

# Optimal Channel Switching in the Presence of Stochastic Signaling

Berkan Dulek and Pramod K. Varshney

Department of Electrical Engineering and Computer Science  
Syracuse University, Syracuse, NY 13244, USA  
Email: {bdulek, varshney}@syr.edu

Mehmet Emin Tutay and Sinan Gezici

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

**Abstract**—Optimal channel switching and detector design is studied for  $M$ -ary communication systems in the presence of stochastic signaling, which facilitates randomization of signal values transmitted for each information symbol. Considering the presence of multiple additive noise channels (which can have non-Gaussian distributions in general) between a transmitter and a receiver, the joint optimization of the channel switching (time-sharing) strategy, stochastic signals, and detectors is performed in order to achieve the minimum average probability of error. It is proved that the optimal solution to this problem corresponds to either (i) switching between at most two channels with deterministic signaling over each channel, or (ii) time-sharing between at most two different signals over a single channel (i.e., stochastic signaling over a single channel). For both cases, the optimal solutions are shown to employ corresponding maximum a posteriori probability (MAP) detectors at the receiver. Numerical results are presented to investigate the proposed approach.

## I. INTRODUCTION

Error performance of some communication systems that are subject to average power constraints can be improved via randomization (time sharing) of signal power [1]–[7]. For example, it is shown in [1] that transmitting a stochastic signal for each symbol instead of a deterministic signal can improve performance of a given receiver in terms of error probability. In particular, it is proved that an optimal stochastic signal can be represented by a randomization of no more than three different signal values under second and fourth moment constraints. Also, sufficient conditions are derived to specify if stochastic signaling can provide performance improvements over deterministic signaling. The study in [2] considers the jointly optimal design of stochastic signals and a detector under an average power constraint, and shows that the optimal solution results in stochastic signaling between at most two distinct signal values and the corresponding maximum a posteriori probability (MAP) detector at the receiver. Detector randomization, which involves the use of multiple detectors with certain probabilities, can also provide improvements in error performance of some communication systems [4], [6]. In other words, a receiver can time-share (randomize) among multiple detectors in order to achieve a lower average probability of error. In [4], an average power constrained binary communication system is considered, and randomization between two deterministic signal pairs and the corresponding MAP detectors is studied. Significant performance improvements are observed as a result of detector randomization in some cases in the presence of non-Gaussian noise. In [6], the results in [4] and [2] are generalized by considering both detector randomization and stochastic signaling, and it is shown that the joint optimization of stochastic signaling and detector randomization results in a randomization between at most two MAP detectors corresponding to two deterministic signals.

In the presence of multiple channels between a transmitter and a receiver, channel switching can be performed (i.e., each channel can be used for a certain fraction of time) in order to improve error performance of a communication system [7]–[9]. In [7], the convexity properties of error probability in terms of signal power are investigated for binary-valued scalar signals over additive noise channels under an average power constraint, and, based on the convexity results, it is concluded that the optimal channel switching strategy involves time-sharing between at most two different channels with antipodal signaling over each channel when the noise in each channel has a unimodal differentiable probability density function (PDF).

To the best of our knowledge, performance benefits due to stochastic signaling (i.e., performing signal randomization for the transmission of each symbol) has not been studied in the presence of multiple channels with generic (non-Gaussian and/or multimodal) noise PDFs. In this study, we investigate the joint optimization of the channel switching (time-sharing) strategy, stochastic signals, and detectors in order to achieve the minimum average probability of error for an average power constrained  $M$ -ary communication system. We prove that the optimal solution to this problem corresponds to either (i) switching between at most two channels with deterministic signaling over each channel and employing the corresponding MAP detector at the receiver, or (ii) time-sharing between at most two different signals over a single channel (i.e., stochastic signaling over a single channel) and employing the corresponding MAP detector at the receiver. Numerical examples are presented to investigate the theoretical results.

## II. OPTIMAL CHANNEL SWITCHING

Consider an  $M$ -ary communication system, in which the information can be conveyed between the transmitter and receiver over  $K$  additive noise channels as illustrated in Fig. 1. The transmitter is allowed to switch or time-share among these  $K$  channels to improve the correct decision performance at the receiver. A relay at the transmitter controls access to the channels so that only one of the channels can be used for symbol transmission at any given time. Furthermore, a stochastic signaling approach is adopted by treating the signals transmitted from each channel for each information symbol as random vectors instead of constant values [1]. In other words, the transmitter can perform randomization of signal values for each information symbol, which also corresponds to a form of constellation randomization [2], [10]. The transmitter and the receiver are assumed to be synchronized so that the receiver knows which channel is currently in use, and employs the optimal decision rule for the corresponding channel and the stochastic signaling scheme.

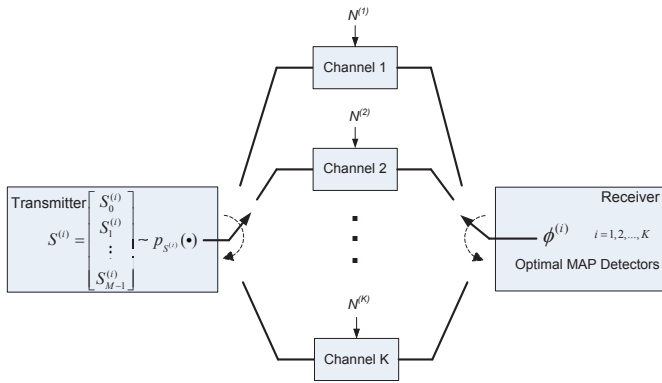


Fig. 1.  $M$ -ary communication system that employs stochastic signaling and channel switching.

We consider generic channels with possibly non-Gaussian/multimodal noise and aim to obtain the optimal channel switching strategy when stochastic signaling is possible over each channel. In this scenario, the noisy observation vector  $\mathbf{Y}$  received by the detector corresponding to the  $i$ th channel can be modeled as follows:

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

where  $\mathbf{S}_j^{(i)}$  represents the  $N$ -dimensional signal vector transmitted for symbol  $j$  over channel  $i$ , and  $\mathbf{N}^{(i)}$  is the noise component of the  $i$ th channel that is independent of  $\mathbf{S}_j^{(i)}$ . It should be emphasized that  $\mathbf{S}_j^{(i)}$  is modeled as a random vector to employ stochastic signaling. Also, the prior probabilities of the symbols, denoted by  $\pi_0, \pi_1, \dots, \pi_{M-1}$ , are assumed to be known. Although the signal model in (1) is in the form of a simple additive noise channel, it is sufficient to incorporate various effects such as thermal noise, multiple-access interference, and jamming [7]. It is also valid in the case of flat-fading channels assuming perfect channel estimation [1]. Note that the probability distribution of the noise component in (1) is not necessarily Gaussian since it is modeled to include the effects of interference and jamming as well [11].

The receiver uses the observation in (1) in order to determine the transmitted information symbol. For that purpose, a generic decision rule (detector) is considered for each channel making a total of  $K$  detectors getting utilized at the receiver. That is, for a given observation vector  $\mathbf{Y} = \mathbf{y}$ , the detector of the  $i$ th channel  $\phi^{(i)}(\mathbf{y})$  is described as

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

for  $j \in \{0, 1, \dots, M-1\}$ , where  $\Gamma_0^{(i)}, \Gamma_1^{(i)}, \dots, \Gamma_{M-1}^{(i)}$  are the decision regions (i.e., a partition of the observation space  $\mathbb{R}^N$ ) for the detector of the  $i$ th channel [12]. The transmitter and the receiver can switch among these  $K$  channels in any manner in order to optimize the probability of error performance. Let  $v_i$  denote the channel switching factor for channel  $i$ , where  $\sum_{i=1}^K v_i = 1$  and  $v_i \geq 0$  for  $i = 1, \dots, K$ . In other words, the transmitter and the receiver communicate over channel  $i$  for  $100v_i$  percent of the time ( $i = 1, \dots, K$ ).

The aim of this study is to jointly optimize the channel switching strategy ( $v_1, \dots, v_K$ ), stochastic signals, and detectors in order to achieve the maximum average probability of correct decision. The average probability of correct decision can be expressed as  $P_c = \sum_{i=1}^K v_i P_c^{(i)}$ , where  $v_i$  is the channel switching factor for channel  $i$ , and  $P_c^{(i)}$  represents the

corresponding probability of correct decision for that channel under  $M$ -ary signaling; that is,

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

for  $i = 1, 2, \dots, K$ , with  $p_j^{(i)}(\mathbf{y})$  denoting the conditional PDF of the observation when the  $j$ th symbol is transmitted over the  $i$ th channel. Since stochastic signaling is considered,  $\mathbf{S}_j^{(i)}$  in (1) is modeled as a random vector. Recalling that the signals and the noise are independent, the conditional PDF of the observation can be calculated as  $p_j^{(i)}(\mathbf{y}) = \int_{\mathbb{R}^N} p_{\mathbf{S}_j^{(i)}}(\mathbf{x}) p_{\mathbf{N}^{(i)}}(\mathbf{y} - \mathbf{x}) d\mathbf{x} = \mathbb{E}\{p_{\mathbf{N}^{(i)}}(\mathbf{y} - \mathbf{S}_j^{(i)})\}$ , where the expectation is over the PDF of  $\mathbf{S}_j^{(i)}$ . Then, the average probability of correct decision can be expressed as

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

In practical systems, there is a constraint on the average power emitted from the transmitter. Under the framework of stochastic signaling and channel switching, this constraint on the average power is expressed in the following form [12]:  $\sum_{i=1}^K v_i \left( \sum_{j=0}^{M-1} \pi_j \mathbb{E}\{\|\mathbf{S}_j^{(i)}\|_2^2\} \right) \leq A$ , where  $A$  denotes the average power limit.

In this study, we primarily concentrate on obtaining the optimal signaling and channel switching strategy in terms of the correct decision probability for an  $M$ -ary communication system. Although the channel switching problem is addressed in various frameworks, the novelty of the problem introduced here arises from the following two aspects: (i) signals transmitted over the channels corresponding to different symbols are modeled as random vectors (stochastic signaling) subject to an overall average power constraint, (ii) no restrictions are imposed on the noise PDFs of the channels available for switching. This formulation, in turn translates into a joint design problem over the channel switching factors  $\{v_i\}_{i=1}^K$ , channel-specific signal PDFs employed at the transmitter  $\{p_{\mathbf{S}_0^{(i)}}, p_{\mathbf{S}_1^{(i)}}, \dots, p_{\mathbf{S}_{M-1}^{(i)}}\}_{i=1}^K$ , and the corresponding optimal detectors used at the receiver  $\{\phi^{(i)}\}_{i=1}^K$ .

$$\begin{aligned} & \max \sum_{i=1}^K v_i \left( \sum_{j=0}^{M-1} \int_{\Gamma_j^{(i)}} \pi_j \mathbb{E}\{p_{\mathbf{N}^{(i)}}(\mathbf{y} - \mathbf{S}_j^{(i)})\} d\mathbf{y} \right) \\ & \text{subject to } \sum_{i=1}^K v_i \left( \sum_{j=0}^{M-1} \pi_j \mathbb{E}\{\|\mathbf{S}_j^{(i)}\|_2^2\} \right) \leq A, \quad (5) \\ & \sum_{i=1}^K v_i = 1, \quad v_i \geq 0, \quad \forall i \in \{1, 2, \dots, K\} \end{aligned}$$

where the maximization is to be performed over  $\{\phi^{(i)}, v_i, p_{\mathbf{S}_0^{(i)}}, p_{\mathbf{S}_1^{(i)}}, \dots, p_{\mathbf{S}_{M-1}^{(i)}}\}_{i=1}^K$ . Included in the above statement are the implicit assumptions stating that each  $p_{\mathbf{S}_j^{(i)}}(\cdot)$  should represent a PDF.

The signals for all the  $M$  symbols that are transmitted over channel  $i$  can be expressed as the elements of a random vector:  $\mathbf{S}^{(i)} \triangleq [\mathbf{S}_0^{(i)} \mathbf{S}_1^{(i)} \dots \mathbf{S}_{M-1}^{(i)}] \in \mathbb{R}^{MN}$ , where  $\mathbf{S}_j^{(i)}$ 's are  $N$ -dimensional row vectors  $\forall j \in \{0, 1, \dots, M-1\}$ . Then, a more compact version of the optimization problem in (5) can be stated as follows:

$$\begin{aligned} & \max_{\{\phi^{(i)}, v_i, p_{\mathbf{S}^{(i)}}\}_{i=1}^K} \sum_{i=1}^K v_i \mathbb{E}\{G_i(\mathbf{S}^{(i)})\} \\ & \text{subject to } \sum_{i=1}^K v_i \mathbb{E}\{H(\mathbf{S}^{(i)})\} \leq A, \\ & \sum_{i=1}^K v_i = 1, v_i \geq 0, \forall i \in \{1, 2, \dots, K\} \end{aligned} \quad (6)$$

where  $G_i(\mathbf{S}^{(i)}) = \sum_{j=0}^{M-1} \int_{\Gamma_j^{(i)}} \pi_j p_{\mathbf{N}^{(i)}}(\mathbf{y} - \mathbf{S}_j^{(i)}) d\mathbf{y}$ ,  $H(\mathbf{S}^{(i)}) = \sum_{j=0}^{M-1} \pi_j \|\mathbf{S}_j^{(i)}\|_2^2$ , and the expectations are taken with respect to  $p_{\mathbf{S}^{(i)}}(\cdot)$ 's, which denote the joint PDFs of the signals for symbols  $\{0, 1, \dots, M-1\}$  that are transmitted over channel  $i$ . Specifically,  $G_i(\mathbf{s}^{(i)})$  represents the average probability of correct decision when the deterministic signal vector  $\mathbf{s}^{(i)}$  is used for the transmission of  $M$  symbols over the additive noise channel  $i$  and the corresponding detector is employed at the receiver. Then,  $\mathbb{E}\{G_i(\mathbf{S}^{(i)})\}$  can be interpreted as a generic stochastic signaling scheme over channel  $i$ . The exact number of signal values employed by the scheme is determined by the number of distinct values that the random vector  $\mathbf{S}^{(i)}$  can take.

Let  $P_c^\dagger$  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. Let  $G(\mathbf{x})$  denote the maximum of the average probabilities of correct decision when the deterministic signal vector  $\mathbf{x}$  is used for the transmission of  $M$  symbols over the additive noise channels  $i = 1, 2, \dots, K$  and the corresponding detectors for all  $K$  channels are employed at the receiver. That is,  $G(\mathbf{x}) \triangleq \max_{i \in \{1, 2, \dots, K\}} G_i(\mathbf{x})$ , from which  $G(\mathbf{x}) \geq G_i(\mathbf{x})$  follows  $\forall i \in \{1, 2, \dots, K\}$  and  $\forall \mathbf{x} \in \mathbb{R}^{MN}$ . This inequality can be applied to the objective function in (6) to obtain a new optimization problem that provides an upper bound on the solution of the optimization problem in (6) as follows.

$$\begin{aligned} & \max_{\{\phi^{(i)}, v_i, p_{\mathbf{S}^{(i)}}\}_{i=1}^K} \sum_{i=1}^K v_i \mathbb{E}\{G(\mathbf{S}^{(i)})\} \\ & \text{subject to } \sum_{i=1}^K v_i \mathbb{E}\{H(\mathbf{S}^{(i)})\} \leq A, \\ & \sum_{i=1}^K v_i = 1, v_i \geq 0, \forall i \in \{1, 2, \dots, K\} \end{aligned} \quad (7)$$

where the expectations are taken with respect to  $p_{\mathbf{S}^{(i)}}(\cdot)$ 's, which denote the joint PDFs of the signals for symbols  $\{0, 1, \dots, M-1\}$  that are transmitted over channel  $i$ .

In order to achieve further simplification of the problem in (7), define  $p_{\check{\mathbf{S}}}(\check{\mathbf{s}}) \triangleq \sum_{i=1}^K v_i p_{\mathbf{S}^{(i)}}(\check{\mathbf{s}})$ , where  $\check{\mathbf{s}} \triangleq [\check{s}_0 \check{s}_1 \dots \check{s}_{M-1}] \in \mathbb{R}^{MN}$ , and  $\check{s}_j$ 's are  $N$ -dimensional row vectors  $\forall j \in \{0, 1, \dots, M-1\}$ . 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  $\mathbb{R}^{MN}$ ,  $p_{\check{\mathbf{S}}}(\check{\mathbf{s}})$  satisfies the conditions to be a PDF. Then, the optimization problem in (7) can be written in the following form:

$$\max_{p_{\check{\mathbf{S}}}, \{\phi^{(i)}\}_{i=1}^K} \mathbb{E}\{G(\check{\mathbf{S}})\} \quad \text{subject to} \quad \mathbb{E}\{H(\check{\mathbf{S}})\} \leq A \quad (8)$$

where  $G(\check{\mathbf{s}}) = \max_{i \in \{1, 2, \dots, K\}} G_i(\check{\mathbf{s}})$  for all  $\check{\mathbf{s}} \in \mathbb{R}^{MN}$ , and the

expectations are taken with respect to  $p_{\check{\mathbf{S}}}(\cdot)$ , which denotes the PDF of the signal constellation employed for transmission of symbols  $\{0, 1, \dots, M-1\}$ . Let  $P_c^*$  denote the maximum average probability of correct decision obtained as the solution to the optimization problem in (8). By definition of function  $G$ ,  $P_c^* \geq P_c^\dagger$  is always satisfied.

In (8),  $G(\check{\mathbf{s}})$  represents the maximum of the average probabilities of correct decision when the deterministic signal vector  $\check{\mathbf{s}}$  is used for the transmission of  $M$  symbols over the additive noise channels  $i = 1, 2, \dots, K$  and the corresponding detectors are employed at the receiver. Therefore,  $\mathbb{E}\{G(\check{\mathbf{S}})\}$  can be interpreted as a randomization (switching) among channels with respect to the PDF  $p_{\check{\mathbf{S}}}(\cdot)$ , where the average probability of correct decision for each signal vector  $\check{\mathbf{s}}$  in the support of  $p_{\check{\mathbf{S}}}$  is maximized by transmitting it over the *most favorable channel* (i.e., the channel with the highest probability of correct decision for the given signal vector  $\check{\mathbf{s}}$ ), and altogether they maximize the total average probability of correct decision.

Optimization problems in the form of (8) have been investigated in various studies in the literature such as [1], [4], [6], [13]. Assume that  $G_i(\mathbf{s})$  in (6) is a continuous function  $\forall i \in \{1, 2, \dots, K\}$  and  $\mathbf{a} \preceq \mathbf{s} \preceq \mathbf{b}$  where  $\mathbf{a}$  and  $\mathbf{b}$  are finite real vectors in  $\mathbb{R}^{MN}$ . Then, the pointwise maximum  $G(\mathbf{s}) = \max_{i \in \{1, 2, \dots, K\}} G_i(\mathbf{s})$  is also continuous on  $[\mathbf{a}, \mathbf{b}]$ , and the optimal solution of (8) can be represented by a randomization of at most two signal levels as a result of Carathéodory's theorem [14]; that is,  $p_{\check{\mathbf{S}}}^{\text{opt}}(\check{\mathbf{s}}) = \lambda \delta(\check{\mathbf{s}} - \mathbf{s}_1) + (1 - \lambda) \delta(\check{\mathbf{s}} - \mathbf{s}_2)$ . Therefore, the problem in (8) can be solved over such signal PDFs resulting in

$$\max_{\{\lambda, \mathbf{s}_1, \mathbf{s}_2, \{\phi^{(i)}\}_{i=1}^K\}} \lambda G(\mathbf{s}_1) + (1 - \lambda) G(\mathbf{s}_2) \quad (9)$$

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

where  $G(\mathbf{s}_k) = \max_{i \in \{1, 2, \dots, K\}} G_i(\mathbf{s}_k)$ ,  $G_i(\mathbf{s}_k) = \sum_{j=0}^{M-1} \int_{\Gamma_j^{(i)}} \pi_j p_{\mathbf{N}^{(i)}}(\mathbf{y} - \mathbf{s}_{k,j}) d\mathbf{y} \quad \forall i \in \{1, 2, \dots, K\}$ ,  $H(\mathbf{s}_k) = \sum_{j=0}^{M-1} \pi_j \|\mathbf{s}_{k,j}\|_2^2$ , and  $\mathbf{s}_k = [\mathbf{s}_{k,0} \mathbf{s}_{k,1} \dots \mathbf{s}_{k,M-1}] \in \mathbb{R}^{MN}$  with  $\mathbf{s}_{k,j}$  being an  $N$ -dimensional row vector  $\forall j \in \{0, 1, \dots, M-1\}$ . Therefore, optimal performance can be achieved by randomizing between at most two signal vectors,  $\mathbf{s}_1$  and  $\mathbf{s}_2$ . From (9), it is deduced that the objective function is maximized under the specified constraints by either one of the following strategies:

- 1) transmitting exclusively over a single channel via deterministic signaling,  $\lambda \in \{0, 1\}$ ,
- 2) time sharing (or randomization) between two signal vectors over a single channel,  $\lambda \in (0, 1)$  and  $\arg \max_{i \in \{1, 2, \dots, K\}} G_i(\mathbf{s}_1) = \arg \max_{i \in \{1, 2, \dots, K\}} G_i(\mathbf{s}_2)$ ,
- 3) switching (or randomization) between two channels and deterministic signaling over each one,  $\lambda \in (0, 1)$  and  $\arg \max_{i \in \{1, 2, \dots, K\}} G_i(\mathbf{s}_1) \neq \arg \max_{i \in \{1, 2, \dots, K\}} G_i(\mathbf{s}_2)$ .

The last step in the simplification of the optimization problem in (9) comes from an observation about the structure of optimal detectors. For a given channel  $i$  and the corresponding signaling scheme over the channel (deterministic or randomization between two signal vectors), the conditional probability of the observation  $\mathbf{y}$  under the hypothesis  $j$  (i.e., when symbol  $j$  is transmitted) is expressed as



$$p_j^{(i)}(\mathbf{y}) = \mathbb{E}\{p_{\mathbf{N}^{(i)}}(\mathbf{y} - \mathbf{S}_j^{(i)})\} = \begin{cases} p_{\mathbf{N}^{(i)}}(\mathbf{y} - \mathbf{s}_j^{(i)}), & \text{if deterministic} \\ \lambda p_{\mathbf{N}^{(i)}}(\mathbf{y} - \mathbf{s}_{1,j}^{(i)}) + (1 - \lambda)p_{\mathbf{N}^{(i)}}(\mathbf{y} - \mathbf{s}_{2,j}^{(i)}), & \text{if randomized} \end{cases} \quad (10)$$

When deciding among  $M$  symbols based on observation  $\mathbf{y}$ , the MAP decision rule selects symbol  $j$  if  $j = \arg \max_{l \in \{0, 1, \dots, M-1\}} \pi_l p_l^{(i)}(\mathbf{y})$ , and it maximizes the average probability of correct decision [12]. Therefore, it is not necessary to search over all decision rules in (9); only the MAP decision rule should be determined for the detector of each channel [2]. 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  $i$ th channel becomes

$$P_{c, \text{MAP}}^{(i)} = \int_{\mathbb{R}^N} \max_{j \in \{0, 1, \dots, M-1\}} \left\{ \pi_j p_j^{(i)}(\mathbf{y}) \right\} d\mathbf{y} \quad (11)$$

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

where  $\mathbb{E}\{p_{\mathbf{N}^{(i)}}(\mathbf{y} - \mathbf{S}_j^{(i)})\}$  is as in (10).

Below, more explicit forms of the optimization problem stated in (9) are given for all possible scenarios mentioned previously.

**1. Transmitting Exclusively over a Single Channel via Deterministic Signaling:** In this case, a single channel is utilized exclusively, and the transmitted signal for each symbol is deterministic; i.e., a fixed signal vector is transmitted over the channel for each symbol. Without loss of generality, channel  $i$  is considered. Using the result given in (11) for the deterministic case, the problem in (9) can be expressed as

$$\max_{\mathbf{s}^{(i)}} \int_{\mathbb{R}^N} \max_{j \in \{0, 1, \dots, M-1\}} \left\{ \pi_j p_{\mathbf{N}^{(i)}}(\mathbf{y} - \mathbf{s}_j^{(i)}) \right\} d\mathbf{y}$$

subject to  $\sum_{j=0}^{M-1} \pi_j \|\mathbf{s}_j^{(i)}\|_2^2 \leq A.$  (12)

**2. Time-Sharing (or Randomization) between at most Two Signal Vectors over a Single Channel:** Similar to the previous case, the transmission occurs over a single channel exclusively, but in this case the transmitted signal for each symbol is a randomization of at most two different signal vectors. Without loss of generality, channel  $i$  is considered. Using the result for the randomized case in (11), the optimization problem in (9) can be expressed as

$$\max_{\{\lambda, \mathbf{s}_1^{(i)}, \mathbf{s}_2^{(i)}\}} \int_{\mathbb{R}^N} \max_{j \in \{0, 1, \dots, M-1\}} \left\{ \pi_j p_j^{(i)}(\mathbf{y}) \right\} d\mathbf{y}$$

subject to  $\lambda \left( \sum_{j=0}^{M-1} \pi_j \|\mathbf{s}_{1,j}^{(i)}\|_2^2 \right) + (1 - \lambda) \left( \sum_{j=0}^{M-1} \pi_j \|\mathbf{s}_{2,j}^{(i)}\|_2^2 \right) \leq A, \lambda \in [0, 1]$  (13)

where  $p_j^{(i)}(\mathbf{y}) = \lambda p_{\mathbf{N}^{(i)}}(\mathbf{y} - \mathbf{s}_{1,j}^{(i)}) + (1 - \lambda) p_{\mathbf{N}^{(i)}}(\mathbf{y} - \mathbf{s}_{2,j}^{(i)})$ . It is noted that the optimization problem in (13) reduces to that of (12) when  $\lambda \in \{0, 1\}$ .

**3. Switching (or Randomization) between at most Two Channels and Deterministic Signaling over each Channel:** In this case, optimum performance is investigated while transmitting over at most two channels and the transmission over

each channel is constrained to be deterministic, i.e., a fixed signal vector is sent over each channel for each symbol but the channels are switched in time. Without loss of generality, channels  $i$  and  $l$  are considered ( $i \neq l$  and  $i, l \in \{1, 2, \dots, K\}$ ). Since deterministic signaling is employed in each channel, the result given in (11) for the deterministic case should be applied for each channel. Then, the optimization problem in (9) becomes

$$\max_{\{\lambda, \mathbf{s}^{(i)}, \mathbf{s}^{(l)}\}} \lambda G_{i, \text{MAP}}(\mathbf{s}^{(i)}) + (1 - \lambda) G_{l, \text{MAP}}(\mathbf{s}^{(l)}) \quad (14)$$

subject to  $\lambda H(\mathbf{s}^{(i)}) + (1 - \lambda) H(\mathbf{s}^{(l)}) \leq A, \lambda \in [0, 1]$

where  $H(\mathbf{s}^{(i)}) = \sum_{j=0}^{M-1} \pi_j \|\mathbf{s}_j^{(i)}\|_2^2$ ,  $G_{i, \text{MAP}}(\mathbf{s}^{(i)}) = \int_{\mathbb{R}^N} \max_{j \in \{0, 1, \dots, M-1\}} \left\{ \pi_j p_{\mathbf{N}^{(i)}}(\mathbf{y} - \mathbf{s}_j^{(i)}) \right\} d\mathbf{y}$ , and  $G_l(\mathbf{s}^{(l)})$  and  $H(\mathbf{s}^{(l)})$  are defined similarly by replacing  $i$  with  $l$  in the corresponding equations.

It is noted that the optimization space is considerably reduced in (12)-(14) compared to that in (9) since there is no need to search over the detectors in (12)-(14).

In the view of the above analysis, the solution of the optimization problem in (9) can be decomposed into two parts. First, time-sharing (or randomization) between at most two signal vectors over a single channel is considered. Let  $P_{c, \text{Opt}}^{(i)}$  be the solution of the optimization problem in (13) for  $i$ th channel; that is,  $P_{c, \text{Opt}}^{(i)}$  denotes the maximum average probability of correct decision that can be achieved by stochastic signaling over channel  $i$  under average power constraint. Secondly, switching (or randomization) between at most two channels with deterministic signaling over each channel is considered. Let  $P_{c, \text{Opt}}^{(i, l)}$  be the solution of the optimization problem in (14) for channels  $i$  and  $l$ ; that is,  $P_{c, \text{Opt}}^{(i, l)}$  denotes the maximum average probability of correct decision that can be achieved by switching between channels  $i$  and  $l$  under average power constraint. Then, the solution of the optimization problem in (9) can be obtained by solving the following set of optimization sub-problems and computing their maximum:

$$P_c^{\text{Stoc}} = \max_{i \in \{1, 2, \dots, K\}} P_{c, \text{Opt}}^{(i)}$$

$$P_c^{\text{CS}} = \max_{i, l \in \{1, 2, \dots, K\} \text{ and } i < l} P_{c, \text{Opt}}^{(i, l)}$$

$$P_c^* = \max \{P_c^{\text{Stoc}}, P_c^{\text{CS}}\} \quad (15)$$

where the superscript Stoc denotes stochastic signaling over a single channel and CS represents channel switching.

Finally, the following proposition is presented.

**Proposition 1:** Let  $P_c^*$  denote the maximum probability of correct decision obtained as the solution of the optimization problem in (6). Then,  $P_c^*$  is equal to  $P_c^*$  in (15).

The proof is omitted due to the space limitation.

Proposition 1 implies that the solution of the original optimization problem stated in (6), which considers the joint optimization of switching factors among channels, channel-specific signal PDFs employed at the transmitter and the corresponding detectors used at the receiver, can be obtained as the solution of the much simpler optimization problem specified in (15). Formally, *when multiple channels are available for signal transmission (i.e.,  $K \geq 2$ ), it is sufficient to either employ switching between two channels with two deterministic signal vectors (i.e., there is no need to employ stochastic signaling over a channel to achieve the optimal solution while switching*

channels); or employ stochastic signaling with at most two signal vectors over a single channel.

### III. NUMERICAL RESULTS AND CONCLUSIONS

For a numerical example, a scalar binary communication system with equiprobable information symbols is considered and the average power limit is set to  $A = 1$ . It is assumed that  $K \geq 2$  channels are available between the transmitter and the receiver for channel switching purposes. The following four schemes are considered for performance comparison.

**Gaussian Solution over a Single Channel:** Antipodal signals in the form of  $(-\sqrt{A}, \sqrt{A})$  are transmitted for binary information symbols, and the corresponding MAP detectors are employed at the receivers. Among the  $K$  channels, the most favorable channel, the one that yields the highest average probability of correct decision, is chosen for transmission.

**Optimal Deterministic Solution over a Single Channel:** Deterministic signaling with corresponding MAP detectors are considered for all the channels.  $K$  optimization problems in the form of (12) are solved and the most favorable channel is chosen for transmission.

**Optimal Stochastic Solution over a Single Channel:** This scheme employs stochastic signaling and corresponding MAP detectors for all the channels. The problem in (13) is solved for all the channels and the most favorable channel is chosen for transmission (see the first expression in (15)).

**Optimal Channel Switching with Deterministic Signaling:** Deterministic signaling with corresponding MAP detectors are employed for all the channels and the optimal switching strategy is obtained. Since optimal switching requires at most two channels,  $K(K-1)/2$  problems in (14) are solved and the channel pair that provides the highest performance is selected (see the second expression in (15)).

For the numerical example, the noise of all the  $K$  channels is modeled by a Gaussian mixture model [5], [11], [13], which is represented by the following PDF:  $p_{N^{(i)}}(n) = (\sqrt{2\pi}\sigma_i L_i)^{-1} \sum_{l=1}^{L_i} \exp\{- (n - \mu_l^{(i)})^2 / (2\sigma_i^2)\}$  for  $i \in \{1 \dots K\}$ , where  $L_i$  is the number of components in the mixture for channel  $i$ . It is noted that the components of the Gaussian mixture noise in each channel are modeled to have the same weight ( $1/L_i$ ) and the same variance ( $\sigma_i^2$ ). For simplicity of notation, the mean values of the Gaussian mixture components for channel  $i$  are collected in a vector as  $\boldsymbol{\mu}^{(i)} = [\mu_1^{(i)} \dots \mu_{L_i}^{(i)}]$ . Then, the average noise power of the  $i$ th channel can be calculated as  $\mathbb{E}\{|N^{(i)}|^2\} = \sigma_i^2 + \frac{1}{L_i} \|\boldsymbol{\mu}^{(i)}\|_2^2$ , where  $\|\boldsymbol{\mu}^{(i)}\|_2$  denotes the  $L_2$  norm of vector  $\boldsymbol{\mu}^{(i)}$ .

We consider a scenario in which all the channels have distinct noise PDFs. For the Gaussian mixture noise model, it is assumed that  $\sigma_i = \sigma$  and  $L_i = L$ ,  $\forall i \in \{1, \dots, K\}$ , and that the component means of the Gaussian mixture are chosen as  $\boldsymbol{\mu}^{(i)} = \sqrt{E} \mathbf{v}_i / \|\mathbf{v}_i\|_2$  for  $i = 1, \dots, K$ , where  $E$  is a constant and  $\mathbf{v}_i$ 's are  $L$ -dimensional distinct vectors. It is noted that  $\|\boldsymbol{\mu}^{(i)}\|_2^2 = E$ . Hence, the average noise power is the same for all the channels. Namely,  $\mathbb{E}\{|N^{(i)}|^2\} = \sigma^2 + E/L$ ,  $\forall i \in \{1, \dots, K\}$ . In Fig. 2, the average probabilities of error for the four strategies are plotted versus  $A/\sigma^2$  for  $K = 3$ ,  $\mathbf{v}_1 = [-3 -2 0 2 3]$ ,  $\mathbf{v}_2 = [-4 -3 0 3 4]$ ,  $\mathbf{v}_3 = [-5 -3 0 3 5]$ , and  $E = 3$ . From Fig. 2, it is observed that the Gaussian solution has the worst performance among all the approaches as expected since it is optimized for Gaussian noise and is not expected to achieve high performance in the presence of multimodal noise PDFs. When optimal deterministic signaling is employed (see (12)), significant gains can be achieved over

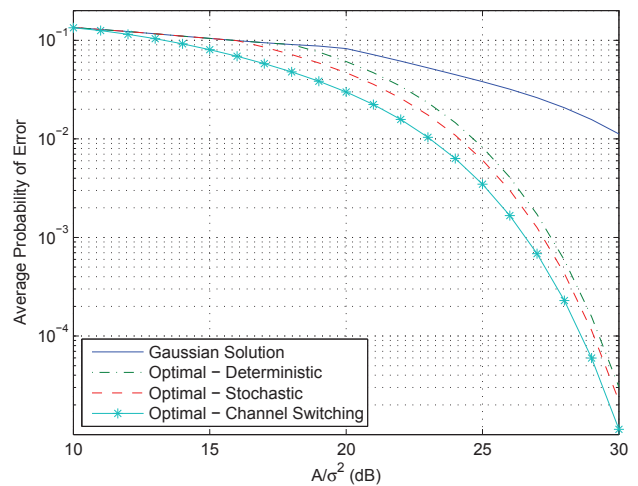


Fig. 2. Average probability of error versus  $A/\sigma^2$  for various approaches, where  $K = 3$ ,  $\mathbf{v}_1 = [-3 -2 0 2 3]$ ,  $\mathbf{v}_2 = [-4 -3 0 3 4]$ ,  $\mathbf{v}_3 = [-5 -3 0 3 5]$ , and  $E = 3$ .

the Gaussian solution in this example. In addition, further improvements are possible when stochastic signaling is used instead of deterministic signaling (see (13)). However, it is observed that the best performance is achieved when optimal channel switching is performed between at most two deterministic signals.

Additional results are obtained when all the channels have identical noise PDFs, and similar observations to those for Fig. 2 are made. However, the results are not presented due to the space limitation.

### REFERENCES

- [1] C. Goken, S. Gezici, and O. Arikan, "Optimal stochastic signaling for power-constrained binary communications systems," *IEEE Trans. Wirel. Commun.*, vol. 9, no. 12, pp. 3650–3661, Dec. 2010.
- [2] —, "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.
- [3] H. Chen and P. K. Varshney, "Theory of the stochastic resonance effect in signal detection: Part II—Variable detectors," *IEEE Trans. Sig. Process.*, vol. 56, no. 10, pp. 5031–5041, Oct. 2007.
- [4] A. Patel and B. Kosko, "Optimal noise benefits in Neyman-Pearson and inequality-constrained signal detection," *IEEE Trans. Sig. Process.*, vol. 57, no. 5, pp. 1655–1669, May 2009.
- [5] S. Bayram, S. Gezici, and H. V. Poor, "Noise enhanced hypothesis-testing in the restricted Bayesian framework," *IEEE Trans. Sig. Process.*, vol. 58, no. 8, pp. 3972–3989, Aug. 2010.
- [6] B. Dulek and S. Gezici, "Detector randomization and stochastic signaling for minimum probability of error receivers," *IEEE Trans. Commun.*, vol. 60, no. 4, pp. 923–928, Apr. 2012.
- [7] M. Azizoglu, "Convexity properties in binary detection problems," *IEEE Trans. Inform. Theory*, vol. 42, no. 4, pp. 1316–1321, July 1996.
- [8] Y. Ma and C. C. Chai, "Unified error probability analysis for generalized selection combining in Nakagami fading channels," *IEEE Jour. Select. Areas Commun.*, vol. 18, no. 11, pp. 2198–2210, Nov. 2000.
- [9] J. A. Ritcey and M. Azizoglu, "Performance analysis of generalized selection combining with switching constraints," *IEEE Commun. Letters*, vol. 4, no. 5, pp. 152–154, May 2000.
- [10] E. G. Larsson, "Improving the frame-error-rate of spatial multiplexing in block fading by randomly rotating the signal constellation," *IEEE Commun. Letters*, vol. 8, no. 8, pp. 514–516, Aug. 2004.
- [11] V. Bhatia and B. Mulgrew, "Non-parametric likelihood based channel estimator for Gaussian mixture noise," *Signal Process.*, vol. 87, pp. 2569–2586, Nov. 2007.
- [12] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1994.
- [13] 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. Process.*, vol. 55, no. 7, pp. 3172–3184, July 2007.
- [14] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ: Princeton University Press, 1968.