ACHIEVING HIGH RESOLUTION FOR SUPER-RESOLUTION VIA REWEIGHTED ATOMIC NORM MINIMIZATION

Zai Yang and Lihua Xie, Fellow, IEEE

School of Electrical and Electronic Engineering
Nanyang Technological University, 639798, Singapore

ABSTRACT

The super-resolution theory developed recently by Candès and Fernandes-Granda aims to recover fine details in a sparse frequency spectrum from coarse scale information. The theory was then extended to the cases of compressive samples and/or multiple measurement vectors. However, the existing atomic norm (or total variation norm) techniques succeed only if the frequencies are sufficiently separated, prohibiting commonly known high resolution. In this paper, a reweighted atomic-norm minimization (RAM) approach is proposed which iteratively carries out atomic norm minimization (ANM) with a sound reweighting strategy that enhances sparsity and resolution. It is demonstrated analytically and via numerical simulations that the proposed method achieves high resolution with application to DOA estimation.

Index Terms— Continuous compressed sensing, high resolution, reweighted atomic norm minimization, super-resolution.

1. INTRODUCTION

Frequency analysis of signals [1] is a classical problem that has broad applications ranging from communications, radar, array processing to seismology and astronomy. *Grid-based* sparse methods have been vastly studied in the past decade with the development of compressed sensing (CS) which exploit signal sparsity–the number of frequency components K is small–but suffer from basis mismatches due to the need of gridding the frequency interval [2, 3]. The research has recently been advanced owing to the mathematical theory of super-resolution introduced by Candès and Fernandes-Granda [4], which refers to the recovery of fine details in a sparse frequency spectrum from coarse scale time-domain samples. They proposed a *gridless* atomic norm (or total variation norm) technique, which can be cast as semidefinite programming (SDP), and proved that a continuous frequency spectrum can be recovered with infinite precision given a set of N regularly spaced samples. The technical method and theoretical result were extended by Tang *et al.* [5] to the case of partial/compressive samples, showing that only a number of $M = O(K \ln K \ln N)$ random samples are sufficient for the recovery with high probability via atomic norm minimization (ANM). Moreover, Yang and Xie [6, 7] studied the multiple-measurement-vector (MMV) case, which arises naturally in array processing applications, with similar results proven using extended MMV atomic norm methods. However, a major problem of the existing atomic norm methods is that the frequency spectrum can be recovered only when the frequencies are sufficiently separated, prohibiting commonly known high resolution–the capability of resolving two closely spaced frequency components. In theory a sufficient frequency separation condition is $\frac{4}{N}$. Empirically, this number can be reduced to $\frac{1}{N}$ (see [5]); it also was shown to depend on K, M and the number of measurement vectors (see [7, 8]). Another gridless sparse method is gridless SPICE (GLS), which was derived from a statistical perspective and is hyperparameter-free but is closely related to the atomic norm methods (see [8–10]).

To break the resolution limit of the existing atomic norm methods, in this paper, we propose a high resolution gridless sparse method for super-resolution (possibly, with compressive data and MMVs). Our method is motivated by the formulations and properties of atomic ℓ_0 norm and the atomic norm in [6, 7]. In particular, the atomic ℓ_0 norm exhibits no resolution limit but is NP hard to compute. To the contrary, as a convex relaxation the atomic norm can be efficiently computed but suffers from a resolution limit as mentioned above. We propose a novel sparse metric and theoretically show that the new metric fills the gap between the atomic ℓ_0 norm and the atomic norm. It approaches the former under appropriate parameter setting. With the sparse metric we formulate a nonconvex optimization problem and present a locally convergent iterative algorithm. The algorithm iteratively carries out ANM with a sound reweighting strategy, which determines preference of frequency selection based on the latest estimate and enhances sparsity and resolution, and is termed as reweighted atomic-norm minimization (RAM). To the best of our knowledge, RAM implements the first reweighting strategy in the continuous dictionary setting whereas existing reweighted algorithms (see, e.g., [11]) can only deal with the discrete setting. Extensive numerical simulations are provided to demonstrate the high resolution performance of RAM with application to DOA estimation compared to existing arts.

2. PRELIMINARY RESULTS

2.1. Problem Formulation

We consider the super-resolution problem in the most general case with compressive samples and MMVs. In particular, we observe the samples of the data matrix $\boldsymbol{Y}^o \in \mathbb{C}^{N\times L}$ on the rows indexed by $\boldsymbol{\Omega} \subset [N] \triangleq \{1, 2, \ldots, N\}$ of size $M = |\boldsymbol{\Omega}| \leq N$, denoted by $\boldsymbol{Y}^o_{\boldsymbol{\Omega}}$. The (j,t)th element of $\boldsymbol{Y}^o$ is (corrupted by noise in practice)

$$y^o_{jt} = \sum_{k=1}^{K} \boldsymbol{a}(f_k)\, \boldsymbol{s}_k, \quad (j,t) \in [N]\times[L], \tag{1}$$

where $\boldsymbol{a}(f) = \left[1, e^{i2\pi f}, \ldots, e^{i2\pi(N-1)f}\right]^T \in \mathbb{C}^N$ denotes a discrete complex sinusoid with frequency $f \in \mathbb{T} \triangleq [0,1]$, and $\boldsymbol{s}_k \in \mathbb{C}^{1\times L}$ is the coefficient vector of the kth sinusoid. That is, each column of $\boldsymbol{Y}^o$ is a linear combination of K discrete sinusoids. We

This research was supported in part by the National Research Foundation of China under grant NSFC 61120106011.

are interested in recovering the frequencies $\{f_k\}$ given $\boldsymbol{Y}^o_{\boldsymbol{\Omega}}$. Meanwhile, it is also of interest to recover the full data matrix $\boldsymbol{Y}^o$. The resulting problem is also known as continuous/off-grid CS, which differs from the existing CS framework in the sense that every frequency f_k can take any continuous value in $\mathbb{T}$ rather than constrained on a finite discrete grid (see [5, 6]). The single-measurement-vector (SMV) case where $L = 1$ is known as line spectral estimation. The MMV case where $L > 1$ is common in array processing. Therein the sampling index set $\boldsymbol{\Omega}$ refers to sensor placement of a linear sensor array and a smaller sample size means use of less sensors. $\boldsymbol{Y}^o_{\boldsymbol{\Omega}}$ consists of measurements of the sensor array and each column vector corresponds to one data snapshot. Each frequency corresponds to the direction of one source. Therefore, the frequency estimation problem is known as direction of arrival (DOA) estimation.

2.2. Existing Gridless Sparse Methods

The super-resolution or continuous CS problem is tackled from the perspective of signal recovery. The frequencies are then retrieved from the computational result. In particular, we seek a *frequency-sparse* candidate $\boldsymbol{Y}$, which is composed of a few frequency components, in a feasible domain defined by the observed samples. To do this, we first define a sparse metric of $\boldsymbol{Y}$ and then optimize the metric over the feasible domain. A direct sparse metric is the smallest number of frequency components composing $\boldsymbol{Y}$, known as the atomic ℓ_0 norm and denoted by $\|\boldsymbol{Y}\|_{\mathcal{A},0}$. It follows from [5–7] that $\|\boldsymbol{Y}\|_{\mathcal{A},0}$ can be characterized as the following rank minimization problem:

$$\begin{aligned} \|\boldsymbol{Y}\|_{\mathcal{A},0} = &\min_{\boldsymbol{u}} \operatorname{rank}\left(T\left(\boldsymbol{u}\right)\right), \\ &\text{subject to } \operatorname{tr}\left(\boldsymbol{Y}^H T\left(\boldsymbol{u}\right)^{-1} \boldsymbol{Y}\right) < +\infty, \\ &T\left(\boldsymbol{u}\right) \succeq \boldsymbol{0}. \end{aligned} \tag{2}$$

The first constraint in (2) imposes that $\boldsymbol{Y}$ lies in the range space of a (Hermitian) Toeplitz matrix $T\left(\boldsymbol{u}\right) \in \mathbb{C}^{N\times N}$ whose first row is specified by the transpose of $\boldsymbol{u} \in \mathbb{C}^N$. The frequencies composing $\boldsymbol{Y}$ are encoded in $T\left(\boldsymbol{u}\right)$. Once an optimizer of $\boldsymbol{u}$, say $\boldsymbol{u}^*$, is obtained the frequencies can be retrieved from $T\left(\boldsymbol{u}^*\right)$ using the Vandermonde decomposition, which says that any positive semidefinite (PSD) Toeplitz matrix $T\left(\boldsymbol{u}^*\right)$ can be decomposed as $T\left(\boldsymbol{u}^*\right) = \sum_{k=1}^{K^*} p_k^* \boldsymbol{a}\left(f_k^*\right)\boldsymbol{a}\left(f_k^*\right)^H$, where the order $K^* = \operatorname{rank}\left(T\left(\boldsymbol{u}^*\right)\right)$ and $p_k^* > 0$ (see [1] and a computational method in [8, Appendix A]). The atomic ℓ_0 norm directly enhances sparsity, however, it is NP-hard to compute and encourages computationally feasible alternatives. In this spirit, the atomic (ℓ_1) norm, denoted by $\|\boldsymbol{Y}\|_{\mathcal{A}}$, is introduced as a convex relaxation of $\|\boldsymbol{Y}\|_{\mathcal{A},0}$ and has the following semidefinite formulation [5–7]:

$$\begin{aligned} \|\boldsymbol{Y}\|_{\mathcal{A}} = &\min_{\boldsymbol{u}} \frac{1}{2\sqrt{N}}\left[\operatorname{tr}\left(T\left(\boldsymbol{u}\right)\right) + \operatorname{tr}\left(\boldsymbol{Y}^H T\left(\boldsymbol{u}\right)^{-1}\boldsymbol{Y}\right)\right], \\ &\text{subject to } T\left(\boldsymbol{u}\right) \succeq \boldsymbol{0}. \end{aligned} \tag{3}$$

From the perspective of low rank matrix recovery (LRMR), (3) attempts to recover the low rank matrix $T\left(\boldsymbol{u}\right)$ by relaxing the pseudo-rank norm in (2) to the nuclear norm (or the trace norm for a PSD matrix). The atomic norm is advantageous in computation compared to the atomic ℓ_0 norm; however, it suffers from a resolution limit due to the relaxation which is not shared by the latter [4, 5, 7].

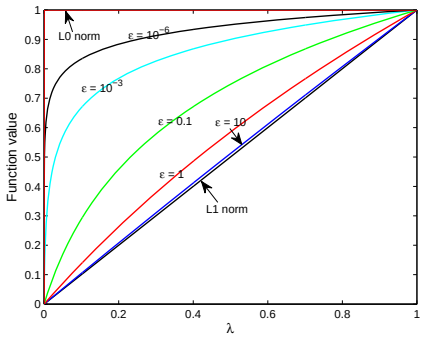

Fig. 1. The sparsity-promoting property of $\mathcal{M}^\epsilon\left(\cdot\right)$ with different ϵ. The plotted curves include the ℓ_0 and ℓ_1 norms corresponding to $\|\cdot\|_{\mathcal{A},0}$ and $\|\cdot\|_{\mathcal{A}}$ respectively, and $\ln|\lambda+\epsilon|$ corresponding to $\mathcal{M}^\epsilon\left(\cdot\right)$ with $\epsilon = 10, 1, 0.1, 10^{-3}$ and 10^{-6}. $\ln|\lambda+\epsilon|$ is translated and scaled such that it equals 0 and 1 at $\lambda = 0$ and 1 respectively for better illustration.

3. ENHANCING SPARSITY AND RESOLUTION VIA A NOVEL SPARSE METRIC

Inspired by the link between continuous CS and LRMR, we propose the following sparse metric of $\boldsymbol{Y}$:

$$\begin{aligned} \mathcal{M}^\epsilon\left(\boldsymbol{Y}\right) = &\min_{\boldsymbol{u}} \ln\left|T\left(\boldsymbol{u}\right) + \epsilon\boldsymbol{I}\right| + \operatorname{tr}\left(\boldsymbol{Y}^H T\left(\boldsymbol{u}\right)^{-1}\boldsymbol{Y}\right), \\ &\text{subject to } T\left(\boldsymbol{u}\right) \succeq \boldsymbol{0}, \end{aligned} \tag{4}$$

where $\epsilon > 0$ is a regularization parameter. Note that the log-det heuristic $\ln|\cdot|$ is a common smooth surrogate of the matrix rank (see, e.g., [12]). From the perspective of LRMR, the atomic ℓ_0 norm minimizes the number of nonzero eigenvalues of $T\left(\boldsymbol{u}\right)$ while the atomic norm minimizes the sum of the eigenvalues. In contrast, the new metric $\mathcal{M}^\epsilon\left(\boldsymbol{Y}\right)$ puts penalty on $\sum_{k=1}^N \ln|\lambda_k+\epsilon|$, where $\{\lambda_k\}_{k=1}^N$ denotes the eigenvalues. The function $h(\lambda) = \ln|\lambda+\epsilon|$ with different ϵ's is plotted in Fig. 1. We expect that the new metric $\mathcal{M}^\epsilon\left(\boldsymbol{Y}\right)$ bridges $\|\boldsymbol{Y}\|_{\mathcal{A}}$ and $\|\boldsymbol{Y}\|_{\mathcal{A},0}$ when ϵ varies from $+\infty$ to 0. Formally, we have the results below and we provide their proofs in the full version of this paper [13].

Theorem 1 *Let $\epsilon \to +\infty$. Then,*

$$\mathcal{M}^\epsilon\left(\boldsymbol{Y}\right) - N\ln\epsilon \sim 2\sqrt{N}\,\|\boldsymbol{Y}\|_{\mathcal{A}}\,\epsilon^{-\frac{1}{2}}, \tag{5}$$

i.e., they are equivalent infinitesimals.

Theorem 2 *Let $\epsilon \to 0$. Then, we have the following results:*

1. *If $\|\boldsymbol{Y}\|_{\mathcal{A},0} \le N-1$, then*

$$\mathcal{M}^\epsilon\left(\boldsymbol{Y}\right) \sim \left(\|\boldsymbol{Y}\|_{\mathcal{A},0} - N\right)\ln\frac{1}{\epsilon}, \tag{6}$$

   i.e., they are equivalent infinities. Otherwise, $\mathcal{M}^\epsilon\left(\boldsymbol{Y}\right)$ is a positive constant depending only on $\boldsymbol{Y}$;

2. *Let $\boldsymbol{u}_\epsilon^*$ be the optimizer of $\boldsymbol{u}$ to the optimization problem in (4). Then, the smallest $N - \|\boldsymbol{Y}\|_{\mathcal{A},0}$ eigenvalues of $T\left(\boldsymbol{u}_\epsilon^*\right)$ are either zero or approach zero as fast as ϵ;*

3. *For any cluster point of $\boldsymbol{u}_\epsilon^*$ at $\epsilon = 0$, denoted by $\boldsymbol{u}_0^*$, there exists an atomic decomposition $\boldsymbol{Y} = \sum_{k=1}^K \boldsymbol{a}\left(f_k\right)\boldsymbol{s}_k$ of order $K = \|\boldsymbol{Y}\|_{\mathcal{A},0}$ such that $T\left(\boldsymbol{u}_0^*\right) = \sum_{k=1}^K \|\boldsymbol{s}_k\|_2^2\,\boldsymbol{a}\left(f_k\right)\boldsymbol{a}\left(f_k\right)^H$.*

Theorem 1 shows that the new metric $\mathcal{M}^{\epsilon}(\boldsymbol{Y})$ plays the same role as $\|\boldsymbol{Y}\|_{\mathcal{A}}$ in the limiting scenario when $\epsilon \to +\infty$, while Theorem 2 says that it is equivalent to $\|\boldsymbol{Y}\|_{\mathcal{A},0}$ as $\epsilon \to 0$. Consequently, it fills the gap between $\|\boldsymbol{Y}\|_{\mathcal{A}}$ and $\|\boldsymbol{Y}\|_{\mathcal{A},0}$ and enhances sparsity and resolution compared to $\|\boldsymbol{Y}\|_{\mathcal{A}}$ as ϵ gets small. Moreover, Theorem 2 characterizes the properties of the optimizer $\boldsymbol{u}_{\epsilon}^{*}$ as $\epsilon \to 0$ including the convergent speed of the smallest $N-K$ eigenvalues and the limiting form of $T(\boldsymbol{u}^{*})$ via the Vandermonde decomposition. In fact, we always observe via simulations that the smallest $N-K$ eigenvalues of $T(\boldsymbol{u}^{*})$ become zero once ϵ is modestly small.

4. REWEIGHTED ATOMIC-NORM MINIMIZATION

4.1. A Locally Convergent Iterative Algorithm

With the proposed sparse metric $\mathcal{M}^{\epsilon}(\boldsymbol{Y})$, we solve the following optimization problem for signal and frequency recovery:

$$\min_{\boldsymbol{Y}} \mathcal{M}^{\epsilon}(\boldsymbol{Y}), \text{ subject to } \boldsymbol{Y} \in \mathcal{D}, \tag{7}$$

or equivalently,

$$\begin{aligned} &\min_{\boldsymbol{Y},\boldsymbol{u}} \ln|T(\boldsymbol{u}) + \epsilon \boldsymbol{I}| + \text{tr}\left(\boldsymbol{Y}^{H} T(\boldsymbol{u})^{-1} \boldsymbol{Y}\right), \\ &\text{subject to } T(\boldsymbol{u}) \succeq \boldsymbol{0} \text{ and } \boldsymbol{Y} \in \mathcal{D}, \end{aligned} \tag{8}$$

where $\mathcal{D}$ denotes the feasible domain of $\boldsymbol{Y}$. For example, in the noiseless case, it is the set $\{\boldsymbol{Y} : \boldsymbol{Y}_{\boldsymbol{\Omega}} = \boldsymbol{Y}^{o}_{\boldsymbol{\Omega}}\}$. Since the log-det term $\ln|T(\boldsymbol{u}) + \epsilon \boldsymbol{I}|$ is a concave function of $\boldsymbol{u}$, the problem is nonconvex and no efficient algorithms can guarantee to obtain the global optimum. A majorization-maximization (MM) algorithm is introduced as follows. Let $\boldsymbol{u}_j$ denote the jth iterate of the optimization variable $\boldsymbol{u}$. Then, at the $(j+1)$th iteration we replace $\ln|T(\boldsymbol{u}) + \epsilon \boldsymbol{I}|$ by its tangent plane at the current value $\boldsymbol{u} = \boldsymbol{u}_j$. As a result, the optimization problem at the $(j+1)$th iteration becomes

$$\begin{aligned} &\min_{\boldsymbol{Y},\boldsymbol{u}} \text{tr}\left[\left(T(\boldsymbol{u}_j) + \epsilon \boldsymbol{I}\right)^{-1} T(\boldsymbol{u})\right] + \text{tr}\left(\boldsymbol{Y}^{H} T(\boldsymbol{u})^{-1} \boldsymbol{Y}\right), \\ &\text{subject to } T(\boldsymbol{u}) \succeq \boldsymbol{0},\ \boldsymbol{Y} \in \mathcal{D}. \end{aligned} \tag{9}$$

Since $\ln|T(\boldsymbol{u}) + \epsilon \boldsymbol{I}|$ is strictly concave in $\boldsymbol{u}$, at each iteration its value decreases by an amount greater than the decrease of its tangent plane. It follows that the objective function in (8) monotonically decreases at each iteration and converges to a local minimum.

4.2. Interpretation as RAM

To interpret the optimization problem in (9), let us define a *weighted continuous dictionary*

$$\mathcal{A}^{w} \triangleq \{\boldsymbol{a}^{w}(f) = w(f)\,\boldsymbol{a}(f) :\ f \in \mathbb{T}\} \tag{10}$$

w.r.t. the original continuous dictionary $\{\boldsymbol{a}(f) : f \in \mathbb{T}\}$, where $w(f) \geq 0$ is a weighting function. For $\boldsymbol{Y} \in \mathbb{C}^{N \times L}$, we define its *weighted atomic norm* w.r.t. $\mathcal{A}^{w}$ as its atomic norm induced by $\mathcal{A}^{w}$:

$$\begin{aligned} \|\boldsymbol{Y}\|_{\mathcal{A}^{w}} &\triangleq \inf_{\boldsymbol{s}_k^{w},\, f_k} \left\{ \sum_k \|\boldsymbol{s}_k^{w}\|_2 : \boldsymbol{Y} = \sum_k \boldsymbol{a}^{w}(f_k)\, \boldsymbol{s}_k^{w} \right\} \\ &= \inf_{\boldsymbol{s}_k,\, f_k} \left\{ \sum_k w(f_k)^{-1} \|\boldsymbol{s}_k\|_2 : \boldsymbol{Y} = \sum_k \boldsymbol{a}(f_k)\, \boldsymbol{s}_k \right\}. \end{aligned} \tag{11}$$

By the definition above, $w(f)$ specifies preference of the atoms $\{\boldsymbol{a}(f)\}$. To be specific, an atom $\boldsymbol{a}(f_0)$, $f_0 \in \mathbb{T}$, is more likely selected if $w(f_0)$ is larger. Moreover, the atomic norm is a special case of the weighted atomic norm with a constant weighting function (i.e., without any preference).

Theorem 3 *Assume that* $w(f) = \frac{1}{\sqrt{\boldsymbol{a}(f)^{H} \boldsymbol{W} \boldsymbol{a}(f)}}$, *where* $\boldsymbol{W} \in \mathbb{C}^{N \times N}$. *Then,*

$$\begin{aligned} \|\boldsymbol{Y}\|_{\mathcal{A}^{w}} = &\min_{\boldsymbol{u}} \frac{\sqrt{N}}{2} tr\left(\boldsymbol{W} T(\boldsymbol{u})\right) + \frac{1}{2\sqrt{N}} tr\left(\boldsymbol{Y}^{H} T(\boldsymbol{u})^{-1} \boldsymbol{Y}\right), \\ &\textit{subject to } T(\boldsymbol{u}) \succeq \boldsymbol{0}. \end{aligned} \tag{12}$$

Let $\boldsymbol{W}_j = \frac{1}{N}\left(T(\boldsymbol{u}_j) + \epsilon \boldsymbol{I}\right)^{-1}$ and $w_j(f) = \frac{1}{\sqrt{\boldsymbol{a}^{H}(f) \boldsymbol{W}_j \boldsymbol{a}(f)}}$. By Theorem 3 we can write the optimization problem in (9) as the following *weighted atomic norm minimization* problem:

$$\min_{\boldsymbol{Y}} \|\boldsymbol{Y}\|_{\mathcal{A}^{w_j}}, \text{ subject to } \boldsymbol{Y} \in \mathcal{D}. \tag{13}$$

As a result, the proposed iterative algorithm can be interpreted as *reweighted atomic-norm minimization* (RAM). If we let $w_0(f)$ be a constant function or equivalently, $\boldsymbol{u}_0 = \boldsymbol{0}$, such that there is no preference of the atoms at the first iteration, then the first iteration coincides with the ANM. From the second iteration on, the preference is defined by the weighting function $w_j(f)$ specified above. Note that $w_j^{2}(f)$ corresponds to the power spectrum of Capon's beamforming (see, e.g., [1]) if $T(\boldsymbol{u}_j)$ is interpreted as the covariance of the noiseless data and ϵ as the noise variance. Therefore, the reweighting strategy makes the frequencies around the estimates of the current iteration preferable at the ensuing iteration and thus enhances sparsity. At the same time, the preference gradually highlights the spectrum near the estimated frequencies and therefore enhances resolution. Since the "noise variance" ϵ can be translated as the confidence level in the current estimate, from this perspective we should gradually decrease ϵ and correspondingly increase the confidence in the solution during the algorithm.

5. NUMERICAL SIMULATIONS

5.1. Sparsity-Separation Phase Transition

In this subsection, we study the success rate of RAM for super-resolution compared to ANM. In particular, we fix $N = 64$ and $M = 30$ with the sampling index set $\boldsymbol{\Omega}$ being generated uniformly at random. We vary the duo (K, Δ_f). At each combination K frequencies are randomly generated such that they are mutually separated by at least Δ_f. The amplitudes $\{s_{kt}\}$ are independently generated from a standard complex normal distribution. After obtaining the noiseless samples, we carry out super-resolution using ANM and RAM, both implemented by an off-the-shelf SDP solver SDPT3 [14]. The recovery is called successful if both the relative MSE of signal recovery and the MSE of frequency recovery are less than 10^{-12}. At each combination (K, Δ_f), the success rate is measured over 20 Monte Carlo runs. In RAM, we first scale the measurements such that $\|\boldsymbol{Y}_{\boldsymbol{\Omega}}\|_{\text{F}}^{2} = M$ and compensate the recovery afterwards. We start with $\boldsymbol{u}_0 = \boldsymbol{0}$ and $\epsilon = 1$ as default. We halve ϵ when beginning a new iteration until $\epsilon = \frac{1}{2^{10}}$. We terminate RAM if the relative change (in Frobenius norm) of the solution $\boldsymbol{Y}^{*}$ at two consecutive iterations is less than 10^{-6} or the maximum number of iterations, set to 20, is reached.

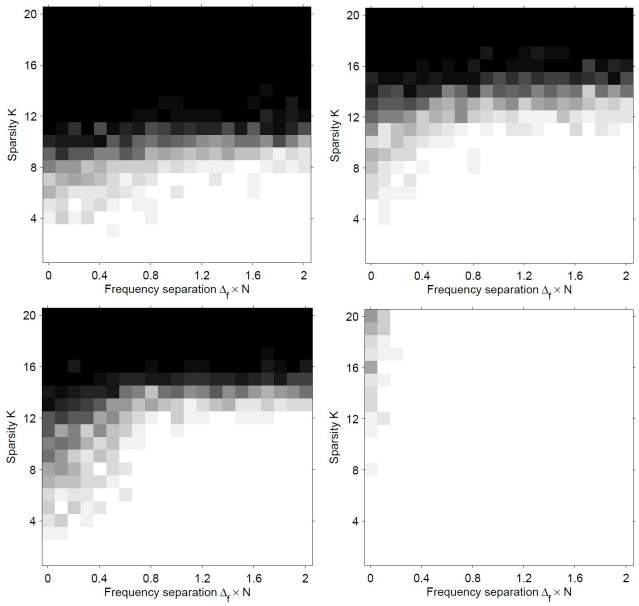

Fig. 2. Sparsity-separation phase transitions of ANM (left) and RAM (right) with $L = 1$ (top) and $L = 5$ (bottom), $N = 64$ and $M = 30$. The grayscale images present the success rates, where white and black colors indicate complete success and complete failure, respectively.

We plot the success rates of ANM and RAM with $L = 1, 5$ in Fig. 2, where it is shown that successful recovery can be obtained with more ease with a smaller K and a larger frequency separation Δ_f, leading to a phase transition in the sparsity-separation domain. It is shown that RAM significantly enlarges the success phase and hence enhances sparsity and resolution compared to ANM. At $L = 5$ we did not find a single failure in our simulation whenever $K \leq 20$ and $\Delta_f \geq \frac{0.3}{N}$. The phase transitions of both ANM and RAM are not sharp since the frequencies are separated by *at least* Δ_f and a set of *well separated* frequencies can be possibly generated at a small value of Δ_f. It is also observed that RAM tends to converge in less iterations with a smaller K and a larger Δ_f.

5.2. Application to DOA Estimation

We apply the proposed RAM method to DOA estimation. In particular, we consider a 10-element sparse linear array (SLA) with sensors' positions indexed by $\mathbf{\Omega} = \{1, 2, 5, 6, 8, 12, 15, 17, 19, 20\}$, where the distance between the first two sensors is half the wavelength. Hence, we have that $N = 20$ and $M = 10$. We consider that $K = 4$ narrowband sources impinge on the sensor array from directions corresponding to frequencies 0.1, 0.11, 0.2 and 0.5, and powers 10, 10, 3 and 1, respectively. It is challenging to separate the first two sources which are separated by only $\frac{0.2}{N}$. Complex normal noise is added to the samples with variance $\sigma^2 = 1$ and $\mathcal{D}$ is defined as $\left\{\boldsymbol{Y} : \|\boldsymbol{Y}_{\mathbf{\Omega}} - \boldsymbol{Y}^o_{\mathbf{\Omega}}\|_{\mathrm{F}} \leq \eta\right\}$, where $\eta^2 = \left(ML + 2\sqrt{ML}\right)\sigma^2$ (mean + twice standard deviation) upper bounds the noise energy with high probability. We consider both the cases of uncorrelated and correlated sources while the latter case is more difficult to dealt with using existing methods such as MUSIC (see, e.g., [1]). In the latter case, sources 1 and 3 are set to be coherent (completely correlated). Assume that $L = 200$ data snapshots are collected which

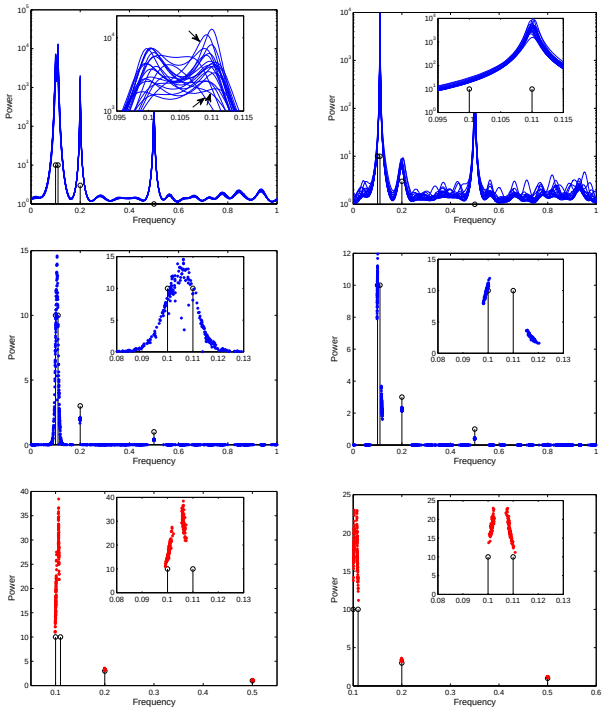

Fig. 3. Results of MUSIC (top), ANM (middle) and RAM (bottom) for DOA estimation with uncorrelated (left) and correlated (right) sources in 100 Monte Carlo runs. The area around the first two sources are zoomed in in each subfigure.

are corrupted by i.i.d. Gaussian noise of unit variance. We propose a dimension reduction technique to reduce the order of the SDP matrix from $L + N$ to $M + N$ and accelerate the computational speed, which is detailed in [13]. We terminate RAM within maximally 10 iterations and consider MUSIC and ANM for comparison.

Our simulation results of 100 Monte Carlo runs are presented in Fig. 3 (for MUSIC only the first 20 runs are presented for better illustration). In the absence of source correlations, MUSIC has satisfactory performance in most scenarios. However, its power spectrum exhibits only a single peak around the first two sources (i.e., the two sources cannot be separated) in at least 3 out of the first 20 runs (indicated by the arrows). Moreover, MUSIC is sensitive to source correlations and cannot detect source 1 when it is coherent with source 3. ANM cannot separate the first two sources in the uncorrelated source case and always produces many spurious sources. In contrast, the proposed RAM always correctly detects 4 sources near the true locations, demonstrating its capabilities in enhancing sparsity and high resolution. ANM and RAM take 0.87s and 7.31s on average, respectively, while these numbers can be greatly decreased using more sophisticated algorithms (see [13]).

6. CONCLUSION

In this paper, we studied the super-resolution problem with compressive samples and MMVs. Motivated by its connection to LRMR, we proposed reweighted atomic-norm minimization (RAM) for achieving high resolution compared to the currently prominent atomic norm minimization (ANM) and validated its performance via numerical simulations.

7. REFERENCES

[1] P. Stoica and R. L. Moses, *Spectral analysis of signals*. Pearson/Prentice Hall Upper Saddle River, NJ, 2005.

[2] Y. Chi, L. Scharf, A. Pezeshki, and A. Calderbank, "Sensitivity to basis mismatch in compressed sensing," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2182–2195, 2011.

[3] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse Bayesian inference," *IEEE Transactions on Signal Processing*, vol. 61, no. 1, pp. 38–43, 2013.

[4] E. J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Communications on Pure and Applied Mathematics*, vol. 67, no. 6, pp. 906–956, 2014.

[5] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7465–7490, 2013.

[6] Z. Yang and L. Xie, "Continuous compressed sensing with a single or multiple measurement vectors," in *IEEE Workshop on Statistical Signal Processing (SSP)*, 2014, pp. 308–311.

[7] Z. Yang and L. Xie, "Exact joint sparse frequency recovery via optimization methods," *Available online at http://arxiv.org/abs/1405.6585*, 2014.

[8] Z. Yang and L. Xie, "On gridless sparse methods for line spectral estimation from complete and incomplete data," *IEEE Transactions on Signal Processing, accepted with minor revisions (AQ)*, 2014.

[9] Z. Yang, L. Xie, and C. Zhang, "A discretization-free sparse and parametric approach for linear array signal processing," *IEEE Transactions on Signal Processing*, vol. 62, no. 19, pp. 4959–4973, 2014.

[10] P. Stoica, G. Tang, Z. Yang, and D. Zachariah, "Gridless compressive-sensing methods for frequency estimation: Points of tangency and links to basics," in *22nd European Signal Processing Conference (EUSIPCO)*, 2014, pp. 1831–1835.

[11] E. J. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5-6, pp. 877–905, 2008.

[12] M. Fazel, H. Hindi, and S. P. Boyd, "Log-det heuristic for matrix rank minimization with applications to Hankel and Euclidean distance matrices," in *American Control Conference*, vol. 3, 2003, pp. 2156–2162.

[13] Z. Yang and L. Xie, "Enhancing sparsity and resolution via a reweighted atomic norm minimization," *IEEE Transactions on Signal Processing, under review, Available online at http://arxiv.org/abs/1408.5750*, 2014.

[14] K.-C. Toh, M. J. Todd, and R. H. Tütüncü, "SDPT3–a MATLAB software package for semidefinite programming, version 1.3," *Optimization Methods and Software*, vol. 11, no. 1-4, pp. 545–581, 1999.