Target Detection Scheme with Optimal Inter-Channel Noncoherent Data Combining

Yuri I. Abramovich
WR Systems
Fairfax, VA 22030 USA

Geoffrey San Antonio
US Naval Research Laboratory
Washington, DC 20375 USA

***Abstract*—This paper introduces a novel non-linear processing technique for reducing the dimensions of a data set without performing hard thresholding while maintaining the detector performance as applied to the original large data set. In particular, the introduced processing technique can be utilized in spatial beamforming, high-range resolution processing, and other scenarios wherein high resolution data must be compressed/compacted to accommodate a lower resolution display and/or processing system.**

***Index Terms*—Beamforming, extreme statistics**

I. Introduction

In a number of radar systems, there still exists the requirement for the radar human operator to remain capable of primary detection, track initialization and maintenance, despite sophisticated automated detection and tracking routines having been implemented. In most cases though, it is impossible for a human operator to process data in the time-frame of the entire radar data flow due to the high resolution in range, Doppler frequency, and azimuth. If for example, visual control is required over a certain region in range, azimuth, and radial velocity, the number of processing resolution cells presented to a human operator must be significantly less than the total number of cells generated by the radar processor given the full radar waveform bandwidth, coherent integration time (CIT), and receive aperture.

The naive approach of simply reducing the radar bandwidth, CIT, or utilized antenna aperture, leads to a significant signal-to-noise ratio (SNR) degradation even in the presence of internal noise only. In the presence of strong clutter returns and/or localized interference, the detection losses might be much more severe. Therefore, more advanced signal processing schemes have to be considered that provide an interface between the high fidelity radar processor and the low resolution human operator (display). In most cases where the requirement on human primary detection capability is enforced, no automated data flow reduction schemes, such as primary detection, are usually acceptable. Under these conditions, the signals at the output of a radar processor multiple resolution cells should be somehow aggregated (combined) into a smaller number of resolution cells, with minimal losses in target detectability.

One such non-coherent processing technique is the selection over the required number of primary resolution cells, of the maximal signal at the output of a standard envelope detector or non-coherent integrator. The rationale behind such a scheme is straightforward. The target signal power within a primary resolution cell has to exceed the power of each of the other $(N-1)$ resolution cells occupied by noise for reliable detection. However, the maximum power over $(N-1)$ resolution cells is much less than the power of their sum. Moreover, when selection occurs over a considerable number of i.i.d. values, the maximum value is much more stable and does not fluctuate as much as the exponentially distributed random values at the output of a standard square law envelope detector. Both these factors suggest that the receiver operating characteristics (ROC) might be superior than for the naive implementation of reduced resolution (lower bandwidth, lower CIT, smaller aperture).

A problem may occur in some degenerate scenarios such as the case when multiple, say P, targets share the same resolution cells in all dimensions (range and Doppler) except the one being compressed (azimuth). Instead of P targets being properly resolved in azimuth, this processing scheme will produce the single (strongest) peak. However such a scenario is not typical for radars with high resolution in range and/or Doppler frequency. Moreover, the naive alternative technique mentioned above suffers from the same problem potentially. Strictly speaking, the linear superposition of several peaks with arbitrary phases may lead to very strong fluctuations of this "synthetic" target response, with a high probability of no target being detected.

In this paper, we analyze the detection losses associated with such "maximum selection" combiner technique. Statistical analysis conducted in section II, is performed under the assumption that the noise values in the cells compared for maximum selection are independent. In this case, analytical expressions could be derived and used for validation of Monte-Carlo simulations. These simulations are the only reliable tool to assess detection performance in the case of correlated noise in compared data. This latter scenario is very important in practical applications. Indeed, in the case of a pure white noise environment, the noise samples at the output of a beamformer are independent across beams only if the beam steering vector themselves are independent. Such is the case for a linear array and an un-tapered DFT beamformer with the standard 3dB beam overlap. However as any practical engineer can attest, 3dB overlapped beams can cause significant scalloping loss, and therefore 0.5-1dB overlap is more typical. Such "oversampling" leads to residual inter-beam noise correlation.

When adaptive beamforming is used in spatially non-white external noise, residual correlation may exist even for 3dB beams. Unfortunately, the reliable statistical description for the distribution of the maximum over N correlated (exponentially distributed) samples [3], does not exist, and one has to rely upon simulations. Results of this analysis are presented in section III. In section IV we conclude the paper.

II. Detection Performance of the "Maximum Selection" Combiner: Independent Samples

Let us assume that N independent noise samples x_j with distribution $f_j(x),\ x \geq 0$ are compared and the sample with maximal value $X_{max} = \max(x_1, x_2, \ldots, x_N)$ is selected. Then the probability that

$$Pr(x_{\max} < \lambda) = Pr\{x_j < \lambda,\ j = 1, \ldots, N\} \tag{1}$$

i.e. the probability that the maximum value $x_{\max}$ does not exceed λ is equal to the probability that none of the samples $x_j,\ j = 1, \ldots, N$ exceeds that threshold. Since the samples are independent

$$Pr(x_{\max} < \lambda) = \prod_{j=1}^{N} F_j(\lambda) \tag{2}$$

$$F_j(\lambda) = \int_0^{\lambda} f_j(x)dx. \tag{3}$$

One can imagine that the samples x_j might not be identically distributed for this analysis. In the case that we do deal with independent and identically distributed samples, then the cumulative distribution function (CDF) of the maximum may be analytically expressed as

$$Pr(x_{\max} < \lambda) \equiv F_N(\lambda) = F_1^N(\lambda), \tag{4}$$

where

$$F_1(\lambda) = \int_0^{\lambda} f(x)dx \tag{5}$$

is the individual sample CDF. Note, the probability of false alarm (P_{FA}) can in this case be expressed analytically as well as

$$P_{FA}^N \equiv Pr(x_{\max} > \lambda) = 1 - F_N(\lambda) = 1 - F_1^N(\lambda). \tag{6}$$

In typical radar applications the P_{FA} is set much less than 1, $P_{FA} \approx 10^{-3} - 10^{-7}$, and therefore $F_1^N(\lambda) \approx 1 - (10^{-3} - 10^{-7})$. For this typical low P_{FA}'s it is true that

$$F_1(\lambda) = 1 - P_{FA}^1(\lambda) \tag{7}$$

where $P_{FA}^1(\lambda)$ is the P_{FA} for a single sample. Since for $P_{FA}^1 \approx 10^{-3} - 10^{-7}$

$$(1 - P_{FA}^1(\lambda))^N \approx 1 - N * P_{FA}^1, \tag{8}$$

we get,

$$P_{FA}^N \approx N * P_{FA}^1. \tag{9}$$

The result in equation (9) means that the probability of false alarm at the output of the "maximum selection" combiner is approximately N-times greater than the probability of false alarm within a single sample. For example, in the case of $N = 10$ and $P_{FA}^N = 10^{-3}$, the P_{FA} within a single sample should be $P_{FA}^1 = 10^{-4}$. When the P_{FA} of the aggregated "maximum selection" output (10 orthogonal beams for example) should be the same as per an individual sample (finger beam), than the detection losses compared with the single sample are associated with the higher threshold for $P_{FA} = 10^{-4}$ instead of $P_{FA} = 10^{-3}$. The actual SNR loss factor associated with the threshold set for $P_{FA} = 10^{-4}$ instead of $P_{FA} = 10^{-3}$ is analyzed in the next section. It is clear though that the detection loss should be much less than the normal 10dB loss associated with intentional radar processing degradation such as a 10-fold aperture reduction. Alternatively, if one considers that the P_{FA} within an entire area of regard covered by all $N = 10$ samples (beams) has to be retained at the same level of $P_{FA}^N = 10^{-3}$ as per the aggregated beam, than the threshold within each individual beam has to be set for $P_{FA}^1 = 10^{-4}$. This consideration means that if the total P_{FA} within the entire area of regard is controlled, the proposed "maximum selection" combiner is equivalent to the optimal multi-beam processing in terms of the probability of detection.

It is important that the distribution of the maximum of N i.i.d. positive random values has a shape that is dependent on N. For this reason, the power (mean value) of the maximum random variable does not entirely characterize the probability of detection and does not characterize the merits of the two detection schemes if compared with the noise power at the output of the beam with shortened aperture.

Now consider the standard case when the noise data at every range/Doppler resolution cell within each beam is a complex Gaussian random variable $CN(0, \sigma^2)$, so that at the output of a square law detector, the distribution is exponential (χ^2).

$$w(x) = \frac{1}{2\sigma^2} e^{-\frac{x}{2\sigma^2}} \tag{10}$$

Then, the distribution of the maximum over N i.i.d. random variables distributed as equation (10) is

$$w_N(x) = N * F_1^{N-1} w(x), \tag{11}$$

where

$$F_1(x) = \frac{1}{2\sigma^2} \int_0^x e^{-\frac{t}{2\sigma^2}} dt = [1 - e^{-\frac{x}{2\sigma^2}}]. \tag{12}$$

Therefore,

$$w_N(x) = \frac{N}{2\sigma^2} \sum_{j=0}^{N-1} (-1)^j C_{N-1}^j e^{-\frac{(1+j)x}{2\sigma^2}} \tag{13}$$

and

$$E_N[x^k] = (2\sigma^2)^k \Gamma(k+1) N \sum_{j=0}^{N-1} \frac{(-1)^j C_{N-1}^j}{(1+j)^K + 1}. \tag{14}$$

For $K = 1$ and $N = \{1, 2, 3\}$ we get: $E_1[x^1] = 2\sigma^2$; $E_2[x^1] = 3\sigma^2$; $E_3[x^1] = \frac{11}{3}\sigma^2$. This demonstrates that the first moment is not linearly dependent on N, unlike the naive resolution spoiling scheme.

III. Detection Performance of "Maximum Selection" Combining for Correlated Data

Unfortunately, for the scenario with correlated (dependent) non-Gaussian data, such as the modulus values of the correlated complex Gaussian random variables, sufficiently accurate analytical results do not exist. An assumption on independence of this data provides the upper bound for the CDF.

$$Pr(M_N > A) \leq N(1 - F(A)) \tag{15}$$

where M_N is the maximum over N real-valued dependent identically distributed variables with continuous distribution function F. Moreover, according to the theorem 2.1 [1], for each N there exists an exponential variable E with

$$-\log N - \log(1 - M_N) < E. \tag{16}$$

Note that in the case of independent random variables according to the theorem 2.2 in [1] the function

$$G(M_N) = -\log(-N \log F(M_N)) \tag{17}$$

has a Gumbel distribution and

$$G \leq -\log(N(1 - F(M_N))) < G + \exp(-G/N). \tag{18}$$

It follows that the limiting distribution of $-\log N - \log(1 - F(M_N))$ is the Gumbel distribution. Moreover, according to [1], E and G are very close in their tail distributions, so there is not much difference between the upper bounds in the independent and dependent cases. In practical applications, the dependence of the noise samples is due to the inter-beam (angular) separation being smaller than the one that warrants strict independence. Note, that in linear uniform arrays and spatially white noise, strict independence takes place for untapered DFT-based beamforming with -3db beam cross-points. None of these conditions are achieved in practice and some level of inter-beam noise correlation is always present. It is clear that if motivated by Nyquist theorem considerations, one can consider a number of independent beams M, with $M < N$, that cover the same area of regard as the N overlapped practical beams. Below we first introduce the results of Monte-Carlo simulations for independent noise (orthogonal beams) and then compare these results with the results of overlapped beams.

We begin by comparing the theoretical PDF's and CDF's derived from the i.i.d. Gaussian noise case using $M = 10$ channels, with the Monte-Carlo results using 10^7 trials. Absolutely identical (figure 1) theoretical and empirical results suggest that we may rely upon the Monte-Carlo simulations for more complicated settings. CDF's and PDF's for cases covering $M = 1 - 100$ are illustrated in (figure 2) which demonstrate strong deformation of both the CDF and PDF with M that leaves the "tail" of the the distribution much less affected in contrast to distribution first moment. This property is more clearly demonstrated by (figure 3(d)), where for different M, we show how the threshold must vary to achieve a certain P_{FA}. One can see that for a $P_{FA} = 10^{-3}$ and $M = 10$, the threshold level has to be set at 9.7dB (noise power is 0dB), while for a single channel ($M = 1$), the same P_{FA} is achieved with only a slightly lower threshold of 8.4dB, or 1.3dB lower. At the same time, if the total P_{FA} over all independent beams $M = 10$ has to be set at the same level of $P_{FA} = 10^{-3}$, the individual P_{FA} within each beam has to be set at the level of $P_{FA} = 10^{-4}$, which for $M = 1$ requires exactly the same threshold of 9.7dB. Therefore, depending on the metric of our comparison, the SNR loss associated with "maximum selection" varies from 1.3dB to 0dB for the same probability of detection and P_{FA}.

In (figure 3(a-c)) we illustrate the dependence on M of the first moment m_1, second moment m_2, and variance $v^2 = (m_2 - m_1^2)$. One can see that while m_1 and m_2 grow with the number of channels M, the variance v^2 stays practically constant $v^2 \approx 2$, which means that m_1/v^2 grows with M, resulting in only 1.3dB threshold growth for $P_{FA} = 10^{-3}$, despite significant growth of the mean power value (4.66dB). Overall, for i.i.d. noise samples, Monte-Carlo results are in full agreement with the theoretical predictions.

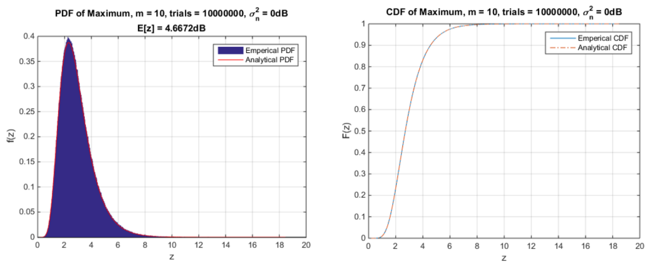
In order to study the effect of correlated noise / dependent beam outputs on the "maximum selection" combiner performance we consider the construction of an oversampled beam space construction. We consider an N element uniform linear array with element spacing $d = \lambda/2$. We may then consider the oversampled beams

$$\mathbf{u}_k = \left[1, e^{j2\pi u(k)}, \ldots, e^{j2\pi(N-1)u(k)}\right]^T / \sqrt{N} \tag{19}$$

$$u(k) = du(k-1)/N, \; 0 < du < 1 \tag{20}$$

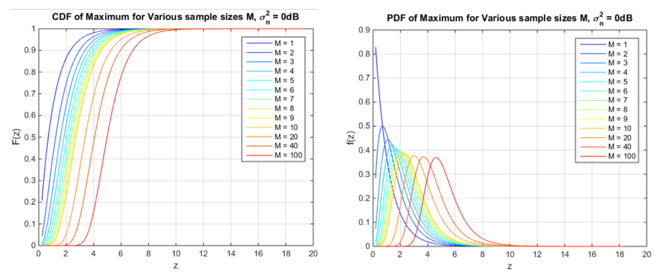
and call the ratio between the number of oversampled and independent beams that cover the same angular sector the oversampling rate f_s. We need to evaluate the impact of the oversampling rate on detection performance of the "maximum selection" combiner.

First we demonstrate that for orthogonal beams with white input noise, the output noise over the beams is uncorrelated indeed. In (figure 4(a,b)) we show results calculated for $M = 10$ i.i.d. complex Gaussian values and for $M = 10$ orthogonal beams with the noise values at the beam output analyzed. No difference between the two is observed. In figure 4(c) we introduce the sample PDF of the maximum values calculated over $M = 164$ strongly oversampled beams that cover the same azimuthal sector as the $M = 10$ orthogonal beam case. The oversampling rate is equal to 16.4, while the inter-beam correlation is equal to 0.99. This is obvious an extreme example with a practically unrealistic oversampling rate. Despite such an extreme oversampling rate, the sample PDF is only slightly shifted to the right with respect to the theoretical PDF for $M - 10$ orthogonal beams. Obviously the theoretical distribution for $M = 162$ independent (orthogonal) samples is shifted significantly further away to the right. This means that for a significant oversampling rate, the "Nyquist" number of



(a) (b)

Fig. 1. Analytic vs. Empirical (a) PDF and (b) CDF, independent case.



(a) (b)

Fig. 2. Analytic (a) PDF and (b) CDF, independent case for $M = 1 - 100$.

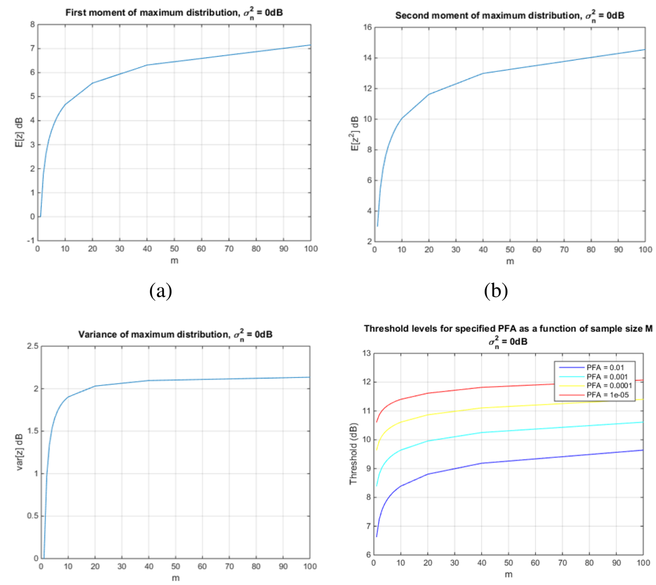
beams ($M = 10$ in this case) provides a much tighter lower bound, than the upper bound derived under the assumption that all M (oversampled) beams are in fact independent. In figure 4(d) we show the variation in the required threshold value as a function of beam oversampling rate to each achieve a $P_{FA} = 10^{-3}$. This shows that required threshold increase due to the use of oversampled beams reaches an asymptotic limit resulting in approximately 0.5dB worst case threshold increase in comparison to completely independent beams. For comparison we also show what the threshold would be if each of the correlated beams were incorrectly assumed to be independent.

IV. Conclusion

The "maximum selection" combiner that selects the range-Doppler resolution cell with the maximal amplitude (power) over a number (M) of finger-beams that cover the required area of angular coverage, may be treated as the optimum combiner. When normalized for the same total probability of false alarm, it provides the same probability of detection as the individually treated M finger-beams that cover the same area of regard. Even when the area of regard is $M = 10$ times broader than a single finger beam coverage, the SNR losses with respect to the single finger beam optimum processing do not exceed 1.8dB with an arbitrary large number of individual finger beams that populate this area of regard. In the case of $M = 10$ independent beams that cover the same area, the SNR losses do not exceed 1.3dB, which means that noise correlation within closely separated beams does not incur more then 0.5dB in additional SNR losses.

Acknowledgment

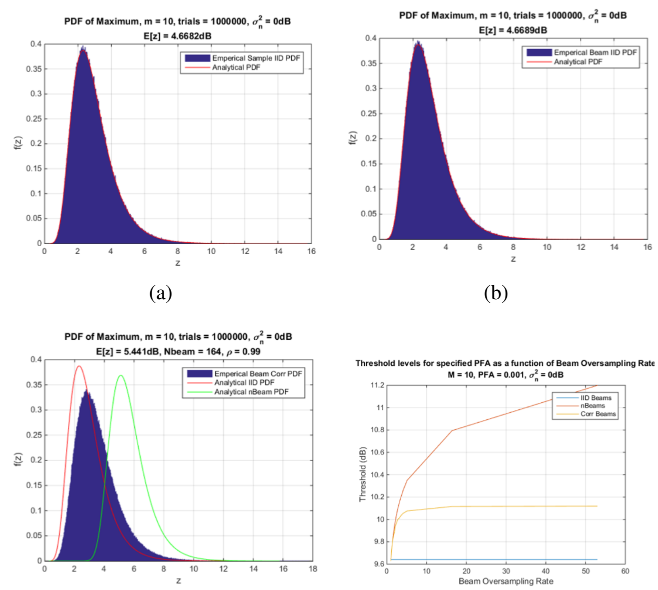
This work was sponsored by the Office of Naval Research under an NRL 6.1 Base Program



(a) (b)

(c) (d)

Fig. 3. Variation of distribution moments with M: (a) $E[z]$, (b) $E[z^2]$, (c) $var[z]$, (d) Variation of P_{FA} with threshold and M, independent case.



(a) (b)

(c) (d)

Fig. 4. Behavior of "maximum selection" distribution with correlated samples M, (a) independent samples, (b) independent beams, (c) correlated beams, (d) correlated beam threshold behavior.

References

[1] E. J. Gumbel, Statistics of Extremes, Echo Point Books and Media, Brattleboro, VT, 1958.

[2] Harry. L. Van Trees, Detection, Estimation, and Modulation Theory: Part I, Wiley Interscience, New York, NY, 2001 edition.

[3] Ranjan K. Mallik, "On Multivariate Rayleigh and Exponential Distributions," *IEEE Trans. Info. Theory*, Vol. 69, No. 6, Pp. 1499-1515, June, 2003.