Knowledge-Aided Bayesian Detection in Heterogeneous Environments

Olivier Besson, *Senior Member, IEEE*, Jean-Yves Tourneret, *Member, IEEE*, and Stéphanie Bidon, *Student Member, IEEE*

***Abstract*—We address the problem of detecting a signal of interest in the presence of noise with unknown covariance matrix, using a set of training samples. We consider a situation where the environment is not homogeneous, i.e., when the covariance matrices of the primary and the secondary data are different. A knowledge-aided Bayesian framework is proposed, where these covariance matrices are considered as random, and some information about the covariance matrix of the training samples is available. Within this framework, the maximum *a priori* (MAP) estimate of the primary data covariance matrix is derived. It is shown that it amounts to colored loading of the sample covariance matrix of the secondary data. The MAP estimate is in turn used to yield a Bayesian version of the adaptive matched filter. Numerical simulations illustrate the performance of this detector, and compare it with the conventional adaptive matched filter.**

***Index Terms*—Bayesian detection, heterogenous environments, knowledge-aided processing, maximum *a posteriori* estimation.**

I. Introduction

A fundamental task of any radar system is to detect the presence of a target, with given space and/or time signature, in a cell under test (CUT), in the presence of noise which consists of clutter, thermal noise, and possibly jamming [1]. In the Gaussian case, when the covariance matrix $\boldsymbol{M}_p$ of the noise in the CUT is known, the optimal processor consists of a whitening step followed by matched filtering [2]. However, the statistics of the noise in the CUT (also referred to as the primary data) are generally unknown and hence $\boldsymbol{M}_p$ must be estimated. This goal is generally achieved through the use of training samples (the so-called secondary data), which consist of noise only, and whose covariance matrix $\boldsymbol{M}_s$ would ideally be $\boldsymbol{M}_p$. Training samples are usually obtained from range cells adjacent to the CUT. When $\boldsymbol{M}_s = \boldsymbol{M}_p$, $\boldsymbol{M}_p$ can be substituted for the sample covariance matrix computed from the secondary data in the optimal detector; this is the principle of the adaptive matched filter (AMF) [3].

However, the assumption of an homogeneous environment, i.e., $\boldsymbol{M}_p = \boldsymbol{M}_s$, is somewhat idealistic. Indeed, it has been experienced with real data that this assumption can be seriously violated [4], [5]. Such heterogeneous environments can be due either to the terrain (highly complex and nonstationary clutter) or to the geometry of the array (e.g., nonlinear arrays or non side-looking configurations). In the case where $\boldsymbol{M}_s$ differs from $\boldsymbol{M}_p$, detectors based upon the assumption of an homogeneous environment can incur a serious loss of performance, see e.g., [6], [7] for thorough theoretical analyses. Despite this performance degradation, there have been very few attempts to design detectors that, from their formulation, take into account the fact that $\boldsymbol{M}_s \neq \boldsymbol{M}_p$. Obviously, prior to that, one must first assume some relation between $\boldsymbol{M}_p$ and $\boldsymbol{M}_s$. Observe that $\boldsymbol{M}_s$ and $\boldsymbol{M}_p$ must be somehow related; otherwise, secondary data would be useless as nothing about $\boldsymbol{M}_p$ could be inferred from observation of the training samples. The most frequently used assumption to depart from an homogeneous environment is to assume that $\boldsymbol{M}_s$ is only proportional to $\boldsymbol{M}_p$; this is often referred to as the partially homogeneous environment. Under this assumption, the adaptive coherence estimator (ACE) is the generalized likelihood ratio test (GLRT) [8], and also the uniformly most powerful invariant test [9]. Note that the ACE was also independently developed in [10] in the case of compound-Gaussian noise, which can also be viewed as an inhomogeneous environment.

The aim of this paper is to present a new approach to model heterogeneous environments, and, accordingly, to derive new detectors based on this model. Towards this end, a Bayesian approach is advocated, in which the covariance matrices $\boldsymbol{M}_p$ and $\boldsymbol{M}_s$ are assumed to be random, with some joint distribution. Moreover, we will assume that some *a priori* information about $\boldsymbol{M}_s$ is available (see below for details). Therefore, our approach enters the framework of knowledge-aided processing [11], which is recognized as one of the potentially most efficient way to handle heterogeneities. Knowledge-aided processing consists of providing conventional adaptive detectors with additional information (such as digital elevation and terrain data, synthetic aperture radar images) so as to improve their performance [12], [13]. This approach was used successfully, e.g., in [14]–[16], where a simplified model for the clutter covariance matrix was used as a good approximation of the actual clutter covariance matrix. Herein, we also assume such knowledge, which is embedded in the *a priori* distribution of the covariance matrix of the secondary data.

II. Problem Statement

In this section, we introduce the detection problem to be solved, as well as the model for heterogeneous environments. The detection problem considered herein is a conventional binary composite hypothesis testing problem, defined as

$$H_0 : \begin{cases} \boldsymbol{z} = \boldsymbol{n}; \\ \boldsymbol{z}_k = \boldsymbol{n}_k; \quad k = 1, \ldots, K \end{cases}$$
$$H_1 : \begin{cases} \boldsymbol{z} = \alpha\boldsymbol{s} + \boldsymbol{n}; \\ \boldsymbol{z}_k = \boldsymbol{n}_k; \qquad k = 1, \ldots, K. \end{cases} \tag{1}$$

Manuscript received June 30, 2006; revised September 6, 2006. This work was supported by the Délégation Générale pour l'Armement (DGA) and by Thales Systèmes Aéroportés. The associate editor coordinating the review of this paper and approving it for publication was Dr. Chong-Yung Chi.

O. Besson and S. Bidon are with the Department of Avionics and Systems, ENSICA, 31056 Toulouse, France (e-mail: besson@ensica.fr; sbidon@ensica.fr)..

J.-Y. Tourneret is with IRIT/ENSEEIHT, 31071 Toulouse, France (e-mail: jean-yves.tourneret@enseeiht.fr).

Color versions of Figs. 1–4 are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2006.888088

In (1), $\boldsymbol{z}$ is the m-length space and/or time snapshot for the CUT, while $\boldsymbol{z}_k$ are the training samples. The $m\times 1$ vector $\boldsymbol{s}$ is the known space and/or time signature of the target and the scalar α stands for its amplitude, which is assumed to be deterministic and unknown.

As for the secondary data, we assume that $\boldsymbol{z}_k$ are proper zero-mean independent and Gaussian distributed noise vectors, with covariance matrix $\boldsymbol{M}_s$. Since the $\boldsymbol{z}_k$'s are independent, the joint density of $\boldsymbol{Z} = [\boldsymbol{z}_1 \quad \cdots \quad \boldsymbol{z}_K]$ is

$$f(\boldsymbol{Z}|\boldsymbol{M}_s) = \pi^{-mK} |\boldsymbol{M}_s|^{-K} \operatorname{etr}\left\{-\boldsymbol{M}_s^{-1}\boldsymbol{S}\right\} \tag{2}$$

where

$$\boldsymbol{S} = \sum_{k=1}^{K} \boldsymbol{z}_k \boldsymbol{z}_k^H \tag{3}$$

denotes the sample covariance matrix of the secondary data, and $\operatorname{etr}\{.\}$ stands for the exponential of the trace of the matrix between braces. Furthermore, we assume that we have some rough knowledge about the average value $\bar{\boldsymbol{M}}_s$ of $\boldsymbol{M}_s$. More precisely, we assume that $\boldsymbol{M}_s$ has an inverse complex Wishart distribution with μ degrees of freedom and mean $\bar{\boldsymbol{M}}_s$ [17]

$$f(\boldsymbol{M}_s) \propto |\boldsymbol{M}_s|^{-(\mu+m)} \operatorname{etr}\left\{-(\mu-m)\boldsymbol{M}_s^{-1}\bar{\boldsymbol{M}}_s\right\} \tag{4}$$

denoted as $\boldsymbol{M}_s \sim \mathcal{CW}_m^{-1}\left((\mu-m)\bar{\boldsymbol{M}}_s, \mu\right)$. Note that (4) is the usual conjugate prior for $\boldsymbol{M}_s$, which will significantly simplify the analysis. A few observations are in order about (4). The matrix $\bar{\boldsymbol{M}}_s$ is the knowledge-aided part of the model and the parameter μ controls the importance of this a priori knowledge. Indeed, as μ increases, the variance of $\boldsymbol{M}_s$ decreases and $\boldsymbol{M}_s$ is closer to $\bar{\boldsymbol{M}}_s$. Therefore, the scalar μ enables us to "tune" the level of *a priori* knowledge we have.

Let us now turn to the assumption about the noise in the primary data. We assume that $\boldsymbol{n}$ is a zero-mean, proper complex-valued Gaussian vector with covariance matrix $\boldsymbol{M}_p$, so that

$$f(\boldsymbol{z}|\boldsymbol{M}_p) = \pi^{-m} |\boldsymbol{M}_p^{-1}| \operatorname{etr}\left\{-\boldsymbol{M}_p^{-1}(\boldsymbol{z}-\bar{\boldsymbol{z}})(\boldsymbol{z}-\bar{\boldsymbol{z}})^H\right\} \tag{5}$$

with $\bar{\boldsymbol{z}} = \boldsymbol{0}$ under H_0 and $\bar{\boldsymbol{z}} = \alpha\boldsymbol{s}$ under H_1. In order to account for an heterogeneous environment, we assume that $\boldsymbol{M}_p$ given $\boldsymbol{M}_s$ has a complex Wishart distribution, with ν degrees of freedom, whose mean is $\boldsymbol{M}_s$, i.e.,

$$f(\boldsymbol{M}_p|\boldsymbol{M}_s) \propto |\boldsymbol{M}_p|^{\nu-m} |\boldsymbol{M}_s|^{-\nu} \operatorname{etr}\left\{-\nu\boldsymbol{M}_p\boldsymbol{M}_s^{-1}\right\}. \tag{6}$$

This distribution will be denoted as $\boldsymbol{M}_p|\boldsymbol{M}_s \sim \mathcal{CW}_m\left(\nu^{-1}\boldsymbol{M}_s, \nu\right)$. The interpretation of (6) is as follows. On the average, $\mathcal{E}\{\boldsymbol{M}_p|\boldsymbol{M}_s\} = \boldsymbol{M}_s$, which means that the two covariance matrices are not too far from one another. The scalar ν controls the distance between the two matrices $\boldsymbol{M}_p$ and $\boldsymbol{M}_s$; as ν increases, this distance decreases. Note however that $\boldsymbol{M}_p \neq \boldsymbol{M}_s$ with probability one, and hence the environment is heterogeneous. Therefore, the proposed model enables us, in a theoretically sound and rather flexible way, to account for nonhomogeneous environments.

III. Detection

As previously explained, we consider the problem of deciding between H_0 and H_1 in (1), under the following assumptions:

$$\boldsymbol{z}|\boldsymbol{M}_p \sim \mathcal{CN}_m(\bar{\boldsymbol{z}}, \boldsymbol{M}_p) \tag{7a}$$

$$\boldsymbol{z}_k|\boldsymbol{M}_s \sim \mathcal{CN}_m(\boldsymbol{0}, \boldsymbol{M}_s) \tag{7b}$$

$$\boldsymbol{M}_p|\boldsymbol{M}_s \sim \mathcal{CW}_m\left(\nu^{-1}\boldsymbol{M}_s, \nu\right) \tag{7c}$$

$$\boldsymbol{M}_s \sim \mathcal{CW}_m^{-1}\left((\mu-m)\bar{\boldsymbol{M}}_s, \mu\right). \tag{7d}$$

When $\boldsymbol{M}_p$ is known, the GLRT consists of comparing

$$\frac{|\boldsymbol{s}^H\boldsymbol{M}_p^{-1}\boldsymbol{z}|^2}{\boldsymbol{s}^H\boldsymbol{M}_p^{-1}\boldsymbol{s}} \tag{8}$$

to an appropriate threshold depending on the probability of false alarm. In an homogeneous environment, the AMF consists of replacing $\boldsymbol{M}_p$ above by its maximum likelihood estimate based on $\boldsymbol{Z}$, namely $K^{-1}\boldsymbol{S}$. We take a similar route here, except that a Bayesian framework is used to estimate $\boldsymbol{M}_p$ differently. More precisely, as it provides a closed-form and simple estimator, we derive the maximum *a posteriori* (MAP) estimate of $\boldsymbol{M}_p$ using $\boldsymbol{Z}$. In order to obtain the latter, the a posteriori distribution $f(\boldsymbol{M}_p|\boldsymbol{Z})$ must be derived. Observe first that

$$\begin{aligned} &f(\boldsymbol{M}_p, \boldsymbol{M}_s|\boldsymbol{Z}) \\ &\propto f(\boldsymbol{Z}|\boldsymbol{M}_p, \boldsymbol{M}_s) f(\boldsymbol{M}_p|\boldsymbol{M}_s) f(\boldsymbol{M}_s) \\ &\propto |\boldsymbol{M}_s|^{-K} \operatorname{etr}\left\{-\boldsymbol{M}_s^{-1}\boldsymbol{S}\right\} |\boldsymbol{M}_p|^{\nu-m} |\boldsymbol{M}_s|^{-\nu} \operatorname{etr}\left\{-\nu\boldsymbol{M}_p\boldsymbol{M}_s^{-1}\right\} \\ &\quad \times |\boldsymbol{M}_s|^{-(\mu+m)} \operatorname{etr}\left\{-(\mu-m)\boldsymbol{M}_s^{-1}\bar{\boldsymbol{M}}_s\right\} \\ &\propto |\boldsymbol{M}_s|^{-(\nu+\mu+m+K)} |\boldsymbol{M}_p|^{\nu-m} \operatorname{etr}\left\{-\boldsymbol{M}_s^{-1}\boldsymbol{B}\right\} \end{aligned} \tag{9}$$

where

$$\boldsymbol{B} = \nu\boldsymbol{M}_p + \boldsymbol{S} + (\mu-m)\bar{\boldsymbol{M}}_s. \tag{10}$$

The *a posteriori* distribution of $\boldsymbol{M}_p$ given $\boldsymbol{Z}$ can thus be written as

$$\begin{aligned} f(\boldsymbol{M}_p|\boldsymbol{Z}) &= \int f(\boldsymbol{M}_p, \boldsymbol{M}_s|\boldsymbol{Z})\, d\boldsymbol{M}_s \\ &= C\, \frac{|\boldsymbol{M}_p|^{\nu-m}}{|\boldsymbol{B}|^{\nu+\mu+K}} \end{aligned} \tag{11}$$

where the constant C is such that $\int f(\boldsymbol{M}_p|\boldsymbol{Z})\, d\boldsymbol{M}_p = 1$. Taking the logarithm of (11), we have

$$\Lambda(\boldsymbol{M}_p|\boldsymbol{Z}) = \text{const.} + (\nu-m)\ln|\boldsymbol{M}_p| - (\nu+\mu+K)\ln|\boldsymbol{B}|. \tag{12}$$

Differentiating the previous equations yields

$$\frac{\partial\Lambda(\boldsymbol{M}_p|\boldsymbol{Z})}{\partial\boldsymbol{M}_p} = (\nu-m)\boldsymbol{M}_p^{-1} - \nu(\nu+\mu+K)\boldsymbol{B}^{-1}. \tag{13}$$

Equating this derivative to zero allows us to show that the MAP estimate of $\boldsymbol{M}_p$ is

$$\widehat{\boldsymbol{M}}_p = \frac{\nu-m}{\nu(\mu+m+K)}\left[\boldsymbol{S} + (\mu-m)\bar{\boldsymbol{M}}_s\right]. \tag{14}$$

We would like to stress that the MAP estimator can be obtained in closed-form and is thus a simple estimator. It is also interesting to note that the MAP estimate is a combination of the a priori information and the information provided by the secondary data. Observe that, when μ increases, more importance is allotted to $\bar{\boldsymbol{M}}_s$, as can be expected. Once the MAP estimate is obtained, it is used to replace $\boldsymbol{M}_p$ in (8), which yields our final detector

$$\frac{|\boldsymbol{s}^H \widehat{\boldsymbol{M}}_p^{-1} \boldsymbol{z}|^2}{\boldsymbol{s}^H \widehat{\boldsymbol{M}}_p^{-1} \boldsymbol{s}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \tag{15}$$

Before closing this section, some remarks are in order.

Remark 1: In what precedes, a MAP approach was advocated as it results in a simple estimate of $\boldsymbol{M}_p$. However, other approaches could be investigated, including the minimum mean-square error (MMSE) estimate of $\boldsymbol{M}_p$. Using (11), the MMSE estimate of $\boldsymbol{M}_p$ is given by

$$\mathcal{E}\left\{\boldsymbol{M}_p|\boldsymbol{Z}\right\} = \frac{\int |\boldsymbol{M}_p|^{\nu-m} |\boldsymbol{B}|^{-(\nu+\mu+K)} \boldsymbol{M}_p \, d\boldsymbol{M}_p}{\int |\boldsymbol{M}_p|^{\nu-m} |\boldsymbol{B}|^{-(\nu+\mu+K)} \, d\boldsymbol{M}_p} \tag{16}$$

where $\mathcal{E}\{.\}$ is the mathematical expectation. Unfortunately, there does not exist any closed-form expression for the above integral, and the MMSE estimate cannot be obtained analytically. In such a case, it is usual to resort to stochastic integration methods such as Markov chain Monte Carlo (MCMC) methods. These methods consist of generating samples distributed according to the posterior distribution $f(\boldsymbol{M}_p|\boldsymbol{Z})$, and to use these samples to approximate the integrals to be computed [18]. However, generating matrices distributed according to (11) is not obvious. An alternative consists of generating matrices $\boldsymbol{M}_p$ and $\boldsymbol{M}_s$ distributed according to the joint distribution $f(\boldsymbol{M}_p, \boldsymbol{M}_s|\boldsymbol{Z})$. This can be achieved by using a Gibbs sampling strategy, generating iteratively matrices $\boldsymbol{M}_p$ and $\boldsymbol{M}_s$ as follows:

$$\boldsymbol{M}_p|\boldsymbol{M}_s, Z \sim \mathcal{CW}_m\left(\nu^{-1}\bar{\boldsymbol{M}}_s, \nu\right) \tag{17a}$$

$$\boldsymbol{M}_s|\boldsymbol{M}_p, Z \sim \mathcal{CW}_m^{-1}\left(\boldsymbol{B}, \nu+\mu+K\right) \tag{17b}$$

where, to obtain the previous distributions, we made use of (9). It is known [18, p. 325] that the matrices $(\boldsymbol{M}_p, \boldsymbol{M}_s)$ generated in this manner are asymptotically distributed according to $f(\boldsymbol{M}_p, \boldsymbol{M}_s|\boldsymbol{Z})$. Therefore, the MMSE estimate can be obtained by averaging the last matrices generated by the Gibbs sampler. However, this approach results in a significantly increased computational complexity. Moreover, we experimented that replacing the MAP estimate by the MMSE estimate in (15) does not result in any improvement in the detection performance. Therefore, only the MAP estimate, due to its simplicity, is retained.

Remark 2: It is interesting to note that the test statistic in (15) is, up to a scaling factor, the output power of a beamformer which corresponds to colored loading of the sample covariance matrix $\boldsymbol{S}$. The loading matrix is $\bar{\boldsymbol{M}}_s$ and the loading level depends on μ, which controls the degree of a priori knowledge. Interestingly, it has the same form as the beamformer designed in [14], [15] in a rather different framework. Indeed, in [14] and [15], a beamformer $\boldsymbol{w}$ is designed under the quadratic constraint that the output power corresponding to the *a priori* clutter covariance matrix, *viz* $\boldsymbol{w}^H \bar{\boldsymbol{M}}_s \boldsymbol{w}$, is small. It turns out that the solution takes the form of colored loading of the sample covariance matrix, with a loading matrix proportional to $\bar{\boldsymbol{M}}_s$. The loading level is chosen so as to satisfy the quadratic constraint. Hence, it has exactly the same form as here, despite the fact that the statistical assumptions are very different. These coincident conclusions suggest that colored loading may be an effective means to incorporate *a priori* knowledge, and to be robust to uncertain environments.

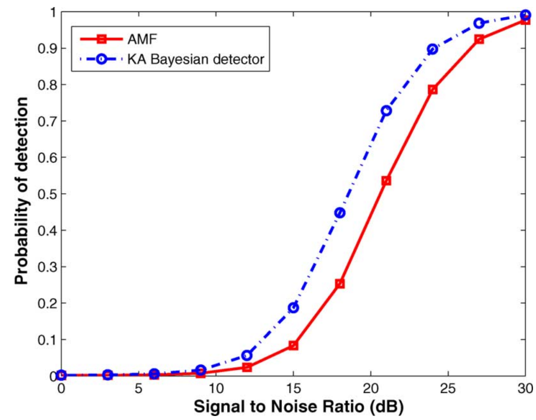

Fig. 1. Probability of detection versus SNR. $\nu = m+1$ and $\mu = m+1$.

IV. Numerical Simulations

In this section, we illustrate the performance of our detector and compare it with that of the conventional adaptive matched filter [3]. In all simulations, we consider an array with $m = 8$ elements, and the signature of the signal of interest is $\boldsymbol{s} = [1 \quad 1 \quad \cdots \quad 1]^T$. The average secondary data covariance matrix is $\bar{\boldsymbol{M}}_s(k,\ell) = 0.9^{|k-\ell|}$. The number of training samples is set to $K = 16$. In all simulations below, the probability of false alarm is set to $P_{fa} = 10^{-3}$. For every simulation, a different (random) matrix $\boldsymbol{M}_s$ is drawn from (4). Then, using this value of $\boldsymbol{M}_s$, a matrix $\boldsymbol{M}_p$ is drawn from (6). Hence the two matrices are different in each run, and therefore the environment is heterogeneous. The thresholds for each detector were obtained from 200 000 simulations. The probability of detection P_d was computed from 100 000 simulations. P_d is plotted as a function of the signal-to-noise ratio, which is defined as $SNR = |\alpha|^2 \boldsymbol{s}^H \boldsymbol{M}_p^{-1} \boldsymbol{s}$.

Figs. 1–4 display the results obtained, with different values of ν and μ. As can observed from these figures, the new detector always improves over the conventional AMF; the improvement is between 2.1 and 2.7 dB for $P_d = 0.7$. This improvement is slightly more pronounced as μ increases, i.e., as the *a priori* knowledge is more significant. In contrast, for fixed μ, the difference between the new detector and the AMF remains nearly constant when ν varies. However, ν has a significant impact on the probability of detection; there is about 5-dB difference between $\nu = m+1$ and $\nu = 2m$. This clearly indicates the effect of nonhomogeneity on the performance of the detectors.

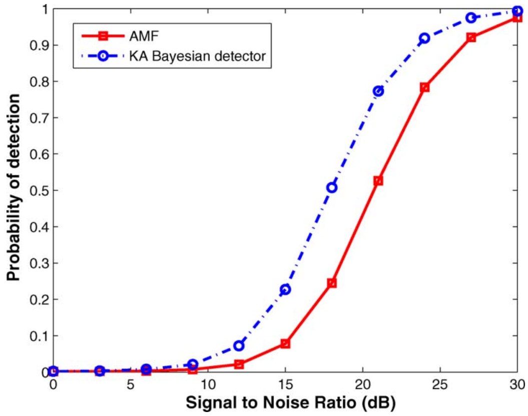

Fig. 2. Probability of detection versus SNR. $\nu = m + 1$ and $\mu = 2m$.

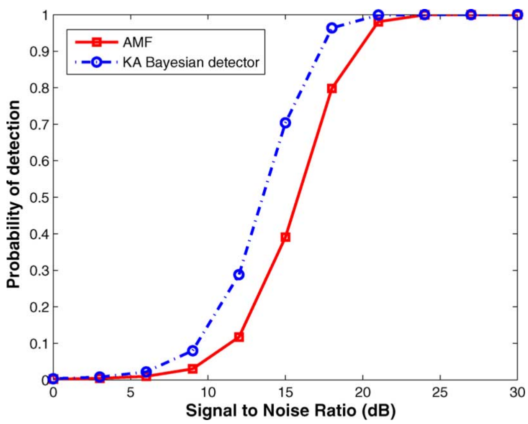

Fig. 3. Probability of detection versus SNR. $\nu = 2m$ and $\mu = m + 1$.

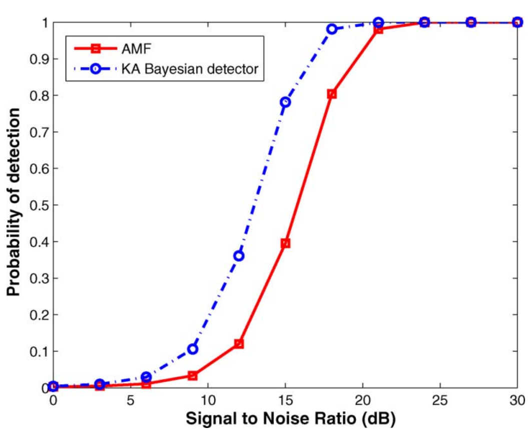

Fig. 4. Probability of detection versus SNR. $\nu = 2m$ and $\mu = 2m$.

V. Conclusions

We introduced a new Bayesian framework for knowledge-aided adaptive detection in nonhomogeneous environments. More precisely, we assumed that the covariance matrices of the primary and secondary data were random, with some joint distribution, and that the average value of the secondary data covariance matrix was known. The distance between the two matrices as well as the importance of the *a priori* knowledge can be tuned through scalar variables. Under this model, a Bayesian version of the adaptive matched filter was derived, where the MAP estimate of the primary data covariance matrix is used in place of the maximum likelihood estimate. It was shown that the new detector amounts to colored loading of the sample covariance matrix. Numerical simulations illustrated the improvement achieved via this knowledge-aided Bayesian detector, as compared to the conventional AMF.

References

[1] A. Farina, *Antenna Based Signal Processing Techniques for Radar Systems*. Norwell, MA: Artech House, 1992.

[2] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation and Time Series Analysis*. Reading, MA: Addison–Wesley, 1991.

[3] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 1, pp. 208–216, Jan. 1992.

[4] R. Klemm, *Principles of Space-Time Adaptive Processing*. London, U.K.: IEE, 2002, vol. 12, IEE Radar, Sonar, Navigation and Avionics Series.

[5] W. L. Melvin, "Space-time adaptive radar performance in heterogeneous clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 2, pp. 621–633, Apr. 2000.

[6] C. D. Richmond, "Performance of a class of adaptive detection algorithms in nonhomogeneous environments," *IEEE Trans. Signal Process.*, vol. 48, no. 5, pp. 1248–1262, May 2000.

[7] K. F. McDonald and R. S. Blum, "Exact performance of STAP algorithms with mismatched steering and clutter statistics," *IEEE Trans. Signal Process.*, vol. 48, no. 10, pp. 2750–2763, Oct. 2000.

[8] S. Kraut and L. L. Scharf, "The CFAR adaptive subspace detector is a scale-invariant GLRT," *IEEE Trans. Signal Process.*, vol. 47, no. 9, pp. 2538–2541, Sep. 1999.

[9] S. Kraut, L. L. Scharf, and R. W. Butler, "The adaptive coherence estimator: A uniformly most powerful invariant adaptive detection statistic," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 427–438, Feb. 2005.

[10] E. Conte, M. Lops, and G. Ricci, "Asymptotically optimum radar detection in compound-Gaussian clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 2, pp. 617–625, Apr. 1995.

[11] "Knowledge-based systems for adaptive radar: Detection, tracking and classification," *Signal Process. Mag. (Special Issue)*, Jan. 2006.

[12] J. R. Guerci and E. J. Baranoski, "Knowledge-aided adaptive radar at DARPA – An overview," *IEEE Signal Process. Mag.*, vol. 23, no. 1, pp. 41–50, Jan. 2006.

[13] in *Proc. KASSPER Workshops*, 2002–2005.

[14] J. S. Bergin, C. M. Teixeira, P. M. Techau, and J. R. Guerci, "Space-time beamforming with knowledge-aided constraints," in *Proc. 11th Adaptive Sensor Array Processing Workshop*, Lexington, MA, Mar. 11–12, 2003.

[15] ——, "STAP with knowledge-aided data pre-whitening," in *Proc. IEEE Radar Conf.*, Philadelphia, PA, Apr. 26–29, 2004, pp. 289–294.

[16] K. Gerlach and M. L. Picciolo, "Airborne/spacebased radar STAP using a structured covariance matrix," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 1, pp. 269–281, Jan. 2003.

[17] J. A. Tague and C. I. Caldwell, "Expectations of useful complex Wishart forms," *Multidim. Syst. Signal Process.*, vol. 5, pp. 263–279, 1994.

[18] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*. London, U.K.: Chapman & Hall, 1995.