OPTIMAL IV-SSF APPROACH TO ARRAY SIGNAL PROCESSING IN COLORED NOISE FIELDS

P. STOICA, M. VIBERG

Department of Applied Electronics
Chalmers University of Technology
S-412 96 Gothenburg, Sweden
Email: viberg@ae.chalmers.se

M. WONG AND Q. WU

Department of ECE
McMaster University
Hamilton, Ontario L8S 4K1, Canada
Email: wong@ece.eng.mcmaster.ca

ABSTRACT

The main goal of this paper is to describe and analyse, in a unifying manner, the *spatial* and *temporal* IV-SSF approaches recently proposed for array signal processing in colored noise fields. (The acronym IV-SSF stands for "Instrumental Variable – Signal Subspace Fitting"). We derive a general, optimally-weighted, IV-SSF direction estimator and show that this estimator encompasses the UNCLE estimator of Wong and Wu, which is a spatial IV-SSF method; and the temporal IV-SSF estimator of Viberg, Stoica and Ottersten. The latter two estimators have seemingly different forms, so their asymptotic equivalence shown in this paper comes as a surprising unifying result.

1. INTRODUCTION

In recent years, a number of methods for direction finding in unknown colored noise-fields have been proposed, e.g. [1]. Herein, the combined Instrumental Variable - Signal Subspace Fitting (IV-SSF) approach is considered. The IV-SSF technique has a number of appealing advantages over other methods, including [2, 3]: 1) The IV-SSF can handle noises with arbitrary spatial covariances, under minor restrictions on the signals or the array. In addition, the IV-SSF do not estimate a noise model – which makes them statistically more robust and computationally simpler than the approaches relying on noise modelling. 2) The IV-SSF approach is applicable to both non-coherent and coherent signal scenarios. 3) The spatial IV-SSF can make use of the information contained in the output of a completely uncalibrated subarray, under certain weak conditions, which other methods cannot.

Depending on the type of "instrumental variables" used, two classes of IV-SSF methods have been proposed:

a) *Spatial IV-SSF*, for which the instrumental variables are derived from the output of a (possibly uncalibrated) subarray whose noise is spatially uncorrelated with the noise in the main calibrated subarray under consideration [2].

b) *Temporal IV-SSF*, which obtain instrumental variables from the delayed versions of the array output, under the assumption that the temporal-correlation length of the noise field is shorter than that of the signals [3].

The previous literature on IV-SSF has treated and analysed the above two classes of spatial and temporal methods separately, ignoring their common basis. In this paper we reveal the common roots of these two classes of DOA estimation methods and study them under the same umbrella. Additionally, we establish the statistical properties of a general weighted IV-SSF method and derive the optimal weights that minimize the DOA estimation errors, In particular, we prove that the optimal four-weight spatial IV-SSF of [2] and the optimal three-weight temporal IV-SSF of [3] are asymptotically equivalent when used under the same conditions.

2. PROBLEM FORMULATION

Consider a scenario in which n narrowband plane waves, generated by point sources, impinge on an array comprising m calibrated sensors. Assume, for simplicity, that the n sources and the array are situated in the same plane. Let $\boldsymbol{a}(\theta)$ denote the array response to a unit-amplitude signal with a DOA parameter equal to θ. Under these assumptions, the output of the array, $\boldsymbol{y}(t) \in \mathcal{C}^{m\times 1}$, can be described by the following well-known equation [4]

$$\boldsymbol{y}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{e}(t) \tag{1}$$

where $\boldsymbol{x}(t) \in \mathcal{C}^{n\times 1}$ denotes the signal vector, $\boldsymbol{e}(t) \in \mathcal{C}^{m\times 1}$ is a noise term, and

$$\boldsymbol{A} = [\boldsymbol{a}(\theta_1)\cdots \boldsymbol{a}(\theta_n)] \tag{2}$$

Hereafter, θ_k denotes the kth DOA parameter.

The following assumptions on the quantities in the array equation, (1), are considered to hold throughout the paper.

A1. The signal vector $\boldsymbol{x}(t)$ is a normally distributed random variable with zero mean and a possibly singular covariance. The signals may be temporally correlated; In fact the temporal IV-SSF approach relies on the assumption that the signals exhibit some form of temporal correlation (see below for details).

A2. The noise $\boldsymbol{e}(t)$ is a random vector that is temporally white, uncorrelated with the signals and circularly symmetric normally distributed with zero mean and unknown covariance matrix[1] $\boldsymbol{Q} > \boldsymbol{O}$,

$$\mathrm{E}\left[\boldsymbol{e}(t)\boldsymbol{e}^*(s)\right] = \boldsymbol{Q}\ \delta_{t,s}\ ; \quad \mathrm{E}\left[\boldsymbol{e}(t)\boldsymbol{e}^T(s)\right] = \boldsymbol{O} \tag{3}$$

[1]Henceforth, the superscript "$*$" denotes the conjugate transpose; whereas the transpose is designated by a superscript "T". The notation $\boldsymbol{A} \geq \boldsymbol{B}$, for two Hermitian matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, is used to mean that $(\boldsymbol{A} - \boldsymbol{B})$ is a nonnegative definite matrix.

A3. The manifold vectors $\{a(\theta)\}$, corresponding to any set of m different values of θ, are linearly independent.

Note that assumption *A1* above allows for coherent signals, and that in *A2* the noise field is allowed to be arbitrarily spatially correlated with an unknown covariance matrix. Assumption *A3* is a well-known condition that, under a weak restriction on m, guarantees DOA parameter identifiability in the case Q is known (to within a multiplicative constant). When Q is completely unknown, DOA identifiability can only be achieved if further assumptions are made on the scenario under consideration. The following assumption is typical of the IV-SSF approach.

A4. There exists a known vector $z(t) \in \mathcal{C}^{\bar{m}\times 1}$, which is normally distributed and satisfies

$$\mathrm{E}\left[z(t)e^*(s)\right] = 0 \quad \text{for } t \leq s \qquad (4)$$

$$\mathrm{E}\left[z(t)e^T(s)\right] = 0 \quad \text{for all } t, s \qquad (5)$$

Furthermore, denote

$$\Gamma = \mathrm{E}\left[z(t)x^*(t)\right] \qquad (\bar{m} \times n) \qquad (6)$$

$$\bar{n} = \mathrm{rank}\,(\Gamma) \leq \bar{m}\,. \qquad (7)$$

It is assumed that no row of Γ is identically zero and that the inequality

$$\bar{n} > 2n - m \qquad (8)$$

holds. (note that a rank-one Γ matrix can satisfy the condition (8) if m is large enough, and hence the condition in question is rather weak). Owing to its (partial) uncorrelatedness with $\{e(t)\}$, the vector $\{z(t)\}$ can be used to eliminate the noise from the array output equation (1), and for this reason $\{z(t)\}$ is called an Instrumental Variable (IV) vector. Below, we briefly describe two possible ways to derive an IV vector from the available data measured with an array of sensors.

Example 2.1 *Spatial IV*
Assume that the n signals, which impinge on the main (sub)array under consideration, are also received by another (sub)array that is sufficiently distanced from the main one so that the noise vectors in the two subarrays are spatially uncorrelated with one another. Then $z(t)$ can be made from the outputs of the sensors in the second subarray (note that those sensors need not be calibrated) [2, 5]. □

Example 2.2 *Temporal IV*
When a second subarray, as described above, is not available but the signals are temporally correlated, one can obtain an IV vector by delaying the output vector: $z(t) = [y^T(t-1)\ y^T(t-2)\ \cdots\ ]^T$. Clearly, such a vector $z(t)$ satisfies (4)–(5) and it also satisfies (8) under weak conditions on the signal temporal correlation. This construction of an IV vector can be readily extended to cases where $e(t)$ is temporally correlated, provided that the signal temporal correlation length is longer than that corresponding to the noise [3]. □

The problem considered in the following sections concerns the estimation of the DOA vector

$$\theta = [\theta_1 \cdots \theta_n]^T \qquad (9)$$

given N snapshots of the array output and of the IV vector, $\{y(t), z(t)\}_{t=1}^N$. The number of signals, n, and the rank of the covariance matrix Γ, $\bar{n}$, are assumed to be given (for the estimation of these integer-valued parameters by means of IV/SSF-based methods. we refer to [6]).

3. THE IV-SSF APPROACH

Define the "weighted IV sample cross-covariance matrix" as

$$\hat{R} = \hat{W}_L \left[\frac{1}{N}\sum_{t=1}^{N} z(t)y^*(t)\right] \hat{W}_R \qquad (\bar{m} \times m) \qquad (10)$$

where $\hat{W}_L$ and $\hat{W}_R$ are two nonsingular Hermitian weighting matrices which are possibly data-dependent (as indicated by the fact that they are roofed). Under the assumptions made, as $N \to \infty$, $\hat{R}$ converges to the matrix:

$$R = W_L \mathrm{E}[z(t)y^*(t)] W_R = W_L \Gamma A^* W_R \qquad (11)$$

where W_L and W_R are the limiting weighting matrices (assumed to be bounded and nonsingular). Owing to assumptions *A2* and *A3*,

$$\mathrm{rank}\,(R) = \bar{n} \qquad (12)$$

Hence, the Singular Value Decomposition (SVD) of R can be written as

$$R = [U\ \ ?] \begin{bmatrix} \Lambda & O \\ O & O \end{bmatrix} \begin{bmatrix} S^* \\ ? \end{bmatrix} = U\Lambda S^* \qquad (13)$$

where $U^*U = S^*S = I$, $\Lambda \in \mathcal{R}^{\bar{n}\times\bar{n}}$ is diagonal and nonsingular, and where the question marks stand for blocks that are of no importance for the present discussion.

The following key equality is obtained by comparing the two expressions for R in (11) and (13) above:

$$S = W_R A C \qquad (14)$$

where $C \triangleq \Gamma^* W_L U \Lambda^{-1} \in \mathcal{C}^{n\times\bar{n}}$ has full column rank. For a given S, the true DOA vector can be obtained as the unique solution to (14) under the parameter identifiability condition (8) (see, e.g., [7]). In the more realistic case when S is unknown, one can make use of (14) to estimate the DOA vector in the following steps.

The IV step: Compute the weighted IV sample cross-covariance matrix $\hat{R}$ in (10), along with its SVD:

$$\hat{R} = \left[\ \hat{U}\ \ ?\ \right] \begin{bmatrix} \hat{\Lambda} & O \\ O & ? \end{bmatrix} \begin{bmatrix} \hat{S}^* \\ ? \end{bmatrix} \qquad (15)$$

Note that $\hat{U}, \hat{\Lambda}$ and $\hat{S}$ are consistent estimates of U, Λ and S in the SVD of R.

The SSF step: Estimate the DOAs by minimizing the following signal subspace fitting criterion:

$$\min_{\theta}\{\min_{C}[\mathrm{vec}\,(\hat{S} - \hat{W}_R A C)]^* \hat{V}[\mathrm{vec}\,(\hat{S} - \hat{W}_R A C)]\} \qquad (16)$$

where $\hat{V}$ is a positive definite weighting matrix, and "vec" is the vectorization operator[2]. Alternatively, one can estimate

[2] If x_k is the kth column of a matrix X, then $\mathrm{vec}\,(X) = [x_1^T\ x_2^T\ \cdots\ ]^T$.

the DOAs as the minimizing arguments of the following criterion:

$$\min_{\boldsymbol{\theta}}\{[\mathrm{vec}\,(\boldsymbol{B}^*\hat{\boldsymbol{W}}_R^{-1}\hat{\boldsymbol{S}})]^*\hat{\boldsymbol{W}}[\mathrm{vec}\,(\boldsymbol{B}^*\hat{\boldsymbol{W}}_R^{-1}\hat{\boldsymbol{S}})]\} \tag{17}$$

where $\hat{\boldsymbol{W}}$ is a positive definite weight, and $\boldsymbol{B} \in \mathcal{C}^{m\times(m-n)}$ is a matrix whose columns form a basis of the null-space of $\boldsymbol{A}^*$ (hence, $\boldsymbol{B}^*\boldsymbol{A} = 0$ and $\mathrm{rank}\,(\boldsymbol{B}) = m - n$). The alternative fitting criterion above is obtained from the simple observation that (14) along with the definition of $\boldsymbol{B}$ imply that

$$\boldsymbol{B}^*\boldsymbol{W}_R^{-1}\boldsymbol{S} = 0 \tag{18}$$

In [8], it is shown that *the classes of DOA estimates derived from (16) and (17), respectively, are asymptotically equivalent.* More exactly, for any $\hat{\boldsymbol{V}}$ in (16) one can choose $\hat{\boldsymbol{W}}$ in (17) so that the DOA estimates obtained by minimizing (16) and, respectively, (17) have the same asymptotic distribution; and vice-versa.

In view of the previous result, in an asymptotical analysis it suffices to consider only one of the two criteria above. In the sections to follow we focus on (17). Compared with (16), the criterion (17) has the advantage that it depends on the DOA only. On the other hand, for a general array there is no known closed-form parameterization of $\boldsymbol{B}$ in terms of $\boldsymbol{\theta}$. However, it turns out that this is no drawback since the optimally weighted criterion (which is the one to be used in applications) is an explicit function of $\boldsymbol{\theta}$ (see the next section).

4. THE OPTIMAL IV-SSF METHOD

In what follows, we deal with the essential problem of choosing the weights $\hat{\boldsymbol{W}}, \hat{\boldsymbol{W}}_R$ and $\hat{\boldsymbol{W}}_L$ in the IV-SSF criterion (17) so as to maximize the DOA estimation accuracy. First, we optimize the accuracy with respect to $\hat{\boldsymbol{W}}$, and then with respect to $\hat{\boldsymbol{W}}_R$ and $\hat{\boldsymbol{W}}_L$.

Define

$$g(\boldsymbol{\theta}) = \mathrm{vec}\,(\boldsymbol{B}^*\hat{\boldsymbol{W}}_R^{-1}\hat{\boldsymbol{S}}) \tag{19}$$

and observe that the criterion function in (17) can be written as,

$$g^*(\boldsymbol{\theta})\hat{\boldsymbol{W}}g(\boldsymbol{\theta}) \tag{20}$$

In [8] we show that $g(\boldsymbol{\theta})$ (evaluated at the true DOA vector) has, asymptotically in N, a circularly symmetric normal distribution with zero mean and the following covariance:

$$G(\boldsymbol{\theta}) = \frac{1}{N}[(\boldsymbol{W}_L\boldsymbol{U}\Lambda^{-1})^*\boldsymbol{R}_z(\boldsymbol{W}_L\boldsymbol{U}\Lambda^{-1})]^T \otimes [\boldsymbol{B}^*\boldsymbol{R}_y\boldsymbol{B}]$$

where $\otimes$ denotes the Kronecker matrix product; and where, for a stationary signal $\boldsymbol{s}(t)$,

$$\boldsymbol{R}_s = \mathrm{E}\,[\boldsymbol{s}(t)\boldsymbol{s}^*(t)]$$

Then, it follows from the ABC (Asymptotically Best Consistent) theory of parameter estimation [9] that the minimum variance estimate, in the class of estimates under discussion, is given by the minimizing argument of the criterion in (20) with $\hat{\boldsymbol{W}} = \hat{\boldsymbol{G}}^{-1}(\boldsymbol{\theta})$, that is

$$f(\boldsymbol{\theta}) = g^*(\boldsymbol{\theta})\hat{\boldsymbol{G}}^{-1}(\boldsymbol{\theta})g(\boldsymbol{\theta}) \tag{21}$$

where

$$\hat{\boldsymbol{G}}(\boldsymbol{\theta}) = \frac{1}{N}[(\hat{\boldsymbol{W}}_L\hat{\boldsymbol{U}}\hat{\Lambda}^{-1})^*\hat{\boldsymbol{R}}_z(\hat{\boldsymbol{W}}_L\hat{\boldsymbol{U}}\hat{\Lambda}^{-1})]^T \otimes [\boldsymbol{B}^*\hat{\boldsymbol{R}}_y\boldsymbol{B}]$$

and where $\hat{\boldsymbol{R}}_z$ and $\hat{\boldsymbol{R}}_y$ are the usual sample estimates of $\boldsymbol{R}_z$ and $\boldsymbol{R}_y$. Furthermore, it is shown in [8] that the minimum variance estimate, obtained by minimizing (21), is asymptotically normally distributed with mean equal to the true parameter vector and the following covariance matrix:

$$\boldsymbol{H} = \frac{1}{2N}\left(\mathrm{Re}\left\{\boldsymbol{\Psi}\odot\boldsymbol{\Omega}^T\right\}\right)^{-1} \tag{22}$$

Here, $\odot$ denotes the Hadamard-Schur matrix product and the matrices $\boldsymbol{\Psi}$ and $\boldsymbol{\Omega}$ are defined as

$$\boldsymbol{\Psi} = \boldsymbol{D}^*\boldsymbol{R}_y^{-1/2}\boldsymbol{\Pi}^{\perp}_{\boldsymbol{R}_y^{-1/2}\boldsymbol{A}}\boldsymbol{R}_y^{-1/2}\boldsymbol{D} \tag{23}$$

$$\boldsymbol{\Omega} = \boldsymbol{\Gamma}^*\boldsymbol{W}_L\boldsymbol{U}(\boldsymbol{U}^*\boldsymbol{W}_L\boldsymbol{R}_z\boldsymbol{W}_L\boldsymbol{U})^{-1}\boldsymbol{U}^*\boldsymbol{W}_L\boldsymbol{\Gamma} \tag{24}$$

where $\boldsymbol{Y}^{-1/2}$ denotes a Hermitian (for notational convenience) square root of the inverse of a positive definite matrix $\boldsymbol{Y}$ and the matrix $\boldsymbol{D}$ is made from the direction vector derivatives,

$$\boldsymbol{D} = [\boldsymbol{d}_1 \ \cdots \ \boldsymbol{d}_n]; \quad \boldsymbol{d}_k = \frac{\partial \boldsymbol{a}(\theta_k)}{\partial \theta_k}$$

Finally, for a full column-rank matrix $\boldsymbol{X}$, the projection operator $\boldsymbol{\Pi}_X^{\perp}$ is defined as

$$\boldsymbol{\Pi}_X^{\perp} = \boldsymbol{I} - \boldsymbol{\Pi}_X; \quad \boldsymbol{\Pi}_X = \boldsymbol{X}(\boldsymbol{X}^*\boldsymbol{X})^{-1}\boldsymbol{X}^* \tag{25}$$

The optimal weights $\hat{\boldsymbol{W}}_R$ and $\hat{\boldsymbol{W}}_L$ are, by definition, those which minimize the limiting covariance matrix $\boldsymbol{H}$ of the DOA estimation errors. It follows from (22)-(24) and the properties of the Hadamard-Schur product, that the optimal weighting matrices are found by maximizing $\boldsymbol{\Omega}$. Since the matrix $\boldsymbol{\Gamma}$ has rank $\bar{n}$, it can be factorized as follows:

$$\boldsymbol{\Gamma} = \boldsymbol{\Gamma}_1\boldsymbol{\Gamma}_2^* \tag{26}$$

where both $\boldsymbol{\Gamma}_1 \in \mathcal{C}^{\bar{m}\times\bar{n}}$ and $\boldsymbol{\Gamma}_2 \in \mathcal{C}^{n\times\bar{n}}$ have full column rank. Insertion of (26) into the equality $\boldsymbol{W}_L\boldsymbol{\Gamma}\boldsymbol{A}^*\boldsymbol{W}_R = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{S}^*$ yields

$$\boldsymbol{W}_L\boldsymbol{\Gamma}_1\boldsymbol{T} = \boldsymbol{U} \tag{27}$$

where $\boldsymbol{T} = \boldsymbol{\Gamma}_2^*\boldsymbol{A}^*\boldsymbol{W}_R\boldsymbol{S}\Lambda^{-1} \in \mathcal{C}^{\bar{n}\times\bar{n}}$ is a nonsingular transformation matrix. By using (27) in (24), we obtain:

$$\boldsymbol{\Omega} = \boldsymbol{\Gamma}_2(\boldsymbol{\Gamma}_1^*\boldsymbol{W}_L^2\boldsymbol{\Gamma}_1)(\boldsymbol{\Gamma}_1^*\boldsymbol{W}_L^2\boldsymbol{R}_z\boldsymbol{W}_L^2\boldsymbol{\Gamma}_1)^{-1}(\boldsymbol{\Gamma}_1^*\boldsymbol{W}_L^2\boldsymbol{\Gamma}_1)\boldsymbol{\Gamma}_2^*$$

Observe that $\boldsymbol{\Omega}$ does not actually depend on $\boldsymbol{W}_R$. Hence, *$\hat{\boldsymbol{W}}_R$ can be arbitrarily selected, as any nonsingular Hermitian matrix, without affecting the asymptotics of the DOA parameter estimates.*

Concerning the choice of $\hat{\boldsymbol{W}}_L$, it is easily verified [8] that

$$\boldsymbol{\Omega} \le \boldsymbol{\Omega}\,|_{\boldsymbol{W}_L=\boldsymbol{R}_z^{-1/2}} = \boldsymbol{\Gamma}_2(\boldsymbol{\Gamma}_1^*\boldsymbol{R}_z^{-1}\boldsymbol{\Gamma}_1)\boldsymbol{\Gamma}_2^* = \boldsymbol{\Gamma}^*\boldsymbol{R}_z^{-1}\boldsymbol{\Gamma} \tag{28}$$

Hence, $\boldsymbol{W}_L = \boldsymbol{R}_z^{-1/2}$ maximizes $\boldsymbol{\Omega}$. We conclude that *the optimal weight $\hat{\boldsymbol{W}}_L$, which yields the best limiting accuracy, is*

$$\hat{\boldsymbol{W}}_L = \hat{\boldsymbol{R}}_z^{-1/2} \tag{29}$$

The (minimum) covariance matrix $\boldsymbol{H}$, corresponding to the above choice, is found to be

$$\boldsymbol{H}_o = \frac{1}{2N}\{\mathrm{Re}[\boldsymbol{\Psi} \odot (\boldsymbol{\Gamma}^* \boldsymbol{R}_z^{-1} \boldsymbol{\Gamma})^T]\}^{-1} \quad (30)$$

Optimal IV-SSF Criteria: The criterion (21) can be expressed in a functional form that depends on the indeterminate $\boldsymbol{\theta}$ in an explicit way (recall that, for most cases, the dependence of $\boldsymbol{B}$ on $\boldsymbol{\theta}$ is not available in explicit form). Because of the arbitrariness in the choice of $\hat{\boldsymbol{W}}_R$, there exists an infinite class of optimal IV-SSF criteria. In what follows, we consider two members of this class.

First, let

$$\hat{\boldsymbol{W}}_R = \hat{\boldsymbol{R}}_y^{-1/2} \quad (31)$$

Insertion of (31), along with (29), into (21) yields after some manipulations the following criterion function

$$f_{WW}(\boldsymbol{\theta}) = \mathrm{tr}\left(\boldsymbol{\Pi}^{\perp}_{\hat{\boldsymbol{R}}_y^{-1/2}\boldsymbol{A}} \tilde{\boldsymbol{S}} \tilde{\boldsymbol{\Lambda}}^2 \tilde{\boldsymbol{S}}^*\right) \quad (32)$$

where $\tilde{\boldsymbol{S}}$ and $\tilde{\boldsymbol{\Lambda}}$ are made from the principal singular right vectors and singular values of the matrix

$$\tilde{\boldsymbol{R}} = \hat{\boldsymbol{R}}_z^{-1/2} \hat{\boldsymbol{R}}_{zy} \hat{\boldsymbol{R}}_y^{-1/2} \quad (33)$$

(with $\hat{\boldsymbol{R}}_{zy}$ defined in an obvious way). The function (36) is the UNCLE (spatial IV-SSF) criterion of Wong and Wu [2].

Next, choose $\hat{\boldsymbol{W}}_R$ as

$$\hat{\boldsymbol{W}}_R = \boldsymbol{I} \quad (34)$$

The corresponding criterion function is

$$f_{VSO}(\boldsymbol{\theta}) = \mathrm{tr}\left(\boldsymbol{\Pi}^{\perp}_{\hat{\boldsymbol{R}}_y^{-1/2}\boldsymbol{A}} \hat{\boldsymbol{R}}_y^{-1/2} \bar{\boldsymbol{S}} \bar{\boldsymbol{\Lambda}}^2 \bar{\boldsymbol{S}}^* \hat{\boldsymbol{R}}_y^{-1/2}\right) \quad (35)$$

where $\bar{\boldsymbol{S}}$ and $\bar{\boldsymbol{\Lambda}}$ are made from the principal singular pairs of

$$\bar{\boldsymbol{R}} = \hat{\boldsymbol{R}}_z^{-1/2} \hat{\boldsymbol{R}}_{zy} \quad (36)$$

The function (35) above is recognized as the optimal (temporal) IV-SSF criterion of Viberg, Stoica and Ottersten [3].

An important consequence of the previous discussion is that the DOA estimation methods of [2] and [3], respectively, which were derived in seemingly unrelated contexts and by means of somewhat different approaches, are in fact asymptotically equivalent when used under the same conditions. These two methods also have very similar computational burdens as is readily seen.

5. CONCLUDING REMARKS

The main points made by the present paper can be summarized as follows.

i) The spatial and temporal IV-SSF approaches can be treated in a unified manner under general conditions.

ii) The optimization of the DOA parameter estimation accuracy, for fixed weights $\hat{\boldsymbol{W}}_L$ and $\hat{\boldsymbol{W}}_R$, can be most conveniently carried out using the ABC theory. (The resulting derivations are more concise than those based on other analysis techniques).

iii) The right-hand (or post-)weight $\hat{\boldsymbol{W}}_R$ has no effect on the asymptotics.

iv) An important corollary of the result in (iii) above is that the optimal IV-SSF methods of [2] and, respectively, [3] are asymptotically equivalent when used on the same data.

Let us reiterate the facts that the IV-SSF approaches can deal with coherent signals, handle noise fields with general (unknown) spatial correlations, and, in their spatial versions, can make use of outputs from completely uncalibrated sensors. They are also comparatively simple from a computational standpoint, since no noise modelling is required. Additionally, the optimal IV-SSF methods provide highly accurate DOA estimates. More exactly, in spatial IV scenarios these DOA estimation methods can be shown to be asymptotically statistically efficient under weak conditions [5]. In temporal scenarios, they are no longer exactly statistically efficient but their accuracy is still quite close to the best possible one [3]. All these features and properties should make the optimal IV-SSF approach appealing for practical array signal processing problems.

6. REFERENCES

[1] V. Nagesha and S. Kay, "Maximum Likelihood Estimation for Array Processing in Colored Noise", In *Proc. ICASSP 93*, volume 4, pages 240–243, Minneapolis, MN, 1993.

[2] Q. Wu and K.M. Wong, "UN-MUSIC and UN-CLE: An Application of Generalized Canonical Correlation Analysis to the Estimation of the Directions of Arrival", *IEEE Trans. SP*, **42**, Sept. 1994.

[3] M. Viberg, P. Stoica, and B. Ottersten, "Array Processing in Correlated Noise Fields Based on Instrumental Variables and Subspace Fitting", *IEEE Trans. SP*, May 1995.

[4] R.O. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation", *IEEE Trans. AP*, **34**:276–280, Mar. 1986.

[5] P. Stoica, M. Viberg, M. Wong, and Q. Wu, "Optimal Direction Finding with Partly Calibrated Arrays in Spatially Correlated Noise Fields", In *Proc. 28th Asilomar Conf. Sig., Syst., Comput.*, Pacific Grove, CA, Oct. 1994.

[6] Q.T. Zhang and K.M. Wong, "Information Theoretic Criteria for the Determination of the Number of Signals in Spatially Correlated Noise", *IEEE Trans. SP*, **SP-41**(4):1652–1663, April 1993.

[7] M. Wax and I. Ziskind, "On Unique Localization of Multiple Sources by Passive Sensor Arrays", *IEEE Trans. on ASSP*, **ASSP-37**(7):996–1000, July 1989.

[8] P. Stoica, M. Viberg, M. Wong, and Q. Wu, "A Unified Instrumental Variable Approach to Direction Finding in Unknown Noise Fields", In V. Madisetti and D.B. Williams, editors, *Digital Signal Processing Handbook*. CRC Press, 1996, To Appear (technical report available from authors).

[9] T. Söderström and P. Stoica, *System Identification*, Prentice-Hall, London, U.K., 1989.