Bayesian Parametric Approach for Multichannel Adaptive Signal Detection

Pu Wang, Hongbin Li, and Braham Himed

Abstract—This paper considers the problem of space-time adaptive processing (STAP) in non-homogeneous environments, where the disturbance covariance matrices of the training and test signals are assumed random and different with each other. A Bayesian detection statistic is proposed by incorporating the randomness of the disturbance covariance matrices, utilizing a priori knowledge, and exploring the inherent Block-Toeplitz structure of the spatial-temporal covariance matrix. Specifically, the Block-Toeplitz structure of the covariance matrix allows us to model the training signals as a multichannel auto-regressive (AR) process and hence, develop the Bayesian parametric adaptive matched filter (B-PAMF) to mitigate the training requirement and alleviate the computational complexity. Simulation using both simulated multichannel AR data and the challenging KASSPER data validates the effectiveness of the B-PAMF in non-homogeneous environments.

Index Terms—Parametric adaptive matched filter, Bayesian detection, space-time adaptive signal processing, non-homogeneous environments.

I. INTRODUCTION

Traditional space-time adaptive processing (STAP) usually deals with homogeneous environments, where the test signal is assumed to share the same covariance matrix with the training signals [1]–[4].

To account non-homogeneous environments, a number of models have been proposed. One is the partially homogenous environment, which assumes the training signals share the covariance matrix with the test signal up to an unknown scaling factor [5], [6]. This model can be considered as a special case of the generalized eigenrelation (GER) [7]. Another one is the compound-Gaussian model, which assumes the training signals are a product of a texture (scaler) and a Gaussian vector. The texture is used to simulate power differences among the signals from range bins [8], [9]. More recently, a new class of non-homogeneous environments for adaptive signal detection emerges. This non-homogeneous environment is characterized by treating disturbance covariance matrices of both the test signal and training signals as random matrices and ensuring that they are different in probability one [10]–[14]. Following the Bayesian non-homogeneous model in [10], an adaptive matched filter has been derived by replacing the exact covariance matrix of the test signal by its maximum a posteriori (MAP) estimate in the matched filter [10]. It has been shown that, by accounting the heterogeneity knowledge, the Bayesian adaptive matched filter (B-AMF) outperforms the standard AMF in the non-homogeneous environment. For applications of this Bayesian non-homogeneous model that employ space-time adaptive processing, the training requirement of the sample covariance matrix (SCM)-based B-AMF cannot be met due to the scarcity of the training signal. For example, with $J = 11$ spatial channels and $N = 32$ coherent pulses of the KASSPER dataset and assuming an instantaneous RF bandwidth of 500 KHz, $K = JN = 352$ training signals calls for the training range over a 200-km range, which is not practical [15], [16].

In this paper, while preserving the Bayesian non-homogeneous environment, we further explore the inherent Block-Toeplitz structure of the spatial-temporal covariance matrix which allows the block LDU decomposition [17], and hence enables the disturbances to be modeled as a multi-channel auto-regressive (AR) process [17]–[23]. The resulting Bayesian parametric adaptive matched filter (B-PAMF) reduces the joint spatial-temporal whitening of the SCM-based B-AMF to successive spatial and temporal whitening. As a result, it facilitates the STAP in the non-homogeneous environments, and reduces the excessive training requirement of the B-AMF. Moreover, the B-PAMF is able to incorporate heterogeneities of the signals and utilize available a priori knowledge to the decision statistic. The effectiveness of the B-PAMF is verified by using the simulated multichannel AR data and the high fidelity KASSPER data [15].

II. SIGNAL MODEL

Assume $J$ spatial channels, $N$ temporal pulses, and $K$ training range cells. The problem of interest is to detect a $JN \times 1$ multichannel signal $s$ with unknown amplitude $\alpha$ in the presence of spatially and temporally correlated disturbance $d_0$:

$$H_0 : x_0(n) = d_0(n), n = 0, 1, \ldots, N - 1,$$

$$H_1 : x_0(n) = \alpha s(n) + d_0(n), n = 0, 1, \ldots, N - 1,$$

where $d_0 = [d_0^T(0), d_0^T(1), \ldots, d_0^T(N - 1)]^T$, and $s$ and $x_0$ are similarly defined. In this paper, the signal model makes the following assumptions:

- **AS1 (Multichannel AR Process):** The disturbances in both test and training signals are modeled as a multi-channel AR process [17]–[19]:

$$d_k(n) = -\sum_{i=1}^{P} A^H(i) d_k(n-i) + \varepsilon_k(n), k = 0, \ldots, K,$$

where $A(i)$ is the $JN \times JN$ autoregressive (AR) coefficient matrix of order $P$. For convenience, we take $A(i)$ to be a Toeplitz matrix, which ensures that the disturbances are temporally correlated.

P. Wang and H. Li are with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030, USA (e-mail: {pwang4, hli}@stevens.edu).

Braham Himed is with AFRL/RYRT, 2241 Avionics Circle, Bldg 620, Dayton, OH45433 (e-mail: Braham.Himed@wpafb.af.mil).

This work was supported by the Air Force Research Laboratory (AFRL) under Contract FA8750-05-2-0001 and the Air Force Office of Scientific Research (AFOSR) under Grant FA9550-09-1-0310.
**Title:** Bayesian Parametric Approach for Multichannel Adaptive Signal Detection

**Abstract:**
This paper considers the problem of space-time adaptive processing (STAP) in non-homogeneous environments where the disturbance covariance matrices of the training and test signals are assumed random and different with each other. A Bayesian detection statistic is proposed by incorporating the randomness of the disturbance covariance matrices, utilizing a priori knowledge, and exploring the inherent Block-Toeplitz structure of the spatial-temporal covariance matrix. Specifically, the Block-Toeplitz structure of the covariance matrix allows us to model the training signals as a multichannel auto-regressive (AR) process and hence, develop the Bayesian parametric adaptive matched filter (B-PAMF) to mitigate the training requirement and alleviate the computational complexity. Simulation using both simulated multichannel AR data and the challenging KASSPER data validates the effectiveness of the B-PAMF in non-homogeneous environments.
where \( A^H = [A^H(1), A^H(2), \ldots, A^H(P)] \) denote the unknown multichannel AR coefficient matrix, and \( \varepsilon_k(n) \) denote the \( J \times 1 \) temporally white but spatially colored noise vectors.

- **AS2 (Random Disturbance Covariance Matrix of Test Signals):** The noise vector \( \varepsilon_k(n) \) is distributed as \( \varepsilon_k(n) \sim \mathbb{C}\mathbb{N}(0, Q) \), and the spatial covariance matrix \( Q \) follows an inverse complex Wishart distribution with degrees of freedom \( \mu \) and mean \( Q_0 \):

\[
p(Q) = \frac{\Gamma(J, \mu)}{\Gamma(J, \nu)} e^{-\frac{(\mu-J)\text{tr}(Q^{-1}Q_0)}}.
\]

where \( \Gamma(J, \mu) = \frac{\Gamma(\mu-J, J/2)}{\Gamma(J/2)} \) with \( \Gamma \) given by the Gamma function [24].

- **AS3 (Random Disturbance Covariance Matrix of Test Signal):** The noise vector in the test signal \( \varepsilon_0(n) \sim \mathbb{C}\mathbb{N}(0, Q_0) \), and \( Q_0 \), given \( Q \), has a complex Wishart distribution with degrees of freedom \( \nu \) and mean \( Q \):

\[
p(Q_0|Q) = \frac{\nu^{J/2}\pi^{J-1}Q_0^{
u-J}}{\Gamma(J, \nu)} e^{-\nu \text{tr}(Q^{-1}Q_0)}.
\]

The multichannel AR process for the disturbances in both the test and training signals consists of two types of unknown parameters: one is the deterministic AR coefficient matrix \( A \), and the other is the random spatial covariance matrices \( Q \) and \( Q_0 \). The available \textit{a priori} knowledge is imposed on the mean of \( Q \), i.e., \( Q_0 \), which can be obtained from sources such as land-use maps, past measurements, etc [15]. It is said that the importance of the \textit{a priori} knowledge \( Q \) is controlled by parameter \( \mu \), while the heterogeneities, i.e., the statistical differences between the test and training signals, are determined by parameter \( \nu \). Most importantly, \( Q \neq Q_0 \) with probability one, which ensures the environment non-homogeneous [10].

### III. BAYESIAN PARAMETRIC ADAPTIVE MATCHED FILTER

By assuming perfect knowledge on parameters \( A \) and \( Q_0 \), the solution to the problem of interest is the classical parametric matched filter (PMF) [17]:

\[
T_{\text{PMF}} = \frac{1}{\sum_{n=P}^{N-1} \text{tr}(\hat{x}_0(n)|H_1| - \text{tr}(\hat{x}_0(n)|H_0))} \geq \gamma_{\text{PMF}}.
\]

where \( \gamma_{\text{PMF}} \) denotes the PMF threshold subject to a selected probability of false alarm, and the whitened steering vector and test signal are obtained by using the true AR coefficient matrix \( A \):

\[
\hat{s}(n) = s(n) + \sum_{p=1}^{P} A^H(p) s(n-p),
\]

\[
\hat{x}_0(n) = x_0(n) + \sum_{p=1}^{P} A^H(p) x_0(n-p).
\]

The parametric AMF replaces the exact AR coefficient matrices \( A \) and the spatial covariance matrix \( Q_0 \) in the PMF statistic by their estimates (e.g., the maximum likelihood estimate (MLE) by using the training signals). For the non-homogeneous environment considered in this paper, due to the randomness of the spatial covariance matrix, we adapt a hybrid parametric AMF, which is denoted as the Bayesian PAMF (BPAMF) by first obtaining the MLE of the deterministic AR coefficient matrix \( A_{\text{ML}} \), then deriving a maximum \textit{a posteriori} estimate (MAP) of the stochastic spatial covariance matrix \( \hat{Q}_0|_{\text{MAP}} \) and finally replacing \( A \) and \( Q_0 \) in the PMF statistic by their estimates.

### A. MLE of \( A \)

According to the signal model, the joint probability density function (pdf) of the training signals can be approximated (ignore the conditionality on the \( P \) temporal vectors) as [18]

\[
f(x_1, \ldots, x_K|A, Q) = \frac{1}{\pi^J |Q|^{-L} \text{tr}(\Sigma^{-1})} e^{-\text{tr}(\Sigma^{-1}A)} K(N-P),
\]

where

\[
\Sigma(A) = \frac{1}{K(N-P)} \sum_{k=1}^{K} \sum_{n=P}^{N-1} \varepsilon_k(n)\varepsilon_k^H(n) \geq 0
\]

From AS2, we can remove the dependence of the above pdf on \( Q \) by integrating over \( Q \):

\[
f(x_1, \ldots, x_K|A) = \frac{1}{\pi^J |Q|^{-L} \text{tr}(\Sigma^{-1})} e^{-\text{tr}(\Sigma^{-1}A)} K(N-P),
\]

\[
= \frac{1}{\pi^J K(N-P) \text{det}(J, \mu)} \int |Q|^{-L} e^{-\text{tr}(\Sigma^{-1}A)} dQ
\]

\[
= \frac{1}{\pi^J K(N-P) \text{det}(J, \mu)} |\Sigma|^{-L+J} \hat{\Sigma} = K(N-P) \Sigma(A) + (\mu-J)\hat{Q}_0
\]

Therefore, minimizing the determinant of \( \hat{\Sigma} \) is equivalent to minimizing the determinant of \( \hat{\Sigma} \). Rewrite the matrix \( \hat{\Sigma} \) as

\[
\hat{\Sigma} = K(N-P) \Sigma(A) + (\mu-J)\hat{Q}_0
\]

\[
= \hat{R}_{xx} + \hat{R}_{yy}^H + \hat{R}_{yy} \hat{R}_{yy}^H \hat{R}_{yy} + (\mu-J)\hat{Q}_0
\]

\[
\hat{R}_{xx} = \sum_{k=1}^{K} \sum_{n=P}^{N-1} x_k(n) x_k^H(n),
\]

\[
\hat{R}_{yy} = \sum_{k=1}^{K} \sum_{n=P}^{N-1} y_k(n) y_k^H(n),
\]

with \( y_k(n) = [x_k^T(n-1), \ldots, x_k^T(n-P)]^T \in \mathbb{C}^{JP \times 1} \). Since \( \hat{R}_{yy} \) is nonnegative definite and the remaining terms \( \hat{R}_{xx} - \hat{R}_{yy}^H \hat{R}_{yy} \) do not depend on \( A \), it follows from (10) that

\[
\hat{\Sigma} \geq \hat{\Sigma}|_{A=A_{\text{ML}}} = \hat{R}_{xx} - \hat{R}_{yy}^H \hat{R}_{yy} + (\mu-J)\hat{Q}_0.
\]
where the MLE of $A$ is given as

$$\hat{A}_{ML} = -\hat{R}_{yy}^{-1}\hat{R}_{yx}^{-1}.$$  \hfill (13)

When $\hat{\Sigma}$ is minimized, the MLE $\hat{A}_{ML}$ will minimize any nondecreasing function including the determinant of $\hat{\Sigma}$.

**B. MAP Estimate of $Q_0$**

The MAP estimate of $Q_0$ requires the computation of the posterior distribution $f(Q_0|x_1, x_2, \ldots, x_K)$:

$$f(Q_0|x_1, x_2, \ldots, x_K) = \int f(Q_0, Q|x_1, x_2, \ldots, x_K) dQ,$$  \hfill (14)

where

$$f(Q_0, Q|x_1, x_2, \ldots, x_K) \propto f(x_1, x_2, \ldots, x_K|Q_0, Q)p(Q_0|Q)p(Q)$$

and

$$\propto |Q_0|^{\nu-J} |Q|^{-(L+\nu)} e^{-\text{tr}(Q^{-1}[\hat{\Sigma}+\nu Q_0])}.$$  \hfill (15)

As a result, (14) can be calculated as

$$\int f(Q_0, Q|x_1, x_2, \ldots, x_K) dQ$$

$$\propto |Q_0|^{\nu-J} \int |Q|^{-(L+\nu)} e^{-\text{tr}(Q^{-1}[\hat{\Sigma}+\nu Q_0])} dQ$$

$$\propto |Q_0|^{\nu-J} |\hat{\Sigma} + \nu Q_0|^{\nu+\nu+K(N-P)},$$  \hfill (16)

Taking the logarithm of the above equation, then taking the derivative with respect to $Q_0$, and equaling to zero, we have

$$\frac{\partial \ln f(Q_0|x_1, x_2, \cdot, x_K)}{\partial Q_0}$$

$$\propto (\nu - J)Q_0^{-1} - \nu (\mu + v + K(N-P)) |\hat{\Sigma} + \nu Q_0|^{-1} = 0,$$  \hfill (17)

which suggests that, given $A$, the estimate of $Q_0$ is

$$Q_0 = \frac{(\nu - J)}{\nu(\mu + J + K(N-P))} \hat{\Sigma}.$$  \hfill (18)

Replacing $A$ with $\hat{A}_{ML}$ of (13) in the above estimate (viz, $\hat{\Sigma}$), the MAP estimate of $Q_0$ is

$$\hat{Q}_{0,\text{MAP}} = \frac{(\nu - J)}{\nu(\mu + J + K(N-P))}$$

$$\times \left[ \hat{R}_{xx} - \hat{R}_{yx}^{-1}\hat{R}_{yy}^{-1}\hat{R}_{yx} + (\mu - J)\hat{Q}_0 \right].$$  \hfill (19)

It is seen that the MAP of $Q$ is a linear combination of a standard estimate of $Q$ as introduced in [18], [19] and the a priori knowledge $\hat{Q}$. This linear combination has been seen before for non-parametric approaches [10], [25].

**C. Bayesian PAMF**

With the ML estimate of $A$ and the MAP estimate of $Q_0$, the adaptive version of the PMF in the heterogeneous environment can be derived as

$$T_{B-PAMF} = \sum_{n=0}^{N-1} \hat{s}(n)Q_{0,\text{MAP}}^{-1}\hat{s}(n) + \sum_{p=1}^{P} \hat{\lambda}_{\text{B-PAMF}} \hat{\lambda}_{\text{B-PAMF}} \hat{s}(n)$$

where $\gamma_{B-PAMF}$ denotes the B-PAMF threshold subject to a selected probability of false alarm, $Q_{0,\text{MAP}}$ is given by (19), and the whitened steering vector and the whitened test signal are obtained by using $\hat{A}_{ML}$ given by (13):

$$\hat{s}(n) = s(n) + \sum_{p=1}^{P} \hat{A}_{ML}^H(p)s(n-p),$$

$$\hat{x}_0(n) = x_0(n) + \sum_{p=1}^{P} \hat{A}_{ML}^H(p)x_0(n-p).$$  \hfill (20)

From (20), on one hand, it is seen that the B-PAMF performs successive whitening, i.e., temporal whitening followed by spatial whitening, as opposed to joint spatio-temporal whitening across all $JN$ dimensions of the Bayesian AMF in [10]. On the other hand, compared with the standard PAMF [17], the B-PAMF incorporates the a priori knowledge, i.e., $\hat{Q}$, and utilizes the heterogeneity parameter $\nu$ and the importance parameter of the a priori knowledge, i.e., $\mu$ into the estimate of the spatial covariance matrix $Q$. Hence, it is seen that the B-PAMF provides computational ef ciency and mitigates training requirement, meanwhile exploiting the a priori knowledge and the heterogeneity to improve the performance of detection.

**IV. Numerical Examples**

In this section, simulation results are provided to illustrate the performance of the B-PAMF. Specifically, we first test the B-PAMF by using simulated data which conforms to $\text{AS1}$, $\text{AS2}$ and $\text{AS3}$, and then using the more challenging KASSPER dataset [15]. The disturbance signal is generated as a multichannel AR(2) process with AR coefficient $A$ and a spatial covariance matrix $Q$. The signal vector $s$ corresponds to a uniform equispaced linear array with randomly selected normalized spatial and Doppler frequencies. The signal-to-interference plus noise ratio (SINR) is defined as

$$\text{SINR} = \frac{|s|^2 s^H R^{-1} s}{\text{Tr}(R)},$$  \hfill (23)

where $R$ corresponds to the assigned AR coefficient matrix $A$ and the mean of the training spatial covariance matrix $Q$. For each Monte-Carlo trial, the spatial covariance matrix $Q$ for the training signal is generated as an inverse Wishart distribution with mean $\hat{Q}$ and then, given $Q$, the spatial covariance matrix $Q_0$ for the test signal is generated as a Wishart distribution with mean $Q$.

We focus here on performance comparison between the B-PAMF and the standard PAMF [17] in the non-homogeneous environment. Figs. 1(a)-(c) show the probability of detection versus the SINR for the B-PAMF and the standard PAMF in cases of different values of $\mu$ and $\nu$, when $P = 2$, $J = 4$, $N = 16$, $K = 1$, and $P_T = 0.01$. It is seen that, in all three cases, the B-PAMF outperforms the standard PAMF. Specifically, for a fixed value of $\mu$, a larger value of $\nu$, more homogeneous environments, results in slightly improved performance of detection of the B-PAMF, from Fig. 1(a) to Fig. 1(b). On the other hand, for a fixed value of $\nu$, increasing $\mu$ means more importance of the a priori knowledge of $Q$, which leads to wider performance gap between the B-PAMF and PAMF, as compared between Fig. 1(a) and Fig. 1(c).
Fig. 1. Probability of detection versus SINR (a) when $P = 2$, $\mu = 9$, and $\nu = 5$; (b) when $P = 2$, $\mu = 9$, and $\nu = 10$; (c) when $P = 2$, $\mu = 14$, and $\nu = 5$.

V. CONCLUSION

A Bayesian parametric adaptive matched lter has been proposed by modeling the disturbances in the test and training signals as a multichannel AR process and simulating the heterogeneity between the training and test signals by introducing random disturbance covariance matrices. The B-PAMF admits successive temporal and spatial whitening, which reduces the computational complexity of the joint spatial-temporal whitening based adaptive detectors. The training requirement is also reduced. Simulation results validate the effectiveness of the B-PAMF.

REFERENCES