

Optimal Channel Switching Strategy for Average Capacity Maximization

Ahmet Dundar Sezer, *Student Member, IEEE*, Sinan Gezici, *Senior Member, IEEE*, and Hazer Inaltekin, *Member, IEEE*

Abstract—In this study, an optimal channel switching strategy is proposed for average capacity maximization in the presence of average and peak power constraints. Necessary and sufficient conditions are derived to determine when the proposed optimal channel switching strategy can or cannot outperform the optimal single channel strategy, which performs no channel switching. Also, it is obtained that the optimal channel switching strategy can be realized by channel switching between, at most, two different channels. In addition, a low-complexity optimization problem is derived to obtain the optimal channel switching strategy. Furthermore, based on some necessary conditions that need to be satisfied by the optimal channel switching solution, an alternative approach is proposed for calculating the optimal channel switching strategy. Numerical examples are provided to exemplify the derived theoretical results and to provide intuitive explanations.

Index Terms—Channel switching, capacity, time sharing.

I. INTRODUCTION

IN recent studies in the literature, benefits of time sharing (“randomization”) have been investigated for various detection and estimation problems [2]–[14]. For instance, in the context of noise enhanced detection and estimation, additive “noise” that is realized by time sharing among a certain number of signal levels can be injected into the input of a suboptimal detector or estimator for performance improvement [2]–[6]. Also, error performance of average power constrained communication systems that operate in non-Gaussian channels can be improved by *stochastic signaling*, which involves time sharing among multiple signal values for each information symbol [9], [10]. It is shown that an optimal stochastic signal can be represented by a randomization (time sharing) among no more than three different signal values under second and fourth moment constraints [9]. In a different context, jammer

systems can achieve improved jamming performance via time sharing among multiple power levels [7], [12], [15]. In [7], it is shown that a weak jammer should employ on-off time sharing to maximize the average probability of error for a receiver that operates in the presence of noise with a symmetric unimodal density. The optimum power allocation policy for an average power constrained jammer operating over an arbitrary additive noise channel is studied in [15], where the aim is to minimize the detection probability of an instantaneously and fully adaptive receiver that employs the Neyman-Pearson criterion. It is proved that the optimum jamming performance is achieved via time sharing between at most two different power levels, and a necessary and sufficient condition is derived for the improvability of the jamming performance via time sharing of the power compared to a fixed power jamming scheme.

Error performance of some communications systems that operate over additive time-invariant noise channels can also be enhanced via time sharing among multiple detectors, which is called *detector randomization* [4], [11], [16]–[18]. In this approach, the receiver employs each detector with a certain time sharing factor (or, probability), and the transmitter adjusts its transmission in coordination with the receiver. In [4], time sharing between two antipodal signal pairs and the corresponding maximum a-posteriori probability (MAP) detectors is studied for an average power constrained binary communication system. Significant performance improvements can be observed as a result of detector randomization in the presence of symmetric Gaussian mixture noise over a range of average power constraint values [4]. In [11], the results in [4] and [10] are extended to an average power constrained M -ary communication system that can employ both detector randomization and stochastic signaling over an additive noise channel with a known distribution. It is obtained that the joint optimization of the transmitted signals and the detectors at the receiver leads to time sharing between at most two MAP detectors corresponding to two deterministic signal constellations. In [13], the benefits of time sharing among multiple detectors are investigated for the downlink of a multiuser communication system and the optimal time sharing strategy is characterized.

In the presence of multiple channels between a transmitter and a receiver, it may be beneficial to perform *channel switching*; that is, to transmit over one channel for a certain fraction of time, and then switch to another channel for the next transmission period [7], [19]–[22]. In [7], the channel switching problem is investigated in the presence of an average

Manuscript received October 1, 2014; revised January 28, 2015 and April 2, 2015; accepted April 7, 2015. Date of publication April 14, 2015; date of current version June 12, 2015. Part of this work was presented at IEEE ICASSP 2014. This research was supported in part by the Distinguished Young Scientist Award of Turkish Academy of Sciences (TUBA-GEBIP 2013). The associate editor coordinating the review of this paper and approving it for publication was A. Khisti.

A. D. Sezer and S. Gezici are with the Department of Electrical and Electronics Engineering, Bilkent University, Ankara 06800, Turkey (e-mail: adsezer@ee.bilkent.edu.tr; gezici@ee.bilkent.edu.tr).

H. Inaltekin is with the Department of Electrical and Electronics Engineering, Antalya International University, Antalya 147004, Turkey (e-mail: hazeri@antalya.edu.tr).

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

Digital Object Identifier 10.1109/TCOMM.2015.2422813

power constraint for the optimal detection of binary antipodal signals over a number of channels that are subject to additive unimodal noise. It is proved that the optimal strategy is either to communicate over one channel exclusively, or to switch between two channels with a certain time sharing factor. In [21], the channel switching problem is studied for M -ary communications over additive noise channels (with arbitrary probability distributions) in the presence of time sharing among multiple signal constellations over each channel. It is shown that the optimal strategy that minimizes the average probability of error under an average power constraint corresponds to one of the following approaches: deterministic signaling (i.e., use of one signal constellation) over a single channel; time sharing between two different signal constellations over a single channel; or switching (time sharing) between two channels with deterministic signaling over each channel [21]. With a different perspective, the concept of channel switching is studied for cognitive radio systems in the context of opportunistic spectrum access, where a number of secondary users aim to access the available frequency bands in the spectrum [23]–[26]. In [26], the optimal bandwidth allocation is studied for secondary users in the presence of multiple available primary user bands and under channel switching constraints, and it is shown that secondary users switching among discrete channels can achieve higher capacity than those that switch among consecutive channels.

In a different but related problem, the capacity of the *sum channel* is presented in [27, p. 525]. The sum channel is defined as a channel whose input and output alphabets are the unions of those of the original channels; that is, there exist multiple available channels between the transmitter and the receiver but only one channel is used at a given time for each possible symbol in the input alphabet. For example, a sum channel can consist of two binary memoryless channels, and the first two elements of the alphabet, say $\{0,1\}$, are transmitted over the first channel whereas the last two elements of the alphabet, say $\{2,3\}$, are transmitted over the second channel. For discrete memoryless channels with capacities C_1, C_2, \dots, C_K , the capacity of the sum channel can be obtained as $\log_2(\sum_{i=1}^K 2^{C_i})$ [27]. The main difference of the sum channel from the channel switching scenario considered in this study (and those in [7], [21]) is that the alphabet is divided among different channels and each channel is used to transmit a certain subset of the alphabet in the sum channel.

In the literature, optimal resource allocation is commonly employed to enhance the capacity of communication systems. In [28], the optimal dynamic resource allocation for fading broadcast channels is studied for code division, time division, and frequency division in the presence of perfect channel side information at the transmitter and the receivers, and ergodic capacity regions are obtained. In [29], an adaptive resource allocation procedure is presented for multiuser orthogonal frequency division multiplexing (MU-OFDM) systems with the consideration of proportional fairness constraints among users. Optimal and suboptimal algorithms are implemented based on sum capacity maximization while satisfying the minimum required data rate constraint for each user. In [30], optimal joint power and channel allocation strategies are investigated for cognitive

radio systems. A near optimal algorithm is presented for the total sum capacity maximization of power-limited secondary users in a centralized cognitive radio network. In [31], capacity maximizing antenna selection is studied for a multiple-input multiple-output (MIMO) system and low-complexity antenna subset selection algorithms are derived. It is shown that near optimal capacity of a full-complexity system is achieved by selecting the number of antennas at the receiver to be at least as large as the number of antennas at the transmitter. In [32], the optimal antenna selection in correlated channels is analyzed for both the transmitter and receiver to reduce the number of radio frequency chains. The proposed algorithm results in a near optimal capacity which is achieved without antenna selection.

Although the optimal channel switching problem is studied thoroughly in terms of average probability of error minimization (e.g., [7], [21], [22]) and in the context of opportunistic spectrum access (e.g., [23]–[26]), no studies in the literature have considered the channel switching problem for maximization of data rates by jointly optimizing time sharing (channel switching) factors and corresponding power levels (please see [1] for the conference version of this study). In this paper, the average Shannon capacity is considered as the main metric since it gives the maximum achievable data rates with low probability of decoding errors. In addition, the data rate targets indicated by the Shannon capacity are achievable in practical communication systems through turbo coding or low density parity check codes [33]. In this study, we formulate the optimal channel switching problem for average Shannon capacity maximization over Gaussian channels in the presence of average and peak power constraints, and derive necessary and sufficient conditions for the proposed channel switching approach to achieve a higher average capacity than the optimal approach without channel switching. In addition, it is obtained that the optimal solution to the channel switching problem results in channel switching between at most two different channels, and an approach is proposed to obtain the optimal channel switching strategy with low computational complexity. Numerical examples are presented to illustrate the theoretical results. The main contributions of this study can be summarized as follows:

- For the first time, the optimal channel switching problem is investigated *for average capacity maximization* in the presence of multiple Gaussian channels and under average and peak power constraints.
- It is shown that the optimal channel switching strategy switches among at most two different channels, and operates at the average power limit.
- Necessary and sufficient conditions are derived to specify when performing channel switching can or cannot provide improvements over the optimal approach without channel switching.
- Optimality conditions are obtained for the proposed channel switching strategy, and an approach with low computational complexity is presented for calculating the parameters of the optimal strategy.

Some of the practical motivations for studying the channel switching problem for data rate maximization can be

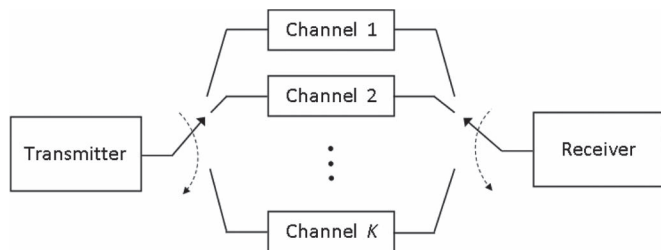


Fig. 1. Block diagram of a communication system in which transmitter and receiver can switch among K channels.

summarized as follows: Firstly, the next-generation wireless communication systems are required to support all IP services including high-data-rate multimedia traffic, with bit rate targets as high as 1 Gbit/s for low mobility and 100 Mbit/s for high mobility [34]. Such high data rate requirements make the capacity (usually measured by using Shannon capacity metric [35], [36]) maximization problems (subject to appropriate operating constraints on power and communication reliability) more relevant for next-generation wireless communication systems, rather than focusing on power or bit error minimization (subject to appropriate operating constraints on rate). Secondly, wireless telecommunication technology is currently on the cusp of a major transition from the traditional carefully planned homogenous macro-cell deployment to highly heterogeneous small cell network architectures. These heterogeneous next generation network architectures (alternatively called HetNets) will consist of multiple tiers of irregularly deployed network elements with diverse range of backhaul connection characteristics, signal processing capabilities and electromagnetic radio emission levels. In such a HetNet scenario, it is expected that more than one radio link such as femto-cell connection, macro-cell connection and Wi-Fi connection (with different operating frequency bands, background noise levels and etc.) will be present to use at each mobile user. From an engineering point of view, this paper provides some fundamental design insights regarding how to time share (randomize) among available radio links to maximize rates of communication for highly heterogeneous wireless environments. Finally, channel switching can be beneficial for secondary users in a cognitive radio system in which there can exist multiple available frequency bands in the spectrum (please see the second paragraph of Section II).

The remainder of the paper is organized as follows: The problem formulation for optimal channel switching is presented in Section II. Section III investigates the solution of the optimal channel switching problem and provides various theoretical results about the characteristics of the optimal channel switching strategy. In Section IV, numerical examples are presented for illustrating the theoretical results, which is followed by the concluding remarks in Section V.

II. PROBLEM FORMULATION

Consider a communication system in which a transmitter and a receiver are connected via K different channels as illustrated in Fig. 1. The channels are modeled as additive Gaussian noise channels with possibly different noise levels and bandwidths. It

is assumed that noise is independent across different channels. The transmitter and the receiver can switch (time share) among these K channels to enhance the capacity of the communication system. A relay at the transmitter controls the access to the channels in such a way that only one of the channels can be used for information transmission at any given time. It is assumed that the transmitter and the receiver are synchronized and the receiver knows which channel is being utilized [7]. In practical scenarios, this assumption can hold in the presence of a communication protocol that notifies the receiver about the numbers of symbols and the corresponding channels to be employed during data communications. This notification information can be sent in the header of a communications packet [11], [21].

In some communication systems, multiple channels with various bandwidth and noise characteristics can be available between a transmitter and a receiver as in Fig. 1. For instance, in a cognitive radio system, primary users are the main owners of the spectrum, and secondary users can utilize the frequency bands of the primary users when they are available [23]–[25], [37], [38]. In such a case, the available bands in the spectrum can be considered as the channels in Fig. 1, and the aim of a secondary user becomes the maximization of its average capacity by performing optimal channel switching under power constraints that are related to hardware constraints and/or battery life. The motivation for using only one channel at a given time is that the transmitter and the receiver are assumed to have a single RF chain each due to complexity/cost considerations. Then, the transmitter-receiver pair can perform time sharing among different channels (i.e., channel switching) by employing only one channel at a given time. In a similar fashion, the proposed system also has a potential to improve data rates in emerging open-access K -tier heterogeneous wireless networks by allowing users to switch between multiple access points and available frequency bands in the spectrum [39], [40].

Let B_i and $N_i/2$ represent, respectively, the bandwidth and the constant power spectral density level of the additive Gaussian noise corresponding to channel i for $i \in \{1, \dots, K\}$. Then, the capacity of channel i is given by

$$C_i(P) = B_i \log_2 \left(1 + \frac{P}{N_i B_i} \right) \text{ bits/sec} \quad (1)$$

where P denotes the average transmit power [41].

The aim of this study is to obtain the optimal channel switching strategy that maximizes the average capacity of the communication system in Fig. 1 under average and peak power constraints. To formulate such a problem, channel switching (time sharing) factors, denoted by $\lambda_1, \dots, \lambda_K$, are defined first, where λ_i is the fraction of time when channel i is used, with $\lambda_i \geq 0$ for $i = 1, \dots, K$, and $\sum_{i=1}^K \lambda_i = 1$.¹ Then, the optimal

¹Channel switching can be implemented in practice by transmitting the first $\lambda_1 N_s$ symbols over channel 1, the next $\lambda_2 N_s$ symbols over channel 2, ..., and the final $\lambda_K N_s$ symbols over channel K , where N_s is the total number of symbols (over which channel statistics do not change), and $\lambda_1, \lambda_2, \dots, \lambda_K$ are the channel switching factors. In this case, suitable channel coding-decoding algorithms can be employed for each channel to achieve a data rate close to the Shannon capacity of that channel.

channel switching problem for average capacity maximization is proposed as follows:

$$\begin{aligned} & \max_{\{\lambda_i, P_i\}_{i=1}^K} \sum_{i=1}^K \lambda_i C_i(P_i) \\ & \text{subject to } \sum_{i=1}^K \lambda_i P_i \leq P_{\text{av}} \\ & P_i \in [0, P_{\text{pk}}], \forall i \in \{1, \dots, K\} \\ & \sum_{i=1}^K \lambda_i = 1, \lambda_i \geq 0, \forall i \in \{1, \dots, K\} \end{aligned} \quad (2)$$

where $C_i(P_i)$ is as defined in (1) with P_i denoting the average transmit power allocated to channel i , P_{pk} represents the peak power limit, and P_{av} is the average power limit for the transmitter. In practical systems, the average power limit is related to the power consumption and/or the battery life of the transmitter whereas the peak power limit specifies the maximum power level that can be generated by the transmitter circuitry; i.e., it is mainly a hardware constraint. Since there exists a single RF unit at the transmitter, the peak power limit is taken to be the same for each channel. It is assumed that $P_{\text{av}} < P_{\text{pk}}$ holds. From (2), it is observed that the design of an optimal channel switching strategy involves the joint optimization of the channel switching factors and the corresponding power levels under average and peak power constraints for the purpose of average capacity maximization.

III. OPTIMAL CHANNEL SWITCHING

In general, it is challenging to find the optimal channel switching strategy by directly solving the optimization problem in (2). For this reason, our aim is to obtain a simpler version of the problem in (2) and to calculate the optimal channel switching solution in a low-complexity manner. To that end, an alternative optimization problem is obtained first. Let $\{\lambda_i^*, P_i^*\}_{i=1}^K$ denote the optimal channel switching strategy obtained as the solution of (2) and define C^* as the corresponding maximum average capacity; that is, $C^* = \sum_{i=1}^K \lambda_i^* C_i(P_i^*)$. Then, the following proposition presents an alternative optimization problem, the solution of which achieves the same maximum average capacity as (2) does.

Proposition 1: The solution of the following optimization problem results in the same maximum value that is achieved by the problem in (2):

$$\begin{aligned} & \max_{\{\nu_i, P_i\}_{i=1}^K} \sum_{i=1}^K \nu_i C_{\text{max}}(P_i) \\ & \text{subject to } \sum_{i=1}^K \nu_i P_i \leq P_{\text{av}} \\ & P_i \in [0, P_{\text{pk}}], \forall i \in \{1, \dots, K\} \\ & \sum_{i=1}^K \nu_i = 1, \nu_i \geq 0, \forall i \in \{1, \dots, K\} \end{aligned} \quad (3)$$

where $C_{\text{max}}(P)$ is defined as

$$C_{\text{max}}(P) \triangleq \max\{C_1(P), \dots, C_K(P)\}. \quad (4)$$

Proof: The proof consists of two steps. Let $\{\nu_i^*, P_i^*\}_{i=1}^K$ represent the solution of (3) and define C^* as the corresponding maximum average capacity; that is, $C^* = \sum_{i=1}^K \nu_i^* C_{\text{max}}(P_i^*)$. First, it can be observed from (2) and (3) that $C^* \geq C^*$ due to the definition in (4), where C^* is the maximum average capacity obtained from (2). Next, define function $g(i)$ and set S_m as follows:²

$$g(i) \triangleq \arg \max_{l \in \{1, \dots, K\}} C_l(P_i^*), \forall i \in \{1, \dots, K\} \quad (5)$$

and

$$S_m \triangleq \{i \in \{1, \dots, K\} \mid g(i) = m\}, \forall m \in \{1, \dots, K\}. \quad (6)$$

Then, the following relations can be obtained for C^* :

$$C^* = \sum_{i=1}^K \nu_i^* C_{\text{max}}(P_i^*) = \sum_{i=1}^K \nu_i^* C_{g(i)}(P_i^*) \quad (7)$$

$$= \sum_{i=1}^K \sum_{k \in S_i} \nu_k^* C_i(P_k^*) \quad (8)$$

$$\leq \sum_{i=1}^K \left(\sum_{k \in S_i} \nu_k^* \right) C_i \left(\frac{\sum_{k \in S_i} \nu_k^* P_k^*}{\sum_{k \in S_i} \nu_k^*} \right) \quad (9)$$

$$= \sum_{i=1}^K \bar{\lambda}_i C_i(\bar{P}_i) \quad (10)$$

where $\bar{\lambda}_i$ and \bar{P}_i are defined as

$$\bar{\lambda}_i \triangleq \sum_{k \in S_i} \nu_k^* \text{ and } \bar{P}_i \triangleq \frac{\sum_{k \in S_i} \nu_k^* P_k^*}{\sum_{k \in S_i} \nu_k^*}. \quad (11)$$

for $i \in \{1, \dots, K\}$. The equalities in (7) and (8) are obtained from the definitions in (5) and (6), respectively, and the inequality in (9) follows from Jensen's inequality due to the concavity of the capacity function [41], [42]. It is noted from (11), based on (5) and (6), that $\bar{\lambda}_i$'s and \bar{P}_i 's satisfy the constraints in (2); that is, $\sum_{i=1}^K \bar{\lambda}_i \bar{P}_i \leq P_{\text{av}}$, $\bar{P}_i \in [0, P_{\text{pk}}]$, $\forall i \in \{1, \dots, K\}$, $\sum_{i=1}^K \bar{\lambda}_i = 1$, and $\bar{\lambda}_i \geq 0$, $\forall i \in \{1, \dots, K\}$. Therefore, the inequality in (7)–(10), namely, $C^* \leq \sum_{i=1}^K \bar{\lambda}_i C_i(\bar{P}_i)$, implies that the optimal solution of (3) cannot achieve a higher average capacity than that achieved by (2); that is, $C^* \leq C^*$. Hence, it is concluded that $C^* = C^*$ since $C^* \geq C^*$ must also hold as mentioned at the beginning of the proof. ■

Based on Proposition 1, the maximum average capacity C^* achieved by the optimal channel switching problem in (2) can also be obtained by solving the optimization problem in (3). Let $\{\nu_i^*, P_i^*\}_{i=1}^K$ denote the optimal solution of (3).

²In the case of multiple maximizers in (5), any maximizing index can be chosen for $g(i)$.

Proposition 1 states that $\sum_{i=1}^K \nu_i^* C_{\max}(P_i^*) = C^*$. In addition, the optimal channel switching strategy corresponding to the channel switching problem in (2) can be obtained, based on the arguments in the proof of Proposition 1, as follows: Once $\{\nu_i^*, P_i^*\}_{i=1}^K$ is calculated from (3), the optimal channel switching strategy can be obtained as $\{\lambda_i^*, P_i^*\}_{i=1}^K$, where $\lambda_i^* = \sum_{k \in S_i} \nu_k^*$ and $P_i^* = (\sum_{k \in S_i} \nu_k^* P_k^*) / (\sum_{k \in S_i} \nu_k^*)$ with S_i being given by (6). It should be emphasized that a low-complexity approach is developed in the remainder of this section for solving (3); hence, it is useful to obtain the optimal channel switching strategy corresponding to the channel switching problem in (2) based on the solution of (3).

The significance of Proposition 1 also lies in the fact that the alternative optimization problem in (3), which achieves the same maximum average capacity as the original channel switching problem in (2), facilitates detailed theoretical investigations of the optimal channel switching strategy, as discussed in the remainder of this section.

Towards the purpose of characterizing the optimal channel switching strategy, the following lemma is presented first, which states that the optimal solutions of (2) and (3) operate at the average power limit.

Lemma 1: Let $\{\lambda_i^*, P_i^*\}_{i=1}^K$ and $\{\nu_i^*, P_i^*\}_{i=1}^K$ denote the solutions of the optimization problems in (2) and (3), respectively. Then, $\sum_{i=1}^K \lambda_i^* P_i^* = P_{\text{av}}$ and $\sum_{i=1}^K \nu_i^* P_i^* = P_{\text{av}}$ hold.

Proof: The proof is provided for the optimization problem in (3) only since the one for (2) can easily be obtained based on a similar approach (cf. Proposition 1 in [22]). Suppose that $\{\nu_i, P_i\}_{i=1}^K$ is an optimal solution of the problem in (3) such that $\sum_{i=1}^K \nu_i P_i < P_{\text{av}}$. Since $P_{\text{av}} < P_{\text{pk}}$, there exist at least one P_i that is strictly smaller than P_{pk} . Let P_l be one of them. Then, consider an alternative solution $\{\nu'_i, P'_i\}_{i=1}^K$, with $\nu'_i = \nu_i, \forall i \in \{1, \dots, K\}$, $P'_i = P_i, \forall i \in \{1, \dots, K\} \setminus \{l\}$, and $P'_l = \min\{P_{\text{pk}}, P_l + (P_{\text{av}} - \sum_{i=1}^K \nu_i P_i) / \nu_l\}$. Note that the alternative solution, $\{\nu'_i, P'_i\}_{i=1}^K$, achieves a larger average capacity than $\{\nu_i, P_i\}_{i=1}^K$ due to the following relation:

$$\sum_{i=1}^K \nu'_i C_{\max}(P'_i) = \sum_{\substack{i=1 \\ i \neq l}}^K \nu'_i C_{\max}(P'_i) + \nu'_l C_{\max}(P'_l) \quad (12)$$

$$> \sum_{\substack{i=1 \\ i \neq l}}^K \nu_i C_{\max}(P_i) + \nu_l C_{\max}(P_l) \quad (13)$$

$$= \sum_{i=1}^K \nu_i C_{\max}(P_i) \quad (14)$$

where the inequality follows from the facts that $C_{\max}(P)$ is a monotone increasing function of P (please see (1) and (4)),³ and that $P'_l > P_l$. Therefore, $\{\nu_i, P_i\}_{i=1}^K$ cannot be an optimal solution of (3), which leads to a contradiction. Hence, any feasible point of the problem in (3) which utilizes an average power strictly smaller than P_{av} cannot be optimal; that is, the optimal solution must operate at the average power limit. ■

³Note that the maximum of a set of monotone increasing functions is also monotone increasing.

A. Optimal Channel Switching versus Optimal Single Channel Strategy

Next, possible improvements that can be achieved via the optimal channel switching strategy over the *optimal single channel* strategy are investigated. The optimal single channel strategy corresponds to the case of no channel switching and the use of the best channel all the time at the average power limit. For that strategy, the achieved maximum capacity can be expressed as $C_{\max}(P_{\text{av}})$, where C_{\max} is as defined in (4), and the best channel is the one with the index $\arg \max_{l \in \{1, \dots, K\}} C_l(P_{\text{av}})$.⁴ It is noted that when a single channel is used (i.e., no channel switching), it is optimal to utilize all the available power, P_{av} since $C_{\max}(P)$ is a monotone increasing and continuous function of P , as can be verified from (1) and (4). In the following proposition, a necessary and sufficient condition is presented for the optimal channel switching strategy to have the same performance as the optimal single channel strategy.

Proposition 2: Suppose that $C_{\max}(P)$ in (4) is first-order continuously differentiable in an interval around P_{av} . Then, the optimal channel switching and the optimal single channel strategies achieve the same maximum average capacity if and only if

$$(P - P_{\text{av}}) \frac{B_{i^*} \log_2 e}{N_{i^*} B_{i^*} + P_{\text{av}}} \geq C_{\max}(P) - C_{\max}(P_{\text{av}}) \quad (15)$$

for all $P \in [0, P_{\text{pk}}]$, where $i^* = \arg \max_{i \in \{1, \dots, K\}} C_i(P_{\text{av}})$.

Proof: The proof consists of the sufficiency and the necessity parts. The sufficiency of the condition in (15) can be proved by employing a similar approach to that in the proof of Proposition 3 in [15]. Under the condition in the proposition, the aim is to prove that the optimal channel switching and the optimal single channel strategies achieve the same maximum average capacity; that is, $\sum_{i=1}^K \nu_i^* C_{\max}(P_i^*) = C_{\max}(P_{\text{av}})$, where $\{\nu_i^*, P_i^*\}_{i=1}^K$ denotes the solution of (3), which achieves the same average capacity as the optimal channel switching strategy corresponding to (2) based on Proposition 1. Due to the assumption in the proposition, the first-order derivative of $C_{\max}(P)$ in (4) exists in an interval around P_{av} and its value at P_{av} is calculated from (1) as

$$C'_{\max}(P_{\text{av}}) = \frac{B_{i^*} \log_2 e}{N_{i^*} B_{i^*} + P_{\text{av}}} \quad (16)$$

where $i^* = \arg \max_{i \in \{1, \dots, K\}} C_i(P_{\text{av}})$. From (16), the condition in (15) can be expressed as $C_{\max}(P) \leq C_{\max}(P_{\text{av}}) + C'_{\max}(P_{\text{av}})(P - P_{\text{av}})$ for all $P \in [0, P_{\text{pk}}]$. Then, for any channel switching strategy denoted as $\{\nu_i, P_i\}_{i=1}^K$, the following inequalities can be obtained:

$$\sum_{i=1}^K \nu_i C_{\max}(P_i) \leq C_{\max}(P_{\text{av}}) + C'_{\max}(P_{\text{av}}) \left(\sum_{i=1}^K \nu_i P_i - P_{\text{av}} \right) \quad (17)$$

$$\leq C_{\max}(P_{\text{av}}) \quad (18)$$

⁴In the case of multiple best channels, any of them can be chosen to achieve $C_{\max}(P_{\text{av}})$.

where $P_i \in [0, P_{\text{pk}}]$ and $\nu_i \geq 0$ for $i \in \{1, \dots, K\}$, $\sum_{i=1}^K \nu_i = 1$, and $\sum_{i=1}^K \nu_i P_i \leq P_{\text{av}}$. It is noted that the inequality in (18) is obtained from the facts that $C'_{\text{max}}(P_{\text{av}})$ in (16) is positive and that $\sum_{i=1}^K \nu_i P_i - P_{\text{av}}$ is non-positive due to the average power constraint. From (17) and (18), it is concluded that when the condition in the proposition holds, channel switching can never result in a higher average capacity than the optimal single channel strategy, which achieves a capacity of $C_{\text{max}}(P_{\text{av}})$. On the other hand, for $\nu_{i^*}^* = 1$, $P_{i^*}^* = P_{\text{av}}$, and $\nu_i^* = P_i^* = 0$ for all $i \in \{1, \dots, K\} \setminus \{i^*\}$, where $i^* = \arg \max_{i \in \{1, \dots, K\}} C_i(P_{\text{av}})$, the $\sum_{i=1}^K \nu_i C_{\text{max}}(P_i)$ term in (17) becomes equal to $C_{\text{max}}(P_{\text{av}})$. Since this possible solution satisfies $\sum_{i=1}^K \nu_i^* P_i^* = P_{\text{av}}$ (cf. Lemma 1) and all the constraints of the optimization problem in (3), it is concluded that $\sum_{i=1}^K \nu_i^* C_{\text{max}}(P_i^*) = C_{\text{max}}(P_{\text{av}})$ under the condition in the proposition.

The necessity part of the proof is contrapositive. Therefore, the aim is to prove that if

$$(P - P_{\text{av}})C'_{\text{max}}(P_{\text{av}}) < C_{\text{max}}(P) - C_{\text{max}}(P_{\text{av}}) \quad (19)$$

for some $P \in [0, P_{\text{pk}}]$, then the optimal channel switching strategy outperforms the optimal single channel strategy in terms of the maximum average capacity. First, assume that there exists $\tilde{P} \in [0, P_{\text{av}}]$ that satisfies the condition in (19) and consider the straight line that passes through the points $(\tilde{P}, C_{\text{max}}(\tilde{P}))$ and $(P_{\text{av}}, C_{\text{max}}(P_{\text{av}}))$. Let φ denote the slope of this line. From (19), the following relation is observed:

$$\varphi \triangleq \frac{C_{\text{max}}(P_{\text{av}}) - C_{\text{max}}(\tilde{P})}{P_{\text{av}} - \tilde{P}} < C'_{\text{max}}(P_{\text{av}}). \quad (20)$$

Due to the assumption in the proposition, the first-order derivative of $C_{\text{max}}(P)$ in (4) is continuous in an interval around P_{av} . Therefore, $C_{\text{max}}(P)$ must correspond to the same channel over an interval around P_{av} ,⁵ which implies the concavity of $C_{\text{max}}(P)$ in that interval as the capacity curves are concave. By definition of the concavity around P_{av} , there exists a point $P_{\text{av}}^+ \triangleq P_{\text{av}} + \epsilon$ for an infinitesimally small positive number ϵ such that

$$\varphi < \frac{C_{\text{max}}(P_{\text{av}}) - C_{\text{max}}(P_{\text{av}}^+)}{P_{\text{av}} - P_{\text{av}}^+} < C'_{\text{max}}(P_{\text{av}}). \quad (21)$$

Then, choose a $\tilde{\lambda}$ such that $\tilde{\lambda}\tilde{P} + (1 - \tilde{\lambda})P_{\text{av}}^+ = P_{\text{av}}$ and consider the following relations:

$$\begin{aligned} & \tilde{\lambda}C_{\text{max}}(\tilde{P}) + (1 - \tilde{\lambda})C_{\text{max}}(P_{\text{av}}^+) \\ & > \tilde{\lambda}C_{\text{max}}(\tilde{P}) + (1 - \tilde{\lambda})((P_{\text{av}}^+ - P_{\text{av}})\varphi + C_{\text{max}}(P_{\text{av}})) \quad (22) \end{aligned}$$

$$\begin{aligned} & = \frac{P_{\text{av}}^+ - P_{\text{av}}}{P_{\text{av}}^+ - \tilde{P}} C_{\text{max}}(\tilde{P}) \\ & + \frac{P_{\text{av}} - \tilde{P}}{P_{\text{av}}^+ - \tilde{P}} ((P_{\text{av}}^+ - P_{\text{av}})\varphi + C_{\text{max}}(P_{\text{av}})) \quad (23) \\ & = C_{\text{max}}(P_{\text{av}}) \quad (24) \end{aligned}$$

⁵If there multiple channels with the same bandwidths and noise levels, they can be regarded as a single channel (i.e., only one of them should be considered) since there is no advantage of switching between such channels.

where the inequality in (22) is obtained from (21), the equality in (23) follows from the definition of $\tilde{\lambda}$, and the final equality is due to the definition of φ in (20). Overall, the inequality in (22)–(24), namely, $\tilde{\lambda}C_{\text{max}}(\tilde{P}) + (1 - \tilde{\lambda})C_{\text{max}}(P_{\text{av}}^+) > C_{\text{max}}(P_{\text{av}})$, implies that the channel switching strategy (specified by channel switching factors $\tilde{\lambda}$ and $(1 - \tilde{\lambda})$ and power levels \tilde{P} and P_{av}^+) achieves a higher average capacity than the optimal single channel strategy.⁶ Since the optimal channel switching strategy always achieves an average capacity that is equal to or larger than the average capacity of any other channel switching strategy, it is concluded that the optimal channel switching strategy outperforms the optimal single channel strategy.

Next, assume that there exists $\bar{P} \in (P_{\text{av}}, P_{\text{pk}}]$ that satisfies the condition in (19). Similar to the previous part of the proof, let ϕ denote the slope of the straight line that passes through the points $(\bar{P}, C_{\text{max}}(\bar{P}))$ and $(P_{\text{av}}, C_{\text{max}}(P_{\text{av}}))$. Then, the following expression is obtained from (19):

$$\phi \triangleq \frac{C_{\text{max}}(P_{\text{av}}) - C_{\text{max}}(\bar{P})}{P_{\text{av}} - \bar{P}} > C'_{\text{max}}(P_{\text{av}}). \quad (25)$$

Similarly, due to the concavity around P_{av} , there exists a point $P_{\text{av}}^- \triangleq P_{\text{av}} - \epsilon$ for an infinitesimally small $\epsilon > 0$ such that

$$\phi > \frac{C_{\text{max}}(P_{\text{av}}) - C_{\text{max}}(P_{\text{av}}^-)}{P_{\text{av}} - P_{\text{av}}^-} > C'_{\text{max}}(P_{\text{av}}). \quad (26)$$

By choosing a $\bar{\lambda} \in (0, 1)$ such that $\bar{\lambda}\bar{P} + (1 - \bar{\lambda})P_{\text{av}}^- = P_{\text{av}}$ and considering the expressions in (25) and (26), the same approach employed in the previous part of the proof (see (22)–(24)) can be applied to show that the optimal channel switching strategy outperforms the optimal single channel strategy. Thus, it is concluded that when the condition in Proposition 2 is not satisfied, the optimal single channel strategy achieves a smaller average capacity than the optimal channel switching strategy, which implies that the condition in the proposition is necessary to achieve the same maximum average capacity for both strategies. ■

A more intuitive description of Proposition 2 can be provided as follows: Based on (16), the condition in (15) is equivalent to having the tangent line to $C_{\text{max}}(P)$ at $P = P_{\text{av}}$ lie completely above the $C_{\text{max}}(P)$ curve [15]. If this condition is satisfied, then channel switching, which performs convex combination of different $C_{\text{max}}(P)$ values (as can be noted from (3)), cannot achieve an average capacity above $C_{\text{max}}(P_{\text{av}})$, which is already achieved by the optimal single channel strategy. Otherwise, a higher average capacity than $C_{\text{max}}(P_{\text{av}})$ is obtained via optimal channel switching.

It is also noted from (15) and (16) that the condition in Proposition 2 corresponds to the subgradient inequality at P_{av} . Therefore, the proposition can also be stated as “the optimal channel switching and the optimal single channel strategies achieve the same maximum average capacity if and only if

⁶Note that the channel switching strategy denoted by channel switching factors $\tilde{\lambda}$ and $(1 - \tilde{\lambda})$ and power levels \tilde{P} and P_{av}^+ must involve switching between two different channels since the inequality $\tilde{\lambda}C_{\text{max}}(\tilde{P}) + (1 - \tilde{\lambda})C_{\text{max}}(P_{\text{av}}^+) > C_{\text{max}}(P_{\text{av}})$ cannot be satisfied for a single channel due to the concavity of the capacity curves.

there exists a sub-gradient at P_{av} ." In addition, it should be emphasized that although concavity of $C_{\max}(P)$ around $P = P_{av}$ is a necessary condition for the scenario in Proposition 2 to hold, it is not a sufficient condition in general.

Based on Proposition 2, it can be determined whether the channel switching strategy can improve the average capacity of the system compared to the optimal single channel strategy. For instance, if $C_{\max}(P)$ in (4) is first-order continuously differentiable in an interval around P_{av} and the condition in (15) is satisfied for all $P \in [0, P_{pk}]$ in a given system, then it is concluded that the optimal single channel strategy has the same performance as the optimal channel switching strategy; that is, there is no need for channel switching. In that case, the maximum average channel capacity is given by $C_{\max}(P_{av})$. On the other hand, if there exist some $P \in [0, P_{pk}]$ for which the condition in (15) is not satisfied, then the optimal channel switching strategy is guaranteed to achieve a higher average capacity than $C_{\max}(P_{av})$.

Remark 1: As a special case, it can be concluded from Proposition 2 that if the bandwidths of the channels are the same, the optimal strategy is to transmit over the least noisy (best) channel exclusively at the average power limit. To make this conclusion, first consider $C_{\max}(P)$ in (4), which becomes equal to the capacity of the least noisy channel, say channel b , when the channels have the same bandwidth (cf. (1)); that is, $C_{\max}(P) \triangleq \max\{C_1(P), \dots, C_K(P)\} = C_b(P)$. Then, from (16), the condition in (15) of Proposition 2 is expressed as $(P - P_{av})C'_b(P_{av}) \geq C_b(P) - C_b(P_{av})$, which always holds for all $P \in [0, P_{pk}]$ due to the concavity of the capacity function, $C_b(P)$ (see (1)). Hence, Proposition 2 applies in this scenario; that is, the optimal single channel strategy (i.e., the use of the best channel all the time at the average power limit) becomes the optimal solution.

In Proposition 2, it is assumed that $C_{\max}(P)$ in (4) is first-order continuously differentiable in an interval around P_{av} . To cover all possible scenarios and to specify the optimal strategy in all cases, the following proposition presents a result for the case of $C_{\max}(P)$ that has a discontinuous first-order derivative at $P = P_{av}$, which states that the optimal channel switching always outperforms the optimal single channel strategy in this scenario.

Proposition 3: If the first-order derivative of $C_{\max}(P)$ in (4) is discontinuous at $P = P_{av}$, then the optimal channel switching strategy outperforms the optimal single channel strategy.

Proof: The aim is to prove that if the condition in Proposition 3 is satisfied, then the channel switching strategy achieves a higher average capacity than the optimal single channel strategy. To that aim, define P_{av}^+ and P_{av}^- as $P_{av} + \epsilon$ and $P_{av} - \epsilon$, respectively, where ϵ is an infinitesimally small positive number. The proof consists of two parts.

First, it is proved that if the first-order derivative, $C'_{\max}(P)$, is discontinuous at $P = P_{av}$, which implies that $C'_{\max}(P_{av}^-) \neq C'_{\max}(P_{av}^+)$, then $C'_{\max}(P_{av}^-) < C'_{\max}(P_{av}^+)$ holds. Due to the discontinuous first-order derivative assumption, $C_{\max}(P_{av}^-)$ and $C_{\max}(P_{av}^+)$ must correspond to different channels since the first-order derivative would be continuous otherwise (please see (1)). Therefore, let channel i and channel j denote the channels corresponding to

the maximum capacities for power levels P_{av}^- and P_{av}^+ , respectively; that is, $C_{\max}(P_{av}^-) = C_i(P_{av}^-)$ and $C_{\max}(P_{av}^+) = C_j(P_{av}^+)$ for $i \neq j$ where $i = \arg \max_{l \in \{1, \dots, K\}} C_l(P_{av}^-)$ and $j = \arg \max_{l \in \{1, \dots, K\}} C_l(P_{av}^+)$. Also, $C_i(P_{av}^-) = C_j(P_{av}^-)$ and $C_i(P_{av}^-) < C_j(P_{av}^+)$ since $C_{\max}(\cdot)$ is a continuous monotone increasing function. Based on Taylor series expansions of $C_i(\cdot)$ and $C_j(\cdot)$ around P_{av}^+ , $C_i(P_{av}^+)$ and $C_j(P_{av}^+)$ can be expressed as follows:

$$C_i(P_{av}^+) = C_i(P_{av}) + C'_i(P_{av})(P_{av}^+ - P_{av}) + R_i(P_{av}^+) \quad (27)$$

$$C_j(P_{av}^+) = C_j(P_{av}) + C'_j(P_{av})(P_{av}^+ - P_{av}) + R_j(P_{av}^+) \quad (28)$$

where $R_i(P_{av}^+)$ and $R_j(P_{av}^+)$ are the second-order remainder terms for $C_i(P_{av}^+)$ and $C_j(P_{av}^+)$, respectively. Based on the remainder theorem, there exist $\kappa \in [P_{av}, P_{av}^+]$ and $\nu \in [P_{av}, P_{av}^+]$ such that

$$R_i(P_{av}^+) = \frac{C''_i(\kappa)(P_{av}^+ - P_{av})^2}{2} \quad (29)$$

$$R_j(P_{av}^+) = \frac{C''_j(\nu)(P_{av}^+ - P_{av})^2}{2} \quad (30)$$

where $C''_i(\cdot)$ and $C''_j(\cdot)$ are the second-order derivatives of $C_i(\cdot)$ and $C_j(\cdot)$, respectively [43]. The second-order derivatives, which can be calculated from (1) as $C''_i(P) = -B_i \log_2 e / (N_i B_i + P)^2$ and $C''_j(P) = -B_j \log_2 e / (N_j B_j + P)^2$, are finite negative numbers for all possible power levels. Since $C_j(P_{av}^+) > C_i(P_{av}^+)$ and $C_i(P_{av}^-) = C_j(P_{av}^-)$ as discussed previously, the following inequality can be obtained based on (27)–(30):

$$C'_j(P_{av}) - C'_i(P_{av}) + \frac{(C''_j(\nu) - C''_i(\kappa))\epsilon}{2} > 0 \quad (31)$$

where $\epsilon = P_{av}^+ - P_{av}$ as defined above. As the second-order derivatives are finite and the relation in (31) should hold for any infinitesimally small ϵ value, it is concluded that $C'_i(P_{av}) < C'_j(P_{av})$. In other words, there is an increase in the first-order derivative of $C_{\max}(P)$ around $P = P_{av}$, which implies that $C'_{\max}(P_{av}^-) < C'_{\max}(P_{av}^+)$.

In the second part, it is proved that when there is an increase in the first-order derivative of $C_{\max}(P)$ around $P = P_{av}$, the optimal channel switching strategy outperforms the optimal single channel strategy. To that aim, consider a channel switching strategy (not necessarily an optimal one) that performs channel switching between channel i and channel j by employing power levels of P_{av}^- and P_{av}^+ , respectively, with equal channel switching factors; i.e., 0.5 each, where i, j, P_{av}^- and P_{av}^+ are as defined in the previous paragraph. Then, that channel switching strategy achieves an average capacity of $0.5 C_i(P_{av}^-) + 0.5 C_j(P_{av}^+)$, which can be expressed via Taylor series expansion as follows:

$$0.5 (C_i(P_{av}) + C'_i(P_{av})(P_{av}^- - P_{av}) + R_i(P_{av}^-)) + 0.5 (C_j(P_{av}) + C'_j(P_{av})(P_{av}^+ - P_{av}) + R_j(P_{av}^+)) \quad (32)$$

where $R_j(P_{av}^+)$ is as in (30) and $R_i(P_{av}^-) = C''_i(\omega)(P_{av}^- - P_{av})^2/2$ for a $\omega \in [P_{av}^-, P_{av}]$. Since $C_i(P_{av}^-) =$

$C_j(P_{\text{av}}) = C_{\max}(P_{\text{av}})$ as mentioned in the previous paragraph, (32) becomes equal to

$$C_{\max}(P_{\text{av}}) + 0.5 \epsilon (C'_j(P_{\text{av}}) - C'_i(P_{\text{av}})) + 0.25 \epsilon^2 (C''_i(\omega) + C''_j(v)). \quad (33)$$

Based on the result obtained in the first part of the proof, namely, $C'_i(P_{\text{av}}) < C'_j(P_{\text{av}})$, (33) implies that there exists an infinitesimally small $\epsilon > 0$ such that the channel switching strategy achieves a larger average capacity than $C_{\max}(P_{\text{av}})$, which is the capacity achieved by the optimal single channel strategy. Hence, based on the first and the second parts of the proof, it is concluded that the optimal channel switching strategy always provides a larger average capacity than the optimal single channel strategy in the case of a discontinuous first-order derivative of $C_{\max}(P)$ at $P = P_{\text{av}}$. ■

As stated in the proof of Proposition 3, the discontinuities in the first-order derivative of $C_{\max}(P)$ are observed when capacity curves intersect. The capacity curves of two channels, say channel k and channel l , can intersect [28] if one of them has a smaller bandwidth and a lower noise level than the other one; i.e., $B_k < B_l$ and $N_k < N_l$. In such a case, channel k has a higher capacity than channel l for small power levels (i.e., in the power-limited regime) since the capacity expression in (1) becomes approximately equal to $(\log_2 e)P/N_k$ and $(\log_2 e)P/N_l$ for channel k and channel l , respectively, when P is close to zero. On the other hand, for high power levels (i.e., in the bandwidth-limited regime), channel l achieves a higher capacity than channel k due to the following reason:

$$\lim_{P \rightarrow \infty} \frac{B_l \log_2 \left(1 + \frac{P}{N_l B_l}\right)}{B_k \log_2 \left(1 + \frac{P}{N_k B_k}\right)} = \frac{B_l}{B_k} > 1. \quad (34)$$

Therefore, the capacity curves can intersect in such scenarios. For example, in cognitive radio systems, there can exist multiple available frequency bands in the spectrum with various bandwidths and noise levels. Hence, such scenarios can be encountered in these systems.

Remark 2: The main reason for the improvements that can be realized via optimal channel switching is related to the fact that the optimal single channel approach can achieve the capacity values specified by $C_{\max}(P)$ in (4) only whereas the upper boundary of the convex hull of $C_{\max}(P)$ can also be achieved via optimal channel switching (cf. (3)). Therefore, the improvements that can be obtained via optimal channel switching over the optimal single channel approach are related to the convexity/concavity properties of $C_{\max}(P)$. Even though each capacity function is concave, their maximum is not necessarily concave. Therefore, opportunities can appear for average power values corresponding to convex regions of $C_{\max}(P)$ as illustrated in Section IV. The proof of Proposition 3 contains the theoretical explanation about this situation by showing that the first-order derivative of $C_{\max}(P)$ increases at the intersection point of two capacity curves, which implies that if two capacity functions intersect at a single point, there always exists a convex region around that intersection due to the mathematical expression for the capacity. Hence, improvements may be realized via channel switching around those intersection points.

B. Solution of Optimal Channel Switching Problem

When the optimal channel switching strategy is guaranteed to achieve a higher average capacity than the optimal single channel strategy (which can be deduced from Proposition 2 or Proposition 3), the optimization problem in (2) or (3) needs to be solved to calculate the maximum average capacity of the system, which involves a search over a $2K$ dimensional space. However, the following proposition states that the optimal strategy can be obtained by switching between no more than two different channels, and the resulting optimal strategy can be found via a search over a two-dimensional space only.

Proposition 4: The optimal solution of (2) results in channel switching between at most two different channels, and the achieved maximum average capacity is calculated as $\lambda^* C_{\max}(P_1^*) + (1 - \lambda^*) C_{\max}(P_2^*)$, where P_1^* and P_2^* are the solutions of the following problem:

$$\max_{\substack{P_1 \in (P_{\text{av}}, P_{\text{pk}}) \\ P_2 \in [0, P_{\text{av}}]}} \frac{P_{\text{av}} - P_2}{P_1 - P_2} C_{\max}(P_1) + \frac{P_1 - P_{\text{av}}}{P_1 - P_2} C_{\max}(P_2) \quad (35)$$

and λ^* is given by

$$\lambda^* = \frac{P_{\text{av}} - P_2^*}{P_1^* - P_2^*}. \quad (36)$$

Proof: As discussed in Proposition 1 and its proof, the optimization problems in (2) and (3) achieve the same maximum average capacity and the optimal channel switching strategy corresponding to (2) can be obtained from the solution of (3). Therefore, the optimization problem in (3) is considered, where the convex combinations of $C_{\max}(P_i)$'s and P_i 's are the two main functions. The set of all possible pairs of $C_{\max}(P)$ and P is defined as set \mathcal{U} ; that is, $\mathcal{U} = \{(C_{\max}(P), P), \forall P \in [0, P_{\text{pk}}]\}$. The convex hull of \mathcal{U} , denoted by \mathcal{V} , is guaranteed to contain the optimal solution of (3) since \mathcal{V} consists of all the convex combinations of the elements of \mathcal{U} by definition. In addition, it can be shown, similarly to [2], that the optimal solution of (3) should be on the boundary of \mathcal{V} since no interior points can be the maximizer of (3). Then, Carathéodory's theorem [44], [45] is invoked, which states that any point on the boundary of the convex hull \mathcal{V} of set \mathcal{U} can be represented by a convex combination of at most D points in set \mathcal{U} , where D is the dimension of space in which \mathcal{U} and \mathcal{V} reside. Hence, in this scenario (where $\mathcal{U} \subset \mathcal{V} \subset \mathbb{R}^2$), Carathéodory's theorem implies that an optimal solution of (3) can be expressed as the convex combination of (i.e., time sharing between) at most two different power levels; that is, $\nu_i \neq 0$ for one or two indices in (3). Therefore, the optimal solution of the channel switching problem in (2) corresponds to channel switching between at most two different channels.

Based on the previous result, the problem in (3) can be expressed as follows:

$$\max_{\lambda, P_1, P_2} \lambda C_{\max}(P_1) + (1 - \lambda) C_{\max}(P_2) \quad (37)$$

$$\text{subject to } \lambda P_1 + (1 - \lambda) P_2 = P_{\text{av}} \quad (38)$$

$$P_1 \in [0, P_{\text{pk}}], P_2 \in [0, P_{\text{pk}}] \quad (39)$$

$$\lambda \in [0, 1] \quad (40)$$

where the average power constraint is imposed with equality based on Lemma 1. Then, by substituting the constraints in (38)–(40) into the objective function and specifying the search space, the optimization problem in (35) can be obtained. ■

Once λ^* , P_1^* , and P_2^* are calculated as in Proposition 4, the optimal strategy can be specified as follows:

- **Case-1 (Channel Switching):** If $\lambda^* \in (0, 1)$, the optimal strategy is to switch between channel i and channel j with channel switching (time sharing) factors λ^* and $1 - \lambda^*$ and power levels P_1^* and P_2^* , respectively, where i and j are given by⁷

$$i = \arg \max_{l \in \{1, \dots, K\}} C_l(P_1^*), \quad (41)$$

$$j = \arg \max_{l \in \{1, \dots, K\}} C_l(P_2^*). \quad (42)$$

- **Case-2 (Single Channel):** If $\lambda^* = 0$, the optimal strategy is to perform communications over channel m all the time with a power level of P_{av} , where m is defined as

$$m = \arg \max_{l \in \{1, \dots, K\}} C_l(P_{av}). \quad (43)$$

Note that, in the case of $\lambda^* \in (0, 1)$, $i = j$ is not possible since time sharing of different power levels over the same channel always reduces the capacity due to the convexity of the capacity function in (1).

A flowchart is provided in Fig. 2 to explain the results obtained in this section. In particular, the optimal strategy can be specified as shown in the flowchart based on the propositions. Depending on the system parameters, either the single channel strategy or the channel switching strategy can be the optimal approach. From Proposition 2 and Proposition 3, the optimal strategy can be classified as *single channel* (case 2) or *channel switching* (case 1) without solving the optimization problem in (35): If the first-order derivative of $C_{\max}(P)$ is continuous at P_{av} (i.e., the condition in Proposition 3 does not hold) and the condition in Proposition 2 is satisfied, then the optimal single channel strategy is optimal (i.e., there is no need for channel switching), as shown in Fig. 2. In that case, the optimal solution of (2) can directly be expressed as $\lambda_{i^*} = 1$, $P_{i^*} = P_{av}$, and $\lambda_j = 0$ for all $j \in \{1, \dots, K\} \setminus \{i^*\}$, where $i^* = \arg \max_{i \in \{1, \dots, K\}} C_i(P_{av})$ (cf. (43)), and the maximum capacity becomes $C_{\max}(P_{av})$. If the condition in Proposition 3 holds or the condition in Proposition 2 is not satisfied, the optimal strategy is to switch between two different channels, and the optimization problem in Proposition 4 (i.e., (35)) can be solved in that case, as illustrated in Fig. 2. (As discussed in the next section, the solution of (35) can also be obtained based on Proposition 5.)

It is noted that the computational complexity of the optimization problem in (35) depends on the number of channels, K , only through C_{\max} in (4), and the dimension of the search space is always two irrespective of the number of channels. Therefore, Proposition 4 can provide a significant simplification over the

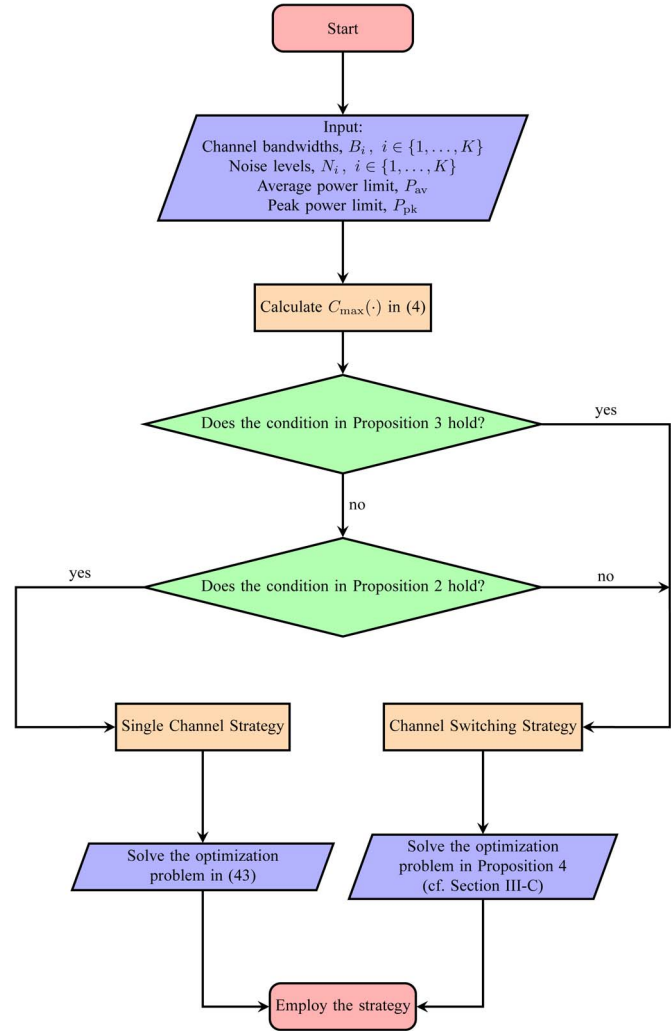


Fig. 2. A flowchart indicating the outline of the proposed optimal channel switching and optimal single channel approaches.

original formulation in (2), which requires a search over a $2K$ dimensional space.

C. Alternative Solution for Optimal Channel Switching

When the optimal strategy involves channel switching, which can be deduced from Proposition 2 and Proposition 3, one way to obtain the solution is to solve the optimization problem in (35). An alternative approach can be developed based on the following proposition:

Proposition 5: Consider a scenario in which channel switching between channel k and channel l is optimal. Let P_1^* and P_2^* denote the optimal transmit powers allocated to channel k and channel l , respectively. Then, the optimal solution satisfies at least one of the following conditions:

- $N_k + \frac{P_1^*}{B_k} = N_l + \frac{P_2^*}{B_l}$, where B_k and $N_k/2$ (B_l and $N_l/2$) are, respectively, the bandwidth and the constant power spectral density level of the additive Gaussian noise corresponding to channel k (channel l).
- $P_1^* = P_{pk}$ and $P_2^* = \frac{P_{av} - \lambda^* P_{pk}}{1 - \lambda^*}$, where $\lambda^* = (P_{av} - P_2^*) / (P_{pk} - P_2^*)$.

⁷In the case of multiple maximizers in (41) or (42), any of them can be chosen for the optimal strategy.

(iii) $P_2^* = P_{\text{pk}}$ and $P_1^* = \frac{P_{\text{av}} - (1 - \lambda^*)P_{\text{pk}}}{\lambda^*}$, where $\lambda^* = (P_{\text{pk}} - P_{\text{av}})/(P_{\text{pk}} - P_1^*)$.

Proof: The results in the proposition can be proved via Karush-Kuhn-Tucker (KKT) conditions [42] based on the optimal channel switching problem formulated in (2). To that aim, the Lagrangian [42] for the optimization problem in (2) is obtained first:

$$L(\lambda, \mathbf{P}, \mu, \gamma, \beta, \theta, \alpha) = - \sum_{i=1}^K \lambda_i C_i(P_i) + \mu \left(\sum_{i=1}^K \lambda_i P_i - P_{\text{av}} \right) - \sum_{i=1}^K \gamma_i P_i + \sum_{i=1}^K \beta_i (P_i - P_{\text{pk}}) + \theta \left(\sum_{i=1}^K \lambda_i - 1 \right) - \sum_{i=1}^K \alpha_i \lambda_i \quad (44)$$

where $\lambda = [\lambda_1 \cdots \lambda_K]$ and $\mathbf{P} = [P_1 \cdots P_K]$ are the optimization variables in (2), and $\mu, \gamma, \beta, \theta$, and α are the KKT multipliers, with $\gamma = [\gamma_1 \cdots \gamma_K]$, $\beta = [\beta_1 \cdots \beta_K]$, and $\alpha = [\alpha_1 \cdots \alpha_K]$. Then, the optimal solution of the problem in (2), denoted by $\{\lambda_i^*, P_i^*\}_{i=1}^K$ (equivalently, by λ^*, \mathbf{P}^*), satisfies the following KKT conditions:

- **Stationarity:** $\frac{\partial L(\lambda^*, \mathbf{P}^*, \mu, \gamma, \beta, \theta, \alpha)}{\partial \lambda_i} = 0$ and $\frac{\partial L(\lambda^*, \mathbf{P}^*, \mu, \gamma, \beta, \theta, \alpha)}{\partial P_i} = 0$ for $i \in \{1, \dots, K\}$, where L is as defined in (44).
- **Complementary slackness:** $\mu \left(\sum_{i=1}^K \lambda_i^* P_i^* - P_{\text{av}} \right) = 0$, $\sum_{i=1}^K \gamma_i P_i^* = 0$, $\sum_{i=1}^K \beta_i (P_i^* - P_{\text{pk}}) = 0$, and $\sum_{i=1}^K \alpha_i \lambda_i^* = 0$.
- **Primal and dual feasibility:** $\mu \geq 0$, $\gamma_i \geq 0$, $\beta_i \geq 0$, and $\alpha_i \geq 0$ for $i \in \{1, \dots, K\}$.

From the stationarity conditions; the following equalities are obtained based on (44):

$$C_i(P_i^*) = \mu P_i^* + \theta - \alpha_i, \quad \forall i \in \{1, \dots, K\}, \quad (45)$$

$$C'_i(P_i^*) = \mu + \frac{\beta_i - \gamma_i}{\lambda_i^*}, \quad \forall i \in \{1, \dots, K\}. \quad (46)$$

Now consider the scenario in the proposition, where channel switching between channel k and channel l is optimal; that is, $\lambda_k^* \neq 0$, $\lambda_l^* \neq 0$, $P_k^* = P_l^* \neq 0$, $P_1^* = P_2^* \neq 0$, and $P_i^* = \lambda_i^* = 0$ for $i \in \{1, \dots, K\} \setminus \{k, l\}$.⁸ Then, $\gamma_k = \gamma_l = 0$ and $\alpha_k = \alpha_l = 0$ can be obtained from the second and fourth complementary slackness conditions. For the optimal power levels, three possible scenarios exist:

- First, it is assumed that $P_1^* < P_{\text{pk}}$ and $P_2^* < P_{\text{pk}}$ hold. Then, $\beta_k = 0$ and $\beta_l = 0$ are satisfied due to the third complementary slackness condition. Combining this result with $\gamma_k = \gamma_l = 0$, $\lambda_k^* \neq 0$, and $\lambda_l^* \neq 0$, the condition in (46) can be expressed as $C'_k(P_k^*) = C'_l(P_l^*) = \mu$, which leads to condition (i) in the proposition based on the first-order derivative expression in (16).

⁸Note that the on-off scheme, in which one power level is equal to zero, cannot be optimal due to the concavity of the capacity curves and the fact that $C_i(0) = 0, \forall i \in \{1, \dots, K\}$.

- Second, it is assumed that $P_1^* = P_{\text{pk}}$ and $P_2^* < P_{\text{pk}}$. Due to Lemma 1, the average power constraint must be satisfied with equality, which leads to $P_2^* = (P_{\text{av}} - \lambda^* P_{\text{pk}})/(1 - \lambda^*)$, where $\lambda^* = (P_{\text{av}} - P_2^*)/(P_{\text{pk}} - P_2^*)$. Hence, condition (ii) in the proposition is obtained. Note that in this case $\beta_k \geq 0$ and $\beta_l = 0$, which implies that $C'_k(P_k^*) \geq C'_l(P_l^*)$ based on (46).
- For the third scenario, the third condition in the proposition can similarly be obtained under the assumption that $P_1^* < P_{\text{pk}}$ and $P_2^* = P_{\text{pk}}$.

Finally, it is noted that $P_1^* = P_2^* = P_{\text{pk}}$ is not possible since it violates the average power constraint as $P_{\text{pk}} > P_{\text{av}}$. Therefore, the optimal solution of the channel switching strategy between two channels satisfies at least one of the three conditions in Proposition 5. ■

Proposition 5 presents necessary conditions that need to be satisfied by the optimal channel switching strategy. Based on this proposition, the optimal solution of the problem in (2) can also be calculated as described in the following. For the scenario in which one of the power levels is set to P_{pk} , the maximum capacity achieved can be calculated from the second and third conditions in Proposition 5 as follows:

$$\tilde{C}_{\text{av}}(i, j) \triangleq \max_{P_j \in [0, P_{\text{av}}]} \frac{P_{\text{av}} - P_j}{P_{\text{pk}} - P_j} C_i(P_{\text{pk}}) + \frac{P_{\text{pk}} - P_{\text{av}}}{P_{\text{pk}} - P_j} C_j(P_j) \quad (47)$$

where $i, j \in \{1, \dots, K\}$ and $i \neq j$. Since one power level is fixed to P_{pk} , it is sufficient to consider the best channel only for that power level in calculating the maximum average capacity. Hence, a new function, which is a function of a single channel index only, is defined in that respect as follows:

$$\tilde{C}_{\text{av}}(j) \triangleq \max_{P_j \in [0, P_{\text{av}}]} \frac{P_{\text{av}} - P_j}{P_{\text{pk}} - P_j} C_{\text{max}}(P_{\text{pk}}) + \frac{P_{\text{pk}} - P_{\text{av}}}{P_{\text{pk}} - P_j} C_j(P_j) \quad (48)$$

where $j \in \{1, \dots, K\} \setminus \{k^*\}$ with $k^* = \arg \max_{i \in \{1, \dots, K\}} C_i(P_{\text{pk}})$ and $C_{\text{max}}(P_{\text{pk}}) = C_{k^*}(P_{\text{pk}})$. Then, in the case of channel switching between two channels where one power level is equal to P_{pk} , the maximum achieved capacity can be calculated as follows:

$$\tilde{C}_{\text{av}} = \max_{\substack{j \in \{1, \dots, K\} \\ j \neq k^*}} \tilde{C}_{\text{av}}(j). \quad (49)$$

It should be noted that \tilde{C}_{av} also includes the maximum capacity that can be achieved by the optimal single channel strategy since $\tilde{C}_{\text{av}}(j)$ in (48) reduces to $C_j(P_{\text{av}})$ for $P_j = P_{\text{av}}$ (which is added to the search space for this purpose). For the scenario in which the optimal power levels are below P_{pk} , the first condition in Proposition 5, namely, $N_i + P_i/B_i = N_j + P_j/B_j$, can be employed to obtain the following formulation for the maximum achieved capacity:

$$\tilde{C}_{\text{av}}(i, j) \triangleq \max_{P_j \in (P_{ij}^{\text{lb}}, P_{ij}^{\text{ub}}]} \frac{P_{\text{av}} - P_j}{P_i - P_j} C_i(P_i) + \frac{P_i - P_{\text{av}}}{P_i - P_j} C_j(P_j) \quad (50)$$

where $P_{ij}^{\text{lb}} \triangleq \max\{0, (P_{\text{av}} \frac{B_j}{B_i} + B_j(N_i - N_j))\}$, $P_{ij}^{\text{ub}} \triangleq \min\{P_{\text{av}}, (P_{\text{pk}} \frac{B_j}{B_i} + B_j(N_i - N_j))\}$, and $P_i = B_i(N_j - N_i) + B_i P_j / B_j$. Note that the search space for P_j (namely, P_{ij}^{lb} and P_{ij}^{ub}) is obtained by the joint consideration of $P_j \in (0, P_{\text{av}}]$ and $P_i = B_i(N_j - N_i) + B_i P_j / B_j \in (P_{\text{av}}, P_{\text{pk}}]$. Then, the maximum capacity that can be achieved by switching between two channels with power levels lower than P_{pk} can be calculated as follows:

$$\bar{C}_{\text{av}} = \max_{\substack{i, j \in \{1, \dots, K\} \\ i \neq j}} \bar{C}_{\text{av}}(i, j). \quad (51)$$

Overall, the solution of the optimal channel switching problem in (2) achieves the following maximum average capacity:

$$C_{\text{av}}^{\text{max}} = \max\{\tilde{C}_{\text{av}}, \bar{C}_{\text{av}}\} \quad (52)$$

where \tilde{C}_{av} and \bar{C}_{av} are as in (49) and (51), respectively. Also, the optimal strategy can be obtained as follows: If $\tilde{C}_{\text{av}} = C_{\text{max}}(P_{\text{av}}) \geq \bar{C}_{\text{av}}$, then the optimal solution corresponds to the single channel strategy, which is to transmit over channel m all the time with power level P_{av} , where $m = \arg \max_{i \in \{1, \dots, K\}} C_i(P_{\text{av}})$. (In fact, based on Proposition 2, the cases in which the single channel strategy is optimal can be determined beforehand, and the efforts in solving (48)–(52) can be avoided.) If $\tilde{C}_{\text{av}} \geq \bar{C}_{\text{av}}$ and $\tilde{C}_{\text{av}} > C_{\text{max}}(P_{\text{av}})$, the optimal strategy is to switch over channel k^* and channel j^* with power levels P_{pk} and $P_{j^*}^*$ and channel switching factors $(P_{\text{av}} - P_{j^*}^*) / (P_{\text{pk}} - P_{j^*}^*)$ and $(P_{\text{pk}} - P_{\text{av}}) / (P_{\text{pk}} - P_{j^*}^*)$, respectively, where $P_{j^*}^*$ denotes the maximizer of the problem in (48), $k^* = \arg \max_{i \in \{1, \dots, K\}} C_i(P_{\text{pk}})$, and $j^* = \arg \max_{j \in \{1, \dots, K\}, j \neq k^*} \tilde{C}_{\text{av}}(j)$, with $\tilde{C}_{\text{av}}(j)$ being as defined in (48). Finally, if $\bar{C}_{\text{av}} > \tilde{C}_{\text{av}}$, then the optimal strategy is to switch between channel j^* and channel i^* with power levels $P_{j^*}^*$ and $P_{i^*}^* = B_{i^*}(N_{j^*} - N_{i^*}) + B_{i^*} P_{j^*}^* / B_{j^*}$ and channel switching factors $(P_{i^*}^* - P_{\text{av}}) / (P_{i^*}^* - P_{j^*}^*)$ and $(P_{\text{av}} - P_{j^*}^*) / (P_{i^*}^* - P_{j^*}^*)$, respectively, where $P_{j^*}^*$ is the maximizer of the problem in (50) and i^* and j^* denote the maximizers of (51).

To compare the approach in the previous paragraph (called the second approach) to the one provided in Proposition 4 (called the first approach) in terms of the computational complexity in obtaining the optimal switching solution, the optimization problems in (35) and in (48)–(52) are considered. In the first approach, the problem in (35) requires a two-dimensional search over $[0, P_{\text{av}}] \times (P_{\text{av}}, P_{\text{pk}}]$. On the other hand, the main operations in the second approach are related to the optimization problem in (48), which requires a one-dimensional search over $[0, P_{\text{av}}]$, and the optimization problem in (50), which requires a one-dimensional search over a subset of $[0, P_{\text{av}}]$. It is observed from (49) and (51) that the problem in (48) is solved for $K - 1$ different channel indices and the one in (50) is solved for $K(K - 1)$ different channel pairs. Therefore, overall, the second approach involves $K^2 - 1$ one-dimensional searches. In fact, instead of K , a smaller number can be considered in many scenarios when some channels outperform other channels in the sense that they have larger or equal capacities for all possible power values. From (1), it

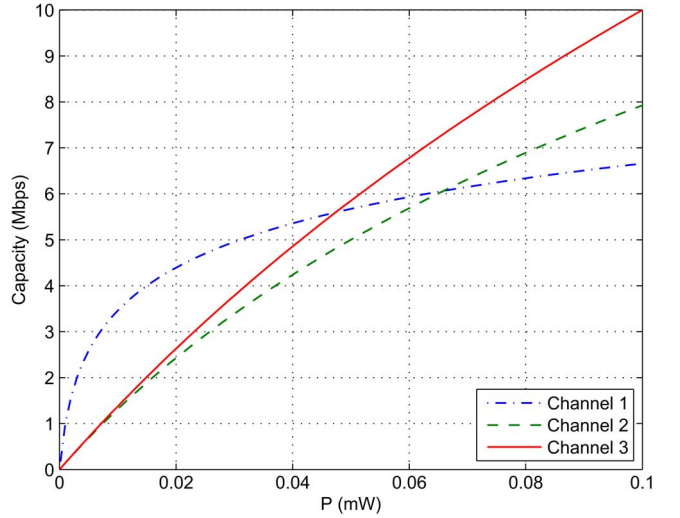


Fig. 3. Capacity of each channel versus power, where $B_1 = 1$ MHz, $B_2 = 5$ MHz, $B_3 = 10$ MHz, $N_1 = 10^{-12}$ W/Hz, $N_2 = 10^{-11}$ W/Hz, and $N_3 = 10^{-11}$ W/Hz.

is observed that, for channel i and channel j , if $N_i \leq N_j$ and $B_i \geq B_j$, then channel i outperforms channel j for all power values. Therefore, channel j can be excluded from the set of channels for the optimal channel switching solution. Hence, based on this observation, it can be stated that the second approach involves $\tilde{K}^2 - 1$ one-dimensional searches, where \tilde{K} is the number of elements in set C , which is defined as $C = \{i \in \{1, \dots, K\} \mid (N_i < N_j \text{ or } B_i > B_j) \forall j \in \{1, \dots, K\} \setminus \{i\}\}$.⁹ Therefore, the computational complexity comparison between the first approach and the second approach depends on the number of channels and their noise levels and bandwidths. In particular, the second (first) approach become more desirable for small (large) values of \tilde{K} .

IV. NUMERICAL RESULTS

In this section, numerical examples are provided to investigate the proposed optimal channel switching strategy and to compare it against the optimal single channel strategy. First, consider a scenario with $K = 3$ channels and the following bandwidths and noise levels (cf. (1)): $B_1 = 1$ MHz, $B_2 = 5$ MHz, $B_3 = 10$ MHz, $N_1 = 10^{-12}$ W/Hz, $N_2 = 10^{-11}$ W/Hz, and $N_3 = 10^{-11}$ W/Hz. Suppose that the peak power limit in (2) is set to $P_{\text{pk}} = 0.1$ mW. In Fig. 3, the capacity of each channel is plotted as a function of power based on the capacity formula in (1). For the scenario in Fig. 3, the proposed optimal channel switching strategy and the optimal single channel strategy are calculated for various average power limits (P_{av}), and the achieved maximum average capacities are plotted in Fig. 4 versus P_{av} . Also, the shaded area in the figure indicates the achievable rates (average capacities) via channel switching that are higher than those achieved by the optimal single channel strategy. As discussed in the previous section, the optimal single channel strategy achieves a capacity of $C_{\text{max}}(P_{\text{av}})$, which is $C_{\text{max}}(P_{\text{av}}) =$

⁹For convenience, it is assumed that the identical channels (the same bandwidth and noise level) are already eliminated.

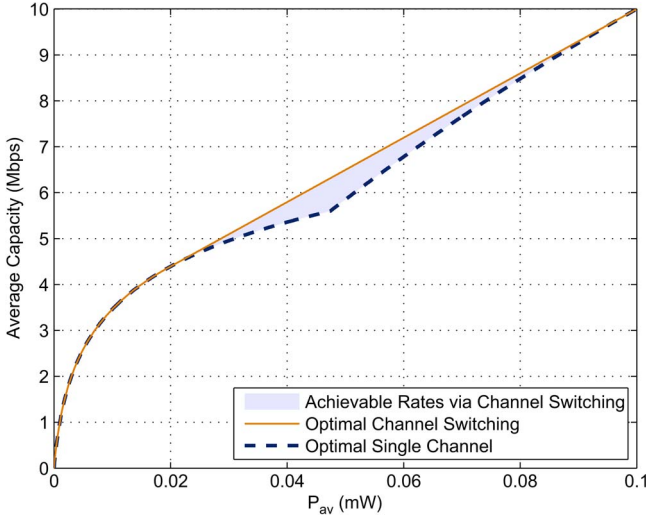


Fig. 4. Average capacity versus average power limit for the optimal channel switching and the optimal single channel strategies for the scenario in Fig. 3, where $P_{pk} = 0.1$ mW. The shaded area indicates the achievable rates via channel switching that are higher than those achieved by the optimal single channel strategy.

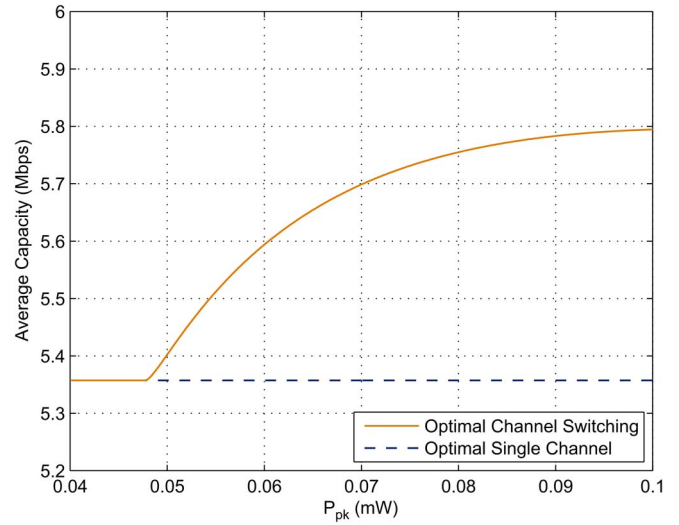


Fig. 5. Average capacity versus peak power limit for the optimal channel switching and the optimal single channel strategies for the scenario in Fig. 3, where $P_{av} = 0.04$ mW.

TABLE I

OPTIMAL STRATEGY FOR THE SCENARIO IN FIG. 3, WHICH EMPLOYS CHANNEL i AND CHANNEL j WITH CHANNEL SWITCHING FACTORS λ^* AND $(1 - \lambda^*)$ AND POWER LEVELS P_1^* AND P_2^* , RESPECTIVELY

P_{av} (mW)	λ^*	P_1^*	i	$(1 - \lambda^*)$	P_2^*	j
0.01	—	—	—	1	0.01	1
0.02	0.005	0.1000	3	0.995	0.0196	1
0.03	0.129	0.1000	3	0.871	0.0196	1
0.04	0.254	0.1000	3	0.746	0.0196	1
0.05	0.378	0.1000	3	0.622	0.0196	1
0.06	0.503	0.1000	3	0.498	0.0196	1
0.07	0.627	0.1000	3	0.373	0.0196	1
0.08	0.751	0.1000	3	0.249	0.0196	1
0.09	0.876	0.1000	3	0.124	0.0196	1

$\max\{C_1(P_{av}), C_2(P_{av}), C_3(P_{av})\}$ in the considered scenario. It is observed from Fig. 3 and Fig. 4 that $C_{\max}(P_{av}) = C_1(P_{av})$ for $P_{av} \in (0, 0.048)$ mW and $C_{\max}(P_{av}) = C_3(P_{av})$ for $P_{av} \in [0.048, 0.1]$ mW; that is, channel 1 is the best channel up to $P_{av} = 0.048$ mW, and channel 3 is the best after that power level. From Fig. 4, it is also noted that the proposed optimal channel switching strategy outperforms the optimal single channel strategy for $P_{av} \in [0.0196, 0.1]$ mW, and the two strategies have the same performance for $P_{av} < 0.0196$ mW. These regions can also be obtained by checking the necessary and sufficient condition in Proposition 2 (see (15)), which is satisfied for all $P \in [0, 0.1]$ mW for $P_{av} < 0.0196$ mW, and is not satisfied for some $P \in [0, 0.1]$ mW for $P_{av} \in [0.0196, 0.1]$ mW. In addition, in accordance with Proposition 3, it is observed that the optimal channel switching strategy outperforms the optimal single channel strategy at $P_{av} = 0.048$ mW, which corresponds to a discontinuity point for the first-order derivative of $C_{\max}(P)$.

To provide a detailed investigation of the optimal channel switching strategy, Table I presents the optimal channel switching solutions for various values of the average power limit, P_{av} . In the table, the optimal solution is represented by parameters λ^* , P_1^* , P_2^* , i , and j , meaning that channel

i is used with channel switching factor λ^* and power P_1^* , and channel j is used with channel switching factor $1 - \lambda^*$ and power P_2^* . It is observed from the table that the optimal solution reduces to the optimal single channel strategy for $P_{av} = 0.01$ mW (in which case channel 1 is used all the time), and it involves switching between channel 1 and channel 3 for larger values of P_{av} . This observation is also consistent with Fig. 4, which illustrates improvements via channel switching for $P_{av} > 0.0196$ mW. It is also observed from the table that the optimal channel switching solution for $P_{av} > 0.0196$ mW satisfies condition (ii) in Proposition 5 since $P_1^* = P_{pk} = 0.1$, $P_2^* = (P_{av} - \lambda^* P_{pk}) / (1 - \lambda^*) = 0.0196$ mW, and $\lambda^* = (P_{av} - P_2^*) / (P_{pk} - P_2^*)$. In addition, as stated in Lemma 1, the optimal solutions always operate at the average power limits.

For the scenario in Fig. 3, the average capacity versus the peak power limit curves are presented for the optimal channel switching and the optimal single channel strategies in Fig. 5, where the average power limit is set to $P_{av} = 0.04$ mW. From the figure, it is observed that the average capacity for the optimal single channel strategy does not depend on the P_{pk} value since this strategy achieves an average capacity of $C_{\max}(P_{av})$ and $P_{pk} > P_{av} = 0.04$ mW in this scenario. On the other hand, increased P_{pk} can improve the average capacity for the optimal channel switching strategy as observed from the figure. The intuition behind this increase can be deduced from Fig. 3 and Table II. In particular, as observed from Table II, when the peak power limit is larger than 0.048 mW, which is the discontinuity point for the first-order derivative of C_{\max} , the optimal channel switching strategy performs time sharing (switching) between channel 1 and channel 3, where channel 3 is operated at the peak power limit, P_{pk} .

Next, a scenario with $K = 4$ channels is considered, where the bandwidths and the noise levels of channels are specified as $B_1 = 0.5$ MHz, $B_2 = 2.0$ MHz, $B_3 = 2.5$ MHz, $B_4 = 5.0$ MHz, $N_1 = 10^{-12}$ W/Hz, $N_2 = 1.5 \times 10^{-11}$ W/Hz, $N_3 = 2.0 \times 10^{-11}$ W/Hz, and $N_4 = 2.5 \times 10^{-11}$ W/Hz. Also, the peak power limit is set to $P_{pk} = 0.25$ mW. In Fig. 6, the

TABLE II
OPTIMAL STRATEGY FOR THE SCENARIO IN FIG. 3, WHICH EMPLOYS CHANNEL i AND CHANNEL j WITH CHANNEL SWITCHING FACTORS λ^* AND $(1 - \lambda^*)$ AND POWER LEVELS P_1^* AND P_2^* , RESPECTIVELY

P_{pk} (mW)	λ^*	P_1^*	i	$(1 - \lambda^*)$	P_2^*	j
0.045	—	—	—	1.000	0.0400	1
0.050	0.465	0.050	3	0.535	0.0313	1
0.055	0.488	0.055	3	0.512	0.0257	1
0.060	0.455	0.060	3	0.545	0.0233	1
0.065	0.419	0.065	3	0.581	0.0220	1
0.070	0.387	0.070	3	0.613	0.0211	1
0.075	0.357	0.075	3	0.643	0.0206	1
0.080	0.331	0.080	3	0.669	0.0202	1
0.085	0.309	0.085	3	0.691	0.0199	1
0.090	0.289	0.090	3	0.711	0.0197	1
0.095	0.271	0.095	3	0.729	0.0196	1
0.100	0.254	0.100	3	0.746	0.0196	1

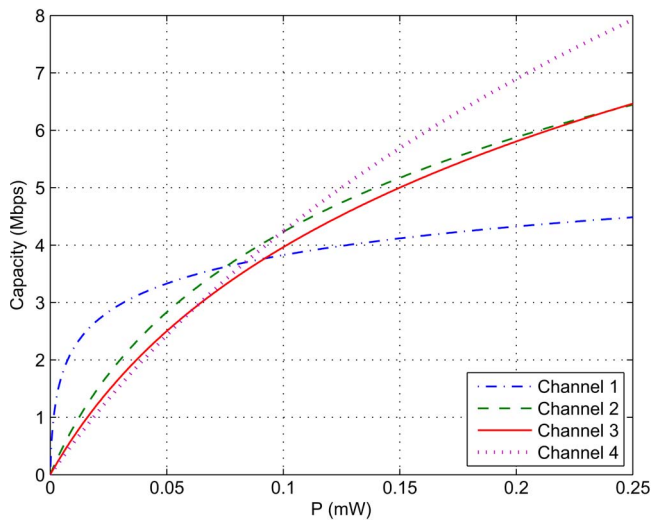


Fig. 6. Capacity of each channel versus power, $B_1 = 0.5$ MHz, $B_2 = 2.0$ MHz, $B_3 = 2.5$ MHz, $B_4 = 5.0$ MHz, $N_1 = 10^{-12}$ W/Hz, $N_2 = 1.5 \times 10^{-11}$ W/Hz, $N_3 = 2.0 \times 10^{-11}$ W/Hz, $N_4 = 2.5 \times 10^{-11}$ W/Hz, and $P_{pk} = 0.25$ mW.

capacity of each channel is plotted versus the transmit power. In Fig. 7, the average capacity versus P_{av} curves are presented for the proposed optimal channel switching strategy and the optimal single channel strategy. In addition, the shaded area in the figure indicates the achievable rates via channel switching that are higher than those achieved by the optimal single channel strategy. From Fig. 7, it is observed that the optimal channel switching strategy outperforms the optimal single channel strategy for $P_{av} \in (0.031, 0.187)$ mW. Also, it can be deduced from Fig. 6 and Fig. 7 that channel 3 is not employed in any strategy since $C_{\max}(P) \neq C_3(P)$ for $P \in [0, 0.25]$ mW. In Table III, the optimal strategies are presented for the scenario in Fig. 6 for various values of P_{av} . As observed from the table, the optimal strategy corresponds to the optimal single channel strategy for small and large values of P_{av} and it corresponds to channel switching between channel 1 and channel 4 for medium range of P_{av} values. These observations are in accordance with Fig. 7. In addition, it is important to emphasize that the channels employed in the optimal channel switching strategy for a given value of P_{av} may not correspond to the channel used in the optimal single channel strategy for the same P_{av}

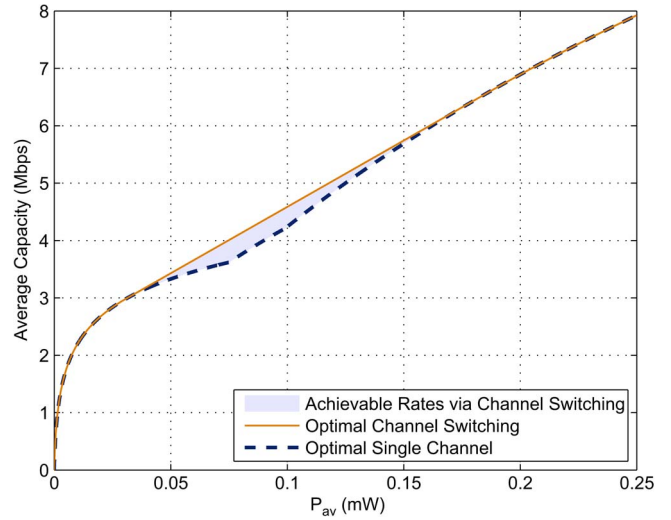


Fig. 7. Average capacity versus average power limit for the optimal channel switching and the optimal single channel approaches for $P_{pk} = 0.25$ mW. The shaded area indicates the achievable rates via channel switching that are higher than those achieved by the optimal single channel strategy.

TABLE III
OPTIMAL STRATEGY FOR THE SCENARIO IN FIG. 6, WHICH EMPLOYS CHANNEL i AND CHANNEL j WITH CHANNEL SWITCHING FACTORS λ^* AND $(1 - \lambda^*)$ AND POWER LEVELS P_1^* AND P_2^* , RESPECTIVELY

P_{av} (mW)	λ^*	P_1^*	i	$(1 - \lambda^*)$	P_2^*	j
0.025	—	—	—	1.0000	0.0250	1
0.050	0.1236	0.1868	4	0.8764	0.0307	1
0.075	0.2838	0.1868	4	0.7162	0.0307	1
0.100	0.4439	0.1868	4	0.5561	0.0307	1
0.125	0.6041	0.1868	4	0.3959	0.0307	1
0.150	0.7643	0.1868	4	0.2357	0.0307	1
0.175	0.9244	0.1868	4	0.0756	0.0307	1
0.200	—	—	—	1.0000	0.2000	4
0.225	—	—	—	1.0000	0.2250	4

value. For example, as can be observed from Fig. 6 and Fig. 7, channel 2 is not employed in the optimal channel switching strategy for $P_{av} \in (0.031, 0.187)$ mW (channel 1 and channel 4 are employed); however, it is the optimal channel for the optimal single channel strategy for $P_{av} \in [0.075, 0.099]$ mW as $C_{\max}(P_{av}) = C_2(P_{av})$. This is mainly due to the fact that the optimal single channel approach achieves the capacity value specified by $C_{\max}(P_{av})$ whereas the upper boundary of the convex hull of $C_{\max}(P)$ is achieved via the optimal channel switching approach.

Finally, for the scenario in Fig. 6, P_{av} is set to $P_{av} = 0.07$ mW, and the effects of the peak power limit, P_{pk} , are investigated. In Fig. 8, the average capacity is plotted versus P_{pk} for the optimal channel switching and optimal single channel strategies. It is observed that the optimal single channel strategy achieves a constant capacity of $C_{\max}(P_{av})$ for all P_{pk} values, where $P_{pk} \in (0.07, 0.25]$ mW. On the other hand, for the optimal channel switching strategy, improvements in the average capacity are observed for when P_{pk} is larger than 0.075 mW. It is also noted that the behavioral changes in the average capacity curve for the optimal channel switching strategy occurs at 0.075 mW and 0.099 mW, which correspond to the discontinuity points for the first-order derivative of C_{\max} , as can be observed from Fig. 6. Similar to Table II, Table IV

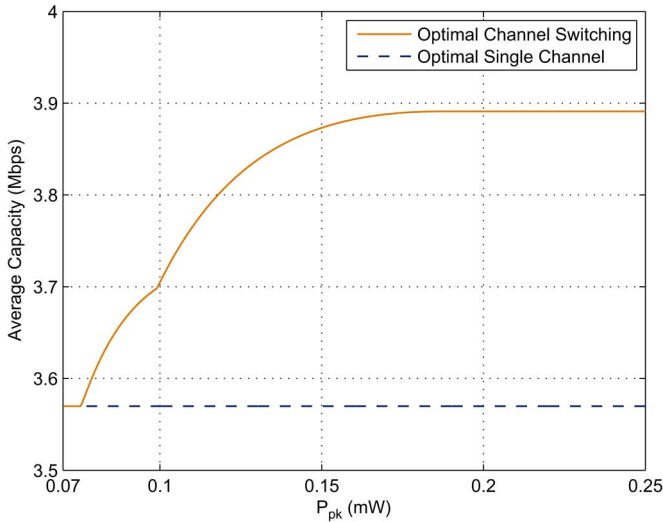


Fig. 8. Average capacity versus peak power limit for the optimal channel switching and the optimal single channel strategies for the scenario in Fig. 6, where $P_{av} = 0.07$ mW.

TABLE IV
OPTIMAL STRATEGY FOR THE SCENARIO IN FIG. 6, WHICH EMPLOYS CHANNEL i AND CHANNEL j WITH CHANNEL SWITCHING FACTORS λ^* AND $(1 - \lambda^*)$ AND POWER LEVELS P_1^* AND P_2^* , RESPECTIVELY

P_{pk} (mW)	λ^*	P_1^*	i	$(1 - \lambda^*)$	P_2^*	j
0.075	—	—	—	1.0000	0.0700	1
0.080	0.6540	0.0800	2	0.3460	0.0511	1
0.085	0.6173	0.0850	2	0.3827	0.0458	1
0.090	0.5745	0.0900	2	0.4255	0.0430	1
0.095	0.5336	0.0950	2	0.4664	0.0414	1
0.100	0.5008	0.1000	4	0.4992	0.0399	1
0.125	0.4009	0.1250	4	0.5991	0.0332	1
0.150	0.3260	0.1500	4	0.6740	0.0313	1
0.175	0.2723	0.1750	4	0.7277	0.0307	1
0.200	0.2518	0.1868	4	0.7482	0.0307	1
0.225	0.2518	0.1868	4	0.7482	0.0307	1
0.250	0.2518	0.1868	4	0.7482	0.0307	1

presents the solutions corresponding to the optimal strategy for various values of P_{pk} . From the table, it is observed that the optimal strategy changes with respect to the peak power limit. In addition, it can be shown that the solutions of the optimal channel switching strategy satisfy condition (i) in Proposition 5 for $P_{pk} \geq 0.1868$ mW and condition (ii) for $P_{pk} \in (0.075, 0.1868)$ mW.

Based on the numerical examples, an intuitive explanation can be provided about the benefits of channel switching and why the optimal channel switching strategy involves switching between no more than two channels. In the absence of channel switching, the maximum capacity is given by $C_{\max}(P_{av})$, whereas via channel switching, the upper boundary of the convex hull of $C_{\max}(P_{av})$ can also be achieved (see, e.g., Fig. 4). Since the upper boundary of the convex hull can always be formed by a convex combination of two different points, no more than two different channels are needed to achieve the optimal capacity. Finally, it is important to note that the optimal solution to the channel switching problem in (3) may not be unique in general; that is, in some cases, two different channel switching strategies or a channel switching strategy and a single channel strategy can be the optimal solutions.

V. CONCLUDING REMARKS

In this study, the optimal channel switching strategy has been proposed for average capacity maximization in the presence of average and peak power constraints. Necessary and sufficient conditions have been derived for specifying whether the proposed optimal channel switching strategy can or cannot outperform the optimal single channel strategy. In addition, the optimal channel switching solution has been shown to be realized by channel switching between at most two different channels, and a low-complexity optimization problem has been formulated to calculate the optimal channel switching solution. Furthermore, based on the necessary conditions that need to be satisfied by the optimal channel switching solution, an alternative approach has been proposed for calculating the optimal channel switching strategy. Numerical examples have been investigated and intuitive explanations about the benefits of channel switching have been provided. Although Gaussian channels have been considered in this study, the results can also be applied to block frequency-flat fading channels in the presence of Gaussian noise when the channel state information is available at the transmitter and the receiver. In that scenario, the proposed channel switching strategy can be adopted for each channel state. As future work, performance improvements that can be achieved by performing both channel switching at each channel state and adaptation over varying channel states can be considered. Another future work involves the consideration of channel switching costs (delays) in the design of optimal channel switching strategies.

ACKNOWLEDGMENT

The authors would like to thank Prof. Erdal Arıkan from Bilkent University for his insightful comments.

REFERENCES

- [1] A. D. Sezer, S. Gezici, and H. Inaltekin, "Optimal channel switching for average capacity maximization," in *Proc. IEEE ICASSP*, Florence, Italy, May 2014, pp. 3503–3507.
- [2] 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, Jul. 2007.
- [3] H. Chen, P. K. Varshney, and J. H. Michels, "Noise enhanced parameter estimation," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5074–5081, Oct. 2008.
- [4] A. Patel and B. Kosko, "Optimal noise benefits in Neyman-Pearson and inequality-constrained signal detection," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1655–1669, May 2009.
- [5] S. Bayram and S. Gezici, "Noise enhanced M-ary composite hypothesis-testing in the presence of partial prior information," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1292–1297, Mar. 2011.
- [6] S. Bayram, S. Gezici, and H. V. Poor, "Noise enhanced hypothesis-testing in the restricted Bayesian framework," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 3972–3989, Aug. 2010.
- [7] M. Azizoglu, "Convexity properties in binary detection problems," *IEEE Trans. Inf. Theory*, vol. 42, no. 4, pp. 1316–1321, Jul. 1996.
- [8] S. Loyka, V. Kostina, and F. Gagnon, "Error rates of the maximum-likelihood detector for arbitrary constellations: Convex/concave behavior and applications," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1948–1960, Apr. 2010.
- [9] C. Goken, S. Gezici, and O. Arıkan, "Optimal stochastic signaling for power-constrained binary communications systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp. 3650–3661, Dec. 2010.
- [10] C. Goken, S. Gezici, and O. Arıkan, "Optimal signaling and detector design for power-constrained binary communications systems over

- non-Gaussian channels,” *IEEE Commun. Lett.*, vol. 14, no. 2, pp. 100–102, Feb. 2010.
- [11] 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.
- [12] B. Dulek, S. Gezici, and O. Arikan, “Convexity properties of detection probability under additive Gaussian noise: Optimal signaling and jamming strategies,” *IEEE Trans. Signal Process.*, vol. 61, no. 13, pp. 3303–3310, Jul. 2013.
- [13] M. E. Tutay, S. Gezici, and O. Arikan, “Optimal detector randomization for multiuser communications systems,” *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 2876–2889, Jul. 2013.
- [14] 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.
- [15] S. Bayram, N. D. Vanli, B. Dulek, I. Sezer, and S. Gezici, “Optimum power allocation for average power constrained jammers in the presence of non-Gaussian noise,” *IEEE Commun. Lett.*, vol. 16, no. 8, pp. 1153–1156, Aug. 2012.
- [16] H. Chen, P. K. Varshney, S. M. Kay, and J. H. Michels, “Theory of the stochastic resonance effect in signal detection: Part II—Variable detectors,” *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5031–5041, Oct. 2007.
- [17] B. Dulek and S. Gezici, “Optimal stochastic signal design and detector randomization in the Neyman-Pearson framework,” in *Proc. 37th IEEE ICASSP*, Kyoto, Japan, Mar. 25–30, 2012, pp. 3025–3028.
- [18] E. L. Lehmann, *Testing Statistical Hypotheses*, 2nd ed. New York, NY, USA: Chapman & Hall, 1986.
- [19] Y. Ma and C. C. Chai, “Unified error probability analysis for generalized selection combining in Nakagami fading channels,” *IEEE J. Sel. Areas Commun.*, vol. 18, no. 11, pp. 2198–2210, Nov. 2000.
- [20] J. A. Ritcey and M. Azizoglu, “Performance analysis of generalized selection combining with switching constraints,” *IEEE Commun. Lett.*, vol. 4, no. 5, pp. 152–154, May 2000.
- [21] B. Dulek, M. E. Tutay, S. Gezici, and P. K. Varshney, “Optimal signaling and detector design for M-ary communication systems in the presence of multiple additive noise channels,” *Digit. Signal Process.*, vol. 26, pp. 153–168, Mar. 2014.
- [22] 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.*, to be published, DOI: 10.1109/TCOMM.2015.2418252.
- [23] F. Gaaloul, H.-C. Yang, R. Radaydeh, and M.-S. Alouini, “Switch based opportunistic spectrum access for general primary user traffic model,” *IEEE Wireless Commun. Lett.*, vol. 1, no. 5, pp. 424–427, Oct. 2012.
- [24] Y. Liu and M. Liu, “To stay or to switch: Multiuser dynamic channel access,” in *Proc. IEEE INFOCOM*, Apr. 2013, pp. 1249–1257.
- [25] L. Chen, S. Iellamo, and M. Coupechoux, “Opportunistic spectrum access with channel switching cost for cognitive radio networks,” in *Proc. IEEE ICC*, Kyoto, Japan, Jun. 2011, pp. 1–5.
- [26] D. Xu, E. Jung, and X. Liu, “Optimal bandwidth selection in multi-channel cognitive radio networks: How much is too much?” in *Proc. 3rd IEEE Symp. New Frontiers DySPAN*, Oct. 2008, pp. 1–11.
- [27] R. G. Gallager, *Information Theory and Reliable Communication*. New York, NY, USA: Wiley, 1968.
- [28] L. Li and A. Goldsmith, “Capacity and optimal resource allocation for fading broadcast channels—Part I: Ergodic capacity,” *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 1083–1102, Mar. 2001.
- [29] Z. Shen, J. Andrews, and B. L. Evans, “Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints,” *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2726–2737, Nov. 2005.
- [30] F. F. Digham, “Joint power and channel allocation for cognitive radios,” in *Proc. IEEE WCNC*, Mar. 2008, pp. 882–887.
- [31] A. F. Molisch, M. Z. Win, Y.-S. Choi, and J. H. Winters, “Capacity of MIMO systems with antenna selection,” *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1759–1772, Jul. 2005.
- [32] L. Dai, S. Sfar, and K. B. Letaief, “Optimal antenna selection based on capacity maximization for MIMO systems in correlated channels,” *IEEE Trans. Commun.*, vol. 54, no. 3, pp. 563–573, Mar. 2006.
- [33] T. J. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2008.
- [34] “Requirements related to technical performance for IMT-advanced radio interface(s),” Int. Telecommun. Union, Geneva, Switzerland, 2008.
- [35] S. R. Bhaskaran, S. V. Hanly, N. Badruddin, and J. S. Evans, “Maximizing the sum rate in symmetric network of interfering link,” *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4471–4487, Sep. 2010.
- [36] H. Inaltekin and S. V. Hanly, “Optimality of binary power control for the single cell uplink,” *IEEE Trans. Inf. Theory*, vol. 58, no. 10, pp. 6484–6498, Oct. 2012.
- [37] J. Mitola and G. Q. Maguire, “Cognitive radio: Making software radios more personal,” *IEEE Pers. Commun. Mag.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.
- [38] S. Gezici, H. Celebi, H. V. Poor, and H. Arslan, “Fundamental limits on time delay estimation in dispersed spectrum cognitive radio systems,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 78–83, Jan. 2009.
- [39] J. G. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. C. Reed, “Femtocells: Past, present, and future,” *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 497–508, Apr. 2012.
- [40] A. Ghosh *et al.*, “Heterogeneous cellular networks: From theory to practice,” *IEEE Trans. Commun. Mag.*, vol. 50, no. 6, pp. 54–64, Jun. 2012.
- [41] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley-Interscience, 1991.
- [42] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge Univ. Press, 2004.
- [43] G. B. Thomas, M. D. Weir, J. Hass, and F. R. Giordano, *Thomas’ Calculus*, 11th ed. Reading, MA, USA: Addison Wesley, 2004.
- [44] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ, USA: Princeton Univ. Press, 1968.
- [45] D. P. Bertsekas, A. Nedic, and A. E. Ozdaglar, *Convex Analysis and Optimization*. Boston, MA, USA: Athena Specific, 2003.



Ahmet Dundar Sezer was born in 1989 in Emet, Kutahya, Turkey. He received the B.S. and M.S. degrees in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2011 and 2013, respectively. He is currently working toward the Ph.D. degree at Bilkent University. His current research interests include signal processing, wireless communications, and optimization.



Sinan Gezici (S’03–M’06–SM’11) received the B.S. degree from Bilkent University, Ankara, Turkey in 2001, and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, USA, in 2006. From 2006 to 2007, he worked at Mitsubishi Electric Research Laboratories, Cambridge, MA, USA. Since 2007, he has been with the Department of Electrical and Electronics Engineering at Bilkent University, where he is currently an Associate Professor.

His research interests are in the areas of detection and estimation theory, wireless communications, and localization systems. Among his publications in these areas is the book *Ultra-wideband Positioning Systems: Theoretical Limits, Ranging Algorithms, and Protocols* (Cambridge University Press, 2008). Dr. Gezici is an Associate Editor for *IEEE TRANSACTIONS ON COMMUNICATIONS*, *IEEE WIRELESS COMMUNICATIONS LETTERS*, and *Journal of Communications and Networks*.



Hazer Inaltekin (S’04–M’06) received the B.S. degree (with high honors) in electrical and electronics engineering from Bogazici University, Istanbul, Turkey, in 2001, and the M.S. and Ph.D. degrees in electrical and computer engineering from Cornell University, Ithaca, NY, USA, in 2006. He is an Associate Professor of Electrical and Electronics Engineering at Antalya International University, Antalya, Turkey. He was a Postdoctoral Research Associate at Cornell University, Ithaca, NY, USA, from 2006 to 2007, and at Princeton University, Princeton, NJ,

from 2007 to 2009. In 2009, he joined the Department of Electrical and Electronic Engineering at the University of Melbourne, Melbourne, Vic., Australia, as a Research Fellow. He was a Senior Research Fellow at the same department between January 2011 and August 2011. Since August 2011, he has been on the faculty at Antalya International University. His research interests include wireless communications, wireless networks, social networks, game theory, and information theory.