Localized Low-Rank Promoting for Recovery of Block-Sparse Signals With Intrablock Correlation

Linxiao Yang, Jun Fang, *Member, IEEE*, Hongbin Li, *Senior Member, IEEE*, and Bing Zeng, *Fellow, IEEE*

***Abstract*—We consider the problem of recovering block-sparse signals with intrablock correlated entries. The block partition of the sparse signal is assumed unknown *a priori*. To exploit the block-sparse structure as well as the local smoothness of the sparse signal, consecutive coefficients of the sparse signal are organized into a number of 2×2 matrices, and the log-determinant function is used to promote the low rankness of these 2×2 matrices. We show that such a log-determinant function has the ability to promote the block-sparsity and local smoothness simultaneously. An iterative reweighted method is developed by iteratively minimizing a surrogate function of the original objective function. Simulation results show that our proposed method offers competitive performance for recovering block-sparse signals with intrablock correlated entries.**

***Index Terms*—Block-sparse signal, compressed sensing, intrablock correlation, localized low-rank promoting.**

I. Introduction

COMPRESSED sensing, a new paradigm for data acquisition and reconstruction, has drawn many researchers' attention over the past few years, e.g., [1]–[3]. Apart from sparsity, real-world signals often exhibit additional patterns, which can be exploited to improve the reconstruction performance. In this letter, we consider the problem of recovering block-sparse signals with intrablock correlated entries. Such signals arise in many practical scenarios. For example, in hyperspectral imaging, unmixed abundance maps of materials have block-sparse structures with correlated coefficients, since the materials generally occupy a continuous region of the scene [4]. Also, in frequency-hopping systems, the slowly time-varying frequency results in a block-sparse time-frequency signal with highly correlated coefficients [5]. A number of algorithms have been proposed for recovering block-sparse signals over the past few years, e.g., [6]–[11]. Specifically, a block-sparse Bayesian learning (BSBL) method was developed in [7], in which the sparse signal is partitioned into predefined blocks and the entries in a block are assumed to have the same sparsity pattern. To promote the block-sparsity, a common sparsity-controlling parameter is assigned to the coefficients in the same block, and a covariance matrix is employed to capture the intrablock correlation. In another two works [8], [9], to promote the intrablock smoothness, a specially designed covariance matrix is employed in the Gaussian prior placed on each block of the sparse signal. However, these methods require the knowledge of the block partition (i.e., the sizes and locations of the blocks) *a priori*, which is usually unknown in practice. To deal with this issue, Zhang and Rao [7] introduced an expanded block-sparse Bayesian learning (EBSBL) method, in which the sparse signal is decomposed into a number of overlapping blocks, and entries in each block are assumed to share a common sparsity pattern such that the BSBL method can be applied. This issue was also addressed in [10], where a pattern-coupled sparse Bayesian learning (PC-SBL) method was proposed via sparsity-controlling hyperparameter association among neighboring coefficients.

In this letter, we develop a LOcalized lOw-rank Promoting method (LOOP) for recovery of block-sparse signals with intrablock correlated entries. Specifically, coefficients of the sparse signal are organized to form a number of 2×2 matrices, and a log-determinant function is used to promote the low rankness of these matrices. We show that this log-determinant function has the potential to promote block-sparsity and local smoothness simultaneously, without requiring the knowledge of the block partition of the sparse signal. Experiment results show that the proposed method presents superior performance in recovering block-sparse signals with correlated entries.

Notations: The notation $|\cdot|$ denotes the determinant of a matrix or the absolute value of a scalar. $\|\boldsymbol{x}\|_2$ denotes the ℓ_2 norm of vector $\boldsymbol{x}$. The superscript $(\cdot)^T$ represents the transpose. $\text{Tr}(\boldsymbol{X})$ denotes the trace of a square matrix $\boldsymbol{X}$.

II. Problem Formulation

We consider the problem of recovering a block-sparse signal $\boldsymbol{x} \in \mathbb{R}^{N\times 1}$ from an underdetermined system

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{w} \quad (1)$$

where $\boldsymbol{y} \in \mathbb{R}^{M\times 1}$, $\boldsymbol{A} \in \mathbb{R}^{M\times N} (M < N)$, and $\boldsymbol{w} \in \mathbb{R}^{M\times 1}$ denote the measurements, the sensing matrix, and the noise, respectively. The signal $\boldsymbol{x}$ has a block-sparse structure with intrablock correlated coefficients, but the block partition is unknown.

We first form a matrix $\boldsymbol{X} \triangleq [0, \boldsymbol{x}^T; \boldsymbol{x}^T, 0]^T \in \mathbb{R}^{(N+1)\times 2}$. Fig. 1 depicts the transpose of $\boldsymbol{X}$. Let $\boldsymbol{X}_i$ $(1 \leq i \leq N)$ denotes a 2×2 matrix, which consists of the ith and $(i+1)$th

Manuscript received June 27, 2016; revised August 8, 2016; accepted August 9, 2016. Date of publication August 11, 2016; date of current version August 31, 2016. This work was supported in part by the National Science Foundation of China under Grant 61428103, Grant 61522104, and Grant U1530154, the National Science Foundation under Grant ECCS-1408182 and Grant ECCS-1609393, and the Air Force Office of Scientific Research under Grant FA9550-16-1-0243. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Deanna Needell.

L. Yang and J. Fang are with the National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: yanglinxiao0929@hotmail.com; JunFang@uestc.edu.cn).

H. Li is with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030 USA (e-mail: Hongbin.Li@stevens.edu).

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

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

Digital Object Identifier 10.1109/LSP.2016.2599525

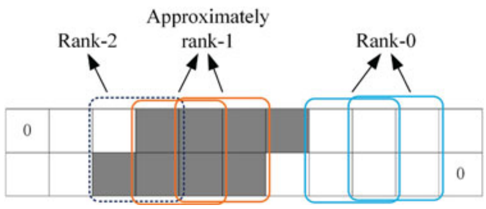

Fig. 1. Example of the transposition of matrix $\boldsymbol{X}$. The shadow and blank square blocks represent the nonzero and zero coefficients, respectively.

rows of $\boldsymbol{X}$, i.e.,

$$\boldsymbol{X}_i \triangleq \begin{bmatrix} x_{i-1} & x_i \\ x_i & x_{i+1} \end{bmatrix}. \tag{2}$$

Clearly, if x_{i-1}, x_i, and x_{i+1} are all equal to zero, then $\boldsymbol{X}_i$ has a zero rank. On the other hand, if x_{i-1}, x_i, and x_{i+1} are nonzero and locally smooth, then $\boldsymbol{X}_i$ is an approximately rank-1 matrix. Inspired by this observation, we seek a block-sparse and locally smooth solution $\boldsymbol{x}$ by promoting low rankness of the matrices $\{\boldsymbol{X}_i\}$. More precisely, the problem can be formulated as

$$\min_{\boldsymbol{x}} \quad \sum_{i=1}^{N} \operatorname{rank}(\boldsymbol{X}_i)$$
$$\text{s.t.} \quad \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 \leq \varepsilon \tag{3}$$

where ε is an error tolerance parameter related to the noise statistics. It is well known that rank minimization is an NP-hard problem. To circumvent this difficulty, we replace $\operatorname{rank}(\boldsymbol{X}_i)$ with a computationally tractable low-rank promoting function $\log|\boldsymbol{X}_i\boldsymbol{X}_i^T| = 2\sum_j \log\nu_j$, where ν_j denotes the jth singular value of $\boldsymbol{X}_i$. Note that this low-rank promoting function was also used in [12] and [13] to encourage low-rank solutions. Replacing $\operatorname{rank}(\boldsymbol{X}_i)$ in (3) with the log-determinant function, we arrive at

$$\min_{\boldsymbol{x}} \quad \sum_{i=1}^{N} \log|\boldsymbol{X}_i\boldsymbol{X}_i^T + \boldsymbol{E}|$$
$$\text{s.t.} \quad \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 \leq \varepsilon \tag{4}$$

where $\boldsymbol{E}$ is a positive-definite matrix to ensure the logarithmic function is well defined. Specifically, we set

$$\boldsymbol{E} \triangleq \epsilon \begin{bmatrix} 1 & \beta \\ \beta & 1 \end{bmatrix} \tag{5}$$

where ϵ is a small positive value, and $\beta(-1 < \beta < 1)$ is a parameter whose choice will be discussed later. The optimization (4) can eventually be formulated as an unconstrained optimization problem as follows:

$$\min_{\boldsymbol{x}} L(\boldsymbol{x}) = \sum_{i=1}^{N} \log|\boldsymbol{X}_i\boldsymbol{X}_i^T + \boldsymbol{E}| + \lambda\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 \tag{6}$$

where λ is a parameter controlling the tradeoff between the low-rankness and the fitting error.

Discussions: To gain insight into our proposed method, we provide an alternative view of the log-determinant function and show its ability to promote block-sparsity and local smoothness of the solution. The 2×2 matrix $\boldsymbol{X}_i\boldsymbol{X}_i^T$ can be written as

$$\boldsymbol{X}_i\boldsymbol{X}_i^T = \begin{bmatrix} x_{i-1}^2 + x_i^2 & x_{i-1}x_i + x_ix_{i+1} \\ x_{i-1}x_i + x_ix_{i+1} & x_i^2 + x_{i+1}^2 \end{bmatrix}. \tag{7}$$

Hence, the log-determinant of $\boldsymbol{X}_i\boldsymbol{X}_i^T$ is given by

$$\log|\boldsymbol{X}_i\boldsymbol{X}_i^T| = \log((x_{i-1}^2 + x_i^2)(x_i^2 + x_{i+1}^2))$$
$$+ \log\left(1 - \frac{(x_{i-1}x_i + x_ix_{i+1})^2}{(x_{i-1}^2 + x_i^2)(x_i^2 + x_{i+1}^2)}\right). \tag{8}$$

Let θ_i denote the angle between the two vectors $[x_{i-1}, x_i]^T$ and $[x_i, x_{i+1}]^T$. The second term on the right-hand side of (8) can be rewritten as

$$\log\left(1 - \frac{(x_{i-1}x_i + x_ix_{i+1})^2}{(x_{i-1}^2 + x_i^2)(x_i^2 + x_{i+1}^2)}\right) = \log(1 - \cos^2\theta_i)$$
$$= 2\log|\sin\theta_i|. \tag{9}$$

Consequently, we have

$$\sum_{i=1}^{N} \log|\boldsymbol{X}_i\boldsymbol{X}_i^T| = 2\sum_{i=2}^{N} \log(x_{i-1}^2 + x_i^2) + 2\sum_{i=1}^{N} \log|\sin\theta_i|$$
$$+ \log(x_1^2) + \log(x_N^2). \tag{10}$$

The first term on the right-hand side of (10) is a log-sum function with overlapping elements. This element-overlapping log-sum function, similar to the overlapping group LASSO [14], promotes block-sparsity of the solution. The second term on the right-hand side of (10) favors solutions with small values of $|\sin\theta_i|, \forall i$, and thus has the potential to encourage local smoothness of the solution.

III. Proposed Iterative Reweighted Algorithm

In this section, we solve the optimization (6) using the majorization–minimization (MM) approach [15]. The MM approach suggests that instead of directly minimizing the objective function, one can downhill the objective function by iteratively minimizing a simple upper-bound. It was shown [12], [13] that a surrogate function for the log-determinant function is given by

$$\log|\boldsymbol{X}_i\boldsymbol{X}_i^T + \boldsymbol{E}|$$
$$\leq \frac{1}{2}\operatorname{Tr}((\boldsymbol{X}_i\boldsymbol{X}_i^T + \boldsymbol{E})\boldsymbol{\Phi}_i^{(t)}) + \log|(\boldsymbol{\Phi}_i^{(t)})0^{-1}| - 1 \tag{11}$$

where $\boldsymbol{\Phi}_i^{(t)} \triangleq (\boldsymbol{X}_i^{(t)}(\boldsymbol{X}_i^{(t)})^T + \boldsymbol{E})0^{-1}$, and the equality is attained when $\boldsymbol{X}_i = \boldsymbol{X}_i^{(t)}$. Consequently, we have

$$\sum_{i=1}^{N} \log|\boldsymbol{X}_i\boldsymbol{X}_i^T + \boldsymbol{E}| \tag{12}$$

$$\leq \sum_{i=1}^{N} \left(\frac{1}{2}\operatorname{Tr}((\boldsymbol{X}_i\boldsymbol{X}_i^T + \boldsymbol{E})\boldsymbol{\Phi}_i^{(t)}) + \log|(\boldsymbol{\Phi}_i^{(t)})o^{-1}|\right) - N \tag{13}$$

$$\triangleq f(\boldsymbol{x}|\boldsymbol{x}^{(t)}). \tag{14}$$

Let

$$\boldsymbol{\Phi}_i^{(t)} \triangleq \begin{bmatrix} a^i & b^i \\ c^i & d^i \end{bmatrix}. \tag{15}$$

Substituting (2) and (15) into (13), we arrive at

$$f\big(\boldsymbol{x}|\boldsymbol{x}^{(t)}\big) = \boldsymbol{x}^T \boldsymbol{W} \boldsymbol{x} + \sum_{i=1}^{N} \big(\log|\boldsymbol{\Phi}_i^{(t)}|\big) - N \tag{16}$$

where $\boldsymbol{W} \in \mathbb{R}^{N\times N}$ is a tridiagonal matrix with its main diagonal, the first diagonal below and above the main diagonal given, respectively, as $W_{i,i} = \frac{1}{2}(a^i + d^i + d^{i-1} + a^{i+1})$, $W_{i+1,i} = \frac{1}{2}(c^i + c^{i+1})$, $W_{i,i+1} = \frac{1}{2}(b^i + b^{i+1})$, where we set $d^0 = a^{N+1} = 0$.

Eventually, a surrogate function majorizing the objective function in (6) can be written as

$$Q\big(\boldsymbol{x}|\boldsymbol{x}^{(t)}\big) = \boldsymbol{x}^T \boldsymbol{W} \boldsymbol{x} + \sum_{i=1}^{N} \big(\log|\boldsymbol{\Phi}_i^{(t)}|\big) + \lambda\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 - N. \tag{17}$$

Hence, the optimization (6) can be solved by iteratively minimizing (17) whose optimal solution is given by

$$\boldsymbol{x} = \big(\boldsymbol{A}^T\boldsymbol{A} + \lambda^{-1}\boldsymbol{W}\big)^{-1}\boldsymbol{A}^T\boldsymbol{y}. \tag{18}$$

Through iteratively minimizing $Q(\boldsymbol{x}|\boldsymbol{x}^{(t)})$, the objective function $L(\boldsymbol{x})$ is guaranteed to be nonincreasing at each iteration.

Remark: We discuss the choice of ϵ and β. Similar to [16], a monotonically decreasing sequence $\epsilon^{(t)}$, instead of a constant ϵ, is used such that the algorithm is less likely to get stuck in undesirable local minima. Regarding the choice of β, our experimental results suggests that a large β (say, $\beta = 0.9$) leads to a more smooth solution. We can also update β at each iteration according to the following rule:

$$\beta = H_{1,2}/H_{1,1} \tag{19}$$

where $\boldsymbol{H} \triangleq \boldsymbol{X}^T\boldsymbol{X}$, and $H_{i,j}$ denotes the (i,j)th entry of $\boldsymbol{H}$. The rationale behind this update rule is that the ratio of (19) offers an estimate of the average correlation or smoothness of the sparse signal.

For clarity, we summarize our algorithm as follows.

Algorithm 1: Localized Low-Rank Promoting Method.

Input: $\boldsymbol{y}$, $\boldsymbol{A}$, λ and β
Output: $\boldsymbol{x}$
Select an initialization $\boldsymbol{x}^{(0)}$, and set $t = 0$.
while not converged **do**
  Calculate $\{\boldsymbol{\Phi}_i^{(t)}\}$ and form $\boldsymbol{W}$.
  Compute a new estimate of the sparse signal, denoted as $\boldsymbol{x}^{(t+1)}$, via (18).
  if $\|\boldsymbol{x}^{(t+1)} - \boldsymbol{x}^{(t)}\|_2 \le \sqrt{\epsilon^{(t)}}/10$ **then**
    $\epsilon^{(t+1)} = \epsilon^{(t)}/10$
  end if
  $t = t + 1$
end while

IV. Simulation Results

We present experimental results to illustrate the performance of our proposed method, which is referred to as a LOOP. Throughout our experiments, λ is set to 10^{12} and β is updated according to (19). The reason we choose a large λ is that the log-determinant term in (6) could become arbitrarily negatively large. Hence, we need to choose a sufficiently large λ to guarantee a small data fitting error. For our proposed algorithm, we continue the iterative process until the difference between the reconstructed signals of successive iterations is negligible, i.e., $\|\boldsymbol{x}^{(t+1)} - \boldsymbol{x}^{(t)}\|_2 < 10^{-7}$. We compare our method with several existing state-of-the-art block-sparse signal recovery methods, namely, the BSBL [7], the EBSBL [7], the PC-SBL [10], and a total-variation-based sparse linear regression method (TV-SLR) [5], which exploits the sparsity of the signal and the smoothness of intrablock entries. The problem of the TV-SLR can be cast as

$$\min_{\boldsymbol{x}} \quad \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \mu_1\|\boldsymbol{x}\|_1 + \mu_2\|\boldsymbol{D}\boldsymbol{x}\|_1 \tag{20}$$

where $\boldsymbol{D}$ is the first-order differential matrix, and μ_1 and μ_2 are parameters set equal to $\lambda_1^*/10$ and $\lambda_2^*/10$, respectively. Here, λ_1^* and λ_2^* are defined in [5, Propositions 1 and 2], respectively.

We examine the recovery performance of respective algorithms on synthetic data. We generate an N-dimensional block-sparse signal, where $N = 400$. The signal is partitioned into a number of blocks, and the size of each block is randomly selected from the set {3,4,5}. Among these blocks, we randomly choose a few blocks as nonzero blocks, which totally contain K nonzero entries. The entries in the ith nonzero block are Gaussian random variables with mean μ_i and variance σ^2, where $\{\mu_i\}$ are randomly generated according to a normal distribution. The measurement matrix $\boldsymbol{A}$ is randomly generated with each entry independently drawn from a normal distribution, and then columns of $\boldsymbol{A}$ are normalized to unit norm. Note that both the BSBL and the EBSBL need to prespecify a block size parameter h, which is carefully chosen to be $h = 3$ in our experiments. Fig. 2 depicts the phase transition curves of respective algorithms under different choices of σ ($\sigma = 0.01$ and $\sigma = 0.2$). The phase transition illustrates the algorithm's success rate under different sparsity level (defined as K/M) and indeterminacy (defined as M/N). Each point in Fig. 2 corresponds to a success rate greater than or equal to 99%. The success rate is computed as the ratio of the number of successful trials to the total number of independent runs. A trial is considered successful if $\|\hat{\boldsymbol{x}} - \boldsymbol{x}\|_2/\|\boldsymbol{x}\|_2 < 10^{-2}$, where $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$ denote the true signal and estimated one, respectively. From Fig. 2, we see that the proposed algorithm presents a substantial performance advantage over the other methods when the intrablock entries are highly correlated (i.e., $\sigma = 0.01$). Note that although the TV-SLR also exploits the intrablock correlation, it performs rather poorly as compared with the LOOP, which suggests that the localized low rank promoting function is more effective than the TV-based regularizer in encouraging block-sparse and locally smooth signals. When $\sigma = 0.2$, the intrablock correlation becomes weak. In this case, both the LOOP and the TV-SLR experience some

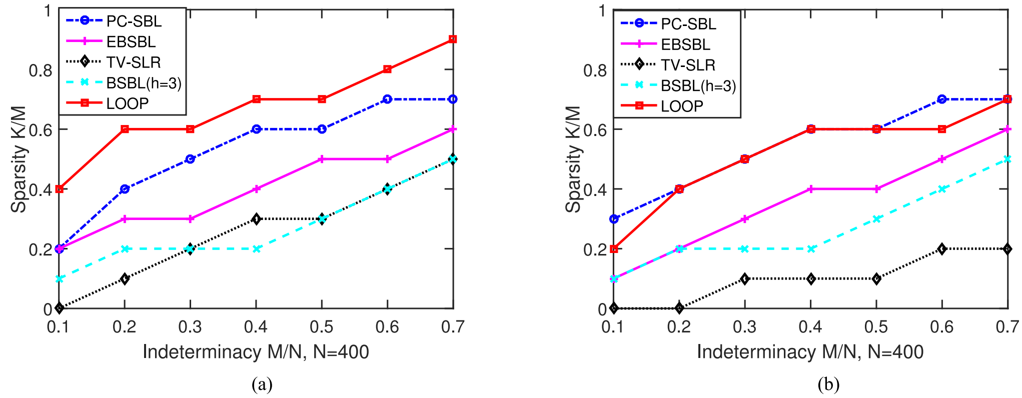

Fig. 2. Empirical 99% phase transitions of respective algorithms, $N = 400$. From left to right: $\sigma = 0.01$ and $\sigma = 0.2$.

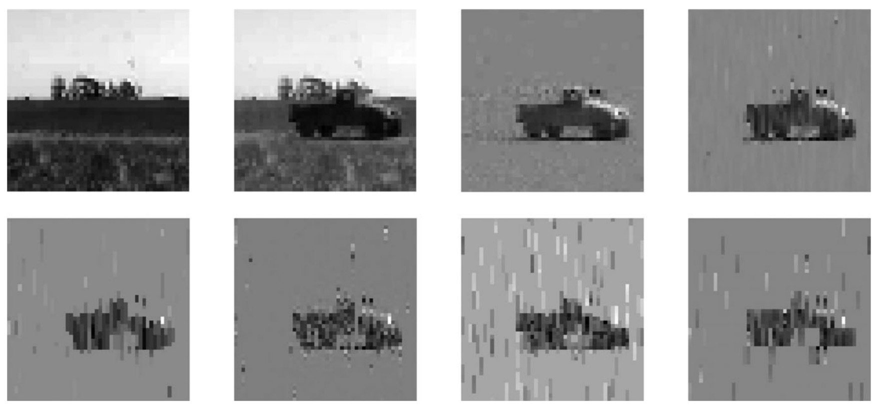

Fig. 3. Top row (from left to right): the background image, the test image (30th frame of the Convoy2 dataset), the true foreground image, the recovered foreground image by LOOP. Bottom row (from left to right): the recovered foreground images by TV-SLR, PC-SBL, EBSBL-BO, and BSBL-BO, respectively.

performance degradation. Nevertheless, the proposed method LOOP still outperforms the BSBL and the EBSBL methods.

We next apply the algorithms to solve the background subtraction problem that can be formulated as a sparse signal recovery problem [17]. Let $\boldsymbol{x}_t$ and $\boldsymbol{x}_b$ denote the test and the background images, respectively. Our object is to reconstruct the foreground image $\boldsymbol{x}_f \triangleq \boldsymbol{x}_t - \boldsymbol{x}_b$ from the compressed measurements $\boldsymbol{y}_t$ and $\boldsymbol{y}_b$, i.e.,

$$\boldsymbol{y}_f = \boldsymbol{y}_t - \boldsymbol{y}_b = \boldsymbol{A}(\boldsymbol{x}_t - \boldsymbol{x}_b) = \boldsymbol{A}\boldsymbol{x}_f. \tag{21}$$

The Convoy2 dataset [18] is used, which contains a video with 260 frames and a background image. The video has a dynamic block-sparse foreground as cars enter and exit the field of view over time. In our experiments, the original images of 480×381 pixels are resized to 50×50 pixels. The measurement matrix is of size 500×2500 and its elements is generated from a normal distribution. For both the BSBL and the EBSBL, the block size h is set to 5. For fair comparison, the 1-D PC-SBL of [10], instead of the 2-D PC-SBL developed in [19], was considered. The 30th frame is simulated and the reconstructed foreground images are shown in Fig. 3. From Fig. 3, we see that LOOP provides a finer foreground image quantity than the other methods. We then choose multiple frames from the 10th to 40th of the video as the test images. Fig. 4 shows the normalized mean squared errors (NMSEs) obtained by respective algorithms. The results are averaged over 50 independent trials. We see that our

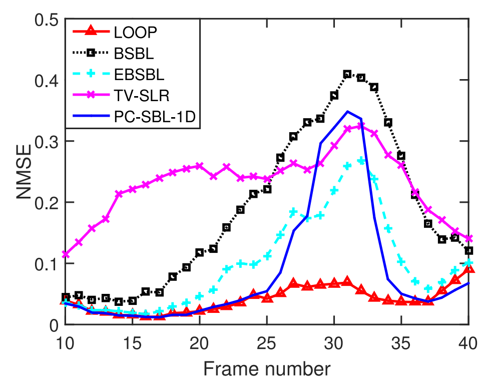

Fig. 4. NMSE versus frame number.

proposed method achieves the best estimation accuracy for most of frames.

V. Conclusion

We proposed a LOOP for recovering block-sparse signals with intrablock correlated entries. Consecutive elements of the signal were organized into a number of 2×2 matrices, and a log-determinant function was employed to promote block-sparsity and local smoothness of the sparse signal simultaneously. An iterative reweighted method was developed by iteratively minimizing a surrogate function of the original objective function. Numerical results show that the proposed method presents superior performance in recovery block-sparse signals with intrablock correlated entries.

REFERENCES

[1] D. L. Donoho, "Compressive sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[2] E. Candés and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[3] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Rev.*, vol. 51, no. 1, pp. 34–81, 2009.

[4] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Total variation spatial regularization for sparse hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4484–4502, Nov. 2012.

[5] D. Angelosante, G. B. Giannakis, and N. D. Sidiropoulos, "Estimating multiple frequency-hopping signal parameters via sparse linear regression," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5044–5056, Oct. 2010.

[6] D. P. Wipf and B. D. Rao, "An empirical Bayesian strategy for solving the simultaneous sparse approximation problem," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3704–3716, Jul. 2007.

[7] Z. Zhang and B. D. Rao, "Extension of SBL algorithms for the recovery of block sparse signals with intra-block correlation," *IEEE Trans. Signal Process.*, vol. 61, no. 8, pp. 2009–2015, Apr. 2013.

[8] M. Luessi, S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Bayesian simultaneous sparse approximation with smooth signals," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5716–5729, Nov. 2013.

[9] Z. Chen, R. Molina, and A. K. Katsaggelos, "Robust recovery of temporally smooth signals from under-determined multiple measurements," *IEEE Trans. Signal Process.*, vol. 63, no. 7, pp. 1779–1791, Apr. 2015.

[10] J. Fang, Y. Shen, H. Li, and P. Wang, "Pattern-coupled sparse Bayesian learning for recovery of block-sparse signals," *IEEE Trans. Signal Process.*, vol. 63, no. 2, pp. 360–372, Jan. 2015.

[11] Y. C. Eldar and M. Mishali, "Robust recovery of signals from a structured union of subspaces," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5302–5316, Nov. 2009.

[12] K. Mohan and M. Fazel, "Iterative reweighted least squares for matrix rank minimization," in *Proc. IEEE 48th Annu. Allerton Conf. Commun., Control, Comput.*, 2010, pp. 653–661.

[13] D. Wipf, "Non-convex rank minimization via an empirical Bayesian approach," *arXiv preprint, arXiv:1207.2440*, 2012.

[14] L. Yuan, J. Liu, and J. Ye, "Efficient methods for overlapping group lasso," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 9, pp. 2104–2116, Sep. 2013.

[15] D. R. Hunter and K. Lange, "A tutorial on MM algorithms," *Amer. Statist.*, vol. 58, no. 1, pp. 30–37, 2004.

[16] R. Chartrand and W. Yin, "Iteratively reweighted algorithms for compressive sensing," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Las Vegas, NV, USA, Mar. 2008, pp. 3869–3872.

[17] V. Cevher, A. Sankaranarayanan, M. F. Duarte, D. Reddy, R. G. Baraniuk, and R. Chellappa, "Compressive sensing for background subtraction," in *Proc. 10th Eur. Conf. Comput. Vision*, Marseille, France, Oct. 2008, pp. 155–168.

[18] G. Warnell, D. Reddy, and R. Chellappa, "Adaptive rate compressive sensing for background subtraction," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Kyoto, Japan, Mar. 2012, pp. 1477–1480.

[19] J. Fang, L. Zhang, and H. Li, "Two-dimensional pattern-coupled sparse Bayesian learning via generalized approximate message passing," *IEEE Trans. Image Process.*, vol. 25, no. 6, pp. 2920–2930, Jun. 2016.