ANALYSIS OF FORWARD-ONLY AND FORWARD-BACKWARD SAMPLE COVARIANCES

Magnus Jansson

Dept. of Signals, Sensors, & Systems
Royal Institute of Technology (KTH)
SWEDEN

Petre Stoica

Systems & Control Group
Uppsala University
SWEDEN

ABSTRACT

In some applications the covariance matrix of the observations is not only symmetric with respect to its main diagonal but also with respect to the anti-diagonal. The standard forward-only sample covariance estimate does not impose this extra symmetry. In such cases one often uses the so-called forward-backward sample covariance estimate. In this paper, a direct comparative study of the relative accuracy of the two sample estimates is performed. An explicit expression for the difference between the estimation error covariance matrices of the two sample estimates is given. The presented results are also useful in the analysis of estimators based on either of the two sample covariances. As an example, spatial power estimation by means of the Capon method is considered. It is shown that Capon based on the forward-only sample covariance (F-Capon) underestimates the power spectrum, and also that the bias for Capon based on the forward-backward sample covariance is half that of F-Capon.

1. INTRODUCTION

Let $\{\mathbf{x}(t)\}$ be an m-dimensional zero-mean stochastic process with covariance

$$\mathbf{R} = \mathrm{E}\{\mathbf{x}(t)\mathbf{x}^*(t)\},$$

where the superscript $*$ denotes conjugate transpose. A commonly used sample estimate of $\mathbf{R}$ is

$$\hat{\mathbf{R}} = \frac{1}{N}\sum_{t=1}^{N}\mathbf{x}(t)\mathbf{x}^*(t),$$

where N is the number of observed samples. The focus of this paper is on cases where $\mathbf{R}$ is *centro-hermitian* (CH), which means that

$$\mathbf{R} = \mathbf{J}\mathbf{R}^T\mathbf{J}, \tag{1}$$

where

$$\mathbf{J} = \begin{bmatrix} 0 & \dots & 0 & 1 \\ 0 & & 1 & 0 \\ \vdots & \iddots & & \vdots \\ 1 & \dots & 0 & 0 \end{bmatrix}.$$

This work was partly supported by the Swedish Research Council for Engineering Sciences (TFR) and the Senior Individual Grant Program of the Swedish Foundation for Strategic Research. Corresponding author: Magnus Jansson, Fax: +46 8 7907329. E-mail: magnus.jansson@s3.kth.se

Note that $\mathbf{R}^T = \mathbf{R}^c$ since $\mathbf{R}$ is hermitian, where $\mathbf{R}^c$ denotes the complex conjugate of $\mathbf{R}$. When $\mathbf{R}$ satisfies the CH property in (1) a natural sample estimate of $\mathbf{R}$ is

$$\hat{\mathbf{R}}_{\mathrm{FB}} = \frac{1}{2}\left(\hat{\mathbf{R}} + \mathbf{J}\hat{\mathbf{R}}^T\mathbf{J}\right). \tag{2}$$

In this paper we will study the relative accuracy of the two estimates $\hat{\mathbf{R}}$ and $\hat{\mathbf{R}}_{\mathrm{FB}}$ of $\mathbf{R}$. In what follows, we will refer to $\hat{\mathbf{R}}$ and $\hat{\mathbf{R}}_{\mathrm{FB}}$ as the forward-only and the forward-backward sample covariance matrix, respectively.

2. MOTIVATION

The main application of interest here is array signal processing (ASP). In ASP $\mathbf{x}(t)$ is the output of a sensor array mathematically described as

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t).$$

Here, $\mathbf{s}(t)$ contains the signals impinging on the array, $\mathbf{A}(\boldsymbol{\theta})$ is the $m \times n$ array response matrix, and $\mathbf{n}(t)$ is noise. The elements of the n-dimensional vector $\boldsymbol{\theta}$ are the directions to the sources that emitted the signals in $\mathbf{s}(t)$. Typically $\mathbf{s}(t)$ and $\mathbf{n}(t)$ are assumed to be independent zero-mean circular Gaussian processes with second moments given by:

$$\begin{aligned} \mathrm{E}\{\mathbf{s}(t)\mathbf{s}^*(s)\} &= \mathbf{P}\delta_{t,s} \\ \mathrm{E}\{\mathbf{s}(t)\mathbf{s}^T(s)\} &= \mathbf{0} \\ \mathrm{E}\{\mathbf{n}(t)\mathbf{n}^*(s)\} &= \mathbf{\Sigma}\delta_{t,s} \\ \mathrm{E}\{\mathbf{n}(t)\mathbf{n}^T(s)\} &= \mathbf{0}, \end{aligned}$$

where $\delta_{t,s}$ is the Kronecker delta. The above assumptions imply that the covariance of $\mathbf{x}(t)$ is

$$\mathbf{R} = \mathbf{A}\mathbf{P}\mathbf{A}^* + \mathbf{\Sigma}.$$

Here, we have omitted the dependence of $\mathbf{A}$ on $\boldsymbol{\theta}$ for notational simplicity.

In ASP, the sample estimate $\hat{\mathbf{R}}_{\mathrm{FB}}$ in (2) is sometimes also used when $\mathbf{R}$ is not centro-hermitian. We do not consider such cases here, as the comparison of $\hat{\mathbf{R}}$ and $\hat{\mathbf{R}}_{\mathrm{FB}}$ would then make no sense. So in what follows we assume that $\mathbf{R}$ is CH. In ASP this happens if (and practically only if) $\mathbf{P}$ is diagonal, that is the signals are uncorrelated, $\mathbf{\Sigma}$ is CH, and the (linear) array is symmetric (for example, a uniform linear array). Typically, one assumes that the noise is *spatially* white so that $\mathbf{\Sigma} = \sigma^2\mathbf{I}$; however, we do not make this assumption here.

It is a well known fact that $\hat{\mathbf{R}}$ is the unstructured maximum likelihood estimate (MLE) of $\mathbf{R}$. It is also known [6] that $\hat{\mathbf{R}}_{\text{FB}}$ is the *structured* MLE of $\mathbf{R}$ when $\mathbf{R}$ is CH. For completeness, in Section 3 we provide a very simple proof of that result.

By the above fact, $\hat{\mathbf{R}}_{\text{FB}}$ should be more accurate than $\hat{\mathbf{R}}$. However, this is a qualitative statement. A more quantitative study of the accuracies of $\hat{\mathbf{R}}_{\text{FB}}$ and $\hat{\mathbf{R}}$ does not appear to be available. Performing this study is one of our goals here.

Some indirect studies of $\hat{\mathbf{R}}_{\text{FB}}$ and $\hat{\mathbf{R}}$ have appeared in [7] and [4]. More exactly, these authors studied the accuracy of MUSIC, ESPRIT, and other similar direction estimates obtained from $\hat{\mathbf{R}}$ and $\hat{\mathbf{R}}_{\text{FB}}$, respectively. Such studies, however, have a lesser interest nowadays for a reason explained next. Statistically efficient methods are now available for estimating the directions by root-techniques of the same complexity as root-MUSIC, ESPRIT etc. (assuming $\mathbf{\Sigma} = \sigma^2\mathbf{I}$, which is anyway required by the methods of [7] and [4] as well) [2]. These optimal methods yield the same asymptotic accuracy, regardless of whether they are applied to $\hat{\mathbf{R}}_{\text{FB}}$ or $\hat{\mathbf{R}}$. Consequently, for the more recent ASP methods (under the assumptions made here) the difference between the performance of $\hat{\mathbf{R}}_{\text{FB}}$ and $\hat{\mathbf{R}}$ is much less significant than for the earlier approaches of [4, 7].

Then why are we interested in a detailed comparative study of $\hat{\mathbf{R}}_{\text{FB}}$ and $\hat{\mathbf{R}}$? One reason is, of course, that the earlier approaches, such as those of [4, 7], are still frequently considered, whereas the newer optimal approach (e.g. [2]) is yet to be widely accepted. Another reason is as follows. The optimal method of [2] assumes that the number of signals n is known/given. To guess the value of n, a most practical idea is to use a non-parametric estimator, such as Capon (see, e.g., [8]). The accuracy with which Capon estimates the signal powers (i.e., the diagonal elements of $\mathbf{P}$) is of most importance for estimating n, and we will look at this aspect in some detail (see Section 5 and Section 6). Additionally, such an estimator will also provide initial direction estimates that may prove useful to the optimal ASP method (indeed, if $\mathbf{\Sigma} \neq \sigma^2\mathbf{I}$, then in general the optimal method requires a search over the whole direction space which, in turn, requires initial estimates.) Another goal of this work is to show that Capon based on $\hat{\mathbf{R}}_{\text{FB}}$ outperforms Capon based on $\hat{\mathbf{R}}$, as might be expected (once we have shown that $\hat{\mathbf{R}}_{\text{FB}}$ is more accurate an estimate than $\hat{\mathbf{R}}$).

3. ML ESTIMATION OF R

The MLE of the centro-hermitian covariance matrix $\mathbf{R}$ is given by the minimizer of the optimization problem

$$\min_{\mathbf{R}} \ln|\mathbf{R}| + \text{Tr}(\mathbf{R}^{-1}\hat{\mathbf{R}}) \quad \text{subject to } \mathbf{R} = \mathbf{J}\mathbf{R}^T\mathbf{J} \tag{3}$$

where ln is the natural logarithm, $|\cdot|$ denotes the determinant function, and Tr is the trace operator. Note that

$$\begin{aligned}\text{Tr}(\mathbf{R}^{-1}\hat{\mathbf{R}}) &= \text{Tr}\big((\mathbf{J}\mathbf{R}^T\mathbf{J})^{-1}\hat{\mathbf{R}}\big) \\ &= \text{Tr}\big((\mathbf{R}^T)^{-1}\mathbf{J}\hat{\mathbf{R}}\mathbf{J}\big) = \text{Tr}\big(\mathbf{R}^{-1}\mathbf{J}\hat{\mathbf{R}}^T\mathbf{J}\big),\end{aligned}$$

which implies that we can rewrite the function in (3) as

$$\ln|\mathbf{R}| + \text{Tr}(\mathbf{R}^{-1}\frac{1}{2}(\hat{\mathbf{R}} + \mathbf{J}\hat{\mathbf{R}}^T\mathbf{J})). \tag{4}$$

Since the minimum of the above function with respect to $\mathbf{R}$, without any constraint, is obtained for

$$\mathbf{R} = \hat{\mathbf{R}}_{\text{FB}} = \frac{1}{2}(\hat{\mathbf{R}} + \mathbf{J}\hat{\mathbf{R}}^T\mathbf{J}), \tag{5}$$

and since (5) turns out to be CH, it follows that (5) is the *centro-hermitian MLE* of $\mathbf{R}$.

4. ANALYTICAL STUDY OF $\hat{\mathbf{R}}_{\text{FB}}$ AND $\hat{\mathbf{R}}$

Let $\mathbf{L}_{mn}$ be a permutation matrix such that

$$\mathbf{L}_{mn}\,\text{vec}(\mathbf{X}) = \text{vec}(\mathbf{X}^T), \tag{6}$$

where $\mathbf{X}$ is any $m \times n$ matrix and $\text{vec}(\cdot)$ is the vectorization operator [1]. By using (6), the relation in (2) between $\hat{\mathbf{R}}_{\text{FB}}$ and $\hat{\mathbf{R}}$ can be expressed as follows (sometimes, for simplicity, we omit the subscripts of $\mathbf{L}$)

$$\begin{aligned}\text{vec}(\hat{\mathbf{R}}_{\text{FB}}) &= \frac{1}{2}\,\text{vec}(\hat{\mathbf{R}}) + \frac{1}{2}\,(\mathbf{J}\otimes\mathbf{J})\,\text{vec}(\hat{\mathbf{R}}^T) \\ &= \underbrace{\frac{1}{2}\Big[\mathbf{I} + (\mathbf{J}\otimes\mathbf{J})\,\mathbf{L}\Big]}_{\stackrel{\text{def}}{=}\mathbf{T}}\,\text{vec}(\hat{\mathbf{R}}),\end{aligned} \tag{7}$$

where we have used the fact that $\text{vec}(\mathbf{ABC}) = \big(\mathbf{C}^T\otimes\mathbf{A}\big)\,\text{vec}(\mathbf{B})$ for matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ of compatible dimensions. The symbol $\otimes$ denotes the Kronecker product [1]. In the comparison of the accuracy of $\hat{\mathbf{R}}$ and $\hat{\mathbf{R}}_{\text{FB}}$, we need the covariance of $\text{vec}(\hat{\mathbf{R}})$. Under the assumption that $\{\mathbf{x}(t)\}$ is a zero-mean circular independent Gaussian process, it is well known that

$$\mathbf{C} \stackrel{\text{def}}{=} \text{E}\{\text{vec}(\tilde{\mathbf{R}})\,\text{vec}^*(\tilde{\mathbf{R}})\} = \frac{1}{N}\left(\mathbf{R}^T\otimes\mathbf{R}\right), \tag{8}$$

where $\tilde{\mathbf{R}} = \hat{\mathbf{R}} - \mathbf{R}$. By defining $\tilde{\mathbf{R}}_{\text{FB}} = \hat{\mathbf{R}}_{\text{FB}} - \mathbf{R}$ and making use of (7), it is easily seen that

$$\mathbf{C}_{\text{FB}} \stackrel{\text{def}}{=} \text{E}\{\text{vec}(\tilde{\mathbf{R}}_{\text{FB}})\,\text{vec}^*(\tilde{\mathbf{R}}_{\text{FB}})\} = \mathbf{T}\mathbf{C}\mathbf{T}^*. \tag{9}$$

In order to study the difference between $\mathbf{C}$ and $\mathbf{C}_{\text{FB}}$ in more detail, we first need some preliminary results. Proofs of these results can be found in [3].

Result R1. *The following two properties hold:*

$$\mathbf{L}_{mn} = \mathbf{L}_{mn}^T \tag{10}$$

$$\mathbf{L}_{km}\big(\mathbf{X}\otimes\mathbf{Y}\big) = \big(\mathbf{Y}\otimes\mathbf{X}\big)\mathbf{L}_{ln}, \tag{11}$$

where $\mathbf{X}$ is $m \times n$ and $\mathbf{Y}$ is $k \times l$.

Result R2. *The matrix $\mathbf{T}$ defined in* (7) *is a projection matrix (i.e., hermitian and idempotent).*

Result R3. *If $\mathbf{R}$ is centro-hermitian then*

$$\mathbf{T}\mathbf{C} = \mathbf{C}\mathbf{T}.$$

We are now equipped with all results needed to compare the accuracies of $\hat{\mathbf{R}}$ and $\hat{\mathbf{R}}_{\text{FB}}$. From (9), **R3** and **R2** we have

$$\mathbf{C} - \mathbf{C}_{\text{FB}} = \mathbf{C} - \mathbf{T}\mathbf{C}\mathbf{T}^* = \mathbf{C}(\mathbf{I} - \mathbf{T}). \tag{12}$$

Define

$$\mathbf{M} = \mathbf{I} - \mathbf{T} \tag{13}$$

and observe that this is a (real-valued) projection matrix since $\mathbf{T}$ is so. Moreover, by making use of **R3**, **R2** and (12) it is seen that

$$\mathbf{C} - \mathbf{C}_{\text{FB}} = \mathbf{MCM}^*, \tag{14}$$

which not only shows that

$$\mathbf{C} - \mathbf{C}_{\text{FB}} \geq \mathbf{0} \tag{15}$$

but also provides an explicit expression for the difference $(\mathbf{C} - \mathbf{C}_{\text{FB}})$. Note that the previously derived expression holds in samples of *any* length $N \geq 1$! A numerical example is given in Section 6 to illustrate this more quantitatively.

Remark 1. Note that there is a non-singular transformation from, e.g., $\text{vec}(\hat{\mathbf{R}})$ to the real and imaginary parts of the independent elements in $\hat{\mathbf{R}}$. This implies that it is sufficient to compare $\mathbf{C}$ and $\mathbf{C}_{\text{FB}}$ to assess the relative accuracy of $\hat{\mathbf{R}}$ and $\hat{\mathbf{R}}_{\text{FB}}$.

As an aside, we note that the relation (7) can also be used in performance analyses even in cases when $\mathbf{R}$ is not CH. However, the comparative study above is only valid when $\mathbf{R}$ is CH.

5. FORWARD-ONLY AND FORWARD-BACKWARD CAPON

As an application, we will consider signal power estimation by the Capon method (see, e.g., [8]). The Capon spatial power spectrum estimate is given by

$$\hat{\alpha}(\theta) = \frac{1}{\mathbf{a}^*(\theta)\hat{\mathbf{R}}^{-1}\mathbf{a}(\theta)}, \tag{16}$$

where $\mathbf{a}(\theta)$ denotes a generic column of $\mathbf{A}(\boldsymbol{\theta})$ and θ is a generic element of $\boldsymbol{\theta}$. The direction estimates are given by the location of the peaks of the spectrum, and the corresponding power estimates are given by the height of the peaks. We will refer to this method as the forward-only Capon (F-Capon) method. The *forward-backward* Capon (FB-Capon) spectrum is simply obtained by replacing $\hat{\mathbf{R}}$ in (16) by $\hat{\mathbf{R}}_{\text{FB}}$,

$$\hat{\alpha}_{\text{FB}}(\theta) = \frac{1}{\mathbf{a}^*(\theta)\hat{\mathbf{R}}_{\text{FB}}^{-1}\mathbf{a}(\theta)}.$$

Let us denote the limiting spectrum by α, i.e.,

$$\alpha(\theta) = \frac{1}{\mathbf{a}^*(\theta)\mathbf{R}^{-1}\mathbf{a}(\theta)}.$$

The Capon spectral estimates are in general biased (unless the signal-to-noise ratio tends to infinity). In the following, we will analyze the errors $\hat{\alpha}(\theta) - \alpha(\theta)$ and $\hat{\alpha}_{\text{FB}}(\theta) - \alpha(\theta)$ due to finite sample effects. The analysis is valid for a general θ and not only for the true directions etc.. By employing a second-order approximation of the error $\hat{\alpha}(\theta) - \alpha(\theta)$ it is possible to show that the bias for large N is given by [3]

$$\text{E}\{\hat{\alpha} - \alpha\} = \frac{1-m}{N}\frac{1}{\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}} = \frac{(1-m)}{N}\alpha. \tag{17}$$

Note that the bias is always negative since $m > 1$.

The bias for FB-Capon can be obtained by following the same steps as in the proof of (17). If the structural relations between the forward-only and the forward-backward cases given in Section 4 are utilized, then it is straightforward to derive the bias expression for FB-Capon

$$\text{E}\{\hat{\alpha}_{\text{FB}} - \alpha\} = \frac{1-m}{2N}\alpha, \tag{18}$$

which is *half* that of F-Capon (see (17)). This result for the spatial Capon spectrum is similar to the result derived in [5] for temporal spectral estimation by the Capon method.

The asymptotic (for large N) variance of $\hat{\alpha}$ can be shown to be the same for F-Capon and FB-Capon and given by [3]

$$\text{E}\{(\hat{\alpha} - \alpha)^2\} = \text{E}\{(\hat{\alpha}_{\text{FB}} - \alpha)^2\} = \frac{1}{N}\frac{1}{(\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a})^2} = \frac{\alpha^2}{N}. \tag{19}$$

The above results indicate that FB-Capon is to be preferred over F-Capon due to the smaller bias. In the next section, we provide a numerical example to illustrate this fact.

6. NUMERICAL EXAMPLES

In this section, we present two simulation examples to illustrate the theoretical results derived in the paper. We begin with an example that shows more quantitatively the difference in accuracy between $\hat{\mathbf{R}}$ an $\hat{\mathbf{R}}_{\text{FB}}$. Then we go on and study the Capon estimates in the second example.

Example 1 The data are generated according to the model in Section 2. The array is uniform and linear and consists of $m = 5$ sensors separated by half of the carrier's wavelength. Two signal wavefronts arrive at the array from the directions $\theta_1 = -15°$ and $\theta_2 = 15°$ measured relative to array broadside. The signal covariance matrix is

$$\mathbf{P} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The noise covariance matrix, $\boldsymbol{\Sigma}$, is a hermitian Toeplitz matrix, with the first row given by

$$\begin{bmatrix} 1 & 0.4 + j0.3 & 0.1 + j0.07 & 0 & 0 \end{bmatrix}.$$

In Figure 1, the difference between $\mathbf{C}$ and $\mathbf{C}_{\text{FB}}$ is illustrated for different data lengths. In the figure, $\text{Tr}(\mathbf{C})$ and $\text{Tr}(\mathbf{C}_{\text{FB}})$ are plotted along with the corresponding sample estimates obtained from 1000 independent simulations. It can be seen that the empirical and the theoretical results are in perfect agreement for all N.

Example 2 In this example, we change the signal and the noise covariances to

$$\mathbf{P} = \begin{bmatrix} 10 & 0 \\ 0 & 1 \end{bmatrix},$$
$$\boldsymbol{\Sigma} = \mathbf{I}.$$

The other parameters are the same as in the previous example. The directions θ_1, θ_2, and the corresponding powers are estimated by the F-Capon and the FB-Capon methods. In Figure 2, we show the "bias" in the power estimates. By bias, we mean the difference between the average power estimates and their limiting values. The

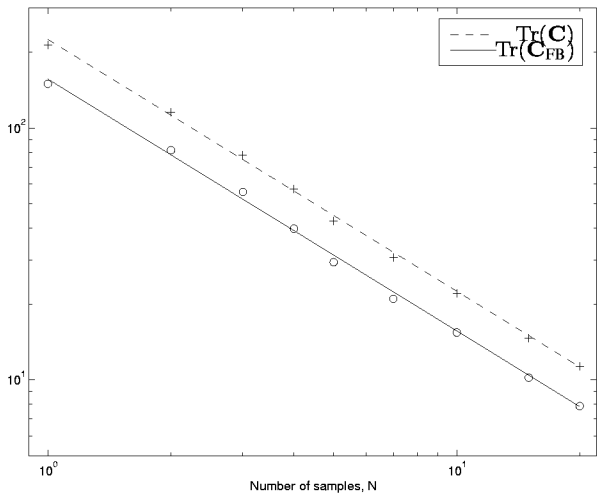

Figure 1: Comparison of the accuracies of the forward-only and the forward-backward sample covariance matrix estimates. The plot shows $\mathrm{Tr}(\mathbf{C})$ and $\mathrm{Tr}(\mathbf{C}_{\mathrm{FB}})$ for different N. The sample estimates of these quantities obtained from the empirical mean-square-error matrices are shown by the symbols + and o, respectively.

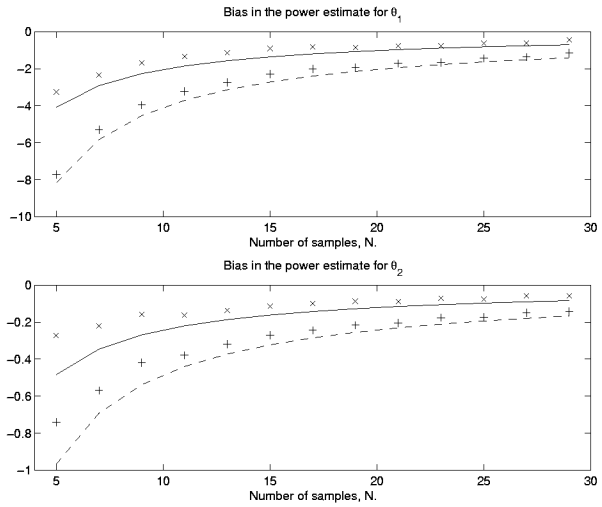

Figure 2: Bias of the F-Capon (+) and the FB-Capon (x) power estimates versus the sample length. The upper and lower plots correspond to θ_1 and θ_2, respectively. The theoretical bias given in (17) and (18) is shown by the dashed and the solid line, respectively.

sample values shown in Figure 2 are the averages over 1000 independent trials. The corresponding theoretical asymptotic bias values computed by (17) and (18) are also shown in Figure 2. Clearly, the asymptotic bias analysis in Section 5 gives a good prediction of the simulation results also for small N. Note that the results derived in Section 5 are true for fixed θ. Here, $\hat{\alpha}$ is computed at the direction estimates which are realization dependent, yet the results apply reasonably well.

The standard deviations of $\hat{\alpha}-\alpha$ and $\hat{\alpha}_{\mathrm{FB}}-\alpha$ in the simulation are similar as predicted by the (asymptotic) result in (19).

One point to be made here is that small peaks in the Capon spectrum shows up more clearly for FB-Capon than for F-Capon due to the smaller bias (see Figure 2). The detection of the number of signals, n, may therefore be easier when using FB-Capon.

7. CONCLUSIONS

We have studied the relative accuracy of the forward-only ($\hat{\mathbf{R}}$) and the forward-backward ($\hat{\mathbf{R}}_{\mathrm{FB}}$) sample covariance matrix estimates. A simple proof that $\hat{\mathbf{R}}_{\mathrm{FB}}$ is the maximum likelihood estimate of a centro-hermitian covariance matrix $\mathbf{R}$ was given. The main contribution is a quantitative study of the estimation error covariances $\mathbf{C}$ and $\mathbf{C}_{\mathrm{FB}}$ of $\hat{\mathbf{R}}$ and $\hat{\mathbf{R}}_{\mathrm{FB}}$, respectively. An explicit expression was given for the difference $\mathbf{C}-\mathbf{C}_{\mathrm{FB}}$ showing that $\hat{\mathbf{R}}_{\mathrm{FB}}$ indeed is a more accurate estimate than $\hat{\mathbf{R}}$ (in cases when the true covariance matrix is centro-hermitian). We also analyzed the spatial power spectrum estimation by the Capon method based on either $\hat{\mathbf{R}}$ (F-Capon) or $\hat{\mathbf{R}}_{\mathrm{FB}}$ (FB-Capon). We showed that the asymptotic bias for FB-Capon is half that of F-Capon, whereas the asymptotic variance is the same. Finally, two numerical examples were presented to illustrate the abovementioned results.

8. REFERENCES

[1] A. Graham. *Kronecker Products and Matrix Calculus with Applications*. Ellis Horwood Ltd, Chichester, UK, 1981.

[2] M. Jansson, B. Göransson, and B. Ottersten. "Analysis of a Subspace-Based Spatial Frequency Estimator". In *Proceedings of ICASSP'97*, pages 4001–4004, München, Germany, April 1997.

[3] M. Jansson and P. Stoica. "Forward-Only and Forward-Backward Sample Covariances – A Comparative Study". Submitted to Signal Processing, 1998.

[4] B. H. Kwon and S. U. Pillai. "A Self Inversive Array Processing Scheme for Angle-of-Arrival Estimation". *Signal Processing*, 17(3):259–277, Jul. 1989.

[5] H. Li, J. Li, and P. Stoica. "Performance Analysis of Forward-Backward Matched-Filterbank Spectral Estimators". *IEEE Trans. on Signal Processing*, 46(7):1954–1966, Jul. 1998.

[6] Q. A. Nguyen. "On the Uniqueness of the Maximum-Likelihood Estimate of Structured Covariance Matrices". *IEEE Trans. on Acoustics, Speech, and Signal Processing*, ASSP-32(6):1249–1251, Dec. 1984.

[7] B. D. Rao and K. V. S. Hari. "Weighted Subspace Methods and Spatial Smoothing: Analysis and Comparison". *IEEE Trans. ASSP*, 41(2):788–803, Feb. 1993.

[8] P. Stoica and R. Moses. *"Introduction to Spectral Analysis"*. Prentice Hall, Upper Saddle River, New Jersey, 1997.