



2017

## Sparse Signal Recovery via Exponential Metric Approximation

Jian Pan

*Department of Electronic Engineering, Tsinghua University, Beijing 100084, China.*


Jun Tang

*Department of Electronic Engineering, Tsinghua University, Beijing 100084, China.*

Wei Zhu

*Department of Electronic Engineering, Tsinghua University, Beijing 100084, China.*

Follow this and additional works at: <https://tsinghuauniversitypress.researchcommons.org/tsinghua-science-and-technology>

 Part of the [Computer Sciences Commons](#), and the [Electrical and Computer Engineering Commons](#)

### Recommended Citation

Jian Pan, Jun Tang, Wei Zhu. Sparse Signal Recovery via Exponential Metric Approximation. *Tsinghua Science and Technology* 2017, 22(1): 104-111.

This Research Article is brought to you for free and open access by Tsinghua University Press: Journals Publishing. It has been accepted for inclusion in *Tsinghua Science and Technology* by an authorized editor of Tsinghua University Press: Journals Publishing.

# Sparse Signal Recovery via Exponential Metric Approximation

Jian Pan\*, Jun Tang, and Wei Zhu

**Abstract:** Sparse signal recovery problems are common in parameter estimation, image processing, pattern recognition, and so on. The problem of recovering a sparse signal representation from a signal dictionary might be classified as a linear constraint  $\ell_0$ -quasinorm minimization problem, which is thought to be a Non-deterministic Polynomial-time (NP)-hard problem. Although several approximation methods have been developed to solve this problem via convex relaxation, researchers find the nonconvex methods to be more efficient in solving sparse recovery problems than convex methods. In this paper a nonconvex Exponential Metric Approximation (EMA) method is proposed to solve the sparse signal recovery problem. Our proposed EMA method aims to minimize a nonconvex negative exponential metric function to attain the sparse approximation and, with proper transformation, solve the problem via Difference Convex (DC) programming. Numerical simulations show that exponential metric function approximation yields better sparse recovery performance than other methods, and our proposed EMA-DC method is an efficient way to recover the sparse signals that are buried in noise.

**Key words:** sparse recovery; exponential metric approximation; sparsity tolerance; DC optimization; signal-to-noise-ratio

## 1 Introduction

The Compressive Sensing (CS)<sup>[1,2]</sup> theory launched a profound revolution in signal processing, machine learning, and statistics over the past ten years, and in its generalization and application, it has brought many inspiring results that have fundamentally changed our understanding of data sampling and storing. In its application, many CS algorithms have been proposed to fast solve the sparse recovery problem. The goal of sparse recovery is to solve the following underdetermined linear system:

$$\mathbf{y} = \mathbf{A}\mathbf{x} \quad (1)$$

where  $\mathbf{y} \in \mathbb{R}^M$  denotes the measurement vector,

• Jian Pan, Jun Tang, and Wei Zhu are with Department of Electronic Engineering, Tsinghua University, Beijing 100084, China. E-mail: panj13@mails.tsinghua.edu.cn.

\* To whom correspondence should be addressed.

Manuscript received: 2016-03-14; revised: 2016-10-24; accepted: 2016-11-11

$\mathbf{A} \in \mathbb{R}^{M \times N}$  is the sensing matrix, and  $\mathbf{x} = [x_1, x_2, \dots, x_N]^T \in \mathbb{R}^N$  refers to the original sparse signal (namely, one vector with only a few nonzero components) to be recovered. With regard to this underdetermined equation, i.e.,  $M < N$ , the problem is of course ill-posed. However, if  $\mathbf{x}$  is known to be sparse then the problem is solvable.

To recover the sparse signal  $\mathbf{x}$ , researchers typically seek the sparsest signal as the optimum solution, which is identical to solving a  $\ell_0$ -quasinorm minimization problem:

$$\underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{x}\|_0, \quad \text{subject to } \mathbf{y} = \mathbf{A}\mathbf{x} \quad (2)$$

where  $\|\mathbf{x}\|_0 \triangleq |\{i | x_i \neq 0\}|$  denotes the  $\ell_0$ -quasinorm of  $\mathbf{x}$ , which counts the number of nonzero elements of  $\mathbf{x}$ . However, adopting this method is not practical since it is usually solved by a combinatorial search, which is Non-deterministic Polynomial-time (NP)-hard<sup>[1]</sup>. Researchers have found that under some assumptions Problem (2) can be exactly solved by replacing the  $\ell_0$ -quasinorm with the  $\ell_1$ -norm<sup>[3]</sup>, i.e.,

$$\operatorname{argmin}_x \|\mathbf{x}\|_1, \quad \text{subject to } \mathbf{y} = \mathbf{A}\mathbf{x} \quad (3)$$

The  $\ell_1$ -norm minimization Problem (3) is a convex optimization problem that can be transformed into a Linear Program (LP) under certain conditions, such as the Restricted Isometry Property (RIP) condition<sup>[4]</sup>, whereby the optimal solution of  $\ell_1$ -norm minimization is exactly equal to that of  $\ell_0$ -quasinorm minimization. Another family of sparse recovery algorithms is based on non-convex optimization, such as  $\ell_p$ -quasinorm ( $0 < p < 1$ ) minimization<sup>[5]</sup>, i.e.,

$$\operatorname{argmin}_x \|\mathbf{x}\|_p^p, \quad \text{subject to } \mathbf{y} = \mathbf{A}\mathbf{x} \quad (4)$$

$\ell_p$ -quasinorm ( $\|\mathbf{x}\|_p^p = \sum_i |x_i|^p$ ) minimization is a nonconvex optimization problem that can be efficiently solved via the iterative hard thresholding algorithm<sup>[6]</sup>. Another nonconvex method known as the iterative reweighted least-squares<sup>[7]</sup> has also been studied. Although these nonconvex methods have shown good sparse recovery ability, there is plenty of room for the study of other sparse metric functions to improve sparse recovery performance, e. g., using fewer samples to perform signal processing while maintaining adequate overall performance.

The main contributions of this paper are twofold. First, a negative exponential metric approximation function is introduced to solve the original sparse recovery problem by transforming the problem into a solvable DC optimization problem via DC programming. Second, simulation results show that our proposed EMA sparse recovery method is with greater sparsity tolerance than the other methods and the proposed method is an efficient way to recover sparse signals that buried in noise.

The rest of this paper is organized as follows. In Section 2, a negative exponential metric function is introduced for approximating the original sparse recovery problem, which transforms the original  $\ell_0$  optimum problem into Exponential Metric Approximation (EMA) optimization problem. In Section 3, some derivations are made to convert the EMA optimization problem into a standard DC programming problem. In Section 4, an iterative DC algorithm is formulated to solve the EMA optimization problem. In Section 5, numerical simulations are completed to test the efficiency of the proposed method. In Section 6, the conclusion is drawn.

## 2 Problem Formulation

In order to solve the original Problem (2), we found the

decreased exponential function is a fine metric function to approximate the  $\ell_0$ -quasinorm

$$f_\alpha(t) = 1 - e^{-|t|/\alpha}, \quad \forall t \in \mathbb{R} \quad (5)$$

$$F_\alpha(\mathbf{x}) = \sum_{i=1}^N f_\alpha(x_i) = N - \sum_{i=1}^N e^{-|x_i|/\alpha} \quad (6)$$

where  $\alpha \in \mathbb{R}_+$  refers to some positive parameters that control the convergence. The figure of this metric function as well as  $\ell_0$ ,  $\ell_{1/2}$ , and  $\ell_1$  are shown in Fig. 1.

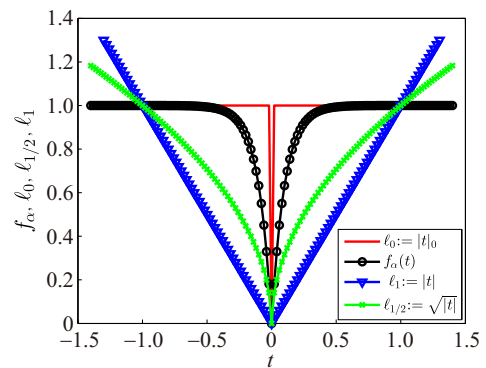
Given that

$$\lim_{\alpha \rightarrow 0^+} f_\alpha(t) = \operatorname{sign}(|t|) = \|t\|_0 = \begin{cases} 1, & t \neq 0; \\ 0, & t = 0 \end{cases} \quad (7)$$

then with  $\alpha \rightarrow 0^+$ ,  $F_\alpha(\mathbf{x}) \rightarrow \|\mathbf{x}\|_0$ . On this basis, we propose the following non-convex optimization problem to approximate Problem (2):

$$\begin{aligned} \text{(EMA)} \quad & \operatorname{argmin}_{\mathbf{x} \in \mathbb{R}^N} F_\alpha(\mathbf{x}) = N - \sum_{i=1}^N e^{-|x_i|/\alpha}, \\ & \text{subject to } \quad \mathbf{y} = \mathbf{A}\mathbf{x} \end{aligned} \quad (8)$$

Note that this is an EMA sparse problem whose constraint domain is convex and object function is nonconvex. Since the object function  $F_\alpha(\mathbf{x})$  is not smooth at the point  $\mathbf{x} = \mathbf{0}$  then the traditional method can not work properly, thus Problem (8) remains difficult to be solved. After much trial and error using many methods, we found that if the sparse signal  $\mathbf{x}$  is nonnegative, i.e.,  $\mathbf{x} \in \mathbf{D} \triangleq \{\mathbf{x} | x_i \geq 0, 1 \leq i \leq N\}$ , then Problem (8) is easy to solve by decomposing the objective function  $F_\alpha(\mathbf{x})$  into a difference of two convex functions, which can be solved via Difference of Convex functions Programming (DCP) method<sup>[8–10]</sup>. In the next section the details of converting the EMA model to standard DC programming are formulated.



**Fig. 1** Relationship between the negative exponential metric  $f_\alpha(\cdot)$  and  $\ell_0, \ell_{1/2}, \ell_1$  when  $\alpha = 0.1$ .

### 3 Converting EMA Model to DC Programing Model

In order to make the problem of minimizing function  $F_\alpha(\mathbf{x})$  solvable, we utilize the nonnegative vector property in the EMA model. Based on our analysis in the section above, we do this by splitting the desired variable  $\mathbf{x}$  into its positive and negative parts. Formally, we introduce vectors  $\mathbf{u} = [u_1, \dots, u_N]^T, \mathbf{v} = [v_1, \dots, v_N]^T$  and make the following substitutions.

Let

$$\begin{cases} \mathbf{u} = \mathbf{x}_+, & \mathbf{u} \in \mathbb{R}_+^N; \\ \mathbf{v} = \mathbf{x}_-, & \mathbf{v} \in \mathbb{R}_+^N \end{cases} \quad (9)$$

where

$$\begin{cases} u_j \triangleq \max\{0, x_j\}, & \forall 1 \leq j \leq N; \\ v_i \triangleq \max\{0, -x_i\}, & \forall 1 \leq i \leq N. \end{cases}$$

Then,  $\mathbf{x}$  can be rewritten as  $\mathbf{x} = \mathbf{u} - \mathbf{v}$  which make Problem (8) identical to the following:

$$\begin{aligned} \operatorname{argmin}_{\mathbf{u}, \mathbf{v} \in \mathbb{R}_+^N} F_\alpha(\mathbf{u} - \mathbf{v}) &= N - \sum_{i=1}^N e^{-(u_i+v_i)/\alpha}, \\ \text{subject to } \mathbf{y} &= A(\mathbf{u} - \mathbf{v}) \end{aligned} \quad (10)$$

Therefore, we can decompose the objective function  $F_\alpha(\mathbf{u}, \mathbf{v})$  into the difference of two convex functions  $g(\mathbf{u}, \mathbf{v})$  and  $h(\mathbf{u}, \mathbf{v})$ , which works due to the following lemma.

**Lemma 1** Let  $g(\mathbf{u}, \mathbf{v}) \triangleq N + \sum_{i=1}^N (u_i + v_i)^2$ ,  $h(\mathbf{u}, \mathbf{v}) \triangleq \sum_{i=1}^N [e^{-(u_i+v_i)/\alpha} + (u_i + v_i)^2]$ , and then functions  $g(\mathbf{u}, \mathbf{v})$  and  $h(\mathbf{u}, \mathbf{v})$  are both convex functions of  $(\mathbf{u}, \mathbf{v})$ , where  $(\mathbf{u}, \mathbf{v}) \in \mathbb{R}_+^N \times \mathbb{R}_+^N$ .

**Proof** The convexity of functions  $g$  and  $h$  is obvious.

With respect to function  $\zeta(x) = e^{-\frac{x}{\alpha}}$  we can obtain its second-order derivative  $\zeta''(x) = \frac{1}{\alpha^2} e^{-\frac{x}{\alpha}}$ . In the same way we can obtain the Hessian matrix of  $h(\mathbf{u}, \mathbf{v})$  as follows:

$$\mathbf{J}_1 \triangleq \nabla^2 h(\mathbf{u}, \mathbf{v}) = \frac{\partial^2 h}{\partial(\mathbf{u}, \mathbf{v})\partial(\mathbf{u}, \mathbf{v})} = \begin{bmatrix} \mathbf{B} & \mathbf{B} \\ \mathbf{B} & \mathbf{B} \end{bmatrix},$$

$$\mathbf{B} = [b_{i,j}]_{N \times N},$$

$$b_{i,j} \triangleq \frac{\partial^2 h}{\partial u_i \partial u_j} = \begin{cases} \lambda_i \triangleq 2 + \frac{1}{\alpha^2} e^{-\frac{(u_i+v_i)}{\alpha}}, & i = j; \\ 0, & i \neq j; \end{cases}$$

where  $\mathbf{B}$  is a diagram matrix and the eigen value of  $\mathbf{J}_1$  is  $\lambda = \lambda_i$ , with  $1 \leq i \leq N$  and  $\lambda = \lambda_j = 0, N \leq j \leq 2N$ . Since  $\lambda(\mathbf{J}_1) \geq 0$ , then the matrix  $\mathbf{J}_1$  is semi-positive, so the function  $h(\mathbf{u}, \mathbf{v})$  is a convex function of  $(\mathbf{u}, \mathbf{v})$ .

Similarly, we can obtain the Hessian matrix of  $g(\mathbf{u}, \mathbf{v})$  as follows:

$$\mathbf{J}_2 \triangleq \nabla^2 g(\mathbf{u}, \mathbf{v}) = \frac{\partial^2 g}{\partial(\mathbf{u}, \mathbf{v})\partial(\mathbf{u}, \mathbf{v})} = \begin{bmatrix} 2\mathbf{I} & 2\mathbf{I} \\ 2\mathbf{I} & 2\mathbf{I} \end{bmatrix}.$$

It is clear that  $\mathbf{J}_2 \geq 0$ , therefore  $g$  is also a convexity function of  $(\mathbf{u}, \mathbf{v})$ . ■

Lemma 1 suggests that the object function of Problem (10) can be decomposed as  $F_\alpha(\mathbf{u} - \mathbf{v}) = g(\mathbf{u}, \mathbf{v}) - h(\mathbf{u}, \mathbf{v})$ , which is the difference of two convex functions. Thus it can be recast to DC programming problem as follows:

$$\begin{aligned} \operatorname{argmin}_{\mathbf{u}, \mathbf{v} \in \mathbb{R}_+^N} F_\alpha(\mathbf{u} - \mathbf{v}) &= g(\mathbf{u}, \mathbf{v}) - h(\mathbf{u}, \mathbf{v}), \\ \text{s.t. } \begin{cases} \mathbf{y} &= A(\mathbf{u} - \mathbf{v}); \\ h(\mathbf{u}, \mathbf{v}) &= \sum_{i=1}^N [e^{-(u_i+v_i)/\alpha} + (u_i + v_i)^2]; \\ g(\mathbf{u}, \mathbf{v}) &= N + \|\mathbf{u} - \mathbf{v}\|^2 \end{cases} \end{aligned} \quad (11)$$

Note that, the first constraint of Problem (11) is a  $2N$ -dim superplane, which is a convex set, so the problem is a DC programming problem.

### 4 Sparse Recovery Algorithm Based on DC Programing

#### 4.1 Preliminary

**Definition 1** ( $\rho$ -convexity<sup>[11]</sup>) A real-valued function  $\phi(t)$  defined on a convex subset  $S \subseteq \mathbb{R}^N$  is called  $\rho$ -convex if there exists some real number  $\rho$  which is the largest real number that the inequality

$$\phi(\lambda \mathbf{t}_1 + (1 - \lambda)\mathbf{t}_2) \leq \lambda \phi(\mathbf{t}_1) + (1 - \lambda)\phi(\mathbf{t}_2) - \rho \lambda(1 - \lambda) \|\mathbf{t}_1 - \mathbf{t}_2\|^2$$

holds for  $\forall \mathbf{t}_1, \mathbf{t}_2 \in S, \forall \lambda \in (0, 1)$ .

**Definition 2** (Subgradient of convex function<sup>[12,13]</sup>) Let  $f(\mathbf{x})$  be a convex function with  $\operatorname{dom}(f) \subseteq \mathbb{R}^N$ , then the set

$\partial f(\mathbf{x}_0) \triangleq \{\boldsymbol{\xi} \in \mathbb{R}^N \mid f(\mathbf{x}) - f(\mathbf{x}_0) \geq \boldsymbol{\xi}^T(\mathbf{x} - \mathbf{x}_0), \forall \mathbf{x}\}$ , is called the subgradient of function  $f(\mathbf{x})$  at  $\mathbf{x}_0$ .

**Theorem 1**<sup>[9,14,15]</sup> Assume a nonconvex function  $f(\mathbf{x})$  defined on  $S \subseteq \mathbb{R}^N$  with  $S$  is a convex set, and  $f$  can be decomposed as  $f = g - h$  with  $g$  and  $h$  are both convex functions. If  $\rho(g), \rho(h) > 0$ , then as  $k \rightarrow \infty$  the problem

$$\begin{cases} \mathbf{x}_{k+1} = \operatorname{argmin}_{\mathbf{x} \in S} \{g(\mathbf{x}) - [h(\mathbf{x}_k) + \boldsymbol{\xi}_k^T(\mathbf{x} - \mathbf{x}_k)]\}, \\ \boldsymbol{\xi}_k \in \partial h(\mathbf{x}_k), \end{cases}$$

is convergent to minimum of problem

$$\mathbf{x}_* = \operatorname{argmin}_{\mathbf{x} \in S} \{f(\mathbf{x})\}.$$

Note that, in order to test whether the optimization Problem (11) meets the requirements of Theorem 1, the weak  $\rho$ -convex coefficients of both  $g$  and  $h$  should be calculated. In fact we have the following lemma to guarantee that Theorem 1 is applicable to Problem (11).

**Lemma 2** The functions  $g(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^N (u_i + v_i)^2$  and  $h(\mathbf{u}, \mathbf{v}) \triangleq \sum_{i=1}^N [e^{-(u_i+v_i)/\alpha} + (u_i + v_i)^2]$  are both  $\rho$ -convex function with  $\rho(g) = 2$  and  $\rho(h) = 2 + \frac{1}{2\alpha^2}$ .

**Proof** Let  $\varphi(t) \triangleq t^2 + e^{-\frac{t}{\alpha}}$ , and then it is easy to achieve  $\varphi''(t) = 2 + \frac{1}{\alpha^2}e^{-\frac{t}{\alpha}}$ , which is the second-order differential of  $\varphi$ . Then, according to Proposition 4.8 from Ref. [11], it can be derived that  $\rho(h) = 2 + \frac{1}{2\alpha^2}$ . With the same method, it can be obtained  $\rho(g) = 2$ . ■

Note that, the decomposition method here differs from that in Ref. [8], in which the authors used  $g(\mathbf{u}, \mathbf{v}) = \mathbf{u} + \mathbf{v}$  to perform decomposition. However, that method does not satisfy the requirements of Theorem 1, since  $\rho(g) = 0$ .

Lemma 2 suggests that the optimization Problem (11) meets all the requirements of the subgradient convergence in Theorem 1, and can thus be solved via DC programming. Therefore, Problem (11) is identical to the following iterative problem as  $k \rightarrow \infty$ ,

$$(\mathbf{u}_{k+1}, \mathbf{v}_{k+1}) = \underset{(\mathbf{u}, \mathbf{v}) \in \mathcal{S}}{\operatorname{argmin}} \{g(\mathbf{u}, \mathbf{v}) - h(\mathbf{u}_k, \mathbf{v}_k) - (\boldsymbol{\xi}_k^T \mathbf{u} + \boldsymbol{\eta}_k^T \mathbf{v})\},$$

$$\boldsymbol{\xi}_k \in \partial h_u(\mathbf{u}_k), \quad \boldsymbol{\eta}_k \in \partial h_v(\mathbf{v}_k),$$

$$\text{s.t. } \begin{cases} h(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^N [e^{-(u_i+v_i)/\alpha} + (u_i + v_i)^2], \\ g(\mathbf{u}, \mathbf{v}) = N + \|\mathbf{u} - \mathbf{v}\|^2, \\ \mathcal{S} \triangleq \{(\mathbf{u}, \mathbf{v}) \in \mathbb{R}_+^N \times \mathbb{R}_+^N \mid \mathbf{y} = \mathbf{A}(\mathbf{u} - \mathbf{v})\} \end{cases} \quad (12)$$

To obtain this objective function, we still have to calculate the subgradients of  $h$ , i.e.,  $\boldsymbol{\xi}_k$  and  $\boldsymbol{\eta}_k$ , which is easy to obtain according to Theorem 2, as follows.

**Theorem 2** (Theorem from Refs. [13,16]) Suppose  $h(\mathbf{s}) : S \rightarrow \mathbb{R}$  is a convex function with  $S \triangleq \operatorname{dom}(h) \subseteq \mathbb{R}^N$ , if  $h$  is differentiable at  $\mathbf{s}_0$ , then the subgradient of function  $h$  is identical to its derivative, i.e.,

$$\partial h(\mathbf{s}_0) = \{h'(\mathbf{s}_0)\} = \left. \frac{dh}{d\mathbf{s}} \right|_{\mathbf{s}=\mathbf{s}_0} \quad (13)$$

## 4.2 DC programming algorithm for sparse recovery

According to the iterative convergence Problem (12) and Theorem 2, let

$$\begin{cases} \mathbf{u}_k = [u_k^{(1)}, u_k^{(2)}, \dots, u_k^{(N)}]^T, \\ \mathbf{v}_k = [v_k^{(1)}, v_k^{(2)}, \dots, v_k^{(N)}]^T, \\ \boldsymbol{\xi}_k = [\xi_k^{(1)}, \xi_k^{(2)}, \dots, \xi_k^{(N)}]^T, \\ \boldsymbol{\eta}_k = [\eta_k^{(1)}, \eta_k^{(2)}, \dots, \eta_k^{(N)}]^T. \end{cases}$$

Then, we can obtain

$$\xi_k^{(i)} = \eta_k^{(i)} = 2(u_k^{(i)} + v_k^{(i)}) - \frac{1}{\alpha} e^{-\frac{1}{\alpha}(u_k^{(i)} + v_k^{(i)})}, \quad \forall 1 \leq i \leq N \quad (14)$$

Let  $\varphi = h$ ,  $\phi = g - N$ , since the optimization Problem (12) is invariant by reducing the elements independent from variable  $(\mathbf{u}, \mathbf{v})$ , then the optimization Problem (11) can be converted to

$$\begin{aligned} (\mathbf{u}_{k+1}, \mathbf{v}_{k+1}) &= \underset{\mathbf{x} \in \mathcal{S}}{\operatorname{argmin}} \{\phi(\mathbf{u}, \mathbf{v}) - (\boldsymbol{\xi}_k^T \mathbf{u} + \boldsymbol{\eta}_k^T \mathbf{v})\}, \\ \boldsymbol{\xi}_k = \boldsymbol{\eta}_k &= 2(u_k^{(i)} + v_k^{(i)}) - \frac{1}{\alpha} e^{-\frac{1}{\alpha}(u_k^{(i)} + v_k^{(i)})}, \\ \text{s.t. } \begin{cases} \varphi(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^N [e^{-(u_i+v_i)/\alpha} + (u_i + v_i)^2], \\ \phi(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} + \mathbf{v}\|^2, \\ \mathcal{S} \triangleq \{(\mathbf{u}, \mathbf{v}) \in \mathbb{R}_+^N \times \mathbb{R}_+^N \mid \mathbf{y} = \mathbf{A}(\mathbf{u} - \mathbf{v})\} \end{cases} \end{aligned} \quad (15)$$

Therefore, the EMA sparse recovery Problem (8) can be solved. Then the sparse recovery algorithm via DC Programming can be summarized in Algorithm 1, which is named as EMA-DC method. In this algorithm,  $k$  refers to the iteration index,  $\varepsilon > 0$  refers to the error tolerance factor, and  $\alpha$  is a small positive number, as is defined in Eq. (5) above.

Note that: (1) The optimization step (line 3) in the algorithm is an iterative reweighted  $\ell_2 - \ell_1$  convex programming problem that can be efficiently solved by the CVX package<sup>[17]</sup>. Based on this algorithm, it can be obtained that  $\hat{\mathbf{x}} \triangleq \hat{\mathbf{u}} - \hat{\mathbf{v}}$  is the solution of optimization Problem (8).

(2) Although Algorithm 1 is derived from noiseless observation Formula (1), it is also applicable to noisy models:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}, \quad \text{s.t. } \|\mathbf{n}\|^2 \leq \sigma^2 \quad (16)$$

---

### Algorithm 1 DC algorithm for EMA model

---

**Require:** initial value  $\mathbf{u}_0 = \mathbf{v}_0 = \mathbf{0} \in \mathbb{R}^N$ ,  $k = 0$  and  $\varepsilon > 0$

1: **repeat**

2:  $\mathbf{w}_k^{(i)} = 2(u_k^{(i)} + v_k^{(i)}) - \frac{1}{\alpha} e^{-\frac{1}{\alpha}(u_k^{(i)} + v_k^{(i)})}$ ,  $1 \leq i \leq N$

3:  $(\mathbf{u}_{k+1}, \mathbf{v}_{k+1}) = \underset{(\mathbf{u}, \mathbf{v}) \in \mathcal{S}}{\operatorname{argmin}} [\|\mathbf{u} + \mathbf{v}\|^2 - \mathbf{w}_k^T (\mathbf{u} + \mathbf{v})]$

where  $\mathcal{S}$  is defined as in Formula (15)

4:  $k \leftarrow k + 1$

5: **until**  $\|\mathbf{u}_{k+1} - \mathbf{u}_k - \mathbf{v}_{k+1} + \mathbf{v}_k\| < \varepsilon$

**Ensure:**  $\hat{\mathbf{u}} = \mathbf{u}_k$ ,  $\hat{\mathbf{v}} = \mathbf{v}_k$

---

where  $\mathbf{n}$  is the observed noise, with its energy constrained by  $\sigma^2$ . In this case, one simply changes the constraint domain of Algorithm 1 as follows:

$$\mathcal{S} \triangleq \{(\mathbf{u}, \mathbf{v}) \in \mathbb{R}_+^N \times \mathbb{R}_+^N \mid \|\mathbf{y} - A(\mathbf{u} - \mathbf{v})\|^2 \leq \sigma^2\} \quad (17)$$

Since the Constraint Domain (17) is a convex set of  $(\mathbf{u}, \mathbf{v})$ , then Algorithm 1 is applicable to Formula (16).

## 5 Experimental Results

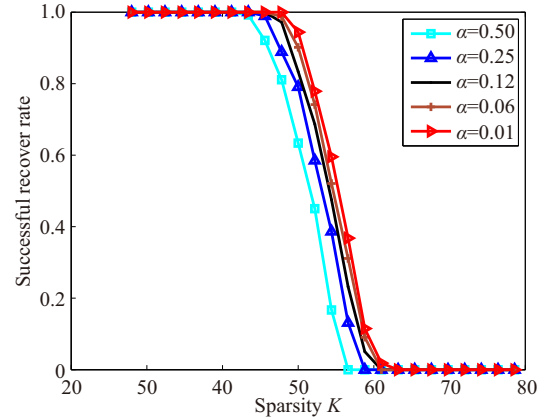
In this section, plenty of simulations are conducted to test the efficiency of the proposed method, and the performance of the proposed method are examined. For all these simulations, the sparsity of signal  $\mathbf{x}$  is defined as  $K$ , which counts the number of all nonzero components of  $\mathbf{x}$ , i.e.,  $K \triangleq \|\mathbf{x}\|_0 = \sum_{i=1}^N \text{sign}(|x_i|)$ , these nonzero components satisfy the zero-mean Gaussian distribution and are randomly chosen among all possible locations. The components of  $A$  are independent identically distributed Gaussian variables with a zero mean and variance  $1/M$ . These simulation results are obtained from the average of 1000 independent trails.

### 5.1 Noiseless case

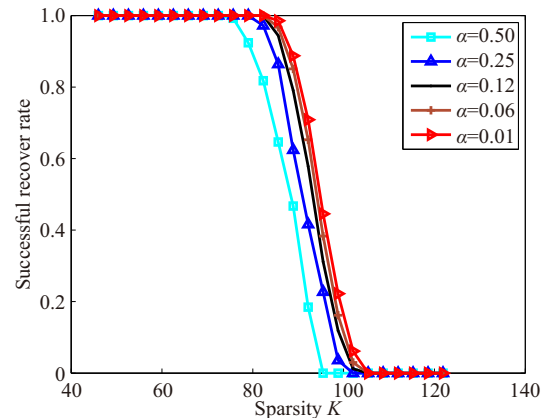
In this case, where  $\sigma^2 = 0$ , the recovery is considered to be exact if the energy of the difference between the reconstructed signal  $\hat{\mathbf{x}}$  and the original signal  $\mathbf{x}$  is smaller than  $10^{-4}$ , that is  $\|\hat{\mathbf{x}} - \mathbf{x}\|^2 < 10^{-4}$ .

In the first experiment, in order to test the relationship between parameter  $\alpha$ , which affects the value of exponential metric function  $F_\alpha(\mathbf{x})$  in Eq. (11), and the sparse recovery performance of our proposed EMA-DC method, different values of  $\alpha$  are chosen to conduct the simulations. The dimension of sensing matrix  $A$  are chosen with typical value in two cases, i.e.,  $M = 128, N = 512$  and  $M = 200, N = 1000$ . It can be seen from Fig. 2 that the smaller the value of  $\alpha$  is, the greater the sparsity of the signal that can be exactly recovered by our proposed EMA-DC method is. This result agrees with the analysis we made in Eq. (7). Meanwhile, it can be seen if  $\alpha < 0.12$ , the recovery performance with respect to different  $\alpha$  values has few discrepancies.

In the second experiment, the recovery results of our proposed method are compared with other sparse recovery methods, including  $\ell_{1/2}$ -quasinorm recovery<sup>[6]</sup>,  $\ell_1$ -norm recovery, Orthogonal Matched Pursuit (OMP)<sup>[18]</sup> method, and Iterative Reweighted Least Squares (IRLS)<sup>[7]</sup> method. For this purpose, the parameter  $\alpha$  is chosen to be  $\alpha = 0.1$  which is



(a)  $M=128$  and  $N=512$



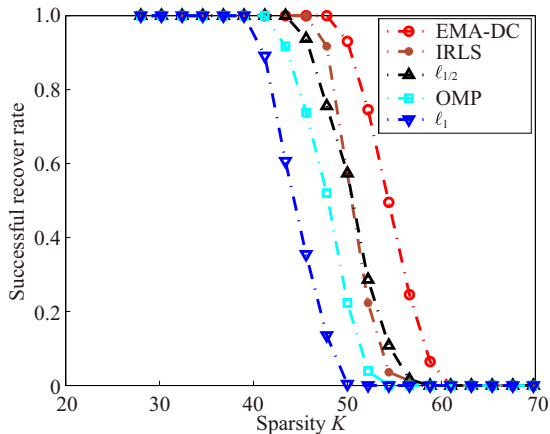
(b)  $M=200$  and  $N=1000$

**Fig. 2 The successful recovery rate of different  $\alpha$  versus sparsity  $K$ .**

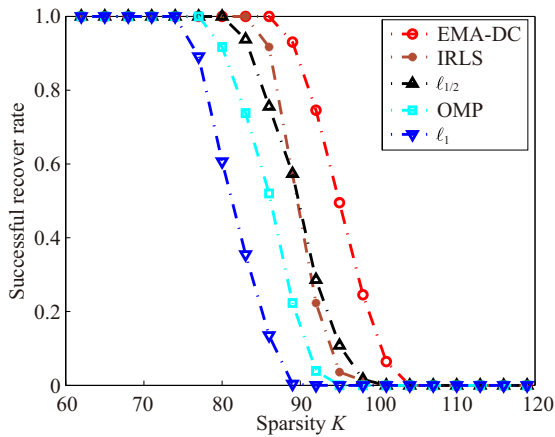
a typical value based on the results of the first experiment. The dimension of sensing matrix  $A$  is chosen corresponding to those in the first experiment. As the simulation results shown in Fig. 3, we can see that the proposed EMA-DC method outperforms the other methods with respect to sparsity tolerance. That is, the proposed method is able to exactly recover the original sparse signal with more nonzero components in  $\mathbf{x}$  than the other methods can.

Although it is difficult to give an elaborate proof of why our proposed EMA-DC method has better sparsity tolerance than other methods, it can be partly explained by the action of the objective functions of these methods. The objective function of the EMA-DC method consists negative exponential metric function  $f_\alpha(\cdot)$ , which generates less error in approximating the  $\ell_0$ -quasinorm than the other methods, which can be seen from Fig. 1.

The time consumed by the EMA-DC method in solving the sparse recovery problem as compared



(a)  $M=128, N=512$ , and  $\alpha=0.1$



(b)  $M=200, N=1000$ , and  $\alpha=0.1$

**Fig. 3** The successful recovery rate of different methods versus sparsity  $K$ .

with  $\ell_1, \ell_{1/2}$ , IRLS, and OMP methods, is shown in Tables 1 and 2, which correspond to Figs. 2a and 2b, respectively. Each of these methods is evaluated 100 times on a PC (Intel Core i5-2500, 3.3 GHz). It can be seen from Tables 1 and 2, the change in the sparsity of signal  $\mathbf{x}$  has little effect on the time consumption of any of these methods except for the OMP method, and if we increase the dimensions of signal  $\mathbf{x}$ , the time consumed by all the methods is greatly increased. The time consumption of our proposed EMA-DC method is more than other methods, i.e., the speed of our proposed EMA-DC method is slower than other methods. This is because the EMA-DC algorithm as proposed in Section 4.2, contains an inner optimization subproblem in each loop that are solved via CVX package<sup>[17]</sup>, then the total computation time is approximately the time consumed by  $\ell_1$  method multiplied by the number of loops in EMA-DC algorithm. In future work, we will work to improve this algorithm to reduce the computational cost

**Table 1** Time consumption of our proposed and other methods when  $M=128, N=512$ . (s)

Sparsity $K$	EMA-DC	IRLS	$\ell_{1/2}$	OMP	$\ell_1$
5	3.315	2.000	1.712	0.002	0.565
10	3.328	1.973	1.720	0.003	0.587
15	3.310	1.995	1.743	0.004	0.541
20	3.355	2.015	1.702	0.006	0.584
25	3.319	1.960	1.691	0.008	0.508
30	3.355	2.030	1.696	0.010	0.540
35	3.338	2.009	1.733	0.012	0.557
40	3.326	1.985	1.741	0.014	0.528
45	3.327	2.000	1.711	0.016	0.530
50	3.315	2.011	1.719	0.019	0.540

**Table 2** Time consumption of our proposed and other methods when  $M=200, N=1000$ . (s)

Sparsity $K$	EMA-DC	IRLS	$\ell_{1/2}$	OMP	$\ell_1$
30	20.228	12.125	10.312	0.020	3.345
35	19.878	11.901	10.335	0.022	3.424
40	20.016	11.986	10.034	0.026	3.259
45	19.942	11.923	10.226	0.030	3.158
50	19.973	11.973	10.256	0.034	3.117
55	19.894	11.899	10.327	0.045	3.449
60	19.987	11.966	10.416	0.047	3.403
65	19.841	11.939	10.321	0.051	3.296
70	20.084	12.066	10.226	0.055	3.343
75	20.097	12.003	10.537	0.064	3.343

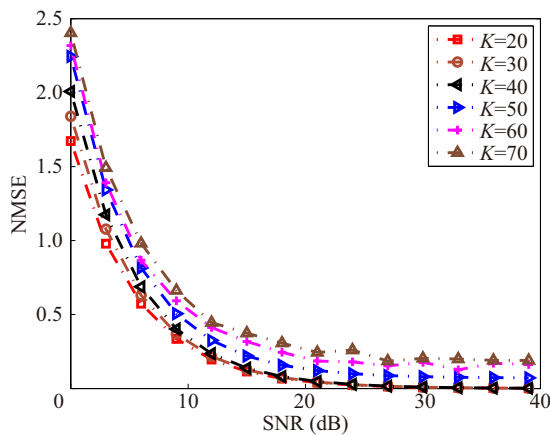
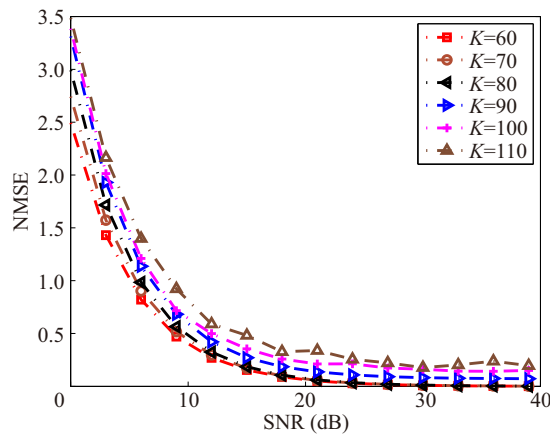
of our proposed EMA-DC method.

### 5.2 Noisy case

In this case, in order to demonstrate the proposed EMA-DC method's ability to recover a sparse signal buried in weak noise, the noise vector  $\mathbf{n}$  is chosen from a zero-mean identically independent Gaussian distribution, i.e.,  $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$ . The error between reconstructed signal  $\hat{\mathbf{x}}$  and the original sparse signal  $\mathbf{x}$  is evaluated by the Normalized Mean Square Error (NMSE), which is defined as  $\text{NMSE} \triangleq E\left(\frac{\|\hat{\mathbf{x}} - \mathbf{x}\|^2}{\|\mathbf{x}\|^2}\right)$ . And the Signal-to-

Noise-Ratio (SNR) is defined as  $\text{SNR} \triangleq \frac{1}{\sigma^2} \|\mathbf{A}\mathbf{x}\|^2$ .

In the third experiment, the dimensions of sensing matrix  $\mathbf{A}$  are chosen with typical value in the two cases  $M = 128, N = 512$  and  $M = 200, N = 1000$ . And  $\alpha$  is chosen as typical value  $\alpha = 0.1$ . Then the NMSE property of different values of  $K$  versus SNR is shown in Fig. 4. It can be seen from Fig.4a that with an increase in the SNR level, the NMSE of recovery results gradually decreases to 0 when the sparsity  $K \leq 30$ . Whereas if  $K \geq 40$ , there is a fixed error that cannot be diminished even if the SNR increases to a very

(a)  $M=128$ ,  $N=512$ , and  $\alpha=0.1$ (b)  $M=200$ ,  $N=1000$ , and  $\alpha=0.1$ **Fig. 4** The NMME property of different  $K$  versus SNR.

high level. This is because the sparsity tolerance of the proposed EMA-DC method is less than 40, as can be seen in Fig. 2a. Similar results can be seen from Fig. 4b. Thus our proposed EMA-DC method can recover a sparse signal buried in noise if the sparse level is at a sufficient level.

## 6 Conclusion

In this paper a new nonconvex exponential metric function is proposed to approximate the original  $\ell_0$ -quasinorm problem, and solving the minimization of exponential metric function can also solve the  $\ell_0$ -quasinorm minimization problem. The nonconvex exponential metric approximation optimization problem is solved via difference convex programming. Numerical simulation results show that the exponential metric function approximation yields better sparsity recovery performance than other methods, and our proposed EMA-DC method can efficiently recover a sparse signal buried in noise. Future work will focus

on improving Algorithm 1 to reduce the computational cost of our proposed EMA-DC method.

## Acknowledgment

This work supported in part by the National Natural Science Foundation of China (Nos. 61171120 and 61501504), the Key National Ministry Foundation of China (No. 9140A07020212JW0101), and the Foundation of Tsinghua University (No. 20141081772).

## References

- [1] D. Donoho, Compressed sensing, *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [2] E. Candes, J. Romberg, and T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, 2006.
- [3] J. A. Tropp, Recovery of short, complex linear combinations via minimization, *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1568–1570, 2005.
- [4] E. Candes, The restricted isometry property and its implications for compressed sensing, *Comptes Rendus Mathematique*, vol. 346, no. 9, pp. 589–592, 2008.
- [5] S. Foucart and M. Lai, Sparsest solutions of underdetermined linear systems via  $\ell_p$ -minimization for  $0 < p \leq 1$ , *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 395–407, 2009.
- [6] Z. Xu, X. Chang, F. Xu, and H. Zhang,  $L_{1/2}$  regularization: A thresholding representation theory and a fast solver, *IEEE Trans. Neural Netw.*, Vol. 23, No. 7, 2012.
- [7] R. Chartrand and W. Yin, Iteratively reweighted algorithms for compressive sensing, in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2008, pp. 3869–3872.
- [8] G. Gasso, A. Rakotomamonjy, and S. Canu, Recovering sparse signals with a certain family of nonconvex penalties and DC programming, *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4686–4698, 2009.
- [9] P. D. Tao and L. T. H. An, Dc optimization algorithms for solving the trust region subproblem, *SIAM J. Optimiz.*, vol. 8, no. 2, pp. 476–505, 1998.
- [10] R. Horst and N. V. Thoai, DC programming: Overview, *Journal of Optimization Theory and Application*, no. 1, pp. 1–43, 1999.
- [11] J. Vial, Strong and weak convexity of sets and functions, *Math. Oper. Res.*, vol. 8, no. 2, pp. 231–259, 1983.
- [12] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge Univ. Press, 2004.
- [13] F. Clarke, Generalized gradients and applications, *Trans. Amer. Math. Soc.*, vol. 202, pp. 247–262, 1975.
- [14] P. D. Tao and L. T. Hoai, DC optimization algorithms for solving the trust region subproblem, *SIAM J. Optim.*, vol. 8, pp. 476–505, 1998.
- [15] H. A. Lethi and P. D. Tao, The DC (Difference of Convex Functions) programming and DCA revisited with DC models of real world nonconvex optimization problems, *Annals of Operations Research*, vol. 133, pp. 23–46, 2005.



- [16] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ, USA: Princeton Univ. Press, 1970.
- [17] M. Grant and S. Boyd, CVX: Matlab software for disciplined convex programming version 2.0 beta,

<http://cvxr.com/cvx>, 2012

- [18] J. Tropp and A. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit, *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.



**Jian Pan** received the BS degree from Beijing Institute of Technology, China, in 2008, and MS degree from Tsinghua University, China, in 2012. He is currently a PhD candidate in Department of Electrical Engineering, Tsinghua University, China. His research interests include compressive sensing application

and radar signal processing.



**Wei Zhu** received the BEng degree from Tsinghua University, China, in 2011. He is currently a PhD candidate at Tsinghua University, China. His current research interests mainly include adaptive radar signal processing and waveform design for MIMO radar.



**Jun Tang** received the PhD degree in electrical engineering from Tsinghua University, China, in 2000. He is currently an professor in the Department of Electronic Engineering, Tsinghua University. His research interest includes array signal processing, information theory and MIMO radar, and compressive sensing

in radar application.