Processing Data for Colored Noise Using a Dynamic State Estimator

MARIUS-CONSTANTIN POPESCU[1] ONISIFOR OLARU[2] NIKOS MASTORAKIS[3]
[1]Faculty of Electromechanical and Environmental Engineering
University of Craiova
[2]Faculty of Engineering
Decebal Bv, No.107, 200440, Craiova
University of Constantin Brancusi
ROMANIA
[3]Technical University of Sofia
BULGARIA
popescu.marius.c@gmail.com onisifor.olaru@yahoo.com mastor@wses.org

Abstract: - This paper treats data reduction in array processing for the spatially colored noise case. The purpose is to reduce the computational complexity of the applied signal processing algorithms by mapping the data into a space of lower dimension by means of a linear transformation. We discuss ways to implement the transformation and show that it suffices to estimate the array covariance matrix instead of the noise covariance matrix in the design process of the optimal transformation. Computer simulations are given that illustrate the problem of interference from out-of-band-sources that result when a beamspace transformation is designed to focus on a particular sector. The presents an dynamic state estimator. The method uses ANN based bus load prediction for the prediction step in the DSE.

Key-words: - Adaptive data reduction, Colored noise, Dynamic state estimation, Load prediction.

1 Introduction

The computational complexity of the algorithms applied to the array signal processing problem is heavily dependent on the number of sensors in the array. Since large arrays with many sensors are preferable from an estimation accuracy viewpoint, accuracy and computational complexity are conicting issues. This has led to diferent schemes for dimension reduction via linear transforms, beamspace transformations, which reduce the computational load. Besides a reduction in computational complexity, a beamspace transformation can have other advantageous ejects, such as reduced bias, reduced sensitivity to directional interference, etc [4], [7]. Diferent criteria can be used when deriving the transform. If one knows from which angular sectors that signals may emanate from, a possible approach is to design a transformation that focuses on these sectors. This can, e.g., be accomplished by a bank of conventional beamformers, the output of which are collected in a new vector, with reduced size. A more sophisticated variation of this theme is to design a beamspace transformation that maximizes the signal-to-noise ratio for signals impinging from the preselected. sectors [12], or minimizes the interference power under the constraint that signals from the sectors are left un-distorted by the transform. The latter method is easily implemented by a bank of linearly constrained minimum variance beamformers, see, e.g., [10], [11]. Yet another method is to design a transformation that preserves the Cramér-Rao bounds (CRB), for the parameters of interest, which is the approach taken in the present paper. This criterion is also considered in [1], [2], [8] for the white-noise case, i.e., an equal amount of un correlated noise at each sensor, and a stochastic signal model.

This paper concentrates on the spatially colored noise case, which is more realistic in any scenario, due to, e.g., directional interference from other sources (‘out-of-band sources'), mutual coupling between sensors, etc. The derived transform depends, in addition to the unknown DOAs, on the color of the noise process. However, it is shown herein that it is possible to implement the proposed transform without knowledge of the noise color and the exact directions of arrival; a fact that has great practical implications in a real scenario.

The paper presents a design approach and computer simulations that support the theoretical results. The simulations evaluate the approach against the method of spheroidal sequences [4], and the method presented

in [1], [8] with respect to the e!ect of out-of-band sources. The outline of the paper is as follows.
Section 2 discusses the signal model, formulates the problem and derives the main results of the paper.
Section 3 treats practical implementation issues and Section 4 presents the computer simulations.
Section 5 treats dynamic load prediction.
Section 6 discusses the analysis of anomalous data.
Section 7 proposed DSE scheme.
Section 8 treats test system simulation.
Finally, conclusions are given in Section 9.

2 The optimal model

Consider an array of m sensors receiving p planar, narrowband, waveforms from the directions $\{\theta_1 \cdots \theta_p\}$. The sensor outputs are modelled by the relation

$$\tilde{y}(t) = \tilde{A}(\theta_0)s(t) + \tilde{e}(t), \quad (1)$$

where $\tilde{A}(\theta_0)$ is the array steering matrix, $\tilde{A}(\theta_0) = [\tilde{a}(\theta_1) \cdots \tilde{a}(\theta_p)]$, and the source signals are collected in the vector

$$s(t) = [s_1(t) \cdots s_p(t)]^T \quad (2)$$

The parameter vector θ_0 is a column vector containing the true directions of arrival. The notation $\tilde{x}$ is used to designate that the quantity, x, belongs to elementspace, i.e., the original data set to which the beamspace transformation is applied. The signal $s(t)$ and the noise $\tilde{e}(t)$ are assumed to be independent, temporally white, zero-mean, complex Gaussian random variables with second-order moments

$$E[s(t)s^*(t)] = R_{ss}, \quad E[\tilde{e}(t)\tilde{e}^*(t)] = \tilde{Q}. \quad (3)$$

The superscript, „*”, denotes complex conjugate transpose. The matrix $\tilde{Q}$ is Hermitian and positive definite, but otherwise arbitrary. The signal covariance matrix, R_{ss}, has rank $d \le p$. The latter condition implies that the signals may be spatially correlated, e.g., due to multipath, etc. The array covariance matrix is given by

$$\tilde{R} = \tilde{A}(\theta_0)R_{ss}\tilde{A}^*(\theta_0) + \tilde{Q}. \quad (4)$$

In general, to avoid ambiguous parameter estimates, some sort of structure must be imposed on the noise covariance matrix, see, e.g., [5], [8]. However, since the main question here is preservation of the CRB of the transformed data, we do not treat the problem of parameter identi"ability herein. Now, we introduce a linear transformation of the data from the complex m-dimensional vector space, $\mathbb{C}^m$ (m being the number of sensors) to the complex n-dimensional vector space $\mathbb{C}^n$ where $n \le m$. Denoting the $m \times n$ beamspace transformation by T, the beamspace signal model becomes

$$y(t) = T^*\tilde{A}(\theta_0)s(t) + T^*\tilde{e}(t) = A(\theta_0)s(t) + e(t) \quad (5)$$

where $A = T^*\tilde{A}$ is the beamspace steering matrix and $e(t)$ is the resulting beamspace noise.
Based on N snapshots of the beamspace data, $y(t)$, the directions of arrival, $\theta_0 = [\theta_1 \cdots \theta_p]^T$ are to be estimated. Before stating and proving the main result of the paper, we briefly discuss the white-noise case, which is addressed in [1]. In this case, the spatial noise covariance matrix is a scaled version of the identity matrix, and the array covariance matrix is given by

$$\tilde{R} = \tilde{A}(\theta_0)R_{ss}\tilde{A}^*(\theta_0) + \sigma^2\tilde{I} \quad (6)$$

where σ^2 is the variance of the additive thermal noise always present in the receiving equipment in a sensor array. Note also that the white-noise model assumes an equal amount of white-noise at each sensor. In order to retain the white-noise model in beamspace, an orthogonality constraint must be imposed on the beamspace transformation matrix: $T^*T=I$. In [1], it is shown that the optimal beamspace transformation matrix satisfies the condition

$$R(T) \supseteq R[\tilde{a}(\theta_1 \cdots \tilde{a}(\theta_p)\tilde{d}(\theta_1) \cdots \tilde{d}(\theta_p))] \quad (7)$$

In the above equation, $R(X)$ denotes the range space of X while $\tilde{d}(\theta)$ is the derivative of the steering vector with respect to θ. Thus, to obtain a beamspace CRB which equals the elementspace CRB, the range space of the beamspace transformation matrix should include the subspace spanned by the steering vectors and the derivative of the steering vectors evaluated at the true directions of arrival. It should be noted that the minimal beamspace dimension that satis"es the condition in Eq.(7) is $n=2p$. We now return to the colored noise case. The following theorem states a condition on the range space of the beamspace transformation that is sufficient to guarantee that the beamspace CRB equals the elementspace CRB.

Theorem 1. The CRB for the estimate of θ_0 in beamspace is equal to the CRB in element space, provided that the beam space transformation matrix T satisfies the following condition:

$$R(T) \supseteq R(\tilde{Q}^{-1}U(\theta_0)) = R(\tilde{R}^{-1}U(\theta_0)),$$

where $U(\theta_0)$ is defined as

$$U(\theta_0) \overset{DEF}{=} \left[\widetilde{A}(\theta_0)\widetilde{D}(\theta_0)\right] = \left[\widetilde{a}(\theta_1)\cdots\widetilde{a}(\theta_p)\widetilde{d}(\theta_1)\cdots\widetilde{d}(\theta_p)\right]$$

Proof. Without loss of generality, partition the beamspace transformation matrix in the following way:

$$T = \left[\widetilde{Q}^{-1}UB : C\right] \overset{DEF}{=} \left[T_1 : T_2\right], \tag{8}$$

where B, with dimension $2p \times 2p$, is a full-rank matrix and C, with dimension $m \times (n\text{-}2p)$, is an arbitrary matrix. For convenience, the dependence on θ_0 is suppressed. The partitioning reflects the condition $R(T) \supseteq R\left(\widetilde{Q}^{-1}U\right)$. The beamspace signal becomes

$$y(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} T_1^* \widetilde{y}(t) \\ T_2^* \widetilde{y}(t) \end{bmatrix}. \tag{9}$$

Concentrating on $T_1^* \widetilde{y}(t)$, we can write

$$T_1^* \widetilde{y}(t) = T_w^* \widetilde{y}_w(t), \tag{10}$$

where

$$T_w = \widetilde{Q}^{-\frac{1}{2}}UB, \quad \widetilde{y}_w = \widetilde{Q}^{-\frac{1}{2}}\widetilde{y}(t), \tag{11}$$

Now, using the result of [1], the optimal transformation that preserves the CRB for the whitened data set, $\widetilde{y}_w(t)$, should satisfy the condition $R(T_w) \supseteq R\left(\widetilde{Q}^{-1/2}U\right)$. This condition is satisfied because B is full rank, see (1). Now, since $y_1(t) = T_1^* \widetilde{y}(t) = T_w \widetilde{y}_w(t)$, we obtain the same CRB for estimates based on the data $y_1(t)$ as for estimates based on elementspace data, $\widetilde{y}(t)$.

The data $y_2(t)$ can, in fact, be ignored since it conveys no more information about the parameters than is already present in $y_1(t)$. We have thus shown that the beamspace CRB equals that of the elementspace provided that $R(T) \supseteq R\left(\widetilde{Q}^{-1}U\right)$, which concludes the first part of the proof. It remains to show that $R\left(\widetilde{Q}^{-1}U\right) \supseteq R\left(\widetilde{R}^{-1}U\right)$. The proof of this is based on the following two claims:

$$\text{I. } \widetilde{R}^{-1}\widetilde{A} = \widetilde{Q}^{-1}\widetilde{A}\left(I + R_{ss}\widetilde{A}^*\widetilde{Q}\widetilde{A}\right)^{-1}, \tag{12}$$

$$\text{II. } \widetilde{Q}^{-1}\widetilde{D} = \widetilde{R}^{-1}\left[\widetilde{A}\widetilde{D}\right]\begin{bmatrix} R_{ss}\widetilde{A}\widetilde{Q}^{-1}\widetilde{D} \\ I \end{bmatrix} \tag{13}$$

Proof of Claim I. Multiplying the right-hand side by $\widetilde{R}$, we have

$$\widetilde{R}^{-1}\widetilde{A} = \widetilde{Q}^{-1}\widetilde{A}\left(I + R_{ss}\widetilde{A}^*\widetilde{Q}\widetilde{A}\right)^{-1} =$$
$$\left(\widetilde{A}R_{ss}\widetilde{A}^* + \widetilde{Q}\right)\widetilde{Q}^{-1}\widetilde{A}\left(I + R_{ss}\widetilde{A}^*\widetilde{Q}^{-1}\widetilde{A}\right)^{-1} =$$
$$\widetilde{A}\left(R_{ss}\widetilde{A}^*\widetilde{Q}^{-1}\widetilde{A} + QI\right)\left(R_{ss}\widetilde{A}^*\widetilde{Q}^{-1}\widetilde{A} + I\right)^{-1} = \widetilde{A}$$

which proves I.

Proof of Claim II. Multiplying the right-hand side by $\widetilde{Q}$, we get

$$\widetilde{Q}\widetilde{R}^{-1}\left[\widetilde{A} \quad \widetilde{D}\right]\begin{bmatrix} R_{ss}\widetilde{A}^*\widetilde{Q}^{-1}\widetilde{D} \\ I \end{bmatrix} =$$
$$\widetilde{Q}\widetilde{R}^{-1}\left(\widetilde{A}R_{ss}\widetilde{A}^* + \widetilde{Q}\right)\widetilde{Q}^{-1}\widetilde{D} = \widetilde{D},$$

which proves II.

From Claim I, it follows that $\widetilde{R}^{-1}\widetilde{A} = \widetilde{Q}^{-1}\widetilde{A}M_1$, where M_1 is the full-rank matrix defined as the inverse of $I + R_{ss}\widetilde{A}^*\widetilde{Q}\widetilde{A}$. This implies that

$$R\left(\widetilde{Q}^{-1}U\right) = R\left(\widetilde{R}^{-1}\widetilde{A}\right). \tag{14}$$

The implication of Claim II is that

$$\widetilde{Q}^{-1}\widetilde{D} = \widetilde{R}^{-1}\left[\widetilde{A}\ \widetilde{D}\right]\begin{bmatrix} R_{ss}\widetilde{A}^*\widetilde{Q}^{-1}\widetilde{D} \\ I \end{bmatrix} = \widetilde{R}^{-1}UM_2 \tag{15}$$

where M_2 is the full-rank matrix defined by

$$M_2 = \begin{bmatrix} R_{ss}\widetilde{A}^*\widetilde{Q}^{-1}\widetilde{D} \\ I \end{bmatrix}. \tag{16}$$

Thus,

$$R\left(\widetilde{Q}^{-1}D\right) \supset R\left(\widetilde{R}^{-1}U\right). \tag{17}$$

Using Eqs. (14) and (17), and noting that the range spaces of $\widetilde{Q}^{-1}U$ and $\widetilde{R}^{-1}U$ have the same dimension, yields the desired result

$$R\left(\widetilde{Q}^{-1}U\right) = R\left(\widetilde{R}^{-1}\widetilde{U}\right). \tag{18}$$

The proof of Theorem 1 is thereby complete. The extension of the result in Theorem 1 to the case where θ contains several parameters, e.g., elevation, is straightforward. The matrix $U(\theta_0)$ is modified to include the derivatives of the steering vector with respect to these parameters as well. The result in

Theorem 1 also applies to other signal models, such as the deterministic parameterized signal model considered in [3]. There is a close similarity of the result in Claim I to the matched filter solution for detecting a single source at a known bearing in a colored noise environment. In that case, one can substitute the noise covariance matrix in the matched filter by the array covariance matrix without loss of detection performance, see, e.g., [7]. The explanation to this is that the matched filter is unique up to a multiplicative scalar, and since it holds that $\widetilde{R}^{-1}a(\theta_0)$ is proportional to $\widetilde{Q}^{-1}a(\theta_0)$, the result follows.

The result of Claim I can thus be seen as a generalization of the single-source case. However, the result in Claim II is more intricate and has no obvious analogy. It is the combination of the claims into a single condition on the range space of the transformation matrix that is important. Thus, the fact that the range space of $\widetilde{R}^{-1}U(\theta_0)$ is the same as that of $\widetilde{Q}^{-1}U(\theta_0)$ is a key point in Theorem 1. If this was not true, an estimate of the noise covariance matrix $\widetilde{Q}$ would be needed. The array covariance matrix, however, can be estimated by a simple time average. In a real scenario, the beamspace transformation can be updated to accommodate for a nonstationary scenario where both the source locations and the noise color may change with time. The matrix inversion lemma may be employed to recursively track the array covariance matrix in such a situation. An LMS - based update of the optimal beamspace transformation is also possible, for further reducing the computational complexity.

3 Implementing

We start this section by suggesting the following algorithm for beamspace DOA estimation:

(1) Determine a set of interesting angle intervals that will be processed in beamspace. Define also the following quantity:

$$U_f \overset{DEF}{=} \left[\widetilde{a}(\theta_{f1})\cdots\widetilde{a}(\theta_{fn})\right], \tag{19}$$

where $\widetilde{a}(\theta_{fi})$, i=1,…, n, is a sufficiently dense set of fictitious array steering vectors located within the selected angle intervals.

(2) Estimate the array covariance matrix as

$$\widetilde{R} = \frac{1}{N}\sum_{t=1}^{N}\widetilde{y}(t)\widetilde{y}^{*}(t) + \varepsilon I,$$

where εI a possible regularization.

(3) Calculate the quantity

$$T_1 = \widetilde{R}^{-1/2}U_f W, \tag{20}$$

where W is a diagonal weighting matrix to be defined subsequently.

(4) Decide on an initial beamspace dimension, denoted n_s, that is large enough to accommodate for the maximum number of expected sources within the sectors.

(5) Apply the singular-value decomposition, SVD, to the quantity in Eq. (20). The transformation is then created as

$$T = \widetilde{R}^{-1/2}\left[q_1\cdots q_{n_s}\right], \tag{21}$$

where n_s is the number determined in Step 4 and q_i, i=1,…,n_s , are the corresponding left singular vectors.

(6) Orthogonalize T (optional).

(7) Optional step. Estimate the number of sources in the reduced beamspace domain using, e.g., the minimum description length (MDL), principle, see [9]. Do Step 4 again where now ns is set to min(n_s,2 $\hat{p}$). Repeat Steps 5 and 6.

(8) Estimate the directions of arrival.

Step 1 can preferably be based on the Capon spectrum, since this is used in the design of the beamspace transformation, see below. In this step, some kind of detection problem must be solved. A possible way to do this is just to look at the spatial spectrum and decide on which sectors to choose. Another way is the automated approach, in which a hypothesis problem can be posted, resulting in a thresholding problem. We do not treat this problem further in the present paper though. In Step 2, a possible regularization of the data covariance estimate is included. This regularization makes the method robust to estimation errors in the data covariance matrix when the number of samples is low. The choice of n_s in Step 4 may, e.g., be based on prior information on the maximum number of sources present in the sectors or such an estimate may be obtained by using, e.g., the MDL principle in elementspace. Another possibility is to choose the number ns as the number of significant singular values in the SVD of Step 5. Step 7 facilitates a possible further reduction in the beamspace dimension, accompanied by a corresponding reduction of the computational load. Finally, we note that the reason why the derivatives of the fictitious steering vectors are not included in U_f is that these can be well approximated by a finite difference of the included fictitious steering vectors, implying that the range space of U_f approximates the range space of $U(\theta_0)$. The introduction of the weighting matrix in Step 3 is pertinent to the success of choosing the proper subspace dimension that the beamspace transformation shall preserve. We will now discuss two different choices of weighting matrices, W. The first is the identity matrix, and the other is created by putting the

Capon spectrum for the fictitious directions of arrival on the main diagonal. Starting with the identity matrix, a column of the matrix defined in Eq. (20) has the magnitude

$$\left\|\hat{\tilde{R}}^{-1/2}\tilde{a}\left(\theta_{fi}\right)\right\|^2 = \tilde{a}^*\left(\theta_{fi}\right)\hat{\tilde{R}}^{-1/2}\tilde{a}^*\left(\theta_{fi}\right) \qquad (22)$$

which is exactly the inverse of the Capon spectrum at the point specified by the fictitious DOA. Thus, truncating the SVD with no weighting can result in noise subspace directions being chosen instead of the signal subspace directions, yielding a loss in performance.

The second weighting is the Capon weighting. With this weighting, a column of the matrix in Eq. (20) has the magnitude

$$\left\|\hat{\tilde{R}}^{-1/2}\frac{1}{\tilde{a}^*\hat{\tilde{R}}^{-1}\tilde{a}^*}\right\|^2 = \frac{1}{\tilde{a}^*\hat{\tilde{R}}^{-1}\tilde{a}^*}, \qquad (23)$$

which equals the Capon spectrum.
This is the recommended weighting, since the problem of choosing the wrong subspace in the truncation of the SVD is eliminated.

4 Numerical examples

In this section we present computer simulations that support the theoretical result. First however, we briefly describe two approaches for finding a beamspace transformation matrix that reduce the dimensionality of the problem. The author of [10] presents a method for calculating a near optimal, in the CRB sense, beamspace transformation matrix for the white-noise case. This method incorporates the use of a set of 'design DOAs' that cover an interval $[\theta_a, \theta_b]$, within which the sources of interest are assumed to be located. The steering vectors of the design DOAs and the respective derivatives are collected in a matrix for which the singular-value decomposition (SVD), is calculated. The final transformation is then created by selecting as columns in the transformation matrix the left singular vectors corresponding to the n most significant singular values. As the singular vectors are orthogonal, the orthogonality constraint of the beamspace transformation matrix is satisfied. This method will be called the 'white-noise method', since it preserves the Cramér-Rao bounds for the spatially white-noise case. In [4], the first n spheroidal sequences are used as columns in the beamspace transformation matrix. The authors select $n=\lfloor 2Bm \rfloor - 2$, where $\lfloor . \rfloor$ denotes the integer part. The method supplies orthonormal beamformer weights that capture most of the energy of a spatially band-limited process with a flat energy spectrum within the sector $[\theta_a, \theta_b]$, see [4], [8] for details. Finally, it should be noted that the parameter values used in Section 4 below result in n=2 if the rule defined above is used. Since the initial beamspace dimension is chosen as 6 in Section 4, the resulting beamspace transformation matrix captures an amont of energy well above the limit implied by the rule aboveu. We employ a uniform linear array with 25 sensors, separated by half a wavelength. Two source signals are located at -7^0 and -4^0 within the interesting sector, defined by $[-10^0:10^0]$ relative broadside of the array. There are three sources located outside the sector, the first at 13^0, the second at

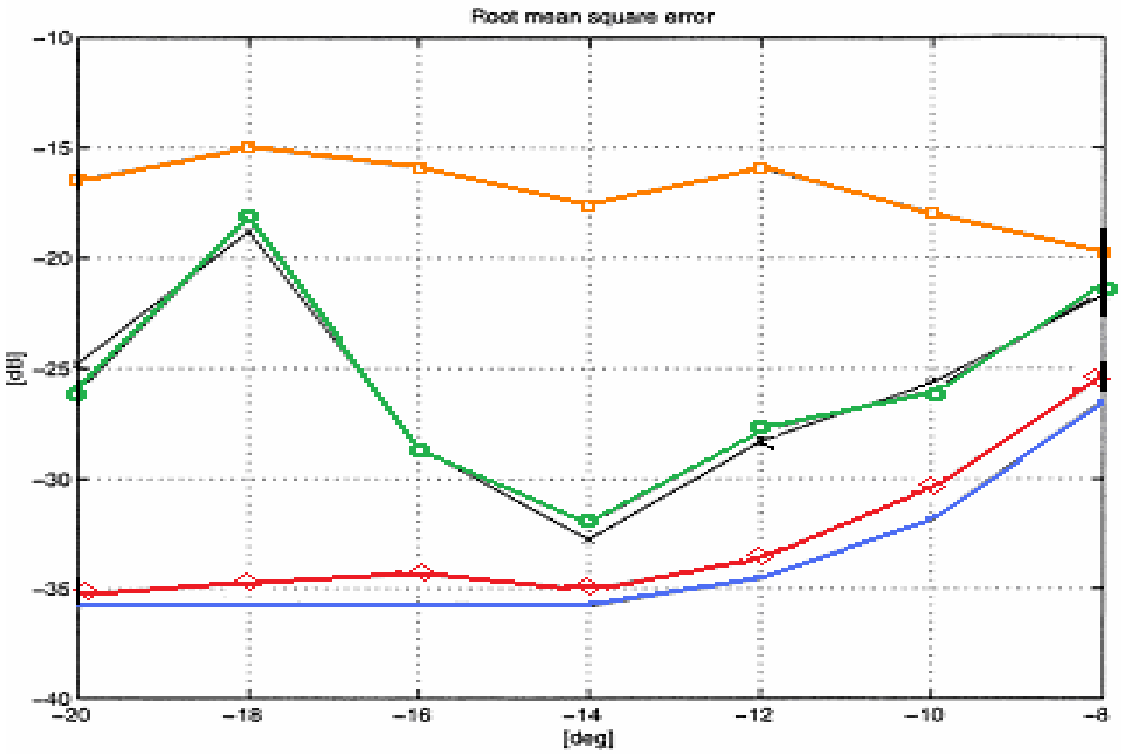

Fig.1: Number of snapshots: 200, ε=0. Solid: square root of the Cramér-Rao bound, ○-solid: the proposed method with Capon weighting, □-solid: the proposed method with no weighting, ×-solid: spheroidal sequences, ◊-solid: the white-noise approach.

18^0 and the location of the third is swept from -20^0 to -8^0 in steps of 2^0. The sources outside the sector is referred to as 'out-of-band' sources. Note that the out-of-band sources act as interference, since we try to estimate only the direction of arrival of the sources within the interesting sector. In effect, we have a spatially colored noise problem.

line is the square root of the element space Cramer-Rao bound.

The proposed method, with the proper weighting, yields an estimate with an accuracy very close to the bound. The other two methods, that have similar performance, produce estimates with considerably higher root-mean-square error. Note the severe

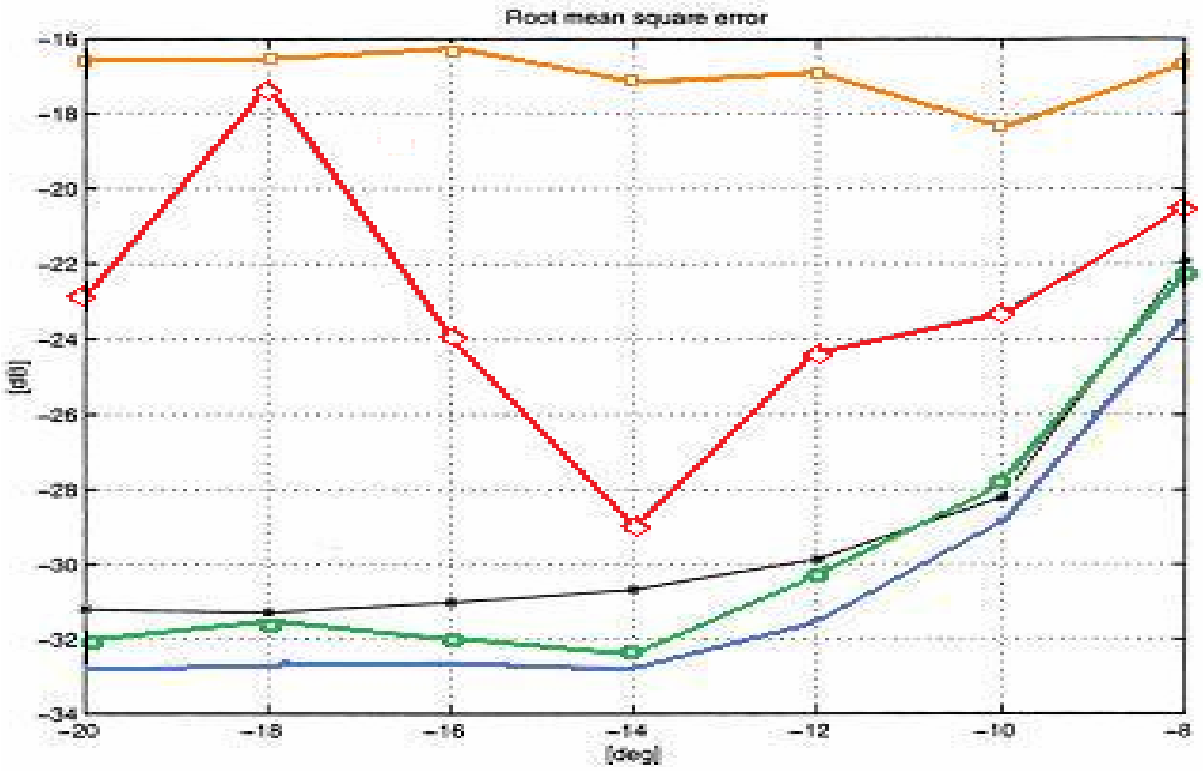

Fig.2: Number of snapshots: 50. Solid: Square root of the Cramér-Rao bound, □-solid: the proposed method with no regularization,*-solid: the proposed method with ε=0.5, ○-solid: the proposed method with ε=1, ◊-solid: the white-noise approach.

The initial beamspace dimension is set to 6, see Steps 4 and 7 of Section 3. The SNR values of the sources are equal and set to 10 dB relative to the noise power in one element. The number of snapshots is either 200 or 50 and the number of Monte Carlo trials is 100. A white-noise model is employed in beamspace together with the MDL estimate of the number of sources. The directions of arrival are estimated using the 'stochastic maximum likelihood' method, see, e.g., [6]. The numerical search is initiated in the following way: if the number of estimated sources in the MDL step is $\hat{p}$, the search is initiated in the $\hat{p}$ true DOAs within and nearest to the interesting sector. In Fig. 1, the root-mean-square error for the source located at -7^0 is plotted as a function of the location of the moving out-of-band source for the proposed method, with the two diferent weightings discussed in 3, and the methods of [4], [1], [8]. The number of snapshots is 200. No regularization of the estimated array covariance matrix is used in this case. As the number of samples is 200, the estimate of the array covariance matrix is of sufficient quality. In the figure, the solid

degradation of the presented method if a proper weighting is not used. In Fig. 2, the approach is evaluated against the white-noise approach when the number of snapshots is decreased from 200 to 50. The result of this is that the quality of the array covariance estimate deteriorates, which in turn degrades the performance of the method. This is clearly seen from the result in Fig. 2. However, the good performance is restored by adding a scaled identity matrix to the estimated data covariance matrix making the method more robust to poor sample support (see Fig. 2). The excellent attenuation of the out-of-bandsources for the proposed method results in that the estimated model order for this method is in general lower than for the other two methods. This fact implies that the resulting multi-dimensional search is simplified for the proposed scheme. Thus, even though the implementation of the proposed transformation is computationally heavier than the other methods, the reduced model order compensates this. Another interesting approach was proposed by Irving and Sterling [9].

They suggested to carry out data validation at the substation level where all the bus couplers and the

circuit breakers are represented in detail. To be able to handle zero impedance branches, they suggested to take as variables the powers flowing through the circuit breakers. Later on, Clewer *et al* [10] extended the approach to the whole system by proposing a 4-stage iterative procedure based on a detailed substation representation together with a bus level network modeling. Unfortunately, the method involves many inter-related steps, which makes it rather complex. In addition, it is not guaranteed to converge [10]. Once the state vector is estimated, the extreme outliers are identified and deleted from the measurement set. Then, one iteration of the algorithm is executed starting from the previous solution. This measurement deletion is performed in order to cancel out the influence of the extreme outliers on the estimates. They are defined as measurements whose weighted residual have an amplitude larger than a given threshold, chosen between 6 and 10. In an attempt to decrease the complexity of the foregoing method while meeting the need of a generalized state estimation raised by Slutsker *et al* [22], [23], Monticelli [11], Alsac *et al.* [4], and Abur *et al* [1] advocated the use of a 2-step procedure that proceeds as follows.

First, a super-node-based state estimation is executed and a residual analysis is performed. In the event that the residuals associated with a branch or a bus are found to be large, then a detailed representation of the suspected substations is carried out and the state vector is expanded accordingly. Finally, the expanded state vector is estimated through either a conventional estimator or a LAV estimator [1]. However, both methods suffer from the aforementioned weaknesses inherent to any post-estimation approach.

5 Dynamic load prediction

The DLP method is based on more realistic and physically meaningful foundation in comparison to the conventional state prediction model. The main arguments which form the basis of DLP method are:

i) It is loads and generation which actually drive the system dynamics which is of concern in DSE (dynamic state estimator) of power system.

ii) Bus loads are more or less independent of each other.

iii) Bus loads follow a reasonably regular pattern and so are easier to predict with reasonable accuracy.

For the purpose of DLP two types of busbars are defined:

a) Load bus bars - where only loads P_L and Q_L are connected. Let, NL be the number of these bus bars, and

b) Generator busbars - where generators are connected. Since generator busbars may also have loads in addition to generation so the net power injection is considered at these busbars. Also, at generating busbars the voltage is controlled, therefore real power injection P_G and voltage magnitude V_G are used as the prediction state variables at these busbars. Let NG be the number of these busbars.

In order to satisfy the active power balance in the system one generating busbars is taken as slack busbar, its active power P_G can not be specified and therefore, is not included in the prediction variables. Thus, the complete prediction vector S is defined as:

$$S \overset{\Delta}{=} \left[P_L^T, Q_L^T, P_G^T, V_G^T\right]^T \tag{24}$$

S consists of $2N$-1 components, where N is the number of busbars and $N=N_L+N_G$. In order to make regular exchange of variables $S <=> X$ possible, the voltage angle of slack busbar is considered as reference angle and is set to zero. This makes the number of variables in both S and X equal to $2N$-1. Active and reactive loads P_{Li}, Q_{Li}; each busbar including the load at generator busbars (if any), is predicted for next time instant (k+1) using the ANN based bus load prediction [9]. Active bus injections (generation - load) for generating busbars can not be directly predicted for the simple reason that generation has to adapt itself to load variation. A simple method for adapting this variation in load to the generation is to use generation participant factors. That is, for the ith busbar, the real power injection at time instant k+1 is given by

$$\tilde{P}_{Gi}(k+1) = \hat{P}_{Gi}(k) + \alpha_i \sum_{j \in loads} \Delta P_{Lj}(k+1) \tag{25}$$

where: $\tilde{P}_{Gi}(k+1) =$ Predicted active power bus injection at busbar i at time instant k+1; $\hat{P}_{Gi}(k) =$ Estimated active power bus injection at busbar i at time instant k; $\alpha_i =$ generation participation factor $\left(0 \le \alpha_i \le 1; \sum \alpha_i = 1\right)$ and

$$\Delta P_{Lj}(k+1) = \tilde{P}(k+1) - \hat{P}_{Lj}(k) \tag{26}$$

The generation participation factors are calculated using the economic load dispatch and are kept constant between two successive economic dispatch period. The DLP method uses bus power injections (P,Q) as prediction variables, while for filtering step complex bus voltage (e,f) are more appropriate. The prediction vector S and filtering step state vector X are related through $2N$-1 nonlineared equations

$$S = g(X). \tag{27}$$

The change of variables from $S \rightarrow X$ at the end of the prediction step is performed using load flow solution [11]. The covariance matrix for injection prediction error is given by

$$P_0^S(k+1) = Cov\left[S(k+1) - \tilde{S}(k+1)\right] = J_0 P_0(k+1) J_0^T \quad (28)$$

where $J_0 = \left.\frac{\partial g}{\partial X}\right|_{X=X(\tilde{K}+1)}$, and

$$P_0(k+1) = \left(J_0^{-1}\right) P_0^S(k+1) \left(J_0^T\right)^{-1} \quad (29)$$

6 Analysis of anomalous data

One of the objectives of state estimation is to detect, identify and remove or correct the anomalous, data from the incoming information, so as, to maintain the integrity of the data base. One major advantage of DSE is the availability of the predicted state *X* which can be very useful in anomaly detection. Pre-filtering scheme for anomaly detection consists of computing the innovation vector:

$$\vartheta(k) = Z(k) - h\left(\tilde{X}(k)\right). \quad (30)$$

In case, any anomalous data is present, the hypothesis $|\lambda_m(k)| < \lambda_{max} \ \forall$m (where m is the measurement index, $\lambda_m(k) = \vartheta_m(k)/\rho_m(k)$ is the normalised innovation for measurement *m* and $\rho_m(k)$ is the *mth* diagonal element of $\Re(k)$ matrix) will not hold. The choice of threshold λ_{max} is based on simulation. Under normal operating condition all the normalised innovation $\lambda_m(k)$ will be small and within threshold value.

Discrimination between Anomalies and Treatment of Anomalous Data. One important property of the innovation vector is that no "smearing" takes place in innovation process, that is, at time sample *k*, $\lambda(k)$ will have abnormal values for only those measurements which correspond to the erroneous (bad) measurements.

Although, the occurrence of bad data can be detected from the abnormal values of $\lambda_i(k)$, the quality of prediction, effect the value of λ, and a poor dynamic model for prediction may lead to λ value comparable to the magnitude of bad data. In such cases, medium size bad data is likely to go undetected. A more reliable method for detection of bad data is based on "skewness" of the innovation vector. Under normal operating condition, the distribution of $\vartheta(k)$ is symmetrical. In case of sudden load change, though the normalised innovation will have large values corresponding to measurements in the vicinity of the busbar where sudden load/generation change has occurred, the innovation vector $\vartheta(k)$ will maintain its symmetrical property. But in presence of bad data, the distribution of $\vartheta(k)$ becomes asymmetrical and its "asymmetry index" (skewness measure) $\gamma(k) = M_3(k)/\sigma^3(k)$ will be large, where, M3 is the third moment and σ is the standard deviation of the normalised innovation vector at time sample *k*. The presence of bad data is determined from the value of asymmetry index. When bad data is present $|\gamma(k)| > \gamma_{max}$, where $\gamma_{\max}$ is a threshold determined through simulation. Once presence of bad data is detected, measurements for which $\lambda_i(k)$ values are greater than the threshold, are considered bad and are eliminated from the measurement set. In case of

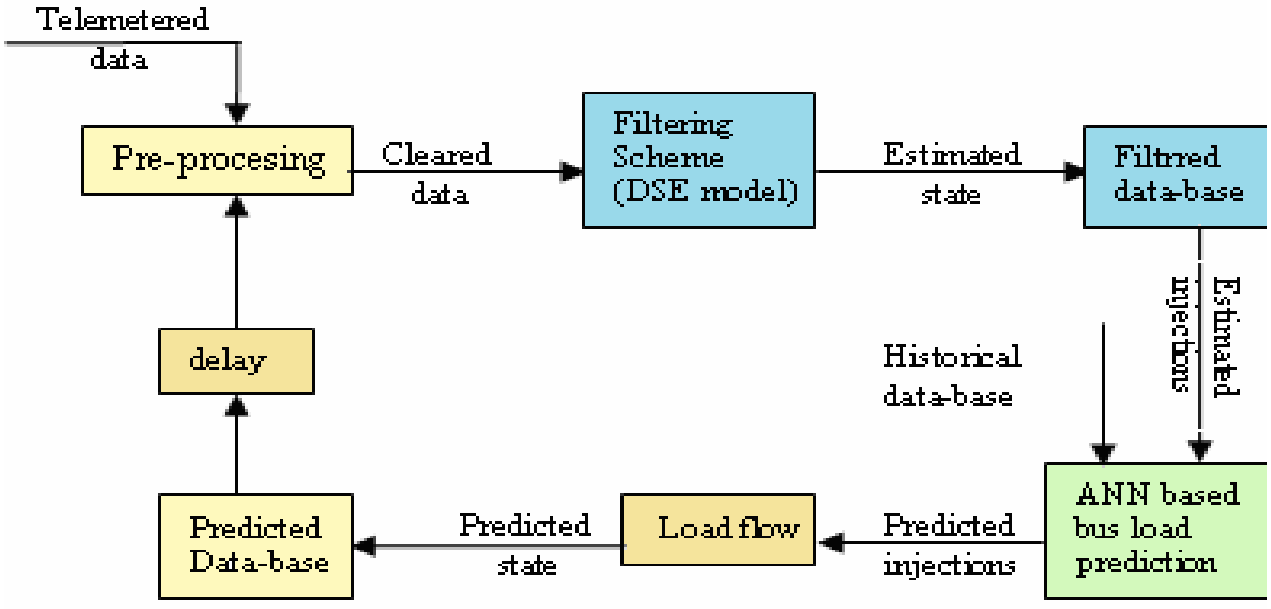

Fig. 3 : ANN based DSE scheme

sudden load/generation change, a number of measurements surrounding the changed injections will exhibit large values for $\lambda_i(k)$.

This is mainly because of the error in prediction of X(k) which has not taken the sudden injection change into account. Such a situation can be taken care of by the proposed filtering scheme as it incorporates the non-linearities of the measurement function.

7 Proposed dse scheme

The overall procedure for the proposed DSE scheme is schematically shown in Figure 3. $\tilde{X}(k+1)$obtained after prediction step provides the complete predictive data base, whereas, $\tilde{X}(k)$ obtained after filtering step provides the complete data base for the current time. The proposed DLP based DSE model faithfully follows the recursive scheme of EKF, except that instead of the state transition based dynamic model it uses a combination of ANN based bus load forecasting and load flow solution to obtain the predicted state $\tilde{X}(k)$. It also incorporates non-linearities of the measurement function the filtering step.

Computational aspect: To reduce the computational burden and speed up the computational process of the proposed DSE scheme, following simplification in the algorithm is introduced:

(i) For computing predicted states from predicted bus loads fast second order load flow [10] is used. This load flow requires computation and factorization of the Jacobian matrix only once in the beginning of the fir iteration. For time step $(k+1)$, the estimated state $\tilde{X}(k)$ is used as the starting value for the load flow computation.

(ii) Since Jacobian matrix is already available from tl load flow at the prediction step, so no computation for Jacobian elements is required for computation $P_0(k+1)$

(iii) Computation of predicted state error covariance $P_0(k+1)$, equation (33), introduces the heaviest comput tional burden in prediction step. Since the Jacobian matrix is very sparse, sparse matrix techniques a: used to obtain $P_0(k+1)$ column by column.

8 Test system simulation

Test Systems: Simulation tests were performed on a number of systems, however due to limitation of space here results for only IEEE 118 bus test system is presented. The simulations were carried out for 30 time samples.

Generation of Bus loads: In the absence of real system data the generation of bus load data was carried out as given below:

(i) 17 sets of hourly load data from the IEEE-24 bus (17 load bus) reliability test system data were used.

(ii) Hourly loads for each set were normalised with respect to the peak load of that data set.

(iii) These 17 load patterns were randomly assigned to various load busbars (one load pattern for each busbar)of the test systems.

(iv) Actual hourly load for each busbar of the test system was obtained by multiplying the normalised hourly load of the assigned load pattern with the peak load of the busbar.

(v) In the absence of any hourly load data for reactive power, hourly reactive power load for each busbar was obtained by keeping the power factor at each busbar constant at the nominal value for that busbar as given in the test system data.

Measurement Generation: The selection of measurements used, was based on random choice and redundancy $(m/n) \cong 2$ was maintained (450 measurements were used). The time evaluation for the system state was simulated by a load curve for each busbar. The power factor for the loads were kept constant. Changes in generation to take care of the load variations were obtained using generation partic-ipation factors. For each time sample, a load flow was carried out using the load and generation values obtained through load curve and generation participation factor, to obtain the true state (X+) and true value of measurements (Z+). The actual measurements (Z) were obtained by adding a normally distributed noise (error) with standard deviation (σ) of 2% for power injections and flow measurements and σ = 1% for voltage measurements, to Z+.

Initialisation: To start the DSE process the initial values for the various parameters were specified as follows: Initial state $\hat{x}(0)$ was obtained using static state estimation from the measurements at k=0 and its error covariance was specified as P+(0) =diag(10-6).

Bad Data: Bad data conditions were simulated by introducing error in the measurement set at 5th time sample in real power at bus 80, at 11th and 16th time sample in P_{80}, P_{92} & Q_{115} at 21st time sample in P_{92}.

Sudden Load Changes: Sudden load change condition was simulated by changing the load at bus number 54 at 11th and 16th time samples.

Performance evaluation: The performance of the algorithms were assessed using the following indices.

Prediction step: The index used for assessing the performance of the prediction model is

$$\zeta_x(k) = \frac{1}{n}\sum_{1}^{n}\left|\tilde{X}_i(k) - X_i^+(k)\right| \tag{31}$$

Filtering step: For assessing the effectiveness of filter, the index used is

$$J_m(k) = \frac{\sum_{i=1}^{m}\left|\hat{Z}_i(k) - Z_i^+(k)\right|}{\sum_{i=1}^{m}\left|\hat{Z}_i(k) - Z_i^+(k)\right|} \tag{32}$$

9 Conclusion

This paper investigates the adaptive data reduction in sensor array processing for the colored noise case. The design criterion for the beamspace transformation is preservation of the Cramér-Rao bounds for the parameter estimates. A design procedure is given that produces a transformation that closely approximates the ideal one. The benefits are shown via computer simulations that focus on the problem of out-of-band sources. These can be viewed as interfering sources in beamspace, causing a loss in performance relative elementspace estimation. The results indicate that significant improvements can be gained in terms of meansquare-error performance if the outlined approach is followed. The proposed ANN based DSE provides better state estimates than the DSE with conventional prediction model.

References

[1] Abur, A., Kim, H., Celik, M. K., "Identifying the Unknown Circuit Breaker Statuses in Power Networks", *IEEE Trans, on Power Systems,* Vol. 10, No. 4, Nov. 1995, pp.2029-2037.

[2] Andersson, S., Optimal dimension reduction for sensor array signal processing, Signal Process. 30 (2), January 1993, pp. 245-256.

[3] Andersson, S., Nehorai, A., "Optimal dimension reduction for array processing-generalized", IEEE Trans. Signal Process. 43(8), August 1995, pp.2025-2027.

[4] Alsac, O., Vempati, N., Stott, B., "Generalized State Estimation," Proceedings of the PICA Conference, Columbus, OH, May 11-16, 1997, pp.90-96.

[5] Bulucea, C.A., Popescu, M.C., Manolea Gh., Patrascu, A., "Interest and Difficulty in Continuous Analysis of Water Quality", Proceedings of the 4th, WSEAS International Conference on Energy & Environment, Feb.22-23, 2009, pp.220-225.

[6] Bulucea, C.A., Popescu, M.C., Manolea Gh., "Real Time Medical Telemonitoring of Sustainable Health Care Measuring Devices", Proceedings of the 8th WSEAS Int. Conf. on Artificial Intelligence, Knowledge Engineering & Data Bases, Feb.22-23, 2009, pp.202-207.

[7] Boteanu, N., Popescu, M.C., Ravigan. F., "Optimal Control by Energetic Criterion of Driving Systems", Proceedings of the 10th WSEAS International Conference on Mathematical and Computational Methods in Science and Engineering, Bucharest, Nov.7-9, 2008, pp.45-51.

[8] Drighiciu, M., Manolea, Gh., Petrisor, A, Popescu, M.C., "On Hybrid Systems Modeling with Petri Nets", Proceedings of the 7th WSEAS International Conference on System Science and Simulation in Engineering, Venice, Nov. 21-23, 2008, pp.73-78.

[9] Eriksson, J.,"Data reduction in sensor array processing using parameterized signals observed in colored noise", Proceedings of the 29th Asilomar Conference on Signals System and Computing, Monterey, CA, 1996.

[10] Forster, P., Vezzosi, G., "Application of spheroidal sequences to array processing". Proceedings of the ICASSP 87, Dallas, TX, pp.2268-2271, 1987.

[11] Monticelli, A. "Modeling Circuit Breakers in Weighted Least Squares State estimation," *IEEE Transactions on Power Systems, Vol.* 8, No. 3, Aug. 1993, pp.1143-1149.

[12] Petrişor, A., Bizdoacă, N.G., Drighiciu, M., Popescu, M.C., "Control Strategy of a 3-DOF Walking Robot", The International Conference on „Computer as a Tool", Warsaw, Polonia, 9-12, sept. 2007, pp.2337-2342.

[13] Popescu, M.C., *Telecomunicații,* Editura Universitaria Craiova, pp.420-428, 2008.

[14] Popescu, M.C., *Modelarea şi simularea proceselor*, Editura Universitaria Craiova, 2008, pp. 544-560.

[15] Popescu, M.C., *Comanda şi optimizarea proceselor*, Editura Universitaria Craiova, 2007, pp. 265-276.

[16] Popescu, M.C., *Estimarea şi identificarea proceselor*, Editura Sitech, pp.228-242, 2006.

[17] Popescu, M.C., *Utilisation des ordinateurs,* Editura Universitaria, pp.258-286, 2004.

[18] Popescu, M.C., Petrişor, A., Drighiciu, M.A, "Algorithm for dynamic state estimation of power systems", Proceedings of 4th international conference on robotics, Bulletin of the *Transilvania* Univ. of Braşov, Vol. 15(50), Series A, Braşov, 2008, pp.295-302.

[19] Popescu, M.C., Petrişor, A., "2D Optimal Control Algorithms Implementation", WSEAS Transactions on System and Control, Issue 1, Volume 1,Veneția, 2006, pg. 94-99.

[20] Popescu, M.C., "Tracking Performance of a Quantized Adaptive Filter Equipped with the Sign Algorithm",The 7th International Carpathian Control Congress, Roznov pod Radhostem, 2006, pp.445-448.

[21] Popescu, M.C., "Approche systemique en genie automatique", Analele Universității din Craiova, Horticultură, Biologie, Tehnologia prelucrării produselor agricole, Ingineria mediului, Vol. XII (XLVIII), Craiova, 26 oct. 2007, pp.415-420.

[22] Slutsker, I. and Mokhtari, S., "Comprehensive Estimation in Power Systems: State, Topology, and Parameter Estimation," Proceedings of the American Power Conference, Vol. 57-1, Chicago, Illinois, April 1995, pp.149-155.

[23] Slutsker, I., Mirheydar, M., Cahagan, H.T. Neisius, G. Schellstede, "Scan Rate Estimation in Energy Management Systems," Proceedings of the American Power Conference, Chicago, Illinois, April 1996, pp. 924-929.