Direction-of-Arrival Estimation for Temporally Correlated Narrowband Signals

Farzan Haddadi, Mohammad M. Nayebi, *Senior Member, IEEE*, and Mohammad R. Aref

***Abstract*—Signal direction-of-arrival (DOA) estimation using an array of sensors has been the subject of intensive research and development during the last two decades. Efforts have been directed to both, better solutions for the general data model and to develop more realistic models. So far, many authors have assumed the data to be independent and identically distributed (i.i.d.) samples of a multivariate statistical model. Although this assumption reduces the complexity of the model, it may not be true in certain situations where signals show temporal correlation. Some results are available on the temporally correlated signal model in the literature. The temporally correlated stochastic Cramér–Rao bound (CRB) has been calculated and an instrumental variable-based method called IV-SSF is introduced. Also, it has been shown that temporally correlated CRB is lower bounded by the deterministic CRB. In this paper, we show that temporally correlated CRB is also upper bounded by the stochastic i.i.d. CRB. We investigate the effect of temporal correlation of the signals on the best achievable performance. We also show that the IV-SSF method is not efficient and based on an analysis of the CRB, propose a variation in the method which boosts its performance. Simulation results show the improved performance of the proposed method in terms of lower bias and error variance.**

***Index Terms*—Array signal processing, Cramér–Rao bound (CRB), direction-of-arrival (DOA) estimation, temporal correlation.**

I. Introduction

DIRECTION-OF-ARRIVAL (DOA) estimation using an array of sensors is widely investigated in the literature in the last decades. Many methods are developed for diverse conditions (see, e.g., [1]–[5]) and their performances are presented via simulation results or analytical calculations (for papers on theoretical analysis of the performances of the methods, see, e.g., [6]–[9]). Data models play an important role in DOA estimation. They facilitate theoretical derivations of various solutions by their inherent simplistic mathematical and statistical nature. At the same time, overseeing many real-world effects may result in modeling errors and therefore suboptimal methods for the problem at hand. Then, there is a trade-off between simplicity of the models and their usefulness in practice. In the context of array signal processing, there has been considerable research to resolve many aspects of the standard model during the last decades. As a parallel line of research, there also have been efforts to introduce better models which are more application-specific or sometimes more general than the standard model.

In the standard DOA estimation model, snapshots are assumed to be independent and identically distributed (i.i.d.) or uncorrelated in time. In practice, i.i.d. assumption is satisfied with narrowband bandpass filtering of the received signal which gives zero time correlation in certain time delays as ideal sampling points. However, assuming i.i.d. snapshots, places a limitation on the applicability of the results in the real world and also forces some practical difficulties. Assuming perfect bandpass filtering of the signal and exact sampling on the zeros of the correlation function, sampling rate should be reduced so that the subsequent samples be uncorrelated. This is a technical difficulty for environments with slowly fluctuating signals when we need large number of samples. Some maximum-likelihood methods require the number of observations be at least equal to the number of sensors in order to have a full rank sample covariance matrix $\hat{\boldsymbol{R}}$ [10], at the same time many available DOA estimation methods are efficient only asymptotically in number of snapshots [6]. Then, assuming i.i.d. snapshots in theory, results in longer observation times which is not always possible due to moving targets, and may expose us to unpredicted errors due to modeling error. Therefore, it seems useful to accommodate time correlation of at least signals in the model.

Recently, there have been some attempts to tackle with this problem. In [11], authors consider the performance of the spatial covariance-based methods for DOA estimation when signal and noise are not guaranteed to be i.i.d. They conclude that only if noise is still uncorrelated in time, most methods are insensitive to the time correlation of the signals. In [12], an instrumental variable approach (IV-SSF) to the DOA estimation problem in the presence of the time correlation of the signals is proposed. The authors improve their previously presented method in [13] to obtain a more reliable approach. The feature of the IV-SSF is that it does not require any knowledge of the noise covariance matrix but its uncorrelatedness in time. In [12], authors also present a statistical performance evaluation for the IV-SSF and calculate some performance bounds for it. In particular, they derive the CRB for the general case of temporally correlated signals and show that it is lower bounded by the well-known deterministic CRB [14]. In another paper in this field [15], authors show the asymptotic equivalence of spatial and temporal IV-SSF methods in a unified framework.

In this paper, we study the properties of the temporally correlated CRB. Although the effect of temporal correlation of the

Manuscript received April 09, 2008; revised September 02, 2008. First published October 31, 2008; current version published January 30, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Andreas Jakobsson. This work was supported in part by the Advanced Communication Research Institute (ACRI), Sharif University of Technology, and in part by the Iranian Telecommunication Research Center (ITRC) under Contract T/500/20613.

The authors are with the Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran (e-mail: farzanhaddadi@ee.sharif.edu; Nayebi@sharif.edu; Aref@sharif.edu).

Digital Object Identifier 10.1109/TSP.2008.2008220

signals on the conventional methods of DOA estimation is investigated in [11], an important remaining question is the role of temporal correlation of the signals in the best achievable performance. Is it helpful or harmful and how much in various conditions? We investigate this issue and show that it is helpful particularly in low SNRs. We also show that the temporally correlated $\mathrm{CRB}^{\mathrm{cor}}$ is decreasing with the number of samples, as we expect. Then we turn to the IV-SSF to show that it is not an efficient method of DOA estimation in the sense that it cannot achieve the $\mathrm{CRB}^{\mathrm{cor}}$. Using an asymptotical analysis of the $\mathrm{CRB}^{\mathrm{cor}}$, we propose an improved version of the IV-SSF which can outperform the existing version. At the end, simulation results confirm the superiority of the new version in terms of lower finite sample bias and estimation error variance.

The remaining of the paper is organized as follows: Section II presents the data model for the temporally correlated signals array processing. Section III is dedicated to the analysis of the CRB properties and a number of new results and comparisons. In Section IV the optimality of the IV-SSF in comparison with the $\mathrm{CRB}^{\mathrm{cor}}$ is investigated and an improvement is proposed. In Section V, simulation results are presented to show the better performance of the proposed method. Finally, Section VI concludes the paper.

Notation:

$\otimes$	Kronecker product.
$\odot$	Hadamard–Schur product.
$\|\cdot\|$	matrix Frobenious norm.
$(\cdot)^*$	conjugate.
$(\cdot)^{\mathrm{T}}$	transpose.
$(\cdot)^{\mathrm{H}}$	conjugate transpose.
$\mathrm{Tr}(\cdot)$	trace.
$\mathrm{vec}(\cdot)$	vectorizing operator.
$\Re$	real part.
δ_{ij}	Kronecker delta.
$\boldsymbol{a}(\theta_i)$	array steering vector.
$\boldsymbol{A}$	$=[\boldsymbol{a}(\theta_1),\ldots,\boldsymbol{a}(\theta_m)]$; steering matrix.
$\boldsymbol{D}$	$=[\boldsymbol{d}_1,\ldots,\boldsymbol{d}_m]$; $\boldsymbol{d}_i = d\boldsymbol{a}(\theta_i)/d\theta_i$.
$\boldsymbol{\Pi}^{\perp}_{\boldsymbol{A}}$	$\boldsymbol{I}-\boldsymbol{A}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{H}}$; orthogonal projection on to the null space of $\boldsymbol{A}$.
$\boldsymbol{A}\geq\boldsymbol{B}$	$\boldsymbol{A}-\boldsymbol{B}$ is positive semi-definite.
$\boldsymbol{\mathcal{P}}^{ij}$	ijth block of matrix $\boldsymbol{\mathcal{P}}$;
$\boldsymbol{\mathcal{P}}_{ij}$	ijth element of matrix $\boldsymbol{\mathcal{P}}$;
$\mathrm{BTr}_m(\boldsymbol{\mathcal{P}})$	$=\sum_i \boldsymbol{\mathcal{P}}^{ii}_{m\times m}$; block trace;
$\mathrm{Block}_{ij}[\boldsymbol{\mathcal{P}}^{ij}]$	a block matrix with blocks $\boldsymbol{\mathcal{P}}^{ij}$.

II. Data Model

Let an array of L sensors receive n samples of the ambient signal and noise. Signal is composed of plane waves from m distant point sources with directions $\boldsymbol{\theta}=[\theta_1,\ldots,\theta_m]$. Spatial correlation between different sources and temporal correlation of the signals are permitted. Noise is assumed to be temporally white. Then, received data can be modeled as

$$\boldsymbol{x}(t)=\boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{s}(t)+\boldsymbol{\nu}(t) \tag{1}$$

where $\boldsymbol{s}(t)_{m\times 1}$ is the sources signals and $\boldsymbol{\nu}(t)_{L\times 1}$ is the sensors noise vector. Data samples are gathered in matrix form as $\boldsymbol{X}_{L\times n}=[\boldsymbol{x}(t_1),\ldots,\boldsymbol{x}(t_n)]$, and the same is done for the signal sequences $\boldsymbol{S}_{m\times n}=[\boldsymbol{s}(t_1),\ldots,\boldsymbol{s}(t_n)]$ and noise $\boldsymbol{V}_{L\times n}=[\boldsymbol{\nu}(t_1),\ldots,\boldsymbol{\nu}(t_n)]$. Using this notation, the model in (1) can be written as

$$\boldsymbol{X}=\boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{S}+\boldsymbol{V}. \tag{2}$$

Now, we consider the statistical properties of the model. Regarding the statistics of the signal, three models can be assumed: deterministic, i.i.d., and temporally correlated signal models.

1) Deterministic Signal Model: Deterministic signal model assumes a constant signal sequence in every realizations of the process. Then the statistical model of received data will be Gaussian with mean of the signal part and covariance matrix due only to noise

$$\boldsymbol{x}^{\mathrm{det}}(t)\sim\mathcal{N}(\boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{s}(t),\boldsymbol{C}) \tag{3}$$

$$\boldsymbol{C}=\mathrm{E}[\boldsymbol{\nu}(t)\boldsymbol{\nu}^{\mathrm{H}}(t)]. \tag{4}$$

2) i.i.d. Signal Model: In the i.i.d. signal model, signal is a random process with the same distribution in every snapshots and no correlation exists between snapshots. Then, data model for the i.i.d. case is

$$\boldsymbol{x}^{\mathrm{iid}}(t)\sim\mathcal{N}(\boldsymbol{0},\boldsymbol{R}) \tag{5}$$

where

$$\boldsymbol{R}=\boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^{\mathrm{H}}+\boldsymbol{C} \tag{6}$$

and

$$\boldsymbol{P}=\mathrm{E}[\boldsymbol{s}(t)\boldsymbol{s}^{\mathrm{H}}(t)]. \tag{7}$$

3) Correlated Signal Model: In the correlated signal model, noise samples are still assumed to be spatially correlated, distributed as a zero-mean Gaussian random vector $\boldsymbol{\nu}(t)\sim\mathcal{N}(\boldsymbol{0},\boldsymbol{C})$, and temporally uncorrelated with $\mathrm{E}[\boldsymbol{\nu}(t_i)\boldsymbol{\nu}^{\mathrm{H}}(t_j)]=\boldsymbol{0}$, $i\neq j$. Then, a space-time distribution for noise can be defined as

$$\mathrm{vec}(\boldsymbol{V})\sim\mathcal{N}(\boldsymbol{0},\boldsymbol{\mathcal{C}}) \tag{8}$$

$$\boldsymbol{\mathcal{C}}=\boldsymbol{I}_n\otimes\boldsymbol{C}. \tag{9}$$

Similarly, a space-time covariance can be defined for signals. Let $\boldsymbol{\mathcal{P}}^{ij}\triangleq\mathrm{E}\left[\boldsymbol{s}(t_i)\boldsymbol{s}^{\mathrm{H}}(t_j)\right]$, then

$$\mathrm{vec}(\boldsymbol{S}^{\mathrm{cor}})\sim\mathcal{N}(\boldsymbol{0},\boldsymbol{\mathcal{P}}) \tag{10}$$

where $\boldsymbol{\mathcal{P}}_{nm\times nm}=\mathrm{Block}_{ij}\left[\boldsymbol{\mathcal{P}}^{ij}_{m\times m}\right]$. Now, the space-time distribution of the data under correlated signal model can be easily shown to be

$$\mathrm{vec}(\boldsymbol{X}^{\mathrm{cor}})\sim\mathcal{N}(\boldsymbol{0},\boldsymbol{\mathcal{R}}) \tag{11}$$

where

$$\boldsymbol{\mathcal{R}}=\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{P}}\boldsymbol{\mathcal{A}}^{\mathrm{H}}+\boldsymbol{\mathcal{C}} \tag{12}$$

and

$$\boldsymbol{\mathcal{A}}=\boldsymbol{I}_n\otimes\boldsymbol{A}(\boldsymbol{\theta}). \tag{13}$$

Note that (12) has a structure similar to the i.i.d. case (6). It is also assumed that $m<L$ and that the array manifold has the property that every set of distinct steering vectors $\{\boldsymbol{a}(\theta_1),\ldots,\boldsymbol{a}(\theta_L)\}$ forms a linearly independent set. Also, $\boldsymbol{a}(\theta)$ is assumed to be a smooth function as it is in real applications which means that $\boldsymbol{d}(\theta)$ exists. These assumptions pave the way for the estimation problem at hand to be identifiable.

III. Analysis of the CRB

In this section, we present new results and analyzes on the temporally correlated CRB to give more insight on the estimation problem. We are specially interested in the role of temporal correlation of the signals in the direction estimation problem in comparison with the uncorrelated signal case or i.i.d. signals.

CRB is a lower bound on the performance of any unbiased estimation method in terms of error variance. Consider a random vector distributed as $\boldsymbol{y}(t)\sim f(\boldsymbol{\psi})$, where $\boldsymbol{\psi}$ is the vector of possibly unknown parameters of the distribution. Given samples of $\boldsymbol{y}(t)$, an unbiased estimator $\hat{\boldsymbol{\psi}}$ satisfies $\mathrm{E}[\hat{\boldsymbol{\psi}}]=\boldsymbol{\psi}$ and its error variance is lower bounded by the CRB as

$$\mathrm{E}[(\hat{\boldsymbol{\psi}}-\boldsymbol{\psi})(\hat{\boldsymbol{\psi}}-\boldsymbol{\psi})^{\mathrm{H}}]\geq\mathrm{CRB}_{\boldsymbol{\psi},\boldsymbol{\psi}}. \tag{14}$$

CRB is an inherent property of the statistical model of the data and can be calculated directly from $f(\boldsymbol{\psi})$. The importance of the CRB also comes from the fact that there exist estimators that at least asymptotically attain the CRB such as the maximum-likelihood estimator. In the following, CRB for three discussed signal models are repeated for the sake of reference. In the deterministic signal model of (3) and (4) we have [14]

$$\mathrm{CRB}^{\mathrm{det}}_{\boldsymbol{\theta},\boldsymbol{\theta}}=\frac{1}{2n}\left[\Re(\boldsymbol{D}^{\mathrm{H}}\boldsymbol{C}^{-(1/2)}\boldsymbol{\Pi}^{\perp}_0\boldsymbol{C}^{-(1/2)}\boldsymbol{D})\odot\boldsymbol{P}^{\mathrm{T}}\right]^{-1} \tag{15}$$

where

$$\boldsymbol{\Pi}^{\perp}_0=\boldsymbol{\Pi}^{\perp}_{\boldsymbol{C}^{-(1/2)}\boldsymbol{A}}. \tag{16}$$

In the i.i.d. model of (5) and (6) the $\mathrm{CRB}^{\mathrm{iid}}$ is [14], [16]

$$\mathrm{CRB}^{\mathrm{iid}}_{\boldsymbol{\theta},\boldsymbol{\theta}}=\frac{1}{2n}\left[\Re(\boldsymbol{D}^{\mathrm{H}}\boldsymbol{C}^{-(1/2)}\boldsymbol{\Pi}^{\perp}_0\boldsymbol{C}^{-(1/2)}\boldsymbol{D})\odot(\boldsymbol{P}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{A}\,\boldsymbol{P})^{\mathrm{T}}\right]^{-1} \tag{17}$$

and in the temporally correlated signal model of (11), (12), and (13) the $\mathrm{CRB}^{\mathrm{cor}}$ is [12, eq. 112]

$$\mathrm{CRB}^{\mathrm{cor}}_{\boldsymbol{\theta},\boldsymbol{\theta}}=\frac{1}{2}\left[\Re(\boldsymbol{D}^{\mathrm{H}}\boldsymbol{C}^{-(1/2)}\boldsymbol{\Pi}^{\perp}_0\boldsymbol{C}^{-(1/2)}\boldsymbol{D})\odot\mathrm{BTr}_m(\boldsymbol{\mathcal{P}}\boldsymbol{\mathcal{G}}^{\mathrm{H}}\boldsymbol{\mathcal{R}}'^{-1}\boldsymbol{\mathcal{G}}\boldsymbol{\mathcal{P}})^{\mathrm{T}}\right]^{-1} \tag{18}$$

where

$$\boldsymbol{\mathcal{G}}=\boldsymbol{\mathcal{C}}^{-(1/2)}\boldsymbol{\mathcal{A}} \tag{19}$$

$$\boldsymbol{\mathcal{R}}'=\boldsymbol{\mathcal{G}}\boldsymbol{\mathcal{P}}\boldsymbol{\mathcal{G}}^{\mathrm{H}}+\boldsymbol{\mathcal{I}}=\boldsymbol{\mathcal{C}}^{-(1/2)}\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{C}}^{-(1/2)} \tag{20}$$

where $\boldsymbol{\mathcal{I}}=\boldsymbol{I}_n\otimes\boldsymbol{I}_L$. It is noteworthy that the expression of the $\mathrm{CRB}^{\mathrm{cor}}$ in [12] is slightly different from (18) in that it contains a factor of $1/n$ in the right-hand-side. This is a direct consequence of the difference in the definition of the CRB between (14) and what is in [12]. As a confirmation for the form of $\mathrm{CRB}^{\mathrm{cor}}$ in (18), consider the case of i.i.d. signals where the space-time matrices turn out to be block-diagonal. then a factor of $1/n$ appears in the right-hand-side to reduce the $\mathrm{CRB}^{\mathrm{cor}}$ in (18) to the $\mathrm{CRB}^{\mathrm{iid}}$ in (17).

To proceed further, we first give a simplified form of the $\mathrm{CRB}^{\mathrm{cor}}$ in (18). Note that

$$\boldsymbol{\mathcal{G}}^{\mathrm{H}}\boldsymbol{\mathcal{R}}'^{-1}\boldsymbol{\mathcal{G}}=\boldsymbol{\mathcal{A}}^{\mathrm{H}}\boldsymbol{\mathcal{R}}^{-1}\boldsymbol{\mathcal{A}} \tag{21}$$

then the $\mathrm{CRB}^{\mathrm{cor}}$ can be written as

$$\mathrm{CRB}^{\mathrm{cor}}_{\boldsymbol{\theta},\boldsymbol{\theta}}=\frac{1}{2}\left[\Re(\boldsymbol{D}^{\mathrm{H}}\boldsymbol{C}^{-(1/2)}\boldsymbol{\Pi}^{\perp}_0\boldsymbol{C}^{-(1/2)}\boldsymbol{D})\odot\mathrm{BTr}_m(\boldsymbol{\mathcal{P}}\boldsymbol{\mathcal{A}}^{\mathrm{H}}\boldsymbol{\mathcal{R}}^{-1}\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{P}})^{\mathrm{T}}\right]^{-1}. \tag{22}$$

Now, we are ready to state our first result on the comparison of the CRB in the signal models introduced.

Theorem 1: For the deterministic signal model in (3) and (4), i.i.d. signal model in (5) and (6), and the temporally correlated signal model in (11) and (12), $\mathrm{CRB}^{\mathrm{cor}}$ is upper bounded by $\mathrm{CRB}^{\mathrm{iid}}$ and lower bounded by $\mathrm{CRB}^{\mathrm{det}}$

$$\mathrm{CRB}^{\mathrm{det}}\leq\mathrm{CRB}^{\mathrm{cor}}\leq\mathrm{CRB}^{\mathrm{iid}}. \tag{23}$$

Proof: The left side inequality is proved in [12], therefore we give a proof for the right side inequality. We make use of a form of the Woodburry identity [17], [18] which states that

$$(\boldsymbol{A}+\boldsymbol{BCD})^{-1}=\boldsymbol{A}^{-1}-\boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{I}+\boldsymbol{CDA}^{-1}\boldsymbol{B})^{-1}\boldsymbol{CDA}^{-1} \tag{24}$$

which is essentially another form of the matrix inversion lemma. Now we expand the space-time matrix in the block trace of (22) substituting from (12) and making use of (24) to have see equation (25) shown at the bottom of the next page. We factor the common terms of (25) from right side to get

$$[\boldsymbol{\mathcal{I}}-\boldsymbol{\mathcal{P}}\boldsymbol{\mathcal{A}}^{\mathrm{H}}\boldsymbol{\mathcal{C}}^{-1}\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}+\boldsymbol{\mathcal{P}}\boldsymbol{\mathcal{A}}^{\mathrm{H}}\boldsymbol{\mathcal{C}}^{-1}\boldsymbol{\mathcal{A}})^{-1}]\boldsymbol{\mathcal{P}}\boldsymbol{\mathcal{A}}^{\mathrm{H}}\boldsymbol{\mathcal{C}}^{-1}\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{P}}. \tag{26}$$

Add and subtract an $\mathcal{I}$ to the $\mathcal{P}\mathcal{A}^{\mathrm{H}}\mathcal{C}^{-1}\mathcal{A}$ in the bracket in (26) to get

$$(\mathcal{I}+\mathcal{P}\mathcal{A}^{\mathrm{H}}\mathcal{C}^{-1}\mathcal{A})^{-1}\mathcal{P}\mathcal{A}^{\mathrm{H}}\mathcal{C}^{-1}\mathcal{A}\mathcal{P}. \tag{27}$$

Do the same for the term outside the bracket in (27) in the following form:

$$(\mathcal{I}+\mathcal{P}\mathcal{A}^{\mathrm{H}}\mathcal{C}^{-1}\mathcal{A})^{-1}(\mathcal{I}+\mathcal{P}\mathcal{A}^{\mathrm{H}}\mathcal{C}^{-1}\mathcal{A}-\mathcal{I})\mathcal{P}. \tag{28}$$

As a useful result, we arrive to the following equality simplifying (28)

$$\mathcal{P}\mathcal{A}^{\mathrm{H}}\mathcal{R}^{-1}\mathcal{A}\mathcal{P}=\mathcal{P}-(\mathcal{P}^{-1}+\mathcal{A}^{\mathrm{H}}\mathcal{C}^{-1}\mathcal{A})^{-1}. \tag{29}$$

Note that the matrix in parenthesis is positive semi-definite, then we will have

$$\mathcal{P}\mathcal{A}^{\mathrm{H}}\mathcal{R}^{-1}\mathcal{A}\mathcal{P}\le\mathcal{P} \tag{30}$$

which is another proof for the left side inequality in (23) (see [12]). Now, we proceed to prove the right side inequality in (23). define

$$\mathcal{P}_d\triangleq\mathrm{Block}_{ij}[\mathcal{P}^{ij}\delta_{ij}] \tag{31}$$

which is the block-diagonalized version of the signal space-time covariance matrix. Note that block-diagonal space-time covariance matrix for signals $\mathcal{P}_d$, represents the temporally uncorrelated signal model (more general than i.i.d. signal model), while $\mathcal{P}$ represents the temporally correlated signal model. Now, assume a matrix $\mathcal{F}$ and its block-diagonalized version $\mathcal{F}_d$. It is well known that

$$(\mathcal{F}^{-1})^{ii}\ge\left(\mathcal{F}^{ii}\right)^{-1}=\left(\mathcal{F}_d^{ii}\right)^{-1}=\left(\mathcal{F}_d^{-1}\right)^{ii}. \tag{32}$$

Using the inequality in (32), besides the fact that $\mathcal{A}^{\mathrm{H}}\mathcal{C}^{-1}\mathcal{A}$ is block-diagonal, we will have

$$\begin{aligned}\mathcal{A}^{\mathrm{H}}\mathcal{C}^{-1}\mathcal{A}-(\mathcal{P}+(\mathcal{A}^{\mathrm{H}}\mathcal{C}^{-1}\mathcal{A})^{-1})^{-1}\\ \le\mathcal{A}^{\mathrm{H}}\mathcal{C}^{-1}\mathcal{A}-(\mathcal{P}_d+(\mathcal{A}^{\mathrm{H}}\mathcal{C}^{-1}\mathcal{A})^{-1})^{-1}.\end{aligned} \tag{33}$$

Now, we multiply $\mathcal{B}\triangleq(\mathcal{A}^{\mathrm{H}}\mathcal{C}^{-1}\mathcal{A})^{-1}$ from right and left of both sides of the inequality to get

$$\mathcal{B}-\mathcal{B}(\mathcal{P}+\mathcal{B})^{-1}\mathcal{B}\le\mathcal{B}-\mathcal{B}(\mathcal{P}_d+\mathcal{B})^{-1}\mathcal{B}. \tag{34}$$

The above expressions can be simplified using matrix inversion lemma to give

$$(\mathcal{P}^{-1}+\mathcal{A}^{\mathrm{H}}\mathcal{C}^{-1}\mathcal{A})^{-1}\le\left(\mathcal{P}_d^{-1}+\mathcal{A}^{\mathrm{H}}\mathcal{C}^{-1}\mathcal{A}\right)^{-1}. \tag{35}$$

Applying the block trace operator and adding a common term results in

$$\begin{aligned}\mathrm{BTr}_m[\mathcal{P}-(\mathcal{P}^{-1}+\mathcal{A}^{\mathrm{H}}\mathcal{C}^{-1}\mathcal{A})^{-1}]\\ \ge\mathrm{BTr}_m\left[\mathcal{P}_d-\left(\mathcal{P}_d^{-1}+\mathcal{A}^{\mathrm{H}}\mathcal{C}^{-1}\mathcal{A}\right)^{-1}\right].\end{aligned} \tag{36}$$

According to (29), the inequality in (36) implies that

$$\mathrm{BTr}_m[\mathcal{P}\mathcal{A}^{\mathrm{H}}\mathcal{R}^{-1}\mathcal{A}\mathcal{P}]\ge\mathrm{BTr}_m\left[\mathcal{P}_d\mathcal{A}^{\mathrm{H}}\mathcal{R}_d^{-1}\mathcal{A}\mathcal{P}_d\right]. \tag{37}$$

which in fact completes the proof of (23) showing that

$$\mathrm{CRB}^{\mathrm{cor}}(\mathcal{P})\le\mathrm{CRB}^{\mathrm{cor}}(\mathcal{P}_d). \tag{38}$$

□

The upper bound and lower bound on the $\mathrm{CRB}^{\mathrm{cor}}$ presented in (23) is very insightful to the estimation problem at hand. It implies that the existence of a temporal correlation in the signals improve the best attainable performance of estimation. The comparison made in theorem 1 is conditioned on the specific spatial covariance matrix of the sources in each sample i.e., with the same diagonal blocks of the space-time signal covariance matrix. It shows that adding a nondiagonal covariance block improves the performance since it simplifies the extraction of the signal part from received data. Most spatial covariance-based DOA estimation methods rely on the different spatial characteristics of the signal and noise (signals are point sources of radiation while noise is uniformly distributed in the space or at least is spread via large areas). This is not the case in the temporally correlated signal model where there is a particular difference between signal and noise in that one is temporally correlated and the other is temporally uncorrelated. This increased distance of the signal model and noise model improves the CRB and the performance of the methods using it. As a result, we can see that the $\mathrm{CRB}^{\mathrm{cor}}$ is lower than the $\mathrm{CRB}^{\mathrm{iid}}$ which implies better performance when noise and signal have different temporal characteristics. Another explanation for the inequalities presented in (23) comes from the degree of predictability of the signal. It is obvious that in the deterministic signal model we have a statistically constant signal which is fully predictable. Therefore, the performance is best in the deterministic signal model. In the correlated signal model, signals are stochastic in nature and vary in each realization, which makes the signals less predictable. Though, the existence of the temporal correlation of the signals offers a limited possibility for coarse signal prediction and extraction from noise which places the $\mathrm{CRB}^{\mathrm{cor}}$ lower than the completely uncorrelated case of $\mathrm{CRB}^{\mathrm{iid}}$.

Although we have confined the $\mathrm{CRB}^{\mathrm{cor}}$ between $\mathrm{CRB}^{\mathrm{det}}$ and $\mathrm{CRB}^{\mathrm{iid}}$ in theorem 1, we are interested to more exactly specify the behavior of the $\mathrm{CRB}^{\mathrm{cor}}$ in different situations. This helps us to get more insight to the role of temporal correlation

---

$$\begin{aligned}\mathcal{P}\mathcal{A}^{\mathrm{H}}\mathcal{R}^{-1}\mathcal{A}\mathcal{P}&=\mathcal{P}\mathcal{A}^{\mathrm{H}}[\mathcal{C}^{-1}-\mathcal{C}^{-1}\mathcal{A}(\mathcal{I}+\mathcal{P}\mathcal{A}^{\mathrm{H}}\mathcal{C}^{-1}\mathcal{A})^{-1}\mathcal{P}\mathcal{A}^{\mathrm{H}}\mathcal{C}^{-1}]\mathcal{A}\mathcal{P}\\ &=\mathcal{P}\mathcal{A}^{\mathrm{H}}\mathcal{C}^{-1}\mathcal{A}\mathcal{P}-\mathcal{P}\mathcal{A}^{\mathrm{H}}\mathcal{C}^{-1}\mathcal{A}(\mathcal{I}+\mathcal{P}\mathcal{A}^{\mathrm{H}}\mathcal{C}^{-1}\mathcal{A})^{-1}\mathcal{P}\mathcal{A}^{\mathrm{H}}\mathcal{C}^{-1}\mathcal{A}\mathcal{P}.\end{aligned} \tag{25}$$

of the signals in DOA estimation. Therefore, we consider approximations of the $\mathrm{CRB}^{\mathrm{cor}}$ in different situations in terms of signal-to-noise ratio (SNR) in the following theorem.

Theorem 2: In the high SNR condition, temporal correlation of the signals makes no improvement on the uncorrelated signal model while in the low SNR condition, the contribution of zero-lag and nonzero-lag covariances are the same, i.e., nonzero-lag covariances improve the CRB.

$$\mathrm{SNR} \gg 1 : \mathrm{CRB}^{\mathrm{cor}} \simeq \mathrm{CRB}^{\mathrm{iid}} \tag{39}$$

$$\mathrm{SNR} \ll 1 : \mathrm{CRB}^{\mathrm{cor}} < \mathrm{CRB}^{\mathrm{iid}}. \tag{40}$$

Proof: Consider the high SNR condition. We make use of the following easily checked approximation for any appropriately sized matrices $\boldsymbol{B}$ and $\boldsymbol{\Delta}$ if $\|\boldsymbol{\Delta}\| \leq \|\boldsymbol{B}\|$

$$(\boldsymbol{B}+\boldsymbol{\Delta})^{-1} \simeq \boldsymbol{B}^{-1} - \boldsymbol{B}^{-1}\boldsymbol{\Delta}\boldsymbol{B}^{-1} + \boldsymbol{B}^{-1}\boldsymbol{\Delta}\boldsymbol{B}^{-1}\boldsymbol{\Delta}\boldsymbol{B}^{-1} - \cdots. \tag{41}$$

Now we can expand the block trace in $\mathrm{CRB}^{\mathrm{cor}}$ as follows

$$\mathrm{BTr}_m[\boldsymbol{\mathcal{P}}\boldsymbol{\mathcal{A}}^{\mathrm{H}}\boldsymbol{\mathcal{R}}^{-1}\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{P}}] = \mathrm{BTr}_m[\boldsymbol{\mathcal{P}}\boldsymbol{\mathcal{A}}^{\mathrm{H}}(\boldsymbol{\mathcal{C}}+\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{P}}\boldsymbol{\mathcal{A}}^{\mathrm{H}})^{-1}\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{P}}]. \tag{42}$$

We use matrix inversion lemma for the inverse in the block trace of (42) to get

$$\mathrm{BTr}_m[\boldsymbol{\mathcal{P}}\boldsymbol{\mathcal{A}}^{\mathrm{H}}(\boldsymbol{\mathcal{C}}^{-1} - \boldsymbol{\mathcal{C}}^{-1}\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{P}}^{-1}+\boldsymbol{\mathcal{A}}^{\mathrm{H}}\boldsymbol{\mathcal{C}}^{-1}\boldsymbol{\mathcal{A}})^{-1}\boldsymbol{\mathcal{A}}^{\mathrm{H}}\boldsymbol{\mathcal{C}}^{-1})\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{P}}]. \tag{43}$$

Now we use the approximation in (41) up to the third term. Note that the high SNR assumption guarantees that $\|\boldsymbol{\mathcal{P}}^{-1}\| \leq \|\boldsymbol{\mathcal{A}}^{\mathrm{H}}\boldsymbol{\mathcal{C}}^{-1}\boldsymbol{\mathcal{A}}\|$. After some calculations, we get

$$\mathrm{BTr}_m[\boldsymbol{\mathcal{P}}\boldsymbol{\mathcal{A}}^{\mathrm{H}}\boldsymbol{\mathcal{R}}^{-1}\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{P}}] \simeq \mathrm{BTr}_m[\boldsymbol{\mathcal{P}} - (\boldsymbol{\mathcal{A}}^{\mathrm{H}}\boldsymbol{\mathcal{C}}^{-1}\boldsymbol{\mathcal{A}})^{-1}] \tag{44}$$

in the high SNR region and to the second order of approximation. The approximation in (44) asserts that in high SNR condition, only block trace of the signal space-time covariance matrix contribute to the Cramér–Rao bound, hence we can conclude that in this case, the temporal correlation of the signals (represented by nondiagonal blocks of $\boldsymbol{\mathcal{P}}$), do not improve the best achievable performance of the DOA estimation and (39) follows.

In the very low SNR region, we use the approximation $\boldsymbol{\mathcal{R}} \simeq \boldsymbol{\mathcal{C}}$. Then expanding the block trace in the CRB gives

$$\mathrm{BTr}_m[\boldsymbol{\mathcal{P}}\boldsymbol{\mathcal{A}}^{\mathrm{H}}\boldsymbol{\mathcal{R}}^{-1}\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{P}}] \simeq \sum_{i,r}(\boldsymbol{A}\boldsymbol{\mathcal{P}}^{ri})^{\mathrm{H}}\boldsymbol{C}^{-1}(\boldsymbol{A}\boldsymbol{\mathcal{P}}^{ri}). \tag{45}$$

We can see from (45) that every blocks of $\boldsymbol{\mathcal{P}}$ contribute the same to the $\mathrm{CRB}^{\mathrm{cor}}$ in the situation of very low SNR. Since each term in the summation of (45) is positive semi-definite, then each nondiagonal block of $\boldsymbol{\mathcal{P}}$ improves the $\mathrm{CRB}^{\mathrm{cor}}$ making the distance from $\mathrm{CRB}^{\mathrm{iid}}$ larger which implies (40). It is noteworthy that improvement of the CRB does not mean the performance improvement of the conventional DOA estimation methods, rather it clarifies the existence of methods that can achieve better performances through making use of the temporal characteristics of the signals and noise. □

Now, after we considered the role of the temporal correlation of the signals in the best achievable performance of DOA estimation, we turn to an assumed characteristic of the CRB. The CRB usually decreases with increased amount of data. This is obvious in the i.i.d. signal models where the CRB for n data samples is $1/n$ of the CRB for one sample. Though, this is not very clear for the temporally correlated signal model where the dependence of the $\mathrm{CRB}^{\mathrm{cor}}$ on n is embedded in the size of the space-time matrices.

Theorem 3: $\mathrm{CRB}^{\mathrm{cor}}$ is decreasing with increasing n:

$$\mathrm{CRB}^{\mathrm{cor}}(n+1) < \mathrm{CRB}^{\mathrm{cor}}(n). \tag{46}$$

Proof: Applying matrix inversion lemma, we can show that

$$\boldsymbol{\mathcal{A}}^{\mathrm{H}}\boldsymbol{\mathcal{R}}^{-1}\boldsymbol{\mathcal{A}} = (\boldsymbol{\mathcal{P}} + (\boldsymbol{\mathcal{A}}^{\mathrm{H}}\boldsymbol{\mathcal{C}}^{-1}\boldsymbol{\mathcal{A}})^{-1})^{-1}. \tag{47}$$

Assuming $\boldsymbol{\mathcal{P}}_{n+1}$ as the block-diagonally augmented version of $\boldsymbol{\mathcal{P}}_n$ and using (47), we will have

$$\mathrm{CRB}^{\mathrm{cor}}\left(\begin{bmatrix}\boldsymbol{\mathcal{P}}_n & \\ & \boldsymbol{P}_{n+1}\end{bmatrix}\right) \leq \mathrm{CRB}^{\mathrm{cor}}(\boldsymbol{\mathcal{P}}_n). \tag{48}$$

since

$$\begin{aligned}\mathrm{BTr}_m&\left[\boldsymbol{\mathcal{P}}_{n+1}\boldsymbol{\mathcal{A}}_{n+1}^{\mathrm{H}}\boldsymbol{\mathcal{R}}_{n+1}^{-1}\boldsymbol{\mathcal{A}}_{n+1}\boldsymbol{\mathcal{P}}_{n+1}\right] \\ &= \mathrm{BTr}_m\left[\boldsymbol{\mathcal{P}}_n\left(\boldsymbol{\mathcal{P}}_n + \left(\boldsymbol{\mathcal{A}}_n^{\mathrm{H}}\boldsymbol{\mathcal{C}}_n^{-1}\boldsymbol{\mathcal{A}}_n\right)^{-1}\right)^{-1}\boldsymbol{\mathcal{P}}_n\right] \\ &\quad + \boldsymbol{P}_{n+1}\left(\boldsymbol{P}_{n+1} + \left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{C}^{-1}\boldsymbol{A}\right)^{-1}\right)^{-1}\boldsymbol{P}_{n+1}\end{aligned} \tag{49}$$

and the second term is positive semi-definite. Following the same steps as in (33) to (38), we can also show the following inequality which completes the proof:

$$\mathrm{CRB}^{\mathrm{cor}}\left(\begin{bmatrix}\boldsymbol{\mathcal{P}}_n & \boldsymbol{\mathcal{Q}}^{\mathrm{H}} \\ \boldsymbol{\mathcal{Q}} & \boldsymbol{P}_{n+1}\end{bmatrix}\right) \leq \mathrm{CRB}^{\mathrm{cor}}\left(\begin{bmatrix}\boldsymbol{\mathcal{P}}_n & \\ & \boldsymbol{P}_{n+1}\end{bmatrix}\right). \tag{50}$$

Note that (50) is a generalization of the right-hand-side inequality in (23). □

In this section, we investigated the general properties of the DOA estimation problem in the presence of the temporally correlated signals. We performed this via Cramér–Rao bound analysis and characterization. In Section IV, we turn to the practical methods to accomplish DOA estimation under temporally correlated signal model.

IV. Suboptimality of the IV-SSF Method

We have considered the best achievable performance in the temporally correlated signal model. In particular, we found that it is possible to improve the performance of the conventional spatial covariance-based methods by devising new methods that can exploit the temporal correlation of the signals. However, the $\mathrm{CRB}^{\mathrm{cor}}$ has been calculated under the assumption of known noise spatial covariance matrix. When this is not true, we do not expect any method to reach the $\mathrm{CRB}^{\mathrm{cor}}$ in performance. In the situation of unknown noise spatial covariance matrix, the instrumental variable subspace fitting (IV-SSF) method for DOA estimation has been proposed in [12]. The method is based on the instrumental variable approach. It makes an instrumental variables vector $\boldsymbol{\phi}(t)$ for each data sample $\boldsymbol{x}(t)$ in such a way that the cross-covariance of $\boldsymbol{\phi}(t)$ and $\boldsymbol{x}(t)$ does not contain the unknown

noise covariance matrix. Then, a sample cross-covariance can be used to extract the signal subspace and parameters of interest. Note that the signal temporal correlation leave a room for the multiplication of noncontemporary data to contain information about the directions of arrival. The instrumental variables vector for each data sample $\boldsymbol{x}(t)$ is formed as

$$\boldsymbol{\phi}(t) \triangleq \begin{bmatrix} \boldsymbol{x}(t-1) \\ \vdots \\ \boldsymbol{x}(t-M) \end{bmatrix} \tag{51}$$

where M is a user defined integer determining the degree of complexity and hence the performance of the method. In general, larger M should result in better estimates, although simulation results show increased bias and decreased error variance when M becomes larger. Let the cross-covariance of the signals at time lag k be defined as

$$\boldsymbol{P}_k \triangleq \mathrm{E}[\boldsymbol{s}(t-k)\boldsymbol{s}^{\mathrm{H}}(t)] \tag{52}$$

and define for convenience

$$\boldsymbol{\mathcal{J}} = \begin{bmatrix} \boldsymbol{P}_1 \\ \vdots \\ \boldsymbol{P}_M \end{bmatrix} \tag{53}$$

then the cross-covariance of the received data and the corresponding instrumental variable will be

$$\boldsymbol{\Sigma} \triangleq \mathrm{E}[\boldsymbol{\phi}(t)\boldsymbol{x}^{\mathrm{H}}(t)] = (\boldsymbol{I}_M \otimes \boldsymbol{A})\boldsymbol{\mathcal{J}}\boldsymbol{A}^{\mathrm{H}} \tag{54}$$

which is independent of the unknown noise spatial covariance matrix $\boldsymbol{C}$. Also define the instrumental variable covariance matrix as

$$\boldsymbol{\Phi} \triangleq \mathrm{E}[\boldsymbol{\phi}(t)\boldsymbol{\phi}^{\mathrm{H}}(t)]. \tag{55}$$

The estimates of the DOAs in IV-SSF method are obtained in the following steps: choose $M > 1$ and compute the sample estimates

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{n-M}\sum_{t=M+1}^{n} \boldsymbol{\phi}(t)\boldsymbol{x}^{\mathrm{H}}(t) \tag{56}$$

$$\hat{\boldsymbol{\Phi}} = \frac{1}{n-M}\sum_{t=M+1}^{n} \boldsymbol{\phi}(t)\boldsymbol{\phi}^{\mathrm{H}}(t). \tag{57}$$

Next, extract $\hat{\boldsymbol{R}}_0$ from $\hat{\boldsymbol{\Phi}}$, as one of the $m \times m$ diagonal blocks. The estimates of the parameters are the minimizer of the following criterion function:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \ \mathrm{Tr}\left(\hat{\boldsymbol{\Pi}}_0^{\perp}\hat{\boldsymbol{R}}_0^{-(1/2)}\hat{\boldsymbol{V}}_s\hat{\boldsymbol{\Phi}}_s^2\hat{\boldsymbol{V}}_s^{\mathrm{H}}\hat{\boldsymbol{R}}_0^{-(1/2)}\right) \tag{58}$$

where

$$\hat{\boldsymbol{\Pi}}_0^{\perp} = \boldsymbol{I} - \hat{\boldsymbol{R}}_0^{-(1/2)}\boldsymbol{A}\left(\boldsymbol{A}^{\mathrm{H}}\hat{\boldsymbol{R}}_0^{-1}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathrm{H}}\hat{\boldsymbol{R}}_0^{-(1/2)} \tag{59}$$

and $\hat{\boldsymbol{V}}_s$ contains the dominant right singular vectors of the matrix $\hat{\boldsymbol{\Phi}}^{(-1/2)}\hat{\boldsymbol{\Sigma}}$, while the associated singular values are gathered in matrix $\hat{\boldsymbol{\Phi}}_s$.

Using the IV-SSF method, the asymptotic error covariance matrix has been shown to be [12]

$$\mathrm{Cov}^{\mathrm{IV-SSF}} = \frac{1}{2n}\left[\Re\left(\boldsymbol{D}^{\mathrm{H}}\boldsymbol{C}^{-(1/2)}\boldsymbol{\Pi}_0^{\perp}\boldsymbol{C}^{-(1/2)}\boldsymbol{D}\right) \odot \left(\boldsymbol{\mathcal{J}}^{\mathrm{H}}\boldsymbol{\mathcal{A}}_M^{\mathrm{H}}\boldsymbol{\Phi}^{-1}\boldsymbol{\mathcal{A}}_M\boldsymbol{\mathcal{J}}\right)^{\mathrm{T}}\right]^{-1} \tag{60}$$

in which $\boldsymbol{\mathcal{A}}_M = \boldsymbol{I}_M \otimes \boldsymbol{A}$. The covariance in (60) is slightly different from the $\mathrm{CRB}^{\mathrm{cor}}$ in (22). We aim to show that the IV-SSF method with above error covariance does not attain the $\mathrm{CRB}^{\mathrm{cor}}$ asymptotically and relying on this analysis, we propose a variation in the method to boost its performance. To this end, we use an asymptotic analysis presented in [12] with some modifications to be able to compare the $\mathrm{CRB}^{\mathrm{cor}}$ and the covariance in (60). Define an estimation problem in which we are interested to estimate the signal $\boldsymbol{s}(t)$ using the data $\boldsymbol{z}(t) = \boldsymbol{s}(t) + \boldsymbol{w}(t)$, in which $\boldsymbol{w}(t)$ is a temporally white noise term with spatial covariance matrix $(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{C}^{-1}\boldsymbol{A})^{-1}$. Note that $\boldsymbol{w}(t)$ is temporally white since it is constructed from i.i.d. noise term $\boldsymbol{\nu}(t)$ by a linear transform in the space domain. Assume that we are to estimate $\boldsymbol{s}(t)$ from M previous samples of $\boldsymbol{z}(t)$

$$\boldsymbol{z}_M(t) = [\,\boldsymbol{z}^{\mathrm{T}}(t-1) \quad \cdots \quad \boldsymbol{z}^{\mathrm{T}}(t-M)\,]^{\mathrm{T}}. \tag{61}$$

with the best linear transform which minimizes the error covariance matrix. The estimation error will be

$$\boldsymbol{s}(t) - \boldsymbol{\mathcal{H}}\boldsymbol{z}_M(t) = \boldsymbol{e}(t) \tag{62}$$

where $\boldsymbol{\mathcal{H}}$ is the best linear transform in the least squares sense. Making use of the orthogonality principle, we can show that

$$\hat{\boldsymbol{\mathcal{H}}} = \boldsymbol{\mathcal{J}}^{\mathrm{H}}\left(\boldsymbol{\mathcal{P}}_M + \left(\boldsymbol{\mathcal{A}}_M^{\mathrm{H}}\boldsymbol{\mathcal{C}}_M^{-1}\boldsymbol{\mathcal{A}}_M\right)^{-1}\right)^{-1} \tag{63}$$

where $\boldsymbol{\mathcal{P}}_M$ is the space-time covariance matrix of the signals in M snapshots and $\boldsymbol{\mathcal{C}}_M = \boldsymbol{I}_M \otimes \boldsymbol{C}$. The resulting minimized error covariance matrix will be

$$\boldsymbol{\Sigma}_e = \boldsymbol{P}_0 - \boldsymbol{\mathcal{J}}^{\mathrm{H}}\left(\boldsymbol{\mathcal{P}}_M + \left(\boldsymbol{\mathcal{A}}_M^{\mathrm{H}}\boldsymbol{\mathcal{C}}_M^{-1}\boldsymbol{\mathcal{A}}_M\right)^{-1}\right)^{-1}\boldsymbol{\mathcal{J}} \tag{64}$$

which along with a readily shown result similar to (47) reduces the second part of the error covariance matrix of IV-SSF method in (60) to

$$\begin{aligned}\boldsymbol{\mathcal{J}}^{\mathrm{H}}\boldsymbol{\mathcal{A}}_M^{\mathrm{H}}\boldsymbol{\Phi}^{-1}\boldsymbol{\mathcal{A}}_M\boldsymbol{\mathcal{J}} &= \boldsymbol{\mathcal{J}}^{\mathrm{H}}\left(\boldsymbol{\mathcal{P}}_M + \left(\boldsymbol{\mathcal{A}}_M^{\mathrm{H}}\boldsymbol{\mathcal{C}}_M^{-1}\boldsymbol{\mathcal{A}}_M\right)^{-1}\right)^{-1}\boldsymbol{\mathcal{J}} \\ &= \boldsymbol{P}_0 - \boldsymbol{\Sigma}_e.\end{aligned} \tag{65}$$

The error covariance matrix in (65) has the implication that the minimum attainable error covariance in the IV-SSF method depends on the minimum error of prediction of the process $\boldsymbol{s}(t)$ in the noise term $\boldsymbol{w}(t)$ using M previous samples of the data $\boldsymbol{z}(t)$. Lowering the prediction error will cause the error covariance of IV-SSF in (60) to reduce. Now we use an asymptotic analysis

on the $\mathrm{CRB}^{\mathrm{cor}}$ to reform it in a form similar to (65) which enables us to understand why IV-SSF method does not reach the $\mathrm{CRB}^{\mathrm{cor}}$ and propose a modification in the method to improve its performance. Assume that large number of data is available $n \to +\infty$. The block trace in the $\mathrm{CRB}^{\mathrm{cor}}$ in (22) can be written as

$$\mathrm{BTr}_m[\mathcal{P}\mathcal{A}^{\mathrm{H}}\mathcal{R}^{-1}\mathcal{A}\mathcal{P}] = \sum_{i=1}^{n} \mathcal{P}^{i\cdot}\mathcal{A}^{\mathrm{H}}\mathcal{R}^{-1}\mathcal{A}\mathcal{P}^{\cdot i} \tag{66}$$

where $\mathcal{P}^{i\cdot}$ and $\mathcal{P}^{\cdot i}$ denote the ith block row and column of $\mathcal{P}$, respectively. We assume a stationary signal model in this section which results in a block Toeplitz signal space-time covariance $\mathcal{P}$. Further, we assume a regular signal random process in which the cross-covariance decreases with increasing time lag. For such asymptotic conditions we can see that block rows $\mathcal{P}^{i\cdot}$ are shifted versions of each other, ignoring the first and last ones. The same is true for block columns $\mathcal{P}^{\cdot i}$ and also for the matrix $\mathcal{R}^{-1}$ and hence $\mathcal{A}^{\mathrm{H}}\mathcal{R}^{-1}\mathcal{A}$, i.e., block rows (and columns) of these matrices are shifted versions of each other. This special matrix multiplication form ensures that the terms in summation in (66) are approximately equal. Then, (66) can be reduced to

$$\mathrm{BTr}_m[\mathcal{P}\mathcal{A}^{\mathrm{H}}\mathcal{R}^{-1}\mathcal{A}\mathcal{P}] \simeq n\mathcal{P}'^{\mathrm{T}}\mathcal{A}^{\mathrm{H}}\mathcal{R}^{-1}\mathcal{A}\mathcal{P}' \tag{67}$$

where

$$\mathcal{P}' = [\cdots \quad \boldsymbol{P}_{-1}^{\mathrm{T}} \quad \boldsymbol{P}_{0}^{\mathrm{T}} \quad \boldsymbol{P}_{+1}^{\mathrm{T}} \quad \cdots]^{\mathrm{T}}. \tag{68}$$

Now, using a similar analysis as in (61) to (65) with a data vector defined as

$$\boldsymbol{z}_{\mathrm{crb}}(t) = [\cdots z^{\mathrm{T}}(t-1) \quad z^{\mathrm{T}}(t) \quad z^{\mathrm{T}}(t+1)\cdots]^{\mathrm{T}} \tag{69}$$

results in an approximation similar to (65) for the second part of the $\mathrm{CRB}^{\mathrm{cor}}$

$$\mathcal{P}'\mathcal{A}^{\mathrm{H}}\mathcal{R}^{-1}\mathcal{A}\mathcal{P}' = P_0 - \Sigma_\epsilon \tag{70}$$

in which $\boldsymbol{\Sigma}_\epsilon$ is the error covariance of the *smoothing* of the random process $\boldsymbol{z}(t)$, i.e., estimating the signal part $\boldsymbol{s}(t)$ from the data described in (69). Now, after that we have transformed the matrix forms of the $\mathrm{CRB}^{\mathrm{cor}}$ and the error of the IV-SSF method to the same formats, it is possible to see that the IV-SSF method will not attain the $\mathrm{CRB}^{\mathrm{cor}}$ since $\boldsymbol{\Sigma}_\epsilon < \boldsymbol{\Sigma}_e$. Obviously, it is easier and results in lower error to estimate the signal $\boldsymbol{s}(t)$ using $\boldsymbol{z}_{\mathrm{crb}}(t)$ in (69) rather than using $\boldsymbol{z}_M(t)$ in (61) because of larger amount of correlated data available. Using this analysis, we can understand what is required to improve the performance of the IV-SSF method. The key observation is that the structure of the estimation problem and $\boldsymbol{z}_M(t)$ is similar to the structure of the $\boldsymbol{\phi}(t)$, the instrumental variable chosen for the IV-SSF. It shows that we could achieve the best performance if we could have defined an instrumental variable containing all data before and after the present signal $\boldsymbol{s}(t)$. Though, we cannot include $\boldsymbol{x}(t)$ in the instrumental variables since it requires the knowledge of the noise covariance matrix. Also, the limited degree of complexity we afford for our method do not permit the inclusion of too many instrumental variables. We are free to choose M instrumental variables and we propose the following:

$$\boldsymbol{\phi}_{pro}(t) \triangleq \begin{bmatrix} \boldsymbol{x}\left(t+\frac{M}{2}\right) \\ \vdots \\ \boldsymbol{x}(t+1) \\ \boldsymbol{x}(t-1) \\ \vdots \\ \boldsymbol{x}\left(t-\frac{M}{2}\right) \end{bmatrix}. \tag{71}$$

In the subsequent section, we will present simulation results which confirm the improvement in the performance of the IV-SSF method using the above two-sided instrumental variables.

V. Simulation Results

In this section, we illustrate the $\mathrm{CRB}^{\mathrm{cor}}$ and the performance of the two methods investigated: the IV-SSF method and our proposed method which we call two-sided IV-SSF. We also present simulation results that confirm the inequality in theorem 1 Performance of the methods is considered in the sense of error variance and bias. Although the IV-SSF method shows to be asymptotically $(n \to +\infty)$ unbiased and this is shown theoretically in [12], in the nonasymptotic region of limited available data, it shows strong bias. We will see that the two-sided IV-SSF outperforms IV-SSF in both lower bias and lower error variance.

We first present some graphs on the Cramér–Rao bounds. The main result we provided on the CRBs is the inequality in (23). We also showed that the temporally correlated CRB is decreasing with the number of snapshots n. These facts are presented via an example of two-source scenario in Fig. 1. Conditions of the simulation are as follows: Two sources at angles $[0 \quad 0.2]$ radians with respect to the array broadside are present. Array number of elements is 3 with $\lambda/2$ spacing and $\mathrm{SNR} = 10$ dB. Uniform linear array is assumed and the signal covariance matrix is defined as

$$\mathcal{P} = \boldsymbol{P}_t \otimes \boldsymbol{P}_s \tag{72}$$

where $[\boldsymbol{P}_t]_{ij} = e^{-0.2|i-j|}$, $[\boldsymbol{P}_s]_{ij} = e^{-0.5|i-j|}$, and $\boldsymbol{C}_{ij} = \sigma^2 e^{-|i-j|}$. Fig. 1 shows the decreasing CRB versus increasing n, it also confirms the inequality in (23) for this special case. It is also noteworthy that the difference between three types of CRBs increases with n since with increased number of snapshots, there are more room for the temporal correlation to improve the DOA estimation performance. In Fig. 2, CRBs are depicted versus SNR. It can be seen that as SNR increases, three types of CRB merge together and decrease linearly. Though, the $\mathrm{CRB}^{\mathrm{cor}}$ joins the $\mathrm{CRB}^{\mathrm{iid}}$ faster than $\mathrm{CRB}^{\mathrm{det}}$. This is better shown in Fig. 3, where the CRBs are normalized to the $\mathrm{CRB}^{\mathrm{iid}}$ to confirm the result of theorem 2 which is stated in (39).

Now, we present simulation results which support the proposed method versus the main IV-SSF. In the simulations, single source located in $\omega \triangleq 2\pi(d/\lambda)\cos(\theta) = 0.8$ is assumed, with a four-element half-wavelength array, number of instrumental variables $M = 2$, and number of trials 10 000. In Fig. 4, performance measures, bias and standard deviation of the estimate of ω are presented versus the number of snapshots

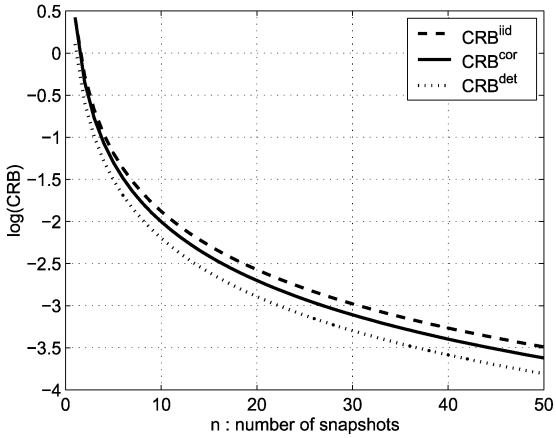

Fig. 1. Example of the CRB versus the number of snapshots in logarithmic scale for two sources. The temporally correlated CRB is upper bounded by the i.i.d. CRB and lower bounded by the deterministic CRB.

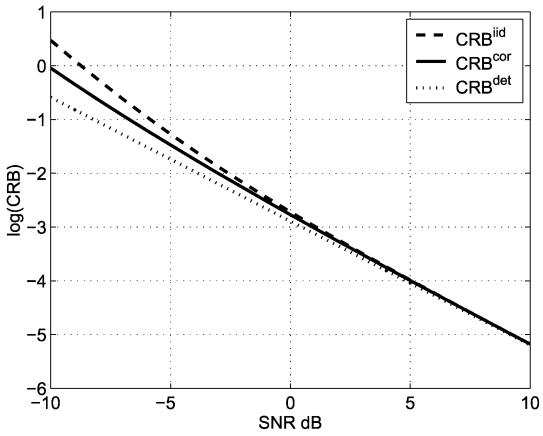

Fig. 2. Example of the CRB versus the SNR in logarithmic scale for two sources. The temporally correlated CRB tends to the i.i.d. CRB as SNR increases.

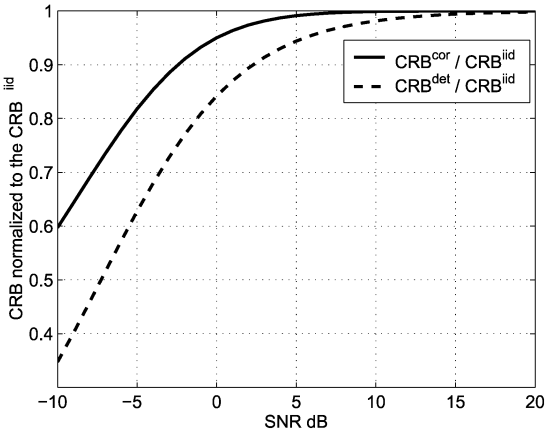

Fig. 3. Comparison of the temporally correlated CRB and deterministic CRB normalized to the i.i.d. CRB. The comparison shows the equivalence of the i.i.d. and temporally correlated CRBs when SNR is high enough.

n, while $\mathrm{SNR} = 0$ dB is constant. Signal temporal correlation is simulated via filtering an i.i.d. random sequence with an FIR filter with relative tap weights

$$f(z) = 1 + 0.5z^{-1} + 0.3z^{-2} + 0.2z^{-3} + 0.1z^{-4} \qquad (73)$$

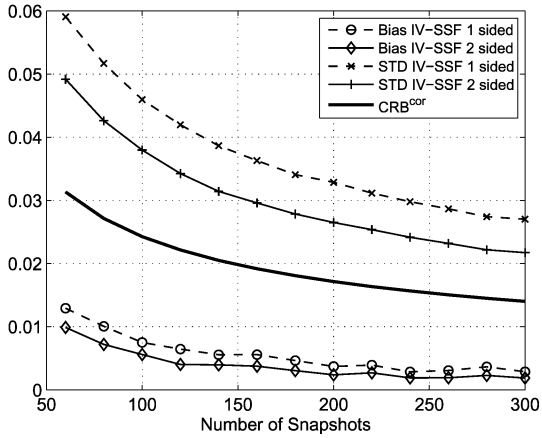

Fig. 4. Bias and standard deviation of one- and two-sided IV-SSF methods versus n. It is clear that the two-sided IV-SSF outperforms the one-sided one in both lower bias and lower standard deviation. The improvement is rather constant with n. Asymptotically in large n, bias is negligible in the overall mean-square error.

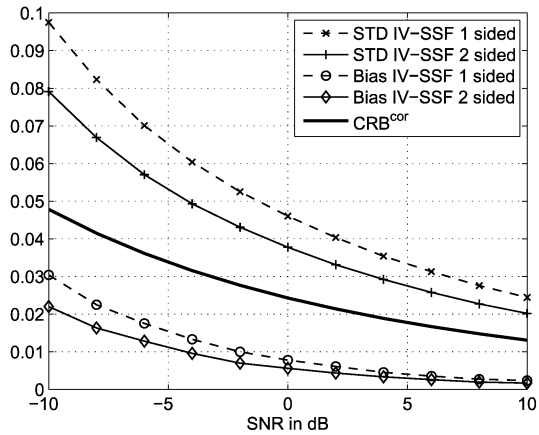

Fig. 5. Bias and standard deviation of one- and two-sided IV-SSF methods versus the SNR. The improvement made by two-sided IV-SSF is larger in lower SNRs. In high SNRs, bias is negligible in the overall error.

which is then normalized to give a unit-energy filter. The estimates are calculated using a two step grid search, first a coarse search with grid size 0.01 and then a finer one with grid size 0.001. Fig. 4 shows the bias and standard deviation of both methods versus n. Although bias is relatively small, it cannot be neglected in small numbers of snapshots. Although the $\mathrm{CRB}^{\mathrm{cor}}$ in (22) has been calculated assuming zero bias, it is still rewarding to compare the performance of the methods to the square root of the $\mathrm{CRB}^{\mathrm{cor}}$, which is depicted in Fig. 4. Here, we used $f(z)$ in (73) to compute the signal space-time covariance matrix $\boldsymbol{\mathcal{P}}$. The improvement made by the two-sided IV-SSF is rather constant with n specially in the standard deviation. It is clear that the two-sided IV-SSF had made roughly 20% improvement in the standard deviation and bias without increasing the computational load of the algorithm.

In Fig. 5, the same comparison is made with constant $n = 100$ and varying SNR. Obviously, the improvement made by two-sided IV-SSF is greater in lower SNR. In high SNR conditions, bias is rather negligible in both methods, while the improvement percentage is nearly constant with SNR. Finally in

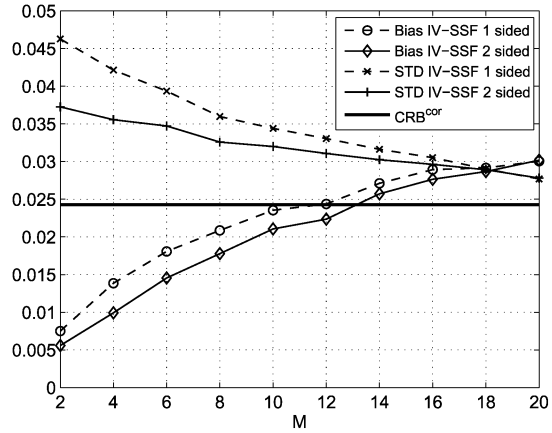

Fig. 6. Bias and standard deviation of one- and two-sided IV-SSF methods versus the number of instrumental variables M. The improvement made by two sided IV-SSF in standard deviation is larger in low M. Also, bias grows with M and finally gets larger than standard deviation for large M.

Fig. 6, the comparison is made in constant $n = 100$ and $\mathrm{SNR} = 0$ dB, while M is changed to see the effect of the method complexity indicator or the number of instrumental variables on the bias and standard deviation of both methods. The interesting point is that in low M, there is considerable improvement from one-sided to two-sided IV-SSF. But this improvement decreases as M increases since for large M, there is enough previous samples to estimate the present signal of interest with sufficient accuracy comparable to the accuracy that can be achieved using the two-sided estimation. Another important point of the Fig. 6 is the increasing relative part of the bias in the overall mean square error of the estimation with increasing M. This phenomena can lead us to limit the number of instrumental variables so as to avoid large unpredictable bias. This selection can be more rational when we consider the increased cost of implementation when M gets larger. It is also evident in Fig. 6 that the $\mathrm{CRB}^{\mathrm{cor}}$ is independent of M and gradually, with increasing M, the portion of error caused by standard deviation decreases while the error caused by bias increases.

VI. Conclusion

In this paper, we presented new theoretical results on the performance of the DOA estimation when signals of interest are possibly temporally correlated. In particular, it was shown that the Cramér–Rao bound in the temporally correlated signal model is upper bounded by the same bound under the i.i.d. signal model. This result implies that temporal correlation of the signals is an additional relevant information for DOA estimation. We showed that the improvement caused by signal temporal correlation is large when SNR is low and little when SNR is high. This is a good news since the high SNR condition is not critical in our systems where we can use many suboptimal methods; while the low SNR performance improvement is valuable to the system. The second part of the paper was devoted to a practical method of DOA estimation in the new signal model. The IV-SSF method was analyzed and compared with the $\mathrm{CRB}^{\mathrm{cor}}$ to show that it is not an efficient method. Then, using a special form of the $\mathrm{CRB}^{\mathrm{cor}}$, we proposed a version of the IV-SSF method (two-sided IV-SSF), that outperform the former method in both lower bias and lower error variance.

References

[1] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Process. Mag.*, vol. 13, pp. 67–94, Jul. 1996.

[2] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Mag.*, vol. 5, no. 2, pp. 4–24, Apr. 1988.

[3] R. O. Schmidt, "Multiple emmiter location and signal parameter estimation," *IEEE Trans. Antennas. Propag.*, vol. AP-34, pp. 276–280, Mar. 1986.

[4] M. Viberg and B. Ottersten, "Sensor array processing based on subspace fitting," *IEEE Trans. Signal Process.*, vol. 39, no. 5, pp. 1110–1121, May 1991.

[5] Y. Bresler and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Trans. Acoust. Speech, Signal Process.*, vol. ASSP-34, pp. 1081–1089, Oct. 1986.

[6] P. Stoica and A. Nehorai, "Music, maximum likelihood, and Cramer–Rao bound," *IEEE Trans. Acoust. Speech, Signal Process.*, vol. 37, pp. 720–741, May 1989.

[7] X. L. Xu and K. M. Buckley, "Bias analysis of the music location estimator," *IEEE Trans. Signal Process.*, vol. 40, no. 10, pp. 2559–2569, Oct. 1992.

[8] P. Tichavsky, "High-SNR asymptotics for signal subspace methods in sinusoidal frequency estimation," *IEEE Trans. Signal Process.*, vol. 41, no. 7, pp. 2448–2460, Jul. 1993.

[9] H. Abeida and J. P. Delmas, "Gaussian Cramer–Rao bound for direction estimation of noncircular signals in unknown noise fields," *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4610–4618, Dec. 2005.

[10] A. G. Jaffer, "Maximum likelihood direction finding of stochastic sources: A separable solution," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing*, 1988, pp. 2893–2896.

[11] J. P. Delmas and Y. Meurisse, "Asymptotic performance analysis of DOA finding algorithms with temporally correlated narrowband signals," *IEEE Trans. Signal Process.*, vol. 48, no. 9, pp. 2669–2674, Sep. 2000.

[12] M. Viberg, P. Stoica, and B. Ottersten, "Array processing in correlated noise fields based on instrumental variables and subspace fitting," *IEEE Trans. Signal Process.*, vol. 43, no. 5, pp. 1187–1199, May 1995.

[13] P. Stoica, M. Viberg, and B. Ottersten, "Instrumental variable approach to array processing in spatially correlated noise fields," *IEEE Trans. Signal Process.*, vol. 42, no. 1, pp. 121–133, Jan. 1994.

[14] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 10, pp. 1783–1795, Oct. 1990.

[15] P. Stoica, M. Viberg, M. Wong, and Q. Wu, "Optimal IV-SSF approach to array signal processing in colored noise fields," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, 1995, pp. 2088–2091.

[16] P. Stoica, E. G. Larsson, and A. B. Gershman, "The stochastic CRB for array processing: A textbook derivation," *IEEE Signal Process. Lett.*, vol. 8, pp. 148–150, May 2001.

[17] G. Bienvenu and L. Kopp, "Optimality of high resolution array processing using the eigensystem approach," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-31, pp. 1235–1248, Oct. 1983.

[18] A. S. Householder, *The Theory of Matrices in Numerical Analysis*. New York: Dover, 1975.

Farzan Haddadi was born in the United States in 1979. He received the B.Sc. and M.Sc. degrees in communication systems from Sharif University of Technology, Tehran, Iran, in 2001 and 2003, respectively, where he is currently working toward the Ph.D. degree.

His main research interests are array signal processing, statistical signal processing, and detection theory.

Mohammad M. Nayebi (S'89–M'94–SM'05) was born in Iran in 1967. He received the B.S.E.E. and M.S.E.E. degrees, both with the honor of first rank, from Sharif University of Technology, Tehran, Iran, in 1988 and 1990, respectively, and the Ph.D. degree in electrical engineering, with the same honor, from Tarbiat Modarres University, Tehran, Iran, in 1994.

He joined the Sharif University of Technology faculty in 1994, and became an Associate Professor in 1998, and Professor in 2003. His main research interests are radar signal processing and detection theory. Also, he has published more than 70 technical papers in international and national journals and conferences.

Mohammad R. Aref was born in Iran in 1951. He received the B.Sc. degree from Tehran University, Tehran, Iran, in 1974 and the M.Sc. and Ph.D. degrees from Stanford University, Stanford, CA, in 1975 and 1979, respectively, all in electrical engineering.

From 1994 to 1997, he was the Chancellor of Tehran University. He is now a Professor of Electrical Engineering with Sharif University of Technology, Tehran, Iran, and has published numerous technical papers on communication theory in international journals and conference proceedings.