MODEL ORDER ESTIMATION FOR ADAPTIVE RADAR CLUTTER CANCELLATION

Muralidhar Rangaswamy[1], *Steven Kay*[2], *Cuichun Xu*[2], *and Freeman C. Lin*[3]

[1]Air Force Research Laboratory, Sensors Directorate, 80 Scott Drive, Hanscom AFB, MA, 01731
[2]Department of Electrical and Computer Engineering, University of Rhode Island,
Kelly Hall, 4 East Alumni Ave. Kingston, RI 02881
[3]ARCON Corporation, 260 Bear Hill Road, Waltham, MA, 02451

ABSTRACT

Adaptive waveform design for radar clutter cancellation requires knowledge of the rank of the clutter subspace. In this paper, we compare the computed clutter subspace rank, r, using three methods: (i) the exponentially embedded family (EEF) estimator, (ii) Rissanen's minimum description length (MDL) estimator, and (iii) the statistical ranking and selection method (CWA).

1. INTRODUCTION

Successful application of the space-time adaptive processing (STAP) for airborne radar systems to detect moving targets depends on the effective use of the available degrees of freedom to ameliorate the problems of training data support and computational cost. These considerations become particularly important in heterogeneous clutter scenarios. One solution to determine the required number of degrees of freedom is to correctly estimate the number of target signals. Under ideal conditions when the power spectrum density (PSD) of the clutter has a sharp locus (clutter ridge) in the angle-Doppler space and PSDs of all targets are away from the clutter ridge, it is simple to obtain the correct number of targets. In practice, the clutter PSD has a wide spread in the angle-Doppler space and some of target PSDs remain buried in the clutter spectrum. This often prohibits the correct estimation of the number of targets. Therefore, another approach is to adaptively design a proper waveform in order to null the clutter. Consequently, adaptive waveform design for radar clutter cancellation requires *a priori* knowledge of rank of the clutter subspace. For airborne radar applications, the clutter rank is a function of two fundamental factors, i.e., the spatio-temporal non-stationarity of the clutter and system parameters such as platform speed, inter-element array spacing, and pulse repetition interval. The KASSPER data sets [1, 2], which are simulated data for airborne linear phased radar system with precisely the above-mentioned system parameters, are used for the estimation of clutter rank.

Designing a model order estimator to compute the clutter rank is essential for waveform design to mitigate clutter and to enhance target detection performance of adaptive detectors, e.g., the normalized low-rank adaptive filter [3]. Since the problem is one of composite hypothesis testing, for which no optimal solution exists, there is no consensus on its solution. One common approach employs a Bayesian philosophy which assumes a noninformative prior in an effort to "integrate out" the unknown model parameters. Then, the effect of the prior is ignored. Along these lines the minimum description length (MDL) has been proposed based on coding arguments [4]. Another approach, based on differential geometric statistical models, is the exponentially embedded family (EEF) [5]. In this paper, we compare EEF and MDL estimators for clutter rank estimation.

The CWA method [6] is based on ranking and a variation of the subset selection approach [7] to develop a screening procedure to select secondary data for radar signal processing. With some effort, the CWA method can be modified for the estimation of clutter rank provided that the clutter covariance matrix is given. In other words, one can reformulate the CWA procedure by replacing δ^* and c with the clutter-to-noise ratio, CNR (in the case of signals embedded in white noise, it will be signal-to-noise ratio, SNR), where δ^* is a preassigned real number to differentiate between good and bad eigenvalues and c is the real number chosen to satisfy the condition that the probability of a correct selection is optimal. After this replacement, the computed clutter rank during each Monte Carlo simulation step is equal to the total number of eigenvectors V_i having the ratio of eigenvalues $\lambda_i/\lambda_p > CNR\,(i = 1, 2, \cdots, p-1)$ where p is the total number of eigenvalues for the clutter covariance matrix and λ_p is the smallest eigenvalue .

In Section 2, we describe radar parameters used to generate simulated L-band and X-band KASSPER data sets and present eigenspectra computed for both data sets. Then, the EEF formulation for covariance rank estimation is given in Section 3. Computed clutter ranks using EEF, MDL, and CWA methods for both data sets are illustrated in the Sections 4 and 5, respectively. Section 6 presents conclusions.

Parameter	Value
Carrier frequency	1240 MHz
Bandwidth	10 MHz
Number of antenna elements (J)	11
Number of pulses (N)	32
Pulse repetition frequency (T^{-1})	1984 Hz
horizontal element spacing (d_h)	0.1092 m
vertical element spacing (d_v)	0.1407 m
1000 range bins	from 35 km to 50 km
Clutter-to-noise ratio (CNR)	40 dB
Platform speed (v_p)	100 m/s

Table 1. Parameters for L-band KASSPER data (set 1)

Parameter	Value
Carrier frequency	10000 MHz
Bandwidth	10 MHz
Number of antenna elements	12
Number of pulses	38
Pulse repetition frequency	2081, 1800, and 1518 Hz
horizontal element spacing	0.015 m
vertical element spacing	0.015 m
1667 range bins	from 30 km to 55 km
Clutter-to-noise ratio	23 dB
Platform speed	150 m/s

Table 2. Parameters for X-band KASSPER data (set 2)

2. EIGENSPECTRA FOR L-BAND AND X-BAND KASSPER DATACUBES

Typically for space-time adaptive processing (STAP), the side-looking linear phased array radar has J array elements and each element transmits N pulses. Within a coherent processing interval (CPI), each receiver processes N pulses. For each pulse repetition interval (PRI), R data samples are also collected and processed to cover all range bins of the illuminated terrain. Thus, the $J \times N \times R$ data is referred as the CPI datacube. The spatio-temporal product for STAP is JN. In this paper, simulations for the estimation of the clutter rank are conducted by using simulated CPI datacubes, namely, L-band KASSPER datacube (referred to set 1) [1] and X-band KASSPER datacubes (referred to set 2) [2]. Parameters for both set 1 and set 2 are shown in Tables 1 and 2. There is one CPI datacube for set 1 and 90 CPI datacubes for set 2. To compare the performance of the clutter rank estimation for three algorithms, namely, EEF, MDL, and CWA, it is necessary to inspect the eigenspectrum for each datacube, which contains the clutter-only data. For set 1, we simulated clutter-only data from the true clutter covariance matrix at range bin 500. We then computed the eigenspectrum for set 1 and plotted it in Figure 1(a), where a 60 dB roll-off in eigenvalue occurs after approximately 50 indices. For set 2, since there are clutter-only data and target-clutter data in each datacube, we used the clutter-only data to compute the eigenspectrum. In Figure 1(b), the eigenspectrum for the CPI 10 datacube of set 2 shows that there is probably a 40 dB roll-off in eigenvalue after approximately 175 eigenvalue indexes, while for other CPIs at 1, 22, 45, 65, 80, and 90, they all have a more gradual roll-off in eigenvalues. The similar characteristic for those eigenspectra indicates that perhaps the CPI 10 datacube is an anomaly. Under ideal conditions for the application of airborne linear phased array radar, the Brennan's rule [8] yields the clutter rank, $r = J + \beta(N-1)$, where $\beta = 2v_pT/d_h$ is the slope of the clutter ridge, v_p the platform velocity, T the pulse repetition interval, and d_h the horizontal element spacing. Since $\beta \sim 1$ for set 1, this leads to $r \approx 42$ according to the Brennan's rule. On the other hand, since β is around 10 for set 2, it yields $r \approx 382$.

3. EEF FOR COVARIANCE RANK ESTIMATION

Suppose q complex spatial sinusoids are embedded in complex white Gaussian noise with a single snapshot data vector given by $\mathbf{x} = [x[0], x[1], \cdots, x[p-1]]^T$, where $p > q$. The data can be described by [9]

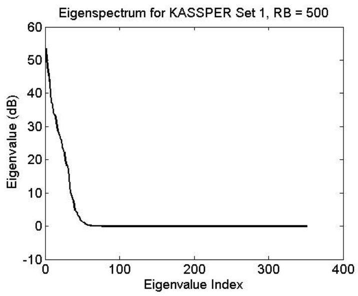
$$\mathbf{x} = \sum_{i=1}^{q} \mathbf{A}(f_i)\, s_i + \mathbf{u} \tag{1}$$

where $\mathbf{A}(f_i) = \left[1, e^{-j2\pi f_i}, \cdots, e^{-j2\pi f_i (p-1)}\right]^T$ is a $p \times 1$ complex steering vector, s_i is a scalar complex Gaussian amplitude, and $\mathbf{u}$ is a $p \times 1$ complex white Gaussian noise vector. It is assumed that $\mathbf{u}$ is independent of the signal amplitudes and its covariance matrix is $\sigma^2\mathbf{I}$. It is also assumed that the signals are incoherent so that $\mathbf{S} = E\left[\mathbf{s}\mathbf{s}^H\right]$ is nonsingular, where $\mathbf{s} = \left[s_1, \cdots, s_q\right]^T$. This is the model used in [9]. Since Eq. (1) can be rewritten in matrix form as $\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{u}$ where $\mathbf{A} = [\mathbf{A}(f_1), \mathbf{A}(f_2), \cdots, \mathbf{A}(f_q)]$, the autocorrelation matrix of $\mathbf{x}$ is $\mathbf{R_x} = \mathbf{A}\mathbf{S}\mathbf{A}^H + \sigma_u^2\mathbf{I}$. The problem is to determine the number of signals q, or equivalently to estimate the rank of $\mathbf{R_x} - \sigma_u^2\mathbf{I}$, given N independent identically distributed (iid) data vectors or snapshots $\mathbf{x}_0$, $\mathbf{x}_1$, $\cdots$, $\mathbf{x}_{N-1}$. Note that the possible rank ranges from 1 to p. Suppose the model order is k. It is convenient to re-parameterize $\mathbf{R_x}$ using the spectral representation theorem. We can therefore express $\mathbf{R}_{\mathbf{x}}^{(k)}$ as $\mathbf{R}_{\mathbf{x}}^{(k)} = \sum_{i=1}^{k} \left(\lambda_i - \sigma_u^2\right)\mathbf{V}_i\mathbf{V}_i^H + \sigma_u^2\mathbf{I}$ where $\lambda_1, \cdots, \lambda_k$ and $\mathbf{V}_1, \cdots, \mathbf{V}_k$ are the eigenvalues and eigenvectors, respectively, of $\mathbf{R}_{\mathbf{x}}^{(k)}$. Denoting by $\boldsymbol{\theta}^{(k)}$ the parameter vector of the model, it is given by

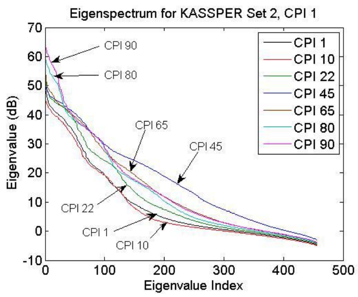
$$\boldsymbol{\theta}^{(k)} = \left[\lambda_1, \cdots, \lambda_k, \sigma_u^2, \mathbf{V}_1^T, \cdots, \mathbf{V}_k^T\right]^T. \tag{2}$$

Since EEF and MDL model order estimators require the maximum likelihood estimator (MLE) of $\boldsymbol{\theta}^{(k)}$, we note that from [10], $\hat{\lambda}_i = l_i$, $\hat{\sigma}^2 = \sum_{i=k+1}^{p} l_i/(p-k)$, and $\hat{\mathbf{V}}_i = \mathbf{C}_i (i = 1, \cdots, k)$ where $l_1 > l_2 \cdots > l_p$ and $\mathbf{C}_1, \cdots, \mathbf{C}_p$ are the eigenvalues and eigenvectors, respectively, of the sample covariance matrix $\hat{\mathbf{R}}$ given by

$$\hat{\mathbf{R}} = \frac{1}{N} \sum_{i=0}^{N-1} \mathbf{x}_i \mathbf{x}_i^H. \tag{3}$$



(a)



(b)

Fig. 1. Eigenspectra of the clutter-only (a) set 1 and (b) set 2

Now with a similar derivation as in [5], it can be shown that for complex data EEF chooses the order that maximizes

$$EEF(k) = \left(L_{G_k}(\mathbf{x}) - n_k \left[\ln\left(\frac{L_{G_k}(\mathbf{x})}{n_k}\right) + 1\right]\right) u\left(\frac{L_{G_k}(\mathbf{x})}{n_k} - 1\right) \quad (4)$$

where $u(x)$ is the unit step function, n_k is the number of free adjustable *real* parameters in the k^{th} model and $L_{G_k}(\mathbf{x})$ is defined as

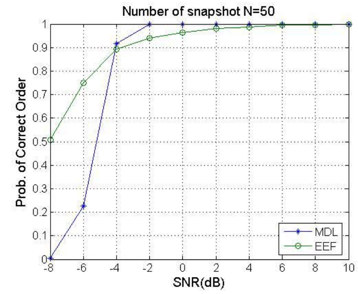
$$L_{G_k}(\mathbf{x}) = 2\ln\frac{p(\mathbf{x};\hat{\boldsymbol{\theta}}^{(k)})}{p(\mathbf{x};\hat{\boldsymbol{\theta}}^{(0)})}. \quad (5)$$

Here $p(\mathbf{x};\boldsymbol{\theta})$ is the probability density function (PDF) of the available data, and $\boldsymbol{\theta}^{(0)}$ is the parameter vector of the reference model. From Eq. (2), the free adjustable parameters in $\boldsymbol{\theta}^{(k)}$ are k real eigenvalues, k complex eigenvectors, and the noise variance. It follows that $\boldsymbol{\theta}^{(k)}$ has $k + 2pk + 1$ parameters. However, the eigenvectors are normalized and mutually orthogonal. As argued in [9], this amounts to a reduction of $k^2 + k$ degrees of freedom, so that $n_k = k(2p-k)+1$. Using this to calculate Eq. (4), it can be shown

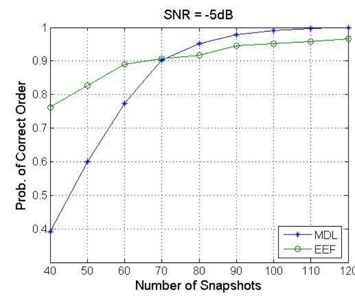
$$\begin{aligned} L_{G_k}(\mathbf{x}) &= 2N\left[\ln\left(\prod_{i=1}^{k} l_i\right) - p\ln\left(\frac{1}{p}tr(\hat{\mathbf{R}})\right)\right. \\ &\quad \left. + (p-k)\ln\left(\frac{1}{p-k}\sum_{i=k+1}^{p} l_i\right)\right]. \end{aligned} \quad (6)$$

Also, it can be shown that the MDL is

$$\begin{aligned} MDL(k) &= \frac{n_k}{2}\ln N + N\left[\ln\left(\prod_{i=1}^{k} l_i\right)\right. \\ &\quad \left. + (p-k)\ln\left(\frac{1}{p-k}\sum_{i=k+1}^{p} l_i\right)\right]. \end{aligned} \quad (7)$$



(a)



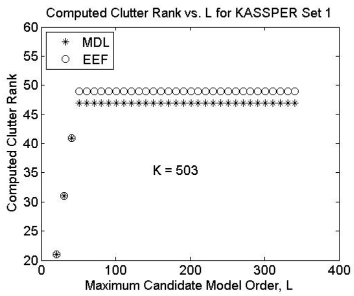
(b)

Fig. 2. Probabilities of correct order for EEF and MDL as functions of (a) SNR ($p = 7$, $N = 50$) and (b) Number of Snapshots ($p = 7$, $SNR = -5$ dB)

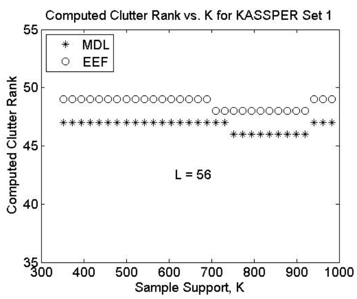
To compare both EEF and MDL methods we assume that three spatial sinusoids are present with true normalized frequencies at [0.2, 0.4, 0.6]. These three sinusoids have the same $SNR \equiv 10\log_{10}\sigma_s^2/\sigma^2$ where σ_s^2 is the variance of each sinusoid. From Monte Carlo simulation, probabilities of correct order selection versus SNR for EEF and MDL are plotted in Figure 2(a) with $p = 7$ and $N = 50$. To compare probabilities of correct order selection versus number of snapshots for EEF and MDL, we also assume three sinusoids are present with true frequencies at [0.2, 0.4, 0.6]. Results from Monte Carlo simulation are plotted in Figure 2(b), with $p = 7$ and $SNR = -5$ dB. Note that in both cases EEF outperforms MDL under low SNR and/or few number of snapshot conditions. This is because EEF can be shown to have a minimax property, in that it estimates the PDF that is closest to the true one in a minimax sense.

4. CLUTTER RANK ESTIMATION USING THE EEF AND THE MDL METHODS

Estimation of the clutter rank using EEF and MDL estimators depends on the proper choice of two parameters: the maximum model order, L, and the sample support, K, used for the subspace estimation. The computed clutter ranks as functions of L and K for set 1 are shown in Figures 3(a) and 3(b),

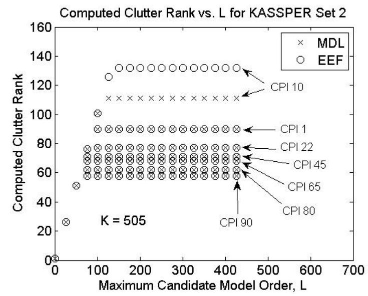


(a)

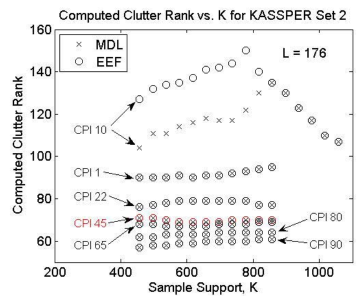


(b)

Fig. 3. Computed clutter ranks using EEF and MDL estimators for set 1 as functions of (a) L and (b) K



(a)



(b)

Fig. 4. Computed clutter ranks using EEF and MDL estimators for set 2 as functions of (a) L and (b) K

while that for set 2, CPI 1, 10, 22, 45, 65, 80, 90 are presented in Figures 4(a) and 4(b). For a fixed K, both EEF and MDL estimators depend linearly on the maximum model order L when the selected L is smaller than the expected clutter rank. However, when the selected L is larger than the expected clutter rank, both estimators become independent of L as shown in Figures 4(a). Except for CPI 10 of set 2 in Figure 4(b), EEF and MDL estimators are weakly dependent on the sample support K. Figures 3(a) and 3(b) show that for set 1, when $L \geq 49$, EEF yields $r = 49$ (48, when $752 < K < 920$) and MDL yields 47 (46, when when $752 < K < 920$). For CPI 10 of set 2, ranges of computed clutter ranks as shown in Figures 4(a) and 4(b) vary from 105 to 150 using the EEF method and from 104 to 135 using the MDL method. On the other hand, for CPI 1, 22, 45, 65, 80, and 90 of set 2, ranges of computed clutter ranks vary from 57 to 96 for both EEF and MDL methods. Although both EEF and MDL estimators provide a better estimate of the clutter rank for set 1, they fail to yield the clutter rank for set 2 according to the Brennan's rule. However, both EEF and MDL estimators can perform better when there is a sharp roll-off in eigenspectrum, e.g., CPI 10 of set 2, of which its eigenspectrum has 40 dB roll-off at eigenvalue index 175.

5. CLUTTER RANK ESTIMATION USING THE CWA METHOD

The CWA method can be reformulated to be dependent on the clutter to noise ratio, CNR, and the sample support, K. From Monte Carlo simulation, we have shown in Figures 5 and 6 that CWA strongly depends on the choice of CNR and K. In Figure 5(a), for fixed $K = 412$, CWA yields $r \sim 42$ for set 1 when $CNR = 40$ dB, which is consistent with the true CNR of set 1 (see Table 1). To confirm the choice of K, if CNR is fixed at 40 dB, the computed clutter rank at $K = 412$ is again about 42 as shown in Figure 5(b). For set 2, if K is fixed at 506, Figure 6(a) shows that CWA for set 2 CPI 1, 10, 22, 45, 65, 80, and 90 gives rise to $r \sim 382$ when $CNR = 20$ dB. On the other hand, if CNR is fixed at 20 dB, the computed clutter rank at $K = 506$ is also close to 382 as shown in Figure 6(b). Private communication with the developers of KASSPER data sets [1, 2] confirms that CNR for KASSPER data set 2 is around 23 dB. This indicates that if one selects the appropriate CNR and K, CWA can properly determine the correct clutter rank, which is governed by the Brennan's rule.

6. CONCLUSIONS

The clutter rank for KASSPER set 1 has the low-rank characteristic because r is about 42 and much less than the spatio-temporal product $JN = 352$. Except for the CPI 10 datacube, the clutter rank for KASSPER set 2 on the average is close to full rank because r is around 382 and close to $JN = 456$. Since both EEF and MDL methods are designed to estimate the model order (namely, a finite number of strong signals) for the measured data, they have better estimation performance of the clutter rank when the clutter has the low-rank character-

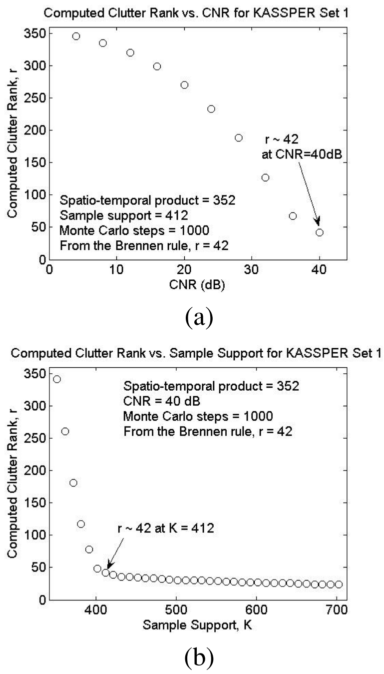


Fig. 5. Computed clutter ranks using CWA method for set 1 as functions of (a) CNR and (b) K

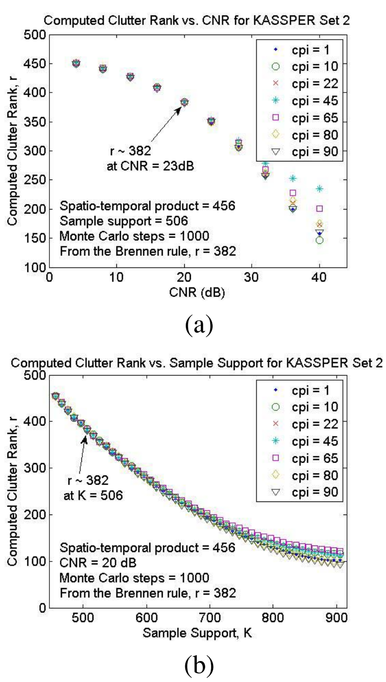


Fig. 6. Computed clutter ranks using CWA method for set 2 as functions of (a) CNR and (b) K

istic. In other words, when there is a rapid roll-off of eigenvalues in the clutter eigenspectrum, both EEF and MDL are efficient to predict the clutter rank. However, when there is a gradual roll-off of eigenvalues in the clutter eigenspectrum, both EEF and MDL yield a estimated clutter rank lower than that given by the Brennan's rule. On the other hand, as long as CNR and K can be "properly" chosen, the CWA method gives rise to the correct clutter rank. Two caveats in the application of CWA are: (i) increased computational cost compared to both EEF and MDL methods and (ii) requirement of $a\ priori$ information about CNR and K.

7. REFERENCES

[1] J. S. Bergin and P. M. Techau. High-fidelity site-specific radar simulation: Kassper'02 workshop datacube. Technical Report ISL-SCRD-TR-02-105, Defense Advanced Research Projects Agency, May 15, 2002.

[2] P. M. Techau J. S. Bergin and G. Chaney. High-fidelity site-specific radar simulation: Kassper data set 2. Technical Report ISL-SCRD-TR-02-106, Defense Advanced Research Projects Agency, October 1, 2002.

[3] M. Rangaswamy, F. C. Lin, and K. Gerlach. Robust adaptive signal processing methods for heterogeneous radar clutter scenario. *special issue of the EURASIP J. Signal Processing: New Trends and Findings in Antenna Array Processing for Radar*, **84**:1653–1665, 2004.

[4] J. Rissanen. Modeling by shortest data description. *Automatica*, pages 465–478, 1978.

[5] S. Kay. Exponentially embedded families-New approaches to model order estimation. *IEEE Trans. on Aerospace and Electronic Systems*, **41**:333–344, 2005.

[6] P. Chen, M. C. Wicks, and R. S. Adve. Development of a statistical procedure for detecting the number of signals in a radar measurement. *IEE Proc. F, Radar, Sonar and Navigation*, **148**:219–226, 2001.

[7] P. Chen, W. L. Melvin, and M. C. Wicks. Screening among multivariate normal data. *Journal of Multivariate Analysis*, **69**:10–29, 1999.

[8] J. Ward. Space-time adaptive processing for airborne radar. Technical Report Technical Report 1015, MIT Lincoln Laboratory, December 1994.

[9] M. Wax and T. Kailath. Detection of signals by information theoretic criteria. *IEEE Trans. on Acoust., Speech, Signal Processing*, **33**:387–392, 1985.

[10] T. W. Anderson. Asymptotic theory for principal component analysis. *The Annals of Mathematical Statistics*, **34**:122–148, 1963.