Optimal Signaling and Detector Design for Power-Constrained Binary Communications Systems over Non-Gaussian Channels

Çağrı Göken, *Student Member, IEEE*, Sinan Gezici, *Member, IEEE*, and Orhan Arikan, *Member, IEEE*

Abstract—In this letter, joint optimization of signal structures and detectors is studied for binary communications systems under average power constraints in the presence of additive non-Gaussian noise. First, it is observed that the optimal signal for each symbol can be characterized by a discrete random variable with at most two mass points. Then, optimization over all possible two mass point signals and corresponding maximum *a posteriori* probability (MAP) decision rules are considered. It is shown that the optimization problem can be simplified into an optimization over a number of signal parameters instead of functions, which can be solved via global optimization techniques, such as particle swarm optimization. Finally, the improvements that can be obtained via the joint design of the signaling and the detector are illustrated via an example.

Index Terms—Stochastic signaling, MAP decision rule.

I. INTRODUCTION AND MOTIVATION

IN binary communications systems over additive white Gaussian noise channels and under average power constraints in the form of $E\{|\mathbf{s}_i|^2\} \le A$ for $i = 0, 1$, the average probability of error is minimized when deterministic antipodal signals ($\mathbf{s}_0 = -\mathbf{s}_1$) are used at the power limit ($|\mathbf{s}_0|^2 = |\mathbf{s}_1|^2 = A$) and a maximum *a posteriori* probability (MAP) decision rule is employed at the receiver [1]. In addition, when the Gaussian noise is colored, the deterministic antipodal signals along the eigenvector of the covariance matrix of the Gaussian noise corresponding to the minimum eigenvalue minimizes the average probability of error [1]. Although the optimal detector and signaling techniques are well-known when the noise is Gaussian, the noise can have significantly different probability distribution from the Gaussian distribution in some cases due to effects such as interference and jamming [2]. In the presence of non-Gaussian noise, stochastic signaling, which models signals $\mathbf{s}_0$ and $\mathbf{s}_1$ as random variables, can result in improved probability of error performance compared to deterministic signaling. In [3], optimal stochastic signaling is studied under second and fourth moment constraints for a *fixed* decision rule (detector) at the receiver, and sufficient conditions are presented to determine whether stochastic signaling can provide performance improvements compared to deterministic signaling. In [4], 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 improvement.

Although optimal stochastic signaling is studied for a fixed detector in [3] and the effects of randomization between two

signaling approaches are considered in [4], no studies have focused on the joint optimization of stochastic signaling and the decision rule (detector). In this letter, this joint optimization problem is formulated, which involves optimization over a function space. Then, theoretical results are provided to show that the optimal solution can be obtained by searching over a number of variables instead of functions, which greatly simplifies the original formulation. In addition, a global optimization approach, namely particle swarm optimization (PSO) [5], is employed to obtain the optimal signals and the decision rule.

The main motivation behind this study is to provide theoretical performance limits on the error probability of communications systems under power constraints. It is assumed that there is a feedback from the receiver to the transmitter so that the joint optimization of the signaling structure and the decision rule can be performed. This scenario is reasonable for cognitive radio systems, and provides theoretical limits on the error performance of other communications systems.

II. OPTIMAL SIGNALING AND DETECTOR DESIGN

Consider a binary communications system, in which the receiver obtains K-dimensional observations over an additive noise channel:

$$\mathbf{y} = \mathbf{s}_i + \mathbf{n} , \quad i \in \{0, 1\} , \tag{1}$$

where $\mathbf{y}$ is the noisy observation, $\mathbf{s}_0$ and $\mathbf{s}_1$ represent the transmitted signal values for symbol 0 and symbol 1, respectively, and $\mathbf{n}$ is the noise component that is independent of $\mathbf{s}_i$. In addition, the prior probabilities of the symbols, represented by π_0 and π_1, are assumed to be known.

The receiver uses the observation in (1) in order to determine the information symbol. A generic decision rule (detector) is considered for that purpose, which estimates the transmitted symbol based on a given observation $\mathbf{y}$ as follows:

$$\phi(\mathbf{y}) = \begin{cases} 0 , & \mathbf{y} \in \Gamma_{\phi_0} \\ 1 , & \mathbf{y} \in \Gamma_{\phi_1} \end{cases} , \tag{2}$$

where Γ_{ϕ_0} and Γ_{ϕ_1} are the decision regions for symbol 0 and symbol 1, respectively [1].

The average probability of error for a decision rule ϕ can be expressed as $P_e = \pi_0 P_{e,0} + \pi_1 P_{e,1}$, where

$$P_{e,i} = \int_{\Gamma_{\phi_{1-i}}} p_i(\mathbf{y}) \, d\mathbf{y} , \tag{3}$$

for $i = 0, 1$, represents the probability of error, with $p_i(\mathbf{y})$ denoting the conditional probability density function (PDF) of the observation, when the ith symbol is transmitted.

Unlike the conventional case, a stochastic signaling framework is adopted in this study [3], and $\mathbf{s}_0$ and $\mathbf{s}_1$ in (1) are modeled as random variables. Since the signals and the noise are independent, the conditional PDFs of the observation can

be calculated as $p_i(\mathbf{y}) = \int_{\mathbb{R}^K} p_{\mathbf{s}_i}(\mathbf{x}) p_{\mathbf{n}}(\mathbf{y} - \mathbf{x}) \, d\mathbf{x}$ for $i = 0, 1$. Then, after some manipulation, (3) can be expressed as

$$\mathrm{P}_{e,i} = \mathrm{E} \left\{ \int_{\Gamma_{\phi_{1-i}}} p_{\mathbf{n}}(\mathbf{y} - \mathbf{s}_i) \, d\mathbf{y} \right\} \triangleq \mathrm{E} \left\{ f(\phi \, ; \mathbf{s}_i) \right\} , \quad (4)$$

where the expectation is taken over the PDF of $\mathbf{s}_i$.

In a practical system, there is a constraint on the average power of the signals, which can be expressed as [1]

$$\mathrm{E} \left\{ |\mathbf{s}_i|^2 \right\} \le A , \quad \text{for } i = 0, 1 , \quad (5)$$

where A is the average power limit. Then, the optimal signaling and detector design problem can be stated as

$$\min_{p_{\mathbf{s}_0}, p_{\mathbf{s}_1}, \phi} \pi_0 \mathrm{P}_{e,0} + \pi_1 \mathrm{P}_{e,1}$$
$$\text{subject to } \mathrm{E} \left\{ |\mathbf{s}_i|^2 \right\} \le A , \quad i = 0, 1 , \quad (6)$$

where $\mathrm{P}_{e,i}$ is as in (4).

The problem in (6) is difficult to solve in general since the optimization needs to be performed over a space of PDFs and decision rules. In the following, a simpler optimization problem over a set of variables (instead of functions) is formulated to obtain optimal signal PDFs and the decision rule. To that aim, the following result is obtained first.

Lemma 1: *Assume $f(\phi \, ; \mathbf{s}_i)$ in (4) is a continuous function of $\mathbf{s}_i$, and each component of $\mathbf{s}_i$ resides in $[-\gamma, \gamma]$ for some finite $\gamma > 0$. Then, for a given (fixed) decision rule ϕ, the solution of the optimization problem in (6) is in the form of*

$$p_{\mathbf{s}_i}(\mathbf{y}) = \lambda_i \delta(\mathbf{y} - \mathbf{s}_{i1}) + (1 - \lambda_i) \delta(\mathbf{y} - \mathbf{s}_{i2}) , \quad (7)$$

for $i = 0, 1$, where $\lambda_i \in [0, 1]$.

Proof: When the decision rule ϕ is given, $f(\phi \, ; \mathbf{s}_i) = \int_{\Gamma_{\phi_{1-i}}} p_{\mathbf{n}}(\mathbf{y} - \mathbf{s}_i) \, d\mathbf{y}$ in (4) can be considered as a function of $\mathbf{s}_i$ only. In other words, $\mathrm{P}_{e,i}$ in (4) can be expressed as $\mathrm{P}_{e,i} = \mathrm{E}\{f(\mathbf{s}_i)\}$ for $i = 0, 1$. Since the objective function in (6) is the sum of $\pi_0 \mathrm{P}_{e,0}$ and $\pi_1 \mathrm{P}_{e,1}$, and the average power constraints are individually imposed on the signals, the optimization problem in (6) can be decoupled into two separate optimization problems as follows:

$$\min_{p_{\mathbf{s}_i}} \mathrm{E}\{f(\mathbf{s}_i)\} , \quad \text{subject to } \mathrm{E} \left\{ |\mathbf{s}_i|^2 \right\} \le A , \quad (8)$$

for $i = 0, 1$. Optimization problems in the form of (8) have been investigated in various studies in the literature [3], [4]. Under the conditions in the lemma, the optimal solution of (8) can be represented by a randomization of at most two signal levels as a result of Carathéodory's theorem [6]. Hence, the optimal signal PDFs can be expressed as in (7). □

Lemma 1 states that, under certain conditions, the optimal stochastic signaling involves randomization among at most four different signal levels (two for symbol "0" and two for symbol "1"). Therefore, the problem in (6) can be solved over the signal PDFs that are in the form of (7). Hence, the search space for the optimization problem is reduced significantly. To achieve further simplification, the following result is obtained.

Proposition 1: *Under the conditions in Lemma 1, the*

optimization problem in (6) can be expressed as follows:

$$\min_{\{\lambda_i, \mathbf{s}_{i1}, \mathbf{s}_{i2}\}_{i=0}^1} \int_{\mathbb{R}^K} \min\{\pi_0 g_0(\mathbf{y}) , \pi_1 g_1(\mathbf{y})\} \, d\mathbf{y}$$
$$\text{subject to } \lambda_i |\mathbf{s}_{i1}|^2 + (1 - \lambda_i)|\mathbf{s}_{i2}|^2 \le A$$
$$\lambda_i \in [0, 1] , \quad i = 0, 1 \quad (9)$$

where $g_i(\mathbf{y}) = \lambda_i p_{\mathbf{n}}(\mathbf{y} - \mathbf{s}_{i1}) + (1 - \lambda_i) p_{\mathbf{n}}(\mathbf{y} - \mathbf{s}_{i2})$.

Proof: For a given signal PDF pair $p_{\mathbf{s}_0}$ and $p_{\mathbf{s}_1}$, the conditional probability of observation $\mathbf{y}$ in (1) can be expressed as $p_i(\mathbf{y}) = \int_{\mathbb{R}^K} p_{\mathbf{s}_i}(\mathbf{x}) p_{\mathbf{n}}(\mathbf{y} - \mathbf{x}) d\mathbf{x}$ for $i = 0, 1$. When deciding between two symbols based on observation $\mathbf{y}$, the MAP decision rule, which selects symbol 1 if $\pi_1 p_1(\mathbf{y}) \ge \pi_0 p_0(\mathbf{y})$ and selects symbol 0 otherwise, minimizes the average probability of error [1]. Therefore, when signal PDFs $p_{\mathbf{s}_0}$ and $p_{\mathbf{s}_1}$ are specified, it is not necessary to search over all the decision rules; only the MAP decision rule should be determined and its corresponding average probability of error should be considered.

From (3), the average probability of error for any decision rule ϕ can be expressed as

$$\mathrm{P}_e = \int_{\Gamma_{\phi_1}} \pi_0 p_0(\mathbf{y}) \, d\mathbf{y} + \int_{\Gamma_{\phi_0}} \pi_1 p_1(\mathbf{y}) \, d\mathbf{y} . \quad (10)$$

Since the MAP decision rule decides symbol 1 if $\pi_1 p_1(\mathbf{y}) \ge \pi_0 p_0(\mathbf{y})$ and decides symbol 0 otherwise, the average probability of error expression in (10) can be expressed for a MAP decision rule as [7]

$$\mathrm{P}_e = \int_{\mathbb{R}^K} \min\{\pi_0 p_0(\mathbf{y}) , \pi_1 p_1(\mathbf{y})\} \, d\mathbf{y} . \quad (11)$$

Since Lemma 1 states that the optimal signal PDFs are in the form of (7), the conditional PDFs $p_i(\mathbf{y}) = \int_{\mathbb{R}^K} p_{\mathbf{s}_i}(\mathbf{x}) p_{\mathbf{n}}(\mathbf{y} - \mathbf{x}) d\mathbf{x}$ can be obtained as $p_i(\mathbf{y}) = \lambda_i p_{\mathbf{n}}(\mathbf{y} - \mathbf{s}_{i1}) + (1 - \lambda_i) p_{\mathbf{n}}(\mathbf{y} - \mathbf{s}_{i2})$, and the average power constraints in (6) become $\lambda_i |\mathbf{s}_{i1}|^2 + (1 - \lambda_i)|\mathbf{s}_{i2}|^2 \le A$, for $i = 0, 1$. Therefore, (11) implies that the optimization problem in (6) can be implemented as the constrained minimization problem in the proposition. □

Comparison of the optimization problems in (6) and (9) reveals that the latter is much simpler than the former since it is over a set of variables instead of a set of functions. However, it is still a non-convex optimization problem in general; hence, global optimization techniques, such as PSO, differential evolution and genetic algorithms, should be employed to obtain the optimal PDF [5]. In this letter, the PSO approach is used in the next section to obtain the solution of (9).

After obtaining the solution of the optimization problem in (9), the optimal signals are specified as $p_{\mathbf{s}_i}^{\mathrm{opt}}(\mathbf{y}) = \lambda_i^{\mathrm{opt}} \delta(\mathbf{y} - \mathbf{s}_{i1}^{\mathrm{opt}}) + (1 - \lambda_i^{\mathrm{opt}}) \delta(\mathbf{y} - \mathbf{s}_{i2}^{\mathrm{opt}})$ for $i = 0, 1$, and the optimal detector becomes the MAP decision rule that decides symbol 1 if $\pi_1 p_{\mathbf{s}_1}^{\mathrm{opt}}(\mathbf{y}) \ge \pi_0 p_{\mathbf{s}_0}^{\mathrm{opt}}(\mathbf{y})$ and decides symbol 0 otherwise.

Finally, it should be noted for symmetric signaling, that is, when $\mathbf{s}_{01} = -\mathbf{s}_{11}$, $\mathbf{s}_{02} = -\mathbf{s}_{12}$ and $\lambda_0 = \lambda_1$, the optimization in (9) can be performed over $\mathbf{s}_{11}$, $\mathbf{s}_{12}$ and λ_1 only.

III. NUMERICAL RESULTS AND CONCLUSIONS

A numerical example is presented to illustrate the improvements that can be obtained via the joint design of the signaling structure and the decision rule for scalar observations. The
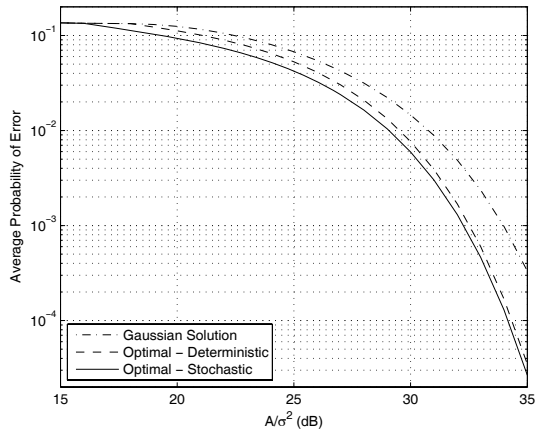
Fig. 1. Average probability of error versus A/σ^2 for the three algorithms.

TABLE I
OPTIMAL STOCHASTIC AND DETERMINISTIC SIGNALS FOR SYMBOL 1.

A/σ^2 (dB)	Stochastic			Deterministic
	λ_1	$\mathbf{s}_{11}$	$\mathbf{s}_{12}$	$\mathbf{s}_1$
15	N/A	1	1	1
20	0.1836	1.648	0.7846	0.7927
25	0.2104	1.614	0.7576	0.7587
30	0.2260	1.586	0.7475	0.7476
35	0.2347	1.568	0.7441	0.8759

noise in (1) is modeled by a Gaussian mixture as in [2] with its PDF being given by $p_n(y) = \frac{1}{\sqrt{2\pi}\,\sigma L}\sum_{i=1}^{L} e^{-\frac{(y-\mu_i)^2}{2\sigma^2}}$, where $L = 6$ and $\boldsymbol{\mu} = [0.27\ 0.81\ 1.08\ -1.08\ -0.81\ -0.27]$ are used. Note that the average power of the noise can be calculated as $\mathrm{E}\{n^2\} = \sigma^2 + 0.6318$. In addition, the average power limit in (5) is set to $A = 1$ and equally likely symbols are considered ($\pi_0 = \pi_1 = 0.5$).

In the following, three different approaches are compared.

Gaussian Solution: In this case, the transmitter is assumed to have no information about the noise PDF and selects the signals as $\mathbf{s}_0 = -\sqrt{A}$ and $\mathbf{s}_1 = \sqrt{A}$, which are known to be optimal in the presence of zero-mean Gaussian noise [1]. On the other hand, the MAP decision rule is used at the receiver.

Optimal – Stochastic: This approach refers to the solution of the most generic optimization problem in (6), which can also be obtained from (9) as studied in the previous section.

Optimal – Deterministic: This is a simplified version of the optimal solution in (9). It assumes that the signals are deterministic; i.e., they are not randomization of two different signal levels. Hence, the optimization problem in (9) becomes

$$\min_{\mathbf{s}_0,\mathbf{s}_1} \int_{\mathbb{R}^K} \min\{\pi_0 p_{\mathbf{n}}(\mathbf{y}-\mathbf{s}_0)\,,\,\pi_1 p_{\mathbf{n}}(\mathbf{y}-\mathbf{s}_1)\}\,d\mathbf{y}$$
$$\text{subject to}\quad |\mathbf{s}_0|^2 \leq A\,,\quad |\mathbf{s}_1|^2 \leq A\,. \qquad (12)$$

In other words, this approach provides the optimal solution when the signals are deterministic.

In Fig. 1, the average probabilities of error are plotted versus A/σ^2 for the three algorithms above by considering symmetric signaling. In obtaining the optimal stochastic solution from (9), the PSO algorithm is employed with 50 particles and 1000 iterations. Please refer to [5] for the details of the PSO algorithm[1]. On the other hand, the optimal deterministic solution in (12) can be obtained via a one-dimensional search due to symmetric signaling. From Fig. 1, it is observed that the Gaussian solution performs significantly worse than the optimal approaches for small σ values. In addition, the optimal approach based on stochastic signaling has the best performance.

In order to explain the results in Fig. 1, Table I presents the solutions of the optimization problems in (6) and (12) for the optimal stochastic and the optimal deterministic approaches,

respectively. Note that the results for symbol 1 are listed in Table I, and the results for symbol 0 are the negatives of the signal values in the table since symmetric signaling is considered. For small A/σ^2 values, such as 15 dB, the optimal solutions are the same as the Gaussian solution, that is, $\mathbf{s}_{11} = \mathbf{s}_{12} = \mathbf{s}_1 = \sqrt{A} = 1$. However, for large A/σ^2's, the Gaussian solution becomes quite suboptimal and choosing the largest possible deterministic signal value, 1, results in higher average probabilities of error, as can be observed from Fig. 1. For example, at $A/\sigma^2 = 30$ dB, the optimal deterministic solution sets $\mathbf{s}_1 = -\mathbf{s}_0 = 0.7476$ and achieves an error rate of 7.66×10^{-3}, whereas the Gaussian one uses $\mathbf{s}_1 = -\mathbf{s}_0 = 1$, which yields an error rate of 0.0146. This seemingly counterintuitive result is obtained since the average probability of error is related to the area under the overlaps of the two shifted noise PDFs as in (12). Although optimal deterministic signaling uses less power than permitted, it results in a lower error probability than Gaussian signaling by avoiding the overlaps between the components of the Gaussian mixture noise more effectively. On the other hand, optimal stochastic signaling further reduces the average probability of error by using all the available power and assigning some of the power to a large signal component that results in less overlapping between the shifted noise PDFs. For example, at $A/\sigma^2 = 30$ dB, the optimal stochastic signal is a randomization of $\mathbf{s}_{11} = -\mathbf{s}_{01} = 1.586$ and $\mathbf{s}_{12} = -\mathbf{s}_{02} = 0.7475$ with $\lambda_0 = \lambda_1 = 0.226$ (cf. (7)), which achieves an error rate of 5.95×10^{-3}.

The results in this letter can be extended to M-ary communications systems as well by noting that the average probability of error expression in (11) becomes $\mathrm{P_e} = 1 - \int \max\{\pi_0 p_0(\mathbf{y}),\ldots,\pi_{M-1} p_{M-1}(\mathbf{y})\}d\mathbf{y}$ for M-ary systems. Then, an optimization problem similar to that in Proposition 1 can be obtained, where the optimization is performed over $\{\lambda_i, \mathbf{s}_{i1}, \mathbf{s}_{i2}\}_{i=0}^{M-1}$.

REFERENCES

[1] H. V. Poor, *An Introduction to Signal Detection and Estimation,* 2nd ed., New York: Springer-Verlag, 1994.
[2] V. Bhatia and B. Mulgrew, "Non-parametric likelihood based channel estimator for Gaussian mixture noise," *Signal Processing,* vol. 87, pp. 2569–2586, Nov. 2007.
[3] C. Goken, S. Gezici, and O. Arikan, "Stochastic signaling under second and fourth moment constraints," submitted, Jan. 2010. [Available: www.ee.bilkent.edu.tr/~gezici/goken.pdf].
[4] 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.
[5] 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.
[6] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis,* Berlin: Springer-Verlag, 2004.
[7] M. Azizoglu, "Convexity properties in binary detection problems," *IEEE Trans. Inform. Theory,* vol. 42, no. 4, pp. 1316–1321, July 1996.

[1]The other parameters are set to $c_1 = c_2 = 2.05$ and $\chi = 0.72984$, and the inertia weight ω is changed from 1.2 to 0.1 linearly with the iteration number [5].