On Peak versus Average Interference Power Constraints for Spectrum Sharing in Cognitive Radio Networks

Rui Zhang

Abstract

This paper considers the spectrum sharing for mobile wireless communications, where a secondary user or cognitive radio (CR) communicates using the same bandwidth originally allocated to an existing primary user link. An effective means for the CR to use to protect the primary transmissions is by applying the so-called *interference-temperature* constraint, whereby the resultant interference power at each primary receiver is kept below some predefined value. For the fading primary and secondary user channels, the interference-power constraint at each primary receiver is usually one of the following two types: One is to regulate the *average-interference-power* (AIP) over all the fading states, while the other is to limit the *peak-interference-power* (PIP) at any of the fading states. From the secondary user's perspective, given the same (average or peak) power-constraint value, the AIP constraint is more favorable than the PIP because of the more flexibility in dynamically allocating transmit powers over the fading states. However, from the perspective of protecting the primary user transmissions, the more restrictive PIP constraint appears at a first glance to be a better option than the AIP. Surprisingly, in this paper it is shown that this seemingly apparent fact is usually untrue for the primary user transmission over the fading channel. This is due to an interesting *interference diversity* phenomenon, i.e., the randomized interference powers from the secondary transmitter in the AIP case can be in fact more advantageous for minimizing the resultant primary user capacity losses as compared to the more deterministic ones in the PIP case. Therefore, the AIP can be superior than the PIP for both the primary and secondary user transmissions.

Index Terms

Cognitive radio, spectrum sharing, interference temperature, interference diversity, fading channel, capacity.

I. INTRODUCTION

This paper is concerned with a typical *spectrum sharing* scenario for wireless mobile communications, where a secondary user, also commonly known as the *cognitive radio* (CR), intends to communicate using the same bandwidth that has already been allocated to an existing primary user link. For such a scenario, a secondary or CR transmitter usually needs to deal with a fundamental tradeoff between maximizing its own transmission throughput and minimizing the amount of interferences it causes to the primary receivers. A commonly adopted means to protect the primary transmission is by adapting the secondary user's transmission such that the resultant interference power, or the so-called *interference-temperature*, at each primary receiver is kept below some prescribed threshold [1], [2]. A general approach for designing the transmission strategies of a CR under the interference-temperature constraints is *dynamic resource allocation*, by which the transmit power, bit-rate, bandwidth and antenna beam of the secondary user are dynamically allocated based upon the fading channel state information (CSI) available at the CR transmitter. For the single-input single-output (SISO) fading primary and secondary user channels with single-antenna terminals, transmit power control for the CR has been studied in [3] by exploiting the CSI on the interfering channel from the secondary transmitter to the primary receiver, and in [4] with availability of the additional CSI on the fading secondary user channel and/or the fading primary user channel. In [5], the authors also proposed both optimal and suboptimal spatial adaptation schemes for the multi-antenna CR transmitter.

In this paper, we focus on the SISO fading primary and secondary user channels. In this case, the interference-power constraint at each primary receiver is usually one of the following two types: One is the *long-term* constraint that regulates the *average-interference-power* (AIP) over all the fading states, while the other is a *short-term* one and limits the *peak-interference-power* (PIP) at any of the fading states. Apparently, the PIP constraint is more restrictive than the AIP given their same (average or peak) power-constraint value. From the secondary user's perspective, it is easy to see that the AIP constraint is more favorable compared to the PIP, since the former provides the CR transmitter the more flexibility in dynamically allocating its transmit powers over the fading states, and as a result maximizes the long-term average throughput. On the other hand, from the perspective of protecting the primary user

transmissions, the more restrictive PIP constraint seems at a first glance to be a better option than the AIP. Surprisingly, in this paper it will be shown that this seemingly apparent fact is indeed untrue for the primary user transmission over the fading channel using some commonly adopted power-control policies. This is due to an interesting *interference diversity* phenomenon briefly explained as follows: Because of the convexity of the capacity function with respect to the noise/interferecne power, the randomized interference powers from the secondary transmitter in the AIP case can be in fact more advantageous for minimizing the resultant primary user capacity losses as compared to the more deterministic ones in the PIP case. Therefore, this paper discovers an important design rule for the CR networks operating using the interference-temperature principle, i.e., the AIP can be indeed superior than the PIP not only from the secondary link throughput maximization viewpoint, but also for the better protection of the primary user transmission.

The rest of this paper is organized as follows. Section II presents the system model of spectrum sharing in the CR network. Section III considers first the secondary link and compares its ergodic capacities subject to the AIP and PIP constraint. Section IV then studies the primary user channel capacities under the AIP or PIP from the secondary transmitter for different types of the primary user power-control policies, and shows the capacity gains under the AIP over the PIP by exploiting the interference diversity. Section V considers both the primary and secondary links jointly and shows the simulation results on their achievable throughput under spectrum sharing over the fading channel. Finally, Section VI concludes this paper.

Notation: $|z|$ denotes the Euclidean norm of a complex number z. $\mathbb{E}[\cdot]$ denotes statistical expectation. The distribution of a circular symmetric complex Gaussian (CSCG) variable with the mean x and the variance y is denoted by $\mathcal{CN}(x, y)$, and $\sim$ means "distributed as". $\max(x, y)$ and $\min(x, y)$ denote, respectively, the maximum and the minimum between two real numbers x and y and for a real number a, $(a)^+ = \max(0, a)$.

II. SYSTEM MODEL

As shown in Fig. 1, a spectrum sharing scenario is considered where a secondary/CR link consisting of a secondary transmitter (ST) and a secondary receiver (SR) shares the same bandwidth for transmission

with an existing primary link consisting of a primary transmitter (PT) and a primary receiver (PR). All terminals are assumed to be equipped with a single antenna. We consider a slow-fading environment and for simplicity, assume a block-fading (BF) channel model for both the primary and secondary links. Furthermore, we assume coherent demodulations at both the primary and secondary receivers, and thus only the fading channel power gain (amplitude square) is of interest. Let e_i denote the channel power gain e for the fading channel from the ST and the SR at the fading state i, $i = 1, 2, \ldots, N$, where N denotes the number of fading states that can be either finite or infinite, and $e = [e_1, e_2, \ldots, e_N]$. Similarly, g_i, f_i, and o_i denote the channel gains at the i-th fading state of g, f, and o, corresponding to the fading channel from the ST to the PR, the PT to the PR, and the PT to the SR, respectively. It is assumed that e_i, g_i, f_i and o_i are independent variables each having a discrete/continuous probability density function (PDF). It is also assumed that the additive noises at both primary and secondary receivers are independent CSCG variables each $\sim \mathcal{CN}(0, 1)$. Since we are interested in the information-theoretic limits of the primary and secondary links, it is assumed that the optimum Gaussian code-books are used by both the primary and secondary transmitters.

First, for the primary link, the transmit powers at different fading states i are denoted as $\{q_i\}$. It is assumed that the primary user is oblivious to the secondary transmission and thus does not attempt to protect the secondary transmission nor cooperate with the secondary user for transmission. Due to the concurrent secondary transmission, the PR may observe an additional interference power, denoted as $I_i = g_i p_i$ at the fading state i, where $\{p_i\}$ are the secondary user transmit powers at different fading states i. The primary user power-control policy $\mathcal{P}_p(\boldsymbol{f}, \boldsymbol{I})$ is then a mapping from f_i and I_i to q_i, where $\boldsymbol{I} = [I_1, I_2, \ldots, I_N]$, subject to an average transmit power constraint Q, i.e., $\mathbb{E}[q_i] \leq Q$. By treating the interference from the ST as the additional Gaussian noise at the PR, the achievable average rate of the primary user channel can be then expressed as

$$R_p = \mathbb{E}\left[\log\left(1 + \frac{f_i q_i}{1 + I_i}\right)\right]. \tag{1}$$

Notice that q_i may also be a function of f_i and I_i.

On the other hand, for the secondary link, since the CR needs to protect the primary transmission, its power-control policy needs to consider both the primary and secondary transmissions. It is assumed

that the interfering channel gain g_i from the ST to the PR is perfectly known at the ST at each fading state i. For convenience, we combine the Gaussian-distributed interferences from the PT with the independent additive Gaussian noises at the SR, and define the equivalent secondary user channel gain as $h_i \triangleq \frac{e_i}{1+q_i o_i}, \forall i$, which is also assumed to be known at the ST at each i. Let $\boldsymbol{h} = [h_1, h_2, \ldots, h_N]$. Thus, the secondary user power-control policy can be expressed as $\mathcal{P}_s(\boldsymbol{h}, \boldsymbol{g})$, subject to an average transmit power constraint P, i.e., $\mathbb{E}[p_i] \leq P$. The achievable average rate of the secondary user channel can be then expressed as

$$R_s = \mathbb{E}\left[\log\left(1 + h_i p_i\right)\right]. \tag{2}$$

Again, notice that p_i maybe a function of both h_i and g_i.

III. SECONDARY LINK CAPACITIES UNDER AIP AND PIP

We first consider the secondary user transmission subject to the AIP or the PIP constraint at the PR as a practical means to protect the primary transmission. The AIP constraint regulates the average-interference-power in the long-term and is thus defined as

$$\mathbb{E}[I_i] \leq \Gamma_a, \tag{3}$$

where Γ_a denotes the predefined AIP value. In contrast, the PIP constraint limits the short-term peak-interference-power at any of the fading states and is thus defined as

$$I_i \leq \Gamma_p, \ \forall i, \tag{4}$$

where Γ_a denotes the predefined PIP value. Notice that the PIP constraint is more restrictive than the AIP. This can be easily seen by observing that given $\Gamma_p = \Gamma_a$, (4) implies (3) but not vice versa. Therefore, from the secondary user's perspective, applying the AIP constraint is more favorable than the PIP because it provides the secondary user power-control policy $\mathcal{P}_s(\boldsymbol{h}, \boldsymbol{g})$ the more flexibility in adapting the transmit powers over the fading states. In this paper, we consider the ergodic capacity as the throughput limit of the secondary user channel. From (2) and (3), the optimal $\mathcal{P}_s(\boldsymbol{h}, \boldsymbol{g})$ for the secondary user fading channel to maximize its ergodic capacity under both the AIP constraint at the PR and the secondary user own transmit-power constraint can be obtained as the solutions of the following

optimization problem (**P1**):

$$\underset{\{p_i\}}{\text{Maximize}} \quad \mathbb{E}\left[\log\left(1 + h_i p_i\right)\right]$$

$$\text{Subject to} \quad \mathbb{E}[g_i p_i] \leq \Gamma_a, \tag{5}$$

$$\mathbb{E}[p_i] \leq P, \tag{6}$$

$$p_i \geq 0, \ \forall i. \tag{7}$$

Notice that the above problem has been studied in [3] without the secondary user own transmit-power constraint (6). It is easy to verify that P1 is a convex optimization problem, and thus by applying the Karush-Kuhn-Tacker (KKT) conditions [6] that are satisfied by the optimal solutions of p_i's for P1, denoted as $\{p_i^{(1)}\}$, we can obtain

$$p_i^{(1)} = \left(\frac{1}{\nu^{(1)} g_i + \mu^{(1)}} - \frac{1}{h_i}\right)^+, \tag{8}$$

where $\nu^{(1)}$ and $\mu^{(1)}$ are the nonnegative optimal dual solutions corresponding to the constraint (5) and (6), respectively, which satisfy the following Complementary Slackness (CS) conditions [6]:

$$\nu^{(1)}\left(\mathbb{E}\left[g_i p_i^{(1)}\right] - \Gamma_a\right) = 0, \tag{9}$$

$$\mu^{(1)}\left(\mathbb{E}\left[p_i^{(1)}\right] - P\right) = 0. \tag{10}$$

Thus, if any of $\nu^{(1)}$ and $\mu^{(1)}$ is strictly positive, the corresponding constraint (5) or (6) must be satisfied with equality by $\{p_i^{(1)}\}$. On the other hand, if any of the constraint (5) and (6) is satisfied with strict inequality, the corresponding dual solution must be zero. Numerically, $\nu^{(1)}$ and $\mu^{(1)}$ can be obtained using, e.g., the ellipsoid method [7].[1] Substituting (8) into (2) or the objective function of P1 yields the secondary user channel ergodic capacity under the AIP constraint expressed as

$$C_{s,a}^{\text{ER}} = \mathbb{E}\left[\left(\log\left(\frac{h_i}{\nu^{(1)} g_i + \mu^{(1)}}\right)\right)^+\right]. \tag{11}$$

Next, if the PIP constraint is applied instead of the AIP, i.e., replacing the AIP constraint (5) in Problem P1 by the PIP constraint (4), which can also be expressed as

$$g_i p_i \leq \Gamma_p, \ \forall i, \tag{12}$$

---

[1]The ellipsoid method applies the sub-gradient $\Gamma_a - \mathbb{E}[g_i p_i[n]]$ and $P - \mathbb{E}[p_i[n]]$ to iteratively update $\nu[n+1]$ and $\mu[n+1]$ until they converge to $\nu^{(1)}$ and $\mu^{(1)}$, respectively, where $\{p_i[n]\}$ are obtained from (8) for some given $\nu[n]$ and $\mu[n]$ at the n-th iteration.

and denoting the resultant new problem as **P2**, the associated optimal power allocations p_i's, denoted as $\{p_i^{(2)}\}$, can be similarly obtained like P1 as

$$p_i^{(2)} = \min\left(\frac{\Gamma_p}{g_i}, \left(\frac{1}{\mu^{(2)}} - \frac{1}{h_i}\right)^+\right), \tag{13}$$

where $\mu^{(2)}$ is the optimal dual solution corresponding to the secondary user own transmit-power constraint (6), which satisfies the similar CS condition like (10), and can be numerically obtained using a simpler version of the ellipsoid method, the bisection method [6]. Substituting (13) into (2) yields the secondary user channel ergodic capacity under the PIP constraint expressed as

$$C_{s,p}^{\text{ER}} = \mathbb{E}\left[\min\left(\log\left(1 + \frac{h_i\Gamma_p}{g_i}\right), \left(\log\left(\frac{h_i}{\mu^{(2)}}\right)\right)^+\right)\right]. \tag{14}$$

Notice that how the optimal power allocations given in (8) for the AIP case are different from (13) for the PIP case. Suppose that $\mathcal{P}_s(\boldsymbol{h}, \boldsymbol{g})$ is such that the AIP constraint (5) for P1 is inactive, while it is not necessarily true that each of the PIP constraint in (12) at any of the fading states is inactive because the latter is a more stringent one than the former. By the CS condition (9), it follows that $\nu^{(1)}$ is zero in (8). Comparing this new power allocation with (13) for the PIP case, it is easy to see that the latter is also more restrictive than the former due to the additional constraint $\frac{\Gamma_p}{g_i}$ at each fading state i.

At last, we provide numerical examples to evaluate the ergodic capacities of the secondary link under the AIP and PIP constraint at the PR. It is assumed that $\Gamma_a = \Gamma_p = 1$, the same as the additive noise power at the PR. It is also assumed that $\boldsymbol{h}$ and $\boldsymbol{g}$ are the squared norms of independent CSCG variables $\sim \mathcal{CN}(0,1)$ and $\mathcal{CN}(0, \sigma_g^2)$, respectively. Fig. 2 and Fig. 3 show the ergodic capacities versus the secondary user own transmit power constraint P for $\sigma_g^2 = 1$ and $\sigma_g^2 = 0.1$, respectively. It is observed that as expected, under the same constraint-power value, the AIP constraint results in larger capacities than the PIP due to the more flexible transmit power adaptations over the fading states. The capacity gain by the AIP over the PIP becomes more substantial as the interfering channel average power σ_g^2 increases. It is also observed that as P increases, the ergodic capacity of the secondary user channel eventually gets saturated under either the AIP or the PIP constraint.

IV. PRIMARY LINK CAPACITIES UNDER AIP AND PIP

From the secondary user's perspective, we have shown that the AIP constraint is more favorable than the PIP for the achievable ergodic capacity of the secondary user fading channel. In this section, we

will compare their effects on the capacities of the primary user fading channel. For fair comparison, we consider the primary user transmission subject to the same peak or average interference-power value Γ from the secondary transmitter, i.e., $\Gamma_a = \Gamma_p = \Gamma$. Furthermore, we assume that both the AIP and PIP constraints are satisfied with equalities, i.e., for the AIP case, $\mathbb{E}[I_i] = \Gamma$; and for the PIP case, $I_i = \Gamma, \forall i$. Next, we consider three well-known power-control policies in literature for the primary transmission and examine the resultant capacities under the AIP and PIP constraints in each case.

A. Constant-Power (CP) Power Control

First, we consider the *constant-power* (CP) power-control policy for $\mathcal{P}_p(\boldsymbol{f}, \boldsymbol{I})$, i.e.,

$$q_i = Q, \ \forall i. \tag{15}$$

CP is an attractive scheme in practice from an implementation viewpoint because it does not require any CSI at the PT, and furthermore, it satisfies the peak transmit-power-constraint at all times. With CP, the ergodic capacity of the primary user channel in the AIP case can be obtained from (1) as

$$C_{p,a}^{\mathrm{ER,CP}} = \mathbb{E}\left[\log\left(1 + \frac{f_i Q}{1 + I_i}\right)\right], \tag{16}$$

and the ergodic capacity under the PIP is

$$C_{p,p}^{\mathrm{ER,CP}} = \mathbb{E}\left[\log\left(1 + \frac{f_i Q}{1 + \Gamma}\right)\right]. \tag{17}$$

The following theorem is then established:

Theorem 4.1: For the CP power-control policy and under the same peak and average interference-power value Γ, $C_{p,a}^{\mathrm{ER,CP}} \geq C_{p,p}^{\mathrm{ER,CP}}$ for the primary user fading channel.

Proof:

$$
\begin{aligned}
C_{p,a}^{\mathrm{ER,CP}} &= \mathbb{E}_f \mathbb{E}_I\left[\log\left(1 + \frac{f_i Q}{1 + I_i}\right)\right] & (18)\\
&\geq \mathbb{E}_f\left[\log\left(1 + \frac{f_i Q}{1 + \mathbb{E}[I_i]}\right)\right] & (19)\\
&= \mathbb{E}_f\left[\log\left(1 + \frac{f_i Q}{1 + \Gamma}\right)\right] & (20)\\
&= C_{p,p}^{\mathrm{ER,CP}}, & (21)
\end{aligned}
$$

where the first equality is due to the independence of f_i and g_i and thus I_i; the inequality is due to the convexity of the function $f(x) = \log\left(1 + \frac{a}{1+x}\right)$ where a is any positive constant and $x \geq 0$, and the Jensen's inequality (e.g., [8]); and the last two equalities are by $\mathbb{E}[I_i] = \Gamma$ and (17), respectively. ■

Theorem 4.1 suggests that, surprisingly, the AIP constraint that results in the randomized interference powers over the fading states at the PR is in fact more advantageous for improving the primary user channel ergodic capacity as compared to the PIP constraint that has the constant interference powers, provided that their average and peak power-constraint values are both equal to Γ. As shown in the above proof, this result is mainly due to the convexity of the capacity function with respect to the noise/interference power. We name this interesting phenomenon as the "interference diversity".

B. Channel-Inversion (CI) Power Control

There are circumstances where the primary user transmission needs to deliver data traffic that has stringent delay requirements. For such scenarios, the delay-limited capacity [9] that is defined as the maximum constant-rate achievable for all the fading states given a long-term average-transmit-power budget can be a relevant measure for the channel capacity limit. For the SISO fading primary user channel, the optimal power-control policy that achieves the fading channel delay-limited capacity is known as the "channel-inversion (CI)" [10], which in the AIP case can be expressed as

$$q_i = \frac{\gamma_a(1 + I_i)}{f_i}, \tag{22}$$

and in the PIP case is expressed as

$$q_i = \frac{\gamma_p(1 + \Gamma)}{f_i}, \tag{23}$$

where γ_a and γ_p are the constant signal-to-interference-plus-noise-ratio (SINR) at the PR for the AIP and PIP case, respectively. Given $\mathbb{E}[q_i] \leq Q$, the maximum value of γ_a and γ_p can be obtained from (22) and (23) as

$$\gamma_a^{\max} = \frac{Q}{\mathbb{E}\left[\frac{1+I_i}{f_i}\right]}, \tag{24}$$

and

$$\gamma_p^{\max} = \frac{Q}{(1 + \Gamma)\mathbb{E}\left[\frac{1}{f_i}\right]}, \tag{25}$$

respectively. Since f_i is independent of I_i, we have

$$\mathbb{E}\left[\frac{1+I_i}{f_i}\right] = \mathbb{E}_f\left[\frac{1+\mathbb{E}[I_i]}{f_i}\right] = (1+\Gamma)\mathbb{E}\left[\frac{1}{f_i}\right],$$

and thus it follows from (24) and (25) that $\gamma_a^{\max} = \gamma_p^{\max}$. Hence, we obtain the following theorem:

Theorem 4.2: For the CI policy and under the same peak and average interference-power value Γ, the AIP and PIP constraints result in the same primary user channel delay-limited capacity equal to

$$C_p^{\mathrm{DL}} = \log\left(1 + \frac{Q}{(1+\Gamma)\mathbb{E}\left[\frac{1}{f_i}\right]}\right). \tag{26}$$

Proof: Substituting (22) and (23) into (1) yields the achievable constant-rates of the primary user under the AIP and PIP are $\log(1+\gamma_a)$ and $\log(1+\gamma_p)$, respectively. Since the channel delay-limited capacity is achievable by the maximum value of γ_a and γ_p for the AIP and PIP case, respectively, using the fact that $\gamma_a^{\max} = \gamma_p^{\max}$ and (25), it follows that (26) holds for both the AIP and PIP cases. ∎

Theorem 4.2 suggests that for the CI policy, the loss of the primary user channel delay-limited capacity due to the randomized interference powers from the secondary transmitter is identical to that by the constant-power interferences, i.e., the AIP is at least no worse than the PIP from the primary user's perspective for delivering delay-limited and constant-rate data traffic.

C. Water-Filling (WF) Power Control

The "water-filling (WF)" power-control policy (e.g., [8], [10]) is designed to achieve the maximum primary user channel ergodic capacity. In the case of the AIP, the optimal power allocations can be obtained as the solutions of the following optimization problem (**P3**):

$$\begin{aligned} \underset{\{q_i\}}{\text{Maximize}} \quad & \mathbb{E}\left[\log\left(1 + \frac{f_i q_i}{1 + I_i}\right)\right] \\ \text{Subject to} \quad & \mathbb{E}[q_i] \leq Q, \tag{27} \\ & q_i \geq 0. \ \forall i. \tag{28} \end{aligned}$$

Similar like P1, the optimal WF solutions of P3 can be obtained as

$$q_i^{(3)} = \left(\frac{1}{\mu^{(3)}} - \frac{1+I_i}{f_i}\right)^+, \tag{29}$$

where $\mu^{(3)}$ is the optimal dual solution associated with the constraint (27), which controls the so-called "water-level" $\frac{1}{\mu^{(3)}}$ with which $\mathbb{E}[q_i^{(3)}] = Q$. Substituting (29) into the objective function of P3 yields the ergodic capacity in the AIP case expressed as

$$C_{p,a}^{\text{ER,WF}} = \mathbb{E}\left[\left(\log\left(\frac{f_i}{\mu^{(3)}(1 + I_i)}\right)\right)^+\right]. \tag{30}$$

Similarly, we can define the ergodic capacity maximization problem for the PIP case, and refer to this problem as **P4**. The optimal WF-based power allocations in this case can be obtained as

$$q_i^{(4)} = \left(\frac{1}{\mu^{(4)}} - \frac{1 + \Gamma}{f_i}\right)^+, \tag{31}$$

where $\mu^{(4)}$ controls the "water-level" with which $\mathbb{E}[q_i^{(4)}] = Q$. The resultant ergodic capacity in the PIP case then becomes

$$C_{p,p}^{\text{ER,WF}} = \mathbb{E}\left[\left(\log\left(\frac{f_i}{\mu^{(4)}(1 + \Gamma)}\right)\right)^+\right]. \tag{32}$$

The direct comparison of $C_{p,a}^{\text{ER,WF}}$ in (30) and $C_{p,p}^{\text{ER,WF}}$ in (32) may be problematic explained as follows. Supposing that $\mu^{(3)} \leq \mu^{(4)}$, from the convexity of the function $g(x) = \left(\log\left(\frac{a}{1+x}\right)\right)^+$ where a is a positive constant and $x \geq 0$, and similar like the proof of Theorem 4.1, it can be shown that $C_{p,a}^{\text{ER,WF}} \geq C_{p,p}^{\text{ER,WF}}$. Unfortunately, the following lemma shows that the opposite inequality is indeed true for $\mu^{(3)}$ and $\mu^{(4)}$, and thus renders the direct comparison of the expressions of $C_{p,a}^{\text{ER,WF}}$ and $C_{p,p}^{\text{ER,WF}}$ difficult to proceed.

Lemma 4.1: For the WF-based primary user power allocations under the same peak and average interference-power Γ from the secondary transmitter, the optimal water-level parameters for the AIP and PIP constraints satisfy that $\mu^{(3)} \geq \mu^{(4)}$.

Proof: Supposing that $\mu^{(3)} < \mu^{(4)}$, we then have

$$\mathbb{E}[q_i^{(3)}] = \mathbb{E}\left[\left(\frac{1}{\mu^{(3)}} - \frac{1+I_i}{f_i}\right)^+\right] \tag{33}$$

$$> \mathbb{E}\left[\left(\frac{1}{\mu^{(4)}} - \frac{1+I_i}{f_i}\right)^+\right] \tag{34}$$

$$= \mathbb{E}_f\mathbb{E}_I\left[\left(\frac{1}{\mu^{(4)}} - \frac{1+I_i}{f_i}\right)^+\right] \tag{35}$$

$$\geq \mathbb{E}_f\left[\left(\frac{1}{\mu^{(4)}} - \frac{1+\mathbb{E}[I_i]}{f_i}\right)^+\right] \tag{36}$$

$$= \mathbb{E}\left[\left(\frac{1}{\mu^{(4)}} - \frac{1+\Gamma}{f_i}\right)^+\right] \tag{37}$$

$$= \mathbb{E}[q_i^{(4)}] \tag{38}$$

$$= Q, \tag{39}$$

where (36) is due to the convexity of the function $z(x) = \left(a - \frac{1+x}{b}\right)^+$ where a, b are positive constants and $x \geq 0$, and the Jensen's inequality. However, since it is known that $\mathbb{E}[q_i^{(3)}] = Q$, which contradicts with the inequality $\mathbb{E}[q_i^{(3)}] > Q$ shown above under the presumption that $\mu^{(3)} < \mu^{(4)}$, it thus concludes that this presumption is untrue and $\mu^{(3)} \geq \mu^{(4)}$ must hold. ∎

From the above discussions, we have to resort to an alternative approach for comparing $C_{p,a}^{\mathrm{ER,WF}}$ and $C_{p,p}^{\mathrm{ER,WF}}$. Here, we provide directly the result for which the proof is provided in the appendix of the paper.

Theorem 4.3: For the WF power-control policy and under the same peak and average interference-power value Γ, $C_{p,a}^{\mathrm{ER,WF}} \geq C_{p,p}^{\mathrm{ER,WF}}$ for the primary user fading channel.

Theorem 4.3 then suggests that for the WF-based power control, the randomized interference powers from the secondary transmission in the AIP case is again superior than the constant interference powers in the PIP case for maximizing the primary user channel ergodic capacity.

D. Numerical Examples

At last, we provide numerical examples to demonstrate the effect of the interference diversity on the capacities of the primary user fading channel. Since it has been shown that the delay-limited capacity of the primary user fading channel is identical under both the AIP and PIP constraint, we consider

here its ergodic capacities by the CP and WF power-control policies. For fair comparison, in the PIP case, it is assumed that $I_i = \Gamma_a = \Gamma, \forall i$, while in the AIP case, I_i's are taken as the squared norms of independent CSCG variables each $\sim (0, \Gamma_p)$, and $\Gamma_p = \Gamma$. It is assumed that f_i's are the squared norms of independent CSCG variables each $\sim \mathcal{CN}(0, 1)$. Fig. 4 and Fig. 5 show the ergodic capacities versus the primary user own transmit power constraint Q for $\Gamma = 10$ and $\Gamma = 1$, respectively. It is observed that unlike the secondary link, the ergodic capacity of the primary user link does not saturate as Q increases due to the protection by the AIP or PIP constraint. As expected, the resultant ergodic capacity under the AIP is always larger than that under the PIP for both CP and WF policies, thanks to the interference diversity. It is also observed that the capacity gains by the AIP over the PIP become more significant as Γ increases.

V. ACHIEVABLE THROUGHPUT UNDER SPECTRUM SHARING OVER FADING CHANNEL

So far, we have studied the effect of the AIP and PIP constraints on the capacities of the secondary and primary link separately. In this section, we will consider a realistic spectrum sharing scenario over the fading channel, and evaluate by simulation the jointly achievable throughput of both the primary and secondary links. It is assumed that $\Gamma_a = \Gamma_p = \Gamma = 1$. It is also assumed that $\boldsymbol{h}$, $\boldsymbol{g}$ and $\boldsymbol{f}$ are the squared norms of independent CSCG variables each $\sim \mathcal{CN}(0, 1)$. We consider the ergodic capacities for both the primary and secondary links, and for the primary link, the WF power-control policy is used. Fig. 6 and Fig. 7 show the jointly achievable ergodic capacities of the primary and secondary links versus the secondary user transmit power constrain P for the primary user transmit power constraint $Q = 10$ and $Q = 1$, respectively. Notice that the secondary user link ergodic capacities in both figures are identical.

First, it is observed that as the secondary user transmit power constraint P increases, its actual average transmit power as well as the ergodic capacity also increase until the AIP or PIP constraint becomes tightened. On the other hand, for some given fixed primary user transmit power constraint Q, the primary user channel ergodic capacities in both the AIP and PIP cases decrease as P increases and get saturated finally as the AIP or PIP constraint becomes tightened.

Secondly, the capacity gains by the AIP over the PIP for both the primary and secondary links are observed in most cases, while in Fig.6 for some low values of P, it is however observed that the

primary link capacity under the PIP can be larger than that under the AIP. Notice that this result does not contradict with Theorem 4.3. This is because for Theorem 4.3, we have assumed that the interference power I_i at the PR at any fading state i in the PIP case is constantly equal to Γ, while in the present simulation, at low P it may be possible that the resultant I_i for some i is strictly less than Γ. Moreover, the resultant AIP value $\mathbb{E}[I_i]$ in such cases is also strictly smaller than Γ. If we instead use a more restrictive PIP value equal to that of the actual resultant AIP in each of these cases, the ergodic capacity under the AIP will be again larger than that under the PIP, as is consistent with Theorem 4.3.

At last, it is observed that for the primary user link, as compared to Fig. 5 with the same value $\Gamma = 1$, the resultant ergodic capacity gain under the AIP over the PIP is slightly larger in the present simulation. This is because for Fig. 5, we have assumed that I_i's are obtained as the squared norms of independent CSCG variables while in the present simulation for a realistic spectrum sharing scenario, the resultant I_i's will have a different distribution due to the WF power-control over the secondary link.

VI. CONCLUSION

This paper studies the fundamental capacity limits for wireless spectrum sharing in the CR networks where the secondary users apply the interference-power/interference-temperature constraint at each primary receiver as a practical meas to protect the primary user transmissions. On the contrary to the traditional viewpoint that the peak-interference-power (PIP) constraint protects better the primary transmission than the average-interference-power (AIP) constraint given their same constraint-power value, this paper shows that the AIP constraint can be in many cases more advantageous for minimizing the resultant capacity losses of the primary user fading channel. This is mainly owing to an interesting interference diversity phenomenon discovered in this paper. This paper thus provides an important design rule for the CR networks in practice, i.e., the AIP constraint should be used for both the purposes of protecting the primary user transmission as well as maximizing the CR link throughput.

This paper assumes that the perfect CSI on the interfering channel from the secondary transmitter to the primary receiver is available at the secondary transmitter for any of the fading states. In practice, it is usually more valid to assume availability of only the statistical channel knowledge. The definition of the AIP constraint in this paper can be extendible to such cases. Furthermore, this paper considers the

fading primary and secondary user channels, but more generally, the results obtained also apply to other channel models consisting of parallel Gaussian channels over which the average and peak interference-power constraints are applicable, e.g., the frequency-selective slow-fading broadband channel that is decomposable into parallel narrow-band channels at each fading state by the well-known orthogonal-frequency-division-multiplexing (OFDM) modulation/demodulation method.

APPENDIX

The proof of Theorem 4.3 is based on the Lagrange dualities of the capacity maximization problems P3 and P4. First, we can rewrite $C_{p,a}^{\text{ER,WF}}$ and $C_{p,p}^{\text{ER,WF}}$ as the solutions of the following min-max optimization problems, respectively:

$$C_{p,a}^{\text{ER,WF}} = \min_{\mu:\mu\geq 0} \max_{\{q_i\}:q_i\geq 0,\forall i} \mathbb{E}\left[\log\left(1 + \frac{f_i q_i}{1 + I_i}\right)\right] - \mu(\mathbb{E}[q_i] - Q), \tag{40}$$

and

$$C_{p,p}^{\text{ER,WF}} = \min_{\mu:\mu\geq 0} \max_{\{q_i\}:q_i\geq 0,\forall i} \mathbb{E}\left[\log\left(1 + \frac{f_i q_i}{1 + \Gamma}\right)\right] - \mu(\mathbb{E}[q_i] - Q). \tag{41}$$

We thus have the following inequalities/equalities:

$$C_{p,p}^{\text{ER}} = \min_{\mu:\mu\geq 0} \mathbb{E}\left[\left(\log\left(\frac{f_i}{(1+\Gamma)\mu}\right)\right)^+\right] - \mathbb{E}\left[\left(1 - \frac{(1+\Gamma)\mu}{f_i}\right)^+\right] + \mu Q \tag{42}$$

$$\leq \mathbb{E}\left[\left(\log\left(\frac{f_i}{(1+\Gamma)\mu^{(3)}}\right)\right)^+\right] - \mathbb{E}\left[\left(1 - \frac{(1+\Gamma)\mu^{(3)}}{f_i}\right)^+\right] + \mu^{(3)} Q \tag{43}$$

$$= \mathbb{E}_f\left[\left(\log\left(\frac{f_i}{(1+\mathbb{E}[I_i])\mu^{(3)}}\right)\right)^+\right] - \mathbb{E}_f\left[\left(1 - \frac{(1+\mathbb{E}[I_i])\mu^{(3)}}{f_i}\right)^+\right] + \mu^{(3)} Q \tag{44}$$

$$\leq \mathbb{E}_f\mathbb{E}_I\left[\left(\log\left(\frac{f_i}{(1+I_i)\mu^{(3)}}\right)\right)^+\right] - \mathbb{E}_f\mathbb{E}_I\left[\left(1 - \frac{(1+I_i)\mu^{(3)}}{f_i}\right)^+\right] + \mu^{(3)} Q \tag{45}$$

$$= \mathbb{E}\left[\left(\log\left(\frac{f_i}{(1+I_i)\mu^{(3)}}\right)\right)^+\right] - \mathbb{E}\left[\left(1 - \frac{(1+I_i)\mu^{(3)}}{f_i}\right)^+\right] + \mu^{(3)} Q \tag{46}$$

$$= C_{p,a}^{\text{ER}}, \tag{47}$$

where (42) is obtained by substituting the optimal power allocations in (31) (replacing the optimal dual solution $\mu^{(4)}$ by any arbitrary dual variable μ) into (41); (43) is due to the fact that $\mu^{(3)}$ is not the optimal dual solution $\mu^{(4)}$ that minimizes the function in (42); (44) is because of $\mathbb{E}[I_i] = \Gamma$; (45) is because of the convexity of the function in (45) with respect to I_i for any arbitrary fixed f_i and the Jensen's

inequality; (46) is due to the independence of f_i and I_i; (47) is due to the fact that $\mu^{(3)}$ and $\{q_i^{(3)}\}$ in (29) are respectively the optimal dual and primal solution of the min-max optimization problem given in (40).

REFERENCES

[1] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201-220, Feb. 2005.

[2] M. Gastpar, "On capacity under receive and spatial spectrum-sharing constraints," *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 471-487, Feb. 2007.

[3] A. Ghasemi and E. S. Sousa, "Fundamental limits of spectrum-sharing in fading environments," *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 649-658, Feb. 2007.

[4] R. Zhang, "Optimal power control over fading cognitive radio channels by exploiting primary user CSI", *submitted to IEEE Global Communications Conference (Globecom)*. Available at http://arxiv.org/abs/0804.1617v1

[5] R. Zhang and Y. C. Liang, "Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks," *IEEE J. S. Topics Sig. Proc.*, vol. 2, no. 1, pp. 88-102, Feb. 2008.

[6] S. Boyd and L. Vandenberghe, *Convex optimization,* Cambridge University Press, 2004.

[7] R. G. Bland, D. Goldfarb, and M. J. Todd, "The ellipsoid method: A survey," *Operations Research,* vol. 29, no. 6, pp. 1039-1091, 1981.

[8] T. Cover and J. Thomas, *Elements of Information Theory,* New York: Wiley, 1991.

[9] S. Hanly and D. Tse, "Multi-access fading channels-Part II: Delay-limited capacities," *IEEE Trans. Inf. Theory*, vol. 44, no.7, pp. 2816-2831, Nov. 1998.

[10] A. 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.
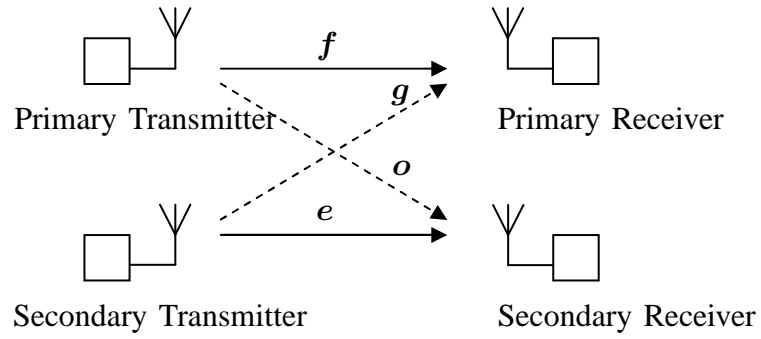
Fig. 1. Wireless spectrum sharing with a single primary and a single secondary/CR transmission link.
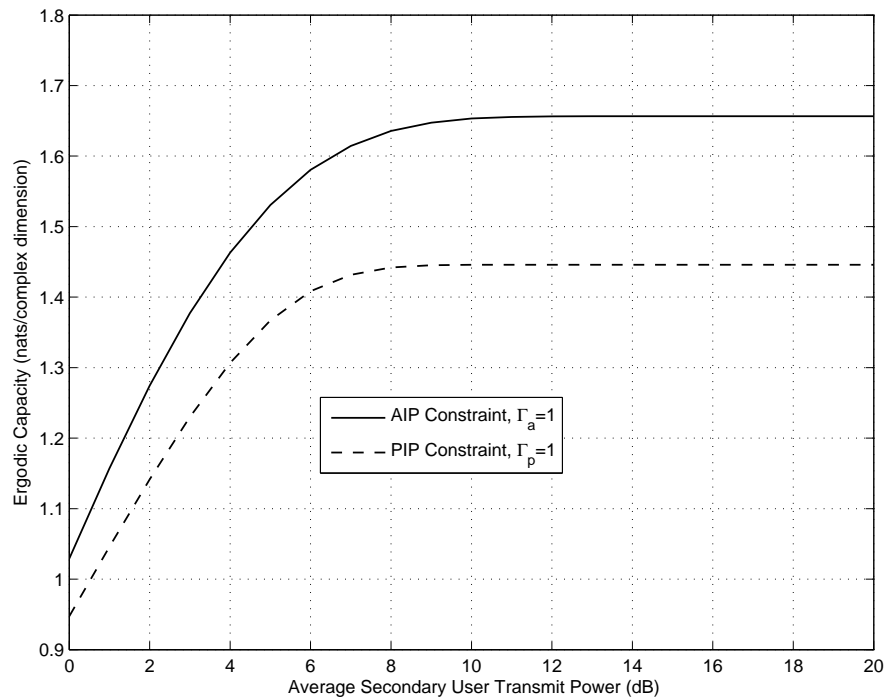

Fig. 2. Comparison of the secondary user channel ergodic capacity under the AIP versus the PIP constraint, $\sigma_g^2 = 1$.
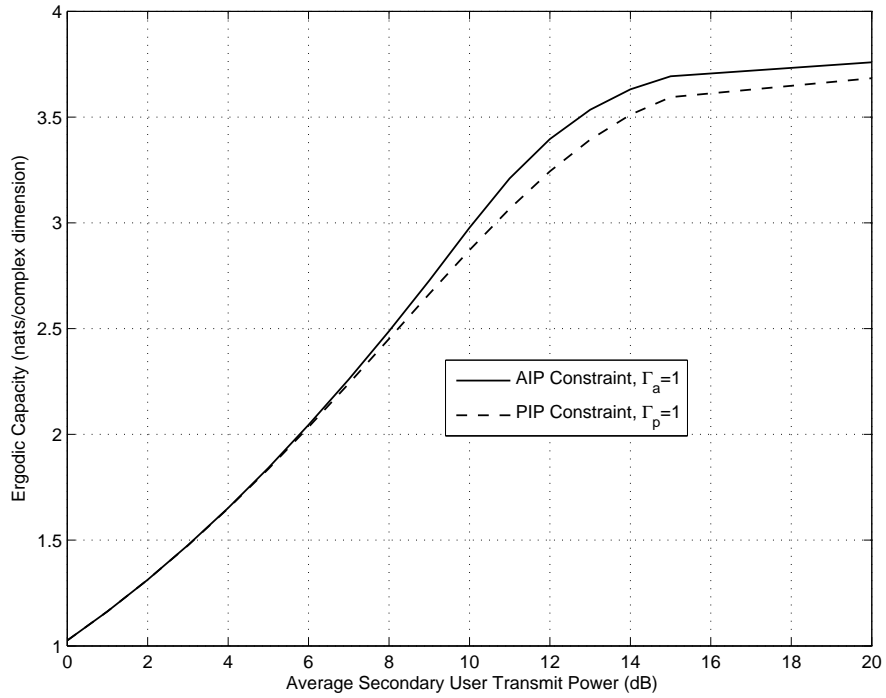
Fig. 3. Comparison of the secondary user channel ergodic capacity under the AIP versus the PIP constraint, $\sigma_g^2 = 0.1$.
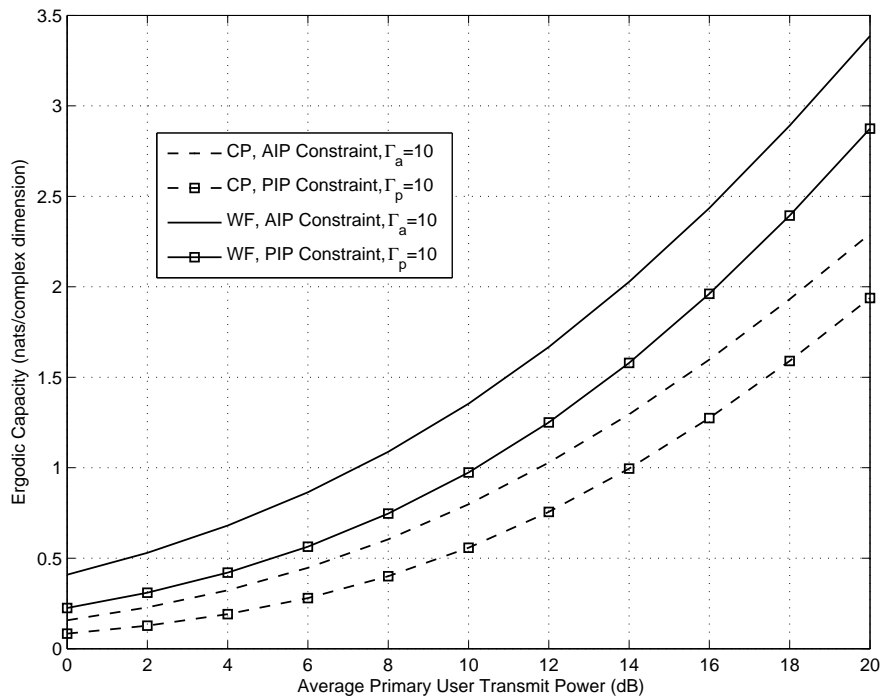

Fig. 4. Comparison of the primary user channel ergodic capacity under the AIP versus the PIP constraint, $\Gamma_a = \Gamma_p = \Gamma = 10$.
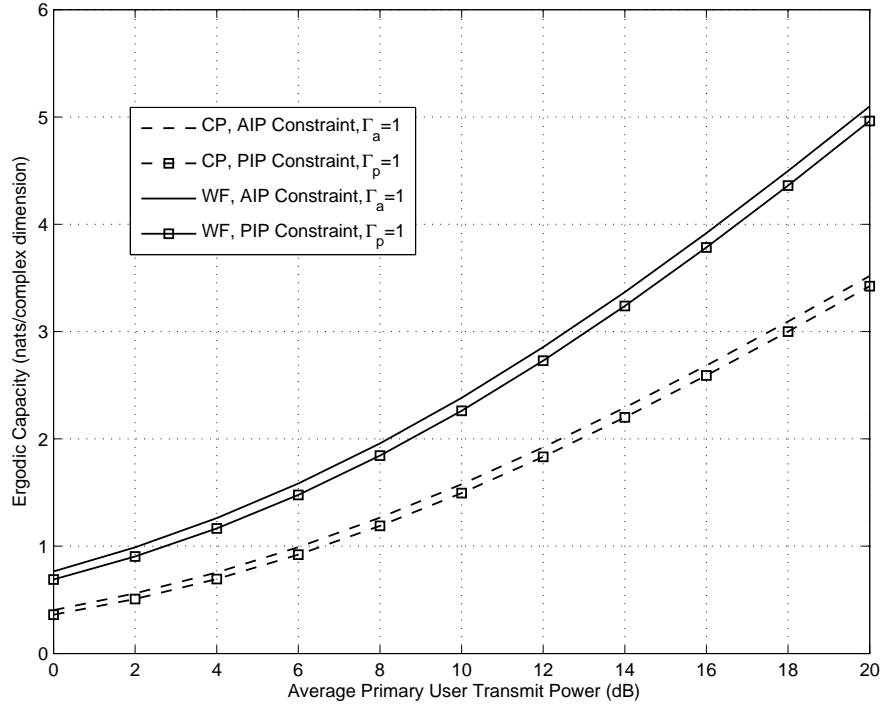
Fig. 5. Comparison of the primary user channel ergodic capacity under the AIP versus the PIP constraint, $\Gamma_a = \Gamma_p = \Gamma = 1$.
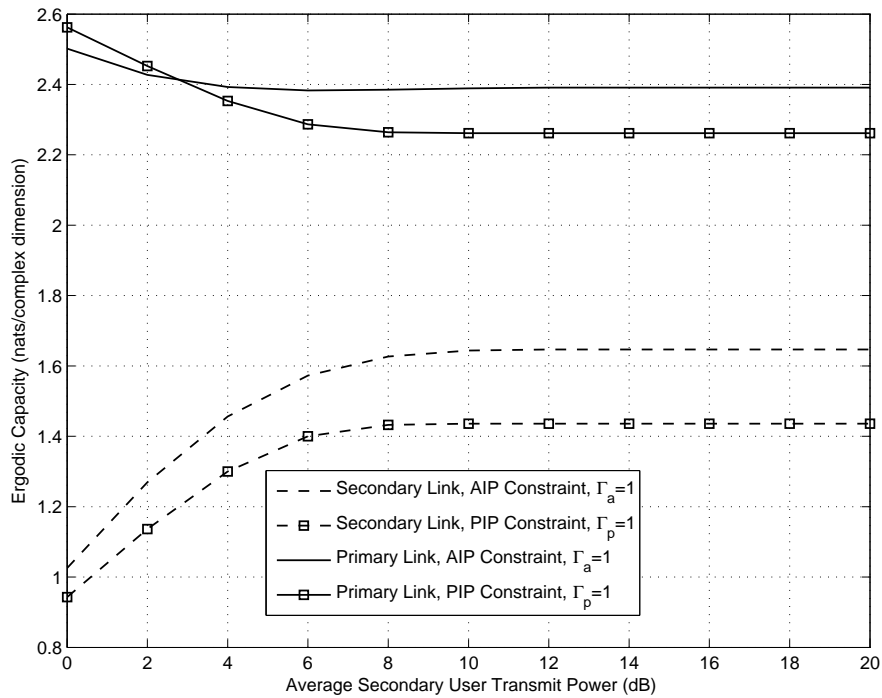

Fig. 6. Comparison of the primary and secondary link ergodic capacities under the AIP versus the PIP constraint, $Q = 10$.
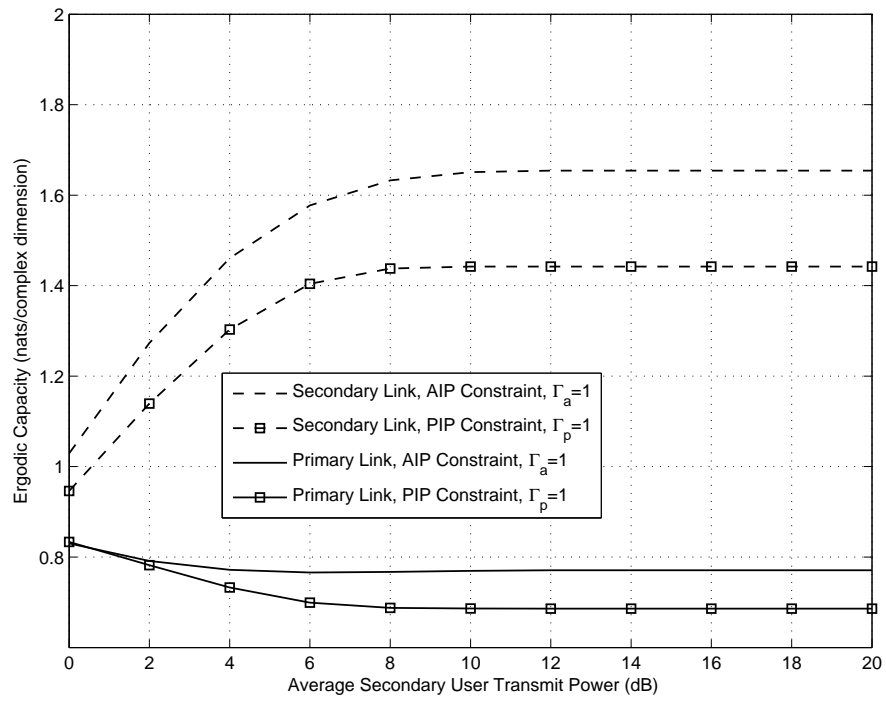
Fig. 7.   Comparison of the primary and secondary link ergodic capacities under the AIP versus the PIP constraint, $Q = 1$.