A Comparison of Iterative Thresholding Algorithms for the MEG Inverse Problem

Vittoria Bruni, Francesca Pitolli, Cristina Pocci
Dept. SBAI, Università di Roma "La Sapienza"
Via Antonio Scarpa 16, 00161 Roma, Italy
{*vittoria.bruni,francesca.pitolli,critina.pocci*}*@sbai.uniroma1.it*

Abstract

The magnetoencephalography (MEG) aims at reconstructing the unknown electric activity in the brain from the measurements of the magnetic field in the outer space. The MEG inverse problem is ill-posed and/or ill-conditioned thus further constraints are needed to guarantee a unique and stable solution. Assuming that neural sources are confined in small regions of the brain, the sparsity constraint can be used as a regularization term. Thus, the solution of the inverse problem can be approximated by iterative thresholding algorithms. In order to identify an efficient inversion method for the MEG problem, we compare the performance - efficiency, accuracy, computational load - of some thresholding algorithms when localizing a single neural source. The numerical tests will give some suggestions on the construction of an efficient algorithm to be used in real life applications.

Keywords: Magnetoencephalography, Sparsity constraint, Iterative thresholding algorithm

1. Introduction

A magnetoencephalographic (MEG) system measures the neuromagnetic field generated outside the head by the electric activity of the brain. MEG devices sample the magnetic field on N sites, say $\vec{\mathbf{q}}_l$, $l = 1, \ldots, N$, located on a 'helmet' external to the head. Here, we consider the case of devices equipped with magnetometers. These sensors measure just the projection of the magnetic field along the direction $\vec{\mathbf{e}}(\vec{\mathbf{q}}_l)$, which is usually the normal with respect to the

magnetometer coil (for a complete review on the magnetoencephalography see, for instance, [4] and [11]).

Actually, the quantity of interest is the electric current distribution flowing inside the brain since its image gives insight on the brain functionality. As a consequence, the use of the magnetoencephalography as a neuroimaging technique requires the solution of an inverse problem.

Although the generation of the neuromagnetic field by a neuroelectric source is a complex phenomenon, the assumption of linear dependence of the observations - the magnetic data - on the quantity of interest - the neuroelectric current distribution - can be feasible. In fact, the projection $B_e(\vec{\mathbf{q}}_l)$ along $\vec{\mathbf{e}}(\vec{\mathbf{q}}_l)$ of the neuromagnetic field measured in $\vec{\mathbf{q}}_l$ is related to the total current $\vec{\mathbf{J}}$ flowing inside the brain volume V_0 by the Biot-Savart law [15]

$$B_e(\vec{\mathbf{q}}_l) = \frac{\mu_0}{4\pi} \int_{V_0} \left(\vec{\mathbf{e}}(\vec{\mathbf{q}}_l) \times \frac{\vec{\mathbf{r}}' - \vec{\mathbf{q}}_l}{|\vec{\mathbf{r}}' - \vec{\mathbf{q}}_l|^3} \right) \cdot \vec{\mathbf{J}}(\vec{\mathbf{r}}\,')\, d\vec{\mathbf{r}}\,' \,. \tag{1}$$

Here $\vec{\mathbf{v}} \times \vec{\mathbf{w}}$ and $\vec{\mathbf{v}} \cdot \vec{\mathbf{w}}$ are the usual cross and scalar products of vectors in $\mathbb{R}^3$, respectively, and $|\vec{\mathbf{v}}|$ is the Euclidean norm. Usually, the total current is modeled as a sum of a finite number of *elementary sources*, i.e.

$$\vec{\mathbf{J}}(\vec{\mathbf{r}}\,') = \sum_{k=1}^{M} \vec{\mathbf{J}}_k \, \psi_k(\vec{\mathbf{r}}\,') \,, \tag{2}$$

where $\mathbf{J}_k = (J_k^x, J_k^y, J_k^z)$ is the current intensity of the k elementary source having spatial distribution ψ_k.

Now, let $T = (t_{lk}^{\ell})$, $l = 1, \ldots, N$, $k = 1, \ldots, M$, $\ell = x, y, z$, be the *linear mapping* relating the measurement vector $G = (g_l)$, $l = 1, \ldots, N$, to the unknown intensity vector $J = (J_k^{\ell})$, $k = 1, \ldots, M$, $\ell = x, y, z$, i.e.

$$G = T\, J \,. \tag{3}$$

Thus, the explicit expression of the entries of T can be obtained by inserting (2) in (1):

$$t_{lk}^{\ell} = \frac{\mu_0}{4\pi} \int_{V_0} \left(\frac{\vec{\mathbf{e}}(\vec{\mathbf{q}}_l) \times (\vec{\mathbf{r}}' - \vec{\mathbf{q}}_l)}{|\vec{\mathbf{r}}' - \vec{\mathbf{q}}_l|^3} \right)_{\ell} \psi_k(\vec{\mathbf{r}}\,')\, d\vec{\mathbf{r}}\,' \,. \tag{4}$$

T is usually called the *lead field matrix* since each row $(t_{lk}^{\ell})_{l=1,\cdots,N,\ell=x,y,z}$ represents the contribution of the k elementary source to the overall magnetic field. Usually, the linear mapping T is not invertible or is ill-conditioned [11]. Moreover, the data G can be incomplete and corrupted by noise. Thus, in order to reconstruct the (unknown) intensity vector J from the measurements G, further constraints on the solution are needed, i.e. a regularization technique is required [1],[7],[13],[16].

Facing the problem of reconstructing localized neuroelectric current, it is rather natural to assume that the electric current distribution is *spatially sparse* with respect to a pre-assigned basis [9]. This means that $\vec{\mathbf{J}}$ in (2) can be well approximated by a series expansion with only a small number of non-vanishing

coefficients [5]. Sparsity also means that only little information is conveyed by $\vec{\mathbf{J}}$. Thus, it is reasonable to expect that only few measurements might be sufficient to characterize and reconstruct the unknown vector J. Assuming that J can be compressed by the basis (ψ_k), sparsity can be modeled by requiring that the vector J is contained in $\ell_1(K)$, where K is a *small* subset of $\{1,\ldots,M\}$. Indeed, the minimization of the $\ell_1(K)$-norm promotes that only a few entries of J are non-zero [14].

The solution of the linear system (3) under sparsity constraints is obtained by minimizing the functional

$$\mathcal{J}(J) = \|G - TJ\|_2^2 + \Psi(J)\,, \tag{5}$$

where $\Psi(J)$ is a sparsity measure. Thus, sparsity acts as a regularization term.

The minimizer of $\mathcal{J}(J)$ with respect to J can be approximated by an iterative thresholding algorithm. The algorithm starts from an initial guess $J^{(0)}$ and iteratively updates the approximation by minimizing $||G - TJ||_2^2$ in the direction of its gradient and promoting the ℓ_1 minimization via thresholding [3],[6]. Denoting by $\mathcal{S}_\sigma$ the *thresholding operator* with threshold σ, the nth iteration step is

$$J^{(n+1)} = \mathcal{S}_\sigma\big(J^{(n)} + \gamma^{(n)} T^*(G - T\,J^{(n)})\big)\,, \qquad n \geq 0\,, \tag{6}$$

where $\gamma^{(n)}$ is a (fixed or adaptive) *descent parameter* that can be used to speed up the convergence rate. The algorithm was proved to be convergent. In practice, the iterations are stopped when $\|G - T\,J^{(n)}\|_2 \leq \eta$, where η is an estimation of the *noise level* of the data.

In the literature many sparsity measures have been proposed (see, for instance, [2], [3],[6],[10],[12]) giving rise to different thresholding operators and algorithms. In this paper we want to compare the performance of some sparsity measures based on the ℓ_1-norm in the solution of the MEG inverse problem. As the numerical tests will show, the sparsity constraints are suitable for the localization of deep focal sources that cannot be localized by the classical Tikhonov regularization based on the ℓ_2-norm [16].

2. Numerical tests

To perform the numerical tests we assume the k elementary source is a point-like current located in $\vec{\mathbf{r}}_k \in V_0$, i.e. $\psi_k(\vec{\mathbf{r}}') = \delta(\vec{\mathbf{r}}_k - \vec{\mathbf{r}}')$. Thus, the entries of the lead field matrix (4) become

$$t_{lk}^{\ell} = \frac{\mu_0}{4\pi}\frac{V_0}{M}\left(\frac{\vec{\mathbf{e}}(\vec{\mathbf{q}}_l)\times(\vec{\mathbf{r}}_k - \vec{\mathbf{q}}_l)}{|\vec{\mathbf{r}}_k - \vec{\mathbf{q}}_l|^3}\right)_{\ell}\,. \tag{7}$$

The set up of the numerical experiment is as follows. The conducting volume V_0 is a homogeneous sphere of radius $R = 8$ *cm*. Since deep sources produce a very small magnetic field in the outer space and cannot be detected by magnetic data, we do not put any elementary source in the core of V_0, i.e. a small sphere

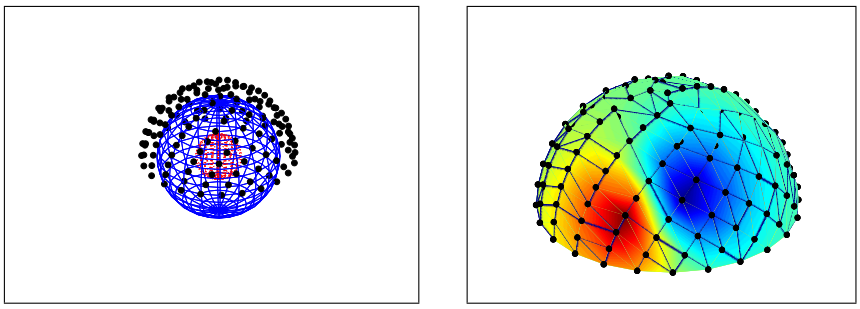

Figure 1: The conducting volume (left) and the magnetic field (right) generated by a single current dipole at depth 1 cm. The conducting volume consists in the spherical shell between the external (blue) sphere and the internal (red) sphere. The black circles represent the magnetometer sites.

of radius 3 *cm* concentric with V_0. The magnetic field in the outer space is generated by a single neuroelectric source which is described as a current dipole located at different depth - between 0.6 to 4.6 cm - below the surface of V_0 (cf. [15]). The magnetic data are obtained by the radial projection of the generated magnetic field sampled in $N = 144$ sites uniformly distributed on a spherical surface of radius 10 *cm* concentric with V_0 (see Fig. 1).

The first test concerns a comparison of the different lead field matrices we obtain when using different distributions of the elementary sources inside the conducting volume V_0. High accuracy reconstruction requires hundreds of sources located in a dense uniform grid [17], usually having a spatial resolution of 3-4 mm. In the following tests we use a grid of 4 mm resolution resulting in about 30000 elementary sources inside V_0. To be precise, $M = 29336$ so that the lead field matrix T has 144 rows and 88008 columns. Since the entries of T corresponding to deep elementary sources have small magnitude, the lead field matrix is preconditioned by a column balancing, i.e. each column is divided by its norm. Fig. 2 shows the structure of the lead field matrix T (left) and of the preconditioned matrix T_c (right).

The use of a uniform source distribution may be computationally demanding when solving the MEG inverse problem using a real head geometry since in this case the resulting inverse problem has very large dimension. To reduce the computational load it could be better to solve a number of inverse problems of small dimension using for each trial a different distribution of few elementary sources. Thus, we locate about 1000 sources, randomly distributed in V_0, and solve 30 different inverse problems. High accuracy can be recovered by collecting together the solutions of all the trials. The structure of the lead field matrix and of the preconditioned matrix for a single small problem is shown in Fig. 3. To give an idea of the conditioning of the lead field matrix in Fig. 4 the amplitude of the singular values of the matrices T and T_c displayed in Figs. 2-3 are shown.

The second test concerns the solution of the inverse problem. In particular, we compare the accuracy of different inversion methods in localizing a current

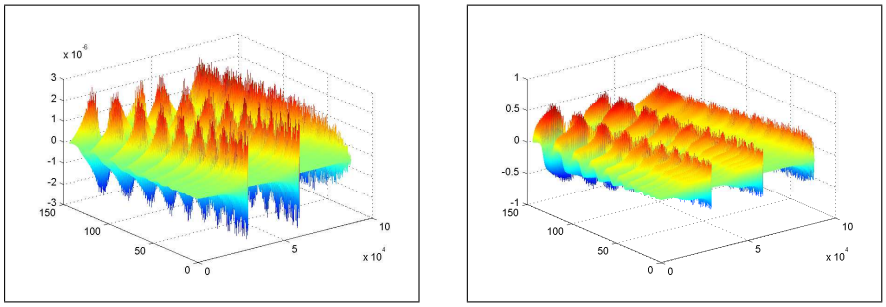

Figure 2: The lead field matrix T (left) and the preconditioned matrix T_c (right) when the elementary sources are distributed on a uniform grid. In the pictures the matrix entries (in arbitrary units) are displayed as a function of the row and column indexes.

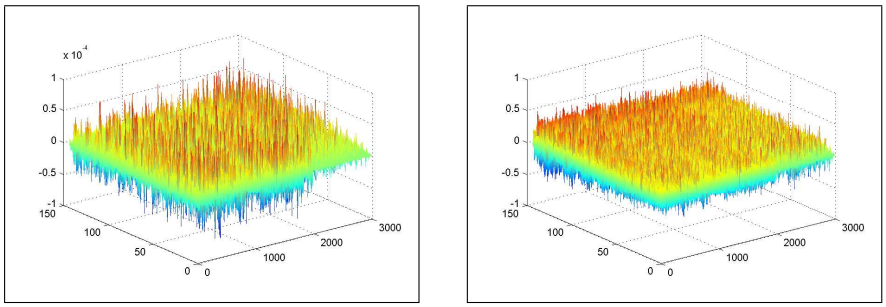

Figure 3: The lead field matrix T (left) and the preconditioned matrix T_c (right) when the elementary sources are randomly distributed. In the pictures the matrix entries (in arbitrary units) are displayed as a function of the row and column indexes.

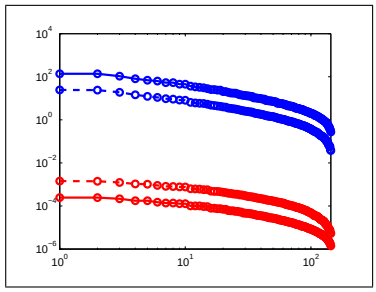

Figure 4: The amplitude (circles) of the singular values of T (bottom) and T_c (top) for the two different elementary source distributions (uniform grid: solid lines; random points: dashed lines). The $N = 144$ singular values are ordered with decreasing amplitude. The graphs are in logarithmic scale.

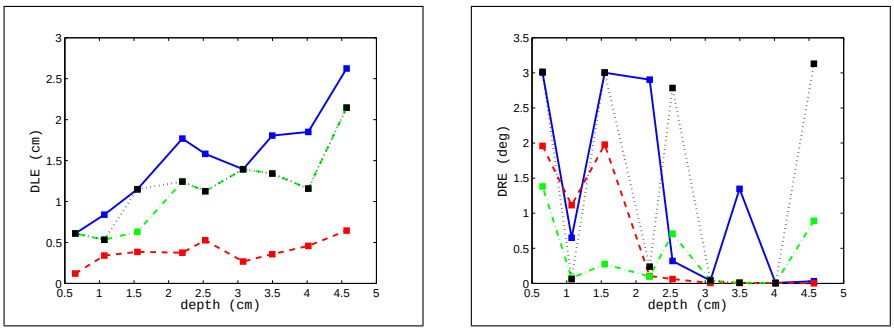

Figure 5: The localization error (left) and the direction error (right) as a function of the depth of the source when the elementary sources are located on a uniform grid and the lead field matrix is preconditioned by column balancing. The four graphs refer to ℓ_2-norm (solid line), hard thresholding (dash-dotted line), joint hard thresholding (dotted line), soft thresholding (dashed line).

dipole with moment $\vec{\mathbf{Q}}_0$ located in a point $\vec{\mathbf{P}}_0$ inside the conducting sphere. For each test we evaluate the distance localization error (DLE) and the direction error (DRE). The DLE is defined as the distance between $\vec{\mathbf{P}}_0$ and the position $\vec{\mathbf{P}}_{max}$ of the maximum of the reconstructed current intensity; the DRE is defined as the angle between $\vec{\mathbf{Q}}_0$ and the direction of the reconstructed current in $\vec{\mathbf{P}}_{max}$.

To solve the inverse problem with sparsity constraints, we use three different thresholding operators $\mathcal{S}_\sigma$, i.e. soft thresholding, hard thresholding and joint hard thresholding (see [8],[14] and references therein for details). For soft and hard thresholding, $\mathcal{S}_\sigma(J^{(n)}) = \big(s_\sigma(J^{(n)})^\ell_k\big)_{k=1,\dots,M,\ell=x,y,z}$ is a component-wise non linear operator acting on each entry $(J^{(n)})^\ell_k$ of the approximated vector $J^{(n)}$ as

$$s_\sigma(J^{(n)})^\ell_k = \begin{cases} (J^{(n)})^\ell_k - \operatorname{sign}(J^{(n)})^\ell_k \dfrac{\sigma}{2}\,, & \text{if} \quad |(J^{(n)})^\ell_k| \geq \dfrac{\sigma}{2}\,, \\ 0\,, & \text{otherwise}\,, \end{cases} \tag{8}$$

for soft thresholding, and as

$$s_\sigma(J^{(n)})^\ell_k = \begin{cases} (J^{(n)})^\ell_k\,, & \text{if} \quad |(J^{(n)})^\ell_k| \geq \dfrac{\sigma}{2}\,, \\ 0\,, & \text{otherwise}\,, \end{cases} \tag{9}$$

for hard thresholding. Instead, the joint hard thresholding $\mathcal{S}_\sigma$ acts on the three-dimensional approximated vector $\vec{J}^{(n)}_k = \big((J^{(n)})^x_k, (J^{(n)})^y_k, (J^{(n)})^z_k\big)$ as follows:

$$\mathcal{S}_\sigma(J^{(n)}) = s_\sigma\big(\vec{J}^{(n)}_k\big)_{k=1,\dots,N}\,, \tag{10}$$

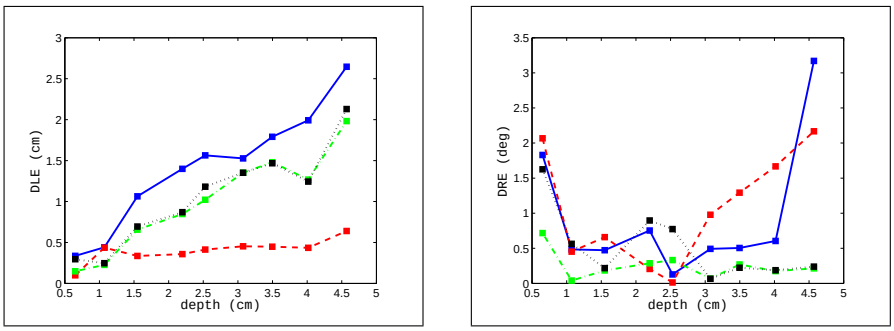

Figure 6: The localization error (left) and the direction error (right) as a function of the depth of the source when the elementary sources are randomly distributed and the lead field matrix is preconditioned by column balancing. The four graphs refer to ℓ_2-norm (solid line), hard thresholding (dash-dotted line), joint hard thresholding (dotted line), soft thresholding (dashed line).

where

$$s_\sigma\big(\vec{J}_k^{(n)}\big) = \begin{cases} \vec{J}_k^{(n)}\,, & \text{if} \quad \|\vec{J}_k^{(n)}\|_2 \geq \dfrac{\sigma}{2}\,, \\ \\ 0\,, & \text{otherwise}\,. \end{cases} \tag{11}$$

In the numerical tests, the threshold σ is adaptively chosen at each iteration as $\sigma = 0.8\, \max_{k,\ell} |(J^{(n)})_k^\ell|$ (soft and hard thresholding) or $\sigma = 0.8\, \max_k \|\vec{J}_k^{(n)}\|_2$ (joint hard thresholding).

The localization error and the direction error obtained by the different thresholding algorithms for the uniform and the random distributions of the elementary sources are shown in Fig. 5 and Fig. 6, respectively. For a comparison, the errors obtained by the classical ℓ_2-norm regularization are also displayed. To see how the preconditioning of the lead field matrix improves the localization results, in Fig. 7 and Fig. 8 the localization and direction errors obtained without preconditioning are shown.

3. Discussion

The numerical results in the previous section show that by using a random distribution of elementary sources we can considerably reduce the computational load while keeping a good accuracy in the localization. Actually, the matrix T (or T_c) has several tens of millions of entries in the case of elementary sources located on a thick uniform grid, while has no more than half a million of entries in the case of few elementary sources randomly distributed. This dramatic reduction of the memory storage allows us to run the latter algorithm on a laptop with a computing time of a few seconds. We notice also that to increase the computational points, that is to increase the accuracy of the localization, it

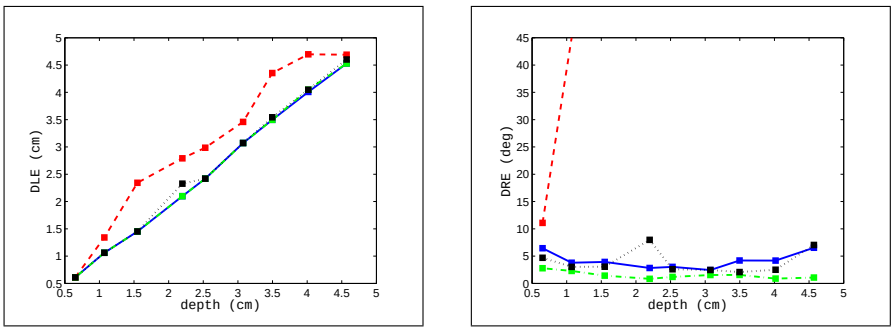

Figure 7: The localization error (left) and the direction error (right) as a function of the depth of the source when the elementary sources are located on a uniform grid and the lead field matrix is not preconditioned. The four graphs refer to ℓ_2-norm (solid line), hard thresholding (dash-dotted line), joint hard thresholding (dotted line), soft thresholding (dashed line).

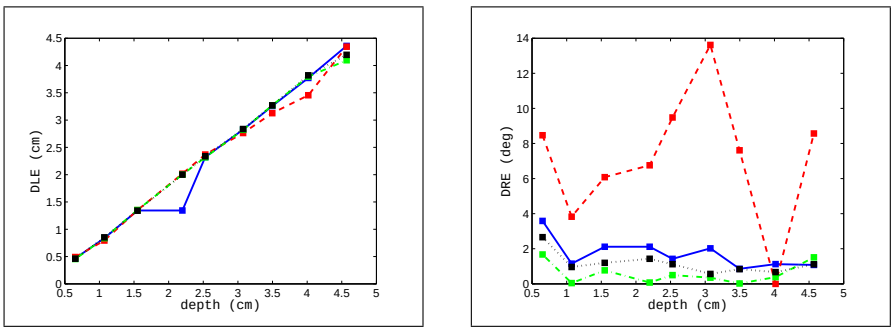

Figure 8: The localization error (left) and the direction error (right) as a function of the depth of the source when the elementary sources are randomly distributed and the lead field matrix is not preconditioned. The four graphs refer to ℓ_2-norm (solid line), hard thresholding (dash-dotted line), joint hard thresholding (dotted line), soft thresholding (dashed line).

is just sufficient to increase the number of trials without increasing the memory storage.

Moreover, the random distribution has better performance in the localization. In fact, the graphs in Figs. 5-6 show that in this case the localization error is slightly lower than the error in the uniform grid case while the direction error is highly reduced. In a next paper we intend to analyze how different random distributions of the elementary sources can affect the localization results.

As for the thresholding algorithms, the localization and direction errors in Fig. 6 show that all the algorithms have a good accuracy in localizing superficial sources while just soft thresholding is able to recover deep sources. On the other hand, hard thresholding is better in recovering the direction of the electric current. Joint hard thresholding instead does not reduce the direction error. We notice that, as expected, classical ℓ_2-regularization has a poor accuracy in localizing deep sources while, more surprisingly, has a rather good accuracy in recovering the direction.

Finally, the graphs in Figs. 7-8 show that the preconditioning of the lead field matrix by a column balancing is mandatory to reduce the localization error while the increase of the direction error obtained without preconditioning is very high just for soft thresholding.

4. Conclusion

We compared different inversion methods for the localization of a neural source inside the brain. Even if we used a simple model for both the head and the source, the numerical tests we performed give some hints on how to construct an efficient inversion algorithm. Summarizing, the soft thresholding algorithm with a random distribution of elementary sources and a column balancing seems the best algorithm to use for the localization of neural sources since it is fast, accurate and has a low computational load. If the recovering of the direction is also required, algorithms combining soft and hard thresholding should be used.

The next step will be the implementation of these inversion methods on real data, i.e. magnetoencephalographic data acquired by a MEG device on human subjects.

References

[1] M. Bertero & P. Boccacci, *Introduction to Inverse Problems in Imaging*, Institute of Physics Publishing, Bristol, 1998.

[2] E.J. Candes, M.B. Wakin & S.P. Boyd, Enhancing sparsity by reweighted ℓ_1 minimization, *Journal of Fourier Analysis and Applications* **16** (2008), 877–905.

[3] I. Daubechies, M. Defrise & C. DeMol, An iterative thresholding algorithm for linear inverse problems, *Comm. Pure Appl. Math.* **57** (1994), 1413–1457.

[4] C. Del Gratta, V. Pizzella, F. Tecchio & G.L. Romani, Magnetoencephalography - A noninvasive brain imaging method with 1ms time resolution, *Rep. Prog. Phys.* **64** (2001) 1759–1814.

[5] D.L. Donoho, Superresolution via Sparsity Constraints, *SIAM J. Math. Anal.* **23** (1992), 1309–1331.

[6] D.L. Donoho, De-noising by soft-thresholding, *IEEE Trans. Inform. Theory* **41** (1995), 613–627.

[7] H.W. Engl, M. Hank & A. Neubauer, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.

[8] M. Fornasier (ed.), *Theoretical Foundations and Numerical Methods for Sparse Recovery*, Radon Series on Computational and Applied Mathematics **9**, 2010.

[9] M. Fornasier & F. Pitolli, Adaptive iterative thresholding algorithms for magnetoencephalography (MEG), *J. Comput. Appl. Math.* **221** (2008), 386–395.

[10] M. Fornasier & H. Rauhut, Iterative thresholding algorithms, *Appl. Comput. Harmon. Anal.* **25** (2008), 187–208.

[11] M.S. Hamalainen, R. Hari, R.J. Ilmoniemi, J. Knuuttila & O.V. Lounasmaa, Magnetoencephalography - Theory, instrumentation and applications to non invasive studies of the working human brain, *Rev. Mod. Phys.* **65** (1993), 413–497.

[12] E. Herrholz & G. Teschke, Compressive sensing principles and iterative sparse recovery for inverse and ill-posed problems, *Inverse Problems* **26** (2010) 125012.

[13] J. Kaipio & E. Somersalo, *Statistical and Computational Inverse Problems*, Springer, New York, 2004.

[14] S. Mallat, *A Wavelet Tour on Signal Processing. The Sparse Way*, 3rd edn. Elsevier/Academic Press, 2009.

[15] J. Sarvas, Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem, *Phys. Med. Biol.* **32** (1987), 11–22.

[16] A.N. Tikhonov & V.A. Arsenin, *Solution of ill-posed problems*, W.H. Winston, Washington D.C., 1977.

[17] K. Wendel et al., EEG/MEG Source Imaging: Methods, Challenges, and Open Issues, *Comput. Intell. Neurosci.* **2009** (2009), 656092.