A Matrix Extension Built Under-Sampled Likelihood Ratio Test with Application to MUSIC Breakdown Prediction and Cure

Ben A. Johnson
RLM Management, Pty. Ltd., Edinburgh, Australia
and Institute of Telecommunications Research, UniSA, Mawson Lakes, Australia
Email: ben.a.johnson@ieee.org

Yuri I. Abramovich
ISR Division, Defence Science and Technology Organisation
Email: yuri.abramovich@dsto.defence.gov.au

***Abstract*— Likelihood ratio tests are used in a range of detection-estimation problems, but normally cannot be extended to cases where training data volume T is smaller than the dimension M of the observations. We propose a non-degenerate normalized LR test that can be used for detection-estimation in such under-sampled training conditions. The LR is formed based on non-degenerate band extension of the original degenerate sample covariance matrix. This LR is then applied within a generalized likelihood ratio test framework to an array processing problem where the presence of closely spaced signal can be robustly detected, but their individual directions of arrival cannot be fully resolved by subspace-based DOA techniques such as MUSIC. In that case, MUSIC produces direction of arrival estimates for some sources with very large errors (outliers). We use the under-sampled likelihood ratio to detect the presence of such MUSIC outliers and provide corrected DOA estimates.**

***Index Terms*— maximum likelihood estimation, direction of arrival estimation, array signal processing, parameter estimation, signal resolution.**

I. INTRODUCTION

Adaptive detection-estimation problems frequently occur when the dimensionality of the observation M is significantly larger than representative training samples T. In such cases, additional *a priori* assumptions are often imposed to improve detection-estimation performance.

In cases where measured signals are sufficiently structured to occupy a finite rank within the observation covariance matrix, a number of well-known signal subspace techniques can be utilized. Specifically, when the number m of the covariance matrix eigenvalues that exceed the minimal eigenvalue (equal to ambient white noise power) is smaller then the matrix dimension M ($m < M$), we can introduce the following form for admissible covariance matrices:

$$R = \sigma_0^2 I_M + R_S;\ \ R_S = \mathcal{U}_m \Lambda_0 \mathcal{U}_m^H;\ \ \Lambda_0 = \Lambda_m - \sigma_0^2 I_m, \tag{1}$$

where $\mathcal{U}_m \in C^{M\times m}$ and $\Lambda_m \in R_+^{m\times m}$ are the $(M\times m)$-variate and $(m\times m)$-variate matrices of signal subspace eigenvectors and (positive) eigenvalues respectively.

For such low-rank covariance matrices, the minimum number of independent identically distributed (i.i.d.) training samples to localize the signal subspace is the dimension m rather than the larger observation length M.

A. Classic Low-Rank Estimation Methods

Since for many problems of interest, m is significantly smaller than M, a number of adaptive filtering techniques exploit the resultant reduced data training volume requirement, including the Hung-Turner fast projection adaptive beamformer [3], [4], and "fast maximum likelihood" [5]. In addition, the well-known loaded sample matrix inversion (LSMI) algorithm in [6], [7] uses diagonal loading in this case to regularize eigenvectors outside the signal subspace and achieve performance (under some mild eigenvalue assumptions) comparable to standard sample matrix inversion techniques, but with considerably reduced sample support requirements, including the case where the number of training samples T is less than M. It has been shown in [6], [8] that average SNR losses for the LSMI technique compared with the clairvoyant filter are equal to approximately 3dB for sample support $T \gtrsim 2m$, while for the traditional SMI technique the required sample support is equal to $T \gtrsim 2M$ for the same average loss [9].

In addition to adaptive filtering, a number of adaptive direction-of-arrival (DOA) estimation techniques exploit the covariance matrix structure given in (1). It is well known that for strong enough signal-to-noise ratio, subspace techniques such MUSIC, ESPRIT or the Minimum-

This paper is based on "A Modified Likelihood Ratio Test for Detection-Estimation in Under-Sampled Training Conditions" by Y. I. Abramovich and B. A. Johnson, which appeared in the Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Toulouse, France, May 2006 © 2006 IEEE [1] as well as "GLRT-based Detection-Estimation of Gaussian Signals in Under-Sampled Training Conditions" by B. A. Johnson and Y. I. Abramovich, which appeared in the Proceedings of the 14th European Signal Processing Conference (EUSIPCO), Florence, Italy, September, 2006 © 2006 EURASIP [2].

This work was conducted under RLM/DSTO Collaboration 290905.

Norm algorithm can provide accurate DOA estimates for cases where the number of training samples T is equal or greater than the number of independent sources m. MLE methods are also available in the under-sampled case, based on formation and maximization of a likelihood function.

B. Likelihood Ratio Formulation

For multi-variate complex Gaussian training data $x_t, t = 1, \ldots, T$, $x_t \sim \mathcal{CN}(0, R_0)$ the likelihood function w.r.t parametric description of its covariance matrix R is:

$$\mathcal{L}(X_T, R) = \left[\frac{1}{\pi \det R} \exp\{-\mathrm{Tr}[R^{-1}\hat{R}]\}\right]^M \qquad (2)$$

where

$$\hat{R} = \frac{1}{T}\sum_{j=1}^{T} \boldsymbol{x}_j \boldsymbol{x}_j^H. \qquad (3)$$

The standard normalization of the likelihood function

$$LR(R) = \frac{\mathcal{L}(X_T, R)}{\max_R \mathcal{L}(X_T, R)} \qquad (4)$$

leads to a likelihood ratio (after the methodology in [10])

$$LR(R) = \left[\frac{\det R^{-1}\hat{R} \exp M}{\exp\{\mathrm{Tr} R^{-1}\hat{R}\}}\right]^M \leqslant 1 \qquad (5)$$

since

$$\max_R \mathcal{L}(X_T, R) = \left[\frac{\exp(-M)}{\pi \det \hat{R}}\right]^M, \quad \text{for } R = \hat{R} \qquad (6)$$

While MLE can be executed on the likelihood function (2) or it's normalized likelihood ratio version (5) in the over-sampled ($T > M$) regime, only the likelihood function (2) can be maximized in the under-sampled ($T < M$) regime, since $\det \hat{R} = 0$ in that case and the normalization used in (5) is no longer available.

The lack of a classically normalized likelihood ratio in the under-sampled regime means that a wide body of likelihood ratio hypothesis testing approaches used in both detection and estimation are not available or require modification in this circumstance. In particular, a variant of generalized likelihood ratio testing (GLRT) demonstrated by one of the authors in [11] is reliant on the invariance property of $LR(R)$ (5):

$$f[LR(R_0)] = f\left[\frac{\det \hat{C} \exp M}{\exp\{\mathrm{Tr}\ \hat{C}\}}\right] \qquad (7)$$

where $\hat{C} \sim \mathcal{CW}(T \geqslant M, M, I_M)$, i.e. $\hat{C}$ is a random matrix with scenario-free complex Wishart p.d.f., fully specified by T and M [12]. In addition, since the estimate with a maximal likelihood by definition has an LR value greater than the true solution (otherwise MLE would be error-free), we can terminate any LR search process once we exceed a threshold based on the scenario-free LR p.d.f. These properties allows one to address many complicated detection-estimation problems that would not be properly addressed by conventional detection-estimation techniques [13], [14]. The key element of these approaches is the ability to compare specific likelihood ratios to bounds based on the underlying scenario-free likelihood-ratio p.d.f., rather than relying on global (and therefore potentially difficult or even impossible) maximization.

C. Likelihood Ratio Tests in the Under-Sampled Case

Obviously, for under-sampled training conditions we would like to have a similar instrument, but the standard approach cannot be used, since the sample covariance matrix is degenerate when $T < M$. For the case with under-sampled training data that belong to the family (1) we would like a likelihood ratio $LR_u(R)$ that satisfies the following conditions.

a) **Normalization condition**:

$$0 < LR_u(R) \leqslant \text{ constant} \qquad (8)$$

b) **Transition behavior**: $LR_u(R)$ should be an "analytic extension" of the $LR(R)$ (5), ie

$$LR_u(R) = LR(R) \ \text{ for } T \geqslant M \qquad (9)$$

c) **Invariance property**:

$$f\left[LR_u(R_0)\right] = f(M, T) \qquad (10)$$

Derivation of a $LR_u(R)$ that meets these requirements is introduced in Section II. In section III, we utilize the undersampled LR for a "prediction and cure" methodology in the presence of MUSIC performance breakdown. And in section IV, we provide simulation results for a particular MUSIC performance breakdown example.

II. LIKELIHOOD RATIO FOR UNDER-SAMPLED GAUSSIAN SCENARIO

The covariance matrix $\hat{R}$ in (3) is rank-deficient when $T < M$ and therefore is described by the anti-Wishart distribution [15]. We wish to form a full rank-extension of $\hat{R}$ for use in the under-sampled likelihood ratio.

A. Formulation of the Under-Sampled LR

In addition to the original sample matrix, the transformed (whitened) sample matrix

$$\hat{C} = R_0^{-\frac{1}{2}} \sum_{j=1}^{T} \boldsymbol{x}_j \boldsymbol{x}_j^H R_0^{-\frac{1}{2}}; \quad \boldsymbol{x}_j \sim \mathcal{CN}(0, R_0) \qquad (11)$$

is described by the anti-Wishart p.d.f. (denoted $\mathcal{ACW}(T < M, M, I_M)$):

$$K_{T,M}\left(\det \hat{C}_{[T]}\right)^{T-M} e^{-\mathrm{Tr}\,\hat{C}} \prod_{l=T+1}^{M} \prod_{p=T+1}^{M} \delta\left(\frac{\det \hat{C}_{[T]lp}}{\det \hat{C}_{[T]}}\right) \qquad (12)$$

Here $K_{T,M}$ is a normalization constant and $\hat{C}_{[T]}$ is the upper left hand $T \times T$ sub-matrix of the original matrix $\hat{C}$:

$$\hat{C} = \left[\begin{array}{c|c} \hat{C}_{[T]} & * \\ \hline * & * \end{array}\right]. \qquad (13)$$

Furthermore, for each $l, p > T$ the $(T+1) \times (T+1)$ matrix $\hat{C}_{[T]lp}$ in (12) is generated by adjoining the l-th row and the p-th column of $\hat{C}$ to $\hat{C}_{[T]}$:

$$\hat{C}_{[T]lm} = \left[\begin{array}{ccc|c} & & & \hat{C}_{1p} \\ & \hat{C}_{[T]} & & \vdots \\ & & & \hat{C}_{Tp} \\ \hline \hat{C}_{l1} & \cdots & \hat{C}_{lT} & \hat{C}_{lp} \end{array}\right]. \tag{14}$$

The number of independent delta-functions in (12) is $(M-T)^2$ and therefore, for $T < M$, there are only $(2MT - T^2)$ real-valued independent entries within matrix $\hat{C}$, namely the first T rows or columns of this matrix (or any set of entries with the same number of real-valued degrees of freedom) which *uniquely specifies* the entire matrix $\hat{C}$ with rank T.

Strictly speaking, any under-sampled likelihood ratio should involve all independent entries within $\hat{C}$ that uniquely specify this matrix, and therefore any test that considers a subset $\Omega_{\hat{C}}$ of the covariance matrix $\hat{C}$ entries with a smaller number of (real-valued) degrees of freedom (DOF):

$$DOF(\Omega_{\hat{C}}) < DOF(\hat{C}) = 2MT - T^2, \tag{15}$$

should be treated as an "information-missing" one.

On the other hand, the "low rank" covariance matrix R_0 in (1) which defines our admissible set of covariance matrices is also described by the limited number of degrees of freedom

$$DOF(\hat{R}_0) = 1 + 2Mm - m^2 \tag{16}$$

where $(2Mm - m^2)$ is the number of DOF that uniquely describe the rank m signal counterpart R_S of the matrix R_0. Therefore, if the number of independent elements in the subset $\Omega_{\hat{R}}$ of $\hat{R}$ considered for hypothesis testing regarding R_S in (1) exceeds $DOF(R_S)$, then one can expect that consistent (with SNR $\rightarrow \infty$) testing is possible, even if some degrees of freedom available in $\hat{C}$ are not utilized. In fact, this statement is just another version of the well-known requirement on a sample support ($T \geqslant m$) for "low-rank" covariance matrix R_0.

Therefore, for $m < T < M$, rather than the first T rows or columns, let we consider a $(2T-1)$ wide band of the matrix $\hat{R}$:

$$\Omega^{\hat{R}} : [\hat{r}_{ij}] \quad |i-j| \leqslant T-1; \quad \hat{R} = [\hat{r}_{ij}] \;\; i,j = 1, \ldots, M \tag{17}$$

Note that the number of real-valued degrees of freedom for this band is equal to

$$DOF(\hat{R}_{B(T)}) = 2MT - T^2 - (M-T) \tag{18}$$

and is only $(M-T)$ degrees short from DOF($\hat{R}$) in (15). Since $\Omega^{\hat{R}}$ does not uniquely specify the rank T matrix $\hat{R}$, the band matrix $[r_{ij}]$ $|i-j| \leqslant T-1$ may be completed in different ways. Thus by giving up a small number of degrees of freedom and no longer fully specifying $\hat{R}$, we open up a series of possible extensions to the band matrix, including the original degenerate matrix $\hat{R}$, but also a number of non-degenerate completions. The band extension we wish to consider is the one with the maximal determinant, which is specified by the Dym-Gohberg band-extension method [16], [17].

Theorem 1: Given an M-variate Hermitian matrix $\hat{R} \equiv \{\hat{r}_{ij}\}$ $i,j = 1, \ldots, M$, suppose that

$$\begin{bmatrix} \hat{r}_{i,i} & \ldots & \hat{r}_{i,i+p} \\ \vdots & & \vdots \\ \hat{r}_{i+p,i} & \ldots & \hat{r}_{i+p,i+p} \end{bmatrix} > 0, \textit{ for } i = 1, \ldots, M-p \tag{19}$$

for $q = 1, \ldots, M$ let

$$\begin{bmatrix} \hat{\mathcal{Y}}_{q,q} \\ \vdots \\ \hat{\mathcal{Y}}_{L(q),q} \end{bmatrix} = \begin{bmatrix} \hat{r}_{q,q} & \ldots & \hat{r}_{q,L(q)} \\ \vdots & & \vdots \\ \hat{r}_{L(q),q} & \ldots & \hat{r}_{L(q),L(q)} \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{20}$$

$$\begin{bmatrix} \hat{\mathcal{Z}}_{\Gamma(q),q} \\ \vdots \\ \hat{\mathcal{Z}}_{qq} \end{bmatrix} = \begin{bmatrix} \hat{r}_{\Gamma(q),\Gamma(q)} & \ldots & \hat{r}_{\Gamma(q),q} \\ \vdots & & \vdots \\ \hat{r}_{q,\Gamma(q)} & \ldots & \hat{r}_{q,q} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{21}$$

where $L(q) = \min\{M, q+p\}$ and $\Gamma(q) = \max\{1, q-p\}$. Furthermore, let the M-variate triangular matrices U and V be defined as

$$\hat{V}_{ij} \equiv \begin{cases} \hat{\mathcal{Y}}_{ij} \hat{\mathcal{Y}}_{jj}^{-\frac{1}{2}} & \textit{for } j \leqslant i \leqslant L(j) \\ 0 & \textit{otherwise} \end{cases} \tag{22}$$

$$\hat{U}_{ij} \equiv \begin{cases} \hat{\mathcal{Z}}_{ij} \hat{\mathcal{Z}}_{jj}^{-\frac{1}{2}} & \textit{for } \Gamma(j) \leqslant i \leqslant j \\ 0 & \textit{otherwise} \end{cases} \tag{23}$$

then the M-variate matrix given by

$$\hat{R}^{(p)} = (\hat{U}^H)^{-1} \hat{U}^{-1} = (\hat{V}^H)^{-1} \hat{V}^{-1} \tag{24}$$

is the unique p.d. Hermitian matrix extension that satisfies the following condition:

$$\begin{cases} \{\hat{R}^{(p)}\}_{ij} = \hat{r}_{ij} & \textit{for } |i-j| \leqslant p, \\ \{(\hat{R}^{(p)})^{-1}\}_{ij} = 0 & \textit{for } |i-j| > p. \end{cases} \tag{25}$$

A MATLAB code snippet is provided in the Appendix which executes this band extension.

In [17], [18] it was proven that of all band extensions, extension (25) has the maximal determinant, and therefore represents a maximum entropy extension. This extension also uniquely has the property

$$\det[\hat{R}^{(p)}]^{-1} = \prod_{q=1}^{M} \hat{\mathcal{Y}}_{qq} = \prod_{q=1}^{M} e_q^T \hat{R}_q^{-1} e_q \tag{26}$$

where $\hat{R}_q$ is the $(L(q)-q+1) \times (L(q)-q+1)$ Hermitian central block matrix in $\hat{R}$, specified in (20). One can see that the Dym-Gohberg band extension method, applied to rank-deficient under-sampled versions of the sample matrix $\hat{R}$ (3), transforms this matrix into a positive definite Hermitian matrix $\hat{R}^{(p)}$ which within the $(2p+1)$-wide band has exactly the same elements as the sample matrix $\hat{R}$. Moreover, this p.d. matrix $\hat{R}^{(p)}$ *is uniquely specified* by all different $(p+1)$-variate central block matrices $\hat{R}_q$, and the only necessary and sufficient condition for such transformations to exist, is the positive definiteness of all $(p+1)$-variate submatrices $\hat{R}_q$ in (19).

Let $p \leqslant T-1$. Then for all m in (1) such that $m < p \leqslant T-1$, the number of degrees of freedom within the signal subspace R_S is less than the degrees of freedom within the transformed sample matrix. In addition, the minimal eigenvalue in all $(p+1)$-variate matrices R_q is equal to the white noise power σ_0^2 in (1), ensuring positive definiteness. For this reason, we can introduce the following likelihood ratio $\Lambda_0^{(p)}(R)$ for our under-sampled scenario:

$$\Lambda_0^{(p)}(R) = \left[\frac{\det(\hat{R}^{(p)}[R^{(p)}]^{-1}) \exp M}{\exp\{\text{Tr } \hat{R}_r R^{-1}\}}\right]^{\frac{1}{M}} \tag{27}$$

where the LR is raised to the power $/frac1M$ rather than M merely to place the LR for typical circumstances into a reasonable range. Here $R^{(p)}$ is the order p Dym-Gohberg band transformation of the tested positive definite covariance matrix model R, which has the properties

$$R^{(p)} = DG_p(R); \quad (R_{ij}^{(p)}) = r_{ij} \quad \text{for } |i-j| \leqslant p$$
$$[(R^{(p)})^{-1}]_{ij} = 0 \quad \text{for } |i-j| > p \tag{28}$$

$$\hat{R}_r = \lim_{\alpha \to 0} \frac{1}{T}(\alpha I + X_T X_T^H); \quad X_T = \{x_1, \ldots, x_T\}. \tag{29}$$

The loading factor α is sufficiently small, such that

$$DG(\hat{R}) = DG(\hat{R}_r) \tag{30}$$

which means that α should be negligible:

$$\alpha \ll \min_q \lambda_{min}(\hat{R}_q) \tag{31}$$

An alternative approach to the infinitesimal loading (not explored here further) would be to use the trace of the whitened matrix $\hat{C}$ as a normalization in the denominator instead. Also note that $\Lambda_0^{(p)}(R)$ is dependent on the determinant of $\hat{R}^{(p)}$ which in (26) is given as a function of $\hat{R}_q$ block submatrices. Therefore, we do not need to explicitly fully construct the Dym-Gohberg extensions for $\Lambda_0^{(p)}(R)$ calculation.

B. Properties of the Under-Sampled LR

Let us now demonstrate that the LR given in (27) meets the requirement $(a)-(c)$ in (8)-(10).

Proper LR Normalisation (requirement a)).

$$\max \Lambda_0^{(p)} < \exp 1; \quad \Lambda_0^{(p)} = \Lambda_0^{(p)}(\hat{R}_r) = 1 \tag{32}$$

Indeed, for $\hat{R}_r$ that satisfies (29)-(31), we have

$$\lim_{\alpha \to 0} \text{Tr } \hat{R}_r[\hat{R}_r + \alpha I]^{-1} = T\left[1 - \alpha \text{Tr } \{(X_T^H X_T)^{-1}\}\right] > 0 \tag{33}$$

$$\lim_{\alpha \to 0} \det[\hat{R}_r^{(p)} DG_p[\hat{R}_r + \alpha I]^{-1}] = 1. \tag{34}$$

Transition to the Conventional LR (requirement b)).

Obviously, for $p = M-1, \quad T \geqslant M, \; DG\{\hat{R}\} = \hat{R}$, while Tr $\hat{R}_r R =$ Tr $\hat{R}R$ for α that satisfies (31).

Scenario Independence (requirement c)).

We have to demonstrate that for the actual covariance matrix $R = R_0$, the p.d.f. for

$$\Lambda_0^{(p)}(R_0) = \left[\frac{\det(\hat{R}^{(p)}(R_0^{(p)})^{-1}) \exp M}{\exp\{\text{Tr } \hat{R}_r (R_0)^{-1}\}}\right]^{\frac{1}{M}} \tag{35}$$

does not depend on R_0, and is fully specified by parameters M, T, and p.

Theorem 2: (see Theorem 2 in [19])

Let R_0 be the true covariance matrix of the training data $X_T \sim \mathcal{CW}_T(0, R_0)$. Then the p.d.f. of $\Lambda_0^{(p)}(R_0)$ does not depend on the scenario, and can be expressed as the p.d.f. of a product of $2M$ independent random numbers α_q and Ω_q:

$$\Lambda_0^{(p)}(R_0) = \exp 1. \left[\prod_{q=1}^{M} \Omega_q \alpha_q\right]^{\frac{1}{M}} \tag{36}$$

where

$$\alpha_q \sim \frac{\alpha_q^{(T-\nu-1)}(1-\alpha_q)^{(\nu-1)}}{B[\nu, T-\nu]} \quad 1 \leqslant \nu \equiv L(q) - q \leqslant p \tag{37}$$

$$\Omega_q = \frac{C_{qq}}{T} \exp\left[-\frac{C_{qq}}{T}\right], \quad C_{qq} \sim \frac{C_{qq}^{T-1}}{\Gamma(T)} \exp(-C_{qq}) \tag{38}$$

The l-th moment of $\Lambda_0^{(p)}(R_0)$ is

$$\varepsilon\left\{\left[\Lambda_0^{(p)}(R_0)\right]^l\right\} = $$
$$= \frac{T^{TM} \exp(l)}{\left[T + \frac{l}{M}\right]^{(TM+l)}} \prod_{q=1}^{M} \frac{\Gamma\left(T + \frac{l}{M} - \nu(q)\right)}{\Gamma(T - \nu(q))} \tag{39}$$

Note that loading factor $\alpha \to 0$ is introduced in $\hat{R}_r$ to secure proper transition to conventional $LR(R)$, so that

$$\lim \text{Tr } \hat{R}_r \hat{R}_r^{-1} = M \tag{40}$$

but it needs to remain small enough for

$$\det \hat{R}^{(p)} \overset{-1}{\det} DG_p(\alpha I_0 + \hat{R}) \to 1. \tag{41}$$

At $\alpha = 0$

$$\text{Tr } \hat{R}[I - X_T(X_T^H X_T)^{-1} X_T^H] = 0, \tag{42}$$

so the term $\exp(M)$ in (35) is not required.

The above LR extends LR-based hypothesis testing techniques, including the GLRT-based detection-estimation framework outlined in Section I, into the important under-sampled domain. It should be noted that this under-sampled likelihood ratio is not the only possible formulation. We have introduced a different LR based on a projection technique in [20], and a number of other test statistics are available in the under-sampled domain (see for example [21]). In the latter example, these non-LR ad-hoc tests usually do not satisfy all three requirements given in (8)-(10) without perturbation, and their justification is asymptotic in nature (either in the classic sense as $T \to \infty$ or as both T and M proceed to ∞ at the same rate). The contribution of this under-sampled likelihood ratio is most significant in applications which leverage off the properties defined in (8)-(10), particular the invariance property which permits the evaluation of the "quality" of any estimate relative to the underlying solution *without any a priori knowledge of that solution*. To illustrate this, in what follows we explore its efficiency in an important under-sampled detection-estimation example, namely prediction of MUSIC "performance breakdown" in the threshold region.

III. GLRT-BASED DETECTION-ESTIMATION APPLICATION TO MUSIC BREAKDOWN

All subspace-based parameter estimation techniques are known to suffer a rapid degradation in performance as the SNR and/or the number of snapshots T drop below certain threshold values [22]–[25]. Because order estimation (using, for example, information theoretic criteria) is still robust in these threshold conditions, the performance breakdown manifests as highly erroneous DOA estimates for one or more of the sources, resulting in "outliers" [26]. The introduction of these outliers (typical of all subspace-based methods) is induced by the interchange of eigenvectors between the estimated signal and noise subspaces ("subspace swap") [25]. Attempts to predict from the data whether or not a subspace swap has actually occurred is given in [25], which also proposes a method of "curing" performance breakdown by comparing the deterministic ("concentrated") likelihood over various partitioning of the signal and noise-subspace eigenvectors.

We adopt a similar "prediction and cure" approach using $\Lambda_0^{(p)}(R)$ in the under-sampled ($T < M$) regime, based on a GLRT-based technique suggested in [27], [28] for conventional ($T > M$) training conditions.

According to this methodology, prediction of the presence of a subspace swap is provided by comparison of $\Lambda_0^{(p)}(R)$ for a covariance matrix model constructed from the estimated source parameters with the scenario-free $\Lambda_0^{(p)}(R_0)$ p.d.f. (39). Scenarios which do not result in a likelihood ratio consistent with the LR distribution predicted by scenario-free parameters T, M, and p are considered to have outliers. This comparison is usually implemented via a pre-computed threshold based on some statistical bound (e.g. the p.d.f median or the 10^{-3} lower extreme).

Specifically, for $\mu = 0, 1, \ldots, m_{max}$ we have to generate an under-sampled maximum likelihood model $\hat{R}_\mu$:

$$\hat{R}_\mu = \hat{\sigma}_0^2 I + \mathcal{S}_\mu(\hat{\theta}_\mu) B_\mu \mathcal{S}_\mu(\hat{\theta}_\mu) \tag{43}$$

based on a separately estimated source order $\hat{m}$ (formed, for example, via information theoretic criteria). Here $\mathcal{S}_\mu(\boldsymbol{\theta}_\mu)$ is the $[M \times \mu]$-variate antenna "manifold" matrix, uniquely specified by a set of μ parameters (DOA's) $\theta_\mu = [\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_\mu]$, B_μ is the $(\mu \times \mu)$-variate Hermitian non-negative definite (n.n.d.) inter-source covariance matrix, and σ_0^2 is the additive white noise power.

Since

$$\max_{\mu \geqslant m} \Lambda_0^{(p)}(R_\mu) \geqslant \Lambda_0^{(p)}(R_0) \tag{44}$$

(i.e. the maximum of the LR exceeds the LR of the (unknown) true solution), the scenario-free p.d.f. for $\Lambda_0^{(p)}(R_0)$ can be used to calculate a threshold ϑ_{FA} for the lower bound of the given probability of false alarm P_{FA}.

$$\int_{\vartheta_{FA}}^{1} f\left[\Lambda_0^{(p)}(R_0)\right] d\Lambda_0^{(p)} = P_{FA} \tag{45}$$

An analytic expression for this p.d.f can be given, but it is cumbersome to calculate and as an alternative, direct Monte-Carlo simulations of (27) may be employed for a given M, T and p to pre-calculate the threshold ϑ_{FA}. It is then used for hypothesis testing of proposed solutions:

$$\Lambda_0^{(p)}(\hat{R}_\mu) \geqslant \vartheta_{FA}. \tag{46}$$

After outlier prediction via the thresholding event given in (46), the outlier can be re-estimated by alternative techniques less susceptible to subspace swap (but presumably more computationally intensive). The method employed in [27], [28] utilizes sequential 1-D maximization of the likelihood ratio, but any of a wide range of MLE-based techniques can be used, as they are much less susceptible than MUSIC to generating estimates with subspace swap. Because sequential 1-D LR can be implemented efficiently and the approach leverages off the capability to evaluate $\Lambda_0^{(p)}(R)$ and truncate the search once the threshold is exceeded, we will continue to employ that technique in this example. The full procedure consists of the following five steps.

Step 1 "Breakdown Prediction".
The covariance matrix model $\hat{R}_\mu$ is tested by the inequality (46). If the threshold in (46) is exceeded, then the solution $\hat{R}_\mu$ is accepted in terms of the LR being statistically as good as the true parameters that specify the covariance matrix R_0. Otherwise, the presence of MUSIC-specific outliers is indicated.

Step 2 "Local refinement".
Local optimization by Gauss-Newton or Neder-Mead (for example) algorithms is performed to handle the case when estimates are within convex proximity to the "proper" solution.

Step 3 "Outlier identification".
When dealing with identifiable scenarios, we have to assume that the LR threshold is not achieved due to some missing DOA estimate(s). Therefore, the source in the model (43) which can be deleted from the model with the minimal degradation in LR, is treated as an outlier.

Step 4 "Outlier replacement".
Instead of the "outlier" excluded at step 3, we now search for the source with DOA estimate

that maximally contributes to the LR. In this case, an exhaustive 1-D search is used.

Step 5 "Final refinement".

Local optimization, as per step 2, is executed in the vicinity of the new set of DOA's.

If the original set includes more than a single outlier, and as a result the threshold is not exceeded, the procedure can be repeated until the threshold is exceeded or a maximum number of iterations is reached.

In [14], [27], [29] this technique was illustrated for uniform linear and circular antenna arrays under "conventional" training conditions with independent Gaussian sources. In Section IV, we provide simulation results that illustrate efficiency of this approach for under-sampled training conditions and both independent and coherent (correlated) Gaussian sources.

IV. SIMULATION RESULTS

Consider measurements from a uniform line array with data sampling at each of M=10 omnidirectional antenna elements, spaced at $\lambda/2$ to ensure independence of each spatial measurement. Fig. 1 shows the mean results of likelihood ratio formation for various levels of training data support. The three key properties of the suggested under-sampled likelihood ratio (27) can be seen in Fig. 1. The LR is normalized between 0 and 1, it transitions properly from the under-sampled likelihood ratio to a standard likelihood ratio at $T = M$, and the analytically derived LR mean (see (39)), which is by definition scenario-free, agrees with both the clairvoyant solution and averaged MUSIC-derived (non-outlier) solutions.

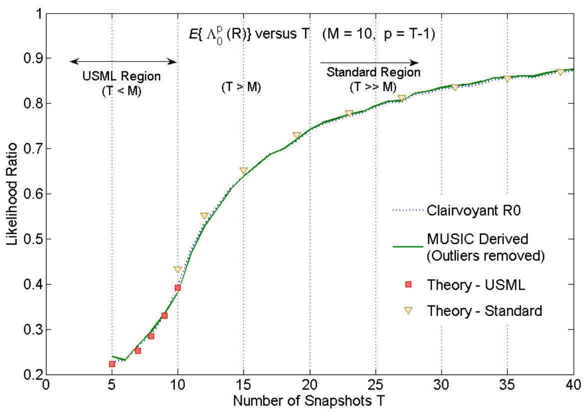

Fig. 1. Theoretical and Observed LRs in the Under and Over-Sampled Training Data Regimes

A. Independent Sources

We consider a three source scenarios with independent Gaussian sources with an input (per antenna element) SNR of 20dB per source:

$$\sin(\boldsymbol{\theta}_S) = \{-.40, 0.0, 0.06\} \tag{47}$$

The level of training support in our example is set to $T = 6$, which is clearly undersampled, but still provides

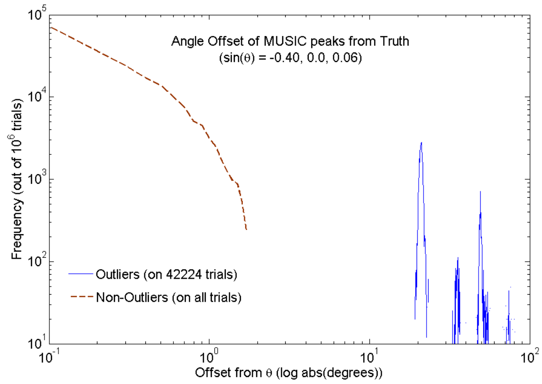

Fig. 2. Observed Angle Errors of MUSIC DOAs for the 3 Independent Source Scenario

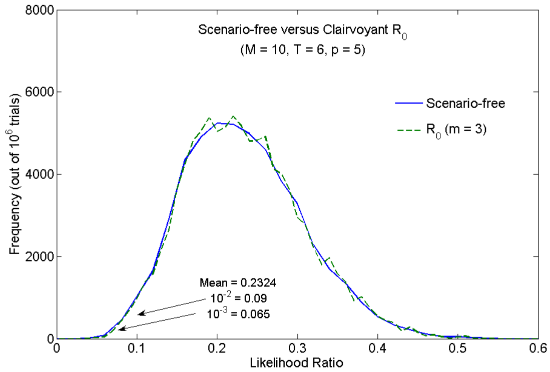

Fig. 3. Pre-Calculable LR PDF Matches Clairvoyant R_0, demonstrating the scenario invariant property of the Under-Sampled LR

a distinct signal and noise subspace in the sample covariance matrix $\hat{R}$.

To generate a "difficult" circumstance, we have selected the third source separation to reside within the MUSIC performance breakdown region mentioned in Section III. Specifically, for the selected scenario, over 40% of MUSIC derived DOA estimates from random draws of this scenario contain severely erroneous estimates ("outliers"). Clairvoyant knowledge of the underlying scenario can be used to show the distribution of the MUSIC generated outliers (Fig. 2). Based on this angular distribution, a value of $\pm 2.0^o$ was used as an association window size with the true signal DOAs while determining whether each trial containing an outlier. Note that this clairvoyant knowledge was used only in evaluating the performance of the LR in outlier detection, not in the outlier detection itself.

The detection step to determine the number of sources is based on information theoretic criteria. This scenario, while problematic for MUSIC, is not pathological, as demonstrated by the fact that the number of sources estimated in each trial using the maximum *a posteriori* probability (MAP) information theoretic criteria agreed with the actual number of independent sources (3 in this case) for all trials.

Results of our GLRT-based scheme that adopts the

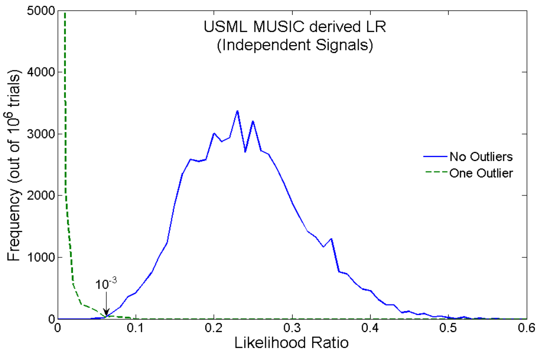

Fig. 4. LR PDFs of Uncorrelated Signals Scenario with and without MUSIC Generated "Outliers"

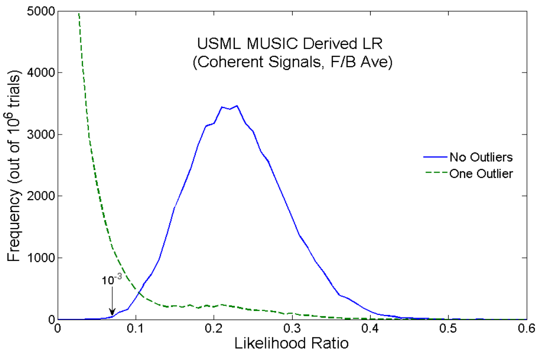

Fig. 5. LR PDFs of Correlated Signals Scenario with and without MUSIC Generated "Outliers"

under-sampled LR (27), are shown in Fig. 3 and 4 and are summarized by Tables I-III. Fig. 3 show that the pre-calculated p.d.f for $\Lambda_0^{(p)}(R_0)$ (which is scenario-free) agrees well with the clairvoyant R_0 LR results seen during the Monte-Carlo trials. Fig. 4 show that the p.d.f's of the "outlier" and "non-outlier" p.d.f.s are well separated and can be properly classified with a thresholding step.

TABLE I
"PRACTICAL THRESHOLD" - INDEPENDENT SIGNALS

GLRT Step	Outlier Detected	"Truth"	Mean LR	P_{FA} 10^{-3} $\alpha = 0.065$
1. Breakdown Prediction	No Yes	56.3% 43.7%	0.2451 0.0015	56.3% 43.7%
2. Local Refinement	No Yes	68.2% 31.8%	0.2382 0.0218	72.4% 27.6%
3/4. Outlier Predict/Cure	No Yes	95.5% 4.5%	0.2295 0.1175	99.0% 1.0%
4. Final Refinement	No Yes	95.5% 4.5%	0.2297 0.1175	99.2% 0.8%

In Table I, we adopted a threshold calculated for a $P_{FA} = 10^{-3}$, to assess "practical" non-clairvoyant performance of our routine. Let us emphasize that $p = T-1$ means that only 5-element antenna covariance array subsets are involved in model $\hat{R}^{(p)}$ reconstruction, yet quite efficient performance is demonstrated here without any diagonal loading or use of other *a-priori* information.

As previously suggested by the well separated p.d.f.s in Fig. 4, Step 1 of Table I shows that around very few non-outlier trials were misclassified (as expected based on the use of a $P_{FA} = 10^{-3}$ threshold). Subsequent steps in the GLRT "prediction and cure" methodology show that virtually all MUSIC-specific "outliers" can be rectified.

B. Correlated Sources

The introduced outlier rectification scheme may also be applied for scenarios with fully correlated sources. There are no modifications to the p.d.f pre-calculation, since it is scenario-free. For uniform linear antenna arrays, "forward-backward" spatial smoothing for each training sample is typically used to provide an M_α-variate sample covariance matrix ($M_\alpha < (M - m_{max}/2)$) that is used for conventional detection-estimation [30]. Here, dependence on T is less critical and in many cases, under-sampled training conditions ($T < M$ or even $T < m_{max}$) are all that is available.

For "forward-backward" averaging, the maximum number of resolvable sources m_{max} is specified as [30]

$$m_{max} < \frac{2}{3}(M+1) \tag{48}$$

with the "sliding window" subarray dimension $M_1 = m_{max} + 1$ that allows for MUSIC application to the M_1-variate sample matrix $\hat{R}_{M_1}$:

$$\hat{R}_{M_1} = \sum_{t=1}^{T} \sum_{j=1}^{M-M_1} \boldsymbol{x}_t^{(j)} \boldsymbol{x}_t^{(j)\mathrm{H}} + \mathcal{J}\, \overline{\boldsymbol{x}}_t^{(j)} \boldsymbol{x}_t^{(j)\mathrm{T}} \mathcal{J} \tag{49}$$

where

$$\boldsymbol{x}_t^{(j)} = [x_{jt}, x_{(j+1)t}, \ldots, x_{(M+j)t}] \tag{50}$$

is the M_1-variate sub-vector of the snapshot x_t, $\mathcal{J}$ is the permutation matrix $\begin{bmatrix} 0 & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & 0 \end{bmatrix}$, and $\overline{x}$ indicates conjugation.

In the case we consider here, where the signal-subspace dimension $\hat{m}$ and DOA estimation are performed by ITC and MUSIC techniques correspondingly, matrix $\hat{R}_{M_1}$ is no longer described by a complex Wishart distribution. Since "Wishart" training conditions $T > M$ or even the less stringent condition $m < T < M$ are no longer realized, in most practical cases that involve large antenna arrays, we have to consider dramatically under-sampled training conditions. Spatial smoothing is clearly a non-asymptotic technique in terms of the training data provided, and therefore the "gap" between performance breakdown threshold conditions of spatially averaged MUSIC and ML estimation are even more profound than in the above analyzed example with independent sources.

Let $T > 1$ and $\hat{\theta}_{\hat{m}}$ be the set of $\hat{m}$ DOA estimates provided by the traditional spatial smoothing technique. Then the "spatial smoothing generated" model of the covariance matrix $\hat{R}_{sp}$ is

$$\hat{R}_{sp} = \hat{\sigma}_0^2 I_M + S_{\hat{m}}(\hat{\boldsymbol{\theta}})_{\hat{m}} \hat{a}_1 \hat{a}_1^{\mathrm{H}} S_{\hat{m}}^{\mathrm{H}}(\hat{\boldsymbol{\theta}})_{\hat{m}} \tag{51}$$

where

$$\hat{a}_1 = [S_{\hat{m}}^{\mathrm{H}}(\hat{\boldsymbol{\theta}})_{\hat{m}} S_{\hat{m}}(\hat{\boldsymbol{\theta}})_{\hat{m}}]^{-1} S_{\hat{m}}^{\mathrm{H}}(\hat{\boldsymbol{\theta}})_{\hat{m}} \hat{\mathcal{U}}_1 \tag{52}$$

and $\hat{\mathcal{U}}_1$ is the first eigenvector of the traditional under-sampled covariance matrix $\hat{R}$.

We now repeat our simulation scenario with a 10 element antenna ($d/\lambda = 0.5$) and three Gaussian sources with an input (per-element) SNR of 20 dB per source and the same DOA set, but with all three sources fully correlated with a randomly selected scalar for the scaling of each source. Furthermore, we limit the training data set to $T = 2$ array snapshots, to demonstrate the performance of the methodology in this highly under-sampled case. Computationally, the GLRT routines were modified slightly to provide optimization across a complex-valued (rank 1) inter-source correlation matrix rather than a real, positive valued diagonal inter-source covariance matrix. Otherwise, the processing remained as before in the uncorrelated signal scenario. Results for a fully coherent 3 source scenario with the same locations given in (47) are summarized by Fig. 5 and Tables II-III.

Fig. 5 show that the p.d.f's of the "outlier" and "non-outlier" p.d.f.s overlap more than in the uncorrelated signal case and therefore are not as well classified with a thresholding step. The results for the fully coherent signal scenario show that improvement can be ultimately achieved via the GLRT-based outlier rectification scheme, but some trials with outliers result in a model LR which exceeds the threshold significantly and becoming indistinguishable in an LR sense from trials without outliers. This is an example of the so-called "maximum-likelihood performance breakdown phenomenon" [14]. Obviously, if a particular model R_μ is close enough to such a ML breakdown condition, local refinement at Step 2 can drive it above the threshold, despite an "outlier" being present in R_μ. It is then excluded from further rectification since it is classified (incorrectly) as outlier-free. Therefore, only the "gap" between MUSIC-specific and maximum likelihood performance breakdown conditions may be rectified by the suggested GLRT-based technique. While in this particular scenario, significantly better performance in this case can be achieved by avoiding the local LR optimization step (Step 2) prior to "outlier prediction and cure" (see Table III), the ML breakdown condition still prevents complete rectification.

TABLE II
"PRACTICAL THRESHOLD" - COHERENT SIGNALS SCENARIO

GLRT Step	Outlier Detected	"Truth"	Mean LR	P_{FA} 10^{-3} $\alpha = 0.065$
1. Breakdown	No	55.5%	0.2294	61.8%
Prediction	Yes	44.5%	0.0346	38.2%
2. Local	No	60.9%	0.2302	86.6%
Refinement	Yes	39.1%	0.1132	13.4%
3/4. Outlier	No	62.7%	0.2280	98.6%
Predict/Cure	Yes	37.3%	0.1321	1.4%
4. Final	No	63.7%	0.2265	99.7%
Refinement	Yes	36.3%	0.1347	0.3%

V. SUMMARY AND CONCLUSION

In this paper we proposed the likelihood ratio test to be used within the GLRT-based adaptive detection-estimation framework for under-sampled ($T < M$) training conditions. This LR involves sample covariance lags within the $(2T - 1)$-wide band of the rank T sample covariance matrix $\hat{R}$, and the maximum entropy (determinant) Dym-Gohberg extension of this band matrix. The introduced LR is normalized, coincides with the conventional LR test on covariance matrices for conventional (Wishart) training conditions ($T \geqslant M$), and most importantly, is described by a scenario-free p.d.f. for the actual covariance matrix. This invariance property, together with the observation that the properly maximized LR value should always exceed the LR value produced by the true covariance matrix, is essential for efficient implementation of GLRT-based adaptive detection-estimation.

TABLE III
NO LOCAL REFINEMENT - COHERENT SIGNAL SCENARIO

GLRT Step	Outlier Detected	"Truth"	Mean LR	P_{FA} 10^{-3} $\alpha = 0.065$
1. Breakdown	No	55.5%	0.2294	61.8%
Prediction	Yes	45.5%	0.0346	38.2%
3/4. Outlier	No	93.6%	0.2268	100%
Predict/Cure	Yes	6.4%	0.1464	0%
4. Final	No	93.6%	0.2268	100%
Refinement	Yes	6.4%	0.1464	0%

We have shown that this LR test for under-sampled conditions can be used to demonstrate significant improvement in detection-estimation performance within a MUSIC-specific breakdown threshold area. Specifically, for scenarios with either independent or fully coherent Gaussian sources, we demonstrated capabilities of our GLRT-based detection-estimation rectification scheme to recover the majority of severely erroneous solutions (outliers) produced by conventional MUSIC (at a level of over 40% of trials in particular scenarios, both correlated and uncorrelated). The previously introduced GLRT-based detection-estimation methodology is now extended to embrace the practically important class of under-sampled training conditions.

REFERENCES

[1] Y. I. Abramovich and B. A. Johnson, "A modified likelihood ratio test for detection-estimation in under-sampled training conditions," in *Proc. ICASSP*. Toulouse, France: IEEE, 2006, pp. 1105–1109.

[2] B. A. Johnson and Y. I. Abramovich, "GLRT-based detection-estimation of Gaussian signals in under-sampled training conditions," in *Proc. European Signal Processing Conference*. Florence, Italy: EURASIP, 2006.

[3] E. Hung and R. Turner, "A fast beamforming algorithm for large arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 19 (4), pp. 589–607, 1983.

[4] C. Gierull, "Performance analysis of fast projections of the Hung-Turner type for adaptive beamforming," *Signal Processing*, vol. 50, no. 1, pp. 17–28, 1996.

[5] M. Steiner and K. Gerlach, "Fast-converging maximum-likelihood interference cancellation," in *Proc. Radar Conference*. Dallas, TX, USA: IEEE, May 11-14 1998, pp. 117–122.

[6] Y. I. Abramovich, "A controlled method for adaptive optimization of filters using the criterion of maximum signal-to-noise ratio," *Radio Eng. Elect. Phys.*, vol. 25, no. 3, pp. 87–95, 1981.

[7] W. Gabriel, "Using spectral estimation techniques in adaptive processing antenna systems," *IEEE Trans. Antennas Propagat.*, vol. 34, no. 3, pp. 291–300, 1986.

[8] O. Cheremisin, "Efficiency of adaptive algorithms with regularised sample covariance matrix," *Radio Eng. Elect. Phys.*, vol. 27, no. 10, pp. 69–77, 1982.

[9] I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 10, no. 6, pp. 853–863, 1974.

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

[11] Y. I. Abramovich, N. K. Spencer, and A. Gorokhov, "A lower bound for the maximum likelihood ratio: Sparse antenna array applications," in *Proc. ICASSP*, vol. 2. Orlando, FL, USA: IEEE, 2002, pp. 1145–1148.

[12] M. Siotani, T. Hayakawa, and Y. Fujikoto, *Modern Multivariate Statistical Analysis*. Cleveland: Amer. Sci. Press, 1985.

[13] Y. I. Abramovich, N. K. Spencer, and A. Gorokhov, "Bounds on maximum likelihood ratio — Part II: Application to antenna array detection-estimation with imperfect wavefront coherence," *IEEE Trans. Signal Processing*, vol. 53, no. 6, pp. 2046–2058, Jun 2005.

[14] ——, "Bounds on maximum likelihood ratio — Part I: Application to antenna array detection-estimation with perfect wavefront coherence," *IEEE Trans. Signal Processing*, vol. 52, no. 6, pp. 1524–1536, Jun 2004.

[15] R. A. Janik and M. A. Nowak, "Wishart and anti-Wishart random matrices," *Journal of Physics A: Mathematics and General*, vol. 36, no. 12, pp. 3629–3637, 2003.

[16] H. Dym and I. Gohberg, "Extensions of band matrices with band inverses," *Linear Algebra Applications*, vol. 36, pp. 1–24, 1981.

[17] H. Woerdeman, "Matrix and operator extensions," Ph.D. dissertation, Vrije Universiteit, Amsterdam, Netherlands, 1989.

[18] R. Grone, C. Johnson, E. de Sa, and H. Wolkowicz, "Positive definite completions of partial Hermitian matrices," in *Proc. Linear Algebra and Applications*, vol. 58, Salt Lake City, Utah, 1984, pp. 109–124.

[19] Y. I. Abramovich, N. K. Spencer, and M. Turley, "Order estimation and discimination between stationary and time-varying autoregressive (TVAR) models," *submitted to IEEE Transactions on Signal Processing*, 2005.

[20] B. A. Johnson and Y. I. Abramovich, "GLRT-based outlier prediction and cure in under-sampled training conditions using a singular likelihood ratio," in *Proc. ICASSP*. Honolulu, HI, USA: IEEE, Apr 15-20 2007, p. 4.

[21] O. LeDoit and M. Wolf, "Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size," *The Annals of Statistics*, vol. 30, no. 4, pp. 1081–1102, 2002.

[22] D. Tufts, A. Kot, and R. Vaccaro, "The threshold effect in signal processing algorithms which use an estimated subspace," in *SVD and Signal Processing II: Algorithms, Analysis and Applications*, R. Vaccaro, Ed. New York: Elsevier, 1991, pp. 301–320.

[23] J. Thomas, L. Scharf, and D. Tufts, "The probability of a subspace swap in the SVD," *IEEE Trans. Signal Processing*, vol. 43, no. 3, pp. 730–736, Mar 1995.

[24] P. Stoica, V. Šimonytė, and T. Söderström, "On the resolution performance of spectral analysis," *Signal Processing*, vol. 44 (6), pp. 153–161, 1995.

[25] M. Hawkes, A. Nehorai, and P. Stoica, "Performance breakdown of subspace-based methods: prediction and cure," in *Proc. ICASSP*, vol. 6. Salt Lake, UT, USA: IEEE, 2001, pp. 4005–4008.

[26] M. Kaveh and A. Barabell, "The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise," *IEEE Trans. Signal Processing*, vol. 34, no. 2, pp. 331–341, 1986.

[27] Y. I. Abramovich, N. K. Spencer, and P. Turcaj, "GLRT-based adaptive detection-estimation of Gaussian sources in coloured noise fields," in *Proc. ISSPIT-2002*, Darmstadt, Germany, 2003.

[28] Y. I. Abramovich, N. K. Spencer, J. Schroeder, and A. Gorokhov, "GLRT-based detection-estimation of uncorrelated Gaussian sources in circular antenna arrays," in *Proc. 37th Asilomar Conference on Signals, Systems and Computers*, vol. 1. Pacific Grove, CA, USA: IEEE, Nov 9-12 2003, pp. 1198–1202.

[29] Y. I. Abramovich and N. K. Spencer, "Performance breakdown of subspace-based methods in arbitrary antenna arrays: GLRT-based prediction and cure," in *Proc. ICASSP*, vol. 2. Montreal, Canada: IEEE, 2004, pp. 117–120.

[30] T. Shan, M. Wax, and T. Kailath, "On spatial smoothing for direction-of-arrival estimation of coherent signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, no. 4, pp. 806–11, 1985.

Ben A. Johnson (SM'04) received a Bachelor of Science (*cum laude*) degree in physics at Washington State University in 1984 and a Master of Science degree in digital signal processing at University of Southern California in 1988.

From 1984 to 1989, he was at Hughes Aircraft Company (now Raytheon) as a systems engineer in airborne radar. From 1989 to 1998, he was with Sensis Corporation as a senior radar engineer in ground-based surveillance systems. Since 1998, he has been with Lockheed Martin (seconded to RLM Management) on the Jindalee over-the-horizon Operational Radar Network (JORN), first as a senior test engineer and then as technical director. He is also currently pursuing a Ph.D. at the Institute of Telecommunications Research, University of South Australia, focusing on application of spatio-temporal adaptive processing in HF radar.

Yuri I. Abramovich 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), U.S.S.R., and in 1981, he received the D.Sc. degree in radar and navigation from the Leningrad Institute for Avionics, Leningrad (Russia), U.S.S.R.

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. Since 1994, he has been 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) as Principal Research Scientist, seconded to CSSIP. 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.

APPENDIX
MATLAB Listing for Dym-Gohberg Band Extension

```
function [DGRhat V] = DGtrans(Rhat,p)
%
% provide Dym-Gohberg extension - note that
% computation of either V or U is necessary,
% but not both.  Only the lower triangular
% matrix V is computed here.  The determinant
% can be computed directly from V.
%
M = size(Rhat,1); V = zeros(M); Y = zeros(M);
for q = 1:M
   Lq = min(M, q+p);
   onetop = zeros(Lq-q+1,1);
   onetop(1,1) = 1;
   subY = inv(Rhat(q:Lq,q:Lq))*onetop;
   Y(q:Lq,q:q) = subY;
end %for
for ii = 1:M
   for jj = 1:M
      if (jj <= ii) && (ii <= min(M,jj+p))
         V(ii,jj) = Y(ii,jj)/sqrtm(Y(jj,jj));
      else
         V(ii,jj) = 0;
      end %if
   end % for
end %for
DGRhat = inv(V')*inv(V);
```