Modified GLRT and AMF Framework for Adaptive Detectors

YURI I. ABRAMOVICH, Senior Member, IEEE
DSTO
Australia

NICHOLAS K. SPENCER
ARI Pty Ltd.
Australia

ALEXEI Y. GOROKHOV, Member, IEEE
Qualcomm, Inc.

The well-known general problem of signal detection in background interference is addressed for situations where a certain statistical description of the interference is unavailable, but is replaced by the observation of some secondary (training) data that contains only the interference. For the broad class of interferences that have a large separation between signal- and noise-subspace eigenvalues, we demonstrate that adaptive detectors which use a diagonally loaded sample covariance matrix or a fast maximum likelihood (FML) estimate have significantly better detection performance than the traditional generalized likelihood ratio test (GLRT) and adaptive matched filter (AMF) detection techniques, which use a maximum likelihood (ML) covariance matrix estimate. To devise a theoretical framework that can generate similarly efficient detectors, two major modifications are proposed for Kelly's traditional GLRT and AMF detection techniques. First, a two-set GLRT decision rule takes advantage of an a priori assignment of different functions to the primary and secondary data, unlike the Kelly rule that was derived without this. Second, instead of ML estimates of the missing parameters in both GLRT and AMF detectors, we adopt expected likelihood (EL) estimates that have a likelihood within the range of most probable values generated by the actual interference covariance matrix. A Gaussian model of fluctuating target signal and interference is used in this study. We demonstrate that, even under the most favorable loaded sample-matrix inversion (LSMI) conditions, the theoretically derived EL-GLRT and EL-AMF techniques (where the loading factor is chosen from the training data using the EL matching principle) gives the same detection performance as the loaded AMF technique with a proper a priori data-invariant loading factor. For the least favorable conditions, our EL-AMF method is still superior to the standard AMF detector, and may be interpreted as an intelligent (data-dependent) method for selecting the loading factor.

Manuscript received January 31, 2005; released for publication March 18, 2006.

IEEE Log No. T-AES/43/3/908408.

Refereeing of this contribution was handled by E. Chornoboy.

Authors' current addresses: Y. I. Abramovich, Intelligence, Surveillance and Reconnaissance Division, Defence Science and Technology Organisation (DSTO), Edinburgh SA 5111, Australia, E-mail: (yuri.abramovich@dsto.defence.gov.au); N. K. Spencer, Adelaide Research & Innovation Pty Ltd. (ARI), C/-200 Labs, ISRD, DSTO, Edinburgh SA 5111, Australia, E-mail: (Nick.Spencer@adelaide.edu.au); A. Y. Gorokhov, QT & V, Qualcomm, Inc., 5775 Morehouse Dr., San Diego, CA 92121-1714, E-mail: (gorokhov@qualcomm.com).

I. INTRODUCTION

Techniques for adaptive signal processing for radar target detection in an unknown interference environment stem from the pioneering work of Reed, Mallet and Brennan (RMB) [1], followed by Kelly's seminal paper [2], and now embrace various scenarios with both Gaussian and non-Gaussian (spherically invariant random process (SIRP)) interference [3–5]. Depending on the practical application, two rather different formulations are considered for the adaptive detection problem.

The first, addressed in RMB [1], is considered when the secondary (training) data is used to design the adaptive filter that is then used to process the entire set of primary data (range cells, say). For example, in adaptive antenna external-noise suppression applications, a limited number of range cells (or even an "inter-dwell gap" that is free of clutter and targets) is typically used to estimate the external-noise covariance matrix and design the adaptive antenna. Since the external noise is supposed to be homogeneous over the set of range cells, this adaptive antenna weight vector is then used to process all operational range cells that usually contain clutter and possible targets. Final target detection is performed downstream after clutter suppression (moving-target indicator or Doppler filtering) and noncoherent integration. Adaptive threshold calculation is done at the output of this signal processing chain using primary range cells. This (adaptive) detection threshold is therefore calculated for this specific (adaptive) antenna weight vector; for Gaussian primary data, the antenna output is also Gaussian with, possibly, unknown output noise power. For this (conditional) Gaussian model, a typical way to calculate adaptive thresholds was introduced by Finn and Johnson [6]. In such schemes, detection performance degradations with respect to the clairvoyant (exact) detector are due to two statistically independent factors. One factor is associated with the adaptive threshold calculations (using primary data), while the other is associated with the signal-to-noise ratio (SNR) degradation in the adaptive antenna (using the secondary data). Therefore, for such applications, analysis of the SNR probability density function (pdf) at the output of the adaptive antenna gives a complete description of the adaptive antenna (filter) design performance. Specifically, receiver operational curves (ROCs), derived for example in Finn and Johnson's paper [6] for a particular adaptive threshold design and a (fluctuating) target with a certain SNR, must now be additionally averaged over the SNR pdf for the given adaptive antenna algorithm. In RMB [1], this pdf for the normalized output SNR was accurately calculated for the case when the unconditional maximum likelihood (ML) covariance matrix estimate (direct sample covariance matrix) is used for adaptive

antenna filter design, drawn from the training (secondary) data. This famous β-distribution has an extremely important invariance property with respect to the observed scenario, namely, it is fully specified by only two parameters: the training sample size (N) and the adaptive antenna (filter) dimension (M). The sample-support requirement

$$N \simeq 2M \tag{1}$$

that ensures approximately 3 dB average SNR losses compared with the clairvoyant solution has become the most quoted requirement in studies on adaptive filters.

Meanwhile, since the early 1980s, considerable research was focused on a specific class of interferences whose covariance matrix has a distinct difference between the size of the m ($< M$) signal-subspace eigenvalues and the $n \equiv M - m$ noise-subspace eigenvalues. The prototypical model that results in this covariance matrix structure is a mixture of m powerful external-noise point sources with (internal) white noise. The minimum eigenvalue here is equal to the noise power, while the sum of the signal eigenvalues is almost equal to the overall power of the external interferences.

Interference-to-noise ratio (INR) values of 20 to 40 dB are not uncommon for practical adaptive antenna applications. In fact, adaptive filter (antenna) design is efficient only for scenarios with a significant ratio of maximum to minimum eigenvalue (λ_1/λ_M). Indeed, if S is the M-variate normalized useful signal (target) array-signal manifold ("steering") vector, and R is the M-variate interference covariance matrix, then the clairvoyant optimum filter

$$W_{\text{opt}} \equiv R^{-1}S, \qquad S^H S = 1 \tag{2}$$

has an SNR improvement over the "white-noise optimum" filter $W_{\text{wn}} \equiv S$ of

$$\eta \equiv (S^H R^{-1} S)(S^H R S). \tag{3}$$

According to the so-called Kantorovich inequality [7]

$$(S^H R^{-1} S)(S^H R S) \leq \frac{(\lambda_1 + \lambda_M)^2}{4\lambda_1 \lambda_M} \qquad \text{if} \quad S^H S = 1. \tag{4}$$

Even an improvement of, say, $\eta = 10$ means that $\lambda_1/\lambda_M \gtrsim 40$ for the interference covariance matrix R.

For the subclass of interferences, whose eigenvalues (sorted in descending order) are such that

$$\lambda_1 > \cdots > \lambda_m \gg \lambda_{m+1} = \cdots \lambda_M \tag{5}$$

it is known that diagonal loading of the direct sample covariance matrix leads to a quite dramatic SNR improvement over the ML covariance matrix estimate considered by RMB [1]. This improvement was described in [8, 9]. In [10, 11], quite accurate analytic derivations showed that for these scenarios the normalized output SNR does not depend on the loading factor, provided it is chosen to be within a certain range. The SNR loss factor, that in RMB was a function of M and N only, was then specified by m and N, irrespective of M.

Specifically, the RMB requirement (1) for the diagonally loaded sample-matrix inversion (LSMI) algorithm was supplanted by the condition

$$N \simeq 2m \tag{6}$$

which means a significant performance improvement for scenarios with $m \ll M$. Since the early 1980s, the properties of the LSMI algorithm have been widely explored and validated in practice for various radar and sonar applications [12, 13]. It is known that diagonal loading offers many other important features that make adaptive antennas robust against numerous inaccuracies in the scenario model [14–16].

The superior SNR performance of LSMI over the RMB sample-matrix inversion (SMI) technique raises concerns about the optimality of the ML criterion for covariance matrix estimation in adaptive antenna/filter design. Nevertheless, the same ML principle was exploited by Kelly in [2], and subsequently by Robey, Fuhrmann, Kelly and Nitzberg (RFKN) in [17], when the significantly different problem of adaptive detection was addressed. This is the second formulation we consider for the adaptive detection problem. This problem is formulated as a hypothesis test, where a decision regarding the presence or absence of a target in a single primary datum (one range cell, say) is made based only on this datum and the auxiliary training (secondary) data. In essence, the training data here must be used to provide all missing information required for decision making, including adaptive antenna/filter design and adaptive thresholding. Since the same training data is used for these two purposes, SNR losses (as in RMB, for example) and adaptive threshold losses (as in Finn) cannot be treated as independent random values, and specific analysis of adaptive detection ROCs is required. This was performed by Kelly in [2] for the generalized likelihood ratio test (GLRT) detector and by RFKN in [17] for the "constant false-alarm rate (CFAR) adaptive matched-filter" detector. In practice, this formulation is satisfactory for scenarios where each primary datum (range cell) is processed by its own (independent) adaptive filter, and the detection threshold must be calculated over the set of all (random) adaptive filters. A typical example is "sliding-window" adaptive processing where a small number of "guard cells" is used for adaptive detection within a specific range cell, and so each range cell is processed by a different adaptive filter with a different set of training guard cells involved. In such cases, the invariance of the output detector statistics for target-free primary data (CFAR) is required for the direct implementation of this approach. Indeed, when

the pdf of the scenario-free output statistics is known, detection can be completed without adaptive threshold calculations (over a set of homogeneous range cells).

Both Kelly's GLRT detector and the CFAR adaptive matched filter (AMF) detector [17] that use the ML covariance matrix estimate enjoy this important invariance property. Due to its importance, the invariance (CFAR) property became a necessary condition for adaptive detector design in one of the modern adaptive detector developments of Scharf, et al. in [18–20]. Nevertheless, for some models invariant solutions may not exist (e.g., adaptive detection for different interference-to-noise ratios (INRs) in primary and secondary data [21]), or come at an excessive performance cost. It is important to emphasize that when this invariance cannot be assured, single primary data detection is strictly impossible. However in practice, a number of homogeneous ranges cells (even processed by different but statistically identical filters) can often be specified and again used for adaptive threshold calculations. Naturally, the additional losses of this method must be taken into consideration.

Various modifications of the adaptive detection problem have been considered since Kelly's paper [4, 5, 22, 3, 23, 24]. However there is frequently confusion of these two quite different problems: adaptive filtering (the same filter for all primary data) and adaptive detection (a different filter for each primary datum). In [25], Lekhovitskii demonstrates that the application of nonnormalized adaptive filters (in fact, randomly normalized) $W_j = \hat{R}_j^{-1} S$ with independent $\hat{R}_j$ for each cell leads to a dramatic degradation in detection performance. Indeed, a single threshold for the various $\hat{R}_j$ has to account for significant output power fluctuations that are due to variations in the norm of the adaptive filter. Clearly this problem does not exist for single adaptive antenna applications. Confusion also occurs when the famous RMB 3 dB SNR average loss factor for $N \simeq 2M$ with the ML covariance matrix estimate is compared with, say, the 6 dB SNR degradation in ROC of the CFAR AMF detector which uses the same estimate [17].

In this paper, we consider the adaptive detection problem in Kelly's context where each range cell is processed by an individually tailored filter, and the (unconditional) output statistics are averaged over a (random) set of adaptive solutions. Our goal is two-fold. First, based on the well-known superior properties of the LSMI algorithm for adaptive filter performance, we want to investigate the class of LSMI or loaded AMF (LAMF) adaptive detectors. We hope and expect that the superiority of LSMI adaptive filters will lead to a superiority of LAMF adaptive detectors for the class of interference scenarios (5). As we shall see, this expectation turns out to be true. The other properties of LAMF detectors (such as CFAR) also need to be specified. Second, this superiority of LAMF over ML-based AMF means that the theoretical framework for adaptive detection (GLRT and AMF) needs to be modified so as to include LAMF detectors. Specifically, we expect that a modified framework would either directly lead to LAMF detectors with some theoretically specified loading factor, or produce different adaptive detectors whose performance is at least as good as the LAMF performance for the class of interference scenarios (5).

In accordance with most papers on adaptive detection [1, 2], we assume that the training data contains only interference that is statistically partly or completely the same as the interference within the primary data. In what follows, we consider the cases where the interferences are exactly the same (homogeneous), or just have a different power (nonhomogeneous). This model comprises "supervised training conditions" since the secondary data does not contain targets or other interference sources. Of course, for "sliding training-window" implementations it is impossible to exclude the possibility that a target or other inhomogeneity is present in the training data. Indeed, the selection of a homogeneous training dataset (inhomogeneity detection) is an important research topic in adaptive radar studies [26–30]. However, here we reconsider the traditional GLRT and AMF solutions under their original assumptions. Recall that the AMF detector in RFKN [17] is derived as a GLRT where the covariance matrix is known. After the test statistics are derived, the ML estimate of the covariance matrix based on the secondary data is substituted for the known covariance matrix. Kelly's GLRT method is to treat the primary and secondary data as a single dataset, then for both of the hypotheses regarding the presence or absence of a target signal, ML estimates of all parameters (including the interference covariance matrix) are used to construct the decision rule. Therefore, both the AMF and Kelly GLRT techniques heavily rely on the same ML principle for interference parameter estimation.

In this regard, it is quite instructive to recall the discussion in RFKN [17] comparing the performance of AMF and GLRT detectors. The AMF detection rule was considered to be inferior to GLRT simply because "the AMF test makes no use of the primary vector to estimate the covariance, therefore poorer detection performance might be expected" compared with the GLRT that "uses all the data (primary and secondary) in the likelihood optimization under each hypothesis." This argument could have been easily justified by comparing the detection performance of Kelly's GLRT technique with N secondary samples against the AMF performance with just $(N + 1)$ secondary samples. If those expectations were correct, then the $(N + 1)$-variate AMF should always outperform the N-variate Kelly GLRT, which is not the case.

Moreover, Kelly's GLRT is effectively based on two ML interference covariance matrix estimates constructed for the hypotheses H_0 and H_1. Specifically [2],

$$\text{const}\ \hat{R} \mid H_0 = X_N X_N^H + YY^H \tag{7}$$

where $X_N \equiv [\mathbf{x}_1,\ldots,\mathbf{x}_N] \in \mathcal{C}^{M\times N} \sim \mathcal{CN}_N(0,R_0)$ is the N-sample secondary (training) data with independent identically distributed (IID) samples $\mathbf{x}_j$ $(j = 1,\ldots,N)$ described by the M-variate complex (circular) Gaussian distribution with covariance matrix R_0, and

$$\text{const}\ \hat{R} \mid H_1 = X_N X_N^H + \left[I_M - \frac{SS^H(X_N X_N^H)^{-1}}{S^H(X_N X_N^H)^{-1}S}\right] YY^H \times \left[I_M - \frac{(X_N X_N^H)^{-1}SS^H}{S^H(X_N X_N^H)^{-1}S}\right]. \tag{8}$$

The primary sample is

$$Y = \begin{cases} X_0 \sim \mathcal{CN}(0,R_0) & \text{for} \quad H_0 \\ X_0 + aS & \text{for} \quad H_1 \end{cases} \tag{9}$$

where $S \in \mathcal{C}^{M\times 1}$ is the target wavefront, and a is the unknown (complex) target amplitude. We can see that for H_1, the primary sample contribution to the covariance matrix estimate in (8) does not retain any interference component that corresponds to the target signal wavefront. This component is rejected by the projection matrix

$$\left[I_M - \frac{SS^H(X_N X_N^H)^{-1}}{S^H(X_N X_N^H)^{-1}S}\right] \tag{10}$$

along with the possible target signal. Since only this component is essential for target detection, the primary sample-produced covariance matrix update in (8) with this component canceled out is nonsensical. As a result, in some situations GLRT was found to be superior to AMF, and inferior in others. Therefore in RFKN [17], the authors rightly conceded that "the generalized likelihood-ratio test is not optimal in the Neyman-Pearson sense as the AMF test has a probability of detection that is higher than that of the GLRT for some situations." For the same reason, AMF is not optimal either. With LAMF and its non-ML covariance matrix estimate, that for scenarios (5) have better detection performance, it became clear that the ML estimation principle needs to be reexamined.

Note that diagonal loading (as well as the fast ML (FML) technique [30]) are quite different from the numerous methods that assume a restricted class of admissible covariance matrices, but still use the ML criterion. Of course, if properly adopted, any valid a priori information on the interference properties that somehow restricts the (ML) search should lead to improved adaptive detection performance, though the CFAR property may not be so easy to maintain. A typical example is restricting to the class of Toeplitz covariance matrices for uniform antenna arrays (or pulse trains). On the contrary, it is straightforward to demonstrate that the ML optimization restricted to the set of diagonally loaded or finite-subspace (FML) covariance matrices results in zero loading and maximal signal subspace, which drives the optimum solution to the same unconstrained (ML) sample covariance matrix estimate with ML.

Hence the main important question is whether the ML principle within the GLRT and AMF framework can be replaced by another general principle that will generate adaptive detectors that are at least not inferior to (say) LAMF for the most favorable scenarios. While these considerations stem mostly from the known discrepancy between SMI and LSMI performance, there are also important theoretical considerations that raise concerns regarding the ML criterion for small sample sizes (which are typical in radar applications).

In this regard, let us consider the traditional likelihood function (LF) for the covariance matrix given $N > M$ IID training samples X_N as in (7) [31]

$$f(R \mid X_N) = \frac{\text{const}}{\det^N R}\exp[-\text{tr}(R^{-1}X_N X_N^H)] \tag{11}$$

then with probability one we have

$$\det(X_N X_N^H) \neq 0 \tag{12}$$

and so the likelihood ratio (LR) [32]

$$LR(R,X_N) \equiv [f(R \mid X_N)\det(X_N X_N^H)]^{1/N} = \frac{\det(R^{-1}\hat{R})\exp M}{\exp[\text{tr}(R^{-1}\hat{R}_N)]} \tag{13}$$

where

$$\hat{R} \equiv \frac{1}{N}X_N X_N^H \tag{14}$$

can also be treated as the LF with respect to R (e.g., [31]). We can see that the (unconstrained) ML solution

$$R_{\text{ML}} = \hat{R} \tag{15}$$

derived by Anderson [33] by direct maximization of the LF $f(R \mid X_N)$ yields the ultimate value for the LR of unity:

$$\max_R LR(R,X_N) = LR(\hat{R},X_N) = 1 \tag{16}$$

irrespective of the sample support N and filter dimension M.

At the same time, for the true (exact) covariance matrix R_0, the pdf of the LR

$$LR(R_0,X_N) = \frac{\det(R_0^{-1/2}\hat{R}R_0^{-1/2})\exp M}{\exp[\text{tr}(R_0^{-1/2}\hat{R}R_0^{-1/2})]} \tag{17}$$

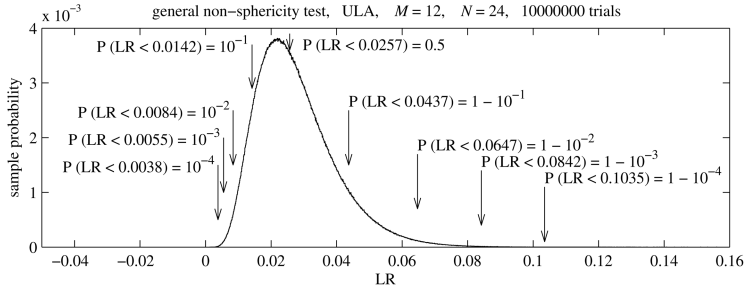


Fig. 1. Sample pdfs for the general test.

does not depend on R, since

$$\hat{C} \equiv R_0^{-1/2} \hat{R} R_0^{-1/2} \sim \mathcal{CW}(N, M, I_M) \tag{18}$$

where $\hat{C}$ is the "white-noise" sample matrix with complex Wishart distribution, is (for $N > M$) specified by two parameters only, namely N and M. Later we introduce moments and series representations for the pdf of this ratio derived for the complex-valued case similarly to the real-valued case in [34]. For the sake of our current discussion, we refer to Fig. 1 where this pdf is shown for the values $M = 12$ and $N = 24$.

Note that according to RMB [1], the average SNR degradation for SMI is about 3 dB here, which is often considered acceptable in practical applications. Yet Fig. 1 shows that there is a huge distinction between the LR values generated by the true covariance matrix R_0 and the ultimate value $LR = 1$ produced by the ML estimate $R_{\mathrm{ML}} = \hat{R}$. Indeed, the median LR is only 0.0257, and with probability 99.99% the LR is less than 0.1051! Based on this pdf, it seems natural to replace the ML estimate by one that generates LR values consistent with what is expected for the true covariance matrix. We call this approach the expected likelihood (EL) estimation criterion and will demonstrate that, unlike the ML criterion, this approach inherently justifies the appropriate selection of parameters (such as loading factor and interference signal subspace dimension) based on direct likelihood matching, rather than on "external considerations" as suggested in [12], [13]. Moreover, we demonstrate that this EL principle, when incorporated into the GLRT and AMF framework, leads to detection rules that are at least not inferior to LAMF with its data-independent loading factor selection.

The issue of ML estimation within the GLRT approach is not the only one that raises concerns within Kelly's "single data" GLRT method. Indeed, Kelly's approach whereby "the decision rule will be formulated in terms of the totality of input data without the a priori assignment of different functions to the primary and secondary input" [2] is not self-evident. It is difficult to accept two different covariance matrix estimates for the same interference, depending on the hypothesis for a single primary snapshot. It is expected that the proper methodology would search for the single interference covariance matrix that is most supportive to the detection problem, possibly dependent on the primary snapshot, but not on the hypothesis. This methodology should suggest both the primary and secondary data processing governed by the detection hypothesis testing criteria on the primary data, with the a priori classification of the secondary data as target-free being respected. The latter means that modification of the GLRT methodology should consider two sets instead of a single-set approach using "the totality of the input data," and for the above reasons use EL rather than ML estimation. This new GLRT framework should deliver adaptive detectors that are at least as efficient as LAMF, even for the scenarios that are most favorable to LAMF (5). Finally, we may hope that if the different adaptive techniques have the same performance, the potential accuracy set by the problem formulation is approached, and then we can favor the simplest technique for practical reasons.

The above-mentioned goals are pursued in this study, where Section II discusses our theoretical framework for the two-set GLRT and AMF detectors that use EL estimation. Section III specifies these techniques for Gaussian interference and a Gaussian (fluctuating) target. Section IV introduces the parametric class of covariance matrix estimates as the class of LSMI and FML estimates, and discusses the "conventional" constant-loaded AMF detector. For the "favorable" scenarios where the interference covariance matrix has an eigendecomposition of the form of (5), we give important features of our detectors (such as the CFAR property). Section V introduces the results of Monte-Carlo simulations that validate the theoretical results and demonstrate the high detection performance of the new adaptive detectors. The main points are summarized in Section VI, while Appendices A–D contain certain analytic derivations.

II. TWO-SET EL AND GLRT AMFs

According to the conventional single-set GLRT criterion, the decision d_1 that a target is present in a

(single) snapshot Y (or absent d_0) is taken according to the rule

$$\Lambda^*(Y) = \frac{\max_{\mu\in\Omega_1} f(Y \mid \mu, H_1)}{\max_{\chi\in\Omega_0} f(Y \mid \chi_1, H_0)} \underset{d_1}{\overset{d_0}{\lessgtr}} h^* \tag{19}$$

where $f(Y \mid \mu_1, H_1)$ is the (primary) data Y pdf for hypothesis H_1 (target signal present), and $\mu \in \Omega_1$ is the set of unknown (nonrandom) parameters that completely specify this pdf. In most cases, random parameters with an unknown (or not accurately known) marginal pdf are treated as $\mu \in \Omega_1$. Similarly, $f(Y \mid \chi_1, H_0)$ is the (primary) data pdf for hypotheses H_0 (target signal absent) specified by the parameters $\chi_1 \in \Omega_0$. The threshold h^* is defined as

$$\int_{\Lambda^*(Y)>h^*} f(Y \mid \chi_1, H_0) dY \le P_{\text{FA}} \qquad \forall \quad \chi \in \Omega_0 \tag{20}$$

where P_{FA} is the desired probability of false alarm.

Note that for a finite-dimensional vector Y, the GLRT criterion does not have a rigorous theoretical justification, similar to the Bayesian rule with given marginal pdfs $f_\mu(\mu)$ and $f_\chi(\chi)$, for example. Only asymptotic considerations ($M \to \infty$) are used to justify the GLRT method, despite its obvious intuitive appeal. In practice, this means that estimates other than the ML ones for μ and χ in (19) could be employed, and could result in better detection performance. We explore this avenue whenever alternative estimates are available, especially for covariance matrix estimation.

To fit into this single-set GLRT framework, Kelly [2] considered a single total data set $\{X_N; Y\}$, and introduced $\chi = R$ and $\mu = (R, a)$ with the two different solutions (7) and (8) for the interference covariance matrix. In our two-set GLRT framework, we introduce

$$f(X_N \mid \eta, \chi_{12}), \qquad \eta \in \Omega_2 \tag{21}$$

$$f(Y \mid \chi_0, \chi_{12} : H_0), \qquad \chi_0 \in \Omega_0 \tag{22}$$

$$f(Y \mid \mu, \chi_{12} : H_1), \qquad \mu \in \Omega_1 \tag{23}$$

where $\chi_{12} \in \Omega_{12}$, and Ω_{12} is the set of common (interference) parameters that describe the pdf for both the primary data Y and the secondary set X_N, and η, χ_0 and μ are now parameters specific to each set and hypothesis.

The a priori classification of the training (secondary) data X_N and the primary snapshot Y means that the estimates χ_{12} do not depend on the hypothesis H_0 versus H_1, and so is the same as appears in (22) and (23). Therefore the following option can be considered

$$\text{GLRT:} \qquad \Lambda_{12} = \max_{\eta,\chi_{12}} f(X_N \mid \eta, \chi_{12}) \times \frac{\max_{\mu\in\Omega_1} f(Y \mid \mu, \chi_{12} : H_1)}{\max_{\chi_0\in\Omega_0} f(Y \mid \chi_0, \chi_{12} : H_0)} \underset{d_1}{\overset{d_0}{\lessgtr}} h^*. \tag{24}$$

Note that this joint optimization over χ_{12} is already different from the standard AMF approach, namely:

$$\text{AMF:} \qquad \Lambda_{12}^{(2)} = \frac{\max_{\mu\in\Omega_1} f(Y \mid \mu, \chi_{12ML} : H_1)}{\max_{\chi_0\in\Omega_0} f(Y \mid \chi_0, \chi_{12ML} : H_0)} \underset{d_1}{\overset{d_0}{\lessgtr}} h^* \tag{25}$$

where

$$\chi_{12\text{ML}} = \arg\max_{\eta,\chi_{12}} f(X_N \mid \eta, \chi_{12}). \tag{26}$$

On the other hand, the joint optimization in (24) should result in a single solution in χ_{12} for both the hypotheses H_0 and H_1. It is important that this solution depends on the actual primary snapshot Y, but not on the hypothesis itself regarding this primary snapshot. Therefore, at least in principle, the GLRT approach (24) differs from Kelly's solution and the AMF technique, even though the same ML principle is used for both GLRT (24) and AMF (25).

However, in most of the cases considered here, we can replace the LF $f(X_N \mid \eta, \chi_{12})$ by the LR

$$LR(X_N \mid \chi_{12}) = \max_{\eta} \frac{f(X_N \mid \eta, \chi_{12})}{f_0(X_N)} \in (0, 1] \tag{27}$$

that has the same ML solution for χ_{12ML}:

$$\arg\max_{\chi_{12}} \max_{\eta,\chi_{12}} f(X_N \mid \eta, \chi_{12}) = \arg\max_{\chi_{12}} LR(X_N \mid \chi_{12}). \tag{28}$$

The most important property of this LR is that, for the actual (true, exact) $\chi_{12}^{(0)}$, the pdf does not depend on $\chi_{12}^{(0)}$, i.e., is scenario-free, depends only on the parameters M and N, and can be precalculated. Hence for a given probability P_0, the upper and lower bounds (α_U and α_L, respectively) can be found such that

$$\int_{\alpha_L}^{1} w[LR(X_N \mid \chi_{12}^{(0)})] dLR = \int_{0}^{\alpha_U} w[LR(X_N \mid \chi_{12}^{(0)})] dLR = P_0 = 1 - \varepsilon, \qquad 0 < \varepsilon \ll 1 \tag{29}$$

so that with the high probability $(1 - 2\varepsilon)$, the exact parameters $\chi_{12}^{(0)}$ generate the LR within the specified bounds. Here $w[\cdot]$ is the scenario-free pdf for the likelihood ratio $LR(X_N \mid \chi_{12}^{(0)})$.

For most cases with relatively small sample size ($N \simeq M$), Fig. 1 shows that

$$\alpha_U \ll 1 \tag{30}$$

and so the ML solution $\chi_{12\text{ML}}$ is far away from the true set of parameters in terms of the LR metric. Of course, very small LRs may be generated not only by the true parameters, but by a variety of completely

erroneous solutions as well. For this reason, we propose an EL approach that is based on LR matching for a certain parameterization of the estimate $\hat{\chi}_{12}(\beta)$ such that

$$\hat{\chi}_{12}(\beta_0) = \chi_{12\mathrm{ML}} \tag{31}$$

and the parameterization corresponds to some valid a priori assumptions regarding the class of covariance matrices.

We can now propose the following two-set GLRT (2S-GLRT) techniques:

$$\text{ML-GLRT:} \qquad \Lambda_{12}^{(3)} = \max_{\beta} \frac{\max_{\mu\in\Omega_1} f(Y \mid \mu, \hat{\chi}_{12}(\beta); H_1)}{\max_{\chi_0\in\Omega_0} f(Y \mid \chi_0, \hat{\chi}_{12}(\beta); H_0)} \underset{d_1}{\overset{d_0}{\lessgtr}} h^* \tag{32}$$

for β such that

$$\alpha_L \le LR(X_N \mid \beta) \le 1 \tag{33}$$

and

$$\text{EL-GLRT:} \qquad \Lambda_{12}^{(4)} = \max_{\beta} \frac{\max_{\mu\in\Omega_1} f(Y \mid \mu, \hat{\chi}_{12}(\beta); H_1)}{\max_{\chi_0\in\Omega_0} f(Y \mid \chi_0, \hat{\chi}_{12}(\beta); H_0)} \underset{d_1}{\overset{d_0}{\lessgtr}} h^* \tag{34}$$

for β such that

$$\alpha_L \le LR(X_N \mid \beta) \le \alpha_U. \tag{35}$$

The only difference between these two methods is that ML-GLRT allows the likelihood ratio $LR(X_N \mid \beta)$ generated by the parameterized estimate $\hat{\chi}_{12}(\beta)$ to arbitrarily approach the upper bound of unity, whereas EL-GLRT restricts the LR to the range of values where the exact parameters are concentrated, according to (29).

Similarly, we can also introduce the EL-AMF detector, where we replace ML estimates by EL ones:

$$\text{EL-AMF:} \qquad \Lambda_{12}^{(5)} = \frac{\max_{\mu\in\Omega_1} f(Y \mid \mu, \hat{\chi}_{12\mathrm{EL}} : H_1)}{\max_{\chi_0\in\Omega_0} f(Y \mid \chi_0, \hat{\chi}_{12\mathrm{EL}} : H_0)} \underset{d_1}{\overset{d_0}{\lessgtr}} h^*. \tag{36}$$

The EL estimate is the estimate which, for given training data X_N, generates the specific precalculated likelihood ratio LR_0, i.e.,

$$LR[X_N \mid \hat{\chi}_{12}(\beta_{\mathrm{EL}})] = LR_0. \tag{37}$$

This value LR_0 is chosen by referring to the scenario-free pdf $LR(X_N \mid \chi_{12}^{(0)})$. For example, the mean or median value of the pdf could be selected.

In the next section, we derive the two-set ML-GLRT, EL-GLRT, and EL-AMF techniques for the typical Gaussian model of interference and target signal. Later we compare the detection performance of these techniques with standard AMF and the clairvoyant case with known interference parameters.

III. TWO-SET GLRT AND AMF DETECTORS FOR GAUSSIAN MODELS

In most GLRT studies, the target is modeled by a vector of given structure (wavefront) with an unknown complex scaling factor that is an additional unknown deterministic parameter (see (9)). We adopt the typical target model here, namely the Swerling 1 model [35], which is the Gaussian model (Rayleigh target) with uniform initial phase and Rayleigh-distributed envelope.

A. Homogeneous Interference Training Conditions; Fluctuating Target with Known Power

In this case, we assume that the only information that is unavailable is the interference covariance matrix, which is identical for both the training and the primary data:

$$\Omega_0 = \emptyset, \qquad \Omega_1 = \emptyset, \qquad \Omega_2 = \emptyset, \qquad \Omega_{12} = \{R\} \tag{38}$$

and

$$f(X_N) = \frac{1}{\pi^N \det^N R} \exp[-\mathrm{tr}(R^{-1} X_N X_N^H)] \tag{39}$$

$$f(Y \mid H_1) = \frac{1}{\pi \det(R + \sigma_s^2 S S^H)} \exp[-Y^H (R + \sigma_s^2 S S^H) Y] \tag{40}$$

where S is the target signal wavefront vector and σ_s^2 is the target power.

Hence, according to (32)–(35), the 2S-GLRT detection problem is

$$\Lambda_{12} = \max_{R(\beta)} \frac{1}{1 + \sigma_s^2 S^H R^{-1}(\beta) S} \exp\left[\frac{\sigma_s^2 |Y^H R^{-1}(\beta) S|^2}{1 + \sigma_s^2 S^H R^{-1}(\beta) S}\right] \tag{41}$$

subject to

$$\text{ML-GLRT:} \qquad \alpha_L \le \frac{\det R^{-1}(\beta) \hat{R}_N \exp M}{\exp \mathrm{tr} R^{-1}(\beta) \hat{R}_N} \le 1 \tag{42}$$

$$\text{EL-GLRT:} \qquad \alpha_L \le \frac{\det R^{-1}(\beta) \hat{R}_N \exp M}{\exp \mathrm{tr} R^{-1}(\beta) \hat{R}_N} \le \alpha_U \tag{43}$$

where

$$\hat{R}_N \equiv \frac{1}{N} X_N X_N^H. \tag{44}$$

The AMF technique is based on

$$\Lambda_{12} = \frac{1}{1 + \sigma_s^2 S^H \hat{R}^{-1}(\beta) S} \exp\left[\frac{\sigma_s^2 |Y^H \hat{R}^{-1}(\beta) S|^2}{1 + \sigma_s^2 S^H \hat{R}^{-1}(\beta) S}\right] \tag{45}$$

subject to

$$\text{standard ML-AMF:} \qquad \hat{R}(\beta = \beta_0) = \hat{R}_N \tag{46}$$

$$\text{EL-AMF:} \qquad \frac{\det \hat{R}^{-1}(\beta_{\mathrm{EL}}) \hat{R}_N \exp M}{\exp \mathrm{tr} \hat{R}^{-1}(\beta_{\mathrm{EL}}) \hat{R}_N} = LR_0. \tag{47}$$

In (42) and (43),

$$\gamma_0^{(1)} \equiv LR(X_N \mid R_0) = \frac{\det N^{-1}\hat{C}\exp M}{\exp \operatorname{tr} N^{-1}\hat{C}} \tag{48}$$

where $\hat{C} \sim \mathcal{CW}(N,M,I_M)$, is described by a scenario-free (complex Wishart) pdf. Indeed, Appendix A shows that the hth moment is given by

$$\mathcal{E}\{\gamma_0^{(1)h}\} = \left(\frac{e}{N}\right)^{Mh} \frac{(N+h)^{-M(N+h)}}{N^{-M(N+h)}} \frac{\prod_{j=1}^{M}\Gamma(N+h+1-j)}{\prod_{j=1}^{M}\Gamma(N+1-j)} \tag{49}$$

$$= N^{MN}e^{Mh}\frac{1}{(N+h)^{M(N+h)}}\frac{\prod_{j=1}^{M}\Gamma(N+h+1-j)}{\prod_{j=1}^{M}\Gamma(N+1-j)}. \tag{50}$$

The pdf for $\gamma_0^{(1)}$, $w(\gamma_0^{(1)})$, can be expressed as an infinite series by applying a Mellin transform, similarly to [34] (see Appendix A).

The bounds are determined from the equations

$$\int_0^{\alpha_U} w(\gamma_0^{(1)})d\gamma_0^{(1)} = P_0, \qquad \int_{\alpha_L}^{1} w(\gamma_0^{(1)})d\gamma_0^{(1)} = P_0 \tag{51}$$

$$\int_0^{LR_0} w(\gamma_0^{(1)})d\gamma_0^{(1)} = 0.5 \tag{52}$$

by using either the analytical expression for $\gamma_0^{(1)}$ or by direct Monte-Carlo simulations.

B. Homogeneous Interference Training Conditions; Fluctuating Target with Unknown Power

In this case

$$\Omega_0 = \emptyset, \qquad \Omega_1 = \{\sigma_s^2\}, \qquad \Omega_2 = \emptyset, \qquad \Omega_{12} = \{R\}. \tag{53}$$

According to (32), we first need to find the ML estimate of the target signal power σ_s^2:

$$\max_{\sigma_s^2} \frac{1}{1+\sigma_s^2 S^H R^{-1}(\beta)S}\exp\left[\frac{\sigma_s^2|Y^H R^{-1}(\beta)S|^2}{1+\sigma_s^2 S^H R^{-1}(\beta)S}\right]. \tag{54}$$

Since $\sigma_s^2 \geq 0$, the solution is

$$\hat{\sigma}_s^2 = \begin{cases} \dfrac{|Y^H R^{-1}(\beta)S|^2 - S^H R^{-1}(\beta)S}{[S^H R^{-1}(\beta)S]^2} \\ \qquad\qquad \text{for} \quad \dfrac{|Y^H R^{-1}(\beta)S|^2}{S^H R^{-1}(\beta)S} \geq 1 \\ 0 \qquad\qquad \text{for} \quad \dfrac{|Y^H R^{-1}(\beta)S|^2}{S^H R^{-1}(\beta)S} < 1. \end{cases} \tag{55}$$

The solution $\hat{\sigma}_s^2 = 0$ clearly means that there is no target signal present in the input data, hence our 2S-GLRT test is

$$\Lambda_{12} = \max_{R(\beta)} \frac{S^H R^{-1}(\beta)S}{|Y^H R^{-1}(\beta)S|^2}\exp\left[\frac{|Y^H R^{-1}(\beta)S|^2}{S^H R^{-1}(\beta)S}\right] \times H\left(\frac{|Y^H R^{-1}(\beta)S|^2}{S^H R^{-1}(\beta)S} - 1\right) \underset{d_1}{\overset{d_0}{\lessgtr}} h^* \tag{56}$$

where $H(x)$ is the unit step function

$$H(x) = \begin{cases} 1 & \text{for} \quad x \geq 0 \\ 0 & \text{for} \quad x < 0. \end{cases} \tag{57}$$

Note that the function

$$f(x) = e^x/x \tag{58}$$

is monotonic for $x \geq 1$, and so this decision rule can be replaced by the more familiar one

$$\Lambda_{12} = \max_{R(\beta)} \frac{|Y^H R^{-1}(\beta)S|^2}{S^H R^{-1}(\beta)S} \underset{d_1}{\overset{d_0}{\lessgtr}} h^* > 1 \tag{59}$$

together with the same constraints on β as in (42) for ML-GLRT and in (43) for EL-GLRT. Note that this maximization can be interpreted as the intuitively appealing maximization of the sample output signal-to-interference ratio (SIR) when the adaptive filter is set to $\hat{W}(\beta) = \hat{R}^{-1}(\beta)S$, and the interference output power is calculated as $\hat{W}^H(\beta)\hat{R}(\beta)\hat{W}(\beta)$.

Actually, the fuller decision rule

$$\Lambda_{12} = \max_{R(\beta)} \frac{S^H R^{-1}(\beta)S}{|Y^H R^{-1}(\beta)S|^2}\exp\left[\frac{|Y^H R^{-1}(\beta)S|^2}{S^H R^{-1}(\beta)S}\right] \underset{d_1}{\overset{d_0}{\lessgtr}} h^* \tag{60}$$

can still be used for maximization, since for $x < 1$ it will lead to a further reduction in x. This means that if the target is absent from Y then a significant number of trials will result in $x < 0$, which means no target detected according to (56) and (58). Of course, a similar maximization result is probable for a weak target. Hence we may use the decision rule

$$\Lambda_{12} = \frac{|Y^H \hat{R}^{-1}(\hat{\beta})S|^2}{S^H \hat{R}^{-1}(\hat{\beta})S} \underset{d_1}{\overset{d_0}{\lessgtr}} h^* > 1 \tag{61}$$

where

$$\hat{R}(\hat{\beta}) = \arg\max_{R(\beta)} \frac{S^H R^{-1}(\beta)S}{|Y^H R^{-1}(\beta)S|^2}\exp\left[\frac{|Y^H R^{-1}(\beta)S|^2}{S^H R^{-1}(\beta)S}\right] \tag{62}$$

subject to the usual ML- (42) or EL-GLRT constraints (43).

Note that if we replace the x-monotonic function (56) or (58) by the function (62) that is unconstrained by the condition $x > 1$, we would expect a significantly different detection performance, even for the clairvoyant case $R = R_0$.

Similarly, we may introduce the AMF decision rule

$$\frac{S^H\hat{R}^{-1}(\beta)S}{|Y^H\hat{R}^{-1}(\beta)S|^2}\exp\left[\frac{|Y^H\hat{R}^{-1}(\beta)S|^2}{S^H\hat{R}^{-1}(\beta)S}\right] \times H\left(\frac{|Y^H\hat{R}^{-1}(\beta)S|^2}{S^H\hat{R}^{-1}(\beta)S}-1\right)\underset{d_1}{\overset{d_0}{\lessgtr}}h^*>1 \quad (63)$$

or

$$\frac{|Y^H\hat{R}^{-1}(\hat{\beta})S|^2}{S^H\hat{R}^{-1}(\hat{\beta})S}\underset{d_1}{\overset{d_0}{\lessgtr}}h^*>1. \quad (64)$$

Here $\hat{R}(\beta)$ should be chosen either in the standard way $\hat{R}(\beta=\beta_0)=\hat{R}_N$ to get the well-known ML-AMF method, or by (47) to get the EL-AMF rule.

C. Arbitrary Scaling Factors for Interference Matrices; Fluctuating Target with Unknown Power

Here we assume that the (total) power of the interference within the training data could be different to that in the primary data, so that the interference covariance matrix is the same up to an arbitrary scaling factor [18].

More specifically,

$$\mathcal{E}\{X_NX_N^H\}=c_2NR,\qquad \mathcal{E}\{YY^H\mid H_0\}=c_1R. \quad (65)$$

Hence we have

$$\Omega_0=\{c_1\},\qquad \Omega_1=\{c_1,\sigma_s^2\},\qquad \Omega_2=\{c_2\},\qquad \Omega_{12}=\{R\}. \quad (66)$$

According to (27) and (32), we first need to find

$$LR(X_N\mid R)=\max_{c_2}\frac{f(X_N\mid c_2)}{f_0(X_N)}. \quad (67)$$

Since

$$\max_{c_2}\frac{1}{\pi^Nc_2^{MN}\det R}\exp\left[-\frac{1}{c_2}\mathrm{tr}(R^{-1}X_NX_N^H)\right] \quad (68)$$

leads to the ML estimate

$$\hat{c}_{\mathrm{ML}}=\frac{1}{M}\mathrm{tr}[R^{-1}\hat{R}_N] \quad (69)$$

we end up with the familiar "sphericity test" in (67):

$$LR(X_N\mid R(\beta))=\left(\frac{\det R^{-1}(\beta)\hat{R}_N}{\left[\frac{1}{M}\mathrm{tr}\,R^{-1}(\beta)\hat{R}_N\right]^M}\right)^N. \quad (70)$$

For the GLRT detection rule, we also find

$$\max_{c_1}f(Y\mid H_0)=\frac{\det R^{-1}(\beta)}{\left[\frac{1}{M}\mathrm{tr}\,R^{-1}(\beta)\hat{R}_N\right]^M} \quad (71)$$

and

$$\max_{c_1,\sigma_s^2}f(Y\mid H_1) =\max_{\substack{\sigma_s^2>0\\ c_1>0}}\frac{\exp\left[-Y^HR^{-1}(\beta)Y+\frac{\sigma_s^2}{c_1}\frac{|Y^HR^{-1}(\beta)S|^2}{1+\frac{\sigma_s^2}{c_1}S^HR^{-1}(\beta)S}\right]}{\det c_1R(\beta)\left[1+\frac{\sigma_s^2}{c_1}S^HR^{-1}(\beta)S\right]}. \quad (72)$$

We first find $\sigma_{s\mathrm{ML}}^2$ by solving the log-likelihood equation

$$\frac{\partial}{\partial\sigma_s^2}\log f(Y\mid H_1)=0 \quad (73)$$

hence

$$1+\frac{\sigma_s^2}{c_1}S^HR^{-1}(\beta)S=\frac{1}{c_1}\frac{|Y^HR^{-1}(\beta)S|^2}{S^HR^{-1}(\beta)S} \quad (74)$$

which leads to the estimate (cf (55))

$$\hat{\sigma}_{s\mathrm{ML}}^2=\frac{|Y^HR^{-1}(\beta)S|^2-c_1S^HR^{-1}(\beta)S}{[S^HR^{-1}(\beta)S]^2} \quad (75)$$

for

$$\frac{|Y^HR^{-1}(\beta)S|^2}{S^HR^{-1}(\beta)S}\geq c_1. \quad (76)$$

Substituting (74) into (72) yields

$$\max_{c_1,\sigma_s^2}f(Y\mid H_1) =\frac{S^HR^{-1}(\beta)S}{c_1^{M-1}\det R(\beta)|Y^HR^{-1}(\beta)Y|^2} \times\exp\left\{\frac{1}{c_1}\left[-Y^HR^{-1}(\beta)Y+\frac{|Y^HR^{-1}(\beta)S|^2}{S^HR^{-1}(\beta)S}\right]\right\}. \quad (77)$$

Therefore the equation

$$\frac{\partial}{\partial c_1}\log\left[\max_{\sigma_s^2}f(Y\mid H_1)\right]=0 \quad (78)$$

leads to the solution

$$\hat{c}_{1\mathrm{ML}}=\frac{1}{M-1}\left[Y^HR^{-1}(\beta)Y-\frac{|Y^HR^{-1}(\beta)S|^2}{S^HR^{-1}(\beta)S}\right]. \quad (79)$$

Schwarz's inequality gives us $\hat{c}_{1\mathrm{ML}}\geq 0$, so substituting the above into (76) gives

$$\frac{|Y^HR^{-1}(\beta)S|^2}{S^HR^{-1}(\beta)SY^HR^{-1}(\beta)Y}\geq\frac{1}{M} \quad (80)$$

so our 2S-GLRT decision rule is

$$\max_R(\beta)\frac{H(\widehat{\cos}^2-\frac{1}{M})}{[1-\widehat{\cos}^2]^{M-1}\widehat{\cos}^2}\underset{d_1}{\overset{d_0}{\lessgtr}}h^*>1 \quad (81)$$

where

$$\widehat{\cos}^2 \equiv \frac{|Y^H R^{-1}(\beta) S|^2}{S^H R^{-1}(\beta) S Y^H R^{-1}(\beta) Y} \tag{82}$$

subject to

$$\text{ML-GLRT:} \qquad \alpha_L \le \frac{\det R^{-1}(\beta)\hat{R}_N}{\left[\frac{1}{M}\text{tr}\, R^{-1}(\beta)\hat{R}_N\right]^M} < 1 \tag{83}$$

$$\text{EL-GLRT:} \qquad \alpha_L \le \frac{\det R^{-1}(\beta)\hat{R}_N}{\left[\frac{1}{M}\text{tr}\, R^{-1}(\beta)\hat{R}_N\right]^M} \le \alpha_U. \tag{84}$$

It is straight forward to show that the function

$$f(\widehat{\cos}^2) = \frac{1}{[1-\widehat{\cos}^2]^{M-1}\widehat{\cos}^2} \tag{85}$$

is monotonic for $\widehat{\cos}^2 \ge 1/M$, so for the same reasons as before we may express the 2S-GLRT decision rule in the traditional form

$$\frac{|Y^H \hat{R}^{-1}(\beta) Y|^2}{S^H \hat{R}^{-1}(\beta) S Y^H \hat{R}^{-1}(\beta) Y} \underset{d_1}{\overset{d_0}{\lessgtr}} h^* \ge \frac{1}{M} \tag{86}$$

where

$$\hat{R}^{-1}(\beta) = \arg\max_{R(\beta)} \frac{1}{[1-\widehat{\cos}^2]^{M-1}\widehat{\cos}^2} \tag{87}$$

subject to the same constraints (83) or (84).

Naturally, the optimized function (85) should not be used directly, in order to avoid the inevitable performance degradation due to the nonmonotonic nature of this function over the entire interval $0 \le \widehat{\cos}^2 \le 1$.

We may now introduce the traditional ML-AMF solution (that is the adaptive coherence estimation (ACE) detector [36, 5]) as

$$\widehat{\cos}^2_{\text{ML}} = \frac{|Y^H \hat{R}_N^{-1} S|^2}{S^H \hat{R}_N^{-1} S Y^H \hat{R}_N^{-1} Y} \underset{d_1}{\overset{d_0}{\lessgtr}} h^* > \frac{1}{M} \tag{88}$$

and the EL-AMF solution as

$$\widehat{\cos}^2_{\text{ML}} = \frac{|Y^H \hat{R}^{-1}(\beta_{\text{EL}}) S|^2}{S^H \hat{R}^{-1}(\beta_{\text{EL}}) S Y^H \hat{R}^{-1}(\beta_{\text{EL}}) Y} \underset{d_1}{\overset{d_0}{\lessgtr}} h^* > \frac{1}{M} \tag{89}$$

where β_{EL} is determined by the condition

$$\frac{\det \hat{R}^{-1}(\beta_{\text{EL}})\hat{R}}{\left[\frac{1}{M}\text{tr}\, \hat{R}^{-1}(\beta_{\text{EL}})\hat{R}\right]^M} = LR_0. \tag{90}$$

As usual, the bounds α_L, α_U, and LR_0 are specified by the scenario-free pdf that has been derived in [37] for $\gamma_0^{(2)}$:

$$\gamma_0^{(2)} = \frac{\det \hat{C}}{\left[\frac{1}{M}\text{tr}\,\hat{C}\right]^M} \tag{91}$$

$$\hat{C} \sim \mathcal{CW}(N, M, I_M) \tag{92}$$

$$w(\gamma_0^{(2)}) = C(M,N)[\gamma_0^{(2)}]^{N-M} G^{M,0}_{M,M}\left(\gamma_0^{(2)}\Big|^{\frac{M^2-1}{M},\frac{M^2-2}{M},\ldots,\frac{M^2-M}{M}}_{0,1,\ldots,M-1}\right) \tag{93}$$

where

$$C(M,N) \equiv (2\pi)^{(M-1)/2} M^{(1-2MN)/2} \frac{\Gamma(MN)}{\prod_{j=1}^{M}\Gamma(N-j+1)} \tag{94}$$

and $G^{M,0}_{M,M}(\cdot)$ is Meijer's G-function [38].

Note that the above well-known models have been elucidated since they permit analytic solutions for the ML estimates $\hat{\sigma}^2_{s\text{ML}}$ and $\hat{c}_{\text{ML}}$. The same methodology can be applied for more complex models where such estimates are found numerically [21].

IV. DIAGONALLY LOADED AND FML ADAPTIVE DETECTORS; "FAVORABLE" SCENARIOS

According to the two-fold goal of this study, we now have to specify the particular parametric families for the covariance matrix estimate $R(\beta)$ within the above 2S-GLRT and AMF detection rules as the diagonally loaded and FML ones, and consider scenarios that are known to be most favorable for LSMI and FML adaptive filter techniques. We also have to specify the LSMI-based LAMF detector with its data-independent (constant) loading factor that is typically recommended for these scenarios. The performance of this conventional LAMF detector must then be compared with that of the new 2S-GLRT and EL-AMF methods that operate with adaptive data-dependent loading-factor selection.

The scenarios most favorable for LSMI treatment (5) are equivalently described by a covariance matrix of the form

$$R_0 = \mu U_s \Lambda_s U_s^H + U_n U_n^H \tag{95}$$

where, for simplicity, the white-noise power σ_n^2 is set to unity; U_s is the $M \times m$ matrix of m signal-subspace eigenvectors; U_n is the $M \times n$ matrix of noise-subspace eigenvectors, and $\mu\Lambda_s \gg I_m$ is the m-variate matrix of signal-subspace eigenvalues. The conditions

$$m < M, \qquad \mu \gg 1 \tag{96}$$

(or $\text{eig}_m(R_0) \gg \sigma_n^2$ in the general case) are favorable, meaning that there is typically tens of dBs difference between the signal- and noise-subspace eigenvalues. For this class of interference covariance matrices, it was demonstrated [10] that the adaptive filter

W_{LSMI} that arises from the diagonally loaded SMI technique

$$W_{\mathrm{LSMI}} \equiv (\beta I_M + \hat{R}_N)^{-1} S \tag{97}$$

where $\hat{R}_N \equiv X_N X_N^H / N$, and the loading factor β is selected within the broad range of values

$$\mu\lambda_m \gg \beta > 1, \qquad (\text{in general, } \mathrm{eig}_m R_0 \gg \beta > \sigma_n^2) \tag{98}$$

gives the normalized output SNR (SNR loss factor)

$$\gamma_{\mathrm{LSMI}} = \frac{[S^H(\beta I_M + \hat{R}_N)^{-1} S]^2}{S^H(\beta I_M + \hat{R}_N)^{-1} R_N (\beta I_M + \hat{R}_N)^{-1} S S^H R_N^{-1} S} \tag{99}$$

that is approximately described by the β-distribution

$$w(\gamma_{\mathrm{LSMI}}) = \frac{N!}{(N-m)!(m-1)!} (1-\gamma_{\mathrm{LSMI}})^{m-1} \gamma_{\mathrm{LSMI}}^{N-m}. \tag{100}$$

This distribution only depends on N and m, not on the loading factor or other scenario parameters. Moreover,

$$\mathcal{E}\{\gamma_{\mathrm{LSMI}}\} \simeq 3 \text{ dB} \qquad \text{for} \quad N \simeq 2m. \tag{101}$$

Later in [39], the same pdf was used to describe the SNR loss factor of the FML (Hung-Turner type) adaptive beamformer

$$W_{\mathrm{FML}} \equiv (\hat{\sigma}_n^2 I_M + \hat{U}_m \hat{A}_m \hat{U}_m^H) S \tag{102}$$

where

$$\hat{\sigma}_n^2 \equiv \frac{1}{M-m} \sum_{j=1}^{M-m} \lambda_{m+j}, \qquad \hat{A}_m \equiv \mathrm{diag}[\hat{\lambda}_j - \hat{\sigma}_n^2] \quad \text{for} \quad j = 1,\ldots,m \tag{103}$$

with U_j and λ_j $(j = 1,\ldots,M)$ coming from the eigendecomposition of the sample (ML) covariance matrix

$$\hat{R}_N = \hat{U}\hat{\Lambda}\hat{U}, \qquad \hat{U} \equiv [\hat{U}_m, \hat{U}_n], \quad \hat{\Lambda} \equiv \mathrm{diag}[\hat{\Lambda}_m, \hat{\Lambda}_n]. \tag{104}$$

Note that unlike the LSMI algorithm, the FML technique requires the signal-subspace dimension (order) m to be specified. For favorable conditions (95)–(96), this order can be estimated with high accuracy by treating $\hat{R}$ with an information-theoretic criterion (ITC) [40]. This approach is actually similar to the EL philosophy, and we show in Section IV that for favorable scenarios, EL matching gives as reliable an order estimate as any ITC. Therefore, for the FML method there is no practical difference between EL-FML and conventional FML (for (95)–(96)). For this reason, we concentrate our comparative analysis on the LAMF detector with constant diagonal loading:

$$\hat{R}_{\mathrm{LSMI}} = \beta_c I_M + \hat{R}_N. \tag{105}$$

According to the conventional AMF methodology, the LAMF detector can be derived from (105) being substituted into the detection test instead of the known covariance matrix [17]. We can now finally specify the adaptive detectors that are to be compared for favorable scenarios.

A. Homogeneous Interference Training Conditions; Fluctuating Target with Unknown Power

2S-GLRT:

$$\max_{\beta} \frac{|Y^H(\beta I_M + \hat{R}_N)^{-1} S|^2}{S^H(\beta I_M + \hat{R}_N)^{-1} S} \underset{d_1}{\overset{d_0}{\lessgtr}} h^* > 1 \tag{106}$$

subject to

$$\alpha_L \le \frac{\det[(\beta I_M + \hat{R}_N)^{-1}\hat{R}_N] \exp M}{\exp \mathrm{tr}[(\beta I_M + \hat{R}_N)^{-1}\hat{R}_N]} \le 1 \qquad \text{(ML-GLRT)} \tag{107}$$

or

$$\alpha_L \le \frac{\det[(\beta I_M + \hat{R}_N)^{-1}\hat{R}_N] \exp M}{\exp \mathrm{tr}[(\beta I_M + \hat{R}_N)^{-1}\hat{R}_N]} \le \alpha_U \qquad \text{(EL-GLRT)} \tag{108}$$

ML-AMF:

$$\frac{|Y^H \hat{R}_N^{-1} S|^2}{S^H \hat{R}_N^{-1} S} \underset{d_1}{\overset{d_0}{\lessgtr}} h^* > 1. \tag{109}$$

EL-AMF:

$$\frac{|Y^H(\hat{\beta} I_M + \hat{R}_N)^{-1} S|^2}{S^H(\hat{\beta} I_M + \hat{R}_N)^{-1} S} \underset{d_1}{\overset{d_0}{\lessgtr}} h^* > 1 \tag{110}$$

where

$$\hat{\beta} \equiv \arg_{\beta} \left\{ \frac{\det[(\beta I_M + \hat{R}_N)^{-1}\hat{R}_N] \exp M}{\exp \mathrm{tr}[(\beta I_M + \hat{R}_N)^{-1}\hat{R}_N]} \equiv L R_0 \right\} \tag{111}$$

LAMF:

$$\frac{|Y^H(\beta_c I_M + \hat{R}_N)^{-1} S|^2}{S^H(\beta_c I_M + \hat{R}_N)^{-1} S} \underset{d_1}{\overset{d_0}{\lessgtr}} h^* > 1 \tag{112}$$

where the constant β_c is about two or three.

B. Nonhomogeneous Interference Training Conditions; Fluctuating Target with Unknown Power

2S-GLRT:

$$\max_{\beta} \frac{|Y^H(\beta I_M + \hat{R}_N)^{-1} S|^2}{S^H(\beta I_M + \hat{R}_N)^{-1} S Y^H(\beta I_M + \hat{R}_N)^{-1} Y} \underset{d_1}{\overset{d_0}{\lessgtr}} h^* > \frac{1}{M} \tag{113}$$

subject to

$$\alpha_L \le \frac{\det[(\beta I_M + \hat{R}_N)^{-1}\hat{R}_N]}{\left\{\frac{1}{M}\mathrm{tr}[(\beta I_M + \hat{R}_N)^{-1}\hat{R}_N]\right\}^M} \le 1 \qquad \text{(ML-GLRT)} \tag{114}$$

or

$$\alpha_L \le \frac{\det[(\beta I_M + \hat{R}_N)^{-1}\hat{R}_N]}{\left\{\frac{1}{M}\mathrm{tr}[(\beta I_M + \hat{R}_N)^{-1}\hat{R}_N]\right\}^M} \le \alpha_U \qquad \text{(EL-GLRT)} \tag{115}$$

ML-AMF ("ACE"):

$$\frac{|Y^H\hat{R}_N^{-1}S|^2}{S^H\hat{R}_N^{-1}SY^H\hat{R}_N^{-1}Y} \underset{d_1}{\overset{d_0}{\lessgtr}} h^* > \frac{1}{M} \tag{116}$$

EL-AMF:

$$\frac{|Y^H(\hat{\beta} I_M + \hat{R}_N)^{-1}S|^2}{S^H(\hat{\beta} I_M + \hat{R}_N)^{-1}SY^H(\hat{\beta} I_M + \hat{R}_N)^{-1}Y} \underset{d_1}{\overset{d_0}{\lessgtr}} h^* > \frac{1}{M} \tag{117}$$

where

$$\hat{\beta} \equiv \arg_\beta \left\{ \frac{\det[(\beta I_M + \hat{R}_N)^{-1}\hat{R}_N]}{\left\{\frac{1}{M}\mathrm{tr}[(\beta I_M + \hat{R}_N)^{-1}\hat{R}_N]\right\}^M} \equiv LR_0 \right\} \tag{118}$$

LAMF:

$$\frac{|Y^H(\beta_c I_M + \hat{R}_N)^{-1}S|^2}{S^H(\beta_c I_M + \hat{R}_N)^{-1}SY^H(\beta_c I_M + \hat{R}_N)^{-1}Y} \underset{d_1}{\overset{d_0}{\lessgtr}} h^* > \frac{1}{M} \tag{119}$$

where the constant β_c is about two or three.

Naturally we are interested in the performance of these detectors compared with the FML-based ones. These can be introduced in a similar way, where now the signal-subspace dimension m in (102) is used as a parameter in 2S-GLRT optimization. As we have already mentioned, EL-FML has the same performance as the FML matrix estimate using the true m for favorable scenarios (95)–(96).

Note that only the familiar ML-AMF (ACE) detectors are known to be strictly CFAR detectors [17, 36]. Indeed, for signal-free primary data Y and no mismatch in the interference properties between the primary and secondary data, the above ML-AMF detectors have pdfs that are functions only of M and N. Such pdfs have been analytically derived in [17, 36], and can be used for false-alarm threshold calculations. For the other detectors introduced above, the strict CFAR property cannot be proven. Yet, for favorable scenarios, we can demonstrate a certain invariance of the output signal-free statistics that is sufficient for practical false-alarm threshold calculations. This invariance is specified by the following two theorems.

THEOREM 1 *Suppose we have a "favorable" interference covariance matrix of the form*

$$R_0 = \mu U_s\Lambda_s U_s^H + U_n U_n^H, \qquad \mu \gg 1, \quad \Lambda_s > I_m \tag{120}$$

and let the loading factor β in the LSMI estimate

$$\hat{R}_{\mathrm{LSMI}} = \beta I_M + \hat{R}_N, \qquad \hat{R}_N = X_N X_N^H/N, \quad X_N \sim \mathcal{CN}_N(0,R_0) \tag{121}$$

be selected within the range $\mu > \beta \gtrsim 1$, then

a) *the test statistics*

$$\hat{t}_1 \equiv \frac{|Y^H(\beta I_M + \hat{R}_N)^{-1}S|^2}{S^H(\beta I_M + \hat{R}_N)^{-1}S}, \qquad Y \sim \mathcal{CN}(0,R_0) \tag{122}$$

can be approximately (as $\mu \to \infty$) represented as

$$\hat{t}_1 \simeq \frac{|[Y_{1n}^H - Y_{1s}^H(Z_sZ_s^H)^{-1}Z_sZ_n^H]\mathcal{Z}\mathbf{e}_1|^2}{\mathbf{e}_1^T\mathcal{Z}\mathbf{e}_1} \tag{123}$$

where

$$\mathcal{Z} \equiv \left\{\beta I_n + \frac{1}{N}Z_n[I_N - Z_s^H(Z_sZ_s^H)^{-1}Z_s]Z_n^H\right\}^{-1} \tag{124}$$

and where

$$\begin{aligned} Y_{1n} &\in \mathcal{CN}^{n\times 1} \sim \mathcal{CN}(0,I_n) \\ Y_{1s} &\in \mathcal{CN}^{m\times 1} \sim \mathcal{CN}(0,I_m) \\ Z_n &\in \mathcal{CN}^{n\times N} \sim \mathcal{CN}_N(0,I_n) \\ Z_m &\in \mathcal{CN}^{m\times N} \sim \mathcal{CN}_N(0,I_m) \\ \mathbf{e}_1 &\in \mathcal{R}^{n\times 1} \equiv [1,0,\ldots,0]^T \end{aligned} \tag{125}$$

and Y_{1n}, Y_{1s}, Z_n, Z_m are mutually independent; and

b) *the test statistics*

$$\hat{t}_2 \equiv \frac{|Y^H(\beta I_M + \hat{R}_N)^{-1}S|^2}{S^H(\beta I_M + \hat{R}_N)^{-1}SY^H(\beta I_M + \hat{R}_N)^{-1}Y}, \qquad Y \sim \mathcal{CN}(0,R_0) \tag{126}$$

can be approximately (as $\mu \to \infty$) represented as

$$\hat{t}_2 \equiv \hat{t}_1 L_{\mathrm{LGIP}}^{-1} \tag{127}$$

where

$$\begin{aligned} L_{\mathrm{LGIP}} &\equiv Y^H(\beta I_M + \hat{R}_N)^{-1}Y \\ &\simeq Y_{1s}^H\left[\frac{1}{N}Z_sZ_s^H - \frac{1}{N^2}Z_sZ_n^H(\beta I_N + \frac{1}{N}Z_nZ_n^H)^{-1}Z_nZ_s^H\right]^{-1}Y_{1s} \\ &\quad - 2\Re[Y_{1s}^H(Z_sZ_s^H)^{-1}Z_sZ_n^H\mathcal{Z}Y_{1n}] + Y_{1n}^H\mathcal{Z}Y_{1n}. \end{aligned} \tag{128}$$

The proof appears in Appendix B.

Despite being rather bulky, these representations mean that for scenarios satisfying (95)–(96) with sufficiently large μ, the test statistics for target-free primary data can be expressed as a function of

white-noise IID data, and so its pdf will depend only upon the parameters N, M, m, and β. For LAMF, these representations can be used directly to calculate false-alarm thresholds, at least, by direct Monte-Carlo simulations. Of course, in this case the order m must be specified, but this is not a problem for favorable scenarios (95)–(96). With respect to these properties, we can treat the LAMF detector as being CFAR in practice.

Here we derived our most accurate representations for $\hat{t}_1$ and $\hat{t}_2$, sufficient for threshold calculations even for reasonably small β. For large values, we get less complicated expressions; specifically for

$$N\beta \gg 1 \tag{129}$$

$$\hat{t}_1 \simeq \frac{1}{\beta}|[Y_{1n}^H - Y_{1s}^H(Z_sZ_s^H)^{-1}Z_sZ_n^H]\mathbf{e}_1|^2 \tag{130}$$

and

$$L_{\mathrm{LGIP}} \simeq NY_{1s}^H(Z_sZ_s^H)^{-1}Y_{1s} + Y_{1n}^HY_{1n}/\beta \tag{131}$$

and with

$$\hat{f} \equiv Y_{1s}^H(Z_sZ_s^H)^{-1}Y_{1s} \sim \frac{\hat{f}^{m-1}}{B(m,N+1-m)(1+\hat{f})^{N+1}} \tag{132}$$

$$\hat{g} \equiv Y_{1n}^HY_{1n} \sim \frac{\hat{g}^{M-m+1}\exp[-\hat{g}]}{\Gamma(M-m)} \tag{133}$$

where B is the incomplete β-function. Note that the statistics L_{LGIP} can be considered to be the loaded version of the generalized inner product (GIP) test that was introduced in [26] for nonhomogeneity detection.

In order to expand these invariance properties to the new 2S-GLRT and EL-AMF techniques, all we have to do now is to demonstrate that (under favorable interference conditions) the LRs in (107) and (114) can be approximately represented as functions of the same random white-noise variables Z_s and Z_n. When dealing with GLRT optimization in (106) and (113), we are unlikely to get analytic expressions for the target-free thresholds. Yet, since both the optimized test statistics and constraints can be represented by white-noise variables, these pdfs can be defined by N, M, m, α_L and α_U, and may be precalculated using Monte-Carlo simulations.

For EL-AMF, such LR representation means that the EL loading factor β is a function of the same white-noise variables Z_s and Z_n, together with the expected median LR value LR_0, and so the target-free test statistics for EL-AMF are a function of N, M, m, and LR_0, and may again be precalculated using Monte-Carlo simulations. The invariance of the LRs is formalized by the following theorem.

THEOREM 2 *Suppose we have a favorable interference covariance matrix of the form*

$$R_0 = \mu U_s\Lambda_sU_s^H + U_nU_n^H, \qquad \mu \gg 1 \tag{134}$$

and let the loading factor β in the LSMI estimate

$$\hat{R}_{\mathrm{LSMI}} = \beta I_M + \hat{R}_N, \qquad \hat{R}_N = X_NX_N^H/N, \quad X_N \sim \mathcal{CN}_N(0,R_0) \tag{135}$$

be selected within the range $\mu > \beta \gtrsim 1$, then the LRs

$$\gamma_{\ell 0}^{(1)} \equiv \frac{\det[\hat{R}_{\mathrm{LSMI}}^{-1}\hat{R}_N]\exp M}{\exp\mathrm{tr}[\hat{R}_{\mathrm{LSMI}}^{-1}\hat{R}_N]} \tag{136}$$

$$\gamma_{\ell 0}^{(2)} \equiv \frac{\det\hat{R}_{\mathrm{LSMI}}^{-1}\hat{R}_N}{\left[\frac{1}{M}\mathrm{tr}\hat{R}_{\mathrm{LSMI}}^{-1}\hat{R}_N\right]^M} \tag{137}$$

may be approximately represented as

$$\gamma_{\ell 0}^{(1)} = \frac{\det[\mathcal{Z}(\mathcal{Z}^{-1}-\beta I_N)]}{\exp[-\beta\mathrm{tr}\,\mathcal{Z}]} \tag{138}$$

$$\gamma_{\ell 0}^{(1)}(\beta = 0) = 1 \tag{139}$$

$$\gamma_{\ell 0}^{(1)}(\beta \gg 1) = \beta^{-(M-m)}\det[\mathcal{Z}^{-1}-\beta I_N]\exp[M-m] \tag{140}$$

and

$$\gamma_{\ell 0}^{(2)} = \frac{\det[\mathcal{Z}(\mathcal{Z}^{-1}-\beta I_N)]}{[1-\beta\mathrm{tr}\,\mathcal{Z}/M]^M} \tag{141}$$

$$\gamma_{\ell 0}^{(2)}(\beta = 0) = 1 \tag{142}$$

$$\gamma_{\ell 0}^{(2)}(\beta \gg 1) = \beta^{-(M-m)}\det[\mathcal{Z}^{-1}-\beta I_N]\left(1-\frac{m}{M}\right)^{-M}. \tag{143}$$

The proof appears in Appendix C.

In the next section, we demonstrate that these approximations are sufficiently accurate for practical false-alarm threshold calculations for "favorable" scenarios.

V. DETECTION PERFORMANCE ANALYSIS OF 2S-GLRT, AMF AND LAMF DETECTORS

A. Favorable Interference Scenario

Consider an $M = 12$-sensor uniform linear antenna array, and $m = 6$ independent Gaussian interference sources, each with 30 dB signal-to-white-noise ratio (SWNR). The interference directions of arrival (DOAs) were chosen to be

$$\mathbf{w}_6 \equiv \sin\boldsymbol{\theta}_6 = [-0.8,-0.4,0.2,0.5,0.7,0.9] \tag{144}$$

so that the eigenspectrum of the interference covariance matrix

$$R_0 = \sigma_n^2I_M + \sum_{j=1}^{6}\sigma_j^2S(w_j)S^H(w_j) \tag{145}$$

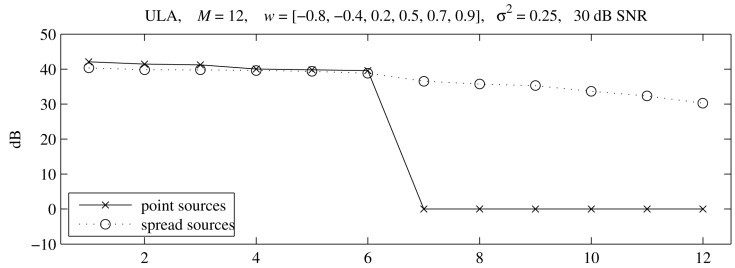


Fig. 2. Eigenspectrum of favorable and unfavorable interference covariance matrices used in the simulations.

where σ_n^2 is the white-noise power, and

$$\sigma_n^2 = 1, \qquad \sigma_j^2 = 1000, \qquad w_j = 2\pi \frac{d}{\lambda} \sin\theta_j \tag{146}$$

as shown in Fig. 2 perfectly meets the favorable interference conditions (95)–(96), since $\lambda_7/\lambda_8 = \lambda_7/\sigma_n^2 \simeq 35$ dB. Note that under our assumption regarding the training data consisting of N IID samples, the above adaptive detectors are applicable to any spatial, temporal or space-time application. In fact, the eigenspectrum in Fig. 2 has been specifically chosen to look like that of the terrain-scattered space-time covariance matrix in a side-looking airborne radar with three antenna sensors and four repetition periods [41].

Two separate target DOAs w_0 have been selected to represent two extreme cases, namely

$$\frac{S_0^H[I - S_6(S_6^H S_6)^{-1} S_6^H]S_0}{S_0^H S_0} = \begin{cases} 0.949 & \text{for} \quad w_0^H = -0.60 \\ 0.040 & \text{for} \quad w_0^L = 0.18. \end{cases} \tag{147}$$

In the first case ("fast target" in STAP (space-time adaptive processing) terminology), total interference mitigation is not accompanied by a significant degradation in target SWNR, whereas in the second case ("slow target"), such interference "nulling" leads to a dramatic signal-power reduction.

Note that for the clairvoyant detector ($R = R_0$), as well as for the standard GLRT and AMF detectors, this distinction does not affect the ROC if the output SNRs are identical:

$$\sigma_{sL}^2 S_0^{LH} R_0^{-1} S_0^L = \sigma_{sH}^2 S_0^{HH} R_0^{-1} S_0^H. \tag{148}$$

Of course, the standard GLRT and AMF detectors are CFAR detectors and therefore their false-alarm rate (thresholds) do not depend on the particular scenario in (147). For the introduced 2S-GLRT, EL-AMF, and LAMF detectors, we have yet to demonstrate that their invariance of the output target-free statistics predicted in Section IV holds sufficiently for constant false-alarm thresholds.

The training sample size N in all our simulations as been chosen according to the RMB rule that ensures 3 dB average SNR losses in the SMI adaptive filter: $N = 2M = 24$.

1) *Homogeneous Interference Training Conditions; Fluctuating Target with Unknown Power*: First, consider the performance of the clairvoyant detector

$$\frac{|Y^H R_0^{-1} S|^2}{S^H R_0^{-1} S} \underset{d_1}{\overset{d_0}{\lessgtr}} h > 1 \tag{149}$$

whose ROC has the well-known analytic expression

$$P_{\rm D} = \exp\left[-\frac{h}{1 + \sigma_s^2 S_0^H R_0^{-1} S_0}\right] \tag{150}$$

where h is the threshold and $(\sigma_s^2 S_0^H R_0^{-1} S_0)$ is the output SNR. For the AMF (ML-AMF) detector, the probability of false-alarm $P_{\rm FA}$ and of target detection can be expressed somewhat differently to RFKN [17] (see Appendix D)

$$\begin{aligned} P_{\rm FA} &= {}_2F_1(N - M + 1, N - M + 2, N + 1; -h) \\ &= \frac{1}{(1+h)^{N-M+1}} {}_2F_1\left(N - M + 1, M - 1, N + 1; -\frac{h}{1+h}\right) \end{aligned} \tag{151}$$

where ${}_2F_1(\alpha, \beta, \gamma; x)$ is the hypergeometric function [38], and

$$\begin{aligned} P_{\rm D} = {} & \left[\frac{1 + \sigma_s^2 S^H R^{-1} S}{1 + \sigma_s^2 S^H R^{-1} S + h}\right]^{N-M+1} \\ & \times F_1\left(M - 1, -(N - M + 1), N - M + 1, N + 1; \right. \\ & \left. \frac{\sigma_s^2 S^H R^{-1} S}{1 + \sigma_s^2 S^H R^{-1} S}, \frac{\sigma_s^2 S^H R^{-1} S + h}{1 + \sigma_s^2 S^H R^{-1} S + h}\right) \end{aligned} \tag{152}$$

where $F_1(\alpha, \beta, \beta', \gamma; x, y)$ is the hypergeometric function of two variables [38]. Note that [38]

$${}_2F_1\left(M - 1, 0, N + 1; \frac{\sigma_s^2}{1 + \sigma_s^2}\right) = 1 \tag{153}$$

hence for $h = 0$ we have $P_{\rm D} = 1$, and for $\sigma_s^2 = 0$ we have $P_{\rm D} = P_{\rm FA}$. These analytical expressions are used to validate the results of our Monte-Carlo simulations. Specifically, we compare the simulated and theoretical ROC for the clairvoyant detector, and use the free software routine gsl_sf_hyperg_2F1 from the GNU Scientific Library (GSL) (http://www.gnu.org/software/gsl/) to calculate ${}_2F_1(\alpha, \beta, \gamma; x)$ and so find the threshold values h for

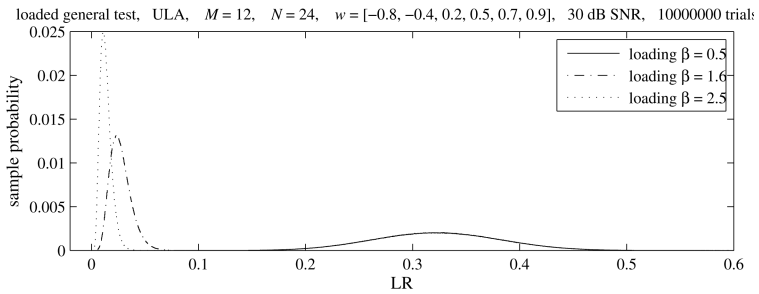


Fig. 3. Sample pdfs for the loaded general test.

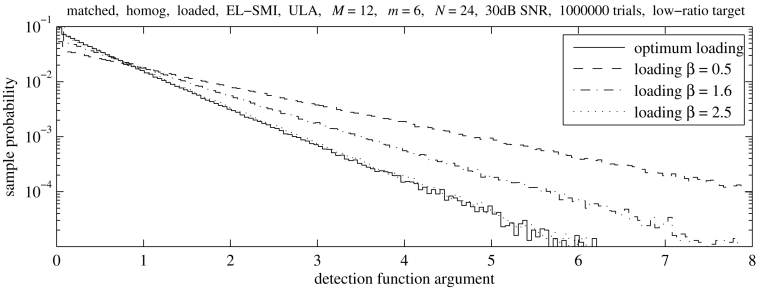


Fig. 4. Sample pdfs for target-free output signal of LAMF detector.

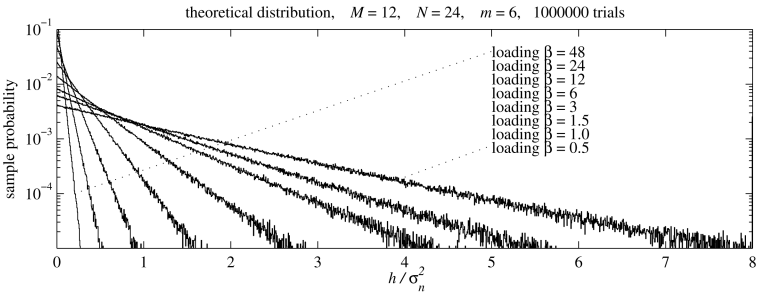


Fig. 5. Sample pdf for white-noise theoretical model for target-free output signal of LAMF detector.

false-alarm rates from 10^{-2} to 10^{-4}. Comparison of the analytically computed values with those computed over one million Monte-Carlo trials demonstrated a perfect match that finally validates the accuracy of the other Monte-Carlo results.

Next we consider the performance of our new detectors, beginning with LAMF which we expect to serve as a benchmark for the 2S-GLRT and EL-AMF detectors with adaptive (data-dependent) loading factors, as suggested by the theoretical framework. Fig. 3 shows sample pdfs calculated over 10^6 Monte-Carlo trials for the LR $\gamma_{0\ell}^{(1)}$ (13) and three fixed loading factors $\beta = 0.5, 1.6, 2.5$. We see that the loading factor $\beta = 1.6$ comes close to matching the pdf of the LR generated by the true covariance matrix R_0, as shown in Fig. 1. Therefore for this scenario we expect $\beta = 1.6$ to be a "sufficient" constant loading factor.

For this relatively low diagonal loading, the accuracy of our white-noise approximation (123) of the output target-free statistics becomes a critical issue that must be addressed first. Fig. 4 illustrates sample pdfs for the output target-free statistics calculated over 10^6 Monte-Carlo trials for the "low-ratio" target in (147). (The pdfs for the "high-ratio" target are not presented here, since they are indistinguishable.) The invariance of false-alarm rate with respect to the target scenario is thus demonstrated.

In order to assess the accuracy of our white-noise theoretical approximations for the output target-free statistics, Fig. 5 shows the sample pdfs for loading factors ranging from $\beta = 0.5$ to 48, and at Fig. 6 we compare the actual threshold values calculated for $P_{\mathrm{FA}} = 10^{-2}$ using the theoretical white-noise representation (123) (10^6 trials) and direct Monte-Carlo simulations (10^4 trials, low-ratio target). We find an absolutely accurate correspondence of threshold values, hence for this scenario the LAMF detector is indeed a CFAR detector, since false-alarm threshold values can be precalculated with sufficient accuracy for any given value of m (the number of dominant covariance matrix eigenvalues). This comparison also proves similar CFAR properties of EL-AMF and 2S-GLRT detectors, since they are

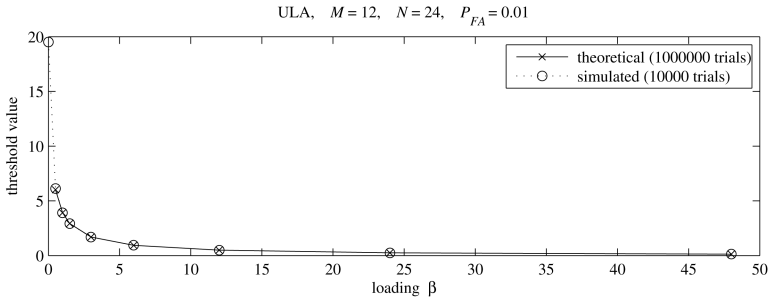
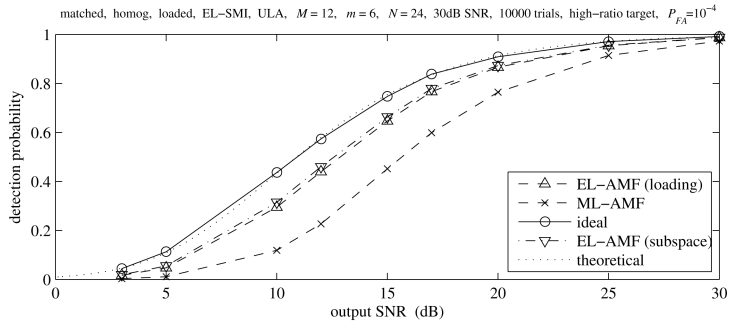
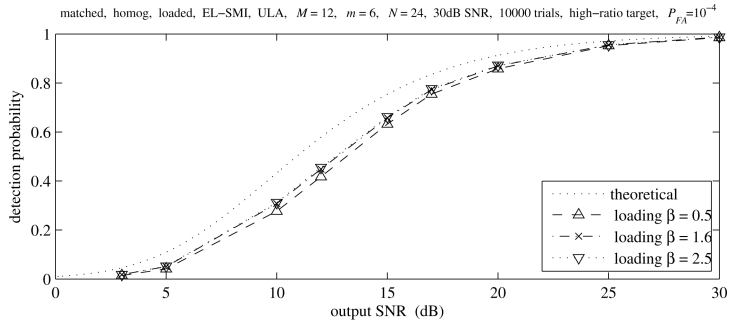


Fig. 6. Comparison of theoretical and actual threshold values for $P_{\mathrm{FA}} = 10^{-2}$ and varying loading factors.



(a)



(b)

Fig. 7. AMF ROCs for high-ratio target for false-alarm probability $P_{\mathrm{FA}} = 10^{-4}$.

based on the same analytical approximation (see Theorem 2).

Now we are in a position to consider the ROC of the clairvoyant, standard AMF, and LAMF detectors, which are presented in Fig. 7 for false-alarm rate set at $P_{\mathrm{FA}} = 10^{-4}$. The ROCs for $P_{\mathrm{FA}} = 10^{-2}$ and 10^{-3} are not presented here, as they are little different. (Again, the ROCs for the low-ratio target are identical to this high-ratio target.) We observe a perfect match between the simulated and analytical ROCs for the clairvoyant detector (149), proving the accuracy of the simulations. Despite the different target model (fluctuating in our study and nonfluctuating in [2, 17]), the standard AMF detector demonstrates performance similar to [17]. Indeed, for $P_{\mathrm{FA}} = 10^{-4}$ and $P_{\mathrm{D}} = 0.5$, our SNR loss factor is about 5 dB, compared with about 3 dB in [1]. A simplistic interpretation of this could be the additional ~ 2 dB losses due to adaptive thresholding.

The most important result following from Fig. 7 is that LAMF does indeed have significantly better performance. Indeed, the SNR loss factor for $P_{\mathrm{FA}} = 10^{-4}$ and $P_{\mathrm{D}} = 0.5$ is 1.6 dB, compared with 5 dB for the standard AMF detector. Less expected is the fact that a very small diagonal loading ($\beta = 0.5$) is only marginally inferior (< 0.1 dB) to LAMF for its "optimal" loading of $\beta \simeq 1.6$. The "optimality" for this fixed loading factor does not need to be very accurately specified, since even for $\beta = 12, 24, 48$ (Fig. 8), we observe the same performance as for $\beta = 1.6$.

Thus our expectations regarding LAMF superiority and performance invariance with respect to the constant loading factor $1 \lesssim \beta < \mu$ are proven correct for this favorable scenario. Its quasi-CFAR properties obviously make LAMF especially attractive in such cases; this and its high performance make it challenging for our "theoretically derived" 2S-GLRT and EL-AMF detectors to be as good. For this reason,

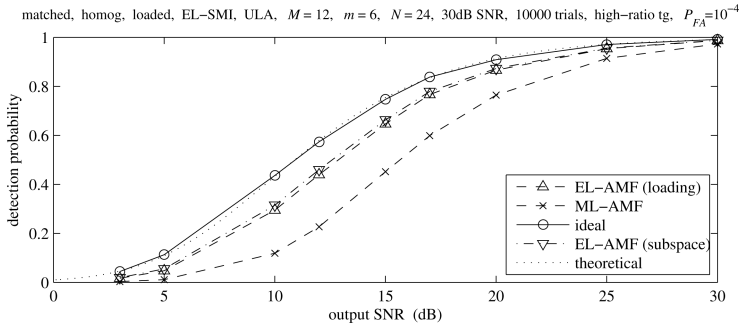


Fig. 8. AMF ROCs for high-ratio target for larger loading factors.

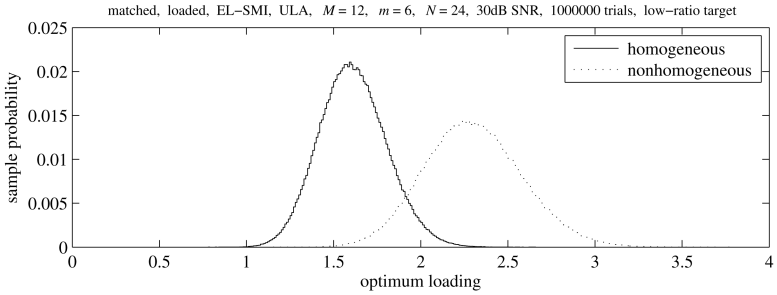


Fig. 9. Sample pdfs for optimum loading in EL-AMF technique.

we now concentrate on EL-AMF performance, wherein for each trial the loading factor is selected adaptively so that the LR $\gamma_{\ell 0}^{(1)}$ for the observed sample matrix $\hat{R}_{\mathrm{LSMI}}$ is equal to $LR_0 = 0.0257$. For the FML algorithm with its limited range of admissible signal subspace dimensions m, in each Monte-Carlo trial the LR $\gamma_{\ell 0}^{(1)}$ closest to $LR_0 = 0.0257$ has been generated for the true number of sources $m = 6$. Hence for the FML class of covariance matrix estimates, no difference can be reported between the adaptive EL-FML and the nonadaptive FML detector with the fixed (true) parameter m.

Fig. 9 shows a sample pdf of the optimum loading factors over 10^6 trials. Unsurprisingly, it spans the familiar range of loading factors $1 \lesssim \beta \lesssim 2.3$. As predicted, the false-alarm rate is found to be invariant with respect to target scenario, with identical threshold values calculated for our high- and low-ratio targets. Moreover, the ROCs are found to depend on the output SNR irrespective of the target type. We see that both EL-AMF (loading) and EL-AMF (FML) techniques have practically indistinguishable detection performance, and most importantly, both are in turn indistinguishable from the "properly" loaded $(1 \leq \beta \leq 48)$ LAMF technique. It is important to note that EL-AMF is not trivially identical to the constantly loaded LAMF. For example, LAMF false-alarm rates of 10^{-2}, 10^{-3}, and 10^{-4} give rise to LR threshold values of 4.00, 6.16, and 8.33, respectively, for $\beta = 1$, while for EL-AMF the same thresholds are 2.90, 4.55, and 6.23. Despite this significant difference in detection statistics, as we have just seen, the EL-AMF and LAMF performance is here practically indistinguishable.

Next, we consider the ROCs of the ML-GLRT method (106) and (107), where only the lower bound for the LR is introduced (that is dependent on the optimized loading factor β or interference subspace dimension m). Fig. 10 shows the ML-GLRT ROCs for the high-ratio/fast target; as before, the ROCs for

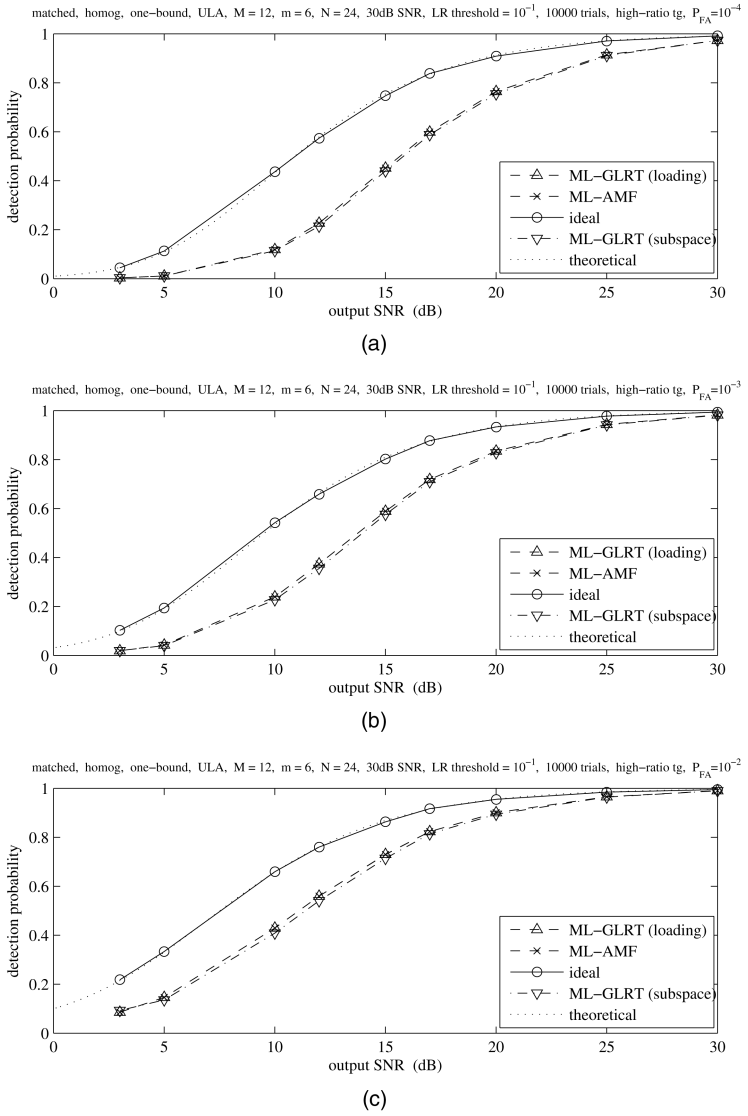


Fig. 10. ML-GLRT ROCs for high-ratio target for three different false-alarm probabilities.

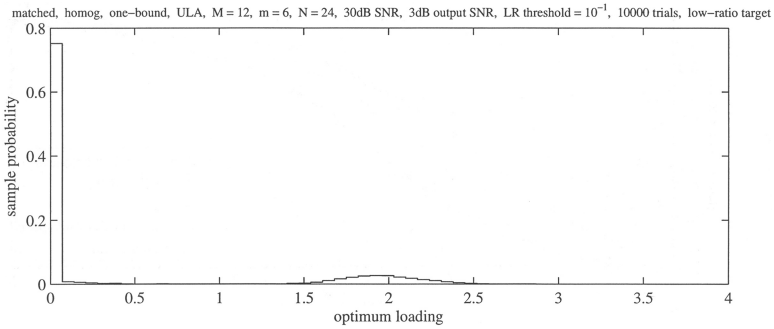


Fig. 11. Sample pdf for optimum loading in ML-GLRT technique.

the low-ratio/slow target are visually identical. The lower bound $\alpha_L = 0.0142$ has been chosen so that the probability of generating a LR below this threshold for the true matrix R_0 is very low: $P(LR < \alpha_L) = 0.1$ (see Fig. 1). As we might expect, we see that the ML-GLRT performance is practically identical to that of the (standard) ML-AMF method with its (unloaded) SMI. Again, this precise coincidence is not caused by a trivial zero loading-factor selection in the detection test optimization (106). The pdf for the selected loading factor for 3 dB output SNR is presented at Fig. 11. We see that, despite their different nature, our 2S ML-GLRT detector has the same performance as the traditional (zero loading) ML-AMF detector. This

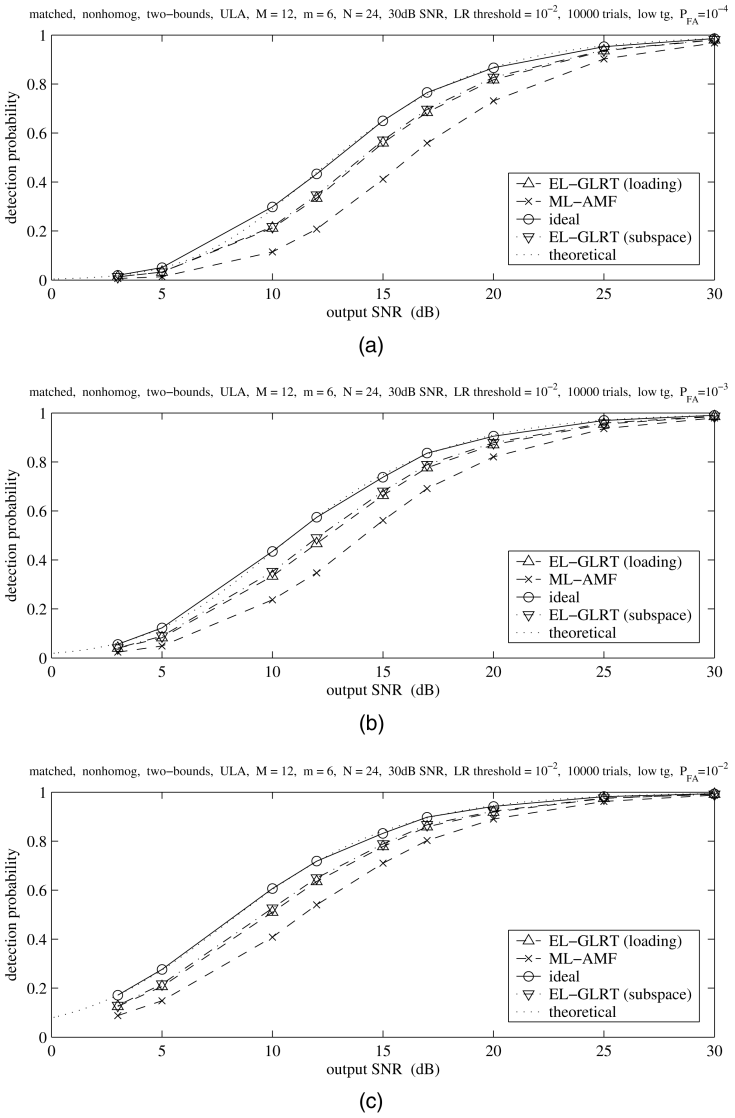


Fig. 12. EL-GLRT ROCs for high-ratio target for three different false-alarm probabilities.

proves our assertion that a single snapshot does not make a significant difference, even for the relatively small sample support studies here. The difference in the covariance matrix estimates for the specific target model analyzed by Kelly [2] is responsible for the slightly better performance of ML-GLRT compared with ML-AMF, whereas for some other models these two techniques are found to be identical (e.g., [18]), or even AMF outperforming GLRT (e.g., [17]).

When both lower and upper bounds on the LR for the optimized loading are introduced in accordance with the EL-GLRT method (106) and (108), the results are completely different. Fig. 12 show the ROCs for the high-ratio target; the ROCs for the low-ratio target are practically identical, and are not presented. Despite the broader area of admissible LR values, $P(LR < \alpha_L) = 10^{-2}$ and $P(LR > \alpha_U) = 10^{-2}$ ($\alpha_L = 0.0084$ and $\alpha_U = 0.0647$), the EL-GLRT ROCs are surprisingly close to those produced by the EL-AMF and LAMF methods. Indeed, the same improvement compared with 2S ML-GLRT and ML-AMF is observed in this case. The same practical performance for the diagonal loading and the interference subspace dimension selection is observed, despite the fact that in this case the optimized signal subspace dimension is far from being always correct in the 2S EL-GLRT algorithm.

This analysis of a typical scenario with fluctuating target in homogeneous interference clearly demonstrates that ML-based ML-GLRT and ML-AMF detectors share the same performance, while a significant and practically identical performance improvement is obtained for both techniques when our new EL approach is substituted for the standard ML criterion for finding the appropriate diagonal loading or interference subspace dimension.

As expected for this favorable scenario, the data-independent loading of LAMF (112) has the same performance as the theoretically derived EL-GLRT and EL-AMF techniques. The precise

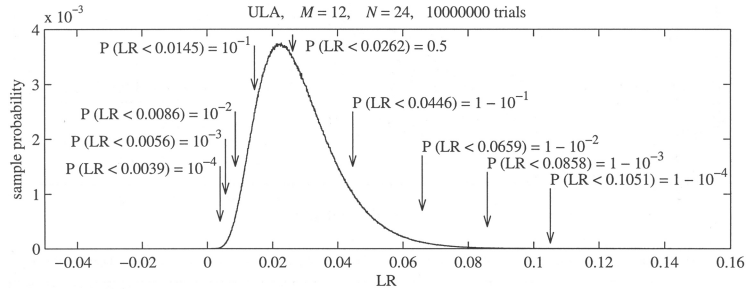


Fig. 13. The pdf for sphericity test.

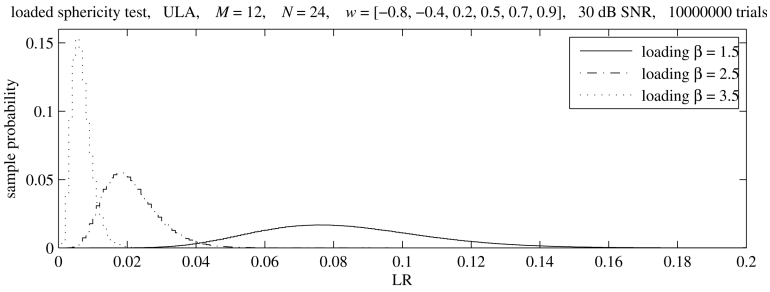


Fig. 14. Sample pdfs for loaded sphericity test.

coincidence of performance for the completely different detectors EL-GLRT, EL-AMF, and LAMF most likely means that they have all approached the ultimate performance set by the adaptive detection problem formulation.

Note that the LR generated by the properly loaded ($\beta = 1.6$) matrix $\hat{R}_{\text{LSMI}}$ is statistically close to that for the true matrix R_0. This match could then be used as a new guide for selecting the diagonal loading factor, but from a theoretical viewpoint could again be treated as another justification of our EL methodology.

2) *Nonhomogeneous Interference Training Conditions; Fluctuating Target with Unknown Power*: Here we briefly introduce our simulation results for this alternate signal model, detection rules, and LRs (113)–(119) which demonstrate the same overall findings as for homogeneous interference training conditions.

We begin with the ROC of the clairvoyant detector:

$$\frac{|Y^H R_0^{-1} S|^2}{S^H R_0^{-1} S Y^H R_0^{-1} Y} \underset{d_1}{\overset{d_0}{\lessgtr}} h \geq \frac{1}{M} \tag{154}$$

that can be analytically computed as [42]

$$P_{\text{D}} = \left[\frac{(1-h)(\sigma_s^2 S^H R^{-1} S + 1)}{(1-h)(\sigma_s^2 S^H R^{-1} S + 1) + h} \right]^{M-1} \tag{155}$$

with $P_{\text{D}} = P_{\text{FA}}$ for $\sigma_s^2 = 0$. This expression will again be used to validate our Monte-Carlo simulation accuracy.

As before, we start with an analysis of the traditional ML-AMF (ACE) detector (116), and compare it with LAMF and EL-AMF. Fig. 13 shows the pdf of the "sphericity test" LR that is used for this model as the EL benchmark. This pdf is only slightly different from the "general (nonsphericity) test" that we saw in Fig. 1, for example, the median LR value is here 0.0262 instead of 0.0257 as before. Fig. 14 shows the sample pdf for the sphericity test generated by the constantly loaded covariance matrix estimate $\hat{R}_{\text{LSMI}}$ with the three example loading factors $\beta = 1.5$, 2.5, and 3.5. The best correspondence with the "EL pdf" in Fig. 13 is for $\beta = 2.5$, and this means for this constant loading we can already expect the best possible detection performance (similar to that of EL-GLRT and EL-AMF).

Similarly to the previous homogeneous case, we checked the accuracy of our theoretical white-noise approximation for the target-free output statistics (Theorem 1) and found a perfect match between these and the directly calculated false-alarm threshold values.

Fig. 15 shows the high-ratio target sample ROCs for the clairvoyant, ML-AMF (ACE) and LAMF detectors, the latter for our three example loading factors. As expected, the theoretical and sample ROCs for the clairvoyant case match perfectly. ML-AMF (ACE) has familiar detection losses of about 3.5 dB for $P_{\text{FA}} = 10^{-4}$ and $P_{\text{D}} = 0.5$ [17, 36]. The practically identical performance of LAMF for all three loading factors (and that is only 1 or 1.5 dB inferior to the clairvoyant detector) now comes as no surprise. Fig. 9 also shows the sample pdf of the EL-AMF loading factor that matches $LR[\hat{R}_{\text{LSMI}}(\beta)]$ of the loaded sample matrix to the median LR value 0.0262. We see that the less accurate a priori assumptions on the interference covariance matrices in the primary and training data has resulted

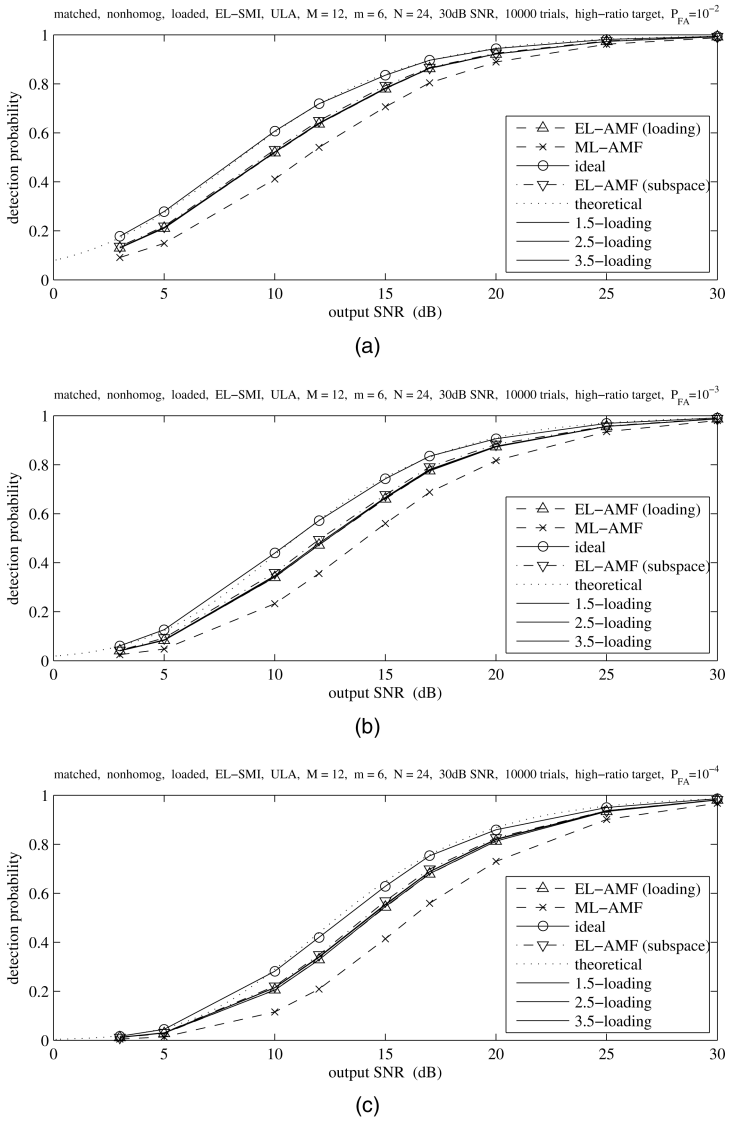


Fig. 15. Nonhomogeneous AMF ROCs for high-ratio target for three different false-alarm probabilities.

in noticeably greater admissible loading factors, from about 1.6 to 3.5. Most importantly, Fig. 15 shows that EL-AMF is as good as the properly loaded LAMF. Again, the FML detector correctly found the number of sources, and is as good as EL-AMF and LAMF.

Now we consider the ML-GLRT (113) and (114) results, with its lower bound on optimized LR: $P(LR < \alpha_L) = 0.01$ ($\alpha_L = 0.0145$ according to Fig. 3). Fig. 16 presents the slow target ROCs (the fast ones are almost identical). Unsurprisingly, standard ML-AMF (ACE) and ML-GLRT have almost the same performance for all P_{FA} and P_{D}, both for diagonal loading and interference subspace dimension selection. Once again, this coincidence does not mean the trivial equality of these routines. In fact, the ultimate interference subspace dimension $\hat{m} = 9$ was selected in about 40% of all trials, and the true interference rank (equal to six) was selected in only 10% of trials.

Similar insight is provided by analyzing the sample pdf of the optimized ML-GLRT loading factor (not shown here). For target-free input data, this pdf is dominated by one peak at $\beta \simeq 0$ and has a second peak at $\beta \simeq 2.8$. The second peak is present for small SNRs such as 3–5 dB, and disappears for sufficiently strong targets ($\gtrsim$15 dB), where zero loading dominates the selection (probability above 0.9 for SNR $>$ 15 dB). It seems remarkable that these "random walks" in the optimized loading factor and interference subspace dimension leave practically no trace on the ROC's behavior, as they are found to be the same as for the standard AMF (ML-AMF) test with constant zero loading.

Finally, let us consider the results of the EL-GLRT technique (113) and (115) for this scenario, where the upper and lower bounds are specified by the conditions $P(LR > \alpha_U) = 10^{-2}$, $P(LR < \alpha_L) = 10^{-2}$ (with $\alpha_L = 0.0056$ and $\alpha_U = 0.0659$ according to Fig. 3). Fig. 17 illustrates the EL-GLRT ROCs that

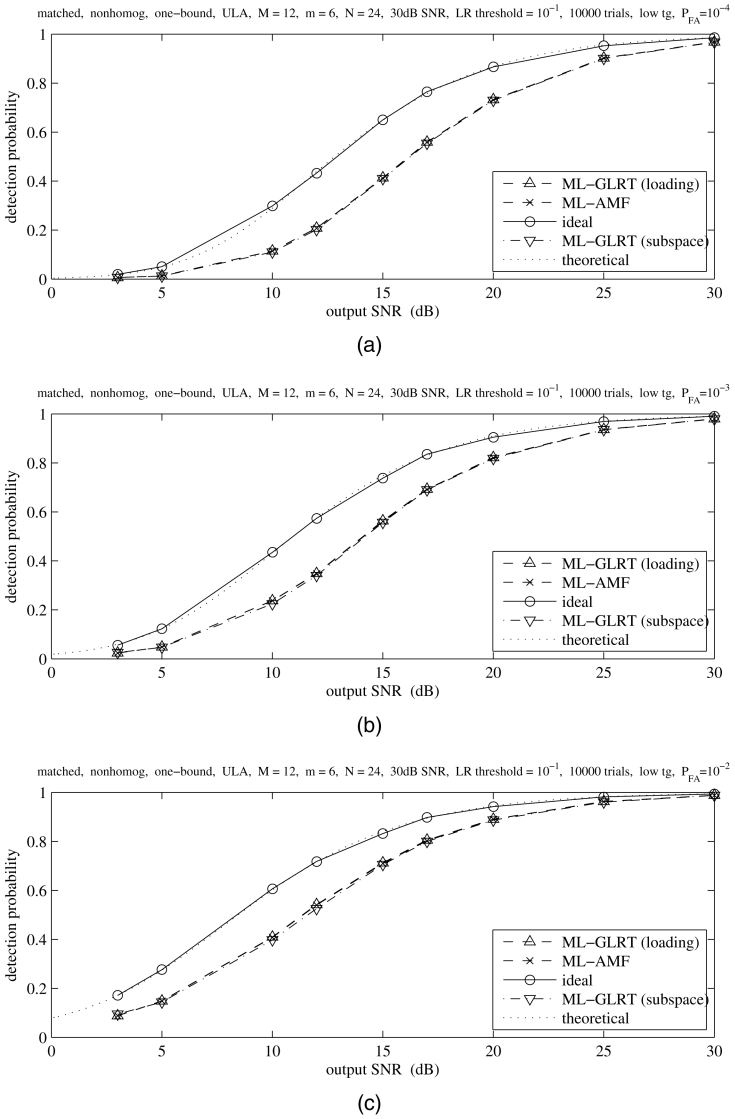


Fig. 16. Nonhomogeneous ML-GLRT ROCs for low-ratio target for three different false-alarm probabilities.

repeat the trend of being practically indistinguishable from those of EL-AMF. Despite this, the EL-GLRT optimum loading factor has strikingly different behavior, and is strongly dependent on the primary data SNR.

In this regard, it is instructive to analyze the sequence of sample loading factor distributions with output SNRs varying from 3–30 dB (not all illustrated here). We found that for the smallest output SNR, the pdf has two distinct peaks at $\beta = 1.5$ and 3.2. As the SNR increases, the second peak decreases, until it disappears at 30 dB SNR. A detailed analysis of the maximized function $\widehat{\cos}^2(\beta)$ of (85) reveals the cause of the two peaks for small output SNR; Fig. 18 shows the sample pdf $f(\widehat{\cos}^2)$ for the slow target and 3 dB output SNR. We see that maximization of this function in the vicinity of $\widehat{\cos}^2 = 1/M$ can "drive" the argument either into the area $\widehat{\cos}^2 < 1/M$ (which means no detection), or into the admissible area with $\widehat{\cos}^2 > 1/M$. In the former case, greater loading factors (within the permitted range) are selected, while in the latter case, smaller loading factors are found to be optimal. Similar behavior is exhibited by the optimal interference subspace dimension m. It is remarkable that, despite the significant redistribution of the optimized $\widehat{\cos}^2$ compared with the clairvoyant (and EL-AMF) cases, the overall ROCs are exactly the same as the EL-AMF ones, and only about 1 dB below the clairvoyant case.

This detailed analysis demonstrates that the accurate equivalence of EL-GLRT, EL-AMF and LAMF performance is not due to the trivial equivalence of these routines for the given model, as in [18] for example. Such an accurate performance equivalence of these three quite different algorithms suggests that in different ways both methods approach the ultimate performance set by the nature of the adaptive detection problem. It is also evident that

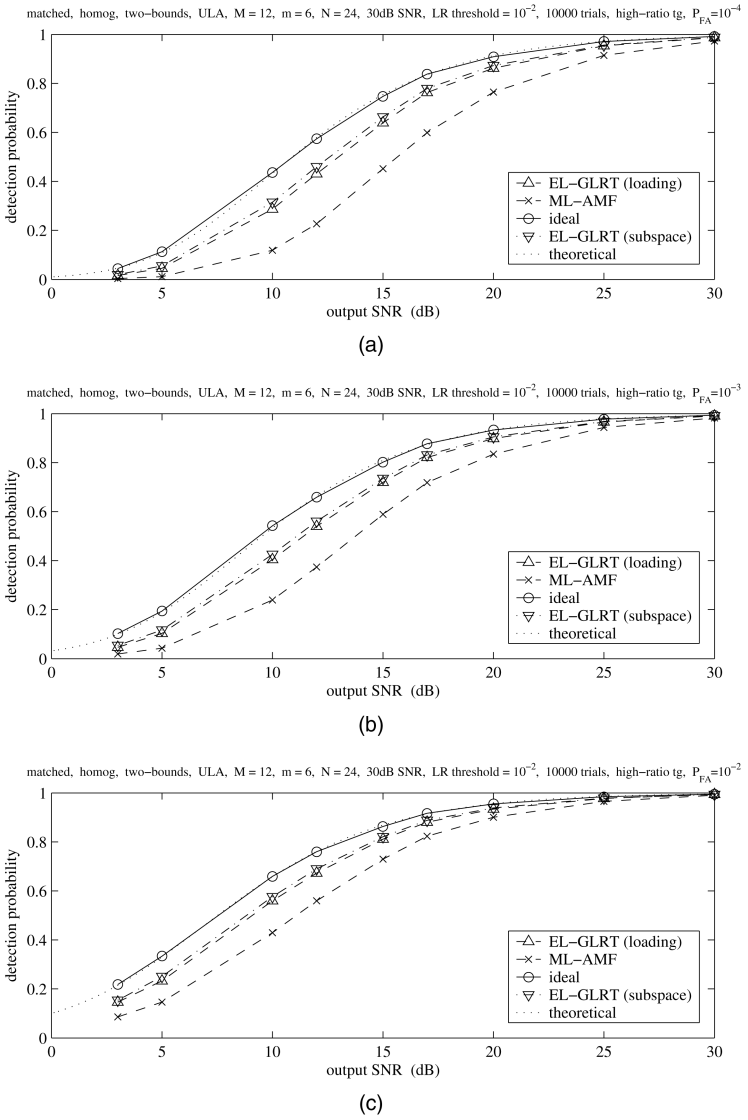


Fig. 17. Nonhomogeneous EL-GLRT ROCs for low-ratio target for three different false-alarm probabilities.

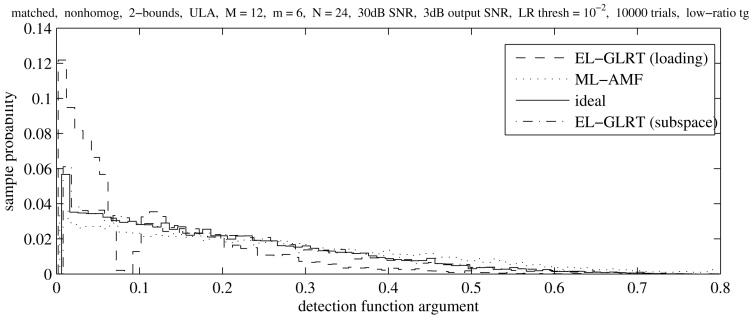


Fig. 18. Sample pdf for $\widehat{\cos}^2$ for slow target scenario and 3 dB output SNR.

the particular parameters used here for EL-GLRT ($P_{\text{bound}} = 10^{-2}$) and EL-AMF ($P_{\text{median}} = 0.5$) have no practical impact upon the demonstrated performance.

B. "Unfavorable" Interference Scenario

It seems quite important to conclude our study by considering the performance of our new detectors for interference models that are unfavorable for the LSMI and/or FML techniques, that is, scenarios with full-rank interference (no noise subspace) and having no obvious abrupt change in the size of the (sorted) covariance matrix eigenvalues.

In fact, adaptive processing does not make sense for small ratios λ_1/λ_M. Even the clairvoyant

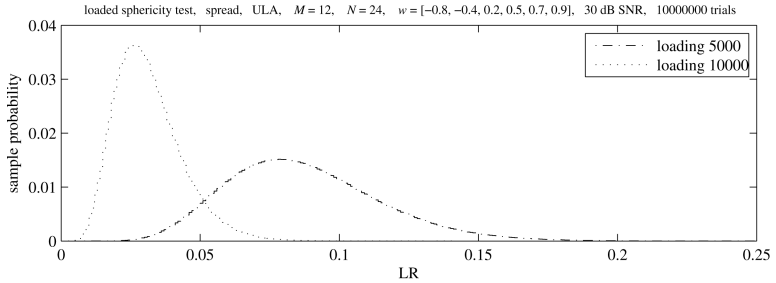


Fig. 19. Sample pdfs for loaded sphericity test for spread (distributed) interferences.

detector has a marginal improvement over the white-noise-matched receiver in this case, hence the loss associated with adaptivity can actually exceed the potential improvement. The ultimate example of such a scenario is input white noise. Of course, nobody should consider adaptive processing for internal white noise, but sometimes the external interference environment can resemble white noise, such as when the number of strong interference sources exceeds the number of antenna sensors (exceeding the degrees of freedom of the array). If the only available information about the external interference is contained in the training data, then a properly designed adaptive detector should succeed for such scenarios as well.

Let us now consider the same six-source interference scenario (144) and (146), except that each source is now equally "spread" (distributed) [43]. Here the interference covariance matrix R_0 can be written as [44]

$$R_0 = B \odot \left[\sigma_n^2 I_M + \sum_{j=1}^{6} \sigma_j^2 S(w_j) S^H(w_j) \right] \quad (156)$$

where [43]

$$B = \{\exp[-\nu|\ell - k|]\}_{\ell,k=1,\ldots,M}, \qquad \beta > 0 \quad (157)$$

is the "spreading matrix" and $\odot$ denotes the Schur-Hadamard (element-wise) matrix product. Similarly to [43], we use a spreading factor of $\nu = 0.25$.

Such spreading "annihilates" the noise subspace of the original point-source covariance matrix. The eigenspectrum shown in Fig. 2 is now characterized by a ratio $\lambda_1/\lambda_M = 10$ compared with the original $\lambda_1/\lambda_M \simeq 35$ dB. Moreover, the clairvoyant optimum filter ($W_{\text{opt}} = R_0^{-1}S$) has an SNR improvement over the white-noise matched filter ($W_{\text{wn}} = S$) of only 0.40 dB for the high-ratio (fast) target and 0.27 dB for the low-ratio (slow) target, compared with the favorable scenario values of 28.5 dB and 26.7 dB, respectively.

We consider nonhomogeneous training conditions with a fluctuating target of unknown power. (Based on our previous results, we expect similar behavior for the homogeneous scenario.) Since the clairvoyant and ACE ROCs do not depend on scenario and are exhaustively specified by N, M, P_{FA}, and output SNR, the ACE detector will again be about 3.5 dB inferior to the clairvoyant detector. We need to investigate whether the EL-AMF and properly loaded LAMF detectors can reduce these losses for diagonally loaded and FML sample covariance matrix estimates. Since the pdf for the sphericity test $LR(R_0)$ is scenario independent and so is the same as already shown in Fig. 13, we can proceed to the pdfs of the sphericity test $LR(\hat{R}_{\text{LSMI}})$ for the three loading factors $\beta = 3$, 5000, and 10000; Fig. 19 shows the latter two. The traditional loading factor of $\beta = 3$ (with respect to unit internal noise power) for favorable scenarios is in this case "as good as" zero loading, as it does not affect the LR whose pdf is concentrated within the range of LRs 0.9996–1. Only for such very high loading factors as five or ten thousand does the LR pdf behave similarly to the EL benchmark presented in Fig. 13. Note that $\beta = 10^4$ means that the diagonal loading here is of the same order of magnitude as the maximum eigenvalue of the true covariance matrix R_0. In fact, such loading drives LAMF towards the white-noise matched detector

$$\frac{|Y^H S|^2}{Y^H Y S^H S} \underset{d_1}{\overset{d_0}{\lessgtr}} h > \frac{1}{M} \quad (158)$$

which is quite understandable given the marginal SNR improvement provided by the clairvoyant optimal filter.

Naturally, the CFAR properties for favorable covariance matrices that were proven in Section IV do not hold for these scenarios and such loading factors. Indeed, for the low-ratio (slow) target and $\beta = 10^4$, the false-alarm rates $P_{\text{FA}} = 10^{-2}, 10^{-3}, 10^{-4}$ are imposed by the threshold values 0.43, 0.56, 0.66, respectively, while for the high-ratio (fast) target they are 0.22, 0.32, 0.42. Despite losing the CFAR properties, an ROC analysis of LAMF still makes sense (see Fig. 20). Whereas the clairvoyant detector ROCs again demonstrate a perfect match between analytical calculations and simulations, we see that the LAMF performance for $\beta = 3$ is the same as for ACE, as predicted. At the same time, the properly loaded LAMF detector ($\beta = 5000, 10000$) is only 0.4 dB inferior to the clairvoyant detector. The similar performance of EL-AMF now comes as no surprise.

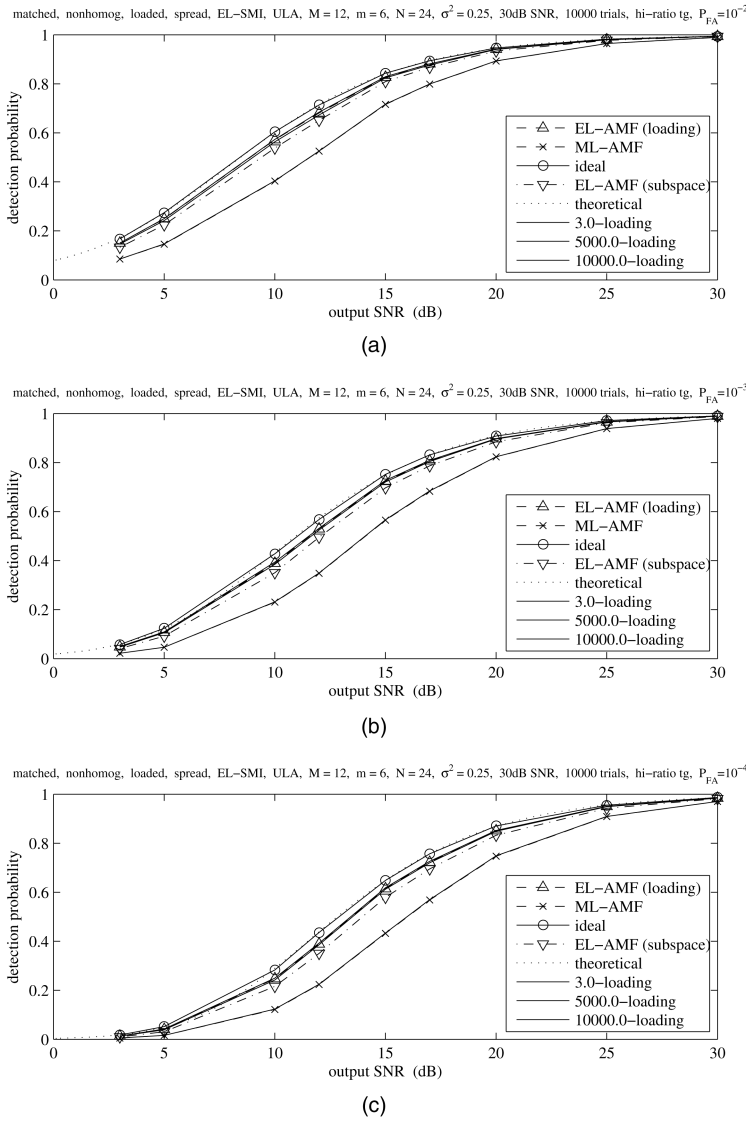


Fig. 20. Nonhomogeneous AMF ROCs for high-ratio target for spread interferences.

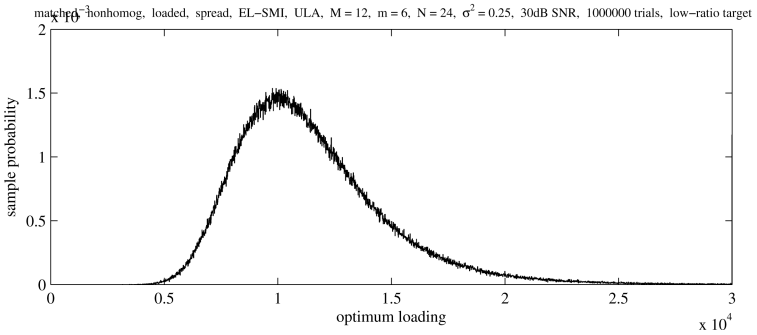


Fig. 21. Sample pdf for optimum loading in EL-AMF technique for spread interferences.

Fig. 21 shows the sample pdf (calculated over one million trials) of the optimum loading factor in the EL-AMF technique that was found by matching the sphericity test LR for the loaded sample covariance matrix $\hat{R}_{\text{LSMI}}$ with the expected LR value $LR_0 = 0.0262$ (see Fig. 13). As we might expect, the optimum loading factor lies mostly in the range $5000 \lesssim \beta \lesssim 10000$. Interestingly, the optimum signal subspace dimension histogram (not illustrated here) is not single-valued: 146 trials chose $\hat{m} = 6$, with the remainder $\hat{m} = 7$. While the loaded EL-AMF detector is as good as LAMF with proper loading, the EL-FML detector (where the signal subspace dimension is found by LR matching (117) and (118))

is marginally inferior (less than 0.2 dB). The coarse discretization of the parameter m could be the reason for this small degradation.

Note that traditional recommendations regarding loading factor selection ($\beta = 3$) or signal-subspace selection ($m = 6$), stemming from favorable conditions, are completely inappropriate here.

Whereas the properly loaded LAMF detector is statistically equivalent to EL-AMF, in practice, when only a sample covariance matrix is available, the only option seems to be to choose the "proper loading factor" (such as $\beta = 10^4$) through LR matching with the expected LR.

Results for the low-ratio (slow) target, as well as the 2S EL-GLRT detector, are similar: 2S ML-GLRT is as good as the conventional ACE, with 2S EL-GLRT, EL-AMF, and properly loaded LAMF being marginally inferior to the clairvoyant detector.

Finally, we wish to comment that the loss of the CFAR property for 2S EL-GLRT, EL-AMF, and LAMF is explainable by the EL matching driving them close to the robust white-noise matched detector, which is not a CFAR detector with respect to the covariance matrix R_0 ($\neq I_M$) and signal S.

VI. SUMMARY AND CONCLUSIONS

In this paper, we have diligently pursued two major goals. Firstly, based on the well-known properties of the LSMI and FML adaptive filter techniques, we introduced adaptive detectors that use the same type of covariance matrix estimates. We did this anticipating that such detectors should have better performance than those that use the conventional ML covariance matrix estimate, at least for (favorable) interferences that have widely separated signal and noise eigensubspaces. Having achieved this, we secondly addressed an important theoretical issue that follows from the now-established superiority of these LSMI and FML detectors. Namely, it seemed necessary to us to suggest modification of the frameworks that were introduced for GLRT by Kelly [2] and for AMF by RFKN [17] to include these new advanced detectors. We sought a modified theoretical framework that would either justify these detectors, and/or generate new ones that are at least as efficient as LAMF and FML for favorable scenarios.

More specifically, with respect to the first goal, we demonstrated that (for favorable interference) the LAMF detector that uses the diagonally loaded sample covariance matrix estimate whose loading factor is chosen from the broad range between the minimum signal and the noise eigenvalues does indeed give a significant detection performance improvement. Moreover, for such interference we demonstrated an important invariance property of the target-free detection statistics. We showed that these statistics are invariant with respect to the true covariance matrix, and can be closely approximated by white-noise generated data, and that the pdf is only a function of the filter dimension, the training sample size, the signal subspace dimension, and the loading factor. The approximation accuracy is sufficiently high to permit the precalculation of false-alarm thresholds, which means that in this case the LAMF detector has, in fact, the CFAR property.

With quite negligible losses demonstrated by LAMF compared with the clairvoyant detector, it then became a challenging problem to propose a theoretical adaptive detection framework capable of giving similar detection performance. In order to achieve this second goal of our study, we reconsidered two important issues in adaptive detection that were postulated within the traditional (Kelly's one-sample) GLRT and AMF techniques. The first concerns the arrangement made by Kelly [2] whereby "the decision rule will be formulated in terms of the totality of input data without the a priori assignment of different functions to the primary and secondary input." Specifically, despite the assumption that the secondary data is free of target component, "any selection rules applied to make this assumption more plausible are ignored" in his technique. We have reformulated the GLRT problem as a two-set adaptive detection problem, where the target-free status of the training data is respected, and so the interference covariance matrix estimate, while depending on the primary sample, does not depend on the hypothesis being tested.

More specifically, we have introduced a new two-set GLRT technique whereby the single covariance matrix estimate for both hypotheses is adaptively selected from a family of admissible solutions so as to maximize the "detection function," which is the associated LR. Unlike most published studies, we used the fluctuating Swerling I model for the target signal, that is specified by a (possibly unknown) positive power. This modeling introduced some important changes into the nature of the optimized LR; most importantly, since a nonpositive ML target power estimate is inadmissible, it must be replaced by a zero estimate that corresponds to a decision that the target is absent.

The second (more important) issue that we have reconsidered here concerns the ML criterion used within the traditional GLRT and AMF techniques. It is important that only asymptotic arguments support the GLRT methodology, so there is no reason why estimates other than the ML one could not be found to be better suited for detection, especially for relatively small training sample sizes. Three straight forward observations alerted us to challenge this dogma. The first stems from the difference in performance between Kelly's GLRT and the AMF methods; for different models, one method was found to be superior [17]. The suggestion made in [17] that

Kelly's technique should be better just because it involves an additional single (primary) snapshot in the covariance matrix estimation could be easily refuted, and so the performance difference should rather be attributed to the specific estimates used.

The second observation arises from the fact that the LSMI or FML sample covariance matrix estimates, that are found to be very successful in LAMF (FML) detectors, are not the ML covariance matrix estimates.

The third theoretical observation arises from a comparison of the actual LR (which is just the normalized LF) produced by the exact (true) covariance matrix and the (unconstrained) ML estimate adopted by both GLRT and AMF. While the ML covariance matrix estimate (sample matrix) always delivers the ultimate value for the maximized LR, equal to one irrespective of the sample size, the exact covariance matrix yields much smaller LRs for relatively small sample support. For the specific example analyzed here, with an $M = 12$-sensor uniform linear antenna array and $N = 24$ samples, the median value of the LR for the exact covariance matrix is found to be only 0.025, and with probability 0.99 the LR lies in the range 0.008–0.07. Thus for relatively small sample support, the ML estimator is extremely far from the true covariance matrix, even in terms of the LR/LF metric.

For this reason, we have introduced an approach called EL, whereby we try to find the estimate that statistically generates the same LR as the exact covariance matrix. This is feasible in practice since the pdf for the LR generated by the exact covariance matrix does not depend on the matrix itself, but only the parameters M and N, and so it can be precalculated.

We have used the above well-known families of covariance matrix estimates: the diagonally loaded sample matrix (i.e., the loaded unconstrained ML solution) and the finite-subspace interference approximation of the ML solution. For these estimates, respectively, the traditional ML criterion drives the loading factor to zero, and the interference subspace dimension to its maximum. For 2S ML-GLRT, the loading factor and interference subspace dimension are constrained only by the lower bound on LR, while the maximum LR value is allowed to reach the ultimate value of unity, as per the ML solution. Despite the new formulation, our Monte-Carlo simulations demonstrated that 2S ML-GLRT detection performance is almost the same as for the traditional ML-AMF (ACE) method that uses the standard ML covariance matrix estimate (zero loading, full interference rank). Thus the detection performance is again proven to be dominated by the type of ML estimate rather than the choice of GLRT or AMF method that uses this estimate.

Our new EL method searches for the diagonal loading or interference subspace dimension so that the modified EL estimate gives an LR value properly within the range of LRs expected for the exact covariance matrix. To be more specific, for the 2S EL-GLRT method we maximize the detection function over the set of (loaded, finite-rank) solutions bounded by the precalculated upper and lower LR bounds. For the EL-AMF (ACE) technique, we seek the loaded solution that generates the median LR value of the exact covariance matrix. For finite-rank approximations that have only a finite number of solutions, we simply find the one that is closest to the upper LR bound, if no solution within the bounds is available.

Our extensive Monte-Carlo simulations for a scenario with signal- and noise-subspace eigenvalues separated by several tens of dBs showed that the EL-GLRT and EL-AMF methods have practically the same performance. For finite-rank approximation in this example, all solutions were of the same subspace dimension, as only the true interference rank estimate was closest to the bounded LR region, making no difference between "adaptive" and "fixed" signal-subspace dimension. While the loading factor fluctuates in the well-known range of 1.5–3 times the white-noise power, depending on the output SNR, the detection performance of EL-GLRT is the same as for the EL-AMF approach, whereby the loading factor is selected based on the training sample only. Most importantly, this performance is significantly better than that of the ML-GLRT or ML-AMF detectors (the loss factor improved to 1–1.5 dB below the clairvoyant case for $P_D = 0.5$ and $P_{FA} = 10^{-2}$–10^{-4} compared with 5 dB for the standard GLRT or ACE techniques), and again is practically the same in this case as for robust selection of the constant loading factor ($\beta = 3\sigma_n^2$) for the LAMF technique.

These results demonstrate that our EL criterion for the proper families (diagonally loaded, finite interference rank) gives a significant improvement in detection performance compared with the ML criterion, which for small sample support produces solutions far away from the exact ones. We emphasize that the introduced families include the standard (unconstrained) ML covariance matrix estimate, while the major distinction stems from the attempt to get a statistically close LR to that of the exact covariance matrix, rather than just the (ultimate) maximum LR value. This is an important distinction from some optimum search over a restricted set of covariance matrices, such as the class of Toeplitz covariance matrices. Any reliable a priori structural information on the covariance matrix should always lead to a detection improvement, however we chose the most generic families specifically to underline the difference in criteria (EL versus ML), rather than any possible difference in covariance matrix description.

More specifically, this approach allowed us to generate solutions that for favorable scenarios and

LSMI/FML applications demonstrated the same detection performance as the constant-loaded LAMF detector. The fact that in this case the modified 2S EL-GLRT and EL-AMF framework had the same performance may be treated as an additional justification for the conventional LSMI/FML technique that is now used in the adaptive detectors. Indeed, we believe that the surprisingly accurate coincidence of ROCs for the three quite different techniques as 2S EL-GLRT, EL-AMF (with adaptively chosen loading) and conventional LAMF (with constant loading) just means that they all approach the best possible performance that is set by the adaptive detection problem formulation itself.

For favorable scenarios, it has been known since [8, 9] that the loading factor can be robustly selected so that the SNR performance of the LSMI adaptive filter does not depend on any particular loading value. Now we have demonstrated that this property extends to include LAMF detection performance under the condition that the false-alarm threshold is adjusted to the selected loading factor.

The other simulation scenario that we considered was specifically selected to not have this favorable property, with its full-rank "spread" interference sources and relatively small separation between signal- and noise-subspace eigenvalues. We demonstrated that for a given sample support, a constant loading factor can still be selected that makes LAMF performance as good as EL-AMF performance with its data-dependent loading factor. Both these detectors are significantly better than the standard AMF (ACE) technique, even though at the expense of losing the CFAR property. Yet, the loading factor for this scenario must be chosen to be comparable to the maximum eigenvalue, unlike our previous favorable scenario whose optimum loading was comparable to the white-noise power (minimum eigenvalue). It is evident that some scenario recognition needs to be considered in order to avoid erroneous loading-factor selection. From this viewpoint, the EL-AMF method that has a universal LR fitting for adaptive loading-factor selection based only on the training data has an important practical advantage.

Finally, we demonstrated that for favorable scenarios, our new EL-AMF and 2S EL-GLRT detectors enjoy the CFAR property, similarly to the LAMF detector. The output target-free statistics for these detectors can be approximated with high accuracy by white noise-only data. The pdfs of these statistics are functions of sample support N, filter (antenna) dimension M, number of dominant eigenvalues m, the lower and upper LR bounds α_L, α_U (for EL-GLRT), and the median LR value (for EL-AMF). While accurate analytic expressions for these pdfs have not been derived, we demonstrated that white-noise Monte-Carlo simulations can be used to give sufficiently accurate false-alarm threshold values.

Note that the framework of our 2S-GLRT approach naturally includes any restriction on the admissible covariance matrix, unlike the traditional GLRT method. Obviously, any proper restrictions include the exact covariance matrix in the admissible set, and so all the introduced scenario-free pdfs and associated bounds remain invariant to such a restriction. For example, autoregressive interference modeling [3] fits into our scheme.

Several important issues, such as the application to non-Gaussian (SIRP) models have not been addressed in this paper. These and many other important facets of the introduced methodology should be addressed in the future.

APPENDIX A. LR DISTRIBUTION FOR COMPLEX GAUSSIAN DATA

Let $\hat{C} \sim \mathcal{CW}(N,M,I_M)$ where

$$\mathcal{CW}(N,M,I_M) = \frac{1}{K(N,M)} \det[\hat{C}^{N-M}] \exp[-\mathrm{tr}\hat{C}] \tag{159}$$

$$K(N,M) \equiv \pi^{M(M-1)/2} \prod_{j=1}^{M} \Gamma(N+1-j). \tag{160}$$

Let us find the hth moment $\mathcal{E}\{\gamma^h\}$ of the LR γ ($= \gamma_0^{(1)}$ in (48))

$$\gamma = \frac{\exp(M)\det(\hat{C})}{\exp[\mathrm{tr}(\hat{C}/N)]N^M}. \tag{161}$$

Correspondingly,

$$\mathcal{E}\{\gamma^h\} = \frac{\exp(Mh)}{N^{Mh}K(N,M)} \int_{\hat{C}} [\det(\hat{C})]^{N+h-M} \times \exp\left[-\left(1+\frac{h}{N}\right)\mathrm{tr}(\hat{C})\right] d\hat{C}. \tag{162}$$

Note that

$$\mathcal{CW}\left(N+h,M,\left(1+\frac{h}{N}\right)^{-1} I_M\right) = \frac{\left(1+\frac{h}{N}\right)^{M(N+h)}}{K(N+h,M)} [\det(\hat{C})]^{N+h-M} \exp\left[-\left(1+\frac{h}{N}\right)\mathrm{tr}(\hat{C})\right]. \tag{163}$$

Therefore

$$\mathcal{E}\{\gamma^h\} = \left(\frac{e}{N}\right)^{Mh} \frac{\prod_{j=1}^{M}\Gamma(N+h+1-j)}{\prod_{j=1}^{M}\Gamma(N+1-j)} \cdot \frac{(N+h)^{-M(N+h)}}{N^{-M(N+h)}} \tag{164}$$

or

$$\mathcal{E}\{\gamma^h\} = N^{MN} e^{Mh} \frac{\prod_{j=1}^{M}\Gamma(N+h+1-j)}{\prod_{j=1}^{M}\Gamma(N+1-j)} \cdot \frac{1}{(N+h)^{M(N+h)}}. \tag{165}$$

According to Mellin's transformation, we get

$$\omega(\gamma) = \frac{N^{MN}}{\prod_{j=1}^{M}\Gamma(N+1-j)} \cdot \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} \frac{\gamma^{-h-1}\prod_{j=1}^{M}\Gamma(N+h+1-j)\exp(Mh)}{(N+h)^{M(N+h)}}dh. \tag{166}$$

Putting $N+h=t$, we get

$$\omega(\gamma) = \left(\frac{e}{N}\right)^{-MN}\frac{\gamma^{N-1}}{\prod_{j=1}^{m}\Gamma(N+j-1)} \times \int_{N-i\infty}^{N+i\infty}\left(\frac{e}{t}\right)^{Mt}\gamma^{-t}\prod_{j=1}^{M}\Gamma(t+1-j)dt. \tag{167}$$

Let us investigate the integral

$$J = \frac{1}{2\pi i}\int_{N-i\infty}^{N+i\infty}\left(\frac{e}{t}\right)^{Mt}\gamma^{-t}\prod_{j=1}^{M}\Gamma(t+1-j)dt. \tag{168}$$

This gamma function has the expansion [38]:

$$\log\Gamma(t+1-j) = \frac{1}{2}\log(2\pi) + (t+1-j)\log t - t - \sum_{r=1}^{m}\frac{(-1)^r B_{r+1}(1-j)}{r(r+1)t^r} + R_{m+1}(t) \tag{169}$$

where $|R_m(t)| \le \Theta|t^m|$, and $B_r(h)$ is the Bernoulli polynomial of degree r and order one.

Correspondingly,

$$\sum_{j=1}^{M}\log\Gamma(t+1-j) \approx \frac{M}{2}\log(2\pi) + \log t\sum_{j=1}^{M}(t+1-j) - Mt - \sum_{r=1}^{m}\frac{1}{t^r}\sum_{j=1}^{M}\frac{(-1)^r B_{r+1}(1-j)}{r(r+1)} + \cdots. \tag{170}$$

Let us define

$$\sum_{j=1}^{M}\frac{(-1)^r B_{r+1}(1-j)}{r(r+1)} = Q_r \tag{171}$$

and since

$$\sum_{j=1}^{M}(t+1-j) = \sum_{j=0}^{M}(t-j) = \frac{M}{2}(2t-M-1)$$

we get

$$\sum_{j=1}^{M}\log\Gamma(t+1-j) \approx \frac{M}{2}\log(2\pi) + Mt\log t - \frac{M(M+1)}{2}\log t - Mt - \sum_{r=1}^{m}\frac{Q_r}{t^r} + \cdots. \tag{172}$$

Correspondingly, we get the following expression for the integral J:

$$J = (2\pi)^{M/2}\frac{1}{2\pi i}\int_{N-i\infty}^{N+i\infty} t^{-M(M+1)/2}\gamma^{-t}\left(1+\sum_{r=1}^{\infty}\frac{B_r}{t^r}\right)dt \tag{173}$$

where the following expansion has been used:

$$\exp\left(-\sum_{r=1}^{m}\frac{Q_r}{t^r}\right) = 1 - \sum_{r=1}^{\infty}\frac{Q_r}{t^r} + \frac{1}{2!}\left(\sum_{r=1}^{\infty}\frac{Q_r}{t^r}\right)^2 - \frac{1}{3!}\left(\sum_{r=1}^{\infty}\frac{Q_r}{t^r}\right)^3 + \cdots = 1 + \sum_{r=1}^{\infty}\frac{B_r}{t^r}. \tag{174}$$

Let us calculate the integral J_2

$$J_2 = \sum_{r=1}^{\infty}B_r\frac{1}{2\pi i}\int_{N-i\infty}^{N+i\infty}t^{-(\nu+r)}\gamma^{-t}dt. \tag{175}$$

Note that $\nu = M(M+1)/2$ is always an integer. Therefore

$$\frac{1}{2\pi i}\int_{N-i\infty}^{N+i\infty}t^{-(\nu+r)}\gamma^{-t}dt = \mathrm{res}[t^{-(\nu+r)}\gamma^{-t}]_{t=0} \tag{176}$$

where

$$\mathrm{res}[t^{-(\nu+r)}\gamma^{-t}] = \frac{1}{(\nu+r-1)!}\frac{d^{\nu+r-1}}{\delta^{\nu+r-1}}\gamma^{-t}\Big|_{t=0} = \frac{1}{\Gamma(\nu+r)}(-\log\gamma)^{\nu+r-1}. \tag{177}$$

Therefore

$$J_2 = \sum_{r=1}^{\infty}B_r\frac{(-\log\gamma)^{\nu+r-1}}{\Gamma(\nu+r)}, \qquad B_0 = 1 \tag{178}$$

and finally we get

$$\omega(\gamma) = \left(\frac{e}{N}\right)^{-MN}\frac{\gamma^{N-1}(2\pi)^{\frac{M}{2}}}{\prod_{j=1}^{M}\Gamma(N+j-1)}\sum_{r=0}^{\infty}\frac{(-\log\gamma)^{\nu+r-1}}{\Gamma(\nu+r)}B_r, \qquad B_0 = 1, \quad \nu = \tfrac{1}{2}M(M+1). \tag{179}$$

Comparing this with [38, eq. (2.13)], we can conclude that the serial part (the integral J_2) is the same as per the real random value case. Therefore, further modifications suggested in [38] for computational convenience could be applied.

APPENDIX B. PROOF OF THEOREM 1

Let us express the favorable exact interference covariance matrix as

$$R_0 \equiv \mu U_s\Lambda_s U_s^H + U_n U_n^H, \qquad \mu \gg 1, \quad \Lambda_s > I_m \tag{180}$$

and investigate the LR (122)

$$\hat{t}_1 \equiv \frac{|Y^H(\beta I_M + \hat{R}_N)^{-1}S|^2}{S^H(\beta I_M + \hat{R}_N)^{-1}S} \tag{181}$$

$$\hat{R}_N = X_N X_N^H/N, \qquad X_N \sim \mathcal{CN}_N(0,R_0), \quad Y \sim \mathcal{CN}(0,R_0). \tag{182}$$

Since [45]

$$Y \equiv R_0^{1/2}Y_1, \qquad X_N \equiv R_0^{1/2}Z_N, \qquad Z_N \sim \mathcal{CN}_N(0,I_M), \quad Y_1 \sim \mathcal{CN}(0,I_M) \tag{183}$$

we have

$$\hat{t}_1 = \frac{|Y_1^H(\beta R_0^{-1} + Z_N Z_N^H/N)^{-1}R_0^{-1/2}S|^2}{S^H R_0^{-1/2}(\beta R_0^{-1} + Z_N Z_N^H/N)^{-1}R_0^{-1/2}S}. \tag{184}$$

Let

$$R_0 = U\Lambda U^H \qquad UU^H = U^H U = I_M \tag{185}$$

then since $Y_1 =^p U^H Y_1$ and $Z_N =^p U^H Z_N$ (where "$=^p$" means statistically equivalent), we get

$$\hat{t}_1 = \frac{|Y_1^H(\beta\Lambda^{-1} + Z_N Z_N^H/N)^{-1}\Lambda^{-1/2}U^H S|^2}{S^H U\Lambda^{-1/2}(\beta\Lambda^{-1} + Z_N Z_N^H/N)^{-1}\Lambda^{-1/2}U^H S}. \tag{186}$$

Let us present $\beta\Lambda^{-1}$ in block form and use (180):

$$\beta\Lambda^{-1} \equiv \begin{bmatrix} \beta/(\mu\Lambda_s) & 0 \\ 0 & \beta I_n \end{bmatrix} \qquad Y_1 \equiv \begin{bmatrix} Y_{1s} \\ Y_{1n} \end{bmatrix} \tag{187}$$

where

$$\beta/(\mu\Lambda_s) \ll I_m. \tag{188}$$

For this reason, the kernel in (186) can be approximated as

$$(\beta\Lambda^{-1} + Z_N Z_N^H/N)^{-1} \approx \begin{bmatrix} Z_s Z_s^H/N & Z_s Z_n^H/N \\ Z_n Z_s^H/N & \beta I_n + Z_n Z_n^H/N \end{bmatrix}^{-1} \tag{189}$$

where

$$Z_n \equiv \begin{bmatrix} Z_s \\ Z_n \end{bmatrix}, \qquad Z_s \sim \mathcal{CN}_N(0,I_m), \quad Z_n \sim \mathcal{CN}_N(0,I_n) \tag{190}$$

and $n \equiv M - m$. Due to (188), we may also consider the approximation

$$\Lambda^{-1/2}U^H S = \begin{bmatrix} \Lambda_s/\mu & 0 \\ 0 & \beta I_n \end{bmatrix} \begin{bmatrix} U_s^H \\ U_n^H \end{bmatrix} S \approx \begin{bmatrix} 0 \\ U_n^H S \end{bmatrix} \tag{191}$$

hence the test statistics are now

$$\hat{t}_1 \approx \frac{\left| [Y_{1s}^H \mid Y_{1n}^H] \begin{bmatrix} Z_s Z_s^H/N & Z_s Z_n^H/N \\ Z_n Z_s^H/N & \beta I_n + Z_n Z_n^H/N \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ U_n^H S \end{bmatrix} \right|^2}{[0 \mid S^H U_n] \begin{bmatrix} Z_s Z_s^H/N & Z_s Z_n^H/N \\ Z_n Z_s^H/N & \beta I_n + Z_n Z_n^H/N \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ U_n^H S \end{bmatrix}}. \tag{192}$$

Let us now introduce

$$\frac{U_n S}{(S^H U_n U_n^H S)^{1/2}} \equiv V_n \mathbf{e}_1, \qquad V_n V_n^H = V_n^H V_n = I_n \tag{193}$$

where $\mathbf{e}_1$ is the unit vector. First, by using the block-matrix inversion formula [46], we get

$$\hat{t}_1 \approx \frac{|[Y_{1n}^H - Y_{1s}^H(Z_s Z_s^H)^{-1}Z_s Z_n^H]\{\beta I_n + N^{-1}Z_n[I_N - Z_s^H(Z_s Z_s^H)^{-1}Z_s]Z_n^H\}^{-1}U_n S|^2}{S^H U_n\{\beta I_n + N^{-1}Z_n[I_N - Z_s^H(Z_s Z_s^H)^{-1}Z_s]Z_n^H\}^{-1}U_n S}. \tag{194}$$

Since $Y_{1n} =^p V_n Y_{1n}$ and $Z_n =^p V_n Z_n$, we finally get

$$\hat{t}_1 \approx \frac{|[Y_{1n}^H - Y_{1s}^H(Z_s Z_s^H)^{-1}Z_s Z_n^H]\{\beta I_n + N^{-1}Z_n[I_N - Z_s^H(Z_s Z_s^H)^{-1}Z_s]Z_n^H\}^{-1}\mathbf{e}_1|^2}{\mathbf{e}_1^T\{\beta I_n + N^{-1}Z_n[I_N - Z_s^H(Z_s Z_s^H)^{-1}Z_s]Z_n^H\}^{-1}\mathbf{e}_1} \tag{195}$$

which is the same as (123).

Similarly to (187) and (189), we get

$$Y^H(\beta I_M + \hat{R}_N)^{-1}Y \approx [Y_{1s}^H \mid Y_{1n}^H] \begin{bmatrix} Z_s Z_s^H/N & Z_s Z_n^H/N \\ Z_n Z_s^H/N & \beta I_n + Z_n Z_n^H/N \end{bmatrix} \begin{bmatrix} Y_{1s} \\ Y_{1n} \end{bmatrix}. \tag{196}$$

By the block-matrix inversion formula,

$$\begin{aligned} &Y^H(\beta I_M + \hat{R}_N)^{-1}Y \\ &\approx Y_{1s}^H[N^{-1}Z_s Z_s - N^{-2}Z_s Z_n^H(\beta I_n + N^{-1}Z_n Z_n^H)^{-1}Z_n Z_s^H]^{-1}Y_{1s} \\ &\quad - 2\Re[Y_{1s}^H(Z_s Z_s^H)^{-1}Z_s Z_n^H \\ &\quad \times \{\beta I_N + N^{-1}Z_n[I_N - Z_s^H(Z_s Z_s^H)^{-1}Z_s]Z_n^H\}^{-1}]Y_{1n} \\ &\quad + Y_{1n}^H\{\beta I_n + N^{-1}Z_n[I_N - Z_s^H(Z_s Z_s^H)^{-1}Z_s]Z_n^H\}^{-1}Y_{1n} \end{aligned} \tag{197}$$

and then it is seen that Y_{1n} and Z_n may be substituted by $V_n Y_{1n}$ and $V_n Z_n$, respectively, without changing the formula.

APPENDIX C. PROOF OF THEOREM 2

Under the same favorable conditions as (180), let us consider the LR

$$\gamma_{\ell 0}^{(1)} \equiv \frac{\det[\hat{R}_{\mathrm{LSMI}}^{-1}\hat{R}]\exp M}{\exp \mathrm{tr}[\hat{R}_{\mathrm{LSMI}}^{-1}\hat{R}]} \tag{198}$$

where $\hat{R}_{\mathrm{LSMI}} = \beta I_M + \hat{R}_N$.

Firstly,

$$\det[\hat{R}_{\mathrm{LSMI}}^{-1}\hat{R}] \approx \det\left(\begin{bmatrix} Z_s Z_s^H/N & Z_s Z_n^H/N \\ Z_n Z_s^H/N & \beta I_n + Z_n Z_n^H/N \end{bmatrix}^{-1} \begin{bmatrix} Z_s Z_s^H/N & Z_s Z_n^H/N \\ Z_n Z_s^H/N & Z_n Z_n^H/N \end{bmatrix}\right). \tag{199}$$

By the block-matrix inversion formula [46], we have

$$\begin{aligned}\det[\hat{R}_{\mathrm{LSMI}}^{-1}\hat{R}] &\approx \{\beta I_n + N^{-1} Z_n [I_N - Z_s^H (Z_s Z_s^H)^{-1} Z_s] Z_n^H\}^{-1} \\ &\quad \times N^{-1} Z_n [I_N - Z_s^H (Z_s Z_s^H)^{-1} Z_s] Z_n^H\end{aligned} \tag{200}$$

and by the approximation (199):

$$\begin{aligned}&\mathrm{tr}[\hat{R}_{\mathrm{LSMI}}^{-1}\hat{R}] \\ &\quad \approx \mathrm{tr}[I_M] - \beta\,\mathrm{tr}\{\beta I_n + N^{-1} Z_n [I_N - Z_s^H (Z_s Z_s^H)^{-1} Z_s] Z_n^H\}\end{aligned} \tag{201}$$

since the product $\hat{R}_{\mathrm{LSMI}}^{-1}\hat{R}$ can be expressed as

$$\hat{R}_{\mathrm{LSMI}}^{-1}\hat{R} = I_M - \begin{bmatrix} Z_s Z_s^H/N & Z_s Z_n^H/N \\ Z_n Z_s^H/N & \beta I_n + Z_n Z_n^H/N \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \beta I_n \end{bmatrix}. \tag{202}$$

Again, Z_n here can be replaced by $V_n Z_n$ without any change in the formulas.

APPENDIX D. ALTERNATIVE ROC DERIVATION

Consider the pdf for the AMF statistics

$$\lambda = \frac{|Y^H \hat{R}^{-1} S|^2}{S^H \hat{R}^{-1} S} \tag{203}$$

in the case when the target signal is absent. Since [1]

$$\hat{R} = R^{1/2} \hat{C} R^{1/2}, \qquad \hat{C} \sim \mathcal{CW}(N, M, I_M) \tag{204}$$

and

$$Y = R^{1/2}\xi, \qquad \xi \sim \mathcal{CN}(0, I_M) \tag{205}$$

$$R^{-1/2} S = (S^H R^{-1} S)^{-1/2} U^H \mathbf{e}_1 \tag{206}$$

where $U^H U = U U^H = I_M$ and $\mathbf{e}_1 \equiv [1, 0, \ldots, 0]^T$, we obtain

$$\lambda = \frac{|\mathbf{e}_1^T U \hat{C}^{-1} \xi|^2}{\mathbf{e}_1^T U \hat{C}^{-1} U^H \mathbf{e}_1}. \tag{207}$$

Since $U\xi \sim \xi$ and $U\hat{C}^{-1}U^H \sim \hat{C}^{-1}$ (statistical equivalence), we get

$$\lambda = \frac{|\mathbf{e}_1^T \hat{C}^{-1} \xi|^2}{\mathbf{e}_1^T \hat{C}^{-1} \mathbf{e}_1}. \tag{208}$$

Using matrix partitioning, similarly to [1]:

$$\hat{C} = \begin{bmatrix} C_{11} & C_{12}^H \\ C_{12} & C_{22} \end{bmatrix}, \qquad \hat{C}^{-1} = \begin{bmatrix} C^{11} & C^{12H} \\ C^{12} & C^{22} \end{bmatrix} \tag{209}$$

where

$$C^{11} \equiv [C_{11} - C_{12}^H C_{22}^{-1} C_{12}]^{-1} \tag{210}$$

and

$$\begin{aligned} C^{12} &\equiv -\left[C_{22} - \frac{C_{12} C_{12}^H}{C_{11}}\right]^{-1} \frac{C_{12}}{C_{11}} \\ &= -\frac{C_{22}^{-1} C_{12}}{[C_{11} - C_{12}^H C_{22}^{-1} C_{12}]^{-1}} \end{aligned} \tag{211}$$

we get

$$\lambda = \frac{|\xi^H [1, -C_{22}^{-1} C_{12}]|^2}{C_{11} - C_{12}^H C_{22}^{-1} C_{12}}. \tag{212}$$

In [1] it was demonstrated that

$$D_{11} \equiv C_{11} - C_{12}^H C_{22}^{-1} C_{12} \qquad \text{and} \qquad E_{12} \equiv C_{22}^{-1} C_{12} \tag{213}$$

are independent, and that D_{11} has a chi-square distribution with $2(N - M + 1)$ degrees of freedom:

$$\begin{aligned} f(D_{11}) &= K_1 D_{11}^{N-M} \exp D_{11} \\ K_1 &\equiv \frac{1}{\Gamma(N - M + 1)} = \frac{1}{(N-M)!} \end{aligned} \tag{214}$$

and

$$f(E_{12}) = \frac{K_4}{[1 + E_{12}^H E_{12}]^{N+1}}, \qquad K_4 \equiv \frac{N!}{\pi^{M-1}(N - M + 1)!} \tag{215}$$

i.e., $f(E_{12})$ is described by a multivariate t-distribution.

Since ξ and E_{12} are mutually independent, the E_{12}-conditional pdf of the product in (212) is Gaussian zero with power $[1 + E_{12}^H E_{12}]$. Hence

$$\lambda(D_{11}, E_{12}) = \frac{1}{D_{11}} |\xi(E_{12}|^2 \xi \sim \mathcal{CN}(0, [1 + E_{12}^H E_{12}]). \tag{216}$$

Now since

$$|\xi(E_{12}|^2 \sim \frac{1}{[1+E_{12}^H E_{12}]}\exp\left(-\frac{x}{[1+E_{12}^H E_{12}]}\right) \tag{217}$$

(chi-square pdf), and due to the β-distribution derived in [1]:

$$\frac{1}{[1+E_{12}^H E_{12}]} \sim \frac{N!}{(N-M+1)!(M-2)!}\rho^{N+1-M}(1-\rho)^{M-2} \tag{218}$$

we finally get

$$f(|\xi|^2) = \frac{N!}{(N-M+1)!(M-2)!} \times \int_0^\infty \rho^{N+2-M}(1-\rho)^{M-2}e^{-\rho x}d\rho \tag{219}$$

which is equal to [38]

$$f(|\xi|^2) = \frac{N-M+2}{N+1}{}_1F_1(N-M+3;N+2,-x). \tag{220}$$

Now we have to find the pdf of the quantity $\lambda = |\xi|^2/D_{11}$:

$$f(\lambda) = \frac{N-M+2}{N+1}\frac{1}{(N-M)!} \times \int_0^\infty {}_1F_1(N-M+3;N+2,-u\lambda)u^{N-M+1}e^{-u}du. \tag{221}$$

According to [38], we have

$$f(\lambda) = \frac{N-M+2}{N+1}\frac{\Gamma(N-M+2)}{(N-M)!}\frac{1}{(1+\lambda)^{N-M+2}} \times {}_2F_1\left(M-1,N-M+2,N+2;\frac{\lambda}{1+\lambda}\right). \tag{222}$$

According to [47, 15.3.5],

$${}_2F_1\left(M-1,N-M+2,N+2;\frac{\lambda}{1+\lambda}\right) = \left[\frac{1}{(1+\lambda)}\right]^{-(N-M+2)} {}_2F_1(N-M+2,N-M+3,N+2;-\lambda). \tag{223}$$

On the other hand, according to [47, 15.2.1]:

$$\frac{(N-M+2)(N-M+1)}{N+1}{}_2F_1(N-M+2,N-M+3,N+2;-\lambda) = \frac{d}{d\lambda}{}_2F_1(N-M+1,N-M+2,N+1;-\lambda) \tag{224}$$

hence

$$P_{\rm FA} = 1 - \int_0^h \left[\frac{d}{d\lambda}{}_2F_1(N-M+1,N-M+2,N+1;-\lambda)\right]d\lambda \tag{225}$$

$$= 1 - \int_{-h}^0 \left[\frac{d}{d\lambda}{}_2F_1(N-M+1,N-M+2,N+1;-\lambda)\right]d\lambda \tag{226}$$

$$= {}_2F_1(N-M+1,N-M+2,N+1;-h) \tag{227}$$

and by [47, 15.3.4],

$$P_{\rm FA} = \frac{1}{(1+h)^{N-M+1}}{}_2F_1\left(N-M+1,M-1,N+1;\frac{h}{1+h}\right) \tag{228}$$

which is our (151).

Now let us consider the pdf for the AMF statistics λ when the signal is present, whereby

$$Y = R^{1/2}\xi + aS, \qquad a \sim \mathcal{CN}(0,\sigma_s^2). \tag{229}$$

We have

$$S^H\hat{R}^{-1}Y = S^H R^{-1/2}\hat{C}^{-1}R^{-1/2}(R^{1/2}\xi + aS) = (S^H R^{-1}S)^{1/2}\mathbf{e}_1^T U\hat{C}^{-1}\xi + (S^H R^{-1}S)\mathbf{e}_1^T U\hat{C}^{-1}U^H\mathbf{e}_1 \tag{230}$$

hence

$$\lambda = \frac{|\mathbf{e}_1^T U\hat{C}^{-1}\xi + a(S^H R^{-1}S)^{1/2}\mathbf{e}_1^T U\hat{C}^{-1}U^H\mathbf{e}_1|^2}{\mathbf{e}_1^T U\hat{C}^{-1}U^H\mathbf{e}_1}. \tag{231}$$

According to (211) and (212),

$$\lambda = \frac{1}{D_{11}}|\xi^H[1,-C_{22}^{-1}C_{12}] + a(S^H R^{-1}S)^{1/2}|^2. \tag{232}$$

Similarly, we get

$$\lambda(D_{11},E_{12}) = \frac{1}{D_{11}}|\xi(E_{12},\sigma_{\rm out}^2)|^2 \tag{233}$$

where

$$\sigma_{\rm out}^2 \equiv \sigma_s^2 S^H R^{-1}S, \qquad \xi(E_{12},\sigma_{\rm out}^2) \sim \mathcal{CN}(0,[1+\sigma_{\rm out}^2 + E_{12}^H E_{12}]). \tag{234}$$

Again, the E_{12}-conditional pdf for $|\xi|^2$ is

$$f[|\xi(E_{12},\sigma_{\rm out}^2)|^2] = \frac{1}{[1+\sigma_{\rm out}^2+E_{12}^H E_{12}]} \times \exp\left[-\frac{x}{1+\sigma_{\rm out}^2+E_{12}^H E_{12}}\right]. \tag{235}$$

Since

$$P_{\rm D} = {\rm Prob}\{|\xi(E_{12},\sigma_{\rm out}^2)|^2 > hD_{11}\} \tag{236}$$

the D_{11}, E_{12}-conditional pdf is

$$P_{\mathrm{D}}(E_{12}, \sigma^2_{\mathrm{out}}, D_{11}) = \exp\left[-\frac{hD_{11}}{1+\sigma^2_{\mathrm{out}}+E^H_{12}E_{12}}\right]. \tag{237}$$

We now need to integrate over the pdf for D_{11} and E_{12}. Integrating (214) gives us

$$P_{\mathrm{D}}(1+\sigma^2_{\mathrm{out}}+E^H_{12}E_{12}) = \left[1+\frac{h}{1+\sigma^2_{\mathrm{out}}+E^H_{12}E_{12}}\right]^{-(N-M+1)}. \tag{238}$$

Using (218), it is straight forward to show that

$$f\left[x=\frac{1}{1+\sigma^2_{\mathrm{out}}+E^H_{12}E_{12}}\right] = \frac{N!}{(N-M+1)!(M-2)!}\frac{x^{N+1-M}[1-(1+\sigma^2_{\mathrm{out}})x]^{M-2}}{(1-x\sigma^2_{\mathrm{out}})^{N+1}} \tag{239}$$

for $0 \le x \le 1/(1+\sigma^2_{\mathrm{out}})$. Therefore

$$P_{\mathrm{D}}(\sigma^2_{\mathrm{out}}) = \frac{N!}{(N-M+1)!(M-2)!} \times \int_0^{1/1+\sigma^2_{\mathrm{out}}} \frac{x^{N+1-M}[1-(1+\sigma^2_{\mathrm{out}})x]^{M-2}}{(1+xh)^{N-M+1}(1-x\sigma^2_{\mathrm{out}})^{N+1}}dx \tag{240}$$

then according to [38],

$$P_{\mathrm{D}} = \left[\frac{1+\sigma^2_s S^H R^{-1} S}{1+\sigma^2_s S^H R^{-1} S+h}\right]^{N-M+1} \times F_1\left(M-1, -(N-M+1), N-M+1, N+1; \frac{\sigma^2_s S^H R^{-1} S}{1+\sigma^2_s S^H R^{-1} S}, \frac{\sigma^2_s S^H R^{-1} S+h}{1+\sigma^2_s S^H R^{-1} S+h}\right) \tag{241}$$

which is our (152).

REFERENCES

[1] Reed, I., Mallett, J., and Brennan, L.
Rapid convergence rate in adaptive arrays.
IEEE Transactions on Aerospace and Electronic Systems, **10**, 6 (Nov. 1974), 853–863.

[2] Kelly, E.
An adaptive detection algorithm.
IEEE Transactions on Aerospace and Electronic Systems, **22**, 1 (Mar. 1986), 115–127.

[3] Michels, J., Rangaswamy, M., and Himed, B.
Performance of parametric and covariance based STAP tests in compound-Gaussian clutter.
Digital Signal Processing, **12** (2002), 307–388.

[4] Conte, E., Lops, M., and Ricci, G.
Asymptotically optimum radar detection in compound-Gaussian clutter.
IEEE Transactions on Aerospace and Electronic Systems, **31** (Apr. 1995), 611–616.

[5] Conte, E., Lops, M., and Ricci, G.
Adaptive matched filter detection in spherically invariant noise.
IEEE Signal Processing Letters, **3** (1996), 248–250.

[6] Finn, H., and Johnson, R.
Adaptive detection mode with threshold control as a function of spatially sampled clutter level estimates.
RCA Review, **29** (Sept. 1968), 414–463.

[7] Alpargu, G., and Styan, G.
Some remarks and a bibliography on the Kantorovich inequality.
In *Proceedings of the Sixth Lukacs Symposium*, 1996, 1–13.

[8] Abramovich, Y.
A controlled method for adaptive optimization of filters using the criterion of maximum SNR.
Radio Engineering and Electronic Physics, **26**, 3 (1981), 87–95.

[9] Abramovich, Y., and Nevrev, A.
An analysis of effectiveness of adaptive maximization of the signal-to-noise ratio which utilizes the inversion of the estimated correlation matrix.
Radio Engineering and Electronic Physics, **26**, 12 (1981), 67–74.

[10] Cheremisin, O.
Efficiency of adaptive algorithms with regularised sample covariance matrix.
Radio Engineering and Electronic Physics, **27**, 10 (1982), 69–77.

[11] Cheremisin, O.
Loading factor selection in the regularised algorithm for adaptive filter optimisation.
Radioteknika i Elektronika, **30**, 12 (1985). English translation should be found in *Soviet Journal of Communication Technology and Electronics*.

[12] Cox, H., Zeskind, R., and Owen, M.
Robust adaptive beamforming.
IEEE Transactions on Acoustics, Speech, and Signal Processing, **35** (1987), 1365–1376.

[13] Carlson, B.
Covariance matrix estimation errors and diagonal loading in adaptive arrays.
IEEE Transactions on Aerospace and Electronic Systems, **24**, 7 (July 1988), 397–401.

[14] Vorobyov, S., Gershman, A., and Zhi-Quan, L.
Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem.
IEEE Transactions on Signal Processing, **51**, 2 (Feb. 2003), 313–324.

[15] Li, J., Stoica, P., and Wang, Z.
On robust Capon beamforming and diagonal loading.
IEEE Transactions on Signal Processing, **51**, 2 (Feb. 2003), 1702–1715.

[16] Shahbazpanahi, S., Gershman, A., Luo, Z.-Q., and Wong, K.
Robust adaptive beamforming for general-rank signal models.
IEEE Transactions on Signal Processing, **51**, 9 (Sept. 2003), 2257–2269.

[17] Robey, F., Fuhrmann, D., Kelly, E., and Nitzberg, R.
A CFAR adaptive matched filter detector.
IEEE Transactions on Aerospace and Electronic Systems, **28**, 1 (Jan. 1992), 208–216.

[18] Kraut, S., and Scharf, L.
The CFAR adaptive subspace detector is a scale-invariant GLRT.
IEEE Transactions on Signal Processing, **47**, 9 (Sept. 1999), 2538–2541.

[19] Kraut, S., Scharf, L., and McWhorter, L.
Adaptive subspace detectors.
IEEE Transactions on Signal Processing, **49**, 1 (Jan. 2001), 1–16.

[20] Besson, O., Scharf, L., and Vincent, F.
Matched direction detectors and estimators for array processing with subspace steering vector uncertainties.
IEEE Transactions on Signal Processing, **53**, 12 (Dec. 2005), 4453–4463.

[21] Liu, B., Chen, B., and Michels, J.
A GLRT for multichannel radar detection in the presence of both SIRP clutter and additive white Gaussian noise.
IEEE Transactions on Signal Processing, to be published.

[22] Gini, F., Greco, M., and Farina, A.
Clairvoyant and adaptive signal detection in non-Gaussian clutter: a data-dependent threshold interpretation.
IEEE Transactions on Signal Processing, **47**, 6 (June 1999), 1522–1531.

[23] Conte, E., De Maio, A., and Galdi, C.
CFAR detection of multidimensional signals: an invariant approach.
IEEE Transations on Signal Processing, **51**, 1 (Jan. 2003), 142–151.

[24] Jin, Y., and Friedlander, B.
A CFAR adaptive subspace detector for second-order Gaussian signals.
IEEE Transactions on Signal Processing, **53**, 3 (Mar. 2005), 871–884.

[25] Lekhovitskii, D., and Bondarenko, M.
Choice of a learning sample in adaptive detectors of signals against the background of Gaussian interferences.
In *Proceedings of CAMSAP-2005*, Puerto Vallarta, Mexico, 2005.

[26] Chen, P., Melvin, W., and Wicks, M.
Screening among multivariate normal data.
Journal of Multivariate Analysis, **69** (1999), 10–29.

[27] Melvin, W.
Space-time adaptive radar performance in heterogeneous clutter.
IEEE Transactions on Aerospace and Electronic Systems, **36**, 2 (Apr. 2000), 621–633.

[28] Rangaswamy, M., Himed, B., and Michels, J.
Statistical analysis of the nonhomogeneity detector.
In *Proceedings of ASILOMAR-2000*, vol. 2, Pacific Grove, CA, 2000, 117–1121.

[29] Rangaswamy, M., Himed, B., and Michels, J.
Performance analysis of the nonhomogeneity detector for STAP applications.
In *Proceedings of IEEE RADAR-2001*, Atlanta, GA, 2001, 193–197.

[30] Gerlach, K.
Outlier resistant adaptive matched filtering.
IEEE Transactions on Aerospace and Electronic Systems, **38**, 3 (July 2002), 885–901.

[31] Porat, B.
Digital Processing of Random Signals (5th ed.).
Englewood Cliffs, NJ: Prentice-Hall, 1994.

[32] Muirhead, R.
Aspects of Multivariate Statistical Theory.
New York: Wiley, 1982.

[33] Anderson, T.
An Introduction to Multivariate Statistical Analysis.
New York: Wiley, 1958.

[34] Nagarsenker, B., and Pillai, K.
Distribution of the likelihood ratio criterion for testing a hypothesis specifying a covariance matrix.
Biometrica, **60**, 2 (1973), 359–361.

[35] Blake, L.
Prediction of Radar Range.
In M. Skolnik (Ed.), *Radar Handbook* (2nd ed.), New York: McGraw-Hill, 1990.

[36] McWhorter, L., Scharf, L., and Griffiths, L.
Adaptive coherence estimation for radar signal processing.
In *Proceedings of ASILOMAR-96*, vol. 1, Pacific Grove, CA, 1996, 536–540.

[37] Abramovich, Y., Spencer, N., and Gorokhov, A.
Bounds on maximum likelihood ratio—Part I: Application to antenna array detection-estimation with perfect wavefront coherence.
IEEE Transactions on Signal Processing, **52**, 6 (June 2004), 1524–1536.

[38] Gradshteyn, I., and Ryzhik, I.
Tables of Integrals, Series, and Products (6th ed.).
New York: Academic Press, 2000.

[39] Gierull, C.
Performance analysis of fast projections of the Hung–Turner type for adaptive beamforming.
Signal Processing (special issue on Subspace Methods, Part I: Array Signal Processing and Subspace Computations), **50**, 1 (1996), 17–28.

[40] Djurić, P.
A model selection rule for sinusoids in white Gaussian noise.
IEEE Transactions on Signal Processing, **44**, 7 (July 1996), 1744–1757.

[41] Klemm, R.
Space-Time Adaptive Processing: Principles and Applications.
UK: IEE, 1998.

[42] Scharf, L.
Statistical Signal Processing.
New York: Addison-Wesley, 1991.

[43] Gershman, A., Mecklenbräucker, C., and Böhme, J.
Matrix fitting approach to direction-of-arrival estimation with imperfect spatial coherence of wavefronts.
IEEE Transactions on Signal Processing, **45**, 7 (July 1997), 1894–1899.

[44] Paulraj, A., and Kailath, T.
Direction-of-arrival estimation by eigenstructure methods with imperfect spatial coherence of wave fronts.
Journal of Acoustical Society of America, **83**, 3 (Mar. 1988), 1034–1040.

[45] Siotani, M., Hayakawa, T., and Fujikoshi, Y.
Modern Multivariate Statistical Analysis.
Ohio: Amer. Sci. Press, 1985.

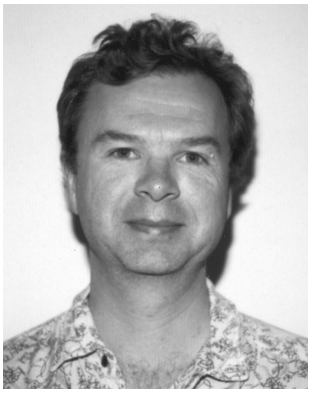
[46] Horn, R., and Johnson, C.
Matrix Analysis.
London: Cambridge University Press, 1990.

[47] Abramowitz, M., and Stegun, I.
Handbook of Mathematical Functions.
Washington, D.C.: National Bureau of Standards, 1964.

Yuri I. Abramovich (M'96—SM'06) received the Dipl.Eng. (Hons.) degree in radio electronics in 1967 and the Cand.Sci. degree (Ph.D. equivalent) in theoretical radio techniques in 1971, both from the Odessa Polytechnic University, Odessa, Ukraine, and in 1981, he received the D.Sc. degree in radar and navigation from the Leningrad Institute for Avionics, Leningrad, Russia.

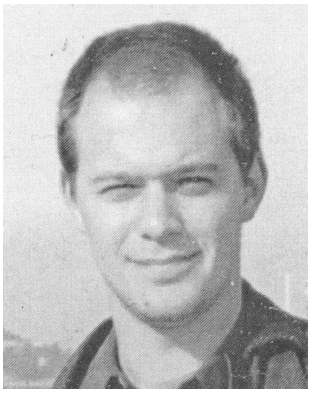
From 1968 to 1994, he was with the Odessa State Polytechnic University, Odessa, Ukraine, as a Research Fellow, professor, and ultimately as Vice-Chancellor of Science and Research. From 1994 to 2006, he was at the Cooperative Research Centre for Sensor Signal and Information Processing (CSSIP), Adelaide, Australia. Since 2000, he has been with the Australian Defence Science and Technology Organisation (DSTO), Adelaide, as principal research scientist, seconded to CSSIP until its closure. His research interests are in signal processing (particularly spatio-temporal adaptive processing, beamforming, signal detection and estimation), its application to radar (particularly over-the-horizon radar), electronic warfare, and communication.

Dr. Abramovich served as Associate Editor of *IEEE Transactions on Signal Processing* from 2002 to 2005.



Nicholas K. Spencer received the B.Sc. (Hons.) degree in applied mathematics in 1985 and the M.Sc. degree in computational mathematics in 1992, both from the Australian National University, Canberra.

He has been with the Australian Department of Defence, Canberra; the Flinders University of South Australia, Adelaide; the University of Adelaide; the Australian Centre for Remote Sensing, Canberra; and the Cooperative Research Centre for Sensor Signal and Information Processing (CSSIP), Adelaide, in the areas of computational and mathematical sciences. He is currently a senior researcher at Adelaide Research & Innovation Pty Ltd (ARI), Australia. His research interests include array signal processing, parallel and supercomputing, software best-practice, human-machine interfaces, multi-level numerical methods, modelling and simulation of physical systems, theoretical astrophysics, and cellular automata.



Alexei Gorokhov (M'96) received the Ms.D. degree in 1993 from the Odessa Polytechnic University (OPU), Odessa, Ukraine, and the Ph.D. degree in 1997 from the École Nationale Supérieure des Télécommunications, Paris, France.

From 1993 to 1994, he served as a research engineer with the research laboratory of the OPU. Since October 1997, he has been affiliated with the Centre National de la Recherche Scientifique (CNRS), Paris, France, where he holds a research position. From January 2000 to January 2004, he served as a senior scientist within the DSP group of Philips Research Laboratories, Eindhoven, The Netherlands, on long-term leave from CNRS. Since January 2004, he is with Qualcomm Inc., San Diego, CA. His general interests cover wireless communications, spectral analysis, statistical signal processing, and information theory. His current research is mainly focused on error control coding, multiuser equalisation, and antenna diversity for wireless communications.