A STATE SPACE MODEL FOR $H^\infty$ TYPE ARRAY SIGNAL PROCESSING

T. Ratnarajah and A. Manikas

Department of Electrical and Electronic Engineering
Imperial College of Science, Technology and Medicine
London SW7 2BT

ABSTRACT

The idea of applying $H^\infty$ estimation techniques to the “array uncertainties problem” is motivated by the fact that $H^\infty$ estimation is robust to model uncertainties and lack of statistical information with respect to noise. In this paper, a new state space model for the received signal of a general array of sensors is developed which, in contrast to existing models, is capable of handling the simultaneous presence of different type of uncertainties (e.g., gain, phase, locations, mutual coupling, etc, uncertainties). Based on this state-space model, formulated in an $H^\infty$ framework, two new robust array signal processing techniques have been proposed which mitigate the degrading effects of array uncertainties.

1. INTRODUCTION

Direction of arrival (DOA) estimation by Signal Subspace type methods requires knowledge of the array covariance matrix and an exact characterisation of the array in terms of geometry (sensor location), sensor gain and phase, mutual coupling between the array elements etc. In practice, however, neither of these quantities is known precisely. Depending on the degree to which they deviate from their nominal values, serious performance degradation may result. Furthermore, all the direction finding based signal-copy algorithms estimate the signals by forming a weighted linear combination of the array outputs. In general, the weight vector involves knowledge of the array response and the directions of arrival of some or all of the signals. Any errors in the array model affect not only the weight vector directly, but also the accuracy of the DOA estimates, and thus can seriously reduce the overall performance [1]. There is, therefore, considerable practical interest in the development of array processing techniques which are able to operate in the presence of array uncertainties and the $H^\infty$ formulation proposed in this paper is an attempt to address this problem.

In this paper, to apply $H^\infty$ estimation techniques to array signal processing, a new state-space model for the received signal of a general array of sensors is developed. The proposed model can cope with directional gain and phase uncertainties (which need not be identical from sensor to sensor), sensor location errors, mutual coupling and noise effects, etc. These uncertainties/errors are defined in the Section-2 of the paper. In Section-3, some of the basic orthogonality and covariance conditions which are induced by the state-space assumptions are identified. The global linear relation between the observations, the initial state vector, the process noise and the measurement noise are then highlighted. In addition, the so-called observability and impulse response matrices associated with our model are introduced and it is shown that the columns of observability matrix is nothing more than the manifold vectors. These concepts enable us to show the agreement between the proposed state-space model and the conventional array data model. In Section-4, the concepts of previous section are then placed in the framework of an $H^\infty$ approach while in Section 5, two representative array processing examples are presented. Finally, in Section 6, the paper is concluded.

2. MODELLING THE ARRAY SIGNAL

Consider an uncalibrated (or partially calibrated) array of $N$ sensors operating in the presence of $M$ narrowband signals. The received signal vector $\mathbf{x}(t)$ can be modelled as

$$\mathbf{x}(t) = (I + \mathbf{C}) \left( (\mathbf{G} + \mathbf{G}) \odot (\mathbf{A} + \mathbf{A}) \right) \mathbf{m}(t) + \mathbf{n}(t) \quad (1)$$

where

- $\odot$ denotes Hadamard product
- $\mathbf{m}(t)$ is the message vector-signal
- $\mathbf{n}(t)$ represents the additive white Gaussian Noise
- $\mathbf{C}$ is the mutual coupling uncertainty matrix
- $\mathbf{G}$ is the array gain and phase error matrix
- $\mathbf{A}$ is the sensor position error matrix

In Equation-(1) $\mathbf{G}$ and $\mathbf{A}$ are $N \times M$ complex matrices with their $r$th column defined as $\gamma(\theta, \phi) \odot \psi(\theta, \phi) \otimes e^{j \psi(\theta, \phi)}$ and $e^{-j \psi(\theta, \phi)}$, respectively where

$$\mathbf{k} = \frac{\pi}{L} \left[ \cos(\theta_i) \cos(\phi_i), \sin(\theta_i) \cos(\phi_i), \sin(\phi_i) \right]^T,$$

$r$ is the array location matrix, and the vectors $\gamma(\theta, \phi)$ and $\psi(\theta, \phi)$ denote the gain and phase response of the array elements for a signal incident from azimuth $\theta_i$ and elevation $\phi_i$. In this investigation Equation-(1) will be used as the starting point to arrive at an alternative modelling and then to propose a new approach for improving the performance of array signal processing algorithms operating in the presence of array uncertainties. The assumption is simply that both the array uncertainties and the noise signals can be considered as bounded energy signals. This implies that the effects due to the presence of array uncertainties can be reduced by using a minimax optimal estimation algorithm, or more specifically, an $H^\infty$ optimal approach.
3. PROPOSED STATE-SPACE MODEL

The proposed state space model is based on the array signal vector $x(t)$ given by Equation-(1). However, by reorganising this equation, the direction gain and phase errors may be grouped together and represented, for the $j^{th}$ sensor, by a scalar $v_j$. Furthermore the array location errors, mutual coupling and noise effects may be grouped together in an $M \times 1$ vector $u_j$. In state-space terminology the scalar $v_j$ is the measurement noise while the vector $u_j$ is the process noise associated with the $j^{th}$ sensor. In this case the elements $x_j, j = 1, \ldots, N,$ of the received signal $x(t)$ at a particular time $t$ obey the following state space model:

$$
\begin{align*}
&x_{j+1} = \text{diag} [a_{j+1} \odot g_j^*] \cdot x_j + B_j \cdot u_j, \quad j \in [0, N-1] \\
&z_j = g_j^T \cdot x_j + v_j, \quad j \in [1, N], \quad z_0 = m(t)
\end{align*}
$$

where

$$
\begin{align*}
&z_j \in \mathbb{C}^{M \times 1} \text{denotes the state of the} \ j^{th}-\text{sensor} \\
&a_j = \left[j^{th} \text{row of } A \right]^T \\
&g_j = \left[j^{th} \text{row of } G \right]^T
\end{align*}
$$

In Equation-(2) the matrix $(\text{diag} [a_{j+1} \odot g_j^*]) \in \mathbb{C}^{M \times M}$ is known as the state transition matrix, while $\mathbf{B} = \mathbf{I} \in \mathbb{R}^{M \times M}$, where $\rho$ is a positive constant which imposes a bound on the process noise $u_j$. It is assumed that the variables $\{u_j\}, \{v_j\}$, and $\{z_0\}$ obey the following relationship

$$
E \left[ \begin{bmatrix} z_0 \\ u_j \\ v_j \end{bmatrix} \right] = \begin{bmatrix} R_{mm} & 0 & 0 \\ 0 & Q_i & P_i \\ 0 & P_i & \alpha_i \delta_{ij} \end{bmatrix}
$$

where $\delta_{ij}$ denotes the Kronecker delta function. Note that if there are no uncertainties in the array system then $u_j = 0, \forall j$, and $v_j$ only represents the additive noise. Various important orthogonalties and covariance properties of the model are illustrated in the following two lemmas. These lemmas enable us to show the agreement between the proposed state-space model and the conventional array data model.

Lemma 1 (Basic Orthogonality Properties)

In the proposed model we can assert that

(i) $E\{u_i \cdot g_i^H\} = 0$ and $E\{u_i \cdot z_j\} = 0^T, \forall \ i \geq j$

(ii) $E\{u_i \cdot z_j^*\} = \begin{cases} P_i, & i = j \\ 0, & i > j \end{cases}$

(iii) $E\{v_i \cdot z_j^*\} = \begin{cases} \alpha_i, & i = j \\ 0, & i > j \end{cases}$

(iv) $E\{u_i \cdot u_j^H\} = 0$ if $0 < \text{diag} [a_{j+1} \odot g_j^*] = \text{nonsingular } \forall \ i$

By introducing the so-called observability ($\mathcal{O} = G \odot A$) and (lower triangular) impulse response ($\Gamma$) matrices associated with the proposed model, the variables $\{z_0, u, v\}$ can be related with the following expression

$$
z = \mathcal{O} \cdot z_0 + \Gamma \cdot u + v = \mathcal{O} \cdot \Gamma \left[ \begin{bmatrix} z_0 \\ u \\ v \end{bmatrix} \right] + v
$$

where

$$
\begin{align*}
&\mathcal{O} = \begin{bmatrix} u_1 \cdots u_N \end{bmatrix}^T, \quad v = [v_1, \ldots, v_N]^T \\
&\Gamma = \begin{bmatrix} \Gamma_{11} \cdots \Gamma_{1N} \\ \Gamma_{21} \cdots \Gamma_{2N} \\ \vdots \\ \Gamma_{N1} \cdots \Gamma_{NN} \end{bmatrix}
\end{align*}
$$

with $\Gamma_{ij} = g_i^T \cdot \Phi(i, j + 1) \cdot B_j$. The following lemma provides the framework for finding a global expression for the covariance matrix $R_{zz}$ based on the linear relation given by Equation-(4).

Lemma 2 (Data Covariance Matrix Modelling)

The covariances of the state variables can be computed as follows

$$
E\{z_i \cdot z_j^*\} = \begin{cases} \Phi(i, j) \cdot \Pi_j, & i > j \\ \Pi_i \cdot \Phi^H(j, i), & i < j \\ \Pi_i, \quad i = j \end{cases}
$$

where

$$
\Pi_{i+1} = \text{diag} \left[ a_{i+1} \odot g_i^* \right] \cdot \Pi_i \cdot \text{diag} \left[ a_{i+1} \odot g_i^* \right] + \mathbf{B}_i \cdot \mathbf{Q}_i \cdot \mathbf{B}_i^H
$$

and

$$
\Phi(i, j) = \text{diag} \left[ a_i \odot g_{i-1}^* \right] \cdots \text{diag} \left[ a_{j+1} \odot g_{j-1}^* \right] \cdot \Phi(i, i) = 1.
$$

Furthermore the covariances of the element of received signal $\{x_j\}$ are

$$
E\{x_i \cdot x_j^*\} = \begin{cases} \\ g_i^T \cdot \Phi(i, j + 1) \cdot D_j, & i > j \\ D_i \cdot \Phi^H(j, i + 1) \cdot g_i^*, & i < j \\ (a_i^2 + \alpha_i^2) \cdot \Pi_i \cdot g_i^*, & i = j \end{cases}
$$

where

$$
D_i = \text{diag} \left[ a_{i+1} \odot g_i^* \right] \cdot \Pi_i \cdot g_i^* + B_i \cdot P_i
$$

Finally the array covariance matrix $R_{zz}$ is given by

$$
R_{zz} = \mathcal{O} \cdot \Pi_0 \cdot \mathcal{O}^H + \Gamma \cdot \mathbf{Q} \cdot \Gamma^H + \mathbf{R}_{zz}
$$

where $\mathbf{Q} = \text{diag}(Q_0, \ldots, Q_{N-1}), \mathbf{P} = \text{diag}(P_0, \ldots, P_{N-1})$.

It is important to point out that Equation-(7) is the discrete space (or time) Lyapunov recursion. Furthermore, if $P_i = 0, i = 0, \ldots, N - 1, \text{then}$

$$
R_{zz} = \mathcal{O} \cdot \Pi_0 \cdot \mathcal{O}^H + \mathbf{Q} \cdot \Gamma^H + \mathbf{R}_{zz}
$$

In addition $\Pi_0 = 0, i = 0, \ldots, N - 1$ (which means that $Q = 0$), then

$$
R_{zz} = \mathcal{O} \cdot \Pi_0 \cdot \mathcal{O}^H + \mathbf{R}_{zz}
$$

In this case $\Pi_0 = R_{mm}$ and the columns of observability matrix are the manifold vectors ($\mathcal{O} = G \odot A$), therefore the output covariance matrix can be written as

$$
R_{zz} = (G \odot A) \cdot R_{mm} \cdot (G \odot A)^H + \mathbf{R}_{zz}
$$

It is clear from the Equation-(12) in conjunction Equation-(13) that the terms $\Gamma \cdot \mathbf{Q} \cdot \Gamma^H$ and $\mathbf{R}_{zz}$ include the array uncertainties.
4. AN $H^\infty$ APPROACH TO THE PROPOSED STATE-SPACE MODEL

The $H^\infty$ estimation methods can be seen as a powerful and robust solution to handle array uncertainties and noise effects with limited statistical information. The idea is to come up with estimators that minimize (or in the suboptimal case, bound) the maximum energy gain from the disturbances to the estimation errors. This will guarantee that if the disturbances are small (in energy) then the estimation errors will be as small as possible (in energy), no matter what the disturbances are. In other words the maximum energy gain is minimized over all possible disturbances. The robustness of the $H^\infty$ estimators arises from this fact.

By using the above state space model the objective is to estimate, using the array data $g(t)$, some linear combination of the states, i.e.,

$$\hat{z}_i = L_i \cdot z_i,$$

where $L_i$ is known. Let $F(.)$ be the functional which represents this estimation process, i.e.,

$$\hat{z}_i = F(x_1, x_2, \ldots, x_i),$$

which indicates the last element of the vector $\hat{z}_i$ (i.e., the $i^{th}$ element) can be estimated as a function of the received signals at time $t$ from sensor $1$ (i.e., $x_1$) up to and including the sensor $i$ (i.e., $x_i$). Let $\Pi_0$ be a given positive-definite matrix and choose any initial estimate for $\hat{z}_0$, which we shall denote by $\hat{z}_0$. Define the weighted initial state error $\hat{\tilde{z}}_0$ as well as the estimation error $\tilde{e}_i$ as follows

$$\begin{align*}
\hat{\tilde{z}}_0 &\equiv \Pi_0^{1/2}(z_0 - \hat{z}_0), \\
\tilde{e}_i &\equiv \hat{z}_i - L_i \cdot \hat{z}_i.
\end{align*}$$

For every sensor $i$, define the ratio:

$$r(i) = \frac{\sum_{j=1}^{i} \tilde{e}_j^T \cdot \tilde{e}_j}{\|z_0\|^2 + \sum_{j=1}^{i} u_j^T \cdot u_j + \sum_{j=1}^{i} v_j^T \cdot v_j}$$

so that $r(i)$ is bounded, for every $\hat{\tilde{z}}_0$, $u(.)$ and $v(.)$, by a given positive constant $\epsilon^2$, say,

$$r(i) < \epsilon^2 \quad \text{for} \quad 1 \leq i \leq N$$

The estimator $F(.)$ that satisfies the Equation-(19) is called $H^\infty$ a posteriori estimator.

Thus assuming that there are $N$ observations available, we collect the estimation error $e_i$ into a column vector, $g(t) \in \mathbb{C}^{(N+1)M \times 1}$, i.e.,

$$g(t) = [e_0^T, e_1^T, \ldots, e_N^T]^T$$

Furthermore the noise sequences and the initial state estimation-error form another column vector, $d(t) \in \mathbb{C}^{((N+1)(M+1)N \times 1)}$, i.e.,

$$d(t) = \left[\left(\Pi_0^{1/2} \tilde{z}_0\right)^T, u_0^T, v_1^T, u_1^T, v_2^T, \ldots, u_{N-1}^T, v_N^T\right]^T$$

It is clear that $d(t)$ contains the disturbance signals (these are signals that we have no control over) while $g(t)$ contains the resulting estimation errors (these are the errors that result from the solution). Now if the estimator $F(.)$ exist, it should induce a mapping, say $T_N(F)$, from $d(t)$ to $g(t)$. The condition $r(N) < \epsilon^2$, when satisfied, will therefore guarantee that the 2-induced norm of $T_N(F)$ is bounded by $\epsilon$. In this case, we say that the level of robustness is $\epsilon$.

The $H^\infty$ a posteriori estimator $F(.)$ can be found by using Theorem-1 from [3], it can be shown that the signal $\hat{z}_j$ can be estimated as follows

$$\hat{z}_j = L_j \cdot \hat{\tilde{z}}_j,$$

where $\hat{\tilde{z}}_j$ is recursively computed using the following expression

$$\hat{\tilde{z}}_{j+1} = \text{diag} \left[\begin{array}{c}
\tilde{e}_{j+1}^T \tilde{e}_j \end{array}\right] + K_{j+1} \begin{pmatrix}
1 - \tilde{e}_{j+1}^T \tilde{e}_{j+1} & \text{diag} \left[\begin{array}{c}
\tilde{e}_{j+1}^T \\
\end{array}\right]
\end{pmatrix} \hat{\tilde{z}}_{j+1},$$

with

$$\begin{pmatrix}
\tilde{e}_0 \\
\end{pmatrix} = (G \otimes A)^T g(t) = \text{initial guess}$$

$$K_{j+1} = \begin{pmatrix}
1 + \text{diag} \left[\begin{array}{c}
\tilde{e}_{j+1}^T \\
\end{array}\right]
\end{pmatrix}^{-1} \begin{pmatrix}
P_{j} \text{diag} \left[\begin{array}{c}
\tilde{e}_{j+1}^T \\
\end{array}\right] + \text{B}_j \text{B}_j^T
\end{pmatrix} \text{R}_{j}^{-1} \text{P}_j \text{diag} \left[\begin{array}{c}
\tilde{e}_{j+1}^T \\
\end{array}\right]^T$$

$$- \text{diag} \left[\begin{array}{c}
\tilde{e}_{j+1}^T \\
\end{array}\right] \text{P}_j \text{diag} \left[\begin{array}{c}
\tilde{e}_{j+1}^T \\
\end{array}\right]^T \text{P}_j \text{diag} \left[\begin{array}{c}
\tilde{e}_{j+1}^T \\
\end{array}\right]^T$$

$$P_0 = \Pi_0$$

Note that for a given $\epsilon > 0$ and if the transition matrices are nonsingular (in our case the matrices $\text{diag} \left[\begin{array}{c}
\tilde{e}_{j+1}^T \\
\end{array}\right]$ are always non-singular) then the above solution exists if and only if

$$\text{P}_j^{-1} + \tilde{z}_j \text{R}_j^{-1} \text{P}_j > 0, \quad j = 0, \ldots, i.$$ 

We shall assume, without loss of generality, that $\Pi_0$ has the special form $\Pi_0 = \mu \mathbf{I}$ (i.e. $\Pi_0$), where $\mu$ is a positive constant.

The above procedure can be used in conjunction with MuSIC algorithm to provide an improved set of direction estimates or in conjunction with a signal-copy beamformer, to provide a good copy of the desired signal. In the first case (MuSIC) the matrix $L_i$ is modelled to represent the gain/phase vector of the i-th sensor (i.e., $L_i = g_i^T$) and Equation-(23) provides the uncorrupted received signal-vector $z(t)$ while the initial directions are provided by MuSIC algorithm operating with the nominal values of an uncalibrated array. In the second case (beamformer) the matrix $L_i$ is the diagonal matrix $\text{diag} \left[\begin{array}{c}
\tilde{e}_1^T, \ldots, \tilde{e}_M^T
\end{array}\right]$ and the copy of the desired message signal is taken to be average of the linear combinations of the states, i.e.,

$$\tilde{m}(t) = \frac{1}{N} \sum_{j=1}^{N} \hat{z}_j.$$ 

5. REPRESENTATIVE EXAMPLES

Consider the planar circular array of six isotropic antennas, with antenna-locations described by the following matrix (in meters):

$$r = \begin{bmatrix}
15.22 & 24.79 & 9.57 & -9.57 & -23.1 & -23.1 \\
19.84 & -3.27 & -23.1 & -23.1 & -9.57 & 9.57 \\
0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}^T$$

Copyright 1997 IEEE
We assume that the array operates in the presence of mutual coupling effects (uncertainties) with each sensor significantly coupled with its nearest neighbours, while the coupling with other sensors can be ignored. Furthermore, consider that there are also gain, phase and location uncertainties (in meters) which can be described as follows:

\[
\tilde{\gamma} = \begin{bmatrix} 0.3935, 0.4674, 0.3558, 0.5704, 0.0331 \end{bmatrix}^T,
\]

\[
\tilde{\psi} = \begin{bmatrix} 0, 11.5623^\circ, 11.4305^\circ, 7.5401^\circ, 10.6857^\circ, 6.7323^\circ \end{bmatrix}^T,
\]

\[
\tilde{\mathbf{r}} = \begin{bmatrix} 0.0668, 0.6872, 0.9305, 0.5266, 0.6543 \end{bmatrix}^T,
\]

\[
\begin{bmatrix} 0.4179, 0.5894, 0.8457, 0.0918, 0.4159 \end{bmatrix}^T\]

We assume that the above array configuration operates in the presence of two equipower uncorrelated sources at \((110^\circ, 0^\circ)\), and \((130^\circ, 0^\circ)\) and the signals are sinusoids of amplitude \(\sqrt{2}\) and normalised frequency of 0.2 (signal frequency is 15.04 MHz). The background noise power is 20dB below the source power and the received signal is formed by 100 snapshots. The nominal gain and phase of the antennas are assumed to be 1 and 0 respectively. The mitigation of the degrading effects of array uncertainties on the performance of the MuSIC is illustrated in Figure-(1), with dashed line representing the MuSIC spectrum applied to the original data collected by uncalibrated array while the solid-line denoting the MuSIC spectrum applied to the "filtered" data after the \(H^\infty\) approach has been used. In this example, for the implementation of the proposed approach the matrix \(\mathbf{L}_i\) becomes a \(M \times 1\) vector representing the gain/phase response \(q_i\), of the \(i^{th}\) sensor. It is clear that the array uncertainty effects are reduced after the application of the proposed algorithm and approximately 15dB improvement is achieved at peak of the MuSIC spectrum.

In the second example we illustrate the robust performance of the signal copy algorithm and assume that the source at \((130^\circ, 0^\circ)\) provides the signal of interest and is fully correlated with the interference at \((110^\circ, 0^\circ)\). In this case, the matrix \(\mathbf{L}_i\) is replaced by \(\text{diag}(e^{j2\pi k_1}, \ldots, e^{j2\pi k_M})\), where nominal locations and the estimated DOAs such as \((128.5^\circ, 0^\circ)\) and \((109.5^\circ, 0^\circ)\) are used. Figure (2) illustrates the result (solid line) of the proposed approach for the above uncalibrated array. The dotted line in this figure illustrates the true signal while on the same figure the results using Wiener-Hopf processor has been also plotted (dashed line) for comparison. It is apparent from this figure that the proposed algorithm is robust to calibration uncertainties and can also handle coherent sources whereas the Wiener-Hopf beamformer totally fails.

![Figure 1. Dashed Line - MuSIC Spectrum BEFORE and Solid line - MuSIC Spectrum AFTER Application of the Proposed Approach](image1.png)

![Figure 2. Solid line - Estimated Signal Amplitude Using the Proposed Algorithm, Dot-Dot line - True Signal Amplitude, and Dashed line - Estimated Signal Amplitude Using Wiener-Hopf.](image2.png)

6. CONCLUSION

In this paper a new array-signal state space model has been developed for a general array geometry which can handle the simultaneous presence of different type of uncertainties. By introducing two lemmas we have shown the agreement between the proposed state-space model and the conventional array data model approach. The proposed approach has been used in conjunction with the MuSIC algorithm and then in conjunction with a beamformer in order to mitigate the simultaneous degrading effects of finite sampling and imprecise modelling of an antenna array and spatial noise statistics.

REFERENCES


