Detector Randomization and Stochastic Signaling for Minimum Probability of Error Receivers

Berkan Dulek, *Student Member, IEEE,* and Sinan Gezici, *Senior Member, IEEE*

Abstract— **Optimal receiver design is studied for a communications system in which both detector randomization and stochastic signaling can be performed. First, 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. Numerical examples are provided to explain the results.**

Index Terms– **Probability of error, detector randomization, stochastic signaling.**

I. INTRODUCTION

Over non-Gaussian communications channels [1], transmitting a stochastic signal for each symbol instead of a deterministic signal can provide performance improvements in terms of the average probability of error [2], [3]. It is shown in [2] that, for a given detector, an optimal stochastic signal can be represented by a randomization of no more than three different signal values under second and fourth moment constraints. In addition, the joint optimal design of stochastic signals and a detector is studied in [3] under an average power constraint, and it is proven that the optimal solution results in stochastic signals with at most two distinct signal values and the corresponding maximum a-posteriori probability (MAP) detector.

Another approach to improve performance of communications systems over non-Gaussian channels is to perform randomization among multiple detectors [4], [5]. In that case, different detectors are employed at the receiver with certain probabilities. In [5], an average power constrained binary communications system is studied, and randomization between two deterministic signal pairs and the corresponding MAP detectors is investigated. It is shown that detector randomization can provide significant performance improvements. In a related work, optimal additive noise components are studied for variable detectors in the context of stochastic resonance, and the optimal randomization between detector and additive noise pairs is investigated [6].

Although the optimal design of stochastic signals and the corresponding MAP detector is analyzed in [3], and the optimal detector randomization and the corresponding MAP detectors are investigated in [5], no studies have considered the joint optimal design of detectors, stochastic signals, and detector randomization. Specifically, the study in [3] did not consider any detector randomization, and that in [5]

The authors are with the Department of Electrical and Electronics Engineering, Bilkent University, Bilkent, Ankara 06800, Turkey, Tel: +90 (312) 290-3139, Fax: +90 (312) 266-4192, e-mails: {dulek,gezici}@ee.bilkent.edu.tr.

assumed deterministic signals (no stochastic signaling). In this study, both detector randomization and stochastic signaling are considered, and the problem of jointly optimizing detectors, stochastic signals, and detector randomization is investigated. First, it is proven that stochastic signaling without detector randomization can never achieve a lower average probability of error than detector randomization with deterministic signaling for the same average power constraint and channel statistics. Then, based on this result and some additional analysis, the solution to the most generic optimization problem is obtained as a randomization between at most two MAP detectors corresponding to two deterministic signal vectors. Two detection examples are provided to compare various optimal and suboptimal signaling schemes.

II. DETECTOR RANDOMIZATION AND STOCHASTIC SIGNALING

Consider an M-ary communications system, in which the receiver acquires N-dimensional observations over an additive noise channel. The receiver is allowed to randomize or time-share among at most K different detectors (decision rules) to improve the detection performance, as shown in Fig. 1. At any given time, only one of those K detectors can be employed at the receiver for the recognition of the transmitted symbol. The transmitter and the receiver are assumed to be synchronized in the sense that the transmitter knows which detector is currently in use at the receiver.[1] Furthermore, a stochastic signaling approach is adopted by treating the transmitted signals for each detector as random vectors. As investigated in [2] and [3] in the absence of detector randomization, employing stochastic signaling; that is, modeling signals for different symbols as random variables instead of deterministic quantities, can provide performance improvements in some scenarios.

Considering both detector randomization and stochastic signaling, the noisy observation vector $\mathbf{Y}$ received by the ith detector can be modeled as follows:

$$\mathbf{Y} = \mathbf{S}_j^{(i)} + \mathbf{N}, \ j \in \{0, 1, \ldots, M-1\}, \ i \in \{1, \ldots, K\} \tag{1}$$

where $\mathbf{S}_j^{(i)}$ represents the transmitted signal vector for symbol j that is intended for detector i, and $\mathbf{N}$ is the noise component that is independent of $\mathbf{S}_j^{(i)}$. It should be emphasized that $\mathbf{S}_j^{(i)}$ is modeled as a random vector to facilitate stochastic signaling. Also, the prior probabilities of the symbols, represented by $\pi_0, \pi_1, \ldots, \pi_{M-1}$, are assumed to be known. In addition,

---

[1] In practice, this can be achieved by employing a communications protocol that allocates the first $N_{s,1}$ symbols in the payload for detector 1, ..., the last $N_{s,K}$ symbols for detector K. The information on the numbers of symbols for different detectors can be included in the header of a communications packet.
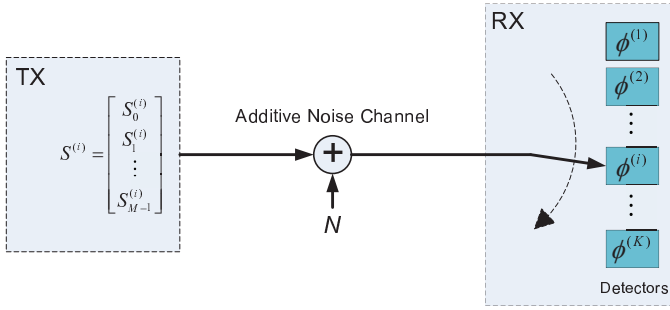
Fig. 1. System model for joint stochastic signaling and detector randomization.

although the signal model in (1) is in the form of a simple additive noise channel, it holds for flat-fading channels as well assuming perfect channel estimation [2].

At the receiver, K generic detectors are utilized to estimate the symbol specified in (1). That is, for a given observation vector $\mathbf{Y} = \mathbf{y}$, the ith detector $\phi^{(i)}(\mathbf{y})$ is described as

$$\phi^{(i)}(\mathbf{y}) = j, \quad \text{if } \mathbf{y} \in \Gamma_j^{(i)} \tag{2}$$

for $i \in \{1, \ldots, K\}$ and $j \in \{0, 1, \ldots, M-1\}$, where $\Gamma_0^{(i)}, \Gamma_1^{(i)}, \ldots, \Gamma_{M-1}^{(i)}$ are the decision regions for the ith detector [7]. The receiver can randomize among these K detectors in any manner in order to optimize its probability of error performance. Let v_i denote the randomization (or time-sharing) factor for detector $\phi^{(i)}$, where $\sum_{i=1}^K v_i = 1$ and $v_i \geq 0$ for $i = 1, \ldots, K$. Then, out of N_s symbols, $v_i N_s$ of them are processed by detector $\phi^{(i)}$ for $i = 1, \ldots, K$.[2]

The aim of this study is to *jointly* optimize the randomization factors, decision regions, and stochastic signals in order to achieve the minimum average probability of error, or equivalently, the maximum average probability of correct decision. The average probability of correct decision can be expressed as $\mathrm{P_c} = \sum_{i=1}^K v_i \mathrm{P_c^{(i)}}$, where $\mathrm{P_c^{(i)}}$ represents the probability of correct decision for the ith detector under M-ary signaling; i.e.,

$$\mathrm{P_c^{(i)}} = \sum_{j=0}^{M-1} \pi_j \int_{\Gamma_j^{(i)}} p_j^{(i)}(\mathbf{y}) \, d\mathbf{y} \tag{3}$$

for $i = 1, 2, \ldots, K$, with $p_j^{(i)}(\mathbf{y})$ denoting the conditional probability density function (PDF) of the observation when the jth symbol that is to be received by the ith detector is transmitted. Recalling that the signals and the noise are independent, the conditional PDF of the observation can be calculated as $p_j^{(i)}(\mathbf{y}) = \int_{\mathbb{R}^N} p_{\mathbf{S}_j^{(i)}}(\mathbf{x}) \, p_{\mathbf{N}}(\mathbf{y} - \mathbf{x}) \, d\mathbf{x} = \mathbb{E}\left\{ p_{\mathbf{N}}(\mathbf{y} - \mathbf{S}_j^{(i)}) \right\}$, where the expectation is over the PDF of $\mathbf{S}_j^{(i)}$. Then, the average probability of correct decision can be expressed as

$$\mathrm{P_c} = \sum_{i=1}^K v_i \sum_{j=0}^{M-1} \int_{\Gamma_j^{(i)}} \pi_j \, \mathbb{E}\left\{ p_{\mathbf{N}}(\mathbf{y} - \mathbf{S}_j^{(i)}) \right\} d\mathbf{y}. \tag{4}$$

[2]It is assumed that $v_i N_s$ is an integer for $i = 1, \ldots, K$. If not, the randomization factors can be achieved approximately. The approximation accuracy improves for larger N_s.

In practical systems, there is a constraint on the average power emitted from the transmitter. Under the framework of stochastic signaling and detector randomization (or time-sharing), this constraint on the average power can be expressed in the following form [7]: $\sum_{i=1}^K v_i \sum_{j=0}^{M-1} \pi_j \, \mathbb{E}\left\{ \|\mathbf{S}_j^{(i)}\|_2^2 \right\} \leq \mathrm{A}$, where A denotes the average power limit.

The probability of correct decision in (4) should be maximized over all possible decision rules (decision regions), randomization factors, and detector-specific signal PDFs that satisfy the average power constraint. For a given detector i and the corresponding signal PDFs, the optimal decision rule is specified by the MAP rule. The MAP decision rule selects symbol k if $k = \arg\max_{j \in \{0, 1, \ldots, M-1\}} \pi_j \, p_j^{(i)}(\mathbf{y})$, and it maximizes the average probability of correct decision [7]. Therefore, it is not necessary to search over all decision rules; only the MAP decision rule should be determined for each detector and its corresponding average probability of correct decision should be considered [3]. The average probability of correct decision for a generic decision rule is given in (3). Using the decision region for the MAP detector; i.e., $\Gamma_j^{(i)} = \{\mathbf{y} \in \mathbb{R}^N \mid \pi_j \, p_j^{(i)}(\mathbf{y}) \geq \pi_l \, p_l^{(i)}(\mathbf{y}), \, \forall l \neq j\}$, the average probability of correct decision for detector i becomes

$$\mathrm{P_{c,\,MAP}^{(i)}} = \int_{\mathbb{R}^N} \max_{j \in \{0, 1, \ldots, M-1\}} \left\{ \pi_j \, p_j^{(i)}(\mathbf{y}) \right\} d\mathbf{y}$$
$$= \int_{\mathbb{R}^N} \max_{j \in \{0, 1, \ldots, M-1\}} \left\{ \pi_j \, \mathbb{E}\{ p_{\mathbf{N}}(\mathbf{y} - \mathbf{S}_j^{(i)}) \} \right\} d\mathbf{y}. \tag{5}$$

Then, the optimal detector randomization and stochastic signal design problem can be stated as

$$\max_{\mathcal{B}} \sum_{i=1}^K v_i \int_{\mathbb{R}^N} \max_{j \in \{0, 1, \ldots, M-1\}} \left\{ \pi_j \, \mathbb{E}\left\{ p_{\mathbf{N}}(\mathbf{y} - \mathbf{S}_j^{(i)}) \right\} \right\} d\mathbf{y}$$

$$\text{subject to} \sum_{i=1}^K v_i \sum_{j=0}^{M-1} \pi_j \, \mathbb{E}\left\{ \|\mathbf{S}_j^{(i)}\|_2^2 \right\} \leq \mathrm{A}$$

$$\sum_{i=1}^K v_i = 1, \quad \mathbf{v} \succeq \mathbf{0} \tag{6}$$

where the optimization space is defined as $\mathcal{B} \triangleq \left\{ v_i, p_{\mathbf{S}_0^{(i)}}, p_{\mathbf{S}_1^{(i)}}, \ldots, p_{\mathbf{S}_{M-1}^{(i)}} \right\}_{i=1}^K$ and $\mathbf{v} \succeq \mathbf{0}$ means that $v_i \geq 0, \, \forall i \in \{1, 2, \ldots, K\}$. Note that there are also implicit constraints in the optimization problem in (6), since each $p_{\mathbf{S}_j^{(i)}}(\cdot)$ in $\mathcal{B}$ represents a PDF. Namely, $p_{\mathbf{S}_j^{(i)}}(\mathbf{x}) \geq 0$, $\forall \mathbf{x} \in \mathbb{R}^N$, and $\int_{\mathbb{R}^N} p_{\mathbf{S}_j^{(i)}}(\mathbf{x}) \, d\mathbf{x} = 1$ should also be satisfied $\forall j \in \{0, 1, \ldots, M-1\}$ and $\forall i \in \{1, \ldots, K\}$ by the optimal solution.

The formulation in (6) generalizes the previous studies in the literature and covers them as special cases. For example, for $K = 1$ (i.e., no detector randomization), it reduces to the problem in [3] (hence, $K \geq 2$ is considered in this study). On the other hand, when deterministic signals are considered; that is, $p_{\mathbf{S}_j^{(i)}}(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{s}_j^{(i)}), \forall i, j$, and when $M = 2$ (binary modulation), the problem in (6) reduces to that in [5].

The optimization problem in (6) provides a generic formulation that is valid for any noise PDF, and it is difficult to solve in general as the optimization needs to be performed over a space of signal PDFs. Let $\mathrm{P_c^\dagger}$ denote the maximum average

probability of correct decision obtained as the solution of the optimization problem in (6). To provide a simpler formulation of this problem, an upper bound on $P_c^\dagger$ will be derived first, and then the achievability of that bound will be investigated. To that aim, the following proposition is presented first.

Proposition 1 (Stochastic Signaling vs. Detector Randomization): *Considering the same average power constraint and the same statistics for the additive noise, stochastic signaling (without detector randomization) can never achieve a larger average probability of correct decision than detector randomization (without stochastic signaling) when optimal MAP detectors are employed in both cases.*

Proof: Consider an M-ary communications system in which the transmitter employs stochastic signaling and the receiver uses the corresponding MAP rule for detection (no detector randomization is performed). Suppose that the transmitted signal for each symbol is characterized with the PDF $p_{\mathbf{X}_j}(\cdot), \ \forall j \in \{0, 1, \ldots, M-1\}$. From (5), the average probability of correct decision for this system is given by

$$P_{\text{c, MAP}} = \int_{\mathbb{R}^N} \max_{j \in \{0,1,\ldots,M-1\}} \{\pi_j \, \mathbb{E}\{p_{\mathbf{N}}(\mathbf{y} - \mathbf{X}_j)\}\} \, d\mathbf{y} \tag{7}$$

where the expectations are taken with respect to the PDFs of $\mathbf{X}_j$'s. The transmitted signals for all the M symbols can be expressed as the elements of a random vector $\mathbf{X}$ as follows: $\mathbf{X} \triangleq [\mathbf{X}_0 \ \mathbf{X}_1 \cdots \mathbf{X}_{M-1}] \in \mathbb{R}^{MN}$, where $\mathbf{X}_j$'s are N dimensional row vectors $\forall j \in \{0, 1, \ldots, M-1\}$. Then, the following inequality follows directly from the definitions of the 'max' and 'expectation' operations.

$$\max_{j \in \{0, 1, \ldots, M-1\}} \{\pi_j \, \mathbb{E}\{p_{\mathbf{N}}(\mathbf{y} - \mathbf{X}_j)\}\} \leq$$
$$\mathbb{E}\left\{\max_{j \in \{0, 1, \ldots, M-1\}} \{\pi_j \, p_{\mathbf{N}}(\mathbf{y} - \mathbf{X}_j)\}\right\} \tag{8}$$

where the expectation on the right-hand-side is taken with respect to the PDF of $\mathbf{X}$. From (7) and (8), it is observed that

$$\int_{\mathbb{R}^N} \max_{j \in \{0,1,\ldots,M-1\}} \{\pi_j \, \mathbb{E}\{p_{\mathbf{N}}(\mathbf{y} - \mathbf{X}_j)\}\} \, d\mathbf{y} \leq$$
$$\mathbb{E}\left\{\int_{\mathbb{R}^N} \max_{j \in \{0,1,\ldots,M-1\}} \{\pi_j \, p_{\mathbf{N}}(\mathbf{y} - \mathbf{X}_j)\} \, d\mathbf{y}\right\} \triangleq \mathbb{E}\{F(\mathbf{X})\}. \tag{9}$$

Looking more closely at (9), it is observed that $F(\mathbf{x})$ represents the average probability of correct decision when the deterministic signal vector $\mathbf{x}$ is used for the transmission of M symbols over the additive noise channel and the corresponding MAP detector is employed at the receiver. Then, $\mathbb{E}\{F(\mathbf{X})\}$ can be interpreted as a randomization (or time-sharing) among MAP detectors. The exact number of MAP detectors is determined by the number of distinct values that the random vector $\mathbf{X}$ can take.[3] Hence, assuming the same average power constraint (see (6)), the average probability of

correct decision obtained by stochastic signaling with PDF $p_{\mathbf{X}}(\cdot)$ is always smaller than or equal to that of deterministic signaling and detector randomization according to the same PDF. ∎

Similarly to the proof of Proposition 1, we can express the transmitted signals for all the M symbols that are to be received by detector i as the elements of a random vector: $\mathbf{S}^{(i)} \triangleq \left[\mathbf{S}_0^{(i)} \ \mathbf{S}_1^{(i)} \ \ldots \ \mathbf{S}_{M-1}^{(i)}\right] \in \mathbb{R}^{MN}$, where $\mathbf{S}_j^{(i)}$'s are N dimensional row vectors. Then, the result in Proposition 1 can be employed to obtain a new optimization problem that provides an upper bound on the problem in (6). Specifically, instead of stochastic signals, consider detector randomization among deterministic signal values according to the joint signal PDF. Then, the inequality in (9) can be applied to the objective function in (6), and the following optimization problem can be obtained.

$$\max_{\{v_i, p_{\mathbf{S}^{(i)}}\}_{i=1}^K} \sum_{i=1}^K v_i$$
$$\times \mathbb{E}\left\{\int_{\mathbb{R}^N} \max_{j \in \{0,1,\ldots,M-1\}} \left\{\pi_j \, p_{\mathbf{N}}(\mathbf{y} - \mathbf{S}_j^{(i)})\right\} d\mathbf{y}\right\}$$
$$\text{subject to } \sum_{i=1}^K v_i \, \mathbb{E}\left\{\sum_{j=0}^{M-1} \pi_j \, \left\|\mathbf{S}_j^{(i)}\right\|_2^2\right\} \leq \mathrm{A},$$
$$\sum_{i=1}^K v_i = 1, \quad \mathbf{v} \succeq \mathbf{0} \tag{10}$$

where the expectations are taken with respect to the PDFs of $\mathbf{S}^{(i)}$'s. Proposition 1 implies that the solution to this optimization problem provides an upper bound on $P_c^\dagger$, which denotes the solution to the optimization problem in (6).

In order to achieve further simplification of the problem in (10), define $p_{\tilde{\mathbf{S}}}(\tilde{\mathbf{s}}) \triangleq \sum_{i=1}^K v_i \, p_{\mathbf{S}^{(i)}}(\tilde{\mathbf{s}})$, where $\tilde{\mathbf{s}} \triangleq [\tilde{\mathbf{s}}_0 \ \tilde{\mathbf{s}}_1 \cdots \tilde{\mathbf{s}}_{M-1}] \in \mathbb{R}^{MN}$. Since $\sum_{i=1}^K v_i = 1$, $v_i \geq 0 \ \forall i$ and $p_{\mathbf{S}^{(i)}}(\cdot)$'s are valid PDFs on $\mathbb{R}^{MN}$, $p_{\tilde{\mathbf{S}}}(\tilde{\mathbf{s}})$ satisfies the conditions to be a PDF. Then, the optimization problem in (10) can be written in the following equivalent form.

$$\max_{p_{\tilde{\mathbf{S}}}} \mathbb{E}\{G(\tilde{\mathbf{S}})\} \quad \text{subject to} \quad \mathbb{E}\{H(\tilde{\mathbf{S}})\} \leq \mathrm{A} \tag{11}$$

with $G(\tilde{\mathbf{S}}) \triangleq \int_{\mathbb{R}^N} \max_{j \in \{0, 1, \ldots, M-1\}} \left\{\pi_j \, p_{\mathbf{N}}(\mathbf{y} - \tilde{\mathbf{S}}_j)\right\} d\mathbf{y}$ and $H(\tilde{\mathbf{S}}) \triangleq \sum_{j=0}^{M-1} \pi_j \, \left\|\tilde{\mathbf{S}}_j\right\|_2^2$, where the expectations are taken with respect to $p_{\tilde{\mathbf{S}}}(\cdot)$, which denotes the joint PDF of transmitted signals for symbols $\{0, 1, \ldots, M-1\}$. In (11), $G(\tilde{\mathbf{s}})$ represents the average probability of correct decision when the deterministic signal vector $\tilde{\mathbf{s}}$ is used for the transmission of M symbols over the additive noise channel and the corresponding MAP detector is employed at the receiver. Therefore, $\mathbb{E}\{G(\tilde{\mathbf{S}})\}$ can be interpreted as a randomization (or time-sharing) among possibly infinitely many MAP detectors.[4] Let $P_c^\star$ denote the maximum average probability of correct decision obtained as the solution to the optimization problem in (11). From Proposition 1, $P_c^\star \geq P_c^\dagger$ is always satisfied.

Optimization problems similar to (11) have been studied before, for example in [2], [3], [5], [8]. Assuming that $G(\mathbf{s})$ in (11) is a continuous function and the components of $\mathbf{s}$ reside

---

[3]In fact, a randomization among two MAP detectors is always sufficient in practice since optimal stochastic signals can be represented by a randomization of at most two different signal values under an average power constraint [3]. In other words, for any stochastic signal PDF, a corresponding discrete probability distribution with at most two mass points can be obtained, and the corresponding MAP detector randomization can be performed according to that distribution.

[4]In the sequel, it will be shown that the optimal solution requires a randomization among at most two MAP detectors.

in finite closed intervals, the optimal solution of (11) can be represented by a randomization of at most two signal levels as a result of Carathéodory's theorem [9]; that is, $p_{\tilde{\mathbf{S}}}^{\text{opt}}(\tilde{\mathbf{s}}) = \lambda\,\delta(\tilde{\mathbf{s}} - \mathbf{s}_1) + (1-\lambda)\,\delta(\tilde{\mathbf{s}} - \mathbf{s}_2)$. Therefore, the problem in (11) can be solved over such signal PDFs, resulting in the following optimization problem:

$$\max_{\{\lambda,\,\mathbf{s}_1,\,\mathbf{s}_2\}} \quad \lambda\,G(\mathbf{s}_1) + (1-\lambda)\,G(\mathbf{s}_2)$$
$$\text{subject to} \quad \lambda\,H(\mathbf{s}_1) + (1-\lambda)\,H(\mathbf{s}_2) \ \leq \mathrm{A}\,, \ \lambda \in [0,1] \quad (12)$$

where

$$G(\mathbf{s}_k) = \int_{\mathbb{R}^N} \max_{j \in \{0,1,\dots,M-1\}} \{\pi_j\,p_{\mathbf{N}}(\mathbf{y} - \mathbf{s}_{k,j})\}\,\mathrm{d}\mathbf{y}\,,$$
$$H(\mathbf{s}_k) = \sum_{j=0}^{M-1} \pi_j\,\|\mathbf{s}_{k,j}\|_2^2\,,$$

and $\mathbf{s}_k = [\,\mathbf{s}_{k,0}\ \ \mathbf{s}_{k,1}\ \ \cdots\ \mathbf{s}_{k,M-1}\,] \in \mathbb{R}^{MN}$, with $\mathbf{s}_{k,j}$ being an N dimensional row vector $\forall j \in \{0,1,\dots,M-1\}$. Therefore, it is observed that the solution of (11) can be obtained by optimizing over a significantly reduced optimization space via (12). Optimization problems in the form of (12) have been investigated in various studies in the literature, such as [3], [5], [8], [10]. The main approaches in solving (12) include the analytical techniques as in [5] and [8], the convex relaxation technique to obtain an approximate solution in polynomial time as employed in [10], and the global optimization algorithms such as differential evolution (DE) and particle swarm optimization (PSO) [10]. In this study, a global optimization technique based on multistart and patternsearch algorithms from MATLAB's Global Optimization Toolbox are used to obtain the solution of (12).

It should be emphasized that detector randomization becomes useful in the presence of the average power constraint [2], [11], [12]. Without the constraint in (11), an optimal solution corresponding to the value of $\tilde{\mathbf{S}}$ that maximizes $G(\tilde{\mathbf{S}})$ can be obtained, meaning that a single detector can be employed to achieve the maximum probability of correct decision.[5]

Finally, the following proposition states that the maximum average probabilities of correct decision achieved by the solutions of the optimization problems in (6) and (12) are equal.

Proposition 2: *The optimization problems in* (6) *and* (12) *result in the same maximum value.*

Proof: First consider the optimization problem in (6) when $K = 2$ detectors are used and deterministic signaling is employed for each detector, that is, $p_{\mathbf{S}^{(1)}}(\mathbf{s}^{(1)}) = \delta(\mathbf{s}^{(1)} - \mathbf{s}_1)$ and $p_{\mathbf{S}^{(2)}}(\mathbf{s}^{(2)}) = \delta(\mathbf{s}^{(2)} - \mathbf{s}_2)$. In that case, (6) reduces to the optimization problem in (12); hence, (6) covers (12) as a special case. Therefore, the maximum value of the objective function in (6) should be larger than or equal to that of (12); namely, $\mathrm{P}_c^\dagger \geq \mathrm{P}_c^\star$. On the other hand, Proposition 1 implies that (11) (equivalently (12)) provides an upper bound on (6); that is, $\mathrm{P}_c^\dagger \leq \mathrm{P}_c^\star$. Therefore, it is concluded that $\mathrm{P}_c^\dagger = \mathrm{P}_c^\star$. ∎

---

[5]In fact, without any power constraints, transmitting at increasingly higher deterministic power levels for each symbol and using a single MAP detector at the receiver would be sufficient to bring the average error probability down to zero.

Proposition 2 implies that the solution of the original optimization problem in (6), which considers the joint optimization of stochastic signals and detector randomization, can be obtained as the solution of the much simpler optimization problem specified in (12). This also means that when multiple detectors are available for randomization (i.e., $K \geq 2$), it is sufficient to employ detector randomization for two deterministic signal vectors; i.e., there is no need to employ stochastic signaling to achieve the optimal solution. On the other hand, when there is only one detector (i.e., $K = 1$), the optimal solution may involve stochastic signaling, as investigated in [3]. All in all, the optimal solution to the most generic optimization problem in (6) results in either detector randomization for two deterministic signal values (for $K \geq 2$) or stochastic signaling without detector randomization (for $K = 1$).

III. Numerical Results and Conclusions

In this section, numerical examples are presented to compare the optimal solution obtained in the previous section against various signaling techniques investigated in [3] in terms of probability of error performance. A communications system specified as in (1) is considered with scalar observations and equal priors. It is assumed that the receiver is able to implement multiple detectors ($K \geq 2$) and to randomize among them. Also, a Gaussian mixture model with equal weights and variances is assumed for the noise, the PDF of which can be expressed as $p_N(n) = \sum_{i=1}^L \exp\{-(n - \mu_i)^2/(2\sigma^2)\}/(\sqrt{2\pi}\,\sigma L)$ [1]. Note that the average power of the noise can be calculated from $\mathbb{E}\{N^2\} = \sigma^2 + (1/L)\sum_{i=1}^L \mu_i^2$. In addition, the average power constraint is denoted as A. Similar to those introduced in [3], three different signaling schemes can be considered:

Gaussian Solution: Lacking any information about the noise PDF, the transmitter employs antipodal signals, $\{-\sqrt{\mathrm{A}}, \sqrt{\mathrm{A}}\}$, for the binary case (Fig. 2), and $\left\{\frac{-3\sqrt{\mathrm{A}}}{\sqrt{5}}, \frac{-\sqrt{\mathrm{A}}}{\sqrt{5}}, \frac{\sqrt{\mathrm{A}}}{\sqrt{5}}, \frac{3\sqrt{\mathrm{A}}}{\sqrt{5}}\right\}$ for the quaternary case (Fig. 3). In addition, the MAP decision rule is used at the receiver.

Optimal−Stochastic: In this scheme, no detector randomization is employed, and the optimal stochastic signals and the corresponding MAP detector are designed jointly [3]. In this case, the optimal signal for each symbol can be characterized by a discrete random variable with at most two mass points, and the solution can be obtained from an optimization problem similar to (9) of [3].

Optimal−Deterministic: This scheme does not employ any detector randomization or stochastic signaling, and obtains the optimal deterministic signal levels and the corresponding MAP detector [3].

In addition to the approaches described above, the following scheme investigated in the previous section is considered as the overall optimal solution:

Optimal Detector Randomization with Deterministic Signaling: This case refers to the solution of the most generic optimization problem in (6), which can be obtained from (12) as studied in the previous section.

In Fig. 2, the average error probabilities of the schemes described above are plotted versus A/σ^2 for $\mathrm{A} = 1$, where the parameters of the Gaussian mixture noise are given by $L = 6$ and $\boldsymbol{\mu} = [-1.08\ -0.81\ -0.27\ 0.27\ 0.81\ 1.08]$ as
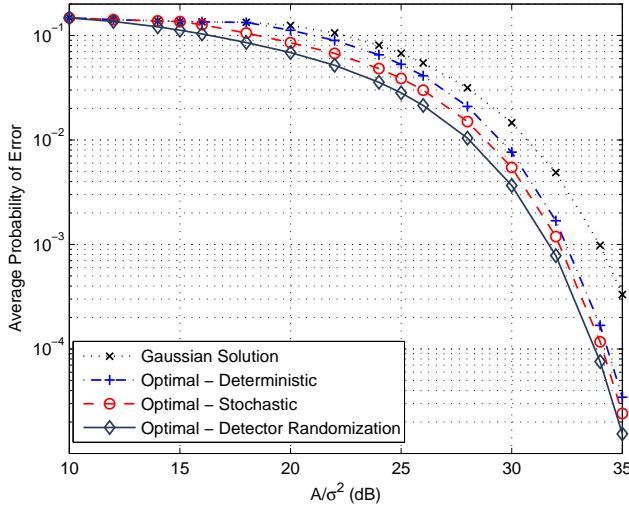
Fig. 2. Average probability of error versus A/σ^2 for various approaches.

Fig. 3. Average probability of error versus A/σ^2 for various approaches.

in [3]. However, unlike [3], symmetric signaling assumption is not employed. From the figure, it is observed that the Gaussian solution has the worst performance as expected since it is optimized for zero-mean Gaussian noise. Optimizing deterministic signal levels improves over the performance of the Gaussian solution, as observed from the Optimal–Deterministic curve. Further performance improvements are obtained when optimal stochastic signals are considered instead of deterministic signals (see Optimal–Stochastic). However, the best probability of error performance is achieved by the optimal solution of the most generic optimization problem investigated in the previous section, which performs optimal detector randomization among two MAP detectors corresponding to two deterministic signal pairs (see Optimal–Detector Randomization). In accordance with Proposition 1, stochastic signaling without detector randomization cannot perform better than detector randomization with deterministic signaling.

In Table I, some optimal signals are presented for the Optimal–Deterministic, Optimal–Stochastic and Optimal–Detector Randomization schemes in Fig. 2. For optimal deterministic signaling, $\mathbf{s}_0$ and $\mathbf{s}_1$ denote the optimal deterministic signal levels for symbol 0 and symbol 1, respectively. On the other hand, the optimal signal for symbol $i \in \{0, 1\}$ has the PDF in the form of $p_{\mathbf{S}}(\mathbf{s}) = \beta\,\delta(\mathbf{s} - \hat{\mathbf{s}}_{1,i}) + (1 - \beta)\,\delta(\mathbf{s} - \hat{\mathbf{s}}_{2,i})$ for optimal stochastic signaling. Finally, the optimal solution obtained in the previous section (Optimal–Detector Randomization) employs the signal pair $[\mathbf{s}_{1,0}\ \mathbf{s}_{1,1}]$ and the corresponding MAP detector with probability λ, and the signal pair $[\mathbf{s}_{2,0}\ \mathbf{s}_{2,1}]$ and the corresponding MAP detector with probability $1 - \lambda$. It is observed that all optimal signaling schemes get close to deterministic signaling for small A/σ^2, which is also verified from Fig. 2. However, the signaling schemes behave differently as A/σ^2 increases from 10 dB, which results in differences in probability of error performance.

Another example is provided to investigate the benefits that can be obtained via detector randomization in M-ary communication syste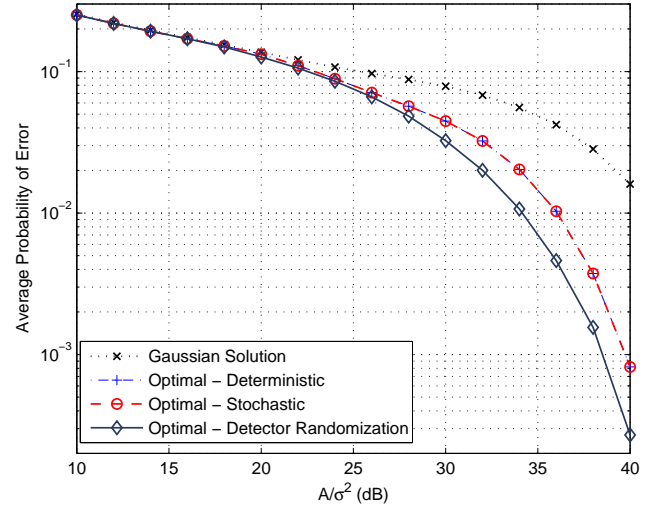ms when $M > 2$. For this purpose, a quaternary communications system ($M = 4$) with symmetric signaling is considered, that is, $\mathbf{s}_2 = -\mathbf{s}_0$ and $\mathbf{s}_3 = -\mathbf{s}_1$. The parameters of the Gaussian mixture noise are given by $L = 6$ and $\boldsymbol{\mu} = [-0.432\ -0.324\ -0.108\ 0.108\ 0.324\ 0.432]$. In Fig. 3, the average error probabilities of the schemes described previously are plotted versus A/σ^2 for $A = 1$. An important observation is that optimal stochastic signaling approach does not improve upon optimal deterministic signaling in this specific example. This is possibly due to the fact that signal space is overcrowded with the PDFs of four distinct symbols (each with six Gaussian mixture components) and there is no room left for any performance improvement via signal randomization after the optimal allocation of deterministic signal values. It is also evident from the figure that optimal detector randomization with deterministic signals achieves the best probability of error performance. Performance improvements among different signaling schemes deteriorate as A/σ^2 drops below 20 dB. The corresponding optimal signals are given in Table II. Note that the results only for symbols 0 and 1 are listed in the table, and the results for symbols 2 and 3 are the negatives of them respectively since symmetric signaling is considered.

The results in this study can also be applied to coded communications systems that employ hard-decision decoding at the receiver [13]. As the average error probability of the hard symbol detection is reduced via detector randomization or stochastic signaling, the overall error performance can be improved. On the other hand, when a soft-decision decoding unit is used at the receiver, the results of this study cannot be directly employed, and an updated formulation of the problem, which considers the average probability of error at the output of the decoding unit, may be required.

REFERENCES

[1] V. Bhatia and B. Mulgrew, "Non-parametric likelihood based channel estimator for Gaussian mixture noise," *Signal Processing*, vol. 87, pp. 2569–2586, Nov. 2007.
[2] 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.

TABLE I

OPTIMAL SIGNALS FOR THE SCENARIO IN FIG. 2

A/σ^2 (dB)	Deterministic sig.		Stochastic signaling					Detector randomization				
	s_0	s_1	β	$\hat{s}_{1,0}$	$\hat{s}_{2,0}$	$\hat{s}_{1,1}$	$\hat{s}_{2,1}$	λ	$s_{1,0}$	$s_{2,0}$	$s_{1,1}$	$s_{2,1}$
10	-0.9992	1.0008	1	-0.9992	N/A	1.0008	N/A	0.0386	-0.6441	-1.0119	0.6441	1.0119
15	-0.9997	1.0003	0.7212	-1.1417	-0.5766	0.9835	0.9836	0.4878	-0.7752	-1.1755	0.7752	1.1755
20	-1.3302	0.2552	0.3927	-1.4543	-0.6388	0.9600	0.9600	0.5417	-0.7756	-1.2127	0.7756	1.2127
25	-0.7622	0.7552	0.4594	-1.4135	-0.5590	0.9556	0.9556	0.4980	-1.1951	-0.7586	1.1951	0.7586
30	-0.8191	0.6761	0.5099	-0.8934	-0.8934	0.6014	1.4408	0.4644	-0.7488	-1.1763	0.7488	1.1763
35	-0.4886	0.9996	0.4892	-1.0150	-1.0150	0.4731	1.2977	0.4399	-0.7456	-1.1513	0.7456	1.1513

TABLE II

OPTIMAL SIGNALS FOR THE SCENARIO IN FIG. 3

A/σ^2 (dB)	Deterministic signaling		Detector randomization				
	s_0	s_1	λ	$s_{1,0}$	$s_{2,0}$	$s_{1,1}$	$s_{2,1}$
12	0.3762	1.3633	1	0.3762	N/A	1.3633	N/A
16	0.3957	1.3571	1	0.3957	N/A	1.3571	N/A
20	0.3066	1.3806	0.5058	0.2323	0.4659	1.2609	1.4654
24	0.3180	1.3780	0.2622	0.4836	0.3081	1.4854	1.3224
28	0.3154	1.3785	0.6570	0.4889	0.0720	1.4814	1.0790
32	0.3051	1.3809	0.7143	0.4826	0.0758	1.4547	1.0553
36	0.3002	1.3820	0.2206	0.0790	0.4745	1.0356	1.4268
40	0.2983	1.3553	0.1720	0.0802	0.4684	1.0204	1.4066

[3] ——, "Optimal signaling and detector design for power-constrained binary communications systems over non-Gaussian channels," *IEEE Commun. Letters*, vol. 14, no. 2, pp. 100–102, Feb. 2010.

[4] E. L. Lehmann, *Testing Statistical Hypotheses*, 2nd ed. New York: Chapman & Hall, 1986.

[5] A. Patel and B. Kosko, "Optimal noise benefits in Neyman-Pearson and inequality-constrained signal detection," *IEEE Trans. Sig. Processing*, vol. 57, no. 5, pp. 1655–1669, May 2009.

[6] H. Chen, P. K. Varshney, S. M. Kay, and J. H. Michels, "Theory of the stochastic resonance effect in signal detection: Part II–Variable detectors," *IEEE Trans. Sig. Processing*, vol. 56, no. 10, pp. 5031–5041, Oct. 2007.

[7] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1994.

[8] H. Chen, P. K. Varshney, S. M. Kay, and J. H. Michels, "Theory of the stochastic resonance effect in signal detection: Part I–Fixed detectors," *IEEE Trans. Sig. Processing*, vol. 55, no. 7, pp. 3172–3184, July 2007.

[9] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ: Princeton University Press, 1968.

[10] S. Bayram, S. Gezici, and H. V. Poor, "Noise enhanced hypothesis-testing in the restricted Bayesian framework," *IEEE Trans. Sig. Processing*, vol. 58, no. 8, pp. 3972–3989, Aug. 2010.

[11] S. Bayram and S. Gezici, "On the improvability and nonimprovability of detection via additional independent noise," *IEEE Sig. Processing Lett.*, vol. 16, no. 11, pp. 1001–1004, Nov. 2009.

[12] M. Azizoglu, "Convexity properties in binary detection problems," *IEEE Trans. Inform. Theory*, vol. 42, no. 4, pp. 1316–1321, July 1996.

[13] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.