Maximizing the Sum Rate in Symmetric Networks of Interfering Links

Sibi Raj Bhaskaran, Stephen V. Hanly, *Member, IEEE*, Nasreen Badruddin, *Member, IEEE*, and Jamie S. Evans, *Member, IEEE*

***Abstract*—We consider the power optimization problem of maximizing the sum rate of a symmetric network of interfering links in Gaussian noise. All transmitters have an average transmit power constraint, the same for all transmitters. This problem has application to DSL, as well as wireless networks. We solve this nonconvex problem by indentifying some underlying convex structure. In particular, we characterize the maximum sum rate of the network, and show that there are essentially two possible states at the optimal solution depending on the cross-gain ($\sqrt{\epsilon}$) between the links, and/or the average power constraint: the first is a wideband (WB) state, in which all links interfere with each other, and the second is a frequency division multiplexing (FDM) state, in which all links operate in orthogonal frequency bands. The FDM state is optimal if the cross-gain between the links is above $1/\sqrt{2}$. If $\sqrt{\epsilon} < \frac{1}{\sqrt{2}}$, then FDM is still optimal provided the SNR of the links is sufficiently high. With $\sqrt{\epsilon} < \frac{1}{\sqrt{2}}$, the WB state occurs when the SNR is low, but as we increase the SNR from low to high, there is a smooth transition from the WB state to the FDM state: For intermediate SNR values, the optimal configuration is a mixture, with some fraction of the bandwidth in the WB state, and the other fraction in the FDM state. We also consider an alternative formulation in which the power is mandated to be frequency flat. In this formulation, the optimal configuration is either all links at full power, or just one link at full power. In this setting, there is an abrupt phase transition between these two states.**

***Index Terms*—Power control, resource allocation, spectrum allocation, sum rate maximization, interference mitigatio.**

I. Introduction

Wireless networks are plagued by two key problems not encountered in wireline networks: multipath fading and interference between links. In this paper, we focus primarily on the management of the second problem using optimized power allocation. The problem of interference also arises in a DSL wireline access network, and our results are applicable to this system as well, perhaps even more so given that we only treat the time-invariant setting in the present paper. We pose a power allocation problem in which the objective is to maximize the total rate achieved in the network. Each link has to choose a transmit power spectrum, but the choice impacts not only the rate achievable on the desired link, but also the rates achievable on the remaining links of the network.

Unlike traditional power control formulations, in which rate targets are constraints of the problem [1], the rate maximization formulation that we consider in this paper provides a more challenging nonlinear, nonconvex optimization problem. In this paper, we focus on networks in which the channel transfer functions are time-invariant and frequency flat; otherwise, the problem is infinite dimensional and computationally intractable [2].

Recently, progress has been made on time-invariant networks characterized by frequency flat channel responses (i.e., those that can be represented by one-parameter channel *gains*), but under the assumption that the power allocation itself is time invariant and frequency flat, with maximum power constraints on the links [3]–[5]. We will also address this problem in Section IV, but first we consider the less constrained version of the problem, in which there are average power constraints on each link, but no peak power constraints, and the frequency response is not mandated to be frequency flat. This problem is known as the "spectrum management problem" or "spectrum balancing problem" in the literature [6], [7].[1]

Although the spectrum balancing problem is not itself convex, we exhibit an underlying convex structure that arises in symmetric networks, and this structure helps us identify the optimal solution. We show that the optimal power spectrum always consists of a relatively small number of modes, where a mode is a chunk of spectrum in which the power spectral density of all links is constant. Thus, the optimal total power spectrum is piecewise constant [8]. In this paper, we characterize the optimal solution precisely for the case of symmetric interfering links: We provide the bandwidths of the modes, and the power allocation for each link in each mode.

The general characteristic of the optimal solution is that it involves at most two states: a frequency division multiplexing (FDM) state, or a wideband (WB) state in which per-link power allocations are flat across the frequency band. In some scenarios, depending on the cross-gain factor and the signal to noise ratio, the optimal configuration is a mixture of these two states.

In Section IV, we impose the constraint that the transmit power spectrum of each link be flat across the frequency band.

Manuscript received July 17, 2009; revised April 20, 2010. Date of current version August 18, 2010. This work was supported in part by ARC Grant DP0984862 and in part by NUS Grant R-263-000-572-133.

S. R. Bhaskaran is with the Department of Electrical Engineering, IIT Bombay, Powai Mumbai 400 076, India (e-mail: bsraj@ee.iitb.ac.in).

S. V. Hanly is with the Department of Electrical and Computer Engineering, NUS, Singapore, 117576 (e-mail: elehsv@nus.edu.sg).

N. Badruddin and J. S. Evans are with the Department of Electrical and Electronic Eng., University of Melbourne, Parkville, Vic. 3010 Australia (e-mail: nasreenb@ee.unimelb.edu.au; jse@ee.unimelb.edu.au).

Communicated by R. A. Berry, Associate Editor for Communication Networks.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2010.2054652

[1]The spectrum balancing problem is usually posed for frequency selective channels, not for the frequency flat channels considered in the present paper, but our approach can still be interpreted as "spectrum balancing" for this specialized problem.

This is in line with assumptions made in several other works in the literature [3]–[5]. Again, exploiting the underlying convex structure of this problem, we provide the complete solution for a symmetric network of N interfering links. Moreover, we identify an interesting phase transition phenomenon for large networks that occurs as the interference cross-gain parameter ($\sqrt{\epsilon}$ in the formulation below) crosses a certain threshold, and we explicitly characterize this threshold.

II. Related Work

The problems investigated in this paper are special cases of "spectrum management" [9], otherwise known as "spectrum balancing" [10]. There has been intensive interest in these problems for about a decade, mostly in the context of digital subscriber lines (DSL). More recently, there has been interest in applications to wireless communications, including cognitive radio [8], wireless mesh networks [11], and similar problems were investigated much earlier for cellular systems [12].

In applications to DSL, the focus is a multi-user communication system, in which each link provides interference to the others. Typically, each link is time-invariant (appropriate for DSL applications), and there is a set of time-invariant transfer functions from each transmitter to each receiver in the network. There are N desired links (transmitter-receiver pairs) but for each receiver, there are $N-1$ undesired cross-links from the other transmitters, each providing interference to the desired receiver. Each transmitter has a power constraint, and all links share the same frequency band. The spectrum management problem is to shape the spectrum of each link to meet the power constraints, but also to maximize some objective function, such as the sum of the rates on each link [2]. Various other objective functions have also been considered in the literature [2].

In the time-invariant setting, one can formulate a general sum-rate maximization problem subject to power constraints

$$\max_{\boldsymbol{P}} \sum_{i=1}^{N} \int_f \log\left(1+\frac{h_{i,i}(f)P_i(f)}{\sigma^2(f)+\sum_j h_{j,i}(f)P_j(f)1_{\{j\neq i\}}}\right) \quad \text{s.t. } 0 \le P_i(f), \int_f P_i(f) \le \bar{P}_i \tag{1}$$

where $\bar{P}_i$ is the average power constraint for the ith link, and $\sigma^2(f)$ is the noise power spectral density at frequency f. Unfortunately, this problem is infinite dimensional, and NP-hard as we increase the number of links [2].

A key reason for the difficulty in solving (1) is the nonconvexity of the objective function. A common approach is to restrict attention to a finite number of subcarriers, as in Orthogonal Frequency Division Multiplexing (OFDM). This make the problem finite dimensional, but does not overcome the computational difficulties: The general problem is NP hard, as we increase the number of subcarriers, for a fixed network of links [2]. It is also NP hard as we increase the number of links, for a fixed number of subcarriers [2]. One is, therefore, motivated to seek suboptimal approaches to the general problem, although one can seek exact solutions in small problem instances [13], [14].

Suboptimal approaches for the general problem include game-theoretic methods, including iterative water filling [15]; high SNR approximations, and the use of Geometric programming methods [16], [17]; methods of successive convex approximation [18]; and dual decomposition/column generation methods [11].

Another approach is to impose further structure on the problem to reduce the inherent complexity. It is shown in [8] that even in the continuous spectrum balancing formulation, the problem is inherently finite dimensional provided that the transfer functions are finite dimensional. For example, in the case of flat fading channels, the dimensionality is no more than $N+2$, where N is the number of links. In the present paper, we restrict the class of problems infinitely more, to the very special case in which all the desired channel gains are unity, and all the cross gains are the same fixed value, $\sqrt{\epsilon}$. Although this symmetric model is far too specialized to be directly relevant to applications, we can completely characterize the optimal solution, for any N, and the solution is elegant and insightful. We hope that techniques developed for the solution will prove more generally useful, perhaps in the development of suboptimal methods for more general classes of problems, or exact methods for more specialized scenarios.

A paper closely related to our own is [13]. Our paper considers an arbitrary number of links, but the channel gains are completely symmetric. In [13], there are only two links, but the channel gains are arbitrary. The symmetric version of the two link problem is also solved in [14].

The discrete version of the sum-rate maximization problem, with a fixed number of subcarriers, has been studied by many authors. The special case of one subcarrier leads to consideration of the following optimization problem:

$$\max_{\boldsymbol{P}} \sum_{i=1}^{N} \frac{1}{2}\log\left(1+\frac{h_{i,i}P_i}{\sigma^2+\sum_j h_{j,i}P_j 1_{\{j\neq i\}}}\right) \quad \text{s.t. } 0 \le P_i \le P_{\max} \tag{2}$$

where N is the number of links, $P_{\max}$ the power constraint, P_i is the power allocated to link i, and $h_{j,i}$ is the channel gain from link j to link i. This is problem P_1' in [2], although our notation is different. We solve the symmetric version of this problem in Section IV, but the general problem, as stated here, is NP-hard [2]. The two link version of this general problem was solved in [3].

The optimization problem (2) is motivated by the fact that the single link capacity of a discrete time Gaussian noise channel is $\frac{1}{2}\log(1+P)$ nats/symbol, when the SNR is P. The sum-rate in (2) can be shown to be achievable using the Gaussian input distribution in a classical random coding argument. Each transmitter selects *i.i.d.* symbols, with transmitter i using the distribution $N(0,P_i)$ for its codeword symbols. This approach is further supported by the mismatched decoding results of [19], which show that these rates are achievable using minimum distance decoding, even if the codebooks of the interferers are not selected randomly as above, which shows that the Gaussian interference model is robust.

The sum-rate in (2) is by no means optimal. One simple extension that can increase capacity is to allow each transmitter to modulate the variance of the Gaussian input over different transmitted symbols, whilst satisfying the long-term average power constraint. This approach is taken in [11], [12]. For linear, time-invariant, continuous-time channels, one can equivalently vary the power spectral density across the frequency band. By allowing the input spectrum to vary across frequency ("spectrum balancing"), rather than restrict it to be frequency flat (white), we can get an improvement in network capacity, even in the case of time-invariant, frequency flat channels [8]. In the present paper, we treat the spectrum balancing problem in Section III, and relegate the flat input spectrum problem to Section IV.

The sum-rate can be increased further, in some cases, by selecting a different, non-Gaussian, random coding distribution. In general, the optimal distribution for a given problem instance is unknown, although it is known to be Gaussian when the interference level is sufficiently low [20]. In the present paper, we will always use the Gaussian distribution to obtain our achievable rates.

The literature we have considered thus far makes the assumption that each decoder treats the signals from the other links as Gaussian noise. As remarked above, this is a robust assumption, and it is reasonable to make this assumption when the decoders do not have access to the codebooks of the interfering links. However, if the codebooks are known, then the decoder may be able to decode and cancel some of the interference, which can be much better than treating it as Gaussian noise, especially when the interference to signal ratio is not so low [21]. This motivates the more fundamental question: what are the absolute limits of communication in networks of interfering links, in the sense of Shannon? It turns out that the capacity region of even the most simple symmetric two link network (of the class investigated in this paper) is unknown, although there has been recent progress in obtaining inner and outer bounds to the capacity region. In particular, the inner and outer bounds are now quite close, with a gap of no more than 1 bit/sec/Hz, irrespective of SNR, in the two link, Gaussian interference channel [21]. One interesting recent result is that it can be optimal to treat the interference as Gaussian noise, when the interference to noise ratio is sufficiently low [20].

In spite of the recent progress on the fundamental information theory, the present paper is focused on the interesting and important problem of spectrum balancing in Gaussian interference networks, in which interference is treated as Gaussian noise.

III. Spectrum Balancing Problem

In this paper, we will adopt the convention that "power spectral density" refers to the one-sided version, and thus we will focus attention on the positive frequencies. In the following, we denote the Shannon capacity of a discrete time, additive white Gaussian noise (AWGN) channel, with SNR P, by $\frac{1}{2}C(P)$, where

$$C(P) = \log(1+P) \quad \text{nats/channel use.}$$

We begin with a spectrum balancing problem for a symmetric network of interfering links, in which the transmitters are band-limited, with symmetric average power constraints: Consider a real, base-band model of N communication links, each of bandwidth W Hz, and each link is individually an additive white Gaussian noise channel (AWGN) with common noise power

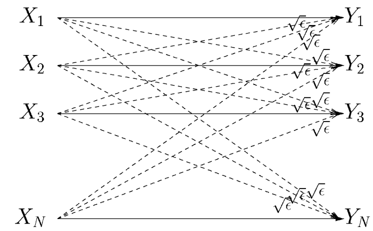

Fig. 1. Symmetric network of interfering links, where $Y_i(t) = X_i(t) + \sum_{j\neq i}\sqrt{\epsilon}X_j(t) + Z_i(t)$.

spectral density of σ^2 at the receivers. By re-scaling powers, we can assume without loss of generality that $\sigma^2 = 1/W$ Watts/Hz. The N links interfere with each other, as depicted in Fig. 1, and we assume an additive model for the interference between links.

We first write down some well known achievable rates for the network. One can compute mutual informations between the transmitter and receiver of each link, assuming that each transmitter uses a stationary Gaussian process to generate the signal. Classical random coding arguments then show that these mutual information rates are achievable. Denote the real, stationary Gaussian process transmitted on link i by $X_i(t)$. The received signal on link i is $Y_i(t)$, where

$$Y_i(t) = X_i(t) + \sum_{j\neq i}\sqrt{\epsilon}X_j(t) + Z_i(t) \tag{3}$$

$Z_i(t)$ is white Gaussian noise of power spectral density $1/W$, and $\sqrt{\epsilon}$ is the cross-gain between the links of the network. If process $X_n(t)$ has power spectral density $\mathcal{P}_n(f)$ then an achievable rate on link i is given by [22]

$$R_i = \int_0^W C\left(\frac{W\mathcal{P}_i(f)}{1+W\epsilon\sum_{j\neq i}\mathcal{P}_j(f)}\right)df.$$

We impose the power constraint that for all i

$$\int_0^W \mathcal{P}_i(f) \le P_{\text{ave}}.$$

The problem we address is that of computing the maximum achievable sum capacity of this network, under the above assumptions, which reduces to finding the optimal input spectra for the links, as expressed in the following optimization problem:

Problem 3.1: Find the input spectra that achieve the maximum in the following program:

$$\max \quad \sum_{i=1}^N \int_0^W C\left(\frac{W\mathcal{P}_i(f)}{1+W\epsilon\sum_{j\neq i}\mathcal{P}_j(f)}\right)df \tag{4}$$

$$\text{s.t.} \quad \int_0^W \mathcal{P}_i(f)df \le P_{\text{ave}}. \tag{5}$$

Clearly this problem is a highly specialized one, with frequency flat channel gains (i.e., no fading) and symmetrical links. Therefore, one should expect a relatively simple solution, perhaps a wide-band (WB) solution with frequency flat power allocations across the links. However, if the cross-gain parameter, $\sqrt{\epsilon}$, is large, it might be better to take a frequency division multiplexing (FDM) approach, to avoid the interference between the

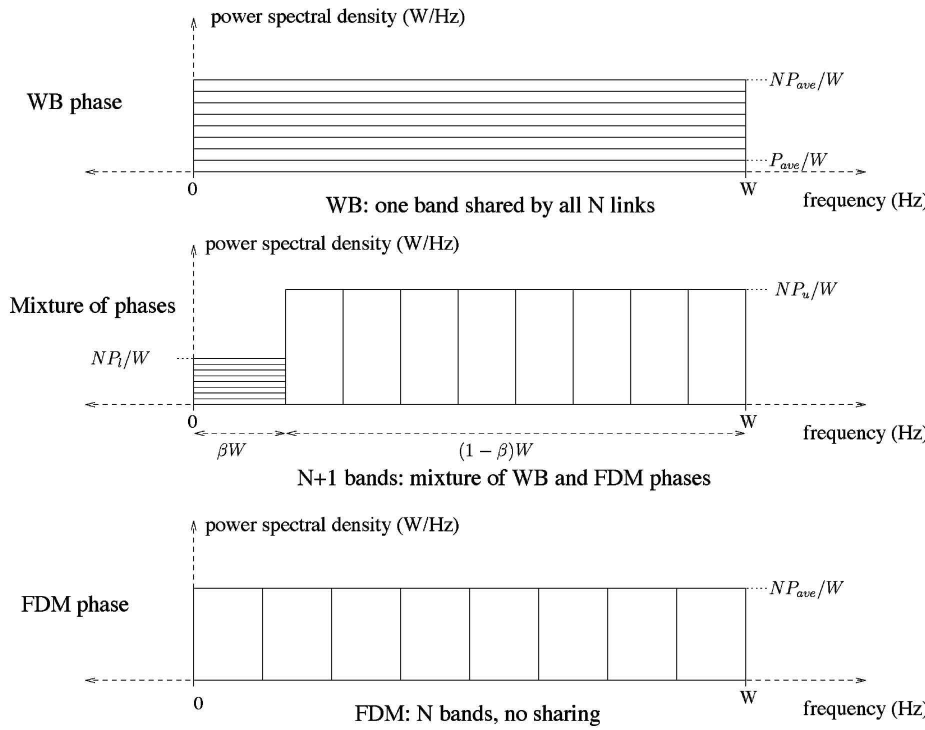

Fig. 2. Two phases and their mixture for the symmetric network of N links.

links. Due to the symmetry of the model, both approaches allocate equal power to all links, and in the FDM approach, each link is allocated a sub-band of width W/N Hz, with a flat power allocation within the sub-band. These two approaches are illustrated in Fig. 2, where one is called the "WB phase", and the other is called the "FDM phase". A natural question to ask is: which phase is the best at a given SNR and level of the cross-gain parameter, $\sqrt{\epsilon}$?. A deeper question is to examine the optimality or otherwise of these two approaches. We will examine both questions in this section, but we start with the first question about the relative performance of WB versus FDM.

A. WB Versus FDM

Comparing the performance of WB and FDM is very natural as these are the two standard approaches used in modern wireless systems to handle multiple access interference. FDM is the classical approach used in traditional radio applications, as well as in narrowband cell-phone radio networks with frequency re-use partitioning between cells. WB is the approach taken in Qualcomm's code division multiple access (CDMA) networks, and is presently used in wideband CDMA 3G networks.

Which approach is better in the symmetric network of interfering links? To answer this question, define the functions f_1 and f_2 by

$$f_1(P) = NC\left(\frac{P}{1+\epsilon(N-1)P}\right) \tag{6}$$

$$f_2(P) = C(NP) \tag{7}$$

Note that $\frac{1}{2}f_2(P)$ is the capacity of a discrete-time AWGN channel at SNR = NP, and $\frac{1}{2}f_1(P)$ is N times the capacity of a discrete time Gaussian link that receives interference from $N-1$ other links, as in the symmetric Gaussian network model. Both are in units of nats per channel use. It follows that $Wf_1(P_{\text{ave}})$ is the achievable sum rate (treating interference as noise) of all links in the WB model, and $Wf_2(P_{\text{ave}})$ is the achievable sum-rate of all links in the FDM model.

Let C_i denote the curve defined by the function $f_i(P)$, $i = 1, 2$. The following two lemmas characterize the relative performance of WB versus FDM, across different values of cross-gain $\sqrt{\epsilon}$ and P.

Lemma 3.1: If $\epsilon > 1/2$ then $f_2(P) > f_1(P)$ for all $P > 0$, i.e., the curve C_2 lies entirely above curve C_1. However, if $\epsilon < 1/2$, then there exists a unique $\tilde{P} > 0$ such that

$$f_1(\tilde{P}) = f_2(\tilde{P}) \tag{8}$$

$$f_2(P) < f_1(P) \quad \forall P < \tilde{P} \quad \text{i.e. WB beats FDM} \tag{9}$$

$$f_1(P) < f_2(P) \quad \forall P > \tilde{P} \quad \text{i.e., FDM beats WB} \tag{10}$$

i.e., C_1 is above C_2 for $P < \tilde{P}$, below C_2 for $P > \tilde{P}$, and $\tilde{P}$ is the point where they cross.

Proof: See Appendix A.

Lemma 3.1 answers the first question concerning the relative merits of WB versus FDM, and is illustrated in Fig. 3 for the case $\epsilon < 1/2$. However, Fig. 3 also illustrates the fact that neither scheme is necessarily optimal: for P in the interval (P_l, P_u), both schemes are beaten by a mixture of the two. This observation holds in general, as shown by the following lemma:

Lemma 3.2: If $\epsilon < 1/2$ then there is a unique tangent curve that touches both C_1 and C_2 at two points, namely $(P_l, f_1(P_l))$ and $(P_u, f_2(P_u))$, with $P_l < \tilde{P} < P_u$ (see Fig. 3). Since both f_1 and f_2 are strictly concave, it follows that for all $P > 0$

$$\max\{f_1(P), f_2(P)\} \le f_1(P_l) + f_1'(P_l)(P - P_l) \tag{11}$$

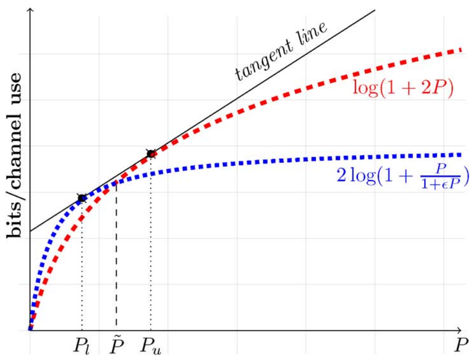

Fig. 3. FDM versus WB when two curves cross, two user case.

with strict inequality for $P \neq P_l, P_u$, and, since the tangent touches C_2 at $(P_u, f_2(P_u))$, we have that

$$f_2(P_u) = f_1(P_l) + f_1'(P_l)(P_u - P_l). \tag{12}$$

Thus, both curves lie below this unique tangent line. For $P^* < P_l$, there is a unique supporting tangent to the curve C_1 at the point $(P^*, f_1(P^*))$, and both C_1 and C_2 lie below this line, i.e., for all $P > 0$

$$\max\{f_1(P), f_2(P)\} \le f_1(P^*) + f_1'(P^*)(P - P^*). \tag{13}$$

Similarly, for $P^* > P_u$, there is a unique supporting tangent to the curve C_2 at the point $(P^*, f_2(P^*))$, both C_1 and C_2 lie below this line, i.e., for all $P > 0$,

$$\max\{f_1(P), f_2(P)\} \le f_2(P^*) + f_2'(P^*)(P - P^*) \tag{14}$$

Proof: See Appendix A.

We conclude that the situation depicted in Fig. 3 holds for all $\epsilon < 1/2$. In other words, when $\epsilon < 1/2$, neither pure phase (FDM or WB) can be optimal when P lies in the interval (P_l, P_u), as one can do better by time sharing between the phases, or by doing the sharing in the frequency domain as illustrated in the mixture plot in Figure 2.

This analysis leads to the more fundamental question about optimality for the symmetric network of interfering links. We have so far highlighted three distinct schemes to handle interference: pure FDM, pure WB, and a hybrid scheme that is a mixture of the two. The following subsection will demonstrate that the optimal scheme in the symmetric model is always one of these three candidates, but the particular one depends on the cross-gain parameter, $\sqrt{\epsilon}$, and/or the SNR.

B. Optimal Scheme

The main result of this section is Theorem 3.1, which supplies the solution to Problem 3.1. In the theorem, the term $C(\epsilon, N, P_{\text{ave}})$ is defined by

$$C(\epsilon, N, P_{\text{ave}}) = \begin{cases} f_2(P_{\text{ave}}) & \epsilon > 1/2 \text{ or } P_{\text{ave}} > P_u \\ f_1(P_{\text{ave}}) & \epsilon < 1/2 \text{ and } P_{\text{ave}} < P_l \\ \beta f_1(P_l) + \\ (1-\beta) f_2(P_u) & \text{o.w.} \end{cases} \tag{15}$$

where, in the last case, i.e., $\epsilon < 1/2$ and $P_l < P_{\text{ave}} < P_u$, we define β to be the unique number in $(0, 1)$ such that $P_{\text{ave}} = \beta P_l + (1-\beta) P_u$.

Theorem 3.1: The optimal value in Problem 3.1 is $WC(\epsilon, N, P_{\text{ave}})$ bits/sec. If $\epsilon > 1/2$ or $P_{\text{ave}} > P_u$, the optimal value is achievable by dividing the spectrum into N equal sized bands, and allocating each link one of the bands (FDM). Each link uses power spectral density $(NP_{\text{ave}})/W$ in its allocated band. If $\epsilon < 1/2$ and $P_{\text{ave}} < P_l$ then the optimal value is achievable via the WB approach: all the links share the entire band, and each link uses power spectral density P_{ave}/W across the wide band. Otherwise, $\epsilon < 1/2$ and $P_{\text{ave}} = \beta P_l + (1-\beta)P_u$ for some unique $\beta \in (0, 1)$. In this case, rate $WC(\epsilon, N, P_{\text{ave}})$ is achievable with $N+1$ bands allocated as follows. One band, of width βW Hz is shared amongst the N links, and each link uses power spectral density P_l/W in this band. The sumrate achieved in this band is $W\beta f_1(P_l)$ bits/sec. The remaining bandwidth, $(1-\beta)W$ Hz is split evenly into N sub-bands, and each sub-band is allocated to one of the links. In its own sub-band, a link uses power spectral density $(NP_u)/W$, and obtains a rate of $W(1-\beta)f_2(P_u)/N$ bits/sec in this sub-band.

In summary, there are essentially two distinct states for the system, FDM or WB. Which one is optimal depends on ϵ and P_{ave}, and, in an intermediate scenario, the optimal configuration is a mixture of the two.

C. Proof of Theorem 3.1

The achievability of $C(\epsilon, N, P_{\text{ave}})$ can be immediately verified. The issue we address in this section is the converse, namely, that there is no other strategy that can beat $C(\epsilon, N, P_{\text{ave}})$.

It can be shown that the optimum can be achieved with spectra that are piece-wise constant: There are at most $N+2$ disjoint intervals in $[0, W]$ with each link having constant power spectral density within each interval [8]. This can be proven via Caratheodory's convexity theorem [23], and is a consequence of the dimensionality of the problem. Here, the dimension of the problem is $N+1$, since there are N links, each of which has to choose a power level, and the sum-rate provides an additional dimension. Thus, the following problem is equivalent to Problem 3.1

Problem 3.2: Let $M = N+2$. Find the normalized bandwidths $(\alpha_0, \alpha_1, \ldots, \alpha_{M-1})$ and power levels $(P_i^{(m)})$ $i = 1, 2, \ldots, N, m = 0, 1, \ldots, M-1$ to solve

$$\max \sum_{i=1}^{N} \sum_{m=0}^{M-1} \alpha_m C\left(\frac{P_i^{(m)}}{1 + \epsilon \sum_{j \neq i} P_j^{(m)}}\right) \tag{16}$$

$$\text{s.t.} \quad \sum_{m=0}^{M-1} \alpha_m P_i^{(m)} \le P_{\text{ave}}, \; P_i^{(m)} \ge 0 \tag{17}$$

$$\sum_{m=0}^{M-1} \alpha_m \le 1, \; 0 \le \alpha_m \le 1. \tag{18}$$

Before attempting to solve this problem, we remark that at first sight it does not appear to be a very simple problem to solve, as it is not a convex problem. For this reason, we begin by formulating a simpler optimization problem that we can directly solve. Suppose each link must choose a fixed, static power level, but the power constraint is on the *sum* of the powers of all the links, rather than on the individual powers of each link. Thus, link i uses a power level P_i, and the sum power constraint is that $\sum_{i=1}^{N} P_i = \hat{P}$. The following function provides the sum-rate for this problem:

$$C_{sum}(\epsilon, N, \boldsymbol{P}) = \sum_{i=1}^{N} C\left(\frac{P_i}{1+\epsilon\sum_{j\neq i} P_j}\right). \tag{19}$$

The following lemma states the optimization problem precisely, and provides its solution:

Lemma 3.3: Consider the optimization problem

$$\max_{\boldsymbol{P}} C_{sum}(\epsilon, N, \boldsymbol{P}) \text{ s.t. } P_i \geq 0 \ \forall i, \quad \sum_{i=1}^{N} P_i = \hat{P} \tag{20}$$

and let $U(\epsilon, N, \hat{P})$ denote the optimal value. Then

$$\begin{aligned} &U(\epsilon, N, \hat{P}) \\ &= \max\left\{NC\left(\frac{\hat{P}/N}{1+\epsilon(N-1)\hat{P}/N}\right), C(\hat{P})\right\} \\ &= \max\{f_1(\hat{P}/N), f_2(\hat{P}/N)\}. \end{aligned} \tag{21}$$

Proof: See Appendix C.

The solution to Problem 3.2 can now be found, using the following two lemmas:

Lemma 3.3: Let $\boldsymbol{\alpha}, \boldsymbol{P}$ be a feasible allocation of normalized bandwidths, and power levels, respectively, for Problem 3.2, and let $C(\epsilon, N, \boldsymbol{\alpha}, \boldsymbol{P})$ denote the corresponding sum-rate. Then $C(\epsilon, N, \boldsymbol{\alpha}, \boldsymbol{P})$ is upper-bounded by the optimal value in the following program:

Problem 3.3:

$$\max_{\alpha, P^{(a)}, P^{(b)}} \alpha f_1\left(P^{(a)}\right) + (1-\alpha) f_2\left(P^{(b)}\right) \tag{22}$$

$$\begin{aligned} \text{s.t.} \quad & 0 \leq \alpha \leq 1, \ P^{(a)} \geq 0, \ P^{(b)} \geq 0 \\ & \alpha P^{(a)} + (1-\alpha) P^{(b)} \leq P_{\text{ave}}. \end{aligned} \tag{23}$$

Proof: Let $\boldsymbol{\alpha}, \boldsymbol{P}$ be a feasible allocation of normalized bandwidths, and power levels, respectively, and let $C(\epsilon, N, \boldsymbol{\alpha}, \boldsymbol{P})$ denote the corresponding sum-rate

$$C(\epsilon, N, \boldsymbol{\alpha}, \boldsymbol{P}) = \sum_{i=1}^{N}\sum_{m=0}^{M-1} \alpha_m C\left(\frac{P_i^{(m)}}{1+\epsilon(N-1)\sum_{j\neq i} P_j^{(m)}}\right).$$

Note that $\boldsymbol{\alpha}$ is of dimension $M = N+2$, and $\boldsymbol{P} = (P_i^{(m)})$ is of dimension NM. Let $\hat{P}^{(m)} = \sum_{i=1}^{N} P_i^{(m)}$. Then the upper bound

$$C(\epsilon, N, \boldsymbol{\alpha}, \boldsymbol{P}) \leq \sum_{m=0}^{M-1} \alpha_m U\left(\epsilon, N, \hat{P}^{(m)}\right) \tag{24}$$

clearly holds, where $U(\epsilon, N, \hat{P}^{(m)})$ is the optimal value in problem (20). By Lemma 3.3, we can write the RHS of (24) as

$$\sum_{m=0}^{M-1} \alpha_m \max\left\{f_1\left(\frac{\hat{P}^{(m)}}{N}\right), f_2\left(\frac{\hat{P}^{(m)}}{N}\right)\right\}. \tag{25}$$

Now re-order the modes so that for modes $m = 0, 1, \ldots, k$, the maximum in (25) is $f_1(\hat{P}^{(m)}/N)$ (if there are no such modes, let $k = -1$) and for modes $k+1, k+2, \ldots, M-1$, the maximum in (25) is $f_2(\hat{P}^{(m)}/N)$ (if there are no such modes, k will equal $M-1$). But the functions f_1 and f_2 are both concave functions, so if we define

$$\alpha = \sum_{m=0}^{k} \alpha_m \tag{26}$$

$$P^{(a)} = \sum_{m=0}^{k} (\alpha_m/\alpha)(\hat{P}^{(m)}/N) \tag{27}$$

$$P^{(b)} = \sum_{m=k+1}^{M-1} (\alpha_m/(1-\alpha))(\hat{P}^{(m)}/N) \tag{28}$$

then

$$C(\epsilon, N, \boldsymbol{\alpha}, \boldsymbol{P}) \leq \alpha f_1(P^{(a)}) + (1-\alpha) f_2(P^{(b)}). \tag{29}$$

Since the initial mode and power allocations $(\boldsymbol{P}, \boldsymbol{\alpha})$ are feasible for Problem 3.2, it must also be true that

$$\alpha P^{(a)} + (1-\alpha) P^{(b)} \leq P_{\text{ave}}.$$

■

Lemma 3.5: The maximum value achieved in Problem 3.3 is $C(\epsilon, N, P_{\text{ave}})$.

Proof: First, consider the case that $\epsilon > 1/2$ and let $\alpha, P^{(a)}, P^{(b)}$ be feasible for Problem 3.3. Then $f_1(P^{(a)}) < f_2(P^{(a)})$ by Lemma 3.1, so

$$\begin{aligned} &\alpha f_1\left(P^{(a)}\right) + (1-\alpha) f_2\left(P^{(b)}\right) \\ &\quad \leq \alpha f_2(P^{(a)}) + (1-\alpha) f_2\left(P^{(b)}\right) \\ &\quad \leq f_2(P_{\text{ave}}) \end{aligned}$$

with the second inequality following from (23) and the concavity of f_2.

Now consider the case that $\epsilon < 1/2$ and $P_{\text{ave}} < P_l$, and let $\alpha, P^{(a)}, P^{(b)}$ be feasible for Problem 3.3. Then

$$\begin{aligned} &\alpha f_1\left(P^{(a)}\right) + (1-\alpha) f_2\left(P^{(b)}\right) \\ &\quad \leq f_1(P_{\text{ave}}) + \alpha f_1'(P_{\text{ave}})\left(P^{(a)} - P_{\text{ave}}\right) \\ &\qquad + (1-\alpha) f_1'(P_{\text{ave}})\left(P^{(b)} - P_{\text{ave}}\right) \\ &\quad \leq f_1(P_{\text{ave}}) \end{aligned}$$

where the first inequality follows from (13), and the second inequality from (23). But clearly $f_1(P_{\text{ave}})$ is achievable if we set

$\alpha = 1$ and $P^{(a)} = P_{\text{ave}}$. The case $\epsilon < 1/2$ and $P_{\text{ave}} > P_u$ follows in the analogous way, using (14) in place of (13).

The remaining case is $\epsilon < 1/2$ and $P_{\text{ave}} = \beta P_l + (1-\beta)P_u$, for $0 < \beta < 1$, with $\alpha, P^{(a)}, P^{(b)}$ feasible for Problem 3.3. Then

$$\begin{aligned}
&\alpha f_1\left(P^{(a)}\right) + (1-\alpha) f_2\left(P^{(b)}\right) \\
&\quad \le \alpha f_1(P_l) + \alpha f_1'(P_l)(P^{(a)} - P_l) \\
&\quad + (1-\alpha) f_1(P_l) + (1-\alpha) f_1'(P_l)(P^{(b)} - P_l) \\
&\quad \le f_1(P_l) + f_1'(P_l)(P_{\text{ave}} - P_l) \\
&\quad = \beta f_1(P_l) + (1-\beta) f_1(P_l) + (1-\beta) f_1'(P_l)(P_u - P_l) \\
&\quad \le \beta f_1(P_l) + (1-\beta) f_2(P_u)
\end{aligned}$$

where the first inequality follows from (11), the second inequality from (23), and the first equality from (12). But $\beta f_1(P_l) + (1-\beta) f_2(P_u)$ is achievable with $N+1$ bands, one of which is of bandwidth β, shared by all links, and the remaining band is partitioned equally amongst the N links using FDM. ■

We summarize this section, with the following conclusion, which completes the proof of Theorem 3.1: $C(\epsilon, N, P_{\text{ave}})$ is the optimal value in Problem 3.2. If $\epsilon > 1/2$ or $P_{\text{ave}} > P_u$, the optimal value is achievable with N modes, $\alpha_m = 1/N$ for $m = 0, 1, \ldots, N-1$, and $P_i^{(m)} = N P_{\text{ave}} 1_{\{i=m+1\}}, i = 1, 2, \ldots N$. If $\epsilon < 1/2$ and $P_{\text{ave}} < P_l$ then the optimal value is achievable with 1 mode, $\alpha_0 = 1$, and $P_i^{(0)} = P_{\text{ave}}, i = 1, 2, \ldots, N$. Otherwise, $\epsilon < 1/2$ and $P_{\text{ave}} = \beta P_l + (1-\beta) P_u$ for some unique $\beta \in (0,1)$. In this case, $C(\epsilon, N, P_{\text{ave}})$ is achievable with $N+1$ modes: $\alpha_0 = \beta$, $\alpha_m = (1-\beta)/N$, $m = 1, 2, \ldots, N$, $P_i^{(0)} = P_l$, $i = 1, 2, \ldots, N$, $P_i^{(m)} = N P_u 1_{\{m=i\}}$, $i = 1, 2, \ldots, N$, $m = 1, 2, \ldots, N$.

IV. Flat Power Constraints

The spectrum balancing problem allows the links to allocate the total power arbitrarily over the available degrees of freedom (frequencies). Another body of work considers the case for which there is only one degree of freedom (a single subcarrier), or equivalently, the case for which the power allocation is mandated to be time invariant and flat across frequency. More precisely, this approach characterizes the achievable rates under random coding under the restriction that the input distribution (after sampling at the maximum frequency) is a fixed Gaussian distribution, with link i using the static power level P_i. This precludes any time-sharing between different power allocation strategies. The symmetric network version of this problem is the following:

Problem 4.1:

$$\max_{\boldsymbol{P}} \sum_{i=1}^{N} C\left(\frac{P_i}{1 + \epsilon \sum_j P_j 1_{\{j \ne i\}}}\right) \text{ s.t. } 0 \le P_i \le P_{\max} \quad (30)$$

which is also the symmetric version of the single subcarrier sum-rate maximization problem (P_1') in [2]. Let $C_{flat}(\epsilon, N, P_{\max})$ denote the value at the optimal solution to Problem 4.1.

The solution of Problem 4.1 will produce a power spectrum that is flat across the frequency spectrum. For applications in which there is a spectral mask on each transmitter, there may be peak constraints on each frequency, in which case the above formulation is relevant. In this section, we will provide a complete solution to Problem 4.1.

We immediately note that Problem 4.1 is a nontrivial optimization problem. It is a nonlinear, nonconvex optimization problem over N continuous variables, and it is not difficult to see that there are many local maxima, causing problems for many standard numerical approaches [4]. However, the two link version of this problem has recently been solved [3], [4], [24]. It has been shown that the optimal solution is particularly simple, and is characterizable in terms of the cross-gain, $\sqrt{\epsilon}$. There is a critical threshold, ϵ_t, such that for $\epsilon < \epsilon_t$, the optimal solution is $(P_1, P_2) = (P_{\max}, P_{\max})$. For $\epsilon > \epsilon_t$, there are two optimal solutions, $(P_{\max}, 0)$ and $(0, P_{\max})$. In other words, either both links blast at full power, or one link switches off, and the other blasts at full power.

In the present section, we solve the more general problem with N links, where N is arbitrary. One might guess from the two link result that the general solution will be a similar all or nothing scheme, in which links either blast at full power or are completely switched off. This turns out to be true, but there is an interesting phase transition effect, in which there is a macroscopic change as the level of interlink interference (the cross-gain parameter, $\sqrt{\epsilon}$) crosses a threshold. If the cross-gain parameter is below the threshold, all links should operate at full power. If the cross-gain parameter is above the threshold, only one link should operate at full power, the rest being switched off. This is expressed in the following theorem:

Theorem 4.1:

$$C_{flat}(\epsilon, N, P_{\max}) = \begin{cases} N C\left(\frac{P_{\max}}{1 + (N-1)\epsilon P_{\max}}\right) & \epsilon < \epsilon_{N,1} \\ C(P_{\max}) & \epsilon \ge \epsilon_{N,1} \end{cases} \quad (31)$$

where $\epsilon_{N,1}$ is given by

$$\epsilon_{N,1} = \frac{(1 + P_{\max}) - (1 + P_{\max})^{\frac{1}{N}}}{(N-1) P_{\max}((1 + P_{\max})^{\frac{1}{N}} - 1)}. \quad (32)$$

An optimal power allocation is:

- $P_i^{opt} = P_{\max} \forall i$, if $\epsilon \le \epsilon_{N,1}$;
- $P_i^{opt} = P_{\max} 1_{\{i=1\}}$ if $\epsilon > \epsilon_{N,1}$.

Theorem 4.1 is proven in Appendix E, using some results from Section IV-B below. The proof has quite a number of cases to consider, so we have relegated it to an Appendix; in the following subsection, we sketch the main argument.

A. Sketch of the Proof of Theorem 4.1

In the first step we impose an additional "binary power constraint": the allowable power levels on each link are to be chosen from the set $\{0, P_{\max}\}$. It is immediate that the right hand side of (31) is an achievable sum-rate under the binary power constraint, and, in Section IV-B, below, we show that it is the optimal sum-rate under binary power constraints.

We then consider a solution to Problem 4.1, in which the power levels can take on arbitrary continuous values between 0 and $P_{\max}$. Let $\boldsymbol{P}^*$ be such a solution, with the entries of $\boldsymbol{P}^*$

ordered in decreasing order. It is immediate that $P_1^* = P_{\max}$, for otherwise we can increase the value of $C_{flat}(\epsilon, N, P_{\max})$ by scaling all elements of $\boldsymbol{P}^*$ by a common factor greater than unity, contradicting the optimality of $\boldsymbol{P}^*$. Thus, without loss of generality, we can assume that $\boldsymbol{P}^*$ has the first k entries equal to $P_{\max}$, for some integer $1 \le k \le N$. If $k = N$, we have a binary power vector, and nothing further to prove: The optimal solution under binary power constraints is completely characterized in Section IV-B, below. If $k < N$, then let $0 \le P_{k+1}^* = P < P_{\max}$. Lemma A.1 in Appendix B can be applied to this case to obtain the following characterization of $\boldsymbol{P}^*$.

Lemma 4.1: If $k < N$ and $0 \le P_{k+1}^* = P < P_{\max}$ then there exists integer l, $1 \le l \le N - k$ such that $P_i^* = P$ for $i = k+1, k+2, \ldots, k+l$, and $P_j^* = 0$ for $j > k+l$ (if $k+l < N$). If $k = N$ then the optimal solution is $P_i^* = P_{\max}$ for $i = 1, 2, \ldots, N$.

Proof: See Appendix E.

Now define the function:

$$J(\epsilon, P_{\max}, k, l, P) = k\, C\left(\frac{P_{\max}}{1+\epsilon(k-1)P_{\max}+\epsilon l P}\right) + l\, C\left(\frac{P}{1+\epsilon k P_{\max}+\epsilon(l-1)P}\right). \quad (33)$$

From the characterization of an optimal solution in Lemma 4.1, one can see that Theorem 4.1 is proven if one can establish, for any integers $k \ge 1, l \ge 1$, and real-valued P, $0 \le P \le P_{\max}$, that

$$J(\epsilon, P_{\max}, k, l, P) \le \begin{cases} NC\left(\frac{P_{\max}}{1+(N-1)\epsilon P_{\max}}\right) & \epsilon < \epsilon_{N,1} \\ C(P_{\max}) & \epsilon \ge \epsilon_{N,1} \end{cases}. \quad (34)$$

It will follow from Lemma 4.5 in Section IV-B that a *sufficient* condition for (34) to hold is

$$J(\epsilon, P_{\max}, k, l, \cdot) \text{ has no local maximum in } (0, P_{\max}). \quad (35)$$

This condition is sufficient because it implies binary power control is optimal, and Lemma 4.5 will show that binary power control can do no better than the RHS of (34). In Appendix E, we will examine the extremal behavior of $J(\epsilon, P_{\max}, k, l, \cdot)$ for any choice of $\epsilon, P_{\max}, k, l$ with $k \ge 1$ and $l \ge 1$ (we know that at the optimal solution $k \ge 1$ and if $l = 0$ there is nothing to prove). We will see that $J(\epsilon, P_{\max}, k, l, \cdot)$ is not in general concave, and it *can* have a local maximum in the interval $(0, P_{\max})$: see Fig. 5. However, we will show that if (35) does not hold at a particular choice of $\epsilon, P_{\max}, k, l$ then (34) does hold for this choice, for all $0 \le P \le P_{\max}$. This is expressed in Lemma 4.2 below, which is proven in Appendix E.

Lemma 4.2: For all $\epsilon, P_{\max}, k, l$, either (35) holds, else (34) holds for all $0 \le P \le P_{\max}$.

Proof: See Appendix E.

We summarize this subsection with the conclusion that the first step in the proof of Theorem 4.1 is to show that the right hand side of (31) cannot be beaten by a scheme restricted to binary power levels. This we do in Section IV-B, via Lemmas 4.3–4.5. After that, it is required to prove Lemmas 4.1 and 4.2, and for this we refer the reader to Appendix E.

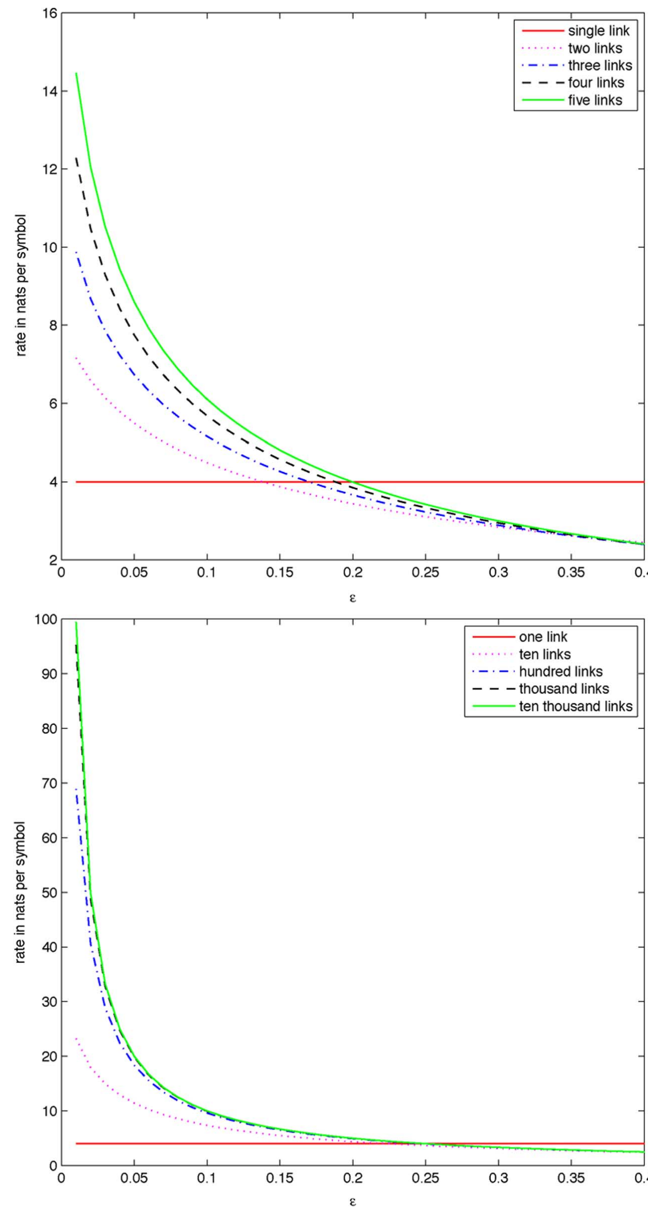

Fig. 4. $R_n(\epsilon)$ as a function of ϵ, for $P_{\max} = \exp(4) - 1$.

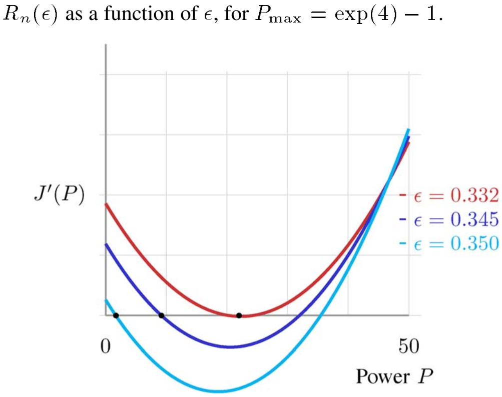

Fig. 5. Existence of local maxima when $k = 2$ users are at $P_{\max} = 50$, and $l = 1$ user is at P.

B. Binary Power Constraints

In this section, we study the maximum sum-rate achievable in (30) when the power levels are further restricted to lie in the set $\{0, P_{\max}\}$. With binary power constraints, power control reduces to deciding if links should be "on" or "off". It is intuitively clear that if the cross-gain parameter, $\sqrt{\epsilon}$, is sufficiently low, then all links should be on. It is also clear that both terms on the right hand side of (31) are achievable with binary power levels, with all links on in the first case, and only one link on in the second case. To show that this is optimal under binary power constraints, we need to consider an arbitrary number of links being on.

For $n = 1, 2, 3, \ldots$, let

$$R_n(\epsilon) = n \log\left(1 + \frac{P_{\max}}{(n-1)\epsilon P_{\max} + 1}\right)$$

and we note that $R_1 = C(P_{\max})$ is independent of ϵ. For $n \geq 2, R_n(0) = nC(P_{\max}) > R_1$ and $\lim_{\epsilon\uparrow\infty} R_n(\epsilon) = 0$. It follows that there is a unique solution in ϵ to the equation $R_n(\epsilon) = C(P_{\max})$, which we denote by $\epsilon_{n,1}$ and it can be explicitly computed

$$\epsilon_{n,1} = \frac{(1+P_{\max}) - (1+P_{\max})^{\frac{1}{n}}}{(n-1)P_{\max}\left((1+P_{\max})^{\frac{1}{n}} - 1\right)} \tag{36}$$

which provides (32).

In this section, we show that the right hand side of (31) cannot be beaten by any binary power allocation scheme. This amounts to showing that if $\epsilon < \epsilon_{N,1}$, then $R_N(\epsilon) > R_n(\epsilon)$ for all $n < N$, but if $\epsilon > \epsilon_{N,1}$, then $R_1(\epsilon) > R_n(\epsilon)$ for all $n > 1$. This is indeed true, as stated in Lemma 4.4 below, and it is illustrated in Fig. 4.

Fig. 4 suggests that we can break up the line into intervals of the form $(\epsilon_{n-1,1}, \epsilon_{n,1})$ and in each such interval $R_n(\epsilon)$ is larger than all the $R_m(\epsilon)$, for $m < n$. Fig. 4 also suggests that the increasing sequence of $\epsilon_{n,1}$ might be approaching a limit, and beyond this limit, the single link rate is dominant. We proceed to verify these statements.

Our first result for binary power control is the following:

Lemma 4.3: $(\epsilon_{n,1})_{n=2}^{\infty}$ is an increasing sequence, with a limiting value of $1/C(P_{\max})$.

Proof: This is a direct corollary of Lemma A.2 in Appendix D.

Note that the value of $1/C(P_{\max})$ for the example depicted in Figure 4 is 0.25.

Lemma 4.3 implies that for any $m \geq 2$ and for any $\epsilon \in (\epsilon_{m-1,1}, \epsilon_{m,1})$, we have that $\epsilon < \frac{1}{C(P_{\max})}$. The following lemma shows that in this case $R_m(\epsilon) > C(P_{\max})$, but also that in the case $\epsilon > \frac{1}{C(P_{\max})}$, the inequality goes the other way. Set $\epsilon_{1,1} = 0$.

Lemma 4.4:
i) $\forall m \geq 2$, and $\forall \epsilon \in (\epsilon_{m-1,1}, \epsilon_{m,1})$

$$C(P_{\max}) < R_m(\epsilon) \tag{37}$$

$$\forall n \geq 1, R_{m+n-1}(\epsilon) < R_{m+n}(\epsilon) \tag{38}$$

$$\lim_{n\uparrow\infty} R_n(\epsilon) = \frac{1}{\epsilon}. \tag{39}$$

ii) For $\epsilon > \frac{1}{C(P_{\max})}$, we have $R_n(\epsilon) < C(P_{\max})$ for all $n \geq 2$.

Proof: A corollary of Lemma A.4 in Appendix D.

Lemma 4.4 implies the following result:

Lemma 4.5:
i) For $\epsilon < \epsilon_{N,1}$, $\max_{1\leq n\leq N} R_n(\epsilon) = R_N(\epsilon)$.
ii) For $\epsilon \geq \epsilon_{N,1}$, $\max_{1\leq n\leq N} R_n(\epsilon) = R_1 = C(P_{\max})$.

Proof: Follows from Lemma 4.4 and the observation that all curves $R_n(\epsilon)$ are strictly decreasing in ϵ, for $n \geq 2$. ■

Lemma 4.5 completes the proof that the right hand side of (31) provides the optimal sumrate under binary power constraints. In summary, for all $n \geq 2$, $R_n(\epsilon)$ decreases with ϵ, and crosses the constant value $C(P_{\max})$ at $\epsilon_{n,1}$. This crossing point increases with n, and on $(0, \epsilon_{N,1})$, R_N dominates all R_n, for $n \leq N$, as depicted in Fig. 4. This figure provides a nice illustration of the sudden switch from having N links on, to having just one link on, as ϵ crosses the critical threshold given in (32).

V. Conclusion

In this paper, we have solved two versions of the sum-rate maximization problem for a symmetric network of an arbitrary number of links. In the first version, each link has an average power constraint, which is the same for all links. We have shown that the critical value of the cross-gain parameter, $\sqrt{\epsilon}$, is $\sqrt{\epsilon} = 1/\sqrt{2}$, above which the optimal spectra consist of N bands, with only one link active in each band, providing a frequency division multiplexing (FDM) characteristic to the solution.

When $\sqrt{\epsilon} < 1/\sqrt{2}$, the FDM configuration is still optimal, provided the SNR is high enough. However, if the SNR is sufficiently low, the optimal power spectra will consist of one band, giving a wideband (WB) characteristic to the solution. For intermediate values of the SNR, the optimal spectra is a mixture of these two states; in these cases there are $N+1$ bands, with N bands of FDM, and one band in which the links interfere with each other.

Although the symmetric network is a very special case of the general problem of interfering links, our solution provides a very clean characterization of the optimal behavior in this particular case, and it may provide insight into more general network problems. The paper [8] shows that the piecewise constant form of the optimal input spectra holds for arbitrary networks with time-invariant, frequency-flat channels. However, the problem of finding the optimal modes and power levels to use in each mode is left completely open. In general, this problem has been shown to be NP-hard [2].

The fact that we need to find the common tangent line to the two curves C_1 and C_2 is a manifestation of the convexity that is a characteristic of all capacity regions. It is well known that capacity regions are always convex: Usually, time-sharing arguments are invoked, but in the present paper, the convexification is obtained in the frequency domain: Our solutions are time-invariant.

An early paper that considered the impact of interference on the capacity of a cellular network is [25]. This paper showed that in some scenarios, pure TDM partitioning of cells into disjoint time-slots, or pure WB strategies, can be beaten by a mixed strategy that they called "fractional intercell time sharing". This is consistent with our findings in the present paper, in which the mixed state is shown to be optimal in some scenarios.

The other problem we have considered in this paper is the rate maximization problem under peak power constraints, in which the power spectrum of the transmitters is mandated to be time-invariant and frequency flat. Each link has a maximum average power constraint, which is the same for all links. We have solved the problem under the constraint that each link must choose a single average power level in the continuous range between 0 and the maximum possible average power level $P_{\max}$. No variation of power in time or frequency is allowed. Under this constraint, we have proven that there is no loss in restricting each link to a binary choice between using zero power, or the full power $P_{\max}$. We have also shown that the optimal choice is either all links on, or just one link on, with a phase transition between these two states as the cross-gain between the links traverses a threshold.

Although the symmetric network is a very special case of the general problem of interfering links, numerical evidence presented in [26] suggests that the binary power control that we have characterized in this paper is likely to be at least close to optimal in many other network scenarios. Furthermore, the techniques developed in the present paper may be useful in analyzing such scenarios. Preliminary work has been undertaken in [27] for a Wyner-type cellular model, in which only neighboring cells interfere with each other.

The assumption of frequency and time-flat power allocation makes this section of the paper directly applicable to CDMA networks. One application is a multiple access channel, with a single receiving node, employing single-user receivers for each user. Our model is directly applicable if all nodes have a maximum average received power constraint, or if they have a maximum average transmit power constraint and they are equidistant from the receiver. The interference parameter, ϵ, is then inversely proportional to the processing gain of the system.

We do not explicitly consider the issue of fairness in the present paper. In Section IV, fairness is inherited from the symmetry of the problem, rather than from an explicit requirement to have fairness. In Section III, some links may need to be switched off, which may be considered unfair. However, fairness can be addressed by higher-layer scheduling algorithms. Alternatively, the assumptions that the power allocation must be frequency flat can be relaxed, as in Section III of this paper.

Finally, we note that the assumption that each link treats the other links as sources of Gaussian noise can be relaxed. A link is instead allowed to know about the codebooks used on the other links. One then enters the difficult territory of the interference channel, although important recent progress has been made [21].

APPENDIX

A. WB Versus FDM: Theoretical Results

This Appendix provides the Proofs of Lemmas 3.1 and 3.2, which allow us to compare WB and FDM, for any number of links, N, any choice of the cross-gain parameter, $\sqrt{\epsilon}$, and any SNR, P.

Proof of Lemma 3.1:

Proof: To account for the dependence of f_1 on ϵ, let us redefine f_1 as a function of two variables

$$f_1(\epsilon, P) = N \log\left(1 + \frac{P}{1+\epsilon(N-1)P}\right)$$

As before, we have

$$f_2(P) = \log(1+NP).$$

The unique $\epsilon^\star(P)$ for which $f_1(\epsilon^\star(P), P) = f_2(P)$ can be explicitly computed as

$$\epsilon^\star(P) = \frac{1+P-(1+NP)^{1/N}}{(N-1)P((1+NP)^{1/N}-1)}$$

which is a decreasing function of P. By the monotonicity of $f_1(\cdot, P)$ (for fixed P), we have that

$$f_1(\epsilon, P) > f_1(\epsilon^\star(P), P) = f_2(P), \text{for } \epsilon < \epsilon^\star(P) \quad (40)$$
$$f_1(\epsilon, P) < f_1(\epsilon^\star(P), P) = f_2(P), \text{for } \epsilon > \epsilon^\star(P). \quad (41)$$

Since $\epsilon^\star(\cdot)$ is decreasing, we can define $\epsilon^\star(0)$ and $\epsilon^\star(\infty)$ by

$$\epsilon^\star(0) = \lim_{P\downarrow 0} \epsilon^\star(P) = 1/2 \text{ and } \epsilon^\star(\infty) = \lim_{P\uparrow\infty} \epsilon^\star(P) = 0.$$

It follows that if $\epsilon > 1/2$ then $f_1(\epsilon, P) < f_2(P)$ for all $P > 0$, but if $\epsilon < 1/2$, then there exists a unique $\tilde{P}(\epsilon) > 0$ such that $\epsilon^\star(\tilde{P}(\epsilon)) = \epsilon$. For $P < \tilde{P}(\epsilon)$, $\epsilon^\star(P) > \epsilon$, and so $f_1(\epsilon, P) > f_2(P)$ by (40). For $P > \tilde{P}(\epsilon)$, $\epsilon^\star(P) < \epsilon$, and so $f_1(\epsilon, P) < f_2(P)$ by (41). ■

Proof of Lemma 3.2:

Proof: If the tangent to C_1 at the point $(P_1, f_1(P_1))$ is to intersect C_2 at $(P, f_2(P))$ then P must solve the equation

$$h(P) = J(P_1) \quad (42)$$

where

$$h(P) := f_2(P) - f_1'(P_1)P$$
$$J(P) := f_1(P) - f_1'(P)P.$$

Since $h''(P) = \frac{-N^2}{(1+NP)^2} < 0$, it follows that $h(P)$ is a concave function, that increases to its maximum value $M(P_1)$, where $M(P)$ is given by

$$M(P) = \log N - \log f_1'(P) - 1 + f_1'(P)/N$$

which is achieved at $P = \frac{1}{f_1'(P_1)} - \frac{1}{N}$, and $h(P)$ decreases on $(\frac{1}{f_1'(P_1)} - \frac{1}{N}, \infty)$. Thus, the following statements of equivalence hold: there are exactly two solutions to (42) iff $M(P_1) > J(P_1)$, there are no solutions to (42) iff $M(P_1) < J(P_1)$, and there is exactly one solution to (42) iff $M(P_1) = J(P_1)$. But

$$M'(P) - J'(P) = \frac{N-1}{N} P f_1''(P) k(\epsilon, N, P)$$

where

$$k(\epsilon, N, P) = 1 - 2\epsilon - \epsilon(1+\epsilon(N-1))P$$

so if $\epsilon < 1/2$ then $M(P) - J(P)$ is decreasing on $(0, \frac{1-2\epsilon}{\epsilon(1+\epsilon(N-1))})$, and increasing on $(\frac{1-2\epsilon}{\epsilon(1+\epsilon(N-1))}, \infty)$. In the following, we assume that $\epsilon < 1/2$, as in the statement of the result. But since $\epsilon < 1/2$, $M(0+) - J(0+) < 0$ and $M(\infty) - J(\infty) > 0$, so there is a unique P_l such that $M(P_l) = J(P_l)$, which implies that the tangent line to C_1 at $(P_l, f_1(P_l))$ touches C_2 at a unique point $(P_u, f_2(P_u))$. For $P < P_l$, $M(P) < J(P)$, and, hence, the tangent line to C_2 at $(P, f_1(P))$ does not intersect C_2 at all. For $P > P_l$, $M(P) > J(P)$, and, hence, the tangent line to C_2 at $(P, f_1(P))$ intersects C_2 at two points.

Equation (12) follows from fact that the tangent to C_1 at $(P_l, f_1(P_l))$ touches C_2 at $(P_u, f_2(P_u))$, as does the equation

$$f_1(P_l) = f_2(P_u) + f_2'(P_u)(P_l - P_u). \tag{43}$$

If $P_u < \tilde{P}$ then $f_2(P_u) < f_1(P_u)$, by (9), but then

$$f_1(P_u) > f_1(P_l) + f_1'(P_l)(P_u - P_l)$$

by (12), which contradicts the concavity of f_1. Similarly, if $\tilde{P} < P_l$ then $f_1(P_l) < f_2(P_l)$, by (10), but then

$$f_2(P_l) > f_2(P_u) + f_2'(P_u)(P_l - P_u)$$

by (43), which contradicts the concavity of f_2. Hence, $P_l < \tilde{P} < P_u$. The remaining statements of the lemma are either straightforward consequences of the strict concavity of the functions f_1 and f_2, or of results proven above. ■

B. A Basic Schur-Concavity Result for Two Links

The functions to be maximized in this paper (*cf* Problem 3.2 in Section III, and Problem 4.1 in Section IV) are neither concave, nor convex, and they possess local maxima, making standard numerical approaches problematic. Nevertheless, using the symmetry in these problems, we have found some interesting underlying structure that enable these problems to be solved. The most basic result we need is expressed in Lemma A.1 below, and this lemma is used in solving both of the above problems.

Consider the following function, that provides the sum rate in the two link case: for $a > 0, \epsilon > 0$

$$g(\epsilon, a, P_1, P_2) = C\left(\frac{P_1}{a + \epsilon P_2}\right) + C\left(\frac{P_2}{a + \epsilon P_1}\right). \tag{44}$$

where a represents the background noise level.

The following lemma considers the function restricted to the segment:

$$\mathcal{P} = \{(P_1, P_2) : P_1 + P_2 = \hat{P}, P_1 \geq 0, P_2 \geq 0\} \tag{45}$$

for some fixed total sum power on the two links, $\hat{P}$.

Lemma A.1: For fixed ϵ, a, the function $g(\epsilon, a, \cdot, \cdot)$ is:

- Schur-concave [28] on $\mathcal{P}$ if $\epsilon \leq \epsilon^*(a, \hat{P})$;
- Schur-convex [28] on $\mathcal{P}$ if $\epsilon \geq \epsilon^*(a, \hat{P})$

where $\epsilon^*(a, \hat{P}) = \sqrt{a}\frac{\sqrt{a+\hat{P}} - \sqrt{a}}{\hat{P}}$.

Proof: We consider only the case where $a = 1$, otherwise we can re-scale the powers. With ϵ and a fixed, and under the constraint (45), we can write (44) as a function of a single variable

$$g(P_1) = g(\epsilon, a, P_1, \hat{P} - P_1). \tag{46}$$

Writing $c = \frac{\hat{P}}{2}$, and employing the change of variables $P_1 = b + c, P_2 = c - b$, (46) becomes

$$g(b) = C\left(\frac{b + c}{1 + \epsilon(c - b)}\right) + C\left(\frac{c - b}{1 + \epsilon(b + c)}\right). \tag{47}$$

and the constraint (45) becomes: $-c \leq b \leq c$. Let $d_1 = 1 + \epsilon c$, $d_2 = \epsilon b$, $d_3 = d_1 + c$ and $d_4 = b - d_2$. Then we can write the derivative of $g(\cdot)$ as

$$g'(b) = 2b\frac{(\epsilon d_3 - (1 - \epsilon)d_1)(\epsilon d_3 + (1 - \epsilon)d_1)}{(d_1^2 - d_2^2)(d_3^2 - d_4^2)} \tag{48}$$

Since d_1 and d_3 are independent of b, if $(\epsilon d_3 - (1 - \epsilon)d_1) \neq 0$, the only root of $g'(b)$ happens at $b = 0$. The only positive ϵ for which $(\epsilon d_3 - (1 - \epsilon)d_1)$ becomes zero is the $\epsilon^*(\cdot, \cdot)$ given in the lemma. This proves that the function $g(\cdot)$ in (46) increases in the interval $(0, \hat{P}/2)$ and then decreases to $\log(1 + \hat{P})$ at $P_1 = \hat{P}$, if $\epsilon \leq \epsilon^*$, and vice versa otherwise. The fact that $g(\cdot)$ is strictly increasing and then strictly decreasing, along with its symmetric property around $\hat{P}/2$ implies that $g(\epsilon, a, \cdot, \cdot)$ in (44) is Schur concave [28] on $\mathcal{P}$ when $\epsilon \leq \epsilon^*$, and Schur convex otherwise. ■

Corollary A.1: For fixed ϵ, a, the maximization of the function $g(\epsilon, a, \cdot, \cdot)$ over $\mathcal{P}$ occurs at:

- $(\hat{P}/2, \hat{P}/2)$ if $\epsilon \leq \epsilon^*(a, \hat{P})$;
- $(0, \hat{P})$ or $(\hat{P}, 0)$ if $\epsilon \geq \epsilon^*(a, \hat{P})$.

Note that Corollary A.1 provides the solution to the two link version of the problem expressed in (20), which is the key result needed in the proof of Theorem 3.1. Problem (20) concerns the maximization of (19) subject to a constraint on the sum of the powers of all the links. One might hope that Lemma A.1 would generalize to an arbitrary number of links, which would be saying that $C_{sum}(\epsilon, N, \cdot)$ is Schur-convex, or Schur-concave, depending on the value of ϵ. However, this turns out not to be the case. Nevertheless, Lemma 3.3 shows that the optimal solution to the problem expressed in (20) is *as if* the function $C_{sum}(\epsilon, N, \cdot)$ has the above Schur-convex/concave structure, even though it does not.

The proof of Lemma 3.3 is given in Appendix C and it uses Lemma A.1 in an iterative manner. Indeed an upper bound to the problem is obtained using an iterative procedure, with links being updated two at a time. The upper bound is then shown to be achieveable. A similar approach can be applied to other problems, including Problem 4.1 in Section IV.

C. Proof of Lemma 3.3

Recall that $C_{sum}(\epsilon, N, \boldsymbol{P})$, as defined in (19), provides the objective function for the problem expressed in (20). Consider an arbitrary feasible power vector $\boldsymbol{P}^{(1)} = (P_1^{(1)}, P_2^{(1)}, \ldots, P_N^{(1)})$, satisfying $\sum_{j=1}^{N} P_j^{(1)} = \hat{P}$. Without loss of generality, we assume the components of all our power vectors are sorted in decreasing order, so that

$$P_1^{(1)} \geq P_2^{(1)} \geq \cdots \geq P_N^{(1)}.$$

Define

$$\bar{P}^{(1)} = \sum_{j=3}^{N} P_j^{(1)}, \; a_1 = 1 + \epsilon \bar{P}^{(1)}, \; \hat{P}^{(1)} = P_1^{(1)} + P_2^{(1)}$$

and consider the function $g^{(1)}(\,\cdot\,) = g(\epsilon, a_1, \cdot)$ (see (44) for the definition of g) restricted to the domain

$$\mathcal{P}^{(1)} = \{(P_1, P_2) : P_1 + P_2 = \hat{P}^{(1)}\}.$$

Lemma A.1 implies that if $\epsilon \leq \epsilon^*(a_1, \hat{P}^{(1)})$ then $g^{(1)}$ is Schur-concave on $\mathcal{P}^{(1)}$, but if if $\epsilon > \epsilon^*(a_1, \hat{P}^{(1)})$ then $g^{(1)}$ is Schur-convex on $\mathcal{P}^{(1)}$. In either case, we can construct a sequence of power vectors that cannot decrease the achieveable sum-rate, as we now show.

Case 1: $\epsilon \leq \epsilon^*(a_1, \hat{P}^{(1)})$.

First, consider another arbitrary, feasible power vector $\boldsymbol{Q}$, ordered in decreasing order as above. For any component i, we can define a new vector $\boldsymbol{Q}'$ by decreasing Q_i and increasing Q_{i+1} by the same amount. Provided the amount swapped between the two vectors is no more than $Q_i - Q_{i+1}$, the vector $\boldsymbol{Q}'$ will also be ordered in decreasing order, and $\boldsymbol{Q} \succ \boldsymbol{Q}'$. Such a transfer is known as a Pigou-Dalton transfer, and it is given the name "elementary Robin Hood operation" in [29]. It is well known [29] that if $\boldsymbol{Q} \succ \boldsymbol{R}$ then one can generate $\boldsymbol{R}$ from $\boldsymbol{Q}$ via a countable sequence of elementary Robin Hood operations.

Now let us denote the vector $(\hat{P}/N, \hat{P}/N, \ldots, \hat{P}/N)$ by $\boldsymbol{P}_{eq}$. Since $\boldsymbol{P}^{(1)} \succ \boldsymbol{P}_{eq}$ it follows that there is a sequence $\boldsymbol{P}^{(n)}$ of feasible power vectors, starting at $\boldsymbol{P}^{(1)}$, converging to $\boldsymbol{P}_{eq}$, where $\boldsymbol{P}^{(n+1)}$ is obtained from $\boldsymbol{P}^{(n)}$ by an elementary Robin Hood operation. Let $i_n, i_n + 1$ denote the components where the transfer takes place at step n of this sequence, and without loss of generality, let $i_1 = 1$. For each $n \in \mathbb{Z}_+$, define

$$\bar{P}^{(n)} = \sum_{j=1}^{N} P_j^{(n)} I_{\{j \neq i_n, i_n+1\}}$$
$$a_n = 1 + \epsilon \bar{P}^{(n)} \text{ and } \hat{P}^{(n)} = P_{i_n}^{(n)} + P_{i_n+1}^{(n)}. \tag{49}$$

It is trivial to show, by induction, that for any $i \in \{1, 2, \ldots, N-1\}$, and any $n \in \mathbb{Z}_+$, we have that

$$P_i^{(n)} + P_{i+1}^{(n)} \leq P_1^{(1)} + P_2^{(1)}$$

so in particular

$$P_{i_n}^{(n)} + P_{i_n+1}^{(n)} \leq P_1^{(1)} + P_2^{(1)}.$$

It follows from (49) that

$$\bar{P}^{(n)} \geq \bar{P}^{(1)}, \; a_n \geq a_1 \text{ and } \hat{P}^{(n)} \leq \hat{P}^{(1)}.$$

Let $g^{(n)}(\,\cdot\,)$ be the function $g(\epsilon, a_n, \cdot)$ restricted to the domain

$$\mathcal{P}^{(n)} = \left\{(P_1, P_2) : P_1 + P_2 = \hat{P}^{(n)}\right\}.$$

Now, since $\epsilon \leq \epsilon^*(a_1, \hat{P}^{(1)})$, it follows that $\epsilon \leq \epsilon^*(a_n, \hat{P}^{(n)})$, for all $n \in \mathbb{Z}_+$, and, hence, $g^{(n)}$ is Schur-concave on $\mathcal{P}^{(n)}$ for all such n. Since the power vectors $\boldsymbol{P}^{(n)}$ decrease in order of majorization, it follows that $C_{sum}(\epsilon, N, \boldsymbol{P}^{(1)}) < C_{sum}(\epsilon, N, \boldsymbol{P}_{eq})$, unless $\boldsymbol{P}^{(1)} = \boldsymbol{P}_{eq}$.

Case 2: $\epsilon > \epsilon^*(a_1, \hat{P}^{(1)})$.

For $n = 2, 3, \ldots, N$, define the power vector $\boldsymbol{P}^{(n)}$ by

$$P_j^{(n)} = \begin{cases} \sum_{i=1}^{n} P_i^{(1)}, & j = 1 \\ P_{j+n-1}^{(1)}, & j = 2, 3, \ldots, N-n+1 \\ 0, & j = N-n+2, N-n+3, \ldots, N \end{cases}$$

which gives us a sequence of power vectors, all satisfying the feasibility constraint that $\sum_{j=1}^{N} P_j^{(n)} = \hat{P}$, and all with the property of decreasing order for the components of each power vector. Further, as a sequence of power vectors, the vectors are increasing in order of majorization, with the final element in the sequence being $\boldsymbol{P}^{(N)} = (\hat{P}, 0, 0, \ldots, 0)$. For each n, define

$$\bar{P}^{(n)} = \sum_{j=2}^{N-n+1} P_j^{(n)} = \sum_{j=n+2}^{N} P_j^{(1)}$$
$$a_n = 1 + \epsilon \bar{P}^{(n)} \text{ and } \hat{P}^{(n)} = \sum_{i=1}^{n+1} P_i^{(1)}.$$

Note that $\bar{P}^{(n)}$ (and, hence, a_n) decrease with n, but $\hat{P}^{(n)}$ increases with n. Define also the function $g^{(n)}(\,\cdot\,) = g(\epsilon, a_n, \cdot)$ restricted to the domain

$$\mathcal{P}^{(n)} = \left\{(P_1, P_2) : P_1 + P_2 = \hat{P}^{(n)}\right\}.$$

Now, since a_n decreases, and $\hat{P}^{(n)}$ increases, it follows that $\epsilon^*(a_n, \hat{P}^{(n)})$ decreases with n. Since $\epsilon > \epsilon^*(a_1, \hat{P}^{(1)})$, it follows that $\epsilon > \epsilon^*(a_n, \hat{P}^{(n)})$, for all $n = 2, 3, \ldots, N$, and, hence, $g^{(n)}$ is Schur-convex on $\mathcal{P}^{(n)}$ for all such n. Since the power vectors $\boldsymbol{P}^{(n)}$ increase in order of majorization, it follows that $C_{sum}(\epsilon, N, \boldsymbol{P}^{(1)}) < C_{sum}(\epsilon, N, \boldsymbol{P}^{(N)})$, unless $\boldsymbol{P}^{(1)} = \boldsymbol{P}^{(N)}$.

Now suppose that the vector $\boldsymbol{P}^{(1)}$ is optimal for the problem expressed in (20). Then either $\boldsymbol{P}^{(1)} = (\hat{P}/N, \hat{P}/N, \ldots, \hat{P}/N))$ (if $\epsilon \leq \epsilon^*(a_1, \hat{P}^{(1)})$) or $\boldsymbol{P}^{(1)} = (\hat{P}, 0, 0, \ldots, 0)$ (if $\epsilon > \epsilon^*(a_1, \hat{P}^{(1)})$), for otherwise, we can improve the objective function, as described above in the two separate cases. We conclude that

$$U(\epsilon, N, \hat{P}) = \max\{f_1(\hat{P}/N), f_2(\hat{P}/N)\}.$$

D. Optimal Sum-Rate Under Binary Power Constraints

In this Appendix, we provide lemmas and their proofs, as required for Section IV-B, which deals with maximizing the sum rate under binary power constraints.

Set $a = 1 + P_{\max}$, and define the function $\phi : (0, 1) \to \mathbb{R}$: $\phi(x) = \frac{x}{1-x}\left(\frac{a - a^x}{a^x - 1}\right)$, so that $\epsilon_{n,1} = \frac{1}{P_{\max}} \phi(1/n)$.

Lemma A.2: ϕ is a decreasing function.

Proof: $\phi'(x) = \frac{\phi^{(1)}(x)}{d(x)}$, where

$$\phi^{(1)}(x) = a^{x+1} - a^{2x} + x a^x \log(a) - a + a^x - x^2 a^x \log(a)$$
$$- a^{x+1} x \log(a) + a^{x+1} x^2 \log(a)$$

and where $d(x) > 0$. The following edge conditions hold:

$$\phi^{(1)}(0) = 0, \phi^{(1)}(1) = 0. \quad (50)$$

Let $\phi^{(2)}(x) = \frac{d}{dx}\phi^{(1)}(x)$ and set $G(x) = \frac{\phi^{(2)}(x)}{a^x \log a}$. Then

$$G(x) = -2a^x + 2 + x\log(a) - 2x - x^2\log(a) - ax\log(a) + ax^2\log(a) + 2ax$$

and

$$G(0) = 0, G(1) = 0. \quad (51)$$

Set

$$G^{(1)}(x) = G'(x) = -2a^x\log(a) + \log(a) - 2 - 2x\log(a) - a\log(a) + 2a + 2ax\log(a).$$

We have

$$G^{(1)}(0) = G^{(1)}(1) = -(a+1)\log(a) - 2 + 2a.$$

Using the inequality $\log(1+x) > \dfrac{2x}{2+x}$ [30], we obtain

$$G^{(1)}(0) < 0, G^{(1)}(1) < 0. \quad (52)$$

Set $G^{(2)}(x) = \frac{d}{dx}G^{(1)}(x)$, and

$$H(x) = \frac{G^{(2)}(x)}{2\log(a)} = -a^x\log(a) - 1 + a.$$

The equation $H(x) = 0$ has a unique solution in x, namely

$$\hat{x} = \frac{1}{\log(a)}\log\left(\frac{a-1}{\log(a)}\right) > 0.$$

Using the inequality $\log(1+x) > \frac{2x}{2+x}$, one can show that

$$\hat{x} < \frac{\log(1+P_{\max}/2)}{\log(1+P_{\max})} < 1$$

but since $\log(1+P_{\max}) < \frac{P_{\max}}{\sqrt{1+P_{\max}}}$ [30], we have $\sqrt{a}\log(a) < a-1$, which implies

$$\log(a) - 2\log\left(\frac{a-1}{\log(a)}\right) - a\log(a) + 2a\log\left(\frac{a-1}{\log(a)}\right) > 0$$

and, hence

$$G^{(1)}(\hat{x}) > 0. \quad (53)$$

From (52), and (53), we have that $G^{(1)}(x) = 0$ has exactly two solutions in $(0,1)$: $\hat{x}_1, \hat{x}_2$, with $0 < \hat{x}_1 < \hat{x} < \hat{x}_2 < 1$, and G decreases on $(0,\hat{x}_1)$, increases on $(\hat{x}_1,\hat{x}_2)$, and decreases on $(\hat{x}_2, 1)$. By (51), it follows that there is a unique x that solves $G(x) = 0$, or equivalently $\phi^{(2)}(x) = 0$, and x is, therefore, the unique minimizer of $\phi^{(1)}$. By (50) it follows that $\phi^{(1)}(x) < 0$ for all $x \in (0,1)$, i.e., ϕ is a decreasing function. ■

Lemma A.3:

$$\forall n \geq 3, \forall\epsilon \in \left(0, \frac{1}{C(P_{\max})}\right), R'_n(\epsilon) < R'_{n-1}(\epsilon) < 0$$

Proof: $\forall n \geq 2$

$$R'_n(\epsilon) = \frac{-n(n-1)P_{\max}^2}{((n-1)\epsilon P_{\max}+1)((n-1)\epsilon P_{\max}+P_{\max}+1)} < 0.$$

Suppose $\epsilon < \frac{1}{C(P_{\max})}$. Then $\epsilon < \frac{1}{2} + \frac{1}{P_{\max}}$, so

$$(n-2)\epsilon P_{\max}(2 + P_{\max} - \epsilon P_{\max}) + 2(1+P_{\max}) > 0.$$

Hence

$$\frac{n}{((n-1)\epsilon P_{\max}+1)((n-1)\epsilon P_{\max}+P_{\max}+1)} > \frac{n-2}{((n-2)\epsilon P_{\max}+1)((n-2)\epsilon P_{\max}+P_{\max}+1)}$$

which implies that $R'_n(\epsilon) < R'_{n-1}(\epsilon)$, and clearly both are negative. ■

Lemma A.4:

$$\forall n \geq 2, \forall\epsilon \in (0, \epsilon_{n,1}), R_n(\epsilon) > R_{n-1}(\epsilon).$$

Proof: The result is clearly true for $n = 2$. For $n \geq 3$, Lemma A.3, and the fact that $R_n(0) > R_{n-1}(0)$, together imply that the equation $R_n(\epsilon) = R_{n-1}(\epsilon)$ has at most one solution for ϵ in the interval $(0, \frac{1}{C(P_{\max})})$. If this solution exists, it cannot lie in the interval $(0, \epsilon_{n-1,1})$, for that would imply $\epsilon_{n,1} < \epsilon_{n-1,1}$, contradicting Lemma 4.3. Thus, $R_{n-1}(\epsilon) < R_n(\epsilon)\ \forall\epsilon \in (0, \epsilon_{n-1,1})$. For $\epsilon_{n-1,1} < \epsilon < \epsilon_{n,1}$, we have $R_{n-1}(\epsilon) < C(P_{\max}) < R_n(\epsilon)$. ■

E. Proof of Theorem 4.1

A sketch of the proof is provided in Section IV-A. This Appendix contains the proofs of the lemmas stated in Section IV-A.

Let $\boldsymbol{P}^*$ be a vector of power levels that provides a solution to Problem 4.1, with the entries of $\boldsymbol{P}^*$ ordered in decreasing order. It is immediate that $P_1^* = P_{\max}$, for otherwise we can increase the value of $C_{sum}(\epsilon, N, \boldsymbol{P}^*)$ (see (19) for the definition) by scaling all elements of $\boldsymbol{P}^*$ by a common factor greater than unity, contradicting the optimality of $\boldsymbol{P}^*$. Thus, without loss of generality, we can assume that $\boldsymbol{P}^*$ has the first k entries equal to $P_{\max}$, for some integer $1 \leq k \leq N$.

If $k = N$, we have a binary power vector, and in this case the statement in Theorem 4.1 follows from Lemma 4.5. Let us, therefore, consider the case that $k < N$: If $k < N$, then let $0 \leq P_{k+1}^* = P < P_{\max}$. Lemma A.1 can be applied to this case to obtain the characterization of $\boldsymbol{P}^*$ expressed in Lemma 4.1, as we now show.

Proof of Lemma 4.1:

Proof: If $k + l = N$ the result holds by definition, so assume $k + l \leq N - 1$. For the purpose of obtaining a contradiction, assume that $0 < P_{k+l+1}^* < P = P_{k+l}^*$. Recall that $P < P_{\max}$. For $i \neq k+l, i \neq k+l+1$, set $P_i = P_i^*$, and define the constant a by

$$a = \epsilon\sum_j P_j^* 1_{\{j \neq k+l, j \neq k+l+1\}}.$$

Then

$$C_{sum}(\epsilon, N, \boldsymbol{P}) = g(\epsilon, a, P_{k+l}, P_{k+l+1}) + \sum_{j\neq k+l, j\neq k+l+1} C\left(\frac{P_j^*}{1+\epsilon\sum_{i\neq j} P_i^*}\right) \tag{54}$$

where the function g is defined in (44). We consider optimizing the expression in (54), but varying only (P_{k+l}, P_{k+l+1}), keeping the sum of these two powers fixed

$$P_{k+l} + P_{k+l+1} = \hat{P} := P_{k+l}^* + P_{k+l+1}^*.$$

Note that this constraint ensures that the second term in (54) is constant.

If $\epsilon < \epsilon^*(a, \hat{P})$ then by Lemma A.1, $g(\epsilon, a, \cdot, \cdot)$ is Schur-concave over the constrained set, so $C_{sum}(\epsilon, N, \boldsymbol{P})$ is maximized at $(P_{k+l}, P_{k+l+1}) = (\hat{P}/2, \hat{P}/2)$, which contradicts the optimality of $\boldsymbol{P}^*$.

If $\epsilon > \epsilon^*(a, \hat{P})$ then by Lemma A.1, $g(\epsilon, a, \cdot, \cdot)$ is Schur-convex over the constrained set. If $\hat{P} > P_{\max}$, then $(P_{\max}, \hat{P} - P_{\max}) \succ (P_{k+l}^*, P_{k+l+1}^*)$, where $\succ$ is the symbol for "majorizes" [28]. If $\hat{P} < P_{\max}$, then $(\hat{P}, 0) \succ (P_{k+l}^*, P_{k+l+1}^*)$. Either way, this contradicts the optimality of $\boldsymbol{P}^*$. ■

It is shown in Section IV-A that the proof of Theorem 4.1 is complete once we establish Lemma 4.2. To this end, recall the function

$$J(\epsilon, P_{\max}, k, l, P) = k\,C\left(\frac{P_{\max}}{1+\epsilon(k-1)P_{\max} + \epsilon l P}\right) + l\,C\left(\frac{P}{1+\epsilon k P_{\max} + \epsilon(l-1)P}\right)$$

and the statement of the lemma in Section IV-A.

Proof of Lemma 4.2: Assume $\epsilon, P_{\max}, k, l$ are fixed, and write $J(P)$ to simplify the notation. By differentiating $J(P)$, we obtain

$$J'(P) = \frac{l(1+\epsilon k P_{\max})}{(\Delta + \epsilon P_{\max} + P)(\Delta + \epsilon P_{\max})} - \frac{l\epsilon k P_{\max}}{(\Delta + \epsilon P + P_{\max})(\Delta + \epsilon P)} \tag{55}$$

where $\Delta = 1 + \epsilon(k-1)P_{\max} + \epsilon(l-1)P$. One can immediately verify that

$$J'(P_{\max}) > 0 \tag{56}$$

from which the next lemma follows.

Lemma A.5: If $J'(P)$ has only one real root in the interval $(0, P_{\max})$ then $J(\cdot)$ does not have a local maximum in the interval $(0, P_{\max})$.

We can write $J'(P)$ as a rational function, with numerator $Q(P) = aP^2 + bP + c$. The coefficients of $Q(P)$ are polynomials in ϵ, and we explicitly compute these coefficients

$$a = kP_{\max} l(2l-1)\epsilon^2(\epsilon - \epsilon_a) \tag{57}$$

$$b = kP_{\max}^2 l 2(k-l)\epsilon\left(\epsilon + \frac{1}{kP_{\max}}\right)(\epsilon - \epsilon_b) \tag{58}$$

$$c = -kP_{\max}^3 l(2k-1)\left(\epsilon + \frac{1}{kP_{\max}}\right)(\epsilon - \epsilon_c)(\epsilon - \epsilon_c^*) \tag{59}$$

where ϵ_a, ϵ_b, and ϵ_c are given by

$$\epsilon_a = \frac{l-1}{2l-1} - \frac{l^2}{kP_{\max}(2l-1)} \tag{60}$$

$$\epsilon_b = \frac{1}{2}\left(1 - \frac{2l}{(k-l)P_{\max}}\right) \tag{61}$$

$$\epsilon_c = \frac{1}{2}\frac{f_1 + \sqrt{f_2}}{(2k-1)P_{\max}};\ \epsilon_c^* = \frac{1}{2}\frac{f_1 - \sqrt{f_2}}{(2k-1)P_{\max}} \tag{62}$$

$$f_1 = k(1+P_{\max}) - P_{\max} - 2 \tag{63}$$

$$f_2 = k^2(1+P_{\max})^2 - 2kP_{\max}^2 + 2kP_{\max} + 4k + P_{\max}^2. \tag{64}$$

If $a = 0$ (i.e., $\epsilon = \epsilon_a$) then $Q(P)$ is linear in P, and, hence, $J'(\cdot)$ has at most one real root. It follows from Lemma A.5 that $J(\cdot)$ does not have a local maximum in the interval $(0, P_{\max})$ in this case.

If $a \neq 0$ then $Q(P)$ has two complex roots (possibly real), namely $\frac{-b\pm\sqrt{D}}{2a}$, where $D = b^2 - 4ac$ is the discriminant of $Q(P)$.

Lemma A.6: If any of the following conditions hold, then $J(\cdot)$ does not have a local maximum in $(0, P_{\max})$: (i) $a < 0$, or (ii) $c < 0$, or (iii) $b > 0$. Equivalently, conditions (i) and (ii) can be written (i) $\epsilon < \epsilon_a$, (ii) $\epsilon > \epsilon_c$, respectively. If $k \le l$ then the condition (iii) is equivalent to $\epsilon < \epsilon_b$. If $k > l$ then the condition (iii) is equivalent to $\epsilon > \epsilon_b$.

Proof: If $a < 0$ then $\lim_{P\to\infty} J'(P) < 0$. But by (56), it follows that $J'(P)$ has a real root in $(P_{\max}, \infty)$, and, hence, at most one real root in $(0, P_{\max})$. If $c < 0$ then $J'(0) < 0$. By (56), it also follows that there is at most one real root in $(0, P_{\max})$. In both cases, the result stated follows from Lemma A.5. If $b > 0$ and $ac > 0$ then $J'(P)$ has no real, positive roots. But if $ac < 0$ then case (i) or case (ii) applies. ■

Lemma A.7: If $k > l$ and $\epsilon_a > \min(\epsilon_b, \epsilon_c)$ then $J(\cdot)$ has no local maximum in $(0, P_{\max})$.

Proof: a direct corollary of Lemma A.6. ■

Now assume that $k \le l$. The value ϵ_c in (62) is a function of k and $P_{\max}$. We shall show that $\epsilon_c(\cdot, P_{\max})$ is increasing in k when $P_{\max} > 1.0$ and decreasing otherwise. By differentiating with respect to k

$$\epsilon_c'(k) = \frac{2\sqrt{f_2}[(2k-1)f_1' - 2f_1] + [(2k-1)f_2' - 4f_2]}{4(2k-1)^2\sqrt{f_2}P_{\max}} \tag{65}$$

$$= \frac{1}{2(2k-1)^2\sqrt{f_2}}((P_{\max}+3)\sqrt{f_2} - 4kP_{\max} + (k-1)P_{\max}^2 - 5k - P_{\max} - 2). \tag{66}$$

The denominator of the above function is always positive when k and $P_{\max}$ are positive; hence, we will focus our attention on the numerator, which is of the form $\zeta_1 + \zeta_2$ with $\zeta_1 = (3 + P_{\max})\sqrt{f_2}$. Since ζ_1 is always positive, the following lemma is immediate.

Lemma A.8: With ζ_i defined as above, if $\zeta_1^2 - \zeta_2^2 > 0$, then $\zeta_1 + \zeta_2 \geq 0$.

But ζ_2 as a function of $P_{\max}$ has one root in the positive real axis, after which it becomes positive. In particular, $\zeta_2 < 0$ when $P_{\max} = 1.0$. Thus, we have the following lemma.

Lemma A.8: With ζ_i defined as above, if $P_{\max} \leq 1$ and $\zeta_1^2 - \zeta_2^2 \leq 0$, then $\zeta_1 + \zeta_2 \leq 0$.

Now we can easily compute

$$\zeta_1^2 - \zeta_2^2 = 16(1+P_{\max})^2(P_{\max}-1)(2k-1)^2 \tag{67}$$

which has the same sign as $(P_{\max} - 1)$. Using the two lemmas above, the function ϵ_c is increasing in k when $P_{\max} \geq 1$ and vice versa. Hence

$$\begin{aligned} \epsilon_c &\leq \max\{\lim_{k\to 0}\epsilon_c, \lim_{k\to\infty}\epsilon_c\} && (68)\\ &= \max\left\{\frac{1}{P_{\max}}, \frac{1}{2} + \frac{1}{2P_{\max}}\right\} && (69)\\ &\leq \frac{1}{2} + \frac{1}{P_{\max}} && (70)\\ &\leq \frac{1}{2} + \frac{l}{(l-k)P_{\max}} && (71)\\ &= \epsilon_b. && (72)\end{aligned}$$

We summarize this inequality in the following lemma:

Lemma A.10: If $k \leq l$ then $\epsilon_b \geq \epsilon_c$.

Lemma A.11: If $k \leq l$ then $J(\cdot)$ does not have a local maximum in $(0, P_{\max})$.

Proof: a direct corollary of Lemmas A.6 and A.10. ■

The remaining case to consider is $k > l$ and $\epsilon_a < \min(\epsilon_b, \epsilon_c)$. For this case, we cannot in general prove that $J(\cdot)$ does not have a local maximum in the interval $(0, P_{\max})$. Indeed, Fig. 5 demonstrates that there do exist choices of $k, l, P_{\max}, \epsilon$ for which the corresponding $J(\cdot)$ *does* have such a local maximum. Moreover, there are subcases in which such a local maximum is in fact a global maximum of the function. The figure suggests that such local maxima may only appear when ϵ is sufficiently large. For the example depicted in Fig. 5, $\epsilon = 0.332$ is the critical value. In fact, this observation is true in general, as proven in the lemma below.

Lemma A.12: If $k > l$ and $\epsilon_a < \min(\epsilon_b, \epsilon_c)$ then there exists a uniquely defined positive number $\epsilon_d(k, l, P_{\max})$ (a function of $k, l, P_{\max}$, which can take the value ∞, and which is defined in the proof below) such that the following three conditions all hold:

i) there is no local maximum of $J(\epsilon, P_{\max}, k, l, \cdot)$ in $(0, P_{\max})$ for any $\epsilon < \epsilon_d$;
ii) if $\epsilon_a < \epsilon_b < \epsilon_c$ then $\epsilon_b < \epsilon_d < \infty$,;
iii) if $\epsilon_a < \epsilon_c < \epsilon_b$ then $\frac{1}{C(P_{\max})} < \epsilon_d$.

Proof: Recall that $D = b^2 - 4ac$ is the discriminant of the quadratic $Q(P)$. For fixed $k, l, P_{\max}$, it is a six degree polynomial in ϵ, which we will denote by $D(\epsilon)$. It has a double root at $\epsilon = 0$, a single root at $\epsilon = -\frac{1}{kP_{\max}}$, so it has three more roots, at least one of which is real. In the following, we assume that $k > l$, as in the statement of the lemma.

We begin by establishing (ii). If $\epsilon_a < \epsilon_b < \epsilon_c$ then Lemma A.6 implies that $D > 0$ on $(0, \epsilon_a]$ and on $[\epsilon_c, \infty)$, but $D(\epsilon_b) < 0$. Since there are at most three positive real roots of $D(\epsilon)$, there must be exactly two, which we denote by ϵ_d^* and ϵ_d, with

$$\epsilon_a < \epsilon_d^* < \epsilon_b < \epsilon_d < \infty$$

which establishes (ii), and provides a definition of ϵ_d for this case.

To define ϵ_d for the case $\epsilon_a < \epsilon_c < \epsilon_b$, note that in this case we again have that $D > 0$ on $(0, \epsilon_a]$ and on $[\epsilon_c, \infty)$. Since there are at most three positive real roots of $D(\epsilon)$, there must be either zero positive roots, in which case we define $\epsilon_d^* = \epsilon_d = \infty$, or there are exactly two positive roots, which we denote by ϵ_d^* and ϵ_d, with

$$\epsilon_a < \epsilon_d^* < \epsilon_d < \epsilon_c.$$

We now proceed to establish (i). If $\epsilon_d^* < \epsilon < \epsilon_d$ then $J(P)$ has no real root, so (35) must hold in this case. At $\epsilon = \epsilon_a$, Lemma A.6 implies that $a = 0, b < 0, c > 0$; thus, $Q(P)$ is linear in P, with $Q(0) = c > 0$, and slope $b < 0$. It follows from (56) that in this case $J'(P)$ has a single root $P \in (P_{\max}, \infty)$. For $\epsilon > \epsilon_a$, we have $\lim_{P\uparrow\infty} J'(P) > 0$, so, for ϵ in a small neighborhood of ϵ_a, with $\epsilon > \epsilon_a$, there are two roots of $J'(P)$ in $(P_{\max}, \infty)$. The roots of $Q(P)$ are continuous functions of ϵ, so it must be that any real, positive roots of $J(P)$ are in $(P_{\max}, \infty)$ for all $\epsilon_a < \epsilon < \epsilon_d^*$; otherwise, there exists $\epsilon_a < \epsilon < \epsilon_d^*$ for which $J'(P_{\max}) = 0$, which contradicts (56). Finally, for $\epsilon < \epsilon_a$, $a < 0$, and Lemma A.6 (i) implies that (35) holds in this case. We conclude that (35) holds for all $\epsilon < \epsilon_d$, establishing (i).

We now proceed to establish (iii). To this end, assume that $\epsilon_a < \epsilon_c < \epsilon_b$. In the following, we write $\epsilon_a(k, l, P_{\max}), \epsilon_b(k, l, P_{\max})$ and $\epsilon_c(k, P_{\max})$ for the functions defined in (60)–(62). For $k \geq 2$, it can be easily seen from (62) that ϵ_c is a decreasing function of $P_{\max}$, and from (61) that ϵ_b is an increasing function of $P_{\max}$ for $k > l$. Thus, for any fixed k and l, the equation $\epsilon_c - \epsilon_b = 0$ will have at most one positive solution in $P_{\max}$. Denote this solution by $\tilde{P}$, where $\tilde{P}$ is a function of k and l

$$\tilde{P}(k,l) = \frac{k^2 + lk + \sqrt{k^4 + 2k^3 l + k^2 l^2 + 4k^2 - 4kl^2 - 4lk^2}}{k-l}. \tag{73}$$

Since $\epsilon_b > \epsilon_c$, it must be that $\tilde{P}(k,l)$ is real and positive and $P_{\max} > \tilde{P}(k,l)$.

We now consider three cases.

Case A: $l = 2, k > 2$

When $l = 2$, $\tilde{P}(k,2)$ is an increasing function in k for all $k \geq 6$, and $\tilde{P}(6,2) = 23.3$. So, for $k \geq 6$, $\frac{1}{C(P_{\max})} < 1/\log(24.3) = 0.3134$. But $0.324 = \epsilon_a(6, 2, 23.3) < \epsilon_a(k, 2, P_{\max})$. Thus, for $k \geq 6$

$$\frac{1}{C(P_{\max})} < \epsilon_a(k, 2, P_{\max}) \tag{74}$$

and one can directly verify that (74) is also true for $k = 3, 4, 5$. But $\epsilon_d > \epsilon_a$, so (iii) is established for case $l = 2$.

Case B: $2 < l < k$

Since $\epsilon_b(k, l, P_{\max})$ is a decreasing function of l (for $l < k$, and k fixed) it follows that $\tilde{P}(k,2) < \tilde{P}(k,l) < P_{\max}$, which implies that $\epsilon_a(k,l,P_{\max}) > \epsilon_a(k,2,P_{\max})$. But (74) holds for all $k \geq 3$, and, hence, (iii) is established for all l, k satisfying $2 < l < k$.

Case C: $l = 1, k > 1$

When $l = 1$ we can explicitly compute the roots of the discriminant

$$D(\epsilon) = \alpha\epsilon^2 \left(\epsilon + \frac{1}{kP_{\max}}\right)^2 (\epsilon - \epsilon_r)(\epsilon - \epsilon_r^*) \qquad (75)$$

(for constant α) where the (possibly complex) roots are given by

$$\epsilon_r = \frac{\beta_1 + \beta_2}{2k};\ \epsilon_r^* = \frac{\beta_1 - \beta_2}{2k} \qquad (76)$$

with

$$\beta_1 = k - 1 - \frac{1}{P_{\max}}; \beta_2 = \sqrt{\frac{14 - 2k}{P_{\max}} + \frac{12 + k}{P_{\max}^2}}. \qquad (77)$$

If ϵ_r^* and ϵ_r are complex conjugates, we have $\epsilon_d = \infty$ and there is nothing to prove, so assume that $\epsilon_d^* = \epsilon_r^*$, $\epsilon_d = \epsilon_r$, both real and positive. For $l = 1$ and $k \geq 2$, we have $P_{\max} > \tilde{P}(k,1) = 11.2$. Thus, $\frac{1}{C(P_{\max})} < 1/\log(12.2) = 0.4$. But $\epsilon_d > \frac{k-1}{2k}$, so if $k \geq 5$, then $\epsilon_d > 0.4$, proving (iii) for this case. But also $\epsilon_d > \frac{k-1-1/P_{\max}+\sqrt{(14-2k)/P_{\max}}}{2k}$, so if $2 \leq k \leq 4$ then

$$\epsilon_d > \frac{P_{\max} - 1 + \sqrt{6P_{\max}}}{4P_{\max}} > \frac{1}{C(P_{\max})}\ \forall P_{\max} > 11.2.$$

■

Lemma A.13: If $k > l$ and $\epsilon_a < \epsilon_b < \epsilon_c$ then for any $\epsilon > 0$, $J(\cdot)$ does not have a local maximum in the interval $(0, P_{\max})$.

Proof: By Lemma A.12 (i), there is no local maximum of $J(\cdot)$ in the interval $(0, P_{\max})$ for any $\epsilon < \epsilon_d$. By Lemma A.6 (iii), there is no local maximum in the same interval for any $\epsilon > \epsilon_b$. By Lemma A.12(ii), $\epsilon_b < \epsilon_d$. ■

The following lemma is concerned with the special case of $\epsilon = \epsilon_{N,1}$, where $\epsilon_{N,1}$ is defined in (32).

Lemma A.14: $C_{flat}(\epsilon_{N,1}, N, P_{\max}) = C(P_{\max})$.

Proof: The preceding results establish that $J(\epsilon_{N,1}, P_{\max}, k, l, \cdot)$ has no local maximum in the interval $(0, P_{\max})$ in all cases apart from $\epsilon_a < \epsilon_c < \epsilon_b$. Let us assume now that $\epsilon_a < \epsilon_c < \epsilon_b$. By Lemma A.12 (i) and (iii), we know that $J(\epsilon, P_{\max}, k, l, \cdot)$ has no local maximum in the interval $(0, P_{\max})$ for all $\epsilon < \frac{1}{C(P_{\max})}$, but by Lemma 4.3, we have that $\epsilon_{N1} < \frac{1}{C(P_{\max})}$, so (35) is established for $\epsilon = \epsilon_{N,1}, \epsilon_a < \epsilon_c < \epsilon_b$. We conclude that binary power control is optimal at $\epsilon = \epsilon_{N,1}$, and the result then follows from Lemma 4.5. ■

Now consider arbitrary ϵ, but with $k, l, P_{\max}$ satisfying the remaining condition that is yet to be resolved in general: $k > l$ and $\epsilon_a < \epsilon_c < \epsilon_b$.

Lemma A.15: If $k > l$ and $\epsilon_a < \epsilon_c < \epsilon_b$ then for any $\epsilon > 0$, and any $P \in (0, P_{\max})$

$$J(\epsilon, P_{\max}, k, l, P) \leq \begin{cases} NC\left(\frac{P_{\max}}{1+(N-1)\epsilon P_{\max}}\right), & \epsilon < \epsilon_{N,1} \\ C(P_{\max}), & \epsilon \geq \epsilon_{N,1}. \end{cases} \qquad (78)$$

Proof: If $\epsilon < \epsilon_d$ then the result follows directly from Lemma A.12 (i), together with Lemma 4.5. If $\epsilon > \epsilon_d$ then $C_{flat}(\epsilon, N, P_{\max}) \leq C_{flat}(\epsilon_d, N, P_{\max})$, by monotonicity of $C_{flat}(\cdot, N, P_{\max})$. But by Lemma 4.3 and Lemma A.12 (iii), we have that $\epsilon_{N1} < \frac{1}{C(P_{\max})} < \epsilon_d$. Monotonicity of $C_{flat}(\cdot, N, P_{\max})$ implies that $C_{flat}(\epsilon_d, N, P_{\max}) \leq C_{flat}(\epsilon_{N,1}, N, P_{\max})$. But by Lemma A.14, $C_{flat}(\epsilon_{N,1}, N, P_{\max}) = C(P_{\max})$. Thus, if $\epsilon > \epsilon_d$ then then

$$J(\epsilon, P_{\max}, k, l, P) \leq C_{flat}(\epsilon, N, P_{\max}) \leq C(P_{\max}).$$

The first inequality is due to $J(\epsilon, P_{\max}, k, l, P)$ being an achieveable sum rate, and the second inequality follows from the above bounds. But by Lemma 4.5, the RHS of (78) equals $C(P_{\max})$ when $\epsilon > \epsilon_{N,1}$. ■

We have now considered all possible cases and shown that if (35) does not hold, then (34) does hold, for all choices of $k, l \geq 1, P_{\max}$ and $0 \leq P \leq P_{\max}$. This concludes the proof of Lemma 4.2, and hence of Theorem 4.1.

ACKNOWLEDGMENT

The authors would like to thank the Australian Research Council for support under Grant DP0984862 and Prof. P. R. Kumar for insightful discussions.

REFERENCES

[1] R. D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 7, pp. 1341–1347, 1995.

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

[3] A. Gjendemsjo, D. Gesbert, G. E. Oien, and S. G. Kiani, "Optimal power allocation and scheduling for two-cell capacity maximization," in *Proc. 4th Int. Symp. Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, Boston, MA, Apr. 2006, pp. 1–6.

[4] M. Ebrahimi, M. A. M. Ali, and A. K. Khandani, "Power allocation and asymptotic achievable sum-rates in single-hop wireless networks," in *Proc. 40th Annu. Conf. Information Sciences and Systems*, Princeton, NJ, Mar. 2006, pp. 498–503.

[5] N. Badruddin, S. R. Bhaskaran, J. S. Evans, and S. V. Hanly, "Maximising the sum rate in symmetric networks of interfering links under flat power constraints," presented at the Allerton Conf., Monticello, IL, Sep. 2008.

[6] K. Song, S. Chung, G. Ginis, and J. Cioffi, "Dynamic spectrum management for next-generation DSLsystems," *IEEE Commun. Mag.*, vol. 40, no. 10, pp. 101–109, Oct. 2002.

[7] R. Cendrillon and M. Moonen, "Iterative spectrum balancing for digital subscriber lines," in *Proc. IEEE Int. Conf. Communications (ICC)*, May 2005, vol. 3, pp. 1937–1941.

[8] R. Etkin, A. Parekh, and D. Tse, "Spectrum sharing for unlicensed bands," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 517–528, Apr. 2007.

[9] R. Cendrillon, W. Yu, M. Moonen, J. Verliden, and T. Bostoen, "Optimal multi-user spectrum management for digital subscriber lines," *IEEE Trans. Commun.*, vol. 54, no. 5, pp. 922–933, May 2006.

[10] R. Cendrillon, J. Huang, M. Chiang, and M. Moonen, "Autonomous spectrum balancing for digital subscriber lines," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 4241–4257, Aug. 2007.
[11] M. Cao, V. Raghanathan, S. Hanly, V. Sharma, and P. Kumar, "Power control and transmission scheduling for network utility maximization in wireless networks," in *Proc. IEEE Conf. Decision and Control (CDC)*, Dec. 2007, pp. 5215–5221.
[12] S. Shamai and A. D. Wyner, "Information theoretic considerations for symmetric, cellular, multiple-access, fading channels—Part 1," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1877–1894, Nov. 1997.
[13] H. Shen, H. Zhou, R. Berry, and M. Honig, "Optimal spectrum allocation in gaussian interference networks," in *Proc. 42nd Asilomar Conf. Signals, Systems and Computers*, Nov. 2008, pp. 2142–2146.
[14] Y. Zhao and G. Pottie, "Optimum spectrum management in two-user symmetric interference channels," presented at the Information Theory and its Applications Workshop, Feb. 2009, Univ. California, San Diego.
[15] W. Yu, G. Ginis, and J. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 5, pp. 1105–1115, Jun. 2002.
[16] M. Chiang, "Balancing transport and physical layers in wireless multihop networks: Jointly optimal congestion control and power control," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 1, pp. 104–116, Jan. 2005.
[17] M. Chiang, C. W. Tan, D. Palomar, D. O'Neill, and D. Julian, "Power control by geometric programming," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2640–2651, Jul. 2005.
[18] J. Papandriopoulos and J. Evans, "SCALE: A low-complexity distributed protocol for spectrum balancing in multi-user DSL networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3711–3724, Dec. 2009.
[19] A. Lapidoth, "Nearest neighbour decoding for non-gaussian noise channels," *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1520–1529, Sep. 1996.
[20] V. Annapureddy and V. Veeravalli, "Gaussian interference networks: Sum capacity in the low interference regime," in *Proc. IEEE Int. Symp. Information Theory*, Toronto, ON, Canada, Jul. 2008, pp. 255–259.
[21] R. Etkin, D. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5534–5562, Dec. 2008.
[22] R. G. Gallager, *Information Theory and Reliable Communication*. Hoboken, NJ: Wiley, 1968.
[23] H. G. Eggleston, *Convexity*. Cambridge, U.K.: Cambridge Univ. Press, 1969.
[24] N. Badruddin, J. S. Evans, and S. V. Hanly, "Maximising sum rate for two interfering wireless links," in *Proc. Australian Comm. Theory Workshop (AusCTW 2008)*, Christchurch, New Zealand, Jan. 2008, pp. 75–81.
[25] S. Shamai and A. Wyner, "Information-theoretic considerations for symmetric, cellular, multiple-access fading channels—Part 1," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1877–1894, 1997.
[26] A. Gjendemsjo, D. Gesbert, G. E. Oien, and S. G. Kianai, "Binary power control for sum rate maximization over multiple interfering links," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 3164–3173, Aug. 2008.
[27] N. Badruddin, J. Evans, and S. Hanly, "Optimal binary power allocation for wireless networks with local interference," presented at the IEEE Int. Conf. Communications, Capetown, South Africa, 2010.
[28] A. Marshall and I. Olkin, *Inequalities : Theory of Majorization and its Applications*, ser. Mathematics in science and engineering. New York: Academic, 1979.
[29] B. Arnold, *Majorization and the Lorenz Order: A Brief Introduction*, ser. Lecture Notes in Statistics. Heidelberg, Germany: Springer-Verlag, 1987, vol. 43.
[30] E. R. Love, "Some logarithm inequalities," *Math. Gazette*, vol. 64, no. 427, pp. 55–57, 1980.

Sibi Raj Bhaskaran received the Ph.D. degree from EPFL, Switzerland, in 2007.

He is currently an Assistant Professor at the Indian Institute of Technology, Bombay. Prior to that, he was a research fellow at the University of Melbourne, Australia, until April 2009. His research interests are in communication and information theory, signal processing, and resource allocation problems.

Stephen V. Hanly (M'98) received the B.Sc. (Hons) and M.Sc. degrees from the University of Western Australia and the Ph.D. degree in mathematics in 1994 from Cambridge University, Cambridge, U.K.

From 1993 to 1995, he was a postdoctoral member of technical staff at AT&T Bell Laboratories. From 1996 to 2009, he was on the research and teaching staff at the University of Melbourne, Australia. He is presently an Associate Professor in the Department of Electrical and Computer Engineering, National University of Singapore. His research interests are in the areas of information theory, signal processing, and wireless networking.

Prof. Hanly was an Associate Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2005–2009 and is guest editor for the IEEE JOURNAL ON SELECTED AREAS special issue on "Cooperative Communications in MIMO Cellular Networks". In 2005, he was the technical co-chair for the IEEE International Symposium on Information Theory held in Adelaide, Australia. He was a co-recipient of the best paper award at the Infocom 1998 Conference and the 2001 Joint IEEE Communications Society and Information Theory Society best paper award, both for his work with D. Tse (Berkeley).

Nasreen Badruddin (M'10) received the B.Eng. (Hons) degree from RMIT University, Australia, in 2000, and the M.Sc. degree in electrical and computer engineering from Carnegie-Mellon University, Pittsburgh, PA, in 2002. She is currently pursuing the Ph.D. degree at the University of Melbourne, Australia.

She served as a lecturer at the Universiti Teknologi PETRONAS, Malaysia, from 2002 to 2006. Her research interests are in wireless networks and multimedia processing.

Jamie S. Evans (M'98) was born in Newcastle, Australia, in 1970. He received the B.S. degree in physics and the B.E. degree in computer engineering from the University of Newcastle, in 1992 and 1993, respectively, where he received the University Medal upon graduation. He received the M.S. and the Ph.D. degrees from the University of Melbourne, Australia, in 1996 and 1998, respectively, both in electrical engineering, and was awarded the Chancellor's Prize for excellence for his Ph.D. thesis.

From March 1998 to June 1999, he was a Visiting Researcher in the Department of Electrical Engineering and Computer Science, University of California, Berkeley. He returned to Australia to take up a position as Lecturer at the University of Sydney, Australia, where he stayed until July 2001. Since that time, he has been with the Department of Electrical and Electronic Engineering, University of Melbourne, where he is now a Professor. His research interests are in communications theory, information theory, and statistical signal processing with a focus on wireless communications networks.