EXPECTED LIKELIHOOD FOR TEMPORALLY CORRELATED (OVER-SAMPLED) TRAINING DATA: EXPERIMENT RESULTS

Scott G. McMillan, Yuri I. Abramovich†, and Ben A. Johnson***

* University of Adelaide, SA 5005, and Lockheed Martin Australia, SA 5085, Australia
email: scott.mcmillan@student.adelaide.edu.au
†ISR Division, DSTO, PO Box 1500, Edinburgh SA 5111, Australia
email: yuri.abramovich@dsto.defence.gov.au
** Lockheed Martin Australia, 45 Third Avenue, Mawson Lakes SA 5085, Australia
email: ben.a.johnson@ieee.org

ABSTRACT

In this study, we analyse sensitivity of the likelihood ratio (LR) p.d.f with respect to the residual temporal correlation of antenna training samples, caused by non-rectangular bandpass filters and over-sampling. This temporal correlation causes deviation of the actual LR p.d.f. from the theoretical one derived for independent (Gaussian) training samples. This in turn affects performance of "expected likelihood" detection-estimation techniques. This paper outlines methodologies for correcting for this temporal correlation, allowing for the application of expected likelihood, and validates the methods using multichannel receiver data collected by the Australian OTHR JORN facility.

1. INTRODUCTION

In [1, 2, 3] we suggested an "expected likelihood" (EL) approach for detection-estimation of a number of sources impinging upon an M-element antenna array. EL operates by directly examining, for a set of estimated parameters that uniquely specify the M-variate spatial covariance matrix, the likelihood ratio (LR) of the reconstructed covariance matrix relative to the input data. EL treats the estimated parameters as appropriate if the LR of the reconstructed model is within the range of LR values expected for the true (actual) covariance matrix and operates in the practical circumstances where the true covariance matrix is not known *a priori*, but the p.d.f. of the LR values for that unknown true solution can be determined nonetheless. More specifically, for independent identically distributed (i.i.d) training samples with (complex) Gaussian p.d.f., this LR distribution can be shown to be exhaustively described by just the antenna dimension M and the number of i.i.d. training samples, N_{iid}, both of which are known *a priori* [3]. This p.d.f may be pre-calculated analytically (through a fairly complex closed-form solution given in [1]), or more conveniently via a single upfront Monte-Carlo simulation. The p.d.f of this LR is highly dependent on the extent to which the input data is i.i.d., and unfortunately, in practice, antenna output signals are never ideally independent.

For real-world systems, to make antenna outputs be fully temporally i.i.d., the bandpass filters (at least) of the multichannel receivers must be ideally rectangular with sample rate accurately equal to the Nyquist rate. This is often an impractical approach, even in theory. Radar systems, for example, usually are subject to stringent out-of-band emission control which mandates a taper on the transmitted waveform, and this taper is often reproduced in receive bandpass filters during the waveform correlation process. Moreover, such systems usually operate with some over-sampling in order to avoid spectral aliasing. In this configuration, even internal noise samples are never ideally independent. The impact of training data temporal correlation on the structure and performance of different detection and direction of arrival (DOA) estimation techniques, has been considered in a number of papers (see, for example, [5]). Note that in most such studies, internal noise samples are still treated as i.i.d, while temporal correlation of source signals is considered.

The focus of this study is rather different. Within the EL approach, a decision regarding any estimated parameters $\hat{\Omega}_{\hat{m}}$ with uniquely reconstructed covariance matrix $R(\hat{\Omega}_{\hat{m}})$ is taken based on comparison of the actual likelihood ratio $LR[R(\hat{\Omega}_{\hat{m}}]$ with the theoretical p.d.f. pre-calculated for the true covariance matrix $R_0, LR[R_0]$. Therefore, if the residual temporal correlation of the training data significantly modifies the p.d.f. of $LR[R_0]$ relative to the one calculated assuming an i.i.d-model, then the performance of the expected likelihood detection-estimation approach may be degraded. In this study, we analyse these modifications to the p.d.f. for the $LR[R_0]$, caused by non-rectangular shape of bandpass filters and oversampling. In most DOA estimation techniques, the accurate description of the additive (internal) white noise is crucial, and in some applications with broadband signals impinging upon an antenna array, the power spectrum of all signals as well as the internal noise power spectrum are specified by bandpass filters and over-sampling rate. While it is clear that for $N \to \infty$, the equivalent number of i.i.d. training samples may be specified as $N_{iid} = T\Delta F$, where T is the observation interval and ΔF is the actual bandwidth of the passband filter, this is of limited assistance in the pre-asymptotic domain, where EL is specifically designed to operate [3]. The sensitivity of different equivalent models that account for sample temporal correlation needs to be carefully evaluated with finite N. In this paper, we introduce our results of this study conducted with the aid of real internal noise data collected at the output of the multiple antenna receivers associated with the Australian JORN OTH Radar in Laverton, WA [4].

2. PROBLEM FORMULATION

The traditional theoretical description of the detection-estimation problem addressed by stochastic (unconditional maximum likelihood (SML), is that a number of N_{iid} independent identically distributed complex Gaussian M-variate training samples $x_j, j = 1, \ldots, N_{iid}$ are observed at the output of an M-element antenna array. In that case, the true (actual) spatial covariance matrix is R_0 :

$$\varepsilon \left\{ x_p x_q^H \right\} = \begin{cases} R_0, & \text{for } p = q \\ 0, & \text{for } p \neq q \end{cases} \tag{1}$$

and for the considered model is uniquely specified by the set of the true parameters Ω_m. Within the EL approach, the set of parameters $\hat{\Omega}_{\hat{m}}$ is treated as the solution of the detection-estimation problem, if

$$\hat{m} = \min_{\mu} \left\{ \mu \forall LR[R(\hat{\Omega}_\mu)] \geq \gamma_0 \right\} \tag{2}$$

where

$$\gamma_0 = \arg_\gamma \left\{ \int_0^\gamma w(x) dx = P_{FA} \right\}, \tag{3}$$

where $w(x)$ is the p.d.f for $LR[R_0]$ and P_{FA} is the given probability to miss a "proper" solution. The normalised likelihood function is used for the $N_{iid} \geq M$ as the likelihood ratio in (2) and (3)

$$LR[R(\hat{\Omega}_\mu] = \left[\frac{\det[\frac{1}{N_{iid}} R^{-1}(\hat{\Omega}_\mu)\hat{R}] \exp M}{\exp \mathrm{Tr}[\frac{1}{N_{iid}} R^{-1}(\hat{\Omega}_\mu)\hat{R}]} \right] \tag{4}$$

$$\text{where } \hat{R} = \sum_{j=1}^{N_{iid}} x_j x_j^H. \tag{5}$$

When the covariance matrix model $R(\hat{\Omega}_\mu)$ may have an arbitrary scaling factor, the "sphericity test" is used as the (condensed) likelihood function [1, 2, 3]:

$$LR_{sp}[R(\hat{\Omega}_\mu)] = \frac{\det[R^{-1}(\hat{\Omega}_\mu)\hat{R}]}{[\frac{1}{M}\mathrm{Tr} R^{-1}(\hat{\Omega}_\mu)\hat{R}]^M} \tag{6}$$

For the actual covariance matrix $R_0 = R(\Omega_\mu)$ in (4) and (6), the p.d.f's for $LR[R_0]$ and $LR_{sp}[R_0)]$ do not depend on the actual covariance matrix, since

$$LR[R_0] = \frac{\det[\frac{1}{N_{iid}}\hat{C}] \exp M}{\exp\ \mathrm{Tr}[\frac{1}{N_{iid}}\hat{C}]}, \tag{7}$$

where for i.i.d. training samples

$$\hat{C} = R_0^{-\frac{1}{2}} \hat{R} R_0^{-\frac{1}{2}} \sim CW(N_{iid}, M, I_M) \tag{8}$$

with $CW(N_{iid}, M, I_M)$ indicating the complex Wishart distribution described by N_{iid} and M.

For the model that reflects a practical scheme, the training samples in the sample covariance matrix

$$\hat{R} = \sum_{j=1}^{N} x(j) x^H(j) \tag{9}$$

are not independent, so the MN-variate covariance matrix of a single (stacked) MN-variate vector of all observed snapshots X_{MN}

$$X_{MN}^T = [x_1^T, \ldots, x_N^T] \in C^{1 \times MN} \tag{10}$$

is defined as the Kronecker product (indicated by $\otimes$) of the $M \times M$ variate matrix R_0 and an $N \times N$-variate Toeplitz matrix T_N which describes the temporal correlation. For the considered stationary Gaussian (noise), the Toeplitz matrix T_N

$$T_N = \mathrm{Toep}(t_0, t_1, \ldots, t_{N-1}), t_0 = 1 \tag{11}$$

is uniquely specified by the power spectrum $f(\omega)$:

$$t_j = \frac{1}{2\pi} \int_0^{2\pi} f(\omega) e^{ij\omega} d\omega \quad (t_0 \equiv \frac{1}{2\pi} \int_0^{2\pi} f(\omega) d\omega = 1) \tag{12}$$

$$0 \leq \omega \equiv 2\pi f/fs \leq 2\pi, \tag{13}$$

where fs is the sampling rate (frequency).

In the over-sampling case, the actual spectrum $f(\omega)$ vanishes outside of the actual filter bandwidth

$$0 < \omega_{min} < \omega_{max} < 2\pi, \tag{14}$$

In this case, some number of eigenvalues in T_N tend to zero, while the rest of them may not be strictly equal, especially for a small N:

$$\lambda_1 \geq \lambda_2 \cdots \geq \lambda_n > \lambda_{n+1} \simeq \cdots \simeq \lambda_N = 0, \tag{15}$$

and therefore, the p.d.f. of the likelihood ratio

$$LR_{sp}(\hat{C}(T_N)) = \frac{\det \hat{C}(T_N)}{[\frac{1}{M}\mathrm{Tr}\hat{C}(T_N)]^M} \tag{16}$$

with $\hat{C}(T_N) = R_0^{-\frac{1}{2}} \hat{R} R_0^{-\frac{1}{2}}$, should be different from $LR_{sp}[\hat{C}_{iid}]$ for the same number of training samples ($N_{iid} = N$). This difference must be investigated and accounted for to retain fidelity of the EL approach.

3. METHODS TO ACCOUNT FOR RESIDUAL TEMPORAL CORRELATION ($T_N \neq I_N$) IN EL METHODOLOGY

$$\text{Let} \qquad Y_N = R_0^{-\frac{1}{2}} X_N \quad ; \quad X_N = [x_1, \ldots, x_N]. \tag{17}$$

Since

$$Y_N Y_N^H = Y_N U_N U_N^H Y_N^H, \tag{18}$$

where $U_N U_N^H = U_N^H U_N$ is the $N \times N$ identity matrix I_N, the p.d.f. for $LR_{sp}(\hat{C}(T_N))$ may be calculated as

$$LR_{sp}(\hat{C}(T_N)) = \frac{\det Z_N \Lambda_N Z_N^M}{[\frac{1}{M}\mathrm{Tr} Z_N \Lambda_N Z_N^H]^M} \tag{19}$$

$$\text{where} \quad Z_N \sim C\mathcal{N}(0, I_M \otimes I_N)$$

$$C\mathcal{N}(0, R_M \otimes R_N) = \frac{1}{(2\pi)^{MN} [\det R_M]^N [\det R_N]^M} \tag{20}$$

$$\exp[-\mathrm{Tr} R_M^{-1} Z_N R_N^{-1} Z_N^H]; \quad R_M, R_N > 0.$$

For the known *a priori* shape of the bandpass filter, covariance matrix T_N and its eigenspectrum may be pre-calculated and used in (19) for pre-calculation of the actual (non-i.i.d) p.d.f of the expected likelihood. Another option is the eigendecomposition-based resampling, with transformation

$$Z_n = Z_N U_n \Lambda_n^{-\frac{1}{2}}, \tag{21}$$

where $U_n \in C^{N\times n}$ is the matrix of n eigenvectors that correspond to the essentially non-zero eigenvalues. This accurate approach now needs to be compared with alternative techniques that are also often applied under these pre-asymptotic circumstances, but are justified by asymptotic considerations.

3.1 Training data "decimation"

The simplest and most obvious approach is to decimate the over-sampled training data, with every k-th sample (of N available) used for covariance estimation (9). While this approach means obvious reduction of the actual training samples used, it still does not guarantee strict independence, since the rate of the"decimated" samples is never accurately equal to the Nyquist rate and some residual correlation remains. While shown to be partially successful in EL applications [6], this brute-force technique is wasteful of training samples and we therefore consider it only for comparison purposes.

3.2 DFT-based resampling

It is well known that the Fourier coefficients of a stationary finite-duration continuous signal are asymptotically uncorrelated in the sense that the cross-correlation between any two Fourier coefficients approaches zero as the time duration T grows infinite [7]. Similar behaviour is observed by discrete-time signals.

Analytically, this statement means that asymptotically, as $N\to\infty$, the Toeplitz matrix T_N maybe approximated by the circulant matrix C_N, so that $\lim_{N\to 0}|T_N - C_N| \to 0$. For a finite N, the bounds for the "weak" norm

$$\Delta^2 = ||T_N - C_N||^2 = \frac{1}{N}\sum_{p=1}^{N}\sum_{q=1}^{N}|(T_N)_{pq} - (C_N)_{pq}|^2, \tag{22}$$

have been derived in [7]. Specifically, for

$$|t(k)| \le M/k; \quad M < \infty\,\forall k; \quad \Delta^2 \le \frac{2M^2}{N}[1+\log(N-1)], \tag{23}$$

while for

$$|t(k)| \le ke^{-\alpha k} \quad 0<\alpha<\infty, \quad \forall k, \tag{24}$$

$$\Delta^2 \le \frac{2k^2}{N}\frac{e^{-\alpha}}{(1-e^{-\alpha})^2}\left[1 - \frac{1}{N}\frac{1+e^{-\alpha}}{1-e^{-\alpha}}\right]. \tag{25}$$

One can see, that as $N\to\infty$, the norm Δ falls like $\sqrt{\frac{1}{N}\log N}$ in (23) or like $\sqrt{\frac{1}{N}}$ in (25), which means that for modest N the errors of this approach should be taken into account. Specifically, we have to investigate the cumulative effect of these errors on the shape of the expected likelihood p.d.f.

Consider DFT transformation of $M\times N$-variate training sample matrix Y_N into an $M\times N$-variate matrix Y_n:

$$Y_n = Y_N F_n^H W_n^{-\frac{1}{2}} \text{ where } F_n = \frac{1}{\sqrt{N}}[e^{i\frac{2\pi}{N}pq}] \in C^{n\times N}; \tag{26}$$

$$p = p_{min},\ldots,p_{max}; \quad p_{max} - p_{min} + 1 = n < N; q = 1,\ldots,N.$$

$$[W_n]_{pq} = \begin{cases} [F_n^H T_n F_n]_{pp} = \sigma_p^2 & \text{for } p = q \\ 0 & \text{for } p \ne q \end{cases}. \tag{27}$$

Here the n DFT frequencies $p_{min},\ldots,p_{max}$ are selected within the essentially non-zero part of the power-spectrum $f(\omega)$. Therefore, the equivalent stochastic representation for the $LR_{sp}[Y_nY_n^H]$

$$LR_{sp}[Y_nY_n^H] = \frac{\det Y_nY_n^H}{[\frac{1}{M}\mathrm{Tr}Y_nY_n^H]^M}, \tag{28}$$

is

$$LR_{sp}[Y_nY_n^H] \sim \frac{\det Z_n\Lambda_nZ_n^H}{[\frac{1}{M}\mathrm{Tr}Z_n\Lambda_nZ_n^H]^M}, \tag{29}$$

where Λ_n is a diagonal matrix of eigenvalues of the Hermitian matrix H_n $(= W_n^{-\frac{1}{2}}F_n^H T_N F_n W_n^{-\frac{1}{2}} \ne I_n)$, and $Z_n \sim \mathscr{CN}(0, I_m\otimes I_n)$.

3.3 Rectangular spectrum approximation

If the spectrum $f(\omega)$ is approximated as

$$f(\omega) = \begin{cases} c & \omega_{min} < \omega \le \omega_{max} \\ 0 & \text{elsewhere in } [0,2\pi] \end{cases},$$

then

$$T_N = \left[\frac{\sin(p-q)(\omega_{max}-\omega_{min})/2}{(p-q)(\omega_{max}-\omega_{min})/2}\right]. \tag{30}$$

The eigenvectors of such a covariance matrix are known to be the discrete prolate spheroidal (Slepian) functions, and for $N\to\infty$, the $n = \frac{N(f_{max}-f_{min})}{fs} < N$ eigenvalues of this matrix are approximately equal, with the rest of them rapidly reaching zero value. Specifically these considerations justify the above mentioned equivalent model

$$LR[Z_nZ_n^H] = \frac{\det Z_nZ_n^H}{[\frac{1}{M}\mathrm{Tr}Z_nZ_n^H]^M}, \quad Z_n \sim C\mathscr{N}(0, I_m\otimes I_n), \tag{31}$$

with $n = TF$ equivalent iid training samples. The ideal rectangular shape assumption, and a medium sample support N volume are the obvious reasons of concern for this technique.

Now, all the mentioned above approximate techniques need to be compared with the accurate equivalent statistical model, when applied to the real-world data.

4. EXPERIMENTAL RESULTS

To verify the relative performance of these techniques, noise data collected at the digitised output of each antenna element (with the receivers terminated into a matched load) have been collected at the Australian OTHR JORN at Laverton [4]. In this collection, the sampling rate fs exceeds the nominal bandwidth $(f_{max} - f_{min})$ by a factor of 1.25.

In order to explore small to medium sample volumes, we considered a $M = 20$ element portion of the antenna array using 400 available channels for statistical averaging. The data have been collected over a large interval with the total number of recorded samples $N_{total} = 250000$ sufficient for reliable statistical averaging and p.d.f / c.d.f. determination.

Apart from the non-rectangular shape of the bandpass filter and over-sampling, we also had to consider possible variations of noise figures across the receivers (i.e. the potential for an "un-balanced" array). Using the entire sample volume for noise power estimation across the receivers, we found variations in the estimate powers p_j. In addition to the training samples temporal correlation phenomenon, we then analysed the impact of this imperfect balance on the expected likelihood p.d.f.

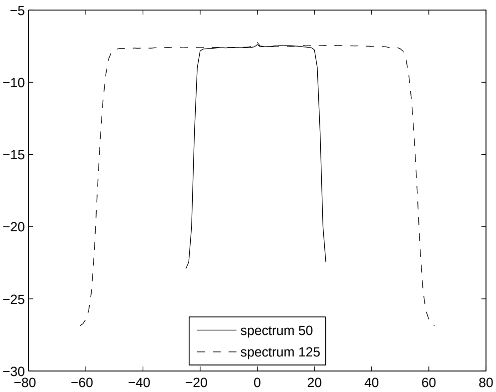


Figure 1: Noise samples spectra (in DFT units)

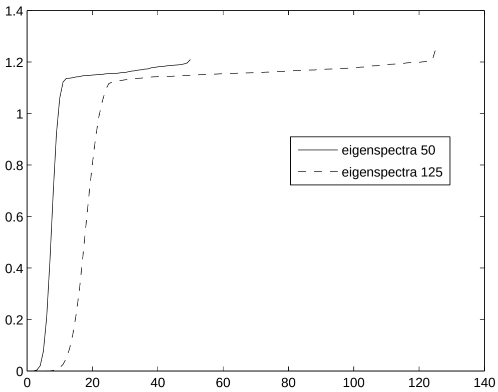


Figure 2: Covariance matrices eigenspectra

To do so, we considered a sphericity test (allowing for an arbitrary scaling of the sample covariance matrix $\hat{R}$).

$$LR_{sp}(\hat{C}) = \frac{[\det \hat{C}]^{\frac{1}{M}}}{[\frac{1}{M}\mathrm{Tr}\hat{C}]}, \quad \hat{C} = R_{mod}^{-\frac{1}{2}} \hat{R} R_{mod}^{-\frac{1}{2}}, \tag{32}$$

using either $R_{mod} = I_M$ ("un-balanced" antenna model) or $R_{mod} = D_M = \mathrm{diag}(p_1, \ldots, p_M)$ ("balanced antenna model"). For the $M = 20$-element antenna arrays, we analysed two training samples volumes, $N_1 = 50$ and $N_2 = 125$, which with respect to the nominal over-sampling rate are (roughly) equivalent to $N_{iid} = 40 (N_{iid}/M = 2)$ and $N_{iid} = 100 (N_{iid}/M = 5)$. These sample volumes span the range of (i.i.d.) training samples per antenna element often used in adaptive antenna array literature: $N_{iid}/M = 2$ to 5. The entire sample volume and averaging across all available 400 channels were used for sufficiently accurate calculation of the $N = 125$-variate Toeplitz matrix and the shape of the bandpass filter transfer function. At Fig. 1 we introduce the FFT spectrum estimate $(F_N^H T_N F_N)_{jj}, \quad j = 1, \ldots, 125$ and eigenspectrum of the covariance matrix T_N (Fig 2). One can see that indeed, in ~ 100 frequency bins the noise power spectrum is approximately equal, and the $n = 100$ largest eigenvalues of the matrix T_N are approximately equal. For $N = 50$ this "rectangular" approximation is slightly less accurate (Figures 1 and 2). We considered both original ("un-balanced") training data X_N, and properly balanced data $Y_N = D_M^{-\frac{1}{2}} X_N$, where D_M is the diagonal matrix with measured noise powers across the receiver outputs (32). Results of our analysis are illustrated by Figures 3 and 4.

At Fig. 3 we provide the averaged over all $m = 20$ available $M = 20$-element antenna arrays ($m = 400/20$) and sample volume (400 Monte-Carlo runs) sample c.d.f.'s for

- original (un-balanced) data X_N,
- equalized (balanced) data Y_N,
- decimated data with the decimation rate of 2 and 5,
- theoretical i.i.d. model (10^6 trials) for $N = 50$ and 125,
- theoretical non-i.i.d. model (19) with eigenvalues Λ_N from Fig. 2.

Comparison of these experimental c.d.f.'s with theoretical modelling leads to several important observations. First of all, we notice that sample c.d.f.'s for original (unbalanced) and balanced data are practically indistinguishable, both for $N = 50$ and $N = 125$. Although not illustrated here, this same correspondence was observed for $M = 200$ element antenna with $N = 1250$. Therefore, variations of the noise figures over 400 JORN antenna receivers are found to be sufficiently small to impose no noticeable impact on the LR c.d.f, and therefore could be ignored. One can also see that while 2:1 decimation "shifts" the LR's c.d.f. closer to the theoretical LR model with $N = 50$, it still does not provide a perfect match that allows the decimated data to be treated as strictly i.i.d. On the contrary, the theoretical non-i.i.d. model (19) that adopts the actual eigenvalues Λ_N (illustrated by Fig. 2) provides the perfect match with the actual data.

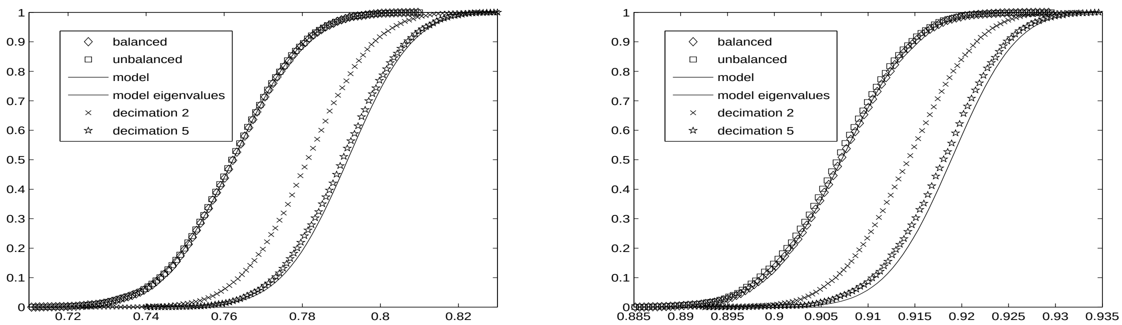
At Fig. 4, we introduce the similarly averaged "experimental" c.d.f's for

- eigendecomposition-based resampling ($n = 40, 100$)
- FFT-based resampling ($n = 40, 100$)
- theoretical i.i.d. model (10^6 trials) for $N_{iid} = 40, 100$

This data shows that for a small sample support ($N = 50$), eigendecomposition-based re-sampling provides a good match between the experimental and theoretical i.i.d.-based c.d.f.'s, outperforming FFT-based re-sampling. For a larger sample support ($N = 125$) the difference between FFT-based and eigendecomposition-based re-sampling is smaller. Note the x scale on Fig.4 is tighter in order to better illustrate the match between eigen whitened samples and the model.

5. CONCLUSIONS AND RECOMMENDATIONS

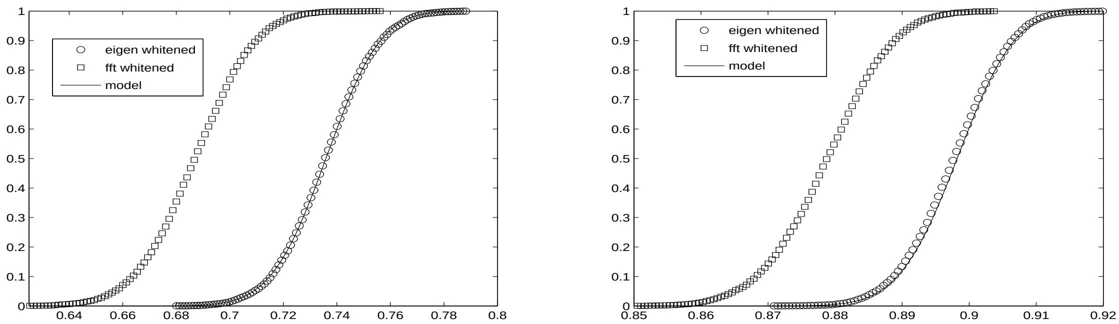
In this paper, we analysed the impact of residual temporal correlation of internal noise data caused by non-rectangular



(a) $N = 50$

(b) $N = 125$

Figure 3: Comparison of LR c.d.f's for decimated and simulated data



(a) $N = 50$

(b) $N = 125$

Figure 4: Comparison of LR c.d.f's for resampled and simulated data

bandpass filters and over-sampling on the accuracy of matching between the actual likelihood ratio c.d.f. and the theoretical (assumed i.i.d.) expected likelihood ratio c.d.f. We demonstrated that in practical systems such as the Australian OTHR "JORN", the i.i.d. noise model is not sufficiently accurate for the "raw" noise data to be directly used for expected likelihood c.d.f. calculations. But accurate theoretical models and/or re-sampling that takes into account the actual shape of the noise power spectrum does allow for the EL matching approach to be applied in practical systems with theoretically predicted performance. The wide-spread FFT-based re-sampling technique can be used for relatively large sample support, while its applicability in particular EL applications with short data records [8] must be carefully verified.

REFERENCES

[1] Y. I. Abramovich, N. K. Spencer, A. Y. Gorokhov, "Bounds on maximum likelihood ratio - Part I. Application to antenna array detection-estimation with perfect waveform coherence", *IEEE Trans. Signal Process.*, vol 52, no. 6 pp 1524-1536, Jun. 2004

[2] Y. I. Abramovich, N. K. Spencer, A. Y. Gorokhov, "GLRT-based threshold detection-estimation performance improvement and application to uniform circular antenna arrays", *IEEE Trans. Signal Process.*, vol 55. no 1. pp 20-31, Jan 2007

[3] B. A. Johnson, Y. I. Abramovich, "DOA estimator performance assessment in the pre-asymptotic domain using the likelihood principle", *Elsevier Journal of Signal Processing*, vol 90 no. 5, pp 1392-1401, May 2010

[4] A. Cameron "The Jindalee Operational Radar Network: Its architecture and surveillance capability" in *Proc. International Radar Conference*, Alexandria, VA: IEEE, May 8-11, 1995, pp 692-697

[5] F. Haddadi, M. H. Nayeti. M. R. Aref, "Direction-of-Arrival estimation for temporally correlated narrowband signals", *IEEE Trans. Signal Process.*, vol 57, no. 2, pp 600-609, Feb 2009

[6] B. A. Johnson, Y. I. Abramovich, "Experimental verification of environmental models for adaptive detection and estimation in HF skywave radar", in *Proc. Radar Conference*, Verona, NY: 24-27 Apr, 2006, pp 782-787

[7] J. Pearl, "On coding and filtering stationary signals by discrete Fourier transforms", *IEEE Trans. On Information Theory*, 17-19, pp 229-232, March 1973

[8] Y. I. Abramovich, B. A. Johnson, "GLRT-based detection estimation for under-sampled training conditions", *IEEE Trans. Signal Process.*, vol 56 no. 8, pp 3100-3612, Aug 2008