Circuits, Systems, and Signal Processing (2024) 43:1862–1888
https://doi.org/10.1007/s00034-023-02543-8

Bayesian Structured-Sparse Modeling Using a Bernoulli–Laplacian Prior

Mayadeh Kouti[1,2] · **Karim Ansari-Asl**[1] · **Ehsan Namjoo**[1]

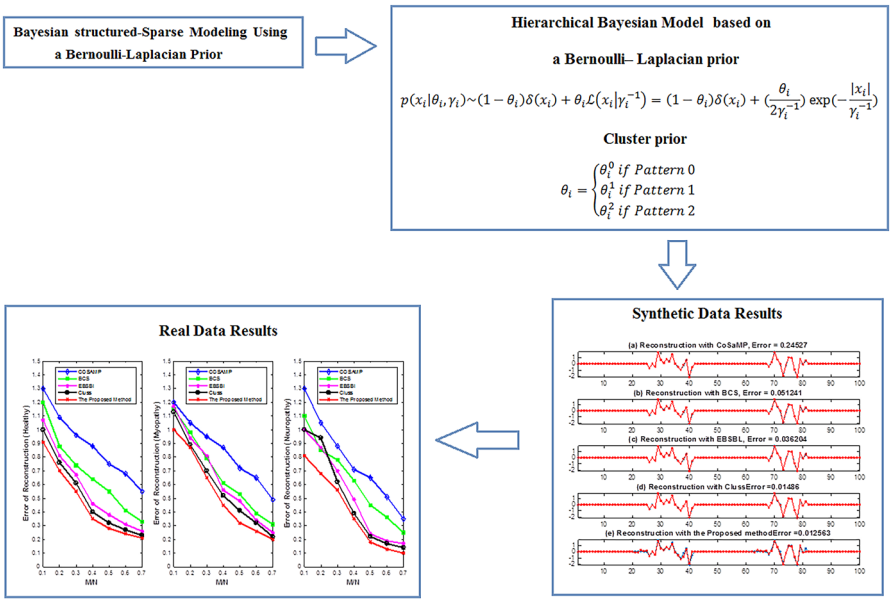
Received: 19 August 2023 / Revised: 13 October 2023 / Accepted: 14 October 2023 /
Published online: 22 November 2023

Abstract

Recently, sparse signal and image recovery have shown significant promise in different biomedical fields. In this paper, we introduce a novel method to recover structured-sparse biomedical signals and images in a hierarchical Bayesian framework. The proposed method promoted sparse distribution using a Bernoulli–Laplace prior. In addition to sparse prior, we consider cluster prior on sparsity patterns. To implement the Bayesian inference, we use an MCMC technique to sample the target posterior distribution. Using generated samples, the model parameters and hyperparameters are estimated in an unsupervised scheme. Finally, the estimation procedure has been completed using the MAP estimator. Our method solves the inverse problem automatically without needing to alter the parameters manually. We have used synthetic data sets with several sparsity scenarios to explore the proposed algorithm, which outperforms the existing recovery methods, *e.g.*, CoSaMP, BCS, EBSBL, and Cluss. Finally, through experiments using different real-life biomedical data (EMG and ECG signals and MR images), the superiority of the proposed method is demonstrated. This study demonstrates that using a combination of Bernoulli–Laplace and structured prior on sparsity patterns can recover structured-sparse biomedical signals and images precisely.

Extended author information available on the last page of the article

Graphical Abstract



Keywords Sparse Bayesian reconstruction · $L_0 + L_1$ Regularization · Block sparse signal · Hierarchical Bayesian model · MCMC methods

1 Introduction

Sparse signal and image recovery have received a great deal of interest in recent years. These approaches assume that signals are sparse in the original domain or a transformed domain, e.g., wavelet transform. The sparse theory is used to constrain the solution, which stabilizes the inverse problem. This idea has been developed in many previous works like Basis Pursuit (BP) [4, 10], orthogonal matching pursuit (OMP) [43], compressive sampling matching pursuit (CoSaMP) [33], iterative hard threshold (IHT) algorithm [2], and Bayesian compressive sensing [27, 41]. Besides sparsity, for modeling some specific signals, one may use structured–sparsity as a priori to enhance the recovery ability. This property allows the nonzero coefficients in sparse signals to be appeared only in clustered blocks. It is shown that structured sparsity can improve results in different applications such as machine learning [25], image analysis [49], audio restoration [29], biomedical signal recovery, e.g., some EMG and ECG signals are block-sparse in the origin domain [15, 37] and EEG source localization.

Several studies have proposed algorithms that use the block-sparsity concept, including greedy pursuit algorithms such as Model-CoSaMp [1], Block-OMP [18], and regularized convex optimizations algorithms like Group Lasso algorithm [30, 52],

Mixed L_1/L_2 Program [17, 35], Group Basis Pursuit [44], Group Orthogonal Matching Pursuit [42, 54], and block-sparse Bayesian learning (BSBL) [48, 53]. It has been theoretically and empirically proved that employing the block-sparsity concept can highly promote the recovery performance of sparse signals [1, 17]. Best performances are accomplished when the supposed structure exactly matches the data. However, in real applications, the block structure is not known. To solve this problem, one can use Bayesian methods which are standard statistical tools to solve inference problems and use different priors to find the block partitions in a probabilistically manner [12, 13, 19, 22, 23, 28, 38, 40, 46, 50, 51]. Indeed, it deals with uncertainties and uses our knowledge, under the form of priors, to solve a given inverse problem. The posterior distribution of the parameters can be calculated by combining the likelihood and the sparse priors.

To this end, a new hierarchical Bayesian model is proposed to.

In this paper, we propose a novel hierarchical Bayesian model based on a Bernoulli–Laplacian prior and structured sparsity. The proposed method can be considered as an extension of the hierarchical Bayesian modeling proposed in [50], which uses the Bernoulli–Laplacian prior introduced in [7]. The principal contribution of this paper is to promote sparsity for source recovery via a structured Bernoulli–Laplace prior. The proposed method solves the inverse problem automatically, while the prior information of the number of clusters and their sizes is unknown. The Bernoulli part guarantees the sources' sparseness, and the Laplacian prior promotes the flexibility of regularization to enforce the sparseness of signal through solving the inverse problem [8, 11, 32, 39]. More specifically, the Laplacian prior can be the Bayesian equivalent of L_1 norm, while a Bernoulli prior is equivalent to the L_0 norm [32]. The L_0 norm handles the nonzero position of coefficients, while the L_1 norm regularizes their amplitude values. In [7, 11, 21], authors suggested combining these norms in a Bayesian framework, while the L_0 penalty was used to find the nonzero positions and the L_1 norm estimated their amplitudes. To promote source recovery, we use the block-structured property of sparse signals, of which the nonzero coefficients appear in blocks. To this end, we consider three kinds of block patterns introduced in [50] to impose sources in a probabilistic framework. Then, a hierarchical Bayesian framework is used to drive posterior distribution that imposes sparse prior as well as the structured sparsity prior at the same time. Our hierarchical Bayesian framework uses hyperparameters to support the sparse coefficients and encourage block patterns. Besides, the method can drive an unsupervised estimation of hyperparameters, *e.g.*, there is no need for any sparse and block prior information. Unfortunately, it is often not possible to achieve closed-form expressions for the Bayesian estimators of the posterior distribution parameters. To address this problem, several methods are available to estimate the posterior parameters. One of these methods is the Markov chain Monte Carlo (MCMC) method [36]. The method is based on a Gibbs sampler, which is applied to draw asymptotically distributed samples based on the posterior of the proposed Bayesian model. Then, the generated samples are employed to estimate the hyperparameters of the Bayesian model in an unsupervised framework.

The performance of the proposed method is assessed using both synthetic and real data, such as EMG and ECG signals and MR images. Results show that in comparison with state-of-the-art methods, e.g., CoSaMP [1], BCS [27], EBSBL [53], and Cluss

[50], our algorithm has advantages, specifically when applied to structured-sparse signals.

The remainder of this text is structured as follows. The second section introduces and formulates the inverse problem. In Sect. 3, the suggested hierarchical Bayesian model is presented. In Sect. 4, posterior calculation using the MCMC method is presented. Results of the simulation and real-world datasets are demonstrated in Sect. 5. In Sect. 6, discussion is presented. Lastly, conclusions and perspectives are drawn in Sect. 7.

2 Sparse Regularization of Inverse Problems

In many generic inverse problems, we want to infer unknowns through observations related between them through an operator. When this relation is linear, we arrive in the relation:

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n} \tag{1}$$

where $\boldsymbol{x} = [x_1, ..., x_N]$ denotes the unknowns, $\boldsymbol{y} = [y_1, ..., y_M]$ the observations, $\boldsymbol{n} = [n_1, ..., n_M]$ the measurement error, and $\boldsymbol{H}$ the matrix of the system response. If the noise could be ignored ($\boldsymbol{n} = \boldsymbol{0}$) and the matrix $\boldsymbol{H}$ invertible($M = N$), the solution $\boldsymbol{x} = \boldsymbol{H}^{-1}\boldsymbol{y}$ is not the best solution, since this solution may be too sensitive to small changes due to the ill-conditioning of this matrix. In the general case of$M < N$, a regularized solution is obtained by defining the optimization problem, including two parts of conditions:

$$\widehat{\boldsymbol{x}} = \boldsymbol{argmin}_{\boldsymbol{x}} \left\{ \boldsymbol{J}(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|_2^2 + \lambda \|\boldsymbol{x}\|_{\boldsymbol{p}}^{\boldsymbol{p}} \right\} \tag{2}$$

where $\|.\|_p^p$ is L_p norm and λ is a regularization parameter which is used to manage a trade-off between data fit and a priori knowledge. In the case of$p = 2$, the L_2 norm is defined as $\|\boldsymbol{x}\|_2^2 = \sum_i |x_i|^2$ and$\widehat{\boldsymbol{x}}$ is obtained as$[\boldsymbol{H}^{\boldsymbol{T}}\boldsymbol{H} + \lambda \boldsymbol{I}]^{-1}\boldsymbol{H}^{\boldsymbol{T}}\boldsymbol{y}$. During the past decades, different regularization terms have been proposed. Particularly, some studies proposed to use the L_0 norm in the regularization term,$L_0(\boldsymbol{x}) = \|\boldsymbol{x}\|_0 = \sum_i \delta(x_i)$, which corresponds to IHT algorithm [2]. Also, the L_1 norm,$L_1(\boldsymbol{x}) = \|\boldsymbol{x}\|_1 = \sum_i |x_i|$, is used to enforce the sparsity of the solution in the place of L_2 norm. This norm corresponds to the LASSO formula problem (IST [14], BPDN, etc.). Furthermore, the parameter λ balances the observation and the sparse prior. If $p \in (0, 1)$, the solution corresponds to the iterative reweighted algorithm [9].

The problem solution of Eq. 1 can be defined as the optimization problem with two terms, which can be integrated to a maximum a posteriori (MAP) solution in a Bayesian scheme. The first term of Eq. 2 corresponds to the likelihood, and the second term is related to a prior model. Once the noise is presumed to be Gaussian, then the MAP strategy can drive the optimization of the cost function $(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|_2^2 + \lambda\varphi(\boldsymbol{x})$, where $\varphi(\boldsymbol{x})$ is related to the prior law. When the prior information is interpreted as a Gaussian probability law, then $(\boldsymbol{x}) = \|\boldsymbol{x}\|_2^2$, when it is understood as a Laplace probability law, then$\varphi(\boldsymbol{x}) = \|\boldsymbol{x}\|_1$, and finally, if it is considered as a Bernoulli probability law, then $\varphi(\boldsymbol{x}) = \|\boldsymbol{x}\|_0$ [31, 32, 34, 47].

3 Hierarchical Bayesian Model

The Bayesian approach has many advantages in handling the hyperparameters in the regularization approach. Here, we suggest a framework for sparse Bayesian regularization using a mixed sparse prior to estimate signals $\boldsymbol{x}$ from the observation vector $\boldsymbol{y}$. Specifically, we approximate a mixed norm regularization with a multivariate Bernoulli–Laplacian prior [11] based on the Bayesian model. To solve this linear observation problem, observation likelihood and priors are described as follows:

3.1 Observation Likelihood

First of all, we suppose that signals are white Gaussian noises with zero mean and variance σ_0^2. The likelihood can be expressed as:

$$p\left(\boldsymbol{y}|\boldsymbol{x},\sigma_0^2\right) \sim \mathcal{N}\left(\boldsymbol{H}\boldsymbol{x},\sigma_0^2\boldsymbol{I}\right) = \left(\frac{1}{2\pi\sigma_0^2}\right)^{\frac{M}{2}} \exp\left(-\frac{||\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}_2^2||}{2\sigma_0^2}\right) \tag{3}$$

The only available knowledge about noise variance is the positivity of this parameter. To infer the conditional posterior distribution of the noise variance σ_0^2, we assume $(\sigma_0^2)^{-1} = \xi$ which has Gamma prior as follows:

$$p(\xi|\alpha,\beta) \sim \text{Gamma}(\alpha,\beta) \tag{4}$$

where α and β are fixed hyperparameters.

3.2 Sparse Priors

The unknown parameter $\boldsymbol{x}$ is denoted by a prior distribution$p(\boldsymbol{x})$, which enforces sparseness on the$\boldsymbol{x}$. Here sparsity prior is modeled by adapting a Bernoulli–Laplacian prior to regularize the nonzero part in a better way and get the zero coefficients of the signal efficiently. So for every $x_i\,(\text{i} = 1,\ldots,N)$, the prior distribution is as follows:

$$p(x_i|\theta_i,\gamma_i) \sim (1-\theta_i)\delta(x_i) + \theta_i\mathcal{L}\left(x_i|\gamma_i^{-1}\right) = (1-\theta_i)\delta(x_i) + \left(\frac{\theta_i}{2\gamma_i^{-1}}\right)\exp\left(-\frac{|x_i|}{\gamma_i^{-1}}\right) \tag{5}$$

where $\gamma_i > 0$ is the parameter of scaling, $\delta(.)$ is the Dirac delta function and $\theta_i \epsilon [0,1]$ is a mixing weight. The mixing weight is the prior probability of nonzero parameters. For the large values of θ_i, a nonzero part has a large probability, whereas the small weights produce a zero entry. The Bernoulli–Laplacian model developed a two-level structured sparsity equivalent to an $L_0 + L_1$ norm. First, the Bernoulli model and the Dirac delta function (equivalent to L_0 norm) enforce sparseness by guaranteeing the solution has few nonzero coefficients. Second, the Laplacian model distribution constrains their amplitude values since it is equivalent to the L_1 norm. So this hybrid

distribution (L_0+L_1) presents more flexibility and supports more signals with different sparseness. In addition to $\boldsymbol{\theta}$, a Gamma prior is considered for the hyperparameter γ_i as follows to obtain a posterior density function:

$$\gamma_i|_{\vartheta,\omega} \sim Gamma(\vartheta, \omega) \tag{6}$$

For vector $\boldsymbol{x}$, a prior is assigned as $p(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{\gamma})$. Considering the independency between x_i (the signal components), the prior distribution of $\boldsymbol{x}$ can be written as:

$$p(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{\gamma}) = \prod_{i=1}^{N}\left[(1-\theta_i)\delta(x_i) + (\frac{\theta_i}{2\gamma_i^{-1}})\exp(-\frac{|x_i|}{\gamma_i^{-1}})\right] \tag{7}$$

3.3 Cluster Prior

In addition to sparse prior, structures on the sparse pattern of the signal have also been investigated. In this paper, we used the cluster-structured-sparse signals introduced in [50], which assumes three cluster patterns for sparseness to model the cluster prior of the unknown $\boldsymbol{x}$. To consider different relations between the current coefficient x_i and its instant neighbors, the cluster structure is presented by three different patterns. Regard $\complement_i^{(k)} = \{(j|D(i,j) \le k, i \ne j)\}$ as the k-th neighborhood of location i over the unknown parameters $\boldsymbol{x}$, where $D(i,j)$ is the Euclidean distance between i and j positions on the vector $\boldsymbol{x}$. Consider $\pounds_N = \{1, \ldots, N\}$ the set of all locations over the unknown parameters $\boldsymbol{x}$ then define $\mathcal{J}_{i,k,\otimes} \triangleq \complement_i^k \cap \pounds_N$ and $\mathcal{J}_{i,k,\odot} \triangleq \complement_i^k \cap \pounds_N \cup \{i\}$. Therefore, we assign $x_{\mathcal{J}_{i,k,\otimes}}$ the set of components placed in the ith coefficient x_i the neighborhood, while $\mathcal{J}_{i,k,\odot}$ is the set of components comprising both the current component and neighbors. Based on these definitions, the categorization of different *block patterns* is as follows*:*

Pattern 0: "strongly eliminate", when $\|x_{\mathcal{J}_{i,k,\otimes}}\|_0 = 0$, presents all neighbors are zero for x_i. This pattern relates to the isolated points.

Pattern 1: "weakly eliminate", when $0 < \|x_{\mathcal{J}_{i,k,\otimes}}\|_0 < \left|\mathcal{J}_{i,k,\otimes}\right|$, shows part of neighbors are nonzero.

Pattern 2: "strongly plump", when $\|x_{\mathcal{J}_{i,k,\otimes}}\|_0 = \left|\mathcal{J}_{i,k,\otimes}\right|$, denotes "all neighbors are nonzero.

Where $|\mathcal{J}|$ denotes the cardinality of the set $\mathcal{J}$. Then the mixing weight θ_i can be selected by the followings:

$$\theta_i = \begin{cases} \theta_i^0 \text{ if Pattern 0} \\ \theta_i^1 \text{ if Pattern 1} \\ \theta_i^2 \text{ if Pattern 2} \end{cases} \tag{8}$$

where θ_i^0,θ_i^1 and θ_i^2 have different Beta distribution, i.e., $\theta_i^{<q>} \sim Beta(\varepsilon^{<q>}, \tau^{<q>})$ and $q \in \{0, 1, 2\}$, corresponding to pattern 0,1, and 2, respectively. Through assigning suitable value for parameters $\boldsymbol{\varepsilon}$,$\boldsymbol{\tau}$ the block pattern selection method improves the blocks and contains the isolates.

4 Markov Chain Monte Carlo

In this section, the Markov Chain Monte Carlo (MCMC) strategy [36] has been used to compute the Bayesian inference of our model. Using a Gibbs sampler, this technique samples the posterior and estimates the model parameters from the generated samples. If we assign the unknown parameters as $\mathcal{Q} \triangleq \{\boldsymbol{x}, \xi, \boldsymbol{\theta}, \boldsymbol{\gamma}\}$ and the model parameters as $\mathcal{K} \triangleq \{\alpha, \beta, \vartheta, \omega, \boldsymbol{\varepsilon}, \boldsymbol{\tau}\}$, the joint posterior distribution of $\mathcal{Q}$ can be computed as $p(\mathcal{Q}|\boldsymbol{y}, \mathcal{K}) \propto p(\boldsymbol{y}|\mathcal{Q})p(\mathcal{Q}|\mathcal{K})$. Using the hierarchical model, $p(\mathcal{Q}|\boldsymbol{y}, \mathcal{K})$ can be written as follows:

$$\boldsymbol{p}(\mathcal{Q}|\boldsymbol{y}, \mathcal{K}) = p(\boldsymbol{x}|\boldsymbol{\gamma}, \boldsymbol{\theta}, \xi, \boldsymbol{y})p(\boldsymbol{\gamma}|\boldsymbol{x}, \vartheta, \omega)p(\boldsymbol{\theta}|\boldsymbol{x}, \boldsymbol{\varepsilon}, \boldsymbol{\tau})p(\xi|\boldsymbol{x}, \boldsymbol{y}, \alpha, \beta) \tag{9}$$

4.1 Posterior Distributions

4.1.1 Noise Variance σ_0^2

As mentioned earlier,$(\sigma_0^2)^{-1} = \xi$ has gamma distribution, so by joining the prior and the likelihood, the posterior distribution can be easily written as:

$$p(\xi|x, y) = Gamma\left(\alpha + \frac{M}{2}, \beta + \frac{\|y - Hx\|_2^2}{2}\right) \tag{10}$$

It also has an inverse gamma distribution, which is easy to sample.

4.1.2 Mixing Weight $\boldsymbol{\theta}$

As mentioned before, each element θ_i of the mixing weight $\boldsymbol{\theta}$ is produced by choosing from three parameters based on its sparsity pattern. Straightforward calculations show the hyperparameter $\theta_i^{<q>}$ with $q \in \{0, 1, 2\}$ has a Beta posterior distribution, which is conjugate to the Bernoulli distribution and chosen due to its helpful mathematical properties. It can be expressed as:

$$p\left(\theta_i^{<q>}|x_{\mathcal{J}_{i,k,\odot}}\right) = Beta\left(\varepsilon^{<q>} + \|x_{\mathcal{J}_{i,k,\odot}}\|_0, \tau^{<q>} + \left|\mathcal{J}_{i,k,\odot}\right| - \|\mathcal{J}_{i,k,\odot}\|_0\right) \tag{11}$$

4.1.3 Hyperparameter $\boldsymbol{\gamma}$ of the Sparse Model

Considering the likelihood $p(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{\gamma})$ and the prior distribution of $\boldsymbol{\gamma}$, one can compute the conditional distribution of the hyperparameter γ of the sparse model as follows:

$$p\left(\gamma_i|\mathcal{J}_{i,k,\odot}, \vartheta, \omega\right) = Gamma\left(\vartheta + \|x_{\mathcal{J}_{i,k,\odot}}\|_0, \omega + \|x_{\mathcal{J}_{i,k,\odot}}\|_1\right) \tag{12}$$

4.1.4 Sparse Signal $\boldsymbol{x}$

Considering the observation, the likelihood, and the prior distribution of x_i, one can obtain the posterior distribution of $\boldsymbol{x}$. Since the prior used for $\boldsymbol{x}$ is not simple, sampling of $p(x_i|\boldsymbol{x}_{-i}, \boldsymbol{y}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \xi)$ is not straightforward. To drive the posterior distribution of $\boldsymbol{x}$, we use the approach given in [16] as follows:

$$p(\boldsymbol{x_i}|\boldsymbol{x_{-i}}, \boldsymbol{y}, \theta, \gamma, \xi) = \theta_{1,i}\delta(x_i) + \theta_{2,i}\mathcal{N}^{+}(\mu_{i+}, \sigma_i^2) + \theta_{3,i}\mathcal{N}^{-}(\mu_i^{-}, \sigma_i^2) \tag{13}$$

where $\mathcal{N}^{+}$ and $\mathcal{N}^{-}$ show the distributions of truncated Gaussian on $\mathbb{R}^{+}$ and $\mathbb{R}^{-}$, respectively. Note that for designing a Gibbs sampler for x_i, we assign $\boldsymbol{x}_{-i}$ as the vector encompassing all but the ith component. Based on the [16], $\mathcal{W} = \{w_1, \ldots, w_N\}$ is the orthonormal basis of $\boldsymbol{x}$, such that $\boldsymbol{x} = \widetilde{\boldsymbol{x}}_{-i} + x_i w_i$ and $\widetilde{\boldsymbol{x}}_{-i}$ is achieved, when ith element of $\boldsymbol{x}$ is set to zero. By defining $\boldsymbol{\lambda}_i = \boldsymbol{y} - \boldsymbol{H}\widetilde{\boldsymbol{x}}_{-i}$ and $\boldsymbol{h}_i = \boldsymbol{H}w_i$, the weights $(\theta_{i,l}^{<q>})_{1\le l\le 3}$ can be written as:

$$\theta_{i,l}^{<q>} = \frac{u_{l,i}}{\sum_{l=1}^{3} u_{l,i}} \tag{14}$$

where

$$\begin{aligned}
u_{l,i} =& 1 - \theta_i^{\langle q\rangle} \\
u_{2,i} =& \frac{\theta_i^{\langle q\rangle}}{2\gamma_i^{-1}} \exp\left(\frac{\mu_{i+}^2}{2\sigma_i^2}\right)\sqrt{2\pi\sigma_i^2}\, C\left(\mu_{i+}, \sigma_i^2\right) \\
u_{3,i} =& \frac{\theta_i^{\langle q\rangle}}{2\gamma_i^{-1}} \exp\left(\frac{\mu_{i-}^2}{2\sigma_i^2}\right)\sqrt{2\pi\sigma_i^2}\, C\left(-\mu_{i-}, \sigma_i^2\right)
\end{aligned} \tag{15}$$

and

$$\begin{aligned}
\sigma_i^2 =& \frac{\sigma_0^2}{\boldsymbol{h}_i^2} \\
\mu_{i+} =& \sigma_i^2\left(\frac{\boldsymbol{h}_i^T\boldsymbol{\lambda}_i}{\sigma_0^2} - \frac{1}{\gamma_i^{-1}}\right) \\
\mu_{i-} =& \sigma_i^2\left(\frac{\boldsymbol{h}_i^T\boldsymbol{\lambda}_i}{\sigma_0^2} + \frac{1}{\gamma_i^{-1}}\right) \\
C\left(\mu, \sigma^2\right) =& \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{\mu}{\sqrt{2\sigma^2}}\right)\right)
\end{aligned} \tag{16}$$

4.2 Gibbs Sampling

The MCMC method gives the Bayesian estimation of the unknown model parameters. Here it uses the Gibbs sampler to produce samples from the joint posterior distributions of Eq. 9. The parameters are jointly estimated in an unsupervised framework. Details of Gibbs sampler are explained as follows:

Initialize with $x^{(0)}$

Repeat

Sample ξ from $p(\xi|\boldsymbol{x}, \boldsymbol{y}, \alpha, \beta)$

For $i = 1, \ldots, N$

Sample x_i from $p(x_i|\boldsymbol{x}_{-i}, \boldsymbol{y}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \xi)$

Sample γ_i from $p(\gamma_i|\boldsymbol{x}, \vartheta, \omega)$

Sample θ_i from $p(\theta_i|\boldsymbol{x}, \boldsymbol{\varepsilon}, \boldsymbol{\tau})$

end for

Until convergence

Algorithm 1. Gibbs sampler

Using the Gibbs sampler, a set of $\mathcal{Q}$, which is asymptotically distributed based on the joint posterior of Eq. 9, is generated. $\mathcal{Q} = \{\mathcal{Q}(j)\}_{j=1,\ldots,t_{Ni},\ldots,t_{MC}}$ is the set of this collection and j is the MCMC sampling steps. Note that t_{Ni} and t_{MC} are the number of burning iterations of the sampler and MCMC iterations, respectively. In the next step, for inferring the $\boldsymbol{x}$ estimation, the MAP estimator is used. The marginal distribution $p(\boldsymbol{x}|\boldsymbol{y})$ can be obtained by considering the full posterior (Eq. 9) and using the integration of the hyperparameters $\boldsymbol{\gamma}$, $\boldsymbol{\theta}$ and ξ:

$$p(\boldsymbol{x}|\boldsymbol{y}) \propto \int p(\boldsymbol{x}|\boldsymbol{\gamma}, \boldsymbol{\theta}, \xi, \boldsymbol{y}) d\boldsymbol{\gamma} d\boldsymbol{\theta} d\boldsymbol{\xi} \tag{17}$$

Hence, by maximizing the posterior (17), the MAP estimator of $\boldsymbol{x}$ can be driven as:

$$\widehat{\boldsymbol{x}} \approx argmax p(\boldsymbol{x}|\boldsymbol{y})_{\boldsymbol{x} \in \mathbf{X}} \tag{18}$$

where $\mathrm{X} = \{\boldsymbol{x}(j)\}_{j=t_{Ni},\ldots,t_{MC}}$

5 Experiments

In this section, we conduct two experiments with synthetic and real data to validate the proposed method. The simulation is performed in MATLAB software. For real data, we use EMG and ECG signals taken from the PhysioBank database [20]. The signals

were compressed with observation operator $\boldsymbol{H}$, which is constructed randomly as in [6], i.e., $\boldsymbol{H}$ is adapted to a standard Gaussian distribution with $\mathcal{N}(0, \frac{1}{\sqrt{M}})$. Besides, results are assessed in terms of reconstruction error (RE) and signal-to-noise ratio (SNR) defined as follows:

$$RE = \frac{\|\widehat{x} - x\|_2}{\|x\|_2} \tag{19}$$

$$SNR = 20\log_{10}\frac{\|x\|_2}{\|\widehat{x} - x\|_2} \tag{20}$$

where $\widehat{\boldsymbol{x}}$ and $\boldsymbol{x}$ are the estimated and the true signals, respectively. For image quality evaluation, we use the peak signal-to-noise ratio (PSNR) which is the most commonly used quality assessment technique to measure the quality of reconstruction of image compression methods. PSNR is expressed as:

$$PSNR = 10\log_{10}\frac{(\text{peakval})^2}{MSE} \tag{21}$$

Here, $peakval$ is the maximal in the image data and MSE (mean square error) is computed by averaging the squared intensity of error between the original and the resultant image pixels. If the image is an 8-bit unsigned integer data type, the peakval is 255 [26]. Also, we compare the proposed method performance with state-of-the-art algorithms, including CoSaMP [1], BCS [27], EBSBL [53], and Cluss [50]. We set the model hyperparameters for the priors as follows $= \beta = \vartheta = \omega = 10^{-6}$, $(\varepsilon^{<0>}, \tau^{<0>}) = (\frac{1}{M}, 1 - \frac{1}{M}) \times \left|\mathcal{J}_{i,k,\odot}\right|$,$(\varepsilon^{<1>}, \tau^{<1>}) = (\frac{1}{M}, \frac{1}{M}) \times \left|\mathcal{J}_{i,k,\odot}\right|$, $(\varepsilon^{<2>}, \tau^{<2>}) = (\frac{1}{M}, 1 - \frac{1}{M}) \times \left|\mathcal{J}_{i,k,\odot}\right|$,where$k = 1$. Also, the initial conditions are fixed to$\theta_i(0) = 0$, $\gamma_i(0) = 1$ and $\sigma_0^2(0) = var(\boldsymbol{y}) \times 10^2$ for all$i \in \pounds_N$.

5.1 Synthetic Data

5.1.1 Signal Recovery Performance

In this section, we generate an artificial signal $\boldsymbol{x}$ with block-sparse nature and length $N = 100$. To this end, we consider three blocks with maximum size block 30 (total number of nonzero values set to 30) and $M = 60$. The observation $\boldsymbol{y}$ is obtained by adding a Gaussian noise n of variance $\sigma_n^2 = 10^{-2}$. Figure 1a and b shows the original (reference) and the observation signals, respectively. Then, we apply the proposed algorithm detailed in Sects. 3 and 4. Figure 2 demonstrates the reconstruction results and errors of the proposed method compared to existing algorithms. Quantitatively speaking, as displayed in Fig. 2, the proposed method has the lowest reconstruction error since the markers are closer to the reference signal. Besides, one may note that the proposed method guarantees a better sparsity reconstruction due to the combination of two sparse priors (Bernoulli–Laplace distribution), which is equivalent to the sum of L_0 and L_1.

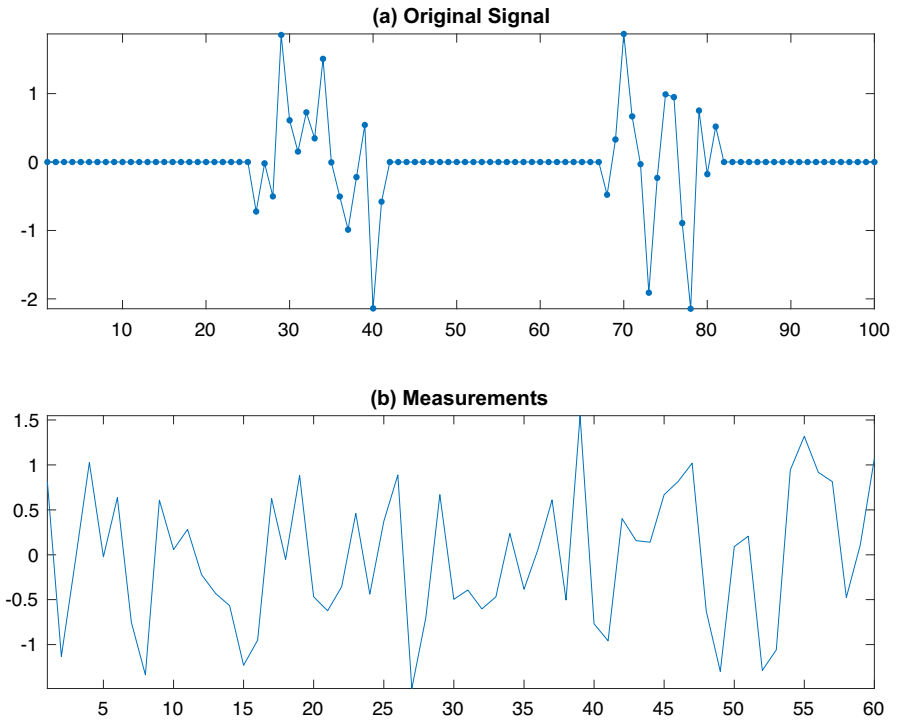


Fig. 1 **a** Original sparse signal with length $N = 100$. **b** measurements with length $M = 60$

5.1.2 Effect of the $\frac{M}{N}$ Ratio

We set another experiment that evaluates the effect of the $\frac{M}{N}$ ratio on the performance of signal recovery. We generate a sparse signal with length $N = 100$, the total number of nonzero element $S = 20$, two blocks ($K = 2$), and M observations varied from 10 to 70. The experiment is repeated 100 times with different observation operator $\boldsymbol{H}$, noise $\boldsymbol{n}$, and signal $\boldsymbol{x}$. Figure 3 shows the reconstruction error of the proposed algorithm compared to CoSaMP, BCS, EBSBL, and Cluss for varying $\frac{M}{N}$ ratios. It can be seen that the proposed approach outperforms existing methods. For example, in $\frac{M}{N} = 0.6$, the reconstruction error of the proposed method is 0.012,while it yields 0.262, 0.053,0.037, and 0.016 for CoSaMP, BCS, EBSBL, and Cluss, respectively. By increasing the $\frac{M}{N}$ ratio, a performance enhancement in all algorithms can be observed. Besides, the proposed approach outperforms previous methods.

5.1.3 Effect of Noise

In this experiment, we assess the effect of noise on the proposed method. We use a sparse signal with length $N = 100$, the total number of nonzero element S = 20, two blocks ($K = 2$), and $M = 50$ observations. Then, the observations are contaminated

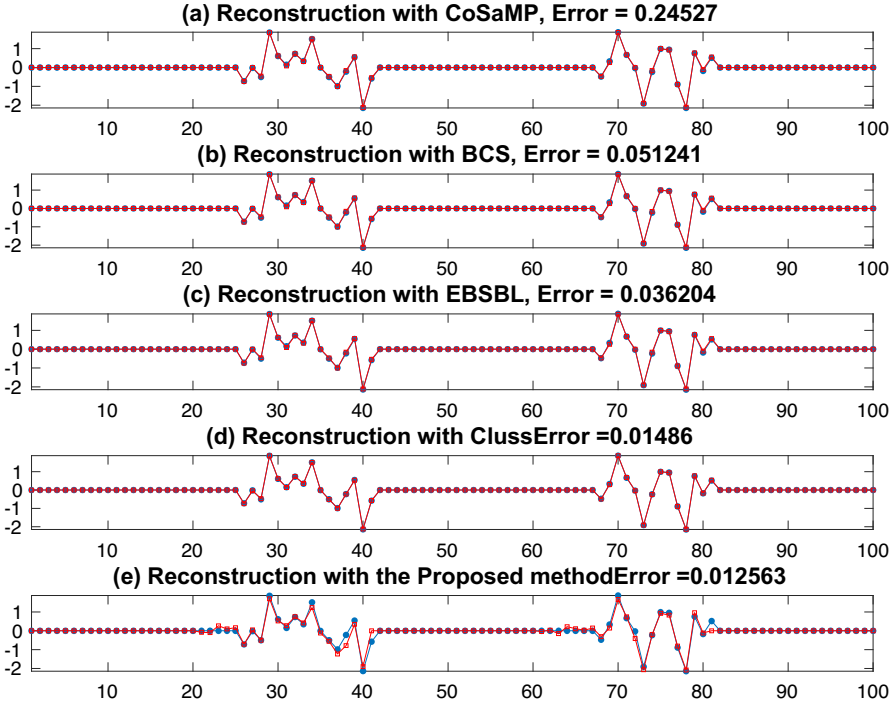


Fig. 2 Reconstruction results and errors in $\frac{M}{N} = 0.6$ of **a** CoSaMP, **b** BCS, **c** EBSBL, **d** Cluss, and **e** the proposed algorithms. The blue ones are the original signals, and the red ones are the reconstructed signals (Color figure online)

by adding a Gaussian noise n of variance $\sigma_n^2 = 0.01$ to 0.15, namely the SNR ranging from 36 to 12 dB. In the next step, the proposed method and the previous methods are employed to recover the sparse signal. The experiment is repeated 100 times with different realizations of observation operator $\boldsymbol{H}$, noise $\boldsymbol{n}$, and signal $\boldsymbol{x}$. Simulation results with different SNRs are displayed in Fig. 4. It can be seen that all reconstruction errors decrease with increasing the noise variance (or decreasing SNR). Furthermore, the proposed algorithm shows better performance in recovering the original sparse signal with achieving the highest SNR of recovery.

5.2 EMG Signals

In this section, the EMG signals obtained from the *PhysioBank* database [20] are chosen as a sample of the block-sparse signals. Data were acquired with a Medelec Synergy N2 EMG Monitoring System (Oxford Instruments Medical, OldWoking, UK), recorded at 50 kHz and before being downsampled to 4 kHz. Then two analog filters were employed: a 20-Hz highpass filter and a 5-kHz lowpass filter.

Figure 5 shows three examples of EMG signals with length $N = 1000$ from (top)

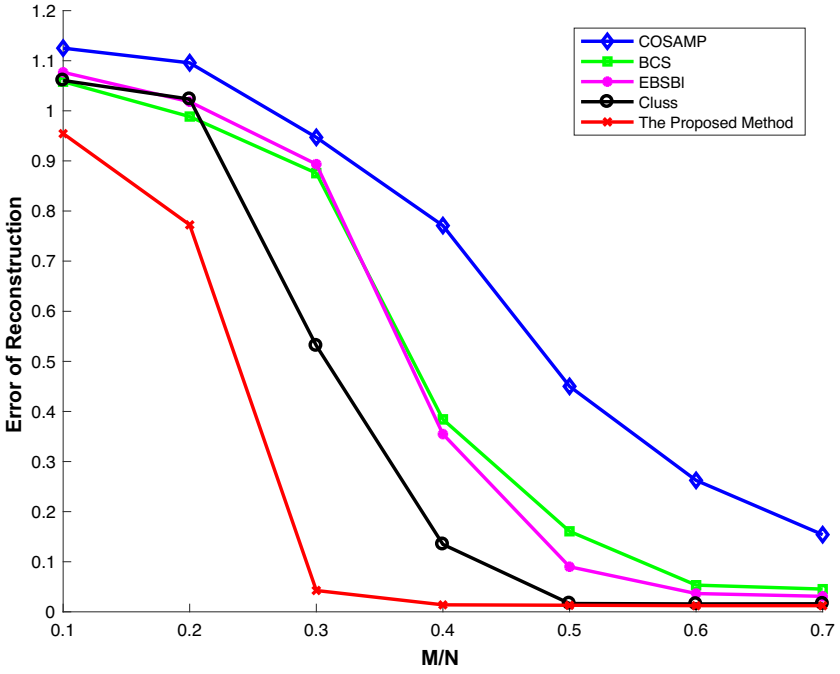


Fig. 3 Comparison of signal recovery performance with varying $\frac{M}{N}$ ratios

a healthy person, (middle) a patient with myopathy, and (bottom) a patient with neuropathy, respectively. It can be observed that all three signals are structured-sparse in the time domain. In this experiment, the recovery performance of the proposed method is evaluated in comparison with the aforementioned methods, e.g., CoSaMP, BCS, EBSBL, and Cluss, for three different EMG signals. The acquired EMG signals are divided into sections with a length of $N = 1000$. To this end, we divided the EMG signals into sections with length of $N = 1000$. Then, one section is selected randomly as the input signal. Following the method described at the beginning of Sect. 5, a sensing matrix can be constructed. Then, the measurements are captured by projecting each EMG signal on the sensing matrix based on Eq. 1.

For the first evaluation, we vary the number of measurements from $M = 100$ to $M = 700$. Figure 6 (from left to right) shows the reconstruction error of healthy, myopathy and neuropathy signals using the proposed method and the other comparing signal recovery approaches. We can see that for different $\frac{M}{N}$ ratios, the error values of the proposed method are lower than other comparing methods. The reconstruction errors of the aforementioned algorithms for each kind of EMG signals in $\frac{M}{N} = 0.6$ are given numerically in Table 1. It can be seen that by increasing the subsampling ratio $\frac{M}{N}$, all estimation errors decrease. As the errors of CosaMP are higher than the other four approaches, this method cannot be a suitable candidate for EMG signal recovery.

In the next step, we measure the average SNR of the recovered signals for different values of the SNR of measurements. For each value of the SNR of measurements, we

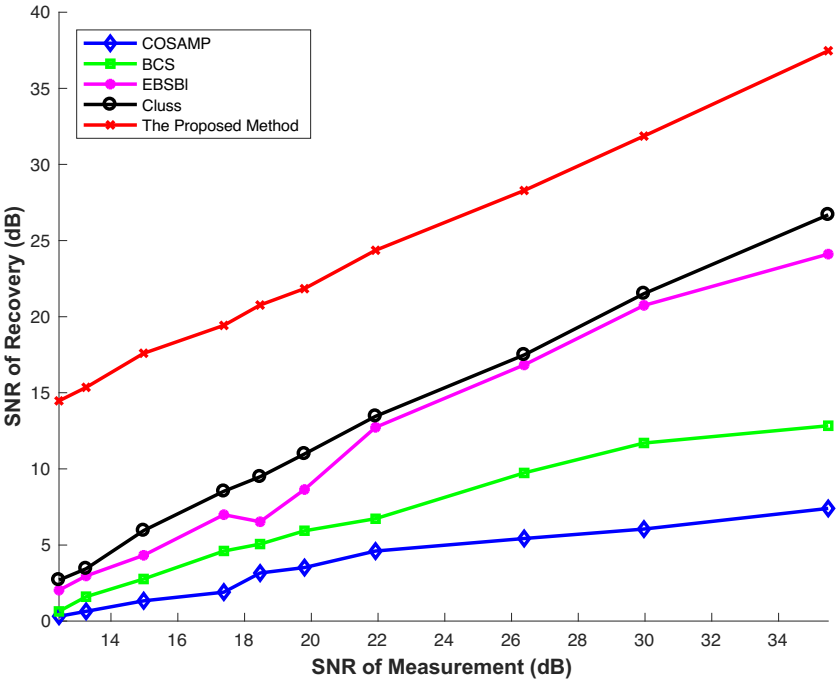


Fig. 4 Comparison of signal recovery performance with different levels of noise

repeated the experiment 100 times, and in each time, the sensing matrix was randomly generated [6]. The results are reported in Fig. 7. Note that for every input SNR, the proposed method yields the highest SNR among all the considered methods.

Figures 8, 9, 10 illustrate examples of the reconstruction of a section of healthy, myopathy, and neuropathy signals when the number of measurements equals to 450 ($M = 450$), respectively. It can be seen that the profile of the Cluss method is better than other compared methods except for the proposed method, which exhibits the best reconstruction performance.

5.3 ECG Signal

Another practical application of the proposed algorithm employs ECG real data set. The data are obtained from the *PhysioBank* database [3, 20]. The original signal is uniformly recorded from ninety-six hundred measurements in 1 h. The acquired ECG signal is divided into sections with a length of $N = 500$. Then, as the input signal, we selected one section randomly. The measurement matrix is generated by using this selected section and the sensing matrix based on Eq. 1. Figure 11 displays one section of the ECG signal. Then, the proposed method and the aforementioned algorithms are employed to implement the reconstruction procedure. Figure 12 demonstrates an example of the reconstruction performance of an ECG section for $M = 180$. According

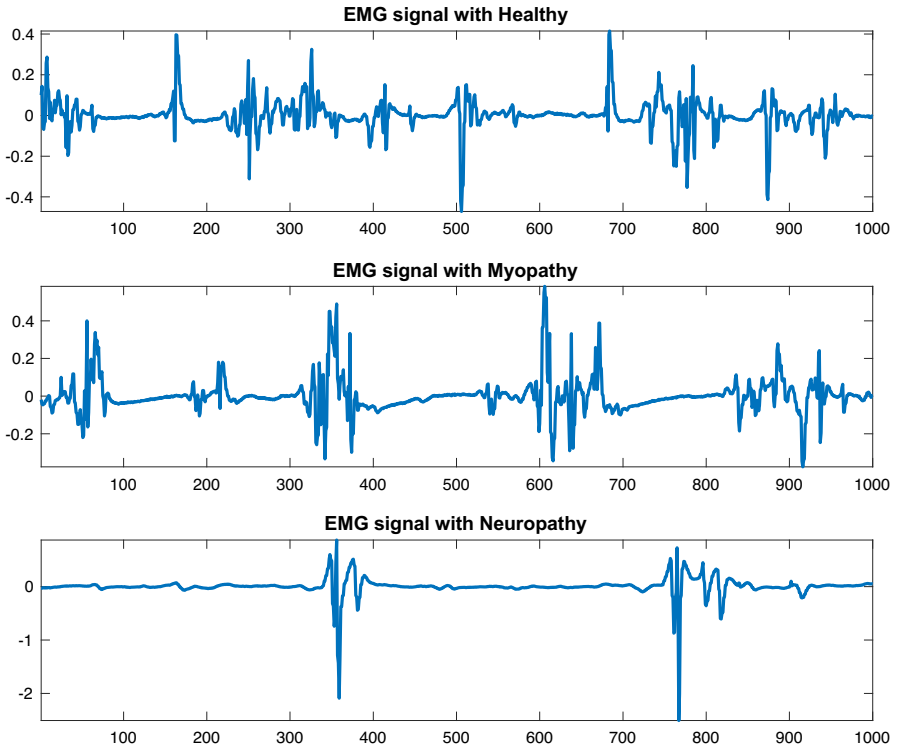


Fig. 5 Example of three kinds of EMG signals

to these results, the profile of the proposed method is found superior compared to the considered approaches.

Figure 13 presents the reconstruction errors obtained by the aforementioned sparse recovery algorithms with varying $\frac{M}{N}$ ratios in 100 trials. Comparing the other signal recovery methods, the proposed method achieves the lowest errors from $M = 50$ to $M = 350$. It can obviously be seen that the performance of CoSaMP is far worse than the other algorithms. For example, in $\frac{M}{N} = 0.6$, the reconstruction error of the proposed method is 0.151, while it yields 0. 396, 0.294, 0.237, and 0.213 for CoSaMP, BCS, EBSBL, and Cluss, respectively. One can see that by increasing the $\frac{M}{N}$ ratio, all estimation errors decrease. Figure 14 reports the SNR of the recovered signal, which is averaged over 100 independent trials, as a function of the SNR of measurements. The obtained results show that there is a meaningful difference in terms of SNR between the proposed algorithm and the existing methods. Depending on the SNR of measurements, the SNRs of recovered signals are improved. Moreover, it can be seen that the neuropathy signal is more compressible than the myopathy signal. With the same input SNR, we obtained a better SNR. This is due to the fact that the neuropathy signal has a better sparsity representation than the myopathy signal.

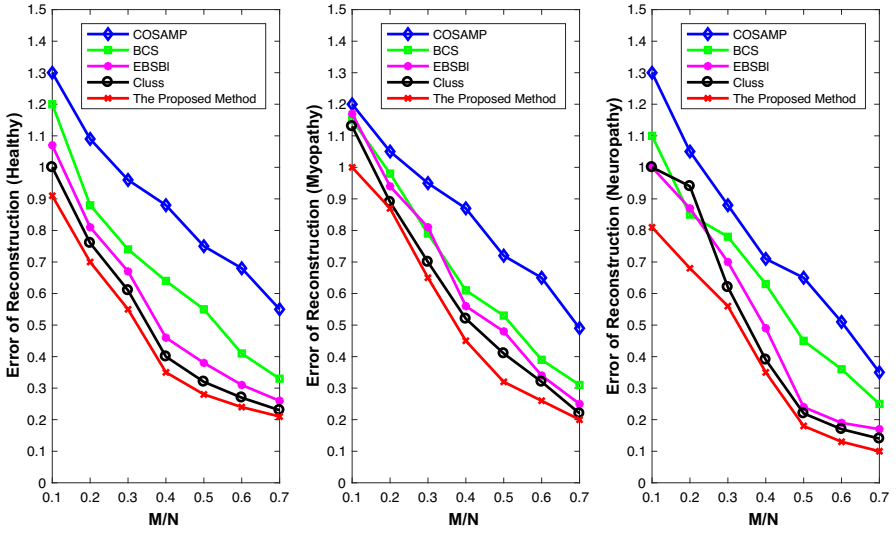


Fig. 6 Errors of reconstruction versus different $\frac{M}{N}$ ratios for three kinds of EMG signals, using CoSaMP, BCS, EBSBL, Cluss, and the proposed algorithms. From left to right, healthy, myopathy and neuropathy signals, respectively

Table 1 Errors of reconstruction in $\frac{M}{N} = 0.6$ for three kinds of EMG signals, using CoSaMP, BCS, EBSBL, Cluss, and the proposed algorithms

EMG signal	Method				
	CoSaMP	BCS	EBSBL	Cluss	The proposed method
Healthy	0.681	0.416	0.316	0.274	0.245
Myopathy	0.655	0.392	0.344	0.328	0.263
Neuropathy	0.515	0.363	0.197	0.173	0.134

5.4 MR Images

In this experiment, we selected four low-rank MRIs [45], which means they have many zero pixels rather than the nonzero pixels, with 100×100 pixels (Fig. 15). The number of measurements is set to $M = 3000$ (30% sampling). The performance of the proposed methods is compared with CoSaMP, BCS, EBSBL, and Cluss. The visual quality image recovery is illustrated in Fig. 15. For all the MRIs, the PSNR performances are also shown in Fig. 15. As can be seen, under the same compression ratio ($\frac{M}{N}$), our proposed algorithm provides a substantial gain of PSNR compared with the existing methods. Indeed, it provides the best performance, while the improvement is not substantial, specifically in comparison with the Cluss algorithm. However, especially in edge details, the visual quality of the recovered images is better than that of Cluss. This is mainly because of the Bernoulli–Laplacian prior introduced in our proposed

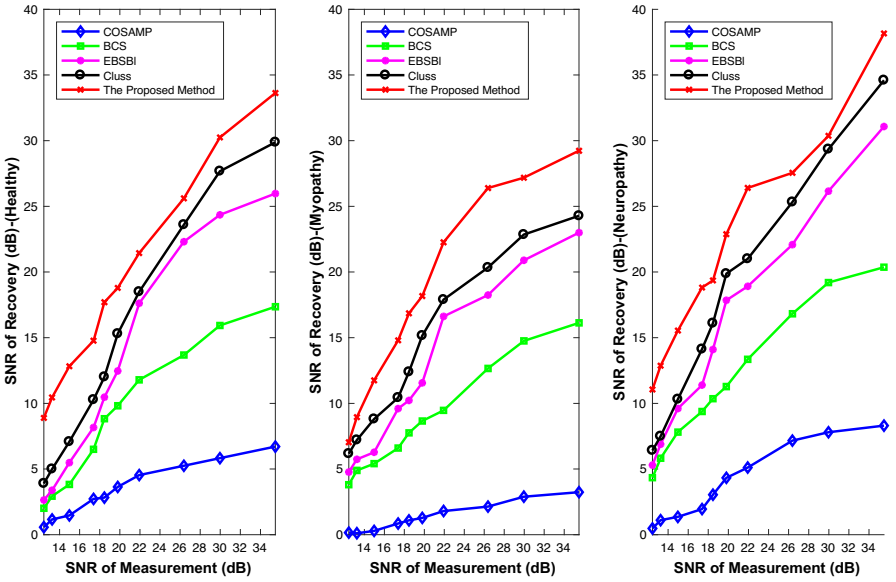


Fig. 7 Comparison of EMG signals recovery performance with different levels of noise, using CoSaMP, BCS, EBSBL, Cluss, and the proposed algorithms. From left to right, healthy, myopathy and neuropathy signals, respectively

algorithm, which is a sparser prior than Bernoulli–Gaussian used in the Cluss scheme. Also, it can obviously be seen that the proposed method has better performance than CoSaMP, which has the worst performance in recovering the images.

6 Discussion

In this work, we focus on block-sparse signal recovery, assuming the significant coefficients of a sparse signal are in clustered blocks. In many biomedical applications, like MRI image reconstruction, EMG/ECG recovery, and EEG source localization in sensor networks, this type of sparse pattern is commonly exploited. The problem is traditionally solved by applying a regularization that imposes realistic constraints on the solution. From this kind of regularization, the L_0 norm is recognized to recover efficaciously sparse signals. Inopportunely, the L_0 pseudo-norm minimization is intractable.

Consequently, it is commonly approximated by the L_1 norm, which is convex and can be solved simply using conventional optimization techniques but does not offer a similar solution [5]. Bayesian models and techniques have become standard statistical tools for solving inference problems, contributed to many applications for signal and image processing. Combining both L_0 and L_1 norms in a Bayesian framework can provide suitable results [7, 11]. It is shown that a Laplacian prior can be the Bayesian equivalent to the L_1 norm, while a Bernoulli prior is equivalent of the L_0 norm. In this article, we suggested a hierarchical Bayesian model using multivariate Bernoulli–Laplace prior, which is equivalent to an $L_0 + L_1$ norm. The L_0 holds the

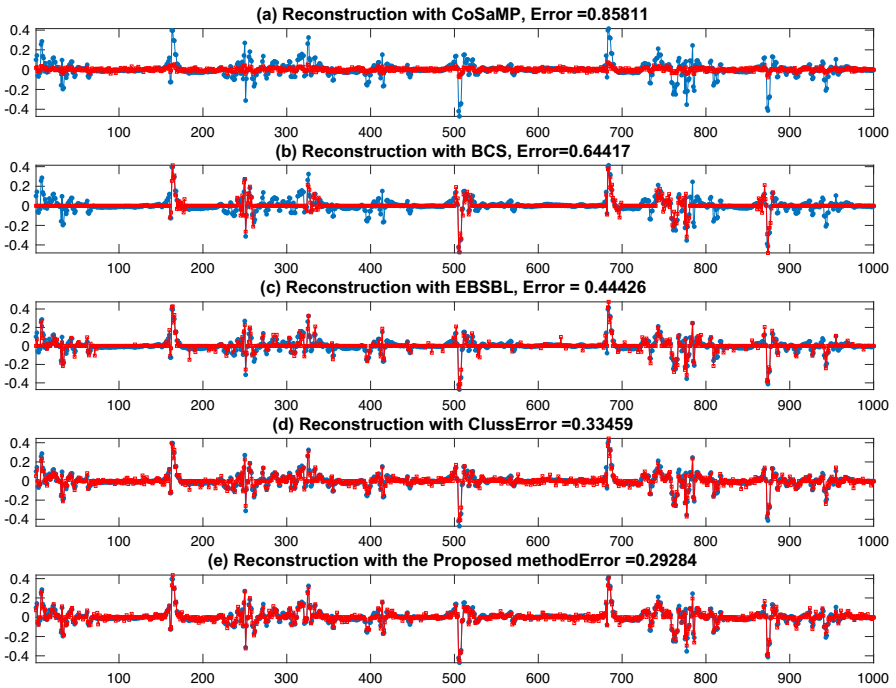


Fig. 8 Reconstruction results and errors of a section of EMG with healthy for different algorithms, **a** CoSaMP, **b** BCS, **c** EBSBL, **d** Cluss, and **e** the proposed algorithms. The blue ones are the original signals and the red ones are the reconstructed signals (Color figure online)

nonzero position of coefficients, while the L_1 norm constrains their amplitude values. Besides, we use the block-structured property of sparse signals, of which the nonzero coefficients appear in blocks. It is proven that block-structured sparsity can improve results by taking advantage of the temporal dimension of the data [24]. The posterior distribution of this model is too complex to derive a closed-form expression using the Bayesian estimator.

Currently, several methods are available for estimating the parameters of the proposed model. In this study, we used a Markov chain Monte Carlo (MCMC) sampling scheme to generate samples according to the posterior. The driven samples were then applied to estimate the model parameters in a fully unsupervised framework. In this paper, we proposed a new hierarchical Bayesian model based on a Bernoulli–Laplacian prior and structured sparsity. The proposed model can be considered as an extension of the hierarchical Bayesian modeling proposed in [50], called "Cluss". The principal contribution of this paper is to promote sparsity for source recovery via a structured Bernoulli–Laplace prior introduced in [7]. The proposed scheme can solve the inverse problem automatically while the prior information of the number of clusters and their sizes is unknown. The Bernoulli part guarantees the sources' sparseness, and the Laplacian prior promotes the flexibility of regularization to enforce the sparseness of the signal by solving the inverse problem [8, 11, 32, 39].

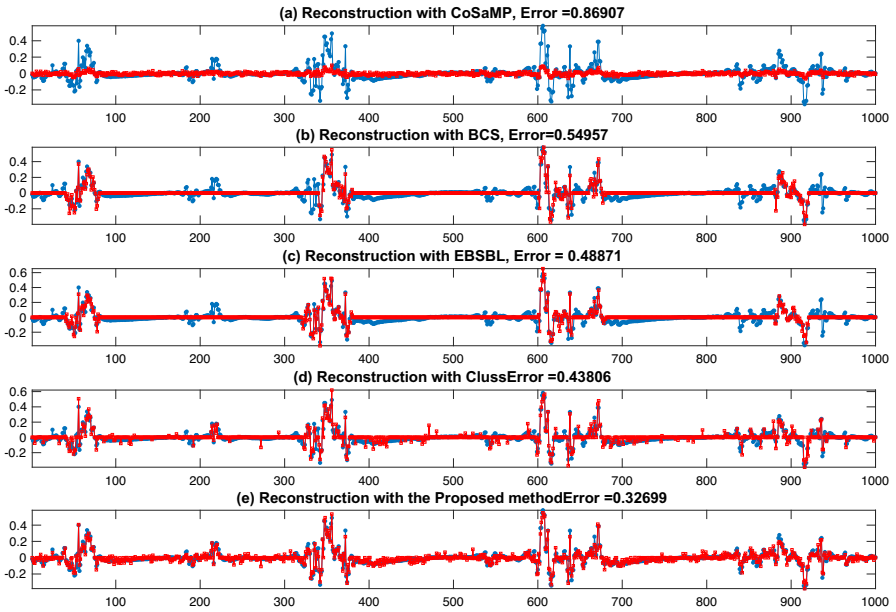


Fig. 9 Reconstruction results and errors of a section of EMG with myopathy for different algorithms, **a** CoSaMP, **b** BCS, **c** EBSBL, **d** Cluss, and **e** the proposed algorithms. The blue ones are the original signals and the red ones are the reconstructed signals (Color figure online)

The performance of the proposed method is investigated by both synthetic and real data compared to the more traditional CoSaMP, BCS, EBSBL, and Cluss methods. For real data, we use EMG and ECG signals taken from the *PhysioBank* database [20] as well as several MRI images [45]. The obtained results show that for synthetic and real data, the proposed method presents good performance compared to the aforementioned approaches in different scenarios. Specifically, the proposed method demonstrates clear advantages over the base method (Cluss) through the use of a Bernoulli–Laplacian prior, which is proved to be a sparser prior compared to the Bernoulli–Gaussian prior employed in the Cluss scheme. The Bernoulli–Laplacian prior makes it a more suitable choice for handling sparse signals, because of its sparsity-inducing properties ($L_0 + L_1$), compared to the Cluss scheme, which is using a combination of the L_0 and L_2.

First, we used synthetic data sets with several sparsity scenarios to explore the proposed method. To this end, synthetic cluster-structured-sparse signals were generated and the relative error was calculated. The objective of the first scenario was to provide a complete viewpoint for the evaluation of the signal recovery performance. Figure 2 illustrates the reconstruction result of the proposed method compared with previous methods, e.g., CoSaMP, BCS, EBSBL, and Cluss. Also, it has shown the proposed method's ability to gain the lowest reconstruction error, which is due to the using combination of two sparse priors (Bernoulli–Laplace distribution) in recovering the cluster-structured-sparse signal.

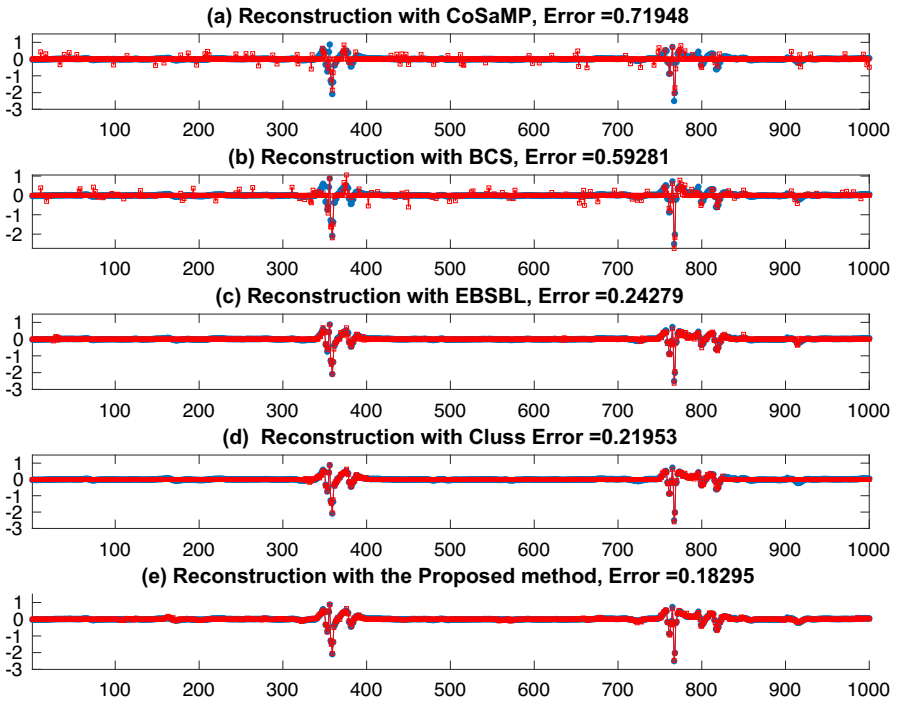


Fig. 10 Reconstruction results and errors of a section of EMG with neuropathy for different algorithms, **a** CoSaMP, **b** BCS, **c** EBSBL, **d** Cluss, and **e** the proposed algorithms. The blue ones are the original signals and the red ones are the reconstructed signals (Color figure online)

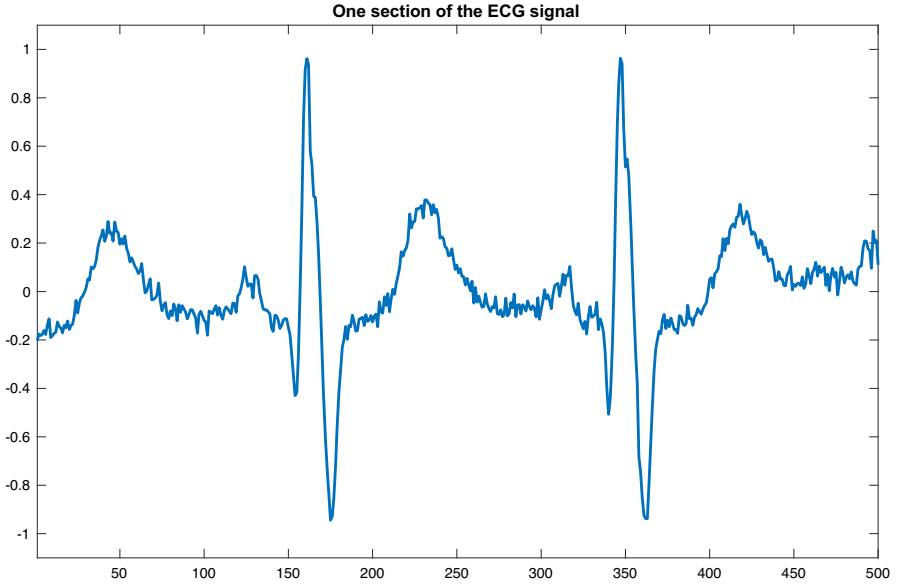


Fig. 11 An example of the used ECG signal

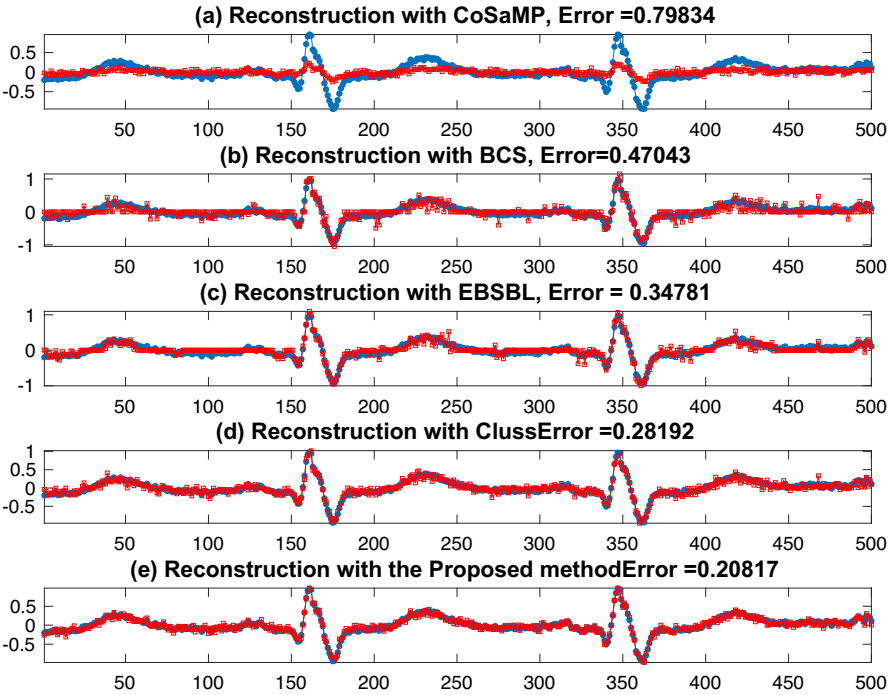


Fig. 12 Reconstruction results and errors of a section of ECG for different algorithms, **a** CoSaMP, **b** BCS, **c** EBSBL, **d** Cluss, and **e** the proposed algorithms. The blue ones are the original signals and the red ones are the reconstructed signals (Color figure online)

Next, the recovery abilities of the aforementioned methods for the various $\frac{M}{N}$ ratios were compared. The experiment is repeated 100 times with different realizations of observation operator, noise, and signal. It is worth noting that by increasing the $\frac{M}{N}$ ratio, the performance of all methods is enhanced. Figure 3 displays that the proposed method obtained the lowest reconstruction error compared to CoSaMP, BCS, EBSBL, and Cluss for the varying $\frac{M}{N}$ ratios. We also evaluated the noise robustness of the proposed algorithm. To this end, measurements are contaminated by white Gaussian noises with variance $\sigma_0^2 = 0.01$ to 0.15, namely the SNR ranging from 36 to 12 dB. For every level of noise, the experiment was run 100 times to achieve the mean and the variance of the recovery SNR (Fig. 4). The result illustrated that only the proposed method could recover the sparse signal with adequate error, confirming the results shown in the previous sections.

In the next step, three kinds of EMG signals, from a healthy participant, and two patients with myopathy and neuropathy are used to demonstrate the performance of the proposed method in comparison with the existing algorithms. To this end, for three kinds of EMG signals, the recovery abilities of the aforementioned methods for the various $\frac{M}{N}$ ratios were compared (Fig. 6). Numerical results indicate that the proposed method outperforms the mentioned algorithms in the structured-sparse signal recovery for the varying $\frac{M}{N}$ ratios. Also, it can be seen that by increasing the subsampling ratio

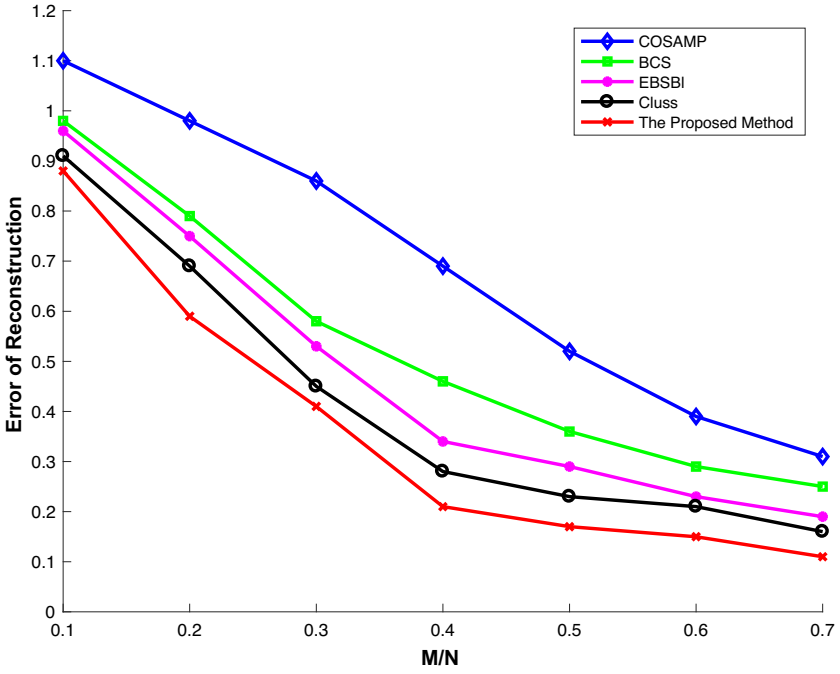


Fig. 13 Errors of reconstruction versus different $\frac{M}{N}$ ratios for ECG signal, using CoSaMP, BCS, EBSBL, Cluss, and the proposed algorithms

$\frac{M}{N}$, all estimation errors decrease. Then, we measure the average SNR of the recovered signals for different values of the SNR of measurements. For each value of the SNR of measurements, we repeated the experiment 100 times. Our results show that for every input SNR, the proposed method yields the highest SNR among all the considered methods (Fig. 7).

Another practical application of the proposed algorithm is presented by using ECG real data set. Reconstruction errors obtained by aforementioned sparse recovery algorithms with varying the $\frac{M}{N}$ ratios in 100 trials are presented in Fig. 13. Comparing the other signal recovery methods, the proposed method achieves the lowest errors from $M = 50$ to $M = 350$. One can see that by increasing the $\frac{M}{N}$ ratio, all estimation errors decrease. Also, we reported the SNR of the recovered signal, which is averaged over 100 independent trials, as a function of the SNR of measurements (Fig. 14). The obtained results show that there is a meaningful difference in terms of SNR between the proposed algorithm and the existing methods.

In the next experiment, we used four low-rank MRIs [45] with 100×100 pixels. The number of measurements is set to $M = 3000$. The visual quality of image recovery and the PSNR values are shown in Fig. 15. It can be seen that under the same compression ratio ($\frac{M}{N}$), our proposed algorithm provides a substantial gain of PSNR compared with the other existing methods, while CoSaMP has the worst performance to recover the images.

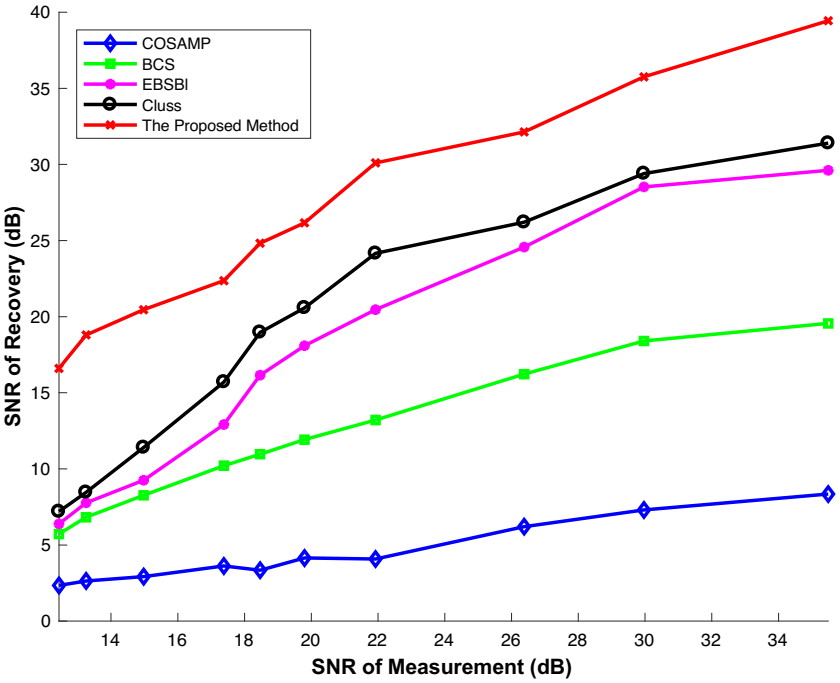


Fig. 14 Comparison of ECG signal recovery performance with different levels of noise, using CoSaMP, BCS, EBSBL, Cluss, and the proposed algorithms

7 Conclusions

In this study, a recovery scheme under a hierarchical Bayesian framework for biomedical signal and image applications is proposed. The proposed algorithm encourages sparse distribution using a two-level prior distribution, *e.g.*, Bernoulli–Laplace prior, which is the Bayesian equivalent of L_0+ L_1 norm regularization. On the other hand, we enhanced the known information of signals by considering the block-sparseness nature of some signals. Several simulation experiments demonstrate the effectiveness of the proposed method compared with state-of -the-art methods, *e.g.*, CoSaMP, BCS, EBSBL, and Cluss. In the case of real-world data, we used ECG signals, healthy, myopathy, and neuropathy EMG signals, and MR images to demonstrate the performance improvement of the proposed method. Using the Bernoulli–Laplacian prior, the proposed approach leverages the strengths of both L_0 and L_1 regularization techniques, leading to more accurate and efficient signal and image recovery. The tested biomedical signals and images, being inherently sparse, align well with this sparsity-inducing prior, enhancing the effectiveness of the proposed method. Numerical results in the simulation and real-world experiments showed the superiority of the newly proposed algorithm. It shows better performance in terms of relative error, SNR, and PSNR of the recovered biomedical signals and images compared to the state-of-the-art methods. In the future, we try to extend the proposed algorithm to be more adapted

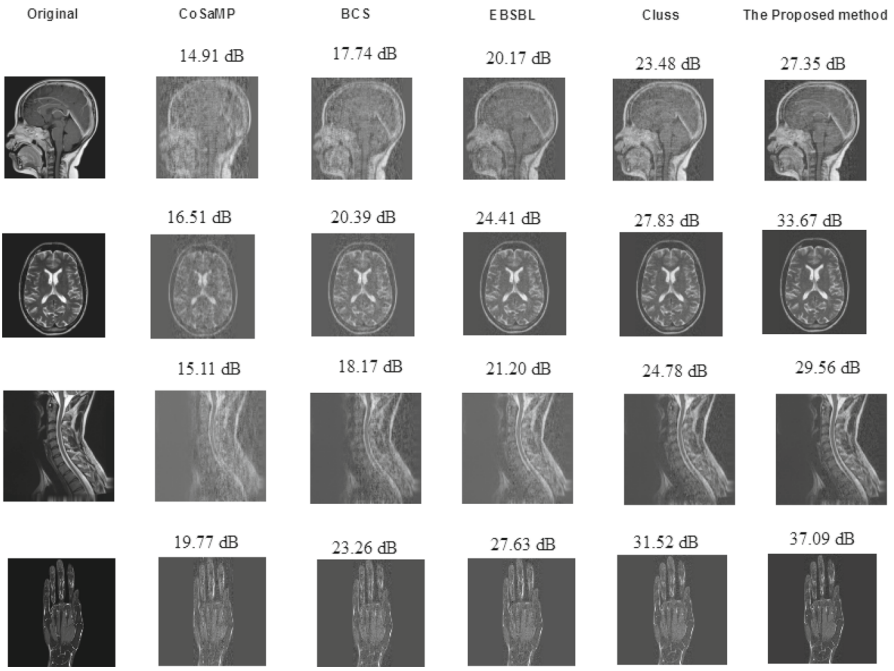


Fig. 15 MRI recovery results obtained by **a** CoSaMP, **b** BCS, **c** EBSBL, **d** Cluss, and **e** the proposed algorithms

to other physiological signals, such as EEG signals; mainly, the hybrid method will be proper for EEG source reconstruction when epileptic seizures occur.

Funding This research was supported in part by Shahid Chamran University of Ahvaz via grant SCU.EE99.82.

Data Availability The ECG database can be downloaded on the following site: https://archive.physionet.org/physiobank/database/ptbdb/. The EMG database can be downloaded on the following site: https://physionet.org/content/emgdb/1.0.0/

Declarations

Conflict of interest The author confirms that there are no known conflicts of interest associated with this publication.

References

1. R.G. Baraniuk, V. Cevher, M.F. Duarte, C. Hegde, Model-based compressive sensing. IEEE Trans. Inf. Theory **56**, 1982–2001 (2010)
2. T. Blumensath, M. Yaghoobi, M.E. Davies, Iterative hard thresholding and l0 regularisation. IEEE Int. Conf. Acoust. Speech Signal Process ICASSP. **3**, 877–880 (2007)
3. R. Bousseljot, D. Kreiseler, A. Schnabel: Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet. 317–318 (1995).

4. E.J. Candes, T. Tao, Decoding by linear programming. IEEE Trans. Inf. Theory **51**, 4203–4215 (2005)
5. E.J. Candes, The restricted isometry property and its implications for compressed sensing. Comptes Rendus Math. **346**, 589–592 (2008)
6. E.J. Candes, M.B. Wakin, An introduction to compressive sampling. IEEE Signal Process. Mag. **25**, 21–30 (2008)
7. L. Chaari, J.-Y. Tourneret, H. Batatia: Sparse bayesian regularization using bernoulli-laplacian priors. In: Signal Process. Conf. (EUSIPCO). 1–5 (2013).
8. L. Chaari, H. Batatia, J.-Y. Tourneret: Sparse Bayesian image restoration with linear operator uncertainties with application to EEG signal recovery. Midd. East Conf. Biomed. Eng. (MECBME). 139–142 (2014).
9. R. Chartrand, W. Yin Iteratively reweighted algorithms for compressive sensing. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP). 3869–3872 (2008).
10. S.S. Chen, D.L. Donoho, M.A. Saunders, Atomic decomposition by basis pursuit. SIAM Rev. **43**, 129–159 (2001)
11. F. Costa, H. Batatia, L. Chaari, J.-Y. Tourneret, Sparse EEG source localization using bernoulli laplacian priors. IEEE Trans. Biomed. Eng. **62**, 2888–2898 (2015)
12. F. Costa, H. Batatia, T. Oberlin, J.-Y. Tourneret EEG source localization based on a structured sparsity prior and a partially collapsed Gibbs sampler. IEEE Int. Work. Comput. Adv. Multi-Sensor Adapt. Process. (CAMSAP). 261–264 (2015).
13. F. Costa, H. Batatia, T. Oberlin, C. Dgiano, J.-Y. Tourneret, Bayesian EEG source localization using a structured sparsity prior. Neuroimage **144**, 142–152 (2017)
14. I. Daubechies, M. Defrise, C. De Mol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. Commun. Pure Appl. Math. **57**, 1413–1457 (2004)
15. A.M. Dixon, E.G. Allstot, D. Gangopadhyay, D.J. Allstot, Compressed sensing system considerations for ECG and EMG wireless biosensors. IEEE Trans. Biomed. Circuits Syst. **6**, 156–166 (2012)
16. N. Dobigeon, A.O. Hero, J.-Y. Tourneret, Hierarchical Bayesian sparse image reconstruction with application to MRFM. IEEE Trans. Image Process. **18**, 2059–2070 (2009)
17. Y.C. Eldar, M. Mishali, Robust recovery of signals from a structured union of subspaces. IEEE Trans. Inf. Theory **55**, 5302–5316 (2009)
18. Y.C. Eldar, P. Kuppinger, H. Bölcskei, Block-sparse signals: Uncertainty relations and efficient recovery. IEEE Trans. Signal Process. **58**, 3042–3054 (2010)
19. J. Fang, Y. Shen, H. Li, P. Wang, Pattern-coupled sparse Bayesian learning for recovery of block-sparse signals. IEEE Trans. Signal Process. **63**, 360–372 (2014)
20. A.L. Goldberger, L.A. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.-K. Peng, H.E. Stanley, Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals. Circulation **101**, 215–220 (2000)
21. A. Gramfort, M. Kowalski, M. Hämäläinen, Mixed-norm estimates for the M/EEG inverse problem using accelerated gradient methods. Phys. Med. Biol. **57**, 1937 (2012)
22. L. He, L. Carin, Exploiting structure in wavelet-based Bayesian compressive sensing. IEEE Trans. Signal Process. **57**, 3488–3497 (2009)
23. L. He, H. Chen, L. Carin, Tree-structured compressive sensing with variational Bayesian analysis. IEEE Signal Process. Lett. **17**, 233–236 (2009)
24. J. Huang, T. Zhang, The benefit of group sparsity. Ann. Stat. **38**, 1978–2004 (2010)
25. J. Huang, T. Zhang, D. Metaxas, learning with structured sparsity. Proc. Ann. Int. Conf. Mach. Learn. 417–424 (2009).
26. Q. Huynh-Thu, M. Ghanbari, Scope of validity of PSNR in image/video quality assessment. Electron. Lett. **44**, 800–801 (2008)
27. S. Ji, Y. Xue, L. Carin, Bayesian compressive sensing. IEEE Trans. Signal Process. **56**, 2346–2356 (2008)
28. M. Korki, J. Zhangy, C. Zhang, H. Zayyani, An iterative Bayesian algorithm for block-sparse signal reconstruction. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP). 2174–2178 (2015).
29. M. Kowalski, K. Siedenburg, M. Dörfler, Social sparsity! neighborhood systems enrich structured shrinkage operators. IEEE Trans. Signal Process. **61**, 2498–2511 (2013)
30. X. Lv, G. Bi, C. Wan, The group lasso for stable recovery of block-sparse signal representations. IEEE Trans. Signal Process. **59**, 1371–1382 (2011)
31. T.J. Mitchell, J.J. Beauchamp, Bayesian variable selection in linear regression. J. Am. Stat. Assoc. **83**, 1023–1032 (1988)

32. A. Mohammad-Djafari, Bayesian approach with prior models which enforce sparsity in signal and image processing. E EURASIP J. Adv. Signal Process. **1**, 1–19 (2012)
33. D. Needell, J.A. Tropp, CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. Appl. Comput. Harmon. Anal. **26**, 301–321 (2009)
34. N.G. Polson, J.G. Scott, Shrink globally, act locally: sparse Bayesian regularization and prediction. Bayesian Stat. **9**, 501–538 (2010)
35. J. Ren, C. Wei, L. Yu, H. Zhang, H. Sun, Dynamic recovery for block sparse signals. Signal Process. **130**, 197–203 (2017)
36. C.P. Robert, G. Casella, G. Casella, *Monte Carlo statistical methods* (Springer, New York, 1999)
37. A. Salman, E. G. Allstot, A. Y. Chen, A. M. Dixon, D. Gangopadhyay, D. J. Allstot, Compressive sampling of EMG bio-signals. IEEE Int. Symp. Circuits Syst. (ISCAS). 2095–2098 (2011).
38. M. Shekaramiz, T. K. Moon, J. H. Gunther AMP-B-SBL, An algorithm for clustered sparse signals using approximate message passing. IEEE Ann. Ubiquit. Comput. Elec. Mob. Comm. Conf. (UEMCON). 1–5 (2016).
39. M. Shekaramiz, T. K. Moon, J. H. Gunther, Sparse Recovery Via Variational Bayesian Inference: Comparing Bernoullis-Gaussians-Inverse Gamma And Gaussians-Inverse Gammas Modeling. IEEE Asilomar Conf. Signal. Syst. Comput. 1969–1973: (2018).
40. M. Shekaramiz, T.K. Moon, J.H. Gunther, Bayesian compressive sensing of sparse signals with unknown clustering patterns. Entropy **21**, 247 (2019)
41. M. Shekaramiz, T. K. Moon, Compressive Sensing via Variational Bayesian Inference. IEEE Intermount. Eng. Tech. Comput. (IETC). 1–6: (2020).
42. G. Swirszcz, N. Abe, A. C. Lozano, Grouped orthogonal matching pursuit for variable selection and prediction. Adv. Neural Inf. Process. Syst. 1150–1158 (2009).
43. J.A. Tropp, A.C. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit. IEEE Trans. Inf. Theory **53**, 4655–4666 (2007)
44. E. Van Den Berg, M.P. Friedlander, Probing the Pareto frontier for basis pursuit solutions. SIAM J. Sci. Comput. **31**, 890–912 (2009)
45. S. Viswanath, M. Ghulyani, M. Arigovindan, Structurally adaptive multi-derivative regularization for image recovery from sparse fourier samples. arXiv preprint arXiv:2105, 12775 (2021).
46. L. Wang, L. Zhao, L. Yu, J. Wang, G. Bi, Structured Bayesian learning for recovery of clustered sparse signal. Signal Process. **166**, 107255 (2020)
47. P.M. Williams, Bayesian regularization and pruning using a Laplace prior. Neural Comput. **7**, 117–143 (1995)
48. D.P. Wipf, B.D. Rao, An empirical Bayesian strategy for solving the simultaneous sparse approximation problem. IEEE Trans. Signal Process. **55**, 3704–3716 (2007)
49. G. Yu, G. Sapiro, S. Mallat, Solving inverse problems with piecewise linear estimators: from Gaussian mixture models to structured sparsity. IEEE Trans. Image Process. **21**, 2481–2499 (2011)
50. L. Yu, H. Sun, J.-P. Barbot, G. Zheng, Bayesian compressive sensing for cluster structured sparse signals. Signal Process. **92**, 259–269 (2012)
51. L. Yu, C. Wei, J. Jia, H. Sun, Compressive sensing for cluster structured sparse signals: variational Bayes approach. IET Signal Process. **10**, 770–779 (2016)
52. M. Yuan, Y. Lin, Model selection and estimation in regression with grouped variables. J. R Stat. Soc. Ser. B Stat. Meth. **68**, 49–67 (2006)
53. Z. Zhang, B.D. Rao, Extension of SBL algorithms for the recovery of block sparse signals with intra-block correlation. IEEE Trans. Signal Process. **61**, 2009–2015 (2013)
54. S. Zheng, C. Ding, Sparse classification using group matching pursuit. Neurocomputing **338**, 83–91 (2019)

Authors and Affiliations

Mayadeh Kouti[1,2] · Karim Ansari-Asl[1] · Ehsan Namjoo[1]

✉ Karim Ansari-Asl
karim.ansari@scu.ac.ir

Mayadeh Kouti
m.kouti@scu.ac.ir

Ehsan Namjoo
ehsan.namjoo@scu.ac.ir

1 Department of Electrical Engineering, Faculty of Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran

2 Department of Electrical Engineering, Shohadaye Hoveizeh Campus of Technology, Shahid Chamran University of Ahvaz, Ahvaz, Iran