Optimal Channel Switching Over Gaussian Channels Under Average Power and Cost Constraints

Mehmet Emin Tutay, Sinan Gezici, *Senior Member, IEEE*, Hamza Soganci, and Orhan Arikan, *Member, IEEE*

***Abstract*—Optimal channel switching that provides the highest performance over a set of Gaussian channels with variable utilization costs is investigated in the presence of average power and average cost constraints. First, generic cost functions are considered, and it is shown that the optimal channel switching strategy performs channel switching (time sharing) among *at most three* different channels and always operates at the average power *and* average cost limits. Also, for channel switching between two channels, relations between the optimal power levels are obtained depending on the average power limit, and it is proved that the ratio of the optimal power levels is upper bounded by the ratio of the larger noise variance to the smaller one. In addition, for logarithmic cost functions, the convexity properties of the error probability are investigated as a function of power and cost, and the optimal channel switching strategy is shown to employ at most two channels, which can easily be determined based on specific formulas, when the average power limit is larger than a certain threshold. Numerical examples are presented to provide illustrations of the theoretical results.**

***Index Terms*—Channel switching, Gaussian channel, time sharing, probability of error.**

I. Introduction

TIME sharing among different power levels, detectors, or channels can provide performance improvements for communication systems that operate under average power constraints and in the presence of additive time-invariant noise [1]–[12]. For example, the average probability of error for some communication systems that are subject to multimodal noise can be reduced by performing time sharing between two different signal levels for each information symbol [2]. In other words, instead of transmitting a constant signal value for each information symbol, performing "randomization" (time sharing) among multiple signal values can result in performance improvements in certain cases [2], [13], [14]. Similarly, jammer systems can achieve improved jamming performance by time sharing among multiple power levels [1], [4], [5]. In [1], it is shown that a weak jammer should employ on-off time sharing to maximize the average probability of error of a receiver that operates in the presence of zero-mean symmetric noise, such as Gaussian noise. The study in [5] investigates the optimum power allocation policy for an average power constrained jammer operating over an arbitrary additive noise channel, where the aim is to minimize the detection probability of an instantaneously and fully adaptive receiver that employs the Neyman-Pearson criterion. It is shown that the optimum jamming performance can be achieved via time sharing between at most two different power levels, and a necessary and sufficient condition is provided for the improvability of the jamming performance via time sharing of the power compared to fixed power jamming schemes.

Time sharing among multiple detectors, which is also called *detector randomization*, presents another approach for improving error performance of average power constrained communication systems that operate over an additive time-invariant noise channel [6]–[9], [15], [16]. In this approach, a receiver has multiple detectors and employs one of them at any given time according to a certain time sharing strategy. In [6], an average power constrained binary communication system is considered, and the optimal time sharing between two antipodal signal pairs and the corresponding maximum *a posteriori* probability (MAP) detectors is investigated. Significant performance improvements can be achieved as a result of the proposed approach in the presence of symmetric Gaussian mixture noise for a certain range of average power limits. In [7], the results in [2] and [6] are generalized by considering an average power constrained M-ary communication system that can employ time sharing among both signal levels and detectors over an additive noise channel with some known distribution. It is proved that the joint optimization of the transmitted signals and the detectors at the receiver results in time sharing between at most two MAP detectors corresponding to two deterministic signal constellations. [9] investigates the benefits of time sharing among multiple detectors for the downlink of a multiuser communication system and characterizes the optimal time sharing strategy. In a related study, the form of the optimal additive noise is obtained for variable detectors in the context of noise enhanced detection under both Neyman-Pearson and Bayesian frameworks [8].

In the presence of multiple channels between a transmitter and a receiver, performing time sharing among different

Manuscript received July 31, 2014; revised January 6, 2015; accepted March 22, 2015. Date of publication March 31, 2015; date of current version May 14, 2015. This research was supported in part by the National Young Researchers Career Development Programme (project no. 110E245) of the Scientific and Technological Research Council of Turkey (TUBITAK) and in part by the Distinguished Young Scientist Award of Turkish Academy of Sciences (TUBA-GEBIP 2013). The associate editor coordinating the review of this paper and approving it for publication was P. Popovski.

M. E. Tutay, S. Gezici, and O. Arikan are with the Department of Electrical and Electronics Engineering, Bilkent University, Ankara 06800, Turkey (e-mail: tutay@ee.bilkent.edu.tr; gezici@ee.bilkent.edu.tr; orikan@ee.bilkent.edu.tr).

H. Soganci is with the Department of Electrical and Electronics Engineering, Bilkent University, Ankara 06800, Turkey, and also with the TUBITAK-SAGE, Group of Electronic Systems and Flight Disciplines, Ankara 06261, Turkey (e-mail: hsoganci@ee.bilkent.edu.tr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCOMM.2015.2418252

channels, which is called *channel switching*, can provide certain performance improvements [1], [10]–[12], [17]. In the channel switching approach, communication occurs over one channel for a certain fraction of time, and then it switches to another channel during the next transmission. In [1], the channel switching problem is studied under an average power constraint for the optimal detection of binary antipodal signals over a number of channels that are subject to additive unimodal noise. It is shown that the optimal solution is either to communicate over one channel exclusively, or to switch between two channels with a certain time sharing factor. In [12], the channel switching problem is investigated for M-ary communication systems in the presence of additive noise channels with arbitrary probability distributions and by facilitating time sharing among multiple signal constellations over each channel. Under an average power constraint, the optimal solution that minimizes the average probability of error is obtained as one of the following strategies: deterministic signaling (i.e., use of one signal constellation) over a single channel; time sharing between two different signal constellations over a single channel; or switching (time sharing) between two channels with deterministic signaling over each channel [12]. In a different context, the concept of channel switching is employed for cognitive radio systems with opportunistic spectrum access, where a number of secondary users try to access the available frequency bands in the spectrum [18]–[20].

Although the channel switching problem has been investigated thoroughly under an average power constraint (e.g., [1], [12]), no studies have considered the cost of communications over different channels in obtaining the optimal channel switching strategy. In practical systems, each channel can be associated with a certain cost depending on its quality [21]–[26]. For example, a channel that presents high signal-to-noise ratio (SNR) conditions has a high cost (price) compared to channels with low SNRs [22], [25]. Therefore, it is important to consider costs of different channels while designing a channel switching strategy. In this study, the optimal channel switching problem is formulated for Gaussian channels in the presence of average power and average cost constraints. First, generic cost values are considered for the channels and the optimal channel switching strategy is characterized. Then, logarithmic cost functions are employed to relate the cost of a channel to its average noise power [26], and specific results are obtained about the optimality of channel switching between two channels or among three channels. Finally, numerical examples are presented to explain the theoretical results. The main contributions of this study can be summarized as follows:

- The optimal channel switching problem over Gaussian channels is investigated under an average cost and average power constraint for the first time.
- For generic cost functions, it is shown that the optimal channel switching strategy is to switch among *at most three* different channels (Proposition 2), and that the optimal strategy must operate at the average cost *and* average power limits (Proposition 1).
- For channel switching between two channels, relations between the optimal power levels are obtained depending on the average power constraint, and it is proved that the ratio of the optimal power levels is upper bounded by the ratio of the larger noise variance to the smaller one under certain conditions (Proposition 3).
- When cost values are related to average noise powers according to a specific logarithmic relation [26], it is shown for sufficiently high power limits that the optimal channel switching strategy involves at most two channels (Proposition 5) and that the optimal channel switching between two channels can easily be specified based on the average cost limit (Proposition 4).

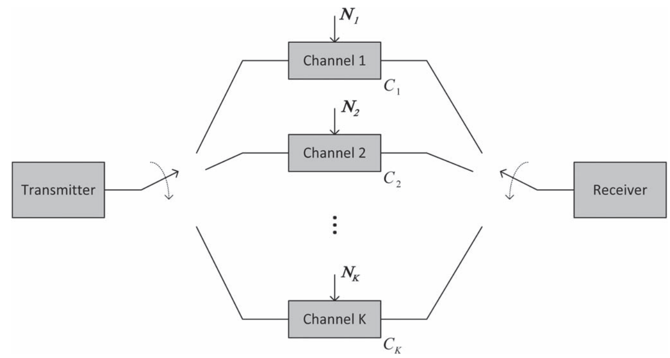

Fig. 1. A communication system that employs channel switching among K additive Gaussian noise channels, where C_i denotes the cost of using channel i, and $\boldsymbol{N}_i$ is the noise component at the ith channel.

A motivating application scenario for the proposed problem is a cognitive radio system in which primary users are the main owners of the spectrum, and secondary users can utilize the frequency bands of primary users under certain conditions. As discussed in [21], the frequency owners (primary users) can sell certain part of their spectrum to secondary users for the aim of maximizing their revenue. From the perspective of a secondary user, there can exist multiple available frequency bands (channels) with different costs in this framework (see Fig. 1). Then, the aim of a secondary user is to optimize its performance under a certain cost constraint (budget). More specifically, among the available channels in the spectrum (which have certain cost values), a secondary user can perform optimal channel switching to minimize its average probability of error under an average cost constraint (together with power constraints that are related to hardware constraints and/or battery life). Hence, the proposed problem formulation is important for cognitive radio systems in terms of performance optimization of secondary users under realistic constraints. In addition, the formulation also carries theoretical significance since the costs of different channels have not been considered in the previous studies on channel switching [1], [10]–[12], [17].

The remainder of the manuscript is organized as follows: The system model and problem formulation are introduced in Section II, and the optimal channel switching problem is studied for generic cost functions in Section III. In Section IV, logarithmic functions are considered for the optimal channel switching problem, and numerical examples are presented in Section V. Finally, various extensions and some concluding remarks are provided in Section VI and Section VII, respectively.

II. System Model and Problem Formulation

Consider a communication system in which K additive noise channels are available between the transmitter and the receiver, as shown in Fig. 1. The transmitter-receiver pair can synchronously switch among those K channels over time; i.e., they can perform time sharing among different channels by employing only one channel at a given time [1], [12]. The channels are corrupted by independent zero-mean Gaussian noise components, denoted by $\boldsymbol{N}_i$ for $i \in \{1,\dots,K\}$. For channel i, the components of $\boldsymbol{N}_i$ are independent and identically distributed with a variance of σ_i^2. In addition, there is a cost associated with the usage of each channel, denoted by C_i for $i \in \{1,\dots,K\}$. The cost values are specified by nonnegative numbers and they satisfy $C_i > C_j$ if $\sigma_i^2 < \sigma_j^2$ for all $i \neq j$. In other words, if a channel has a smaller (larger) average noise power, it has a higher (lower) cost. Assigning costs to different channels or measurement devices has various motivations and implications, as discussed for example in [21]–[26].

A generic M-ary communication system is considered, where the received signal corresponding to the ith channel is expressed as

$$\boldsymbol{Y} = \sqrt{P_i}\; \boldsymbol{s}_i^{(j)} + \boldsymbol{N}_i,\ j \in \{0,1,\dots,M-1\},\ i \in \{1\dots K\}, \quad (1)$$

where $\boldsymbol{s}_i^{(0)}, \boldsymbol{s}_i^{(1)}, \dots, \boldsymbol{s}_i^{(M-1)}$ represent the signal constellation employed for communication over channel i, P_i denotes the average power of the transmitted signal (assuming normalization for the average energy of the signal constellation), and $\boldsymbol{N}_i$ is the Gaussian noise over channel i, which is independent of $\boldsymbol{s}_i^{(j)}$. It is assumed that the symbols are equally likely; that is, the prior probabilities of the signals are equal to $1/M$ for each channel.

Let λ_i denote the fraction of time during which channel i is employed for transmission, which is called the *channel switching factor* for channel i. The channel switching factors satisfy $\sum_{i=1}^K \lambda_i = 1$ and $\lambda_i \geq 0\ \forall\, i \in \{1\dots K\}$. In practice, channel switching is performed by utilizing the ith channel for $100\lambda_i$ percent of time for $i = 1,\dots,K$. The aim of this study is to jointly optimize the channel switching factors and signal powers to minimize the average probability of error (symbol error rate) under average power and cost constraints. The error probability of channel i for a transmit power of P is represented by $g_i(P)$, and the following assumptions are employed.

Assumption 1: The error probability over channel i, denoted by $g_i(P)$, is a convex and a monotone decreasing function of P.

Assumption 2: For $C_i > C_j$ (equivalently, for $\sigma_i^2 < \sigma_j^2$), $g_i(P) < g_j(P), \forall\, P > 0$.

As studied in [3], for maximum likelihood (ML) detection over additive Gaussian channels, the error probability is a convex function of the signal power for all 1-dimensional and 2-dimensional constellations (such as BPSK, PAM, QPSK, and QAM, which are commonly employed in practice), and it is convex also for higher dimensional constellations at high SNRs. In addition, for Gaussian channels, the error probability is a monotone decreasing function of the signal power and a monotone increasing function of the noise power. Therefore, Assumption 1 and Assumption 2 are applicable in practical scenarios.

When power P_i is allocated to channel i, the average probability of error is expressed as $\sum_{i=1}^K \lambda_i\, g_i(P_i)$, where λ_i's are the channel switching factors. In practical systems, there exist an average power constraint and a peak power constraint, which can be expressed as $\sum_{i=1}^K \lambda_i\, P_i \leq A_{\rm p}$ and $P_i \in [0, P_{\max}]$, where $A_{\rm p}$ and $P_{\max}$ represent the average power limit and the peak power limit, respectively. It is assumed that $P_{\max} > A_{\rm p} > 0$. In addition, an average transmission cost constraint can be stated as $\sum_{i=1}^K \lambda_i\, C_i \leq A_{\rm c}$, where $A_{\rm c}$ denotes the average cost limit (budget). Then, the proposed optimization problem can be expressed as

$$\begin{aligned} &\min_{\{\lambda_i, P_i\}_{i=1}^K} \sum_{i=1}^K \lambda_i\, g_i(P_i) \\ &\text{subject to} \sum_{i=1}^K \lambda_i\, P_i \leq A_{\rm p}, \quad \sum_{i=1}^K \lambda_i\, C_i \leq A_{\rm c}, \\ &\qquad\qquad \sum_{i=1}^K \lambda_i = 1,\ \lambda_i \geq 0,\ P_i \in [0, P_{\max}],\ \forall\, i \in \{1\dots K\} \end{aligned} \quad (2)$$

In other words, the aim is to obtain the optimal channel switching strategy that minimizes the average probability of error under constraints on the average power, average cost, and peak power.

In the remainder of the manuscript, it is assumed that the noise variances σ_i^2's of the channels are distinct without loss of generality. This is mainly because of the fact that if there are multiple channels with the same noise variances, it is always better to employ only one of them due to the convexity of the error probability.[1] Hence, the problem formulation that considers only the channels with distinct noise variances is sufficient to achieve the overall optimal solution.

In the proposed problem formulation in (2), stochastic signaling [12], [13] is not considered and the power levels, P_i's, are modeled as deterministic quantities for each channel. This is mainly due to the convexity of the error probability $g_i(P)$ with respect to P, which implies that stochastic signaling (i.e., time sharing among different power levels) over a given channel increases the error probability under an average power constraint. For example, instead of performing stochastic signaling over channel i via time sharing between power levels $P_{i,1}$ and $P_{i,2}$ and time sharing factors λ_i and $(1-\lambda_i)$, respectively, performing deterministic signaling with power $\lambda_i P_{i,1} + (1-\lambda_i)P_{i,2}$ yields a lower error probability since $\lambda_i g_i(P_{i,1}) + (1-\lambda_i) g_i(P_{i,2}) > g_i(\lambda_i P_{i,1} + (1-\lambda_i)P_{i,2})$. Based on this argument for each channel, it is concluded that the proposed formulation

[1]To verify this statement, let channel i and channel j have the same noise variances specified by $\sigma_i^2 = \sigma_j^2 = \sigma^2$, and let $g(P)$ denote their error probability expression. Then, it can be shown that instead of employing channel i and channel j with powers P_i and P_j and channel switching (time sharing) factors of λ_i and λ_j, respectively, it is always better to employ only one of these channels with power $(\lambda_i P_i + \lambda_j P_j)/(\lambda_i + \lambda_j)$ and a time sharing factor of $(\lambda_i + \lambda_j)$. This is because of the convexity of $g(P)$ for $P > 0$, which implies that $\lambda_i g(P_i) + \lambda_j g(P_j) > (\lambda_i + \lambda_j) g((\lambda_i P_i + \lambda_j P_j)/(\lambda_i + \lambda_j))$ for all $\lambda_i, \lambda_j \in (0,1)$ and $P_i, P_j > 0$.

in (2) covers the scenarios in the presence of stochastic signaling as well since the joint optimization of channel switching and stochastic signaling results in channel switching with deterministic signaling in the considered scenario.

Finally, it is worth mentioning that the results in this study can also be applied to multipath channels with block *frequency-flat* fading under the assumption of perfect channel estimation at the receiver. In that case, the proposed channel switching approach can be employed for each fading state.

III. Optimal Channel Switching

In this section, a detailed theoretical investigation of the optimal channel switching problem in (2) is presented. In the following analysis, it is assumed without loss of generality that the noise variances of the channels satisfy $\sigma_1^2 < \sigma_2^2 < \cdots < \sigma_K^2$, which implies that the cost values are ordered as $C_1 > C_2 > \cdots > C_K$. In addition, the average cost limit A_c in (2) is assumed to be larger than or equal to the minimum of the cost values; i.e., $A_\mathrm{c} \geq C_K$, since (2) would yield no solution otherwise. Then, the following remark, which specifies two simple cases, is presented first.

Remark 1: (**i**) If $A_\mathrm{c} = C_K$, the optimal solution of (2) is to transmit over channel K exclusively with power A_p.

(**ii**) If $A_\mathrm{c} \geq C_1$, the optimal solution of (2) is to transmit over channel 1 exclusively with power A_p.

Proof: The first part is obvious since the use of another channel apart from channel K would violate the average cost constraint in (2) as $C_1 > C_2 > \cdots > C_K = A_\mathrm{c}$. Also, since $g_i(P)$ is a monotone decreasing function of P, the optimal strategy operates at the average power limit A_p in (2).

To prove the second part, consider a generic strategy that employs channel switching with power P_i and channel switching factor λ_i for channel i, which achieves an average probability of error given by $\sum_{i=1}^K \lambda_i\, g_i(P_i)$. Then, the following inequalities can be obtained:

$$\sum_{i=1}^{K} \lambda_i\, g_i(P_i) > \sum_{i=1}^{K} \lambda_i\, g_1(P_i) > g_1\left(\sum_{i=1}^{K} \lambda_i\, P_i\right) \quad (3)$$

The first inequality follows from the fact that $g_1(P) < g_i(P)$ for all $P \geq 0$ and $i \in \{2, \ldots, K\}$ since channel 1 has the smallest noise variance (the largest cost). On the other hand, the second inequality follows from the strict convexity of g_1 for positive arguments. It is noted that the expression on the right-hand-side of (3) is the probability of error that is achieved by employing channel 1 exclusively with power $\sum_{i=1}^K \lambda_i\, P_i$. Therefore, it is concluded from (3) that employing channel 1 exclusively always achieves a smaller average probability of error than any strategy that employs channel switching. (Since $A_\mathrm{c} \geq C_1$, it is possible to employ channel 1 exclusively.) In addition, since $g_1(P)$ is a monotone decreasing function of P, the optimal strategy operates at the average power limit A_p. ■

Remark 1 presents intuitive results for two simple cases, which can be summarized as follows: If the budget (average cost limit) allows the use of the worst (cheapest) channel only, then the only feasible approach is to employ that channel exclusively, which becomes the optimal solution of (2). On the other hand, if the budget allows the use of any channel with any switching factors, then the optimal solution is to employ the best channel all the time by using all the available power; that is, channel switching can only degrade the performance in this scenario. Since the solutions in these two special cases are obtained in a simple manner, we focus on the other scenarios for which the average cost limit satisfies $C_K < A_\mathrm{c} < C_1$ in the remainder of this study.

Instead of trying to solve the problem in (2) directly for obtaining the optimal channel switching strategy, the properties of the optimal solution are investigated first to propose alternative approaches that yield the optimal channel switching strategy with reduced computational complexity. To that aim, the following proposition states that the optimal solution of (2) always satisfies the average power and average cost constraints with equality.

Proposition 1: Assume that $C_K < A_\mathrm{c} < C_1$ and let $\{\lambda_i^*, P_i^*\}_{i=1}^K$ denote the solution of the optimization problem in (2). Then, $\sum_{i=1}^K \lambda_i^* P_i^* = A_\mathrm{p}$ and $\sum_{i=1}^K \lambda_i^* C_i = A_\mathrm{c}$; that is, the optimal channel switching strategy utilizes the maximum average power and the maximum average cost.

Proof: The claims in the proposition can be proved via contradiction. To prove the claim about the utilization of the maximum average power, suppose that $\{\lambda_l, P_l\}_{l=1}^K$ is an optimal solution of (2) with $\sum_{l=1}^K \lambda_l P_l < A_\mathrm{p}$ and channel i is one of the employed channels (i.e., $\lambda_i > 0$) with $P_i < P_{\max}$. Note that such a channel must exist since $A_\mathrm{p} < P_{\max}$. Next, define another solution as $\{\lambda_l, P_l'\}_{l=1}^K$, where $P_l' = P_l, \forall\, l \neq i$ and $P_i' = \min\{P_i + (A_\mathrm{p} - \sum_{l=1}^K \lambda_l P_l)/\lambda_i, P_{\max}\}$. It is noted that $P_i' > P_i$. Then, the following relation can be derived:

$$\begin{aligned}\sum_{l=1}^{K} \lambda_l\, g_l(P_l) &= \lambda_i\, g_i(P_i) + \sum_{\substack{l=1\\ l\neq i}}^{K} \lambda_l\, g_l(P_l) \\ &> \lambda_i\, g_i\left(P_i'\right) + \sum_{\substack{l=1\\ l\neq i}}^{K} \lambda_l\, g_l\left(P_l'\right) = \sum_{l=1}^{K} \lambda_l\, g_l\left(P_l'\right) \quad (4)\end{aligned}$$

where the inequality is obtained due to the facts that g_i is monotone decreasing, $P_l' = P_l, \forall\, l \neq i$, and $P_i' > P_i$. From (4), it is concluded that the solution $\{\lambda_l, P_l\}_{l=1}^K$, which operates at an average power below A_p, has a higher average probability of error than $\{\lambda_l, P_l'\}_{l=1}^K$. This leads to a contradiction since $\{\lambda_l, P_l\}_{l=1}^K$ was assumed to be an optimal solution of (2). Therefore, a solution that operates at an average power below A_p cannot be optimal. In other words, an optimal solution must utilize all the available power; i.e., operate at the average power limit, A_p.

To prove the claim about the operation at the maximum average cost, first suppose that the optimal solution employs at least two different channels, say channel i and channel j with powers P_i and P_j and channel switching factors λ_i and λ_j, respectively, where $i < j$ (hence, $C_i > C_j$), and it operates at an average cost of A_c', which is strictly less than A_c; that is, $A_\mathrm{c}' < A_\mathrm{c}$. For notational convenience, define $P_{ij} \triangleq \lambda_i P_i + \lambda_j P_j$ and $C_\lambda \triangleq \lambda_i C_i + \lambda_j C_j$. Then, consider an alternative solution which employs a similar strategy to the optimal solution except that it uses channel i with power P_i' and channel switching factor γ, and channel j with power P_j and channel switching factor $\lambda_i+\lambda_j-\gamma$, where $\lambda_i<\gamma<\lambda_i+\lambda_j$ with $\gamma P_i'+(\lambda_i+\lambda_j-\gamma)P_j=P_{ij}$ (the same average power as the optimal one) and $\gamma C_i+(\lambda_i+\lambda_j-\gamma)C_j=C_\lambda+A_{\rm c}-A_{\rm c}'$ (larger average cost than the optimal one). By equating the average power terms (that is, $\gamma P_i'+(\lambda_i+\lambda_j-\gamma)P_j=\lambda_iP_i+\lambda_jP_j$), P_i' can be obtained as $P_i'=\lambda_iP_i/\gamma+(1-\lambda_i/\gamma)P_j$. Then, the following relations can be obtained:

$$\begin{aligned}
&\gamma\, g_i(P_i')+(\lambda_i+\lambda_j-\gamma)g_j(P_j)\\
&=\gamma\, g_i\left(\lambda_iP_i/\gamma+(1-\lambda_i/\gamma)P_j\right)+(\lambda_i+\lambda_j-\gamma)g_j(P_j) \quad (5)\\
&<\lambda_i\, g_i(P_i)+(\gamma-\lambda_i)\, g_i(P_j)+(\lambda_i+\lambda_j-\gamma)g_j(P_j) \quad (6)\\
&\le\lambda_i\, g_i(P_i)+\lambda_j\, g_j(P_j) \quad (7)
\end{aligned}$$

where the first inequality is obtained from the strict convexity of g_i and the second inequality follows from the fact that $g_i(P_j)\le g_j(P_j)$ since channel i has a smaller noise variance (higher cost) than channel j. The inequality in (5)–(7), namely, $\lambda_i\, g_i(P_i)+\lambda_j\, g_j(P_j)>\gamma\, g_i(P_i')+(\lambda_i+\lambda_j-\gamma)g_j(P_j)$, leads to a contradiction since the optimal solution results in a higher average probability of error than the alternative solution, which uses the same average power but operates at the maximum average cost. Therefore, it is concluded that a solution that employs at least two channels and operates below the average cost limit $A_{\rm c}$ cannot be optimal. To complete the proof, suppose that an optimal solution employs a single channel (say, channel i) and operates below $A_{\rm c}$; that is, channel i is employed exclusively with power P_i and its cost C_i is strictly smaller than $A_{\rm c}$; that is, $C_i<A_{\rm c}<C_1$. Next, consider an alternative solution that employs channel i and channel 1 with channel switching factors λ_i' and $1-\lambda_i'$, respectively, and with the same power P_i, where $\lambda_i'\in(0,1)$. Then,

$$\begin{aligned}
&\lambda_i'g_i(P_i)+(1-\lambda_i')\,g_1(P_i)\\
&\qquad<\lambda_i'g_i(P_i)+(1-\lambda_i')\,g_i(P_i)=g_i(P_i) \quad (8)
\end{aligned}$$

where the inequality follows from the fact that $g_1(P)<g_i(P)$, $\forall\, P$, by definition (note that $C_1>C_i$). The inequality in (8) leads to a contradiction since the alternative solution achieves a smaller average probability of error than the optimal one by using the same average power. Therefore, a solution that employs a single channel and operates below the maximum average cost cannot be optimal. Overall, since any channel switching strategy either uses a single channel or switches among multiple channels, the previous arguments prove that an optimal channel switching strategy must always operate at the maximum average cost. ■

Proposition 1 states that the optimal channel switching strategy utilizes all the average power and average cost. Therefore, the optimization problem in (2) can be solved by considering equality constraints (instead of inequality constraints) for the average power and average cost, which leads to an important reduction in computational complexity. Another implication of Proposition 1 is presented in the following corollary.

Corollary 1: Assume that $C_K<A_{\rm c}<C_1$. If $C_i\ne A_{\rm c}, \forall\, i\in\{1,\dots,K\}$, then the optimal solution of (2) involves channel switching among multiple channels; that is, transmission over a single channel is not optimal.

Proof: Let $\{\lambda_i^*,P_i^*\}_{i=1}^K$ denote the solution of the optimization problem in (2). Proposition 1 states that $\sum_{i=1}^K\lambda_i^*C_i=A_{\rm c}$ must hold. If $C_i\ne A_{\rm c}, \forall\, i\in\{1,\dots,K\}$, then the condition of $\sum_{i=1}^K\lambda_i^*C_i=A_{\rm c}$ cannot be satisfied unless at least two of λ_i^*'s are nonzero, which implies switching among multiple channels. ■

It should be noted that the converse of Corollary 1 is not necessarily true. That is, when $C_i=A_{\rm c}$ for some $i\in\{1,\dots,K\}$, the structure of the optimal solution depends on the cost values and the average power constraint. In other words, either transmission over a single channel or channel switching can be optimal depending on the system parameters.

Although the optimization problem in (2) is formulated to search over strategies that involve channel switching among up to K channels, a similar approach to those in [8], [12], [13] can be employed to restrict the optimal solution to a smaller subset of strategies. Namely, the following proposition states that the optimal solution of (2) can be expressed as channel switching among $\min\{K,3\}$ or fewer channels.

Proposition 2: The optimal channel switching strategy is to switch among at most $\min\{K,3\}$ channels.

Proof: If $K\le3$, the statement in the proposition is satisfied trivially. Assume that $K>3$ and define the following sets:

$$\mathcal{W}=\left\{\left(\sum_{i=1}^K\lambda_i\,P_i,\ \sum_{i=1}^K\lambda_i\,g_i(P_i),\sum_{i=1}^K\lambda_i\,C_i\right),\ \forall\,\lambda_i\ge0,\ \sum_{i=1}^K\lambda_i=1,\ \forall\,P_i\in[0,P_{\max}]\right\} \quad (9)$$

$$\mathcal{U}=\{(P,g_i(P),C_i)\,,\ \forall\, i\in\{1,\dots,K\},\ \forall\, P\in[0,P_{\max}]\} \quad (10)$$

It is noted that set U is the set of all triples $(P,g_i(P),C_i)$, for $i\in\{1,\dots,K\}$ and $P\in[0,P_{\max}]$, which consists of infinitely many elements. Also, by definition, set $\mathcal{W}$ contains the optimal solution of (2) since it consists of all possible average power, average probability of error and average cost triples. In addition, it is observed from (9) and (10) that $\mathcal{W}$ is a subset of the convex hull of set $\mathcal{U}$; i.e., $\mathcal{W}\subset{\rm hull}(\mathcal{U})$. This is because of the fact that all the triples in $\mathcal{W}$ can be obtained as the convex combinations of K elements in $\mathcal{U}$ whereas some convex combinations of the elements of $\mathcal{U}$, which involve the use of at least one channel multiple times,[2] are not included in $\mathcal{W}$. Since $\mathcal{W}$ is contained in the convex hull of set $\mathcal{U}$, any element of $\mathcal{W}$ can be expressed as a convex combination of $\dim(\mathcal{U})+1=4$ elements in $\mathcal{U}$ as a result of Carathéodory's theorem [27], where $\dim(\mathcal{U})$ denotes the dimension of the space in which $\mathcal{U}$ resides. (Note that $\mathcal{U}\subset\mathbb{R}^3$.) In addition, since the aim is to achieve the minimum average probability of error (see (2)), the optimal solution

[2]For example, the convex combination of $(P_1,g_1(P_1),C_1)$ and $(P_2,g_1(P_2),C_1)$ is not included in $\mathcal{W}$, which involves the use of channel 1 twice.

corresponds to a point on the boundary of $\mathrm{hull}(\mathcal{U})$, which can be achieved by a convex combination of $\dim(\mathcal{U}) = 3$ elements in $\mathcal{U}$ by Carathéodory's theorem [27]. Finally, it is noted that all such convex combinations are guaranteed to be elements of set $\mathcal{W}$ due to the following reason: The difference of $\mathrm{hull}(\mathcal{U})$ from $\mathcal{W}$ (that is, $\mathrm{hull}(\mathcal{U}) \setminus \mathcal{W}$) consists of the points corresponding to strategies that use at least one channel multiple times. However, such strategies cannot be optimal solutions since the use of a channel multiple times always increases the average probability of error compared to the use of that channel once with the same average power (which can be proved by an argument similar to that in Footnote 1). Therefore, the optimal solution cannot be in $\mathrm{hull}(\mathcal{U}) \setminus \mathcal{W}$; i.e., it is always in $\mathcal{W}$, which implies that the optimal solution can be expressed as a convex combination of up to 3 elements in $\mathcal{U}$ that correspond to different channel indices (see index i in (10)). Hence, channel switching among up to 3 different channels is optimal. ■

Based on Proposition 1 and Proposition 2, the optimal channel switching corresponds to one of the following three strategies:

Strategy 1—Transmission Over a Single Channel: In this case, one of the channels is employed exclusively. Based on Corollary 1, this strategy cannot be an optimal solution of (2) unless there exists a channel with cost A_{c}. If there exists such a channel and i^* denotes the index of that channel (that is, $C_{i^*} = A_{\mathrm{c}}$), then the minimum average probability of error achieved by this strategy is given by $g_{i^*}(A_{\mathrm{p}})$, which corresponds to transmission over channel i^* exclusively by utilizing the maximum available power (cf. (2)). Note that this strategy may or may not be the optimal solution of the problem in (2) depending on the system parameters.

Strategy 2—Channel Switching Between Two Channels: In this strategy, channel switching is performed between two different channels. Let channel i and channel j denote those channels. Then, based on Proposition 1, the problem in (2) can be formulated under Strategy 2 as

$$\begin{aligned} \min_{\lambda,\, P_i,\, P_j} \; & \lambda\, g_i(P_i) + (1-\lambda)\, g_j(P_j) \\ \text{subject to } & \lambda\, P_i + (1-\lambda) P_j = A_{\mathrm{p}}, \\ & \lambda\, C_i + (1-\lambda)\, C_j = A_{\mathrm{c}},\ \lambda \in [0,1], \\ & P_i \in [0, P_{\max}],\ P_j \in [0, P_{\max}]. \end{aligned} \tag{11}$$

It is observed from the average cost constraint in (11) that, for the optimal channel switching between two channels, one of the channels should have a cost higher than A_{c} and the other channel should have a cost lower than A_{c}. Therefore, to obtain the optimal solution for Strategy 2, the problem in (11) should be solved for $K_{\mathrm{s}}K_{\mathrm{g}}$ channel pairs, where K_{s} (K_{g}) is the number of channels the costs of which are lower (higher) than A_{c}. In other words, the problem in (11) should be solved for all channel pairs $(i,j) \in \mathcal{S}_2$, where $\mathcal{S}_2 = \{(i,j) : C_i > A_{\mathrm{c}} > C_j \text{ and } i,j \in \{1,\ldots,K\}\}$.

To investigate the properties of the solution of (11), a specific expression is considered for the error probability over each channel.

Assumption 3: The error probability for channel i is expressed as

$$g_i(P) = \eta\, Q\left(\frac{\kappa\sqrt{P}}{\sigma_i}\right) \tag{12}$$

where Q denotes the Q-function, P is the average symbol energy, σ_i^2 is the noise variance, and η and κ are some positive constants that depend on the modulation type and order [28].

As discussed in [28], the error probability for many coherent modulation schemes can be represented either exactly or approximately (at high SNRs) in the form of (12); hence, Assumption 3 provides a generic expression that can represent various different scenarios.

Assume, without loss of generality, that $C_i > A_{\mathrm{c}} > C_j$ for the problem in (11). Then, the optimal value of λ can be obtained from the average cost constraint as $\lambda^* = (A_{\mathrm{c}} - C_j)/(C_i - C_j)$. Also, suppose that $P_{\max}$ is sufficiently large so that the optimal power levels for Strategy 2 are always below $P_{\max}$. (The condition for this assumption is specified in Proposition 3 below.) Then, due to the average power constraint, the powers are related as $P_j = (A_{\mathrm{p}} - \lambda^* P_i)/(1-\lambda^*)$. From (12), the optimization problem in (11) can then be expressed as follows:

$$\begin{aligned} \min_{P_i \in (0,\ A_{\mathrm{p}}/\lambda^*)} \; & \lambda^* \eta\, Q\left(\frac{\kappa\sqrt{P_i}}{\sigma_i}\right) \\ & + (1-\lambda^*)\eta\, Q\left(\frac{\kappa}{\sigma_j}\sqrt{\frac{A_{\mathrm{p}} - \lambda^* P_i}{1-\lambda^*}}\right) \end{aligned} \tag{13}$$

where the constraint for P_i is obtained from the relation $\lambda^* P_i + (1-\lambda^*) P_j = A_{\mathrm{p}}$. From (13), it is observed that the optimal solution for Strategy 2 requires a search over a one-dimensional space only (for each possible channel pair). In addition, it can be shown that the objective function in (13) is strictly convex for $P_i \in (0,\ A_{\mathrm{p}}/\lambda^*)$.[3] Therefore, convex optimization algorithms can be employed to obtain the result in polynomial time [29]. In fact, as stated in the following proposition, the structure of the objective function also leads to additional properties, which result in further simplifications.

Proposition 3: Suppose that $C_i > A_{\mathrm{c}} > C_j$, $P_{\max} > A_{\mathrm{p}}\sigma_j^2/\sigma_i^2$, and define $A_{ij} \triangleq \frac{\sigma_i^2\sigma_j^2}{\kappa^2(\sigma_j^2 - \sigma_i^2)} \log\left(\frac{\sigma_j^2}{\sigma_i^2}\right)$, where log denotes the natural logarithm. Then, the optimal solution of (11), denoted by $\{\lambda^*,\ P_i^*,\ P_j^*\}$, satisfies the following relations depending on the average power limit:

(i) If $A_{\mathrm{p}} = A_{ij}$, then $P_i^* = P_j^* = A_{ij}$.
(ii) If $A_{\mathrm{p}} > A_{ij}$, then $P_j^* > A_{\mathrm{p}} > P_i^* > A_{ij}$.
(iii) If $A_{\mathrm{p}} < A_{ij}$, then $A_{ij} > P_i^* > A_{\mathrm{p}} > P_j^*$.

In addition, the ratio between the optimal power levels cannot exceed σ_j^2/σ_i^2; that is,

$$\max\left\{\frac{P_j^*}{P_i^*}, \frac{P_i^*}{P_j^*}\right\} < \frac{\sigma_j^2}{\sigma_i^2}. \tag{14}$$

[3] The first-order derivative of the objective function is presented in (25), which is a monotone increasing function of P_i for $P_i \in (0,\ A_{\mathrm{p}}/\lambda^*)$.

Proof: Please see Appendix A.

Under the conditions in Proposition 3, the search space for the optimization problem in (13) can be reduced. Specifically, for each channel pair (i, j) with $C_i > C_j$, the value of A_{ij} is calculated first, as defined in the proposition. Then, the optimal power levels are obtained as follows:

- If $A_\mathrm{p} = A_{ij}$, the optimal solution is given by $P_i^* = P_j^* = A_{ij}$.
- If $A_\mathrm{p} > A_{ij}$, the optimization problem in (13) is solved for $P_i \in (\max\{A_{ij}, \sigma_i^2 A_\mathrm{p}/\sigma_j^2\}, A_\mathrm{p})$, which is obtained from (14) and the relation in the second part of the proposition.
- If $A_\mathrm{p} < A_{ij}$, the problem in (13) is solved for $P_i \in (A_\mathrm{p}, \min\{A_{ij}, A_\mathrm{p}/\lambda^*, \sigma_j^2 A_\mathrm{p}/\sigma_i^2\})$, which is obtained from (14) and the relation in the third part of the proposition.

Once the optimal value of P_i, denoted by P_i^*, is obtained, the optimal value of P_j is calculated as $P_j^* = (A_\mathrm{p} - \lambda^* P_i^*)/(1 - \lambda^*)$, where $\lambda^* = (A_\mathrm{c} - C_j)/(C_i - C_j)$.

Strategy 3—Channel Switching Among Three Channels: In this strategy, channel switching is performed among three different channels. Let channel i, channel j, and channel k denote those channels. Then, based on Proposition 1, the problem in (2) can be formulated under Strategy 3 as

$$\begin{aligned} \min_{\lambda_i,\lambda_j,\lambda_k,P_i,P_j,P_k} \quad & \lambda_i\, g_i(P_i) + \lambda_j\, g_j(P_j) + \lambda_k\, g_k(P_k) \\ \text{subject to } & \lambda_i\, P_i + \lambda_j\, P_j + \lambda_k\, P_k = A_\mathrm{p}, \\ & \lambda_i\, C_i + \lambda_j\, C_j + \lambda_k\, C_k = A_\mathrm{c}, \\ & \lambda_i + \lambda_j + \lambda_k = 1,\ \lambda_i, \lambda_j, \lambda_k \geq 0, \\ & P_i, P_j, P_k \in [0, P_{\max}]. \end{aligned} \tag{15}$$

Due to the strict average cost constraint, it is required that at least one of the channels must have a cost lower than A_c and at least one of the channels must have a cost higher than A_c. Therefore, to obtain the optimal solution for Strategy 3, the problem in (15) should be solved for $K_\mathrm{s} K_\mathrm{g}(K - 2)$ channel triples, where K_s (K_g) is the number of channels the costs of which are lower (higher) than A_c, and K is the total number of channels. In other words, the problem in (15) should be solved for all channel triples $(i, j, k) \in \mathcal{S}_3$, where $\mathcal{S}_3 = \{(i, j, k) : C_i > A_\mathrm{c} > C_j \text{ and } i, j, k \in \{1, \ldots, K\}\}$. In addition, it is observed that the solution of (15) can be obtained via optimization over a *three-dimensional* space instead of six by utilizing the three equality constraints.

It is noted from Proposition 2 that Strategy 3 is guaranteed to provide the optimal solution of the channel switching problem in (2). In addition, it covers Strategy 2 and Strategy 1 as special cases, which may be suboptimal in general. Therefore, to obtain the optimal channel switching solution, it can be necessary in general to solve the optimization problem in (15), which is computationally more complex than obtaining the optimal solutions under Strategy 1 and Strategy 2. However, in some cases (see Proposition 5), it is guaranteed that Strategy 1 or Strategy 2 can provide the optimal solution of the channel switching problem in (2); that is, it is not necessary to solve the optimization problem in (15) for obtaining the optimal channel switching solution. Therefore, whenever the conditions under which Strategy 1 or Strategy 2 is optimal are satisfied, the optimal channel switching solution can be obtained in a low-complexity manner as follows: If there exist no channels with cost A_c, then Strategy 2 provides the optimal solution. If there exists a channel with cost A_c, then the optimal solution is either to employ that channel exclusively with the maximum power (Strategy 1), or to switch between two channels as specified by the solution of (11) (Strategy 2). In that case, the strategy that achieves the smaller average probability of error becomes the optimal solution of (2).

IV. Optimal Channel Switching for Logarithmic Cost Functions

In this section, specific theoretical results are obtained by considering a suitable cost function for the channels. Since each channel can be regarded as a measurement device, a cost function similar to that proposed in [26] can be adopted for relating the noise power of each channel to a cost value as follows:

$$C_i = \log\left(1 + \frac{b}{\sigma_i^2}\right), \quad i \in \{1 \ldots K\}, \tag{16}$$

where $b > 0$ is a given system parameter (a constant). It is noted that the function in (16) has the desirable property that it assigns higher (lower) cost values to less (more) noisy channels; that is, $\sigma_i^2 < \sigma_j^2$ implies $C_i > C_j$. In addition, $\lim_{\sigma_i \to \infty} C_i = 0$ and $\lim_{\sigma_i \to 0} C_i = \infty$. As in the previous section, it is assumed without loss of generality that the noise variances of the channels satisfy $\sigma_1^2 < \sigma_2^2 < \cdots < \sigma_K^2$, which implies that the cost values are ordered as $C_1 > C_2 > \cdots > C_K$. In addition, the error probability expression in (12) is considered.

Based on the cost function in (16), the following result is obtained first.

Lemma 1: Consider infinitely many channels and assume that the channels take a continuum of cost values in the interval $[C_{\min}, C_{\max}]$ based on the cost function in (16), where $0 < C_{\min} < C_{\max} < \infty$. Let $g(P, C)$ denote the error probability when transmission is performed by utilizing a power level of P over a channel with cost C. Then, $g(P, C)$ is a strictly convex function over set $\mathcal{S}_c$, which is a convex set defined as $\mathcal{S}_c \triangleq \{(P, C) : P > b/(\kappa^2(\mathrm{e}^C + 1)),\ C \in (C_{\min}, C_{\max})\}$.

Proof: Please see Appendix B.

Lemma 1 describes the convexity properties of the error probability, which is considered as a function of power and cost. Based on Lemma 1, the solutions of the optimal channel switching problem can be specified in certain scenarios. To that aim, the following proposition presents the optimal solution when channel switching is performed between two channels (i.e., Strategy 2).

Proposition 4: Suppose there exist K channels and each channel has a cost value obtained from the cost function in

(16). If the power limits satisfy $A_{\mathrm{p}} \geq \frac{b\ \sigma_K^4}{\kappa^2\sigma_1^2(2\sigma_K^2+b)}$ and $P_{\max} > A_{\mathrm{p}}\sigma_K^2/\sigma_1^2$, then the optimal solution for Strategy 2 employs channel i and channel j, where

$$i = \underset{k\in\{1,\ldots,K\}}{\arg\min}\ C_k \text{ subject to } C_k > A_{\mathrm{c}}, \tag{17}$$

$$j = \underset{k\in\{1,\ldots,K\}}{\arg\max}\ C_k \text{ subject to } C_k < A_{\mathrm{c}}. \tag{18}$$

Proof: From Proposition 1, it is known that the optimal channel switching solution utilizes the maximum average cost. Therefore, for Strategy 2, the optimal pair of channels, say (k, l), must satisfy $k \leq i$ and $l \geq j$, where i and j are as defined in (17) and (18), respectively. (Note that the cost values are ordered as $C_1 > C_2 > \cdots > C_K$.) For simplicity of notation, define $\mathbf{z}_k \triangleq (P_k, C_k)$, for $k = 1, \ldots, K$, and $\boldsymbol{a} \triangleq (A_{\mathrm{p}}, A_{\mathrm{c}})$. To prove that the optimal channel pair for Strategy 2 is (i, j), first consider channel pair (k, l), where $k = i$ and $l > j$. The optimal solution for channel pair (i, l) must utilize the maximum average power and cost due to Proposition 1. In addition, consider an alternative solution that employs channel pair (i, j) and operates at the average power and cost limits. Then, the following inequalities are obtained:

$$\lambda\, \mathbf{z}_i + (1-\lambda)\, \mathbf{z}_j = \boldsymbol{a} \text{ and } \gamma\, \mathbf{z}_i + (1-\gamma)\, \mathbf{z}_l = \boldsymbol{a}, \tag{19}$$

where $\lambda = (A_{\mathrm{c}} - C_j)/(C_i - C_j)$ and $\gamma = (A_{\mathrm{c}} - C_l)/(C_i - C_l)$, which are obtained from the average cost constraint. Since $C_i > A_{\mathrm{c}} > C_j > C_l$, it can be shown that $\gamma > \lambda$. Therefore, $\mathbf{z}_j$ can be expressed as

$$\mathbf{z}_j = \frac{\gamma-\lambda}{1-\lambda}\, \mathbf{z}_i + \frac{1-\gamma}{1-\lambda}\, \mathbf{z}_l. \tag{20}$$

Then, it is shown in the following that channel pair (i, l) cannot be optimal since it results in a higher average probability of error than channel (i, j):

$$\lambda\, g(\mathbf{z}_i) + (1-\lambda) g(\mathbf{z}_j)$$

$$= \lambda\, g(\mathbf{z}_i) + (1-\lambda)\, g\left(\frac{\gamma-\lambda}{1-\lambda}\, \mathbf{z}_i + \frac{1-\gamma}{1-\lambda}\, \mathbf{z}_l\right) \tag{21}$$

$$< \lambda\, g(\mathbf{z}_i) + (1-\lambda)\left(\frac{\gamma-\lambda}{1-\lambda}\, g(\mathbf{z}_i) + \frac{1-\gamma}{1-\lambda}\, g(\mathbf{z}_l)\right) \tag{22}$$

$$= \gamma\, g(\mathbf{z}_i) + (1-\gamma)\, g(\mathbf{z}_l) \tag{23}$$

where $g(\mathbf{z}_i) = g(P_i, C_i)$ denotes the average probability of error as a function of power and cost, as defined in (31). In obtaining the equality in (21), the expression in (20) is employed, and the inequality in (22) follows from the strict convexity of g, which is guaranteed under the conditions in the proposition, namely, $A_{\mathrm{p}} \geq \frac{b\ \sigma_K^4}{\kappa^2\sigma_1^2(2\sigma_K^2+b)}$ and $P_{\max} > A_{\mathrm{p}}\sigma_K^2/\sigma_1^2$. To verify the convexity of g in this scenario, Lemma 1 is considered first, which states that the power levels should satisfy $P > b/(\kappa^2(\mathrm{e}^{C}+1))$ for strict convexity. Since the cost values are ordered as $C_1 > C_2 > \cdots > C_K$ (equivalently, $\sigma_1^2 < \sigma_2^2 < \cdots < \sigma_K^2$), $P > b/(\kappa^2(\mathrm{e}^{C_K}+1))$ is required to guarantee that function g is strictly convex for all channels. From Proposition 3, it is concluded that the optimal power levels under Strategy 2 always satisfy $\min\{P_k, P_l\} > \sigma_k^2 A_{\mathrm{p}}/\sigma_l^2$ for channel pair (k, l) with $C_k > A_{\mathrm{c}} > C_l$.[4] Therefore, if $\sigma_1^2 A_{\mathrm{p}}/\sigma_K^2 > b/(\kappa^2(\mathrm{e}^{C_K}+1))$ holds, then it is guaranteed that the optimal power levels for any channel pair for Strategy 2 satisfy the convexity condition in Lemma 1. Mathematically stated,

$$\min\{P_k, P_l\} > \frac{\sigma_k^2 A_{\mathrm{p}}}{\sigma_l^2} > \frac{\sigma_1^2 A_{\mathrm{p}}}{\sigma_K^2} > \frac{b}{\kappa^2(\mathrm{e}^{C_K}+1)} = \frac{b}{\kappa^2(2+b/\sigma_K^2)} \tag{24}$$

for all $k \leq i$ and $l \geq j$, where first inequality is obtained from Proposition 3, the second one follows from the relation $\sigma_1^2 < \sigma_2^2 < \cdots < \sigma_K^2$, the third one is imposed to guarantee that the power levels satisfy the convexity condition in Lemma 1, and the equality is obtained from (16). From (24), it is deduced that the condition $A_{\mathrm{p}} > b\ \sigma_K^2/(\kappa^2(2\ \sigma_1^2 + b\ \sigma_1^2/\sigma_K^2))$ guarantees the strict convexity of g.

Similar arguments to those in (19)–(23) can be used to prove that channel pair (k, j) with $k < i$ results in a larger average probability of error than channel pair (i, j). Then, it can be concluded that channel pair (k, l) cannot be optimal if $k < i$ and/or $l > j$. Hence, the optimal channel pair for Strategy 2 is shown to be the channel pair (i, j) as defined in the proposition when the power limits are larger than the specified values. ∎

Proposition 4 states that if the average and peak power limits are larger than certain values, then the optimal solution for Strategy 2 is to switch between the two channels, one of which has the lowest cost among the channels with costs higher than A_{c}, and the other has the highest cost among the channels with costs lower than A_{c}. In other words, among all the channel pairs, where each pair has one channel with a cost higher than A_{c} and another channel with a cost lower than A_{c}, the one that has the *minimum cost difference* is selected to achieve the minimum average probability of error, which is mainly due to the convexity of the error probability, as specified in Lemma 1. Thanks to Proposition 4, it is not necessary to search over all feasible channel pairs to obtain the optimal solution for Strategy 2 under the conditions in Proposition 4.

Remark 2: Under the condition in Proposition 4, if there exists a channel with cost A_{c}, then it outperforms the channel pair (i, j) specified in (17) and (18); that is, Strategy 1 outperforms Strategy 2 in that scenario. This is due to the strict convexity of g, which results in $\lambda\ g(P_i, C_i) + (1-\lambda)\ g(P_j, C_j) > g(A_{\mathrm{p}}, A_{\mathrm{c}})$. In other words, if $A_{\mathrm{p}} \geq \frac{b\ \sigma_K^4}{\kappa^2\sigma_1^2(2\sigma_K^2+b)}$ and $P_{\max} > A_{\mathrm{p}}\sigma_K^2/\sigma_1^2$, transmission over a single channel with cost A_{c} at the maximum power level A_{p} achieves a smaller average probability of error than performing optimal channel switching between two channels.

[4]This result is obtained by combining the inequality in (14) with the three possible scenarios in Proposition 3. Note that $P_{\max} > A_{\mathrm{p}}\sigma_K^2/\sigma_1^2$ guarantees that the assumption in Proposition 3 holds for all channel pairs.

Based on Lemma 1, it is also possible to describe scenarios in which Strategy 1 or Strategy 2 is the optimal solution of the channel switching problem; that is, switching among more than two channels is not needed. The following proposition presents such a scenario:

Proposition 5: Consider the optimal channel switching problem in (2) with the cost values as defined in (16), and assume that $P_{\max} \to \infty$. Then, the optimal channel switching strategy involves at most two channels if the average power limit satisfies $A_{\rm p} \geq \frac{2\, b\, \sigma_K^4}{\kappa^2 \sigma_1^2 (2\sigma_K^2 + b)}$.

Proof: Please see Appendix C.

Proposition 5 states that in the absence of peak power constraints, if the average power limit is larger than a certain value, then the optimal channel switching strategy is to use a single channel exclusively or to switch between two channels; that is, Strategy 3 is not optimal. In such a scenario, the optimal solution is either to transmit over a single channel with cost $A_{\rm c}$ if such a channel exists, or to switch between channel i and channel j as specified in Proposition 4 if there exists no channels with cost $A_{\rm c}$.

Remark 3: Based on the results in Section III and Section IV, the following algorithm can be described for obtaining the optimal channel switching solution:

- If $A_{\rm c} = C_K$, the optimal channel switching strategy is to transmit over channel K exclusively with power $A_{\rm p}$ (see Remark 1-(i)).
- If $A_{\rm c} \geq C_1$, the optimal channel switching strategy is to transmit over channel 1 exclusively with power $A_{\rm p}$ (see Remark 1-(ii)).
- If $C_K < A_{\rm c} < C_1$,
  - — if the cost function is the logarithmic cost function in (16), $A_{\rm p} \geq \frac{2\, b\, \sigma_K^4}{\kappa^2 \sigma_1^2 (2\sigma_K^2 + b)}$, and no peak power constraints exist,
    - ∗ if there exists a channel with cost $A_{\rm c}$, transmission over that channel at the maximum power level $A_{\rm p}$ is the optimal strategy (see Proposition 5 and Remark 2).
    - ∗ otherwise, the optimal strategy is to perform time sharing between channel i and channel j specified in (17) and (18) (see Proposition 4), and the optimal solution can be obtained based on (11).
  - — otherwise, the optimal channel switching strategy is obtained based on the optimization problem in (15).

V. Numerical Examples

In this section, various numerical examples are presented to provide illustrations of the theoretical results and to investigate performance gains that can be achieved via channel switching. The following strategies are compared in the numerical examples:

Optimal Single Channel: In this strategy, channel switching is not allowed, and only one channel is employed exclusively. The optimal solution for this approach is obtained by using Strategy 1 in Section III.

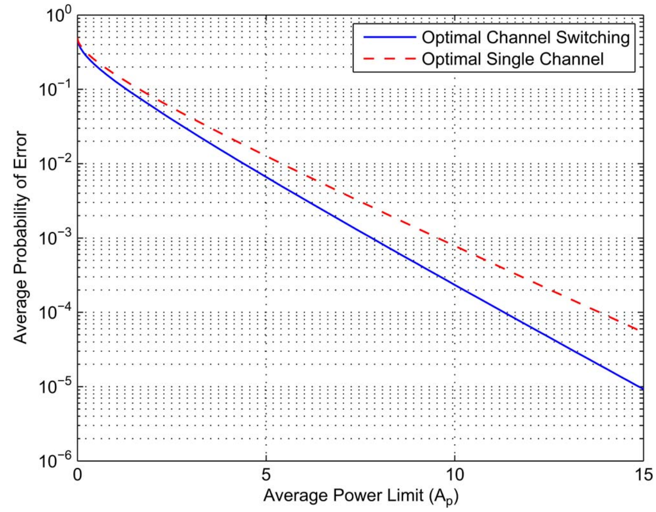

Fig. 2. Average probability of error versus $A_{\rm p}$ for the optimal single channel and optimal channel switching strategies, where $K = 4$, $\boldsymbol{\sigma} = [0.4\ 0.6\ 0.8\ 1]$, $\boldsymbol{C} = [7\ 5\ 3\ 1]$, and $A_{\rm c} = 2$.

Optimal Channel Switching: In this strategy, channel switching is allowed, and the optimal solution of the channel switching problem in (2) is obtained based on Strategy 3. (Since Strategy 3 covers Strategy 2 as a special case, Strategy 2 is not considered separately.)

A scenario with K Gaussian channels is considered, and the standard deviations and the costs of the channels are represented, for notational simplicity, in the vector form as $\boldsymbol{\sigma} = [\sigma_1 \cdots \sigma_K]$ and $\boldsymbol{C} = [C_1 \cdots C_K]$, respectively. For all the examples, the peak power limit in (2) is set to $P_{\max} = 10 A_{\rm p}$, where $A_{\rm p}$ is the average power limit. In addition, binary antipodal signaling is considered, which corresponds to an error probability expression as in (12) with $\eta = \kappa = 1$. First, a four-channel system is studied, where $\boldsymbol{\sigma} = [0.4\ 0.6\ 0.8\ 1]$, $\boldsymbol{C} = [7\ 5\ 3\ 1]$, and the average cost limit is equal to 2; that is, $A_{\rm c} = 2$. In Fig. 2, the average probabilities of error are plotted versus the average power limit $A_{\rm p}$ for the optimal single channel and optimal channel switching approaches. It is observed that the optimal channel switching strategy outperforms the optimal single channel strategy for *all* values of $A_{\rm p}$. This is an expected result since the optimal single channel approach cannot be the optimal solution of the channel switching problem in this scenario as there exists no channel with a cost of $A_{\rm c}$ and $A_{\rm c} < C_1$ (Corollary 1). To provide further investigations of the results in Fig. 2, the parameters of the optimal channel switching strategy are presented in Table I for some values of $A_{\rm p}$. In the table, the optimal channel switching solution is represented by channel switching factors $(\lambda_i, \lambda_j, \lambda_k)$ and power levels (P_i, P_j, P_k), where $i < j < k$. The channels that are not employed in the optimal solution are marked with "-" in the table. Since at most three channels can be utilized in the optimal solution according to Proposition 2, only two of the channel switching factors are shown in the table, and the remaining one can be calculated as $\lambda_k = 1 - \lambda_i - \lambda_j$. It should be noted that λ_i, λ_j, and λ_k correspond to the channel switching factors of the

TABLE I
PARAMETERS OF THE OPTIMAL CHANNEL SWITCHING STRATEGY IN FIG. 2

A_p	λ_i	λ_j	P_1	P_2	P_3	P_4
0.1	0.1667	0.8333	0.1821	–	–	0.0836
0.2	0.1667	0.8333	0.2663	–	–	0.1867
2	0.25	0.75	–	1.3461	–	2.2180
4	0.5	0.5	–	–	3.4117	4.5883
6	0.5	0.5	–	–	4.9898	7.0102
8	0.5	0.5	–	–	6.5601	9.4399
10	0.5	0.5	–	–	8.1270	11.873

employed channels with the smallest index, the second smallest index, and the third smallest index, respectively. For example, for $A_p = 0.1$, channel 1 is employed with channel switching factor 0.1667 and power 0.1821 and channel 4 is employed with channel switching factor 0.8333 and power 0.0836. (In this case, $\lambda_k = 0$, meaning that only two channels are employed in the optimal solution). It is observed from Table I that the optimal channel switching strategy performs channel switching between two channels, which in compliance with Proposition 2. In addition, the calculations show that the optimal channel switching solution utilizes the maximum average power and maximum average cost as claimed in Proposition 1. In addition, the statements in Proposition 3 are verified, which can be exemplified as follows: Parameter A_{ij} in Proposition 3 can be calculated for channel 3 and channel 4 as $A_{34} = 0.7934$. As observed from Table I, when channel 3 and channel 4 are employed, $A_p > A_{34}$ is satisfied and the conditions in Part (ii) of Proposition 3 hold; that is, $P_4 > A_p > P_3 > A_{34}$. In addition, the ratio of the optimal power levels is always smaller than the ratio of the noise variances, $1/(0.8)^2 = 1.5625$, as stated in (14) in Proposition 3. (Note that $P_{\max} = 10A_p > A_p\sigma_4^2/\sigma_3^2 = 1.5625A_p$ is also satisfied.) Compared to the optimal channel switching strategy, which performs channel switching between channel 4 and another channel, the optimal single channel solution always utilizes channel 4 at the maximum power limit A_p since it is the only channel with a cost that is lower than the average cost limit A_c. However, as observed from Fig. 2 and Table I, performing time sharing between channel 4 and a channel with a higher cost (lower error probability) reduces the average probability of error in this scenario.

Next, the same channel configuration is considered with a different average cost limit, which is given by $A_c = 5$, and the average probability of error curves are presented in Fig. 3. In this case, since there is a channel with a cost that is equal to A_c, Corollary 1 does not apply; i.e., channel switching is not necessarily optimal. As observed from the figure, for small values of the average power limit A_p, the optimal channel switching strategy outperforms the optimal single channel strategy (please see Fig. 4 for a zoomed-in version of Fig. 3 for $A_p \in [0, 1]$), whereas both strategies achieve the same performance as A_p increases. Table II presents the parameters of the optimal channel switching solution, which indicates that employing channel 2 exclusively at the power limit (which is the optimal single channel solution) becomes optimal when A_p is larger than a certain value whereas switching between channel 1 and channel 4 is optimal for small values of A_p. Hence, it is concluded that when there exists a channel with a cost equal to A_c, employing a single channel exclusively may or may not be the optimal solution depending on the system parameters.

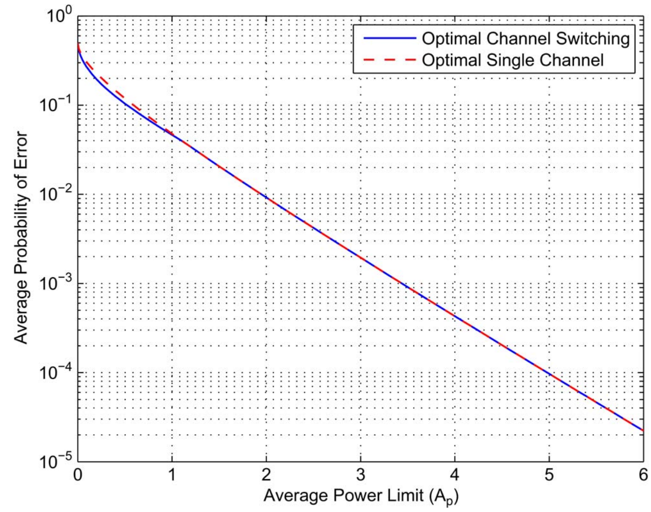

Fig. 3. Average probability of error versus A_p for the optimal single channel and optimal channel switching strategies, where $K = 4$, $\boldsymbol{\sigma} = [0.4\ 0.6\ 0.8\ 1]$, $\boldsymbol{C} = [7\ 5\ 3\ 1]$, and $A_c = 5$.

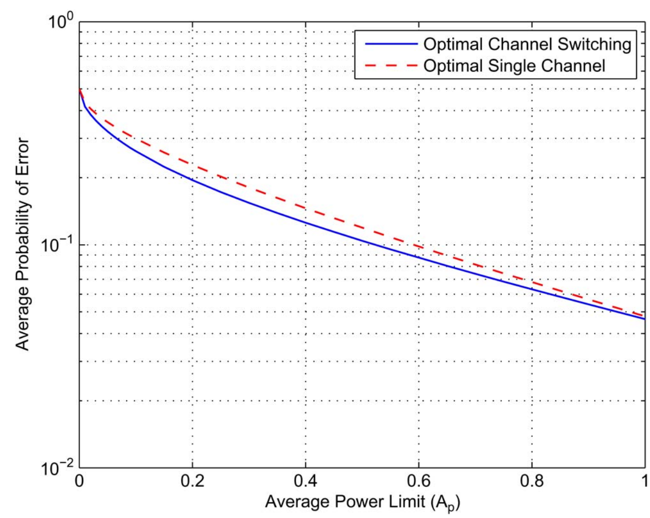

Fig. 4. A closer look at Fig. 3 for $A_p \in [0, 1]$.

TABLE II
PARAMETERS OF THE OPTIMAL CHANNEL SWITCHING STRATEGY IN FIG. 3

A_p	λ_i	λ_j	P_1	P_2	P_3	P_4
0.1	0.6667	0.3333	0.1281	–	–	0.0437
0.2	0.6667	0.3333	0.2314	–	–	0.1371
1	0.6667	0.3333	0.6894	–	–	1.6212
2	1	–	–	2	–	–
3	1	–	–	3	–	–
4	1	–	–	4	–	–
5	1	–	–	5	–	–

To investigate the effects of the average cost limit in more detail, the average probabilities of error are plotted versus A_c in Fig. 5 for various values of A_p based on the same channel configuration as in the previous scenario. As expected, the average probability of error is a non-increasing function of the

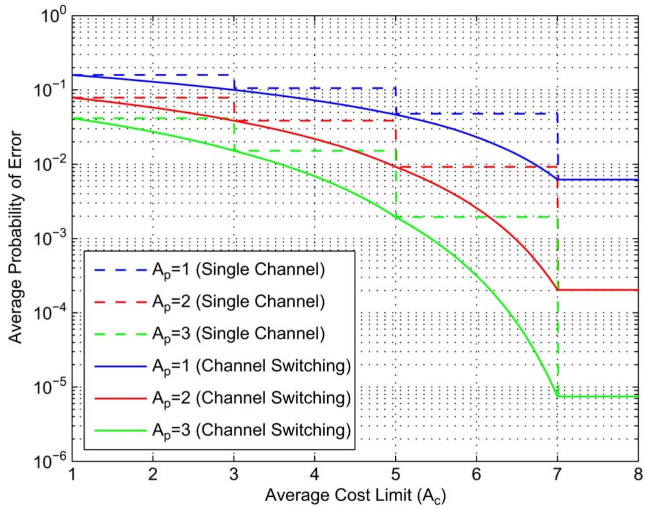

Fig. 5. Average probability of error versus $A_{\rm c}$ for the optimal single channel and optimal channel switching strategies, where $K = 4$, $\boldsymbol{\sigma} = [0.4\ 0.6\ 0.8\ 1]$, and $\boldsymbol{C} = [7\ 5\ 3\ 1]$.

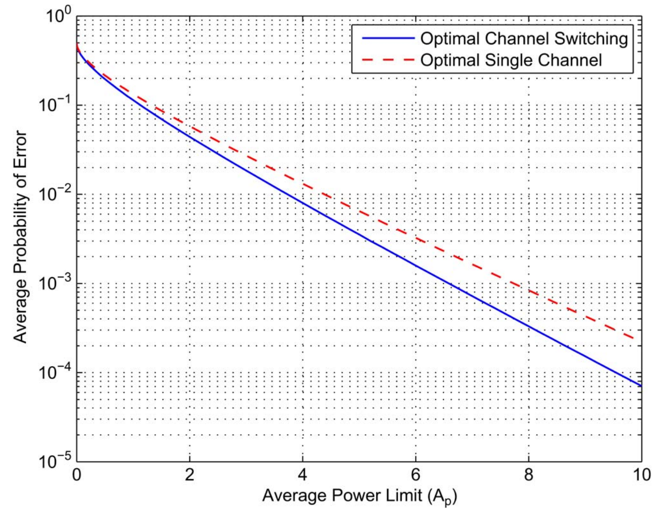

Fig. 6. Average probability of error versus $A_{\rm p}$ for the optimal single channel and optimal channel switching strategies, where $K = 5$, $\boldsymbol{\sigma} = [0.6\ 0.7\ 0.8\ 0.9\ 1]$, $\boldsymbol{C} = [1.329\ 1.112\ 0.941\ 0.804\ 0.6931]$, and $A_{\rm c} = 0.9$.

average cost limit $A_{\rm c}$. Also, in accordance with Part (ii) of Remark 1, the average probability of error converges to the error probability of the best channel (channel 1) at the average power limit $A_{\rm p}$ when $A_{\rm c}$ is larger than or equal to the cost of the best channel; i.e., when $A_{\rm c} \geq 7$. In addition, it is observed that the optimal single channel strategy results in piecewise constant average probabilities of error, which is due to the fact that the optimal single channel solution corresponds the use of the best channel that has a cost lower than or equal to $A_{\rm c}$. Specifically, the optimal single channel strategy achieves the error probabilities of $g_4(A_{\rm p})$, $g_3(A_{\rm p})$, $g_2(A_{\rm p})$, and $g_1(A_{\rm p})$ for $A_{\rm c} \in [1,3)$, $A_{\rm c} \in [3,5)$, $A_{\rm c} \in [5,7)$, and $A_{\rm c} \geq 7$, respectively, where $g_i(A_{\rm p}) = Q(\sqrt{A_{\rm p}}/\sigma_i)$ denotes the error probability of channel i at power level $A_{\rm p}$. Furthermore, Fig. 5 verifies the argument in Corollary 1 that, for $C_K < A_{\rm c} < C_1$, channel switching is guaranteed to outperform the optimal single channel strategy if $A_{\rm c}$ is not equal to the cost of one of the channels.

As another scenario, a five-channel system is considered, and the cost values are calculated based on the logarithmic cost function in (16) with $b = 1$. The standard deviations of the channels are set to $\boldsymbol{\sigma} = [0.6\ 0.7\ 0.8\ 0.9\ 1]$, and the average cost limit is given by $A_{\rm c} = 0.9$. In Fig. 6, the average probability of error is plotted versus $A_{\rm p}$ for the optimal single channel and optimal channel switching strategies. Similar to the scenario in Fig. 2, it is observed that channel switching outperforms the single channel approach for all values of $A_{\rm p}$ as a consequence of Corollary 1 as there exists no channel with a cost equal to $A_{\rm c}$. The parameters of the optimal channel switching strategy in Fig. 6 are presented in Table III for some values of $A_{\rm p}$. It is noted that the optimal solution performs channel switching among at most three channels in compliance with Proposition 2. Also, numerical calculations show that the results in Proposition 1 and Proposition 3 are satisfied. In addition, as stated in Proposition 4, when $A_{\rm p} \geq b\ \sigma_5^4/(\sigma_1^2(2\ \sigma_5^2 + b)) = 0.926$,[5] the optimal channel switching between two channels is performed between channel 3 and channel 4, which is in accordance with (17) and (18). Furthermore, channel switching among three channels is not optimal for $A_{\rm p} \geq 2\,b\,\sigma_5^4/(\sigma_1^2(2\,\sigma_5^2 + b)) = 1.852$.

[5] $P_{\max} > A_{\rm p}\sigma_K^2/\sigma_1^2 = 2.778A_{\rm p}$ is always satisfied in this scenario since $P_{\max} = 10A_{\rm p}$.

TABLE III
PARAMETERS OF THE OPTIMAL CHANNEL SWITCHING STRATEGY IN FIG. 6

$A_{\rm p}$	λ_i	λ_j	P_1	P_2	P_3	P_4	P_5
0.05	0.3028	0.0347	0.0811	0.0652	–	–	0.0352
0.2	0.3254	0.6746	0.2652	–	–	–	0.1685
1	0.7007	0.2993	–	–	0.9881	1.0278	–
2	0.7007	0.2993	–	–	1.9310	2.1617	–
3	0.7007	0.2993	–	–	2.8648	3.3164	–
4	0.7007	0.2993	–	–	3.7955	4.4788	–
5	0.7007	0.2993	–	–	4.7247	5.6447	–

Finally, the average probabilities of error are plotted versus $A_{\rm c}$ in Fig. 7 for various values of $A_{\rm p}$ based on the scenario in Fig. 6. Similar observations to those related to Fig. 5 can be made. Namely, if $A_{\rm c}$ is smaller than the cost of the best channel, which is equal to 1.329, the optimal channel switching strategy outperforms the single channel one when $A_{\rm c}$ is not equal to the cost of a channel. On the other hand, for $A_{\rm c} \geq 1.329$, both strategies achieve an average probability of error that is equal to the error probability of the best channel at the power limit.

VI. EXTENSIONS

The problem formulation in Section II assumes that there is a single RF chain at the transmitter and the receiver; hence, only one channel is employed at any time during channel switching. Also, the transmitter has an average power constraint denoted by $A_{\rm p}$, which can, for example, be determined according to the hardware constraints and/or the battery life of the communication system. This power constraint specifies a restriction on the transmit powers that can be used over different channels. An important extension is the scenario in which there exist multiple RF chains at the transmitter and the receiver, and multiple channels can be used simultaneously. For that

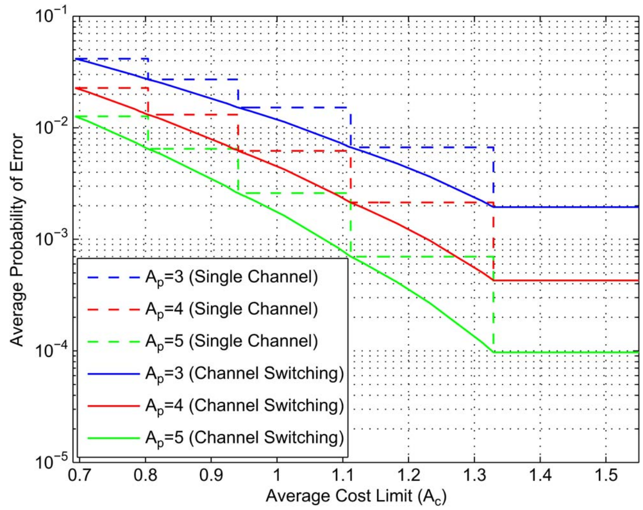

Fig. 7. Average probability of error versus A_c for the optimal single channel and optimal channel switching strategies, where $K = 5$, $\boldsymbol{\sigma} = [0.6\ 0.7\ 0.8\ 0.9\ 1]$, and $\boldsymbol{C} = [1.329\ 1.112\ 0.941\ 0.804\ 0.6931]$.

scenario, two different cases can be considered. In the first case, the transmitter has the same average power constraint as in the single RF chain scenario, and multiple RF chains share the power at each time. In that case, it can be shown that only one of the RF chains should be used with all the available power for minimizing the average probability of error. Therefore, this case reduces to the case with single RF chains investigated in the previous sections. For example, consider two RF chains at the transmitter and the receiver, and assume that power P is allocated for transmission over two channels simultaneously, which are denoted as channel i and channel j with average noise powers of σ_i^2 and σ_j^2, respectively. If power P_i is allocated for the RF chain that operates over channel i, and power $(P - P_i)$ for the RF chain over channel j, then the SNR at the receiver becomes $P_i/\sigma_i^2 + (P - P_i)/\sigma_j^2$ via optimal processing [28], which is maximized by setting $P_i = P$ if $\sigma_i < \sigma_j$ and $P_i = 0$ otherwise; that is, all the power is used for the RF chain corresponding to the best channel. In the second case, it is considered that the same power level can be used for all the RF chains, which corresponds to an increased total power due to the use of multiple RF chains. In that case, an equivalent model can be developed and the proposed model in Section II can still be employed as follows: Suppose that there exist R RF chains at the transmitter and the receiver. Then, all R combinations of K channels can be considered as new *combined channels*; that is, $\binom{K}{R}$ combined channels exist. For the ith combined channel, let the average noise powers and the costs of the corresponding R channels be denoted by $\sigma_{i,1}^2, \ldots, \sigma_{i,R}^2$ and $C_{i,1}, \ldots, C_{i,R}$, respectively. Then, this combined channel can be considered as a single channel as in Fig. 1 with an average noise power of $\tilde{\sigma}_i^2 = (\sigma_{i,1}^{-2} + \cdots + \sigma_{i,R}^{-2})^{-1}$ and a cost of $\tilde{C}_i = C_{i,1} + \cdots + C_{i,R}$. Hence, the same problem formulation as in Section II is obtained, and all the results in the previous sections apply. It should be emphasized that the expression for $\tilde{\sigma}_i^2$ is obtained by considering the SNR at the receiver after optimal processing, which is expressed as $P/\tilde{\sigma}_i^2 = P/\sigma_{i,1}^2 + \cdots + P/\sigma_{i,R}^2$ [28].

VII. Concluding Remarks

In this study, optimal channel switching has been investigated for Gaussian channels in the presence of average power and average cost constraints. For generic cost functions, it has been shown that the optimal channel switching strategy performs time sharing among at most three channels and operates at the average power and average cost limits. Also, for channel switching between two channels, it has been proved that the ratio of the optimal power levels is upper bounded by the ratio of the larger noise variance to the smaller one under certain conditions. In addition, for logarithmic cost functions, the convexity properties of the error probability have been characterized as a function of power and cost, and the optimal channel switching strategy has been shown to employ at most two channels, which can be determined based on specific formulas, in certain scenarios. Numerical examples have provided illustrations of the theoretical results. Future work involves the incorporation of switching costs [20] in the design of optimal channel switching strategies.

Appendix

A. Proof of Proposition 3

First, consider the problem in (11) without the peak power constraints. Then, it can be solved based on (13). The first-order derivative of the objective function in (13) with respect to P_i is expressed as

$$\frac{-\lambda^* \kappa\, \eta}{2\sqrt{2\pi}\ \sigma_i \sqrt{P_i}}\, \mathrm{e}^{-\frac{\kappa^2 P_i}{2\sigma_i^2}} + \frac{\lambda^* \sqrt{1-\lambda^*}\ \kappa\, \eta}{2\sqrt{2\pi}\ \sigma_j \sqrt{A_\mathrm{p} - \lambda^* P_i}}\, \mathrm{e}^{-\frac{\kappa^2 (A_\mathrm{p} - \lambda^* P_i)}{2(1-\lambda^*)\sigma_j^2}}. \tag{25}$$

It is observed that the first-order derivative in (25) is a monotone increasing function of P_i for $P_i \in (0,\ A_\mathrm{p}/\lambda^*)$, which starts from $-\infty$ at $P_i = 0$ and increases monotonically towards infinity as P_i goes to A_p/λ^*. Therefore, there is a unique minimizer P_i^* for the optimization problem in (13), which corresponds to the point at which the first-order derivative is zero. Equating the first-order derivative in (25) to zero yields the following necessary and sufficient condition for the optimal solution of (13):

$$\mathrm{e}^{\frac{\kappa^2 P_i}{2\sigma_i^2} - \frac{\kappa^2 (A_\mathrm{p} - \lambda^* P_i)}{2(1-\lambda^*)\sigma_j^2}} = \frac{\sigma_j \sqrt{A_\mathrm{p} - \lambda^* P_i}}{\sqrt{1-\lambda^*}\ \sigma_i \sqrt{P_i}}. \tag{26}$$

Since $\lambda^* P_i + (1 - \lambda^*) P_j = A_\mathrm{p}$, the condition in (26) can also be expressed as

$$\mathrm{e}^{\kappa^2 \left(\frac{P_i}{\sigma_i^2} - \frac{P_j}{\sigma_j^2} \right)} = \frac{\sigma_j^2 P_j}{\sigma_i^2 P_i}. \tag{27}$$

If $A_\mathrm{p} = A_{ij}$, it can be shown, by using the definition of A_{ij} in the proposition, that $P_i = P_j = A_{ij}$ satisfies the condition in (27). Since the solution of (27) is unique, the optimal solution of (13) is obtained as $P_i^* = P_j^* = A_{ij}$. In addition, as $P_\mathrm{max} > A_\mathrm{p} = A_{ij}$, the optimal solution of the problem in (11) is the same as that of (13) in this case; hence, the first part of Proposition 3 is obtained.

To prove the second part of the proposition, it is first observed that the first-order derivative in (25) is a monotone decreasing function of $A_{\rm p}$ and a monotone increasing function of P_i. Therefore, the value of P_i at which the first-order derivative becomes zero gets larger as $A_{\rm p}$ increases. Since the first-order derivative becomes zero at $P_i = A_{ij}$ when $A_{\rm p} = A_{ij}$ (as proved in the first part), the first-order derivative becomes zero at a value larger than A_{ij} when $A_{\rm p} > A_{ij}$. Hence, the optimal solution of (13) satisfies $P_i^* > A_{ij}$ for $A_{\rm p} > A_{ij}$. In addition, it is concluded from (27) that as P_i^* increases, the optimal value of P_j should also increase for the optimality condition in (27) to be satisfied. In other words, $P_i^* > A_{ij}$ also implies $P_j^* > A_{ij}$ based on the relation in (27). Next, the ordering between P_i^* and P_j^* should be determined. To that aim, the optimal signal values are expressed as $P_i^* = \alpha A_{ij}$ and $P_j^* = \beta A_{ij}$, where α and β are some positive numbers that are larger than one. Then, the optimality condition in (27) becomes $\mathrm{e}^{\kappa^2 A_{ij}(\alpha/\sigma_i^2 - \beta/\sigma_j^2)} = \beta\sigma_j^2/(\alpha\sigma_i^2)$. From the definition of A_{ij} in the proposition, σ_j^2/σ_i^2 can be expressed as $\sigma_j^2/\sigma_i^2 = \mathrm{e}^{\kappa^2 A_{ij}(1/\sigma_i^2 - 1/\sigma_j^2)}$. Then, the optimality condition is stated as

$$\frac{\alpha}{\beta} = \mathrm{e}^{\kappa^2 A_{ij}\left(\frac{\beta-1}{\sigma_j^2} - \frac{\alpha-1}{\sigma_i^2}\right)}. \tag{28}$$

If it is assumed that $\alpha > \beta$, then (28) implies that $\frac{\beta-1}{\alpha-1} > \frac{\sigma_j^2}{\sigma_i^2}$. However, since $\sigma_i^2 < \sigma_j^2$ (as $C_i > C_j$), this inequality leads to a contradiction. Therefore, α cannot be larger than β. On the other hand, if it is assumed that $\alpha < \beta$, then (28) becomes $\frac{\alpha-1}{\beta-1} > \frac{\sigma_i^2}{\sigma_j^2}$, which is not a contradiction. Therefore, it is obtained that $\alpha < \beta$, that is, $P_i^* < P_j^*$, when $A_{\rm p} > A_{ij}$. Furthermore, due to the average power constraint, $\lambda^* P_i^* + (1 - \lambda^*)P_j^* = A_{\rm p}$, it is concluded that $P_j^* > A_{\rm p} > P_i^*$. Combining this result with the first result in this paragraph, it is obtained that when $A_{\rm p} > A_{ij}$, the optimal signal values satisfy $P_j^* > A_{\rm p} > P_i^* > A_{ij}$. Hence, the second part of Proposition 3 is proved. The third part of the proposition can be proved in a similar manner to the proof of the second part, and it can be shown that $A_{ij} > P_i^* > A_{\rm p} > P_j^*$ based on (25) and (27).

The final statement in the proposition can be proved as follows: For $A_{\rm p} > A_{ij}$, it is obtained in the previous paragraph that $\frac{\alpha-1}{\beta-1} > \frac{\sigma_i^2}{\sigma_j^2}$, where $\beta > \alpha > 1$ with $P_i^* = \alpha A_{ij}$ and $P_j^* = \beta A_{ij}$. The inequality can be manipulated as follows:

$$\frac{\sigma_j^2}{\sigma_i^2} > \frac{\beta-1}{\alpha-1} > \frac{\beta}{\alpha} = \frac{P_j^*}{P_i^*} \tag{29}$$

where the second inequality is obtained from the relation $\beta > \alpha > 1$. For $A_{\rm p} < A_{ij}$, the second part of the proposition states that $P_i^* > P_j^*$. Since $\sigma_i^2 < \sigma_j^2$ by definition (as $C_i > C_j$), it is obtained that $P_i^*/\sigma_i^2 > P_j^*/\sigma_j^2$. Therefore, the relation in (27) yields

$$\frac{\sigma_j^2 P_j^*}{\sigma_i^2 P_i^*} = \mathrm{e}^{\frac{P_i^*}{\sigma_i^2} - \frac{P_j^*}{\sigma_j^2}} > 1, \tag{30}$$

which results in $P_i^*/P_j^* < \sigma_j^2/\sigma_i^2$. Finally, for $A_{\rm p} = A_{ij}$, $P_i^*/P_j^* = 1$ as stated in the first part of the proposition. Overall, the ratio between the optimal power levels is upper bounded by σ_j^2/σ_i^2 for any value of $A_{\rm p}$, as stated in the proposition.

Based on the three conditions in the proposition and the inequality in (14), it can be shown that the optimal power levels P_i^* and P_j^* are always smaller than $P_{\max}$ since $P_{\max} > A_{\rm p}\sigma_j^2/\sigma_i^2$. Hence, the properties of the solution of (11) obtained without the peak power constraints (via the solution of (13)) also hold for the solution of the problem in (11) in the presence of peak power constraints. ■

B. *Proof of Lemma 1*

The error probability for a transmission power of P is expressed as $\eta\, Q(\kappa\sqrt{P}/\sigma)$ as in (12), where σ is the standard deviation of the channel noise. Based on the cost function in (16), σ is expressed as $\sigma = \sqrt{b/(\mathrm{e}^C - 1)}$, which leads to the following expression for the error probability:

$$g(P, C) = \eta\, Q\left(h(P) f(C)\right), \tag{31}$$

where $h(P) \triangleq \kappa\sqrt{P}$ and $f(C) \triangleq \sqrt{(\mathrm{e}^C - 1)/b}$. To investigate the convexity of (31), the derivatives of h and f are calculated first, which are expressed as

$$h' = \frac{\kappa^2}{2h}, \qquad f' = \frac{bf^2+1}{2bf},$$
$$h'' = -\frac{\kappa^4}{4h^3},\ f'' = \frac{(bf^2-1)(bf^2+1)}{4b^2f^3}. \tag{32}$$

Then, the first-order partial derivatives of $g(P, C)$ with respect to P and C are given by

$$\frac{\partial g(P,C)}{\partial P} = \eta\, h' f\, Q'(hf) \text{ and } \frac{\partial g(P,C)}{\partial C} = \eta\, h f'\, Q'(hf), \tag{33}$$

where $Q'(x)$ denotes the first derivative of the Q-function. (From (33), it is observed that the error probability is a monotone decreasing function of power and cost as expected since Q-function is monotone decreasing.) Next, the second-order partial derivatives are calculated as

$$\begin{aligned}\frac{\partial^2 g(P,C)}{\partial P^2} &= \eta\,(h'f)^2 Q''(hf) + \eta\, h'' f\, Q'(hf)\\ &= \eta\left(fh'' - f^3 h (h')^2\right) Q'(hf)\end{aligned} \tag{34}$$

$$\begin{aligned}\frac{\partial^2 g(P,C)}{\partial C^2} &= \eta\,(hf')^2 Q''(hf) + \eta\, hf'' Q'(hf)\\ &= \eta\left(hf'' - h^3 f (f')^2\right) Q'(hf)\end{aligned} \tag{35}$$

$$\begin{aligned}\frac{\partial^2 g(P,C)}{\partial P \partial C} &= \eta\, hh'ff' Q''(hf) + \eta\, h'f' Q'(hf)\\ &= \eta(1 - h^2f^2) h'f' Q'(hf)\end{aligned} \tag{36}$$

where the relation $Q''(x) = -x\, Q'(x)$ is employed to obtain the final expressions. From (34)–(36), the 2×2 Hessian matrix can be formed for $g(P,C)$, and the convexity of $g(P,C)$ can be investigated based on the positive definiteness of the Hessian matrix, which requires the leading principal minors to be positive [30]. It noted from (34) that the second-order derivative with respect to P is always positive. Therefore, the only condition for positive definiteness becomes the determinant of the Hessian matrix to be positive, which leads, after some manipulation, to the following inequality:

$$P > \frac{b}{\kappa^2(bf^2+2)} = \frac{b}{\kappa^2(\mathrm{e}^C+1)}, \tag{37}$$

where the final expression is obtained based on the definition of f; i.e., $f(C) = \sqrt{(\mathrm{e}^C-1)/b}$. Therefore, the convexity of $g(P,C)$ requires that power P should be larger than $b/(\kappa^2(\mathrm{e}^C+1))$ as stated in the lemma.

Finally, it is shown that $\mathcal{S}_c$, as defined in the lemma, is a convex set. Let (P_i, C_i) and (P_j, C_j) denote any two elements from set $\mathcal{S}_c$. Then, their convex combination is given by $(\lambda\, P_i + (1-\lambda)P_j, \lambda\, C_i + (1-\lambda)C_j)$, where $\lambda \in [0,1]$. Since both C_i and C_j are in $(C_{\min}, C_{\max})$, their convex combination resides in the same interval as well. In addition, the convex combination of the powers satisfies the condition for set $\mathcal{S}_c$ due to the following inequalities:

$$\begin{aligned}\lambda\, P_i + (1-\lambda)P_j &> \lambda\, \frac{b/\kappa^2}{\mathrm{e}^{C_i}+1} + (1-\lambda)\, \frac{b/\kappa^2}{\mathrm{e}^{C_j}+1} \\ &> \frac{b/\kappa^2}{\mathrm{e}^{\lambda\, C_i + (1-\lambda)\, C_j}+1}\end{aligned} \tag{38}$$

where the second inequality follows from the strict convexity of $b/(\kappa^2(\mathrm{e}^C+1))$. Therefore, $\mathcal{S}_c$ is a convex set, and $g(P,C)$ is a strictly convex function over set $\mathcal{S}_c$. ■

C. Proof of Proposition 5

The statement in the proposition can be proved via contradiction. Suppose that the optimal solution is to switch among three different channels, and let the channel indices, channel switching factors, and power levels for that optimal solution be denoted by (i,j,k), $(\lambda_i, \lambda_j, \lambda_k)$ and (P_i, P_j, P_k), respectively, where $C_i > C_j > C_k$ without loss of generality. Since the optimal solution must utilize the maximum average cost A_c (see Proposition 1), either $C_i > A_\mathrm{c} > C_j > C_k$ or $C_i > C_j > A_\mathrm{c} > C_k$ must hold. Assume that $C_i > A_\mathrm{c} > C_j > C_k$. (The proof for the other scenario can be obtained in a similar manner.) As stated in Proposition 1, the optimal solution operates at the maximum average power and cost, which leads to the following equality:

$$\lambda_i\, \mathbf{z}_i + \lambda_j\, \mathbf{z}_j + (1-\lambda_i-\lambda_j)\, \mathbf{z}_k = \boldsymbol{a}, \tag{39}$$

where $\mathbf{z}_i = (P_i, C_i)$ and $\boldsymbol{a} = (A_\mathrm{p}, A_\mathrm{c})$, as in the proof of Proposition 4. Consider an alternative solution that switches between two channels, channel i and channel j, with channel switching factors γ and $(1-\gamma)$ and powers P_i and P_j, respectively, and utilizes the maximum average power and cost; that is,

$$\gamma\, \mathbf{z}_i + (1-\gamma)\, \mathbf{z}_j = \boldsymbol{a}. \tag{40}$$

Based on (39) and (40), γ and λ_i can be obtained from the average cost constraint as $\gamma = (A_\mathrm{c} - C_j)/(C_i - C_j)$ and $\lambda_i = (A_\mathrm{c} - C_k - \lambda_j(C_j - C_k))/(C_i - C_k)$. First, it is shown that $\lambda_i > \gamma$. To that aim, the following inequality is obtained from the condition $\lambda_i > \gamma$ based on the definition of λ_i and γ: $(A_\mathrm{c} - C_k - \lambda_j(C_j - C_k))(C_i - C_j) > (A_\mathrm{c} - C_j)(C_i - C_k)$, which reduces, after some manipulation, to $\lambda_j\, C_j + (1-\lambda_j)\, C_i > A_\mathrm{c}$. Since $(1-\lambda_j) = \lambda_i + \lambda_k$, $C_i > C_k$, and $\lambda_i\, C_i + \lambda_j\, C_j + \lambda_k\, C_k = A_\mathrm{c}$, the inequality $\lambda_j\, C_j + (1-\lambda_j)\, C_i > A_\mathrm{c}$ always holds, which verifies that $\lambda_i > \gamma$. Then, from (39) and (40), $\mathbf{z}_j$ can be expressed as

$$\mathbf{z}_j = \frac{\lambda_i - \gamma}{1-\gamma-\lambda_j}\, \mathbf{z}_i + \frac{1-\lambda_i-\lambda_j}{1-\gamma-\lambda_j}\, \mathbf{z}_k. \tag{41}$$

The remaining part of the proof depends on the values of powers P_i, P_j, and P_k.

Case 1: If all the power levels satisfy the convexity condition in Lemma 1, then the following inequality can be obtained:

$$\begin{aligned}&\gamma\, g(\mathbf{z}_i) + (1-\gamma)\, g(\mathbf{z}_j) \\ &= \gamma\, g(\mathbf{z}_i) + \lambda_j\, g(\mathbf{z}_j) + (1-\gamma-\lambda_j)\, g(\mathbf{z}_j) &(42)\\ &< \gamma\, g(\mathbf{z}_i) + \lambda_j\, g(\mathbf{z}_j) + (1-\gamma-\lambda_j) \\ &\quad \times \left(\frac{\lambda_i-\gamma}{1-\gamma-\lambda_j}\, g(\mathbf{z}_i) + \frac{1-\lambda_i-\lambda_j}{1-\gamma-\lambda_j}\, g(\mathbf{z}_k)\right) &(43)\\ &= \lambda_i\, g(\mathbf{z}_i) + \lambda_j\, g(\mathbf{z}_j) + (1-\lambda_i-\lambda_j)\, g(\mathbf{z}_k) &(44)\end{aligned}$$

where $g(\mathbf{z}_i) = g(P_i, C_i)$ denotes the average probability of error as a function of power and cost, as defined in (31). In obtaining the inequality in (43), the definition of $\mathbf{z}_j$ in (41) and the strict convexity of g are employed. Note that g is strictly convex when the power levels satisfy the condition in Lemma 1, which is the assumption in Case 1. In addition, it is noted that the $(1-\gamma-\lambda_j)$ term in (42) is never negative since $\gamma < \lambda_i$ as proved in the previous paragraph and $\lambda_i < 1-\lambda_j$ by definition. The inequality in (42)–(44) implies that the channel switching between channel i and channel j with channel switching factors γ and $(1-\gamma)$, respectively, achieves a lower average probability of error than the optimal solution, which switches among channels i, j, and k with channel switching factors λ_i, λ_j, and λ_k, respectively. Hence, a contradiction arises. Therefore, the strategy that switches among three channels cannot be optimal. In other words, for any strategy that switches among three channels, there exist a strategy that performs channel switching between two channels and achieves a smaller average probability of error.

Case 2: Suppose that some of the power levels do not satisfy the convexity condition in Lemma 1. Since the average power should be equal to A_p due to Proposition 1, at least one power level should be below A_p. Assume without loss of generality that $P_i < A_\mathrm{p}$. Then, the average probability of error for the optimal solution that switches among three different channels can be bounded from below as follows:

$$\lambda_i\, g(\mathbf{z}_i) + \lambda_j\, g(\mathbf{z}_j) + \lambda_k\, g(\mathbf{z}_k)$$

$$= \lambda_i\, g(\mathbf{z}_i) + (\lambda_j + \lambda_k)\left(\tilde{\lambda}_j\, g(\mathbf{z}_j) + \tilde{\lambda}_k\, g(\mathbf{z}_k)\right) \tag{45}$$

$$\geq \lambda_i\, g(\mathbf{z}_i) + (\lambda_j + \lambda_k)\left(\nu_j\, g(\tilde{\mathbf{z}}_j) + \nu_k\, g(\tilde{\mathbf{z}}_k)\right) \tag{46}$$

$$= \lambda_i\, g(\mathbf{z}_i) + \tilde{\nu}_j\, g(\tilde{\mathbf{z}}_j) + \tilde{\nu}_k\, g(\tilde{\mathbf{z}}_k) \tag{47}$$

with $\tilde{\lambda}_j \triangleq \lambda_j/(\lambda_j + \lambda_k)$, $\tilde{\lambda}_k \triangleq \lambda_k/(\lambda_j + \lambda_k)$, $\tilde{\nu}_j \triangleq (\lambda_j + \lambda_k)\nu_j$, $\tilde{\nu}_k \triangleq (\lambda_j + \lambda_k)\nu_k$, $\tilde{\mathbf{z}}_j \triangleq (\tilde{P}_j, C_j)$, and $\tilde{\mathbf{z}}_k \triangleq (\tilde{P}_k, C_k)$, where $\tilde{P}_j$ and $\tilde{P}_k$ are the optimal power levels and ν_j and ν_k are the corresponding optimal channel switching factors when the channel switching is performed between channel j and channel k *only* under the average cost limit $\tilde{A}_\mathrm{c} \triangleq \tilde{\lambda}_j\ C_j + \tilde{\lambda}_k\ C_k$ and the average power limit $\tilde{A}_\mathrm{p} \triangleq \tilde{\lambda}_j\ P_j + \tilde{\lambda}_k\ P_k > A_\mathrm{p}$.[6] Since $(\tilde{P}_j, \tilde{P}_k)$ and (ν_j, ν_k) are the solution of the optimal channel switching problem in the presence of channel j and channel k only under the average power limit $\tilde{A}_\mathrm{p}$ and the average cost limit $\tilde{A}_\mathrm{c}$, the average probability of error is bounded from below by the expression in (46). From Proposition 3, $\min\{\tilde{P}_j, \tilde{P}_k\} > \sigma_j^2 \tilde{A}_\mathrm{p}/\sigma_k^2$. Then, based on a similar argument to that in (24), it can be shown that the convexity condition in Lemma 1 is satisfied for power levels $\tilde{P}_j$ and $\tilde{P}_k$ if $\tilde{A}_\mathrm{p} > b\ \sigma_K^4/(\kappa^2(2\ \sigma_1^2\sigma_K^2 + b\ \sigma_1^2))$, which always holds due to the assumption in the proposition and the fact that $\tilde{A}_\mathrm{p} > A_\mathrm{p}$. Next assume without loss of generality that $\tilde{\nu}_j\ \tilde{P}_j \geq \tilde{\nu}_k\ \tilde{P}_k$. Then, the lower bound in (47) can be improved as follows:

$$\lambda_i\, g(\mathbf{z}_i) + \tilde{\nu}_j\, g(\tilde{\mathbf{z}}_j) + \tilde{\nu}_k\, g(\tilde{\mathbf{z}}_k)$$

$$= (\lambda_i + \tilde{\nu}_j)\left(\lambda_i^*\, g(\mathbf{z}_i) + \lambda_j^*\, g(\tilde{\mathbf{z}}_j)\right) + \tilde{\nu}_k\, g(\tilde{\mathbf{z}}_k) \tag{48}$$

$$\geq (\lambda_i + \tilde{\nu}_j)\left(\nu_i^*\, g(\mathbf{z}_i^*) + \nu_j^*\, g(\mathbf{z}_j^*)\right) + \tilde{\nu}_k\, g(\tilde{\mathbf{z}}_k) \tag{49}$$

$$= \hat{\nu}_i\, g(\mathbf{z}_i^*) + \hat{\nu}_j\, g(\mathbf{z}_j^*) + \tilde{\nu}_k\, g(\tilde{\mathbf{z}}_k) \tag{50}$$

with $\lambda_i^* \triangleq \lambda_i/(\lambda_i + \tilde{\nu}_j)$, $\lambda_j^* \triangleq \tilde{\nu}_j/(\lambda_i + \tilde{\nu}_j)$, $\hat{\nu}_i \triangleq (\lambda_i + \tilde{\nu}_j)\nu_i^*$, $\hat{\nu}_j \triangleq (\lambda_i + \tilde{\nu}_j)\nu_j^*$, $\mathbf{z}_i^* \triangleq (P_i^*, C_i)$, and $\mathbf{z}_j^* \triangleq (P_j^*, C_j)$, where P_i^* and P_j^* are the optimal power levels and ν_i^* and ν_j^* are the corresponding optimal channel switching factors when the channel switching is performed between channel i and channel j *only* under the average cost limit $A_\mathrm{c}^* \triangleq \lambda_i^*\ C_i + \lambda_j^*\ C_j$ and the average power limit $A_\mathrm{p}^* \triangleq \lambda_i^*\ P_i + \lambda_j^*\ \tilde{P}_j > 0.5\ A_\mathrm{p}$.[7] From Proposition 3, $\min\{P_i^*, P_j^*\} > \sigma_i^2 A_\mathrm{p}^*/\sigma_j^2$. Then, similar to (24), it can be shown that the convexity condition in Lemma 1 is satisfied for power levels P_i^* and P_j^* if $A_\mathrm{p}^* > b\,\sigma_K^4/(\kappa^2(2\,\sigma_1^2\sigma_K^2 + b\ \sigma_1^2))$, which is true due to the assumption in the proposition and the fact that $A_\mathrm{p}^* > 0.5 A_\mathrm{p}$. From (45)–(50), it is concluded that when the assumption in the proposition holds, for any strategy that performs channel switching among three different channels, there exists another strategy that switches among the same channels with power levels that satisfy the convexity condition in Lemma 1, and achieves a smaller average probability of error. Therefore, the arguments in the previous part of the proof (Case 1) can be employed to show that there exists a strategy that performs channel switching between two channels and achieves a smaller average probability of error than the lower bound in (50). Therefore, channel switching among three channels cannot be optimal under the condition in the proposition. ■

[6]The inequality $\tilde{A}_\mathrm{p} > A_\mathrm{p}$ follows from the assumption that $P_i < A_\mathrm{p}$.

[7]The inequality $A_\mathrm{p}^* > 0.5\ A_\mathrm{p}$ follows from the assumptions that $\tilde{\nu}_j\ \tilde{P}_j \geq \tilde{\nu}_k\ \tilde{P}_k$ and $P_i < A_\mathrm{p}$.

ACKNOWLEDGMENT

The authors would like to thank Musa Furkan Keskin from Bilkent University for his insightful comments.

REFERENCES

[1] M. Azizoglu, “Convexity properties in binary detection problems,” *IEEE Trans. Inf. Theory*, vol. 42, no. 4, pp. 1316–1321, Jul. 1996.

[2] C. Goken, S. Gezici, and O. Arikan, “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.

[3] S. Loyka, V. Kostina, and F. Gagnon, “Error rates of the maximum-likelihood detector for arbitrary constellations: Convex/concave behavior and applications,” *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1948–1960, Apr. 2010.

[4] B. Dulek, S. Gezici, and O. Arikan, “Convexity properties of detection probability under additive Gaussian noise: Optimal signaling and jamming strategies,” *IEEE Trans. Signal Process.*, vol. 61, no. 13, pp. 3303–3310, Jul. 2013.

[5] S. Bayram, N. D. Vanli, B. Dulek, I. Sezer, and S. Gezici, “Optimum power allocation for average power constrained jammers in the presence of non-Gaussian noise,” *IEEE Commun. Lett.*, vol. 16, no. 8, pp. 1153–1156, Aug. 2012.

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

[7] 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.

[8] H. Chen and P. K. Varshney, “Theory of the stochastic resonance effect in signal detection: Part II-Variable detectors,” *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5031–5041, Oct. 2007.

[9] M. E. Tutay, S. Gezici, and O. Arikan, “Optimal detector randomization for multiuser communications systems,” *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 2876–2889, Jul. 2013.

[10] A. Sezer, S. Gezici, and H. Inaltekin, “Optimal channel switching for average capacity maximization,” in *Proc. IEEE ICASSP*, May 2014, pp. 3503–3507.

[11] J. A. Ritcey and M. Azizoglu, “Performance analysis of generalized selection combining with switching constraints,” *IEEE Commun. Lett.*, vol. 4, no. 5, pp. 152–154, May 2000.

[12] B. Dulek, M. E. Tutay, S. Gezici, and P. K. Varshney, “Optimal signaling and detector design for M-ary communication systems in the presence of multiple additive noise channels,” *Digit. Signal Process.*, vol. 26, pp. 153–168, Mar. 2014.

[13] C. Goken, S. Gezici, and O. Arikan, “Optimal stochastic signaling for power-constrained binary communications systems,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp. 3650–3661, Dec. 2010.

[14] M. Tutay, S. Gezici, and O. Arikan, “Optimal randomization of signal constellations on the downlink of a multiuser DS-CDMA system,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 4878–4891, Oct. 2013.

[15] B. Dulek and S. Gezici, “Optimal stochastic signal design and detector randomization in the Neyman-Pearson framework,” in *Proc. 37th IEEE ICASSP*, Kyoto, Japan, Mar. 25–30, 2012, pp. 3025–3028.

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

[17] Y. Ma and C. C. Chai, “Unified error probability analysis for generalized selection combining in Nakagami fading channels,” *IEEE J. Sel. Areas Commun.*, vol. 18, no. 11, pp. 2198–2210, Nov. 2000.

[18] F. Gaaloul, H.-C. Yang, R. Radaydeh, and M.-S. Alouini, "Switch based opportunistic spectrum access for general primary user traffic model," *IEEE Wireless Commun. Lett.*, vol. 1, no. 5, pp. 424–427, Oct. 2012.

[19] Y. Liu and M. Liu, "To stay or to switch: Multiuser dynamic channel access," in *Proc. IEEE INFOCOM*, Apr. 2013, pp. 1249–1257.

[20] L. Chen, S. Iellamo, and M. Coupechoux, "Opportunistic spectrum access with channel switching cost for cognitive radio networks," in *Proc. IEEE ICC*, Kyoto, Japan, Jun. 2011, pp. 1–5.

[21] D. Niyato and E. Hossain, "Spectrum trading in cognitive radio networks: A market-equilibrium-based approach," *IEEE Wireless Commun.*, vol. 15, no. 6, pp. 71–80, Dec. 2008.

[22] J. Huang, R. Berry, and M. L. Honig, "Auction-based spectrum sharing," *ACM/Springer Mobile Netw. Appl.*, vol. 11, no. 3, pp. 405–418, Jun. 2006.

[23] D. Niyato and E. Hossain, "Competitive pricing for spectrum sharing in cognitive radio networks: Dynamic game, inefficiency of Nash equilibrium, and collusion," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 192–202, Jan. 2008.

[24] D. Niyato and E. Hossain, "Competitive spectrum sharing in cognitive radio networks: A dynamic game approach," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2651–2660, Jul. 2008.

[25] Z. Ji and K. Liu, "Dynamic spectrum sharing: A game theoretical overview," *IEEE Commun. Mag.*, vol. 45, no. 5, pp. 88–94, May 2007.

[26] A. Ozcelikkale, H. M. Ozaktas, and E. Arikan, "Signal recovery with cost-constrained measurements," *IEEE Trans. Signal Process.*, vol. 58, no. 7, pp. 3607–3617, Jul. 2010.

[27] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ, USA: Princeton Univ. Press, 1968.

[28] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[29] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[30] G. Strang, *Introduction to Linear Algebra*, 4th ed. Cambridge, MA, USA: Wellesley-Cambridge, 2009.

Mehmet Emin Tutay received the B.S., M.S., and Ph.D. degrees from the Department of Electrical and Electronics Engineering, Bilkent University, Ankara, Turkey, in 2008, 2010, and 2013, respectively. His main research interests are in the fields of statistical signal processing and wireless communications.

Sinan Gezici (S'03–M'06–SM'11) received the B.S. degree from Bilkent University, Ankara, Turkey, in 2001, and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, USA, in 2006. From 2006 to 2007, he worked at Mitsubishi Electric Research Laboratories, Cambridge, MA. Since 2007, he has been with the Department of Electrical and Electronics Engineering at Bilkent University, where he is currently an Associate Professor. His research interests are in the areas of detection and estimation theory, wireless communications, and localization systems. Among his publications in these areas is the book *Ultra-wideband Positioning Systems: Theoretical Limits, Ranging Algorithms, and Protocols* (Cambridge University Press, 2008). He is an Associate Editor for IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE WIRELESS COMMUNICATIONS LETTERS, and *Journal of Communications and Networks*.

Hamza Soganci received the B.Sc., M.Sc. and Ph.D. degrees in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2007, 2009 and 2015, respectively. He is currently working at TÜBİTAK SAGE. His research interests are detection and estimation theory, wireless localization, and ultra-wideband systems.

Orhan Arikan (M'91) was born in 1964 in Manisa, Turkey. He received the B.Sc. degree in electrical and electronics engineering from the Middle East Technical University, Ankara, Turkey, in 1986. He received the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Illinois, Urbana-Champaign, IL, USA, in 1988 and 1990, respectively. Following his graduate studies, he was employed as a Research Scientist at Schlumberger-Doll Research Center, Ridgefield, CT. In 1993, he joined the Electrical and Electronics Engineering Department of Bilkent University, Ankara, Turkey. Since 2011, he is serving as the Department Chairman. His current research interests include statistical signal processing, time-frequency analysis, and remote sensing. He has served as Chairman of IEEE Signal Processing Society Turkey Chapter and President of IEEE Turkey Section.