

2017 MAGNETIC RESONANCE IMAGING PROJECT - PROJECT REPORT -

Project group: *Incoherence*

Authors:

- Savinien Bonheur
- Anirudh Puligandla
- Darja Stoeva
- Shubham Wagh

Abstract

Decrease time of exposure to magnetization in MR imaging is challenging task. One method that reduces the measuring time is undersampling the measuring data, that leads to an ill-posed inverse problem. In this project three main problems were analyzed. Reconstruction of the image, knowing the sensitivity maps, reconstructing the sensitivity maps, knowing the true image, and joint estimation when they are both unknown. For each of the methods, we proposed a data fidelity and a regularization term, which form a convex or biconvex minimization problem. Then we propose an algorithm that solves the minimization problem iteratively.

1 Introduction

MR imaging has wide applications in the field anatomical fields. This procedure typically involves a few transformations to obtain visible images from mainly, anatomical parts of interest. There have been recent advancements in the way of sampling and reconstructing the received signals. The main goal of this project is to show that MR images can be reconstructed almost perfectly, using under-sampled signals. The project was divided into three broad categories; where, the image was to be reconstructed, knowing the true sensitivity maps; or, the sensitivity maps were to be generated for four sensitivity coils, considering the true reconstructed image is known; or, the third case where both the reconstructed image and the sensitivity maps are unknown and have to be estimated simultaneously. All the methods had to be performed on sub-sampled k -space data.

This report is divided into 4 broad categories, imaging method, calibration method, joint calibration method, each detailing about the 3 goals for the project mentioned in the previous paragraph, and, the 4th section, i.e, section 5, compares the results from our method with the state-of-the-art, ESPIRiT method. The results for each approach are mentioned and detailed in their corresponding sections. While, the last section discusses our conclusions.

2 Fundamentals of magnetic resonance imaging

2.1 Signal Generation

MR imaging techniques are widely used in various fields across the world to study anatomical parts of interest to try to diagnose deadly diseases at an early stage. The fundamentals of magnetic resonance imaging lie in the magnetic properties of subatomic particles. The magnetic spin of subatomic particles has been widely studied and recorded. It is known that the nuclei of an atom have random spins that align either parallel or anti parallel to an externally applied magnetic field. This alignment produces a net micro-magnetic field around the spinning particle, when in the presence of an external magnetic field B_0 . The magnitude of the bulk magnetization \vec{M} produced by the object under and external magnetic field is directly proportional to the applied magnetic field B_0 and the number of spins, N_s , while, it is inversely proportional to the absolute temperature T_s . Generally, the object under scan is measured at room temperature and the N_s is a property of the object. Therefore, one is restricted only to change B_0 to change the magnitude of the bulk magnetization \vec{M} .

RF Excitations are used to induce rotating magnetization about the transverse plane. The nuclei spin randomly when there is no magnetic excitation, and this random spinning results in null transverse magnetic field. Resonance condition is created by establishing phase coherence among the spinning nuclei through RF excitations. The nuclei start to get back to their original orientations once the external magnetic field is removed. This step is called relaxation and the bulk magnetization is observed. When the longitudinal bulk magnetization is recovered, it is called longitudinal magnetization, while, the destruction of transverse magnetic field is called transverse magnetization. Detecting the signal follows the basic principle of Faraday's law to convert the magnetism into electrical signals. The electrical signals are received at the Larmor frequency (low frequency signals) and the high frequency signals are filtered out [1].

2.2 Spatial Localization and sampling

Spatial localization has to be done to locate map the bulk magnetization to the appropriate locations of the object. The main methods for spatial localization are selective excitation and spatial encoding and both the methods involve the use of gradient field. Selective excitation is done through selecting a slice and exciting only the nuclei present inside that slice. This can be done by augmenting the homogenous B_0 field with a linear gradient field during excitation period, that in turn makes the resonance frequency to be position-dependent. The simplest approach to generate slice selective RF pulses is through the Fourier transform. While, spatial information encoding approaches tend to encode the activated MR signal either by frequency encoding or by phase encoding. The frequency-encoding gradient field, by assigning local signals to different frequencies, gradually pushing the local signals out of phase coherence. Hence, the received frequency encoded signal is expected to decay at a faster rate than its non-encoded counterpart. While the phase encoded signal has the form of nonencoded signal with a position-dependent initial phase angle. The role of the frequency encoding gradient is to map a time signal to a k -space signal. The k -space, here, would refer to the dimensions of the desired slice. If constant gradients are used for frequency encoding, the k -space sampling trajectory of a frequency-encoded signal will, simply, be a straight line. While, phase encoding influences only the starting point but not the shape and form of the k -space trajectory for an evolving transient MR signal, and, the shape and form remain unimportant as long as the area under the gradient pulse remains constant. These encoding techniques weight the regions in the k -space that are relevant to the respective coil (sensitivity maps) [1].

In the case of 2-D imaging, a set of identical signals are generated using repetitive excitations and then excite each properly so that k -space will be filled with multiple lines. But, MR imaging is inherently a three-dimensional technique, hence, it is easier to generate signals from the entire volume than from a limited region. The simplest way to sort all the signal components from different spatial locations is to use slice selective excitations for localization in the third dimension while leaving the other two to be done with encoding methods. Typically, in MR imaging, the x-direction is frequency encoded while the y- and z- directions are phase-encoded. Lastly, sampling of a k -space signal can be done either through uniform sampling, interlaced sampling or sampling of bandlimiting periodic functions. The sampling techniques generally follow nyquist criterion. But, there have been recent advancements to generate perfect reconstructions from sub-sampled k -space data.

3 Imaging problem

The imaging problem refers to the estimation of the unknown image, \mathbf{X} , when it is assumed that the sensitivity maps, $(\mathbf{S}_c)_{1 \leq c \leq n_c}$, are known. The inverse problem then becomes ill-posed since the measurement is sub-sampled:

$$(\mathbf{Y}_c)_{1 \leq c \leq n_c} = \Phi^\dagger (\bar{\mathbf{S}}_c)_{1 \leq c \leq n_c} \cdot \mathbf{X} + \mathbf{N}_c, \quad \text{where } \Phi^\dagger = \mathcal{MF} \quad (1)$$

where $(\mathbf{Y}_c)_{1 \leq c \leq n_c}$ is the noisy incomplete Fourier measurement of n_c coils and Φ is the sensing matrix, which in the case of MR imaging is a selection operator \mathcal{M} and a 2D Fourier transform operator \mathcal{F} . Assuming that the additive noise \mathbf{N}_c is an independent identically distributed Gaussian noise, we propose the following data fidelity term:

$$\|\Phi^\dagger (\bar{\mathbf{S}}_c)_{1 \leq c \leq n_c} \cdot \mathbf{X} - (\mathbf{Y}_c)_{1 \leq c \leq n_c}\|_2^2 \leq \varepsilon, \quad \text{where } \|(\mathbf{N}_c)_{1 \leq c \leq n_c}\|_2 < \varepsilon \quad (2)$$

In order to reconstruct the image, we propose two different regularization terms ℓ_1 norm and TV norm. So finally we have two constrained analysis convex minimization problems:

$$\min_{\mathbf{X}} \lambda \|\boldsymbol{\Psi}^\dagger \mathbf{X}\|_1 \quad \text{subject to} \quad \sum_{c=1}^{n_c} \|\boldsymbol{\Phi}^\dagger \bar{\mathbf{S}}_c \cdot \mathbf{X} - \mathbf{Y}_c\|_2^2 \leq \varepsilon \quad (3)$$

$$\min_{\mathbf{X}} \lambda \|\mathbf{X}\|_{TV} \quad \text{subject to} \quad \sum_{c=1}^{n_c} \|\boldsymbol{\Phi}^\dagger \bar{\mathbf{S}}_c \cdot \mathbf{X} - \mathbf{Y}_c\|_2^2 \leq \varepsilon \quad (4)$$

and two unconstrained analysis convex minimization problems:

$$\min_{\mathbf{X}} \lambda \|\boldsymbol{\Psi}^\dagger \mathbf{X}\|_1 + \frac{1}{2} \sum_{c=1}^{n_c} \|\boldsymbol{\Phi}^\dagger \bar{\mathbf{S}}_c \cdot \mathbf{X} - \mathbf{Y}_c\|_2^2 \quad (5)$$

$$\min_{\mathbf{X}} \lambda \|\mathbf{X}\|_{TV} + \frac{1}{2} \sum_{c=1}^{n_c} \|\boldsymbol{\Phi}^\dagger \bar{\mathbf{S}}_c \cdot \mathbf{X} - \mathbf{Y}_c\|_2^2 \quad (6)$$

where λ is a regularization parameter and $\boldsymbol{\Psi}$ is the sparsity basis, the sparse domain of the image. The regularization parameter depends on the gradient of the rough estimate of the image from the measurement and the sensitivity maps. We assume that the image is sparse in the wavelet domain and use the multilevel 2D Daubechies wavelet transform as the sparsity basis. Knowing the MR image will contain a lot of edges and details, the high pass filter is considered when transforming the image. Next we propose an iterative algorithm that solves the unconstrained analysis problem for both regularization terms shown in eq. 5 and 6. Since both of the norms are convex, the regularization term is convex but not differentiable, while the data fidelity term is convex and differentiable. Therefore, we purpose the forward-backward algorithm to solve the forementioned convex optimization problems. The forward-backward algorithm estimates \mathbf{X}^* at each iteration as the minimizer of the objective function (eq. 2) by performing forward step, using the gradient of the data fidelity term and a backward step, by computing the proximity operator of the regularization term. Therefore the estimate \mathbf{X}^* can be found as:

$$\mathbf{X}^* = \arg \min_{\mathbf{X}} g(\mathbf{X}) + f(\mathbf{X}) + \langle \nabla f(\mathbf{X}), \mathbf{X} - \mathbf{X}^* \rangle + \frac{1}{2\delta_t} \|\mathbf{X} - \mathbf{X}^*\|_2^2 \quad (7)$$

where $g(\mathbf{X})$ is the regularization term, $f(\mathbf{X})$ is the data fidelity term and $\delta_t \leq 1/L(\nabla f)$, $L(\nabla f)$ being the Lipschitz constant. The proximity operator of the regularization term is given in eq. (8).

$$\mathbf{X}^* = \text{prox}_{\delta_t g}(\mathbf{X} - \delta_t \nabla f(\mathbf{X})) \quad (8)$$

The gradient step for both minimization problems (eq. 5 and 6) will have the same form and its explicit formula is given in eq. 9

$$\nabla f(\mathbf{X}) = \sum_{c=1}^{n_c} \bar{\mathbf{S}}_c^* \cdot \boldsymbol{\Phi}(\boldsymbol{\Phi}^\dagger \bar{\mathbf{S}}_c \cdot \mathbf{X} - \mathbf{Y}_c) \quad (9)$$

while the proximal step for the ℓ_1 norm is the soft thresholding shown in eq. 10

$$\text{prox}_g = \boldsymbol{\Psi} \text{shrink}(\boldsymbol{\Psi}^\dagger(\mathbf{X} - \delta_t \nabla f(\mathbf{X})), \lambda \delta_t) \quad (10)$$

The proximal step for the TV norm is computed iteratively. The algorithm runs for maximum 1000 iterations, otherwise it stops when the difference between the current and previous estimation of the minimization problem is smaller than five decimal points. The algorithm converges for tolerance values to the four or fifth decimal point for both ℓ_1 and TV norm. The pseudocode of the algorithm is shown in Algorithm 1.

Algorithm 1: Imaging problem algorithm

input : Measurements $((\mathbf{Y}_c)_{1 \leq c \leq n_c})$ and sensitivity maps $((\bar{\mathbf{S}}_c)_{1 \leq c \leq n_c})$
output: Reconstructed image (\mathbf{X}^*)
 Initialize \mathbf{X}_0
 Initialize $h(\mathbf{X}) = g(\mathbf{X}) + f(\mathbf{X})$
for $i \leftarrow 1$ **to** *max iterations* **do**
 $\mathbf{X}_{i+1} = \text{prox}_{\delta_i, g}(\mathbf{X}_i - \delta_i \nabla f(\mathbf{X}_i))$
 if $\frac{|h(\mathbf{X}_{i+1}) - h(\mathbf{X}_i)|}{h(\mathbf{X}_{i+1})} \leq \epsilon$ **then**
 $\mathbf{X}^* = \mathbf{X}_{i+1}$
 break
 end
end

The algorithm was run for 10 different test estimating the unknown image with 50%, 34% and 25% of the measurement data. The average computation time and the average SNR values and the standard deviation of each is given in Table 1. From the results shown in the table it can be concluded that the TV norm gives significantly better reconstruction for all cases with respect to the SNR measure. However, the ℓ_1 norm showed that is less costly to compute. The reason TV is more costly could be due to the iterative computation of the proximity operator. The performance of both methods decreases with decreasing the number of available measured data, while the computational time increases. From the images in figures 1, 2 and 3 it can also be noted that the TV outperforms ℓ_1 , independently of the available measured data. This is due to the detail preservation that the TV norm provides, which is also clear in the difference images. In the case of the ℓ_1 norm, if one looks at the images, it can be seen that the gray matter in the brain is still noise, while the TV shows more smoothed reconstruction.

		Avrg. Time (s)	s.d	Avrg. SNR	s.d.	best SNR
50% data	TV	44.66	9.53	30.51	0.99	31.27
	ℓ_1	22.68	6.68	18.87	0.59	20.09
34% data	TV	111.58	27.29	25.48	1.62	27.97
	ℓ_1	28.78	5.55	16.39	0.60	17.28
25% data	TV	110.25	4.88	19.16	2.21	22.49
	ℓ_1	32.01	10.98	13.26	0.71	14.18

Table 1: Table showing average values for 10 different tests

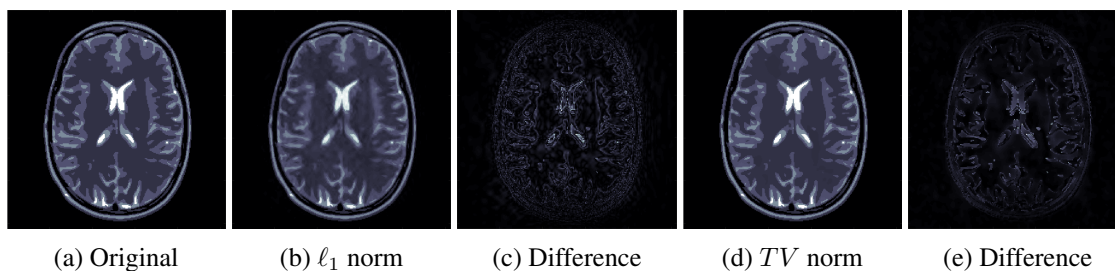


Figure 1: The reconstructed images with the best estimate for both regularization terms and their difference images for using 50% of the measurement data

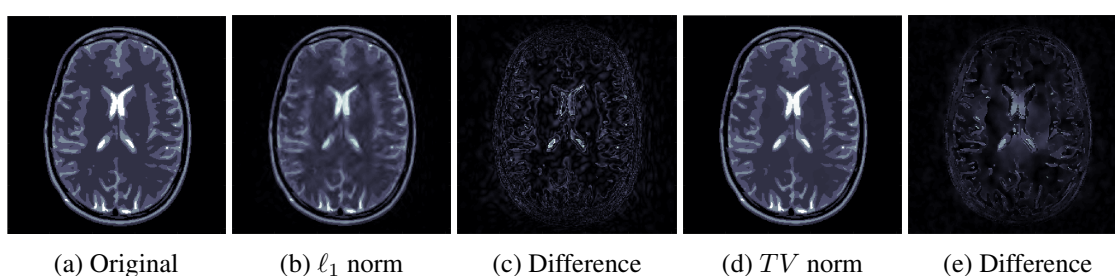


Figure 2: The reconstructed images with the best estimate for both regularization terms and their difference images for using 34% of the measurement data

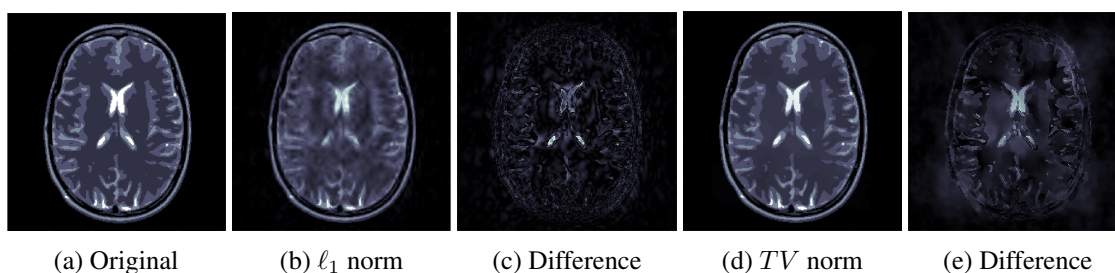


Figure 3: The reconstructed images with the best estimate for both regularization terms and their difference images for using 25% of the measurement data

4 Calibration problem

Calculating coil sensitivities can be considered to be the most important step in any MR imaging technique. The importance of a sensitivity map estimation as a pre-calibration step is clearly explained in [2]. This estimation can broadly be classified into two different approaches. An example of a few state-of-the-art methods can be considered to understand the different approaches. SENSE [2] and ASSET [3], are two state-of-the-art methods that operate in the image space after reconstruction of data from the individual coils. This approach contrasts with GRAPPA [4] method that operates primarily on k-space data before image reconstruction. Most of the major MR vendors offer some version of SENSE technique under different names.

The calibration step can be briefly understood with reference to SENSE. In this step low resolution images are initially acquired separately from each surface coil at full FOV. These images are then normalized by dividing with the low resolution body coil image. Filtering, thresholding and point estimation are then applied to the data to generate coil sensitivity maps. These maps quantify the relative weights for signals originating from different points within the reception area of each coil. A general approach is to pre-calibrate the device before actually taking measurements. This can be achieved by performing a separate 20 second acquisition before the beginning of the actual imaging. Alternatively, automated calibration can be integrated into the pulse sequence itself. This method of auto-calibration is less sensitive to motion artifacts. The calibration method for this project can be generalized as a pre-calibration step considering the true image is known. Having these assumptions at hand, the unconstrained convex minimization problem to reconstruct the sensitivity maps from the sub-sampled k-space data can be written as in eq. 11 and eq. 12, corresponding to two different regularization terms. We have considered ℓ_1 norm and TV norm of each of the sensitivity map \mathbf{S}_c as the two different regularization terms.

$$\min_{(\mathbf{S}_c)_{1 \leq c \leq n_c}} \beta_c \|\Psi^\dagger \mathbf{S}_c\|_1 + \frac{1}{2} \|\Phi^\dagger \mathbf{S}_c \cdot \bar{\mathbf{X}} - \mathbf{Y}_c\|_2^2 \quad (11)$$

$$\min_{(\mathbf{S}_c)_{1 \leq c \leq n_c}} \beta_c \|\mathbf{S}_c\|_{TV} + \frac{1}{2} \|\Phi^\dagger \mathbf{S}_c \cdot \bar{\mathbf{X}} - \mathbf{Y}_c\|_2^2 \quad (12)$$

Where, $\bar{\mathbf{X}}$ corresponds to the true image, β_c is the regularization parameter and ϵ is a known bound on the noise. We have selected 2D Daubechies wavelet transform as the sparsity basis Ψ for the signal. With the selection of the mentioned regularization terms and the sparsity basis, 11 and 12, become sparse denoising problems. The choice of Daubechies wavelets as sparsity basis is an appropriate choice because, the sensitivity maps are smooth in image domain. Both the proposed regularization terms are convex and have a closed form solution that can be computed using the forward-backward algorithm. It is known that forward-backward algorithm provides guaranteed convergence for convex objective functions when the regularization parameter β_c is chosen correctly. We select a different β_c for each coil from the gradient of the measurements in the image domain. The gradient step for both the minimization problems will be the same and can be given by eq. 13.

$$\nabla f((\mathbf{S}_c)_{1 \leq c \leq n_c}) = \bar{\mathbf{X}}^* \cdot \Phi (\Phi^\dagger \mathbf{S}_c \cdot \bar{\mathbf{X}} - \mathbf{Y}_c) \quad (13)$$

While the proximity step for the TV and ℓ_1 norm minimization problems is given as shown in eq. 14 and 15 respectively.

$$\mathbf{S}_c^* = \text{prox}_{\delta_{tg}}(\mathbf{S}_c - \delta_t \nabla f(\mathbf{S}_c)) \quad (14)$$

$$\text{prox}_g = \Psi_{\text{shrink}}(\Psi^\dagger(\mathbf{S}_c - \delta_t \nabla f(\mathbf{S}_c)), \lambda \delta_t) \quad (15)$$

The minimization problem estimates each of the coil sensitivities individually. Hence, the algorithm needs to be repeated for the number of coils n_c to estimate all of the coils sensitivities individually. The pseudocode is the same as in the previous section because, the same algorithm and similar minimization problems are being used for the estimation of the sensitivity maps also. The only difference would be that now the input would be the true image instead of the true sensitivity along with the measurements Y and we would be reconstructing the sensitivities of each coil, \mathbf{S}_c instead of the image, \mathbf{X} (see algorithm 2).

Algorithm 2: Calibration problem algorithm

input : Measurements $((\mathbf{Y}_c)_{1 \leq c \leq n_c})$ and true image $(\bar{\mathbf{X}})$
output: Reconstructed sensitivity maps $((\mathbf{S}_c^*)_{1 \leq c \leq n_c})$
Initialize pair $(\mathbf{S}_{0,c})_{1 \leq c \leq n_c}$
Initialize $h(\mathbf{X}) = g(\mathbf{X}) + f(\mathbf{X})$
for $i \leftarrow 1$ **to** max iterations **do**
 $(\mathbf{S}_{i+1,c})_{1 \leq c \leq n_c} = \text{prox}_{\delta_i, h}((\mathbf{S}_{i,c})_{1 \leq c \leq n_c} - \delta_i \nabla_{(\mathbf{S}_c)_{1 \leq c \leq n_c}} f(\bar{\mathbf{X}}, (\mathbf{S}_{i,c})_{1 \leq c \leq n_c}))$
 if $\frac{|g(\mathbf{S}_{i+1,c})_{1 \leq c \leq n_c} - g((\mathbf{S}_{i,c})_{1 \leq c \leq n_c})|}{g((\mathbf{S}_{i+1,c})_{1 \leq c \leq n_c})} \leq \epsilon$ **then**
 $(\mathbf{S}_c^*)_{1 \leq c \leq n_c} = (\mathbf{S}_{i+1,c})_{1 \leq c \leq n_c}$
 break
 end
end

		Avrg. Time (s)	s.d	Avrg. SNR	s.d.	best SNR
50% data	TV	55.23	5.47	37.87	0.51	38.51
	ℓ_1	59.83	12.04	26.62	0.03	26.67
34% data	TV	58.14	5.52	37.42	0.25	37.81
	ℓ_1	65.88	11.62	26.28	0.10	26.39
25% data	TV	56.08	2.29	36.84	0.23	37.24
	ℓ_1	63.39	2.64	26.04	0.02	26.08

Table 2: Table showing average values for 10 different tests

We have selected DWT and DCT as the sparsity bases, Ψ , for the TV and ℓ_1 minimization problems respectively. The sensitivity maps are sparser in the DCT domain than in DWT domain, and, it was also noted that choosing DCT to be the sparsity basis increases the accuracy of the reconstruction manifold. The tests were run for the TV and ℓ_1 norm regularization terms with different convergence criteria. Differences f , between the values of the minimization problems for the current and previous iterations, was used to stop the algorithm. Through many trials it was noticed that the algorithm can be said to converge for relative tolerance values of $5e-4$ and $1e-5$ for the TV and ℓ_1 norm minimization problems, respectively. It was also noticed that the relative differences start to oscillate beyond the mentioned tolerance limits. The average SNR values increase, an average by about 4dB for the minimization TV problem, and, by about 2dB for the ℓ_1 minimization problem, if the algorithm is allowed to run for 1000

iterations. But, this increase in performance comes at the cost of a 2-fold and a 3-fold increase in the computational time for the two respective cases. The results for the tests are shown in Table 2 and the reconstructed images are shown in Figures 4, 5 and 6 for acceleration values 2,3 and 4 respectively. From the results, we can clearly see that the TV reconstruction outperforms the ℓ_1 reconstruction in all aspects. This is due to the fact that the sensitivity maps are smooth in the image domain and TV reconstruction works on the gradient of the values.

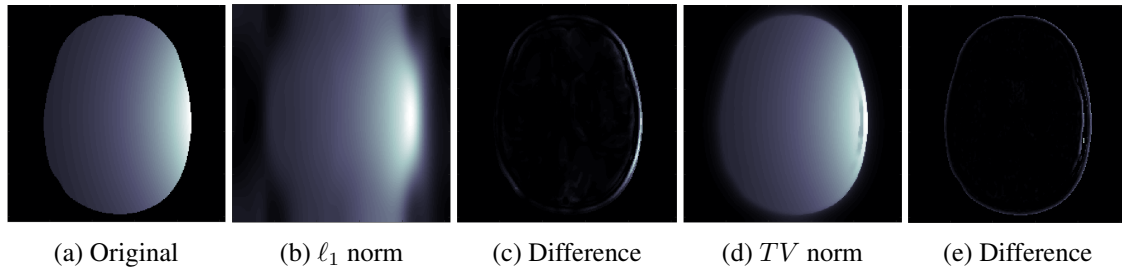


Figure 4: The reconstructed images with the best estimate for both regularization terms and their difference images for using 50% of the measurement data

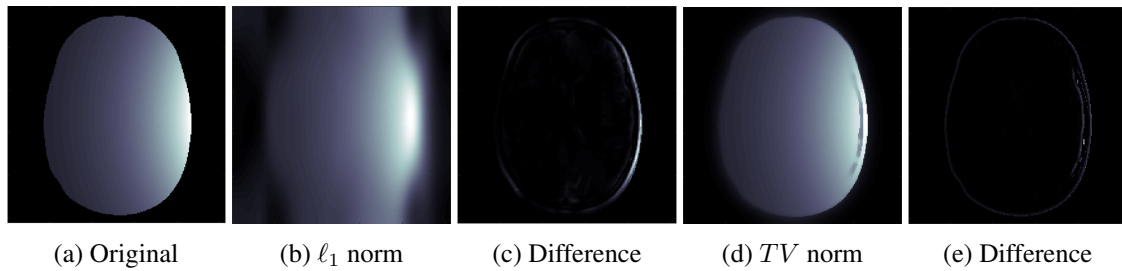


Figure 5: The reconstructed sensitivity map for 1st coil with the best estimate for both regularization terms and their difference images for using 34% of the measurement data

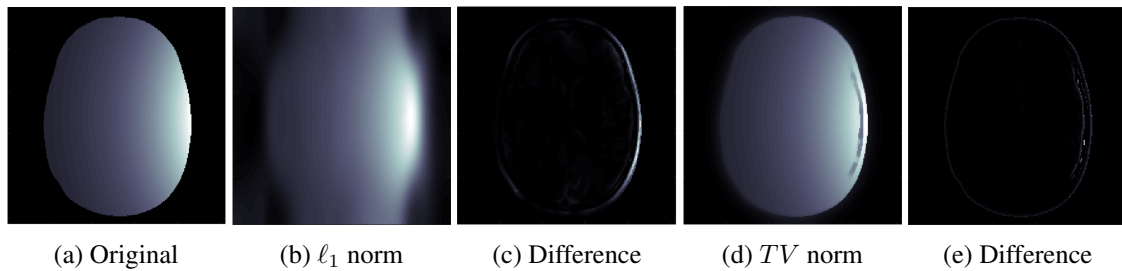


Figure 6: The reconstructed images with the best estimate for both regularization terms and their difference images for using 25% of the measurement data

5 Joint imaging and calibration problem

The joint imaging and calibration problem refers to joint estimation of unknown image \mathbf{X} and sensitivity maps $(\mathbf{S}_c)_{1 \leq c \leq n_c}$, given the measurements for each coil $(\mathbf{Y}_c)_{1 \leq c \leq n_c}$. Taking into account the objective functions minimized in imaging and calibration problems, we propose the following Lipschitz differentiable data fidelity term $f: \mathbb{R}^{N \times N} \times (\mathbb{C}^{N \times N})^{n_c} \rightarrow]-\infty, +\infty]$

$$f(\mathbf{X}, \mathbf{S}_c) = \frac{1}{2} \sum_{c=1}^{n_c} \|\Phi^\dagger \mathbf{S}_c \cdot \mathbf{X} - \mathbf{Y}_c\|_2^2 \quad (16)$$

f is a convex (17) and differentiable (18) function in \mathbf{X} if $(\mathbf{S}_c)_{1 \leq c \leq n_c}$ is fixed as it is quadratic with respect to \mathbf{X} and vice-versa i.e if \mathbf{X} is fixed then the function is quadratic with respect to $(\mathbf{S}_c)_{1 \leq c \leq n_c}$, hence convex (19) and differentiable (20) function in $(\mathbf{S}_c)_{1 \leq c \leq n_c}$ as well.

$$f_{(\mathbf{S}_c)_{1 \leq c \leq n_c}}(\mathbf{X}) := f(\mathbf{X}, \bullet) : \mathbb{R}^{N \times N} \rightarrow]-\infty, +\infty] \quad (17)$$

$$\nabla_{\mathbf{X}} f(\mathbf{X}, \bullet) = \sum_{c=1}^{n_c} \mathbf{S}_c^* \cdot \Phi(\Phi^\dagger \mathbf{S}_c \cdot \mathbf{X} - \mathbf{Y}_c) \quad (18)$$

$$f_{\mathbf{X}}((\mathbf{S}_c)_{1 \leq c \leq n_c}) := f(\bullet, (\mathbf{S}_c)_{1 \leq c \leq n_c}) : (\mathbb{C}^{N \times N})^{n_c} \rightarrow]-\infty, +\infty] \quad (19)$$

$$\nabla_{(\mathbf{S}_c)_{1 \leq c \leq n_c}} f(\bullet, (\mathbf{S}_c)_{1 \leq c \leq n_c}) = \mathbf{X}^* \cdot \Phi(\Phi^\dagger \mathbf{S}_c \cdot \mathbf{X} - \mathbf{Y}_c) \quad (20)$$

Based on the above equations, we can say that it is a biconvex optimization problem. But the function jointly is non-convex [5] in $(\mathbf{X}, (\mathbf{S}_c)_{1 \leq c \leq n_c})$. The quadratic term in \mathbf{X} and $(\mathbf{S})_{1 \leq c \leq n_c}$ inside the l_2 norm square of the data fidelity function f makes it non-convex. They may have lots of local minima, maxima or even saddle point. Therefore, solving for all the variables jointly is an intractable non-convex problem. So, without knowing more about a special case, global optimization is the only alternative. The standard, obvious heuristic for a biconvex problem is to fix $(\mathbf{S})_{1 \leq c \leq n_c}$ and minimize over \mathbf{X} , then fix \mathbf{X} and minimize over $(\mathbf{S})_{1 \leq c \leq n_c}$, and repeat. Hence, the image and the sensitivity maps can be solved iteratively, as :

$$\arg \min_{\mathbf{X}} f(\mathbf{X}, (\mathbf{S}_c)_{1 \leq c \leq n_c}) \quad (21)$$

is solved assuming $(\mathbf{S})_{1 \leq c \leq n_c}$ is fixed, then

$$\arg \min_{(\mathbf{S}_c)_{1 \leq c \leq n_c}} f(\mathbf{X}, (\mathbf{S}_c)_{1 \leq c \leq n_c}) \quad (22)$$

is solved assuming \mathbf{X} is fixed (using the estimate of the image obtained in (21)). As mentioned before, problem (16) is non-convex and there is no guarantee that solving it by iterating between (21) and (22) will converge. However, if sparsity is imposed on the model based on prior information of image and sensitivity maps, the problem can be solved similarly in an iterative fashion to get a sub-optimal solution.

Based on the previous simulations in imaging and calibration methods, we observe that both image and coil sensitivities estimated gave good results for the sparsity base in their respective gradient domains. Likewise, we propose the global unconstrained minimization problem to find an estimate $(\mathbf{X}^*, (\mathbf{S}_c^*)_{1 \leq c \leq n_c})$ of $(\mathbf{X}, (\mathbf{S}_c)_{1 \leq c \leq n_c})$

$$E(\mathbf{X}, (\mathbf{S}_c)_{1 \leq c \leq n_c}) = \arg \min_{\mathbf{X}, (\mathbf{S}_c)_{1 \leq c \leq n_c}} \lambda \|\mathbf{X}\|_{TV} + \sum_{c=1}^{n_c} \beta_c \|\mathbf{S}_c\|_{TV} + \frac{1}{2} \sum_{c=1}^{n_c} \|\Phi^\dagger \mathbf{S}_c \cdot \mathbf{X} - \mathbf{Y}_c\|_2^2 \quad (23)$$

The above problem of interest, solving for the mixing variables \mathbf{X} and $(\mathbf{S}_c)_{1 \leq c \leq n_c}$ can be seen as a blind deconvolution problem. In this framework, it is assumed that both image and coil sensitivities are sparse in their respective gradient domain. Solving for the MR image and the sensitivity maps jointly is an intractable non-convex problem. Following the approach proposed in (21) and (22), we propose to iteratively solve the optimization problem. Therefore the proposed global unconstrained minimization problem can be solved by alternately solving the following two optimization problems.

$$\arg \min_{\mathbf{X}} \lambda \|\mathbf{X}\|_{TV} + \frac{1}{2} \sum_{c=1}^{n_c} \|\Phi^\dagger \mathbf{S}_c \cdot \mathbf{X} - \mathbf{Y}_c\|_2^2 \quad (24)$$

$$\arg \min_{(\mathbf{S}_c)_{1 \leq c \leq n_c}} \beta_c \|(\mathbf{S}_c)_{1 \leq c \leq n_c}\|_{TV} + \frac{1}{2} \|\Phi^\dagger (\mathbf{S}_c)_{1 \leq c \leq n_c} \cdot \mathbf{X} - (\mathbf{Y}_c)_{1 \leq c \leq n_c}\|_2^2 \quad (25)$$

In (24), the image is solved assuming the sensitivity maps are known, by total variation regularization. In (25) the sensitivity map is estimated for each of the coil assuming the image is known, again by total variation regularization. It is seen that each sub-problem in (24) and (25) is convex and thus an optimal solution can be found for each sub-problem. However, these problems require reasonably good initial estimates of both the sensitivity maps and the image.

The problem mentioned in (24) and (25) can be solved in inexact form, by simply performing a single iteration of forward-backward:

Algorithm 3: Proposed algorithm

input : Measurements $((\mathbf{Y}_c)_{1 \leq c \leq n_c})$
output: Reconstructed image (\mathbf{X}^*) and sensitivity maps $(\mathbf{S}_c^*)_{1 \leq c \leq n_c}$
Initialize pair $(\mathbf{X}_0, (\mathbf{S}_{0,c})_{1 \leq c \leq n_c})$
Initialize $E(\mathbf{X}_0, (\mathbf{S}_{0,c})_{1 \leq c \leq n_c})$
for $i \leftarrow 1$ **to** *max iterations* **do**
 $\mathbf{X}_{i+1} \approx \text{prox}_{\delta_{i1}, g}(\mathbf{X}_i - \delta_{i1} \nabla_{\mathbf{X}} f(\mathbf{X}_i, (\mathbf{S}_{i,c})_{1 \leq c \leq n_c}))$
 $(\mathbf{S}_{i+1,c})_{1 \leq c \leq n_c} \approx \text{prox}_{\delta_{i2}, h}((\mathbf{S}_{i,c})_{1 \leq c \leq n_c} - \delta_{i2} \nabla_{(\mathbf{S}_c)_{1 \leq c \leq n_c}} f(\mathbf{X}_{i+1}, (\mathbf{S}_{i,c})_{1 \leq c \leq n_c}))$
 if $\frac{|E(\mathbf{X}_{i+1}, (\mathbf{S}_{i+1,c})_{1 \leq c \leq n_c}) - E(\mathbf{X}_i, (\mathbf{S}_{i,c})_{1 \leq c \leq n_c})|}{E(\mathbf{X}_{i+1}, (\mathbf{S}_{i+1,c})_{1 \leq c \leq n_c})} \leq \epsilon$ **then**
 $\mathbf{X}^* = \mathbf{X}_{i+1}$
 $(\mathbf{S}_c^*)_{1 \leq c \leq n_c} = (\mathbf{S}_{i+1,c})_{1 \leq c \leq n_c}$
 break
 end
end

where, ϵ is tolerance limit

$g(\mathbf{X}) = \lambda \|\mathbf{X}\|_{TV}$

$h((\mathbf{S}_c)_{1 \leq c \leq n_c}) = \beta_c \|(\mathbf{S}_c)_{1 \leq c \leq n_c}\|_{TV}$

$$\nabla_{\mathbf{X}} f = \sum_{c=1}^{n_c} \mathbf{S}_c^* \cdot \Phi(\Phi^\dagger \mathbf{S}_c \cdot \mathbf{X} - \mathbf{Y}_c)$$

$$\delta_{i_1} \leq 1/L(\nabla_{\mathbf{X}} f)$$

$$L(\nabla_{\mathbf{X}} f) = \sum_{c=1}^{n_c} \|\mathbf{S}_c^* \cdot \Phi \Phi^\dagger \mathbf{S}_c\|_2$$

$L(\nabla_{\mathbf{X}} f)$ being the Lipschitz constant of $\nabla_{\mathbf{X}} f$

$L(\nabla_{(\mathbf{S}_c)_{1 \leq c \leq n_c}} f)$ being the Lipschitz constant of $\nabla_{(\mathbf{S}_c)_{1 \leq c \leq n_c}} f$.

$$\nabla_{(\mathbf{S}_c)_{1 \leq c \leq n_c}} f = \mathbf{X}^* \cdot \Phi(\Phi^\dagger \mathbf{S}_c \cdot \mathbf{X} - \mathbf{Y}_c)$$

$$\delta_{i_2} \leq 1/L(\nabla_{(\mathbf{S}_c)_{1 \leq c \leq n_c}} f)$$

$$L(\nabla_{(\mathbf{S}_c)_{1 \leq c \leq n_c}} f) = \|\mathbf{X}^* \cdot \Phi \Phi^\dagger \mathbf{X}\|_2$$

The proximal step for the TV norm is computed iteratively. The algorithm runs for maximum 1000 iterations, otherwise it stops when the relative difference between the current and the previous estimation of the global minimization problem is smaller than five decimal points.

As mentioned earlier, the initial estimates of both the sensitivity maps and image play important role in the global unconstrained minimization problem. We initialize the sensitivity maps and image as follows:

$$(\mathbf{S}_{0,c})_{1 \leq c \leq n_c} = \Phi(\mathbf{Y}_c)_{1 \leq c \leq n_c} \quad (26)$$

$$(\mathbf{X}_c^s)_{1 \leq c \leq n_c} = (\mathbf{S}_{0,c}^*)_{1 \leq c \leq n_c} \cdot \Phi(\mathbf{Y}_c)_{1 \leq c \leq n_c} \quad (27)$$

$$\mathbf{X}_0 = \sqrt{\sum_{c=1}^4 (\mathbf{X}_c^s)^2} \quad (28)$$

where, $(\mathbf{X}_c^s)_{1 \leq c \leq n_c}$ is MRI from each of the coils.

We initialize sensitivity maps as sensitivity encoded image for each of the coils then compute image from each of the coil. Based on this, we get initial estimate of MRI by taking square root on sum of squares of individual images from each coil. The regularization parameters λ and $(\beta_c)_{1 \leq c \leq n_c}$ in (23) control the tradeoff between data fidelity and TV prior. Although automatic optimization of regularization parameters remains an open problem in the context of compressed sensing, we compute λ and $(\beta_c)_{1 \leq c \leq n_c}$ based on the initial estimates of image and sensitivity maps such that both depend on the gradient of the initial estimates of the image and the sensitivity maps respectively.

The algorithm was run for 10 different test estimating both the unknown image and sensitivity maps with 50% of the measurement data using 8 coils. The average computation time and the average SNR values and the standard deviation is given in Table:

		Avg. Time (s)	s.d	Avg. SNR	s.d.	best SNR
50% data	TV	658.69	18	17.49	0	17.60

Table 3: Table showing average values for 10 different tests for Image reconstruction

		Avg. Time (s)	s.d	Avg. SNR	s.d.	best SNR
50% data	TV	658.69	18	1.2926	0.0011	1.2948

Table 4: Table showing average values for 10 different tests for Sensitivity map reconstruction

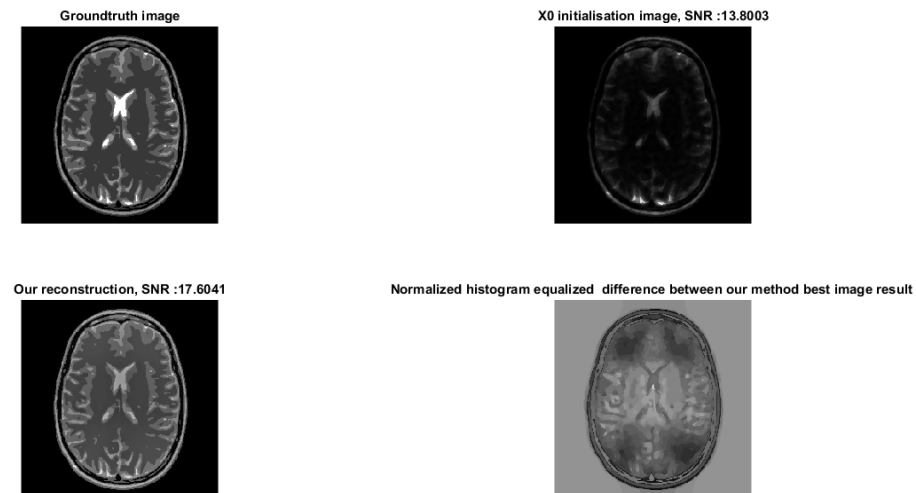


Figure 7: Comparison between the image generated with our method and the groundtruth

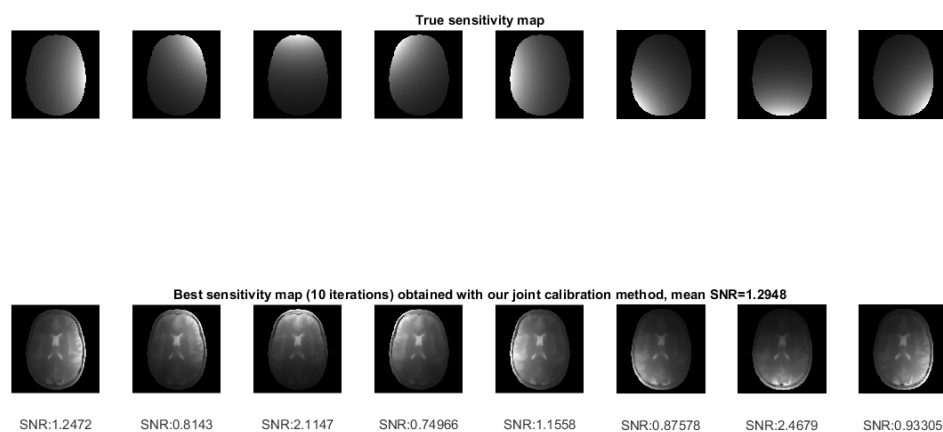


Figure 8: Comparison between the sensitivity map generated with our method and the groundtruth

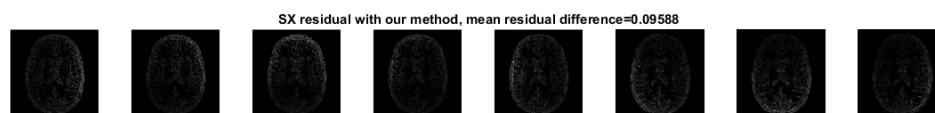


Figure 9: Residuals of the coils measurement generated by our method and the groundtruth

6 Comparison with ESPIRiT

The last part of this report is dedicated to the comparison of our results with a state of the art approach, ESPIRiT.

First published in 2013, ESPIRiT linked theoretically and practically two previous approaches: SPIRiT which uses coil sensitivities, and GRAPPA, which uses learned correlations in k-space.

The first comparisons between ESPIRiT and our joint calibration and imaging method will assess the following : the reconstituted MRI images, the reconstituted sensitivity map, the residual of each method and finally, their run time.// To produce the following plot, we run each algorithms 10 times (N=256, M=32768) with 8 coils.

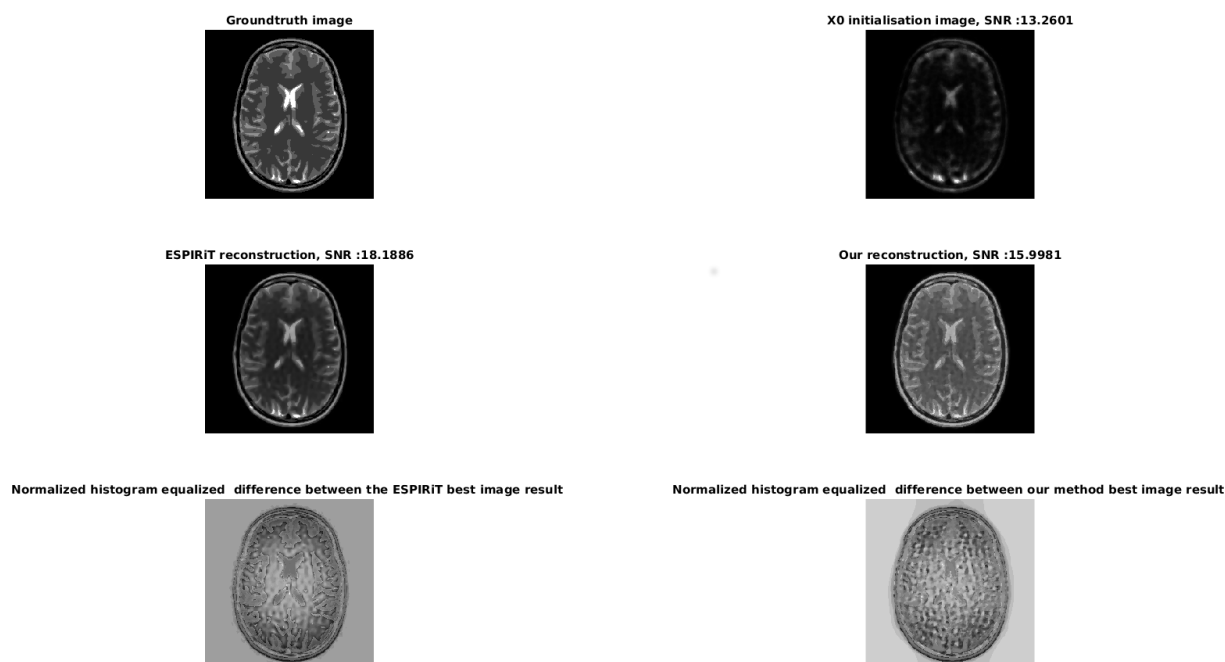


Figure 10: Comparison between the image generated with our method and with ESPIRiT

To compute the SNR between the ground truth and the brain we first applied, for each image, a normalization (to bring the intensity of each image within the same interval) followed by an histogram equalization (to force the intensity of each region to match). Although ESPIRiT leads to a bigger SNR (i.e a closer reconstruction) and to a better contrast, it is affected by the same noise (the difference in contrast is due to a better noise removal from ESPIRiT). Interestingly, the very same noise is also present in the original image. As this noise is composed of strong gradient, it seems amplified by the tv norm (the tv norm promotes sharp gradient). Removing this noise (by pre-processing the initialisation X0) on the original image might lead to better results.

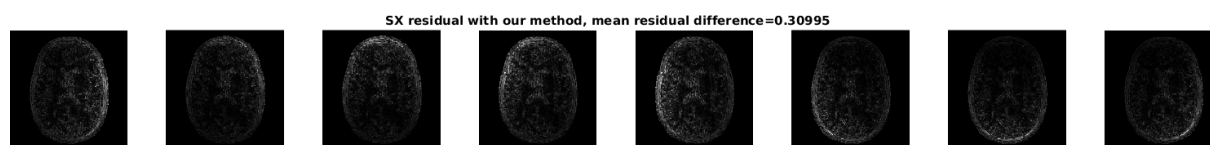


Figure 11: Residuals of the coils measurement generated by our method and by the ESPIRiT method and the groundtruth

This plot only confirm the observation made before. As the noise affect more our method, the features are less visible on our residual.

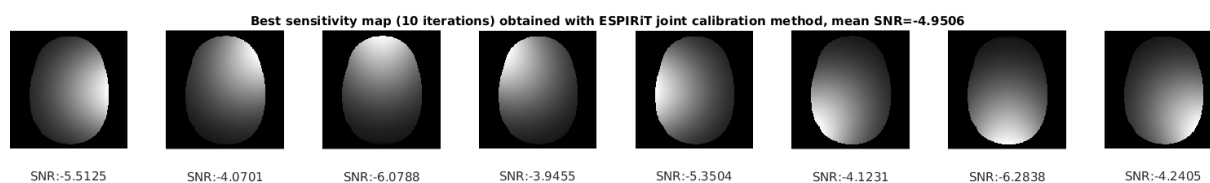
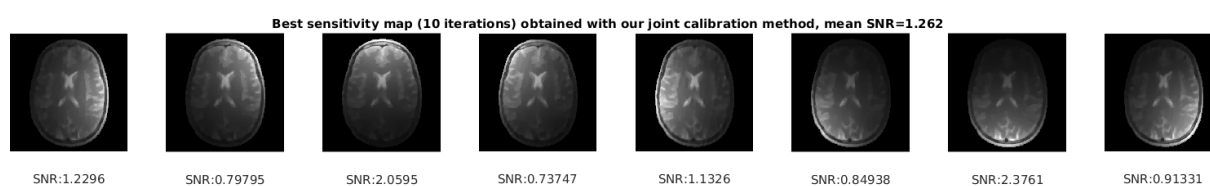
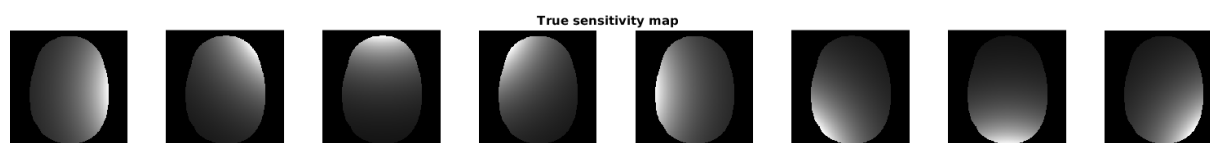


Figure 12: Comparison between the sensitivity map generated with our method and with ESPIRiT

In these image we compare the real part of the cropped (the sensitivity is not important where there is no measurement) reconstructed sensitivity maps. The complex part was ignored due to the complex SNR being the same for our method and for ESPIRiT. We can observe that the SNR of the reconstructed ESPIRiT sensitivity map are negative, however, due to time imperative, we weren't able to understand

why.

Due to our use of the TV norm both for the calibration and the imaging, our algorithm cannot discriminate effectively between the reconstructed image and the sensitivity map (thus the feature of the brain appearing in the sensitivity maps).

On a Intel i7-3630QM we obtained the following performance :

	ESPIRiT	our method
Average Computation time (s)	20.2824	369.346
STD Average Computation time (s)	0.2254	14.1129
mean SNR image	18.1886	15.6914
STD SNR image	0	0
mean SNR sensitivity	-4.9506	18.1886
STD SNR sensitivity	0	0.0013

Figure 13: Comparison table between our method and ESPIRiT

We can observe that, while being about 18 times faster, leads to results about twice as close (3 dB of difference for the SNR) to the original image.

Our second comparison will compare the influence of the calibration surface influence depending on its size (x axis)

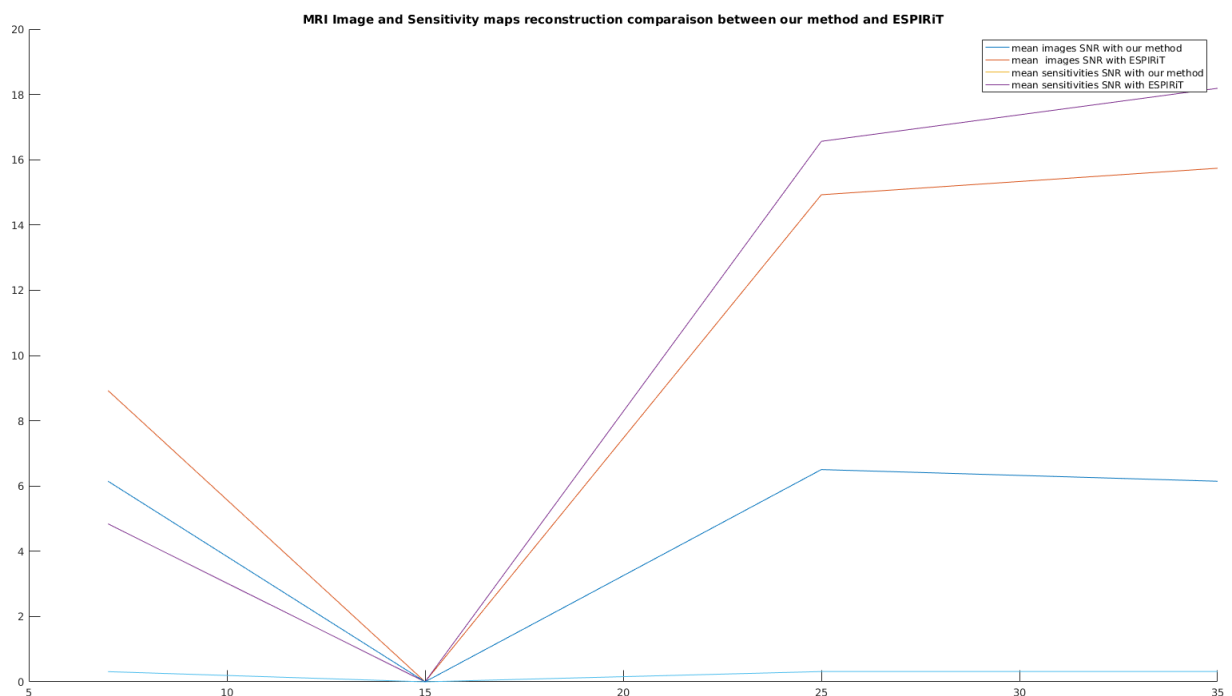


Figure 14: Comparison between ESPIRiT and our method for different windows size

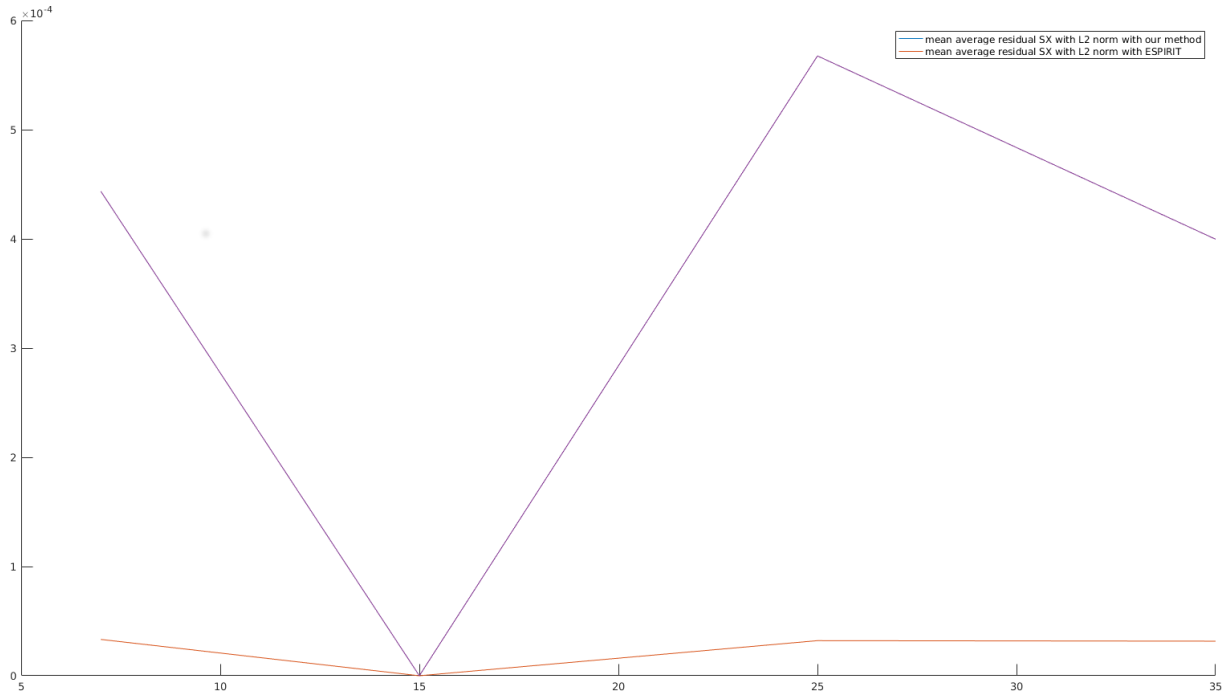


Figure 15: Comparison between ESPIRiT and our method for different windows size

Those two plots demonstrate the superiority of ESPIRiT independently of the windows size.

7 Conclusions and discussion

Algorithms, to individually reconstruct either the image or the sensitivity maps, and also to jointly estimate both the image and sensitivity maps simultaneously, were developed and tested successfully. Forward-Backward algorithm was used for the first two problems. A heuristic approach to solve bi-convex problems was also proposed for the third problem. Two different regularization parameters, l_1 and tv norm, were also proposed for the first two problems. Two different sparsity bases, DCT and DWT, were tested for the calibration method.

During the tests, it was also noticed that the initial estimates for the image and the sensitivity maps greatly affect the reconstruction, and an automatic selection of the initial estimates was used for the joint calibration method. The tests were done 10 consecutive times for 50%, 34% and 25% of the data, respectively, and satisfactory results were obtained. On comparison with the state of the art, ESPIRiT method, it was noticed that ESPIRiT method outperforms, although by a very small margin, from our proposed method.

References

- [1] Liang, Z.P., Lauterbur, P.C.: Principles of Magnetic Resonance Imaging: A Signal Processing Perspective. Wiley-IEEE Press (1999)

- [2] Pruessmann, K., Weiger, M., Scheidegger, M., Boesiger, P.: SENSE: Sensitivity encoding for fast MRI. *Magn. Reson. Med.* **42**, 952–962 (1999)
- [3] King, K.F.: Assetparallel imaging on the ge scanner. 2nd international workshop on parallel MRI **October**, pp. 15–17 (2004)
- [4] Griswold, M., Jakob, P., Heidemann, R.M., Nittka, M., Jellus, V., Wang, J., Kiefer, B., Haase, A.: Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magn. Reson. Med.* **47**, 1202–1210 (2002)
- [5] Ying, L., Sheng, J.: Joint image reconstruction and sensitivity estimation in SENSE (JSENSE). *Magn. Reson. Med.* **57**, 1196–1202 (2007)