IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 3, 2019, accepted April 28, 2019, date of publication May 7, 2019, date of current version May 20, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2915311

Efficient Sparse Recovery and Demixing Using Nonconvex Regularization

JUNHUI MEI AND JUNTONG XI
School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
Corresponding author: Junhui Mei (junhuimei@163.com)

ABSTRACT Sparse demixing aims to separate signals that are sparse in some general dictionary, which has wide applications in signal and image processing, such as in super-resolution, image inpainting, robust sparse recovery, source separation, interference cancellation, saturation, and clipping restoration. For sparsity promotion in sparse demixing, the convex ℓ_1 norm is of the most popular but it has a bias problem. In comparison, nonconvex regularization can mitigate the bias problem and can be expected to yield significantly better performance. In this paper, we employ the nonconvex ℓ_q-norm ($0 \leq q < 1$) for sparsity promotion and consider a linearly constrained ℓ_q-minimization formulation for the sparse demixing problem. Since the ℓ_q-minimization formulation is nonconvex and nonsmoothing, the standard alternative direction method of multipliers (ADMM) often fails to converge. To address this problem, we develop an iteratively reweighted ADMM algorithm which solves convex subproblems in each iteration and is convergent. Further, for the application of color image inpainting, we extend the new algorithm for multi-channel (RGB) joint recovery. The experimental results showed that the new algorithms can achieve significantly better performance than the ℓ_1 algorithm.

INDEX TERMS Alternative direction method of multipliers, inpainting, nonconvex optimization, sparse demixing, sparse recovery, signal separation.

I. INTRODUCTION

Sparse recovery has attracted much research attention in the past decade, which has found wide applications in signal/image processing, statistics and machine learning [1]. In this paper, we consider the problem of recovering two sparse vectors $\mathbf{x}_1 \in \mathbb{R}^{n_1}$ and $\mathbf{x}_2 \in \mathbb{R}^{n_2}$, from the linear measurements $\mathbf{y} \in \mathbb{R}^m$ expressed as

$$\mathbf{y} = \mathbf{A}_1\mathbf{x}_1 + \mathbf{A}_2\mathbf{x}_2 \tag{1}$$

where $\mathbf{A}_1 \in \mathbb{R}^{m\times n_1}$ and $\mathbf{A}_2 \in \mathbb{R}^{m\times n_2}$ are known (deterministic) dictionaries, on which the two components can be sparsely (or approximately sparsely) represented, e.g., with sparse coefficients $\mathbf{x}_1$ and $\mathbf{x}_2$. The goal is to recover and demix the two components via exploiting their underlying sparsity structure. This sparse recovery and separation problem has many applications in signal and image processing, such as the following applications.

1) Super-resolution and inpainting of image, audio and video signals: For example, in the super-resolution and inpainting of image, audio, and video signals [2]–[5]. In these applications, a part of the signal is missing and the task is to retrieve the missing part from an available subset of the desired signal $\mathbf{y}_0 = \mathbf{A}_1\mathbf{x}_1$. In these applications, $\mathbf{A}_2 = \mathbf{I}_m$ and $\mathbf{x}_2$ represents the missing part to be retrieved.

2) Robust sparse signal recovery in the presence of outliers: This problem considers sparse recovery in the presence of impulsive noise. In realistic applications, impulsive noise may come from missing data in the measurement process, transmission problems, faulty memory locations, buffer overflow, reading out from unreliable memory [6]–[10]. In this problem, $\mathbf{A}_2 = \mathbf{I}_m$ and $\mathbf{x}_2$ represents the impulsive noise which is sparse [19].

3) Source separation: For example, in the separation of image texture [11], [12] and in the separation of neuronal calcium transients in calcium imaging [13]. In these applications, $\mathbf{A}_1$ and $\mathbf{A}_2$ are two dictionaries allowing for sparse representation of the two distinct components, and $\mathbf{x}_1$ and $\mathbf{x}_2$ are the corresponding sparse coefficients [14]–[16]. The objective is to demix the two distinct components $\mathbf{A}_1\mathbf{x}_1$ and $\mathbf{A}_2\mathbf{x}_2$ in $\mathbf{y}$.

4) Narrow-band interference signal cancellation: In some communication applications, it is desired to recover a signal contaminated by narrowband interference, e.g., electric

The associate editor coordinating the review of this manuscript and approving it for publication was Guan Gui.

hum [16]. As a typical interference can be sparsely represented in the frequency domain, $\mathbf{A}_2$ can be an inverse discrete Fourier transform matrix on which the interference can be sparsely represented.

5) Restoration of saturation and clipping signals: Saturation arises in some practical systems where the measurements are quantized to a finite number of bits, which can cause significant nonlinearity and potentially unbounded errors [16]–[18]. In the saturation and clipping restoration problem, the objective is to restore $\mathbf{y}_0 = \mathbf{A}_1\mathbf{x}_1$ from its saturated measurement $\mathbf{y}$, with $\mathbf{x}_2$ accounts for the saturation errors.

In the above mentioned applications, $\mathbf{x}_1$ and $\mathbf{x}_2$ can be reasonably assumed to be sparse. To recover (demix) the two components from $\mathbf{y}$ via exploiting the sparsity of $\mathbf{x}_1$ and $\mathbf{x}_2$, a natural formulation is the ℓ_0 norm minimization problem

$$\begin{aligned} &\underset{\mathbf{x}_1,\mathbf{x}_2}{\text{minimize}} \ \{\lambda \|\mathbf{x}_1\|_0 + \|\mathbf{x}_2\|_0\} \\ &\text{subject to } \mathbf{A}_1\mathbf{x}_1 + \mathbf{A}_2\mathbf{x}_2 = \mathbf{y} \end{aligned} \tag{2}$$

where $\|\mathbf{x}\|_0$ is the ℓ_0 norm which counts the number of nonzero elements in $\mathbf{x}$. λ is a positive balance parameter which takes the statistic difference between the two components into consideration. The optimal value of λ depends on the statistical information of the true signals $\mathbf{x}_1$ and $\mathbf{x}_2$. However, the ℓ_0 minimization problem is highly nonconvex and difficult to solve, known as NP hard. A popular manner is to use convex relaxation that replaces the ℓ_0 norm by its convex envelop (i.e., the ℓ_1 norm) as [15]

$$\begin{aligned} &\underset{\mathbf{x}_1,\mathbf{x}_2}{\text{minimize}} \ \{\lambda \|\mathbf{x}_1\|_1 + \|\mathbf{x}_2\|_1\} \\ &\text{subject to } \mathbf{A}_1\mathbf{x}_1 + \mathbf{A}_2\mathbf{x}_2 = \mathbf{y}. \end{aligned} \tag{3}$$

For sparsity promotion, the ℓ_1-norm regularization is of the most popular due to its convexity, since well-established convex optimization algorithms can be directly applied or can be applied after some extension in solving the involved convex problems. However, the ℓ_1-norm regularization would yield a biased estimation, as the corresponding soft-thresholding operation imposes a constant shrinkage on the parameters which would result in biased estimate for large coefficients [1], [32], [35]. Moreover, it has been demonstrated both theoretically and empirically that, the ℓ_1-regularization cannot achieve reliable recovery with the least measurements [20]. To alleviate the bias problem of the ℓ_1-regularization, nonconvex regularization can be used. Typical nonconvex penalties include the ℓ_q-norm ($0 \leq q < 1$), and the smoothly clipped absolute deviation (SCAD) penalty [21]. For example, it has been shown that ℓ_q-regularization with $q < 1$ can attain better sparse recovery performance than ℓ_1-regularization [22]–[27]. More specifically, ℓ_q-regularization requires fewer measurements and weaker sufficient conditions for reliable reconstruction than ℓ_1-regularization. It has been demonstrated that, in terms of restricted isometry property (RIP), the sufficient conditions of ℓ_q-regularization for reliable reconstruction are weaker than those of ℓ_1-regularization [22].

In this paper we consider a sparse demixing formulation using the ℓ_q-norm with $0 \leq q < 1$ for sparsity inducing as

$$\begin{aligned} &\underset{\mathbf{x}_1,\mathbf{x}_2}{\text{minimize}} \ \left\{\lambda \|\mathbf{x}_1\|_{q_1}^{q_1} + \|\mathbf{x}_2\|_{q_2}^{q_2}\right\} \\ &\text{subject to } \mathbf{A}_1\mathbf{x}_1 + \mathbf{A}_2\mathbf{x}_2 = \mathbf{y} \end{aligned} \tag{4}$$

where $0 \leq q_1, q_2 < 1$, $\|\cdot\|_q$ is the ℓ_q quasi-norm defined as $\|\mathbf{x}\|_q = (\sum_{i=1}^n |x_i|^q)^{1/q}$.

A. RELATED WORK

Using convex regularization for sparse demixing has been considered in [14], [15]. Specifically, the particular case of $q_1 = q_2 = 1$ has been considered in [15] for source separation, in this case the formulation becomes

$$\begin{aligned} &\underset{\mathbf{x}_1,\mathbf{x}_2}{\text{minimize}} \ \{\lambda \|\mathbf{x}_1\|_1 + \|\mathbf{x}_2\|_1\} \\ &\text{subject to } \mathbf{A}_1\mathbf{x}_1 + \mathbf{A}_2\mathbf{x}_2 = \mathbf{y}. \end{aligned} \tag{5}$$

Meanwhile, the case of $\lambda = 1$ and $q_1 = q_2 = 1$ has been considered in [14] for sparse demixing. Moreover, the case with $\mathbf{A}_2 = \mathbf{I}_m$ and $q_1 = q_2 = 1$, problem (4) becomes

$$\begin{aligned} &\underset{\mathbf{x}_1,\mathbf{x}_2}{\text{minimize}} \ \{\lambda \|\mathbf{x}_1\|_1 + \|\mathbf{x}_2\|_1\} \\ &\text{subject to } \mathbf{A}_1\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{y} \end{aligned}$$

which reduces to the ℓ_1-regularized least-absolute problem for robust sparse recovery [28]

$$\underset{\mathbf{x}_1}{\text{minimize}} \ \left\{\lambda \|\mathbf{x}_1\|_1 + \|\mathbf{A}_1\mathbf{x}_1 - \mathbf{y}\|_1\right\}. \tag{6}$$

The ℓ_1-loss is more robust to impulsive noise than the ℓ_2-loss. An extended version for robust sparse recovery with $\mathbf{A}_2 = \mathbf{I}_m$, $q_1 = 1$ and $0 \leq q_2 < 2$ has been considered in [29]. Furthermore, the ℓ_q-regularized least-squares formulation in [27] is also a special cases of (4) with $\mathbf{A}_2 = \mathbf{I}_m$, $0 \leq q_1 < 1$ and $q_2 = 2$.

More recently, the sparse demixing formulation (4) with $0 \leq q_1, q_2 < 1$ has been considered in [5]. Instead of directly solving (4), a relaxed version has been used in [5] as

$$\min_{\mathbf{x}_1,\mathbf{x}_2} \left\{\frac{1}{\beta}\|\mathbf{A}_1\mathbf{x}_1 + \mathbf{A}_2\mathbf{x}_2 - \mathbf{y}\|_2^2 + \lambda \|\mathbf{x}_1\|_{q_1}^{q_1} + \|\mathbf{x}_2\|_{q_2}^{q_2}\right\} \tag{7}$$

where $\beta > 0$ is a penalty parameter. For this unconstrained formulation, two first-order algorithms have been proposed in [5] based on the block coordinate descent (BCD) and alternative direction method of multipliers (ADMM) frameworks. With a sufficiently small value of β, ideally $\beta \to 0$, the solution of (7) satisfy $\|\mathbf{A}_1\mathbf{x}_1 + \mathbf{A}_2\mathbf{x}_2 - \mathbf{y}\|_2 \to 0$ and hence accurately approaches the solution of (4). However, with a very small value of β, problem (7) becomes ill-conditioned and the algorithms would converge very slowly and become impractical.

B. CONTRIBUTION

In this work we directly solve problem (4) using ADMM. However, the standard ADMM applied to problem (4) is not guaranteed to converge in the nonconvex case. Numerical

studies showed that the standard ADMM algorithm often fails to converge (see Fig. 2 in section IV).

First, to derive a convergent algorithm, we employ an iterative reweighting strategy in the ADMM algorithm. The proposed iteratively reweighted ADMM algorithm solves a convex subproblem in each step and is convergent under some mild conditions.

Second, for the application of color image inpainting, the new algorithm has been extended to exploit the feature correlation between the RGB channels of a color image.

Finally, experimental results have been presented to show the effectiveness and efficiency of the new algorithm, with special application interesting on color image inpainting.

C. OUTLINE

Section II introduces the proposed algorithm for sparse vector demixing. Section III extends the proposed algorithm to the multi-channels joint sparse recovery application. Numerical experiments are provided in Section IV. Finally, Section V ends the paper with concluding remarks.

II. PROPOSED ALGORITHM

ADMM is a powerful framework that can be used to efficiently solve many high-dimensional problems in signal processing, image processing and machine learning [30]. The core idea of ADMM is to employ a *splitting-coordination* procedure to decouple the variables and make an intackable global problem easy to tackle.

A. STANDARD ADMM ALGORITHM

For problem (4), the standard ADMM can be applied. Specifically, the augmented Lagrangian function is

$$\begin{aligned} L(\mathbf{x},\mathbf{w}) &= \lambda\|\mathbf{x}_1\|_{q_1}^{q_1}+\|\mathbf{x}_2\|_{q_2}^{q_2}+\langle \mathbf{w}, \mathbf{A}_1\mathbf{x}_1+\mathbf{A}_2\mathbf{x}_2-\mathbf{y}\rangle \\ &\quad +\frac{\rho}{2}\|\mathbf{A}_1\mathbf{x}_1+\mathbf{A}_2\mathbf{x}_2-\mathbf{y}\|_2^2 \\ &= \lambda\|\mathbf{x}_1\|_{q_1}^{q_1}+\|\mathbf{x}_2\|_{q_2}^{q_2}+\frac{\rho}{2}\left\|\mathbf{A}_1\mathbf{x}_1+\mathbf{A}_2\mathbf{x}_2-\mathbf{y}+\frac{\mathbf{w}}{\rho}\right\|_2^2 \\ &\quad -\frac{\|\mathbf{w}\|_2^2}{2\rho} \end{aligned}$$

where $\mathbf{w}$ is the dual variable, ρ is a positive penalty parameter. Then, the primal and dual variables are alternatively updated as follows

$$\mathbf{x}_1^{k+1}=\arg\min_{\mathbf{x}\in R^{n_1}}\left(\lambda\|\mathbf{x}\|_{q_1}^{q_1}+\frac{\rho}{2}\left\|\mathbf{A}_1\mathbf{x}+\mathbf{A}_2\mathbf{x}_2^k-\mathbf{y}+\frac{\mathbf{w}^k}{\rho}\right\|_2^2\right) \tag{8}$$

$$\mathbf{x}_2^{k+1}=\arg\min_{\mathbf{x}\in R^{n_2}}\left(\|\mathbf{x}\|_{q_2}^{q_2}+\frac{\rho}{2}\left\|\mathbf{A}_1\mathbf{x}_1^{k+1}+\mathbf{A}_2\mathbf{x}-\mathbf{y}+\frac{\mathbf{w}^k}{\rho}\right\|_2^2\right) \tag{9}$$

$$\mathbf{w}^{k+1}=\mathbf{w}^k+\rho(\mathbf{A}_1\mathbf{x}_1^{k+1}+\mathbf{A}_2\mathbf{x}_2^{k+1}-\mathbf{y}). \tag{10}$$

The $\mathbf{x}_1$-subproblem is a nonconvex regularized least-square problem which is difficult to solve directly. A standard trick is to use a proximal linearization for the quadratic term. Specifically, let $\mathbf{u}^k=\mathbf{A}_2\mathbf{x}_2^k-\mathbf{y}+\mathbf{w}^k/\rho$, use a quadratic majorization of the second term in the $\mathbf{x}_1$-subproblem as

$$\begin{aligned} &\frac{1}{2}\left\|\mathbf{A}_1\mathbf{x}_1+\mathbf{u}^k\right\|_2^2 \\ &\approx \frac{1}{2}\left\|\mathbf{A}_1\mathbf{x}_1^k+\mathbf{u}^k\right\|_2^2+\left\langle \mathbf{x}_1-\mathbf{x}_1^k, d_1(\mathbf{x}_1^k)\right\rangle+\frac{\mu_1}{2}\left\|\mathbf{x}_1-\mathbf{x}_1^k\right\|_2^2 \end{aligned} \tag{11}$$

where $d_1(\mathbf{x}_1^k)=\mathbf{A}_1^T(\mathbf{A}_1\mathbf{x}_1^k+\mathbf{u}^k)$, $\mu_1>0$ is a proximal parameter. In this manner, the $\mathbf{x}_1$-subproblem becomes

$$\mathbf{x}_1^{k+1}=\arg\min_{\mathbf{x}\in R^{n_1}}\left(\lambda\|\mathbf{x}\|_{q_1}^{q_1}+\frac{\rho\mu_1}{2}\left\|\mathbf{x}-\mathbf{x}_1^k+\frac{d_1(\mathbf{x}_1^k)}{\mu_1}\right\|_2^2\right) \tag{12}$$

which can be efficiently solved via element-wise proximity operator. Recall the proximity operator for the ℓ_q-norm function

$$T_{q,\eta}(t)=\arg\min_x\left\{\|x\|_q^q+\frac{\eta}{2}(x-t)^2\right\} \tag{13}$$

where $\eta>0$. For $q=0$, it reduces to the hard-thresholding

$$T_{0,\eta}(t)=\begin{cases}0, & |t|<\sqrt{2/\eta}\\ \{0,t\}, & |t|=\sqrt{2/\eta}\\ t, & \text{otherwise}\end{cases}$$

while for $q=1$, it becomes the soft-thresholding operator

$$T_{1,\eta}(t)=\text{sign}(t)\max\{|t|-1/\eta,0\} \tag{14}$$

For $0<q<1$, it can be computed as [31]

$$T_{q,\eta}(t)=\begin{cases}0, & |t|<\tau\\ \{0,\text{sign}(t)\beta\}, & |t|=\tau\\ \text{sign}(t)z, & |t|>\tau\end{cases}$$

where $\beta=[2(1-q)/\eta]^{1/(2-q)}$, $\tau=\beta+q\beta^{q-1}/\eta$, $z\in(\beta,|t|)$ satisfies $h(z)=qz^{q-1}+\eta z-\eta|t|=0$. As $h(z)$ is a convex function, when $|t|>\tau$, z can be efficiently solved using a Newton's method. The subproblem (12) is separable and can be solved in an element-wise manner as (13).

Similarly, let $\mathbf{v}^k=\mathbf{A}_1\mathbf{x}_1^{k+1}-\mathbf{y}+\mathbf{w}^k/\rho$, use a quadratic majorization of the second term in the $\mathbf{x}_2$-subproblem as

$$\begin{aligned} &\frac{1}{2}\left\|\mathbf{A}_2\mathbf{x}_2+\mathbf{v}^k\right\|_2^2 \\ &\approx \frac{1}{2}\left\|\mathbf{A}_2\mathbf{x}_2^k+\mathbf{v}^k\right\|_2^2+\left\langle \mathbf{x}_2-\mathbf{x}_2^k, d_2(\mathbf{x}_2^k)\right\rangle+\frac{\mu_2}{2}\left\|\mathbf{x}_2-\mathbf{x}_2^k\right\|_2^2 \end{aligned} \tag{15}$$

where $d_2(\mathbf{x}_2^k)=\mathbf{A}_2^T(\mathbf{A}_2\mathbf{x}_2^k+\mathbf{v}^k)$, $\mu_2>0$ is a proximal parameter. Then, the $\mathbf{x}_2$-subproblem becomes

$$\mathbf{x}_2^{k+1}=\arg\min_{\mathbf{x}\in R^{n_2}}\left(\|\mathbf{x}\|_{q_2}^{q_2}+\frac{\rho\mu_2}{2}\left\|\mathbf{x}-\mathbf{x}_2^k+\frac{d_2(\mathbf{x}_2^k)}{\mu_2}\right\|_2^2\right) \tag{16}$$

which can be efficiently solved via the proximal minimization (13).

However, the above ADMM algorithm is not guaranteed to converge, which often fails to converge in numerical experiments (see Fig. 2 in section IV). In the following, we propose

an iteratively reweighted algorithm to approximately solve the problem (4).

B. PROPOSED ITERATIVELY REWEIGHTED ADMM ALGORITHM

To develop a convergent algorithm for the linearly constrained nonconvex minimization problem (4), we use the iteratively reweighted method to approximately solve (4). Our algorithm can be viewed as an extension of the work [33] on sparse recovery to sparse demixing. Besides, while the method in [33] solves a sequence of ℓ_1 minimization subproblems, our algorithm is a first-order algorithm as in each iteration the dominant computational complexity is matrix-vector multiplication, hence is highly efficient and scales well to high-dimensional problems.

Specifically, at the $(k+1)$-th iteration, the ℓ_q norm is approximated by

$$\|\mathbf{x}\|_{q,\varepsilon}^q = \sum_i \left(\left|x_i^k\right| + \varepsilon\right)^{q-1} |x_i| \tag{17}$$

where $\varepsilon > 0$ is a proximal parameter.

With this approximation, the $\mathbf{x}_1$ and $\mathbf{x}_2$ subproblems become convex problems as

$$\mathbf{x}_1^{k+1} = \arg\min_{\mathbf{x}\in R^{n_1}} \left(\lambda \left\| \omega_1^k \circ \mathbf{x} \right\|_1 + \frac{\rho}{2} \left\| \mathbf{A}_1\mathbf{x} + \mathbf{A}_2\mathbf{x}_2^k - \mathbf{y} + \frac{\mathbf{w}^k}{\rho} \right\|_2^2 \right) \tag{18}$$

$$\mathbf{x}_2^{k+1} = \arg\min_{\mathbf{x}\in R^{n_2}} \left(\left\| \omega_2^k \circ \mathbf{x} \right\|_1 + \frac{\rho}{2} \left\| \mathbf{A}_1\mathbf{x}_1^{k+1} + \mathbf{A}_2\mathbf{x} - \mathbf{y} + \frac{\mathbf{w}^k}{\rho} \right\|_2^2 \right) \tag{19}$$

where $\circ$ denotes the Hadamard product and ω_i^k is weight at the k-th iteration given by

$$\omega_i^k = \left[\left(\left|x_i^k(1)\right| + \varepsilon \right)^{q_i-1}, \cdots, \left(\left|x_i^k(n_i)\right| + \varepsilon \right)^{q_i-1} \right]. \tag{20}$$

Then, with the use of the linearization (11) and (15), it follows that

$$\mathbf{x}_1^{k+1} = \arg\min_{\mathbf{x}\in R^{n_1}} \left(\lambda \left\| \omega_1^k \circ \mathbf{x} \right\|_1 + \frac{\rho\mu_1}{2} \left\| \mathbf{x} - \mathbf{x}_1^k + \frac{d_1(\mathbf{x}_1^k)}{\mu_1} \right\|_2^2 \right) \tag{21}$$

$$\mathbf{x}_2^{k+1} = \arg\min_{\mathbf{x}\in R^{n_2}} \left(\left\| \omega_2^k \circ \mathbf{x} \right\|_1 + \frac{\rho\mu_2}{2} \left\| \mathbf{x} - \mathbf{x}_2^k + \frac{d_2(\mathbf{x}_2^k)}{\mu_2} \right\|_2^2 \right) \tag{22}$$

which can be efficiently solved via the soft-thresholding operator (14).

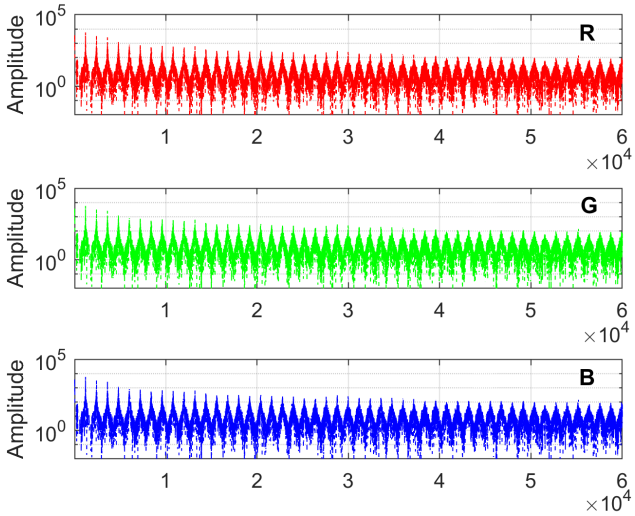

FIGURE 1. DCT coefficients of the RGB channels of a 500 × 318 color image used in the inpainting experiment in section IV (see Fig. 4). For clarity, only the first 25% coefficients of each channel are plotted.

Extensive numerical studies show that, this iteratively reweighted algorithm can achieve satisfactory with an initialization obtained by this algorithm with $q_1 = q_2 = 1$. When $q_1 = q_2 = 1$, the $\mathbf{x}_1$ and $\mathbf{x}_2$ subproblems degenerate to

$$\mathbf{x}_1^{k+1} = \arg\min_{\mathbf{x}\in R^{n_1}} \left(\lambda \|\mathbf{x}\|_1 + \frac{\rho\mu_1}{2} \left\| \mathbf{x} - \mathbf{x}_1^k + \frac{d_1(\mathbf{x}_1^k)}{\mu_1} \right\|_2^2 \right) \tag{23}$$

$$\mathbf{x}_2^{k+1} = \arg\min_{\mathbf{x}\in R^{n_2}} \left(\|\mathbf{x}\|_1 + \frac{\rho\mu_2}{2} \left\| \mathbf{x} - \mathbf{x}_2^k + \frac{d_2(\mathbf{x}_2^k)}{\mu_2} \right\|_2^2 \right). \tag{24}$$

In the new algorithm, the dominant computation is matrix-vector multiplication with complexity $O(mn_1 + mn_2)$, thus it scales well to large-scale problems.

III. EXTENSION TO MULTICHANNEL JOINT SPARSE RECOVERY PROBLEMS

This section extends the proposed algorithm to the problem of joint sparse recovery. An interesting application of such problem is color image inpainting. For a color image with three channels (RGB image), each channel can be recovered separately. Since the three color channels (also the corruption in the three channels) usually have similar (roughly the same) sparsity pattern (as illustrated in Fig. 1), performance improvement can be expected via exploiting the sparsity pattern similarity among different channels. This is also referred to as group or joint sparse recovery. In this section, we extend the above ADMM algorithm to joint sparse recovery.

In the joint sparse recovery problem, assume that there are L channels, the linear measurements $\mathbf{Y} \in R^{m\times L}$ can be expressed as

$$\mathbf{Y} = \mathbf{A}_1\mathbf{X}_1 + \mathbf{A}_2\mathbf{X}_2 \tag{25}$$

where $\mathbf{X}_k \in R^{n_k\times L}$, $k = 1, 2$, contain the sparse coefficients of the two components. To exploit the joint sparsity among

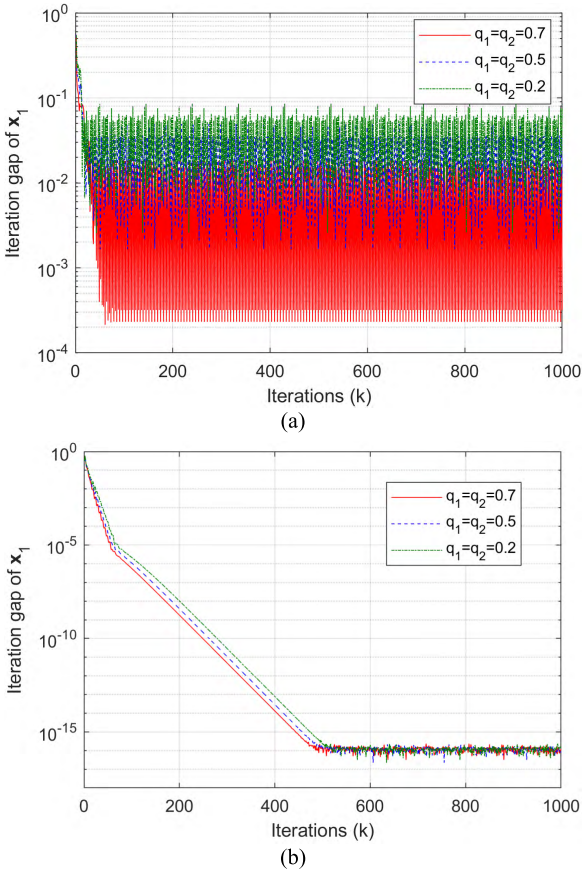

FIGURE 2. Typical convergence behavior of ADMM in the nonconvex case with $q_1 = q_2 = q < 1$. (a) Standard ADMM. (b) Proposed iteratively reweighted ADMM.

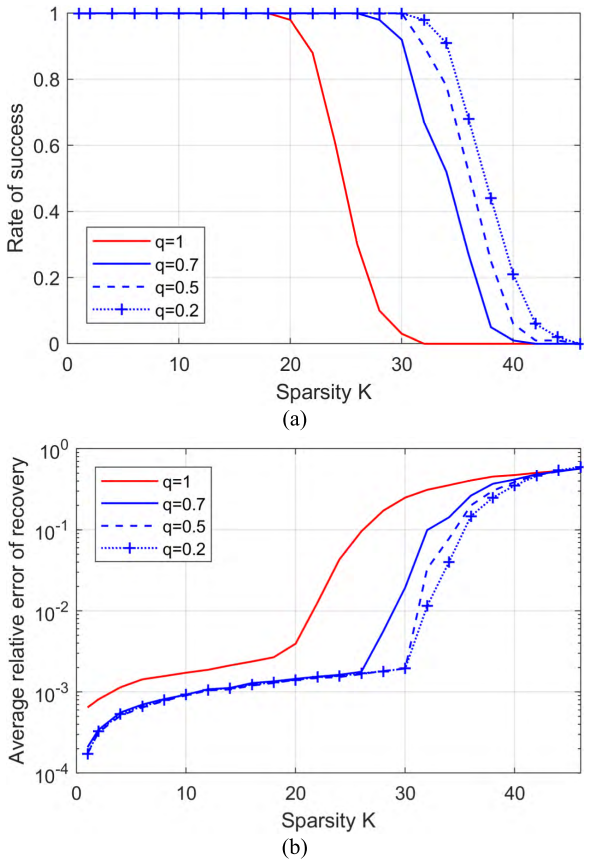

FIGURE 3. Rate of successful recovery and averaged relative error of recovery versus sparsity K. $\mathbf{A}_1$ is a DCT matrix and $\mathbf{A}_2$ be a Gaussian matrix, with $q_1 = q_2 = q \in \{0.2, 0.5, 0.7, 1\}$. (a) Rate of successful recovery. (b) Averaged relative error of recovery.

the L channels, we reformulate the problem (4) as

$$\begin{aligned} &\underset{\mathbf{X}}{\text{minimize}} \ \lambda \ \|\mathbf{X}_1\|_{1,q_1}^{q_1} + \|\mathbf{X}_2\|_{1,q_2}^{q_2} \\ &\text{subject to } \mathbf{A}_1\mathbf{X}_1 + \mathbf{A}_2\mathbf{X}_2 = \mathbf{Y}. \end{aligned} \tag{26}$$

For $q \geq 0$, the matrix norm $\|\mathbf{X}\|_{1,q}^q$ is defined as

$$\|\mathbf{X}\|_{1,q}^q = \sum\nolimits_i \|\mathbf{X}[i,:]\|_1^q = \sum\nolimits_i \left(\sum\nolimits_j |\mathbf{X}[i,j]|\right)^q.$$

Similar to (17), at the $(k+1)$-th iteration, we approximate the ℓ_q norm by

$$\|\mathbf{X}\|_{1,q,\varepsilon}^q = \sum\nolimits_i \left[\left(\varepsilon + \sum\nolimits_j |\mathbf{X}^k[i,j]|\right)^{q-1} \sum\nolimits_j |\mathbf{X}[i,j]|\right] \tag{27}$$

with $\varepsilon > 0$ be a proximal parameter. Then, similar to the iteratively reweighted ADMM algorithm in section II-B, the ADMM algorithm for joint sparse recovery consists of the following steps

$$\begin{aligned} \mathbf{X}_1^{k+1} = \arg \min_{\mathbf{X} \in R^{n_1 \times L}} \Bigg(&\lambda \left\| \Sigma_1^k \circ \mathbf{X} \right\|_{1,1} \\ &+ \frac{\rho}{2} \left\| \mathbf{A}_1\mathbf{X} + \mathbf{A}_2\mathbf{X}_2^k - \mathbf{Y} + \frac{\mathbf{W}^k}{\rho} \right\|_F^2 \Bigg) \end{aligned} \tag{28}$$

$$\begin{aligned} \mathbf{X}_2^{k+1} = \arg \min_{\mathbf{X} \in R^{n_2 \times L}} \Bigg(&\left\| \Sigma_2^k \circ \mathbf{X} \right\|_{1,1} \\ &+ \frac{\rho}{2} \left\| \mathbf{A}_1\mathbf{X}_1^{k+1} + \mathbf{A}_2\mathbf{X} - \mathbf{Y} + \frac{\mathbf{W}^k}{\rho} \right\|_F^2 \Bigg) \end{aligned} \tag{29}$$

$$\mathbf{W}^{k+1} = \mathbf{W}^k + \rho(\mathbf{A}_1\mathbf{X}_1^{k+1} + \mathbf{A}_2\mathbf{X}_2^{k+1} - \mathbf{Y}) \tag{30}$$

where $\mathbf{W}$ is the dual variable, $\Sigma_i^k \in R^{n_i \times L}$ (with $i = 1, 2$) is a weighting matrix in the $(k+1)$-th iteration given by

$$\Sigma_i^k = \begin{bmatrix} \left(\varepsilon + \sum_j |\mathbf{X}^k[1,j]|\right)^{q_i - 1} & \cdots & \left(\varepsilon + \sum_j |\mathbf{X}^k[1,j]|\right)^{q_i - 1} \\ & \vdots & \\ \left(\varepsilon + \sum_j |\mathbf{X}^k[n_i,j]|\right)^{q_i - 1} & \cdots & \left(\varepsilon + \sum_j |\mathbf{X}^k[n_i,j]|\right)^{q_i - 1} \end{bmatrix}$$

Similar to the algorithm in section II, we further use a linearization for the $\mathbf{X}_1$ and $\mathbf{X}_2$ subproblems. Let $\mathbf{U}^k = \mathbf{A}_2\mathbf{X}_2^k - \mathbf{Y} + \mathbf{W}^k/\rho$, a quadratic majorization of the second term in the $\mathbf{X}_1$-subproblem is

$$\begin{aligned} &\frac{1}{2} \left\| \mathbf{A}_1\mathbf{X}_1 + \mathbf{U}^k \right\|_F^2 \\ &\approx \frac{1}{2} \left\| \mathbf{A}_1\mathbf{X}_1^k + \mathbf{U}^k \right\|_F^2 + \left\langle \mathbf{X}_1 - \mathbf{X}_1^k, D_1(\mathbf{X}_1^k) \right\rangle + \frac{c_1}{2} \left\| \mathbf{X}_1 - \mathbf{X}_1^k \right\|_F^2 \end{aligned}$$

where $D_1(\mathbf{X}_1^k) = \mathbf{A}_1^T(\mathbf{A}_1\mathbf{X}_1^k + \mathbf{U}^k)$, $c_1 > 0$ is a proximal parameter. In this manner, the $\mathbf{X}_1$-subproblem

FIGURE 4. Recovery performance of the proposed algorithms on three 600 × 400 color images corrupted by salt-and-pepper noise.

becomes

$$\mathbf{X}_1^{k+1} = \arg\min_{\mathbf{X}\in R^{n_1\times L}} \left(\lambda \left\| \Sigma_1^k \circ \mathbf{X} \right\|_{1,1} + \frac{\rho c_1}{2} \left\| \mathbf{X} - \mathbf{X}_1^k + \frac{D_1(\mathbf{X}_1^k)}{c_1} \right\|_F^2 \right). \quad (31)$$

Moreover, let $\mathbf{V}^k = \mathbf{A}_1\mathbf{X}_1^{k+1} - \mathbf{Y} + \mathbf{W}^k/\rho$, a quadratic majorization of the second term in the $\mathbf{X}_2$-subproblem is

$$\frac{1}{2}\left\| \mathbf{A}_2\mathbf{X}_2 + \mathbf{V}^k \right\|_F^2 \approx \frac{1}{2}\left\| \mathbf{A}_2\mathbf{X}_2^k + \mathbf{V}^k \right\|_F^2 + \left\langle \mathbf{X}_2 - \mathbf{X}_2^k, D_2(\mathbf{X}_2^k) \right\rangle + \frac{c_2}{2}\left\| \mathbf{X}_2 - \mathbf{X}_2^k \right\|_F^2$$

where $D_2(\mathbf{X}_2^k) = \mathbf{A}_2^T(\mathbf{A}_2\mathbf{X}_2^k + \mathbf{V}^k)$, $c_2 > 0$ is a proximal parameter. In this manner, the $\mathbf{X}_2$-subproblem becomes

$$\mathbf{X}_2^{k+1} = \arg\min_{\mathbf{X}\in R^{n_2\times L}} \left(\left\| \Sigma_2^k \circ \mathbf{X} \right\|_{1,1} + \frac{\rho c_2}{2} \left\| \mathbf{X} - \mathbf{X}_2^k + \frac{D_2(\mathbf{X}_2^k)}{c_2} \right\|_F^2 \right). \quad (32)$$

Then the $\mathbf{X}_1$ and $\mathbf{X}_2$ subproblems (31) and (32) can be solved in an element-wise manner. Specifically, the subproblems (31) and (32) are separable with respect to each element in the varable $\mathbf{X}$, and each element is a proximal optimization problem of the form

$$\min_x \left\{ \omega|x| + \frac{\eta}{2}(x-t)_2^2 \right\} \quad (33)$$

whose solution is explicitly given by the soft thresholding $\text{sign}(t)\max\{|t| - \omega/\eta, 0\}$. Thus, the solution to the subproblems (31) and (32) are given by

$$\mathbf{X}_1^{k+1} = \text{sign}(\mathbf{T}_1) \circ \max\left\{ |\mathbf{T}_1| - \frac{\lambda}{\rho c_1}\Sigma_1^k, 0 \right\} \quad (34)$$

$$\mathbf{X}_2^{k+1} = \text{sign}(\mathbf{T}_2) \circ \max\left\{ |\mathbf{T}_2| - \frac{1}{\rho c_2}\Sigma_2^k, 0 \right\} \quad (35)$$

with

$$\mathbf{T}_i = \mathbf{X}_i^k - D_i(\mathbf{X}_i^k)/c_i.$$

For the special case of $L = 1$, this algorithm degenerates to the ADMM algorithm for sparse vector recovery in section II-B.

IV. NUMERICAL EXPERIMENTS

This section evaluates the performance of the new algorithms via numerical experiments, including a synthetic experiment using simulated sparse signals and a real life experiment on color image inpainting. The regularization parameter λ is selected by providing the best performance of the algorithm in terms of the lowest relative error (RelErr) of recovery. For the new algorithm, the proximal parameter is selected as $\varepsilon = 10^{-4}$, and we first run it with $q_1 = q_2 = 1$ and $\lambda = 1$ to obtain an initial estimation.

A. SYNTHETIC DATA RECOVERY

In the first experiment, we use simulated data as: $\mathbf{A}_1 \in \mathbb{R}^{128\times 128}$ is a DCT matrix, $\mathbf{A}_2 \in \mathbb{R}^{128\times 128}$ is an orthonormal Gaussian random matrix, the sparse vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ have a same sparsity of K. The positions of nonzeros elements in $\mathbf{x}_1$ and $\mathbf{x}_2$ are uniformly distributed, while the amplitude of the nonzero elements Gaussian distributed. The sparsity K is varied from 1 to 45. We use $q_1 = q_2 = q$ with $q \in \{0.2, 0.5, 0.7, 1\}$.

Fig. 2 plots typical convergence behavior comparison between the standard ADMM using the steps (8)–(10) and the proposed iteratively reweighted ADMM algorithm using the steps (18), (19) and (10) in the considered nonconvex cases. It can be seen that, the standard ADMM using (8)–(10) does not converge, while the proposed iteratively reweighted ADMM algorithm is convergent. The new algorithm can converge to an accuracy with tolerance lower than 10^{-15} within 500 iterations.

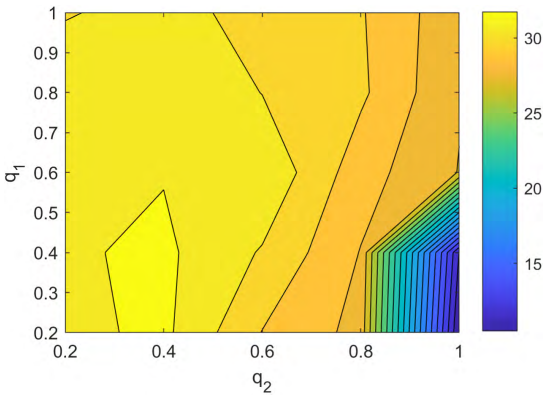

FIGURE 5. Recovery performance versus q_1 and q_2 in salt-and-pepper noise, in terms of PSNR in dB.

Fig. 3 shows the recovery performance for different sparsity K in terms of success rate of recovery and averaged relative error of recovery. Let $\hat{\mathbf{x}}_1$ denote the recovered result of the ground-truth $\mathbf{x}_1^o$, then the recovery is regarded as successful if the relative error of recovery is smaller than 10^{-2}, i.e.,

$$\frac{\left\|\hat{\mathbf{x}}_1 - \mathbf{x}_1^o\right\|_2}{\left\|\mathbf{x}_1^o\right\|_2} < 10^{-2}.$$

The result in Fig. 3 is averaged over 200 independent runs. It can be seen from Fig. 3 that, using $q < 1$ can achieve significantly better performance than using the popular convex penalty with $q = 1$. The performance gain of nonconvex regularization is especially conspicuous when the sparsity is in the region $K \in [20, 35]$. For example, for $K = 25$, the success rate given by $q = 1$ is about 40% while that given by $q \in \{0.2, 0.5, 0.7\}$ is 100%. Moreover, for $K = 25$, the averaged recovery error given by $q = 1$ is about 80 times larger than that given by $q = 0.2$. On the whole, $q = 0.2$ yields the best performance.

B. REALISTIC EXPERIMENT ON COLOR IMAGE INPAINTING

In the second experiment, we consider color image inpainting using the multichannel version of the new algorithm. It is worth noting that, there exist a number of inpainting methods, such as the method using the Field of Experts (FoE) model [34], which can attain better performance than the proposed algorithm in the following considered inpainting experiment. However, such methods require the exact support-set knowledge (mask) of the corruption, while our algorithm does not use such prior information. The main interesting here is evaluate the new algorithm with $q < 1$ in comparison with the convex ℓ_1 method.

The objective is to separate the original image from text overwriting or salt-and-pepper noise corruption. For this

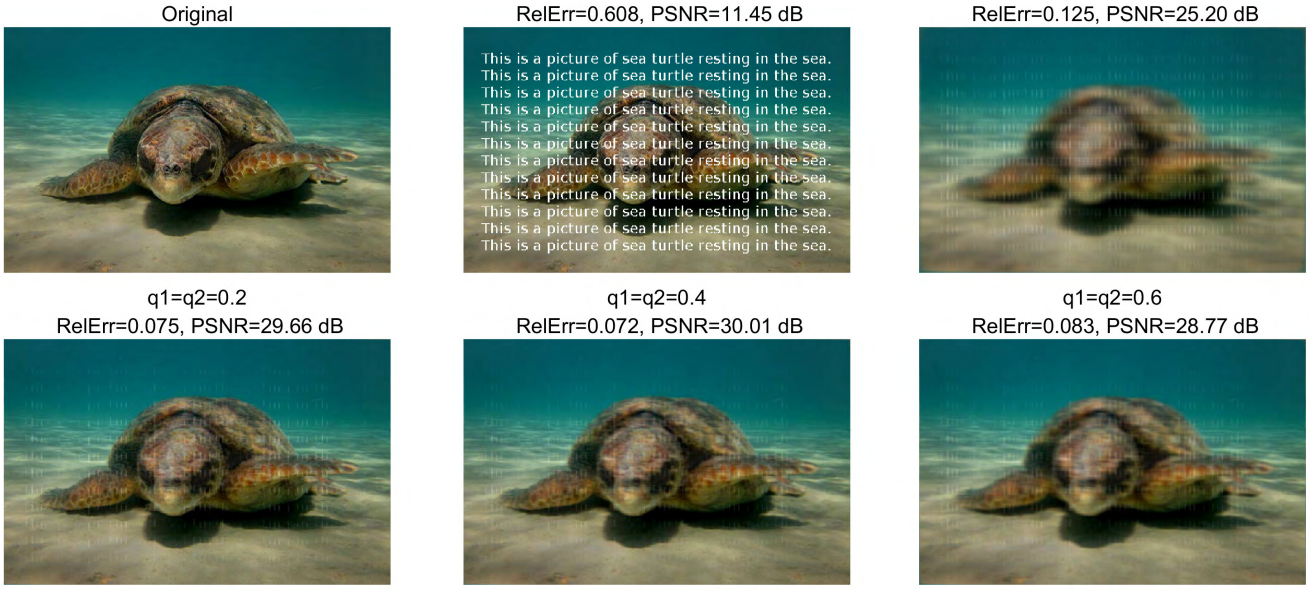

FIGURE 6. Recovery performance of the proposed algorithms on a 500 × 318 color image corrupted by overwriting with text.

TABLE 1. Recovery performance of the new method on color image inpainting in the presence of salt-and-pepper noise ($q_1 = q_2 = q$).

	Image 1		Image 2		Image 3	
	RelErr	PSNR	RelErr	PSNR	RelErr	PSNR
q= 1	0.092	28.23	0.136	23.63	0.105	26.32
q= 0.2	0.065	31.29	0.122	24.60	0.086	28.07
q= 0.4	**0.062**	**31.76**	0.119	24.82	**0.084**	**28.23**
q= 0.6	0.065	31.34	**0.115**	**25.08**	0.090	27.66

application, $\mathbf{A}_1$ is a basis of the image and $\mathbf{A}_2 = \mathbf{I}$. $\mathbf{A}_1$ is selected as an inverse discrete cosine transformation (IDCT) matrix. In this case, $\mathbf{X}_1$ contains the DCT coefficients of the image. The IDCT matrix enables fast computation of the multiplication of $\mathbf{A}_1$ (or $\mathbf{A}_1^T$) with a vector.

Fig. 4 shows the results in color image inpainting with the images be corrupted by salt-and-pepper noise, where 30% of the pixels of each image are corrupted. The presented results include the recovered images, the relative error (RelErr) of the estimated DCT coefficients and the peak-signal noise ratio (PSNR) of the restored image.

The results demonstrate that, the proposed algorithm with $q_1 < 1$ and $q_2 < 1$ can yield distinctly better performance than that of $q_1 = 1$ and $q_2 = 1$. For example, as shown by the results in Table I, for the three images, the PSNR with $q_1 = q_2 = 0.4$ are respectively 3.53, 1.19 and 1.91 dB higher than that of $q_1 = q_2 = 1$. Fig. 5 shows the PSNR for the first image of the proposed method for different values of q_1 and q_2. The result indicates that, the new algorithm can achieve satisfactory performance by roughly selecting $q_1 \leq 0.8$ and $q_2 \leq 0.5$. This is reasonable, since the DCT coefficients of a real-life image are not strictly sparse but rather approximately follow an exponential decay, while the salt-and-pepper noise is strictly sparse.

The final experiment considers color image inpainting when the image is corrupted by overwriting with text. The result is shown in Fig. 6. It can be seen that, with properly chosen values of q_1 and q_2, the proposed method has significantly better performance than the ℓ_1 method.

V. CONCLUSIONS

This paper proposed an efficient algorithm for sparse demixing based on nonconvex regularization. The new algorithm uses a reweighted ℓ_1 method to approximately solve the ℓ_q-norm regularized nonconvex formulation based on ADMM. Further, the new algorithm has been extended to the multichannel joint recovery problem for color image inpainting. Both synthetic and realistic experiments demonstrated that, the new algorithm can achieve considerable better performance than the popular ℓ_1 method in sparse demixing problems.

REFERENCES

[1] F. Wen, L. Chu, P. Liu, and R. C. Qiu, "A survey on nonconvex regularization-based sparse and low-rank recovery in signal processing, statistics, and machine learning," *IEEE Access*, vol. 6, pp. 69883–69906, 2018.
[2] M. Elad and Y. Hel-Or, "A fast super-resolution reconstruction algorithm for pure translational motion and common space-invariant blur," *IEEE Trans. Image Process.*, vol. 10, no. 8, pp. 1187–1193, Aug. 2001.
[3] S. Mallat and Y. Guoshen, "Super-resolution with sparse mixing estimators," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2889–2900, Nov. 2010.
[4] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *Proc. 27th Annu. Conf. Comput. Graph. Interact. Techn.*, Jul. 2000, pp. 417–424.
[5] F. Wen, L. Adhikari, L. Pei, R. F. Marcia, P. Liu, and R. C. Qiu, "Nonconvex regularization-based sparse recovery and demixing with application to color image inpainting," *IEEE Access*, vol. 5, pp. 11513–11527, 2017.

[6] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.
[7] B. Popilka, S. Setzer, and G. Steidl, "Signal recovery from incomplete measurements in the presence of outliers," *Inverse Problems Imag.*, vol. 1, no. 4, pp. 661–672, Nov. 2007.
[8] E. J. Candes and P. A. Randall, "Highly robust error correction byconvex programming," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 2829–2840, Jul. 2008.
[9] R. Chan, C.-W. Ho, and M. Nikolova, "Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization," *IEEE Trans. Image Process.*, vol. 14, no. 10, pp. 1479–1485, Oct. 2005.
[10] T. Hashimoto, "Bounds on a probability for the heavy tailed distribution and the probability of deficient decoding in sequential decoding," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 990–1002, Mar. 2005.
[11] M. Elad, J.-L. Starck, P. Querre, and D. L. Donoho, "Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)," *Appl. Comput. Harmon. Anal.*, vol. 19, no. 3, pp. 340–358, 2005.
[12] J.-F. Cai, S. Osher, and Z. Shen, "Split Bregman methods and frame based image restoration," *Multiscale Model. Simul.*, vol. 8, no. 2, pp. 337–369, 2010.
[13] W. Gobel and F. Helmchen, "In vivo calcium imaging of neural network function," *Physiology*, vol. 22, no. 6, pp. 358–365, Dec. 2007.
[14] C. Studer, P. Kuppinger, G. Pope, and H. Bölcskei, "Recovery of sparsely corrupted signals," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 3115–3130, May 2012.
[15] M. B. McCoy, V. Cevher, Q. T. Dinh, A. Asaei, and L. Baldassarre, "Convexity in source separation: Models, geometry, and algorithms," *IEEE Signal Process. Mag.*, vol. 31, no. 3, pp. 87–95, May 2014.
[16] C. Studer and R. G. Baraniuk, "Stable restoration and separation of approximately sparse signals," *Appl. Comput. Harmon. Anal.*, vol. 37, no. 1, pp. 12–35, 2014.
[17] A. Adler, V. Emiya, M. G. Jafari, M. Elad, R. Gribonval, and M. D. Plumbley, "Audio inpainting," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 3, pp. 922–932, Mar. 2012.
[18] J. N. Laska, P. T. Boufounos, M. A. Davenport, and R. G. Baraniuk, "Democracy in action: Quantization, saturation, and compressive sensing," *Appl. Comput. Harmon. Anal.*, vol. 31, no. 3, pp. 429–443, Nov. 2011.
[19] F. Wen, P. Liu, Y. Liu, R. C. Qiu, and W. Yu, "Robust sparse recovery in impulsive noise via Lp-L1 optimization," *IEEE Trans. Signal Process.*, vol. 65, no. 1, pp. 105–118, Jan. 2017.
[20] R. Chartrand and V. Staneva, "Restricted isometry properties and nonconvex compressive sensing," *Inverse Problems*, vol. 24, May 2008, Art. no. 035020.
[21] J. Fan and R. Li, "Variable selection via nonconcave penalized likelihood and its oracle properties," *J. Amer. Stat. Assoc.*, vol. 96, pp. 1348–1360, Dec. 2001.
[22] S. Foucart and M.-J. Lai, "Sparsest solutions of underdetermined linear systems via ℓq-minimization for $0 < q \leq 1$," *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 395–407, May 2009.
[23] R. Chartrand and W. Yin, "Iteratively reweighted algorithms for compressive sensing," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Apr. 2008, pp. 3869–3872.
[24] M.-J. Lai, Y. Xu, and W. Yin, "Improved iteratively reweighted least squares for unconstrained smoothed ℓq minimization," *SIAM J. Numer. Anal.*, vol. 51, no. 2, pp. 927–957, Mar. 2013.
[25] R. Saab, R. Chartrand, and O. Yilmaz, "Stable sparse approximations via nonconvex optimization," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Apr. 2008, pp. 3885–3888.
[26] F. Wen, L. Pei, Y. Yang, W. Yu, and P. Liu, "Efficient and robust recovery of sparse signal and image using generalized nonconvex regularization," *IEEE Trans. Comput. Imag.*, vol. 3, no. 3, pp. 566–579, Dec. 2017.
[27] G. Marjanovic and V. Solo, "ℓ_q Sparsity penalized linear regression with cyclic descent," *IEEE Trans. Signal Process.*, vol. 62, no. 6, pp. 1464–1475, Mar. 2014.
[28] J. F. Yang and Y. Zhang, "Alternating direction algorithms for ℓ1-problems in compressive sensing," *SIAM J. Sci. Comput.*, vol. 33, pp. 250–278, Feb. 2011.
[29] Y. Li, Y. Lin, X. Cheng, Z. Xiao, F. Shu, and G. Gui, "Nonconvex penalized regularization for robust sparse recovery in the presence of SαS noise," *IEEE Access*, vol. 6, pp. 25474–25485, 2018.
[30] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.
[31] G. Marjanovic and V. Solo, "On ℓ_q optimization and matrix completion," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 5714–5724, Nov. 2012.
[32] L. Yunyi, D. Fei, C. Xiefeng, X. Li, and G. Guan, "Multiple-prespecified-dictionary sparse representation for compressive sensing image reconstruction with nonconvex regularization," *J. Franklin Inst.*, vol. 356, no. 4, pp. 2353–2371, Mar. 2019.
[33] E. Candes, M. Wakin, and S. Boyd, "Enhancing sparsity by reweighted L1 minimization," *J. Fourier Anal. Appl.*, vol. 14, no. 5, pp. 877–905, Dec. 2008.
[34] S. Roth and M. J. Black, "Fields of experts: A framework for learning image priors," in *Proc. IEEE Comput. Soc. Conf. Vis. Pattern Recognit.*, Jun. 2005, pp. 860–867.
[35] Y. Li, X. Cheng, and G. Gui, "Co-robust-ADMM-net: Joint ADMM framework and DNN for robust sparse composite regularization," *IEEE Access*, vol. 6, pp. 47943–47952, 2018.

JUNHUI MEI is currently pursuing the Ph.D. degree with the School of Mechanical Engineering, Shanghai Jiao Tong University. His major research interests include precision measurement, image processing, and machine vision.

JUNTONG XI is currently a Professor with the School of Mechanical Engineering, Shanghai Jiao Tong University. His major research interests include computer integrated manufacturing, precision measurement, micro-jet manufacturing, and medical device manufacturing.