Throughput and Delay Scaling of General Cognitive Networks

Wentao Huang, Xinbing Wang

Department of Electronic Engineering, Shanghai Jiao Tong University, China

{yelohuang,xwang8}@sjtu.edu.cn

Abstract—There has been recent interest within the networking research community to understand how performance scales in cognitive networks with overlapping n primary nodes and m secondary nodes. Two important metrics, i.e., throughput and delay, are studied in this paper. We first propose a simple and extendable decision model, i.e., the hybrid protocol model, for the secondary nodes to exploit spatial gap among primary transmissions for frequency reuse. Then a framework for general cognitive networks is established based on the hybrid protocol model to analyze the occurrence of transmission opportunities for secondary nodes. We show that in the case that the transmission range of the secondary network is smaller than that of the primary network in order, as long as the primary network operates in a generalized round-robin TDMA fashion, the hybrid protocol model suffice to guide the secondary network to achieve the same throughput and delay scaling as a standalone network, without harming the transmissions of the primary network. Our approach is general in the sense that we only make a few weak assumptions on both networks, and therefore obtain a wide variety of results. We show secondary networks can obtain the same order of throughput and delay as standalone networks when primary networks are classic static networks, networks with random walk mobility, hybrid networks, multicast networks, hierarchically cooperative networks or clustered networks. Our work presents a relatively complete picture of the performance scaling of cognitive networks and provides fundamental insight on the design of them.

I. Introduction

The electromagnetic radio spectrum is a natural resource, the use of which by transmitters and receivers is licensed by governments. Today, as wireless applications demand ever more bandwidth, efficient usage of spectrum is becoming necessary. However, recent measurement [1] observed a severe under-utilization of the licensed spectrum, implying the non-optimality of the current scheme of spectra management. As a remedy, the Federal Communications Commission (FCC) has recently recommended [1], [2] more flexibility in spectrum assignment so that new regulations would allow for devices which are able to sense and adapt to their spectral environment, such as *cognitive radios*, to become *secondary* or *cognitive users*. Cognitive users could opportunistically access the spectrum originally licensed to *primary users*, in a manner that their transmissions will not affect the performance of primary users. Primary users have a higher priority to the spectrum; they may be legacy devices and may not cooperate with secondary users. The overlapping *primary network* and *secondary network* together form the *cognitive network*.

This paper focuses on the performance scaling analysis of cognitive networks. The fundamental scaling laws in ad hoc networks has attracted tremendous interest in the networking community for long. This track of research is initiated by Gupta and Kumar, whose landmark paper [3] showed that generally, the per-node throughput capacity of a wireless ad hoc network with n users only scales as $O(1/\sqrt{n})$[1]. Following works have covered a wide variety of ad hoc networks with different features, such as mobile ad hoc networks (MANETs) [5], hybrid networks [6], [7], multicast networks [8], [9], hierarchically cooperative networks [10], clustered networks [11], [12], etc. Performance metrics other than capacity are also studied, among which delay and its optimal tradeoff with throughput are of critical importance [4], [13].

As most related works, under Gaussian channel model, Jeon et al. [14] considered the capacity scaling of a cognitive network where the number of secondary users, m, is larger than n in order. Under similar assumption, Yin et al. [15] developed the throughput-delay tradeoff of both primary and secondary networks and Wang et al. [16] studied the cases of multicast traffic pattern. Interestingly, all these works showed that both primary and secondary networks can achieve similar or same performance bounds as they are standalone networks.

All previous works on cognitive networks [14], [15], [16] considered some particular scenarios. They first assumed some particular primary networks with specific scheduling and routing protocols, then proposed the communication schemes for secondary users accordingly, and lastly showed such schemes suffice to achieve the same performance bounds as standalone networks. However, a key principle of cognitive networks is that primary users are spectrum license holders and may operate at their own will without considering secondary nodes. Therefore, though assuming a specific primary network can simplify the problem, the results will heavily depend on the communication schemes of the primary network, which is often unmanageable.

That motivates us to study a general cognitive network in this paper. Our major contributions are three folds. First, we characterize the regime that cognitive networks can achieve the same order of throughput and delay scaling as standalone networks. Secondly, we propose a simple decision model for secondary users to identify transmission opportunities and based on it establish a framework with which schemes of standalone networks can be readily extended to secondary

---

[1]We use the standard order notations [4].

networks. Thirdly, we apply the framework to various specific scenarios and show that secondary networks can obtain the same order of throughput and delay as standalone networks when primary networks are classic static networks, networks with random walk mobility, hybrid networks, multicast networks, hierarchically cooperative networks or clustered networks.

In particular, the following conditions are sufficient for a general cognitive network to achieve the same throughput and delay bounds as standalone networks.

A1) The cognitive network is subject to the physical interference model. The primary network operates at a SINR level larger than the threshold for successful reception by some small allowance.

A2) The primary network is scheduled in a generalized round-robin TDMA manner.

A3) Scheduling schemes of secondary network follow $r_{\max} = o(R_{\min})$ and $r_{\max}^{\gamma-2} = o(R_{\min}^{\gamma}/R_{\max}^2)$ with high probability, where R and r are the transmission ranges of primary and secondary networks, and γ is the path loss exponent.

Intuitively, condition A1 ensures that primary transmission links are neither too dense nor too vulnerable so that there exist opportunities for secondary users. Such opportunities will frequently appear, as a consequence of A2. The first equation of A3 is the generalization of the condition $m = \omega(n)$ in related works, while the second equation is more technical. It characterizes the case that the scheduling of primary networks is somewhat "homogeneous" such that there exists a simple rule for opportunity decision.

We note this paper is not merely a generalization of results from previous works. Our work shows the fact that cognitive networks, and especially secondary networks, can achieve the same throughput and delay scaling as standalone networks, is mainly determined by the underlying interference model, and only weakly relies on the specific settings such as scheduling and routing protocols of primary networks. Such insight is fundamental and implies that for quite general cases, "cognitive" will not be a handicap to performance scaling.

The paper is organized as follows. In Section 2 we introduce system models and formalize the operation rules of cognitive networks. We propose the hybrid protocol model and establish its physical feasibility in Section 3. Section 4 identifies the conditions under which the secondary network will have plenty transmission opportunities if scheduled according to the hybrid protocol model. We apply the general results to several specific cognitive networks in Section 5, and Section 6 concludes the paper.

II. SYSTEM MODEL

Throughout this paper we denote the probability of an event E as $\Pr(E)$ and say E happens with high probability (*w.h.p.*) if $\lim_{n\to\infty} \Pr(E) = 1$. A number of parameters and constants will be needed and by convention we use $\{c_i\}$ to denote some positive constants and $\{C_j\}$ some parameters dependent on n.

A. Network Topology

We define the network extension $\mathcal{O}$ to be a unit square. Two kinds of nodes, i.e., the *primary nodes* and the *secondary nodes*, overlap in $\mathcal{O}$. They share the same time, space and frequency dimensions. Unless further specifications are made, by convention we assume n primary nodes are independently and identically distributed (i.i.d.) in $\mathcal{O}$ according to uniform distribution, and so do the m secondary users. Their positions are $\{X_i\}_{i=1}^n$ and $\{Y_j\}_{j=1}^m$. $\forall i, j, X_i, Y_j \in \mathcal{O}$. At times we may denote a node by its position, i.e., we refer to primary node i and secondary node j as X_i and Y_j, and let $|X_i - Y_j|$ be the distance between them. Two types of nodes form their respective networks, the *primary network* and the *secondary network*. In each network nodes are randomly grouped into source-destination (S-D) pairs, such that every node is both source and destination, with traffic rate λ. Equivalently we can describe the traffic pattern in matrix form $\lambda\Lambda$, where $\Lambda = [\lambda_{sd}]$ is a random permutation matrix [2] with $\lambda_{sd} \in \{0, 1\}$. Note that we do not consider cross-network traffic. We use index p and s to distinguish quantities between primary nodes and secondary nodes when needed, for example, λ_p and λ_s.

B. Communication Model

We assume all nodes share a wireless channel with bandwidth W bps. Assume that path loss exponent is $\gamma > 2$, then the normalized channel gain is $G(|X - Y|) = |X - Y|^{-\gamma}$. Besides, wireless transmission may be subject to failures or collisions caused by noise or interference. To judge whether a direct wireless link is feasible, we have the following physical model, whose well-known prototype is proposed in [3]:

The Physical Model: Let $\{X_i; i \in \mathcal{T}^{(p)}\}$ and $\{Y_j; j \in \mathcal{T}^{(s)}\}$ be the subsets of nodes simultaneously transmitting at some time instant. Let P be the uniform power level of primary network, and P_j be the power chosen by secondary node Y_j, for $j \in \mathcal{T}^{(s)}$. Then, For primary network, the transmission from node X_i is successfully received by node X_j if

$$\frac{PG(|X_i - X_j|)}{N + \sum\limits_{k \in \mathcal{T}^{(p)} \setminus \{i\}} PG(|X_k - X_j|) + \sum\limits_{l \in \mathcal{T}^{(s)}} P_l G(|Y_l - X_j|)} \geq \alpha \quad (1)$$

where N is ambient noise and constant α characterize the minimum signal-to-interference-plus-noise ratio (SINR) necessary for successful receptions for primary nodes. For secondary network, the transmission from node Y_i is successfully received by node Y_j if

$$\frac{P_i G(|Y_i - Y_j|)}{N + \sum\limits_{k \in \mathcal{T}^{(s)} \setminus \{i\}} P_k G(|Y_k - Y_j|) + \sum\limits_{l \in \mathcal{T}^{(p)}} PG(|X_l - Y_j|)} \geq \beta$$

where constant β is the minimum required SINR for secondary network. Note that we allow secondary users to have more flexible power control ability. This is in accord with the design principle of cognitive radios.

---

[2] $\Lambda = [\lambda_{sd}]$ is a permutation matrix if $\forall s, d, \lambda_{sd} \in \{0, 1\}; \forall d, \sum_s \lambda_{sd} = 1; \forall s, \sum_d \lambda_{sd} = 1$

We call a couple of nodes a link if they form a transmitter-receiver pair, e.g., (X_i, X_j). Given an interference model, in general there is a number of subsets of links that can be active simultaneously. We call such subsets of links *feasible states*, and define the set of all feasible states as *feasible family*. We use $\mathscr{P}(\alpha, \beta)$ to denote the feasible family of physical model.

C. Operation Rules

The operation rules are the key that make cognitive networks different from normal ad hoc networks. Though primary and secondary users overlap and share the channel, they are different essentially because of their behavior. In principle, primary nodes are spectrum license holders and have the priority to access the channel. It is followed by two important implications. First, primary nodes may operate at their own will without considering secondary nodes. They may be legacy devices running on legacy protocols, which are fixed and unmanageable. Therefore the assumptions made about primary networks should be as few and general as possible. Besides, the secondary network, which is opportunistic in nature, should control its interference to the primary network and prevent deteriorating the performance of primary users. The challenge is, primary scheduler may not alter its protocol due to the existence of secondary network and its decision model could be different from physical model (1), i.e., the interference term from secondary network in the denominator is not available. But in order to leave some margin for secondary nodes, it is necessary for the decision model to operate at a SINR larger than α by an allowance ϵ.

Operation Rule 1. *Decision model for primary network:* The primary scheduler considers the transmission from X_i to X_j to be feasible if:

$$\frac{PG(|X_i - X_j|)}{N + \sum_{\substack{k \in \mathcal{T}^{(p)} \\ k \neq i}} PG(|X_k - X_j|)} \geq \alpha + \epsilon$$

The feasible family of primary decision model is denoted as $\mathscr{D}(\alpha + \epsilon)$.

Then, as the operation rule, secondary nodes should guarantee that feasible state under decision model $\mathscr{D}$ above should be indeed feasible under physical model.

Operation Rule 2. Let $\mathcal{S}^{(p)}$ and $\mathcal{S}^{(s)}$ be the sets of active primary links and active secondary links. If $\mathcal{S}^{(p)} \in \mathscr{D}(\alpha + \epsilon)$, then $\mathcal{S}^{(p)} \cup \mathcal{S}^{(s)} \in \mathscr{P}(\alpha, \beta)$, *w.h.p.*.

D. Capacity Definition

Definition 1. *Feasible throughput*: Per-node throughput $g(n)$ of primary network is said to be feasible if there exists a spatial and temporal scheme for scheduling transmissions, such that by operating the primary network in a multi-hop fashion and buffering at intermediate nodes when awaiting transmission opportunities, every primary source can send $g(n)$ bps to its destination on average.

Definition 2. *Asymptotic per-node capacity* $\lambda_p(n)$ of the primary network is said to be $\Theta(g(n))$ if there exist two

TABLE I: Important Notations

Notation	Definition
X_i	position of primary user i
Y_j	position of secondary user j
$\mathscr{P}(\alpha, \beta)$	feasible family of physical model
$\mathscr{D}(\alpha + \epsilon)$	feasible family of primary scheduler
$\mathscr{Q}_p(\Delta_p), \mathscr{Q}_s(\Delta_s)$	feasible family of protocol model
$\mathscr{H}(\Delta_p, \Delta_{ps}, \Delta_{sp}, \Delta_s)$	feasible family of hybrid protocol model
$\mathcal{S}^{(p)}$	set of active primary links
$\mathcal{S}^{(s)}$	set of active secondary links
$\mathcal{S}$	$\mathcal{S}^{(p)} \cup \mathcal{S}^{(s)}$
R_i	Tx range of active link $(X_i, X_{\text{Rx}(i)})$
r_j	Tx range of active link $(Y_j, Y_{\text{Rx}(j)})$
P	Tx power of primary network
P_j	Tx power of link $(Y_j, Y_{\text{Rx}(j)})$

positive constant c and c' such that:

$$\begin{cases} \lim_{n \to \infty} \Pr\{\lambda_p(n) = cg(n) \text{ is feasible}\} = 1 \\ \lim_{n \to \infty} \Pr\{\lambda_p(n) = c'g(n) \text{ is feasible}\} < 1 \end{cases}$$

Similarly we can define the asymptotic per-node capacity $\lambda_s(m)$ of the secondary network.

III. IDENTIFYING OPPORTUNITIES: THE HYBRID PROTOCOL MODEL

In this section we consider the problem of how to schedule links in the cognitive network under interference constraint. Recall from operation rules that primary nodes are unmanageable, so in fact the key issue is the schedule strategy for the secondary network. In specific, we will face two challenges: first, how to ensure that secondary transmissions are harmless to the primary network and; secondly, how to establish a secondary link given uncontrollable interference from the primary network. Our goal is to design a practical decision model for the secondary users to address these two seemingly contradictory challenges at the same time. Intuitively, that is to say we should find simple rules for secondary nodes to hunt and exploit opportunities in the network.

A. Hybrid Protocol Model

Since we assume the primary network to be a general network which operates according to decision model $\mathscr{D}(\alpha + \epsilon)$, it is our starting point. $\mathscr{D}$ is of physical concern and cares about the aggregate interference and SINR, but the following lemma relates it to a simpler pairwise model. This alternative model is known as *protocol model* in literature and often plays the role as interference model. But here we use it as a tool to characterize the relative position of active primary nodes.

Definition 3. *Protocol Model* for primary network: A transmission from X_i to X_j is feasible if

$$|X_k - X_j| \geq (1 + \Delta_p)|X_i - X_j|, \quad \forall k \in \mathcal{T}^{(p)}$$

where Δ_p defines the guard zone for the primary network. The corresponding feasible family is noted as $\mathscr{Q}_p(\Delta_p)$. Likewise we define protocol model $\mathscr{Q}_s(\Delta_s)$ for the secondary network.

Lemma 1. If $\mathcal{S}^{(p)} \in \mathscr{D}(\alpha + \epsilon)$ and $\Delta_p \leq (\alpha + \epsilon)^{\frac{1}{\gamma}} - 1$, then $\mathcal{S}^{(p)} \in \mathscr{Q}_p(\Delta_p)$.

    Proof: Let $\mathcal{S}^{(p)} \in \mathscr{D}$, $\forall (X_i, X_j) \in \mathcal{S}^{(p)}$ and $\forall k \neq i, k \in \mathcal{T}^{(p)}$, holds,

$$\frac{|X_i - X_j|^{-\gamma}}{|X_k - X_j|^{-\gamma}} \geq \frac{P|X_i - X_j|^{-\gamma}}{N + P \sum_{\substack{l \in \mathcal{T}^{(p)} \\ l \neq i}} |X_l - X_j|^{-\gamma}} \geq \alpha + \epsilon$$

therefore $|X_k - X_j| \geq (\alpha + \epsilon)^{\frac{1}{\gamma}}|X_i - X_j|$, set $\Delta_p \leq (\alpha + \epsilon)^{\frac{1}{\gamma}} - 1$, then $\mathcal{S}^{(p)} \in \mathscr{Q}_p$. ∎

Since $\mathscr{Q}_p \supseteq \mathscr{D}$, i.e., $\mathscr{Q}_p$ captures all degrees of freedom of the primary network, and considering the simplicity of the form of protocol model, it motivates us to define a new *hybrid protocol model* $\mathscr{H}$ based on $\mathscr{Q}_p$ and $\mathscr{Q}_s$, to be the decision model for secondary network.

Definition 4. *The Hybrid Protocol Model* with feasible family $\mathscr{H}(\Delta_p, \Delta_{ps}, \Delta_{sp}, \Delta_s)$: $\forall \mathcal{S} \in \mathscr{H}$, let $\mathcal{S}^{(p)} = \{(X_i, X_j) \in \mathcal{S}\}$ and $\mathcal{S}^{(s)} = \{(Y_i, Y_j) \in \mathcal{S}\}$, then $\mathcal{S}^{(p)} \in \mathscr{Q}_p(\Delta_p)$, $\mathcal{S}^{(s)} \in \mathscr{Q}_s(\Delta_s)$. Further, $\forall (X_i, X_j) \in \mathcal{S}^{(p)}$,

$$|Y_k - X_j| \geq (1 + \Delta_{sp})|X_i - X_j|, \forall k \in \mathcal{T}^{(s)} \quad (2)$$

and $\forall (Y_i, Y_j) \in \mathcal{S}^{(s)}$,

$$|X_k - Y_j| \geq (1 + \Delta_{ps})|Y_i - Y_j|, \forall k \in \mathcal{T}^{(p)} \quad (3)$$

where Δ_{sp} and Δ_{ps} define the inter-network guard zone.

The hybrid protocol model only depends on pairwise distance between transmitters and receivers. Such simplicity will facilitate our analysis in the next section. Besides, it is compatible with the classic protocol interference model. Thus rich communication schemes and results based on protocol model can be easily extended to cognitive networks, as will be shown in Section 5.

In the following we should prove that if $\mathscr{H}$ is used as decision model for secondary nodes, it will comply with Operation Rule 2. This involves correctly tuning the parameters Δ_p, Δ_{ps}, Δ_{sp}, Δ_s and $\{P_j\}_{j \in \mathcal{T}^{(s)}}$.

B. Interference at Primary Nodes

We first address the challenge that primary transmissions should not be interrupted by secondary nodes. The main task is to bound the interference from the secondary network. We start with a useful property of the hybrid protocol model.

Lemma 2. Given arbitrary $Z_i, Z_j, Z_k, Z_l \in \mathcal{O}$, if (Z_i, Z_j), (Z_k, Z_l) are active links (primary or secondary), and $|Z_k - Z_j| \geq (1 + \Delta_1)|Z_i - Z_j|$, $|Z_i - Z_l| \geq (1 + \Delta_2)|Z_k - Z_l|$, then the $\Delta_1|Z_i - Z_j|/2$ and $\Delta_2|Z_k - Z_l|/2$ neighborhood of the line segment joining Z_i, Z_j and Z_l, Z_k are disjoint.

    Proof: Let E and F be two arbitrary points on line segment Z_iZ_j and Z_kZ_l, by triangle inequity,

$$|Z_i - E| + |E - F| + |F - Z_l| + |Z_k - F| +$$
$$|F - E| + |E - Z_j| \geq |Z_i - Z_l| + |Z_k - Z_j|$$

Since Z_i, E, Z_j are collinear, substituting lemma condition

$$|Z_i - Z_j| + |Z_k - Z_l| + 2|E - F| \geq |Z_i - Z_l| + |Z_k - Z_j|$$
$$\geq (1 + \Delta_2)|Z_k - Z_l| + (1 + \Delta_1)|Z_i - Z_j|$$

therefore $|E - F| \geq \frac{\Delta_1}{2}|Z_i - Z_j| + \frac{\Delta_2}{2}|Z_k - Z_l|$. Now we prove the lemma by contradiction. Suppose the two neighborhoods overlap, then there exist point X on Z_iZ_j, Y on Z_kZ_l and Z such that $|Z - X| < \frac{\Delta_1}{2}|Z_i - Z_j|$ and $|Z - Y| < \frac{\Delta_2}{2}|Z_k - Z_l|$. Then $|X - Y| < |Z - X| + |Z - Y| < \frac{\Delta_1}{2}|Z_i - Z_j| + \frac{\Delta_2}{2}|Z_k - Z_l|$, which is a contradiction. ∎

Corollary 1. Under hybrid protocol model,

- If (X_i, X_j) and (X_k, X_l) are active primary links, the $\Delta_p|X_i - X_j|/2$ neighborhood of line segment X_iX_j and $\Delta_p|X_k - X_l|/2$ neighborhood of X_kX_l are disjoint.
- If (Y_i, Y_j) and (Y_k, Y_l) are active secondary links, the $\Delta_s|Y_i - Y_j|/2$ neighborhood of line segment Y_iY_j and $\Delta_s|Y_k - Y_l|/2$ neighborhood of Y_kY_l are disjoint.
- If (X_i, X_j) is active primary link and (Y_k, Y_l) is active secondary link, the $\Delta_{sp}|X_i - X_j|/2$ neighborhood of line segment X_iX_j and $\Delta_{ps}|Y_k - Y_l|/2$ neighborhood of Y_kY_l are disjoint.

For active link $(X_i, X_{\text{Rx}(i)})$ and $(Y_j, Y_{\text{Rx}(j)})$, where function Rx indicates the index of receiver, let $R_i = |X_i - X_{\text{Rx}(i)}|$ and $r_j = |Y_j - Y_{\text{Rx}(j)}|$. Let $R_{\max} = \max R_i$, $R_{\min} = \min R_i$, $r_{\max} = \max r_j$ and $r_{\min} = \min r_j$. We say the secondary network adopts *power assignment scheme* $\mathcal{A}(C)$ if for $i \in \mathcal{T}^{(s)}, P_i = Cr_i^2 P$.

Theorem 1. Under power assignment $\mathcal{A}(C_1)$ and hybrid protocol model, if $\Delta_{ps} > \Delta_s$, then for any active primary link $(X_i, X_{\text{Rx}(i)})$, the interference suffered by $X_{\text{Rx}(i)}$ from the secondary network is upper bounded by $C_2 R_i^{2-\gamma} P$, for some $C_2 = \Theta(C_1)$.

    Proof: Let $B(X, r)$ be the disk centered at X with radius r. Then all $B(Y_j, \Delta_s r_j/2), j \in \mathcal{T}^{(s)}$ should be mutually disjoint according to Corollary 1. As well, $B(Y_j, \Delta_{ps} r_j/2), j \in \mathcal{T}^{(s)}$ are disjoint with $B(X_{\text{Rx}(i)}, \Delta_{sp} R_i/2)$. Since $\Delta_{ps} > \Delta_s$, then all $B(Y_j, \Delta_s r_j/2), j \in \mathcal{T}^{(s)}$, $B(X_{\text{Rx}(i)}, \Delta_{sp} R_i/2)$ are pairwise disjoint. Denote $D_{ij} = B(X_{\text{Rx}(i)}, |X_{\text{Rx}(i)} - Y_j|) \cap B(Y_j, \Delta_s r_j/2)$, it is clear that all D_{ij} are disjoint (See Figure 1). Denote by E, F the two points where $B(X_{\text{Rx}(i)}, |X_{\text{Rx}(i)} - Y_j|)$ intersects $B(Y_j, \Delta_s r_j/2)$. It is clear that $\angle FY_jX_{\text{Rx}(i)} = \angle EY_jX_{\text{Rx}(i)} \geq \pi/3$ because $|X_{\text{Rx}(i)} - Y_j| > \Delta_s r_j/2$. So the area of D_{ij} is at least one third of $B(Y_j, \Delta_s r_j/2)$. Let $I_{sp}(i)$ denote the interference at receiver $X_{\text{Rx}(i)}$ from secondary network, and $\mathrm{d}A$ be the area element,

$$I_{sp}(i) = \sum_{j \in \mathcal{T}^{(s)}} \frac{P_j}{|Y_j - X_{\text{Rx}(i)}|^\gamma}$$
$$= \sum_{j \in \mathcal{T}^{(s)}} \frac{4C_1 P}{\pi \Delta_s^2} \int_{B(Y_j, \Delta_s r_j/2)} \frac{\mathrm{d}A}{|Y_j - X_{\text{Rx}(i)}|^\gamma}$$
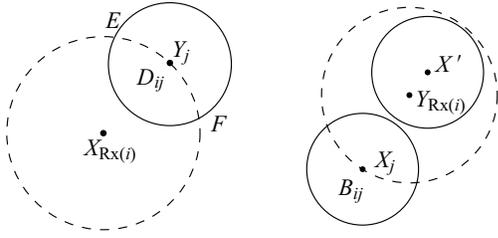$$\leq \sum_{j \in \mathcal{T}^{(s)}} \frac{12C_1 P}{\pi \Delta_s^2} \int_{D_{ij}} \frac{\mathrm{d}A}{|Y_j - X_{\text{Rx}(i)}|^\gamma}$$

Fig. 1: Analyzing the interference. Left plot shows an example for D_{ij} and right plot for B_{ij}.

$$\leq \sum_{j \in \mathcal{T}^{(s)}} \frac{12C_1 P}{\pi \Delta_s^2} \int_{D_{ij}} \frac{dA}{|X - X_{\mathrm{Rx}(i)}|^\gamma}$$

(where X is the position vector of dA)

$$\leq \frac{12C_1 P}{\pi \Delta_s^2} \int_{\cup_{j \in \mathcal{T}^{(s)}} D_{ij}} \frac{dA}{|X - X_{\mathrm{Rx}(i)}|^\gamma}$$

Since $(\cup_{j \in \mathcal{T}^{(s)}} D_{ij}) \cap B(X_{\mathrm{Rx}(i)}, \Delta_{sp} R_i/2) = \emptyset$, we have,

$$I_{sp}(i) \leq \frac{12C_1 P}{\pi \Delta_s^2} \int_{|X - X_{\mathrm{Rx}(i)}| \geq \Delta_{sp} R_i/2} \frac{dA}{|X - X_{\mathrm{Rx}(i)}|^\gamma}$$
$$= \frac{12C_1 P}{\pi \Delta_s^2} \int_{\Delta_{sp} R_i/2}^\infty \frac{2\pi r \, dr}{r^\gamma}$$
$$= \frac{24C_1 P}{\Delta_s^2(\gamma - 2)} \left(\frac{2}{\Delta_{sp} R_i}\right)^{\gamma - 2} = C_2 P R_i^{2 - \gamma}$$

∎

C. Interference at Secondary Nodes

Now we focus on the interference at secondary nodes. The main challenge is to bound the uncontrollable interference from the primary network.

Theorem 2. Under power assignment $\mathcal{A}(C_1)$ and hybrid protocol model, for any active link $(Y_i, Y_{\mathrm{Rx}(i)})$, the interference at $Y_{\mathrm{Rx}(i)}$ from the primary network is upper bounded by $c_3 P R_{\min}^{-\gamma}$, for some constant c_3.

Proof: Denote by $I_{ps}(i) = \sum_{j \in \mathcal{T}^{(p)}} \frac{P}{|X_j - Y_{\mathrm{Rx}(i)}|^\gamma}$ the interference at $Y_{\mathrm{Rx}(i)}$ from primary network. Pick X' as the interfering primary transmitter closest to $Y_{\mathrm{Rx}(i)}$. From Corollary 1, distance between any primary transmitter and $Y_{\mathrm{Rx}(i)}$ should be larger than $\Delta_{sp} R_{\min}/2 + \Delta_{ps} r_i/2$; distance between any two primary transmitter is larger than $\Delta_p R_{\min}$.

Now consider the case that $\Delta_{sp} < \Delta_p$. (Note that if $\Delta_{sp} > \Delta_p$, $I_{ps}(i)$ will be smaller and the upper bound still holds.) Then all $X_j, j \in \mathcal{T}^{(p)}$ is at least $\Delta_p R_{\min}/2$ away from $Y_{\mathrm{Rx}(i)}$ except X'. First consider the interference contributed by X',

$$\frac{P}{|X' - Y_{\mathrm{Rx}(i)}|^\gamma} \leq P \left(\frac{\Delta_{sp}}{2} R_{\min}\right)^{-\gamma}$$

Next consider the interference from some other primary transmitter X_j. Let $B_{ij} = B(X_j, \Delta_p R_{\min}/2) \cap B(Y_{\mathrm{Rx}(i)},$ $|X_j - Y_{\mathrm{Rx}(i)}|)^c$, as shown in Figure 1, then,

$$\frac{P}{|X_j - Y_{\mathrm{Rx}(i)}|^\gamma} \leq \int_{B_{ij}} \frac{P}{|B_{ij}|} \frac{dA}{|X_j - Y_{\mathrm{Rx}(i)}|^\gamma}$$

(where $|B_{ij}|$ is the area of disk B_{ij})

$$\leq \int_{B_{ij}} \frac{P}{|B_{\min}|} \frac{dA}{|X_j - Y_{\mathrm{Rx}(i)}|^\gamma}$$

(where $|B_{\min}| = \min_j |B_{ij}|$)

$$\leq \int_{B_{ij}} \frac{P}{|B_{\min}|} \frac{dA}{(\frac{1}{2}|X - Y_{\mathrm{Rx}(i)}|)^\gamma}$$

(because $|X_j - Y_{\mathrm{Rx}(i)}| > |\Delta_p R_{\min}/2|$)

$$= \frac{2^\gamma P}{|B_{\min}|} \int_{B_{ij}} \frac{dA}{|X - Y_{\mathrm{Rx}(i)}|^\gamma}$$
$$\leq \frac{2^{\gamma + 3} P}{\pi \Delta_p^2 R_{\min}^2} \int_{B_{ij}} \frac{dA}{|X - Y_{\mathrm{Rx}(i)}|^\gamma}$$

To sum up, let $I'_{ps}(i) = \sum_{j \in \mathcal{T}^{(p)} \setminus \{X'\}} \frac{P}{|X_j - Y_{\mathrm{Rx}(i)}|^\gamma}$,

$$I'_{ps}(i) \leq \frac{2^{\gamma + 3} P}{\pi \Delta_p^2 R_{\min}^2} \sum_{j \in \mathcal{T}^{(p)} \setminus \{X'\}} \int_{B_{ij}} \frac{dA}{|X - Y_{\mathrm{Rx}(i)}|^\gamma}$$
$$= \frac{2^{\gamma + 3} P}{\pi \Delta_p^2 R_{\min}^2} \int_{\cup_{j \in \mathcal{T}^{(p)} \setminus \{X'\}} B_{ij}} \frac{dA}{|X - Y_{\mathrm{Rx}(i)}|^\gamma}$$

(since $\{B_{ij}\}$ are disjoint)

$$\leq \frac{2^{\gamma + 3} P}{\pi \Delta_p^2 R_{\min}^2} \int_{|X - Y_{\mathrm{Rx}(i)}| > \frac{\Delta_p R_{\min}}{2}} \frac{dA}{|X - Y_{\mathrm{Rx}(i)}|^\gamma}$$
$$= \frac{2^{\gamma + 3} P}{\pi \Delta_p^2 R_{\min}^2} \int_{\frac{\Delta_p R_{\min}}{2}}^\infty \frac{2\pi r \, dr}{r^\gamma}$$
$$= \frac{2^{2\gamma + 2} P}{(\gamma - 2) \Delta_p^\gamma} R_{\min}^{-\gamma}$$

Combining the contribution from X',

$$I_{ps}(i) \leq \left(\frac{2^\gamma}{\Delta_{sp}^\gamma} + \frac{2^{2\gamma + 2}}{(\gamma - 2) \Delta_p^\gamma}\right) P R_{\min}^{-\gamma}$$

∎

We should also take into account the interference between secondary links. Power assignment scheme $\mathcal{A}$ is well designed so that it not only restricts the interference from the secondary network to the primary network, but also that between secondary links, as shown by the next theorem. Its proof is similar to Theorem 1 and is omitted for space concern.

Theorem 3. Under power assignment scheme $\mathcal{A}(C_1)$ and hybrid protocol model, for any active secondary link $(Y_i, Y_{\mathrm{Rx}(i)})$, the interference at $Y_{\mathrm{Rx}(i)}$ from all other simultaneously active secondary links is upper bounded by $I_{ss}(i) \leq C_4 P r_i^{2 - \gamma}$, where $C_4 = \frac{24 \cdot 2^{\gamma - 2}}{(\gamma - 2) \Delta_s^\gamma} C_1$.

D. Physical Feasibility of the Hybrid Protocol Model

Last we show under appropriate conditions, hybrid protocol model is indeed physically feasible. We begin with a lemma.

Lemma 3. Given $A, B, C, a, b > 0$, if $\frac{C}{A} \leq \frac{b}{a(a + b)}$ and $\frac{A}{B} \geq a + b$, then $\frac{A}{B + C} \geq a$.

Then first consider the primary network.

Lemma 4. If $\mathcal{S}^{(p)} \in \mathscr{D}(\alpha + \epsilon)$, $(X_i, X_{\mathrm{Rx}(i)}) \in \mathcal{S}^{(p)}$, and $I_{sp}(i) \le PR_i^{-\gamma} \frac{\epsilon}{\alpha(\alpha+\epsilon)}$, then $(X_i, X_{\mathrm{Rx}(i)})$ is feasible under physical model $\mathscr{P}(\alpha, \beta)$. That is to say, $\frac{PR_i^{-\gamma}}{N+I_{sp}(i)+I_{pp}(i)} \ge \alpha$, where $I_{pp}(i)$ is interference at $X_{\mathrm{Rx}(i)}$ from other simultaneously active primary links.

Proof: Let $A = PR_i^{-\gamma}$, $B = N + I_{pp}(i)$, $C = I_{sp}(i)$, $a = \alpha$, $b = \epsilon$. Note that $\mathcal{S}^{(p)} \in \mathscr{D}(\alpha+\epsilon)$ implies $A/B \ge a+b$, then the result holds from Lemma 3. ∎

Lemma 5. If $\Delta_{ps} > \Delta_s$, $\mathcal{S} \in \mathscr{H}$ and $\mathcal{S}^{(p)} \in \mathscr{D}(\alpha + \epsilon)$, then under power assignment $\mathcal{A}(C_1)$ such that $C_2 \le \frac{\epsilon}{\alpha(\alpha+\epsilon)R_{\max}^2}$, all primary links are feasible under physical model $\mathscr{P}(\alpha, \beta)$.

Proof: For any $(X_i, X_{\mathrm{Rx}(i)}) \in \mathcal{S}^{(p)}$, from Theorem 1, $I_{sp}(i) \le C_2 R_i^{2-\gamma} P$, thus the SIR at $X_{\mathrm{Rx}(i)}$ satisfies,

$$\frac{PR_i^{-\gamma}}{I_{sp}(i)} \ge \frac{PR_i^{-\gamma}}{C_2 PR_i^{2-\gamma}} \ge \frac{1}{C_2 R_{\max}^2} \ge \frac{\alpha(\alpha+\epsilon)}{\epsilon}$$

Then from Lemma 4 we have the assertion. ∎

Now turn to the secondary network. Similarly, it can be shown that Lemma 6 follows from Theorem 2 and Lemma 7 from Theorem 3,

Lemma 6. Under power assignment $\mathcal{A}(C_1)$ with $C_1 \ge \frac{c_3 c_5 R_{\min}^{-\gamma}}{r_i^{2-\gamma}}$, if $\mathcal{S}^{(p)} \in \mathscr{H}$, then for any $(Y_i, Y_{\mathrm{Rx}(i)}), i \in \mathcal{T}^{(s)}$, it holds:

$$\frac{C_1 Pr_i^{2-\gamma}}{I_{ps}(i)} \ge c_5 \qquad (4)$$

Lemma 7. Under the condition of Lemma 6, if $\Delta_s \ge \left(48 \frac{2^{\gamma-2}}{\gamma-2} c_6\right)^{\frac{1}{\gamma}}$, for any $(Y_i, Y_{\mathrm{Rx}(i)}), i \in \mathcal{T}^{(s)}$, follows:

$$\frac{C_1 Pr_i^{2-\gamma}}{N + I_{ss}(i)} \ge c_6$$

Then we are ready to prove the final result.

Theorem 4. If $r_{\max}^{\gamma-2} = o\left(\frac{R_{\min}^\gamma}{R_{\max}^2}\right)$, $\Delta_p \le (\alpha + \epsilon)^{\frac{1}{\gamma}} - 1$, and $\Delta_{ps} \ge \Delta_s \ge \left(24 \frac{2^\gamma}{\gamma-2} \beta\right)^{\frac{1}{\gamma}}$, then there exists power assignment $\mathcal{A}(C_1)$, such that for any $\mathcal{S}^{(p)} \in \mathscr{D}(\alpha + \epsilon)$, holds $\mathcal{S}^{(p)} \in \mathscr{H}(\Delta_p, \Delta_{ps}, \Delta_{sp}, \Delta_s)$. And if we schedule secondary network transmissions in the way such that $\mathcal{S}^{(p)} \cup \mathcal{S}^{(s)} \in \mathscr{H}$, holds $\mathcal{S}^{(p)} \cup \mathcal{S}^{(s)} \in \mathscr{P}(\alpha, \beta)$.

Proof: The first claim follows from Lemma 1. To prove the second claim, first notice every active primary link is physically feasible if the condition of Lemma 5 is verified, i.e.,

$$C_2 = \Theta(C_1) = o\left(1/R_{\max}^2\right) \qquad (5)$$

On the other hand, consider the secondary network, if

$$C_1 = \omega\left(R_{\min}^{-\gamma}/r_i^{2-\gamma}\right) \qquad (6)$$

then according to Lemma 6, (4) holds for any positive constant c_5. In combination with Lemma 7, it is clear that SINR at any secondary receiver is greater than $\frac{1}{2}(\min\{c_5, c_6\}) = \beta$. Since

$r_{\max}^{\gamma-2} = o\left(\frac{R_{\min}^\gamma}{R_{\max}^2}\right)$, we can indeed find C_1 and C_2, such that (5) and (6) hold, proving the theorem. ∎

The condition $r_{\max}^{\gamma-2} = o(R_{\min}^\gamma/R_{\max}^2)$ characterizes the regime that primary links are homogeneous in range. In other words, if this condition fails, it implies that the scheduling of the primary network is somewhat "chaotic" and simple decision model like $\mathscr{H}$ does not suffice to identify transmission opportunities. Fortunately, this condition usually holds because $R_{\max}$ and $R_{\min}$ typically do not differ much in order and we tend to employ a small r.

IV. AVAILABILITY OF TRANSMISSION OPPORTUNITIES

Section III addresses the problem of how to identify transmission chances for secondary networks: given a set $\mathcal{S}^{(p)}$ of simultaneously active primary links, we allow a set $\mathcal{S}^{(s)}$ of simultaneously active secondary links according to hybrid protocol model $\mathscr{H}$. This section, on the other hand, considers the problem that for those secondary links which desire to transmit, how frequently do these chances occur. Of particular interest is to compare this result with an identical standalone network. Standalone networks provide trivial performance upper bounds since cognitive secondary networks will suffer from additional transmission constraints imposed by primary networks. To alleviate the performance loss due to such constraints, it is intuitive that one should reduce the range of secondary links, and this fact is indeed verified by hybrid protocol model and Theorem 4. This section will further show if $r_{\max} = o(R_{\min})$, then for quite general cases, such performance loss is insignificant and has no impact in order sense. In other words, all secondary links have a constant ratio of time to be *unconstrained* as if they were in a standalone network. Besides, note it is well known that to achieve better scalability, a smaller range is also favorable. This coincidence implies that secondary networks can reach the optimal scaling performance of a standalone network. Now we formally introduce the concept of unconstraint, and analyze the unconstrained time in the following subsections.

Definition 5. Given arbitrary $\mathcal{S}_{\mathrm{s.a.}}^{(s)} \in \mathscr{Q}_s(\Delta_s)$ and arbitrary $\mathcal{S}^{(p)} \in \mathscr{Q}_p(\Delta_p)$, there exists an unique maximal $\mathcal{S}^{(s)} \subset \mathcal{S}_{\mathrm{s.a.}}^{(s)}$ such that $\mathcal{S}^{(p)} \cup \mathcal{S}^{(s)} \in \mathscr{H}(\Delta_p, \Delta_{ps}, \Delta_{sp}, \Delta_p)$. We say a link $(Y_i, Y_{\mathrm{Rx}(i)}) \in \mathcal{S}_{\mathrm{s.a.}}^{(s)}$ is *unconstrained* if $(Y_i, Y_{\mathrm{Rx}(i)}) \in \mathcal{S}^{(s)}$.

Note the fraction of time that the link is constrained characterizes the performance loss relative to the corresponding standalone network.

A. Cell Partitioning Round-Robin Mode

We consider the case that primary networks operate according to a common scheduling paradigm: the cell partitioning round-robin active scheme. It first spatially tessellates the network into cells, then assigns color to each cell, such that cells with the same color, if limit their transmissions to neighbors, will not interfere with each other. Then we allow cells with the same color to transmit simultaneously, and let different colors take turns to be active. A simple TDMA scheme will suffice. This very widely employed scheme [4],

[13], [3], [7] features a high degree of spatial concurrency and thus frequency reuse. It is deterministic and therefore simple. To the best of our knowledge, all previous works on asymptotic analysis of cognitive networks focused on some particular variants of such TDMA sheme. We now show for a generic primary scheduling policy which operates in the round-robin fashion, the unconstrained time fraction for any short range secondary link is constant, as a simple consequence of the hybrid protocol model.

Definition 6. A *network tessellation* is a set of disjoint cells $\{V_i \subset \mathcal{O}\}$. A *round-robin TDMA scheme* is a scheduling scheme that i) tessellates the network into cells such that every cell is contained in a disk of radius $\rho(n)$, ii) allows non-interfering cells to be simultaneously active and transmit to neighbor cells, where two cells V_i, V_j are non-interfering if $\sup\{|E - F| : E \in V_i, F \in V_j\} \geq (2 + \Delta_p)4\rho(n)$, and iii) activates different groups of cells in a round-robin TDMA fashion, and guarantees every cell can be active for at least c_7 fraction of time in one round, for some constant $c_7 > 0$.

The existence of round-robin TDMA schemes is a consequence of the well-known theorem about vertex coloring of graphs. The next theorem shows such scheme is favorable to secondary networks because it deterministically ensures transmission opportunities not only for every primary cell, but also for every secondary link.

Theorem 5. If the primary network operates according to a round-robin TDMA scheme and $\Delta_p > 2$, $\Delta_{sp} \leq \frac{\Delta_p - 2}{2}$, then every secondary link with range $r = o(R_{\min})$ has at least c_7 fraction of time to be unconstrained in one round.

Proof: Consider a generic link $(Y_i, Y_{\mathrm{Rx}(i)})$, pick a point X such that $|X - Y_i| = (4 + 2\Delta_p)\rho(n)$, and denote by V the cell X belongs to, we claim whenever V is scheduled to be active, $(Y_i, Y_{\mathrm{Rx}(i)})$ is unconstrained. To that end, we first verify transmitter Y_i will not upset transmissions in V. Indeed, any point E belongs to V should lie within distance $2\rho(n)$ from X, thus any point F belongs to a neighbor cell of V should lie within distance $4\rho(n)$ from X, then

$$
\begin{aligned}
|Y_i - F| &\geq (4 + 2\Delta_p)\rho(n) - 4\rho(n) \\
&= 2\Delta_p\rho(n) \geq (1 + \Delta_{sp})4\rho(n) \\
&\geq (1 + \Delta_{sp})|E - F| \quad\quad (7)
\end{aligned}
$$

Now consider another simultaneous active cell V', it is clear that any point $X' \in V'$ is at least $(2 + \Delta_p)4\rho(n)$ away from X, then $|X' - Y_i| \geq (4 + 2\Delta_p)\rho(n) = |X - Y_i|$. Together with (7), condition (2) is verified. Besides, since $r = o(R_{\min})$, condition (3) is obvious. This completes the proof. ∎

We observe that under hybrid protocol model $\mathscr{H}$, $\Delta_p > 2$ is critical to guarantee transmission opportunities for secondary nodes, as shown in Theorem 5. Equivalently, it implies $\alpha + \epsilon \geq 2\gamma$. This is an assumption about primary networks and we assume it always holds from now on. However, we conjecture this assumption is not fundamental and can be relaxed by introducing a criterion with more flexible form, i.e., allowing Δ_{sp} and Δ_{ps} to be dependent on n. Such decision models

may have a better capability of digging into the potential of available gaps, at the cost of complexity. We leave for future work a more in-depth analysis of such cases.

We are ready to show that cognitive networks can achieve the same throughput and delay scaling as standalone networks.

Theorem 6. If the primary network operates in the round-robin TDMA or the independent relay fashion, for any protocol interference model based scheme that schedules and routes the secondary network such that $r_{\max} = o(R_{\min})$, $r_{\max}^{\gamma-2} = o(R_{\min}^\gamma / R_{\max}^2)$ w.h.p., and achieves per-node throughput λ_s and delay D_s in the case that secondary network is standalone, there exists a corresponding scheme which can achieve per-node throughput $\Theta(\lambda_s)$ and delay $\Theta(D_s)$ when primary network is present and Operation Rules 1 and 2 apply.

Proof: First we hypothesize the secondary network is standalone, and denote by $\mathfrak{c}_{\mathrm{s.a.}}^{ij}$ the throughput rate of link (Y_i, Y_j), then $\{\mathfrak{c}_{\mathrm{s.a.}}^{ij}\}$ is determined by the scheduling scheme. For example, if we assume slotted time, then a deterministic scheduling scheme is characterized by a sequence $(\mathcal{S}_{\mathrm{s.a.}}^t)_{t=1}^T$, $\mathcal{S}_{\mathrm{s.a.}}^t \in \mathcal{Q}_s$, and

$$
\mathfrak{c}_{\mathrm{s.a.}}^{ij} = \frac{W}{T} \sum_{t=1}^T \mathbb{I}_{(Y_i, Y_j) \in \mathcal{S}_{\mathrm{s.a.}}^t}
$$

The network can be mapped to a graph $\mathcal{G}$, where m vertices stand for secondary nodes, and $\{\mathfrak{c}_{\mathrm{s.a.}}^{ij}\}$ compose edges. The network traffic is represented as a multicommodity flow instance on $\mathcal{G}$ [17], and the routing scheme is defined by $\{f_{ij}^{sd}\}$, the average fraction of traffic from Y_s to Y_d which is routed through link (Y_i, Y_j). Because the overall scheme achieves per-node throughput λ_s, holds,

$$
\lambda_s \sum_s \sum_d \lambda_{sd} f_{ij}^{sd} \leq \mathfrak{c}_{\mathrm{s.a.}}^{ij} \quad 1 \leq i, j \leq m
$$

Now let the primary network joins, we stick the secondary network to the prior scheme except for only allowing the unconstrained links to be active. According to Theorem 4, such scheduling is physically feasible. Denote the corresponding throughput rate of link (Y_i, Y_j) as $\mathfrak{c}_{\mathrm{c.r.}}^{ij}$, from Theorem 5, $\mathfrak{c}_{\mathrm{c.r.}}^{ij} \geq c_{12} \mathfrak{c}_{\mathrm{s.a.}}^{ij}$, where constant $c_{12} = \min\{c_7, c_{11}\}$. Let $\lambda_s' = c_{12}\lambda_s = \Theta(\lambda_s)$, follows,

$$
\lambda_s' \sum_s \sum_d \lambda_{sd} f_{ij}^{sd} \leq \mathfrak{c}_{\mathrm{c.r.}}^{ij} \quad 1 \leq i, j \leq m
$$

Therefore no edge is overloaded and throughput λ_s' is feasible.

As to delay, the definition and calculation of it depend on specific network settings, such as packet size or mobility patterns [4], [13]. However, we note that a general baseline in prior works is that per-hop delay for a packet is at least $\Omega(1)$ (this may include transmission delay, queueing delay and delay incurred by mobility, etc.). In secondary networks, packets will suffer from extra delay because at each hop, or at each time they are transmitted by a link, they must wait until the link is unconstrained. According to Theorem 5, such delay penalty is upper bounded by the duration of one round, which is $\Theta(1)$. Therefore the order of overall delay is preserved. ∎

V. Optimal Performance Scaling

This section presents results on optimal performance scaling of specific cognitive networks. With Theorem 6 we can conveniently extend optimal schemes and results of standalone networks to cognitive networks. The optimality is preserved in cognitive networks unless we allow cooperation between primary and secondary nodes, which is beyond the scope of this work. The following two corollaries are straightforward from [4], [13]. For clarity, we assume by convention that the networks are static unless further specifications are made.

Corollary 2. The optimal throughput delay tradeoff is $D_p = \Theta(n\lambda_p)$, $\lambda_p \leq \Theta(1/\sqrt{n})$ for primary network and $D_s = \Theta(m\lambda_s)$, $\Theta(n\lambda_p/m) < \lambda_s \leq \Theta(1/\sqrt{m})$ for secondary network, if $\Theta(1/\sqrt{m}) > \Theta(n\lambda_p/m)$.

Remark 1. In the primary networks (classic static wireless ad hoc networks), optimal throughput delay tradeoff can be achieve by [4, Scheme 1], which is a cell partitioned TDMA scheme with squarelet of side length $l_p(n)$. Then $D_p(n) = \Theta(1/l_p(n))$ and $\lambda_p(n) = \Theta(1/nl_p(n))$. Notice that primary transmission range $w.h.p.$ is $l_p(n)$. Let the secondary networks also follow [4, Scheme 1], with $l_s(n) = o(l_p(n))$, therefore $\lambda_s(n) = \Theta(1/ml_s(n)) = \Omega(n\lambda_p(n)/m)$. On the other hand, $\Theta(1/\sqrt{n})$ and $\Theta(1/\sqrt{m})$ is the maximal achievable throughput in primary and secondary networks, respectively.

Corollary 3. If primary nodes move according to random walk model, then the optimal throughput delay tradeoff for primary network is $D_p = \Theta(n\lambda_p)$ if $\lambda_p \leq \Theta(1/\sqrt{n})$, $D_p = \Theta(n\log n)$ if $\Theta(1/\sqrt{n}) < \lambda_p \leq \Theta(1)$. And the optimal throughput delay tradeoff for secondary network is $D_s = \Theta(m\lambda_s)$, $\Theta(n\min(1/\sqrt{n}, \lambda_p)/m) < \lambda_s \leq \Theta(1/\sqrt{m})$, if $\Theta(1/\sqrt{m}) > \Theta(n\min(1/\sqrt{n}, \lambda_p)/m)$.

Remark 2. Our framework could well accommodate node mobility. Notice that neither the hybrid protocol model nor the unconstraint time analysis rely on mobility pattern. Specifically, if primary users move according to random walk model, optimal throughput delay tradeoff in the primary network can be achieved by cell partitioned TDMA scheme [4, Scheme 3]. Therefore the results in Corollary 3 can be obtained following similar argument as in Remark 1.

We can extend the theorem to other variations of ad hoc networks, such as hybrid networks [7].

Corollary 4. If the primary network is equipped with $k = \Omega(\sqrt{n})$ base stations, the capacity of it is $\lambda_p = \Theta(k/n)$, and the optimal throughput delay tradeoff for the secondary network is $D_s = \Theta(m\lambda_s)$, $\Theta(\sqrt{\lambda_p n}/m) < \lambda_s \leq \Theta(1/\sqrt{m})$, if $\Theta(1/\sqrt{m}) > \Theta(\sqrt{\lambda_p n}/m)$.

Remark 3. Though base stations are wired together, they still need to access the wireless channel to communication with primary users. In these kind of hybrid networks, [7] shows that the throughput capacity can be achieved by tessellating the network into hexagons and assign one base station for each hexagon. The diameter of the hexagon, $\Theta(\sqrt{1/k})$, is the average primary transmission range. Different groups of hexagons are scheduled in a cell partitioned TDMA fashion.

Now we try to apply the results to networks with multicast traffic pattern.

Corollary 5. If each primary user wishes to send information to k independently chosen destinations, with $k = o(n/\log n)$, then the per-node multicast capacity of the primary network is $\lambda_p = \Theta(\frac{1}{\sqrt{kn\log n}})$. The optimal throughput delay tradeoff for secondary network is $D_s = \Theta(m\lambda_s)$, $\Theta(\sqrt{n}/m\sqrt{\log n}) < \lambda_s \leq \Theta(1/\sqrt{m})$, if $m = \Omega(n^M)$ for some constant $M > 1$.

Remark 4. Capacity of the primary network can be achieved by a cell partitioned TDMA scheme [8, Section V]. The primary transmission range $w.h.p.$ is $\Theta(\sqrt{\log n/n})$. Notice that our analysis can be directly extended to multicast traffic pattern since the capacity achieving scheme does not require multi-reception or broadcasting.

A. Gaussian Channel Model

The physical interference model is a fixed rate on-off channel model. An alternative, i.e., the Gaussian channel model, generalizes data rate to be continuous in SINR, based on Shannon's capacity formula for the additive Gaussian noise channel. We now briefly extend our results to this model. Specifically, the communication rate of a primary link (X_i, X_j) is given by

$$W = \log\left(1 + \frac{P|X_i - X_j|^{-\gamma}}{N + I_{pp} + I_{sp}}\right)$$

where I_{pp} and I_{sp} are interference from primary network and secondary network, respectively (cf. Section 3). Similarly we define the channel model for secondary links. We limit our interest to the case that W is bounded between two positive constants, i.e., $W_1 < W < W_2$. We note this is also the most realistic case and suffice to generalize the results of prior works [15], [14]. Let $\lambda_p^{\text{s.a.}}$ be the per-node throughput of the primary network in the absence of the secondary network.

Operation Rule 3. The secondary scheduling should ensure that $\frac{\lambda_p}{\lambda_p^{\text{s.a.}}} \geq 1 - \delta_{\text{loss}}$.

Theorem 7. Theorem 6 holds under Gaussian channel model if we substitute Operation Rule 3 for Operation Rules 1 and 2, with $\delta_{\text{loss}} \in (0, 1)$.

Proof: We claim that Operation Rules 1 and 2 with appropriate parameters are sufficient conditions for Operation Rule 3. Indeed, set $\alpha \leq e^{W_1} - 1$ and $\epsilon \leq ((1 + \alpha) - (1 + \alpha)^{1 - \delta_{\text{loss}}})^{1/1 - \delta_{\text{loss}}}$, it is easy to show that

$$\frac{\log\left(1 + \frac{P|X_i - X_{\text{Rx}(i)}|^{-\gamma}}{N + I_{pp} + I_{sp}}\right)}{\log\left(1 + \frac{P|X_i - X_{\text{Rx}(i)}|^{-\gamma}}{N + I_{pp}}\right)} \geq \frac{\log(1 + \alpha)}{\log(1 + \alpha + \epsilon)} \geq 1 - \delta_{\text{loss}}$$

for any active link $(X_i, X_{\text{Rx}(i)})$. On the other hand, to ensure the scheduling of secondary network under the Gaussian channel model is feasible under $\mathscr{P}$, we set $\beta \leq e^{W_1} - 1$. The rest part of the theorem is obvious. ∎

Based on Theorem 7, we study the case that primary users are capable of performing hierarchical MIMO cooperation.

Corollary 6. If the primary users employ hierarchical cooperation as proposed in [10], for arbitrary $\varepsilon > 0$, there exists a scheme such that the primary per-node capacity is $\lambda_p = \Theta(n^{-\varepsilon})$. The optimal throughput delay tradeoff for secondary network is $D_s = \Theta(m\lambda_s)$, $\Theta(\sqrt{n}/m\sqrt{\log n}) < \lambda_s \leq \Theta(1/\sqrt{m})$, if $m = \Omega(n^M)$ for some constant $M > 1$.

Remark 5. The capacity achieving scheme in the primary network consists of three phases [10, Section III]. Among them the first and the third phase eventually will boil down to a cell partitioned TDMA scheme, with transmission range $w.h.p.$ being $\Theta(1/\sqrt{n})$. Note that long range MIMO links will be activated in phase 2, and during this time all secondary users should be muted.

Next we discuss the class of clustered networks [11], [12].

Corollary 7. If the distribution of the primary users follows shot-noise Cox process, the primary network can achieve capacity following the scheme proposed in [12]. The asymptotic throughput and delay of the secondary network is the same as standalone networks if the communication scheme satisfies the condition of Theorem 7.

Remark 6. Since the model of shot-noise Cox processes and the corresponding results are complicated, we refer interested readers to [11], [12] for details. We emphasize that even in these heterogeneous clustered networks, cell partitioned TDMA schemes are still highly efficient. Though similar to Corollary 6, the cell partitioned scheme may be hierarchical to handle inhomogeneities in the network, it is well compatible with our analysis framework.

Lastly, the results that can be obtained are not limited to the cases listed above. Since our framework only relies on a few general conditions, it is flexible and is able to accommodate various cognitive networks with different specific forms. For instance, one can otherwise let both the networks or only the secondary network be mobile.

VI. Conclusions

This paper studies the throughput and delay scaling of general cognitive networks and characterizes the conditions for them to achieve the same throughput and delay scaling as standalone networks. We propose a hybrid protocol model for secondary nodes to identify transmission opportunities and show that based on it communication schemes of standalone networks can be easily extended to secondary networks, without harming the performance of primary networks. In particular, we show that secondary networks can obtain the same optimal performance as standalone networks when primary networks are classic static networks, networks with random walk mobility, hybrid networks, multicast networks, hierarchically cooperative networks or clustered networks. Our work provides fundamental insight on the understanding and design of cognitive networks.

VII. Acknowledgment

References

[1] Federal Communications Commission Spectrum Policy Task Force, "Report of the spectrum efficiency working group," Nov. 2002.

[2] *Cognitive Radio Technologies Proceeding.* Federal Communications Commission, federal Communications Commission, Cognitive Radio Technologies Proceeding, No. 03-108, http://www.fcc.gov/oet/cognitiveradio/.

[3] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. on Information Theory*, vol. 46, no. 2, pp. 388–404, 2000.

[4] A. El Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Optimal throughput-delay scaling in wireless networks: part i: the fluid model," *IEEE/ACM Trans. on Networking*, vol. 14, no. SI, pp. 2568–2592, 2006.

[5] M. Grossglauser and D. N. C. Tse, "Mobility increases the capacity of ad hoc wireless networks," *IEEE/ACM Trans. on Networking*, vol. 10, no. 4, pp. 477–486, 2002.

[6] A. Agarwal and P. R. Kumar, "Capacity bounds for ad hoc and hybrid wireless networks," *SIGCOMM Comput. Commun. Rev.*, vol. 34, no. 3, pp. 71–81, 2004.

[7] B. Liu, Z. Liu, and D. Towsley, "On the capacity of hybrid wireless networks," in *IEEE INFOCOM*, vol. 2, San Franciso, 2003, pp. 1543–1552.

[8] X.-Y. Li, "Multicast capacity of wireless ad hoc networks," *IEEE/ACM Transactions on Networking*, vol. 17, no. 3, pp. 950–961, 2009.

[9] X.-Y. Li, Y. Liu, S. Li, and S. Tang, "Multicast capacity of wireless ad hoc networks under gaussian channel model," *IEEE/ACM Transactions on Networking*, vol. 18, no. 4, pp. 1145–1157, 2010.

[10] A. zgr, O. Lvłque, and D. N. C. Tse, "Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks," *IEEE Trans. Inf. Theory*, vol. 53, pp. 3549–3572, 2007.

[11] G. Alfano, M. Garetto, and E. Leonardi, "Capacity scaling of wireless networks with inhomogeneous node density: Upper bounds," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1147–1157, 2009.

[12] ——, "Capacity scaling of wireless networks with inhomogeneous node density: Lower bounds," in *IEEE INFOCOM*, 2009, pp. 1890–1898.

[13] A. El Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Optimal throughput-delay scaling in wireless networks-part ii: Constant-size packets," *IEEE Trans. on Information Theory*, vol. 52, no. 11, pp. 5111–5116, 2006.

[14] S.-W. Jeon, N. Devroye, M. Vu, S.-Y. Chung, and V. Tarokh, "Cognitive networks achieve throughput scaling of a homogeneous network," submitted to *IEEE Trans. on Information Theory*, 2009.

[15] C. Yin, L. Gao, and S. Cui, "Scaling laws for overlaid wireless networks: A cognitive radio network vs. a primary network," accepted by *IEEE/ACM Trans. on Networking*, 2011.

[16] C. Wang, S. Tang, X.-Y. Li, and C. Jiang, "Multicast capacity of multihop cognitive networks," in *Proc. of IEEE MASS*, Macau SAR, 2009, pp. 274–283.

[17] Y. Aumann and Y. Rabani, "An o(log k) approximate min-cut max-flow theorem and approximation algorithm," *SIAM J. Comput.*, vol. 27, no. 1, pp. 291–301, 1998.

[18] C. Wang, X.-Y. Li, C. Jiang, S. Tang, Y. Liu, and J. Zhao, "Scaling laws on multicast capacity of large scale wireless networks," in *IEEE INFOCOM*, 2009, pp. 1863–1871.

[19] P. Li and Y. Fang, "The capacity of heterogeneous wireless networks," in *Proc. IEEE INFOCOM*, 2010, pp. 1–9.

[20] P. Li, Y. Fang, and J. Li, "Throughput, delay, and mobility in wireless ad hoc networks," in *Proc. IEEE INFOCOM*, 2010, pp. 1–9.