Multi-structural Signal Recovery for Biomedical Compressive Sensing

Yipeng Liu*, *Member, IEEE*, Maarten De Vos, *Member, IEEE*, Ivan Gligorijevic, Vladimir Matic, Yuqian Li, *Student Member, IEEE*, and Sabine Van Huffel, *Fellow, IEEE*

***Abstract*—Compressive sensing has shown significant promise in biomedical fields. It reconstructs a signal from sub-Nyquist random linear measurements. Classical methods only exploit the sparsity in one domain. A lot of biomedical signals have additional structures, such as multi-sparsity in different domains, piecewise smoothness, low rank, etc. We propose a framework to exploit all the available structure information. A new convex programming problem is generated with multiple convex structure-inducing constraints and the linear measurement fitting constraint. With additional *a priori* information for solving the underdetermined system, the signal recovery performance can be improved. In numerical experiments, we compare the proposed method with classical methods. Both simulated data and real-life biomedical data are used. Results show that the newly proposed method achieves better reconstruction accuracy performance in term of both L1 and L2 errors.**

***Index Terms*—Biomedical signal reconstruction, compressive sensing (CS), low rank, piecewise smoothness, sparsity.**

I. Introduction

CURRENT biomedical signals usually ask large amount of data to be sampled, transmitted, stored, and processed. This results in large-scale devices, time, and power consumption [1]–[4]. Most of the current compression techniques sample the analog signal at the Nyquist rate, and then compress the data with different kinds of encoders. This acquisition process leads to a huge amount of irrelevant samples which are discarded during the compression stage of the signals. Besides, the high sampling rate requires a highly power-consuming analog-to-digital converter (ADC) with a large number of bits.

Compressive sensing (CS) can offer a solution. Rather than first sampling at a high rate and then compressing, it prefers to directly "sense" (acquire) the data in a compressive form at a much lower sampling rate [5]. CS has attracted considerable attention in signal processing. It employs linear projections that preserve the structure of the signal as much as possible; the signal is then reconstructed from these projections using nonlinear signal recovery methods. It provides a new promising framework for acquiring signals.

Signal recovery is one of the key aspects of CS. Convex optimization is a popular way, due to its high recovery accuracy, guarantee of successful recovery, and the high availability of efficient algorithms. In the early stage of CS research, sparsity has been exploited by formulating an L1-norm-based optimization problem. Sparsity is assumed in one domain as the key constraint to recover the signal [5]. Recently, progress shows that other structure information can be exploited to recover signals [6], [7], such as piecewise smoothness [8], low-rank property [9], [10], orthogonality [7], permutation [7].

Much literature exists on CS applied to biomedical signal processing, such as MRI, EMG, EEG, ECG [2], [3], [11]–[20]. However, most papers only exploit the sparsity in one domain, while many biomedical signals are sparse in more than one domain. Even more generally, some biomedical signals have structural features other than sparsity. For example, some EMG signals are sparse in both time and frequency domains [16], [17]; multichannel EMG signals are highly correlated with each other [18], which can lead to a low-rank structure in the data matrix; MRI data have both a piecewise smoothness structure and a low rank structure [2], [10].

In this paper, we give a framework for exploiting multiple structures of biomedical signals for the recovery of the signal from sub-Nyquist samples with CS. First, we generalize the sparse signal model by allowing different kinds of possible structures applicable to biomedical signals. Then, we incorporate all the available information about the data structures of the signal, by adding multiple convex structure-inducing constraints to enforce the corresponding structures in all the corresponding domains. By jointly constraining the multiple structure-inducing

Manuscript received December 5, 2012; revised March 14, 2013; April 19, 2013; accepted May 15, 2013. Date of publication May 23, 2013; date of current version September 14, 2013. This work was supported in part by the Research Council KUL: GOA MaNet, PFV/10/002 (OPTEC), IDO 08/013 Autism, several Ph.D./postdoctoral and fellow grants; in part by the Flemish Government—FWO: Ph.D./postdoctoral grants, projects: G.0427.10N (Integrated EEG-fMRI), G.0108.11 (Compressed Sensing) G.0869.12N (Tumor imaging) G.0A5513N (Deep brain stimulation), IWT: TBM070713-Accelero, TBM080658-MRI (EEG-fMRI), TBM110697-NeoGuard, Ph.D. Grants, iMinds 2013, Flanders Care: Demonstratieproject Tele-Rehab III (2012–2014); in part by the Belgian Federal Science Policy Office: IUAP P719/ (DYSCO, "Dynamical systems, control and optimization," 2012–2017), ESA AO-PGPF-01, PRODEX (CardioControl) C4000103224; in part by the EU: RECAP 209G within INTERREG IVB NWE programme, EU HIP Trial FP7-HEALTH/2007–2013 (no. 260777), EU MC ITN Transact 2012 # 316679; and in part by the Alexander von Humboldt stipend. This paper was presented at the 34th Annual International Conference of the Engineering in Medicine and Biology Society (IEEE EMBC 2012), San Diego, CA, USA, Aug. 28th–Sep. 1, 2012. *Asterisk indicates corresponding author.*

*Y. Liu is with SCD-SISTA, Department of Electrical Engineering (ESAT) and iMinds Future Health Department, KU Leuven, Heverlee 3001, Belgium (e-mail: yipeng.liu@esat.kuleuven.be).

M. De Vos is with the Department of Psychology, University of Oldenburg, Oldenburg 26129, Germany (e-mail: maarten.de.vos@uni-oldenburg.de).

I. Gligorijevic, V. Matic, and S. Van Huffel are with SCD-SISTA, Department of Electrical Engineering (ESAT) and iMinds Future Health Department, KU Leuven, Heverlee 3001, Belgium (e-mail: ivan.gligorijevic@esat.kuleuven.be; vladimir.matic@esat.kuleuven.be; sabine.vanhuffel@esat.kuleuven.be).

Y. Li is with the School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: yuqianli@uestc.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TBME.2013.2264772

0018-9294 © 2013 IEEE

norms and the data fitting, a new convex programming problem is presented for multi-structural signal recovery, which can be efficiently solved. As more *a priori* information is used to solve the largely underdetermined system, the recovery performance is expected to be enhanced. Numerical experiments show the better performance of the proposed method compared to previous methods exploiting only one sparsity constraint, with both simulated and real-life biomedical data, such as block-sparse signals, ECG, EMG, MRI.

The major contributions of this paper can be summarized as below. First, sparsity was originally regarded as one of the two fundamental premises underlying CS. Classical methods only exploit the sparsity in one domain. Here we propose a novel signal recovery framework to exploit as many kinds of data structures as possible. In addition to the sparsity in one signal domain, other data structures are taken advantage too, such as sparsity in other domains, piecewise smoothness, low rank, etc. In CS, a small number of measurements are used to recover large-scale data, which results in a largely underdetermined linear system. Hence, the signal recovery performance should be improved, provided the added regularizations are in accordance with the criterion used to judge the efficacy of a model. Second, we give a brief summary of the biomedical data structures and their representations. In the newly proposed framework, we propose three convex optimization models in CS applied to biomedical signals: L1-TV optimization for ECG signals, L1-L1 optimization for EMG signals, and L1-nuclear optimization for MRI. Generally, the used structure-inducing constraints, such as L1 norm minimization, total variation minimization (TVM), nuclear norm minimization, are based on previous investigations. They should be in accordance with the criterion used to judge the efficacy of a model. Numerical experiments also show that the proposed methods outperform the classical ones. Third, as far as we know, it is the first time that the cosparse signal recovery methods are used to recover biomedical signals from subsampled random measurements in CS. Besides their convenience to represent signals in the multi-structural signal recovery formulation, they have some other advantages, such as super-resolution, no incoherence requirement for the measurement matrix. Fourth, it is the first time that total variation (TV) optimization is used to recover ECG signals. We show that the performance outperforms that of the classical sparse signal recovery methods.

The rest of the paper is organized as follows. Section II presents the multi-structural signal model. In Section III, different kinds of norm regularizations for signal structures are discussed. In Section IV, the convex programming problem for multi-structural signal recovery is presented. Numerical results are demonstrated in Section V. In Section VI, we draw the conclusion.

II. Multi-structural Signal Model

In a practical CS system, the analog baseband signal $x(t)$ is sampled using an analog-to-information converter (AIC) [21]. The AIC can be conceptually modeled as an ADC operating at Nyquist rate, followed by a sub-Nyquist linear operation. The random sub-Nyquist measurement vector $\mathbf{y} \in \mathbf{R}^{M\times 1}$ is obtained directly from the continuous-time signal $x(t)$ by the AIC. For demonstration convenience, we formulate the sampling in discrete form as

$$\mathbf{y} = \mathbf{\Phi x} \tag{1}$$

where $\mathbf{\Phi} \in \mathbf{R}^{M\times N}$ is the measurement matrix (sampling matrix) with $M \ll N$, and $\mathbf{x} \in \mathbf{R}^{N\times 1}$ is the sampled signal which can be regarded as the original signal obtained at Nyquist sampling rate.

Because in practice, noise cannot be avoided, the obtained sampling model with noise is

$$\mathbf{y} = \mathbf{\Phi x} + \mathbf{n} \tag{2}$$

where $\mathbf{n}$ is the additive white Gaussian noise (AWGN) with zero mean and variance σ^2.

To enable CS, the measurement matrix $\mathbf{\Phi}$ should satisfy some sufficient conditions, such as the restricted isometry property [22], the coherence condition [23], the null space property [24], the constrained minimal singular values condition [25], etc. Usually one of the three types of measurement matrices is used: Gaussian matrix, Bernoulli matrix, or partial Fourier matrix.

The signal recovery from sub-Nyquist measurements is obviously an ill-posed inverse problem. The incorporation of prior information with a convex regulator is a popular way to deal with it. Such prior information specifies some simple signal structures. For biomedical signals, there are several common structures, such as sparsity, piecewise smoothness, low-rank property of the data matrix.

Sparsity exists in many biomedical signals. It means that many of the representation coefficients are close to or equal to zero, when the signal is represented in a certain domain. Traditionally, a representation model decomposes the signal into a linear combination of a few columns chosen from a predefined dictionary (representation matrix). Recently, a new signal model, called cosparse analysis model, was proposed [26]. In this new representation, an analysis operator multiplying the measurements leads to a sparse outcome. Let the signal in discrete form be expressed as

$$\boldsymbol{\theta} = \mathbf{\Psi x} \tag{3}$$

where $\mathbf{\Psi} \in \mathbf{R}^{L\times N}$ is the analysis operator (representation matrix/dictionary); $\boldsymbol{\theta} \in \mathbf{R}^{L\times 1}$ is the resulting sparse representative vector, i.e., most of the elements of $\boldsymbol{\theta}$ are zero or almost zero. Here $L \geq N$.

Besides sparsity, the processed signal has a piecewise smoothness structure in many biomedical signal processing applications [8]. The signal can be divided into several parts, and the adjacent elements of inner parts of every subsection are approximately smooth, while the elements on the boundaries of adjacent subsections can be quite different. For example, in MRI, an image often consists of several zones with abrupt boundaries between the zones.

Low rank is also a typical simple structural property of a data matrix [9], [10], [27], as originating from MRI, or multichannel EMG. The rank of a matrix is its maximum number of linearly independent columns or rows. An $L \times R$ matrix $\mathbf{\Theta}$ of rank K,

is called low-rank when $K \ll \min(L, R)$. Its singular value decomposition (SVD) is given by

$$\mathbf{\Psi X} = \mathbf{\Theta} = \mathbf{U\Sigma V}^H = \sum_{k=1}^{K} \delta_k \mathbf{u}_k \mathbf{v}_k^H \tag{4}$$

where $\mathbf{\Psi}$ is an $L \times N$ analysis operator, $\mathbf{X}$ is the $N \times R$ data matrix; $\mathbf{U} = [\, \mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_L \,]$ is an $L \times L$ unitary matrix with $\mathbf{u}_l$ being an $L \times 1$ column vector, the matrix $\mathbf{\Sigma}$ is an $L \times R$ diagonal matrix with nonnegative real numbers $\delta_k, k = 1, 2, \ldots, K$ on the diagonal, and the $R \times R$ unitary matrix $\mathbf{V}^H$ denotes the conjugate transpose of $\mathbf{V} = [\, \mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_R \,]$ with $\mathbf{v}_r$ being an $R \times 1$ column vector. Recovering it from limited information is also a problem that has received considerable attention.

Considering the fact that some biomedical signals have more than one structures simultaneously, their multi-structural model can be formulated as

$$\begin{aligned} \boldsymbol{\theta}_1 &= \mathbf{\Psi}_1 \mathbf{x} \\ \boldsymbol{\theta}_2 &= \mathbf{\Psi}_2 \mathbf{x} \\ &\vdots \\ \boldsymbol{\theta}_P &= \mathbf{\Psi}_P \mathbf{x} \end{aligned} \tag{5}$$

where P is the number of analysis linear transformation matrices $\mathbf{\Psi}_p, p = 1, 2, \ldots, P$. The corresponding expression in matrix form is

$$\begin{aligned} \mathbf{\Theta}_1 &= \mathbf{\Psi}_1 \mathbf{X} \\ \mathbf{\Theta}_2 &= \mathbf{\Psi}_2 \mathbf{X} \\ &\vdots \\ \mathbf{\Theta}_P &= \mathbf{\Psi}_P \mathbf{X}. \end{aligned} \tag{6}$$

By means of different linear transformation matrices, the resulting vectors $\boldsymbol{\theta}_p, p = 1, 2, \ldots, P$ and matrices $\mathbf{\Theta}_p, p = 1, 2, \ldots, P$ have some simple and typical structural properties, such as sparsity, piecewise smoothness, low rank property, orthogonality.

III. Structure-Inducing Constraints

After obtaining the random samples from AIC as in (1), the samples are processed in the digital signal processor (DSP) to recover the signal. Since $M \ll N$, it is an ill-posed linear inverse problem. Since many biomedical signals have simple algebraic structures, such as the ones mentioned in Section II, some corresponding structure-inducing constraints can help to successfully recover the signals in combination with the linear measurement fitting error constraint. The problem can be formulated as

$$\begin{aligned} &\min_{\mathbf{x}} f(\mathbf{x}) \\ &\text{s.t.} \quad \|\mathbf{y} - \mathbf{\Phi x}\| \le \varepsilon \end{aligned} \tag{7}$$

where $f(\mathbf{x})$ measures the degree of the structure of interest, and ε bounds the power of the AWGN in the measurements.

A. Sparsity-Inducing Constraint

A number of sparsity measures exist, such as L0 norm, L1 norm, normalized kurtosis, the Hoyer measure, Gini index, and so on [28]. Minimization/maximization of one of them can encourage sparse structure in the recovered signal. The most commonly used and studied ones are the minimization of the L0 norm and L1 norm of the estimated signal. The L0 norm is defined as $\|\mathbf{x}\|_0 = \#\{n : x_n \neq 0, n = 1, 2, \ldots, N\}$, which equals the number of nonzero elements of the vector $\mathbf{x} = [\, x_1 \quad x_2 \quad \cdots \quad x_N \,]^T$.

Using the L0 norm minimization to impose a sparse constraint in signal recovery yields

$$\begin{aligned} &\min_{\mathbf{x}} \|\mathbf{\Psi x}\|_0 \\ &\text{s.t.} \quad \mathbf{y} = \mathbf{\Phi x}. \end{aligned} \tag{8}$$

However, (8) is NP-hard unfortunately. One of the most popular ways to solve it is the basis pursuit (BP). It replaces the L0 norm with the L1 norm to yield a convex programming problem

$$\begin{aligned} &\min_{\mathbf{x}} \|\mathbf{\Psi x}\|_1 \\ &\text{s.t.} \quad \mathbf{y} = \mathbf{\Phi x} \end{aligned} \tag{9}$$

where $\|\boldsymbol{\theta}\|_1 = \sum_{n=1}^{N} |\theta_n|$ is the L1 norm of the vector $\boldsymbol{\theta} = [\, \theta_1 \quad \theta_2 \quad \cdots \quad \theta_N \,]^T$. Equation (9) can be solved efficiently by an interior-point method, subgradient algorithm [33], alternating direction method of multipliers (ADMM) [30], and so on. Because it is a convex programming problem, it can guarantee efficient computation and global optimality.

To suppress the noise in measurements as shown in (2), the linear measurement fitting error constraint can be relaxed as done in the basis pursuit denoising (BPDN). It can be formulated as

$$\begin{aligned} &\min_{\mathbf{x}} \|\mathbf{\Psi x}\|_1 \\ &\text{s.t.} \quad \|\mathbf{y} - \mathbf{\Phi x}\|_2^2 \le \varepsilon. \end{aligned} \tag{10}$$

For block-sparse signals, L2/L1 optimization, which is a general case of the BPDN, is usually used to recover the signal [34], [35]. It can be formulated as

$$\begin{aligned} &\min_{\mathbf{x}} \sum_{d=1}^{D} \|\mathbf{\Psi}_d \mathbf{x}\|_2 \\ &\text{s.t.} \quad \|\mathbf{y} - \mathbf{\Phi x}\|_2^2 \le \varepsilon \end{aligned} \tag{11}$$

where $\|\mathbf{x}\|_2 = \sqrt{\sum_{n=1}^{N} |x_n|^2}$ is the L2 norm of the vector $\mathbf{x} = [\, x_1 \quad x_2 \quad \cdots \quad x_N \,]^T$; $\mathbf{\Psi}_d, d = 1, 2, \ldots, D$ is the dth block subdictionary, which gives birth to the dth block in the sparse signal, i.e.,

$$\boldsymbol{\theta}_d = \mathbf{\Psi}_d \mathbf{x}; \tag{12}$$

$$\boldsymbol{\theta}^T = [\, \boldsymbol{\theta}_1^T \quad \boldsymbol{\theta}_2^T \quad \cdots \quad \boldsymbol{\theta}_D^T \,] \tag{13}$$

where $\boldsymbol{\theta}_d, d = 1, 2, \ldots, D$ is the dth block of $\boldsymbol{\theta}$. When $D = 1$, L2/L1 optimization reduces to BPDN.

B. Piecewise-Smoothness-Inducing Constraint

The piecewise smooth signal can have a sparse representation in the wavelet dictionary. However, TV minimization is more popular to impose a piecewise smoothness constraint. Two TV formulations exist, i.e., TV1 and TV2 [31], [32]:

$$\|\mathbf{x}\|_{TV1} = \|\mathbf{D}\mathbf{x}\|_1 \tag{14}$$

$$\|\mathbf{x}\|_{TV2} = \|\mathbf{D}\mathbf{x}\|_2 \tag{15}$$

where $\mathbf{D}$ is one of $\mathbf{D}_i, i = 1, 2, \ldots, N$ as follows:

$$\mathbf{D}_i = \begin{bmatrix} \mathbf{D}_{i,F} \\ \mathbf{D}_{i,B} \end{bmatrix} \tag{16}$$

$$\mathbf{D}_{i,F} = \begin{bmatrix} -\mathbf{1} & \mathbf{1} & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\mathbf{1} & \mathbf{1} & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & -\mathbf{1} & \mathbf{1} \\ 0 & 0 & 0 & \cdots & 0 & 0 & -\mathbf{1} \end{bmatrix} \tag{17}$$

$$\mathbf{D}_{i,B} = \begin{bmatrix} \mathbf{1} & -\mathbf{1} & 0 & \cdots & 0 & 0 & 0 \\ 0 & \mathbf{1} & -\mathbf{1} & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \mathbf{1} & -\mathbf{1} \\ 0 & 0 & 0 & \cdots & 0 & 0 & \mathbf{1} \end{bmatrix} \tag{18}$$

where $\mathbf{D}_{i,F}$ and $\mathbf{D}_{i,B}$ are the *i*th order forward and backward differential matrices; $\mathbf{1}$ is a $1 \times i$ row vector with all elements being one; and $-\mathbf{1}$ is a $1 \times i$ row vector with all elements being -1. Usually the lengths of the vectors $\mathbf{1}$ and $-\mathbf{1}$ are set to 1. When used in TV1 (14) and TV2 (15), actually $\mathbf{D}_{i,F}$ and $\mathbf{D}_{i,B}$ would result into a similar expression. Hence, usually we only need to use one of them.

Generally, TV1 is used more frequently than TV2. TV1 can be regarded as one kind of sparse constraint with the dictionary $\mathbf{D}$.

Incorporation of the TVM constraints into the optimization model (7) for signal recovery, yields

$$\min_{\mathbf{x}} \|\mathbf{x}\|_{TV} \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{\Phi}\mathbf{x}\|_2^2 \le \varepsilon. \tag{19}$$

Here we call (19) the TV optimization.

C. Low-Rank-Inducing Constraint

To force a matrix to be of low rank, we can minimize the number of nonzero singular values. Inspired by the sparsity-inducing constraint, here we define

$$\|\mathbf{X}\|_{GS-p} = \begin{cases} \left(\sum_{i=1}^{\min(M,N)} \delta_i^p \right)^{1/p}, p \in (0, +\infty) \\ \#\{\delta_i \neq 0, i = 1, 2, \ldots, \min(M,N)\}, p = 0 \end{cases} \tag{20}$$

where $\delta_i, i = 1, 2, \ldots, \min(M,N)$ are the singular values of the matrix $\mathbf{X}$. It is similar to the Schatten *p*-norm which is defined as

$$\|\mathbf{X}\|_{S-p} = \left(\sum_{k=1}^{K} \delta_k^p \right)^{1/p}, p \in [1, +\infty). \tag{21}$$

We call (20) the generalized Schatten *p*-norm, though it is not a real norm for $0 \le p < 1$. $\|\mathbf{X}\|_{GS-0}$ is the best for measuring a low-rank structure, but it is NP-hard. To improve efficiency when using a low-rank constraint, $\|\mathbf{X}\|_{GS-0}$ is relaxed to $\|\mathbf{X}\|_{GS-1}$ which is the well-known nuclear norm $\|\mathbf{X}\|_*$ [9], [10].

Combining the minimization of the nuclear norm with the data fitting error constraint, we write the problem as

$$\min_{\mathbf{X}} \|\mathbf{X}\|_* \quad \text{s.t.} \quad \mathbf{Y} = \mathbf{\Phi}\mathbf{X} \tag{22}$$

(22) is a convex programming problem, and it can be solved efficiently. This nuclear norm-based convex programming problem is often used for solving the matrix completion problem [10], [36].

Besides the constraints mentioned above, some other ones can be used, such as L2 norm minimization for Gaussian distribution structure, $L\infty$ norm minimization for uniform distribution structure, spectral norm minimization for orthogonal matrix structure, and so on [5], [7], [37].

IV. Multi-structural Signal Recovery

To improve the recovery of compressively sampled biomedical signals, we can exploit the property that some biomedical signals have multiple structures simultaneously. For example, the ECG signal is piecewise smooth as can be seen in Fig. 5 and sparse in the wavelet domain [14], [15]; EMG is sparse in both time domain and frequency domain [16], [17]; MRI has a sparse representation and low-rank property [8], [10], multichannel EMG signals are sparse in some dictionaries and of low-rank [18], [27]. If properly used, the additional *a priori* information can be helpful to improve the signal recovery performance.

Here we propose a new optimization model for multi-structural signal recovery as

$$\min_{\mathbf{x}} \sum_{p=1}^{P} \lambda_p f_p(\mathbf{x}) \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{\Phi}\mathbf{x}\|_{\diamondsuit}^2 \le \varepsilon \tag{23}$$

where P is the number of analysis operators which generate structural outcomes; $\lambda_p, p = 1, 2, \ldots, P$, is the parameter balancing the different structure-inducing constraints, which can be tuned using cross validation [38]. $\|\mathbf{x}\|_{\diamondsuit}$ is the L2 norm if $\mathbf{x}$ is a vector; and the Frobenius norm if $\mathbf{x}$ is a matrix. It is obvious that λ_1 can be set to be 1. Here we call (23) multi-structure optimization. It is a scalarized formulation of multi-criterion optimization [33]. Because all the used constraints $f_p(\mathbf{x}), p = 1, 2, \ldots, P$ are convex, efficient solutions exist, such as subgradient methods [29], decomposition methods [33], ADMM [30], and so on. Compared to the traditional

ways which only use one kind of structural information, we give some examples when more *a priori* information is used, and we expect to achieve better reconstruction performance.

We should note that the obtained optimal value is Pareto optimal. Since the multi-structural signal recovery is a multi-criterion optimization problem, we know that the optimal values could not be the same for all criterion. In practice, the choice of the λ_p may be dependent. Similarly to the choice of the parameter ε in BPDN (10), the optimal choice is dependent on the true solution $\mathbf{x}$, and is therefore difficult to obtain. For this reason, various suboptimal approaches exist for the selection. One way can be to select a fixed one based on experience. The other way is learning. In numerical experiments, the training data can be generated. Cross validation is a simple and widely used learning way. Instead of using the entire dataset when training a learner, some of the data are removed prior to training. After training, the removed data can be used to test the performance of the learned model on "new" data [38].

A. L1-TV Optimization

For piecewise smooth and sparse signals, we combine the L1 norm minimization constraint and TVM constraint in the multi-structure optimization problem (23). We set $P = 2$, $f_1(\mathbf{x}) = \|\mathbf{x}\|_{TV}$, and $f_2(\mathbf{x}) = \|\mathbf{\Psi x}\|_1$. The multi-structure optimization problem (23) reduces to

$$\min_{\mathbf{x}} \left(\|\mathbf{x}\|_{TV} + \lambda_2 \|\mathbf{\Psi x}\|_1\right) \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{\Phi x}\|_2^2 \le \varepsilon. \tag{24}$$

We call (24) the L1-TV optimization. Here we generalize the TV constraint by taking a linear combination of TV1 and TV2 constraints as $\|\mathbf{x}\|_{TV1} + \lambda \|\mathbf{x}\|_{TV2}$, where λ is a scalar balancing the two constraints.

In the special case of block-sparse signals, another form of multi-structure signal recovery can be formulated by replacing the L1 norm with L2/L1 mixed norm. The obtained optimization is

$$\min_{\mathbf{x}} \left(\|\mathbf{x}\|_{TV} + \lambda_2 \sum_{d=1}^{D} \|\mathbf{\Psi}_d \mathbf{x}\|_2\right) \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{\Phi x}\|_2^2 \le \varepsilon. \tag{25}$$

We call (25) the L2/L1-TV optimization.

B. L1-L1 Optimization

The multi-structure optimization can also be applied to the signals that are nearly sparse in multiple domains. For example, to reconstruct the EMG signals which are sparse in both time and frequency domains, we set $P = 2$, $f_1(\mathbf{x}) = \|\mathbf{x}\|_1$, and $f_2(\mathbf{x}) = \|\mathbf{Fx}\|_1$. $\mathbf{F}$ is the discrete Fourier transformation (DFT) matrix. Generally a signal cannot be sparse in both time and frequency domains. But the sparsity here does not strictly refer to the number of nonzero elements but to the number of significantly small elements. The multi-structure optimization for EMG signal recovery can be reformulated as [20]

$$\min_{\mathbf{x}} \left(\|\mathbf{x}\|_1 + \lambda_2 \|\mathbf{Fx}\|_1\right) \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{\Phi x}\|_2^2 \le \varepsilon \tag{26}$$

(26) is called L1–L1 optimization.

C. L1-Nuclear Optimization

Many biomedical images are sparse in some domains and of low rank simultaneously, such as different kinds of MRI. It also applies to many multichannel biomedical signals with highly correlated channels, such as multichannel EMG, etc. To recover this kind of signal, we set $P = 2$, $f_1(\mathbf{X}) = \|\mathrm{vec}(\mathbf{\Psi X})\|_1$, and $f_2(\mathbf{X}) = \|\mathbf{X}\|_*$, and we can obtain

$$\min_{\mathbf{X}} \left(\|\mathrm{vec}(\mathbf{\Psi X})\|_1 + \lambda_2 \|\mathbf{X}\|_*\right) \quad \text{s.t.} \quad \|\mathbf{Y} - \mathbf{\Phi X}\|_F^2 \le \varepsilon \tag{27}$$

where $\mathrm{vec}(\mathbf{X})$ puts all the columns of $\mathbf{X}$ into one vector; and $\|\mathbf{X}\|_F$ is the Frobenius norm of the matrix $\mathbf{X}$. Equation (27) is called the L1-nuclear optimization.

The general formulation for convex optimization can be written as

$$\min \quad f_0(\mathbf{x}), \quad \text{s.t.} \quad f_i(\mathbf{x}) \le b_i, i = 1, \ldots, M \tag{28}$$

where the variable $\mathbf{x}$ is of length N. The computational time is roughly proportional to $\max\{N^3, N^2M, G\}$, where G is the cost for evaluating the functions f_i and their first and second derivatives [33]. Additional regularizers ask for more computations, which makes G larger. But compared with the single structure-inducing constraint-based optimization problem, the additional computational complexity should not be significant. Considering the accuracy performance improvement, it should be worthwhile. For example, if we used the subgradient methods to solve the convex optimization problem, one subgradient of an L1 norm $\|\mathbf{Ax}\|_1$ is $\mathbf{A}^T \mathrm{sgn}(\mathbf{Ax})$, where $\mathbf{A} \in \mathbf{R}^{M \times N}$ and $\mathbf{x} \in \mathbf{R}^{N \times 1}$ [33]. The additional computation time should be approximately proportional to $4MN$ for each iteration step. Compared with M and N, the number of iteration steps to convergence should be smaller, because the length of the signal N should be considerably large where CS is used. Therefore, compared with $\max\{N^3, N^2M, G\}$, the additional computation time is not very large and can be acceptable.

V. Numerical Experiments

To quantify the performance of signal recovery, the estimation errors are calculated via the following formulas:

$$e = \frac{1}{C} \sum_{c=1}^{C} \|\mathbf{x}_c - \hat{\mathbf{x}}_c\|_b \tag{29}$$

for vectors $\mathbf{x}_c$ and $\hat{\mathbf{x}}_c$; and

$$e = \frac{1}{C} \sum_{c=1}^{C} \|\mathbf{vec}(\mathbf{X}_c - \hat{\mathbf{X}}_c)\|_b \tag{30}$$

for matrices $\mathbf{X}_c$ and $\hat{\mathbf{X}}_c$, where $\mathbf{x}_c$ and $\mathbf{X}_c$ are the original signals in the cth experiment; $\hat{\mathbf{x}}_c$ and $\hat{\mathbf{X}}_c$ are the estimated signals in the cth experiment; C is the number of experiments. $b \in \{1, 2\}$ indicates the criteria. When $b = 1$, it represents the mean L1 error; and when $b = 2$, it represents the mean L2 error.

To demonstrate the performance improvement of the proposed multi-structure optimization for biomedical signals, we perform four groups of numerical experiments. The first group uses L1-TV optimization and L2/L1-TV1 optimization to recover some simulated signals; the second group uses L1-TV1-TV2 optimization to reconstruct the ECG signals; the third group uses L1-L1 optimization to recover the EMG signals; in the fourth group, MRIs are reconstructed by L1-nuclear optimization. In each group of experiments, some related methods are used for comparison, such as the least squares (LS) methods, BPDN, nuclear norm-based matrix recovery.

In numerical experiments, the K-fold cross validation is used to learn the parameters $\lambda_p, p = 1, 2, \ldots, P$ [38]. The training data can be generated, because the original data are available, and the compressive measurements can be obtained by the product of the measurement matrix and the original data as (1). We can generate T groups of data. T iterations of training and validation are performed. In each iteration, we only use $T-1$ groups of data (training subset) for training, and use the remaining group of data for validation. In the training of each interation, we choose the optimal $\lambda_{p,t}, p = 1, 2, \ldots, P, t = 1, 2, \ldots, T-1$ to achieve a smallest residual $r_t = \|\hat{\mathbf{x}}_t - \mathbf{x}_t\|_2$ by exhaustive searching method. Then, we obtain the average $\bar{\lambda}_p = \frac{1}{T-1}\sum_{t=1}^{T-1} \lambda_{p,t}$, and $r_{\text{training}} = \frac{1}{T-1}\sum_{t=1}^{T-1} r_t$. With the learned $\bar{\lambda}_p, p = 1, 2, \ldots, P$, we can use the remaining group of data to test whether the testing residual approximately equals the average training residual, i.e., $|r_{\text{testing}} - r_{\text{training}}| \leq \delta |r_{\text{training}}|$, where $\delta \geq 0$ is a small scalar. The ten-fold cross validation ($T = 10$) is the most common, and is used here.

EEG signals are another typical class of biomedical signals. However, to our knowledge, they do not have any data structure except sparsity. In fact, even the sparsity of EEG is controversial. Zhang *et al.* [19] claims that EEG is nonsparse in the time domain and also nonsparse in transformed domains (such as the wavelet domain). This is also verified in our experiments based on our EEG data. Therefore, currently we cannot use the proposed method to recover an EEG signal from its compressive measurements.

A. Simulated Signals

In the first group of numerical experiments, we simulated a series of signals which are sparse and piecewise smooth simultaneously. The length of the signal is $N = 500$. The number of measurements ranges from $M = 10$ to $M = 100$. The measurement matrix consists of the entries sampled from an i.i.d. Gaussian distribution. The signal-to-noise ratio (SNR) of the signal is 5. Every signal is normalized by its L2 norm. Fig. 1 gives several examples of the signals. As shown, there is a nonzero block, randomly positioned in the signal. The width of the block is 50. Inside the block, the elements can be constant, linearly increasing, or sinusoidal. The number of Monte Carlo simulations

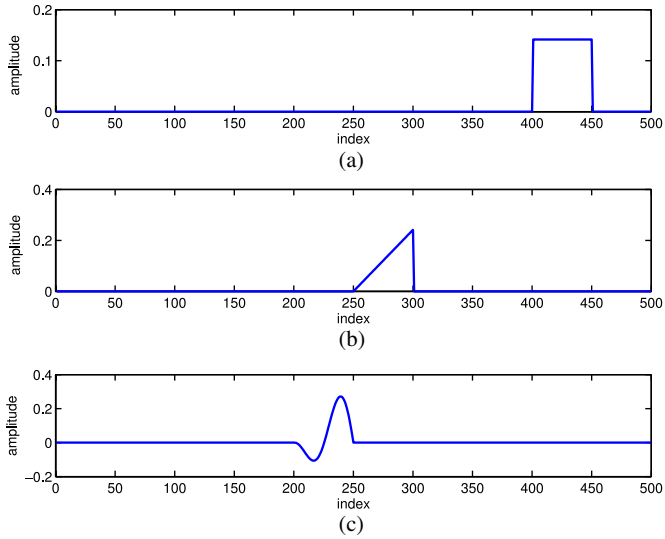


Fig. 1. Three types of block sparse signals with different inner block structures.

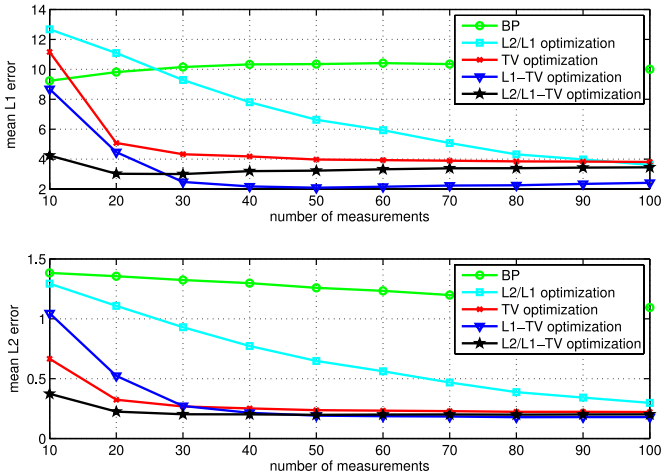


Fig. 2. Mean L1 and L2 errors versus the number of measurements with different kinds of recovery methods when the block is a rectangle as in Fig. 1(a).

is set to be 1000, i.e., $C = 1000$. Five methods are employed to recover the signals. They are BP, L2/L1 optimization, TV optimization, L1-TV optimization, and L2/L1-TV optimization. The TV herein refers to TV1.

Fig. 2 gives the mean L1 and L2 errors for simulated constant block-sparse signals in 1000 Monte Carlo simulations; Fig. 3 gives the mean L1 and L2 errors for simulated triangle block-sparse signals in 1000 Monte Carlo simulations; and Fig. 4 gives the mean L1 and L2 errors for simulated sine block-sparse signals in 1000 Monte Carlo simulations. It is obvious that the two multi-structural optimization methods, L1-TV optimization and L2/L1-TV optimization, outperform the others. The two multi-structural optimization methods almost achieve the same mean L2 error performance. The L1-TV optimization is better with mean L1 error performance.

B. ECG Signals

The used ECG data are obtained from the *Physiobank* database [39], [40]. Ninety-six hundred measurements are

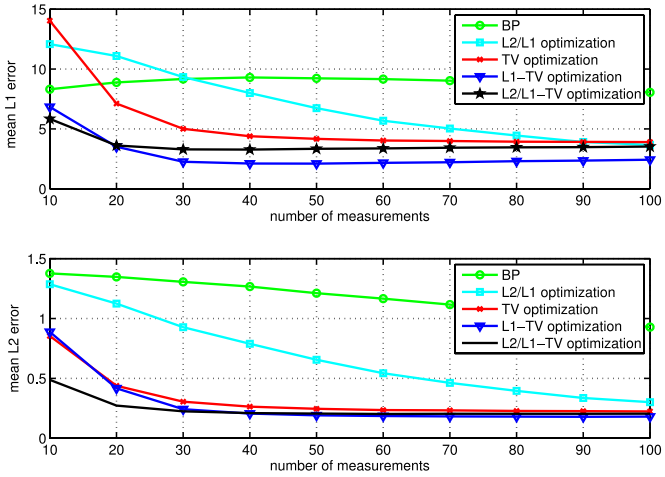


Fig. 3. Mean L1 and L2 errors versus the number of measurements with different kinds of recovery methods when the block is a triangle as in Fig. 1(b).

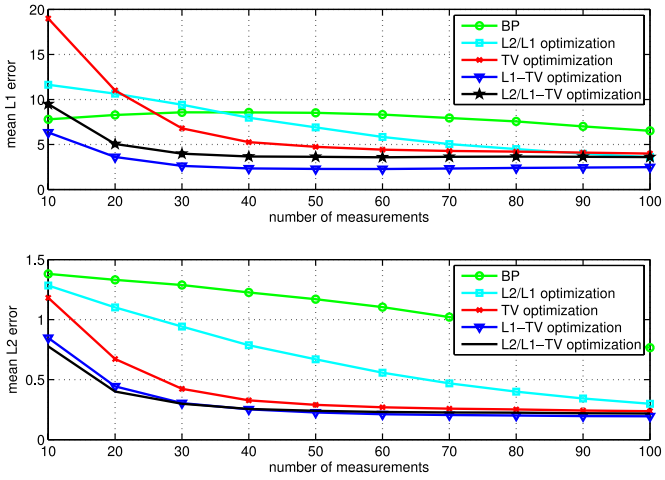


Fig. 4. Mean L1 and L2 errors versus the number of measurements with different kinds of recovery methods when the block is a period of sine waveform as in Fig. 1(c).

uniformly obtained in 1 hour and used as the original signal. Previously used signal recovery methods are mainly based on sparse signal recovery methods, such as BP, and orthogonal matching pursuit [14]. Here we propose to exploit ECG signal's piecewise smoothness property by using the TV optimization. We use the L1-TV optimization to make use of both piecewise smoothness and sparsity in the wavelet domain. The L1-TV optimization for ECG signals is compared with other single structure constraint methods, such as BP, TV optimization. The utilized dictionary is given by the orthogonal Daubechies wavelets (db 10) which is reported to be the most popular wavelet family for ECG compression [14]. Here we divide the obtained ECG signal into sections. The length of every section is $N = 512$. Fig. 5 shows one section of the ECG signals. The number of measurements used ranges from $M = 20$ to $M = 300$. The elements of the measurement matrix are i.i.d. sampled from a Gaussian distribution. Every section of the signal is normalized by its L2 norm.

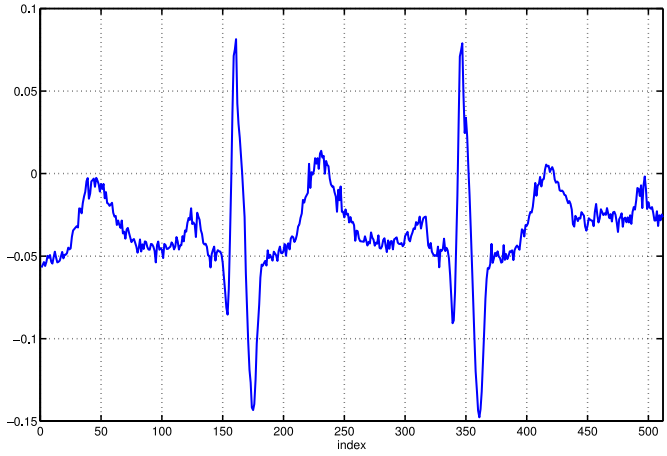


Fig. 5. One section of the used ECG signal.

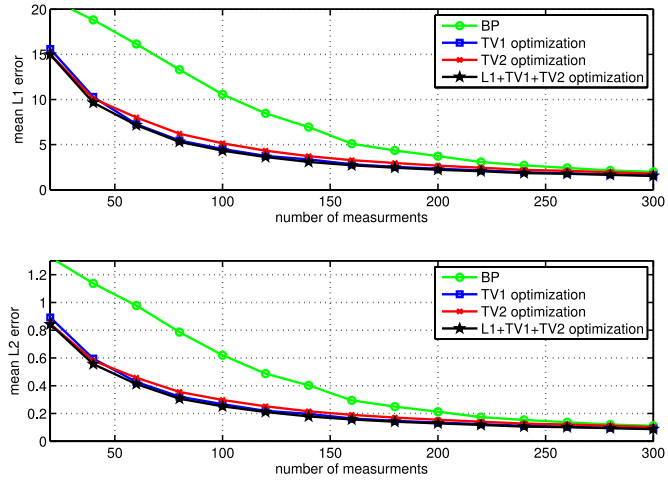


Fig. 6. Mean L1 and L2 errors versus the number of measurements with different kinds of methods for ECG signal recovery.

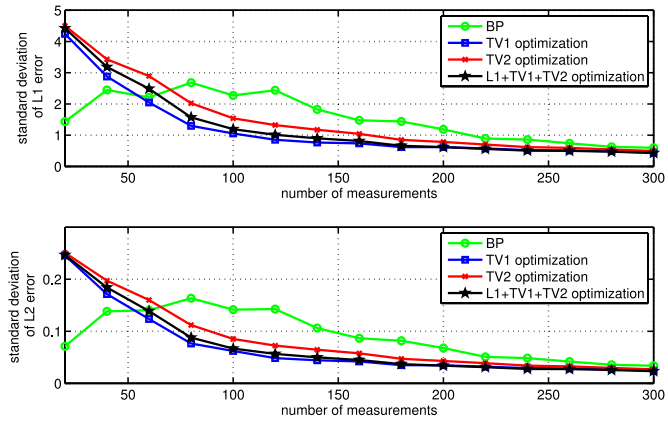


Fig. 7. Standard deviation of the L1 and L2 errors versus the number of measurements with different kinds of methods for ECG signal recovery.

Fig. 6 gives the mean L1 and L2 errors and Fig. 7 gives the standard deviation of L1 and L2 errors with the number of measurements ranging from $M = 20$ to $M = 300$ in $C = 280$ experiments. We can see that the performance of BP is far worse than the others. Comparing the other three methods,

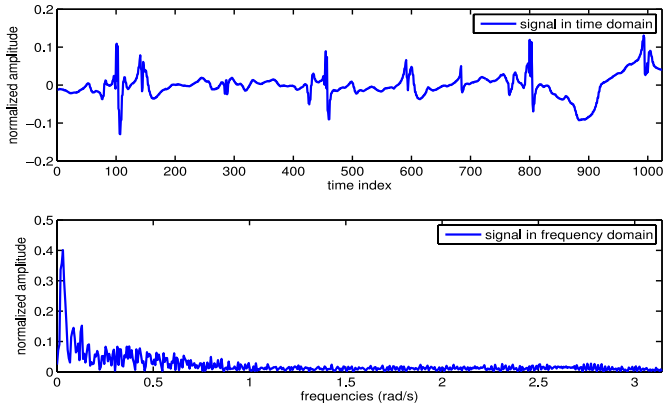


Fig. 8. Example of EMG data from a healthy person: *EMG—healthy*.

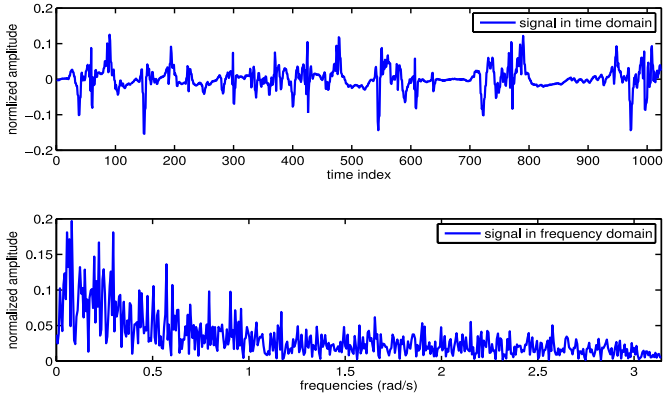


Fig. 9. Example of EMG data from a patient with myopathy: *EMG—myopathy*.

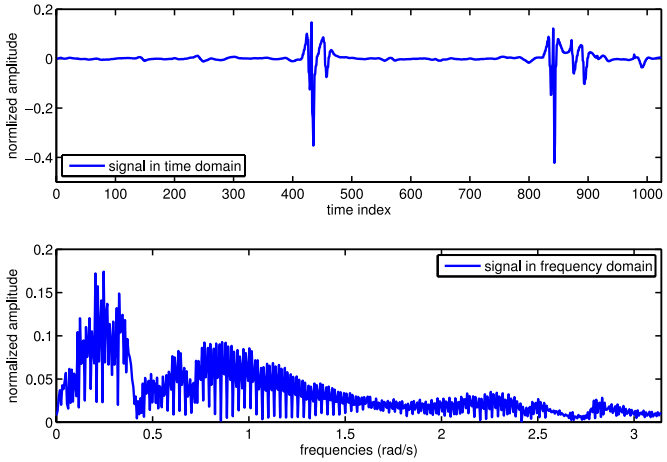


Fig. 10. Example of EMG data from a patient with neuropathy: *EMG—neuropathy*.

the L1-TV1-TV2 optimization achieves the smallest mean L1 and L2 errors from $M = 20$ to $M = 300$. It gives the best performance indeed, though the improvement is not significant. The L1-TV1-TV2 optimization and the TV1 optimization have nearly the same standard deviation performance when the number of measurements is larger than 100, i.e., when the mean errors are considerably acceptable.

C. EMG Signals

The EMG signals are obtained from the *Physiobank* database [39] too. Data were collected with a Medelec Synergy N2 EMG Monitoring System (Oxford Instruments Medical, Old Woking, U.K.). A 25-mm concentric needle electrode was placed into the tibialis anterior muscle of each subject. The patient was then asked to dorsiflex the foot gently against resistance. The needle electrode was repositioned until motor unit potentials with a rapid rise time were identified. Data were then collected for several seconds, after which the patient was asked to relax and the needle removed. Fig. 8 shows three examples of EMG data from: 1) a 44-year old man without history of neuromuscular disease; 2) a 62-year old man with chronic low back pain and neuropathy due to a right L5 radiculopathy; and 3) a 57-year old man with myopathy due to longstanding history of polymyositis, treated effectively with steroids and low-dose methotrexate. The data were recorded at 50 kHz and then downsampled to 4 kHz. During the recording process two analog filters were used: a 20-Hz high-pass filter and a 5-kHz low-pass filter.

In [16], the static thresholding algorithm is used to reconstruct the EMG signals. But those thresholding methods are proved to be worse than convex relaxation. The measurement matrix $\mathbf{\Phi}$ is formed by sampling the i.i.d. entries from a white Gaussian distribution. Here four signal recovery methods, namely the LS methods with $\min_{\mathbf{x}}\|\mathbf{x}\|_2$, s.t. $\|\mathbf{y}-\mathbf{\Phi x}\| \leq \varepsilon$, BPDN with dictionary the identity matrix (T-L1 optimization), BPDN with dictionary the DFT matrix (F-L1 optimization), and the newly proposed L1-L1 optimization with both the identity matrix and DFT matrix as the dictionaries, are used to reconstruct the EMG signals. Both T-L1 optimization and F-L1 optimization are in the form of BPDN. λ_2 is chosen to be 0.05; ε is chosen to be 5% of the measurement power, i.e., $\varepsilon = 0.05\|\mathbf{y}\|_2$. Because the amount of available data is limited, the number of experiments C is chosen to be 40 here. Every section of the signal is normalized by its L2 norm.

Figs. 8–10 show three sections of EMG signals of a healthy person (*EMG—healthy*), a patient with myopathy (*EMG—myopathy*), and a patient with neuropathy (*EMG—neuropathy*), respectively. We can see that all three signals are sparse in the time domain. In the frequency domain, *EMG—healthy* and *EMG—myopathy* signals are sparse but the *EMG—neuropathy* signal is not.

Figs. 11–13 show the recovery performance of the three different EMG signals. Here the length of the original EMG signal sections is equal to $N = 512$. All estimation errors decrease with increasing subsampling ratio *M/N*. When the subsampling ratio reaches 1, the perfect reconstruction with $e = 0$ is still not achieved. This is due to the relaxation of the constraint from $\mathbf{y} = \mathbf{\Phi x}$ to $\|\mathbf{y}-\mathbf{\Phi x}\|_2 \leq \varepsilon$. It may be the price for robustness. Besides, because all the EMG data are noisy, and the noiseless signal is not available in (29) and (30), the performance may be better than demonstrated.

Figs. 14–16 show the standard deviations of EMG signal reconstruction. We can see that the proposed method has a better standard deviation performance than those of the other sparse

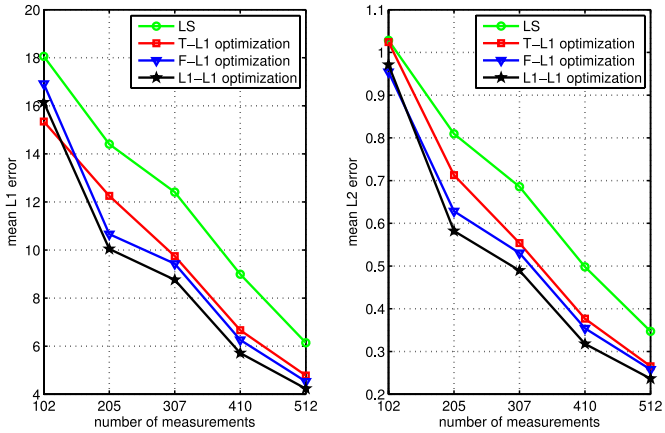


Fig. 11. Mean L1 and L2 errors versus the number of measurements with the data *EMG—healthy*.

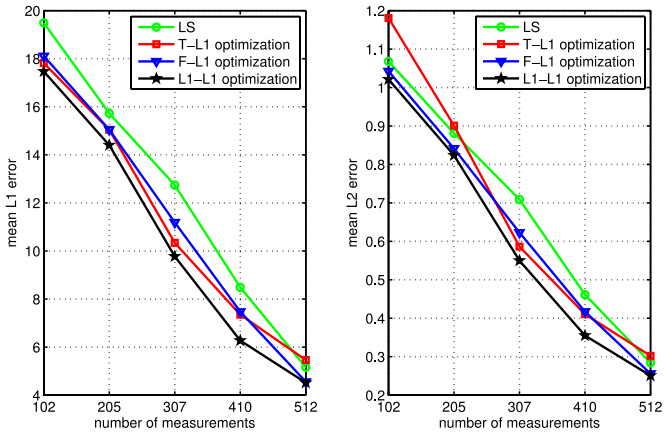


Fig. 12. Mean L1 and L2 errors versus the number of measurements with the data *EMG—myopathy*.

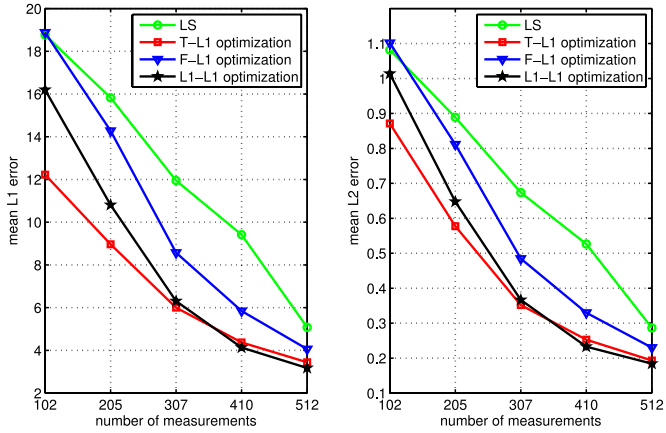


Fig. 13. Mean L1 and L2 errors versus the number of measurements with the data *EMG—neuropathy*.

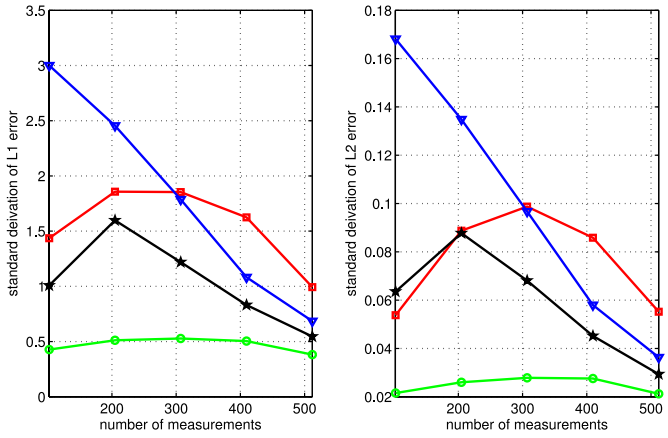


Fig. 14. Standard deviation of L1 and L2 errors versus the number of measurements with the data *EMG—healthy*.

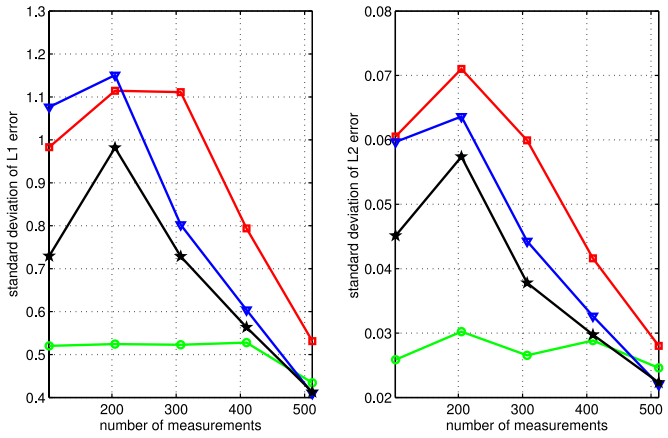


Fig. 15. Standard deviation of L1 and L2 errors versus the number of measurements with the data *EMG—myopathy*.

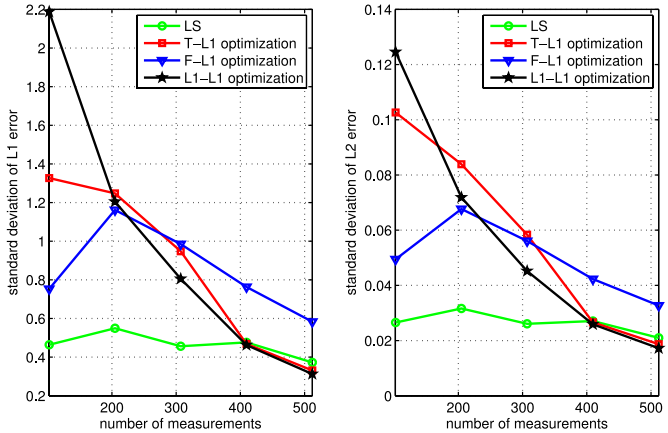


Fig. 16. Sandard deviation of L1 and L2 errors versus the number of measurements with the data *EMG—neuropathy*.

signal recovery methods in Figs. 14 and 15. In Fig. 16, we can see that the values of the standard deviation of the L1-L1 optimization are smaller than those of the other sparse signal recovery methods except when the number of measurements is smaller than about 200. Despite the fact that the proposed standard deviations of the L1-L1 optimization are larger than those of the LS, the LS method cannot be a good candidate for EMG signal recovery from compressive measurements, since the mean L1 and L2 errors of LS are much larger than that of the other three methods.

To illustrate the recovery performance more directly, Fig. 17 shows an example of the reconstruction of a section of *EMG—*

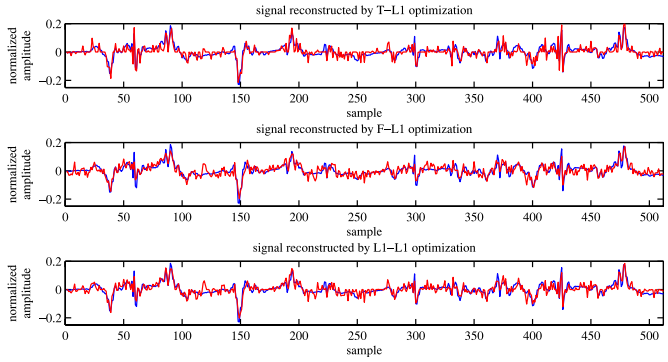


Fig. 17. Example of the reconstruction of a section of *EMG—myopathy* signal with sub-sampling ratio equals to 0.50. The red ones are the original signals; and the blues ones are the estimated signals.

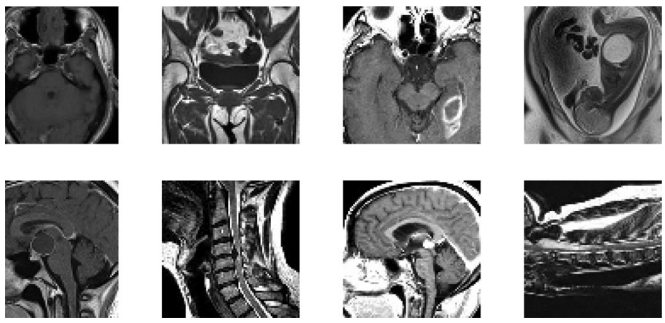


Fig. 18. Used medical images [2], [3], [8], [32], [41].

myopathy signal with subsampling ratio equals to 0.50. We can see that the profile of the signal is well reconstructed.

In Fig. 11, T-L1 optimization performs better than F-L1 optimization; but in Fig. 12, F-L1 optimization is better than T-L1 optimization. However, L1-L1 optimization is superior in both Figs. 11 and 12. In Fig. 13, we can see that L1-L1 optimization is better than F-L1 optimization, but worse than T-L1 optimization. The reason is that the EMG signal here is not sparse in the frequency domain, which is evident from Fig. 10.

In summary, if the EMG signal is approximately sparse in both time and frequency domains, L1-L1 optimization is the best candidate for compressive EMG signal recovery. Moreover, if the signal is likely to be sparse in multiple domains with a certain degree of uncertainty, the L1-L1 optimization is also a robust choice, because it can at least avoid the worst performance.

D. MRI

In the experiments, we selected $C = 8$ MRIs with 81×81 pixels, as shown in Fig. 18. The measurement matrix $\boldsymbol{\Phi}$ is formed by sampling the i.i.d. entries from a white Gaussian distribution. Because the TV1 is able to recover MRIs, and the images have low rank structure, BP (9), nuclear norm-based recovery (22), and L1-nuclear optimization (27) are used to reconstruct the images. Here the dictionary for sparse representation is $\boldsymbol{D}$ with $i = 1$ in (16). In L1-nuclear optimization, λ_2 is chosen to be 3.

Figs. 19 and 20 show the L1 and L2 errors with different number of subsampled measurements when the images in Fig. 18 are reconstructed. Every image is normalized by its maximum element. In Figs. 19 and 20, we can see that both the L1 and

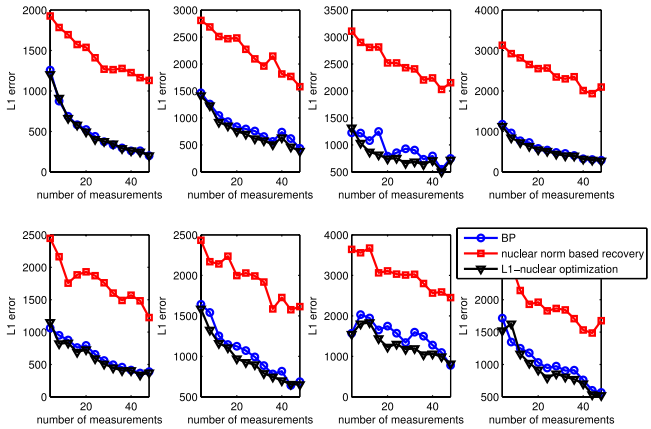


Fig. 19. L1 errors versus the number of measurement when the images in Fig. 18 are reconstructed.

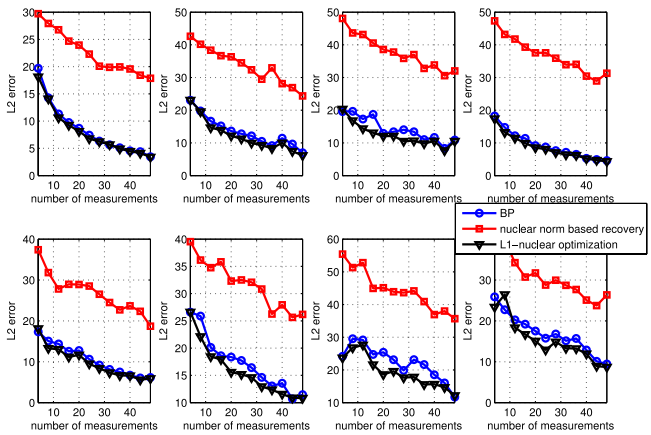


Fig. 20. L2 errors versus the number of measurement when the images in Fig. 18 are reconstructed.

L2 errors of nuclear norm-based recovery are much larger than the ones of BP, which agrees with the fact that nuclear norm minimization constraint is not used for CS, but for matrix completion [10]. In these figures, obviously we can see that the proposed L1-nuclear optimization is better than BP. Although the nuclear norm-based recovery has bad accuracy to recover the signal, the nuclear norm minimization for encouraging low rank structure in the estimated matrix can improve the performance of BP which only exploits the sparse structure.

In fact, L1-nuclear optimization has already been used to dynamic MRI. Its performance was shown in [41].

VI. Conclusion

In this paper, we give a novel framework for multi-structural signal recovery for CS. The newly proposed methods impose different data structures which are common in biomedical signals. Since more *a priori* information is exploited, the signal recovery performance is enhanced. Numerical experiments confirm the performance improvement.

Acknowledgment

The scientific responsibility of this paper is assumed by the authors.

REFERENCES

[1] C. Bachmann, M. Ashouei, V. Pop, M. Vidojkovic, H. De Groot, and B. Gyselinckx, "Low-power wireless sensor bodes for ubiquitous long-term biomedical signal monitoring," *IEEE Commun. Mag.*, vol. 50, no. 1, pp. 20–27, Jan. 2012.

[2] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn. Reson. Med.*, vol. 58, no. 6, pp. 1182–1195, Dec. 2007.

[3] M. Lustig, D.L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed sensing MRI," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 72–82, Mar. 2008.

[4] Y. Guo, S. Ruan, J. Landre, and J. Constans, "A sparse representation method for magnetic resonance spectroscopy quantification," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 7, pp. 1620–1627, Jul. 2010.

[5] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[6] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, "The convex algebraic geometry of linear inverse problems," in *Proc. 48th Annu. Allerton Conf. Commun. Contr. Comput.*, Allerton, IL, USA, Sep./Oct. 2010, pp. 699–703.

[7] V. Chandrasekaran, "Convex optimization methods for graphs and statistical modeling" Ph.D. dissertation, Massachusetts Inst. Technol., MA, USA, Apr. 2011.

[8] J. Romberg, "Imaging via compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 14–20, Mar. 2008.

[9] M. Fazel, "Matrix rank minimization with applications" Ph.D. dissertation, Electr. Eng. Dept., Stanford Univ., CA, USA, Mar. 2002.

[10] E. J. Candes and Y. Plan, "Matrix completion with noise," *Proc. IEEE*, vol. 98, no. 6, pp. 925–936, Jun. 2010.

[11] A. J. Casson and E. Rodriguez-Villegas, "Signal agnostic compressive sensing for body area networks: Comparison of signal reconstructions," in *Proc. IEEE 34th Int. Conf. Eng. Med. Biol. Soc.*, San Diego, CA, USA, Aug.–Sep. 2012, pp. 4497–4500.

[12] S. Aviyente, "Compressed sensing framework for EEG compression," in *Proc. IEEE/SP 14th Workshop Stat. Signal Process.*, Aug. 2007, pp. 181–184.

[13] V. Matic, M. De Vos, B. Mijovic, and S. Van Huffel, "Sparse approximation of the neonatal EEG," in *Proc. 4th Workshop Signal Process. Adapt. Sparse Struct. Represent.*, Edinburgh, U.K., Jun. 2011.

[14] H. Mamaghanian, N. Khaled, D. Atienza, and P. Vandergheynst, "Compressed sensing for real-time energy-efficient ECG compression on wireless body sensor nodes," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 9, pp. 2456–2466, Sep. 2011.

[15] P. S. Addison, "Wavelet transforms and the ECG: A review," *Physiol. Meas.*, vol. 26, no. 5, pp. 155–199, Aug. 2005.

[16] A. Salman, E. G. Allstot, A. Y. Chen, A. M. R. Dixon, D. Gangopadhyay, and D. J. Allstot, "Compressive sampling of EMG bio-signals," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2011, Brazil, pp. 2095–2098.

[17] A. M. R. Dixon, E. G. Allstot, D. Gangopadhyay, and D. J. Allstot, "Compressed sensing system considerations for ECG and EMG wireless biosensors," *IEEE Trans. Biomed. Circuits Syst.*, vol. 6, no. 2, pp. 156–166, May 2012.

[18] M. J. Zwarts and D. F. Stegeman, "Multichannel surface EMG: Basic aspects and clinical utility," *Muscle Nerve*, vol. 28, no. 1, pp. 1–17, Jul. 2003.

[19] Z. Zhang, T. Jung, S. Makeig, and B. D. Rao, "Compressed sensing of EEG for wireless telemonitoring with low energy consumption and inexpensive hardware," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 1, pp. 221–224, Jan. 2013.

[20] Y. Liu, I. Gligorijevic, V. Matic, M. De Vos, and S. Van Huffel, "Multi-sparse signal recovery for compressive sensing," in *Proc. 34th Annu. Int. Conf. Eng. Med. Biol. Soc.*, San Diego, CA, USA, Aug./Sep. 2012, pp. 1053–1056.

[21] J. Laska, S. Kirolos, M. F. Duarte, T. S. Ragheb, R. G. Baraniuk, and Y. Messoud, "Theory and implementation of an analog-to-information converter using random demodulation," in *Proc. IEEE Int. Symp. Circuits Syst.*, New Orleans, LI, USA, May 2007, pp. 1959–1962.

[22] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constr. Approx.*, vol. 28, no. 3, pp. 253–263, Dec. 2008.

[23] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[24] A. Cohen, W. Dahmen, and R. Devore, "Compressed sensing and best k-term approximation," *J. Amer. Math. Soc.*, vol. 22, no. 1, pp. 211–231, Jan. 2009.

[25] G. Tang and A. Nehorai, "Performance analysis of sparse recovery based on constrained minimal singular values," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5734–5745, Dec. 2011.

[26] S. Nam, M. E. Davies, M. Elad, and R. Gribonval, "Cosparse analysis modeling—uniqueness and algorithms," in *Proc. Int. Conf. Acoust. Speech, Signal Process.*, Prague, Czech Republic, May 2011, pp. 5804–5807.

[27] I. Gligorijevic, J. Van Dijk, B. Mijovic, S. Van Huffel, J. Blok, and M. De Vos, "A new and fast approach towards sEMG decomposition," *Edical Biol. Eng. Comput.*, vol. 51, no. 5, pp. 593–605, May 2013.

[28] N. Hurley and S. Rickard, "Comparing measures of sparsity," *IEEE Trans. Inf. Thoer.*, vol. 55, no. 10, pp. 4723–4741, Oct. 2009.

[29] J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems," *Proc. IEEE*, vol. 98, no. 6, pp. 948–958, Jun. 2010.

[30] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.

[31] Y. Liu and Q. Wan, "Robust beamformer based on total variation minimisation and sparse-constraint," *Electron. Lett.*, vol. 46, no. 25, pp. 1697–1699, Dec. 2010.

[32] D. Needell, and R. Ward, "Near-optimal compressed sensing guarantees for total variation minimization," arXiv:1210.3098, Oct. 2012. [Online]. Available: http://arxiv.org/abs/1210.3098.

[33] S. Boyd and L. Vanderberghe, *Convex Optimization*. New York, NY, USA: Cambridge Univ. Press, 2004.

[34] M. Stojnic, F. Parvaresh, and B. Hassibi, "On the reconstruction of block-sparse signals with an optimal number of measurements," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3075–3085, Aug. 2009.

[35] Y. C. Eldar, P. Kuppinger, and H. Bolcskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3042–3054, Jun. 2010.

[36] B. Hunyadi, M. Signoretto, S. Van Huffel, J. Suykens, and M. De Vos, "Incorporating structural information from the multichannel EEG improves patient-specific seizure detection," *Clin. Neurophysiol.*, vol. 123, no. 12, pp. 2352–2361, Dec. 2012.

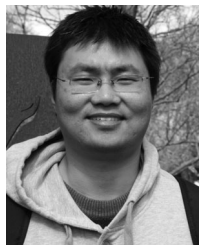
[37] M. De Vos, D. Nion, S. Van Huffel, and L. De Lathauwer, "A combination of parallel factor and independent component analysis," *Signal Process.*, vol. 92, no. 12, pp. 2990–2999, Jul. 2012.

[38] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, (Series in Statistics).. New York, NY, USA: Springer, 2001.

[39] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P. Ivanov, R. Mark, J. Mietus, G. Moody, C. Peng, and H. Stanley, "PhysioBank, Physiotoolkit, and Physionet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. 215–220, Jun. 2000.

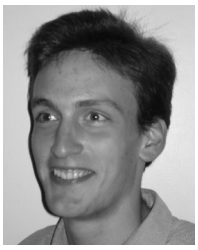
[40] R. Bousseljot, D. Kreiseler, and A. Schnabel, "Nutzung der EKG-Signaldatenbank CARDIODAT der PTB ber das Internet," *Biomed. Tech. Biomed. Eng.*, vol. 40, no. 1, pp. 317–318, Jul. 2009.

[41] S. G. Lingala, Y. Hu, E. DiBella, and M. Jacob, "Accelerated dynamic MRI exploiting sparsity and low-rank structure: k-t SLR," *IEEE Trans. Med. Imag.*, vol. 30, no. 5, pp. 1042–1054, May 2011.



Yipeng Liu (S'08–M'13) was born in Chengdu, China, in 1983. He received the B.Eng. degree in biomedical engineering and the Ph.D. degree in information and communication engineering from the University of Electronic Science and Technology of China, Chengdu, China, in June 2006 and June 2011, respectively.

From July 2007 to September 2008, he was a Research Assistant with the 10th Institute of China Electronics Technology Group Corporation (CETC-10). From November 2010 to May 2011, he was a visiting Ph.D. student at Tsinghua University. From June 2011 to November 2011, he was a Technical Staff with Huawei Technologies. Since November 2011, he has been a Postdoctoral Research Fellow with ESAT-SCD-SISTA, KU Leuven. His current research interests include compressed sensing of biomedical signals.



Maarten De Vos (M'09) received the M.Sc. degree in electrotechnical–mechanical engineering, with specialization in biomedical techniques, and the Ph.D. degree in engineering from Katholieke Universiteit Leuven, Leuven, Belgium, in 2005 and 2009, respectively.

Since 2009, he has been a Postdoctoral Researcher in the Electrical Engineering Department, SCD Division, KU Leuven, and in the Psychology Department of the University of Oldenburg. Since 2013, he has been Junior Professor at the University of Oldenburg, heading the research group "methods for cognitive psychology." His current research interests include linear and multilinear algebra, decomposition techniques for biomedical signals, and developing Human Machine Interfacing solutions.



Ivan Gligorijevic received the M.Sc. degree in electrical engineering from the University of Belgrade, Belgrade, Serbia, in July 2008 and the Ph.D. degree from Biomed Group, ESAT-SCD, KU Leuven, Leuven, Belgium, in April 2013.

Since 2013, he has been a Postdoctoral Researcher in ESAT-SCD, KU Leuven. His current research interests include developing algorithms for processing of neural and muscle signals recorded by microarray probes.



Vladimir Matic was born in Belgrade, Serbia, in 1984. He received the B.Eng. degree and M.Sc. degrees from the School of Electrical Engineering, University of Belgrade, in 2007 and 2010, respectively. Since October 2010, he has been working toward the Ph.D. degree at the ESAT-SCD-SISTA, KU Leuven, Leuven, Belgium.

From December 2007 to April 2008, he was an Intern at the Novel Mechatronic Department at AD HILTI company, Liechtenstein. His research interests include neonatal EEG signal processing.



Yuqian Li (S'11) received the B.Eng. degree in electronic engineering from University of Electronic Science and Technology of China (UESTC), China, in June 2009. Since September 2009, she has been working toward the Ph.D. degree at UESTC, Chengdu, China. From August 2010 to August 2012, she was a visiting Ph.D. student at KU Leuven, Belgium.

Her current research interests include pattern recognition, multimodal data fusion, and biomedical and remote data processing.



Sabine Van Huffel (M'96–A96–SM'99–F'09) received the M.D. degree in computer science engineering, the M.D. degree in biomedical engineering, and the Ph.D. degree in electrical engineering all from Katholieke Universiteit Leuven, Leuven, Belgium, in June 1981, July 1985, and June 1987, respectively.

She is currently a Full Professor at the Department of electrical Engineering, KU Leuven. Her research interests include numerical (multi)linear algebra and software, system identification, parameter estimation, and biomedical data processing.