A New Non-Convex Approach for Compressive Sensing MRI

Huihui Yue and Xiangjun Yin*

Abstract—Compressive sensing (CS) is an effective method for reconstructing magnetic resonance imaging (MRI) image from under-determined linear system (ULS). However, how to improve the accuracy of MRI image reconstructed by CS is still a serious problem, especially in noisy conditions. To solve this problem, in this paper, we propose a novel approach, dubbed as regularized maximum entropy function (RMEF) minimization algorithm. Specifically, motivated by the entropy function in information theory, we propose a maximum entropy function (MEF) to approximate $L^q$-norm ($0 < q < 1$) as sparsity promoting objectives, and then the regularization mechanism for improving the de-noising performance is adopted. Combining the above two ideas, a new objective function of RMEF method is proposed, and the global minimum is iteratively solved. We further analyze the convergence to verify the robustness of the RMEF algorithm. Experiments demonstrate the state-of-the-art performances of the proposed RMEF algorithm and show that the RMEF achieves higher PSNR and SSIM than other widely-adopted methods in MRI image recovery.

1. INTRODUCTION

MRI is a noninvasive imaging technology which plays an important role in clinical diagnosis. MRI uses different characteristics of energy attenuation released by different material structures in the environment, detects the information emitted by gradient magnetic fields, and realizes the mapping of the internal structure of the object, so as to achieve imaging. However, the limitation of the Shannon-Nyquist sampling theorem and the physical mechanism of magnetic resonance data encoding leads to the slow speed of conventional MRI. This is mainly because the MRI scanner captures not the original image pixel, but the global Fourier transform image in the frequency domain. Each frequency domain pixel is a linear combination of time domain pixels, which makes MRI data redundancy more obvious and thus increases the sampling time. In view of this, the accurate reconstruction of MRI with low sampling rate is of great significance.

As an undersampling MRI technique, compressive sensing magnetic resonance imaging (CS-MRI) breaks through the limitation of Shannon-Nyquist sampling theorem and offers us an efficient framework to deal with the above challenge. CS-MRI technology first undersamples data in $k$-space (i.e., Fourier space), then obtains image using CS theory [1–3]. Therefore, CS tasks eventually boil down to the MRI image recovery problem in the following ULS [4]:

$$y = Ax + w,$$

where $y \in \mathbb{R}^m$ is an under-sampled $k$-space data, $x \in \mathbb{R}^n$ the MRI image to be reconstructed, and $w$ the additive Gaussian noise. $A = \Phi D \in \mathbb{R}^{m \times n}$ is a measurement matrix, $\Phi \in \mathbb{R}^{m \times n}$ an undersampling matrix, and $D \in \mathbb{R}^{n \times n}$ a Fourier transform, $m \ll n$.

In CS, the MRI image $x$ is recovered from given $y$ and $A$. For this, $A$ has more columns than rows, which results in multiple solutions to recover $x$, thus making recovery of $x$ an ill-posed problem.
In addition, since the image of interest itself is sparse in $k$-space, the most straight-forward way is to search for the solution that also shares this sparse property [4]. Combining the above two points, we apply the $L_0$-norm to choose the sparsest one:

$$\min_x \|x\|_0, \quad s.t. \ y = Ax.$$  \hspace{1cm} (2)

This fairly simple form of Equation (2) is actually supported by theories [4]. According to these theories, when $x$ is sparse and $A$ satisfies the restricted isometry property (RIP) [4, 5], $x$ can be well restored. In fact, solving the Equation (2) is a nonconvex NP-hard problem whose solutions require an intractable combinatorial search [6]. To solve this problem, scholars have developed two alternative solutions:

- Greedy search method when $\|x\|_0 \leq k$;
- Relaxation method for $\|x\|_0$.

The sparsity of $x$ is $k$, which represents the upper bound on the number of nonzero entries in $x$. Greedy search method is mainly composed of a series of algorithms focusing on matching pursuit algorithms [7–10]. These algorithms have good results in noiseless environments, but they do not perform well in noisy environments. The main purpose of relaxation method for $\|x\|_0$ is to convert $L_0$-norm into other forms [11–13] equivalently so as to avoid the NP-hard problem. This kind of method is relatively scalable. Here this paper focuses on relaxation method that tries to solve the following unconstrained recovery problem:

$$\min_x \|y - Ax\|^2_2 + \lambda g(x).$$  \hspace{1cm} (3)

where $\lambda$ is the positive parameter used to balance the trade-off between data error $\|y - Ax\|^2_2$ and the sparsity regularizer $g(x)$. The role of $g(x)$ is to promote sparsity, and the function of $\|y - Ax\|^2_2$ is to ensure that the recovered image has certain fidelity whether it is noisy or noiseless. The combination of the two can improve the reconstruction accuracy, especially in the noisy case.

For relaxation methods, some convex relaxation methods are proposed, such as basis pursuit de-noise (BPDN) [14], SpaRSA [15], FISTA [16], and D-AMP [17, 18], which relax the $L_0$-norm regularization into the $L_1$-norm regularization. In other words, let $g(x)$ be the $L_1$-norm. This $L_1$-regularized problem can be efficiently tackled by convex optimization techniques if they meet some conditions, such as the RIP, null space property (NSP) [19], and incoherence condition (IC) [20]. However, due to the relaxation, the recovery accuracy is often degraded in noiseless case, e.g., it often introduces extra bias [21] and cannot reconstruct an image with the least observations [22]. Furthermore, for some applications, the result of the $L_1$-norm minimization is not sparse enough, and the original images cannot be recovered [23, 24].

To solve this problem, a number of state-of-the-art algorithms is proposed, which replace $g(x)$ with the $L_q$-norm ($0 < q < 1$) [25]. Algorithms in this category include StSALq [26], UALp [27], and $L_q$-RLS [28]. All these algorithms use a de-noising model in Equation (3) and solve an unconstrained $L_q$-norm regularized least squares ($L_q$-LS) problem. Compared with the $L_1$-norm, the $L_q$-norm is a closer approximation of the $L_0$-norm. It has been shown in [29] that under certain RIP condition of $A$, $L_q$-norm minimization algorithms require fewer sampling data but gain a better recovery performance than $L_1$-norm minimization algorithms. Moreover, the sufficient conditions in terms of RIP for $L_q$-norm minimization are weaker than those for $L_1$-norm minimization [30, 31]. However, in general, relative to $L_1$-norm minimization, $L_q$-norm minimization is more difficult to directly tackle due to its nonsmoothing. To solve this problem, $L_q$-RLS algorithm, which uses an efficient conjugate gradient (CG) method to solve a sequence of smooth subproblems, makes it succeed to handle large scale problem.

In fact, most of algorithms solve the $L_q$-norm minimization problem via smoothing (approximating) it, and for example, the works [32, 33] use an approximation of $\|x\|^q_0$ as

$$\|x\|_{q,\epsilon}^q = \sum_{i=1}^n (x_i^2 + \epsilon^2)^{q/2},$$  \hspace{1cm} (4)

where $\epsilon$ is a small enough positive constant. Furthermore, the iteratively re-weighted (IR) algorithm [34]
uses the following two penalties
\[ \|x\|_{q, \epsilon}^q = \sum_{i=1}^{n} (|x_i| + \epsilon)^{q-1} |x_i|, \]
\[ \|x\|_{q, \epsilon}^q = \sum_{i=1}^{n} (|x_i|^2 + \epsilon^2)^{q/2-1} |x_i|^2, \]
which explicitly relate to the \(L_q\)-norm approximation. Based on this, we obtain a smoothing function, MEF, to approximate \(L_q\)-norm, and the sparsity regularizer \(g(x)\) in Equation (3) is then replaced with MEF to form a new objective function. Moreover, iterative optimization method is used to tackle this nonsmooth problem. Hence, this paper develops a highly efficient numerical method for the nonconvex \(L_q\)-minimization problem.

The rest of the paper is organized as follows: Section 2 introduces the maximum entropy function for nonconvex \(L_q\)-norm functions and then proposes a new RMEF minimization algorithm to solve the ULS. Then the convergence of the RMEF minimization algorithm is proved in Section 3. In Section 4, the effect of the RMEF minimization algorithm in MRI image reconstruction with noise is verified by comparing it with L2-SL0 [35], UALp, and \(L_q\)-RLS approaches. Section 5 concludes this paper.

2. RELATED WORK

2.1. Maximum Entropy Function to Approximate \(L_q\)-Norm

Constructing a smooth \(L_q\)-norm is of great significance for CS, as shown in Equations (4–6), and \(L_q\)-norm can be represented as \(\|t\|_q^q = \sum_{i=1}^{n} |t_i|^q\), where \(t \in \mathbb{R}^n\), \(t \in \mathbb{R}\), and
\[ |t| = \begin{cases} t, & t > 0 \\ -t, & t \leq 0 \end{cases}, \]
so it can be converted as
\[ \tau(t) = |t| = \max\{t, -t\}. \]

Here, we introduce a maximum entropy function which can be given by
\[ \tau(t, \alpha) = \frac{1}{\alpha} \left\{ \log \left( 10^{\alpha t} + 10^{-\alpha t} \right) \right\}, \alpha > 0. \]

Then, for any \(\alpha > 0\),
\[ \tau(t, \alpha) = \tau(t) + \frac{1}{\alpha} \left\{ \log \left( 10^{\alpha(t-\tau(t))} + 10^{\alpha(-t-\tau(t))} \right) \right\}, \]
where \(0 \leq \left\{ \log(10^{\alpha(t-\tau(t))} + 10^{\alpha(-t-\tau(t))}) \right\} \leq \log 2\), which implies
\[ \tau(t) \leq \tau(t, \alpha) \leq \tau(t) + \frac{\log 2}{\alpha}. \]

From Equation (11), when \(\alpha \rightarrow \infty\), \(\tau(t, \alpha) \rightarrow \tau(t)\). Thus, the function \(\tau(t, \alpha)\) defined in Equation (9) is a smoothing approximation of the function \(\tau(t)\), which is an MEF.

Here let \(x = [x_1, x_2, \ldots, x_n]^T\), and moreover we make
\[ \Gamma(x, \alpha) = \lim_{\alpha \rightarrow \infty} \sum_{i=1}^{n} \tau(x_i, \alpha)^q = \lim_{\alpha \rightarrow \infty} \sum_{i=1}^{n} \left( \frac{1}{\alpha} \left\{ \log \left( 10^{\alpha x_i} + 10^{-\alpha x_i} \right) \right\} \right)^q, \]
and rewrite the \(L_q\)-norm as:
\[ \Gamma(x) = \sum_{i=1}^{n} \max\{x_i, -x_i\}^q. \]

Property 1 \(\forall x \in \mathbb{R}^n, \alpha > 0, \Gamma(x, \alpha)\) is an axisymmetric function, that is to say, \(\Gamma(x, \alpha) = \Gamma(-x, \alpha).\)
Theorem 1

\(\Gamma(\beta)\) is small enough, the iteration can be stopped. Thus, the algorithm is well reasonable.

\[\Gamma(x, \alpha) = \lim_{a \to -\infty} \sum_{i=1}^{n} \left( \frac{1}{\alpha} \log \left( 10^{a(x_i)} + 10^{-\alpha(x_i)} \right) \right)^q \]

\[\Gamma(x, \alpha) - \Gamma(-x, \alpha) = \lim_{a \to -\infty} \sum_{i=1}^{n} \left( \frac{1}{\alpha} \log \left( 10^{a(x_i)} + 10^{-\alpha(x_i)} \right) \right)^q - \lim_{a \to -\infty} \sum_{i=1}^{n} \left( \frac{1}{\alpha} \log \left( 10^{a(-x_i)} + 10^{-\alpha(-x_i)} \right) \right)^q \]

\[= \Gamma(x, \alpha) - \Gamma(-x, \alpha) = 0.\]

Therefore, \(\Gamma(x, \alpha) = \Gamma(-x, \alpha)\). To summarize, Property 1 is proved.

\[\text{Proof:} \] Please see Appendix A.

From Theorem 1, it can be seen that

\[0 \leq \Gamma(x, \alpha) - \Gamma(x) \leq n \left( \frac{\log 2}{\alpha} \right)^q. \tag{14}\]

Then we let \(\epsilon = n \left( \frac{\log 2}{\alpha} \right)^q\), and \(\epsilon\) is a small enough constant, so

\[0 \leq \Gamma(x, \alpha) - \Gamma(x) \leq \epsilon. \tag{15}\]

Therefore, \(\Gamma(x, \alpha)\) is equivalent to \(\Gamma(x)\), that is to say, the MEF can replace \(L_q\)-norm. Therefore, the proposed maximum entropy function model can be used for MRI image recovery.

2.2. Regularized Maximum Entropy Function Minimization Algorithm for MRI

Performing MRI image recovery under noise conditions is an important observational measure of the application of CS technology in practical fields. Traditional CS algorithms are extremely sensitive to noise due to the effects of noise folding [36, 37], thus result in poor recovery performance. Regularization mechanism can offset the impact of noise, so as to achieve the purpose of suppressing noise. This paper adopts this idea, and then the corresponding compression-aware objective function can be improved to

\[L_q(x, \lambda, \alpha) = \frac{1}{2} \| Ax - y \|^2 + \lambda \Gamma(x, \alpha), \tag{16}\]

where \(\lambda > 0\). The above equation is the model of the proposed RMEF minimization algorithm. As shown in this equation, this minimization problem must have a solution because \(L_q(x, \lambda, \alpha)\) is continuous with respect to \(x\), and it can achieve the minimum over a bounded set \(\{ x \mid \| x \|_2 \leq \Delta \}\), where \(\Delta\) is a positive constant. In addition, \(L_q(x, \lambda, \alpha)\) blows up as \(\| x \|_2 \to \infty\). Let \(\beta = \frac{1}{\alpha} > 0\), and \(x_{\beta, \lambda, q}\) denotes a critical point and it satisfies the first-order optimality condition

\[\lambda \frac{q}{\alpha^q-1} \frac{10^\beta - 10^-\beta}{10^\beta + 10^-\beta} \log^{q-1} \left( 10^\beta + 10^-\beta \right) + A_T (Ax - y) = 0. \tag{17}\]

Due to the underdetermination, there is no straightforward method to solve the above system of equations unless for specific instances, such as \(A^T A\) is a diagonal matrix. Combining with [32], we approximate solve the system with a sequence of \(\beta\), and the method is summarized as follows.

In the steps of proposed algorithm above, \(\beta\) is a descending order execution sequence formed by annealing mechanism. If \(| x^{(t+1)} - x^{(t)} | < \xi\), we choose \(x^{(t+1)}\) to be an approximation of the MRI image and stop the iteration. Otherwise, we stop the computation within a reasonable time and return the last \(x^{(t+1)}\). It is easy to know that the ULS in Equation (1) is invertible for any \(x^{(t+1)}\) if \(\beta > 0\). Once \(\beta\) is small enough, the iteration can be stopped. Thus, the algorithm is well reasonable.
Algorithm 1: Regularized Maximum Entropy Function (RMEF) Minimization Algorithm for MRI Image Reconstruction

Input: under-sampled k-space data $y$ and matrix $A$.
Output: MRI image $x$.

Choose appropriate parameters $\lambda > 0, q \in (0, 1)$.

Initialize $x^{(0)}$ such that $Ax^{(0)} = y$ and $\beta_0 = (\frac{1}{q})^\frac{1}{q}/\log 2$, $\beta = \beta_0, \beta_T = 10^{-6}$.

For $t = 0, 1, 2, \ldots, T$

Solve the following linear system for $x^{(t+1)}$
\[
q \cdot 10^\frac{x^{(t)}}{\beta} - 10^\frac{-x^{(t)}}{\beta} \log^q - 1 \left( 10^\frac{x^{(t)}}{\beta} + 10^\frac{-x^{(t)}}{\beta} \right) + \frac{1}{\lambda} A^T (Ax^{(t+1)} - y) = 0
\]

or equivalently
\[
x^{(t+1)} = \frac{1}{\lambda^2 A} (A^T y - \frac{q\lambda}{\alpha^q - 1} \cdot 10^{-x^{(t)}} + 10^{-x^{(t)}}) - 10^\frac{-x^{(t)}}{\beta} \log^q - 1 \left( 10^\frac{x^{(t)}}{\beta} + 10^\frac{-x^{(t)}}{\beta} \right)
\]

Update $\beta = \beta_0 10^{-\mu(t+1)}$, where $\mu = \frac{\log(\beta_0/\beta_T)}{\beta_T + 1}$.

if $|x^{(t+1)} - x^{(t)}| < \xi$ or $\beta = \beta_T$, then ends the cycling. $\xi$ is a small enough positive constant.

3. CONVERGENCE OF THE PROPOSED RMEF MINIMIZATION ALGORITHM

In this section, we discuss the convergence of the proposed RMEF minimization algorithm. Before analysis, we derive two simple properties of the first derivative of the function $\tau(x, \alpha)$, which will be used in the following analysis.

Proposition 1 Given $\alpha \in \mathbb{R}^+$, we have that

(i) $|\tau'(x, \alpha)| > 0.5(1 - 10^{-2})$ holds for any $x \in \mathbb{R}$ satisfying $|x| > \beta$;

(ii) $\lim_{\beta \to 0} \tau'(x, \alpha) = \text{sign}(x)$ holds for any $x \in \mathbb{R}$.

where $\beta = \frac{1}{\alpha}$, and $x$ is a scalar.

Proof sketch. (i) For any $|x| > \beta$, there will be either $x > \beta$ or $x < -\beta$.

- If $x > \beta$, there will be $-x < -\beta$, hence, $\frac{x}{\beta} > 1$ and $\frac{-x}{\beta} < -1$. Furthermore,

\[
10^{\frac{x}{\beta}} > 10^1 \quad \text{and} \quad 0 < 10^{\frac{-x}{\beta}} < 10^{-1},
\]

which imply that $10^{\frac{x}{\beta}} - 10^{\frac{-x}{\beta}} > 10^1 - 10^{-1} > 0$. Thus,

\[
|\tau'(x, \alpha)| = \frac{10^{\frac{x}{\beta}} - 10^{\frac{-x}{\beta}}}{10^\frac{x}{\beta} + 10^\frac{-x}{\beta}} > \frac{10^{\frac{x}{\beta}} - 10^{1(-1)}}{2 \times 10^{\frac{x}{\beta}}}
\]

\[
= 0.5 \left( 1 - \frac{10^{1(-1)}}{10^{\frac{x}{\beta}}} \right) > 0.5 \left( 1 - 10^{-2} \right).
\]

- If $x < -\beta$, there will be $-x > \beta$. By the fact that $\tau'(x, \alpha)$ is odd function and the above inequality, we have that

\[
|\tau'(x, \alpha)| = |\tau'(-x, \alpha)| > 0.5 \left( 1 - 10^{-2} \right).
\]

(ii) If $x > 0$, there will be $10^{\frac{x}{\beta}} \to \infty$ and $10^{\frac{-x}{\beta}} \to 0$ when $\beta \to 0$. Hence,

\[
\lim_{\beta \to 0} \tau'(x, \alpha) = \lim_{\beta \to 0} \frac{10^{\frac{x}{\beta}} - 10^{\frac{-x}{\beta}}}{10^\frac{x}{\beta} + 10^\frac{-x}{\beta}} = \lim_{\beta \to 0} \left\{ 1 - \frac{2 \times 10^{\frac{x}{\beta}}}{10^\frac{x}{\beta} + 10^\frac{-x}{\beta}} \right\} = 1.
\]
we can get
\[ \lim_{\beta \to 0} \tau'(x, \alpha) = - \lim_{\beta \to 0} \tau'(-x, \alpha) = -1. \]

- If \( x = 0 \), there will be \( 10^{\frac{q}{2}} = 10^{-\frac{q}{2}} = 1 \), hence \( \tau'(x, \alpha) = 0 \). Therefore, the Proposition 1 is proved.

In addition, we need the following notation and results. Denote
\[ \Gamma'(x, \alpha) := \nabla \Gamma(x, \alpha) = \left( \frac{\partial \Gamma(x, \alpha)}{\partial x_1}, \frac{\partial \Gamma(x, \alpha)}{\partial x_2}, \ldots, \frac{\partial \Gamma(x, \alpha)}{\partial x_n} \right)^T, \]
\[ \frac{\partial \Gamma(x, \alpha)}{\partial x_i} = q\tau^{q-1}(x_i, \alpha)\tau'(x_i, \alpha), i = 1, \ldots, n. \]

The first-order necessary condition of \( L_q(x, \lambda, \alpha) \) is given in Equation (17), which means \( \chi \) that satisfied with the above equation is a stationary point of \( L_q(x, \lambda, \alpha) \) if it satisfies Equation (17).

Moreover, for given \( q \in (0, 1) \) and \( \lambda > 0 \), the definition of \( L_q \) regularized problem is shown as follows:
\[ L_q(x, \lambda) = \lambda \| x \|_q^q + \| Ax - y \|_2^2, \] (18)
then \( x \in \mathbb{R}^n \) is called a stationary point of \( L_q \) regularized problem if it satisfies [38]
\[ \lambda q |x|^q + \begin{bmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & x_n \end{bmatrix} A^T(Ax - y) = 0. \] (19)

**Theorem 2** Suppose that the sequence \( \{x_{\beta t}\} \) generated by proposed RMEF minimization algorithm is contained in the level set \( \{x| L_q(x, \lambda, \alpha) \leq L_q(x^0, \lambda, \alpha)\} \) for an arbitrary given initial point \( x^0 \).
(i) For any accumulation point of \( \{x_{\beta t}\} \) is a stationary point of \( L_q(x, \lambda) \).
(ii) Let \( \{x_{\beta t}\} \) be a global minimizer of \( \{x| L_q(x, \alpha, \lambda)\} \) for any given \( t \), then any accumulation point of \( \{x_{\beta t}\} \) is a global minimizer of \( L_q(x, \lambda) \).

**Proof** Please see Appendix B.

4. RESULT AND ANALYSIS

Here we verify the MRI image recovery performance of the proposed RMEF minimization algorithm by comparing it with the state-of-the-art L2-SLO, UALp, and \( L_q \)-RLS algorithms on Brain \((256 \times 256)\) datasets. The numerical simulation platform is MATLAB 2017b, which is installed on the WINDOWS 10, 64-bit operating system. The type of CPU is Inter (R) Core (TM) i5-3230M, and the frequency is 2.6 GHz.

Given a pair of \( \{y, \Phi, D, w\} \), we try to obtain the original MRI image. The recovery process of the MRI image by the RMEF minimization algorithm is shown in Figure 1, while the image recovery performance is valuated by Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index (SSIM). PSNR is defined as
\[ \text{PSNR} = 10 \log \left( \frac{255^2}{\text{MSE}} \right), \] (20)
where \( \text{MSE} = \| x - \hat{x} \|_2^2 \), and SSIM is defined as
\[ \text{SSIM}(p, q) = \frac{(2\mu_p + \mu_q + c_1)(2\sigma_{pq} + c_2)}{(\mu_p^2 + \mu_q^2 + c_1)(\sigma_p^2 + \sigma_q^2 + c_2)}, \] (21)
where \( \mu_p \) is the mean of image \( p \), \( \mu_q \) the mean of image \( q \), \( \sigma_p \) the variance of image \( p \), \( \sigma_q \) the variance of image \( q \), and \( \sigma_{pq} \) the covariance between image \( p \) and image \( q \). Parameters \( c_1 = z_1L \) and \( c_2 = z_2L \),
in which \( z_1 = 0.01, z_2 = 0.03 \) and \( L \) is dynamic range of pixel values. Range of SSIM is \([-1, 1]\), and if the two compared images are same, the SSIM equals 1.

Figures 2, 3, and 4 show the MRI image recovery effect of the selected algorithms when compression ratio (CR, which is defined as \( m/n \)) is respectively 0.4, 0.5, 0.6, and noise intensity \( \delta_N \) equals 0.01. As shown in these figures, under any same CR, the proposed RMEF minimization algorithm can recover a clearer MRI image than the other three algorithms. However, the differences of each algorithm in image recovery performance are not obvious, so the recovery performance cannot be determined. Further, Table 1 shows the difference in detail through scientific data. From this table, it can be seen that the proposed RMEF minimization algorithm performs better than the other three methods in Brain recovery, which is consistent with the results in Figures 2, 3, and 4. Compared with \( L_q \)-RLS algorithm, which is the best one among the other three algorithms, the proposed RMEF minimization algorithm can improve approximately 0.7038 dB, 0.6027 dB, and 0.0249 dB on PSNR and 0.5%, 0.31%, and 0.01%
Figure 4. *Brain* image recovery effect by different algorithms with $\delta_N = 0.01$, $CR = 0.6$.

Table 1. PSNR and SSIM analysis of *Brain* recovered by the L2-SL0, UALp, $L_q$-RLS, and proposed RMEF minimization algorithms with $CR = 0.4, 0.5, 0.6$ and $\delta_N = 0.01$.

<table>
<thead>
<tr>
<th>CR</th>
<th>PSNR (dB)</th>
<th>SSIM (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L2-SL0</td>
<td>UALp</td>
</tr>
<tr>
<td>0.4</td>
<td>25.4844</td>
<td>25.9210</td>
</tr>
<tr>
<td>0.5</td>
<td>26.6514</td>
<td>26.9360</td>
</tr>
<tr>
<td>0.6</td>
<td>26.9159</td>
<td>27.0416</td>
</tr>
</tbody>
</table>

Table 2. PSNR and SSIM analysis of *Brain* recovered by the proposed RMEF minimization algorithm under different $\delta_N$.

<table>
<thead>
<tr>
<th>$\delta_N$</th>
<th>MRI image</th>
<th>PSNR (dB)</th>
<th>SSIM (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td><em>Brain</em></td>
<td>30.9261</td>
<td>98.85</td>
</tr>
<tr>
<td>0.05</td>
<td><em>Brain</em></td>
<td>27.8016</td>
<td>97.64</td>
</tr>
<tr>
<td>0.1</td>
<td><em>Brain</em></td>
<td>24.5298</td>
<td>95.12</td>
</tr>
<tr>
<td>0.2</td>
<td><em>Brain</em></td>
<td>18.4506</td>
<td>82.59</td>
</tr>
<tr>
<td>0.3</td>
<td><em>Brain</em></td>
<td>13.6842</td>
<td>60.92</td>
</tr>
<tr>
<td>0.5</td>
<td><em>Brain</em></td>
<td>8.1555</td>
<td>29.49</td>
</tr>
</tbody>
</table>

In conclusion, experiments verify the most advanced MRI image recovery performance of the RMEF minimization algorithm. Therefore, it is feasible for the RMEF minimization algorithm to be applied to MRI image recovery.
Figure 5. Brain image recovery effect by the proposed RMEF minimization algorithm with \(\delta_N = [0, 0.05, 0.1, 0.2, 0.3, 0.5]\).

5. CONCLUSIONS

In this paper, a non-convex RMEF minimization algorithm is proposed to reconstruct MRI images with noise. We firstly propose a MEF and use the property that MEF approximates to \(L_q\)-norm to promote the sparsity of the recovered images. On this basis, we replace sparsity regularizer \(g(x)\) with proposed MEF and form a new approach called RMEF minimization algorithm. We also prove that the proposed RMEF minimization algorithm can converge to optimal solution of MRI images. Finally, experiments show that the proposed RMEF performs better than the popular \(\|x\|_q^q\) regularization algorithms in MRI image recovery. Additionally, in the future, we would also like to apply the RMEF minimization algorithm to other CS applications such as the Blind Source Separation (BSS) [39], Robust Principal Component Analysis (RPCA) [40], and Dictionary Learning [41, 42].

ACKNOWLEDGMENT

Thanks to Huihui Yue for her work in paper layout, paper writing, paper revision and experimental simulation, and Xiangjun Yin for his work in paper supervision, paper revision and experiment correction. Finally, thanks to the 531 Laboratory of School of Electrical and Information Engineering of Tianjin University for its support.

APPENDIX A. PROOF OF THEOREM 1

Proof Known by Equation (11), \(\forall x > 0, \alpha > 0\), there will be

\[
\max\{x, -x\} \leq \frac{1}{\alpha} \left\{ \log \left( 10^{\alpha x} + 10^{-\alpha x} \right) \right\} \leq \max\{x, -x\} + \frac{\log_2 2}{\alpha}.
\]
According to the definitions of $\Gamma(x, \alpha)$ and $\Gamma(x)$, we can get
\[
\Gamma(x, \alpha) - \Gamma(x) = \lim_{\alpha \to \infty} \sum_{i=1}^{n} \left( \frac{1}{\alpha} \{ \log (10^{\alpha x_i} + 10^{-\alpha x_i}) \} \right)^q - \sum_{i=1}^{n} \max\{x_i, -x_i\}^q
\leq \sum_{i=1}^{n} \left( \frac{1}{\alpha} \{ \log (10^{\alpha x_i} + 10^{-\alpha x_i}) \} - \max\{x_i, -x_i\} \right)^q
\leq \sum_{i=1}^{n} \left( \frac{\log 2}{\alpha} \right)^q
= n \left( \frac{\log 2}{\alpha} \right)^q,
\]
and according to Equation (11), we can obtain
\[
\Gamma(x) \leq \Gamma(x, \alpha) \leq \Gamma(x) + n \left( \frac{\log 2}{\alpha} \right)^q. \tag{A2}
\]

Moreover, according to Property 1, $\Gamma(x, \alpha)$ is an axisymmetric function. For any $x_i \leq 0$, the result of Equation (A2) is still obtained. Therefore, the Theorem 1 is proved.

APPENDIX B. PROOF OF THEOREM 2

**Proof** Since $\{x_{\beta t}\} \subseteq \{x | L_q(x, \lambda, \alpha) \leq L_q(x^0, \lambda, \alpha)\}$, we always have $L_q(x_{\beta t}, \lambda, \alpha) \leq L_q(x^0, \lambda, \alpha)$ for any $t$. It can be obtained that the sequence $\{x_{\beta t}\}$ is bounded. In fact, suppose that $\{x_{\beta t}\}$ is unbounded, then for any given $q \in (0, 1)$, $\{\|x_{\beta t}\|_q\}$ is unbounded. Thus, by combining $L_q(x, \lambda, \alpha)$, we have that for any given $q \in (0, 1)$ and $\lambda > 0$, $L_q(x_{\beta t}, \lambda, \alpha) \to +\infty$ when $t \to \infty$, which contradicts the result that $L_q(x_{\beta t}, \lambda, \alpha) \leq L_q(x^0, \lambda, \alpha)$ for all $t$.

Thus, there exists at least a convergent subsequence of $\{x_{\beta t}\}$. Let $\tilde{x}$ be an accumulation point of $\{x_{\beta t}\}$, and denote $\lim_{t \to \infty} \{x_{\beta t}\} = \tilde{x}$ without loss of generality. Here we denote $\hat{x}_{\beta t} = \text{diag}(x_{\beta t})$ and $\tilde{x} = \text{diag}(\tilde{x})$.

(i) We can see that $x_{\beta t}$ satisfies the first-order necessary condition of $L_q(x, \alpha, \lambda)$. Thus, by Equation (17), we have
\[
\nabla L_q(x_{\beta t}, \alpha, \lambda) = \lambda \Gamma'(x_{\beta t}, \alpha) + A^T (Ax_{\beta t} - y) = 0.
\]
where
\[
\Gamma'(x_{\beta t}, \alpha) = \left( \frac{\partial \Gamma(x_{\beta t}, \alpha)}{\partial (x_{\beta t})_1}, \ldots, \frac{\partial \Gamma(x_{\beta t}, \alpha)}{\partial (x_{\beta t})_n} \right)^T,
\]
and for any $i = 1, \ldots, n$,
\[
\frac{\partial \Gamma(x_{\beta t}, \alpha)}{\partial (x_{\beta t})_i} = q \tau^{q-1}((x_{\beta t})_i, \alpha) \tau'((x_{\beta t})_i, \alpha).
\]
Moreover,
\[
\hat{x}_{\beta t} \nabla L_q(x_{\beta t}, \alpha, \lambda) = \lambda \hat{x}_{\beta t} \Gamma'(x_{\beta t}, \alpha) + \hat{x}_{\beta t} A^T (Ax_{\beta t} - y) = 0,
\]
and
\[
[\hat{x}_{\beta t} \Gamma'(x_{\beta t}, \alpha)]_i = q ((x_{\beta t})_i, \alpha) \tau^{q-1}((x_{\beta t})_i, \alpha) \\
\rightarrow q \hat{x}_i \tau^{q-1}(\hat{x}_i) \text{sign}(\hat{x}_i) \\
= q |\hat{x}_i|^q,
\]
where the second step follows from Equation (8) and Proposition 1. Therefore,
\[
0 = \lim_{t \to \infty} \lambda \hat{x}_{\beta t} \Gamma'(x_{\beta t}, \alpha) + \lim_{t \to \infty} \hat{x}_{\beta t} A^T (Ax_{\beta t} - y) = \lambda q |\hat{x}|^q + \hat{x} A^T (A \hat{x} - y).
\]
i.e., $\tilde{x}$ satisfies Equation (18), which implies that $\tilde{x}$ is a stationary point of $L_q(x, \lambda)$.

(ii) Let $x^*$ be a global minimizer of $L_q(x, \lambda)$, then, we have
\[
L_q(x^*, \lambda) \leq L_q(x_{\beta t}, \lambda) \leq L_q(x^*, \alpha, \lambda) \leq L_q(x^*, \alpha) \leq L_q(x^*, \lambda) + \lambda n(\beta \log 2)^q.
\]

Take $t \to \infty$, we obtain that $L_q(\tilde{x}, \lambda) = L_q(x^*, \alpha)$. Hence, any accumulation point of $\{x_{\beta t}\}$ is a global minimizer of $L_q(x, \lambda)$. So the proof is complete. According to the Theorem, we can see that the convergence of $L_q(x, \alpha, \lambda)$ and $L_q(x, \lambda)$ is equivalent. Hence, our proposed algorithm can converge to global minimizer of $L_q(x_{\beta t}, \lambda)$. So, the convergence is simply analyzed.

REFERENCES

