GLRT-BASED OUTLIER PREDICTION AND CURE IN UNDER-SAMPLED TRAINING CONDITIONS USING A SINGULAR LIKELIHOOD RATIO

Ben A. Johnson and Yuri I. Abramovich †*

* RLM Management, Pty. Ltd. and University of South Australia/ITR, e-mail: ben.a.johnson@ieee.org
†DSTO, ISR Division, Edinburgh, Australia, e-mail: yuri.abramovich@dsto.defence.gov.au

ABSTRACT

For cases where the number of training samples T does not exceed the number of antenna elements M, we consider a detection-estimation problem for Gaussian sources occupying a low-rank m-dimensioned signal subspace within the associated covariance matrix ($m < T < M$). We derive a likelihood ratio that for the null hypothesis is described by a probability function that does not depend on a scenario, and investigate a (non-trivial) correspondence between the likelihood function and the derived likelihood ratio with respect to maximization performance. Practical application of this technique is illustrated for under-sampled ($T < M$) conditions for the purpose of MUSIC performance enhancement in the "threshold" region.

Index Terms— Array signal processing, Maximum likelihood estimation, Adaptive estimation.

1. INTRODUCTION AND PROBLEM FORMULATION

When a covariance matrix of a Gaussian mixture, impinging upon an M-variate antenna array, belongs to an *a priori* restricted class, specified by a limited number of parameters, one can consider addressing the detection-estimation problem having less training samples T than the antenna dimension M (the "under-sampled" regime). One of the well-known families of this kind, considered in this paper, is the "low-rank" signal subspace one where an admissible covariance matrix could be described as:

$$R = \sigma_0^2 I_M + R_S; \;\; R_S = U_m \Lambda_0 U_m^{\rm H}; \;\; \Lambda_0 = \Lambda_m - \sigma_0^2 I_m \quad (1)$$

where $U_m \in C^{M\times m}$ 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 strong enough signal-to-noise ratio (SNR), the MUltiple SIgnal Classification (MUSIC) algorithm can provide accurate DOA estimates, based on T training snapshots, equal to the number of independent sources m. Similarly, the well-known Wax-Kailath ITC criteria may address the detection problem (i.e. estimation of the number of sources m), if $T > M$ [1]. Yet, below some "threshold" SNR, the MUSIC "breaks down", which means that the DOA estimation accuracy rapidly departs from the Cramer-Rao Lower Bound (CRLB), due to generation of erroneous DOA estimates ('"outliers").

The mechanism of this breakdown is well investigated in [2] and is proven to be MUSIC-specific, associated with the so-called "subspace swap". Maximum likelihood estimation (MLE) also breaks down at the point where solutions with completely erroneous DOA estimates (outliers) generate a likelihood function (LF) value that exceeds the LF value produced by the true covariance matrix, or even the local LF extremum in the vinicity of the true covariance matrix [3].

Investigations of ML "performance breakdown" (threshold) conditions are continuing [4, 5], but it has already been demonstrated that there typically is a large "gap" in required SNR, source separation or sample support between MUSIC-specific and ML-intrinsic threshold conditions [6, 7]. Moreover, in [6], we were able to propose a ML-based technique for detection of MUSIC-specific outliers and rectification of the "broken-down" MUSIC solution by replacing the outlier with a proper DOA estimate that increases the ML value beyond a certain expected (threshold value). Specifically, we considered a normalized version of the LF, such that for the null hypothesis (.e. when the model coincides with the actual covariance matrix), its probability function does not depend on the actual covariance matrix and is fully specified by T and M only. This invariance property allowed for precalculation of a threshold value that must be exceeded by this normalized LF (likelihood ratio, LR) of the actual covariance matrix, with a certain (high) probability. We then demonstrated the high efficiency of this practical LR-based MUSIC "performance breakdown prediction and cure" routine, referred to in this paper as GLRT-PAC.

The most important limitation of this routine is that it may operate only in the properly sampled (Wishart) training condition ($T \geqslant M$), since for this condition the LF could be *accurately* normalized. Of course, for multi-variate training data $x_t, t = 1, \ldots, T$ $x_t \sim CN(0, R_0)$, where R_0 is the underlying actual (true) covariance matrix, the likelihood function exists and is non-degenerate even under under-sampled ($T < M$) conditions:

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

where the sample covariance matrix $\hat{R} = \frac{1}{T}\sum_{j=1}^{T} x_j x_j^{\rm H}$.

For ($T \geqslant M$), the normalized version of the LF (2) is given by:

$$LR(R) = \frac{\mathcal{L}(X_T, R)}{\max_R \mathcal{L}(X_T, R)} = \left[\frac{\det \hat{R}R^{-1} \exp M}{\exp\{\text{Tr}\,\hat{R}R^{-1}\}}\right]^T < 1, \quad (3)$$

which, as a function of R, is *identical* to the LF (2), and at the same time is described by the required scenario invariant p.d.f. for $R = R_0$. However, for $T < M$, a unique and identical to the LF (2) normalized likelihood ratio $LR_u(R)$ does not exist (since $\det(\hat{R}) = 0$), preventing the direct application of the GLRT-PAC routine to the under-sampled regime.

To address this shortfall, in [8] we introduced an appropriately invariant likelihood ratio, based on considering a subset of the sample covariance matrix $\hat{R}$ entries and application of a specific Dym-Gohberg transformation to the matrix $[\hat{R}R^{-1}]$. In [9] we showed

This work was conducted under DSTO/RLM Collaboration Agreement 290905

that this approach for a practically important case allows for MUSIC performance improvement in the threshold area with under-sampled (anti-Wishart) training conditions.

Yet, this particular *ad-hoc* criterion has not been directly derived for the model (1) and most importantly, distinctions between the likelihood function (2) maximization and the under-sampled likelihood ratio LR_u has not been explored. In this paper, we derive a different likelihood ratio, and analyze distinctions between the maximum likelihood function and maximum likelihood ratio solutions for $T < M$. Finally, we investigate the gap between the theoretical ML-based "performance breakdown" and breakdown conditions for our practical GLRT-PAC routine.

2. LIKELIHOOD RATIO FOR UNDER-SAMPLED TRAINING CONDITION

Let the eigen-decomposition of the sample covariance matrix $\hat{R}$ for $T < M$ be presented as

$$\hat{R} = \hat{\mathcal{U}}_T \hat{\Lambda}_T \hat{\mathcal{U}}_T^{\mathrm{H}},\ \hat{\mathcal{U}}_T^{\mathrm{H}} \hat{\mathcal{U}}_T = I_T,\ \ \hat{\Lambda}_T = diag(\hat{\lambda}_1, \ldots, \hat{\lambda}_T)). \quad (4)$$

Then any inference about the covariance matrix model R_{mod} may be provided only regarding its "projection" unto the linear subspace spanned by the $[M \times T]$-variate matrix of eigenvectors $\hat{\mathcal{U}}_T$ associated with the T non-zero eigenvalues.

Therefore, for any given R_{mod}, we have to define such a projection, i.e. we have to find the rank T Hermitian matrix

$$R_T = \hat{\mathcal{U}}_T D_T \hat{\mathcal{U}}_T^{\mathrm{H}},\ \ 0 \leqslant D_T \in \mathcal{H}^{T \times T} \quad (5)$$

which in a certain sense is the "closest" to the model R_{mod}. The way we specify D_T, given R_{mod}, actually specifies distinctions between the original (unnormalized) likelihood function and the under-sampled likelihood ratio.

Let us consider the matrix pencil

$$(\hat{\mathcal{U}}_T D_T \hat{\mathcal{U}}_T^{\mathrm{H}}) V_j = \mu_j R_{mod} V_j,\ \ V_j \neq 0 \quad (6)$$

with $$\det[\hat{\mathcal{U}}_T D_T \hat{\mathcal{U}}_T^{\mathrm{H}} - \mu_j R_{mod}] = 0 \quad (7)$$

and $$V_j^{\mathrm{H}} (\hat{\mathcal{U}}_T D_T \hat{\mathcal{U}}_T^{\mathrm{H}}) V_j = \mu_j V_j^{\mathrm{H}} R_{mod} V_j \quad (8)$$

$$V_j^{\mathrm{H}} R_{mod} V_l = \delta_{jl},\ (\delta_{jl} = \text{ kroneker delta}). \quad (9)$$

For all $\mu_j > 0$ ($j = 1, \ldots, T$), the set of corresponding T linearly independent eigenvectors V_j span the linear subspace that is identified by the matrix $\hat{\mathcal{U}}_T$. Therefore, if we can find a solution D_T such that in (6), $\mu_1 = \mu_2 = \cdots = \mu_T = 1$, then the two matrices $\hat{\mathcal{U}}_T D_T \hat{\mathcal{U}}_T^{\mathrm{H}}$ and R_{mod} *within the subspace spanned by* $\mathcal{U}_T$, may be treated as being properly normalized.

It is straightforward to show that such a solution uniquely exists:

$$D_T = (\hat{\mathcal{U}}_T^{\mathrm{H}} R_{mod}^{-1} \hat{\mathcal{U}}_T)^{-1}, \quad (10)$$

and therefore, given an arbitrary p.d. covariance matrix model R_{mod}, we can make an inference about its "projection"

$$R_T = \hat{\mathcal{U}}_T (\hat{\mathcal{U}}_T^{\mathrm{H}} R_{mod}^{-1} \hat{\mathcal{U}}_T)^{-1} \hat{\mathcal{U}}_T^{\mathrm{H}} \quad (11)$$

based on the provided sufficient statistics $\hat{R}$.

It is important to note, that for any R_{mod} in (11), R_T is now an admissible *singular covariance matrix* for the input data $x_j, j = 1, \ldots, T$. In fact, the main distinction for the LF (2) that *always considers* a p.d. admissible covariance matrix is expanded for $T < M$ into the domain of *singular* admissible p.d.f's.

Indeed, a singular complex Gaussian p.d.f. is specified by a singular covariance matrix $R_S = \mathcal{U}_S \Lambda_S \mathcal{U}_S^{\mathrm{H}}$ when a random M-variate vector x fully resides in the subspace spanned by $\mathcal{U}_S$; i.e. when

$$[I - \mathcal{U}_S \mathcal{U}_S^{\mathrm{H}}] x = 0, \quad (12)$$

and the p.d.f. $w(x)$ is introduced as

$$w(x) = \frac{1}{\pi^M \det \Lambda_S} \exp\left[-x^{\mathrm{H}} \mathcal{U}_S \Lambda_S^{-1} \mathcal{U}_S^{\mathrm{H}} x\right] \quad (13)$$

For $X_T = [x_1, \ldots, x_T]$ such that $[I - \mathcal{U}_S \mathcal{U}_S^{\mathrm{H}}] x_j = 0$ for all j, we have

$$w(X_T) = \left[\frac{1}{\pi^M \det \Lambda_S} \exp\left[-\mathrm{Tr}\, \mathcal{U}_S \Lambda_S^{-1} \mathcal{U}_S^{\mathrm{H}} \hat{R}\right]\right]^T \quad (14)$$

and therefore, for our specific case we now may consider the *singular* likelihood function

$$\mathcal{L}_S(X_T | R_{mod}) = \left[\frac{\exp\left[-\mathrm{Tr}\, \hat{\mathcal{U}}_T (\hat{\mathcal{U}}_T^{\mathrm{H}} R_{mod}^{-1} \hat{\mathcal{U}}_T) \hat{\mathcal{U}}_T^{\mathrm{H}} \hat{R}\right]}{\pi^M \det[\hat{\mathcal{U}}_T^{\mathrm{H}} R_{mod}^{-1} \hat{\mathcal{U}}_T]^{-1}}\right]^T \quad (15)$$

One can see that the main distinction between $\mathcal{L}(X_T, R_{mod})$ in (2) and $\mathcal{L}_S$ (15) is that instead of $\det(R_{mod})$ in the denominator of (2), we now have the determinant of the T-variate matrix $[\hat{\mathcal{U}}_T^{\mathrm{H}} R_{mod}^{-1} \hat{\mathcal{U}}_T]^{-1}$.

Now we can find (see Lemma 3.2.2 from Anderson [10])

$$\max_{R_{mod}>0} \mathcal{L}_S(X_T | R_{mod}) \leqslant \left[\frac{\det[(\hat{\mathcal{U}}_T^{\mathrm{H}} \hat{R} \hat{\mathcal{U}}_T)^{-1}] \exp(-T)}{\pi^M}\right]^T \quad (16)$$

with the maximum obtained for $[\hat{\mathcal{U}}_T^{\mathrm{H}} R_{mod}^{-1} \hat{\mathcal{U}}_T] = \hat{\Lambda}_T^{-1}$ that finally leads to the properly normalized likelihood ratio

$$LR_u(X_T | R_{mod}) = \frac{\mathcal{L}_u(X_T | R_{mod})}{\max_{R_{mod}} \mathcal{L}_u(X_T | R_{mod})}$$
$$= \left[\frac{\det \hat{\Lambda}_T (\hat{\mathcal{U}}_T^{\mathrm{H}} R_{mod}^{-1} \hat{\mathcal{U}}_T) \exp T}{\exp\left[\mathrm{Tr}\ (\hat{\mathcal{U}}_T^{\mathrm{H}} R_{mod}^{-1} \hat{\mathcal{U}}_T) \hat{\Lambda}_T\right]}\right]^T$$
$$\equiv \left[\frac{\det(X_T^{\mathrm{H}} R_{mod}^{-1} X_T / T) \exp T}{\exp\left[\mathrm{Tr}\ (X_T^{\mathrm{H}} R_{mod}^{-1} X_T / T)\right]}\right]^T,\ T \leqslant M \quad (17)$$

Now the introduced likelihood ratio has a very straight-forward interpretation. Indeed, the likelihood ratio (17) now tests the hypothesis

$$H_0: \quad \mathcal{E}\{X_T^{\mathrm{H}} R_{mod}^{-1} X_T / T\} = I_T \quad \text{versus} \quad (18)$$

$$H_1: \quad \mathcal{E}\{X_T^{\mathrm{H}} R_{mod}^{-1} X_T / T\} \neq I_T \quad (19)$$

with $\mathcal{E}\{\cdot\}$ as the expectation operator. In fact, the $LR_u(X_T | R_{mod})$ tests the "quality" of the pre-whitened (degenerate) matrix $\hat{C}_{mod}$

$$\hat{C}_{mod} = R_{mod}^{-\frac{1}{2}} X_T X_T^{\mathrm{H}} R_{mod}^{-\frac{1}{2}} \quad (20)$$

More specifically that this test checks how close to unity are the *non-zero* eigenvalues of the "pre-whitened" matrix $\hat{C}$.

It is quite clear that for the null-hypothesis in (18) when $R_{mod} = R_0$, the distribution of $LR_u(X_T | R_0)$ does not depend on R_0. This p.d.f. has been previously derived in [11] via Mellin's transform of the moment function:

$$\mathcal{E}\{[LR_u(X_T | R_0)]^h\} =$$
$$\frac{M^{TM} \exp[Th]}{(M+h)^{T(M+h)}} \prod_{j=1}^{T} \frac{\Gamma(M+h+1-j)}{\Gamma(M+1-j)},\ \ T \leqslant M. \quad (21)$$

where $\Gamma(x)$ represents the Gamma function.

3. MAXIMUM LIKELIHOOD VERSUS MAXIMUM SINGULAR LIKELIHOOD RATIO COMPARISON

It seems quite obvious that the introduced singular likelihood ratio (17) and, in general, any reliable inference about a *positive definite* model R_{mod} could be successful only if a $(T < M)$-variate subspace spanned by the training data *accurately enough* spans the unknown (estimated) subspace of the covariance matrix model R_{mod}.

In our problem, we deal with the model

$$R_{mod} = \sigma_0^2 I + S_m B_m S_m^{\mathrm{H}} \qquad (22)$$

where $B_m \geqslant 0$ is the m-variate inter-source Hermitian correlation matrix ($B_m = diag\{\sigma_1^2, \ldots, \sigma_m^2\}$ for independent sources) and $S_m = [S(\theta_1), \ldots, S(\theta_m)] \in C^{M\times m}$ is the set of antenna steering vectors, specified uniquely by the set of DOA parameters $\{\theta_1, \ldots, \theta_m\}$, and $T \geqslant m$.

The condition $T \geqslant m$, while clearly necessary, may only be (practically) sufficient if the complementary $(M - T)$-variate subspace in the model (22) is fixed, which means that the white noise power σ_0^2 is known *a priori*. Otherwise, the sample support T should significantly exceed the (maximal) number of sources to allow for proper estimation of the noise subspace of the model.

For this reason, let us consider the case when the white noise power σ_0^2 is known *a priori*, in order to secure sufficiently high potential DOA estimation performance. Still, even under the known white noise power, a T-variate sample subspace does not entirely span the m-variate $(T \geqslant m)$ antenna manifold subspace S_m. Therefore, the strict equivalence between the likelihood function and the under-sampled likelihood ratio (17) does not exist, but a "sufficiently accurate" statistical equivalence may.

From a practical viewpoint, we must be interested whether the (global) maximization of the likelihood function (2) can provide significantly better DOA estimation accuracy than the (global) LR_u (17) optimization. The estimation accuracy metric, as usual, means the proximity to the CRLB, and involves both an "asymptotic" regime (no outliers) as well as a "threshold" regime where the ML criterion starts to break. Apart from theoretical interest, this analysis has quite a straight-forward practical application.

Indeed, the central idea of the GLRT-PAC technique [6, 9] is based on the invariance property of the $LR_u(X_T|R_0)$, which allows for the setup of pre-calculated thresholds, above which the likelihood ratio should be driven. Naturally, we must be sure that by driving the likelihood ratio above such values produced (statistically) by the true covariance matrix R_0, we simultaneously drive the likelihood function beyond the clairvoyant value as well. For $T \geqslant M$, the likelihood ratio (3) is precisely a scaled version of the likelihood function (2), while for $T < M$, this property needs to be verified for the introduced $LR_u(X_T|R_{mod})$ (17).

Unfortunately, since we are concerned with the "gap" between MUSIC-specific and ML-intrinsic breakdown, global LF or LR maximization cannot be easily implemented or substituted by an approximately convex problem. For this reason, in our analysis of the Monte-Carlo simulations, we have adopted the following approach. For a given scenario, specified by the (true) covariance matrix R_0 (with independent sources)

$$R_0 = \sigma_0^2 I + \sum \sigma_j^2 S(\theta_j) S^{\mathrm{H}}(\theta_j) \qquad (23)$$

for every Monte-Carlo trial that result in the sample matrix $\hat{R}$ $(T < M)$, we calculated $\mathcal{L}(X_T, R_{MUSIC})$ and $\mathcal{L}(X_T, R_L)$ where R_{MUSIC} is the covariance matrix model (22) restored using the MUSIC DOA estimates $\hat{\theta}_j, j = 1, \ldots, m$. The number of sources m in this study is assumed to be known *a priori*, while the source power estimates that match $\hat{R}$ for the known white noise power σ_0^2 and DOAs $\hat{\theta}_j$ are calculated in the traditional way.

The likelihood function $\mathcal{L}(X_T, R_L)$ represents the LF found by an optimization routine converging to a local extremum after being initiated by the actual (true) parameters θ_j, σ_j^2 in the model R_0. Note that the asymptotic theory of ML estimation considers this local extremum as the global one.

In the same way, we also calculate $LR_u(X_T, R_{MUSIC})$ and $LR_u(X_T, R_L)$. Then for every result R_{MUSIC} such that

$$\mathcal{L}(X_T|R_{MUSIC}) < \mathcal{L}(X_T, R_L), \qquad (24)$$

we apply the particular LF optimization routine derived in [6, 9]. That routine consists of outlier identification, outlier replacement by a DOA estimate that finds the LF maximum via a 1-D search, and finally local refinement of all the DOA estimates.

In what follows, we consider for further assessment only the set of successful trials, whereby

$$\mathcal{L}(X_T, R_{opt}) \geqslant \mathcal{L}(X_T, R_L). \qquad (25)$$

Clearly, if the strict inequality in (25) takes place, R_L is not a global extremum, and if R_{opt} in this case still contains an outlier, then an ML "breakdown" instance is realized. The ML DOA estimation accuracy is now averaged over the set of successful trials R_{opt} that passed the comparison (25).

Similar optimization and analysis has also been performed for the $LR_u(X_T|R_{mod})$ values, ie. for every R_{MUSIC} with

$$LR_u(X_T, R_{MUSIC}) < LR_u(X_T, R_L) \qquad (26)$$

we tried to find a "successful" solution with

$$LR_u(X_T, R_{opt}) \geqslant LR_u(X_T, R_L) \qquad (27)$$

Finally, using the invariance property of the p.d.f. for $LR_u(X_T, R_0)$ (see (21)), we find the threshold α, such that

$$prob\{LR_u(X_T, R_0) \geqslant \alpha\} = P_0, \qquad (28)$$

and for all MUSIC results where $LR_u(X_T, R_{MUSIC}) < \alpha$ we use the GLRT-PAC routine [9], to find an solution such that

$$LR_u(X_T, R_{pac}) \geqslant \alpha \qquad (29)$$

Clearly, the DOA estimation accuracy averaged over the set of R_{opt} that meet the condition (25), represent the a proxy for MLE, even if these solutions retain a number of outliers.

Estimation accuracy, averaged over the set of trials that passed the likelihood ratio comparison (27) allows us to assess the potential estimation performance losses associated with adopting the practical under-sampled $LR_u(X_T, R_{mod})$ instead of the unnormalized likelihood function (2).

Finally, the set of solutions R_{pac} that meet condition (28) illustrates the practically delivered improvement in MUSIC DOA estimation accuracy by the GLRT-PAC routine. Comparison of this practically available performance with the routines utilizing clairvoyantly determined completion thresholds demonstrates the performance gap that can be closed by more powerful (but computationally intensive) ML optimization routines.

Results of this analysis is illustrated below in Fig. 1 for the DOA set $\theta_m = \{-20^o, -15^o, 35^o, 37^o\}$, with $T = 15$ training samples and a uniform linear antenna with $M = 20$ elements, spaced at $\lambda/2$. The mean square error (based on offset of the estimated DOAs from the true DOAs) for the two closely spaced sources $(35^o, 37^o)$ and the average CRLB associated with the two sources is shown for source SNRs ranging from -15 dB to +25 dB. Results are shown for standard MUSIC, the practical GLRT-PAC algorithm, the GLRT-PAC algorithm based on the "strict" threshold (27) and finally the predict and cure search algorithm based on the likelihood function (2) and the completion threshold (25).

The need for the threshold in the latter two algorithms is to ensure that failures of the routine itself to find the global maximum are not considered. The results are therefore a proxy for global MLE using both the LR_u (17) and the LF (2). It should be noted, however, that in this case, the predict and cure routine success rate is quite high. For example, for the worst-case result at -3dB SNR, we were able to get 78% of the solutions to meet condition (25), 94% of the solutions to meet condition (27), and 100% of the solutions to meet the practical threshold comparison (28). Even better statistics are observed at other SNRs and therefore the conclusions that follow are statistically reliable.

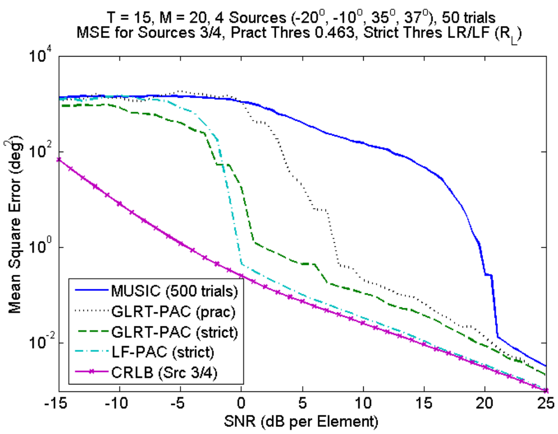

Fig. 1: Mean Square Error for Practical and Strict GLRT-PAC

At the high SNR = 25 dB, all the above mentioned techniques, including MUSIC, demonstrate high accuracy consistent with the CRLB (although, as expected, the ML optimization provides improvement relative to MUSIC, since the sample support is limited). At SNR = 21 dB, MUSIC starts to break, and at SNR = 7 dB all trials contained an outlier with an error that exceeded 2^o.

The best performance is delivered by the "predict and cure" routine based on the LF (labeled LF-PAC (strict)), which is not surprising given that it is the closest proxy to global MLE of the routines compared. Here the "ML breakdown" phenomena is not observed until SNR = 1 dB, which is 20dB beyond the MUSIC breakdown threshold in this small sample case. This demonstrates the considerable gap that exists between MUSIC-specific and ML-intrinsic breakdown when operating in this small sample regime.

It is very pleasing to confirm that there is no significant difference between the "MLE proxy" routines which employ the likelihood function (2) and LF local extremum threshold (25) versus those which employ the under-sampled likelihood ratio (17) and LR_u local extremum threshold (27). The observed breakdown threshold is practically indistinguishable, while the actual accuracy is only mildly degraded (and still improved relative to MUSIC).

Finally, we are able to demonstrate a significant performance improvement w.r.t. the conventional MUSIC, delivered by the practical (statistical) LR_u-based GLRT-PAC routine. Indeed, the SNR breakdown point has been moved from 21 dB down to around 8 dB.

At the same time, we can see that potential ML-intrinsic "threshold" performance is not fully achieved by the GLRT-PAC routine due to the statistical nature of the completion threshold in (29). Further reduction of the breakdown point from 8 dB to 1 dB could serve as a motivation for more use of computationally involved techniques.

4. SUMMARY AND CONCLUSION

In this paper we introduced the new under-sampled likelihood ratio $LR_u(X_T, R_{mod})$ that can be used in GLRT-based detection-estimation techniques such as GLRT-PAC when $T < M$. This LR_u is properly normalized and most importantly, for $R_{mod} = R_0$ is described by a p.d.f. invariant w.r.t. the scenario R_0, being fully specified by M and T. We demonstrated that optimization of this LR_u is practically equivalent to the accurate LF optimization. At the same time, the above mentioned p.d.f. invariance enables the statistical assessment of the "quality" of any solution derived by MUSIC, and furthermore, supports significant DOA estimation accuracy improvement. The potential threshold capabilities of accurate ML estimation is also demonstrated and shown to be still above practically available performance.

5. REFERENCES

[1] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, no. 2, pp. 387–92, 1985.

[2] 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.

[3] Y. Abramovich, N. Spencer, and A. Gorokhov, "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.

[4] F. Athley, "Performance analysis of DOA estimation in the threshold region," in *Proc. ICASSP*, vol. 3. Orlando, FL, USA: IEEE, 2002, pp. 3017–3020.

[5] C. D. Richmond, "On the threshold region mean-squared error performance of maximum-likelihood direction-of arrival estimation in the presence of signal model mismatch," in *Sensor Array and Multichannel Signal Processing Workshop*. Boston, MA, USA: IEEE, 2006, pp. 268–272.

[6] 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.

[7] N. Spencer and Y. Abramovich, "Performance analysis of DOA estimation using uniform circular antenna arrays in the threshold region," in *Proc. ICASSP*, vol. 2. Montreal, Canada: IEEE, 2004, pp. 233–236.

[8] 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.

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

[10] T. Anderson, *An Introduction to Multivariate Statistical Analysis*. New York: John Wiley and Sons, 1958.

[11] B. Nagarsenker and K. Pillai, "Distribution of the likelihood ratio criterion for testing a hypothesis specifying a covariance matrix," *Biometrica*, vol. 60, no. 2, pp. 359–361, 1973.