



## Hybrid Reweighted Optimization Method for Gridless Direction of Arrival Estimation in Heteroscedastic Noise Environment

Lama Zien Alabideen\*, Oumayma Al-Dakkak, Khaldoun Khorzom

Department of Telecommunication, Higher Institute for Applied Sciences and Technology (HIAST), Damascus 31983, Syria

Corresponding Author Email: [lama.zienalabideen@hiast.edu.sy](mailto:lama.zienalabideen@hiast.edu.sy)

<https://doi.org/10.18280/mmep.080116>

### ABSTRACT

**Received:** 18 July 2020

**Accepted:** 6 November 2020

#### Keywords:

*Gridless DoA estimation, atomic norm, covariance fitting, co-prime array, heteroscedastic noise*

In this paper, we present a hybrid optimization framework for gridless sparse Direction of Arrival (DoA) estimation under the consideration of heteroscedastic noise scenarios. The key idea of the proposed framework is to combine global and local minima search techniques that offer a sparser optimizer with boosted immunity to noise variation. In particular, we enforce sparsity by means of reformulating the Atomic Norm Minimization (ANM) problem through applying the nonconvex Schatten-p quasi-norm ( $0 < p < 1$ ) relaxation. In addition, to enhance the adaptability of the relaxed ANM in more practical noise scenarios, it is combined with a covariance fitting (CF) criterion resulting in a locally convergent reweighted iterative approach. This combination forms a hybrid optimization framework and offers the advantages of both optimization approaches while balancing their drawbacks. Numerical simulations are performed taking into account the configuration of co-prime array (CPA). The simulations have demonstrated that the proposed method can maintain a high estimation resolution even in heteroscedastic noise environments, a low number of snapshots, and correlated sources.

## 1. INTRODUCTION

DoA estimation is one of the most addressed topics in recent researches, which is playing a significant role in several applications such as smart antennas, wireless communication, and radar. The major objective of DoA estimation is to acquire the direction information of multiple sources from the outputs of the sensor array. Many methods have been proposed for this purpose, including conventional subspace-based techniques such as MUSIC, ESPRIT, and their variants [1-3]. In general, these methods are subject to certain well-known disadvantages. For example, they require the number of sources and sufficient numbers of snapshots to attain accurate estimation, that may not be available in practice. Moreover, they are affected considerably by the correlations of sources.

Major advances have been made in the DoA estimation over the last decade with the evolution of sparse representation and compressed sensing (CS) [4, 5], which yields to a variety of sparse methods for estimating DoA. Early grid-based sparse approaches have been devoted to discrete linear systems and encounter the basis or grid mismatch problem [6, 7]. Many off-grid solutions have been proposed to address this problem [8, 9]. However, these methods impose high complexity and maintain the grid determination issue. Recently, the gridless sparse methods have been presented [10, 11]. These methods remove gridding by directly working in a continuous dictionary and estimate the DoAs by solving a semidefinite programming (SDP) problem.

One of the distinct gridless sparse approaches in DoA estimation depends on the notion of atomic norm minimization (ANM) introduced by Chandrasekaran et al. [12]. ANM is a deterministic approach that seeks the minimum number of

“atoms” from a manifold required to reconstruct a signal. The theoretical performance of the noiseless ANM problem was investigated by Candès and Fernandez-Granda [13] without any statistical assumptions. Mainly, it has been successfully used to estimate a spectrally sparse signal, provided that the frequency components are adequately separated by minimum  $\frac{4}{n}$ , where  $n$  indicates the dimension of full data. Later, this condition was minimized to  $\frac{2.52}{n}$  by Fernandez-Granda [14]. The ANM problem was expanded to the compressive data case and referred to as continuous CS [15]. Meanwhile, several theoretical works [16-19] have investigated the robustness of ANM in the case of corrupted measurements. Atomic norm denoising approach was presented for line spectral estimation application by Bhaskar et al. [17]. This approach was further extended to the multiple measurement vectors (MMVs) case [8, 18]. Tang et al. [19] proposed a nearly optimal method to denoise a mixture of complex sinusoids from full data case, which was designated as atomic norm soft thresholding (AST). Atomic norm denoising is applied to test the efficiency of super-resolution under the consideration of white noise, it offered theoretical guarantees for parameter recovery [20].

One competitive alternative approach is the structured covariance fitting (CF) [10, 21-23], which takes advantage of the second-order statistical properties of the directions coefficients. In particular, CF-based methods exploit the Toeplitz and Hermitian structure of the covariance matrix. As well, they exhibit high estimation accuracy without requiring to angle separations as long as the statistical assumptions of a sufficient number of snapshots and non-correlated sources are respected.

Although gridless DoA estimators enjoy remarkable

advantages, they have considerable drawbacks. In general, CF-based approaches are statistically compatible in the total count of snapshots and do not provide adequate results in the case of the correlated sources. While ANM-based methods suffer from restricted resolution due to the separation condition and unsatisfying performance in the modest signal-to-noise ratio (SNR). Furthermore, regarding noise, ANM-based estimators assume that the noises of sensor array are white with identical variances (i.e. homoscedastic noise). This assumption helps to set the regularization parameter in the optimization problem. However, this may be violated in a range of practical applications due to the imperfect real implementation or non-idealities of the communication channel. Further, the noise may be heteroscedastic and exhibit non-stationary manners by changing its variance in both space and time. It is worth mentioning that CF is recommended in the absence of the noise level or in the presence of heteroscedastic noise, while ANM approaches are preferred when the sources are strongly correlated or when the noise is considered to have the upper bound power [11].

Reweighted optimization is common in the literature on sparse DoA estimation. Typically, it is used to reduce the limitation of estimation resolution in ANM-based methods [24, 25], or to encounter solution deviation from the prime problem (i.e. rank minimization) when other assumptions are violated (e.g. sources separation limit, signals correlations) -as in the case of CF-based methods [23, 26]. Moreover, some reweighted methods are derived using nonconvex optimization. Compared to the convex problem, it has been shown that applying nonconvex surrogate functions, such as  $l_p$  norm, Logarithm, Laplace, and Schatten-p quasi-norm [23, 27, 28], can achieve better sparse parameter estimation and signal recovery within a few iterations. In the notable work of [24], the so-called reweighted ANM (RAM) was suggested based on a nonconvex log-det heuristic, RAM provides sparsity and resolution enhancement, with applicability to DoA estimation. Our earlier work [29] presented a high-resolution reweighted CF approach for gridless DoA estimation based on nonconvex Schatten-p quasi-norm that provided better performance compared to the state-of-the-art methods. The major drawback of the methods derived from nonconvex penalties is that they suffer from local convergence. Table 1 summarizes the gridless DoA estimation methods based on nonconvex optimization with their penalties.

**Table 1.** The gridless DoA estimation methods and their dependent approaches and nonconvex penalties

Method	Approach	Penalty	Formula $\mathbf{X} \in \mathbb{C}^{n \times m}, x =$ $\sigma(\mathbf{X}) \geq 0, \varepsilon > 0$
RAM [24]	ANM	Log-det	$\ln \mathbf{X} + \varepsilon \mathbf{I} $
[26]	CF	Laplace	$\sum (1 - e^{-\frac{x}{\varepsilon}})$
ICMRA [23]		Logarithm	$\sum \ln(x + \varepsilon)$
Rwp-GLS [29]		Schatten-p	$\sum (x + \varepsilon)^p$

## 2. MAIN CONTRIBUTIONS

In this paper, we study the problem of denoising the gridless DoA estimation, from compressive and corrupted

measurements of multiple sparse signals, in the spatial domain. Our motivation is twofold: (i) take advantage of the attractive properties of the recent nonconvex optimization techniques in low-rank matrix recovery. Accordingly, we have introduced the nonconvex Schatten-p quasi-norm surrogate function as a sparsity metric. Then the resulting objective function is linearized by the use of Taylor expansion leading to a locally converging reweighted iterative method. (ii) Cast the ANM-based DoA estimation in the background of anonymous practical noise cases (i.e. heteroscedastic noise) by incorporating the covariance fitting criterion into the proposed ANM problem, which has not been analyzed in the literature to the best of our knowledge. Furthermore, with regard to the main role of array geometry, and with a view to achieving a high degree of freedom, with a reasonable number of physical sensors, we have followed the co-prime array (CPA) structure [30, 31].

The remainder of the paper will be arranged as follows. Section 3 demonstrates the signal model in the DoA estimation problem. Section 4 revisits the preliminary results of ANM and CF frameworks that motivate this paper. Section 5 presents the suggested hybrid method for gridless DoA estimation. Section 6 provides numerical simulations to validate the performance of the proposed method. Finally, the paper is concluded in Section 7.

Through this paper, we use bold letters for vectors and matrices.  $\mathbb{C}$  and  $\mathbb{R}$  stand for the sets of complex and real numbers respectively. The transpose is symbolized by  $(\cdot)^T$ , and the complex conjugate or Hermitian is symbolized by  $(\cdot)^H$ .  $rank(\cdot)$  and  $tr(\cdot)$  denote matrix rank and trace respectively.  $\|\cdot\|_F$  states to Frobenius norm, and  $\mathbf{A} \geq 0$  signifies that  $\mathbf{A}$  is a Positive SemiDefinite (PSD) matrix.

## 3. SIGNAL MODEL

Consider  $K$  narrow-band far-field uncorrelated sources  $s_k$  impinge on a sparse linear CPA array from the directions of  $\boldsymbol{\theta} = [\theta_1, \dots, \theta_K]$ . Specifically, for a co-prime pair  $N > M$ , CPA consists of two uniform linear subarrays: the first composed of  $M$  sensors with an inter-element spacing of  $Nd$ . The other comprises of  $N$  sensors with an inter-element spacing of  $Md$ , where  $d$  is picked to be half of the wavelength  $\lambda/2$ . Note that the CPA contains a total of  $M + N - 1$  physical sensors in total. Further, the CPA output data matrix  $\mathbf{Y}_\Omega = [\mathbf{y}_\Omega(1), \dots, \mathbf{y}_\Omega(L)] \in \mathbb{C}^{(M+N-1) \times L}$  can be viewed as the output of virtual  $N \times M$  sensors of a uniform linear array by keeping the sensors indexed by  $\Omega = [l_1, l_2, \dots, l_{M+N-1}] \subset \{1, \dots, NM\}$ . The observation model at  $t$  snapshot of CPA can be described as follows,

$$\begin{aligned} \mathbf{y}_\Omega(t) &= \sum_{k=1}^K \mathbf{a}_\Omega(\theta_k) s_k(t) + \mathbf{n}_\Omega(t) \\ &= \mathbf{A}_\Omega(\boldsymbol{\theta}) \mathbf{s}(t) + \mathbf{n}_\Omega(t), \quad t = 1, \dots, L \end{aligned} \quad (1)$$

where,  $L$  is the total number of snapshots, and  $\mathbf{s}(t) \in \mathbb{C}^{K \times 1}$ ,  $\mathbf{y}_\Omega(t) \in \mathbb{C}^{(M+N-1) \times 1}$  and  $\mathbf{n}_\Omega(t) \in \mathbb{C}^{(M+N-1) \times 1}$  are the vectors of sources signals, the array output, and the noise, respectively.  $\mathbf{A}_\Omega(\boldsymbol{\theta}) = [\mathbf{a}_\Omega(\theta_1), \dots, \mathbf{a}_\Omega(\theta_K)] \in \mathbb{C}^{(M+N-1) \times K}$  indicates the steering matrix, where  $\mathbf{a}_\Omega(\theta_k)$  is the steering vector of the  $k^{th}$  source which is specified by the geometry of the sensor array and is set as,

$$\mathbf{a}_\Omega(\tilde{\theta}_k) = [e^{i2\pi\frac{d}{\lambda}l_1\tilde{\theta}_k}, e^{i2\pi\frac{d}{\lambda}l_2\tilde{\theta}_k}, \dots, e^{i2\pi\frac{d}{\lambda}l_{M+N-1}\tilde{\theta}_k}]^T \quad (2)$$

where,  $\tilde{\theta}_k = \sin(\theta_k)$  is  $k^{\text{th}}$  normalized ungrided direction, and  $l_i$  is the physical position of the  $i^{\text{th}}$  sensor.

Let  $\{\mathbf{s}\}_1^K$  be temporarily and spatially uncorrelated, this implies that  $E[\mathbf{s}\mathbf{s}^H] = \text{diag}(\mathbf{P})$ , where  $\mathbf{P} = [P_1, P_2, \dots, P_K] \in \mathbb{R}_+^{1 \times K}$  is the vector of the sources' powers. In addition,  $\mathbf{n}_\Omega(t)$  is independent of  $\mathbf{s}(t)$  and fulfils  $E[\mathbf{n}_\Omega(t_i)\mathbf{n}_\Omega^H(t_j)] = \text{diag}(\boldsymbol{\sigma}_\Omega)$  for any  $i \neq j$ , where  $\boldsymbol{\sigma}_\Omega$  represents the noise variance vector. In this case the sample covariance matrix is given by,

$$\mathbf{R}_\Omega = E[\mathbf{y}_\Omega\mathbf{y}_\Omega^H] = \mathbf{A}_\Omega(\tilde{\boldsymbol{\theta}})\text{diag}(\mathbf{P})\mathbf{A}_\Omega^H(\tilde{\boldsymbol{\theta}}) + \text{diag}(\boldsymbol{\sigma}_\Omega) \quad (3)$$

$\mathbf{R}_\Omega$  is a PSD Toeplitz matrix and has rank  $K$  under the assumption that  $N > K$ , while  $\tilde{\boldsymbol{\theta}} = [\tilde{\theta}_1, \dots, \tilde{\theta}_K]$ .

Shortly, the output of the array can be rewritten as,

$$\mathbf{Y}_\Omega = \mathbf{A}_\Omega(\tilde{\boldsymbol{\theta}})\mathbf{S} + \mathbf{N}_\Omega \quad (4)$$

where,  $\mathbf{S} = [\mathbf{s}(1), \dots, \mathbf{s}(L)] \in \mathbb{C}^{K \times L}$  is the sources signals matrix, and  $\mathbf{N}_\Omega = [\mathbf{n}_\Omega(1), \dots, \mathbf{n}_\Omega(L)] \in \mathbb{C}^{(M+N-1) \times L}$  is the matrix of the measurements noise.

Referring to the key objective of the DoA estimation, the purpose of this paper is to estimate the parameters  $(\boldsymbol{\theta}, \boldsymbol{\sigma}_\Omega)$  provided the compressed corrupted measurement matrix  $\mathbf{Y}_\Omega$ .

## 4. PRELIMINARY RESULTS

### 4.1 ANM approach for gridless DoA estimation

Principally, the ANM approach has strong theoretical guarantees. The main merits of ANM are embodied in exploiting the signal sparsity and resolving the grid mismatch problem by working directly on a continuous dictionary. As well, ANM preserves the signal structure, which is critical in the noisy cases. Further, it can be used in a single or limited number of snapshots cases.

Inspired by atomic norm representation, the objective of DoA problem is reduced to describe the spatial sparsity as atomic decomposition of  $\mathbf{Y}_\Omega$  with the smallest number of atoms. The atomic  $l_0$  norm offers the proper description of this problem [12, 15], and it is defined as,

$$\|\mathbf{Y}\|_{A,0} = \inf_{\tilde{\boldsymbol{\theta}}, \mathbf{s}_k} \left\{ K: \mathbf{Y} = \sum_{k=1}^K \mathbf{a}(\tilde{\theta}_k)\mathbf{s}_k: \mathbf{a}(\tilde{\theta}_k) \in \mathbf{A}, \|\mathbf{s}_k\|_2 \geq 0 \right\} \quad (5)$$

where,  $\mathbf{Y}$  refers to the full data matrix, which is a linear combination of a number of atoms  $\mathbf{a}(\tilde{\theta}_k) = [1, e^{i2\pi\tilde{\theta}_k}, \dots, e^{i2\pi(MN-1)\tilde{\theta}_k}]^T \in \mathbb{C}^{MN \times 1}$  from the continuous dictionary  $\mathbf{A} := \{\mathbf{a}(\tilde{\boldsymbol{\theta}}) : \tilde{\boldsymbol{\theta}} \in [0,1], \}$ . While,  $\mathbf{s}_k = [s_k(1), \dots, s_k(L)] \in \mathbb{C}^{1 \times L}$  represents the multiple snapshots coefficients vector of the  $k^{\text{th}}$  source. Following from [12], in the noiseless case,  $\|\mathbf{Y}\|_{A,0}$  can be characterized by the next rank minimization problem,

$$\|\mathbf{Y}\|_{A,0} = \min_{\mathbf{X}, \mathbf{u}} \text{rank}(\mathbf{T}(\mathbf{u})) \quad (6)$$

subject to  $\begin{bmatrix} \mathbf{X} & \mathbf{Y}^H \\ \mathbf{Y} & \mathbf{T}(\mathbf{u}) \end{bmatrix} \geq 0.$

$\mathbf{T}(\mathbf{u}) \in \mathbb{C}^{NM \times NM}$  is a Toeplitz Hermitian PSD matrix, with vector  $\mathbf{u}$  being its first row. Since rank minimization problem is nonconvex and NP-hard to compute; then the nonconvexity is avoided by using convex relaxation, through replacing  $\|\mathbf{Y}\|_{A,0}$  by the atomic  $l_1$  norm -denoted briefly as atomic norm-, which is identified as the gauge function of the convex hull of  $\mathbf{A}$  or  $\text{conv}(\mathbf{A})$  [12].

$$\|\mathbf{Y}\|_A = \inf\{t > 0: \mathbf{Y} \in t \cdot \text{conv}(\mathbf{A})\} = \inf_{\tilde{\boldsymbol{\theta}}, \mathbf{s}_k} \left\{ \sum_{k=1}^K \|\mathbf{s}_k\|_2 : \mathbf{Y} = \sum_{k=1}^K \mathbf{a}(\tilde{\theta}_k)\mathbf{s}_k: \tilde{\theta}_k \in [0,1] \right\} \quad (7)$$

According to ref. [15, 18],  $\|\mathbf{Y}\|_A$  has the following efficient computation SDP formulation,

$$\|\mathbf{Y}\|_A = \min_{\mathbf{X}, \mathbf{u}} \frac{1}{2\sqrt{NM}} (\text{tr}(\mathbf{T}(\mathbf{u})) + \text{tr}(\mathbf{X})) \quad (8)$$

subject to  $\begin{bmatrix} \mathbf{X} & \mathbf{Y}^H \\ \mathbf{Y} & \mathbf{T}(\mathbf{u}) \end{bmatrix} \geq 0.$

Notice the interesting directions are embedded in  $\mathbf{T}(\mathbf{u})$ . As a result, whenever an optimizer of  $\mathbf{u}$  is found, the directions can be recovered either by using Vandermonde decomposition of  $\mathbf{T}(\mathbf{u})$  [14, 15] (this decomposition is unique if  $K < NM$ ), or through conventional covariance-based subspace methods, depending on the fact that  $\mathbf{T}(\mathbf{u})$  is the data covariance of  $\mathbf{Y}$  after removing the sources' correlations.

In the presence of noise, ANM assumes given bounded energy noise. Hence, the SDP in (8) can be cast to the next regularized optimization formula,

$$\|\mathbf{Y}\|_A = \min_{\mathbf{X}, \mathbf{u}} \frac{\mu}{2\sqrt{NM}} (\text{tr}(\mathbf{T}(\mathbf{u})) + \text{tr}(\mathbf{X})) + \|\mathbf{Y}_\Omega - \mathbf{Y}_\Omega^{\text{measured}}\|_F^2 \quad (9)$$

subject to  $\begin{bmatrix} \mathbf{X} & \mathbf{Y}^H \\ \mathbf{Y} & \mathbf{T}(\mathbf{u}) \end{bmatrix} \geq 0.$

where,  $\mathbf{Y}_\Omega^{\text{measured}} \in \mathbb{C}^{(M+N-1) \times L}$  is the compressive corrupted measured matrix on CPA, and  $\mu$  is the regularization parameter required to harmonize the data fidelity and the signal sparsity. On the basis of i.i.d Gaussian noise with variance  $\sigma$ , and in the case of a single measurement vector and a sparse linear array, it has been shown by Yang and Xie [21] that  $\mu \approx \sqrt{(N+M-1)\log(NM\sigma)}$  leads to a consistent estimator.

### 4.2 Gridless covariance fitting

The key concept of a gridless CF is to manipulate the data covariance matrix structure and to make estimates in an ungrided parameterized domain. Depending on the statistical considerations set out in Section 3, and a reasonable number of snapshots  $L \geq M + N - 1$ ; in which the inverse of sample covariance matrix  $\tilde{\mathbf{R}}_\Omega = \frac{1}{L}\mathbf{Y}_\Omega\mathbf{Y}_\Omega^H$  exists with probability one. Consequently, we adopt the CF criterion [21, 32] which is given by,

$$f(\mathbf{u}, \boldsymbol{\sigma}_\Omega) = \left\| \mathbf{R}_\Omega^{-\frac{1}{2}} (\tilde{\mathbf{R}}_\Omega - \mathbf{R}_\Omega) \tilde{\mathbf{R}}_\Omega^{-\frac{1}{2}} \right\|_F^2 \quad (10)$$

$$= \text{tr}(\mathbf{R}_\Omega^{-1} \tilde{\mathbf{R}}_\Omega) + \text{tr}(\tilde{\mathbf{R}}_\Omega^{-1} \mathbf{R}_\Omega) - 2NM$$

where,  $\mathbf{R}_\Omega$  is the covariance matrix that is modeled as:

$$\mathbf{R}_\Omega = E\{\mathbf{y}_\Omega(t)\mathbf{y}_\Omega^H(t)\} \quad (11)$$

$$= \mathbf{T}_{\mathbf{u},\Omega} + \text{diag}(\boldsymbol{\sigma}_\Omega), \quad \mathbf{T}_{\mathbf{u},\Omega} \geq 0, \boldsymbol{\sigma}_\Omega > 0$$

where,  $\mathbf{T}_{\mathbf{u},\Omega} = \boldsymbol{\Gamma}_\Omega \mathbf{T}(\mathbf{u}) \boldsymbol{\Gamma}_\Omega^T = \mathbf{A}_\Omega(\tilde{\boldsymbol{\theta}}) \text{diag}(\mathbf{P}) \mathbf{A}_\Omega^H(\tilde{\boldsymbol{\theta}})$ .  $\boldsymbol{\Gamma}_\Omega \in \{0,1\}^{(M+N-1) \times NM}$  represents a selection matrix, so that the  $i^{\text{th}}$  row of  $\boldsymbol{\Gamma}_\Omega$  comprises zeros except a single 1 at the  $l_i$  position. Note that in the case of heteroscedastic noise, the variances  $\{\boldsymbol{\sigma}_\Omega\}$  are distinct. In contrary to homoscedastic noise in which  $\{\boldsymbol{\sigma}_\Omega\}$  are identical.

From Eq. (10), by applying the linear parameterization of  $\mathbf{R}_\Omega$ , then CF can be accomplished by minimizing  $f$ , which yields to the next equivalences:

$$\min_{\mathbf{u}, \{\boldsymbol{\sigma} \geq 0\}} \text{tr} \left( \tilde{\mathbf{R}}_\Omega^{\frac{1}{2}} (\mathbf{T}_{\mathbf{u},\Omega} + \text{diag}(\boldsymbol{\sigma}_\Omega))^{-1} \tilde{\mathbf{R}}_\Omega^{\frac{1}{2}} \right) \quad (12)$$

$$+ \text{tr}(\tilde{\mathbf{R}}_\Omega^{-1} \mathbf{T}_{\mathbf{u},\Omega})$$

$$+ \text{tr}(\tilde{\mathbf{R}}_\Omega^{-1} \text{diag}(\boldsymbol{\sigma}_\Omega)),$$

subject to  $\mathbf{T}_{\mathbf{u},\Omega} \geq 0$

$$\Leftrightarrow \min_{\mathbf{X}, \mathbf{u}, \{\boldsymbol{\sigma} \geq 0\}} \text{tr}(\mathbf{Z}) + \text{tr}(\tilde{\mathbf{R}}_\Omega^{-1} \mathbf{T}_{\mathbf{u},\Omega}) \quad (13)$$

$$+ \text{tr}(\tilde{\mathbf{R}}_\Omega^{-1} \text{diag}(\boldsymbol{\sigma}_\Omega)),$$

subject to  $\mathbf{T}_{\mathbf{u},\Omega} \geq 0$

and  $\mathbf{Z} \geq \tilde{\mathbf{R}}_\Omega^{\frac{1}{2}} (\mathbf{T}_{\mathbf{u},\Omega} + \text{diag}(\boldsymbol{\sigma}_\Omega))^{-1} \tilde{\mathbf{R}}_\Omega^{\frac{1}{2}}$

$$\Leftrightarrow \min_{\mathbf{X}, \mathbf{u}, \{\boldsymbol{\sigma} \geq 0\}} \text{tr}(\mathbf{Z}) + \text{tr}(\tilde{\mathbf{R}}_\Omega^{-1} \mathbf{T}_{\mathbf{u},\Omega}) \quad (14)$$

$$+ \text{tr}(\tilde{\mathbf{R}}_\Omega^{-1} \text{diag}(\boldsymbol{\sigma}_\Omega)),$$

subject to  $\begin{bmatrix} \mathbf{Z} & \tilde{\mathbf{R}}_\Omega^{\frac{1}{2}} \\ \tilde{\mathbf{R}}_\Omega^{\frac{1}{2}} & \mathbf{T}_{\mathbf{u},\Omega} + \text{diag}(\boldsymbol{\sigma}_\Omega) \end{bmatrix} \geq 0$

Minimizing  $f$  is convex, and can be formulated as SDP. Problem 14 was derived under the development of the gridless sparse iterative covariance-based estimation (GLS) method [21]. The strength of the GLS method relates to the transformation of the problem of direction estimation into the estimation of the PSD Toeplitz matrix wherein the DoAs parameters are encoded. This agrees with ANM which tends to estimate the matrix  $\mathbf{T}(\mathbf{u})$ , then extract the directions.

## 5. THE PROPOSED FRAMEWORK

Despite the impressive theoretical guarantees of the ANM system, it has substantial weaknesses, such as restricted resolution and unsatisfactory performance in the moderate SNR range. However, we believe that ANM performance might be further enhanced by combining with CF criterion and using the nonconvex penalty. In fact, the sparse continuous-field and super-resolution method (SCSM) [33] combined the CF criterion with the atomic norm of  $\mathbf{u}$  vector. However, this paper differs from SCSM in the following aspects. First, in this paper, the sparsity is enforced by means of nonconvex

relaxation of the atomic norm of signal vector  $\mathbf{Y}$ , while SCSM was based on the conventional ANM of  $\mathbf{u}$ . Second, in this paper, we seek to generalize the ANM framework to more realistic noise scenarios, while this issue has not been considered in SCSM. Regarding source localization and power estimation, SCSM used Prony's method [34], which is sensitive to measurement noise, while our work is based on the high-resolution Root-MUSIC method [35].

### 5.1 Sparsity promoting by nonconvex relaxation

Inspired by the significant advantages of the nonconvex penalties, we reformulate Problem 8 in a new relaxed ANM using the nonconvex Schatten- $p$  quasi-norm penalty. Recall the Schatten- $p$  quasi-norm of a matrix  $\mathbf{Y} \in \mathbb{C}^{NM \times L}$  is given by,

$$\|\mathbf{Y}\|_p = \left( \sum_{k=1}^{\min(NM,L)} \sigma_k^p(\mathbf{Y}) \right)^{\frac{1}{p}} \quad (15)$$

$$= \left( \text{tr}((\mathbf{Y}^T \mathbf{Y})^{\frac{p}{2}}) \right)^{\frac{1}{p}}, \quad 0 < p < 1$$

where,  $\sigma_k(\mathbf{Y})$  are the singular values of  $\mathbf{Y}$ . Noting that Schatten- $p$  quasi-norm is equal to the trace or nuclear norm when  $p = 1$ , thus, the nuclear norm is a special case of Schatten- $p$  quasi-norm. As well, if  $p \rightarrow 0$ , then  $\|\mathbf{Y}\|_p \rightarrow \text{rank}(\mathbf{Y})$ , which implies that Schatten- $p$  quasi-norm can provide a low-rank solution when  $p$  has a small value.

Next, we are introducing our method by substituting the term  $\text{tr}(\mathbf{T}(\mathbf{u}))$  in Problem 8 by its Schatten- $p$  quasi-norm, resulting in the subsequent optimization problem,

$$\min_{\mathbf{X}, \mathbf{u}} \frac{1}{2\sqrt{NM}} (\|\mathbf{T}(\mathbf{u})\|_p^p + \text{tr}(\mathbf{X})) \quad (16)$$

subject to  $\begin{bmatrix} \mathbf{X} & \mathbf{Y}^H \\ \mathbf{Y} & \mathbf{T}(\mathbf{u}) \end{bmatrix} \geq 0$ .

for,

$$\|\mathbf{T}(\mathbf{u})\|_p^p = \left( \sum_{k=1}^{NM} (\sigma_k(\mathbf{T}(\mathbf{u})) + \xi)^p \right) \quad (17)$$

where,  $\{\sigma_k \geq 0\}_1^{NM}$  are descendingly sorted singular values of  $\mathbf{T}(\mathbf{u})$ ,  $\xi > 0$  is a smoothing parameter and  $0 < p < 1$ .

Since Problem 16 has a nonconvex objective function, we have to linearize it. Thus, we use Taylor expansion for Schatten- $p$  term. In the  $l^{\text{th}}$  iteration, the expansion will be as follows:

$$\left( \sigma_k(\mathbf{T}^l(\mathbf{u})) + \xi \right)^p \quad (18)$$

$$+ \frac{p}{(\sigma_k(\mathbf{T}^l(\mathbf{u})) + \xi)^{1-p}} (\sigma_k(\mathbf{T}(\mathbf{u})) - \sigma_k(\mathbf{T}^l(\mathbf{u})))$$

As  $\mathbf{T}^l(\mathbf{u})$  is fixed, then the related terms can be excluded from the optimization problem. As a result, Problem 16 can be reformulated as:

$$\min_{\mathbf{X}, \mathbf{u}} \frac{1}{2\sqrt{NM}} (\text{tr}(\mathbf{W}^l \mathbf{T}(\mathbf{u})) + \text{tr}(\mathbf{X})) \quad (19)$$

subject to  $\begin{bmatrix} \mathbf{X} & \mathbf{Y}^H \\ \mathbf{Y} & \mathbf{T}(\mathbf{u}) \end{bmatrix} \geq 0$ .

where,  $tr(\mathbf{W}^l \mathbf{T}(\mathbf{u})) = \sum_{k=1}^{NM} (w_k^l \sigma_k(\mathbf{T}(\mathbf{u})))$  is the weighted nuclear norm of matrix  $\mathbf{T}(\mathbf{u})$ , and  $\mathbf{W}^l = \mathbf{V} \text{diag}(w_1^l, w_2^l, \dots, w_{NM}^l) \mathbf{V}^H$  is the reweighting matrix whose elements  $w_k^l = p / (\sigma_k(\mathbf{T}^l(\mathbf{u})) + \xi)^{1-p}$  are updated depending on the previous problem solution.  $\mathbf{V}$  is acquired by using the eigendecomposition of  $\mathbf{T}^l(\mathbf{u})$ .

## 5.2 The hybrid reweighted method

Motivated by the properties of the covariance fitting (CF) technique and in order to generalize the proposed ANM framework to more realistic noise scenarios, we develop a hybrid method by combining Problem 19 with CF minimization in Problem 14 as a noise metric. This leads to the following SDP problem:

$$\begin{aligned} \min_{\mathbf{u}, \mathbf{Y}, \{\sigma_{\Omega} \neq 0\}} & \frac{\alpha}{2\sqrt{NM}} tr(\mathbf{X}) + tr(\mathbf{W}_h^l \mathbf{T}(\mathbf{u})) \\ & + tr(\mathbf{Z}) + tr(\tilde{\mathbf{R}}_{\Omega}^{-1} \text{diag}(\sigma_{\Omega})) \\ \text{subject to} & \begin{bmatrix} \mathbf{X} & \mathbf{Y}^H \\ \mathbf{Y} & \mathbf{T}(\mathbf{u}) \end{bmatrix} \geq 0, \\ & \begin{bmatrix} \mathbf{Z} & \tilde{\mathbf{R}}_{\Omega}^{-1} \\ \tilde{\mathbf{R}}_{\Omega}^{-1} & \Gamma_{\Omega} \mathbf{T}(\mathbf{u}) \Gamma_{\Omega}^T + \text{diag}(\sigma_{\Omega}) \end{bmatrix} \geq 0 \end{aligned} \quad (20)$$

where,  $\mathbf{W}_h^l = \frac{\alpha}{2\sqrt{NM}} \mathbf{W}^l + \Gamma_{\Omega}^T \tilde{\mathbf{R}}_{\Omega}^{-1} \Gamma_{\Omega}$  is the hybrid reweighting matrix and  $\alpha$  is a regularization parameter that balances the sparsity and the covariance fitting criterion.

We refer to Problem 20 as denoising reweighted ANM (DRAM) method. Mainly, DRAM inherits the merits of the CF technique and promotes sparsity as well. In addition, DRAM automatically estimates the noise variance(s)  $\sigma_{\Omega}$  with the ability to handle both homoscedastic and heteroscedastic noise models, without the knowledge of noise level, which is not available in ANM.

In DRAM, we switch our focus on estimating the covariance matrix  $\mathbf{T}(\mathbf{u})$  rather than the direction parameters. Thus, when  $\mathbf{T}(\hat{\mathbf{u}})$  is acquired by solving the SDP in Problem 20 using off-the-shelf SDP solvers, the next step is to estimate the parameters  $\{\hat{\boldsymbol{\theta}}\}$ . Recognize that DRAM presumes that  $K$  is unknown, which must be specified in order to complete the final stage of the estimation. Since  $rank(\mathbf{T}(\hat{\mathbf{u}})) = K$ , then  $K$  could be specified as the number of eigenvalues of  $\mathbf{T}(\hat{\mathbf{u}})$  that are higher than a pre-assigned threshold. Consequently,  $\{\hat{\boldsymbol{\theta}}\}$  are acquired from  $\mathbf{T}(\hat{\mathbf{u}})$  either by applying the Vandermonde decomposition lemma (see, e.g., [36]) or through the conventional covariance-based subspace techniques such as Root-MUSIC. In this paper, the latter technique is applied.

The brief outline of the DRAM method for DoAs estimation is presented in following:

Method: Denoising Reweighted ANM (DRAM)	
<b>Input</b>	The compressed corrupted measurement matrix $\mathbf{Y}_{\Omega}^{measured}$ , $\alpha$ , $p$ , $\xi$ , $N$ , $M$ , $\Omega$ , the acceleration parameter $\eta$ and the maximum number of iterations $\beta$ .
<b>Output</b>	The estimated directions $\{\hat{\boldsymbol{\theta}}\}$ .
<b>Procedure</b>	1. Calculate $\tilde{\mathbf{R}}_{\Omega}^{-1}$ , and initialize $\mathbf{W}^1 = \mathbf{I}$ and $l = 1$ . 2. Repeat: – Solve Problem 20 and get $\mathbf{T}(\hat{\mathbf{u}})$ .

- 
- Calculate SVD of  $\mathbf{T}(\hat{\mathbf{u}})$  and estimate  $K$ .
  - Get  $\{\hat{\boldsymbol{\theta}}\}$  by applying Root-MUSIC( $\mathbf{T}(\hat{\mathbf{u}})$ ,  $K$ ).
  - Set  $p = \frac{p}{\eta}$  to accelerate the convergence, and  $l = l + 1$ .
  - Update the weighting matrix  $\mathbf{W}^{l+1}$ .
- 
3. Until  $\frac{\|\hat{\mathbf{u}}^l - \hat{\mathbf{u}}^{l-1}\|_F}{\|\hat{\mathbf{u}}^{l-1}\|_F} \leq 10^{-4}$  or  $l = \beta$ .
- 

## 5.3 Convergence analysis

Problem 20 consists of convex trace terms and weighted nuclear norm (WNN) term. In general, the WNN minimization problem is not convex, hence it is difficult to analyze the convergence of the proposed method. However, we observe that  $\{\sigma_k\}_1^{NM}$  are positive monotonically shrinking, so the non-descending order of the weights will be preserved during the reweighting process. According to Gu et al. [37], it was proved that WNN converges weakly to the solution if the weights are non-descendingly ordered through a weighted soft-thresholding operator. Thus, we can conclude that our methods converge weakly to the optimizer.

## 5.4 Computational issues

The underlying computational cost of DRAM involves two major sections, the first one is associated with the reweighted iterative process for  $\mathbf{T}(\hat{\mathbf{u}})$  computation, and the other is the cost of  $\{\hat{\boldsymbol{\theta}}\}$  estimation. In order to make the next discussion more relevant, we set  $I = M + N - 1$  and  $J = MN$ .

Regarding the iterative part, DRAM requires computing  $\tilde{\mathbf{R}}_{\Omega}^{-1}$  and  $\tilde{\mathbf{R}}_{\Omega}^{-1}$  for once. This, involves  $O(I^2L + I^3)$  and  $O(I^3)$  floating point operations per second (flops), respectively. Then, in each iteration, DRAM needs computing  $\mathbf{W}_h^l$  at a cost of  $O(I^3 + IJ^2 + I^2J)$  flops. Solving the SDP is performed by using, for example, the interior-point SDPT3 solver [38], which in turn needs  $O(n_1^2 n_2^{2.5})$  flops, where  $n_1$  and  $n_2$  denote the variables size and the dimension of the PSD matrices, respectively. In Problem 20,  $n_1$  is on the order of  $(J + I^2 + L^2)$  while  $n_2$  is of  $(I + J + L)$ . Then, the complexity of SDP will be  $O(L^{6.5} + I^2L^{4.5} + J^2L^{2.5})$ . Finally, applying the Root-MUSIC method to estimate  $\{\hat{\boldsymbol{\theta}}\}$  consumes up to  $O(J^3)$ .

## 6. NUMERICAL RESULTS

In this section, we evaluate the performance of the DRAM method via numerical simulations and compare the performance with the state-of-the-art methods using CVX toolbox [39] in MATLAB v.14. on a PC, with Windows 10 system and a 2 GHz CPU. In particular, we examine the impact of the suggested hybrid framework on the performance of DoA estimation in terms of resolution, and applicability in the cases of a limited number of snapshots, correlation of sources, and under heteroscedastic noise scenario.

### 6.1 Simulation setup

All the simulations are performed for uncorrelated signals of equal powers. The CPA array consists of 7-elements with  $M = 3$  and  $N = 5$  with half-wavelength inter-element spacing. Assume, both signals and noise are i.i.d. Gaussian. The mean squared error (MSE) of the DoA estimation is figured as

$\frac{1}{TK} \|\tilde{\theta} - \hat{\theta}_K\|_2^2$ , where  $\tilde{\theta}$  is the set of true normalized directions, and  $\hat{\theta}_K$  indicates the estimated ones. While,  $T$  is the number of Monte Carlo experiments, and  $K$  is the number of DoAs.  $K$  is assumed unknown and detected by estimating the largest eigenvalues of  $T(\hat{\mathbf{u}})$  according to the appropriate threshold.

As a noise metric, the worst noise power ratio (WNPR) as  $WNPR = \frac{\sigma_{max}}{\sigma_{min}}$  is used to quantify the fluctuation of the noise on the CPA array. In addition, the SNR is defined for Heteroscedastic noise and equal-power sources ( $P_1 = P_2 = \dots = P_K = P$ ) as [40],

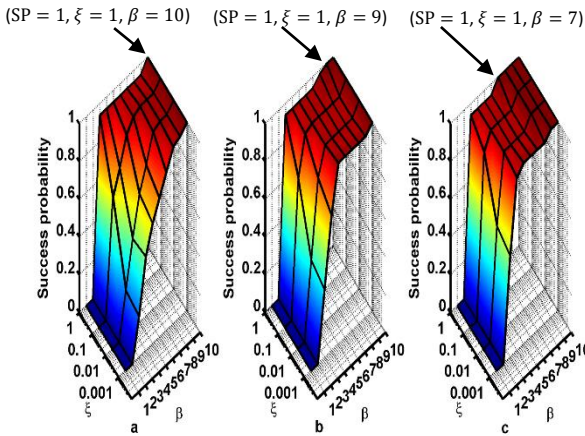
$$SNR = 10 \log_{10} \frac{P}{M+N-1} \sum_{i=1}^{M+N-1} \sigma_{\Omega_i}^{-1} \quad (21)$$

where,  $\{\sigma_{\Omega_i}\}_{i=1}^{M+N-1}$  are spatially uncorrelated. Recall, the minimum distance between any two sources  $\Delta\tilde{\theta}$  is defined as:

$$\Delta\tilde{\theta} = \min_{\tilde{\theta}_k, \tilde{\theta}_l, \tilde{\theta}_k \neq \tilde{\theta}_l} |\tilde{\theta}_k - \tilde{\theta}_l| \quad (22)$$

For well initialization of DRAM, we analyze the success probability (SP) of DRAM in terms of  $\xi \in \{10^{-3}, 10^{-2}, 10^{-1}, 1\}$ , and the number of iterations that needed to converge. Initially, we set  $p = 0.1$  since  $\mathbf{W}^l$  yields to a sparser optimizer when  $p \rightarrow 0$  according to [29]. As well, we assume  $p$  is decreased in each iteration by a factor  $\eta$ , while the iteration process stops if the maximum number of iterations  $\beta (\leq 10)$  is achieved, or if the relative variation of  $\hat{\mathbf{u}}$  at two successive iterations is less than  $10^{-4}$ . Besides, in the first iteration, a constant weighting matrix  $\mathbf{W}^1 = \mathbf{I}$  is applied. Figure 1 presents the SP of DRAM for three different choices of  $\eta \in \{2, 5, 10\}$ , considering  $\tilde{\theta} = [0.1, 0.167]$  (i.e.  $\Delta\tilde{\theta} = \frac{1}{NM}$ ), SNR = 0dB with  $\sigma_{\Omega} = [1 \ 4 \ 7.3 \ 10 \ 8 \ 6 \ 5]$  (i.e. WNPR = 10) and  $L = 15$  snapshots. While, the regularization parameter is set as  $\alpha = 2 \times 10^{-2} \sqrt{NM}$ , which gives a good performance empirically.

As shown in Figure 1, SP will be greater than 0.9 for  $(\xi = 1, \beta = 3)$  regardless of  $\eta$ . For the best estimation performance (SP=1), setting  $\eta = 10$  and  $\xi = 1$  will lead to the fastest convergence speed ( $\beta = 7$ ) when compared to the other choices.



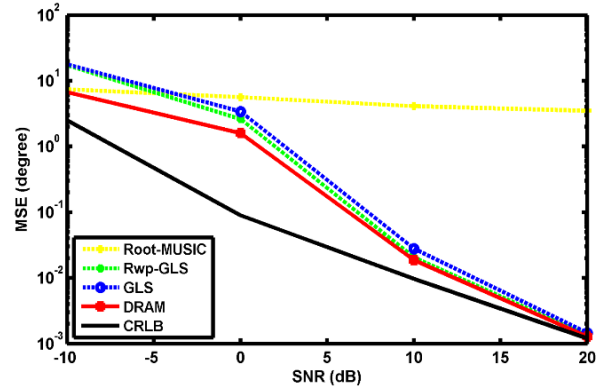
**Figure 1.** Success probability with the number of iterations  $\beta$  and  $\xi$  for: (a)  $\eta = 2$ , (b)  $\eta = 5$ , (c)  $\eta = 10$

Accordingly, in the subsequent simulations, the results are evaluated after 100 Monte Carlo experiments, taking into account the following settings:  $p = 0.1$ ,  $\xi = 1$ ,  $\eta = 10$ , and  $L = 15$  snapshots, unless otherwise is mentioned.

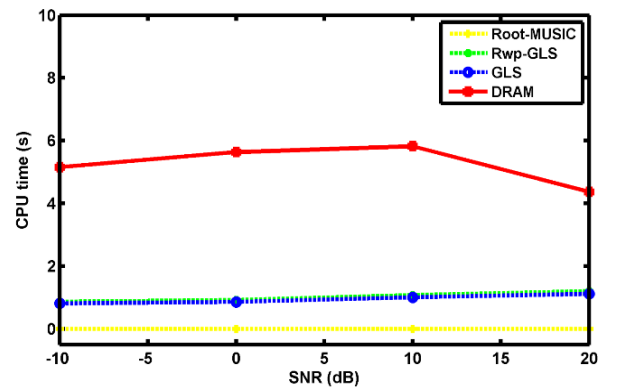
## 6.2 Comparisons with prior arts

We examine and compare the DoA estimation performance of the DRAM with the high-resolution Root-MUSIC method, and to the two remarkable gridless methods, GLS [21] and Rwp-GLS [29], since they adopt the same CF criterion as that used in DRAM. Moreover, GLS has theoretically been shown to be equivalent to weighted ANM, thus we expect to have competitive performances in the presence of finite snapshots and correlated sources. In order to provide a benchmark for the evaluation, the stochastic Cramer-Rao lower bound (CRLB) for heteroscedastic noise [41] is incorporated. All over the following simulations, we set WNPR = 10 and  $L = 15$  unless otherwise is mentioned.

We first test the performance of DRAM versus SNR for  $\sigma_{\Omega} = [1, 4, 7.3, 10, 8, 3.5, 6]$ . Consider three uncorrelated sources imping on CPA from  $\tilde{\theta} = [0.1, 0.167, 0.3]$  with SNR ranging from -10 dB to 20 dB. The results are shown in Figure 2. Note that the first two DoAs are separated by only  $\Delta\tilde{\theta} = \frac{1}{NM}$ . As it is shown, DRAM has the lowest MSE among all methods, which indicates its superiority in separating spatially adjacent signals under violated separation conditions and heteroscedastic noise cases.



**Figure 2.** MSEs of Root-MUSIC, Rwp-GLS, GLS, DRAM, and CRLB with respect to SNR



**Figure 3.** CPU times comparisons of Root-MUSIC, Rwp-GLS, GLS and DRAM with respect to SNR.

Figure 3 demonstrates the CPU times of the experiments related to Figure 2. DRAM is the most time-consuming due to the SVD decomposition in each iteration. This drawback is balanced by the estimation accuracy. Particularly, the computational complexity increases slightly till SNR = 10 dB, because of the  $\sigma_\Omega$  estimation, and when SNR > 10 dB, DRAM converges more rapidly, leading to a decrease in the number of iterations and the time usage.

We then evaluate the MSE with respect to  $15 \leq L \leq 50$ . The simulation is carried out assuming  $\tilde{\theta} = [0.1, 0.167, 0.3]$ , SNR = 10 dB and  $\sigma_\Omega = [8, 9, 2, 5, 7, 4, 10]$ . The results are shown in Figure 4. It can be concluded that DRAM performs the best throughout the examined range, and that only a few snapshots are required to resolve the sources with  $\Delta\tilde{\theta} = \frac{1}{NM}$  precision. While Root-MUSIC has the highest snapshot threshold.

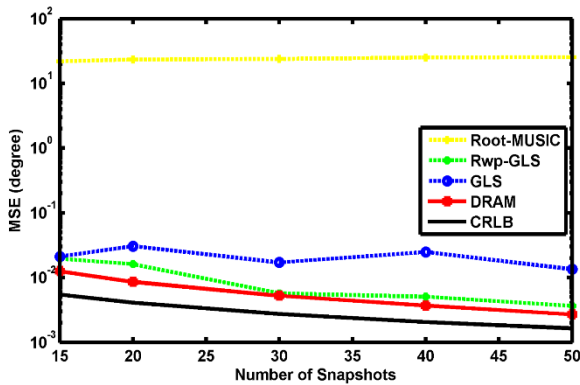


Figure 4. MSEs of Root-MUSIC, Rwp-GLS, GLS, DRAM, and CRLB with respect to the number of snapshots

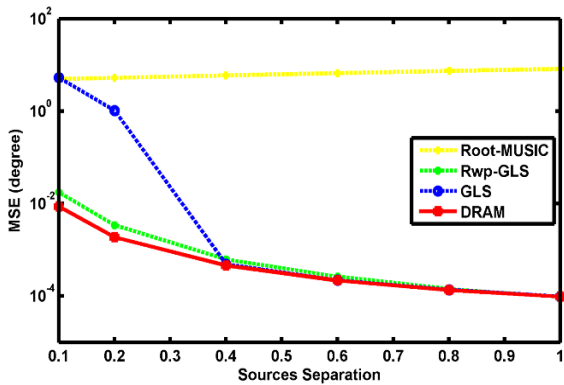


Figure 5. MSEs of Root-MUSIC, Rwp-GLS, GLS and DRAM with respect to the sources separations

To evaluate the performance of DRAM in terms of resolution, we test its ability in resolving closely spaced directions. Consider two signals impinge onto the CPA from  $\tilde{\theta} = [0.1, 0.1 + \Delta\tilde{\theta}]$  with  $\Delta\tilde{\theta} \in [\frac{0.1}{NM}, \frac{1}{NM}]$ , this corresponds to only  $0.382^\circ$  to  $3.8^\circ$  difference in angle. Our simulation results are presented in Figure 5 for SNR = 10 dB and  $\sigma_\Omega = [8, 9, 2, 5, 7, 4, 10]$ . The results indicate the superiority of DRAM over other methods in isolating spatially adjacent signals as close as  $\frac{0.1}{NM}$  and thus exhibit super-resolution capability in DoA estimation.

In addition, to verify the robustness of the DRAM under heteroscedastic noise, the MSEs versus the WNPR are tested

in Figure 6, in which WNPR ranges from 20 to 200, where SNR  $\in \{0, 10, 20\}$  dB and  $\tilde{\theta} = [0.1, 0.3]$ . It can be seen that for SNR = 0 dB, the MSEs of both GLS and Rwp-GLS get larger than CRLB. In contrast to the MSEs of DRAM which coincides with CRLB for all values of SNR over the considered range of WNPR. This illustrates again the robustness of DRAM against the heteroscedastic noise.

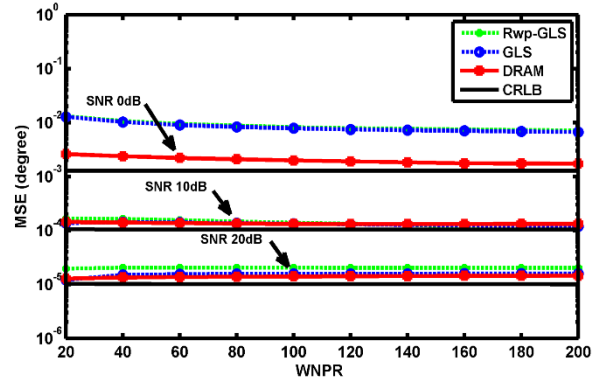


Figure 6. MSEs of Rwp-GLS, GLS, DRAM and CRLB with respect to WNPR

In the following, we provide an example to verify the ability of DRAM in resolving DoAs of correlated sources. Figure 7 shows the MSEs versus the correlation factor  $\rho$ , when SNR=10dB,  $\sigma_\Omega = [8, 9, 2, 5, 7, 4, 10]$  and  $\tilde{\theta} = [0.1, 0.167]$ . As shown, when the correlation factor increases DRAM shows immunity to the integrated correlations among sources and gains the best results.

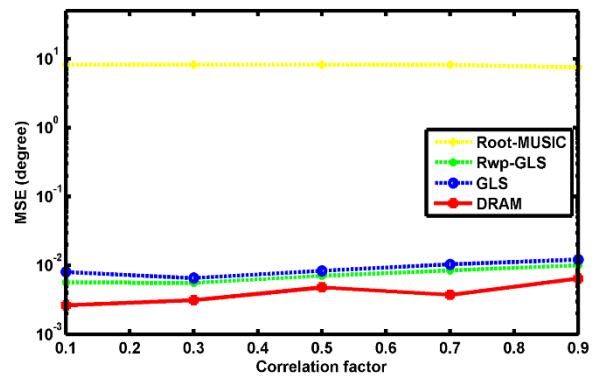


Figure 7. MSEs of Root-MUSIC, Rwp-GLS, GLS and DRAM with respect to correlation factor

## 7. CONCLUSION

The problem of sparse gridless DoA estimation from compressive corrupted measurements was addressed in this paper. We present a hybrid optimization framework based on introducing CF criterion and nonconvex relaxation into the ANM approach. Then, an iterative reweighted minimization method is developed.

The effectiveness of the proposed method is verified through numerical simulations. The results have demonstrated the superiority of DRAM in isolating spatially adjacent signals as close as  $\Delta\tilde{\theta} = \frac{0.1}{NM}$ . As well, the results have shown that DRAM has the lowest snapshot threshold, and the best

immunity against sources correlation and noise fluctuation among the compared methods. In the future, we may try more computationally efficient combining approach for noisy gridless DoA.

## REFERENCES

- [1] Swindlehurst, A.L., Jeffs, B.D., Seco-Granados, G., Li, J. (2014). Academic Press Library in Signal Processing: Volume 3 - Array and Statistical Signal Processing. Academic Press Library in Signal Processing. <https://doi.org/10.1016/c2013-0-00294-0>
- [2] Schmidt, R.O. (1981). A signal subspace approach to multiple emitter location and spectral estimation. PhD Thesis, Stanford University, California. Schmidt, R.
- [3] Roy, R., Kailath, T. (2009). ESPRIT-Estimation of signal parameters via rotational invariance techniques. *Adaptive Antennas for Wireless Communications*, 37(7): 984-995. <https://doi.org/10.1109/9780470544075.ch2>
- [4] Donoho, D.L. (2006). Compressed sensing. *IEEE Transactions on Information Theory*, 52(4): 1289-306. <https://doi.org/10.1109/TIT.2006.871582>
- [5] Baraniuk, R.R.G. (2007). Compressive sensing. *Signal Processing Magazine*, 24(4). <https://doi.org/10.1109/msp.2007.4286571>
- [6] Chae, D.H., Sadeghi, P., Kennedy, R.A. (2010). Effects of basis-mismatch in compressive sampling of continuous sinusoidal signals. *Proceedings of the 2010 2nd International Conference on Future Computer and Communication, ICFCC 2010*, pp. V2-739. <https://doi.org/10.1109/icfcc.2010.5497605>
- [7] Chi, Y., Scharf, L.L., Pezeshki, A., Calderbank, A.R. (2011). Sensitivity to basis mismatch in compressed sensing. *IEEE Transactions on Signal Processing*, 59(5): 2182-2195. <https://doi.org/10.1109/TSP.2011.2112650>
- [8] Li, Y., Chi, Y. (2016). Off-the-grid line spectrum denoising and estimation with multiple measurement vectors. *IEEE Transactions on Signal Processing*, <https://doi.org/10.1109/TSP.2015.2496294>
- [9] Tan, Z., Nehorai, A. (2014). Sparse direction of arrival estimation using co-prime arrays with off-grid targets. *IEEE Signal Processing Letters*. <https://doi.org/10.1109/LSP.2013.2289740>
- [10] Pal, P., Vaidyanathan, P.P. (2014). A grid-less approach to underdetermined direction of arrival estimation via low rank matrix denoising. *IEEE Signal Processing Letters*, 21(6): 737-741. <https://doi.org/10.1109/LSP.2014.2314175>
- [11] Yang, Z., Xie, L. (2017). On gridless sparse methods for multi-snapshot direction of arrival estimation. *Circuits, Systems, and Signal Processing*, 36(8): 3370-3384. <https://doi.org/10.1007/s00034-016-0462-9>
- [12] Chandrasekaran, V., Recht, B., Parrilo, P.A., Willsky, A.S. (2012). The convex geometry of linear inverse problems. *Foundations of Computational Mathematics*, 12(6): 805-849. <https://doi.org/10.1007/s10208-012-9135-7>
- [13] Candès, E.J., Fernandez-Granda, C. (2014). Towards a mathematical theory of super-resolution. *Communications on Pure and Applied Mathematics*, 67(6): 906-956. <https://doi.org/10.1002/cpa.21455>
- [14] Fernandez-Granda, C. (2015). Super-resolution of point sources via convex programming. 2015 IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, CAMSAP 2015, pp. 41-44. <https://doi.org/10.1109/camsap.2015.7383731>
- [15] Tang, G., Bhaskar, B.N., Shah, P., Recht, B. (2013). Compressed sensing off the grid. *IEEE Transactions on Information Theory*, 59(11): 7465-7490. <https://doi.org/10.1109/TIT.2013.2277451>
- [16] Candès, E.J., Fernandez-Granda, C. (2013). Super-Resolution from Noisy Data. *Journal of Fourier Analysis and Applications*. <https://doi.org/10.1007/s00041-013-9292-3>
- [17] Bhaskar, B.N., Tang, G., Recht, B. (2013). Atomic norm denoising with applications to line spectral estimation. *IEEE Transactions on Signal Processing*, 61(23): 5987-5999. <https://doi.org/10.1109/TSP.2013.2273443>
- [18] Yang, Z., Xie, L. (2016). Exact joint sparse frequency recovery via optimization methods. *IEEE Transactions on Signal Processing*, 64(19): 5145-5157. <https://doi.org/10.1109/TSP.2016.2576422>
- [19] Tang, G., Bhaskar, B.N., Recht, B. (2015). Near minimax line spectral estimation. *IEEE Transactions on Information Theory*. <https://doi.org/10.1109/TIT.2014.2368122>
- [20] Li, Q., Tang, G. (2020). Approximate support recovery of atomic line spectral estimation: A tale of resolution and precision. *Applied and Computational Harmonic Analysis*. <https://doi.org/10.1016/j.acha.2018.09.005>
- [21] Yang, Z., Xie, L. (2015). On gridless sparse methods for line spectral estimation from complete and incomplete data. *IEEE Transactions on Signal Processing*, 63(12): 3139-3153. <https://doi.org/10.1109/TSP.2015.2420541>
- [22] Wu, X., Zhu, W.P., Yan, J., (2017). A toeplitz covariance matrix reconstruction approach for direction-of-arrival estimation. *IEEE Transactions on Vehicular Technology*, pp. 8223-8237. <https://doi.org/10.1109/tvt.2017.2695226>
- [23] Wu, X., Zhu, W.P., Yan, J. (2018). A high-resolution DOA estimation method with a family of nonconvex penalties. *IEEE Transactions on Vehicular Technology*, 67(6): 4925-4938. <https://doi.org/10.1109/TVT.2018.2817638>
- [24] Yang, Z., Xie, L. (2016). Enhancing sparsity and resolution via reweighted atomic norm minimization. *IEEE Transactions on Signal Processing*, 64(4): 995-1006. <https://doi.org/10.1109/TSP.2015.2493987>
- [25] Zien Alabideen, L., Al-Dakkak, O., Khorzom, K. (2019). Frequency recovery resolution enhancement using generalized reweighted nonconvex relaxation. *U.P.B. Sci. Bull., Series C*, 81(1): 81-98.
- [26] Tan, W., Feng, X. (2019). Covariance matrix reconstruction for direction finding with nested arrays using iterative reweighted nuclear norm minimization. *International Journal of Antennas and Propagation* 2019. <https://doi.org/10.1155/2019/7657898>
- [27] Lu, C., Tang, J., Yan, S., Lin, Z. (2014). Generalized nonconvex nonsmooth low-rank minimization. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 4130-4137. <https://doi.org/10.1109/cvpr.2014.526>
- [28] Malek-Mohammadi, M., Babaie-Zadeh, M., Skoglund, M. (2015). Performance guarantees for Schatten-p quasi-norm minimization in recovery of low-rank matrices. *Signal Processing*, 114(C): 225-230. <https://doi.org/10.1016/j.sigpro.2015.02.025>



- [29] Alabideen, L.Z., Al-Dakkak, O., Khorzom, K. (2020). Reweighted covariance fitting based on nonconvex Schatten-p minimization for gridless direction of arrival estimation. *Mathematical Problems in Engineering*, 2020: 1–11. <https://doi.org/10.1155/2020/3012952>
- [30] Zhou, C., Gu, Y., Zhang, Y.D., Shi, Z., Jin, T., Wu, X. (2017). Compressive sensing-based coprime array direction-of-arrival estimation. *IET Communications*, 11(11): 7465-7490. <https://doi.org/10.1049/iet-com.2016.1048>
- [31] Tan, Z., Eldar, Y.C., Nehorai, A. (2014). Direction of arrival estimation using co-prime arrays: A super resolution viewpoint. *IEEE Transactions on Signal Processing*, 62(21): 5565-5576. <https://doi.org/10.1109/TSP.2014.2354316>
- [32] Stoica, P., Babu, P., Li, J. (2011). New method of sparse parameter estimation in separable models and its use for spectral analysis of irregularly sampled data. *IEEE Transactions on Signal Processing*. <https://doi.org/10.1109/TSP.2010.2086452>
- [33] Lin, B., Zhu, J. (2017). Sparse continuous-field and super-resolution method for direction-of-arrival estimation. *Multidimensional Systems and Signal Processing*. <https://doi.org/10.1007/s11045-016-0451-y>
- [34] Blu, T., Dragotti, P.L., Vetterli, M., Marziliano, P., Coulot, L. (2008). Sparse sampling of signal innovations. *IEEE Signal Processing Magazine*. <https://doi.org/10.1109/msp.2007.914998>
- [35] Barabell, A.J. (1983). Improving the resolution performance of eigenstructure-based direction-finding algorithms. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pp. 336-339. <https://doi.org/10.1109/icassp.1983.1172124>
- [36] Yang, Z., Li, J., Stoica, P., Xie, L. (2017). Sparse methods for direction-of-arrival estimation. *Academic Press Library in Signal Processing, Volume 7: Array, Radar and Communications Engineering*, pp. 509-581. <https://doi.org/10.1016/b978-0-12-811887-0.00011-0>
- [37] Gu, S., Xie, Q., Meng, D., Zuo, W., Feng, X., Zhang, L. (2017). Weighted nuclear norm minimization and its applications to low level vision. *International Journal of Computer Vision*, 121(2): 183-208. <https://doi.org/10.1007/s11263-016-0930-5>
- [38] Toh, K.C., Todd, M.J., Tütüncü, R.H. (1999). SDPT3 - a MATLAB software package for semidefinite programming, version 1.3. *Optimization Methods and Software*. <https://doi.org/10.1080/10556789908805762>
- [39] Grant, M., Boyd, S. (2017). CVX: Matlab software for disciplined convex programming, ver 2.1. Available at [Http://Cvxr.Com/Cvx/](http://cvxr.com/cvx/).
- [40] Pesavento, M., Gershman, A.B. (2001). Maximum-likelihood direction-of-arrival estimation in the presence of unknown nonuniform noise. *IEEE Transactions on Signal Processing*, 49(7): 1310-1324. <https://doi.org/10.1109/78.928686>
- [41] He, Z.Q., Shi, Z.P., Huang, L. (2014). Covariance sparsity-aware DOA estimation for nonuniform noise. *Digital Signal Processing: A Review Journal*, 28(1): 75-81. <https://doi.org/10.1016/j.dsp.2014.02.013>