A MODIFIED LIKELIHOOD RATIO TEST FOR DETECTION-ESTIMATION IN UNDER-SAMPLED TRAINING CONDITIONS

Yuri I. Abramovich and Ben A. Johnson

Defence Science and Technology Organisation (DSTO), ISR Division
e-mail: yuri.abramovich@dsto.defence.gov.au
RLM Management, Pty. Ltd. *and* University of South Australia (ITR)
e-mail: ben.a.johnson@ieee.org
Edinburgh, SA 5111, Australia

ABSTRACT

For a number of training samples T that do not exceed the number of antenna elements M, we propose a non-degenerate normalized LR test that can be used in various detection-estimation problems. For the null hypothesis this test is described by a scenario-free probability density function.

1. INTRODUCTION

In many practical problems in adaptive detection-estimation, the number of representative training samples is very limited and therefore additional *a priori* assumptions are usually imposed to make smaller sample support sufficient for acceptable detection-estimation performance.

One of the well-known families of this kind is the one with low signal subspace dimension. Here the number m of the covariance matrix eigenvalues that exceed the minimal eigenvalue (equal to ambient white noise power) is significantly smaller then the matrix dimension M $(m < M)$.

In the most general case an admissible covariance matrix could be introduced in the form

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

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 localization of the "signal subspace" of such a "low rank" covariance matrix, the minimum sample support (ie the number of independent identically distributed training samples) is equal to m. This fact has been heavily exploited for justification of the well-known loaded sample matrix inversion (LSMI) algorithm in [1][2], the Hung-Turner fast projection adaptive beamformer [3][4], and "fast maximum likelihood" [5].

For example, in [6] it has been first analytically proven that under rather mild assumptions regarding covariance matrix eigenvalues in (1), average SNR losses for the LSMI technique, compared with the clairvoyant filter are equal to approximately 3dB for sample support T, where

$$T \gtrsim 2m \quad (2)$$

Work funded under DSTO/RLM R&D CA 290905

while for the traditional SMI technique the required sample support is equal to $T \gtrsim 2M$ for the same average loss [7].

Similarly, it is well known that for strong enough signal-to-noise ratio, multiple signal classification (MUSIC) can provide accurate DOA estimates for the number of training samples equal to the number of independent sources m.

While adaptive filters and DOA estimation exist for under-sampled $(T < M)$ training conditions, the modern GLRT-based detection-estimation techniques do not embrace this scenario, mainly because proper likelihood ratios have not been yet introduced.

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 exists and is non-degenerate even under under-sampled training conditions:

$$\mathcal{L}(X_T, R) = \left[\frac{1}{\pi \det R} \exp\{-\mathrm{Tr}[R^{-1}\hat{R}]\}\right]^T \quad (3)$$

where

$$\hat{R} = \frac{1}{T}\sum_{j=1}^{T} x_j x_j^H, \;\; T < M, \;\; R > 0 \quad (4)$$

But the standard approach to form the likelihood ratio (see [8]) as

$$LR(R) = \frac{\mathcal{L}(X_T, R)}{\max_R \mathcal{L}(X_T, R)} \quad (5)$$

can not be used here, since $\det \hat{R} = 0$ in this case. For $T \geqslant M$, this approach leads to (after the methodology in [8])

$$LR(R) = \left[\frac{\det R^{-1}\hat{R}\exp M}{\exp\{\mathrm{Tr}R^{-1}\hat{R}\}}\right]^T \leqslant 1 \quad (6)$$

since

$$\max_R \mathcal{L}(X_T, R) = \left[\frac{\exp(-M)}{\pi \det \hat{R}}\right]^T, \;\; \text{for } R = \hat{R} \quad (7)$$

In [9], we demonstrated that the $LR(R)$ (6) as well as it's variants such as the "sphericity test" [8] are instrumental for GLRT-based detection-estimation due to the invariance property:

$$f[LR(R_0)] = f\left[\frac{\det \hat{C}\exp M}{\exp\{\mathrm{Tr}\;\hat{C}\}}\right] \quad (8)$$

where $\hat{C} \sim \mathcal{CW}(T \geqslant M, M, I_M)$, ie $\hat{C}$ is a random matrix with scenario-free complex Wishart p.d.f., fully specified by M and T [10]. This invariance property, along with the straightforward observation

$$\max_{R \in \mathcal{R}} LR(R) > LR(R_0), \quad R_0 \in \mathcal{R} \tag{9}$$

allows one to address many complicated detection-estimation problems that would not be properly addressed by conventional detection-estimation techniques [11],[12].

Obviously, for under-sampled training conditions we would like to have a similar instrument, and thus for the under-sampled training data that belong to the family (1) we need the likelihood ratio $LR_u(R)$ that satisfies the following conditions.

a) **Normalization condition**:

$$0 < LR_u(R) \leqslant \text{ constant} \tag{10}$$

b) **Transition behavior**: $LR_u(R)$ should be an "analytic extension" of the $LR(R)$ (6), ie

$$LR_u(R) = LR(R) \text{ for } T \geqslant M \tag{11}$$

c) **Invariance property**:

$$f\left[LR_u(R_0)\right] = f(M, T) \tag{12}$$

Derivation of the $LR_u(R)$ that meets these requirements are introduced in Section 2 and summarized in Section 3.

2. LIKELIHOOD RATIO FOR UNDER-SAMPLED GAUSSIAN SCENARIO

The under-sampled covariance matrix $\hat{R}$ in (4) is rank-deficient and therefore is described by the anti-Wishart distribution [13]. Specifically, the sample matrix

$$\hat{C} = R_0^{-\frac{1}{2}} \sum_{j=1}^{T} x_j x_j^H R_0^{-\frac{1}{2}}; \quad x_j \sim \mathcal{CN}(0, R_0) \tag{13}$$

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

$$K_{T,M} \left(\det \hat{C}_{[T]}\right)^{T-M} e^{-\operatorname{Tr} \hat{C}} \prod_{\mathrm{l=T+1}}^{M} \prod_{\mathrm{p=T+1}}^{M} \delta\left(\frac{\det \hat{C}_{[T]lp}}{\det \hat{C}_{[T]}}\right) \tag{14}$$

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

$$\hat{C} = \left[\begin{array}{c|c} \hat{C}_{[T]} & * \\ \hline * & * \end{array}\right]. \tag{15}$$

Furthermore, for each $l, p > T$ the $(T+1) \times (T+1)$ matrix $\hat{C}_{[T]lp}$ in (14) is obtained by adjoining the l-th row and the p-th column of $\hat{C}$ to $\hat{C}_{[T]}$:

$$\hat{C}_{[T]lm} = \left[\begin{array}{ccc|c} & & & \hat{C}_{1p} \\ & \hat{C}_{[T]} & & \vdots \\ & & & \hat{C}_{Tp} \\ \hline \hat{C}_{l1} & \dots & \hat{C}_{lT} & \hat{C}_{lp} \end{array}\right]. \tag{16}$$

The number of independent delta-functions in (14) 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 (columns) of this matrix that *uniquely specifies* the entire matrix $\hat{C}$. Obviously, one can select another set of covariance matrix $\hat{C}$ entries with the same number of real-valued degrees of freedom, that uniquely describe $\hat{C}$ with rank T.

Strictly speaking, the "under-sampled" likelihood ratio should involve all independent entries within $\hat{C}$ that uniquely specify this matrix, and 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:

$$DOF(\Omega_{\hat{C}}) < DOF(\hat{C}) = 2MT - T^2, \tag{17}$$

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

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

$$DOF(\hat{R}_0) = 1 + 2Mm - m^2 \tag{18}$$

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 $\Omega_{\hat{R}}$, considered for hypothesis testing regarding R_S in (1) exceeds $DOF(R_S)$, then one can expect that consistent (with SNR $\to \infty$) testing is possible.

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 $m < T < M$, we consider a $(2T-1)$ wide band of the matrix $\hat{R}$:

$$\Omega^{\hat{R}} : [\hat{r}_{ij}] \quad |i-j| \leqslant T-1; \quad \hat{R} = [\hat{r}_{ij}] \; i,j = 1, \dots M. \tag{19}$$

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

$$DOF(\hat{R}_{B(T)}) = 2MT - T^2 - (M-T) \tag{20}$$

and is only $(M-T)$ degrees short from DOF($\hat{R}$) in (17). 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, including the rank T completion $\hat{R}$. At the same time, dealing with $\Omega^{\hat{R}}$, we may consider different non-degenerate completions, including the one with the maximal(non-zero) determinant, specified by the Dym-Gohberg band-extension method [14], [15].

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

$$\begin{bmatrix} \hat{r}_{i,i} & \dots & \hat{r}_{i,i+p} \\ \vdots & & \vdots \\ \hat{r}_{i+p,i} & \dots & \hat{r}_{i+p,i+p} \end{bmatrix} > 0, \; for \; i = 1, \dots, M-p \tag{21}$$

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

$$\begin{bmatrix} \hat{\mathcal{Y}}_{q,q} \\ \vdots \\ \hat{\mathcal{Y}}_{L(q),q} \end{bmatrix} = \begin{bmatrix} \hat{r}_{q,q} & \dots & \hat{r}_{q,L(q)} \\ \vdots & & \vdots \\ \hat{r}_{L(q),q} & \dots & \hat{r}_{L(q),L(q)} \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{22}$$

$$\begin{bmatrix} \hat{\mathcal{Z}}_{\Gamma(q),q} \\ \vdots \\ \hat{\mathcal{Z}}_{qq} \end{bmatrix} = \begin{bmatrix} \hat{r}_{\Gamma(q),\Gamma(q)} & \dots & \hat{r}_{\Gamma(q),q} \\ \vdots & & \vdots \\ \hat{r}_{q,\Gamma(q)} & \dots & \hat{r}_{q,q} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \quad (23)$$

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

$$\hat{V}_{ij} \equiv \begin{cases} \hat{\mathcal{Y}}_{ij}\hat{\mathcal{Y}}_{jj}^{-\frac{1}{2}} & \text{for } j \leqslant i \leqslant L(j) \\ 0 & \text{otherwise} \end{cases} \quad (24)$$

$$\hat{U}_{ij} \equiv \begin{cases} \hat{\mathcal{Z}}_{ij}\hat{\mathcal{Z}}_{jj}^{-\frac{1}{2}} & \text{for } \Gamma(j) \leqslant i \leqslant j \\ 0 & \text{otherwise} \end{cases} \quad (25)$$

then the M-variate matrix given by

$$\hat{R}^{(p)} = (\hat{U}^H)^{-1}\hat{U}^{-1} = (\hat{V}^H)^{-1}\hat{V}^{-1} \quad (26)$$

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

$$\begin{cases} \{\hat{R}^{(p)}\}_{ij} = \hat{r}_{ij} & \text{for } |i-j| \leqslant p, \\ \{(\hat{R}^{(p)})^{-1}\}_{ij} = 0 & \text{for } |i-j| > p. \end{cases} \quad (27)$$

In [15],[16] it was proven that amongst all band extensions, extension (27) has the maximal determinant. This extension also has the unique property among possible extensions that, according to (26)

$$\det[\hat{R}^{(p)}]^{-1} = \prod_{q=1}^{M} \hat{\mathcal{Y}}_{qq} = \prod_{q=1}^{M} e_q^T \hat{R}_q^{-1} e_q \quad (28)$$

where $\hat{R}_q$ is the $(L(q)-q+1) \times (L(q)-q+1)$ Hermitian central block matrix in $\hat{R}$, specified in (22). One can see that the Dym-Gohberg band extension method, applied to the rank-deficient sample matrix $\hat{R}$ (4), 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 $\hat{R}$, 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 (21). Let $p \leqslant T-1$. Then for all m in (1) such that $m < p \leqslant T-1$, we have

$$DOF(R_S) < DOF(\hat{R}^{(p)}), \quad (29)$$

while the minimal eigenvalue in all $(p+1)$-variate matrices R_q is equal to the white noise power σ_0^2 in (1). For this reason, we can introduce the following likelihood ratio $\Lambda_0^{(p)}(R)$ for our under-sampled scenario:

$$\Lambda_0^{(p)}(R) = \left[\frac{\det(\hat{R}^{(p)}[R^{(p)}]^{-1})\exp M}{\exp\{\text{Tr }\hat{R}_r R^{-1}\}}\right]^{\frac{1}{M}} \quad (30)$$

Here $R^{(p)}$ is the Dym-Gohberg p-band transformation of the tested positive definite covariance matrix model R, which has the properties

$$R^{(p)} = DG_p(R); \quad \begin{matrix} (R_{ij}^{(p)}) = r_{ij} & \text{for } |i-j| \leqslant p \\ [(R^{(p)})^{-1}]_{ij} = 0 & \text{for } |i-j| > p \end{matrix} \quad (31)$$

$$\hat{R}_r = \lim_{\alpha \to 0} \frac{1}{T}(\alpha I + X_T X_T^H); \;\; X_T = \{x_1, \dots, x_T\}. \quad (32)$$

The loading factor α is sufficiently small, such that

$$DG(\hat{R}) = DG(\hat{R}_r) \quad (33)$$

which means that α should be negligible:

$$\alpha \ll \min_q \lambda_{min}(\hat{R}_q) \quad (34)$$

Note that $\Lambda_0^{(p)}(R)$ is dependent on the determinant of $\hat{R}^{(p)}$ and $R^{(p)}$ which in (28) is given as a function of $\hat{R}_q$ block submatrices. Therefore, we do not need to explicitly construct the Dym-Gohberg extensions for $\Lambda_0^{(p)}(R)$ calculation.

Let us demonstrate that the LR given in (30) meets the requirement $(a)-(c)$ in (10)-(12).

Proper LR Normalisation (requirement a)).

$$\max \Lambda_0^{(p)} < \exp 1; \;\; \Lambda_0^{(p)} = \Lambda_0^{(p)}(\hat{R}_r) = 1 \quad (35)$$

Indeed, for $\hat{R}_r$ that satisfies (32)-(34), we have

$$\lim_{\alpha \to 0} \text{Tr } \hat{R}_r[\hat{R}_r + \beta I]^{-1} = T\left[1 - \beta \text{Tr }\{(X_T^H X_T)^{-1}\}\right] > 0 \quad (36)$$

$$\lim_{\alpha \to 0} \det[\hat{R}_r^{(p)} DG_p[\hat{R}_r + \beta I]^{-1}] = 1; \;\; \beta < \min_q \lambda_{min}(\hat{R}_q) \quad (37)$$

Transition to the Conventional LR (requirement b)).

Obviously, for $p = M-1$, $T \geqslant M$, $DG\{\hat{R}\} = \hat{R}$, while $\text{Tr } \hat{R}_r R = \text{Tr } \hat{R}R$ for α that satisfies (34).

Scenario Independence (requirement c)).

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

$$\Lambda_0^{(p)}(R_0) = \left[\frac{\det(\hat{R}^{(p)}(R_0^{(p)})^{-1})\exp M}{\exp\{\text{Tr }\hat{R}_r(R_0)^{-1}\}}\right]^{\frac{1}{M}} \quad (38)$$

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

Theorem 2 *(see Theorem 2 in [17])*
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:

$$\Lambda_0^{(p)}(R_0) = \exp 1. \left[\prod_{q=1}^{M} \Omega_q \alpha_q\right]^{\frac{1}{M}} \quad (39)$$

where

$$\alpha_q \sim \frac{\alpha_q^{(T-\nu-1)}(1-\alpha_q)^{(\nu-1)}}{B[\nu, T-\nu]} \quad 1 \leqslant \nu \equiv L(q) - q \leqslant p \quad (40)$$

$$\Omega_q = \frac{C_{qq}}{T} \exp\left[-\frac{C_{qq}}{T}\right], \quad C_{qq} \sim \frac{C_{qq}^{T-1}}{\Gamma(T)} \exp(-C_{qq}) \tag{41}$$

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

$$\varepsilon\left\{\left[\Lambda_0^{(p)}(R_0)\right]^l\right\} = \frac{T^{TM}\exp(l)}{\left[T + \frac{l}{M}\right]^{(TM+l)}} \prod_{q=1}^{M} \frac{\Gamma\left(T + \frac{l}{M} - \nu(q)\right)}{\Gamma(T - \nu(q))} \tag{42}$$

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

$$\lim \mathrm{Tr}\ \hat{R}_r \hat{R}_r^{-1} = M \tag{43}$$

but it needs to remain small enough for

$$\det \hat{R}^{(p)} \overset{-1}{\det} DG_p(\alpha I_0 + \hat{R}) \to 1. \tag{44}$$

At $\alpha = 0$

$$\mathrm{Tr}\ \hat{R}[I - X_T(X_T^H X_T)^{-1} X_T^H] = 0, \tag{45}$$

so the term $\exp(M)$ in (38) is not required.

The introduced likelihood ratio (30) also allows us to specify the projection matrix

$$\lim_{\alpha\to 0} \hat{R}_r^{-1} = \left[I - X_T(X_T^H X_T)^{-1} X_T^H\right] = [I - P_{X_T}] \tag{46}$$

as an under-sampled maximum likelihood (USML) inverse covariance matrix estimate, "complementary" to the rank-deficient (USML) covariance matrix estimate

$$\lim_{\alpha\to 0} \alpha^{-1} \hat{R}_r = X_T X_T^H. \tag{47}$$

3. 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.

A. REFERENCES

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

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

[3] E.K.L. Hung and R.M. Turner, "Fast beamforming algorithm for large arrays," *IEEE Trans. AES*, vol. 19, no. 4, pp. 598–607, 1983.

[4] C. Gierull, "Performance analysis of fast projections of the Hung-Turner type for adaptive beamforming," in *Signal Processing*, 1996, vol. 50, pp. 17–28.

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

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

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

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

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

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

[11] 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. SP*, vol. 52, no. 6, pp. 1524–1536, 2004.

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

[13] R. Janik and M. Nowak, "Wischart and anti-Wishart random matrices," *Journal of Physics A: Mathematical and General*, vol. 36, no. 4, pp. 3629–3637, 2003.

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

[15] H. Woerdeman, *Matrix and operator extensions*, Ph.D. thesis, Vrije Universiteit, Amsterdam, 1989.

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

[17] Y.I. Abramovich, N.K. Spencer, and M.D.E. Turley, "Order estimation and discimination between stationary and time-varying autoregressive (TVAR) models," *submitted to IEEE Trans. SP*, 2005.