

Research Article

Virtual Cooperation for Throughput Maximization in Distributed Large-Scale Wireless Networks

Jamshid Abouei,¹ Alireza Bayesteh,² Masoud Ebrahimi,² and Amir K. Khandani²

¹Department of Electrical Engineering, Yazd University, P.O. Box 98195-741, Yazd, Iran

²Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada N2L 3G1

Correspondence should be addressed to Jamshid Abouei, abouei@yazduni.ac.ir

Received 28 May 2010; Revised 12 September 2010; Accepted 29 October 2010

Academic Editor: Robert Schober

Copyright © 2011 Jamshid Abouei et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A distributed wireless network with K links is considered, where the links are partitioned into M clusters each operating in a subchannel with bandwidth W/M . The subchannels are assumed to be orthogonal to each other. A general shadow-fading model described by the probability of shadowing α and the average cross-link gains $\bar{\omega} \leq 1$ is considered. The main goal is to find the maximum network throughput in the asymptotic regime of $K \rightarrow \infty$, which is achieved by: (i) proposing a distributed power allocation strategy, where the objective of each user is to maximize its best estimate (based on its local information) of the average network throughput and (ii) choosing the optimum value for M . In the first part, the network throughput is defined as the *average sum-rate* of the network, which is shown to scale as $\Theta(\log K)$. It is proved that the optimum power allocation strategy for each user for large K is a threshold-based on-off scheme. In the second part, the network throughput is defined as the *guaranteed sum-rate*, when the outage probability approaches zero. It is demonstrated that the on-off power scheme maximizes the throughput, which scales as $(W/\alpha\bar{\omega}) \log K$. Moreover, the optimum spectrum sharing for maximizing the average sum-rate and the guaranteed sum-rate is achieved at $M = 1$.

1. Introduction

A primary challenge in wireless networks is to use available resources efficiently so that the network throughput is maximized. Throughput maximization in multiuser wireless networks has been addressed from different perspectives, resource allocation [1–3], routing by using relay nodes [4], exploiting mobility of the nodes [5], and exploiting channel characteristics (e.g., power decay versus distance law [6–8], geometric path loss and fading [9]).

Among different resource allocation strategies, power and spectrum allocation have long been regarded as efficient tools to mitigate the interference and improve the network throughput. In recent years, power and spectrum allocation schemes have been extensively studied in cellular and multihop wireless networks [1, 2, 10–12]. In [11], the authors provide a comprehensive survey in the area of resource allocation, in particular in the context of spectrum assignment. Much of these works rely on centralized

and cooperative algorithms. Clearly, centralized resource allocation schemes provide a significant improvement in the network throughput over decentralized (distributed) approaches. However, they require extensive knowledge of the network configuration. In particular, when the number of nodes is large, deploying such centralized schemes may not be practically feasible. Due to significant challenges in using centralized approaches, the attention of researchers has been drawn to the decentralized resource allocation schemes [13–18].

In decentralized schemes, the decisions concerning network parameters (e.g., rate and/or power) are made by the individual nodes based on their local information. The local decision parameters that can be used for adjusting the rate are the Signal-to-Interference-plus-Noise Ratio (SINR) and the direct channel gain. Most of the works on decentralized throughput maximization target the SINR parameter by using iterative algorithms [15–17]. This leads to the use of game theory concepts [19] where the main challenge is the

convergence issue. For instance, Etkin et al. [17] develop power and spectrum allocation strategies by using game theory. Under the assumptions of the omniscient nodes and strong interference, the authors show that Frequency Division Multiplexing (FDM) is the optimal scheme in the sense of throughput maximization. They use an iterative algorithm that converges to the optimum power values. In [16], Huang et al. propose an iterative power control algorithm in an ad hoc wireless network, in which receivers broadcast adjacent channel gains and interference prices to optimize the network throughput. However, this algorithm incurs a great amount of overhead in large wireless networks.

A more practical approach is to rely on the channel gains as local decision parameters and avoid iterative schemes. Motivated by this consideration, we study the throughput maximization of a distributed single-hop wireless network with K links, operating in a bandwidth of W . Wireless networks using unlicensed spectrum (e.g., Wi-Fi systems based on IEEE 802.11b standard [20]) are a typical example of such networks. To mitigate the interference, the links are partitioned into a fixed number (M) of clusters, each operating in a subchannel with bandwidth W/M , where the subchannels are orthogonal to each other. The cross-link channel gains are assumed to be Rayleigh-distributed with shadow fading, described by parameters $(\alpha, \bar{\omega})$, where α denotes the probability of shadowing and $\bar{\omega}$ ($\bar{\omega} \leq 1$) represents the statistical average of the Rayleigh distribution. The above configuration differs from the geometric models proposed in [5–8, 21]. Unlike the studies in [14–17] which rely on iterative algorithms using SINR, we assume that each transmitter adjusts its power solely based on its direct channel gain.

If each user maximizes its rate selfishly, the optimum power allocation strategy for all users is to transmit with full power. This strategy results in excessive interference, degrading the average network throughput. To prevent this undesirable effect, one should consider the negative impact of each user's power on other links. A reasonable approach for each user is to choose a noniterative power allocation strategy to maximize its best local estimate of the network throughput. In fact, the network nodes aim to cooperative unselfishly to improve the network throughput. We call this unselfish action in the proposed distributed wireless network as a *virtual cooperation* without broadcasting information from one link to the other links.

The network throughput in this paper is defined in two ways: (i) *average sum-rate* and (ii) *guaranteed sum-rate*. It is established that the average sum-rate in the network scales at most as $\Theta(\log K)$ in the asymptotic case of $K \rightarrow \infty$. This order is achievable by the distributed *threshold-based on-off scheme* (i.e., links with a direct channel gain above certain threshold transmit at full power and the rest remain silent). In addition, the on-off power allocation scheme is always optimal for maximizing the guaranteed sum-rate in the network, which is shown to scale as $(W/\alpha\bar{\omega}) \log K$. These results are different from the result in [22] where the authors use a similar on-off scheme for $M = 1$ and prove its optimality only among all *on-off schemes*, and from that in [18] where the authors use a distributed power allocation for

two users. This work also differs from the studies in [23–25] in terms of the network model. We use a distributed power allocation strategy in a single-hop network, while the studies in [23, 24] consider an ad hoc network model with random connections and relay nodes.

We optimize the average network throughput in terms of the number of the clusters, M . It is proved that the maximum average sum-rate and the guaranteed sum-rate of the network for every value of α and $\bar{\omega}$ are achieved at $M = 1$. In other words, splitting the bandwidth W into M orthogonal subchannels does not increase the throughput.

The rest of the paper is organized as follows. In Section 2, the network model and objectives are described. The distributed on-off power allocation strategy and the network average sum-rate are presented in Section 3. We analyze the network guaranteed sum-rate in Section 4. Finally, in Section 5, an overview of the results and some conclusion remarks are presented.

1.1. Notations. For any functions $f(n)$ and $g(n)$ [26] we have the following:

- (i) $f(n) = O(g(n))$ means that $\lim_{n \rightarrow \infty} |f(n)/g(n)| < \infty$;
- (ii) $f(n) = o(g(n))$ means that $\lim_{n \rightarrow \infty} |f(n)/g(n)| = 0$;
- (iii) $f(n) = \omega(g(n))$ means that $\lim_{n \rightarrow \infty} f(n)/g(n) = \infty$;
- (iv) $f(n) = \Omega(g(n))$ means that $\lim_{n \rightarrow \infty} f(n)/g(n) > 0$;
- (v) $f(n) = \Theta(g(n))$ means that $\lim_{n \rightarrow \infty} f(n)/g(n) = c$, where $0 < c < \infty$;
- (vi) $f(n) \sim g(n)$ means that $\lim_{n \rightarrow \infty} f(n)/g(n) = 1$;
- (vii) $f(n) \lesssim g(n)$ means that $\lim_{n \rightarrow \infty} f(n)/g(n) \leq 1$;
- (viii) $f(n) \approx g(n)$ means that $f(n)$ is approximately equal to $g(n)$, that is, if we replace $f(n)$ by $g(n)$ in the equations, the results still hold.

Throughout the paper, we use $\log(\cdot)$ as the natural logarithm function and $\mathbb{P}\{\cdot\}$ denotes the probability of the given event. Boldface letters denote vectors; and for a random variable x , \bar{x} means $\mathbb{E}[x]$, where $\mathbb{E}[\cdot]$ represents the expectation operator. $\text{RH}(\cdot)$ represents the right hand side of the equations.

2. Network Model and Objectives

2.1. Network Model. In this work, we consider a single-hop wireless network consisting of K pairs of nodes indexed by $\{1, \dots, K\}$, operating in bandwidth W . The term “pair” is used to describe a transmitter and its corresponding receiver, while the term “user” is used only for the transmitter. All the nodes in the network are assumed to have a single antenna. The links are assumed to be randomly divided into M clusters denoted by \mathbb{C}_j , $j = 1, \dots, M$ such that the number of links in all clusters are the same. Without loss of generality, we assume that $\mathbb{C}_j \triangleq \{(j-1)n+1, \dots, jn\}$, where $n \triangleq K/M$ denotes the cardinality of the set \mathbb{C}_j which is assumed to be known to all users. It is assumed that K is divisible by M , and hence, $n = K/M$ is an integer number. To eliminate the mutual interference among the clusters, we assume an

M -dimensional orthogonal coordinate system in which the bandwidth W is split into M disjoint subchannels each with bandwidth W/M . It is assumed that the links in \mathbb{C}_j operate in subchannel j . We also assume that M is fixed, that is, it does not scale with K . The power of Additive White Gaussian Noise (AWGN) at each receiver is $(N_0W)/M$, where N_0 is the noise power spectral density.

The channel model is assumed to be Rayleigh flat fading with the shadowing effect. The channel gain, defined as the square magnitude of the *channel coefficient*, between transmitter k and receiver i is represented by the random variable \mathcal{L}_{ki} . For $k = i$, the *direct channel gain* is defined as $\mathcal{L}_{ki} \triangleq h_{ii}$, where h_{ii} is exponentially distributed with unit mean (and unit variance). For $k \neq i$, the *cross channel gains* are defined based on a shadowing model as follows:

$$\mathcal{L}_{ki} \triangleq \begin{cases} \beta_{ki}h_{ki}, & \text{with probability } \alpha, \\ 0, & \text{with probability } 1 - \alpha, \end{cases} \quad (1)$$

where h_{ki} 's have the same distribution as h_{ii} 's, $0 \leq \alpha \leq 1$ is a fixed parameter, and the random variable β_{ki} , referred to as the *shadowing factor*, is independent of h_{ki} and satisfies the following conditions:

- (i) $\beta_{\min} \leq \beta_{ki} \leq \beta_{\max}$, where $\beta_{\min} > 0$ and β_{\max} is finite;
- (ii) $\mathbb{E}[\beta_{ki}] \triangleq \omega \leq 1$.

It is also assumed that $\{\mathcal{L}_{ki}\}$ and $\{\beta_{ki}\}$ are mutually independent random variables for different (k, i) .

All the channels in the network are assumed to be quasi static block fading, that is, the channel gains remain constant during one block and change independently from block to block. In addition, we assume that each transmitter knows its direct channel gain.

We assume a homogeneous network in the sense that all the links have the same configuration and use the same protocol. We denote the transmit power of user i by p_i , where $p_i \in \mathcal{P} \triangleq [0, \mathbf{P}_{\max}]$. The vector $\mathbf{P}^{(j)} = (p_{(j-1)n+1}, \dots, p_{jn})$ represents the power vector of the users in \mathbb{C}_j . Also, $\mathbf{P}_{-i}^{(j)}$ denotes the vector consisting of elements of $\mathbf{P}^{(j)}$ other than the i th element, $i \in \mathbb{C}_j$. To simplify the notations, we assume that the noise power $(N_0W)/M$ is normalized by \mathbf{P}_{\max} . Therefore, without loss of generality, we assume that $\mathbf{P}_{\max} = 1$. Assuming that the transmitted signals are Gaussian, the interference term seen by link $i \in \mathbb{C}_j$ will be Gaussian with power

$$I_i = \sum_{\substack{k \in \mathbb{C}_j \\ k \neq i}} \mathcal{L}_{ki} p_k. \quad (2)$$

Due to the orthogonality of the allocated subchannels, no interference is imposed from links in \mathbb{C}_k on links in \mathbb{C}_j , $k \neq j$. Under these assumptions, the achievable data rate of each link $i \in \mathbb{C}_j$ is expressed as

$$R_i(\mathbf{P}^{(j)}, \mathcal{L}_i^{(j)}) = \frac{W}{M} \log \left(1 + \frac{h_{ii} p_i}{I_i + (N_0W)/M} \right), \quad (3)$$

where $\mathcal{L}_i^{(j)} \triangleq (\mathcal{L}_{((j-1)n+1)i}, \dots, \mathcal{L}_{jn i})$. To analyze the performance of the underlying network, we use the following performance metrics

(i) *Network Average Sum-Rate:*

We define the network average sum-rate as

$$\bar{R}_{\text{ave}} \triangleq \mathbb{E} \left[\sum_{j=1}^M \sum_{l \in \mathbb{C}_j} R_l(\mathbf{P}^{(j)}, \mathcal{L}_l^{(j)}) \right], \quad (4)$$

where the expectation is computed with respect to $\mathcal{L}_l^{(j)}$. This metric is used when there is no decoding delay constraint, that is, decoding is performed over arbitrarily large number of blocks.

(ii) *Network Guaranteed Sum-Rate:*

We define the network guaranteed sum-rate as

$$\bar{R}_g \triangleq \sum_{j=1}^M \sum_{l \in \mathbb{C}_j} \mathbb{E}_{h_{ll}} [R^*(h_{ll})], \quad (5)$$

in which for all h_{ll} , $l \in \mathbb{C}_j$, we have

$$R^*(h_{ll}) \triangleq \sup R(h_{ll}), \quad (6)$$

such that

$$\mathbb{P} \{ R_l(\mathbf{P}^{(j)}, \mathcal{L}_l^{(j)}) < R(h_{ll}) \} \rightarrow 0. \quad (7)$$

This metric is useful when there exists a stringent decoding delay constraint, that is, decoding must be performed over each separate block, and a single-layer code is used. In this case, as the transmitter does not have any information about the interference term, an outage event may occur. Network guaranteed throughput is the average sum-rate of the network which is guaranteed for all channel realizations.

2.2. Objectives

Part I: Maximizing the Network Average Sum-Rate. The main objective of the first part of this paper is to maximize the network average sum-rate. This is achieved by the following.

- (i) Proposing a distributed and noniterative power allocation strategy, where each user maximizes its best estimate (based on its local information, that is, direct channel gain) of the average network sum-rate.
- (ii) Choosing the optimum value for M .

To address this problem, we first define a utility function for link $i \in \mathbb{C}_j$ ($j = 1, \dots, M$) that describes the average sum-rate of the links in cluster \mathbb{C}_j as follows:

$$u_i(p_i, h_{ii}) \triangleq \mathbb{E} \left[\sum_{l \in \mathbb{C}_j} R_l(\mathbf{P}^{(j)}, \mathcal{L}_l^{(j)}) \right], \quad (8)$$

where the expectation is computed with respect to $\{\mathcal{L}_{kl}\}_{k,l \in \mathbb{C}_j}$ excluding $k = l = i$ (namely, h_{ii}). As mentioned

earlier, h_{ii} is considered as the local (known) information for link i however, all the other gains are unknown to user i which is the reason behind statistical averaging over these parameters in (8). User i selects its power using

$$\hat{p}_i = \arg \max_{p_i \in \mathcal{P}} u_i(p_i, h_{ii}). \quad (9)$$

Given the optimum power vector $\hat{\mathbf{P}}^{(j)} = (\hat{p}_{(j-1)n+1}, \dots, \hat{p}_{jn})$ obtained from (9), the network average sum-rate is then computed as (4). Next, we choose the optimum value of M such that the network average sum-rate is maximized, that is,

$$\hat{M} = \arg \max_M \bar{R}_{\text{ave}}. \quad (10)$$

Part II: Maximizing the Network Guaranteed Sum-Rate. The main objective of the second part is finding the maximum achievable network guaranteed sum-rate in the asymptotic case of $K \rightarrow \infty$. For this purpose, a lower bound and an upper bound on the network guaranteed sum-rate are presented and shown to converge to each other as $K \rightarrow \infty$. Also, the optimum value of M is obtained.

3. Network Average Sum-Rate

In order to maximize the average sum-rate of the network, we first find the optimum power allocation policy. Using (8), we can express the utility function of link $i \in \mathbb{C}_j$, $j = 1, \dots, M$, as

$$u_i(p_i, h_{ii}) = \bar{R}_i(p_i, h_{ii}) + \sum_{\substack{l \in \mathbb{C}_j \\ l \neq i}} \bar{R}_l(p_i), \quad (11)$$

where

$$\bar{R}_i(p_i, h_{ii}) = \mathbb{E} \left[\frac{W}{M} \log \left(1 + \frac{h_{ii} p_i}{I_i + (N_0 W)/M} \right) \right] \quad (12)$$

with the expectation computed with respect to I_i defined in (2), and

$$\begin{aligned} \bar{R}_l(p_i) &= \mathbb{E} \left[R_l(\mathbf{P}^{(j)}, \mathcal{L}_l^{(j)}) \right] \\ &= \mathbb{E} \left[\frac{W}{M} \log \left(1 + \frac{h_{ll} p_l}{I_l + (N_0 W)/M} \right) \right] \end{aligned} \quad (13)$$

$$\begin{aligned} &= \mathbb{E} \left[\frac{W}{M} \log \left(1 + \frac{h_{ll} p_l}{I_l + (N_0 W)/M} \right) \right] \\ &= \mathbb{E} \left[\frac{W}{M} \log \left(1 + \frac{h_{ll} p_l}{\mathcal{L}_{il} p_i + \sum_{k \neq l, i} \mathcal{L}_{kl} p_k + (N_0 W)/M} \right) \right], \\ &\quad k, l \in \mathbb{C}_j, l \neq i, \end{aligned} \quad (14)$$

with the expectation computed with respect to $\mathbf{P}_{-i}^{(j)}$ and $\{\mathcal{L}_{kl}\}_{k, l \in \mathbb{C}_j}$ excluding $l = i$. Note that the power of the users are random variables, since they are a deterministic function of their corresponding direct channel gains, which

are random variables. It is worth mentioning that the power p_i in (15) prevents the i th user from selfishly maximizing its average rate given in (12) displaying a virtual cooperation in the network. Using the fact that all users follow the same power allocation policy, and since the channel gains \mathcal{L}_{kl} are random variables with the same distributions, $\bar{R}_l(p_i)$ becomes independent of l . Thus, by dropping the index l from $\bar{R}_l(p_i)$, the utility function of link i can be simplified as

$$u_i(p_i, h_{ii}) = \bar{R}_i(p_i, h_{ii}) + (n-1)\bar{R}(p_i). \quad (16)$$

Noting that p_i depends only on the channel gain h_{ii} , in the sequel we use $p_i = g(h_{ii})$.

Lemma 3.1. *Let assume $0 < \alpha \leq 1$ is fixed and $\mathbb{E}[p_k] \triangleq q_n$. Then with probability one (w. p. 1), we have*

$$I_i \sim (n-1)\hat{\alpha}q_n, \quad (17)$$

as $K \rightarrow \infty$ (or equivalently, $n \rightarrow \infty$), where $\hat{\alpha} \triangleq \alpha\omega$. More precisely, substituting I_i by $(n-1)\hat{\alpha}q_n$ does not change the asymptotic average sum-rate of the network.

Proof. See Appendix A. \square

Lemma 3.2. *For large values of n , the links with a direct channel gain above $h_{\text{Th}} = c \log n$, where $c > 1$ is a constant, have negligible contribution in the network average sum-rate.*

Proof. See Appendix B. \square

From Lemma 3.2 and for large values of n , we can limit our attention to a subset of links for which the direct channel gain h_{ii} is less than $c \log n$, $c > 1$.

Theorem 3.3. *Assuming K is large, the optimum power allocation policy for (9) is $\hat{p}_i = g(h_{ii}) = U(h_{ii} - \tau_n)$, where $\tau_n > 0$ is a threshold level which is a function of n and $U(\cdot)$ is the unit step function. Also, the maximum network average sum-rate in (4) is achieved at $M = 1$ and is given by*

$$\bar{R}_{\text{ave}} \sim \frac{W}{\hat{\alpha}} \log K. \quad (18)$$

Proof. The steps of the proof are as follows: First, we derive an upper bound on the utility function given in (16). Then, we prove that the optimum power allocation strategy that maximizes this upper bound is $\hat{p}_i = g(h_{ii}) = U(h_{ii} - \tau_n)$. Based on this power allocation policy, in Lemma 3.5, we derive the optimum threshold level τ_n . We then show that, using this optimum threshold value, the maximum value of the utility function in (16) becomes asymptotically the same as the maximum value of the upper bound obtained in the first step. Finally, the proof of the theorem is completed by showing that the maximum network average sum-rate is achieved at $M = 1$.

Step 1 (Upper Bound on the Utility Function). Let us assume that $\mathbb{E}[p_k] = q_n$. Using the results of Lemma 3.1, $\bar{R}_i(p_i, h_{ii})$ in (16) can be expressed as

$$\bar{R}_i(p_i, h_{ii}) \approx \frac{W}{M} \mathbb{E} \left[\log \left(1 + \frac{h_{ii} p_i}{(n-1)\hat{\alpha}q_n + (N_0 W)/M} \right) \right] \quad (19)$$

$$\stackrel{(a)}{=} \frac{W}{M} \log \left(1 + \frac{h_{ii} p_i}{\lambda} \right), \quad (20)$$

as $K \rightarrow \infty$, where

$$\lambda \triangleq (n-1)\hat{\alpha}q_n + \frac{N_0 W}{M}. \quad (21)$$

In the above equations, (a) follows from the fact that h_{ii} is a known parameter for user i and $p_i = g(h_{ii})$ is the optimization parameter. With a similar argument, (15) can be simplified as

$$\bar{R}(p_i) \approx \frac{W}{M} \mathbb{E} \left[\log \left(1 + \frac{h_{ll} p_l}{\mathcal{L}_{il} p_i + (n-2)\hat{\alpha}q_n + (N_0 W)/M} \right) \right], \quad i \neq l, \quad (22)$$

$$\stackrel{(a)}{=} \alpha \frac{W}{M} \times \mathbb{E} \left[\log \left(1 + \frac{h_{ll} p_l}{\beta_{il} h_{il} p_i + (n-2)\hat{\alpha}q_n + (N_0 W)/M} \right) \right] + (1-\alpha) \frac{W}{M} \times \mathbb{E} \left[\log \left(1 + \frac{h_{ll} p_l}{(n-2)\hat{\alpha}q_n + (N_0 W)/M} \right) \right] \quad (23)$$

$$= \frac{\alpha W}{M} \mathbb{E} \left[\log \left(1 + \frac{h_{ll} p_l}{\beta_{il} h_{il} p_i + \lambda'} \right) \right] + (1-\alpha) \frac{W}{M} \mathbb{E} \left[\log \left(1 + \frac{h_{ll} p_l}{\lambda'} \right) \right], \quad (24)$$

as $K \rightarrow \infty$, where the expectation is computed with respect to h_{ll} , h_{il} , p_l and β_{il} , and $\lambda' \triangleq (n-2)\hat{\alpha}q_n + (N_0 W)/M$. Also, (a) comes from the shadowing model described in (1). Using (20), (24), and the inequality $\log(1+x) \leq x$, $\forall x \geq 0$, the utility function in (16) is upper bounded as

$$u_i(p_i, h_{ii}) \leq \frac{W}{M} \frac{h_{ii}}{\lambda} p_i + n \frac{\alpha W}{M} \mathbb{E} \left[\frac{h_{ll} p_l}{\beta_{il} h_{il} p_i + \lambda'} \right] + n(1-\alpha) \frac{W}{M \lambda'} \mathbb{E}[h_{ll} p_l]. \quad (25)$$

Note that the factor $(n-1)$ in (16) is replaced by n in (25), which does not affect the validity of the equation. Noting that h_{ll} is independent of h_{il} , $i \neq l$, we have

$$\mathbb{E} \left[\frac{h_{ll} p_l}{\beta_{il} h_{il} p_i + \lambda'} \mid \beta_{il} \right] = \mu \int_0^\infty \frac{e^{-y}}{y \beta_{il} p_i + \lambda'} dy = -\frac{\mu}{\beta_{il} p_i} e^{\lambda'/(\beta_{il} p_i)} \text{Ei} \left(-\frac{\lambda'}{\beta_{il} p_i} \right), \quad (26)$$

where

$$\mu \triangleq \mathbb{E}[h_{ll} p_l], \quad (27)$$

and $\text{Ei}(x) \triangleq -\int_{-x}^\infty e^{-t}/dt$, $x < 0$ is the *exponential-integral function* [27]. Thus, the right hand side of (25) is simplified as

$$u_i(p_i, h_{ii}) \leq \frac{W}{M} \frac{h_{ii}}{\lambda} p_i - n \frac{\alpha \mu W}{M} \mathbb{E} \left[\frac{1}{\beta_{il} p_i} e^{\lambda'/(\beta_{il} p_i)} \text{Ei} \left(-\frac{\lambda'}{\beta_{il} p_i} \right) \right] + n(1-\alpha) \frac{W}{M} \frac{\mu}{\lambda'}, \quad (28)$$

where the expectation is computed with respect to β_{il} . An asymptotic expansion of $\text{Ei}(x)$ can be obtained as [27, page 951]

$$\text{Ei}(x) = \frac{e^x}{x} \left[\sum_{k=0}^{L-1} \frac{k!}{x^k} + O(|x|^{-L}) \right]; \quad L = 1, 2, \dots, \quad (29)$$

as $x \rightarrow -\infty$. Setting $L = 4$, we can rewrite (28) as

$$u_i(p_i, h_{ii}) \leq \frac{W}{M} \frac{h_{ii}}{\lambda} p_i + n \frac{\alpha W \mu}{M \lambda'} \times \mathbb{E} \left[\left(1 - \frac{\beta_{il} p_i}{\lambda'} + 2 \left(\frac{\beta_{il} p_i}{\lambda'} \right)^2 - 6 \left(\frac{\beta_{il} p_i}{\lambda'} \right)^3 \right) \right] + n \frac{\alpha W \mu}{M \lambda'} \mathbb{E} \left[O \left(\left| \frac{\beta_{il} p_i}{\lambda'} \right|^4 \right) \right] + n(1-\alpha) \frac{W \mu}{M \lambda'} \quad (30)$$

$$\stackrel{(a)}{\approx} \frac{W}{M} \frac{h_{ii}}{\lambda} p_i + n \frac{\alpha W \mu}{M \lambda'} \left(1 - \frac{\omega p_i}{\lambda'} + 2\kappa \left(\frac{p_i}{\lambda'} \right)^2 - 6\eta \left(\frac{p_i}{\lambda'} \right)^3 \right) + n(1-\alpha) \frac{W \mu}{M \lambda'}, \quad (31)$$

$$\triangleq \Xi_i(p_i, h_{ii}) \quad (32)$$

as $\lambda' \rightarrow \infty$, where $\kappa \triangleq \mathbb{E}[\beta_{il}^2]$ and $\eta \triangleq \mathbb{E}[\beta_{il}^3]$, and (a) follows from the fact that, for large values of λ' , the term $\mathbb{E}[O(|\beta_{il} p_i|/\lambda'^4)]$ can be ignored.

Step 2 (Optimum Power Allocation Policy for $\Xi_i(p_i, h_{ii})$). Using the fact that $p_i \in [0, 1]$, the second-order derivative of (31) in terms of p_i , $\partial^2 \Xi_i(p_i, h_{ii}) / \partial p_i^2 = n(\alpha W \mu / M \lambda') (4\kappa / \lambda'^2 - (36\eta / \lambda'^3) p_i)$, is positive as $\lambda' \rightarrow \infty$. It is observed from (29) and (31) that for any value of $L > 4$, the second-order derivative of (31) in terms of p_i is positive too. Thus, (31) is a *convex* function of p_i . It is known that a convex function attains its maximum at one of its extreme points of its domain [28]. In other words, the optimum power that maximizes (31) is $\hat{p}_i \in \{0, 1\}$. To show that this optimum power is in the form of a unit step function, it is sufficient to prove that $p_i = g(h_{ii})$ is a monotonically increasing function of h_{ii} .

Suppose that the optimum power that maximizes $\Xi_i(p_i, h_{ii})$ is $p_i = 1$. Also, let us define $h'_{ii} \triangleq h_{ii} + \delta$, where $\delta > 0$. From (31), it is clear that $\Xi_i(p_i, h_{ii})$ is a monotonically increasing function of h_{ii} , that is,

$$\Xi_i(p_i = 1, h'_{ii}) > \Xi_i(p_i = 1, h_{ii}). \quad (33)$$

On the other hand, since the optimum power is $p_i = 1$, we conclude that

$$\Xi_i(p_i = 1, h_{ii}) > \Xi_i(p_i = 0, h_{ii}). \quad (34)$$

Using the fact that $\Xi_i(p_i = 0, h_{ii}) = \Xi_i(p_i = 0, h'_{ii})$, we arrive at the following inequality

$$\Xi_i(p_i = 1, h'_{ii}) > \Xi_i(p_i = 0, h'_{ii}). \quad (35)$$

From (33)–(35), it is concluded that $g(h_{ii})$ is a monotonically increasing function of h_{ii} . Consequently, the optimum power allocation strategy that maximizes $\Xi_i(p_i, h_{ii})$ is a unit step function, that is,

$$\hat{p}_i = \begin{cases} 1 & \text{if } h_{ii} > \tau_n, \\ 0 & \text{otherwise,} \end{cases} \quad (36)$$

where τ_n is a threshold level to be determined. We call this the *threshold-based on-off power allocation strategy*. It is observed that the optimum power \hat{p}_i is a Bernoulli random variable with parameter q_n , that is,

$$f(\hat{p}_i) = \begin{cases} q_n, & \hat{p}_i = 1, \\ 1 - q_n, & \hat{p}_i = 0, \end{cases} \quad (37)$$

where $f(\cdot)$ is the probability mass function (pmf) of \hat{p}_i . We conclude from (36) and (37) that the probability of link activation in each cluster is $q_n \triangleq \mathbb{P}\{h_{ii} > \tau_n\} = e^{-\tau_n}$ which is a function of n .

Step 3 (Optimum Threshold Level τ_n). From Step 1, it is observed that for every value of p_i we have

$$u_i(p_i, h_{ii}) \leq \Xi_i(p_i, h_{ii}). \quad (38)$$

The above inequality is also valid for the optimum power \hat{p}_i obtained in Step 2. Thus, using the fact that for $X \leq Y$, $\mathbb{E}[X] \leq \mathbb{E}[Y]$, we conclude

$$\mathbb{E}[u_i(\hat{p}_i, h_{ii})] \leq \mathbb{E}[\Xi_i(\hat{p}_i, h_{ii})], \quad (39)$$

where the expectations are computed with respect to h_{ii} . In the following lemmas, we first derive the optimum threshold level τ_n that maximizes $\mathbb{E}[\Xi_i(\hat{p}_i, h_{ii})]$, and then prove that this quantity is asymptotically the same as the optimum threshold level maximizing $\mathbb{E}[u_i(\hat{p}_i, h_{ii})]$, assuming an on-off power scheme. In fact, since the threshold τ_n is fixed and does not depend on a specific realization of h_{ii} , finding the optimum value of τ_n requires averaging the utility function over all realizations of h_{ii} . We also show that the maximum value of $\mathbb{E}[u_i(\hat{p}_i, h_{ii})]$ (assuming an on-off power scheme) is the same as the optimum value of $\mathbb{E}[\Xi_i(\hat{p}_i, h_{ii})]$, proving the desired result.

Lemma 3.4. *For large values of n and given $0 < \alpha \leq 1$, the optimum threshold level that maximizes $\mathbb{E}[\Xi_i(\hat{p}_i, h_{ii})]$ is computed as*

$$\hat{\tau}_n \sim \log n. \quad (40)$$

Also, the maximum value of $\mathbb{E}[\Xi_i(\hat{p}_i, h_{ii})]$ scales as $(W/M\hat{\alpha}) \log n$.

Proof. See Appendix C. \square

Lemma 3.5. *For large values of n and given $0 < \alpha \leq 1$,*

- (i) *the optimum threshold level that maximizes $\mathbb{E}[u_i(\hat{p}_i, h_{ii})]$ is computed as*

$$\hat{\tau}_n = \log n - 2 \log \log n + O(1), \quad (41)$$

- (ii) *the probability of link activation in each cluster is given by*

$$q_n = \delta \frac{\log^2 n}{n}, \quad (42)$$

where $\delta > 0$ is a constant,

- (iii) *the maximum value of $\mathbb{E}[u_i(\hat{p}_i, h_{ii})]$ scales as $(W/M\hat{\alpha}) \log n$.*

Proof. See Appendix D. \square

Step 4 (Optimum Power Allocation Strategy that Maximizes $u_i(p_i, h_{ii})$). In order to prove that the utility function in (16) is asymptotically the same as the upper bound $\Xi_i(p_i, h_{ii})$ obtained in (31), it is sufficient to show that the low SINR conditions in (20) and (24) are satisfied. Using (20), (21), and (42), the SINR is equal to $h_{ii} p_i / \lambda$, where

$$\lambda \approx \hat{\alpha} \delta \log^2 n + \frac{N_0 W}{M}. \quad (43)$$

It is observed that λ goes to infinity as $n \rightarrow \infty$. On the other hand, since we are limiting our attention to links with $h_{ii} < h_{Th} = c \log n$, we have

$$\frac{h_{ii} p_i}{\lambda} = O\left(\frac{1}{\log n}\right), \quad (44)$$

when $n \rightarrow \infty$. Thus, for large values of n , the low SINR condition, $h_{ii}p_i/\lambda \ll 1$, is satisfied. With a similar argument, the low SINR condition for (24) is satisfied. Hence, we can use the approximation $\log(1+x) \approx x$, for $x \ll 1$, to simplify (20) and (24) as follows:

$$\bar{R}_i(p_i, h_{ii}) \approx \frac{W}{M} \frac{h_{ii}}{\lambda} p_i, \quad (45)$$

$$\bar{R}(p_i) \approx \frac{\alpha W}{M} \mathbb{E} \left[\frac{h_{ii} p_i}{\beta_{ii} h_{ii} p_i + \lambda'} \right] + (1 - \alpha) \frac{W}{M \lambda'} \mathbb{E}[h_{ii} p_i]. \quad (46)$$

Consequently, the utility function $u_i(p_i, h_{ii})$ is the same as the upper bound $\Xi_i(p_i, h_{ii})$ obtained in (31), when $n \rightarrow \infty$. Thus, the optimum power allocation strategy for (9) is the same as the optimum power allocation policy that maximizes $\Xi_i(p_i, h_{ii})$.

Step 5 (Maximum Average Network Sum-rate). Using (8), the average utility function of each user i , $\mathbb{E}[u_i(\hat{p}_i, h_{ii})]$, $i \in \mathbb{C}_j$, is the same as the average sum-rate of the links in cluster \mathbb{C}_j represented by

$$\bar{R}_{\text{ave}}^{(j)} \triangleq \sum_{i \in \mathbb{C}_j} \mathbb{E} \left[R_i(\hat{\mathbf{P}}^{(j)}, \mathcal{L}_i^{(j)}) \right], \quad j = 1, \dots, M. \quad (47)$$

where $\hat{\mathbf{P}}^{(j)}$ is the on-off powers vector of the links in cluster \mathbb{C}_j . In this case, the network average sum-rate defined in (4) can be written as

$$\bar{R}_{\text{ave}} = \sum_{j=1}^M \bar{R}_{\text{ave}}^{(j)}, \quad (48)$$

$$\stackrel{(a)}{\approx} \frac{W \hat{\tau}_n}{\hat{\alpha}}, \quad (49)$$

where (a) follows from (D.14) of Appendix D. Using (41), and noting that $n = K/M$, we have

$$\bar{R}_{\text{ave}} \sim \frac{W}{\hat{\alpha}} \log \frac{K}{M}. \quad (50)$$

Step 6 (Optimum Spectrum Allocation). According to (49), the network average sum-rate is a monotonically increasing function of $\hat{\tau}_n$. Rewriting (D.10) of Appendix D, which gives the optimum threshold value for the on-off scheme,

$$-e^{-\hat{\tau}_n} \log \left(1 + \frac{\hat{\tau}_n e^{\hat{\tau}_n}}{n \hat{\alpha}} \right) + \frac{1 + \hat{\tau}_n}{n \hat{\alpha} + \hat{\tau}_n e^{\hat{\tau}_n}} = 0, \quad (51)$$

it can be shown that

$$\hat{\tau}_n^2 e^{\hat{\tau}_n} \approx n \hat{\alpha}, \quad (52)$$

which implies that $\hat{\tau}_n$ is an increasing function of n . In deriving (52), we have used the fact that $\hat{\tau}_n e^{\hat{\tau}_n}/n \hat{\alpha} \ll 1$, which is feasible based on the solution given in (41). Therefore, the average sum-rate of the network is an increasing function of n and consequently, noting that $n = K/M$, is a decreasing function of M . Hence, the maximum average sum-rate of the network for large K and $0 < \alpha < 1$ is obtained at $M = 1$ and this completes the proof of the theorem. \square

Motivated by Theorem 3.3, we describe the proposed threshold-based on-off power allocation strategy for single-hop wireless networks. Based on this scheme, all users perform the following steps during each block.

(i) Based on the direct channel gain, the transmission policy is

$$\hat{p}_i = \begin{cases} 1 & \text{if } h_{ii} > \tau_n \\ 0 & \text{Otherwise.} \end{cases} \quad (53)$$

(ii) Knowing its corresponding direct channel gain, each active user i transmits with full power and rate

$$R_i = \log \left(1 + \frac{h_{ii}}{(n-1)\hat{\alpha}e^{-\tau_n} + (N_0 W)/M} \right). \quad (54)$$

(iii) Decoding is performed over sufficiently large number of blocks, yielding the average rate of $(W/\hat{\alpha}K) \log K$ for each user, and the average sum-rate of $W/\hat{\alpha} \log K$ in the network.

Remark 1. Theorem 3.3 states that the average sum-rate of the network for fixed M depends on the value of $\hat{\alpha} = \alpha \hat{\omega}$ and scales as $(W/\hat{\alpha}) \log(K/M)$. Also, for values of M such that $\log M = o(\log K)$, the network average sum-rate scales as $(W/\hat{\alpha}) \log K$.

Remark 2. Let m_j denote the number of active links in \mathbb{C}_j . Lemma 3.5 states that the optimum selection of the threshold value yields $\mathbb{E}[m_j] = n q_n = \Theta(\log^2 n)$. More precisely, it can be shown that the optimum number of active users scales as $\Theta(\log^2 n)$, with probability one.

Theorem 3.6. *Let us assume that K is large and M is fixed. Then,*

- (i) *for the moderate interference, that is, $\mathbb{E}[I_i] = \Theta(1)$, the network average sum-rate is bounded by $\bar{R}_{\text{ave}} \leq \Theta(\log n)$;*
- (ii) *for the weak interference, that is, $\mathbb{E}[I_i] = o(1)$, the network average sum-rate is bounded by $\bar{R}_{\text{ave}} \leq o(\log n)$.*

Proof. (i) From (4), we have

$$\bar{R}_{\text{ave}} = \sum_{j=1}^M \sum_{l \in \mathbb{C}_j} \mathbb{E} \left[\frac{W}{M} \log \left(1 + \frac{h_{ll} \hat{p}_l}{I_l + \frac{N_0 W}{M}} \right) \right] \quad (55)$$

$$\stackrel{(a)}{\leq} \sum_{j=1}^M \sum_{l \in \mathbb{C}_j} \frac{W}{M} \mathbb{E} \left[\log \left(1 + \frac{\hat{p}_l c \log n}{I_l + (N_0 W)/M} \right) \right] \quad (56)$$

$$\leq \sum_{j=1}^M \sum_{l \in \mathbb{C}_j} \frac{W}{M} \mathbb{E} \left[\log \left(1 + \frac{\hat{p}_l c \log n}{(N_0 W)/M} \right) \right] \quad (57)$$

$$\stackrel{(b)}{\leq} \sum_{j=1}^M \sum_{l \in \mathcal{C}_j} \frac{W}{M} \log \left(1 + \frac{cq_n \log n}{(N_0 W)/M} \right) \quad (58)$$

$$\stackrel{(c)}{\leq} \frac{cM}{N_0} nq_n \log n \quad (59)$$

where (a) follows from Lemma 3.2, which implies that the realizations in which $h_{ll} > c \log n$ for some $c > 1$ have negligible contribution in the network average sum-rate, (b) results from the *Jensen's inequality*, $\mathbb{E}[\log x] \leq \log(\mathbb{E}[x])$, $x > 0$. Also, (c) follows from the fact that $\log(1+x) \leq x$, $x \geq 0$. Since for the moderate interference, $\mathbb{E}[I_i] = \hat{\alpha}nq_n = \Theta(1)$, and using the fact that M is fixed, we come up with the following inequality:

$$\begin{aligned} \bar{R}_{\text{ave}} &\leq \frac{cM}{\hat{\alpha}N_0} \Theta(1) \log n \\ &= \Theta(\log n). \end{aligned} \quad (60)$$

(ii) For the weak interference scenario, where $\mathbb{E}[I_i] = \hat{\alpha}nq_n = o(1)$, and similar to the part (i), it is concluded from (59) that

$$\begin{aligned} \bar{R}_{\text{ave}} &\leq \frac{cM}{\hat{\alpha}N_0} o(1) \log n \\ &= o(\log n). \end{aligned} \quad (61)$$

□

Remark 3. It is concluded from Theorems 3.3 and 3.6 that the maximum average sum-rate of the proposed network is scaled as $\Theta(\log K)$.

So far, we have assumed that M is fixed, that is, it does not scale with K . In the following, we present some results for the case that M scales with K . Obviously, we consider the values of M which are in the interval $[1, K]$. It should be noted that the results for $M = o(K)$ are the same as the results in Theorem 3.3.

Theorem 3.7. *In the network with the on-off power allocation strategy, if $M = \Theta(K)$ and $0 < \alpha < 1$, then the maximum network average sum-rate in (4) is less than that of $M = 1$. Consequently, the maximum average sum-rate of the network for every value of $1 \leq M \leq K$ is achieved at $M = 1$.*

Proof. See Appendix E. □

Remark 4. According to the shadow-fading model proposed in (1), it is seen that for $\alpha = 0$, with probability one, $\mathcal{L}_{ki} = 0, k \neq i$. This implies that no interference exists in each cluster. In this case, the maximum average sum-rate of the network is clearly achieved by all users in the network transmitting at full power. It can be shown that for every value of $1 \leq M \leq K$, the maximum network average sum-rate for $\alpha = 0$ is achieved at $M = 1$ (See Appendix F for the proof).

Remark 5. Noting that for $M = K$ only one user exists in each cluster, all the users can communicate using an interference free channel. It can be shown that for $M = K$ and every value

of $0 \leq \alpha \leq 1$, the network average sum-rate is asymptotically obtained as

$$\bar{R}_{\text{ave}} \approx W(\log K - \log N_0 W - \gamma), \quad (62)$$

where γ is Euler's constant (See Appendix G for the proof). Therefore, for every value of $0 < \alpha < 1$, it is observed that the average sum-rate of the network in (62) is less than that of $M = 1$ obtained in (18).

Remark 6. Note that for $M = 1$, in which the average number of active links scales as $\Theta(\log^2 K)$ (in the optimum on-off scheme), we have significant energy saving in the network as compared to the case of $M = K$, in which all the users transmit with full power.

3.1. Numerical Results. So far, we have analyzed the average sum-rate of the network in terms of M and $\hat{\alpha}$, in the asymptotic case of $K \rightarrow \infty$. For finite number of users, we have evaluated the network average sum-rate versus the number of clusters (M) through simulation. For this case, we assume that all the users in the network follow the threshold-based on-off power allocation policy, using the optimum threshold value. In addition, the shadowing effect is assumed to be *lognormal* distributed with mean $\bar{\omega} \leq 1$ and variance 1. Figure 1 shows the average sum-rate of the network versus M for $K = 20$ and $K = 40$ and different values of α and $\bar{\omega}$. It is observed from this figure that the average sum-rate of the network is a monotonically decreasing function of M for every value of $(\alpha, \bar{\omega})$, which implies that the maximum value of \bar{R}_{ave} is achieved at $M = 1$. This result confirms our claim in Theorem 3.7.

Based on the above arguments, we have plotted the average sum-rate of the network versus K for $M = 1$ and different values of $(\alpha, \bar{\omega})$. It is observed from Figure 2 that the network average sum-rate depends strongly on the values of $(\alpha, \bar{\omega})$. In addition, we can see that the average sum-rate of the network increases logarithmically in terms of n .

In addition, Figure 3 illustrates the average sum-rate of the network with the optimized on-off power allocation strategy compared to the centralized power allocation algorithm and the case that all the links transmit with full power. In the centralized scheme, it is assumed that the central node knows all the network information. For each channel realization and through exhaustive search, the central node selects the optimum powers for all the links such that the maximum average sum-rate is achieved. It is seen that the performance of the proposed on-off power allocation strategy is better than that of the full power scheme. Also, the highest average sum-rate is achieved by the centralized scheme. However in the network with a large number of links, deploying centralized power allocation schemes becomes computationally intractable, while in the on-off power scheme, the average sum-rate is achieved without coordination among the links.

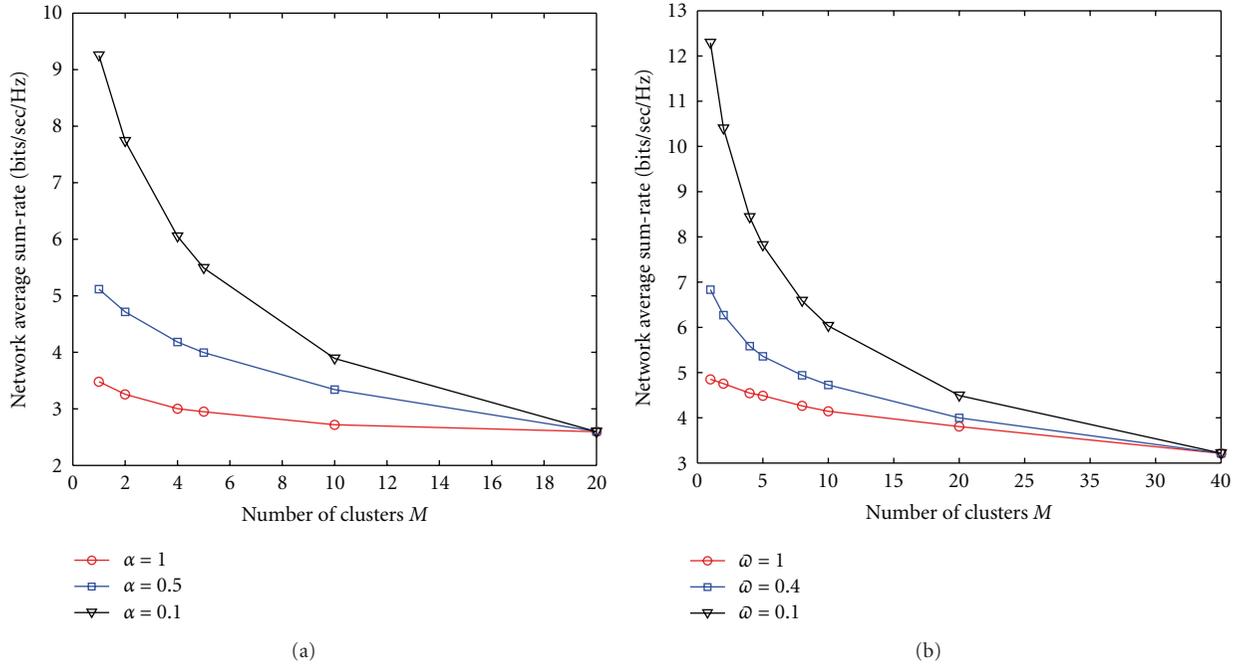


FIGURE 1: Network average sum-rate versus M for (a) $K = 20$, $\alpha = 1, 0.5, 0.1$, and shadowing model with $\omega = 0.5$ and variance 1 and for (b) $K = 40$, $\alpha = 0.5$, and shadowing model with $\omega = 1, 0.4, 0.1$ and variance 1.

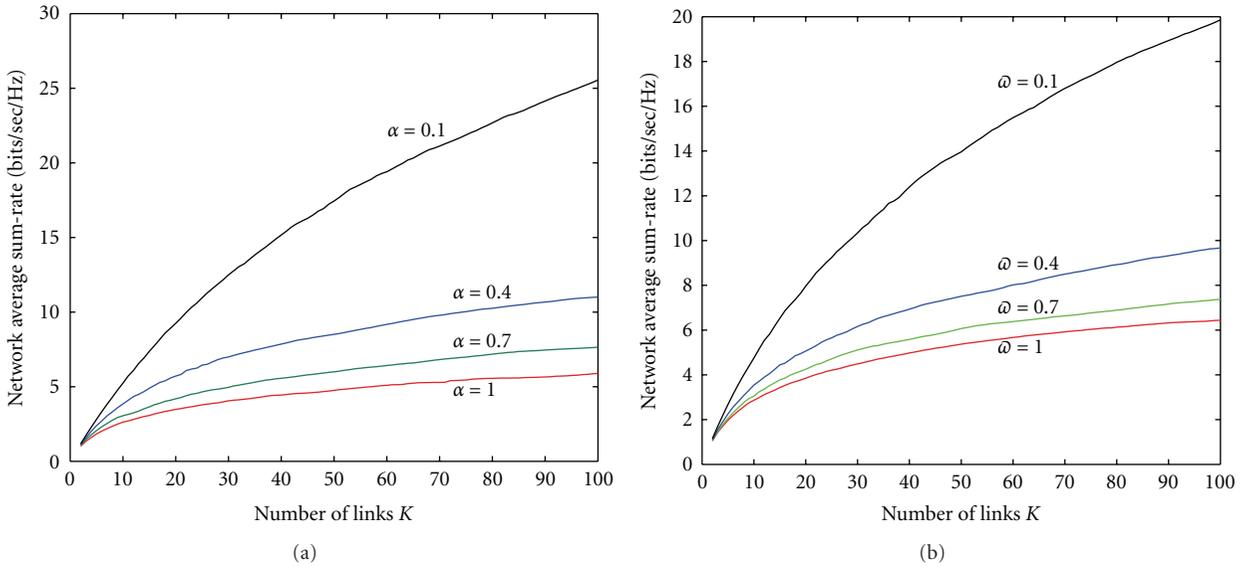


FIGURE 2: Network average sum-rate versus K for $M = 1$, (a) shadowing model with $\omega = 0.5$ and variance 1 and $\alpha = 1, 0.7, 0.4, 0.1$, and b) shadowing model with $\omega = 1, 0.7, 0.4, 0.1$, variance 1, and $\alpha = 0.5$.

4. Network Guaranteed Sum-Rate

Recalling the definition of the network guaranteed sum-rate in (5), in this section we aim to find the maximum achievable guaranteed sum-rate of the network, as well as the optimum power allocation scheme and the optimum value of M .

Theorem 4.1. *The guaranteed sum-rate of the underlying network in the asymptotic case of $K \rightarrow \infty$ is obtained by*

$$\bar{R}_g \sim \frac{W}{\alpha} \log K, \quad (63)$$

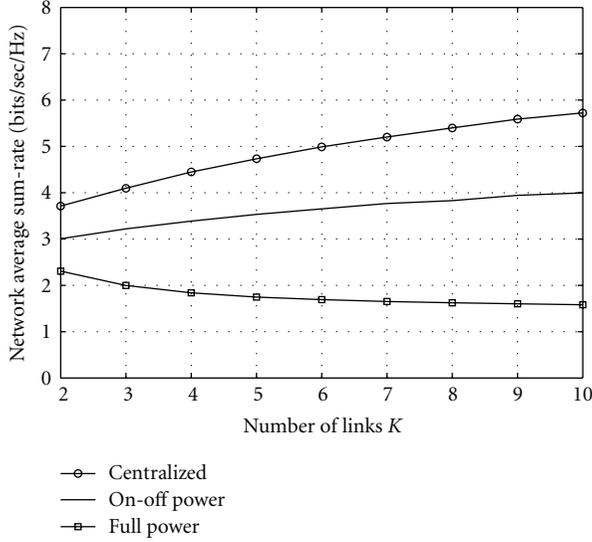


FIGURE 3: Average sum-rate of the network versus the number of links K for different power allocation schemes.

which is achievable by the decentralized on-off power allocation scheme.

Proof. In order to compute the guaranteed rate for link $l \in \mathbb{C}_j$, we first define the corresponding outage event as follows:

$$\begin{aligned} \mathcal{O}_l^{(j)} &\equiv \{R_l(\mathbf{p}^{(j)}, \mathcal{L}_l^{(j)}) < R(h_{ll})\} \\ &\equiv \left\{ \log \left(1 + \frac{p_l h_{ll}}{I_l + (N_0 W)/M} \right) < R(h_{ll}) \right\}. \end{aligned} \quad (64)$$

In the following, we give an upper bound and a lower-bound for \bar{R}_g and show that these bounds converge to each other as $K \rightarrow \infty$ (or equivalently, $n \rightarrow \infty$).

Upper Bound. An upper bound on the guaranteed sum-rate can be given by lower-bounding the outage probability as follows:

$$\mathbb{P}\{\mathcal{O}_l^{(j)}\} \geq \mathbb{P}\left\{ \frac{p_l h_{ll}}{I_l + N_0 W/M} < R(h_{ll}) \right\} \quad (65)$$

$$= \mathbb{P}\left\{ p_l h_{ll} - \frac{N_0 W}{M} R(h_{ll}) < I_l R(h_{ll}) \right\}, \quad (66)$$

in which we have used the fact that $\log(1+x) \leq x$. Denoting $\nu = h_{ll}$, we can write

$$\mathbb{P}\{\mathcal{O}_l^{(j)}\} \stackrel{(a)}{\geq} \mathbb{P}\left\{ e^{-I_l \xi(\nu) R(\nu)} \leq e^{\xi(\nu) ((N_0 W/M) R(\nu) - p_l \nu)} \right\} \quad (67)$$

$$\stackrel{(b)}{\geq} 1 - e^{-\xi(\nu) ((N_0 W/M) R(\nu) - p_l \nu)} \mathbb{E}\left[e^{-I_l \xi(\nu) R(\nu)} \right], \quad (68)$$

for some positive $\xi(\nu)$. In the above equation, (a) results from (66), noting that $\xi(\nu) > 0$, and (b) follows from Markov's inequality [29, page 77], and the expectation is taken with respect to I_l . The above equation implies that finding an upper bound for $\mathbb{E}[e^{-I_l \xi(\nu) R(\nu)}]$ is sufficient for the lower-bounding the outage probability. For this purpose, using (2), we can write

$$\mathbb{E}\left[e^{-I_l \xi(\nu) R(\nu)} \right] = \mathbb{E}\left[e^{-\xi(\nu) R(\nu) \sum_{k \in \mathbb{C}_j, k \neq l} \mathcal{L}_{kl} p_k} \right], \quad (69)$$

$$\stackrel{(a)}{=} \prod_{\substack{k \in \mathbb{C}_j \\ k \neq l}} \mathbb{E}\left[e^{-\xi(\nu) R(\nu) \mathcal{L}_{kl} p_k} \right], \quad (70)$$

$$\stackrel{(b)}{=} \prod_{\substack{k \in \mathbb{C}_j \\ k \neq l}} \mathbb{E}\left[e^{-\xi(\nu) R(\nu) u_{kl} \beta_{kl} h_{kl} p_k} \right], \quad (71)$$

$$\stackrel{(c)}{=} \left(\mathbb{E}\left[e^{-\xi(\nu) R(\nu) u_{kl} \beta_{kl} h_{kl} p_k} \right] \right)^{n-1}, \quad k \neq l. \quad (72)$$

In the above equation, (a) follows from the fact that $\{\mathcal{L}_{kl}\}_{k \in \mathbb{C}_j}$ with $k \neq l$, and $\{p_k\}_{k \in \mathbb{C}_j}$ are mutually independent random variables, (b) results from writing \mathcal{L}_{kl} as $u_{kl} \beta_{kl} h_{kl}$ (from (1)), in which u_{kl} is an indicator variable which takes zero when $\mathcal{L}_{kl} = 0$ and one, otherwise. (c) follows from the symmetry which incurs that all the terms $\mathbb{E}[e^{-\xi(\nu) R(\nu) u_{kl} \beta_{kl} h_{kl} p_k}]$, $k \in \mathbb{C}_j$, are equal. Noting that u_{kl} , β_{kl} , h_{kl} , and p_k are independent of each other, we have

$$\mathbb{E}\left[e^{-\xi(\nu) R(\nu) u_{kl} \beta_{kl} h_{kl} p_k} \right] = \mathbb{E}_{\beta_{kl}} \left[\mathbb{E}_{h_{kl}} \left[\mathbb{E}_{u_{kl}} \left[\mathbb{E}_{p_k} \left[e^{-\xi(\nu) R(\nu) u_{kl} \beta_{kl} h_{kl} p_k} \right] \right] \right] \right], \quad (73)$$

$$\stackrel{(a)}{\leq} \mathbb{E}_{\beta_{kl}} \left[\mathbb{E}_{h_{kl}} \left[\mathbb{E}_{u_{kl}} \left[(1 - q_n) + q_n e^{-\xi(\nu) R(\nu) u_{kl} \beta_{kl} h_{kl}} \right] \right] \right], \quad (74)$$

$$\stackrel{(b)}{=} \mathbb{E}_{\beta_{kl}} \left[\mathbb{E}_{h_{kl}} \left[(1 - q_n) + q_n (1 - \alpha + \alpha e^{-\xi(\nu) R(\nu) \beta_{kl} h_{kl}}) \right] \right], \quad (75)$$

$$\stackrel{(c)}{=} \mathbb{E}_{\beta_{kl}} \left[1 - \alpha q_n + \frac{\alpha q_n}{1 + \beta_{kl} \xi(\nu) R(\nu)} \right] \quad (76)$$

$$= \mathbb{E}_{\beta_{kl}} \left[1 - \frac{\alpha q_n \beta_{kl} \xi(\nu) R(\nu)}{1 + \beta_{kl} \xi(\nu) R(\nu)} \right] \quad (77)$$

$$\stackrel{(d)}{\leq} 1 - \frac{\alpha q_n \omega \xi(\nu) R(\nu)}{1 + \beta_{\max} \xi(\nu) R(\nu)}, \quad (78)$$

$$\stackrel{(e)}{\leq} e^{-\frac{\hat{\alpha} q_n \xi(\nu) R(\nu)}{1 + \beta_{\max} \xi(\nu) R(\nu)}}. \quad (79)$$

In the above equation, (a) follows from the fact that $e^{-\theta x} \leq (1-x) + xe^{-\theta}$, $\forall \theta \geq 0$, and $0 \leq x \leq 1$, noting that $\mathbb{E}[p_k] = q_n$. (b) results from the definition of u_{kl} , which is an indicator variable taking zero with probability $1 - \alpha$ and one, with probability α . (c) follows from the fact that h_{kl} is exponentially distributed, we have $\mathbb{E}_{h_{kl}}[e^{-\xi(\nu) R(\nu) \beta_{kl} h_{kl}}] = 1/(1 + \beta_{kl} \xi(\nu) R(\nu))$. (d) results from the facts that $\beta_{kl} \leq \beta_{\max}$ and $\mathbb{E}[\beta_{kl}] = \omega$. Finally, (e) follows from the fact that $1 - x \leq e^{-x}$, $\forall x$, and noting that $\alpha \omega = \hat{\alpha}$. Combining (72) and (79) and substituting into (68) yields

$$\begin{aligned} \mathbb{P}\{\mathcal{O}_l^{(j)}\} &\geq 1 - e^{-\xi(\nu)((N_0 W/M)R(\nu) - p_l \nu)} e^{-((n-1)\hat{\alpha}q_n \xi(\nu)R(\nu))/(1+\beta_{\max} \xi(\nu)R(\nu))} \\ &= 1 - e^{-\xi(\nu)R(\nu)((n-1)\hat{\alpha}q_n)/(1+\beta_{\max} \xi(\nu)R(\nu)) + (N_0 W/M)(1-t(\nu)/R(\nu))}, \end{aligned} \quad (80)$$

where $t(\nu) \triangleq (p_l \nu)/(((n-1)\hat{\alpha}q_n)/(1+\beta_{\max} \xi(\nu)R(\nu)) + (N_0 W/M))$.

Consider the cases of $\mathbb{E}\{I_l\} = \omega(1)$ (strong interference) or $\mathbb{E}\{I_l\} = \Theta(1)$ (moderate interference). Let us

define $\gamma \triangleq \min(1, (M(n-1)q_n \hat{\alpha})/N_0 W)$. Setting $\xi(\nu) \triangleq (\gamma/2)(N_0 W/M)/(\beta_{\max} R(\nu)((n-1)\hat{\alpha}q_n - \gamma/2 (N_0 W/M))$, we have $((n-1)\hat{\alpha}q_n)/(1+\beta_{\max} \xi(\nu)R(\nu)) + (N_0 W/M) = (n-1)\hat{\alpha}q_n + (1-\gamma/2)(N_0 W/M)$, and as a result,

$$\begin{aligned} \mathbb{P}\{\mathcal{O}_l^{(j)}\} &\geq 1 - e^{-((\gamma/2)N_0 W/M)[(n-1)\hat{\alpha}q_n + (1-\gamma/2)(N_0 W/M)]/(\beta_{\max}[(n-1)\hat{\alpha}q_n - ((\gamma/2)N_0 W/M)])(1-t(\nu)/R(\nu))} \\ &\geq 1 - e^{-(\gamma N_0 W)/(2M\beta_{\max} (1-t(\nu))/(R(\nu)))}. \end{aligned} \quad (81)$$

Since $(\gamma N_0 W)/(2M\beta_{\max}) = \Theta(1)$, it follows that the necessary condition to have $\mathbb{P}\{\mathcal{O}_l^{(j)}\} \rightarrow 0$ is having $R(\nu) \lesssim t(\nu) = (p_l \nu)/((n-1)\hat{\alpha}q_n + (1-\gamma/2)(N_0 W/M))$. In other words,

$$R^*(\nu) \lesssim \frac{p_l \nu}{(n-1)\hat{\alpha}q_n + (1-\gamma/2)(N_0 W/M)}, \quad (82)$$

which implies that \bar{R}_g defined in (5) is upper bounded by

$$\bar{R}_g \lesssim nW \mathbb{E}_\nu \left[\frac{p_l \nu}{(n-1)\hat{\alpha}q_n + (1-\gamma/2)(N_0 W/M)} \right] \quad (83)$$

$$= \frac{nW \mathbb{E}_\nu [p_l \nu]}{(n-1)\hat{\alpha}q_n + (1-\gamma/2)(N_0 W/M)}. \quad (84)$$

Now, defining $\Psi_n \triangleq \log n + 2 \log \log n$, we have

$$\mathbb{E}[p_l \nu] \leq \mathbb{E}[p_l \nu \mid \nu \leq \Psi_n] \mathbb{P}\{\nu \leq \Psi_n\} + \mathbb{E}[p_l \nu \mid \nu > \Psi_n] \mathbb{P}\{\nu > \Psi_n\}, \quad (85)$$

$$\stackrel{(a)}{\leq} q_n \Psi_n + \mathbb{E}[\nu \mid \nu > \Psi_n] \mathbb{P}\{\nu > \Psi_n\}, \quad (86)$$

$$\stackrel{(b)}{\leq} q_n \Psi_n + (\Psi_n + 1)e^{-\Psi_n}, \quad (87)$$

$$\stackrel{(c)}{\lesssim} q_n \log n. \quad (88)$$

In the above equation, (a) comes from the facts that

$$\begin{aligned} \mathbb{E}[p_l \nu \mid \nu \leq \Psi_n] \mathbb{P}\{\nu \leq \Psi_n\} &\leq \Psi_n \mathbb{E}[p_l \mid \nu \leq \Psi_n] \mathbb{P}\{\nu \leq \Psi_n\} \\ &\leq \Psi_n \mathbb{E}[p_l] = \Psi_n q_n, \end{aligned} \quad (89)$$

and $0 \leq p_l \leq 1$. (b) results from the fact that ν is exponentially distributed. (c) follows from the facts that (i) as we are considering the strong and moderate interference scenarios, it yields that $(n-1)\hat{\alpha}q_n = \Omega(1)$, or equivalently, $q_n = \Omega(1/n)$, and (ii) the term $(\Psi_n + 1)e^{-\Psi_n}$ scales as $1/n \log n$

(due to the definition of Ψ_n) which is negligible with respect to the first term $q_n \Psi_n$. Combining (84) and (88) yields

$$\bar{R}_g \lesssim \frac{W n q_n \log n}{(n-1)\hat{\alpha}q_n + (1-\gamma/2)(N_0 W/M)} \quad (90)$$

$$\lesssim \frac{W}{\hat{\alpha}} \log n \quad (91)$$

$$\lesssim \frac{W}{\hat{\alpha}} \log K. \quad (92)$$

In the case of weak interference, we have

$$\begin{aligned} \bar{R}_g &\leq nW \frac{\mathbb{E}[p_l \nu]}{(N_0 W/M)} \\ &= \frac{Mn}{N_0} \mathbb{E}[p_l \nu]. \end{aligned} \quad (93)$$

Rewriting (87), we obtain

$$\mathbb{E}[p_l \nu] \leq q_n \Psi_n + (\Psi_n + 1)e^{-\Psi_n}, \quad \forall \Psi_n > 0. \quad (94)$$

Selecting $\Psi_n = \log(q_n^{-2})$ and defining $\varepsilon \triangleq nq_n$, we have

$$\bar{R}_g \lesssim \frac{2M\varepsilon}{N_0} (\log n - \log(\varepsilon^{-1})). \quad (95)$$

As in the weak interference scenario we have $\varepsilon = o(1)$, it follows from the above equation that $\bar{R}_g = o(W \log n)$ in this scenario. Comparing with (92), it follows that

$$\bar{R}_g \lesssim \frac{W}{\hat{\alpha}} \log K. \quad (96)$$

Lower Bound. For the lower-bound, we consider the on-off power allocation scheme with $\tau_n = \log n - 2 \log \log n$. Also, assume that $M = 1$ (or equivalently, $n = K$). Noting $q_n = e^{-\tau_n}$, we obtain

$$\mathbb{E}[I_l] = (n-1)\hat{\alpha}q_n = \Theta(\log^2 n). \quad (97)$$

Therefore, using the result of Lemma 3.1, it is realized that with probability one $(n-1)\hat{\alpha}q_n(1-\epsilon) \leq I_l \leq (n-1)\hat{\alpha}q_n(1+\epsilon)$, for some $\epsilon = o(1)$. In other words, defining

$$\Phi(\nu) \triangleq \log\left(1 + \frac{p_l \nu}{(n-1)\hat{\alpha}q_n(1+\epsilon) + (N_0 W/M)}\right), \quad (98)$$

it follows that

$$\mathbb{P}\{R_l(\mathbf{P}^{(j)}, \mathcal{L}_l^{(j)}) < \Phi(\nu)\} = o(1), \quad (99)$$

which implies that $R^*(\nu) \geq \Phi(\nu)$. As a result,

$$\begin{aligned} \bar{R}_g &\geq nW\mathbb{E}[\Phi(\nu)] \\ &= nW\mathbb{E}\left[\log\left(1 + \frac{p_l \nu}{(n-1)\hat{\alpha}q_n(1+\epsilon) + (N_0 W/M)}\right)\right] \\ &\stackrel{(a)}{=} nW \int_{\tau_n}^{\infty} \log\left(1 + \frac{\nu}{(n-1)\hat{\alpha}q_n(1+\epsilon) + (N_0 W/M)}\right) e^{-\nu} d\nu \\ &\geq nW \int_{\tau_n}^{\Psi_n} \log\left(1 + \frac{\nu}{(n-1)\hat{\alpha}q_n(1+\epsilon) + (N_0 W/M)}\right) e^{-\nu} d\nu, \end{aligned} \quad (100)$$

where $\Psi_n \triangleq \log n + 2 \log \log n$ and (a) follows from the on-off power allocation assumption. As $(n-1)\hat{\alpha}q_n(1+\epsilon) = \Theta(\log^2 n)$, it follows that $\nu/((n-1)\hat{\alpha}q_n(1+\epsilon) + (N_0 W)/M) = o(1)$ in the interval $[\tau_n, \Psi_n]$, which implies that

$$\begin{aligned} &\log\left(1 + \frac{\nu}{(n-1)\hat{\alpha}q_n(1+\epsilon) + (N_0 W/M)}\right) \\ &\sim \frac{\nu}{(n-1)\hat{\alpha}q_n(1+\epsilon) + (N_0 W/M)}, \end{aligned} \quad (101)$$

in the interval of integration $[\tau_n, \Psi_n]$. Hence,

$$\begin{aligned} \bar{R}_g &\gtrsim nW \int_{\tau_n}^{\Psi_n} \frac{\nu}{(n-1)\hat{\alpha}q_n(1+\epsilon) + (N_0 W/M)} e^{-\nu} d\nu \\ &= \frac{nW}{(n-1)\hat{\alpha}q_n(1+\epsilon) + (N_0 W/M)} \int_{\tau_n}^{\Psi_n} \nu e^{-\nu} d\nu \\ &= \frac{nW}{(n-1)\hat{\alpha}q_n(1+\epsilon) + (N_0 W/M)} \\ &\quad \times \left((\tau_n + 1)e^{-\tau_n} - (\Psi_n + 1)e^{-\Psi_n} \right) \\ &\stackrel{(a)}{\sim} \frac{nW\tau_n q_n}{(n-1)\hat{\alpha}q_n(1+\epsilon) + (N_0 W/M)} \\ &\sim \frac{W}{\hat{\alpha}} \log n \\ &= \frac{W}{\hat{\alpha}} \log K, \end{aligned} \quad (102)$$

where (a) results from the facts that $(\Psi_n + 1)e^{-\Psi_n} \ll (\tau_n + 1)e^{-\tau_n}$ and $e^{-\tau_n} = q_n$. Combining the above equation with (96), the proof of Theorem 4.1 follows. \square

Remark 7. Similar to the proof steps of Theorem 3.3, it can be shown that the optimum value of M is equal to one. In fact, since the maximum guaranteed sum-rate of the network is achieved in the strong interference scenario in which the interference term scales as $n\hat{\alpha}q_n$ with probability one, it follows that the maximum network average sum-rate and the network guaranteed sum-rate are equal. Therefore, the optimum spectrum sharing for maximizing the network guaranteed sum-rate is the same as the one maximizing the average sum-rate of the network ($M = 1$).

5. Conclusion

In this paper, a distributed single-hop wireless network with K links was considered, where the links were partitioned into a fixed number (M) of clusters each operating in a subchannel with bandwidth W/M . The subchannels were assumed to be orthogonal to each other. A general shadow-fading model, described by parameters (α, ω) , was considered, where α denotes the probability of shadowing and ω ($\omega \leq 1$) represents the average cross-link gains. The maximum achievable network throughput was studied in the asymptotic regime of $K \rightarrow \infty$. In the first part of the paper, the network throughput is defined as the *average sum-rate* of the network, which is shown to scale as $\Theta(\log K)$. Moreover, it was proved that the optimum power allocation strategy for each user was a threshold-based on-off scheme, when K is large. To achieve this performance metric, each user chooses a noniterative power allocation strategy based on its direct channel gain as a local information. This approach prevents imposing more interference on the other links when the channel condition is poor. The main advantage of this virtual cooperation is that the network nodes cooperate unselfishly to improve the network throughput instead of solely increasing their rates. In the second part, the network throughput is defined as the *guaranteed sum-rate*, when the outage probability approaches zero. In this scenario, it was demonstrated that the on-off power allocation scheme maximizes the network guaranteed sum-rate, which scales as $(W/\hat{\alpha}) \log K$. Moreover, the optimum spectrum sharing for maximizing the average sum-rate and guaranteed sum-rate is achieved at $M = 1$.

The optimum power allocation policy proposed in this paper maximizes the throughput of the network under the assumption of a Rayleigh fading channel with the shadowing effect, while ignoring the effect of the distance-based propagation loss. The proposed channel model can be considered as a special case of a multiple access channel, where the distance between each user and its corresponding receiver (or with an access point) is the same as that of the other links. In this case, the distance-based propagation loss only changes the scaling factor in the throughput maximization, and we have the same scaling $\Theta(K)$ for the average sum-rate of the network. Our future research involves considering the effect of the path-loss channel model on the optimum power allocation policy and the throughput maximization, where we assume that the distance between nodes in each link is not necessarily the same.

Appendices

A. Proof of Lemma 3.1

Let us define $\chi_k \triangleq \mathcal{L}_{ki} p_k$, where \mathcal{L}_{ki} is independent of p_k , for $k \neq i$. Under a quasi static Rayleigh fading channel model, it is concluded that χ_k 's are independent and identically distributed (i.i.d.) random variables with

$$\begin{aligned} \mathbb{E}[\chi_k] &= \mathbb{E}[\mathcal{L}_{ki} p_k] = \hat{\alpha} q_n, \\ \text{Var}[\chi_k] &= \mathbb{E}[\chi_k^2] - \mathbb{E}^2[\chi_k] \\ &\stackrel{(a)}{\leq} 2\alpha\kappa q_n - (\hat{\alpha} q_n)^2, \end{aligned} \quad (\text{A.1})$$

where $\mathbb{E}[h_{ki}^2] = 2$ and $\hat{\alpha} \triangleq \alpha\omega$. Also, (a) follows from the fact that $p_k^2 \leq p_k$. Thus, $\mathbb{E}[p_k^2] \leq \mathbb{E}[p_k] = q_n$. The interference $I_i = \sum_{k \in \mathbb{C}_j, k \neq i} \chi_k$ is a random variable with mean μ_n and variance ϑ_n^2 , where

$$\begin{aligned} \mu_n &\triangleq \mathbb{E}[I_i] = (n-1)\hat{\alpha} q_n, \\ \vartheta_n^2 &\triangleq \text{Var}[I_i] \leq (n-1)(2\alpha\kappa q_n - (\hat{\alpha} q_n)^2) \leq (n-1)(2\alpha\kappa q_n). \end{aligned} \quad (\text{A.2})$$

Using the *Central Limit Theorem* [30, page 183] we obtain

$$\begin{aligned} \mathbb{P}\{|I_i - \mu_n| < \psi_n\} &\approx 1 - Q\left(\frac{\psi_n}{\vartheta_n}\right) \\ &\stackrel{(a)}{\geq} 1 - e^{-(\psi_n^2)/(2\vartheta_n^2)} \end{aligned} \quad (\text{A.3})$$

for all $\psi_n > 0$ such that $\psi_n = o(n^{1/6}\vartheta_n)$. In the above equation, the $Q(\cdot)$ function is defined as $Q(x) \triangleq 1/\sqrt{2\pi} \int_x^\infty e^{-u^2/2} du$, and (a) follows from the fact that $Q(x) \leq e^{-x^2/2}$, $\forall x > 0$. Selecting $\psi_n = (nq_n)^{1/8} \sqrt{2}\vartheta_n$, we obtain

$$\mathbb{P}\{|I_i - \mu_n| < \psi_n\} \geq 1 - e^{-(nq_n)^{1/4}}. \quad (\text{A.4})$$

Therefore, defining $\varepsilon \triangleq \psi_n/\mu_n = O((nq_n)^{-3/8})$, we have

$$\mathbb{P}\{\mu_n(1-\varepsilon) \leq I_i \leq \mu_n(1+\varepsilon)\} \geq 1 - e^{-(nq_n)^{1/4}}. \quad (\text{A.5})$$

Noting that $nq_n \rightarrow \infty$, it follows that $I_i \sim \mu_n$, with probability one. Now, we show a stronger statement, which is, the contribution of the realizations in which $|I_i - \mu_n| > \psi_n$ in the network average sum-rate is negligible. For this purpose, we give a lower-bound and an upper bound for the network average sum-rate and show that these bounds converge to each other when $nq_n \rightarrow \infty$. A lower-bound denoted by $\bar{R}_{\text{ave}}^{(L)}$, can be given by

$$\begin{aligned} \bar{R}_{\text{ave}}^{(L)} &\triangleq nW \mathbb{E} \left[\log \left(1 + \frac{\hat{p}_i h_{ii}}{I_i + (N_0 W/M)} \right) \mid |I_i - \mu_n| < \psi_n \right] \\ &\quad \times \mathbb{P}\{|I_i - \mu_n| < \psi_n\} \\ &\geq nW \mathbb{E} \left[\log \left(1 + \frac{\hat{p}_i h_{ii}}{\mu_n(1+\varepsilon) + (N_0 W/M)} \right) \right] \\ &\quad \times \left[1 - e^{-(nq_n)^{1/4}} \right], \end{aligned} \quad (\text{A.6})$$

which scales as $W/\hat{\alpha} \log n$ (as shown in the proof of Theorem 3.3, by optimizing the power allocation function). An upper bound for the network average sum-rate, denoted by $\bar{R}_{\text{ave}}^{(U)}$, can be given as

$$\begin{aligned} \bar{R}_{\text{ave}}^{(U)} &= nW \mathbb{E} \left[\log \left(1 + \frac{\hat{p}_i h_{ii}}{I_i + (N_0 W/M)} \right) \mid |I_i - \mu_n| < \psi_n \right] \\ &\quad \times \mathbb{P}\{|I_i - \mu_n| < \psi_n\} \\ &\quad + nW \mathbb{E} \left[\log \left(1 + \frac{\hat{p}_i h_{ii}}{I_i + (N_0 W/M)} \right) \mid |I_i - \mu_n| \geq \psi_n \right] \\ &\quad \times \mathbb{P}\{|I_i - \mu_n| \geq \psi_n\} \\ &\leq \bar{R}_{\text{ave}}^{(L)} + nW \mathbb{E} \left[\log \left(1 + \frac{\hat{p}_i h_{ii}}{(N_0 W/M)} \right) \right] e^{-(nq_n)^{1/4}} \\ &\stackrel{(a)}{\leq} \bar{R}_{\text{ave}}^{(L)} + nW \mathbb{E} \left[\frac{\hat{p}_i h_{ii}}{(N_0 W/M)} \right] e^{-(nq_n)^{1/4}} \\ &\stackrel{(b)}{\leq} \bar{R}_{\text{ave}}^{(L)} + WO(nq_n \log n) e^{-(nq_n)^{1/4}} \\ &\stackrel{(c)}{\lesssim} \bar{R}_{\text{ave}}^{(L)} \end{aligned} \quad (\text{A.7})$$

In the above equation, (a) follows from the fact that $\log(1+x) \leq x$, $\forall x \geq 0$, (b) comes from the facts that $\mathbb{E}\{\hat{p}_i h_{ii}\} \lesssim q_n \log n$ (this is shown in the proof of Theorem 4.1) and $(N_0 W)/M$ is fixed, and finally, (c) results from the fact that as $nq_n \rightarrow \infty$, $nq_n e^{-(nq_n)^{1/4}} \rightarrow 0$. The above equation implies that substituting I_i by its mean $((n-1)\hat{\alpha} q_n)$ does not affect the analysis of the network average sum-rate in the asymptotic case of $K \rightarrow \infty$.

B. Proof of Lemma 3.2

Denoting $\mathbb{T}_j \triangleq \{l \in \mathbb{C}_j \mid h_{ll} > h_{Th}\}$, the cardinality of the set \mathbb{T}_j is a binomial random variable with the mean $n\mathbb{P}\{h_{ll} > h_{Th}\}$. From (4), we have

$$\bar{R}_{\text{ave}} = \sum_{j=1}^M \mathbb{E} \left[\sum_{l \in \mathbb{C}_j} R_l(\hat{\mathbf{P}}^{(j)}, \boldsymbol{\mathcal{L}}_l^{(j)}) \right], \quad (\text{B.1})$$

where

$$\begin{aligned} \mathbb{E} \left[\sum_{l \in \mathbb{C}_j} R_l(\hat{\mathbf{P}}^{(j)}, \boldsymbol{\mathcal{L}}_l^{(j)}) \right] &= \mathbb{E} \left[\sum_{l \in \mathbb{T}_j} R_l(\hat{\mathbf{P}}^{(j)}, \boldsymbol{\mathcal{L}}_l^{(j)}) \right] \\ &\quad + \mathbb{E} \left[\sum_{l \in \mathbb{T}_j^c} R_l(\hat{\mathbf{P}}^{(j)}, \boldsymbol{\mathcal{L}}_l^{(j)}) \right], \end{aligned} \quad (\text{B.2})$$

in which \mathbb{T}_j^C denotes the complement of \mathbb{T}_j . Note that

$$\begin{aligned} & \mathbb{E} \left[\sum_{l \in \mathbb{T}_j^C} R_l(\hat{\mathbf{P}}^{(j)}, \mathcal{L}_l^{(j)}) \right] \\ &= n \frac{W}{M} \mathbb{E} \left[\log \left(1 + \frac{h_{ll} \hat{p}_l}{I_l + (N_0 W)/M} \right) \mid h_{ll} > h_{Th} \right] \\ & \quad \times \mathbb{P}\{h_{ll} > h_{Th}\} \end{aligned} \quad (\text{B.3})$$

$$\begin{aligned} & \leq n \frac{W}{M} \mathbb{E} \left[\log \left(1 + \frac{h_{ll}}{(N_0 W)/M} \right) \mid h_{ll} > h_{Th} \right] \\ & \quad \times \mathbb{P}\{h_{ll} > h_{Th}\} \end{aligned} \quad (\text{B.4})$$

$$\stackrel{(a)}{\leq} \frac{n}{N_0} e^{-h_{Th}} \mathbb{E}[h_{ll} \mid h_{ll} > h_{Th}] \quad (\text{B.5})$$

$$= \frac{n}{N_0} e^{-h_{Th}} (1 + h_{Th}), \quad (\text{B.6})$$

where (a) follows from $\log(1+x) \leq x$, for $x \geq 0$. It is observed that for $h_{Th} = c \log n$, where $c > 1$, the right hand side of (B.6) tends to zero as $n \rightarrow \infty$. Thus,

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\sum_{l \in \mathbb{T}_j^C} R_l(\hat{\mathbf{P}}^{(j)}, \mathcal{L}_l^{(j)}) \right] = 0. \quad (\text{B.7})$$

Consequently,

$$\lim_{n \rightarrow \infty} \sum_{j=1}^M \mathbb{E} \left[\sum_{l \in \mathbb{T}_j} R_l(\hat{\mathbf{P}}^{(j)}, \mathcal{L}_l^{(j)}) \right] = 0, \quad (\text{B.8})$$

and this completes the proof of the lemma.

C. Proof of Lemma 3.4

Using (31), we have

$$\begin{aligned} \mathbb{E}[\Xi_i(\hat{p}_i, h_{ii})] &\approx \frac{W}{M\lambda} \mathbb{E}[h_{ii} \hat{p}_i] + n \frac{\alpha W \mu}{M\lambda'} \\ &\quad \times \left(1 - \frac{\omega}{\lambda'} \mathbb{E}[\hat{p}_i] + \frac{2\kappa}{\lambda'^2} \mathbb{E}[\hat{p}_i^2] - \frac{6\eta}{\lambda'^3} \mathbb{E}[\hat{p}_i^3] \right) \\ &\quad + n(1-\alpha) \frac{W\mu}{M\lambda'}, \end{aligned} \quad (\text{C.1})$$

$$\begin{aligned} &\stackrel{(a)}{=} \frac{W}{M\lambda} (1 + \tau_n) q_n - \frac{n\hat{\alpha}W}{M\lambda'^2} (1 + \tau_n) q_n^2 \\ &\quad + \frac{n\alpha W 2\kappa}{M\lambda'^3} (1 + \tau_n) q_n^2 \\ &\quad - \frac{n\alpha W 6\eta}{M\lambda'^4} (1 + \tau_n) q_n^2 + \frac{nW}{M\lambda'} (1 + \tau_n) q_n \end{aligned} \quad (\text{C.2})$$

$$\stackrel{(b)}{\approx} \frac{W}{M\hat{\alpha}} \left(1 + \tau_n + \frac{\xi_1}{n^2} (1 + \tau_n) e^{\tau_n} - \frac{\xi_2}{n^3} (1 + \tau_n) e^{2\tau_n} \right), \quad (\text{C.3})$$

where $\xi_1 \triangleq 2\kappa/(\omega\hat{\alpha})$ and $\xi_2 \triangleq 6\eta/(\omega\hat{\alpha}^2)$. In the above equation, (a) follows from the fact that $\mathbb{E}[h_{ii}\hat{p}_i] = \mu = (1 + \tau_n)q_n$, and (b) results from (i) $\lambda = (n-1)\hat{\alpha}q_n + (N_0W/M) \approx n\hat{\alpha}q_n$ and $\lambda' \approx n\hat{\alpha}q_n$ incurred by the fact that $\lambda \gg 1$ and (ii) $q_n = e^{-\tau_n}$. Since $n\hat{\alpha}q_n \rightarrow \infty$, it follows that the right hand side of (C.3) is a monotonically increasing function of τ_n , which attains its maximum when τ_n takes its maximum feasible value. The maximum feasible value of τ_n , denoted as $\hat{\tau}_n$, can be obtained as

$$n\hat{\alpha}e^{-\hat{\tau}_n} \rightarrow \infty \implies \hat{\tau}_n \sim \log n. \quad (\text{C.4})$$

Thus, the maximum achievable value for $\mathbb{E}[\Xi_i(\hat{p}_i, h_{ii})]$ scales as $W/(M\hat{\alpha}) \log n$.

D. Proof of Lemma 3.5

(i) Using (8) and assuming that all users follow the on-off power allocation policy, $\mathbb{E}[u_i(\hat{p}_i, h_{ii})]$ can be expressed as

$$\mathbb{E}[u_i(\hat{p}_i, h_{ii})] = \sum_{l \in \mathbb{C}_j} \mathbb{E}[R_l(\hat{\mathbf{P}}^{(j)}, \mathcal{L}_l^{(j)})], \quad j = 1, \dots, M, \quad (\text{D.1})$$

where the expectation is computed with respect to h_{ll} and I_l . Noting that $q_n = \mathbb{P}\{h_{ll} > \tau_n\}$, we have

$$\begin{aligned} \mathbb{E}[R_l(\hat{\mathbf{P}}^{(j)}, \mathcal{L}_l^{(j)})] &= \mathbb{E}[R_l(\hat{\mathbf{P}}^{(j)}, \mathcal{L}_l^{(j)}) \mid h_{ll} > \tau_n] \\ &\quad \times \mathbb{P}\{h_{ll} > \tau_n\} \\ &\quad + \mathbb{E}[R_l(\hat{\mathbf{P}}^{(j)}, \mathcal{L}_l^{(j)}) \mid h_{ll} \leq \tau_n] \\ &\quad \times \mathbb{P}\{h_{ll} \leq \tau_n\} \\ &= q_n \mathbb{E}[R_l(\hat{\mathbf{P}}^{(j)}, \mathcal{L}_l^{(j)}) \mid h_{ll} > \tau_n] \\ &\quad + (1 - q_n) \mathbb{E}[R_l(\hat{\mathbf{P}}^{(j)}, \mathcal{L}_l^{(j)}) \mid h_{ll} \leq \tau_n]. \end{aligned} \quad (\text{D.2})$$

Since for $h_{ll} \leq \tau_n$, $\hat{p}_l = 0$, it is concluded that

$$\begin{aligned} &\mathbb{E}[R_l(\hat{\mathbf{P}}^{(j)}, \mathcal{L}_l^{(j)})] \\ &= \frac{q_n W}{M} \mathbb{E} \left[\log \left(1 + \frac{h_{ll}}{I_l + (N_0 W)/M} \right) \mid h_{ll} > \tau_n \right]. \end{aligned} \quad (\text{D.3})$$

For large values of K , we can apply Lemma 3.1 to obtain

$$\begin{aligned} & \mathbb{E}\left[R_l(\hat{\mathbf{P}}^{(j)}, \mathcal{L}_l^{(j)})\right] \\ & \approx \frac{q_n W}{M} \mathbb{E}\left[\log\left(1 + \frac{h_{ll}}{(n-1)\hat{\alpha}q_n + (N_0 W)/M}\right) \mid h_{ll} > \tau_n\right] \end{aligned} \quad (\text{D.4})$$

$$= \frac{q_n W}{M} \mathbb{E}\left[\log\left(1 + \frac{h_{ll}}{\lambda}\right) \mid h_{ll} > \tau_n\right], \quad (\text{D.5})$$

where the expectation is computed with respect to h_{ll} . Using the Taylor series for $\log(1+x)$, (D.5) can be written as

$$\begin{aligned} \mathbb{E}\left[R_l(\hat{\mathbf{P}}^{(j)}, \mathcal{L}_l^{(j)})\right] & \approx \frac{q_n W}{M} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k\lambda^k} \mathbb{E}\left[h_{ll}^k \mid h_{ll} > \tau_n\right] \\ & \stackrel{(a)}{\approx} \frac{q_n W}{M} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k(n\hat{\alpha}q_n)^k} \mathbb{E}\left[h_{ll}^k \mid h_{ll} > \tau_n\right] \\ & \stackrel{(b)}{\approx} \frac{q_n W}{M} \sum_{k=1}^{\infty} \frac{(-1)^{k-1} \tau_n^k}{k(n\hat{\alpha}q_n)^k} \\ & = \frac{q_n W}{M} \log\left(1 + \frac{\tau_n}{n\hat{\alpha}q_n}\right) \\ & \stackrel{(c)}{=} \frac{e^{-\tau_n} W}{M} \log\left(1 + \frac{\tau_n e^{\tau_n}}{n\hat{\alpha}}\right), \end{aligned} \quad (\text{D.6})$$

where (a) follows from the fact that for large values of n , $\lambda \approx n\hat{\alpha}q_n$. Also, (b) results from the fact that under a Rayleigh fading channel model,

$$\mathbb{E}[h_{ll} \mid h_{ll} > \tau_n] = 1 + \tau_n, \quad (\text{D.7})$$

$$\mathbb{E}[h_{ll}^k \mid h_{ll} > \tau_n] = \tau_n^k + k\mathbb{E}[h_{ll}^{k-1} \mid h_{ll} > \tau_n].$$

Since $\lambda \gg 1$, the term $\mathbb{E}[h_{ll}^{(k-1)} \mid h_{ll} > \tau_n]/\lambda^k \ll \mathbb{E}[h_{ll}^{k-1} \mid h_{ll} > \tau_n]/\lambda^{k-1}$, which implies that we can neglect this term and simply write $\mathbb{E}[h_{ll}^k \mid h_{ll} > \tau_n] \approx \tau_n^k$. (c) results from $q_n = e^{-\tau_n}$. Thus, (D.1) can be simplified as

$$\mathbb{E}[u_i(\hat{p}_i, h_{ii})] \approx \frac{ne^{-\tau_n} W}{M} \log\left(1 + \frac{\tau_n e^{\tau_n}}{n\hat{\alpha}}\right). \quad (\text{D.8})$$

In order to find the optimum threshold value:

$$\hat{\tau}_n = \arg \max_{\tau_n} \mathbb{E}[u_i(\hat{p}_i, h_{ii})], \quad (\text{D.9})$$

we set the derivative of the right hand side of (D.8) with respect to τ_n to zero

$$-e^{-\hat{\tau}_n} \log\left(1 + \frac{\hat{\tau}_n e^{\hat{\tau}_n}}{n\hat{\alpha}}\right) + \frac{1 + \hat{\tau}_n}{n\hat{\alpha} + \hat{\tau}_n e^{\hat{\tau}_n}} = 0, \quad (\text{D.10})$$

which after some manipulations yields

$$\hat{\tau}_n = \log n - 2 \log \log n + O(1). \quad (\text{D.11})$$

(ii) Using (D.11), it is concluded that

$$q_n = e^{-\tau_n} = \delta \frac{\log^2 n}{n}, \quad (\text{D.12})$$

where δ is a constant.

(iii) Using (D.11), we have

$$\frac{\hat{\tau}_n e^{\hat{\tau}_n}}{n\hat{\alpha}} = \Theta\left(\frac{1}{\log n}\right), \quad (\text{D.13})$$

which implies that the right hand side of (D.8) can be written as

$$\text{RH(D-8)} \approx \frac{W\hat{\tau}_n}{M\hat{\alpha}}. \quad (\text{D.14})$$

Thus, the maximum value for $\mathbb{E}[u_i(\hat{p}_i, h_{ii})]$ in (D.8) scales as $W/(M\hat{\alpha}) \log n$.

E. Proof of Theorem 3.7

Let us define \mathbb{A}_j as the set of active links in cluster j . The random variable m_j denotes the cardinality of the set \mathbb{A}_j . Noting that for $M = \Theta(K)$, $\lim_{K \rightarrow \infty} (M/K)$ is constant, it is concluded that n and $m_j \in [1, n]$ do not grow with K . To obtain the network average sum-rate, we assume that among M clusters, Γ clusters have $m_j = 1$ and the rest have $m_j > 1$. We first obtain an upper bound on the average sum-rate in each cluster when $m_j = 1$, $1 \leq j \leq M$. Clearly, since only one user in each cluster activates its transmitter, $I_i = 0$. Thus, by using (47), the maximum achievable average sum-rate of cluster \mathbb{C}_j is computed as

$$\bar{R}_{\text{ave}}^{(j)} = \frac{W}{M} \mathbb{E}\left[\log\left(1 + \frac{M}{N_0 W} h_{\max}\right)\right], \quad (\text{E.1})$$

where $h_{\max} \triangleq \max\{h_{ii}\}_{i \in \mathbb{C}_j}$ is a random variable. Since $\log x$ is a concave function of x , an upper bound of (E.1) is obtained through *Jensen's inequality*, $\mathbb{E}[\log x] \leq \log(\mathbb{E}[x])$, $x > 0$. Thus,

$$\bar{R}_{\text{ave}}^{(j)} \leq \frac{W}{M} \log\left(1 + \frac{M}{N_0 W} \mathbb{E}[h_{\max}]\right). \quad (\text{E.2})$$

Under a Rayleigh fading channel model and noting that $\{h_{ii}\}$ is a set of i.i.d. random variables over $i \in \mathbb{C}_j$, we have

$$\begin{aligned} F_{h_{\max}}(y) &= \mathbb{P}\{h_{\max} \leq y\}, \quad y > 0 \\ &= \prod_{i \in \mathbb{C}_j} \mathbb{P}\{h_{ii} \leq y\} \\ &= (1 - e^{-y})^n, \end{aligned} \quad (\text{E.3})$$

where $F_{h_{\max}}(\cdot)$ is the cumulative distribution function (CDF) of h_{\max} . Hence,

$$\mathbb{E}[h_{\max}] = \int_0^{\infty} n y e^{-y} (1 - e^{-y})^{n-1} dy. \quad (\text{E.4})$$

Since $(1 - e^{-y})^{n-1} \leq 1$, we arrive at the following inequality:

$$\mathbb{E}[h_{\max}] \leq \int_0^\infty ny e^{-y} dy = n. \quad (\text{E.5})$$

Consequently, the upper bound of (E.2) can be simplified as

$$\bar{R}_{\text{ave}}^{(j)} \leq \frac{W}{M} \log\left(1 + \frac{K}{N_0 W}\right). \quad (\text{E.6})$$

For $m_j > 1$ and due to the shadowing effect with parameters (α, ϖ) , the average sum-rate of cluster \mathbb{C}_j can be written as

$$\bar{R}_{\text{ave}}^{(j)} = \sum_{i \in \mathbb{A}_j} \frac{W}{M} \mathbb{E} \left[\log \left(1 + \frac{h_{ii}}{\sum_{k \in \mathbb{A}_j, k \neq i} u_k \beta_{ki} h_{ki} + (N_0 W)/M} \right) \right], \quad (\text{E.7})$$

where u_k 's are Bernoulli random variables with parameter α . Thus,

$$\begin{aligned} \bar{R}_{\text{ave}}^{(j)} &= \frac{W}{M} \sum_{i \in \mathbb{A}_j} \sum_{l=0}^{m_j-1} \binom{m_j-1}{l} \alpha^l (1-\alpha)^{m_j-1-l} \\ &\quad \times \mathbb{E} \left[\log \left(1 + \frac{h_{ii}}{\Sigma_l + (N_0 W)/M} \right) \right] \\ &= \frac{W}{M} \sum_{i \in \mathbb{A}_j} (1-\alpha)^{m_j-1} \mathbb{E} \left[\log \left(1 + \frac{h_{ii}}{(N_0 W)/M} \right) \right] \quad (\text{E.8}) \\ &\quad + \frac{W}{M} \sum_{i \in \mathbb{A}_j} \sum_{l=1}^{m_j-1} \binom{m_j-1}{l} \alpha^l (1-\alpha)^{m_j-1-l} \\ &\quad \times \mathbb{E} \left[\log \left(1 + \frac{h_{ii}}{\Sigma_l + (N_0 W)/M} \right) \right], \end{aligned}$$

where Σ_l is the sum of l i.i.d random variables $\{Z_i\}_{i=1}^l$, where $Z_i \triangleq \beta_{ki} h_{ki}$, $k \neq i$. For $m_j > 1$, Σ_l is greater than the interference term caused by one interfering link. Thus, an upper bound on the average sum-rate of cluster \mathbb{C}_j is computed as

$$\begin{aligned} \bar{R}_{\text{ave}}^{(j)} &\leq \frac{W}{M} m_j (1-\alpha)^{m_j-1} \mathbb{E} \left[\log \left(1 + \frac{Y}{(N_0 W)/M} \right) \right] \\ &\quad + \frac{W}{M} \sum_{i \in \mathbb{A}_j} \sum_{l=1}^{m_j-1} \binom{m_j-1}{l} \alpha^l (1-\alpha)^{m_j-1-l} \quad (\text{E.9}) \\ &\quad \times \mathbb{E} \left[\log \left(1 + \frac{Y}{Z_i + (N_0 W)/M} \right) \right], \end{aligned}$$

where $Y \triangleq h_{\max} = \max\{h_{ii}\}_{i \in \mathbb{C}_j}$. According to binomial formula, we have

$$\sum_{l=1}^{m_j-1} \binom{m_j-1}{l} \alpha^l (1-\alpha)^{m_j-1-l} = 1 - (1-\alpha)^{m_j-1}. \quad (\text{E.10})$$

Thus,

$$\begin{aligned} \bar{R}_{\text{ave}}^{(j)} &\leq \frac{W}{M} m_j (1-\alpha)^{m_j-1} \mathbb{E} \left[\log \left(1 + \frac{Y}{(N_0 W)/M} \right) \right] \\ &\quad + \frac{W}{M} m_j (1 - (1-\alpha)^{m_j-1}) \quad (\text{E.11}) \\ &\quad \times \mathbb{E} \left[\log \left(1 + \frac{Y}{\beta_{ki} h_{ki} + (N_0 W)/M} \right) \right]. \end{aligned}$$

We have

$$\mathbb{E} \left[\log \left(1 + \frac{Y}{\beta_{ki} h_{ki} + (N_0 W)/M} \right) \right] \leq \mathbb{E} \left[\log \left(1 + \frac{Y}{\beta_{\min} h_{ki}} \right) \right]. \quad (\text{E.12})$$

Defining $Z \triangleq \beta_{\min} h_{ki}$ and $X \triangleq Y/Z$, the CDF of X can be evaluated as

$$\begin{aligned} F_X(x) &= \mathbb{P}\{X \leq x\}, \quad x > 0 \\ &= \mathbb{P}\{Y \leq Zx\} \\ &= \int_0^\infty \mathbb{P}\{Y \leq Zx \mid Z = z\} f_Z(z) dz \quad (\text{E.13}) \\ &= \int_0^\infty (1 - e^{-zx})^n \frac{1}{\beta_{\min}} e^{-z/\beta_{\min}} dz \\ &= \int_0^\infty (1 - e^{-t\beta_{\min}x})^n e^{-t} dt. \end{aligned}$$

Thus, the probability density function of X can be written as

$$\begin{aligned} f_X(x) &= \frac{dF_X(x)}{dx} \\ &= \beta_{\min} \int_0^\infty n t e^{-t(1+\beta_{\min}x)} (1 - e^{-t\beta_{\min}x})^{n-1} dt \quad (\text{E.14}) \\ &\leq \beta_{\min} \int_0^\infty n t e^{-t(1+\beta_{\min}x)} dt \\ &= \frac{n\beta_{\min}}{(1 + \beta_{\min}x)^2}. \end{aligned}$$

Using the above equation, the right hand side of (E.12) can be upper bounded as

$$\begin{aligned} \mathbb{E} \left[\log \left(1 + \frac{Y}{\beta_{\min} h_{ki}} \right) \right] &= \int_0^\infty f_X(x) \log(1+x) dx \\ &\leq n\beta_{\min} \int_0^\infty \frac{\log(1+x)}{(1 + \beta_{\min}x)^2} dx \quad (\text{E.15}) \\ &= \frac{-n \log \beta_{\min}}{1 - \beta_{\min}} \\ &= \Theta(1), \end{aligned}$$

where the last line follows from the fact that $0 < \beta_{\min} \leq 1$. Substituting the above equation in (E.11) yields

$$\begin{aligned} \bar{R}_{\text{ave}}^{(j)} &\leq \frac{W}{M} m_j (1 - \alpha)^{m_j - 1} \mathbb{E} \left[\log \left(1 + \frac{Y}{\frac{N_0 W}{M}} \right) \right] \\ &\quad + \frac{W}{M} m_j (1 - (1 - \alpha)^{m_j - 1}) \Theta(1) \\ &\stackrel{(a)}{\leq} \frac{W}{M} m_j (1 - \alpha)^{m_j - 1} \log \left(1 + \frac{K}{N_0 W} \right) + \Theta \left(\frac{W}{M} \right) \\ &= \frac{W}{M} m_j (1 - \alpha)^{m_j - 1} \log \left(1 + \frac{K}{N_0 W} \right) [1 + o(1)], \end{aligned} \quad (\text{E.16})$$

where (a) follows from (E.6) and the fact that $m_j \in \{2, \dots, n\}$ does not scale with K .

Let us assume, that among M clusters, Γ clusters have $m_j = 1$ and for the $M - \Gamma$ of the rest, the number of active links in each cluster is greater than one. By using (E.6) and (E.16), an upper bound on the network average sum-rate is obtained as

$$\begin{aligned} \bar{R}_{\text{ave}} &\leq \frac{\Gamma W}{M} \log \left(1 + \frac{K}{N_0 W} \right) \\ &\quad + (M - \Gamma) \frac{W}{M} m_j (1 - \alpha)^{m_j - 1} \\ &\quad \times \log \left(1 + \frac{K}{N_0 W} \right) [1 + o(1)]. \end{aligned} \quad (\text{E.17})$$

To compare this upper bounded with the computed network average sum-rate in the case of $M = 1$, we note that as $\omega \leq 1$ and $\alpha < 1$, we have $\hat{\alpha} < 1$, and consequently,

$$\frac{\Gamma W}{M} \log \left(1 + \frac{K}{N_0 W} \right) < \frac{\Gamma W}{M \hat{\alpha}} \log \left(1 + \frac{K}{N_0 W} \right). \quad (\text{E.18})$$

To prove that the maximum network average sum-rate obtained in (E.17) is less than that value obtained for $M = 1$ from (18), it is sufficient to show that

$$\begin{aligned} (M - \Gamma) \frac{W}{M} m_j (1 - \alpha)^{m_j - 1} \log \left(1 + \frac{K}{N_0 W} \right) \\ < (M - \Gamma) \frac{W}{M \hat{\alpha}} \log \left(1 + \frac{K}{N_0 W} \right) \end{aligned} \quad (\text{E.19})$$

or

$$m_j (1 - \alpha)^{m_j - 1} < \frac{1}{\hat{\alpha}}. \quad (\text{E.20})$$

Since $\hat{\alpha} \leq \alpha$, it is sufficient to show that $m_j (1 - \alpha)^{m_j - 1} < \frac{1}{\alpha}$.

Defining $\Lambda(\alpha) = \alpha m_j (1 - \alpha)^{m_j - 1}$, we have

$$\frac{\partial \Lambda(\alpha)}{\partial \alpha} = m_j (1 - \alpha)^{m_j - 2} (1 - \alpha m_j). \quad (\text{E.21})$$

Thus, the extremum points of $\Lambda(\alpha)$ are located at $\alpha = 1$ and $\alpha = 1/m_j$, where $m_j \in \{2, \dots, n\}$. It is observed that

$$\begin{aligned} \Lambda(1) &= 0 < 1, \\ \Lambda \left(\frac{1}{m_j} \right) &= \left(\frac{m_j - 1}{m_j} \right)^{m_j - 1} < 1. \end{aligned} \quad (\text{E.22})$$

Since $\Lambda(\alpha) < 1$, we conclude (E.19), which implies that the maximum average sum-rate of the network for $M = \Theta(K)$ is less than that of $M = 1$. Knowing the fact that for $M = o(K)$, similar to the result of Theorem 3.3, one can show that the maximum average sum-rate of the network is achieved at $M = 1$, it is concluded that using the on-off allocation scheme the maximum average sum-rate of the network is achieved at $M = 1$, for all values of $1 \leq M \leq K$.

F. Proof of Remark 4

Using (3) and (4) and for every value of $1 \leq M \leq K$ and $\alpha = 0$, the average sum-rate of the network is simplified as

$$\bar{R}_{\text{ave}} = \sum_{j=1}^M \sum_{i \in \mathcal{C}_j} \mathbb{E} \left[\frac{W}{M} \log \left(1 + \frac{h_{ii}}{(N_0 W)/M} \right) \right], \quad (\text{F.1})$$

where the expectation is computed with respect to h_{ii} . Under a Rayleigh fading channel condition and using the fact that $n = K/M$, (F.1) can be written as

$$\bar{R}_{\text{ave}} = nW \int_0^\infty e^{-x} \log \left(1 + \frac{M}{N_0 W} x \right) dx \quad (\text{F.2})$$

$$= \frac{KW}{M} e^{(N_0 W)/M} \text{E}_1 \left(\frac{N_0 W}{M} \right) \quad (\text{F.3})$$

$$= \frac{KW}{M} e^{(N_0 W)/M} \int_1^\infty \frac{e^{-t(N_0 W)/M}}{t} dt, \quad (\text{F.4})$$

where $\text{E}_1(x) = -\text{Ei}(-x) = \int_1^\infty (e^{-tx}/t) dt$, $x > 0$. Taking the first-order derivative of (F.4) in terms of M yields

$$\frac{\partial \bar{R}_{\text{ave}}}{\partial M} = -\frac{KW}{M^2} e^{(N_0 W)/M} \left(1 + \frac{N_0 W}{M} \right) \text{E}_1 \left(\frac{N_0 W}{M} \right) + \frac{KW}{M^2}. \quad (\text{F.5})$$

Since for every value of $N_0 W$, $(\partial \bar{R}_{\text{ave}})/\partial M$ is negative, it is concluded that the network average sum-rate is a monotonically decreasing function of M . Consequently, the maximum average sum-rate of the network for $\alpha = 0$ and every value of $1 \leq M \leq K$ is achieved at $M = 1$.

G. Proof of Remark 5

From (3) and (4), the average sum-rate of the network is given by

$$\bar{R}_{\text{ave}} = \mathbb{E} \left[\sum_{i=1}^K R_i(\hat{\mathbf{p}}^{(j)}, \mathcal{L}_i^{(j)}) \right] \quad (\text{G.1})$$

$$= \frac{W}{K} \sum_{i=1}^K \mathbb{E} \left[\log \left(1 + \frac{h_{ii}}{(N_0 W)/K} \right) \right],$$

where the expectation is computed with respect to h_{ii} . Under a Rayleigh fading channel condition, we have a

$$\bar{R}_{\text{ave}} = W \int_0^\infty e^{-x} \log \left(1 + \frac{K}{N_0 W} x \right) dx \quad (\text{G.2})$$

$$= W e^{(N_0 W)/K} \text{E}_1 \left(\frac{N_0 W}{K} \right). \quad (\text{G.3})$$

To simplify (G.3), we use the following series representation for $E_1(x)$,

$$E_1(x) = -\gamma + \log\left(\frac{1}{x}\right) + \sum_{s=1}^{\infty} \frac{(-1)^{s+1} x^s}{s \cdot s!}, \quad x > 0, \quad (\text{G.4})$$

where γ is Euler's constant and is defined by the limit [27]

$$\gamma \triangleq \lim_{s \rightarrow \infty} \left(\sum_{k=1}^s \frac{1}{k} - \log s \right) = 0.577215665 \dots \quad (\text{G.5})$$

Thus, (G.3) can be simplified as

$$\begin{aligned} \bar{R}_{\text{ave}} &= W e^{(N_0 W)/K} \\ &\times \left(-\gamma + \log\left(\frac{K}{N_0 W}\right) + \sum_{s=1}^{\infty} \frac{(-1)^{s+1}}{s \cdot s!} \left(\frac{N_0 W}{K}\right)^s \right). \end{aligned} \quad (\text{G.6})$$

In the asymptotic case of $K \rightarrow \infty$,

$$\begin{aligned} e^{(N_0 W)/K} &\approx 1, \\ \sum_{s=1}^{\infty} \frac{(-1)^{s+1}}{s \cdot s!} \left(\frac{N_0 W}{K}\right)^s &\approx 0. \end{aligned} \quad (\text{G.7})$$

Consequently, the network average sum-rate for $M = K$ is asymptotically obtained by

$$\bar{R}_{\text{ave}} \approx W (\log K - \log N_0 W - \gamma). \quad (\text{G.8})$$

Acknowledgments

The work of J. Abouei was performed when he was with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON Canada. A. Bayesteh and M. Ebrahimi are now with RIM. This work is financially supported by funds from the Natural Sciences and Engineering Research Council of Canada (NSERC) and Ontario Centers of Excellence (OCE). The material in this paper was presented in part at the IEEE International Symposium on Information Theory (ISIT), Nice, France, June 24–29, 2007 [31], and at the IEEE Conference on Information Sciences and Systems (CISS), Johns Hopkins University, Baltimore, USA, March 2007 [32].

References

- [1] Y. Liang, V. V. Veeravalli, and H. V. Poor, "Resource allocation for wireless fading relay channels: max-min solution," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3432–3453, 2007.
- [2] K. Kumaran and H. Viswanathan, "Joint power and bandwidth allocation in downlink transmission," *IEEE Transactions on Wireless Communications*, vol. 4, no. 3, pp. 1008–1015, 2005.
- [3] J. Abouei, *Delay-throughput analysis in distributed wireless networks*, Ph.D. thesis, University of Waterloo, 2009.
- [4] E. M. Yeh and R. A. Berry, "Throughput optimal control of cooperative relay networks," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3827–3833, 2007.
- [5] M. Grossglauser and D. N. C. Tse, "Mobility increases the capacity of ad hoc wireless networks," *IEEE/ACM Transactions on Networking*, vol. 10, no. 4, pp. 477–486, 2002.
- [6] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 388–404, 2000.
- [7] S. R. Kulkarni and P. Viswanath, "A deterministic approach to throughput scaling in wireless networks," *IEEE Transactions on Information Theory*, vol. 50, no. 6, pp. 1041–1049, 2004.
- [8] L.-L. Xie and P. R. Kumar, "A network information theory for wireless communication: scaling laws and optimal operation," *IEEE Transactions on Information Theory*, vol. 50, no. 5, pp. 748–767, 2004.
- [9] F. Xue, L.-L. Xie, and P. R. Kumar, "The transport capacity of wireless networks over fading channels," *IEEE Transactions on Information Theory*, vol. 51, no. 3, pp. 834–847, 2005.
- [10] T. ElBatt and A. Ephremides, "Joint scheduling and power control for wireless ad hoc networks," *IEEE Transactions on Wireless Communications*, vol. 3, no. 1, pp. 74–85, 2004.
- [11] I. Katzela and M. Naghshineh, "Channel assignment schemes for cellular mobile telecommunication systems: a comprehensive survey," *IEEE Personal Communications*, vol. 3, no. 3, pp. 10–31, 1996.
- [12] S. G. Kiani and D. Gesbert, "Maximizing the capacity of large wireless networks: optimal and distributed solutions," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '06)*, pp. 2501–2505, Seattle, Wash, USA, July 2006.
- [13] R. D. Yates, "Framework for uplink power control in cellular radio systems," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 7, pp. 1341–1347, 1995.
- [14] G. J. Foschini and Z. Miljanic, "Simple distributed autonomous power control algorithm and its convergence," *IEEE Transactions on Vehicular Technology*, vol. 42, no. 4, pp. 641–646, 1993.
- [15] C. U. Saraydar, N. B. Mandayam, and D. J. Goodman, "Efficient power control via pricing in wireless data networks," *IEEE Transactions on Communications*, vol. 50, no. 2, pp. 291–303, 2002.
- [16] J. Huang, R. A. Berry, and M. L. Honig, "Distributed interference compensation for wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 5, pp. 1074–1084, 2006.
- [17] R. Etkin, A. Parekh, and D. Tse, "Spectrum sharing for unlicensed bands," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 3, pp. 517–528, 2007.
- [18] S. G. Kiani, D. Gesbert, A. GjendemsjØ, and G. Øien, "Distributed power allocation for interfering wireless links based on channel information partitioning," *IEEE Transactions on Wireless Communications*, vol. 8, no. 6, pp. 3004–3015, 2009.
- [19] M. J. Osborne, *An Introduction to Game Theory*, Oxford University Press, Oxford, UK, 2004.
- [20] F. Ohrtman and K. Roeder, *Wi-Fi Handbook: Building 802.11b Wireless Networks*, McGraw-Hill, New York, NY, USA, 2003.
- [21] N. Jindal, J. Andrews, and S. Weber, "Bandwidth partitioning in decentralized wireless networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 5408–5419, 2008.
- [22] M. Ebrahimi, M. A. Maddah-Ali, and A. K. Khandani, "Throughput scaling laws for wireless networks with fading channels," *IEEE Transactions on Information Theory*, vol. 53, no. 11, pp. 4250–4254, 2007.
- [23] R. Gowaikar, B. Hochwald, and B. Hassibi, "Communication over a wireless network with random connections," *IEEE*

- Transactions on Information Theory*, vol. 52, no. 7, pp. 2857–2871, 2006.
- [24] R. Gowaikar and B. Hassibi, “Achievable throughput in two-scale wireless networks,” *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1169–1179, 2009.
- [25] J. Abouei, H. Bagheri, and A. K. Khandani, “An efficient adaptive distributed space-time coding scheme for cooperative relaying,” *IEEE Transactions on Wireless Communications*, vol. 8, no. 10, pp. 4957–4962, 2009.
- [26] D. E. Knuth, “Big omicron and big omega and big theta,” *ACM SIGACT News*, vol. 8, pp. 18–24, 1967.
- [27] I. S. Gradshteyn, I. M. Ryzhik, and A. Jeffrey, *Table of Integrals, Series, and Products*, Academic Press, 1994.
- [28] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, 2nd edition, 1999.
- [29] S. M. Ross, *Introduction to Probability Models*, Academic Press, Amsterdam, The Netherlands, 8th edition, 2003.
- [30] V. V. Petrov, *Limit Theorems of Probability Theory: Sequences of Independent Random Variables*, Oxford University Press, Oxford, UK, 1995.
- [31] J. Abouei, A. Bayesteh, M. Ebrahimi, and A. K. Khandani, “Sum-rate maximization in single-hop wireless networks with the on-off power scheme,” in *Proceedings of the IEEE International Symposium on Information Theory (ISIT’07)*, pp. 2761–2765, Nice, France, June 2007.
- [32] J. Abouei, M. Ebrahimi, and A. K. Khandani, “A new decentralized power allocation strategy in single-hop wireless networks,” in *Proceedings of the 41st Annual Conference on Information Sciences and Systems (CISS ’07)*, pp. 288–293, Johns Hopkins University, Baltimore, Md, USA, March 2007.