Adaptive Iterative Thresholding Algorithms for Magnetoenceophalography (MEG)

Massimo Fornasier* and Francesca Pitolli †

January 8, 2007

Abstract

We provide fast and accurate adaptive algorithms for the spatial resolution of current densities in MEG. We assume that vector components of the current densities possess a sparse expansion with respect to preassigned wavelets. Additionally, different components may also exhibit common sparsity patterns. We model MEG as an inverse problem with joint sparsity constraints, promoting coupling of non-vanishing components. We show how to compute solutions of the MEG linear inverse problem by iterative thresholded Landweber schemes. At each iteration an adaptive application of a compressed matrix associated to the Biot-Savart operator in wavelet coordinates is introduced. The resulting adaptive scheme is fast, robust, and significantly outperforms the classical Tikhonov regularization in resolving sparse current densities. Numerical examples are included.

AMS subject classification: 65J22, 65K10, 65T60, 52A41, 49M30, 68U10
Key Words: Magnetoenceophalography, Inverse problems, Iterative thresholding, Adaptive algorithms, Matrix compression, Wavelets

---

*Program in Applied and Computational Mathematics, Princeton University, Fine Hall, Washington Road, 08544-1000 Princeton NJ, USA. Email: mfornasi@math.princeton.edu. Massimo Fornasier acknowledges the financial support provided by the European Union's Human Potential Programme under contract MOIF-CT-2006-XXXX. He also thanks Ingrid Daubechies and Ignace Loris for the fruitful discussions on sparse recovery during the preparation of this work.

†Dipartimento di Metodi e Modelli Matematici per le Scienze Applicate, Università di Roma "La Sapienza", Via A. Scarpa 16, I-00161 Roma, Italy. Email: pitolli@dmmm.uniroma1.it

1 Introduction

The aim of magnetoencephalography (MEG) is the analysis of brain functionality through the measurements of the tiny magnetic fields generated by neuronal currents [2, 3]. As a matter of fact, since neuron cells functioning is mediated by electric currents, to understand brain functionality it is important to gain knowledge about the current distribution within the head. Therefore, looking from the physics side, the final goal of MEG is to accurately determine (in combination with other measurements, e.g., EEG) the current density flowing within the volume of the head in the working human brain. Among the available functional imaging techniques, such as Positron Emission Tomography (PET) or functional Magnetic Resonance Imaging (fMRI), MEG is a noninvasive technique with a high temporal resolution, typically in the order of milliseconds [9]. For this reason MEG technology appears particularly attractive. Nevertheless, the neuromagnetic field is only in the order of 10^{-13} Tesla in magnitude, that is more than one billion times smaller than the Earth's steady magnetic field and about one million times smaller than the magnetic fields generated by even distant moving metal objects (e.g., cars and elevators) or electric power lines. To successfully resolve the current density flowing within the brain it is mandatory to use low noise superconducting magnetometers as well as sophisticated signal processing. In addition, the magnetic field measured externally the head has a poor spatial resolution - at present MEG devices have at most a few hundred sensors - and can be generated by several possible current density configuration. Hence, the identification of a specific electric source configuration from the measured distribution of magnetic fields is an *ill-conditioned inverse problem*, and one of the challenging aspects of this technology. The main goal of this paper is to give an efficient and stable numerical scheme to compute the solution of the MEG inverse problem assuming that the currents flowing in the brain satisfy a sparsity constraint. In fact, according to Barlow [1], an important characteristic of sensory processing in the brain is the "redundancy reduction": the brain activity is represented as a sum of vectors of neurons, weighted by their activity, under the assumption that only a small number of neuronal areas are activated at the same time. For this reason, the sparsity assumption in the MEG model seems sufficiently realistic. Since our aim is to recover the current density, which is a vector-valued function, the joint sparsity constraint introduced in [8] to reconstruct multi-channel signals seems particularly attractive.

The paper is organized as follows. In Section 2 we describe the physical model for the MEG problem in a simplified geometry. This gives rise to an inverse problem involving the Biot-Savart operator whose properties are recalled in Section 3 [10]. In Section 4 we show how to transform the MEG problem into an inverse problem with joint sparsity constraints and in Section 5 we describe the algorithm for its solution and prove its convergence. In order to obtain an efficient algorithm, we need a compressed wavelet representation of the Biot-Savart operator. The compressibility properties of the Biot-Savart operator with respect to *wavelet bases* as well as the numerical stability of the whole scheme are the subject of Section 6.

2 The MEG Model

The neuromagnetic phenomena can be modelled by the Maxwell's equations for a polarizable and magnetizable macroscopic media (see [9] and references therein for details). In particular, bioelectric and biomagnetic fields can be described by the quasi-static Maxwell's equations which, for the electric field $\mathbf{E}$ and the magnetic induction $\mathbf{B}$, are given by

$$\begin{aligned} \operatorname{curl} \mathbf{E} &= 0, \\ \operatorname{curl} \mathbf{B} &= \mu_0 \, \mathbf{J}, \\ \operatorname{div} \mathbf{B} &= 0. \end{aligned}$$

Here μ_0 is the magnetic permeability of the vacuum and $\mathbf{J}$ is the current density in the medium. Due to the conservation of the charge, the current density satisfies

$$\operatorname{div} \mathbf{J} = 0. \tag{2.1}$$

The currents $\mathbf{J}$ flowing within the brain are generated, on the one hand, by active ion displacements due to the electrochemical cell activity - *impressed currents*, and, on the other hand, by passive displacement of free ions under the action of the electric field generated in the surrounding tissue by cell activity - *volume currents*. If we denote by $\mathbf{J}^i$ and $\mathbf{J}^v$ the density of the impressed and volume currents, respectively, the current density at any point in the conductor will be given by

$$\mathbf{J} = \mathbf{J}^i + \mathbf{J}^v = \mathbf{J}^i + \sigma \mathbf{E}, \tag{2.2}$$

where σ is a scalar isotropic conductivity function.

In a realistic model we should consider a conductivity function σ that may take different values inside the volume conductor, corresponding to, e.g., scalp, skull, cerebrospinal fluid, and brain. Let us denote by V_k, $k = 0, 1, \ldots, K$, the different regions of the conductor such that

$$V_0 \subset V_1 \subset \ldots \subset V_K \,.$$

We assume that impressed currents are confined to the innermost region and that each region V_k has constant conductivity σ_k. Thus, the contributions to $\mathbf{B}$ originates, on the one hand, from the impressed current in V_0 and, on the other hand, from the discontinuities of σ at the each interface $S_k = \partial V_k$ between V_k and V_{k+1}. Let us denote by $\mathbf{n}_k$ the unit normal vector to S_k. Moreover, we assume that $\sigma_{K+1} \equiv 0$ outside the volume conductor. Under these assumptions the quasi-stationary Maxwell's equation give [9]

$$\mathbf{B}(x) = \mathbf{B}_\infty(x) - \frac{\mu_0}{4\pi}\sum_{k=0}^{K}(\sigma_{k+1} - \sigma_k)\int_{S_k} \frac{\Phi(y)\,\mathbf{n}_k(y) \times (x-y)}{|x-y|^3}\, d\Sigma_k(y)\,, \tag{2.3}$$

where $\mathbf{B}_\infty$ is the magnetic induction in an infinite homogeneous medium, i.e.

$$\mathbf{B}_\infty(x) = \frac{\mu_0}{4\pi}\int_{V_0} \mathbf{J}^i(y) \times \frac{x-y}{|x-y|^3}\, d(y)\,, \tag{2.4}$$

Φ is the electric potential. The electric potential Φ is related to the electric field by $\mathbf{E} = -\nabla\Phi$, and cannot be derived explicitly. Nevertheless, it solves a fixed point integral equation [9, Formula (3.47)] which depends again on $\mathbf{J}^i$.

The MEG technology does *not* furnish a full measurement of the external magnetic field, but only of its normal component on an external surface to the head. In order to obtain a model easily tractable, but still sufficiently realistic, we assume that the volumes V_k are confined in spheres S_k of increasing radius and centered at the origin. If $\mathbf{e}_r(x)$ is the radially oriented unit vector at point x, then $B_r(x) := \mathbf{B}(x)\cdot\mathbf{e}_r(x)$ is given by

$$\begin{aligned} B_r(x) \;&:=\; B_r(x, \mathbf{J}^i) = \mathbf{B}_\infty(x)\cdot\mathbf{e}_r(x) - \\ &\quad \underbrace{- \frac{\mu_0}{4\pi}\sum_{k=0}^{K}(\sigma_{k+1} - \sigma_k)\int_{S_k} \frac{\Phi(y)\,\mathbf{n}_k(y) \times (x-y)\cdot\mathbf{e}_r(x)}{|x-y|^3}\, d\Sigma_k(y)}_{=0} \\ &= \frac{\mu_0}{4\pi}\int_{V_0} \frac{\mathbf{J}^i(y) \times (x-y)\cdot\mathbf{e}_r(x)}{|x-y|^3}\, d(y). \end{aligned} \tag{2.5}$$

This formula relates the normal component of the magnetic field directly to the impressed current density $\mathbf{J}^i$. Thanks to well-known results in potential theory [10, Theorem 2.21], the normal component of $\mathbf{B}$ on an external surface uniquely determines the magnetic field out of the head as soon as the current flux on the scalp is known. In our case the current flux is null. Thus, the measurements of B_r on an external surface of the head are a sufficient information to determine the whole magnetic field.

3 The MEG Inverse Problem

The MEG inverse problem consists in deriving the impressed current density $\mathbf{J}^i$ from the measured components of the magnetic field. From (2.5) we see that the problem involves the so-called *Biot-Savart operator*

$$\mathcal{B}(x) := \mathcal{B}(x, \mathbf{J}) = \frac{\mu_0}{4\pi} \int_{V_0} \frac{\mathbf{J}(y) \times (x-y)}{|x-y|^3} d(y) \tag{3.1}$$

which is known to have non-trivial kernel [10]. This means that there exist *infinitely many* current densities $\mathbf{J}$ which generate the *same* B_r. In addition, due to the fact that $\mathcal{B}$ is a compact operator, its generalized inverse is unbounded.

A well known technique to stabilize the inversion is by Tikhonov regularization, i.e., by minimizing the functional

$$\mathcal{K}_\alpha(\mathbf{J}) := \|\mathcal{B}(\cdot, \mathbf{J}) - \mathcal{B}_0\|_2^2 + \alpha \|\mathbf{J}\|_2^2 \,,$$

where $\mathcal{B}_0$ is the observed magnetic field. The minimization has the unique solution

$$\mathbf{J}_\alpha := (\alpha \mathbf{I} + \mathcal{B}^* \mathcal{B})^{-1} \mathcal{B}^* \mathcal{B}_0 \,,$$

with the property that

$$\lim_{\alpha \to 0} \mathbf{J}_\alpha \in \ker(\mathcal{B})^\perp,$$

and $\|\mathcal{B}(\cdot, \mathbf{J}_\alpha) - \mathcal{B}_0\|_2^2$ is a non-decreasing function of α. Since we would like to reconstruct the impressed current $\mathbf{J}^i$ by minimizing the discrepancy $\|\mathcal{B}(\cdot, \mathbf{J}_\alpha) - \mathcal{B}_0\|_2^2$, we should question whether $\mathbf{J}^i \in \ker(\mathcal{B})^\perp$ is a meaningful property for our problem. While it is know [10, Section 2.3.2] that the volume current $\mathbf{J}^v$ in V_0 (where $\sigma = \sigma_0$ is assumed constant) belongs to $\ker(\mathcal{B})^\perp$, nothing can be said about $\mathbf{J}^i$. Therefore, in this case, the use of a simple

Tikhonov regularization will risk to reconstruct incorrect current densities, while the sparsity assumption is more realistic.

Then we may assume that $\mathbf{J}^i$ is a function with compact support which is the union of few local disjoint compact volumes. By expanding the components of $\mathbf{J}^i$ with respect to a suitable basis of compactly supported functions $\Psi = \{\psi_\lambda\}_{\lambda\in\Lambda}$ we may obtain the series

$$\mathbf{J}^i = \left(\sum_{\lambda\in\Lambda} j_\lambda^\ell \psi_\lambda\right)_{\ell=1}^3 ,$$

where only *few* scalar coefficients $\mathbf{j} = (j_\lambda^\ell)_{\lambda\in\Lambda}^{\ell=1,2,3}$ are non-vanishing. Here the index ℓ denotes the label for the vector components and λ denotes the basis index. We will say that $\mathbf{J}^i$ is *sparsely represented* by the basis elements $\Psi = \{\psi_\lambda\}_{\lambda\in\Lambda}$. The sparsity with respect to a basis can be indeed measured directly by the ℓ_p norm for $0 < p \le 1$: as smaller $\|(\mathbf{j}_\lambda)_{\lambda\in\Lambda}\|_{\ell_p} := \left(\sum_\lambda \|\mathbf{j}_\lambda\|_{\mathbb{R}^3}^p\right)^{1/p}$ is, as fewer non-vanishing coefficients we have. With this characterization we can solve the problem of finding $\mathbf{J}^i$ by minimizing the discrepancy with respect to the given measured magnetic field (its normal component on an external surface, to be precise) *and* the sparsity of $\mathbf{J}^i$ with respect of a suitable basis of compactly supported functions.

4 The MEG Problem With Joint Sparsity Constraints

Before starting our discussion let us briefly introduce some of the spaces we will use in the following.

Notation. For some countable index set Λ we denote by $\ell_p = \ell_p(\Lambda)$, $1 \le p \le \infty$, the space of real sequences $u = (u_\lambda)_{\lambda\in\Lambda}$ with norm

$$\|u\|_p = \|u|\ell_p\| := \left(\sum_{\lambda\in\Lambda} |u_\lambda|^p\right)^{1/p}, \quad 1 \le p < \infty ,$$

and $\|u\|_\infty := \sup_{\lambda\in\Lambda} |u_\lambda|$ as usual. If $v = (v_\lambda)_{\lambda\in\Lambda}$ is a sequence of positive weights, then we define the weighted spaces $\ell_{p,v} = \ell_{p,v}(\Lambda) = \{u, (u_\lambda v_\lambda)_{\lambda\in\Lambda} \in$

$\ell_p(\Lambda)\}$ with norm

$$\|u\|_{p,v} = \|u|\ell_{p,v}\| = \|(u_\lambda v_\lambda)_{\lambda\in\Lambda}\|_p = \left(\sum_{\lambda\in\Lambda} v_\lambda^p |u_\lambda|^p)\right)^{1/p}$$

(with obvious modification for $p = \infty$). If the entries u_λ are actually vectors in a Banach space X with norm $\|\cdot\|_X$ then we denote

$$\ell_{p,v}(\Lambda, X) := \{(u_\lambda)_{\lambda\in\Lambda}, u_\lambda \in X, (\|u_\lambda\|_X)_{\lambda\in\Lambda} \in \ell_{p,v}(\Lambda)\}$$

with norm $\|u|\ell_{p,v}(\Lambda, X)\| = \|(\|u_\lambda\|_X)_{\lambda\in\Lambda}|\ell_{p,v}(\Lambda)\|$. Usually X will be $\mathbb{R}^N$ endowed with the Euclidean norm, or the N-dimensional space ℓ_q^N, i.e., $\mathbb{R}^N$ endowed with the ℓ_q-norm. By $\mathbb{R}_+$ we denote the non-negative real numbers. With $\mathcal{L}(\mathcal{H})$ we denote the space of bounded linear operators on a Hilbert space $\mathcal{H}$.

Now, let us describe the MEG inverse problem with joint sparsity constraints in details. The measured data $g = (g_1, \ldots, g_N)^T \in \mathbb{R}^N$ provided by MEG are the normal components B_r of the magnetic field at a distribution of points on a surface $\partial\Omega$ external to the scalp. Let $V_0 \subset V_K \subset \Omega$ and denote with $x_1, ..., x_N \in \partial\Omega$ the sampling points. Let us assume that $\delta := \mathrm{dist}(V_0, \partial\Omega) > 0$.

For a given current density $\mathbf{J} = (J_1, J_2, J_3)$ we define the operator $A : L_2(V_0; \mathbb{R}^3) \to \mathbb{R}^N$ given by

$$A\,\mathbf{J} = (B_r(x_1, \mathbf{J}), \ldots, B_r(x_N, \mathbf{J}))^T \in \mathbb{R}^N . \tag{4.1}$$

From (2.5) it follows

$$\begin{aligned} B_r(x_k, \mathbf{J}) &:= \frac{\mu_0}{4\pi}\int_{V_0}\left(\frac{\mathbf{e}_r(x_k)\times(x_k - y)}{|x_k - y|^3}\right)\cdot\mathbf{J}(y)\,d(y) \\ &= \sum_{\ell=1}^{3}\frac{\mu_0}{4\pi}\int_{V_0}\left(\frac{\mathbf{e}_r(x_k)\times(x_k - y)}{|x_k - y|^3}\right)_\ell J_\ell(y)\,d(y) \\ &= \sum_{\ell=1}^{3} A_{\ell,k}\,J_\ell, \quad k = 1, \ldots, N , \end{aligned} \tag{4.2}$$

where $A_{\ell,k} : L_2(V_0; \mathbb{R}) \to \mathbb{R}$ is the operator

$$A_{\ell,k}\,J_\ell := \frac{\mu_0}{4\pi}\int_{V_0}\left(\frac{\mathbf{e}_r(x_k)\times(x_k - y)}{|x_k - y|^3}\right)_\ell J_\ell(y)\,d(y) . \tag{4.3}$$

Here we have used the relation $\mathbf{v} \times \mathbf{w} \cdot \mathbf{z} = \mathbf{z} \times \mathbf{w} \cdot \mathbf{v}$, for $\mathbf{v}, \mathbf{w}, \mathbf{z} \in \mathbb{R}^3$. Due to the fact that $\delta > 0$, one can easily check that the operator A is bounded from $L_2(V_0; \mathbb{R}^3)$ to $\mathbb{R}^N$.

Now, we assume that the space $L_2(V_0; \mathbb{R})$ disposes of a stable *wavelet basis* $(\psi_\lambda)_{\lambda \in \Lambda}$ [6, 15]. Stability means that there exist constants $C_1, C_2 > 0$ such that

$$C_1 \|f\|^2_{L_2(V_0;\mathbb{R})} \leq \sum_{\lambda \in \Lambda} |\langle f, \psi_\lambda \rangle|^2 \leq C_2 \|f\|^2_{L_2(V_0;\mathbb{R})} \tag{4.4}$$

for all $f \in L_2(V_0; \mathbb{R})$.
We do not enter into the details of these bases; we will just recall some of their useful properties:

- The index $\lambda = (|\lambda|, k, e)$ encodes several different properties, respectively, the scale $|\lambda|$, the spatial location $k \in \mathbb{R}^3$, and the wavelet label e;
- $\Omega_\lambda := \operatorname{supp}(\psi_\lambda)$, $|\Omega| \sim 2^{-|\lambda|}$;
- $\int_{V_0} y^\alpha \, \psi_\lambda(y) \, dy = 0$, $\alpha = 0, .., d^* \in \mathbb{N}$;
- $\psi_\lambda \in C^\beta(V_0)$;
- $|\psi_\lambda| \leq C \, 2^{3/2|\lambda|}$.

Refer to [15] for more details on the construction of multivariate wavelets. Thus, every $\mathbf{J} \in L_2(V_0; \mathbb{R}^3)$ can be decomposed in a stable way, i.e.

$$\mathbf{J} = \left(\sum_{\lambda \in \Lambda} j^\ell_\lambda \, \psi_\lambda \right)^3_{\ell=1} . \tag{4.5}$$

Remark. By (2.1) and by the fact that $\mathbf{J}^v \in \ker(\mathcal{B})^\perp$ (see [10, Section 2.2.2])

we have $\operatorname{div} \mathbf{J}^i = 0$. Therefore, we may restrict our problem to the space

$$V(\operatorname{div}, V_0) := \{ \mathbf{J} \in L_2(V_0; \mathbb{R}^3) : \operatorname{div} \mathbf{J} = 0 \} .$$

Also the space $V(\operatorname{div}, V_0)$ disposes of a Riesz basis composed by *divergence free wavelets* $(\psi^{\text{div}}_\lambda := (\psi^{\text{div}}_{\lambda,1}, \psi^{\text{div}}_{\lambda,2}, \psi^{\text{div}}_{\lambda,3}))_{\lambda \in \Lambda}$ [14, 15]. The additional constraint $\operatorname{div} \psi^{\text{div}}_\lambda = 0$ enforces in this case a direct coupling of the vector values. Nevertheless, by the definition appearing in [14], these wavelets have at most *two* non-vanishing components $\psi^{\text{div}}_{\lambda,\ell}$ and again it is necessary to consider a joint sparsity constraint to couple *all* the components.

Let us denote by $F : \ell_2(\Lambda; \mathbb{R}) \to L_2(V_0; \mathbb{R})$ the *synthesis operator*

$$F\, u := \sum_{\lambda \in \Lambda} u_\lambda \, \psi_\lambda \,, \tag{4.6}$$

where $u = (u_\lambda)_{\lambda \in \Lambda} \in \ell_2(\Lambda)$. Due to the stability inequalities (4.4) the synthesis operator F is bounded.

We define the operator $T_{\ell,k} : \ell_2(\Lambda; \mathbb{R}) \to \mathbb{R}$ as $T_{\ell,k} := A_{\ell,k} F$, i.e.

$$T_{\ell,k}(j^\ell_\lambda)_{\lambda \in \Lambda} = \sum_{\lambda \in \Lambda} j^\ell_\lambda \left(\frac{\mu_0}{4\pi} \int_{V_0} \left(\frac{\mathbf{e}_r(x_k) \times (x_k - y)}{|x_k - y|^3} \right)_\ell \psi_\lambda(y)\, d(y) \right) \,, \tag{4.7}$$

and

$$T\mathbf{j} := \left(\sum_{\ell=1}^{3} T_{\ell,k}(j^\ell_\lambda)_{\lambda \in \Lambda} \right)_{k=1}^{N} \,, \tag{4.8}$$

where $\mathbf{j} := (j^\ell_\lambda)^{\ell=1,2,3}_{\lambda \in \Lambda}$ are the wavelet coefficients associated to the components of $\mathbf{J}$.

Given a set of data g we would like to compute $\mathbf{J} \in L_2(V_0; \mathbb{R}^3)$ such that

$$\sum_{\ell=1}^{3} T_{\ell,k}(j^\ell_\lambda)_{\lambda \in \Lambda} \sim g_k, \qquad k = 1, \dots, N. \tag{4.9}$$

As mentioned, we can assume that $\mathbf{J}$ is a function with compact support which is the union of few local disjoint compact volumes. Since also ψ_λ are compactly supported, only few non-vanishing coefficients $\mathbf{j}_\lambda = (j^\ell_\lambda)^3_{\ell=1}$ should be relevant to reconstruct $\mathbf{J} = F\mathbf{j}$. Additionally, different components of $\mathbf{J}$ should be assumed mutually correlated. A way to incorporate such correlation is the assumption of *joint sparsity* [13]. By this we mean that the pattern of non-zero coefficients representing $\mathbf{J}$ is (approximately) the same for all the channels, i.e.

$$J_\ell \approx \sum_{\lambda \in \Lambda_0} j^\ell_\lambda \, \psi_\lambda \,, \qquad \ell = 1, 2, 3 \,, \tag{4.10}$$

where $\Lambda_0 \subset \Lambda$ is the same *finite* set for all J_ℓ's.

Thus, following [8], the solution of our problem can be modeled as the minimizer of the functional

$$\begin{aligned}\mathcal{J}(\mathbf{j},v) &= \mathcal{J}^{(q)}_{\theta,\rho,\omega}(\mathbf{j},v) := \|T\mathbf{j}-g\|^2_{\mathbb{R}^N}+ \\ &+\left(\sum_{\lambda\in\Lambda} v_\lambda \|\mathbf{j}_\lambda\|_q + \sum_{\lambda\in\Lambda}\omega_\lambda \|\mathbf{j}_\lambda\|_2^2 + \sum_{\lambda\in\Lambda}\theta_\lambda(\rho_\lambda - v_\lambda)^2\right),\end{aligned} \tag{4.11}$$

restricted to $v_\lambda \geq 0$. Here, $(\theta_\lambda)_{\lambda\in\Lambda}$, $(\rho_\lambda)_{\lambda\in\Lambda}$, and $(\omega_\lambda)_{\lambda\in\Lambda}$ are some suitable positive parameter sequences. The task is to minimize $\mathcal{J}(\mathbf{j},v)$ jointly with respect to both the variables u, v. The first belongs to the space of signals (current densities) to be reconstructed, the second belongs to the space of sparsity indicator weights. Here, $q > 1$ and in particular, $q = 2$ or $q = \infty$, represent the interesting cases. In fact if q is large and some $|j^\ell_\lambda|$ is large then the channel entries $|j^{\ell'}_\lambda|$ are also allowed to be large for $\ell' \neq \ell$, without increasing significantly the norm $\|\mathbf{j}_\lambda|\ell^N_q\|$. The minimization of the above norm promotes that all entries of the 'interchannel' vector $\mathbf{j}_\lambda$ may become significant, once at least one of the components $|j^\ell_\lambda|$ is large. Moreover, for the minimizer $(\mathbf{j}, v)$ we will have $v_\lambda = 0$ (or close to 0) if $\|\mathbf{j}_\lambda\|_q$ is large so that $v_\lambda\|\mathbf{j}_\lambda\|_q$ gets small. On the other hand, if $\|\mathbf{j}_\lambda\|_q$ is small then the term $\theta_\lambda(\rho_\lambda - v_\lambda)^2$ dominates and forces v_λ to be close to ρ_λ. Thus, v_λ serves indeed as an indicator of large values of $\|\mathbf{j}_\lambda\|_q$. It has the effect, that if v_λ is chosen small due to one large j^ℓ_λ then also the other coefficients $j^{\ell'}_\lambda$, $\ell' \neq \ell$ can be chosen large without making the functional considerably bigger. The quadratic term $\sum_{\lambda\in\Lambda}\omega_\lambda\|\mathbf{j}_\lambda\|_2^2$ serves in order to make the overall functional convex, depending on the suitable choice of ω_λ.

Observe that at the minimizer we will always have $0 \leq v_\lambda \leq \rho_\lambda$. Therefore, we can assume the domain of $\mathcal{J}$ to be $\ell_2(\Lambda, \mathbb{R}^N) \times \ell_{\infty,\rho^{-1}}(\Lambda)_+$, where $\ell_{\infty,\rho^{-1}}(\Lambda)_+$ denotes the (convex) cone of all non-negative sequences $(v_\lambda)_{\lambda\in\Lambda} \in \ell_{\infty,\rho^{-1}}(\Lambda)$.

5 An Efficient Numerical Minimization

An algorithm for the minimization of $\mathcal{J}(\mathbf{j}, v)$ is described as follows. First, let us recall the convexity conditions for the functional $\mathcal{J}$.

Proposition 5.1. *[8] Let $q \in \{1, 2, \infty\}$. If $\omega_\lambda\theta_\lambda \geq \frac{\kappa}{4}$ for all $\lambda \in \Lambda$, where $\kappa = N$ for $q = 1$, and $\kappa = 1$ for $q \in \{2, \infty\}$, then $\mathcal{J}$ is convex. In case of a strict inequality $\omega_\lambda\theta_\lambda > \frac{\kappa}{4}$, then $\mathcal{J}$ is strictly convex.*

For some initial choice $v^{(0)}$, for example $v^{(0)} = (\rho_\lambda)_{\lambda\in\Lambda}$, the algorithm for the computation of the minimizer $(\mathbf{j}^*, v^*)$ of the functional $\mathcal{J}(\mathbf{j}, v) = J^{(q)}_{\theta,\rho,\omega}(\mathbf{j}, v)$ defined in (4.11) is given by

$$\begin{aligned} \mathbf{j}^{(n)} &:= \arg\min\nolimits_{\mathbf{j}\in\ell_2(\Lambda,\mathbb{R}^N)} \mathcal{J}(\mathbf{j}, v^{(n-1)}) \,, \\ v^{(n)} &:= \arg\min\nolimits_{v\in\ell_{\infty,\rho^{-1}}(\Lambda)_+} \mathcal{J}(\mathbf{j}^{(n)}, v). \end{aligned} \tag{5.1}$$

The minimization of $\mathcal{J}(\mathbf{j}, v^{(n-1)})$ with respect to $\mathbf{j}$ can be done by means of a fast thresholded Landweber algorithm, similar to that presented in [7]. The minimizer $v^{(n)}$ of $\mathcal{J}(\mathbf{j}^{(n)}, v)$ for fixed $\mathbf{j}^{(n)}$ can be computed explicitly. Indeed, it follows from elementary calculus that

$$v^{(n)}_\lambda = \begin{cases} \rho_\lambda - \frac{1}{2\theta_\lambda}\|\mathbf{j}^{(n)}{}_\lambda\|_q \,, & \text{if } \|\mathbf{j}^{(n)}{}_\lambda\|_q < 2\theta_\lambda\rho_\lambda \,, \\ 0 \,, & \text{otherwise.} \end{cases} \tag{5.2}$$

From [8] we have the following result about the convergence of the above algorithm.

Theorem 5.2. *[8] Let $1 \le q \le \infty$ and assume that $\mathcal{J}$ is strictly convex (see Proposition 5.1). Moreover, we assume that $\ell_{2,\omega^{1/2}}(\Lambda, \mathbb{R}^N)$ is embedded into $\ell_2(\Lambda, \mathbb{R}^N)$, i.e., $\omega_\lambda \ge \gamma > 0$ for all $\lambda \in \Lambda$. Then, the sequence $(\mathbf{j}^{(n)}, v^{(n)})_{n\in\mathbb{N}}$ converges to the unique minimizer $(\mathbf{j}^*, v^*) \in \ell_2(\Lambda, \mathbb{R}^N) \times \ell_{\infty,\rho^{-1}}(\Lambda)_+$ of $\mathcal{J}$. The convergence of $\mathbf{j}^{(n)}$ is weak in $\ell_2(\Lambda, \mathbb{R}^N)$ and that of $v^{(n)}$ holds componentwise.*
For the most interesting cases $q \in \{1, 2, \infty\}$, if in addition $\theta_\lambda\omega_\lambda \ge \sigma > \phi_q/4$ for all $\lambda \in \Lambda$, where $\phi_1 = N$, $\phi_2 = 1$, $\phi_\infty = \sqrt{N}$, then the convergence of $\mathbf{j}^{(n)}$ to $\mathbf{j}^*$ is also strong in $\ell_2(\Lambda, \mathbb{R}^N)$ and $v^{(n)} - v^*$ converges to 0 strongly in $\ell_{2,\theta}(\Lambda)$.

To implement the algorithm above we need a method to approximate the solution of the minimization with respect to the first variable $\mathbf{j}$. To this end we introduce the *thresholding operators.*

Definition 5.3. *Let $1 \le q \le \infty$ and $u \in L_2(V_0; \mathbb{R})$. We define the thresholding operator by*

$$(U^{(q)}_{v,\omega}(u))_\lambda \;:=\; (1+\omega_\lambda)^{-1} S^{(q)}_{v_\lambda}(u_\lambda),$$

where

$$S^{(q)}_v(x) \;=\; \operatorname{argmin}_{z\in\mathbb{R}^N} \|z - x\|_2^2 + v\|z\|_q, \quad x \in \mathbb{R}^N. \tag{5.3}$$

The operator $S_v^{(q)}$ is given by

$$S_v^{(q)}(x) \;=\; x - P_{v/2}^{q'}(x), \tag{5.4}$$

where $P_{v/2}^{q'}$ denotes the orthogonal projection onto the norm ball of radius $v/2$ with respect to the dual norm of $\|\cdot\|_q$, i.e., the $\|\cdot\|_{q'}$-norm with q' denoting the dual index, $1/q + 1/q' = 1$.

In [8] the explicit form of $S_v^{(q)}$ is given for $q = 1, 2, \infty$ and $v \geq 0$. For $q = 1$ we have $S_v^{(1)}(x) = (s_v^{(1)}(x_\ell))_{\ell=1}^N$ where for $y \in \mathbb{R}$

$$s_v^{(1)}(y) \;=\; \begin{cases} 0 & \text{if } |y| \leq \frac{v}{2}, \\ \operatorname{sign}(y)(|y| - \frac{v}{2}) & \text{otherwise.} \end{cases} \tag{5.5}$$

For $q = 2$ it holds

$$S_v^{(2)}(x) := \begin{cases} 0 & \text{if } \|x\|_2 \leq \frac{v}{2}, \\ \frac{(\|x\|_2 - v/2)}{\|x\|_2} x & \text{otherwise.} \end{cases} \tag{5.6}$$

Finally, for $q = \infty$ we order the entries of x by magnitude such that $|x_{i_1}| \geq |x_{i_2}| \geq \ldots \geq |x_{i_N}|$ and we let

i. $S_v^{(\infty)}(x) = 0 \quad \text{if} \quad \|x\|_1 < v/2;$

ii. $(S_v^{(\infty)}(x))_{i_j} \;=\; \frac{\operatorname{sign}(x_{i_j})}{n} \left(\sum_{k=1}^n |x_{i_k}| - \frac{v}{2} \right), \qquad j = 1, \ldots, n,$

$(S_v^{(\infty)}(x))_{i_j} \;=\; x_{i_j}, \quad j = n+1, \ldots, N,$

if $\|x\|_1 > v/2$, where $n \in \{1, \ldots, N\}$ is the largest index satisfying

$$|x_{i_n}| \;\geq\; \frac{1}{n-1} \left(\sum_{k=1}^{n-1} |x_{i_k}| - \frac{v}{2} \right). \tag{5.7}$$

With the thresholding operator at hand we can formulate our numerical scheme to compute

$$\mathbf{j}^{(n)} := \arg \min_{\mathbf{j} \in \ell_2(\Lambda, \mathbb{R}^N)} \mathcal{J}(\mathbf{j}, v^{(n-1)}),$$

given by

$$\mathbf{j}^{(n,m+1)} := U_{v^{(n)},\omega}^{(q)} \left(\mathbf{j}^{(n,m)} + T^* g - T^* T \mathbf{j}^{(n,m)} \right). \tag{5.8}$$

In [8, Theorem 4.9] it is shown that for a fixed $v^{(n)}$ the previous iteration is convergent in $\ell_2(\Lambda, \mathbb{R}^N)$ to the minimizer $\mathbf{j}^{(n)}$. Nevertheless, observe that each iteration involves an application of T^*T and of the thresholding operator $U^{(q)}_{v,\omega}$. The latter can clearly be applied fast on finite sequences. Since in general T^*T is represented as an infinite matrix,its evaluation might not be exactly numerically implementable. In this paper we deal in fact with the case $\#\Lambda = \infty$ and the treatment of sparse (approximate) evaluations of infinite matrices in order to realize fast and convergent schemes also in this situation (compare [4, 5, 12]). This issue will be addressed in the following section. For now, we assume that we have the following procedure **APPLY** at our disposal:

- **APPLY**$[\varepsilon, \mathcal{N}, \mathbf{j}] \to \mathbf{j}_\varepsilon$: determines for $\mathcal{N} \in \mathcal{L}(\ell_2(\Lambda, \mathbb{R}^N))$ and for a finitely supported $\mathbf{j} \in \ell_2(\Lambda, \mathbb{R}^N)$ a finitely supported $\mathbf{j}_\varepsilon$ such that

$$\|\mathcal{N}\mathbf{j} - \mathbf{j}_\varepsilon\|_{\ell_2(\Lambda,\mathbb{R}^N)} \leq \varepsilon; \tag{5.9}$$

Moreover, in the following we can assume without loss of generality that T^*g is also a finitely supported vector. We want to substitute the exact iteration (5.8) with the following inexact one

$$\mathbf{j}^{(n,m+1)} := U^{(q)}_{v^{(n)},\omega}\left(\mathbf{j}^{(n,m)} + T^*g - \mathbf{APPLY}[\frac{\varepsilon^{(n,m)}}{K}, T^*T, \mathbf{j}^{(n,m)}]\right). \tag{5.10}$$

For a suitable choice of the approximations $\varepsilon^{(n,m)} > 0$ and of the constant $K > 0$, we will derive a fully implementable and convergent scheme.

Proposition 5.4. *Assume that $\omega_\lambda \geq \gamma > 0$ for all $\lambda \in \Lambda$ and $\|T\| < 1$. Let us denote $\alpha := (1+\gamma)^{-1}(\frac{1}{K} + \|I - T^*T\|) < 1$ for $K > 0$ large enough, and*

$$\mathbf{j}^{(n,\infty)} := U^{(q)}_{v^{(n)},\omega}\left(\mathbf{j}^{(n,\infty)} + T^*(g - T\mathbf{j}^{(n,\infty)})\right).$$

If $\|\mathbf{j}^{(n,\infty)} - \mathbf{j}^{(n,m)}\|_2 \leq \varepsilon^{(n,m)}$ and $\varepsilon^{(n,m+1)} := \alpha\varepsilon^{(n,m)} < \varepsilon^{(n,m)}$, then

$$\mathbf{j}^{(n,m+1)} := U^{(q)}_{v^{(n)},\omega}\left(\mathbf{j}^{(n,m)} + T^*g - \mathbf{APPLY}[\frac{\varepsilon^{(n,m)}}{K}, T^*T, \mathbf{j}^{(n,m)}]\right),$$

is such that

$$\|\mathbf{j}^{(n,\infty)} - \mathbf{j}^{(n,m+1)}\|_2 \leq \varepsilon^{(n,m+1)}. \tag{5.11}$$

Proof. By non-expansiveness of $S_v^{(q)}$ (see [8, Lemma 4.4] and its proof) we obtain

$$
\begin{aligned}
&\|\mathbf{j}^{(n,\infty)} - \mathbf{j}^{(n,m+1)}\|_2 \\
\leq & \|U^{(q)}_{v^{(n)},\omega}\left(\mathbf{j}^{(n,\infty)} + T^*(g - T\mathbf{j}^{(n,\infty)})\right) - U^{(q)}_{v^{(n)},\omega}\left(\mathbf{j}^{(n,m)} + T^*(g - T\mathbf{j}^{(n,m)})\right)\|_2 \\
+ & \|U^{(q)}_{v^{(n)},\omega}\left(\mathbf{j}^{(n,m)} + T^*(g - T\mathbf{j}^{(n,m)})\right) - U^{(q)}_{v^{(n)},\omega}\left(\mathbf{j}^{(n,m)} + T^*g - \mathbf{APPLY}[\frac{\varepsilon^{(n,m)}}{K}, T^*T, \mathbf{j}^{(n,m)}]\right)\|_2 \\
\leq & \left(\sum_{\lambda\in\Lambda}(1+\omega_\lambda)^{-2}\|S^{(q)}_{v_\lambda}((\mathbf{j}^{(n,\infty)} + T^*(g - T\mathbf{j}^{(n,\infty)})_\lambda) - S^{(q)}_{v_\lambda}((\mathbf{j}^{(n,m)} + T^*(g - T\mathbf{j}^{(n,m)})_\lambda)\|_2^2\right)^{1/2} \\
& + \sup_{\lambda\in\Lambda}(1+\omega_\lambda)^{-1}\frac{\varepsilon^{(n,m)}}{K} \\
\leq & \sup_{\lambda\in\Lambda}(1+\omega_\lambda)^{-1}\left(\|(I - T^*T)(\mathbf{j}^{(n,\infty)} - \mathbf{j}^{(n,m)})\|_2 + \frac{\varepsilon^{(n,m)}}{K}\right) \\
\leq & (1+\gamma)^{-1}\left(\frac{\varepsilon^{(n,m)}}{K} + \|I - T^*T\|\,\|\mathbf{j}^{(n,\infty)} - \mathbf{j}^{(n,m)}\|_2\right) \\
\leq & \varepsilon^{(n,m+1)}.
\end{aligned}
$$

This establishes the claim. □

Let $q \in \{1, 2, \infty\}$. Assume that $\theta_\lambda\omega_\lambda \geq \sigma > \phi_q/4$ for all $\lambda \in \Lambda$, where $\phi_1 = N$, $\phi_2 = 1$, $\phi_\infty = \sqrt{N}$, implying that $\mathcal{J}(u,v)$ is strictly convex, see Proposition 5.1. Moreover, let us assume that $\omega_\lambda \geq \gamma > 0$ for all $\lambda \in \Lambda$. Suppose $\|T\| < 1$ resulting in $\|I - T^*T\| \leq 1$. Denote

$$\beta := \sup_{\lambda\in\Lambda}\frac{\phi_q}{4\theta_\lambda\omega_\lambda + 4\theta_\lambda(1 - \|I - T^*T\|)} \leq \frac{\phi_q}{4\sigma} < 1. \tag{5.12}$$

With this previous result and notations, we formulate the fully implementable and convergent scheme.

JOINTSPARSE

Input: Data vector $(g_j)_{j=1}^N$, initial sequences $\mathbf{j}^{(0)} = 0 \in \ell_2(\Lambda, \mathbb{R}^N)$, $v^{(0)}$ with $0 \leq v_\lambda^{(0)} \leq \rho_\lambda$, $\varepsilon^{(0)} := \sqrt{\frac{\mathcal{J}(0,v^{(0)})}{\gamma}}$, number $n_{\max}$ of outer iterations, number of inner iterations L_n, $n = 1, \dots, n_{\max}$.

Parameters: $q \in [1,\infty]$, positive weights (θ_λ), (ρ_λ), (ω_λ) with $\omega_\lambda \geq c > 0$, such that $\Phi^{(q)}$ and hence $\mathcal{J}$ are convex, see Proposition 5.1

Output: Approximation $(\mathbf{j}^*, v^*)$ of the minimizer of $\mathcal{J}_{\theta,\rho,\omega}^{(q)}$

$\mathbf{j}^{(0,0)} := \mathbf{j}^{(0)}$;
$\varepsilon^{(0,0)} := \varepsilon^{(0)}$;
`for` $n := 0$ `to` $n_{\max}$ `do`
  Choose $L_n > 0$ *such that* $\alpha^{L_n} + \beta < 1$;
  `for` $m := 0$ `to` L_n `do`
    $\mathbf{j}^{(n,m+1)} := U_{v^{(n)},\omega}^{(q)} \left(\mathbf{j}^{(n,m)} + T^* g - \mathbf{APPLY}[\frac{\varepsilon^{(n,m)}}{K}, T^*T, u^{(n,m)}] \right)$;
    $\varepsilon^{(n,m+1)} := \alpha \varepsilon^{(n,m)}$;
  `endfor`
  $\mathbf{j}^{(n+1,0)} := \mathbf{j}^{(n,L_n)}$;
  $v^{(n+1)} := \left(\begin{cases} \rho_\lambda - \frac{1}{2\theta_\lambda} \|\mathbf{j}^{(n+1,0)}{}_\lambda\|_q, & \|\mathbf{j}^{(n+1,0)}{}_\lambda\|_q < 2\theta_\lambda \rho_\lambda \\ 0, & \textit{otherwise} . \end{cases} \right)_{\lambda \in \Lambda}$;
  $\varepsilon^{(n+1,0)} := \varepsilon^{(n+1)} := (\alpha^{L_n} + \beta)\varepsilon^{(n)}$;
`endfor`
$\mathbf{j}^* := \mathbf{j}^{(n_{\max}, L_{n_{\max}})}$;
$v^* := v^{(n_{\max})}$.

Theorem 5.5. *Under the assumptions so far considered, the algorithm converges in norm and in particular we have*

$$\|\mathbf{j}^{(n,0)} - \mathbf{j}^*\|_2 \leq \varepsilon^{(n)}.$$

Proof. We prove the statement by induction. We have

$$\begin{aligned} \|\mathbf{j}^{(0,0)} - \mathbf{j}^*\|_2^2 &= \|\mathbf{j}^*\|_2^2 \leq \frac{\sum_\lambda \omega_\lambda \|\mathbf{j}_\lambda^*\|_2^2}{\gamma} \\ &\leq \frac{\mathcal{J}(\mathbf{j}^*, v^*)}{\gamma} \leq \frac{\mathcal{J}(0, v^{(0)})}{\gamma} \\ &= (\varepsilon^{(0)})^2. \end{aligned}$$

Assume then that

$$\|\mathbf{j}^{(n-1,0)} - \mathbf{j}^*\|_2 \leq \varepsilon^{(n-1)}.$$

From Proposition 5.4 and [8, Proposition 5.4] we have

$$\begin{aligned}\|\mathbf{j}^{(n,0)} - \mathbf{j}^*\|_2 &\leq \|\mathbf{j}^{(n,0)} - \mathbf{j}^{(n-1,\infty)}\|_2 + \|\mathbf{j}^{(n-1,\infty)} - \mathbf{j}^*\|_2 \\ &\leq \alpha^{L_{n-1}}\varepsilon^{(n-1)} + \beta\|\mathbf{j}^{(n-1,0)} - \mathbf{j}^*\|_2 \\ &\leq \left(\alpha^{L_{n-1}} + \beta\right)\varepsilon^{(n-1)} \\ &= \varepsilon^{(n)}.\end{aligned}$$

This concludes the proof. □

In the following sections we prove that the Biot-Savart operator induces compressible matrices in *wavelet coordinates*. This will allow for the efficient implementation of the procedure **APPLY** and of the whole algorithm **JOINTSPARSE** for the MEG problem.

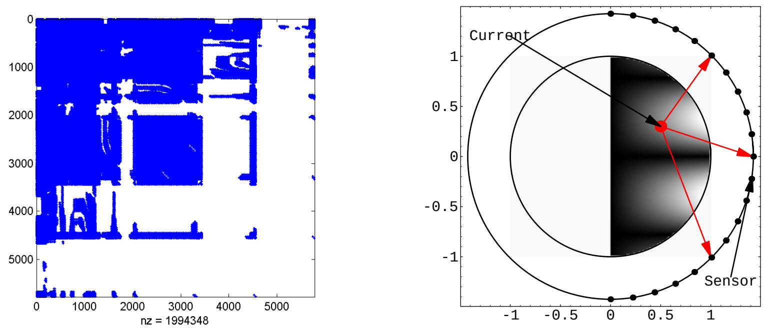

Figure 1: **Right**: Compressed matrix $\mathcal{M}$ (6.4) associated to the Biot-Savart operator in 2D, i.e., $\mathcal{B}(x,j) := \int_{V_0} \frac{(x-y)^\perp}{\|x-y\|^2_{\mathbb{R}^2}} j(y)dy$, with respect to Daubechies wavelets with $d^* = 4$ vanishing moments. We retain only the entries which exceed 10^{-6}. Only the 5.97% of the matrix is non-zero. **Left**: 2D model of MEG. Sensors are distributed on a semicircle around the area where current densities are distributed.

6 Compressibility of the Biot-Savart Operator in Wavelet Coordinates

The minimization of $\mathcal{J}(\mathbf{j}, v)$ can be implemented only if the action of the operator T^*T can be efficiently approximated by a suitable procedure **APPLY**.

For this we need to study the compressibility properties of the Biot-Savart operator with respect to wavelet coordinates, see also [11]. First, we should express T^*T in matrix form and clarify its structure. Straightforward computations give the following results.

Lemma 6.1. *The operator* $T : \ell_2(\Lambda;\mathbb{R}^3) \to \mathbb{R}^N$ *as in* (4.8) *is bounded and is expressed in wavelet coordinates by the formula*

$$T\mathbf{j} = \left(\sum_{\lambda\in\Lambda} \sum_{\ell=1}^{3} j_\lambda^\ell \, A_{\ell,k}\, \psi_\lambda \right)_{k=1}^{N}. \tag{6.1}$$

Similarly, its adjoint $T^* : \mathbb{R}^N \to \ell_2(\Lambda;\mathbb{R}^3)$ *is given by*

$$T^* g = \left(\sum_{k=1}^{N} g_k \, A_{\ell,k}\, \psi_\lambda \right)_{\lambda\in\Lambda}^{\ell=1,2,3}. \tag{6.2}$$

By combining (6.1) and (6.2) we find the expression of $T^*T : \ell_2(\Lambda;\mathbb{R}^3) \to \ell_2(\Lambda;\mathbb{R}^3)$, i.e.

$$T^*T\,\mathbf{j} = \left(\sum_{\mu\in\Lambda} \sum_{m=1}^{3} \left(\sum_{k=1}^{N} (A_{\ell,k}\,\psi_\lambda)(A_{m,k}\psi_\mu) \right) j_\mu^m \right)_{\lambda\in\Lambda}^{\ell=1,2,3}. \tag{6.3}$$

Therefore, let us denote by $\mathcal{M}$ the matrix associated to T^*T in wavelet coordinates whose entries are

$$(\mathcal{M})_{(\lambda,\ell),(\mu,m)} := \left(\sum_{k=1}^{N} (A_{\ell,k}\psi_\lambda)(A_{m,k}\psi_\mu) \right). \tag{6.4}$$

Theorem 6.2. *Let* $r = -2$. *Then we have*

$$\begin{aligned} |\mathcal{M}_{(\lambda,\ell),(\mu,m)}| \;&\leq\; C 2^{-(|\lambda|+|\mu|)(3/2+d^*+1)} \times \\ &\times \begin{cases} N \operatorname{dist}(\Omega_\lambda,\Omega_\mu)^{-(3+r+d^*+2)}, & \operatorname{dist}(\Omega_\lambda,\Omega_\mu) > 0 \\ \sum_{k=1}^{N} (\operatorname{dist}(x_k,\Omega_\lambda)\operatorname{dist}(x_k,\Omega_\mu))^{-(3+r+d^*+2)}, & \textit{otherwise.} \end{cases} \end{aligned}$$

Proof. Let us express the Biot-Savart operator as

$$\mathcal{B}(x,\mathbf{J}) \;=\; \frac{\mu_0}{4\pi} \int_{V_0} \frac{\mathbf{J}(y) \times (x-y)}{|x-y|^3} d(y) = \int_{V_0} \nabla_x \Phi(x,y) \times \mathbf{J}(y) d(y),$$

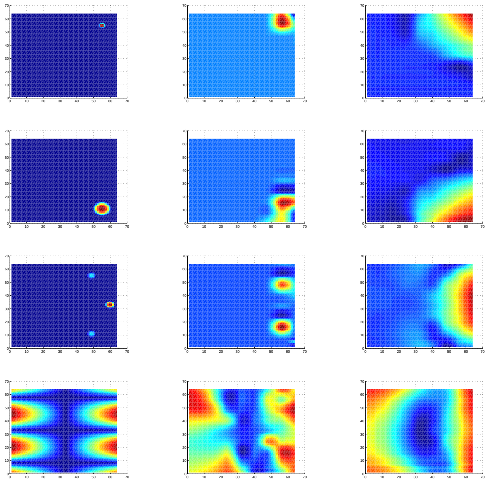

Figure 2: In the first column we illustrate some fictitious spatially localized (sparse) current densities as in the 2D model of Figure 1. In the second column we illustrate the result of the reconstruction due to our algorithm for suitable choices of the parameters. The third column shows the corresponding reconstruction due to Tikhonov regularization. The use of **JOINTSPARSE** significantly outperforms the classical Tikhonov regularization, improving the spatial resolution of the current densities.

where Φ is the potential kernel $\Phi(x,y) = \frac{\mu_0}{4\pi}\frac{1}{|x-y|}, \quad x \neq y$. This implies

$$\begin{aligned} A_{\ell,k}J_\ell &= \int_{V_0} \left(\nabla_x \Phi(x_k,y) \times \mathbf{e}_r(x_k)\right)_\ell J_\ell(y)dy \\ &= \int_{V_0} \left(\pm\partial_x^\alpha \Phi(x_k,y)e_{r,p}(x_k) \pm \partial_x^\beta \Phi(x_k,y)e_{r,q}(x_k)\right) J_\ell(y)dy, \end{aligned}$$

for suitable multi-index $\alpha = (\alpha_1,\alpha_2,\alpha_3)$ and $\beta = (\beta_1,\beta_2,\beta_3)$ of length 1 and for suitable $p,q = 1,2,3$ depending on ℓ.

Since we want to estimate $|A_{\ell,k}\psi_\lambda|$, it is sufficient to have a bound for $\left|\int_{V_0} \partial_x^{\alpha_0}\Phi(x,y)\psi_\lambda(y)dy\right|$, for any multi-index α_0 of length 1. Let us recall first [11] that for $r=-2$ we have

$$|\partial_x^\alpha \partial_y^\beta \Phi(x,y)| \leq c_{\alpha,\beta}|x-y|^{-(3+r+|\alpha|+|\beta|)},$$

with $3+r+|\alpha|+|\beta| > 0$ for any multi-index α,β. This means that

$\Psi_{\alpha_0}(x,y) := \partial_x^{\alpha_0}\Phi(x,y)$ has the property

$$|\partial_x^\alpha \partial_y^\beta \Psi_{\alpha_0}(x,y)| \le c_{\alpha+\alpha_0,\beta}|x-y|^{-(3+r+|\alpha|+|\beta|+1)}.$$

We develop Ψ_{α_0} into power series around the point $y_0 \in \Omega_\lambda$:

$$\Psi_{\alpha_0}(x,y) = \sum_{|\beta|\le d^*+1} c_\beta(x,y_0)(y-y_0)^\beta + R_{d^*+1}(x,y_0,y),$$

where

$$R_{d^*+1}(x,y_0,y) = \sum_{|\beta|=d^*+1} \frac{d^*+1}{\beta!}(y-y_0)^\beta \int_0^1 (1-t)^{d^*} \partial_y^\beta \Psi_{\alpha_0}(x, y_0+t(y-y_0))dt.$$

For $y \in \Omega_\lambda$, we have the estimates

$$\begin{aligned}
|R_{d^*+1}(x,y_0,y)| &\le \left(\sum_{|\beta|=d^*+1} \frac{d^*+1}{\beta!}|y-y_0|^{d^*+1} \sup_{y\in\Omega_\lambda} |\partial_y^\beta \Psi_{\alpha_0}(x,y)| \right) \\
&\le c_{\alpha_0,\beta} \sup_{y\in\Omega_\lambda} |x-y|^{-(3+r+d^*+2)} \left(\sum_{|\beta|=d^*+1} \frac{d^*+1}{\beta!}|y-y_0|^{d^*+1} \right) \\
&\le c_{\alpha_0,\beta}\, \mathrm{dist}(x,\Omega_\lambda)^{-(3+r+d^*+2)} \left(\sum_{|\beta|=d^*+1} \frac{d^*+1}{\beta!}|y-y_0|^{d^*+1} \right).
\end{aligned}$$

By the latter estimates, we have

$$\left| \int_{V_0} \partial_x^{\alpha_0}\Phi(x,y)\psi_\lambda(y)dy \right| = \left| \int_{\Omega_\lambda} \Psi_{\alpha_0}(x,y)\psi_\lambda(y)dy \right|$$

$$\le c_{\alpha_0,\beta}\, \mathrm{dist}(x,\Omega_\lambda)^{-(3+r+d^*+2)} \int_{\Omega_\lambda} \sum_{|\beta|=d^*+1} \frac{d^*+1}{\beta!}|y-y_0|^{d^*+1}|\psi_\lambda(y)|dy.$$

By the basic properties of wavelets

$$\begin{aligned}
\int_{\Omega_\lambda} |y-y_0|^{d^*+1}|\psi_\lambda(y)|dy &= \int_{\mathbb{R}^3} |y-y_0|^{d^*+1}|\psi_\lambda(y)|dy \\
&\le C \int_{[0,1]^3} |2^{-|\lambda|}x - y_0|^{d^*+1} 2^{3/2|\lambda|} 2^{-3|\lambda|} dx \\
&\le C 2^{-|\lambda|(3/2+d^*+1)}.
\end{aligned}$$

Finally this implies

$$\left| \int_{V_0} \partial_x^{\alpha_0} \Phi(x,y)\psi_\lambda(y)dy \right| \leq C_{\alpha_0,\beta} 2^{-|\lambda|(3/2+d^*+1)} \operatorname{dist}(x,\Omega_\lambda)^{-(3+r+d^*+2)},$$

and

$$|A_{\ell,k}\psi_\lambda| \leq C'_{\alpha_0,\beta} 2^{-|\lambda|(3/2+d^*+1)} \operatorname{dist}(x_k,\Omega_\lambda)^{-(3+r+d^*+2)}.$$

Thus, by using the fact that $\frac{1}{|x-z||y-z|} \leq C_{z,\delta}\frac{1}{|x-y|}$, for $|x-z| \geq \delta, |y-z| \geq \delta$, we show immediately

$$\begin{aligned} |\mathcal{M}_{(\lambda,\ell),(\mu,m)}| &= |\sum_{k=1}^{N} (A_{\ell,k}\psi_\lambda)(A_{m,k}\psi_\mu)| \\ &\leq C'_{\alpha_0,\beta} 2^{-(|\lambda|+|\mu|)(3/2+d^*+1)} \\ &\quad \sum_{k=1}^{N} \operatorname{dist}(x_k,\Omega_\lambda)^{-(3+r+d^*+2)} \operatorname{dist}(x_k,\Omega_\mu)^{-(3+r+d^*+2)} \\ &\leq NC'_{\alpha_0,\beta} 2^{-(|\lambda|+|\mu|)(3/2+d^*+1)} \operatorname{dist}(\Omega_\lambda,\Omega_\mu)^{-(3+r+d^*+2)}. \end{aligned}$$

□

The estimates provided by the previous theorem allow for the formulation of an efficient **APPLY** routine, compare [11]. In Figure 1 we show the compressibility of the matrix $\mathcal{M}$ for the corresponding model in 2D. Figure 2 shows some numerical results and the comparison with the classical Tikhonov regularization.

References

[1] H. B. Barlow, *What is the computational goal of the neocortex?*, Large-scale neuronal theories and the brain (C. Koch and L. J. Davis, eds.), MIT Press, Cambridge, MA, 1994.

[2] D. Cohen, *Magnetonecephalography: evidence of magnetic fields produced by alpha-rhythm currents*, Science **161** (1968), no. 784-6.

[3] ______, *Magnetonecephalography: detection of the brain's electrical activity with a superconducting magnetometer*, Science **175** (1972), no. 664-6.

[4] S. Dahlke, M. Fornasier, and T. Raasch, *Adaptive frame methods for elliptic operator equations*, to appear in Adv. Comput. Math. (2007).

[5] S. Dahlke, M. Fornasier, T. Raasch, R. Stevenson, and M. Werner, *Adaptive frame methods for elliptic operator equations: The steepest descent approach*, preprint, 2005.

[6] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, 1992.

[7] I. Daubechies, M. Defrise, and C. DeMol, *An iterative thresholding algorithm for linear inverse problems*, Comm. Pure Appl. Math. **57** (2004), no. 11, 1413–1457.

[8] M. Fornasier and H. Rauhut, *Recovery algorithms for vector valued data with joint sparsity constraints*, Report 2006-27, Johann Radon Institute for Computational and Applied Mathematics (RICAM), 2006, http://www.ricam.oeaw.ac.at/publications/reports/06/rep06-27.pdf.

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

[10] L. Kühn, *Magnetic Tomography. On the Nullspace of the Biot-Savart Operator and Point Sources for Field and Domain Reconstruction*, Ph.D. thesis, Georg-August-Universität Göttingen, Germany, 2005.

[11] R. Schneider, *Multiskalen- und Wavelet-Matrixkompression: Analysisbasierte Methoden zur effizienten Lösung grosser vollbesetzter Gleichungssysteme*, Advances in Numerical Mathematics, Teubner Stuttgart, 1998.

[12] R. Stevenson, *Adaptive solution of operator equations using wavelet frames*, SIAM J. Numer. Anal. **41** (2003), no. 3, 1074–1100.

[13] J. Tropp, *Algorithms for simultaneous sparse approximation. Part II: Convex relaxation*, Signal Processing **86** (2006), 589–602.

[14] K. Urban, *On divergence-free wavelets*, Adv. Comput. Math. **4** (1995), 51–82.

[15] ______, *Wavelets in Numerical Simulation: Problem Adapted Construction and Applications*, Springer–Verlag, 2002.