

On the Optimality of Stochastic Signaling under an Average Power Constraint

Cagri Goken, Sinan Gezici, and Orhan Arikan
Department of Electrical and Electronics Engineering
Bilkent University, Bilkent, Ankara 06800, Turkey
E-mails: {goken,gezici,oarikan}@ee.bilkent.edu.tr

Abstract—In this paper, stochastic signaling is studied for scalar valued binary communications systems over additive noise channels in the presence of an average power constraint. For a given decision rule at the receiver, the effects of using stochastic signals for each symbol instead of conventional deterministic signals are investigated. First, sufficient conditions are derived to determine the cases in which stochastic signaling can or cannot outperform the conventional signaling. Then, statistical characterization of the optimal signals is provided and it is obtained that an optimal stochastic signal can be represented by a randomization of at most two different signal levels for each symbol. In addition, via global optimization techniques, the solution of the generic optimal stochastic signaling problem is obtained, and theoretical results are investigated via numerical examples.

Index Terms— Probability of error, additive noise channels, stochastic signaling, global optimization.

I. INTRODUCTION

In this study, optimal signaling strategies are investigated for minimizing the average probability of error in binary communications systems under an average power constraint. In the literature, optimal signaling in the presence of zero-mean Gaussian noise and under individual average power constraints in the form of $E\{|S_i|^2\} \leq A$ for $i = 0, 1$ has been studied extensively, and it is known that the average probability of error is minimized when deterministic antipodal signals ($S_0 = -S_1$) are used at the power limit ($|S_0|^2 = |S_1|^2 = A$) and a maximum *a posteriori* probability (MAP) decision rule is employed at the receiver [1]. However, in some scenarios, instead of individual average power constraints on each signal, there may be an average power constraint in the form of $\sum_{i=1}^2 \pi_i E\{|S_i|^2\} \leq A$, as considered in [2], where π_i represents the prior probability of symbol i . In [2], the optimal deterministic signaling is investigated for nonequal prior probabilities under such an average power constraint, when the noise is zero-mean Gaussian and the MAP decision rule is employed at the receiver. It is shown that the optimal signaling strategy is on-off keying for coherent receivers when the signals have nonnegative correlation and for noncoherent receivers with any arbitrary correlation value. In addition, it is also concluded from [2] that, for coherent systems, the best performance is achieved when the signals have a correlation of -1 and the power is distributed among the signals in such a way that the Euclidean distance between them is maximized under the given power constraint.

Although the optimal detector and signaling structures are well-known when the noise is Gaussian, the noise can have significantly different probability distribution than the Gaussian distribution in some cases due to effects such as jamming and interference [3]. When the noise is non-Gaussian, stochastic signaling can provide improved average probability of error performance as compared to the conventional deterministic signaling. The main difference of the optimal stochastic signaling approach from the conventional approach [1] is that signals S_0 and S_1 are considered as random variables in the former whereas they are regarded as deterministic quantities in the latter. In [4], the convexity properties of the average probability of error are investigated for binary-valued scalar signals in additive noise channels under an average power constraint and the problem of maximizing the average probability of error is studied for an average power constrained jammer. It is obtained that the optimal solution can be achieved when the jammer randomizes its power between at most two power levels. In [5], optimal stochastic signaling is studied under second and fourth moment constraints for a given decision rule (detector) at the receiver, and sufficient conditions are presented to determine whether stochastic signaling can provide performance improvements compared to deterministic signaling. Also, [6] investigates the joint design of the optimal stochastic signals and the detector, and illustrates the improvements that can be obtained via stochastic signaling. In addition, in [7], randomization between two deterministic signal pairs and the corresponding MAP decision rules is studied under the assumption that the receiver knows which deterministic signal pair is transmitted. It is shown that power randomization can result in significant performance improvements.

Although the effects of stochastic signaling have been investigated in [5] and [6] based on individual average power constraints for different symbols (that is, $E\{|S_0|^2\} \leq A$ and $E\{|S_1|^2\} \leq A$), no studies have considered stochastic signaling based on an average power constraint in the form of $\sum_{i=1}^2 \pi_i E\{|S_i|^2\} \leq A$. In this study, we consider this less strict average power constraint, and provide a generic formulation of the optimal stochastic signaling problem under that constraint. The formulation is valid for any given detector structure and noise probability distribution. First, sufficient conditions are derived to specify the cases in which the conventional signaling is optimal; that is, stochastic signal-

ing cannot provide any performance improvements over the conventional one. Then, sufficient conditions, under which the average probability of error performance of the conventional signaling can be improved via stochastic signaling, are obtained. In addition, the statistical structure of the optimal stochastic signals is investigated and it is shown that an optimal stochastic signal can be represented by a randomization between at most two signal levels for each symbol. In addition, by using a global optimization technique, named particle swarm optimization (PSO) [8], optimal stochastic signals are calculated, and numerical examples are presented to illustrate the theoretical results.

II. SYSTEM MODEL AND MOTIVATION

Consider a scalar binary communications system, as in [4] and [9], in which the received signal is given by

$$Y = S_i + N, \quad i \in \{0, 1\}, \quad (1)$$

where S_0 and S_1 denote the transmitted signal values for symbol 0 and symbol 1, respectively, and N is the noise component that is independent of S_i . In addition, the prior probabilities of the symbols, which are denoted by π_0 and π_1 , are supposed to be known [5].

Note that the probability distribution of the noise component in (1) is not necessarily Gaussian. Due to interference, such as multiple-access interference, the noise component can have a probability distribution that is different from the Gaussian distribution [3], [4].

A generic decision rule is considered at the receiver to estimate the symbol in (1). Specifically, for a given observation $Y = y$, the decision rule $\phi(y)$ is expressed as

$$\phi(y) = \begin{cases} 0, & y \in \Gamma_0 \\ 1, & y \in \Gamma_1 \end{cases}, \quad (2)$$

where Γ_0 and Γ_1 are the decision regions for symbol 0 and symbol 1, respectively [1].

In this study, the aim is to design signals S_0 and S_1 in (1) in order to minimize the average probability of error for a given decision rule, which is given by

$$P_{\text{avg}} = \pi_0 P_0(\Gamma_1) + \pi_1 P_1(\Gamma_0), \quad (3)$$

where $P_i(\Gamma_j)$ is the probability of selecting symbol j when symbol i is transmitted. In practical systems, the signal are commonly subject to an average power constraint, which can be expressed as

$$\pi_0 E\{|S_0|^2\} + \pi_1 E\{|S_1|^2\} \leq A, \quad (4)$$

where A is the average power limit. Therefore, the problem is to calculate the optimal probability density functions (PDFs) for signals S_0 and S_1 that minimize the average probability of error in (3) under the average power constraint in (4).

The main motivation for the optimal stochastic signaling problem is to enhance the error performance of a communications system by considering the signals at the transmitter

as random variables and obtaining the optimal probability distributions for those signals [4]-[7].

In the conventional signal design, S_0 and S_1 are considered as deterministic signals and they are designed in such a way that the Euclidean distance between them is maximized under the constraint in (4). In fact, when the effective noise has a zero-mean Gaussian PDF and the receiver employs the MAP decision rule, the probability of error is minimized when the Euclidean distance between the signals is maximized for a given average power constraint [1]. To that aim, S_0 and S_1 can conventionally be set to

$$S_0 = -\sqrt{A}/\alpha \text{ and } S_1 = \alpha\sqrt{A}, \quad (5)$$

where $\alpha \triangleq \sqrt{\pi_0/\pi_1}$ by considering the average power constraint in (4) (see [2] for the derivation). Then, the average probability of error in (3) becomes

$$P_{\text{avg}}^{\text{conv}} = \pi_0 \int_{\Gamma_1} p_N(y + \sqrt{A}/\alpha) dy + \pi_1 \int_{\Gamma_0} p_N(y - \alpha\sqrt{A}) dy, \quad (6)$$

where $p_N(\cdot)$ is the PDF of the noise in (1). Although the conventional signal design is optimal for certain classes of noise PDFs and decision rules, in some cases, the use of stochastic signals instead of deterministic ones can improve the system performance, as studied in the next section.

III. OPTIMAL STOCHASTIC SIGNALING

Instead of using constant levels for S_0 and S_1 as in the conventional case, one can consider a more generic scenario in which the signals can be stochastic. Then, the aim is to calculate the optimal PDFs for S_0 and S_1 in (1) that minimize the average probability of error under the constraint in (4).

Let $p_{S_0}(\cdot)$ and $p_{S_1}(\cdot)$ denote the PDFs for S_0 and S_1 , respectively. Then, from (3), the average probability of error for the decision rule in (2) is given by

$$P_{\text{avg}}^{\text{stoc}} = \sum_{i=0}^1 \pi_i \int_{-\infty}^{\infty} p_{S_i}(t) \int_{\Gamma_{1-i}} p_N(y-t) dy dt. \quad (7)$$

Therefore, the optimal stochastic signal design problem can be expressed as

$$\begin{aligned} \min_{p_{S_0}, p_{S_1}} P_{\text{avg}}^{\text{stoc}} \\ \text{subject to } \pi_0 E\{|S_0|^2\} + \pi_1 E\{|S_1|^2\} \leq A. \end{aligned} \quad (8)$$

After some manipulation, the objective function in (7) can be expressed as

$$P_{\text{avg}}^{\text{stoc}} = \pi_0 \int_{-\infty}^{\infty} p_{S_0}(x)(1-G(x))dx + \pi_1 \int_{-\infty}^{\infty} p_{S_1}(x)G(x)dx, \quad (9)$$

where $G(x)$ is defined as

$$G(x) \triangleq \int_{\Gamma_0} p_N(y-x) dy. \quad (10)$$

Then the expression in (9) can be written in terms of the expectation of $G(S_1)$ over S_1 and that of $G(S_0)$ over S_0 as

$$P_{\text{avg}}^{\text{stoc}} = \pi_0 - \pi_0 E\{G(S_0)\} + (1 - \pi_0) E\{G(S_1)\}. \quad (11)$$

Signals S_0 and S_1 can be expressed as the elements of a vector random variable \mathbf{S} as $\mathbf{S} \triangleq [S_0 \ S_1]$. Then the final form of optimization problem in (8) can be formulated as

$$\begin{aligned} \min_{p_{\mathbf{S}}} E\{F(\mathbf{S})\} \\ \text{subject to } E\{H(\mathbf{S})\} \leq A, \end{aligned} \quad (12)$$

where the expectations are taken over \mathbf{S} , $p_{\mathbf{S}}(\cdot)$ denotes the joint PDF of S_0 and S_1 ,

$$F(\mathbf{S}) \triangleq (1 - \pi_0) G(S_1) - \pi_0 G(S_0) + \pi_0, \quad (13)$$

and

$$H(\mathbf{S}) \triangleq (1 - \pi_0)|S_1|^2 + \pi_0|S_0|^2. \quad (14)$$

Note that there are also implicit constraints in the optimization problem in (12), since $p_{\mathbf{S}}(\mathbf{s})$ is a joint PDF.

A. On the Optimality of Conventional Signaling

In some cases, the conventional signaling is the optimal approach; that is, setting $p_{\mathbf{S}}(\mathbf{s}) = \delta(\mathbf{s} - \mathbf{S}_A)$, where $\mathbf{S}_A = [-\sqrt{A}/\alpha \ \alpha\sqrt{A}]$ with $\alpha = \sqrt{\pi_0/\pi_1}$, can solve the optimization problem in (12). In this section, we derive sufficient conditions that guarantee the optimality of the conventional signaling scheme.

Proposition 1: Assume that $G(x)$ in (10) is twice continuously differentiable. Then, $p_{\mathbf{S}}(\mathbf{s}) = \delta(\mathbf{s} - \mathbf{S}_A)$ is a solution of the optimization problem in (12), if the following three conditions are satisfied:

- $G(x)$ is a strictly decreasing function.
- $x G''(x) > 0, \forall x \neq 0$, and $G''(0) = 0$.
- For every (x_0, x_1) that satisfies

$$\pi_1 [G(\alpha\sqrt{A}) - G(x_1)] > \pi_0 [G(-\sqrt{A}/\alpha) - G(x_0)], \quad (15)$$

$$\pi_0 x_0^2 + \pi_1 x_1^2 > A \text{ is satisfied as well.}$$

Proof: In this proof, it is shown by contradiction that, when the conditions in the proposition are satisfied, there exist no signal PDFs that can result in a lower probability of error than the conventional signal \mathbf{S}_A under the given average power constraint. To that aim, it is first assumed that there exists a PDF $p_{\mathbf{S}}(\mathbf{s})$ for signal $\mathbf{S} = [S_0 \ S_1]$ such that $E\{F(\mathbf{S})\} < F(\mathbf{S}_A)$ and $E\{H(\mathbf{S})\} \leq A$. In other words, suppose that there exists a signal \mathbf{S} , with PDF $p_{\mathbf{S}}(\mathbf{s})$, which is better than the conventional signaling (see (12)). In addition, it is assumed without loss of generality that S_0 is a nonpositive and S_1 is a nonnegative random variable. [This assumption does not reduce the generality of the proof as $G(x)$ is a strictly decreasing function; hence, $F(\mathbf{S})$ in (13) is a strictly

increasing (decreasing) function of S_0 (S_1). Since the average power depends only on the absolute value of the signals, choosing nonpositive S_0 and nonnegative S_1 always achieves the minimum average probability of error. In other words, for each positive (negative) value of S_0 (S_1), its negative (positive) can be used instead, which results in smaller average probability of error and the same average power value.]

Under the assumptions above, if it is shown that there can exist no PDF $p_{\mathbf{S}}(\mathbf{s})$ for the signal $\mathbf{S} = [S_0 \ S_1]$, with S_0 being nonpositive and S_1 being nonnegative, that satisfies the three conditions in the proposition and $E\{F(\mathbf{S})\} < F(\mathbf{S}_A)$ under the average power constraint, it means that there can exist no signal PDF $p_{\mathbf{S}}(\mathbf{s})$ (for any signs of S_0 and S_1) that has lower probability of error than the conventional signal under the average power constraint. For that purpose, it is shown in the following that $F(\mathbf{x})$ in (13) is a convex function. Since $F(\mathbf{x}) = (1 - \pi_0)G(x_1) - \pi_0 G(x_0) + \pi_0$, its Hessian matrix can be obtained as

$$\begin{bmatrix} \frac{\partial^2 F}{\partial x_0^2} & \frac{\partial^2 F}{\partial x_0 \partial x_1} \\ \frac{\partial^2 F}{\partial x_1 \partial x_0} & \frac{\partial^2 F}{\partial x_1^2} \end{bmatrix} = \begin{bmatrix} -\pi_0 G''(x_0) & 0 \\ 0 & \pi_1 G''(x_1) \end{bmatrix}. \quad (16)$$

Since S_0 is a nonpositive random variable, x_0 can take only nonpositive values and similarly since S_1 is a nonnegative random variable, x_1 can take only nonnegative values. Therefore, under the second condition in the proposition, namely, $x G''(x) > 0, \forall x \neq 0$, and $G''(0) = 0$, the Hessian matrix is always positive semidefinite; hence, $F(\mathbf{x})$ is a convex function.

Since $F(\mathbf{S})$ is a convex function, Jensen's inequality implies that $E\{F(\mathbf{S})\} \geq F(E\{\mathbf{S}\}) = F([E\{S_0\} \ E\{S_1\}])$. Then, $E\{F(\mathbf{S})\} < F(\mathbf{S}_A)$ requires that $F([E\{S_0\} \ E\{S_1\}]) < F(\mathbf{S}_A)$, which can be expressed from (13) as

$$\begin{aligned} \pi_1 G(E\{S_1\}) - \pi_0 G(E\{S_0\}) < \pi_1 G(\alpha\sqrt{A}) \\ - \pi_0 G(-\sqrt{A}/\alpha). \end{aligned} \quad (17)$$

In addition, Jensen's inequality also implies that $E\{|S_0|^2\} \geq (E\{S_0\})^2$ and $E\{|S_1|^2\} \geq (E\{S_1\})^2$. Therefore, $E\{|S_0|^2\} + E\{|S_1|^2\} \geq (E\{S_0\})^2 + (E\{S_1\})^2$ is obtained. At this point, defining $x_0 = E\{S_0\}$ and $x_1 = E\{S_1\}$, and plugging them into (17) yields $\pi_1 [G(\alpha\sqrt{A}) - G(x_1)] > \pi_0 [G(-\sqrt{A}/\alpha) - G(x_0)]$, which is the first inequality in the third condition of the proposition. According to the third condition, whenever this inequality is satisfied for any (x_0, x_1) , $\pi_0 x_0^2 + \pi_1 x_1^2 > A$, equivalently, $\pi_0 E\{|S_0|^2\} + \pi_1 E\{|S_1|^2\} > A$, is also satisfied. Therefore, $E\{H(\mathbf{S})\} > A$ always holds, which indicates that the average power constraint in (12) is violated. Hence, it is concluded that when the conditions in Proposition 1 are satisfied, no PDF can achieve $E\{F(\mathbf{S})\} < F(\mathbf{S}_A)$ under the average power constraint. \square

As an example application of Proposition 1, consider a zero mean and unit variance Gaussian noise N in (1) with $p_N(x) = \exp\{-x^2/2\}/\sqrt{2\pi}$, and assume equal priors ($\pi_0 = \pi_1 = 0.5$). Also, the average power constraint A in (12) is taken to be 1. In this case, the conventional signaling becomes the antipodal signaling with $S_0 = -1$ and $S_1 = 1$, and a decision rule

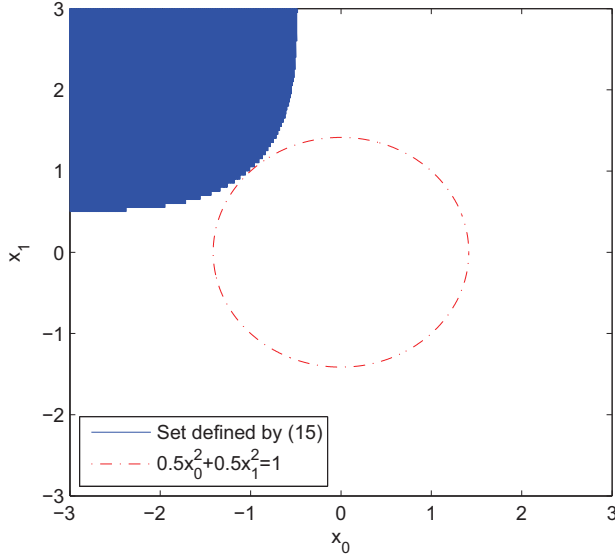


Fig. 1. The region in which the inequality $Q(x_1) - Q(x_0) < Q(1) - Q(-1)$ is satisfied is outside of the circle $0.5x_0^2 + 0.5x_1^2 = 1$.

of the form $\Gamma_0 = (-\infty, 0]$ and $\Gamma_1 = [0, \infty)$; that is, the sign detector, is the optimal MAP decision rule. Then, $G(x)$ in (10) can be calculated as $G(x) = Q(x)$, where $Q(x) = (\int_x^\infty e^{-t^2/2} dt) / \sqrt{2\pi}$ defines the Q -function. Since $Q(x)$ is a monotone decreasing function and $xQ''(x) > 0, \forall x \neq 0$ with $Q''(0) = 0$, the first two conditions in Proposition 1 are satisfied. For the third condition, we need to check the region in which $Q(x_1) - Q(x_0) < Q(1) - Q(-1) = -0.6827$. Then, as $Q(x)$ is a decreasing function, if one can find (a, b) such that $Q(a) = Q(b) - 0.6827$, then for every $x_1 > a$ and $x_0 = b$, $Q(x_1) - Q(x_0) < -0.6827$ and $0.5x_0^2 + 0.5x_1^2 > 0.5a^2 + 0.5b^2$. Also, since the Q -function takes values only between 0 and 1, $b < -0.475$ should hold. A simple search on this region reveals that $0.5a^2 + 0.5b^2 \geq 1$, where the equality holds only at $(a, b) = (-1, 1)$. This fact can be observed from Fig. 1 as well. The geometrical interpretation of the third condition in Proposition 1 is that the set of all (x_0, x_1) pairs that satisfy $\pi_1[G(\alpha\sqrt{A}) - G(x_1)] > \pi_0[G(-\sqrt{A}/\alpha) - G(x_0)]$ should be completely outside of the elliptical region whose boundary is $\pi_0 x_0^2 + \pi_1 x_1^2 = A$. In Fig. 1, this is shown for this example and it is observed that every point that satisfies the inequality $Q(x_1) - Q(x_0) < Q(1) - Q(-1)$, is located outside of the circle $0.5x_0^2 + 0.5x_1^2 = 1$. Thus, the third condition in Proposition 1 holds as well. Therefore, it is guaranteed that the conventional signaling is optimal in this scenario.

B. Sufficient Conditions for Improvability

In this section, we obtain sufficient conditions under which the performance of the conventional signaling approach can be improved via stochastic signaling.

Proposition 2: Assume that $G(x)$ in (10) is twice continuously differentiable. Then, $p_S(s) = \delta(s - S_A)$ is not an optimal

solution of (12) if

$$G''(\alpha\sqrt{A}) < \frac{G'(\alpha\sqrt{A})}{\alpha\sqrt{A}}, \quad (18)$$

or, alternatively,

$$G''(-\sqrt{A}/\alpha) > \frac{G'(-\sqrt{A}/\alpha)}{-\sqrt{A}/\alpha}. \quad (19)$$

Proof: In order to prove the suboptimality of the conventional solution $p_S(s) = \delta(s - S_A)$, it is shown that, under the conditions in the proposition, there exist $\lambda \in (0, 1)$, Δ_1 , Δ_2 , Δ_3 , and Δ_4 such that¹

$$p_{S_2}(s) = \lambda \delta(s - (S_A + \epsilon_1)) + (1 - \lambda) \delta(s - (S_A + \epsilon_2)), \quad (20)$$

where $\epsilon_1 = [\Delta_1 \ \Delta_2]$ and $\epsilon_2 = [\Delta_3 \ \Delta_4]$, yields a lower probability of error than $p_S(s)$ and satisfies the constraint in (12). Specifically, proving the existence of $\lambda \in (0, 1)$, Δ_1 , Δ_2 , Δ_3 , and Δ_4 that satisfy

$$\lambda F(S_A + \epsilon_1) + (1 - \lambda) F(S_A + \epsilon_2) < F(S_A) \quad (21)$$

and

$$\begin{aligned} & \pi_0 [\lambda (-\sqrt{A}/\alpha + \Delta_1)^2 + (1 - \lambda) (-\sqrt{A}/\alpha + \Delta_3)^2] + \\ & \pi_1 [\lambda (\alpha\sqrt{A} + \Delta_2)^2 + (1 - \lambda) (\alpha\sqrt{A} + \Delta_4)^2] = A \end{aligned} \quad (22)$$

is sufficient to prove that the conventional signaling is not optimal. From (22), the following equation is obtained:

$$\begin{aligned} & [\pi_0 (\lambda \Delta_1^2 + (1 - \lambda) \Delta_3^2) + \pi_1 (\lambda \Delta_2^2 + (1 - \lambda) \Delta_4^2)] / \sqrt{A} \\ & = -2 \left[\pi_1 (\lambda \Delta_2 \alpha + (1 - \lambda) \Delta_4 \alpha) \right. \\ & \quad \left. - \pi_0 \left(\frac{\Delta_1 \lambda}{\alpha} + \frac{(1 - \lambda) \Delta_3}{\alpha} \right) \right]. \end{aligned} \quad (23)$$

Since the left-hand-side of the equality in (23) is always positive, the term on the right-hand-side should also be positive, which leads to the following inequality since $\alpha = \sqrt{\pi_0/\pi_1}$:

$$\lambda \Delta_2 + (1 - \lambda) \Delta_4 < \lambda \Delta_1 + (1 - \lambda) \Delta_3. \quad (24)$$

In addition, from (13) and (21), the following inequality is obtained:

$$\begin{aligned} & \lambda \pi_1 G(\alpha\sqrt{A} + \Delta_2) + (1 - \lambda) \pi_1 G(\alpha\sqrt{A} + \Delta_4) \\ & - \lambda \pi_0 G(-\sqrt{A}/\alpha + \Delta_1) - (1 - \lambda) \pi_0 G(-\sqrt{A}/\alpha + \Delta_3) \\ & < \pi_1 G(\alpha\sqrt{A}) - \pi_0 G(-\sqrt{A}/\alpha). \end{aligned} \quad (25)$$

For infinitesimally small Δ_1 , Δ_2 , Δ_3 and Δ_4 , the first three terms of the Taylor series expansion for $G(\alpha\sqrt{A} + \Delta_2)$, $G(\alpha\sqrt{A} + \Delta_4)$, $G(-\sqrt{A}/\alpha + \Delta_1)$ and $G(-\sqrt{A}/\alpha + \Delta_3)$ can

¹It is assumed that Δ_1 , Δ_2 , Δ_3 , and Δ_4 are not all zeros, since that would result in the conventional signaling.

be used to approximate (25) as

$$\begin{aligned} & G'(\alpha\sqrt{A})[\lambda\pi_1\Delta_2 + (1-\lambda)\pi_1\Delta_4] \\ & + G'(-\sqrt{A}/\alpha)[- \lambda\pi_0\Delta_1 - (1-\lambda)\pi_0\Delta_3] \\ & + \frac{G''(\alpha\sqrt{A})}{2}[\lambda\pi_1\Delta_2^2 + (1-\lambda)\pi_1\Delta_4^2] \\ & + \frac{G''(-\sqrt{A}/\alpha)}{2}[- \lambda\pi_0\Delta_1^2 - (1-\lambda)\pi_0\Delta_3^2] < 0. \end{aligned} \quad (26)$$

For $\Delta_1 = \Delta_3 = 0$, (24) becomes $\lambda\Delta_2 + (1-\lambda)\Delta_4 < 0$ and (23) becomes $\pi_1(\lambda\Delta_2^2 + (1-\lambda)\Delta_4^2) = -2\sqrt{A}\pi_1\pi_0(\lambda\Delta_2 + (1-\lambda)\Delta_4)$. Then, (26) simplifies to

$$\begin{aligned} & G'(\alpha\sqrt{A})[\lambda\pi_1\Delta_2 + (1-\lambda)\pi_1\Delta_4] \\ & + G''(\alpha\sqrt{A})[-\sqrt{A\pi_0\pi_1}(\lambda\Delta_2 + (1-\lambda)\Delta_4)] < 0. \end{aligned} \quad (27)$$

Since $\lambda\Delta_2 + (1-\lambda)\Delta_4 < 0$, (27) implies that $G'(\alpha\sqrt{A})\pi_1 - G''(\alpha\sqrt{A})\sqrt{A\pi_0\pi_1} > 0$, which is equivalent to $G'(\alpha\sqrt{A}) - G''(\alpha\sqrt{A})(\alpha\sqrt{A}) > 0$; that is, the first condition in the proposition.

Similarly, for $\Delta_2 = \Delta_4 = 0$, (24) becomes $\lambda\Delta_1 + (1-\lambda)\Delta_3 > 0$ and (23) becomes $\pi_0(\lambda\Delta_1^2 + (1-\lambda)\Delta_3^2) = 2\sqrt{A}\pi_1\pi_0(\lambda\Delta_1 + (1-\lambda)\Delta_3)$. Then, (26) can be rewritten as follows:

$$\begin{aligned} & G'(-\sqrt{A}/\alpha)[- \lambda\pi_0\Delta_1 - (1-\lambda)\pi_0\Delta_3] \\ & + G''(-\sqrt{A}/\alpha)[- \sqrt{A\pi_0\pi_1}(\lambda\Delta_1 + (1-\lambda)\Delta_3)] < 0. \end{aligned} \quad (28)$$

Since $\lambda\Delta_1 + (1-\lambda)\Delta_3 > 0$, (28) becomes $G'(-\sqrt{A}/\alpha)\pi_0 + G''(-\sqrt{A}/\alpha)\sqrt{A\pi_0\pi_1} > 0$, which is equivalent to $G'(-\sqrt{A}/\alpha) + G''(-\sqrt{A}/\alpha)(\sqrt{A}/\alpha) > 0$. Hence, the second condition in the proposition is obtained.

This proof indicates that that $p_{S_2}(\mathbf{s})$ in (20) can result in a lower probability of error than the conventional signaling for infinitesimally small Δ_2 and Δ_4 values along with $\Delta_1 = \Delta_3 = 0$, or, for infinitesimally small Δ_1 and Δ_3 values along with $\Delta_2 = \Delta_4 = 0$, which satisfy (23). \square

Proposition 2 provides simple sufficient conditions to determine if stochastic signaling can improve the probability of error performance of a given detector. A practical example is presented in Section IV on the use of the results in the proposition.

C. Statistical Characteristics of Optimal Signals

The optimization problem in (12) may be difficult to solve in general since the optimization needs to be performed over a space of PDFs. However, by using the following result, that optimization problem can be formulated over a set of variables instead of functions, hence can be simplified to a great extent.

Lemma 1: *Assume that $G(x)$ in (10) is a continuous function and possible signal values for S_0 and S_1 reside in $[-\gamma, \gamma]$ for some finite $\gamma > 0$. Then, the solution of the optimization problem in (12) is in the form of*

$$p_{\mathbf{S}}(\mathbf{s}) = \lambda\delta(\mathbf{s} - \mathbf{s}_1) + (1-\lambda)\delta(\mathbf{s} - \mathbf{s}_2), \quad (29)$$

where $\lambda \in [0, 1]$ and \mathbf{s}_i is two-dimensional vector for $i = 1, 2$.

Proof: Optimization problems in the form of (12) have been investigated in various studies in the literature [7], [10], [11], [12]. Under the conditions in the lemma, the optimal solution of (12) can be represented by a randomization of at most two signal levels as a result of Carathéodory's theorem [13], [14]. Hence, the optimal signal PDF can be expressed as in (29). \square

Lemma 1 states that the optimal signal PDF that solves the optimization problem in (12) can be represented by a discrete probability distribution with at most two mass points. Therefore, the optimization problem in (12) can be simplified as follows:

$$\begin{aligned} & \min_{\lambda, \mathbf{s}_1, \mathbf{s}_2} \lambda F(\mathbf{s}_1) + (1-\lambda)F(\mathbf{s}_2) \\ & \text{subject to } \lambda H(\mathbf{s}_1) + (1-\lambda)H(\mathbf{s}_2) \leq A. \end{aligned} \quad (30)$$

In other words, instead of optimization over functions, an optimization over a five-dimensional space (two two-dimensional mass points, \mathbf{s}_1 and \mathbf{s}_2 , plus the weight, λ) can be considered for the optimal signaling problem as a result of Lemma 1.

Although (30) is significantly simpler than (12), it can still be a nonconvex optimization problem in general. Therefore, global optimization techniques such as particle-swarm optimization (PSO) [8], [15], [16], genetic algorithms and differential evolution [17], can be used to obtain the optimal solution [11], [18]. In the next section, the PSO algorithm is used to calculate the optimal stochastic signals in the numerical examples. For the details of the PSO algorithm, please refer to [8] and for the PSO parameters used in PSO approach on this paper, please refer to [6].

IV. NUMERICAL RESULTS

In this section, a numerical example is presented to show the improvements over conventional signaling via optimal stochastic signaling. For this example, a binary communications system with priors $\pi_0 = 0.2$ and $\pi_1 = 0.8$ is considered [2]. Hence $\alpha = \sqrt{\pi_0/\pi_1}$ is equal to 0.5 in this case. Also, the average power constraint A is set to 1. It is assumed that the receiver employs a simple threshold detector such that $\Gamma_0 = (-\infty, \tau)$ and $\Gamma_1 = (\tau, \infty)$, where $\tau = (2\sigma^2 \ln(0.25) - 3.75)/5$. In fact, this is the optimal MAP decision rule for given the prior probabilities and the average power constraint, when the conventional signaling is performed and the noise is zero-mean Gaussian noise with variance σ^2 .

In this example, the effective noise in (1) is modeled by Gaussian mixture noise [3], whose PDF can be expressed as

$$p_N(y) = \frac{1}{\sqrt{2\pi}\sigma} \sum_{l=1}^L \mathbf{v}_l e^{-\frac{(y-\mu_l)^2}{2\sigma^2}}. \quad (31)$$

By using this noise model, and the receiver structure specified above, $G(x)$ in (10) can be obtained as

$$G(x) = \sum_{l=1}^L \mathbf{v}_l Q\left(\frac{-\tau + x + \mu_l}{\sigma}\right). \quad (32)$$

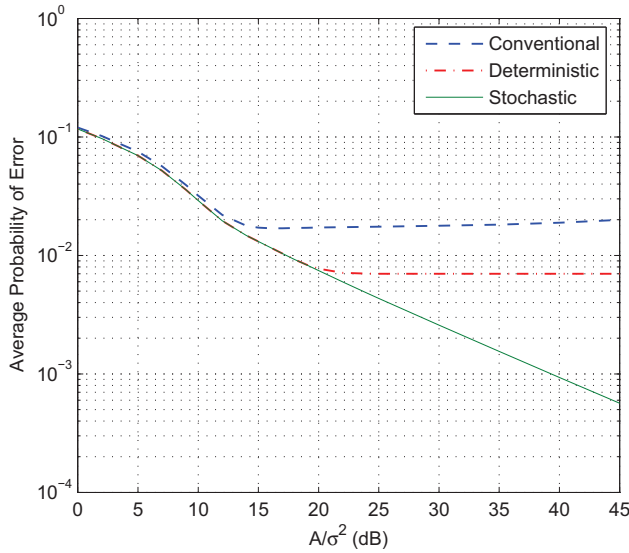


Fig. 2. Average probability of error versus A/σ^2 for conventional, optimal deterministic, and optimal stochastic signaling.

In the numerical example, $\mathbf{v} = [0.465 \ 0.035 \ 0.035 \ 0.465]$ and $\boldsymbol{\mu} = [-1.251 \ -0.7 \ 0.7 \ 1.251]$ are used. Gaussian mixture noise is encountered in practical systems in the presence of interference [3]. Note that the variance of each component of the Gaussian mixture noise is set to σ^2 and the average power of the noise can be calculated as $E\{N^2\} = \sigma^2 + 1.4898$ for the given values.

In this example, three different signaling schemes are considered:

Conventional Signaling: In this case, the transmitter selects the signals as $S_0 = -\sqrt{A}/\alpha = -2$ and $S_1 = \sqrt{A}\alpha = 0.5$, which are known to be optimal if the noise is zero-mean Gaussian and the receiver structure is as specified above [1].

Stochastic Signaling: In this case, the solution of the most generic optimization problem in (8) is obtained. Since that problem can be reduced to the optimization problem in (30), the optimal stochastic signals are calculated via PSO based on the formulation (30) in this scenario.

Deterministic Signaling: In this case, it is assumed that the signals are deterministic, and the optimization problem in (12) is solved under that assumption. That is, the optimal signal PDF is given by $p_S(\mathbf{s}) = \delta(\mathbf{s} - \mathbf{s}^*)$, where \mathbf{s}^* is the solution of the following optimization problem:

$$\begin{aligned} \min_{\mathbf{s}} F(\mathbf{s}) \\ \text{subject to } H(\mathbf{s}) \leq A. \end{aligned} \quad (33)$$

In other words, this solution provides a simplified version of the optimal solution in (12). Indeed, there are two optimization variables (two signal levels, S_0 and S_1) in this case, instead of the five optimization variables in the stochastic signaling case (see (30)).

TABLE I
OPTIMAL STOCHASTIC SIGNALING.

A/σ^2 (dB)	λ	S_{11}	S_{12}	S_{21}	S_{22}
0	1	-1.8221	0.6480	N/A	N/A
15	1	-1.8424	0.6336	N/A	N/A
30	0.3149	-1.5467	0.5782	-2.0607	0.5782
45	0.0733	-1.4702	0.5185	-2.0159	0.5185

TABLE II
OPTIMAL DETERMINISTIC SIGNALING.

A/σ^2 (dB)	S_0	S_1
0	-1.8221	0.6480
15	-1.8424	0.6336
30	-1.6911	0.7314
45	-1.6249	0.7306

In Fig. 2, the average probabilities of error are plotted versus A/σ^2 for the three signaling schemes. In order to calculate both the stochastic signaling and the deterministic signaling solutions, the PSO approach is used. From Fig. 2, it is observed that for low values of σ , the conventional signaling performs worse than the others, and the stochastic signaling achieves the lowest probabilities of error. Specifically, after A/σ^2 exceeds 30 dB, significant improvements can be obtained via stochastic signaling over the conventional and deterministic signaling approaches. Indeed, improvements are expected based on Proposition 2 as well. For example, at 30 dB, $G''(-2) = 0.6514$ and $G'(-2) = -0.441$, and at 40 dB, $G''(-2) = 13.84$ and $G'(-2) = -1.389$, which results in $G''(-2) > -G'(-2)/2$ for both of the cases. Therefore, the second sufficient condition in Proposition 2 (i.e., the inequality in (19)) is satisfied and improvements over the conventional solution are guaranteed in those scenarios.

Moreover, it should be noted that the average probability of error does not monotonically decrease for the conventional and deterministic solutions as A/σ^2 increases. This is because of the fact that average probability of error is related to the area under the two shifted noise PDFs as in (6). Since the noise PDF has a multimodal PDF in this example, and the amount of shifts that can be imposed on the noise PDFs is restricted by the average power constraint, that area may increase or remain same as A/σ^2 increases in some cases.

In order to provide further explanations of the results, Table 1 and 2 present the solutions of the stochastic and deterministic signaling schemes for some A/σ^2 values. In Table 1, the optimal \mathbf{s}_1 and \mathbf{s}_2 in (30) are expressed as $\mathbf{s}_1 = [s_{11} \ s_{12}]$ and $\mathbf{s}_2 = [s_{21} \ s_{22}]$ for each A/σ^2 value. For small A/σ^2 values, such as 0 dB and 15 dB, the deterministic solutions are the same as the stochastic ones. In fact, the performance of the deterministic and the stochastic signaling is same for A/σ^2 values less than 20 dB, as can be observed from Fig. 2. Also, their performance is very close to the performance of conventional signaling at high σ values. For example, at 0 dB, the average probability of error for the conventional signaling is 0.120, and it is 0.117 for the other schemes.

Furthermore, it can be observed from Table 1 that as A/σ^2 increases, the randomization between two signal vectors becomes more effective and this helps reduce the average probability of error as compared with the other signaling schemes. For example, at $A/\sigma^2 = 45$ dB, the average probability of error for the stochastic signaling is 5.66×10^{-4} , whereas it is 0.007 and 0.02 for the deterministic signaling and the conventional signaling schemes, respectively.

V. CONCLUDING REMARKS

The optimal stochastic signaling problem has been studied under an average power constraint. It has been shown that, under certain conditions, the conventional signaling approach, which maximizes the Euclidean distance between the signals, is the optimal signaling strategy. Also, sufficient conditions have been obtained to specify when randomization between different signal values may result in improved performance in terms of the average probability of error. In addition, the discrete structure of the optimal stochastic signals has been specified, and a global optimization technique, called PSO, has been used to solve the generic stochastic signaling problem under the average power constraint. Finally, numerical examples have been presented to illustrate some applications of the theoretical results.

REFERENCES

- [1] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1994.
- [2] I. Korn, J. P. Fonseka, and S. Xing, "Optimal binary communication with nonequal probabilities," *IEEE Trans. Commun.*, vol. 51, no. 9, pp. 1435–1438, Sep. 2003.
- [3] V. Bhatia and B. Mulgrew, "Non-parametric likelihood based channel estimator for Gaussian mixture noise," *Signal Processing*, vol. 87, pp. 2569–2586, Nov. 2007.
- [4] M. Azizoglu, "Convexity properties in binary detection problems," *IEEE Trans. Inform. Theory*, vol. 42, no. 4, pp. 1316–1321, July 1996.
- [5] C. Goken, S. Gezici, and O. Arikan, "Stochastic signaling under second and fourth moment constraints," in *Proc. IEEE International Workshop on Signal Processing Advances for Wireless Communications (SPAWC)*, Marrakech, Morocco, June 2010.
- [6] —, "Optimal signaling and detector design for power-constrained binary communications systems over non-Gaussian channels," *IEEE Commun. Lett.*, vol. 14, no. 2, pp. 100–102, Feb. 2010.
- [7] 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.
- [8] K. E. Parsopoulos and M. N. Vrahatis, *Particle swarm optimization method for constrained optimization problems*. IOS Press, 2002, pp. 214–220, in *Intelligent Technologies—Theory and Applications: New Trends in Intelligent Technologies*.
- [9] S. Shamai and S. Verdú, "Worst-case power-constrained noise for binary-input channels," *IEEE Trans. Inform. Theory*, vol. 38, pp. 1494–1511, Sep. 1992.
- [10] 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.
- [11] 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, Aug. 2010.
- [12] L. Huang and M. J. Neely, "The optimality of two prices: Maximizing revenue in a stochastic network," in *Proc. 45th Annual Allerton Conference on Commun., Control, and Computing*, Monticello, IL, Sep. 2007.
- [13] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*. Berlin:Springer-Verlag, 2004.
- [14] D. P. Bertsekas, A. Nedic, and A. E. Ozdaglar, *Convex Analysis and Optimization*. Boston, MA: Athena Specific, 2003.
- [15] S. Koziel and Z. Michalewicz, "Evolutionary algorithms, homomorphous mappings, and constrained parameter optimization," *Evolutionary Computation*, vol. 7, no. 1, pp. 19–44, 1999.
- [16] X. Hu and R. Eberhart, "Solving constrained nonlinear optimization problems with particle swarm optimization," in *Proc. Sixth World Multiconference on Systemics, Cybernetics and Informatics 2002 (SCI 2002)*, Orlando, FL, 2002.
- [17] K. V. Price, R. M. Storn, and J. A. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization*. New York: Springer, 2005.
- [18] S. Bayram and S. Gezici, "Noise-enhanced M -ary hypothesis-testing in the minimax framework," in *Proc. International Conference on Signal Processing and Commun. Systems*, Omaha, Nebraska, Sep. 2009, pp. 31–36.