Coorbit theory, multi-α-modulation frames and the concept of joint sparsity for medical multi-channel data analysis

Stephan Dahlke[*] Gerd Teschke[†]

December 3, 2007

Abstract

This paper is concerned with the analysis and decomposition of medical multi-channel data. We present a signal processing technique that reliably detects and separates signal components such as mMCG, fMCG or MMG by involving the spatio-temporal morphology of the data provided by the multi-sensor geometry of the so-called multi-channel superconducting quantum interference device (SQUID) system. The mathematical building blocks are Coorbit theory, multi-α-modulation frames and the concept of joint sparsity measures. Combining the ingredients, we end up with an iterative procedure (with component dependent projection operations) that delivers the individual signal components.

MSC: 57S25, 42C15, 42C40, 46E15
Keywords: Coorbit-theory, α-Modulation frames, medical multi-channel data, joint sparsity

1 Introduction

One focus in the field of prenatal diagnostics is the investigation of fetal developmental brain processes that are limited by the inaccessibility of the fetus. Currently there exist two techniques for the study of fetal brain function in utero namely functional magnetic resonance imaging (fMRI) [15, 17] and fetal magnetoencephalography (fMEG) [9, 10, 16, 21]. There are several advantages and disadvantages of both techniques. The fMEG, for instance, is a completely passive and non-invasive method with superior temporal resolution and is currently measured by a multi-channel superconducting quantum interference device (SQUID) system, see Figure 1. However, the fMEG is measured in the presence of environmental noise and various near-field biological signals and other interference as for example, maternal magnetocardiogram (mMCG), fetal magnetocardiogram (fMCG), uterine smooth muscle (magnetomyogram=MMG), and motion artifacts [19, 28]. After the removal of environmental noise [27], the emphasis is on the detection and separation of mMCG, fMCG and MMG. To solve this detection problem seriously is the main prerequisite for observing and analyzing the fMEG. In the majority of reported work the MCG was reduced by adaptive filtering and/or noise estimation techniques [20, 22]. In [20] different algorithms for elimination of MCG from MEG recordings are considered, e.g. direct subtraction (DS) of a MCG signal, adaptive interference cancellation (AIC), and orthogonal signal projection algorithms (OSPA). All these approaches and their slightly modified versions

[*]Philipps-University of Marburg, FB 12 Mathematics and Computer Sciences, Hans-Meerwein-Str., Lahnberge, 35032 Marburg, Germany

[†]University of Applied Sciences Neubrandenburg, Brodaer Str. 2, 17033 Neubrandenburg and as Junior Konrad-Zuse-Fellow with the Zuse-Institute Berlin, Germany.

Figure 1: Multi-channel superconducting quantum interference device (SQUID) system.

are used for fMEG detection. In this paper, we present a different data processing technique that reliably detects both, the mMCG+fMCG and MMG+"motion artifacts" by involving the spatio-temporal morphology of the data given by the multi-sensor geometry information. Mathematically, the main ingredients of our procedure are so-called *multi-α-modulation frames* (for which the construction relies on the theory of Coorbit spaces) for an *optimal/sparse signal expansion* and the concept of *joint sparsity measures*.

A *sparse representation* of an element in a Hilbert or Banach space is a series expansion with respect to an orthonormal basis or a frame that has only a small number of large/nonzero coefficients. Several types of signals appearing in nature admit sparse frame expansions and thus, sparsity is a realistic assumption for a very large class of problems. Recent developments have shown the practical impact of sparse signal reconstruction (even the possibility to reconstruct sparse signals from incomplete information [2, 3, 7]). This is in particular the case for the medical multi-channel data under consideration that usually consist of pattern representing specific biomedical information (mMCG and fMCG). But multi-channel signals (i.e., vector valued functions) may not only possess sparse frame expansions for each channel individually, but additionally (and this is the novelty) the different channels can also exhibit common sparsity patterns. The mMCG and fMCG exhibiting a very rich morphology that appear in all the channels at the same temporal locations. This will be reflected, e.g., in sparse wavelet/Gabor expansions [1, 8] with relevant coefficients appearing at the same labels, or in turn in sparse gradients with supports at the same locations. Hence, an adequate sparsity constraint is a so-called common or joint sparsity measure that promotes patterns of multi-channel data that do not belong only to one individual channel but to all of them simultaneously.

In order to sparsely represent the MCG data we propose the usage of *multi-α-modulation frames*. These frames have only been recently developed as a mixture of Gabor and wavelet frames. Wavelet frames are optimal for piecewise smooth signals with isolated singularities, whereas Gabor frames have been very successfully applied to the analysis of periodic structures. The α-modulation frames therefore have the potential to detect both features at the same time, and therefore they seem to be extremely well-suited for the problems studied in this paper. Indeed, the numerical experiments presented here definitely confirm this conjecture.

This paper is organized as follows. In Section 2, we briefly recall the setting of α-modulation frames as far as this is needed for our purposes. Then, in Section 3, we explain how these frames can be used in multi-channel data processing involving joint sparsity constraints. Finally, in the last section, we present the numerical experiments.

2 Coorbit theory and α-modulation frames

In this section, we review the basic that provide so-called α-modulation frames. We propose to treat the medical data analysis problem with this specific kind of frame expansions since varying the parameter α allows to switch between completely different frame expansions highlighting different features of the signal to be analyzed while having to manage only one frame construction principle. The focus is not yet on multi-channel data approximation but rather on the basic methodologies that apply for single-channel signals but can (in the next section) simply be extended to multi-channel data.

In general, the motivation (and central issue in applied analysis) is the problem of analyzing and approximating a given signal. The first step is always to decompose the signal with respect to a suitable set of building blocks. These building blocks may, e.g., consist of the elements of a basis, a frame, or even of the elements of huge dictionaries. Classical examples with many important practical applications are wavelet bases/frames and Gabor frames, respectively. The wavelet transform is very useful to analyze piecewise smooth signals with isolated singularities, whereas the Garbor transform is well–suited for the analysis of periodic structures such as textures. Quite surprisingly, there is a common thread behind both transforms, and that is group theory. In general, a unitary representation U of a locally compact group G in a Hilbert space $\mathcal{H}$ is called *square integrable* if there exists a function $\psi \in \mathcal{H}$ auch that

$$\int_G |\langle \psi, U(g)\psi\rangle_{\mathcal{H}}|^2 d\mu(g) < \infty,$$

where $d\mu$ denotes the (left) Haar measure on G. In this case, the *voice transform*

$$V_\psi f(g) := \langle f, U(g)\psi\rangle_{\mathcal{H}}$$

is well-defined and invertible on its range by its adjoint. It turns out that the Gabor transform can be interpreted as the voice transform associated with a representation of the Weyl-Heisenberg group in L_2, whereas the wavelet transform is related with a square-intergrable representation of the affine group in L_2.

Since both transforms have their specific advantages, it is quite natural to try to combine them in a joint transform. One way to achieve this would be to use the *affine Weyl-Heisenberg group* G_{aWH} which is the set $\mathbb{R}^{2+1} \times \mathbb{R}_+$ equipped with group law

$$(q, p, a, \varphi) \circ (q', p', a', \varphi') = (q + aq', p + a^{-1}p', aa', \varphi + \varphi' + paq').$$

This group has the *Stone-Von-Neumann representation* on $L_2(\mathbb{R})$

$$U(q, p, a, \varphi)f(x) = a^{-1/2}e^{2\pi i(p(x-q)+\varphi)} f\left(\frac{x-q}{a}\right) = e^{2\pi i\varphi} T_x M_\omega D_a f(t), \tag{1}$$

where

$$M_\omega f(t) = e^{2\pi i\omega t} f(t), \quad T_x f(t) = f(t - x), \quad \text{and} \quad D_a f(t) = |a|^{-1/2} f(t/a),$$

which obvious contains all three basic operation, i.e., dilations, modulations and translations. Unfortunately, U is not square integrable. One way to overcome this problem is to work with representations modulo quotients. In general, given a locally compact group G with closed subgroup H, we consider the quotient group $X = G/H$ and fix a section $\sigma : X \to G$. Then, we define the generalized voice transform:

$$V_\psi f(x) := \langle f, U(\sigma(x))\psi \rangle_{\mathcal{H}}. \tag{2}$$

In the case of the affine Weyl-Heisenberg group, it has been shown in [4] that by using the specific group $H := \{(0, 0, a, \varphi) \in G_{aWH}\}$ and the specific section $\sigma(x, \omega) = (x, \omega, \beta(x, \omega), 0), \beta(x, \omega) = (1 + |\omega|)^{-\alpha}, \alpha \in [0, 1)$, the associated voice transform (2) is indeed well-defined and invertible on its range. Hence, it gives rise to a mixed form of the wavelet and the Gabor transform, and it also provides some kind of homotopy between both cases. Indeed, for $\alpha = 0$, we are in the classcial Gabor setting, whereas the case $\alpha = 1$ is very close to the wavelet setting, see, e.g., [4] for details.

Once a square-integrable representation modulo quotients is established, there is also natural way to define associated smoothnness spaces, the so-called *coorbit spaces*, by collecting all functions for which the voice transform has a certain decay, see [11, 12, 13]. More precisely, given some positive measurable weight function v on X and $1 \leq p \leq \infty$, let

$$L_{p,v}(X) := \{f \text{ measurable } : \ fv \in L_p(X)\}.$$

Then, for suitable ψ, we define the spaces

$$\mathcal{H}_{p,v} := \{f : V_\psi(A_\sigma^{-1} f) \in L_{p,v}\}, \quad A_\sigma f := \int_X \langle f, U(\sigma(x))\psi \rangle_{\mathcal{H}} U(\sigma(x))\psi d\mu, \tag{3}$$

where $d\mu$ denotes a quasi-invariant measure on X. In the classical cases, i.e., for the affine group and the Weyl-Heisenberg group, one obtains the Besov spaces and the modulation spaces, respectively. In the setting of the affine Weyl-Heisenberg group and the specific case $v_s(\omega) = (1 + |\omega|)^s$, the following theorem has been shown in [4]:

Theorem 1 *Let $1 \leq p \leq \infty$, $0 \leq \alpha < 1$ and $s \in \mathbb{R}$. Let $\psi \in L_2$ with $\mathrm{supp}\hat{\psi}$ compact and $\hat{\psi} \in C^2$. Then the coorbit spaces $\mathcal{H}_{p,v_{s-\alpha(1/p-1/2)},\alpha}$ are well-defined and can be identified with the α–modulation spaces $M^{s,\alpha}_{p,p}$, which are defined by*

$$M^{s+\alpha(1/q-1/2),\alpha}_{p,p}(\mathbb{R}) = \{f \in \mathcal{S}'(\mathbb{R}) : \langle f, U(\sigma(x, \omega))\psi \rangle \in L_{p.v_s}(\mathbb{R}^2)\}. \tag{4}$$

Consequently, the α–modulation spaces are the natural smoothness spaces associated with representations modulo quotients of the affine Weyl–Heisenberg group.

When it comes to practical applications, then one can only work with discrete data, and therefore it is necessary to discretize the underlying representation in a suitable way. Indeed, in a series of papers [11, 12, 12] Feichtinger and Gröchenig have shown that a judicious discretization gives rise to frame decompositions. The general setting can be described as follows. Given an Hilbert space $\mathcal{H}$, a countable set $\{f_n : n \in \mathbb{N}\}$ is called a *frame* for $\mathcal{H}$ if

$$\|f\|^2_{\mathcal{H}} \sim \sum_{n \in \mathbb{N}} |\langle f, f_n \rangle_{\mathcal{H}}|^2 \quad \text{for all} \quad f \in \mathcal{H}. \tag{5}$$

As a consequence of (5), the corresponding operators of analysis and synthesis given by

$$F \ : \ \mathcal{H} \to \ell_2(\mathbb{N}), \quad f \mapsto (\langle f, f_n \rangle_{\mathcal{H}})_{n \in \mathbb{N}} \tag{6}$$

$$F^* \; : \; \ell_2 \to \mathcal{H}, \quad \mathbf{c} \mapsto \sum_{n\in\mathbb{N}} c_n f_n \tag{7}$$

are bounded. The composition $S := F^*F$ is boundedly invertible and gives rise to the following decomposition and reconstruction formulas:

$$f = SS^{-1}f = \sum_{n\in\mathbb{N}} \langle f, S^{-1}f_n\rangle_{\mathcal{H}} f_n = S^{-1}Sf = \sum_{n\in\mathbb{N}} \langle f, f_n\rangle_{\mathcal{H}} S^{-1}f_n. \tag{8}$$

The Feichtinger-Gröchenig theory gives rise to frame decompositions of this type, not only for the underlying representation space $\mathcal{H}$ but also for the associated coorbit spaces. Indeed, it is possible to decompose any element in the coorbit space with respect to the frame elements (atomic decomposition), and it is also possible to reconstruct it from its sequence of moments. For the case of the α-modulation spaces, these results can be summarized as follows.

Theorem 2 *Let* $1 \le p \le \infty$, $0 \le \alpha < 1$ *and* $s \in \mathbb{R}$. *Let* $\psi \in L_2$ *with* $\mathrm{supp}\hat{\psi}$ *compact and* $\hat{\psi} \in C^2$. *Then there exists* $\varepsilon_0 > 0$ *with the following property: Let* $\Lambda(\alpha) := \{(x_{j,k}, \omega_j)\}_{j,k\in\mathbb{Z}}$ *denote the point set* $\omega_j := p_\alpha(\varepsilon_j)$, $x_{j,k} := \varepsilon\beta(\omega_j)k$, $0 < \varepsilon \le \varepsilon_0$ *where*

$$p_\alpha(\omega) := \mathrm{sgn}(\omega)\left((1 + (1-\alpha)|\omega|)^{1/(1-\alpha)} - 1\right),$$

then the following holds true.

i) (Atomic decomposition) *Any* $f \in M^{s,\alpha}_{p,p}$ *can be written as*

$$f \;=\; \sum_{(j,k)\in\mathbb{Z}^2} c_{j,k}(f) T_{x_{j,k}} M_{\omega_j} D_{\beta_\alpha(\omega_j)}\psi$$

and there exist constants $0 < C_1, C_2 < \infty$ (*independent of* p) *such that*

$$C_1\|f\|_{M^{s,\alpha}_{p,p}} \;\le\; \left(\sum_{(j,k)\in\mathbb{Z}^2} |c_{j,k}(f)|^p (1 + (1-\alpha)|j|)^{\frac{s-\alpha(1/p-1/2)}{1-\alpha}p}\right)^{1/p} \;\le\; C_2\|f\|_{M^{s,\alpha}_{p,p}}.$$

ii) (Banach Frames) *The set of functions* $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}} := \{T_{x_{j,k}} M_{\omega_j} D_{\beta_\alpha(\omega_j)}\psi\}_{j,k\in\mathbb{Z}^2}$ *forms a Banach frame for* $M^{s,\alpha}_{p,p}$. *This means that:*

1) There exist constants $0 < C_1, C_2 < \infty$ *(independent of* p*) such that*

$$C_1\|f\|_{M^{s,\alpha}_{p,p}} \;\le\; \left(\sum_{(j,k)\in\mathbb{Z}^2} |\langle f, \psi_{j,k}\rangle|^p (1 + (1-\alpha)|j|)^{\frac{s-\alpha(1/p-1/2)}{1-\alpha}p}\right)^{1/p} \;\le\; C_2\|f\|_{M^{s,\alpha}_{p,p}}.$$

2) There is a bounded, linear reconstruction operator $\mathcal{S}$ *such that*

$$\mathcal{S}\left((\langle f, \psi_{j,k}\rangle_{\mathcal{H}'_{1,v_{s-\alpha(1/p-1/2)}} \times \mathcal{H}_{1,v_{s-\alpha(1/p-1/2)}}})_{j,k\in\mathbb{Z}}\right) = f.$$

In what follows, we apply the concept of α-modulation frames according to Theorem 2 to our multi-channel data. As we have mentioned in this section, we expect that these frame provide a mixture of Gabor- und wavelet frames: for small α, the frames are similar to Gabor frames and therefore suitable for texture detection (e.g. the detection oscillatory/swinging components), whereas for α close to one, the frames are similar to wavelet frames and therefore suitable to extract signal components that contain singularities (e.g. rapid jumps as they appear in heart beat pattern). By varying the parameter α, it is possible to pass from one case to the other.

3 Multi-channel data, ℓ_q-joint sparsity and recovery model

Within this section, we focus now on multi-channel data and its representation by different α-modulation frames, the concept of joint sparsity (detection of common pattern) and, finally, on establishing the signal recovery model.

The aspect of common sparsity patterns was quite recently under consideration e.g. in [25, 26]. In the framework of inverse problems/signal recovery this issue was discussed in [14]. In the latter paper the authors proposed an algorithm for solving vector valued linear inverse problems with common sparsity constraints. In [24] this approach was generalized to nonlinear ill-posed inverse problems. In what follows, we revise this specific iterative thresholding scheme for solving the MCG signal recovery problem with joint sparsity constraints. We refer the interested reader to [24] in which the vector-valued joint sparsity concept is discussed and for the projection and thresholding techniques used therein to [5, 6, 18].

In order to cast the recovery problem as an inverse problem leading to some variational functional with a suitable sparsity constraint (forcing the detection of common signal pattern), we firstly have to realize that we want to act on channels of frame coefficient sequences since we aim to identify those coefficients at labels where specific medical patterns appear. To this end, we assume we are given n channels containing m components we wish to recover, i.e. we measure data

$$y = (y_1, \ldots, y_n) \in \bigotimes_{j=1}^{n} \mathcal{Y} = \mathcal{Y}^n \ ,$$

where each channel can be represented as a sum of m different components,

$$y_j = \sum_{i=1}^{m} f_j^i \quad .$$

Suppose f_j^i belongs for $j = 1, \ldots, n$ to some Hilbert space $\mathcal{X}_i$ and that each $\mathcal{X}_i$ is spanned by one individual α_i-modulation frame $\Psi_{\alpha_i} = \{\psi_\lambda^i : \lambda \in \Lambda(\alpha_i)\}$ such that each $f_j^i \in \mathcal{X}_i$ can be expressed by

$$f_j^i = \sum_{\lambda \in \Lambda(\alpha_i)} (\boldsymbol{f}_j^i)_\lambda \psi_\lambda^i.$$

The index λ is a shorthand notation for (j,k) and $\Lambda(\alpha_i)$ for the index set corresponding to the specific choice α_i. This construction allows the choice of different smoothness spaces that are spanned by differently structured frames (different choice of α_i) and involves therewith the fact that fMCG, mMCG and MMG are of completely different nature. If we denote with $F_i : \mathcal{X}_i \to \ell_2(\Lambda_{\alpha_i})$ the associated α_i-modulation analysis operator, compare with (6), and with $\mathrm{id}_i : \mathcal{X}_i \to \mathcal{Y}$ the embedding operator, we may define the relationship between the data of the j-th channel y_j and the frame coefficients $\boldsymbol{f}_j = (\boldsymbol{f}_j^1, \ldots, \boldsymbol{f}_j^m)$ of the m associated components,

$$y_j = A\boldsymbol{f}_j = A(\boldsymbol{f}_j^1, \ldots, \boldsymbol{f}_j^m) = \sum_{i=1}^{m} \mathrm{id}_i F_i^* \boldsymbol{f}_j^i \ ,$$

where $\boldsymbol{f}_j^i \in \ell_2(\Lambda(\alpha_i))$, i.e. $\boldsymbol{f}_j = (\boldsymbol{f}_j^1, \ldots, \boldsymbol{f}_j^m) \in \bigotimes_{i=1}^m \ell_2(\Lambda_{\alpha_i})$. Consequently,

$$\begin{aligned} A \ &: \ \bigotimes_{i=1}^{m} \ell_2(\Lambda_{\alpha_i}) \to \mathcal{Y} \ \text{ via } \ (\boldsymbol{f}^1, \ldots, \boldsymbol{f}^m) \mapsto \sum_{i=1}^{m} \mathrm{id}_i F_i^* \boldsymbol{f}^i \qquad \text{and} \\ A^* \ &: \ \mathcal{Y} \to \bigotimes_{i=1}^{m} \ell_2(\Lambda_{\alpha_i}) \ \text{ via } \ y \mapsto (F_1 \mathrm{id}_1^* y, \ldots, F_m \mathrm{id}_m^* y) \ . \end{aligned}$$

Following the arguments in [14, 25] on joint sparsity and denoting with $\boldsymbol{f}^i = (\boldsymbol{f}_1^i, \ldots, \boldsymbol{f}_n^i)$ the vector of frame coefficient sequences of all n channels with respect to one specific signal component, a reasonable measure that forces a coupling of non–vanishing frame coefficients through all n channels (representing a common morphology) is of the form

$$\Phi(\boldsymbol{f}^i) = \Phi_{p_i,q_i,\omega^i}(\boldsymbol{f}^i) = \sum_{\lambda \in \Lambda(\alpha_i)} \omega_\lambda^i \|(\boldsymbol{f}_\cdot^i)_\lambda\|_{q_i}^{p_i} \tag{9}$$

with $q_i \in [1, \infty]$, $p_i \in \{1, q_i\}$, $\omega_\lambda^i \geq c > 0$ and where the q_i-norm is taken with respect to the channel index, i.e.

$$\|(\boldsymbol{f}_\cdot^i)_\lambda\|_{q_i} = \left(\sum_{j=1}^{n} |(\boldsymbol{f}_j^i)_\lambda|^{q_i} \right)^{1/q_i} .$$

Forcing for a common sparsity pattern (e.g. common heart beats) a coupling of the different channels is advantageous and can be achieved when setting, e.g., $q_i = 2$ and $p_i = 1$.

Summarizing the findings, an m component signal recovery model in a variational formulation reads as

$$J_{\mu,p,q}(\boldsymbol{f}) = J_{\mu,p,q}(\boldsymbol{f}^1, \ldots, \boldsymbol{f}^m) = \sum_{j=1}^{n} \|y_j - A\boldsymbol{f}_j\|_{\mathcal{Y}}^2 + 2\sum_{i=1}^{m} \mu_i \Phi_{p_i,q_i,\omega^i}(\boldsymbol{f}^i) \tag{10}$$

or in compact form

$$J_{\mu,p,q}(\boldsymbol{f}) = \|y - \tilde{A}\boldsymbol{f}\|_{\mathcal{Y}^n}^2 + 2\sum_{i=1}^{m} \mu_i \Phi_{p_i,q_i,\omega^i}(\boldsymbol{f}^i) \ ,$$

where we have defined the following shorthand notations

$$\tilde{A}y = (Ay_1, \ldots, Ay_n), \ \mu = (\mu_1, \ldots, \mu_m), \ p = (p_1, \ldots, p_m), \ q = (q_1, \ldots, q_m) \ .$$

An approximation to the original m different signal components (mMCG, fMCG, MMG, ...) is now computed by means of the minimizer $\boldsymbol{f} \in (\bigotimes_{i=1}^m \ell_2(\Lambda_{\alpha_i}))^n$ of (10). Unfortunately, a direct approach towards its minimization leads to a nonlinear optimality system where the frame coefficients are coupled. Instead, we propose to replace (10) by a sequence of functionals that are much easier to minimize and for which the sequence of the corresponding minimizers converges at least to a critical point of (10). To be explicit, for $\boldsymbol{f} \in (\bigotimes_{i=1}^m \ell_2(\Lambda_{\alpha_i}))^n$ and some auxiliary $\boldsymbol{a} \in (\bigotimes_{i=1}^m \ell_2(\Lambda_{\alpha_i}))^n$, we define a surrogate functional

$$J_{\mu,p,q}^s(\boldsymbol{f}, \boldsymbol{a}) := J_{\mu,p,q}(\boldsymbol{f}) + C\|\boldsymbol{f} - \boldsymbol{a}\|_{(\bigotimes_{i=1}^m \ell_2(\Lambda_{\alpha_i}))^n}^2 - \|\tilde{A}\boldsymbol{f} - \tilde{A}\boldsymbol{a}\|_{\mathcal{Y}^n}^2 \tag{11}$$

and create an iteration process by:

1. Pick some initial guess $[\boldsymbol{f}]_0 \in (\bigotimes_{i=1}^m \ell_2(\Lambda_{\alpha_i}))^n$ and some proper constant $C > 0$.

2. Derive a sequence $([\boldsymbol{f}]_k)_{k=0,1,\ldots}$ by the iteration:

$$[\boldsymbol{f}]_{k+1} = \arg \min_{\boldsymbol{f} \in (\bigotimes_{i=1}^m \ell_2(\Lambda_{\alpha_i}))^n} J_{\mu,p,q}^s(\boldsymbol{f}, [\boldsymbol{f}]_k) \qquad k = 0, 1, 2, \ldots \tag{12}$$

It will turn out that the minimizers of the surrogate functionals are easily computed. In particular, the problem decouple, and every frame coefficient can be treated separately. In order to ensure the existence of global minimizers, norm convergence of the iterates $[\boldsymbol{f}]_k$, and regularization properties, some weak assumptions (exhibiting no significant restriction) have to be made, see for details [23] and [24] and references therein.

4 Algorithmic implementation and numerical experiments

In order specify the numerical algorithm, we have to setup the constant C and to derive the necessary condition for a minimum of $J^s_{\mu,p,q}(\boldsymbol{f},\boldsymbol{a})$ yielding the concrete proceeding of iteration (12).

The constant C can be easily determined, see [23]. For $\boldsymbol{f} \in (\bigotimes_{i=1}^m \ell_2(\Lambda_{\alpha_i}))^n$, we have

$$\langle \tilde{A}\boldsymbol{f}, \tilde{A}\boldsymbol{f}\rangle_{\mathcal{Y}^n} = \sum_{j=1}^n \|A\boldsymbol{f}_j\|^2_{\mathcal{Y}} \ .$$

Since A is bounded, it holds $\|A\| = \|A^*\|$, and we may estimate

$$\langle A^*y, A^*y\rangle_{\bigotimes_{i=1}^m \ell_2(\Lambda_{\alpha_i})} = \sum_{i=1}^m \|F_i \mathrm{id}_i^* y\|^2_{\ell_2(\Lambda_{\alpha_i})} \le \sum_{i=1}^m \|F_i\|^2 \|\mathrm{id}_i^*\|^2 \|y\|^2_{\mathcal{Y}}.$$

Therefore,

$$\|\tilde{A}\boldsymbol{f}\|^2 \le \sum_{j=1}^n \sum_{i=1}^m \|F_i\|^2 \|\mathrm{id}_i^*\|^2 \|\boldsymbol{f}_j\|^2_{\bigotimes_{i=1}^m \ell_2(\Lambda_{\alpha_i})} \le \sum_{i=1}^m \|F_i\|^2 \|\mathrm{id}_i^*\|^2 \|\boldsymbol{f}\|^2_{\mathcal{Y}^n}$$

and consequently, C must be chosen such that $\|\tilde{A}\|^2 \le \sum_{i=1}^m \|F_i\|^2 \|\mathrm{id}_i^*\|^2 < C$. In order to specify the algorithm, we firstly rewrite (10),

$$J^s_{\mu,p,q}(\boldsymbol{f},\boldsymbol{a}) = \|C^{-1}\tilde{A}^*y + \boldsymbol{a} - C^{-1}\tilde{A}^*\tilde{A}\boldsymbol{a} - \boldsymbol{f}\|^2_{(\bigotimes_{i=1}^m \ell_2(\Lambda_{\alpha_i}))^n} + \frac{2}{C}\sum_{i=1}^m \mu_i \Phi_{p_i,q_i,\omega_i}(\boldsymbol{f}^i) + \text{rest} \ ,$$

where the "rest" does not depend on $\boldsymbol{f}$. The righthand side without the "rest" can be rewritten as follows

$$\begin{aligned}
& J^s_{\mu,p,q}(\boldsymbol{f},\boldsymbol{a}) - \text{rest} \\
&= \sum_{j=1}^n \|C^{-1}A^*y_j + \boldsymbol{a}_j - C^{-1}A^*A\boldsymbol{a}_j - \boldsymbol{f}_j\|^2_{\bigotimes_{i=1}^m \ell_2(\Lambda_{\alpha_i})} + \frac{2}{C}\sum_{i=1}^m \mu_i \Phi_{p_i,q_i,\omega_i}(\boldsymbol{f}^i) \\
&= \sum_{j=1}^n \sum_{i=1}^m \|C^{-1}F_i \mathrm{id}_i^*(y_j - A\boldsymbol{a}_j) + \boldsymbol{a}_j^i - \boldsymbol{f}_j^i\|^2_{\ell_2(\Lambda_{\alpha_i})} + \frac{2}{C}\sum_{i=1}^m \mu_i \Phi_{p_i,q_i,\omega_i}(\boldsymbol{f}^i) \\
&= \sum_{i=1}^m \left\{ \sum_{j=1}^n \|C^{-1}F_i \mathrm{id}_i^*(y_j - A\boldsymbol{a}_j) + \boldsymbol{a}_j^i - \boldsymbol{f}_j^i\|^2_{\ell_2(\Lambda_{\alpha_i})} + \frac{2\mu_i}{C}\Phi_{p_i,q_i,\omega_i}(\boldsymbol{f}^i) \right\} \\
&= \sum_{i=1}^m \sum_{\lambda\in\Lambda(\alpha_i)} \left\{ \sum_{j=1}^n |\left(C^{-1}F_i \mathrm{id}_i^*(y_j - A\boldsymbol{a}_j) + \boldsymbol{a}_j^i - \boldsymbol{f}_j^i\right)_\lambda|^2 + \frac{2\mu_i}{C}\omega_\lambda^i \|(\boldsymbol{f}_\cdot^i)_\lambda\|^{p_i}_{q_i} \right\} \\
&= \sum_{i=1}^m \sum_{\lambda\in\Lambda(\alpha_i)} \left\{ \|(C^{-1}F_i \mathrm{id}_i^*(y_\cdot - A\boldsymbol{a}_\cdot) + \boldsymbol{a}_\cdot^i)_\lambda - (\boldsymbol{f}_\cdot^i)_\lambda\|_2^2 + \frac{2\mu_i}{C}\omega_\lambda^i \|(\boldsymbol{f}_\cdot^i)_\lambda\|^{p_i}_{q_i} \right\} \ .
\end{aligned}$$

For $p_i = q_i$, the variational equations completely decouple and a straightforward minimization with respect to $(\boldsymbol{f}_j^i)_\lambda$ yields the necessary conditions. For $p_i = 1$, the term within the brackets is of the following general structure

$$\|y - x\|_2^2 + \nu\|x\|_q \ ,$$

with $x, y \in \mathbb{R}^n$ and some $\nu \in \mathbb{R}_+$. The minimizing element x^* of this functional is easily obtained, see [14, 24],

$$x^* = (I - P_{B_{q'}(\nu)})(y) \ , \tag{13}$$

where $P_{B_{q'}(\nu)}$ is the orthogonal projection onto the ball $B_{q'}(\nu)$ with radius ν in the dual norm of $\| \cdot \|_q$ (i.e. $1/q + 1/q' = 1$). In general, the evaluation of $P_{B_{q'}(\nu)}$ is rather difficult and only for a few individual choices of q given, see [14, 23]. For the case $q_i = 2$ (on which we shall focus), the projection is explicitly given by

$$P_{B_{q'}(\nu)}(y) = \begin{cases} y & \text{if } \|y\|_2 \leq \nu \\ \nu \frac{y}{\|y\|_2} & \text{otherwise} \end{cases} \ . \tag{14}$$

In what follows, we adapt now the algorithm to our concrete medical signal analysis problem. The 155-channel SQUID data consist (beside biological background noise) essentially of four components: fMCG, mMCG, MMG and "motion artifacts". We aim to split the multi-channel signal into fMCG+mMCG and MMG+"motion artifacts". Therefore, we set $n = 155$ and $m = 2$. Since the fMCG+mMCG is assumed to be coupled trough all the 155 channels, we put on this signal component ($i = 1$) the joint sparsity constraint. This ensures the natural condition that heart beat patterns appear in all the channels at the same (temporal) location. On the other hand, since the MMG+"motion artifacts" component ($i = 2$) can be arbitrarily (but sparsely) localized, we do not put a common sparsity constraint on this signal component. These constraint setup can be realized when choosing $p_1 = 1$, $q_1 = 2$ and $p_2 = q_2 = 1$. Finally, we have to select adequate α_i-modulation frames. Since the fMCG+mMCG component is allowed to consist of rapid jumps (being close to singularities), we prefer α_1 close to one. In contrast, the MMG+"motion artifacts" component is supposed to be much smoother, we prefer α_2 close to zero. For this particular situation, the variational functional reads as

$$\begin{aligned} & J^s_{(\mu_1,\mu_2),(1,1),(2,1)}(\boldsymbol{f}, \boldsymbol{a}) - \text{rest} \\ & = \sum_{\lambda \in \Lambda(\alpha_i)} \Big\{ \|(C^{-1} F_1 \text{id}_1^*(y_\cdot - A\boldsymbol{a}_\cdot) + \boldsymbol{a}^1_\cdot)_\lambda - (\boldsymbol{f}^1_\cdot)_\lambda\|_2^2 + \frac{2\mu_1}{C}\omega^1_\lambda \|(\boldsymbol{f}^1_\cdot)_\lambda\|_2 \\ & \qquad + \|(C^{-1} F_2 \text{id}_2^*(y_\cdot - A\boldsymbol{a}_\cdot) + \boldsymbol{a}^2_\cdot)_\lambda - (\boldsymbol{f}^2_\cdot)_\lambda\|_2^2 + \frac{2\mu_2}{C}\omega^2_\lambda \|(\boldsymbol{f}^2_\cdot)_\lambda\|_1 \Big\} \ . \end{aligned}$$

Defining

$$M^i(y_j, \boldsymbol{a}_j) := C^{-1} F_i \text{id}_i^*(y_j - A\boldsymbol{a}_j) + \boldsymbol{a}^i_j \ ,$$

the individual α_1-modulation frame coefficients of signal component 1 are given thanks to (13) and (14) by

$$(\boldsymbol{f}^1)_\lambda = ((\boldsymbol{f}^1_1)_\lambda, \ldots, (\boldsymbol{f}^1_{155})_\lambda) = (I - P_{B_2(\mu_1\omega^1_\lambda/C)}) \left((M^1(y_1, \boldsymbol{a}_1))_\lambda, \ldots, (M^1(y_{155}, \boldsymbol{a}_{155}))_\lambda\right) \tag{15}$$

for all $\lambda \in \Lambda(\alpha_1)$, whereas the α_2-modulation frame coefficients of signal component 2 are given by

$$(\boldsymbol{f}^2)_\lambda = ((\boldsymbol{f}^2_1)_\lambda, \ldots, (\boldsymbol{f}^2_{155})_\lambda) = S_{\mu_1\omega^1_\lambda/C} \left((M^2(y_1, \boldsymbol{a}_1))_\lambda, \ldots, (M^2(y_{155}, \boldsymbol{a}_{155}))_\lambda\right) \tag{16}$$

for all $\lambda \in \Lambda(\alpha_2)$ and where $S_{\mu_1\omega^1_\lambda/C}$ denotes the well-known nonlinear soft-shrinkage operator (acting on each channel individually).

With the help of (15) and (16), the iterates (12) that approximate the minimizer of (10) can finally be written as

$$\begin{bmatrix} (\boldsymbol{f}^1)_\lambda \\ (\boldsymbol{f}^2)_\lambda \end{bmatrix}_{k+1} = \begin{pmatrix} (I - P_{B_2(\mu_1\omega^1_\lambda/C)}) \left((M^1(y_1, [\boldsymbol{f}_1]_k))_\lambda, \ldots, (M^1(y_{155}, [\boldsymbol{f}_{155}]_k))_\lambda\right) \\ S_{\mu_1\omega^1_\lambda/C} \left((M^2(y_1, [\boldsymbol{f}_1]_k))_\lambda, \ldots, (M^2(y_{155}, [\boldsymbol{f}_{155}]_k))_\lambda\right) \end{pmatrix} \ . \tag{17}$$

Procedure (17) is now applied to the SQUID multi-channel data. The original data (for sake of simple illustration restricted to two channels) at different zoom level can be seen in Figure 2. One clearly observes similarities and differences of the two channels. The similarities are given by the fMCG and mMCG (fetal and maternal heart beats) signal component whereas the differences due to biological background noise, the MCG and "motion artifacts".

The results that are obtained with the application of iteration (17) (setting $\alpha_1 = 0.9$, $\alpha_2 = 0$ and $\mu_1 = \mu_2 = 0.001$) to the SQUID data are visualized in Figure 3.

Acknowledgement

Both authors thank H. Preissl (MEG Zentrum Tübingen) for encouraging discussions on the signal processing problem and for providing the SQUID data. G. T. gratefully acknowledges the partial support by Deutsche Forschungsgemeinschaft Grants TE 354/1-2, TE 354/3-1, TE 354/4-1, TE 354/5-1. S. D. gratefully acknowledges the partial support by Deutsche Forschungsgemeinschaft, Grant Da 360/4-3.

References

[1] E.J. Candès and D.L. Donoho. New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities. *Commun. Pure Appl. Math.*, 57(2):219–266, 2004.

[2] E.J. Candès, J. Romberg, and T. Tao. Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory*, 52(2):489–509, 2006.

[3] E.J. Candès and T. Tao. Near optimal signal recovery from random projections and universal encoding strategies. *IEEE Trans. Inf. Theory*, to appear.

[4] S. Dahlke, M. Fornasier, H. Rauhut, G. Steidl, and G. Teschke. Generalized Coorbit theory, Banach frames, and the relations to α–modulation spaces. *to appear in: Proc. Lond. Math. Soc.*, 2007.

[5] I. Daubechies, G. Teschke, and L. Vese. Iteratively solving linear inverse problems with general convex constraints. *Inverse Problems and Imaging*, 1(1):29–46, 2007.

[6] I. Daubechies, G. Teschke, and L. Vese. On some iterative concepts for image restoration. *Advances in Imaging and Electron Physics*, 2008.

[7] D.L. Donoho. Compressed Sensing. *IEEE Trans. Inf. Theory*, 52(4):1289–1306, 2006.

[8] M. Elad, J.-L. Starck, P. Querre, and D.L. Donoho. Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA). *Appl. Comput. Harmon. Anal.*, 19:340–358, 2005.

[9] H. Eswaran, H. Preissl, J.D. Wilson, P. Murphy, S.E. Robinson, D. Rose, J. Vrba, and C. Lowery. Short-term serial magnetoencephalography recordings of fetal auditory evoked responses. *Neurosci. Lett.*, (331):128–132, 2002.

[10] H. Eswaran, J.D. Wilson, H. Preissl, S.E. Robinson, J. Vrba, P. Murphy, D. Rose, and C.L. Lowery. First report on the magnetoencephalographic recordings of visual evoked brain activity from the human fetus. *Lancet*, (360):779–780, 2002.

[11] H. G. Feichtinger and K. Gröchenig. A unified approach to atomic decompositions via integrable group representations. *Proc.Conf. Function Spaces and Applications, Lund 1986, Lecture Notes in Math.*, 1302:52–73, 1988.

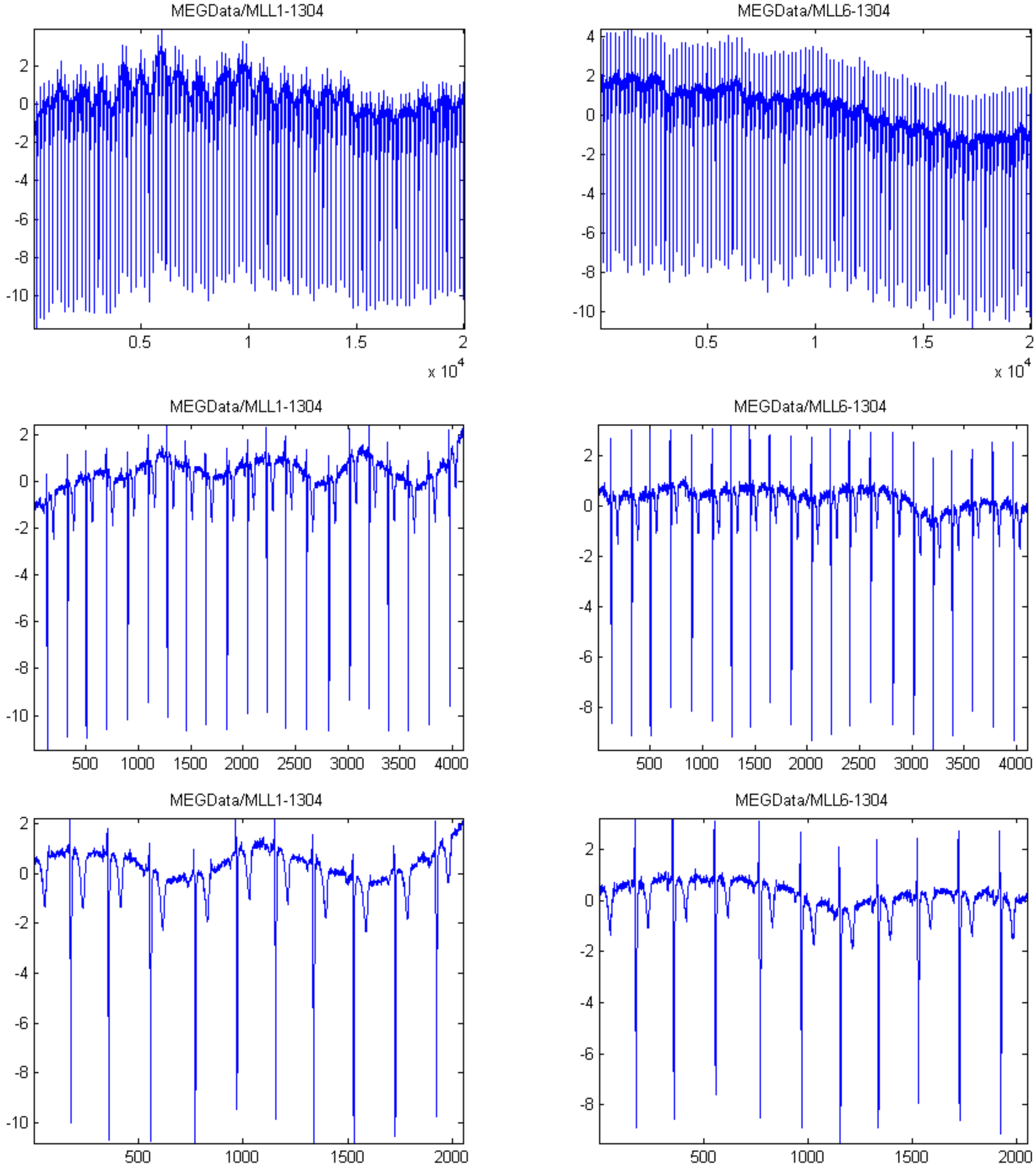

Figure 2: Two particular channels of the SQUID multi-channel data (MLL1-1304 left and MLL6-1304 right) at different zoom level.

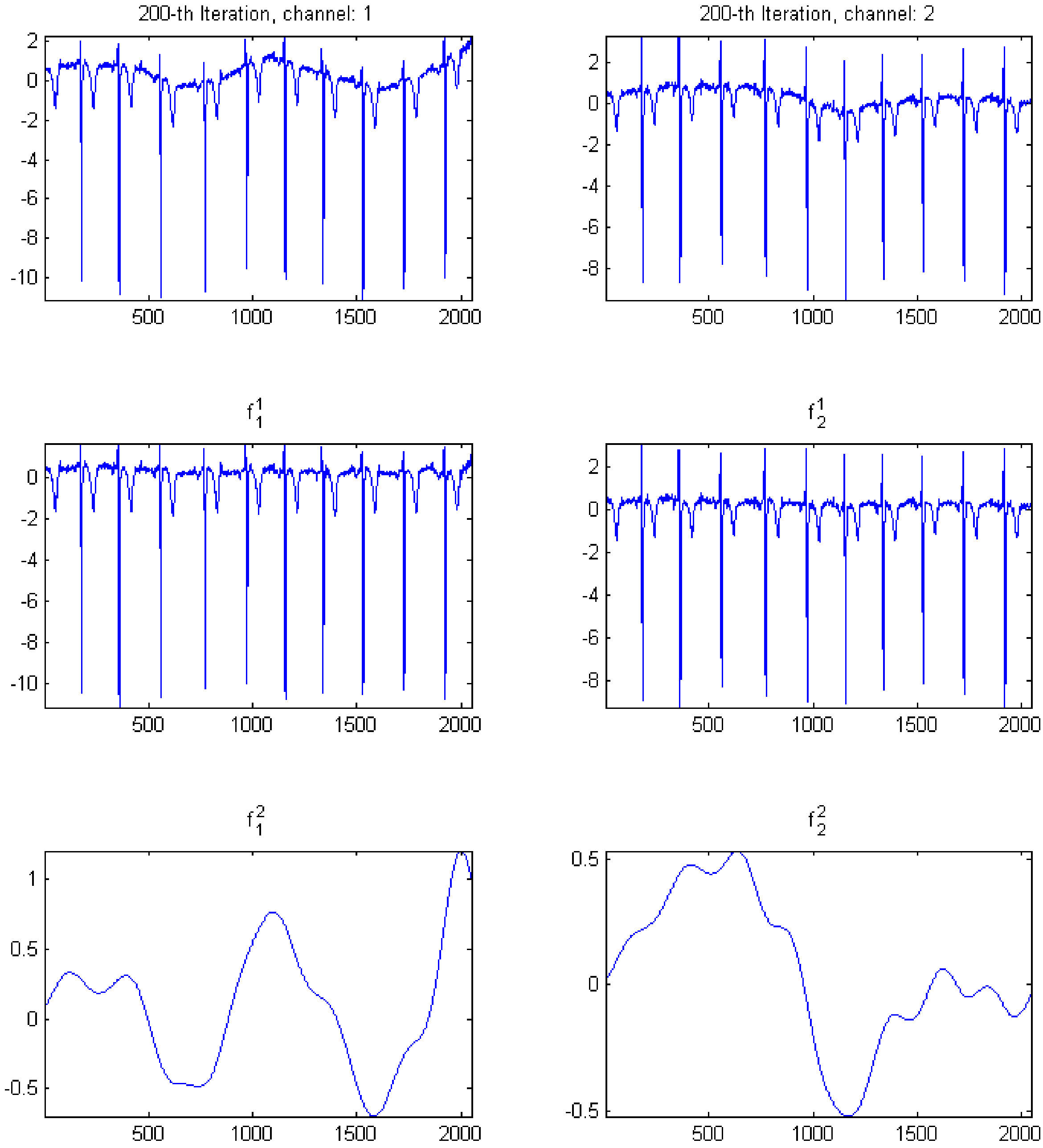

Figure 3: The reconstruction/decomposition of two particular channels of the SQUID multi-channel data (MLL1-1304 left and MLL6-1304 right). Top row: reconstructions $\boldsymbol{f}_1^1+\boldsymbol{f}_1^2$ (left) and $\boldsymbol{f}_2^1+\boldsymbol{f}_2^2$ (right); middle row: fMCG+mMCG reconstructed component $\boldsymbol{f}_1^1$ (left) and $\boldsymbol{f}_2^1$ (right); bottom row: MMG+"motion artifacts" reconstructed component $\boldsymbol{f}_1^2$ (left) and $\boldsymbol{f}_2^2$ (right).

[12] H. G. Feichtinger and K. Gröchenig. Banach spaces related to integrable group representations and their atomic decomposition I. *J. Funct. Anal.*, 86:307–340, 1989.

[13] H. G. Feichtinger and K. Gröchenig. Banach spaces related to integrable group representations and their atomic decomposition II. *Monatsh. Math.*, 108:129–148, 1989.

[14] M. Fornasier and H. Rauhut. Recovery algorithms for vector valued data with joint sparsity constraints. *SIAM J. Numer. Anal.*, 2007. to appear.

[15] J. Hykin, R. Moore, K. Duncan, S. Clare, P. Baker, I. Johnson, R. Bowtell, P. Mansfield, and P. Gowland. Fetal brain activity demonstrated by functional magnetic resonance imaging. *Lancet*, (354(9179)):645–646, 1999.

[16] J.M. Lengle, M. Chen, and R.T. Wakai. Improved neuromagnetic detection of fetal and neonatal auditory evoked response. *Clin. Neurophysiol.*, (112):785–792, 2001.

[17] R.J. Moore, J. Vadevar, S.and Fulford, D.J. Tyler, C. Gribben, P.N. Baker, D. James, and P.A. Gowland. Antenatal determination of fetal brain activity in response to an acoustic stimulus using functional magnetic resonance imaging. *Hum. Brain Mapp.*, (12(2)):94–97, 2001.

[18] R. Ramlau and G. Teschke. A projection iteration for nonlinear operator equations with sparsity constraints. *Numer. Math.*, 104:177–203, 2006.

[19] S.E. Robinson, J. Vrba, and J. McCubbin. Separating fetal meg signals from the noise, in: Nowak, h. et al. (eds.), biomag 2002, vde verlag gmbh, berlin and offenbach, germany. pages 665–667, 2002.

[20] Petrou M. Samonas, M. and A. A. Ioannides. Identification and elimination of cardiac contribution in single-trial magnetoencephalographic signals. *IEEE Trans. Biomed. Eng.*, (44):386–350, 1997.

[21] U. Schneider, E. Schleussner, J. Haueisen, H. Nowak, and H. J. Seewald. Signal analysis of auditory evoked cortical fields in fetal magnetoencephalography. *Brain Topogr.*, (14):69–80, 2001.

[22] P. Strohbach, K. Abraham-Fuchs, and W. Hrer. Event-synchronous cancellation of the heart interference in biomedical signals. *IEEE Trans. Biomed. Eng.*, (41):343–350, 1994.

[23] G. Teschke. Multi-frame representations in linear inverse problems with mixed multi-constraints. *Appl. Computat. Harmon. Anal.*, 22:43 – 60, 2007.

[24] G. Teschke and R. Ramlau. An iterative algorithm for nonlinear inverse problems with joint sparsity constraints in vector valued regimes and an application to color image inpainting. *Inverse Probl.*, 23:1851–1870, 2007.

[25] J. Tropp. Algorithms for simultaneous sparse approximation. part i: Convex relaxation. *Signal Process., special issue Sparse approximations in signal and image processing*, 86, 2006.

[26] J. Tropp, A. C. Gilbert, and M. J. Strauss. Algorithms for simultaneous sparse approximation. part i: Greedy pursuit. *Signal Process., special issue Sparse approximations in signal and image processing*, 86, 2006.

[27] J Vrba, S.E. Robinson, J. McCubbin, C.L. Lower, H. Preissl, H. Eswaran, D. Wilson, and P. Murphy. Fetal meg redistribution by projection operators. *IEEE Trans. Biomed. Eng. Submitted for publication*, 2003.

[28] R.T. Wakai and W.J. Lutter. Matched-filter template generation via spatial filtering: application to fetal biomagnetic recordings. *IEEE Trans. Biomed. Eng.*, (49):1214–1217, 2002.