Cramér–Rao Bound for Sparse Signals Fitting the Low-Rank Model with Small Number of Parameters

Mahdi Shaghaghi, Student Member, IEEE, and Sergiy A. Vorobyov, Senior Member, IEEE

Abstract—In this letter, we consider signals with a low-rank covariance matrix which reside in a low-dimensional subspace and can be written in terms of a finite (small) number of parameters. Although such signals do not necessarily have a sparse representation in a finite basis, they possess a sparse structure which makes it possible to recover the signal from compressed measurements. We study the statistical performance bound for parameter estimation in the low-rank signal model from compressed measurements. Specifically, we derive the Cramér-Rao bound (CRB) for a generic low-rank model and we show that the number of compressed samples needs to be larger than the number of sources for the existence of an unbiased estimator with finite estimation variance. We further consider the applications to direction-of-arrival (DOA) and spectral estimation which fit into the low-rank signal model. We also investigate the effect of compression on the CRB by considering numerical examples of the DOA estimation scenario, and show how the CRB increases by increasing the compression or equivalently reducing the number of compressed samples.

Index Terms—Compressed sensing, Cramér–Rao bound, DOA estimation, low-rank model, spectral estimation.

I. INTRODUCTION

S IGNALS with sparse representations can be recovered from much less number of measurements than the number of samples given by the Nyquist rate using compressed sensing (CS) methods [1]–[3]. Such measurements can be obtained by correlating the signal with a number of sensing waveforms [4]–[9]. The recovering algorithms for the signals from such measurements exploit the sparsity of the signals in a proper basis (see [3], [10]–[17] to mention a few existing algorithms).

There are signals which inherently possess a sparse structure meaning that they can be defined by a small number of parameters. However, such signals may not necessarily be represented as sparse signals using a proper finite basis, i.e., there may not exist or be known a finite basis such that the transformation of the signal to that basis results in a small number of non-zero coefficients. For example, consider a signal composed of a linear combination of sinusoids. Such a signal generates sparse coefficients by the discrete-time Fourier transform (DTFT), but

Manuscript received December 23, 2014; revised February 26, 2015; accepted March 02, 2015. Date of publication March 04, 2015; date of current version March 11, 2015. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Michael B. Wakin.

M. Shaghaghi was with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4 Canada. He is now with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4 Canada (e-mail: mahdi.shaghaghi@ualberta.ca).

S. A. Vorobyov is with Department of Signal Processing and Acoustics, Aalto University, FI-00076 Aalto, Finland (e-mail: svor@ieee.org).

Digital Object Identifier 10.1109/LSP.2015.2409896

its representation in the Fourier basis obtained by the discrete Fourier transform (DFT) exhibits frequency leakage [18]. Although the DTFT is a proper transform, as it results in a small number of non-zero coefficients for the considered signal, it is not a finite basis and cannot be used in conventional CS recovery methods which rely on a finite sparsity basis. Such methods have poor performance for such signals if the DFT basis is used [19]. We consider a general class of sparse signals which are represented by a small number of parameters in a low-rank signal model. Our goal is to study the performance bounds for the estimation of unknown parameters and also the reconstruction of this class of signals from compressed measurements.

The Cramér-Rao bound (CRB) [20] for estimating a sparse parameter vector from compressed measurements has been studied in [21]. However, the signal model in [21] considers signals which can be represented by a finite sparsity basis. The CRB has been computed using approaches from the theory of constrained CRB in [22] and [23]. The constrained CRB for estimating a low-rank matrix from compressed measurements has been studied in [24]. Here we consider a different signal model which does not involve the constraint on the rank of a matrix. The CRB for parameter estimation in compressed sensing has been also studied in [25]-[27]. In [25], the signal is assumed to be a function of real-valued parameters, and it is not necessarily sparse in a finite basis. The CRB has been computed and bounded for different realizations of the measurement matrix. The signal model in [26] and [27] is different from the one studied here in two aspects. First, in [26] and [27], a noiseless signal is first compressed and then white noise is added to the compressed signal. In contrast, we first add the noise to the signal and then compress. This results in a different distribution for the compressed measurements. Second, in [26] and [27], the signal is a vector which depends on a number of parameters, whereas in this letter, the signal is composed of a parametrized matrix multiplied by a vector of coefficients. This structure of the signal enables us to derive a closed-form expression for the CRB of the parameters.

In this letter, we extend the results of [25] for a low-rank signal model. We derive the CRB for real and complex-valued parameters. Furthermore, multiple signal snapshots are considered, whereas in [25], the signal model consists of only a single signal snapshot. We also study the minimum number of compressed samples required for unbiased estimation with finite variance. Furthermore, the applications to direction-of-arrival (DOA) and spectral estimation which fit into the low-rank signal model are also studied. Finally, numerical examples for the DOA estimation problem are given to illustrate the effect of compression on the CRB.

^{1070-9908 © 2015} IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

II. SIGNAL MODEL

Consider the signal $\boldsymbol{x}(t) \in \mathbb{C}^{N_x \times 1}$ at time instant t to be of the form

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{d}(t) \tag{1}$$

where $\boldsymbol{A} \in \mathbb{C}^{N_x \times K}$ is a tall matrix (the number of rows is much larger than the number of columns), $\boldsymbol{d}(t) \in \mathbb{C}^{K \times 1}$ is a vector containing unknown amplitudes, and $1 \leq t \leq N$. A practically important example of \boldsymbol{A} is given in Section V. Since \boldsymbol{A} is a tall matrix, the covariance matrix of the signal is a low-rank matrix. Therefore, such a signal is called low-rank. Matrix \boldsymbol{A} can be fully known, known up to a number of unknown parameters, or completely unknown. In this letter, we study the second case where matrix \boldsymbol{A} has a known structure, but it contains P number of unknown parameters $\boldsymbol{\Omega} \triangleq [\omega_1, \dots, \omega_P]^T \in \mathbb{R}^{P \times 1}$ where $(\cdot)^T$ stands for the transposition operator.

Let the vector of the measurements $\boldsymbol{y} \in \mathbb{C}^{N_y \times 1}$ be given by

$$\boldsymbol{y}(t) = \boldsymbol{\Phi} \left(\boldsymbol{x}(t) + \boldsymbol{w}(t) \right) = \boldsymbol{\Phi} \boldsymbol{x}(t) + \boldsymbol{n}(t) \tag{2}$$

where $\mathbf{\Phi} \in \mathbb{R}^{N_y \times N_x}$ is the measurement matrix with $N_y \leq N_x$. The additive noise $\mathbf{w}(t) \in \mathbb{C}^{N_x \times 1}$ is assumed to have the circularly-symmetric complex jointly-Gaussian distribution $\mathcal{N}_C(\mathbf{0}, \sigma^2 \mathbf{I}_{N_x})$ where \mathbf{I}_{N_x} is the identity matrix of size N_x and σ^2 is the noise power. No specific structure for the measurement matrix $\mathbf{\Phi}$ needs to be considered in our derivations. It is because $\mathbf{\Phi}$ is assumed to be known at the signal reconstruction stage, and therefore, it is treated as a deterministic matrix in our derivations. As a result, irrespective to how $\mathbf{\Phi}$ is generated, the measurement noise $\mathbf{n}(t) \in \mathbb{C}^{N_y \times 1}$ has Gaussian distribution $\mathcal{N}_C(\mathbf{0}, \mathbf{R})$ where $\mathbf{R} = \sigma^2 \mathbf{\Phi} \mathbf{\Phi}^T$.

III. DERIVATION OF THE CRB

In this section, we derive the CRB for the signal model given by (1) and (2).

First, let the vector of parameters be defined as

$$\boldsymbol{\vartheta} \triangleq \left[\bar{\boldsymbol{d}}^T(1), \tilde{\boldsymbol{d}}^T(1), \cdots, \bar{\boldsymbol{d}}^T(N), \tilde{\boldsymbol{d}}^T(N), \boldsymbol{\Omega}^T \right]^T \qquad (3)$$

where $\bar{d}(t)$ and $\tilde{d}(t)$ represent the real and imaginary parts of d(t), respectively.

The likelihood function of the compressed measurements (2) is given by

$$p(\mathbf{y}(1), \cdots, \mathbf{y}(N) | \boldsymbol{\vartheta}) = \frac{1}{\pi^{N_y N} |\boldsymbol{R}|^N}$$
$$\times \exp\{-\sum_{t=1}^N (\boldsymbol{y}(t) - \boldsymbol{B}\boldsymbol{d}(t))^H \boldsymbol{R}^{-1} (\boldsymbol{y}(t) - \boldsymbol{B}\boldsymbol{d}(t))\} (4)$$

where $\boldsymbol{B} \triangleq \boldsymbol{\Phi} \boldsymbol{A}$ and $(\cdot)^H$ stands for the Hermitian transposition operator. The log-likelihood function can be found by taking the natural logarithm of (4) as

$$LL \triangleq \ln p \left(\mathbf{y}(1), \cdots, \mathbf{y}(N) | \boldsymbol{\vartheta} \right) = -N_y N \ln \pi - N \ln |\mathbf{R}|$$
$$-\sum_{t=1}^{N} \left(\mathbf{y}(t) - \mathbf{B}\mathbf{d}(t) \right)^H \mathbf{R}^{-1} \left(\mathbf{y}(t) - \mathbf{B}\mathbf{d}(t) \right).$$
(5)

For brevity, the notation LL will be used in the rest of the letter to refer to the log-likelihood function (5). The Fisher information matrix (FIM) is given by

$$\boldsymbol{I}(\boldsymbol{\vartheta}) = E\left\{\boldsymbol{\psi}\boldsymbol{\psi}^T\right\} \tag{6}$$

where $\psi \triangleq \partial LL / \partial \vartheta$. The CRB covariance matrix for the vector of parameters ϑ is then given by

$$\operatorname{CRB}\left(\boldsymbol{\vartheta}\right) = \boldsymbol{I}^{-1}(\boldsymbol{\vartheta}). \tag{7}$$

The derivatives of the *LL* with respect to $\overline{d}(t)$ and $\overline{d}(t)$ are given by

$$\frac{\partial LL}{\partial \bar{\boldsymbol{d}}(t)} = \boldsymbol{B}^{H} \boldsymbol{R}^{-1} \boldsymbol{n}(t) + \left(\boldsymbol{n}^{H}(t) \boldsymbol{R}^{-1} \boldsymbol{B}\right)^{T}$$
$$= 2Re \left\{ \boldsymbol{B}^{H} \boldsymbol{R}^{-1} \boldsymbol{n}(t) \right\}$$
(8)

and

$$\frac{\partial LL}{\partial \tilde{\boldsymbol{d}}(t)} = -j\boldsymbol{B}^{H}\boldsymbol{R}^{-1}\boldsymbol{n}(t) + j\left(\boldsymbol{n}^{H}(t)\boldsymbol{R}^{-1}\boldsymbol{B}\right)^{T}$$
$$= 2Im\left\{\boldsymbol{B}^{H}\boldsymbol{R}^{-1}\boldsymbol{n}(t)\right\}$$
(9)

where $Re\{\cdot\}$ and $Im\{\cdot\}$ stand for the real part and imaginary part operators, respectively. Recall that $\boldsymbol{n}(t) = \boldsymbol{y}(t) - \boldsymbol{Bd}(t)$ is the measurement noise introduced in (2).

Note that A has a known structure and contains P unknown parameters $\omega_1, \dots, \omega_P$. Therefore, the derivative of the LL with respect to ω_p for $1 \le p \le P$ can be found as

$$\frac{\partial LL}{\partial \omega_p} = \sum_{t=1}^{N} \boldsymbol{d}^{H}(t) \frac{\partial \boldsymbol{B}^{H}}{\partial \omega_p} \boldsymbol{R}^{-1} \boldsymbol{n}(t) + \boldsymbol{n}^{H}(t) \boldsymbol{R}^{-1} \frac{\partial \boldsymbol{B}}{\partial \omega_p} \boldsymbol{d}(t)$$

$$= 2 \sum_{t=1}^{N} Re \left\{ \boldsymbol{d}^{H}(t) \frac{\partial \boldsymbol{B}^{H}}{\partial \omega_p} \boldsymbol{R}^{-1} \boldsymbol{n}(t) \right\}$$

$$= 2 \sum_{t=1}^{N} Re \left\{ \boldsymbol{d}^{H}(t) \frac{\partial \boldsymbol{A}^{H}}{\partial \omega_p} \boldsymbol{\Phi}^{T} \boldsymbol{R}^{-1} \boldsymbol{n}(t) \right\}.$$
(10)

The derivatives of the LL with respect to the whole vector Ω can be then written in matrix form as

$$\frac{\partial LL}{\partial \boldsymbol{\Omega}} = 2 \sum_{t=1}^{N} Re \left\{ \boldsymbol{D}^{H}(t) \boldsymbol{\Phi}^{T} \boldsymbol{R}^{-1} \boldsymbol{n}(t) \right\}$$
(11)

where the matrix $D(t) \in \mathbb{C}^{N_x \times P}$ is given by

$$\boldsymbol{D}(t) \triangleq \left[\frac{\partial \boldsymbol{A}}{\partial \omega_1} \boldsymbol{d}(t), \cdots, \frac{\partial \boldsymbol{A}}{\partial \omega_P} \boldsymbol{d}(t) \right] \\ = \left[\frac{\partial \boldsymbol{A}}{\partial \omega_1}, \cdots, \frac{\partial \boldsymbol{A}}{\partial \omega_P} \right] (\boldsymbol{I}_P \otimes \boldsymbol{d}(t))$$
(12)

with \otimes standing for the Kronecker product.

To proceed, we use the following identities [28]. For two arbitrary complex vectors p and q, we have

$$Re(\boldsymbol{p})Re\left(\boldsymbol{q}^{T}\right) = \frac{1}{2} \left(Re\left(\boldsymbol{p}\boldsymbol{q}^{T}\right) + Re\left(\boldsymbol{p}\boldsymbol{q}^{H}\right) \right)$$
(13)

$$Im(\boldsymbol{p})Im(\boldsymbol{q}^{T}) = -\frac{1}{2}\left(Re\left(\boldsymbol{p}\boldsymbol{q}^{T}\right) - Re\left(\boldsymbol{p}\boldsymbol{q}^{H}\right)\right) \quad (14)$$

$$Re(\boldsymbol{p})Im\left(\boldsymbol{q}^{T}\right) = \frac{1}{2}\left(Im\left(\boldsymbol{p}\boldsymbol{q}^{T}\right) - Im\left(\boldsymbol{p}\boldsymbol{q}^{H}\right)\right).$$
 (15)

This paper previously published in IEEE Signal Processing Letters

Using (13), (14), (15), and the fact that for $1 \le r, s \le N$

$$E\left\{\boldsymbol{n}(r)\boldsymbol{n}^{T}(s)\right\} = \boldsymbol{0}$$
(16)

$$E\left\{\boldsymbol{n}(r)\boldsymbol{n}^{H}(s)\right\} = \delta_{r,s}\boldsymbol{R}$$
(17)

where $\delta_{r,s}$ denotes the Kronecker delta, we can compute the submatrices of $I(\vartheta)$ as

$$E\left(\frac{\partial LL}{\partial \bar{\boldsymbol{d}}(r)}\right)\left(\frac{\partial LL}{\partial \bar{\boldsymbol{d}}(s)}\right)^{T} = 2Re\{E\{\boldsymbol{B}^{H}\boldsymbol{R}^{-1}\boldsymbol{n}(r) \times \boldsymbol{n}^{H}(s)\boldsymbol{R}^{-1}\boldsymbol{B}\}\}$$
$$= 2Re\left\{\boldsymbol{B}^{H}\boldsymbol{R}^{-1}\boldsymbol{B}\right\}\delta_{r,s}$$
(18)

$$E\left(\frac{\partial LL}{\partial \bar{\boldsymbol{d}}(r)}\right)\left(\frac{\partial LL}{\partial \tilde{\boldsymbol{d}}(s)}\right)^{T} = -2Im\left\{\boldsymbol{B}^{H}\boldsymbol{R}^{-1}\boldsymbol{B}\right\}\delta_{r,s}$$
(19)

$$E\left(\frac{\partial LL}{\partial \bar{\boldsymbol{d}}(r)}\right)\left(\frac{\partial LL}{\partial \boldsymbol{\Omega}}\right)^{T} = 2Re\left\{\boldsymbol{B}^{H}\boldsymbol{R}^{-1}\boldsymbol{\Phi}\boldsymbol{D}(r)\right\}$$
(20)

$$E\left(\frac{\partial LL}{\partial \tilde{\boldsymbol{d}}(r)}\right)\left(\frac{\partial LL}{\partial \tilde{\boldsymbol{d}}(s)}\right)^{T} = 2Re\left\{\boldsymbol{B}^{H}\boldsymbol{R}^{-1}\boldsymbol{B}\right\}\delta_{r,s}$$
(21)

$$E\left(\frac{\partial LL}{\partial \tilde{\boldsymbol{d}}(r)}\right)\left(\frac{\partial LL}{\partial \boldsymbol{\Omega}}\right)^{T} = 2Im\left\{\boldsymbol{B}^{H}\boldsymbol{R}^{-1}\boldsymbol{\Phi}\boldsymbol{D}(r)\right\}$$
(22)

$$E\left(\frac{\partial LL}{\partial \boldsymbol{\Omega}}\right)\left(\frac{\partial LL}{\partial \boldsymbol{\Omega}}\right)^{T} = 2\sum_{t=1}^{N} Re\left\{\boldsymbol{D}^{H}(t)\boldsymbol{\Phi}^{T}\boldsymbol{R}^{-1}\boldsymbol{\Phi}\boldsymbol{D}(t)\right\}.$$
(23)

Then, $I(\vartheta)$ can be found as

$$I(\vartheta) = \begin{bmatrix} H & -H & & \Delta_1 \\ \tilde{H} & \bar{H} & 0 & \tilde{\Delta}_1 \\ & \ddots & & \vdots \\ 0 & & \bar{H} & -\tilde{H} & \bar{\Delta}_N \\ & & \tilde{H} & \bar{H} & \tilde{\Delta}_N \\ \bar{\Delta}_1^T & \tilde{\Delta}_1^T & \cdots & \bar{\Delta}_N^T & \tilde{\Delta}_N^T & \Gamma \end{bmatrix}$$
(24)

where $\overline{(\cdot)}$ and $\overline{(\cdot)}$ stand for the real and imaginary parts of a matrix, and

$$\boldsymbol{H} \triangleq 2\boldsymbol{B}^{H}\boldsymbol{R}^{-1}\boldsymbol{B}$$
(25)

$$\boldsymbol{\Delta}_{r} \triangleq 2\boldsymbol{B}^{H}\boldsymbol{R}^{-1}\boldsymbol{\Phi}\boldsymbol{D}(r) \tag{26}$$

$$\boldsymbol{\Gamma} \triangleq 2\sum_{t=1}^{T} Re\left\{\boldsymbol{D}^{H}(t)\boldsymbol{\Phi}^{T}\boldsymbol{R}^{-1}\boldsymbol{\Phi}\boldsymbol{D}(t)\right\}.$$
 (27)

It is shown in [28] that for FIM with the structure given in (24), the CRB covariance matrix for Ω is given by

$$\operatorname{CRB}^{-1}(\mathbf{\Omega}) = \mathbf{\Gamma} - \sum_{t=1}^{N} \operatorname{Re}\left\{\mathbf{\Delta}_{t}^{H} \mathbf{H}^{-1} \mathbf{\Delta}_{t}\right\}.$$
(28)

Using (25)–(28), the CRB for Ω can be found in closed-form as

$$CRB^{-1}(\boldsymbol{\Omega}) = 2\sum_{t=1}^{N} Re\{\boldsymbol{D}^{H}(t)\boldsymbol{\Phi}^{T}\boldsymbol{R}^{-1}(\boldsymbol{I}_{N_{y}} - \boldsymbol{B}(\boldsymbol{B}^{H}\boldsymbol{R}^{-1}\boldsymbol{B})^{-1}\boldsymbol{B}^{H}\boldsymbol{R}^{-1})\boldsymbol{\Phi}\boldsymbol{D}(t)\}.$$
(29)

Given $I^{-1}(\vartheta)$, the covariance matrix of any unbiased estimator of $\boldsymbol{x}(t)$, i.e., $C_{\hat{\boldsymbol{x}}(t)}$, satisfies the inequality

$$\boldsymbol{C}_{\hat{\boldsymbol{x}}(t)} - \frac{\partial \boldsymbol{x}(t)}{\partial \boldsymbol{\vartheta}} \boldsymbol{I}^{-1}(\boldsymbol{\vartheta}) \frac{\partial \boldsymbol{x}^{H}(t)}{\partial \boldsymbol{\vartheta}} \geq \boldsymbol{0}.$$
 (30)

The signal $\boldsymbol{x}(t)$ can be written as

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{d}(t) = \boldsymbol{A}\boldsymbol{\bar{d}}(t) + j\boldsymbol{A}\boldsymbol{\bar{d}}(t). \tag{31}$$

The derivative of $\boldsymbol{x}(t)$ with respect to the vector of unknown parameters $\boldsymbol{\vartheta}$ is given by

$$\frac{\partial \boldsymbol{x}(t)}{\partial \boldsymbol{\vartheta}} = \left[\boldsymbol{e}_t \otimes \left[\boldsymbol{A}, j\boldsymbol{A}\right], \boldsymbol{D}(t)\right]$$
(32)

where e_t is a row vector of length N with all its elements equal to zero except for the t-th element which is equal to 1. Finally, by summing over the diagonal elements of (30), we obtain

$$E\left\{\|\hat{\boldsymbol{x}}(t) - \boldsymbol{x}(t)\|^{2}\right\} \geq \operatorname{Tr}\left\{\frac{\partial \boldsymbol{x}(t)}{\partial \boldsymbol{\vartheta}}\boldsymbol{I}^{-1}(\boldsymbol{\vartheta})\frac{\partial \boldsymbol{x}^{H}(t)}{\partial \boldsymbol{\vartheta}}\right\}.$$
 (33)

Similar to (29), the results in (32) and (33) can also be regarded as closed-form as they can be used for analysis, fast computations, and getting insights without requiring Monte-Carlo simulations as shown in the next section. It is worth noting that the derived CRB (especially (29)) can be also used for selecting/optimizing the measurement matrix $\mathbf{\Phi}$ as the CRB depends on a specific selection of $\mathbf{\Phi}$.

IV. MINIMUM NUMBER OF COMPRESSED SAMPLES

In this section, we show that if the number of compressed samples is less than or equal to the number of sources $(N_y \leq K)$, the FIM $I(\vartheta)$ is singular. It is shown in [29] that a singular FIM means that unbiased estimation of the entire parameter vector with finite variance is impossible.

Let us start with the case that $N_y < K$. In this case, we have $rank\{B\} < K$ since $B \in \mathbb{C}^{N_y \times K}$. As a result, we also have $rank\{H\} < K$ (see (25)), and therefore, H is singular. Consequently, there exists a nonzero vector $\boldsymbol{u} = \bar{\boldsymbol{u}} + j\tilde{\boldsymbol{u}} \in \mathbb{C}^{K \times 1}$ such that $H\boldsymbol{u} = \boldsymbol{0}$. Therefore, $(\bar{\boldsymbol{H}} + j\tilde{\boldsymbol{H}})(\bar{\boldsymbol{u}} + j\tilde{\boldsymbol{u}}) = \boldsymbol{0}$, which can be written in matrix form as

$$\begin{bmatrix} \bar{H} & -\tilde{H} \\ \tilde{H} & \bar{H} \end{bmatrix} \begin{bmatrix} \bar{u} \\ \tilde{u} \end{bmatrix} = 0.$$
(34)

Let $\boldsymbol{v} \triangleq [\boldsymbol{\bar{u}}^T, \boldsymbol{\tilde{u}}^T, \boldsymbol{0}]^T \in \mathbb{R}^{(2NK+P)\times 1}$. Finally, using (24), we have $\boldsymbol{v}^T \boldsymbol{I}(\boldsymbol{\vartheta}) \boldsymbol{v} = \boldsymbol{0}$, which means that $\boldsymbol{I}(\boldsymbol{\vartheta})$ has a zero eigenvalue, and therefore, it is singular. For the case that $N_y = K$, if $rank\{\boldsymbol{B}\} < K$, the singularity of the FIM follows from the discussion above.

Now, consider the case that **B** is full-rank. Thus, **H** is invertible. Consider the structure of $I(\vartheta)$ in (24) and let the block of all the real and imaginary parts of **H** be denoted by **T**. It is shown in [28] that for an invertible matrix **H**, matrix **T** is also invertible. The Schur complement of **T** denoted by $I(\vartheta)/T$ is equal to the inverse of the CRB covariance matrix for Ω as given in (29). Matrix **B** is invertible since it is square and full-rank. Therefore, we have

$$\boldsymbol{I}_{N_y} - \boldsymbol{B} \left(\boldsymbol{B}^H \boldsymbol{R}^{-1} \boldsymbol{B} \right)^{-1} \boldsymbol{B}^H \boldsymbol{R}^{-1} = \boldsymbol{0}.$$
(35)

As a result, $I(\vartheta)/T = 0$ (see (29)). According to the rank additivity formula [30], we have

$$rank \{ I(\vartheta) \} = rank \{ T \} + rank \{ I(\vartheta) / T \} = rank \{ T \}.$$
(36)

Therefore, $I(\vartheta)$ is rank-deficient or equivalently singular.

This paper previously published in IEEE Signal Processing Letters

Remark: As shown above, if the number of compressed samples is less than or equal to the number of sources, the FIM is necessarily singular. However, if the number of compressed samples increases, it does not necessarily result in a non-singular FIM for a few more samples. Thus, the converse does not hold in general. The minimum number of compressed samples for satisfactory performance depends on a specific performance criterion and the estimation method used. For example, the minimum number of compressed samples can be chosen to bound the probability of a subspace swap [31] or to bound the error of signal subspace estimation [32]. The required number of compressed samples can also be studied from a geometric point of view [33].

V. APPLICATION EXAMPLES

For the problems of DOA and spectral estimation, d(t) consists of the amplitudes of K number of sources at time instant t. The number of parameters in A is also equal to the number of sources, i.e., P = K. Furthermore, A has the structure given by

$$\boldsymbol{A} \triangleq \left[\boldsymbol{a}\left(\omega_{1}\right), \cdots, \boldsymbol{a}\left(\omega_{K}\right)\right]$$
(37)

where $\mathbf{a}(\omega_k)$ for $1 \leq k \leq K$ is the steering vector corresponding to the k-th source. Let us define $\mathbf{c}(\omega)$ as the derivative of $\mathbf{a}(\omega)$ with respect to ω , i.e., $\mathbf{c}(\omega) \triangleq d\mathbf{a}(\omega)/d\omega$. Then, $\mathbf{D}(t)$ given by (12) can be simplified to

$$D(t) = [c(\omega_1) d_1(t), \cdots, c(\omega_K) d_K(t)]$$

= [c(\omega_1), \cdots, c(\omega_K)] diag {d(t)} (38)

where $d_k(t)$ is the k-th element of d(t) and the diag $\{\cdot\}$ operator converts a vector into a diagonal matrix.

VI. NUMERICAL RESULTS

In this section, the application of the derived CRB formulas for the problem of DOA estimation is illustrated. Our goal is to investigate the performance bounds for unbiased estimators when the signal is compressed at different rates.

Consider K = 11 equally spaced sources impinging on a uniform linear array of $N_x = 50$ antenna elements from directions $\omega_1 = 20^\circ \times (\pi/180), \omega_2 = 23^\circ \times (\pi/180), \cdots, \omega_{11} =$ $50^\circ \times (\pi/180)$. The steering vector of the array $\boldsymbol{a}(\omega)$ is given by

$$\boldsymbol{a}(\omega) \triangleq \left[1, e^{-j2\pi(d/\lambda)\sin(\omega)}, \cdots, e^{-j2\pi(N_x-1)(d/\lambda)\sin(\omega)}\right]^T$$
(39)

where d is the interelement spacing of the array and λ is the wavelength of the plane wave impinging on the array. In our numerical example, d/λ is set to 0.5. The number of snapshots is also set to N = 10. Each source vector d(t) is considered to be independent from the source vectors at other time instances and is drawn from the circularly-symmetric complex jointly-



Fig. 1. CRB for estimating $\omega_6 = 35^\circ \times (\pi/180)$.

Gaussian distribution $\mathcal{N}_C(\mathbf{0}, \sigma_s^2 \boldsymbol{I}_K)$. The signal-to-noise ratio (SNR) is set to SNR $\triangleq 10 \log_{10}(\sigma_s^2/\sigma^2) = 10$ dB. The source vectors are drawn once and kept unchanged.

Fig. 1 shows the CRB for estimating $\omega_6 = 35^\circ \times (\pi/180)$ versus the number of compressed samples N_y . For the case when $N_y = N_x = 50$, the measurement matrix $\mathbf{\Phi}$ is set to the identity matrix. Then, $\mathbf{\Phi}$ is initialized for $N_y = 49$ by drawing samples from the Gaussian distribution $\mathcal{N}(0, 1/49)$. For the rest of N_y values, the first N_y rows of the initial matrix $\mathbf{\Phi}$ are scaled by $\sqrt{49/N_y}$ and used to obtain the CRB.

As expected, it can be seen in Fig. 1 that the CRB increases as the number of compressed samples N_y reduces. The minimum number of compressed samples is set to $N_y = 12$ which is equal to the number of sources plus one (K + 1). As shown in Section IV, if the number of compressed samples is equal to or less than the number of sources, there can be no unbiased estimator with a finite estimation variance. Otherwise, if the CRB exists, there also exist estimators [18] that achieve it.

VII. CONCLUSION

The class of signals fitting a low-rank signal model has been considered in this letter. Such signals are inherently sparse according to the signal model and can be recovered from compressed measurements. We have studied the performance bounds for unbiased estimators of parameters of such low-rank signal model from compressed samples. The Cramér-Rao bound has been derived for a generic low-rank model and it has been shown that the number of compressed samples needs to be at least larger than the number of sources for the existence of an unbiased estimator with finite variance. Furthermore, the applications to DOA and spectral estimation have been considered. Numerical examples have been also given to illustrate the effect of compression on the CRB. It has been shown how the CRB increases until the point where the number of compressed samples is larger than the number of sources. For lower number of compressed samples, the CRB becomes unbounded.

This paper previously published in IEEE Signal Processing Letters

REFERENCES

- D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [2] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, pp. 489–509, Feb. 2006.
- [3] E. J. Candès and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies?," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
- [4] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, pp. 21–30, Mar. 2008.
- [5] J. N. Laska, S. Kirolos, M. F. Duarte, T. S. Ragheb, R. G. Baraniuk, and Y. Massoud, "Theory and implementation of an analog-to-information converter using random demodulation," in *Proc. IEEE ISCAS*, New Orleans, LA, May 2007, pp. 1959–1962.
- [6] O. Taheri and S. A. Vorobyov, "Segmented compressed sampling for analog-to-information conversion: Method and performance analysis," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 554–572, Feb. 2011.
- [7] H. Fang, S. A. Vorobyov, and H. Jiang, "Performance limits of segmented compressive sampling: Correlated samples versus bits," *preprint arXiv: 14115178*, 2014.
- [8] M. Mishali and Y. C. Eldar, "Sub-Nyquist sampling: Bridging theory and practice," *IEEE Signal Process. Mag.*, vol. 28, no. 6, pp. 98–124, Nov. 2011.
- [9] M. Mishali, Y. C. Eldar, and A. Elron, "Xampling: Signal acquisition and processing in union of subspaces," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4719–4734, Oct. 2011.
- [10] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 4, pp. 586–597, Dec. 2007.
- [11] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.
- [12] E. J. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.*, vol. 59, pp. 1207–1223, Aug. 2006.
- [13] J. Haupt and R. Nowak, "Signal reconstruction from noisy random projections," *IEEE Trans. Inf. Theory*, vol. 52, pp. 4036–4048, Sep. 2006.
- [14] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
- [15] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comput. Harmon Anal.*, vol. 26, no. 3, pp. 301–321, May 2009.
- [16] E. J. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted l₁ minimization," *J. Fourier Anal. Applicat.*, vol. 14, no. 5, pp. 877–905, Dec. 2008.
- [17] O. Taheri and S. A. Vorobyov, "Reweighted l₁-norm penalized LMS for sparse channel estimation and its analysis," *Signal Process.*, vol. 104, pp. 70–79, May 2014.

- [18] M. Shaghaghi and S. A. Vorobyov, "Improved model-based spectral compressive sensing via nested least squares," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, Prague, Czech Republic, May 2011, pp. 3904–3907.
- [19] Y. Chi, L. L. Scharf, A. Pezeshki, and A. R. Calderbank, "Sensitivity to basis mismatch in compressed sensing," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2182–2195, May 2011.
- [20] H. L. Van Trees, K. L. Bell, and Z. Tian, *Detection, Estimation, and Modulation Theory*. Hoboken, NJ, USA: Wiley, 2005.
- [21] Z. Ben-Haim and Y. C. Eldar, "The Cramér-Rao bound for estimating a sparse parameter vector," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3384–3389, Jun. 2010.
- [22] J. D. Gorman and A. O. Hero, "Lower bounds for parametric estimation with constraints," *IEEE Trans. Inf. Theory*, vol. 26, no. 6, pp. 1285–1301, Nov. 1990.
- [23] P. Stoica and B. C. Ng, "On the Cramér-Rao bound under parametric constraints," *IEEE Signal Process. Lett.*, vol. 5, pp. 177–179, Jul. 1998.
- [24] G. Tang and A. Nehorai, "Lower bounds on the mean-squared error of low-rank matrix reconstruction," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4559–4571, Oct. 2011.
- [25] P. Pakrooh, L. L. Scharf, A. Pezeshki, and Y. Chi, "Analysis of Fisher information and the Cramér-Rao bound for nonlinear parameter estimation after compressed sensing," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, Vancouver, BC, USA, May 2013, pp. 6630–6634.
- [26] D. Ramasamy, S. Venkateswaran, and U. Madhow, "Compressive estimation in AWGN: General observations and a case study," in *Proc.* 46th Asilomar Conf. Signals, Syst., Comput., Nov. 2012, pp. 953–957.
- [27] D. Ramasamy, S. Venkateswaran, and U. Madhow, "Compressive parameter estimation in AWGN," *IEEE Trans. Signal Process.*, vol. 62, no. 8, pp. 2012–2027, Apr. 2014.
- [28] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 5, pp. 720–741, May 1989.
- [29] P. Stoica and T. L. Marzetta, "Parameter estimation problems with singular information matrices," *IEEE Trans. Signal Process.*, vol. 49, no. 1, pp. 87–90, Jan. 2001.
- [30] F. Zhang, The Schur Complement and Its Applications. Berlin, Germany: Springer-Verlag, 2005.
- [31] P. Pakrooh, A. Pezeshki, and L. L. Scharf, "Threshold effects in parameter estimation from compressed data," in *Proc. 1st IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Austin, TX, Dec. 2013, pp. 997–1000.
- [32] W. Mantzel and J. Romberg, "Compressed subspace matching on the continuum," *preprint arXiv*:1407.5234, 2014.
- [33] R. Baraniuk and M. Wakin, "Random projections of smooth manifolds," *Found. Comput. Math.*, vol. 9, no. 1, pp. 51–77, Feb. 2009.