

Nonconvex is Attractive: $L_{2/3}$ Regularized Thresholding Algorithm using Multiple Sub-Dictionaries

Yunyi Li¹, Yue Hao², Fei Dai², Yue Yin², Shangang Fan², Jie Yang², Guan Gui², Fumiyuki Adachi³

¹College of Electronic and Optical Engineering & College of Microelectronics, Nanjing University of Posts and Telecommunications, Nanjing 210023, China

²College of Telecommunication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

³Research Organization of Electrical Communication, Tohoku University, Sendai 980-8577, Japan
E-mail: guiguan@njupt.edu.cn

Abstract—The $L_{2/3}$ -regularization is a typical nonconvex and nonsmooth optimization method, which can obtain more powerful performance than L_1 regularization in some applications, such as computational imaging, sparse signal recovery and low-rank matrix completion, etc. This paper proposes an adaptive iteratively-weighted thresholding algorithm for $L_{2/3}$ -regularized problem based on the multiple analysis sub-dictionaries (MD) sparsifying transform strategy, the MD strategy can be employed to further exploit the prior knowledge of estimated signal for sparse recovery. What's more, we propose an adaptive updating scheme for regularization parameter to weight the contribution of each sub-dictionary. Experiments confirm that the proposed method could obtain higher image quality and achieve faster convergence than some corresponding prior work.

Index terms—Nonconvex; $L_{2/3}$ regularization; multiple sub-dictionaries sparsifying transform; iteratively thresholding; image restoration.

I. INTRODUCTION

Compressed sensing (CS) is a sparse signal acquisition and recovery technology which has been widely used in image processing [1]–[3], wireless communications [4]–[8], vehicular systems [9] and the Internet of Thing [10]. Among such methodologies, the regularized methods are certainly the most successful to reconstruct sparse signal, such as the well-known L_1 -norm regularization (Lasso) [11]. Recently, the nonconvex regularization methods have attracted more and more attentions because of the unbiased property and hence can obtain a sparser solution, typically the L_p ($0 < p < 1$)-norm. Unlike the convex regularized algorithms, this kind of nonconvex and nonsmooth optimization problem cannot be solved effectively and efficiently. Among the existing algorithms, the iterative thresholding algorithm (ITA) is one of the most effective and efficient CS reconstruction methods to solve such L_p ($0 < p < 1$) regularization optimization problems.

Existing pioneering work shows that the $L_{2/3}$ regularization model is a typical representative among L_p ($0 < p < 1$) regularization and can obtain a superior performance than L_1 [12]. Sparse representation (SR) theory shows that a given signal $\mathbf{x} \in R^{N \times 1}$ can be represented by different dictionary, and

motivated by the sparse representation (SR) strategy, some new sparsifying transform strategies have emerged to enhance the sparsity for L_1 -norm regularization problems by exploiting the sparsity structure of signals and images, and obtain better performance than previous algorithms [13][14]. A wavelet-based sparsity-induced composite regularization method (Co-L1) has been proposed for the well-known LASSO problem [13], which can improve the reconstruction performance significantly by weighting the composite regularization iteratively and respectively. In addition, a shearlet-based multiple level sparse representation algorithm was also proposed for enhancing the image reconstruction quality [5][6], which is adapted to the structure of multiscale coefficients by incorporating the iteratively reweighted shrinkage step. However, these work only focus on the L_1 -norm regularized optimization problem.

To further improve the recovery performance in sparse recovery of the nonconvex $L_{2/3}$ regularization, this paper focus on investigating the multiple sub-dictionaries (MD) sparsifying transform strategy to exploit more prior knowledge for the $L_{2/3}$ regularized optimization problem. Experimental results show that the MD sparsifying transform strategy based algorithms combining with the prior work can significantly improve the recovery performance.

The rest of this paper is organized as follows. In Sect. II, the problem formulation is given. Then, the proposed algorithm is presented in Section III. Experiment results are conducted in Section IV. Finally, Section IV offers some concluding remarks.

II. PROBLEM FORMULATION

Suppose a given signal $\mathbf{x} \in R^{N \times 1}$ is not sparse, if we conduct a transform $\Psi\mathbf{x}$ by a given analysis dictionary $\Psi \in R^{N \times N}$, then the transform coefficients $\Psi\mathbf{x}$ is sparse. The shearlet transform [16] and the wavelet transform [17] are two typical transforms for sparsifying transform. In this paper, we particularly choose the wavelet transform as the idea transform because of it is very effective to compress the natural images and the wavelet coefficients of \mathbf{x} are truly sparse. It is well known that the structure of $\Psi\mathbf{x}$ will vary with different given analysis dictionaries Ψ , such as the sparsity. It is well known

that the sparsity is often regarded as prior knowledge for optimization problem. Hence, as a motive, we can utilize the MD strategy to further exploit the sparse structure of estimated signal as prior knowledge for $L_{2/3}$ -norm regularized sparse recovery problem. Here are the main steps:

Firstly, we construct the analysis dictionary (or sparsifying transform matrix) Ψ by

$$\Psi = \begin{bmatrix} \Psi_1 \\ \Psi_2 \\ \vdots \\ \Psi_i \\ \vdots \\ \Psi_{nN} \end{bmatrix} \in R^{(nN) \times N}, \quad (1)$$

where $\Psi_i = [\psi_{i1}, \psi_{i2}, \dots, \psi_{iN}] \in R^{1 \times N}$, $i = 1, \dots, nN$ denotes the i -th row of Ψ .

Secondly, we select L_d rows form Ψ as the sub-dictionaries Ψ_d , $d = 1, 2, \dots, D$ by the form

$$\Psi_d = \begin{bmatrix} \Psi_{L_{d-1}+1} \\ \Psi_{L_{d-1}+2} \\ \vdots \\ \Psi_{L_d} \end{bmatrix} \in R^{L_d \times N}, \quad (2)$$

where L_d , $d = 1, 2, \dots, D$ denote the number of the rows of each sub-dictionaries Ψ_d , and D represent the number of sub-dictionaries, then

$$L_1 + L_2 + \dots + L_d = nN, \quad d = 1, 2, \dots, D. \quad (3)$$

By splitting the matrix Ψ into several sub-dictionaries Ψ_d , that is to say, we convert the sparsifying transform $\Psi \mathbf{x}$ to several different $\Psi_d \mathbf{x}$, $d = 1, 2, \dots, D$, hence the sparse structures of $\Psi_d \mathbf{x}$ are different.

Thirdly, to make fully use of the sparse structures for sparse recovery, we proposed the MD based $L_{2/3}$ regularized optimization problem, which can be described as following

$$\hat{\mathbf{x}}_{l_{d,2/3}} = \arg \min_{\hat{\mathbf{x}}} \left\{ f(\mathbf{x}) = \|\Phi \mathbf{x} - \mathbf{y}\|_2^2 + \sum_{d=1}^D \lambda_{d,2/3} \|\Psi_d \mathbf{x}\|_{2/3}^{2/3} \right\} \quad (4)$$

where $\Phi \in R^{M \times N}$ ($M \ll N$) is a given measurement matrix, \mathbf{x} and \mathbf{y} denotes the raw signal and the measured signal respectively, Ψ_d ($d = 1, 2, \dots, D$) are the designed sub-dictionaries, and $\lambda_{d,2/3}$, $d = 1, 2, \dots, D$ denote a series of regularization parameters for each corresponding regularization term.

Hence, if we adjust these regularization parameters $\lambda_{1,2/3}$, $\lambda_{2,2/3}$, \dots , $\lambda_{D,2/3}$ to weight each term of the $L_{2/3}$ -norm based regularizer $\sum_{d=1}^D \lambda_{d,2/3} \|\Psi_d \mathbf{x}\|_{2/3}^{2/3}$, we will control the contribution of each sub-dictionary for sparse recovery. In this paper, we choose $L_1 = L_2 = \dots = L_d = N$, then $\Psi_d \in R^{N \times N}$ are orthogonal matrixes.

III. THE PROPOSED ALGORITHM

The major disadvantage of the $L_{2/3}$ minimization problem is nonconvex. Hence, it is difficult to be solved efficiently. We develop the MD based iteratively weighted $L_{2/3}$ regularization algorithm based on the existing algorithm [12][13].

A. The iterative thresholding algorithm for analysis $L_{2/3}$ regularization

To solve this problem, we first consider the analysis optimization problem as

$$\hat{\mathbf{x}}_{l_{2/3}} = \arg \min_{\hat{\mathbf{x}}} \left\{ f(\mathbf{x}) = \|\Phi \mathbf{x} - \mathbf{y}\|_2^2 + \lambda_{2/3} \|\Psi' \mathbf{x}\|_{2/3}^{2/3} \right\} \quad (5)$$

where $\Psi' \in R^{N \times N}$ denotes the sparse representation dictionary. Then the first order optimality condition of \mathbf{x} is described as

$$\nabla f(\mathbf{x}) = 2\Psi' \Phi^T (\Phi \mathbf{x} - \mathbf{y}) + \lambda_{2/3} \left(\nabla (\|\Psi' \mathbf{x}\|_{2/3}^{2/3}) \right), \quad (6)$$

The operator $\nabla(\cdot)$ denotes the gradient operator. Let $\nabla f(\mathbf{x}) = 0$, then

$$\Psi' \Phi^T (\mathbf{y} - \Phi \mathbf{x}) = \frac{1}{2} \nabla (\lambda_{2/3} \|\Psi' \mathbf{x}\|_{2/3}^{2/3}). \quad (7)$$

Adding $\Psi' \mathbf{x}$ and multiplying by any parameter τ in both sides of (18), then

$$\Psi' \mathbf{x} + \tau \Psi' \Phi^T (\mathbf{y} - \Phi \mathbf{x}) = \Psi' \mathbf{x} + \frac{\tau}{2} \nabla (\lambda_{2/3} \|\Psi' \mathbf{x}\|_{2/3}^{2/3}) \quad (8)$$

To this end, there is the resolvent operator [12] as

$$\mathbf{H}_{\lambda_{2/3}}(\cdot) = \left(\mathbf{I} + \frac{\lambda_{2/3}}{2} \nabla (\|\cdot\|_{2/3}^{2/3}) \right)^{-1}, \quad (9)$$

where λ is the regularization parameter. Then it implies

$$\begin{aligned} \Psi' \mathbf{x} &= \left(\mathbf{I} + \frac{\tau \lambda_{2/3}}{2} \nabla (\|\cdot\|_{2/3}^{2/3}) \right)^{-1} \left(\Psi' (\mathbf{x}^n + \tau \Phi^T (\mathbf{y} - \Phi \mathbf{x}^n)) \right) \\ &= H_{\lambda_{2/3}} \left(\Psi' (\mathbf{x}^n + \tau \Phi^T (\mathbf{y} - \Phi \mathbf{x}^n)) \right), \end{aligned} \quad (10)$$

where $\tau > 0$ controls the step size in each iteration (e.g., $\tau = 0.99 / \|\Phi\|_2^2$). Owing to $\Psi' \in R^{N \times N}$ are orthogonal matrix, thus we have

$$\begin{aligned} \mathbf{x}^{t+1} &= (\Psi')^{-1} H_{\lambda_{2/3}} \left(\Psi' (\mathbf{x}^t + \tau \Phi^T (\mathbf{y} - \Phi \mathbf{x}^t)) \right) \\ &= (\Psi')^{-1} H_{\lambda_{2/3}} (\theta(\mathbf{x}^t)), \end{aligned} \quad (11)$$

in which

$$\theta(\mathbf{x}) = \Psi' (\mathbf{x}^t + \tau \Phi^T (\mathbf{y} - \Phi \mathbf{x}^t)), \quad (12)$$

and the component-wise thresholding operator $H_{\lambda_{2/3}}(\cdot)$ is defined as

$$H_{\lambda_{2/3}}(\mathbf{x}) = (h_{\lambda_{2/3}}(x_1), h_{\lambda_{2/3}}(x_2), \dots, h_{\lambda_{2/3}}(x_N))^T, \quad (13)$$

in which, the $h_{\lambda_{2/3}}(x_i)$ is defined by:

$$h_{\lambda_{2/3}}(x_i) = \begin{cases} \varphi_{\lambda_{2/3}}(x_i), & |x_i| > T \\ 0, & \text{otherwise} \end{cases}, \quad (14)$$

where $T = \frac{2^{\frac{4\sqrt{3}}{3}}}{3}(\lambda_{2/3}\tau)^{3/4}$ is the threshold value, and

$$\varphi_{\lambda_{2/3}}(x_i) = \frac{1}{8} \left(\vartheta_{\lambda_{2/3}}(x_i) + \sqrt{\frac{2|x_i|}{|\vartheta_{\lambda_{2/3}}(x_i)|} - |\vartheta_{\lambda_{d,2/3}}(x_i)|^2} \right)^3 \text{sgn}(t), \quad (15)$$

where $\text{sgn}(\cdot)$ denotes sign function and $\vartheta_{\lambda_{d,2/3}}(x_i)$ is denoted as

$$\begin{aligned} & \vartheta_{\lambda_{2/3}}(x_i) \\ &= \frac{2}{\sqrt{3}} (\lambda_{2/3})^{1/4} (\cosh(\frac{1}{3} \text{arccosh}(\frac{27}{16} (\lambda_{2/3})^{-3/2} (x_i)^2)))^{1/2} \end{aligned} \quad (16)$$

The iterative thresholding algorithm for analysis $L_{2/3}$ can be described as in Table I.

TABLE I. THE ANALYSIS THRESHOLDING ALGORITHM FOR $L_{2/3}$ REGULARIZATION

<p>Algorithm 1: The analysis $L_{2/3}$ thresholding algorithm for solving</p> $\mathbf{x}^{t+1} = \arg \min \gamma \ \Phi \mathbf{x} - \mathbf{y}\ _2^2 + \lambda_{2/3} \ \Psi' \mathbf{x}\ _{2/3}^{2/3}$ <p>1: Input: The analysis dictionary Ψ', the measurement \mathbf{y}, the measurement matrix Φ; $\gamma = 1$; $\varepsilon = 0.01$; $\tau = \frac{1-\varepsilon}{\ \Phi\ _2^2}$; $\lambda_{2/3}$;</p> <p>2: Initialization: $t = 0$; $\mathbf{x}^0 = 0$;</p> <p>3: for $t = 1, 2, 3, \dots$;</p> <p style="padding-left: 20px;">While not converged do</p> <p style="padding-left: 40px;">Step 1: Compute $\boldsymbol{\theta}(\mathbf{x}^t) = \Psi'(\mathbf{x}^t + \tau \Phi^T(\mathbf{y} - \Phi \mathbf{x}^t))$;</p> <p style="padding-left: 40px;">Step 2: Apply the close-form thresholding operator to obtain solution:</p> $\mathbf{x}^{t+1} = (\Psi')^{-1} H_{\lambda_{2/3}}(\boldsymbol{\theta}(\mathbf{x}^t));$ <p style="padding-left: 40px;">Step 3: Update the value of μ using</p> $\mu^{t+1} = \frac{1 + \sqrt{1 + 4(\mu^t)^2}}{2};$ <p style="padding-left: 40px;">Step 4: Update the solution</p> $\mathbf{x}^{t+1} = \mathbf{x}^t + \frac{\mu^t - 1}{\mu^{t+1}} (\mathbf{x}^t - \mathbf{x}^{t-1});$ <p style="padding-left: 20px;">End</p> <p>4: End</p>
--

B. Updating iteratively weighting rules of $\lambda_{d,2/3}$

Comparing with the optimization problem between (4) and (5), a main improvement is the adaptively-updating regularization parameter. For the given $\lambda_{d,2/3}$, $d = 1, 2, \dots, D$, the problem (4) can be easily resolved by the iterative

thresholding algorithm for analysis $L_{2/3}$ in table I. As mentioned before, the parameters $\lambda_{d,2/3}$ play a crucial role in recovery performance, in order to weight the contribution of each sub-dictionary and make a tradeoff between error and prior knowledge, this paper proposes an adaptively-updating parameters $\lambda_{d,2/3}$ to weight the contribution of each regularizer by

$$\lambda_{d,2/3}^{(t+1)} = \frac{L_d}{(\varepsilon + \|\Psi_d \mathbf{x}^{(t)}\|_2^2)^{\frac{1-p}{2}}}, \quad p = \frac{2}{3}, \quad (17)$$

where L_d denotes the sub-dictionary size, $\varepsilon > 0$ is a small constant to prevent the denominator from zero, thus $\lambda_{d,2/3}$ will vary with Ψ_d adaptively, we could appropriately weight the contribution from every Ψ_d by adjusting $\lambda_{d,2/3}$. The proposed multiple sub-dictionary based iterative thresholding algorithm for $L_{2/3}$ regularized optimization problem can be described as in Table II.

TABLE II. THE PROPOSED MD- $L_{2/3}$ ALGORITHM

<p>Algorithm 2: The proposed MD-$L_{2/3}$ algorithm for solving</p> $\mathbf{x}^{t+1} = \arg \min \gamma \ \Phi \mathbf{x} - \mathbf{y}\ _2^2 + \sum_{d=1}^D \lambda_d \ \Psi_d \mathbf{x}\ _{2/3}^{2/3}$ <p>1: Input: The sub-wavelet-dictionaries $\{\Psi_d\}_{d=1}^D$, the measured data \mathbf{y}, the measurement matrix Φ; L_d; $\gamma = 1$; $\varepsilon = 10^{-3}$; $\varepsilon = 0.01$; $\tau = \frac{1-\varepsilon}{\ \Phi\ _2^2}$;</p> <p>2: Initialization: $t = 0$; $\mathbf{x}^0 = 0$; $\lambda_d^0 = 1$;</p> <p>3: for $t = 1, 2, 3, \dots$;</p> <p>4: Call algorithm in TABLE I:</p> $\mathbf{x}^{t+1} \leftarrow \arg \min \gamma \ \Phi \mathbf{x} - \mathbf{y}\ _2^2 + \sum_{d=1}^D \lambda_d \ \Psi_d \mathbf{x}\ _{2/3}^{2/3}$ <p>5: Compute $\lambda_{d,2/3}^{t+1} = \frac{L_d}{(\varepsilon + \ \Psi_d \mathbf{x}^t\ _2^2)^{\frac{1-p}{2}}}$, $p = \frac{2}{3}$ in (17);</p> <p>6: end</p> <p>7: Output \mathbf{x}^t.</p>
--

IV. EXPERIMENT STUDIES

To demonstrate the superior performance of the proposed algorithm, in this section, we evaluate its performance in compressive imaging problem under different scenarios. We first use the concatenating undecimated ‘*db1*’ and ‘*db2*’ wavelets as the basis function to construct the analysis dictionary $\Psi \in R^{8N \times N}$ for wavelet transform with 2-level decomposition, then we can obtain eight sub-dictionaries of $\Psi_1, \Psi_2, \dots, \Psi_8 \in R^{N \times N}$. Consider a real-value image $\mathbf{x} \in R^{N \times 1}$ (Image can be vectorizing into a one-dimensional vector), and $\Psi_d \mathbf{x}$ is sparse under the given dictionary Ψ_d . We use the ‘‘Spread Spectrum’’ operator as the measurement matrix Φ [18], so the typical measured data has the following linear form

$$\mathbf{y} = \Phi \mathbf{x} + \mathbf{n} \quad (18)$$

where \mathbf{n} denotes the additive noise. Our goal is to recovery \mathbf{x} from \mathbf{y} . We utilize the well-known ‘Cameraman’ figure for experiment, which is shown in Fig 1(a), in order to reduce the computation, we choose a part of the figure shown in (b). We

define the noise level in the measurement by $m\text{SNR} = \|\mathbf{y}\|_2^2 / (M\sigma^2)$, where M, σ^2 denotes the number of \mathbf{y} and the variance of the white Gaussian noise respectively. We evaluate the performance of all algorithms by the popular recovery SNR (RSNR) by $-20\log(\|\mathbf{x} - \hat{\mathbf{x}}\|_2 / \|\mathbf{x}\|_2)$, where $\hat{\mathbf{x}}$ denotes the estimated sparse image. The prior work of ‘Co-L1’ algorithm [13] and the SD based iterative $L_{2/3}$ thresholding algorithm [12] will be performed as comparisons. To make it fair, we choose the fixed regularization parameter for $L_{2/3}$ thresholding algorithm as

$$\lambda_{2/3} = \frac{1}{(\epsilon + \|\Psi\mathbf{x}\|_2^2)^{\frac{1-p}{2}}}, \quad p = \frac{2}{3}. \quad (19)$$

where $\Psi \in R^{8N \times N}$ is not orthogonal, and the resulting analysis sparse recovery problem can be resolved by an smoothing and decomposition method [19].

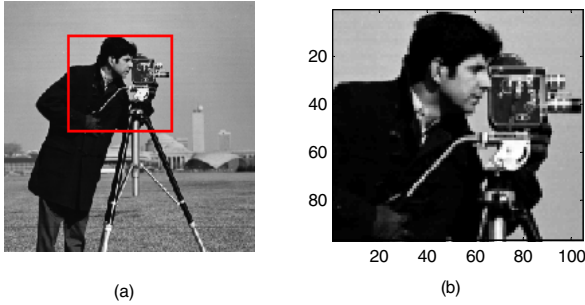


Fig. 1. Left: the Cameraman image of size $N = 256 \times 256$; Right: the selected part image of size $N = 96 \times 104$.

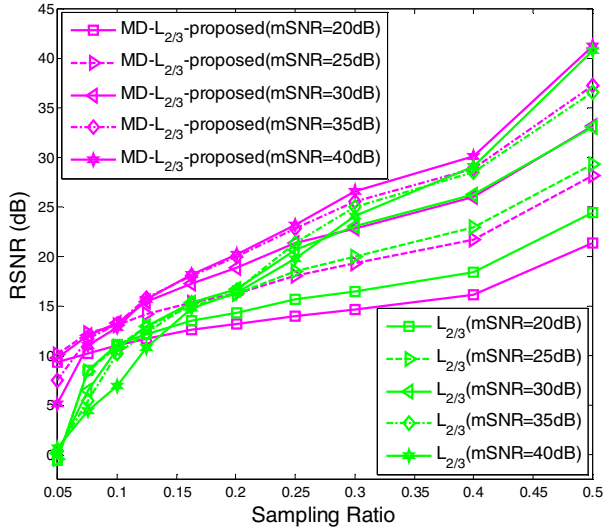


Fig. 2. The RSNR of proposed MD- $L_{2/3}$ algorithm and the $L_{2/3}$ algorithm versus sampling ratio with four mSNRs $\in \{20, 25, 30, 35, 40\}$ dB.

A. The Recovery SNR performances versus Sampling Ratio

In the subsection A, we will first evaluate the superiority and robustness of the proposed algorithm versus the sampling ratio M/N by considering the scenarios where the measurements are contaminated with noise, we set five measurements SNR levels by 20dB, 25dB, 30dB, 35dB and

40dB. Fig. 2 depict the RSNR of the proposed MD- $L_{2/3}$ algorithm and the $L_{2/3}$ iterative thresholding algorithm versus sampling ratio. The results show that the proposed MD based algorithm perform better than the SD based $L_{2/3}$ algorithm, especially under the scenarios of lower sampling ratios M/N and the stronger measurements noises, and the robustness of the proposed algorithm is better than prior work.

B. The recovery SNR performances versus measurement SNR

For our second experiment, we investigate the influence of different level noises on the proposed algorithm. We evaluate the performance by the RSNR versus lower mSNR (20dB~40dB) with four different sampling ratios. The results in Fig. 3 shown that the proposed MD based algorithm can obtain a higher RSNR, and has a better robustness than other two algorithms. With the increasing of the mSNR value, the SD based $L_{2/3}$ algorithm will deteriorate evidently, while the proposed MD- $L_{2/3}$ algorithm is still robust.

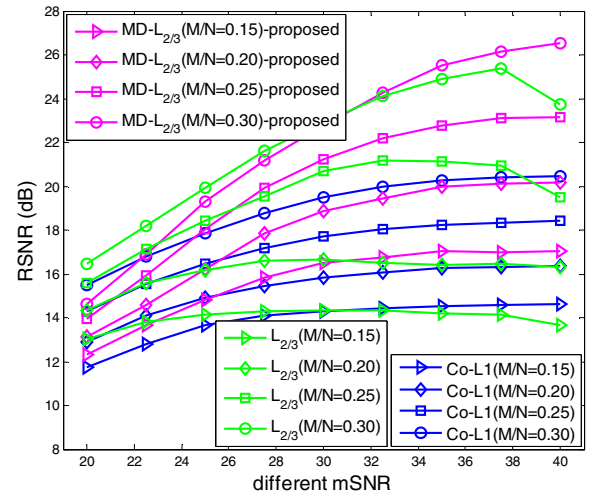


Fig. 3. The RSNR of proposed MD- $L_{2/3}$ algorithm, Co-L1 and the $L_{2/3}$ algorithm versus measurement SNR with $M/N \in \{0.15, 0.20, 0.25, 0.30\}$.

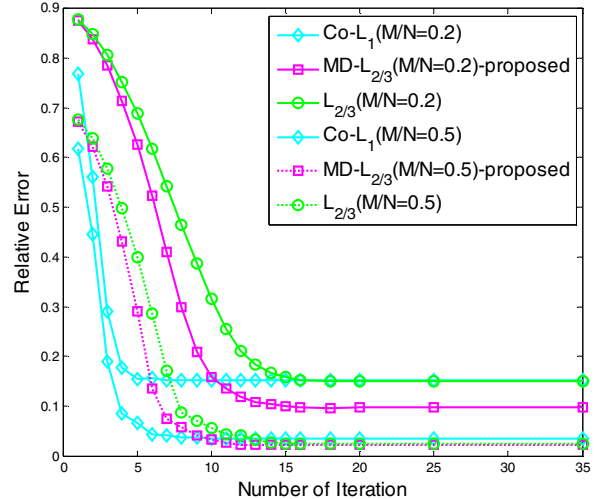


Fig. 4 The Relative Error of proposed MD- $L_{2/3}$, Co-L1 and $L_{2/3}$ algorithms versus iteration numbers with sampling ratio $M/N \in \{0.2, 0.5\}$.

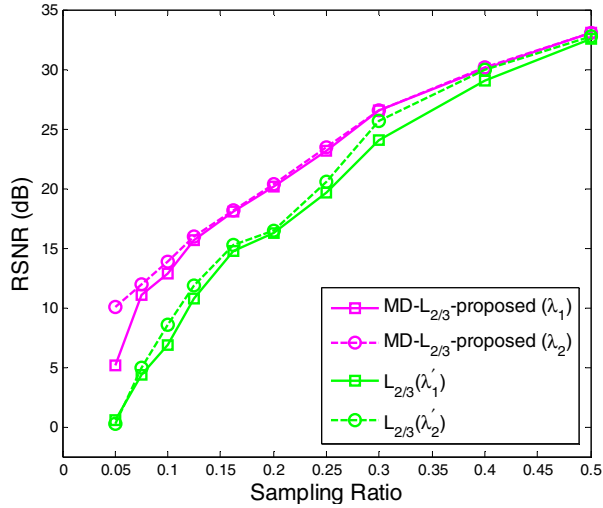


Fig. 5. The RSNR of proposed MD- $L_{2/3}$ algorithm and $L_{2/3}$ -algorithm versus Sampling ratio using two updating schemes for regularization parameter.

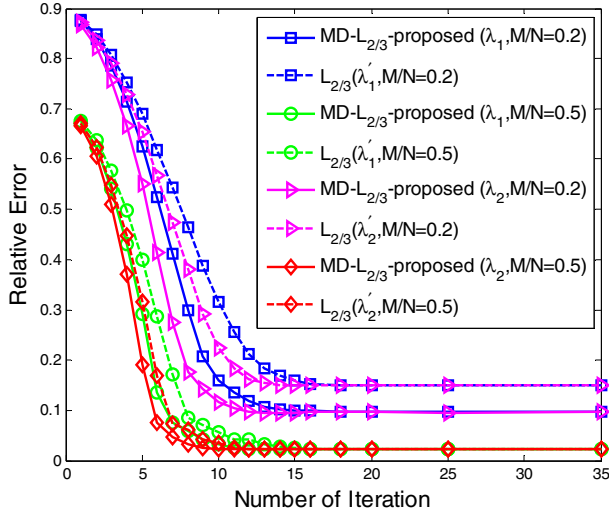


Fig. 6. The relative error of proposed MD- $L_{2/3}$ algorithm and $L_{2/3}$ algorithms versus iteration numbers using two updating schemes for regularization parameter.

C. The relative error performances versus the number of iterations

We study the convergence and the reconstruction errors versus every number of iterations via the relative error performances, the relative error formula is given by $\|\mathbf{x} - \hat{\mathbf{x}}\|_2 / \|\mathbf{x}\|_2$, where $\hat{\mathbf{x}}$ denotes the estimated sparse image. Fig. 4 shows the relative error of three algorithms versus the number of iteration with two sampling ratios. The result shows that the proposed algorithm converges faster than the SD based $L_{2/3}$ algorithm. Moreover, the relative error of the proposed algorithms are significant smaller than the SD based algorithm when the sampling ratio is 0.2. Compared to Co-L1 [13], our proposed MD based $L_{2/3}$ algorithm can obtain the lower

relative error when the sampling ratio $M/N \in \{0.2, 0.5\}$ significantly, while be at a disadvantage on the convergence.

D. Performance evaluation using different regularization parameters

In the fourth experiment, we evaluate the reconstruction performance of the proposed MD based $L_{2/3}$ algorithms using another adaptively-updating scheme for the regularization parameters,

$$\lambda_{d,2/3}^{t+1} = \frac{L_d}{(\epsilon + \|\Psi_d \mathbf{x}^t\|_{2/3}^{2/3})} \quad (20)$$

The corresponding fixed regularization parameter for $L_{2/3}$ regularization algorithm is

$$\lambda_{2/3} = \frac{1}{(\epsilon + \|\Psi \mathbf{x}\|_{2/3}^{2/3})} \quad (21)$$

As a convenience, we define the two regularization parameters of (17) and (20) as λ_1 and λ_2 , and the corresponding fixed regularization parameters of (19) and (21) as λ'_1 and λ'_2 respectively. Fig. 5 and Fig. 6 present the recovery performance of the proposed algorithm and the prior $L_{2/3}$ algorithm. From the results we can see that the proposed MD based algorithm both outperform the prior algorithm using the two different parameters.

VI. CONCLUSION

In this paper, we proposed a novel iteratively weighted $L_{2/3}$ thresholding algorithm by incorporating with the MD sparse representation strategy. We evaluate the proposed algorithm in compressive image recovery problem. Experiments have shown that the proposed algorithms performs better than the SD based analysis $L_{2/3}$ algorithms by using the proposed MD sparsifying transforms strategy. In addition, the $L_{2/3}$ norm minimization makes a closer approximation to the L_0 norm minimization than L_1 norm, hence can obtain a sparser solution and need fewer measurement data, moreover, the experimental results also indicate that the L_p ($0 < p < 1$) may be a potentially powerful new method to the sparse recovery problem.

ACKNOWLEDGEMENT

This work was supported by National Natural Science Foundation of China grants (No. 61401069, No. 61271240, No. 61501254); Jiangsu Specially Appointed Professor Grant (RK002STP16001); Innovation and Entrepreneurship of Jiangsu High-level Talent Grant (CZ0010617002); High-level talent startup grant of Nanjing University of Posts and Telecommunications (XK0010915026); “1311 Talent Plan” of Nanjing University of Posts and Telecommunications as well as the Postgraduate Research Innovation Program, Jiangsu (KYLX16_0647).

REFERENCES

- [1] Michael Lustig; David L. Donoho; Juan M. Santos; John M. Pauly, “Compressed Sensing MRI: A look at how CS can improve on current

- imaging techniques,” *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 72–82, 2008.
- [2] M. a T. Figueiredo, J. M. Bioucas-dias, and R. D. Nowak, “Majorization – Minimization Algorithms for Wavelet-Based Image Restoration,” *IEEE Trans. image Process.*, vol. 16, no. 12, pp. 2980–2991, 2007.
- [3] X. Cong et al., “Quadratic compressed sensing based SAR imaging algorithm for phase noise mitigation,” *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, vol. E99A, no. 6, 2016.
- [4] S. Han, Z. Gong, W. Meng, C. Li, D. Zhang, and W. Tang, “Automatic Precision Control Positioning for Wireless Sensor Network,” *IEEE Sens. J.*, vol. 16, no. 7, pp. 2140–2150, 2016.
- [5] J. Zhang, C. Pan, F. Pei, G. Liu, and X. Cheng, “Three-Dimensional Fading Channel Models: A Survey of Elevation Angle Research,” *China Commun.*, vol. 52, no. 6, pp. 218–226, 2014.
- [6] S. Han, Z. Gong, W. Meng, C. Li, and X. Gu, “Future alternative positioning, navigation, and timing techniques: A survey,” *IEEE Wirel. Commun.*, vol. 23, no. 6, pp. 154–160, 2016.
- [7] G. Gui, W. Peng, and F. Adachi, “Sub-Nyquist rate ADC sampling-based compressive channel estimation,” *Wirel. Commun. Mob. Comput.*, vol. 15, no. 4, pp. 639–648, 2015.
- [8] G. Gui, L. Dai, S. Kumagai, and F. Adachi, “Variable earns profit: Improved adaptive channel estimation using sparse VSS-NLMS algorithms,” in *IEEE International Conference on Communications (ICC)*, Sydney, Australia, 10–14 June, 2014, pp. 1–5.
- [9] J. Guo, B. Song, Y. He, F. R. Yu, and M. Sookhak, “A Survey on Compressed Sensing in Vehicular Infotainment Systems,” *IEEE Commun. Surv. Tutorials*, vol. 19, no. 4, pp. 2662–2680, 2017.
- [10] Z. Li, H. Huang, and S. Misra, “Compressed Sensing via Dictionary Learning and Approximate Message Passing for Multimedia Internet of Things,” *IEEE Internet Things J.*, vol. 4, no. 2, pp. 505–512, 2017.
- [11] D. L. Donoho and Y. Tsaig, “Fast solution of L1-norm minimization problems when the solution may be sparse,” *Inf. Theory, IEEE Trans.*, vol. 54, no. 11, pp. 4789–4812, 2008.
- [12] W. Cao, J. Sun, and Z. Xu, “Fast image deconvolution using closed-form thresholding formulas of L_q (q = 1/2, 2/3) regularization,” *J. Vis. Commun. Image Represent.*, vol. 24, no. 1, pp. 31–41, 2013.
- [13] R. Ahmad and P. Schniter, “Iteratively Reweighted L1 Approaches to Sparse Composite Regularization,” *IEEE Trans. Comput. Imaging*, vol. 1, no. 4, pp. 220–235, 2015.
- [14] J. Ma and M. März, “A multilevel based reweighting algorithm with joint regularizers for sparse recovery,” <http://arxiv.org/abs/1604.06941>, pp. 1–18, 2016.
- [15] J. Ma et al., “Shearlet-based compressed sensing for fast 3D cardiac MR imaging using iterative reweighting,” <http://arxiv.org/abs/1705.00463>, pp. 1–10, 2017.
- [16] W. Q. Lim, “The discrete shearlet transform: A new directional transform and compactly supported shearlet frames,” *IEEE Trans. Image Process.*, vol. 19, no. 5, pp. 1166–1180, 2010.
- [17] I. Daubechies, “Ten Lectures on Wavelets,” *IEEE Commun. Lett.*, vol. 19, no. 3, pp. 499–502, 2015.
- [18] G. Puy, P. Vandergheynst, R. Gribonval, and Y. Wiaux, “Universal and efficient compressed sensing by spread spectrum and application to realistic Fourier imaging techniques,” *EURASIP J. Adv. Signal Process.*, vol. 2012, no. 1, p. 6, 2012.
- [19] Z. Tan, Y. C. Eldar, A. Beck, and A. Nehorai, “Smoothing and decomposition for analysis sparse recovery,” *IEEE Trans. Signal Process.*, vol. 62, no. 7, pp. 1762–1774, 2014.