LOCALLY OPTIMAL MAXIMUM-LIKELIHOOD ESTIMATE OF A TOEPLITZ MATRIX OF GIVEN RANK

Y. I. Abramovich

Surveillance Systems Division (SSD),
Defence Science and Technology
Organisation (DSTO), PO Box 1500,
Salisbury SA 5108, Australia
`yuri@cssip.edu.au`

N. K. Spencer

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

ABSTRACT

We derive an algorithm to compute a maximum-likelihood (ML) estimate of a Toeplitz covariance matrix T whose rank is known *a priori* that is "locally optimal", by maximising the likelihood ratio in the neighborhood of T. This problem arises, for example, in the detection and/or direction-of-arrival (DOA) estimation of m uncorrelated plane-wave sources using a nonuniform (sparse) M-sensor linear antenna array, where $m \geq M$ (the "superior" case), but the problem is important in its own right, and has application to other areas of signal processing and communications. The algorithm relies upon the solution of a convex linear programming (LP) problem, whose feasibility is guaranteed.

1. INTRODUCTION

While this ML estimation problem could be considered in isolation, it is more instructive to present the problem in one of its natural contexts, for example detection and DOA estimation with nonuniform (sparse) linear antenna arrays [1, 2]. Such arrays attract considerable attention, especially when the number of sensors (M) is limited, since they can offer a significant improvement over the standard M-sensor uniform array in terms of both estimation accuracy and the maximum estimated number of uncorrelated sources. Most DOA estimation studies assume that the number of uncorrelated Gaussian sources m is known or somehow accurately estimated. Naturally, in practical applications where nothing is known about the signal environment *a priori*, this step is not trivial, both for *conventional* scenarios, where $m < M$, and especially for *superior* scenarios, where $m \geq M$.

In the conventional case, the detection problem can be addressed by the traditional AIC and MDL techniques [3] or their later variants [4]. In the superior case, the standard AIC and MDL methods are inappropriate simply because the covariance matrix R has no noise subspace. A forthcoming paper [2] presents and discusses the overall detection technique for superior scenarios of independent Gaussian sources for the class of *fully augmentable* nonuniform linear arrays (NLA's).

Fully augmentable NLA's belong to the family of "minimum redundancy" arrays [5], and are defined as having no gaps in the set of $M(M-1)/2$ intersensor separations, in other words, having a full or complete co-array that is identical to the corresponding (M_α-sensor) uniform linear array (ULA). Such geometries consequently permit DOA estimation of a superior number

$$1 \leq m < M_\alpha \gg M \tag{1}$$

of independent sources, and traditionally, the direct augmentation approach (DAA) of Pillai *et al*. [6] has been considered adequate. However, we have demonstrated [7] that the DOA estimates provided by MUSIC applied to the M_α-variate augmented sample covariance matrix $\hat{T}$ are far from asymptotically optimal. This is important because a naive attempt to directly apply existing AIC/MDL criteria to $\hat{T}$ also fails; indeed, the DAA matrix $\hat{T}$ is not statistically p.d. for any reasonable sample size N [2]. For this reason, the set of estimated spatial covariance lags forming the M_α-variate Toeplitz matrix $\hat{T}$ comprises only the first initialization step of the proposed detection scheme.

The philosophy of this scheme is quite simple. Given the set of all M_α measured spatial covariance lags provided by the standard M-variate direct data covariance (DDC) matrix $\hat{R}$, we find a set $\{\mathcal{T}_\mu\}$ ($\mu = 1, \ldots, M_\alpha - 1$) of p.d. Toeplitz covariance matrices, each having $(M_\alpha - \mu)$ equalized minimum eigenvalues. A selection of the most appropriate model (*ie*. number of sources) from the candidates $\mathcal{T}_\mu$ may then be made by the traditional information criteria or Bayesian model techniques.

Here we specifically present the derivation of a "locally optimal" ML estimate of a positive-definite (p.d.) Toeplitz covariance matrix, which is a crucial piece of the entire detection scheme.

2. PROBLEM FORMULATION

Consider an M-element NLA with sensors located at positions

$$\boldsymbol{d} \equiv [d_1 \equiv 0,\ d_2,\ d_3,\ \ldots,\ d_M \equiv M_\alpha - 1] \tag{2}$$

restricted to integer values of d_j measured in half-wavelength units. *Fully augmentable* arrays have the property that the set of all intersensor distances

$$\mathcal{D} \equiv \{d_j - d_k \mid j, k = 1, \ldots, M;\ j \geq k\} \tag{3}$$

is complete, *ie*. $\mathcal{D} \equiv \{0, \ldots, d_M\}$. Recall that the *co-array* $\boldsymbol{c}$ of a linear array $\boldsymbol{d}$ is the sorted set of nonduplicated elements of $\mathcal{D}$, thus the M_α-element co-array corresponding to every fully-augmentable NLA is uniform. Furthermore, for independent Gaussian sources, this property indicates that up to $(M_\alpha - 1)$ sources may be identified (detected and estimated), which reveals the potential ability to identify up to $(M_\alpha - 1)$ sources, given sufficient statistics via the DDC matrix R. We assume that Gaussian processes are observed as a mixture of m uncorrelated plane waves with DOA's $\boldsymbol{\theta} \equiv [\theta_1, \ldots, \theta_m]^T$, powers $P \equiv \mathrm{diag}\,[p_1, \ldots, p_m]$

and Gaussian white noise of power p_0:

$$\boldsymbol{y}(t) = S(\boldsymbol{\theta})\boldsymbol{x}(t) + \boldsymbol{\eta}(t) \qquad \text{for} \quad t = 1, \ldots, N \tag{4}$$

where $\boldsymbol{y}(t) \in \mathcal{C}^{M\times 1}$ is the vector of observed sensor outputs (the "snapshot"), $\boldsymbol{x}(t) \in \mathcal{C}^{m\times 1}$ are the Gaussian signal amplitudes

$$\mathcal{E}\left\{\boldsymbol{x}(t_1)\,\boldsymbol{x}^H(t_2)\right\} = \begin{cases} P & \text{for} \quad t_1 = t_2 \\ 0 & \text{for} \quad t_1 \neq t_2 \,, \end{cases} \tag{5}$$

$\boldsymbol{\eta}(t) \in \mathcal{C}^{M\times 1}$ is additive white Gaussian noise, $\mathcal{C}^{p\times q}$ is the space of $p\times q$ complex-valued matrices, and $\mathcal{E}\{\cdot\}$ is the expectation operator. The array manifold matrix is $S(\boldsymbol{\theta}) \equiv [\boldsymbol{s}(\theta_1), \ldots, \boldsymbol{s}(\theta_m)] \in \mathcal{C}^{M\times m}$, where each

$$\boldsymbol{s}(\theta_j) = \left[1,\, \exp\left(i\pi d_2 \sin\theta_j\right),\, \ldots,\, \exp\left(i\pi d_M \sin\theta_j\right)\right]^T \tag{6}$$

is a so-called "steering" vector. The set of independent snapshots $\boldsymbol{y}(t) \in \mathcal{C}^{M\times 1}$ originates from a complex Gaussian distribution $\mathcal{CN}(M, 0, R)$, where

$$R = S(\boldsymbol{\theta})PS(\boldsymbol{\theta})^H + p_0 I_M \,. \tag{7}$$

Given N independent snapshots, the sufficient statistic for DOA estimation is the DDC (sample) matrix

$$\hat{R} = \frac{1}{N}\sum_{t=1}^{N} y(t)\, y^H(t)\,. \tag{8}$$

For the corresponding M_α-element ULA, the array manifold matrix $A(\boldsymbol{\theta}) \equiv [\boldsymbol{a}(\theta_1), \ldots, \boldsymbol{a}(\theta_m)] \in \mathcal{C}^{M_\alpha\times m}$ is of Vandermonde structure, with

$$\boldsymbol{a}(\theta_j) = \left[1,\, \exp(i\pi w_j),\, \ldots,\, \exp\left(i\pi[M_\alpha - 1]w_j\right)\right]^T \tag{9}$$

where the spatial frequency is $w_j \equiv \sin\theta_j$. By definition, the M-element NLA can be viewed as a subarray of the M_α-element ULA. Thus the M-variate snapshot vector $\boldsymbol{y}(t)$ can be presented as a subset of the M_α-variate ULA output $\boldsymbol{Y}(t) \in \mathcal{C}^{M_\alpha\times 1}$:

$$\boldsymbol{y}(t) = L\boldsymbol{Y}(t) \tag{10}$$

where L is the $M \times M_\alpha$ binary selection (or incidence) matrix with L_{jk} equal to unity in the j^{th} row and d_k^{th} column, and zero otherwise. Correspondingly, the M-variate Hermitian covariance matrix R is linked to the M_α-variate Toeplitz covariance matrix T for the M_α-element ULA by the linear transformation

$$R = LTL^T \tag{11}$$

with

$$T = A(\boldsymbol{\theta})PA(\boldsymbol{\theta})^H + p_0 I_{M_\alpha}\,. \tag{12}$$

The philosophy of the DAA [6] is to estimate the Toeplitz matrix covariance lags

$$\hat{T} = \{t_{j-k}\}_{j,k=1}^{M_\alpha} \tag{13}$$

by simply averaging over the set of corresponding (possibly redundant) covariance lags taken from $\hat{R}$:

$$t_{j-k=\kappa} = \frac{\sum_{j,k=1}^{M} \hat{R}_{jk}\,\delta(\kappa,\, d_j - d_k)}{\sum_{j,k=1}^{M} \delta(\kappa,\, d_j - d_k)}\,, \qquad j > k \tag{14}$$

where $\delta(a, b)$ is the generalized Kronecker delta function. Clearly the augmented covariance matrix $\hat{T}$ is not necessarily p.d., and the distribution of its $(M_\alpha - m)$ smallest (noise-subspace) eigenvalues is of major concern.

Detection cannot be based on the direct testing of the equality of the $(M_\alpha - m)$ smallest eigenvalues of the DAA matrix $\hat{T}$, since these noise-subspace eigenvalues fluctuate considerably and, moreover, $\hat{T}$ is generally not p.d. for any reasonable sample size N. We require some new technique that exploits the sufficient statistic $\hat{R}$ in a more appropriate fashion.

3. ESTIMATION OF $\hat{T}$ WITH A GIVEN RANK

Given the DDC matrix $\hat{R}$, we now wish to compute a p.d. Toeplitz matrix T_μ, whose $(M_\alpha - \mu)$ smallest eigenvalues are equal, that is "sufficiently close" to $\hat{R}$, via the transformation $\hat{R}_\mu = LT_\mu L^T$, in the ML sense.

The main idea behind our algorithm is quite simple. Since small perturbations in the sample covariance lags of $\hat{R}$ (with respect to the exact values in R) lead to significant fluctuations in the noise-subspace eigenvalues $\boldsymbol{\sigma}_n$ of the matrix $\hat{T}$, "inverse perturbations" in $\hat{T}$ that equalize up to the $(M_\alpha - m)$ smallest eigenvalues should not involve significant changes to the sample lags. Thus for $\mu = m$ we expect the matrix $\hat{R}_\mu = LT_\mu L^T$ to be close to $\hat{R}$, as measured by the likelihood ratio (LR). On the contrary, if some signal eigenvalue was attempted to be equalized with the noise eigenvalues (for $\mu > m$), there would be a noticeable degradation in LR.

Since the DAA matrix $\hat{T}$ is (generally) not p.d., we cannot use it as an initial value directly; rather the covariance lags generated by the DDC matrix $\hat{R}$ should be modified appropriately, in a manner as close as possible to $\hat{R}$ in the ML sense [8]. For example, the simplest idea is to properly load the DAA matrix:

$$T_0 = \hat{T} - \hat{\sigma}_{M_\alpha} I_{M_\alpha}\,. \tag{15}$$

On the other hand, even if the DAA matrix $\hat{T}$ is p.d., it could still be far away from the ML Toeplitz estimate for the given sufficient statistic $\hat{R}$. For this reason, it would be appropriate to instead choose the initial point to be a "locally optimal" ML estimate $\hat{T}_{ML}$

$$T_0 = \hat{T}_{ML}\,. \tag{16}$$

The matrix $\hat{T}_{ML}$ is defined as that which has maximum LR, and by "locally optimal", we mean that this applies in the neighborhood of the DAA matrix $\hat{T}$.

In principle, for sufficiently small perturbations in T_0, we may apply a first-order expansion of the LR

$$\gamma(T_\mu) = \left(\frac{\det({R_\mu}^{-1}\hat{R})}{\left[\frac{1}{M}\mathcal{T}r({R_\mu}^{-1}\hat{R})\right]^M}\right)^N \tag{17}$$

then directly maximize the LR in this neighborhood. Instead, we suggest a slightly different optimization criterion that is closely related to the ML one. Let

$$\hat{G}(R) = \hat{R}^{-\frac{1}{2}}R\hat{R}^{-\frac{1}{2}} \tag{18}$$

with

$$R = LTL^H = LH^HQHL^H \tag{19}$$

then the LR can be expressed as $\gamma(T) = [\gamma_0(T)]^N$, where

$$\gamma_0(T) = \frac{\prod_{k=1}^{M} \lambda_k^{-1}}{\left[\frac{1}{M}\sum_{k=1}^{M} \lambda_k^{-1}\right]^M} \tag{20}$$

where λ_k is the k^{th} eigenvalue of $\hat{G}(R)$. It should now be clear that the LR reaches its maximum of $\gamma_0(T) = 1$ only when $\hat{G}(R)$ is the identity matrix, so that $\lambda_1 = \lambda_2 = \ldots = \lambda_M$. Thus, instead of directly maximizing the LR $\gamma_0(T)$, we propose minimizing the difference between the eigenvalues for both the direct ($\hat{G}$) and inverse ($\hat{G}^{-1}$) matrices. Specifically, in a sufficiently small neighborhood of the initial p.d. Toeplitz matrix T_0, we iteratively solve two separate problems:

$$\text{Find} \qquad \min(\lambda_1 - \lambda_M) \tag{21}$$

and

$$\text{Find} \qquad \min(\lambda_M^{-1} - \lambda_1^{-1}) \tag{22}$$

where the eigenvalues are sorted in descending order $\lambda_1 > \lambda_2 > \ldots > \lambda_M > 0$; then simply select the solution with greater LR. The numerator of (20) makes this dual approach necessary, since a "one-sided" minimization (such as (21) alone) would not preclude $\lambda_M \to 0$ and a correspondingly tiny LR.

Instead of these two separate problems (21) and (22), we could consider the simultaneous minimization of the eigenspectra of $\hat{G}$ and $\hat{G}^{-1}$, or some alternating rule. While such options could be explored in the future, we have found the above dual approach to be quite satisfactory.

Since the optimization criteria (21) and (22) are not exactly the ML criteria, we must treat our results as suboptimal in the ML sense; nevertheless, for simplicity we call the result the "locally optimal ML Toeplitz estimate", based on efficiency shown in [2].

Let us now describe the iterative routine that solves the direct problem (21). We start from some admissible (p.d.) M_α-variate Toeplitz matrix T_0 (that transforms into the M-variate p.d. Hermitian matrix R_0 via (19)), and at each iteration t solve the LP problem:

$$\text{Find} \qquad \min(\alpha - \beta) \qquad \text{subject to} \tag{23}$$

$$0 < \boldsymbol{\lambda}_t + \mathcal{D}_t \boldsymbol{x} < \alpha \mathbf{1} \tag{24}$$

$$\boldsymbol{\lambda}_t + \mathcal{D}_t \boldsymbol{x} > \beta \mathbf{1} \tag{25}$$

$$-\varepsilon < x_k < \varepsilon \qquad \text{for} \quad k = 1, \ldots, 2(M_\alpha - 1) \tag{26}$$

$$\lambda_M(\hat{G}(T_0)) < \alpha, \beta < \infty \tag{27}$$

where $\boldsymbol{\lambda}_t$ is the M-variate vector of eigenvalues of $\hat{G}_t$, with

$$\hat{G}_{t+1} = \hat{G}_t + \sum_{k=1}^{2(M_\alpha-1)} x_{t+1,k} \hat{R}^{-\frac{1}{2}} L H^H F_k H L^H \hat{R}^{-\frac{1}{2}}. \tag{28}$$

The inequalities (26) ensure the perturbations are sufficiently small, so that the following first-order expansion is valid

$$\lambda_j(\hat{G}_{t+1}) = \lambda_j(\hat{G}_t) + \sum_{k=1}^{2(M_\alpha-1)} x_{t+1,k} \boldsymbol{g}_j^H \hat{R}^{-\frac{1}{2}} L H^H F_k H L^H \hat{R}^{-\frac{1}{2}} \boldsymbol{g}_j \tag{29}$$

where $\boldsymbol{g}_j$ is the j^{th} eigenvector of the matrix $\hat{G}_t$. The $M \times 2(M_\alpha - 1)$ matrix $\mathcal{D}_t$ is defined according to the first-order expansion (29) as

$$\mathcal{D}_t = \left[\boldsymbol{g}_j^H \hat{R}^{-\frac{1}{2}} L H^H F_k H L^H \hat{R}^{-\frac{1}{2}} \boldsymbol{g}_j\right]_{j=1,\ldots,M}^{k=1,\ldots,2(M_\alpha-1)}. \tag{30}$$

The inequalities (27) do not allow the minimum eigenvalue of $\hat{G}(T_t)$ to fall below its initial value $\lambda_M(\hat{G}(T_0))$. The overall result of this (direct problem) optimization is that the minimum eigenvalue of $\hat{G}_t$ (and hence R_t) remains positive, while the eigenvalue spread $\lambda_1(\hat{G}_t) - \lambda_M(\hat{G}_t)$ is minimized.

It is important to note that the controlled positive-definiteness of R_t *does not* guarantee that the M_α-variate Toeplitz matrix T_t remains p.d. throughout this optimization procedure. In fact, controlled perturbations in T_0 can lead to a higher LR for R_t, while T_t itself can become non-p.d. Interestingly, this property allows us to determine whether the DAA matrix $\hat{T}$ is the best (in the LR sense) *non-positive* Toeplitz estimate. In order to find a proper ML estimate, we have to ensure the positive-definiteness of the optimized Toeplitz matrix, and so we must augment the linear inequalities of our direct optimization problem (23)–(27) with the following:

$$-\mathcal{A}'(\boldsymbol{x}_t)\boldsymbol{x} - \boldsymbol{\sigma}(\boldsymbol{x}_t) < -p_0 \mathbf{1} \tag{31}$$

where $\mathcal{A}'(\boldsymbol{x}_t)$ is defined by

$$\mathcal{A}(\boldsymbol{x}_0) = \left\{\boldsymbol{v}_j^H(\boldsymbol{x}_0)\, F_k\, \boldsymbol{v}_j(\boldsymbol{x}_0)\right\}_{j=1,\ldots,M_\alpha}^{k=1,\ldots,2(M_\alpha-1)} \tag{32}$$

and where p_0 is the minimum eigenvalue of the initial matrix T_0.

Finally, step-size management is now defined by whichever of the first-order expansions

$$0 < \boldsymbol{\sigma}_\mu(\boldsymbol{x}_1) \simeq \boldsymbol{\sigma}_\mu(\boldsymbol{x}_0) + \mathcal{A}(\boldsymbol{x}_0)\boldsymbol{x}^*. \tag{33}$$

or (29) is most sensitive (typically the former).

The "inverse" optimization problem concerning $\hat{G}^{-1}(T_t)$ in (22) is solved in a similar fashion [2].

As was mentioned above, the final step is to simply select the fully iterated solution from the "direct" (23)–(27)+(31) and the "inverse" LP problems with the greater LR as the best approximation to the locally optimal ML estimate T_{ML} for the M_α-variate Toeplitz covariance matrix, given the M-variate sufficient statistics $\hat{R}$ and the initial solution T_0.

It should be clear that in many applications this estimate could be important in its own right. Moreover, it could be directly treated by the traditional AIC/MDL criteria [3]. In [2], we described in detail a method of subsequently minimizing the greatest of the set of noise-subspace eigenvalues, in other words, noise-subspace equalization, by a linear programming (LP) technique. This generates the desired sequence of p.d. Toeplitz matrices T_μ ($\mu = 1, \ldots, M_\alpha - 2$), each of the proper μ-plane-waves-plus-noise structure, that then permits detection of the correct number of sources m based on information theoretic criteria.

4. SIMULATION RESULTS

Fig. 1 shows the sample LR distribution $\gamma_0(T)$ corresponding to the simulation of $\boldsymbol{d}_5 = [0,\ 2,\ 5,\ 8,\ 9]$ conducted over a set of 1000 Monte-Carlo trials. We indicate the LR distributions for (a) the DAA matrix $\hat{T}$, (b) the directly loaded DAA matrix ("loaded $\hat{T}$"), (c) the "unconstrained" (not necessarily p.d.) solution T_{ML}^u, (d) the "constrained" p.d. Toeplitz ML estimate T_{ML}, and (e) the exact covariance matrix R. Strictly speaking, the distribution $\gamma(\hat{T})$ should not be calculated, since in 988 out of our 1000 Monte-Carlo trials, the DAA matrix $\hat{T}$ was not p.d. Nevertheless, the corresponding "contracted" matrix $L\hat{T}L^H$ may well be p.d.

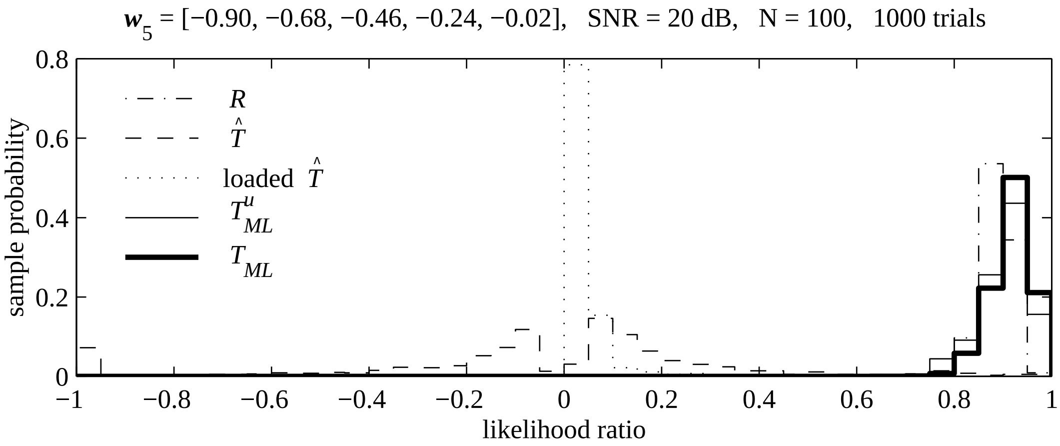

Fig. 1. Sample probability distribution of the likelihood ratio.

First, comparison of $\gamma(T^u_{ML})$ with $\gamma(\hat{T})$ proves that within the class of non-p.d. Toeplitz matrices (that yield M-variate p.d. Hermitian matrices LTL^H), LR maximization gives a result that is usually even higher than the LR of the true covariance matrix, $\gamma(R)$, and significantly better than that of the DAA matrix or its variants. We can see that $\gamma(\hat{T})$ spills over into the inappropriate negative domain, and even when the LR is positive, it lies in a region of very small values compared with $\gamma(R)$. Direct diagonal loading brings the LR distribution back into the positive domain, but also has small LR values compared with $\gamma(R)$. This once again demonstrates that both redundancy averaging and direct augmentation are actually very poor estimates (in the ML sense) of a structured (Toeplitz) covariance matrix.

Further LR maximization is shown by the distribution $\gamma(T_{ML})$, which leads to the remarkable observation that, while $\gamma(R)$ is not surprisingly in the high range 0.80 to 0.95, local ML optimization generally exceeds this by a significant margin. Indeed, $\gamma(T_{ML})$ has a noticeably right-skewed distribution that approaches the ultimate limit of unity. Note that this $\gamma(T_{ML}) > \gamma(R)$ behavior does not necessarily mean that T_{ML} really is a better estimate than the true matrix R; it is simply a better estimate that the ML criterion can suggest. Thus despite the local nature of our optimizations and the use of an indirect optimization criterion, any further attempts to improve the LR beyond $\gamma(T_{ML})$ would not necessarily give a significant performance improvement in either detection or estimation. Effectively, comparison of $\gamma(T_{ML})$ and $\gamma(R)$ allows us to conclude, from a practical viewpoint, that the problem of ML estimation of a structured (Toeplitz) covariance matrix has been solved by the proposed technique.

5. SUMMARY AND CONCLUSIONS

We have introduced a new technique for computing a ML estimate of a Toeplitz covariance matrix T. Rather than the true and exact ML estimate (that is unknown as yet), we have proposed a technique that yields a "local ML estimate" T_{ML}. This technique is initialized by the direct augmentation approach (DAA), then uses an iterated linear programming (LP) optimization, with sufficiently small perturbations at each iteration providing sufficient accuracy in the first-order expansion of the eigenvalues. This scheme guarantees local optimality of the perturbations at each iteration.

We stress that this proposed technique is applicable for the general case $M_\alpha = M$, that makes it useful for a range of problems far beyond the limited scope of NLA processing [9].

6. REFERENCES

[1] Y.I. Abramovich, N.K. Spencer, and A.Y. Gorokhov, "DOA estimation for noninteger linear antenna arrays with more uncorrelated sources than sensors," *IEEE Trans. Sig. Proc.*, vol. 48 (4), pp. 943–955, 2000.

[2] Y.I. Abramovich, N.K. Spencer, and A.Y. Gorokhov, "Detection-estimation of more uncorrelated Gaussian sources than sensors in nonuniform linear antenna arrays — Part I: Fully augmentable arrays," *IEEE Trans. Sig. Proc.*, To appear 2001.

[3] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust. Sp. Sig. Proc.*, vol. 33, pp. 387–392, 1985.

[4] P.M. Djurić, "A model selection rule for sinusoids in white Gaussian noise," *IEEE Trans. Sig. Proc.*, vol. 44 (7), pp. 1744–1757, 1996.

[5] A.T. Moffet, "Minimum-redundancy linear arrays," *IEEE Trans. Ant. Prop.*, vol. 16 (2), pp. 172–175, 1968.

[6] S. Pillai, Y. Bar-Ness, and F. Haber, "A new approach to array geometry for improved spatial spectrum estimation," *Proc. IEEE*, vol. 73 (10), pp. 1522–1524, 1985.

[7] Y.I. Abramovich, D.A. Gray, A.Y. Gorokhov, and N.K. Spencer, "Positive-definite Toeplitz completion in DOA estimation for nonuniform linear antenna arrays — Part I: Fully augmentable arrays," *IEEE Trans. Sig. Proc.*, vol. 46 (9), pp. 2458–2471, 1998.

[8] K.M. Grigoriadis, A.E. Frazho, and R.E. Skelton, "Application of alternating convex projection methods for computation of positive definite Toeplitz matrices," *IEEE Trans. Sig. Proc.*, vol. 42 (7), pp. 1873–1975, 1994.

[9] J.P. Burg, D.G. Luenberger, and D.L. Wenger, "Estimation of structured covariance matrices," *Proc. IEEE*, vol. 70 (9), pp. 963–974, 1982.