Alternative Extended Block Sparse Bayesian Learning for Cluster Structured Sparse Signal Recovery

Lu Wang[1], Lifan Zhao[1], Guoan Bi[1(✉)], and Xin Liu[2]

[1] School Electrical and Electronic Engineering, Nanyang Technological University, Singapore, Singapore
{wanglu,zhao0145,egbi}@ntu.edu.sg

[2] School of information and communication, Dalian University of Technology, Dalian, China
liuxinstar1984@dlut.edu.cn

Abstract. Clustered sparse signals recovery with unknown cluster sizes and locations is considered in this paper. An improved alternative extended block sparse Bayesian learning algorithm (AEBSBL) is proposed. The new algorithm is motivated by the graphic models of the extended block sparse Bayesian learning algorithm (EBSBL). By deriving the graphic model of EBSBL, an equivalent cluster structured prior for sparse coefficients is obtained, which encourages dependencies among neighboring coefficients. With the sparse prior, other necessary probabilistic modelings are constructed and Expectation and Maximization (EM) is applied to infer all the unknowns. The alternative algorithm reduces the unknowns of EBSBL. Numerical simulations are conducted to demonstrate the effectiveness of the proposed method.

Keywords: Clustered structure · Bayesian sparse learning · Expectation · Maximization method

1 Introduction

Traditional sparse representation attempts to find a parsimonious coefficient vector $\mathbf{x} \in \mathbb{R}^{N\times 1}$ from noisy observation $\mathbf{y} \in \mathbb{R}^{M\times 1}$

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}, \tag{1}$$

where $\mathbf{A} \in \mathbb{R}^{M\times N}$ is an representative dictionary and n is the noise. To further include the inherent structure onto the signal, structured sparse recovery is proposed with improved recovery performance. Among which, block structure is one of the mostly observed structures in practice [1,2,4,5]. In block structure, elements of sparse coefficient $\mathbf{x}$ are likely to be nonzero or zeros in blocks. Particularly, when the block sizes or the localizations are unknown, the structure is referred to cluster structure.

M. Jia et al. (Eds.): WiSATS 2019, LNICST 280, pp. 3–12, 2019.
https://doi.org/10.1007/978-3-030-19153-5_1

Algorithms dealing with the simple block sparse signal with known block partition have been widely studied and can be generally divided into three categories. The greedy pursuit algorithms are one of the categories including Model-CoSaMp [6], Block-OMP [7], and their variations. The second category is the convex optimization, such as GLasso algorithm [8], GBasis Pursuit [9], l_2/l_1 Programming [10], etc. The third one is the sparse Bayesian learning method known as block-sparse Bayesian learning (BSBL) [11]. They performs competitively with each other in ideal situation with known block partition and size. Obvious performance degradation is observed when such algorithms are applied to the cluster structure where knowledge on the block partition is unavailable. Signal with cluster structure is generally much harder to recover. However, under such a circumstance, it is observed that neighboring elements of the sparse coefficients are statistical dependent on each other, which may be potentially used to capture the underlying structure of the signal. This fact has been studied by recently proposed methods, such as extended block-sparse Bayesian learning (EBSBL) in [11] as well as pattern-coupled sparse Bayesian learning in PC-SPL [12].

In EBSBL, an augmented vector is artificially constructed and linearly combined to represent the sparse coefficient to introduce interactions among its neighboring elements. Due to the augmented vector, an extended problem with higher dimensionality has to be solved in EBSBL. This paper tries to derive an explicit interaction expression of the neighboring coefficients and reduce the problem size of EBSBL. By studying the underlying relationship of the hidden variables in the graphic model of EBSBL, a clustered-sparsity imposed prior can be derived by integrating out the hidden augmented vector artificially constructed. The newly derived prior of one certain sparse coefficient introduces interactions of its own hyperparameter and those of its neighbors, thus imposing dependencies among neighbors. By constructing structured prior and other proper probabilistic models for all hidden variables, the original sparse representation problem is easily treated by Bayesian method. The expectation-maximization is used to estimate all the unknowns. The relationship between the alternative EBSBL with the original EBSBL is also revealed. Simulations are conducted to demonstrate its effectiveness.

2 The Review of the Idea of EBSBL

Since the proposed alternative method is motivated by the EBSBL, we first briefly review the basic idea of EBSBL.

2.1 EBSBL

The basic idea of EBSBL to recover clustered sparse signal is to try to reformulate the signal with unknown location and size by expanding its sparse coefficients into known blocks. The processing can be illustrated by Fig. 1. The constructed augmented vector is denoted by $\mathbf{z} = [\cdots \mathbf{z}_{i-1}\mathbf{z}_i\mathbf{z}_{i+1}\cdots]^T$, where block

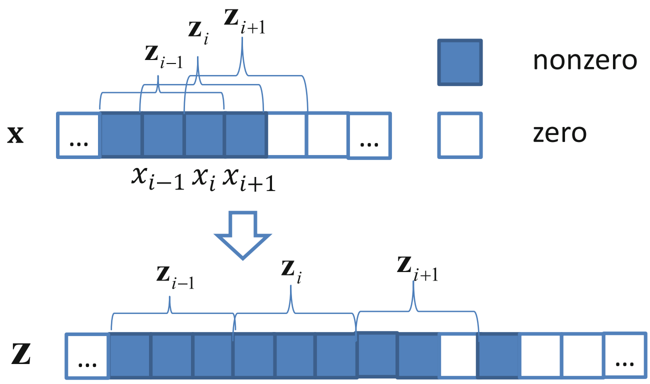

Fig. 1. The illustration of the expanding procedure.

$\mathbf{z}_i = [x_{i-1}x_ix_{i+1}]^T$. Since merely neighboring information is used, the block size in Fig. 1 can be reasonably set to 3. For clustered signal, due to the fact that neighboring coefficients are likely to behave in the same way, it is highly probably that the augmented vector is of block sparse with equally partitioned blocks of size 3. Therefore, $\mathbf{z}$ can be effectively recovered by the BSBL method. With the augment vector $\mathbf{z}$, the original problem can be approximately transformed into a block sparse recovery problem with known block partition:

$$\mathbf{y} = \sum_{i=1}^{g} \mathbf{A}\mathbf{E}_i\mathbf{z}_i + \mathbf{n} = \mathbf{\Phi}\mathbf{z} + \mathbf{n} \tag{2}$$

where $\mathbf{\Phi} = [\mathbf{\Phi}_1, \cdots, \mathbf{\Phi}_g]$ with sub-dictionary $\mathbf{\Phi}_i = \mathbf{A}\mathbf{E}_i$. $\mathbf{E}_i \in \mathcal{R}^{N\times 3}$ is a zero matrix with its rows from i-th row to $i+h-1$-th row being an identity matrix. The original coefficient $\mathbf{x}$ is then a linear transformation of $\mathbf{z}$:

$$\mathbf{x} = \sum\nolimits_{i=1}^{g} \mathbf{E}_i\mathbf{z}_i. \tag{3}$$

2.2 Discussion of EBSBL

It should be emphasized that the problem of cluster sparse signal recovery in (2) can only be approximately solved by the BSBL, since the block sparsity of each block $\mathbf{z}_i$ does not exactly hold. This is the fact with great probability. However, for those block $\mathbf{z}_i$ containing the edge of the clusters of $\mathbf{x}$, the block sparsity cannot be applied. This approximation will result in inaccuracies in the edge recovery of the clusters of $\mathbf{x}$.

3 Graphic Model of EBSBL

In this section, the graphic model of EBSBL is constructed to explicitly show the relationship of each hidden variable. And a new structured prior of $\mathbf{x}$ is constructed based on the graphic model.

3.1 Interaction Among the Hidden Variables of EBSBL

To apply the BSBL method to infer the hidden variable $\mathbf{z}$, $\mathbf{z}$ has to be hierarchically modeled:

$$p\left(\mathbf{z};\{\gamma_i\}_i,\mathbf{B}\right)=\mathcal{N}\left(0,\mathbf{\Sigma}_0\right) \tag{4}$$

with $\boldsymbol{\Sigma}_0=\mathrm{diag}\left(\gamma_1\mathbf{B},\cdots,\gamma_g\mathbf{B}\right)$. Each block $\mathbf{z}_i$ follows Gaussian distribution as below:

$$p\left(\mathbf{z}_i;\gamma_i,\mathbf{B}\right)=\mathcal{N}\left(0,\gamma_i,\mathbf{B}\right), \tag{5}$$

where γ_i is the parameter controlling the block sparsity degree. With those probabilistic model in (4) and (5) and the relationship between $\mathbf{x}$ and $\mathbf{z}$ in (3), The interactions among all the hidden variables of x, z and γ can be explicitly shown in the Graphic model of EBSBL in Fig. 2.

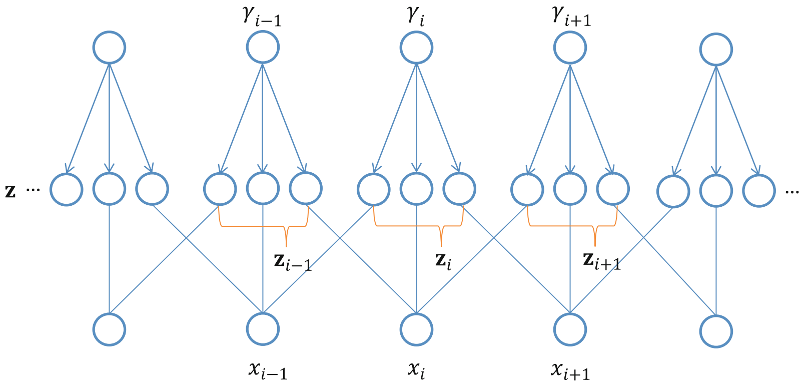

Fig. 2. The graphic model of EBSBL.

In Fig. 2, the ith element of x_i directly is connected to $\mathbf{z}_{i-1}\{3\}$, $\mathbf{z}_i\{2\}$ and $\mathbf{z}_{i+1}\{1\}$ according to the linear transformation in (3). Since $\mathbf{z}_{i-1}\{3\}$, $\mathbf{z}_i\{2\}$ and $\mathbf{z}_{i+1}\{1\}$ depends on γ_{i-1}, γ_i and γ_{i+1}, then x_i is implicitly connected to γ_{i-1}, γ_i and γ_{i+1}. Similarly, it is concluded that x_{i-1} depends on γ_{i-1}, γ_i and γ_{i+1}, and x_{i+1} depends on γ_{i-1}, γ_i and γ_{i+1}. Therefore, x_i implicitly interacts with its neighbors x_{i-1} and x_{i+1}. By this hierarchical probabilistic modeling, EBSBL easily exert the dependency on neighboring sparse coefficients in a probabilistic way. That statistical dependency will be used in EBSBL to enhance the performance of sparse recovery for signals with cluster structure.

It is noted that EBSBL requires to infer an extended vector $\mathbf{z}$ of size 3 times that of $\mathbf{x}$. To maintain the problem size to that of $\mathbf{x}$, we try to derive an explicit structured prior for the sparse coefficients, which is motivated by the graphic model of EBSBL.

3.2 Structured Prior for Clustering Sparse Signal

It should be emphasized that the graphic model in Fig. 2 only tells us that x_i is implicitly and jointly depending on γ_{i-1}, γ_i and γ_{i+1} through the intermediate hidden variables of $\mathbf{z}_{i-1}\{3\}$, $\mathbf{z}_i\{2\}$ and $\mathbf{z}_{i+1}\{1\}$. Actually, the explicit relationship of x_i on γ_{i-1}, γ_i and γ_{i+1} can possibly given by integrating out the intermediate variables $\mathbf{z}_{i-1}\{3\}$, $\mathbf{z}_i\{2\}$ and $\mathbf{z}_{i+1}\{1\}$. Since we previously assume that blocks of $\mathbf{z}_{i-1}$ for different i are independent as shown in the probabilistic model in (4), $\mathbf{z}_{i-1}\{3\}$, $\mathbf{z}_i\{2\}$ and $\mathbf{z}_{i+1}\{1\}$ are independent Gaussian random variables with variances γ_{i-1}, γ_i and γ_{i+1}. Therefore, $x_i = \mathbf{z}_{i-1}(3) + \mathbf{z}_i(2) + \mathbf{z}_{i+1}(1)$ according to (3) follows normal distribution with a variance of $\gamma_i + \gamma_{i+1} + \gamma_{i-1}$, i.e.,

$$p\left(x_i; \gamma_i, \gamma_{i-1}, \gamma_{i+1}\right) = \mathcal{N}\left(0, \left(\gamma_i + \gamma_{i-1} + \gamma_{i+1}\right)\right). \tag{6}$$

The overall prior of $\mathbf{x}$ also depends on the expression of $\mathbf{B}$ in (5). If $\mathbf{B}$ is an identity matrix, elements inside each $\mathbf{z}_i$ are independent. Therefore, x_i will be independent for all i and an explicit prior of $\mathbf{x}$ over $\boldsymbol{\gamma}$ can be given as below

$$p\left(\mathbf{x}; \boldsymbol{\gamma}\right) = \prod_i p\left(x_i; \gamma_i + \gamma_{i-1} + \gamma_{i+1}\right). \tag{7}$$

The probabilistic graphic model of prior in (7) can be given by Fig. 3. As illustrated in Fig. 3, prior for coefficient x_i directly involves its own hyperparameter γ_i and those of its neighbors x_{i-1} and x_{i+1}. The hyperparameters are coupled together to introduce the dependency.

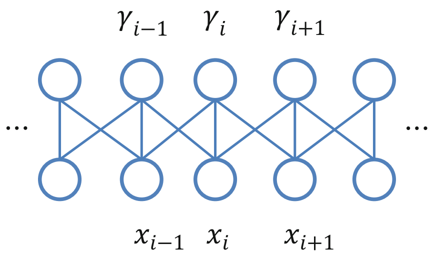

Fig. 3. The graphic model of the proposed structured prior.

It should be noted that if $\mathbf{B}$ in (5) has nonzeros on its off-diagonal, elements inside each $\mathbf{z}_i$ will not be independent. Under this case, neighboring x_i will be correlated with each other and the distribution of $\mathbf{x}$ over $\boldsymbol{\gamma}$ can not be written explicitly. For simplicity, we always assume an identity matrix for $\mathbf{B}$. In contrast, EBSBL theoretically allows arbitrary form of $\mathbf{B}$.

3.3 Discussions on the Constructed Prior

The hyperparamter coupled prior in (7) coincidently shows a similar form of the pattern-coupled prior proposed in [12]. However, it should be noted that coupling of the hyperparameter is modeled as the precision of x_i in [12], while coupling of the hyperparameter in our case is the variance of sparse coefficient x_i and statistically implies x_i of a linear combination of three independent normal variables.

The assumption of an identity matrix for $\mathbf{B}$, though exerts more restriction, is reasonable. The independence of elements in $\mathbf{z}_i$ somehow decreases the aforementioned inaccuracy of edges in the non-zero cluster, since it allow totally different elements in $\mathbf{z}_i$. To see this, first let us assume each element of $\mathbf{B}$ is 1. Consider $x_i = \mathbf{z}_{i-1}(3) + \mathbf{z}_i(2) + \mathbf{z}_{i+1}(1) \neq 0$ is on the edge of the nonzero cluster and $x_{i-1} = \mathbf{z}_{i-2}(3) + \mathbf{z}_{i-1}(2) + \mathbf{z}_i(1) \neq 0$ and $x_{i+1} = \mathbf{z}_i(3) + \mathbf{z}_{i+1}(2) + \mathbf{z}_{i+2}(1) = 0$. In this situation, block $\mathbf{z}_i$ should be nonzero and $\mathbf{z}_i(3) \approx \mathbf{z}_i(2) \approx \mathbf{z}_i(1) \gg 0$ with the most probability according EBSBL, thus leading to a large inaccuracy in estimation of x_{i+1}. On the other hand, if elements in $\mathbf{z}_i$ are independent, $\mathbf{z}_i(3)$ is allowed to be close to 0, which makes it easy to estimate x_{i+1}.

4 Alternative to EBSBL for Cluster Sparse Signal Recovery

In order to formulate our problem under Bayesian framework, in addition to the cluster-structured prior of $\mathbf{x}$ in (7), necessary proper probabilistic models are need to be constructed for other hidden variables in Bayesian treatment. After that, Expectation and Maximization algorithm can be applied to infer all the hidden variables and unknown parameters.

4.1 Updating Rules for Unknowns

The noise $\mathbf{n}$ is assumed to be Gaussian distributed with variance α_0, so that the likelihood function of $\mathbf{y}$ in (1) is as follows:

$$\mathbf{y}\,|\mathbf{x};\alpha_0 \sim \mathcal{N}\left(\mathbf{A}\mathbf{x}, \alpha_0 \mathbf{I}_M\right), \tag{8}$$

where $\mathbf{I}_M$ is an identity matrix with size M.

We can then easily calculate the posterior of $\mathbf{x}$ given likelihood in (8) and its prior of (7), which is of a Gaussian distribution with mean and covariance

$$\begin{aligned} \tilde{\mathbf{x}} = \boldsymbol{\mu} &= \alpha_0^{-1} \boldsymbol{\Sigma} \mathbf{A}^T \mathbf{y} \\ &= \mathbf{D}\mathbf{A}^T \left(\alpha_0 \mathbf{I} + \mathbf{A}\mathbf{D}\mathbf{A}^T\right)^{-1} \mathbf{y}, \end{aligned} \tag{9}$$

and

$$\boldsymbol{\Sigma} = \left(\alpha_0^{-1} \mathbf{A}^H \mathbf{A} + \mathbf{D}\right)^{-1}. \tag{10}$$

In (10), $\mathbf{D}$ is a diagonal matrix and the i-th element on its diagonal is $(\alpha_{i-1} + \alpha_i + \alpha_{i+1})^{-1}$.

For the unknown parameters estimation in Bayesian treatment, the algorithm of Expectation-Maximization is mostly used. In EM, $\boldsymbol{\gamma}$ and α_0 can be found by maximizing

$$E_{\mathbf{x}|\mathbf{y},\gamma,\alpha_0} \left\{ p\left(\mathbf{y}|\mathbf{x};\alpha_0\right) p\left(\mathbf{x};\boldsymbol{\gamma}\right) \right\}, \tag{11}$$

where $E_{\mathbf{x}|\mathbf{y};\gamma,\alpha_0}\{\cdot\}$ denotes an expectation with respect to the posterior of $\mathbf{x}$.

According to (11), $\boldsymbol{\gamma}$ can be given by maximizing

$$E_{\mathbf{x}|\mathbf{y};\boldsymbol{\alpha}}\{\log p(\mathbf{x};\boldsymbol{\alpha})\} = \\ -\sum_{i=1}^{N}\log(\alpha_{i-1}+\alpha_i+\alpha_{i+1}) - \sum_{i=1}^{N}\frac{E_{\mathbf{x}|\mathbf{y};\boldsymbol{\alpha}}\{x_i^H x_i\}}{\alpha_{i-1}+\alpha_i+\alpha_{i+1}}, \tag{12}$$

with $E_{\mathbf{x}|\mathbf{y};\boldsymbol{\alpha}}\{x_i^H x_i\} = \mu_i^2 + \boldsymbol{\Sigma}_{i,i}$ and $\boldsymbol{\Sigma}_{i,i}$ being the i-th diagonal component of $\boldsymbol{\Sigma}$. Let $v_i = \alpha_{i-1}+\alpha_i+\alpha_{i+1}$ and $w_i = \mu_i^2 + \boldsymbol{\Sigma}_{i,i}$. $\boldsymbol{\gamma}$ can be found by letting the gradient of (12) with respect to γ_i to zero:

$$\frac{\partial E_{\mathbf{x}|\mathbf{y};\boldsymbol{\alpha}}\{\log p(\mathbf{x};\boldsymbol{\alpha})\}}{\partial \alpha_i} \\ = \frac{w_i}{v_i^2} + \frac{w_{i-1}}{v_{i-1}^2} + \frac{w_{i+1}}{v_{i+1}^2} - \frac{1}{v_i} - \frac{1}{v_{i-1}} - \frac{1}{v_{i+1}} = 0. \tag{13}$$

The root of (13) can not be given in a close-form, since γ_i is entangled with other unknowns of γ_{i-1} and γ_{i+1}. There are three possible ways to find the estimation of γ_i. One simple way is to iteratively find the root γ_i of (13) by using the previously estimated γ_{i-1} and γ_{i+1}. This will be time-consuming. The other possible method is to find the lower and upper bound of (13) following [12] to analyze the approximation of the root.

Let us define

$$\bar{\gamma}_{\max} = \max\{\bar{\gamma}_{i-2}, \bar{\gamma}_{i-1}, \bar{\gamma}_i, \bar{\gamma}_{i+1}, \bar{\gamma}_{i+2}\}$$

and

$$\bar{\gamma}_{\min} = \min\{\bar{\gamma}_{i-2}, \bar{\gamma}_{i-1}, \bar{\gamma}_i, \bar{\gamma}_{i+1}, \bar{\gamma}_{i+2}\},$$

where $\bar{\gamma}_{i-2}, \bar{\gamma}_{i-1}, \bar{\gamma}_i, \bar{\gamma}_{i+1}, \bar{\gamma}_{i+2}$ are the optimal solution to (13). Therefore, we have the following equation:

$$\frac{w_i}{\bar{v}_i^2} + \frac{w_{i-1}}{\bar{v}_{i-1}^2} + \frac{w_{i+1}}{\bar{v}_{i+1}^2} = \frac{1}{\bar{v}_i} + \frac{1}{\bar{v}_{i-1}} + \frac{1}{\bar{v}_{i+1}}. \tag{14}$$

Since $\bar{v}_i^2 = (\bar{\gamma}_{i-1}+\bar{\gamma}_i+\bar{\gamma}_{i+1})^2 \le 9\bar{\gamma}_{\max}^2$, $\bar{v}_{i-1}^2 = (\bar{\gamma}_{i-2}+\bar{\gamma}_{i-1}+\bar{\gamma}_i)^2 \le 9\bar{\gamma}_{\max}^2$, and $\bar{v}_{i+1}^2 = (\bar{\gamma}_{i+2}+\bar{\gamma}_{i+1}+\bar{\gamma}_i)^2 \le 9\bar{\gamma}_{\max}^2$, we have the left hand side of (14) satisfy

$$\frac{w_i}{\bar{v}_i^2} + \frac{w_{i-1}}{\bar{v}_{i-1}^2} + \frac{w_{i+1}}{\bar{v}_{i+1}^2} \ge \frac{w_i + w_{i-1} + w_{i+1}}{9\bar{\gamma}_{\max}^2}. \tag{15}$$

And because $\bar{v}_i = \bar{\gamma}_{i-1}+\bar{\gamma}_i+\bar{\gamma}_{i+1} \ge 3\bar{\gamma}_{\min}$, $\bar{v}_{i+1} = \bar{\gamma}_{i+2}+\bar{\gamma}_{i+1}+\bar{\gamma}_i \ge 3\bar{\gamma}_{\min}$ and $\bar{v}_{i-1} = \bar{\gamma}_{i-2}+\bar{\gamma}_{i-1}+\bar{\gamma}_i \ge 3\bar{\gamma}$, the right hand side of (14) should satisfy

$$\frac{1}{\bar{v}_i} + \frac{1}{\bar{v}_{i-1}} + \frac{1}{\bar{v}_{i+1}} \le \frac{1}{\bar{\gamma}_{\min}}. \tag{16}$$

Therefore, it is easy to show

$$\frac{w_i + w_{i-1} + w_{i+1}}{9\bar{\gamma}_{\max}^2} \leq \frac{1}{\bar{\gamma}_{\min}}. \tag{17}$$

Since $\bar{\gamma}_{\min} \leq \bar{\gamma}_i \leq \bar{\gamma}_{\max}$, if we make additional constraint that $\bar{\gamma}_{\max} \to \bar{\gamma}_{\min}$, then we have an approximation to the optimal root of (13) from (17) as follows:

$$\bar{\gamma}_i = \frac{w_i + w_{i-1} + w_{i+1}}{9}. \tag{18}$$

Similarly, α_0 is calculated according to (11) by maximizing

$$\begin{aligned} &E_{\mathbf{x}|\mathbf{y};\alpha}\left\{\log p\left(\mathbf{y}\,|\mathbf{x};\alpha_0\,\right)\right\} \\ &\propto -\frac{M}{2}\log\alpha_0 - \frac{E_{\mathbf{x}|\mathbf{y};\alpha}\left\{\|\mathbf{y}-\mathbf{A}\mathbf{x}\|_2^2\right\}}{2\alpha_0} \\ &= -\frac{M}{2}\log\alpha_0 - \frac{\|\mathbf{y}-\mathbf{A}\mu\|_2^2 + \mathrm{Tr}\left(\Sigma\mathbf{A}^T\mathbf{A}\right)}{2\alpha_0}. \end{aligned} \tag{19}$$

Setting the derivative of (19) with respect to α_0 to be 0, α_0 is estimated by the root:

$$\tilde{\alpha}_0 = \frac{\|\mathbf{y}-\mathbf{A}\boldsymbol{\mu}\|_2^2 + \mathrm{Tr}\left(\Sigma\mathbf{A}^T\mathbf{A}\right)}{M}. \tag{20}$$

The algorithm iteratively estimates $\mathbf{x}$, α_0, and γ_i by $\boldsymbol{\mu}$ in (9), $\tilde{\alpha}_0$ in (20) and $\tilde{\gamma}_i$ in (18) till convergence.

4.2 Discussion

To find the close-form sub-optimal solution to (12), a constraint that $\bar{\gamma}_{\max} \to \bar{\gamma}_{\min}$ is assumed. This is somehow similar to the assumption that x_{i-1}, x_i and x_{i+1} share variance $\gamma_{i-2} = \gamma_{i-1} = \gamma_{i+1} = \gamma_{i+2} = \gamma_i$ as used in [3]. Same updating rule for γ_i is obtained as that in [3].

5 Simulations

The performance of the alternative to the EBSBL algorithm has been evaluated and comprehensively compared in [3]. For completeness, we quote the main performance evaluation results here. Two measures of support recovery rate and normalized root mean squared error (NRMSE) [3] are used for performance evaluation under the signal to noise ratio of 25 dB. The achieved support recovery and NRMSE are shown in Figs. 4 and 5, respectively, using 100 independent trials.

Figure 4 gives the support recovery rates against the sampling ratio while Fig. 5 gives the achieved NRMSE against the sparsity. As shown in sub-figures (b) of both figures, the NRMSEs of all algorithms decrease as the under-sampling

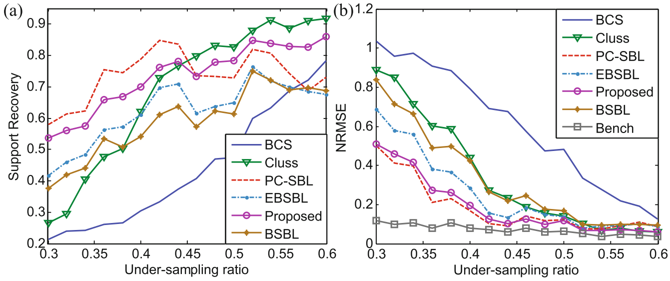

Fig. 4. Support recovery and NRMSE against sampling ratio when the signal is of a sparse degree of 25. (a) the support recovery; (b) the NRMSE.

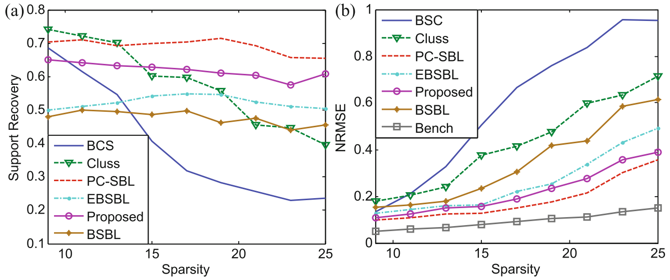

Fig. 5. Support recovery and NRMSE against sparsity level using 35 measurements. (a) the support recovery; (b) the NRMSE.

ratio increases and the sparsity level decreases. When the under-sampling ratio or the sparsity increases as shown in sub-figures (a) of both figures, the achieved support recovery rates of different algorithms all decrease.

It is easy to note that the structured sparse recovery algorithms generally achieve better NRMSE than that of BCS as shown in Fig. 4(a). The inaccuracy aforementioned in Sect. 2.2 can be noticed by the result that as the observations increase, the support recovery rates of all structure considered sparse representations hardly increase. This inaccuracy can also be noted in Fig. 5(a). We see that as the sparsity decreases, the support recovery rate achieved by either BSBL or EBSBL is lower than that achieved by conventional BCS [14]. It is noted that Cluss in [13] achieves a better support recovery rate as there are enough observations since prevention of structure over fitting is considered in [13].

6 Conclusion

This paper serves mainly as a supplementary for our prior work in [3]. The alternative extended block sparse Bayesian learning algorithm is re-investigated. More discussions on the relationship among the original problem, EBSBL and the alternative EBSBL are made for better understanding. A new interpretation of the derivation of close-form updating rules for the disentangling of the unknowns is provided. Problems of the edge inaccuracy in both the EBSBL and its alternative will be further investigated.

References

1. Mishali, M., Eldar, Y.C.: Blind multi-band signal reconstruction: compressed sensing for analog signals. IEEE Trans. Signal Process. **57**(3), 993–1009 (2009)
2. Wang, L., Zhao, L., Bi, G., Wan, C., Zhang, L., Zhang, H.: Novel wideband DOA estimation based on sparse bayesian learning with dirichlet process priors. IEEE Trans. Signal Process. **64**(2), 275–289 (2016)
3. Wang, L., Zhao, L., Rahardja, S., Bi, G.: Alternative to extended block sparse Bayesian learning and its relation to pattern-coupled sparse Bayesian learning. IEEE Trans. Signal Process. **66**(10), 275–289 (2018)
4. Wang, L., Zhao, L., Bi, G., Wan, C., Yang, L.: Enhanced ISAR imaging by exploiting the continuity of the target scene. IEEE Trans. Geosci. Remote Sens. **52**(9), 5736–5750 (2014)
5. Wang, L., Zhao, L., Bi, G., Wan, C.: Sparse representation based ISAR imaging using Markov random fields. IEEE J. Sel. Top. Appl. Earth Observ. **8**(8), 3941–3953 (2015)
6. Baraniuk, R.G., Cevher, V., Duarte, M.F., Hegde, C.: Model-based compressive sensing. IEEE Trans. Inf. Theory **56**(4), 1982–2001 (2010)
7. Eldar, Y.C., Kuppinger, P., Bolcskei, H.: Block-sparse signals: uncertainty relations and efficient recovery. IEEE Trans. Signal Process. **58**(6), 3042–3054 (2010)
8. Yuan, M., Lin, Y.: Model selection and estimation in regression with grouped variables. J. R. Statist. Soc. B **68**(8), 49–67 (2006)
9. Van Den Berg, E., Friedlander, M.: Probing the pareto frontier for basis pursuit solutions. SIAM J. Sci. Comput. **31**(2), 890–912 (2008)
10. Eldar, Y.C., Mishali, M.: Robust recovery of signals from a structured union of subspaces. IEEE Trans. Inf. Theoy **55**(11), 5302–5316 (2009)
11. Zhang, Z., Rao, B.D.: Extension of SBL algorithms for the recovery of block sparse signals with intra-block correlation. IEEE Trans. Signal Process. **61**(8), 2009–2015 (2013)
12. Fang, J., Shen, Y., Li, H., Wang, P.: Pattern-coupled sparse Bayesian learning for recovery of block-sparse signals. IEEE Trans. Signal Process. **63**(2), 360–372 (2015)
13. Yu, L., Sun, H., Barbot, J.P., Zheng, G.: Bayesian compressive sensing for cluster structured sparse signals. Signal Process. **92**(7), 259–269 (2012)
14. Ji, S., Xue, Y., Carin, L.: Bayesian compressive sensing. IEEE Trans. Signal Process. **55**(6), 2346–2356 (2008)