

OPTIMAL STOCHASTIC SIGNAL DESIGN AND DETECTOR RANDOMIZATION IN THE NEYMAN-PEARSON FRAMEWORK

Berkan Dulek and Sinan Gezici

Dept. of Electrical and Electronics Engineering, Bilkent University, Ankara, Turkey
{dulek,gezici}@ee.bilkent.edu.tr

ABSTRACT

Power constrained on-off keying communications systems are investigated in the presence of stochastic signaling and detector randomization. The joint optimal design of decision rules, stochastic signals, and detector randomization factors is performed. It is shown that the solution to the most generic optimization problem that employs both stochastic signaling and detector randomization can be obtained as the randomization among no more than three Neyman-Pearson (NP) decision rules corresponding to three deterministic signal vectors. Numerical examples are also presented.

Index Terms– Detection, stochastic signaling, detector randomization, Neyman-Pearson, on-off keying.

1. INTRODUCTION AND MOTIVATION

Optimal signaling and detector design have been studied in elaborate detail under various frameworks (e.g., Bayes, minimax, Neyman-Pearson) for communications systems corrupted by additive white Gaussian noise (AWGN) [1, 2]. Recently, there has been a renewed interest in this subject to improve the performance of communication systems operating under various system constraints. In [3], convexity properties of error probability are studied in the optimal detection of binary-valued scalar signals corrupted by additive noise under an average power constraint. It is shown that the average probability of error is a nonincreasing convex function of the signal power when the channel has a continuously differentiable unimodal noise probability density function (PDF) with finite variance. This discussion is extended from binary modulations to arbitrary signal constellations in [4] by concentrating on the maximum likelihood (ML) detection for AWGN channels. It is proven that an average power-limited transmitter cannot improve its error performance via time-sharing between different power levels in low dimensions (1-D and 2-D) as opposed to the situation for some M -D constellations, $M \geq 3$.

Despite its analytical tractability, the actual noise at the receiver is rarely Gaussian distributed due to effects such as non-linear filtering, intersymbol and multiuser interference [5]. As a result, a comparable amount of effort has been devoted to analyze performance improvements due to randomized signaling and detection techniques over non-Gaussian channels. More specifically, two methods have proven effective in reducing the average probability of error for power constrained communications systems over additive noise channels with multimodal PDFs: either the signals for transmitted symbols are modeled as random variables instead of deterministic quantities [6, 7] (so called *stochastic signaling*), or different detectors are employed at the receiver with certain probabilities corresponding to antipodal signals [8] (so called *detector randomization*). Recently, the authors have studied optimal receiver design for a vector-valued M -ary communications system in which both detector randomization and stochastic signaling can be performed [9]. It is proven that stochastic signaling without detector randomization cannot achieve a smaller average probability of error than detector randomization with deterministic signaling for the same average power constraint and noise statistics. Then, it is shown that the optimal receiver

design results in a randomization between at most two maximum a-posteriori probability (MAP) detectors corresponding to two deterministic signal vectors.

Until recently, the scope of the benefits obtained from stochastic signaling and detector randomization approaches were limited to the average probability of error criterion. However, in some cases the probabilities of detection and false alarm become the main performance metrics as in the Neyman-Pearson (NP) approach. In [10], a power constrained on-off keying communications system is considered in the NP framework, and the problem of designing the optimal stochastic signals is addressed using a single detector in order to maximize the probability of detection without violating the constraints on the probability of false alarm and the average signal power. Based on a theoretical analysis, it is shown that the optimal solution can be obtained by employing randomization between at most two signal vectors for the on-signal (symbol 1) and using the corresponding NP-type likelihood ratio test (LRT) at the receiver.

In this study, we investigate power constrained on-off keying communications systems in the presence of *multiple detectors at the receiver*. Specifically, we consider the joint optimal design of decision rules, stochastic signals, and detector randomization factors. Adopting a similar analysis strategy to [9], it is proven that the solution to the most generic optimization problem (i.e., employing both stochastic signaling and detector randomization) can be obtained as the randomization among no more than three NP decision rules corresponding to three deterministic signal vectors. As a result, the optimal parameters can be computed over a significantly reduced set instead of an infinite space of functions using global optimization techniques. Numerical simulations are conducted to corroborate our theoretical results.

2. DETECTOR RANDOMIZATION AND STOCHASTIC SIGNALING

We consider an average power constrained on-off keying communications system operating over an additive noise channel. The receiver can randomize among at most K different detectors (decision rules) in any manner to improve the average detection performance, as shown in Fig. 1. At any given time, only a single detector is employed at the receiver to conclude the presence/absence of a signal level embedded in noise. Via a communications protocol, the transmitter is informed of the detector currently active at the receiver. As pointed out in [10], in the absence of detector randomization, employing stochastic signaling; that is, modeling the on-signal as a random variable instead of assuming a constant level, can help improve the detection performance without violating the constraints on the false alarm probability and average signal power.

Given an N -dimensional observation vector, the receiver has to decide between two hypotheses H_0 or H_1 specified as

$$H_0 : \mathbf{Y} = \mathbf{N}, \quad H_1 : \mathbf{Y} = \mathbf{S}^{(i)} + \mathbf{N}, \quad i \in \{1, \dots, K\} \quad (1)$$

where $\mathbf{Y}$ is the noisy observation vector, $\mathbf{S}^{(i)}$ represents the transmitted signal vector for the on-signal destined for detector i , and $\mathbf{N}$ is the noise component that is independent of $\mathbf{S}^{(i)}$.

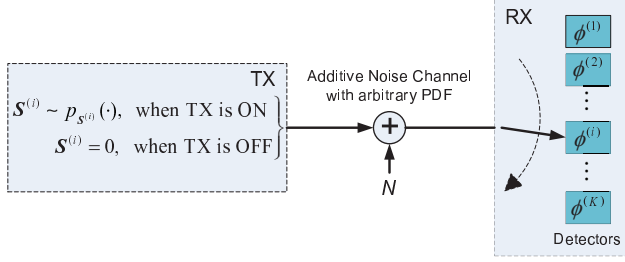


Fig. 1. On-off keying communications system model for joint stochastic signaling and detector randomization.

Furthermore, $\mathbf{S}^{(i)}$ is modeled as a random vector to facilitate stochastic signaling [10]. It should be emphasized that the noise component in (1) is not necessarily Gaussian distributed. Due to interference, such as inter-symbol and multiple-access interference, the effective noise in the channel can deviate significantly from the Gaussian case [5].

Let v_i denote the randomization factor for detector i , where $\sum_{i=1}^K v_i = 1$ and $v_i \geq 0$ for $i = 1, \dots, K$. The two probabilities of interest in the NP framework, the average probability of detection P_D and the average probability of false alarm P_{FA} , can be calculated as $P_D = \sum_{i=1}^K v_i P_D^{(i)}$ and $P_{FA} = \sum_{i=1}^K v_i P_{FA}^{(i)}$. $P_D^{(i)}$ and $P_{FA}^{(i)}$ represent the detection and false alarm probabilities for detector i , respectively; and are specified by $P_D^{(i)} = \int_{\mathbb{R}^N} \phi^{(i)}(\mathbf{y}) p_1^{(i)}(\mathbf{y}) d\mathbf{y}$ and $P_{FA}^{(i)} = \int_{\mathbb{R}^N} \phi^{(i)}(\mathbf{y}) p_N(\mathbf{y}) d\mathbf{y}$, where $\phi^{(i)}$ is the decision rule for detector i , and $p_1^{(i)}(\mathbf{y})$ denotes the conditional PDF of the observation received by detector i under the alternative hypothesis H_1 . Recalling that signal and noise are independent, $p_1^{(i)}(\mathbf{y}) = \int_{\mathbb{R}^N} p_{S^{(i)}}(\mathbf{s}) p_N(\mathbf{y} - \mathbf{s}) d\mathbf{s} = \mathbb{E}\{p_N(\mathbf{y} - \mathbf{S}^{(i)})\}$, where the expectation is taken over the PDF of $\mathbf{S}^{(i)}$. Similarly, under the framework of stochastic signaling and detector randomization, the constraint on the average signal power can be expressed as [1]: $\sum_{i=1}^K v_i \mathbb{E}\{\|\mathbf{S}^{(i)}\|_2^2\} \leq A$, where A denotes the average power limit.

For a given detector i and the corresponding signal PDFs, the probability of detection is maximized under the false alarm constraint using the NP decision rule [1, 2], which takes the form of an LRT

$$\phi_{NP}^{(i)}(\mathbf{y}) = \begin{cases} 1, & \text{if } p_1^{(i)}(\mathbf{y}) \geq \eta^{(i)} p_N(\mathbf{y}) \\ 0, & \text{if } p_1^{(i)}(\mathbf{y}) < \eta^{(i)} p_N(\mathbf{y}) \end{cases}, \quad (2)$$

where the decision threshold $\eta^{(i)} \geq 0$ is chosen such that the probability of false alarm satisfies $P_{FA}^{(i)} = \int_{\mathbb{R}^N} \phi_{NP}^{(i)}(\mathbf{y}) p_N(\mathbf{y}) d\mathbf{y} = \alpha_i$ for some value $\alpha_i \in (0, 1)$. Then, the NP rule is the optimal one among all α_i -level decision rules for detector i , i.e., $P_D^{(i)} = \int_{\mathbb{R}^N} \phi_{NP}^{(i)}(\mathbf{y}) p_1^{(i)}(\mathbf{y}) d\mathbf{y}$ is maximized [1, 2]. Therefore, it is not necessary to search over all decision rules; only the NP decision rule should be determined for each detector and the corresponding average detection and false alarm probabilities should be considered [9, 10]. Using the decision region for the NP detector, $\Gamma_{NP}^{(i)}(p_{S^{(i)}}, \eta^{(i)}) = \{\mathbf{y} \in \mathbb{R}^N : \mathbb{E}\{p_N(\mathbf{y} - \mathbf{S}^{(i)})\} \geq \eta^{(i)} p_N(\mathbf{y})\}$, detection and false alarm probabilities for detector i can be expressed as $P_{D,NP}^{(i)} = \int_{\Gamma_{NP}^{(i)}} \mathbb{E}\{p_N(\mathbf{y} - \mathbf{S}^{(i)})\} d\mathbf{y}$ and $P_{FA,NP}^{(i)} = \int_{\Gamma_{NP}^{(i)}} p_N(\mathbf{y}) d\mathbf{y}$.

By adapting stochastic signaling and detector randomization into the NP framework, we aim to *jointly* optimize the randomization factors, decision thresholds and signal PDFs in order to maximize the average probability of detection under the constraints on the average probability of false alarm and average signal power (Joint optimization can be facilitated via a feedback mechanism from the receiver to the transmitter, such as those in cognitive radio (CR) systems). Then, by denoting the optimization space as $\mathcal{S} \triangleq \{v_i, \eta^{(i)}, p_{S^{(i)}}\}_{i=1}^K$, the optimal design problem can be solved from

$$\begin{aligned} \max_{\mathcal{S}} \quad & \sum_{i=1}^K v_i \int_{\Gamma_{NP}^{(i)}} \mathbb{E}\{p_N(\mathbf{y} - \mathbf{S}^{(i)})\} d\mathbf{y} \\ \text{subject to} \quad & \sum_{i=1}^K v_i \int_{\Gamma_{NP}^{(i)}} p_N(\mathbf{y}) d\mathbf{y} \leq \alpha \\ & \sum_{i=1}^K v_i \mathbb{E}\{\|\mathbf{S}^{(i)}\|_2^2\} \leq A, \quad \sum_{i=1}^K v_i = 1, \quad \mathbf{v} \succeq \mathbf{0} \end{aligned} \quad (3)$$

where $\alpha \in (0, 1)$ is the average false alarm constraint, $\mathbf{v} \succeq \mathbf{0}$ means that $v_i \geq 0 \forall i \in \{1, 2, \dots, K\}$, and expectations are taken over the signal PDFs $p_{S^{(i)}}$. Implicit constraints are also present in (3) due to each $p_{S^{(i)}}$ representing a PDF [9, 10].

A direct evaluation of (3) requires an exhaustive search over the space of randomization factors, decision thresholds and signal PDFs, which is inherently a difficult procedure. Let $P_D^\dagger$ denote the maximum average probability of detection obtained from the solution of (3). In the sequel, an upper bound on this problem with a simpler solution is derived, and then the achievability of this bound is demonstrated. To that aim, the following observations are stated first.

Suppose that the decision rule $\tilde{\phi}_{NP}$ (i.e., threshold $\tilde{\eta}$) and the signal PDF $\tilde{p}_S(\cdot)$ are specified for one of the detectors employed at the receiver. The corresponding detection probability can be written as $\tilde{P}_D = \int_{\mathbb{R}^N} \tilde{\phi}_{NP}(\mathbf{y}) \mathbb{E}\{p_N(\mathbf{y} - \mathbf{S})\} d\mathbf{y} = \mathbb{E}\{\int_{\mathbb{R}^N} \tilde{\phi}_{NP}(\mathbf{y}) p_N(\mathbf{y} - \mathbf{S}) d\mathbf{y}\}$, where the linearity of the expectation operator is imposed over the fixed decision rule $\tilde{\phi}_{NP}$. Recall that the expression inside the expectation operator is the probability of detection when the deterministic signal vector $\mathbf{s}$ is used for the transmission of on-symbol over the additive noise channel and decision rule $\tilde{\phi}_{NP}$ is employed at the receiver. Although the detector $\tilde{\phi}_{NP}$ is in the optimal form for the signal distribution $\mathbb{E}\{p_N(\mathbf{y} - \mathbf{S})\}$, it can be sub-optimal for each component $p_N(\mathbf{y} - \mathbf{s})$. By applying the NP criterion to each signal component $p_N(\mathbf{y} - \mathbf{s})$ that make up the received signal distribution for the on-symbol, the probability of detection can be increased even further without violating the false alarm constraint. More specifically,

$$\hat{\phi}_{NP}(\mathbf{y}, \mathbf{s}) = \begin{cases} 1, & \text{if } p_N(\mathbf{y} - \mathbf{s}) \geq \eta(\mathbf{s}) p_N(\mathbf{y}) \\ 0, & \text{if } p_N(\mathbf{y} - \mathbf{s}) < \eta(\mathbf{s}) p_N(\mathbf{y}) \end{cases} \quad (4)$$

where $\eta(\mathbf{s}) \geq 0$ is determined as a function of $\mathbf{s}$ from the false alarm constraint via $\int_{\mathbb{R}^N} \hat{\phi}_{NP}(\mathbf{y}, \mathbf{s}) p_N(\mathbf{y}) d\mathbf{y} = \int_{\mathbb{R}^N} \tilde{\phi}_{NP}(\mathbf{y}) p_N(\mathbf{y}) d\mathbf{y}$. As a result, the decision rule $\tilde{\phi}_{NP}$ for the given detector can be replaced with a set of decision rules $\hat{\phi}_{NP}$ indexed by parameter $\mathbf{s}$ such that

$$\mathbb{E}\left\{\int_{\mathbb{R}^N} \hat{\phi}_{NP}(\mathbf{y}, \mathbf{S}) p_N(\mathbf{y} - \mathbf{S}) d\mathbf{y}\right\} \geq \int_{\mathbb{R}^N} \tilde{\phi}_{NP}(\mathbf{y}) \mathbb{E}\{p_N(\mathbf{y} - \mathbf{S})\} d\mathbf{y} \quad (5)$$

is always satisfied while guaranteeing the false alarm constraint due to the increased number of optimal NP decision

rules in the new formulation (in contrast with the limited number of detectors in the original problem).

In accordance with the terminology in [9, 10], the left side of the inequality in (5) can be interpreted as a randomization among NP detectors corresponding to deterministic signal vectors, while the right hand side can be understood as stochastic signaling using a single detector. Hence, assuming the same average power and false alarm constraints, the average probability of detection obtained by stochastic signaling with PDF $\tilde{p}_{\mathbf{S}}(\cdot)$ is always smaller than or equal to that of deterministic signaling and detector randomization according to the same PDF when optimal NP detectors are employed in both cases under the same statistics for the additive noise.

Notice that a new decision rule is added for each $\mathbf{s}$ in the support of $\tilde{p}_{\mathbf{S}(i)}$ to obtain the upper bound for a given detector i in the previous analysis. This procedure can be extended safely across multiple detectors by assuming that the supports of $\tilde{p}_{\mathbf{S}(i)}$, $i = 1, 2, \dots, K$ are non-overlapping. If there were overlapping supports, then $\exists \tilde{\mathbf{s}} \in \mathbb{R}^N$ such that $\tilde{p}_{\mathbf{S}(i)}(\tilde{\mathbf{s}}) \neq 0$ and $\tilde{p}_{\mathbf{S}(j)}(\tilde{\mathbf{s}}) \neq 0$ for $i \neq j$. After applying the procedure described above, there would be contributions in the overall average false alarm probability as $\tilde{v}_i \tilde{p}_{\mathbf{S}(i)}(\tilde{\mathbf{s}}) \int_{\mathbb{R}^N} \hat{\phi}_{\text{NP}}^{(i)}(\mathbf{y}, \tilde{\mathbf{s}}) p_{\mathbf{N}}(\mathbf{y}) d\mathbf{y} + \tilde{v}_j \tilde{p}_{\mathbf{S}(j)}(\tilde{\mathbf{s}}) \int_{\mathbb{R}^N} \hat{\phi}_{\text{NP}}^{(j)}(\mathbf{y}, \tilde{\mathbf{s}}) p_{\mathbf{N}}(\mathbf{y}) d\mathbf{y} \triangleq \alpha_{ij}$. Similarly, the contributions from these terms to the average detection probability would be $\tilde{v}_i \tilde{p}_{\mathbf{S}(i)}(\tilde{\mathbf{s}}) \int_{\mathbb{R}^N} \hat{\phi}_{\text{NP}}^{(i)}(\mathbf{y}, \tilde{\mathbf{s}}) p_{\mathbf{N}}(\mathbf{y} - \tilde{\mathbf{s}}) d\mathbf{y} + \tilde{v}_j \tilde{p}_{\mathbf{S}(j)}(\tilde{\mathbf{s}}) \int_{\mathbb{R}^N} \hat{\phi}_{\text{NP}}^{(j)}(\mathbf{y}, \tilde{\mathbf{s}}) p_{\mathbf{N}}(\mathbf{y} - \tilde{\mathbf{s}}) d\mathbf{y}$. Then, the contributions from detectors i and j can be replaced in the respective expressions with a single term corresponding to the NP decision rule $\hat{\phi}_{\text{NP}}(\mathbf{y}, \tilde{\mathbf{s}})$ with the false alarm probability $\alpha_{ij}/(\tilde{v}_i \tilde{p}_{\mathbf{S}(i)}(\tilde{\mathbf{s}}) + \tilde{v}_j \tilde{p}_{\mathbf{S}(j)}(\tilde{\mathbf{s}}))$ and the corresponding weight coefficient would become $\tilde{v}_i \tilde{p}_{\mathbf{S}(i)}(\tilde{\mathbf{s}}) + \tilde{v}_j \tilde{p}_{\mathbf{S}(j)}(\tilde{\mathbf{s}})$. Since the receiver operating characteristics (ROC) curve corresponding to an NP decision rule is concave for any given $\tilde{\mathbf{s}}$, the resulting system would have an even higher average detection probability while possessing the same average false alarm probability and average signal power as the former case [2].

In the light of these observations and the inequality in (5), an upper bound on the problem in (3) can be obtained as

$$\begin{aligned} & \max_{\mathbf{p}_{\mathbf{S}, \eta}} \mathbb{E}\{D(\mathbf{S}, \eta)\} \\ & \text{subject to } \mathbb{E}\{F(\mathbf{S}, \eta)\} \leq \alpha \text{ and } \mathbb{E}\{\|\mathbf{S}\|_2^2\} \leq A \end{aligned} \quad (6)$$

with $D(\mathbf{S}, \eta) \triangleq \int_{\Gamma(\mathbf{S}, \eta)} p_{\mathbf{N}}(\mathbf{y} - \mathbf{S}) d\mathbf{y}$, and $F(\mathbf{S}, \eta) \triangleq \int_{\Gamma(\mathbf{S}, \eta)} p_{\mathbf{N}}(\mathbf{y}) d\mathbf{y}$, where $\Gamma(\mathbf{s}, \eta) = \{\mathbf{y} \in \mathbb{R}^N : p_{\mathbf{N}}(\mathbf{y} - \mathbf{s}) \geq \eta p_{\mathbf{N}}(\mathbf{y})\}$ and the expectations are taken with respect to the joint PDF $p_{\mathbf{S}, \eta}(\mathbf{s}, \eta)$ by treating both $\mathbf{S}$ and η as random variables. Let P_{D}^* denote the maximum average probability of detection obtained as the solution to the optimization problem in (6). Since this is an upper bound, $P_{\text{D}}^* \geq P_{\text{D}}^\dagger$ is always satisfied.

Optimization problems in the form of (6) have been studied in various contexts [6, 9, 10]. Assuming that $D(\mathbf{s}, \eta)$ and $F(\mathbf{s}, \eta)$ are continuous functions defined on a compact subset of $\mathbb{R}^{N+1}$, then an optimal solution to (6) can be expressed by a convex combination among at most three components due to Carathéodory's theorem [11]; that is, $p_{\mathbf{S}, \eta}^{\text{opt}}(\mathbf{s}, \eta) = \lambda_1 \delta(\mathbf{s} - \mathbf{s}_1, \eta - \eta_1) + \lambda_2 \delta(\mathbf{s} - \mathbf{s}_2, \eta - \eta_2) + \lambda_3 \delta(\mathbf{s} - \mathbf{s}_3, \eta - \eta_3)$. Motivated by this observation, we state the following proposition.

Proposition 1: *The solution of the optimization problem in (3) can be obtained as follows:*

$$\begin{aligned} & \max_{\{\lambda_i, \mathbf{s}_i, \eta_i\}_{i=1}^3} \sum_{i=1}^3 \lambda_i \int_{\Gamma(\mathbf{s}_i, \eta_i)} p_{\mathbf{N}}(\mathbf{y} - \mathbf{s}_i) d\mathbf{y} \\ & \text{subject to } \sum_{i=1}^3 \lambda_i \int_{\Gamma(\mathbf{s}_i, \eta_i)} p_{\mathbf{N}}(\mathbf{y}) d\mathbf{y} \leq \alpha \\ & \sum_{i=1}^3 \lambda_i \|\mathbf{s}_i\|_2^2 \leq A, \quad \sum_{i=1}^3 \lambda_i = 1 \\ & \lambda_i \geq 0 \text{ and } \eta_i \geq 0 \quad \forall i \in \{1, 2, 3\} \end{aligned} \quad (7)$$

where $\Gamma(\mathbf{s}_i, \eta_i) = \{\mathbf{y} \in \mathbb{R}^N : p_{\mathbf{N}}(\mathbf{y} - \mathbf{s}_i) \geq \eta_i p_{\mathbf{N}}(\mathbf{y})\}$ $\forall i \in \{1, 2, 3\}$, and $\alpha \in (0, 1)$.

Proof: The optimization problem in (7) is obtained by substituting the form of the optimal PDF $p_{\mathbf{S}, \eta}^{\text{opt}}(\mathbf{s}, \eta)$ into the optimization problem in (6). Now, we show that the optimization problems in (3) and (7) result in the same maximum value. Since (6) and equivalently (7), provide an upper bound on (3), $P_{\text{D}}^\dagger \leq P_{\text{D}}^*$. Next, consider the optimization problem in (3) when $K = 3$ detectors are used and deterministic signaling is employed for each detector, that is, $p_{\mathbf{S}(i)}(\mathbf{s}) = \delta(\mathbf{s} - \mathbf{s}_i)$, $i = 1, 2, 3$. In that case, (3) reduces to the optimization problem in (7). As a result, the maximum value of the objective function in (3) should be larger than or equal to that of (7); namely, $P_{\text{D}}^\dagger \geq P_{\text{D}}^*$. Therefore, $P_{\text{D}}^\dagger = P_{\text{D}}^*$ must be satisfied. ■

A few conclusions can be drawn from Proposition 1. Firstly, when multiple detectors are available for randomization ($K \geq 3$), it is sufficient to employ detector randomization among three deterministic signal vectors; i.e., there is no need to employ stochastic signaling to achieve the optimal solution. Secondly, the solution of (3) can be obtained by optimizing over a significantly reduced optimization space via (7). Despite the simplification, the solution still requires global optimization algorithms such as particle swarm optimization (PSO), or convex relaxation techniques can be utilized to obtain an approximate solution [6]. In this paper, MATLAB's *multistart* method is employed with 500 random start points and *sqp* algorithm is used together with the local solver *fmincon*. The extrema returned by the method are cross-checked with the results from the *patternsearch* algorithm.

3. SIMULATION RESULTS AND CONCLUSIONS

In this section, the numerical example presented in [10] is revisited to compare the detection performance of the optimal solution obtained in the previous section against various signalling techniques studied in [10]. It is assumed that the receiver is capable of randomizing among multiple detectors ($K \geq 3$). The noise N in (1) is assumed to have a symmetric Gaussian mixture distribution with equal variances as follows: $p_{\mathbf{N}}(\mathbf{n}) = \sum_{i=1}^L l_i \exp\{-(\mathbf{n} - \boldsymbol{\mu}_i)^2 / (2\sigma^2)\} / (\sqrt{2\pi}\sigma)$, where $\mathbf{l} = [0.1492 \ 0.1088 \ 0.2420 \ 0.2420 \ 0.1088 \ 0.1492]$, and $\boldsymbol{\mu} = [-1.211 \ -0.755 \ -0.3 \ 0.3 \ 0.755 \ 1.211]$. The average signal power and average false alarm constraints are selected as $A = 1$ and $\alpha = 0.05$, respectively. The following signaling schemes, which employ a single detector at the receiver, were already discussed in [10]:

Conventional Solution: Lacking any information about the noise PDF, the transmitter employs deterministic signaling at the maximum permitted power level ($S = \sqrt{A}$). The receiver is fully informed of the channel statistics, and designs the corresponding optimal NP decision rule [10, Eq. 11].

Optimal-Deterministic: This scheme searches for the optimal deterministic signal level and the corresponding NP decision rule to maximize the detection probability [10, Eq. 12].

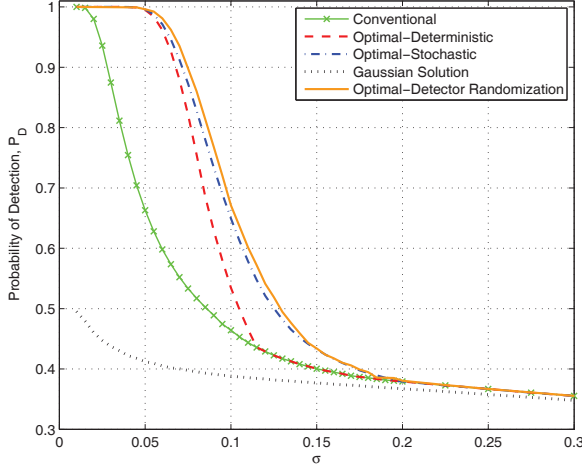


Fig. 2. Probability of detection P_D versus σ for different approaches when $A = 1$ and $\alpha = 0.05$.

Optimal–Stochastic: The optimal stochastic signal and the corresponding NP detector are designed jointly from [10, Eq. 9]. The optimal signal for the on-symbol consists of a randomization between at most two different signal vectors.

Also, a worst case scenario is considered where we assume that the receiver has a limited capability in the sense that it can measure only the mean and variance of the channel noise instead of a complete statistical knowledge of the noise.

Gaussian Solution: The transmitter employs signaling at the maximum power level ($S = \sqrt{A}$), and the receiver uses the corresponding NP detector assuming that the noise present in the channel is Gaussian distributed. Specifically, the α -level NP test at the receiver is given by $\hat{\phi}_{NP}(y) = 1$ if $y \geq \hat{\sigma} \mathcal{Q}^{-1}(\alpha)$ and $\hat{\phi}_{NP}(y) = 0$ otherwise, where $\hat{\sigma}^2 = \sigma^2 + \sum_{i=1}^L l_i \mu_i^2$ is the average power of the zero-mean channel noise, and $\mathcal{Q}(x) = (\int_x^\infty e^{-t^2/2} dt) / \sqrt{2\pi}$ is the tail probability of the standard normal distribution.

Finally, the following scheme is considered as the overall optimal solution when detector randomization is allowed at the receiver as well:

Optimal Detector Randomization with Deterministic Signaling: This case refers to the solution of the most generic optimization problem in (3), which can be obtained from (7) as studied in the previous section.

In Fig. 2, the detection probabilities of the schemes described above are plotted versus $\sigma \in \{0.01 : 0.005 : 0.30\}$. From the figure, it is observed that the Gaussian solution has the worst performance as expected since neither the signaling scheme nor the detector is optimized according to the multimodal channel noise PDF. Respectively, the conventional solution presents poor performance as well since no optimization is performed for the signaling scheme employed at the transmitter even though the detector is optimized by taking into account the actual noise PDF. Optimizing deterministic signal levels improves over the performance of the conventional solution for $0.01 \leq \sigma \leq 0.115$, as observed from the Optimal–Deterministic curve. Further performance improvements are obtained when randomization is performed between two signal levels instead of a deterministic signal (see Optimal–Stochastic for $0.04 \leq \sigma \leq 0.20$). However, the highest detection performance is achieved by the solution of the most generic joint optimization problem given in (7), which performs randomization among NP detectors corresponding to three or fewer deterministic signal values for the on-symbol (see Optimal–Detector Randomization). For

σ	Detector I			Detector II			Detector III		
	λ_1	S_1	η_1	λ_2	S_2	η_2	λ_3	S_3	η_3
0.010	0.609	1.216	2.556	0.389	0.201	3.433	0.002	0.169	1.931
0.050	0.609	1.267	0.021	0.391	0.237	0.328	0	N/A	N/A
0.075	0.617	1.258	2.511	0.383	0.251	1.274	0	N/A	N/A
0.100	0.665	1.211	3.380	0.335	0.265	3.014	0	N/A	N/A
0.125	0.639	1.228	2.972	0.218	0.319	3.254	0.143	0.315	2.879
0.150	0.551	1.212	2.362	0.449	0.651	2.226	0	N/A	N/A
0.175	0.686	1.153	1.993	0.314	0.530	2.552	0	N/A	N/A
0.200	0.724	1.101	1.863	0.247	0.594	2.644	0.029	1.118	1.194
0.250	0.979	1.007	1.917	0.021	0.636	3.482	0	N/A	N/A
0.300	0.751	1.005	2.033	0.249	0.984	2.190	0	N/A	N/A

Table 1. Optimal parameters for Detector Randomization

example, at $\sigma = 0.1$, a detection probability of 0.671 can be achieved by transmitting $s_1 = 1.211$ with probability $\lambda_1 = 0.665$ and $s_2 = 0.265$ with probability $\lambda_2 = 0.335$, and employing the corresponding NP detectors with false alarm probabilities $\alpha_1 = 0.0368$ and $\alpha_2 = 0.0763$ (see Table 1 for more results). On the other hand, the optimal stochastic solution results in a randomization between $s_1 = 1.246$ and $s_2 = 0.273$ with $\lambda = 0.626$, achieving a detection probability of 0.645 using a single detector [10, Table 1]. It is seen in Table 1 that randomization between two NP decision rules achieves the highest detection performance for most values of σ . Since Proposition 1 states that at most three detectors are sufficient to obtain the optimal solution via randomization, one can find examples where optimal performance can be achieved using fewer detectors as in this case. On the contrary, there may be cases where randomization among three detectors becomes a necessity for optimality (e.g., some multivariate noise PDFs, $N > 1$). It is also observed that all signaling schemes get closer for large values of σ due to increasing overlap among mixture components which renders randomized approaches ineffective over the conventional ones.

4. REFERENCES

- [1] H. V. Poor, *An Introduction to Signal Detection and Estimation*, Springer-Verlag, New York, 1994.
- [2] H. L. V. Trees, *Detection, Estimation, and Modulation Theory: Part I*, John Wiley & Sons, Inc., New York, NY, 2001.
- [3] M. Azizoglu, “Convexity properties in binary detection problems,” *IEEE Trans. Info. Theory*, vol. 42, no. 4, pp. 1316–1321, July 1996.
- [4] S. Loyka, V. Kostina, and F. Gagnon, “Error rates of the maximum-likelihood detector for arbitrary constellations: Convex/concave behavior and applications,” *IEEE Trans. Info. Theory*, vol. 56, pp. 1948–1960, Apr. 2010.
- [5] V. Bhatia and B. Mulgrew, “Non-parametric likelihood based channel estimator for Gaussian mixture noise,” *Signal Processing*, vol. 87, pp. 2569–2586, Nov. 2007.
- [6] C. Goken, S. Gezici, and O. Arikan, “Optimal stochastic signaling for power-constrained binary communications systems,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp. 3650–3661, Dec. 2010.
- [7] C. Goken, S. Gezici, and O. Arikan, “Optimal signaling and detector design for power-constrained binary communications systems over non-Gaussian channels,” *IEEE Commun. Letters*, vol. 14, pp. 100–102, Feb. 2010.
- [8] A. Patel and B. Kosko, “Optimal noise benefits in Neyman-Pearson and inequality-constrained signal detection,” *IEEE Trans. Sig. Processing*, vol. 57, pp. 1655–1669, May 2009.
- [9] B. Dulek and S. Gezici, “Detector randomization and stochastic signaling for minimum probability of error receivers,” *IEEE Trans. Commun.*, to appear, 2012.
- [10] B. Dulek and S. Gezici, “Optimal signaling and detector design for power constrained on-off keying systems in Neyman-Pearson framework,” in *Proc. IEEE Statistical Signal Process. Workshop (SSP)*, June 2011, pp. 93–96.
- [11] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1968.