STOCHASTIC SIGNALING UNDER SECOND AND FOURTH MOMENT CONSTRAINTS

Cagri Goken, Sinan Gezici, Orhan Arikan

Department of Electrical and Electronics Engineering
Bilkent University, Bilkent, Ankara 06800, Turkey
{goken,gezici,oarikan}@ee.bilkent.edu.tr

ABSTRACT

Stochastic signaling is investigated under second and fourth moment constraints for the detection of scalar-valued binary signals in additive noise channels. Sufficient conditions are derived to determine when the use of stochastic signals instead of deterministic ones can or cannot enhance the error performance of a given binary communications system. Also, a convex relaxation approach is employed to obtain approximate solutions of the optimal stochastic signaling problem. Finally, numerical examples are presented, and extensions of the results to M-ary communications systems and to other criteria than the average probability of error are discussed.
***Index Terms*–** Probability of error, additive noise channels, stochastic signaling, convex optimization.

1. INTRODUCTION

In this study, the optimal signaling approach is investigated for minimizing the average probability of error of a binary communications system under second and fourth moment constraints. Optimal signaling in the presence of zero-mean Gaussian noise has been studied extensively [1], [2]. It is known that deterministic antipodal signals; that is, $S_1 = -S_0$, minimize the average probability of error of a binary communications system in additive Gaussian noise channels. Also, for vector observations, selecting the deterministic signals along the eigenvector of the covariance matrix of the Gaussian noise corresponding to the minimum eigenvalue minimizes the average probability of error under power constraints in the form of $\|\mathbf{S}_0\|^2 \leq A$ and $\|\mathbf{S}_1\|^2 \leq A$ [2]. Although the average probability of error expressions and optimal signaling techniques have been investigated for Gaussian noise, the noise can have significantly different probability distribution than the Gaussian distribution in some cases due to effects such as multiuser interference and jamming [3], [4]. In [5], additive noise channels with binary inputs and scalar outputs are studied, and it is proven that the least-favorable noise distribution that maximizes the average probability of error and minimizes the channel capacity is a mixture of discrete lattices [5]. A similar problem is investigated in [6] for a binary communications system in the presence of an additive jammer, and the properties of optimal jammer distribution and signal distribution are obtained.

The convexity properties of the average probability of error are investigated in [3] for binary-valued scalar signals in additive noise channels under an average power constraint. It is proven that the average probability of error is a convex non-increasing function for unimodal differentiable noise probability density functions (PDFs) and for maximum likelihood (ML) receivers. Then, it is concluded that random-

ization of signal values (or, stochastic signaling) cannot improve error performance for the considered communications system. Also, the problem of maximizing the average probability of error is studied for an average power constrained jammer, and it is obtained that the optimal solution can be achieved when the jammer randomizes its power between at most two power levels [3]. In a related study [7], optimal randomization of signal amplitude is investigated for an average power constrained antipodal binary communications system that employs an ML receiver. Similar to [3], the optimal signal is shown to be a randomization of at most two signal levels.

Although the average probability of error for a binary communications system is minimized by *deterministic* antipodal signals in additive Gaussian noise channels [2], the studies in [3], [6], [7] imply that *stochastic* signaling can provide lower average probabilities of error in some cases when the noise is non-Gaussian. Hence, a generic formulation of the optimal signaling problem for binary communications systems can be stated as the calculation of optimal probability distributions for signals S_0 and S_1 such that the average probability of error of the system is minimized under certain constraints on the moments of S_0 and S_1. The main difference of this optimal stochastic signaling approach from the conventional (deterministic) approach [1], [2] is that signals S_0 and S_1 are modeled as random variables in the former whereas they are considered as deterministic quantities in the latter.

In this paper, a generic formulation of the optimal stochastic signaling problem is considered, which is valid for any receiver structure and noise probability distribution. Also, both average power and peakedness constraints are imposed on the signals. In addition, sufficient conditions are derived to determine if the error performance of a receiver can or cannot be improved by using stochastic signaling instead of conventional signaling. Furthermore, an optimization theoretic approach is proposed for approximately solving the generic optimal signaling problem via a convex relaxation technique [8]. Finally, it is mentioned that the results obtained for minimizing the average probability of error for a binary communications system can be extended to M-ary systems, as well as to other performance criteria than the average probability of error, such as the Bayes risk [2], [9].

2. SYSTEM MODEL AND MOTIVATION

Consider a scalar binary communications system, as in [3] and [5], in which the received signal is given by

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

where S_0 and S_1 denote the transmitted signal values for symbol 0 and symbol 1, respectively, and N is the noise compo-

nent that is independent of S_i. In addition, the prior probabilities of the symbols, which are denoted by π_0 and π_1, are assumed to be known.

As stated in [3], the scalar channel model in (1) presents an abstraction for a continuous-time system that processes the received signal by a linear filter and samples it once per symbol interval. Also, although the signal model in (1) is in the form of a simple additive noise channel, it also holds for flat-fading channels assuming perfect channel estimation. In that case, the signal model in (1) can be obtained after appropriate equalization [1]. 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

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

where Γ_0 and Γ_1 are the decision regions for symbol 0 and symbol 1, respectively [2].

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 calculated as

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

where $P_i(\Gamma_j)$ is the probability of selecting symbol j when symbol i is transmitted. In practical systems, there exist constraints on the average power and the peakedness of signals, which can be expressed as

$$E\{|S_i|^2\} \le A, \qquad E\{|S_i|^4\} \le \kappa A^2, \qquad (4)$$

for $i = 0, 1$, where A is the average power limit and the second constraint imposes a limit on the peakedness of the signal depending on the $\kappa \in (1, \infty)$ parameter [10]. Therefore, the problem is to calculate the optimal PDFs for signals S_0 and S_1 that minimize the average probability of error in (3) under the second and fourth moment constraints 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 [3], [7]. Therefore, the generic problem can be formulated as the calculation of the optimal probability distributions for signals S_0 and S_1 for a given decision rule at the receiver under the average power and peakedness constraints in (4).

Since the optimal signal design is performed at the transmitter, the transmitter is supposed to have the knowledge of the statistics of the noise at the receiver and the channel state information. If this information is not available, the probability of error expression obtained via the optimal stochastic signal design (cf. (6)-(7)) provides a lower bound on the probability of error. Although this information may not be available in some cases, there exist certain scenarios in which it can be valid. For example, consider the downlink of a multiple-access communications system, in which the received signal

is modeled as $Y = S^{(1)} + \sum_{k=2}^{K} S^{(k)} + \eta$, where $S^{(k)}$ is the signal of the kth user and η is a zero-mean Gaussian noise component. For the desired signal component $S^{(1)}$, $N = \sum_{k=2}^{K} S^{(k)} + \eta$ constitutes the total noise, which has Gaussian mixture distribution. When the receiver sends via feedback the variance of noise η and the signal-to-noise ratio (SNR) to the transmitter, the transmitter can fully characterize the PDF of the total noise N, as it already knows the transmitted signal levels of all the users.

In the conventional signal design, S_0 and S_1 are considered as deterministic signals, and set to $S_0 = -\sqrt{A}$ and $S_1 = \sqrt{A}$ [1], [2]. Then, the average probability of error in (3) becomes

$$P_{avg}^{conv} = \pi_0 \int_{\Gamma_1} p_N(y + \sqrt{A})\,dy + \pi_1 \int_{\Gamma_0} p_N(y - \sqrt{A})\,dy \tag{5}$$

where $p_N(\cdot)$ is the PDF of the noise in (1). As studied in Section 3.1, the conventional signal design is optimal for certain classes of noise PDFs and decision rules. However, in some cases, use of stochastic signals instead of deterministic ones can improve the system performance, as studied next.

3. 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 constraints 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

$$P_{avg}^{stoc} = \sum_{i=0}^{1} \pi_i \int_{-\infty}^{\infty} p_{S_i}(t) \int_{\Gamma_{1-i}} p_N(y - t)\, dy\, dt. \tag{6}$$

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

$$\min_{p_{S_0}, p_{S_1}} P_{avg}^{stoc} \tag{7}$$

subject to $E\{|S_i|^2\} \le A$, $E\{|S_i|^4\} \le \kappa A^2$, $i = 0, 1$.

Note that there are also implicit constraints in the optimization problem in (7), since $p_{S_0}(t)$ and $p_{S_1}(t)$ are PDFs. Namely, $p_{S_i}(t) \ge 0 \; \forall t$ and $\int_{-\infty}^{\infty} p_{S_i}(t)dt = 1$ for $i = 0, 1$.

Because the aim is to obtain optimal stochastic signals for a given receiver, the decision rule in (2) is fixed. Therefore, the structure of the objective function P_{avg}^{stoc} in (6) and the individual constraints on each signal imply that the optimization problem in (7) can be stated as two decoupled optimization problems. Specifically, the optimal signal for symbol 1 can be obtained from the solution of the following optimization problem:

$$\min_{p_{S_1}} \int_{-\infty}^{\infty} p_{S_1}(t) \int_{\Gamma_0} p_N(y - t)\, dy\, dt$$

subject to $E\{|S_1|^2\} \le A$, $E\{|S_1|^4\} \le \kappa A^2$. \tag{8}

A similar problem can also be formulated for S_0. As the signals can be designed separately, the remainder of this study focuses on the design of optimal S_1 according to (8).

The objective function in (8) can be expressed as the expectation of $G(S_1)$ over S_1, where

$$G(S_1) \triangleq \int_{\Gamma_0} p_N(y - S_1)\, dy \ . \tag{9}$$

Then, the optimization problem in (8) can be stated as

$$\min_{p_{S_1}} \mathrm{E}\{G(S_1)\}$$

$$\text{subject to } \mathrm{E}\{|S_1|^2\} \le A \ , \quad \mathrm{E}\{|S_1|^4\} \le \kappa A^2 \ . \tag{10}$$

In the following, the signal subscripts are dropped for notational simplicity.

3.1. On the Optimality of Conventional Signals

In some cases, the conventional signaling is an optimal approach; that is, setting $S = \sqrt{A}$ [or, $p_S(x) = \delta(x - \sqrt{A})$] can solve the optimization problem in (10). For example, if $G(x)$ in (9) achieves its minimum at $x = \sqrt{A}$; that is, $\arg\min_x G(x) = \sqrt{A}$, then $p_S(x) = \delta(x - \sqrt{A})$ is the optimal solution as it provides the minimum value for $\mathrm{E}\{G(S_1)\}$ under the constraints. However, the definition of $G(x)$ in (9) reveals that it is the probability of deciding symbol 0 instead of symbol 1 when signal S_1 takes a constant value of x; hence, it is commonly a decreasing function of x, as larger signal values can lead to smaller error probabilities. Therefore, a more generic condition is obtained in the following proposition for the optimality of the conventional algorithm.

Proposition 1: *If $G(x)$ is a strictly convex and monotone decreasing function, then $p_S(x) = \delta(x - \sqrt{A})$ is a solution of the optimization problem in* (10).

Proof: The result can be obtained by showing, via Jensen's inequality, that no signal PDF can satisfy $\mathrm{E}\{G(S)\} < G(\sqrt{A})$ and the constraints in (10) at the same time when $G(x)$ is a strictly convex and monotone decreasing function. $\square$

As an example, consider zero-mean Gaussian noise N in (1) with $p_N(x) = \exp\{-x^2/(2\sigma^2)\}/\sqrt{2\pi}\,\sigma$, and a decision rule of the form $\Gamma_0 = (-\infty, 0]$ and $\Gamma_1 = [0, \infty)$; that is, the sign detector. Then, $G(x)$ in (9) can be calculated as $G(x) = Q(x/\sigma)$, where $Q(x) = (1/\sqrt{2\pi}) \int_x^\infty e^{-t^2/2} dt$ defines the Q-function. Since $G(x)$ is a monotone decreasing and strictly convex function for $x > 0$,[1] the optimal signal can be specified by $p_S(x) = \delta(x - \sqrt{A})$ based on Proposition 1. Similarly, the optimal signal for symbol 0 can be calculated as $p_S(x) = \delta(x + \sqrt{A})$. Hence, the conventional signaling is optimal in this scenario.

3.2. Sufficient Conditions for Improvability

In this section, we study the conditions under which the performance of the conventional signaling approach can be improved via stochastic signaling. A simple observation from

[1] It is sufficient to consider the positive signal values only, because $G(x)$ is monotone decreasing and the constraints x^2 and x^4 are even functions.

(10) reveals that if $G'(\sqrt{A}) > 0$, where $G'(x)$ is the first derivative of $G(x)$, a signal PDF in the form of $p_{S2}(x) = \delta(x - \sqrt{A} + \epsilon)$ provides a smaller average probability of error than the conventional solution for infinitesimally small $\epsilon > 0$. Hence, the conventional signaling is suboptimal in that case. Although this condition is sufficient for the improvability of the conventional solution, it is rarely met in practice since $G(x)$ is commonly a decreasing function of x as discussed before. Therefore, a sufficient condition is derived for more generic and practical $G(x)$ functions in the following.

Proposition 2: *Assume that $G(x)$ is twice continuously differentiable. If $G''(\sqrt{A}) < G'(\sqrt{A})/\sqrt{A}$, then $p_S(x) = \delta(x - \sqrt{A})$ is not an optimal solution of* (10).

Proof: In order to prove the suboptimality of the conventional solution $p_S(x) = \delta(x - \sqrt{A})$, it is shown that, under the conditions in the proposition, there exist $\lambda \in (0, 1)$, $\epsilon > 0$ and $\Delta > 0$ such that $p_{S2}(x) = \lambda\,\delta(x - \sqrt{A} + \epsilon) + (1 - \lambda)\,\delta(x - \sqrt{A} - \Delta)$ yields a lower error probability than $p_S(x)$ and satisfies the constraints in (10). Specifically, the existence of $\lambda \in (0, 1)$, $\epsilon > 0$ and $\Delta > 0$ that satisfy

$$\lambda G(\sqrt{A} - \epsilon) + (1 - \lambda)\,G(\sqrt{A} + \Delta) < G(\sqrt{A}) \tag{11}$$

$$\lambda(\sqrt{A} - \epsilon)^2 + (1 - \lambda)(\sqrt{A} + \Delta)^2 = A \tag{12}$$

$$\lambda(\sqrt{A} - \epsilon)^4 + (1 - \lambda)(\sqrt{A} + \Delta)^4 \le \kappa A^2 \tag{13}$$

is *sufficient* to prove the suboptimality of the conventional signaling. From (12), the following relation is obtained.

$$\lambda \epsilon^2 + (1 - \lambda)\Delta^2 = -2\sqrt{A}\,[(1 - \lambda)\Delta - \lambda \epsilon] \ . \tag{14}$$

For infinitesimally small ϵ and Δ, the first three terms of the Taylor series expansions for $G(\sqrt{A} - \epsilon)$ and $G(\sqrt{A} + \Delta)$ can be used to approximate (11) as

$$G'(\sqrt{A})\,[(1 - \lambda)\Delta - \lambda \epsilon] + \frac{G''(\sqrt{A})}{2}\,[\lambda \epsilon^2 + (1 - \lambda)\Delta^2] < 0. \tag{15}$$

Based on the relation in (14), (15) can be expressed as

$$[(1 - \lambda)\Delta - \lambda \epsilon]\left(G'(\sqrt{A}) - \sqrt{A}\,G''(\sqrt{A})\right) < 0 \ . \tag{16}$$

Since $(1 - \lambda)\Delta - \lambda \epsilon$ is always negative, which can be observed from (14), the $G'(\sqrt{A}) - \sqrt{A}\,G''(\sqrt{A})$ term in (16) must be positive to satisfy the condition. In other words, when $G''(\sqrt{A}) < G'(\sqrt{A})/\sqrt{A}$, $p_{S2}(x)$ can have a smaller error value than the conventional solution for infinitesimally small ϵ and Δ values that satisfy (14).

Finally, the condition in (13) can be verified in a similar fashion, which is not shown here due to space limitations. $\square$

The reasoning behind Proposition 2 is explained as follows. Since the optimization problem in (10) aims to minimize $\mathrm{E}\{G(S)\}$ while keeping $\mathrm{E}\{S^2\}$ and $\mathrm{E}\{S^4\}$ below thresholds A and κA^2, respectively, a better solution than $p_S(x) = \delta(x - \sqrt{A})$ can be obtained with multiple mass points if $G(x)$ is decreasing at an increasing rate (i.e., with a negative second derivative) such that an increase from $x = \sqrt{A}$ causes a fast decrease in $G(x)$ but a relatively slow increase in x^2 and x^4, and a decrease from $x = \sqrt{A}$ causes a fast decrease in x^2

and x^4 but a relatively slow increase in $G(x)$. Then, it becomes possible to use a PDF with multiple mass points and to achieve a smaller $E\{G(S)\}$ while satisfying $E\{S^2\} \leq A$ and $E\{S^4\} \leq \kappa A^2$.

3.3. Calculation of Optimal Signals

In order to obtain the PDF of an optimal signal, the constrained optimization problem in (10) should be solved. In this section, a convex optimization approach is studied in order to provide approximate solutions for that optimization problem. We consider a scenario in which the PDF of the signal is modeled as

$$p_S(x) = \sum_{j=1}^{K} \tilde{\lambda}_j \, \delta(x - \tilde{x}_j), \qquad (17)$$

where $\tilde{x}_j$'s are the *known* mass points of the PDFs, and $\tilde{\lambda}_j$'s are the weights (probabilities) to be estimated. In other words, it is assumed that there are a finite number of possible signal values, and the aim is to determine the probabilities of those values. Of course, the PDF model in (17) provides an approximation to the optimal solution, which can also take values different from $\tilde{x}_j$'s. However, as the number of possible values increases, the approximate solution can get closer to the exact solution. In addition, in practical systems, the signals are digital; hence, they can only take finitely many possible values as in (17). Therefore, the model would be exact for such digital systems.

Based on the model in (17), the optimal signal design problem in (10) can be expressed as the following convex optimization problem:[2]

$$\min_{\tilde{\lambda}} \mathbf{g}^T \tilde{\lambda} \qquad (18)$$

$$\text{subject to } \mathbf{B}\tilde{\lambda} \preceq \mathbf{C}, \quad \mathbf{1}^T\tilde{\lambda} = 1, \quad \tilde{\lambda} \succeq \mathbf{0},$$

where $\mathbf{g} \triangleq [G(\tilde{x}_1) \cdots G(\tilde{x}_K)]^T$, with $G(x)$ as in (9),

$$\mathbf{B} \triangleq \begin{bmatrix} \tilde{x}_1^2 & \cdots & \tilde{x}_K^2 \\ \tilde{x}_1^4 & \cdots & \tilde{x}_K^4 \end{bmatrix}, \quad \mathbf{C} \triangleq \begin{bmatrix} A \\ \kappa A^2 \end{bmatrix}, \qquad (19)$$

and $\mathbf{1}$ and $\mathbf{0}$ denote vectors of ones and zeros, respectively.

It is observed from (18) that the optimal weight assignments can be obtained from the solution of a convex optimization problem; specifically, a linearly constrained linear programming problem. Therefore, the problem can be efficiently solved by interior-point methods, which are polynomial time in the worst case, and are very fast in practice [8].

4. SIMULATION RESULTS

In this section, numerical examples are presented for a binary communications system with equal priors; that is, $\pi_0 = \pi_1 = 0.5$. The decision rule at the receiver is specified by $\Gamma_0 = (-\infty, 0]$ and $\Gamma_1 = [0, \infty)$ (i.e., a sign detector).

A communications system in the presence of interference is considered, and the noise in (1) is modeled as Gaussian
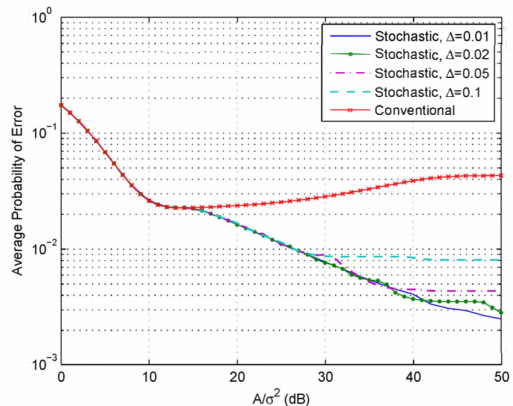
Fig. 1. Average probability of error versus A/σ^2.

mixture noise, which is specified by $p_N(y) = \sum_{l=1}^{L} v_l \, \psi_l(y - y_l)$, where $\psi_l(y) = \exp\{-y^2/(2\sigma_l^2)\}/(\sqrt{2\pi}\,\sigma_l)$. It should be noted that such Gaussian mixture noise can be encountered in practical communications systems in the presence of co-channel interference [4]. Then, $G(x)$ in (9) is obtained as $G(x) = \sum_{l=1}^{L} v_l \, Q((x + y_l)/\sigma_l)$. In the following, the variance parameter for each mass point of the Gaussian mixture is set to σ^2 (i.e., $\sigma_l^2 = \sigma^2 \,\forall l$), the average power constraint A is set to 1, and $\kappa = 1.5$ is used. Note that the average power of the noise can be calculated as $E\{N^2\} = \sigma^2 + \sum_{l=1}^{L} v_l \, y_l^2$.

In Fig. 1, the average probabilities of error are plotted against A/σ^2 for the conventional and stochastic signaling approaches for a symmetric Gaussian mixture noise that has its mass points at $\pm[0.105\ 0.275\ 1.013]$ with corresponding weights $[0.129\ 0.328\ 0.043]$. In the implementation of the convex solution in Section 3.3, the mass points $\tilde{x}_j$ in (17) are selected uniformly over the interval $[0, 2]$ with a step size of Δ, and the results for $\Delta = 0.01, 0.02, 0.05, 0.1$ are illustrated.[3] It is observed from Fig. 1 that the conventional signaling, which uses a constant signal value of 1, has a large error floor compared to the stochastic signaling at high A/σ^2 values. In addition, the average probability of error of the conventional signaling increases as A/σ^2 increases after a certain value. This seemingly counterintuitive result is observed since the average probability of error is related to the area under the two shifted noise PDFs as in (5). Since the noise has a multi-modal PDF, that area is a non-monotonic function of A/σ^2 and can increase in some cases as A/σ^2 increases. Moreover, Fig. 1 shows that the stochastic signaling provides significant performance improvements over the conventional signaling, especially for densely spaced possible signal values. In addition, it is observed that decreasing the value of Δ below a certain value does not result in significant reductions in the average probability of error. For example, $\Delta = 0.01$ does not provide much performance improvement compared to $\Delta = 0.02$. Hence, a reasonably small Δ can be chosen in practice in order to obtain close-to-optimal performance.

Another observation from Fig. 1 is that improvements over the conventional algorithm disappear as σ^2 increases (that is, for small A/σ^2 values), which can be explained from Propositions 1 and 2, based on the plots of $G(x)$ at various A/σ^2 values. As an example, Fig. 2 illustrates the plots of $G(x)$ at

---

[2]For K-dimensional vectors $\mathbf{x}$ and $\mathbf{y}$, $\mathbf{x} \preceq \mathbf{y}$ means that the ith element of $\mathbf{x}$ is smaller than or equal to the ith element of $\mathbf{y}$ for $i = 1, \ldots, K$.
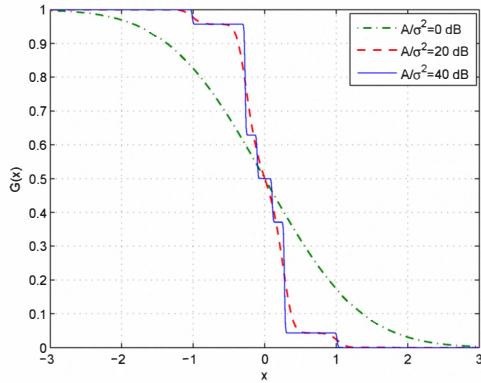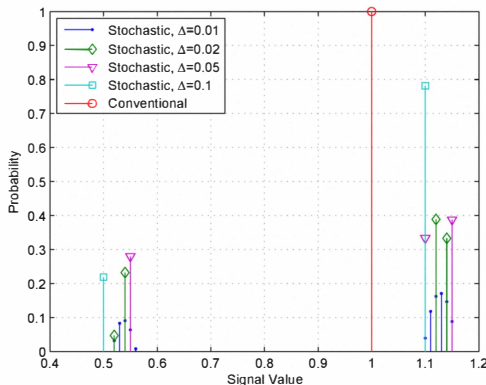
[3]The signal values with zero probabilities are not marked in the figures to clarify the illustrations.

Fig. 2. $G(x)$ in (9) for various values of A/σ^2.

Fig. 3. Signal PMFs for various schemes at $A/\sigma^2 = 20\,\text{dB}$.

A/σ^2 of 0, 20 and 40 dB. The function is decreasing and convex for 0 dB for the positive signal values, which are practically the domain of optimization as $G(x)$ is a decreasing function and the constraint functions x^2 and x^4 are even functions. Hence, Proposition 1 implies that the conventional algorithm that uses a constant signal value of 1 is optimal in this case, as observed in Fig. 1. On the other hand, at 20 dB and 40 dB, the calculations show that the condition in Proposition 2 is satisfied. Namely, $G''(1) = -0.221$ and $G'(1) = -0.170$ at 20 dB, and $G''(1) = -95.8$ and $G'(1) = -0.737$ at 40 dB. Therefore, the conventional algorithm cannot be optimal in that case, and improvements are observed in Fig. 1 at $A/\sigma^2 = 20$ dB and $A/\sigma^2 = 40$ dB.

For the scenario in Fig. 1, the probability mass functions (PMFs) of the conventional and stochastic signals are shown in Fig. 3 for $A/\sigma^2 = 20$ dB. It is observed that the stochastic signaling performs randomization of signal amplitudes mainly around two values ($S \approx 0.54$ and $S \approx 1.13$). Depending on the resolution; that is, the value of Δ, various numbers of mass points are obtained. As Δ increases, the convex optimization approach does not provide sufficient resolution for the signal values, and the resulting error probability becomes higher, especially for small σ's, as observed from Fig. 1.

5. CONCLUSIONS AND EXTENSIONS

The stochastic signaling problem has been studied for binary communications systems under second and fourth moment constraints. It has been shown that, the conventional signaling approach, which employs deterministic signals at the average power limit, is optimal under certain monotonicity and convexity conditions. On the other hand, in certain cases, a smaller average probability of error can achieved by using a signal that is obtained by a randomization of multiple signal values. In addition, a convex relaxation approach has been proposed to perform close-to-optimal signal design.

The results in this study can be extended to a generic binary hypothesis-testing problem in the Bayesian framework [2], [9]. In that case, the average probability of error expression in (3) is generalized to the *Bayes risk*, which is defined as $\pi_0[C_{00}P_0(\Gamma_0) + C_{10}P_0(\Gamma_1)] + \pi_1[C_{01}P_1(\Gamma_0) + C_{11}P_1(\Gamma_1)]$, where $C_{ij} \geq 0$ represents the cost of deciding the ith hypothesis when the jth one is true. Then, all the results are still valid when function G in (9) is replaced by $G(x) = C_{01}\int_{\Gamma_0} p_N(y - x)dy + C_{11}\int_{\Gamma_1} p_N(y - x)dy$. Moreover, it can be shown that the results in this study can also be extended to M-ary communications systems for $M > 2$.

6. REFERENCES

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

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

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

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

[5] S. Shamai and S. Verdu, "Worst-case power-constrained noise for binary-input channels," *IEEE Trans. Inform. Theory*, vol. 38, pp. 1494–1511, Sep. 1992.

[6] M. A. Klimesh and W. E. Stark, "Worst-case power-constrained noise for binary-input channels with varying amplitude signals," in *Proc. IEEE Int. Symp. on Inform. Theory (ISIT)*, July 1994, p. 381.

[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] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.

[9] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Upper Saddle River, NJ: Prentice Hall, Inc., 1998.

[10] M. C. Gursoy, H. V. Poor, and S. Verdu, "Efficient signaling for low-power Rician fading channels," in *Proc. Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Oct. 2002.