Optimal Correlators and Waveforms for Mismatched Detection

Neri Merhav

The Andrew & Erna Viterbi Faculty of Electrical and Computer Engineering
Technion - Israel Institute of Technology
Technion City, Haifa 32000, ISRAEL
E–mail: `merhav@ee.technion.ac.il`

Abstract

We consider the classical Neymann–Pearson hypothesis testing problem of signal detection, where under the null hypothesis ($\mathcal{H}_0$), the received signal is white Gaussian noise, and under the alternative hypothesis ($\mathcal{H}_1$), the received signal includes also an additional non–Gaussian random signal, which in turn can be viewed as a deterministic waveform plus zero–mean, non-Gaussian noise. However, instead of the classical likelihood ratio test detector, which might be difficult to implement, in general, we impose a (mismatched) correlation detector, which is relatively easy to implement, and we characterize the optimal correlator weights in the sense of the best trade-off between the false-alarm error exponent and the missed-detection error exponent. Those optimal correlator weights depend (non-linearly, in general) on the underlying deterministic waveform under $\mathcal{H}_1$. We then assume that the deterministic waveform may also be free to be optimized (subject to a power constraint), jointly with the correlator, and show that both the optimal waveform and the optimal correlator weights may take on values in a small finite set of typically no more than two to four levels, depending on the distribution of the non-Gaussian noise component. Finally, we outline an extension of the scope to a wider class of detectors that are based on linear combinations of the correlation and the energy of the received signal.

Index terms: hypothesis testing, signal detection, correlation–detection, error exponent.

1 Introduction

The topic of detection of signals corrupted by noise has a very long history of active research efforts, as it has an extremely wide spectrum of engineering applications in the areas communications and signal processing. These include radar, sonar, light detection and ranging (LIDAR), object recognition in images and video streams, diagnosis based on biomedical signals, watermark detection in images and audio signals, seismological signal detection related to geophysical activity, and

object detection using multispectral/hyperspectral imaging, just to name a few. One of the most problematic and frequently encountered issues in signal detection scenarios is mismatch between the signal model and the detector design, which is based upon certain assumptions on that model. Accordingly, the topic of mismatched signal detection has received considerable attention in the literature, see, e.g., [1], [10], [11], [18], [19], [20], [21], [28], and [30], for a non-exhaustive list of relevant references. The common theme in most of these works is the possible presence of uncertainties in the desired signal to be detected, in the steering vector, in the transfer function of the propagation medium, and/or in the distributions of the various kinds of noise, interference and clutter. Accordingly, adaptive detection mechanisms with tunable parameters have been developed and proposed in order to combat those types of mismatch.

Another line of earlier relevant research activity is associated with the notion of *robust detection techniques*, where the common theme is generally directed towards a worst-case design of the detector against small non-parametric uncertainties around some nominal noise distribution, most notably, a Gaussian distribution. See, e.g., [2], [4], [5], [9], [12], [13], [15], [16], [17], [22], [24], and [25]. See also [14] for a survey on the subject.

Last but not least, when the uncertainty is only in a finite number of parameters of the model, the problem is normally treated in the framework of composite hypothesis testing, where the popular approach is the well–known generalized likelihood ratio test (GLRT) [27], which is often (but not always) asymptotically optimal in the error exponent sense, see, for example, [3], [6], [7], and [29]. The GLRT is applied also in some of the above cited articles on mismatched detection, among many others. Another approach to composite hypothesis testing is the competitive minimax approach, proposed in [8].

Our objective in this work is partially related to those studies, but it is different. It is associated with mismatched detection, except that the origin of this mismatch is not quite due to uncertainty in the signal-plus-noise model, but it comes from practical considerations: the optimal likelihood ratio test (LRT) detector might be difficult to implement in many application examples, especially in the case of sensors that are built on small, mobile devices which are subjected to severe limitations on power and computational resources. In such situations, it is desirable that the detector would be as simple as possible, e.g., a correlation detector, or a detector that is based on correlation and energy.

Within this framework, the number of arithmetic operations (especially the multiplications) should be made as small as possible. Clearly, a detector from this class cannot be optimal, unless the noise is Gaussian, hence the mismatch. Nonetheless, we would like to find the best correlator weights in the sense of optimizing the trade-off between the false-alarm (FA) and the missed–detection (MD) rates. This would partially compensate for the mismatch in case the noise is not purely Gaussian.

More precisely, consider the following signal detection problem, of distinguishing between two hypotheses:

$$\mathcal{H}_0: \quad Y_t = N_t, \qquad t = 1, 2, \ldots, n \tag{1}$$

$$\mathcal{H}_1: \quad Y_t = X_t + N_t, \qquad t = 1, 2, \ldots, n \tag{2}$$

where $\{N_t\}$ is an independently identically distributed (IID), zero-mean Gaussian noise process with variance σ_N^2, independent of $\{X_t\}$, which is another random process, that we decompose as $X_t = s_t + Z_t$, with $s_t = \boldsymbol{E}\{X_t\}$ being a deterministic waveform and $Z_t = X_t - s_t$ being an IID, zero-mean noise process, which is not necessarily Gaussian in general. The non–Gaussian noise component, $\{Z_t\}$, can be thought of as signal–induced noise (SIN), which may stem from several possible mechanisms, such as: echos of the desired signal, multiplicative noise, cross-talks from parallel channels conveying correlated signals, interference by jammers, and in the case of optical detection using avalanche photo-diodes (APDs), it corresponds to shot noise plus multiplicative noise due to the random gain of the device (see, e.g., [23] and references therein, for more details). In general, $\{Z_t\}$ may also designate randomness that could be attributed to uncertainty associated with the transmitted signal.

As mentioned above, the optimal LRT detector might be considerably difficult to implement in practice since the probability density function (PDF) of $\{Y_t\}$ under $\mathcal{H}_1$ involves the convolution between the Gaussian PDF of N_t and the (non-Gaussian) PDF of Z_t, which is typically complicated. As said, a reasonable practical compromise, valid when the underlying signal $\{s_t\}$ is not identically zero, is a correlation detector, which compares the correlation, $\sum_{t=1}^{n} w_t Y_t$, to a threshold, where $w_1, \ldots, w_n$ are referred to as the *correlator weights*, and the threshold controls the trade-off between the FA probability and the MD probability. Our first objective is to characterize the best correlator weights, $w_1^*, \ldots, w_n^*$, in the sense of the optimal trade-off between the FA probability and the MD probability, or more precisely, the optimal trade-off between the asymptotic exponential rates of

decay of these probabilities as functions of the sample size n, i.e., the *FA exponent* and the *MD exponent*. Clearly, the optimal correlation detector is, in general, not as good as the optimal, LRT detector, but it is the best compromise between performance and the practical implementability in the framework of correlation detectors. As very similar study was already carried out in [23], in the context of optical signal detection using photo-detectors, where the optimal correlator waveform was characterized in terms of the optical transmitted signal in continuous time, and was found to be given by a certain non-linear function of the optical signal.

Here, we study the problem in a more general framework, in the sense that the PDF of the SIN, Z_t, is arbitrary. Moreover, we expand the scope in several directions, in addition to the study that is directly parallel to that of [23].

1. We consider the possibility of limiting the number of levels $\{w_t\}$ to be finite (e.g., binary, ternary, etc.), with the motivation of significantly reducing the number of multiplications needed to calculate the correlation, $\sum_t w_t Y_t$.

2. We jointly optimize both the signal $\{s_t\}$ and the correlator, $\{w_t\}$. Interestingly, here both the optimal signal and the optimal correlation weights turn out to have a *finite number of levels* even if this number is not restricted a-priori. The number of levels depends on the PDF of Z_t, and it is typically very small (e.g., two to four levels). Moreover, the optimal $\{s_t\}$ and $\{w_t\}$ turn out to be proportional to each other, in contrast to the non-linear relation resulting when only $\{w_t\}$ is optimized while $\{s_t\}$ is given.

3. We outline an extension to a wider class of detectors that are based on linear combinations of the correlation, $\sum_t w_t Y_t$, and the energy, $\sum_t Y_t^2$, with the motivation that it is, in fact, the structure of the optimal detector when Z_t is Gaussian noise, and that it is reasonable regardless, since the under $\mathcal{H}_1$ the power (or the variance) of the received signal is larger than under $\mathcal{H}_0$.[1] We also address the possibility of replacing the energy term by the sum of absolute values, $\sum_t |Y_t|$, which is another measure of signal intensity, with the practical advantage that its calculation does not require multiplications.

The outline of the remaining part of this work is as follows. In Section 2, we formalize the

---

[1]In fact, when $s_t \equiv 0$, the correlation term becomes useless altogether and the energy term becomes necessary.

problem rigorously and spell out our basic assumptions. In Section 3, we characterize the optimal correlator, $\{w_t^*\}$, for a given signal, $\{s_t\}$, subject to the power constraint. In Section 4, we address the problem of joint optimization of both $\{w_t\}$ and $\{s_t\}$, both under power constraints, and finally in Section 5, we outline extensions to wider classes of detectors that are based on correlation and energy.

2 Assumptions and Preliminaries

Consider the signal detection model described in the fifth paragraph of the Introduction. We assume that $Z_1, \ldots, Z_n$ are independent copies of a zero–mean random variable (RV), Z, that has a symmetric[2] PDF, $f_Z(z)$, around the origin, and that it has a finite cumulant generating function (CGF),

$$C(v) \stackrel{\triangle}{=} \ln \boldsymbol{E}\{e^{vZ}\}, \tag{3}$$

at least in a certain interval of the real valued variable v. Note that since $f_Z(\cdot)$ is assumed symmetric around the origin, then so is $C(\cdot)$. We also assume that $C(\cdot)$ is twice differentiable within the range it exists. It is well known to be a convex function, because its second derivative cannot be negative, as it can be viewed as the variance of Z under the PDF that is proportional to $f_Z(z)e^{vz}$. Further assumptions on Z and its CGF will be spelled out in the sequel, at the places they will be needed. The following simple special cases will accompany our derivations and discussions repeatedly in the sequel:

Case 1. Z is zero-mean, Gaussian RV with variance σ_Z^2:

$$C(v) = \frac{\sigma_Z^2 v^2}{2}. \tag{4}$$

Case 2. Z is a Laplacian RV with parameter q, i.e., $f_Z(z) = \frac{q}{2}e^{-q|z|}$:

$$C(v) = -\ln\left(1 - \frac{v^2}{q^2}\right). \tag{5}$$

Case 3. Z in a binary RV, taking values in $\{-z_0, +z_0\}$ with equal probabilities:

$$C(v) = \ln\cosh(z_0 v). \tag{6}$$

---

[2]The symmetry assumption is imposed mostly for convenience, but the results can be extended to address also non-symmetric PDFs.

Case 4. Z is a uniformly distributed RV across the interval $[-z_0, +z_0]$:

$$C(v) = \ln\left(\frac{\sinh(z_0 v)}{z_0 v}\right). \tag{7}$$

The signal vector, $\boldsymbol{s} = (s_1, \ldots, s_n)$, $s_t \in \mathbb{R}$, $t = 1, \ldots, n$, is assumed known, and we denote its power by $P(\boldsymbol{s})$, that is,

$$P(\boldsymbol{s}) \triangleq \frac{1}{n}\sum_{t=1}^{n} s_t^2. \tag{8}$$

Consider the class of correlation detectors, i.e., detectors that compare the correlation, $\sum_{t=1}^{n} w_t Y_t$, to a certain threshold T, where $\boldsymbol{w} = (w_1, \ldots, w_n)$ is a vector of real valued correlator coefficients, henceforth referred to as the *correlator*, for short. The decision rule is as follows: If $\sum_{t=1}^{n} w_t Y_t < T$, accept the null hypothesis, $\mathcal{H}_0$, otherwise, accept the alternative, $\mathcal{H}_1$. The threshold, T, controls the trade-off between the FA probability and the MD probability of the detector. To allow exponential decay (as n grows without bound) of both types of error probabilities, we let T vary linearly with n, and denote $T = \theta n$, where θ is a real valued constant, independent of n.

In order to have a well-defined asymptotic FA exponent, we assume that the correlator, $\boldsymbol{w}$, has a fixed power,

$$P(\boldsymbol{w}) = \frac{1}{n}\sum_{t=1}^{n} w_t^2, \tag{9}$$

which is independent of n, or, more generally, that the right–hand side (RHS) of eq. (9) tends to a certain fixed positive power level, as $n \to \infty$ (otherwise, the normalized logarithm of the FA probability would oscillate indefinitely, without a limit). Indeed, the FA probability of the correlation detector is given by

$$P_{\text{FA}} = \Pr\left\{\sum_{t=1}^{n} w_t N_t \geq \theta n\right\} = Q\left(\frac{\theta n}{\sigma\|\boldsymbol{w}\|}\right) \doteq \exp\left\{-\frac{\theta^2 n}{2\sigma_N^2 P(\boldsymbol{w})}\right\}, \tag{10}$$

where Q is the well-known Q-function,

$$Q(u) \triangleq \frac{1}{\sqrt{2\pi}}\int_u^{\infty} e^{-u^2/2}\mathrm{d}u, \tag{11}$$

and $\doteq$ denotes equivalence in the exponential scale, in other words, the notation, $a_n \doteq b_n$, for two positive sequences $\{a_n\}$ and $\{b_n\}$, means that $\lim_{n\to\infty}\frac{1}{n}\log\frac{a_n}{b_n} = 0$. It follows from (10) that the FA exponent is given by

$$E_{\text{FA}}(\theta) = \frac{\theta^2}{2\sigma_N^2 P(\boldsymbol{w})}. \tag{12}$$

Thus, the FA exponent depends on $\boldsymbol{w}$ only via $P(\boldsymbol{w})$. It follows that for a given θ, if we wish to achieve a given, prescribed FA exponent, $E_{\text{FA}}(\theta) \geq E_{\text{FA}}$ (where E_{FA} is a given positive number), we must have

$$P(\boldsymbol{w}) \leq P_w \stackrel{\Delta}{=} \frac{\theta^2}{2\sigma_N^2 E_{\text{FA}}}. \tag{13}$$

In other words, a constraint on the FA exponent means a corresponding constraint on the asymptotic power of $\boldsymbol{w}$ to be no larger than P_w.

In order to have a well–defined MD exponent, our assumptions concerning the asymptotic behavior of $\boldsymbol{w}$ and $\boldsymbol{s}$ will have to be more restrictive: We will assume that as $n \to \infty$, the pairs $\{(w_t, s_t)\}_{t=1}^n$ obey a certain joint PDF, $f_{WS}(w, s)$, in the following sense: For every $\lambda \geq 0$,

$$\lim_{n \to \infty} \left\{ \lambda \left(\cdot \frac{1}{n} w_t s_t - \theta \right) - \frac{1}{n} \sum_{t=1}^n C(\lambda w_t) - \frac{\lambda^2 \sigma_N^2}{2} \cdot \frac{1}{n} \sum_{t=1}^n w_t^2 \right\}$$
$$= \boldsymbol{E}_{WS} \left\{ \lambda (\boldsymbol{E}\{W \cdot S\} - \theta) - \boldsymbol{E}\{C(\lambda W)\} - \frac{\lambda^2 \sigma_N^2}{2} \cdot \boldsymbol{E}\{W^2\} \right\}, \tag{14}$$

where $\boldsymbol{E}_{WS}\{\cdot\}$ denotes expectation with respect to (w.r.t.) f_{WS}. Whenever there is no room for confusion, the subscript WS will be omitted and the expectation will be denoted simply by $\boldsymbol{E}\{\cdot\}$. The function $f_{WS}(\cdot, \cdot)$ will be referred to as the *asymptotic empirical joint PDF* of $\boldsymbol{w}$ and $\boldsymbol{s}$.[3]

The MD probability is now upper bounded, exponentially tightly, by the Chernoff bound, as follows. Denoting the Gaussian random variable, $U \stackrel{\Delta}{=} \sum_{t=1}^n w_t N_t$, we have

$$
\begin{aligned}
P_{\text{MD}} &= \Pr\left\{ \sum_{t=1}^n w_t s_t + \sum_t w_t Z_t + U \leq \theta n \right\} \\
&\leq \inf_{\lambda \geq 0} \boldsymbol{E}\left(\exp\left\{ \lambda \left[\theta n - \sum_{t=1}^n w_t s_t - \sum_t w_t Z_t - U \right] \right\} \right) \\
&= \inf_{\lambda \geq 0} \exp\left\{ \lambda \left[\theta n - \sum_{t=1}^n w_t s_t \right] \right\} \cdot \boldsymbol{E} \exp\{-\lambda U\} \cdot \boldsymbol{E} \exp\left\{ -\lambda \sum_{t=1}^n w_t Z_t \right\} \\
&= \inf_{\lambda \geq 0} \exp\left\{ \lambda \left[\theta n - \sum_{t=1}^n w_t s_t \right] \right\} \cdot \exp\left\{ \frac{n\lambda^2 \sigma_N^2 P(\boldsymbol{w})}{2} \right\} \cdot \prod_{t=1}^n \boldsymbol{E} \exp\{-\lambda w_t Z_t\} \\
&= \inf_{\lambda \geq 0} \exp\left\{ \lambda \left[\theta n - \sum_{t=1}^n w_t s_t \right] \right\} \cdot \exp\left\{ \frac{n\lambda^2 \sigma_N^2 P(\boldsymbol{w})}{2} \right\} \cdot \prod_{t=1}^n \exp\{C(-\lambda w_t Z_t)\} \\
&= \inf_{\lambda \geq 0} \exp\left\{ \lambda \left[\theta n - \sum_{t=1}^n w_t s_t \right] \right\} \cdot \exp\left\{ \frac{n\lambda^2 \sigma_N^2 P(\boldsymbol{w})}{2} \right\} \cdot \prod_{t=1}^n \exp\{C(\lambda w_t Z_t)\}, \tag{15}
\end{aligned}
$$

---

[3]In the sequel, we will encounter one scenario where the asymptotic empirical PDF will be irrelevant, but this scenario will be handled separately, in the original domain of vectors of dimension n.

where the last step is due to the symmetry of the $C(\cdot)$. The resulting MD exponent is therefore given by

$$E_{\mathrm{MD}}(\theta) = \sup_{\lambda \geq 0} \left\{ \lambda(\boldsymbol{E}\{W \cdot S\} - \theta) - \boldsymbol{E}\{C(\lambda W)\} - \frac{\lambda^2 \sigma_N^2}{2} \cdot \boldsymbol{E}\{W^2\} \right\}, \tag{16}$$

which is a functional of f_{WS}.

The problem of optimal correlator design for a given $\boldsymbol{s}$, is equivalent to the problem of finding a conditional density, $f_{W|S}$, that maximizes the MD exponent subject to the power constraint, $\boldsymbol{E}\{W^2\} \leq P_w$. The problem of joint design of both $\boldsymbol{w}$ and $\boldsymbol{s}$ is asymptotically equivalent to the problem of maximizing the MD exponent over $\{f_{WS}\}$ subject to the power constraints, and $\boldsymbol{E}\{W^2\} \leq P_w$ and $\boldsymbol{E}\{S^2\} \leq P_s$, for some given $P_s > 0$. The first problem is relevant when the detector designer has no control of the transmitted signal, for example, when the transmitter and the receiver are hostile parties, which is typically the case in military applications. The second problem is relevant when the transmitter and the receiver cooperate. In radar applications, for example, the transmitter and the receiver are the same party. In sections 3 and 4, we address the first problem and the second problem, respectively.

Comment 1. Instead of maximizing the MD exponent for a fixed threshold, θ, and a fixed power constraint, P_w, in order to fit a prescribed FA exponent, in principle, there is an alternative approach: maximize the MD exponent directly for a given FA exponent, by substituting $\theta = \sigma\sqrt{2P_w E_{\mathrm{FA}}}$ in the MD exponent expression. Not surprisingly, in this case, the MD exponent would become invariant to scaling in W (as any scaling of W can be absorbed in λ for all terms of the MD exponent), and so, there would be no need for the P_w-constraint, but this invariance property holds only after maximizing over λ, not for a given λ. However, maximizing over λ as a first step of the calculation, does not seem to lend itself to closed form analysis, in general, and consequently, it would make the subsequent optimization extremely difficult, if not impossible, to carry out. We therefore opt to fix both θ and P_w throughout our derivations.

3 Optimum Correlator for a Given Signal

In view of the discussion in Section 2, we wish to find the optimal conditional density, $f_{W|S}$, in the sense of maximizing

$$\lambda \boldsymbol{E}\{W \cdot S\} - \boldsymbol{E}\{C(\lambda W)\} - \frac{\lambda^2 \sigma_N^2}{2} \boldsymbol{E}\{W^2\} = \int_\infty^{+\infty} f_S(s) \cdot \boldsymbol{E} \left\{ \lambda s W - C(\lambda W) - \frac{\lambda^2 \sigma_N^2}{2} \cdot W^2 \Big| S = s \right\} \mathrm{d}s, \tag{17}$$

subject to the power constraint,

$$\boldsymbol{E}\{W^2\} \equiv \int_\infty^{+\infty} f_S(s) \boldsymbol{E}\{W^2 | S = s\} \mathrm{d}s \leq P_w. \tag{18}$$

At this stage, we carry out this optimization for a given $\lambda \geq 0$, but with the understanding that eventually, λ will be subjected to optimization as well. To this end, let us denote the derivative of $C(v)$ by $\dot{C}(v)$, and for a given $\rho \geq 0$, define the function

$$g(w|\rho, \lambda) \stackrel{\Delta}{=} \dot{C}(\lambda w) + \left(\frac{\rho}{\lambda} + \sigma_N^2 \lambda \right) \cdot w. \tag{19}$$

Observe that since C is convex, $\dot{C}$ is monotonically non-decreasing, and so, $g(\cdot|\rho, \lambda)$ is monotonically strictly increasing, which in turn implies that it has an inverse. We denote the inverse of $g(\cdot|\rho, \lambda)$ by $g^{-1}(\cdot|\rho, \lambda)$. Also, since Z is assumed zero mean, then $\dot{C}(0) = 0$, and hence also $g(0|\rho, \lambda) = 0$ and $g^{-1}(0|\rho, \lambda) = 0$. Note also that $g(\cdot|\rho, \lambda)$ (and hence also $g^{-1}(\cdot|\rho, \lambda)$) is a linear function if and only if Z is Gaussian. The following lemma characterizes the optimal $f_{W|S}$.

Theorem 1 *Let the assumptions of Section 2 hold. Assume further that P_w is such that there exists $\rho \geq 0$ (possibly, depending on λ), with $\boldsymbol{E}\{[g^{-1}(S|\rho, \lambda)]^2\} = P_w$. Otherwise, if $\boldsymbol{E}\{[g^{-1}(S|0, \lambda)]^2\} < P_w$, set $\rho = 0$. Then, the optimal conditional density, $f_{W|S}$, is given by*

$$f_{W|S}^*(w|s) = \delta(w - g^{-1}(s|\rho, \lambda)), \tag{20}$$

where $\delta(\cdot)$ is the Dirac delta function.

The theorem tells that the best correlator, $\boldsymbol{w}^* = (w_1^*, \ldots, w_n^*)$, for a given $\boldsymbol{s} = (s_1, \ldots, s_n)$, is obtained by the relation,

$$w_t^* = g^{-1}(s_t|\rho, \lambda), \qquad t = 1, \ldots, n, \tag{21}$$

which means that w_t^* is given by a function of s_t, which is non-linear unless Z is Gaussian. To gain an initial insight regarding the condition on ρ, consider the Gaussian example (Case 1, eq. (4)). In this case, $g(W|\rho, \lambda) = [(\sigma_N^2 + \sigma_Z^2)\lambda + \rho/\lambda]W$, and so, $g^{-1}(S|\rho, \lambda) = \lambda S/[(\sigma_N^2 + \sigma_Z^2)\lambda^2 + \rho]$, whose power is P_w for $\rho = \lambda\sqrt{\boldsymbol{E}\{S^2\}/P_w} - (\sigma_N^2 + \sigma_Z^2)\lambda^2$, which is non-negative as long as $P_w \le \boldsymbol{E}\{S^2\}/[(\sigma_N^2 + \sigma_Z^2)^2\lambda]$. In general, the exact choice of P_w is not crucial, as the prescribed FA exponent can be achieved by adjusting θ proportionally to $\sqrt{P_w}$. However, once P_w is chosen, we will keep it fixed throughout (see Comment 1 above).

Proof of Theorem 1. Consider the following chain of inequalities and inequalities.

$$\sup_{\{f_{W|S}: \, \boldsymbol{E}\{W^2\} \le P_w\}} \int_{\infty}^{+\infty} f_S(s) \cdot \boldsymbol{E}\left\{\lambda s W - C(\lambda W) - \frac{\lambda^2 \sigma_N^2}{2} \cdot W^2 \Big| S = s\right\} ds$$

$$= \sup_{f_{W|S}} \inf_{\varrho \ge 0} \left\{ \int_{\infty}^{+\infty} f_S(s) \cdot \boldsymbol{E}\left\{\lambda s W - C(\lambda W) - \frac{\lambda^2 \sigma_N^2}{2} \cdot W^2 \Big| S = s\right\} ds + \right.$$
$$\left. \frac{\varrho}{2}\left[P_w - \int_{\infty}^{+\infty} f_S(s) \boldsymbol{E}\{W^2|S=s\} ds\right]\right\}$$

$$\overset{(a)}{=} \inf_{\varrho \ge 0} \sup_{f_{W|S}} \left\{ \int_{\infty}^{+\infty} f_S(s) \cdot \boldsymbol{E}\left\{\lambda s W - C(\lambda W) - \left(\frac{\lambda^2 \sigma_N^2}{2} + \frac{\varrho}{2}\right) \cdot W^2 \Big| S = s\right\} ds + \frac{\varrho P_w}{2}\right\}$$

$$\overset{(b)}{=} \inf_{\varrho \ge 0} \left\{ \int_{\infty}^{+\infty} f_S(s) \cdot \sup_{w}\left\{\lambda s w - C(\lambda w) - \left(\frac{\lambda^2 \sigma_N^2}{2} + \frac{\varrho}{2}\right) \cdot w^2\right\} ds + \frac{\varrho P_w}{2}\right\}$$

$$\overset{(c)}{=} \inf_{\varrho \ge 0} \left\{ \int_{\infty}^{+\infty} f_S(s) \cdot \left\{\lambda s g^{-1}(s|\varrho, \lambda) - C(\lambda g^{-1}(s|\varrho, \lambda)) - \left(\frac{\lambda^2 \sigma_N^2}{2} + \frac{\varrho}{2}\right) \cdot [g^{-1}(s|\varrho, \lambda)]^2\right\} ds + \frac{\varrho P_w}{2}\right\}$$

$$= \inf_{\varrho \ge 0} \left\{ \int_{\infty}^{+\infty} f_S(s) \cdot \left\{\lambda s g^{-1}(s|\varrho, \lambda) - C(\lambda g^{-1}(s|\varrho, \lambda)) - \frac{\lambda^2 \sigma_N^2}{2} \cdot [g^{-1}(s|\varrho, \lambda)]^2\right\} ds + \right.$$
$$\left. \frac{\varrho}{2}\left[P_w - \int_{\infty}^{+\infty} f_S(s) \cdot [g^{-1}(s|\varrho, \lambda)]^2 ds\right]\right\}$$

$$\overset{(d)}{\le} \int_{\infty}^{+\infty} f_S(s) \cdot \left\{\lambda s g^{-1}(s|\rho, \lambda) - C(\lambda g^{-1}(s|\rho, \lambda)) - \frac{\lambda^2 \sigma_N^2}{2} \cdot [g^{-1}(s|\rho, \lambda)]^2\right\} ds + $$
$$\frac{\rho}{2}\left[P_w - \int_{\infty}^{+\infty} f_S(s) \cdot [g^{-1}(s|\rho, \lambda)]^2 ds\right]$$

$$\overset{(e)}{=} \int_{\infty}^{+\infty} f_S(s) \cdot \left\{\lambda s g^{-1}(s|\rho, \lambda) - C(\lambda g^{-1}(s|\rho, \lambda)) - \frac{\lambda^2 \sigma_N^2}{2} \cdot [g^{-1}(s|\rho, \lambda)]^2\right\} ds$$

$$= \boldsymbol{E}\left\{\lambda S g^{-1}(S|\rho, \lambda) - C(\lambda g^{-1}(S|\rho, \lambda)) - \frac{\lambda^2 \sigma_N^2}{2} \cdot [g^{-1}(S|\rho, \lambda)]^2\right\}, \tag{22}$$

where (a) is since the objective is affine in both $f_{W|S}$ and in ρ (and hence is concave in $f_{W|S}$ and convex in ϱ), (b) is since the unconstrained maximum of the conditional expectation of a function of W given $S = s$ is attained when $f_{W|S}$ puts all its mass on the maximum of this function, (c)

is because the maximum is over a concave function of w, which is attained at the point of zero-derivative, $w = g^{-1}(s|\varrho, \lambda)$, (d) is by the postulate that $\rho \geq 0$, and (e) is because either $\rho = 0$ or $P_w - \mathbf{E}\{[g^{-1}(S|\rho, \lambda)]^2\} = 0$. The upper bound on the constrained maximum in the first line of the above chain is therefore attained by $W = g^{-1}(S|\rho, \lambda)$ with probability one, which is equivalent to (20). This completes the proof of Theorem 1. $\square$

Optimal correlator weights within a finite set. There is a practical motivation to consider the case where $\boldsymbol{w} = (w_1, \ldots, w_n)$ is restricted to be a binary vector with bipolar components, taking the values $+\sqrt{P_w}$ and $-\sqrt{P_w}$ only. The reason is that in such a case, the implementation of the correlation detector involves no multiplications at all, as it is equivalent to the comparison of the difference

$$\sum_{\{t:\ w_t=\sqrt{P_w}\}} Y_t - \sum_{\{t:\ w_t=-\sqrt{P_w}\}} Y_t$$

to $\theta n/\sqrt{P_w}$. Here, the maximization over w (step (b) in the proof of Theorem 1) is carried out just over its two allowed values, $+\sqrt{P_w}$ and $-\sqrt{P_w}$. As $C(\cdot)$ is symmetric, the maximum is readily seen to be attained by $W = \sqrt{P_w} \cdot \mathrm{sgn}(S)$, which means $w_t^* = \sqrt{P_w} \cdot \mathrm{sgn}(s_t)$.

Suppose, more generally, that $\{w_t\}$ is constrained to take on values in a finite set whose cardinality k is fixed, independent of n. This can be considered as a compromise between the above two extremes of performance and computational complexity, since the number of multiplications need not be larger than $k-1$. The design of such a signal is very similar to the scalar quantizer design problem: A finite–alphabet signal w_t is defined as follows. Let $s_{\min} \equiv a_0 < a_1 < \ldots < a_{k-1} < a_k \equiv s_{\max}$, where $s_{\min} = \min_t s_t$, $s_{\max} = \max_t s_t$, and let $\mathcal{I}_i \overset{\Delta}{=} [a_i, a_{i+1})$, $i = 0, 1, \ldots, k-1$, be given. Define

$$W = \sum_{i=0}^{k-1} \omega_i \cdot 1\{S \in \mathcal{I}_i\}, \tag{23}$$

for some given $\omega_0, \omega_1, \ldots, \omega_{k-1}$. We wish to minimize

$$\Delta = \sum_{i=0}^{k-1} \int_{a_i}^{a_{i+1}} ds \cdot f_S(s) \left[\lambda \omega_i s - C(\lambda \omega_i) - \frac{1}{2}\lambda^2 \sigma_N^2 \omega_i^2 + \frac{\rho}{2}(P_w - \omega_i^2) \right] \tag{24}$$

over $\{a_i\}$, $i = 1, \ldots, k-1$, and $\{\omega_i\}$, $i = 0, 1, \ldots, k-1$. The necessary conditions for optimality are obtained by equating to zero all partial derivatives w.r.t. $\{a_i\}$, $i = 1, \ldots, k-1$, and $\{\omega_i\}$,

$i = 0, 1, \ldots, k - 1$. This results in the following sets of equations:

$$\lambda \omega_{i-1} a_i - C(\lambda \omega_{i-1}) - \left(\frac{\rho}{2} + \frac{\lambda^2 \sigma_N^2}{2} \right) \omega_{i-1}^2 = \lambda \omega_i a_i - C(\lambda \omega_i) - \left(\frac{\rho}{2} + \frac{\lambda^2 \sigma_N^2}{2} \right) \omega_i^2, \quad i = 1, 2, \ldots, k - 1$$

$$\dot{C}(\lambda \omega_i) + \left(\frac{\rho}{\lambda} + \lambda \sigma_N^2 \right) \omega_i = \boldsymbol{E}\{S | S \in \mathcal{I}_i\}, \quad i = 0, 1, \ldots, k - 1.$$

Alternatively, we may represent these equations as:

$$a_i = \frac{C(\lambda \omega_i) - C(\lambda \omega_{i-1}) + (\rho + \lambda^2 \sigma_N^2)(\omega_i^2 - \omega_{i-1}^2)/2}{\lambda(\omega_i - \omega_{i-1})}, \quad i = 1, 2, \ldots, k - 1 \tag{25}$$

$$\omega_i = g^{-1}[\boldsymbol{E}\{S | S \in \mathcal{I}_i\} | \rho, \lambda]], \quad i = 0, 1, \ldots, k - 1, \tag{26}$$

where ρ is tuned such that

$$\sum_{i=0}^{k-1} P(\mathcal{I}_i) \cdot (g^{-1}[\boldsymbol{E}\{S | S \in \mathcal{I}_i\} | \rho, \lambda]])^2 \leq P_w \tag{27}$$

as before. The first set of equations is parallel to the nearest-neighbor condition in optimal quantizer design, and the second set corresponds to the centroid condition. Optimal signal design can be conducted iteratively, like in quantizer design, by alternating between the two sets of equations.

Example 1. Consider the case where $Z \sim \mathcal{N}(0, \sigma_Z^2)$ (i.e., Case 1). In this case, $C(v) = \sigma_Z^2 v^2 / 2$, and so, $\dot{C}(v) = \sigma_Z^2 v$, which leads to

$$g(w | \rho, \lambda) = \sigma_Z^2 \lambda w + \left(\frac{\rho}{\lambda} + \sigma_N^2 \lambda \right) \cdot w = \left[(\sigma_N^2 + \sigma_Z^2) \lambda + \frac{\rho}{\lambda} \right] \cdot w, \tag{28}$$

and so,

$$g^{-1}(s | \rho, \lambda) = \frac{\lambda s}{(\sigma_N^2 + \sigma_Z^2) \lambda^2 + \rho}. \tag{29}$$

Choosing

$$\rho = \lambda \sqrt{\frac{\boldsymbol{E}\{S^2\}}{P_w}} - \lambda^2 \sigma_Z^2, \tag{30}$$

yields

$$w_t^* = \sqrt{\frac{P_w}{\boldsymbol{E}\{S^2\}}} \cdot s_t, \tag{31}$$

which results in

$$E_{\text{MD}}(\theta) = \begin{cases} \frac{(\sqrt{P_w \boldsymbol{E}\{S^2\}} - \theta)^2}{2(\sigma_N^2 + \sigma_Z^2) P_w} & \theta < \sqrt{P_w \boldsymbol{E}\{S^2\}} \\ 0 & \theta \geq \sqrt{P_w \boldsymbol{E}\{S^2\}} \end{cases} \tag{32}$$

12

If w_t is constrained to be binary, then as we already saw, $w_t^* = \sqrt{P_w} \cdot \mathrm{sgn}(s_t)$ and then

$$E_{\mathrm{MD}}(\theta) = \begin{cases} \frac{(\sqrt{P_w} \cdot \boldsymbol{E}\{|S|\} - \theta)^2}{2(\sigma_N^2 + \sigma_Z^2)P_w} & \theta < \sqrt{P_w} \cdot \boldsymbol{E}\{|S|\} \\ 0 & \theta \geq \sqrt{P_w} \cdot \boldsymbol{E}\{|S|\} \end{cases} \tag{33}$$

For the more general quantization, we obtain

$$a_i = \frac{(\sigma_Z^2 \lambda^2/2 + \rho/2)(\omega_i^2 - \omega_{i-1}^2)}{\lambda(\omega_i - \omega_{i-1})} = \left(\sigma_Z^2 \lambda + \frac{\rho}{\lambda}\right) \cdot \frac{\omega_i + \omega_{i-1}}{2}. \tag{34}$$

For simplicity, let us assume that f_S is uniform across the interval $[-A, +A]$. In this case, $\boldsymbol{E}\{S | S \in \mathcal{I}_i\} = (a_i + a_{i+1})/2$, and so,

$$\omega_i = \frac{\lambda(a_i + a_{i+1})}{2[(\sigma_N^2 + \sigma_Z^2)\lambda^2 + \rho)]}. \tag{35}$$

It follows that $\{a_i\}$ have uniform spacings across the support of S, that is, $a_i = (2i/k - 1)S$, $i = 0, 1, \ldots, k$. Accordingly,

$$\omega_i = \frac{\lambda A[(2i+1)/k - 1]}{(\sigma_N^2 + \sigma_Z^2)\lambda^2 + \rho}, \tag{36}$$

where ρ is chosen such that

$$\frac{1}{k} \sum_{i=0}^{k-1} \frac{\lambda^2 A^2[(2i+1)/k - 1]^2}{[(\sigma_N^2 + \sigma_Z^2)\lambda^2 + \rho]^2} = P_w. \tag{37}$$

The binary case considered above is the special case of $k = 2$. This concludes Example 1. $\square$

If $\{s_t\}$ is itself a finite–alphabet signal, then the optimal $\{w_t^*\}$ is also a finite-alphabet signal with the same alphabet size, even without restricting it to be so in the first place. If this alphabet is small enough and/or there is a strong degree of symmetry, one might as well optimize the levels of $\{w_t\}$ directly subject to the power constraint. Consider, for example, the case of a 4-ASK signal, $s_t \in \{-3a, -a, +a, +3a\}$, for some given $a > 0$. Then, since the PDF of Z is assumed symmetric, the alphabet of the optimal $\{w_t\}$ must be of the form $\{-\beta, -\alpha, +\alpha, +\beta\}$ for some $0 < \alpha < \beta$. Assuming that $s_t = \pm a$ along half of the time and $s_t = \pm 3a$ along the other half, then $w_t = \pm\alpha$ and $w_t = \pm\beta$ in the corresponding halves, and so, $\frac{1}{2}\alpha^2 + \frac{1}{2}\beta^2 = P_w$, or $\beta = \sqrt{2P_w - \alpha^2}$. Thus, the MD exponent should be maximized over one parameter only (beyond the optimization over λ), which is $\alpha \in [0, \sqrt{2P_w}]$. In particular,

$$\begin{aligned} E_{\mathrm{MD}}(\theta) &= \sup_{\lambda \geq 0} \max_{0 \leq \alpha \leq \sqrt{2P_w}} \left\{ \frac{1}{2}\lambda a\alpha + \frac{3}{2}\lambda a\sqrt{2P_w - \alpha^2} - \right. \\ &\left. \frac{1}{2}C(\lambda\alpha) - \frac{1}{2}C(\lambda\sqrt{2P_w - \alpha^2}) - \lambda\theta - \frac{\lambda^2 \sigma_N^2 P_w}{2} \right\}. \end{aligned} \tag{38}$$

13

We next examine this expression in several examples.

Example 2. Let Z be a binary symmetric source, taking values $\pm z_0$ for some $z_0 > 0$ (Case 3). Then, owing to eq. (6), eq. (38) becomes

$$
\begin{aligned}
E_{\mathrm{MD}}(\theta) &= \sup_{\lambda \geq 0} \max_{0 \leq \alpha \leq \sqrt{2P_w}} \left\{ \frac{1}{2} \lambda a \alpha + \frac{3}{2} \lambda a \sqrt{2P_w - \alpha^2} - \right. \\
&\qquad \left. \frac{1}{2} \ln \cosh(z_0 \lambda \alpha) - \frac{1}{2} \ln \cosh(z_0 \lambda \sqrt{2P_w - \alpha^2}) - \lambda \theta - \frac{\lambda^2 \sigma_N^2 P_w}{2} \right\} \qquad (39) \\
&= \frac{1}{2} \sup_{\lambda \geq 0} \max_{0 \leq \alpha \leq \sqrt{2P_w}} \left\{ \lambda a \alpha + 3 \lambda a \sqrt{2P_w - \alpha^2} - \right. \\
&\qquad \left. \ln \cosh(z_0 \lambda \alpha) - \ln \cosh(z_0 \lambda \sqrt{2P_w - \alpha^2}) - 2\lambda\theta - \lambda^2 \sigma_N^2 P_w \right\} \qquad (40)
\end{aligned}
$$

The 'classical' correlator, where $w_t \propto s_t$, corresponds to the choice $\alpha = \sqrt{P_w/5}$ instead of maximizing over α. In Fig. 1, we compare the two curves of the MD exponent as functions of θ. Since they share the same level of P_w, the FA exponents are the same for a given θ. As can be seen, the optimal correlator significantly outperforms the classical one, which is optimal in the Gaussian case only. This concludes Example 2. $\square$
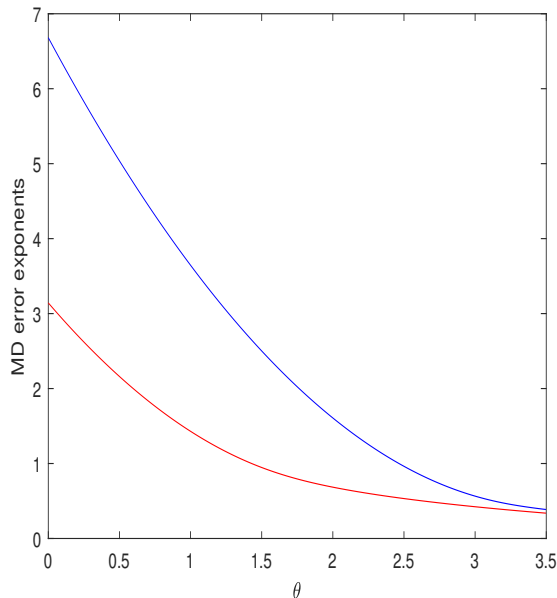

Figure 1: Graphs for binary interference: MD error exponents as functions of θ pertaining to the classical correlator (red curve) and the optimal correlator (blue curve) for the following parameter values: $P_w = 1$, $z_0 = 7$, $a = 4$, and $\sigma_N^2 = 1$.

Example 3. We conduct a similar comparison for the case where Z is distributed uniformly over $[-z_0, +z_0]$ (Case 4), which corresponds to eq. (7). The results are displayed in Fig. 2, and as can be seen, here too, the optimal correlator significantly improves upon the classical one. This concludes Example 3. $\square$
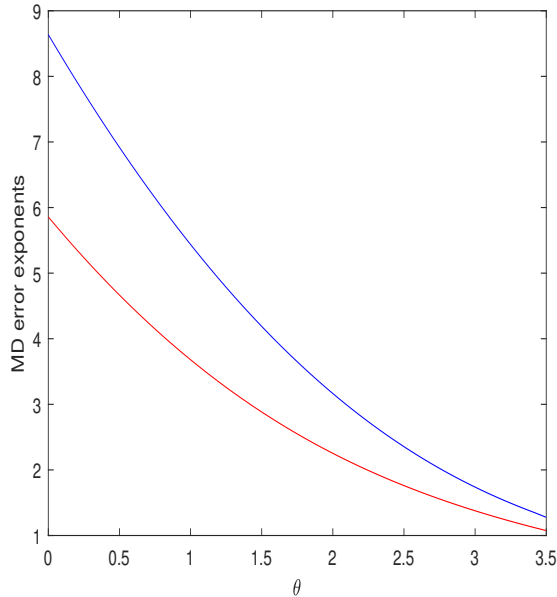

Figure 2: Graphs for uniformly distributed interference: MD error exponents as functions of θ pertaining to the classical correlator (red curve) and the optimal correlator (blue curve) for the following parameter values: $P_w = 1$, $z_0 = 7$, $a = 4$, and $\sigma_N^2 = 1$.

It is interesting to note that in both Examples 2 and 3, for large θ, the two graphs approach each other faster than they approach zero. A possible intuitive explanation is that for large θ, what counts is the behavior of the PDF of $\sum_t w_t Z_t$, fairly close to its peak, where the regime of the central limit theorem is quite relevant, and so, there is no significant difference from Case 1, where Z is Gaussian and the classical correlator is good. Mathematically, as θ grows, the optimum λ decreases, and so, it 'samples' the function $C(\lambda w_t)$ in the vicinity of the origin, where it is well approximated by a quadratic function, just like in the Gaussian case (Case 1).

Example 4. Finally, consider the case where Z is Laplacian (Case 2). In this case, the differences turned out to be rather minor – see Fig. 3. A plausible intuition is that the Laplacian PDF is much

'closer' to the Gaussian PDF, relative to the binary distribution and the uniform distribution of Examples 2 and 3. This concludes Example 4. □
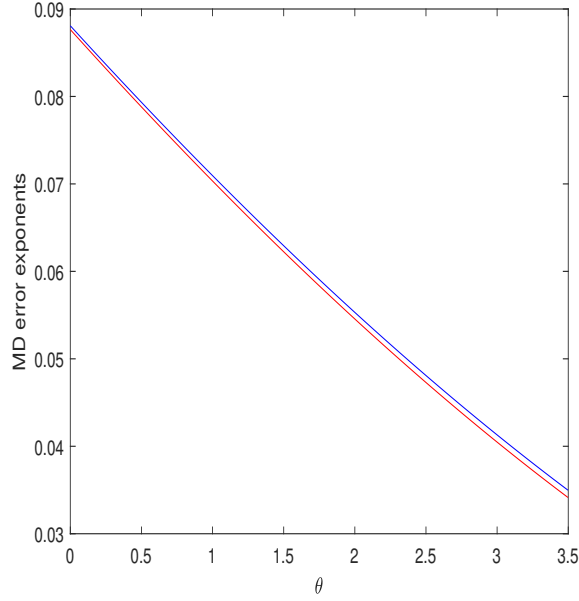

Figure 3: Graphs for Laplace-distributed interference: MD error exponents as functions of θ pertaining to the classical correlator (red curve) and the optimal correlator (blue curve) for the following parameter values: $P_w = 1$, $q = 0.1$, $a = 4$, and $\sigma_N^2 = 1$.

The loss relative to the optimal LRT detector depends on the relative intensity of the process $\{Z_t\}$ compared to the Gaussian noise component.

4 Joint Optimization of the Correlator and the Signal

So far, we have concerned ourselves with the optimization of the correlator waveform, $\{w_t\}$ for a given signal, $\{s_t\}$. But what would be the optimal signal $\{s_t\}$ (subject to a power constraint) when it is jointly optimized with $\{w_t\}$? Mathematically, we are interested in the problem,

$$
\sup_{\{f_S:\ \boldsymbol{E}\{S^2\}\leq P_s\}} \sup_{\{f_{W|S}:\ \boldsymbol{E}\{W^2\}\leq P_w\}} E_{\mathrm{MD}}(\theta)
$$

$$
= \sup_{\{f_S:\ \boldsymbol{E}\{S^2\}\leq P_s\}} \sup_{\{f_{W|S}:\ \boldsymbol{E}\{W^2\}\leq P_w\}} \sup_{\lambda\geq 0} \left[\boldsymbol{E}\left\{ \lambda(W\cdot S - \theta) - C(\lambda W) - \frac{\lambda^2\sigma_N^2 W^2}{2} \right\} \right]
$$

$$
= \sup_{\{f_W:\ \boldsymbol{E}\{W^2\}\leq P_w\}} \sup_{\lambda\geq 0} \sup_{\{f_{S|W}:\ \boldsymbol{E}\{S^2\}\leq P_s\}} \left[\lambda\boldsymbol{E}\{W\cdot S\} - \boldsymbol{E}\{C(\lambda W)\} - \lambda\theta - \frac{\lambda^2\sigma_N^2\boldsymbol{E}\{W^2\}}{2} \right]
$$

16

$$\overset{(a)}{=} \sup_{\{f_W:\ \boldsymbol{E}\{W^2\}\leq P_w\}} \sup_{\lambda\geq 0} \left[\lambda\boldsymbol{E}\left\{ W\cdot\sqrt{\frac{P_s}{\boldsymbol{E}\{W^2\}}}\cdot W \right\} - \boldsymbol{E}\{C(\lambda W)\} - \lambda\theta - \frac{\lambda^2\sigma_N^2\boldsymbol{E}\{W^2\}}{2} \right\}$$

$$= \sup_{\{f_W:\ \boldsymbol{E}\{W^2\}\leq P_w\}} \sup_{\lambda\geq 0} \left\{ \lambda\sqrt{P_s\boldsymbol{E}\{W^2\}} - \boldsymbol{E}\{C(\lambda W)\} - \lambda\theta - \frac{\lambda^2\sigma_N^2\boldsymbol{E}\{W^2\}}{2} \right\}$$

$$= \sup_{\lambda\geq 0} \sup_{P\leq P_w} \left\{ \lambda\sqrt{P_s P} - \min_{\{f_W:\ \boldsymbol{E}\{W^2\}=P\}} \boldsymbol{E}\{C(\lambda W)\} - \lambda\theta - \frac{\lambda^2\sigma_N^2 P}{2} \right\}, \tag{41}$$

where in (a) we have used the simple fact that, for a given W and P_s, the correlation, $\boldsymbol{E}\{W\cdot S\}$ is maximized by $S = \sqrt{P_s/\boldsymbol{E}\{W^2\}}\cdot W$. Earlier, we maximized the MD exponent w.r.t. W for a given S and found that the optimal W is given by a function, $g^{-1}(S|\rho,\lambda)$, which is, in general, non–linear (unless Z is Gaussian). Now, on the other hand, the optimal S for a given W turns out to be given by a linear function. These two findings settle together if and only if W takes values only in the set of solutions, $\mathcal{S}(\zeta)$, to the equation

$$\dot{C}(\lambda W) + \frac{\rho}{\lambda}\cdot W = \zeta\cdot W, \tag{42}$$

for some $\zeta > 0$ (and then S takes the corresponding values according to their relationship). The two sides of the equation represent the non-linear and the linear relations, respectively. Note that $\mathcal{S}(\zeta)$ always includes at least the solution $W = 0$. Once ζ is chosen, W is allowed to take on values only within $\mathcal{S}(\zeta)$. The inner minimization over f_W in the last line of (41) is obviously lower bounded by $\tilde{C}_\lambda(P)$, which is defined as

$$\tilde{C}_\lambda(P) \overset{\Delta}{=} \inf_{\zeta>0} \inf_{\{\mu(\cdot):\ \int_{\mathcal{S}^2(\zeta)} p\cdot\mu(p)\mathrm{d}p=P\}} \int_{\mathcal{S}^2(\zeta)} \mu(p)C(\lambda\sqrt{p})\mathrm{d}p, \tag{43}$$

where $\mathcal{S}^2(\zeta) = \{w^2:\ w\in\mathcal{S}(\zeta)\}$, and $\mu(\cdot)$ is understood to be a weight function over $\mathcal{S}^2(\zeta)$, i.e., $\mu(p)\geq 0$ for all $p\in\mathcal{S}^2(\zeta)$ and $\int_{\mathcal{S}^2(\zeta)}\mu(p)\mathrm{d}p = 1$. While this expression appears complicated, there are two facts that help to simplify it significantly. The first is that, in most cases, $\mathcal{S}(\zeta)$ is a finite set (unless $C(\cdot)$ is linear, or contains linear segments), and the second is that only two members of $\mathcal{S}(\zeta)$ suffice, i.e., eq. (43) simplifies to

$$\tilde{C}_\lambda(P) \overset{\Delta}{=} \inf_{\zeta>0} \min_{\{p_0,p_1\in\mathcal{S}^2(\zeta),\ \alpha\in[0,1]:\ (1-\alpha)p_0+\alpha p_1=P\}} \{(1-\alpha)C(\lambda\sqrt{p_0}) + \alpha C(\lambda\sqrt{p_1})\}. \tag{44}$$

The function $\tilde{C}_\lambda(P)$ has the flavor of a lower convex envelope for the function $C(\lambda\sqrt{\cdot})$, but with the exception that the support of the convex combinations is limited to $\mathcal{S}^2(\zeta)$. Finally, the optimal

17

MD exponent is given by

$$E_{\mathrm{MD}}(\theta) = \sup_{\lambda \geq 0} \sup_{P \leq P_w} \left\{ \lambda(\sqrt{P_s P} - \theta) - \tilde{C}_\lambda(P) - \frac{\lambda^2 \sigma_N^2 P}{2} \right\}. \tag{45}$$

The optimal W is one that achieves $\tilde{C}_\lambda(P)$ for the maximizing λ and P, that is, the components of $\{|w_t|\}$ take only two values in $\mathcal{S}(\zeta^*)$, with relative frequencies given by α^* and $1 - \alpha^*$, where ζ^* and α^* are the achievers of $\tilde{C}_\lambda(P^*)$, P^* being the optimal P. In other words, the optimal signal has at most four levels, $\pm a$ and $\pm b$, for some $a \geq 0$ and $b > 0$.

Comment 2. By a simple change of variables, $q = \lambda^2 p$, it is readily seen that $\tilde{C}_\lambda(P)$ depends on λ and P only via the quantity $\lambda\sqrt{P}$, and so, it might as well be denoted as $\tilde{C}(\lambda\sqrt{P})$. $\square$

Observe that while the function $C(\cdot)$ is always convex, nothing general can be asserted regarding convexity or concavity properties of the function $C(\lambda\sqrt{\cdot})$, as the internal square root function, which is concave, may or may not destroy the convexity of the composite function, depending on the function $C(\cdot)$. In other words, $C(\lambda\sqrt{\cdot})$ may either be convex, or concave, or neither. For example, if $Z \in \{-z_0, +z_0\}$ with equal probabilities (as in Case 3), then $C(\lambda\sqrt{p}) = \ln\cosh(z_0\lambda\sqrt{p})$ which is concave in p. On the other hand, if Z is Laplacian with parameter q (Case 2), then $C(\lambda\sqrt{p}) = -\ln(1 - \lambda^2 p/q^2)$, which is convex in p. By mixing these two distributions, we can also make it neither convex, nor concave, as will be shown in the sequel.

Let us examine now several special cases, where the form of $\tilde{C}_\lambda(P)$ can be determined more explicitly.

1. Consider first the Gaussian case (Case 1), where $C(\lambda\sqrt{p}) = \frac{1}{2}\sigma_Z^2 \lambda^2 p$, namely, it is linear in p. In this case, for $\zeta = \sigma_Z^2 \lambda + \rho/\lambda$, $\mathcal{S}(\zeta) = \mathbb{R}^+$, the choice of μ is immaterial, and $\tilde{C}_\lambda(P) = \frac{1}{2}\sigma_Z^2 \lambda^2 P$. In this case, any signal $\boldsymbol{w}$ with power P is equally good, as expected.

2. Consider next the case where $C(\lambda\sqrt{\cdot})$ is is convex. Then,

$$\begin{aligned}
\boldsymbol{E}\{C(\lambda W)\} &= \boldsymbol{E}\left\{C\left(\lambda\sqrt{W^2}\right)\right\} \tag{46} \\
&\geq C\left(\lambda\sqrt{\boldsymbol{E}\{W^2\}}\right) \tag{47} \\
&= C(\lambda\sqrt{P}), \tag{48}
\end{aligned}$$

where the inequality is achieved with equality whenever $W^2 = \text{const}$ with probability one, and then this constant must be P. So, here $\tilde{C}_\lambda(P) = C(\lambda\sqrt{P})$, $\mathcal{S}^2(\zeta) = \{0, P\}$ and $\mu(p) = \delta(p - P)$, which

18

is expected, because in the convex case, there is no need for any non-trivial convex combinations. The optimal signal vector $\boldsymbol{w}$ is any member of $\{-\sqrt{P^*}, +\sqrt{P^*}\}^n$, and then $\boldsymbol{s}$ is the corresponding member of $\{-\sqrt{P_s}, +\sqrt{P_s}\}^n$. It is interesting to note that they both turn out to be DC or bipolar signals, which is good news from the practical point of view, as discussed in Section 1.

3. We now move on to the case where $C(\lambda\sqrt{\cdot})$ is concave. In this case, it is instructive to return temporarily to the original domain of vectors $\{\boldsymbol{w}\}$ of finite dimension n, find the optimal solution in that domain, and finally, take the limit of large n (see footnote no. 3). We therefore wish to minimize $\frac{1}{n}\sum_{t=1}^{n} C(\lambda w_t)$ s.t. $\sum_{t=1}^{n} w_t^2 = nP$. Since $C(\lambda\sqrt{0}) = 0$ and each w_t^2 is limited to the range $[0, nP]$, we can lower bound the function $C(\lambda\sqrt{w_t^2})$ (which is concave as a function of w_t^2), by a linear function of w_t^2, as follows:

$$C\left(\lambda\sqrt{w_t^2}\right) \geq \frac{C(\lambda\sqrt{nP})}{nP} \cdot w_t^2, \tag{49}$$

with equality at $w_t^2 = 0$ and $w_t^2 = nP$. Consequently,

$$\frac{1}{n}\sum_{t=1}^{n} C(\lambda w_t) = \frac{1}{n}\sum_{t=1}^{n} C\left(\lambda\sqrt{w_t^2}\right) \tag{50}$$

$$\geq \frac{1}{n}\sum_{t=1}^{n} \frac{C(\lambda\sqrt{nP})}{nP} \cdot w_t^2 \tag{51}$$

$$= \frac{C(\lambda\sqrt{nP})}{nP} \cdot P \tag{52}$$

$$= \frac{C(\lambda\sqrt{nP})}{n}, \tag{53}$$

with equality if one of the components of $\boldsymbol{w}$ is equal to $\pm\sqrt{nP}$ and all other components vanish, and then, the same component of $\boldsymbol{s}$ is $\pm\sqrt{nP_s}$ (and, of course, all other vanish), correspondingly. Here, we have $\mathcal{S}^2(\zeta) = \{0, nP\}$ and $\mu(p) = \left(1 - \frac{1}{n}\right)\delta(p) + \frac{1}{n}\delta(p - nP)$. Asymptotically, as n grows without bound, $\tilde{C}_\lambda(P) = \lim_{n\to\infty} C(\lambda\sqrt{nP})/n$, and the limit exists since $C(\lambda\sqrt{nP})/n$ is monotonically non-increasing by the assumed concavity of $C(\lambda\sqrt{\cdot})$. If this limit happens to vanish (like in Case 3, for instance), then the interference $\{Z_t\}$ has no impact whatsoever on the MD exponent for the optimal $\boldsymbol{s}$ and $\boldsymbol{w}$. Here too, the optimal signaling is binary.

We now summarize our findings, in this section so far, in the following theorem.

Theorem 2 *Let the assumptions of Section 2 hold. Then, w_t^* and s_t^* are proportional to each*

other with $|w_t^*|$ and $|s_t^*|$ taking values in a finite set of size at most two $(t = 1, \ldots, n)$, and the MD exponent is given by eq. (45).

1. *If the function $C(\lambda\sqrt{\cdot})$ is convex, then both $\boldsymbol{w}^*$ and $\boldsymbol{s}^*$ are either DC or bipolar, and the MD exponent is given by*

$$E_{MD}(\theta) = \sup_{\lambda \geq 0} \sup_{P \leq P_w} \left\{ \lambda(\sqrt{P_s P} - \theta) - C\left(\lambda\sqrt{P}\right) - \frac{\lambda^2 \sigma_N^2 P}{2} \right\}. \tag{54}$$

2. *If the function $C(\lambda\sqrt{\cdot})$ is concave, then the components of both $\boldsymbol{w}^*$ and $\boldsymbol{s}^*$ are all zero, except for one component which exploits their entire energy. The MD exponent is given by*

$$E_{MD}(\theta) = \sup_{\lambda \geq 0} \sup_{P \leq P_w} \left\{ \lambda(\sqrt{P_s P} - \theta) - \lim_{n \to \infty} \frac{C\left(\lambda\sqrt{Pn}\right)}{n} - \frac{\lambda^2 \sigma_N^2 P}{2} \right\}. \tag{55}$$

Finally, we should consider the case where $C(\lambda\sqrt{\cdot})$ is neither convex, nor concave. Here, we will not carry out the full calculations needed, but we will demonstrate that $\mathcal{S}(\zeta)$ may include more than one positive solution, in addition to the trivial solution at the origin. Consider, for example a mixture of the binary PDF and the Laplacian PDF with weights δ and $1-\delta$, respectively ($\delta \in (0, 1)$). In this case,

$$C(\lambda w) = C\left(\lambda\sqrt{w^2}\right) = \ln\left[\delta \cdot \cosh\left(z_0 \lambda\sqrt{w^2}\right) + \frac{1-\delta}{1 - \lambda^2 w^2/q^2}\right], \tag{56}$$

If δ is close to 1, the hyperbolic cosine term is dominant for small and moderate values of w, where $C(\lambda(\sqrt{\cdot})$ is concave. In contrast, when w^2 approaches q^2/λ^2, the second term tends steeply to infinity and hence must be convex. So in this example, $\hat{C}$ is concave in a certain range of relatively small w^2 and at some point it becomes convex. Now, the derivative w.r.t. w is given by

$$\dot{C}(\lambda w) = \frac{\delta z_0 \sinh(z_0 \lambda w) + (1-\delta) \cdot 2\lambda q^2/(q^2 - \lambda^2 w^2)^2}{\delta \cosh(z_0 \lambda w) + (1-\delta)q^2/(q^2 - \lambda^2 w^2)}. \tag{57}$$

As discussed above, the first step is to solve the equation

$$\dot{C}(\lambda w) = \left(\zeta - \frac{\rho}{\lambda}\right) w. \tag{58}$$

As there is no apparent analytical closed-form solution to this equation, here we demonstrate the solutions graphically. In Fig. 4, we plot the functions $\dot{C}(\lambda w)$ and $(\zeta - \rho/\lambda) \cdot w$ vs. w for the following parameter values: $\delta = 0.95$, $q = 5$, $z_0 = 0.5$, $\lambda = 1$, and $\zeta - \rho/\lambda = 0.13$. As can be

20

seen, in this example, there are two positive solutions (in addition to the trivial solution, $w_0 = 0$), which are approximately, $w_1 = 3.71$ and $w_2 = 4.58$. Thus, in this case, $\mathcal{S}(\zeta) = \{0, 3.71, 4.58\}$, which corresponds to the set of power levels, $\mathcal{S}^2(\zeta) = \{0, 13.7641, 20.9764\}$. According to the above discussion, optimal signaling is associated with time-sharing between two out of these three signal levels. Given this simple fact, the optimal signal levels, say, $a \geq 0$ and $b > 0$, and the optimal weight parameter, α, can also be found directly by maximizing the MD error exponent expression with respect to these parameters, subject to the power constraint, $(1 - \alpha)a^2 + \alpha b^2 = P_w$, similarly as was done earlier in the example of the 4-ASK signal in Section 3.
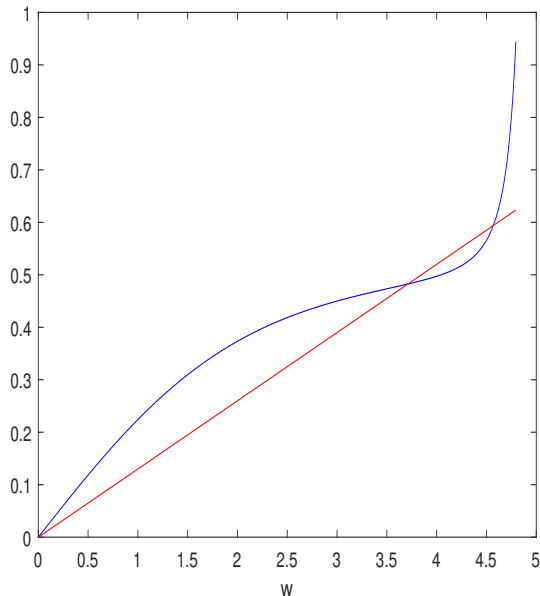

Figure 4: The functions $\dot{C}(\lambda w)$ (blue curve) and $(\zeta - \rho/\lambda) \cdot w$ (red straight line) for the example of eqs. (56) and (57) with parameter values: $\delta = 0.95$, $q = 5$, $z_0 = 0.5$, $\lambda = 1$, and $\zeta - \rho/\lambda = 0.13$. As can be seen, these two graphs meet at three points, $w_0 = 0$, $w_1 \approx 3.71$ and $w_2 \approx 4.58$.

5 Detectors Based on Linear Combinations of Correlation and Energy

In this section, we provide a brief outline of a possible extension of the scope to a broader class of detectors that compare the test statistic

$$\sum_{t=1}^{n} w_t Y_t + \alpha \sum_{t=1}^{n} Y_t^2$$

21

to a threshold, $T = \theta n$. The motivation stems from the fact that the two hypotheses, $\mathcal{H}_0$ and $\mathcal{H}_1$, differ not only in the presence of the signal, $\{s_t\}$, but also in the presence of the SIN, $\{Z_t\}$, which adds to the energy (or the variance) of the received signal. In the extreme case, where $s_t \equiv 0$, the simple correlation detector we examined so far (corresponding to $\alpha = 0$) would be useless, but still, one expects to be able to distinguish between the two hypotheses thanks to the different energies of the received signal. Indeed, if $\{Z_t\}$ is Gaussian white noise (Case 1), the optimal LRT detector obeys this structure with $\alpha = \sigma_Z^2 / [2(\sigma_N^2 + \sigma_Z^2)]$.

For practical reasons, it would also be relevant to consider detectors that are based on

$$\sum_{t=1}^{n} w_t Y_t + \alpha \sum_{t=1}^{n} |Y_t|,$$

where the second term is another measure of the signal intensity, but with the advantage that its calculation does not require multiplications. We shall consider both classes of detectors, but provide merely the basic derivations of the MD exponent, without attempt to arrive at full, explicit solutions. Nevertheless, we will make an attempt to make some observations on the solutions.

We begin with the first class of detectors mentioned above. The FA probability is readily bounded by

$$
\begin{aligned}
P_{\text{FA}}(\theta) &= \Pr\left\{\sum_{t=1}^{n} w_t N_t + \alpha \sum_{t=1}^{n} N_t^2 \geq \theta n\right\} \\
&\leq \exp\left\{-n \sup_{\lambda \geq 0}\left[\lambda\theta - \frac{\lambda^2 \sigma_N^2 P_w}{2(1 - 2\alpha\lambda\sigma_N^2)} + \frac{1}{2}\ln(1 - 2\alpha\lambda\sigma_N^2)\right]\right\},
\end{aligned}
\tag{59}
$$

which depends on $\boldsymbol{w}$ only via P_w, as before.

As for the MD probability, we define

$$A = \frac{1}{n}\sum_{t=1}^{n}(w_t s_t + \alpha s_t^2) \tag{60}$$

and

$$u_t = w_t + 2\alpha s_t, \qquad t = 1, 2, \ldots, n. \tag{61}$$

Then,

$$P_{\text{MD}}(\theta) = \Pr\left\{\sum_{t=1}^{n} w_t(s_t + Z_t + N_t) + \alpha \sum_{t=1}^{n}(s_t + Z_t + N_t)^2 < \theta n\right\}$$

$$
\begin{aligned}
&= \Pr\left\{ nA + \sum_{t=1}^{n} u_t(Z_t + N_t) + \alpha \sum_{t=1}^{n}(Z_t + N_t)^2 < n\theta \right\} \\[2mm]
&\le \boldsymbol{E}\left\{ \exp\left[\lambda n(\theta - A) - \lambda \sum_{t=1}^{n} u_t(Z_t + N_t) - \alpha\lambda \sum_{t=1}^{n}(Z_t + N_t)^2 \right] \right\} \\[2mm]
&= \boldsymbol{E}\left\{ \exp\left[\lambda n(\theta - A) - \lambda \sum_{t=1}^{n} u_t(Z_t + N_t) \right] \times \right. \\[1mm]
&\qquad\qquad \left. \prod_{t=1}^{n} \exp\left[-\alpha\lambda(Z_t + N_t)^2 \right] \right\} \\[2mm]
&\overset{(a)}{=} \boldsymbol{E}\left\{ \exp\left[\lambda n(\theta - A) - \lambda \sum_{t=1}^{n} u_t(Z_t + N_t) \right] \times \right. \\[1mm]
&\qquad\qquad \left. \prod_{t=1}^{n} \left[(4\pi\alpha\lambda)^{-1/2} \int_{-\infty}^{\infty} \exp\left\{ -jq_t(Z_t + N_t) - \frac{q_t^2}{4\alpha\lambda} \right\} \mathrm{d}q_t \right] \right\} \\[2mm]
&= e^{\lambda n(\theta - A)} \prod_{t=1}^{n} \left[(4\pi\alpha\lambda)^{-1/2} \int_{-\infty}^{\infty} \boldsymbol{E}\left\{ \exp\left[-(\lambda u_t + jq_t)(Z_t + N_t) \right] \right\} \exp\left(-\frac{q_t^2}{4\alpha\lambda} \right) \mathrm{d}q_t \right] \\[2mm]
&= e^{\lambda n(\theta - A)} \prod_{t=1}^{n} \left[(4\pi\alpha\lambda)^{-1/2} \int_{-\infty}^{\infty} \boldsymbol{E}\left\{ \exp\left[-(\lambda u_t + jq_t)Z_t \right] \right\} \times \right. \\[1mm]
&\qquad\qquad \left. \boldsymbol{E}\left\{ \exp\left[-(\lambda u_t + jq_t)N_t \right] \right\} \cdot \exp\left(-\frac{q_t^2}{4\alpha\lambda} \right) \mathrm{d}q_t \right] \\[2mm]
&= e^{\lambda n(\theta - A)} \prod_{t=1}^{n} \left[(4\pi\alpha\lambda)^{-1/2} \int_{-\infty}^{\infty} \boldsymbol{E}\left\{ \exp\left(-\lambda u_t Z_t \right) e^{-jq_t Z_t} \right\} \times \right. \\[1mm]
&\qquad\qquad \left. \exp\left(\frac{1}{2}\sigma_N^2[\lambda u_t + jq_t]^2 \right) \cdot \exp\left(-\frac{q_t^2}{4\alpha\lambda} \right) \mathrm{d}q_t \right] \\[2mm]
&= e^{\lambda n(\theta - A)} \prod_{t=1}^{n} \left[(4\pi\alpha\lambda)^{-1/2} \int_{-\infty}^{\infty} \boldsymbol{E}\left\{ \exp\left(-\lambda u_t Z_t \right) e^{-jq_t Z_t} \right\} \times \right. \\[1mm]
&\qquad\qquad \left. \exp\left(\frac{1}{2}\sigma_N^2\lambda^2 u_t^2 \right) e^{j\sigma_N^2 \lambda u_t q_t} \cdot \exp\left\{ -\left(\frac{\sigma_N^2}{2} + \frac{1}{4\alpha\lambda} \right) q_t^2 \right\} \mathrm{d}q_t \right] \\[2mm]
&= \exp\left\{ \lambda n(\theta - A) + \frac{1}{2}\sigma_N^2\lambda^2 \sum_{t=1}^{n} u_t^2 \right\} \times \\[1mm]
&\qquad\qquad \prod_{t=1}^{n} \left[(4\pi\alpha\lambda)^{-1/2} \int_{-\infty}^{\infty} \boldsymbol{E}\left\{ \exp\left(-\lambda u_t Z_t \right) \cos((Z_t - \sigma_N^2\lambda u_t)q_t) \right\} \times \right. \\[1mm]
&\qquad\qquad \left. \exp\left\{ -\left(\frac{\sigma_N^2}{2} + \frac{1}{4\alpha\lambda} \right) q_t^2 \right\} \mathrm{d}q_t \right],
\end{aligned}
\tag{62}
$$

where $j = \sqrt{-1}$ and (a) is due to the identity

$$
e^{-ax^2} = (4\pi a)^{-1/2} \int_{-\infty}^{\infty} e^{-jqx} \exp\left\{ -\frac{q^2}{4a} \right\} \mathrm{d}q, \qquad a > 0,
\tag{63}
$$

which is the characteristic function of a zero-mean Gaussian random variable with variance $2a$.[4]
We now define

$$C_\alpha(v) \triangleq \ln \left[\frac{1}{\sqrt{4\pi\alpha\lambda}} \int_{-\infty}^{\infty} \boldsymbol{E}\left\{ \exp\left(-vZ\right) \cos((Z - \sigma_N^2 v)q) \right\} \cdot \exp\left\{ -\left(\frac{\sigma_N^2}{2} + \frac{1}{4\alpha\lambda} \right) q^2 \right\} \mathrm{d}q \right],$$

and we arrive at the following expression of the MD exponent:

$$
\begin{aligned}
E_{\mathrm{MD}}(\theta) &= \sup_{\lambda \geq 0} \lim_{n \to \infty} \left\{ \lambda(A - \theta) - \frac{1}{2}\lambda^2\sigma_N^2 \cdot \frac{1}{n}\sum_{t=1}^{n} u_t^2 - \frac{1}{n}\sum_{t=1}^{n} C_\alpha(\lambda u_t) \right\} \\
&= \sup_{\lambda \geq 0} \left\{ \lambda\left(\boldsymbol{E}\{S \cdot U\} - \alpha P_s - \theta\right) - \frac{1}{2}\lambda^2\sigma_N^2 \cdot \boldsymbol{E}\{U^2\} - \boldsymbol{E}\{C_\alpha(\lambda U)\} \right\},
\end{aligned}
\tag{64}
$$

where $U = W + 2\alpha S$. Note that this expression is of the same form of the one we had earlier, except that W is replaced by U, θ is replaced by $\theta + \alpha P_s$, and C is replaced by C_α.[5] This expression should now be jointly maximized w.r.t. f_{US} subject to the power constraints, $\boldsymbol{E}\{S^2\} \leq P_s$, $\boldsymbol{E}\{(U - 2\alpha S)^2 \leq P_w$. Using the same techniques as before, it is not difficult to infer that the optimal S for a given U is linear in U, whereas the optimal U for a given S is given by a non-linear equation. Whenever the number of simultaneous solutions to both equations is finite, the signal levels can be optimized directly, as before. As for the optimization of α, among all pairs $\{(\alpha, P_w)\}$ that give rise to the same value of the FA exponent, one chooses the one that maximizes the MD exponent.

Moving on to the second class of detectors, the analysis can be carried out using the same technique as above, where the this time, we use the Fourier transform identity,

$$e^{-a|x|} = \frac{a}{\pi} \int_{-\infty}^{+\infty} \frac{e^{-jqx}\mathrm{d}q}{q^2 + a^2}, \quad a > 0 \tag{65}$$

which, as before, enables to exploit the independence between Z_t and N_t once the expectation operator is commuted with the inverse Fourier transform integral over q. Equipped with this identity, we have

$$
\begin{aligned}
P_{\mathrm{MD}}(\theta) &= \Pr\left\{ \sum_{t=1}^{n} w_t(s_t + Z_t + N_t) + \alpha\sum_{t=1}^{n} |s_t + Z_t + N_t| < \theta n \right\} \\
&\leq \boldsymbol{E}\left\{ \exp\left[\lambda\left(n\theta - \sum_{t=1}^{n} w_t(s_t + Z_t + N_t) - \alpha\sum_{t=1}^{n} |s_t + Z_t + N_t| \right) \right] \right\}
\end{aligned}
$$

---

[4] Alternatively, it can be viewed as the Fourier transform relation between two Gaussians, one in the domain of x and one in the domain of q.

[5] It is easy to verify that $\frac{1}{2}\lambda^2\sigma_N^2 u^2 + C_\alpha(\lambda u)$ is convex in u simply because $\ln \boldsymbol{E}\left\{\exp\left[\lambda(\theta + \alpha P_s) - su) - \lambda u(Z + N) - \alpha\lambda(Z + N)^2\right]\right\}$ is such. Therefore, its derivative is monotonically non-decreasing.

$$= e^{\lambda n\theta} \prod_{t=1}^{n} \left(\boldsymbol{E}\left\{ \exp\left[-\lambda w_t(s_t + Z_t + N_t) \right] \exp\left[-\alpha\lambda|s_t + Z_t + N_t| \right] \right\} \right)$$

$$= e^{\lambda n\theta} \prod_{t=1}^{n} \left(\boldsymbol{E}\left\{ \exp\left[-\lambda w_t(s_t + Z_t + N_t) \right] \cdot \exp\left[-\alpha\lambda|s_t + Z_t + N_t| \right] \right\} \right)$$

$$= e^{\lambda n\theta} \prod_{t=1}^{n} \left(\boldsymbol{E}\left\{ \exp\left[-\lambda w_t(s_t + Z_t + N_t) \right] \cdot \frac{\alpha\lambda}{\pi} \int_{-\infty}^{\infty} \frac{e^{-jq_t(s_t+Z_t+N_t)}\mathrm{d}q_t}{q_t^2 + \alpha^2\lambda^2} \right\} \right)$$

$$= \exp\left\{ \lambda\left(n\theta - \sum_{t=1}^{n} w_t s_t \right) \right\} \prod_{t=1}^{n} \left(\boldsymbol{E}\left\{ \frac{\alpha\lambda}{\pi} \int_{-\infty}^{\infty} \frac{e^{-jq_t s_t} e^{-(\lambda w_t + jq_t)(Z_t+N_t)}\mathrm{d}q_t}{q_t^2 + \alpha^2\lambda^2} \right\} \right)$$

$$= \exp\left\{ \lambda\left(n\theta - \sum_{t=1}^{n} w_t s_t \right) \right\} \prod_{t=1}^{n} \left(\frac{\alpha\lambda}{\pi} \int_{-\infty}^{\infty} \frac{e^{-jq_t s_t}\boldsymbol{E}\left\{ e^{-(\lambda w_t + jq_t)(Z_t+N_t)} \right\}\mathrm{d}q_t}{q_t^2 + \alpha^2\lambda^2} \right)$$

$$= \exp\left\{ \lambda\left(n\theta - \sum_{t=1}^{n} w_t s_t \right) \right\} \times$$

$$\prod_{t=1}^{n} \left(\frac{\alpha\lambda}{\pi} \int_{-\infty}^{\infty} \frac{e^{-jq_t s_t}\boldsymbol{E}\left\{ e^{-(\lambda w_t + jq_t)Z_t} \right\}\boldsymbol{E}\left\{ e^{-(\lambda w_t + jq_t)N_t} \right\}\mathrm{d}q_t}{q_t^2 + \alpha^2\lambda^2} \right)$$

$$= \exp\left\{ \lambda\left(n\theta - \sum_{t=1}^{n} w_t s_t \right) \right\} \times$$

$$\prod_{t=1}^{n} \left(\frac{\alpha\lambda}{\pi} \int_{-\infty}^{\infty} \frac{e^{-jq_t s_t}\boldsymbol{E}\left\{ e^{-(\lambda w_t + jq_t)Z_t} \right\}\exp\left\{ \frac{1}{2}(\lambda w_t + jq_t)^2\sigma_N^2 \right\}\mathrm{d}q_t}{q_t^2 + \alpha^2\lambda^2} \right)$$

$$= \exp\left\{ \lambda\left(n\theta - \sum_{t=1}^{n} w_t s_t \right) + \frac{1}{2}\lambda^2\sigma_N^2\sum_{t=1}^{n} w_t^2 \right\} \times$$

$$\prod_{t=1}^{n} \left(\frac{\alpha\lambda}{\pi} \int_{-\infty}^{\infty} \frac{\boldsymbol{E}\left\{ e^{-\lambda w_t Z_t}\exp\{ jq_t(Z_t + \sigma_N^2\lambda w_t - s_t)\} \right\} e^{-q_t^2\sigma_N^2/2}\mathrm{d}q_t}{q_t^2 + \alpha^2\lambda^2} \right)$$

$$= \exp\left\{ \lambda\left(n\theta - \sum_{t=1}^{n} w_t s_t \right) + \frac{1}{2}\lambda^2\sigma_N^2\sum_{t=1}^{n} w_t^2 \right\} \times$$

$$\prod_{t=1}^{n} \left(\frac{\alpha\lambda}{\pi} \int_{-\infty}^{\infty} \frac{\boldsymbol{E}\left\{ e^{-\lambda w_t Z_t}\cos((Z_t + \sigma_N^2\lambda w_t - s_t)q_t) \right\} e^{-q_t^2\sigma_N^2/2}\mathrm{d}q_t}{q_t^2 + \alpha^2\lambda^2} \right). \qquad (66)$$

Thus, defining

$$C_\alpha(v, s) = \ln\left[\frac{\alpha\lambda}{\pi} \int_{-\infty}^{\infty} \frac{\boldsymbol{E}\left\{ e^{-vZ}\cos((Z + \sigma_N^2 v - s)q) \right\} e^{-q^2\sigma_N^2/2}\mathrm{d}q}{q^2 + \alpha^2\lambda^2} \right], \qquad (67)$$

the MD exponent is

$$E_{\mathrm{MD}}(\theta) = \sup_{\lambda \geq 0}\left\{ \lambda\left(\boldsymbol{E}\{W \cdot S\} - \theta \right) - \frac{1}{2}\lambda^2\sigma_N^2\boldsymbol{E}\{W^2\} - \boldsymbol{E}\{C_\alpha(\lambda W, S)\} \right\}. \qquad (68)$$

However, in this case, there is an additional complication, which stems from the fact that the FA exponent depends on $\boldsymbol{w}$ not only via P_w. A standard Chernoff-bound analysis yields

$$
\begin{aligned}
E_{\mathrm{FA}}(\theta) \;=\; & \sup_{\lambda \geq 0} \left(\lambda\theta - \frac{1}{2}\lambda^2\sigma_N^2(\boldsymbol{E}\{W^2\} + \alpha^2) - \right. \\
& \left. \boldsymbol{E}\left\{ \ln\left[e^{\lambda\alpha W}\left[1 - Q\left(\frac{\lambda(W+\alpha)}{\sigma}\right)\right] + e^{-\lambda\alpha W}Q\left(\frac{\lambda(W-\alpha)}{\sigma}\right)\right]\right\}\right).
\end{aligned}
\tag{69}
$$

Therefore, the maximization of the MD exponent will have to incorporate the full asymptotic PDF of W and not just its second moment.

References

[1] F. Bandiera, D. Orlando, and G. Ricci, *Advanced Radar Detection Schemes Under Mismatched Signal Models*, Synthesis Lectures in Signal Processing, Morgan & Claypool Publishers, 2009.

[2] J. Capon, "On the asymptotic efficiency of locally optimum detectors," *IRE Trans. Inform. Theory*, pp. 67–71, 1961.

[3] E. Conte and G. Ricci, "Sensitivity study of GLRT detection in compound Gaussian clutter," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 34, no. 1, pp. 308–316, January 1998.

[4] A. H. El-Sawy and V. D. Vandelinde, "Robust detection of known signals," *IEEE Transactions on Information Theory*, vol. IT–23, no. 6, pp. 722–727, November 1977.

[5] A. H. El-Sawy and V. D. Vandelinde, "Robust sequential detection of signals in noise," *IEEE Transactions on Information Theory*, vol. IT–25, no. 3, pp. 346–353, November 1979.

[6] E. Erez and M. Feder, "The generalized likelihood ratio decoder can be uniformly improved for Gaussian intersymbol interference channels," *Proc. ISITA 2000*, Honolulu, Hawaii, November 2000.

[7] E. Erez and M. Feder, "Uniformly improving the generalized likelihood ratio test for linear Gaussian channels," *Proc. Allerton Conference on Communications Control and Computing*, pp. 498–507, September 2000.

[8] M. Feder and N. Merhav, "Universal composite hypothesis testing: a competitive minimax approach," *IEEE Trans. Inform. Theory*, special issue in memory of Aaron D. Wyner, vol. 48, no. 6, pp. 1504–1517, June 2002.

[9] E. A. Geraniotis, "Performance bounds for discrimination problems with uncertain statistics," *IEEE Transactions on Information Theory*, vol. IT–31, no. 5, pp. 703–707, September 1985.

[10] F. Gini, M. V. Greco, A. Farina, and P. Lombardo, "Optimum and mismatched detection against K-distributed plus Gaussian clutter," *IEEE Trans. Aerospace and Electronic Systems*, vol. 34, no. 3, pp. 860–876, July 1998.

[11] C. Hao, B. Liu, S. Yan, and L. Cai1, "Parametric adaptive Radar detector with enhanced mismatched signals rejection capabilities," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, Article ID 375136, 2010.

[12] S. A. Kassam, "Robust hypothesis testing for bounded classes of probability densities," *IEEE Transactions on Information Theory*, vol. IT–27, no. 2, pp. 242–247, March 1981.

[13] S. A. Kassam, G. Moustakides, and J. G. Shin, "Robust detection of known signals in asymmetric noise," *IEEE Transactions on Information Theory*, vol. IT–28, no. 1, pp. 84–91, January 1982.

[14] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing: a survey," *Proc. IEEE*, vol. 73, no. 3, pp. 433–481, March 1985.

[15] S. A. Kassam and J. B. Thomas, "Asymptotically robust detection of a known signal in contaminated non-Gaussian noise," *IEEE Transactions on Information Theory*, vol. IT–22, no. 1, pp. 22–26, January 1976.

[16] S. M. Kay, "Robust detection by autoregressive spectrum analysis," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. ASSP–30, no. 2, pp. 256–269, April 1982.

[17] V. M. Krasnenker, "Stable (robust) detection methods for signal against a noise background," *Automatation Remote Control*, vol. 41, no. 5, pt. 1, pp. 640–659, May 1980.

[18] J. Liu and J. Li, "Robust detection in MIMO Radar with steering vector mismatches," *IEEE Trans. Signal Processing*, vol. 67, no. 20, October 2019.

[19] W. Liu, J. Liu, Y. Gao, G. Wang, and Y. Wang, "Multichannel signal detection in interference and noise when signal mismatch happens," *Signal Processing*, vol. 166, 107268 2020.

[20] W. Liu, W. Xie, R. Li, F. Gao, X. Hu, and Y. Wang, "Adaptive detection in the presence of signal mismatch," *Journal of Systems Engineering and Electronics*, vol. 26, no. 1, pp. 38–43, February 2015.

[21] W. Liu, W. Xie, and Y. Wang, "Parametric detector in the situation of mismatched signals," *IET Radar, Sonar and Navigation*, vol. 8, no. 1, pp. 48-–53, 2014.

[22] R. D. Martin and S. C. Schwartz, "Robust detection of a known signal in nearly Gaussian noise," *IEEE Transactions on Information Theory*, vol. IT–17, no. 1, pp. 50–56, January 1971.

[23] N. Merhav, "Optimal correlators for detection and estimation in optical receivers," *IEEE Trans. Inform. Theory*, vol. 67, no. 8, pp. 5200–5210, August 2021.

[24] G. V. Moustakides, "Robust detection of signals: a large deviations approach," *IEEE Transactions on Information Theory*, vol. IT–31, no. 6, pp. 822–825, November 1985.

[25] G. V. Moustakides and J. B. Thomas, "Min-max detection of weak signals in phi-mixing noise," *IEEE Transactions on Information Theory*, vol. IT–30, no. 3, pp. 529–537, May 1984.

[26] X. Shuwen, S. Xingyu, and S. Penglang, "An adaptive detector with mismatched signals rejection in compound Gaussian clutter," *Journal of Radars*, vol. 8, no. 3, pp. 326-334, 2019.

[27] H. van Trees, *Detection, Estimation and Modulation Theory*, part I, John Wiley & Sons, New York 1968.

[28] L. Wei-jian, W. Li-cai, D. Yuan-shui, J. Tao, X, Dang, and W. Yong-liang, "Adaptive energy detector and its application for mismatched signal detection," *Journal of Radars*, vol. 4, no. 2, pp. 149–159, April 2015.

[29] O. Zeitouni, J. Ziv, and N. Merhav, "When is the generalized likelihood ratio test optimal?" *IEEE Trans. Inform. Theory*, vol. 38, no. 5, pp. 1597–1602, September 1992.

[30] D. Zhao, W. Gu, L. Jie, and L. Jing, "Weighted detector for distributed target detection in the presence of signal mismatch," *2020 Proc. IET International Radar Conference (IET IRC 2020)*, pp. 960–963, doi: 10.1049/icp.2021.0816.