

KNOWLEDGE-AIDED PARAMETRIC ADAPTIVE MATCHED FILTER WITH AUTOMATIC COMBINING FOR COVARIANCE ESTIMATION

Pu Wang*, Hongbin Li, and Zhe Wang†

Braham Himed

Stevens Institute of Technology
Hoboken, NJ 07030

AFRL/RVMD
Dayton, OH 45433

ABSTRACT

In this paper, a knowledge-aided parametric adaptive matched filter (KA-PAMF) is proposed that utilizing both observations (including the test and training signals) and *a priori* knowledge of the spatial covariance matrix. Unlike existing KA-PAMF methods, the proposed KA-PAMF is able to automatically adjust the combining weight of *a priori* covariance matrix, thus gaining enhanced robustness against uncertainty in the prior knowledge. Meanwhile, the proposed KA-PAMF is significantly more efficient than its KA non-parametric counterparts when the amount of training signals is limited. One distinct feature of the proposed KA-PAMF is the inclusion of both the test and training signals for automatic determination of the combining weights for the prior spatial covariance matrix and observations. Numerical results are presented to demonstrate the effectiveness of the proposed KA-PAMF, especially in the limited training scenarios.

Index Terms— STAP, parametric adaptive matched filter.

1. INTRODUCTION

Traditional space-time adaptive processing (STAP) methods such as the Kelly's generalized likelihood ratio test (GLRT) [1] and the adaptive matched filter (AMF) [2] usually require excessive homogeneous training (secondary) data to obtain an accurate estimate of the disturbance covariance matrix for adaptive detection of targets. For example, it is known that for these methods at least $K \geq JN$ homogeneous training signals are required for a full-rank covariance matrix estimator, where J is the number of channels and N is the number of temporal observations.

Knowledge-aided detectors have been introduced to reduce the demanding need of training signals by fusing some prior knowledge in the estimation of the disturbance covariance matrix [3]. One approach toward this goal is based on the Bayesian framework, which embeds the *a priori* knowledge via a prior distribution of the disturbance covariance matrix [4–10]. Another approach was based on the regularized method [11–14], which usually linearly shrinks the eigenvalues of the sample covariance matrix towards a targeted covariance matrix, e.g., the identity matrix up to a scaling factor [11], a diagonal matrix consisting of the diagonal entries of the sample covariance matrix [12], or the *a priori* covariance matrix [14]. Interestingly, both approaches result in a colored loading form between the *a priori* matrix and the sample covariance matrix. While the weights

in the Bayesian approach are determined by the hyper-parameters of the statistical model, the regularized method uses the (training) signals to determine the amount of regularization.

The regularized method has been considered for STAP detection, which employs the loaded covariance matrix for signal whitening and test statistic calculation. Specifically, [14] introduces the knowledge-aided AMF (KA-AMF) which first linearly combines the sample covariance matrix and the *a priori* covariance matrix [11], which is then used in the conventional AMF for adaptive detection. The linear combining weights are determined from the training signals. Results obtained with the KASSPER dataset show that with $J = 11$ channels and $N = 32$ pulses, the proposed KA-AMF offers good detection performance by using $K = 50$ training signals. Still, it may be difficult to obtain $K = 50$ homogeneous training signals in a non-homogeneous environment, where a more efficient solution with less training data is desirable. Moreover, the computational complexity of the KA-AMF is still high, since it needs compute to the inverse of the $JN \times JN$ covariance matrix.

In this paper, we aim to address both issues of limited homogeneous training signals and the complexity by extending the parametric adaptive matched filter (PAMF) [15, 16] and integrate knowledge-aided processing. As shown with numerous simulated and measured STAP datasets [16, 17], the parametric framework using a multichannel auto-regressive (AR) process can effectively and efficiently capture the correlation structure of the disturbance in STAP. Furthermore, we develop a regularized method which automatically determines the combining weights *jointly* from the test signal and training signals. Our scheme is different from other regularized methods, such as the KA-AMF [14], which uses only the training signals to determine the combining weights. It appears that the inclusion of the test signal for weight calculation is critical to achieving a robust performance in scenarios where the number of training data is limited. Our proposed knowledge-aided PAMF with automatic combining (referred to as the KA-AC-PAMF hereafter) is derived in a three-step approach. First, conditioned on the given AR temporal correlation matrices, a partially adaptive detector is derived according to the GLRT principle which yields an estimate of the spatial covariance matrix. Then, the estimate of the spatial covariance matrix is linearly combined with the prior knowledge in an adaptive way from the test and training signals. Finally, the fully adaptive KA-AC-PAMF is obtained by replacing the AR temporal correlation matrices in the partially adaptive detector by its maximum likelihood (ML) estimate.

2. SIGNAL MODEL

Consider the problem of detecting a *known* multi-channel signal with *unknown* amplitude in the presence of spatially and temporally cor-

*The first author performed the work while at Stevens Institute of Technology. He is now at the Schlumberger-Doll Research, Cambridge, MA 01239 (e-mail: pwang@ieee.org)

†The work was supported in part by a subcontract with Dynetics, Inc. for research sponsored by the Air Force Research Laboratory (AFRL) under Contract FA8650-08-D-1303.

related disturbance: (e.g., [18]):

$$\begin{aligned} H_0: \quad \mathbf{x}_0(n) &= \mathbf{d}_0(n), n = 0, 1, \dots, N-1, \\ H_1: \quad \mathbf{x}_0(n) &= \alpha \mathbf{s}(n) + \mathbf{d}_0(n), n = 0, 1, \dots, N-1, \end{aligned} \quad (1)$$

where all vectors are $J \times 1$ vectors obtained from J spatial channels/receivers, and N is the number of temporal observations/snapshots. The subindex of $\mathbf{x}_0(n)$ is referred to the range bin of interest, and $\{\mathbf{x}_0(n)\}_{n=1}^N$ forms the test signal from J receivers and N pulses. The steering vector $\{\mathbf{s}(n)\}_{n=1}^N$ takes into account of the array geometry with spatial frequency ω_s and the Doppler frequency ω_d . For a uniformly equi-spaced linear array, the (normalized) steering vector is given as $\mathbf{s}(n) = e^{j\omega_d(n-1)} [1, \dots, e^{j\omega_s(J-1)}]^T / \sqrt{JN}$. In addition, α denotes the *unknown*, deterministic and complex-valued signal amplitude, and $\mathbf{d}_0(n)$ is the disturbance signal that is correlated in space and time. Besides the test signal $\mathbf{x}_0(n)$, there may be a set of *target-free* training signals $\mathbf{x}_k(n)$, $k = 1, 2, \dots, K$:

$$\mathbf{x}_k(n) = \mathbf{d}_k(n), k = 1, \dots, K. \quad (2)$$

Denote the following $JN \times 1$ space-time vectors $\mathbf{s} \triangleq [\mathbf{s}^T(0), \dots, \mathbf{s}^T(N-1)]^T$, $\mathbf{d}_k \triangleq [\mathbf{d}_k^T(0), \dots, \mathbf{d}_k^T(N-1)]^T$, and $\mathbf{x}_k \triangleq [\mathbf{x}_k^T(0), \dots, \mathbf{x}_k^T(N-1)]^T$. It is assumed that the disturbance signals \mathbf{d}_k , $k = 0, 1, \dots, K$, are independent and identically distributed (i.i.d.) with the complex Gaussian distribution $\mathbf{d}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{R})$, where \mathbf{R} is the unknown space-time covariance matrix [1]. The parametric framework further assumes that the disturbance signals $\{\mathbf{d}_k\}_{k=0}^K$ in the test and training signals follow the assumption below [16]:

- **AS — Multi-Channel AR Model:** The disturbance signals $\mathbf{d}_k(n)$, $k = 0, \dots, K$, in the test and training signals are modeled as J -channel AR(P) processes with model order P :

$$\mathbf{d}_k(n) = -\sum_{i=1}^P \mathbf{A}^H(i) \mathbf{d}_k(n-i) + \boldsymbol{\varepsilon}_k(n), \quad (3)$$

where $\{\mathbf{A}(i)\}_{i=1}^P$ denote the *unknown* $J \times J$ AR coefficient matrices, \mathbf{A}^H denotes the conjugate transpose of \mathbf{A} , $\boldsymbol{\varepsilon}_k(n)$ denote the $J \times 1$ spatial noise vectors that are temporally white but spatially colored Gaussian noise: $\{\boldsymbol{\varepsilon}_k(n)\}_{k=0}^K \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q})$, and \mathbf{Q} denotes the *unknown* $J \times J$ spatial covariance matrix.

In other words, the disturbance covariance matrix \mathbf{R} is parametrized in **AS** with P AR coefficient matrices $\mathbf{A}(p)$ and the spatial covariance matrix \mathbf{Q} . In many cases, it is possible to have some *a priori* knowledge on \mathbf{R} which can be utilized for improved detection performance. Such knowledge can be obtained from previously acquired database, e.g., digital terrain maps, synthetic aperture radar (SAR) images, as well as from real-time information including the transmit/receive array configurations, beam patterns, etc. [3]. In this paper, we only assume *a priori* knowledge of the spatial covariance matrix, denoted as $\tilde{\mathbf{Q}}$, as it may be easier to obtain compared with the *a priori* knowledge $\tilde{\mathbf{R}}$ on the spatial-temporal covariance matrix.

3. PROPOSED APPROACH

In this section, we consider a three-step approach to develop the KA-AC-PAMF detector. First, assuming that the AR coefficient matrix $\mathbf{A} \triangleq [\mathbf{A}^T(1), \mathbf{A}^T(2), \dots, \mathbf{A}^T(P)]^T$ is known, a partially adaptive PAMF is derived by finding the ML estimates of unknown parameters \mathbf{Q} and α that maximize the joint likelihood function of the test signal \mathbf{x}_0 and the training signals \mathbf{x}_k . Then, the ML estimate of

\mathbf{Q} is regularized with the prior $\tilde{\mathbf{Q}}$ according to the minimum mean squared error (MMSE) criterion. Third, the AR coefficient matrix \mathbf{A} in the partially adaptive PAMF is replaced by its ML estimate of \mathbf{A} leading to the fully adaptive KA-AC-PAMF.

3.1. Partially Adaptive KA-PAMF

Assuming **AS** with a known \mathbf{A} , the partially adaptive PAMF takes the form of a likelihood ratio test

$$T = \frac{\max_{\alpha, \mathbf{Q}} f_1(\alpha, \mathbf{Q})}{\max_{\mathbf{Q}} f_0(\mathbf{Q})}, \quad (4)$$

where $f_i(\alpha, \mathbf{Q})$ is the joint (asymptotic) likelihood function of \mathbf{x}_0 and \mathbf{x}_k under H_i , $i = 0, 1$,

$$f_i(\alpha, \mathbf{Q}) = \left[\frac{1}{\pi^J |\mathbf{Q}|} e^{-\text{tr}(\mathbf{Q}^{-1} \boldsymbol{\Gamma}(\alpha))} \right]^{(K+1)(N-P)}, \quad (5)$$

with $\alpha = 0$ when $i = 0$,

$$\boldsymbol{\Gamma}(\alpha) = \left(\tilde{\mathbf{X}}_0 - \alpha \tilde{\mathbf{S}} \right) \left(\tilde{\mathbf{X}}_0 - \alpha \tilde{\mathbf{S}} \right)^H + \sum_{k=1}^K \tilde{\mathbf{X}}_k \tilde{\mathbf{X}}_k^H, \quad (6)$$

and

$$\tilde{\mathbf{S}} = [\tilde{\mathbf{s}}(P), \dots, \tilde{\mathbf{s}}(N-1)] \in^{J \times (N-P)}, \quad (7)$$

$$\tilde{\mathbf{X}}_k = [\tilde{\mathbf{x}}_k(P), \dots, \tilde{\mathbf{x}}_k(N-1)] \in^{J \times (N-P)}, \quad (8)$$

$$\tilde{\mathbf{x}}_k(n) = \mathbf{x}_k(n) + \mathbf{A}^H \mathbf{y}_k(n), \quad (9)$$

$$\tilde{\mathbf{s}}(n) = \mathbf{s}(n) + \mathbf{A}^H \mathbf{t}(n), \quad (10)$$

$$\mathbf{y}_k(n) = [\mathbf{x}_k^T(n-1), \dots, \mathbf{x}_k^T(n-P)]^T, \quad (11)$$

$$\mathbf{t}(n) = [\mathbf{s}^T(n-1), \dots, \mathbf{s}^T(n-P)]^T, \quad (12)$$

for $k = 0, 1, \dots, K$. Note that (9) and (10) perform the temporal whitening for $\mathbf{x}_k(n)$ and $\mathbf{s}(n)$. It can be shown in [19] that the partially adaptive GLRT takes the form of

$$T = \frac{\left| \text{tr} \left\{ \tilde{\mathbf{S}}^H \boldsymbol{\Psi}^{-1} \tilde{\mathbf{X}}_0 \right\} \right|^2}{\left| \text{tr} \left\{ \tilde{\mathbf{S}}^H \boldsymbol{\Psi}^{-1} \tilde{\mathbf{S}} \right\} \right|} = \frac{\left| \sum_{n=P}^{N-1} \tilde{\mathbf{s}}^H(n) \boldsymbol{\Psi}^{-1} \tilde{\mathbf{x}}_0(n) \right|^2}{\sum_{n=P}^{N-1} \tilde{\mathbf{s}}^H(n) \boldsymbol{\Psi}^{-1} \tilde{\mathbf{s}}(n)}, \quad (13)$$

where

$$\boldsymbol{\Psi} = \frac{1}{L} \left(\tilde{\mathbf{X}}_0 \mathbf{P}^\perp \tilde{\mathbf{X}}_0^H + \sum_{k=1}^K \tilde{\mathbf{X}}_k \tilde{\mathbf{X}}_k^H \right), \quad (14)$$

with $L = K(N-P) + N - P - 1$ and $\mathbf{P}^\perp = \mathbf{I} - \mathbf{P} = \mathbf{I} - \tilde{\mathbf{S}}^H (\tilde{\mathbf{S}}^H)^\dagger$ denoting the projection matrix projecting to the orthogonal complement of the range of $\tilde{\mathbf{S}}^H$.

It is seen that the partially adaptive PAMF first performs the temporal whitening process to obtain the test and training signals $\{\tilde{\mathbf{x}}_k(n)\}_{k=0}^K$ and the steering vector $\tilde{\mathbf{s}}(n)$ via (9) and (10), and then performs the spatial whitening process in (13) with $\boldsymbol{\Psi}$ of (14). More importantly, the spatial whitening matrix $\boldsymbol{\Psi}$ of (14) includes contribution from the temporally whitened test signal $\tilde{\mathbf{x}}_0(n)$ (after projecting onto the orthogonal complement of the range space of $\tilde{\mathbf{S}}^H$) and the temporally whitened training signals $\tilde{\mathbf{x}}_k(n)$. It is also noted that

Ψ of (14) is an unbiased estimator of \mathbf{Q} under the two hypotheses, as summarized in the following proposition:¹

Proposition: Given the signal model of $\{\mathbf{x}_k(n)\}_{k=0}^K$ and $\mathbf{s}(n)$ and assuming the multichannel AR model in \mathbf{AS} , the estimate Ψ of (14) is an unbiased estimate of \mathbf{Q} under the two hypotheses

$$E\{\Psi\} = \mathbf{Q}, \quad \text{under } H_0 \text{ and } H_1. \quad (15)$$

3.2. Automatic Weighting Between Ψ and $\bar{\mathbf{Q}}$

We consider a linear combination scheme between the unbiased Ψ and the prior $\bar{\mathbf{Q}}$ according to the minimum mean squared error (MMSE) criterion [11, 14]. Consistent with the fact that the estimate Ψ consists of both the test and training signals, we propose to use the test and training signals for automatic determination of the linear combining weights, thus extending the regularized method in [11] and [14] which uses only the training signals. As shown in later numerical examples, the inclusion of the test signal leads to improved performance when the number of training signals is limited.

Specifically, we consider a convex combination between $\bar{\mathbf{Q}}$ and the estimate Ψ ²

$$\tilde{\mathbf{Q}} = \beta \bar{\mathbf{Q}} + (1 - \beta) \Psi, \quad (16)$$

where $\beta \in [0, 1]$ is the combining weight to be determined. This scheme is to balance the contribution from the prior knowledge and the observed signals. One can replace $\bar{\mathbf{Q}}$ by an identity matrix in the case that $\bar{\mathbf{Q}}$ is unavailable or has a large amount of uncertainty.

As shown in [11, 14] and thanks to the proposition of the unbiasedness, the optimal weight β in (16) is determined as

$$\beta = \frac{E\{\|\Psi - \mathbf{Q}\|^2\}}{E\{\|\Psi - \bar{\mathbf{Q}}\|^2\}} = \frac{E\{\|\Psi - \mathbf{Q}\|^2\}}{E\{\|\Psi - \mathbf{Q}\|^2\} + \|\bar{\mathbf{Q}} - \mathbf{Q}\|^2}. \quad (17)$$

Define $\rho \triangleq E\{\|\Psi - \mathbf{Q}\|^2\}$ and $\nu \triangleq \|\bar{\mathbf{Q}} - \mathbf{Q}\|^2$. Since ρ and ν depend on the true but unobservable \mathbf{Q} , the optimal combining weight β has to be estimated from the observations. In our case, we use both the test and training signals to achieve this purpose.

First, regarding the estimate of ρ , we show in the following that Ψ can be considered as equivalently the sample covariance matrix from a set of L i.i.d. Gaussian vectors, among which $N - P - 1$ vectors are obtained from the test signal and the remaining $K(N - P)$ vectors are from the training signals. From (14), Ψ consists of two components:

$$L\Psi = \tilde{\mathbf{X}}_0 \mathbf{P}^\perp \tilde{\mathbf{X}}_0^H + \sum_{k=1}^K \tilde{\mathbf{X}}_k \tilde{\mathbf{X}}_k^H = \mathbf{E}_0 \mathbf{P}^\perp \mathbf{E}_0^H + \sum_{k=1}^K \mathbf{E}_k \mathbf{E}_k^H,$$

where $\mathbf{E}_k = [\epsilon_k(P), \dots, \epsilon_k(N - 1)]$, $k = 0, 1, 2, \dots, K$ with columns $\epsilon_k(n)$ distributed as i.i.d. complex Gaussian vectors with zero mean and covariance matrix \mathbf{Q} .

Then the $(N - P) \times (N - P - 1)$ orthogonal projection matrix \mathbf{P}^\perp can be decomposed to

$$\mathbf{P}^\perp = \mathbf{U}_P \mathbf{U}_P^H, \quad (18)$$

where \mathbf{U}_P is an $(N - P) \times (N - P - 1)$ matrix with $N - P - 1$ orthonormal columns. Together with the $N - P$ i.i.d. Gaussian

¹The proof is straightforward by noting that temporally whitened test signal $\tilde{\mathbf{x}}_0(n)$ is statistically equivalent to the spatial noise vector $\epsilon_0(n)$ plus $\alpha \tilde{\mathbf{s}}(n)$, while the training signals $\tilde{\mathbf{x}}_k(n)$ are statistically equivalent to $\epsilon_k(n)$.

²Other linear combination such as the generalized linear combination (GLC) can be derived similarly, which has a similar performance.

columns in \mathbf{E}_0 and the $(N - P - 1)$ orthonormal columns of \mathbf{U}_P , the resulting $N - P - 1$ columns of

$$\mathbf{Z}_0 = \mathbf{X}_0 \mathbf{U}_P = \mathbf{E}_0 \mathbf{U}_P = [\mathbf{z}_0(P), \dots, \mathbf{z}_0(N - P - 1)], \quad (19)$$

are i.i.d. Gaussian vectors with zero mean and covariance matrix \mathbf{Q} , i.e., $\mathbf{z}_0(n) \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q})$. As a result,

$$\tilde{\mathbf{X}}_0 \mathbf{P}^\perp \tilde{\mathbf{X}}_0^H = \mathbf{E}_0 \mathbf{P}^\perp \mathbf{E}_0^H = \mathbf{Z}_0 \mathbf{Z}_0^H = \sum_{n=P}^{N-P-1} \mathbf{z}_0(n) \mathbf{z}_0^H(n). \quad (20)$$

For the training signal component, we define $\mathbf{z}_k(n) = \tilde{\mathbf{x}}_k(n)$, where $\mathbf{z}_k(n)$ are i.i.d. Gaussian vectors with zero mean and covariance matrix \mathbf{Q} . Stacking all $\mathbf{z}_k(n)$, we have a set of L i.i.d. Gaussian vectors

$$\mathbf{z}(n) = \begin{cases} \mathbf{z}_0(l), & n = 1, 2, \dots, N - P - 1, l = n + P - 1, \\ \mathbf{z}_k(l), & n = N - P, \dots, L, k = \left\lfloor \frac{n - (N - P)}{N - P} \right\rfloor, \\ & l = n - (k + 1)(N - P) + P. \end{cases} \quad (21)$$

Therefore, $\Psi = \frac{1}{L} \sum_{n=1}^L \mathbf{z}(n) \mathbf{z}^H(n)$ can be considered as the sample covariance matrix from L i.i.d. Gaussian vectors $\mathbf{z}(n)$ with zero mean and covariance matrix \mathbf{Q} .

With the above results, the estimate of ρ reduces to the estimation of β from L i.i.d. random vectors with zero mean and covariance matrix \mathbf{Q} . Then, the coefficient ρ can be adaptively estimated as

$$\tilde{\rho} = \frac{1}{(L - 1)L} \sum_{n=1}^L \|\mathbf{z}(n) \mathbf{z}^H(n) - \Psi\|^2. \quad (22)$$

For the second quantity ν , we simply replace the true \mathbf{Q} by the unbiased estimate Ψ

$$\tilde{\nu} = \|\mathbf{Q}_0 - \Psi\|^2, \quad (23)$$

which leads to the knowledge-aided spatial covariance matrix estimate

$$\tilde{\mathbf{Q}} = \frac{\tilde{\rho}}{\tilde{\rho} + \tilde{\nu}} \bar{\mathbf{Q}} + \frac{\tilde{\nu}}{\tilde{\rho} + \tilde{\nu}} \Psi. \quad (24)$$

3.3. Fully Adaptive KA-PAMF

Finally, an adaptive estimate of \mathbf{A} is needed to enable a fully adaptive KA-AC-PAMF. In the following, the ML estimate of \mathbf{A} from the training signals, derived in [20], is used. Specifically, the ML estimate of \mathbf{A} can be computed as follows

$$\hat{\mathbf{A}}_{\text{ML}} = -\hat{\mathbf{R}}_{yx,K}^H \hat{\mathbf{R}}_{yy,K}^{-1}, \quad (25)$$

where

$$\hat{\mathbf{R}}_{yy,K} = \sum_{k=1}^K \sum_{n=P}^{N-1} \mathbf{y}_k(n) \mathbf{y}_k^H(n), \quad (26)$$

$$\hat{\mathbf{R}}_{yx,K} = \sum_{k=1}^K \sum_{n=P}^{N-1} \mathbf{y}_k(n) \mathbf{x}_k^H(n), \quad (27)$$

with $\mathbf{y}_k(n)$ is defined in (11).

Using the ML estimate of \mathbf{A} in (13), the proposed KA-AC-PAMF detector takes the form of

$$T_{\text{KA-PAMF}} = \frac{\left| \sum_{n=P}^{N-1} \hat{\mathbf{s}}^H(n) \hat{\mathbf{Q}}^{-1} \hat{\mathbf{x}}_0(n) \right|^2}{\sum_{n=P}^{N-1} \hat{\mathbf{s}}^H(n) \hat{\mathbf{Q}}^{-1} \hat{\mathbf{s}}(n)}, \quad (28)$$

where the fully adaptively temporally whitened vectors

$$\begin{aligned} \hat{\mathbf{x}}_k(n) &= \mathbf{x}_k(n) + \sum_{p=1}^P \hat{\mathbf{A}}_{\text{ML}}^H(p) \mathbf{x}_k(n-p), \\ \hat{\mathbf{s}}(n) &= \mathbf{s}(n) + \sum_{p=1}^P \hat{\mathbf{A}}_{\text{ML}}^H(p) \mathbf{s}(n-p), \end{aligned}$$

and, correspondingly,

$$\hat{\mathbf{X}}_k = [\hat{\mathbf{x}}_k(P), \dots, \hat{\mathbf{x}}_k(N-1)], \hat{\mathbf{S}} = [\hat{\mathbf{s}}(P), \dots, \hat{\mathbf{s}}(N-1)].$$

Then, the vectors $\hat{\mathbf{z}}(n)$ is formed from the columns of the following matrices

$$\hat{\mathbf{Z}}_0 = \hat{\mathbf{X}}_0 \hat{\mathbf{U}}_{\hat{\mathbf{P}}}, \quad \hat{\mathbf{Z}}_k = \hat{\mathbf{X}}_k, \quad k = 1, 2, \dots, K,$$

where $\hat{\mathbf{U}}_{\hat{\mathbf{P}}}$ given as $\hat{\mathbf{U}}_{\hat{\mathbf{P}}} = \text{Null}(\hat{\mathbf{P}})$ with $\hat{\mathbf{P}} = \hat{\mathbf{S}}^H (\hat{\mathbf{S}})^{\dagger}$. Finally, the spatial covariance matrix is estimated as

$$\hat{\mathbf{Q}} = \frac{\hat{\rho}}{\hat{\rho} + \hat{\nu}} \hat{\mathbf{Q}} + \frac{\hat{\nu}}{\hat{\rho} + \hat{\nu}} \hat{\Psi},$$

where $\hat{\Psi} = L^{-1} \sum_{n=1}^L \hat{\mathbf{z}}(n) \hat{\mathbf{z}}^H(n)$ and

$$\hat{\rho} = \frac{1}{(L-1)L} \sum_{n=1}^L \|\hat{\mathbf{z}}(n) \hat{\mathbf{z}}^H(n) - \hat{\Psi}\|^2, \quad \hat{\nu} = \|\mathbf{Q}_0 - \hat{\Psi}\|^2.$$

4. NUMERICAL RESULTS

In this section, numerical results are provided to compare the proposed KA-AC-PAMF with other conventional and knowledge-aided parametric detectors in terms of the detection performance versus the signal-to-interference-plus-noise ratio (SINR) for a probability of false alarm $P_f = 0.01$. Specifically, we consider 1) the conventional PAMF [16], 2) the Bayesian KA-PAMF [19] with a random guess of the hyper-prior parameter, and 3) the KA-AMF [14]. The analytical performance of the optimal matched filter is also shown to provide the benchmark. In simulations, the disturbance signals \mathbf{d}_k are generated from the AR(3) process with a given \mathbf{A} and \mathbf{Q} . To account for the uncertainty of prior knowledge about \mathbf{Q} , a perturbed version of \mathbf{Q} is used as $\tilde{\mathbf{Q}}$ [14]

$$\tilde{\mathbf{Q}} = \mathbf{Q} \odot \mathbf{t}_s \mathbf{t}_s^H, \quad (29)$$

where \mathbf{t}_s is a $J \times 1$ vector of i.i.d. Gaussian random vectors with mean 1 and variance σ_t^2 , and \odot denotes the Hadamard matrix product. Specifically, we refer to the σ_t^2 as the uncertainty variance (UV). As the UV increases, the prior knowledge $\tilde{\mathbf{Q}}$ is on average away from the true \mathbf{Q} . The SINR is defined as

$$\text{SINR} = |\alpha|^2 |\mathbf{s}^H \mathbf{R}^{-1} \mathbf{s}|, \quad (30)$$

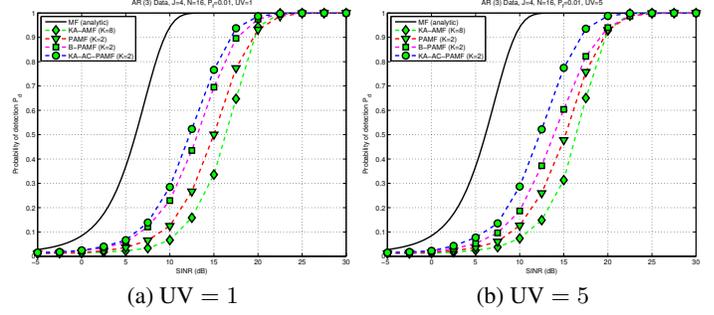


Fig. 1. Probability of detection when $K = 2$, $J = 4$, $N = 16$, $P_f = 0.01$

where \mathbf{R} is the spatial-temporal covariance matrix corresponding to the multi-channel AR process with coefficients \mathbf{A} and \mathbf{Q} . Then the prior $\tilde{\mathbf{R}}$ for the KA-AMF is generated by using the same perturbation model (29) as $\tilde{\mathbf{R}} = \mathbf{R} \odot \mathbf{t} \mathbf{t}^H$, where \mathbf{t} is a $JN \times 1$ vector of i.i.d. Gaussian random vectors with mean 1 and the same variance σ_t^2 . For the synthesized AR dataset, the number of channel is $J = 4$ and the number of temporal observations is $N = 16$.

We first consider the case of UV = 1, e.g., a case with a relatively reliable prior ($\tilde{\mathbf{Q}}$ for the parametric detectors and $\tilde{\mathbf{R}}$ for the KA-AMF) and the number of training signals $K = 2$. Without using the prior knowledge, the conventional PAMF fully relies on the $K = 2$ training signals and, as shown in Fig. 1 (a), its performance is the worst among the parameter detectors. Although utilizing the prior knowledge, the KA-AMF shows worse performance than the parametric detectors since the limited $K = 8$ training signals, compared with the total dimension $JN = 64$, are unable to obtain a good covariance matrix estimate without exploiting the structural information of \mathbf{R} . In contrast, by using the prior $\tilde{\mathbf{Q}}$ and exploiting the multi-channel AR structure, the knowledge-aided parametric detectors, i.e., the B-PAMF and the proposed KA-AC-PAMF, give improved performance than the conventional PAMF and the KA-AMF. In addition, the proposed KA-AC-PAMF has an SINR improvement of about 1 dB over the B-PAMF.

Next, we increase the prior uncertainty to UV = 5. Since the conventional PAMF uses no prior knowledge, its performance is independent of the UV as shown in Fig. 1 (b). Comparison between Fig. 1 (a) and Fig. 1 (b) reveals that the increased UV has little impacts on the KA-AMF and the proposed KA-AC-PAMF since they both determine the combining weight adaptively from the observations. There is a noticeable performance degradation for the B-PAMF which uses a non-adaptive weight on $\tilde{\mathbf{Q}}$, regardless of the prior uncertainty.

5. CONCLUSION

In this paper, a new knowledge-aided PAMF is proposed which automatically determines the linear combining weights between the prior covariance matrix and the conventional covariance estimate using both the test and training signals. On one hand, the proposed detector is more robust against uncertainty in the prior knowledge than the existing knowledge-aided PAMF. On the other hand, it also outperforms the knowledge-aided non-parametric detector in scenarios with limited training signals. Simulation results confirm the effectiveness of the proposed detector.

6. REFERENCES

- [1] E. J. Kelly, "An adaptive detection algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 22, no. 1, pp. 115–127, March 1986.
- [2] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, no. 1, pp. 208–216, January 1992.
- [3] M. C. Wicks, M. Rangaswamy, R. Adve, and T. B. Hale, "Space-time adaptive processing: a knowledge-based perspective for airborne radar," *IEEE Signal Processing Magazine*, vol. 23, no. 1, pp. 51–65, January 2006.
- [4] A. De Maio and A. Farina, "Adaptive radar detection: a Bayesian approach," in *Proceedings of the 2006 International Radar Symposium*, Krakow, Poland, May 2006, pp. 85–88.
- [5] A. De Maio, A. Farina, and G. Foglia, "Adaptive radar detection: a Bayesian approach," in *Proceedings of the 2007 IEEE International Conference on Radar*, Waltham, MA, April 2007, pp. 624–629.
- [6] O. Besson, J. Y. Tournet, and S. Bidon, "Knowledge-aided bayesian detection in heterogeneous environment," *IEEE Signal Processing Letters*, pp. 355–358, May 2007.
- [7] O. Besson, J.-Y. Tournet, and S. Bidon, "Knowledge-aided Bayesian detection in heterogeneous environment," *IEEE Signal Processing Letters*, vol. 14, no. 5, pp. 355–358, May 2007.
- [8] S. Bidon, O. Besson, and J.-Y. Tournet, "A Bayesian approach to adaptive detection in non-homogeneous environments," *IEEE Transactions on Signal Processing*, vol. 56, no. 1, pp. 205–217, January 2008.
- [9] A. De Maio, A. Farina, and G. Foglia, "Knowledge-aided Bayesian radar detectors & their application to live data," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 46, no. 1, pp. 170–183, February 2010.
- [10] S. Bidon, O. Besson, and J. Y. Tournet, "Knowledge-aided STAP in heterogeneous clutter using a hierarchical bayesian algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 3, pp. 1863–1879, July 2011.
- [11] O. Ledoit and M. Wolf, "A well-conditioned estimator for large-dimensional covariance matrices," *Journal of Multivariate Analysis*, vol. 88, no. 2, pp. 365 – 411, 2004.
- [12] J. P. Hoffbeck and D. A. Landgrebe, "Covariance matrix estimation and classification with limited training data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 7, pp. 763–767, July 1996.
- [13] P. J. Bickel and E. Levina, "Regularized estimation of large covariance matrices," *Ann. Statist.*, vol. 36, no. 1, pp. 199–227, 2008.
- [14] P. Stoica, J. Li, X. Zhu, and J. R. Guerci, "On using a priori knowledge in space-time adaptive processing," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2598–2602, June 2008.
- [15] M. Rangaswamy and J. H. Michels, "A parametric multichannel detection algorithm for correlated non-Gaussian random processes," in *Proceedings of the 1997 IEEE National Radar Conference*, Syracuse, NY, May 1997, pp. 349–354.
- [16] J. R. Román, M. Rangaswamy, D. W. Davis, Q. Zhang, B. Himed, and J. H. Michels, "Parametric adaptive matched filter for airborne radar applications," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 36, no. 2, pp. 677–692, April 2000.
- [17] J. H. Michels, B. Himed, and M. Rangaswamy, "Robust STAP detection in a dense signal airborne radar environment," *Signal Processing*, special issue on New Trends and Findings in Antenna Array Processing for Radar, vol. 84, pp. 1625–1636, 2004.
- [18] J. Ward, "Space-time adaptive processing for airborne radar," Technical Report 1015, Lincoln Laboratory, MIT, December 1994.
- [19] P. Wang, H. Li, and B. Himed, "Knowledge-aided parametric tests for multichannel adaptive signal detection," *IEEE Trans. Signal Processing*, vol. 59, no. 12, pp. 5970–5982, December 2011.
- [20] K. J. Sohn, H. Li, and B. Himed, "Parametric Rao test for multichannel adaptive signal detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 3, pp. 920–933, July 2007.