HYPERSPECTRAL PERFORMANCE PREDICTION OF THE ADAPTIVE COSINE ESTIMATOR

Eric Truslow[⋆†], *Dimitris Manolakis*[⋆], *Michael Pieper*[⋆], *Thomas Cooley*[†] *and Michael Brueggeman*[†]

[⋆†] Northeastern University,
360 Huntington Ave.
Boston, MA 02115

[⋆] MIT-Lincoln Laboratory,
244 Wood Street
Lexington, MA 02421

[†] Air Force Research Laboratory,
Space Vehicles Directorate,
3550 Alberdeen Ave SE
Kirtland AFB, NM 87117

ABSTRACT

The adaptive cosine estimator is a popular and effective algorithm for detecting materials in hyperspectral images. To predict the performance of this algorithm in real hyperspectral scenes, a statistical model using a mixture of multivariate t-distributions for the background and a Gaussian distribution for the target is utilized. In this paper, two methods for finding the response of the adaptive cosine estimator (ACE) and Beta-detector when applied to a statistical model. To verify that the proposed techniques work as expected, t-distribution and F-distribution quantiles are computed and compared to standard values. Finally, a preliminary validation with Monte Carlo simulation based on real hyperspectral data is presented. We build on previous work for the matched filter and extends it to use two more detectors.

***Index Terms*—** Hyperspectral imaging, detection algorithms, signal detection, matched filters

1. INTRODUCTION

The detection of materials in hyperspectral imagery (HSI) is a useful, but challenging area of research; the large amount of data that must be analyzed, necessitates the use of automated detection algorithms [1]; while a variety of detection algorithms are available, the absolute performance of these algorithms is less well understood. The end goal is to be able to predict how well a hyperspectral detection system will perform when searching for a material in a scene.

To this end, the performance prediction model herein consists of a statistical characterization of the background, as discussed in [2] and [3], and a linear mixture for the target as in [4]. In this setup, the background of the scene is statistically described using multiple elliptical t-distributions, while the target is modeled by a single Gaussian distribution. As shown in Fig. 1 and following [5], a performance estimate is obtained by propagating the background and target statistics through a detector to find the cumulative distribution function (CDF) of the output; the corresponding probability of false alarm (PFA) and probability of detection (PD) for any threshold determine the expected performance.

In [5] the performance of the matched filter (MF) is considered in the same context, while here techniques for propagating the background and target statistics through the adaptive cosine estimator (ACE) and beta-detector are presented. Throughout, we will use the term beta-detector to refer to the squared version of ACE, but both are referred to as ACE in the literature. The ACE and beta-detector are two well known and widely used detectors, but because they are non-linear detectors, they are often difficult to statistically characterize. As discussed more thoroughly at the end of Section 2 the model we use does not result in a simple solution.

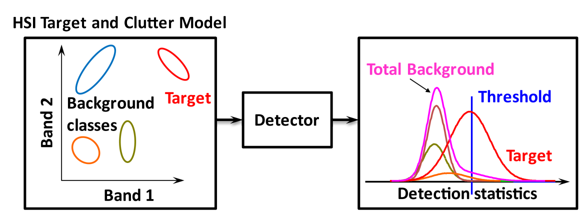

Fig. 1. Illustration of prediction model with background and target.

Although the model is multi-modal, each background cluster passes through the detector independent of the others. Therefore, to use ACE or any other detector in place of the MF, only a single class needs to be examined at a time. Thus, we can consider a single component at the input as shown in Fig. 2 and summarized as: given some input $\mathbf{x}$ with probability density function (PDF) $f(\mathbf{x})$ and a detector $y = D(\mathbf{x})$, find the distribution of the output $f(y)$. In this paper, $D(\mathbf{x})$ is either the ACE or beta-detector and $\mathbf{x}$ is an elliptically distributed vector.

The rest of the paper is structured as follows: in Section 2 the statistical model for the input and the MF, ACE and Beta-

This work is sponsored by the Air Force Research Laboratory under Air Force Contract #FA8721-05-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the United States Government.

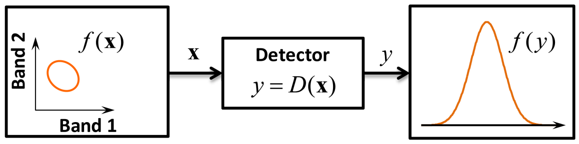

Fig. 2. Operation of detector with unimodal distribution at the input.

detectors are introduced; in Section 3 a technique for finding the CDF of ACE is proposed; in Section 4 a similar technique for finding the CDF of the Beta-detector is proposed; in Section 6 our techniques are compared to standard F-distribution and t-distribution CDFs and a preliminary validation with real data is conducted. In Section 7 we give a short summary.

2. BACKGROUND

Restricting our attention to Gaussian input. The input to the target detection system may be defined by a p-dimensional column vector $\mathbf{x}_0$. The target signature is also a p-dimensional vector denoted $\mathbf{s}_0$. Assume that the background mean vector $\mathbf{m}_\text{b}$ and positive definite background covariance matrix $\mathbf{C}_\text{b}$ are known (non-random), and are used to design the detector; then let the input $\mathbf{x}_0$ be

$$\mathbf{x}_0 \sim \mathcal{N}\left(\mathbf{m}_0, \mathbf{C}\right)$$

where $\mathbf{m}_0$ and $\mathbf{C}$ are the mean vector and covariance matrix of the input. Prior to detection, the input data and target signature are mean-subtracted, yielding

$$\mathbf{x} = \mathbf{x}_0 - \mathbf{m}_\text{b} \qquad \text{and} \qquad \mathbf{s} = \mathbf{s}_0 - \mathbf{m}_\text{b}.$$

Thus, $\mathbf{x} \sim \mathcal{N}\left(\mathbf{m}, \mathbf{C}\right)$ where $\mathbf{m} = \mathbf{m}_0 - \mathbf{m}_\text{b}$. Then, defining the whitened vectors

$$\tilde{\mathbf{x}} = \mathbf{C}_\text{b}^{-1/2}\mathbf{x} \qquad \text{and} \qquad \tilde{\mathbf{s}} = \mathbf{C}_\text{b}^{-1/2}\mathbf{s}$$

the detectors of interest – the Matched Filter (MF), the Adaptive Cosine Estimator (ACE), and the Beta-detector – are defined as

$$y_\text{MF} = \frac{\tilde{\mathbf{x}}^\text{T}\tilde{\mathbf{s}}}{\sqrt{\tilde{\mathbf{s}}^\text{T}\tilde{\mathbf{s}}}} \tag{1}$$

$$y_\text{ACE} = \frac{\tilde{\mathbf{x}}^\text{T}\tilde{\mathbf{s}}}{\sqrt{\tilde{\mathbf{s}}^\text{T}\tilde{\mathbf{s}}}\sqrt{\tilde{\mathbf{x}}^\text{T}\tilde{\mathbf{x}}}} \tag{2}$$

$$y_\beta = \frac{(\tilde{\mathbf{x}}^\text{T}\tilde{\mathbf{s}})^2}{(\tilde{\mathbf{s}}^\text{T}\tilde{\mathbf{s}})(\tilde{\mathbf{x}}^\text{T}\tilde{\mathbf{x}})} \tag{3}$$

These versions of y_ACE and y_β are commonly used in practice, but are not convenient for statistical results; instead, introducing the projection matrices

$$\mathbf{P}_{\tilde{\mathbf{s}}} = \tilde{\mathbf{s}}(\tilde{\mathbf{s}}^\text{T}\tilde{\mathbf{s}})^{-1}\tilde{\mathbf{s}}^\text{T} \tag{4}$$
$$\mathbf{P}_{\tilde{\mathbf{s}}}^\perp = \mathbf{I} - \mathbf{P}_{\tilde{\mathbf{s}}}$$

two statistics equivalent to Eq. 2 and Eq. 3 are the "cotangent" and F-detector:

$$y_\text{cot} = \frac{y_\text{ACE}}{\sqrt{1 - y_\text{ACE}^2}} = \frac{\tilde{\mathbf{s}}^\text{T}\mathbf{P}_{\tilde{\mathbf{s}}}\tilde{\mathbf{x}}}{\sqrt{\tilde{\mathbf{s}}^\text{T}\tilde{\mathbf{s}}}\sqrt{\tilde{\mathbf{x}}^\text{T}\mathbf{P}_{\tilde{\mathbf{s}}}^\perp\tilde{\mathbf{x}}}} \tag{5}$$

$$y_\text{F} = \frac{y_\beta}{1 - y_\beta} = \frac{(\tilde{\mathbf{x}}^\text{T}\tilde{\mathbf{s}})^2}{(\tilde{\mathbf{s}}^\text{T}\tilde{\mathbf{s}})(\tilde{\mathbf{x}}^\text{T}\mathbf{P}_{\tilde{\mathbf{s}}}^\perp\tilde{\mathbf{x}})} \tag{6}$$

Both of these statistics are ratios of orthogonal components and throughout the rest of this paper, these versions will be used instead of Eq. 2 and Eq. 3.

To predict the performance of these algorithms, the distribution of the output needs to be found. Finding the distribution, or equivalently the CDF, of the output is required to find the PD and PFA, which are used to measure performance. The distribution of the ACE and Beta-detector have been well studied in [6] and [7], but the statistical results are derived using two critical assumptions about the input:

$$\mathbf{m} = a\mathbf{s} \qquad \text{and} \qquad \mathbf{C} = \sigma^2\mathbf{C}_\text{b}$$

where a and σ^2 are scalars. The importance of these assumptions is made clear by examining the distribution of $\tilde{\mathbf{x}}$. In general, the input to the detectors becomes

$$\tilde{\mathbf{x}} \sim \mathcal{N}\left(\tilde{\mathbf{m}}, \tilde{\mathbf{C}}\right) \tag{7}$$

where

$$\tilde{\mathbf{m}} = \mathbf{C}_\text{b}^{-1/2}(\mathbf{m}_0 - \mathbf{m}_\text{b}) \text{ and} \tag{8}$$
$$\tilde{\mathbf{C}} = \mathbf{C}_\text{b}^{-1/2}\mathbf{C}\mathbf{C}_\text{b}^{-1/2} \tag{9}$$

The first assumption constrains the mean of the input to be in the direction of $\tilde{\mathbf{s}}$; the second assumption constrains the covariance matrix of the input to be a scalar multiple of the background covariance, and yields uncorrelated components of equal variance: $\tilde{\mathbf{C}} = \sigma^2\mathbf{I}$. In the performance prediction model we adopt [5], neither of these constraints is enforced; therefore, in this paper, a set of alternative techniques is proposed.

3. THE DISTRIBUTION OF ACE

In general we wish to find a general expression for the CDF of y_cot for as outlined in Fig. 3. From Eq. 5, an equivalent

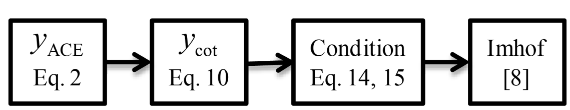

Fig. 3. Illustration of steps in computing CDF of the Beta-detector.

way to finding the CDF of y_{ACE} is

$$\Pr\{y_{\text{cot}} \le \eta_{\text{cot}}\} = \Pr\left\{\frac{y_{\cos}}{\sqrt{1-y_{\cos}}} \le \eta_{\text{cot}}\right\} \tag{10}$$

where η_{cot} is the threshold for the detector. Using the definition of the MF in Eq. 1, and defining

$$q = \tilde{\mathbf{x}}^{\mathrm{T}}\mathbf{P}_{\tilde{\mathbf{s}}}^{\perp}\tilde{\mathbf{x}} \tag{11}$$

the cotangent in Eq. 5 can be written as

$$y_{\text{cot}} = \frac{y_{\text{MF}}}{\sqrt{q}} \tag{12}$$

Substituting into Eq. 10 and rearranging, yields

$$\Pr\{y_{\text{cot}} \le \eta_{\text{cot}}\} = \Pr\{y_{\text{MF}} \le \eta_{\text{cot}}\sqrt{q}\} \tag{13}$$

Conditioning on q, the CDF of y_{cot} becomes

$$\Pr\{y_{\text{cot}} \le \eta_{\text{cot}}\} = \int_0^{\infty} \Pr\{y_{\text{MF}} \le \eta\sqrt{q}|q\}\, f_q(q)dq \tag{14}$$

where $\Pr\{y_{\text{MF}} \le \eta\sqrt{q}|q\}$ is the CDF of the MF given q. Two functions remain unknown: the marginal probability density function (PDF) of q $f_q(.)$ and $\Pr\{y_{\text{MF}} \le \eta\sqrt{q}|q\}$. See Section 3.1 for the former and Section 3.2 for the latter.

3.1. The Probability Density Function of q

The PDF of the quadratic form q in Eq. 11 is not readily accessible in closed form, but the characteristic function is [8]. Denoting the characteristic function of q as $\phi_q(t)$, the PDF of q is obtained via an inverse Fourier transform

$$f_q(q) = \mathcal{F}^{-1}\left\{\overline{\phi}_q(t)\right\} \tag{15}$$

where $\overline{\phi}_q(.)$ is the complex conjugate of the characteristic function [9]. The variable q is a sum of squares of the components of $\tilde{\mathbf{x}}$, which can be represented as a sum of independent chi-squared variables [10]. To calculate the characteristic function $\phi_q(t)$ of q, the non-centrality parameters δ_i and scaling parameters λ_i ($i = 1, \ldots, p$) of the constituent variables can be found through the technique in [8].

3.2. The Conditional Distribution of the MF

To find the CDF of the MF given q there are two cases to consider:

1. the MF and q are uncorrelated and independent, and
2. the MF and q are correlated.

In the former case, the conditional distribution of the MF is the marginal distribution; it is well known that the MF is Gaussian distributed when the input $\mathbf{x}$ is Gaussian distributed [5]; thus $\Pr\{y_{\text{MF}} \le \eta\sqrt{q}|q\}$ is the CDF of a Gaussian random variable; when the MF and q are nearly uncorrelated, this is a reasonable approximation. In future work, the latter case will be addressed more thoroughly.

4. THE DISTRIBUTION OF THE BETA-DETECTOR

In this section, an expression for the CDF of the Beta-detector is sought. The processing chain used is shown in Fig. 4 Recall

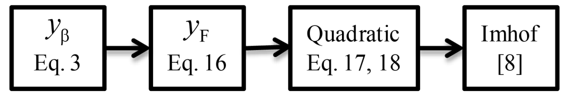

Fig. 4. Illustration of steps in computing CDF of the Beta-detector.

that for statistical analysis it is preferable to use y_{F} of Eq. 6 instead of Eq. 3. However, the two detectors are equivalent since

$$\Pr\{y_{\text{F}} \le \eta_F\} = \Pr\left\{\frac{y_\beta}{1-y_\beta} \le \eta_F\right\} \tag{16}$$

From Eq. 6 and the projection matrix of Eq. 4, the F-detector can be written conveniently in terms of projection matrices as

$$y_{\text{F}} = \frac{(\tilde{\mathbf{x}}^{\mathrm{T}}\tilde{\mathbf{s}})^2}{(\tilde{\mathbf{s}}^{\mathrm{T}}\tilde{\mathbf{s}})(\tilde{\mathbf{x}}^{\mathrm{T}}\mathbf{P}_{\tilde{\mathbf{s}}}^{\perp}\tilde{\mathbf{x}})} = \frac{\tilde{\mathbf{x}}^{\mathrm{T}}\mathbf{P}_{\tilde{\mathbf{s}}}\tilde{\mathbf{x}}}{\tilde{\mathbf{x}}^{\mathrm{T}}\mathbf{P}_{\tilde{\mathbf{s}}}^{\perp}\tilde{\mathbf{x}}}.$$

Following the work in [11] this allows the CDF of y_{F} in Eq. 16 to be rewritten as

$$\Pr\{y_{\text{F}} \le \eta_F\} = \Pr\left\{\tilde{\mathbf{x}}^{\mathrm{T}}(\mathbf{P}_{\tilde{\mathbf{s}}} - \eta_F\mathbf{P}_{\tilde{\mathbf{s}}}^{\perp})\tilde{\mathbf{x}} \le 0\right\} \tag{17}$$

$$= \Pr\left\{\tilde{\mathbf{x}}^{\mathrm{T}}\mathbf{B}(\eta_F)\tilde{\mathbf{x}} \le 0\right\}. \tag{18}$$

where $\mathbf{B}(\eta) = \mathbf{P}_{\tilde{\mathbf{s}}} - \eta_F\mathbf{P}_{\tilde{\mathbf{s}}}^{\perp}$; this expression is now in terms of a single quadratic form that is a function of threshold η_F. The CDF of this quadratic form is generally unavailable in closed form, but, as with q in Eq. 11, its characteristic function can be found using the technique in [8], and inverted to find its CDF.

5. HANDLING HEAVY-TAILED INPUT

Up to this point the input to the detector has been Gaussian distributed, but the background model used in [5] uses the heavy-tailed multivariate t-distribution (MVT). Letting

$$\mathbf{x} \stackrel{d}{=} \mathbf{m} + \mathbf{z}(s/\nu)^{-1/2} \tag{19}$$

where $s \sim \chi^2_\nu$ and $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{C})$ then $\mathbf{x}$ follows a MVT distribution with mean $\mathbf{m}$, shape matrix $\mathbf{C}$ and ν degrees of freedom [12].

Consider the case where $\mathbf{m} = \mathbf{0}$; substituting Eq. 19 into Eq. 2 or Eq. 5, the scaling (s/ν) cancels. This property makes the ACE and Beta-detector insensitive to ν in some cases, and less sensitive to ν when the mean $\mathbf{m}$ is small [7]. Therefore, in many cases, assuming Gaussian input is a good approximation for these detectors, and the results derived previously are applicable.

In situations where an exact result is desired, conditional probability can be employed; by conditioning on s it is apparent from Eq. 19 that $\mathbf{x} \sim \mathcal{N}\left(\mathbf{m}, \mathbf{C}(s/\nu)^{-1}\right)$, which means that by conditioning on s the results for Gaussian inputs can be applied. The total CDF of a detector in this case is

$$F(y) = \int_0^\infty F(y|s) f(s) ds$$

where $F(y|s)$ is the CDF of the detector given a value of s. This adds a level of complexity to the Gaussian results, but is unnecessary in many cases, especially when $\mathbf{m}$ is negligibly small.

6. RESULTS

In this section, we verify that the algorithms outlined in Fig. 4 and Fig. 3 produce the expected output. Experiments were conducted using Eq. 5 and Eq. 6. Here we enforce the constraints mentioned earlier: $\mathbf{m} = a\mathbf{s}$ and $\mathbf{C} = \mathbf{C}_{\mathrm{b}}$. For this experiment the number of dimensions $p = 140$ and $\mathbf{C} = \mathbf{I}$. Then, the variables $\sqrt{p-1} y_{\mathrm{cot}} \sim t_{p-1}(\delta)$ and $p y_{\mathrm{F}} \sim \mathrm{F}_{1,p}(\delta^2)$ follow non-central t and F-distributions respectively. The non-centrality parameter is $\delta = \sqrt{\mathbf{m}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{s}}$. In Fig. 5 and Fig. 6 the MATLAB implementation of the noncentral-F and noncentral-t exceedances are shown (exceedance $= 1-$ CDF). As shown, there is good agreement between the two techniques to about 10^{-8}.

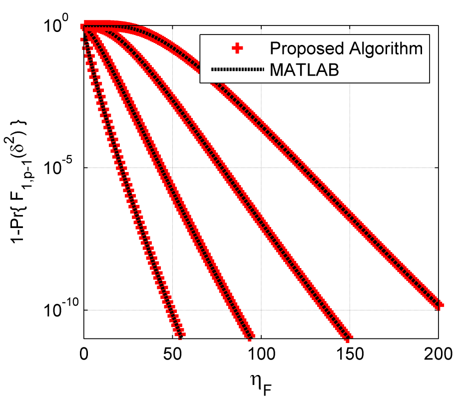

Fig. 5. Noncentral F-distribution exceedance plots with $\delta = 0, 2, 4, 8$ (left to right).

6.1. Algorithm Verification with Monte Carlo Simulation

In this experiment, a hyperspectral background model was created using real hyperspectral data, and the linear mixing model was used to simulate a target of various fill fractions. For verification 10^5 pixels were generated from both the background and that target, and fill fractions of 3.0%, 5.0%, 6.5% and 8.5% were selected for the target. The results of these simulations for the Beta-detector are shown in Fig. 7. While

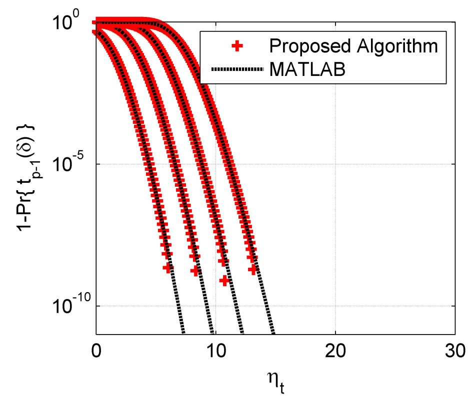

Fig. 6. Noncentral t-distribution exceedance plots with $\delta = 0, 2, 4, 8$ (left to right).

there is good agreement between the Monte Carlo estimated Receiver Operating Characteristic (ROC) curve and the predicted ROC curve [13]. Near a PFA of 10^{-5} there is more variability in the estimate because there are few samples in this region.

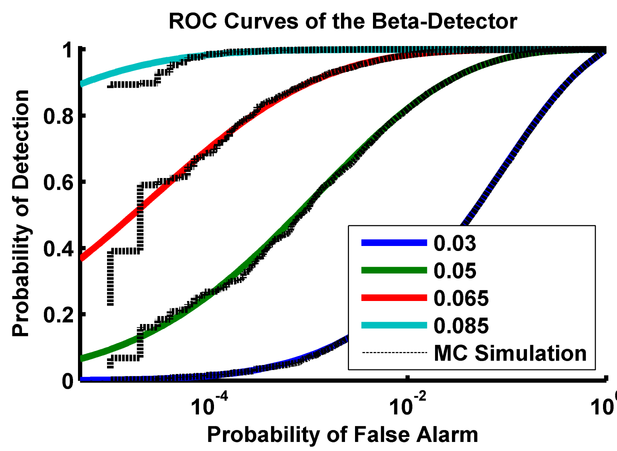

Fig. 7.

7. CONCLUSION

In this paper we have presented techniques for computing the CDF of two detection algorithms. These techniques do not rely on Monte Carlo simulation and fit well with previous work on performance prediction that was related to the matched filter. Finally, we demonstrated that this technique can be used to compute t-distribution quantiles and F-distribution quantiles. In the future we hope to discuss implementation details and to test the output of such a model against real data.

8. REFERENCES

[1] D.G. Manolakis, D.B. Marden, and G.A. Shaw, "Hyperspectral image processing for automatic target detection applications," *Lincoln Laboratory Journal*, vol. 14, no. 1, pp. 79–116, 2003.

[2] D.B. Marden and D.G. Manolakis, "Modeling hyperspectral imaging data using elliptically contoured distributions," in *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery*, Sylvia S. Shen and Eds. Paul E. Lewis, Eds., Orlando, FL, 2003, SPIE.

[3] D. Manolakis, M. Rossacci, J. Cipar, R. Lockwood, T. Cooley, and J. Jacobson, "Statistical characterization of natural hyperspectral backgrounds," in *Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006. IEEE International Conference on*, 312006-aug.4 2006, pp. 1624 –1627.

[4] D. Manolakis, C. Siracusa, and G. Shaw, "Hyperspectral subpixel target detection using the linear mixing model," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 39, no. 7, pp. 1392 –1409, jul 2001.

[5] Dimitris Manolakis, "Realistic matched filter performance prediction for hyperspectral target detection," *Optical Engineering*, vol. 44, no. 11, pp. 116401–116401–7, 2005.

[6] L.L. Scharf and B. Friedlander, "Matched subspace detectors," *Signal Processing, IEEE Transactions on*, vol. 42, no. 8, pp. 2146 –2157, Aug. 1994.

[7] S. Kraut, L.L. Scharf, and L.T. McWhorter, "Adaptive subspace detectors," *Signal Processing, IEEE Transactions on*, vol. 49, no. 1, pp. 1–16, 2001.

[8] J. P. Imhof, "Computing the distribution of quadratic forms in normal variables," *Biometrika*, vol. 48, no. 3/4, pp. 419–426, 1961.

[9] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, 1984.

[10] Norman Johnson, *Continuous Univariate Distributions, Vol. 2*, J. Wiley, New York, 1970.

[11] S. Broda and M.S. Paolella, "Evaluating the density of ratios of noncentral quadratic forms in normal variables," *Computational Statistics & Data Analysis*, vol. 53, no. 4, pp. 1264 – 1270, 2009.

[12] R.J. Muirhead, *Aspects of Multivariate Statistical Theory*, Wiley series in probability and mathematical statistics: Probability and mathematical statistics. Wiley, 1982.

[13] W.H. Tranter, *Principles of Communication Systems Simulation With Wireless Applications*, Prentice Hall Communications Engineering and Emerging Techno Series. Prentice Hall, 2004.