions. Application to the joint detection-estimation of brain activity in fMRI, Journal of Signal Processing Systems 60 (1).

[88] I. Ekeland, R. Témam, Convex analysis and variational problems, Society for Industrial and Applied Mathematics, Philadelphia, 1999.