Underdetermined DOA Estimation for Wideband Signals Using Robust Sparse Covariance Fitting

Zhen-Qing He, Student Member, IEEE, Zhi-Ping Shi, Member, IEEE, Lei Huang, Member, IEEE, and Hing Cheung So, Senior Member, IEEE

Abstract—From the co-array perspective, sparse spatial sampling can significantly increase the degrees-of-freedom (DOFs), enabling us to perform underdetermined direction-of-arrival (DOA) estimation. By leveraging the increased DOFs from the sparse spatial sampling, we develop a new underdetermined DOA estimation method for wideband signals, named wideband sparse spectrum fitting (W-SpSF) estimator. In W-SpSF, we formulate a sparse reconstruction problem that includes a quadratic \( \{ \mathbf{e}_2 \} \) weighted covariance fitting term added to a sparsity-promoting \( (\mathbf{e}_{2,1}) \) regularizer. Meanwhile, the optimal regularization parameter of W-SpSF is studied to ensure robust sparse recovery. Numerical results enabled nested arrays demonstrate that the W-SpSF estimator outperforms the spatial smoothing based MUSIC algorithm and works well in nonuniform noise environment.

Index Terms—Co-array, direction-of-arrival (DOA) estimation, sparse spectrum fitting (SpSF), wideband signal.

I. INTRODUCTION

The topic of target localization with less sensors than sources, i.e., underdetermined direction-of-arrival (DOA) estimation, has been receiving considerable interest in recent years [1]–[9]. An effective approach for this problem is to construct a new virtual array with a higher degrees-of-freedom (DOFs) than that obtained from the physical array. From the co-array perspective, sparse spatial sampling can provide a significant improvement in DOFs, and typical schemes of which include nested arrays [2] and co-prime arrays [3], [4]. Indeed, since the nested/co-prime arrays share a basic co-array, their combinations with the standard MUSIC [10] have been extended to various scenarios, e.g., 2\(q\)th-order cumulant statistics [5], two-dimensional arrays [6], [7], vector-sensor arrays [8] and wideband (WB) signals [9]. Most of them, however, resort to the spatial smoothing technique for establishing a rank-aware covariance matrix, thereby resulting in some aperture loss.

More recently, by vectorizing the covariance matrix, a series of sparse signal recovery (SSR) based approaches [11]–[18] have also been devised for DOA estimation, which have their roots in the sparse spectrum/covariance fitting criterion. These methods, explicitly or implicitly, become aware of the increased DOFs from co-array. For instance, to tackle the underdetermined case, the SSR based technique is tailored in [14] for nested arrays and in [15]–[18] for co-prime arrays, respectively. However, these techniques (albeit without aperture loss) concern only the narrowband signals and the literature for WB signals is less abundant. Although WB DOA estimation using SSR is not new (see, e.g., [19]–[21]), previous works seldom address the underdetermined case from the difference co-array perspective.

In this letter, inspired by the increased DOFs from the difference co-array in sparse spatial sampling, we reformulate the underdetermined DOA estimation problem for WB signals in the SSR framework. We first devise a unified Gaussian-noise-aware sparse model by prewhitening the perturbed errors resulting from finite samples (snapshots). Then an approach called Wideband Sparse Spectrum Fitting (W-SpSF) is developed using \( \ell_{2,1} \)-norm sparse regularization plus a weighted covariance fitting (least squares) criterion. In the W-SpSF estimator, the choice of the regularization parameter is also discussed to ensure robust sparse recovery. The W-SpSF approach has many merits such as without a priori knowledge on the source number, suppressing the spatial aliasing and being applicable to nonuniform noise. Simulation results corroborate the effectiveness of the W-SpSF.

II. SIGNAL MODEL AND PROBLEM STATEMENT

Consider a linear array of \( M \) sensors whose locations are given by the set \( \mathcal{D} = \{ d_m : m \in [M] \}, [M] \triangleq \{1, \ldots, M\}. \) And suppose \( Q \) (possibly \( Q \geq M \)) WB signals impinging on this array from the far-field directions \( \mathbf{\theta} = [\theta_1, \ldots, \theta_Q]^T \) where \( (\cdot)^T \) is the transpose. Each sensor signal after time-sampling is partitioned into \( L \) segments, and a \( K \)-point discrete Fourier transform (DFT) is applied to each segment. The WB array output can be modeled as [22], [23]:

\[
\mathbf{x}_l(f_k) = \mathbf{A}_k(\mathbf{\theta}) \mathbf{s}_l(f_k) + \mathbf{n}_l(f_k), \quad k \in [K], \quad l \in [L] \tag{1}
\]

where \( \mathbf{x}_l(f_k) \in \mathbb{C}^M, \mathbf{s}_l(f_k) \in \mathbb{C}^Q \) and \( \mathbf{n}_l(f_k) \in \mathbb{C}^M \) are the DFT coefficients of the received data, source signals and additive noise, respectively, \( \mathbf{A}_k(\mathbf{\theta}) \triangleq A(f_k(\mathbf{\theta})) = [a_k(\theta_1), \ldots, a_k(\theta_Q)] \in \mathbb{C}^{M \times Q} \).
is the steering matrix at the frequency bin $f_k$ with the $M \times 1$ steering vector

$$a_k(\theta) = \left[ e^{-j2\pi f_k d_1 \sin \theta}, \ldots, e^{-j2\pi f_k d_M \sin \theta} \right]^T$$

where $j = \sqrt{-1}$, and $c$ is the propagation speed.

The signals and noises, within different frequency bins, are assumed to be complex circular Gaussian distributed and independent of each other [9]. Then the $k$th narrowband output covariance matrix is

$$R_k = \mathbb{E} \{ s_k(f_k)s_k^H(f_k) \} = A_k(\theta)\mathbb{R}_k A_k^H(\theta) + \mathbb{Q}_k$$

where $\mathbb{E}\{\cdot\}$ and $(\cdot)^H$ represent the expectation operator and conjugate transpose, respectively, $\mathbb{R}_k = \mathbb{E} \{ s_k(\theta)s_k^H(\theta) \} = \text{diag}(\rho_k)$ is the $k$th signal covariance matrix with $\rho_k = [\rho_k^1, \ldots, \rho_k^m]^T$ being the source power vector, and $\mathbb{Q}_k = \mathbb{E} \{ n_k(\theta)n_k^H(\theta) \} = \text{diag}(\sigma_k)$ is the $k$th noise covariance matrix with $\sigma_k = [\sigma_k^1, \ldots, \sigma_k^m]^T$ being the noise power vector. Vectorizing (3), we get

$$\nu_k \doteq \text{vec}(R_k) - A_k(\theta)\rho_k + 1_k, k \in [K]$$

where $1_k = [\sigma_k^1, e^1, \ldots, \sigma_k^m, e^m]^T$, $e_m$ is the $m$th vector in the canonical basis of $\mathbb{C}^M$, $A_k(\theta) \doteq A_k^1(\theta) \otimes A_k^2(\theta) = [a_k^1(\theta) \otimes a_k(\theta), \ldots, a_k^m(\theta) \otimes a_k(\theta)] \in \mathbb{C}^{M^2 \times 2}$ where $\otimes$ and $\odot$ represent the complex conjugate, Khatri-Rao product and Kronecker product, respectively. The increased DOFs directly rely on the distinct rows of $A_k(\theta)$, which is equivalent to the cardinality of the location set, $\mathcal{D}_k = \{ d_i - d_j : i, j \in [M] \}$, of the difference co-array [2]. The problem at hand is to determine $\theta$ without the need of the knowledge on the source number $Q$ and noise variances $\{\sigma_k\}_{k=1}^K$.

III. WIDEBAND DOA ESTIMATION USING ROBUST SPARSE COVARIANCE FITTING

A. Noise Reduction and Prewhitening

Note that the noise-like component $1_k$ in (4) consists of the unknown terms $\{\sigma_k^m\}_{m=1}^M$ and 0 otherwise. In fact, the indices of $\sigma_k^m$ in $1_k$ correspond exactly with the zero locations of the difference co-array. Furthermore, these zero locations have occurrences (i.e., redundancy) and inherently lead to only one DOF instead of $M$ DOFs. To mitigate the effect of $1_k$, we therefore eliminate the $M$ entries of $\nu_k$ indexed by the positions of $\{\sigma_k^m\}_{m=1}^M$ within $1_k$. This can be achieved by the following linear operation:

$$\tilde{\nu}_k = J\nu_k - B_k(\theta)\rho_k + 1_k, k \in [K]$$

where $J = [J_1, \ldots, J_{M-1}]^T \in \mathbb{R}^{M(M-1) \times M^2}$, $J_m = [E_m(\tau), \ldots, E_m(M+1)] \in \mathbb{R}^{m \times M^2}$ with $E_m = m[1, \ldots, m] - M - 1, \ldots, m(M+1)$, being the $m$th vector in the canonical basis of $\mathbb{C}^M$, and $B_k(\theta) = J A_k(\theta) \in \mathbb{C}^{M^2 \times 2}$ is the virtual steering matrix.

Having the $L$ available segments (frequency snapshots), $\nu_k$ is estimated by $\tilde{\nu}_k = \nu_k + \Delta \nu_k = \text{vec}(R_k)$ where $R_k = \frac{1}{L} \sum_{l=1}^L s_l(f_l)s_l^H(f_l)$ and $\Delta \nu_k = \nu_k - \nu_k$ is the estimation error. Accordingly, we have

$$\tilde{\nu}_k = J\tilde{\nu}_k + \Delta \nu_k$$

where $\Delta \nu_k = J\Delta \nu_k$ is the corresponding estimation error of $\nu_k$. Actually, $\Delta \nu_k$ is asymptotically (in $L$) complex normal ($CN$) distributed (see, e.g., [21], [24], [25]), viz.,

$$\Delta \nu_k \sim CN(0, \frac{1}{L} R_k \otimes R_k)$$

which further leads to

$$\Delta \nu_k \sim CN(0, W_k), W_k \doteq \frac{1}{L} J (R_k \otimes R_k) J^T$$

where $0$ is a zero vector/matrix with proper dimension which, for simplicity, is determined from the context. Therefore, the error $\Delta \nu_k$ in (6) can be, at least approximately, viewed as an additive spatially correlated (colored) Gaussian noise. To further alleviate this correlation, a prewhitening procedure is considered here, that is,

$$\tilde{\nu}_k = W_k^{-\frac{1}{2}} \tilde{\nu}_k = W_k^{-\frac{1}{2}} B_k(\theta) \rho_k + \epsilon_k, k \in [K]$$

where $\epsilon_k \doteq W_k^{-\frac{1}{2}} \Delta \nu_k \sim CN(0, I_{M(M-1)})$, and $I_{M(M-1)}$ is the $M(M-1) \times M(M-1)$ identity matrix.

The new observation (8), reminiscent of the initial model (1), can be viewed as a long received model with standard Gaussian noise. In the following, we will adopt (8) to address the problem of WB DOA estimation because compared to the ideal model (4), it is not corrupted by the noise variances anymore, thereby allowing easy access to sparse modeling without introducing additional parameters.

B. Wideband Sparse Spectrum Fitting

Following sparse localization framework (see, e.g., [12]–[21]), all $K$ vectors $\{\tilde{\nu}_k\}_{k=1}^K$ in (8) can be readily rewritten as a unified noise-aware sparse model

$$\tilde{\nu} = \Phi(\Omega)p^\nu + \epsilon$$

where $\epsilon \doteq W_k^{-\frac{1}{2}} \tilde{\nu}_k \sim CN(0, I_{M(M-1)})$, and $I_{M(M-1)}$ is the $M(M-1) \times M(M-1)$ identity matrix. The above vectorial representation is the overcomplete basis in which each column vector is normalized by its norm, $\Omega \doteq \{\Omega_i : i \in [N']\}$ is the sampling grid set covering all the potential direction domain, and $P^\nu = \{p^\nu_1, \ldots, p^\nu_K\}$ where $p^\nu_k \in \mathbb{R}^N$ is the $k$th narrowband sparse spatial spectrum. When $N \gg Q$ such that $\{\theta_k \in [Q]\} \subset \Omega$, $\{p^\nu_k\}_{k=1}^K$ share the same sparse support (i.e., nonzero index) and hence $P^\nu \doteq \{p^\nu_1, \ldots, p^\nu_K\}$ is a joint row $Q$-sparse matrix. That is, the sparsity of $P^\nu$ (or $P^\nu$) behaves only in the spatial domain instead of the frequency domain.

This motivates us to employ the prevalent $\ell_2,1$-norm minimization formulation for joint sparse recovery [19], [26]–[28]:

$$\min_{p^\nu} \frac{1}{2} \| \tilde{\nu} - \Phi(\Omega)p^\nu \|_2^2 + \tau \| p^\nu_2,1, p \geq 0$$

where $\| p \|_2,1 \doteq \sum_{i=1}^N \| p_i \|_2$ is the $\ell_2,1$-norm of $p$, $p^\nu = \{p^\nu_1, \ldots, p^\nu_K\}$ \in $\mathbb{R}^{N \times K}$ is the $t$th entry of $p^\nu$, $\tau$ denotes $\geq$ with an elementwise operation, and $\tau$ is a nonnegative regularization parameter. The convex formulation (10), also referred to as the Wideband Sparse Spectrum Fitting (W-SpSF) estimator, can be solved by interior-point algorithm based software packages such as SeDuMi [29] and CVX [30], with a complexity of $\mathcal{O}(NK^3)$. Actually, the quadratic term of (10), i.e., $\frac{1}{2} \| \tilde{\nu} - \Phi(\Omega)p^\nu \|_2^2$, is essentially a weighted least squares fitting criterion which can

1Here we use its rudiment from the narrowband SpSF estimator [13].
induce an asymptotically unbiased estimator [31] and hence is desirable to debias the solution. The sparse-inducing solution of (10), however, is highly sensitive to \( \tau \) as this parameter balances the sparsity and the data fidelity (least squares term) and its poor choice may lead to a non-sparse solution. Therefore, an appropriate selection of \( \tau \) is required to guarantee robust sparse recovery, which will be investigated in the next subsection.

**Remark on Parameter Identifiability:** Based on Corollary 1 of [13], the maximum number of sources, which can be uniquely identified by W-SpSF, is \( N_{\text{id}} = \frac{N^2}{M} \) if any \( N \) columns of \( \Theta \) are linearly independent (assume here that \( \tau \) has been well chosen to promote sparse solution). With no ambiguity in \( \Phi(\Omega) \), we have \( N = K \cdot DOF_A \) where \( DOF_A \) denotes the DOFs of \( A_\alpha(\theta) \). This indicates that more dimensions mean better identifiability, which agrees with our intuition. As a result, the proposed W-SpSF estimator is able to accommodate the enhanced DOFs from arbitrary array geometries (e.g., nested arrays [2], co-prime arrays [3], [4] and spatial compressive sensing based random arrays [32]) to improve its parameter identifiability considerably.

### C. Regularization Parameter Selection

Here we analyze how to choose a proper \( \tau \) via duality theory. Unlike [26], [27] which determine an appropriate choice of \( \tau \) by revisiting the minimization (10) from its dual maximization, we address the selection of \( \tau \) directly via the Karush-Kuhn-Tucker (KKT) conditions [33], [34], under an additional non-negative constraint. We start with the real-valued reformulation of problem (10):

\[
\begin{align*}
\min_{p} \frac{1}{2} \| \tilde{z} \|_2^2 + \tau \sum_{i=1}^{N} \| p^i \|_2 \quad & \text{ s.t. } \quad z = \tilde{z} - \Phi(\Omega) p, \quad p \succeq 0 \tag{11}
\end{align*}
\]

where \( \Phi(\Omega) = [R(\Phi^T(\Omega)), \text{Im}(\Phi^T(\Omega))]^T \), \( \tilde{z} = [R(\tilde{y}), \text{Im}(\tilde{y})]^T \), \( R(\cdot) \) and \( \text{Im}(\cdot) \) represent the real and imaginary parts, respectively. The associated Lagrangian is

\[
\begin{align*}
\frac{1}{2} \| z \|_2^2 + \gamma \sum_{i=1}^{N} \| p^i \|_2 + \mu^T (z - \tilde{z} - \Phi(\Omega) p) - \lambda^T p \\
- \frac{1}{2} \| z \|_2^2 + \mu^T (z - \tilde{z} - \Phi(\Omega) p) + \sum_{i=1}^{N} h_i(p^i)
\end{align*}
\]

where \( \mu \) and \( \lambda \) are vectors of Lagrangian multipliers associated with \( z = \tilde{z} - \Phi(\Omega) p \) and \( p \succeq 0 \), respectively, \( h_i(p^i) \Delta [\| p^i \|_2 + (\eta^i - \lambda^i)^T p^i] \). \( \eta^i \Delta \Phi(\Omega)^T(\Omega) \mu^i \) and \( \lambda^i \Delta \Phi(\Omega) \) are sub-vector and sub-matrix extracted, respectively, from \( \lambda \) and columns of \( \Phi(\Omega) \), corresponding to \( p^i \) within \( p \).

By the KKT conditions [33], [34], we get

\[
\begin{align*}
z &= -\mu, \tag{12a} \\
0 &= \partial h_i(p^i); i \in [N], \tag{12b} \\
z &= \tilde{z} - \Phi(\Omega) p, \tag{12c} \\
p^i \succeq 0; \lambda^i \geq 0, \tag{12d}
\end{align*}
\]

where \( \partial h_i(p^i) \) denotes the subgradient set of \( h_i \) at \( p^i \) (see [34] for a definition). \( p^i \) and \( \lambda^i \) represent the \( i \)th element of \( p^i \) and \( \lambda^i \), respectively. The condition (12a) is due to the gradient operation: \( \nabla L(z) = 0 \). In addition, \( 0 \in \partial h_i(p^i) \), means that \( h_i(p^i) \) reaches its minimum at \( p^i \), for which the following lemma provides a necessary and sufficient condition.

**Lemma 1:** The convex function, \( h_i(p^i) \), has a minimum value \( (\sigma^i > 0) \), if and only if \( [\lambda^i - \eta^i]^T p^i \leq \| p^i \|^2 \).

**Proof:** Note that \( h_i(p^i) \) consists of a cone (conical surface) \( \tau \| p^i \|_2 \), and a hyperplane \( (\eta^i - \lambda^i)^T p^i \), both passing through the origin. Thus, the necessary and sufficient condition for \( h_i(p^i) \), reaching its minimum \( (\sigma^i > 0) \), is that the hyperplane, \( (\eta^i - \lambda^i)^T p^i \), must be located below the cone, \( \tau \| p^i \|_2 \), i.e., \( \lambda^i(p^i) \geq 0 \) and hence \( \lambda^i(p^i) \leq \| p^i \|_2 \).

Combining Lemma 1, (12a), (12c) and (12d), the condition (12b) can be readily simplified as

\[
\tau \| p^i \|_2 \geq \| \tilde{z} - \Phi(\Omega) p^i \|_2, i \in [N]. \tag{13}
\]

Since \( \tilde{z} - \Phi(\Omega) p^i \) is equivalent to the random noise \( \tilde{e} \Delta [R(e^T), \text{Im}(e^T)] \), \( \tau \) should be chosen such that the inequality (13) holds with high probability as nearly a sure event. The condition (13), however, involves the unknown parameter \( p^i \) (or \( \| p^i \|_2 \)), which raises difficulty for choosing a proper \( \tau \). To circumvent this issue, we introduce the following theorem which characterizes the condition (13) in a probabilistic sense and hence provides a turning point for the regularization parameter selection.

**Theorem 1:** Denote the \( j \)th column of \( \Phi(\Omega) \) by \( \Phi_{\alpha,j} \), and define \( N \) events:

\[
\Delta_i : \| \tilde{z} - \Phi(\Omega) p^i \|_2 \leq \| p^i \|_2, \quad i \in [N],
\]

\[
\Delta'_i : \sum_{i=1}^{K} \| \Phi_{\alpha,j}(\tilde{z} - \Phi(\Omega) p^i) \|_2 \leq \tau^2, \quad i \in [N]. \tag{14}
\]

then \( \Pr(\Delta'_i) \leq \Pr(\Delta_i), i \in [N], \) where \( \Pr(.) \) represents the probability of event.

**Proof:** Using the Cauchy–Schwarz inequality

\[
\| \tilde{z} - \Phi(\Omega) p^i \|^2 \| \Phi_{\alpha,j} \|^2 \leq \| \tilde{z} - \Phi(\Omega) p^i \|_2 \| \Phi_{\alpha,j} \|^2,
\]

i.e., \( \Delta'_i \Rightarrow \Delta_i \) and hence \( \Delta'_i \subset \Delta_i \), which ultimately leads to

\[
\Pr(\Delta'_i) \leq \Pr(\Delta_i), i \in [N].
\]

From Theorem 1, if \( \Pr(\Delta'_i) \rightarrow 1 \), then \( \Pr(\Delta_i) \rightarrow 1 \). Therefore, we select the parameter \( \tau \) such that the \( N \) events, \( \{ \Delta'_i : i \in [N] \} \), i.e., all inequalities in (14), hold with high probability. Note that the signals and noises are statistically independent of each other. We thus have

\[
\Phi_{\alpha,j}(\tilde{z} - \Phi(\Omega) p^i) \sim N(0, K_{i,M,M(i)}),
\]

and hence \( \tilde{e} \sim \mathcal{CN}(0, I_{K,M,M(i)}) \) which further yields \( \tilde{e} \sim [R(e^T), \text{Im}(e^T)] \sim N(0, \frac{1}{2} K_{M,M(i)}) \) where \( N(\cdot) \) denotes the normal distribution. Recall that all the column vectors of \( \Phi(\Omega) \) have been normalized to one. We naturally have

\[
\Phi_{\alpha,j}(\tilde{z} - \Phi(\Omega) p^i) \sim N(0, \frac{1}{2} I_{K,M,M(i)}),
\]

where \( I_{K,M,M(i)} \) denotes the chi-square distribution with \( K_{i,M,M(i)} \) DOFs. For simplicity, we simultaneously set all \( N \) independent events \( \{ \Delta'_i : i \in [N] \} \) with high probability, defined by \( \Phi_{\alpha,j} \).
and each event $\Delta'$, with equal probability $P_\Delta$. Eventually, the choice of $\tau$ is
\[
\tau = \sqrt{\frac{\xi}{2}}, \quad P_{\frac{\chi^2}{2}}(K) \leq \xi
\] (15)
where $\mathcal{F}_{\chi^2}(K)$ denotes the cumulative distribution function of $\chi^2(K)$.

### IV. SIMULATION RESULTS

In our simulations, the signals and noises within different frequency bins are assumed to be complex Gaussian distributed and with power vectors $\rho = [\rho_1, \ldots, \rho_Q]^T$ and $\sigma = [\sigma_1^2, \ldots, \sigma_M^2]^T$, respectively. Accordingly, the WB stochastic Cramér-Rao bound (CRB) is derived in the Appendix, which can be applied to the underdetermined case. The ideal weighted matrix $W_k$, defined in (7), is replaced by $W_k \triangleq \frac{1}{\sqrt{2}} J \{ H_k^H \otimes R_k \} J^T$. We further assume $\rho_1^2 = \ldots = \rho_Q^2 = \rho^2$ and define the signal-to-noise ratio (SNR) as $\text{SNR} \triangleq \frac{\rho^2}{\sigma^2} \sum_{m=1}^{M} \frac{1}{\tau_m}$ [35]. Throughout the simulations, we consider a 2-level nested array of 6 sensors with locations $\mathcal{D} = \{0, 2d, 3d, 7d, 11d\}$ where $d$ is the basic element-spacing. The probability $P_s$ is set to be 0.999. In addition, $c = 340$ m/s and each signal has the common center frequency of $f_c = 2.5$ kHz.

**Case 1: Aliasing-free Test.** In this case, there are $Q = 9$ signals with a common bandwidth of 1 kHz from directions $[-68^\circ, -45^\circ, -30^\circ, -15^\circ, 0^\circ, 12^\circ, 24^\circ, 42^\circ, 63^\circ]$. The uniform noise case with variance 10 is considered here and the grid spacing is set as $1^\circ$ within $[-90^\circ, 90^\circ]$. The basic element-spacing, $d$, is set as 8 cm, i.e., the half-wavelength of 2.125 kHz. As such, the spatial aliasing may occur when $f > 2.125$ kHz for the SS-MUSIC [9], which, however, can be well suppressed by the W-SpSF approach as long as $f$ is less than the half-wavelength of the lowest frequency [20, Theorem 2]. The frequency-bearing images are plotted in Fig. 1 at $L = 400$. 0.1° around the estimated peaks, but the initial grid remains unchanged as in Case 1. In Fig. 2, we plot the RMSEs versus the SNR with $L = 400$ (left subplot) and versus the number of snapshots with SNR = 0 dB (right subplot), under a common condition of $K = 10$. It is observed that the proposed W-SpSF estimator has a lower RMSE which is much closer to the CRB than that of the SS-MUSIC.

### V. CONCLUSION

This letter presents the Wideband Sparse Spectrum Fitting (W-SpSF) estimator which can efficiently accommodate the increased DOFs from sparse spatial sampling to perform underdetermined DOA estimation. To guarantee robust sparse recovery against the parameter sensitivity of W-SpSF, the best regularization parameter is obtained by the KKT conditions. Theoretical analysis and simulation results illustrate that W-SpSF exhibits excellent estimation performance and is superior to the SS-MUSIC. Nevertheless, due to modeling error (outlier) W-SpSF cannot provide reliable DOA estimation for correlated signals, which will be addressed as our future work.

### APPENDIX

### STOCHASTIC CRB FOR WIDEBAND SIGNALS

Stack all the array outputs $\{x_i(f_k)\}_{k=1}^{K}$ as a vector $x \triangleq [x^T_1, x^T_2, \ldots, x^T_K]^T$. With the statistic assumptions in Section II, the resultant covariance matrix of $x$ is
\[
R = \mathbb{E} \{ x \cdot x^T \} = \text{blkdiag} \{ R_1, \ldots, R_K \}. \tag{16}
\]
Define the parameter vector as $\psi \triangleq [\theta^T, \delta^T, \beta^T, \gamma^T]^T$. Recall that the CRB is the inverse of the Fisher information matrix (FIM). That is, we only need to derive its FIM. Using the block diagonal structure of $R$, the $(i,j)$th entry of the associated FIM, denoted by $F$, is given by [36]
\[
[F]_{i,j} = L \cdot \text{tr} \left( R^{-1} \frac{\partial R}{\partial \psi_i} R^{-1} \frac{\partial R}{\partial \psi_j} \right) = \sum_{k=1}^{K} \left[ L \cdot \text{tr} \left( R_{k}^{-1} \frac{\partial R_{k}}{\partial \psi_i} R_{k}^{-1} \frac{\partial R_{k}}{\partial \psi_j} \right) \right] \tag{17}
\]
where $\text{tr}(\cdot)$ is the trace operator and $\psi_i$ is the $i$th element of $\psi$. Thus, the WB FIM, $F$, is the sum of all the associated narrowband FIMs, $\{F_k\}_{k=1}^{K}$. Then using the result in [14] on the narrowband FIM, we can easily obtain the relevant WB CRB.
REFERENCES


