

Power Control and Asymptotic Throughput Analysis for the Distributed Cognitive Uplink

Ehsan Nekouei, *Student Member, IEEE*, Hazer Inaltekin, *Member, IEEE*,
and Subhrakanti Dey, *Senior Member, IEEE*

Abstract—This paper studies optimum power control and sum-rate scaling laws for the distributed cognitive uplink. It is first shown that the optimum distributed power control policy is in the form of a threshold based water-filling power control. Each secondary user executes the derived power control policy in a distributed fashion by using local knowledge of its direct and interference channel gains such that the resulting aggregate (average) interference does not disrupt primary’s communication. Then, the tight sum-rate scaling laws are derived as a function of the number of secondary users N under the optimum distributed power control policy. The fading models considered to derive sum-rate scaling laws are general enough to include Rayleigh, Rician and Nakagami fading models as special cases. When transmissions of secondary users are limited by both transmission and interference power constraints, it is shown that the secondary network sum-rate scales according to $\frac{1}{e^{n_h}} \log \log(N)$, where n_h is a parameter obtained from the distribution of direct channel power gains. For the case of transmissions limited only by interference constraints, on the other hand, the secondary network sum-rate scales according to $\frac{1}{e^{\gamma_g}} \log(N)$, where γ_g is a parameter obtained from the distribution of interference channel power gains. These results indicate that the distributed cognitive uplink is able to achieve throughput scaling behavior similar to that of the centralized cognitive uplink up to a pre-log multiplier $\frac{1}{e}$, whilst primary’s quality-of-service requirements are met. The factor $\frac{1}{e}$ can be interpreted as the cost of distributed implementation of the cognitive uplink.

Index Terms—Cognitive radio, multiple access channels, power control, throughput scaling, distributed algorithms.

I. INTRODUCTION

A. Background and Motivation

COGNITIVE radio technology has recently emerged as an aspirant solution for the problem of spectrum scarcity [1]-[3]. Unlike the traditional static command-and-control approach, it provides a more dynamic means for spectrum management and utilization. More specifically, cognitive radio protocols such as those in IEEE 802.22 allow the cognitive

users, alternatively called secondary users (SUs), to dynamically share the underutilized frequency bands with primary users (PUs) both in time and space under various forms of primary quality-of-service (QoS) protections [4]-[8]. In practice, channel state information (CSI) is one of the main requisites for successful implementation of such dynamic cognitive radio protocols. However, its availability is often sidelined in most previous works [9]-[19] by either assuming a centralized band manager or perfect instantaneous CSI feedback between primary and secondary networks.

For example, both interference management and resource allocation tasks in cognitive radio networks heavily depend on the availability of CSI at the secondary network. This requirement is especially more pronounced for multiuser cognitive radio networks. The assumption of the existence of a *centralized* entity having *global* knowledge of CSI may not be realistic in certain multiuser cognitive communication scenarios, depending on the physical characteristics of wireless channels, infrastructure limitations, number of SUs and etc. In these cases, distributed utilization of CSI is the key for successful implementation of cognitive radio protocols. To this end, the current paper explores the design of optimum distributed power control mechanisms for the cognitive uplink, allowing each SU to adjust its transmission power level based *only* on local knowledge of its CSI. It also investigates multiuser diversity gains for the distributed cognitive uplink by deriving tight sum-rate capacity scaling laws under the optimum distributed power control mechanisms.

In the centralized uplink, the secondary base-station (SBS) is primarily responsible for the power control task *e.g.*, see [10]-[13]. That is, it first acquires global knowledge of direct (from SUs to the SBS) and interference (from SUs to PUs) channel gains via a feedback mechanism, and then exploits this knowledge to obtain the optimum transmission power level for each SU by respecting primary QoS requirements. Finally, the allocated transmission power levels are broadcasted to SUs by the SBS at each fading block. Although required for the centralized power control at the cognitive uplink per above discussion, the assumption of availability of direct and interference channel gains at the SBS within the channel coherence time can often be too restrictive for practical cognitive multiple access networks consisting of large numbers of SUs.¹ On the other hand, unlike the centralized operation, SUs only need to have local access to their direct and interference channel gains in the distributed operating mode. Under the

¹The terms *cognitive uplink* and *cognitive multiple access network* are used interchangeably throughout the paper.

Manuscript received July 4, 2013; revised September 25, 2013. The editor coordinating the review of this paper and approving it for publication was E. Au.

This work was partially supported by the Scientific and Technological Research Council of Turkey (TUBITAK) under Grant 112E024.

E. Nekouei is with the Department of Electrical and Electronic Engineering, The University of Melbourne, VIC 3010, Australia (e-mail: e.nekouei@pgrad.unimelb.edu.au).

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

S. Dey was with the Department of Electrical and Electronic Engineering, The University of Melbourne, Australia. He is now with the Division of Signals & Systems, Uppsala University, Uppsala, SE 75121, Sweden (e-mail: Subhra.Dey@signal.uu.se).

Digital Object Identifier 10.1109/TCOMM.2013.112413.130510

assumption of uplink-downlink channel reciprocity (also used in [22] for distributed primary multiple access networks), it is easy for each SU to obtain local knowledge of its direct and interference channel gains using pilot signals transmitted periodically by the SBS and primary base-station (PBS)². These observations motivate the current paper, and lead to the following research questions of interest here: (i) what is the structure of optimum distributed power control mechanisms for the cognitive uplink?, and (ii) what are the fundamental throughput scaling laws of such decentralized cognitive multiple access networks under optimum power control subject to various forms of power and interference constraints?

This paper provides important insights into these questions by studying the optimum distributed power control mechanisms for the cognitive uplink that enable SUs to accomplish the power control task in a distributed fashion while the interference at the primary network is successfully regulated. More importantly, we evaluate the performance of our distributed power control mechanisms, in terms of the secondary network sum-rate, as the number of SUs becomes large. Our results signify the fact that distributed cognitive multiple access networks are capable of achieving throughput scaling behavior similar to that of centralized cognitive multiple access networks.

B. Contributions

This paper has two main contributions to the cognitive radio literature. First, we derive the structure of the optimum distributed power control policy, maximizing the secondary network sum-rate for two network types: (i) distributed total power and interference limited (DTPIL) networks and (ii) distributed interference limited (DIL) networks. In DTPIL networks, transmission powers of SUs are limited by a constraint on the average total transmission power of SUs and a constraint on the average interference power at the PBS. To confine the collision level, a transmission probability constraint is also considered for each SU. In DIL networks, transmission powers of SUs are limited by a constraint on the average interference power at the PBS as well as transmission probability constraints. For each network type, we show that the optimum distributed power control policy is in the form of a *threshold* based water-filling power control with changing water levels.

Secondly, we study the sum-rate scaling behavior of DTPIL and DIL networks, under the optimum distributed power control policy, when distributions of direct and interference channel gains belong to a fairly large class of distribution functions called class- \mathcal{C} distributions. In DTPIL networks, it is shown that the secondary network throughput scales according to $\frac{1}{e^{\gamma_g n_h}} \log \log(N)$ when the transmission probability is set to $\frac{1}{N}$ for all SUs. Here, N is the number of SUs, and n_h is a parameter obtained from the distribution of direct channel power gains. The choice of transmission probability adds an extra dimension to the optimization problems studied in this paper. To this end, we show that although $\frac{1}{N}$ may not be the

optimum transmission probability selection for the secondary network sum-rate maximization for finite values of N , it is *asymptotically* optimum in the sense that the same throughput scaling behavior holds even under the optimum transmission probability selection.

Analogous results are also obtained for DIL networks. In particular, it is shown that the secondary network sum-rate scales according to $\frac{1}{e^{\gamma_g}} \log(N)$ when the transmission probability is set to $\frac{1}{N}$ for all SUs, and the optimum distributed power control policy is employed. γ_g is a parameter obtained from the distribution of interference channel power gains. It is also shown that $\frac{1}{N}$ is the *asymptotically* optimum transmission probability selection for DIL networks, too. Unlike the known results for the Rayleigh fading model, our results here indicate that the pre-log factor does indeed heavily depend on the parameters of fading channel (direct and interference) distributions. From an engineering point of view, these results indicate that the optimum distributed power control at the cognitive uplink is capable of achieving aggregate data rates similar to those achieved through a centralized scheduler up to a pre-log multiplier $\frac{1}{e}$ [12]. Here, $\frac{1}{e}$ has the economic interpretation of the cost of avoiding feedback signals between primary and secondary networks. Our main results are summarized in Table I. The fundamentally different throughput scaling laws for the DTPIL and DIL network types also justify their separate consideration and analysis.

II. RELATED WORK

This section briefly reviews the papers that are most relevant to ours. In this paper, we are mainly motivated by exploiting distributed techniques for optimum resource/power allocation in the cognitive uplink and the corresponding sum-rate capacity scaling via multiuser diversity.

Optimum allocation of transmission powers in a cognitive radio setup has recently been investigated in [9]-[13]. In [9], Ghasemi and Sousa showed that the optimum power control maximizing the ergodic capacity of a point-to-point cognitive radio link under average interference power constraints is in the form of a water-filling power control policy with changing water levels. In [10], this result was extended to the cognitive uplink. Particularly, they showed that, under average transmission power and average interference power constraints, the optimum power allocation policy for a cognitive uplink is in the form of an opportunistic water-filling power allocation policy. That is, the SBS schedules the SU with the best joint direct and interference channel state, and the scheduled SU employs a water-filling power allocation policy for its transmission.

Similar results have also been obtained by considering total power and reduced CSI constraints in [11] and [12]. In [13], Inaltekin and Hanly established the binary structure of the optimum power control for the cognitive uplink operating under interference limitations without successive interference cancellation, *i.e.*, see Section VI of [13]. They showed that the set of transmitting SUs always corresponds to the ones having better joint channel states. Although the single-user decoding assumption in [13] simplifies the decoder, it complicates the power optimization problem. The resulting optimization problem, in contrast to the one in [9]-[12], is no longer convex.

²SUs can obtain the knowledge of their interference channel gains by overhearing primary's transmission [20], [21]. For example, when the primary network is operating at the time-division-duplex (TDD) mode, PBS transmits pilot signals assisting primary users to estimate their channel gains. SUs can use these pilot signals to estimate their interference channel gains.

TABLE I
THROUGHPUT SCALING BEHAVIOR OF DISTRIBUTED COGNITIVE RADIO NETWORKS

Network Model	Transmission probability	
	$p_N = \frac{1}{N}$	p_N^*
Distributed Total Power And Interference Limited	$\lim_{N \rightarrow \infty} \frac{R_N^\dagger}{\log \log(N)} = \frac{1}{e n_h}^\ddagger$	$\lim_{N \rightarrow \infty} \frac{R_N}{\log \log(N)} = \frac{1}{e n_h}$
Distributed Interference Limited	$\lim_{N \rightarrow \infty} \frac{R_N}{\log(N)} = \frac{1}{e \gamma_g}^\S$	$\lim_{N \rightarrow \infty} \frac{R_N}{\log(N)} = \frac{1}{e \gamma_g}$

* p_N^* is the optimum transmission probability.

† R_N is the secondary network sum-rate under the optimum distributed power control policy.

‡ n_h is parameter determined from the asymptotic tail behavior of the distribution of direct channel power gains.

§ γ_g is a parameter determined from the behavior of the distribution of interference channel power gains around the origin.

This paper differs from above previous work in two important aspects. Firstly, we focus on the distributed cognitive uplink in this paper, whereas [9]-[13] analyzed the centralized power control with perfect or reduced CSI at the SBS. The distributed operation requires contention control by constraining channel access probabilities, which in turn makes the studied power optimization problem here non-convex. Secondly, similar to these previous work, our analysis in this paper starts with the consideration of optimum power control policies. However, different from them, we also investigate multiuser diversity gains in the distributed cognitive uplink as a function of the number of SUs.

Multiuser diversity gains for the centralized cognitive radio networks with global knowledge of CSI at the SBS have also been studied in the literature extensively, *e.g.*, see [11], [14]-[17], under various types of constraints on the transmission powers of SUs. In [14], the authors established logarithmic and double-logarithmic throughput scaling behavior of the cognitive uplink for Rayleigh fading channels under joint peak transmission and interference power constraints. These results were extended to cognitive multiple access, cognitive broadcast and cognitive parallel access channels in [15]. The authors in [11], different from [14] and [15], considered average power limitations (both transmission and interference), and obtained parallel ergodic sum-rate scaling results for cognitive multiple access networks under optimum power control.

The main point of difference between this paper and above previous work is the utilization of more practical distributed approaches for the cognitive uplink here. Specifically, different from them, SUs in our setup independently decide to transmit (with power control) based on local knowledge of their CSI. This provides a more practical framework to study multiuser diversity in the cognitive uplink, but at the expense of a more complicated optimum power control analysis (*e.g.*, see Appendix A) and the corresponding estimates on the tails of joint channel states (*e.g.*, see Appendices B and E).

In [16], the scheduling gain in a cognitive uplink was considered for a hybrid scheduling policy under peak transmission and interference power constraints. All SUs transmit with the same fixed power level. Under this setup, it was shown that the secondary network throughput scales logarithmically (as a function of the number of SUs) with a pre-log factor depending on the number of PUs. Similar results

were extended to cognitive radio networks with multiple antennas at the SBS and PBS in [17]. They showed that the secondary network throughput scales logarithmically with a pre-log factor depending on the operating modes (*i.e.*, multiple access versus broadcast) and the number of antennas at the SBS and PBS.

Other related work also includes secondary network capacity scaling in a multi-band setup such as [18] and [19]. In [18], Wang *et al.* studied the multiuser and multi-spectrum diversity gains for a cognitive broadcast network sharing multiple orthogonal frequency bands with a primary network. Assuming Rayleigh fading channels, they analytically derived capacity expressions for the secondary network when the transmission power at each band is limited by a constraint on the peak interference power at the primary network. For a similar setup in [18], the authors in [19] considered N secondary transmitter-receiver pairs sharing M frequency bands with a primary network. Under the optimum matching of SUs with primary frequency bands, they derived a double-logarithmic scaling law for the secondary network capacity for Rayleigh fading channels. They also considered a contention-free distributed scheduling algorithm in which SUs decide to transmit (without any power control) if their received signal-to-interference-plus-noise-ratio in a frequency band is greater than a threshold level.

Unlike [16]-[19], this paper considers general fading models including Rayleigh fading as a special case (*i.e.*, see Table II). Further, all sum-rate scaling laws are derived for the contention-limited distributed cognitive uplink under optimum allocation of transmission powers to SUs, rather than assuming fixed transmission power levels as in [16]-[19]. The distributed power control mechanisms are designed as such they provide stringent QoS guarantees for the primary network under a collision channel model. Hence, some parts of our analysis in this paper are expected to find greater applicability to extend multiuser diversity results derived for multi-band and multi-antenna networks in [16]-[19] to fading models beyond Rayleigh fading and to more practical distributed communication scenarios with optimum resource allocation.

It is also important to note that multiuser diversity gains in distributed primary multiple access networks were also studied in the literature, *e.g.*, see [22]. However, these results are not applicable to the cognitive uplink as they do not account for

the impact of SUs' transmissions on the primary's QoS. What is needed in a cognitive setup is a more advanced distributed power management mechanism that can harvest multiuser diversity gains in both direct and interference channels simultaneously, whilst respecting primary's QoS requirements. This often results in solving non-convex optimization problems as in Section IV, and using more complicated techniques to obtain tail estimates of joint channel states as in Appendices B and E.

Finally, the authors in [23], a precursor to the current paper, studied the jointly optimum power allocation and scheduling policies, maximizing individual data rates, in a distributed cognitive uplink. They also studied the scaling behavior of the secondary network aggregate data rate under optimum resource allocation. The novelty of the current paper lies in deriving the optimum distributed power control maximizing the secondary network sum-rate, and the corresponding sum-rate scaling behavior for different network types. The asymptotic optimality of the transmission probability per user being chosen as $\frac{1}{N}$ is also an additional result that was not obtained in [23].

The rest of the paper is organized as follows. Section III describes our system model and modeling assumptions. Section IV derives the optimum distributed power control policies and their corresponding sum-rate scaling laws for DPIL and DIL networks. Section V presents our numerical studies. Section VI concludes the paper.

III. SYSTEM MODEL

We consider a cognitive uplink in which N SUs communicate with an SBS and simultaneously cause interference to a PBS as depicted in Fig. 1. Let h_i and g_i represent the i th direct and interference channel power gains, respectively. We consider the classical ergodic block fading model [24] to model the statistical variations of all direct and interference channel gains. $\{h_i\}_{i=1}^N$ and $\{g_i\}_{i=1}^N$ are assumed to be collections of i.i.d. random variables, with unit mean, distributed according to distribution functions F_h and F_g , respectively. F_h can be different than F_g . The random vectors $\mathbf{h} = [h_1, h_2, \dots, h_N]^\top$ and $\mathbf{g} = [g_1, g_2, \dots, g_N]^\top$ are assumed to be independent from each other. We assume that the uplink and downlink channels are reciprocal and each SU has access to its direct and interference channel gains by means of pilot training signals periodically transmitted by the SBS and PBS (*i.e.*, see [22], [25] for similar assumptions in the primary multiple-access networks). This type of channel reciprocity assumption is valid for wireless networks where time-division-duplexing (TDD) schemes are used. Note that due to various advantages of TDD such as (*i*) flexibility in controlling asymmetric uplink and downlink bandwidth (for wireless Internet services), (*ii*) more efficient use of the spectrum and (*iii*) reduced complexity of receiver design, TDD is the duplexing scheme of choice for the majority of future WiMAX based broadband wireless networks [29]. For similar reasons, TDD is expected to feature alongside frequency-division-duplexing (FDD) in the LTE based wireless networks as well. TDD transmission schemes are also a part of IEEE standards for cognitive radio networks, *e.g.*, IEEE 802.22.

Definition 3.1: We say that the cumulative distribution function (CDF) of a random variable X , denoted by F_X , belongs to the class- \mathcal{C} distributions if it satisfies the following properties:

- $F_X(x)$ is continuous.
- $F_X(x)$ has a positive support, *i.e.*, $F_X(x) = 0$ for $x \leq 0$.
- $F_X(x)$ is strictly increasing, *i.e.*, $F_X(x_1) < F_X(x_2)$ for $0 < x_1 < x_2$.
- The tail function of $F_X(x)$ defined as $1 - F_X(x)$ decays to zero *double exponentially*, *i.e.*, there exist constants $\alpha > 0$, $\beta > 0$, $n > 0$, $l \in \mathbb{R}$ and a slowly varying function $H(x)$ satisfying $H(x) = o(x^n)$ as $x \rightarrow \infty$ such that³

$$\lim_{x \rightarrow \infty} \frac{1 - F_X(x)}{\alpha x^l e^{(-\beta x^n + H(x))}} = 1. \quad (1)$$

- $F_X(x)$ varies *regularly* around the origin, *i.e.*, there exist constants $\eta > 0$ and $\gamma > 0$ such that

$$\lim_{x \downarrow 0} \frac{F_X(x)}{\eta x^\gamma} = 1.$$

The class- \mathcal{C} distribution functions comprise distribution functions that behave polynomially around the origin and their tail functions decay to zero double exponentially. This family of distribution functions can be used to quantify and obtain insights into the effects of fading model parameters on the capacity of different wireless networks. We assume that the CDFs of all fading power gains in this paper belong to the class- \mathcal{C} distributions. Table II illustrates the parameters characterizing the behavior of the distribution of fading power gains around zero and infinity for the commonly used fading models in the literature. To avoid any confusion, these parameters are represented by subscript h for direct channel gains and by subscript g for interference channel gains in the rest of paper.

In what follows, a wireless channel is said to be a Rayleigh fading channel if the channel magnitude gain is Rayleigh distributed, or equivalently the channel power gain is exponentially distributed. It is said to be Rician- K fading channel if the channel magnitude gain is Rician distributed with a Rician factor K . By a Nakagami- m distributed wireless fading channel, we mean the channel magnitude gain is Nakagami- m distributed, or equivalently the channel power gain is Gamma distributed. Finally, a wireless fading channel is said to be Weibull- c distributed if the channel magnitude gain is Weibull distributed with a Weibull parameter c .⁴ Interested readers are referred to [26], [27] and [28] for more details regarding fading distributions.

Each SU exploits knowledge of its direct and interference channel gains to *locally* perform the task of power allocation, *independent* of other SUs without any feedback from the SBS. A collision channel model is assumed for the resolution of concurrent transmissions from SUs at the SBS. That is, if more than one SUs transmit concurrently, data transmissions from all of them collide, and the resulting throughput at the SBS is set to zero. In the next section, we derive the structure of the optimum distributed power control policy maximizing

³By $p(x) = o(q(x))$, we mean that $p(x)$ and $q(x)$ are two positive functions such that $\lim_{x \rightarrow \infty} \frac{p(x)}{q(x)} = 0$.

⁴The definition of the c parameter for Weibull fading channels is adapted from [26].

TABLE II
COMMON FADING CHANNEL MODELS AND THEIR PARAMETERS

Channel Model	Parameters						
	α	l	β	n	$H(x)$	η	γ
Rayleigh	1	0	1	1	0	1	1
Rician- K	$\frac{1}{2\sqrt{\pi}e^{K/2}\sqrt{K(K+1)}}$	$-\frac{1}{4}$	$K+1$	1	$2\sqrt{K(K+1)}x$	$\frac{K+1}{e^K}$	1
Nakagami- m	$\frac{m^{m-1}}{\Gamma(m)}$	$m-1$	m	1	0	$\frac{m^{m-1}}{\Gamma(m)}$	m
Weibull- c	1	0	$\Gamma^{\frac{c}{2}}\left(1+\frac{2}{c}\right)$	$\frac{c}{2}$	0	$\Gamma^{\frac{c}{2}}\left(1+\frac{2}{c}\right)$	$\frac{c}{2}$

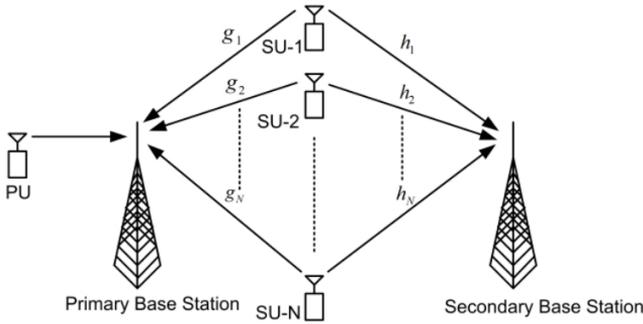


Fig. 1. N SUs forming a cognitive uplink to the SBS and interfering with signal reception at the PBS.

the secondary network sum-rate under the aforementioned assumptions for two different network types: (i) distributed total power and interference limited (DTPIL) networks, and (ii) distributed interference limited (DIL) networks. After obtaining the optimum distributed power control policy, we also derive throughput scaling laws for these network types when each SU controls its transmission power optimally.

IV. THE STRUCTURE OF THE OPTIMUM DISTRIBUTED POWER CONTROL POLICY AND THROUGHPUT SCALING LAWS

In this section, we will first present and solve the sum-rate maximization problems in DTPIL and DIL networks. Then, each problem will be followed by the corresponding throughput scaling results along with detailed insights into the observed throughput scaling behavior. All proofs are relegated to the appendices for the sake of paper fluency. We start our discussions by formulating the sum-rate maximization problem for DTPIL networks.

A. Optimum Power Control and Throughput Scaling in DTPIL Networks

In DTPIL networks, transmission powers of SUs are limited by an average total transmission power constraint and a constraint on the average total interference power of SUs at the PBS, *e.g.*, see [11], [30] and [31] for similar assumptions. Transmission probabilities of SUs are also constrained to avoid excessive collisions. The average interference power constraint is considered to protect long-term QoS at the PBS by keeping the average interference power of the secondary

network at the PBS below a predetermined level. The average total transmission power constraint confines the total transmission power of the secondary network as the number of SUs becomes large. Hence, our throughput scaling results in DTPIL networks reveal the fundamental capacity scaling limits of distributed cognitive radio uplinks in energy limited communication scenarios.

We define the power allocation policy in DTPIL networks, $P_{\text{DTPIL}}(\cdot, \cdot)$, as a mapping from \mathbb{R}_+^2 to \mathbb{R}_+ , where $P_{\text{DTPIL}}(h_i, g_i)$ represents the transmission power of the i th SU at the joint channel state (h_i, g_i) . The power allocation policy $P_{\text{DTPIL}}(\cdot, \cdot)$ is designed such that the transmission probability is equal to p_N , $p_N \in (0, 1)$, for all SUs, *i.e.*, $\Pr\{P_{\text{DTPIL}}(h_i, g_i) > 0\} = p_N$ for $i \in \{1, \dots, N\}$. Here, p_N can be considered as a design degree-of-freedom parameter helping us to keep the collision rate below some certain level. Under these modeling assumptions, the secondary network sum-rate for a given power control policy P_{DTPIL} , $R_{\text{DTPIL}}(p_N, N, P_{\text{DTPIL}})$, can be expressed as (2) where h and g are two independent generic random variables distributed according to F_h and F_g , respectively. Similarly, the average total transmission power and the average interference power at the PBS can be written as

$$\mathbb{E}_{h,g} \left[\sum_{i=1}^N P_{\text{DTPIL}}(h_i, g_i) \right] = N \mathbb{E}_{h,g} [P_{\text{DTPIL}}(h, g)]$$

and

$$\mathbb{E}_{h,g} \left[\sum_{i=1}^N g_i P_{\text{DTPIL}}(h_i, g_i) \right] = N \mathbb{E}_{h,g} [g P_{\text{DTPIL}}(h, g)]$$

respectively. In DTPIL networks, transmission powers of SUs are allocated according to the solution of the following functional optimization problem:

$$\begin{aligned} & \underset{P_{\text{DTPIL}}(h,g) \geq 0}{\text{maximize}} && R_{\text{DTPIL}}(p_N, N, P_{\text{DTPIL}}) \\ & \text{subject to} && N \mathbb{E}_{h,g} [P_{\text{DTPIL}}(h, g)] \leq P_{\text{ave}} \\ & && N \mathbb{E}_{h,g} [g P_{\text{DTPIL}}(h, g)] \leq Q_{\text{ave}} \\ & && \Pr\{P_{\text{DTPIL}}(h, g) > 0\} = p_N \end{aligned} \quad (3)$$

The power optimization problem in (3) is not necessarily convex due to the transmission probability constraint. However, in the next theorem, we show that the optimum power control policy solving (3) is in the form of a threshold based water-filling power control when the number of SUs is large enough. We prove Theorem 1 by obtaining the convex relaxation of problem (3) first, and then showing that the

$$\begin{aligned}
R_{\text{DTPIL}}(p_N, N, P_{\text{DTPIL}}) &= \mathbb{E}_{\mathbf{h}, \mathbf{g}} \left[\sum_{i=1}^N \log(1 + h_i P_{\text{DTPIL}}(h_i, g_i)) \prod_{j \neq i} \mathbf{1}_{\{P_{\text{DTPIL}}(h_j, g_j) = 0\}} \right] \\
&= N(1 - p_N)^{N-1} \mathbb{E}_{\mathbf{h}, \mathbf{g}} [\log(1 + h P_{\text{DTPIL}}(h, g))].
\end{aligned} \tag{2}$$

upper and lower bounds coincide with each other for N large enough.

Theorem 1: Let $P_{\text{DTPIL}}^*(h, g)$ be the solution of (3). Then, for $p_N = \frac{1}{N}$ and N large enough, $P_{\text{DTPIL}}^*(h, g)$ is given by (4) where λ_N and μ_N are power control parameters adjusted such that the average total transmission power and the average interference power constraints in (3) are met, and $F_{\lambda_N, \mu_N}^{-1}$ is the functional inverse of the CDF of $\frac{h_i}{\lambda_N + \mu_N g_i}$, which is shown as F_{λ_N, μ_N} .

Proof: See Appendix A. ■

Theorem 1 pinpoints that the jointly optimum scheduling and power control strategy is in the form of a threshold-based water-filling power control policy. That is, the i th SU first decides against or in favor of transmission by comparing the value of its observed joint direct and interference channel state $\frac{h_i}{\lambda_N + \mu_N g_i}$ with the threshold value of $F_{\lambda_N, \mu_N}^{-1}(1 - \frac{1}{N})$. Upon a positive decision in favor of transmission, it transmits by using a water-filling power allocation policy, which is embodied by the $(\frac{1}{\lambda_N + \mu_N g_i} - \frac{1}{h})^+$ term in (4). We note that the computation of λ_N and μ_N does not violate the distributed operation. These power control parameters do not vary with channel realizations, and can be computed off-line at each SU by solving the dual problem associated with the optimization problem (12) in Appendix A under the local statistical channel knowledge. This computation needs to be performed only once (for each N) to initialize the communication. Only global knowledge required here is N and p_N , which can be broadcasted by the SBS to SUs. The rest of the communication proceeds in a distributed manner according to the result of Theorem 1.

In the centralized case, direct and interference channel gains of all SUs are available at the SBS, and in order to maximize the secondary network sum-rate, the SBS schedules the SU having the maximum of $\left\{ \frac{h_i}{\lambda_N + \mu_N g_i} \right\}_{i=1}^N$ [11]. The scheduled SU employs a water-filling power allocation policy with changing power levels. Hence, the multiuser diversity gain with a centralized scheduler depends on the maximum of $\left\{ \frac{h_i}{\lambda_N + \mu_N g_i} \right\}_{i=1}^N$, which concentrates around $F_{\lambda_N, \mu_N}^{-1}(1 - \frac{1}{N})$ as the number of SUs becomes large (*i.e.*, see Lemma 2 in [12] for more details). Based on this observation and Theorem 1, we conclude that in DTPIL networks, the i th SU transmits if the likelihood of its being the SU with the maximum of $\left\{ \frac{h_i}{\lambda_N + \mu_N g_i} \right\}_{i=1}^N$ is high. Hence, throughput scaling laws similar to those obtained in [11] and [12] are expected to hold for DTPIL networks when $p_N = \frac{1}{N}$. Later, we show that this choice of transmission probability is asymptotically optimal. That is, the secondary network sum-rate under $p_N = \frac{1}{N}$ serves as an upper bound on aggregate communication rates that we would otherwise achieve through other choices of p_N when N is large enough.

Under the collision channel model, the SBS can decode

the received signal successfully *if and only if* just one SU transmits. Otherwise, a collision happens and no data is delivered to the SBS. In our setup with $p_N = \frac{1}{N}$ and N large enough, this observation implies that the received signal will be decoded successfully if and only if just the SU with the maximum of $\left\{ \frac{h_i}{\lambda_N + \mu_N g_i} \right\}_{i=1}^N$ transmits. Let $X_N^*(\lambda_N, \mu_N)$ and $X_N^\circ(\lambda_N, \mu_N)$ be the largest and the second largest elements among the collection of i.i.d. random variables $\{X_i(\lambda_N, \mu_N)\}_{i=1}^N$, respectively, where $X_i(\lambda_N, \mu_N) = \frac{h_i}{\lambda_N + \mu_N g_i}$. Let $R_{\text{DTPIL}}^*(p_N, N)$ be the sum-rate in DTPIL networks under the optimum distributed power control policy with the transmission probability equal to p_N . Then, Theorem 1 implies that, for $p_N = \frac{1}{N}$ and N large enough, we have

$$R_{\text{DTPIL}}^*\left(\frac{1}{N}, N\right) = \mathbb{E}_{\mathbf{h}, \mathbf{g}} [\log(X_N^*(\lambda_N, \mu_N)) \mathbf{1}_{A_N}], \tag{5}$$

where A_N is given by (6). In the next theorem, we derive the scaling behavior of $R_{\text{DTPIL}}^*\left(\frac{1}{N}, N\right)$.

Theorem 2: The secondary network sum-rate $R_{\text{DTPIL}}^*\left(\frac{1}{N}, N\right)$ for $p_N = \frac{1}{N}$ under the optimum distributed power control policy scales according to

$$\lim_{N \rightarrow \infty} \frac{R_{\text{DTPIL}}^*\left(\frac{1}{N}, N\right)}{\log \log(N)} = \frac{1}{en_h}.$$

Proof: See Appendix B. ■

Theorem 2 formally establishes the double logarithmic scaling behavior of the secondary network sum-rate for DTPIL networks. Further, it shows that the pre-log multiplier in this scaling behavior is equal to $\frac{1}{en_h}$. n_h is equal to 1 for Rayleigh, Rician- K and Nakagami- m distributed direct channel gains, and is equal to $\frac{c}{2}$ for Weibull- c distributed direct channel gains.

The result of Theorem 2 has the following intuitive explanation. The event A_N in (5) represents a successful transmission event. For N large enough, $\Pr(A_N)$ represents the fraction of time that only the SU with the maximum of $\left\{ \frac{h_i}{\lambda_N + \mu_N g_i} \right\}_{i=1}^N$ transmits. In Appendix B, we show that, for $p_N = \frac{1}{N}$, $\Pr(A_N)$ converges to $\frac{1}{e}$ as N becomes large. Hence, as the number of SUs becomes large, the fraction of time that just the best SU transmits is approximately equal to $\frac{1}{e}$. Also, in Appendix B, we show that $\log(X_N^*(\lambda_N, \mu_N))$ term in (5) scales according to $\frac{1}{n_h} \log \log(N)$. These observations suggest that the secondary network sum-rate (under the optimum distributed power control) scales according to $\frac{1}{en_h} \log \log(N)$ as N becomes large. It should be noted that this is just an intuitive explanation, a rigorous proof of which is provided in Appendix B.

We also identify the second order determinants of $R_{\text{DTPIL}}^*\left(\frac{1}{N}, N\right)$ in Appendix B. Formally, we show that it can be expressed as (7). The first term in (7) converges to $\frac{1}{e} \log(P_{\text{ave}})$ as N becomes large. This finding displays the logarithmic effect of the power constraint on the secondary sum-rate in DTPIL networks. The second term in (7) gives rise to the scaling of secondary sum-rate according to

$$P_{\text{DTPIL}}^*(h, g) = \left(\frac{1}{\lambda_N + \mu_N g} - \frac{1}{h} \right)^+ \mathbf{1}_{\left\{ \frac{h}{\lambda_N + \mu_N g} > F_{\lambda_N, \mu_N}^{-1} \left(1 - \frac{1}{N} \right) \right\}} \quad (4)$$

$$A_N = \left\{ X_N^*(\lambda_N, \mu_N) > F_{\lambda_N, \mu_N}^{-1} \left(1 - \frac{1}{N} \right), X_N^\circ(\lambda_N, \mu_N) \leq F_{\lambda_N, \mu_N}^{-1} \left(1 - \frac{1}{N} \right) \right\}. \quad (6)$$

$$R_{\text{DTPIL}}^* \left(\frac{1}{N}, N \right) = \log \left(\frac{1}{\lambda_N} \right) \Pr(A_N) + \mathbb{E}_{h, g} \left[\log \left(X_N^* \left(1, \frac{\mu_N}{\lambda_N} \right) \right) \mathbf{1}_{A_N} \right]. \quad (7)$$

$\frac{1}{en_h} \log \log(N)$.

So far, we have assumed that p_N is equal to $\frac{1}{N}$. One may speculate that DTPIL networks may obtain a better throughput scaling behavior if the transmission probability is optimally adjusted, rather than to be set to $\frac{1}{N}$. To investigate this idea, we study the throughput scaling behavior of DTPIL networks under the optimum transmission probability selection in the next theorem.

Theorem 3: For each $N \in \mathbb{N}$, let p_N^* be an optimum transmission probability selection maximizing $R_{\text{DTPIL}}^*(p_N, N)$, i.e., $p_N^* \in \arg \max_{0 \leq p_N \leq 1} R_{\text{DTPIL}}^*(p_N, N)$. Then,

$$\lim_{N \rightarrow \infty} \frac{R_{\text{DTPIL}}^*(p_N^*, N)}{\log \log(N)} = \frac{1}{en_h}.$$

Proof: See Appendix C. ■

Theorem 3 indicates that the secondary network sum-rate under the optimum transmission probability also scales according to $\frac{1}{en_h} \log \log(N)$. Thus, the choice of transmission probability as $p_N = \frac{1}{N}$ is asymptotically optimal, and the secondary network achieves the same throughput scaling under p_N^* and $p_N = \frac{1}{N}$. However, it should be noted that the optimum transmission probability might be different from $\frac{1}{N}$ for any finite N . Identical throughput scaling behavior of DTPIL networks under p_N^* and $p_N = \frac{1}{N}$ gives rise to the following question: Does the optimum transmission probability asymptotically behave as $\frac{1}{N}$? The next lemma gives an affirmative answer to this question.

Lemma 1: For each $N \in \mathbb{N}$, let p_N^* be an optimum transmission probability selection in DTPIL networks. Then, $\lim_{N \rightarrow \infty} N p_N^* = 1$.

Proof: See Appendix D. ■

Lemma 1 shows that the optimum transmission probability in DTPIL networks should scale according to $\frac{1}{N}$. This scaling behavior of p_N^* can be intuitively considered as the origin of identical throughput scaling behavior of DTPIL networks under p_N^* and $p_N = \frac{1}{N}$ (i.e., see Appendix D for more details).

Finally, it is instructive to compare the throughput scaling laws obtained by using completely decentralized transmission strategies with those obtained through a centralized scheduler. In [12], it has been shown that the secondary network throughput with a centralized scheduler (usually, the SBS) scales according to $\frac{1}{m_h} \log \log(N)$ when the optimum power allocation policy is employed. Hence, compared to the centralized case, the factor $\frac{1}{e}$ here can be interpreted as the price of avoiding feedback signals between primary and secondary networks, which are the key parameters required by the centralized scheduler to perform optimum power control and scheduling. We note that the above results obtained for DTPIL networks

can be also extended to the individual power limited case using the techniques in [12]. However, we skip these calculations for the sake of brevity.

B. Optimum Power Control and Throughput Scaling in DIL Networks

In this case, transmission powers of SUs are only limited by a constraint on the total average interference power that SUs cause to the PBS (essentially, the average total transmission power constraint is inactive, which can happen when there is a large amount of average transmission power at disposal) and a transmission probability constraint. We define the power allocation policy in DIL networks, $P_{\text{DIL}}(\cdot, \cdot)$, as a mapping from \mathbb{R}_+^2 to \mathbb{R}_+ , where $P_{\text{DIL}}(h_i, g_i)$ denotes the transmission power of the i th SU at the joint channel state (h_i, g_i) . The power allocation policy in DIL networks is designed such that the transmission probability for all SUs is equal to p_N , i.e., $\Pr\{P_{\text{DIL}}(h_i, g_i) > 0\} = p_N$ for all $i \in \{1, \dots, N\}$. For a given power control policy P_{DIL} , the secondary network sum-rate in DIL networks can be expressed as $R_{\text{DIL}}(p_N, N, P_{\text{DIL}}) = N(1 - p_N)^{N-1} \mathbb{E}_{h, g} [\log(1 + h P_{\text{DIL}}(h, g))]$. In this case, transmission powers of SUs are allocated according to the solution of the following functional optimization problem:

$$\begin{aligned} & \text{maximize} && R_{\text{DIL}}(p_N, N, P_{\text{DIL}}) \\ & \text{subject to} && N \mathbb{E}_{h, g} [g P_{\text{DIL}}(h, g)] \leq Q_{\text{ave}} \\ & && \Pr\{P_{\text{DIL}}(h, g) > 0\} = p_N \end{aligned} \quad (8)$$

The next theorem establishes the structure of the optimum power allocation policy in DIL networks. The proof of Theorem 4 is similar to that of Theorem 1, and therefore, it is skipped to avoid repetition.

Theorem 4: Let $P_{\text{DIL}}^*(h, g)$ be the solution of (8). Then, for $p_N = \frac{1}{N}$ and N large enough, we have

$$P_{\text{DIL}}^*(h, g) = \left(\frac{1}{\mu_N g} - \frac{1}{h} \right)^+ \mathbf{1}_{\left\{ \frac{h}{g} > F_{\frac{h}{g}}^{-1} \left(1 - \frac{1}{N} \right) \right\}},$$

where μ_N is the power control parameter adjusted such that the average interference power constraint in (8) is met with equality, and $F_{\frac{h}{g}}^{-1}$ is the functional inverse of the CDF of $\frac{h}{g}$.

Theorem 4 implies that, for $p_N = \frac{1}{N}$ and N large enough, the optimum power allocation policy for the i th SU is to transmit by using a water-filling power allocation policy if its joint power and interference channel state, i.e., $\frac{h_i}{g_i}$, is greater than the threshold value of $F_{\frac{h}{g}}^{-1} \left(1 - \frac{1}{N} \right)$. In the centralized case, in order to maximize the secondary network sum-rate, the SBS schedules the SU having the maximum

of $\left\{\frac{h_i}{g_i}\right\}_{i=1}^N$, and the scheduled SU employs a water-filling power allocation policy, *i.e.*, see [11] and [12]. Moreover, the multiuser diversity gain with a centralized scheduler heavily depends on the maximum of $\left\{\frac{h_i}{g_i}\right\}_{i=1}^N$, and as the number of SUs becomes large, the maximum of $\left\{\frac{h_i}{g_i}\right\}_{i=1}^N$ takes values around $F_b^{-1}\left(1 - \frac{1}{N}\right)$ with high probability. In this regard, Theorem 4 further shows that, in DIL networks, a SU transmits with positive power if it has a high chance of being the SU with the maximum of $\left\{\frac{h_i}{g_i}\right\}_{i=1}^N$. Thus, we expect to observe throughput scaling behavior similar to that observed with a centralized scheduler, which is indeed the case as shown next.

Let Y_N^* and Y_N° be the largest and the second largest elements of the collection of random variables $\{Y_i\}_{i=1}^N$, where $Y_i = \frac{h_i}{g_i}$. Also, let $R_{\text{DIL}}^*(p_N, N)$ be the sum-rate in DIL networks under the optimum distributed power control policy with transmission probability equal to p_N . Then, for $p_N = \frac{1}{N}$ and N large enough, we have

$$R_{\text{DIL}}^*\left(\frac{1}{N}, N\right) = \mathbb{E}_{h,g} \left[\log \left(\frac{Y_N^*}{\mu_N} \right) \mathbf{1}_{B_N} \right],$$

where B_N is given by (9). The next theorem establishes the sum-rate scaling behavior of DIL networks.

Theorem 5: The secondary network sum-rate $R_{\text{DIL}}^*\left(\frac{1}{N}, N\right)$ for $p_N = \frac{1}{N}$ under the optimum distributed power control policy scales according to

$$\lim_{N \rightarrow \infty} \frac{R_{\text{DIL}}^*\left(\frac{1}{N}, N\right)}{\log(N)} = \frac{1}{e\gamma_g}.$$

Proof: See Appendix E. ■

Theorem 5 reveals the logarithmic scaling behavior of the secondary network sum-rate as a function of the number of SUs in DIL networks. It also shows that the pre-log multiplier in this scaling behavior is equal to $\frac{1}{e\gamma_g}$. The $\frac{1}{e}$ stems from the probability of successful transmission, whereas $\frac{1}{\gamma_g}$ term stems from the throughput scaling behavior on the event of successful transmission. γ_g is equal to 1 for Rayleigh and Rician- K distributed interference channel gains, and is equal to m and $\frac{c}{2}$ when interference channel gains are Nakagami- m and Weibull- c distributed, respectively.

Our analysis in Appendix E also reveals some second order effects on the secondary network throughput. In particular, we show that $R_{\text{DIL}}^*\left(\frac{1}{N}, N\right)$ can be written as (10). It is shown in Appendix E that μ_N converges to $\frac{1}{Q_{\text{ave}}}$ as N grows large. Hence, the first term in (10) converges to $\frac{1}{e} \log(Q_{\text{ave}})$ as N becomes large, implying the logarithmic effect of Q_{ave} on the secondary network sum-rate in DIL networks. Further, it is also shown that the second term in (10) scales according to $\frac{1}{e\gamma_g} \log(N)$, signifying the logarithmic effect of the number of SUs on the secondary network sum-rate in DIL networks.

It is also instructive to compare the result of Theorem 5 with the throughput scaling behavior that can be obtained by means of a centralized scheduler. The secondary network throughput scales according to $\frac{1}{\gamma_g} \log(N)$ when the optimum transmission power control is performed by a centralized scheduler [12]. This observation suggests that the throughput scaling law obtained through distributed implementation differs from that obtained in the centralized case only in the observed pre-log factors. Similar to the previous case, $\frac{1}{e}$ can be interpreted

as the cost of decentralized implementation of the cognitive uplink.

It might be hypothesized that the capacity scaling behavior obtained in Theorem 5 can be improved if the optimum transmission probability is employed instead of $p_N = \frac{1}{N}$. The next theorem disproves this hypothesis.

Theorem 6: For each $N \in \mathbb{N}$, let p_N^* be an optimum transmission probability selection maximizing $R_{\text{DIL}}^*(p_N, N)$, *i.e.*, $p_N^* \in \arg \max_{0 \leq p_N \leq 1} R_{\text{DIL}}^*(p_N, N)$. Then,

$$\lim_{N \rightarrow \infty} \frac{R_{\text{DIL}}^*(p_N^*, N)}{\log(N)} = \frac{1}{e\gamma_g}. \quad (11)$$

Proof: See Appendix F. ■

Theorem 6 establishes the logarithmic throughput scaling behavior of DIL networks under the optimum transmission probability. Hence, the choice of $p_N = \frac{1}{N}$ is asymptotically optimal, and one cannot obtain better throughput scaling by other choices of p_N . In the next lemma, we study the asymptotic behavior of the sequence of optimum transmission probabilities in DIL networks as the number of SUs becomes large.

Lemma 2: For each $N \in \mathbb{N}$, let p_N^* be an optimum transmission probability selection in DIL networks. Then, $\lim_{N \rightarrow \infty} N p_N^* = 1$.

Proof: See Appendix G. ■

Lemma 2 indicates that the sequence of optimum transmission probabilities in DIL networks decays to zero at the rate of $\frac{1}{N}$. In other words, $\frac{1}{N}$ serves as a good approximation for the optimum transmission probability when N is large enough.

As a final remark before numerical results, we note that the power optimization problems studied in (3) and (8) can also be solved for the case of non-i.i.d. fading distributions, albeit with a slight sacrifice from the distributed operation. In this case, the power optimization problems cannot be decomposed into smaller local optimization problems since SUs do not know each other's channel statistics. Hence, optimum power allocation policies need to be computed by the SBS under the statistical knowledge of direct and interference channel gains, and the resulting policies need to be broadcasted to all SUs. However, this computation is performed only once to initialize the communication, and the rest of the operation again proceeds in a distributed manner as above by using the *centrally* computed power allocation policies. Further, throughput scaling laws in the non-i.i.d. case can also be derived for specific fading distributions as in [11], but deriving general throughput scaling laws for arbitrary non-i.i.d. class- \mathcal{C} fading distributions is prohibitively hard and out of scope of the current paper.

V. NUMERICAL RESULTS

In this section, we numerically evaluate the sum-rate performance of DTPIL and DIL networks as a function of the number of SUs. We also compare their sum-rate scaling behavior with that of orthogonal channel access networks, *i.e.*, time division multiple access (TDMA) and frequency division multiple access (FDMA) networks. In the considered orthogonal channel access schemes, the global CSI is not available at the SBS either, and communication resources (*i.e.*, either time or frequency) are periodically allotted to SUs regardless of their channel conditions.

$$B_N = \left\{ Y_N^* > F_{\frac{h}{g}}^{-1} \left(1 - \frac{1}{N} \right), Y_N^\diamond \leq F_{\frac{h}{g}}^{-1} \left(1 - \frac{1}{N} \right) \right\}. \quad (9)$$

$$R_{\text{DIL}}^* \left(\frac{1}{N}, N \right) = \log \left(\frac{1}{\mu_N} \right) \Pr(B_N) + E_{h,g} [\log(Y_N^*) \mathbf{1}_{B_N}]. \quad (10)$$

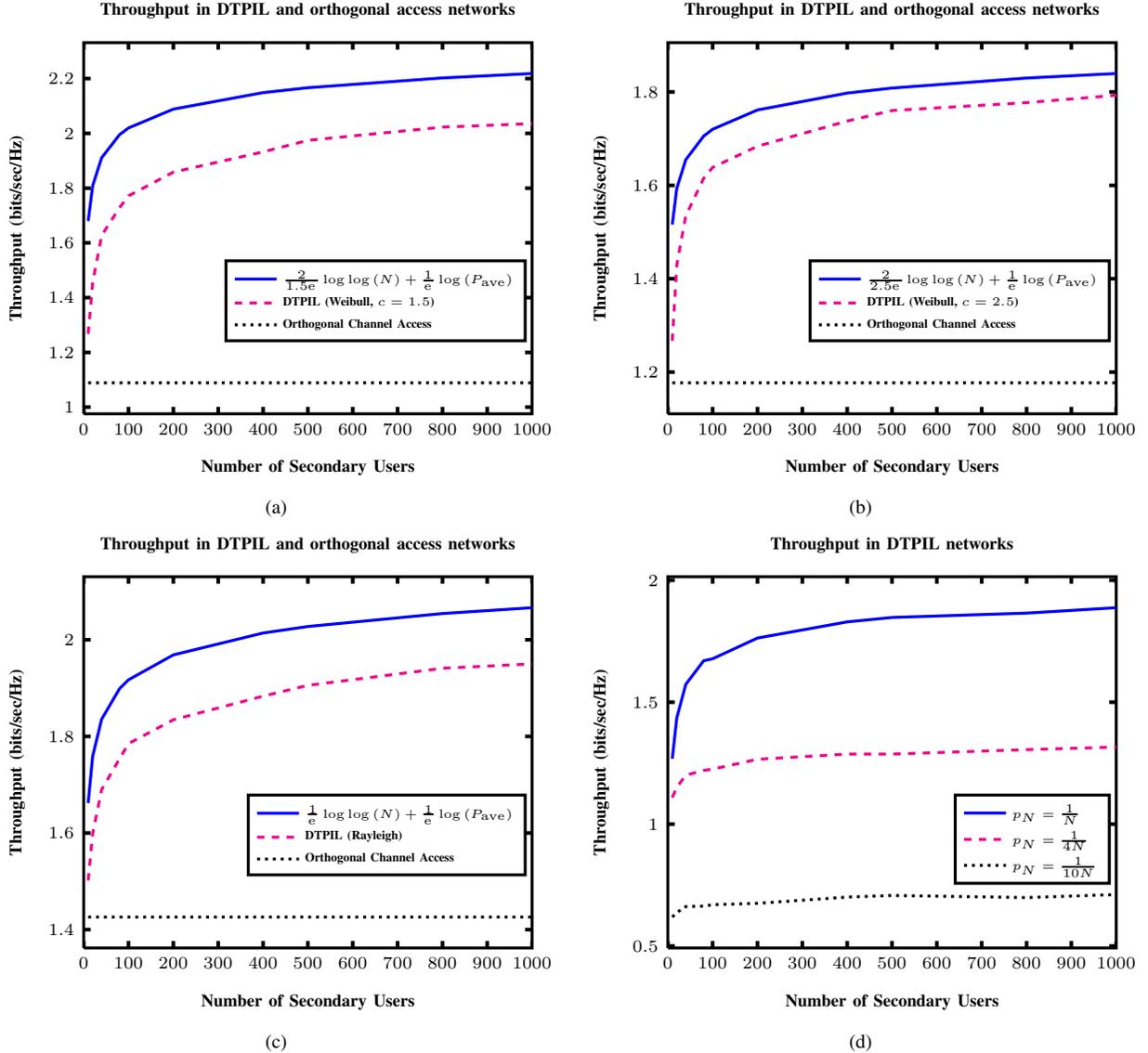


Fig. 2. Secondary network throughput in DTPIL and orthogonal channel access networks as a function of the number of SUs for different communication environments (a)-(c). Throughput in DTPIL networks as a function of the number of SUs for different choices of p_N (d). P_{ave} and Q_{ave} are set to 15dB and 0dB, respectively.

In TDMA networks, time is divided into equal length time slots, and a time slot is allocated to each SU. In FDMA networks, the total frequency band is divided into narrow-band frequency chunks, and each frequency chunk is allocated to a SU. In orthogonal channel access networks, we assume that SUs have access to their direct and interference channel gains, and upon being scheduled for transmission, each SU adjusts its transmission power level according to a single-user water-filling power allocation policy based on the local knowledge of its channel gains. The same average total transmission power and average interference power constraints are considered for

DTPIL, DIL and orthogonal channel access networks.

Figures 2(a)-(c) demonstrate the sum-rate scaling behavior of DTPIL and orthogonal channel access networks as a function of the number of SUs. In these figures, P_{ave} and Q_{ave} are set to 15dB and 0dB, respectively. Similar qualitative behavior continues to hold for other values of P_{ave} and Q_{ave} . In Figs. 2(a)-(c), p_N is set to $\frac{1}{N}$ for DTPIL networks. Also, identical fading models are considered for both DTPIL and orthogonal channel access networks.

More specifically, in Fig. 2(a), direct channel gains are distributed according to the Weibull- c fading model with

$c = 1.5$ and interference channel gains are distributed according to the Rayleigh fading model. In Fig. 2(b), direct channel gains are Weibull- c distributed with $c = 2.5$ and interference channel gains are Rayleigh distributed. As Fig. 2(a) and Fig. 2(b) show, the secondary network sum-rate in DTPIL networks scales according to $\frac{2}{ec} \log \log(N)$ with the number of SUs, *i.e.*, $\frac{2}{1.5e} \log \log(N)$ for $c = 1.5$ and $\frac{2}{2.5e} \log \log(N)$ for $c = 2.5$, when direct channel gains are Weibull- c distributed, a behavior predicted by Theorem 2. Also, closeness of the simulated data rates of DTPIL networks to the curves of $\frac{2}{1.5e} \log \log(N) + \frac{1}{e} \log(P_{ave})$ and $\frac{2}{2.5e} \log \log(N) + \frac{1}{e} \log(P_{ave})$ in Fig. 2(a) and Fig. 2(b), respectively, indicates the logarithmic effect of P_{ave} on the secondary network throughput in DTPIL networks.

In Fig. 2(c), direct channel gains are Rayleigh distributed and interference channel gains are Weibull- c distributed with $c = 1.5$. As Fig. 2(c) shows, the secondary network sum-rate in DTPIL networks scales according to $\frac{1}{e} \log \log(N)$ when direct channel gains are Rayleigh distributed, which is also in accordance with Theorem 2. Also, proximity of simulated data rates of DTPIL network to the $\frac{1}{e} \log \log(N) + \frac{1}{e} \log(P_{ave})$ curve in this figure again shows the logarithmic effect of P_{ave} on the secondary network sum-rate in DTPIL networks.

Moreover, as Figs. 2(a)-(c) show, the secondary network throughput in orthogonal channel access networks does not scale with the number of SUs since SUs are scheduled for transmission regardless of their channel gains, rather than being scheduled opportunistically, in these networks. Furthermore, DTPIL networks achieve much higher throughputs when compared to those achieved by orthogonal channel access networks, even with possibly suboptimal choice of transmission probability (*i.e.*, $p_N = \frac{1}{N}$) for small numbers of SUs. This is due to the fact that DTPIL networks can harvest multiuser diversity gains, in a distributed fashion, without any global knowledge of CSI at the SBS.

In Fig. 2(d), we demonstrate the secondary network throughput scaling in DTPIL networks as a function of the number of SUs when p_N is set to $\frac{1}{N}$, $\frac{1}{4N}$ and $\frac{1}{10N}$. In this figure, direct and interference channel gains are distributed according to the Rayleigh fading model. As Fig. 2(d) shows, the secondary network asymptotically achieves much higher throughputs with $p_N = \frac{1}{N}$ when compared to other choices of p_N that do not scale according to $\frac{1}{N}$. This finding signifies the importance of setting p_N correctly to maximize secondary network sum-rates in DTPIL networks.

Figure 3 shows the change of the secondary network sum-rate in DIL and orthogonal channel access networks as a function of the number of SUs for different communication environments. In this figure, Q_{ave} is set to 0dB. Similar qualitative behavior continues to hold for other values of Q_{ave} . The transmission probability is set to $\frac{1}{N}$ for DIL networks. In Fig. 3(a), direct channel gains are distributed according to the Rayleigh fading model and interference channel gains are distributed according to the Weibull- c fading model with $c = 1.5$. In Fig. 3(b), direct channel gains are Rayleigh distributed and interference channel gains are Weibull- c distributed with $c = 2.5$. From these figures, we can clearly observe that the secondary network throughput scales according to $\frac{2}{ec} \log(N)$ as a function of the number of SUs when interference channel

gains are Weibull distributed with different values of c .

In Figs. 3(c) and 3(d), direct channel gains are Rayleigh distributed and interference channel gains are Nakagami- m distributed with m set to 0.5 and 1.2, respectively. As these figures indicate, the secondary network throughput scales according to $\frac{1}{em} \log(N)$ with the number of SUs in DIL networks for Nakagami- m distributed interference channel gains. All simulated capacity curves in Fig. 3 concur with the capacity scaling laws established in Theorem 5. We know that $p_N = \frac{1}{N}$ may not be the optimum choice of transmission probability for N small enough, but from Fig. 3, we still observe that DIL networks with $p_N = \frac{1}{N}$ outperform orthogonal channel access networks largely, in terms of the sum-rate performance, even for small numbers of SUs.

In Fig. 4(a), we plot the normalized throughputs in DIL networks as a function of the number of SUs to further illustrate the accuracy of our scaling results. In this figure, direct channel gains are Rayleigh distributed and interference channel gains are Weibull- c distributed with $c = 1.5, 2.5$. As Fig. 4(a) shows, the sum-rate in DIL networks scales according to $\frac{2}{ce} \log(N)$, which is in harmony with Theorem 5.

Figure 4(b) depicts the throughput scaling behavior of DIL networks for different selections of the transmission probability. In this figure, direct and interference channel gains are Rayleigh distributed, and p_N is set to $\frac{1}{N}$, $\frac{1}{4N}$ and $\frac{1}{10N}$. We observe that the sum-rate performance of DIL networks under $p_N = \frac{1}{N}$ is asymptotically much higher when compared to that of DIL networks under $p_N = \frac{1}{4N}$ and $p_N = \frac{1}{10N}$. Similar to DTPIL networks, this observation indicates the importance of correct calibration of p_N to maximize secondary network sum-rates in DIL networks.

Figure 5 illustrates the throughput scaling behavior of DTPIL and DIL networks, with the number of SUs, for different values of P_{ave} and Q_{ave} , respectively. Direct and interference channel gains are distributed according to the Rayleigh fading model in this figure. In Fig. 5(a), Q_{ave} is set to zero dB. As Fig. 5(a) and Fig. 5(b) show, the secondary network throughput increases with the number of SUs due to multiuser diversity gains predicated by Theorems 2 and 5. Furthermore, as P_{ave} or Q_{ave} increase, SUs transmit with higher transmission powers which in turn results in higher throughputs as shown in Fig. 5(a) and Fig. 5(b).

VI. CONCLUSIONS

In this paper, we have studied the optimum distributed power control problem and the throughput scaling laws for the distributed cognitive uplink. First, we have shown that the optimum distributed power control policy for the cognitive uplink is in the form of a threshold based water-filling power control. The derived optimum distributed power control policy maximizes the secondary network sum-rate subject to transmission and interference power limitations, whilst guaranteeing primary QoS requirements without any feedback signals. Second, we have obtained tight throughput scaling laws for the distributed cognitive uplink by considering fading models general enough to include Rayleigh, Rician and Nakagami fading as special cases. In particular, it has been shown that the secondary network sum-rate, under the optimum distributed power control policy, scales according

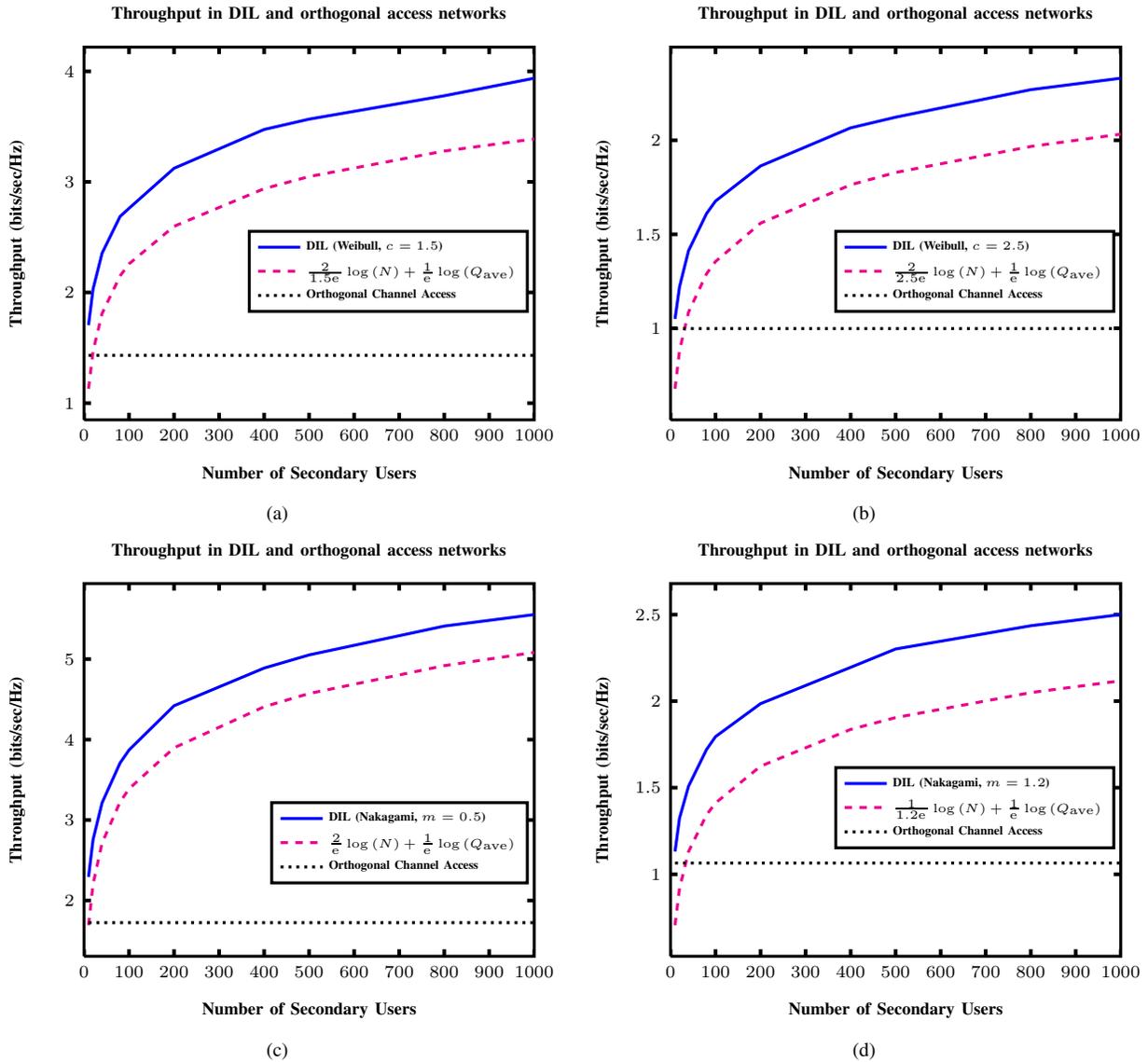


Fig. 3. Secondary network throughput in DIL and orthogonal channel access networks as a function of the number of SUs for different communication environments (a)-(d). Q_{ave} is set to 0dB.

to $\frac{1}{en_h} \log \log(N)$ when transmission powers of SUs are limited by a total average transmission power constraint and a constraint on the average interference power of SUs at the PBS. Here, n_h is a parameter obtained from the distribution of direct channel power gains, and N is the number of SUs. It has also been shown that the secondary network sum-rate, under the optimum distributed power control policy, scales according to $\frac{1}{e\gamma_g} \log(N)$ when transmission powers of SUs are only limited by an average interference power constraint. Here, γ_g is a parameter obtained from the distribution of interference channel power gains. Our throughput scaling results demonstrate that the cognitive uplink operating according to the derived optimum distributed power control policy is able to harvest multiuser diversity gains, even in a distributed fashion without any feedback between SUs and the SBS. The pre-log multiplier $\frac{1}{e}$ is the cost of distributed implementation of the cognitive uplink.

APPENDIX A PROOF OF THEOREM 1

To prove Theorem 1, we form a new functional optimization problem as follows

$$\begin{aligned}
 & \underset{\tilde{P}(h,g), W(h,g)}{\text{maximize}} && E_{h,g} \left[W(h,g) \log \left(1 + h\tilde{P}(h,g) \right) \right] \\
 & \text{subject to} && E_{h,g} \left[W(h,g) \tilde{P}(h,g) \right] \leq \frac{P_{ave}}{N} \\
 & && E_{h,g} \left[W(h,g) g\tilde{P}(h,g) \right] \leq \frac{Q_{ave}}{N} \\
 & && E_{h,g} [W(h,g)] = \frac{1}{N} \\
 & && 0 \leq W(h,g) \leq 1
 \end{aligned} \quad (12)$$

where $\tilde{P}(h,g)$ is a mapping from \mathbb{R}_+^2 to \mathbb{R}_+ . For $\tilde{P}(h,g) = P(h,g)$ and $W(h,g) = 1_{\{P(h,g) > 0\}}$, the optimization problem in (12) reduces to the one in (3). Thus, the optimal value of (12) serves as an upper bound for the optimal value of (3). Later, we show that this upper bound is achievable for N large enough. Using the change of variable $\Pi(h,g) = \tilde{P}(h,g) W(h,g)$, (12) can be transformed into the following

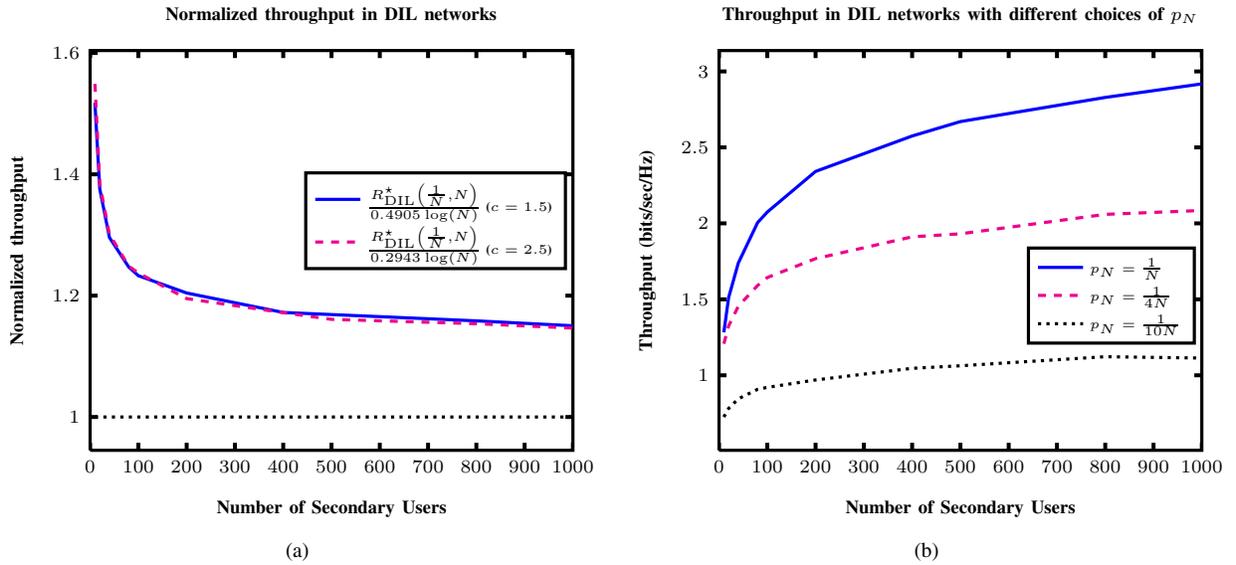


Fig. 4. Normalized throughput in DIL networks as a function of number of SUs (a). Secondary network throughput as a function of the number of SUs for different choices of p_N (b). Q_{ave} is set to 0dB.

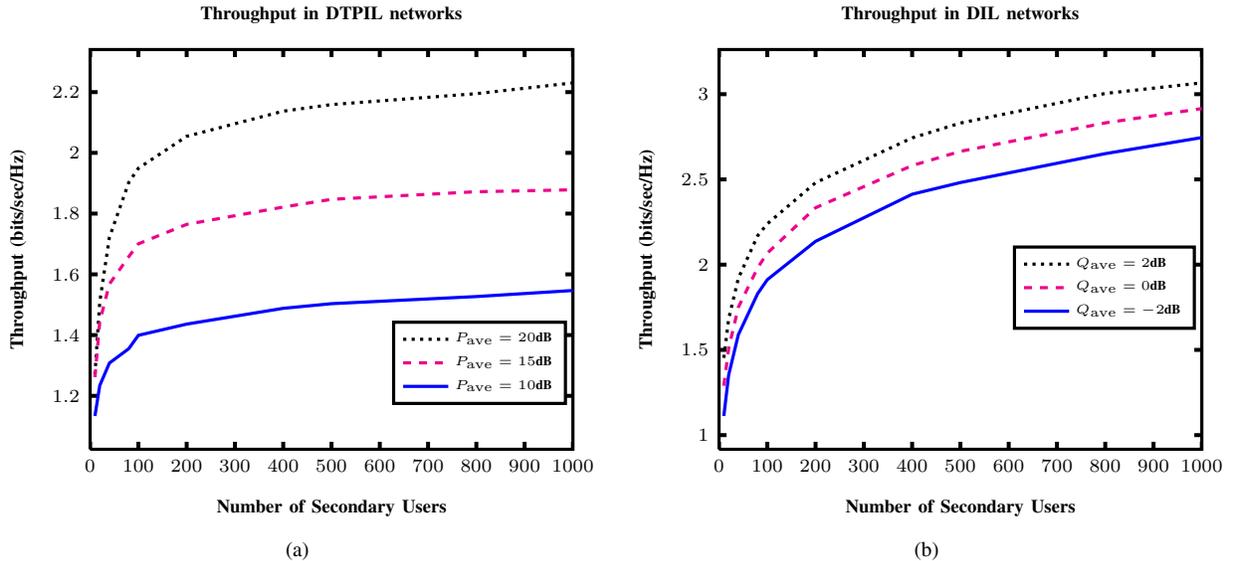


Fig. 5. Throughput in DTPIL networks versus number of SUs for different values of P_{ave} with $Q_{\text{ave}} = 0$ dB (a). Throughput in DIL networks versus number of SUs for different values of Q_{ave} (b).

convex optimization problem:

$$\begin{aligned}
 & \underset{\Pi(h,g), W(h,g)}{\text{maximize}} && \mathbb{E}_{h,g} \left[W(h,g) \log \left(1 + \frac{h\Pi(h,g)}{W(h,g)} \right) \right] \\
 & \text{subject to} && \mathbb{E}_{h,g} [\Pi(h,g)] \leq \frac{P_{\text{ave}}}{N} \\
 & && \mathbb{E}_{h,g} [g\Pi(h,g)] \leq \frac{Q_{\text{ave}}}{N} \\
 & && \mathbb{E}_{h,g} [W(h,g)] = \frac{1}{N} \\
 & && 0 \leq W(h,g) \leq 1 \\
 & && \Pi(h,g) \geq 0
 \end{aligned} \quad (13)$$

It can be shown that the objective function in (13) as a function of Π and W is concave on \mathbb{R}_+^2 . The Lagrangian for (13) can be written as (14) where $\lambda_N \geq 0$, $\mu_N \geq 0$ and η_N are Lagrange multipliers associated with the average transmission power, average interference power and transmission probability constraints, respectively. Let $\Pi^*(h,g)$ and $W^*(h,g)$ be a pair of solutions of (13). Let also let $A = \{(h,g) \in \mathbb{R}_+^2 : \Pi^*(h,g) > 0\}$

and $B = \{(h,g) \in \mathbb{R}_+^2 : W^*(h,g) > 0\}$. Using generalized Karush-Kuhn-Tucker (KKT) conditions [32], [33], we have (15). This result implies that $\Pi^*(h,g)$ and $W^*(h,g)$ satisfy the first order KKT conditions if and only if $\Pi^*(h,g) = \left(\frac{1}{\lambda_N + \mu_N g} - \frac{1}{h} \right)^+ W^*(h,g)$. We let $\tilde{P}^*(h,g) = \left(\frac{1}{\lambda_N + \mu_N g} - \frac{1}{h} \right)^+$. From KKT conditions, we also need to have (16). Let $f(x) = \log(1+x) + \frac{1}{1+x}$ (note that $f(x)$ is a monotonically increasing function for $x \geq 0$), $C = \{(h,g) \in \mathbb{R}_+^2 : W^*(h,g) = 1\}$ and $D = \{(h,g) \in \mathbb{R}_+^2 : 0 < W^*(h,g) < 1\}$. First consider the case $\eta_N > 0$. Using KKT conditions, we have $\frac{h}{\lambda_N + \mu_N g} \geq 1 + f^{-1}(1 + \eta_N)$ on the set C . Similarly, on D , we have $\frac{h}{\lambda_N + \mu_N g} = 1 + f^{-1}(1 + \eta_N)$. Hence, the choice of $W^*(h,g)$ as $W^*(h,g) = 1_{\left\{ \frac{h}{\lambda_N + \mu_N g} > 1 + f^{-1}(1 + \eta_N) \right\}} + \epsilon \cdot$

$$L(\Pi, W, \lambda_N, \mu_N, \eta_N) = W(h, g) \log \left(1 + \frac{h\Pi(h, g)}{W(h, g)} \right) - \lambda_N \Pi(h, g) - \mu_N g \Pi(h, g) - \eta_N W(h, g). \quad (14)$$

$$\begin{aligned} \frac{\partial L(\Pi, W^*, \lambda_N, \mu_N, \eta_N)}{\partial \Pi(h, g)} \Big|_{\Pi=\Pi^*} &= \left(\frac{h}{1 + \frac{h\Pi^*(h, g)}{W^*(h, g)}} - \lambda_N - \mu_N g \right) \mathbf{1}_B - (\lambda_N + \mu_N g) \mathbf{1}_{B^c} \begin{cases} = 0 & \Pi^*(h, g) > 0 \\ \leq 0 & \Pi^*(h, g) = 0 \end{cases}. \end{aligned} \quad (15)$$

$$\begin{aligned} \frac{\partial L(\Pi^*, W, \lambda_N, \mu_N, \eta_N)}{\partial W(h, g)} \Big|_{W=W^*} &= \left(\log \left(1 + h\tilde{P}^*(h, g) \right) + \frac{1}{1 + h\tilde{P}^*(h, g)} - 1 - \eta_N \right) \mathbf{1}_A - \eta_N \mathbf{1}_{A^c} \begin{cases} = 0 & 0 < W^*(h, g) < 1 \\ \leq 0 & W^*(h, g) = 0 \\ \geq 0 & W^*(h, g) = 1 \end{cases}. \end{aligned} \quad (16)$$

$\mathbf{1}_{\left\{ \frac{h}{\lambda_N + \mu_N g} = 1 + f^{-1}(1 + \eta_N) \right\}}$ for some $\epsilon \in (0, 1)$ satisfies the KKT conditions. Since h and g have continuous distributions, we can also write $W^*(h, g) = \mathbf{1}_{\left\{ \frac{h}{\lambda_N + \mu_N g} > 1 + f^{-1}(1 + \eta_N) \right\}}$ with probability one. We set η_N such that $1 + f^{-1}(1 + \eta_N) = F_{\lambda_N, \mu_N}^{-1} \left(1 - \frac{1}{N} \right)$ to satisfy the constraint on $E_{h, g} [W^*(h, g)]$. We note that $F_{\lambda_N, \mu_N}^{-1} \left(1 - \frac{1}{N} \right)$ must be larger than 1 for $\eta_N > 0$. Similar arguments show that $W^*(h, g) = \mathbf{1}_{\left\{ \frac{h}{\lambda_N + \mu_N g} > F_{\lambda_N, \mu_N}^{-1} \left(1 - \frac{1}{N} \right) \right\}}$ and $F_{\lambda_N, \mu_N}^{-1} \left(1 - \frac{1}{N} \right) \leq 1$ for $\eta_N = 0$. For $\eta_N < 0$, there does not exist any $W^*(h, g)$ satisfying KKT conditions. Hence, combining all cases, we conclude that $\Pi^*(h, g)$ and $W^*(h, g)$ can be chosen as $W^*(h, g) = \mathbf{1}_{\left\{ \frac{h}{\lambda_N + \mu_N g} > F_{\lambda_N, \mu_N}^{-1} \left(1 - \frac{1}{N} \right) \right\}}$ and $\Pi^*(h, g) = \tilde{P}^*(h, g) W^*(h, g)$ with probability one. It is well known that a one-to-one mapping of optimization variables to convert one optimization problem to another preserves the solvability of the KKT conditions. In other words, if the new variables solve the KKT conditions of the transformed problem, the old variables solve the KKT conditions of the original problem. This implies that $\tilde{P}^*(h, g) = \frac{\Pi^*(h, g)}{W^*(h, g)}$ and $W^*(h, g)$ constitute a solution for Problem (12). For $\lambda_N > 0$, we have

$$\begin{aligned} P_{\text{ave}} &= N E_{h, g} \left[\left(\frac{1}{\lambda_N + \mu_N g} - \frac{1}{h} \right)^+ W^*(h, g) \right] \\ &\leq \frac{1}{\lambda_N} N E_{h, g} [W^*(h, g)] \\ &= \frac{1}{\lambda_N}. \end{aligned}$$

Thus, we have $\lambda_N \leq \frac{1}{P_{\text{ave}}}$. Similarly, it can be shown that $\mu_N \leq \frac{1}{Q_{\text{ave}}}$. Hence, we have $F_{\lambda_N, \mu_N}^{-1} \left(1 - \frac{1}{N} \right) \geq F_{\frac{1}{P_{\text{ave}}}, \frac{1}{Q_{\text{ave}}}}^{-1} \left(1 - \frac{1}{N} \right) \geq 1$ for N large enough. This implies that $P_{\text{DTPIL}}(h, g) = \tilde{P}^*(h, g) W^*(h, g)$ is also a feasible solution for (3) when N is large enough. For $P_{\text{DTPIL}}(h, g) = \tilde{P}^*(h, g) W^*(h, g)$, the value of objective function in (3) is equal to the optimal value of (12), implying the upper bound is achieved, which completes the proof.

APPENDIX B THROUGHPUT SCALING IN DTPIL NETWORKS

In this appendix, we first establish some preliminary results. Then, we use these results to prove Theorem 2. Lemma 3 below establishes the asymptotic behavior of $F_{\lambda, \mu}^{-1}(x)$, which is the *functional* inverse of the common CDF of joint channel states $\frac{h_i}{\lambda + \mu g_i}$, $i = 1, 2, \dots, N$, as x becomes close to one.

Lemma 3: Let $F_{\lambda, \mu}(x)$ be the common CDF of joint channel states $\frac{h_i}{\lambda + \mu g_i}$, $i = 1, 2, \dots, N$, where $\lambda > 0$ and $\mu \geq 0$ are constants. Then, as x becomes close to one, its functional inverse $F_{\lambda, \mu}^{-1}(x)$ scales according to $\lim_{x \uparrow 1} \frac{F_{\lambda, \mu}^{-1}(x)}{\frac{1}{x} \left(-\frac{1}{\beta_h} \log(1-x) \right)^{\frac{1}{n_h}}} = 1$.

Proof: We only focus on the case where both λ and μ are strictly positive. The proof of the remaining case in which $\lambda > 0$ and $\mu = 0$ is easier and follows along the same lines. To prove the desired result, we first obtain the asymptotic behavior of $F_{\lambda, \mu}(x)$ as x becomes large. Note that $F_{\lambda, \mu}(x)$ is the CDF of the product of two independent random variables, *i.e.*, h_i and $\frac{1}{\lambda + \mu g_i}$, and the asymptotic tail behavior for the product of two independent random variables was studied in [34] for the case of $H(x) = 0$. Since $H(x)$ is not necessarily equal to zero for the class- \mathcal{C} distributions, *i.e.*, see the Rician fading model in Table II, we need to upper and lower bound the tail of $F_{\lambda, \mu}(x)$ by using distribution functions with $H(x) = 0$. To this end, we let $h_{+\epsilon}$ and $h_{-\epsilon}$ be two random variables, independent of g_i , with respective CDFs $F_{+\epsilon}(x)$ and $F_{-\epsilon}(x)$ satisfying $\lim_{x \rightarrow \infty} \frac{1 - F_{+\epsilon}(x)}{\alpha_h x^{l_h} e^{-(\beta_h - \epsilon)x^{n_h}}} = \lim_{x \rightarrow \infty} \frac{1 - F_{-\epsilon}(x)}{\alpha_h x^{l_h} e^{-(\beta_h + \epsilon)x^{n_h}}} = 1$ for $\epsilon > 0$ small enough.

Let $F_{+\epsilon, \lambda, \mu}(x)$ and $F_{-\epsilon, \lambda, \mu}(x)$ be the CDFs of $\frac{h_{+\epsilon}}{\lambda + \mu g_i}$ and $\frac{h_{-\epsilon}}{\lambda + \mu g_i}$, respectively. Let also $F_h(x)$ be the CDF of h_i . Observing that $F_{+\epsilon}(x) \leq F_h(x) \leq F_{-\epsilon}(x)$ for x large enough, we can upper and lower bound $F_{\lambda, \mu}(x)$ as (17) for x large enough, where expectations are taken over interference channel states. Using Theorem 3 in [34], the asymptotic tail behavior of $F_{+\epsilon, \lambda, \mu}(x)$ can be shown to satisfy $\lim_{x \rightarrow \infty} \frac{1 - F_{+\epsilon, \lambda, \mu}(x)}{C x^{l_h - n_h} \gamma_g e^{-(\beta_h - \epsilon)(\lambda x)^{n_h}}} = 1$, where $C = \eta_g \alpha_h \Gamma(\gamma_g + 1) \left(\frac{\lambda^2}{\mu(\beta_h - \epsilon)n_h} \right)^{\gamma_g} \left(\frac{1}{\lambda} \right)^{n_h \gamma_g + \gamma_g - l_h}$ and $\Gamma(\cdot)$

$$F_{+\epsilon, \lambda, \mu}(x) = \mathbb{E}_{g_i} [F_{+\epsilon}((\lambda + \mu g_i)x)] \leq F_{\lambda, \mu}(x) = \mathbb{E}_{g_i} [F_h((\lambda + \mu g_i)x)] \leq F_{-\epsilon, \lambda, \mu}(x) = \mathbb{E}_{g_i} [F_{-\epsilon}((\lambda + \mu g_i)x)]. \quad (17)$$

is the Gamma function [37]. This result implies that the functional inverse $F_{+\epsilon, \lambda, \mu}^{-1}(x)$ of $F_{+\epsilon, \lambda, \mu}(x)$ behaves according to $\lim_{x \uparrow 1} \frac{F_{+\epsilon, \lambda, \mu}^{-1}(x)}{\frac{1}{\lambda} \left(-\frac{1}{(\beta_h - \epsilon)} \log(1-x) \right)^{\frac{1}{n_h}}} = 1$ as x becomes close to one. Following the same steps, we also have $\lim_{x \uparrow 1} \frac{F_{-\epsilon, \lambda, \mu}^{-1}(x)}{\frac{1}{\lambda} \left(-\frac{1}{(\beta_h + \epsilon)} \log(1-x) \right)^{\frac{1}{n_h}}} = 1$. Using (17), $F_{\lambda, \mu}^{-1}(x)$ can be upper and lower bounded as $F_{-\epsilon, \lambda, \mu}^{-1}(x) \leq F_{\lambda, \mu}^{-1}(x) \leq F_{+\epsilon, \lambda, \mu}^{-1}(x)$ for x close enough to one. Since ϵ can be chosen arbitrarily close to zero, we have $\lim_{x \uparrow 1} \frac{F_{\lambda, \mu}^{-1}(x)}{\frac{1}{\lambda} \left(-\frac{1}{\beta_h} \log(1-x) \right)^{\frac{1}{n_h}}} = 1$ which completes the proof. ■

Next, by using Lemma 3, we establish the asymptotic behavior for the extreme order statistic of the collection of random variables $\left\{ \frac{h_i}{\lambda + \mu g_i} \right\}_{i=1}^N$. The derived convergence behavior will be helpful for studying the asymptotic behavior of λ_N , and in turn, for proving Theorem 2.

Lemma 4: Let $X_N^*(\lambda, \mu) = \max_{1 \leq i \leq N} \frac{h_i}{\lambda + \mu g_i}$ for $\lambda > 0$ and $\mu \geq 0$. Then, $\frac{X_N^*(\lambda, \mu)}{\left(\frac{1}{\beta_h} \log(N) \right)^{\frac{1}{n_h}}} \xrightarrow{i.p.} \frac{1}{\lambda}$ as N tends to infinity, where *i.p.* stands for convergence in probability.

Proof: Let $F_{\lambda, \mu}$ be the CDF of $\frac{h_i}{\lambda + \mu g_i}$ as in Lemma 3. Using Lemma 2 in [12], the concentration behavior of $X_N^*(\lambda, \mu)$ can be given as (18) for all $\epsilon > 0$ small enough. Using Lemma 3 above and (18), we have (19) which implies the convergence of $\frac{X_N^*(\lambda, \mu)}{\left(\frac{1}{\beta_h} \log(N) \right)^{\frac{1}{n_h}}}$ to $\frac{1}{\lambda}$ in probability. ■

In the next lemma, we show that λ_N converges to $\frac{1}{P_{\text{ave}}}$ as N becomes large. This lemma will be used to quantify the effect of the average total power constraint P_{ave} on the secondary network throughput in DTPIL networks.

Lemma 5: Let λ_N be one of the power control parameters in DTPIL networks, as described in (4). Then, $\lim_{N \rightarrow \infty} \lambda_N = \frac{1}{P_{\text{ave}}}$.

Proof: First, we show $\liminf_{N \rightarrow \infty} \lambda_N > 0$. To obtain a contradiction, we assume that λ_N can be arbitrarily close to zero as N becomes large. This implies that for all $\epsilon > 0$, we can find a subsequence of N, N_j , such that $\lambda_{N_j} \leq \epsilon$ for all N_j large enough. Let $X_{N_j}^*(\lambda, \mu) = \max_{1 \leq i \leq N_j} \frac{h_i}{\lambda + \mu g_i}$ (as in Lemma 4), $h_{N_j}^* = \max_{1 \leq i \leq N_j} h_i$ and $I_N = \arg \max_{1 \leq i \leq N} \frac{h_i}{\lambda_N + \mu_N g_i}$. Let also $P_N^{\text{DTPIL}}(\mathbf{h}, \mathbf{g}) = \sum_{i=1}^N \left(\frac{1}{\lambda_N + \mu_N g_i} - \frac{1}{h_i} \right)^+ \mathbb{1}_{\left\{ \frac{h_i}{\lambda_N + \mu_N g_i} > F_{\lambda_N, \mu_N}^{-1} \left(1 - \frac{1}{N} \right) \right\}}$ be the instantaneous total power consumed by the secondary network. Then, the average power consumption, for all N_j large enough, can be lower bounded as (20) where (a) follows from observing that $\mu_N \leq \frac{N P_N}{Q_{\text{ave}}}$ and $p_N = \frac{1}{N}$ in this case. Using Lemma 4, we have $\frac{X_{N_j}^*(\epsilon, \frac{1}{Q_{\text{ave}}})}{\left(\frac{1}{\beta_h} \log(N_j) \right)^{\frac{1}{n_h}}} \xrightarrow{i.p.} \frac{1}{\epsilon}$ and $\frac{h_{N_j}^*}{\left(\frac{1}{\beta_h} \log(N_j) \right)^{\frac{1}{n_h}}} \xrightarrow{i.p.} 1$ as N_j tends to infinity.

Also, it is easy to see that $\mathbb{1}_{\left\{ X_{N_j}^*(\lambda_{N_j}, \mu_{N_j}) > F_{\lambda_{N_j}, \mu_{N_j}}^{-1} \left(1 - \frac{1}{N_j} \right) \right\}} \xrightarrow{i.d.} \text{Bern} \left(1 - \frac{1}{e} \right)$ as N_j tends to infinity, where $\text{Bern}(p)$ denotes a 0-1 Bernoulli random variable with mean p , and *i.d.* stands for convergence in distribution. Hence, by using Slutsky's Theorem [35], we have (21). Applying Fatou's Lemma to (20), we obtain $\liminf_{N_j \rightarrow \infty} \mathbb{E}_{\mathbf{h}, \mathbf{g}} \left[P_{N_j}^{\text{DTPIL}}(\mathbf{h}, \mathbf{g}) \right] \geq \frac{1}{e} \left(1 - \frac{1}{e} \right)$, which implies that the average power consumption can be made arbitrarily large, violating the power constraint, for ϵ small enough and N_j large enough. Thus, $\liminf_{N \rightarrow \infty} \lambda_N > 0$.

Now, by using the fact that λ_N cannot be arbitrarily close to zero, we show that $\lim_{N \rightarrow \infty} \lambda_N = \frac{1}{P_{\text{ave}}}$. Note that $\lambda_N \leq \frac{1}{P_{\text{ave}}}$ for $p_N = \frac{1}{N}$, which implies that $\limsup_{N \rightarrow \infty} \lambda_N \leq \frac{1}{P_{\text{ave}}}$. Hence, showing that $\liminf_{N \rightarrow \infty} \lambda_N \geq \frac{1}{P_{\text{ave}}}$ will conclude the proof. To this end, the average total power consumed by the secondary network can be lower bounded as (22) where (a) follows from observing that $\lambda F_{\lambda, \mu}^{-1}(x) = F_{\frac{1}{\lambda}, \mu}^{-1}(x)$, and (b) follows from observing that $\mu_N \leq \frac{1}{Q_{\text{ave}}}$ and $F_{\lambda, \mu}^{-1}(x)$ decreases with increasing values of μ . Using (22), λ_N can be lower bounded as (23). Using Lemma 3 and the fact that λ_N cannot be arbitrarily close to zero, we have $\lim_{N \rightarrow \infty} \frac{F_{\frac{1}{\lambda_N}, \frac{1}{Q_{\text{ave}}}}^{-1} \left(1 - \frac{1}{N} \right)}{\left(\frac{1}{\beta_h} \log(N) \right)^{\frac{1}{n_h}}} = 1$, which implies $\frac{\left(F_{\frac{1}{\lambda_N}, \frac{1}{Q_{\text{ave}}}}^{-1} \left(1 - \frac{1}{N} \right) - \lambda_N \right)^+}{\left(\frac{1}{\beta_h} \log(N) \right)^{\frac{1}{n_h}}} \xrightarrow{i.p.} 1$ as N tends to infinity.

Let $S_N = \sum_{i=1}^{h_{N_j}^*} \mathbb{1}_{\left\{ \frac{h_i}{1 + \frac{\mu_N}{\lambda_N} g_i} > F_{\frac{1}{\lambda_N}, \frac{\mu_N}{\lambda_N}}^{-1} \left(1 - \frac{1}{N} \right) \right\}}$. S_N has a Binomial distribution with parameters N and $\frac{1}{N}$. Hence, using Poisson approximation for Binomial distributions, we conclude that S_N converges in distribution to $\text{Po}(1)$, where $\text{Po}(p)$ represents a Poisson random variable with mean p . Using Slutsky's Theorem, we have (24) as N grows large. Applying Fatou's Lemma to (23), we have $\liminf_{N \rightarrow \infty} \lambda_N \geq \frac{1}{P_{\text{ave}}}$. ■

Now, we are ready to prove Theorem 2 by utilizing above auxiliary results. Note that the sum-rate under the optimum distributed power control in DTPIL networks for $p_N = \frac{1}{N}$ can be written as (25). Note that $\Pr(A_N)$ can be written as (26). Hence, we have $\lim_{N \rightarrow \infty} \Pr(A_N) = \frac{1}{e}$ by the selection of transmission probabilities. This gives us the logarithmic effect of P_{ave} on the secondary network throughput since λ_N converges to $\frac{1}{P_{\text{ave}}}$. Using Lemma 4, we have $\frac{\log(X_N^*(1, \frac{\mu_N}{\lambda_N}))}{\log \log(N)} \xrightarrow{i.p.} \frac{1}{n_h}$ as N tends to infinity since λ_N is bounded away from zero and $\mu_N \leq \frac{1}{Q_{\text{ave}}}$. Also, we have $\mathbb{1}_{A_N}$ converging in distribution to $\text{Bern} \left(\frac{1}{e} \right)$ as N tends to infinity. As a result, applying Slutsky's Theorem, we conclude that $\frac{\log(X_N^*(1, \frac{\mu_N}{\lambda_N}))}{\log \log(N)} \mathbb{1}_{A_N} \xrightarrow{i.d.} \frac{1}{n_h} \text{Bern} \left(\frac{1}{e} \right)$. This final result almost completes the proof of Theorem 2 up to a slight technicality. That is, convergence in distribution does not always imply convergence in mean [36].

$$\lim_{N \rightarrow \infty} \Pr \left\{ F_{\lambda, \mu}^{-1} (1 - N^{-\epsilon-1}) \leq X_N^* (\lambda, \mu) \leq F_{\lambda, \mu}^{-1} (1 - N^{-\epsilon-1}) \right\} = 1. \quad (18)$$

$$\lim_{N \rightarrow \infty} \Pr \left\{ \frac{1}{\lambda} (1 - \epsilon)^{\frac{1}{n_h}} \leq \frac{X_N^* (\lambda, \mu)}{\left(\frac{1}{\beta_h} \log(N) \right)^{\frac{1}{n_h}}} \leq \frac{1}{\lambda} (1 + \epsilon)^{\frac{1}{n_h}} \right\} = 1. \quad (19)$$

$$\begin{aligned} \mathbb{E}_{\mathbf{h}, \mathbf{g}} \left[P_{N_j}^{\text{DTPIL}} (\mathbf{h}, \mathbf{g}) \right] &= \mathbb{E}_{\mathbf{h}, \mathbf{g}} \left[\sum_{i=1}^{N_j} \left(\frac{1}{\lambda_{N_j} + \mu_{N_j} g_i} - \frac{1}{h_i} \right)^+ \mathbf{1}_{\left\{ \frac{h_i}{\lambda_{N_j} + \mu_{N_j} g_i} > F_{\lambda_{N_j}, \mu_{N_j}}^{-1} \left(1 - \frac{1}{N_j} \right) \right\}} \right] \\ &\geq \mathbb{E}_{\mathbf{h}, \mathbf{g}} \left[\frac{1}{h_{I_{N_j}}} \left(X_{N_j}^* (\lambda_{N_j}, \mu_{N_j}) - 1 \right)^+ \mathbf{1}_{\left\{ X_{N_j}^* (\lambda_{N_j}, \mu_{N_j}) > F_{\lambda_{N_j}, \mu_{N_j}}^{-1} \left(1 - \frac{1}{N_j} \right) \right\}} \right] \\ &\stackrel{(a)}{\geq} \mathbb{E}_{\mathbf{h}, \mathbf{g}} \left[\frac{1}{h_{N_j}^*} \left(X_{N_j}^* \left(\epsilon, \frac{1}{Q_{\text{ave}}} \right) - 1 \right)^+ \mathbf{1}_{\left\{ X_{N_j}^* (\lambda_{N_j}, \mu_{N_j}) > F_{\lambda_{N_j}, \mu_{N_j}}^{-1} \left(1 - \frac{1}{N_j} \right) \right\}} \right]. \end{aligned} \quad (20)$$

$$\frac{1}{h_{N_j}^*} \left(X_{N_j}^* \left(\epsilon, \frac{1}{Q_{\text{ave}}} \right) - 1 \right)^+ \mathbf{1}_{\left\{ X_{N_j}^* (\lambda_{N_j}, \mu_{N_j}) > F_{\lambda_{N_j}, \mu_{N_j}}^{-1} \left(1 - \frac{1}{N_j} \right) \right\}} \xrightarrow{i.d.} \frac{1}{\epsilon} \text{Bern} \left(1 - \frac{1}{e} \right). \quad (21)$$

$$\begin{aligned} P_{\text{ave}} &= \mathbb{E}_{\mathbf{h}, \mathbf{g}} \left[\sum_{i=1}^N \left(\frac{1}{\lambda_N + \mu_N g_i} - \frac{1}{h_i} \right)^+ \mathbf{1}_{\left\{ \frac{h_i}{\lambda_N + \mu_N g_i} > F_{\lambda_N, \mu_N}^{-1} \left(1 - \frac{1}{N} \right) \right\}} \right] \\ &\stackrel{(a)}{=} \frac{1}{\lambda_N} \mathbb{E}_{\mathbf{h}, \mathbf{g}} \left[\sum_{i=1}^N \left(\frac{1}{1 + \frac{\mu_N}{\lambda_N} g_i} - \frac{1}{h_i} \right)^+ \mathbf{1}_{\left\{ \frac{h_i}{1 + \frac{\mu_N}{\lambda_N} g_i} > F_{1, \frac{\mu_N}{\lambda_N}}^{-1} \left(1 - \frac{1}{N} \right) \right\}} \right] \\ &\stackrel{(b)}{\geq} \frac{1}{\lambda_N} \mathbb{E}_{\mathbf{h}, \mathbf{g}} \left[\sum_{i=1}^N \frac{\left(F_{1, \frac{1}{\lambda_N Q_{\text{ave}}}}^{-1} \left(1 - \frac{1}{N} \right) - \lambda_N \right)^+}{h_i} \mathbf{1}_{\left\{ \frac{h_i}{1 + \frac{\mu_N}{\lambda_N} g_i} > F_{1, \frac{\mu_N}{\lambda_N}}^{-1} \left(1 - \frac{1}{N} \right) \right\}} \right] \\ &\geq \frac{1}{\lambda_N} \mathbb{E}_{\mathbf{h}, \mathbf{g}} \left[\frac{\left(F_{1, \frac{1}{\lambda_N Q_{\text{ave}}}}^{-1} \left(1 - \frac{1}{N} \right) - \lambda_N \right)^+}{h_N^*} \sum_{i=1}^N \mathbf{1}_{\left\{ \frac{h_i}{1 + \frac{\mu_N}{\lambda_N} g_i} > F_{1, \frac{\mu_N}{\lambda_N}}^{-1} \left(1 - \frac{1}{N} \right) \right\}} \right]. \end{aligned} \quad (22)$$

$$\lambda_N \geq \frac{1}{P_{\text{ave}}} \mathbb{E}_{\mathbf{h}, \mathbf{g}} \left[\frac{\left(F_{1, \frac{1}{\lambda_N Q_{\text{ave}}}}^{-1} \left(1 - \frac{1}{N} \right) - \lambda_N \right)^+}{h_N^*} \sum_{i=1}^N \mathbf{1}_{\left\{ \frac{h_i}{1 + \frac{\mu_N}{\lambda_N} g_i} > F_{1, \frac{\mu_N}{\lambda_N}}^{-1} \left(1 - \frac{1}{N} \right) \right\}} \right]. \quad (23)$$

$$\frac{\left(F_{1, \frac{1}{\lambda_N Q_{\text{ave}}}}^{-1} \left(1 - \frac{1}{N} \right) - \lambda_N \right)^+}{h_N^*} \sum_{i=1}^N \mathbf{1}_{\left\{ \frac{h_i}{1 + \frac{\mu_N}{\lambda_N} g_i} > F_{1, \frac{\mu_N}{\lambda_N}}^{-1} \left(1 - \frac{1}{N} \right) \right\}} \xrightarrow{i.d.} \text{Po} (1). \quad (24)$$

$$R_{\text{DTPIL}}^* \left(\frac{1}{N}, N \right) = \log \left(\frac{1}{\lambda_N} \right) \Pr (A_N) + \mathbb{E}_{\mathbf{h}, \mathbf{g}} \left[\log \left(X_N^* \left(1, \frac{\mu_N}{\lambda_N} \right) \right) \mathbf{1}_{A_N} \right]. \quad (25)$$

$$\Pr (A_N) = \sum_{i=1}^N \Pr \left\{ X_i (\lambda_N, \mu_N) > F_{\lambda_N, \mu_N}^{-1} \left(1 - \frac{1}{N} \right) \right\} \prod_{j \neq i} \Pr \left\{ X_j (\lambda_N, \mu_N) \leq F_{\lambda_N, \mu_N}^{-1} \left(1 - \frac{1}{N} \right) \right\}. \quad (26)$$

To show that convergence in mean does also hold in our case, we let $\hat{X}_N(1, \frac{\mu_N}{\lambda_N}) = \frac{\log(X_N^*(1, \frac{\mu_N}{\lambda_N}))}{\log \log(N)} 1_{A_N}$. It is enough to show that the collection of random variables $\left\{ \hat{X}_N(1, \frac{\mu_N}{\lambda_N}) \right\}_{N=1}^{\infty}$ is uniformly integrable, *i.e.*,
$$\lim_{C' \rightarrow \infty} \sup_{N \geq 1} E_{h,g} \left[\left| \hat{X}_N(1, \frac{\mu_N}{\lambda_N}) \right| 1_{\left\{ |\hat{X}_N(1, \frac{\mu_N}{\lambda_N})| \geq C' \right\}} \right] = 0$$
 to conclude the proof. We can upper bound the random variable $\frac{\log(X_N^*(1, \frac{\mu_N}{\lambda_N}))}{\log \log(N)} 1_{A_N}$ as $\frac{\log(X_N^*(1, \frac{\mu_N}{\lambda_N}))}{\log \log(N)} 1_{A_N} \leq \frac{\log(X_N^*(1, \frac{\mu_N}{\lambda_N}))}{\log \log(N)} 1_{\{X_N^*(1, \frac{\mu_N}{\lambda_N}) \geq 1\}}$. Using proof techniques similar to those used in the proof of Lemma 3 in [12], it can be shown that $\left\{ \frac{\log(X_N^*(1, \frac{\mu_N}{\lambda_N}))}{\log \log(N)} 1_{\{X_N^*(1, \frac{\mu_N}{\lambda_N}) \geq 1\}} \right\}_{N=1}^{\infty}$ is uniformly integrable, which implies the uniform integrability of $\left\{ \frac{\log(X_N^*(1, \frac{\mu_N}{\lambda_N}))}{\log \log(N)} 1_{A_N} \right\}_{N=1}^{\infty}$.

APPENDIX C PROOF OF THEOREM 3

Note that $R_{\text{DTPIL}}^*(p_N^*, N) \geq R_{\text{DTPIL}}^*(\frac{1}{N}, N)$. Hence, it is enough to show that $\limsup_{N \rightarrow \infty} \frac{R_{\text{DTPIL}}^*(p_N^*, N)}{\log \log(N)} \leq \frac{1}{en_h}$. To this end, let $\tilde{X}_N = \frac{\log(X_N^*(\lambda_N, \mu_N))}{\log \log(N)}$, where $X_N^*(\lambda, \mu)$ is defined as in Lemma 4. For all $\epsilon > 0$, we have (27).

As in the proof of Theorem 2, we have $\tilde{X}_N 1_{\{X_N^*(\lambda_N, \mu_N) \geq 1\}} \xrightarrow{i.p.} \frac{1}{n_h}$ and $1_{\left\{ \left| \tilde{X}_N - \frac{1}{n_h} \right| > \epsilon \right\}} \xrightarrow{i.p.} 0$ as N tends to infinity. Hence, $\tilde{X}_N 1_{\{X_N^*(\lambda_N, \mu_N) \geq 1\}} 1_{\left\{ \left| \tilde{X}_N - \frac{1}{n_h} \right| > \epsilon \right\}}$ converges to zero in probability. Using techniques similar to those used in the proof of Theorem 2, it can also be shown that the collection of random variables $\left\{ \tilde{X}_N 1_{\{X_N^*(\lambda_N, \mu_N) \geq 1\}} 1_{\left\{ \left| \tilde{X}_N - \frac{1}{n_h} \right| > \epsilon \right\}} \right\}_{N=1}^{\infty}$ is uniformly integrable, which implies that $\lim_{N \rightarrow \infty} E_{h,g} \left[\tilde{X}_N 1_{\{X_N^*(\lambda_N, \mu_N) \geq 1\}} 1_{\left\{ \left| \tilde{X}_N - \frac{1}{n_h} \right| > \epsilon \right\}} \right] = 0$.

For N large enough, $\Pr(A_N)$ can be upper bounded as

$$\begin{aligned} \Pr(A_N) &= N p_N^* (1 - p_N^*)^{N-1} \\ &\stackrel{(a)}{\leq} \left(1 - \frac{1}{N}\right)^{N-1}, \end{aligned}$$

where (a) follows from the fact that $N p_N^* (1 - p_N^*)^{N-1}$ is maximized at $p_N^* = \frac{1}{N}$. Hence,

$$\limsup_{N \rightarrow \infty} \frac{R_{\text{DTPIL}}^*(p_N^*, N)}{\log \log(N)} \leq \frac{1}{en_h} + \frac{\epsilon}{e},$$

which completes the proof since ϵ is arbitrary.

APPENDIX D PROOF OF LEMMA 1

Assume that a is a limit point of the sequence $a_N = N p_N^*$, where $N \geq 1$. We only consider the case $a \in (0, \infty)$. For $a = 0$, it can be shown that the probability of successful transmission, and hence the secondary network throughput, goes to zero due to lack of enough transmission attempts. For $a = \infty$, the probability of successful transmission, and

hence the secondary network throughput, goes to zero due to excessive simultaneous transmission attempts.

Let N_j be a subsequence of N such that $\lim_{N_j \rightarrow \infty} a_{N_j} = a$. As argued in the proof of Theorem 2, it can be shown that the probability of successful transmission on this subsequence converges to $\frac{a}{e^a}$, *i.e.*, $\lim_{N_j \rightarrow \infty} \Pr(A_{N_j}) = \frac{a}{e^a}$. Hence, using techniques similar to those employed in the proof of Theorem 2, we can further show that $\lim_{N_j \rightarrow \infty} \frac{R_{\text{DTPIL}}^*(p_{N_j}^*, N_j)}{\log \log(N_j)} = \frac{a}{e^a n_h}$, which is maximized at $a = 1$. This implies that p_N^* must be chosen such that $\lim_{N \rightarrow \infty} N p_N^* = 1$ to obtain optimal secondary network throughput scaling behavior.

APPENDIX E THROUGHPUT SCALING IN DIL NETWORKS

To obtain the throughput scaling behavior in DIL networks, we will first provide a preliminary lemma establishing the convergence behavior of μ_N . This lemma will also be helpful to study the effect of average total interference power, Q_{ave} , on the secondary network throughput in DIL networks.

Lemma 6: Let μ_N be the power control parameter in DIL networks, as stated in Theorem 4. Then, $\lim_{N \rightarrow \infty} \mu_N = \frac{1}{Q_{\text{ave}}}$.

Proof: First, we show that μ_N is upper bounded by $\frac{1}{Q_{\text{ave}}}$ for all N . To this end, we have

$$\begin{aligned} Q_{\text{ave}} &= E_{h,g} \left[\sum_{i=1}^N g_i \left(\frac{1}{\mu_N g_i} - \frac{1}{h_i} \right)^+ 1_{\left\{ \frac{h_i}{g_i} > F^{-1}\left(1 - \frac{1}{N}\right) \right\}} \right] \\ &\leq E_{h,g} \left[\sum_{i=1}^N \frac{1}{\mu_N} 1_{\left\{ \frac{h_i}{g_i} > F^{-1}\left(1 - \frac{1}{N}\right) \right\}} \right] \\ &= \frac{1}{\mu_N}, \end{aligned}$$

which implies that $\mu_N \leq \frac{1}{Q_{\text{ave}}}$. Hence, to complete the proof, it is enough to show that $\liminf_{N \rightarrow \infty} \mu_N \geq \frac{1}{Q_{\text{ave}}}$. We can lower bound μ_N as (28). Since μ_N is bounded above by $\frac{1}{Q_{\text{ave}}}$ and $F_{\frac{h}{g}}^{-1}\left(1 - \frac{1}{N}\right)$ tends to infinity as N grows large, (28) implies that $\liminf_{N \rightarrow \infty} \mu_N = \frac{1}{Q_{\text{ave}}}$. ■

Now, we are ready to prove Theorem 5. The sum-rate under the optimum distributed power control in DIL networks can be written as (29).

It is easy to see that $\lim_{N \rightarrow \infty} \Pr(B_N) = \frac{1}{e}$. Thus, the first term on the right-hand side of (29) converges to $\frac{1}{e} \log(Q_{\text{ave}})$ as N tends to infinity, which indicates the logarithmic effect of the average interference power constraint, Q_{ave} , on the secondary network sum-rate in DIL networks. It can also be shown that $\frac{\log(Y_N^*)}{\log(N)} \xrightarrow{i.p.} \frac{1}{\gamma_g}$ (*i.e.*, see Lemma 8 in [12]) and $1_{B_N} \xrightarrow{i.d.} \text{Bern}\left(\frac{1}{e}\right)$ as N tends to infinity. Therefore, using Slutsky's theorem, we have $\frac{\log(Y_N^*)}{\log(N)} 1_{B_N} \xrightarrow{i.d.} \text{Bern}\left(\frac{1}{e}\right)$ as N grows large. Since convergence in distribution does not always imply convergence in mean, we need to show that the collection of random variables $\left\{ \frac{\log(Y_N^*)}{\log(N)} 1_{B_N} \right\}_{N=1}^{\infty}$ is uniformly integrable. For N large enough, we have $\frac{\log(Y_N^*)}{\log(N)} 1_{B_N} \leq \frac{\log(Y_N^*)}{\log(N)} 1_{\{Y_N^* \geq 1\}}$. Using Lemma 8 in [12], we conclude that $\left\{ \frac{\log(Y_N^*)}{\log(N)} 1_{\{Y_N^* \geq 1\}} \right\}_{N=1}^{\infty}$ is uniformly integrable, which im-

$$\begin{aligned} \frac{R_{\text{DTPL}}^*(p_N^*, N)}{\log \log(N)} &= \mathbf{E}_{\mathbf{h}, \mathbf{g}} \left[\tilde{X}_N \mathbf{1}_{A_N} \mathbf{1}_{\{|\tilde{X}_N - \frac{1}{n_h}| > \epsilon\}} \right] + \mathbf{E}_{\mathbf{h}, \mathbf{g}} \left[\tilde{X}_N \mathbf{1}_{A_N} \mathbf{1}_{\{|\tilde{X}_N - \frac{1}{n_h}| \leq \epsilon\}} \right] \\ &\leq \mathbf{E}_{\mathbf{h}, \mathbf{g}} \left[\tilde{X}_N \mathbf{1}_{\{X_N^*(\lambda_N, \mu_N) \geq 1\}} \mathbf{1}_{\{|\tilde{X}_N - \frac{1}{n_h}| > \epsilon\}} \right] + \left(\frac{1}{n_h} + \epsilon \right) \Pr(A_N). \end{aligned} \quad (27)$$

$$\begin{aligned} \mu_N &= \frac{1}{Q_{\text{ave}}} \mathbf{E}_{\mathbf{h}, \mathbf{g}} \left[\sum_{i=1}^N \left(1 - \frac{g_i \mu_N}{h_i} \right)^+ \mathbf{1}_{\left\{ \frac{h_i}{g_i} > F_{\frac{h}{g}}^{-1} \left(1 - \frac{1}{N} \right) \right\}} \right] \\ &\geq \frac{1}{Q_{\text{ave}}} - \frac{\mu_N}{Q_{\text{ave}}} N \mathbf{E}_{h_1, g_1} \left[\frac{g_1}{h_1} \mathbf{1}_{\left\{ \frac{h_1}{g_1} > F_{\frac{h}{g}}^{-1} \left(1 - \frac{1}{N} \right) \right\}} \right] \\ &\geq \frac{1}{Q_{\text{ave}}} - \frac{\mu_N}{Q_{\text{ave}} F_{\frac{h}{g}}^{-1} \left(1 - \frac{1}{N} \right)}. \end{aligned} \quad (28)$$

$$R_{\text{DIL}}^* \left(\frac{1}{N}, N \right) = \log \left(\frac{1}{\mu_N} \right) \Pr(B_N) + \mathbf{E}_{\mathbf{h}, \mathbf{g}} [\log(Y_N^*) \mathbf{1}_{B_N}]. \quad (29)$$

plies uniform integrability of $\left\{ \frac{\log(Y_N^*)}{\log(N)} \mathbf{1}_{B_N} \right\}_{N=1}^{\infty}$. Hence, we have $\lim_{N \rightarrow \infty} \mathbf{E}_{\mathbf{h}, \mathbf{g}} \left[\frac{\log(Y_N^*)}{\log(N)} \mathbf{1}_{B_N} \right] = \frac{1}{e^{\gamma_g}}$, which concludes the proof.

APPENDIX F

PROOF OF THEOREM 6

Since $R_{\text{DIL}}^*(p_N^*, N) \geq R_{\text{DIL}}^*\left(\frac{1}{N}, N\right)$, we have $\liminf_{N \rightarrow \infty} \frac{R_{\text{DIL}}^*(p_N^*, N)}{\log(N)} \geq \frac{1}{e^{\gamma_g}}$. To show the other direction, let $\tilde{Y}_N = \frac{\log\left(\frac{Y_N^*}{\mu_N}\right)}{\log(N)}$. For all $\epsilon > 0$, we have (30).

Recall from the proof of Theorem 5 that $\frac{Y_N^*}{\log(N)}$ converges in probability to $\frac{1}{\gamma_g}$. This implies that $\tilde{Y}_N \mathbf{1}_{\{Y_N^* \geq \mu_N\}} \xrightarrow{i.p.} \frac{1}{\gamma_g}$ and $\mathbf{1}_{\{|\tilde{Y}_N - \frac{1}{\gamma_g}| > \epsilon\}} \xrightarrow{i.p.} 0$ as N tends to infinity. Thus, $\tilde{Y}_N \mathbf{1}_{\{Y_N^* \geq \mu_N\}} \mathbf{1}_{\{|\tilde{Y}_N - \frac{1}{\gamma_g}| > \epsilon\}}$ converges in probability to 0. Using techniques similar to those used in the proof of Theorem 5, we can show that the collection of random variables $\left\{ \tilde{Y}_N \mathbf{1}_{\{Y_N^* \geq \mu_N\}} \mathbf{1}_{\{|\tilde{Y}_N - \frac{1}{\gamma_g}| > \epsilon\}} \right\}_{N=1}^{\infty}$ is uniformly integrable. This implies that $\lim_{N \rightarrow \infty} \mathbf{E}_{\mathbf{h}, \mathbf{g}} \left[\tilde{Y}_N \mathbf{1}_{\{Y_N^* \geq \mu_N\}} \mathbf{1}_{\{|\tilde{Y}_N - \frac{1}{\gamma_g}| > \epsilon\}} \right] = 0$.

For N large enough, $\Pr(B_N)$ can be upper bounded as

$$\begin{aligned} \Pr(B_N) &= N p_N^* (1 - p_N^*)^{N-1} \\ &\leq \left(1 - \frac{1}{N} \right)^{N-1}. \end{aligned}$$

Hence, $\limsup_{N \rightarrow \infty} \frac{R_{\text{DIL}}^*(p_N^*, N)}{\log(N)} \leq \frac{1}{e^{\gamma_g}} + \frac{\epsilon}{e}$ which completes the proof since ϵ is arbitrary.

APPENDIX G

PROOF OF LEMMA 2

The proof of this lemma is similar to the proof of Lemma 1. Assume that a is a limit point of the sequence $a_N = N p_N^*$, $N \geq 1$, and let N_j be a subsequence of N achieving a . For

$a = 0$, it can be shown that the probability of successful transmission, and hence the secondary network throughput, goes to zero due to lack of enough transmission attempts. For $a = \infty$, the probability of successful transmission, and hence the secondary network throughput, goes to zero due to excessive simultaneous transmission attempts.

For $a \in (0, \infty)$, as argued in the proof of Theorem 5, it can be shown that the probability of successful transmission on N_j converges to $\frac{a}{e^a}$, which, in turn, leads to $\lim_{N_j \rightarrow \infty} \frac{R_{\text{DIL}}^*(p_{N_j}^*, N_j)}{\log(N_j)} = \frac{a}{e^a}$. Since $\frac{a}{e^a}$ is maximized at $a = 1$, we have $\lim_{N \rightarrow \infty} N p_N^* = 1$.

REFERENCES

- [1] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: an information theoretic perspective," *Proc. IEEE*, vol. 97, no. 5, pp. 894–914, May 2009.
- [2] FCC Spectrum Policy Task Force, "Report of the spectrum efficiency working group," Tech. Rep. 02-135, Nov. 2002. Available: http://www.fcc.gov/sptf/files/SEWGFinalReport_1.pdf.
- [3] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: a survey," *Computer Networks*, vol. 50, no. 13, pp. 2127–2159, Sept. 2006.
- [4] B. Wang and K. J. R. Liu, "Advances in cognitive radio networks: a survey," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 1, pp. 5–23, Feb. 2011.
- [5] Y. Liang, K. Chen, G. Y. Li, and P. Mahonen, "Cognitive radio networking and communications: an overview," *IEEE Trans. Veh. Technol.*, vol. 60, no. 7, pp. 3386–3407, Sept. 2011.
- [6] C. Claudia and K. R. Chowdhury, "A survey on MAC protocols for cognitive radio networks," *Ad Hoc Netw.* no. 7, pp. 1315–1329, 2009.
- [7] J. Mitola III and G. Q. Maguire Jr., "Cognitive radio: making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.
- [8] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.
- [9] 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.
- [10] R. Zhang, S. Cui, and Y.-C. Liang, "On ergodic sum capacity of fading cognitive multiple-access and broadcast channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5161–5178, Nov. 2009.
- [11] E. Nekouei, H. Inaltekin, and S. Dey, "Throughput scaling in cognitive multiple access with average power and interference constraints," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 927–946, Feb. 2012.

$$\begin{aligned} \frac{R_{\text{DIL}}^*(p_N^*, N)}{\log(N)} &= \mathbf{E}_{h,g} \left[\tilde{Y}_N \mathbf{1}_{B_N} \mathbf{1}_{\left\{ \left| \tilde{Y}_N - \frac{1}{\gamma_g} \right| > \epsilon \right\}} \right] + \mathbf{E}_{h,g} \left[\tilde{Y}_N \mathbf{1}_{B_N} \mathbf{1}_{\left\{ \left| \tilde{Y}_N - \frac{1}{\gamma_g} \right| \leq \epsilon \right\}} \right] \\ &\leq \mathbf{E}_{h,g} \left[\tilde{Y}_N \mathbf{1}_{\{Y_N^* \geq \mu_N\}} \mathbf{1}_{\left\{ \left| \tilde{Y}_N - \frac{1}{\gamma_g} \right| > \epsilon \right\}} \right] + \left(\frac{1}{\gamma_g} + \epsilon \right) \Pr(B_N). \end{aligned} \quad (30)$$

- [12] E. Nekouei, H. Inaltekin, and S. Dey, "Multiuser diversity for the cognitive uplink with generalized fading and reduced primary's cooperation." Available: <http://arxiv.org/abs/1209.1424>, Tech. Rep., Sept. 2012.
- [13] H. Inaltekin and S. V. Hanly, "Optimality of binary power control for the single cell uplink," *IEEE Trans. Inf. Theory*, vol. 58, no. 10, pp. 6484–6498, Oct. 2012.
- [14] T. W. Ban, W. Choi, B. C. Jung, and D. K. Sung, "Multi-user diversity in a spectrum sharing system," *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 102–106, Jan. 2009.
- [15] R. Zhang and Y.-C. Liang, "Investigation on multiuser diversity in spectrum sharing based cognitive radio networks," *IEEE Commun. Lett.*, vol. 14, no. 2, pp. 133–135, Feb. 2010.
- [16] L. Yang and A. Nosratinia, "Hybrid opportunistic scheduling in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 1, pp. 328–337, Jan. 2012.
- [17] L. Yang and A. Nosratinia, "Capacity limits of multiuser multi-antenna cognitive networks," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4493–4508, Jul. 2012.
- [18] H. Wang, J. Lee, S. Kim, and D. Hong, "Capacity of secondary users exploiting multispectrum and multiuser diversity in spectrum-sharing environments," *IEEE Trans. Veh. Technol.*, vol. 59, no. 2, pp. 1030–1036, Feb. 2010.
- [19] A. Tajer and X. Wang, "Multiuser diversity gain in cognitive networks," *IEEE/ACM Trans. Netw.*, vol. 18, no. 6, pp. 1766–1779, Dec. 2010.
- [20] Y. J. Zhang and A. M. So, "Optimal spectrum sharing in MIMO cognitive radio networks via semidefinite programming," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 2, pp. 362–373, Feb. 2011.
- [21] F. Gao, R. Zhang, Y. Liang, and X. Wang, "Design of learning-based MIMO cognitive radio systems," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1707–1720, May 2010.
- [22] X. Qin and R. A. Berry, "Distributed approaches for exploiting multiuser diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 392–413, Feb. 2006.
- [23] E. Nekouei, H. Inaltekin, and S. Dey, "Distributed cognitive multiple access networks: power control, scheduling and multiuser diversity," in *Proc. 2013 IEEE International Symposium on Information Theory*.
- [24] D. N. C. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [25] H. Boche and M. Wiczanowski, "Optimization-theoretic analysis of stability-optimal transmission policy for multiple-antenna multiple-access channel," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2688–2702, June 2007.
- [26] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels*, 2nd ed. Wiley-Interscience, 2005.
- [27] G. Stüber, *Principles of Mobile Communication*. Kluwer Academic Publishers, 1996.
- [28] H. Bertoni, "Coverage prediction for mobile radio systems operating in the 800/900 MHz frequency range," *IEEE Trans. Veh. Technol.*, vol. 37, no. 1, pp. 57–60, Feb. 1988.
- [29] J. G. Andrews, A. Ghosh, and R. Muhamed, *Fundamentals of WiMAX*. Prentice Hall, 2007.
- [30] W. Yu, "Uplink-downlink duality via minimax duality," *IEEE Trans. Inf. Theory*, vol. 52, pp. 361–374, Feb. 2006.
- [31] G. A. Gupta and S. Toumpis, "Power allocation over parallel Gaussian multiple access and broadcast channels," *IEEE Trans. Inf. Theory*, vol. 52, pp. 3274–3282, Jul. 2006.
- [32] D. G. Luenberger, *Optimization by Vector Space Methods*. Wiley, 1968.
- [33] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [34] A. Sarantsev, "Tail asymptotic of sum and product of random variables with applications in the theory of extremes of conditionally Gaussian processes," Master thesis, Lomonosov Moscow State University, Moscow,

July, 2010. Available: <http://arxiv.org/pdf/1107.3869v1.pdf>.

- [35] A. Gut, *Probability: A Graduate Course*. Springer, 2005.
- [36] P. Billingsley, *Probability and Measure*, 3rd ed. John Wiley and Sons, 1995.
- [37] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, 6th ed. Academic Press, 2000.



Ehsan Nekouei (S'11) received B.S. from Shahid Bahonar University of Kerman, Kerman, Iran, in 2003 and M.S. from TarbiatModares University, Tehran, Iran, in 2006. He received Ph.D. degree in Electrical Engineering from University of Melbourne in July 2013. His research interests include communications and information theory, mechanism design theory and smart grid.



Hazer Inaltekin (S'04, M'06) is an Assistant Professor of Electrical and Electronics Engineering at Antalya International University. He received his B.S. degree (High Hons) in electrical and electronics engineering from Bogazici University, Istanbul, Turkey, in 2001, and his M.S./Ph.D. degree in electrical and computer engineering from Cornell University, Ithaca, NY, in 2006. He was a Post-doctoral Research Associate at Cornell University from 2006 to 2007, and at Princeton University, Princeton, NJ, from 2007 to 2009. In 2009, he joined

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



Subhrakanti Dey (SM'06) was born in India in 1968. He received the B.Tech. and M.Tech. degrees from the Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology, Kharagpur, India, in 1991 and 1993, respectively, and the Ph.D. degree from the Department of Systems Engineering, Research School of Information Sciences and Engineering, Australian National University, Canberra, Australia, in 1996.

Since June 2013, he is a Professor in the Division of Signals and Systems, Dept. of Eng. Sciences, Uppsala University in Sweden. Prior to this, he was with the Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, Australia, since February 2000, where he was a full Professor since 2007. From September 1995 to September 1997 and September 1998 to February 2000, he was a postdoctoral Research Fellow with the Department of Systems Engineering, Australian National University. From September 1997 to September 1998, he was a post-doctoral Research Associate with the Institute for Systems Research, University of Maryland, College Park. His current research interests include networked control systems, wireless communications and networks, signal processing for sensor networks, and stochastic and adaptive estimation and control. Prof. Dey currently serves on the Editorial Board of *Elsevier Systems and Control Letters*. He was also an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE TRANSACTIONS ON AUTOMATIC CONTROL. He is a Senior Member of IEEE.