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 $l_q(0<q<1)$ 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] 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 $l_q(0<p<1)$ 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) $x^* \in \mathbb{R}^N$ from measurement $y \in \mathbb{C}^N$ of the form

$$y = Qx^* + w \quad x^* \geq 0$$ (1)

where $x^*$ is a sparse vector with non-negative entries, $w$ is the measurement noise and $Q \in \mathbb{C}^{N \times N}$ is the measurement matrix. Here, $Q$ represents a low-pass filter such that $y$ only
3. NON-CONVEX POSITIVE SUPER-RESOLUTION VIA $l_{1/2}$ QUASINORM MINIMIZATION

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

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

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

$$\min_h \|h\|_2 \quad (\tilde{P}_1)$$
$$s.t. \quad \|y - Q(h \circ h)\|_p \leq \delta_p, h \geq 0$$

Without loss of generality, we can assume $h$ 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 $h$ and $x$, and the optimal $h$ has the same support as the optimal $x$. As evident from $(\tilde{P}_1)$, minimizing $l_1$ norm of $x$ is equivalent to minimizing the $l_2$ norm of $h$. A natural question to ask is: what happens if we enforce sparsity of $h$ by replacing its $l_2$ norm with $l_1$ norm in the objective function? In other words, we consider the following problem

$$\min_h \|h\|_1 \quad (\tilde{P}_2)$$
$$s.t. \quad \|y - Q(h \circ h)\|_p \leq \delta_p, h \geq 0$$

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

$$\min_x \|x\|_{1/2} \quad s.t. \quad \|y - Qx\|_p \leq \delta_p, x \geq 0 \quad (P_2)$$

where we use the fact that $\|x\|_{1/2} = \|h\|_1$. The problem $(P_2)$ minimizes the non-convex $l_{1/2}$ quasinorm of $x$ over a convex feasible set. It is well known that minimizing the $l_{1/2}$ quasinorm favors even sparser solutions over minimizing $l_1$ norm $[19, 23, 29, 30]$. 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 $Q$ 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 $y$, parameters $\delta_p$ and $\varepsilon > 0$

**Output:** An estimate $x^\#$ of $x^*$.

1. **Initialization:** An initial feasible guess $x_0$ such that $|y - Qx_0|_p \leq \delta_p, x_0 \geq 0$, and a sequence of non-increasing positive numbers $\{\epsilon_n\}$ such that $\lim_{n \to \infty} \epsilon_n = 0$.

2. **Iteration:** Given $x_n$, obtain $x_{n+1}$ as

$$x_{n+1} = \arg\min_{z \in \mathbb{R}^N} \sum_{i=0}^{N-1} \frac{z_i}{(x_{n,i} + \epsilon_n)^{1/2}}$$

subject to $|y - Qz|_p \leq \delta_p, \ z \geq 0$.

3. **Stopping Criterion:** Stop when $|x_n - x_{n+1}|_1 \leq \varepsilon$. Return $x_{n+1}$ as the estimate of $x^*$.

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

$$\min_{z \geq 0} \left\{ g(x_n) + \sum_{i=0}^{N-1} \frac{1}{2} \frac{z_i - x_{n,i}}{\sqrt{x_{n,i} + \epsilon}} \right\}, \text{ s.t. } |y - Qz|_p \leq \delta_p. \quad (3)$$

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 $z$, 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 $Q$, and develop explicit bounds on the estimation error $\|x^* - x^\#\|_1$. 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 obeying separation condition [10, 9].

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

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

where $\rho(\cdot, \cdot)$ is a wrap-around distance function [6] such that for $\forall u_1, u_2 \in [0, 1]$, we have $\rho(u_1, u_2) \leq \min(|u_1 - u_2|, (u_1 + 1 - u_2), (u_2 + 1 - u_1))$.

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

**Theorem 1.** Given any non-increasing positive sequence $\{\epsilon_n\}$ and a feasible initial point $x_0$, the solution sequence $(x_n)$ of Algorithm 1 has a convergent subsequence which converges to a feasible point $x^\#$ of $(P_2)$. Furthermore, if $x^* \in \Delta^+_{sep}$, the limit $x^\#$ obeys

$$\|x^\# - x^*\|_1 \leq C \left( \frac{N}{n - 1} \right)^2 \delta_1$$

where $C$ is a positive constant.

**Proof.** Notice that

$$\sum_{i=0}^{N-1} \left( \frac{|x_{n+1,i} + \epsilon_n|}{(x_{n,i} + \epsilon_n)^{1/4}} \right)^{1/2} \leq \sum_{i=0}^{N-1} \left( \frac{|x_{n,i} + \epsilon_n|}{(x_{n,i} + \epsilon_n)^{1/4}} \right)^{1/2}$$

where (a) is due to $\epsilon_{n+1} \leq \epsilon_n$, (b) follows from Hölder’s inequality and (c) is true because $x_{n+1}$ is the optimal solution of $(P_3)$. We further use the fact that $\epsilon_1$.

$$\|x_n\|_\infty \leq \left( \sum_{i=0}^{N-1} \left( \frac{|x_{n,i} + \epsilon_n|}{(x_{n,i} + \epsilon_n)^{1/4}} \right)^{1/2} \right)^{1/2} \leq \left( \sum_{i=0}^{N-1} (x_{0,i} + \epsilon_0)^{1/2} \right)^{1/2}$$

This shows that the sequence $(x_n)$ 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 $(x_n)$ will be feasible [23].

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

Given the existence of such a $q$, we have

$$\|Q^T v\| = \|Qv\| = \|Q^T Qv\| \leq \|Qv\|_1 \leq (1 - \eta) \|Q^T Qv\|_1 \leq (1 - \eta) \|Q^T Qv\|_1 \leq (1 - \eta) \|Qv\|_1$$

$\bullet$
On the other hand, we also have
\[ |q^T v| = \sum_{l=0}^{N-1} q_l v_l = \sum_{l=0}^{N-1} q_l v_l \geq \eta \|v\|_1 \]
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 \( x^* \) both in terms of sparsity and smaller estimation error. We choose \( N = 64, n = 21 \), and the true sparsity is set at \( \|x^*\|_0 = 6 \). The non-zero entries of \( x^* \) are produced by first generating uniform random variables between 1 and 2 and then normalizing the entries so that \( \|x^*\|_1 = 1 \). Similarly, the measurement noise \( w \) is produced by generating complex standard Gaussian random variables and then normalizing \( w \) such that \( \|w\|_1 = \delta_1 \). We will compare the performance of different algorithms by varying \( \delta_1 \). To implement Algorithm 1, we set the stopping parameter to \( \epsilon = 0.001 \) and select \( \epsilon_n = \frac{10^{-4}}{n} \).

In Fig. 1 we compare the performance of \( (P_{\text{den}}) \), \( (P_1) \), and \( (P_2) \) by varying \( \delta_1 \) which represents the \( l_1 \) norm of the noise \( w \). The Signal-to-Noise ratio (SNR) is defined as \( 20 \log \left( \frac{\|x^0\|_1}{\|w\|_1} \right) \). It can be seen that the proposed algorithm produces the smallest normalized estimation error \( \|x^# - x^*\|_1 / \|x^*\|_1 \) and outperforms both \( (P_{\text{den}}) \) and \( (P_1) \).

In Fig. 2 we also plot NAF \( = \frac{\|x^# - x^*\|_1}{\|x^\#\|_1} \), 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 \( (2) \).

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 \( Q \) 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 \( w \) so that \( \|w\|_1 = 0.1 \). 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.

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


