Optimal Power Allocation for Average Detection Probability Criterion over Flat Fading Channels

Serkan Sarıtaş, Berkan Dulek, Ahmet Dundar Sezer, Sinan Gezici, and Serdar Yüksel

***Abstract*—In this paper, the problem of optimal power allocation over flat fading additive white Gaussian noise channels is considered for maximizing the average detection probability of a signal emitted from a power constrained transmitter in the Neyman-Pearson framework. It is assumed that the transmitter can perform power adaptation under peak and average power constraints based on the channel state information fed back by the receiver. Using results from measure theory and convex analysis, it is shown that this optimization problem, which is in general nonconvex, has an equivalent Lagrangian dual that admits no duality gap and can be solved using dual decomposition. Efficient numerical algorithms are proposed to determine the optimal power allocation scheme under peak and average power constraints. Furthermore, the continuity and monotonicity properties of the corresponding optimal power allocation scheme are characterized with respect to the signal-to-noise ratio for any given value of the false alarm probability. Simulation examples are presented to corroborate the theoretical results and illustrate the performance improvements due to the proposed optimal power allocation strategy.**

***Index Terms*– Detection probability, power allocation, fading, Neyman-Pearson, power constraint.**

I. Introduction

Due to time-varying nature of wireless channels, dynamic allocation of communication resources has a significant impact on the performance of communication systems. In particular, the use of dynamic power allocation instead of a fixed strategy can lead to significant performance improvements in the presence of fading. In the literature, dynamic power allocation has mostly been employed for enhancing the channel capacity of communication systems (e.g., [1]–[3]). For a fading additive white Gaussian noise (AWGN) channel with perfect channel state information (CSI) available at both the transmitter and the receiver, the optimal power allocation problem is studied in order to maximize the ergodic capacity subject to an average power constraint in [1]. It is shown that the optimal power allocation policy corresponds to the water-filling solution, in which no power is transmitted until the received signal-to-noise ratio (SNR) exceeds a certain threshold, and higher power levels are allocated as the channel condition improves. In [2], the optimal power allocation strategies are obtained to maximize the ergodic capacity and the outage capacity of secondary users in cognitive radio networks. In the presence of average/peak transmit and interference power constraints, it is demonstrated that the average power constraints are more flexible than the peak power constraints in terms of the capacity improvements for the secondary users. In a similar context, the optimal power allocation is investigated in [3] for energy-efficient capacity maximization over fading cognitive radio channels and similar results to those in [2] are obtained.

Although less numerous in the literature, dynamic power allocation is also considered for performance metrics other than the channel capacity in order to utilize the communication channel effectively. In [4], optimal power allocation strategies are derived in order to minimize the average bit error rate (BER) for secondary users in a cognitive radio network. In [5], the optimal power allocation over space and time is considered for BER minimization of multiple-input single-output (MISO) communications over Rayleigh fading channels subject to an average power constraint and a peak-to-average power ratio constraint. In [6], the optimal power allocation strategy is presented for the minimization of outage probability in fading channels and it is shown that the optimal power allocation policy is to employ the "save-then-transmit" protocol, that is, no power transmission occurs until the accumulated power becomes sufficiently high and then transmission is performed continuously with non-decreasing power. In [7], the optimal power adaptation is considered for a frequency-selective fading channel in the context of energy efficiency. Similarly, energy-efficient optimal power allocation is also studied in [8] for an orthogonal frequency division multiple access (OFDMA) system in which flat fading channels exist. In [9], the optimal power allocation is considered for multiple-input multiple-output (MIMO) flat fading channels in order to maximize the effective SNR under sum energy and total block length constraints. In [10], an energy-efficient power allocation method is proposed for Nakagami-m flat-fading channels in the presence of delay-outage probability constraint. For cooperative wireless sensor networks, an optimized dynamic power control approach is proposed in [11] with the consideration of quality of service (QoS) requirements. In [12], the optimal power and rate adaptation scheme is investigated in order to maximize the spectral efficiency of a communication system where multilevel quadrature amplitude modulation (MQAM) is considered over Rayleigh flat-fading channels. The common thread in this line of research is to devise power allocation algorithms that can adapt to varying channel conditions in an efficient, preferably optimal, manner and hence improve the system performance beyond that of the conventional uniform power allocation approach.

Although the problem of optimal power allocation over fading channels has been considered under various criteria such as channel capacity (e.g., [1]–[3]), BER (e.g., [4], [5], [13]–[15]), outage probability (e.g., [6], [16], [17]), and energy efficiency (e.g., [3], [7], [8]), no studies in the literature have investigated the optimal power allocation problem over fading channels within the context of the Neyman-Pearson framework. This can, in part, be attributed to the lack of closed form solutions and the nonconvex nature of the optimization problem for practical values of the false alarm rate. In particular, results for the convexity properties of the detection probability are established in [18] for the problem of determining the presence

S. Sarıtaş, A. D. Sezer, and S. Gezici are with the Department of Electrical and Electronics Engineering, Bilkent University, 06800, Ankara, Turkey, emails: {serkan,adsezer,gezici}@ee.bilkent.edu.tr. B. Dulek is with the Department of Electrical and Electronics Engineering, Hacettepe University, 06800, Ankara, Turkey, email: berkan@ee.hacettepe.edu.tr. S. Yüksel is with the Department of Mathematics and Statistics, Queen's University, Kingston, Ontario, Canada, K7L 3N6, email: yuksel@mast.queensu.ca. B. Dulek acknowledges the support by the National Young Researchers Career Development Program (project no. 215E118) of the Scientific and Technological Research Council of Turkey (TUBITAK). A. D. Sezer is supported by ASELSAN Graduate Scholarship for Turkish Academicians.

of a transmitted signal immersed in additive Gaussian noise in the absence of fading. In addition to the convexity analysis, the optimal power allocation strategy is derived for an average power constrained jammer trying to reduce the detection probability at the receiver. In a related study [19], the detection probability is analyzed for cooperative spectrum sensing over Rayleigh fading channels in cognitive radio systems, and it is concluded that cooperation among secondary users improves the detection performance.

In this paper, the optimal power allocation problem is considered for maximizing the average detection probability over a flat fading AWGN channel subject to average and peak power constraints. In order to obtain the optimal power allocation policy, the convexity properties of the detection probability are analyzed with respect to the received SNR, which is subject to flat fading. Then, it is shown that a dual problem that admits dual decomposition with no duality gap can be constructed. Based on the primal and dual formulations, various algorithms are proposed for the optimal power allocation. Furthermore, the optimal power allocation strategies are characterized according to the desired false alarm probability and it is obtained that the optimal power allocation scheme for the maximization of average detection probability is completely different from the uniform power allocation strategy. Numerical examples are presented to investigate the validity of the theoretical results. The main contributions of this study can be summarized as follows:

- For the first time in the literature, solutions to the optimal power allocation problem are proposed for average detection probability maximization in the presence of flat fading AWGN channels.
- The formulated optimization problem is generic in the sense that it takes into account both average and peak power constraints and it applies to any continuous fading distribution (Sec. II).
- Using results from measure theory and convex analysis, it is shown that the Lagrangian dual problem admits no duality gap. This, in turn, provides an efficient framework for the solution of the original optimization problem, which is nonconvex in general (Secs. III-A–III-C).
- The computational complexity of the problem is reduced significantly by applying dual decomposition. The resulting subproblems are coupled only through a single parameter (Sec. III-D).
- The proposed algorithms are guaranteed to converge to the global optimum with desired accuracy (Secs. III-E–III-F).
- For various probabilities of false alarm, the continuity and monotonicity properties of the optimal power allocation scheme are investigated and the conditions, which must be satisfied by any optimal power allocation policy, are derived (Theorem 1 and Theorem 2 in Sec. III-H).

The rest of the paper is organized as follows: Sec. II presents the problem formulation for the optimal power allocation subject to the average and peak power constraints. In Sec. III, the solution of the optimization problem and the optimal power allocation algorithms are provided together with the theorems characterizing optimal power allocation. In Sec. IV, numerical examples are presented to corroborate the theoretical results. Finally, Sec. V concludes the paper with remarks.

II. Problem Formulation

Consider a transmitter and a receiver that communicate over a flat fading AWGN channel. The task of the receiver is to decide between two hypotheses, $\mathcal{H}_0$ and $\mathcal{H}_1$, which correspond to the absence and presence of a signal, respectively. The observation model under each hypothesis is expressed as follows:

$$\mathcal{H}_0 : Y = \sigma N, \quad \mathcal{H}_1 : Y = \sqrt{P_t}\, s\, h + \sigma N \tag{1}$$

where Y denotes a real-valued scalar observation,[1] N is a standard Gaussian random variable with zero mean and unit variance; i.e., $N \sim \mathcal{N}(0,1)$, $\sigma > 0$ is the standard deviation of the noise at the receiver, $\sqrt{P_t}\, s$ denotes the transmitted signal, and h denotes the scalar channel gain after carrier phase synchronization at the receiver, which is assumed to be nonzero. Without loss of generality, it is assumed that $s = 1$ in (1); hence, P_t represents the power allocated by the transmitter. It is noted that the scalar observation model in (1) provides an abstraction for a continuous-time system which processes the received signal by down-conversion, matched filtering and sampling at the symbol rate with precise symbol timing; hence, the effects of modulator, additive noise channel, fading, and receiver front-end processing are taken into account in the discrete-time baseband model [18], [20]. In addition, it is assumed that the receiver has the knowledge of the channel coefficient h (i.e., perfect CSI) and the standard deviation of the noise, σ.

In this work, the Neyman-Pearson (NP) criterion is considered; i.e., the receiver implements the optimal NP decision rule which maximizes the probability of detection subject to a constraint on the probability of false alarm, denoted by α [21].[2] In accordance with the NP criterion, the likelihood ratio test (LRT) corresponding to (1) is obtained as follows:

$$\mathcal{L}(Y) = \frac{\frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-\left(Y-\sqrt{P_t}h\right)^2}{2\sigma^2}}}{\frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-Y^2}{2\sigma^2}}} = e^{\frac{\sqrt{P_t}hY}{\sigma^2} - \frac{P_t h^2}{2\sigma^2}} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta\,, \tag{2}$$

which can be simplified into

$$\operatorname{sgn}(h)\, Y \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \frac{\sigma^2 \ln \eta}{|h|\sqrt{P_t}} + \frac{\sqrt{P_t}|h|}{2} \triangleq \tilde{\eta}\,. \tag{3}$$

The optimum NP decision rule satisfies the constraint on the probability of false alarm with equality [21]. From (1) and (3), the probability of false alarm can be obtained as $\mathrm{P_F} = \Pr[\mathcal{L}(Y) \geqslant \eta \,|\, \mathcal{H}_0] = \Pr[\operatorname{sgn}(h)\, Y \geqslant \tilde{\eta} \,|\, \mathcal{H}_0] = \mathcal{Q}(\tilde{\eta}/\sigma)$, where $\mathcal{Q}(\cdot)$ denotes the $\mathcal{Q}$-function; i.e., $\mathcal{Q}(x) = (\sqrt{2\pi})^{-1} \int_x^\infty e^{-t^2/2} \mathrm{d}t$. Setting the probability of false alarm equal to α, the threshold is calculated as $\tilde{\eta} = \sigma \mathcal{Q}^{-1}(\alpha)$, where $\mathcal{Q}^{-1}(\cdot)$ is the inverse $\mathcal{Q}$-function. Hence, the NP decision rule is given by

$$\operatorname{sgn}(h)\, Y \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \sigma \mathcal{Q}^{-1}(\alpha)\,. \tag{4}$$

The detection probability corresponding to the decision rule in (4) can be obtained from (1) as

$$\begin{aligned} \mathrm{P_D} &= \Pr[\operatorname{sgn}(h)\, Y \geqslant \sigma \mathcal{Q}^{-1}(\alpha) \,|\, \mathcal{H}_1] \\ &= \mathcal{Q}\left(\mathcal{Q}^{-1}(\alpha) - \frac{\sqrt{P_t}|h|}{\sigma}\right) \triangleq \mathcal{Q}\left(\mathcal{Q}^{-1}(\alpha) - \sqrt{P_t \gamma}\right) \end{aligned} \tag{5}$$

[1] The results can also be extended to vector observations (see Sec. V).

[2] The NP framework is well-suited for applications in which the detection and false alarm events have different importance levels. As an example, consider a scenario in which the transmitter, equipped with some sensors, sends a signal to the receiver whenever it detects the presence of a person in a restricted area (or, fire in a forest).

where $\gamma \triangleq h^2/\sigma^2$. In the presence of a signal, γ determines the signal-to-noise ratio (SNR) at the receiver since $|h|$ represents the channel gain and σ^2 denotes the average noise power (see (1)). In the sequel, it is assumed that γ takes values in an interval $\Gamma \subset \mathbb{R}^+$ and that the transmitter has the knowledge of γ, which is commonly provided via feedback from the receiver in practice [22, Ch. 4]. Equipped with the knowledge of γ, it is assumed that the transmitter can perform power adaptation; i.e., the transmit power can be adjusted based on the current value of γ according to the power adaptation strategy given by $P_t(\gamma) : \Gamma \to [0, \infty)$. Consequently, the detection probability in (5) can be written as

$$\begin{aligned} \mathrm{P_D}(P_t(\gamma),\gamma) &= \mathcal{Q}\left(\mathcal{Q}^{-1}(\alpha) - \sqrt{P_t(\gamma)\gamma}\right) \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathcal{Q}^{-1}(\alpha)-\sqrt{P_t(\gamma)\gamma}}^{\infty} e^{-\frac{x^2}{2}}\, \mathrm{d}x. \end{aligned} \quad (6)$$

Although the optimal power allocation problem in the presence of CSI at the transmitter has been investigated for various metrics such as Shannon capacity, outage capacity, and average probability of error (e.g., [1]–[17]), the optimal power allocation problem for the maximization of average detection probability over flat fading AWGN channels has not been considered to the best of authors' knowledge. The aim in this work is to obtain the optimal power allocation strategy that maximizes the average detection probability under an average power constraint, i.e., to solve the following optimization problem:

$$\sup_{P_t(\gamma)} \ \mathbb{E}_\gamma[\mathrm{P_D}(P_t(\gamma),\gamma)] \quad \text{s.t.} \ \ \mathbb{E}_\gamma[P_t(\gamma)] \le \overline{P}, \quad (7)$$

where $\mathbb{E}_\gamma[\cdot]$ denotes the expectation over the continuous random variable γ, $\overline{P}$ denotes the average transmit power limit, and $P_t(\gamma)$ is a Lebesgue-measurable function with $0 \le P_t(\gamma) \le P_{peak} \ \ \forall \gamma \in \Gamma$, and P_{peak} denotes the peak power constraint satisfying $P_{peak} > \overline{P}$. More explicitly,

$$\begin{aligned} \sup_{0\le P_t(\gamma)\le P_{peak}} & \int_{\gamma\in\Gamma} \mathcal{Q}\left(\mathcal{Q}^{-1}(\alpha) - \sqrt{P_t(\gamma)\gamma}\right) p(\gamma)\, \mathrm{d}\gamma \\ \text{s.t.} \quad & \int_{\gamma\in\Gamma} P_t(\gamma) p(\gamma)\, \mathrm{d}\gamma \le \overline{P}, \end{aligned} \quad (8)$$

where $p(\gamma)$ is the probability density function (PDF) of γ and satisfies the conditions for a valid PDF, i.e., $p(\gamma) \ge 0 \ \ \forall \gamma \in \Gamma$, and $\int_{\gamma\in\Gamma} p(\gamma)\, \mathrm{d}\gamma = 1$.

Remark 1: It is noted from (8) that the transmitter calculates the average detection probability and the average power by using the PDF of γ, which must be estimated in practice. Such an estimation process can be performed when the channel characteristics are constant for a sufficiently long time interval (e.g., when the transmitter and the receiver stay in the same environment for some time and do not move very rapidly). In the presence of imperfect estimation, the results in this study can be regarded as theoretical upper bounds on the average detection probability.

III. Optimal Power Allocation

In this section, first, the convexity properties of the detection probability are analyzed with respect to the transmitted signal power. Then, the dual of the optimal power allocation problem is formulated and it is shown that the duality gap between the original problem and the dual problem is zero. In order to solve the dual problem, the dual decomposition method is presented, and the related algorithms are provided in order to obtain the optimal power allocation strategy numerically. Finally, the properties of the optimal power allocation strategy are investigated for various probabilities of false alarm.

A. Convexity/Concavity Properties

To obtain the optimal power allocation policy based on the optimization problem in (7) (equivalently, (8)), the convexity properties of the detection probability are discussed with respect to the transmitted signal power based on the results obtained in [18].

Lemma 1. *For $\alpha \in [\mathcal{Q}(2), 1)$, $\mathrm{P_D}(P_t(\gamma),\gamma)$ is a monotonically increasing and strictly concave function of $P_t(\gamma) \in (0,\infty)$ for any given value of $\gamma \in \Gamma$. For $\alpha \in (0, \mathcal{Q}(2))$, $\mathrm{P_D}(P_t(\gamma),\gamma)$ is a monotonically increasing function with two inflection points $I_1(\gamma) < I_2(\gamma)$ such that $\mathrm{P_D}(P_t(\gamma),\gamma)$ is strictly concave for $P_t(\gamma) \in (0, I_1(\gamma))$, strictly convex for $P_t(\gamma) \in (I_1(\gamma), I_2(\gamma))$, and strictly concave for $P_t(\gamma) \in (I_2(\gamma), \infty)$ for any given value of $\gamma \in \Gamma$.*

Proof: The proof is similar to that of Proposition 1 in [18], which was derived in the absence of fading (hence, no power allocation with respect to fading). First, the limits of the detection probability are noted. For any fixed value of $\gamma \in \Gamma$, $\lim_{P_t(\gamma)\to 0} \mathrm{P_D}(P_t(\gamma),\gamma) = \alpha$ and $\lim_{P_t(\gamma)\to\infty} \mathrm{P_D}(P_t(\gamma),\gamma) = 1$. Furthermore, the first-order derivative of the detection probability with respect to $P_t(\gamma)$ is given by

$$\begin{aligned} \frac{\partial \mathrm{P_D}(P_t(\gamma),\gamma)}{\partial P_t(\gamma)} &= \frac{\sqrt{\gamma}}{2\sqrt{2\pi}\sqrt{P_t(\gamma)}} \\ &\quad \times \exp\left\{-\frac{1}{2}\left(\mathcal{Q}^{-1}(\alpha) - \sqrt{P_t(\gamma)\gamma}\right)^2\right\}, \end{aligned} \quad (9)$$

which is positive for all values of $P_t(\gamma) > 0$ and $\gamma \in \Gamma$. Hence the detection probability is a strictly increasing function of the transmit power $P_t(\gamma)$. Similarly, the limits of the first-order derivative are given as $\lim_{P_t(\gamma)\to 0} \frac{\partial \mathrm{P_D}(P_t(\gamma),\gamma)}{\partial P_t(\gamma)} = \infty$ and $\lim_{P_t(\gamma)\to\infty} \frac{\partial \mathrm{P_D}(P_t(\gamma),\gamma)}{\partial P_t(\gamma)} = 0$. Differentiating once more with respect to $P_t(\gamma)$ yields

$$\begin{aligned} \frac{\partial^2 \mathrm{P_D}(P_t(\gamma),\gamma)}{\partial \left(P_t(\gamma)\right)^2} &= \left(\gamma P_t(\gamma) - \mathcal{Q}^{-1}(\alpha)\sqrt{P_t(\gamma)\gamma} + 1\right) \\ &\times \underbrace{\frac{-\sqrt{\gamma}}{4\sqrt{2\pi}(P_t(\gamma))^{3/2}} \exp\left\{-\frac{1}{2}\left(\mathcal{Q}^{-1}(\alpha) - \sqrt{P_t(\gamma)\gamma}\right)^2\right\}}_{\triangleq A(P_t(\gamma),\gamma)}. \end{aligned} \quad (10)$$

Since $A(P_t(\gamma),\gamma)$ is negative for all $P_t(\gamma) > 0$ and $\gamma \in \Gamma$, the sign of the second-order derivative is determined by the first term, $\left(\gamma P_t(\gamma) - \mathcal{Q}^{-1}(\alpha)\sqrt{P_t(\gamma)\gamma} + 1\right)$. Let $x \triangleq \sqrt{P_t(\gamma)\gamma}$. Then, in order to identify the sign of the second-order derivative, we need to check the sign of $f(x) \triangleq x^2 - \mathcal{Q}^{-1}(\alpha)x + 1$ for $x > 0$, which can be determined from its discriminant, $\Delta = \left(\mathcal{Q}^{-1}(\alpha)\right)^2 - 4$. For $\alpha \in (\mathcal{Q}(2), \mathcal{Q}(-2))$, the discriminant is negative, which indicates that $f(x) > 0$ $\forall x > 0$. For $\Delta = (\mathcal{Q}^{-1}(\alpha))^2 - 4 > 0$, the real roots of $f(x)$ occur at $x_{1,2} = \left(\mathcal{Q}^{-1}(\alpha) \pm \sqrt{\left(\mathcal{Q}^{-1}(\alpha)\right)^2 - 4}\right)/2$. If $\mathcal{Q}^{-1}(\alpha) \ge 2$, we have $\alpha \le \mathcal{Q}(2)$ and both roots are positive. Thus, $f(x) > 0$ for $0 \le x < x_1$ and $x > x_2$, whereas $f(x) < 0$ for $x_1 < x < x_2$. On the other hand, if $\mathcal{Q}^{-1}(\alpha) \le -2$, that

is, if $\alpha \geq \mathcal{Q}(-2)$, then both roots are negative, which implies that $f(x) > 0$ for all $x \geq 0$.

From the analysis above, it follows that $\mathrm{P_D}(P_t(\gamma), \gamma)$ is a monotonically increasing and strictly concave function of $P_t(\gamma) \in (0, \infty)$ for $\alpha \in (\mathcal{Q}(2), 1)$. For $\alpha \in (0, \mathcal{Q}(2))$, $\mathrm{P_D}(P_t(\gamma), \gamma)$ is a monotonically increasing function with two inflection points $I_1(\gamma) < I_2(\gamma)$, where

$$I_1(\gamma) = \frac{1}{4\gamma}\left(\mathcal{Q}^{-1}(\alpha) - \sqrt{\left(\mathcal{Q}^{-1}(\alpha)\right)^2 - 4}\right)^2$$
$$I_2(\gamma) = \frac{1}{4\gamma}\left(\mathcal{Q}^{-1}(\alpha) + \sqrt{\left(\mathcal{Q}^{-1}(\alpha)\right)^2 - 4}\right)^2 \qquad (11)$$

such that $\mathrm{P_D}(P_t(\gamma), \gamma)$ is strictly concave for $P_t(\gamma) \in (0, I_1(\gamma))$, strictly convex for $P_t(\gamma) \in (I_1(\gamma), I_2(\gamma))$, and strictly concave for $P_t(\gamma) \in (I_2(\gamma), \infty)$ for any given value of $\gamma \in \Gamma$. ■

Based on Lemma 1, when $\alpha \in [\mathcal{Q}(2), 1)$, the optimization problem in (8) becomes a convex optimization problem since $\mathrm{P_D}(P_t(\gamma), \gamma)$ is a concave function of $P_t(\gamma)$ for all values of $P_t(\gamma) > 0$. However, in many practical applications, the required values for the probability of false alarm are smaller than $\mathcal{Q}(2) \approx 0.02275$. In such cases, i.e., for $\alpha < \mathcal{Q}(2)$, the optimization problem in (8) is in general nonconvex since $\mathrm{P_D}(P_t(\gamma), \gamma)$ is no longer a concave function of $P_t(\gamma)$ for all values of $P_t(\gamma) > 0$. Nonetheless, based on the results established in [23], it can be shown that the duality gap of the optimization problem is zero (Sec. III-C). This, in turn, leads to efficient numerical algorithms for the solution of the nonconvex optimization problem in the dual domain, as discussed in the sequel.

B. Dual Problem

For the optimization problem in (8), the corresponding Lagrangian function is expressed as

$$\mathcal{L}\left(P_t(\gamma), \lambda\right) = \int_{\gamma\in\Gamma} \mathcal{Q}\left(\mathcal{Q}^{-1}(\alpha) - \sqrt{P_t(\gamma)\gamma}\right) p(\gamma)\,\mathrm{d}\gamma \qquad (12)$$
$$+ \lambda\left(\overline{P} - \int_{\gamma\in\Gamma} P_t(\gamma) p(\gamma)\,\mathrm{d}\gamma\right)$$

$$= \int_{\gamma\in\Gamma} \left(\mathcal{Q}\left(\mathcal{Q}^{-1}(\alpha) - \sqrt{P_t(\gamma)\gamma}\right) - \lambda P_t(\gamma)\right) p(\gamma)\,\mathrm{d}\gamma + \lambda\overline{P},$$

and the dual function is given by

$$g(\lambda) \triangleq \sup_{P_t(\gamma)} \quad \mathcal{L}\left(P_t(\gamma), \lambda\right)$$
$$\text{s.t.} \quad 0 \leq P_t(\gamma) \leq P_{peak},\ \forall\gamma \in \Gamma$$
$$P_t(\cdot) \text{ is Lebesgue measurable} \qquad (13)$$

Then, the Lagrangian dual problem of (8) is defined as

$$\min_{\lambda}\ g\left(\lambda\right) \ \text{ s.t. } \lambda \geq 0\,. \qquad (14)$$

Let P^* and D^* denote the optimal values obtained as the solutions of the original problem in (8) and its dual in (14). It should be noted that the latter optimization problem is convex whereas the former is not necessarily so. From weak duality, it follows that $P^* \leq D^*$ [24]. In general, the primal in (8) is not equivalent to the dual in (14). In the following, it is shown that the duality gap is zero when γ takes values in an interval Γ. Hence, strong duality holds and the solution of (8) can be obtained from the solution of its dual in (14).

C. Strong Duality

In order to show that the duality gap between (8) and (14) is zero, we follow a similar approach to that employed in [23] and [25], which relies on a variant of Lyapunov theorem due to Blackwell [26], [27].

Lemma 2. *[23, Lemma 1], [25, Theorem 1] Let ν be a nonatomic[3] measure on a Borel field $\mathcal{B}$ generated from subsets of a space Γ. Let $g_i\left(x\left(\cdot\right), \cdot\right)$ be a $\mathcal{B}$-measurable function whenever $x\left(\cdot\right)$ is $\mathcal{B}$-measurable for $i = 1, 2, \ldots, m$. Then,*

$$\left\{ \begin{pmatrix} \int_\Gamma g_1\left(x\left(\cdot\right), \cdot\right) \mathrm{d}\nu \\ \int_\Gamma g_2\left(x\left(\cdot\right), \cdot\right) \mathrm{d}\nu \\ \vdots \\ \int_\Gamma g_m\left(x\left(\cdot\right), \cdot\right) \mathrm{d}\nu \end{pmatrix} \middle| \, x \text{ is } \mathcal{B}\text{-measurable } \& \ x \in [0, x_{\max}] \right\}$$

is a convex set.

It should be emphasized that no assumption is imposed in Lemma 2 on the convexity of functions g_i or the set Γ. The convexity of the image of the mapping stems from the nonatomic property of measure ν. This condition is satisfied if an absolutely continuous probability measure with PDF $p(\gamma)$ is assumed in the problem formulation. Then, the following result is obtained.

Proposition 1. *Let $v\left(\overline{P}\right)$ denote the solution of* (8) *for an absolutely continuous probability measure with PDF $p(\gamma)$. Then, $v\left(\overline{P}\right)$ is a concave function of $\overline{P}$.*

Proof: The statement in the proposition can be proved based on similar arguments to those in Theorem 7 of [23]. Let $\overline{P}^1$ and $\overline{P}^2$ represent two average power limits, and let $P_t^1\left(\gamma\right)$ and $P_t^2\left(\gamma\right)$ denote $\varepsilon-$optimal solutions of the optimization problem in (8) under average power limits $\overline{P}^1$ and $\overline{P}^2$, respectively, so that $v\left(\overline{P}^i\right) \leq \overline{P}_D{}^i = \int_{\gamma\in\Gamma} \mathcal{Q}\left(\mathcal{Q}^{-1}(\alpha) - \sqrt{P_t^i(\gamma)\gamma}\right) p(\gamma)\,\mathrm{d}\gamma + \varepsilon$. Then, Lemma 2 implies that for each $0 \leq \theta \leq 1$, there exists a nonnegative Lebesgue measurable function $\widehat{P}_t\left(\gamma\right)$ such that $\int_{\gamma\in\Gamma} \mathcal{Q}\left(\mathcal{Q}^{-1}(\alpha) - \sqrt{\widehat{P}_t\left(\gamma\right)\gamma}\right) p(\gamma)\,\mathrm{d}\gamma = \theta\overline{P}_D{}^1 + (1-\theta)\overline{P}_D{}^2 - \varepsilon$ and $\int_{\gamma\in\Gamma} \widehat{P}_t\left(\gamma\right) p(\gamma)\,\mathrm{d}\gamma = \theta\int_{\gamma\in\Gamma} P_t^1\left(\gamma\right) p(\gamma)\,\mathrm{d}\gamma + (1-\theta)\int_{\gamma\in\Gamma} P_t^2\left(\gamma\right) p(\gamma)\,\mathrm{d}\gamma \leq \theta\overline{P}^1 + (1-\theta)\overline{P}^2$. This holds for every $\varepsilon > 0$; therefore, the supremum of (8) under an average power limit $\theta\overline{P}^1 + (1-\theta)\overline{P}^2$ satisfies $v\left(\theta\overline{P}^1 + (1-\theta)\overline{P}^2\right) \geq \theta\overline{P}_D{}^1 + (1-\theta)\overline{P}_D{}^2 = \theta v\left(\overline{P}^1\right) + (1-\theta)v\left(\overline{P}^2\right)$, from which the concavity of $v\left(\overline{P}\right)$ with respect to $\overline{P}$ follows. ■

In Proposition 1, it is stated that the optimal value $v\left(\overline{P}\right)$ of the objective function in (8) is a concave function of the average power limit $\overline{P}$ for absolutely continuous $p(\gamma)$. Then, it follows that the Lagrangian dual problem admits no duality gap with the original problem [28, Theorem 2], [29]. Hence, the following corollary is obtained.

Corollary 1. *The duality gap between the solutions of* (8) *and* (14) *is zero.*

[3]A measure is nonatomic if every set of nonzero measure has a subset with strictly less nonzero measure. The standard Lebesgue measure is nonatomic. The uniform measure on a finite set is atomic [23].

D. Dual Decomposition

Since the equivalence of the primal and dual formulations is now established, the solution of the optimization problem can be investigated based on the dual problem. The dual function in (13) involves the maximization of Lagrangian function $\mathcal{L}\left(P_t(\gamma),\lambda\right)$ for a given value of λ. It is observed from (12) that the Lagrangian function $\mathcal{L}\left(P_t(\gamma),\lambda\right)$ can be decomposed into

$$\mathcal{L}\left(P_t(\gamma),\lambda\right)=\int_{\gamma\in\Gamma}\mathcal{L}_\gamma\left(P_t(\gamma),\lambda\right)p(\gamma)\,\mathrm{d}\gamma+\lambda\overline{P}\,, \qquad (15)$$

where $\mathcal{L}_\gamma\left(P_t(\gamma),\lambda\right)\triangleq\mathcal{Q}\left(\mathcal{Q}^{-1}(\alpha)-\sqrt{P_t(\gamma)\gamma}\right)-\lambda P_t(\gamma)$. Evidently, the optimal power allocation that maximizes $\mathcal{L}\left(P_t(\gamma),\lambda\right)$ obtained from (13) should also maximize $\mathcal{L}_\gamma\left(P_t(\gamma),\lambda\right)$ for each given value of γ. This is known as dual decomposition and it facilitates the decomposition of the dual problem into suboptimization problems which are coupled only through λ [24]. More explicitly, we need to compute

$$\begin{aligned}&\sup_{0\le P_t(\gamma)\le P_{peak}}\mathcal{L}_\gamma\left(P_t(\gamma),\lambda\right)\\&=\sup_{0\le P_t(\gamma)\le P_{peak}}\mathcal{Q}\left(\mathcal{Q}^{-1}(\alpha)-\sqrt{P_t(\gamma)\gamma}\right)-\lambda P_t(\gamma)\end{aligned} \qquad (16)$$

for each value of $\gamma\in\Gamma$. It is also required to search through values of λ which place sufficient emphasis on the power constraint term in $\mathcal{L}_\gamma\left(P_t(\gamma),\lambda\right)$ so that the average power constraint in (8) is satisfied.

E. Algorithms

In this part, two algorithms are presented for the proposed optimum power allocation problem over flat fading AWGN channels. Both algorithms contain a loop that searches over λ. The first algorithm employs a subgradient method to iteratively update λ, whereas the second algorithm employs a bisection method [4], [30], [31]. In both methods, the search direction for λ suggests that λ should increase if the constraint is exceeded; i.e., $\int_{\gamma\in\Gamma}P_t(\gamma)p(\gamma)\,\mathrm{d}\gamma>\overline{P}$, and decrease otherwise. This is because a larger value of λ places more emphasis on the power constraint in the Lagrangian and results in a lower average power.

Algorithm 1. *Optimal Power Allocation Algorithm - Subgradient Method*

Initialize $\lambda_1,\ k=1$
do
solve ${P_t}^*(\gamma)=\underset{x\in[0,P_{peak}]}{\operatorname{argsup}}\ \mathcal{Q}\left(\mathcal{Q}^{-1}(\alpha)-\sqrt{x\gamma}\right)-\lambda_k\ \forall\gamma\in\Gamma$
$\lambda_{k+1}=\left[\lambda_k+\alpha_k\left(\int_{\gamma\in\Gamma}{P_t}^*(\gamma)p(\gamma)\,\mathrm{d}\gamma-\overline{P}\right)\right]^+$
$k=k+1$
while $|\lambda_{k+1}-\lambda_k|>\epsilon$

In Algorithm 1, k is the iteration number, $\alpha_k>0$ is the step size for the kth iteration (a decreasing sequence of k), $[\cdot]^+\triangleq\max\{\cdot,0\}$, and $\epsilon>0$ is a small number used to signal convergence. The subgradient update is guaranteed to converge to the optimal value of λ as long as α_k is chosen to be sufficiently small [32]. As mentioned in [31, Sec. IV-A], when the norm of the subgradient is bounded, the choice of $\alpha_k=\beta/k$ is guaranteed to converge to the optimal for some constant β.

The second algorithm, which relies on a bisection search to update λ and converges in general faster than the subgradient method [30], [31], is described next.

Algorithm 2. *Optimal Power Allocation Algorithm - Bisection Method*

Initialize $\lambda_{min}=0,\ \lambda_{max}$ (described in Algorithm 3)
do
$\lambda=\frac{\lambda_{min}+\lambda_{max}}{2}$
solve ${P_t}^*(\gamma)=\underset{x\in[0,P_{peak}]}{\operatorname{argsup}}\ \mathcal{Q}\left(\mathcal{Q}^{-1}(\alpha)-\sqrt{x\gamma}\right)-\lambda x\ \forall\gamma\in\Gamma$
if $\int_{\gamma\in\Gamma}{P_t}^*(\gamma)p(\gamma)\mathrm{d}\gamma\le\overline{P}$, **then** $\lambda_{max}=\lambda$, **else** $\lambda_{min}=\lambda$
while $|\lambda_{max}-\lambda_{min}|>\epsilon$

In the initialization stage of the bisection based algorithm, it is necessary to find a value of λ_{max} that guarantees that the average power constraint is satisfied. Algorithm 3 tackles this problem.

Algorithm 3. *How to compute λ_{max}*

$\lambda_{max}=1$
while $\int_{\gamma\in\Gamma}{P_t}^*(\gamma)p(\gamma)\,\mathrm{d}\gamma>\overline{P}$
$\lambda_{max}=2\lambda_{max}$
solve ${P_t}^*(\gamma)=\underset{x\in[0,P_{peak}]}{\operatorname{argsup}}\ \mathcal{Q}\left(\mathcal{Q}^{-1}(\alpha)-\sqrt{x\gamma}\right)-\lambda_{max}x\ \forall\gamma\in\Gamma$
end

Although we have decoupled the original optimization problem across different values of γ (for fixed λ) via dual decomposition, we still need to solve a nonconvex optimization problem (for $\alpha<\mathcal{Q}(2)$) at each iteration to compute ${P_t}^*(\gamma)=\underset{x\in[0,P_{peak}]}{\operatorname{argsup}}\ \mathcal{Q}\left(\mathcal{Q}^{-1}(\alpha)-\sqrt{x\gamma}\right)-\lambda x$ for all $\gamma\in\Gamma$. Fortunately, the optimal solution for ${P_t}^*(\gamma)$ can be obtained with desired accuracy using tools from convex optimization. This is explained in the next part.

F. Subroutines

In Sec. III-A, it is shown that $\mathrm{P_D}(x)\triangleq\mathcal{Q}\left(\mathcal{Q}^{-1}(\alpha)-\sqrt{x\gamma}\right)$ is a monotonically increasing and strictly concave function of $x\in(0,\infty)$ for all $\alpha\ge\mathcal{Q}(2)$ and $\gamma>0$. Therefore, the following optimization problem

$$\begin{aligned}{P_t}^*(\gamma)&=\underset{x\in[0,P_{peak}]}{\operatorname{argsup}}\ \mathrm{P_D}(x)-\lambda x\\&=\underset{x\in[0,P_{peak}]}{\operatorname{argsup}}\ \mathcal{Q}\left(\mathcal{Q}^{-1}(\alpha)-\sqrt{x\gamma}\right)-\lambda x\end{aligned} \qquad (17)$$

is convex for the specified range of parameter values. If $\mathrm{P_D}'(P_{peak})=\frac{\sqrt{\gamma}}{2\sqrt{2\pi}\sqrt{P_{peak}}}\exp\left\{-\frac{1}{2}\left(\mathcal{Q}^{-1}(\alpha)-\sqrt{P_{peak}\gamma}\right)^2\right\}\ge\lambda$, then ${P_t}^*(\gamma)=P_{peak}$. If $\mathrm{P_D}'(P_{peak})<\lambda$, we need to numerically evaluate $\mathrm{P_D}'(x)=\lambda$ or more explicitly $\frac{\sqrt{\gamma}}{2\sqrt{2\pi}\sqrt{x}}\exp\left\{-\frac{1}{2}\left(\mathcal{Q}^{-1}(\alpha)-\sqrt{x\gamma}\right)^2\right\}=\lambda$. Since $\mathrm{P_D}'(x)$ is a monotonically decreasing function (from infinity to 0) of x and λ is a constant, there is a unique ${P_t}^*(\gamma)$ which can be calculated based on a simple bisection search, which is described as follows:

Algorithm 4. *Solution for concave $\mathrm{P_D}(x)$*

$x_{min}=0$
$x_{max}=P_{peak}$
do

$x = \frac{x_{min}+x_{max}}{2}$
if $\mathrm{P_D}'(x) > \lambda$, **then** $x_{min} = x$, **else** $x_{max} = x$
while $|x_{max} - x_{min}| > \epsilon$

On the other hand, for $\alpha \in (0, \mathcal{Q}(2))$, it is shown in Sec. III-A that $\mathrm{P_D}(x) = \mathcal{Q}\left(\mathcal{Q}^{-1}(\alpha) - \sqrt{x\gamma}\right)$ is a monotonically increasing function with two inflection points $I_1(\gamma)$ and $I_2(\gamma)$ (as specified in (11)) such that $\mathrm{P_D}(x)$ is strictly concave for $x \in (0, I_1(\gamma))$, strictly convex for $x \in (I_1(\gamma), I_2(\gamma))$, and strictly concave for $x \in (I_2(\gamma), \infty)$ for any given value of γ. Fig. 1 presents a generic graphical description of $\mathrm{P_D}(x)$ as a function of x for an arbitrary value of $\gamma > 0$ when $\alpha \in (0, \mathcal{Q}(2))$. Consequently, the optimization problem in (17) is not convex for $\alpha \in (0, \mathcal{Q}(2))$.

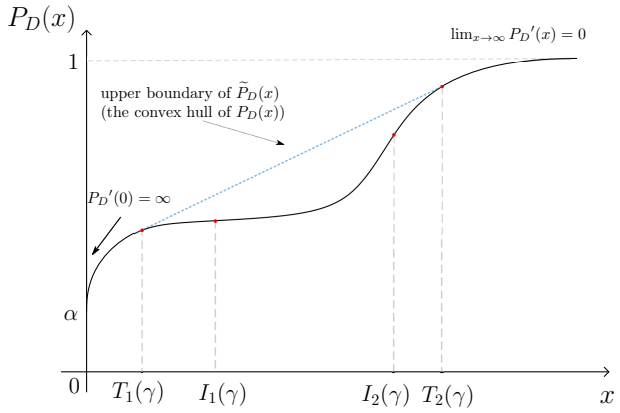


Fig. 1. An illustrative description of $\mathrm{P_D}(x)$ for an arbitrary value of $\gamma > 0$ when $\alpha \in (0, \mathcal{Q}(2))$. The tangent points $\{T_1(\gamma), T_2(\gamma)\}$ and the inflection points $\{I_1(\gamma), I_2(\gamma)\}$ are shown on the graph.

Based on a careful analysis of the behavior of $\mathrm{P_D}(x)$ in Fig. 1, efficient numerical methods are proposed for the solution of the optimization problem in (17) under different cases. Prior to the description of the proposed methods, the following lemmas are presented.

Lemma 3. *Let $\alpha \in (0, \mathcal{Q}(2))$, and $I_1(\gamma)$ and $I_2(\gamma)$ be the inflection points of $\mathrm{P_D}(x)$ as given in* (11)*. There exist unique points $T_1(\gamma) \in [0, I_1(\gamma)]$ and $T_2(\gamma) \geq I_2(\gamma)$ such that the tangent to $\mathrm{P_D}(x)$ at $T_1(\gamma)$ is also tangent at $T_2(\gamma)$ and this tangent lies above $\mathrm{P_D}(x)$ for all $\gamma > 0$.*

Proof: Similar to [18, Appendix A]. ■

Lemma 4. *Let $\lambda > 0$ and $\widetilde{P}_D(x)$ denote the upper boundary of the convex hull of $\mathrm{P_D}(x)$ (as depicted in Fig. 1). Then,* $\underset{x>0}{\text{argsup}}\ \mathrm{P_D}(x) - \lambda x = \underset{x>0}{\text{argsup}}\ \widetilde{P}_D(x) - \lambda x$.

Proof: Since $\mathrm{P_D}(x) \leq \widetilde{P}_D(x)$ for all $x > 0$, we get $\sup_{x>0} \mathrm{P_D}(x) - \lambda x \leq \sup_{x>0} \widetilde{P}_D(x) - \lambda x$ for all $x > 0$. Furthermore, $\widetilde{P}_D(x)$ is concave and the maximum occurs at $(\widetilde{P}_D)'(x^*) = \lambda$, where $x^* \in (0, T_1(\gamma)] \cup [T_2(\gamma), \infty)$ for all values of $\lambda > 0$. But noting that $\mathrm{P_D}(x) = \widetilde{P}_D(x)$ over $x \in (0, T_1(\gamma)] \cup [T_2(\gamma), \infty)$, the result follows. ■

Lemma 4 is the key observation in the development of our methods for the solution of (17). It indicates that the maximum of the nonconvex optimization problem $\underset{x>0}{\text{argsup}}\ \mathrm{P_D}(x) - \lambda x$ coincides with the maximum of the convex optimization problem $\underset{x>0}{\text{argsup}}\ \widetilde{P}_D(x) - \lambda x$, which can be computed easily by obtaining the solutions $x_1 = \underset{x\in(0,T_1(\gamma)]}{\text{argsup}}\ \mathrm{P_D}(x) - \lambda x$ and $x_2 = \underset{x\in[T_2(\gamma),\infty)}{\text{argsup}}\ \mathrm{P_D}(x) - \lambda x$, and selecting the solution with the highest score $x^* = \underset{\{x_1, x_2\}}{\text{argsup}}\ \mathrm{P_D}(x) - \lambda x$.[4] To this end, the tangent points $T_1(\gamma)$ and $T_2(\gamma)$ should be computed first. This can be achieved with desired accuracy using the following numerical method. (For a detailed explanation, see [18, Algorithm 1].)

Algorithm 5. *Computation of tangent points $T_1(\gamma)$ and $T_2(\gamma)$ when $\alpha \in (0, \mathcal{Q}(2))$*

$\beta_{min} = \mathrm{P_D}'(I_1(\gamma))$, $\beta_{max} = \mathrm{P_D}'(I_2(\gamma))$
$x_{min,1} = 0$, $x_{max,1} = I_1(\gamma)$
$x_{min,2} = I_2(\gamma)$, $x_{max,2} = \infty$
do
  $\beta = \frac{\beta_{min}+\beta_{max}}{2}$
  $x_1 = \underset{x\in(x_{min,1},x_{max,1})}{\text{argsup}}\ \mathrm{P_D}(x) - \beta x$
  $x_2 = \underset{x\in(x_{min,2},x_{max,2})}{\text{argsup}}\ \mathrm{P_D}(x) - \beta x$
  if $\mathrm{P_D}(x_1) - \beta x_1 > \mathrm{P_D}(x_2) - \beta x_2$
    then $\beta_{max} = \beta$, $x_{min,1} = x_1$, $x_{min,2} = x_2$
    else $\beta_{min} = \beta$, $x_{max,1} = x_1$, $x_{max,2} = x_2$
while $|\beta_{max} - \beta_{min}| > \epsilon$

At convergence, the tangent points and the slope of the tangent line that constitutes the upper boundary of the convex hull of $\mathrm{P_D}(x)$ corresponding to γ can be obtained as $T_1(\gamma) \approx x_1$, $T_2(\gamma) \approx x_2$, and $\mathrm{P_D}'(T_1(\gamma)) \approx \mathrm{P_D}'(T_2(\gamma)) \approx \beta$. Although $I_1(\gamma)$, $I_2(\gamma)$, $T_1(\gamma)$, and $T_2(\gamma)$ should be computed for all $\gamma \in \Gamma$ separately, they do not depend on the value of λ employed in (17). Consequently, they can be computed offline before either Algorithm 1 or Algorithm 2 is employed to find the optimal power allocation.

It should be noted that the peak power constraint is not employed in Lemma 4. In the sequel, we first present the proposed numerical method for the solution of (17) in the absence of a peak power constraint, i.e.,

$$P_t^*(\gamma) = \underset{x\geq 0}{\text{argsup}}\, \mathrm{P_D}(x) - \lambda x = \underset{x\geq 0}{\text{argsup}}\, \mathcal{Q}\left(\mathcal{Q}^{-1}(\alpha) - \sqrt{x\gamma}\right) - \lambda x$$

Algorithm 6. *Numerical method to compute $P_t^*(\gamma)$ in* (17) *without peak power constraint when $\alpha \in (0, \mathcal{Q}(2))$*

if $\mathrm{P_D}'(T_1(\gamma)) \leq \lambda$
  $P_t^*(\gamma) = \underset{x\in(0,T_1(\gamma)]}{\text{argsup}}\ \mathrm{P_D}(x) - \lambda x$
else
  $P_t^*(\gamma) = \underset{x\in[T_2(\gamma),\infty)}{\text{argsup}}\ \mathrm{P_D}(x) - \lambda x$

When a peak power constraint is imposed on the transmitter power as in (17), we can obtain the solution to (17) with some modifications depending on the relationship between $I_1(\gamma)$, $T_2(\gamma)$, and P_{peak}.

Case 1. $P_{peak} \leq I_1(\gamma)$: Since $\mathrm{P_D}(x)$ is concave for $x \leq I_1(\gamma)$, the optimization problem in (17) is convex and the following algorithm computes the solution with desired accuracy.

Algorithm 7. *Numerical method to compute $P_t^*(\gamma)$ in* (17) *for $P_{peak} \leq I_1(\gamma)$ when $\alpha \in (0, \mathcal{Q}(2))$*

[4] As will be seen in Algorithm 6, it is possible to improve on this result as well by noting that the optimal point x^* satisfies $\mathrm{P_D}'(x^*) = \lambda$ and $\mathrm{P_D}'(x)$ monotonically decreases over the intervals $(0, T_1(\gamma)]$ and $[T_2(\gamma), \infty)$ with $\mathrm{P_D}'(T_1(\gamma)) = \mathrm{P_D}'(T_2(\gamma))$.

if $\mathrm{P_D}'(P_{peak}) \geq \lambda$
$P_t{}^*(\gamma) = P_{peak}$
else
$P_t{}^*(\gamma) = \underset{x\in[0,P_{peak}]}{\text{argsup}} \ \mathrm{P_D}(x) - \lambda x$

Case 2. $P_{peak} \geq T_2(\gamma)$: In this case, the solution can be obtained with a slight modification to the one obtained assuming that no peak power constraint is imposed. This is because the convex hull of the upper boundary of $\mathrm{P_D}(x)$ is unchanged with respect to that scenario.

Algorithm 8. *Numerical method to compute* $P_t{}^*(\gamma)$ *in* (17) *for* $P_{peak} > T_2(\gamma)$ *when* $\alpha \in (0, \mathcal{Q}(2))$

if $\mathrm{P_D}'(P_{peak}) \geq \lambda$
$P_t{}^*(\gamma) = P_{peak}$
else if $\mathrm{P_D}'(P_{peak}) < \lambda < \mathrm{P_D}'(T_2(\gamma))$
$P_t{}^*(\gamma) = \underset{x\in[T_2(\gamma),P_{peak}]}{\text{argsup}} \ \mathrm{P_D}(x) - \lambda x$
else $\lambda \geq \mathrm{P_D}'(T_2(\gamma))$
$P_t{}^*(\gamma) = \underset{x\in[0,T_1(\gamma)]}{\text{argsup}} \ \mathrm{P_D}(x) - \lambda x$

Case 3. $I_1(\gamma) < P_{peak} < T_2(\gamma)$: Since the transmitter power x cannot take values greater than P_{peak}, the upper boundary of the convex hull of $\mathrm{P_D}(x)$ over the interval $[0,\, P_{peak}]$ is different from the previous cases. In order to present the solution of the optimization problem in (17) under this scenario, we need the following lemma.

Lemma 5. *Let* $\alpha \in (0,\, \mathcal{Q}(2))$*, and* $I_1(\gamma)$ *and* $I_2(\gamma)$ *be the inflection points of* $\mathrm{P_D}(x)$*. Suppose also that* $T_1(\gamma)$ *and* $T_2(\gamma)$ *be the contact points of the tangent line as described in Lemma 3. Given a point* $P_{peak} \in [I_1(\gamma),\, T_2(\gamma)]$*, there exists a unique point* $C(\gamma) \in [T_1(\gamma),\, I_1(\gamma)]$ *such that the tangent at* $C(\gamma)$ *passes through the point* $(P_{peak},\, \mathrm{P_D}(P_{peak}))$ *and lies above* $\mathrm{P_D}(x)$ *for all* $x \in (0,\, P_{peak})$*.*

A graphical description of the tangent point $C(\gamma)$ is presented in Fig. 2.

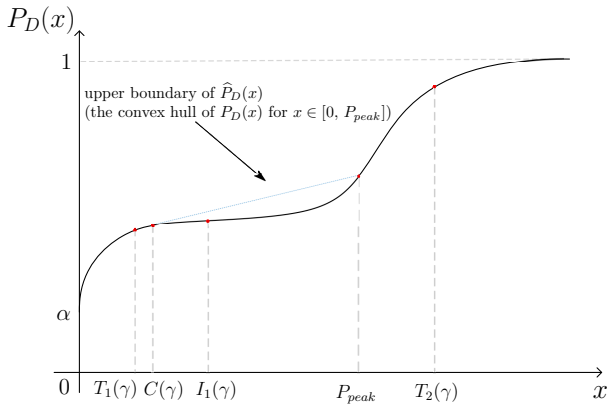


Fig. 2. $\mathrm{P_D}(x)$ and the upper boundary of the convex hull of $\mathrm{P_D}(x)$ for $x \in (0,\, P_{peak})$ for an arbitrary value of $\gamma > 0$ when $\alpha \in (0, \mathcal{Q}(2))$ and $P_{peak} \in (I_1(\gamma),\, T_2(\gamma))$. The corresponding tangent point $C(\gamma)$ is also shown on the graph.

Based on a similar argument to that presented in Lemma 4, it can be shown that $\underset{x\in[0,\, P_{peak}]}{\text{argsup}} \ \mathrm{P_D}(x) - \lambda x = \underset{x\in[0,\, P_{peak}]}{\text{argsup}} \ \widehat{P}_D(x) - \lambda x$, where $\widehat{P}_D(x)$ denotes the upper boundary of the convex hull of $\mathrm{P_D}(x)$ for $x \in [0,\, P_{peak}]$, which is obtained such that the values of $\mathrm{P_D}(x)$ for $x > P_{peak}$ are not taken into account. This observation in conjunction with Fig. 2 suggest that when $P_{peak} \in (I_1(\gamma),\, T_2(\gamma))$, the solution of the nonconvex optimization problem can be obtained via the following algorithm.

Algorithm 9. *Numerical method to compute* $P_t{}^*(\gamma)$ *in* (17) *for* $P_{peak} \in (I_1(\gamma),\, T_2(\gamma))$ *when* $\alpha \in (0, \mathcal{Q}(2))$

if $\mathrm{P_D}'(C(\gamma)) > \lambda$
$P_t{}^*(\gamma) = P_{peak}$
else
$P_t{}^*(\gamma) = \underset{x\in[0,(C(\gamma)]}{\text{argsup}} \ \mathrm{P_D}(x) - \lambda x$

Obviously, the value of $C(\gamma)$ is required to implement Algorithm 9. To that aim, Algorithm 10 provides an effective bisection search method.

Algorithm 10. *Computation of tangent point* $C(\gamma)$ *for* $I_1(\gamma) < P_{peak} < T_2(\gamma)$ *when* $\alpha \in (0, \mathcal{Q}(2))$

$\beta_{min} = \mathrm{P_D}'(I_1(\gamma))$, $\beta_{max} = \mathrm{P_D}'(T_1(\gamma))$
$x_{min} = T_1(\gamma)$, $x_{max} = I_1(\gamma)$
do
$\beta = \frac{\beta_{min}+\beta_{max}}{2}$
$x = \underset{x\in(x_{min},x_{max})}{\text{argsup}} \ \mathrm{P_D}(x) - \beta x$
if $\mathrm{P_D}(x) + \mathrm{P_D}'(x)(P_{peak} - x) > \mathrm{P_D}(P_{peak})$
then $\beta_{max} = \beta$, $x_{min} = x$
else $\beta_{min} = \beta$, $x_{max} = x$
while $|\beta_{max} - \beta_{min}| > \epsilon$

At convergence, the tangent point and the slope of the tangent line that constitutes the upper boundary of the convex hull of $\mathrm{P_D}(x)$ for $x \in [0,\, P_{peak}]$ can be obtained as $C(\gamma) \approx x$ and $\mathrm{P_D}'(C(\gamma)) \approx \beta$. Although $C(\gamma)$ must be computed for all $\gamma \in \Gamma$ separately, it does not depend on the value of λ employed in (17). Consequently, it can be computed offline together with $I_1(\gamma)$, $I_2(\gamma)$, $T_1(\gamma)$, and $T_2(\gamma)$ prior to the start of the power adaptation algorithm.

At this point, it should be emphasized that all the subroutines that are proposed to obtain the solution of the optimization problem in (17) under different cases involve convex problems. Furthermore, the bisection search described in Algorithm 4 at the beginning of Sec. III-F can be employed to solve all the problems that are of the general form $P_t{}^*(\gamma) = \underset{x\in[x_{min},\, x_{max}]}{\text{argsup}} \ \mathrm{P_D}(x) - \lambda x$ (as seen in Algorithms 5-10) due to the fact that the interval $[x_{min},\, x_{max}]$ is so arranged that $\mathrm{P_D}(x)$ is concave over the specified interval.

G. Implementation and Complexity

In this section, the implementation of the proposed power allocation approach is discussed. Based on the dual decomposition method, the optimal power allocation strategy can be obtained by solving the optimization problem in (16) for each given value of γ. Since the optimal value of λ is not known at the beginning of the iterations, λ is initialized to a certain value and updated at each iteration based on Algorithm 1 and Algorithm 2. In order to calculate the optimal power level for a given γ value and a fixed λ value, the subroutines are provided in Section III-F. For different values of the false alarm probability (α), the following statements specify the algorithms that can be used to calculate the solution of (16):

1) If $\alpha \geq \mathcal{Q}(2)$, the optimization problem in (16) becomes convex and Algorithm 4 addresses the problem.
2) If $\alpha \in (0, \mathcal{Q}(2))$ and there is no peak power constraint, then Algorithm 6 can be used.
3) When there is a peak power constraint for $\alpha \in (0, \mathcal{Q}(2))$, the optimization problem in (16) can be solved by using one of the algorithms described in Algorithm 7, Algorithm 8, and Algorithm 9.

For complexity comparisons, suppose that there exist finitely many possible values of γ, and let N_γ denote the number of different γ values. Also, consider the subroutines (i.e., Algorithms 6, 7, 8, and 9), each of which solves a 1-dimensional convex optimization problem, and assume that each of those algorithms has a computational complexity of $\mathcal{O}(C_{1D})$, where $\mathcal{O}(C_{1D})$ denotes the computational complexity of a 1-dimensional convex optimization problem with bound constraints. The main algorithm, Algorithm 1 or Algorithm 2, in Section III-E checks the convergence of the λ value and decides whether the optimal power allocation strategy obtained by the subroutines for a fixed λ value satisfy the average power constraint. In those algorithms, the corresponding optimal power levels for all γ values are calculated in each iteration. For that reason, in each iteration, the main algorithm calls a total of N_γ subroutines in order to calculate the optimal power levels for all γ values. In the context of the convergence of λ, the subgradient method in Algorithm 1 requires $\mathcal{O}(1/\epsilon^2)$ iterations in order to achieve a given tolerance level of ϵ, whereas the bisection method employed in Algorithm 2 requires $\mathcal{O}(\log((\lambda_{max} - \lambda_{min})/\epsilon^2))$ iterations, where $\lambda_{min} = 0$ and λ_{max} is a parameter used in Algorithm 2 that can be obtained by employing Algorithm 3. As a result, if Algorithm 1 is employed to obtain the optimal power allocation strategy, the overall complexity of the proposed solution is in the order of $\mathcal{O}(N_\gamma \times 1/\epsilon^2) \times \mathcal{O}(C_{1D})$. Otherwise, if Algorithm 2 is used to find the optimal strategy, the overall complexity is $\mathcal{O}(N_\gamma \times \log((\lambda_{max} - \lambda_{min})/\epsilon^2)) \times \mathcal{O}(C_{1D})$.

H. Characterization of Optimal Power Allocation

In this section, the properties of the optimal power allocation strategy are analyzed. To that aim, first, it can be shown that the average power constraint must hold with equality for the solution of (8) since any power allocation policy with $\int_\Gamma P_t(\gamma)p(\gamma)\mathrm{d}\gamma < \overline{P}$ cannot be optimal as it can be improved by allocating higher power levels for some values of γ such that $\int_\Gamma P_t(\gamma)p(\gamma)\mathrm{d}\gamma = \overline{P}$ (due to the monotone increasing nature of the probability of detection). Then, the Karush-Kuhn-Tucker (KKT) conditions [24] can be stated for the optimization problem in (8) as follows:

$$\frac{\partial\big(\mathcal{Q}\big(\mathcal{Q}^{-1}(\alpha) - \sqrt{P_t(\gamma)\gamma}\big)\big)}{\partial\left(P_t(\gamma)\right)} p(\gamma) - \lambda p(\gamma) + \mu(\gamma) - \nu(\gamma) = 0 \tag{18}$$

$$\int_{\gamma\in\Gamma} P_t(\gamma)p(\gamma)\,\mathrm{d}\gamma = \overline{P},\ P_t(\gamma) \geq 0,\ \gamma \in \Gamma \tag{19}$$

$$\mu(\gamma) \geq 0,\ \nu(\gamma) \geq 0,\ \gamma \in \Gamma \tag{20}$$

$$\mu(\gamma)P_t(\gamma) = 0,\ \gamma \in \Gamma \tag{21}$$

$$\nu(\gamma)\left(P_{peak} - P_t(\gamma)\right) = 0,\ \gamma \in \Gamma \tag{22}$$

where λ, $\mu(\gamma)$, and $\nu(\gamma)$ are the KKT multipliers. The stationarity condition in (18) can also be written as

$$\frac{\sqrt{\gamma}}{2\sqrt{2\pi}\sqrt{P_t(\gamma)}}\, e^{-\frac{\left(\mathcal{Q}^{-1}(\alpha) - \sqrt{P_t(\gamma)\gamma}\right)^2}{2}} = \lambda + \frac{\nu(\gamma) - \mu(\gamma)}{p(\gamma)}\,. \tag{23}$$

From (19)–(22), one of the following cases must be satisfied for each γ: (i) $P_t(\gamma) = 0$, $\mu(\gamma) \geq 0$, and $\nu(\gamma) = 0$, or (ii) $0 < P_t(\gamma) < P_{peak}$, $\mu(\gamma) = 0$, and $\nu(\gamma) = 0$, or (iii) $P_t(\gamma) = P_{peak}$, $\mu(\gamma) = 0$, and $\nu(\gamma) \geq 0$. In Case (i), the left-hand-side (LHS) of (23) becomes infinity for any $\gamma > 0$; hence, λ must be infinity in that case since $\mu(\gamma) \geq 0$ and $p(\gamma) > 0\ \forall\gamma \in \Gamma$ and $\mu(\gamma) = 0$ for all γ such that $P_t(\gamma) = 0$. On the other hand, in Case (ii), the LHS of (23) is finite for any γ and it must be equal to λ since $\mu(\gamma) = 0$ and $\nu(\gamma) = 0$. Therefore, if Case (i) holds for any $\gamma > 0$ (meaning that λ becomes infinity), then Case (ii) cannot hold for any value of $\gamma > 0$, leading to the violation of the average power constraint in (19). Hence, Case (i) cannot hold for any $\gamma > 0$; that is, $P_t(\gamma) > 0$ must be satisfied for $\gamma > 0$. (Since $p(\gamma)$ is a continuous random variable, this implies that for an optimal power allocation policy, $P_t(\gamma) > 0$ almost surely.) For the case with $P_t(\gamma) = P_{peak}$ for some $\gamma \in \Gamma$ (i.e., Case (iii)), the statement $\mathrm{P_D}'(P_{peak}, \gamma) \geq \lambda$ is satisfied since $\mu(\gamma) = 0$ and $\nu(\gamma) \geq 0$ for that $\gamma \in \Gamma$. Based on these cases, the solution of (8) must satisfy

$$\mathrm{P_D}'(P_t(\gamma), \gamma) = \tag{24}$$
$$\begin{cases} \frac{\sqrt{\gamma}}{2\sqrt{2\pi}\sqrt{P_t(\gamma)}}\, e^{-\frac{\left(\mathcal{Q}^{-1}(\alpha) - \sqrt{P_t(\gamma)\gamma}\right)^2}{2}} = \lambda^*, \text{ if } 0 < P_t(\gamma) < P_{peak} \\ \frac{\sqrt{\gamma}}{2\sqrt{2\pi}\sqrt{P_{peak}}}\, e^{-\frac{\left(\mathcal{Q}^{-1}(\alpha) - \sqrt{P_{peak}\gamma}\right)^2}{2}} \geq \lambda^*, \text{ if } P_t(\gamma) = P_{peak} \end{cases}$$

and $\int_\Gamma P_t(\gamma)p(\gamma)\,\mathrm{d}\gamma = \overline{P}$ (cf. (6) and (9)).[5]

The following lemma specifies γ values for which the optimal power level is equal to P_{peak}; that is, $P_t^*(\gamma) = P_{peak}$.

Lemma 6. *For $\mathcal{Q}(2) < \alpha < 1$, if $\mathrm{P_D}'(P_{peak}, \gamma) \geq \lambda^*$ for some $\gamma \in \Gamma$, then $P_t^*(\gamma) = P_{peak}$ for those values of γ.*

Proof: Consider that $\mathcal{Q}(2) < \alpha < 1$ and $\mathrm{P_D}'(P_{peak}, \gamma) \geq \lambda^*$ for some $\gamma \in \Gamma$. Then, suppose that $P_t^*(\gamma) \neq P_{peak}$ for those values of $\gamma \in \Gamma$; that is, $0 < P_t^*(\gamma) < P_{peak}$. Since $\mathrm{P_D}'(P_t(\gamma), \gamma)$ is monotone decreasing for $P_t(\gamma) \in (0, P_{peak})$ in the case of $\alpha \in (\mathcal{Q}(2), 1)$, $\mathrm{P_D}'(P_t(\gamma), \gamma)$ satisfies for all $P_t(\gamma) \in (0, P_{peak})$ that $\mathrm{P_D}'(P_t(\gamma), \gamma) > \mathrm{P_D}'(P_{peak}, \gamma) \geq \lambda^*$. However, $\mathrm{P_D}'(P_t^*(\gamma), \gamma) = \lambda^*$ for $0 < P_t^*(\gamma) < P_{peak}$ based on (24), which contradicts with the inequality that $\mathrm{P_D}'(P_t^*(\gamma), \gamma) > \lambda^*$. Therefore, $P_t^*(\gamma) = P_{peak}$ if there exist $\gamma \in \Gamma$ which satisfy $\mathrm{P_D}'(P_{peak}, \gamma) \geq \lambda^*$. ∎

Based on Lemma 6 and the expression in (24), it can be stated for $\mathcal{Q}(2) < \alpha < 1$ that the optimal power level is $P_t^*(\gamma) = P_{peak}$ if and only if there exists a γ such that $\mathrm{P_D}'(P_{peak}, \gamma) \geq \lambda^*$.

To provide a further analysis, the expression in (24) can also be motivated based on dual decomposition. As discussed in Sec. III-D, the optimal power allocation policy can be determined by choosing the optimum power $P_t(\gamma)$ for each value of $\gamma \in \Gamma$ based on the dual decomposition approach. Let the minimizer λ of the dual problem in (14) be denoted by

[5]From (19)–(23), it can be shown that $P_t(\gamma) = 0$ for $\gamma = 0$ in the optimal solution. In addition, via (19) and (24), $P_t(\gamma) \to 0$ as $\gamma \to 0$, implying that the optimal power allocation policy is continuous at $\gamma = 0$.

λ^*. Then, from the dual decomposition approach, the optimal power allocation is specified as

$$P_t^*(\gamma) = \underset{0\leq P_t(\gamma)\leq P_{peak}}{\text{argsup}} \; \mathrm{P_D}(P_t(\gamma),\gamma) - \lambda^* P_t(\gamma) \qquad (25)$$

for any given value of $\gamma \in \Gamma$ (cf. (5) and (16)). This implies that the optimum power $P_t^*(\gamma)$ must satisfy (24), that is, $\mathrm{P_D}'(P_t^*(\gamma),\gamma) = \lambda^*$ if $0 < P_t^*(\gamma) < P_{peak}$ and $\mathrm{P_D}'(P_t^*(\gamma),\gamma) \geq \lambda^*$ if $P_t^*(\gamma) = P_{peak}$. Recall that, by Lemma 1, for $\alpha \in (0, \mathcal{Q}(2))$, $\mathrm{P_D}'(P_t(\gamma),\gamma)$ is monotone decreasing for $P_t(\gamma) \in (0, I_1(\gamma))$, monotone increasing for $P_t(\gamma) \in (I_1(\gamma), I_2(\gamma))$, and monotone decreasing for $P_t(\gamma) \in (I_2(\gamma), \infty)$ for any given value of $\gamma \in \Gamma$, where $I_1(\gamma)$ and $I_2(\gamma)$ are the two inflection points of $\mathrm{P_D}(P_t(\gamma),\gamma)$ with $I_1(\gamma) < I_2(\gamma)$ (see (11)). Thus, if $\lambda^* > \mathrm{P_D}'(I_2(\gamma),\gamma)$ or $\lambda^* < \mathrm{P_D}'(I_1(\gamma),\gamma)$, then $\mathrm{P_D}'(P_t^*(\gamma),\gamma) = \lambda^*$ has a unique solution $P_t^\star(\gamma)$; otherwise, there exist three (or, two) candidates for the optimal power level. From (6), (9), and (11), it can be shown that the inflection points $I_1(\gamma)$ and $I_2(\gamma)$ decrease as γ increases; however, the value of $\mathrm{P_D}'$ at the inflection points increases with γ. Let γ_l and γ_u be defined such that $\lambda^* = \mathrm{P_D}'(I_2(\gamma_l),\gamma_l)$ and $\lambda^* = \mathrm{P_D}'(I_1(\gamma_u),\gamma_u)$, respectively. From (9) and (11), $\lambda^* = \mathrm{P_D}'(I_2(\gamma_l),\gamma_l) = \frac{\gamma_l}{\sqrt{2\pi}\left(\mathcal{Q}^{-1}(\alpha)+\sqrt{(\mathcal{Q}^{-1}(\alpha))^2-4}\right)} \exp\left\{ -\frac{1}{2}\left(\frac{\mathcal{Q}^{-1}(\alpha)}{2} - \frac{\sqrt{(\mathcal{Q}^{-1}(\alpha))^2-4}}{2}\right)^2\right\}$ is obtained, which results in $\gamma_l = \lambda^*\sqrt{2\pi}\left(\mathcal{Q}^{-1}(\alpha)+\sqrt{(\mathcal{Q}^{-1}(\alpha))^2-4}\right)\exp\left\{\frac{1}{2}\left(\frac{\mathcal{Q}^{-1}(\alpha)}{2} - \frac{\sqrt{(\mathcal{Q}^{-1}(\alpha))^2-4}}{2}\right)^2\right\}$. Similarly, $\lambda^* = \mathrm{P_D}'(I_1(\gamma_u),\gamma_u) = \frac{\gamma_u}{\sqrt{2\pi}\left(\mathcal{Q}^{-1}(\alpha)-\sqrt{(\mathcal{Q}^{-1}(\alpha))^2-4}\right)} \exp\left\{ -\frac{1}{2}\left(\frac{\mathcal{Q}^{-1}(\alpha)}{2} + \frac{\sqrt{(\mathcal{Q}^{-1}(\alpha))^2-4}}{2}\right)^2\right\}$ implies that $\gamma_u = \lambda^*\sqrt{2\pi}\left(\mathcal{Q}^{-1}(\alpha) - \sqrt{(\mathcal{Q}^{-1}(\alpha))^2-4}\right)\exp\left\{\frac{1}{2}\left(\frac{\mathcal{Q}^{-1}(\alpha)}{2} + \frac{\sqrt{(\mathcal{Q}^{-1}(\alpha))^2-4}}{2}\right)^2\right\}$. Hence, $\lambda^* > \mathrm{P_D}'(I_2(\gamma),\gamma)$ for every $\gamma < \gamma_l$ and $\lambda^* < \mathrm{P_D}'(I_1(\gamma),\gamma)$ for every $\gamma > \gamma_u$, which imply that $\mathrm{P_D}'(P_t(\gamma),\gamma) = \lambda^*$ has a unique solution $P_t^\star(\gamma)$, and consequently, $P_t^*(\gamma) = \min\{P_t^\star(\gamma), P_{peak}\}$. Therefore, it is concluded that the optimal power allocation policy is a continuous function of γ for $\gamma < \gamma_l$ and for $\gamma > \gamma_u$. However, the behavior of the the optimal power allocation for values of γ between γ_l and γ_u depends on the false alarm level, α, as specified in the following theorems.

Theorem 1. *Let $\mathcal{Q}(2) < \alpha < 1$. Then, the optimal power level according to* (8) *is a continuous function of γ, which satisfies one of the following conditions:*

- *(i) It increases with γ up to some unique value and then decreases as γ increases.*
- *(ii) It increases up to P_{peak} as γ goes to $\bar{\gamma}_l > 0$, stays at P_{peak} for a certain interval of $\gamma \in [\bar{\gamma}_l, \bar{\gamma}_u]$, and then decreases as $\gamma > \bar{\gamma}_u$ increases.*

Proof: Please see Appendix A. ∎

From Theorem 1 and Footnote 3, it is concluded for $\alpha > \mathcal{Q}(2)$ that only two possible scenarios exist for the optimal power allocation policy. In the first scenario, the optimal power level starts from zero at $\gamma = 0$ and increases monotonically with γ up to a unique value, after which it decreases monotonically. In the second one, the optimal power level starts from zero at $\gamma = 0$ and increases monotonically with γ up to P_{peak}, and then remains at P_{peak} for a certain interval after which it decreases monotonically for higher values of γ. Based on these scenarios, the characterization of the optimal power allocation policy in Theorem 1 can be interpreted as follows: For low values of γ (i.e., for unfavorable channel conditions), the transmitter employs low power levels for the transmitted signal, and it increases the power level as γ increases. However, after a certain value of γ, it becomes more preferable to transmit with lower power levels since high detection probabilities can already be achieved with lower power levels (as the channel condition is very favorable), which leads to savings in the average transmit power.

Theorem 2. *Let $0 < \alpha < \mathcal{Q}(2)$. Then, the optimal power allocation policy is continuous everywhere except at one point, and there exists a positive jump at the discontinuity point. Further, in the absence of the peak power constraint, the optimal power level can never take values between $I_1(\gamma)$ and $I_2(\gamma)$; i.e., either $P_t^*(\gamma) < I_1(\gamma)$ or $P_t^*(\gamma) > I_2(\gamma)$.*

Proof: Please see Appendix B. ∎

Theorem 2 specifies the discontinuous nature of the optimal power allocation strategy for low false alarm levels, i.e., for $\alpha < \mathcal{Q}(2)$. The statements in Theorems 1 and 2 are investigated via numerical examples in Sec. IV.

IV. Numerical Examples

In this section, the proposed optimal power allocation strategy for the maximization of the average detection probability is investigated via numerical examples. In the examples, both exponential distribution (corresponding to Rayleigh fading channels) and uniform distribution are considered for parameter γ in (8), which is defined as $\gamma = h^2/\sigma^2$. For comparison purposes, the results for the uniform power allocation strategy are also presented. In the simulations, the average power limit is taken as one; i.e., $\overline{P} = 1$ in (8), and the peak power limit in (8) is set to $P_{peak} = 20$. For the maximization of the average detection probability according to the proposed approach, the solution of (16) is obtained for a given λ for every α and γ; then, the bisection-based update method is used to obtain the optimal λ and the corresponding power allocation strategy.

In Fig. 3, the average detection probabilities of the proposed optimal power allocation strategy and the uniform power allocation strategy are plotted versus the probability of false alarm, α, for exponentially distributed γ, where the average values of γ are specified by $\bar{\gamma} = 1$ and $\bar{\gamma} = 2$. In addition, Fig. 4 illustrates the region of low false alarm rates in more detail by zooming into Fig. 3 for $\alpha \in [0, 0.1]$. From the figures, it is observed that the proposed power allocation strategy achieves higher detection probabilities than the uniform power allocation strategy for all values of the probability of false alarm, which indicates that employing a constant power level is not an optimal strategy for the considered problem. In particular, significant gains are achieved in the average detection probability for small values of α in this example (see Fig. 4). In addition, as expected, improved detection performance is achieved as the mean of γ increases as it leads to a more favorable distribution for the SNR.

Next, uniform distribution is employed for γ, and the average detection probabilities of the proposed optimal power allocation strategy and the uniform power allocation strategy are plotted versus the probability of false alarm in Fig. 5, where the intervals $[0, 2]$ and $[0, 4]$ are considered for the uniform distribution. Also, Fig. 6 zooms into Fig. 5 for $\alpha \in [0, 0.1]$. As in the exponentially distributed case, the proposed power allocation strategy leads to higher detection

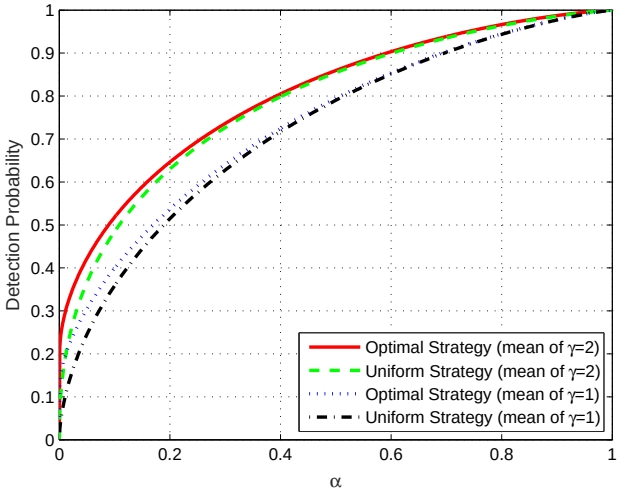


Fig. 3. Average detection probability versus the probability of false alarm for the proposed optimal power allocation strategy and the uniform power allocation strategy, where γ is exponentially distributed with mean 1 or 2.

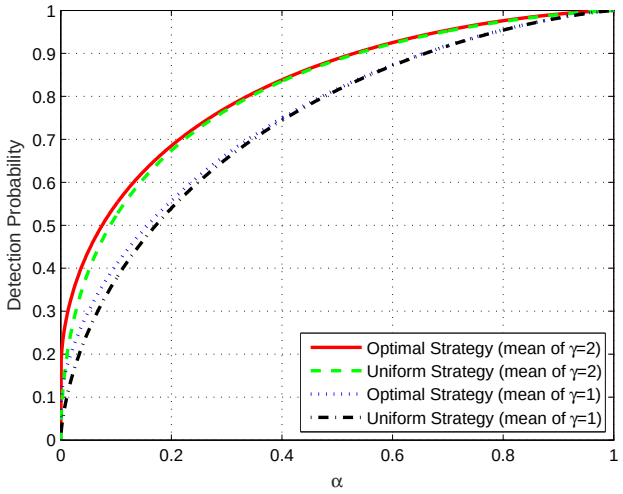


Fig. 5. Average detection probability versus the probability of false alarm for the proposed optimal power allocation strategy and the uniform power allocation strategy, where γ is uniformly distributed over $[0, 2]$ or $[0, 4]$.

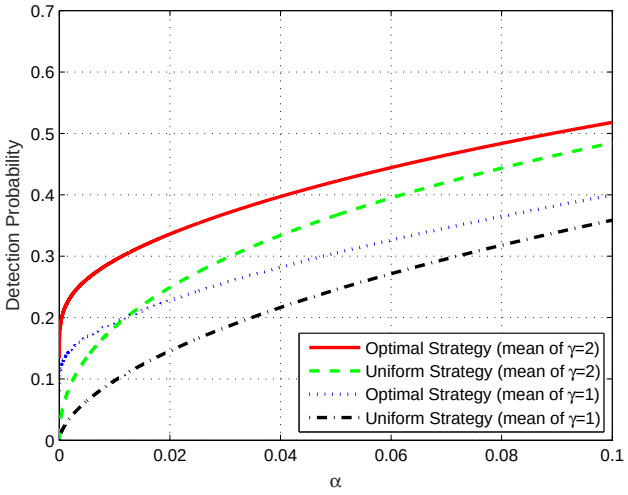


Fig. 4. The zoomed version of Fig. 3 for $\alpha \in [0, 0.1]$.

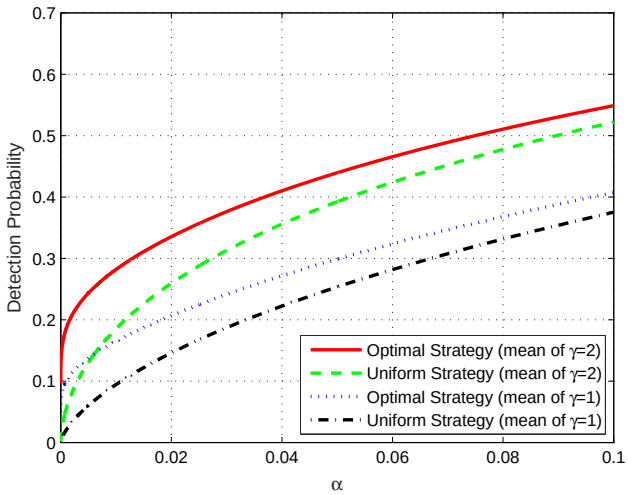


Fig. 6. The zoomed version of Fig. 5 for $\alpha \in [0, 0.1]$.

probabilities than the uniform power allocation strategy, as expected. In addition, higher detection probabilities are observed when γ is distributed between 0 and 4.

To illustrate the results in Section III-H, the transmitted power levels are plotted versus γ for the proposed optimal power allocation strategy in Fig. 7, where γ is exponentially distributed with a mean of 1, α is set to 0.001, 0.01, 0.03, and 0.1, and the peak power limit is given by $P_{peak} = 3$. Also, the transmitted power according to the uniform power allocation strategy is shown in the figure for comparison purposes. (In addition, Fig. 8 zooms into Fig. 7 for $\gamma \in [0, 20]$.) In accordance with Theorem 1, the optimal transmitted power is a continuous function of γ for $\alpha = 0.1$ and $\alpha = 0.03$ in Fig. 7, where $\alpha > \mathcal{Q}(2)$. In addition, the optimal power allocation policy for $\alpha = 0.1$ satisfies condition (i) in Theorem 1, which first increases up to a unique value of γ (namely, $\gamma = 1.636$) and then decreases monotonically. For $\alpha = 0.03$, condition (ii) in Theorem 1 holds, which states that the optimal power level increases as γ increases to $\bar{\gamma}_l = 1.61$, stays at $P_{peak} = 3$ for $\bar{\gamma}_l \leq \gamma \leq \bar{\gamma}_u$ where $\bar{\gamma}_u = 1.97$, and then decreases for $\gamma > \bar{\gamma}_u$. On the other hand, for $\alpha = 0.01$ and $\alpha = 0.001$, the condition in Theorem 2 is satisfied; i.e., $\alpha < \mathcal{Q}(2)$, and discontinuities are observed in the optimal transmitted power curves. In particular, the transmitted power level is continuous before and after a certain value of γ, and there exists one positive jump in the optimal power level, which are in compliance with Theorem 2. To specify the application of Theorem 2 in more detail, $\alpha = 0.001$ is considered as an example, for which parameters γ_l and γ_u are obtained as $\gamma_l = 0.488$ and $\gamma_u = 2.508$. As stated in the theorem, the optimal power allocation policy for $\alpha = 0.001$ is continuous for $\gamma \leq \gamma_l = 0.488$ and $\gamma \geq \gamma_u = 2.508$, and there exists a positive jump for $\gamma_l < \gamma < \gamma_u$, which is at $\gamma = 1.11$. Another observation from Fig. 7 is that as α decreases, the optimal transmission strategy becomes more peaky in order to satisfy the false alarm constraint while maximizing the average probability of detection. Regarding the uniform power allocation policy, it is noted that the employed power allocation strategy is significantly different from the optimal one.

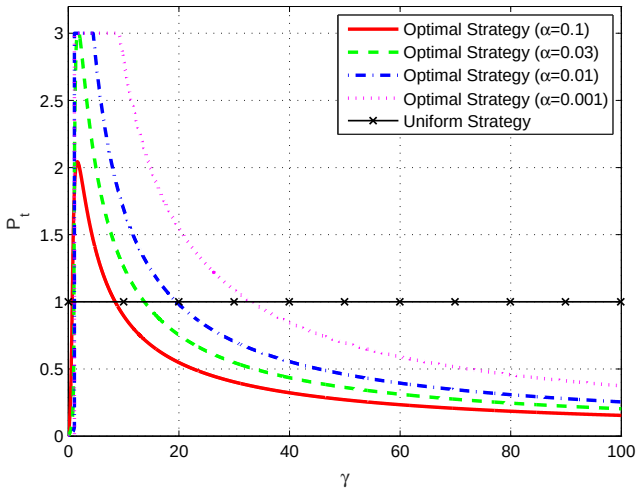


Fig. 7. The transmitted power level versus γ for the proposed optimal power allocation strategy and the uniform power allocation strategy, where γ is exponentially distributed with mean 1.

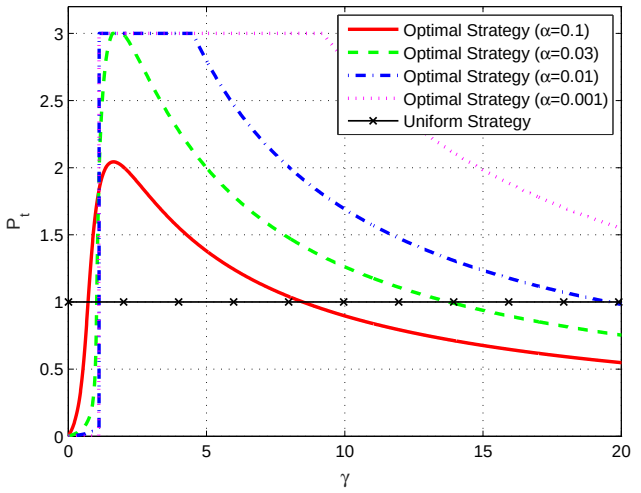


Fig. 8. The zoomed version of Fig. 7 for $\gamma \in [0, 20]$.

In Fig. 9, the transmitted power levels are plotted versus γ for the proposed optimal power allocation strategy and the uniform power allocation strategy, where γ is uniformly distributed between 0 and 2 and the peak power limit is set to $P_{peak} = 5$. Similar to the previous scenario, the statements in Theorem 1 and Theorem 2 can be verified based on the transmitted power levels of the optimal power allocation strategy for various values of α. For example, for $\alpha = 0.1$, the optimal power level increases until $\gamma = 1.917$ and decreases after that value in accordance with Theorem 1. In addition, as the false alarm limit decreases, the transmitter employs higher power levels for some values of γ while sending very low powers at other values, leading to a more peaky transmission strategy as in the previous scenario.

Finally, the concavity property of the optimal average detection probability with respect to the average power limit, $\overline{P}$, is illustrated in Fig. 10, where both uniform distribution (between 0 and 2) and exponential distribution (with a mean of 1) are considered. As stated in Proposition 1, the average probability of detection corresponding to the solution of (8) is

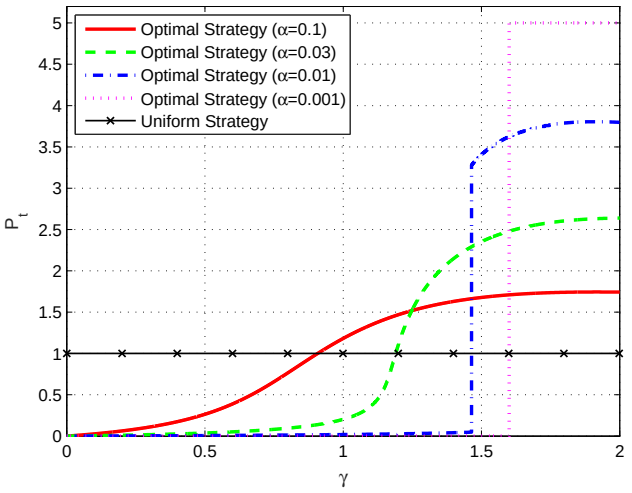


Fig. 9. The transmitted power level versus γ for the proposed optimal power allocation strategy and the uniform power allocation strategy, where γ is uniformly distributed between 0 and 2.

a concave function of the average power limit for any value of α.

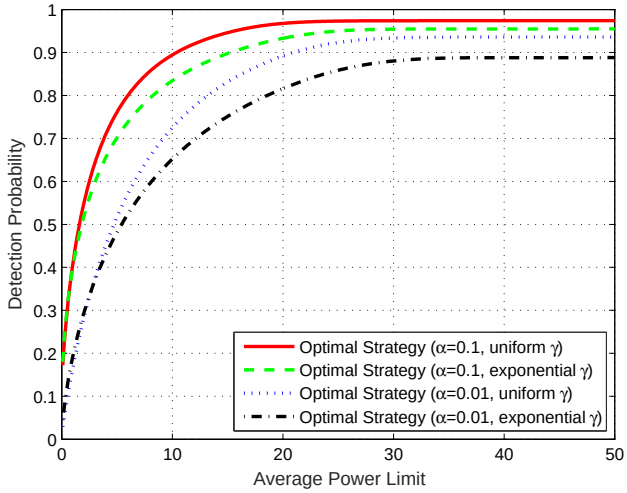


Fig. 10. The optimal average detection probability versus the average power limit, $\overline{P}$.

V. Conclusions and Extensions

In this study, the optimal power allocation problem has been proposed to maximize the average detection probability for detecting the presence of a signal in an AWGN channel with flat fading. An optimization problem has been formulated under average and peak power constraints when perfect CSI is available at the transmitter and the receiver. Utilizing the analytical properties of the detection probability, a dual problem with no duality gap with the original problem has been obtained. The dual decomposition approach has been employed and various algorithms and subroutines have been proposed to specify the optimal power allocation scheme under average and peak power constraints. In addition, for all values of the false alarm probability, the continuity and monotonicity properties of the optimal power allocation scheme have

been characterized with respect to γ, the ratio between the channel power gain and the noise power. Numerical examples have provided some examples of the theoretical results and illustrated the improvements achieved via the optimal power allocation approach.

Although scalar observations are considered in (1), the results can also be extended to vector observations in the presence of AWGN since the detection probability can be expressed similarly to (6) by updating the definition of γ.

APPENDIX

A. Proof of Theorem 1

The proof consists of two parts. In the first part of the proof, the aim is to prove that if ${\rm P_D}'(P_{peak},\gamma) < \lambda^*$ for all $\gamma \in \Gamma$, then the optimal power allocation policy is a continuous function of γ, which increases with γ up to some unique value and then decreases as γ increases. Since ${\rm P_D}'(P_{peak},\gamma) < \lambda^*$ for all $\gamma \in \Gamma$, $P_t^*(\gamma) \neq P_{peak}$ for all $\gamma \in \Gamma$ based on (24); that is, $P_t^*(\gamma)$ satisfies $0 < P_t^*(\gamma) < P_{peak}$ for all $\gamma \in \Gamma$. First, the limiting cases of the equation in (24) are investigated for $0 < P_t^*(\gamma) < P_{peak}$. Namely, it is observed that as $\gamma P_t(\gamma)$ goes to zero, $P_t(\gamma)/\gamma$ converges to a constant. Similarly, as $\gamma P_t(\gamma)$ goes to infinity, $P_t(\gamma)/\gamma$ converges to zero. Let $x \triangleq \sqrt{\gamma}$, $y \triangleq \sqrt{P_t(\gamma)}$, and $\mathcal{G}(x,y) \triangleq \frac{x}{2\sqrt{2\pi}\lambda^*}\exp\left\{-\frac{1}{2}\left(\mathcal{Q}^{-1}(\alpha)-xy\right)^2\right\}$. Then, from (24), the following relation is obtained:

$$y = \mathcal{G}(x,y) = \frac{x}{2\sqrt{2\pi}\lambda^*}\exp\left\{-\frac{1}{2}\left(\mathcal{Q}^{-1}(\alpha)-xy\right)^2\right\}. \quad (26)$$

Now suppose that $\frac{{\rm d}y}{{\rm d}x}$ exists. Then, $\frac{{\rm d}y}{{\rm d}x} = \frac{\partial \mathcal{G}}{\partial y}\frac{{\rm d}y}{{\rm d}x} + \frac{\partial \mathcal{G}}{\partial x}$, which leads to $\frac{{\rm d}y}{{\rm d}x} = \frac{\partial \mathcal{G}/\partial x}{1-\partial \mathcal{G}/\partial y} \triangleq \mathcal{F}(x,y)$. These derivative expressions are calculated, from (26), as follows:

$$\frac{\partial \mathcal{G}(x,y)}{\partial x} = \frac{1}{2\sqrt{2\pi}\lambda^*}\exp\left\{-\frac{1}{2}\left(\mathcal{Q}^{-1}(\alpha)-xy\right)^2\right\}$$
$$\times\left(1+xy\big(\mathcal{Q}^{-1}(\alpha)-xy\big)\right) = \frac{y}{x}\left(1+xy\left(\mathcal{Q}^{-1}(\alpha)-xy\right)\right)$$
$$\frac{\partial \mathcal{G}(x,y)}{\partial y} = \frac{x^2}{2\sqrt{2\pi}\lambda^*}\exp\left\{-\frac{1}{2}\left(\mathcal{Q}^{-1}(\alpha)-xy\right)^2\right\}$$
$$\times\left(\mathcal{Q}^{-1}(\alpha)-xy\right) = xy\left(\mathcal{Q}^{-1}(\alpha)-xy\right)$$
$$\mathcal{F}(x,y) = \frac{\frac{y}{x}\left(1+xy\left(\mathcal{Q}^{-1}(\alpha)-xy\right)\right)}{1-xy\left(\mathcal{Q}^{-1}(\alpha)-xy\right)}$$
$$= -\frac{y}{x}\left(1+\frac{2}{xy\left(\mathcal{Q}^{-1}(\alpha)-xy\right)-1}\right)$$
$$\frac{\partial \mathcal{F}(x,y)}{\partial y} = -\frac{1}{x}\left(1+\frac{2}{xy\left(\mathcal{Q}^{-1}(\alpha)-xy\right)-1}\right)$$
$$+\frac{2y\left(\mathcal{Q}^{-1}(\alpha)x-2x^2y\right)}{x\left(xy\left(\mathcal{Q}^{-1}(\alpha)-xy\right)-1\right)^2}.$$

Note that $\mathcal{F}(x,y)$ and $\partial\mathcal{F}(x,y)/\partial y$ are continuous functions for $x > 0$ and $y > 0$ except at the points that satisfy $xy\left(\mathcal{Q}^{-1}(\alpha)-xy\right)-1=0$. Let $t \triangleq xy$. Then, the solutions of $h(t) \triangleq t^2 - \mathcal{Q}^{-1}(\alpha)t+1=0$ are sought. In Lemma 1, it is shown that if $\alpha > \mathcal{Q}(2)$, $h(t) > 0$ for $t > 0$, and if $\alpha < \mathcal{Q}(2)$, there are two roots $t_1 < t_2$ where $t_1 = \frac{\mathcal{Q}^{-1}(\alpha)-\sqrt{(\mathcal{Q}^{-1}(\alpha))^2-4}}{2}$ and $t_2 = \frac{\mathcal{Q}^{-1}(\alpha)+\sqrt{(\mathcal{Q}^{-1}(\alpha))^2-4}}{2}$ with $h(t) > 0$ for $t < t_1$ and $t > t_2$, and $h(t) < 0$ for $t_1 < t < t_2$. Thus, if $\alpha > \mathcal{Q}(2)$, $\mathcal{F}(x,y)$ and $\partial\mathcal{F}(x,y)/\partial y$ are continuous functions for $x > 0$ and $y > 0$, which implies that $\frac{{\rm d}y}{{\rm d}x} = \mathcal{F}(x,y)$ has a unique solution by the existence and uniqueness theorems for first-order ordinary differential equations (ODEs) [33]. Hence, y is differentiable in x, which implies that y is a continuous function of x, or, equivalently, $P_t(\gamma)$ is a continuous function of γ. This proves the continuity of the power allocation policy for $\alpha > \mathcal{Q}(2)$.

Since $x = \sqrt{\gamma}$ and $y = \sqrt{P_t(\gamma)}$, ${\rm d}x = {\rm d}\gamma/(2\sqrt{\gamma})$ and ${\rm d}y = {\rm d}P_t(\gamma)/(2\sqrt{P_t(\gamma)})$ are obtained. Then, from $\frac{{\rm d}y}{{\rm d}x} = -\frac{y}{x}\left(1+\frac{2}{xy(\mathcal{Q}^{-1}(\alpha)-xy)-1}\right)$ in the previous paragraph, the following relation is achieved:

$$\frac{{\rm d}P_t(\gamma)}{{\rm d}\gamma} = \frac{\sqrt{P_t(\gamma)}}{\sqrt{\gamma}}\frac{{\rm d}y}{{\rm d}x}$$
$$= -\frac{P_t(\gamma)}{\gamma}\left(1+\frac{2}{\sqrt{P_t(\gamma)\gamma}\left(\mathcal{Q}^{-1}(\alpha)-\sqrt{P_t(\gamma)\gamma}\right)-1}\right).$$

Based on the observations at the beginning of the proof, $\lim_{\gamma P_t(\gamma)\to 0} P_t(\gamma)/\gamma = c_1$, where c_1 is some positive constant and $\lim_{\gamma P_t(\gamma)\to\infty} P_t(\gamma)/\gamma = 0$. Thus, $\lim_{\gamma P_t(\gamma)\to 0}\frac{{\rm d}P_t(\gamma)}{{\rm d}\gamma} = c_2$, where c_2 is some positive constant and $\lim_{\gamma P_t(\gamma)\to\infty}\frac{{\rm d}P_t(\gamma)}{{\rm d}\gamma} = 0^-$. From the analysis of the limit cases above, the optimal power linearly increases with respect to γ for small values of γ and decreases with respect to γ for large values of γ.

Up to this point, it is shown that the optimal power increases linearly for small γ, and decreases for large γ. Hence, there must be a maximum power value, which can be found by setting $\frac{\partial P_t(\gamma)}{\partial\gamma}$ to zero.

$$\frac{\partial P_t(\gamma)}{\partial\gamma} = 0 \Rightarrow 1+\frac{2}{\left(\mathcal{Q}^{-1}(\alpha)-\sqrt{\gamma P_t(\gamma)}\right)\sqrt{\gamma P_t(\gamma)}-1} = 0$$
$$\Rightarrow \gamma P_t(\gamma) - \mathcal{Q}^{-1}(\alpha)\sqrt{\gamma P_t(\gamma)} - 1 = 0$$
$$\Rightarrow \sqrt{\gamma P_t(\gamma)} = \frac{\mathcal{Q}^{-1}(\alpha)+\sqrt{(\mathcal{Q}^{-1}(\alpha))^2+4}}{2}. \quad (27)$$

From (24) and (27), the $(\gamma, P_t(\gamma))$ pair that specifies the power level and γ for which the maximum power is employed can be obtained uniquely as follows:

$$\gamma = \lambda^*\sqrt{2\pi}\left(\mathcal{Q}^{-1}(\alpha)+\sqrt{(\mathcal{Q}^{-1}(\alpha))^2+4}\right) \quad (28)$$
$$\times\exp\left\{\frac{1}{8}\left(\mathcal{Q}^{-1}(\alpha)-\sqrt{(\mathcal{Q}^{-1}(\alpha))^2+4}\right)^2\right\}$$
$$P_t(\gamma) = \frac{1}{4\lambda^*\sqrt{2\pi}}\left(\mathcal{Q}^{-1}(\alpha)+\sqrt{(\mathcal{Q}^{-1}(\alpha))^2+4}\right)$$
$$\times\exp\left\{-\frac{1}{8}\left(\mathcal{Q}^{-1}(\alpha)-\sqrt{(\mathcal{Q}^{-1}(\alpha))^2+4}\right)^2\right\}.$$

In the second part, the aim is to prove that if ${\rm P_D}'(P_{peak},\gamma) \geq \lambda^*$ for some $\gamma \in \Gamma$, the optimal power allocation policy is a continuous function of γ, which increases up to P_{peak} as γ increases to $\bar{\gamma}_l$, stays at P_{peak} for a certain interval (i.e., $\bar{\gamma}_l \leq \gamma \leq \bar{\gamma}_u$), and then decreases as $\gamma > \bar{\gamma}_u$ increases. To that aim, consider ${\rm P_D}'(P_{peak},\gamma)$ for all $\gamma \in \Gamma$ and note that ${\rm P_D}'(P_{peak},\gamma)$ is a continuous function of γ, which increases up to a certain value and decreases for higher values of γ. Then, it is stated that if ${\rm P_D}'(P_{peak},\gamma) \geq \lambda^*$ for some $\gamma \in \Gamma$,

then there exists only a unique interval, $\gamma \in [\bar{\gamma}_l, \bar{\gamma}_u]$, such that $\mathrm{P_D}'(P_{peak}, \gamma) \geq \lambda^*$ for all $\gamma \in \bar{\Gamma}$ where $\bar{\Gamma} = [\bar{\gamma}_l, \bar{\gamma}_u]$ and $\bar{\Gamma} \subseteq \Gamma$. Based on the statement in Lemma 6, it is obtained that $P_t^*(\gamma) = P_{peak}$ for all $\gamma \in \bar{\Gamma}$. Next, consider $\gamma \in \Gamma \setminus \bar{\Gamma}$ and note that $\mathrm{P_D}'(P_{peak}, \gamma) = \lambda^*$ for $\gamma = \bar{\gamma}_l$ and $\gamma = \bar{\gamma}_u$, and consequently, $P_t^*(\bar{\gamma}_l) = P_{peak}$ and $P_t^*(\bar{\gamma}_u) = P_{peak}$. Then, from (24) and Lemma 6, $\mathrm{P_D}'(P_t^*(\gamma), \gamma) = \lambda^*$ for $\gamma \in (0, \bar{\gamma}_l]$ or $\gamma \in [\bar{\gamma}_u, \infty)$ where $P_t^*(\gamma) \in (0, P_{peak}]$. Based on a similar approach to that in the first part of the proof, $P_t^*(\gamma)$ is a continuous function of γ for $\gamma \in (0, \bar{\gamma}_l]$, which increases with γ up to $P_t^*(\gamma) = P_{peak}$ for $\gamma = \bar{\gamma}_l$. Similarly, $P_t^*(\gamma)$ is a continuous function of γ for $\gamma \in [\bar{\gamma}_u, \infty)$, which reduces as γ increases. Overall, $P_t^*(\gamma)$ is a continuous function of γ where $\gamma \in \Gamma$ and satisfies the second condition in Theorem 1 if $\mathrm{P_D}'(P_{peak}, \gamma) \geq \lambda^*$ for some $\gamma \in \Gamma$. ■

B. Proof of Theorem 2

Consider the relation in (24), which must be satisfied for the optimum power levels. Since the inflection points $I_1(\gamma)$ and $I_2(\gamma)$ decrease as γ increases, and the value of $\mathrm{P_D}'(P_t(\gamma), \gamma)$ at the inflection points increases as γ increases, there exist 5 different cases with respect to λ^* and the value of $\mathrm{P_D}'(P_t(\gamma), \gamma)$ at the inflection points for a given γ. These cases occur in the order of *1-2-3-4-5* as γ increases.

1) $\lambda^* > \mathrm{P_D}'(I_2(\gamma), \gamma)$: This case is valid for $\gamma \in (0, \gamma_l)$. In this case, all $P_t(\gamma)$ values that satisfy $\mathrm{P_D}'(P_t(\gamma), \gamma) \geq \lambda^*$ cannot exceed $I_1(\gamma)$; that is, $P_t(\gamma) < I_1(\gamma)$. Therefore, $P_t^*(\gamma) < I_1(\gamma)$ for all $\gamma \in (0, \gamma_l)$. Also, $\mathrm{P_D}'(P_t(\gamma), \gamma)$ is monotone decreasing for $P_t(\gamma) \in (0, I_1(\gamma))$. Then, based on the equation in (24), one of the following conditions holds:

- If $\mathrm{P_D}'(P_{peak}, \gamma) < \lambda^*$ for all $\gamma \in (0, \gamma_l)$, $\mathrm{P_D}'(P_t^*(\gamma), \gamma) = \lambda^*$ for all $\gamma \in (0, \gamma_l)$. Also, $\mathrm{P_D}'(P_t(\gamma), \gamma) = \lambda^*$ has only one solution (optimal power level) $P_t^*(\gamma)$, where $P_t^*(\gamma) < I_1(\gamma)$. Based on the definitions in the proof of Theorem 1, if $P_t(\gamma) < I_1(\gamma)$, then $t = xy = \sqrt{\gamma P_t(\gamma)} < \sqrt{\gamma I_1(\gamma)} = \sqrt{\gamma \frac{\left(\mathcal{Q}^{-1}(\alpha) - \sqrt{(\mathcal{Q}^{-1}(\alpha))^2 - 4}\right)^2}{4\gamma}} = \frac{\mathcal{Q}^{-1}(\alpha) - \sqrt{(\mathcal{Q}^{-1}(\alpha))^2 - 4}}{2} = t_1$. We also have $h(t) > 0$ for $t < t_1$. Thus, due to a similar reasoning to that in the proof of Theorem 1, $P_t^*(\gamma)$ constitutes a continuous function of γ for $\gamma \in (0, \gamma_l)$.
- If $\mathrm{P_D}'(P_{peak}, \gamma) \geq \lambda^*$ for some $\gamma \in (0, \gamma_l)$, $\mathrm{P_D}'(P_{peak}, \gamma) \geq \lambda^*$ for $\gamma \in [\bar{\gamma}_l, \bar{\gamma}_u]$ due to the function properties of $\mathrm{P_D}'(P_{peak}, \gamma)$ where $\bar{\gamma}_l$ and $\bar{\gamma}_u$ are positive finite values. If $\bar{\gamma}_l < \bar{\gamma}_u < \gamma_l$ holds, $P_t^*(\gamma) = P_{peak}$ for $\gamma \in [\bar{\gamma}_l, \bar{\gamma}_u]$ and $P_t^*(\gamma)$ is continuous function of γ for $\gamma \in (0, \gamma_l)$, which increases for $\gamma \in (0, \bar{\gamma}_l)$ and decreases for $\gamma \in (\bar{\gamma}_u, \gamma_l)$ based on a similar approach to that in the first condition. Otherwise, if $\bar{\gamma}_l = \gamma_l \leq \bar{\gamma}_u$, $P_t^*(\gamma)$ is a continuous function of γ for $\gamma \in (0, \gamma_l)$, which increases for $\gamma \in (0, \bar{\gamma}_l)$ and becomes equal to P_{peak} for $\gamma \in [\bar{\gamma}_l, \gamma_l]$.

2) $\lambda^* = \mathrm{P_D}'(I_2(\gamma), \gamma)$: This case is valid for $\gamma = \gamma_l$. In this case, $\mathrm{P_D}'(P_t(\gamma), \gamma) = \lambda^*$ has two solutions; i.e., there are two candidates for the optimal power level. Let these candidates be $P_t^{*,1}(\gamma)$ and $P_t^{*,2}(\gamma)$ with $P_t^{*,1}(\gamma) < I_1(\gamma) < I_2(\gamma) = P_t^{*,2}(\gamma)$. Also, there is another candidate, $P_t^{*,3}(\gamma) = P_{peak}$. If $P_t^{*,3}(\gamma) = P_{peak} \leq P_t^{*,1}(\gamma) < P_t^{*,2}(\gamma)$; then, the optimal power level is $P_t^*(\gamma) = P_t^{*,3}(\gamma) = P_{peak}$ due to the peak power constraint. Also, $P_t^{*,3}(\gamma) = P_{peak}$ must satisfy $\mathrm{P_D}'(P_t^{*,3}(\gamma), \gamma) \geq \lambda^*$ based on (24) if the optimal power level is $P_t^{*,3}$. Since $\mathrm{P_D}'(P_t(\gamma), \gamma) < \lambda^*$ for $P_t(\gamma) \in (P_t^{*,1}(\gamma), P_t^{*,2}(\gamma))$, $P_t^{*,3}(\gamma) = P_{peak}$ cannot be optimal if $P_t^{*,1}(\gamma) < P_{peak} < P_t^{*,2}(\gamma)$ and the optimal power level is $P_t^*(\gamma) = P_t^{*,1}(\gamma)$. If $P_t^{*,1}(\gamma) < P_t^{*,2}(\gamma) < P_{peak}$, $P_t^{*,3}(\gamma) = P_{peak}$ cannot be optimal due to the similar reason. Then, the maximizer of $\mathrm{P_D}(P_t^*(\gamma), \gamma) - \lambda^* P_t^*(\gamma)$ is either $P_t^{*,1}(\gamma)$ and $P_t^{*,2}(\gamma)$, which can be found by the comparison below:

$$\mathrm{P_D}(P_t^{*,1}(\gamma), \gamma) - \lambda^* P_t^{*,1}(\gamma) \underset{P_t^{*,2}}{\overset{P_t^{*,1}}{\gtrless}} \mathrm{P_D}(P_t^{*,2}(\gamma), \gamma) - \lambda^* P_t^{*,2}(\gamma) \tag{29}$$

$$\frac{\mathrm{P_D}(P_t^{*,1}(\gamma), \gamma) - \mathrm{P_D}(P_t^{*,2}(\gamma), \gamma)}{P_t^{*,1}(\gamma) - P_t^{*,2}(\gamma)} \underset{P_t^{*,2}}{\overset{P_t^{*,1}}{\lessgtr}} \lambda^* \tag{30}$$

Since $\mathrm{P_D}'(P_t(\gamma), \gamma) < \lambda^*$ for $P_t(\gamma) \in (P_t^{*,1}(\gamma), P_t^{*,2}(\gamma))$, $\frac{\mathrm{P_D}(P_t^{*,1}(\gamma), \gamma) - \mathrm{P_D}(P_t^{*,2}(\gamma), \gamma)}{P_t^{*,1}(\gamma) - P_t^{*,2}(\gamma)} < \lambda^*$. Then, $P_t^{*,2}(\gamma)$ cannot be optimal. Lastly, if $P_t^{*,1}(\gamma) < P_t^{*,2}(\gamma) = P_{peak}$, $P_t^{*,2}(\gamma) = P_{peak}$ cannot be optimal due to the comparison of the candidates as in (30). Thus, the optimal power level for $\gamma = \gamma_l$ can be selected as $P_t^*(\gamma) = \min\{P_t^{*,1}(\gamma), P_{peak}\}$.

3) $\mathrm{P_D}'(I_2(\gamma), \gamma) > \lambda^* > \mathrm{P_D}'(I_1(\gamma), \gamma)$: This case is valid for $\gamma \in (\gamma_l, \gamma_u)$. In this case, $\mathrm{P_D}'(P_t^*(\gamma), \gamma) = \lambda^*$ has three solutions; i.e., there are three candidates for the optimal power level. Let these candidates be $P_t^{*,1}(\gamma)$, $P_t^{*,2}(\gamma)$ and $P_t^{*,3}(\gamma)$ with $P_t^{*,1}(\gamma) < P_t^{*,2}(\gamma) < P_t^{*,3}(\gamma)$. Note that $P_t^{*,1}(\gamma) \in [0, I_1(\gamma))$, $P_t^{*,2}(\gamma) \in (I_1(\gamma), I_2(\gamma))$ and $P_t^{*,3}(\gamma) \in (I_2(\gamma), \infty)$. Also notice that $\mathrm{P_D}(P_t(\gamma), \gamma) - \lambda^* P_t(\gamma)$ is an increasing function of $P_t(\gamma)$ for $P_t(\gamma) \in (0, P_t^{*,1}(\gamma))$, a decreasing function for $P_t(\gamma) \in (P_t^{*,1}(\gamma), P_t^{*,2}(\gamma))$, increasing for $P_t(\gamma) \in (P_t^{*,2}(\gamma), P_t^{*,3}(\gamma))$, and decreasing for $P_t(\gamma) \in (P_t^{*,3}(\gamma), \infty)$. Thus, from the candidates $P_t^{*,1}(\gamma)$, $P_t^{*,2}(\gamma)$ and $P_t^{*,3}(\gamma)$, either $P_t^{*,1}(\gamma)$ or $P_t^{*,3}(\gamma)$ is a maximizer for $\mathrm{P_D}(P_t(\gamma), \gamma) - \lambda^* P_t(\gamma)$. Since $P_t^{*,1}(\gamma) < I_1(\gamma) < I_2(\gamma) < P_t^{*,3}(\gamma)$, the optimal power level cannot take any values between $I_1(\gamma)$ and $I_2(\gamma)$ for the case that no peak power constraint is considered. On the other hand, the optimal power level may not be $P_t^{*,1}(\gamma)$ or $P_t^{*,3}(\gamma)$ due to the peak power constraint. If $P_{peak} \leq P_t^{*,1}(\gamma) < P_t^{*,3}(\gamma)$; then, the optimal power level is $P_t^*(\gamma) = P_{peak}$ since $\mathrm{P_D}'(P_t(\gamma), \gamma)$ is monotone decreasing for $P_t(\gamma) \in (0, I_1(\gamma))$ and $P_{peak} \leq P_t^{*,1}(\gamma) < I_1(\gamma)$. Otherwise, if $P_t^{*,1}(\gamma) < P_{peak}$, the optimal power level is either $P_t^{*,1}(\gamma)$ or $\min\{P_t^{*,3}(\gamma), P_{peak}\}$. Note that $P_t^{*,1}(\gamma)$ and $P_t^{*,3}(\gamma)$ are differentiable functions of γ for $\gamma \in [\gamma_l, \gamma_u]$ as shown in Case-1 and Case-5, respectively, since $P_t^{*,1}(\gamma) < I_1(\gamma)$ and $P_t^{*,3}(\gamma) > I_2(\gamma)$. Then, $\min\{P_t^{*,3}(\gamma), P_{peak}\}$ is a continuous function of γ for $\gamma \in [\gamma_l, \gamma_u]$ and may not be differentiable at a γ value where $P_t^{*,3}(\gamma) = P_{peak}$. Let $\tilde{P}_t^{*,3} \triangleq \min\{P_t^{*,3}(\gamma), P_{peak}\}$ and $\mathcal{S}(\gamma) \triangleq \left(\mathrm{P_D}(P_t^{*,1}(\gamma), \gamma) - \lambda^* P_t^{*,1}(\gamma)\right) - \left(\mathrm{P_D}(\tilde{P}_t^{*,3}(\gamma), \gamma) - \lambda^* \tilde{P}_t^{*,3}(\gamma)\right)$, which are continuous in γ for $\gamma \in [\gamma_l, \gamma_u]$. Then, via (29), the optimal power level can be chosen by comparing $\mathcal{S}(\gamma)$ against zero; that is,

$\mathcal{S}(\gamma) \underset{\tilde{P}_t^{*,3}}{\overset{P_t^{*,1}}{\gtrless}} 0$. In addition, Case-2 (the equivalence case can be obtained by comparing $P_t^{*,1}(\gamma_l)$ and $P_t^{*,2}(\gamma_l) = \tilde{P}_t^{*,3}(\gamma_l)$) and Case-4 (the equivalence case can be obtained by comparing $\tilde{P}_t^{*,3}(\gamma_u)$ and $P_t^{*,2}(\gamma_u) = P_t^{*,1}(\gamma_u)$) imply $\mathcal{S}(\gamma_l) > 0$ and $\mathcal{S}(\gamma_u) < 0$, respectively. Note the following identity:

$$\begin{aligned}
&\frac{\partial \mathrm{P_D}(P_t(\gamma),\gamma)}{\partial\left(P_t(\gamma)\right)} = \lambda^* \\
&= \frac{\sqrt{\gamma}}{2\sqrt{2\pi}\sqrt{P_t(\gamma)}} \exp\left\{-\frac{1}{2}\left(\mathcal{Q}^{-1}(\alpha) - \sqrt{P_t(\gamma)\gamma}\right)^2\right\} \\
&= \frac{\gamma}{P_t(\gamma)} \frac{\sqrt{P_t(\gamma)}}{2\sqrt{2\pi}\sqrt{\gamma}} \exp\left\{-\frac{1}{2}\left(\mathcal{Q}^{-1}(\alpha) - \sqrt{P_t(\gamma)\gamma}\right)^2\right\} \\
&= \frac{\gamma}{P_t(\gamma)} \frac{\partial \mathrm{P_D}(P_t(\gamma),\gamma)}{\partial\gamma} \\
&\Rightarrow \frac{\partial \mathrm{P_D}(P_t(\gamma),\gamma)}{\partial\gamma} = \lambda^* \frac{P_t(\gamma)}{\gamma}. \qquad (31)
\end{aligned}$$

Then, from (31),

$$\begin{aligned}
&\frac{\mathrm{d}\mathcal{S}(\gamma)}{\mathrm{d}\gamma} \\
&= \left(\frac{\partial \mathrm{P_D}(P_t^{*,1}(\gamma),\gamma)}{\partial\left(P_t^{*,1}(\gamma)\right)} \frac{\mathrm{d}P_t^{*,1}(\gamma)}{\mathrm{d}\gamma} + \frac{\partial \mathrm{P_D}(P_t^{*,1}(\gamma),\gamma)}{\partial\gamma} - \lambda^* \frac{\mathrm{d}P_t^{*,1}(\gamma)}{\mathrm{d}\gamma}\right) \\
&- \left(\frac{\partial \mathrm{P_D}(\tilde{P}_t^{*,3}(\gamma),\gamma)}{\partial\left(\tilde{P}_t^{*,3}(\gamma)\right)} \frac{\mathrm{d}\tilde{P}_t^{*,3}(\gamma)}{\mathrm{d}\gamma} + \frac{\partial \mathrm{P_D}(\tilde{P}_t^{*,3}(\gamma),\gamma)}{\partial\gamma} - \lambda^* \frac{\mathrm{d}\tilde{P}_t^{*,3}(\gamma)}{\mathrm{d}\gamma}\right) \\
&= \left(\lambda^* \frac{\mathrm{d}P_t^{*,1}(\gamma)}{\mathrm{d}\gamma} + \lambda^* \frac{P_t^{*,1}(\gamma)}{\gamma} - \lambda^* \frac{\mathrm{d}P_t^{*,1}(\gamma)}{\mathrm{d}\gamma}\right) \\
&- \left(\lambda^* \frac{\mathrm{d}\tilde{P}_t^{*,3}(\gamma)}{\mathrm{d}\gamma} + \lambda^* \frac{\tilde{P}_t^{*,3}(\gamma)}{\gamma} - \lambda^* \frac{\mathrm{d}\tilde{P}_t^{*,3}(\gamma)}{\mathrm{d}\gamma}\right) \\
&= \lambda^* \frac{P_t^{*,1}(\gamma) - \tilde{P}_t^{*,3}(\gamma)}{\gamma} < 0 \qquad (32)
\end{aligned}$$

is obtained for γ such that $\mathcal{S}(\gamma)$ is differentiable. The expression in (32) implies that $\mathcal{S}(\gamma)$ is a monotone decreasing continuous function of γ for $\gamma \in [\gamma_l, \gamma_u]$, and since $\mathcal{S}(\gamma_l) > 0$ and $\mathcal{S}(\gamma_u) < 0$, there must exist a unique value $\widetilde{\gamma} \in (\gamma_l, \gamma_u)$ such that $\mathcal{S}(\widetilde{\gamma}) = 0$. The optimal power allocation strategy is indifferent between the power levels $P_t^{*,1}(\widetilde{\gamma})$ and $\tilde{P}_t^{*,3}(\widetilde{\gamma})$ for $\gamma = \widetilde{\gamma}$, and the optimal power level is $P_t^{*,1}(\gamma)$ for $\gamma_l < \gamma < \widetilde{\gamma}$, whereas the optimal power level is $\tilde{P}_t^{*,3}(\gamma)$ for $\widetilde{\gamma} < \gamma < \gamma_u$. Hence, there exists a positive jump from $P_t^{*,1}(\widetilde{\gamma})$ to $\tilde{P}_t^{*,3}(\widetilde{\gamma})$ at $\gamma = \widetilde{\gamma}$.

4) $\underline{\lambda^* = \mathrm{P_D}'(I_1(\gamma),\gamma)}$: This case is valid for $\gamma = \gamma_u$. In this case, $\mathrm{P_D}'(P_t(\gamma),\gamma) = \lambda^*$ has two solutions; i.e., there are two candidates for the optimal power level. Let these candidates be $P_t^{*,1}(\gamma)$, and $P_t^{*,2}(\gamma)$ with $P_t^{*,1}(\gamma) = I_1(\gamma) < I_2(\gamma) < P_t^{*,2}(\gamma)$. If $P_{peak} \leq P_t^{*,2}(\gamma)$, $P_t^*(\gamma) = P_{peak}$ since $\mathrm{P_D}'(P_t(\gamma),\gamma) \geq \lambda^*$ for all $P_t(\gamma) \in (0, P_{peak}]$. Otherwise, if $P_t^{*,2}(\gamma) < P_{peak}$; then, the optimal power level is either $P_t^{*,1}(\gamma)$ or $P_t^{*,2}(\gamma)$. Since $\mathrm{P_D}'(P_t(\gamma),\gamma) > \lambda^*$ for $P_t(\gamma) \in (P_t^{*,1}(\gamma), P_t^{*,2}(\gamma))$, $\frac{\mathrm{P_D}(P_t^{*,1}(\gamma),\gamma) - \mathrm{P_D}(P_t^{*,2}(\gamma),\gamma)}{P_t^{*,1}(\gamma) - P_t^{*,2}(\gamma)} > \lambda^*$. Then, $P_t^{*,1}(\gamma)$ cannot be optimal based on (30). Thus, the optimal power level can be selected as $P_t^*(\gamma) = \min\{P_t^{*,2}(\gamma), P_{peak}\}$.

5) $\underline{\lambda^* < \mathrm{P_D}'(I_1(\gamma),\gamma)}$: This case is valid for $\gamma \in (\gamma_u, \infty)$. In this case, $\mathrm{P_D}'(P_t(\gamma),\gamma) > \lambda^*$ for $P_t(\gamma) \leq I_2(\gamma)$ and $\mathrm{P_D}'(P_t(\gamma),\gamma)$ is a monotone decreasing function of $P_t(\gamma)$ for $P_t(\gamma) \in (\gamma_u, \infty)$. Based on (24), one of the following conditions holds:

- If $\mathrm{P_D}'(P_{peak},\gamma) < \lambda^*$ for all $\gamma \in (\gamma_u, \infty)$, $\mathrm{P_D}'(P_t^*(\gamma),\gamma) = \lambda^*$ for all $\gamma \in (\gamma_u, \infty)$. Also, $\mathrm{P_D}'(P_t(\gamma),\gamma) = \lambda^*$ has only one solution (optimal power level) $P_t^*(\gamma)$, where $P_t^*(\gamma) > I_2(\gamma)$. By using the definitions in the proof of Theorem 1, if $P_t(\gamma) > I_2(\gamma)$, then $t = xy = \sqrt{\gamma P_t(\gamma)} > \sqrt{\gamma I_2(\gamma)} = \sqrt{\gamma \frac{\left(\mathcal{Q}^{-1}(\alpha) + \sqrt{(\mathcal{Q}^{-1}(\alpha))^2 - 4}\right)^2}{4\gamma}} = \frac{\mathcal{Q}^{-1}(\alpha) + \sqrt{(\mathcal{Q}^{-1}(\alpha))^2 - 4}}{2} = t_2$. We also have $h(t) > 0$ for $t > t_2$. Thus, due to a similar reasoning to that in the proof of Theorem 1, $P_t^*(\gamma)$ constitutes a continuous function in γ for $\gamma \in (\gamma_u, \infty)$.
- If $\mathrm{P_D}'(P_{peak},\gamma) \geq \lambda^*$ for some $\gamma \in (\gamma_u, \infty)$, $\mathrm{P_D}'(P_{peak},\gamma) \geq \lambda^*$ for $\gamma \in [\bar{\gamma}_l, \bar{\gamma}_u]$ due to the function properties of $\mathrm{P_D}'(P_{peak},\gamma)$ where $\bar{\gamma}_l$ and $\bar{\gamma}_u$ are positive finite values. If $\gamma_u < \bar{\gamma}_l < \bar{\gamma}_u$ holds, $P_t^*(\gamma) = P_{peak}$ for $\gamma \in [\bar{\gamma}_l, \bar{\gamma}_u]$ and $P_t^*(\gamma)$ is continuous function of γ for $\gamma \in (\gamma_u, \infty)$, which increases for $\gamma \in (\gamma_u, \bar{\gamma}_l)$ and decreases for $\gamma \in (\bar{\gamma}_u, \infty)$ based on a similar approach to that in the first condition. Otherwise, if $\bar{\gamma}_l = \gamma_u \leq \bar{\gamma}_u$, $P_t^*(\gamma)$ is a continuous function of γ for $\gamma \in (\gamma_u, \infty)$, which becomes equal to P_{peak} for $\gamma \in [\gamma_u, \bar{\gamma}_u]$ and decreases for $\gamma \in (\bar{\gamma}_u, \infty)$.

Considering the analyses of the different cases above, the following summary can be stated:

1) For $\gamma < \widetilde{\gamma}$, the optimal power allocation policy is continuous in γ and the optimal power level is always lower than $I_1(\gamma)$, where $\widetilde{\gamma} \in (\gamma_l, \gamma_u)$.

2) At $\gamma = \widetilde{\gamma}$, the optimal power allocation policy has a positive jump.

3) For $\gamma > \widetilde{\gamma}$, the optimal power allocation policy is continuous in γ. ■

REFERENCES

[1] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1986–1992, Nov. 1997.

[2] X. Kang, Y. C. Liang, A. Nallanathan, H. K. Garg, and R. Zhang, "Optimal power allocation for fading channels in cognitive radio networks: Ergodic capacity and outage capacity," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 940–950, Feb. 2009.

[3] F. Zhou, N. C. Beaulieu, Z. Li, J. Si, and P. Qi, "Energy-efficient optimal power allocation for fading cognitive radio channels: Ergodic capacity, outage capacity, and minimum-rate capacity," *IEEE Trans. Wireless Commun.*, vol. 15, no. 4, pp. 2741–2755, Apr. 2016.

[4] D. Xu, Z. Feng, and P. Zhang, "Minimum average BER power allocation for fading channels in cognitive radio networks," in *IEEE Wireless Commun. and Networking Conf. (WCNC)*, Mar. 2011, pp. 78–83.

[5] P. H. Reddy, V. N. Kumar, M. Z. A. Khan, and M. B. Srinivas, "Optimal power allocation in space and time in MISO Rayleigh fading channels with peak to average power ratio constraint," in *IEEE Sarnoff Symposium*, Apr. 2007, pp. 1–6.

[6] C. Huang, R. Zhang, and S. Cui, "Optimal power allocation for outage probability minimization in fading channels with energy harvesting constraints," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 1074–1087, Feb. 2014.

[7] A. Helmy, L. Musavian, and T. Le-Ngoc, "Energy-efficient power adaptation over a frequency-selective fading channel with delay and power constraints," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4529–4541, Sep. 2013.

[8] G. Miao, N. Himayat, Y. Li, and D. Bormann, "Energy efficient design in wireless OFDMA," in *IEEE International Conference on Communications (ICC)*, May 2008, pp. 3307–3312.

[9] I. Hadj-Kacem, N. Sellami, I. Fijalkow, and A. Roumy, "Joint training interval length and power allocation optimization for MIMO flat fading channels," in *6th International Symposium on Wireless Communication Systems*, Sep. 2009, pp. 16–20.

[10] L. Musavian and T. Le-Ngoc, "Energy-efficient power allocation over Nakagami-m fading channels under delay-outage constraints," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4081–4091, Aug. 2014.

[11] Z. Feng and I. Wassell, "Dynamic power control and optimization scheme for QoS-constrained cooperative wireless sensor networks," in *IEEE International Conference on Communications (ICC)*, May 2016, pp. 1–6.

[12] A. Olfat and M. Shikh-Bahaei, "Optimum power and rate adaptation for MQAM in Rayleigh flat fading with imperfect channel estimation," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2622–2627, July 2008.

[13] L. Goldfeld, V. Lyandres, and D. Wulich, "Minimum BER power loading for OFDM in fading channel," *IEEE Trans. Commun.*, vol. 50, no. 11, pp. 1729–1733, Nov. 2002.

[14] A. P. Iserte, M. A. L. Hernandez, and A. I. Perez-Neira, "Robust power allocation for minimum BER in a SISO-OFDM system," in *IEEE Int. Conf. on Commun. (ICC)*, vol. 2, May 2003, pp. 1263–1267 vol.2.

[15] P. Zhang and H. C. Yang, "Minimum-BER power allocation for multicarrier systems with outdated channel state information," in *2007 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, Apr. 2007, pp. 180–183.

[16] L. Li and A. Goldsmith, "Minimum outage probability and optimal power allocation for fading multiple-access channels," in *IEEE International Symposium on Information Theory (ISIT)*, June 2000, pp. 305–.

[17] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1468–1489, July 1999.

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

[19] J. Zhu and B. Zheng, "Detection probability analysis of cooperative spectrum sensing in Rayleigh fading channels," in *IEEE/ACIS International Conference on Computer and Information Science (ICIS)*, June 2009, pp. 177–182.

[20] M. Azizoglu, "Convexity properties in binary detection problems," *IEEE Trans. Inf. Theory*, vol. 42, no. 4, pp. 1316–1321, July 1996.

[21] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York, NY, USA: Springer-Verlag, 1994.

[22] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[23] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 57–73, Feb. 2008.

[24] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[25] Z. Q. Luo and S. Zhang, "Duality gap estimation and polynomial time approximation for optimal spectrum management," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2675–2689, July 2009.

[26] D. Blackwell, "On a theorem of Lyapunov," *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 112–114, 03 1951. [Online]. Available: http://dx.doi.org/10.1214/aoms/1177729699

[27] A. Liapounoff, "Sur les fonctions-vecteurs complètement additives," *Izv. Akad. Nauk SSSR Ser. Mat.*, vol. 4, pp. 465–478, 1940.

[28] K. Seong, M. Mohseni, and J. Cioffi, "Optimal resource allocation for OFDMA downlink systems," in *IEEE International Symposium on Information Theory*, July 2006, pp. 1394–1398.

[29] R. T. Rockafellar, *Convex analysis*, ser. Princeton Mathematical Series. Princeton, N. J.: Princeton University Press, 1970.

[30] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, "Optimal multiuser spectrum balancing for digital subscriber lines," *IEEE Trans. Commun.*, vol. 54, no. 5, pp. 922–933, May 2006.

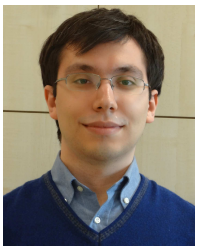
[31] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, July 2006.

[32] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA: Athena Scientific, 1999.

[33] E. A. Coddington, *An Introduction to Ordinary Differential Equations*. Courier Corporation, 2012.

Serkan Sarıtaş (S'16) received his B.Sc. degree in Electrical and Electronics Engineering and M.S. degree in Computer Engineering from Bilkent University in 2010 and 2013, respectively. Currently he is working towards the Ph.D. degree in the Department of Electrical and Electronics Engineering of Bilkent University. His research interests include communication systems, game theory and information theory.

Berkan Dulek received the B.S., M.S. and Ph.D. degrees in electrical and electronics engineering from Bilkent University in 2003, 2006 and 2012, respectively. From 2007 to 2010, he worked at industry. From 2012 to 2013, he was a postdoctoral research associate at the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY. Since 2014, he has been with the Department of Electrical and Electronics Engineering at Hacettepe University, where he is currently an Associate Professor. His research interests are in statistical signal processing, detection and estimation theory, and communication theory.



Ahmet Dundar Sezer was born in 1989 in Emet, Kutahya, Turkey. He received both his 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 towards 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, Turkey in 2001, and the Ph.D. degree in Electrical Engineering from Princeton University in 2006. From 2006 to 2007, he worked at Mitsubishi Electric Research Laboratories, Cambridge, MA. Since 2007, he has been with the Department of Electrical and Electronics Engineering at Bilkent University, where he is currently an Associate Professor. Dr. Gezicis 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.

Serdar Yüksel (M'11) received his B.Sc. degree in electrical and electronics engineering from Bilkent University in 2001; M.S. and Ph.D. degrees in electrical and computer engineering from the University of Illinois at Urbana-Champaign in 2003 and 2006, respectively. He was a post-doctoral researcher at Yale University before joining Queen's University as an Assistant Professor of Mathematics and Engineering in the Department of Mathematics and Statistics, where he is now an Associate Professor. He has held visiting positions at KTH Royal Institute of Technology, Sweden and Bilkent University, Turkey. He has been awarded the 2013 CAIMS/PIMS Early Career Award in Applied Mathematics and is currently an associate editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL. His research interests are on stochastic control, information theory, probability, and networked control.