A Machine Learning Approach to Cognitive Radar Detection

Justin Metcalf[1,2], Shannon D. Blunt[1], and Braham Himed[2]
[1]Radar Systems Lab (RSL), University of Kansas
[2]Sensors Directorate, Air Force Research Laboratory

Abstract— **We consider the requirements of cognitive radar detection in the presence of non-Gaussian clutter. A pair of machine learning approaches based on non-linear transformations of order statistics are examined with the goal of adaptively determining the optimal detection threshold within the low sample support regime. The impact of these algorithms on false alarm rate is also considered. It is demonstrated that the adaptive threshold estimate is effective even when the distribution in question is unknown to the machine learning algorithm.**

I. Introduction

In addition to the familiar passive sensing modalities, such as sight and hearing, the natural world is full of examples of active sensing modalities. The most notable example is biosonar employed by bats, dolphins, and whales, the operation and characteristics of which bears an understandable resemblance to the systems that have been developed by the radar and sonar communities [1], [2]. Using these natural biological systems as inspiration the concept of cognitive, or fully adaptive, radar has been an active area of recent research [3]–[11]. In this spirit, we consider the needs of a cognitive radar detector.

The concept of a cognitive radar, particularly in a bioinspired sense, lends itself naturally to target tracking and automatic target recognition tasks [3], [4], [6], [7]. The cognitive aspect manifests in the form of persistent memory and a perception-action cycle. The perception-action cycle requires the ability to adapt waveforms to the non-stationary environment. In addition, cognition for a single radar may be expanded into a network of multiple radars intelligently cooperating to accomplish the mission goals [8], [11]. However, while the main foci in the literature has been on adaptive waveform design or cognitive target recognition strategies, here we concentrate on the detection statistics needed to maintain an acceptable false alarm rate over a multitude of operational environments.

A radar detector must incorporate previously determined knowledge to estimate the statistical characteristics of the operating environment. The estimated statistics are then used to provide the highest possible probability of detection, while maintaining a constant, acceptable probability of false alarm. The environmental characteristics are governed by two primary phenomenon: clutter and thermal noise. Typically, the superposition of the backscattered echoes received by the radar detector may be considered to be the sum of a large number of independent, identically distributed (I.I.D.) random variables.

Therefore, the Central Limit Theorem may be invoked and the clutter is commonly assumed to be Gaussian distributed [12]. Assuming a large transmit power, the magnitude of the clutter is assumed to be much greater than the thermal noise, such that the thermal noise statistics may be ignored.

The statistical distribution of the clutter provides the context necessary to form the optimal Neyman-Pearson detector, or the detector that has the maximum probability of detecting a target with an acceptable false alarm rate [13]. If the clutter is distributed as complex Gaussian, the optimum detector requires knowledge of the first and second moments. With the Gaussian assumption, there have been efforts to use prior estimates of the covariance matrix of the environment [14], [15]. However, it has also been established that clutter may follow a number of non-Gaussian distributions, most notably the K, Pareto, and Weibull distributions [12], [16]–[24]. Because they are heavy-tailed, the occurrence of such distributions as these in practice invariably leads to increased false alarms when the Gaussian assumption is made.

By definition, a radar detector is sampling a physical process. The parameters of the clutter process depend on multiple factors such as the physical environment, the grazing angle of the radar, and the range resolution. Therefore, given the sheer number of influences on the clutter statistics, the existence of clutter distributions that differ to some degree from the existing models [12], [16]–[24] remains a possibility. However, cognitive systems in nature are adept at adapting to new situations by leveraging prior knowledge. Therefore, a true cognitive radar should adaptively estimate an accurate threshold regardless of the distribution to which the clutter process belongs. Building from the results of [25], here we consider the framework of a cognitive detector and explore one method of implementing such a system.

II. Modeling Radar Clutter

The characterization of radar clutter is dependent on the interplay of the design parameters of the radar (*e.g.* bandwidth, grazing angle, etc.) as well as the operating environment. The complicated relationship between these parameters requires that care be taken when assuming the characteristics of the clutter response. The overall clutter response is often considered to be composed of contributions from two different types of clutter: distributed and discrete.

For distributed clutter each range cell is assumed to contain a large number of individual, electrically small elementary

scatterers. The backscattered radiation from the range cell may then be characterized as the sum of a large number of independent, identically distributed random variables. Under this assumption, the Central Limit Theorem may be invoked and the clutter response follows a complex Gaussian distribution with average power proportional to the product of the transmit power and the backscatter coefficient of the illuminated terrain.

Discrete clutter results from specular returns, often from objects that resemble plate or corner reflectors. From a physical perspective, when the size of a range cell is decreased via a smaller beamwidth or a transmitted waveform with an increased bandwidth (with an unchanged pulse-width), the distributed clutter power in each new range cell decreases proportionally. However, discrete clutter in the smaller range cells may come to dominate the clutter return, leading to outliers in the measured clutter.

Observing the trend of increases in measured outliers as range resolution becomes finer, alternative statistical distributions with heavier tails have been examined (*e.g.* [12], [16]–[24] and the references therein). The class of spherically invariant random processes (SIRPs) arises when the number of elementary scatterers is assumed to result from a mixture of Poisson random variables, even if the expected value of the number of scatterers is very large. In addition, most of the commonly fitted clutter distributions (notably the K, Weibull, and Pareto distributions) belong to the class of SIRPs. Therefore, SIRPs may be justified from both an empirical and a theoretical point of view.

A multidimensional sample of a SIRP yields a spherically invariant random vector (SIRV). Crucially, a SIRV may be expressed as a Gaussian random vector that is modulated by a positive random variable. From a physical perspective, this model considers the temporally correlated L slow time returns from a single range cell to be locally distributed as complex Gaussian. However, when a collection of the length L vectors are considered (corresponding to a homogoneous group of range cells) there exists a power modulation from cell to cell. The distribution of the modulating random variable then controls the length of the tail of the SIRV.

A zero mean SIRV can be characterized by the quadratic form

$$q = \mathbf{y}^H \mathbf{\Sigma}^{-1} \mathbf{y}, \tag{1}$$

where $\mathbf{y}$ is a length L complex SIRV, $\mathbf{\Sigma}$ is the covariance matrix of $\mathbf{y}$, and $(\bullet)^H$ denotes the complex-conjugate transpose. The pdf of the random vector $\mathbf{y}$ can then be expressed in terms of the quadratic form q and the modulating random variable V as

$$f_{\mathbf{Y}}(\mathbf{y}) = (\pi)^{-L} |\mathbf{\Sigma}|^{-1} \int_0^\infty v^{-2L} \exp\left(-\frac{q}{v^2}\right) f_V(v) \mathrm{d}v, \tag{2}$$

where $|\mathbf{\Sigma}|$ denotes the determinant of the covariance matrix. Note that (2) illustrates that the pdf of V is the only degree of freedom between different SIRV distributions with identical covariance structure. The distribution of V is often parametrized by a shape parameter, ν. It is easily shown that the Gaussian distribution is obtained from (2) by setting the

modulating random variable $f_V(v) = \delta(v-1)$, where $\delta(v)$ is the impulse function [17].

SIRVs provide an attractive model for radar clutter for several reasons. First, as has been established, the SIRV model can be justified from both a theoretical and an empirical basis. By considering the entire class of SIRVs, a cognitive radar detector relaxes statistical assumptions of Gaussianity to provide an additional degree of freedom over which to adapt while maintaining a physical justification for the model. Second, SIRVs are closed under linear transforms [17]. Therefore, under any linear transformation only the mean and covariance matrix will change. The closure property allows for the detection in SIRV clutter to take the familiar form of a whitening filter (*i.e.* via the quadratic form of (1)) followed by an operation on the data-dependent threshold governed by (2) [12]. Third, exploitation of the closure property and separability of the modulating random variable V and the underlying Gaussian process allows for convenient simulation of SIRV radar clutter.

It should be noted that there are several challenges when using SIRVs as a model for clutter. First, there is often no closed form solution available, leading to reliance on numerical and Monte Carlo methods to evaluate theoretical performance [26]. Also, for low shape parameter values (*i.e.* very heavy tailed) the numerical estimation of the pdf and cdf of SIRVs can suffer from numerical instability [27]. Second, due to the individual scaling induced on each sample random vector, estimating the covariance matrix is a rather difficult open problem [28]–[30]. Finally, multiple SIRV distributions may fit measured data equally well [19]. Therefore, the possibility of heretofore unknown distributions must be considered. The problem of identifying the true underlying SIRV distribution is directly related to the difficulty in estimating the sample covariance, as the maximum likelihood estimation of the covariance matrix requires knowledge of the pdf of (2).

III. CONSEQUENCES OF NON-GAUSSIAN CLUTTER

A radar detector must estimate the statistics of a cell-under-test from the returns of the surrounding (assumed homogenous) range cells. The Neyman-Pearson detector is then the detector that maximizes the detection probability while maintaining a desired probability of false alarm P_{fa} [13]. The detection threshold is set according the the tail of the distribution of the null (target absent) hypothesis. Clearly, heavier tailed distributions directly lead to an increased threshold to maintain the same P_{fa}.

Section II discussed the dependence of the optimal detector on properly estimating the whitening stage of (1), and knowledge of (2) in order to set the correct data-dependent threshold. However, these two estimates are intrinsically linked via the estimation of the covariance matrix [27]. For the sake of brevity, here we consider the covariance matrix to be clairvoyantly known.

As an example, consider the K distribution, whose tail is governed by a shape parameter ν. The K distribution is notable for providing a particularly good fit to sea clutter [12]. Small
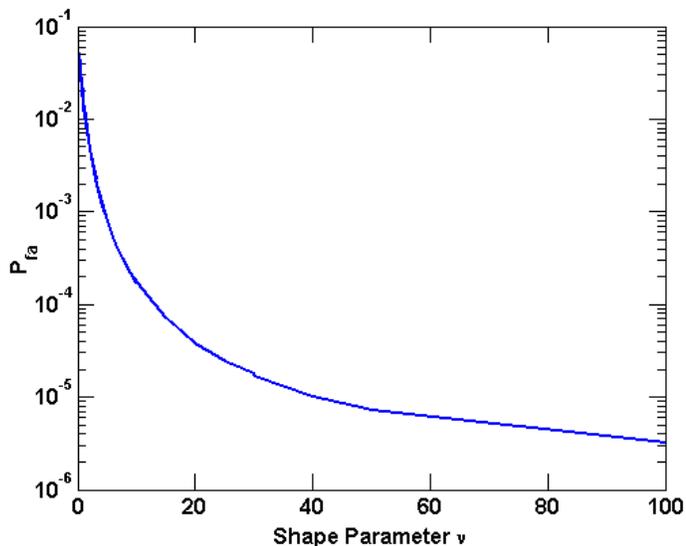
Fig. 1. Increased P_{fa} from K distributed clutter

values of ν lead to a heavy tailed distribution, but the K distribution converges to Gaussian as $\nu \to \infty$. To illustrate the consequences of the K distribution, consider a radar detector with clairvoyant knowledge of the covariance matrix but with the assumption of Gaussian clutter for dimensionality $L = 4$ slow time samples. The radar detector can then determine the optimal detection threshold $\mathcal{T}_\mathrm{G}$ in the presumed presence of Gaussian clutter for a $P_\mathrm{fa} = 10^{-6}$. Figure 1 shows the resultant actual P_fa when K distributed clutter of varying shape parameters is present but the Gaussian-presumed threshold $\mathcal{T}_\mathrm{G}$ is used. Noting that the K distributed clutter requires a unique threshold for each value of ν, it is clear that a wrongly assumed distribution can yield a significant increase in false alarms.

IV. CONSIDERATIONS FOR A COGNITIVE DETECTOR

It has been established that SIRVs are an attractive model for radar clutter. Not only are most known non-Gaussian distributions admissible as SIRVs, the Gaussian distribution is also a SIRV. However, predicting which SIRV is the best model for the radar clutter encountered in any given scenario remains an open problem. One possible approach is to construct a database of distributions identified for previously encountered geographical areas. However, such a database suffers from several problems. First, the distribution measured is strongly dependent on transmit beamwidth, time-bandwidth product, and grazing angle. Therefore, dependence on such a database necessarily imposes a tradeoff between on-the-fly emission controls and the validity of previous measurements. Second, in a long term temporal sense geography may be non-stationary. Human influence can change geographic features or introduce discrete clutter. For instance, the clutter response of farmland will vary greatly in the winter compared to during a harvest when large radar cross-section farm machinery is scattered throughout a scene.

For these reasons, the goal of a cognitive radar should be to

adapt to the present conditions based on *suggestions* from past experiences. One method would be to construct a Bayesian approach to distribution identification. The prior distribution can be formed in a knowledge-aided approach based on past measurements of similar geography and transmit parameters. The best estimate of the current distribution would then be formed based on the prior and current measurements. Once the distribution is determined, an optimal threshold may be derived. A second method is to utilize machine learning to suggest directly a threshold based on the measured data and knowledge of the commonly encountered threshold.

Methods of distribution identification for SIRVs were presented in [16], [17], [31]–[33] and expanded upon in [25]. In [16], [17], [25], [31]–[33] the primary goal was to use a non-linear operation based on order statistics to identify the SIRV distributions that best fit a set of sample data by comparing the test statistics to a precomputed library of known distributions. The non-linear transformations used in [16], [17], [25], [31]–[33] provided a unique, one-to-one mapping between distribution/shape parameter pairs and the transformation space. However, the algorithms required large sample support to classify the distribution correctly. Recall from (2) that the distinction between SIRVs of the same dimension is encapsulated in the positive random variable V. However, distributions such as the K, Pareto, and Weibull all have continuously defined shape parameters. Therefore, it is inevitable that ambiguities will arise. Due to the difficulties in matching measured data having low sample support to the generating SIRV, the Bayesian approach is not further considered here.

V. WEIGHTED ORDER STATISTICS

In [25], it was suggested that the ambiguity inherent in SIRV distributions would allow the distribution identification algorithm of [31]–[33], denoted as the Ozturk Algorithm after the lead author, to be adapted to direct threshold estimation. Here we expand these results and consider a new formulation.

To form a library of distributions, we find the expected value of a test statistic for a series of known SIRV distributions and shape parameter pairs. The basis of this test statistic is formed from a set of N length L SIRVs. In the context of radar detection, this set corresponds to N homogeneous, target-free range cells measured over L pulses. For the purpose of forming the library, the covariance matrix is known. The set of vectors are compressed to a set of quadratic forms $\mathbf{q} = [q_1, \ldots, q_N]$ via (1). The order statistics of $\mathbf{q}$ are then formed by sorting the samples such that

$$q_{(1)} \leq q_{(2)} \leq \cdots \leq q_{(N)}. \tag{3}$$

The studentized order statistics are given as

$$z_{(i)} = \frac{q_{(i)} - \bar{q}}{\hat{\sigma}}, \quad i = 1, 2, \ldots, N \tag{4}$$

where $\bar{q}$ is the sample mean and $\hat{\sigma}$ is the sample standard deviation. The extended Ozturk algorithm (EOA), first introduced in [25], forms the expected points in the library by

finding the expected value of the summation of the magnitude of studentized order statistics that have been multiplied by a weighting function. In other words, the point associated with distribution j with shape parameter ν is found as

$$X_{\text{EOA},j}(\nu) = \text{E}\left[\sum_{i=1}^{N} w_i |z_{j,(i)}(\nu)|\right] \quad (5)$$

where $|\bullet|$ denotes absolute value and w_i is some weighting function.

Here we consider forming a library with the weighted sum of order statistics (WSOS), found similarly as

$$X_{\text{WSOS},j}(\nu) = \text{E}\left[\sum_{i=1}^{N} w_i q_{j,(i)}(\nu)\right]. \quad (6)$$

In other words, we remove the studentization and the absolute value operations and only compare weightings of the raw order statistics. In Section VI we compare and contrast the two approaches.

Each point in the library is calculated offline, allowing a threshold to also be calculated offline and associated with each point. The radar may then form the same test statistic from measured data and find the closest point in the library matching the sample test statistic. The associated threshold is then used as a detection threshold for the cell under test.

Upon examination of (5) and (6), it is clear that proper selection of weighting functions is crucial to the performance of the algorithm. In [25] sets of weights were used to form a multidimensional search space. The weighting functions were selected from trigonometric and hyperbolic functions, as well as the squares of those functions. Here we examine two weighting functions in a single-dimension space. Similar to the original Ozturk algorithm [31]–[33], we consider the sine and cosine weighting functions, uniformly parametrized on the open set $(0, \pi)$. These functions were selected due to their orthogonal nature.

By using weighted sums of order statistics, the weighting functions serve to enhance or suppress segments of the pdf. For example, Figure 2 shows the values of the sine and cosine weights for a set of $N = 64$. Notice that the sine provides more emphasis to ordered points near the median values, while suppressing values near the extremes. In contrast, the cosine provides positive emphasis to the minimum values, negative emphasis to the maximum values, and de-emphasizes the values close to the median.

The WSOS and EOA provide a method to pre-compute a lookup table of thresholds paired with SIRV distributions that is low complexity. The primary computational expense comes from computing the inverse of the covariance matrix, which must be performed regardless of algorithm choice. In addition, the order statistic approach induces structure on the sample points that are used to enable effective operation in low sample support regimes. For example, it was found that the original Ozturk algorithm was robust to sample covariance matrix estimation error [17]. The points in the library do depend on the dimensionality of the SIRV (and therefore the
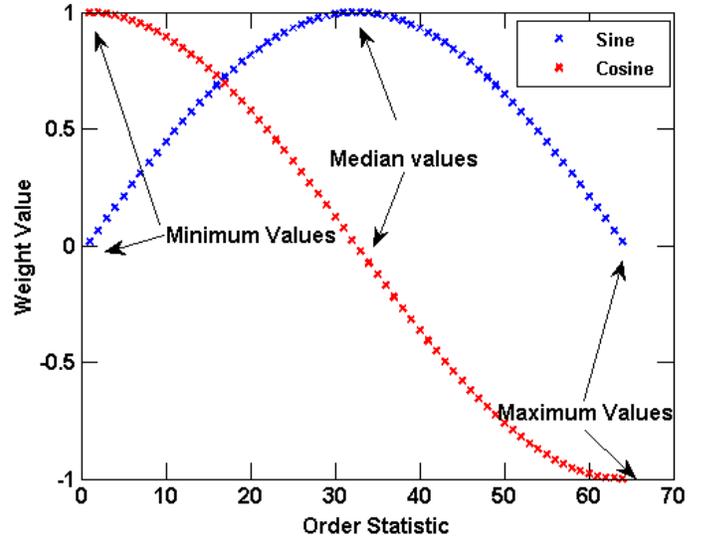

Fig. 2.   Weight Values for Sine and Cosine Weightings

number of pulses in a CPI) as well as the number of range cells used to provide the order statistics. The number of range cells is assumed to be equal to the number of cells that would be used to estimate the covariance matrix.

Comparing the threshold estimation accuracy of the WSOS and EOA methods allows one to consider the value of the studentization/magnitude operations that encapsulate the difference between the EOA method and the new WSOS method. In addition, we also consider a new method of combining the weighting functions. In [25] the combination of multiple endpoints in a multidimensional search space was considered. Here we examine the combination of the endpoints after threshold has been estimated by taking the average of the estimates given by the sine and cosine weighting functions.

VI. SIMULATION RESULTS

To examine the EOA and WSOS methods, a library was constructed with K distributed data with shape parameters $1.3 \le \nu \le 100$, Weibull distributed data with shape parameters $1.05 \le \nu \le 1.95$, Pareto distributed data with shape parameters $3.2 \le \nu \le 40$, and Gaussian distributed data. The lower values of each shape parameter was chosen to give a detection threshold approximately 10 dB greater than the threshold needed for the Gaussian distribution with equal clutter power. The SIRV dimensionality was $L = 16$, and the number of order statistics used was $N = 4L = 64$. The thresholds were estimated from 10^7 Monte Carlo runs with a desired $P_{\text{fa}} = 10^{-5}$. Each point in the library was found from 10^5 Monte Carlo runs.

A. Threshold Estimation Accuracy

Figures 3 and 4 show the average threshold error in dB for both the WSOS and EOA algorithms. The WSOS results are denoted as solid lines, while the EOA results are dashed lines. All threshold errors were found via 10^4 Monte Carlo runs. Note that the 0 dB line corresponds to a correct threshold
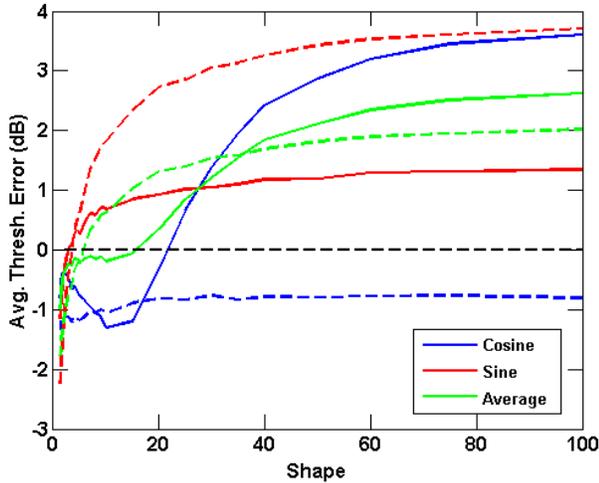
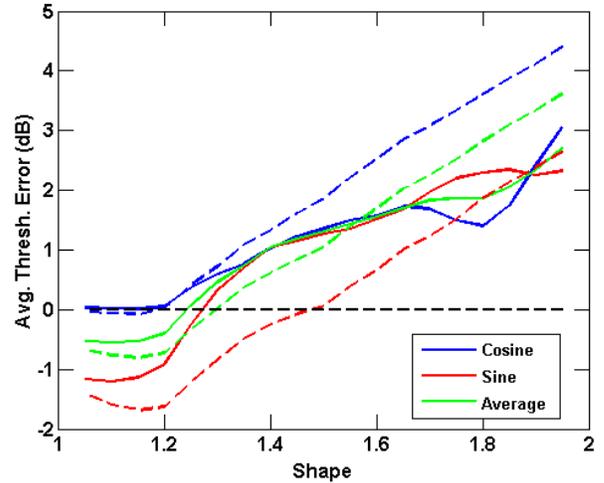Fig. 3. Threshold estimation accuracy for K data, WSOS (solid) and EOA (dashed) v. shape parameter

Fig. 4. Threshold estimation accuracy for Weibull data, WSOS (solid) and EOA (dashed) v. shape parameter

TABLE I

THRESHOLD ESTIMATION FOR GAUSSIAN DISTRIBUTED DATA

Weighting	WSOS	EOA
Cosine	0.22 dB	2.21 dB
Sine	2.99 dB	7.14 dB
Average	1.82 dB	5.34 dB

estimate for each value of the shape parameter, but the true threshold varies greatly by shape parameter. The threshold at the smallest shape parameter is $\approx 7 - 9$ dB greater than the threshold at the largest shape parameter for both distributions. Any estimated threshold below the 0 dB line is associated with an increase in false alarm rate, while any estimated threshold above the line corresponds to a detection loss. A total of six combinations are considered. First, the endpoints given by the sine and cosine weighting functions are used to estimate the threshold. As the sine and cosine are orthogonal functions, we considered the combination of the two weightings in a way different than has previously been reported. For each Monte Carlo, the threshold reported by the sine and the cosine test statistics were averaged to yield a combined threshold estimate. The average of this new estimate is denoted by the green lines. Finally, each of these tests is performed for both the WSOS and EOA transformation methods.

Figure 3 shows average threshold error when the test data is K distributed. For the WSOS method, the cosine weighting exhibits a bias to a greater threshold when compared to the sine weighting function. However, the average of the two threshold estimates is rather accurate for low to mid values of the shape parameter, with an accuracy of ± 1 dB for $1.3 \leq \nu < 30$. The average does suffer a loss of 2.6 dB for $\nu = 100$. In contrast, when the average of the two weighting functions is used in conjunction with the EOA method the estimated threshold is only better than that given by the WSOS method at very high shape parameters. At the extreme of $\nu = 100$ the EOA average yields a detection loss ≈ 0.6 dB less than the WSOS average. Note that even at the lowest shape parameter value, the average estimated threshold is still 7 dB greater than the threshold would be if the clutter was assumed to be Gaussian. Therefore, the improved threshold estimate will lead to a much lower false alarm rate in the face of low shape parameter K distributed clutter.

For the Weibull distribution, shown in Figure 4, the WSOS method outperforms the EOA method. In general, the cosine weighting provides the best results, albeit with increasing detection loss as the tail becomes small.

For many radar detection scenarios, the Gaussian distribution should be considered the default distribution. Therefore, the performance of non-Gaussian oriented techniques should still be tested with Gaussian data. The detection loss associated with using the combinations of weighting function and transformation methods are summarized in Table I. Note that there is no point in the library associated with a distribution with a lighter tail than the Gaussian. Therefore, any selection of a threshold associated with a non-Gaussian distribution can only bias the overall estimate up, yielding an average detection loss. In general, the WSOS method yields more accurate thresholds in the presence of Gaussian clutter. Therefore, of the considered weighting functions, the EOA method should only be used in conjunction with the cosine weighting unless the cognitive radar has prior knowledge to indicate the likely presence of non-Gaussian clutter.

A key capability of a cognitive radar is to estimate thresholds accurately in an unfamiliar environment. To test this scenario, the points in the library associated with the K distribution were removed. The same threshold estimation was then performed with K distributed test data. The average threshold error is shown in Figure 5. Note that with the exception of the cosine weighting, the EOA method suffers little to no estimation loss compared to Figure 3 despite the removal of the distribution under test.
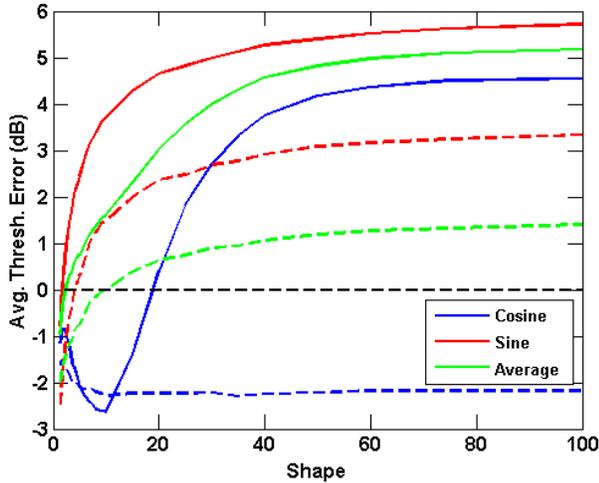
Fig. 5. Threshold estimation accuracy for K data without K in library, WSOS (solid) and EOA (dashed) v. shape parameter
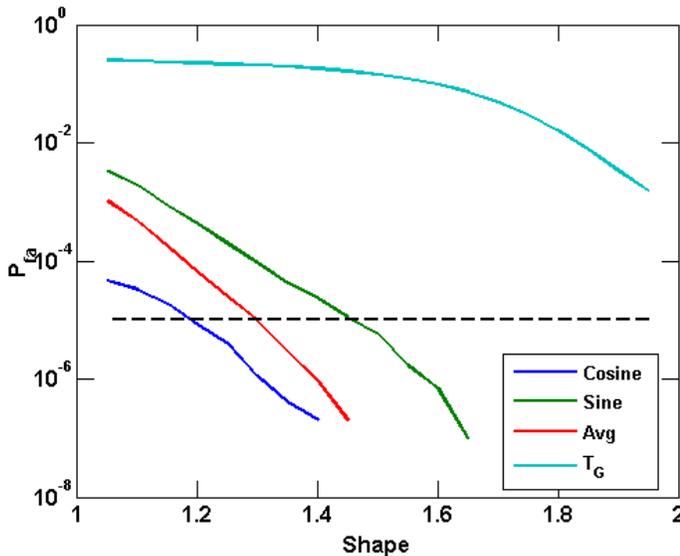

Fig. 6. P_{fa} analysis for Weibull data using EOA

Finally, the impact of the threshold estimation on the probability of false alarm is examined. The average threshold estimates associated with the EOA method were employed with simulated Weibull data. The false alarms over 10^7 Monte Carlo runs were estimated and the results shown in Figure 6. Note that the false alarm rate associated with $\mathcal{T}_{\mathrm{G}}$ is the expected false alarm rate if the optimal Gaussian threshold was used in the presence of Weibull clutter. The desired P_{fa} of 10^{-5} is shown by the dashed black line.

VII. Conclusion

The requirements for cognitive radar detection were considered. A low-complexity, robust approach was implemented to estimate detection thresholds adaptively in non-Gaussian clutter with low sample support. True to the needs of a

cognitive radar, the method was successfully applied when the true distribution of the data was unknown to the estimation algorithm. Ongoing work is also considering the impact of covariance estimation.

References

[1] C. Baker, G. Smith, A. Balleri, M. Holderied, and H. Griffiths, "Biomimetic echolocation with application to radar and sonar sensing," *Proceedings of the IEEE*, vol. 102, no. 4, pp. 447–458, April 2014.

[2] J. Simmons and J. Gaudette, "Special section on biologically-inspired radar and sonar systems - biosonar echo processing by frequency-modulated bats," *Radar, Sonar Navigation, IET*, vol. 6, no. 6, pp. 556–565, July 2012.

[3] S. Haykin, "Cognitive radar: a way of the future," *Signal Processing Magazine, IEEE*, vol. 23, no. 1, pp. 30–40, 2006.

[4] S. Haykin, Y. Xue, and P. Setoodeh, "Cognitive radar: Step toward bridging the gap between neuroscience and engineering," *Proceedings of the IEEE*, vol. 100, no. 11, pp. 3102–3130, 2012.

[5] P. Setlur and M. Rangaswamy, "Signal dependent clutter waveform design for radar stap," in *Radar Conference, 2014 IEEE*, May 2014, pp. 1311–1316.

[6] K. Bell, C. Baker, G. Smith, J. Johnson, and M. Rangaswamy, "Fully adaptive radar for target tracking part i: Single target tracking," in *Radar Conference, 2014 IEEE*, May 2014, pp. 0303–0308.

[7] ——, "Fully adaptive radar for target tracking part ii: Target detection and track initiation," in *Radar Conference, 2014 IEEE*, May 2014, pp. 0309–0314.

[8] H. Griffiths and C. Baker, "Towards the intelligent adaptive radar network," in *Radar Conference (RADAR), 2013 IEEE*, April 2013, pp. 1–5.

[9] S. Blunt, P. McCormick, T. Higgins, and M. Rangaswamy, "Spatially-modulated radar waveforms inspired by fixational eye movement," in *Radar Conference, 2014 IEEE*, May 2014, pp. 0900–0905.

[10] A. Amiri and S. Haykin, "Improved sparse coding under the influence of perceptual attention," in *Neural Computation*, Feb 2014.

[11] S. Haykin, "Cognitive radar networks," in *Computational Advances in Multi-Sensor Adaptive Processing, 2005 1st IEEE International Workshop on*, Dec 2005, pp. 1–3.

[12] K. Sangston and K. Gerlach, "Coherent detection of radar targets in a non-gaussian background," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 30, no. 2, pp. 330 –340, apr 1994.

[13] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Prentice Hall PTR, 1998.

[14] J. Guerci and E. Baranoski, "Knowledge-aided adaptive radar at darpa: an overview," *Signal Processing Magazine, IEEE*, vol. 23, no. 1, pp. 41 – 50, jan. 2006.

[15] J. Bergin and P. Techau, "High-fidelity site-specific radar simulation: Kassper '02 workshop datacube," ISL, Tech. Rep. ISLSCRD-TR-02-105, 2002.

[16] M. Rangaswamy, D. Weiner, and A. Ozturk, "Non-gaussian random vector identification using spherically invariant random processes," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 29, no. 1, pp. 111 –124, jan 1993.

[17] M. Rangaswamy, P. Chakravarthi, D. Weiner, L. Cai, H. Wang, and A. Ozturk, "Signal detection in correlated gaussian and non-gaussian radar clutter," Rome Laboratory, Tech. Rep. 93-17391, 1993.

[18] E. Conte and M. Longo, "Characterisation of radar clutter as a spherically invariant random process," *Communications, Radar and Signal Processing, IEE Proceedings F*, vol. 134, no. 2, pp. 191 –197, april 1987.

[19] E. Conte, A. De Maio, and C. Galdi, "Statistical analysis of real clutter at different range resolutions," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 40, no. 3, pp. 903 – 918, july 2004.

[20] ——, "Statistical analysis of real clutter at different range resolutions," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 40, no. 3, pp. 903 – 918, july 2004.

[21] M. Farshchian and F. Posner, "The pareto distribution for low grazing angle and high resolution x-band sea clutter," in *Radar Conference, 2010 IEEE*, May 2010, pp. 789–793.

[22] G. Weinberg, "Assessing pareto fit to high-resolution high-grazing-angle sea clutter," *Electronics Letters*, vol. 47, no. 8, pp. 516–517, April 2011.

[23] T. Barnard and F. Khan, "Statistical normalization of spherically invariant non-gaussian clutter," *Oceanic Engineering, IEEE Journal of*, vol. 29, no. 2, pp. 303–309, April 2004.

[24] G. Weinberg, "Constant false alarm rate detectors for pareto clutter models," *Radar, Sonar Navigation, IET*, vol. 7, no. 2, pp. 153–163, February 2013.

[25] J. Metcalf, S. Blunt, and B. Himed, "A machine learning approach to distribution identification in non-gaussian clutter," in *Radar Conference, 2014 IEEE*, May 2014, pp. 0739–0744.

[26] A. Zoubir, V. Koivunen, Y. Chakhchoukh, and M. Muma, "Robust estimation in signal processing: A tutorial-style treatment of fundamental concepts," *Signal Processing Magazine, IEEE*, vol. 29, no. 4, pp. 61–80, July.

[27] M. Rangaswamy, "Statistical analysis of the nonhomogeneity detector for non-gaussian interference backgrounds," *Signal Processing, IEEE Transactions on*, vol. 53, no. 6, pp. 2101 – 2111, june 2005.

[28] F. Pascal, Y. Chitour, J. Ovarlez, P. Forster, and P. Larzabal, "Covariance structure maximum-likelihood estimates in compound gaussian noise: Existence and algorithm analysis," *Signal Processing, IEEE Transactions on*, vol. 56, no. 1, pp. 34–48, Jan 2008.

[29] E. Conte, A. De Maio, and G. Ricci, "Recursive estimation of the covariance matrix of a compound-gaussian process and its application to adaptive cfar detection," *Signal Processing, IEEE Transactions on*, vol. 50, no. 8, pp. 1908–1915, Aug 2002.

[30] M. Greco, P. Stinco, F. Gini, and M. Rangaswamy, "Impact of sea clutter nonstationarity on disturbance covariance matrix estimation and cfar detector performance," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 46, no. 3, pp. 1502 –1513, july 2010.

[31] A. Ozturk, "A general algorithm for univariate and multivariate goodness-of-fit tests based on graphical representation," *Communications in statistics - theory and methods*, vol. 20, no. 10, pp. 3111 –3137, 1991.

[32] A. Ozturk and E. Dudewicz, "A new statistical goodness-of-fit test based on graphical representation," *Biometrical Journal*, vol. 34, no. 4, pp. 403 –427, 1992.

[33] A. Ozturk, "An application of a distribution identification algorithm to signal detection problems," in *Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on*, nov 1993, pp. 248 –252 vol.1.