A NON-CONVEX APPROACH TO NON-NEGATIVE SUPER-RESOLUTION: THEORY AND ALGORITHM

Heng Qiao and Piya Pal

Dept. of Electrical and Computer Engineering,
University of California, San Diego, USA
E-mail: h1qiao@eng.ucsd.edu, pipal@eng.ucsd.edu

ABSTRACT

This paper considers the problem of super-resolution reconstruction by casting it as an optimization problem with positive constraints and non-convex objective function. Enforcing the solution to be simultaneously sparse and non-negative naturally leads to a non-convex $l_{1/2}$ quasinorm minimization problem. A reweighted l_1 norm minimization algorithm is proposed to solve this problem, which is tailored for $l_{1/2}$ quasinorm minimization using the idea of Majorization-Minimization. Although the problem is non-convex and non-smooth, and the measurement matrix does not satisfy restricted isometry conditions, we are able to obtain deterministic stable reconstruction guarantees in presence of bounded noise by using the structure of the measurement matrix and non-negativity of the signal. Numerical results demonstrate that $l_{1/2}$ minimization promotes sparser solution and outperforms l_1 minimization.

***Index Terms*—** Positive Super-Resolution, $l_{1/2}$ Minimization, Non-Convex Optimization, Reweighted l_1 Norm Minimization.

1. INTRODUCTION

The problem of super-resolution is fundamental across imaging applications such as astronomy [1], medical imaging [2], microscopy [3] and radar [4]. In these systems, the resolution of the captured image is always limited by the physical measurement process which necessitates the use of sophisticated signal processing techniques to retrieve finer details that are apparently lost. The problem of super-resolution with noisy measurements was analyzed in the pioneering work by Donoho [7] and further developed in recent works [6, 8] where total-variation (TV) and l_1 norm based convex algorithms were used for promoting sparse structure in super-resolution reconstruction. The analysis technique of [6, 8] involves an explicit construction of a certain dual polynomial (based on the Fejér kernel), whose properties can be exploited to analyze the performance of convex super-resolution algorithms for noisy line spectrum estimation [9] and low-rank Toeplitz covariance estimation [10].

More recently, the role of positive constraints in super resolution was analyzed in [11] by imposing a new notion of Rayleigh regularity on the underlying signal. Using the same dual polynomial as [6, 8], the authors in [11] established stability guarantees for a simple l_1 norm based denoising problem with non-negative constraint. In another recent work [5], the author established robust recovery guarantees of positive streams of spikes by imposing strong structural constraints on the admissible blurring kernel. It should be noted that existing analysis of noisy super-resolution focus on solving *convex* problems.

In this paper, we show that positive constraints on the unknown target signal can be exploited in a suitable way alongside its sparsity, leading to a *non-convex* super-resolution problem which minimizes the $l_{1/2}$ quasinorm of the signal. Such $l_q(0 < q \leqslant 1)$ norm based non-convex constrained optimization problems and corresponding algorithms to (approximately) solve them, have been studied in recent literature [13, 14, 19, 15, 21]. However, in order to establish stable reconstruction guarantees, these techniques either rely on exploiting certain properties of the measurement matrix such as the Restricted Isometry Property (RIP) [16]) [19, 15, 23, 24, 25, 26, 31, 32], or Kruskal rank [14]. However, the measurement matrix arising in super-resolution imaging is a *deterministic rank-deficient* matrix composed of DFT matrices for which RIP cannot be established. Besides, $l_{1/2}$ quasinorm is not differentiable and recent advances in non-convex gradient descent based algorithms are inapplicable [20, 23, 24].

Our Contributions. In this work, we propose a non-convex $l_{1/2}$ quasinorm minimization problem for non-negative super-resolution reconstruction and provide theoretical guarantees for stable reconstruction. Inspired by Majorization Minimization techniques [21], we propose an iterative reweighted l_1 minimization algorithm to approximate the $l_{1/2}$ quasinorm, and analyze its convergence and the reconstruction error. Owing to the special deterministic low-pass structure of our measurement matrix, we cannot use existing RIP-based analysis of $l_p(0 < p \leqslant 1)$ minimization problems [19]. Instead, we borrow tools from recent analysis of *convex* super-resolution problems [11] and show that they can be used to derive upper bounds on the reconstruction error (in terms of the so-called super-resolution factor) for the non-convex problem as well.

2. PROBLEM FORMULATION

The goal of discrete positive super-resolution [11, 5, 12] is to reconstruct a signal (or image) $\mathbf{x}^\star \in \mathbb{R}^N$ from measurement $\mathbf{y} \in \mathbb{C}^N$ of the form [11]

$$\mathbf{y} = \mathbf{Q}\mathbf{x}^\star + \mathbf{w} \qquad \mathbf{x}^\star \geqslant \mathbf{0} \tag{1}$$

where $\mathbf{x}^\star$ is a sparse vector with non-negative entries, $\mathbf{w}$ is the measurement noise and $\mathbf{Q} \in \mathbb{C}^{N\times N}$ is the measurement matrix. Here, $\mathbf{Q}$ represents a low-pass filter such that $\mathbf{y}$ only retains the low-frequency components of $\mathbf{x}^\star$, and the high-frequency components are lost. This imparts the following special structure to $\mathbf{Q}$ [11],

$$\mathbf{Q} = \mathbf{F}_N^H \mathbf{\Lambda}_n \mathbf{F}_N \quad (2)$$

where $\mathbf{F}_N \in \mathbb{C}^N$ is given by $[\mathbf{F}_N]_{k,l} = \frac{1}{\sqrt{N}} e^{-j2\pi kl/N}$, $-N/2+1 \leqslant k \leqslant N/2$, $0 \leqslant l \leqslant N-1$ and $\mathbf{\Lambda}_n = \text{diag}([\lambda_{-N/2+1}, \cdots, \lambda_{N/2}])$ with

$$\lambda_k = \begin{cases} 1, & k = -\frac{n-1}{2}, \cdots, \frac{n-1}{2} \\ 0, & \text{otherwise} \end{cases}$$

[1]. Hence, $\mathbf{Q}$ only collects the n low-frequency DFT coefficients of $\mathbf{x}^\star$. The goal of super-resolution is to accomplish the difficult task of recovering the *lost high frequency components* of $\mathbf{x}^\star$ by utilizing its sparsity.

In recent efforts to solve the positive superresolution problem *with provable guarantees*, the authors in [11] proposed the following convex optimization problem to estimate sparse non-negative $\mathbf{x}^\star$

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{Q}\mathbf{x}\|_1 \quad s.t. \quad \mathbf{x} \geqslant \mathbf{0} \quad (P_{\text{den}})$$

Inspite of its simple formulation, (P_{den}) is quite effective in finding $\mathbf{x}^\star$ with provable guarantees. In fact, it is shown that if $\mathbf{x}^\#$ is an optimal solution to (P_{den}), then the l_1 norm of the estimation error $\|\mathbf{x}^\# - \mathbf{x}^\star\|_1$ gets amplified by a factor of $(\frac{N}{n-1})^2$ where $\frac{N}{n-1}$ is the so-called super-resolution factor (SRF).

Notice that the formulation (P_{den}) does not explicitly enforce any sparsity penalty on $\mathbf{x}$, and only uses the prior that it is non-negative. If we assume that $\|\mathbf{w}\|_1 \leqslant \delta_1$, we can further promote sparsity by using the l_1 norm of $\mathbf{x}$ as a convex surrogate for its sparsity [16]. This will be equivalent to adding a non-negative constraint to the convex super-resolution problem proposed in [6, 8]:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad s.t. \quad \|\mathbf{y} - \mathbf{Q}\mathbf{x}\|_p \leqslant \delta_p, \mathbf{x} \geqslant \mathbf{0} \quad (P_1)$$

where p is usually chosen as $p = 1, 2$. Although (P_1) is reminiscent of standard l_1 minimization problem from compressed sensing, conventional analysis tools such as Restricted Isometry Property (RIP) [16] or neighborly-polytope conditions [17, 18] are inapplicable in this case. This is because $\mathbf{Q}$ is a *deterministic* rank-deficient matrix composed of DFT matrices, for which neither RIP nor neighborly-polytope properties can be readily established. The problem (P_1) without the positivity constraint and for $p = 1$ was analyzed in [8, 6] using a different analysis technique that constructs a certain dual certificate in the form of a trigonometric polynomial, and obtained similar error bounds as [11].

3. NON-CONVEX POSITIVE SUPER-RESOLUTION VIA $L_{1/2}$ QUASINORM MINIMIZATION

We will now show how utilizing the positivity of $\mathbf{x}^\star$ actually leads to a non-convex quasinorm minimization problem, which can promote higher sparsity and exhibit better performance than l_1 minimization.

[1]For ease of presentation, we assume that the ambient dimension N is even and n is odd [11]

3.1. Motivation for using $l_{1/2}$ quasinorm in positive super-resolution

As a simple fact, any non-negative vector $\mathbf{x}$ can be represented as $\mathbf{x} = \mathbf{h} \circ \mathbf{h}$, where $\circ$ represents the Hadamard product. Thus, the convex l_1 norm minimization problem (P_1) can be equivalently rewritten in terms of $\mathbf{h}$ as

$$\begin{aligned} &\min_{\mathbf{h}} \|\mathbf{h}\|_2^2 \qquad (\tilde{P}_1) \\ &s.t. \quad \|\mathbf{y} - \mathbf{Q}(\mathbf{h} \circ \mathbf{h})\|_p \leqslant \delta_p, \mathbf{h} \geqslant \mathbf{0} \end{aligned}$$

Without loss of generality, we can assume $\mathbf{h}$ is also non-negative. The formulation $(\tilde{P}_1)$ has convex objective and non-convex constraints. Clearly, $(\tilde{P}_1)$ is equivalent to the convex problem (P_1) due to a one-to-one mapping between $\mathbf{h}$ and $\mathbf{x}$, and the optimal $\mathbf{h}$ has the same support as the optimal $\mathbf{x}$. As evident from $(\tilde{P}_1)$, minimizing l_1 norm of $\mathbf{x}$ is equivalent to minimizing the l_2 norm of $\mathbf{h}$. A natural question to ask is: what happens if we enforce sparsity of $\mathbf{h}$ by replacing its l_2 norm with l_1 norm in the objective function? In other words, we consider the following problem

$$\min_{\mathbf{h}} \|\mathbf{h}\|_1 \quad s.t. \quad \|\mathbf{y} - \mathbf{Q}(\mathbf{h} \circ \mathbf{h})\|_p \leqslant \delta_p, \mathbf{h} \geqslant \mathbf{0} \quad (\tilde{P}_2)$$

Using $\mathbf{x} = \mathbf{h} \circ \mathbf{h}$, $(\tilde{P}_2)$ can be rewritten as

$$\min_{\mathbf{x}} \|\mathbf{x}\|_{\frac{1}{2}} \quad s.t. \quad \|\mathbf{y} - \mathbf{Q}\mathbf{x}\|_p \leqslant \delta_p, \mathbf{x} \geqslant \mathbf{0} \quad (P_2)$$

where we use the fact that $\|\mathbf{x}\|_{\frac{1}{2}}^{0.5} = \|\mathbf{h}\|_1$. The problem (P_2) minimizes the non-convex $l_{1/2}$ quasinorm of $\mathbf{x}$ over a convex feasible set. It is well known that minimizng the $l_{1/2}$ quasinorm favors even sparser solutions over minimizing l_1 norm [19, 28, 29, 30]. While $l_{1/2}$ quasinorm minimization has been explored and analyzed as a better alternative to l_1 norm for promoting sparsity, the corresponding theoretical guarantees (which are based on RIP) [19, 26, 31, 32] do not apply to $\mathbf{Q}$ which represents a low-pass filter in super-resolution imaging. We bridge this gap by first proposing an iterative reweighted l_1 norm minimization algorithm (for approximating the $l_{1/2}$ quasinorm) and developing theoretical guarantees under which this algorithm can provide stable solution in presence of noise.

3.2. Iterative Algorithm to approximate $l_{1/2}$ quasinorm minimization

Since (P_2) is non-convex and has non-differentiable objective function, recent advances in non-convex gradient descent based algorithms [20, 23, 24, 25] are not applicable. Inspired by [19], we propose an iterative reweighted l_1 norm minimization algorithm to solve (P_2), by explicitly enforcing positivity of the desired signal.

Algorithm 1: Non-negative Reweighted l_1 Minimization

Input: Noisy measurements $\mathbf{y}$, parameters δ_p and $\varepsilon > 0$
Output: An estimate $\mathbf{x}^{\#}$ of $\mathbf{x}^{\star}$.

1. **Initialization:** An initial feasible guess $\mathbf{x}_0$ such that $\|\mathbf{y} - \mathbf{Q}\mathbf{x}_0\|_p \leqslant \delta_p, \mathbf{x}_0 \geqslant \mathbf{0}$, and a sequence of non-increasing positive numbers $\{\epsilon_n\}$ such that $\lim_{n\to\infty} \epsilon_n = 0$.

2. **Iteration:** Given $\mathbf{x}_n$, obtain $\mathbf{x}_{n+1}$ as

$$\mathbf{x}_{n+1} = \arg\min_{\mathbf{z}\in\mathbb{R}^N} \sum_{i=0}^{N-1} \frac{z_i}{(x_{n,i} + \epsilon_n)^{\frac{1}{2}}} \qquad (P_3)$$
$$s.t. \quad \|\mathbf{y} - \mathbf{Q}\mathbf{z}\|_p \leqslant \delta_p, \quad \mathbf{z} \geqslant \mathbf{0}$$

3. **Stopping Criterion:** Stop when $\|\mathbf{x}_n - \mathbf{x}_{n+1}\|_1 \leqslant \varepsilon$. Return $\mathbf{x}_{n+1}$ as the estimate of $\mathbf{x}^{\star}$.

The problem (P_3) in Algorithm 1 can be identified as a reweighted l_1 minimization problem, where the weights are given by $(x_{n,i} + \epsilon_n)^{-0.5}, i = 0, \cdots, N-1$. [2] The motivation stems from prior works in Majorization-Minimization (MM) algorithms that iteratively minimizes simple (possibly convex) surrogates for a given objective function [21]. In our case, we want to minimize the non-convex $l_{1/2}$ quasinorm $g(\mathbf{z}) \triangleq \|\mathbf{z} + \epsilon\|_{1/2}^{0.5} = \sum_{i=0}^{N-1} \sqrt{z_i + \epsilon}$, for $\mathbf{z} \geqslant \mathbf{0}$. We instead iteratively minimize the first-order linear approximation of $g(\mathbf{z})$ at $\mathbf{z} = \mathbf{x}_n$, giving rise to the following formulation

$$\min_{\mathbf{z}\geqslant\mathbf{0}} \left\{ g(\mathbf{x}_n) + \sum_{i=0}^{N-1} \frac{1}{2} \frac{z_i - x_{n,i}}{\sqrt{x_{n,i} + \epsilon}} \right\}, \text{ s.t. } \|\mathbf{y} - \mathbf{Q}\mathbf{z}\|_p \leqslant \delta_p, \quad (3)$$

Here, $\sum_{i=0}^{N-1} \frac{z_i}{\sqrt{x_{n,i}+\epsilon}}$ can be identified as the weighted l_1 norm of non-negative $\mathbf{z}$, implying that (3) identical to (P_3).

4. ANALYSIS OF $L_{1/2}$ MINIMIZATION: CONVERGENCE AND ERROR BOUND

A main contribution of our paper is to analyze Algorithm 1 given the special structure of the low-pass filter $\mathbf{Q}$, and develop explicit bounds on the estimation error $\|\mathbf{x}^{\star} - \mathbf{x}^{\#}\|_1$. Although Algorithm 1 does not solve a convex problem, we show that the error bound behaves similar to the convex non-negative superresolution algorithm proposed in [11] and gets amplified by SRF^2.

We begin by defining the set of signals obyeing separation condition[10, 9].

Definition 1. *(Set of Non-Negative Signals Obeying Separation Condition) Given N and n, the set Δ^+_{sep} is given by*

$$\{\mathbf{x} \in \mathbb{C}^N, \mathbf{x} \geqslant \mathbf{0} \mid \rho(\frac{k}{N}, \frac{l}{N}) \geqslant \frac{4}{n-1} \quad \forall k \neq l \in supp(\mathbf{x})\}$$

where $\rho(\cdot,\cdot)$ is a wrap-around distance function [6] such that for $\forall \mu_1, \mu_2 \in [0, 1]$, we have $\rho(\mu_1, \mu_2) \triangleq \min(|\mu_1 - \mu_2|, |\mu_1 + 1 - \mu_2|, |\mu_2 + 1 - \mu_1|)$

The following theorem shows that the sequence of iterates produced by Algorithm 1 has a converging subsequence, and whenever $\mathbf{x}^{\star} \in \Delta^+_{\text{sep}}$, the limit of this convergent subsequence produces a stable estimate of $\mathbf{x}^{\star}$ (and in particular, exactly recovers $\mathbf{x}^{\star}$ in absence of noise).

Theorem 1. *Given any non-increasing positive sequence $\{\epsilon_n\}$ and a feasible initial point $\mathbf{x}_0$, the solution sequence $\{\mathbf{x}_n\}$ of Algorithm 1 has a convergent subsequence which converges to a feasible point $\mathbf{x}^{\#}$ of (P_2). Furthermore, if $\mathbf{x}^{\star} \in \Delta^+_{sep}$, the limit $\mathbf{x}^{\#}$ obeys*

$$\|\mathbf{x}^{\#} - \mathbf{x}^{\star}\|_1 \leqslant C \left(\frac{N}{n-1}\right)^2 \delta_1 \qquad (4)$$

where C is a positive constant.

Proof. Notice that

$$\begin{aligned}
&\sum_{i=0}^{N-1} (x_{n+1,i} + \epsilon_{n+1})^{1/2} \\
&\overset{(a)}{\leqslant} \sum_{i=0}^{N-1} \frac{(x_{n+1,i} + \epsilon_n)^{1/2}}{(x_{n,i} + \epsilon_n)^{1/4}} (x_{n,i} + \epsilon_n)^{1/4} \\
&\overset{(b)}{\leqslant} \left[\sum_{i=0}^{N-1} \frac{(x_{n+1,i} + \epsilon_n)}{(x_{n,i} + \epsilon_n)^{1/2}}\right]^{1/2} \left[\sum_{i=0}^{N-1} (x_{n,i} + \epsilon_n)^{1/2}\right]^{1/2} \\
&\overset{(c)}{\leqslant} \left[\sum_{i=0}^{N-1} \frac{(x_{n,i} + \epsilon_n)}{(x_{n,i} + \epsilon_n)^{1/2}}\right]^{1/2} \left[\sum_{i=0}^{N-1} (x_{n,i} + \epsilon_n)^{1/2}\right]^{1/2} \\
&= \sum_{i=0}^{N-1} (x_{n,i} + \epsilon_n)^{1/2}
\end{aligned}$$

where (a) is due to $\epsilon_{n+1} \leqslant \epsilon_n$, (b) follows from Hölder's inequality and (c) is true because $\mathbf{x}_{n+1}$ is the optimal solution of (P_3). We further use the fact that [19]

$$\|\mathbf{x}_n\|_\infty \leqslant \left[\sum_{i=0}^{N-1} (x_{n,i} + \epsilon_n)^{1/2}\right]^2 \leqslant \left[\sum_{i=0}^{N-1} (x_{0,i} + \epsilon_0)^{1/2}\right]^2$$

This shows that the sequence $\{\mathbf{x}_n\}$ is bounded, and thus it has a converging subsequence. Additionally, the feasible set of (P_3) is the intersection of non-negative orthant and closed l_p ball (where $p = 1, 2$) and hence any cluster point of $\{\mathbf{x}_n\}$ will be feasible [22].

To prove the second part, we use the following fact about $\mathbf{Q}$ from [11]. Let $\mathbf{v} = \mathbf{x}^{\#} - \mathbf{x}^{\star}$, and $\mathcal{T}_{\mathbf{v}} = \{l | v_l < 0, 0 \leqslant l \leqslant N-1\}$. If $\mathbf{x}^{\star} \in \Delta^+_{sep}$, there exists $\mathbf{q} \in \mathbb{R}^N$ and $c\left(\frac{n-1}{N}\right)^2 \leqslant \eta < 1$ where $c = 0.0036$, such that $\mathbf{Q}\mathbf{q} = \mathbf{q}$ and [11]

$$q_l = -\eta \quad \text{If } l \in \mathcal{T}_{\mathbf{v}}; \quad \eta < q_l < 1 - \eta \quad \text{otherwise}$$

Given the existence of such a $\mathbf{q}$, we have

$$\begin{aligned}
&|\mathbf{q}^T\mathbf{v}| = |(\mathbf{Q}\mathbf{q})^T\mathbf{v}| = |\mathbf{q}^T\mathbf{Q}\mathbf{v}| \leqslant \|\mathbf{q}\|_\infty \|\mathbf{Q}\mathbf{v}\|_1 \\
&\leqslant (1-\eta)\|\mathbf{Q}\mathbf{x}^{\#} - \mathbf{Q}\mathbf{x}^{\star}\|_1 \\
&\leqslant (1-\eta)\left(\|\mathbf{Q}\mathbf{x}^{\#} - \mathbf{y}\|_1 + \|\mathbf{Q}\mathbf{x}^{\star} - \mathbf{y}\|_1\right) \leqslant 2(1-\eta)\delta_1
\end{aligned}$$

[2]The positive parameter ϵ_n is used to avoid zero denominator [21].

On the other hand, we also have

$$|\mathbf{q}^T \mathbf{v}| = |\sum_{l=0}^{N-1} q_l v_l| = \sum_{l=0}^{N-1} q_l v_l \geqslant \eta \|\mathbf{v}\|_1$$

The proof completes by using $\eta = c\left(\frac{n-1}{N}\right)^2$. □

5. NUMERICAL RESULTS

We now conduct numerical experiments to demonstrate that the proposed reweighted iterative algorithm for approximating $l_{1/2}$ quasinorm minimization can produce better estimate of $\mathbf{x}^\star$ both in terms of sparsity and smaller estimation error. We choose $N = 64$, $n = 21$, and the true sparsity is set at $\|\mathbf{x}^\star\|_0 = 6$. The non-zero entries of $\mathbf{x}^\star$ are produced by first generating uniform random variables between 1 and 2 and then normalizing the entries so that $\|\mathbf{x}^\star\|_1 = 1$. Similarly, the measurement noise $\mathbf{w}$ is produced by generating complex standard Gaussian random variables and then normalizing $\mathbf{w}$ such that $\|\mathbf{w}\|_1 = \delta_1$. We will compare the performance of different algorithms by varying δ_1. To implement Algorithm 1, we set the stopping parameter to $\varepsilon = 0.001$ and select $\epsilon_n = \frac{10^{-4}}{n}$.

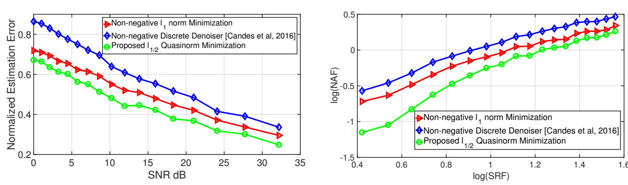

Fig. 1. (Left) Comparative performance of different algorithms as a function of Signal-to-Noise ratio (SNR). The results are averaged over 1200 Monte Carlo runs. (Right) Noise Amplification Factor (NAF) of (P_{den}), (P_1) and proposed $l_{1/2}$ minimization, as a function of Super-Resolution Factor (SRF) $\frac{N}{n-1}$. The results are averaged over 800 runs.

In Fig. 1, we compare the performance of (P_{den}), (P_1), and (P_2) [3] by varying δ_1 which represents the l_1 norm of the noise $\mathbf{w}$. The Signal-to-Noise ratio (SNR) is defined as $20\log(\frac{\|\mathbf{x}^\star\|_1}{\|\mathbf{w}\|_1})$. It can be seen that the proposed algorithm produces the smallest normalized estimation error $\|\mathbf{x}^\# - \mathbf{x}^\star\|_1/\|\mathbf{x}^\star\|_1$ and outperforms both (P_{den}) and (P_1).

In Fig.1, we also plot NAF $\triangleq \frac{\|\mathbf{x}^\# - \mathbf{x}^\star\|_1}{\|\mathbf{w}\|_1}$ which represents the factor by which the estimation error is amplified with respect to the input noise, as a result of super-resolution reconstruction. As before, the proposed $l_{1/2}$ based algorithm shows minimum noise amplification. Moreover, the noise amplification of the proposed algorithm follows similar trend as [11], as predicted by the noise bound (4).

We finally demonstrate the reconstruction quality of the Algorithm 1 for 2D super-resolution. The proposed algorithm can be readily extended to two dimensions by choosing $\mathbf{Q}$ as a 2D DFT matrix. Fig. 2 shows the performance of Algorithm 1 and (P_1) on synthetic $2D$ data. The ground truth is a sparse $N \times N$ image where $N = 24$. We generate the low-frequency measurements by only retaining the 49 low frequency DFT coefficients. We further normalize $\mathbf{w}$ so that $\|\mathbf{w}\|_1 = 0.1$. It can be clearly seen that Algorithm 1 exactly recovers the true support while (P_1) produces several false peaks. This further corroborates the fact that the proposed $l_{1/2}$ minimization framework indeed favors and identifies sparser solutions.

[3] we choose $p = 1$ for solving (P_{den}) and (P_1)

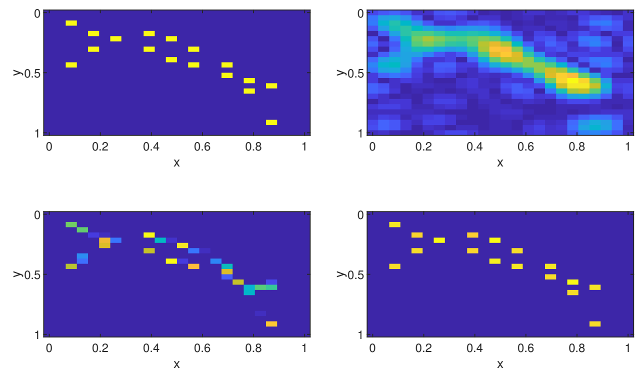

Fig. 2. (Top Left) Ground truth image with positive emitters. (Top Right) Measured image consisting of only low frequency components. (Bottom Left) Estimate produced by solving convex problem (P_1) (Bottom Right) Estimate produced by the proposed iterative $l_{1/2}$ minimization algorithm.

6. CONCLUSION

In this paper, we analyzed the problem of super-resolution where the desired signal is both sparse and non-negative. We proposed a constrained non-convex $l_{1/2}$ quasinorm minimization problem to promote sparsity in the reconstructed signal. Such a formulation naturally stems from exploiting non-negative constraints on the signal. Although $l_{1/2}$ quasinorm is non-convex and non-differentiable and the measurement matrix does not satisfy RIP, the stability of the solution can still be guaranteed by constructing appropriate dual certificates. An iterative reweighted l_1 minimization algorithm is proposed to approximate the $l_{1/2}$ quasinorm and simulations show that it has better performance than l_1 norm minimization, in terms of both accuracy and sparsity of the solution. A complete analysis of the iterative algorithm and derivation of tighter error upper bounds will be left as future work.

7. REFERENCES

[1] K. G. Puschmann and F. Kneer, "On super-resolution in astronomical imaging", *Astronomy and Astrophysics*, vol. 436, no. 1, pp. 373-378, 2005.

[2] H. Greenspan, "Super-resolution in medical imaging", *Comput. J.*, vol. 52, no. 1, pp. 43-63, 2009.

[3] C. W. McCutchen, "Superresolution in microscopy and the Abbe resolution limit", *J. Opt. Soc. Am.*, vol. 57, no. 10, pp. 1190-1192, 1967.

[4] R. Heckel, V. I. Morgenshtern and M. Soltanolkotabi, “Super-resolution radar”, vol. 5, no. 1, pp. 22-75, 2016.

[5] T. Bendory, “Robust Recovery of Positive Stream of Pulses”, *IEEE Transactions on Signal Processing*, vol. 65, no. 8, pp. 2114-2122, April 2017.

[6] E. J. Candès and C. Fernandez-Granda,“Super-resolution from noisy data”, *J. Fourier Anal. Appl.*, vol. 19, no. 6, pp. 1229-1254, 2013.

[7] D. L. Donoho, “Superresolution via sparsity constraints”, *SIAM J. Math. Anal.*, vol. 23, no. 5, pp. 1309-1331, Sep. 1992.

[8] E. J. Candès and C. Fernandez-Granda, “Towards a mathematical theory of super-resolution”, *Commun. Pure Appl. Math.*, vol. 67, no. 6, pp. 906-956, June 2014.

[9] G. Tang, B. N. Bhaskar and B. Recht, “Near minimax line spectral estimation”, *IEEE Trans. Inf. Theory*, vol. 61, no. 1, pp. 499 -512, Jan. 2015.

[10] H. Qiao and P. Pal, “Gridless Line Spectrum Estimation and Low-Rank Toeplitz Matrix Compression Using Structured Samplers: A Regularization-Free Approach”, *IEEE Transactions on Signal Processing*, vol. 65, no. 9, pp. 2211-2226, May 2017.

[11] V. I. Morgenshtern and E. J. Candès, “Super-Resolution of Positive Sources: the Discrete Setup,” *SIAM J Imaging Sciences*, vol. 9, no. 1, pp. 412-444, 2016.

[12] Q. Denoyelle, V. Duval and G. Peyré, “Support Recovery for Sparse Super-Resolution of Positive Measures”, *J. Fourier Anal. Appl.*, vol. 23, no. 5, pp. 1153-1194, 2017.

[13] I. F. Gorodnitsky and B. D. Rao, “Sparse signal reconstruction from limited data using FOCUSS: A reweighted minimum norm algorithm”, *IEEE Transactions on Signal Processing*, vol. 45, no. 3, pp. 600 - 616, March 1997.

[14] Md. M. Hyder and K. Mahata, “An improved smoothed l_0 approximation algorithm for sparse representation” *IEEE Transactions on Signal Processing*, vol. 58, no. 4, pp. 2194 - 2205, April 2010.

[15] T. Blumensath and M. E. Davies, “Normalized iterative hard thresholding: guaranteed stability and performance” *IEEE J. Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 298-309, April 2010.

[16] S. Foucart and H. Rauhut, “A Methematical Introduction to Compressive Sensing”, Birkhäuser/Springer, New York, 2013.

[17] D. L. Donoho and J. Tanner, “Sparse nonnegative solution of underdetermined linear equations by linear programming”, *Proc. Natl. Acad. Sci.*, vol 102, no. 27, pp. 9446-9451, July 2005.

[18] M. Wang, W. Xu and A. Tang, “A Unique 'Nonnegative' Solution to an Underdetermined System: From Vectors to Matrices”, *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 1007-1016, Mar. 2011.

[19] S. Foucart and M. Lai, “Sparsest solutions of underdetermined linear systems via l_p-minimization for $0 < q \leqslant 1$, *Appl. Comput. Harmon. Anal.*, vol. 26, pp. 395-407, 2009.

[20] Y. Nesterov, “Introductory lectures on convex optimization: a basic course”, Springer, 2004.

[21] E. J. Candès, M. B. Wakin and S. P. Boyd, “Enhancing sparsity by reweighted l_1 minimization”, *J. Fourier Anal. Appl.*, vol. 14, pp. 877-905, 2008.

[22] W. Rudin, “Principles of mathematical analysis”, McGraw-Hill Education, 1976.

[23] E. J. Candès, X. Li and M. Soltanolkotabi, “Phase retrieval via Wirtinger flow: theory and algorithms”, *IEEE Trans. Info. Theory*, vol. 61, no. 4, pp. 1985-2007, 2015.

[24] R. Ge, J . D. Lee and T. Ma, “Matrix completion has no spurious local minima”, in Advances in Neural Information Processing Systems (NIPS), 2016.

[25] C. Ma, K. Wang, Y. Chi and Y. Chen,“Implicit Regularization in Nonconvex Statistical Estimation: Gradient Descent Converges Linearly for Phase Retrieval, Matrix Completion, and Blind Deconvolution”, *arXiv:*1711.10467, 2017.

[26] I. Daubechies, R. Devore, M. Fornasier and C. S. Güntürk, “Iteratively reweighted least squares minimization for sparse recovery” *Communications on Pure and Applied Mathematics*, vol. LXIII, 0001-0038, 2010.

[27] M. S. Lobo, M. Fazel and S. Boyd, “Portfolio optimization with linear and fixed transaction costs”, *Ann. Oper. Res.*, vol. 152, no. 1, pp. 341-365, July 2007.

[28] K. Mohan and M. Fazel, “Iterative Reweighted Algorithms for Matrix Rank Minimization”, *Journal of Machine Learning Research*, vol. 13, pp. 3441-3473, Nov. 2012.

[29] Y. Xie, S. Gu, Y. Liu, W. Zuo, W. Zhang and L. Zhang, “Weighted schatten p-norm minimization for image denoising and Background subtraction”, *IEEE Transactions on Signal Processing*, vol. 25, no. 10, pp. 4842-4857, Oct. 2016.

[30] M. E. Davies and R. Gribonval, “Restricted isometry constants where l^p sparse recovery can fail for $0 < p \leqslant 1$”, *IEEE Trans. Info. Theory*, vol. 55, no. 5, pp. 2203-2214, Feb. 2009.

[31] R. Chartrand, “Exact reconstruction of sparse signals via nonconvex minimization”, *IEEE Signal Processing Letters*, vol. 14, no. 10, pp. 707-710, Oct. 2007.

[32] M. Fornasier, H. Rauhut and R. Ward, “Low-rank matrix recovery via iteratively reweighted least squares minimization”, *SIAM J. Optim.*, vol. 21, no. 4, pp. 1614-1640, 2011.