PERFORMANCE ANALYSIS OF DOA ESTIMATION USING UNIFORM CIRCULAR ANTENNA ARRAYS IN THE THRESHOLD REGION

N. K. Spencer

Cooperative Research Centre for Sensor Signal & Information Processing (CSSIP), SPRI Building, Technology Park Adelaide, Mawson Lakes SA 5095, Australia
nspencer@cssip.edu.au

Y. I. Abramovich

Intelligence, Surveillance & Reconnaissance Division, Defence Science & Technology Organisation (DSTO), PO Box 1500, Edinburgh SA 5111, Australia
yuri@cssip.edu.au

ABSTRACT

We extend our investigations into the region between "performance breakdown" (threshold) conditions for subspace-based direction-of-arrival (DOA) estimation techniques and maximum likelihood (ML) DOA estimation for uniform circular antenna arrays. Using a new outlier mitigation technique that can be applied to arbitrary geometry arrays, we demonstrate that subspace-based DOA estimation outliers could be reliably rectified up to the point where erroneous DOA estimates start to generate optimally high likelihood ratios (LRs). This is a manifestation of the onset of a discontinuity in the ML estimation, which cannot be rectified within the ML paradigm.

1. INTRODUCTION

While the fact that all subspace-based estimation methods suffer an abrupt degradation in performance as either the signal-to-noise ratio (SNR) or the number of available snapshots N drops below certain threshold values has been known for a long time [1, 2], only recently was it emphasised that true ML estimation suffers from a similar "performance breakdown" phenomenon, but at a different threshold. Specifically, in [3] a study of ML threshold conditions was conducted for a single source and a single snapshot, where the ML estimator is equivalent to the conventional beamformer. In [4, 5], we suggested a technique based on the generalised likelihood-ratio test (GLRT), whereby a set of DOA estimates is treated as a set of proper ML estimates if the LR generated by this set exceeds that generated by the exact (true) parameters for the given sufficient statistics.

For performance investigations within Monte-Carlo simulations, this comparison can be directly performed since the exact covariance matrix of the simulated data is known. On the other hand, for practical applications, this approach is still valid since the exact LR value generated by the true covariance matrix is replaced by some threshold value. Fortunately, the exact LR is a random number with a *scenario-free* p.d.f. that is completely specified by the number of the antenna sensors M and the number of independent training snapshots N [see [8](15)], hence a threshold can be pre-calculated based on any desired probability of false identification as an outlier for some true ML solution.

This technique was applied to both linear uniform and sparse antenna arrays (with a uniform linear co-array) in [4, 5], where we were able to demonstrate that subspace-based method-specific outliers could be reliably identified by this technique, since LRs generated by sets of estimates containing outlier(s) have significantly smaller LRs than the (scenario-free) threshold value. For linear antenna arrays, the LR optimisation technique, which heavily relied upon Toeplitz properties of the relevant covariance matrices, was able to provide ML DOA estimates without outliers, even though MUSIC mostly failed [4, 5]. Yet, if the SNR and/or sample support continued to degrade, we observed the ML "performance breakdown" whereby sets of DOA estimates containing severely erroneous estimates still generated high LRs, in many cases exceeding the exact LR bound. Clearly, only the intermediate region between the threshold conditions of some subspace-based technique (*eg*. MUSIC) and the ML threshold conditions could be considered for rectification; ML performance breakdown obviously cannot be neither detected nor rectified within the ML paradigm.

In this paper, we are interested in the similar phenomenon in uniform circular antenna arrays (UCAs). Recently, two techniques have been reported (the Global Matched filter approach [6] and the Generalised Rectification scheme [7]) that in the threshold area provided better performance than MUSIC for UCA. Yet, analysis of the "gap" between MUSIC and ML threshold conditions within our approach became possible only since an appropriate outlier mitigation scheme, that does not rely upon the particular (Toeplitz) properties, was now developed [8].

2. DATA AND ALGORITHM DESCRIPTION

Consider a UCA with M omnidirectional sensors located at positions $\{\rho\cos(k\delta), \rho\sin(k\delta)\}$, $(k = 0, \ldots, M-1,\ \delta = 2\pi/M)$. For simplicity, we assume that the array sensors and the sources are co-planar; ρ is the radius of the UCA, measured in wavelength λ units. The steering vector associated with the azimuthal angle (DOA) $\theta \in [0, 2\pi]$ is:

$$\boldsymbol{s}(\theta) = \left[s(\theta),\ s(\theta - \frac{2\pi}{M}),\ \ldots,\ s(\theta - \frac{2\pi(M-1)}{M})\right]^T \tag{1}$$

where $s(\theta) \equiv \exp[2\pi i \frac{\rho}{\lambda}\cos\theta]$. Assuming $m < M$ uncorrelated Gaussian sources, we may express the vector of observed sensor outputs (the "snapshot") at time t as

$$\boldsymbol{y}(t) \ = \ S(\boldsymbol{\theta})\,\boldsymbol{x}(t) + \boldsymbol{\eta}(t) \qquad \text{for} \quad t = 1, \ldots, N \tag{2}$$

where $\boldsymbol{x}(t) \in \mathcal{C}^{m\times 1}$ are the Gaussian signal amplitudes with DOAs $\boldsymbol{\theta} \equiv [\theta_1, \ldots, \theta_m]^T$ and powers $P \equiv \operatorname{diag}[p_1, \ldots, p_m]$, the array-signal manifold matrix is $S(\boldsymbol{\theta}) \equiv [\boldsymbol{s}(\theta_1), \ldots, \boldsymbol{s}(\theta_m)] \in \mathcal{C}^{M\times m}$, and $\boldsymbol{\eta}(t) \in \mathcal{C}^{M\times 1}$ is Gaussian white noise of power p_0:

$$\boldsymbol{x}(t) \sim \mathcal{CN}(m,\,0,\,P), \quad \boldsymbol{\eta}(t) \sim \mathcal{CN}(M,\,0,\,p_0 I_M) \tag{3}$$

where $\mathcal{CN}(M,\, 0,\, R)$ denotes a complex (circular) Gaussian distribution of dimension M with zero mean and covariance matrix R. Therefore the input data is described by the complex Gaussian distribution $\mathcal{CN}(M,\, 0,\, R)$, where $R = S(\boldsymbol{\theta})\, P\, S(\boldsymbol{\theta})^H + p_0\, I_M$. We assume that the snapshots are statistically independent and so the sufficient statistic for inferences regarding this data is the direct data covariance (DDC) matrix $\hat{R} = \frac{1}{N}\sum_{t=1}^{N} \boldsymbol{y}(t)\, \boldsymbol{y}^H(t)$ where $N\hat{R}$ is described by the complex Wishart distribution $\mathcal{CW}(N, M; R)$.

Given the antenna array geometry, the sufficient statistic $\hat{R}$ and the known (or previously estimated) number of sources m, we implement the general GLRT-based algorithm described in our companion paper [8]. A key role in this algorithm is played by the sphericity test and its LR analysis of the subspace-derived (*eg*. MUSIC) solutions, and direct optimisation of this LR over the set of DOA and power parameters. (The source powers and white-noise power are estimated by traditional means, *eg*. [9].) For Gaussian mixtures, the LR for this test is

$$\gamma(\tilde{R}) = \left(\frac{\det(\tilde{R}^{-1}\hat{R})}{\left[\frac{1}{M}\,\mathrm{tr}\,(\tilde{R}^{-1}\hat{R})\right]^M} \right)^N \equiv \gamma_0^N(\tilde{R}) \tag{4}$$

($0 < \gamma(\tilde{R}) \leq 1$), with the set of estimated parameters $\{\hat{\boldsymbol{\theta}},\, \hat{\boldsymbol{p}},\, \hat{p}_0\}_{ML}$ being those that yield the global maximum of the LR $\gamma(\tilde{R})$, which coincides with the global maximum of the (stochastic) likelihood function [4, 5].

Since the ML-generated covariance matrix $R_{ML}(\hat{\boldsymbol{\theta}},\, \hat{\boldsymbol{p}},\, \hat{p}_0)$ belongs to the same admissible set as the true covariance matrix R, we obviously have

$$\gamma(R_{ML}) > \gamma(R)\,. \tag{5}$$

Therefore any set of parameters estimates $\{\hat{\boldsymbol{\theta}},\, \hat{\boldsymbol{p}},\, \hat{p}_0\}$ (including our MUSIC-generated ones) that are deemed to be sufficiently close to ML estimates must obey the similar condition

$$\gamma(\tilde{R}) \geq \gamma(R)\,. \tag{6}$$

However, when $\gamma(\tilde{R}) < \gamma(R)$, then the set of (MUSIC-generated, say) estimates is not statistically close enough to the ML estimates, and should be disregarded, since it contains one or more outliers.

Naturally, the strict comparison in (6) can be performed only for Monte-Carlo simulations where the exact covariance matrix R is known. For practical applications we adopt the thresholding

$$\gamma_0(\tilde{R}) > \alpha \tag{7}$$

where the threshold value α is pre-calculated using the scenario-free p.d.f. $f(\gamma_0)$ for the exact LR, and any prescribed probability that some true ML solution would be wrongly identified as an outlier [8].

3. SIMULATION RESULTS AND DISCUSSION

Consider a ten-sensor UCA ($M = 10$) with $\rho = (\lambda/4)\sin(\pi/M)$ and $\rho = 0.809\lambda$ so that the distance between two neighbouring sensors is $\lambda/2$. Suppose there are $m = 5$ independent sources with 20dB SNR per source, and $N = 300$ snapshots. Four sources are uniformly distributed in azimuth, while the fifth DOA is varied in separation from the fourth:

$$\boldsymbol{\theta}_5 = [0^\circ,\, 30^\circ,\, 60^\circ,\, 90^\circ,\, \theta_5]\,. \quad \theta_5 = \{93^\circ,\, 91.6^\circ,\, 91^\circ\}\ . \tag{8}$$

These three separate DOA values for the fifth source θ_5 have been specifically selected to demonstrate the transition from the MUSIC-specific performance-breakdown conditions, that can be efficiently rectified by our GLRT-based technique ($\theta_5 = 93^\circ$), to the ML performance-breakdown conditions that could not be rectified within the ML paradigm ($\theta_5 = 91^\circ$).

The Cramér–Rao bound (CRB) for the fifth source is 0.36°, 0.75° and 1.46° respectively. Thus for $\theta_5 = 93^\circ$, the CRB predicts that the ultimate DOA estimation accuracy is sufficient for reliable separation of the fourth and fifth sources, since it is significantly smaller than the inter-source separation of 3°. At the other extreme, the CRB exceeds the 1° separation. Thus reliable separation of the last two sources is impossible. Still, all three locations of the fifth source are close enough to the fourth source to cause a significant number of MUSIC outliers. Outlier analysis has been conducted first by the impractical "strict" condition (6) in order to evaluate the potential capabilities of ML DOA estimation, but also in order to assess additional degradations caused by practical thresholding (7) with probabilities of incorrect identification set at 10^{-2} and 10^{-3}. We computed the scenario-free p.d.f. $f(\gamma_0)$ by direct Monte-Carlo simulation with 10^5 trials, which leads to the thresholds $P(LR < 0.796) = 10^{-2}$ and $P(LR < 0.778) = 10^{-3}$. The function (not illustrated here) is very well localised within the range $0.8 \lesssim LR \lesssim 0.9$.

Fig. 1 shows the sample p.d.f.'s of the LR for the two scenarios with $\theta_5 = 93^\circ$ and $\theta_5 = 91^\circ$. Fig. 1(a) and (b) show the exact LR $\gamma_0(R)$ ("threshold"), the MUSIC LR $\gamma_0(\tilde{R})$ ("MUSIC"), and the LR after the first optimisation (using the MUSIC DOAs and estimated powers as initialisers) ("optimisation", see [8]). We see that for $\theta_5 = 93^\circ$, the vast majority of trials (in fact, 91.9%) resulted in MUSIC LRs that were extremely poor compared with the "proper ML" bound represented by the exact LR threshold, situated between about 0.8 and 0.9. The application of this LR maximisation routine in the vicinity of the MUSIC-generated estimates improved matters only slightly, since the MUSIC estimates were generally so far from ML ones. In fact, only 24.4% of trials resulted in LRs after optimisation that exceeded the ML bound.

For $\theta_5 = 91^\circ$ (Fig. 1(b), introduced to illustrate ML performance breakdown) the MUSIC estimates appear to have an improved LR, however, 98.6% of MUSIC trials were below the ML bound (and hence would be declared to contain outliers), and only 6.9% of optimisation trials exceeded the ML bound. This behaviour is a typical manifestation of the onset of ML performance breakdown, whereby severely erroneous DOA estimates may generate optimally high LRs.

Analysis of the maximum DOA estimation error supports this conclusion. For $\theta_5 = 93^\circ$, the p.d.f. of such errors (not introduced here) shows MUSIC and optimisation concentrated around 150°, while after refinement the maximum DOA errors are concentrated around 1°. For $\theta_5 = 91^\circ$, the distinction is similar but less marked (approximately 130° and 4° respectively). While here optimisation added 55 trials to the 14 originally exceeding the ML bound, in all cases severe outliers remained. The physical explanation of such behavior is clear: with such a small separation between the two last sources, even only four properly estimated DOAs with the last one being between the two close sources results in a quite high LR, while the fifth DOA estimate is selected in insignificant spurious LR maximum. By the way, this means that in most cases the detection-estimation procedure should identify such a scenario as four-source, rather than five-source. Yet, in our problem we are forced to select an erroneous DOA estimate, while the correct four

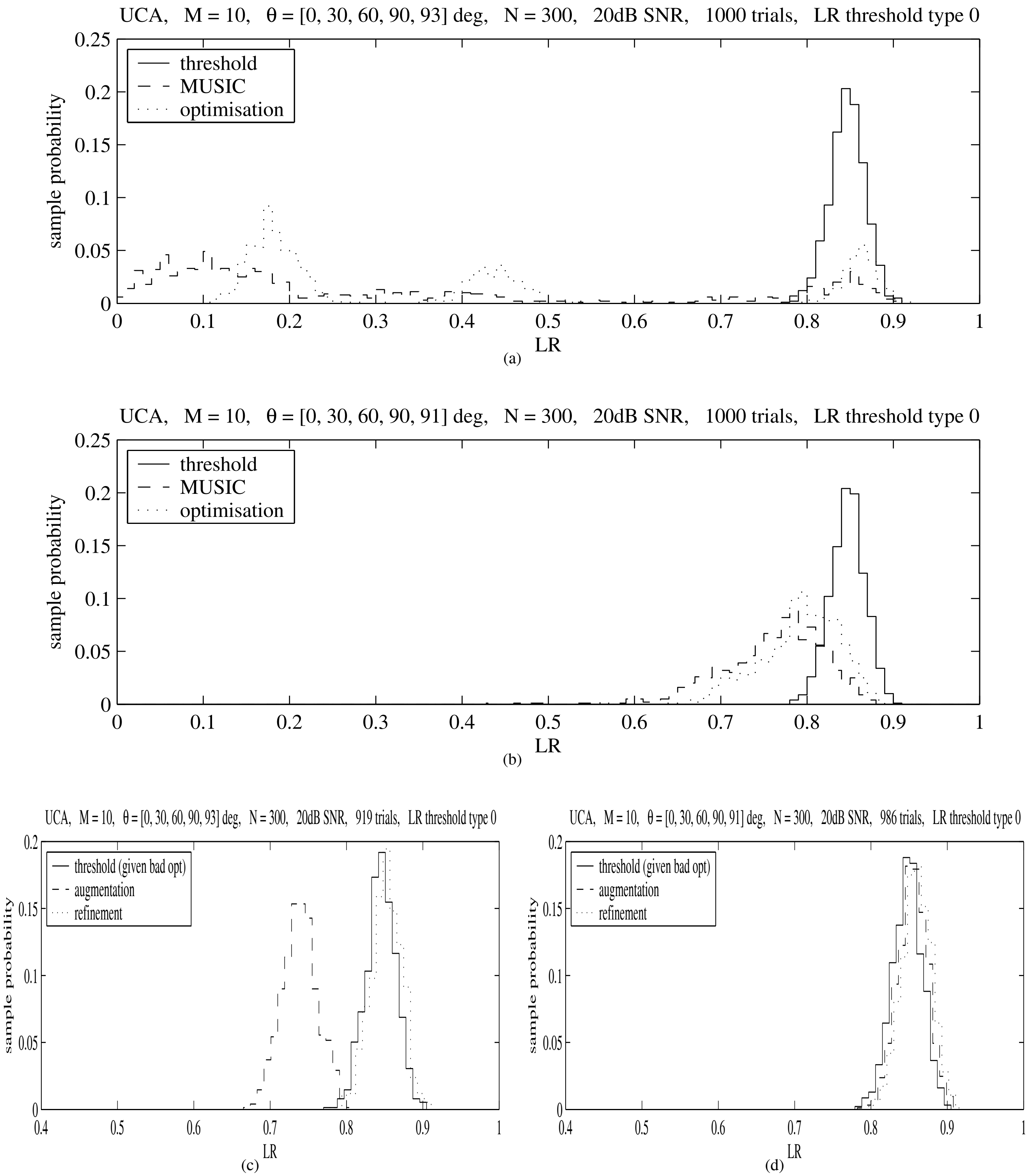

Fig. 1. Sample probability distributions of the likelihood ratio for $\theta_5 = 93^\circ$ (a,c), and $\theta_5 = 93^\circ$ (b,d) for an exact LR threshold.

θ_5	93°	91.6°	91°	93°	91.6°	91°	93°	91.6°	91°
LR(MUSIC) $<$ LR(threshold)	91.9%	100%	98.6%	84.2%	99.5%	70.8%	83.1%	98.4%	56.3%
LR(opt) $\geq$ LR(threshold)	24.4%	0.2%	6.9%	24.4%	1.7%	51.2%	24.5%	6.9%	68.9%
LR(MUSIC) $<$ LR(threshold) & good opt	16.3%	0.2%	5.5%	8.6%	1.2%	22.0%	7.6%	5.3%	25.2%
max CRB in $\boldsymbol{\theta}$ (degrees)	0.36	0.75	1.46	0.36	0.75	1.46	0.36	0.75	1.46
RMSE in $\hat{\boldsymbol{\theta}}$ after refinement (degrees)	0.68	1.82	2.94	1.32	1.99	3.35	1.39	2.01	4.20

Table 1. Sample five-source performance statistics for an exact LR threshold (left three columns); for a LR threshold equal to 1% probability of false alarm (centre three columns), and for a LR threshold equal to 0.1% probability of false alarm (right three columns).

of them already resulted in a sufficiently high LR.

Fig. 1(c) and (d) show similar LR p.d.f.'s, but for the successive stages in the outlier mitigation method. Each subfigure shows the exact LR $\gamma_0(R)$ distribution for those trials where MUSIC estimates did not exceed the ML bound ("threshold (given bad opt)"), the LR after replacing one outlier by the most likely new DOA [8] ("augmentation"), and the LR after the second optimisation (using the "augmented" DOAs and estimated powers as initialisers) ("refinement").

In this case, *all trials* were (if necessary) correctly rectified by our approach, *ie*. the LR after refinement (or after the first optimisation, if successful) always exceeded the ML bound defined by the LR of the exact covariance matrix. In spite of this, the sample RMSE of 0.68° was significantly greater than the CRB of 0.36° (see Table 1). This result indicates that the usual assumptions made regarding the second-order expansion of the likelihood function in the vicinity of the exact parameters is not sufficiently accurate for this "threshold" region.

Analysis of the maximum DOA estimation error demonstrates that though in all "rectified" trials, severe outliers have been removed, they have been replaced by the DOA estimates in the vicinity of the true ones, but this vicinity was found to be too large to provide sufficiently accurate DOA estimates that meet the CRB prediction. Therefore, in these two cases $\theta_5 = 93^\circ$ and $\theta_5 = 91^\circ$, despite extremely high LR maximisation efficiency, with all 1000 trials found to result in both cases in "better than the true" DOA estimates (in terms of the maximised LR), the actual performance is significantly different. For the first scenario, the GLRT-based approach provided reliable outlier mitigation and accurate identification of all five sources. The second scenario, far beyond the resolution capabilities of any ML technique illustrates "ML performance breakdown" behavior, when severely erroneous solutions still generate optimally high LRs.

Table 1 also introduces results for "practical" thresholding (7). As one would expect, such a thresholding let some MUSIC generated solutions that were sufficiently close (in terms of LR) to $\gamma(R)$ to be treated as ML-optimal, while the strict inequality (6) still did not occur. Remarkably, for both threshold values and $\theta_5 = 93^\circ$, only 244 or 245 scenarios were accepted where previously direct LR optimisation in the strict case resulted in "better than the true" solutions. This means that for this practically important scenario, all outliers have been properly identified and replaced by sufficiently accurate DOA estimates, despite the fact that accuracy in the "refined" data has been degraded compared with the ideal case (1.39° instead of 0.68°).

4. CONCLUSIONS

Analysis of the MUSIC-derived estimates and the estimates provided by our proposed GLRT-based technique has revealed a significant "gap" between the performance breakdown (threshold) conditions for this subspace-based DOA estimation technique and ML estimation in uniform circular arrays. We have demonstrated that MUSIC-produced outliers could be reliably identified and rectified by the introduced technique, unless conditions for the ML performance breakdown are met. In the latter case, our GLRT-based technique produces solutions that have optimally high LRs, being still severely erroneous DOA estimates. These conditions constitute the ultimate limit, that could not be overcome by any technique within the ML paradigm.

5. REFERENCES

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

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

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

[4] Y.I. Abramovich and N.K. Spencer, "DOA estimation performance breakdown: a new approach to prediction and cure," in *Proc. EUSIPCO-2002*, Toulouse, France, 2002.

[5] Y.I. Abramovich, N.K. Spencer, and A.Y. Gorokhov, "Bounds on maximum likelihood ratio — Part I: Application to antenna array detection-estimation with perfect wavefront coherence," *IEEE Trans. Sig. Proc.*, submitted 10 Mar 2002, accepted 31 Jul 2003, scheduled to appear Jun 2004.

[6] J.-J. Fuchs, "On the application of the global matched filter to DOA estimation with uniform circular arrays," *IEEE Trans. Sig. Proc.*, vol. 49, no. 4, pp. 702–709, Apr. 2001.

[7] P. Forster, "Generalized rectification of cross spectral matrices for arrays of arbitrary geometry," *IEEE Trans. Sig. Proc.*, vol. 49, no. 5, pp. 972–978, May 2001.

[8] 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-2004*, Montreal, Canada, 2004, accepted.

[9] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai, "*Exact and large sample maximum likelihood techniques for parameter estimation and detection in array processing*," in *Radar Array Processing*, S. Haykin, J. Litva, and T.J. Shepherd, Eds., pp. 99–151. Springer-Verlag, Berlin, 1993, chapter 4, Springer Series in Information Sciences, vol 25.