

Maximization of Average Number of Correctly Received Symbols over Multiple Channels in the Presence of Idle Periods

Musa Furkan Keskin, Mehmet Necip Kurt, Mehmet Emin Tutay, Sinan Gezici, and
Orhan Arikan

Abstract

In this study, optimal channel switching (time sharing) strategies are investigated under average power and cost constraints for maximizing the average number of correctly received symbols between a transmitter and a receiver that are connected via multiple flat-fading channels with additive Gaussian noise. The optimal strategy is shown to correspond to channel switching either among at most three different channels with full channel utilization (i.e., no idle periods), or between at most two different channels with partial channel utilization. Also, it is stated that the optimal solution must operate at the maximum average power and the maximum average cost, which facilitates low-complexity approaches for obtaining the optimal strategy. For two-channel strategies, an upper bound is derived, in terms of the parameters of the employed channels, on the ratio between the optimal power levels. In addition, theoretical results are derived for characterizing the optimal solution for channel switching between two channels, and for comparing performance of single channel strategies. Sufficient conditions that depend solely on the systems parameters are obtained for specifying when partial channel utilization cannot be optimal. Furthermore, the proposed optimal channel switching problem is investigated for logarithmic cost functions, and various theoretical results are obtained related to the optimal strategy. Numerical examples are presented to illustrate the validity of the theoretical results.

Index Terms– Channel switching, Gaussian channel, fading, probability of correct decision, partial transmission, logarithmic cost.

I. INTRODUCTION

Time sharing (randomization) has attracted a significant deal of interest in the literature due to its capability to provide performance improvements for communication systems [1]–[11]. In [2], it is demonstrated

The authors are with the Department of Electrical and Electronics Engineering, Bilkent University, Bilkent, Ankara 06800, Turkey, Tel: +90-312-290-3139, Fax: +90-312-266-4192, Email: {keskin,mnkurt,tutay,gezici,oarikan}@ee.bilkent.edu.tr. This research was supported in part by the Distinguished Young Scientist Award of Turkish Academy of Sciences (TUBA-GEBIP 2013). Part of this work will be presented at IEEE Wireless Communications and Networking Conference (WCNC) 2016.

that the average probability of error over additive noise channels with arbitrary noise probability density functions (PDFs) can be reduced via optimal stochastic signaling, which performs time sharing among at most three different signal levels for each information symbol. The study in [5] investigates performance gains that can be achieved by detector randomization and stochastic signaling, and proves that the optimal receiver design is realized by time sharing (randomization) between at most two maximum a-posteriori probability (MAP) detectors corresponding to two deterministic signal vectors. For the downlink of a multiuser communications system, [8] performs joint optimization of signal amplitudes, detectors, and detector randomization factors in order to reduce the worst-case average probability of error. Similarly, jamming performance of average power constrained jammers can be enhanced by performing time sharing among different power levels [3], [6], [7]. In [3], the optimal time sharing strategy for a jammer that operates over a channel with a symmetric unimodal noise density is shown to correspond to on-off jamming when the average power constraint is below a certain threshold. The optimum jamming strategy that minimizes the probability of detection in the Neyman-Pearson framework is considered in [7], where it is stated that power randomization between at most two different power levels can result in the highest jamming performance over an additive noise channel with a generic PDF.

Performance enhancements via time sharing can also be realized in communication systems where the transmitter and the receiver are connected through multiple channels [3], [10]–[13]. In such a scenario, *channel switching* can be performed by transmitting over a channel during a certain period of time and switching to another channel during the next period. In [3], the optimal channel switching strategy is studied for minimizing the average probability of error over a set of channels with additive unimodal noise under an average power constraint, and it is proved that the optimum performance can be achieved via time sharing between at most two channels and power levels. An average power constrained M -ary communication system in the presence of multiple additive noise channels with generic noise PDFs is studied in [11] in the context of minimizing the average probability of error by joint optimization of channel switching, stochastic signaling, and detection strategies. It is demonstrated that the optimal strategy is to employ deterministic signaling or time sharing between at most two signal constellations over a single channel, or to perform channel switching between two channels with deterministic signaling. The benefits of channel switching are investigated in [13] for additive Gaussian noise channels under average and peak power constraints, where the objective is to maximize average channel capacity. It is proved that the optimal solution performs channel switching between at most two different channels. The study in [10] formulates the channel switching problem by incorporating channel costs associated with the usage of each channel for transmission and imposing an average cost constraint. The optimal channel switching strategy

over a set of Gaussian channels and under average power and cost constraints is shown to correspond to time sharing among at most three different channels [10].

The previous studies on the channel switching problem mainly employ the average probability of error [3], [10], [11] or the average channel capacity [13] as the objective functions, and assume that the channels are fully utilized; i.e., there always exists transmission over one of the channels and there are no idle periods. In this manuscript, the channel switching problem is investigated for maximizing the average number of correctly received symbols in the absence of the full transmission/utilization constraint. More specifically, the optimal channel switching strategies are designed over a set of flat-fading channels under average power and cost constraints for the maximization of the average number of correctly received symbols. Rather than forcing full utilization of channels (i.e., no idle periods) as in [10], a novel and more general formulation is developed for channel switching, where communication may not occur during a certain period of time, which, in some scenarios, is shown to attain a higher average number of correctly received symbols than full channel utilization. In addition, unlike the no fading assumption in [10], Rayleigh fading channels are also considered in designing the optimal channel switching strategies. Furthermore, the proposed optimal channel switching problem is studied for logarithmic cost functions, where a logarithmic relation is employed between the signal-to-noise ratio (SNR) of each channel and its utilization cost, which is in compliance with the cost functions in the literature [10], [14], [15], [16]. The main contributions and novelty of the study in this manuscript can be summarized as follows:

- The optimal channel switching problem is formulated in the presence of partial data transmission for the first time.
- It is shown that the optimum solution is achieved via channel switching either among at most three channels with full transmission or between at most two channels with partial transmission (Proposition 1).
- Theoretical results are obtained for characterizing the optimal solution for channel switching between two channels, and for comparing the performance of single channel strategies (Propositions 2 and 3).
- Sufficient conditions for the optimality of full data transmission are derived in terms of channel costs, standard deviations of channel noise, and channel fading statistics (Proposition 4). Under these conditions, partial transmission strategies are guaranteed to be not optimal, which facilitates a low complexity solution for the optimal channel switching problem.
- For logarithmic cost functions, it is shown that the partial transmission strategies are not optimal if the average power limit is higher than a certain threshold, which depends on the parameters of the worst and best channels, as well as the parameters of the cost function and the probability of correct

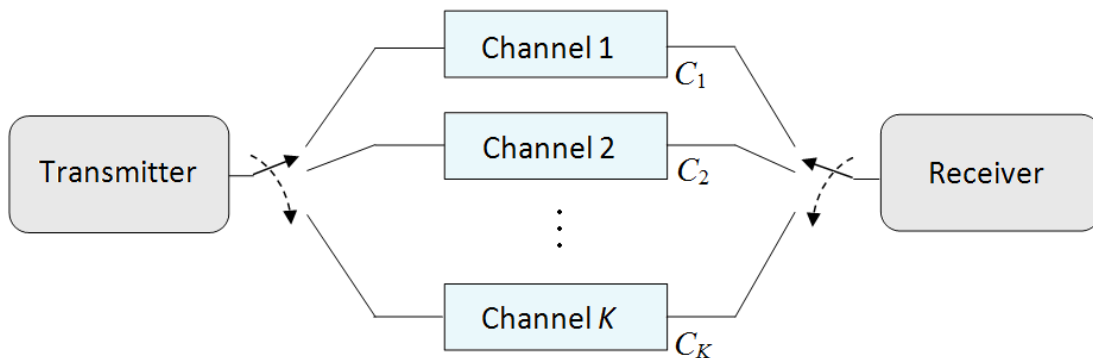


Fig. 1. Channel switching among K channels, where C_i denotes the cost of using channel i .

decision (Proposition 5).

In addition, numerical examples are presented for the demonstration of the theoretical results.

An important practical application of the channel switching problem considered in this manuscript is a cognitive radio (CR) system, in which primary users are regarded as owners of the frequency spectrum, and secondary users can utilize the frequency bands of primary users under certain conditions [10], [17]. In the spectrum trading framework proposed in [18], primary users can sell certain part of their spectrum to secondary users for the aim of revenue maximization. In that case, there can exist multiple available frequency bands (channels) with different costs for the use of secondary users, and a secondary user can perform channel switching among different available channels to improve its communication performance [10]. Similar to the effort of secondary users in CR networks for obtaining the best performance over the available bands, the aim of this study is to design optimal channel switching strategies for an arbitrary signal constellation to maximize the average number of correctly received symbols between the transmitter and the receiver under average cost and power constraints over multiple fading channels corrupted by additive white Gaussian noise.

The remainder of the manuscript is organized as follows: The system model and the problem formulation are presented in Section II. The solution of the optimal channel switching problem is characterized and theoretical results are obtained in Section III. The optimal channel switching problem is studied for logarithmic cost functions in Section IV. In Section V, numerical examples are provided, which is followed by the concluding remarks in Section VI.

II. SYSTEM MODEL AND PROBLEM FORMULATION

The channel switching problem is formulated for an M -ary communication system with K channels between the transmitter and the receiver, as shown in Fig. 1. Channel switching is performed over a

communication interval that consists of a sufficiently large number of symbols. Depending on fading conditions, the following two cases are considered:

- **Case 1:** In this case, it is assumed that slow fading occurs and the channel coefficient of each channel is fixed over the whole communication interval during which channel switching is performed. Also, the transmitter is assumed to have the channel state information (CSI) for all the channels, which can be provided in practice via feedback from the receiver.
- **Case 2:** In this case, block fading is considered, where each block consists of a number of symbols and the block duration is significantly shorter than the communication interval during which channel switching is performed. It is assumed that the channel coefficients change from block to block (independently) but the statistics of the channel coefficients are fixed for each channel in the communication interval. Also, the transmitter has the channel distribution information (CDI) for all the channels but it does not have CSI.

In both cases, channels are assumed to be frequency non-selective (i.e., flat fading) to eliminate inter-symbol interference.

To enhance system performance, the transmitter performs channel switching among K channels over time in perfect synchronization with the receiver, that is, time sharing is performed among different channels by using only one channel in a certain fraction of time [3], [11].¹ Fraction of time when transmission is performed over channel i is denoted by λ_i , which is called the *channel switching factor* for channel i . The channel switching factors satisfy $\sum_{i=1}^K \lambda_i \leq 1$ and $\lambda_i \geq 0, \forall i \in \{1 \dots K\}$. Thus, unlike the previous studies such as [3], [10], [11], it is possible to have idle periods of communications where symbol transmission/reception is not performed (in the case of $\sum_{i=1}^K \lambda_i < 1$), which can provide performance improvements in certain conditions as compared to full utilization of channels (see Proposition 1 and Section V).

Remark 1: *For the implementation of channel switching, the transmitter and the receiver are assumed to be synchronized so that the receiver knows which channel is currently in use or if it is the idle period. Then, the receiver employs a decision rule for the corresponding channel or does not perform any decision for the idle period. In practice, this assumption can be realized by employing a communications protocol that allocates the first $N_{s,1}$ symbols in the payload for channel 1, the next $N_{s,2}$ symbols in the payload for channel 2,..., and the last $N_{s,K+1}$ symbols for the idle period. The information on the number of symbols for different channels and for the idle period can be included in the header of a communications packet.*

¹It is assumed that the transmitter and the receiver have single RF units so that they can use only one channel at a given time [3], [11].

Generic M -ary modulation with an arbitrary one-dimensional or two-dimensional signal constellation² is considered for communication over each channel. The complex received signal corresponding to channel i can be expressed as

$$y = \sqrt{P_i} \alpha_i s_i^{(j)} + n_i \quad (1)$$

for $j \in \{0, 1, \dots, M-1\}$ and $i \in \{1, \dots, K\}$, where $s_i^{(0)}, s_i^{(1)}, \dots, s_i^{(M-1)}$ denote the set of (complex) transmitted signals (with unit average energy) employed for M -ary communications over channel i , P_i determines the average power of the transmitted signal for channel i , α_i is the complex fading coefficient of the i th channel, and n_i is circularly-symmetric complex Gaussian noise for channel i with mean zero and variance $2\sigma_i^2$. It is assumed that the noise components are independent across the channels and they are also independent of the fading coefficients and the transmitted signals. In addition, equally likely symbols are considered; hence, the prior probability of each symbol $s_i^{(j)}$ for $j \in \{0, 1, \dots, M-1\}$ is equal to $1/M$. It is assumed that α_i 's are perfectly estimated at the receiver. The signal-to-noise ratio (SNR) per symbol is defined as

$$\gamma_i = \frac{P_i |\alpha_i|^2}{2\sigma_i^2}. \quad (2)$$

For optimum coherent demodulation, a generic expression for the probability of symbol error corresponding to the SNR in (2) over Gaussian channels can be stated exactly or approximately (depending on the modulation type and order) as [19]

$$P_s(\gamma_i) = \eta Q(\kappa \sqrt{\gamma_i}) \quad (3)$$

where $Q(\cdot)$ denotes the Q -function, γ_i is as in (2), and η and κ are constant parameters that depend on the modulation type and order. It should be noted that the expression in (3) is exact for several types of modulations such as BPSK, BFSK and M-PAM, and it holds approximately for other types of modulation at high SNRs [19].

In Case 2, the fading coefficient α_i for channel i is modeled (over the fading blocks) as a zero-mean, circularly-symmetric complex Gaussian random variable with variance ς_i^2 , which corresponds to Rayleigh fading. Then, γ_i in (2) becomes an exponential random variable, and the average probability of symbol error can be obtained by calculating the expected value of (3) over that exponential distribution, which

²One-dimensional and two-dimensional signal constellations (e.g., PAM, PSK, QAM) are employed in almost all practical digital communications systems.

yields [19]

$$g_i(P) = \tilde{\eta} \left(1 - \sqrt{\frac{\tilde{\kappa} P}{\tilde{\kappa} P + \sigma_i^2/\zeta_i^2}} \right) \quad (4)$$

where $g_i(P)$ represents the average probability of symbol error over channel i for a power level of P , $\tilde{\eta} \triangleq \eta/2$ and $\tilde{\kappa} \triangleq \kappa^2/2$. It is noted that $g_i(P)$ is a convex and monotone decreasing function of P for $P \geq 0$.

In Case 1, the fading coefficients of the channels are fixed during the channel switching operation and they are known by the transmitter and the receiver. Since the probability of symbol error depends on α_i and σ_i only through the $|\alpha_i|^2/\sigma_i^2$ term (see (2)), $|\alpha_i| = \sqrt{2}$ can be employed for $i = 1, \dots, K$ without loss of generality, and the differences between the channel coefficients can be reflected to the σ_i^2 terms accordingly. Then, based on (2), (3) can be expressed for Case 1 as

$$g_i(P) = \eta Q \left(\kappa \sqrt{\frac{P}{\sigma_i^2}} \right) \quad (5)$$

which holds exactly for BPSK, BFSK and M-PAM modulations and approximately for other types of modulation at high SNRs [19]. For rectangular M-QAM and QPSK constellations, the exact error probability of symbol error for Case 1 can be stated as

$$g_i(P) = 1 - \left(1 - \eta Q \left(\kappa \sqrt{\frac{P}{\sigma_i^2}} \right) \right)^2 \quad (6)$$

where η and κ are determined by the modulation type.

The analysis in this study is generic to a certain extent since it employs (6) for QPSK and M-QAM modulations in Case 1, (5) for other types of modulation in Case 1, and (4) in Case 2. Although [10] considers (5) in Case 1 for scenarios with full channel utilization, there exist no studies in the literature that investigate the channel switching problem based on (6) in Case 1 (i.e., for QPSK and M-QAM modulations with slow fading) and based on (4) in Case 2 (i.e., for Rayleigh fading). In addition, the scenario with partial channel utilization is proposed for channel switching for the first time in this study.

In the considered system model in Fig. 1, there exists K channels for transmission, and each channel has a cost value, denoted by C_i for $i \in \{1, \dots, K\}$, which represents the cost of utilizing a channel per unit time [10], [18], [20]. Cost values are nonnegative, and the relation between costs of different channels is given by $C_i > C_j$ if $\sigma_i^2/\zeta_i^2 < \sigma_j^2/\zeta_j^2$ in Case 2 and if $\sigma_i^2 < \sigma_j^2$ in Case 1, $\forall j \neq i$. This is motivated by the fact that a channel with a higher ζ_i^2/σ_i^2 value, or a lower σ_i^2 value (equivalently, higher SNR) yields



Fig. 2. A communication interval of T seconds during which channel switching is performed between the transmitter and the receiver. Communication occurs during the transmission period and stops during the idle period. λ_i shows the percentage of time channel i is employed for transmission with $i \in \{1, \dots, K\}$, and λ_{K+1} denotes the percentage of the idle period in the interval. The symbol rate of the communication link is assumed to be R in symbols per second.

a lower average probability of symbol error as suggested by (4)-(6), which requires such a channel to have a higher cost [20], [21]. In the remainder of the manuscript, β_i is employed to denote the *channel parameter* of the i th channel for both Case 1 and Case 2, which is defined as

$$\beta_i \triangleq \begin{cases} \sigma_i^2, & \text{Case 1} \\ \sigma_i^2/\zeta_i^2, & \text{Case 2} \end{cases} \quad (7)$$

The channel parameters satisfy $\beta_i < \beta_j$ for $C_i > C_j$.

In this study, several assumptions/properties are stated regarding the probability of error function $g_i(\cdot)$ in order to derive generic theoretical results which are valid for various types of modulations. The following assumptions state the convexity and monotonicity properties of the error function.

Assumption 1: $g_i(P)$ is a convex function of P for $P > 0$.

Assumption 2: $g_i(P)$ is a monotone decreasing function of P/β_i , that is, $g_i(P) = \tilde{g}(P/\beta_i)$ where \tilde{g} is a monotone decreasing function.

The convexity assumption is satisfied for all three types of error functions in (4)-(6). Similarly, Assumption 2 is valid for all types of error probability functions, which are actually functions of SNR rather than power.

The aim is to perform joint optimization of channel switching factors and signal powers in order to maximize the average (expected) number of correctly received symbols per unit of time between a transmitter and a receiver under average power and cost constraints. It is assumed that the transmitter knows the fading coefficients of the channels, α_i 's, in Case 1. On the other hand, the transmitter has the knowledge of the ζ_i^2/σ_i^2 term for each channel but does not know the fading coefficient for each symbol in Case 2. A communication interval for the channel switching operation is shown in Fig. 2, where T denotes the duration of the interval in seconds and R denotes the symbol rate over a given communication link in symbols per second, which is the same for all the K channels. According to Fig. 2, for the communication interval of T seconds, the average (expected) number of correctly received symbols over the i th channel

can be expressed as

$$N_{c,i} = \lambda_i T R P_{c,i} \quad (8)$$

where λ_i is the channel switching factor for channel i and $P_{c,i}$ is the average probability of correct decision over channel i for a power level of P_i , which can be calculated as

$$P_{c,i} = 1 - g_i(P_i) \quad (9)$$

with $g_i(P_i)$ denoting the average probability of symbol error as computed in (4)-(6) for different types of modulations and cases. The expression in (8) corresponds to the average number of symbols that are correctly received during the communication interval of length $\lambda_i T$. Extending (8) to all the channels yields the average number of correctly received symbols during an interval of T seconds over all channels:

$$N_c = T R \sum_{i=1}^K \lambda_i P_{c,i}. \quad (10)$$

For a communication interval of T seconds, the objective function to maximize is given by the expression in (10). Since T and R can be assumed to be constant design/system parameters, the maximization of (10) is equivalent to maximizing

$$\bar{P}_c = \sum_{i=1}^K \lambda_i P_{c,i}. \quad (11)$$

Defining $h_i(P) \triangleq 1 - g_i(P)$ as the correct decision probability over channel i for a power level of P , (11) can be expressed based on (9) as

$$\bar{P}_c = \sum_{i=1}^K \lambda_i h_i(P_i). \quad (12)$$

Remark 2: *The weighted sum of the correct decision probabilities in (12) represents the average probability of correct decision if the sum of the channel switching factors equals to 1; otherwise, it corresponds to the “normalized” average number of correctly received symbols, normalized by TR , which is the number of symbols transmitted during a communication interval, assuming a fixed symbol duration. In the rest of the manuscript, the objective function in (12) will be referred to as the “average probability of correct decision”, regardless of whether the channel switching factors sum to 1 or not.*

The reasoning behind the choice of the average probability of correct decision instead of the average probability of error as the optimization criterion can be explained as follows: When the probability of

error metric, i.e., $\sum_{i=1}^K \lambda_i g_i(P_i)$, is employed in partial utilization of channels, the transmitter may choose not to send any symbols, that is, $\lambda_i = 0, \forall i$ (which is not possible in full utilization of channels), thus leading to zero average probability of error, which is the minimum that can be achieved. On the other hand, for the probability of correct decision metric in partial utilization, if no symbol transmission occurs during a certain period of time, the average probability of correct decision, $\sum_{i=1}^K \lambda_i h_i(P_i)$, turns out to be zero during that period, which is undesirable. Hence, the average probability of correct decision, as opposed to the average probability of error, as the optimization criterion, forces the transmitter to exploit the communication channels as efficiently as possible. (In this context, the term *goodput* can be used to replace the average probability of correct decision (with appropriate scaling) when it refers to the ratio of the total number of correctly received symbols to the total transmission time at a system level without taking into account encoding/decoding, the packet-by-packet transmission scheme and the layered concept of networking.)

It is noted that for any two channels, the one with a higher cost always results in a higher probability of correct decision for the same power level; that is, if $C_i > C_j$ (which implies $\beta_i < \beta_j$), then $h_i(P) > h_j(P)$ for all $P > 0$ (cf. (4)-(7)). Several constraints must be imposed while maximizing the average probability of correct decision in order for the channel switching strategies to be applicable in practical settings. Namely, there exists an average power constraint, which can be stated as $\sum_{i=1}^K \lambda_i P_i \leq A_p$, where A_p represents the average power limit. Also, an average transmission cost constraint can be expressed as $\sum_{i=1}^K \lambda_i C_i \leq A_c$, where A_c denotes the average cost limit [10]. Then, the following optimization problem is proposed:

$$\begin{aligned}
 & \max_{\{\lambda_i, P_i\}_{i=1}^K} \sum_{i=1}^K \lambda_i h_i(P_i) \\
 & \text{subject to } \sum_{i=1}^K \lambda_i P_i \leq A_p, \quad \sum_{i=1}^K \lambda_i C_i \leq A_c, \\
 & \sum_{i=1}^K \lambda_i \leq 1, \quad \lambda_i \geq 0, \quad \forall i \in \{1 \dots K\}.
 \end{aligned} \tag{13}$$

The optimization problem in (13) searches over both *full transmission* strategies (i.e., $\sum_{i=1}^K \lambda_i = 1$) and *partial transmission* strategies (i.e., $\sum_{i=1}^K \lambda_i < 1$) in order to achieve the maximum probability of correct symbol decision over available channels under average power and cost constraints. As investigated in the remainder of the study, the partial transmission strategy may yield higher average probabilities of correct decision in certain scenarios than the full transmission strategy and can be the solution of the

optimization problem in (13). In such scenarios, no transmission during a certain period of time facilitates a more efficient usage of the cost budget. Hence, by generalizing the concept of channel switching to scenarios with possible idle periods, the average number of correctly received symbols can be improved in a communication system that is subject to average cost and power constraints. In fact, this improvement can be achieved without any significant complexity increase compared to the channel switching systems in the literature [10]. In addition, the average number of correctly received symbols can be considered as an important parameter in practical systems.

For the theoretical analyses, it is assumed without loss of generality that the channel parameter β_i is distinct for each channel. This is based on the fact that if there are multiple channels with the same channel parameter, channel switching between such channels can never increase the average probability of correct decision compared to employing only one of them at the same average power for the total duration of time, which is due to the concavity of the correct decision probability expressions, $h_i(\cdot)$. For this reason, the problem formulation that considers only the channels with distinct channel parameters is sufficient to obtain the overall optimal solution.

III. OPTIMAL CHANNEL SWITCHING – GENERAL ANALYSIS

In this section, the optimal channel switching problem in (13) is examined in detail. In particular, the problem in (13) is reduced to a simpler equivalent form and the optimal strategies are obtained based on low-complexity calculations. The assumption made about the ordering of channel costs without loss of generality is that the cost values satisfy $C_1 > C_2 > \dots > C_K$, thus the channel parameters are ordered as $\beta_1 < \beta_2 < \dots < \beta_K$. In this case, the probability of correct decision functions satisfy $h_1(P) > h_2(P) > \dots > h_K(P)$ for all $P \geq 0$.

Based on the ordering of the channel costs, it is clear that if $A_c \geq C_1$, the optimal solution of (13) is to transmit over channel 1 exclusively with power A_p . In other words, since transmission over channel 1 results in the highest probability of correct decision among all the channels, the optimal approach becomes the use of the best channel (channel 1) all the time at the maximum power limit when the cost budget allows it.

Since (13) can easily be solved for $A_c \geq C_1$, the case of $A_c < C_1$ is considered in the remainder of the study. It is straightforward to show that the solution of the problem in (13) always satisfies the average power constraint with equality since $h_i(P)$ is a monotone increasing function of P for all $i \in \{1, \dots, K\}$. Mathematically speaking, if $\{\lambda_i^*, P_i^*\}_{i=1}^K$ denotes the solution of the optimization problem in (13), then $\sum_{i=1}^K \lambda_i^* P_i^* = A_p$. Furthermore, based on a similar approach to the proof of Proposition 1 in [10] with

slight modifications to consider the partial transmission strategies, it can be inferred that the optimal channel switching solution operates at the average cost limit, that is, $\sum_{i=1}^K \lambda_i^* C_i^* = A_c$. Hence, an optimal channel switching strategy must utilize all the available average power and average cost for $A_c < C_1$. Therefore, the optimization problem in (13) can be solved by considering equality constraints (instead of inequality constraints) for the average power and average cost, which provides an important reduction in computational complexity.

The following remark is presented to reveal the reasoning behind the use of partial transmission.

Remark 3: *Partial data transmission could not be an optimal strategy if the average cost constraint did not exist in the optimization problem in (13); that is, the optimal solution of (13) satisfies $\sum_{i=1}^K \lambda_i = 1$ in the absence of the cost constraint.*

Proof: The proof can be obtained by contradiction. Let $\{\lambda_i^*, P_i^*\}_{i=1}^K$ denote the solution of the optimization problem in (13). Suppose that the average cost constraint does not exist and the optimal solution satisfies $\sum_{i=1}^K \lambda_i^* < 1$. Define the idle period as $\lambda_{K+1}^* \triangleq 1 - \sum_{i=1}^K \lambda_i^*$. Then, following relations can be established:

$$\sum_{i=1}^K \lambda_i^* h_i(P_i^*) < \sum_{i=1}^K \lambda_i^* h_1(P_i^*) \quad (14)$$

$$< \sum_{i=1}^K \lambda_i^* h_1(P_i^*) + \lambda_{K+1}^* h_1(0) \quad (15)$$

$$< h_1 \left(\sum_{i=1}^K \lambda_i^* P_i^* + \lambda_{K+1}^* 0 \right) \quad (16)$$

$$= h_1 \left(\sum_{i=1}^K \lambda_i^* P_i^* \right) \quad (17)$$

where the first inequality follows from the fact that $h_1(P) > h_i(P)$, $\forall P$ (since $C_1 > C_2 > \dots > C_K$), the second inequality uses the facts that $\lambda_{K+1}^* > 0$ (due to partial transmission) and $h_1(0) = 1/M$ with M denoting the modulation order, and the third inequality is obtained from the strict concavity of h_1 . The inequality in (14)-(17), namely, $\sum_{i=1}^K \lambda_i^* h_i(P_i^*) < h_1 \left(\sum_{i=1}^K \lambda_i^* P_i^* \right)$, indicates that the optimal solution yields a lower average probability of correct decision than the solution which utilizes channel 1 exclusively (i.e. with channel switching factor 1) with the same average power $\sum_{i=1}^K \lambda_i^* P_i^*$, but operates at an average cost C_1 that satisfies $C_1 > \left(\sum_{i=1}^K \lambda_i^* \right) C_1 > \sum_{i=1}^K \lambda_i^* C_i$ since $C_1 > C_i$ for $i \in \{2 \dots K\}$. Hence, in the absence of the cost constraint, there always exists a full transmission strategy that achieves a higher average probability of correct decision than a partial transmission strategy, which implies that $\sum_{i=1}^K \lambda_i^* = 1$ must hold, leading to a contradiction. Therefore, partial transmission cannot be optimal in the absence of

the average cost constraint. ■

Remark 3 points to an important fact about the conditions under which partial transmission strategy can be applied instead of full utilization of channels. Remark 3 states that partial transmission can be reasonable only if the budget is limited. The optimal strategy for an unlimited budget is to utilize exclusively the channel with the highest cost (i.e., the highest correct decision probability) at the maximum average power. Thus, partial data transmission may be optimal when the budget should be used efficiently to maximize the probability of correct decision (e.g., instead of using a low-cost channel exclusively, it may be better to use a high-cost channel partially).

In the following proposition, it is stated that the optimal channel switching strategy, which is obtained as the solution of (13), corresponds to channel switching either among at most $\min\{K, 3\}$ channels with full transmission or between at most $\min\{K, 2\}$ channels with partial transmission.

Proposition 1: *Assume that the power levels satisfy $P_i \in [0, P_{\max}]$ for some finite P_{\max} . Then, the optimal channel switching strategy is to switch either among at most $\min\{K, 3\}$ channels with full transmission, or between at most $\min\{K, 2\}$ channels with partial transmission.*

Proof: The proof is based on Carathéodory's theorem [22], and similar arguments to those in [10], [23] can be employed. Assume that $K \geq 3$ since the statement in the proposition already holds otherwise. First, sets \mathcal{U} and \mathcal{W} are defined as

$$\mathcal{U} = \{(P, h_i(P), C_i), \forall i \in \{1, \dots, K\}, \forall P \in [0, P_{\max}]\} \cup \{(0, 0, 0)\} \quad (18)$$

$$\mathcal{W} = \left\{ \left(\sum_{i=1}^K \lambda_i P_i, \sum_{i=1}^K \lambda_i h_i(P_i), \sum_{i=1}^K \lambda_i C_i \right), \forall \lambda_i \geq 0, \sum_{i=1}^K \lambda_i \leq 1, \forall P_i \in [0, P_{\max}] \right\}. \quad (19)$$

From (19), it is observed that the solution of (13) must be an element of \mathcal{W} . Also, it can be concluded that \mathcal{W} is a subset of the convex hull of \mathcal{U} according to the definitions in (18) and (19). In addition, due to the aim of maximization in (13), the solution can be shown to correspond to an element of \mathcal{W} that lies on the boundary of the convex hull of \mathcal{U} based on similar arguments to those in [10], [23]. Then, Carathéodory's theorem states that any element of on the boundary of the convex hull of \mathcal{U} (including the solution of (13)) can be expressed as the convex combination of at most $\dim(\mathcal{U}) = 3$ elements in \mathcal{U} [22]. Therefore, the solution of (13) corresponds to channel switching either (i) among at most 3 channels with full transmission if $(0, 0, 0)$ is not one of the elements of \mathcal{U} employed in the convex combination for the solution, or (ii) between at most 2 channels with partial transmission otherwise. ■

Based on Proposition 1, an optimal channel switching solution corresponds to one of the following

strategies: Partial/full transmission over a single channel, partial/full transmission over two channels, and full transmission over three channels. The following sections explore the details of those strategies.

A. Single Channel Strategies

The optimal solutions for full and partial transmission over a single channel are investigated in this section.

Strategy 1P – Partial Transmission over a Single Channel: In this case, one of the channels is employed partially; that is, a single channel is used during the busy period, and an idle period exists, as well. A partial transmission strategy that employs a single channel with a cost smaller than A_c cannot be optimal (i.e., the solution of (13)) since the optimal solution must operate at the average cost limit A_c , as discussed above (the third paragraph of Section III).

In some cases, the optimal channel switching strategy corresponds to Strategy 1P. In those scenarios, the optimal solution must be searched among the channels with costs higher than A_c . Let $\mathcal{S}_g \triangleq \{l \in \{1, \dots, K\} : C_l > A_c\}$. Assume that channel $i \in \mathcal{S}_g$ is employed with channel switching factor λ_i and power P_i . Then, $\lambda_i P_i = A_p$ and $\lambda_i C_i = A_c$. Therefore, the optimal solution for channel i is obtained as $\lambda_i^* = A_c/C_i$ and $P_i^* = A_p C_i/A_c$. Hence, the average probability of correct decision is given by

$$\lambda_i^* h_i(P_i^*) = \frac{A_c}{C_i} h_i\left(A_p \frac{C_i}{A_c}\right) \quad (20)$$

and the channel that yields the optimal solution under Strategy 1P is obtained as

$$i^* = \arg \max_{i \in \mathcal{S}_g} \frac{A_c}{C_i} h_i\left(A_p \frac{C_i}{A_c}\right). \quad (21)$$

Strategy 1F – Full Transmission over a Single Channel: In this case, one of the channels is employed all the time. This strategy may be the optimal channel switching strategy if there exists a channel with cost A_c since otherwise the average cost cannot be equal to A_c .

B. Two-Channel Strategies

There exist two strategies for channel switching between two channels: Partial transmission over two channels and full transmission over two channels.

Strategy 2P – Channel Switching between Two Channels with Partial Transmission: In this strategy, channel switching is performed between two different channels and the sum of channel switching factors is smaller than 1, i.e., there exists an idle period with no data transmission. Let channel i and channel

j denote the channels employed in this strategy. Then, the problem in (13) can be formulated under Strategy 2P as

$$\begin{aligned}
& \max_{\lambda_i, \lambda_j, P_i, P_j} \lambda_i h_i(P_i) + \lambda_j h_j(P_j) \\
& \text{subject to } \lambda_i P_i + \lambda_j P_j = A_p, \\
& \lambda_i C_i + \lambda_j C_j = A_c, \\
& \lambda_i + \lambda_j < 1, \quad \lambda_i, \lambda_j \in [0, 1).
\end{aligned} \tag{22}$$

It is noted that Strategy 1P is covered as a special case of Strategy 2P. It is observed from the average cost constraint in (22) that, for the optimal channel switching between two channels, at least one of the channels should have a cost greater than A_c . Therefore, in order to obtain the optimal solution for Strategy 2P, the problem in (22) should be solved for $K_g(K-1)$ channel pairs, where K_g is the number of channels the costs of which are greater than A_c and K is the total number of channels.

Based on the argument in the previous paragraph, assume, without loss of generality, that $C_i > C_j$ for the problem in (22). From the average power and cost constraints in (22), the optimal value of λ_j and P_j can be expressed in terms of the optimal values of λ_i and P_i as $\lambda_j = (A_c - \lambda_i C_i)/C_j$ and $P_j = (A_p - \lambda_i P_i)/\lambda_j$. Therefore, the optimization problem in (22) can be simplified significantly by optimizing over two variables instead of four variables based on the two equality constraints. The optimization problem in (22) can then be expressed as follows:

$$\max_{\substack{\lambda_i \in [0, A_c/C_i] \\ P_i \in [0, A_p/\lambda_i]}} \lambda_i h_i(P_i) + \lambda_j h_j \left(\frac{A_p - \lambda_i P_i}{\lambda_j} \right) \tag{23}$$

where $\lambda_j = (A_c - \lambda_i C_i)/C_j$ and the constraints for λ_i and P_i are obtained from the relations $\lambda_i C_i + \lambda_j C_j = A_c$ and $\lambda_i P_i + \lambda_j P_j = A_p$. From (23), it is observed that the optimal solution for Strategy 2P requires a search over a two-dimensional space only (for each possible channel pair). This two-dimensional search must be executed by first determining a value for λ_i and then finding the optimal P_i for the current λ_i value since the search interval for P_i depends on the value of λ_i . Finally, the maximum for all those (λ_i, P_i) pairs is calculated and the pair that yields the maximum value of the objective function is determined to be optimal.

Strategy 2F – Channel Switching between Two Channels with Full Transmission: In this strategy, channel switching is performed between two different channels and the sum of channel switching factors

is equal to 1. The formulation of the problem in (13) under Strategy 2F is as follows:

$$\begin{aligned}
& \max_{\lambda_i, \lambda_j, P_i, P_j} \lambda_i h_i(P_i) + \lambda_j h_j(P_j) \\
& \text{subject to } \lambda_i P_i + \lambda_j P_j = A_p, \\
& \lambda_i C_i + \lambda_j C_j = A_c, \\
& \lambda_i + \lambda_j = 1, \quad \lambda_i, \lambda_j \in [0, 1].
\end{aligned} \tag{24}$$

In this case, the optimization can be performed over a single variable since the sum of channel switching factors forms a new equality. Strategy 2F reduces to Strategy 1F if one of the channel switching factors is equal to 1.

The following assumptions are made about the error functions $g_i(\cdot)$ to provide a basis for the next proposition.

Definition 1: Assume $C_i > C_j$ and let P_{ij} be defined as the solution to equation $g'_i(x) - g'_j(x) = 0$, that is, $g'_i(P_{ij}) = g'_j(P_{ij})$.

Assumption 3:

(i) $g'_i(P) > g'_j(P)$ if $P > P_{ij}$.

(ii) $g'_i(P_{ij}) = g'_j(P_{ij})$.

(iii) $g'_i(P) < g'_j(P)$ if $0 < P < P_{ij}$.

(iv) $g'_i(0) = -\infty \quad \forall i \in \{1, \dots, K\}$.

Assumption 4: If $g'_i(P_i) = g'_j(P_j)$ is satisfied, then

(i) $\frac{P_j}{P_i} < \frac{\beta_j}{\beta_i}$ for $P_j > P_i > P_{ij}$

(ii) $\frac{P_i}{P_j} < \frac{\beta_j}{\beta_i}$ for $P_j < P_i < P_{ij}$

Assumption 3 is valid for all types of modulations whose error probability expressions are given by (4)-(6). On the other hand, Assumption 4 is satisfied for the error functions in (4) and (5), but not satisfied for (6).

The following proposition derives upper and lower bounds for the optimal power levels obtained for Strategy 2P and Strategy 2F and their ratios.

Proposition 2: Suppose that Assumptions 1, 3, and 4 hold. Let the solution of the optimization problem in (13) under the two-channel strategies be denoted by $\{\lambda_i^*, P_i^*, \lambda_j^*, P_j^*\}$ and suppose that $\lambda_i^* > 0$, $\lambda_j^* > 0$, and $C_i > C_j$. Then, the optimal power levels and the channel switching factors satisfy the following relations depending on the average power limit:

(i) If $A_p = P_{ij}(\lambda_i^* + \lambda_j^*)$, then $P_i^* = P_j^* = P_{ij}$.

(ii) If $A_p > P_{ij}(\lambda_i^* + \lambda_j^*)$, then $P_j^* > P_i^* > P_{ij}$.

(iii) If $A_p < P_{ij}(\lambda_i^* + \lambda_j^*)$, then $P_j^* < P_i^* < P_{ij}$.

where P_{ij} is as in Definition 1. In addition, the ratio between the optimal power levels cannot exceed β_j/β_i ; that is,

$$\max \left\{ \frac{P_j^*}{P_i^*}, \frac{P_i^*}{P_j^*} \right\} < \frac{\beta_j}{\beta_i}. \quad (25)$$

Proof: The optimization problems in (22) and (24) can be solved together by re-writing the constraint on the sum of the channel switching factors as $\lambda_i + \lambda_j \leq 1$. In this case, the optimization problem again becomes the one in (23). Consider the first-order derivative of the objective function in (23) with respect to P_i :

$$\lambda_i \left(h'_i(P_i) - h'_j \left(\frac{A_p - \lambda_i P_i}{\lambda_j} \right) \right) \quad (26)$$

or, equivalently,

$$\lambda_i (h'_i(P_i) - h'_j(P_j)). \quad (27)$$

Due to Assumption 1, $g'_i(P)$ is an increasing function of P and $h'_i(P) = -g'_i(P)$ is a decreasing function of P for $P > 0$. Hence, the expression in (26) is a decreasing function of P_i for $P_i > 0$, starting from ∞ at $P_i = 0$ and decreasing monotonically towards $-\infty$ at $P_i = A_p/\lambda_i$ (due to Assumption 3). Therefore, given the value of λ_i , there is a unique maximizer P_i^* for the optimization problem in (23), which corresponds to the point at which the first-order derivative is zero. Equating the first-order derivative in (27) to zero and setting $h'_i(P) = -g'_i(P)$ yields the following necessary and sufficient condition for the optimal solution of (23):

$$g'_i(P_i) = g'_j(P_j). \quad (28)$$

If $A_p = P_{ij}(\lambda_i + \lambda_j)$, it is obvious, by using the definition of P_{ij} in Definition 1 and Assumption 3, that $P_i = P_j = P_{ij}$ satisfies the condition in (28). Since the solution of (28) is unique, the optimal solution of (23) is obtained as $P_i^* = P_j^* = P_{ij}$, as stated in the first part of Proposition 2.

In order to prove the second part of the proposition, it is first observed that the first-order derivative in (26) is a monotone decreasing function of A_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_p increases. Since the first-order derivative becomes zero at $P_i = P_{ij}$ when $A_p = P_{ij}(\lambda_i + \lambda_j)$ (as proved in the first part), the first-order derivative becomes zero at a value larger than P_{ij} when $A_p > P_{ij}(\lambda_i + \lambda_j)$. Hence, the optimal

solution of (23) satisfies $P_i > P_{ij}$ for $A_p > P_{ij}(\lambda_i + \lambda_j)$. In addition, it is concluded from (28) that as P_i increases, the optimal value of P_j should also increase since $g'_i(P)$ is an increasing function of P for $P > 0$, as stated above. In other words, $P_i^* > P_{ij}$ also implies $P_j^* > P_{ij}$ based on the relation in (28). Next, the ordering between P_i^* and P_j^* should be determined. Due to Assumption 3 and the condition (28), $g'_j(P_j^*) = g'_i(P_i^*) > g'_i(P_i^*)$ since $P_i^* > P_{ij}$, which requires $P_j^* > P_i^*$ due to the monotone increasing property of $g'_j(P)$. Therefore, for $A_p > P_{ij}(\lambda_i + \lambda_j)$, the inequality $P_j^* > P_i^* > P_{ij}$ is obtained. Similarly, for $A_p < P_{ij}(\lambda_i + \lambda_j)$, the inequality $P_j^* < P_i^* < P_{ij}$ can be obtained.

Since the optimal power levels satisfy the condition in (28), the final statement in the proposition can be reached by using Assumption 4 and the results obtained in the previous parts of the proposition related to the ordering of the optimal power levels. For $A_p = P_{ij}(\lambda_i + \lambda_j)$, $P_i^*/P_j^* = 1$ as stated in the first part of the proposition. Overall, the ratio between the optimal power levels is upper bounded by β_j/β_i for any value of A_p , as stated in the proposition. \blacksquare

The search space of the optimization for Strategy 2P can be reduced based on Proposition 2 as follows. For each $\lambda_i \in (0, A_c/C_i)$, λ_j is calculated as $\lambda_j = (A_c - \lambda_i C_i)/C_j$ and the optimal power levels are obtained as follows:

- If $A_p = A_{ij}(\lambda_i + \lambda_j)$, the optimal solution is given by

$$P_i^* = P_j^* = P_{ij}. \quad (29)$$

- If $A_p > A_{ij}(\lambda_i + \lambda_j)$, the optimization problem in (23) is solved for

$$P_i \in \left(\max \{P_{ij}, \beta_i A_p / \beta_j\}, \min \{A_p / \lambda_i, \beta_j A_p / (\beta_i \lambda_j)\} \right) \quad (30)$$

which is obtained from (25) and the relation in the second part of the proposition.

- If $A_p < A_{ij}(\lambda_i + \lambda_j)$, the problem in (23) is solved for

$$P_i \in \left(A_p, \min \{P_{ij}, A_p / \lambda_i, \beta_j A_p / (\beta_i \lambda_j)\} \right) \quad (31)$$

which is obtained from (25) and the relation in the third part of the proposition.

Applying the procedure above, the optimal value of P_i and the corresponding value of the objective function in (23) can be determined for a given $\lambda_i \in (0, A_c/C_i)$. The optimal value of λ_i , denoted by λ_i^* , is the value that yields the maximum of the objective function in (23) and the corresponding value of P_i^* gives the optimal value of P_i . Once the optimal pair $\{\lambda_i^*, P_i^*\}$ is obtained, the optimal values of λ_j and P_j are calculated as $\lambda_j^* = (A_c - \lambda_i^* C_i)/C_j$ and $P_j^* = (A_p - \lambda_i^* P_i^*)/\lambda_j^*$, respectively.

C. Three-Channel Strategies

Based on Proposition 1, there exists only one strategy for channel switching among three channels.

Strategy 3 – Channel Switching among Three Channels: In this strategy, transmission is performed by switching among three different channels and the channel switching factors add up to 1 (i.e., full transmission). The formulation of the optimization problem in (13) under Strategy 3 is given as follows:

$$\begin{aligned}
 & \max_{\lambda_i, \lambda_j, \lambda_k, P_i, P_j, P_k} \lambda_i h_i(P_i) + \lambda_j h_j(P_j) + \lambda_k h_k(P_k) \\
 & \text{subject to } \lambda_i P_i + \lambda_j P_j + \lambda_k P_k = A_p, \\
 & \lambda_i C_i + \lambda_j C_j + \lambda_k C_k = A_c, \\
 & \lambda_i + \lambda_j + \lambda_k = 1, \quad \lambda_i, \lambda_j, \lambda_k \in [0, 1]
 \end{aligned} \tag{32}$$

where i , j and k are the employed channels. Due to the average cost constraint in (32), at least one of the channels should have a cost greater than A_c and at least one of them should have a cost smaller than A_c . Thus, the optimization problem in (32) must be solved for $K_g K_s (K - 2)$ channel triples and the triple that yields the highest average probability of correct decision is determined to be optimal. Here, K_g and K_s denote, respectively, the number of channels with costs greater than A_c and smaller than A_c . It is observed from (32) that the optimization can be performed over three variables instead of six variables by imposing the three equality constraints.

It should be noted that Strategy 1F and Strategy 2F are covered as special cases of Strategy 3. Therefore, in order to obtain the optimal solution in case of full data transmission, the optimization problem in (32) can be solved first, which reveals the type of the transmission strategy to be applied.

D. Comparison of Channel Switching Strategies

In this section, theoretical results are obtained for comparing the performance of the different strategies. First, the single channel strategies are examined in terms of the probability of correct decision to put forward a suboptimal solution when only a single channel is employed. Conditions are investigated under which full or partial transmission over a single channel (Strategy 1P or Strategy 1F) is optimal. Strategy 1F can be optimal only if there exists a channel with cost A_c ; otherwise, the cost budget would be used partially and the solution would not be optimal. Hence, the comparison of Strategy 1F versus Strategy 1P as candidates for the overall optimal solution can be made by

$$h_{i^*}(A_p) \geq \max_{j \in \mathcal{S}_g} \frac{A_c}{C_j} h_j \left(A_p \frac{C_j}{A_c} \right) \tag{33}$$

where $\mathcal{S}_g \triangleq \{l \in \{1, \dots, K\} : C_l > A_c\}$, i^* is the index of the channel satisfying $C_{i^*} = A_c$, and the left-hand-side and the right-hand-side of (33) represent the average probabilities of correct decision corresponding to Strategy 1F and Strategy 1P, respectively. For partial transmission over channel j with $j \notin \mathcal{S}_g$ (i.e., $C_j < A_c$), the average probability of correct decision can be expressed as

$$\lambda_j h_j(P_j) < \lambda_j h_j(P_j) + (1 - \lambda_j) h_j(0) \quad (34)$$

$$< h_j(\lambda_j P_j + (1 - \lambda_j) 0) \quad (35)$$

$$\leq h_j(A_p) \quad (36)$$

$$< h_{i^*}(A_p) \quad (37)$$

where the second inequality results from the concavity of h_j , the third inequality is due to the average power constraint, and the last inequality is obtained from $C_j < A_c = C_{i^*}$. The inequalities in (34)-(37), namely, $\lambda_j h_j(P_j) < h_{i^*}(A_p)$, demonstrate why channels not in \mathcal{S}_g need not be included in (33).

The following proposition presents a sufficient condition for deciding between two channels in terms of optimality under the single channel strategies, Strategy 1P or Strategy 1F.

Proposition 3: *Suppose that Assumptions 1 and 2 hold. Consider a channel pair (i, j) such that $C_i > C_j \geq A_c$. If the condition*

$$\frac{\beta_j}{\beta_i} \leq \frac{1 - \frac{C_j}{C_i}}{g_i(0)} + \frac{C_j}{C_i} \quad (38)$$

is satisfied, then partial/full transmission over channel j achieves a higher probability of correct decision than partial transmission over channel i .

Proof: Assume that the inequality in (38) is satisfied. First, the inequality in (38) is stated as

$$\frac{A_p C_j / A_c}{\beta_i \left(\frac{1 - \frac{C_j}{C_i}}{g_i(0)} + \frac{C_j}{C_i} \right)} \leq \frac{A_p C_j / A_c}{\beta_j}. \quad (39)$$

Since \tilde{g} is a monotone decreasing function (see Assumption 2), (39) leads to the following inequality:

$$\tilde{g} \left(\frac{A_p C_j / A_c}{\beta_i \left(\frac{1 - \frac{C_j}{C_i}}{g_i(0)} + \frac{C_j}{C_i} \right)} \right) \geq \tilde{g} \left(\frac{A_p C_j / A_c}{\beta_j} \right). \quad (40)$$

Then, based on the relation $g_i(P) = \tilde{g}(P/\beta_i)$ stated in Assumption 2, the following inequality is obtained:

$$g_i \left(\frac{A_p C_j / A_c}{\frac{1 - \frac{C_j}{C_i}}{g_i(0)} + \frac{C_j}{C_i}} \right) \geq g_j \left(\frac{A_p C_j}{A_c} \right). \quad (41)$$

Since $g_i(P)$ represents a value of probability, $g_i(0) \leq 1$ is always satisfied. Hence,

$$\frac{1 - \frac{C_j}{C_i}}{g_i(0)} + \frac{C_j}{C_i} \geq 1 \quad (42)$$

is obtained, which, together with (41), leads to the following relation:

$$\left(\frac{1 - \frac{C_j}{C_i}}{g_i(0)} + \frac{C_j}{C_i} \right) g_i \left(\frac{A_p C_j / A_c}{\frac{1 - \frac{C_j}{C_i}}{g_i(0)} + \frac{C_j}{C_i}} \right) \geq g_j \left(\frac{A_p C_j}{A_c} \right). \quad (43)$$

Exploiting the convexity of the g_i function for non-negative arguments, the following inequality is obtained:

$$\left(\frac{1 - \frac{C_j}{C_i}}{g_i(0)} \right) g_i(0) + \frac{C_j}{C_i} g_i \left(\frac{A_p C_i}{A_c} \right) > g_j \left(\frac{A_p C_j}{A_c} \right). \quad (44)$$

Then, rearranging the terms in (44) results in

$$\frac{C_j}{C_i} \left(1 - g_i \left(\frac{A_p C_i}{A_c} \right) \right) < 1 - g_j \left(\frac{A_p C_j}{A_c} \right). \quad (45)$$

Finally, inserting the definition of the function h_i ,

$$\frac{C_j}{C_i} h_i \left(A_p \frac{C_i}{A_c} \right) < h_j \left(A_p \frac{C_j}{A_c} \right) \quad (46)$$

$$\frac{A_c}{C_i} h_i \left(A_p \frac{C_i}{A_c} \right) < \frac{A_c}{C_j} h_j \left(A_p \frac{C_j}{A_c} \right) \quad (47)$$

are obtained, which demonstrate that the probability of correct decision attained by partial/full (partial if $C_j > A_c$ and full if $C_j = A_c$) transmission over channel j is higher than that over channel i if the condition in (38) is satisfied. ■

Note that $g_i(0)$ in (38) is equal to $\frac{M-1}{M}$, where M is the modulation order. A simple condition which does not involve the calculations of g_i (or, h_i) is provided in Proposition 3 as compared to (21) and (33) for determining whether Strategy 1F achieves a higher probability of correct decision than Strategy 1P if there exists a channel with cost A_c , and for deciding between two channels in terms of probability of correct decision under Strategy 1P otherwise. The inverse of Proposition 3 may not be valid as it puts forward only a sufficient condition for deciding between the two cases. As a reasonable approach, the condition in

(38) can be checked first, and if it is not satisfied, then the necessary and sufficient conditions in (20) and (33) can be examined. Proposition 3 can especially be useful for applications where transmitters do not have sufficient time or capability (due to hardware, complexity, etc. limitations) to switch among different channels, thereby constraining themselves to use only a single channel.

One of the main results in this study is the following proposition, which presents a sufficient condition under which partial transmission (Strategy 2P or Strategy 1P) cannot be optimal. That is, it is guaranteed under the stated conditions that partial transmission over a single channel or two channels is outperformed by a full transmission strategy.

Proposition 4: *Suppose that Assumptions 1 and 2 hold, and that there exists a channel $k \in \{1, \dots, K\}$ satisfying the conditions*

$$C_k \leq A_c \text{ and } \frac{\beta_k}{\beta_i} \leq \frac{1 - \frac{C_k}{C_i}}{g_i(0)} + \frac{C_k}{C_i}, \quad \forall i \in \mathcal{S}_g = \{l : C_l > A_c\}. \quad (48)$$

Then, partial data transmission is not optimal.

Proof: There exist two possible strategies for partial data transmission, as discussed previously. First, consider Strategy 2P, where partial transmission is performed via channel switching between two channels. Let i and j denote the channels employed for partial transmission over two channels, and $\{\lambda_i^*, \lambda_j^*, P_i^*, P_j^*\}$ represent the solution of the optimization problem in (22) for the channel pair (i, j) . Since $\lambda_i^* C_i + \lambda_j^* C_j = A_c$ from Proposition 1 and $\lambda_i^* + \lambda_j^* < 1$, at least one of the channels should have a cost higher than A_c . Assume without loss of generality that $C_i > C_j$; then, $C_i > A_c$ is obtained. Let k be the channel that satisfies the conditions in (48). Define $\lambda_{K+1}^* \triangleq 1 - (\lambda_i^* + \lambda_j^*)$. If $\lambda_i^* C_i \geq (\lambda_i^* + \lambda_{K+1}^*) C_k$, then the alternative solution that employs the channel triple (i, j, k) with channel switching factors $\{(\lambda_i^* - \gamma), \lambda_j^*, \nu_k\}$ and powers $\{P_i^*, P_j^*, P_k\}$, respectively, results in a higher average probability of correct decision than the two-channel partial transmission solution, denoted by $\{\lambda_i^*, \lambda_j^*, P_i^*, P_j^*\}$, over channels i and j :

$$\lambda_i^* h_i(P_i^*) + \lambda_j^* h_j(P_j^*) < (\lambda_i^* - \gamma) h_i(P_i^*) + \lambda_j^* h_j(P_j^*) + \nu_k h_k(P_k) \quad (49)$$

where ν_k is defined as $\nu_k \triangleq \gamma + \lambda_{K+1}^*$ so that the channel switching factors in the alternative solution sum to 1. The alternative solution corresponds to Strategy 3, which is the three-channel full transmission strategy. The inequality in (49) can be expressed as

$$\gamma h_i(P_i^*) < \nu_k h_k(P_k). \quad (50)$$

Based on the power and cost constraint equations for the optimal solutions according to Proposition 1,

the following relations are obtained for the two cases:

$$\lambda_i^* P_i^* + \lambda_j^* P_j^* = (\lambda_i^* - \gamma) P_i^* + \lambda_j^* P_j^* + \nu_k P_k = A_p \quad (51)$$

$$\lambda_i^* C_i + \lambda_j^* C_j = (\lambda_i^* - \gamma) C_i + \lambda_j^* C_j + \nu_k C_k = A_c \quad (52)$$

From (52), $\gamma C_i = \nu_k C_k$ is obtained. Hence, γ is computed as $\gamma = \frac{\lambda_{K+1}^* C_k}{C_i - C_k}$, which is greater than zero since $C_i > A_c \geq C_k$. From (51), $\gamma P_i^* = \nu_k P_k$ is obtained and P_k is computed as $P_k = P_i^* C_k / C_i$. Then, the inequality in (50) can be re-written as

$$\gamma(1 - g_i(P_i^*)) < (\gamma + \lambda_{K+1}^*) (1 - g_k(P_k)) \quad (53)$$

which, after some manipulation and inserting the values of γ and P_k , reduces to

$$\left(1 - \frac{C_k}{C_i}\right) + \frac{C_k}{C_i} g_i(P_i^*) > g_k\left(\frac{P_i^* C_k}{C_i}\right). \quad (54)$$

Exploiting the convexity of the g_i function, the left-hand side of (54) can be bounded from below:

$$\left(\frac{1 - \frac{C_k}{C_i}}{g_i(0)}\right) g_i(0) + \frac{C_k}{C_i} g_i(P_i^*) > \left(\frac{1 - \frac{C_k}{C_i}}{g_i(0)} + \frac{C_k}{C_i}\right) g_i\left(\frac{C_k P_i^* / C_i}{\frac{1 - \frac{C_k}{C_i}}{g_i(0)} + \frac{C_k}{C_i}}\right). \quad (55)$$

If the lower bound in (55) is greater than or equal to the right-hand side of the inequality in (54), then the proof of the inequality in (49) is completed:

$$\left(\frac{1 - \frac{C_k}{C_i}}{g_i(0)} + \frac{C_k}{C_i}\right) g_i\left(\frac{C_k P_i^* / C_i}{\frac{1 - \frac{C_k}{C_i}}{g_i(0)} + \frac{C_k}{C_i}}\right) \geq g_k\left(\frac{P_i^* C_k}{C_i}\right). \quad (56)$$

Since $g_i(0) \leq 1$, the multiplying factor before g_i on the left-hand side of (56) is greater than or equal to 1. Then, based on Assumption 2, it is sufficient for the proof of the inequality in (49) that the following inequality is satisfied:

$$\tilde{g}\left(\frac{C_k P_i^* / C_i}{\beta_i \left(\frac{1 - \frac{C_k}{C_i}}{g_i(0)} + \frac{C_k}{C_i}\right)}\right) \geq \tilde{g}\left(\frac{P_i^* C_k}{\beta_k C_i}\right). \quad (57)$$

From the condition in (48), the argument of the \tilde{g} function on the left-hand side of (57) is readily found to be smaller than that on the right-hand side. Since the \tilde{g} function is monotone decreasing, the expression on the left-hand side is greater than that on the right-hand side. Hence, the inequality in (57) is satisfied, which completes the proof of (49). As for the legitimacy of the alternative solution whose probability of

correct decision expression is the right-hand side of (49), $\lambda_i^* \geq \gamma$ must be satisfied. Inserting the expression for γ , this inequality becomes equivalent to $\lambda_i^* C_i \geq (\lambda_i^* + \lambda_{K+1}^*) C_k$, which was the initial assumption.

In order to prove the second part of the argument related to Strategy 2P, assume that $\lambda_i^* C_i < (\lambda_i^* + \lambda_{K+1}^*) C_k$. In this case, the alternative solution proposed in the first part is not valid. Consider an alternative two-channel solution with channel switching factors $\{\lambda_j^*, \lambda_k\}$ and power levels $\{P_j^*, P_k\}$. It will be shown that this alternative solution outperforms the original two-channel partial solution:

$$\lambda_i^* h_i(P_i^*) + \lambda_j^* h_j(P_j^*) < \lambda_j^* h_j(P_j^*) + \lambda_k h_k(P_k). \quad (58)$$

From the power and cost constraint equations as in (51) and (52), the channel switching factor and the power level for channel k are obtained as $\lambda_k = \lambda_i^* C_i / C_k$ and $P_k = P_i^* C_k / C_i$, respectively. Inserting these into (58) yields the inequality

$$h_i(P_i^*) < \frac{C_i}{C_k} h_k(P_i^* \frac{C_k}{C_i}) \quad (59)$$

which is the same as the inequality in (54). Hence, the inequality in (58) is proved. Since $\lambda_k = \lambda_i^* \frac{C_i}{C_k} < \lambda_i^* + \lambda_{K+1}^*$, $\lambda_j^* + \lambda_k < \lambda_j^* + \lambda_i^* + \lambda_{K+1}^* = 1$, which shows that the alternative two-channel solution is also partial. Define $\tilde{\lambda}_{K+1} \triangleq 1 - (\lambda_j^* + \lambda_k)$, and note that $C_j \leq A_c$ leads to the contradiction $\lambda_j^* C_j + \lambda_k C_k \leq A_c (\lambda_j^* + \lambda_k) < A_c$. In addition, if $\lambda_j^* C_j < (\lambda_j^* + \tilde{\lambda}_{K+1}) C_k$, the inequality $\lambda_j^* C_j + \lambda_k C_k < (\lambda_j^* + \tilde{\lambda}_{K+1} + \lambda_k) C_k = C_k \leq A_c$ is obtained, which is a contradiction. Hence, $C_j > A_c$ and $\lambda_j^* C_j \geq (\lambda_j^* + \tilde{\lambda}_{K+1}) C_k$ are satisfied. Then, based on the proof of the inequality in (49), following inequality can be written by considering a channel pair (j, k) instead of (i, j) :

$$\lambda_j^* h_j(P_j^*) + \lambda_k h_k(P_k) < (\lambda_j^* - \tilde{\gamma}) h_j(P_j^*) + \lambda_k h_k(P_k) + \tilde{\nu}_k h_k(\tilde{P}_k). \quad (60)$$

Based on the concavity of the h_k function,

$$\lambda_j^* h_j(P_j^*) + \lambda_k h_k(P_k) < (\lambda_j^* - \tilde{\gamma}) h_j(P_j^*) + (\lambda_k + \tilde{\nu}_k) h_k \left(\frac{\lambda_k}{\lambda_k + \tilde{\nu}_k} P_k + \frac{\tilde{\nu}_k}{\lambda_k + \tilde{\nu}_k} \tilde{P}_k \right) \quad (61)$$

is obtained. Note that the average power and the average cost of the three-channel solution (two channels being the same) in the right-hand side of (60) are the same as those of the two-channel solution in the right-hand side of (61), which indicates that the latter solution satisfies the average power and cost constraints. Combining the inequalities (58) and (61), it is demonstrated that the two-channel partial solution with channel switching factors $\{\lambda_i^*, \lambda_j^*\}$ and power levels $\{P_i^*, P_j^*\}$ achieves a lower average probability of correct decision than the two-channel full solution with channel switching factors $\{(\lambda_j^* - \tilde{\gamma}), (\lambda_k + \tilde{\nu}_k)\}$

and power levels $\{P_j^*, \frac{\lambda_k}{\lambda_k + \tilde{\nu}_k} P_k + \frac{\tilde{\nu}_k}{\lambda_k + \tilde{\nu}_k} \tilde{P}_k\}$. Therefore, a two-channel partial solution cannot be optimal if the conditions presented in Proposition 4 are satisfied.

Secondly, consider Strategy 1P, where a single channel is employed partially. Let i denote the channel employed for partial transmission over a single channel. $C_i > A_c$ must be satisfied since a partial transmission strategy that employs a single channel with a cost smaller than A_c cannot be optimal (cf. Section III-A). Let $\{\lambda_i^*, P_i^*\}$ represent the solution of the partial transmission strategy that uses channel i . Then, similar to the proof of the inequality in (49), the following inequality can be obtained:

$$\lambda_i^* h_i(P_i^*) < (\lambda_i^* - \gamma) h_i(P_i^*) + \nu_k h_k(P_k) \quad (62)$$

where ν_k is defined as $\nu_k \triangleq \gamma + \lambda_{K+1}^*$ and $\lambda_{K+1}^* \triangleq 1 - \lambda_i^*$. Here, $\lambda_i^* C_i = A_c = (\lambda_i^* + \lambda_{K+1}^*) A_c \geq (\lambda_i^* + \lambda_{K+1}^*) C_k$ leads to the inequality $\lambda_i^* C_i \geq (\lambda_i^* + \lambda_{K+1}^*) C_k$, which, together with $C_i > A_c$, completes the proof of (62). Thus, under the conditions stated in (48), given a one-channel partial solution, there always exists a two-channel full solution that attains a higher average probability of correct decision.

Overall, since both Strategy 1P and Strategy 2P (which are the possible optimal solutions for partial transmission according to Proposition 1) are outperformed by the full transmission strategies, partial transmission is not optimal under the conditions stated in the proposition. ■

Proposition 4 is highly crucial since it provides a condition that definitely removes the computational burden of solving the optimization problem in (22), which involves both Strategy 2P and Strategy 1P. Hence, it suffices to solve the optimization problem in Strategy 3 only in order to obtain the optimal solution of (13), thereby greatly reducing the computational complexity. In addition, the condition derived in Proposition 4 does not depend on the optimal power levels or channel switching factors; it depends only on the system parameters such as the channel costs, the noise variances, and the statistics of the fading coefficients. Therefore, given a set of communication channels with assigned costs and known noise and fading statistics, if the condition in (48) is satisfied, it can be stated beforehand that partial data transmission is not optimal.

IV. OPTIMAL CHANNEL SWITCHING FOR LOGARITHMIC COST FUNCTION

In this section, a suitable cost function is employed for the channels, and specific theoretical results are obtained regarding the optimality of various channel switching strategies and the characterization of the optimal channel switching solution. For the analysis in this section, Case 1 in Section II is considered. In addition, the probability of symbol error corresponding to the optimum coherent detection over channel i

is expressed as

$$g_i(P) = \eta Q \left(\kappa \sqrt{\frac{P}{\beta_i}} \right) \quad (63)$$

where $\beta_i \triangleq 2\sigma_i^2/|\alpha_i|^2$ is the channel parameter corresponding to the i th channel (cf. (7)), and η and κ are constant parameters that depend on the modulation type and order [19]. As noted in Section II, the expression in (63) holds exactly for several types of modulations such as BPSK, BFSK and M-PAM, and it is approximate for other types of modulations at high SNRs [19].

In practical systems, each communication channel can be regarded as a measurement device that has a cost related to the quality of the measurement [14], [10], [18], [20]. Hence, a cost function based on that in [14] can be adopted for defining the cost of channel i as follows:

$$C_i = \log \left(1 + \frac{b}{\beta_i} \right), \quad i \in \{1 \dots K\} \quad (64)$$

where β_i is the channel parameter for the i th channel and $b > 0$ is a constant system parameter. The main motivations behind the use of (64) are that it satisfies $C_i > C_j$ for $\beta_i < \beta_j$ (i.e., a channel with a larger SNR has a higher cost since $\beta_i = 2\sigma_i^2/|\alpha_i|^2$), and the cost of a channel converges to zero (infinity) as β_i goes to infinity (zero).

As in Section III, it is assumed that the channel parameters satisfy $\beta_1 < \beta_2 < \dots < \beta_K$ (i.e., $C_1 > C_2 > \dots > C_K$) without loss of generality. It should be noted that the channels with the same costs (channel parameters) can be considered as a single channel since switching among them does not improve the system performance due to the concavity of the probability of correct decision with respect to power (cf. (13) and (63)). In addition, $A_c < C_1$ is assumed in the remainder of this section since (13) leads to the trivial solution of using the best channel (channel 1) exclusively for $A_c \geq C_1$.

In order to facilitate theoretical analyses, the properties of the probability of correct decision should be specified first. For the cost function in (64), β is given by $\beta = b/(e^C - 1)$. Hence, from (63), the probability of correct decision can be expressed as a function of power and cost as follows:

$$h(P, C) = 1 - \eta Q \left(\kappa \sqrt{\frac{P(e^C - 1)}{b}} \right). \quad (65)$$

Then, the concavity property of $h(P, C)$ in (65) is stated in the following lemma.

Lemma 1 [10]: *Consider infinitely many channels and suppose that the channels take a continuum of cost values in the interval $[C_{\min}, C_{\max}]$ based on the cost function in (64), where $0 < C_{\min} < C_{\max} < \infty$. Then, $h(P, C)$ is a strictly concave function over set \mathcal{S}_c , which is a convex set defined as $\mathcal{S}_c \triangleq$*

$$\{(P, C) : P > b/(\kappa^2(e^C + 1)), C \in (C_{\min}, C_{\max})\}^3$$

The following lemma presents an important result that proves to be highly useful in deriving optimality conditions for the full transmission strategies.

Lemma 2: *Let $h(P, C)$ denote the probability of correct decision as a function of power and cost as defined in (65). The inequality*

$$\lambda h(P, C) < h(\lambda P, \lambda C) \quad (66)$$

is satisfied $\forall \lambda \in (0, 1)$ if

$$(P, C) \in \mathcal{S}_i \triangleq \left\{ (P, C) : P > \max \left\{ \frac{b}{\kappa^2(e^C + 1)}, \frac{bC}{v} \right\}, C \in (C_{\min}, C_{\max}) \right\} \quad (67)$$

where v is a positive constant that depends on η , as defined in (85) based on (83). In addition, \mathcal{S}_i is a convex set.

Proof: Please see Appendix A.

Lemma 2 identifies the region in the power-cost plane where the inequality $\lambda h(P, C) < h(\lambda P, \lambda C)$ is satisfied $\forall \lambda \in (0, 1)$. One of the main results in this manuscript is derived based on Lemma 2 in the following proposition, which presents a condition for the optimality of the full transmission strategies when the logarithmic cost function in (64) is employed.

Proposition 5: *Consider K channels and suppose that each channel has a cost value based on the cost function in (64). If the average power limit satisfies*

$$A_p \geq \max \left\{ \frac{b \beta_K^2}{\kappa^2 \beta_1 (2\beta_K + b)}, \frac{b \beta_K \log \left(1 + \frac{b}{\beta_1} \right)}{v \beta_1} \right\}, \quad (68)$$

then partial data transmission cannot be optimal.

Proof: Please see Appendix B.

Proposition 5 sets an upper bound on the average power limit, above which the partial transmission strategies cannot yield the optimal solution of the optimization problem in (13) in the case of the logarithmic cost functions. This simplifies the solution under the condition in (68) in Proposition 5 by making it possible to eliminate the partial transmission strategies, namely, Strategy 1P and Strategy 2P. Therefore, it is sufficient to solve the optimization problem under Strategy 3 in order to obtain the optimal channel switching strategy in such a scenario (as Strategy 3 covers Strategy 1F and Strategy 2F as special

³The concavity property can be shown to hold also for Rayleigh fading channels, where the symbol error is as given by (4), if $C > \log 2$ is satisfied.

cases).

The following proposition states that the solution for Strategy 2F can be obtained in a simple manner for sufficiently large values of the average power limit.

Proposition 6: *Consider K channels and assume that each channel has a cost value based on the cost function in (64). If the average power limit satisfies*

$$A_p \geq \frac{b\beta_K^2}{\kappa^2\beta_1(2\beta_K + b)}, \quad (69)$$

then the optimal solution for Strategy 2F uses channel i and channel j , where

$$i = \arg \min_{k \in \{1, \dots, K\}} C_k \text{ subject to } C_k > A_c \quad (70)$$

$$j = \arg \max_{k \in \{1, \dots, K\}} C_k \text{ subject to } C_k < A_c. \quad (71)$$

Proof: The proof can be obtained based on similar arguments to those employed in the proof of Proposition 4 in [10]. In particular, the concavity of the probability of correct decision function h can be demonstrated for $A_p \geq \frac{b\beta_K^2}{\kappa^2\beta_1(2\beta_K + b)}$ as in the proof of Proposition 5 in Appendix B, and the probability of correct decision can be shown to be maximum when the channels that are closest to A_c from above and below are employed for Strategy 2F. ■

The final proposition presents a condition under which the optimal channel switching solution involves no more than two channels. Hence, there is no need to consider Strategy 3 under the specified condition.

Proposition 7: *For the optimal channel switching problem in (13) with the cost function in (64), the optimal channel switching strategy involves at most two channels if the average power limit satisfies*

$$A_p \geq \frac{2b\beta_K^2}{\kappa^2\beta_1(2\beta_K + b)}. \quad (72)$$

Proof: The argument in the proposition can be proved in a similar way to the proof of Proposition 5 in [10]. Namely, by exploiting the concavity of the probability of correct decision function h , it can be shown that there always exists a Strategy 2F solution that yields a higher average probability of correct decision than the optimal solution of Strategy 3 under the condition stated in Proposition 7. Hence, the optimal solution can be shown to employ at most two channels if the average power limit exceeds the specified threshold. ■

Remark 4: *If the average power limit satisfies*

$$A_p \geq \max \left\{ \frac{2b\beta_K^2}{\kappa^2\beta_1(2\beta_K + b)}, \frac{b\beta_K \log \left(1 + \frac{b}{\beta_1}\right)}{v\beta_1} \right\}, \quad (73)$$

the optimal solution corresponds to either Strategy 1F or Strategy 2F. This is due to the fact that since the inequality in (73) implies (72) and (68), Strategy 3 and the partial data transmission strategies cannot be optimal, which means that the optimal strategy can be either Strategy 1F or Strategy 2F. Under the condition in (73), Strategy 2F becomes the optimal strategy if there exist no channels with cost A_c , since Strategy 1F cannot be optimal in this case (due to the fact that an optimal solution must operate at the average cost limit). As (73) also implies (69), the optimal strategy is to switch between the channels closest to A_c as described in (70) and (71). On the other hand, if there is a channel with cost A_c , Strategy 1F is the optimal one due to the following inequality:

$$h(A_p, A_c) = h(\lambda P_i + (1 - \lambda) P_j, \lambda C_i + (1 - \lambda) C_j) > \lambda h(P_i, C_i) + (1 - \lambda) h(P_j, C_j) \quad (74)$$

where the strict concavity of h for the optimal power levels of Strategy 2F is guaranteed under the condition in (69). Therefore, if the average power limit is above the threshold in (73), the optimal strategy is either Strategy 1F or Strategy 2F, depending on whether a channel with cost A_c exists or not.

V. NUMERICAL EXAMPLES

In this section, the theoretical results are demonstrated via numerical examples. Various scenarios are studied to investigate when the full or the partial transmission strategy outperforms the other one and when channel switching leads to a higher average probability of correct decision than employing a single channel for transmission. Comparisons of the following strategies are performed in the numerical examples:

Partial Transmission: In this approach, it is possible to have idle periods where no data transmission occurs. One or two channels should be employed for partial transmission due to Proposition 1. The optimal solutions for this approach are obtained based on Strategy 1P and Strategy 2P, which converges to Strategy 1P when one of the optimal channel switching factors equals to zero.

Full Transmission: In this approach, there are no idle periods during transmission, and one, two, or three channels are employed due to Proposition 1. Strategy 1F, Strategy 2F, and Strategy 3 are employed to find the optimal solution in this case. Strategy 3 converges to Strategy 2F when one of the optimal channel switching factors equals to zero, and to Strategy 1F when two of the optimal channel switching factors are zero.

A. Examples for General Analysis

In this part, numerical examples are presented for the general analysis in Section III. Three simulation scenarios with different types of modulations are presented to explore the performance improvements that can be achieved via partial/full transmission and channel switching. In the scenarios, there exist K Gaussian channels, the channel parameters and the costs of which are represented, for notational simplicity, in the vector form as $\beta = [\beta_1 \cdots \beta_K]$ and $C = [C_1 \cdots C_K]$, respectively. The results for the optimal channel switching solution for various values of the average power limit A_p are presented by using both the plots of the average probability of correct decision and the tables containing the channel switching factors and the power levels corresponding to the optimal strategy, an example of which is shown in Table II. In this table, λ_i , λ_j , and λ_k represent, respectively, the channel switching factor for the first, the second, and the third channel employed by the optimal channel switching strategy where $i < j < k$, and the non-zero power levels, the indices of which denote the employed channels, are the corresponding optimal power levels. For instance, for $A_p = 0.05$ in Table II, Strategy 2P is the optimal strategy since the sum of the channel switching factors is less than 1, and channels 1 and 2 are employed with channel switching factors 0.5248 and 0.1935, and power levels 0.0357 and 0.1615, respectively. On the other hand, for $A_p = 500$, Strategy 2F that uses channels 2 and 4 is the optimal strategy, where the channel switching factors are 0.7 and 0.3, and the power levels are 1.2539 and 1497.5, corresponding to channels 2 and 4, respectively. As for the plots of the average probability of correct decision, it should be noted that the performance of Strategy 1F and Strategy 2F is never higher than that of Strategy 1P and Strategy 2P, respectively, for any value of A_p . This is due to the fact that the optimal solution of the partial transmission strategies (Strategy 1P and Strategy 2P) can converge to that of the full transmission strategies (Strategy 1F and Strategy 2F) in cases where full transmission is optimal, and can lead to higher average probabilities of correct decision in cases where partial transmission is optimal.

In the first scenario, BPSK modulation is employed with the following parameters: $\beta = [0.51 \ 0.52 \ 0.53 \ 1]$, $C = [3 \ 1.1 \ 1.01 \ 0.01]$, and the average cost limit is equal to 1; that is, $A_c = 1$. The parameters for BPSK in the error function (5) are computed as $\eta = 1$ and $\kappa = \sqrt{2}$. In Fig. 3, the average probabilities of correct symbol decision are plotted versus the average power limit A_p for the optimal solutions of the five possible strategies, namely, Strategy 1P, Strategy 1F, Strategy 2P, Strategy 2F, and Strategy 3.⁴ The parameters of the optimal channel switching strategy are presented in Table I. It is observed from Fig. 3 that the optimal

⁴As explained in Section II, the ‘‘average probability of correct decision’’ is a scaled version of the average number of correctly received symbols (see (10) and Remark 2). Hence, if a strategy achieves a higher average probability of correct decision, it also yields a higher average number of correctly received symbols.

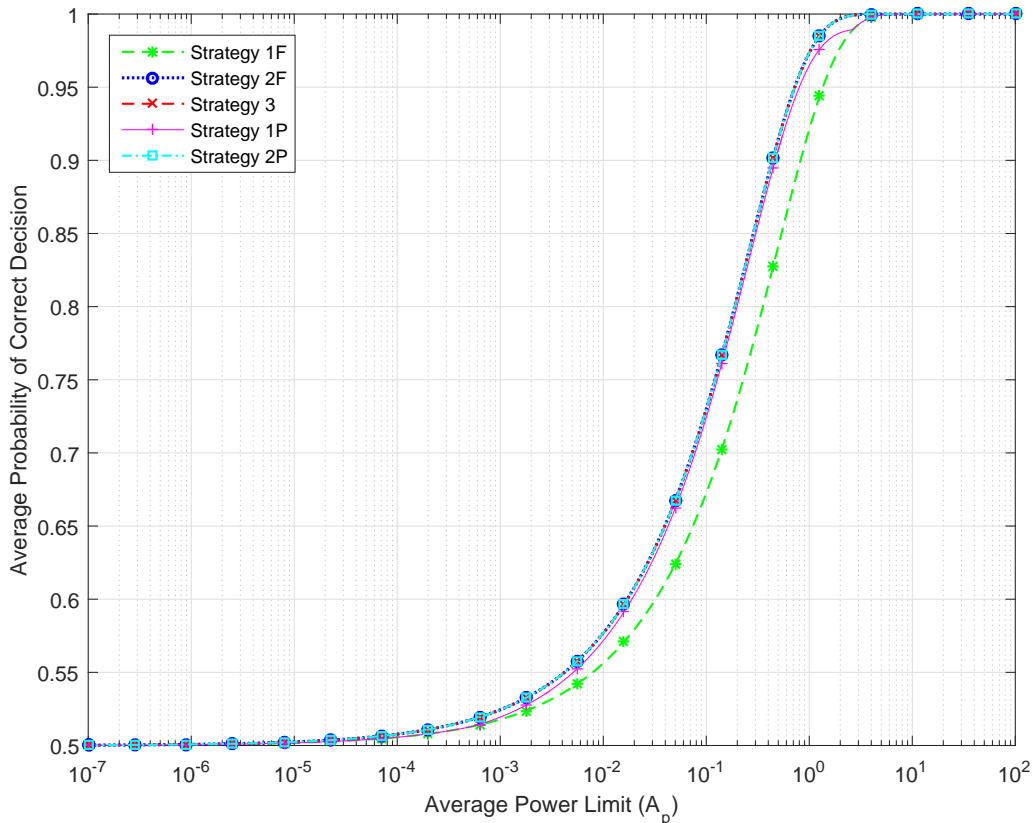


Fig. 3. Average probability of correct decision versus A_p for BPSK modulation for the optimal solution of the five possible strategies, where $\beta = [0.51 \ 0.52 \ 0.53 \ 1]$, $C = [3 \ 1.1 \ 1.01 \ 0.01]$, and $A_c = 1$.

TABLE I
PARAMETERS OF OPTIMAL CHANNEL SWITCHING STRATEGY IN FIG. 3.

A_p	λ_i	λ_j	λ_k	P_1	P_2	P_3	P_4
0.001	0.99	0.01	—	—	—	0.0010	0.0005
0.005	0.99	0.01	—	—	—	0.0050	0.0027
0.01	0.99	0.01	—	—	—	0.0100	0.0055
0.05	0.99	0.01	—	—	—	0.0502	0.0303
0.1	0.99	0.01	—	—	—	0.1003	0.0678
0.5	0.99	0.01	—	—	—	0.4994	0.5639
1	0.99	0.01	—	—	—	0.9960	1.3938
10	0.99	0.01	—	—	—	9.9183	18.092

channel switching strategies achieve higher average probabilities of correct decision than the optimal single channel strategies for all values of A_p . In addition, Strategy 2P, Strategy 2F, and Strategy 3 yield the same probabilities for all values of A_p , meaning that the optimal strategy is to switch between two channels with full transmission. The maximum gains in the average probability of correct decision provided by employing the optimal channel switching strategy (Strategy 2F in this example) and the optimal single channel strategy (Strategy 1P in this example), are given by 9.0% and 8.1%, respectively, as compared

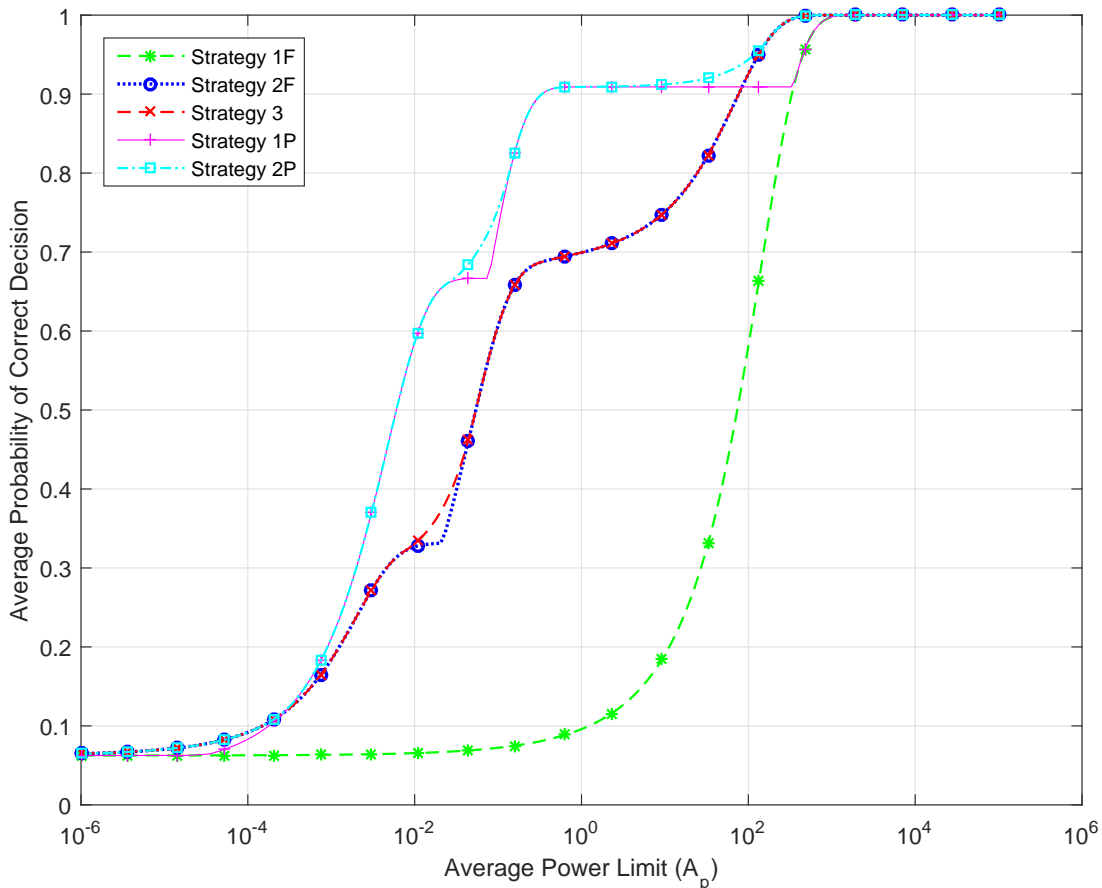


Fig. 4. Average probability of correct decision versus A_p for 16-QAM modulation for the optimal solution of the five possible strategies, where $\beta = [0.001 \ 0.01 \ 20 \ 30]$, $C = [15 \ 11 \ 10 \ 8]$, and $A_c = 10$.

to employing a single channel without idle periods (i.e., Strategy 1F). Hence, partial data transmission is not optimal in this scenario, which can be validated by Proposition 4, as well: Channel 4 satisfies the conditions stated in Proposition 4; that is, $C_4 \leq A_c$ and $\beta_4/\beta_i \leq (1 - C_4/C_i)/g_i(0) + C_4/C_i$ for $i \in \{1, 2, 3\}$, where $g_i(0) = 0.5$ can be obtained from (5). Therefore, it is concluded that partial transmission strategies cannot be optimal. In addition, since there exist no channels with cost A_c , Strategy 1F cannot be optimal either. Hence, it is theoretically concluded that either Strategy 2F or Strategy 3 is the optimal strategy in this scenario. On the other hand, Proposition 2 can also be tested with the results of this scenario. For $A_p = 0.05$, $P_3/P_4 = 1.6568 < \beta_4/\beta_3 = 1.8868$ and for $A_p = 10$, $P_4/P_3 = 1.8241 < \beta_4/\beta_3 = 1.8868$, which confirms the validity of Proposition 2 for this scenario.

The second scenario utilizes 16-QAM modulation, where the channel parameters, the channel costs, and the average cost limit are given by $\beta = [0.001 \ 0.01 \ 20 \ 30]$, $C = [15 \ 11 \ 10 \ 8]$, and $A_c = 10$, respectively, and the modulation parameters in (6) are computed as $\eta = 1.5$ and $\kappa = 0.4472$. The plot of the average probability of correct decision versus A_p is shown in Fig. 4. An important observation

is that the optimal partial transmission strategy outperforms the optimal full transmission strategy for $A_p \in (0.0002, 195.6)$, in which there exist sub-intervals where both the partial transmission over a single channel and the channel switching between two channels with partial utilization can be the overall optimal strategy. For very small and very large values of A_p , all the strategies converge to each other, indicating that Strategy 1F is the optimal one, which is theoretically possible since there exists a channel with cost A_c in this scenario. The maximum gains in terms of the average probability of correct decision achieved by employing the optimal channel switching with partial utilization (i.e., Strategy 2P) are calculated to be 76.1%, 97.2%, and 877.6% compared to Strategy 3, Strategy 2F and Strategy 1F, respectively. In addition, several turning points can be observed in Fig. 4, which generally reflect the changes in the set of employed channels as A_p increases. For instance, for A_p around 0.022, Strategy 3 outperforms Strategy 2F since Strategy 3 always employs channels 1, 2, and 4 whereas Strategy 2F uses channels 1 and 4 for $A_p < 0.022$ and channels 2 and 4 for $A_p > 0.022$. Hence, Strategy 2F changes the set of employed channels for performance improvement as A_p increases while Strategy 3 can always use the optimal set of three channels, which provides an improvement over Strategy 2F around the turning point. The parameters of the overall optimal strategy are presented in Table II for some values of A_p . Table II demonstrates that the optimal strategy may employ a single channel or two channels, and perform full or partial utilization of channels for transmission, as stated in Proposition 1. It is observed from Fig. 4 and Table II that for $A_p \in (0.023, 0.157)$, the channel switching between channel 1 and channel 2 with partial utilization outperforms the single channel strategies Strategy 1P and Strategy 1F, which employs channel 1 and channel 3, respectively, and Strategy 2F, which switches between channel 2 and channel 4. The ratio of the optimal power levels for the solutions involving two channels is calculated to confirm the validity of Proposition 2 for some values of the average power limit. For $A_p = 0.05$, $P_2/P_1 = 4.5238 < \beta_2/\beta_1 = 10$ and for $A_p = 500$, $P_4/P_2 = 1194.3 < \beta_4/\beta_2 = 3000$, which are in compliance with Proposition 2.

TABLE II
PARAMETERS OF OPTIMAL CHANNEL SWITCHING STRATEGY IN FIG. 4.

A_p	λ_i	λ_j	λ_k	P_1	P_2	P_3	P_4
0.0001	0.2857	0.7143	–	0.00035	–	–	6.1×10^{-10}
0.001	0.6667	–	–	0.0015	–	–	–
0.01	0.6667	–	–	0.015	–	–	–
0.05	0.5248	0.1935	–	0.0357	0.1615	–	–
1	0.9091	–	–	–	1.1	–	–
100	0.7724	0.1879	–	–	0.8999	–	528.3665
500	0.7	0.3	–	–	1.2539	–	1497.5
10000	1	–	–	–	–	10000	–

To demonstrate the effectiveness of channel switching with partial utilization in the presence of fading (i.e., Case 2), 8-PAM modulation with five Rayleigh fading channels is studied in the third scenario, and the parameters are given by $\beta = [0.02 \ 0.05 \ 0.1 \ 10 \ 100]$, $\mathbf{C} = [10 \ 8 \ 6 \ 4 \ 2]$, and $A_c = 5$. The modulation parameters in (4) are determined to be $\tilde{\eta} = 0.8750$ and $\tilde{\kappa} = 0.0476$. Fig. 5 illustrates the average probability of correct symbol decision with respect to A_p , and Table III presents the parameters of the optimal channel switching strategy. It is observed that the optimal partial transmission strategies can outperform the optimal full transmission strategies when the transmission is performed over Rayleigh fading channels. Strategy 1F achieves the lowest probability of correct decision for most A_p 's whereas Strategy 2P and Strategy 1P turn out to be the optimal strategy in distinct intervals of the considered A_p region. Employing the optimal partial channel utilization strategies provides a maximum gain of 4.9% and 170.4% as compared to using Strategy 2F and Strategy 1F, respectively. In addition, in Fig. 5, there are crossing and overlapping points where different strategies converge to and diverge from each other. For A_p around 100, for instance, Strategy 2P, which employs channels 2 and 4, yields a higher average probability of correct decision than Strategy 1P, which employs channel 2, whereas these two strategies have the same performance for $A_p < 80$. The reason is that Strategy 2P can use the optimal set of two channels to maximize its performance while Strategy 1P can use a single channel, which can be changed (which occurs at $A_p = 382$ in this case) only if the utilization of another channel provides a higher performance. As an example to the validity of Proposition 2, the following cases can be examined: For $A_p = 75$, $P_4/P_3 = 6.75 < \beta_4/\beta_3 = 10$ and for $A_p = 1000$, $P_3/P_4 = 9.2824 < \beta_4/\beta_3 = 10$. That is, the ratio between the optimal power levels is limited by the ratio between the channel parameters as specified in (25) in Proposition 2.

TABLE III
PARAMETERS OF OPTIMAL CHANNEL SWITCHING STRATEGY IN FIG. 5.

A_p	λ_i	λ_j	λ_k	P_1	P_2	P_3	P_4	P_5
0.001	0.375	0.625	–	0.0027	–	–	–	5.48×10^{-7}
0.01	0.375	0.625	–	0.0267	–	–	–	6.41×10^{-6}
0.1	0.375	0.625	–	0.2663	–	–	–	0.00023
1	0.8333	–	–	–	–	1.2	–	–
5	0.8333	–	–	–	–	6	–	–
10	0.8333	–	–	–	–	12	–	–
75	0.7192	0.1711	–	–	–	40.0137	270.091	–
1000	0.5	0.5	–	–	–	194.508	1805.5	–

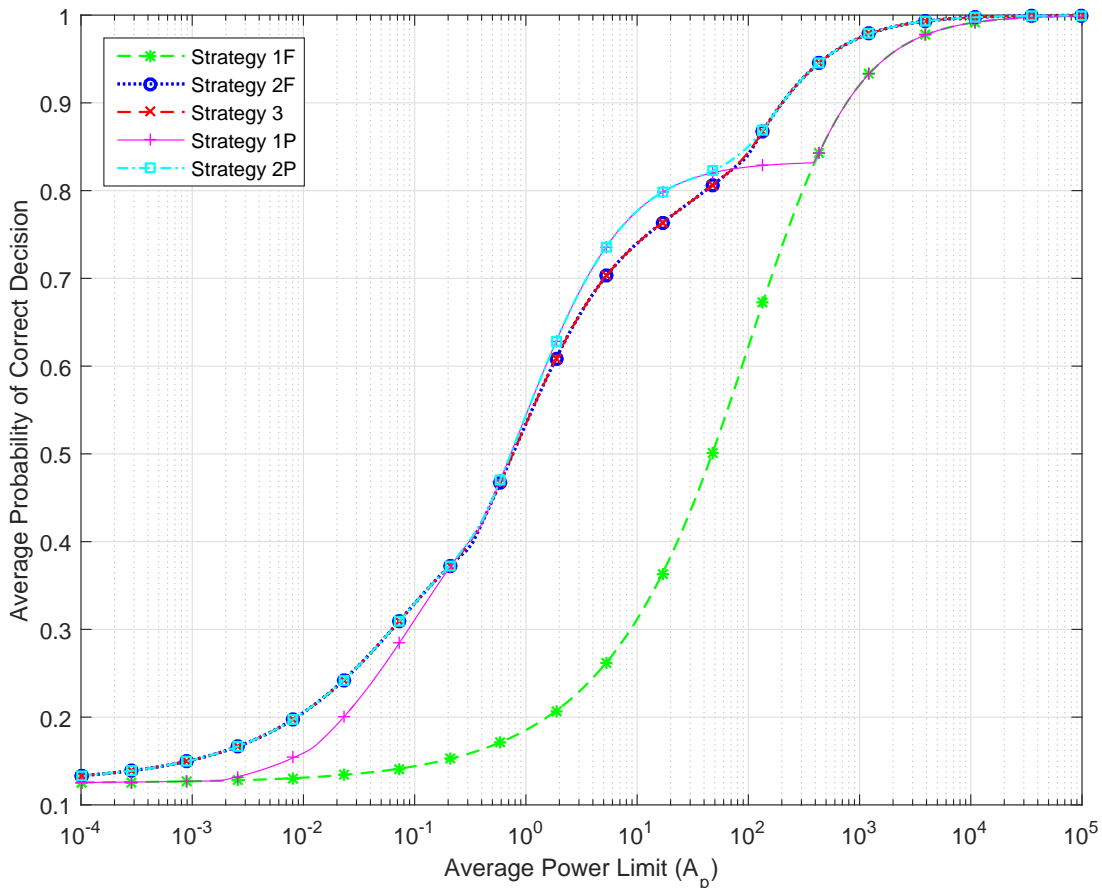


Fig. 5. Average probability of correct decision versus A_p for 8-PAM modulation for the optimal solution of the five possible strategies, where $\beta = [0.02 \ 0.05 \ 0.1 \ 10 \ 100]$, $\mathbf{C} = [10 \ 8 \ 6 \ 4 \ 2]$, and $A_c = 1$.

B. Examples for Logarithmic Cost Function

In this part, numerical examples are presented for the logarithmic cost function studied in Section IV. In the numerical examples, the magnitudes of the fading coefficients are set to $|\alpha_i| = \sqrt{2}$ for $i = 1, \dots, K$; hence, the channel parameters are given by $\beta_i = 2\sigma_i^2/|\alpha_i|^2 = \sigma_i^2$. This does not cause any loss of generality since the differences among the fading coefficients can be reflected to the variance terms, σ_i^2 , appropriately.

In the first example, four channels are available for channel switching, and BPSK modulation is employed. The BPSK parameters are calculated to be $\eta = 1$ and $\kappa = \sqrt{2}$ (cf. (63)). Also, the average cost limit is set to $A_c = 7$, and the costs of the channels are equal to $\mathbf{C} = [C_1 \ C_2 \ C_3 \ C_4] = [9 \ 8 \ 6 \ 5]$. Based on the logarithmic cost function in (64) with $b = 1$, the corresponding channel parameters are given by $\beta = [\beta_1 \ \beta_2 \ \beta_3 \ \beta_4] = [0.0001234 \ 0.0003356 \ 0.002485 \ 0.006784]$. In Fig. 6, the average probabilities of correct decision are plotted versus A_p for the five possible strategies. It is observed that the single channel strategies are outperformed by the channel switching strategies for all values of A_p in this example. Also,

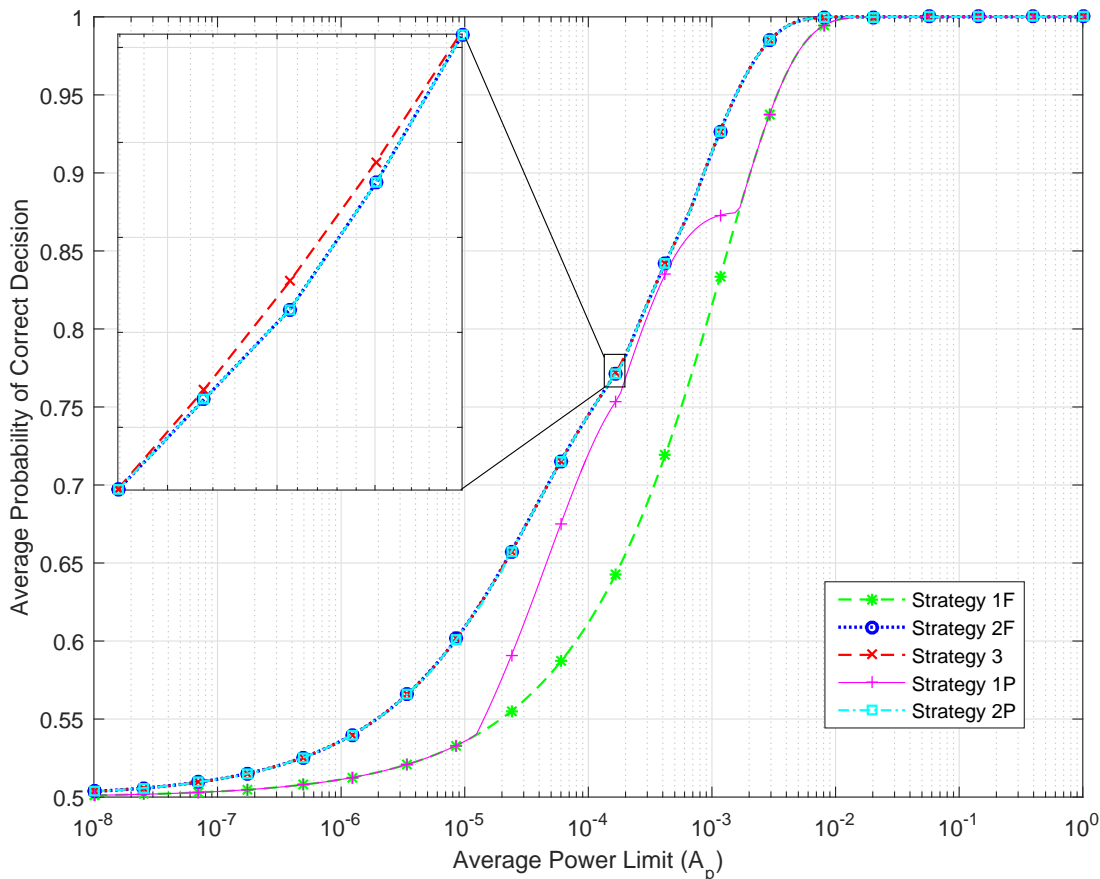


Fig. 6. Average probability of correct decision versus A_p for BPSK modulation for the optimal solutions of the five possible strategies, where $C = [9 \ 8 \ 6 \ 5]$, $b = 1$, and $A_c = 7$.

the full data transmission is always optimal, and Strategy 3 achieves the highest average probabilities of correct decision for a small interval $(0.0001517, 0.0002196)$ of A_p values (as shown in the zoomed window in Fig. 6), where the maximum gain provided by using Strategy 3 instead of the two-channel strategies is given by 0.2%. The parameters of the optimal strategy are presented in Table IV. In the table, λ_i , λ_j , and λ_k represent, respectively, the channel switching factor for the first, second, and third channel employed by the optimal channel switching strategy, and the non-zero power levels, the indices of which denote the employed channels, are the corresponding optimal power levels. For example, for $A_p = 0.0005$, channel 2 is employed with channel switching factor 0.6667 and power 0.00051, and channel 4 is employed with channel switching factor 0.3333 and power 0.00047. Table IV indicates that the optimal approach for this example is to switch either between two channels or among three channels with no idle periods.

For comparison of the channel switching strategies, it is also of interest to evaluate the performance with respect to the average cost limit, A_c . Fig. 7 considers the same scenario as in Fig. 6, and shows the average probability of correct decision for $A_p = 10^{-5}$ as A_c changes from 4 to 10. As observed from the figure, for $A_c < 5$, the full transmission strategies cannot be employed (therefore, are not optimal)

since Strategy 2F and Strategy 3 require at least one channel to have a cost less than A_c and Strategy 1F must use a channel exclusively. (Remember that the costs are given by $\mathbf{C} = [9 \ 8 \ 6 \ 5]$.) However, for the partial transmission strategies, it is possible to adjust the channel switching factors to optimize the average probability of correct decision while conforming to the constraints. An important observation from Fig. 7 is that the channel cost values are marked by some jumps in the correct decision probability in Strategy 1F, which selects the best channel for transmission as dictated by the cost constraint A_c . Strategy 1P, on the other hand, is observed to move to the best channel at a lower value of A_c as A_c increases since it can use that channel partially, which may lead to a higher performance than using a lower cost channel exclusively. Finally, the channel switching strategies outperform the single channel strategies for $5 \leq A_c \leq 9$ and all the strategies converge to Strategy 1F for $A_c \geq 9$, as expected.

In the second example, a scenario with the following parameters is considered: $\beta = [0.000306 \ 0.0126 \ 0.2035 \ 0.2249] \times 10^{-3}$, $\mathbf{C} = [15 \ 11.28 \ 8.5 \ 8.4]$, $b = 1$, and $A_c = 8.6967$. In this scenario, BFSK modulation is employed, which corresponds to parameters $\eta = 1$ and $\kappa = 1$ (cf. (63)). The average probability of correct decision curves for all the strategies are shown in Fig. 8, and the parameters of the optimal channel switching strategy are presented in Table V. It is observed that the partial transmission strategies outperform the full transmission strategies for a certain range of the average power limit values, where Strategy 1P and Strategy 2P achieve the highest average probabilities of correct decision. The maximum gains achieved by employing the optimal channel switching with partial utilization (i.e., Strategy 2P) in terms of the average probability of correct decision are given by 10.1% and 14.4%, respectively, as compared to the optimal channel switching strategy with full utilization (Strategy 3) and the optimal single channel strategy with full utilization. Since no channel has a cost value equal to A_c , Strategy 1F cannot be optimal, which is verified by the numerical example.

In the final example, the channel costs and the channel parameters are set as $\mathbf{C} = [7.0791 \ 6.5 \ 5.98 \ 5.942]$

TABLE IV
PARAMETERS OF OPTIMAL CHANNEL SWITCHING STRATEGY IN FIG. 6.

A_p	λ_i	λ_j	λ_k	P_1	P_2	P_3	P_4
0.0001	0.5	0.5	–	0.00016	–	–	0.000039
0.00018	0.3051	0.2598	0.4351	0.0002	0.0003	–	0.00009
0.0005	0.6667	0.3333	–	–	0.00051	–	0.00047
0.001	0.5	0.5	–	–	0.00064	0.00135	–
0.005	0.5	0.5	–	–	0.00171	0.00828	–
0.01	0.5	0.5	–	–	0.00293	0.01706	–
0.05	0.5	0.5	–	–	0.0139	0.0860	–
0.1	0.3333	0.6667	–	0.00609	–	0.1469	–

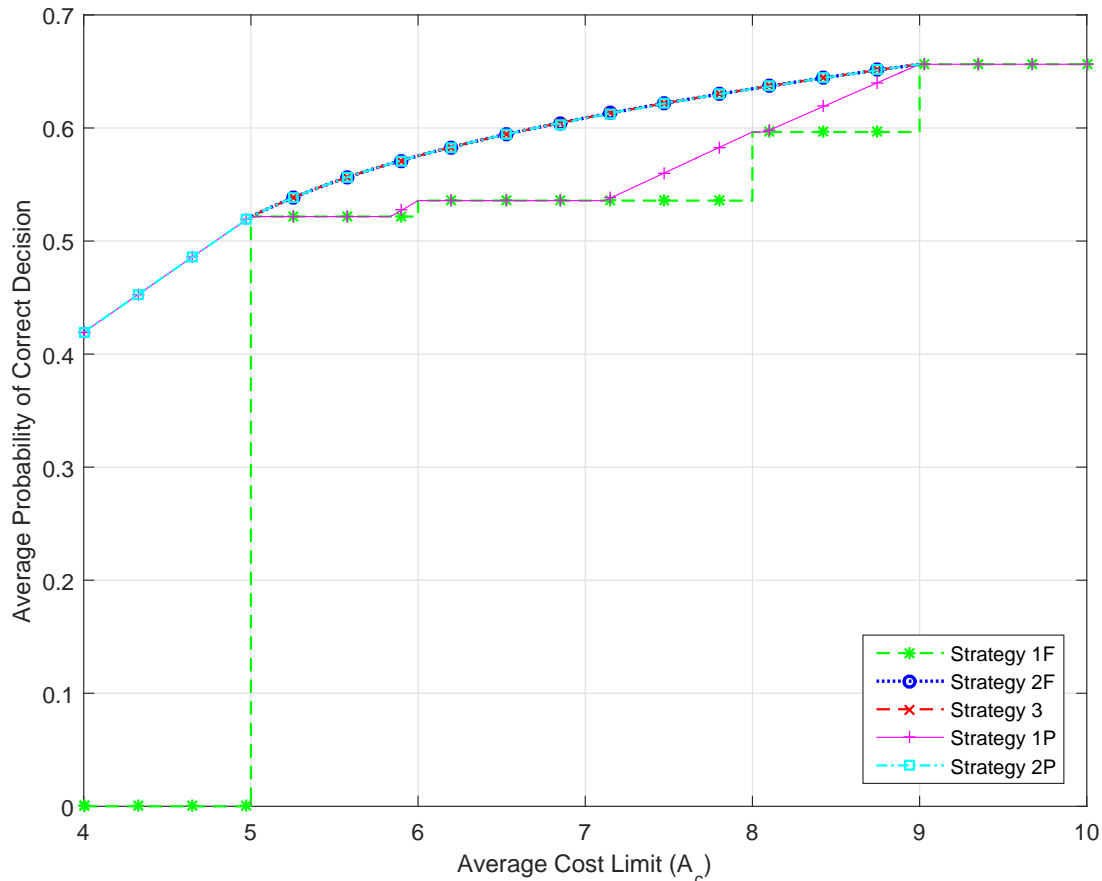


Fig. 7. Average probability of correct decision versus A_c for BPSK modulation for the optimal solutions of the five possible strategies, where $\mathbf{C} = [9 \ 8 \ 6 \ 5]$, $b = 1$, and $A_p = 10^{-5}$.

TABLE V
PARAMETERS OF OPTIMAL CHANNEL SWITCHING STRATEGY IN FIG. 8.

A_p	λ_i	λ_j	λ_k	P_1	P_2	P_3	P_4
0.000001	0.5798	–	–	0.0000017	–	–	–
0.00001	0.0578	0.6942	–	0.000002	0.000014	–	–
0.0001	0.5017	0.3574	–	–	0.000062	0.000193	–
0.001	0.0708	0.9292	–	–	0.000128	0.001066	–
0.005	0.0708	0.9292	–	–	0.000399	0.005350	–
0.01	0.0708	0.9292	–	–	0.000731	0.010705	–
0.1	0.0303	0.9697	–	0.01	–	0.102808	–
1	0.0303	0.9697	–	0.1	–	1.028085	–

and $\beta = [0.0008432 \ 0.001506 \ 0.002535 \ 0.002634]$, the average cost limit is $A_c = 6.002$, and $b = 1$. In this example, BFSK modulation is employed with the corresponding parameters $\eta = 1$ and $\kappa = 1$ (cf. (63)). The average probabilities of correct decision for the optimal full transmission strategy (which corresponds to the best of Strategy 1F, Strategy 2F, and Strategy 3) and for the optimal partial transmission strategy (which corresponds to the best of Strategy 1P and Strategy 2P) are plotted versus the average power limit in Fig. 9. It is observed from the figure that the optimal full and partial transmission solutions converge to one another

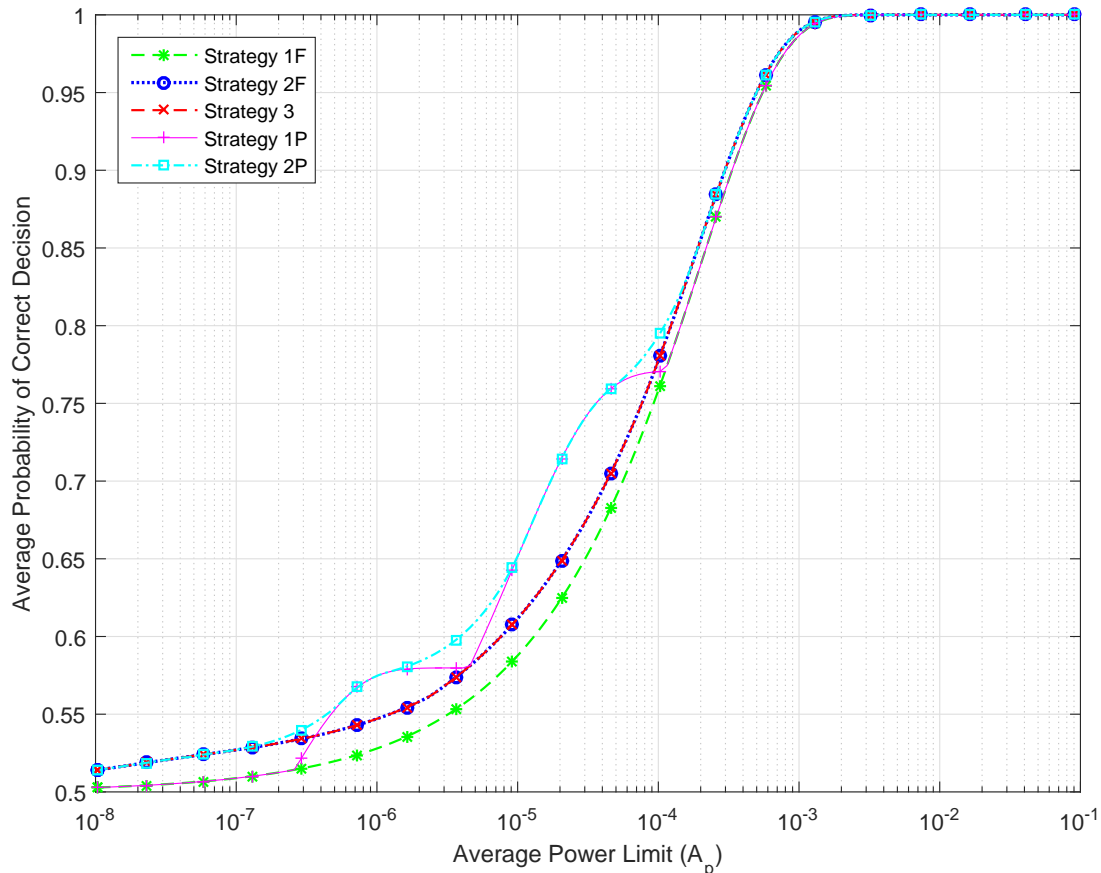


Fig. 8. Average probability of correct decision versus A_p for BFSK modulation for the optimal solutions of the five possible strategies, where $C = [15 \ 11.28 \ 8.5 \ 8.4]$, $b = 1$, and $A_c = 8.6967$.

for sufficiently low and sufficiently high values of A_p , thereby implying the optimality of the full data transmission, whereas the partial transmission achieves higher average probabilities of correct decision than the full transmission within a certain range of A_p values (please also see Fig. 10), for which the optimal partial transmission can provide a maximum gain of 1.3% in terms of the average probability of correct decision. Table VI shows the optimal channel switching factors and the power levels for the scenario in Fig. 9. According to Proposition 5, when $A_p \geq \max \{b \beta_4^2 / (\kappa^2 \beta_1 (2\beta_4 + b)), b \beta_4 \log(1 + b/\beta_1) / (v \beta_1)\} = 0.0098$, the partial data transmission cannot be optimal, which complies with the results in Fig. 9 and Table VI. On the other hand, the optimal channels leading to the highest average probability of correct decision for Strategy 2F are found to be channel 2 and channel 3 for $A_p \geq b \beta_4^2 / (\kappa^2 \beta_1 (2\beta_4 + b)) = 0.0082$, as claimed in (70) and (71) in Proposition 6. In addition, as indicated in Remark 4, when $A_p \geq \max \{2b \beta_4^2 / (\kappa^2 \beta_1 (2\beta_4 + b)), b \beta_4 \log(1 + b/\beta_1) / (v \beta_1)\} = 0.0164$, Strategy 2F is the optimal strategy since there exists no channel with a cost equal to A_c .

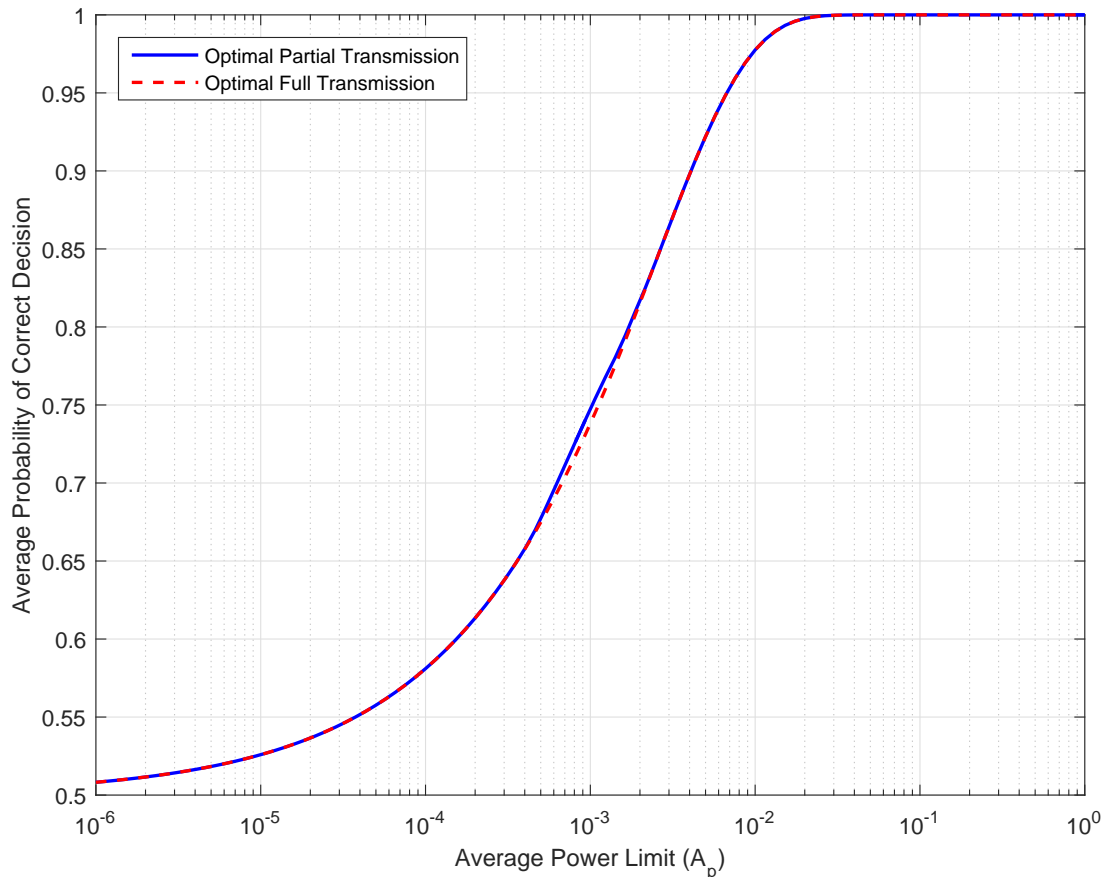


Fig. 9. Average probability of correct decision versus A_p for BFSK modulation for the optimal solutions of the full and partial transmission strategies, where $\mathcal{C} = [7.0791 \ 6.5 \ 5.98 \ 5.9420]$, $b = 1$, and $A_c = 6.002$.

VI. CONCLUSION

Optimal channel switching strategies over Gaussian noise channels have been studied under average power and cost constraints in the presence of partial and full utilization of channels for the purpose of maximizing the average probability of correct symbol decision. A generic optimization scheme has been developed to cover arbitrary signal constellations and Rayleigh fading channels, which improves the

TABLE VI
PARAMETERS OF OPTIMAL CHANNEL SWITCHING STRATEGY IN FIG. 9.

A_p	λ_i	λ_j	λ_k	P_1	P_2	P_3	P_4
0.00001	0.0528	0.9472	–	0.000027	–	–	0.000009
0.00005	0.0528	0.9472	–	0.000125	–	–	0.000045
0.0001	0.0528	0.9472	–	0.000228	–	–	0.000092
0.0005	0.7498	0.1169	–	0.000611	–	–	0.000353
0.001	0.8478	–	–	0.001179	–	–	–
0.005	0.0423	0.9577	–	–	0.004091	0.00504	–
0.01	0.0423	0.9577	–	–	0.007288	0.010119	–
0.05	0.0423	0.9577	–	–	0.0316724	0.050809	–
0.1	0.0423	0.9577	–	–	0.061925	0.101681	–

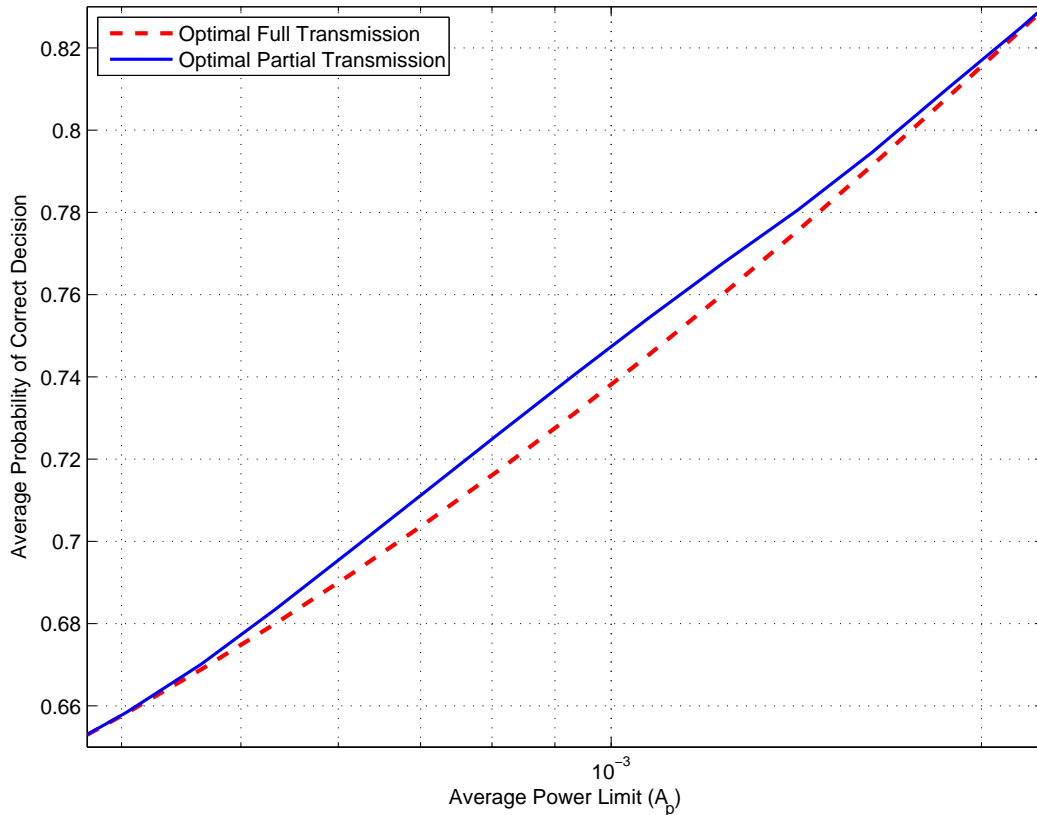


Fig. 10. The zoomed version of Fig. 9 for $A_p \in [3.75 \times 10^{-4}, 2.25 \times 10^{-3}]$.

applicability of the results to various types of modulations in the presence of slow or block fading. It has been demonstrated that the optimal channel switching strategy employs at most three channels in the full transmission case and at most two channels in the partial transmission case. In addition, it has been stated that the optimal solution must operate at the average power and the average cost limits. For the two-channel strategies, the ratio between the optimal power levels has been shown to be upper bounded by the ratio of the parameters of the employed channels. Conditions that depend only on the system parameters, namely, the channel costs, the standard deviations, and the fading statistics, have been derived under which partial data transmission cannot be optimal. The optimal channel switching problem has also been investigated for logarithmic cost functions. It has been shown that full data transmission is optimal when the average power limit exceeds a certain threshold, which is related to the parameters of the best and the worst channels. Also, the optimal channel switching with the full utilization of two channels has been specified for sufficiently high values of the average power limit. In addition, it has been stated that the use of more than two channels is not necessary in obtaining the optimal channel switching strategy when the average power limit is larger than a specified threshold. Improvements via channel switching

and partial channel utilization have been illustrated via numerical results, which demonstrate that full utilization of channels does not always yield a higher average probability of correct decision than partial utilization and that the optimal single channel strategy can be outperformed by the proposed optimal channel switching approach.

Remark 5: *In order to emphasize the differences of this study from the recent works in the literature, Table VII is presented. As can be noted from the table, one of the main contributions is related to generalizing the concept of channel switching to scenarios with idle periods. In this way, the average number of correctly received symbols, which is an important parameter in practical systems, can be improved in certain communication systems. In addition, the results are provided for block Rayleigh fading channels, as well. Regarding the theoretical contributions, Propositions 1, 2, 6, and 7 generalize and improve the results in the literature while Propositions 3, 4, and 5 present completely new results.*

APPENDIX

A. Proof of Lemma 2

Consider the surface $h(P, C)$ and let (P^*, C^*) denote a given power and cost pair. For the inequality in (66) to hold for the pair $(P^*, C^*) \forall \lambda \in (0, 1)$, the line passing through the point $(0, 0, 0)$, i.e., through the origin, and the point $(P^*, C^*, h(P^*, C^*))$ should not intersect with the surface $h(P, C)$ for any $P \in (0, P^*)$ and $C \in (0, C^*)$. Let $m \triangleq C^*/P^*$ represent the slope of the two-dimensional projection onto the xy -plane of the line passing through the origin and the point $(P^*, C^*, h(P^*, C^*))$. Since $h(P, C)$ is concave for $P > b/(\kappa^2(e^C + 1))$ due to Lemma 1, if the line tangent to the surface $h(P, C)$ at $P = b/(\kappa^2(e^C + 1))$, passing through the z -axis, and whose projection onto the xy -plane has a slope of m , intersects with the z -axis above the xy -plane, then the inequality $\lambda h(P^*, C^*) < h(\lambda P^*, \lambda C^*)$ holds $\forall \lambda \in (0, 1)$ if $P^* > b/(\kappa^2(e^{C^*} + 1))$. The proof of the argument in the previous sentence can be explained more clearly by referring to Fig. 11 and Fig. 12, which illustrate slices of the surface $h(P, C)$ cut along the z -axis and in the direction connecting the origin and the point (P^*, C^*) on the xy -plane, where the slope is $m = 500$

TABLE VII
COMPARISON OF THIS STUDY WITH SOME RECENT WORKS.

	[11]	[13]	[10]	This Study
Objective function	Average prob. of correct decision	Average capacity	Average prob. of error	Average # correctly received symbols
Idle periods	No	No	No	Yes
Channel costs	No	No	Yes	Yes
Rayleigh fading	No	No	No	Yes

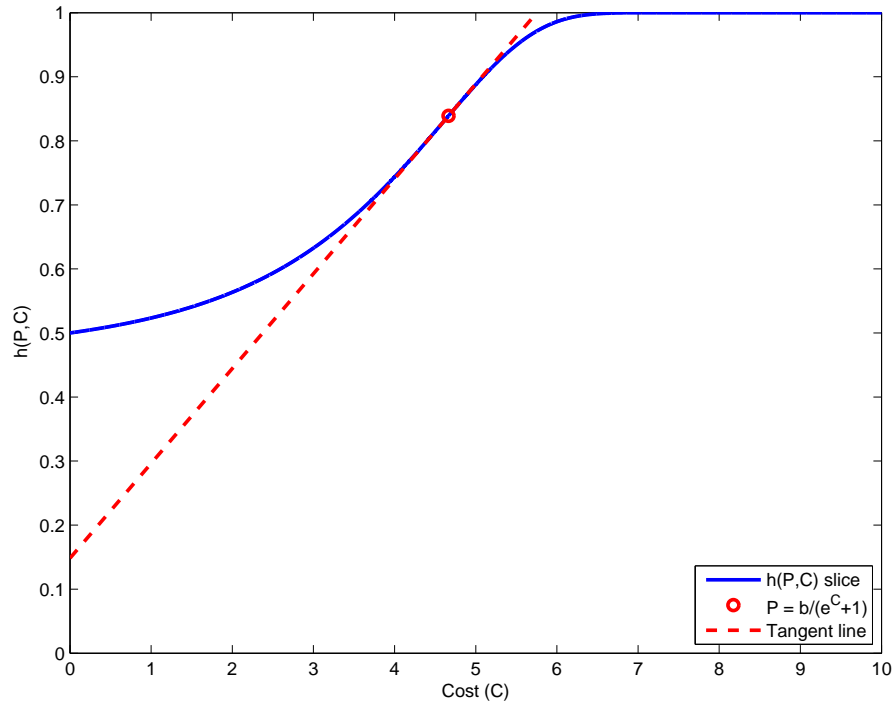


Fig. 11. Slice of the surface $h(P, C)$ cut along the z -axis, where $b = 1$, $\kappa = 1$, and the slope of the slice on the xy -plane is $m = C/P = 500$. The intersection of the line tangent to the resulting curve at $P = b/(\kappa^2(e^C + 1))$ with the z -axis is positive.

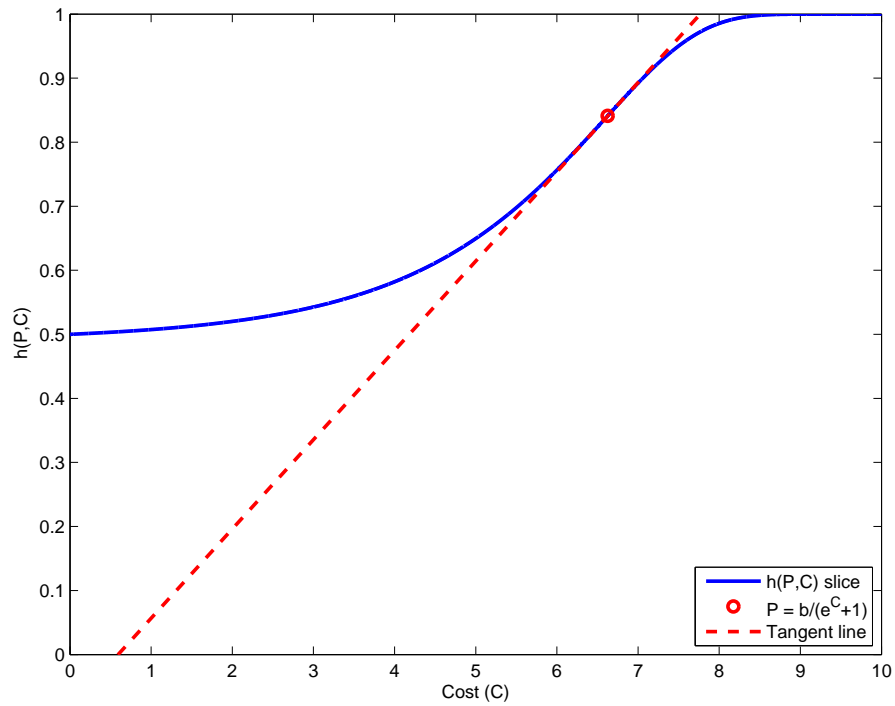


Fig. 12. Slice of the surface $h(P, C)$ cut along the z -axis, where $b = 1$, $\kappa = 1$, and the slope of the slice on the xy -plane is $m = C/P = 5000$. The intersection of the line tangent to the resulting curve at $P = b/(\kappa^2(e^C + 1))$ with the z -axis is negative.

in the first figure and $m = 5000$ in the latter one. The proof can be obtained via contradiction in two steps. In the proof, consider only those cost values C^* that satisfy $P^* > b/(\kappa^2(e^{C^*} + 1))$, meaning that $C^* > \tilde{C}$, where (\tilde{P}, \tilde{C}) is the power-cost pair satisfying $\tilde{P} = b/(\kappa^2(e^{\tilde{C}} + 1))$ for the slope $\tilde{C}/\tilde{P} = m$. Consider a case as depicted in Fig. 11, where the intersection of the tangent line with the z -axis is positive, and assume that the line connecting the origin $(0, 0)$ and the point $(C^*, h(P^*, C^*))$ intersects with the curve at a cost value $C' \in (0, C^*)$. First, the case with $C' \in (0, \tilde{C})$ is analyzed. The slope of the line passing through the origin and the point $(\tilde{C}, h(\tilde{P}, \tilde{C}))$ is higher than that of the tangent line and lower than that of the line passing through the origin and the point $(C', h(P', C'))$, where $C'/P' = m$. Due to the concavity of the curve for $C > \tilde{C}$, the tangent line never crosses the curve for $C > \tilde{C}$. Thus, the line passing through the origin and the point $(C', h(P', C'))$ does not intersect the curve for $C > \tilde{C}$, which contradicts with the initial assumption. Therefore, $C' \in [\tilde{C}, C^*)$ must hold. The concavity of the curve for $C \geq \tilde{C}$ indicates that any line connecting two points on the curve in the region $C \geq \tilde{C}$ has a slope lower than that of the tangent line. Hence, the slope of the line passing through the points $(C^*, h(P^*, C^*))$ and $(C', h(P', C'))$ is lower than that of the tangent line and the intersection of that line with the z -axis is positive since the tangent line crosses the z -axis above the x -axis (“Cost”-axis). This means that the line connecting the points $(C^*, h(P^*, C^*))$ and $(C', h(P', C'))$ never passes through the origin, thereby leading to a contradiction with the initial assumption. Therefore, the line connecting the origin $(0, 0)$ and the point $(C^*, h(P^*, C^*))$ does not intersect the curve except at the end point, which completes the proof of the inequality $\lambda h(P^*, C^*) < h(\lambda P^*, \lambda C^*) \forall \lambda \in (0, 1)$ for $P^* > b/(\kappa^2(e^{C^*} + 1))$ when the intersection of the tangent line with the z -axis is positive, an example of which is presented in Fig. 11. For $P^* < b/(\kappa^2(e^{C^*} + 1))$, the validity of the inequality in (66) cannot be guaranteed as the convexity-concavity test is inconclusive for $P < b/(\kappa^2(e^C + 1))$, as noted in Lemma 1. Similarly, when the intersection of the tangent line with the z -axis is negative as in Fig. 12, the inequality in (66) may not hold for some (P, C) satisfying $P > b/(\kappa^2(e^C + 1))$.

In order to complete the proof of Lemma 2, the range of values of m is determined for which the intersection of the line tangent to the surface $h(P, C)$ at $P = b/(\kappa^2(e^C + 1))$ with the z -axis is positive, where m is the slope of the projection of that line onto the xy -plane (i.e., the PC -plane). Let $D(P, C, z)$ and (P_0, C_0, z_0) denote, respectively, the equation of the surface $h(P, C)$ and the point of tangency to the

surface $h(P, C)$ at $P = b/(\kappa^2(e^C + 1))$, leading to the following relations:

$$D(P, C, z) = z + \eta Q \left(\kappa \sqrt{\frac{P(e^C - 1)}{b}} \right) = 1 \quad (75)$$

$$P_0 = \frac{b}{\kappa^2(e^{C_0} + 1)} \quad (76)$$

$$z_0 = h(P_0, C_0) \quad (77)$$

The equation of the tangent plane to the surface $h(P, C)$ at $P = b/(\kappa^2(e^C + 1))$ can be expressed as

$$aP + bC + cz = aP_0 + bC_0 + cz_0 \quad (78)$$

where

$$a = \left. \frac{\partial D}{\partial P} \right|_{(P,C,z)=(P_0,C_0,z_0)}, \quad b = \left. \frac{\partial D}{\partial C} \right|_{(P,C,z)=(P_0,C_0,z_0)}, \quad c = \left. \frac{\partial D}{\partial z} \right|_{(P,C,z)=(P_0,C_0,z_0)}. \quad (79)$$

The intersection point of the z -axis with the tangent plane, denoted by z_{int} , can be calculated by inserting $(P, C) = (0, 0)$ into (78) as

$$z_{int} = \frac{a}{c}P_0 + \frac{b}{c}C_0 + z_0 \quad (80)$$

which, after inserting (75)-(77) and (79) and performing algebraic manipulations, yields the following equality:

$$z_{int} = 1 - \frac{\eta}{2\sqrt{2\pi}} e^{-\frac{\kappa^2 P_0 (e^{C_0} - 1)}{2b}} \sqrt{\frac{\kappa^2 P_0}{(e^{C_0} - 1)b}} (C_0 e^{C_0} + e^{C_0} - 1) - \eta Q \left(\sqrt{\frac{\kappa^2 P_0 (e^{C_0} - 1)}{b}} \right). \quad (81)$$

Define $x \triangleq \kappa \sqrt{P_0/b}$. Then, C_0 is computed as $C_0 = \log(1/x^2 - 1)$ due to (76). Also, z_{int} can be rewritten as a function of x as

$$z_{int} = w(x) \triangleq 1 - \frac{\eta}{2\sqrt{2\pi}} e^{x^2 - \frac{1}{2}} \left(\sqrt{1 - 2x^2} + \frac{1 - x^2}{\sqrt{1 - 2x^2}} \log \left(\frac{1 - x^2}{x^2} \right) \right) - \eta Q \left(\sqrt{1 - 2x^2} \right). \quad (82)$$

Parameter η , as defined in (65), takes values in the interval $(0, 4]$ for different modulations and $w(x)$ is a monotone increasing function for any $\eta \in (0, 4]$. Therefore, $w(x) = 0$ is satisfied for $x = r^{-1}(1/\eta)$ where the function $r(x)$ is defined as

$$r(x) \triangleq \frac{1}{2\sqrt{2\pi}} e^{x^2 - \frac{1}{2}} \left(\sqrt{1 - 2x^2} + \frac{1 - x^2}{\sqrt{1 - 2x^2}} \log \left(\frac{1 - x^2}{x^2} \right) \right) - Q \left(\sqrt{1 - 2x^2} \right). \quad (83)$$

Hence, $w(x) \leq 0$ for $x \in [0, r^{-1}(1/\eta)]$ and $w(x) > 0$ for $x \in (r^{-1}(1/\eta), \kappa/\sqrt{2}]$. Note that the maximum

value of x is given by $\kappa\sqrt{P_0/b} = \kappa\sqrt{1/(e^{C_0} + 1)} \leq \kappa/\sqrt{2}$. Therefore, it can be concluded that if $\kappa\sqrt{P_0/b} > r^{-1}(1/\eta)$, $z_{int} > 0$. The problem is to determine the range of values of m for which the point of tangency (P_0, C_0, z_0) satisfies $\kappa\sqrt{P_0/b} > r^{-1}(1/\eta)$, or equivalently, $C_0 < \log([r^{-1}(1/\eta)]^{-2} - 1)$ by (76). Here, the upper bound for the slope m so that $z_{int} > 0$ holds is provided by the following expression:

$$m = \frac{C}{P} = \frac{C_0}{P_0} = \frac{C_0(e^{C_0} + 1)}{b} < \frac{(r^{-1}(1/\eta))^{-2} \log\left((r^{-1}(1/\eta))^{-2} - 1\right)}{b}. \quad (84)$$

Defining

$$v \triangleq (r^{-1}(1/\eta))^{-2} \log\left((r^{-1}(1/\eta))^{-2} - 1\right) \quad (85)$$

as a parameter that depends on η , if

$$P > \frac{bC}{v}, \quad (86)$$

the tangent line to the surface $h(P, C)$ at $P = b/(\kappa^2(e^C + 1))$, which passes through the z -axis and whose projection onto the xy -plane has a slope that satisfies (84), and thus (86), intersects with the z -axis above the xy -plane. Therefore, the inequality $\lambda h(P, C) < h(\lambda P, \lambda C)$ is satisfied $\forall \lambda \in (0, 1)$ for $P > b/(\kappa^2(e^C + 1))$ if (86) is satisfied. In other words,

$$P > \max\left\{\frac{b}{\kappa^2(e^C + 1)}, \frac{bC}{v}\right\} \quad (87)$$

guarantees the validity of the inequality $\lambda h(P, C) < h(\lambda P, \lambda C) \forall \lambda \in (0, 1)$.

Finally, the convexity of set \mathcal{S}_i is demonstrated in a manner similar to the proof of convexity of set \mathcal{S}_c in Lemma 1 of [10]. Let (P_i, C_i) and (P_j, C_j) denote any two elements from set \mathcal{S}_i , whose convex combination is expressed as $(\lambda P_i + (1 - \lambda)P_j, \lambda C_i + (1 - \lambda)C_j)$, where $\lambda \in [0, 1]$. $\lambda C_i + (1 - \lambda)C_j$ lies in the interval (C_{\min}, C_{\max}) . The convexity of \mathcal{S}_i related to the power component is shown via the following inequality:

$$\lambda P_i + (1 - \lambda)P_j > \lambda \frac{bC_i}{v} + (1 - \lambda) \frac{bC_j}{v} = \frac{b(\lambda C_i + (1 - \lambda)C_j)}{v}. \quad (88)$$

Combining the inequalities in eqn. (38) of [10] and in (88) yields

$$\lambda P_i + (1 - \lambda)P_j > \max\left\{\frac{b}{e^{(\lambda C_i + (1 - \lambda)C_j)} + 1}, \frac{b(\lambda C_i + (1 - \lambda)C_j)}{v}\right\} \quad (89)$$

which completes the proof that \mathcal{S}_i is a convex set. ■

B. Proof of Proposition 5

The partial transmission strategies either use one channel or two channels, corresponding to Strategy 1P and Strategy 2P, respectively, as defined in Section II. Assume that there exist no channels with cost A_c (the case in which there exists such a channel is studied at the end). The proof of the proposition consists of two main parts:

Proof for non-optimality of Strategy 2P: Firstly, consider Strategy 2P, and let i and j denote the channels employed for channel switching between two channels. Then, the channel switching problem in (13) can be expressed as in (22). Assume without loss of generality that $C_i > C_j$. Let the solution of the optimization problem in (22) for the channel pair (i, j) under Strategy 2P be denoted by $\{\lambda_i^*, \lambda_j^*, P_i^*, P_j^*\}$. Assume that $\lambda_i^* > 0$, $\lambda_j^* > 0$, and $\lambda_i^* + \lambda_j^* < 1$ so that the optimal solution does not involve the two-channel full (Strategy 2F), one-channel partial (Strategy 1P) and one-channel full (Strategy 1F) cases. The relations $\lambda_i^* C_i + \lambda_j^* C_j = A_c$ and $\lambda_i^* + \lambda_j^* < 1$ imply that at least one of the channels should have a cost higher than A_c ; thus, $C_i > A_c$. Then, the possible solutions for Strategy 2P are handled in two cases: $C_j > A_c$ and $C_j < A_c$.

For $C_j > A_c$, the channel costs are ordered as $C_i > C_j > A_c$. It will be demonstrated that Strategy 1P, which employs channel j , achieves a higher average probability of correct decision than the optimal solution obtained for Strategy 2P, which employs channels i and j , if (68) is satisfied; that is,

$$\lambda_i^* h(P_i^*, C_i) + \lambda_j^* h(P_j^*, C_j) < \tilde{\lambda}_j h(\tilde{P}_j, C_j) \quad (90)$$

where $\tilde{\lambda}_j = A_c/C_j$ and $\tilde{P}_j = A_p C_j/A_c$ are, respectively, the channel switching factor and the power level obtained by employing channel j for Strategy 1P. Since an optimal solution must operate at the average power limit and at the average cost limit, the following relations are obtained:

$$\lambda_i^* (P_i^*, C_i) + \lambda_j^* (P_j^*, C_j) = \tilde{\lambda}_j (\tilde{P}_j, C_j) = (A_p, A_c). \quad (91)$$

Assume that the inequality $\lambda_i^* + \lambda_j^* \geq \tilde{\lambda}_j$ holds. Then, $\lambda_i^* \geq A_c/C_j - \lambda_j^*$ is obtained, which, after inserting into (91), yields $A_c = \lambda_i^* C_i + \lambda_j^* C_j \geq C_i A_c/C_j - \lambda_j^* (C_i - C_j)$. Rearranging the terms, the inequality $\lambda_j^* \geq A_c/C_j$ is reached, which leads to a contradiction since $\lambda_j^* C_j < \lambda_i^* C_i + \lambda_j^* C_j = A_c$ noting that $\lambda_i^* > 0$ and $C_i > 0$. Hence, the optimal channel switching factors satisfy $\lambda_i^* + \lambda_j^* < \tilde{\lambda}_j$. Meanwhile, the optimal power levels satisfy the inequality $\min\{P_i^*, P_j^*\} > A_p \beta_i/\beta_j$, which can be obtained based on the upper bound in Proposition 2 (see (25)) and the fact that at least one of the power levels should be larger than A_p . The lower bound for the optimal power levels can be computed from (68) and the relation

$\beta_1 < \beta_2 < \dots < \beta_K$ as

$$\min\{P_i^*, P_j^*\} > \frac{A_p \beta_i}{\beta_j} > \frac{A_p \beta_1}{\beta_K} \geq \max \left\{ \frac{b}{\kappa^2(2 + b/\beta_K)}, \frac{b \log \left(1 + \frac{b}{\beta_1}\right)}{v} \right\} = \max \left\{ \frac{b}{\kappa^2(e^{C_K} + 1)}, \frac{b C_1}{v} \right\}. \quad (92)$$

Since the ordering of the channel costs is given by $C_1 > C_2 > \dots > C_K$, (92) guarantees that both (P_i^*, C_i) and (P_j^*, C_j) are elements of set \mathcal{S}_i , as defined in Lemma 2. Exploiting the strict concavity of $h(P, C)$ for $P > b/(\kappa^2(e^C + 1))$, the average probability of correct decision for Strategy 2P can be expressed as

$$\lambda_i^* h(P_i^*, C_i) + \lambda_j^* h(P_j^*, C_j) < (\lambda_i^* + \lambda_j^*) h \left(\frac{\lambda_i^* P_i^* + \lambda_j^* P_j^*}{\lambda_i^* + \lambda_j^*}, \frac{\lambda_i^* C_i + \lambda_j^* C_j}{\lambda_i^* + \lambda_j^*} \right). \quad (93)$$

Since \mathcal{S}_i is a convex set by Lemma 2, the convex combination of (P_i^*, C_i) and (P_j^*, C_j) is also an element of \mathcal{S}_i . Hence,

$$\frac{(\lambda_i^* + \lambda_j^*)}{\tilde{\lambda}_j} h \left(\frac{\lambda_i^* P_i^* + \lambda_j^* P_j^*}{\lambda_i^* + \lambda_j^*}, \frac{\lambda_i^* C_i + \lambda_j^* C_j}{\lambda_i^* + \lambda_j^*} \right) < h \left(\frac{\lambda_i^* P_i^* + \lambda_j^* P_j^*}{\tilde{\lambda}_j}, \frac{\lambda_i^* C_i + \lambda_j^* C_j}{\tilde{\lambda}_j} \right) \quad (94)$$

$$= h(\tilde{P}_j, C_j) \quad (95)$$

where the inequality results from Lemma 2 and the inequality $\lambda_i^* + \lambda_j^* < \tilde{\lambda}_j$, and the equality is due to (91). Combining the relations in (93)-(95) yields the inequality in (90). Thus, when $C_j > A_c$, Strategy 2P cannot be optimal as it always attains a lower average probability of correct decision than Strategy 1P that employs channel j .

The other case to consider for Strategy 2P is $C_j < A_c$. In this case, $C_i > A_c > C_j$. It will be proved that the optimal channel switching solution for Strategy 2P leads to a lower average probability of correct decision than a specific solution for Strategy 2F with channel switching factors $\{(\lambda_i^* - \gamma), (\lambda_j^* + \gamma + \lambda_{K+1}^*)\}$ and power levels $\{P_i^*, P_j\}$; that is,

$$\lambda_i^* h(P_i^*, C_i) + \lambda_j^* h(P_j^*, C_j) < (\lambda_i^* - \gamma) h(P_i^*, C_i) + (\lambda_j^* + \gamma + \lambda_{K+1}^*) h(P_j, C_j) \quad (96)$$

where $\lambda_{K+1}^* \triangleq 1 - (\lambda_i^* + \lambda_j^*)$ is the idle period, and $\gamma = \frac{\lambda_{K+1}^* C_j}{C_i - C_j}$ and $P_j = \frac{\gamma P_i^* + \lambda_j^* P_j^*}{\lambda_j^* + \gamma + \lambda_{K+1}^*}$ are obtained from the fact that an optimal solution must operate at the average power and average cost limits; i.e.,

$$\lambda_i^* (P_i^*, C_i) + \lambda_j^* (P_j^*, C_j) = (\lambda_i^* - \gamma) (P_i^*, C_i) + (\lambda_j^* + \gamma + \lambda_{K+1}^*) (P_j, C_j) = (A_p, A_c). \quad (97)$$

The switching factor $(\lambda_i^* - \gamma)$ is legitimate due to the inequality $0 < \gamma < \lambda_i^*$. Assuming $\gamma \geq \lambda_i^*$ yields

$(\lambda_{K+1}^* + \lambda_i^*) C_j \geq \lambda_i^* C_i$, which is equivalent to $(1 - \lambda_j^*) C_j \geq \lambda_i^* C_i$. This leads to a contradiction since it leads to $C_j \geq \lambda_i^* C_i + \lambda_j^* C_j = A_c$. Hence, $\gamma < \lambda_i^*$ is obtained. Since $C_i > C_j$, $\gamma > 0$; hence, the proof of the inequality $0 < \gamma < \lambda_i^*$ is complete. After rearranging the terms, the inequality in (96) can be expressed as

$$\gamma h(P_i^*, C_i) + \lambda_j^* h(P_j^*, C_j) < (\lambda_j^* + \gamma + \lambda_{K+1}^*) h(P_j, C_j). \quad (98)$$

Since $(P_i^*, C_i) \in \mathcal{S}_i$ and $(P_j^*, C_j) \in \mathcal{S}_i$, the following inequality is obtained from the strict concavity of $h(P, C)$ for $P > b/(\kappa^2(e^C + 1))$:

$$\gamma h(P_i^*, C_i) + \lambda_j^* h(P_j^*, C_j) < (\gamma + \lambda_j^*) h\left(\frac{\gamma P_i^* + \lambda_j^* P_j^*}{\gamma + \lambda_j^*}, \frac{\gamma C_i + \lambda_j^* C_j}{\gamma + \lambda_j^*}\right). \quad (99)$$

The convex combination of (P_i^*, C_i) and (P_j^*, C_j) lies also in \mathcal{S}_i due to the convexity of \mathcal{S}_i . Therefore,

$$\frac{\gamma + \lambda_j^*}{\gamma + \lambda_j^* + \lambda_{K+1}^*} h\left(\frac{\gamma P_i^* + \lambda_j^* P_j^*}{\gamma + \lambda_j^*}, \frac{\gamma C_i + \lambda_j^* C_j}{\gamma + \lambda_j^*}\right) < h\left(\frac{\gamma P_i^* + \lambda_j^* P_j^*}{\gamma + \lambda_j^* + \lambda_{K+1}^*}, \frac{\gamma C_i + \lambda_j^* C_j}{\gamma + \lambda_j^* + \lambda_{K+1}^*}\right) \quad (100)$$

$$= h(P_j, C_j) \quad (101)$$

where the inequality is based on Lemma 2 and the equality results from (97). Combining the relations in (99)-(101) leads to the inequality (98), and hence, to the inequality in (96). Therefore, for $C_j < A_c$, there always exists a full transmission strategy employing two channels that achieves a higher average probability of correct decision than the optimal solution for Strategy 2P. This completes the proof of the proposition for partial transmission strategies that use two channels.

Proof for non-optimality of Strategy 1P: Secondly, Strategy 1P is considered and shown to be outperformed by Strategy 2F if the condition in (68) is satisfied. Let i denote the channel employed for partial transmission over a single channel. Since an optimal solution must operate at the average cost limit and since no channels with cost A_c are assumed to exist, $C_i > A_c$ must be satisfied. Let $\{\lambda_i^*, P_i^*\}$ denote the optimal channel switching factor and the optimal power level for Strategy 1P when channel i is employed. Then, $\lambda_i^* = A_c/C_i$ and $P_i^* = A_p C_i/A_c$ are obtained due to the fact that an optimal solution must operate at the average cost and power limits. In the following, the optimal solution under Strategy 1P is proved to achieve a lower probability of correct decision than a specific solution under Strategy 2F which uses channels i and j with channel switching factors and power levels of $\{(\lambda_i^* - \gamma), (\gamma + \lambda_{K+1}^*)\}$ and $\{P_i^*, P_j\}$, respectively, where $\lambda_{K+1}^* \triangleq 1 - \lambda_i^*$ is defined as the idle period, and $\gamma = \frac{\lambda_{K+1}^* C_j}{C_i - C_j}$ and $P_j = \frac{\gamma P_i^*}{\gamma + \lambda_{K+1}^*}$ are

derived based on the property of an optimal solution to operate at the average power and cost limits; i.e.,

$$(\lambda_i^* - \gamma) (P_i^*, C_i) + (\gamma + \lambda_{K+1}^*) (P_j, C_j) = \lambda_i^* (P_i^*, C_i) = (A_p, A_c). \quad (102)$$

Here, the cost of channel j satisfies $C_j < A_c$ as $C_i > A_c$, and the inequality to prove is expressed as

$$\lambda_i^* h(P_i^*, C_i) < (\lambda_i^* - \gamma) h(P_i^*, C_i) + (\gamma + \lambda_{K+1}^*) h(P_j, C_j), \quad (103)$$

where the left and right hand sides are the average probabilities of correct decision corresponding to Strategy 1P and Strategy 2F, respectively. Note that $(\lambda_i^* - \gamma)$ is a legitimate switching factor since $0 < \gamma < \lambda_i^*$. $C_i > A_c > C_j$ implies that $\gamma > 0$. Assume $\gamma \geq \lambda_i^*$, which, after manipulations, leads to the following inequality: $A_c = \lambda_i^* C_i \leq C_j$. This contradicts with the inequality $C_j < A_c$. Therefore, the inequality $0 < \gamma < \lambda_i^*$ is verified. The inequality in (103) reduces to the following form after some algebraic operations:

$$\frac{\gamma}{\gamma + \lambda_{K+1}^*} h(P_i^*, C_i) < h(P_j, C_j). \quad (104)$$

Exploiting the equation in (102) to derive a relation between the power-cost pairs (P_i^*, C_i) and (P_j, C_j) , one obtains

$$\frac{\gamma}{\gamma + \lambda_{K+1}^*} (P_i^*, C_i) = (P_j, C_j). \quad (105)$$

In the following, it will be demonstrated that $(P_i^*, C_i) \in \mathcal{S}_i$, where \mathcal{S}_i is as defined in Lemma 2. Since $P_i^* > A_p$, it is sufficient to prove that

$$A_p \geq \max \left\{ \frac{b}{\kappa^2 (e^{C_i} + 1)}, \frac{b C_i}{v} \right\} \quad (106)$$

for (P_i^*, C_i) to be an element of set \mathcal{S}_i . From (68), it is known that $A_p \geq \frac{b \beta_K^2}{\kappa^2 \beta_1 (2\beta_K + b)}$. The inequality

$$\frac{b \beta_K^2}{\kappa^2 \beta_1 (2\beta_K + b)} \geq \frac{b}{\kappa^2 (e^{C_i} + 1)} \quad (107)$$

is reduced to the following form by using the cost function in (64) and performing several algebraic steps:

$$2 \beta_i \beta_K (\beta_K - \beta_1) \geq b (\beta_1 \beta_i - \beta_K^2). \quad (108)$$

Since the channel parameters are ordered in the ascending order of the channel indices, $\beta_K > \beta_1$ and $\beta_K^2 \geq \beta_i^2 \geq \beta_i \beta_1$, which make the left-hand side of (108) greater than zero and the right-hand side of

(108) smaller than or equal to zero. Hence, the inequality in (108) and thus the inequality in (107) is proved to be correct. Therefore, the average power limit satisfies

$$A_p \geq \frac{b}{\kappa^2(e^{C_i} + 1)}. \quad (109)$$

On the other hand, the inequality

$$\frac{b\beta_K \log\left(1 + \frac{b}{\beta_1}\right)}{v\beta_1} \geq \frac{bC_i}{v} \quad (110)$$

can be reduced to

$$\frac{\beta_K}{\beta_1} \geq \frac{C_i}{C_1} \quad (111)$$

after some manipulation and based on the cost function in (64). Since $\beta_K > \beta_1$ and $C_i < C_1$, the inequalities (111) and (110) are satisfied, which, together with the inequality in (68), yields the following result:

$$A_p \geq \frac{bC_i}{v}. \quad (112)$$

Combining the inequalities in (109) and (112) produces the inequality in (106). Therefore, the proof of the argument $(P_i^*, C_i) \in \mathcal{S}_i$ is finalized. Now, one can use Lemma 2 and the equation in (105) to write

$$\frac{\gamma}{\gamma + \lambda_{K+1}^*} h(P_i^*, C_i) < h\left(\frac{\gamma P_i^*}{\gamma + \lambda_{K+1}^*}, \frac{\gamma C_i}{\gamma + \lambda_{K+1}^*}\right) = h(P_j, C_j) \quad (113)$$

which completes the proof of the inequality in (104) and the equivalent one in (103). Hence, given the optimal solution for Strategy 1P, there always exists a full transmission solution performing channel switching between two channels, which leads to a higher average probability of correct decision. Therefore, under the assumption that there exist no channels with cost A_c , the claim that partial transmission is not optimal when the average power limit satisfies the condition in (68) is verified.

For the final part of the proof, assume that there exists a channel with cost A_c . Let $\{\lambda_i^*, P_i^*\}_{i=1}^K$ denote the solution of the optimization problem in (13) under the partial transmission strategies, Strategy 1P or Strategy 2P. Here, at most two of the channel switching factors are non-zero, since only single channel or two-channel strategies are employed. The average probability of correct decision achieved by the optimal

solution $\{\lambda_i^*, P_i^*\}_{i=1}^K$ can be bounded from above based on the strict concavity of $h(P, C)$:

$$\sum_{i=1}^K \lambda_i^* h(P_i^*, C_i) < \left(\sum_{i=1}^K \lambda_i^* \right) h \left(\frac{\sum_{i=1}^K \lambda_i^* P_i^*}{\sum_{i=1}^K \lambda_i^*}, \frac{\sum_{i=1}^K \lambda_i^* C_i}{\sum_{i=1}^K \lambda_i^*} \right). \quad (114)$$

From that part of the proof above which is related to the case in which there exist no channels with cost A_c , it is known that the power-cost pairs obtained from the optimal solution under both Strategy 1P and Strategy 2P are elements of set \mathcal{S}_i defined in Lemma 2. Hence, the power-cost pairs (P_i^*, C_i) corresponding to the employed channels satisfy the concavity condition in Lemma 1, which verifies the inequality in (114). Since \mathcal{S}_i is a convex set, the convex combination of the power-cost pairs (P_i^*, C_i) is also an element of \mathcal{S}_i . Noting that $\sum_{i=1}^K \lambda_i^* < 1$ and using Lemma 2, the following inequality is obtained:

$$\left(\sum_{i=1}^K \lambda_i^* \right) h \left(\frac{\sum_{i=1}^K \lambda_i^* P_i^*}{\sum_{i=1}^K \lambda_i^*}, \frac{\sum_{i=1}^K \lambda_i^* C_i}{\sum_{i=1}^K \lambda_i^*} \right) < h \left(\sum_{i=1}^K \lambda_i^* P_i^*, \sum_{i=1}^K \lambda_i^* C_i \right) = h(A_p, A_c) \quad (115)$$

where the equality follows from the fact that an optimal solution must operate at the average power and the average cost limits. Combining the inequalities in (114) and (115) yields the desired inequality

$$\sum_{i=1}^K \lambda_i^* h(P_i^*, C_i) < h(A_p, A_c) \quad (116)$$

where the left and right hand sides represent, respectively, the average probabilities of correct decision achieved by the optimal partial transmission strategy (whether it be Strategy 1P or Strategy 2P) and Strategy 1F that employs the channel with cost A_c at the maximum average power A_p . Therefore, if there is a channel with cost A_c , partial channel utilization cannot be optimal. In summary, with the verification of the inequalities in (90), (96), (103), and (116), it is concluded that when the average power limit is above a certain threshold as stated in the proposition, the partial transmission strategies cannot be optimal.

■

REFERENCES

- [1] S. Loyka, V. Kostina, and F. Gagnon, "Error rates of the maximum-likelihood detector for arbitrary constellations: Convex/concave behavior and applications," *IEEE Trans. Inform. Theory*, vol. 56, no. 4, pp. 1948–1960, April 2010.
- [2] 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.
- [3] M. Azizoglu, "Convexity properties in binary detection problems," *IEEE Trans. Inform. Theory*, vol. 42, no. 4, pp. 1316–1321, July 1996.

- [4] C. Goken, S. Gezici, and O. Arikan, "Optimal signaling and detector design for power-constrained binary communications systems over non-Gaussian channels," *IEEE Commun. Letters*, vol. 14, no. 2, pp. 100–102, Feb. 2010.
- [5] 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.
- [6] 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, July 2013.
- [7] 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. Letters*, vol. 16, no. 8, pp. 1153–1156, Aug. 2012.
- [8] 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, July 2013.
- [9] B. Dulek, N. D. Vanli, S. Gezici, and P. K. Varshney, "Optimum power randomization for the minimization of outage probability," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4627–4637, Sep. 2013.
- [10] M. E. Tutay, S. Gezici, H. Soganci, and O. Arikan, "Optimal channel switching over Gaussian channels under average power and cost constraints," *IEEE Trans. Commun.*, vol. 63, pp. 1907–1922, May 2015.
- [11] B. Dulek, M. E. Tutay, S. Gezici, and O. Arikan, "Optimal signaling and detector design for M-ary communication systems in the presence of multiple additive noise channels," *Digital Signal Processing*, vol. 26, pp. 153–168, Mar. 2014.
- [12] J. A. Ritcey and M. Azizoglu, "Performance analysis of generalized selection combining with switching constraints," *IEEE Commun. Letters*, vol. 4, no. 5, pp. 152–154, May 2000.
- [13] A. D. Sezer, S. Gezici, and H. Inaltekin, "Optimal channel switching strategy for average capacity maximization," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2143–2157, June 2015.
- [14] A. Ozcelikkale, H. M. Ozaktas, and E. Arikan, "Signal recovery with cost-constrained measurements," *IEEE Trans. on Signal Process.*, vol. 58, no. 7, pp. 3607–3617, 2010.
- [15] T. Alpcan, T. Basar, R. Srikant, and E. Altman, "CDMA uplink power control as a noncooperative game," in *Proceedings of the 40th IEEE Conference on Decision and Control*, vol. 1, 2001, pp. 197–202.
- [16] M. Sayin, N. Vanli, and S. Kozat, "A novel family of adaptive filtering algorithms based on the logarithmic cost," *IEEE Transactions on Signal Processing*, vol. 62, no. 17, pp. 4411–4424, Sep. 2014.
- [17] H. A. (editor), *Cognitive Radio, Software Defined Radio, and Adaptive Wireless Systems*. Springer, 2007.
- [18] 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.
- [19] A. Goldsmith, *Wireless Communications*. Cambridge Univ. Press, 2005.
- [20] J. Huang, R. Berry, and M. L. Honig, "Auction-based spectrum sharing," *ACM/Springer Mobile Networks and Apps.*, pp. 405–418, 2006.
- [21] Z. Ji and K. Liu, "Dynamic spectrum sharing: A game theoretical overview," *IEEE Commun. Mag.*, vol. 45, no. 5, pp. 88–94, May 2007.
- [22] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ: Princeton University Press, 1968.
- [23] H. Chen, P. K. Varshney, S. M. Kay, and J. H. Michels, "Theory of the stochastic resonance effect in signal detection: Part I—Fixed detectors," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3172–3184, July 2007.